Journal of Advances in Computer Research
Quarterly pISSN: 2345-606x eISSN: 2345-6078
Sari Branch, Islamic Azad University, Sari, I.R.Iran
(Vol. 8, No. 4, November 2017), Pages: 59-72
www.jacr.iausari.ac.ir

A Novel Method for Content - Based Image Retrieval Using Combination of Local and Global Features

Elham Heidari[✉1], **Sajjad Tavassoli**[2]
1) University College of Rouzbahan, Sari, Iran
2) Department of Computer Engineering, Sari Branch, Islamic Azad University, Sari, Iran
heidari.elham3@gmail.com; tavassoli_5@yahoo.com

Received: 2016/08/16; Accepted: 2016/11/12

Abstract

Content-based image retrieval (CBIR) has been an active research topic in the last decade. In this paper we proposed an image retrieval method using global and local features. Firstly, for local features extraction, SURF algorithm produces a set of interest points for each image and a set of 64-dimensional descriptors for each interest points and then to use Bag of Visual Words model, a clustering algorithm is used to obtain the visual vocabulary and each resulted centroid represent a visual word. Then, images are viewed as BoVW represented as histogram. In order to improve retrieval performance, global feature is extracted by HSV color feature. Finally, this approach uses the combined local and global features as feature vectors to provide image retrieval. The COREL image database has been used for our experimental results. The experimental results show that the performance of the combination of both local and global features is much higher than each of them, which is applied separately.

Keywords: *SURF, K-Means, Bag-of-Visual Word, Color Features, Local Features*

1. Introduction

Image retrieval is the field of the study that concerned with looking, browsing, and recovering digital images from an extensive database. Most of the CBIR system uses the global features such as color, texture and shape to extract the features from the images [1]. Image features descriptors can be either global or local .The global feature descriptors describe the visual content of the entire image. Global features cannot describe all the characteristics of the image [2].Global features related to color, texture or shape are usually used to provide a low level description of image content. The global features cannot capture all parts of the image having different characteristics. This can be overcome by local features. Then the locale of interest is encoded overall, which results in the image descriptor [3]. Local feature describes describe a patch within an image of the image content. Local descriptors depict a pixel in an image through its local neighborhood content which should be distinctive and robust to changes in viewing environment or deformations or localization errors. Local Features find the corresponding pixel locations in images which compute the same amount of information about the spatial intensity patterns under different conditions [4],[5]. The superiority of the global descriptor extraction is the increased speed for both feature extraction and computing similarity. Content-Based Image Retrieval (CBIR) approaches query the images with their real contents instead of their annotated metadata such as keywords, tags or text descriptions [6]. Primary CBIR approaches with low-level visual features such color, texture and shape easily perceived low-level features [7], [8]. Color features are considered to be stable, robust and invariant to scaling, translation and rotation regarding other visual features [10]. The interest point detectors and descriptors [11] are employed in many Content-based image retrieval (CBIR) systems. SURF algorithm [12], or Speeded-Up Robust Features, is a robust image local features detector which detects interest points and produces their descriptors. It is widely used in most of the computer vision applications. The speeded up robust features algorithm is a scale and rotation-invariant interest point detector and descriptor which are computationally very fast. The detector locates the interest points in the image, and the descriptor depicts the features of the interest points and constructs a distribution of Haar-wavelet responses within the interest point neighborhood as feature vectors of the interest points [12].

The feature vectors are extracted from the images in the database and described by multidimensional feature vectors, which form a feature database. To retrieve images, the feature vectors are extracted from the given query image. The similarities between the feature vectors of the query image and the feature vectors of the database images are then calculated. And the retrieval is performed with the aid of an indexing scheme and matching strategy, which provide an efficient way to search the image database .SURF is one of the best interest point detectors and descriptors currently available.

In this work, SURF algorithm is used to extract the features and the first order and second order Color features of images are extracted by using HSV color space. To improve the performance of the system the SURF is combined with the global features (color, texture and shape).Usually, for each image, there would be hundreds of detected interest points and regions. Besides, the length of the feature vector is large. This led to augment the computational complexity of the image matching. Hence, we implemented a popular technique, Bag-of- Visual-Words (BoVW), to give a compact representation of image features.

The rest of this paper is organized as follows: Section 2 describes the Related Work. Feature Extraction explained in more details in Section 3. Section 4 introduces our CBIR system based on global and local features. The similarity measures for the proposed features are described in Section 5. Also, Experimental setup and results are described in Section 6. Finally, the conclusion and future works are discussed in Section 7.

2. Related Work

The subject of image retrieval is discussed intensively in the literature. The success of using BOVW model had also contributed to increasing the number of researchers and studies.

For example, Thomee [13] proposed TOP-SURF which is an image descriptor that combines interest points with visual words, and offers the flexibility to vary descriptor size. This is open source software which provides the freedom to modify and re-distribute the code. Additionally,

They provide a high level API and very large pre-computed.

Zhang et al. [14] proposed a bag of images for CBIR schemes. They supposed that the image collection composed of image bags rather than independent individual images.

They contain some relevant images that have same perceptual meaning. The image bags were built before image retrieval.

Giveki et al. [16] offered two methods for implementing SIFT features in CBIR. These methods are based on applying k-means clustering on the extracted SIFT feature matrix and are aimed to minimize the SIFT feature matrix dimension.

Liu [17] reviewed BoVW model in image retrieval system. He provided details about BoVW model and explained different building strategies based on this model. First, he presented several procedures that can be taken in BOVW model. Then, he explained some popular key point detectors and descriptors. Finally, he looked at strategies and libraries to generating vocabulary and do the search.

Alfanindya et al. [18] presented a method for CBIR by using SURF with BoVW. First, they used SURF to computed interest points and descriptors. Then, they created a visual dictionary for each group in the COREL database. They concluded from their experiments that their method outperforms some other methods in terms of accuracy. The major challenge in their work was that the proposed method is highly supervised. It means that they n need to determine the number of groups before they perform classification.

Karakasis et al. [19] proposed an image retrieval structure that relies on utilizing image affine moment invariants as descriptors of salient image patches. BoVW concept is used for indexing and retrieval. Authors considered three setup designs in their experimental study. First, color affine moment invariants are computed. Second, the invariant moments are computed over all chromatic ties of the original image, whereas in the third design a normalization method is performed.

The primary aim of this paper is to design a system for image retrieval based on local feature descriptors and global feature descriptors. Most of the previous image retrieval using BoVW systems used only one local descriptor. Whereas, our proposed system uses both global and local feature descriptors.

3. Feature Extraction

3.1 Speeded Up Robust Features (SURF)

Herbert Bay et.al.[12] first introduced the SURF algorithm as a novel scale- and rotation-invariant interest point detector and descriptor. SURF produces a set of interest points for each image and a set of 64-dimensional descriptors for each interest points. To detect interest points, SURF algorithm is based on the Hessian Matrix, but uses a very basic accurate approximation of Hessian determinant using the Difference-of-Gaussian (DoG). The descriptor uses a distribution of Haar-wavelet responses around the interest point's neighborhood. SURF algorithm is very similar to SIFT algorithm [20] they are both an Interest Point detector and descriptors as image features. In SIFT, these features are identified by using a staged filtering approach. The first stage identifies key locations in scale space by looking for locations that are maxima or minima of a DoG function. Each point is used to generate a feature vector that describes the local image region sampled relative to its scale-space coordinate frame. The major difference between SIFT and SURF is that, in the implementation of scale-space, SIFT typically implemented image pyramid where the input image is iteratively convolved with Gaussian kernel and repeatedly sub-sampled [21]; while SURF created scale-space by applying kernels of increasing size to the original image. Another difference is that SURF descriptor has 128 dimensions while SURF descriptor only has 64 dimensions.

Some comparison papers such as [22] and [23] have stated that SURF outperforms SIFT in terms of result and computational time, thus we chose SURF instead of SIFT as our feature extractor.

The major computational steps of SURF algorithm is as follows:

Step 1: Interest Point Detection.

In order to extract features quickly based on SURF algorithm. The SURF feature detector is based on the Hessian matrix. The determinant of the Hessian matrix is used to determine the location and scale of the descriptor.

The Hessian matrix is defined as H(x, σ) for a given point X = (x, y) in an image as follows:

$$H(X,\sigma)=\begin{bmatrix} Lxx(X,\sigma) & Lx,y(X,\sigma) \\ Lx,y(X,\sigma) & Ly,y(X,\sigma) \end{bmatrix} \tag{1}$$

Where Lx (x, σ) is the convolution of the Gaussian second order derivative $\frac{\partial^2}{\partial x^2}g(\partial)$ with the image I in point x and similarly for Lxy (x, σ) and Lyy (x, σ). The SURF approximates second order derivatives of the Gaussian with box filters. Image convolutions with these box filters can be computed rapidly by using integral images.

The determinant of the Hessian matrix is written as:

$$Det(Happrox)=D_{XX}D_{yy}-(w\ D_{xy})^2 \tag{2}$$

Where, w ≈ 0.9. The relative weight w of the filter responses is used to balance the expression for the Hessian’s determinant.

Step 2: Interest Point Description

The SURF descriptor depicts the distribution of the intensity content within the interest point neighborhood. The first order is assigning an orientation based on the information of a circular region around the detected interest points. The orientation is computed using Haar-wavelet responses in horizontal and vertical directions are called dx and dy, respectively. Once the Haar-wavelet responses are computed and they are weighted with a Gaussian with σ = 3.3s centered at the interest points. In a next step the dominant orientation is estimated by summing the horizontal and vertical wavelet responses within a rotating wedge which covering an angle π/3 in the wavelet response space. The resulting maximum is then chosen to describe the orientation of the interest point descriptor.

In a second order, the region is split up regularly into smaller 4 x 4 square sub-regions. For each sub-region, the Haar wavelet responses are computed at 5x5 regularly spaced sample points.

The horizontal and vertical wavelet responses are summed up over each sub-region to form a first set of entries to the feature vector. The responses of the Haar-wavelets are weighted with a Gaussian centered at the interest point in order to increase robustness to geometric deformations and the wavelet responses dx and dy are summed up over each sub-region. Moreover, the absolute values |dy| and |dy|are summed in order to obtain information about the polarity of the image intensity changes. Therefore each sub-region has a four-dimensional descriptor vector

$$V = (\Sigma\, d_x, \Sigma\, d_y, \Sigma|d_x|, \Sigma|d_y|) \quad (3)$$

The resulting descriptor vector for all 4 by 4 sub-regions is of length 64.

Figure 1 shows an example of SURF Interest points in image number 460 (Dinosaur) from COREL 1000 database:

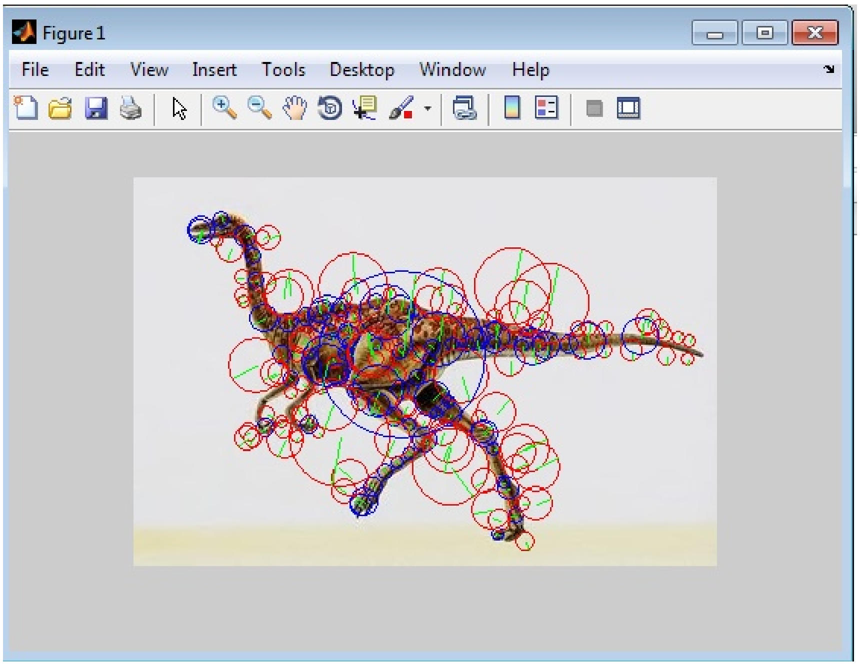


Figure 1. Example of SURF Interest Point

3.2 Color Features

The image retrieval problem is motivated by the need to search the exponentially increasing space of image and video databases efficiently and effectively. The visual content of an image is analyzed in terms of low-level features extracted from the image.

If the color pattern is unique compared with the rest of the data set then the color histogram serves as an efficacious representation of the color content of an image [24].

The color histogram is easy to compute as well as effective in characterizing both the global and local distribution of color in an image [25]. There are many color models like RGB, GLHS, HSV and HIS used to represent the color features. RGB space, three-based color (r, g ,b) presents not only color but also luminance[26]. GLHS space as mapping space because this color model is robust against different types of skin. GLSH' color space has a non-linear relationship with the RGB space, so it is a complex process from RGB color space to GLHS' color space [26]. HSV color space is more similar to human vision hence this colorimetric approach is used in this work. The image in general is in RGB color space is converted into HSV color space.

In this proposed scheme the color features are extracted using Color Histogram in HSV color space. The steps to find the color feature are as follows.

Step 1. Convert the images from RGB space to HSV space.

Step 2. The HSV space the H is quantized into 16 bins, S is quantized into 2 bins and V is quantized into 2 bins. Thus the quantized results are coded as in

$$C = 16\times H + 2\times S + 2\times V \tag{4}$$

Step 3. Count each feature value.

Step 4. The similarity between color feature values of query image and that of the database image is calculated using Euclidean distance similarity measure given in (5).

$$D(Q,T) = (\sum_{i=0}^{l-1}(\mathrm{Q}i - \mathrm{T}i)^2)^{(1/2)} \tag{5}$$

3.3 Bag-of-Visual Word Model

The BoVW model is one of the most widely used ways that represents images as a collection of local features. These local features are typically grouped of local descriptors. The total number of local descriptors that is extracted for each image may be enormous [27].

BoVW was proposed as an approach to tackling this issue by quantizing descriptors into "visual words," which decreases the descriptors' sum drastically. BoVW makes the descriptor more robust to change. This model is very close to the traditional description of texts in information retrieval, but it is considered for images retrieval [28], [29].

BoVW is the de facto standard of image features for retrieval and recognition.

Generally, the BoVW consists of 3 main steps:

1. Automatically extract the interest points and descriptor from the images.
   - ✓ BoVW is usually defining the training dataset as S including images represented by S = s1, s2,…, S_n, where S is the extracted interest points
2. Quantize the interest points and descriptors to form the visual dictionary.
   - ✓ used clustering algorithm like K-Means, which is based on a fixed number to visual words K represented by k = k1, k2,..., k_n, where K is the cluster number
3. Find the occurrences of each visual words in the image in order to build the BoVW histogram.
   - ✓ the data is summarized in a K×N occurrence table of counts $N_{ij} = n(k_i, s_j)$, where $n(k_i, s_j)$ denotes how often the word k_i is occurred in an image s_j

4. Proposed Method

The general methodology used in this proposed system is shown in figure 2.We proposed a system for image retrieval using a combination of local and global features.

For local features extraction using BoVW model and global features extraction using HSV color space. The system uses SURF techniques to extract key points and compute the descriptor for those key points. To apply the Bag of Visual Words model, a clustering algorithm (e.g. K-Means) is applied on the visual descriptions of regions of interest, and the each resulted centroid represent a visual word. The set of visual words are called visual vocabulary. Then, images are viewed as bags of visual words represented as histograms. The system for Color features of images are extracted by using HSV color space. When user supplies the query image through user interface.

First of all query image is preprocessed to increase the contrast of image and then feature extraction is applied. HSV color space is used as a global descriptor and SURF is used as a local descriptor. After applying these techniques, a feature vector of global and local features is created. A final feature vector is created by combining the global feature vector and local feature vector. By using this feature vector, histogram is generated and then similarity matching is performed to retrieve the result from database

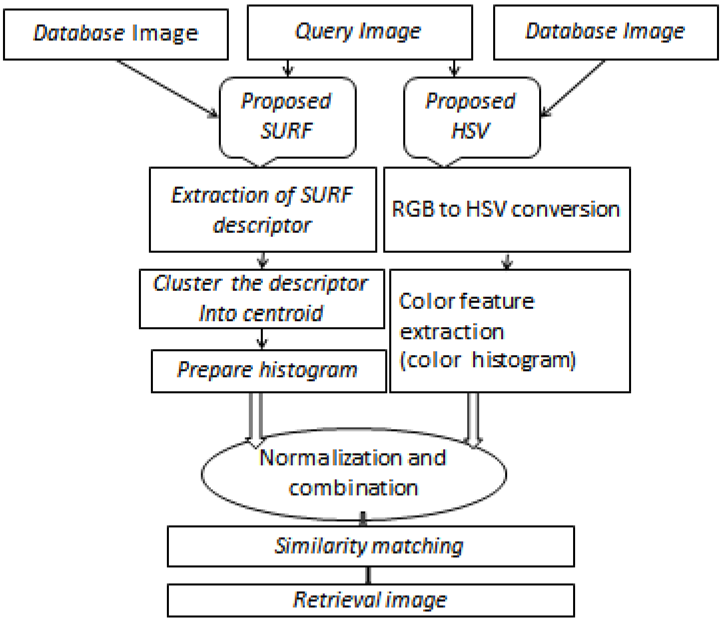


Fig2. The Proposed CBIR system

5. Similarity Matching

In CBIR systems, retrieval accuracy and recall depends on both the performance of feature descriptor and the choice of similarity measures. So it is also a key step to select an appropriate similarity measure for proposed system in image retrieval. In our proposed system, we compare several common similarity measures such as the Euclidean distance, the L1 distance, the Canberra distance, the weighted L1 distance, their similarity measures can be represented as:

$$L1distance\ :E(Q,T)=\sum_{i=1}^{n}\left|(Qi-Ti)\right| \tag{6}$$

Weighted L1 distance:

$$E(Q,T)=\sum_{i=1}^{n}\frac{|Qi-Ti|}{(w+Qi+Ti)} \tag{7}$$

Canberra distance

$$E(Q,T)=\sum_{i=1}^{n}\frac{|(Qi-Ti)|}{(Qi+Ti)} \tag{8}$$

Euclidean distance

$$E(Q,T)=\sum_{i=1}^{n}\sqrt[2]{(Qi-Ti)}\,^\wedge 2 \tag{9}$$

To check the performance of proposed technique the precision and accuracy of the retrieval and classification performance, we used the confusion matrix. Confusion matrix is a table used to evaluate the performance of machine learning classifier during supervised learning. From the confusion matrix we could calculate the accuracy of the classification using the Equation(9) And we could calculate the precision of the classification using the Equation(10)

$$Accuracy=\frac{\text{True postive}}{\begin{array}{c}True\ Positives+False\ Negatives+\\ True\ Negatives+False\ Positivese\end{array}} \tag{10}$$

$$Precision=\frac{\text{True postive}}{\text{True Positives}+\text{False Negatives}} \tag{11}$$

The Figure 3 is a graph showing the precision obtained by proposed system with different similarity measures. It can be seen that the Weighted L1 distance performs better than other similarity measures. Euclidean distance is one of the most commonly used similarity measures, but not always the best one because the distances put too much emphasis on features that are greatly dissimilar. Both weighted L1 distance and Canberra distance can be considered as a weighted L1 distance with different weights. So we choose weighted L1distance as similarity measure for our proposed system in this paper. In our experiment, L1 distance weighted chosen is w= 0.15.

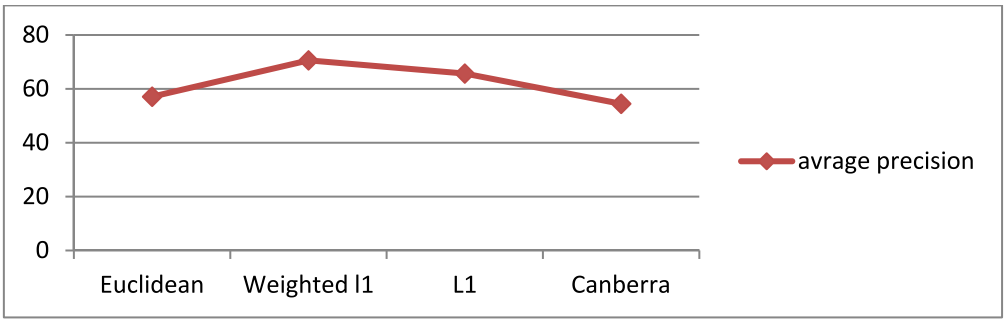


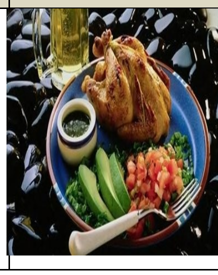
Figure 3 .The comparison of average precision among different similarity measure

6. Experiment and Results

We presented approach for image retrieval those combination local and global features.

The system uses SURF (local features) techniques to extract key points and compute the descriptor for those key points. K-Means algorithm is used to obtain the visual vocabulary. Color features (global features) of images are extracted by using HSV color space. We tested our program with the highly diverse COREL1000 database [30]. It consists of 1000 images in which they are divided into 10 classes consisting of 100 images for each class. The classes are highly diverse, which consists of the classes: African People, Beaches, Buildings, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains, and Food. Table 1: Shows a sample database of 10 images by randomly selecting one image from each category of the COREL-1000 database.

Table 1. Example of Images from each class

African People	Beaches	Buildings	Buses	Dinosaurs
				
Elephants	Flowers	Horses	Mountains	Food
				

For a query image, each extracted descriptor is mapped into its nearest cluster centroid. It will extract the interest points from each image with its respective 64 dimensions descriptors .The interest points will then clustered into k clusters using k-means algorithm, using Euclidean distance, with respect to their descriptors. For this experiment, we choose k = 100. We chose k = 100 because from our experiment, k = 100 have the best accuracy, precision, and computational time ratio. We could see the comparison of different k in term of accuracy and precision in Figure 4. The computational time increases significantly every time the value of k increases. We could see from the result explained using k = 100, our method still outperforms the other methods.

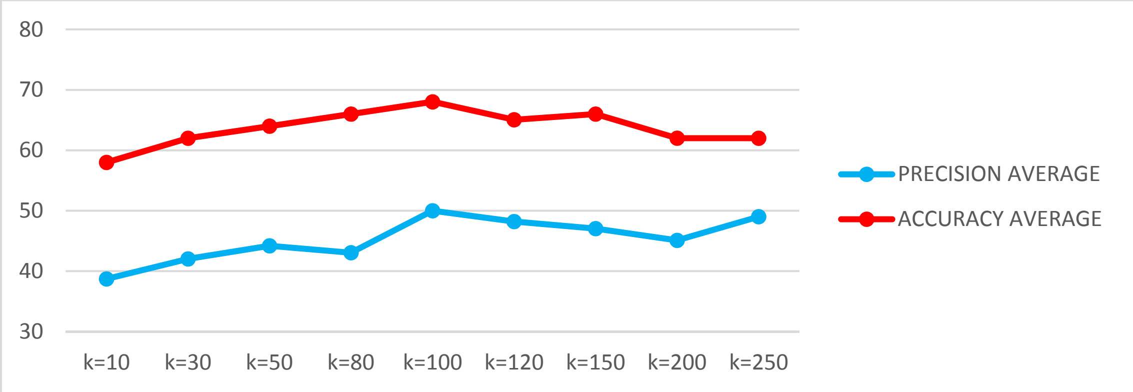


Figure 4. K Values Comparison Graph

In order to explain the well performance of proposed technique, we give the result of the experiments by the same platform and the same image database. To check the performance of proposed technique the precision is used .Table 2 lists the average precision for each image class using proposed technique .gives number of total relevant images in the set of first 20 retrieved images for all 10 categories. The percentage of Precision for all categories of the resultant images are shown in Table3.The Average Precision for all 10 categories is 71.45 %.

Table 2 .average precision for each image class using proposed technique

Image Categories	**Vocabulary size of proposed technique**		
	K=50	**K=100**	**K=150**
African People	72.40	75.20	74.65
Beaches	28.25	35.20	30.45
Buildings	62.30	68.40	64.50
Buses	78.10	82.60	80.70
Dinosaurs	99.80	99.95	99
Elephants	53.00	56.40	56.95
Flowers	83.30	85.80	86.05
Horses	89.25	91.80	89.45
Mountains	37.60	43.60	42.75
Food	73.85	75.60	74.19
average precision	**67.78**	**71.45**	**69.86**

In the retrieval process, we input a query image, should have 20 minimum distances, representing the distance of the query image to database. We will then choose the smallest minimum distance and classify the query image to the database with the smallest distance to the query image. We can obtain a ranked set of most similar images based on the Weighted L1 distance. The sample result for image retrieval by proposed method, Color HSV and SURF features are shown in Fig 5, 6 and 7 respectively. The image at the top of left corner is the query image and other 20 images are the retrieval results. It shows that our proposed descriptor is more outperforms than the local features alone and global features alone.

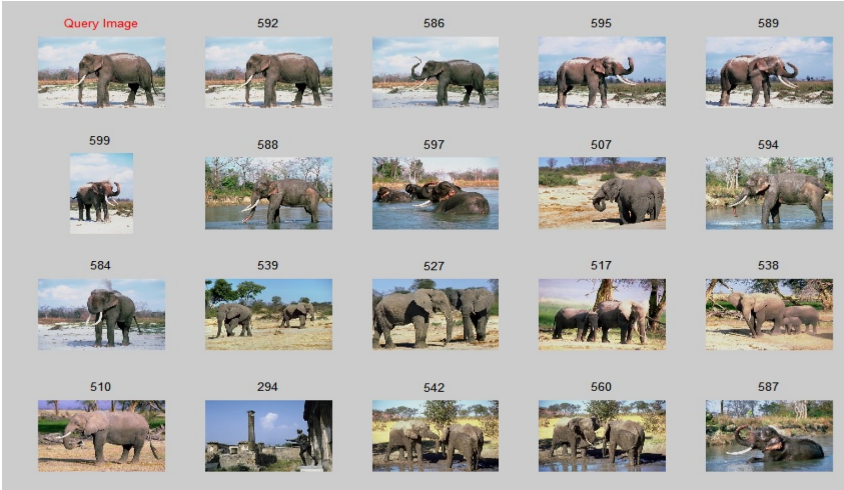


Figure 5 .Image retrieval using proposed technique

Figure 6.Image retrieval using Color HSV

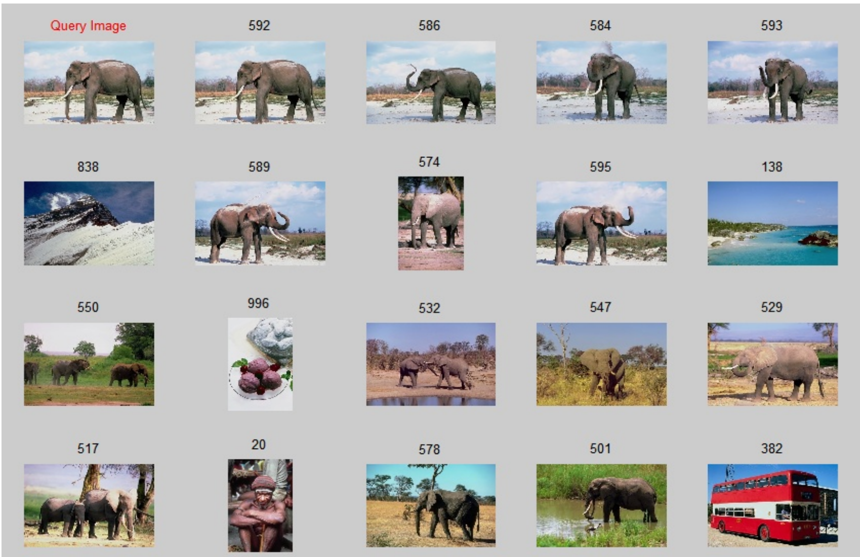


Figure 7.Image retrieval using SURF Features

To evaluate the performance of the proposed method, it has been compared with traditional CBIR systems such as co-occurrence matrix-based CBIR[31] , Image retrieval by wavelet and color features[32] and structure elements descriptor(SED)[33] From Corel database, all images have been used as query images (the tested 10 semantic class includes African People, Beaches, Buildings, Buses ,Dinosaurs, Elephants, Flowers, Horses, Mountains, and Food) and then the first 20 most similar images are retrieved. For each class of image, the both average normal precision and recall are computed for all 100 query images in each class. Results for all image categories is show in Table 3.

Table 3.precision and recall for proposed and other methods

Image categories	**Proposed CBIR**		**co-occurrence matrix-based CBIR[31]**		**wavelet and color features[32]**		**structure elements descriptor(SED)[33]**	
	Precision	**recall**	**Precision**	**recall**	**Precision**	**recall**	**Precision**	**recall**
African People	75.20	15.04	45.30	11.50	44.80	10.90	75.10	12.90
Beaches	35.20	6.64	39.80	12.10	47.20	11.90	25.20	3.40
Buildings	68.40	13.28	37.40	12.70	53.40	13.00	45.30	9.70
Buses	82.60	16.52	74.10	9.20	73.40	11.80	69.33	10.60
Dinosaurs	99.95	19.99	91.50	7.2	99.80	19.90	90.10	18.10
Elephants	56.40	11.08	30.04	13.20	56.80	13.40	50.20	10.60
Flowers	85.80	17.16	85.20	8.7	87.50	17.50	67.50	11.30
Horses	91.80	18.16	56.80	10.20	70.70	12.10	57.30	11.20
Mountains	43.60	8.32	29.30	13.50	39.30	16.60	35.20	8.20
Food	75.60	14.92	36.90	12.90	61.00	12.30	59.17	12.10
precision average	**71.45**	**14.11**	**52.93**	**11.10**	**63.39**	**13.90**	**57.44**	**10.81**

Our method's precision average comes down to 71.45% and it's higher than co-occurrence matrix-based CBIR the precision average (52.93%), structure elements descriptor(SED) (57.44%), Image retrieval by wavelet and color features (63.39%). The table comparisons show that our method outperforms the other methods in term of accuracy and precision. We observed that the chosen feature extractor plays a big role in determining the precision and recall of the method. Our chosen method to extract image features, SURF, has been proven to be superior to other feature extractor used in other methods.

7. Conclusion and Future Work

In this paper, we presented a new approach for image retrieval that combination works in SURF (local feature), color (global features). The new approach extracts image salient points to be combination with the proposed global features. The experimental results showed that this combination provide more accurate results than other compared techniques. The combination of individual features improves the retrieval accuracy when compared to individual features used alone.

In the future work we plan to improve the efficiency of the search in term of accuracy by Grouped BoVW is different with the normal BoVW. The normal BoVW only has 1 global dictionary and GBOVW has a dictionary for each group or class in test database, which make in method more discriminative and results in higher accuracy.

Reference

[1] Yong-Hwan Lee, Hyochang Ahn and Sang-Burm Rhee “Efficient Image Retrieval using Advanced Clustering SURF,“15th International Conference on Network-Based Information Systems 2012 .

[2] N. Sebe, Q. Tian, E. Loupias, M.S. Lew and T.S. Huang, “Salient Points for Content BasedRetrieval”.

[3] Atefe Aghaei , Sajjad Tavassoli ,” Human Action Recognition Based Sparse Representation “ journal of Advances in Computer Research ,Pages: 117-130 ,2016

[4] Alfanindya, A., Hashim, N., & Eswaran, C. “Content-Based Image Retrieval andC lassification using speeded-up robust features (SURF) and grouped bag-of-visual-words (GBoVW) “. In Technology, Informatics, Management, Engineering, and Environment (TIME-E), 2013 International Conference on (pp. 77-82). IEEE. 2013.

[5] Yasmin, M., M. Sharif and S. Mohsin. Use of low-level features for content-based image retrieval: Survey. Res. J. Recent Sci., 2: 65-75. , 2013

[6] Ashish Mumar, “Image Retrieval using SURF Features”, Master Thesis, Thapar University, 2011.

[7] Kodituwakku, S.R. and S. Selvarajah,. “Comparison of color features for image retrieval”. Ind. J. Comput. Sci. Eng., 1: 207-211. 2010

[8] Barcellos, R; Oliani , R. S.; Lorenzi, L. T.; Gonzaga, A. "Content Based Image Retrieval Using Color Autocorrelograms in HSV Color Space", In Proceedings Of Xviii Brazilian Symposium on Computer Graphics and Image Processing - Sibgrapi ,2005.

[9] Ji-quan Ma, "Content-Based Image Retrieval with HSV Color Space and Texture Features," International Conference on Web Information Systems and Mining, 2009

[10] Manuel Cedillo-Hernandez ,et al “Mexican Archaeological Image Retrieval Based on Object Matching and a Local Descriptor”, (ICCCI -2015), Jan. 08 – 10, 2015

[11] Herbert Bay, Tinne Tuytelaars, Luc Van Gool, "Speeded-Up Robust Features," Computer vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346-359, EECV, 2008.

[12] Bart Thomee , Erwin M. Bakker and Michael S. Lew, “TOP-SURF: a Visual Words Toolkit”

[13] Zhang, J., & Ye, L. “Image retrieval based on a bag of images“. In Image Processing (ICIP), 16th IEEE International Conference on (pp. 1865-1868). IEEE. 2009.

[14] Giveki, D., M.A. Soltanshahi, F. Shiri and H. Tarrah, b. A new SIFT-based image descriptor applicable for content-based image retrieval.J.Comput.Commun., , 2015

[15] Liu, J. “Image retrieval based on bag-of-words model “. Cornell University Library:1304.5168, 2013.

[16] Alfanindya, A., Hashim, N., & Eswaran, C. “Content-Based Image Retrieval and Classification using speeded-up robust features (SURF) and grouped bag-of-visual-words (GBoVW) “. In Technology, Informatics, Management, Engineering, and Environment (TIME-E), 2013 International Conference on (pp. 77-82). IEEE. 2013.

[17] Karakasis, E.G., A. Amanatiadis, A. Gasteratos and S.A. Chatzichristo. Image moment invariants as local features for content-based image retrieval using the bag-of-visual-words model, 2015

[18] David G. Lowe "Object Recognition from Local Scale-Invariant Features," The Proceedings of the Seventh IEEE International Conference on Computer Vision, Vol. 2, pp. 1150-1157, 1999.

[19] Christopher Evans "Notes on the OpenSURF Library," University of Bristol,2009.

[20] Shihua He, Chao Zhang, Pengwei Hao, "Comparative Study of Features for Fingerprint Indexing," 161h IEEE International Conference on Image Processing (ICIP), pp. 2749;2752, November 2009.

[21] Luo Juan, Oubong Gwun, "SURF Applied in Panorama Image Stitching," 2nd International Conference on Image Processing Theory, Tools, and Applications (IPTA), pp. 495;499, July 2010.

[22] Dr. Fuhui Long, Dr. Hongjiang Zang and Prof. David Dagan Feng, "Fundamentals of Content Based Image Retrieval".

[23] JunYue,Zhenbo Li,Lu Liu,Zetian Fu, "Content-based image retrieval using color and texture fused features", Mathematical and Computer Modelling, 2011.

[24] Farima Jafari, Reza Javidan ,"A New Compression Method based on Jpeg2000 and Contourlet Transform", Journal of Advances in Computer Research Iran , Pages: 27-43 ,2015

[25] David G. Lowe "Object Recognition from Local Scale-Invariant Features," The Proceedings of the Seventh IEEE International Conference on Computer Vision, Vol. 2, pp. 1150-1157, 1999.

[26] Csurka G., Dance C., Fan L., Willamowski J., and Bray C. "Visual categorization with bags of keypoints, " Workshop on Statistical learning in compute. Vis., ECCV. Vol. 1. No. pp.1-22. 2004.

[27] Lowe D.G., "Distinctive image features from scale-invariant keypoints, " Int. journal of compute. Vis., vol. 60, pp. 91–110. 2004.

[28] R. Brunelli and O. Mich, "Histograms Analysis for Image Retrieval," Pattern Recognition, ,1625-1637,2001

[29] Jhanwar, N., et al. "Content based image retrieval using motif cooccurrence matrix." Image and Vision Computing , 1211-1220, 2004

[30] A Rashno, et al , "An Efficient content-based image retrieval with ant colony base on wavlet and color features", AISP march 2015

[31] X. Wang, Z. Wang, A novel method for image retrieval based on structure elements' descriptor, J. Vis. Commun. Image Represent. , 63–74. 2013