Expert Systems with Applications 39 (2012) 2385–2396

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Efficient content-based image retrieval using Multiple Support Vector Machines Ensemble

Ela Yildizer [a], Ali Metin Balci [b], Mohammad Hassan [c], Reda Alhajj [d,*]

[a] Microsoft Inc., Bellevue, WA, USA
[b] Turkish Navy Research Center Command, Perndik / İstanbul, Turkey
[c] Zarqa University, Zarqa 13110, Jordan
[d] Department of Computer Science, University of Calgary, Calgary, Alberta, Canada

ARTICLE INFO

Keywords:
Content based image retrieval
Support Vector Machine
Wavelet transformation
Multimedia

ABSTRACT

With the evolution of digital technology, there has been a significant increase in the number of images stored in electronic format. These range from personal collections to medical and scientific images that are currently collected in large databases. Many users and organizations now can acquire large numbers of images and it has been very important to retrieve relevant multimedia resources and to effectively locate matching images in the large databases. In this context, content-based image retrieval systems (CBIR) have become very popular for browsing, searching and retrieving images from a large database of digital images with minimum human intervention. The research community are competing for more efficient and effective methods as CBIR systems may be heavily employed in serving time critical applications in scientific and medical domains. This paper proposes an extremely fast CBIR system which uses Multiple Support Vector Machines Ensemble. We have used Daubechies wavelet transformation for extracting the feature vectors of images. The reported test results are very promising. Using data mining techniques not only improved the efficiency of the CBIR systems, but they also improved the accuracy of the overall process.

1. Introduction

The evolution in digital technology has dictated a shift in handling images and there has been a significant increase in the number of images stored in electronic form. In other words, it may be argued that printed images are turning into history after the introduction of digital images capturing facilities. Libraries and collections of digital images are dominating different aspects of our daily life. Users are browsing and searching multimedia databases (Powell, 2006), to retrieve the most relevant multimedia resources they are looking for Wang, Li, and Wiederhold (2001). Therefore, the research community has realized the need to develop an approach which finds similar multimedia resources in a reasonable amount of time and with high accuracy. In this paper, we concentrate on image databases and hence tackle the problem of developing effective and efficient ways for image retrieval. With printed images people employed different techniques that could facilitate effective manual retrieval. These techniques include classifying and annotating images. The same ideas have been adopted to deal with electronic images and hence efficient and effective algorithms are the target.

In general, images could be classified into two classes, texture and non-texture. Texture images form an important class, where an object within the image is repeated periodically throughout the image. Some medical images such as X-rays and some topographic images fall under this category. Non-texture images tend to have objects of interest clustered in one or more regions of an image. Most of the real world images that people are familiar with fall under the second category. In this study, we focus on non-texture images; they are more challenging to handle.

Most traditional and common methods of image retrieval utilize some method of adding meta data such as captioning, keywords, or descriptions to the images so that retrieval can be performed over the annotation words. However, adding meta data to every picture in a large database is done by human experts and it is very time consuming. Moreover, these keywords may not reflect the real semantics of the image and may mostly reflect one viewpoint which is that of the expert involved in the process. Consequently, a legitimate question to raise here could be articulated as follows: do we really need to have the search capabilities limited by what is reflected into the meta data? Of course no; rather an image can describe itself more effectively than 100 keywords can describe it. Regarding these disadvantages of traditional methods, content-based image retrieval (CBIR) systems have attracted the attention of many researchers. The term "Content-based"

* Corresponding author.
E-mail addresses: elyildiz@microsoft.com (E. Yildizer), alhajj@ucalgary.ca (R. Alhajj).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.08.086

means that the search will analyze the actual contents of the image. The word 'content' in this context might refer to color, shape, texture, or any other information that can be derived from the image itself. While performing content-based image retrieval, comparing every pixel value with other picture's pixel value is very costly besides being wasteful. It is very probable that we find large distance measures between similar images, when the orientation of the image is changed or small details of the pictures are different. Therefore, it is very important to extract a compact, informative and accurate feature vector from each image. The search algorithm can use the feature vectors to locate the target image(s).

As we have already mentioned, the size of multimedia databases has been increasing continuously. It is clear that making a linear search in the whole database for similarity search during the image retrieval process is not acceptable. It is very important to narrow down the search space without missing relevant images. In this paper, we are proposing two different methods in which we are integrating data mining (Maimon & Rokach, 2005), wavelet transform (Li, Li, Zhu, & Ogihara, 2002; Timoney, 2000) and indexing techniques (Lightstone, Teorey, & Nadeau, 2007; Zhang & Alhajj, 2010) into the automated search process leading to effective and efficient content-based image retrieval systems. Wavelets are employed for feature extraction and hence to help construct the feature vector per image.

Our motivation for developing the proposed approach was to find a good similarity measure between images, which is one of the major difficulties of image retrieval systems. Similarity between two images is a subjective decision and many researchers have used class labels of the images during the evaluation of image retrieval systems. We have the class labels of images and can make use of this valuable information. If two images are said to be similar when they belong to same class, obviously we can say that similar images belong to predefined classes with close probabilities. In our approach, we are using Support Vector Regression (SVR) model to find the class probabilities. Since SVR is mainly defined for 2 class problems, we are using multi-class SVR ensemble to handle multi-class image databases. This approach reduces dimensions of feature vectors dramatically while enabling us to find a good similarity measure between images.

Thus, the main target of this research project is to develop a general purpose CBIR technique that can handle large image databases and can be smoothly embedded into different image retrieval systems. The following list outlines the main contributions of this paper.

1. The proposed approach is decomposable and the ideas in this paper can be embedded into other projects. For example, one can use different feature extraction algorithms and use our models for narrowing down the search space or for reducing the dimensions of the feature vector.
2. The proposed approach is scalable. When we developed our system efficiency was one of our main priorities and we reduced the retrieval cost by minimizing the I/O costs, by narrowing the search space and by reducing the dimensionality of the feature vector.
3. Multiple Support Vector Machines (Gidudu, Hulley, & Marwala, 2007; Wang, 2001; Wang & Kecman, 2005; Zhang, Liu, Zhu, & Hu, 2007) are used for classification and for reducing the dimensionality of the feature vectors. Our approach to classify images has given excellent results while we report a huge performance gain when we reduce the dimensionality of the feature vectors.

The rest of the paper is organized as follows. Section 2 provides the background information necessary to understand the proposed approaches. We mainly cover the algorithms and data structures necessary to understand the different aspects of the CBIR systems we have developed; this section also includes an overview of related past work on CBIR systems. Section 3 covers the proposed CBIR system that we have developed using Multiple Support Vector Machines Ensemble. Accuracy and experimental results of this system are discussed in Section 4. Section 5 is conclusions future research directions.

2. Related work

The development of automated CBIR systems has been an attractive research area due to its wide range of applications in critical fields like bioinformatics and medical imaging, space images, etc. There are many CBIR approaches described in the literature, e.g., (Chang & Jay Kuo, 1993; Howarth & Rüger, 2004; Kubo, Aghbari, & Makinouchi, 2003; Sun & Ozawa, 2003; Sklansky, 1978); the two papers (Veltkamp & Tanase, 2000; Zachary & Iyengar, 1999) include good surveys of CBIR systems. In this section, we will briefly overview some of the already developed systems as described in the literature.

QBIC (Niblack, Barber, Equitz, Flickner, Glasman, et al. (1993)) is one of the most-well known and earliest content-based image retrieval systems. It works both with image and video databases. It uses the contents of the images and videos such as color, shape, layout and texture information. One of the most important properties of QBIC is that it supports contents of the images and user drawn graphics in its queries. In other words, a query is not restricted with images themselves but their contents can also be queried. For example, a user can enter a query image and search for the images who have the similar texture features to the query image's texture. QBIC technology has been developed at IBM and now being used commercially.

The VIR Image Engine (Bach, Fuller, Gupta, Hampapur, Gorowitz, et al. (1996)), developed by Virage Inc., is similar to QBIC in the sense that it supports querying by color, shape, layout and texture. As in the QBIC, user drawn sketch queries are also supported. It provides a GUI where developers can sketch the images or make modifications on the images. It also supports integration of image keywords into the query (Uhlmann, 1991). VIR Image Engine has been integrated into several databases; it is a component of the Oracle DBMS.

Multimedia Analysis and Retrieval System (MARS) (Huang, Mehrotra, & Ramchandran, 1996) project was started at the University of Illinois to develop an effective multimedia database management system. As a first step in that project, they had developed an image retrieval system. In this retrieval project, image features are extracted from the content of the image and the textual annotation of the images. Content of the images are stored as a global color histogram, a texture histogram and shape features. Shape features are obtained after segmentation of the image and applying Fourier descriptor to each of the segments. User can query for any combination of the image features. For example a user can enter a query which contains an image's color histogram and another image's shape feature.

Photobook (Pentland, Picard, & Sclaroff, 1994) project was developed at MIT and uses the features of the images for comparison. Features of the images contain information about the image's color, shape and texture. Features are compared according to any linear combination of the following distance measurements: Euclidean, mahalanobis, divergence, vector space angle, histogram, Fourier peak, and wavelet tree distances. Photobook also lets developers to load their own distance measurement.

In WBIIS Wang et al. (2001) developed an image retrieval system using Daubechies wavelet transformation. They first represent the Red (R), Green (G) and Blue (B) components of the images in a different domain which reflects human perception better. Then, they apply 4-level and 5-level Daubechies wavelet transformation on the images and store the low frequency components of the resulting