Available online at www.sciencedirect.com
ScienceDirect
Knowledge-Based Systems 20 (2007) 266–276
Knowledge-Based SYSTEMS
www.elsevier.com/locate/knosys

Using multiple uncertain examples and adaptative fuzzy reasoning to optimize image characterization

Luigi Lancieri *, Larbi Boubchir

France Telecom R&D, University of Caen, France

Received 20 October 2005; accepted 6 May 2006
Available online 12 September 2006

Abstract

This article proposes an automatic characterization method by comparing unknown images with examples more or less known. Our approach allows to use uncertain examples but easy to obtain (e.g. by automatic retrieval on the Internet). The use of fuzzy logic and adaptive clustering makes it possible to reduce automatically the noise from this database by preserving only the examples having a strong level of redundancy in the dominant shapes. To validate this method, we compared our artificial process of recognition with the estimation of human operators. The tests show that the automatic process gives an average accuracy of the characterization near to 95%.

Keywords: Uncertainty; Data mining; Collective intelligence; Probabilistic reasoning; Vision

1. Introduction

The increasing success of image-based multimedia content (images, video) on the information systems focuses on the need to improve the characterization methods of such content. As a comparison, the textual contents indexing technique that dates from the sixties with vector space [1,2] oriented work still has some limitations. For example, even if actual search engines become more and more effective, one can still experience difficulties in rapidly finding good results, since they are often numerous and not always pertinent. The problem is much higher with image-oriented objects since their semantic translation is more difficult and carries much more inaccuracy and ambiguity. Furthermore, artificial perception involves higher algorithmic complexity which is time and computer resources consuming. In order to shortcut this difficult step of semantic translation (i.e. from image to words), some researchers have imagined the "query by example" method where the intermediate word space is less necessary.

This method compares an unknown image with images where the content is perfectly known (human-based semantic translation). According to a measurement of visual similarity, one considers that the potential example image resembles or not the model. The many works already completed in this field (see state of the art) start most of the time from the need of having a well-controlled and purified base of examples. This stage which is primarily manual is often difficult and expensive since the contents of the images are often difficult to characterize.

More generally, the proper characterization of an object of information, textual or not, needs to take into account the interaction between the user and the object. In other words, the question of what the object "says" (expressive capacity) to the observer is as important as what the observer can understand from what the object says (perception limits). The ambiguity has born from that situation of description–perception gap. Starting from one's own experience each individual can potentially have his/her own perception.

* Corresponding author. Tel.: +33 02 31 75 91 16; fax: +33 02 31 73 56 26.
E-mail address: luigi.lancieri@orange-ft.com (L. Lancieri).

0950-7051/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2006.05.018

The practical impact of this well-known philosophical discussion in our study is the necessity to take into account the different possible points of view of an object description (i.e. the most common meaning extractable from several sets of similar shapes representing the same object). A good description of an object (with stronger reasons for a complex image) should take into account all points of view as a "network of perception". We introduce this concept in order to manage the ambiguity problem in artificial perception. A network of perception is a set of images or shapes having a reasonable probability to be a specific object. Even if some of these images are ambiguous (i.e. could be associated with other objects), its association with the reference object is confirmed by the other images from the network.

In basic words, an image has a better probability of really being, for example, a cylinder (see Fig. 1) if it is similar to at least one image having a good probability of being a cylinder. In the first part of the Fig. 1 the unknown object still has the probability of being a cylinder where as this probability is very low with a single perception (second part of the figure). This approach consistent with the human cognitive process [3], replaces the difficulty of having one precise visual description of an object (which is impossible from our point of view) to the necessity of having a set of visual even imprecise description. So that, and contrary to the classical approach, we think that characterization or classification by examples could be improved using a wide set of examples.

Our method makes it possible, automatically, to retrieve reference images in rough databases via key words (e.g. through a search engine on the Internet). This operation is easy to manage but the result is poor, for the collected images contain a high level of noise. Some do not even correspond to the target keywords. One of the features of our work is to be able to exploit these basic kinds of vague examples with good results.

The weak control that we have of these reference contents forces us to use adaptive operation and comparison algorithms. In particular, we adapted the principles of fuzzy clustering to take into account a non-fixed number of input variables (component of the perception network) for the fuzzification. The goal here is to automatically identify the most representatives examples (i.e. the most redundant shape) for the comparison with the unknown images. In other words we try characterizing unknown images thanks to a set of references a few less unknown.

To validate this approach through a human assessment, we automatically collected 200 images supposedly (see Section 3.1) corresponding to eight different topics (key words). Each group of 25 images, more or less ambiguous, was used as examples for its topic (rough network of perception). Then, the comparison process took these 200 images and for each one evaluated the level of membership with each of the 8 topics. The results were then compared with those provided by a group of human testers. To be efficient, a method of characterization should generate a membership for the 200 images to the eight topics consistent with the human assessment.

Fig. 2 introduces the different stages of the characterization process. Stages 1–4 correspond to the automatic constitution of the example database for each topic. This database finally (stage 4) contains the images describing the best possible each topic.

Stages 5 and 6 correspond to the presentation of the unknown image for the comparison (stage 7). The result of the comparison (stage 8) provides the unknown image with a membership probability with each topic. On the right side of Fig. 2 (stages 9 and 10), we can see the validation phase per comparison with the human estimate. We applied this global process to five experiments each one corresponding to stages 3, 4 and 7 different, the other stages remaining identical. The five experiments are thus characterized by more or less complex methods of noise reduction in the set of examples and comparisons with the unknown images (Fig. 3).

In Section 2, we present the general philosophy of our approach including a survey on similar works. Section 3 describes in details the stages as showed in Fig. 2. The results of our experiment and the method of validation are presented in Section 4 before to conclude in Section 5.

2. State of the art and discussion

The problem studied in this article seems close to that posed by the field of "Query By Visual Examples" (QBVE) or that of "Content Based Image Retrieval" (CBIR). These methods employ an image as reference to seek results that express a good visual similarity. In fact, our study poses the opposite problem: from a whole set of statistically known

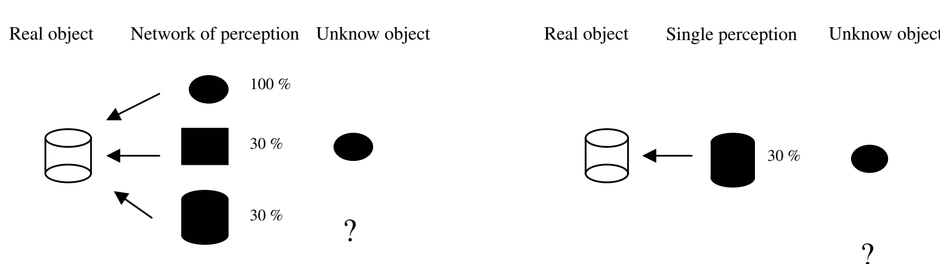

Fig. 1. The advantage of the network of perception.