Analysis techniques of Content based Image Retrieval based on Confusion Matrix

Jaison Singh[1], Harshdeep Trehan[2], Varinderjit Kaur[3], Dr. Naveen Dhillon[4]

Student, CSE Department, RIET, Phagwara, India[1]

Asst. Prof, CSE Department, RIET, Phagwara, India[2]

HOD, CSE Department, RIET, Phagwara, India[3]

Principal, RIET, Phagwara, India[4]

Abstract: In today's epoch of Digital technology, the need of content based image retrieval is increasing day by day. Image retrieval system based on image's visual content becomes more desirable. Hence, content based image retrieval has emerged as one of the most active research areas in the past few years. There are a lot of approaches have been proposed for content based image retrieval. The literature addresses the issues involved in content based image retrieval and presents solution to them. To this end, an interface is designed for content based image retrieval and in order to key out best approach a comparative study is performed

Keyword: Content Based Image Retrieval, Fuzzy Semantic Relevance Matrix, Histogram of Oriented Gradients, Local Binary Pattern, Negative Predictive Value, Positive Predictive Value.

I. INTRODUCTION

In recent emerging years, digital images have become a remarkable array in each and every field. With the advancement in internet and multimedia technologies, a huge amount of images has been used in many fields like medical treatment, satellite data, still images repositories, digital forensics. This has created an ongoing demand of systems that can store and retrieve images in an effective way. Many multimedia information storage and retrieval systems have been developed till now for catering these demands.

The one of the intuitive and user friendly way of retrieving images is Content Based Image Retrieval system. A CBIR system uses visual contents of the images described in the form of low level features such as color, texture and shape to represent digital images [1].CBIR system basically consists of two parts i.e. feature extraction and similarity matching. In CBIR many techniques [2] [3] [4] [5] [6] are used to evaluate different impact on the results of retrieving images. The retrieval results should be robust to find out the similar images irrespective of their features color, shape and texture. In earlier findings different CBIR techniques evaluates different results for different database on the basis of various parameters.

II. ENVIRONMENT

To evaluate the performance of CBIR algorithms, a database of five classes tribal, beach, building, bus and dinosaur has been taken from WANG dataset [7]. The implementation of the system is done in MATLAB® 2015a. The results of algorithms under study are evaluated on texture and shape analysis with fuzzy semantic relevance matrix. For HOG based CBIR, we will work upon CBIR based on LBP-HOG and Gradient field HOG descriptor based techniques under this performance evaluation survey.

III. EXPERIMENT

There are different approaches of content based image retrieval process. The approaches are categorized into two parts, first is CBIR based on relevance feedback [8] [9] and second is CBIR based on HOG [10] [12] features. These approaches have been studied in detail as follows:-

A. CBIR-Relevance Feedback
The relevance feedback technique in CBIR is based on user interaction with the retrieval system. This technique is designed to bridge the semantic gap between low level and high level features. The basic idea of relevance feedback is to shift the burden of finding the right query formulation from the user to the system [8]. To retrieve the image from the database, we first extract feature vectors from images (the features can be shape, color, texture etc), then store feature vectors into another database for future use.

When given query image, we similarly extract its feature vectors, and match those features with database image features. If the distance between two images feature vectors is small enough; we consider the corresponding image in the database similar to the query. The search is usually based on similarity rather than on exact match, and the retrieval results are given to the user.
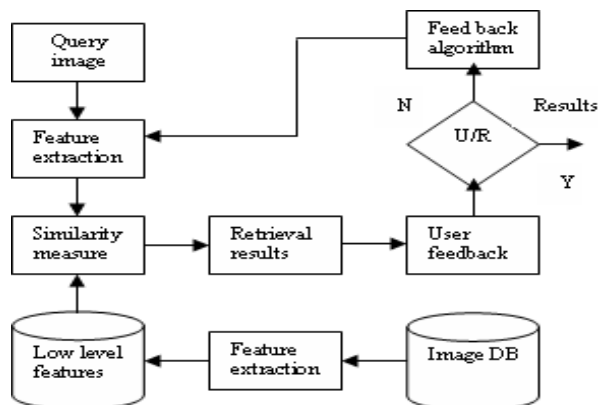
Fig 1 Architecture of Relevance Feedback Model

B. CBIR-HOG

Histogram of Oriented Gradients (HOG) is the basic idea behind the HOG descriptor is that local object appearance and shape within an image can be described by distribution of intensity gradients or edge detections. The HOG is window based descriptor which is divided into rectangular shaped N*N grid cells. The frequency of each grid cell is computed and represented in form of distributed histogram.

1) Working of Hog descriptor

The working of HOG descriptor can be achieved by dividing an image into small connected regions which are called as cells and for each cell creating a histogram for gradient direction or edge orientations for the pixels within the cell. The combination of these histograms then represents the descriptor. The working of HOG descriptor is explained by following steps:

2) Gradient Computation

The magnitude of the gradient is $|G| = \sqrt{I^2 x + I^2 y}$

The orientation of Gradient is given by $\theta = \arctan \frac{I_y}{I_x}$

where I_x and I_y are derivatives of an image I.

3) Block Normalization :

Let v be the non- normalized vector containing all histograms in a given block. There are three different methods of block normalization which are as follows:-

$$\text{L2-norm: } f = \frac{v}{\sqrt{\|v\|^2 + e^2}}$$

$$\text{L1-norm: } f = \frac{v}{\|v\|_1 + e}$$

$$\text{L1-sqrt: } f = \sqrt{\frac{v}{\|v\|_1 + e}}$$

C. CBIR-Local Binary Pattern

Local Binary Pattern is a feature descriptor like HOG. It has been used for texture classification [14]. The LBP has been widely used in different applications such as human detection, face recognition. The performance of HOG descriptor is very poor in case of image backgrounds with noisy edges. But LBP improves the performance by integrating with HOG descriptor. The value of LBP is computed by calculating the gray level value of the center pixel and its neighbors in the give image. The following equation is used for computation of LBP descriptor.

$$LBP_{P,R} = p = 1P2 \ (p-1) \ f1 \ (g_{p} - g_c)$$

$f1(x) = 1, x \geq 00$, otherwise

where P is no of neighbors and R is center of the neighborhood.

g_c is gray value of center pixel and g_p is gray value of its neighbors.

D. CBIR-FSRM

(FSRM), is constructed to describe the semantically relevance between the images in the database. The weights in the FSRM are adjusted according to user's feedback in each feedback and the FSRM are modified by learning more time. The algorithm does not need a priori knowledge of specific problem because it based on FSRM. A fuzzy set is a class of objects with a continuum of grades of membership. Fuzzy sets characteristic a set 0, 1 expand to the interval [0, 1]. Therefore, we use the value of interval [0, 1] represents the "grade of membership" of an object of the concept

IV. PERFORMANCE PARAMETERS

The parameters selection for evaluating effectiveness of the approaches under study is based on default evaluation criterion. The following performance metrics are used for measuring the efficiency of the system.

- True positive (TP) are images correctly classified as belonging to the class.
- True negative (TN) are images correctly classified as not belonging to the class.
- False positive (FP) or (Type-II error) are images incorrectly classified as belonging to the class.
- False negative (FP) or (Type-I error) are images incorrectly classified as not belonging to the class.

A. Predicitive Value

Precision or Positive Predictive Value (PPV) measures the ratio of relevant output instances to the total instances obtained from output.

$$\text{Precision} = \frac{TP}{TP + FP} \qquad \text{----Equation (1)}$$

B. Sensitivity

Sensitivity or recall is ratio of relevant output instances to the total instances. It can be represented as in equation....

$$\text{Recall} = \frac{TP}{TP + FN} \qquad \text{---- Equation (2)}$$

C. Negative Predictive value

Negative predictive value is defined as the ratio of number correctly labeled negative images to the total number of negatively labeled images.

$$\text{Negative Predictive value} = \frac{TN}{TN+FN} \quad \text{---- Equation (3)}$$

D. Specificity

Specificity is the ratio of total number of correctly labeled negative images to the total number of negative images.

$$\text{Specificity} = \frac{TN}{TN+FP} \quad \text{---- Equation (4)}$$

E. Accuracy

The accuracy is the proportion of the test results that is true positive and false negatives among total number of cases.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad \text{---- Equation (5)}$$

F. Overall Accuracy

The accuracy assessment provides us with the overall accuracy of the confusion matrix. Overall accuracy sums up the total correctness of image retrieval mechanism of all the images for which equation 2 and 3 has been used

$$\omega = \sum_{i=1}^{nc} e_{ii} / NT \qquad \text{--Equation(6)}$$

$$NT = \sum_{i=1}^{nc} \sum_{j=1}^{nc} e_{ij} \qquad \text{--Equation(7)}$$

Where, ω: denotes the overall accuracy to be calculated, NT: is the sum of all the off diagonal elements in the confusion matrix, e_{ii} is the sum of total numbers of correct cells along the major diagonal of confusion matrix, and nc: is total number of columns.

G. Kappa Analysis

Kappa coefficient is generally used as the measure of inter-rater reliability as overall accuracy is usually calculated along the major diagonal in confusion matrix but kappa takes all columns and rows in account while calculating [29]. We have used the kappa $\hat{k}$ to find the overall reliability of CBIR mechanism and the results were analyzed on the standard of agreement for the kappa purposed by Landis and Koch (1977) (≤ 0 = poor, .01-.20 = slight, .21 - .40 = fair, .41 - .60 = moderate, .61 - .80 = substantial, .81 – 1 = almost perfect) [30]. The equation used to estimate the weighted kappa is:

$$\hat{k} = N \sum_{i=1}^{nc} X_{ii} - \sum_{i=1}^{nc} (X_{i+} \times X_{+i})/N^2 - \sum_{i=1}^{nc} (X_{i+} \times X_{+i}) \quad (3.8)$$

Where, N: is total number of cell in matrix, nc: is total number of columns in confusion matrix, i+: is sum of column i, +i: is sum of row i, and X_{ii}: is total number of correct cells in matrix.

IV. RESULT OF THE INDIVIDUAL CLASS OF EACH ALGORITHM

The results of individual class i.e. tribe, beach, building, bus and dinosaur have been discussed in this section. In confusion matrix, the columns represented as five different categories and rows represented as retrieved images corresponding to their categories. Each non-diagonal element in matrix represents the number of non-similar image of representative category and diagonal elements represent the similar images of corresponding category. The performance parameters are calculated on the basis of obtained results and their analysis has also been done for each class.
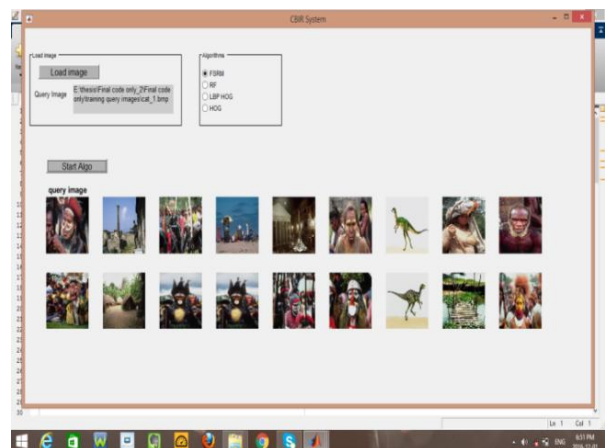
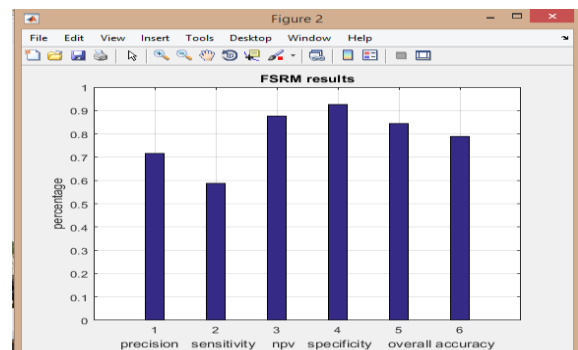### A. CBIR -FSRM

Fig 2 Results of FSRM (tribal)

Fig 3 Results of FSRM (graph)

The results of the FSRM are evaluated by using confusion matrix. In confusion matrix, 10 similar images of first category retrieved and 12, 13, 12 and 12 images of other categories i.e. beach, building, bus, dinosaurs.
The average value of sensitivity or recall is calculated as 69.40% which is fairly good. The average precision is 71.47% which is good but could be better by retrieving more relevant images. The overall accuracy is calculated by using equation 6 which comes out to be 69.41% which

is good but could be improved by getting more true positive values which are calculated diagonally. But the accuracy in this case is 85.46.78 that is quite good. Kappa analysis is computed using equation 8. The kappa statistics comes out to be 0.735, which according to Landis and Koch(1997) standard of agreement is substantial

B. CBIR-Relevance Feedback

The results of the Relevance feedback are evaluated by using confusion matrix. In confusion matrix, 12 similar images of tribal category are retrieved and similarly 13, 14, 13 and 14 similar images retrieved of beach, building, bus and dinosaur categories respectively.

The average value of sensitivity or recall is calculated as 77.64% which is good but less than FSRM sensitivity value. The average precision is 79.18% which is enough good .The overall accuracy is calculated by using equation 6 which comes out to be 70.64% which is good. But the accuracy in this case is 89.80% that is very good. Kappa analysis is computed by using equation 8. The kappa statistics comes out to be 0.732, which according to Landis and Koch(1997) standard of agreement is substantial.
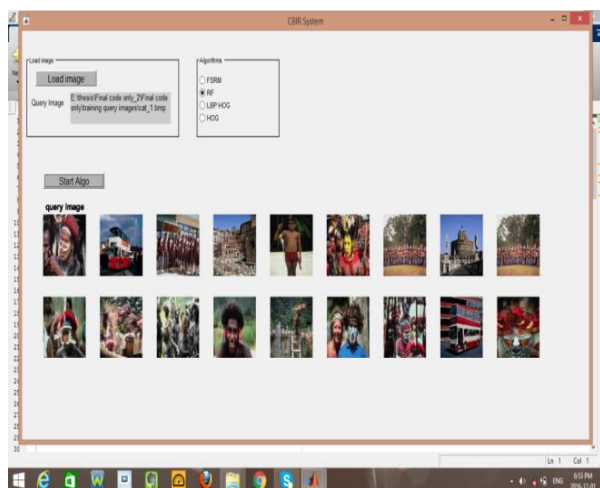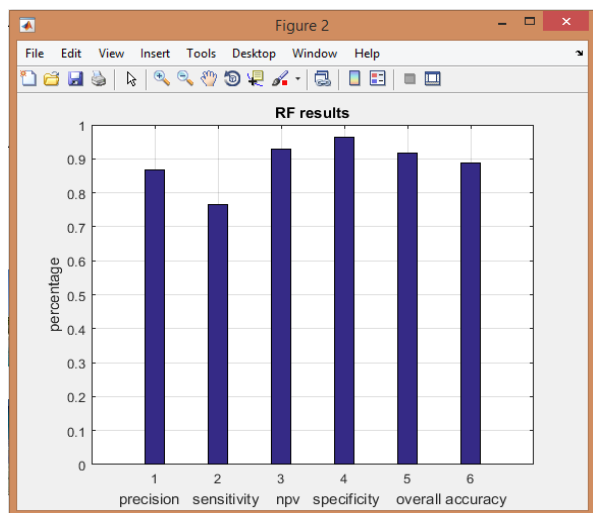

Fig: 4 Results of Relevance Feedback

Fig: 5 Results of RF (graph)
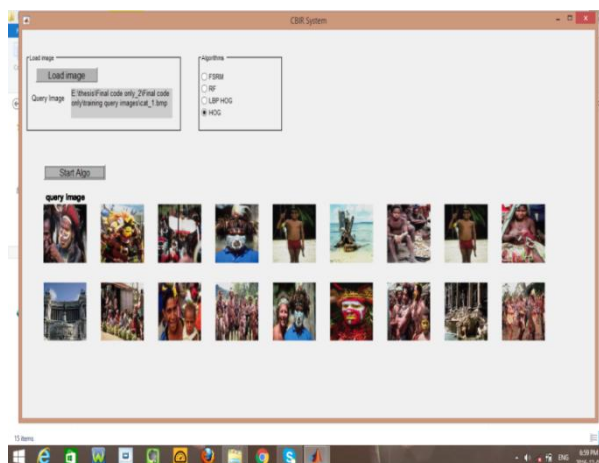
C. CBIR-HOG
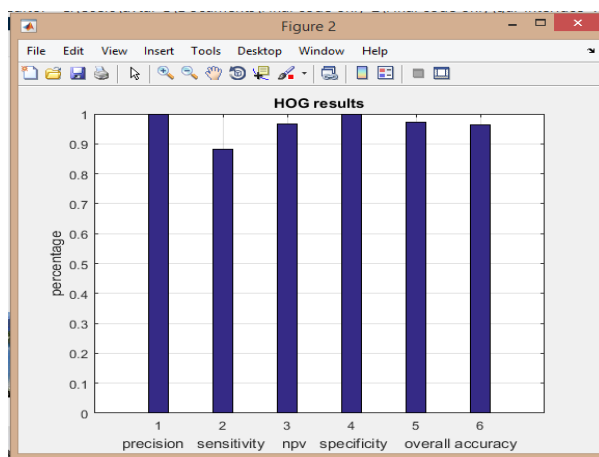

Fig: 6 Results of HOG (tribal class)

Fig: 7 Results of HOG (graph)

The results of CBIR-HOG algorithm are also evaluated by using confusion matrix.14 similar images retrieved in tribal category and 15, 15, 15, 15 similar images retrieved in other categories such as beach, building, bus, dinosaur respectively. This algorithm exhibits higher performance as compared to other algorithms. The confusion matrix The average value of sensitivity or recall is calculated as 87.05% which is better as compare to other algorithms. The average precision is 87.86% which is also good .The overall accuracy is calculated by using equation 6 which comes out to 87.05% which is extremely good and the accuracy is 94.43% which is almost perfect. Kappa analysis is computed by using equation 8. The kappa statistics comes out to be 0.779, which according to Landis and Koch (1997) standard of agreement is substantial.

D. CBIR-LBP HOG

The confusion matrix shows that 12 similar images retrieved from tribal category and 12, 11, 13, 12 images retrieved from beach, building, bus, and dinosaur categories respectively. The performance of LBP-HOG algorithm is also good but it could be improved in case of beach and building.
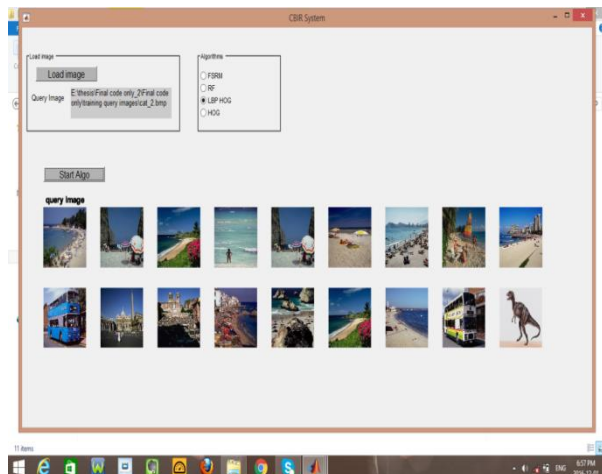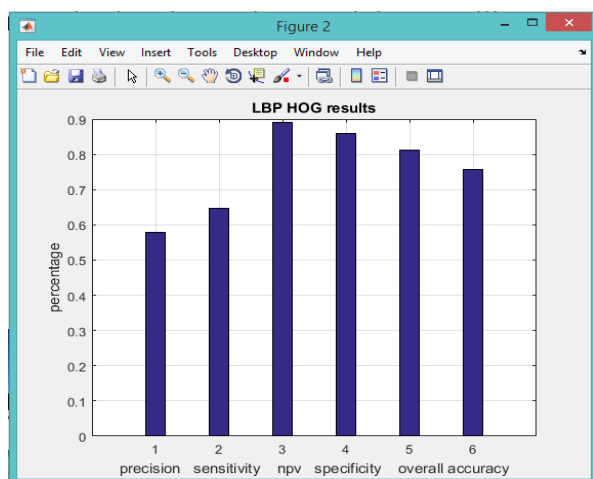
Fig: 8Results of LBP-HOG (beach class)

Fig: 9 Results of LBP-HOG (graph)

The average value of sensitivity or recall is calculated as 70.58% which is fairly good. The average precision is 70.13% which could be better. The overall accuracy is calculated by using equation 3.6 which comes out to 70.58 which is better than relevance feedback algorithm and the accuracy is 94.43%. The kappa statistics comes out to be 0.791, which according to Landis and Koch (1997) standard of agreement is substantial

V. OVERALL RESULT

The combined results have been illustrated in Table 2. Among the four algorithms CBIR-HOG depicts the better performance in terms of all parameters as compared to other three algorithms. According to Kappa analysis all the four algorithms attains substantial performance. The table 2 summarizes the performance of algorithms in terms of Precision, Sensitivity, Specificity, Accuracy, Overall accuracy and Kappa analysis. The average values of precision, sensitivity, specificity and accuracy have been calculated to evaluate the overall performance of the system

The average precision, specificity, sensitivity, NPV and overall accuracy of each algorithm have been illustrated in figure 3 with the help of line graph. The average precision is highest for HOG and while RF holds the lowest value. The sensitivity or recall is also highest in case of HOG. There is very little variation in case of sensitivity and NPV, but HOG attains highest value for sensitivity. The overall accuracy is highest for HOG and LBP-HOG attains less overall accuracy whereas FSRM and RF hold almost same value with little difference.
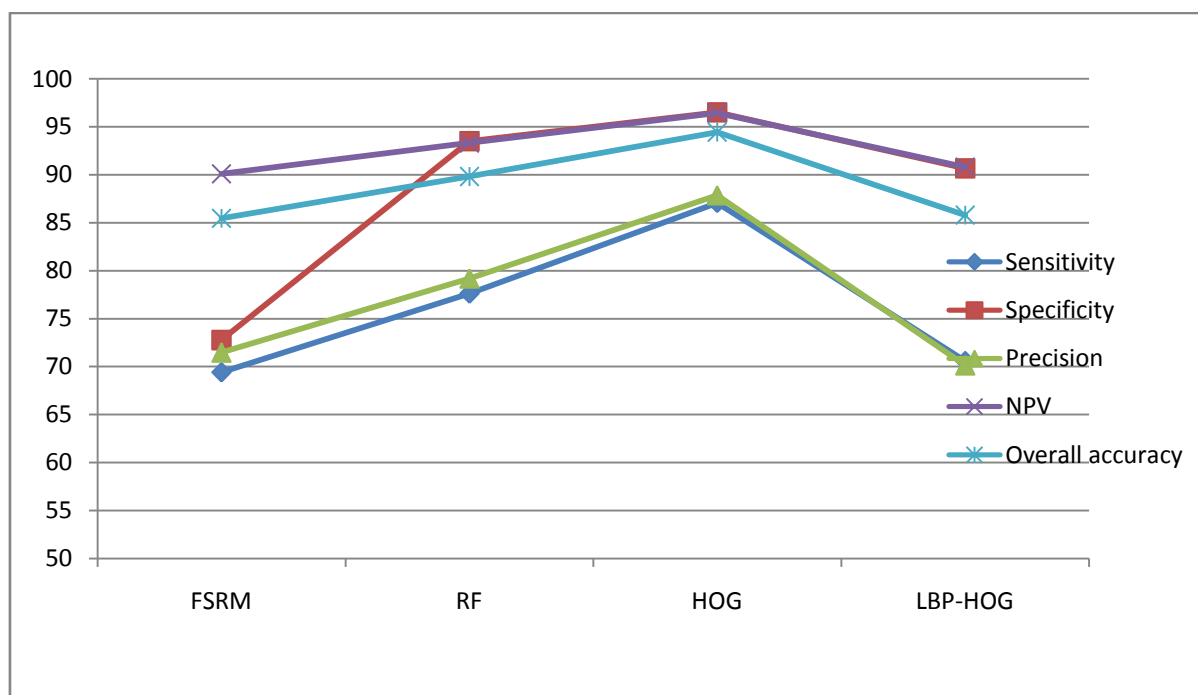

Figure 3 Overall results of all algorithms

Table 2 Overall performance of CBIR algorithms

Algorithm	Sensitivity	Specificity	Precision	Negative Predictive value	Accuracy	Overall Accuracy	Kappa Analysis
FSRM	69.40%	72.77%	71.47%	90.08%	85.46%	69.41%	0.735
RF	77.6%	93.48%	79.18%	93.31%	89.80%	77.64%	0.732
HOG	87.05%	96.49%	87.86%	96.42%	94.43%	87.05%	0.779
LBP-HOG	70.58%	90.66%	70.13%	90.78%	85.78%	85.78%	0.751

VI. CONCLUSIONS

Content based image retrieval is a dominant way of retrieving images. In recent years, many researchers have worked out various techniques for the content based image retrieval. A considerable number of different feature extraction techniques and similarity measures have been proposed. The presented algorithms for CBIR have different impact on the results of image retrieval. These algorithms for content based image retrieval are tested on five classes i.e. tribal, beach, building, bus and dinosaur. The result of the implementation fortified that the HOG and LBP-HOG feature extraction algorithm is more efficient than Relevance Feedback and FSRM algorithms

REFERENCES

[1] S. Shete and M. S. Chavan, "Content Based Image Retrieval : Review," Int. J. Emerg. Technol. Adv. Eng., vol. 2, no. 9, pp. 85–90, 2012.
[2] H. Narasimhan and P. Ramraj, "Contribution-based clustering algorithm for content-based image retrieval," 2010 5th International Conference on Industrial and Information Systems, 2010, pp. 442–447
[3] Y. Gao, M. Wang, Z. J. Zha, J. Shen, X. Li, and X. Wu, "Visual-textual joint relevance learning for tag-based social image search," IEEE Trans. Image Process., vol. 22, no. 1, pp. 363–376, 2013.
[4] H. Aboulmagd, N. El-Gayar, and H. Onsi, "A new approach in content-based image retrieval using fuzzy," Telecommun. Syst., vol. 40, no. 1–2, pp. 55–66, 2008.
[5] R. Tronci, G. Murgia, M. Pili, L. Piras, and G. Giacinto, "ImageHunter: A novel tool for Relevance Feedback in Content Based Image Retrieval," CEUR Workshop Proc., vol. 771, pp. 1–12, 2011.
[6] J. L. James Z. Wang, "http://wang.ist.psu.edu/docs/related/," International
[7] Journal of Advanced Research in Computer and Communication Engineering, 2013.
[8] L. Pinjarkar, M. Sharma, and K. Mehta, "Comparison and Analysis of Content Based Image Retrieval Systems Based On Relevance Feedback," J. Emerg. Trends Comput. Inf. Sci., vol. 3, no. 6, pp. 833–837, 2012.
[9] L. Wang and W. Lin, "A Semantic Subspace Learning Method to Exploit Relevance Feedback Log Data for Image Retrieval," Computational Intelligence and Data Mining (CIDM), 2013 IEEE Symposium, 2013, pp. 178–183.
[10] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in CVPR "05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR"05) - Volume 1, 2005, pp. 886–893.
[11] S. Y. -, X. L. -, and U. B. -, "A Pedestrian Detection Method Based on the HOG-LBP Feature and Gentle AdaBoost," Int. J. Adv. Comput. Technol., vol. 4, no. 19, pp. 553–560, 2012.
[12] R. Hu and J. Collornosse, "A performance evaluation of gradient field HOG descriptor for sketch based image retrieval," Comput. Vis. Image Underst., vol. 117, no. 7, pp. 790–806, 2013.
[13] X. Wang, T. X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," in Computer Vision, 2009 IEEE 12th International Conference on, 2009.
[14] J. Zhang, K. Huang, Y. Yu, and T. Tan, "Boosted local structured HOG-LBP for object localization," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2011, pp. 1393–1400.
[15] H. Sebai and A. Kourgli, "An adaptive CBIR system for remote sensed data," in Content-Based Multimedia Indexing (CBMI), 2014 12th International Workshop, 2014, pp. 1–6.
[16] R. Huitl, G. Schroth, S. Hilsenbeck, F. Schweiger, and E. Steinbach, "Virtual reference view generation for CBIR-based visual pose estimation,"Proceedings of the 20th ACM international conference on Multimedia - MM '12, 2012,pp. 993.
[17] M. M. Rahman, S. K. Antani, and G. R. Thoma, "Biomedical CBIR using 'bag of keypoints' in a modified inverted index," in Proceedings - IEEE Symposium on Computer-Based Medical Systems, 2011, pp. pp. 304–311.
[18] 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
[19] X. Yuan, J. Yu, Z. Qin, and T. Wan, "A SIFT-LBP image retrieval model based on bag of features," in International Conference on Image Processing, 2011, pp. 1061–1064.
[20] Valle and M. Cord, "Advanced techniques in CBIR local descriptors, visual dictionaries and bags of features," Tutorials SIBGRAPI 2009 - 22nd Brazilian Symp. Comput. Graph. Image Process., pp. 72–78, 2009

BIOGRAPHY

Jaison Singh is a student of Computer Engineering Department, Ramgardia Engg. She received Diploma in Computer Engineering in 2014 from MSBTE. Her research interests are IOT, Software Testing, and Cloud Computing etc