J. Vis. Commun. Image R. 33 (2015) 104–114

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Global Correlation Descriptor: A novel image representation for image retrieval ☆

Lin Feng [a,b], Jun Wu [a,b], Shenglan Liu [c], Hongwei Zhang [d,*]

[a] School of Innovation and Entrepreneurship, Dalian University of Technology, Dalian, Liaoning 116024, China
[b] School of Computer Science and Technology, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, Liaoning 116024, China
[c] School of Control Science and Engineering, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, Liaoning 116024, China
[d] School of Mathematical Sciences, Dalian University of Technology, Dalian, Liaoning 116024, China

ARTICLE INFO

Article history:
Received 23 March 2015
Accepted 7 September 2015
Available online 25 September 2015

Keywords:
Image retrieval
HSV color space
Texton detection
Global Correlation Descriptor (GCD)
Global Correlation Vector (GCV)
Global Correlation Vector (DGCV)
Feature fusion
Structure element correlation (SEC)

ABSTRACT

The image descriptors based on multi-features fusion have better performance than that based on simple feature in content-based image retrieval (CBIR). However, these methods still have some limitations: (1) the methods that define directly texture in color space put more emphasis on color than texture feature; (2) traditional descriptors based on histogram statistics disregard the spatial correlation between structure elements; (3) the descriptors based on structure element correlation (SEC) disregard the occurring probability of structure elements. To solve these problems, we propose a novel image descriptor, called Global Correlation Descriptor (GCD), to extract color and texture feature respectively so that these features have the same effect in CBIR. In addition, we propose Global Correlation Vector (GCV) and Directional Global Correlation Vector (DGCV) which can integrate the advantages of histogram statistics and SEC to characterize color and texture features respectively. Experimental results demonstrate that GCD is more robust and discriminative than other image descriptors in CBIR.

1. Introduction

In recent years, with the rapid development of the Internet and mobile terminals, processing image data effectively and efficiently is still an open issue in image research. Due to the complexity and diversity of images, it is difficult for users to retrieve desired images in the huge image datasets. Generally speaking, there are mainly three categories of image retrieval systems: text-based, content-based and semantic-based [1,2]. However, text-based and semantic-based image retrieval have some obvious limitations. Traditional text-based image retrieval systems need to annotate the images manually in advance. But it's time-consuming due to the amount of images, and people may perceive the same image differently. Besides, semantic-based image retrieval is still an open problem because of the limitations in studying about mechanisms of the primary visual cortex and artificial intelligence. So content-based image retrieval (CBIR) systems [3–6] have attracted more and more attentions in practical applications. These systems can generally be divided into three steps: first, the low-level features of the query image and all the images in dataset need to be extracted; second, choosing appropriate similarity measure, we can obtain the similarities between the query image and the images in dataset; third, sorting images by similarity, these systems return the top images that are most similar to the query image. Thus, the performance of the low-level feature has a great influence on CBIR systems.

Currently, most of CBIR systems are concentrated on extracting substantial features of an image, such as color, texture, shape, and fusion of two or more such features. So far, there have been many great research in extracting a single low-level feature. Due to the different description methods, these features can be classified into global and local features. The image representation based on histogram is one of the most common means of global features extraction, such as color histogram [7], Local Binary Pattern (LBP) [8] and Histogram of Oriented Gradient (HOG) [9], which see the whole image as visual information. However, these features tend to lose the spatial correlation among pixels. To overcome the problem, many researchers proposed their own visual models in succession with different ideas. Sticker and Orengo proposed the concept of color moments [10], which used the first three central moments called mean, standard deviation and skewness. Color correlogram [11] and color coherence vector (CCV) [12] also characterizes the color distributions of pixels and the spatial correlation between pair of colors. In addition, the gray co-occurrence matrix

☆ This paper has been recommended for acceptance by M.T. Sun.

* Corresponding author.
E-mail address: hwzhang@dlut.edu.cn (H. Zhang).

http://dx.doi.org/10.1016/j.jvcir.2015.09.002
1047-3203/© 2015 Elsevier Inc. All rights reserved.

[13] and co-occurrence histograms of oriented gradients (CoHOG) [14] characterize the co-occurrence relationship between the values of two pixels. Different from Global feature-based algorithms, local feature-based algorithms focus mainly on key points and salient patches. In [15,16], scale-invariant feature transform (SIFT) was proposed to detect and describe local key points in scale space and has been widely applied to computer vision fields. Gabor wavelets [17], whose kernels are similar to the receptive field of the mammalian cortical simple cells, has been introduced to image analysis models. Based on the biological properties of Gabor wavelets, Liu and Wechsler [18] proposed the Gabor feature representation. And Serre et al. [19] introduced HMAX model based on the hierarchical visual processing in the primary visual cortex (V1). In this model, S1 units also take the form of Gabor filters in different scales and orientations. In [20], the comparison of some local descriptors for image processing can be seen.

Color, texture and shape are significant properties of an image, but simple feature usually has some limitations. To overcome the problem, some researchers proposed image representation based on multi-features fusion, which characters simultaneously two or more low-level features. These methods also can be classified two strategies. One strategy is to extract several features respectively and combine them into an integral vector [21–24], which can express more image information than simple one feature. In [21], Dubey et al. proposed the combination of four feature extraction methods namely color histogram, color moment, Texture and Edge Histogram Descriptor. Jalali et al. [23,24] studied the color processing in the high-level visual area of the primate brain and proposed the Color Quantization Hierarchical Max (CQ-HMAX) model, which use color quantization cores instead of Gabor filters to encode color information. And then combined with HMAX, the model can obtain color and shape information simultaneously. The other strategy is to extract texture and shape in color space directly. In [25], Liu and Yang introduced texton co-occurrence matrix (TCM) in RGB color space. Multi-texton histogram (MTH) [26] and color difference histograms (CDH) [27] integrates the advantages of co-occurrence matrix and histogram and encode color, texture and shape information. Micro-structure descriptor (MSD) [28] extracts micro-structures based on an edge orientation similarity and the underlying colors and structure element correlation statistics characterizes the spatial correlation of micro-structures. In [29], Wang and Wang introduced the concept of structure elements' descriptor (SED) in HSV color space which can describe color and texture information. And then they propose multi-factors correlation (MFC) namely structure element correlation (SEC) [30], gradient value correlation (GVC) and gradient direction correlation (GDC).

However, the strategy that extracts texture and shape in color space directly tends to enhance the characteristics of image color features. Thus, the color information may be dominant in these descriptors and lead to impair the performance of these descriptors in CBIR. In this paper, we propose a novel image representation, called Global Correlation Descriptor (GCD), which can extract color and texture feature of color image respectively.

The rest of this paper is organized as follow. In Section 2, the structure of our GCD model is introduced. In Sections 3 and 4, the color feature and texture feature of GCD model are presented respectively. The experimental results in CBIR are showed in Section 5. Section 6 is the conclusion of the paper.

2. Global Correlation Descriptor (GCD)

Color and texture of color image are important features in CBIR. Texture histogram which extracts some regular textons of gray image is a common approach, such as LBP and gray co-occurrence matrix which are introduced based on the different expression of texton. To combine with color information, some researchers defined texton of a color image in color space and extracted texture information based on this color space. However, though merging color and texture simultaneously, these descriptors can get worse performance in CBIR due to enhance the characteristics of image color features. As shown in Fig. 1, there are three flowers with similar texture and shape and different color significantly. The feature vector of these descriptors may be different significantly because the texton of the three flowers focus on their own color layer. To avoid this problem, we propose a new image feature model, called Global Correlation Descriptor (GCD). With this model, color and texture feature are extracted respectively and then are merged into an integral vector.

The structure of our GCD model is shown in Fig. 2. First, we construct image pyramid in color space to obtain color information of multiresolution images. To characterize the color distributions of pixels and the spatial correlation between pair of colors, we propose a new method, called Global Correlation Vector (GCV), to extract color feature in image pyramid. Second, in order to extract texture information unaffected by color, the color image is converted to grayscale. And then we choose Gabor filters in different scales and orientations to process the image. On this basis, we propose a new texture descriptor by defining several directional texton types and introduce Directional Global Correlation Vector (DGCV) to extract texture information. Finally, combining GCV and DGCV, we can get the Global Correlation Descriptor (GCD). The performance of color and texture feature may differ in CBIR in different image datasets, but they are of equal importance for most of the images.

As shown in Fig. 3, in the process of color feature extraction, we can choose appropriate scales for image pyramid. And then combining the GCV of each scale of the pyramid, we can get a more discriminative color feature vector, which is denoted H_{color}. Similarly, we choose multiple Gabor filters in different scales and obtain DGCV of each filtered image in the process of texture feature extraction. Then combining these vectors, we can get a more discriminative texture feature vector, which is denoted $H_{texture}$. Finally, GCD can be represented as:

$$H = [H_{color} \quad H_{texture}]^T \tag{1}$$

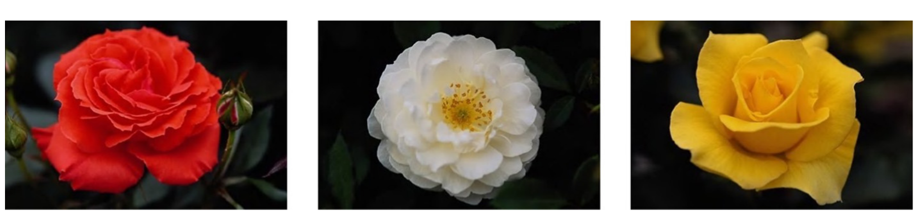

Fig. 1. Three flowers with different color and similar texture.