International Journal of Imaging and Robotics,
[Formerly known as the "International Journal of Imaging" (ISSN 0974-0627)]
Volume 18; Issue No. 2; Year 2018; Int. J. Imag. Robot.
ISSN 2231–525X; Copyright © 2018 [International Journal of Imaging and Robotics]

CBIR based on Weighted Multi-feature Voting Technique

Walaa E. Elhady[1], **Abdelwahab k. ALSammak** [2] **and Shady Y. El-Mashad**[3]

[1]Faculty of Engineering, October 6 University, Egypt
Walaa.elhady@o6u.edu.eg

[2]Faculty of Engineering at Shoubra, Benha University, Egypt
asammak@feng.bu.edu.eg

[3]Faculty of Engineering at Shoubra, Benha University, Egypt
shady.elmashad@feng.bu.edu.eg

ABSTRACT

Content-Based Image Retrieval (CBIR) has received a comprehensive attention from researchers due to the quickly growing and the diffusion of image databases. Despite the huge research efforts consumed for CBIR, the completely promising results have not yet been presented. In this paper, a novel weighted multi-feature voting technique is proposed which incorporates various types of low-level visual features such as texture, shape and color in retrieval process. The color feature is described by color histogram and hierarchical annular histogram whereas shape feature is described by edge histogram and edge direction histogram while texture feature is described by gabor filter and co-occurrence matrix. Each feature has certain weight computed based on its precision to reflect its importance in retrieval procedure. Furthermore, different distance measures are implemented to get the highest precision of each feature. The results indicate that by applying multi-features and multi-distance measures, the obtained retrieval system outperforms other existing methods with accuracy 89.5% for Wang database, 91.5% for Caltech101 database and 89% for UW database.

Keywords: CBIR, Feature extraction, Weighted average, Matching measures, Weighted multi-feature voting.

Mathematics Subject Classification: 62H35, 68U10.

1 Introduction

An image retrieval framework is a computerized scheme developed to manage (browse, search and retrieve) digital images within huge databases. Currently, the size of digital image collection increments rapidly due to the huge extension of the internet as well as the approachability of image capturing devices as digital cameras and image scanners. Thus, there is a great motivation to develop efficient and effective tools for searching, browsing and retrieving images by users from various areas, including medicine, remote sensing, publishing, architecture, crime prevention ... etc. To attain this goal, research efforts guided to develop various general-purpose image retrieval schemes. Nowadays, practically all human life applications use images to get efficient services. A massive collection of these images is indicated as an image database. An image database is an organized structure of digital images where a large number of images are stored and queried.

Over the last few years, many researchers have been performed on image retrieval. These investigations can be classified into three different domains based on the type of the applied methodology; text-based approach, context-based approach and content-based approach. In text-based methodology, retrieval procedure is achieved by adding metadata like captions, keywords or text to the images so that retrieval can be accomplished over the annotation words. Images are manually annotated and subsequently retrieved in the same fashion as text documents using a database management system. Moreover, conventional annotation has three disadvantages: Manual annotation requires significant level of human effort; the annotation is inexact due to the subjectivity of human perceptiveness, in addition to the Polysemy problem which means that the same word can indicate more than one object (Zhang, Islam and Lu, 2012; Markkula and Sormunen, 2000).

These problems drew attentiveness to image retrieval approaches based on the content.

Content-Based Image Retrieval (CBIR) approaches query the images with their actual contents instead of their annotated metadata such as keywords, tags or text descriptions. Elementary CBIR methods automatically indexed and retrieved with low-level visual features like texture, shape spatial information or color (Yasmin, Sharif and Mohsin, 2013; Danish, Rawat and Sharma, 2013; El-Mashad and Shoukry, 2014b; El-Mashad and Shoukry, 2015).

In this paper, a novel CBIR system is proposed using a weighted feature voting technique with multi-features and multi distance measures. The proposed system provides a considerable enhancement in the overall performance compared to the existing methods.

The rest of the paper is organized as follows. In Section 2, content based image retrieval fundamentals and related works are presented. Section 3, describes the architecture of the proposed system. The experimental results are carried out on three databases. Comparisons between the proposed system and other existing methods are presented in Section 4. Finally, the conclusion and future work as well are provided in Section 5.

2 Related Work

Youness et al. (Youness, Mohammed and Brahim, 2016) presented approach of content based image retrieval based on extraction content frequency from image. 2-D ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) method is used to extract the frequency content from the image for constructing the vector descriptor. This method can be classified in the category of texture. This approach has tested to the Coil-100 database, and the experimental results showed that the average precision is 96.94%.

CBIR technique that relies on extracting Speeded Up Robust Features (SURF) and Maximally Stable Extremal Regions (MSER) feature descriptors as well as the color features; color correlograms and Improved Color Coherence Vector (ICCV) is offered by Heba Elnemr (Elnemr, 2016). These features are joined and used to build a multidimensional feature vector. Bag-of-Visual-Words (BoVW) technique is utilized to quantize the extracted feature vector. Then, a multiclass Support Vector Machine (SVM) is implemented to classify the query images. This method is tested on two benchmark datasets; Corel-1000 and COIL-100 datasets. The system achieved average precisions of 88% and 93% for the Corel-1000 and COIL-100 datasets, respectively.

feature extraction technique by binarisation with Sauvolas local threshold selection algorithm has been proposed by the Das et al.(Das, Thepade and Ghosh, 2017). this technique has evaluated on 17,021 images for performance assessment. The precision results for classification and retrieval have shown an increment of 17% and 13.1% respectively when compared to state-of-the-art techniques.

Content based image retrieval using gray scale weighted average method for reducing the feature vector dimension has been proposed by Kumar et al. (Kumar, Li, Shaikh et al., 2016). This method has been divided into six steps, firstly, RGB color image is converted into Gray scale Image. Then image histogram is generated and the sum of the occurrence of all unique gray shades is calculated. The probability distribution function for finding the occurrence of individual unique value in the grayscale image array is applied. Moreover, the weighted average of unique values and their corresponding probabilities is calculated. Finally, the feature vector value for grayscale image is generated. The Euclidean distance between the query image feature vector and image feature vector in the database is computed for retrieving most similar images. To evaluate this system Wang and Amsterdam Library of Texture Images (A LOT) for color and texture have been used and the experimental results showed that the average precision is 70%, 61% for Wang and A LOT datasets, respectively.

Sorted Block Truncation Coding (SBTC) method of feature extraction is proposed by Sudeep Thepade et.al (Thepade, Das and Ghosh, 2015). Proposed Sorted BTC has implemented to near weighted mode and the similarity measure is calculated by Euclidean distance. Extracted feature vectors from the color components Red (R), Green (G) and Blue (B), where each color component has been considered as a block. Each block has been divided into bins of sorted intensity values. The average of sorted intensity values in each bin has been considered to form the feature vector of that block. The generated feature vectors of the blocks have been combined to create the feature vector of the image. The process represented the intensity values of an image within a single dimensional array. The single dimensional array has been then sorted in ascending order. The sorted array has been divided into N blocks to calculate the average of intensities in each block to generate the feature vectors. Two different classifiers have been used for the comparison of performance of classification processes, which are K-Nearest Neighbor (KNN) classifier and Neural Network (NN) Classifier. The classification process has been carried out by measuring the distance between the query image and the database images. The classification has been done to the category which has the minimum distance from the query image. The method is evaluated through experiments on Wang and Caltech101 datasets and gives a precision of 78% for Wang dataset and 68.2% for Caltech101 dataset.

Anandh et.al (Anandh, Mala and Suganya, 2016) proposed a technique for the generation of image content descriptor with three features which are Color auto-Correlogram, Gabor Wavelet and Wavelet Transform. Color Auto-Correlogram feature is associated with color information of an image which is derived from the RGB color space of an image. The Gabor Wavelet feature has texture information to extract textural features associated with the image and the Wavelet Transform feature is linked with shape information in the extraction of edges in an image. The extracted features are stored in a feature dataset. The Manhattan distance is applied on the user given query image and feature vector computed from database images for measuring similarity. Finally, the proposed technique retrieves the meaningful image from the image database which satisfies the user expectation. The performance of the retrieval system has been analyzed by the performance measures Precision and Recall. The efficiency of the feature descriptor is tested for CBIR system using Wang database, Li database and Caltech-101 database. The method achieved an average accuracy rate of 83% for corel database, whereas 88% for Li database and 70% for Caltech-101 database in Content Based Image Retrieval system.

3 Proposed System

In weighted multi-feature voting mechanism, each feature retrieves images of different classes. The number of retrieved images in each class represents the vote of this class. Then the votes of all features are aggregated for a certain class through the weights which reflect the relative importance of the features. For example, if we have two features f1 and f2 with weights 0.8 and 0.7 respectively, and the votes of f1 and f2 for class c are 5 and 4 respectively, then the overall votes for this class will be (0.8*5)+(0.7*4).

The CBIR proposed system is depicted in Fig. 1. In this system, the features of the database images are extracted and stored. Then the weight of each feature is calculated separately. The features are extracted from the query image and compared to those of the database images through the matching measures to calculate the distances. Based on these distances and the feature weights, the class of the query image is identified using the weighted multi-feature voting mechanism as elaborated in Section 3.3.1. Finally, the related images are retrieved based on the feature that returned the maximum number of images in the identified class.

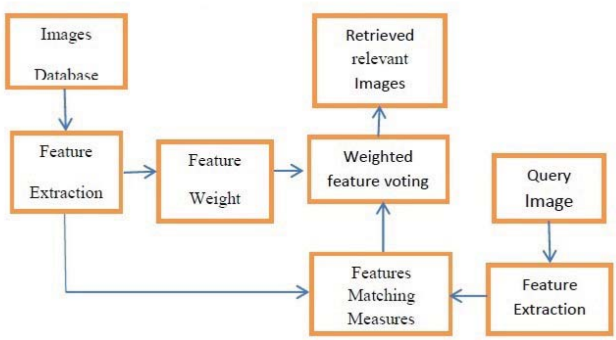

Figure 1: The proposed CBIR system.

3.1 Feature Extraction

The first step in the retrieval process concerns with extracting the most distinguishable image features. The selection of these features to represent an image is one of the keys of a CBIR system(El-Mashad and Shoukry, 2014a). In this system, features are extracted using color histogram, edge histogram, edge direction histogram, hierarchical annular histogram, Gabor filter, and co-occurrence matrix vector (Manjunath, Ohm, Vasudevan and Yamada, 2001; Jain and Vailaya, 1996; Yang, Qi, Xing, Kurc, Saltz and Foran, 2013; Daugman, 1988; Haralick, Shanmugam et al., 1973).

Color is one of the most widely used features. Color histogram is the most popular technique, which represents the global distribution of colors in an image. However In large database, it doesn't have good performance because it stores only color information and lacks spatial information, so images with very different appearances can have similar histograms (Fadaei and Sortrakul, 2014). To enhance the performance, hierarchical annular histogram (HAH) (Yang et al., 2013) is used, which is rotation invariant and can capture the spatial configuration of pixel intensities throughout the image. The image is segmented into consecutive concentric rectangles, within the rest of each rectangular ring the intensity histogram for RGB channels is calculated and concatenated together as a feature vector called HAH. Because HAH takes into consideration the spatial configuration of the features, it can differentiate between images with similar total intensity distribution, but different in spatial intensity configurations (Qi, Gensure, Foran and Yang, 2013). This technique was originally developed for medical applications (Yang et al., 2013; Qi et al., 2013). HAH gaves a relatively high precision in the proposed system.

Depending only on color features makes the system confused with classes which having the same color. So, to reduce the effect of color features, the texture and shape features are used. Edges in images constitute an important feature to represent their content. Also, human eyes are sensitive to edge features for image perception. Therefore, we combined the pervious implemented features with edge histogram and edge direction histogram. An edge histogram in the image space represents the frequency and the directionality of the brightness changes in the image (Won, Park and Park, 2002). It represents the local edge distribution in the image which is obtained by subdividing the whole image into 4x4 sub-images. For each of these sub-images the histogram is computed for five directional edges; vertical, horizontal, 45 degree, 135 degree, and one non-directional edge. Thus the obtained vector is 16x5 for each image(Manjunath et al., 2001; Nandagopalan, Adiga and Deepak, 2008).

In addition, an edge direction histogram is used. The Sobel edge operator is firstly applied to retrieve the global edge points and then a histogram of the directions of the edge points is used to represent the shape attributes.

Texture refers to the visual patterns that have properties of homogeneity. It does not result from the existence of only a single color or intensity (Smith and Chang, 1996).

A co-occurrence matrix is created from a grayscale image. The co-occurrence matrix was originally proposed by R.M. Haralick (Haralick et al., 1973). This technique constructs a co-occurrence matrix on the basis of orientation and the distance between the grey level values. Then statistical features are extracted from the co-occurrence matrix to obtain the texture representation from an image. The advantage of the co-occurrence matrix is that the co-occurring pairs of pixels can be spatially related in various orientations with reference to distance and angular spatial relationships, as on considering the relationship between two pixels at a time. As a result the combination of gray levels and their positions are exhibited apparently. Therefore, it is defined as a two dimensional histogram of grey levels for pair of pixels, which are separated by a fixed spatial relationship (Lingadalli and Ramesh, 2015). From the co-occurrence matrix we extracted contrast, energy, correlation and homogeneity texture feature.

On the other hand, Gabor filter consists of a group of wavelets each of which capturing energy at a specific resolution and orientation. Therefore, Gabor filter can capture the local energy of the entire signal or image. The Gabor filter has been widely used to extract image features, especially texture features(Daugman, 1988). Daugman discovered that Gabor filter provides optimal Heisenberg joint resolution in visual space and spatial frequency. For this reason, Gabor filter has been successfully employed in many image processing applications. However, Gabor filter produces information redundancy that can be reduced by downsampling the feature images (Liu and Wechsler, 2002; Shen, Bai and Fairhurst, 2007).

The precision of each feature is calculated to determine the role of each feature in extracting relevant images. Features with higher precision play important role in determining the relevant images. The precision is calculated as in Eq.3.1 (Bala and Sharma, n.d.):

$$precision = \frac{No.of relavent\ images\ retrieved}{Total\ No.of\ images\ retrieved} \tag{3.1}$$

3.2 Matching Measures

Different features of each image are represented by vectors. These vectors are calculated offline. A user can use an image as a query to retrieve similar images from the database. The feature vector of the query image is computed. Then, the similarity between the query image and all other images in the database are calculated using distance metrics between the query feature vector and the database feature vectors. Small distances mean more similarity. Distance metrics used in the proposed system for comparison are Histogram intersection(Smith, 1997), Euclidean distance(Sergyan, 2008), and Cosine distance (Kaur and Aggarwal, 2013).

3.3 Feature Weight

The proposed fusion mechanism uses weights to determine the role of each feature in retrieving the relevant image. Therefore, these weights reflect the importance of each individual feature. In this paper, the proposed CBIR system calculates these weights offline once for the entire training set features. These weights are used later as shown in the next section to address the relevant images. Each feature weight is calculated as the average precision of this feature over the entire database. The weights are arranged in a row vector w as in Eq.3.2:-

$$\mathbf{w} = [w_1 \;\; w_2 \;\; w_3 \;\; w_4 \;\; w_5 \;\; w_6] \tag{3.2}$$

Where w_i is the weight of the i^{th} feature. w_i is calculated as in Eq.3.3

$$w_i = \frac{1}{N}\sum_{j=1}^{N} p_{ij} \tag{3.3}$$

Where N is no. of images in the database and p_{ij} is the accuracy of the i^{th} feature to retrieve the j^{th} image in the database which is the precision at $k = 10$ as shown in Eq 3.4(Zakariya and Akhtar, 2014):

$$p_(ij) = \frac{no.of\ relevant\ images\ at\ k = 10}{k = 10} \tag{3.4}$$

Where, k represents the top ranked images.

3.3.1 Weighted Multi-feature Voting

In the proposed voting mechanism, each feature has a specific precision for retrieving each image in the database. To increase the overall precision of the retrieval system, multi-feature mechanism is utilized such that the overall precision is higher than the individual feature precision. The importance of each feature is reflected by its weight, which is its average precision all over the database.

This multi-feature mechanism is inspired by how a decision is taken in a parliament by voting of the members. The members are the features. Each feature returns ten images which are the closest to the query from its point of view. These ten images of each feature are arranged in the voting matrix V as in Eq.3.5:

$$\mathbf{V} = \begin{bmatrix} c_{11} & \dots & c_{1C} \\ & . & \\ & . & \\ c_{F1} & \dots & c_{FC} \end{bmatrix} \tag{3.5}$$

Where c_{ij} is the number of images that returned by the i^{th} feature that belong to the j^{th} class in the database. F is the number of features and C is the number of classes in the database. Therefore, the voting matrix V reflects how each feature classifies the query from its point of view. As an interpretation of the voting matrix, c_{ij} is considered as how likely the query belongs to the j^{th} class according to the opinion of the i^{th} feature. All features votes are fused into score vector s as shown in Eq. 3.6:

$$\mathbf{s} = \mathbf{w}.\mathbf{V} = \begin{bmatrix} s_1 & \dots & s_C \end{bmatrix} \tag{3.6}$$

Where w is the weight vector and V is the voting matrix. s_i is the i^{th} component of the score vector, which reflects how likely the query belong to the i^{th} class taking into consideration all the features as per their importance revealed by weights. The class of the query is determined as in Eq. 3.7:

$$Class = \underset{i}{\operatorname{argmax}}\ s_i,\ \ j = 1, \dots, C \tag{3.7}$$

In other words, this mechanism determines in the first place the class in the database to which the query belongs. In the last step, images from the determined class are retrieved. In this paper, it is chosen to retrieve the top ranked ten images that belong to the determined class according to the feature that returned the maximum number of images in this class in voting matrix. If the query belongs to the i^{th} class, then from the voting matrix the feature that gives the highest vote to this class is calculated as in Eq.3.8:

$$Feature = \underset{j}{\operatorname{argmax}}\ c_{j\ Class},\ \ j = 1, \dots, F \tag{3.8}$$

Then the retrieved images are the highest ranked ten images returned by the feature $Feature$ that belong to the class $Class$.

4 Experimental Results

The proposed image retrieval system is tested and evaluated on three widespread images datasets. The first dataset is Wang database(Wang, Li and Wiederhold, 2001) , it contains 1000 images categorized into 10 classes which are African people, beach, building, bus, dinosaur, elephant, flower, horse, mountain and food as shown in Fig.2. Each class includes 100 images. For each class, 80 images are utilized to train the system and 20 images are exploited to test the system (i.e., 800 and 200 images for training and testing, respectively).
The second one is Caltech 101 database (Fei-Fei, Fergus and Perona, 2007). It contains 101 categories which have huge variances in shape, color and texture. Each object category contains about 31 to 800 images. The size of each image is roughly 300 x 200 pixels. These categories are airplane, bonsai, panther, dalmatian, dolphin, faces, flamingo, deer, piano, skates, metronome, minar, motorbike, panda, football, stopsign, sunflower, trees, monument, watches, . . . etc as shown in Fig. 3. 20 images from each group are selected for training while 10 sample images are selected for testing. Thus, there is a set of 2020 images are reserved for training while 1010 images are used for testing.
Finally, UW dataset is tested(Shapiro, 2005). The UW dataset consists of 855 images belonging to 19 categories. The database is created at the University of Washington. The images are of various sizes and mainly include vacation pictures from various locations, for example spring owers, Barcelona, and Iran. Some example images are shown in Fig.4. 30 images from each group are used for training and 15 sample images are selected for testing. Hence, 570 images are assigned for the training task while 285 images are kept for the testing task.

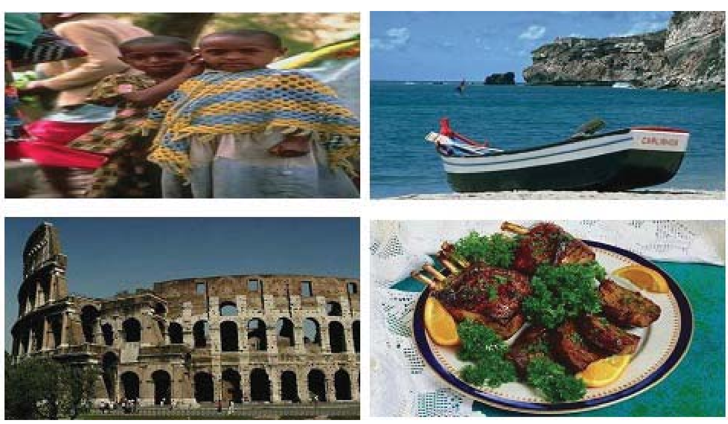

Figure 2: Example images from the WANG database.

Table 1 represents the average accuracy for each individual feature. It's evident that, using certain matching measure with all features is not efficient. However, depending on color histogram alone is not enough to build a robust retrieval system because it is based on color only (Huang,

Figure 3: Example images from the Caltech 101 database.

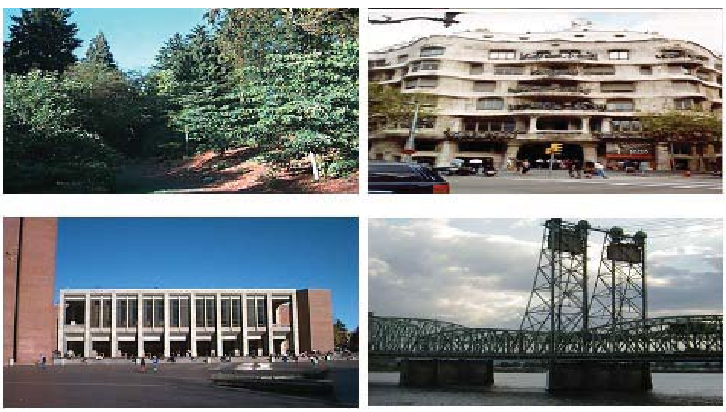

Figure 4: Example images from the UW database.

Table 1: Comparison between the accuracy of individual features by using different matching distance measures

Matching Measures	Color histogram	Edge direction	Edge histogram	Hierarchical annular histogram	Gabor filter	Co-occurrence matrix
Histogram intersection	**0.7632**	0.2572	0.1774	**0.6647**	0.2408	0.3289
Cosine distance	0.6855	0.3710	0.3909	0.6557	**0.3212**	**0.4163**
Euclidean distance	0.6610	**0.4670**	**0.4869**	0.6145	0.2870	0.3861

Shu, Ma and Gong, 2015; Anandh et al., 2016; Walaa, Abdulwahab and Shady, 2018) . So, the system cannot distinguish between images which contain objects with the same color. For example, images contain sky and clouds are not distinguishable from images contain sea. The proposed multi-feature system utilizes the features through the weighted multi-feature voting technique to identify the class of the query image. The results are compared to the weighted average technique as shown in Table 2. The weighted average technique is used to fuse distances between the database images and the query image as shown in Eq.4.1.

$$\mathbf{d}_{av} = \frac{\mathbf{w} \times \mathbf{M}}{\sum_{i=1}^{6} w_i} \tag{4.1}$$

Where M is distance matrix

$$\mathbf{M} = \begin{bmatrix} d_{11} & \dots & d_{1N} \\ & . & \\ & . & \\ d_{61} & \dots & d_{6N} \end{bmatrix}_{6 \times N} \tag{4.2}$$

N is the number of images in the database and d_{ij} is the i^{th} feature distance between the j^{th} image and the query image.

Table 2: Comparison between average accuracy of weighted multi-feature voting technique and weighted average technique by different matching distance measures.

Weight techniques	Histogram intersection	Cosine distance	Euclidean distance
Weighted average	59.0800	77.7300	58.7600
Weighted multi-feature voting	67.5000	85.2014	83.0000

As shown in Table 2, the results achieved using weighted multi-feature voting technique are better than that using the weighted average technique.

A further enhancement is obtained by using each feature with the most suitable matching measure according to Table1. For example, color histogram is used with histogram intersection distance, Gabor filter used with cosine distance and etc.
Wang, Caltech 101 and UW databases are used to test and evaluate the proposed system as mentioned previously. Table3, illustrates the precision for each class in the Wang database, whereas the average accuracy for the utilized databases are listed in Table 4. From Table 3, we can observe that a precision of 100% has been obtained for classes African people, Bus, Dinosaur, Flower and Horse. Also, Table 4 stated that the overall average precision are 89.5%, 91.5% and 89% for Wang, Caltech101 and UW datasets, respectively.
Table 5 illustrates the results of the experiments, performed on the Wang database and compared to the results of existing methods in (Fadaei and Sortrakul, 2014; Thepade et al., 2015; Das, Thepade, Bhattacharya and Ghosh, 2016; Anandh et al., 2016; Das et al., 2017; Walaa et al., 2018). It's evident that, the proposed system has better accuracy than other existing methods for individual classes of Wang database. The overall average precision has been raised to 89.5%. Table 6, illustrates the average accuracy of the proposed approach and other existing methods using Caltech101 dataset. The results prove that, the proposed system outperforms other techniques with a precision of 91.5%. On the other hand, Table 7, represents

Table 3: the precision for each class in the Wang database using weighted multi-feature voting technique.

Classes	Description	Proposed Method
1	African people	100%
2	Beach	65%
3	Building	95%
4	Bus	100%
5	Dinosaur	100%
6	Elephant	85%
7	Flower	100%
8	Horse	100%
9	Mountain	60%
10	Food	90%

Table 4: Comparison between the average accuracy for Wang, Caltech 101 and UW database using the proposed system.

Databases	Wang	Caltech 101	UW
Average accuracy	89.5%	91.5%	89%

the overall average precision of the proposed system and other existing methods as well using UW database. The proposed system has better performance than other existing methods with accuracy of 89%.

Table 5: Comparison among the average accuracy for each class in the Wang database using proposed system and other existing methods

Classes	Description	(Fadaei et al.,2014)	(Thepade et al.,2105)	(Ghosh et al.,2016)	(Anandh et al.,2016)	(Das et al.,2017)	(Walaa et al.,2018)	proposed system
1	African people	72.10%	80%	80%	80%	80%	100%	100%
2	Beach	61.20%	80%	60%	84%	60%	60%	65%
3	Building	52.25%	80%	60%	82%	40%	75%	95%
4	Bus	91.30%	80%	80%	90%	80%	100%	100%
5	Dinosaur	96.40%	100%	100%	100%	100%	100%	100%
6	Elephant	66.40%	80%	80%	72%	60%	80.5%	85%
7	Flower	87.05%	80%	80%	98%	100%	100%	100%
8	Horse	94.85%	100%	100%	88%	100%	100%	100%
9	Mountain	52.35%	40%	80%	66%	60%	60%	60%
10	Food	73.20%	60%	60%	70%	60%	85%	90%
Average accuracy		74.71%	78%	78%	83%	74%	86.5%	**89.5**%

Table 6: Comparison between the average accuracy for Caltech 101 database using proposed system and other existing methods.

Method	(Zhu et al.,2014)	(Thepade et al.,2015)	(Anandh et al.,2016)	(Neelima et al.,2016)	(Das et al.,2017)	(Walaa et al.,2018)	Proposed system
Average accuracy	48.6%	68.2%	70%	72%	84.8%	85%	**91.5**%

Table 7: Comparison between the average accuracy for UW database using proposed system and other existing methods.

Method	(Mahajan et al.,2014)	(Yang et al.,2014)	(Walaa et al.,2018)	proposed system
Average accuracy	80%	67.46 %	86.5%	**89**%

5 Conclusion

In this paper, a novel CBIR system is proposed using a weighted multi-feature voting technique with multi-features and multi-distance measures. As noticed, from the previous approaches, the majority features used are color features, so the system confused between classes which share the same color. The proposed system depends on various types of low level visual features such as color, shape and texture. The proposed system provides a considerable enhancement in the overall performance compared to traditional weighted average techniques and other existing methods. The overall average precision is 89.5%, 91.5% and 89% for Wang, Caltech101 and UW datasets, respectively.
As a future work, the weight of each feature can be optimized using heuristic optimization technique for better performance. In addition, a feedback learning algorithm can be adopted for more improvement.

References

Anandh, A., Mala, K. and Suganya, S. 2016. Content based image retrieval system based on semantic information using color, texture and shape features, *Computing Technologies and Intelligent Data Engineering (ICCTIDE), International Conference on*, IEEE, pp. 1–8.

Bala, R. and Sharma, R. n.d.. Efficient method of cbir using combination of shape, color and texture feature with fuzzy.

Danish, M., Rawat, R. and Sharma, R. 2013. A survey: content based image retrieval based on color, texture, shape & neuro fuzzy, *image* **2**: 3.

Das, R., Thepade, S., Bhattacharya, S. and Ghosh, S. 2016. Retrieval architecture with classified query for content based image recognition, *Applied Computational Intelligence and Soft Computing* **2016**: 2.

Das, R., Thepade, S. and Ghosh, S. 2017. Novel feature extraction technique for content-based image recognition with query classification, *International Journal of Computational Vision and Robotics* **7**(1-2): 123–147.

Daugman, J. G. 1988. Complete discrete 2-d gabor transforms by neural networks for image analysis and compression, *IEEE Transactions on Acoustics, Speech, and Signal Processing* **36**(7): 1169–1179.

El-Mashad, S. Y. and Shoukry, A. 2014a. Evaluating the robustness of feature correspondence using different feature extractors, *Methods and Models in Automation and Robotics (MMAR), 2014 19th International Conference On*, IEEE, pp. 316–321.

El-Mashad, S. Y. and Shoukry, A. 2014b. A more robust feature correspondence for more accurate image recognition, *Computer and Robot Vision (CRV), 2014 Canadian Conference on*, IEEE, pp. 181–188.

El-Mashad, S. Y. and Shoukry, A. 2015. Towards a robust scale invariant feature correspondence, *International Conference on Pattern Recognition and Machine Intelligence*, Springer, pp. 33–43.

Elnemr, H. A. 2016. Combining surf and mser along with color features for image retrieval system based on bag of visual words, *Journal of Computer Science* **12**(4): 213–222.

Fadaei, H. and Sortrakul, T. 2014. Content-based image retrieval system with combining color features and gradient feature, *Int. Conference on Advanced Computational Technologies & Creative Media*.

Fei-Fei, L., Fergus, R. and Perona, P. 2007. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories, *Computer vision and Image understanding* **106**(1): 59–70.

Haralick, R. M., Shanmugam, K. et al. 1973. Textural features for image classification, *IEEE Transactions on systems, man, and cybernetics* **3**(6): 610–621.

Huang, M., Shu, H., Ma, Y. and Gong, Q. 2015. Content-based image retrieval technology using multi-feature fusion, *Optik-International Journal for Light and Electron Optics* **126**(19): 2144–2148.

Huang, Z.-C., Chan, P. P., Ng, W. W. and Yeung, D. S. 2010. Content-based image retrieval using color moment and gabor texture feature, *Machine Learning and Cybernetics (ICMLC), 2010 International Conference on*, Vol. 2, IEEE, pp. 719–724.

Jain, A. K. and Vailaya, A. 1996. Image retrieval using color and shape, *Pattern recognition* **29**(8): 1233–1244.

Kaur, S. and Aggarwal, D. 2013. Image content based retrieval system using cosine similarity for skin disease images, *Advances in Computer Science: an International Journal* **2**(4): 89–95.

Kumar, K., Li, J.-P., Shaikh, R. A. et al. 2016. Content based image retrieval using gray scale weighted average method, *IJACSA) International Journal of Advanced Computer Science and Applications* **7**(1).

Lingadalli, R. and Ramesh, N. 2015. Content based image retrieval using color, shape and texture, *International Advanced Research Journal in Science, Engineering and Technology* **2**(6): 40–46.

Liu, C. and Wechsler, H. 2002. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition, *IEEE Transactions on Image processing* **11**(4): 467–476.

Long, F., Zhang, H. and Feng, D. D. 2003. Fundamentals of content-based image retrieval, *Multimedia Information Retrieval and Management*, Springer, pp. 1–26.

Manjunath, B. S., Ohm, J.-R., Vasudevan, V. V. and Yamada, A. 2001. Color and texture descriptors, *IEEE Transactions on circuits and systems for video technology* **11**(6): 703–715.

Markkula, M. and Sormunen, E. 2000. End-user searching challenges indexing practices in the digital newspaper photo archive, *Information retrieval* **1**(4): 259–285.

Nandagopalan, S., Adiga, D. B. and Deepak, N. 2008. A universal model for content-based image retrieval, *jiP* **1**: 5.

Pass, G. and Zabih, R. 1996. Histogram refinement for content-based image retrieval, *Applications of Computer Vision, 1996. WACV'96., Proceedings 3rd IEEE Workshop on*, IEEE, pp. 96–102.

Qi, X., Gensure, R. H., Foran, D. J. and Yang, L. 2013. Content-based white blood cell retrieval on bright-field pathology images, *SPIE Medical Imaging*, International Society for Optics and Photonics, pp. 86760L–86760L.

Sergyan, S. 2008. Color histogram features based image classification in content-based image retrieval systems, *Applied Machine Intelligence and Informatics, 2008. SAMI 2008. 6th International Symposium on*, IEEE, pp. 221–224.

Shapiro, L. G. 2005. Object and concept recognition for content-based image retrieval, *Dept of Computer Science and Engineering, University of Washington, NSF grant IIS-0097329* .

Shen, L., Bai, L. and Fairhurst, M. 2007. Gabor wavelets and general discriminant analysis for face identification and verification, *Image and Vision Computing* **25**(5): 553–563.

Smith, J. R. 1997. *Integrated spatial and feature image systems: Retrieval, analysis and compression*, PhD thesis, Columbia University.

Smith, J. R. and Chang, S.-F. 1996. Automated binary texture feature sets for image retrieval, *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on*, Vol. 4, IEEE, pp. 2239–2242.

Thepade, S., Das, R. and Ghosh, S. 2015. Novel technique in block truncation coding based feature extraction for content based image identification, *Transactions on Computational Science XXV*, Springer, pp. 55–76.

Walaa, E., E., Abdulwahab, K., A. and Shady, Y., E.-M. 2018. Weighted feature voting technique for content-based image retrieval, *Int. J. Computational Vision and Robotics* **X**(Y): 1–17.

Wang, J. Z., Li, J. and Wiederhold, G. 2001. Simplicity: Semantics-sensitive integrated matching for picture libraries, *IEEE Transactions on pattern analysis and machine intelligence* **23**(9): 947–963.

Won, C. S., Park, D. K. and Park, S.-J. 2002. Efficient use of mpeg-7 edge histogram descriptor, *ETRI journal* **24**(1): 23–30.

Yang, L., Qi, X., Xing, F., Kurc, T., Saltz, J. and Foran, D. J. 2013. Parallel content-based sub-image retrieval using hierarchical searching, *Bioinformatics* p. btt623.

Yasmin, M., Sharif, M. and Mohsin, S. 2013. Use of low level features for content based image retrieval: Survey, *Research Journal of Recent Sciences, ISSN* **2277**: 2502.

Youness, C., Mohammed, O. and Brahim, A. 2016. Cbir using the 2-d esprit method: Application to coil 100 database, *International Journal of Imaging and Robotics* **16**(2): 66–77.

Zakariya, S. and Akhtar, N. 2014. Analysis of unsupervised content-based image retrieval by fusing the image features, *Proc. of Int. Conf. on Recent Trends in Information, Telecommunication and Computing* pp. 328–337.

Zhang, D., Islam, M. M. and Lu, G. 2012. A review on automatic image annotation techniques, *Pattern Recognition* **45**(1): 346–362.