

Fusion et sélection des traits visuels et de leur hétérogénéité dépendant du mot-clé par LDA approximée pour filtrer l'indexation d'images

Word dependant fusion and selection of visual features and their heterogeneity by approximation of LDA for image keyword filtering

Sabrina Tollari(1), Hervé Glotin(1), Pascale Giraudet(2)

(1)Laboratoire LSIS UMR CNRS 6168

(2)Département de biologie

Université du Sud Toulon-Var

BP 20132, F-83957 La Garde cedex

{tollari,glotin,giraudet}@univ-tln.fr

Pour améliorer les performances des systèmes automatiques d'annotations d'images à partir de données réalistes, nous proposons tout d'abord une méthode permettant de sélectionner les traits visuels les plus discriminants pour un mot-clé donné. Cette méthode est basée sur une analyse factorielle discriminante (LDA), mais étendue par approximation (ALDA) au cas de bases mal annotées comme le sont les grandes bases d'images. Nous validons l'ALDA théoriquement et expérimentalement sur COREL (10000 images, 200 mots).

D'autre part, nous proposons de dériver pour chaque trait leur hétérogénéité (ou entropie) dans chaque image associée à chaque concept visuel (ou mot clé). Nous appliquons l'ALDA sur ce nouveau jeu de traits, et nous mesurons pour chaque mot les gains de classification après sélection dans ces deux espaces réduits. Nous montrons que les gains sont forts pour la majorité des mots, et que les concepts sont plus ou moins associés à des traits classiques ou à leur hétérogénéité. Nous exposons alors les concepts de consistance visuelle et de caractère contextuel des concepts (voir Fig. 1).

Enfin nous appliquons notre système de classification adaptative au filtrage de l'indexation d'image (Fig. 2). Nous montrons que les deux informations visuelles (classique et hétérogénéité) sont complémentaires. Nous comparons donc des méthodes de fusion précoces et tardives pour améliorer ce filtrage, et obtenons un gain de l'ordre de +60% de classification pour une réduction de 80% du nombre de dimensions.

Mots-clés

Recherche d'images, LDA, problème des grandes dimensions, hétérogénéité, classification, filtrage

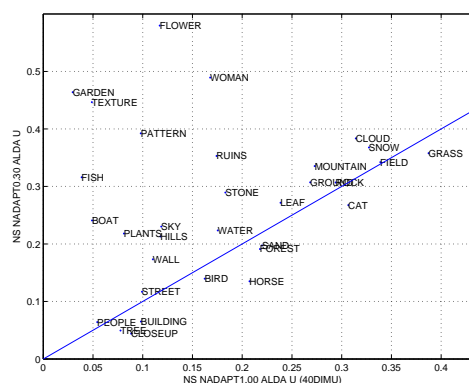


FIG. 1 – Représentation de la consistance visuelle (score NS) des mots obtenus par classification sans sélection de traits visuels en abscisse, et avec sélection des traits visuels par ALDA en ordonnée.



Image 172052 (10 blobs)

Légende (3 sur 3)
water(OK) mountain(OK)
coast(OK)
sensi=1.00 specif=0.65
preci=0.15 NS=0.65

20 mots associés par le système

desert(7) water(6) sky(6) wave(6) hills(6)
closeup(6) mountain(6) coast(6) tree(6)
beach(6) boat(5) branch(5) temple(5) fish(4)
sand(4) forest(4) cloud(4) people(4)
horizon(3) valley(3)

32 mots non associés

snow(2) statue(2) vegetable(1) rock(1) bird(1)
wall(1) flower(1) head(1) building(1)
window(1) woman(1) street(1) plants(1)
field(1) cat cougar food fungus garden grass
ground horse house ice leaf mushroom ocean
pattern ruins stone texture

Total: 52 mots

FIG. 2 – Exemple de filtrage d'indexation textuelle d'images. Les 52 mots les plus fréquents sont associés à cette image de TEST, puis filtrés par les classes visuelles construites sur l'espace réduit par sélection des traits visuels avec l'ALDA. Les mots *wave*, *hills*, *boat*, *cloud* associés par le système à cette image pourraient annoter cette image.