Improvement of Accuracy of Content-Based Image Retrieval using Local and Statistical Methods

Narges Savoj[1], Bijan Shoushtarian[2*]
1- Department of Computer Engineering, Sheikhbahaee University, Isfahan, Iran.
Email: savoj.n@gmail.com
2- Department of Computer Engineering, Sheikhbahaee University, Isfahan, Iran.
Email: shoushtarian@aim.com(Corresponding author)

Received: June 2019 Revised: July 2019 Accepted: August 2019

ABSTRACT:
Content-based image retrieval (CBIR) system plays an important role in retrieving desired images from a large database of images. These programs in all areas, including hospitals, regulatory applications (surveillance), architecture, journalism and many other items found in the role. In initial research text-based image retrieval was performed, but with the advent of great challenges in text-based retrieval (eg spelling errors), content-based image retrieval has been introduced by researchers, which is by far the most effective method for image retrieval. Content-based image retrieval system uses features such as color, shape and texture. To extract the tissue properties local binary patterns and edge filtering methods are of particular popularity among researchers. A review of the methods presented so far shows that despite the quality of the descriptors and categories and retrieval methods, none of these methods can meet the needs and challenges of the present, so to improve the accuracy of image retrieval, in this study, a method introduced. To extract feature from the images, five color histogram descriptors, color moment, edge histogram, gradient oriented histogram and MRELBP were used. To classify the attributes extracted by the descriptors, three categories of support vector machine and k nearest neighbour and random forest are used. In the method, the features extracted by the five descriptors are combined and after classifying and identifying the test image class, using the Kmeans cluster, the closest images to the test image are retrieved from the identified class. Experimental results method on three databases Corel 4k, Wang and Corel 5k show We have accomplished the highest precision rate of 86% using proposed CBIR system.

KEYWORDS: Image Retrieval, Content-Based Image Retrieval methods, Local Image Features, Statistical Features, Distance-based Identification.

1. INTRODUCTION

Nowadays, with the advent of machine learning science, digital image processing has become a major focus for timely and accurate use of these visual content. Image retrieval is one of the most effective ways to access visual content quickly and accurately. Before discussing the definition of image retrieval, we must discuss the main reason of this science. With the rapid advancement of technology, especially in the past decade, we are seeing an increasing volume of images stored on the World Wide Web or memory devices. This huge volume of images and visual content is stored and imaging in cyberspace, hospitals, surveillance systems, etc. and timely and accurate access to these images is a constant concern of users of these systems. Therefore, the need for an efficient system in order to retrieve the desired image of the user from the huge volume of image is inevitable. So the purpose of retrieval in these systems is to search for digital images between a large number of images and display similar images. Image retrieval, like many other sciences, has its own challenges. One of the challenges facing of image retrieval, data retrieval accuracy and indexing in databases with a large number of classes [1]. Image retrieval was introduced for the first time in year 2 with the introduction of text-based image retrieval [2] so that images were annotated with queries and the search was performed based on the entered phrase. Two decades later, content-based retrieval was invented [3], where the search is based on the extracted features from the images is performed. This type of retrieval was quickly replaced by the previous method and then used in medical fields, digital libraries, fingerprint recognition, crime prevention (Recover the faces of criminals or their fingerprints), search websites, etc. These methods are based on low-level feature extraction from images. But given the semantic gap between human perception and these low-level features, and on the other hand with the increasing volume of images, the need for an efficient system that accurately carries out the retrieval process and the semantic gap between human perception and reduce the extracted features are felt.

Two traditional methods of image retrieval are text-based retrieval (TBIR) and content-based retrieval (CBIR) [4]. In TBIR, the user enters a query, such as the image name, image creation date, and any other keywords, and the system retrieves the image based on user input information. There were some difficulties in using this method, such as that it was not possible to express emotion in the form of words or sometimes misspellings or words that had similar spelling but different meanings had reduced retrieval accuracy [5]. To solve this problem, content-based retrieval was invented so that the system receives the input image from the user and retrieves images that are similar in content (such as color, texture, and shape) to the on-demand image. The benefits of CBIR include the possibility of automatic retrieval, and this leads to superiority over the TBIR method, which it comes with time consuming annotations. In this method, a feature vector is extracted from each image and comparing images with comparing feature vectors is done. A review of recent years' research shows that machine vision methods have performed well in this area, and among them, machine learning algorithms have been remarkably successful. Because good features can be extracted from the image. So in this article, we present a new machine-based approach to image retrieval. For this purpose, one of the newest machine learning algorithms in the field of machine learning called MRELBP has been used along with other descriptors based on colour and edge features and so on for image retrieval. Section 2 provides an overview of the work done and the research background. Section 3 is devoted to introducing the algorithms used. In Section 4, a proposed model for image retrieval is presented, and Section 5 compares the results of this method with previous methods.

2. RESEARCH BACKGROUND

Yuan et al [1]. [6] used colour co-occurrence matrix to extract texture property and measure similarity of two colour images. In this paper, the color information such as components and color distribution is considered and the resulting feature not only reflects the texture correlation but also the color information. Liu and Yang[2] [7] used from color difference histograms (CDH) to recover the image. Most histogram techniques only count the number or frequency of pixels, but the unique feature of CDH considers the uniform random color difference between two points in different contexts with respect to the colors and orientations of the edge in the color space L * a * b they take. Giri and Meena[3] [8] used color features using color histograms and color moments[4] and texture features using Wavelet transform and Gabor to recover Euclidean distance-based image. As well as to classify the extracted features, support vector machine is used. Huneiti et al[5]. [9] Performed image retrieval using color and texture features based on artificial neural networks (SOM) and violet conversion (DWT). At the time of retrieval, the similarity of images is calculated using the Euclidean distance. In addition, another method for SOM-based retrieval has been explored. kaur and Sohi[6] [10] a method based on features such as HSV, Color Moment, HSV, Gabor wavelet and wavelet transform, Edge Gradient used for retrieval. The texture algorithms used are robust against scaling and rotation, and the gradient of the edges is used to incorporate the shape properties. Fadai et al[7]. [11] used color descriptors (DCDs), wavelet features[8], and curves[9] to introduce a content-based image retrieval system. In the proposed CBIR scheme, the DCD attributes are first extracted as color attributes and then an appropriate similarity criterion is applied. Mahindra et al. [12] used a combination of DCT and SVD conversion methods for image retrieval.

3. THE PROPOSED METHOD

One of the most useful and effective ways to improve identification accuracy in machine learning research and pattern recognition is the use of integration techniques. Among the various merging methods, cascade feature vectors with regard to achieving effective results is particularly popular.

For this purpose, in cascade feature vectors, the features of each image are extracted using five MRELBP descriptors, color histogram, color moment, edge histogram, and HOG. Then, using elemental component analysis, the effective attributes extracted from each descriptor are selected and cascaded together. These steps produce a single feature vector for each image. Next, the test feature vector is classified using the support vector machine class clusters, random forest and k nearest neighbor, and the images with the test image in a cluster, from the class predicted by the class clusters, are provided to the user. Placed. This clustering is done by k-means. Figure (1) is a block diagram of the proposed method.

[1] . WangXing-yuan, et al
[2] . Guang-Hai Liu and Jing-Yu Yang
[3] . Aditi Giri and Yogesh Kumar Meena
[4] . color moment
[5] . Ammar M Huneiti and Maisa Daoud
[6] . ManpreetKaur and NeelofarSohi
[7] . Sadegh Fadaei, et al.
[8] . Wavelet
[9] . curvelet

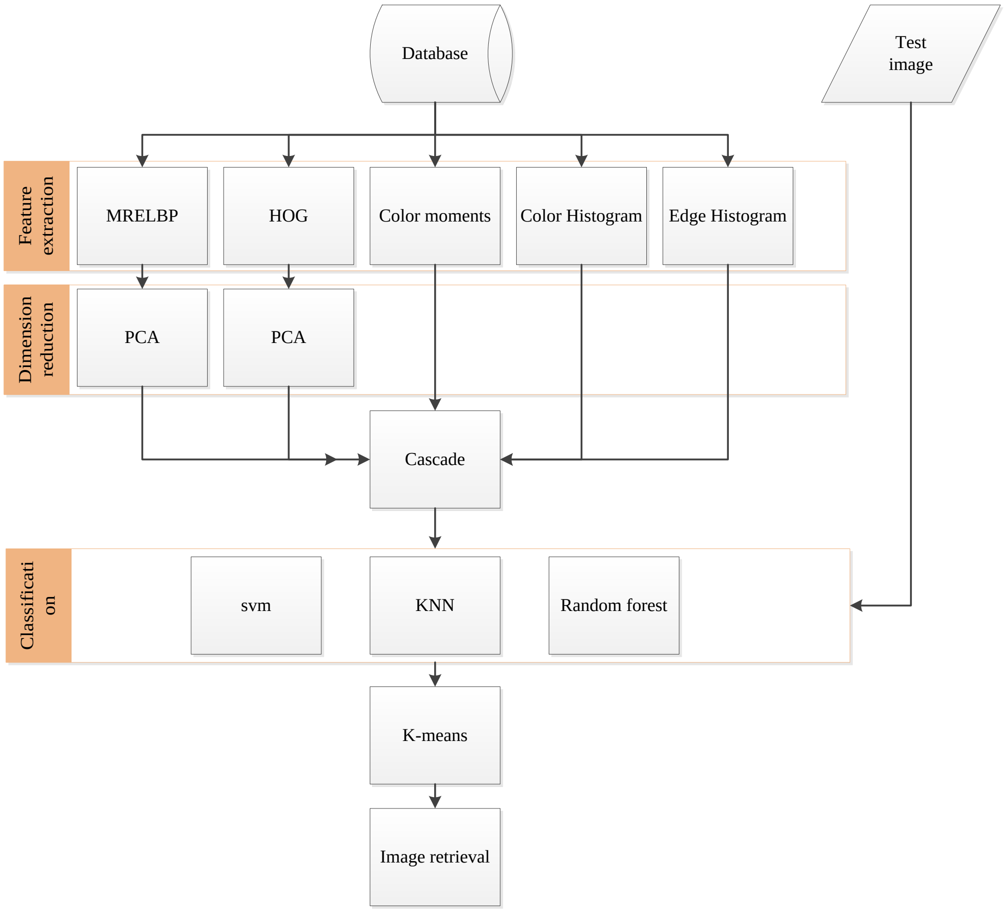

Fig.1. Block diagram of the proposed method.

3.1. Feature Extraction

At the feature extraction stage, the features of the database images are extracted using the descriptors listed and the corresponding feature vectors are generated for each image.

3.1.1. Color Histogram Descriptor

Color histograms are one of the most effective and widely used statistical methods that can be used to differentiate color images. This descriptor is often used in image retrieval and its effect is to make a distinction based on the image environment such as urban environment, forest, sea and so on. That is why it is widely used in image retrieval. Figure (3) is a sample histogram output of a database image.

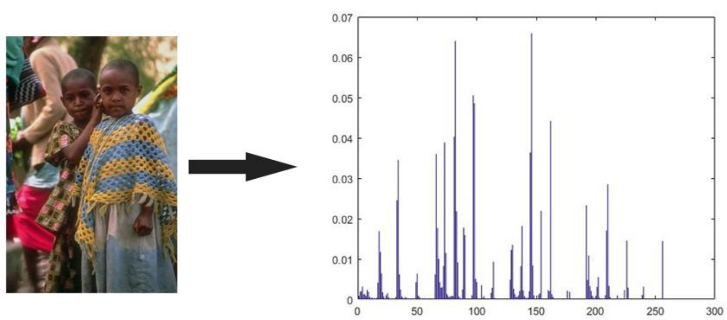

Fig .2. Output Color Histogram Descriptor

3.1.2. Color moment descriptor

As mentioned, color is typically one of the most valuable information sources in the process of extracting features from color images. In this paper, RGB color spaces are used for feature extraction. The method of using these color spaces is that the first order (mean) and second order (variance) moments of the color distribution of the images are calculated for each color layer in different color spaces and the number of these features is 6.

3.1.3. Edge histogram descriptor

Edge histogram descriptors are one of the methods widely used in tissue imagery to detect shape. Basically, this method represents the relative abundance of 5 edge types 0, 45, 90, 135 and without direction in each local area. Figure (4) shows an example of an edge histogram output.

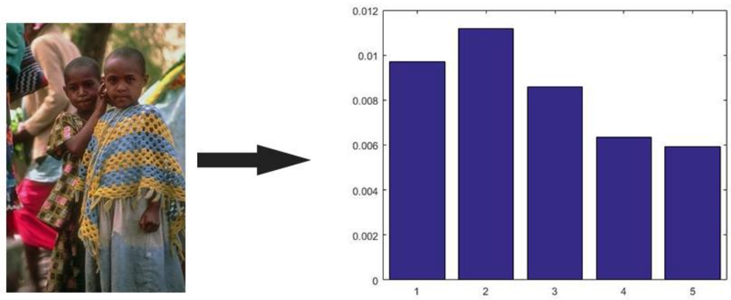

Fig. 3. The output of the edge histogram descriptor

3.1.4. Oriented Gradient Histogram Descriptor

In this method, the image gradient is first calculated in the x and y directions. Then locally, each part of the image is divided into blocks (4 × 4) without overlap, and each block is divided into cells (8 × 8). The angle of the gradient is limited to 0-180 ° and this range is divided into 90 two-degree bars. To calculate the histogram, each pixel within a cell votes on one of the histogram ranges based on its gradient angle. These votes are weighted by the size of the gradient in that pixel. After calculating the gradient histogram in each cell, this histogram is assigned to the central pixel of the cell. Therefore, for each pixel position, an n-dimensional vector is obtained, which represents the histogram of the neighborhood gradient around it. Figure 5 illustrates this description.

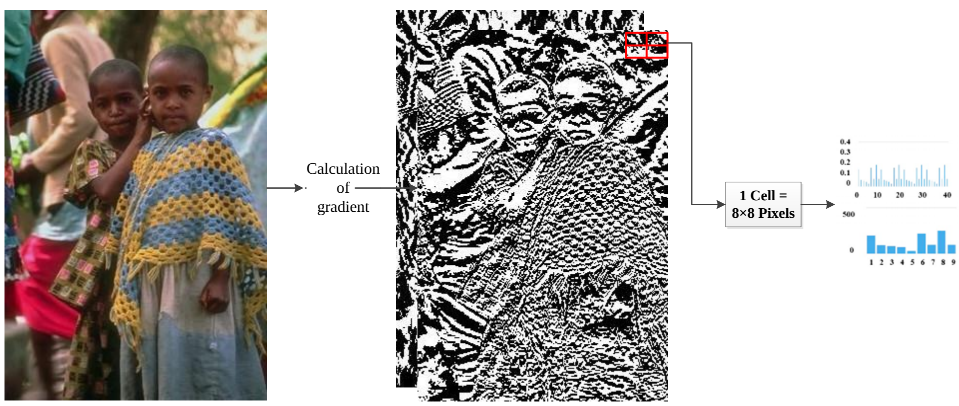

Fig. 4. HOG descriptor output

3.1.5. MRELBP descriptor

This texture descriptor was introduced in 2016 by Lee et al. All input parameters of this algorithm are similar to those mentioned in Lee et al.'s article and the only different case is the number of image blocks in which the images are divided into blocks (5 x 5) without overlap. Figure 6 shows how to feature extraction by this algorithm.

Finally, with the completion of the implementation of any of the descriptors, feature vectors are produced for each of the images.

3.2. Feature Selection

After examining the dimensions of the feature vectors produced by the two descriptors MRELBP and HOG it can be predicted that the classification of feature vectors with these dimensions is time-consuming as well as having higher computational complexity. For example, the feature vectors produced by the MRELBP descriptor for each image are very high (1×20000). Therefore, the principal component analysis has been used to removal this challenge. Attribute vectors, after being analyzed by this attribute selector, are much smaller by maintaining the quality of the attribute vector and are ready to enter the cluster for identification (950×898).

3.3. Classification and Identification

Output feature vectors from the dimension reduction step are categorized by clusters of support vector machines, k nearest neighbor, and random forest, and the accuracy of identifying each of the classifications is calculated for each descriptor.

The support vector machine used in this study has a linear kernel and contains the entries of this clause, the image feature database, the training and test image feature database, and the class determination matrix of each training sample.

In cluster k, the number of nearest neighbors is set to 1 (k = 1). Also the criterion for calculating the distance in this cluster, Euclidean distance is considered. These two parameters, together with the training imagery database and the test imagery database, and the class assignment matrix of each sample are considered as inputs to this cluster.

In the random forest cluster, the parameter of the number of trees for the forest is first determined. This parameter is estimated to be 500 in this study. The cluster in the training phase uses the mentioned parameter and the training data and the training data class design a model. This model predicts the test data class in the prediction phase.

3.4. K-means Cluster

After the test image class is identified by the methods mentioned, it is time to recover the image from the specified class. For this purpose, a clustering is first performed on the images of each class using the k-means method. Then the distance between the test image and each cluster head is calculated and the selected images are selected to recover from the cluster that has the least distance to the test image.

4. EXPERIMENTAL RESULTS

In this section, we evaluate the results of the proposed method using the Wang and Coral databases. Therefore, the databases used in the proposed method are discussed first, and then the results of the proposed methods are discussed.

4.1. Databases

The data used in this study include two databases of Wang and Corel [54] [55]. The Wang dataset consists of 1000 images in 10 classes, each containing 100 images in 384×256 dimensions. The Corel-4k dataset consists of 4,000 images in 40 classes, available in each class, 100 images, 128 x 192 or 192 x 128.

4.2. Results of Combinations of Descriptors and Clusters

This article is based on the integration of descriptive features from the descriptors and voting results from the clusters. Therefore, in order to evaluate these two methods, the results of each descriptor and cluster should first be examined and finally these results compared with the results of the proposed methods. Tables 2 to 4 report the results of the descriptor and classifier combinations in the three databases used in this study. The attributes extracted from each descriptor are categorized by three categories and the identification rate is recorded.

Table 1. Compares the results of the feature vector classification of each image in the Wang database.

Descriptor	Support Vector Machine	*k Nearest neighbor*	Random forest
	Precision	Precision	Precision
Color histogram	62%	70%	82%
Color Moment	38%	64%	60%
MRELBP	82%	66%	58%
HOG	52%	36%	58%
edge histogram	24%	40%	44%

Table 2. Compares the results of the feature vector classification of each image in the Corel4k database.

Descriptor	Support Vector Machine	*k Nearest neighbor*	Random forest
	Precision	Precision	Precision
Color histogram	Lack of convergence	45%	75%

Color Moment	Lack of convergence	53%	62/5%
MRELBP	56/5%	56%	51%
HOG	38/5%	48/5%	34/5%
edge histogram	Lack of convergence	37%	40%

Table 3. Ccompares the results of the feature vector classification of each image in the Corel5k database.

Descriptor	Support Vector Machine	k Nearest neighbor	Random forest
	Precision	Precision	Precision
Color histogram	Lack of convergence	71/5%	84/4%
Color Moment	Lack of convergence	56%	34/4%
MRELBP	57/2%	55/30%	46/8%
HOG	38/8%	56/8%	34%
edge histogram	Lack of convergence	36/4%	40/8%

The results of this section show that the support vector machine cluster has the lowest detection rate compared to the other two clusters. Because in many cases the feature vectors of different clusters are so close that the cluster does not converge. After cluster the support vector machine, random forest with color and edge histogram features and K had the closest neighborhood to the HOG, MRELBP and color moments with the highest classification rates. In sum, it should be concluded that none of the results obtained are in a good level for classification.

4.3. Results of the Proposed Method in the Feature-Level Integration Section

As noted in the proposed method section, the features extracted from the image are merged using five feature-level descriptors. The important point in this section is the type of attributes extracted from each descriptor. Color histogram descriptors, edge histograms, HOGs and MRELBPs are all histogram types, so merging these features into the final feature vector does not create a coherent state. Therefore, this method is expected to produce relatively acceptable results. Table 5 compares the results of integration at the feature level.

Table 4. Compares the results of the combination of the five descriptors in image retrieval.

Database	Support Vector Machine	k Nearest neighbor	Random forest
	Accuracy	Precision	Recall
Wang	86%	66%	80%
Corel4k	60%	56%	84%
Corel5k	68/4%	85/9%	85%

As can be seen from the results in Table 4, cascade vector feature has been able to achieve a better identification rate than the previous segment in all classification methods. In this case, the backing vector machine classifier, with the distinction created by the merged feature vectors, has been able to achieve a rate of nearly 86% in classification and consequently in image retrieval. One can definitely say that the reason for this is the homogeneity of the final feature vector. This correlation in the final vector has made the merged feature vector more distinct than the descriptive feature vector.

REFERENCES

[1] A. N. Tikle, C. Vaidya, and P. Dahiwale, **"Notice of Retraction A Survey of Indexing Techniques For Large Scale Content-Based Image Retrieval,"** in *2015 International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO)*, IEEE, pp. 1-5, 2015.

[2] P. Mueller, **"Color image retrieval from monochrome transparencies,"** *Applied optics*, Vol. 8, No. 10, pp. 2051-2057, 1969.

[3] V. N. Gudivada and V. V. Raghavan, **"Content based image retrieval systems,"** *Computer*, Vol. 28, No. 9, pp. 18-22, 1995.

[4] J. A. da Silva Júnior, R. E. Marçal, and M. A. Batista, **"Image retrieval: importance and applications,"** in *Workshop de Vis~ ao Computacional-WVC*, 2014.

[5] M. S. Lew, N. Sebe, and J. P. Eakins, **"Challenges of image and video retrieval ",** **i**n *International Conference on Image and Video Retrieval*, Springer, pp. 1-6, 2002.

[6] X.-y. Wang, Z.-f. Chen, and J.-j. Yun, **"An effective method for color image retrieval based on texture,"** *Computer Standards & Interfaces*, Vol. 34, No. 1, pp. 31-35, 2.012

[7] G.-H. Liu and J.-Y. Yang, **"Content-based image retrieval using color difference histogram,"** *Pattern recognition*, Vol. 46, No. 1, pp. 188-198, 2013.

[8] A. Giri and Y. K. Meena, **"Content based image retrieval using integration of color and texture features,"** *International Journal of Advanced Research in Computer Engineering & Technology*, Vol. 3, 2014.

[9] A. Huneiti and M. Daoud, **"Content-based image retrieval using SOM and DWT,"** *Journal of Software Engineering and Applications*, Vol. 8, No. 2, pp. 510, 2015.

[10] M. Kaur and N. Sohi, **"A novel technique for content based image retrieval using color, texture and edge features,"** in *2016 international conference on communication and electronics systems (ICCES)*, IEEE, pp. 1-7, 2016.

[11] S. Fadaei, R. Amirfattahi, and M. R. Ahmadzadeh, **"New content-based image retrieval system based on optimised integration of DCD, wavelet and curvelet features,"** *IET Image Processing*, Vol. 11, No. 2, pp. 89-98, 2016.

[12] N. Sai and R. C. Patil, **"DCT-SVD domain feature vector for image retrieval,"** in *2017 International Conference on Signal Processing and Communication (ICSPC)*, 2017: IEEE, pp. 477-482, 2017.

[13] N. Jain and S. Salankar, **"Content based image retrieval using improved gabor wavelet transform and linear discriminant analysis,"** in *2018 3rd International Conference for Convergence in Technology (I2CT)*, IEEE, pp. 1-4, 2018.

[14] X. Tian, Q. Zheng, and J. Xing, **"Content-Based Image Retrieval System Via Deep Learning Method,"** in *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, IEEE, pp. 1257-1261, 2018.

[15] L. Liu, S. Lao, P. W. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, **"Median robust extended local binary pattern for texture classification,"** *IEEE Transactions on Image Processing* , Vol. 25, No. 3, pp. 1368-1381, 2016.