# Weighted feature voting technique for content-based image retrieval


## Walaa E. Elhady*

Faculty of Engineering,
October 6 University, Egypt
Email: Walaa.elhady@o6u.edu.eg
*Corresponding author

## Abdulwahab K. Alsammak and Shady Y. El-Mashad

Faculty of Engineering (Shoubra),
Benha University, Egypt
Email: asammak@feng.bu.edu.eg
Email: shady.elmashad@feng.bu.edu.eg


**Abstract:** A content-based image retrieval process is used to retrieve most similar images to a query from a large database of images on the basis of extracted features. Matching measures are used to find similar images by measuring how the query features are close to the features of other images in the database. In this paper, a multi-features system is proposed which incorporates more than one feature in the retrieval process. The weights of these features are calculated based on the precision of each feature to reflect its importance in the retrieval process. These weights are used in a weighted feature voting technique to incorporate the role of each feature in extracting the relevant images. Also, different distance measures are used to get the highest precision of each feature. The results of applying the multi-features and multi-distances measures technique outperform other existing methods with accuracy 86.5% for Wang database, 86.5% for UW database and 85% for Caltech101 database.

**Keywords:** content based image retrieval; computational vision; feature extraction; hierarchical annular histogram; weighted average; matching measures; weighted feature voting.

**Reference** to this paper should be made as follows: Elhady, W.E., Alsammak, A.K. and El-Mashad, S.Y. (2018) 'Weighted feature voting technique for content-based image retrieval', *Int. J. Computational Vision and Robotics*, Vol. 8, No. 3, pp.283–299.

**Biographical notes:** Walaa E. Elhady received her BE in Computer Engineering from the October 6 University in 2006 and her Master's in Computer Engineering from the Arab Academy for Science and Maritime Transport (AASTMT) in 2012. She has more than ten years of teaching experience. Currently, she is preparing a PhD at the Benha University in the Faculty of Engineering at Shoubra in the area of content-based image retrieval. She has publications in image processing and computer vision. In addition, she is interested in pattern recognition, image retrieval and video processing.

Abdulwahab K. Alsammak is an Associate Professor in the Electrical Engineering Department at the Faculty of Engineering (Shoubra) at the Benha University. He is the Head of Computer Systems Engineering Program. He obtained his PhD, MSc and BSc from the Zagazig University in 1992, 1987, 1982, respectively. He also works as the Consultant of the Egyptian Universities Portal Project, Ministry of Higher Education in Egypt since 2012. His research interests include artificial intelligence, natural language processing, data mining and software engineering.

Shady Y. El-Mashad received his PhD in Computer Systems Engineering from the E-JUST in 2015, MSc in Computer Systems Engineering from the Benha University in 2011 and BSc in Computer Systems Engineering from the Benha University in 2005. He has about 12 years of experience in teaching research. Currently, he is an Assistant Professor at the Computer Systems Engineering Department in the Faculty of Engineering at Shoubra at the Benha University. In addition, he is the Director of Portal Project – Benha University.

## 1 Introduction

Retrieval is the wide topic of research for decades. The concept behind is to get desired data from the database. It may be image, text, audio or video according to user requirement (Sawant et al., 2012).

In this computer age, virtually all spheres of human life, including commerce, government, academics, hospitals, crime prevention, surveillance, engineering, architecture, journalism, fashion and graphic design and historical research use images for efficient services. In a small collection of images, simple browsing can identify an image. This is not the case for large and varied collection of images, where user encounters an image retrieval problem. The image retrieval problem is the problem encountered when searching and retrieving images that are relevant to a user's request from a database. To solve this problem, text-based and content-based are the two techniques adapted for search and retrieval in an image database (Salamah, 2010).

### *1.1 Text-based image retrieval*

In text-based image retrieval, images are indexed using keywords, subject headings or classification codes known as the metadata of an image (Gerard and Buckely, 1988). Although, text-based approach can offer much flexibility in query formulation, image retrieval based only on text information is not sufficient, since it cannot capture visual content; such as colour, texture, or shape, etc. Furthermore, the amount of labour required to manually annotate every single image, in addition to the difference in human perception when describing images, may lead to inaccuracies during the retrieval process. Hence, there is an urge the demand for efficient and effective image retrieving technique. Thus, the focus of researchers has been shifted to content-based image retrieval (CBIR) (Madugunki et al., 2011; Smeulders et al., 2000; Rui et al., 1998).

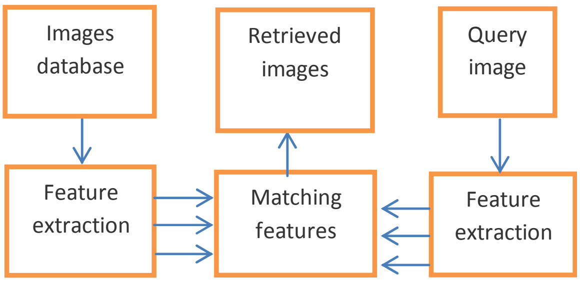
### *1.2 Content-based image retrieval*

CBIR is a well-known technology used for retrieval of images from large databases. Content-based means that the search construes the contents of the image rather than the metadata such as keywords, tags or descriptions associated with the image. The term ‘content’ in this context might point to colours, shapes, textures or any other information that can be derived from the image itself.

To avoid manual annotation, an alternative approach is CBIR, by which images would be indexed by their visual content such as colour, texture, shape, etc. The desired images are retrieved from a large collection on the basis of features that can be automatically extracted from the images themselves (Smeulders et al., 2000; Rui et al., 1998). These features are extracted and stored in a feature database. Similarly, the low level features are extracted from the query image and the query image features are compared with the database image features using a distance measure (Rui et al., 1998; El-Mashad and Shoukry, 2014; El-Mashad and Shoukry, 2015).

Images having the least distance to the query image are displayed as the result. The block diagram of the basic CBIR system is shown in Figure 1 (Singh and Hemachandran, 2012).

**Figure 1** The block diagram of basic CBIR (see online version for colours)




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 organised as follows. In Section 2, some issues about CBIR and related works are presented. In Section 3, describe the architecture of the proposed system. The experimental results are carried out on two databases and the comparisons with existing methods are presented in Section 4. Finally, we provide the conclusions and future work in Section 5.

## 2 Related works

CBIR method is introduced by Tian et al. (2014). Firstly, edge orientation difference histogram (EODH) and colour-sift are extracted in each channel, respectively then, the weighted distribution coding is performed for each channel. Finally, the coding result in each channel is normalised and the normalised code in each channel is integrated as image representation. The similarity measure between the query image and database images is obtained by using Euclidean distances. Experiments showed that the average precision for this system was 72.7%.

Fadaei and Sortrakul (2014) developed a technique for CBIR system based on colour and gradient features. In colour feature, colour histogram and colour moments are used while in gradient feature, edge histogram feature was adopted based on local binary pattern. The principle component analysis (PCA) has been used for decreasing the dimension of features. After performing PCA upon each feature, the weight was assigned to the colour feature and gradient feature and the similarity was also calculated based on normalising Canberra distance. Wang dataset was used to test this system and the average precision was 74.71%.

A CBIR system based on fusion of colour, texture and shape features has been proposed by Sakthivel et al. (2014). In colour feature, colour histogram and block histogram are used. The two colour-based image retrieval approach is used for results which showed that the block colour histogram is better than the global colour histogram. So, block colour histogram alone used in the fused CBIR system. Texture feature through her transform is better than the texture through the co-occurrence matrix. So, Haar transform alone used in the fused CBIR system. Shape feature extracted using fuzzy c-means algorithm. Finally, the best method of each feature is fused by linear weighted mode and the similarity measure is calculated by Euclidean distance.

A CBIR system that merges the retrieval results of two short-term learning (STL) algorithms using the Borda count fusion method to improve the accuracy of the system is presented by Bagheri et al. (2013). To do this properly, the types of outputs of the STL methods are assumed to be in the form of ranked lists. Borda count is used as the combination method that maps a set of individual rankings to a combined ranking. Then, the system presented 25 top-ranked images in the merged result for labelling by the user.

A CBIR system was implemented in which each session consists of four rounds of relevance feedback and Corel data set with 10,000 colour images from 82 different semantic groups are used for evaluation. The experimental results on 100 test images revealed that the combination method gave precision of 70%.

An image retrieval method based on colour moments and Gabor texture features was proposed by Singh and Hemachandran (2012). To improve the discriminating power of colour indexing techniques, a minimal amount of spatial information was encoded in the index by extracting features from the regions of the image divided horizontally into three equal non-overlapping regions. From each region in the image, the first three moments of the colour distribution were extracted and store the 27 floating point numbers (or three regions, each region being represented by a vector of nine floating point numbers) of the image in the index. As its texture feature, Gabor texture descriptors are adopted. Then, the weights were assigned to each feature, respectively and calculate the similarity with combined features of colour and texture using Canberra distance as a similarity measure. The average precision of this method was calculated which gave 61%.

Min et al. (2015) proposed an image retrieval method using multi-feature fusion. In this method, the colour moment in RGB colour space in combination with the colour histogram in HSV colour space is used for colour feature extraction. For shape feature extraction, Zernike moments are utilised. In addition, the gray level co-occurrence matrix is used for texture feature extraction. Then, all these three features are fused. The experimental data are 400 selected images from the Corel image database. The experimental results showed that the image retrieval method based on multi-feature fusion has better retrieval performance. However, the difference between the object and the background is less obvious, the retrieval accuracy decreases and the retrieval time increases. Consequently, this method needs to be improved.

Sorted block truncation coding (SBTC) method of feature extraction is proposed by Thepade et al. (2015). Proposed SBTC has implemented to near weighted mode and the similarity measure is calculated by Euclidean distance. Extract feature vectors from each of the colour component red (R), green (G) and blue (B), where each of this colour 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 neighbour (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.

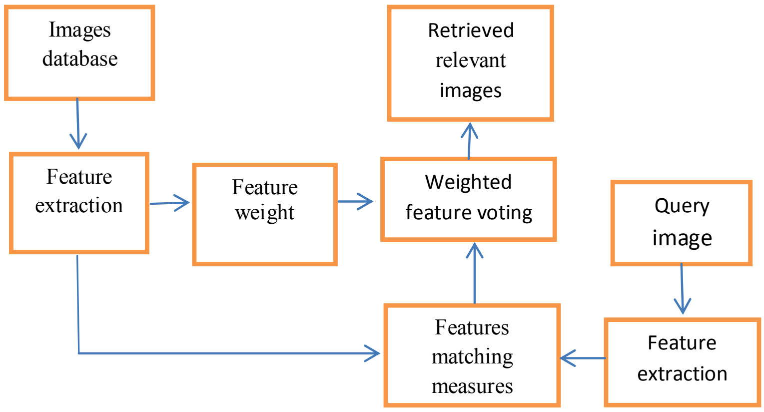
Anandh et al. (2016) proposed a technique for the generation of image content descriptor with three features colour auto-correlogram, Gabor wavelet and wavelet transform. colour auto-correlogram feature is associated with colour information of an image which is derived from the RGB colour space of an 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 analysed by the performance measures precision and recall. The efficiency of the feature descriptor is tested for CBIR system using Wang database, Li database and Caltech101 database. The method achieved an average accuracy rate of 83% for Corel database, whereas 88% for Li database and 70% for Caltech101 database in CBIR system.

## 3 Proposed system

In weighted 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, we have two features f1 and f2 with weights 0.8 and 0.7, respectively. The votes of f1 and f2 are five and four for class c respectively, then the overall votes for this class will be (0.8 * 5) + (0.7 * 4).

The CBIR proposed system is depicted in Figure 2. 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, which are 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 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 2** Proposed CBIR system (see online version for colours)




## *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 (Shoukry and El-Mashad, 2014). In this system, features are extracted using colour histogram (Manjunath et al., 2001), edge histogram (Manjunath et al., 2001), edge direction histogram (Jain and Vailaya, 1996), hierarchical annular histogram (HAH) (Yang et al., 2013), colour moments (Long et al., 2003) and colour coherence vector (CCV) (Pass and Zabith, 1996).

Colour is one of the most widely used features. Colour histogram is the most popular technique, which represents the global distribution of colours in an image. However, in large database, it does not have good performance because it stores only colour information and lacks spatial information, so images with very different appearances can have similar histograms (Fadaei and Sortrakul, 2014). To enhance the performance, HAH (Yang et al., 2013) is used which is rotation invariant and can capture the spatial configuration of pixel intensities throughout the image. In HAH, 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 the HAH. Because the 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 et al., 2013). This technique was originally developed for medical applications (Yang et al., 2013; Qi et al., 2013). HAH gave an interesting relatively high precision in the proposed system in this paper.

On the other hand, colour moments have the lowest feature vector dimension and lower computational complexity (Huang et al., 2010). Therefore, it is more suitable for image retrieval. In this system, the first-order and second-order colour moments are considered (Long et al., 2003).

Moreover, we implemented CCV (Pass and Zabith, 1996) that takes into account some of the spatial information between the pixels within the same colour coherence region. CCV is an improved conventional histogram. The pixel points belonging to colour bucket of the same colour histogram are divided into coherence and incoherence, which refers to pixels of the same colour distributed over the same area. Therefore, both CCV and colour moments not only takes into account the number of pixel points in different colours, but also the relative position information of the pixels.

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 et al., 2002). It represents the local edge distribution in the image which is obtained by subdividing the whole image into 4 × 4 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 16 × 5 for each image (Manjunath et al., 2001; Nandagopalan et al., 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.

he 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 equation (1) (Bala and Sharma, 2014):

$$\text{precision} = \frac{\text{No. of relevant images retrieved}}{\text{Total no. of images retrieved}}. \tag{1}$$

### *3.2 Matching measures*

Different features of an image are represented by vectors. Feature vectors of images in the database are calculated offline. A user can give an image as a query to retrieve similar images from the database. The feature vector of the query image is computed. Similarity measurement is an important key of a CBIR system in which the query image is compared with other database images. The similarity between the query image and other database images is 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 (Szabolcs, 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 equation (2):

$$\mathrm{w} = \begin{bmatrix} \mathrm{w}_1 & \mathrm{w}_2 & \mathrm{w}_3 & \mathrm{w}_4 & \mathrm{w}_5 & \mathrm{w}_6 \end{bmatrix} \tag{2}$$

where $\mathrm{w}_\mathrm{i}$ is the weight of the $\mathrm{i}^{\mathrm{th}}$ feature, $\mathrm{w}_\mathrm{i}$ is calculated as in equation (3):

$$\mathrm{w}_\mathrm{i} = \frac{1}{\mathrm{N}} \sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{p}_{\mathrm{ij}} \tag{3}$$

where N is the 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 equation (4) (Zakariya and Akhtar, 2014):

$$p_{ij} = \frac{\text{no. of relevant images at } k = 10}{k = 10} \tag{4}$$

where, $k$ represents the top ranked images.

#### 3.3.1 Weighted 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 utilised 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 equation (5):

$$\boldsymbol{V} = \begin{bmatrix} c_{11} & \dots & c_{1C} \\ & . & \\ & . & \\ c_{F1} & \dots & c_{FC} \end{bmatrix} \tag{5}$$

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, we can consider $c_{ij}$ 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 vectors as shown in equation (6):

$$\boldsymbol{s} = \boldsymbol{w}.\boldsymbol{V} = [s_1 \ldots s_C] \tag{6}$$

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 by taking into consideration all the features as per their importance revealed by weights. The class of the query is determined as in equation (7):

$$Class = \arg\max_i s_i, i = 1, \ldots, C \tag{7}$$

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 equation (8):

$$Feature = \arg\max_j c_{j\,Class}, j = 1, \ldots, F \tag{8}$$

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

## 4 Experimental results

In this part, the focus is on the effects of the proposed technique on the retrieval process in the term of accuracy to illustrate the robustness of the proposed method. The proposed image retrieval system is tested and evaluated on three images datasets. The first dataset is Wang database (Wang et al., 2001), it contains 1,000 images categorised into ten classes which are African people, beach, building, bus, dinosaur, elephant, flower, horse, mountain and food as shown in Figure 3. Each class includes 100 images. For each group, 80 images are utilised 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 Caltech101 database (Li et al., 2007). Caltech101 contains 101 categories which have huge variances in shape, colour and texture. Each object category contains about 31 to 800 images. The size of each image is roughly 300 × 200 pixels. The categories such as airplane, bonsai, panther, dalmatian, dolphin, faces, flamingo, deer, piano, skates, metronome, minar, motorbike, panda, football, stopsign, sunflower, trees, monument, watches, etc. as shown in Figure 4. To examine this system, 20 query images from each group are selected for training while ten sample images are selected for testing. Thus, there is a set of 2,020 images are reserved for training while 1,010 images are earmarked 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 flowers’, ‘Barcelona’ and ‘Iran’. Some example images with annotations are shown in Figure 5. Thirty query 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 3** Example images from the WANG database (see online version for colours)

**Figure 4** Example images from the Caltech101 database (see online version for colours)

**Figure 5** Example images from the UW database (see online version for colours)

Table 1 represents the average accuracy for each individual feature. It is evident that, there are some features that give the highest precision when used certain matching measure. For example, colour histogram gives the highest result with histogram intersection distance measure. But colour histogram alone not enough to build robust retrieval system because it is based on colour only, so the images that contain objects with the same colour, the system cannot distinguish between them. For example, the images contain sky and clouds are not distinguishable from the images contain sea.

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

| *Matching measures* | *Colour histogram* | *Edge direction* | *Edge histogram* | *Hierarchical annular histogram* | *Colour moments* | *Colour coherent vector* |
|---|---|---|---|---|---|---|
| Histogram intersection | *0.7632* | 0.2572 | 0.1774 | *0.6647* | 0.2378 | 0.2077 |
| Cosine distance | 0.6855 | 0.3710 | 0.3909 | 0.6557 | *0.5973* | 0.55290 |
| Euclidean distance | 0.6610 | *0.4670* | *0.4869* | 0.6145 | 0.58709 | *0.5861* |

The proposed multi-feature system utilises the features through the weighted feature voting mechanism to identify the class of the query image. The results are compared to the weighted average technique as shown in Table 2, which is used to fuse distances between the database images and the query image as shown in equation (9).

$$\boldsymbol{d}_{av} = \frac{\boldsymbol{w} \times \boldsymbol{M}}{\sum_{i=1}^{6} w_i} \tag{9}$$

where $M$ is distance matrix

$$\boldsymbol{M} = \begin{bmatrix} d_{11} & \dots & d_{1N} \\ & . & \\ & . & \\ d_{61} & \dots & d_{6N} \end{bmatrix}_{\boldsymbol{6 \times N}} \tag{10}$$

$N$ 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 feature voting technique and weighted average technique by using different matching distance measures

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

As shown in Table 2, the results achieved by weighted feature voting technique are better than the weighted average technique. A further enhancement is obtained by using each feature with the most suitable matching measure according to Table 1. For example, colour histogram is used with histogram intersection distance, colour moments used with cosine distance and etc.

**Table 3** Comparison between the average accuracy for each class in the Wang database using weighted feature voting technique and the existing methods

| *Classes* | *Description* | *Singh and Hemachandran (2012)* | *Fadaei and Sortrakul (2014)* | *Tian et al. (2014)* | *Thepade et al. (2015)* | *Anandh et al. (2016)* | *Das et al. (2017)* | *Proposed method* |
|---|---|---|---|---|---|---|---|---|
| 1 | African people | 74% | 72.10% | 74.6% | 80% | 80% | 80% | 100% |
| 2 | Beach | 38% | 61.20% | 37.8% | 80% | 84% | 60% | 60% |
| 3 | Building | 36% | 52.25% | 53.9% | 80% | 82% | 40% | 75% |
| 4 | Bus | 77% | 91.30% | 96.7% | 80% | 90% | 80% | 100% |
| 5 | Dinosaur | 95% | 96.40% | 99% | 100% | 100% | 100% | 100% |
| 6 | Elephant | 44% | 66.40% | 65.9% | 80% | 72% | 60% | 80.5% |
| 7 | Flower | 69% | 87.05% | 91.2% | 80% | 98% | 100% | 100% |
| 8 | Horse | 67% | 94.85% | 86.9% | 100% | 88% | 100% | 100% |
| 9 | Mountain | 69% | 52.35% | 58.9% | 40% | 66% | 60% | 60% |
| 10 | Food | 41% | 73.20% | 62.2% | 60% | 70% | 60% | 85% |
| Average accuracy | | 61% | 74.71% | 72.7% | 78% | 83% | 74% | *86.5%* |

Wang, UW and Caltech101 databases are used to test and evaluate our system as mentioned previously. The results of experiments, performs on the Wang database are listed in Table 3 and compared to the existing methods in Singh and Hemachandran (2012), Tian et al. (2014), Fadaei and Sortrakul (2014), Thepade et al. (2015), Das et al. (2017) and Anandh et al. (2016). It is evident that the proposed system has better accuracy than the existing methods for individual classes for Wang database and the overall precision, which is up to 86.5%.

**Table 4** Comparison between the average accuracy for Caltech101 database databases using weighted feature voting technique and the existing methods

| Method | *Krishnamoorthi and Sathiya Devi (2013)* | *Zhu et al. (2014)* | *Thepade et al. (2015)* | *Anandh et al. (2016)* | *Neelima et al. (2016)* | *Proposed system* |
|---|---|---|---|---|---|---|
| Average accuracy | 72.3% | 48.6% | 68.2% | 70% | 72% | *85%* |

Table 4 illustrates the result of average accuracy between the proposed system using Caltech101 and other existing methods. The results reflect that, the proposed system gives a precision as high as 85%. On the other hand, Table 5 represents 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 where the precision is 86.5%. Figure 6 represents sample of a query image of Wang dataset, while the retrieved images using the weighted feature voting technique of for the query image represented in Figure 6 are illustrated in Figure 7.

**Table 5** Comparison between the average accuracy for UW database using proposed system and other existing methods

| *Method* | *Mahajan and Patil (2014)* | *Yang and Cai (2014)* | *Proposed system* |
|---|---|---|---|
| Average accuracy | 80% | 67.46% | *86.5%* |

**Figure 6** Query image from Wang database (see online version for colours)

Although, the weighted feature voting mechanism considers all the features in the decision making process, in some cases the decision is falsely influenced by the high weighted feature. For example, this case is happened in class 2 and class 9 in Wang database, where the system is confused between these two classes because of the colour similarity and hence the high weighted colour features take a wrong decision. Unfortunately, a highly weighted feature sometimes takes over the decision; therefore, the final result may be wrong because of powerful features.

**Figure 7** Sample of retrieved images using weighted feature voting technique (see online version for colours)

## 5 Conclusions

In this paper, a novel CBIR system is proposed using a weighted feature voting technique with multi-features and multi distance measures. The implemented features are colour histogram, edge histogram, edge direction histogram, HAH, colour moments and CCV. While the distance measures used are histogram intersection, Euclidean distance and cosine distance. The results indicate that the suitable distance measure for features colour histogram and HAH is histogram intersection, whereas the appropriate distance measure for edge direction, edge histogram and CCV is the Euclidean distance. On the other hand, cosine distance is the best distance measure for CCV. The proposed system provides a considerable enhancement in the overall performance compared to the weighted average technique and existing methods where the achieved accuracy reaches up to 86.5% for Wang database, 86.5% for UW and 85% for Caltech101.

As a future work, a shape-based feature should be incorporated 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', *International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16), IEEE explorer*, pp.1–8.

Bagheri, B., Pourmahyabadi, M. and Nezamabadi-pour, H. (2013) 'Novel content based image retrieval approach by fusion of short-term learning methods', *5th Conference on Information and Knowledge Technology (IKT) IEEE*, pp.355–358.

Bala, R. and Sharma, R. (2014) 'Efficient method of CBIR using combination of shape, color and texture feature with fuzzy technique', *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, Vol. 3, No. 8, pp.909-913.

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*, Vol. 7, Nos.1–2, pp.123–147.

El-Mashad, S.Y. and Shoukry, A. (2014) 'A more robust feature correspondence for more accurate image recognition', in *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', in *International Conference on Pattern Recognition and Machine Intelligence*, Springer International Publishing, pp.33–43.

Fadaei, H. and Sortrakul, T. (2014) 'Content-based image retrieval system with combining color features and gradient feature', *Int'l. Conference on Advanced Computational Technologies & Creative Media (ICACTCM'2014)*, pp.15–20.

Gerard, S. and Buckely, C. (1988) 'Term-weighting approaches in automatic text retrieval', *Information Processing and Management*, Vol. 24, No. 5, pp.513–523.

Huang, Z., Chan, P., Ng, W. and Yeung, D. (2010) 'Content based image retrieval using color moment and Gabor texture feature', *The Ninth International Conference on Machine Learning and Cybernetics IEEE*, pp.719–724.

Jain, A. and Vailaya, A. (1996) 'Image retrieval using color and shape', *Pattern Recognition*, Vol. 29, No. 8, pp.1233–1244.

Kaur, S. and Aggarwal, D. (2013) 'Image content based retrieval system using cosine similarity for skin disease images', *ACSIJ Advances in Computer Science: an International Journal*, Vol. 2, No. 5, pp.89–95.

Krishnamoorthi, R. and Sathiya Devi, S. (2013) 'A simple computational model for image retrieval with weighted multifeatures based on orthogonal polynomials and genetic algorithm', *Seventh International Conference on Intelligent Computing*, pp.165–181.

Li, F-F., Fergus, R. and Perona, P. (2007) 'Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories', *Comput. Vis. Image Underst.*, Vol. 106, No. 1, pp.59–70.

Long, F., Zhang, H. and Feng, D. (2003) 'Fundamentals of content based images retrieval', *Multimedia Information Retrieval and Management*, pp.1–26, Springer Berlin Heidelberg.

Madugunki, M., Bormane, D., Bhadoria, S. and Dethe, C. (2011) 'Comparison of different CBIR techniques', *3rd Conference on Electronics Computer Technology (ICECT) IEEE*, Vol. 4, pp.372–375.

Mahajan, S. and Patil, D. (2014) 'Image retrieval using contribution-based clustering algorithm with different feature extraction techniques', *Conference on IT in Business, Industry and Government (CSIBIG), IEEE*, pp.1–7.

Manjunath, B., Ohm, J., Vasudevan, V. and Yamada, A. (2001) 'Color and texture descriptors', *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, No. 6, pp.703–715.

Min, H., Huazhong, S., Yaqiong, M. and Qiuping, G. (2015) 'Content-based image retrieval technology using multi- feature fusion', *Optik-International Journal for Light and Electron Optics*, Vol. 126, No. 19, pp.2144–2148.

Nandagopalan, S., Adiga, B. and Deepak, N. (2008) 'A universal model for content-based image retrieval', *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, Vol. 2, No. 10, pp.3436–3439.

Neelima, N., Sreenivasa, E. and Kalpitha, N. (2016) 'An integrated approach to CBIR using multiple features and HSV histogram', *International Journal of Engineering and Technology (IJET)*,Vol. 8, No. 5, pp.1946–1953.

Pass, G. and Zabith, R. (1996) 'Histogram refinement for content-based image retrieval', *IEEE Workshop on Applications of Computer Vision*, pp.96–102.

Qi, X., Gensure, R., Foran, D. and Yang, L. (2013) 'Content-based white blood cell retrieval on bright-field pathology images', *Medical Imaging 2013: Digital Pathology, Proc. of SPIE*, pp.1–6.

Rui, Y., Huang, T., Ortega, M. and Mehrotra, S. (1998) 'Relevance feedback: a power tool for interactive content-based image retrieval', *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 8, No. 5, pp.644–655.

Sakthivel, K., Nallusamy, R. and Kavitha, C. (2014) 'Image retrieval using fused features', *International Journal of Computer, Information, Systems and Control Engineering*, Vol. 8, No. 9, pp.1463–1467.

Salamah, R. (2010) *Efficient Content Based Image Retrieval*, Master thesis, Islamic University, Gaza.

Sawant, A., Bharadi, V. and Markarkandy, B. (2012) 'Color &texture based image retrieval using fusion of modified block truncation coding (MBTC) and kekre transform patterns', *International Conference & Workshop On Emerging Trends In Technology*, pp.12–17.

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.

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

Singh, S. and Hemachandran, K. (2012) 'Content-based image retrieval using color moment and Gabor texture feature', *IJCSI International Journal of Computer Science Issues*, Vol. 9, No. 5, pp.299–309.

Smeulders, A., Worring, M., Santini, S., Gupta, A. and Jain, R. (2000) 'Content-based image retrieval at the end of the early years', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 2, pp.1349–1380.

Smith, J. (1997) *Integrated Spatial and Feature Image Systems: Retrieval, Analysis and Compression*, PhD thesis, Columbia Univ.

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

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*, Vol. 9030, pp.55–76, Springer Berlin Heidelberg.

Tian, X. Jiao, L., Liu, X. and Zhang, X. (2014) 'Feature integration of EODH and color-SIFT: application to image retrieval based on codebook', *Signal Processing: Image Communication*, Vol. 29, No. 4, pp.530–545.

Wang, J.Z., Li, J. and Wiederhold, G. (2001) 'SIMPLIcity: semantics-sensitive integrated matching for picture libraries', *The IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 9, pp.947–963 .

Won, C., Park, D. and Park, S. (2002) 'Efficient use of MPEG-7 edge histogram descriptor', *ETRI Journal*, Vol. 24, No. 1, pp.23–30.

Yang, L., Qi, X., Xing, F., Kurc, T., Saltz, J. and Foran, D. (2013) 'Parallel content based sub-image retrieval using hierarchical searching', *Bioinformatics Advance*, pp.1–7 [online] http://bioinformatics.oxfordjournals.org (accessed 9 November 2013).

Yang, X. and Cai, L. (2014) 'Adaptive region matching for region-based image retrieval by constructing region importance index', *IET Computer Vision*, Vol. 8, No. 2, pp.141–151.

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

Zhu, L., Jin, H., Zheng, R. and Feng, X. (2014) 'Weighting scheme for image retrieval based on bag-of-visual-words', *IET Image Processing*, Vol. 8, No. 9, pp.509–518.