Neurocomputing 116 (2013) 165–181

Contents lists available at SciVerse ScienceDirect

Neurocomputing

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

A simple computational model for image retrieval with weighted multifeatures based on orthogonal polynomials and genetic algorithm

R. Krishnamoorthi *, S. Sathiya Devi

Vision Lab, Department of Computer Science and Engineering, Anna University of Technology, Thiruchirappalli, Tamilnadu, India

ARTICLE INFO

Available online 8 October 2012

Keywords:
Orthogonal polynomials model
Shape
Texture
Color
Multifeature set
Genetic algorithm
Pseudo-Zernike moment
Autocorrelogram

ABSTRACT

This paper proposes a simple and new image retrieval method with weighted multifeature set based on multiresolution enhanced orthogonal polynomials model and genetic algorithm. In the proposed method, initially the orthogonal polynomials model coefficients are computed and reordered into multiresolution subband like structure. Then the statistical, directional, perceptual and invariant texture, shape and color features are directly extracted from the subband coefficients. The extracted texture, shape and color features are integrated into linear multifeature set and the significance of each feature in the multifeature set is determined by assigning appropriate weight. This paper also proposes a method to compute the optimized weight for each feature in the integrated linear multifeature multi feature set using genetic algorithm. Then the obtained optimized weight is multiplied with the corresponding features in the multifeature set and the weighted Manhattan distance metric is used for retrieving similar images. The efficiency of the proposed method is experimented on the standard subset of Corel and Caltech database images. The performance of the proposed method is compared with other existing retrieval methods such as Haar wavelet and Contourlet Transform based retrieval schemes. The proposed method yields high average recall and precision of 92.6% and 71% for Corel database and 90.5% and 72.3% of Caltech database images when compared with other existing methods.

1. Introduction

With the rapid growth of digital and information technologies, more and more multimedia data are generated and made available in digital form. Searching and retrieving relevant images in this huge volume of data is a difficult task and has created an urgent need to develop new tools and techniques. One such solution is the Content Based Image Retrieval (CBIR). As the image databases grow larger, the traditional keyword-based approach for retrieving a particular image becomes inefficient and suffers from the following limitations: (i) vast amount of labor is required for manual image annotation and (ii) limited capacity for retrieving the visual content of the image and subjectivity of human perception. Hence, to overcome these difficulties of manual annotation approach, content based image retrieval has emerged. CBIR is a collection of techniques and algorithms which enable querying the image databases with low level image content such as color, texture, objects and their geometries rather than textual attributes such as image name or other keywords [1]. Many image retrieval systems have been developed using all or some of these features. It includes Chabot [2], Photobook [3], QBIC [4], Virage [5], VisualSeek [6], MARS [7], Netra [8] and Excalibur [9]. The extensive literature and the state of art methods about content based image retrieval can be found in [10–15]. Though some of the image retrieval applications such as trademark retrieval, character recognition and leaf image retrieval [16,17] are implemented based on single feature (either texture, shape or color), but the single feature is found to be insufficient for natural, web based image retrieval applications as it affects the retrieval performance. Hence recently general-purpose CBIR systems concentrate on multiple features such as color, texture and shape along with some domain specific features for improving the performance of the image retrieval.

Sklansky [18] defined the texture as a set of local properties in the image region with a constant, slowly varying or approximately periodic pattern and it is measured using its distinct properties such as periodicity, coarseness, directionality and pattern complexity for efficient image retrieval particularly on the aspects of orientation and scale [19,20]. In a typical CBIR system, identification of the proper features that maximizes the differentiation of the texture is an important step. There are many categories of methods that exist for identifying and manipulating the texture: (i) statistical methods (Gray level Co occurrence matrix (GLCM) [21]), (ii) model based methods such as Markov Random Fields (MRF) [22], Simultaneous Auto Regression (SAR) [23], Wold decomposition [24] and (iii) signal processing methods (Gabor filters [25], Wavelet Transforms [26,27]). Some of these techniques depend on the comparison values of

* Corresponding author.
E-mail addresses: rkrish26@hotmail.com (R. Krishnamoorthi), sathyadevi_s@yahoo.com (S. Sathiya Devi).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2012.05.030

second order statistics obtained from query and stored images [28,27] for measuring the texture similarity.

The shape of an image is effectively perceived by the human eye than color or texture. Hence shape-based searching and retrieving has gained much attention in CBIR. Shape-based retrieval involves three primary issues: shape representation, shape similarity measure and shape indexing. Among them, shape representation is the most important issue in shape based image retrieval. The shape representation methods reported in the literature can be classified into two categories: region based and contour based [29]. Region based techniques have frequently used moment descriptors to obtain shape representation [30–33] which considers all the pixels inside the shape to compute the shape features. Contour based shape representation [29] only exploits shape boundary information and are classified into continuous approach (global) and discrete approach (structural). Both region and contour based representation methods compute shape features either in spatial or frequency domain. Spatial domain descriptors are sensitive to noise and not robust. In addition, it requires intensive computation during similarity calculation, due to the hard normalization of rotation invariance [34]. As the result, these spatial representations need further processing using spectral transform such as Fourier transform and wavelet transform. Wavelet descriptors have the advantage over Fourier descriptors in that they achieve localization of shape features in joint-space, i.e., in both spatial and frequency domains.

Early color based retrieval systems have used the global RGB histogram information such as the Local Color histogram [35], histogram difference approach [36], histogram intersection [37–39] and quadratic histogram comparison [40]. Though the color histogram based approaches are extremely easy to compute and insensitive to small changes in viewing positions and partial occlusion, they do not capture local spatial color information. Hence this approach is liable to false positives and is not robust to large appearance changes. Several recent schemes viz, Color Coherence Vector [41], Color Correlogram [42] and Binary Color Set [43,44] incorporate spatial correlation of color regions as well as the global distribution of local spatial correlation of colors to improve upon the histogram method. Though these techniques perform better than traditional histograms, they require intensive computation. Generally color spatial techniques are classified into three categories [45]: (i) partition based approach (ii) signature based approach and (iii) cluster based approach.

Saber and Tekalp [46] have introduced a region based algorithm for automatic image annotation and retrieval with color, edge, shape and texture features. The regions are extracted based on edge information and Bayesian techniques with texture and shape features are computed on these regions. Region-wise feature extraction becomes computationally intensive and degrades the retrieval performance. The combination of structure, color and texture features for image retrieval is reported in [47]. But in this method the automatic adjustment of weight for the features are missing and obtaining texture feature using gabor filter becomes computationally difficult. Howe and Huttenlocher [48] have used a technique that integrates diverse and expandable set of image properties such as color, texture and location in a retrieval framework. The retrieval performance of this method is fairly better (68.3%) compared to other methods and the substantial control is given to end user for retrieving relevant images. Wavelet transform based color, shape and texture feature extraction for image retrieval is reported in [49,50]. The multifeatures such as color, shape and texture are extracted from low and high frequency band of wavelet transform and the weight for each feature is determined using relevance feedback technique. Katare et al. [51] have integrated the shape and color feature for multi object image retrieval. In this method, the object with different orientation could not be identified with active contour based shape feature representation and it degrades the retrieval performance. Xiaojuan et al. [52] have established a method for image retrieval based on multifeatures with color, shape and texture and also introduced a method for normalization of the multifeature set. In this work, the multifeature normalization is performed in two stages: (i) Internal Normalization and (ii) Exterior Normalization and the authors themselves claimed that it is computationally intensive. An online application called garment image retrieval has used multifeatures for retrieving similar garment images [53]. Since the physical meaning, importance and value ranges of each feature are different in the linear combination of multifeature set, the similarity score computation with single distance metric becomes a series problem and degrades the retrieval accuracy. This problem can be solved using two methods viz., (i) relevance feedback [54–56] and (ii) appropriate weight generation [57,58]. Relevance feedback is computationally high demanding and difficult to incorporate human into the loop. The latter method can be viewed as an optimization problem and a suitable optimization technique has to be incorporated to generate the weight in an adaptive manner for effective discrimination and retrieval with less computational cost. Hence this paper proposes a method for optimized weight generation using genetic algorithm for multifeature representation with multiresolution enhanced orthogonal polynomials model for efficient image retrieval. This paper is organized as follows. The orthogonal polynomials model and reordering the transformed coefficients into multiresolution subband like structure have been described in Section 2. The multi feature extraction is presented in Section 3. The evolution and the process of genetic algorithm are presented in Section 4. The genetic algorithm based optimized weight generation process is presented in Section 5. The performance metric is described in Section 6. Experiments and results are discussed in Section 7 and conclusion is drawn in Section 8.

2. Multiresolution reordering with orthogonal polynomials model coefficients

In this section the orthogonal polynomials model and the reordering of the orthogonal polynomials coefficients into multiresolution subband structure is presented. Multiresolution analysis plays a vital role in Human Visual System (HVS) and the experimental studies have shown that the eye's sensitivity to a visual stimulus strongly depends upon the spatial frequency contents of this stimulus. Hence, HVS motivates the use of multiscale image decompositions as a front end to complex image processing algorithms. In general the multiscale decomposition of an image can be viewed as a linear transformation of the original pixel values into a set of bases of the transformed coefficients. The connection between multistage decompositions and the bases for image representation shows that images are sparse linear combinations of elementary images. Consider the following representation of a signal $x(t)$ defined over some domain $\Im$:

$$x(t) = \sum_k a_k \varphi_k(t), \qquad t \in \Im \tag{1}$$

Here, $\varphi_k(t)$ is termed as basis function and a_k are the coefficients of the signal $x(t)$ in the basis $\alpha = \{\varphi_k(t)\}$. For a discrete $(P \times Q)$ image, let the variable t in Eq. (1) be the pair of integers (n_1, n_2) and the domain of x be $\Im = \{0,1,....N-1\} \times \{0,1,,....M-1\}$,then the basis t is said to be discrete. Hence the subband decomposition of an image can be viewed as a linear transformation of the original PQ pixel values $x(t)$ into a set of PQ subband coefficients a_k. In this proposed work, the image analysis is considered as a linear transformation of original pixel values with a set of orthogonal polynomials functions. The orthogonal polynomials that have already been well established for image coding [59–61], have been extended to this proposed CBIR system. In order to analyze the content of an image for efficient proposal of CBIR system, a linear