

Rehaussement de la classification textuelle d'images par leur contenu visuel

S. Tollari H. Glotin J. Le Maitre
Laboratoire SIS, Université de Toulon et du Var, France
Mail: {tollari, glotin, lemaitre}@univ-tln.fr

Thème: indexation, recherche d'images par le contenu, classification, Kullback-Leibler.

Problème: Rehausser le résultat d'une classification textuelle d'images étiquetées en fonction de leur contenu visuel.

Contributions: Nous testons l'existence d'une cohérence entre l'indexation textuelle (ensemble de mots-clés) d'une image et son indexation visuelle (attributs de couleur et de forme). Ces expériences sont menées sur un corpus de 600 photos de presse indexées manuellement par des mots-clés extraits d'un thésaurus structuré hiérarchiquement. Nous constituons une classification de référence à partir de leur indexation textuelle. Chaque classe contient une dizaine d'images, qui serviront de référence visuelle. Après extraction de traits visuels simples, nous construisons des vecteurs de traits visuels pour chaque images de référence et de test. Nous évaluons alors les performances d'une recherche d'images combinant description textuelle et visuelle en se basant sur les mesures de distances entre les différents vecteurs. Nous obtenons par fusion tardive des deux modes de classification (visuelle et textuelle), 54% de gain par rapport à l'information textuelle seule.

Enfin, nous appliquons ces méthodes sur une autre base (COREL, 600 images), pour filtrer les images trouvées par un moteur de recherche d'images (« Google »). La requête textuelle permet d'accéder directement à la classe de référence visuelle la plus proche dans notre système, à partir de laquelle les comparaisons visuelles sont effectuées. La distribution des distances à la référence (Fig 2) permet alors de filtrer les images les plus pertinentes, c'est-à-dire celles appartenant au groupe des distances les plus petites. Ces images sont alors affichées comme résultat du filtre (Bas de la Fig. 2).

Image search :



Figure 1. Haut : images résultant de la requête textuelle « Black Bear Snow » sur Google. Bas : images filtrées par notre système visuel.

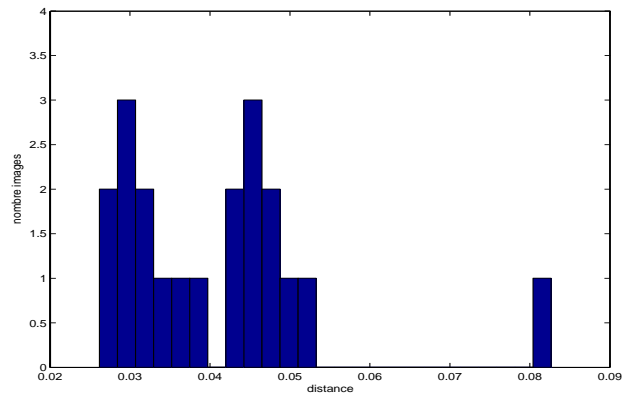


Figure 2: distribution des distances visuelles, à la classe de référence, des 18 images de la requête textuelle sur Google. Le groupe de gauche définit alors les 9 images conservées et affichées en bas de la figure 1.