International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9	**VOL**	01
Website	www.iciems.in	**eMail**	iciems@asdf.res.in
Received	10 - July - 2015	**Accepted**	31- July - 2015
Article ID	ICIEMS018	**eAID**	ICIEMS.2015.018

A Review on Feature Extraction Techniques for CBIR System

Kavita Chauhan[1], Shanu Sharma[1]
[1]Department of Computer Science & Engineering,
ASET, Amity University, Uttar Pradesh, Noida, India

ABSTRACT: *Ongoing expansion of digital images requires improved and efficient methods for sorting, browsing and searching operations through ever-growing image databases. Content Based Image Retrieval (CBIR) systems are search engines for image databases, which perform indexing on images according to their content and features. This paper presents the systematic review of various existing CBIR systems and their feature extraction techniques. Further the performance analysis and limitations of these systems have been discussed.*

KEYWORDS: CBIR, Image Feature Extraction, Similarity Measurement, Neural Network, Support Vector Machine.

I. INTRODUCTION

The advancement in computer technologies produces huge volume of multimedia data, specifically on image data. The greatest challenge of the World Wide Web is that the more information available about a given topic, the more difficult it is to locate accurate and relevant information. Generally users know that which information they need, but are unsure where to find it. Search engines can facilitate the ability of the users to locate such relevant information.

Content Based Image Retrieval is a technique which uses visual content to search and compares images from large scale image databases according to the interest of the users [1]. In this process firstly, the user submits a query image or a series of images and the system is required to retrieve images from the database as similar as possible. It also includes another task which is a support for browsing through large image databases, where the images are supported to be grouped or organised in accordance with similar properties [2]. During the past few years, CBIR has gained much devotion for its potential application in multimedia management. The term 'content' in this context might refer to colors, shapes, texture, or any other information that can be derived from the image itself. Basically, there are two ways of image retrieval in CBIR- Query by tag and Query by example. In the former method, the query is submitted in the form of tag like square is used for searching square shape in image, and in the later method the query is given with the image example such that the results resemble the given query [2]. CBIR is also known as Query by Image Content (QBIC) and Content-Based Visual Information Retrieval (CBIR).

This poses two main challenges for the Image Retrieval researchers and practitioners:

a. Low level features extracted and their semantic meanings may differ thus forming a gap known as the "Semantic Gap"[3]

b. Granularity of classification, this granularity is closely related to the level of invariant that the CBIR system should guarantee [3].

This paper is divided into four sections. Section I contains the introduction of the CBIR system and their challenges. Section II discusses the result of the various feature extraction techniques in the form of a table defining the various attributes and how they differ from each other. Section III concludes this paper followed by future scope.

Cite this article as: Kavita Chauhan, Shanu Sharma. "A Review on Feature Extraction Techniques for CBIR System." *International Conference on Information Engineering, Management and Security (2015)*: 110-114. Print.

II. RELATED WORK

A thorough and systematic literature has been done on various CBIR systems. The origin of major studies vary from listed repositories (IEEE, ACM Library, IJCA, Science Direct etc.) to common purpose search engines like Google. Various research papers with the search string 'Feature Extraction Techniques and CBIR System' has been searched and finally 65 papers have been downloaded out of which 17 papers were considered of most relevant to the objective. Review for the selected paper is further presented: Ligade *et al.* [3] provides a review on techniques of Neural Network, Interactive Genetic Algorithm and Relevance Feedback where the characteristics of the above 3 techniques has been described along with their current achievement, uses and advantages in image retrieval. The experimental evaluation is done by them using the convergence ratio, precision and recall parameters. Walia *et al.* [4] proposed a new similarity measure that improves the efficiency of the CBIR system by using the dominant color descriptor. Their work compares two images and consider their no. of dominant colors and their distance and thus improves the performance using the colors and also the results were verified on two different image databases. Sarangi *et al.* [5] proposed an automatic contrast enhancement technique using differential evolution for gray-scale images. The technique attempts to demonstrate the methods adaptability and effectiveness for searching global optima solutions to enhance the contrast and detail in gray scale images. Agarwal *et al.* [6] proposed a novel feature descriptor for CBIR system by integrating Cooccurence of Haar like Wavelet Filter (CHLWF) with Color Histogram. It extracts the image properties from different visual perspectives to give the image representation almost similar to the human interpretation and hence improves the effectiveness of retrieval system. Jeyabharathi *et al.* [7] analyzed the performance of the most useful feature extraction techniques which includes Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Independent Component Analysis (ICA) and also the performance of the most renowned classification techniques, i.e., Support Vector Machine (SVM) and Nearest Neighbour (NN). The performance metric used by them were Recognition Rate and F-Score. Based on the performance evaluation models, they concluded that PCA with SVM provide more recognition accuracy than others. Ezekiel *et al.* [8] proposed a CBIR technique based on multi-scale feature extraction scheme. They designed a Pulse Coupled Neural Network (PCNN) based fusion of a fast wavelet transformation and Contourlet transformation coefficients applied on Rescaled Range(R/S) analysis techniques. The method highlights the edges, segment edges and finds control points to answer the image retrieval query. Zhang *et al.* [9] introduced a method of color principal feature extraction called ColorPCA which works in color image space extracting the principal features directly from the color images. It considers only one parameter known as reduced dimension to estimate the projection axes. Syam *et al.* [10] proposed a CBIR system based on GA for Medical image retrieval using the feature extraction of color, texture and shape. They used Squared Euclidean Distance (SED) for the computation of similarity measure for efficient retrieval of images. Their work assures the benefit of the shape feature in addition to the other features. Ligade *et al.* [11] proposed an image extractor method on multi-feature similarity synthesis using the GA for efficient image retrieval in CBIR systems. The method extends the GA algorithm by using the methods of relevance feedback for magnified retrieval performance, also the method uses both the implicit and explicit feedback technique. Ho *et al.* [12] proposed a novel system architecture for CBIR system in which the well-known techniques are combined like content-based image, color analysis and data mining techniques for better performance and efficiency. It combines the segmentation and grid module, feature extraction module, K-means clustering and neighbourhood module to build the CBIR system. Chadha *et al.* [13] proposed an improved technique of image retrieval by incorporating the Query modification through image cropping. This feature identifies the user's interest region in a particular image and thus resulted into more precise and personalized search results. This technique results into 28% improvement in accuracy. Rashedi *et al.* [14] proposed a method for the CBIR system improving its precision by Feature selection using the Binary Gravitational Search Algorithm. It selects the most relevant features from the query image thus leading to more accurate results by reducing the semantic gap. The results are examined in the Corel database. Also they compared the work of GA and BPSO and found BGSA to be the best among them. Pighetti *et al.* [15] proposed a new architecture combining the multi-objective interactive genetic algorithm and Support Vector Machine. The multi-objective IGA is used for its capability to converge towards global optima and SVM for its capabilities to learn user evaluations required by IGA. Madugunki *et al.* [16] described the detail classification of the CBIR system and also about their efficiency. The work compares the CBIR and TBIR technique and discusses about the effect of different matching techniques on retrieval process. Euclidean Distance Method, City Block Distance and Canberra Distance Method are used to calculate the matching distance and found Canberra Distance Method best among others. Selvarajah *et al.* [17] introduced a descriptor called Combined Feature Descriptor for the CBIR system to enhance the retrieval performance. It uses the concept of Haar Wavelet and color histogram moment. The descriptor works for the application which includes traditional color moments, 2D-Discrete Wavelet Transform. Abubacker *et al.* [18] proposed a CBIR technique based on query and extracts the most vital attributes, i.e., color, shape and texture. It includes the automatic extraction of spatial based color feature using invariant Fourier descriptor and texture feature using the Gabor filter. The distance metrics were used for distance calculation and their weightage were determined. Based on the data, the output images are sorted and ranked so that most similar images can be displayed to the user. Omar *et al.* [19] proposed a WhatAreYouLOOKing4(WAY-LOOK4) system using the Local descriptor and Image Signatures. It contains 3 system components: feature extraction, image database indexing and similarity retrieval. The system maintains reasonable storage and computational costs. The system is simple because of no iterations for clustering or complex wavelet transformation.

Cite this article as: Kavita Chauhan, Shanu Sharma. "A Review on Feature Extraction Techniques for CBIR System." *International Conference on Information Engineering, Management and Security (2015)*: 110-114. Print.

The analysis of different techniques is presented in the following table.

Table 1: Analysis of Different Feature Extraction Techniques

S.No.	Author	Year	Source	Major Findings/Limitations/Future Scope
1	Ligade *et al.*	2014	IJCSIT	Reviews NN, IGA and RF using the convergence ratio, precision and recall and provides the various CBIR system based RF techniques.
2	Walia *et al.*	2014	IEEE	1. Proposed a new similarity measure that improves the efficiency of the CBIR system by using the dominant color descriptor. 2. In future, it can be used for color based retrieval of videos.
3	Sarangi *et al.*	2014	IEEE	Proposed an automatic contrast enhancement technique using differential evolution for gray-scale images
4	Agarwal *et al.*	2014	IEEE	Proposed a novel feature descriptor for CBIR system by integrating Cooccurence of Haar like Wavelet Filter (CHLWF) with Color Histogram
5	Jeyabharathi *et al.*	2013	IEEE	Analyses the performance of feature extraction techniques (PCA, LDA & ICA) and classification techniques (SVM & NN) using the Recognition Rate and F-Score and finds PCA with SVM to the best.
6	Ezekiel *et al.*	2013	IEEE	Proposed a multi-scale feature extraction scheme based on wavelet and Contourlet transform
7	Zhang *et al.*	2013	IEEE	1. Proposed a method of color principal feature extraction called ColorPCA 2. This method contains three main problems: a. Nonlinear structures remains unclear on how to extend the ColorPCA algorithm. b. To extend the ColorPCA idea in manifold structures or discriminant information of samples for locality preservation c. To determine the optimal reduced dimensions for dimension reduction algorithms using ColorPCA.
8	Syam *et al.*	2013	IEEE	Proposed a CBIR System using the GA and SED with feature extraction process.
9	Ligade *et al.*	2013	TJPRC	1. Proposed an image extractor method on multi-feature similarity synthesis using the Genetic Algorithm for efficient image retrieval in CBIR systems 2. They consider only the occurrence frequencies of image in result and not the location of the retrieval result which directly reflects the similarity of it and the query image.
10	Ho *et al.*	2012	IEEE	Proposed an architecture for CBIR system combining content-based image, color analysis and data mining.

Cite this article as: Kavita Chauhan, Shanu Sharma. "A Review on Feature Extraction Techniques for CBIR System." *International Conference on Information Engineering, Management and Security (2015)*: 110-114. Print.

11	Chadha *et al.*	2012	IJCA	Proposes an improved technique for image retrieval using the Query modification through image cropping.
12	Rashedi *et al.*	2012	IEEE	Proposed a method to improve the precision of CBIR system by Feature selection using the Binary Gravitational Search Algorithm
13	Pighetti *et al.*	2012	ICPR	Proposed a new architecture combining the multi-objective IGA and SVM.
14	Madugunki *et al.*	2011	IEEE	Compares the CBIR and TBIR techniques of image retrieval and also the matching technique and found Canberra Distance Method best among other matching technique.
15	Selvarajah *et al.*	2011	IEEE	Introduced Combined Feature Descriptor for the CBIR system to enhance the retrieval performance using Wavelet and Color Histogram concepts.
16	Abubacker *et al.*	2010	IEEE	1.Proposes a CBIR technique based on query and vital attributes. 2.Scanner& camera support can be aimed. 3. Different query formats support incorporating filtering & indexing mechanism can be used for quick image access from the database.
17	Omar *et al.*	2009	IEEE	1. Proposes a WhatAreYouLOOKing4(WAY-LOOK4) system using the Local descriptor and Image Signatures. 2. System performance degrades in case of fine details like buildings and food.

III. CONCLUSION

In this work, overview of various feature extraction techniques for Content Based Image Retrieval System is explained briefly. This paper classifies the current methods and summarizes their features. After having a review, it can be considered that no single method can be resulted as best or very good for all types of images or all the methods uniformly good for a specific image type. Considering all these limitations and major findings, CBIR system remains to be a challenging problem in image processing. Feature extraction techniques for CBIR system is still a pending problem in the world and more research need to be carried out for better estimation.

REFERENCES

[1] R. Choudhary, N. Raina, N. Chaudhary, R. Chauhan and R H Goudar," - An Integrated Approach to Content Based Image Retireval", International Conference on Advances in Computing, Communications and informatics (ICACC1), pp. 2402-2410, IEEE, 2014.

[2] B. Syam and Y. Rao, " An Effective Similarity Measure via Genetic Algorithm for Content Based Image Retrieval with Extensive Features", The International Arab Journal of Information Technology, Volume 10, pp. 143-151, 2013.

[3] A. N. Ligade and M. R. Patil, " Content Based Image Retrieval Using Interactive Genetic Algorithm with Relevance Feedback Technique – Survey", International Journal of Computer Science and Information Technologies, Vol. 5(4), pp. 5610-5613, IJCSIT, 2014.

[4] E. Walia, P. Saigal and A. Pal," Enhanced Linear Block Algorithm with Improved Similarity Measure", Canadian Conference on Electrical and Computer Engineering, pp. 1-7 , IEEE, 2014.

[5] P.P. Sarangi, B.S.P. Mishra, B. Majhi and S. Dehuri," Gray-level Image Enhancement Using Differential Evolution Optimization Algorithm", International Conference on Signal Processing and Integrated Networks (SPIN), IEEE, 2014.

[6] M. Agarwal," Integrated Features of Haar-like Wavelet Filters," 7th International Conference on Contemporary Computing, pp. 370-375, IEEE, 2014.

[7] D. Jeyabharathi and A. Suruliandi," Performance Analysis of Feature Extraction and Classification Techniques in CBIR", International Conference on Circuits, Power and Computing Technologies, pp. 1211-1214, IEEE, 2013.

[8] S. Ezekiel, M.G. Alford, D. Ferris and E. Jones," Multi-Scale Decomposition Tool for Content Based Image Retrieval", Applied Imagery Pattern Recognition Workshop: Sensing of Control and Augmentation (AIPR), pp. 1-5, IEEE, 2013.

Cite this article as: Kavita Chauhan, Shanu Sharma. "A Review on Feature Extraction Techniques for CBIR System." *International Conference on Information Engineering, Management and Security (2015)*: 110-114. Print.

[9] Z. Zhang, M. Zhao, B. Li and P. Tang," ColorPCA: Color Principal Feature Extraction Technique for COlor Image Reconstruction and Recognition", International Joint Conference on Neural Networks (IJCNN), pp. 1-7, IEEE, 2013.
[10] B. Syam, S. R. Victor J and Y.S. Rao," Efficient Similarity Measure via Genetic Algorithm for Content Based Medical Image Retrieval with Extensive Features", International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing, pp. 704-711, IEEE, 2013.
[11] A. N. Ligade and M. R. Patil," Optimized Content Based Image Retrieval Using Genetic Algorithm with Relevance Feedback Technique", International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR), pp. 49-54, TJPRC, 2013
[12] J.M. Ho, S.Y. Lin, C.W. Fann, Y.C. Wang and R.I Chang," A Novel Content Based Image Retrieval System using K-means with Feature Extraction", International Conference on Systems and Informatics (ICSAI), pp. 785-790, IEEE, 2012.
[13] A. Chadha, S. Mallik and R. Johar," Comparative Study and Optimization of Feature Extraction Techniques for Content Based Image Retrieval", International Journal of Computer Applications, Vol. 52-No. 20, pp. 35-42, IJCA, 2012.
[14] E. Rashedi and H.Nezamabadi-pour," Improving the Precision of CBIR Systems by Feature Selection Using Binary Gravitational Search Algorithm", International Symposium on Artificial Intelligence and Signal Processing (AISP), pp. 039-042, IEEE, 2012.
[15] R. Pighetti, D. Pallez and F. Precioso," Hybrid Content Based Image Retrieval combining Multi-objective Interactive Genetic Algorithm and SVM", 21st International Conference on Pattern Recognition (ICPR), pp. 2849-2852. ICPR, 2012.
[16] M. Madugunki, D.S. Bormane, S. Bhadoria and C.G. Dethe," Comparison of Different CBIR Techniques", International Conference on Electronics Computer Technology, pp. 372-375, IEEE, 2011.
[17] S. Selvarajah and S. R. Kodithuwakku," Combined Feature Descriptor for Content Based Image Retrieval", 6th International Conferencec on Industrial and Information Systems (ICIIS), pp. 164-168, IEEE, 2011.
[18] K.A.S. Abubacker and L.K. Indumathi," Atrribute Associated Image Retrieval and Similarity Reranking", Proceedins of the International Conference on Communication and Computational Intelligence, pp. 235-240, IEEE, 2010.
[19] S.G. Omar, M.A. Ismail and S.M. Ghanem," WAY-LOOK4: A CBIR System Based on Class Signature of the Images' Color and Texture Features", International Conference on Computer Systems and Applications, pp. 464-471, IEEE, 2009.

Cite this article as: Kavita Chauhan, Shanu Sharma. "A Review on Feature Extraction Techniques for CBIR System." *International Conference on Information Engineering, Management and Security (2015)*: 110-114. Print.