J. Vis. Commun. Image R. 25 (2014) 727–739

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Statistical distributional approach for scale and rotation invariant color image retrieval using multivariate parametric tests and orthogonality condition

K. Seetharaman *, M. Jeyakarthic [1]

Department of Computer Science and Engineering, Annamalai University, Annamalainagar 608 002, Tamilnadu, India

A R T I C L E I N F O

Article history:
Received 3 April 2013
Accepted 3 January 2014
Available online 18 January 2014

Keywords:
Variance–Covariance
Mean vectors
Position vector
Skewness
Kurtosis
Query image
Target image
Tests of hypothesis

A B S T R A C T

This paper proposes a unified framework for color image retrieval, based on statistical multivariate parametric tests, namely test for equality of covariance matrices, test for equality of mean vectors, and the orthogonality test. The proposed method tests the variation between the query and target images; if it passes the test, then it proceeds to test the spectrum of energy of the two images; otherwise, the test is dropped. If the query and target images pass both the tests then it is concluded that the two images belong to the same class, i.e., both the images are same; otherwise, it is assumed that the images belong to different classes, i.e., both the images are different. The obtained test statistic values are indexed in ascending order and the image corresponds to the least value is identified as same or similar images. Here, either the query image or target image is treated as sample; the other is treated as population. Also, some other features such as Coefficient of Variation, Skewness, Kurtosis, Variance–Covariance, spectrum of energy, and number of shapes in the images are compared between the query and target images color-wise. Furthermore, to emphasize the efficiency of the proposed system, the geometrical structure, viz. test for orthogonality between the query and target images, is examined. In the case of structure images, the number of shapes in the query and target images are compared; if it matches, then the contents in the shapes are compared color-wise. The proposed system is invariant for scaling, and rotation, since the system adjusts itself and treats either the query image or the target image is the sample of other. The proposed framework provides hundred percent accuracy if the query and target images are same, whereas there is a slight variation for similar, scaled, and rotated images.

1. Introduction

The advances in computer vision systems and the advent of the World Wide Web, the usage of the digital libraries and the access of image data through on-line have got exponential growth during the past decade. Many areas such as medical, educational, traffic management, law enforcement, automatic target recognition, and geographic information systems effectively utilize these advances in vision technology. Handling images, that is, manipulation, storing, analyzing, indexing, matching, retrieval, display etc., is very complicated when compared with text manipulation. Thus, it needs proper image database systems, which can support the aforesaid image manipulations. It is observed from the literature that remarkable progress has been made in both theoretical research and system development. However, still it is a challenging problem for the researchers in the area of visual data mining to design an automatic retrieval system, because real-world images usually contain complicated objects and color information. Mainly, images are retrieved based on the content of the image.

The content-based image retrieval (CBIR) system has attracted many researchers for over a decade. In CBIR system, the researchers concentrate on developing low-level global visual features such as color properties, shape, texture, and spatial relationship etc., which are used as query for the retrieval process [1–4]. Jing et al. [5] suggested that a single signature computed for the entire image cannot sufficiently capture the important features of individual objects, and there is a gap between the visual features and semantic concepts of images. To overcome this problem, region-based system is developed [6–9], which represents the focus of the user's perceptions of the image contents. The method proposed in [4,10–14] classifies or segments the entire image into various regions according to the objects or structures present in the image, and then the region-to-region comparison is made to measure

* Corresponding author. Fax: +91 4144 238090.
E-mail addresses: kseethadde@yahoo.com (K. Seetharaman), jeya_karthic@yahoo.com (M. Jeyakarthic).
[1] Fax: +91 4144 238090.

http://dx.doi.org/10.1016/j.jvcir.2014.01.004
1047-3203/

the similarity between two images [3,11,12]. In a region-based system, the user has to provide one or more regions from the query image to start a query session. Automatic and precise extraction of image objects is still beyond the ability of the retrieval system available with the computer vision [15]. Therefore, the above system tends to partition one object into several regions, but none of them is representative of the semantic object. Edge-oriented segmentation algorithms are available for both region-based and content-based retrieval systems. This type of technique uses features of local edges, and the edges are classified based on two factors such as orientation and correlation between neighboring edges [17]. Hence, it contains information of continuous edges and lines of images, and describes the major shape characteristics of images. The main drawback of this system is to determine the region, which is to be used as feedback query image example for retrieval [5]. In this type of systems, when either the target image or the query image is rotated by a particular angle, the corresponding regions of both images may not match exactly. Hence, it is not robust for the same image with different angles.

Su et al. [18] suggest that the content-based retrieval system yields results with low accuracy and slow response time, because there is a big gap between semantic concepts and low-level image features. A concept, relevance feedback, has been developed to bridge the gap [19–22]. In [18], a new relevance feedback approach is proposed, which uses Bayesian classifier and treats positive and negative feedback images with different strategies. In relevance feedback method, the user has to provide positive and negative feedback images to improve the performance of the system. Again the problem arises, as discussed in [5], how to incorporate the negative and positive examples to refine the query and how to adjust the similarity measures according to the feedback [18]. The degree of relevance (to the query) of each of the $\boldsymbol{n}$ positive feedback images is determined by the user at each feedback iteration. Since the positive feedback images are chosen by the user with the use of iterative technique, this technique demands computational complexity. Minka and Picard [22] proposed the FourEyes system, which has two disadvantages: (i) it uses the region-to-region similarity measure; (ii) the re-clustering of all the features when a new image is added. Thus, it is not very scalable [5]. To overcome this, Jing et al. [5] proposed a system with the features: (i) it computes probabilistic interpretation and it is used in region matching; (ii) region codebook is used; (iii) the SVM based classifier and clustering techniques are adopted, but it requires high computational effort. Above all these, it requires positive and negative query image examples. Again the problem is how to select the negative and positive image examples. To avoid this problem, many researchers [8,15,23–26] have taken efforts to integrate the information from all regions in an image and then the image-to-image similarity measure (integrated information) is used to compare the images. For instance, SIMPLIcity system [15] uses an integrated region matching as the similarity measure, and in recent works [27,28], image-to-image comparison is made. In summary, the review highlights that the region-based technique again tends to the concept of content-based retrieval using the integrated information. Moreover, the region-based technique demands high computational efforts.

Theoharatos et al. [16] proposed a system, based on multivariate non-parametric test, namely Wald–Wolfowitz test (WW-test), and graph theoretic framework of minimal-spanning-tree (MST). In this work, first, the MST is constructed based on the sample identities of the points taken from the images. Based on the consecutive sequence of identical sample identities, runs of the sample points are computed and the WW-test is employed to identify whether the query and target images are same or not. In this work, the drawbacks are

(i) Construction of the MST demands computational overhead.
(ii) Based on the sample identities of the points, run length of each sample identical identities is computed and then the WW-test is used to identify whether the query and target images are same or not.

The pixel values in the images are quantized, so that the parametric test can be applied in straight without converting the content of the image as qualitative values that are either 0 or 1 as in [16]. By applying the parametric test in straight, the additional computational headache can also be avoided. Also, the non-parametric tests do not yield good results compared to those of parametric tests. Generally, the decision makers or statisticians choose non-parametric tests in rare case, while there is no option to apply the parametric tests.

In this paper, a unified scheme is proposed for automatic image retrieval, based on the multivariate parametric tests, namely test for equality of means, test for covariance, and geometrical interpretation (GI). In the proposed technique, mean and covariance (first and second moments of the sample points) are used as representatives of both query and target images. The GI is used to give geometrical interpretation of two normal populations [29]. In addition to that the statistical features such as Coefficient of Variation (CV), Skewness (Skw), Kurtosis (Kur), Variance–Covariance (VaCov), spectrum of energy (SoE), and number of shapes in the images are used. The methods proposed in [27,28] retrieve only the texture images with intensity values ranging from 0 to 255, i.e., gray-scale images. This motivated us to develop a new technique which retrieves color images; both texture and structure images; and also invariant for rotation and scaling.

1.1. Overview of the proposed work

First, the colors, viz. red, green and blue are segregated, and various features such as CV, Skw, Kur, VaCov, SoE, and number of shapes in the images are extracted for each color. In the case of structure images, first shapes are segregated and then the features are extracted color-wise. The extracted features are formed as vector. The feature vectors are clustered into various classes using k-means algorithm with confidence interval [30]. The feature vectors in each class are indexed in ascending order. The centre value of each vector of each class is computed and indexed in ascending order. A link is established between the feature vector and its corresponding image in the image database. The indexed feature vectors are stored as feature database. Now, the images are also stored as in the order of the feature vector database. To include a new image in the image database, feature vectors of the image are computed, and it is compared to that of the center feature value of each class, and wherever it matches, then the image is incorporated in the corresponding class. If it does not match with any class in the feature database, the feature vectors of that image are categorized into a new class, and indexed in the feature vector database. The new feature vectors are incorporated in the appropriate indexed location. Similarly, feature vectors are computed for the query image submitted by the user. The feature vectors are compared with the median (index of the classes) value of each class; if it matches, then it is identified that the query image belongs to that class; otherwise, it forms a new class. Then, the system compares the feature vectors with all feature vectors in the class, and retrieves the images which are nearer to the query image. Furthermore, to confirm the image to be retrieved, the angle between the query and target images is computed, and are indexed in ascending order. The image, which corresponds to the nearest feature vector in the image feature database, is retrieved. The entire retrieving procedure is depicted in Fig. 1.