A Similarity Retrieval Algorithm for Natural Images

Natalia Vassilieva, Boris Novikov

St.Petersburg State University, Russia
vnat@nv10381.spb.edu, borisnov@acm.org

Abstract. Content-Based Image Retrieval (CBIR) has become one of the most active research areas in the past years and still is known as a difficult task. Using a retrieval system is generally frustrating for users, due to a gap between low-level features managed by system and image semantics. We propose in this paper a general point of view for introducing a bridge between the user and the system. This includes a textual query, image features, based on visual perception models, a relevence feedback.

Keywords: CBIR, Image retrieval, Content analysis, Visual perception

1. Introduction

The use of images for illustration has always had a wide distribution in human activities. Thanks to recent technological advances the existence of large digital image databases became possible. The growing size of the contemporary image collections has created the necessity of image retrieval systems. This reinforced the efforts of researches and leaded to appearance of various approaches. The earlier retrieval methods are based largely on the semantic keywords attached to the images. This can be done either by manual annotation or by automatically extracting the keyword from the context, when it exists. The first one is subjective and very time consuming, the second one is not always possible. More recent works propose automated Content-Based Image Retrieval (CBIR), based on the "low level features" (color, texture, etc.) extracted from pixel values.

The preferred mode of querying in image database is semantic. For example, we might search for images of a road in a forest. To satisfy such a query, the system must be able to recognize roads and forests in the images. But this level of interpretation in CBIR systems is still out of the question. As the system can estimate the similarity of images based on their low level features, there is an important "semantic gap" between the image features, which have been extracted, and the semantics of the image.

Most widely used kind of query in CBIR systems is a query by example (QBE), formulated by providing an example of a similar image. The system returns a set of images estimated as similar by image features.

2. Related work and motivation

Literature shows a huge amount of various techniques that have been applied to CBIR in the recent years. (In [4] you can find a good review of them.) The great number of different approaches can be explained by a wide variety of application domains and by the fact that the existing CBIR systems are still not as efficient as the user want them to be.

Typically, a CBIR system finds images from a large data collection that visually match to a given query. The most systems propose to define the query by providing one or more example images. Another possibility is to provide a rough sketch of the desired image [2]. But both of the ways leads to the inconsistency between the semantic query that the user has in mind and its description which makes it hard for the user to specify the query and for the system to return the correct images.

Existing approaches based on treatment of low level features can be classified by:

- considered features,
- used data models,
- applied similarity metrics.

Among the features the color is the most extensively used. It's meaningful in human perception, easy to extract and robust to noise, scaling and rotation. Majority of data models for representation of color features vary between different kinds of histograms [8, 9, 5] and statistic models of color distribution. Texture features are also widely used in image retrieval. For texture analysis [2, 3, 6] use wavelets, other approaches use synthesized banks of filters for texture extraction (see [10] for a full review). Shape features attract less attention of scientific community. Retrieval by shape is useful only in specific collections (e.g. items on a homogeneous background or geometrical images). See [1] for review of shape representation and retrieval approaches.

Many systems provide the possibility to combine or select between one or more models, based on different features. But they typically process the features independently and use several indices, which can increase both space and time requirements.

Among more recent works there are those that propose a kind of "semantic bridge" between the user and the system. The widely used mechanism of "relevance feedback" ([5], for example) takes into account user's satisfaction by iterative process of user-system interaction. During the retrieval process, the user high-level query and perceptual subjectivity are captured by dynamically refined queries based on the user's feedback.

3. Problem definition

We consider a database of natural images, where no additional semantic information about images is available. Research is performed in the following directions:

- provide a textual way of query definition, which is more clear and intuitive to the user than the query by example;

- select the most meaningful features for natural images and propose the unique model for its representation;
- design of indexing algorithms, that agree with human visual perception;
- use clustering to reduce the number of images to process in the image retrieval phase and accelerate the retrieval.

4. Proposed solution

The complete image storage and retrieval process is divided into three phases. During the first one we define the training image set, which is expected to be much smaller than the image database itself. This set is divided into the groups of similar images by performing the classification task. For each of the groups we compute an average feature-vector from feature-vectors of group images. Feature-vector for an image is computed from its low-level features. We describe each group also by textual keywords.

The second phase serves to compute feature-vectors for the rest set of images and perform data clustering in order to the distance between an image feature-vector and average feature-vectors of the groups.

In the third phase of image retrieval we compare textual query of the user with textual descriptions of the clusters and define the most suitable cluster of images. Predefined number of images randomly selected from this cluster and returned to the user. After that the iterative process of "relevance feedback" is used: in each iteration user marks right and wrong images among the returned set. The system uses this information for query refinement.

5. Conclusions

The paper describes general ideas for centered user approach for image retrieving. These are the first steps, which are very promising and must be developed to obtain further results.

References

[1] Fan Shuang. Shape representation and Retrieval Using Distance Histograms. *PhD thesis, University of Alberta*. October 2001.

[2] Jacobs C. E., Finkelstein A., and Salesin D. H. Fast multiresolution image querying. *ACM International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH), pp 277–286*, Los Angeles, CA, August 1995.

[3] Kingsbury Nick. Image processing with complex wavelets. *Phil. Trans. Royal Society London A, vol. 357, pp. 2543-2560*. 1999.

[4] Rubner Y. Perceptual Metrics for Image Database Navigation. *PhD thesis, Stanford University*. May 1999

[5] Rui Yong, Huang Tomas, Ortega Michael, Mehrotra Sharad. Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval. *IEEE Transactions on Circuits and Video Technology*. 1998.

[6] Siggelkow Sven. Feature histograms for content based image retrieval. *PhD thesis, Albert-Ludwigs-Universitat Freiburg im Breisgau*. 2002.
[7] Simoncelli Eero, Portilla Javier. Texture characterization via joint statistics of wavelet coefficient magnitudes. *Proceedings of Fifth International Conference on Image Processing, vol I*. Chicago, IL, 4-7 October 1998.
[8] Stricker Markus, Orengo Markus. Similarity of color images. *In Storage and Retrieval for Image and Video Databases III, SPIE 2420, pages 381--392, San Jose, CA*. February 1995.
[9] Swain M. J. and Ballard D. H. Color indexing. *Intern. Journal of Computer Vision 7(1), pp. 11-32*, 1991
[10] Tuceryan Mihran, Jain Anil. *Texture analysis. The Handbook of Pattern Recognition and Computer Vision (2nd Edition), by C. H. Chen, L. F. Pau, P. S. P. Wang (eds.), pp. 207-248*, World Scientific Publishing Co., 1998.