Computers and Electrical Engineering 46 (2015) 288–302

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

Rotation and scale invariant hybrid image descriptor and retrieval ☆

Shiv Ram Dubey *, Satish Kumar Singh, Rajat Kumar Singh

Indian Institute of Information Technology, Allahabad, India

ARTICLE INFO

Article history:
Received 4 April 2014
Received in revised form 16 April 2015
Accepted 16 April 2015
Available online 3 June 2015

Keywords:
Structure element
Rotation-invariant descriptor
Scale-invariant descriptor
Texture analysis
Content-based image retrieval
Color quantization

ABSTRACT

Accurate image retrieval is required to index and retrieve large number of images from huge databases. In this paper, an efficient approach is presented to encode the color and textural features of images from the local neighborhood of each pixel. The color features are extracted by quantizing the RGB color space into a single channel with reduced number of shades. The texture information is encoded with structuring patterns generated from the locally structured elements chosen as a basis. Color and textural features are fused together to construct the inherently rotation and scale-invariant hybrid image descriptor (RSHD). This fusion is carried out by extracting textural cues over each shade independently. RSHD has been tested on the Corel dataset and experimental results suggest that RSHD outperforms state-of-the-art descriptors. The performance of the RSHD is promising under rotation and scaling. It can also be effectively used under more complex image transformations.

1. Introduction

1.1. Motivation

The demand for efficient image retrieval is rapidly increasing. In the early days, text based approaches were being used for image retrieval, but since the scope of such methods got reduced upon the existence of content-based image retrieval (CBIR), because retrieving images from its content is more visually accurate. In the published literature, content-based approaches describe more objectively and effectively, than text based approaches [1]. The main aim of CBIR is to facilitate efficient searching, browsing and matching over large datasets either offline or online opens an active research area in the field of computer vision and image processing from more than decades. Some recently reported typical applications of image retrieval are computerized facial diagnosis and retrieving human actions from realistic video databases [2]. The efficiency of the any CBIR system primarily depends upon the discriminating power present in their image feature descriptions. A CBIR system must be able to retrieve the images having details oriented randomly in nature. Recently, semantic based approaches became popular for image retrieval because it copes with the problem of CBIR [3,4]. To describe the semantic concepts and to be able to get the results close to human perception, relevance feedback algorithms are used by some researchers [5] in semantic image retrieval. They considered the priority provided by the users and bridge the semantic

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Associate Editor Dr. Ferat Sahin.

* Corresponding author.
E-mail addresses: shivram1987@gmail.com (S.R. Dubey), sk.singh@iiita.ac.in (S.K. Singh), rajatsingh@iiita.ac.in (R.K. Singh).

http://dx.doi.org/10.1016/j.compeleceng.2015.04.011
0045-7906/

gap, but drawback of such methods is that it is not fully automated and it requires user intervention to provide the feedback. Several methods to encode the information present in the images are proposed through published literature for image retrieval [6–11]. These methods used pixel level details (i.e. low level feature descriptions) of the image including color, texture, shape, gradient, orientation, etc. in the form of pattern to represent the whole image and images matched with their pixel level details. These features have been used efficiently in each type of CBIR systems such as based on global feature, region based feature, local feature, and structuring feature. The main shortcoming of these methods is with the utilization of different more robust features and lack of multiple type of information because by using just one type of features the method may not be able to characterize the image more accurately.

1.2. Related works

Global feature-based CBIR [12–14] does not divide the image into multiple regions. Chen et al. [12] used only color information to represent the image features by using the image color distributions. The color distributions are preserved up to the third moment in their method. The results are better than other color based features, but it should be noted that they have not considered any features other than color while the proposed method extracts structuring patterns over quantized shades distributions which increases the distinctiveness of proposed descriptor. Color difference histogram (CDH) is designed for color image analysis [13]. Color, texture and shape features are used for global image representation effectively by Wang et al. [14]. They have used only dominant set of color information with steerable filter decomposition and pseudo-Zernike moments in their descriptor construction. These methods have not taken care of local neighboring structures required to encode the relationship among the neighboring pixels.

Region based approaches [15,16] also used by several researchers to integrate the spatial location with the feature description. Hsiao et al. [15] partitioned images into five regions with fixed absolute locations. Similar to the case of semantic retrieval, their approach also needs user intervention in the middle of the retrieval process. On the other hand, proposed method considers only local neighborhoods of a given pixel which boost it with local discriminative power. In order to represent the image's spatial and color arrangements, Lin et al. [16] introduced three kinds of feature descriptors. To extract these features, they used K-means clustering approach to partition the whole image into different groups (i.e. clusters) using its intensity values. These regions based approaches have shown promising results with the expense of large dimensional descriptions and high computation.

Over the year's local binary pattern (LBP) [17] and various LBP based approaches [18–23] have been reported and became more popular because of their highly accurate performance and simplicity. LBP is constructed by comparing the each pixel in the image with its neighboring pixels and according to the sign of comparison the LBP is generated [17]. To reduce the size of LBP, only orthogonal pairs are compared in the center-symmetric local binary pattern (CS-LBP) [18]. Based on the multi-resolution, Zhu and Wang [19] have used multiple local patterns to encode the textural feature. Their method is also promising in the case of rotation and scaling, but lack of color information restricts their use in image retrieval. The complete local binary pattern is used for the fruit disease recognition [20]. To achieve invariance toward monotonic intensity change an order based descriptor local intensity order pattern (LIOP) is proposed [21]. Promising results using LIOP have been reported in the case of a monotonic intensity change, but this type of methods fails if there is a change in the objects of the image because using orders only does not consider actual scenarios. Dubey et al. [22] have extended the concept of LIOP over interleaved neighboring sets and designed the interleaved intensity order based local descriptor (IOLD) for image matching. They also proposed an illumination compensation mechanism to cope with the varying illumination for brightness-invariant image retrieval [23]. LBP and LBP based descriptors can be used efficiently to match the images having some geometric and photometric transformations. Most of these methods have not considered the fine structures of the images which can be seen as a basic building block of the image and the performance of such methods can be boosted effectively using fine structures in their framework.

Images are also represented by different types of structures present in the image [24–26]. Liu et al. [24] represented the co-occurrence matrix properties using the histogram to compound the advantages of histogram with co-occurrence matrix and proposed a multi-texton histogram (MTH) as a feature descriptor for image retrieval. Liu et al. [25] have introduced microstructure descriptors (MSD). MSD integrates color, texture, shape and spatial layout properties of the image for efficient content-based image retrieval. An efficient Structure Element Histogram (SEH) is presented by Xingyuan and Zongyu [26] which integrates texture with color feature. These structures based methods shown promising results in image retrieval, but their performance degrades under rotation and scaling. In the image retrieval and image classification problems, it is not possible to encode the exact information contained by an image using only one type of features such as color or texture. Therefore, it becomes highly desirable to merge these features in such a way that dimensionality should not increase too much.

Color and texture information are used by Wang et al. [27] to design a CBIR system. They used Zernike chromaticity distribution moments to capture the color features from the opponent chromaticity space which is a rotation and flip-invariant. They also used the Contourlet transform to encode the texture feature which is a rotation and scale-invariant. In [27], the color and texture features are first encoded separately and then combined to form the final feature vector, whereas we formed the descriptor by simultaneous encoding of color and texture in a hybrid manner. The main difference between the approach in [27] and proposed is that the color and texture features are processed independently in [27], whereas we processed the texture feature in conjunction with the color feature. Curvelet transform is also adopted by Youssef and