A Novel Image Retrieval System Based On Dual Tree Complex Wavelet Transform and Support Vector Machines

Adeel Mumtaz, S.A.M. Gilani, Tahir Jameel
Faculty of Computer Science and Engineering
Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi, Pakistan
{adeel, asif, teejay}@giki.edu.pk

Abstract— **This paper presents a novel image retrieval system (SVMBIR) based on dual tree complex wavelet transform (CWT) and support vector machines (SVM). We have shown that how one can improve the performance of image retrieval systems by assuming two attributes. Firstly, images that user needs through query image are similar to a group of images with same conception. Secondly, there exists non-linear relationship between feature vectors of different images and can be exploited very efficiently with the use of support vector machines. At first level, for low level feature extraction we have used dual tree complex wavelet transform because recently it is proven to be one of the best for both texture and color based features. At second level to extract semantic concepts, we grouped images of typical classes with the use of one against all support vector machines. We have also shown how one can use a correlation based distance metric for comparison of SVM distance vectors. The experimental results on standard texture and color datasets show that the proposed approach has superior retrieval performance over the existing linear feature combining techniques.**

***Index Terms*—complex wavelet transform (CWT), support vector machine (SVM), content based image retrieval, texture image retrieval.**

I. Introduction

Digitization has made a profound effect in our everyday life from HDTV to digital camera. Global village concept is truly visualized in todays filed of communication where boundaries of many independent technologies have been merged to a common technology. With the rapid development of computing hardware, digital acquisition of information has become one popular method in recent years. Every day, G-bytes of images are generated by both military and civilian equipment. Large set of medical images, architectural and engineering designs, journalism and advertising, are worth mentioning. Consequently, how to make use of this huge amount of images effectively becomes a highly challenging problem [1]. Historically, only way to search through these collections was text based. Images were first annotated using text and then traditional Database Management Systems (DBMS) were used to retrieve relevant images when required.

There were two main problems with this approach; at first the amount of labor involved in manually annotating these images and secondly the inherent complexity and richness of image contents were making the annotation process difficult. For example it was not easy to label two images containing similar objects but with different meanings. In order to overcome these problems and to automate the retrieval process researchers from the field of computer vision proposed a new idea [2], which is known as content based image retrieval (CBIR). CBIR is also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR). In these days many commercial and research CBIR systems are available. IBM's QBIC [7] and UC Berkeley's Blobworld are well known. Detailed comparison of such systems can be found at [3].

In content based image retrieval, we automatically extract features from images and then compare images using these features. Images having similar features would have similar contents. Basic block diagram of a content based image retrieval system is shown in Fig. 1.

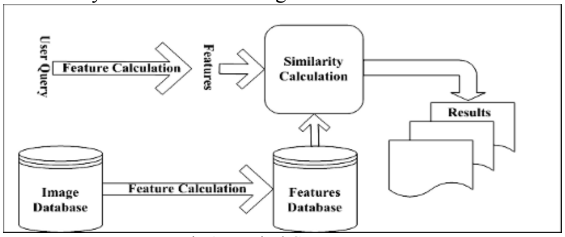

Fig.1. Typical CBIR system

At first features from all images in a database are extracted and stored into a feature database. This process is also known as indexing. When a user tries to search some images from the collection, he provides the system with a query image. Different options for query image are possible *i.e.*

- Image
- Rough Sketch
- Color or textural layout
- Verbal or semantic description

Features from query image are extracted by the same indexing mechanism. Then these query image features are matched with feature database using a similarity metric and finally similar images are retrieved.

A majority of indexing techniques are based on pixel

T. Sobh (ed.), Innovations and Advanced Techniques in Computer and Information Sciences and Engineering, 67–72.

domain features such as color [4], texture [5], and shape [6]. However, with recent advancements in image compression, compressed domain indexing techniques gained popularity due to their less complexity. Some frequency domain techniques include wavelet domain features, Gabor transform and Fourier domain features. Comprehensive survey of existing CBIR techniques can be found in [1, 2].

Researchers have shown that texture is one of the most important features for CBIR. Texture refers to the visual patterns that have properties of homogeneity that do not result from presence of only single color or intensity. It is an innate property of virtually all surfaces, including clouds, trees, bricks, hairs, fabric, etc. It contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment. There are many review papers exist on texture based image retrieval. Manjunath and Ma [8] evaluated the texture image annotation by various wavelet transform representations and found that Gabor transform was the best among the tested candidates.

Kingsbury [13] proposed a new complex wavelet transform which gives a fast way of computing Gabor-like wavelets. Peter and Kingsbury [14] in his paper have shown that how one can use this new transform to speedup and enhance the image retrieval process. Kokare et al. [15] have proposed even better extension of this work. Janney et al. [17] have shown how we can enhance the texture extraction capabilities of CWT for color image retrieval. They have shown that we can achieve almost the same precession for color image retrieval as well. These properties of CWT have motivated us to use it as feature extraction for our proposed system.

Mostly during the comparison phase of features, linear metrics like Euclidean distance etc. were used. Recently Han et al. [9] has shown how one can improve the performance of image retrieval systems by assuming non-linear relationship among feature vectors and grouping the images into similar classes. We have applied a similar idea for retrieving of texture images. We have used support vector machines for classification of images in the database.

The paper is organized as follows: Section II provides a brief introduction to dual tree complex wavelet transform and some of its applications. Section III provides an overview of support vector machines and how SVM can be used for classification. Section IV explains feature extraction process. Section V describes proposed and implemented CBIR system. Section VI discusses the results of our technique in comparison with existing techniques. Section VII gives the concluding remarks.

II. Dual tree complex wavelet transform

Kingsbury's [13] dual-tree complex wavelet transform (CWT) is an enhancement to the discrete wavelet transform (DWT), with important additional properties. The main advantages as compared to the DWT are that the complex wavelets are approximately shift invariant (meaning that our texture features are likely to be more robust to translations in the image) and that the complex wavelets have separate sub-bands for positive and negative orientations. Conventional separable real wavelets only have sub-bands for three different orientations at each level, and cannot distinguish between lines at 45° and -45°.

The complex wavelet transform attains these properties by replacing the tree structure of the conventional wavelet transform with a dual tree. At each scale one tree produces the real part of the complex wavelet coefficients, while the other produces the imaginary parts. A complex-valued wavelet $\psi(t)$ can be obtained as:

$$\psi(t) = \psi_h(t) + j\psi_g(t) \tag{1}$$

where $\psi_h(t)$ and $\psi_g(t)$ are both real valued wavelets.

CWT like Gabor transform have six orientations at each of four scales (any number of scales can be used, but the number of orientations is built into the method). The main advantage as compared to the Gabor transform is speed of computation. It has a redundancy of only 4 in 2-dimensions and so the post-processing stages (of calculating mean and standard deviation) are also faster as it has less redundancy than the Gabor wavelets. Fig. 2 shows magnitudes of CWT coefficients for a texture image, one can see more details about orientation and scales. Each row represents one scale and columns represent angles within that scale.

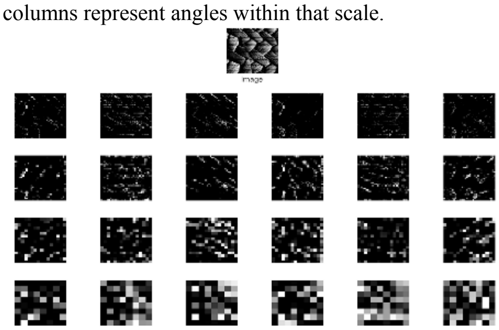

Fig.2. Four-scale CWT of a texture image

III. Support vector machines

There exists many pattern matching and machine learning tools and techniques for clustering and classification of linearly separable and non-separable data. Support vector machine (SVM) is a relatively new classifier and is based on strong foundations from broad area of statistical learning theory [11]. Since its inception in early 90's, it is being used in many application areas such as character recognition, image classification, bioinformatics, face detection, financial time series prediction etc.

SVM offers many advantages as compared with other classification methods such as neural networks. Kashif and Nasir [11] highlights many of advantages of support vector