

Incremental indexing of objects in pictorial databases

G. Castellano, A.M. Fanelli, M.A. Torsello

Computer Science Department

University of Bari A. Moro

Via Orabona, 4 - 70126 Bari, Italy

(giovanna.castellano, annamaria.fanelli, mariaalessandra.torsello)@uniba.it

Abstract

Object indexing is a challenging task that enables the retrieval of relevant images in pictorial databases. In this paper, we present an incremental indexing approach of picture objects based on clustering of object shapes. A semi-supervised fuzzy clustering algorithm is used to group similar objects into a number of clusters by exploiting a-priori knowledge expressed as a set of pre-labeled objects. Each cluster is represented by a prototype that is manually labeled and used to annotate objects. To capture eventual updates that may occur in the pictorial database, the previously discovered prototypes are added as pre-labeled objects to the current shape set before clustering. The proposed incremental approach is evaluated on a benchmark image dataset, which is divided into chunks to simulate the progressive availability of picture objects during time.

1. Introduction

The extensive use of image digital capturing systems in several fields has generated massive amount of digital images that are typically collected in pictorial databases [1]. Most of the past projects on pictorial databases focus on content-based approaches searching images that are visually similar to the query image [2]. Such approaches do not have the capability of assigning textual descriptions automatically to pictures, i.e. they do not perform linguistic indexing.

Linguistic indexing is a difficult task due to the semantic gap problem, i.e. the lack of coincidence among the visual content of images represented by automatically extracted features and the human visual interpretation of the picture content typically expressed by high-level concepts [3]. Learning concepts from images and automatically translating the content of images to linguistic terms can bridge the semantic gap thus resulting in one of the most influential factors in successful image retrieval [4], [5] consequently

broadening the possible usages of pictorial databases.

Different machine-learning methods have been applied to learn associations between the low-level features and the linguistic concepts in a pictorial database [15]. In particular, learning techniques can be used to annotate objects clearly identifiable by linguistic cues. A common approach is to perform classification on the collection of picture objects [6], [14], so that visually similar objects are grouped into the same class and a textual label is associated to each class. Thus, each object is indexed by classifying it into one of the identified classes.

Classification of picture objects can be performed by means of supervised or unsupervised learning methods. Supervised techniques require a lot of training data, and providing these data is a very tedious and error-prone task, especially for large image database. Unsupervised learning techniques overcome these limitations but often they generate inconsistent classes including objects that, although having a similar shape, actually represent different linguistic cues. The presence of objects with ambiguous shapes motivates the use of semi-supervised clustering algorithms that can improve classification by using a combination of both labeled and unlabeled data. In [11] we proposed the use of a semi-supervised clustering algorithm called SSFCM (Semi-Supervised Fuzzy C-Means) to create object classes and prototypes useful for indexing images in a database. However, when new images are added to the database, this static indexing scheme requires rebuilding the prototypes starting from scratch by reprocessing the whole set of objects, i.e. it does not take advantage of the previously created prototypes.

To overcome this limitation, in this paper we propose the use of an incremental version of the SSFCM clustering algorithm, that we call Incremental SSFCM (ISSFCM). The ISSFCM applies SSFCM to chunks of picture objects that are periodically added to the database, thus providing an incremental scheme for picture object indexing.

The paper is organized as follows. Section 2 describes the proposed indexing scheme for pictorial object annota-

tion. In section 3 we provide some preliminary simulation results on a benchmark data set containing picture objects of different shapes. Finally, section 4 concludes the paper.

2. Incremental scheme for object indexing

We assume that a collection of pictorial objects is available. Each object is described by the contour of its shape. Different shape descriptors could be used to represent object shapes. In this work, each object shape is represented by means of Fourier descriptors that are well-recognized to provide robustness and invariance, obtaining good effectiveness in shape-based indexing and retrieval [8]. The shape of each pictorial object is described by means of M Fourier descriptors and denoted by a numerical vector $\mathbf{x} = (x_1, x_2, \dots, x_M)$.

The proposed scheme for incremental indexing of objects is based on the assumption that sets of object shapes belonging to different semantic classes are available during time and processed as chunks, that is, a chunk of N_1 object shapes is available at time t_1 , a chunk of N_2 shapes is available at t_2 and so on. We denote by X_t the chunk of picture objects available at time t . For a correct application of the proposed incremental scheme, all semantic classes should be represented in the early chunks. The chunks of objects are processed as they are added to the database, by applying incrementally the Semi-Supervised FCM (SSFCM) algorithm [11] described in section 2.1. The resulting scheme, called ISSFCM (Incremental SSFCM), is shown in fig. 1. It enables the update of previously derived prototypes when new shapes are continuously available over time. Each time a new chunk of shapes is available, previously created cluster prototypes are used as pre-labeled shapes for the new run of SSFCM. At the end of each SSFCM run, the derived labeled prototypes are used to index all available shapes accumulated in the pictorial database.

The overall scheme of the proposed incremental indexing approach is summarized in algorithm 1. Each time a chunk is available, it is clustered by SSFCM and the resulting clusters are represented by K prototypes that are manually annotated by textual labels (step 4-7). Then each object is added to the cluster corresponding to the best matching prototype and labeled with the related label (step 8-9). Matching is based on computing Euclidean distance between the Fourier descriptors of the object and the descriptors of prototypes. We chose the Euclidean distance since it is one of the most popular distances in literature that permits to obtain accurate results when matching shapes represented by Fourier descriptors with a low-cost and simple computation [16]. To take into account the evolution of the database, the prototypes discovered from one chunk are added as pre-labeled objects to the next chunk (step 10, step 4).

Precisely, when the first chunk of pictorial objects is

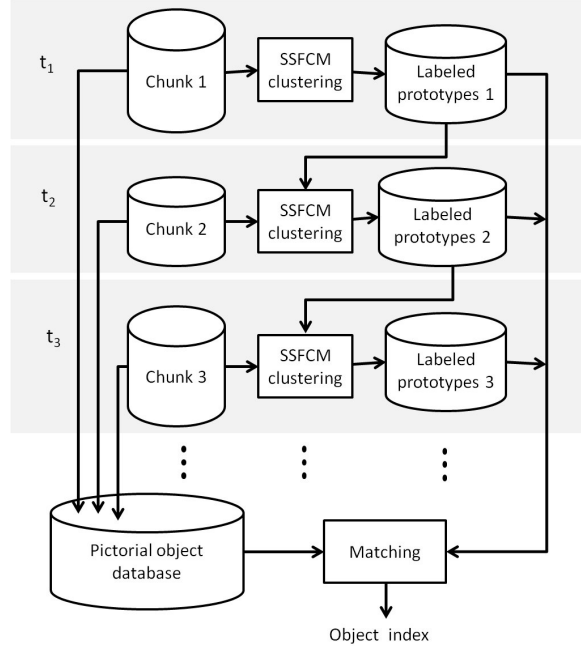


Figure 1. The scheme of the incremental indexing approach

available, the algorithm will cluster the chunk into K clusters and it will derive a set of K object prototypes that are manually labeled. When a second or later chunk of objects is available, it will be clustered with the labeled prototypes derived from the previous clustered chunks¹.

Summarizing, our incremental indexing scheme generates a structure of clusters on the basis of chunks which capture the availability of new picture objects during time and reflect physical evolution of the database. The indexing mechanism is incremental in the sense that the cluster prototypes derived from one chunk are used not only for current indexing but also as a starting point for the clustering of successive chunks. The derived prototypes offer an intermediate indexing mechanism that enables automatic linguistic indexing of pictorial objects by requiring manual annotation of a very limited number of objects (namely the prototypes).

2.1. Clustering by SSFCM

The SSFCM algorithm works in the same manner as FCM (Fuzzy C-Means) [9], i.e. it iteratively derives K clusters by minimizing an objective function. Unlike FCM, that performs a completely unsupervised clustering, SSFCM performs a semi-supervised clustering, i.e. it uses

¹How many chunks of history to use for clustering with a new chunk is predefined by the user.

Algorithm 1 Incremental SSFCM (ISSFCM)

Require: Chunks of unlabeled objects X_1, X_2, \dots **Ensure:** P : set of labeled prototypes; X : set of annotated objects

- 1: $H \leftarrow \emptyset$ /* Initialization of history */
 - 2: $t \leftarrow 1$ /* Initialization of time step */
 - 3: **while** \exists non empty chunk X_t **do**
 - 4: $X_t \leftarrow X_t \cup H$ /* Add history to current chunk */
 - 5: Cluster X_t using SSFCM
 - 6: Derive the set P of prototypes
 - 7: Annotate manually each prototype in P
 - 8: Annotate each object in $\bigcup_{\tau=1}^t X_\tau$ using the best-matching prototype in P
 - 9: Update X with annotated objects
 - 10: Update H with P
 - 11: $t := t + 1$
 - 12: **end while**
 - 13: **return** P, X
-

a set of pre-labeled data to improve clustering results. To embed partial supervision in the clustering process, the objective function of SSFCM includes a supervised learning component, as follows:

$$J = \sum_{k=1}^K \sum_{j=1}^{N_t} u_{jk}^m d_{jk}^2 + \alpha \sum_{k=1}^K \sum_{j=1}^{N_t} (u_{jk} - b_j f_{jk})^m d_{jk}^2 \quad (1)$$

where

$$b_j = \begin{cases} 1 & \text{if object } \mathbf{x}_j \text{ is pre-labeled} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

f_{jk} denotes the true membership value of the pre-labeled object \mathbf{x}_j to the cluster k , d_{jk} represents the Euclidean distance between the object shape \mathbf{x}_j and the center of the k -th cluster, m is the fuzzification coefficient ($m \geq 2$) and α is a parameter that serves as a weight to balance the supervised and unsupervised components of the objective function. The higher the value of α , the higher the impact coming from the supervised component is. The second term of J captures the difference between the true membership f_{jk} and the membership u_{jk} computed by the algorithm. The aim to be reached is that, for the pre-labeled objects, these values should coincide.

As described in [12], the problem of optimizing the objective function J is converted into the form of unconstrained minimization using the standard technique of Lagrange multipliers. By setting the fuzzification coefficient m equal to 2, the objective function is minimized by updating membership values u_{jk} according to:

$$u_{jk} = \frac{1}{1 + \alpha} \left[\frac{1 + \alpha(1 - b_j \sum_{l=1}^K f_{lk})}{\sum_{l=1}^K d_{jk}^2 / d_{lk}^2} \right] + \alpha b_j f_{jk} \quad (3)$$

and the centers of clusters according to:

$$\mathbf{c}_k = \frac{\sum_{j=1}^{N_t} u_{jk}^m \mathbf{x}_j}{\sum_{j=1}^{N_t} u_{jk}^m} \quad (4)$$

The clustering process ends when the difference between the values of J in two consecutive iterations drops below a prefixed threshold or when the established maximum number of iterations is achieved.

Once the clustering process is completed, a prototype is identified for each cluster by selecting the object shape belonging with the highest membership to that cluster. Then, each prototype is manually associated to a label corresponding to a specific linguistic cue or semantic class.

Summarizing, the result of SSFCM applied to each chunk is a set of K labeled prototypes $P = \{p_1, p_2, \dots, p_K\}$ that are used to index objects in the database. Namely, all objects belonging to cluster k are associated with the text label assigned to prototype p_k .

3. Experimental results

To assess the suitability of the proposed incremental indexing approach, we considered the MPEG-7 Core Experiment CE-Shape-1 data set [8] containing 1400 binary images of object shapes grouped into 70 different classes with each class including 20 samples. Fig. 2 shows a sample image for each class of the considered data set. In order to apply ISSFCM, all images were processed to extract boundaries of shapes and compute Fourier descriptors. Each object shape was represented by a vector of 32 Fourier coefficients (this number was set in our previous experiments on the same dataset).

To evaluate the clustering results we used the average purity error, as in [10], defined as follows:

$$pur = 1 - \frac{1}{K} \times \sum_{k=1}^K \frac{|C_k^{d_1}|}{|C_k|}$$

where K denotes the number of clusters, $|C_k^{d_1}|$ denotes the number of objects with the dominant class label in cluster k and $|C_k|$ denotes the total number of objects in cluster k . Intuitively, the purity error measures the purity of the clusters with respect to the true cluster (class) labels that are known for the MPEG-7 dataset.

We performed a suite of experiments in order to analyze the behavior of ISSFCM when varying the percentage p of pre-labeled shapes ($p = 20\%$ and $p = 30\%$) and following two different pre-labeling schema:

- scheme A: we assume that each chunk contains a percentage p of pre-labeled shapes;

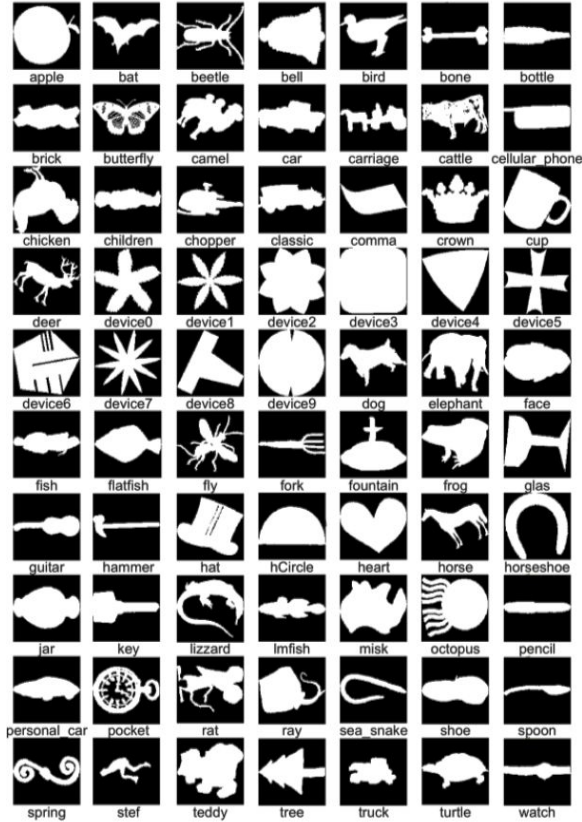


Figure 2. Sample images from the MPEG-7 Core Experiment CE-Shape-1 data set

- scheme B: we assume that only the first chunk contains a percentage p of pre-labeled shapes; in the next chunks the previously derived prototypes represent the pre-labeled shapes.

In all the experiments, the parameters of the ISSFCM algorithm were set as follows: the number of cluster K was set to the number of classes in the dataset (i.e. $K = 70$), the size of a chunk was set to 280 shapes (hence 5 chunks were built from the whole dataset), the history was set to 1, meaning that only the prototypes extracted from the previous chunk were considered as pre-labeled shapes in the current chunk. Since SSFCM is not deterministic (due to the random initialization of the cluster centers) 10 different runs were performed and the average results are presented.

At the first time step, the SSFCM was applied to the union of the first two chunks in order to obtain more stable and significant initial prototypes to be exploited in the next steps of the incremental clustering process. In this way 4 different time steps were simulated. After clustering a chunk, 70 prototypes were derived and each prototype was manually annotated by a label descriptive of a semantic

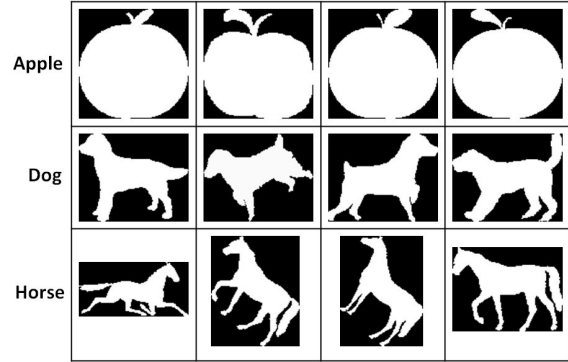


Figure 3. Prototypes derived in each time step for three semantic classes

Table 1. Average purity error values

scheme	percentage of pre-labeled shapes	
	20%	30%
A	0.29	0.18
B	0.28	0.17

class. These prototypes were used to annotate all shapes included in the previous chunks on the basis of a top-matching score. To perform matching we computed the Euclidean distance between descriptors of each shape and descriptors of each prototype. Each shape was annotated with the label of the best-matching prototype. As an example, in fig. 3, we show the prototypes derived for three semantic classes at the end of each time step by applying ISSFCM with the 30% of pre-labeled shapes following scheme B.

The annotation results were evaluated by computing the average purity error. Table 1 reports the average values of the purity error obtained by varying the percentage of pre-labeled shapes and the pre-labeling scheme. It can be seen that, as expected, when the percentage of pre-labeled shapes increases, the quality of the obtained clusters improves. Regardless the pre-labeling percentage, the two pre-labeling schema provide comparable values of the purity error.

The effectiveness of the proposed incremental approach was evaluated by comparing the average purity error obtained in the last step of ISSFCM and the average purity error obtained by applying the SSFCM algorithm in a one-shot way (following the experimental setting described in [11]). To apply the SSFCM in one-shot way the data set was divided into a training set (composed of the shapes included in the first 4 chunks) and a test set (including the 280 remaining shapes). The training set was used to derive the shape prototypes whilst the test set was used to perform annotation by exploiting the derived prototypes. Figure 4a compares the average purity error obtained by applying ISS-

FCM (varying the pre-labeling scheme and the percentage of pre-labeled shapes) and the one-shot SSFCM. We observe that ISSFCM obtains results that are comparable to those obtained by the static one-shot SSFCM with the additional advantage to exploit and update the knowledge discovered in the previous time steps. Finally, we evaluated the annotation accuracy in terms of Precision and Recall and we compared the results obtained by applying the incremental SSFCM the static SSFCM. Figures 4b and 4c show the comparative values of precision and recall, respectively. It can be seen that the incremental indexing approach achieves better annotation accuracy with respect the static one-shot approach thus confirming the benefit of exploiting previously acquired knowledge whenever new picture objects have to be added to the pictorial database.

4. Conclusions

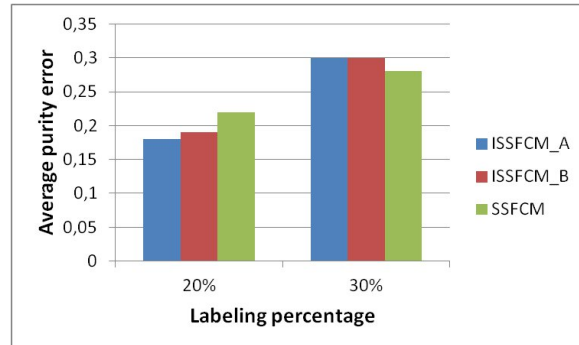
In this paper an incremental scheme for pictorial object indexing has been proposed. The approach exploits a semi-supervised fuzzy clustering algorithm to derive a set of prototypes representative of a number of semantic categories. The derived prototypes are manually annotated by attaching labels related to semantic categories. The use of shape prototypes, which represent an intermediate level of visual signatures, facilitates the annotation process, since only a reduced number of objects need to be manually annotated. Moreover, the use of prototypes simplifies the search process in a pictorial database by reducing time needed to retrieve similar shapes. Indeed, a query is matched only with shape prototypes, thus avoiding unnecessary comparisons with all objects in the database. Annotation results on the MPEG-7 benchmark dataset show that our incremental scheme obtains results which are very similar to those obtained by the one-shot approach with the additional advantage to exploit the previously discovered prototypes thus avoiding the reprocessing of the whole database. These preliminary results encourage the application of the proposed approach to real-world contexts requiring the indexing of evolving collections of pictorial objects.

References

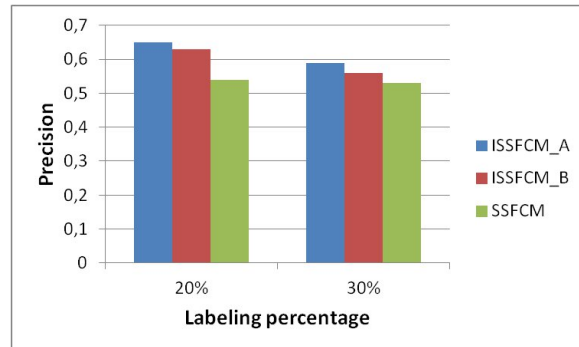
[1] S.K. Chang and T.L. Kunii. Pictorial data-base systems, IEEE 23. G. Strang, Linear Algebra and Its Applications. Harcourt, Brace, and Computer 14, pp. 13-21, 1981.

[2] Y. Rui, T. Huang, and S. Chang, Image retrieval: current techniques, promising directions and open issues, J. Visual Commun. Image R. 10(4):39-62, 1999.

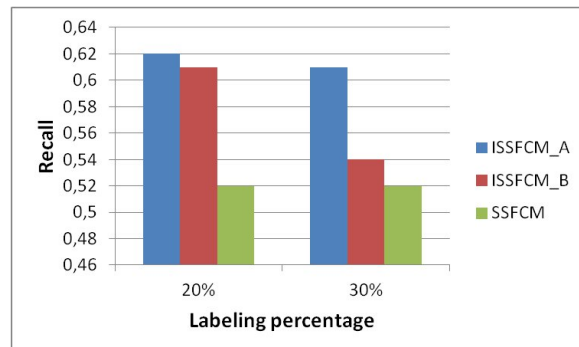
[3] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, Content-based image retrieval at the end



(a)



(b)



(c)

Figure 4. Comparison between Incremental SSFCM and one-shot SSFCM

- of the early years, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22:1349-1380, 2000.
- [4] W.I. Grosky, and R. Mehrotra. Index-based object recognition in pictorial data management. *Computer Vision, Graphics, and Image Processing*, 52(3):416-436, 1990.
- [5] J. Li, and J.Z. Wang. Automatic linguistic indexing of pictures by a statistical modeling approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25(9):1075-1088, 2003.
- [6] H. Nezamabadi-Pour, and S. Saryazdi. Object-based image indexing and retrieval in DCT domain using clustering techniques. In *Proc. of World Academy of Science Engineering and Technology*, pp. 207-210, 2005.
- [7] S. Aghabozorgi, M.R. Saybani, and T.Y. Wah. Incremental clustering of time-series by fuzzy clustering. *Journal of Information Science and Engineering* 28(4):671-688, 2012.
- [8] I. Bartolini, P. Ciaccia, and M. Patella. WARP: Accurate retrieval of shapes using phase of Fourier descriptors and Time warping distance. *IEEE Trans. on Pattern Analysis and machine Intelligence*, 27(1): 142-147, 2005.
- [9] J.C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, New York, 1981.
- [10] F. Cao, M. Ester, W. Qian and A. Zhou. Density-based clustering over an evolving data stream with noise. In *2006 SIAM Conference on Data Mining*, pp. 328-339, 2006.
- [11] G. Castellano, A.M. Fanelli and M.A. Torsello. Shape annotation by semi-supervised fuzzy clustering. *Information Sciences*, 289(24):148-161, 2014.
- [12] W. Pedrycz and J. Waletzky. Fuzzy clustering with partial supervision. *IEEE Transactions on System Man and Cybernetics*, 27(5): 787-795, 1997.
- [13] S. Guha, A. Meyerson, N. Mishra, R. Motwani and L. O'Callaghan. Clustering data streams: Theory and practice, *IEEE Trans. on Knowledge and Data Engineering*, 15(3):515-528, 2003.
- [14] D. Stan, and I.K. Sethi. Mapping low-level image features to semantic concepts. In *Proc. of the SPIE*, pp. 172-179, 2001.
- [15] D. Zhang, M.M. Islam, and G. Lu, A review on automatic image annotation techniques, *Pattern Recogn.* 45(1):346-362, 2011.
- [16] D. Zhang and G. Lu, Shape-based image retrieval using generic Fourier descriptor, *Signal Processing: Image Communication*, 17(10):825-848, 2002.