RESEARCH ARTICLE OPEN ACCESS

Evaluation of Euclidean and Manhanttan Metrics In Content Based Image Retrieval System

Gunjan Khosla[1], Dr. Navin Rajpal[2] and Jasvinder Singh[3]
[1]M.Tech student, University School of Information and Communication Technology, Guru Gobind Singh Indraprastha University, Sector 16-C Dwarka, New Delhi-110078, India
[2]Professor, University School of Information and Communication Technology, Guru Gobind Singh Indraprastha University, Sector 16-C Dwarka, New Delhi-110078, India
[3]Ph.D Research Scholar, University School of Information and Communication Technology, Guru Gobind Singh Indraprastha University, Sector 16-C Dwarka, New Delhi-110078, India

Abstract
Content-based Image Retrieval is all about generating signatures of images in database and comparing the signature of the query image with these stored signatures. Color histogram can be used as signature of an image and used to compare two images based on certain distance metric. Distance metrics Manhattan distance (L1 norm) and Euclidean distance (L2 norm) are used to determine similarities between a pair of images. In this paper, Corel database is used to evaluate the performance of Manhattan and Euclidean distance metrics. The experimental results showed that Manhattan showed better precision rate than Euclidean distance metric. The evaluation is made using Content based image retrieval application developed using color moments of the Hue, Saturation and Value(HSV) of the image and Gabor descriptors are adopted as texture features.
Keywords:-Content-Based Image Retrieval, CBIR systems, Image Database, Color histograms, color scale images, Euclidean Distance method, Manhattan distance, Gabor wavelet.

I. INTRODUCTION

Content-based image retrieval (**CBIR**), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision to the image retrieval problem, that is, the problem of searching for digital images in large databases. Content based image retrieval (CBIR) is a lively discipline, expanding in both depth and breadth and an important research topic covering a large number of research domains like image processing, computer vision, very large databases and human computer interaction [6, 10].

"Content-based" means that the search will analyse the actual contents of the image [1]. The term 'content' in this context might refer colors, shapes, textures, or any other information that can be derived from the image itself. Due to inability to examine image content, searches must rely on metadata such as captions or keywords. Such metadata must be generated by a human and stored alongside each image in the database. The design and development of effective and efficient CBIR systems are still a research problem, because the nature of digital images involves two well-known problems: the semantic gap and the computational load to manage large file collections. The semantic gap is the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation [7]. It has linguistic and contextual consequences, and mainly depends on the domain knowledge to represent images. On the other hand, the computation load, when large image collections are managed, may make impractical use of CBIR systems [4].

One of the main tasks for (CBIR) systems is similarity comparison, extracting feature signatures of every image based on its pixel values and defining rules for comparing images [8]. These features become the image representation for measuring similarity with other images in the database. Distance metric or matching criteria is the main tool for retrieving similar images from large image databases for all the above categories of search. Several distance metrics, such as the L1 metric (Manhattan Distance), the L2 metric (Euclidean Distance) and the Vector Cosine Angle Distance (VCAD) have been proposed in the literature for measuring similarity between feature vectors [18]. In content-based image retrieval systems, Manhattan distance and Euclidean distance are typically used to determine similarities between a pair of images. In image processing applications, components of a feature vector (e.g., color histogram) are usually normalized by the size of the image and as a result, the Manhattan, Euclidean, the cosine angle based distance and Histogram Intersection distance metrics produce different ordering of retrieved images. In this paper, Corel database is used to evaluate the performance of

Manhattan and Euclidean distance metrics. The experimental results showed that Manhattan showed better precision rate than Euclidean distance metric. Content based image retrieval application developed use color moments of the features [12].

Hue, Saturation and Value (HSV) of the image and Gabor descriptors are adopted as texture

In the next section of the paper, color and texture representation and different distance metrics used for comparison have been described. In section 3, we explain the details of the content based image retrieval application. Section 4 contains the experimental results and we draw conclusions from our work in the last section of the paper.

II. Color and Texture Representation

A. *Color Representation*

Color is one of the important features for recognizing the images by humans and it is most commonly used feature in image retrieval. Color is an important dimension of human visual observation that allows discrimination and recognition of visual information [3] .Color features are also easy to extract, and have been found to be effective during indexing and searching of images in image databases. One of the main aspects of color feature extraction is the choice of a color space.

Commonly used color spaces are:

- RGB (Red , Green, Blue)
- CIE (abbreviated for its French name, Commission internationale de l'éclairage)
- HSV (Hue, Saturation, Value)

In RGB color space it assigns to each pixel a three element vector giving the color intensities of the three colors, red, green and blue[15,16]. The space spanned by the R, G, and B values describe visible colors, which are represented as vectors in the 3D RGB color space. As a result, the RGB color space provides a useful starting point for representing color features of images. However, the RGB color space is not perceptually uniform. The RGB color space can be transformed to generate other color spaces. The idea color space transformation is to develop a model of color space that is perceptually similar with human color vision.

CIE color space was created by The International Commission on Illumination (usually abbreviated CIE for its French name, Commission internationale de l'éclairage). The CIE color space encompasses all color sensations that an average person can experience. In CIE color space it judges the relative luminance (brightness) of different colors in well-lit situations. Humans tend to perceive light within the green parts of the spectrum as brighter than red or blue light of equal power. A CIE Color space is generated by nonlinear transformation of the RGB space [12]. However, the CIE color spaces are inconvenient because of the calculation complexities of the transformation to and from the RGB color space.

The HSV color space (Hue, Saturation, Value) corresponds better to how people experience color than the RGB color space does. The HSV color space is widely used in the field of color vision. The chromatic components Hue, Saturation and Value correspond closely with the categories of human color perception. As hue varies from 0 to 1.0, the corresponding colors vary from red, through yellow, green, cyan, blue, and magenta, back to red, so that there are actually red values both at 0 and 1.0. As saturation varies from 0 to 1.0, the corresponding colors (hues) vary from unsaturated (shades of gray) to fully saturated (no white component). As value, or brightness, varies from 0 to 1.0, the corresponding colors become increasingly brighter.

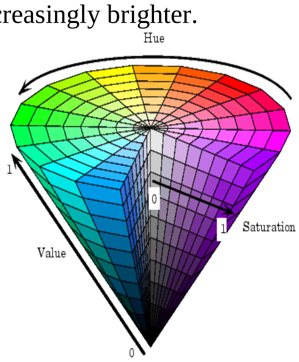

Figure 1: HSV Cone

HSV color space is also a nonlinear transformation of the RGB, but it is easily invertible. The HSV color space is perceptually uniform. In this project, HSV color space has been used to extract color features. The HSV values of a pixel can be transformed from its RGB representation according to the following formula

$$H = \cos^{\wedge}(-1)[(1/2\,[(R-G)+(R-B)\,])/(\sqrt{}((R-G)^{\wedge}2\,)+(R-B)(G-B)\,)]$$

$$S = 1 - \left[\frac{3[\min(R,G,B)]}{R+G+B}\right]$$

$$V = \left[\frac{R+G+B}{3}\right]$$

B. *Color Feature Extraction*

The objective of color indexing is to retrieve all the images whose color compositions are similar to the color composition of the image in the query. Histograms are useful because they are relatively insensitive to position and orientation changes. Besides they are sufficiently accurate [11]. However, they do not capture spatial relationship of color regions and they have limited discriminating power.

Many publications focus on color indexing techniques based on global color distributions. These global distributions have limited discriminating power because they are unable to capture local color information. Color correlogram and color coherence vector can combine the spatial correlation of color regions as well as the global distribution of local spatial correlation of colors. These techniques perform better than traditional color histograms when used for content-based image retrieval. But they require very expensive computation. Color moments have been successfully used in content based image retrieval systems. It has been shown [12] that characterizing one dimensional color distributions with the first three moments is more robust and runs faster than the histogram based methods.

C. TEXTURE REPRESENTATION

Texture is defined as structure of surface formed by repeating element or several elements in different relative spatial position. Texture is an inherent property of all surfaces that describes visual pattern, each having homogeneity [10].Gabor wavelet is widely adopted to extract texture from the images for retrieval and has been shown to be very efficient. Basically Gabor filters are a group of wavelet capturing energy at specific frequency and specific orientation. The scale and orientation tunable property of Gabor filter especially useful for texture analysis [2]. The Gabor filter design is done as follows:

For a given image I(x,y) with size P x Q, its Gabor wavelet transform is given by a convolution:

$$G_{mn} = \sum\sum I(x-s, y-t)\, \ddot{Y}^{n}{}_{mn}(s,t)$$

Where, t and s are the filter mask size variables, and *mn is a complex conjugate of mn. After applying Gabor filters on the image with different orientation at different scale, we obtain an array of magnitudes which is given as:

$$E(m,n) = \sum |G_{mn}(x,y)|$$

Where m =0,1,……, M-1; n=0,1,…… N-1.

These magnitudes represent the energy content at different scale and orientation of the image. The main purpose of texture-based retrieval is to find images or regions with similar texture [16]. It is assumed that one is interested in images or regions that have homogenous texture; therefore the following mean μ_{mn} and standard deviation mn of the magnitude of the transformed coefficients are used to represent the homogenous texture feature of the region:

$$\mu_{mn} = E(m,n)\frac{\sigma_{mn}}{P\ X\ Q}$$

$$= \sqrt{}\sum\sum\frac{(|G_{mn}(x,y)| - \mu_{mn})2}{P\ X\ Q}$$

A feature vector f_g (texture representation) is created using μ_{mn} and mn as the feature components. Four scales and 6 orientations are used in common implementation and the feature vector of length 48 is given by:

$$f_g = \{\mu_{00}, \sigma_{00,}\, \mu_{01}, \sigma_{01,} \ldots\ldots\ldots\ldots, \mu_{35}, \sigma_{35,}\}$$

III. DISTANCE METRICS

In the domain of image retrieval from large databases using signatures like color histogram, each 'n' dimensional feature vector may be considered as a point in the 'n' dimensional vector space. Thus, a feature vector is mapped to a point in the n-dimensions [18]. This mapping helps us to perceive the images (represented by their feature vectors) as high-dimensional points. The advantage of this representation is that one can now use different distance metrics for (i) finding similarity between two images and (ii) ordering a set of images based on their distances from a given image. This enables one to do a nearest neighbor search on a large database of images and retrieve a result set containing images that are closest matches to a user-specified query. It is evident that the images and their ordering depend both on the feature extraction method as well as on the distance metric used. In this work, Manhattan distance and Euclidean distance metrics have been used as a similarity rule.

After the color, shape or texture information is extracted from an image, it normally encoded into a feature vector. Given two feature vectors, x_1 and x_2, a distance function computes the difference between them. It is hoped that this difference will accurately measure the dissimilarity between the images from which the features were extracted. The greater the distance, the less the similarity. Distance functions Euclidean (L2) norm and Manhattan distance or L1 norm (also known as city block metric) equations are as follows:

The Euclidean distance is given by the following mathematical expression.

$$d_E(x_1, x_2) = \sqrt{\sum_{i=1}^{i=n}(x_1(i) - x_2(i))^2}$$

Where, x_1, x_2 are the coordinates where two pixels p_1 and p_2 are located.

The Manhattan distance is given by following mathematical

$$d_M(x_1, x_2) = \sum_{i=1}^{i=n}(x_1(i) - x_2(i))$$

Distance functions or metrics follows following properties:

$$d(p,q) \geq 0 \ \ (d(p,p) = 0)$$
$$d(p,q) = d(q,p)$$
$$d(p,z) \leq d(p,q) + d(q,z)$$

IV. CONTENT BASED IMAGE RETRIEVAL SYSTEM

A content based application for large-scale study of image retrieval algorithms has been developed as shown in Fig. 2. The system retrieves images that are similar to a user-specified query from an image database. The goal is to present the user with a subset of images that are more similar to the query image. First of all, signatures of all the images are stored in a database and then an image is queried with its signature as input. In this application, color histogram of an image is used as the signature. The rule for similarity measure is distance metrics[9]. Two different distance metrics, via, Euclidean distance and Manhattan distance, are used in two color spaces, RGB and HSV. Firstly apply the color image conversion, i.e., in the form of RGB. One must have to apply the color conversion technique first with the help of color model, i.e., HSV (Hue, Saturation, and Value). The same is done using MATLAB for converting the image color in HSV format. After converting, one selects the blocks of the picture, which is helpful for finding the distance measure. The system has been developed on Windows OS, Matlab 7.9 as front end and Ms Excel as backend. The WANG dataset, that is a subset of 1000 images of the Corel dataset is used (http://savvash.Blogspot.com/2008/12/benchmark-databases-for-cbir.html). The Corel dataset is a large dataset of photographs with annotation. The images in WANG dataset were selected manually to form a dataset of 10 classes of 100 images each. The database is being expanded by adding more images from time to time. The image features i.e. color histograms are stored in a file structure. The number of bins is 64 for all the histogram types. L1 and L2 distance measures have been used to select nearest neighbors and both have given almost identical results. The top similar images, ranked on the basis of the distance are displayed as thumbnails. The graphical user interface displays the query image and the results for browsing to the user. A snapshot of the user interface is shown in the Figure 2 and Figure 3.

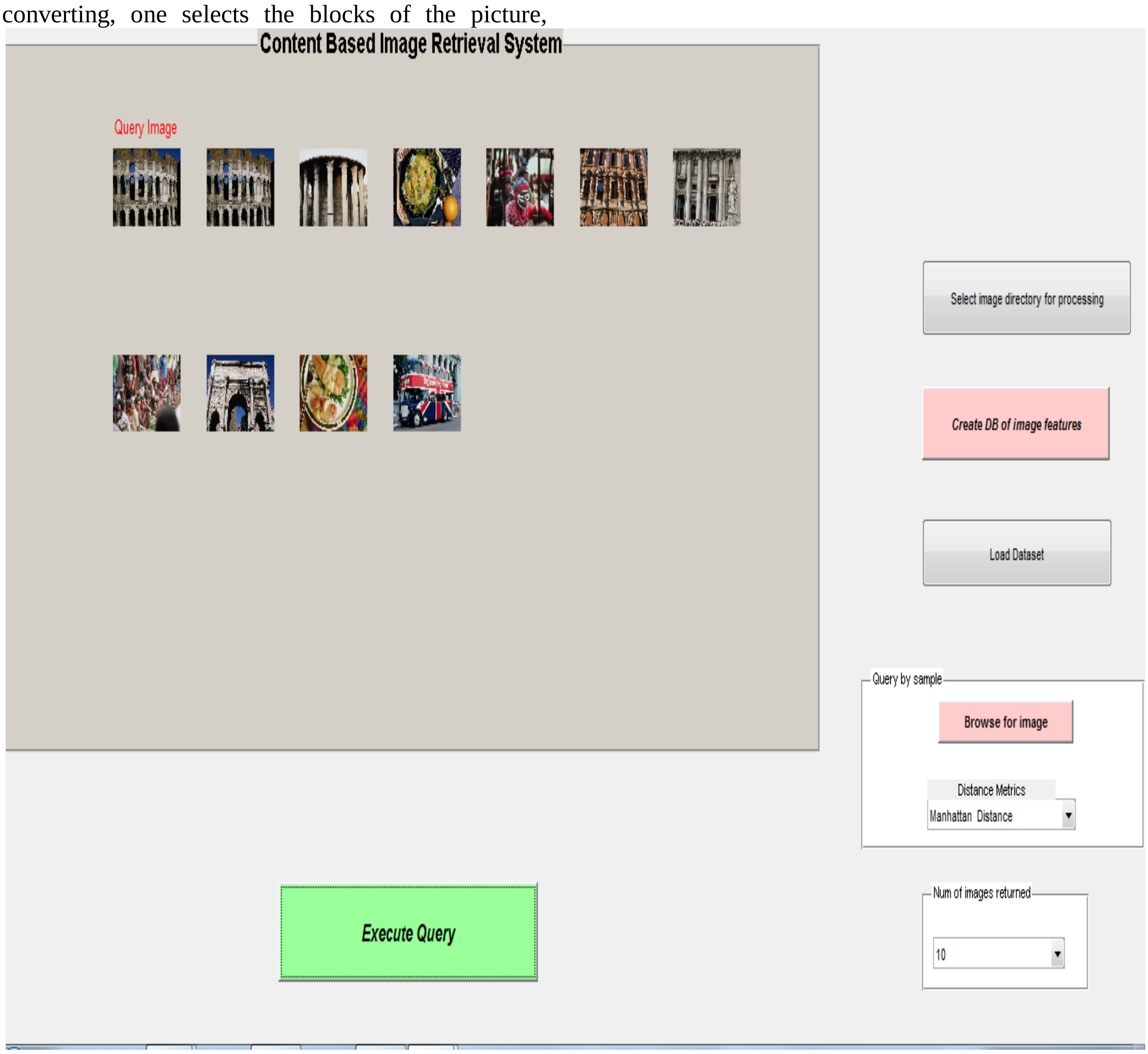

Figure 2: Resultset display of Manhattan Distance in the application

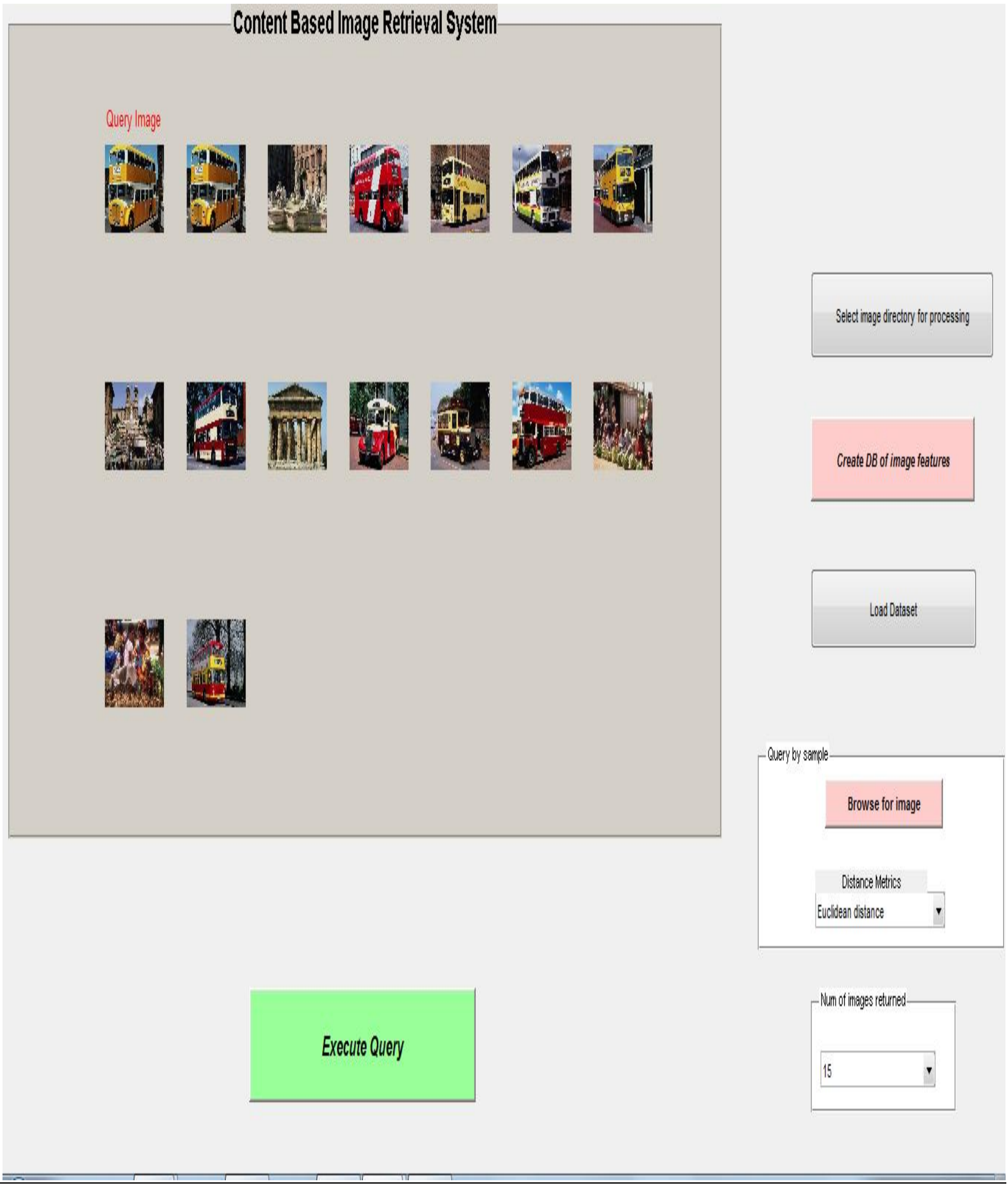

Figure 3: Resultset display of Euclidean Distance in the application

V. Experimental Results

In this section, experimental results have been reported. All experiments were performed on an Intel core i3 machine.

Processor 2.20 GHz with 2 GBytes of RAM. The system was implemented in Matlab 7.9. the COREL database for the experiment purpose. consists of 1000 images divided in 10 categories with 100 images in each category. The categories of the images are people, beach, monuments, buses, dinosaurs, elephants, flowers, horses, mountains and cuisines.

For the performance evaluation, 5 random images from each of the 10 category as the query image have been arbitrarily chosen. An ideal image retrieval system is expected to retrieve all the relevant images. One of the popular measures is the precision rate that is the ratio of the number of relevant images retrieved and total number of images in the collection [13].

$$Precision = \frac{Number\ of\ Relevant\ Images\ retrieved}{Total\ Number\ of\ images\ retrieved}$$

The precision of the distance metric is calculated by varying the number of retrieved images. Once, the distance of query image is calculated with all the images in the database, it is sorted. The order of sorting depends on the type of distance metric. So, the denominator for calculating precision (Total No. of images retrieved) is varied by considering 5, 10, 15 and 20 nearest neighbors (NN). Out of these nearest neighbors (NN), how many of them belong to the same category as the query image, that's precision.

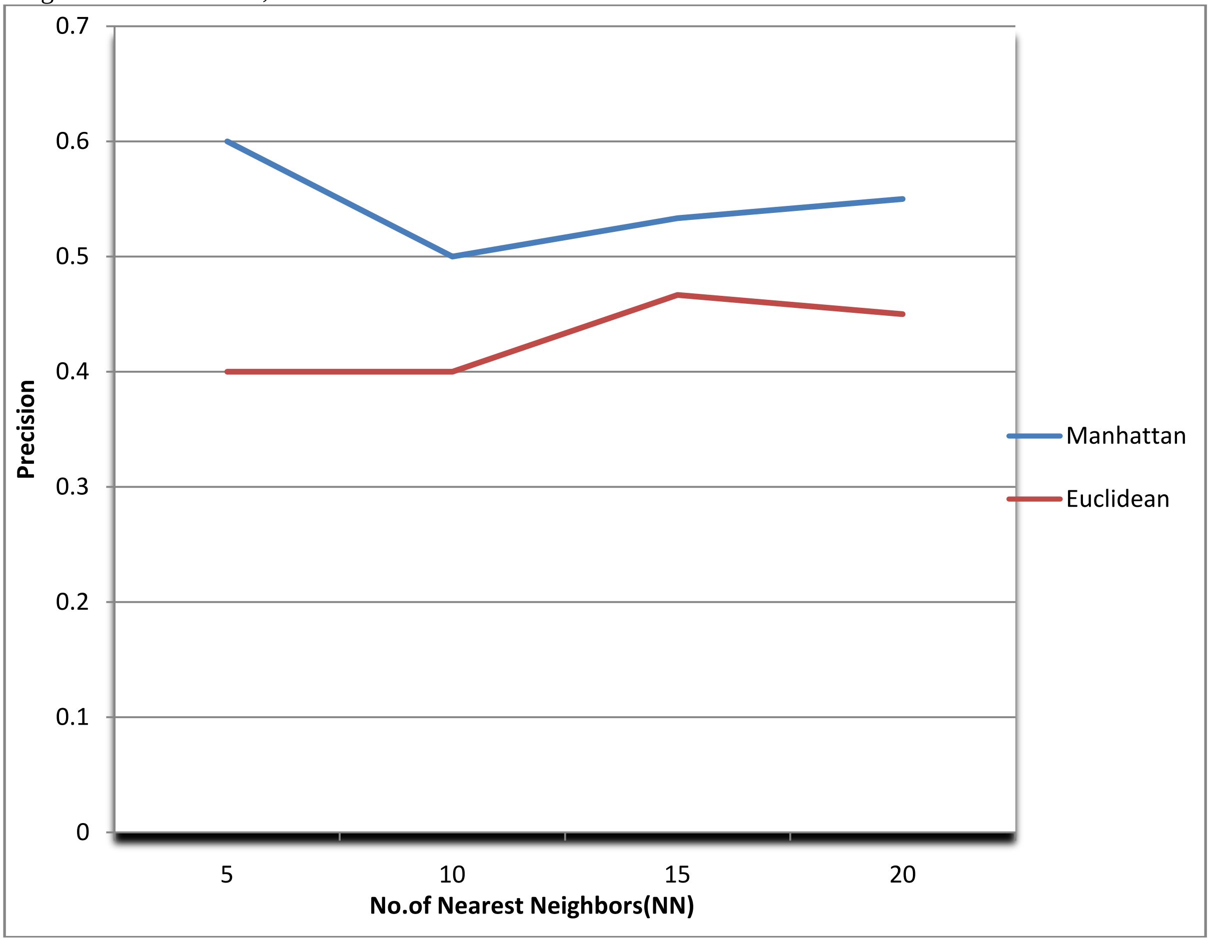

Figure 3: Performance comparison of Euclidean and Manhattan distance metrics

The performance of two similarity measures, L1 and L2 distance have been compared, in the experiments. In the figure 3 ,it has been shown precision for the first 5, 10, 15 and 20 nearest neighbor images of the result set. From the above shown figure, one can observe that Manhattan shows better result than Euclidean. However, with higher number of nearest neighbors spurious images keep coming in the result set.

VI. Conclusion

Content based image retrieval (CBIR) is a lively discipline, expanding in both depth and breadth and an important research topic covering a large number of research domains like image processing, computer vision, very large databases and human computer interaction. One of the main tasks for (CBIR) systems is similarity comparison, extracting feature signatures of every image based on its pixel values and defining rules for comparing images. Distance metric or matching criteria is the main tool for retrieving similar images from large image databases for all the above categories of search [17]. The Manhattan distance and Euclidean distance have been used to determine similarities between a pair of images in the content based image retrieval application. In this paper, the comparison of the performance of Manhattan (L1) and Euclidean(L2) distance metric has been done.. According to the precision graph, one can conclude Manhattan distance metric shows better precision than Euclidean metric.

Developments and studies are going on for further improvements in design and performance of "CONTENT BASED IMAGE RETRIEVAL SYSTEMS". As continuation of this work we propose to perform analysis of other distance metrics like Earth Mover’s distance, weighted Euclidean distance , Vector Cosine Angle distance and Histogram Intersection. In addition to this we would like to performance comparison of these metrics using confusion metrics.

REFERENCES

[1] Ja-Hwung Su, Wei-Jyun Huang, Philip S. Yu, and Vincent S. Tseng, ”*Efficient Relevance Feedback for Content-Based Image Retrieval by Mining User Navigation Patterns*” IEEE transactions on knowledge and data engineering, vol. 23, no. 3, march 2011.

[2] Wasim Khan, Shiv Kumar. Neetesh Gupta, Nilofar Khan, “*A Proposed Method for Image Retrieval using Histogram values and Texture Descriptor Analysis*” , International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-I Issue-II, May 2011.

[3] P Liu, K. Jia, Z. Wang and Z. Lv, “*A New and Effective Image Retrieval Method Based on Combined Features*”, *Proc. IEEE Int. Conf. on Image and Graphics,* vol. I, pp. 786-790, August 2007.

[4] Kambiz Jarrah, Ling Guan “*Content-Based Image Retrieval via Distributed Databases*” CIVR’08, July 7-9, 2008, Niagara Falls, Ontario, Canada. 2008 ACM.

[5] Ritendra Datta, Dhiraj Joshi, Jia Li and James Wang, “*Image Retrieval: Ideas, Influences, and Trends of the New Age*”, Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval, November 10-11, 2005, Hilton, Singapore.

[6] Neetesh Gupta, R.K.Singh, “*A New Approach for CBIR Feedback based Image Classifier*”, International Journal of Computer Applications.(0975 – 8887) Volume 14– No.4, January 2011.

[7] Quack, U. MÄonich, L. Thiele, and B. S. Manjunath. Cortina: *a system for large-scale, content-based web image retrieval.* In *MULTIMEDIA '04: Proceedings of the 12th annual ACM international conference on Multimedia*, pages 508{511, New York, NY, USA, 2004. ACM Press.

[8] Jun Luo and Hang Kuang *Content-based image retrieval using combination features.* Computer Engineering and Applications 2009 45(1):153-155.

[9] A.M.W. Smeulders, M. Worring, S. Santini, A.Gupta, and R. Jain, “*Content-based image retrieval at the end of early years*, “IEEE Trans. On Pattern Analysis and machine intelligence, vol. 22, no. 12, December 2000.

[10] Sanjay Patil, Sanjay Talbar (2012), ‘*Content Based Image Retrieval Using Various Distance Metrics*’, Data Engineering and Management, Lecture Notes in Computer Science, Vol. 6411, pp 154-161.

[11] Chaobing Huang, Yarong Han, Yu Zhang (2012),‘*A Method for Object-based Color Image Retrieval’*, Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on , pp:1659-1663

[12] Rasli, R.M.Muda, T.Z.T., Yusof, Y.,Bakar, J.A. (2012)‘*Comparative Analysis of Content Based Image Retrieval Technique using Color Histogram. A Case Study of GLCM and K-Means Clustering’*, Intelligent Systems, Modelling and Simulation (ISMS), Third International Conference on, pp: 283 – 286.

[13] Xiang-Yang Wang & Jun-Feng Wu & Hong-Ying Yang (2010), ‘*Robust image retrieval based on colour histogram of local feature regions*’, Multimedia Tools Apply vol. 49, pp. 323–345.

[14] Yong-Hwan Lee, Sang-Burm Rhee, Bonam Kim, ‘*Content-based Image Retrieval Using Wavelet Spatial-Color and Gabor Normalized Texture in Multi-resolution Database’*, Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference on , pp: 371 – 377, 2012.

[15] Thomas Deselaers, Daniel Keysers, and Hermann Ney (2007), Features for Image Retrieval: An Experimental Comparison. Information Retrieval, 11(2): 77-107, 2008.

[16] P. S. Hiremath, Jagadeesh Pujari, “Content Based Image Retrieval using Color, Texture and Shape Features,” 15th International Conference on Advanced Computing and Communications, IEEE Computer Society, 2007, pp. 780-784.

[17] Vadivel, A.K. Majumdar, and S. Sural, “Performance comparison of distance metrics in content-based image retrieval applications”, Proc. of Int’l. Conf. on Information Technology, Bhubaneswar, India, pp. 159–164, 2003.

[18] Qian, G., Sural, S., Gu, Y., and Pramnik, S., “Similarity between Euclidean and Cosine Angle Distance for Nearest Neighbor Queries,” ACM Symposium on Applied Computing, pp. 1232-1237, 2004.