

Multi-step approach for Speeding-up Region Based Image Retrieval using Indexing technique

Miss. Rutuja N. Pathare¹, Prof. K. W. Kimmatkar²

¹ Student, Department of Computer Engineering, Zeal College of Engineering & Research, Narhe, Pune.

² Assistant Professor, Department of Computer Engineering, Zeal College of Engineering & Research, Narhe, Pune.

Abstract - In the recent year the database of the images is growing rapidly and there is huge demand in the enhancement of the intelligent retrieval of the images. Color, shape, texture, spatial layout are the main attributes to represent as well index the images. In this paper we proposed a novel approach which integrates Indexing techniques, support vector machine based learning with new random algorithm called firefly algorithm as a relevance feedback technique into region based image retrieval system (RBIR). The Region Based Image Retrieval system (RBIR) system uses the segmentation on image into regions. For approaching towards the gathering of similar images in the flying of the track space based on user's evaluation, the firefly optimizer guides the swarm agents. With the firefly algorithm, this research design has a focused concern to step-up the performance by optimizing division feature. Set of optical characteristics is used to show each region of the image and the equality between regions and is measured using a particular metric function on such characteristics. The proposed system efficiency is tested on the standard subset of Pascal database images and performance is compared with the existing image retrieval approaches with various parameters.

Key Words: Content based image retrieval, K-Means Algorithm, firefly optimizer, Region Based Image Retrieval, Support Vector Machine, Firefly algorithm.

1. INTRODUCTION

Existing image retrieval approaches were built on textual annotation of images. Images can be set by current or syntactic grouping to ease navigation and browsing on the basis of standard Boolean queries by using text descriptions. It was well accepted that a more skilled and direct method to exhibit and index optical information would be based upon the basic features of the images themselves. The optical elements of an image such to exhibit and index the image Shape, color, spatial layout, and texture is used to apply CBIR. In ideal content based image retrieval systems, multi-dimensional feature vectors are used to extract and illustrate the visual features of the images in the database. The feature vectors of the images exist in the database result in formation of a feature database. For image retrieval, the user's inputs sample images or sketched figures to the retrieval system. Sample images are converted into its

internal stereotype of feature vectors by system.

The similarities/distances among the feature vectors of the query sample or sketch & those of the images accessible in the database are then estimated and retrieval is performed using an indexing scheme. To examine the image database indexing scheme provides an efficient approach. Current retrieval systems have enclosed user's relevance feedback to settle the retrieval process in order to create perceptually and semantically more precise retrieval results. The optical content descriptor is either local or global. The optical features of the whole image & optical features of objects are use in local descriptor to describe the image content is used by global descriptor. An image is often divided into parts first for developing local visual descriptors. Content based image retrieval systems compute the visual similarities between a query image & the images in a database instead of exact matching, e. The result for the input query image is not just a single image but, a set of images organized according to their similarities with the query image. Various similarity/distance measures will affect the performances of an image retrieval system substantially. The most commonly applied similarity measures are: Euclidean Distance and Bhattacharyya Distance.

Efficient indexing and quicker image retrieval on the basis of visual features is one of the main issues in CBIR. As the feature vectors of images mostly have high dimensionality and hence they are not eligible for conventional indexing structures, hence we are going to do dimension reduction which will set up an effective indexing scheme. One common approach is Principal component analysis for dimension reduction.

In PCA we linearly depict input data to a coordinate space, so that the axes are aligned to point the maximum variations in the data. After dimension reduction, the indexing of multi-dimensional data is done.

The remainder of this paper is organized as follows. We provide a deep literature study of existing system in Section II. Section III describes the architecture of the proposed system and working of the proposed system with the implementation details is explain in Section IV and Conclusion is then given in Section V.

2. RELATED WORK

The traditional keyword based stratagem for image retrieval becomes embarrassed from the following limitations that are; Requires an tremendous amount of

member of the working class operated work [1,2] and particularity of human approach leads to untrue results. To take over these limitations content-based image retrieval or (CBIR) [3] has been used. CBIR is a set of methods and algorithms which especially focuses on silent level features of image content such as shape, color and texture for retrieving the images from an image database based on the search image query given by user. The accomplishment of CBIR is far from satisfaction on user's ambition such as high level concepts because its primary attract is on silent level features of the images.

Using Content based Image Retrieval; only low-level features are instantly extracted from images, to capture the shape, color and texture impression of the image. Not complete of them efficient to reflect high level concepts in the user's mind, as a result the retrieval performance is absolutely far from user satisfaction. To study this obstruction [4], two generally used techniques are: Region based image retrieval or

Region-based Image Retrieval that equal to the representation of the user's approach of perception item into segmented point features and is specially to verify a user's desire.

Region-based Image Retrieval which is based on segmentation of image to gain the performance of an image retrieval process. The region based methods are used to calculate features of the segmented point of the image at the object level and the compare similarity at the granularity of the point whereas earlier traditional techniques liberate images specifically based on complete features of the image. Relevance reply is a well known of the useful tool that retrieves complementary word as by the user's regard by iterating in the contour [5]. The reply which is combined from the user allows reevaluating the input search or helps in useful partition of the relevant image cluster. As the technology in the machine learning is extended, a location of machine learning techniques have been familiar to the relevance feedback problem in image retrieval system. Some of them are boosting techniques [6], dimension reduction [7], ensemble learning [8], Bayesian learning [9], decision tree [10], discriminant analysis [11] etc. Recently, Support Vector Machines or (SVM) have been hugely express in equipment learning as its performance in the real-world applications is barely higher in pattern classification.

In Support Vector Machine, the training word is not linearly divisible; SVM introduces the work which maps data into a higher-dimensional point where the data is partitioned into two classes. Such mapping stage creates problems. The results of Support Vector Machine can be forced tougher by adjusting the internal parameters, one that they decrease the error. So incidentally do thus, it could be adventitious to synthesize the advantages of SVM by all of well-mannered, theoretical problem solving algorithms, a well known as Meta-heuristic algorithms [12], which fulfill situation between randomization and local search. Support vector machine learning way is set into a meta-heuristics firefly algorithm as a relevance feedback course in point based image retrieval. Thus, decrease the performance of the

Support Vector Machine or SVM classifier is gain by assist the new developed meta-heuristic algorithms, the firefly algorithm. The firefly optimizer[12] not only allow as a helpful optimization engine, but further examine the track space around and prevent converging to local minima. Also, for the advance of retrieval accuracy, random walk method based on Gaussian distribution is applied at the end of each iteration in firefly algorithm.

Various methods and broad literature about content-based image retrieval are demonstrated in [13, 14]. [15] Present a detailed study existing relevance feedback techniques for image retrieval. CBIR systems that employ global representations of images have integrated SVM as an efficient tool for its better performance [16], Interactive Genetic Algorithm (IGA) is employed in [17] and SVM to track the user perception. [18] constitutes two levels ideal i.e. query refinement & feature re-weighting based on intuitive heuristics. Later on iPURE [19] employed relevance feedback in RBIR system which uses a revised intra query learning method of [17] to reduce the weight bias problem.

[20] and [21] showed that multimodal optimization problems can be efficiently developed and figure out by firefly algorithm. In [22] it is showed that FA created a reliable and improved performance on base of time and optimality for feature selection when compared with other algorithms.

In [23], Integrated Region Matching, a new similarity measure for region-based image similarity or closeness is proposed. In this Image retrieval based on segmented regions using the similarity study mutually robust integrated region-matching course and higher accuracy is studied. In this paper image retrieval program segments images using the k-means algorithm which is based on color and frequency features. For general purpose images such photo library or the images on World Wide Wed. Shape, Color and texture attributes of the regions is used to measure the distance between a region pair.

In [24], newly emerged Content Based Image Retrieval (CBIR) is used. Shape is one of the primary silent level image features. Contour-based methods return shape boundary features at the same time ignore shape inner content. Region-based methods capture shape inside features while do not emphasize on boundary features. So neither one of this method can produce ideal retrieval results. In this paper, integrated shape descriptor, which combines contour shape descriptors and the region shape descriptors, is used. The integrated shape descriptors allow interactivity between retrieval program and users and can increase retrieval performance.

In [25], multiple-region level image retrieval algorithm which is developed on region-level image segmentation and their spatial relationship. To nab spatial uniformity, we used Hausdorff Distance to our region-based image retrieval program for finding region in the pictures. Compare to other object or thousand and one region-based retrieval systems, author updated classic Hausdorff Distance to extricate similar regions about to their spatial signification, addition, and deletion. Therefore we connect relevance feedback to

copy the user's high-level search and especially to the program and for performance degradation merit to inexact image segmentation.

In [26], we studied a unique approach that is firefly algorithm as a relevance feedback approach to a region based image retrieval program which combine support vector machine based study by all of an evolutionary logical algorithm. This program overcomes the semantic point through contemporary iterative learning and further produces a better exploration of solution space.

Studying the existing system we find out that the retrieval process needs to be faster to improve the performance of the system. So we proposed a multi-step mechanism which will use indexing to speed up the retrieval process.

3. PROPOSED SYSTEM

2.1 System Architecture

We are using multi-step approach to pass region-based techniques for enormous image collections. We used general techniques for image retrieval are shape, color and texture once Visual similarity evaluation and Support Vector Machine along with firefly algorithm.

Region-based image retrieval (RBIR) is an extension to the content-based image retrieval (CBIR). In RBIR system automatically segments images into a diverse number of regions and extracts a set of features for each region. Then dissimilarity field calculates the distance between set of recommendation regions and database image. Therefore, the enormous evaluation costs of the dissimilarity field are restricting RBIR to simply small databases. In this paper, we used multistep stratagem to pass region-based techniques for large image collections. Firstly we used applying general techniques for image retrieval are shape, color and texture previously Visual similarity evaluation.

To calculate a color histogram, each image added to the group is analyzed to calculate color histogram, shows the purity of pixels of each color in the image. The color histogram is previously stored in the database for each image. At the time of searching addict can specify the desired proportion by all of each color or user can submit an example image from which course of action can calculate color histogram. Whose color histograms correlate those of the query approximately similar and previously matching manner retrieves those images. The matching technique most consistently used histogram intersection. A current Content Based Image Retrieval system uses alternative of this technique. This research design has a focused approach to improve the performance by optimizing region feature with the firefly algorithm. The performance of the eventual technique is compared with other existing retrieval methods like genetic algorithm, particle swarm optimization, query point movement and support vector machine to notice the high quality with regard to the ideal in terms of recall and precision.

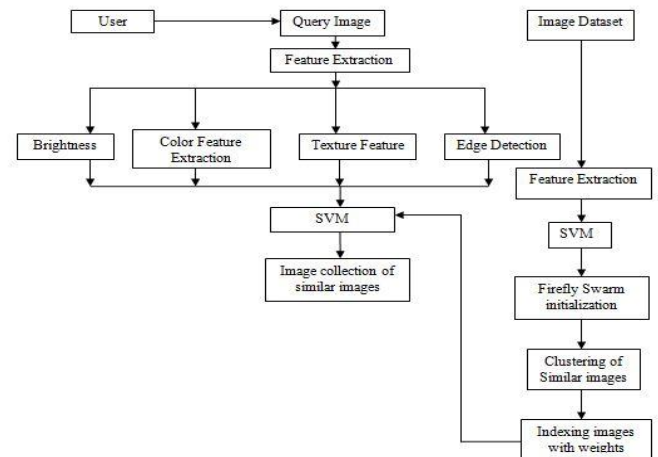


Fig: 1. System Architecture

2.2 Algorithm (cluster Generation)

Input:

- 1) Dataset containing relevant and irrelevant images $img_1 \dots img_n$

Output:

- 1) Clusters of n most relevant images

- 1: Read_Image(q_image)
- 2: Features_q_image[] = Extract_Features(q_image)
Where Features = {Brightness, Color feature, Texture Features}
- 3: for each(i=0 to n)
- 4: Features_img_i[] = Extract_Features(img_i)
- 5: SVM(Features_q_img, Features_img_i)
- 6: Sort fireflies with light intensity
- 7: Perform image indexing with weights
- 8: Rank relevant images
- 9: Generate clusters of relevant images
- 10: End if

2.3 Algorithm (Relevant Image Extraction)

Input:

- 1)Query image q_img
- 2)Cluster containing relevant images $img_1 \dots img_n$

Output:

- 1)Order list L of n most relevant images

- 1: Read_Image(q_image)
- 2: Features_q_image[] = Extract_Features(q_image)
Where Features = {Brightness, Color feature, Texture Features}
- 3: for each(i=0 to n)
- 4:Features_img_i[] = Extract_Features(img_i)
- 5:SVM(Features_q_img, Features_img_i)
- 6:Show images from related cluster
- 7:End for

4. RESULTS AND ANALYSIS

The common retrieval performance study precision and the recall are used as the analysis of the search results. Precision P is marked as the ratio of the number of retrieved relevant images r to the total number of images n, that is $P = r/n$. Precision P point to the accuracy and time of the retrieval. Recall R is defined as the ratio of the number of relevant images retrieved r to the total number m of relevant images in the entire database, that is $R = r/m$. Recall R indicates the robustness of the retrieval performance. For each query, the precision of the retrieval at each level of the recall is obtained.

The analysis of the proposed system is done by giving input as Pascal database images from which the relevant images are retrieved speedy with accuracy. Fig 1 depicts the graph of obtained the results.

- Accuracy: It is the degree to which the result of a measurement, calculation, or specification conforms to the correct value (true). The proposed system gives the 95 % Accuracy.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

Where,

TP (True Positive) – Correctly Identified.

FP (False Positive) – Incorrectly Identified.

TN (True Negative) – Correctly Rejected.

FN (False Negative) – Incorrectly Rejected.

- Sensitivity: It is a True Positive Rate (TPR). It measures propagation of actual positives which are correctly identified. It measures the percentage of correctly identified. The TPR is 65%

$$\text{Sensitivity} = TP / (TP + FN)$$

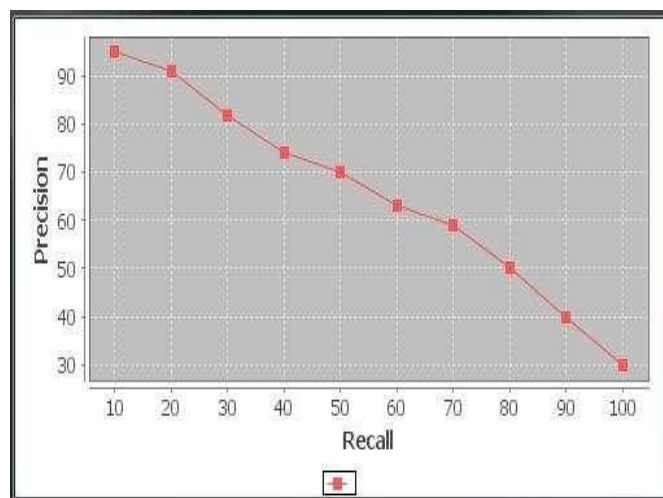


Fig 2: Precision and Recall

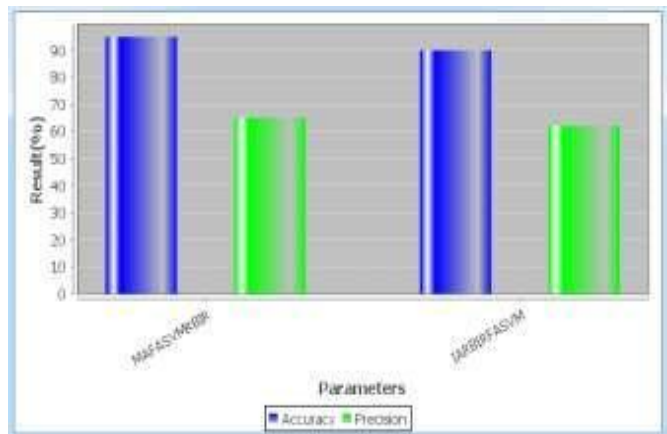


Fig 3: Comparative Result Analysis

TABLE 1: Performance Analysis

Performance measure	Performance evaluation in %
Accuracy(Proposed System)	95
Sensitivity((Proposed System))	65
Accuracy(Existing System)	90
Sensitivity(Existing System)	62

The results obtained in fig 2 and the comparative analysis done in fig 3 and table 1 proves that the our proposed system architecture is much more efficient than the existing system as the we are not only getting the results more accurately but we have also speed up the retrieval process by applying the indexing technique.

5. CONCLUSION

Region-based Image Retrieval (RBIR) is subset of classical Content-based Image Retrieval. Region-based Image Retrieval systems provide new approach to search for objects embedded in an arbitrary environment. For multi-region queries, we deployed system which is combination of Color Feature Based Retrieval Firefly algorithm and Support Vector Machine which is more intuitive and efficient compared to alternative models. The result gained shows the improvement in the retrieval of the RBIR images with faster speed and great accuracy. The accuracy is measured is 95% and the TPR is 65% of the proposed system.

ACKNOWLEDGEMENT

Our sincere thanks go to ZCOER for providing a strong platform to develop our skill and capabilities. I would like to thanks all those who directly or indirectly help us in presenting the paper. I hereby take this opportunity to express our heartfelt gratitude towards the people whose help is very useful to complete our project. I would like to express our heartfelt thanks to my guide Prof. K. W. Kimmatkar whose experienced guidance became very

valuable for us.

REFERENCES

- [1] V.S. Tseng, J.H. Su, J.H. Huang, C.J. Chen, "Integrated mining of visual features, speech features and frequent patterns for semantic video annotation", *IEEE Trans. Multimed*, 2008, pp. 260-267.
- [2] V.S. Tseng, J.H. Su, B.W. Wang, Y.M. Lin, "Web image annotation by fusing visual features and textual information", in: *Proceedings of the 22nd ACM Symposium on Applied Computing*, 2007, pp. 1056-1060.
- [3] R. Datta, D. Joshi, J. Li, J.Z. Wang, "Image retrieval: ideas, influences, and trends of the new age", *ACM Comput. Surv.* 2008, <http://dx.doi.org/10.1145/1348246.1348248> (Article 5, 60 pp.).
- [4] F. Jing, M. Li, H.J. Zhang, B. Zhang, "Region-based relevance feedback in image retrieval", *IEEE Trans. Circuits Syst. Video Technol*, 2004, pp. 672-681.
- [5] Y. Rui, T.S. Huang, M. Ortega, S. Mehrotra, "Relevance feedback: a power tool in interactive content-based image retrieval", *IEEE Trans. Circuits Syst. Video Technol*, (1998, pp. 644-655.
- [6] K. Tieu, P. Viola, "Boosting image retrieval", in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2000, pp. 228- 235.
- [7] D. Tao, X. Tang, "Nonparametric discriminant analysis in relevance feedback for content-based image retrieval", in: *Proceedings of the IEEE International Conference Pattern Recognition*, 2004, pp. 1013-1016.
- [8] C.H. Hoi, M.R. Lyu, "Group-Based relevance feedback with support vector machine ensembles", in: *Proceedings of 17th International Conference Pattern Recognition*, 2004, pp. 874-877.
- [9] N. Vasconcelos, A. Lippman, "Learning from user feedback in image retrieval systems", in: *Proceedings of NIPS*, 1999, pp. 977-983.
- [10] S. MacArthur, C. Brodley, C. Shyu, "Relevance feedback decision trees in content-based image retrieval", in: *Proceedings of the IEEE Workshop Content-Based Access of Image and Video Libraries*, 2000, pp. 68-72.
- [11] X.S. Zhou, T.S. Huang, "Small sample learning during multimedia retrieval using biasmap", in: *Proceedings of the IEEE International Conference Computer Vision and Pattern Recognition*, 2001, pp. 11-17.
- [12] X.S. Yang, "Firefly algorithms for multimodal optimization, in: *Stochastic Algorithms: Foundations and Applications*", SAGA 2009, *Lecture Notes in Computer Sciences*, 2009, pp. 169-178.
- [13] Rui, Yong, Thomas S. Huang, and Shih-Fu Chang. "Image retrieval: Current techniques, promising directions, and open issues." *Journal of visual communication and image representation*, 1999, pp. 39-62.
- [14] Smeulders, A. W., Worring, M., Santini, S., Gupta, A., & Jain, R., "Content-based image retrieval at the end of the early years. *Pattern Analysis and Machine Intelligence*", *IEEE Transactions on*, 2000, pp. 1349-1380.
- [15] Zhou, Xiang Sean, and Thomas S. Huang. "Relevance feedback in image retrieval: A comprehensive review", *Multimedia systems*, 2003, pp. 536-544.
- [16] Zhang, Lei, Fuzong Lin, and Bo Zhang. "Support vector machine learning for image retrieval", *Image Processing*, 2001. *Proceedings. 2001 International Conference on*. Vol. 2. IEEE, 2001.
- [17] Wang, Shang-Fei, Xu-Fa Wang, and Jia Xue. "An improved interactive genetic algorithm incorporating relevant feedback", *Machine Learning and Cybernetics*, 2005. *Proceedings of 2005 International Conference on*. Vol. 5. IEEE, 2005.
- [18] Rui, Yong, Thomas S. Huang, and Sharad Mehrotra, "Content-based image retrieval with relevance feedback in MARS", *Image Processing*, 1997. *Proceedings*., *International Conference on*. Vol. 2. IEEE, 1997.
- [19] Aggarwal, Geeta, et al., "ipure: Perceptual and user-friendly retrieval of images", *Multimedia and Expo, IEEE International Conference on*. Vol. 2. IEEE, 2000.
- [20] Kazemzadeh, Azad S. "Optimum Design Of Structures Using an Improved Firefly algorithm", 2011, pp. 327-340.
- [21] Gandomi, A. H., Yang, X. S., & Alavi, A. H, "Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems", *Engineering with computers*, 2013, pp. 17-35.
- [22] Banati, Hema, and Monika Bajaj. "Firefly based feature selection approach", *IJCSI International Journal of Computer Science Issues* 8.4, 2011.
- [23] Jia Li, J.Wang, G. Wiederhold, "IRM: Integrated Region Matching for Image Retrieval", *Stanford University Stanford*, 2000.
- [24] Dengsheng Zhang and Guojun Lu, "An Integrated Approach to Shape Based Image Retrieval", *5th Asian Conference on Computer Vision*, 2002.
- [25] ByoungChul Ko, Hyeran Byun, "Integrated Region-Based Image Retrieval Using Region's Spatial Relationships", *IEEE*, 2002.
- [26] T.Kanimozhi n, K.Latha. "An integrated approach to region based image retrieval using firefly algorithm and support vector machine", *ScienceDirect*, www.elsevier.com/locate/neucom, 2014, pp. 1099-1111.