IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-10, Oct 2018 E-ISSN: 2347-2693

Hybrid Features For Content Based Image Retrieval System

A.D. Mahajan[1*], S. Chaudhary[2]

[1] Computer Engineering, Madhav University, Abu road, Shirohi, Rajasthan, India

[*]*Corresponding Author: mahajan.arpana@yahoo.com, Tel.: +91-7069024891*

Available online at: www.ijcseonline.org

Accepted: 15/Oct/2018, Published: 31/Oct/2018

***Abstract*—** The speedy progress in multimedia and imaging technology, the numbers of images uploaded and shared on the internet have increased. It leads to develop the highly effective image retrieval system to satisfy the human needs. The content-Based image retrieval (CBIR) system which retrieves the image based on the Low level features such as color, texture and shape which are not sufficient to describe the user's high level perception for images. Therefore reducing this semantic gap problem of image retrieval is challenging task. Some of the most important notions in image retrieval are keywords, terms or concepts. Terms are used by humans to describe their information need and it also used by system as a way to represent images. Here in this paper different types of features their advantage and disadvantages are described. We have carried out comparative analysis of different techniques used in our system to determine best suitable technique to be used for our proposed system. We have analyze the our proposed system on large image dataset and our approach gives high precision and required less computations which proves efficiency of our system. In our proposed system we have evaluated the performance of our feature extraction techniques i.e. FCH and GWT using precision and recall metric and compared the result with existing feature extraction approaches i.e. color moment and GWT. Implementation results show that the feature extraction techniques for the proposed system are better than the existing techniques. SVM Classifier also gives good accuracy using these feature extraction techniques.

***Keywords*—**CBIR, Color Moment, Fuzzy Color Histrogram, Gabor Wavelate, Support Vector Machine

I. INTRODUCTION

"Content" means Context allude to Shapes, Texture, Color or some other Information that can be gotten from Image itself. "Content Based" means Search examines the Contents of the picture as opposed to the metadata, for example, Tags, Description or Keywords related with the Image. An "Image Retrieval System" is a Computer System for Searching and Retrieving image from a Database of Digital Image. Many Features of Content Based Image Retrieval but Four of them are considered such as Color, Texture, Shape and Spatial Properties. "Low Level Feature" means Image like color, Texture, Shape can be extracted from the Image.

Owing to the popularity of social networking and media sharing websites numbers of images uploaded and shared on the internet have increased. It leads to the availability of extremely large quantities images that are tagged by users. Social media sharing websites such as flickr, facebook, instagram, twitter, pinterest etc has given freedom of sharing and tagging images to users. So the development of the highly effective image retrieval system to satisfy the human needs is required in spite of large scale of image data.

Image retrieval systems have developed from text based image retrieval (TBIR) to content based image retrieval (CBIR). In TBIR images are retrieved from the database based upon the text associated with images. While In CBIR systems are based on the visual properties of the images. It uses image low level visual features such as color, texture, shape, spatial information to retrieve images from large set of database.

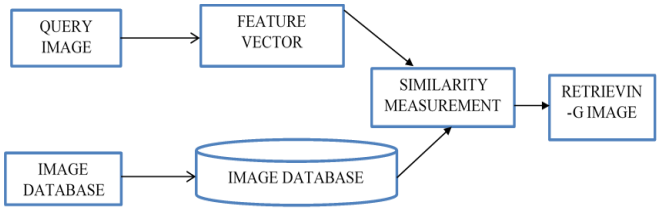

Figure 1. Bsis of CBIR System

II. RELATED WORK

A.Ramesh Kumar, D.Saravanan describe CBIR using color histograms technique with help of grid technique to improve the image retrieval performance. The color histogram for an image is constructed by quantizing the colors within the image and counting the number of pixels of each color. The feature vector of an image can be derived from the histograms of its color components and finally can

© 2018, IJCSE All Rights Reserved

set the number of bins in the color histogram to obtain the feature vector of desired size [1].

Sushant Shrikant Hiwale, Dhanraj Dhotre, Dr. G.R.Bamnote describe a CBIR framework which separates the components of computerized picture to recover comparable pictures from tremendous databases We have utilized Color Histogram, Color Auto-Correlogram, Color Moment, Gabor Wavelet and Discrete Wavelet change to concentrate picture highlights. The pictures are ordered utilizing Support Vector Machine (SVM) classifier which viably recognizes significant and superfluous pictures. The outcomes portray that proposed technique has better Accuracy and review rate contrasted with different strategies [5].

Ahmad Alzu'bi*, Abbes Amira, Naeem Ramzan presented a comprehensive survey on different techniques and recent research works in the CBIR domain. Firstly, this study has discussed the general retrieval framework which most of CBIR systems have adopted over the last 15 years. Secondly, the proposed approaches in each block of CBIR framework have been presented along with the impact of research shift from low-level to high-level processing. Thirdly, this study has identified the most recent advances that contribute in reducing the semantic gap [9].

Valentina Franzoni propose Context-based Image Similarity, a scheme for discovering and evaluating image similarity in terms of the associated groups of concepts. Several semantic proximity/similarity among image concepts and different concept ontology - WordNet Distance, Wikipedia Distance, Flickr Distance, Confidence, Normalized Google Distance (NGD), Pointwise Mutual Information (PMI) and PMING, have been considered as elementary metrics for the context. Comparing to Content Based Image Retrieval (CBIR), which measures the image content similarities by low level features, the proposed Context based Image Similarity outperformed CBIR in measuring the deep concept similarity and relationship of images. Experimental results, obtained in the domain of images semantic similarity using search engine based tag similarity, show the adequacy of the proposed approach in order to reflect the collective notion of semantic similarity [10].

Nidhi Goel, Priti Sehgal interest in the recent years has progressed to improve the performance of image retrieval (IR) systems by reducing the semantic gap between the low-level features and the high-level concept. In this paper, we proposed an approach to combine the two modalities in IR systems, i.e., content and text, while considering the semantics between the query image and the textual query provided by the user. For content matching, color feature is extracted and is represented using fuzzy color histogram (FCH). Using correlation, we combined the two modalities with late fusion approach. The proposed approach is assessed on standard annotated database. Higher values of precision and recall show better performance of the proposed approach. Moreover, the use of correlation helps in reducing the semantic gap and providing good results through better ranking of the similar images [15].

Subrahmanyam Murala, Anil Balaji Gonde, R. P. Maheshwari has combines color and texture features for content based image retrieval (CBIR). The color and texture features are obtained by computing the mean and standard deviation on each color band of image and sub-band of different wavelets. The standard Wavelet and Gabor wavelet transforms are used for decomposing the image into sub-bands. The retrieval results obtained by applying color histogram (CH) + Gabor wavelet transform(GWT) to a 1000 image database demonstrated significant improvement in precision and recall, compared to the color histogram (CH), wavelet transform (WT), wavelet transform and color histogram (WT + CH) and Gabor wavelet transform (GWT) [19].

III. PROPOSED APPROACH

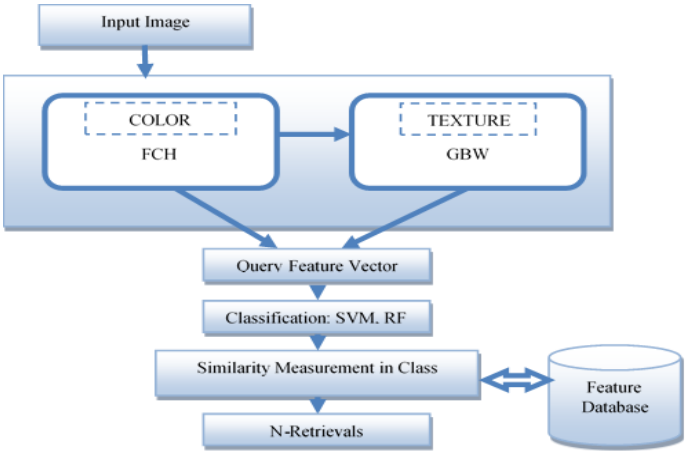

Figure 2. Proposed Block diagram

Propose hybrid image retrieval system block diagram shows in figure 2. Which will retrieve images based on image low level features. Workflow of our proposed image retrieval system is decided to use SVM as a classified as it can give accurate classification compare to other classifier. In low level feature extraction for color feature FCH required less storage, less computation hence faster and also invariant to rotation, scaling and translation, insensitive to noise so we have extracted color feature using FCH. So that it can bridge the semantic gap between human perception and machine understanding between human perceptions.

For Color feature extraction as we have decided to use Fuzzy Color Histogram (FCH). We have use FCH and extracted the color features from the images. For FCH we have used the

© 2018, IJCSE All Rights Reserved

27 fuzzy rules and for that used Mamdani type fuzzy inference system given in [19]. It gives the output as fuzzy histogram consists of only 10 bins. Output of the 10 bins histogram consist of black, dark grey, red, brown, yellow, green, blue, cyan, magenta and white respectively. So we got 1-D feature vector of size for each image as a color feature of an image.

For texture feature extraction using Gabor wavelet transform first of all color image is converted in to grayscale image. Then Gabor filter is constructed using different number of frequencies and different phase angles. In last step, mean and standard deviation are calculated and store in feature vector. Here we have use 5 different scales and 6 orientations so we obtained total 60 features per image as a texture feature.
Initially in training phase using FCH and Gabor wavelet feature are extracted of database images then combined resultant feature vectors have given to SVM for training. SVM is trained using predefined classes of image database. After that in testing phase in our system we have taken query image as well as number of related tags or keywords from the use. So first using FCH and Gabor wavelet query image feature are extracted after that combined feature vector has given to SVM for finding the appropriate class for the query image.

IV. RESULTS AND DISCUSSION

As in portion evaluated the performance of feature extraction techniques which are used in our proposed system. To evaluate the performance first of all we have extracted the color and texture features of the query image as well as database image as discussed in previous section. So we get feature vector of length 70 for each image.

To measure the performance of the feature extraction technique we have randomly taken 5 images from each class as query image. Then calculated the precision and recall of each 5 query image of every class then calculated the average precision and recall of each class by taking average of precision and recall of 5 query image of each class. The evaluation was carried out with the number N of the retrieved images set as 10 to compute Precision and Recall.

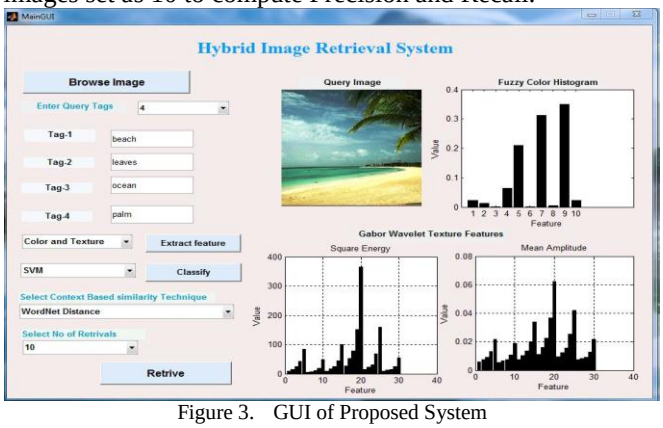

Figure 3. GUI of Proposed System

Figure 4. Retrival Images

Table 1: Distance Analysis

Query Image Name	Image Retrieval Time using Euclidean Distance (S)	
	GWT + Color Moment [11]	GWT + FCH
coast_n203069.jpg	3.58	5.72
forest_for22.jpg	2.18	5.28
highway_bost169.jpg	2.08	5.17
insidecity_bo109.jpg	2.36	5.13
mountain_moun44.jpg	2.19	5.30
opencountry_file37.jpg	2.91	5.10
street_art764.jpg	1.99	5.37
Tallbuilding_a487092.jpg	2.25	5.05
Average	**2.44**	**5.26**

From the table we can see that the feature extraction techniques given in [11] required less time than the feature extraction techniques used in our system. But the difference is not major. Our proposed system's feature extraction technique required only 2.8 seconds more than existing one and give better precision and recall as we have already shown in Table 4.3. Therefore despite of our proposed system's feature extraction technique required on average 2.8 seconds more than existing [11], it gives nearly 15 % more precision than existing techniques. Hence our proposed system's feature extraction techniques are proved more efficient than existing.

Table 2: Precision and Recall Analysis

Categories	N=10		N=30		N=50	
	Precision	Recall	Precision	Recall	Precision	Recall
Cost	0.82	0.12	0.81	0.37	0.79	0.60
Forest	0.94	0.11	0.91	0.32	0.89	0.52
Highway	0.82	0.13	0.81	0.39	0.78	0.65
Inside City	0.80	0.13	0.78	0.37	0.76	0.80
Mountain	0.86	0.14	0.80	0.39	0.76	0.61
Open Country	0.88	0.12	0.86	0.35	0.84	0.57
Street	0.82	0.14	0.76	0.38	0.74	0.65
Tall Building	0.98	0.19	0.94	0.32	0.92	0.51
Symbols	1.0	1.0	0.33	1.0	0.20	1.0
Average	**0.865**	**0.135**	**0.834**	**0.360**	**0.810**	**0.6137**

© 2018, IJCSE All Rights Reserved

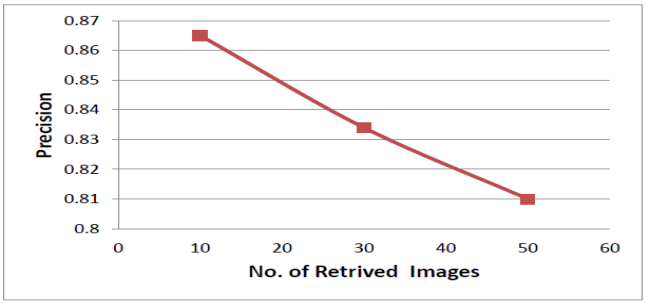

Figure 5. Precision graph

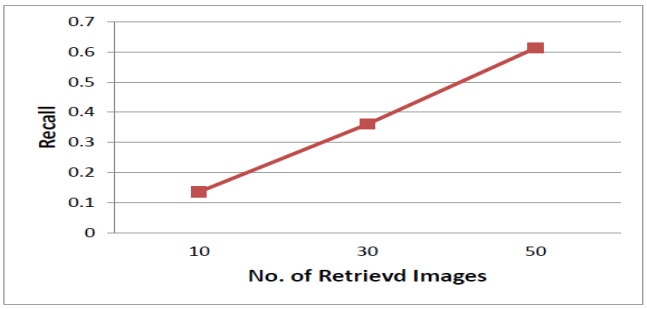

Figure 6. Recall graph

Here we have considered Euclidean distance which is most widely used similarity measure used for measure the low level feature based similarity. Following Figure shows retrieval result of proposed system with context based similarity techniques and without context based similarity techniques.

V. CONCLUSION AND FUTURE SCOPE

We have studied various image retrieval techniques and identified the issues in existing image retrieval system. Most of the image retrieval systems have semantic gap problem that is the lack the understanding of user's intention behind the query image. To overcome this issue and provide accurate retrieval of images we have proposed hybrid image retrieval system which is based on image color features as well as image texture features. We have evaluated the performance of our feature extraction techniques i.e. FCH and GWT using precision and recall metric. From the results, we have analysed that our proposed system gives high precision and recall value despite of large size of image database. Our proposed system easily understand the human requirement behind query image using query tags and higher precision is obtained compare to image existing similarity methods.

REFERENCES

[1] ElAlami, M.E."A new matching strategy for content based image retrieval system." Applied Soft Computing 14 (2014): 407-418.

[2] Murala, Subrahmanyam, Anil Balaji Gonde, and Rudra Prakash Maheshwari. "Color and texture features for image indexing and retrieval." In Advance Computing Conference, 2009. IACC 2009. IEEE International, pp. 1411-1416. IEEE, 2009.

[3] Zhang, Dengsheng, Aylwin Wong, Maria Indrawan, and Guojun Lu. "Content-based image retrieval using Gabor texture features." In IEEE Pacific-Rim Conference on Multimedia, University of Sydney, Australia. 2000.

[4] Howarth, Peter, and Stefan Rüger. "Evaluation of texture features for content-based image retrieval." In Image and Video Retrieval, pp. 326-334. Springer Berlin Heidelberg, 2004.

[5] Lin, Chuen-Horng, Rong-Tai Chen, and Yung-Kuan Chan. "A smart content-based image retrieval system based on color and texture feature." Image and Vision Computing 27, no. 6 (2009): 658-665.

[6] Saad, Michele. "Low-level color and texture feature extraction for content-based image retrieval." Final Project Report, EE K 381 (2008): 20-28.

[7] Jhanwar, N., Subhasis Chaudhuri, Guna Seetharaman, and Bertrand Zavidovique. "Content based image retrieval using motif cooccurrence matrix."Image and Vision Computing 22, no. 14 (2004): 1211-1220.

[8] Yue, Jun, Zhenbo Li, Lu Liu, and Zetian Fu. "Content-based image retrieval using color and texture fused features." Mathematical and Computer Modelling54, no.3 (2011): 1121-1127.

[9] Xianzhe Cao and Shimin Wang, "Research about Image Mining Technique," in proc. Springer ICCIP,2012, pp.127-134.

[10] Ahmad Alzu'bi, Abbes Amira and Naeem Ramzan, "Semantic content-based image retrieval: A comprehensive study," in Elseveir Journal of Visual Communication and Image Representation, Vol. 32, pp. 20-54 ,2015

[11] V. Franzoni, A. Milani, S. Pallottelli, C. H. C. Leung and Yuanxi Li, "Context-based image semantic similarity," in proc. IEEE twelveth international conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2015, pp. 1280-1284.

[12] Valentina Franzoni, Clement H.C. Leung, Yuanxi Li,Paolo Mengoni and Alfredo Milani," Set Similarity Measures for Images Based on Collective Knowledge," in proc. Springer ICCSA,2015,pp.408-417

[13] Elalami,"A New Matching Strategy for Content Based Image Retrieval System," in ACM Appl. Soft Comput., vol. 14,pp. 407-418,2014

[14] Mohsen Sardari Zarchi, Amirhasan Monadjemi and Kamal Jamshidi, " A concept- based model for image retrieval systems," in Elsevier Computers & Electrical Engineering, vol. 46 , pp. 303-313, 2015

[15] N. Goel and P. Sehgal, "Weighted semantic fusion of text and content for image retrieval," in proc. IEEE International Conference Advances in Computing, Communications and Informatics (ICACCI), 2013 , pp. 681-687.

[16] K. Singh, K. J. Singh and D. S. Kapoor, "Image Retrieval for Medical Imaging Using Combined Feature Fuzzy Approach," in proc. IEEE International Conference on Devices, Circuits and Communications (ICDCCom), 2014, pp. 1-5

[17] N. Goel and P. Sehgal," Image Retrieval Using Fuzzy Color Histogram and Fuzzy String Matching: A Correlation-Based Scheme to Reduce the Semantic Gap", in proc. Springer Intelligent Computing, Networking, and Informatics,2014, pp. 327- 341.

[18] C.-H. Lin, R.-T. Chen and Y.-K. Chan, "A smart content-based image retrieval system based on color and texture feature", in Elsevier Image and Vision Computing, Vol. 27 , pp. 658–665,2009.

[19] Zhi-chun huang, Patrick P. K. Chan, Wing W. Y. Ng, D aniel s. Yeung" Content- based image retrieval using color moment and Gabor texture feature", Proc. IEEE Ninth international Conference on Machine Learning and Cybernetics, 2010, pp 719-724.

[20] Nizampatnam Neelima and E. Sreenivasa Reddy, "An Efficient Multi Object Image Retrieval System Using Multiple Features and

© 2018, IJCSE All Rights Reserved

SVM", in proc. Springer Advances in Intelligent Systems and Computing, Vol. 425,2015, pp 257-265.
[21] Subrahmanyam Murala, Anil Balaji Gonde and R.P. Maheshwari, "Color and Texture Features for Image Indexing and Retrieval", in proc. IEEE international advanced computing conference , 2009, pp. 1411-1416.
[22] L. Wu, X. Hua, N. Yu, W. Ma, and S. Li, ''Flickr distance: A relationship measure for visual concepts,'' in IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, pp. 863– 875, 2012.
[23] L. Wu, X.-S. Hua, N. Yu, W.-Y. Ma, and S. Li, "Flickr Distance," in Proc. 16th ACM International Conf. Multimedia, 2008, pp. 31-40
[24] G. A. Miller, "Wordnet: a lexical database for english", in ACM Communications of the, 38(11):39–41, 1995.
[25] A. Budanitsky and G. Hirst, "Semantic Distance in Wordnet: An Experimental, Application-Oriented Evaluation of Five Measures," Proc. WordNet and Other Lexical Resources, 2001
[26] S. S. Hiwale, D. Dhotre and G. R. Bamnote, "Quick interactive image search in huge databases using Content-Based image retrieval," in proc. IEEE International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2015, pp. 1-5.
[27] K. Konstantinidis, A. Gasteratos, I. Andreadis, "Image Retrieval Based on Fuzzy Color Histogram Processing", in Elsevier Optics Communications, Vol. 248, pp. 375-386, 2005.
[28] J. Liu, Z. Li, J. Tang, Y. Jiang and H. Lu, "Personalized Geo-Specific Tag Recommendation for Photos on Social Websites," in IEEE Transactions on Multimedia, vol. 16, pp. 588-600, 2014.
[29] Oliva and A. Torralba. Modeling the shape of the scene: a holistic representation of the spatial envelope. Int. Journal of Computer Vision., 42, 2001.
[30] X. Li, T. Uricchio, L. Ballan, M. Bertini, C. Snoek, and A. Del Bimbo, "Socializing the semantic gap: A comparative survey on image tag assignment, refinement and retrieval," in ACM Computing Surveys, 2016, in press.
[31] Cinemast,"Semantic Gap in image analysis",
[32] Internet: https://en.wikipedia.org/wiki/Semantic_gap, Apr. 2016.
[33] Dong-Chul Park, "Image Classification Using Naïve Bayes Classifier", in International Journal of Computer Science and Electronics Engineering, vol. 4 , pp. 135-139, 2016
[34] MATLAB and Statistics Toolbox Release 2013a, The MathWorks, Inc., Natick, Massachusetts, United States
[35] Finlayson and Mark Alan "Java Libraries for Accessing the Princeton Wordnet: Comparison and Evaluation" in Proceedings of the 7th International Global WordNet Conference, 2014, pp. 78-85 .
[36] Ted Pedersen, WordNet::Similarity. [Online]
[37] Available at: http://wn- similarity.sourceforge.net/ [Accessed on 15 may 2017].

Authors Profile

Mrs. Arpana D. Mahajan Pursuing Ph.D. in Computer Engineering from Madhav University, Abu Road, Sirohi, Rajasthan, India in year 2018. She is currently working as Assistant Professor Computer Engineering Department in, Sigma Institute of Engineering, Vadodara, India since 2013. She has published more than 22 research papers in reputed international journals and 2 in National conferences including Scopus and conferences including IEEE, Springer and it's also available online. Her main research work focuses on Image Processing,Big Data Analytics and Cloud Computing. She has 5 years of teaching experience and 5 years of Research Experience.

Dr. Sanjay Chaudhary is working as a Professor, Department of Computer Science at Madhav University, Abu Road, Sirohi, Rajasthan. He has 20 years of experience in his field. He has been previously working as a Principal in S.S. Polytechnic College, Udaipur and as Head of Department in S.S. College of Engineering, Udaipur. He holds an experience of 10 years in teaching and another 10 years in industries, serving as Technical Manager and Senior Consultant. He has conducted his research in finding out the applications, perspectives and prospects of E-commerce in Rajasthan and have supervised 7 scholars (awarded) and 8 scholars (pursuing) in their Doctoral degree. There are more than 25 research papers in various National and International Journals in his name. He have attended numerous Seminars, Conferences and Faculty Development Programmes. He is famous for his Motivational, Academic and Administrative skills. His motive is to spread practical knowledge in order to flourish the students and the institute as a whole.

© 2018, IJCSE All Rights Reserved