Pattern Recognition 43 (2010) 485 – 493

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Interactive unsupervised classification and visualization for browsing an image collection

Pierrick Bruneau[a,b], Fabien Picarougne[a], Marc Gelgon[a,b,*]

[a]Nantes University, LINA (UMR CNRS 6241), Polytech'Nantes rue C.Pauc, La Chantrerie, 44306 Nantes cedex 3, France
[b]INRIA Atlas Project-Team, France

ARTICLE INFO

Article history:
Received 1 July 2008
Received in revised form 7 January 2009
Accepted 2 March 2009

Keywords:
Variational Bayes
Image browsing
Interactive visualization

ABSTRACT

In this paper, we propose an approach to interactive navigation in image collections. As structured groups are more appealing to users than flat image collections, we propose an image clustering algorithm, with an incremental version that handles time-varying collections. A 3D graph-based visualization technique reflects the classification state. While this classification visualization is itself interactive, we show how user feedback may assist the classification, thus enabling a user to improve it.

1. Introduction

Content-based image retrieval has been a matter of much attention and the literature in the last decade [13]. Distinct goals may be distinguished. On the one hand, one may aim at identifying various occurrences of a single physical scene or object, making for variability of appearance [24]. On the other hand, there exists numerous needs where the target images are not initially accurately expressed, either because they are difficult to express, or because the user rather wants to explore the content of an unknown image collection. Our paper fits the latter setting, in related to a surveillance need. More precisely, our goal[1] is to provide a scheme that assists a human operator to monitor the flow of images traveling through network routers from/to a set of users that have previously identified as suspect. The elementary solution, involving display of all images, proves too tedious for the operator, thus an organized and more concise view is sought. We hence put forward a solution based on the following principles:

- The image set is organized into groups of visually similar groups, supplying a graph where image groups are nodes and vertices reflect inter-group similarity. Fig. 1 summarizes this phase.
- The collection is displayed in a way that reveals its visual structure, by means of the above-mentioned graph. Fig. 7 supplies an example screenshot.

A major point of the contribution is that the algorithmic solutions proposed for clustering and visualization can efficiently accommodate change in structure due to time-varying image collections or user feedback via the interface, i.e. low computational cost and temporal consistency of the structure is ensured.

The novelty of the system is not in the image features that are extracted, hence their description is reported with the experimental results. Rather, we select the setting of mixture models in a multivariate feature space for describing images, and the contribution operates in this context, which can in practice capture the probability distribution of feature in many applicative situations.

For instance in Ref. [9], the authors modeled color and texture segments with a Gaussian mixture. Users can build a query by choosing significant segments from an image, obtaining sorted relevant results from the database. The framework presented in Ref. [11] is closer to our work in that it builds a local cluster structure in the neighborhood of a query image. In Ref. [27], the authors rely on appropriate choice of salient features and vector quantization learning to assess hierarchical classification of images in a number of predefined semantic classes (indoor/outdoor, landscape/city, etc.). Our contribution rather proposes a method to cluster a set of images, which is about discovering some information and structure in an unstructured and unlabelled data set. Our clustering algorithm operates on Gaussian mixtures (i.e. a Gaussian mixture being an item to cluster), and to our knowledge, clustering of continuous densities has been little addressed in the literature.

* Corresponding author at: Nantes University, LINA (UMR CNRS 6241), Polytech'Nantes rue C.Pauc, La Chantrerie, 44306 Nantes cedex 3, France.
E-mail addresses: pierrick.bruneau@univ-nantes.fr (P. Bruneau), fabien.picarougne@univ-nantes.fr (F. Picarougne), marc.gelgon@univ-nantes.fr (M. Gelgon).

[1] This work is funded under the ANR Safeimage/Systems and Tools for Global Security Program.

0031-3203/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2009.03.024

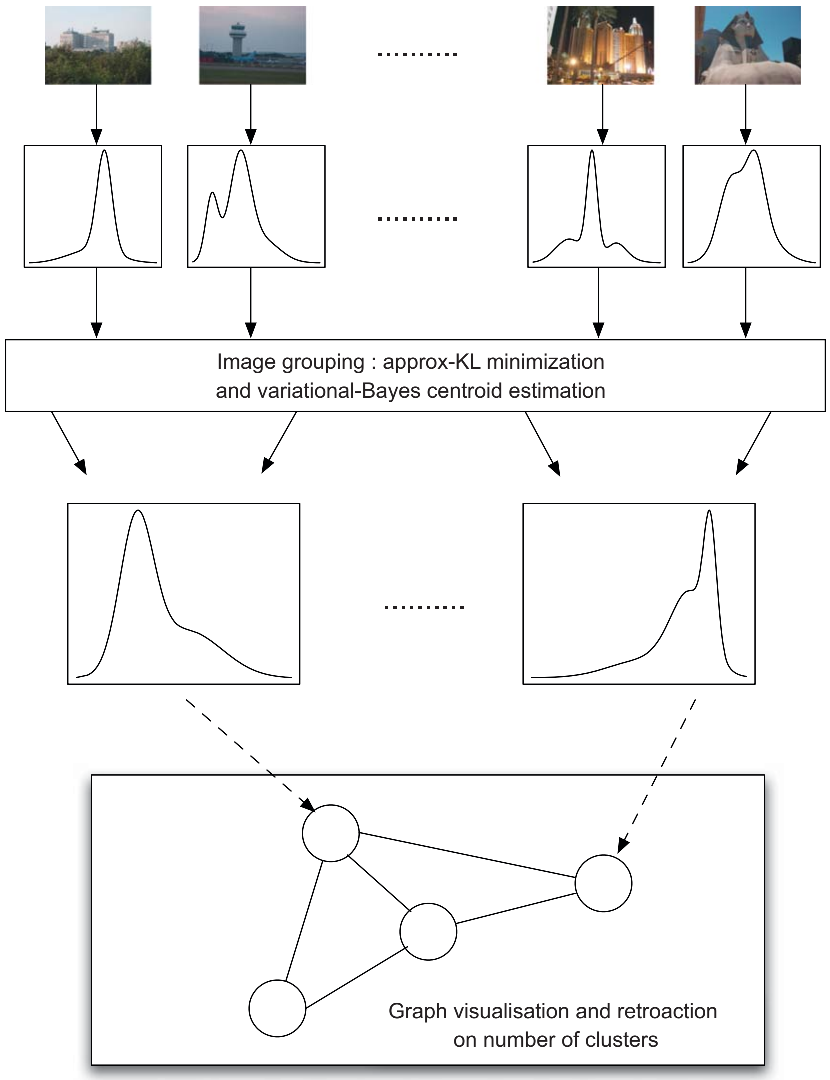

Fig. 1. Summary of the proposed approach: for each image, a mixture model is estimated. Then, images are grouped via variational-Bayes parameter-based clustering among mixture models, in an unsupervised fashion. A graph of visually similar images groups is finally built, based on distance between centroid densities of these groups. As new images flow in, the cluster structure may be updated at low cost, as updates only rely on a concise set of model parameters rather than low-level image features. Layout and interaction with the graph are discussed further down in this paper.

In the present work, user feedback does not appear in terms of improving a query, but rather capture user's intentions regarding the classifier itself. This pertains to semi-supervised learning techniques. Only few examples will be presented here, for a more comprehensive survey see Ref. [10]. Nigam et al. [25] initialized their cluster structure with a set of user-labelled examples, and estimated a mixture model on whole data using EM algorithm. Basu et al. [4] modeled constraints with hidden Markov random fields, and described a modified objective function for the k-means clustering algorithm [23] that takes these constraints into account. Their scheme also includes a learning step for a flexible distortion measure.

To reduce the cognitive load of the user, we only show, at the beginning of the visualization process, the evolution of the distance between each cluster. We construct a complete graph where nodes represent clusters. To give users an explicit view to the content of each cluster, and meanwhile limit the amount of information on the screen, each node displays few thumbnails of its most representative images. The weights of edges relate directly to the distance between clusters. During the clustering process, as the number of clusters and the distance between clusters evolve, the graph is updated accordingly by a spring-based algorithm [19,15].

Visualizing a complete graph is a difficult task. However, since the number of cluster produced by our algorithm is relatively small (maximum 20 clusters), we allow the user to set a distance threshold for edge filtering, as long as the graph remains connected. We also let the user interact with the visualization by viewing all the images of a cluster, rotating the visualization in order to better understand the relations between clusters, and provide feedback to the clustering algorithm by removing or adding nodes.

The remainder of this paper is organized as follows. In Section 2, we describe mechanisms and baseline algorithms used to build clusters from an image set. In Section 3, we expose a dynamic visualization system that provides user-appealing interaction with the cluster structure. Section 4 contains some experimental evaluations on a real image set. Finally, in Section 5 we draw some conclusions and perspectives.