Semantic Learning Methods: Application to Image Retrieval

Philippe H. Gosselin and Matthieu Cord

ETIS / CNRS UMR 8051
6, avenue du Ponceau, 95014 Cergy-Pontoise, France

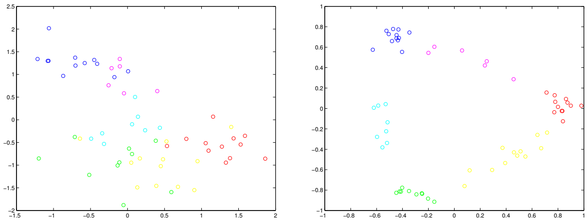
Introduction. Indexing, retrieval and classification tools are usefull to process large digital document collections. Digital documents can be automatically gathered into *clusters* or *concepts*. For instance, concepts like keywords are useful for text database organization. Concepts do not necessarily form a clustering, *i.e.* a document can belong to several concepts. Documents are usually represented by low-level features, computed automatically, and learning techniques based on relevance feedback are used to retrieve the concepts (Tong & Koller, 2000; Gosselin & Cord, 2004).

According to an incomplete set of partial labels, we propose two semantic learning methods to improve the representation of the document collection, even if the size, the number and the structure of the concepts are unknown. These methods may learn a lot of concepts with many mixed information. We build these methods in a general framework, thus powerful learning or semi-supervised learning methods may be used to retrieve, classify, or browse data.

Challenge. Suppose that we have a set of documents, each of them represented by a vector $\mathbf{x}_i \in \mathbb{R}^{N_d}$ of $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_{N_x}\}$, and a set of labels $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_{N_y}\}$. For instance, X can be the set of feature vectors of an image database, and each $\mathbf{y}_p$ contains the labels provided by a user during the retrieval session p. We suppose that labels are sampled from a hidden set of concepts. The documents are gathered in a finite (but unknown) number N_c of concepts, and these concepts do not necessarily form a clustering. Thus, a document represented by a vector $\mathbf{x}_i$ can belong to several concepts. For instance on an image database, one can find buildings, cars, houses, or landscapes, but also cars in front of a building or a house, or houses in a landscape.

A vector $\mathbf{y}_p \in [-1, 1]^{N_x}$ is a partial labeling of the set X, according to one of the concepts. Every positive value y_{ip} means that the document represented by $\mathbf{x}_i$ is in this concept, as much as y_{ip} is close to 1. Every negative value y_{ip} means that the document represented by $\mathbf{x}_i$ is not in this concept, as much as y_{ip} is close to -1. Every value y_{ip} close to zero means that there is no information for $\mathbf{x}_i$ about this concept. We also suppose that the number of non-zero values in $\mathbf{y}_p$ is small against the size of the concept. Thus, from only one $\mathbf{y}_p$, it is impossible to build the corresponding concept. The challenge is to use this set of partial labeling in order to learn the concepts.

Vector-based update. The repartition of the centers in space is important in the case of mixed concepts. As we wish to represent any possible combination of memberships, centers should be at the same distance from each other. The building of equidistant cen-



(a) (b)

FIG. 1 – Toy example with 3 mixed concepts, blue = concept 1, red = concept 2, green = concept 3, magenta = concept 1 and 2, cyan = concept 1 and 3, yellow = concept 2 and 3. (a) Initial set. (b) Concept Vector Learning method.

ters has implications on the dimension N_d of vectors. A theorem also shows, for N_c concepts in a space of $N_c - 1$ dimensions, that distance between equidistant concept centers is unique, modulo a scaling. It is easy to see that in higher dimension, this property is no longer true. In the computation of possible centers, we exploit this property in order to get equidistant centers.

Kernel-based update (Gosselin & Cord, 2005). The knowledge contained in the matrix Y can be used to update the similarity matrix. The similarity matrix is the matrix of all similarities between image pairs. The strategy is based on a kernel matrix adaptation, and is designed to model mixed concepts. We also manage the complexity constraint using efficient eigenvalue matrix decomposition ; the method has a $O(N_x)$ complexity and memory need, and so it is applicable to large databases.

Experiments. Tests are carried out on the generalist COREL photo database. Results show that the proposed method, and especialy the vector-based one, increase significally the performances of the system in comparison to distance learning methods.

Conclusion. We introduced a feature-based and a kernel-based semantic learning methods, to improve the performances of an image retrieval system. These methods deal with the constraints of the CBIR framework, and are able to enhance the database representation with a partial and incomplete knowledge about the structure of the database. Tests carried out on a generalist database show that the data representation may be improved by these learning strategies. Using the proposed learning protocol, the vector-based technique gives the best results.

Références

GOSSELIN P. & CORD M. (2004). RETIN AL : An active learning strategy for image category retrieval. In *IEEE International Conference on Image Processing*, Singapore.

GOSSELIN P. & CORD M. (2005). Semantic kernel learning for interactive image retrieval. In *IEEE International Conference on Image Processing*, Genova, Italy.

TONG S. & KOLLER D. (2000). Support vector machine active learning with applications to text classification. *International Conference on Machine Learning*, p. 999–1006.