2018 International Conference on Computing, Mathematics and Engineering Technologies - iCoMET 2018

Content Based Image Retrieval System by using HSV Color Histogram, Discrete Wavelet Transform and Edge Histogram Descriptor

Atif Nazir
Department Of Computer Science
National Textile University
Faisalabad,Pakistan
atif.nazir1225@gmail.com

Rehan Ashraf
Department Of Computer Science
National Textile University
Faisalabad,Pakistan
rehan@ntu.edu.pk

Talha Hamdani
Department Of Textile Engineering
National Textile University
Faisalabad,Pakistan
talha@ntu.edu.pk

Nouman Ali
Department of Software Engineering
Mirpur University of Science and Technology
Azad Kashmir, Pakistan
nouman.ali@live.com

***Abstract*—In last few decades, Content Based Image Retrieval System (CBIR) is an emerging field to retrieve relevant images from a database. It utilizes the visual contents of an image for the local and global features. Local feature includes spatial domain which presents the significance of the image as well as the index of an image. Global feature includes shape descriptors, contour representations and texture features. Segmentation process is required in global feature extraction technique. It is a challenging task to simulate visual information in CBIR system. CBIR strategy combines the local and global features to deal with the low level information. In this paper, we proposed new CBIR technique to fuse color and texture features. Color Histogram (CH) is used to extract a color information. Texture features are extracted by Discrete Wavelet Transform (DWT) and Edge Histogram Descriptor (EDH). The features are created for each image and stored as a feature vector in the database. We evaluated our work using Corel 1-k dataset. To examine the accuracy with the other proposed systems, precision and recall methods are used that provides competitive and efficient result. The experimental results show that our proposed method outperforms with existing CBIR systems.**

***Index Terms*—Content Based Image Retrieval, Image Retrieval, Color Histogram, Discrete Wavelet Transforms, Edge Histogram Descriptor**

I. Introduction

Due to increase of visual data like videos and images over the internet, we need an alternate solution to store such data. The storage size of the database is also very difficult to handle and it requires a lot of processing time. To handle such data, it is essential to improve our application to cope with these type of data collections. CBIR system is used to find similar images from a database where the user provides a query image and system provides most similar images. In this context, there are two types of approaches are used for image retrieval. First one is text based image retrieval that provides results with the semantic similarities and the second one is Content based Image Retrieval (CBIR) that provide results with the visual information. CBIR system overcomes the problems of text based image retrieval because manual annotation is not used in this system. Image retrieval system is based on a feature extraction technique. Features are extracted from the query image then similarity measure is used to find the most similar image from a database. Feature extraction technique is classified into two categories, local and global features. Local feature technique includes color, shape and texture features discussed in $[1-4]$ literature and these feature used for object identification and recognition. Recognition is a process to identify objects such as recognizing a person or identification of the existing objects of an image. This process used in many application like toll booth monitoring, factory automation, and security surveillance. SIFT [5], SURF [5], MSER [6] and LBP [7] are the example of local descriptor. Color is an effective technique as compared to other features such as texture and shape. Color feature techniques are Color Histogram [8], Fuzzy Color Histogram [9], Color Moments etc. Texture are the visual patterns and they are prominent in natural scenes such as fabrics, grasslands and bricks wall. Texture feature extraction techniques are Discrete Wavelet Transform [10], Gabour Wavelet Transform [10], Co-occurrence Matrix [11] etc. Global feature techniques are used to retrieve object classification and object detection in image. These are shape, texture and local feature descriptor of an image. It describes the whole image and segmentation process is applied to find the appropriate features. Histogram Oriented Gardients (HOG)

978-1-5386-1370-2/18/$31.00 ©2018 IEEE

[7], CO-HOG [12], Invariant Moment (Zernike, Hue) [13] are the example of global descriptor. From the Literature review, mostly all CBIR systems used color, shape, texture and spatial information. In last few years combine a multiple features gives an efficient result as compared to single feature. There are many processes are proposed such as indexing, classification and extracting visual information using color, shape, and texture [14]. Color features are consist of histogram bins in MPEG-7, such us color layout descriptor, saleable color descriptor and dominant color descriptor [15,16]. The Fusion of color and texture based images reterival system is explained in [4,11,17,18]. Region of interest (ROI) point are determined by gabor filter describe in [19]. Texture features are extracted by gabor features, shape features are based on zernike moments and color moments, color histogram is used to extract the color features [19]. Color feature requires a low complexity of image retrieval, it does not require the distribution of color information. For this reason in Content Based Image Retrieval we require a texture feature. Texture feature has been widely used in machine vision, pattern recognition and image retrieval [20]. In this paper [8] researcher combines the local and global features using color correlogram and dominant color. Local features are extracted by color correlogram and Dominant color are used to extract the global features.

In this paper texture and color features are combined to extract a relevant image from a large database. Color feature is extracted by Color Histogram and texture features are extracted by Discrete Wavelet Transform and Edge Histogram Descriptor. The rest of paper is divided into following sections.

II. Feature Extraction

Feature extraction is a basic and fundamental step in a CBIR system which depends on the how the researchers define the visual content or visual signature (composition of multiple features). In Content Based Image Retrieval system an appropriate representation of image is required. An image is a group of pixels which is used to resolve the computational tasks in computer vision. Images represent their feature on the base of low level and high level features. Low level feature includes color, shape and texture features and high level is based on a machine learning techniques. High Level technique is based on a classification or interpretation of a whole image that includes object detection [21], face detection [22] and human body pose classification [21].

A. Color Histogram

In CBIR system color histogram [23] is an effective approach to retrieve an image. The 3D image is potentially superior to 2-D images. In order to extract the color features, we primarily find the color space of an image which describes the array of a color. For example, in RGB color space there are three values in color components. In our proposed method we have used an HSV color space because hue and saturation are close to human visual system reported as $[24-26]$. In image block, color quantization process shrinks the total number of different colors. RGB image can be easily transformed into HSV color space [27], by using this formula

$$H = cos^{-1}\frac{\frac{1}{2}[R-G]+[R-B]}{\sqrt{(R-G)^2-(G-B)(R-B)}} \quad (1)$$

$$S = 1-\left(\frac{3[min(R,G,B)]}{R+G+B})\right) \quad (2)$$

$$V = \left(\frac{R+G+B}{3}\right) \quad (3)$$

Here H represents hue, which is the prominent wavelength in the collection of waves. S represents the saturation which represents the white light. V represents values in color space which describes the intensity value in the image: range in between 0 and 1. Here 0 represent black and 1 represents a white color.

B. Discrete Wavelet Transform

Wavelets are a little wave or a function that is concentrated in time as well as the frequency around a certain point. Fourier transforms only deal with the frequency component in a signal while temporal detail is not available, so we used a Discrete Wavelet Transform (DWT) [10] because it is appropriate for non-stationary signals and vary for both the frequency range and spatial range. It is used to compose the signal, approximation and detail information of an image. In DWT query image is divided into rows and columns by Low pass (L) which are basically the average intensity value of the image and High temporal resolution (H) for high pass frequency component which are the edges of an image. The basic step to find the Discrete Wavelet Transform is to use a mother wavelet such as coiflets, Haar, Daubechies available in Matlab. In this paper, we use a Daubechies db1 wave name because it gives an efficient result as compared to other wavelets. First, we have to apply a low pass filter to preserve

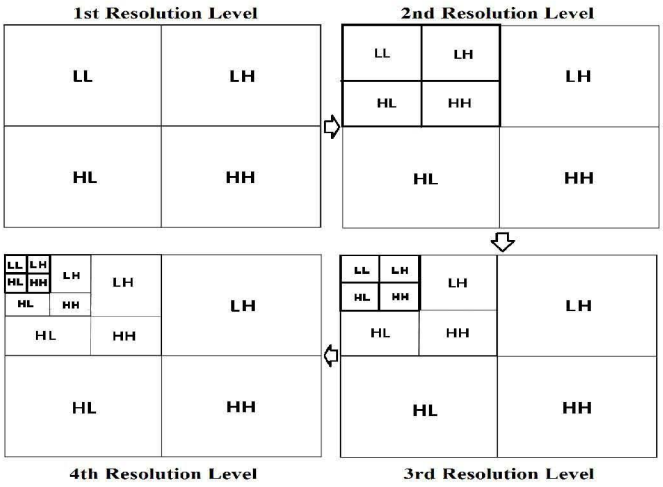

Fig. 1. One To Four-Level Decomposition In DWT

the low frequency and high pass filter to preserve the high frequency of an image. After applying a low pass filter we get a horizontal approximation and by applying a high pass filter we get a horizontal detail. Same process is applied to the column to get first approximation like (LL, LH, HL, and HH) magnitudes then calculate decomposition of DWT at four levels. A block diagram of four level decomposition is shown in Fig. 1.

C. Edge Histogram Descriptor

Edge Histogram Descriptor [28] represents the geometry of an Image and it is designed to depict only the distribution of local edges in image. Edges are explained the significant attribute to visualize the content of images and the histogram is used to represent the Edge. The features described of an images using EDH cannot be duplicated by homogeneous color histogram and texture feature. In MPEG-7 [18,29] we use these features to represent its attribute, and also there is a huge description of the edge distribution in image. In the spatial domain, image block is extracted by a digital filter. The Edge Histogram Descriptor represents five type of edges [30] includes non-directional edge and four directional edges. Directional edges consist of 135-degree edge, 45-degree edge, horizontal and vertical edge as shown in Fig. 2. If the sub block image are neither directional nor arbitrary, then this type of edge is fall in non-directional category.

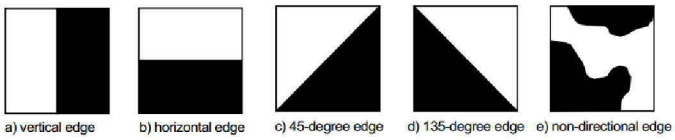

Fig. 2. Five Type Of Edges [30]

The main purpose of EDH is used in MPEG-7 to find the relevant images efficiently [28]. Firstly, Image is converted into gray level. The image is divided into sub-block to find the edges of each block. Edges are extracted by the direction and non-directional edge feature. The division of an image is based on a fixed size of image blocks. Each sub-block image size is 2 x 2 matrix, this algorithm [30] perform to masking for each image block. Finally, EDH is calculated for a query image.

III. Similarities Measurement

In literature eight distance method have been proposed [31]. In this research, manhaten distance (d) is used.

$$d(F_{query}, F_{dbase}) = \sum_{i=0}^{N-1} \left| f_{query}[i] - f_{dbase}[i] \right| \qquad (4)$$

Here $F_{query}[i]$ is a feature vector for a query image and $F_{dbase}[i]$ is a feature vector for a database image. i denotes the range $1.....N$ and N represent the total numbers of descriptor in each vector.

IV. Proposed Research

The overall working flow is shown in Fig. 3 and proposed research has following steps.

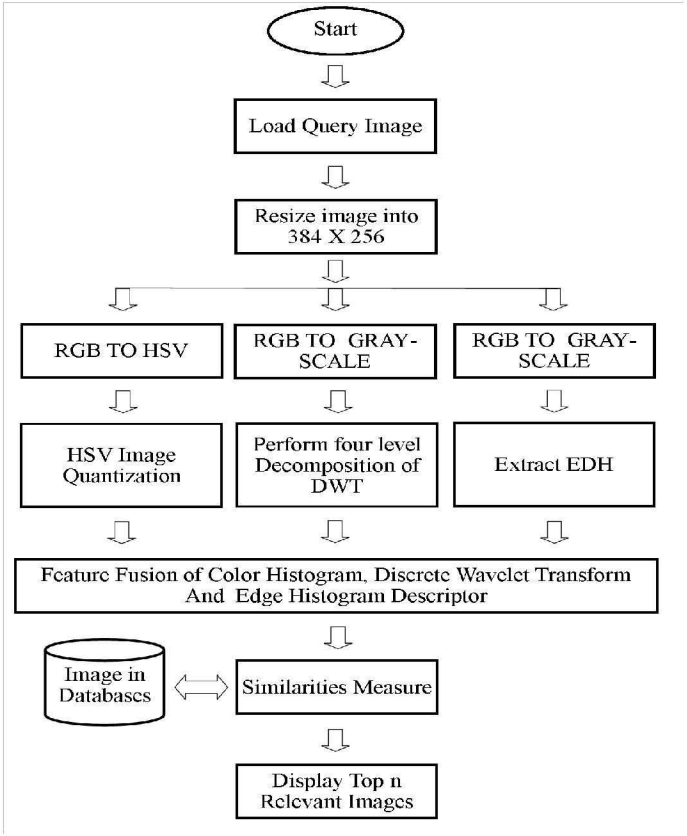

Fig. 3. Block Diagram OF Proposed Research

1) In first step image is resized into 384×256 pixels.
2) Color image is converted into HSV color space to extract color histogram. Color Histogram is constructed by quantization of HSV image into feature vector to construct a normalized histogram which is $F_{ch} = \{f_1 + f_2 + f_3.....f_{128}\}$.
3) Discrete Wavelet Transforms and Edge Histogram Descriptor is deployed for a grayscale image. Perform four level discrete wavelet transform to image and construct a feature vector $F_{dwt} = \{f_1+f_2+f_3.....f_{32}\}$ for the first two moments those are mean and standard deviation. Edge Histogram Descriptor features are extracted by MPEG-7 Descriptor. Feature vector is constructed by EDH is $F_{edh} = \{f_1 + f_2 + f_3.....f_{150}\}$.
4) The computed features are extracted from color and texture features represents an image as a combination of three features which is $F_{total} = \{f_1+f_2+f_3.....f_{310}\}$.
5) Query image is taken and repeat the step 1-4 to build a feature vector based on color and texture feature.
6) Distance between the query image and database image

is calculated using Manhattan distance [31] is

$$d(F_{query}, F_{dbase}) = \sum_{i=0}^{309} \frac{f_{query}[i] - f_{dbase}[i]}{1 + f_{query}[i] + f_{dbase}[i]} \quad (5)$$

The value of equation 5 is always in-between 0 and 1. Here 0 indicates that images are same and 1 indicates the images are not same.

V. Experimental Result and Performance Evaluation

In this section we briefly describe materials used, performance evaluation method and experimental result of our proposed method.

A. Materials

Several datasets of images are used, such as Caltech 256, OLIVIA, Corel Stock, Caltech101 image set. In the domain of CBIR, Corel dataset are famous to evaluate performance [33,34,35,36]. In our paper, we have used a Corel 1-k dataset and compare our results with [4,7,18,32]. This data set contains 1000 images which are divided into 10 classes. Each class contains 100 images named are Africa, Beach, Bus, Building, Dinosaur, Flower, Elephant, Horse, Mountain, and Food. All of the images were characterized by RGB color-space. Some sample images of Corel 1-k are shown in Fig. 5. For each category, we selected one image with the size of 384 x 256 pixels.

Fig. 4. Sample Images For Corel-1k Data Set

B. Performance Evaluation

In a CBIR system find an accuracy of a system is an important task. Precision-recall is the most commonly used methods to evaluate the accuracy of retrieved image [4,7,18,32]. Precision P (also known as positive predictive value) is derived as the ratio of the number of relevant images retrieved r by the total number of images retrieved n. Mathematical definition of Precision is

$$P = \frac{NumberOfReleventImageReterieved}{TotalNumberOfImagesReterieved} \quad (6)$$

Precision is used to find the accuracy of image retrieval. Recall R (also known as sensitivity) is termed as the ratio of the number of related instances retrieved r by the total number of relevant instances m in the database, here instances referred to an image in the database. Mathematical definition of recall is

$$R = \frac{NumberOfReleventImageReterieved}{TotalNumberOfReleventImagesInDatabase} \quad (7)$$

Recall is used to find the robustness of image retrieval. Precision and Recall method provides very competitive and efficient results. To find the significance of proposed features, our experimental scheme is divided into four classes. In the first three cases, only single color histogram, discrete wavelet transform, and edge histogram descriptor is mentioned. After combining all three features which were proposed in our method, we achieved better results as shown in Table I. It shows that single feature gives a less result as compared to combination of multiple features. The comparison of precision and recall with state of the art system shown in Table II and Table III.

TABLE I
PRECISION FOR TOP-20 IMAGE RETRIEVALS WITH SINGLE AND FUSED FEATURES

Category	CH	4D-DWT	CH+4D-DWT	EDH	Proposed Research
Africa	0.65	0.55	0.55	0.65	0.85
Beach	0.35	0.25	0.30	0.20	0.50
Building	0.45	0.55	0.75	0.50	0.75
Bus	0.75	0.70	0.80	0.85	1
Dinasour	0.95	0.95	1	0.90	1
Elephant	0.35	0.35	0.80	0.15	0.55
Flower	0.70	0.70	0.95	0.65	0.95
Horse	0.75	0.85	0.90	0.60	0.90
Mountain	0.75	0.30	0.40	0.25	0.30
Food	0.35	0.65	0.50	0.30	0.55
Average	0.605	0.585	0.695	0.505	0.735

TABLE II
COMPARISON OF PRECISION WITH STATE OF THE ART SYSTEM

Category	M.E. Elalami [32]	J.Yue, etl [4]	J. Yu etl [7]	S.Somnug etl [18]	Proposed Research
Africa	0.58	0.53	0.57	0.676	0.85
Beach	0.41	0.45	0.58	0.598	0.50
Building	0.42	0.46	0.43	0.58	0.75
Bus	0.71	0.84	0.93	0.94	1
Dinosaur	0.74	0.90	0.98	0.998	1
Elephant	0.65	0.72	0.666	0.58	0.55
Flower	0.83	0.74	0.83	0.886	0.95
Horse	0.69	0.72	0.68	0.938	0.90
Mountain	0.44	0.53	0.46	0.478	0.30
Food	0.44	0.46	0.53	0.492	0.55
Average	0.595	0.641	0.650	0.725	0.735

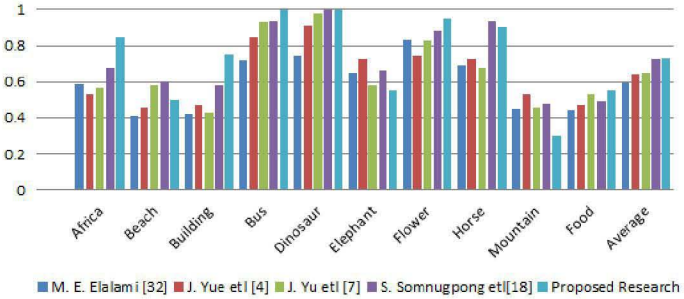

Fig. 5. Comparison of Precision Graph With State Of The Art System

TABLE III
COMPARISON OF RECALL WITH STATE OF THE ART SYSTEM

Category	M.E. Elalami [32]	J.Yue, etl [4]	J. Yu etl [7]	S.Somnug etl [18]	Proposed Research
Africa	0.12	0.11	0.11	0.13	0.17
Beach	0.08	0.09	0.12	0.12	0.10
Building	0.08	0.09	0.08	0.12	0.15
Bus	0.14	0.17	0.19	0.19	0.20
Dinosaur	0.15	0.18	0.19	0.19	0.20
Elephant	0.13	0.15	0.12	0.13	0.11
Flower	0.17	0.15	0.16	0.18	0.19
Horse	0.14	0.14	0.13	0.19	0.18
Mountain	0.09	0.11	0.09	0.09	0.06
Food	0.09	0.09	0.10	0.10	0.11
Average	0.119	0.128	0.129	0.144	0.147

Fig. 6. Top-20 Retrieval Result For Bus Image

VI. CONCLUSION AND FUTURE DIRECTION

A color and texture based image retrieval is explained in our work by using low level features with the combination of local and global features. Combining two or more features give a better result as compared to one feature because color and texture feature gives an efficient and appropriate result in the human visual system. Local features are extracted by edge histogram descriptor and global features are extracted by color histogram and discrete wavelet transform. In this paper, a more appropriate color and texture features are used to form a feature vector and similarities are matched by manhattan distance. Results are compared with the state of the art systems which shows an outperform result. In future work, researchers use different features along with existing features and machine learning techniques such as ANN for further improvement.

REFERENCES

[1] C.-H. Lin, R.-T. Chen, and Y.-K. Chan, A smart content-based image retrieval system based on color and texture feature, Image Vis. Comput., vol. 27, no. 6, pp. 658–665, 2009.

[2] X.-Y. Wang, H.-Y. Yang, and D.-M. Li, A new content-based image retrieval technique using color and texture information, Comput. Electr. Eng., vol. 39, no. 3, pp. 746–761, 2013.

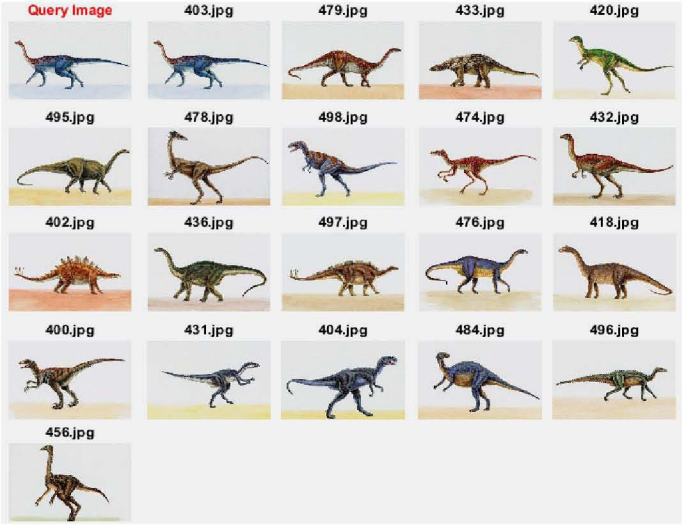

Fig. 7. Top-20 Retrieval Result For Dinosaur Image

Fig. 8. Top-20 Retrieval Result For Flower Image

Fig. 9. Top-20 Retrieval Result For Horse Image

[3] J. Xu, B. Xu, and S. Men, Feature-based similarity retrieval in Content-based image retrieval, Proc. - 7th Web Inf. Syst. Appl. Conf. WISA 2010, Work. Semant. Web Ontol. SWON 2010, Work. Electron. Gov. Technol. Appl. EGTA 2010, pp. 215–219, 2010.
[4] J. Yue, Z. Li, L. Liu, and Z. Fu, Content-based image retrieval using color and texture fused features, Math. Comput. Model., vol. 54, no. 34, pp. 1121–1127, 2011.
[5] N. Ali et al., A novel image retrieval based on visual words integration of SIFT and SURF, PLoS One, vol. 11, no. 6, 2016.
[6] K. Mikolajczyk, K. Mikolajczyk, C. Schmid, and C. Schmid, A performance evaluation of local descriptors, IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 10, pp. 1615–1630, 2005.
[7] J. Yu, Z. Qin, T. Wan, and X. Zhang, Feature integration analysis of bag-of-features model for image retrieval, Neurocomputing, vol. 120, pp. 355–364, 2013.
[8] A. N. Fierro-Radilla, M. Nakano-Miyatake, H. Perez-Meana, M. Cedillo-Hernandez, and F. Garcia-Ugalde, An efficient color descriptor based on global and local color features for image retrieval, 2013 10th Int. Conf. Electr. Eng. Comput. Sci. Autom. Control. CCE 2013, no. Dc, pp. 233–238, 2013.
[9] S. A. Chatzichristofis and Y. S. Boutalis, FCTH: Fuzzy Color and texture histogram a low level feature for accurate image retrieval, WIAMIS 2008 - Proc. 9th Int. Work. Image Anal. Multimed. Interact. Serv., pp. 191–196, 2008.
[10] H. Farsi and S. Mohamadzadeh, Colour and texture feature-based image retrieval by using hadamard matrix in discrete wavelet transform, IET Image Process., vol. 7, no. 3, pp. 212–218, 2013.
[11] M. E. Elalami, A new matching strategy for content based image retrieval system, Appl. Soft Comput. J., vol. 14, no. PART C, pp. 407–418, 2014.
[12] C. Hans, C. McCollum, M. Bondesson, S. Shah, and F. Merchant, Gradient weighted CO-HOG for analysis of caudal vein structural changes in toxin exposed zebrafish embryo, in Proceedings - International Symposium on Biomedical Imaging, 2015, vol. 2015July, pp. 524–527.
[13] A. Goyal and E. Walia, Variants of dense descriptors and Zernike moments as features for accurate shape-based image retrieval, Signal, Image Video Process., vol. 8, no. 7, pp. 1273–1289, 2014.
[14] A. Marakakis, G. Siolas, N. Galatsanos, a. Likas, and a. Stafylopatis, Relevance feedback approach for image retrieval combining support vector machines and adapted Gaussian mixture models, IET Image Process., vol. 5, no. 6, p. 531, 2011.
[15] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, and A. Yamada, Color and texture descriptors, IEEE Trans. Circuits Syst. Video Technol., vol. 11, no. 6, pp. 703–715, 2001.
[16] T. Sikora, The MPEG-7 visual standard for content description - An overview, IEEE Trans. Circuits Syst. Video Technol., vol. 11, no. 6, pp. 696–702, 2001.
[17] H. A. Jalab, Image retrieval system based on color layout descriptor and Gabor filters, 2011 IEEE Conf. Open Syst. ICOS 2011, pp. 32–36, 2011.
[18] S. Somnugpong and K. Khiewwan, Content-Based Image Retrieval using a Combination of Color Correlograms and Edge Direction Histogram, 2016.
[19] R. S. Chora, T. Andrysiak, and M. Chora, Integrated color, texture and shape information for content-based image retrieval, Pattern Anal. Appl., vol. 10, no. 4, pp. 333–343, 2007.
[20] C.-H. Lin, C.-W. Liu, and H.-Y. Chen, Image retrieval and classification using adaptive local binary patterns based on texture features, IET Image Process., vol. 6, no. 7, p. 822, 2012.
[21] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, A survey of deep neural network architectures and their applications, Neurocomputing, vol. 234, pp. 11–26, 2017.
[22] R. Ashraf, K. Bashir, and T. Mahmood, Content-based image retrieval by exploring bandletized regions through support vector machines, J. Inf. Sci. Eng., vol. 32, no. 2, pp. 245–269, 2016.
[23] K. Hemachandran, a. Paul, and M. Singha, Content-based image retrieval using the combination of the fast wavelet transformation and the colour histogram, IET Image Process., vol. 6, no. 9, pp. 1221–1226, 2012.
[24] M. Singha and K. Hemachandran, Performance analysis of color spaces in image retrieval, Assam Univ. J. , vol. 7, no. 2, 2011.
[25] X. Wan and L. Angeles, Color Distribution Analysis and Quantization for Image Retrieval, vol. 2670, no. 213, pp. 8–16.
[26] S. Sural, Gang Qian, and S. Pramanik, Segmentation and histogram generation using the HSV color space for image retrieval, Proceedings. Int. Conf. Image Process., vol. 2, p. II–589–II–592, 2002.
[27] C.-H. Su, H.-S. Chiu, and T.-M. Hsieh, An efficient image retrieval based on HSV color space, 2011 Int. Conf. Electr. Control Eng., pp. 5746–5749, 2011.
[28] D. K. Park, Y. S. Jeon, and C. S. Won, Efficient use of local edge histogram descriptor, Multimed. 00 Proc. 2000 ACM Work. Multimed., pp. 51–54, 2000.
[29] S. Agarwal, Content Based Image Retrieval using Discrete Wavelet Transform and Edge Histogram Descriptor, IEEE, Int. Conf. Inf. Syst. Comput. Networks, pp. 19–23, 2013.
[30] C. S. Won, D. K. Park, and S. J. Park, Efficient use of MPEG-7 edge histogram descriptor, ETRI J., vol. 24, no. 1, pp. 23–30, 2002.
[31] D. Zhang and G. Lu, Evaluation of similarity measurement for image retrieval, Proc. 2003 Int. Conf. Neural Networks Signal Process. ICNNSP03, vol. 2, pp. 928–931, 2003.
[32] M. E. Elalami, A novel image retrieval model based on the most relevant features, Knowledge-Based Syst., vol. 24, no. 1, pp. 23–32, 2011.
[33] Ali N, Mazhar DA, Iqbal Z, Ashraf R, Ahmed J, Khan FZ. Content-Based Image Retrieval Based on Late Fusion of Binary and Local Descriptors. arXiv preprint arXiv:1703.08492. 2017.
[34] Z. Mehmood, S. M. Anwar, N. Ali, H. A. Habib, and M. Rashid, A Novel Image Retrieval Based on a Combination of Local and Global Histograms of Visual Words, Math. Probl. Eng., vol. 2016, 2016.
[35] R. Ashraf, K. Bashir, A. Irtaza, and M. T. Mahmood, Content Based Image Retrieval Using Embedded Neural Networks with Bandletized Regions, pp. 3552–3580, Entropy 2015.
[36] R. Ashraf, T. Mahmood, A. Irtaza, and K. Bajwa, A novel approach for the gender classification through trained neural networks, J. Basic Appl. Sci. Res, vol. 4, pp. 136–144, 2014.