Combination of Local, Global and K-Mean using Wavelet Transform for Content Base Image Retrieval

Ekta Gupta
Deptt. ofComputer Science andEngg.
Institute of Technology &Management, Gwalior- India

Rajendra Singh Kushwah
Deptt.ofComputer Science andEngg.
Institute of Technology &Management, Gwalior- India

ABSTRACT
With the ever expanding database and advancement of technology in the fields of Data mining, remote sensing and management of Earth resources, Crime prevention, Weather Forecasting, E-commerce, Medical Imaging, and soon. The Content Based Image Retrieval Technique is becoming more and more indispensable and vital. The paper proposes Content Based Image Retrieval technique incorporating WBCHIR (Wavelet Based Color Histogram Image Retrieval) which utilizes features of an image like Color and Texture.

The shape and shade features are extracted in the course of Wavelet Transform and Color Histogram, and the arrangement of these features is the vital the scaling and conversion of objects into an image. Now, it is being presented for the first time in our era that techniques such as Feature Extraction, segmentation and Grid, K-means module and k-nearest neighborhood module are integrated together to build the CBIR System. It is a hybrid of Global and Local Features method with K-means Clustering algorithm. Given a set of instruction images, a K-means Clustering Algorithm is applied to cluster the regions on the basis of these features. These features, which they identify as "Blobs", compose the expressions for the set of images. Each of these "blobs" is assigned an exclusive integer to serve as its identifier (analogous to a word's ASCII representation.

In this paper, we present a technique for integration of Wavelet Based Color Histogram Image Retrieval (WBCHIR) using color and texture features into Content Based Image Retrieval. The Evaluation between the images is ascertained by means of a Distance Function. The concept proposed in this paper will provide better results as compared to other retrieval methods in terms of average accuracy. Moreover, the computational steps are summarily consistent with the use of Wavelet transformation.

Keywords
CBIR, K-means, DWT, Global Feature, Local Feature.

1. INTRODUCTION
With the ever increasing number of digital images, development of an effectively optimum digital management system and retrieval method for a large number of files is becoming inevitable. Ordinarily, an image is specified by adhering a text and notation (a set of keywords) to it. Then the image retrieval procedure is to match the text and notation of a query image with those images in the image databases [1][2]. Subsequently, the method exhibits candidate images with similar text annotations.

Content Based Image Retrieval (CBIR) is browsing, penetrating, and direction-finding of images from huge databases based on their illustration/visual contents. Consequently, CBIR has been an area of attraction for innovative spirits for beyond a decade. Many CBIR systems have been evolved, QBIC, Simplicity and Blob World to name a few. A comprehensive survey of CBIR methodology is evident in the Conventional CBIR System's application of law of low level features like color, texture, shape and spatial location of objects to guide and retrieve images from databases [3][4]. Low level features can be global or local (region based). Global feature based CBIR falters in evaluating the regions or objects in which a user may be specifically concerned. Then the Region Based Image Retrieval (RBIR) becomes highly optimal in delivering user's obligation.

Galvanized by the accomplishments of text-based document retrieval techniques, where an upturned indexing arrangement is built to obtain highly efficient penetration; BAG-OF-VISUAL (BOV) words illustration is considered to represent an Image that is analogous to the text-based indexing arrangement and is extensively employed in CBIR applications[5][6]. Therefore, different text-based indexing in sequence retrieval tech-inquests[6], such as the vector space model, language-string model and Okami-BM25 can be readily exploited with the CBIR. There are two key issues which influence the recital of BOV based CBIR applications. The first is how to select a healthy ranking model for so many mature text information retrieval techniques for CBIR [7][8], convinced ranking model may be more effectual for some image databases and less for the others. Not a single ranking model has been able to maintain its superiority over the others consistently. Secondly, many kinds of such features have been designed and can be utilized to boost the presentation, assortment of visual features is again a key issue in CBIR.

The primary set of facial appearances is BOV based ranking features which are position scores resulting from several adhoc ranking models. Since ranking features in text in sequence retrieval have always resulted from dissimilar fields of documents, such as designation, theoretical, and a corpse, we split images into more than a few regions according to their most important degrees to generate image ''fields'', and BOV-based ranking features are extracted from each image ''meadow''. For each block, BOV-based ranking features are computed and all the blocks' features are concatenated. Finally, image similarities of global features are modified as a complement to the BOV-based features. To maintain the system's scalability, the global features are rehabilitated into binary codes with LSH [9] for efficient similarity calculation. Global features are also entrenched into lower dimensional features with a linear multi view, embedding (LME) [10] to compute the visual Page Rank [11] of database images.

In this paper, we suggest a hybrid method for CBIR that takes into account global, local and k-mean features of an image. Towards this, Stationary Wavelet Transform (SWT) is applied to the input image to take out horizontal, vertical and diagonal detail matrices. SWT is used because of its translational invariant stuff. After this global textural features are mined using Gray level Co-occurrence Matrix for each of these sub-matrices. To help the retrieval process, a local descriptor is also computed by splitting the image into sub-regions. Finally Euclidean distance is applied to retrieve the relevant results from the given sample of unknown class into one of n classes.

Wavelet transforms can diminish the magnitude of images and permit good declaration in time and frequency, and multi-resolution techniques can be used in wavelet transforms, so the scheme that color feature is extracted from a low frequency band of wavelet transform and texture feature from high incidence bands is not only straightforward and rapid to put into practice but also can lead to good retrieval consequences. Color feature is one of the most imperative low-level illustration features, so color index is one of the basic methods of image retrieval. Texture feature responds to the aspect in sequence of local image and measure the relation of pixels in local regions, so the texture feature is also very significant. Only one of the two features is used in most of the former retrieval systems, those using color feature only pay no attention to the local details, and those using only texture feature ignored the universal color information, so these systems cannot give acceptable consequences.

2. RELATED WORKS

Journalism survey is significant for understanding and advancing more information about an exact area of the subject. Towards this, a brief overview of the dissimilar obtainable techniques for CBIR is presented. In order to overcome the drawbacks of previous methodologies, content based image retrieval (CBIR) was primarily presented by Kato in 1992. He proposed a color-texture and color-histogram based image retrieval system (CTCHIR)[12]. They proposed three image facial appearances, based on color, texture and shape, a color co-occurrence matrix (CCM), difference between pixels of scan pattern (DBPSP) and color histogram for K-mean (CHKM) respectively, and a method for image retrieval by integrating CCM, DBPSP and CHKM together to improve image discovery rate and simplify scheming of image retrieval. From the experimental results, they found that, their wished-for method outpaces the two[13][14]methods. Raghupathi et al. made a relative study of image retrieval techniques (methods utilizing diverse features) like color histogram, Gabor Transform, color histogram+ gab-we transform, Contour-let Transform and color histogram + contour-let transform. Another technique[15] that proposed CBIR system based on the color, texture and shape features by partitioning the image into tiles. The features computed on tiles serve as local descriptors of color and texture features. The color and texture investigations are analyzed by means of two level grid frameworks and the shape distinction by Gradient Vector Flow. The comparison of investigation result of the proposed method with other system[16][17] found that, the anticipated retrieval system gives better recital than the others[18]. Proposed CTDCIRS (color-texture and dominant color based image retrieval system), they employed collectively three of the features like Motif co-occurrence matrix (MCM), Difference between Pixels of Scan Pattern (DBPSP) which describes the texture features and Dynamic Dominant Color (DDC) to take out color feature for image retrieval. They compared its outcomes with the work of[12] and They [14] found that their method gives improved retrieval results than others.

3. METHODS

Feature extraction is a significant step in any CBIR system. The retrieval precision depends profoundly upon the extent to which the given feature vector is able to represent the image under consideration. In our technique, we have measured both local & global features and bag of visual (BOV) words.

3.1. Computation of Global Features

3.1.1 Stationary Wavelet Transformation

Discrete Wavelet Transformation (DWT)[19] is utilized for altering an image from a spatial sphere into recurrent space. The wavelet transform represents an occupation as superposition of relatives of basic functions called wavelets. Wavelet transforms take out information from the signal at diverse scales by passing the signal through low pass and high pass filters. Wavelets make available multi-resolution potential and good power compaction. Wavelets are fiery with a high feeling to shading focus moves and can confine both composition and shape data creatively. The wavelet transforms can be computed linearly with time and thus allowing for very quick algorithms[20]. DWT decomposes a signal into a set of foundation Functions and Wavelet Functions. The wavelet change count of a two-dimensional picture is likewise a multi-determination move toward, which applies recursive sifting and sub-examining. At each level (scale), the image is fragmented into four incidence sub-bands, LL, LH, HL, and HH where L denotes low frequency and H denotes high frequency as shown in Figure 1.

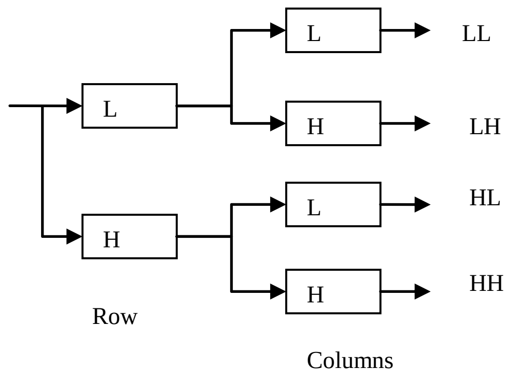

Fig.1. Stationary Wavelet Decomposition of a two-dimensional image

3.1.2 Gray level Co-occurrence Matrix

A communal technique in texture psychoanalysis involves the computation of GLCM as a second-order texture amount [21]. GLCM defines the incidence of one gray tone apparent in a particular spatial linear relationship with another gray tone, within the area under study. Several arithmetical constraints can be extracted from the GLCM. Some of these constraints are connected to specific first-order statistical concepts, such as variance and contrast, and have an apparent textural connotation. Other parameters hold textural statistics, but are associated with supplementary than a single explicit textural meaning.

Six textural constraints are reflected: energy, inverse difference moment, variance, correlation, entropy and contrast. The purpose of this study was to evaluate categorical differences among these six constraints and to assign a textural meaning to each constraint, in order to carry out an image-independent feature selection among them. An initial image fragmentation is carried out from 8 x 8 pixel segments. It must be emphasized that GLCM computation must be performed within a realistic classification method. For instance, texture analysis should not be applied to those parts of the image that could be confidential by a dependable non textural (e.g., spectral) investigation.

3.2 Computation of Local Features

A global descriptor uses the visual features of the complete image, while a local descriptor takes into account the regions or objects to portray the image. To compute local features we first fragment the processed image into blocks and attain a descriptor for each block. Fig. 2 shows the method of splitting the image into three different sub-regions.

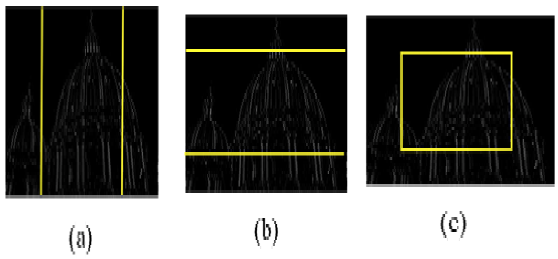

Fig.2. Different templates for splitting images into (a) Vertical crop;(b) Horizontal crop;(c) Cenntral crop

After splitting the image into sub-regions, two statistical measures are computed for each region. These measures are mean (μ) and standard deviation (σ) [22].

$$\mu = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} f(i,j)}{M \times N}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{M}(|f(i,j)| - \mu)^2}{M \times N}}$$

3.3 Bag of Words (BOW)

Prearranged situate of training images, a K-means clustering algorithm is serviceable to cluster the regions on the foundation of these features.These clusters which they call ‘‘blobs’’ constitute the glossary for the set of images. Each blob is assigned an explicit and a unique integer to hand round as its identifier (akin to a word’s ASCII depiction) [23].

All images in the training set can now be exemplified as a set of blobs from these expressions. Given a new test image, it can be segmented into regions and region features can be figured out. The blob which is the neighboring to it in the bunch space is assigned to it. The essential thought of Bag of Words is to portray each image as an order fewer anthology of local features. For dense illustration, a visual vocabulary is usually fabricated to define BOW through the clustering of features. With the visual statements, we can describe the image as a feature vector according to the attendance or count of each visual word.Underneath the directed learning stage, the peculiarity vector structures the fundamental visual prompt for article and scene classification. In a Bow approach, the classification stage transforms into a histogram based ordering, and in spite of the fact that the sample is basic, it doesn't limit any geometrical data.

3.4 Features Vector Generation

Having computed the local as well as global features[24] for the query image, the next step is to amalgamate all these features into a single feature vector which will be used for comparison during similarity matching. We acquire a total of 12 global features, 4 from each GLCM computed over horizontal, vertical and diagonal sub-matrices. Thus the dimension of global feature vector f(G) is 1×12 and is represented as in (Fig. 1). In addition, we consider 2 local measures for each region of image (Fig. 2) Yielding in total a 6 element local feature vector for each of horizontal, vertical and diagonal sub-matrices. The dimension of feature vector for each sub-matrix is 1×6. Hence the final dimensions of the local feature vector f(L) will be 1×18 and is represented as in (Fig. 2). The next step is to concatenate both the feature vectors to generate a single feature vector f(Query) which will represent the query image. The dimensions of this main feature vector will be 1×30 and is represented as in (Fig. 3). A similar feature vector f(Database) having identical dimensions to query image’s feature vector is computed for every database image and is used during similarity matching process.

4. PROPOSED WORK

The color of an image is represented using any of the conventional color spaces like RGB, XYZ, YIQ, L*a*b*, U*V*W*, YUV and HSV. Apply DWT convert at each level (scale), the image is alienated into four frequency sub-bands, LL, LH, HL, and HH where L denotes low frequency and H denotes high frequency. Partition the image into sub-regions: horizontal, vertical and center region:

$$fG=(f_{GH}, f_{GV}, f_{GD})$$

$$f_{GH}=(f_{GHene}, f_{GHcor}, f_{GHcon}, f_{GHidm})$$

$$f_{GV}=(f_{GVene}, f_{GVcor}, f_{GVcon}, f_{GVidm})$$

$$f_{GD}=(f_{GDene}, f_{GDcor}, f_{GDcon}, f_{GDidm})$$

$$fL=(f_{LH}, f_{LV}, f_{LD})$$

$$f_{LH}=(f_{LHh\_crop}, f_{LHv\_crop}, f_{LHc\_crop})$$

$$f_{LV}=(f_{LVh\_crop}, f_{LVv\_crop}, f_{LVc\_crop})$$

$$f_{LD}=(f_{LDh\_crop}, f_{LDv\_crop}, f_{LDc\_crop})$$

After separating the image into sub-regions, two arithmetic actions are computed for each region. These measures are mean (μ) and standard deviation (σ). Gray Level Co-occurrence Matrix (GLCM) evaluated over horizontal, vertical and diagonal sub-matrices. Given a set of training images, a K-means clustering algorithm is functional to cluster the regions on the foundation of these features (Fig.6).

$$f_{QUERY}=(fG,fL)$$

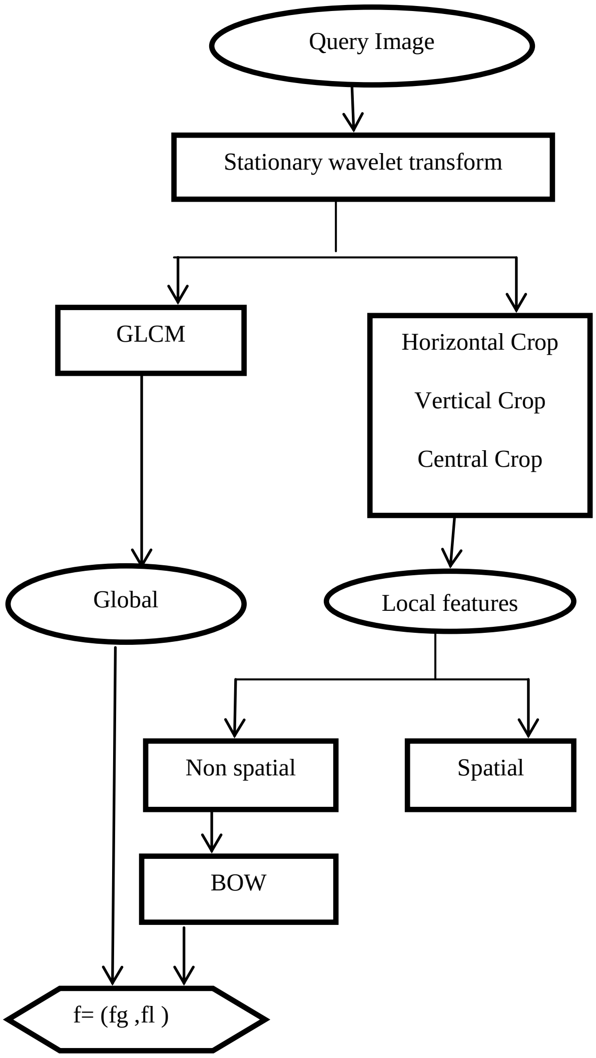

Fig.3.Combination of local, global and k-mean features

5. RESULT ANALYSIS

5.1 Performance Evaluation

For the purpose of computation, we have utilized two extensively used metrics of Precision and Recall. Precision is a measure of aptitude of CBIR algorithm to get back only pertinent images, while Recall decides the ability of CBIR algorithm to retrieve all pertinent images as defined respectively.

$$P = \frac{No.\ of relevantimageretrieved}{Totalnumberofimageretrieved}$$

$$R = \frac{No.\ of relevantimageretrieved}{numberofimageinthedatebese}$$

5.2 Similarity Measurement

The quality skin tone of the inquiry image and all the images in database are stored in conditions of feature vectors. Based on these vectors the comparison or variation between the images is computed. Methods used to organize images, moreover, calculate the distinction or comparison between two vectors. Two vectors with small differentiation will have large resemblance. We have used Euclidean distance which is the most conventional metric for determining the detachment between two vectors. Given two vectors *Q* and *D*, where

$$Q = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} and D = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$$

Then the Euclidean distance flanked by them is given by

$$ED = \sqrt{\sum_{i=1}^{n} (Q_i - D_i)^2}$$

6. ACKNOWLEDGMENT

I express my sincere gratitude to the RGPV University, Bhopal for giving me the opportunity to work on the thesis during my final year of M.Tech. Thesis work is an important aspect in the field of engineering.

I would like to thank ***RajendraKushwah***, Associate Professor in the Department of Computer Science,Institue Of Technology & Management College, Gwalior for his kind Support.

7. CONCLUSION

In this paper, we open usual methodology for Content Based Picture Recovery by amalgamating the shading and surface gimmicks called Wavelet-Based Shading Histogram Picture Recovery (WBSHIR). The judgment between the images is ascertained by means of aloofness meaning. We present a technique for integration of Wavelet Based Color Histogram Image Retrieval (WBCHIR) using color and texture features into Content Based Image Retrieval. The Evaluation between the images is ascertained by means of a Distance Function. The concept proposed in this paper will provide better results as compared to other retrieval methods in terms of average accuracy. Moreover, the computational steps are summarily consistent with the use of Wavelet transformation. The proposed paper results better than most the other retrieval methods in terms of average accuracy. Moreover, the computational steps are effectively abridged with the use of Wavelet transformation.

8. REFERENCES

[1] J.Philbin, O.Chum, M.Isard, J. Sivic, A. Zisserman, Object retrieval with large vocabularies and fast spatial matching, in: IEEE Con- faience on Computer Vision and Pattern Recognition, 2007 (CVPR’07), IEEE, 2007, pp-1–8.

[2] J.Yu, D.Liu, D.Tao, H.Seah, Complex object correspondence construction in two-dimensional animation, IEEE Transactions on Image Processing 11 (2012) pp. 3257–3269.

[3] M.Subrahmanyam,R.Maheshwari,R. Balasubramanian, Local maximum edge binary patterns: a new descriptor for image retrieval and object tracking, Signal Processing 92 (6) (2012) pp-1467–1479.

[4] X. Tian, D. Tao, X. Hua, X. Wu, Active reranking for web image search, IEEE Transactions on Image Processing 19 (3) (2010) pp-805–820.

[5] Tian, X.,Tao D. and Rui, Y. 2011. Sparse Transfer Learning for Interactive Video Search Reranking.

ACM Trans. Multimedia Computing, Communications and Applications.

[6] C.Manning, P.Raghavan, H.Schutze, Introduction to Information Retrieval, Cambridge University Press, Cambridge, 2008. Pp- 256-280.

[7] B. Geng, L. Yang, C. Xu, A study of language model for image retrieval, in: IEEE International Conference on Data Mining Work- shops, 2009 (ICDMW'09), IEEE, 2009, pp-158–163.

[8] B.Geng, L.Yang, C.Xu, X.Hua, Ranking model adaptation for domain- specific search, in: Proceedings of the 18th ACM Conference on Information and Knowledge Management, ACM, 2009, pp-197–206.

[9] M. Datar, N.Immorlica, P. Indyk, V. Mirrokni, Locality-sensitive hashing scheme based on p-stable distributions, in: Proceedings of the 20th Annual Symposium on Computational Geometry, ACM, 2004, pp-253–262.

[10] Y.Li, B.Geng, Z.Zha, D.Tao, L.Yang, C.Xu, Difficulty guided image retrieval using linear multiview embedding, in: Proceedings of the 19th ACM International Conference on Multimedia, ACM, 2011, pp-1169–1172.

[11] Y. Jing, S. Baluja, Pagerank for product image search, in: Proceeding of the 17th International Conference on World Wide Web, ACM, 2008, pp-307–316.

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

[13] N. Jhanwar, S. Chaudhurib, G. Seetharamanc and B. Zavidovique, "Content based image retrieval using motif co-occurrence matrix", Image and Vision Computing, Vol.22, pp-1211–1220, 2004.

[14] P.W. Huang and S.K. Dai, "Image retrieval by texture similarity", Pattern Recognition, Vol. 36, pp- 665–679, 2003.

[15] P.S. Hiremath and J. Pujari, "Content Based Image Retrieval based on Color, Texture and Shape features using Image and its complement", 15th International Conference on Advance Computing and Communications. IEEE. 2007. pp- 262-282.

[16] Y.Chen and J.Z. Wang, "A Region-Based Fuzzy Feature Matching Approach to Content Based Image Retrieval", IEEE Transactions on Pattern Analysis and Machine Intelligence.Vol. 24, No.9, pp-1252-1267, 2002.

[17] Y.Rubner, L.J.Guibas and C.Tomasi, "The earth mover's distance, multidimensional scaling, and color-based image retrieval" , Proceedings of DARPA Image understanding Workshop. pp-661-668, 1997.

[18] M.B. Rao, B. P. Rao, and A. Govardhan, "CTDCIRS: Content based Image Retrieval System based on Dominant Color and Texture Features", International Journal of Computer Applications, Vol. 18– No.6, pp-0975-8887, 2011.

[19] A.Natsev, R.Rastogi and K.Shim, "WALRUS: A SimilarityRetrieval Algorithm for Image Databases", In Proceeding. ACMSIGMOD Int. Conf. Management of Data, pp-395–406, 1999.

[20] R.C.Gonzalez, R.E.Woods and S.L, Eddins.Digital Image Processing Using MALAB, By Pearson Education, 2008.pp-426-437.

[21] R. M. Haralick, K. Shanmugam, I. Dinstein, "Textural Features for Image Classification", IEEE Transactions on Systems, Man, and Cybernetics, pp.610-621, 1973.

[22] Rafael C.Gonzalez, Richard E.Woods, Digital Image Processing, Pearson Education, Third Edition, Copyright © 2008.pp-372-378.

[23] Structured representations in a content based image retrieval context RomainRaveaux, Jean-Christophe Burie , Jean-Marc Ogier. pp-1252-1268

[24] Fusion of Local and Global Features using Stationary Wavelet Transform for Efficient Content Based Image Retrieval 2014 IEEE Students' Conference on Electrical, Electronics and Computer Science.