Information Sciences 289 (2014) 148–161

Contents lists available at ScienceDirect

Information Sciences

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

Shape annotation by semi-supervised fuzzy clustering

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

Department of Informatics, University of Bari "A. Moro", Via Orabona, 4, 70126 Bari, Italy

ARTICLE INFO

Article history:
Received 29 February 2012
Received in revised form 20 May 2014
Accepted 8 June 2014
Available online 6 August 2014

Keywords:
Fuzzy shape annotation
Fuzzy shape clustering
Image annotation
Semi-supervised clustering
Shape representation

ABSTRACT

Image annotation is an important and challenging task when managing large image collections. In this paper, a fuzzy shape annotation approach for semi-automatic image annotation is presented. A fuzzy clustering process guided by partial supervision is applied to shapes represented by Fourier descriptors in order to derive a set of shape prototypes representative of a number of semantic categories. Next, prototypes are manually annotated by attaching textual labels related to semantic categories. Based on the labeled proto-types, a new shape is automatically labeled by associating a fuzzy set that provides membership degrees of the shape to all semantic categories. The proposed annotation approach provides an innovative indexing method for shape-based image retrieval. Indeed, shape prototypes represent an inter-mediate indexing level that allows a faster retrieval process since a query is matched against prototypes, instead of the whole shape database, resulting in a speed up of the retrieval. The proposed approach is tested on synthetic and real-word images in order to show its suitability.

1. Introduction

The exponential diffusion of large digital image collections has given rise to an increasing interest for automated tools able to efficiently store, organize and retrieve images. This has led to the growth of a promising research area referred as Content-Based Image Retrieval (CBIR) that involves the design and the development of systems able to exploit low-level visual features such as texture, color and shape for the indexing and retrieval of images [18,9].

The extraction of low-level features related to the visual content of images is quite an easy and fairly direct task. Moreover, the use of low-level features enables the definition of suitable similarity measures to develop effective matching methods useful in CBIR processes. However, many users find it difficult to formulate search queries by directly specifying low-level visual attributes. Indeed, humans tend to recognize images and to express their content relying on high-level concepts, i.e. they usually formulate their queries in natural language by employing semantic concepts. Thus, one of the main problems that heavily affects performance of CBIR systems is represented by the "semantic gap", defined as "the lack of coincidence between the information automatically extracted from the visual data and the semantic meaning, i.e. the interpretation that the same visual data have for a user in a given situation" [31].

In the literature a variety of approaches have been proposed in order to uncover the hidden correlation between low-level visual features and high-level concepts [21,15]. Among these, a straightforward approach is image annotation also known as image labeling that consists in associating textual descriptions to visual content of images providing the fundamental ability to index images semantically [13,10]. Manual annotation of images is a subjective, time-consuming and error-prone process.

* Corresponding author.
E-mail addresses: castellano@di.uniba.it (G. Castellano), fanelli@di.uniba.it (A.M. Fanelli), torsello@di.uniba.it (M.A. Torsello).

http://dx.doi.org/10.1016/j.ins.2014.06.025
0020-0255/

Hence, increasing research efforts have been addressed towards the definition of automatic methods for image annotation based on low-level visual content [35,11,22]. Broadly speaking, image annotation can be defined as the process that assigns textual labels to images in order to describe the semantic categories expressed in the image. In image annotation several issues require to be faced such as the choice of low-level features to describe the image visual content, the definition of the semantic categories and the association of a new image to the detected categories.

In particular, the choice of low-level features is a central tenet in image annotation. Several works have proved that visual features such as color, texture, and positioning, though important, are insufficient to convey the information that could be obtained through shape analysis of objects contained into images [3,19,30,32]. Shape plays a critical role for the representation of objects contained into images becoming a key feature exploited in many applications of computer vision and image understanding for indexing and retrieval purposes. Indeed a considerable amount of information is contained in the boundaries of the objects, thus a definite opinion of the scientific community is that shapes should be considered as an essential mean to describe objects in an image. Moreover, it is recognized that shape is an important feature strictly related to human perception. In fact, most users tend to perceive images as composed of individual objects identified by their shapes and they usually retrieve images based on shapes of objects contained in images by giving examples or by sketching [23,4].

As concerns the definition of semantic categories, in many cases the number of semantic categories is not known in advance. Experts often may have some kind of knowledge about the image domain such as the membership of a number of shapes to specific semantic categories. Of course, the effectiveness of the overall image annotation process can benefit from methods that are able to exploit such kind of domain knowledge in order to automatically discover a number of semantic categories along with the corresponding visual prototypes. Typically, the association of a shape to the semantic categories is a crisp process, i.e. a shape is exactly classified into one semantic category [27,33]. However, due to the presence of noise and ambiguity in image features, it is very difficult or even impossible to classify shapes precisely into one category. Therefore, the uncertainty underlying the annotation process can be suitably captured by labeling shapes with multiple keywords leading to a fuzzy annotation process that assigns multiple labels to a shape together with values encoding the membership degree of the same shape to each semantic category [8,7,1].

To address all these aspects, in this paper we propose a fuzzy shape labeling approach for image annotation. The key idea on the basis of our approach is to equip fuzzy clustering with a partial supervision mechanism that is able to provide a useful guide during the process of automatic discovery of visual shape prototypes representative of a number of semantic categories. Semi-supervised fuzzy clustering allows to identify semantic categories by grouping together similar shapes taking advantage from the domain knowledge expressed in terms of few labeled shapes. Such knowledge may be acquired in different ways. For example, this can be explicitly supplied by a domain expert or automatically derived from the available shape collection as better explained later in this work. After clustering, for each cluster a prototypical shape is determined. Then, each prototype is manually associated to a textual label corresponding to a specific semantic category. Finally, to label a new shape, its visual descriptors are matched with visual descriptors of all prototypes and the similarity values are used to create a fuzzy set expressing the membership degrees of the shape to each identified semantic category. In this way, the proposed approach creates a two-level index structure where the first layer is represented by the database shapes and the second layer is represented by shape prototypes representative of the different clusters formed by similar shapes. This gives an advantage in terms of speed when performing shape-based image retrieval. Actually, the prototype layer acts as a primary filter that reduces the search space quickly in the retrieval process avoiding comparison between the query and all shapes. A shape query is matched only with the derived prototypes (and not with all the database shapes) and shapes belonging to clusters corresponding to the top-n similar prototypes are provided as retrieval result. Therefore the proposed annotation approach provides an innovative and efficient indexing method for shape-based image retrieval.

The remainder of the paper is organized as follows. In Section 2 we provide a brief overview of works adopting shape clustering approaches for annotation and classification. Section 3 describes the proposed approach by detailing the involved steps. Experimental results obtained by testing our approach on synthetic and real-world images are presented in Section 4. Finally, conclusive remarks are drawn in Section 5.

2. Shape clustering: a brief overview

Annotation of images on the basis of shapes of objects requires an automatic mechanism to associate high-level concepts (i.e. semantic categories) to low-level shape features. The association of semantic concepts to images can be essentially viewed as a classification process. For each shape, the membership to a semantic category has to be derived according to certain similarity measures. Then the shape has to be classified into one of the considered categories. According to the current literature, this process can be performed by means of supervised or unsupervised learning algorithms. The main goal of supervised learning consists in predicting a semantic category label for an object based on its shape descriptors by exploiting information about previously labeled shapes. In unsupervised learning no outcome label is available and the aim is to discover how shape descriptors can be clustered on the basis of some similarity criterium. At the end, each resulting cluster can be associated to a category label. The success of supervised techniques is affected by several factors. To obtain good classification performance, they require a large number of labeled training data, but providing these data is a very tedious and