QUANTITATIVE ANALYSIS OF CONTENT BASED IMAGE RETRIEVAL USING HOG, LBP AND GABOR FEATURE DESCRIPTORS IN COREL DATASET

Anto.A.Micheal[1], Pradeepthi.K.V[2]
[1,2] Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telengana, India

Abstract

Objectives: The evolution of multimedia technology and rapidly increasing image collections on the Internet has attracted significant research efforts in image retrieval. Difficulties faced by text-based image retrieval motivated the researchers to develop new solutions for representation and indexing of visual information. This paper proposes a Content Based Image retrieval using HOG, LBP and Gabor features retrieve query related images. Method: This proposed approach extracts various features such as Gabor features, Histogram of Orientation Gradient (HOG) and Local Binary Pattern (LBP) from the images and they are formed as feature vectors. The relevance distance between the query image and images in the database are measured using similarity distance metrics known as Euclidean Distance, Mahalanobis Distance and Minkowski-Form Distance. Based on the similarity metrics, the images relevant to the query image are extracted. The proposed approach is experimented on COREL dataset and its performance is evaluated using statistical parameters such as, the precision and recall. Findings: The statistical results show that the precision and recall of the proposed system are 80.06% and 19.26% respectively. Novelty: The comparative study between the proposed content based image retrieval method and other existing methods indicates that proposed method proves its better feature extraction capability than the other image retrieval methods.

Keywords: Content Based Image Retrieval, Corel Dataset, HOG, LBP, Gabor, Similarity Metric

1. Introduction

Nowadays images are broadly utilized because of its visual representation advantage. Due to the rapid advancement of computers and networks, the transmission and storage capacity of ample number of images have become possible. In earlier days, the image retrieval was widely required instead of text retrieval. Content-Based Image Retrieval (CBIR) is a standout amongst the best methods for getting into visual information [1]. CBIR deals with image content, such as color, shape and structure instead of annotated text. In order to implement CBIR, the framework needs to comprehend and interpret the content of stored images. The retrieval index should be created automatically, which provides a more visual retrieval interface to users. The fundamental idea of CBIR is to analyze image information by low level features of an image [2] and to set up feature vectors of an image as its index. CBIR has exceptionally wide and essential applications in many areas including military affairs, medical science, education, architectural design, the justice department and agriculture, etc.

The advancement of CBIR exploration was clearly summarized at high level in [2, 3]. Features are the basics for CBIR, for the whole image or locally for a small group of pixels. As per the techniques utilized for CBIR, features can be grouped into low-level features and high-level features. The most practical CBIR system depends on the color, shape, texture and other low-level features. Some researchers aim at reducing the semantic gap between visual features and the richness of human semantics [4]. With a particular ultimate goal to derive the high-level semantic features for CBIR, object-ontology [5] was used to characterize high-level concepts. Supervised or unsupervised learning methods were used to associate low-level features with the query concepts. Relevance feedback was introduced into the retrieval loop for learning user's intentions and semantic templates, it was generated to support high-level image retrieval. As there is inconsistency in comprehension, the gap between semantics in the visual data for different is difficult to eliminate.

Research on CBIR can be partitioned into two groups on the basis of the features used to retrieve the required image. Prior approaches utilized features such as shape, texture, color and region to retrieve the required image. The current methodologies use a distinctive combination of visual features to retrieve the required image [1,6]. The shape descriptor gives prevalent data in image retrieval because the shape is the main source through which people can perceive objects. These shape features are retrieved by two strategies, boundary based shape feature and region based shape feature. The boundary based shape feature extraction technique is based on the outer boundary of an object, while the region based shape feature technique is based on the whole region of an image. The different techniques in view of texture features have been proposed in the literature. This includes both statistical and spectral approaches. The greater part of these strategies were not able to capture the required information. Color is the most reliable feature, it is easier to implement and robustness to background compilation. It is not influenced by image size and its orientation. The most generally perceived methodology for color feature extraction is histogram. Color histogram illustrates its distribution in an image and it involves low computational cost. The main disadvantage of color histogram is, it cannot completely consider spatial information and it is not exclusive.

In spite of utilizing extracted information from an image, the majority of the CBIR frameworks yields imprecise outcomes. The semantic gap is defined as to relate the low-level features with the high-level user semantics. The relevance feedback method was used to over-come this semantic gap [7]. In content-based image retrieval framework, a distinctive feature of the queried image are explored in search for equivalent image features in the database [8]. Frequently, it is observed that there is a semantic gap between visual features and the semantic content of an image. Extracting more effective image features can decrease this semantic gap, this is a challenging task in CBIR research. Moreover, different machine learning techniques are used to reduce this semantic gap. In [9] SVM was utilized to extract image features and it retrieves the querying image efficiently. The hierarchical methodology [10] retrieve an image, where two separate features are explored to extract the contents and texture of an image. The proposed technique is also assessed over noisy images.

In this paper, pixel based feature descriptors (Gabor, LBP and HOG) are adopted for a faithful representation of images. The motive is to extract complimentary information from various descriptors from an image. The different similarity distance metrics such as Euclidean, Mahalanobis and Minkowski-Form distance metrics are used to retrieve images from databases. The similar image retrieval using the pixel based feature descriptor and the similarity metric are evaluated using precision and recall.

2. Methodology

The proposed content based image retrieval methodology is shown in Fig. 1. The proposed methodology involves image database, query image, feature extraction, similarity measures and the extracted images. The image database used for the proposed approach is Corel dataset and it consists of different categories, namely like Africans, buses, dinosaurs, etc. The datasets were collected and fed as input to the feature extraction process. The pixel based feature descriptors such as, Gabor [11], LBP [12] and HOG feature descriptors [13] are extracted and stored as a feature set. The image to be retrieved is given as a query image, for which the features such as, Gabor, LBP and HOG are extracted. These extracted feature sets were compared to the existing feature sets using a similarity measure, namely Euclidean distance, Mahalanobis distance and Minkowski-Form distance. This proposed model is done to achieve the goal to extract images similar to the query image.

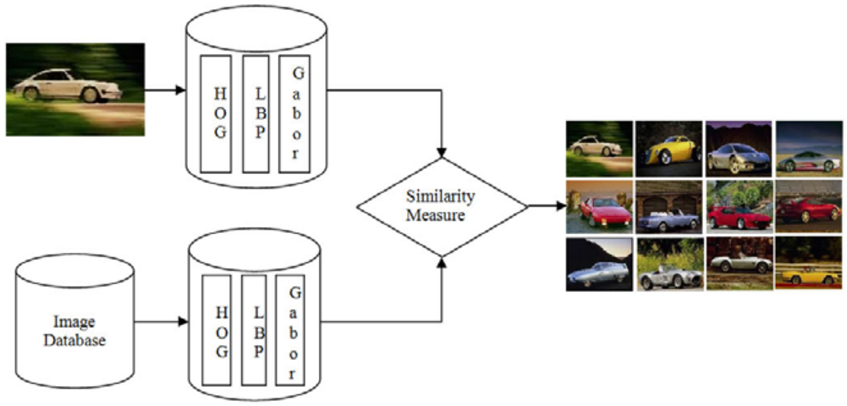

Figure. 1. Pictorial Representation of the Content Based Image Retrieval Architecture

2.1 HOG Features

HOG features, which were connected for human detection [14] are computed by counting the occurrences of gradient orientation in localized portions of an image. HOG features are based on the fact that the appearance and shape of the facial features can be depicted by the distribution of intensity gradients. The features so acquired are profoundly discriminative and represent an image characteristic faithfully [15]. Motivated by the significant results for feeling acknowledgment, we have adopted HOG features for image retrieval.

2.2 LBP Features

LBP features are observed to be a capable technique for texture feature extraction and have been popularly accepted for image representation [16, 17]. The most important properties of LBP are computation simplicity and illumination invariance. Moore and Bowden [18] have explored numerous variants of LBP for multi-view image representation to investigate the importance of multi-resolution and orientation analysis for feature representation. We have used extended LBP, which is rotation invariant. Rotation invariant LBP utilizes fewer bins as compared to consistent LBP and reduces the quantity of components used for binary pattern representation.

Feature extraction is implemented as follows: to start with, the image is divided into several non-overlapping blocks. Then, LBP histograms are calculated for each block. Finally, the block LBP histograms are concatenated into a single vector. Therefore, the images are represented by the LBP and the shape is recovered by the concatenation of different local histograms.

2.3 Gabor Features

Gabor filters have been utilized widely in image processing, texture analysis for their excellent properties: optimal joint spatial\spatial-frequency localization and the capacity to simulate the receptive fields of simple cells in the visual cortex [19]. Two-dimensional Gabor filter is a complex sinusoidal regulated Gaussian function with the response in the spatial domain and in the spatial-frequency domain.

2.4 Euclidean Distance

The squared distance between two vectors $x = [x_1, x_2]$ and $y = [y_1, y_2]$ is the sum of squared differences in their coordinates. To denote the distance between vectors x and y can use the notation, so that the last result can be written as an equation (1).

$$d_{x,y} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad - \quad (1)$$

2.5 Mahalanobis Distance

Mahalanobis distance is constantly utilized for data clustering, computed by measuring two data points in the space defined by relevant features, clustering technique group data into clusters so that the data objects within a cluster have high similarity to one another. This process is difficult as the data objects will be dissimilar to those in other clusters. As it portrays unequal variances and correlations between features, it will sufficiently assess the distance by giving different weights to the features of data points. This term originates from Mahalanobis [20] and originally refers to a distance measure that incorporates the correlation between features.

2.6 Minkowski-Form Distance

Many similarity measures on shapes, they are based on the $Dist_{xy}$ distance between two points. For two points the $Dist_{xy}$ Minkowski Distance is defined as an equation-(2).

$$Dist_{xy} = \left(\sum_{k=1}^{d} \left| x_{ik} - y_{ik} \right|^{\frac{1}{p}} \right)^{p} \qquad - \qquad (2)$$

2.7 Evaluation Metrics

The performance of the proposed CBIR framework is measured in three aspects, namely, efficiency, effectiveness and computational complexity. The effectiveness of a framework is related to the retrieval accuracy of the framework and is measured using the most widely used precision (percentage of retrieved images that are also relevant) and recall (percentage of relevant images that are retrieved) [21] and is defined as follows in equation (3) – (4):

$$\Pr ecision = \frac{R_i}{T_i} \qquad - \qquad (3)$$

$$\operatorname{Re} call = \frac{R_i}{T} \qquad - \qquad (4)$$

Where R_i is the number of relevant retrieved images, T is the total number of relevant images in the image database, and T_i is the total number of all retrieved images. The proposed system's effectiveness is measured in terms of average recognition rate (ARR). This is defined as the percentage of retrieved images in top matches, that belongs to the same class as a query image.

3. Results and Discussions

The tests are carried out on COREL photo database. The COREL database contains more than 5000 pictures organized in categories. Each category has about 100 images. There are 50 categories and the corresponding database is composed of 5000 images. The results presented here are five categories directly extracted from the 50 categories.

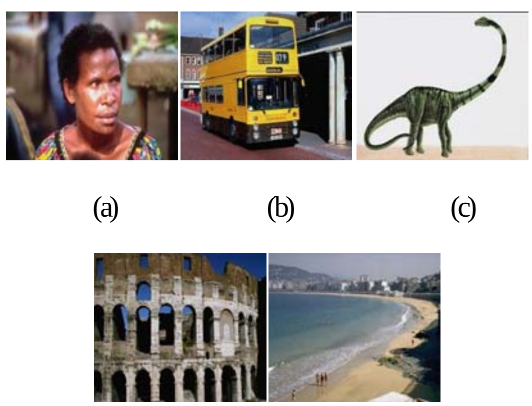

Figure 2. Image Datasets used from Corel database a) African dataset b) Buses c) Dinosaur d) Monuments and e) Beaches

Fig. 2 shows that the images are taken from different categories in Corel database, such as African, buses, dinosaur, monuments and beaches, etc. The dinosaur image retrieved using Euclidean Distance, Mahalanobis Distance and Minkowski-Form Distance are shown in Fig. 3.

(a)

(b)

(c)

(d)

Figure 3. Results of proposed CBIR on Dinosaur dataset using different similarity metrics (a) Query Image, (b) Images retrieved using Euclidean Distance, (c) Mahalanobis Distance and (d) Minkowski-Form Distance.

	Lin et al. [22]	Elalami [23]	Poursistani et al. [24]	Guo et al. [25]	Subra et al. [26]	Walia et al. [27]	Irtaza et al. [28]	Elalami [29]	Zeng et al. [30]	Proposed Method
Africa	68.30	70.30	70.20	84.70	69.75	51.00	65.00	72.60	72.50	73.30
Buses	88.80	87.60	76.30	85.30	89.65	78.00	85.00	89.10	89.20	90.05
Dinosaurs	99.25	98.70	100.00	99.30	98.70	100.00	93.00	99.30	100.00	99.35
Monuments	56.15	57.10	70.80	67.80	63.95	58.00	62.00	58.70	70.60	73.25
Beaches	54.00	56.10	44.40	45.40	54.25	44.00	60.00	59.30	65.20	64.35

Table 1. Comparison of proposed method precision with other method's precision metric

Table 1 shows the comparison of proposed method‘s precision with the other existing methods. The precision of the proposed image retrieval method achieves on average 80.06%, wherein Lin et al [22] achieves 73.3%, Elalami [23] achieves 73.96%, Poursistani et al. [24] achieves 72.34%, Guo et al.[25] achieves 76.5%, Subra et al [26] achieves 75.26, Walia et al. [27] achieves 66.2%, Irtaza et al. [28] achieves 73.00%, Elalami [29] achieves 75.8% and Zeng et al. [30] achieves 79.5%. The statistical results shows that the precision value of the proposed method is more than the existing methods, which implies that the proposed method retrieves more relevant images than the existing methods. The average recall value of the proposed method is 19.26%, wherein Lin et al [22] achieves 14.58%, Elalami [23] achieves 14.94%, Poursistani et al [24] achieves 14.46%, Guo et al. [25] achieves 15.3%, Subra et al [26] achieves 15.05%, Walia et al. [27] achieves 15.08%, Irtaza et al. [28] achieves 14.6%, Elalami [29] achieves 15.8%, Zeng et al. [30] achieves 15.9%. The comparison of precision value between the proposed and the other existing methods is shown in Fig 4.

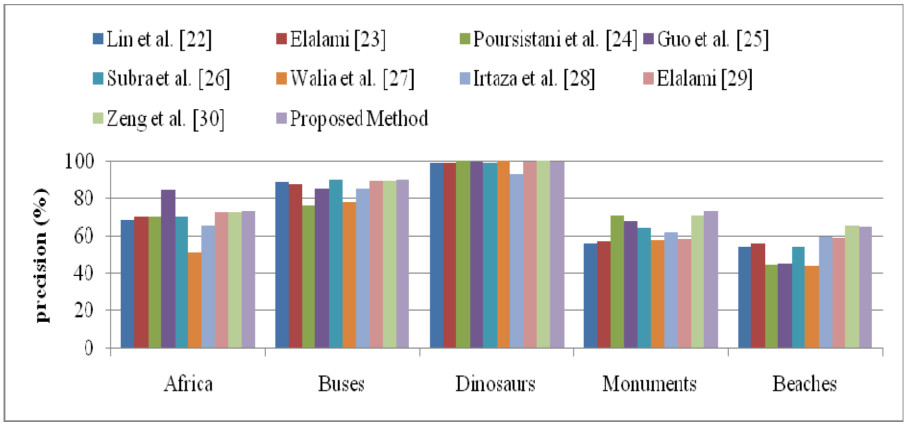

Figure 4. Comparison of precision of proposed approach to the other methods for different semantic classes in the Corel database.

The observation from statistics shows that the recall value of the proposed method is higher than the existing methods and this indicates that the proposed method retrieves more relevant images from the image database. Hence, from the obtained precision and recall values the proposed content based image retrieval method using HOG, LBP and Gabor feature descriptors retrieves more relevant images from the database that the existing methods.

4. Conclusion

This study proposes a content based image retrieval using HOG, LBP and Gabor feature descriptors. There is a significant difference between the proposed image retrieval performance and other existing method's performance. The study is experimented on Corel dataset and its retrieval performance is evaluated using precision and recall. This research discloses the effectiveness of HOG, LBP and Gabor feature descriptors and suggests that the proposed model can be useful in retrieving relevant images. The statistical results shows that the precision and recall of the proposed approach are 80.06% and 19.26% respectively. The comparative study shows that the proposed content based image retrieval method proves its feature extraction capability than the other image retrieval methods. The future enhancement will be a design of improving the system and utilizing the same to predict crime prevention, medical diagnosis, intellectual property and textile industry.

5. References

1. Muneesawang P, Guan L. An Interactive Approach for CBIR Using a Network of Radial Basis Functions, IEEE Transactions on Multimedia, 2004, 6 (5), pp 703–716.

2. Datta R, Joshi D, Li D, Wang JZ. Image Retrieval: Ideas, Influences, and Trends of the New Age, ACM Computing Surveys, 2008, 40, (2), pp 1-60.

3. Smeulders AW, Worring M, Santini S, Gupta A, Jain R. Content-Based Image Retrieval at the End of the Early Years, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22, (12), pp 1349-1380.

4. Lu Y, Hu C, Zhu X, Zhang H, Yang Q. A Unified Framework for Semantics and Feature Based Relevance Feedback in Image Retrieval Systems, Proc. Eighth ACM international conference on Multimedia, 2000, pp 31-37.

5. Mezaris V, Kompatsiaris I, Strintzis MG. An Ontology Approach to Object-Based Image Retrieval, *Proc. International Conference on Image Processing*, 2003, 2, pp 511-514.

6. Chun YD, Kim NC, Jang IH. Content-Based Image Retrieval Using Multiresolution

Color and Texture Features, IEEE Transactions on Multimedia, 2008, 10, (6), pp 1073-1084.

7. Bugatti PH, Traina C, Traina AJM. Improving Content-Based Retrieval of Medical Images Through Dynamic Distance with Relevance Feedback, *Proc. 24th International Symposium on Computer-Based Medical Systems (CBMS)*, 2011, pp 1–6.

8. Iqbal K, Odetayo MO, James A. Content-Based Image Retrieval Approach for Biometric Security Using Colour, Texture and Shape Features Controlled by Fuzzy Heuristics, Journal of Computer and System Sciences, 2012, 78, (4), pp 1258-1277.

9. Royal M, Chang R, Qi X. Learning from Relevance Feedback Sessions Using a K-Nearest-Neighbor-Based Semantic Repository, *Proc. IEEE International Conference on Multimedia and Expo (ICME07)*, 2007, pp 1994–1997.

10. Rao Y, Mundur P, Yesha Y. Fuzzy SVM Ensembles for Relevance Feedback in Image Retrieval, LNCS, 2006, 4071, pp 350-359.

11. Hamamoto Y, Uchimura S, Watanabe M, Yasuda, Tetsuya, Mitani Y, Tomita S. A Gabor Filter-Based Method for Recognizing Handwritten Numbers, Pattern Recog, 1998, 31, (4), pp 395–400.

12. Kaltwang S, Rudovic O, Pantic M. Continuous Pain Intensity Estimation from Facial Expressions, Advances in Visual Computing, Springer, 2012, 7432, pp 368–377.

13. Khan RA, Meyer A, Konik H, Bouakaz S. Pain Detection Through Shape and Appearance Features, *Proc. IEEE International Conference on Multimedia and Expo (ICME)*, 2013, pp 1–6.

14. Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection, *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR*, 2005, 1, pp 886–893.

15. Mavadati S, Mahoor M, Bartlett K, Trinh P, Cohn, J. DISFA: A Spontaneous Facial Action, Intensity Database, IEEE Transactions on Affective Computing, 2013, 4, (2), pp 151–160.

16. Feng X, Pietikainen M, Hadid A. Facial Expression Recognition with Local Binary Patterns and Linear Programming, Pattern Recognition and Image Analysis, 2005, 15, (2), pp 546-548.

17. Shan C, Gong S, McOwan PW. Facial Expression Recognition Based on Local Binary Patterns: A Comprehensive Study, Image and Vision Computing, 2009, 27, (6), pp 803–816.

18. Moore S, Bowden R. Local Binary Patterns for Multi-View Facial Expression Recognition, Computer Vision and Image Understanding. 2011, 115, (4), pp 541–558.

19. Rashedi SSE, Nezamabadi-pour H, Saryazdi S. A Simultaneous Feature Adaptation and Feature Selection Method for Content-Based Image Retrieval Systems, Knowledge-Based Systems, 2013, 39, pp 85-94.

20. Mahalanobis PC. On The Generalized Distance in Statistics, *Proc. National Institute of Science of India*, 1936, 2, (1), pp 49–55.

21. Huang J, Kumar SR, Mitra M, Zhu W, Zabih R. Image Indexing Using Color Correlograms, *Proc. Conference on Computer Vision and Pattern Recognition*, 1997, pp. 762–768.

22. Lin CH, Chen RT, Chan YK. A Smart Content-Based Image Retrieval System Based on Color and Texture Feature, Image and Vision Computing, 2009, 27, (6), pp 658-665.

23. Elalami ME. A Novel Image Retrieval Model Based on the Most Relevant Features, Knowledge-Based Systems, 2011, 24, (1), pp 23–32.

24. Poursistani P, Nezamabadi-pour H, Moghadam RA, Saeed M. Image Indexing and Retrieval on JPEG Compressed Domain Based on Vector Quantization, Mathematical and Computer Modelling, 2013, 57, (5–6), pp 1005–1017.

25. Guo JM., Prasetyo H, Su HS. Image Indexing Using The Color and Bit Pattern Feature Fusion, Journal of Visual

Communication and Image Representation, 2013, 24, (8), pp 1360–1379.

26. Subrahmanyam M, Wu QMJ, Maheshwari RP, Balasubramanian R. Modified Color Motif Co-Occurrence Matrix for Image Indexing and Retrieval, Computers and Electrical Engineering, 2013, 39, (3), pp 762-774.

27. Walia E, Pal A. Fusion Framework for Effective Color Image Retrieval, Journal of Visual Communication and Image Representation, 2014, 25, (6), pp 1335–1348.

28. Irtaza A, Jaffar MA, Aleisa E, Choi TS. Embedding Neural Networks for Semantic Association in Content Based Image Retrieval, Multimedia Tools and Applications, 2014, 72, (2), pp 1911–1931.

29. ElAlami ME. A New Matching Strategy for Content Based Image Retrieval System, Applied Soft Computing, 2014, 14, pp 407–418.

30. Zeng S, Huang R, Wang H, Kang Z. Image Retrieval Using Spatiograms of Colors Quantized by Gaussian Mixture Models', Neurocomputing, 2016, 171, (1), pp 673–684.