An Analytical Study on Query Integration in Image Retrieval System

J.Esther*
Department of Computer Science
Mano College Govindaperi-627 414, India
esther_obed@yahoo.co.in

Dr.M.Mohamed Sathik
Principal Sadakathulla Appa College
Tirunelveli,India
mmdsadiq@gmail.com

Abstract: Content Based Image Retrieval(CBIR) consists of retrieving the most visually similar images to a given query image from a database of images. CBIR in medical image databases are used to assist physician in diagnosis the diseases and also CBIR techniques are used to aid diagnosis by identifying similar past cases. The medical and related health professions use and store visual information in the form of X-rays, CT scanned images, MRI scanned images, for diagnosis and monitoring purposes. In this paper we analysis the several techniques viz., Color Histogram, Gabor, and wavelet were applied in a novel way for Brain image and also in HSV Color space.

Keywords: Color Histogram, Gabor, Wavelet, HSV, Brian image.

I. INTRODUCTION

Retrieval of required-query-similar images from abundantly available / accessible digital images is a challenging need of today. The image retrieval techniques based on visual image content has been in-focus for more than a decade. Medical image databases have grown immensely in the past few years. Imaging studies, such as magnetic resonance (MR) and computed tomography (CT) result in a large volume of data. In the medical domain, the goal of a query system is to aid doctors to diagnose a patient by retrieving images with Known pathologies that are similar to the patient's image(s). The medical query system belongs to the content-based image retrieval (CBIR), and is aimed at efficient retrieval of relevant images from large image database based on automatically derived imagery features.

In this paper, we introduce our system briefly including the brain image data collecting, database structure and how to integrate the query. We believe this ability is critical to image retrieval and to progress for object recognition in general purpose and provide good enough performance to yield clinical use for brain CT/MRI images. The motivation is to get the best technique to be used in further image retrieval application. The techniques are color histogram, Gabor Transform, and wavelet Transform. The basic block diagram of CBIR is shown in the Figure 1.
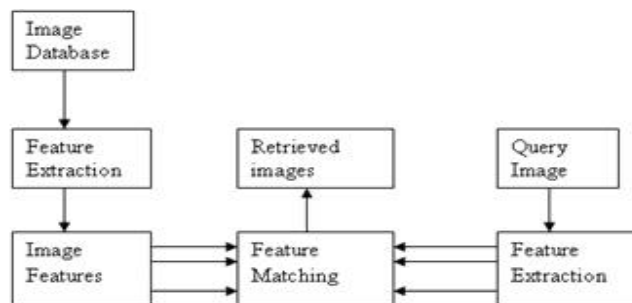

Figure 1.   Block diagram of CBIR

This paper is organized as follows. The next section briefly describes the System Description, Feature Extraction methods are discussed in Section III, Result and Discussions are in Section IV, finally conclusion of the paper and presenting future research directions are in Section V.

II. SYSTEM DESCRIPTION

Most of the current image indexing practices mainly rely on color, texture or shape features. The performance of the image retrieval technique improves if these features are combined and considered together. In case of color and texture features combination, the HSV planes are considered separately and then some texture features are extracted from these color planes. The Data flow diagram of our medical image retrieval is shown in the Figure 2. The medical image search system combines advanced image processing algorithms with direct access to a large medical image database to reliably return similar images to those submitted as queries by the user. The database matching system relies on a very large, well-organized database of images. The CT/MRI brain images are used here for the retrieval of medical image dataset. We implement it in the HSV color space.
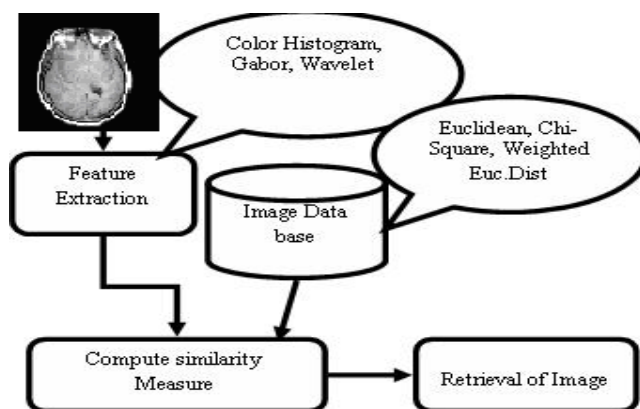

Figure 2.   Data Flow diagram of Medical image retrieval

III. FEATURE EXTRACTION METHODS

The steps of the system are described in the Figure 1 contains two phases.
A. Feature Extraction Technique
B.. Retrieval of medical image.

A. *Feature Extraction Technique:*

CBIR systems use visual content such as color, texture, and simple shape properties to search images from large image databases[1]. In our system we used the feature extraction techniques of Color Histogram (Color Feature), Gabor Transform (Texture feature) and Wavelet transform (Texture Feature)[2].Extract the features of CT brain image and MRI scan brain image in the data set and also for the query image. For retrieve the similarity images we used Distance metric measures.

a. *Color Histogram:*

In this paper, three dimensional color space HSV is investigated and the color histogram based image retrieval method is used. The color histogram for an image is constructed by counting the number of pixels of each color. This approach is more frequently adopted for CBIR systems is based on the conventional color histogram (CCH), which contains occurrences of each color obtained counting all image pixels having that color. Each pixel is associated to a specific histogram bin only on the basis of its own color, and color similarity across different bins or color dissimilarity in the same bin is not taken into account [3]. Since any pixel in the image can be described by three components in a HSV color space (hue, saturation and v value in HSV space), a histogram, i.e., the distribution of the number of pixels for each quantized bin, can be defined for each component. By default the maximum number of bins one can obtain using the histogram function in MatLab is 256. The conventional color histogram (CCH) of an image indicates the frequency of occurrence of every color in an image.

b. *Gabor Transform:*

Gabor transform is a technique that extracts texture information from an image. It's a multi-scale, multi-resolution filter. The two-dimensional Gabor filter can be represented as a complex sinusoidal signal modulated by Gaussian function as $\psi(x, y; \sigma, \lambda, \theta k)$ [4],[5],[6] and it can be formulated as follows:

$$\psi(x, y; \sigma, \lambda, \theta_k) = g(x, y; \sigma) \exp\left(\frac{2\pi x \theta k}{\lambda} i\right) \qquad (1)$$

$$g(x, y, \sigma) = \exp\left(-\frac{x^2 \theta k + \gamma^2 y^2 \theta k}{2\sigma^2}\right) \qquad (2)$$

Is a Gaussian function where

$$x\theta_k = x\cos(\theta_k) + y\sin(\theta_k)$$
$$y\theta_k = -x\sin(\theta_k) + y\sin(\theta_k)$$

Here σ is the standard deviation of the Gaussian function with x, y dimensions and λ and θ_k are wavelength and orientation. The parameter γ is usually equal to 0.5. We design a Gabor wavelet for 4 scales and 5 orientations. We have conducted retrieval test both on texture images and natural images. Here we define the angle θ_k by 15, 45, 75,135 and 180 respectively and the wavelengths are 60, 80,120 and 130. For the first wavelength 60 we calculate θ_k, where k=1 to 5 with 15, 45, 75, 135 and 180. Then for the 2nd wavelength 80 we calculate θ_k, with 15, 45, 75,135 and 180 and so on. After applying the gabor filter, we extract the texture features of image by using standard deviation function. Then retrieving the image from the dataset we calculate the distance metric measures for every image. The minimum distance value signifies an exact match with the query.

c. *Wavelet Transform:*

Wavelet is the multiresolution analysis of the image. The simplest orthogonal filter bank is Haar filter bank[7].The content based image retrieval (CBIR) is done using the image feature set extracted from Haar Wavelets applied on the image at various levels of decomposition. It applies two channel filter banks namely, high pass filter and low pass filter. The high pass filter identifies the texture of the image. The low pass filter identifies the structure of the image.(i.e) edges, curves etc.

B. *Retrieval Of Medical Image:*

The most common method for comparing two images in content based image retrieval (typically an example image and an image from the database) is using an image distance measure. An image distance measure compares the similarity of two images in various dimensions such as color, texture, shape, and others. For example a distance of 0 signifies an exact match with the query, with respect to the dimensions that were considered. For retrieve the similarity medical image form the large medical image dataset we used the distance metric measure. The distance metric is a function which defines the distance between two images. Here we used the distance metric measures like Euclidean Distance,[8], Chi-Square Distance, and Weighted Euclidean Distance.

a) *Euclidean Distance:*

The formula of Euclidean distance is

$$D = \sqrt{\sum (x_i - y_i)^2} \qquad (3)$$

The minimum distance value signifies an exact match with the query.

b) *Chi-Square Distance:*

The formula is

$$D^2 = \left(\frac{(x_i - y_i)^2}{y_i}\right) \qquad (4)$$

x_i is the training image values

y_i is the query image values.

The minimum distance value signifies an exact match with the query.

c) *Weighted Euclidean Distance:*

The formula is

$$D = \sqrt{\sum W(x_i - y_i)^2} \qquad (5)$$

The minimum distance value signifies an exact match with the query. Here for the value of weight, which is denoted by W is assigned by 2.

IV. EXPERIMENTS & RESULTS

In this experiment we used 800 images, which consists of CT Brain images and MRI Scan brain images[9]. The experiments were conducted using Matlab 7.3 on an Intel Pentium-D 2.0GHz processor with 2GB memory. To measure the retrieval effectiveness, we have used the precision and recall as statistical comparison parameters for our proposed technique of CBIR[10]. The standard definitions of these measures are given in equation (6) & (7).

$$\Pr ecision = \frac{No.of\_relevant\_images\_retrieved}{Total\_number\_of\_images\_retrieved} \quad (6)$$

$$\mathrm{Re} call = \frac{No.of\_relevant\_images\_retrieved}{Total\_number\_of\_relevant\_images\_in\_dataset} \quad (7)$$

By applying these Measures in the Feature Extraction Techniques like Color Histogram, Gabor and Wavelet Transform with three different Distance Metrics Measures like Euclidean Distance, Chi-Square Distance, and Weighted Euclidean Distance are given in equation (3),(4),(5). to all 800 images in the dataset.

To ease the work of testing and analyzing the images a graphical user interface (GUI) was developed using Matlab. It consists of two main panels, one is for selection of feature extraction technique and another one is for the Result panel.
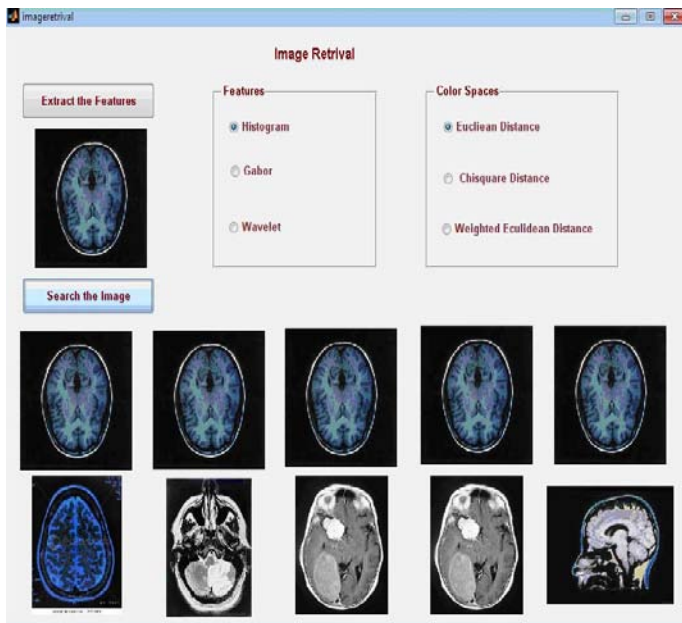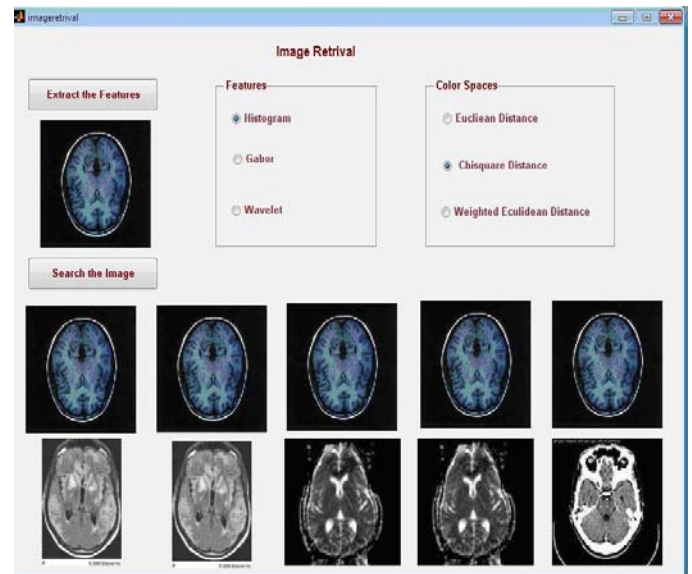

Figure 4.   Retrieval example of Color Histogram with Chi-Square Distance

Figure.3 shows an example of retrieval results obtained by Color Histogram technique with Euclidean distance metric measures. Figure 4 shows an example of retrieval results obtained by color histogram technique with Chi-square distance metric measures. Figure 5 color histogram technique with Weighted Euclidean distance metric measures.

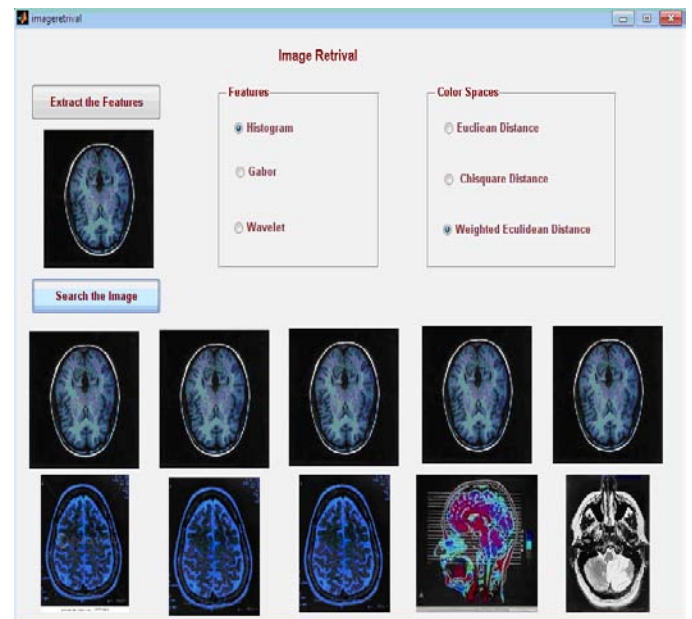Figure 3.   Retrieval example of Color Histogram with Euclidean Distance

Figure 5.   Retrieval example of Color Histogram with Weighted Euclidean Distance

Figure 6 shows that the retrieval example of Gabor transform using Euclidean distance. Figure 7 shows that the retrieval example of Gabor transform using Chi-Square Distance. Figure 8 shows that the retrieval example of Gabor transform using Weighted Euclidean Distance metric measures.
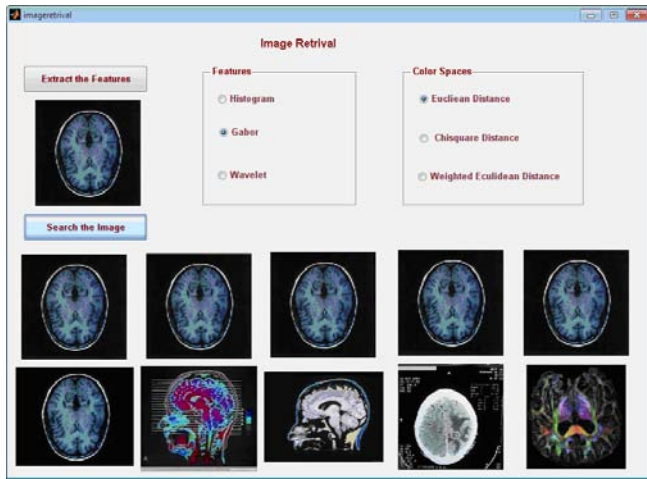
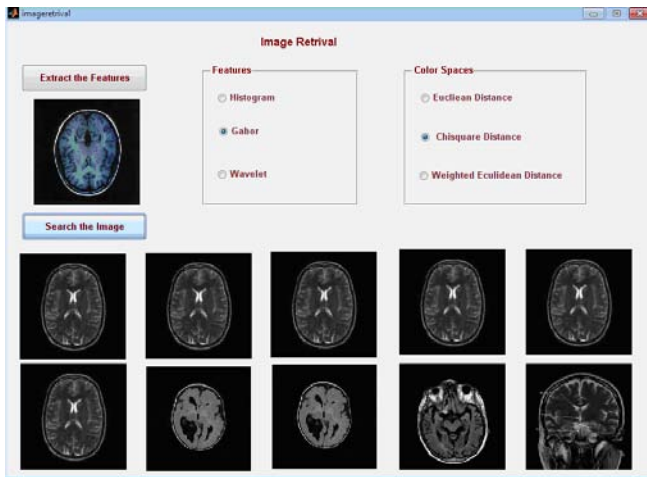Figure 6.   .Retrieval example of Gabor Transform with Euclidean Distance

Figure 7.   Retrieval example of Gabor Transform with Chi-Square Distance
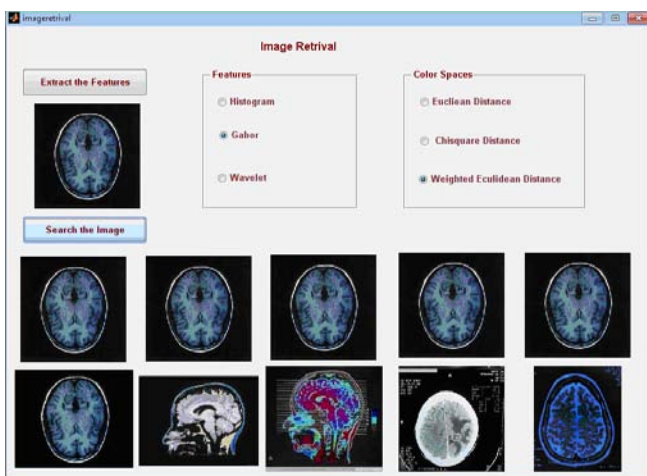

Figure 8.   Retrieval example of Gabor Transform with Weighted Euclidean Distance

Figure 9 shows that the retrieval example of Wavelet transform using Euclidean distance. Figure 10 shows that the retrieval example of Wavelet transform using Chi-Square Distance. Figure 11 shows that the retrieval example of Wavelet transform using Weighted Euclidean Distance metric measures.
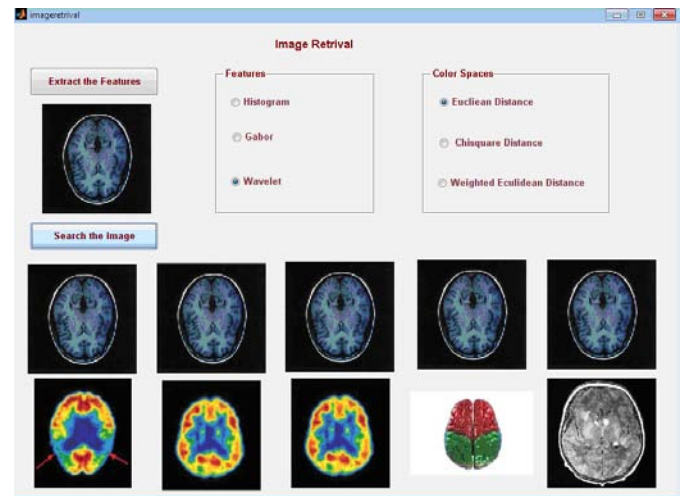

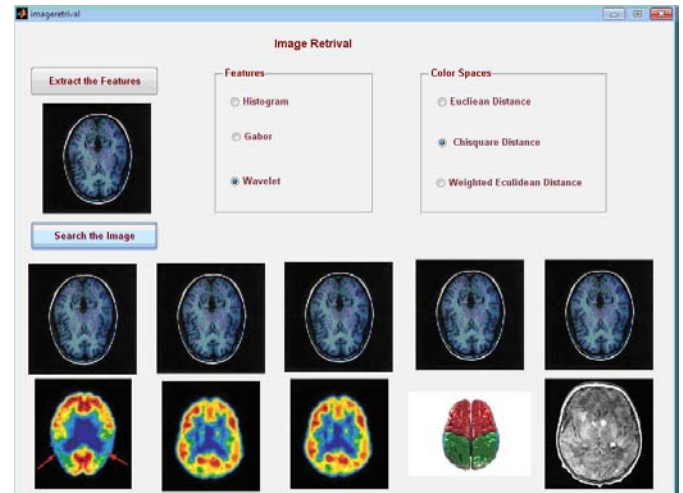Figure 9.   Retrieval example of Wavelet Transform with Euclidean Distance

Figure 10. Retrieval example of Wavelet Transform with Chi-Square
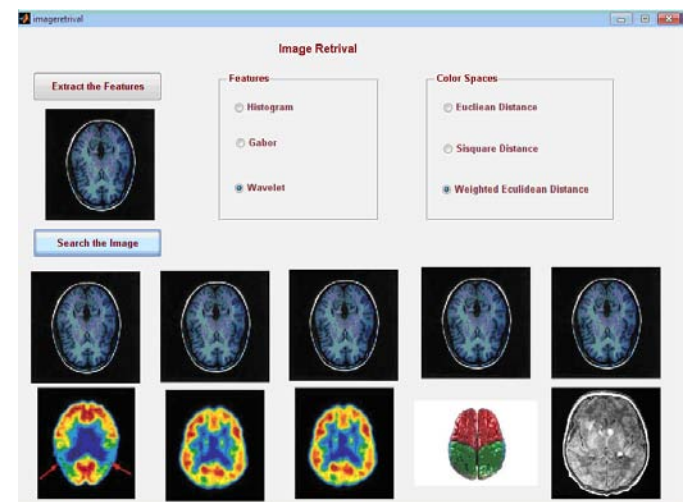

Figure 11. Retrieval example of Wavelet Transform with Weighted Euclidean Distance

From this analysis, the Gabor Transform with weighted Euclidean Distance Metric Measure gives the better performance. The average execution time taken by three techniques with three different distances metric measures are given in the comparison chart shown in Figure 12.
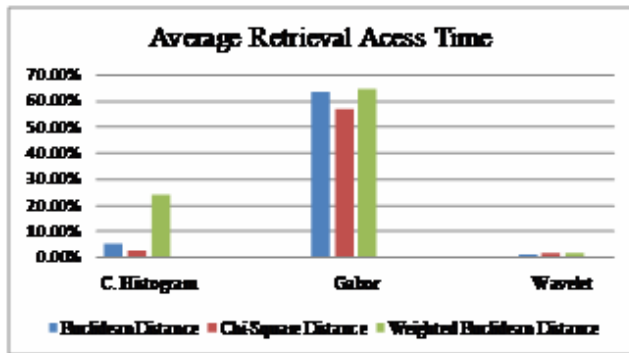

Figure 12. Retrieval access time with all distance measures

The comparative chart table of Precision and Recall in percentage is in Table 1.

Table I Comparative Chart Of Precision & Recall

Feature Extraction Technique	Distance Metric Measure	Precision	Recall
Color Histogram	Euclidean Distance	81.46%	75.11%
	Chi – Square Distance	66.38%	70.71%
	Weighted Euclidean Distance	76.90%	78.72%
Gabor Transform	Euclidean Distance	86.43%	87.59%
	Chi – Square Distance	63.28%	67.04%
	Weighted Euclidean Distance	88.77%	89.67%
Wavelet Transform	Euclidean Distance	62.98%	68.823%
	Chi – Square Distance	65.97%	67.97%
	Weighted Euclidean Distance	65.16%	68.22%

From the equation 6, the average recognition rate using precision is shown in Figure 13.
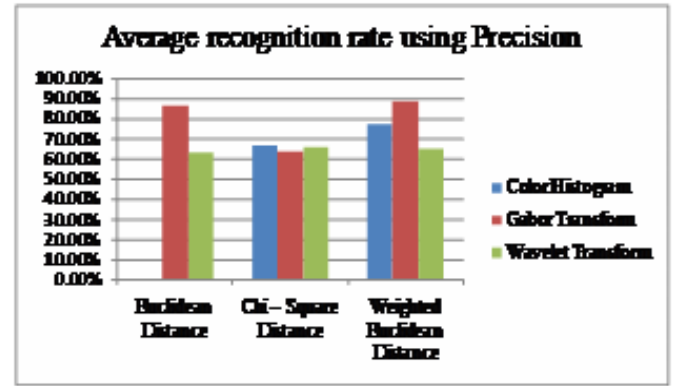

Figure 13. Average recognition rate using precision

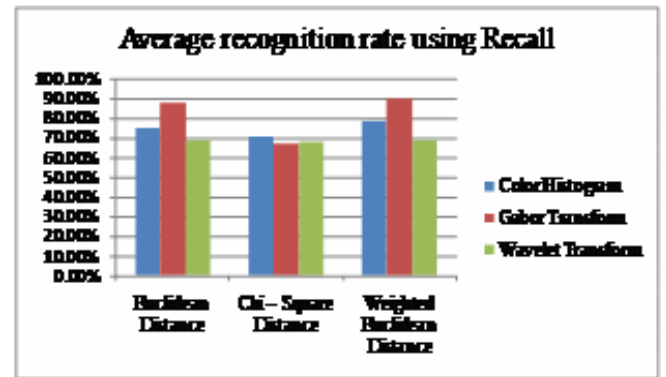From the equation 7, the average recognition rate using recall is shown in Figure 14.

Figure 14. Average recognition rate using Recall

Wavelet transform retrieve the image at very fast compared it with other techniques; but gives the poor performance, similarly color histogram also. But Gabor transform takes much time but gives the best result.

V. CONCLUSIONS

In this paper, an analytical study has been carried out on three feature extraction techniques and extracts the feature of texture, Color of CT/MRI brain image with three different distance metrics measures are taken for retrieving the similar images from the data set. From this experiment more images are retrieved by using Gabor than the other two feature extraction techniques like Color Histogram, Wavelet Transform. The future work is, doing the feature extraction of the image with another technique and with different distance measure to give the better result. And also extract the shape features from the image.

VI. ACKNOWLEDGEMENT

VII. REFERENCES

[1]   V. N. Gudivada and V. V. Raghavan, *Content based image retrieval systems*, IEEE Computer, vol. 28, Sept. 1995.

[2]   Young Deok chun,Nam Chul kim, "Content-Based Image Rerieval Using Multiresolution color and Texture features", Vol. 10,No 6, IEEE Transaction on multimedia,October 2008.

[3]   Wei-Ta Chen, Wei-Chuan Liu, and Ming-Syan Chen*, Fellow,* "Adaptive Color Feature Extraction Based onImage Color Distributions" IEEE Transactions On Image Processing, Vol. 19, No. 8, August 2010.

[4]   Tomasz Andrysiak, Michał Chora´, "Image Retrieval Based On Hierarchical Gabor Filters", Int. J. Appl. Math. Comput. sci., , Vol. 15, No. 4, 471–480,2005.(Article in a journal)

[5]   B. S. Manjunath and W. Y. Ma, *Texture features for browsing and retrieval of image data*, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 18, pp. 837-842, Aug. 1996.

[6]   Tomasz Andrysiak, Michał Chora´S," Image Retrieval Based On Hierarchical Gabor Filters", Int. J. Appl. Math. Comput. Sci.,Vol.15,No.4, 471–480,2005.(Article in a journal)

[7]   Rafael C.Gonzalez, Richard E Woods, Steven L.Eddins,"Wavelet Transforn in Digital Image Processing using MATLAB," pp.256-294, Pearson Education.

[8]   Wan Siti Halimatul Munirah Wan Ahmad and Mohammad Faizal Ahmad Fauzi," Comparison of Different Feature Extraction Techniques in Content-Based Image Retrieval for CT Brain Images", 978-1-4244-2295-1/08,IEEE 2008.(Article in a Conference Proceedings)

[9]   "The Whole Brain Atlas" www.harvard.edu

[10]   N.Krishnan, M..Sheerin Banu,C.Callins Christiyana, "Content Based Image Retrieval using Dominant Color Identification Based on Foreground Objects", 0-7695-3050-8/07 IEEE 2007 (Article in a Conference Proceedings).