Journal of Engineering and Sustainable Development
http://jeasd.uomustansiriyah.edu.iq
https://doi.org/10.31272/jeasd.24.5.6

Vol. 24, No. 05, September 2020
ISSN 2520-0917

SPECTRAL ALGORITHM FOR CONTENT-BASED IMAGE RETRIEVAL

*Hanan A. Al-Jubouri[1] Sawsan M. Mahmmod[1]

1) Lecturer, Computer Engineering Department, Mustansiriyah University, Baghdad, Iraq

Received 09/12/2019 *Accepted in revised form 04/02/2020* *Published online 01/09/2020*

Abstract: Colour images are rich in visual information. The process of searching for the most similar images in large-scale database based on visual features of query image is still a challenge in Content-Based Image Retrieval (CBIR) due to a semantic gap issue. In this paper, we proposed a fusing retrieval method to diminish the gap between high-level and low-level meanings by involving two aspects. The first aspect is increasing the effectiveness of image representation. Hence, data-level fusion features were suggested, a local feature from Discrete Cosine Transform (DCT) and Local Binary Patterns (LBP) in frequency and spatial domains respectively that was applied by a spectral clustering algorithm (graph-based) in addition to a global weighted LBP feature. The second aspect is fusing multiple retrieved similarity measures (scores/evidences) obtained from above global (LBP) and local features (DCTLBP) in terms of score-level fusion. The method is evaluated in WANG standard publically dataset.

Keywords: *Content-Based Image Retrieval (CBIR), Local Binary Pattern (LBP), Discrete Cosine Transform (DCT), data-level fusion, and score-level fusion*

1. Introduction

Extensive use of digital photographic devices has resulted in large volumes of digital images being acquired and stored in databases. Whether it is for scientific research, forensic analysis or social networking, there is a growing demand for effective retrieval of digital images based on their visual content. Content-Based Image Retrieval (CBIR) systems are developed to meet this demand. In its very essence, a CBIR system indexes an image in the database by extracting a feature vector that reflects the visual content of the image. Upon request, the system extracts a feature vector from a query image in the same way and compares it with the feature vectors of the images in the database using a similarity measure. The most similar images are ranked and returned to the user (Fig. 1). The main challenge for CBIR systems is how to bridge the *semantic gap* between human conceptual meaning for images and machines such as a computer. In other words, how the CBIR system can extract effective features that represent the image in the database and retrieve in terms of relevant images.

The aim of this paper is to develop a retrieval method that addresses the semantic gap problem in CBIR by firstly increasing a discrimination of image representation and secondly aggregating multiple retrieved evidences from different features to enhance the performance image retrieval. Hence, local features from DCT and LBP were concatenated to produce a new robust colour-texture feature from frequency and

**Corresponding Author*: hanansalman@yahoo.com

spatial domains respectively to represent images. A spectral clustering algorithm was applied to quantize the number of produced features. Also global LBP features from *Y*, *Cb*, and *Cr* channels were concatenated to increase the discriminative of image representation. Then evidences obtained from retrieval process using mentioned features were fused to increase the performance of image retrieval. Section 2 will state literature review. Section 3 will explain the methodology of DCT, LBP, clustering algorithms, and the framework. Section 4 will illustrate an experiment setup. Section 5 will presnt results and discussion. Finally, conclusions will be in Section 6.

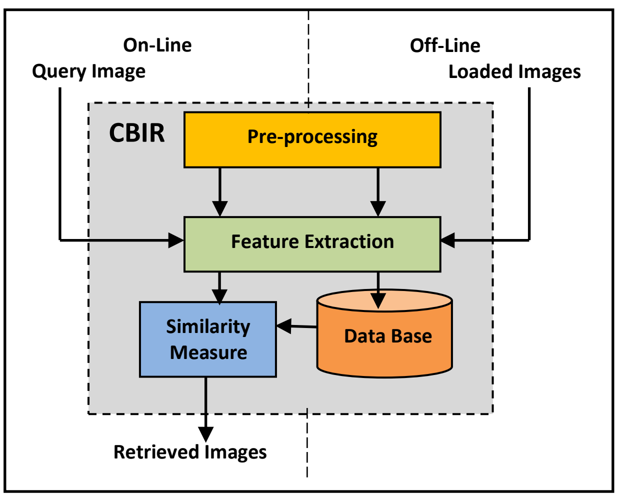


Figure 1. CBIR system

2. Literature Review

Over the last two decades, many approaches and methods in CBIR have been proposed [1, 2]. Image classification methods are supervised learning with external labelling and annotations such as Support Vector Machine (SVM) and K-Nearest Neighbour (K-nn). Image retrieval methods are unsupervised learning without using labels such as clustering, region of interest (ROI) [3-6], Bag-Of-Visual-Words (BOVW) [41, 42]], relevance-feedback (RF) [7-11], and Browsing [44]. Feature extraction and similarity measures are the core processes of mentioned approaches. We interested clustering and extracted features in this paper.

The idea of the clustering approach is developing an algorithm that groups feature vectors into clusters based on similarity measure. The effectiveness of the process is mainly depends on data and cluster characteristics, details will be in Sec. 3.3. Meanwhile, the base of the ROI approach is defining an interested area from where features are extracted to index images. However, one major limitation is specifying the ROI during a retrieval session by the user. The principle of the FR approach is refining the retrieved images by the user to determine relevant and irrelevant images, which may cause burden to the user. The concept of the BOVW approach is dividing the image into patches from where visual features are extracted and then quantized by using a clustering algorithm. Resulted clusters correspond to vocabularies and their centroids to words. The number of generated clusters is very large such that 500 up to 10000 which is an issue in terms of retrieval efficiency. Whilst the most mentioned approaches use an example image as query to the CBIR system, the browsing approach displays images to navigate and choose the ones of the interest. Hence, the difficulty is visualizing the image collections on small screen size in addition to developing an effective and efficient tool to navigate through many database images. Recently, a deep learning technique attracts researchers of CBIR to reduce the semantic gap issue [42, 43]. Convolution neural networks are used to learn feature representations to index images in the database. Different algorithms are tested on several large image databases such as Caltech256 and ImageNet. Results have satisfied better levels of retrieval accuracy.

3. Methodology

3.1 Discrete Cosine Transform (DCT)

DCT is one type of transformation methods and can be computed by the following operation that is executed iteratively on the pixel intensity values of an 8 x 8 block window on the image:

$$C(u,v) = \frac{1}{4}k(u)k(v)\sum_{i=0}^{7}\sum_{j=0}^{7} f(i,j)\cos\left(\frac{(2i+1)u\pi}{16}\right)\cos\left(\frac{(2j+1)v\pi}{16}\right) \quad (1)$$

$$k(u),k(v) = \begin{cases} 1/\sqrt{2} & if\ u\ and\ v = 0 \\ 1 & otherwise \end{cases}$$

where 0 ≤ u, v ≤ 7 and *f*(i, j) is the pixel intensity value at location i, j. C(0, 0) represents a low frequency DC and the remaining represents high frequency (ACs). The DC coefficient tends to capture the colour of the block and the AC coefficients the textures of the block. For colour images, DCT can be applied to each colour channel.

Consequently, DCT was applied to extract image features in the frequency domain from the *YCbCr* colour space [22, 23, 24]. In addition, the DCT features are investigated using *RGB*, *YCgCb*, *YUV*, *YIQ*, *XYZ*, *LUV*, and *YCbCr* colour spaces, the performance of the features in *YCbCr* is better than others [25].

Researchers exploited DCT coefficients in different order such as a zigzag manner, where the coefficients are arranged from low to high frequencies for 8x8 blocks. Consequently, a resulted feature is high in dimension and possibility of making the vector susceptible for *over-fitting* means the feature vector has too much specific details of the local block. Therefore, the DCT-zigzag feature vector was reduced to the first 10 most significant DCT coefficients. Meanwhile, the DCT coefficients in a (8x8) block are ordered like multi-resolution decomposition discrete wavelet transform coefficients in three levels sub-bands. Means and standard deviations are calculated accordingly to create the local feature vector that showed a good performance [26].

The DCT colour texture (F_{DCT}) feature vector [27] takes the standard deviations of the coefficients in C_4, C_5, C_6, C_7, C_8 and C_9 sub-blocks, capturing the multi-resolution textural information (i.e. variations) in all high frequency bands, and at the same time maintaining the robustness of the feature vector with only 12 dimensions, i.e. ($C_Y(0,0)/8$, $C_{Cb}(0,0)/8$, $C_{Cr}(0,0)/8$, $C_Y(0,1)$, $C_Y(1,0)$,$C_Y(1,1)$, $std(C_Y4)$, $std(C_Y5)$, ..., $std(C_Y9)$). The performance of F_{DCT} was better than DCT coefficients in a traditional zigzag order and DWT itself [28]. Therefore, the same F_{DCT} feature was adopted in this paper (Fig. 2).

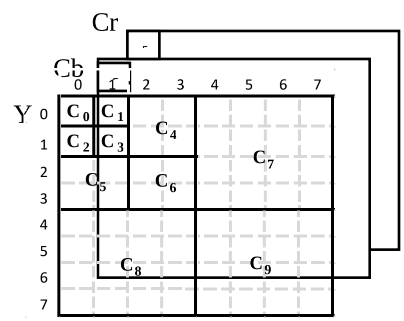


Figure 2. DCT-CT Feature from 8x8 block using *YCbCr*

3.2 Local Binary Pattern (LBP)

LBP is a measurement of a local texture using the relationships between a pixel and its neighbouring pixels. The process starts by subtracting the central pixel value from the neighbouring pixel values. Then a binary number is used to refer to each corresponding result. Consequently, these binary numbers create the local binary pattern that are multiplied by weights of locations and summed to obtain the new value of the central pixel according to the following formula:

$$LBP_{N,R}(P_c) = \sum_{n=0}^{N-1} s(P_n - P_c)2^n \quad (2)$$

$$s(P) = \begin{cases} 1 \; if \; P \geq 0 \\ 0 \; if \; P < 0 \end{cases}$$

where $N$and R are the number of neighbouring pixels and radius of neighbouring pixels. Fig. 3 illustrates above process for 3x3 block of the image.

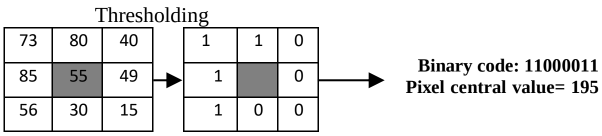


Figure 3. LBP for *N*=8 and *R*=1

Suppose Z is a X x Y image, the resulting LBP$_{N,R}$ code image can be represented by a histogram h of length S, where $0 \leq s \leq S-1$ and $S = 2^p$ is the number of all the LBP codes. For instance, if p=8 neighbours, then$S$=256. Feature h has good properties such as grey-scale invariance, low complexity, few parameters, and satisfactory discriminating power [12]. However, the h is a long histogram (2^p distinct values). Consequently, the LBP codes are reduced because not all of the local patterns are necessary to make texture analysis and suggested using just “uniform” patterns $LBP_{N,R}^{u2}$ [13]. The uniform patterns contain at most two bitwise transitions from 0 to 1 or vice versa when the binary string is considered as circular 11000011. Uniform patterns consist of useful texture features compared to non-uniform binary patterns. Therefore, all occurrences of non-uniform patterns are aggregated to a single bin of the histogram. As a result, the number of bins in h is reduced to 59-bins (i.e. 58 uniform patterns and 1 for non-uniform patterns).

The first application for Local Binary Patterns texture feature was in face recognition [14-16]. Basically, face images are divided into regions (3x3, 5x5, or 7x7) block to compute LBP codes. Histograms are then calculated for each region and concatenated to represent a descriptor (feature) of the face images. A Principle Component Analysis (PCA) and Linear Discriminate Analysis (LDA) can be used to reduce a dimensionality feature.

Researchers in CBIR field exploited the LBP texture feature to represent other types of images such as texture and natural scene images. For example, Different colour spaces (*RGB*, *HSV*, and *CIE* $L^*a^*b^*$) were used to apply LBP. $LBP_{N,R}^{u2}$ feature was extracted from greyscale images as well as colour images, where (N= 8, 16, and 24) circular neighbourhoods and (R = 1, 2, 3, and 5) radii were tested [17]. In addition, $LBP_{N,R}^{u2}$ feature was extracted from (128 x 128) blocks on one time and (96 x 96) blocks on the other time in two cases overlapping and non-overlapping [18]. Meanwhile, a completed local binary pattern (CLBP) was proposed to combine three features. The first refers to center gray level values (CLBP-C). The second refers to signs (CLBP-S) such that [-1, -1, 1, 1, 1, 1, -1, -1] vector where negative signs mean neighbourhood pixel values are less than the center pixel value. The third refers to magnitude values of neighbourhood pixels such that [-49, -40, 80, 73, 85, 56, -30, -15]. The complement of mentioned features improved representing images [19].

More recently, the idea of LBP was extended to regard differences between neighbourhoods therefore the feature was called a local neighbourhood difference pattern (LNDP). Both LNDP and LBP features were combined to support image retrieval because each feature reflects different information of local pixel intensity [20]. A combination between features was also applied in [21] to enhance image retrieval. Where LBPC and LBPH features were derived from *RGB* images and from hue component in the *HSI* respectively, in addition to colour histogram CH.

3.3 Clustering Algorithms

Basically, clustering is a process of grouping elements into homogeneous clusters according to their similarities. The desirable result is a high degree of intra-cluster similarity and a high degree of inter-cluster differences [29]. Different categories of clustering algorithms have been developed according to the meaning of the clusters produced [30], and have been used for CBIR. For instance, the *k*-means, Expectation Maximization/Gaussian Mixture Model, Mean Shift, and Normalized Laplacian Spectral are clustering algorithms from *prototype-based*, *model-based*, *density-based*, and *graph-based* categories respectively. In this paper, we implemented the Normalized Laplacian Spectral algorithm [31] that is shown in Fig 4.

***Step 1*:** Suppose a set points X={$x_1, x_2, \ldots, x_n$} in R^l then create the affinity matrix

$$A \in R^{n\times n} \text{ by } A_{ij} = exp\left(\frac{-D_{L_2}(s_i,s_j)}{2\sigma^2}\right) \text{ for}$$

$i \neq j$ and $A_{ii} = 0$

***Step 2*:** Calculate D to be a diagonal matrix where $D_{ii} = \sum_{j=1}^{n} A_{ij}$ and construct the matrix

$$L = D^{-1/2}AD^{-1/2}.$$

Step 3: Find k eigenvectors such that $V = [v_1, v_2, \ldots, v_k] \in R^{n\times k}$ with largest magnitude eigenvalues of the matrix L.

Step 4: Construct matrix Y by normalise each of V's rows to have unit length. $Y_{ij} = V_{ij}/\left(\sum_j V_{ij}^2\right)^{1/2}$

***Step 5*:** Regard each row of Y as a point in R^k and cluster using K-means.

***Step 6*:** Assign the original point x_i to cluster k if and only if the corresponding row i of the matrix Y was assigned to cluster k.

Where the affinity matrix A is calculated based on the Gaussian distance. Then a normalized Laplacian L matrix is constructed based on A and a degree diagonal matrix D. The spectrum of the matrix is determined.

Figure 4. Normalized Laplacian Spectral Algorithm

3.4 Framework

Fig. 5 shows stages of our framework that was followed to investigate colour texture features. Colour images were converted from *RGB* into *YCbCr* colour space at the first stage. An extracting process of global LBP, local LBP and DCT features was made at the second stage. The global and local features were represented as data-level fusion. The local DCTLBP features were clustered by Normalized Laplacian Spectral algorithm. City block distance function was used to match between features of query image and those of the database images at the third stage. Images were ranked in ascending order as retrieved list based on dissimilarity values at the final stage. Two types of fusion so-called data and score-level were applied to reduce the semantic gap challenge. Consequently, the accuracy of image retrieval will be increased.

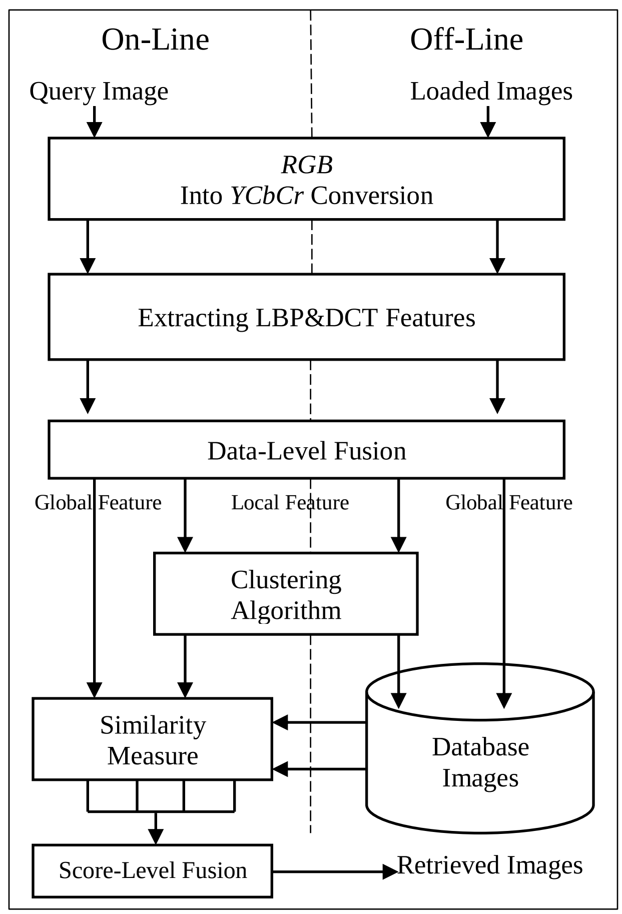


Figure 5. Frame work of Retrieval Method

4. Experiments Set up

This section clarifies a database that was used in conducted experiments and performance evaluation.

4.1 Database

WANG standard colour database images [32] which is part of Corel database and publically available for researchers. The standard database contains 1000 images in JPEG format that are divided into 10 categories (Elephants, Flowers, Buses, Foods, Horses, Mountains, African people, Beach, Buildings, and Dinosaurs). Each category has 100 images in (256x384) and (384 x 256) size. Fig. 6 shows sample of image categories.

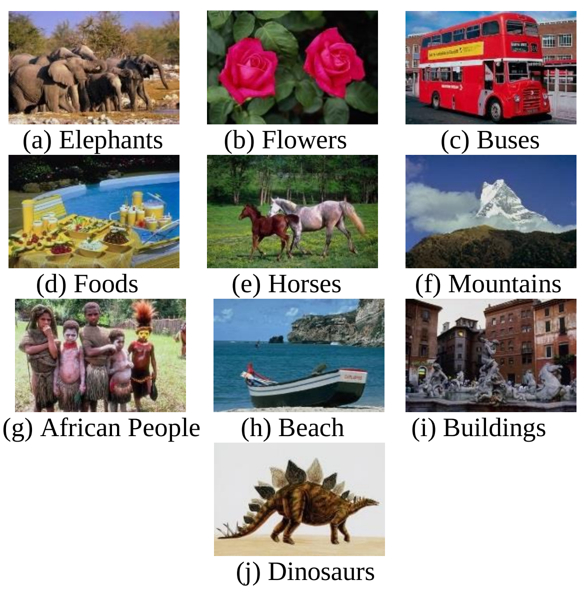


Figure 6. Sample of WANG database

4.2 Performance Evaluation

A precision is a commonly measure to evaluate a proposed method or system in terms CBIR, where relevant images are base for a computation.

$$P = \frac{NIM}{TRIM} \tag{3}$$

where, P is the precision of image retrieval, NIM is number of relevant retrieved images and $TRIM$ is total number of retrieved images.

$$AP = \frac{\sum_{i=1}^{n} P_i}{n} \tag{4}$$

where, AP is average precision of image retrieval, P_i is precision of i image in the class, and n is total number of images in the class.

$$MAP = \frac{\sum_{j=1}^{m} AP_j}{m} \tag{5}$$

where, MAP is mean average precision of image retrieval, AP_j is average precision of j class image, and m is total number of classes in the database.

5. Results and Discussions

Originally, data and score-level fusions were applied in information retrieval [34, 35]. Both fusion types were exploited in our experiments. The data-level fusion is concatenating features into one feature to increase the robustness. Hence, local DCT and LBP features from *YCbCr* colour spaces were integrated to be F_{DCTLBP} using Normalized Laplacian Spectral algorithm to represent images. In addition, a global LBP from *Y*, *Cb*, and *Cb* channels were integrated to be $F_{LBPYCbCr}$ and then weighted to be $F_{WLBPYcbCr}$. Meanwhile, the score-level fusion is combining scores/evidences from different runs and then making image retrieval. Therefore, produced outcomes from the image retrieval using global $F_{WLBPYcbCr}$ and local $F_{DCTLBPYcbCr}$ features were combined to improve the accuracy. All experiments were conducted using MatLab version 2013a and CBIR system was designed by GUI in the MatLab. Details will be demonstrated in this section.

5.4.1 Global Feature

In this experiment, images were converted from *RGB* to *YCbCr* colour space. Then Local Binary Patterns was calculated for *Y*, *Cb*, and *Cr* channels using *N*=8 neighbours and *R*=1 radius. The uniform histogram (59D) was computed to represent gray-scale images as F_{LBPG} feature. Results of MAPs along Top10-50 are illustrated in Table 1. Then the uniform histogram was extracted from *Y*, *Cb*, and *Cr* channels and concatenated to represent colour images by $F_{LBPYCbCr}$ (177D) in terms of data-level fusion. We can see how the discriminate between class images is increased. That means integrating colour texture visual information has a positive effect.

Moreover, we weighted $F_{LBPYCbCr}$ by 0.6, 0.3, and 0.1 weights for *Y*, *Cb*, and *Cr* respectively. The *Y* panel was weighted by highest value because it contains visual information like that human eye recognize. Consequently, obtained feature $F_{WLBPYCbCr}$ satisfied improvements about 4% as reported in Table 1.

Table 1. MAPs of image retrieval using LBP features with and without weights

Retrieved Images	F_{LBPG}	$F_{LBPYCbCr}$	$F_{WLBPYCbCr}$
Top10	0.59	0.69	0.72
Top20	0.55	0.64	0.66
Top30	0.52	0.60	0.63
Top40	0.49	0.56	0.60
Top50	0.47	0.53	0.57

Further, the *t*-test was used to evaluate the significance of the performance between F_{LBPG} and $F_{WLBPYCbCr}$ features. The *t*-test is commonly statistical method and can be computed as [33]:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{(N_1-1)S_1^2+(N_2-1)S_2^2}{N_1+N_2-2}\right)\left(\frac{1}{N_1}+\frac{1}{N_2}\right)}} \quad (6)$$

where $\bar{X}_1$ and $\bar{X}_2$ are the sample precision rates (P), S_1 and S_2 are standards deviations, and N_1 and N_2 are the sample sizes.

Two hypotheses are regarded and determined based on *t*-test, the null hypothesis (H_0) where $\bar{X}_1 - \bar{X}_2 = 0$ and alternative hypothesis (H_A) where $\bar{X}_1 - \bar{X}_2 \neq 0$. For each class in the WANG database the test was computed. This means the size of each sample is 100 elements (i.e. precision values). Hence, the first sample is precision rates of Top10 retrieved images from using $F_{WLBPYCbCr}$ feature and the second sample is from using F_{LBPG} feature (Table 2).

Table 2. MAPs of image retrieval using LBP features from gray and *YCbCr* images

Classes	$F_{WLBPYCbCr}$	F_{LBPG}
Elephants	0.6	0.34
Flowers	0.91	0.87
Buses	0.97	0.94
Foods	0.75	0.44
Horses	0.81	0.66
Mountains	0.42	0.3
People	0.67	0.55
Beach	0.49	0.46
Buildings	0.63	0.41
Dinosaurs	0.98	0.97
MAP	0.72	0.59

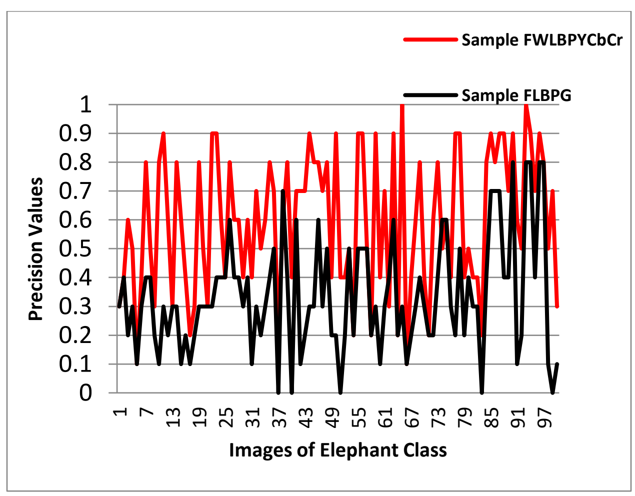


Figure 7. Precision values along Elephants class images

Fig. 7 illustrates the two samples for Elephants class. It is clear that rates of relevant images are increased using $F_{WLBPYCbCr}$ and the significant of increment is proved by *t*-test, where accepting the hypothesis H for 8 out of the total 10 classes (Fig. 8).

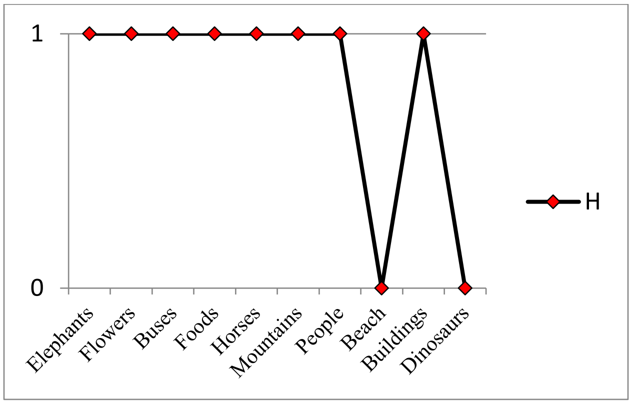


Figure 8. Hypotheses values on WANG classes

5.4.2 Local Feature

In this experiment, the same above image conversion was followed. Then the image was divided into 8x8 blocks and Discrete Cosine Transform was applied to extract local F_{DCT} features (12D). The Normalized Laplacian Spectral algorithm (NS) was used to quantize local features into *K* clusters. The number of clusters was adapted using EM/GMM algorithm that determines the value of *K* according to the Rissanen's Minimum Description Length (MDL) estimator [33].

In addition, local $F_{LBPYCbCr}$ features were extracted from 8x8 blocks after converting images into Local Binary Patterns. The same spectral algorithm was applied on resulted local features. Table 3 illustrates MAPs as long as Top 10 to 50 retrieved images. $F_{LBPYCbCr}$ produced rates higher than F_{DCT} about 2% because F_{DCT} is a frequency feature 12 dimensional in length, 9 coefficients from *Y* channel and just 2 coefficients from Cb an Cr based on coefficients at high frequency are very small than those at low frequency as clarified in Sec. 3. Meanwhile, $F_{LBPYCbCr}$ is a spatial feature used three uniform histograms (i.e. 177D) for *Y*, *Cb*, and *Cr* channels means the feature can capture more information about the image.

Hence, we integrated F_{DCT} and $F_{LBPYCbCr}$ features into one F_{DCTLBP} feature (189D) in terms of data-level fusion to represent images. Averages of image retrieval are raised about (7-8) %. This is evidence that integrated features increased image discrimination.

Table 3. MAPs of image retrieval using local F_{DCT}, $F_{LBPYCbCr}$, F_{DCTLBP} features

Retrieved Images	$F_{LBPYCbCr}$	F_{DCT}	F_{DCTLBP}
Top10	0.58	0.56	0.66
Top20	0.53	0.51	0.60
Top30	0.50	0.48	0.57
Top40	0.47	0.45	0.54
Top50	0.45	0.43	0.51

5.4.3 Global and Local Features

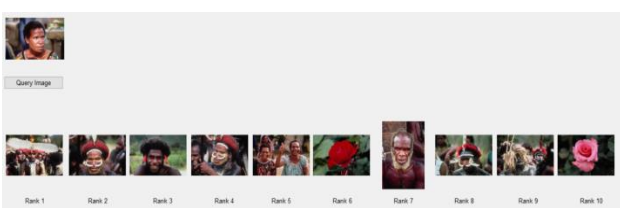
At the beginning, the fusion was applied in information retrieval and affected positively on results [34, 35]. Hence, researchers in CBIR field benefit from the idea. Therefore, we used the score-level fusion in this experiment aiming to restrict the semantic gap and raising the accuracy of image retrieval.

The global and local features were used to retrieved images after computing dissimilarity measures; values were normalized to make score-level fusion by giving weight 0.7 for weighted $F_{LBPYCbCr}$ and 0.3 for F_{DCTLBP}. We can see that MAPs increased about 5%.

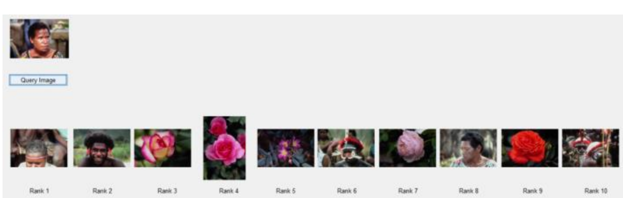
Taken some examples to clarify and explain the reason to use the fusion. Fig.9 (a) displays Top 10 retrieved images using local F_{DCTLBP} features which clustered using NS algorithm for African people image. We can see two images are irrelevant to the query image at rank 6 and 10 compared to those are obtained from using global Weighted $F_{LBPYCbCr}$ which brings five irrelevant images at rank 3, 4, 5, 7,and 9 (Fig.8 (b)). At the same time, the global feature retrieves three relevant images at rank 1, 8, and 10 which are different from images using local feature. Therefore, fusion results support to increase the accuracy of image retrieval. Another example for a building query image in Fig.10 (a-b) that shows the local feature brings two different relevant images from the global one. Hence, the fusion is useful to integrate results.

Table 4. MAPs of image retrieval using local F_{DCTLBP}, F_{LBPHu2} features, and Score-level fusion

Retrieved Images	F_{DCTLBP}	$WF_{LBPYCbCr}$	Score-LevelFusion
Top10	0.66	0.72	0.76
Top20	0.60	0.66	0.71
Top30	0.57	0.63	0.67
Top40	0.54	0.60	0.64
Top50	0.51	0.57	0.61

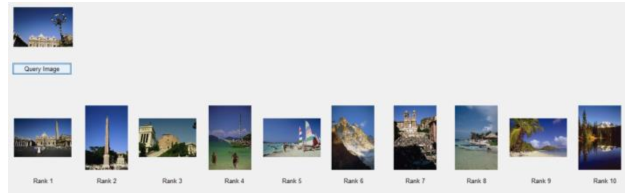


(a) Retrieval using local F_{DCTLBP} feature

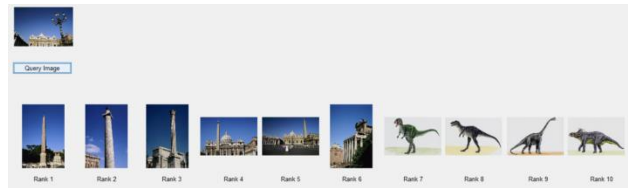


(b) Retrieval using global $F_{LBPYCbCr}$ feature

Figure 9. Top10 retrieved images for African people query image



(a) Retrieval using local F_{DCTLBP} feature



(b) Retrieval using global $F_{LBPYCbCr}$ feature

Figure 10. Top10 retrieved images for building query image

We also used *t*-test to judge the significant of differences between the performance before and after the fusing. *P*-value of the test is the probability of observing a test. Small values of *p* support the alternative hypothesis. Results of *t*-test for Elephants, Horses, Mountains, Beach, and Dinosaurs classes indicated respective significance levels $p = 0.000437$, $p = 0.039218$, $p = 0.001603$, $p = 0.00001$, and $p = 0.030122$ respectively. This means that the null hypothesis is rejected for mentioned classes as is shown in Fig. 11. Hence, the fusing between global and local features from spatial and frequency domains with appropriate weights helped to recognize more relevant images. In other words, resulted evidences from the local F_{DCTLBP} feature support those from the global $F_{LBPYCbCr}$ feature to make a decision about the image, if it is relevant or irrelevant.

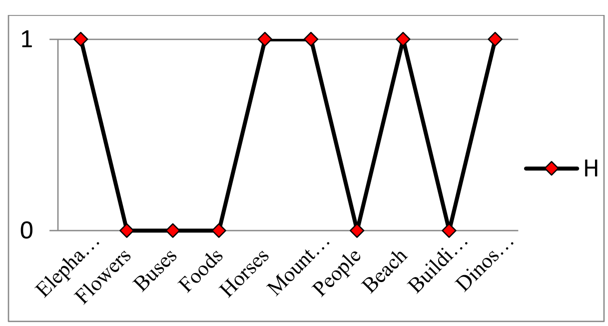


Figure 11. Hypotheses values on WANG classes

To sum up, image representation is one of core components in any Content-Based Image Retrieval system/approach and reflected by global and local features. The global feature represents the entire image by a single vector, and hence retrieval by using the global feature is more appropriate with obvious scenes. On the other hand, the local features represent the image by a set of vectors, therefore complicated images are better to represent. Consequently, each feature can capture different visual information about images. Hence, we suggested new colour texture global and local features from spatial and frequency domains which are $F_{WLBPYCbCr}$ and F_{DCTLBP} respectively to represent colour images in terms of data-level fusion. To explore clustering algorithms, we used a Normalize Spectral algorithm (*graph-based*) to aggregate local F_{DCTLBP} features. $F_{WLBPYCbCr}$ satisfied 13% more comparing with F_{LBPG} and F_{DCTLBP} satisfied (10-8) % more comparing with F_{DCT} and $F_{LBPYCbCr}$. Consequently, each feature type can bring different scores of retrieval. Therefore, we developed a method that uses both $WF_{LBPYCbCr}$ and F_{DCTLBP} features and combines retrieved scores with appropriate weights in terms of score-level fusion.

6. Conclusions

In this paper, we proposed a method that uses global and local features using a spectral clustering algorithm (*graph-base*). The method improved the performance of image retrieval compared to other methods from the CBIR literature for Top 10 retrieved images as illustrated in Table 5.

We can conclude that the spectral clustering algorithm is benefit for CBIR, bringing accurate results particularly for images with complicated scenes. The algorithm participates to increase the effective of representing images.

On the other hand, above method is unsupervised means a class label is not involved to determine relevant images. Hence, future investigation is a supervised method such as Convolutional Neural Network which is a deep learning technique to learn image features. Resulted features can be tested in CBIR.

Table 5. Comparison of MAP of different methods for Top10 retrieved images

Method	Top10
Proposed Method	0.76
[36]	0.73
[37]	61.99
[38]	0.74
[39]	0.66
[40]	0.58

7. References

1. Smeulders, A. W. *et al.* (2000). *"Content-based image retrieval at the end of the early years"*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22(12), pp. 1349-1380.
2. Datta, R. Joshi, D. , Li J. & Wang, J. Z. (2008). *"Image retrieval: Ideas, influences, and trends of the new age"*. ACM Computing Surveys (CSUR), 40(2), p. 5E.
3. Yung-Gi, W. (2006). *"Region of interest image indexing system by DCT and entropy"*. GVIP Journal, 6(4), pp. 7-16.
4. Ren, Z., Gao, S., Chia, L.T. and Tsang, I.W.H. (2014). *"Region-based saliency detection and its application in object recognition"*. IEEE Transactions on Circuits and Systems for Video Technology, 24(5), pp.769-779.
5. Grycuk, R., Gabryel, M., Korytkowski, M. and Scherer, R. (2014). *"Content-based image indexing by data clustering and inverse document frequency"*. In International Conference: Beyond

Databases, Architectures and Structures (pp. 374-383), Springer.

6. Sakji-Nsibi, S. and Benazza-Benyahia, A. (2016). *"Region-based image retrieval using a joint scalable Bayesian segmentation and feature extraction"*. In Signal Processing Conference(EUSIPCO), 2016 24th European (pp. 1272-1276), IEEE.
7. Pinjarkar, L., Sharma, M. and Mehta, K. (2012). *"Relevance feedback techniques in content based image retrieval: A survey"*. Int. J. Eng, 1(9), pp.19-23.
8. Zhou, X. S. & Huang, T. S. (2003). *"Relevance feedback in image retrieval: A comprehensive review"*. Multimedia Systems, 8(6), pp. 536-544.
9. Yasmin, M., Mohsin, S., Irum, I. and Sharif, M. (2013). *"Content based image retrieval by shape, color and relevance feedback"*. Life Science Journal, 10(4s), pp.593-598.
10. Papadopoulos, G.T., Apostolakis, K.C. and Daras, P. (2014). *"Gaze-based relevance feedback for realizing region-based image retrieval"*. IEEE Transactions on Multimedia, *16*(2), pp.440-454.
11. Kundu, M.K., Chowdhury, M. and Bulò, S.R. (2015). *"A graph-based relevance feedback mechanism in content-based image retrieval"*. Knowledge-Based Systems, *73*, pp.254-264.
12. Ojala, T. Pietikainen, M. and Harwood, D. (1994) "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions". Proceedings of the 12th International Conference on Pattern Recognition Computer Vision \& Image Processing, pp. 582-585.
13. Ojala, T., Pietikainen, M. and Maenpaa, T. (2002). *"Multiresolution gray-scale and rotation invariant texture classification with local binary patterns"*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 24(7), pp. 971-987.
14. Ahonen, T., Hadid, A. and Pietikäinen, M. (2004). "*Face recognition with local binary patterns*". European conference on computer vision. Springer, Berlin, Heidelberg.
15. Ahonen, T., Hadid, A. and Pietikäinen, M. (2006). "*Face description with local binary patterns: Application to face recognition*". IEEE Transactions on Pattern Analysis & Machine Intelligence 12 (2006): 2037-2041.
16. Chan, C.H., Kittler, J. and Messer, K. (2007). "*Multi-scale local binary pattern histograms for face recognition*". In International conference on biometrics (pp. 809-818). Springer, Berlin, Heidelberg.
17. Mäenpää, T. & Pietikainen, M. (2004). "*Classification with color and texture: jointly or separately*?". Pattern Recognition, 37(8), pp. 1629-1640.
18. Takala, V., Ahonen, T. & Pietikäinen, M. (2005). "*Block-based methods for image retrieval using local binary patterns*". In Scandinavian Conference on Image Analysis (SCIA), Springer Berlin Heidelberg, pp. 882-891.
19. Guo, Z. & Zhang, L. & Zhang, D. (2010). "*A Completed Modeling of Local Binary Pattern Operator for Texture Classification*". IEEE transactions on image processing : a publication of the IEEE Signal Processing Society. 19. 1657-63. 10.1109/TIP.2010.2044957.
20. Verma, M. & Raman, B. (2017). "*Local neighborhood difference pattern: A new feature descriptor for natural and texture image retrieval*". Multimedia Tools and Applications. 1-25. 10.1007/s11042-017-4834-3.
21. Singh, C., Walia, E. and Kaur, K.P. (2018). "*Color texture description with novel local binary patterns for effective image retrieval*". Pattern recognition, 76, pp.50-68.

22. Lay, G.A. and Guan, L. (1999). "*Image retrieval based on energy histograms of the low frequency DCT coefficients*". Proceedings of 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 3009-3012.
23. Schaefer, G. (2011). "*Content-based image retrieval: Advanced topics*". Man-Machine Interactions 2. Springer Berlin Heidelberg, pp. 31-37.
24. Abd-Elhafiez, W. M. and Gharibi, W. (2012). "*Color Image Compression Algorithm Based on the DCT Blocks*", arXiv preprint arXiv:1208.3133.
25. Kekre, H. B., Thepade, S. D., Chaturvedi, R. N. & Gupta, S. (2012). "*Walsh, Sine, Haar & Cosine Transform with various color spaces for 'Color to Gray and Back'*". Internetional Journal of Image Processing (IJIP), Vol.6, No.5, pp349-356.
26. Huang, Y-L. and Chang, R-F. (1999). "*Texture features for DCT-coded image retrieval and classification*". IEEE International Conference on Acoustics, Speech, and Signal Processing, Phoenix, AZ , pp. 3013-3016.
27. Nezamabadi-Pour, H. and Saryazdi, S. (2007). "Object-based image indexing and retrieval in DCT domain using clustering techniques". International Journal of Computer and Information Engineering, Vol.1, No.3, pp. 98-101.
28. Al-Jubouri, H. A. (2015). "Multi-evidence fusion scheme for content-based image retrieval by clustering localised colour and texture features ". Doctoral thesis, University of Buckingham, 2015.
29. Du, H. (2010). "*Data mining techniques applications: an introduction*". *Cengage Learning EMEA.*
30. Jain, A. (2010). "*Data clustering beyond K-means*". Pattern Recognition Letters, Vol 31, pp. 651-666.
31. Ng, A. Y., Jordan, M. I. & Weiss, Y. (2001). *On Spectral Clustering: Analysis and an algorithm*". Cambrige, USA, MIT Press, pp. 849-856.
32. Wang, J. Z., Li, J. & Wiederhold, G. (2001). "*SIMPLIcity: Semantics-sensitive integrated matching for picture libraries*". Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23(9), pp. 947-963.
33. Graziano, A. M. and Raulin, M. L. (2012). "*Research Mwthods: A Process of Inquiry*". 8th ed., New York, HarperCollins College Publishers.
34. Fox, E. A. and Shaw, J. A. (1994). "*Combination of Multiple Searches*". NIST SPECIAL PUBLICATION SP, pp. 243-246.
35. Lee, J. (1997). "*Analyses of Multiple Evidence Combination*". New York, NY, USA, Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 267-276.
36. Karpagam, V. and Rangarajan, R. (2012). "*A Simple and Competent System for Content Based Retrieval of Images using Color Indexed Image Histogram Combined with Discrete Wavelet Decomposition*". European Journal of Scientific Research, 73(2), pp. 278-190.
37. Salmi, M. and Boucheham, B. (2014). "Content based image retrieval based on Cell Color Coherence Vector (Cell-CCV)". Proceedings of 4th International Symposium ISKO-Maghreb: Concepts and Tools for knowledge Management, pp. 1-5.
38. Chaudhary, M. D. and Upadhyay, A. B. (2014). "Fusion of local and global features using Stationary Wavelet Transform for efficient Content Based Image Retrieval". IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1-6.

39. Feng, L., Wu, J. Liu, S. and Zhang, H. (2015). "*Global correlation descriptor: a novel image representation for image retrieval*". Journal of Visual Communication and Image Representation, Vol.*33*, pp.104-114.
40. Lokoč, J. D., Novák, Batko, M. and Skopal, T. (2012). "*Visual image search: feature signatures or/and global descriptors*". LNCS on Similarity Search and Applications, Springer-Verlag, Berlin, Heidelberg, pp. 177-191.
41. Sarwar, A. *et al.* (2018). "*A novel method for content-based image retrieval to improve the effectiveness of the bag-of-words model using a support vector machine*". Journal of Information Science.
42. Wan, J., *et al*. (2014). "*Deep learning for content-based image retrieval: A comprehensive study*". *Proceedings of the 22nd ACM international conference on Multimedia*.
43. Xiang, J., *et al*. (2019) "*Fabric image retrieval system using hierarchical search based on deep convolutional neural network*". *IEEE Access* 7.
44. Plant, W. and Schaefer, G., (2009). "*Navigation and Browsing of Image Databases*". International Conference of Soft Computing and Pattern Recognition, Malacca, pp. 750-755.