

Comparaisons noyaux sur sac et approches dictionnaires pour la classification d'images et de vidéos

Stage Master 2 Recherche 2010 - UPMC LIP6

1 Informations pratiques

Mots-clés : Sciences pour l'ingénieur, Traitement des images, Apprentissage, Indexation
Durée : 5 à 6 mois
Localisation : LIP6, 104 ave du President Kennedy, 75016 Paris
Encadrement : David Picard, Matthieu Cord, Nicolas Thome
Contact : david.picard@lip6.fr, Matthieu.Cord@lip6.fr, nicolas.thome@lip6.fr
Rémunération LIP6 : 2k euros pour le stage

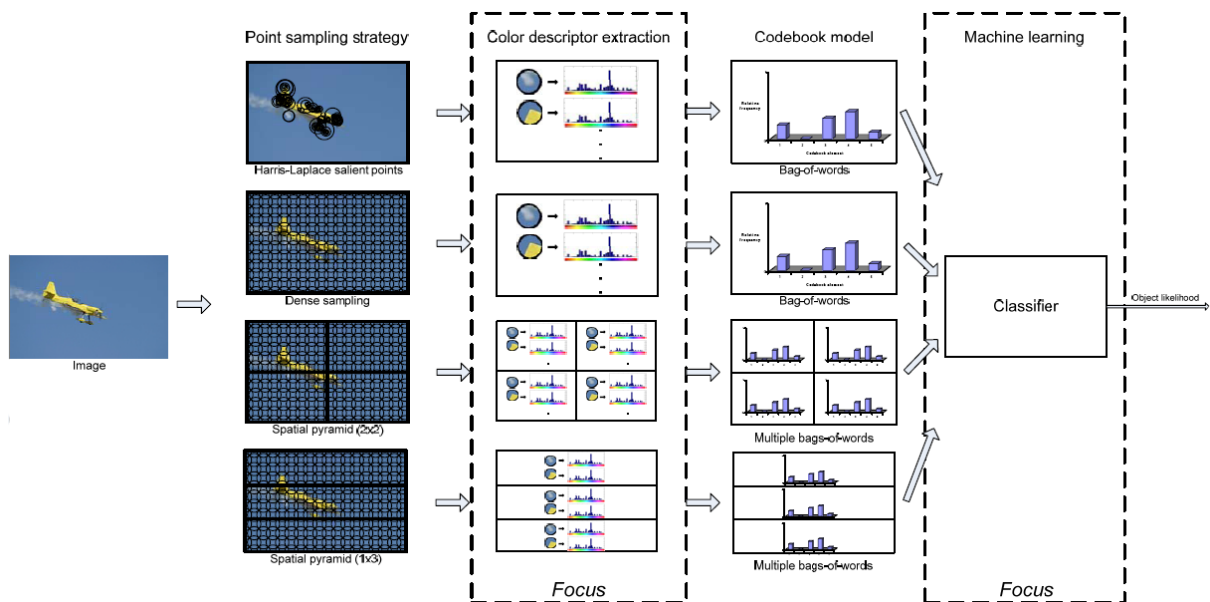


FIG. 1 – Pipeline de l'approche par dictionnaire.

2 Contexte

Pour opérer une classification d'images basée contenu, les méthodes proposées dans l'état de l'art repose aujourd'hui presque toutes sur l'extraction d'un ensemble descripteurs locaux (typiquement des descripteurs SIFT[3]), l'image étant alors représentée par un sac contenant ses descripteurs. Afin de comparer des images sur la base de ces descripteurs, plusieurs approches ont été proposées. La première consiste à quantifier l'espace des descripteurs de manière à produire un dictionnaire de mots visuels. Chaque descripteur du sac de l'image est alors projeté sur le(s) mot(s) le(s) plus proche(s) afin de constituer l'histogramme des occurrences des mots visuels dans l'image. La représentation de l'image de vient alors vectorielle (vecteur de fréquence d'appartenance des mots visuels dans l'image), facilement

exploitable par les outils de classification [6] (figure 1). La seconde consiste à établir une mesure de similarité entre les sacs de descripteurs directement, sous la forme de noyaux (noyau somme [5], noyau somme-puissance[4], etc). Cette méthode à l'avantage de fournir directement une mesure exploitable par les outils d'apprentissage statistique largement utilisés aujourd'hui, tels que les SVM[1].

3 Objectifs

Dans le cadre de la reconnaissance de catégories d'objets, la première méthode dite de *bag-of-features* (*BoF*) est aujourd'hui très populaire. Elle permet de passer d'une représentation complexe de l'image sous forme d'ensemble de descripteurs à une représentation vectorielle compacte à l'aide d'une étape de quantification de l'espace des descripteurs. La similarité entre deux images est effectuée sur les vecteurs des occurrences des descripteurs quantifiés. Dans cette méthode, on approxime les descripteurs par les centroïdes issus de la quantification. Ainsi, l'appariement ne se fait plus entre les descripteurs extraits des images, mais entre les centroïdes, présents dans les images.

A contrario, les méthodes définissant directement des mesures de similarités sur les sacs de descripteurs conservent un appariement point à point. **L'objectif de ce stage est de déterminer dans quelle mesure la conservation ou non d'un appariement point à point affecte les résultats de classification.** Pour cela, on comparera sur plusieurs bases d'images, couramment utilisées dans la littérature telles que PASCAL [2], les approches *BoF* et les approches *Kernels on Bags*. On pourra, à ce titre, proposer de nouveaux noyaux sur sacs et les inclure dans la comparaison. En outre, on essaiera de formaliser le rapprochement entre ces deux méthodes.

4 Compétences requises

Des connaissances minimales en traitement d'images et en apprentissage statistique sont nécessaires. Un bon niveau de programmation est demandé (le développement se fera en Java). Ce sujet comporte un fort côté expérimental, aussi, de la rigueur dans la conduite du plan expérimental sera-t-elle exigée.

Références

- [1] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3) :273–297, 1995.
- [2] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes Challenge 2009 (VOC2009) Results. <http://www.pascal-network.org/challenges/VOC/voc2009/workshop/index.html>.
- [3] D. Lowe. Distinctive image features from scale-invariant keypoints. In *International Journal of Computer Vision*, volume 20, pages 91–110, 2003.
- [4] Siwei Lyu. Mercer kernels for object recognition with local features. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 2 :223–229, 2005.
- [5] John Shawe-Taylor and Nello Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, New York, NY, USA, 2004.
- [6] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (in press), 2010.