ISSN 2250-3153

Image Database Classification Using Neural Network with Gabor Filter and CBIR Technique

Swathi Rao G[*] , Manipriya Singh[**], Ramnath panwar Singh[***]

[*]Digital Communication, Dehradun Institute of technology, Uttarakhand Technical University, Dehradun-248006, India
[**]Information Technology, SRM University
[***]Electrical and Electronics Engineering, Roorkee college of engineering

Abstract- Content based image retrieval system is the technique which uses visual contents to search images from large scale image databases according to the user's interest. The term content refers to color, shape, texture that can be derived from the image. In this paper an image retrieval system using artificial neural network (ANN) in MATLAB with the help of Gabor filter features is contemplated. In the proposed system, mean and standard deviation of the images are calculated later to the filtering process of the images using Gabor filter. Using the neural network classifier the system is trained and tested and classifies the images from a vast database relevant to the requirement. A database having 1000 images spread across ten categories is taken for the implementation purpose. Net average precision and recall values are computed for the database query. The obtained results show the performance improvement with higher precision and recall values.

Index Terms- ANN, CBIR, Classification, Feature extraction, Gabor feature vector, neural network, Similarity measurements.

I. Introduction

In today's scenario image classification and retrieval has become the most challenging and important research work for a wide range of applications like architectural and engineering design, art collections, crime prevention, geographical information and remote sensing systems, intellectual property, medical diagnosis, military applications etc. Locating a desired image in a large and varied collection of database is a considerable impediment for the researchers. Early work on image retrieval can be dated back to the 1970's but they were not based on visual features but textual annotation. Hence it couldn't support task dependant queries. Traditional methods posed such problems which have led to the image retrieving techniques based on content or features such as texture, color and shape. This is called the content based image retrieval (CBIR) technique. Since early 1990's, many systems have been proposed and developed, like QBIC, Virage, Pichunter, Visual SEEK, Chabot, Excalibur, Photo book, Jacob, UC Berkeley Digital Library Project. Most of the above mentioned systems and much of the past research have procured the CBIR from its infancy to the matured stage. Even though in some cases these systems exhibit substantial outcomes but still have limited efficiency. They have concentrated on the extraction of the low-level features. They emphasized on the explicit features of the images. These features are automatic but omit to study the implicit meaning behind the image. The hidden meaning or the semantic idea of the image and video can be interpreted solely by analyzing its contents. This leads to the semantic gap. Until and unless, we study the various implicit meanings of images and videos, which cannot be discernible by the content the semantic gap will not reduce. The performance of these image retrieval systems can be improved if features extracted are used to train artificial neural networks and classify the image based on trained variables. Hence an automatic technique is proposed to classify and categorise images based on supervised learning where collection of trained images are given. But, the problem is to identify a new upcoming unlabeled image. Each instance of the training will contain class specific labeling of the images and their descriptive feature vectors. Hence this paper combines the approach of both the CBIR technique along with the neural network training. The ANN network is trained and the set of image database are labeled using the feedback weight vectors from the users which would distinguish the image as appropriate or inappropriate. The image labeling is done once and the neural network is trained for the concerned database and therefore it can group and classify the images according to the feature vectors.
In section 2 theoretical contemplation is elaborated, section 3 explains the feature vector extraction, section 4 brings out the proposed CBIR technique and results and conclusion is revealed in section 5 & 6 respectively.

II. THEORETICAL CONCEPTS

A.Content Based Image Retrieval

The term content-based image retrieval (CBIR) seems to have originated in 1992, when it was used by T. Kato to describe experiments into automatic retrieval of images from a database, based on the colors and shapes present. Since then, the term has been used to describe the process of retrieving desired images from a large collection on the basis of syntactical image features. The techniques, tools and algorithms that are used originate from fields such as statistics, pattern recognition, signal processing, and computer vision. There is a growing interest in CBIR because of the limitations inherent in metadata-based systems, as well as the large range of possible uses for efficient image retrieval. Textual information about images can be easily searched using existing technology, but requires humans to personally describe every image in the database. This is impractical for very large databases, or for images that are generated automatically, e.g. from surveillance cameras. It is also possible to miss images that use different synonyms in their descriptions. Systems based on categorizing images in semantic classes like "cat" as a subclass of "animal" avoid this problem but still face the same scaling issues. Different implementations of CBIR make use of different types of user queries.

Semantic retrieval

The ideal CBIR system from a user perspective would involve what is referred to as semantic retrieval, where the user makes a request like "find pictures of Abraham Lincoln". This type of open-ended task is very difficult for computers to perform - pictures of Chihuahuas and Great Danes look very different, and Lincoln may not always be facing the camera or in the same pose. Current CBIR systems therefore generally make use of lower-level features like texture, color, and shape, although some systems take advantage of very common higher-level features like faces. Not every CBIR system is generic. Some systems are designed for a specific domain, e.g. shape matching can be used for finding parts inside a CAD-CAM database.

Query by example is a query technique that involves providing the CBIR system with an example image that it will then base its search upon. The underlying search algorithms may vary depending on the application, but result images should all share common elements with the provided example. Options for providing example images to the system include:

A preexisting image may be supplied by the user or chosen from a random set. The user draws a rough approximation of the image they are looking for, for example with blobs of color or general shapes. This query technique removes the difficulties that can arise when trying to describe images with words.

Since textual annotations are not available for most images, searching for particular pictures becomes an inherently difficult task. Luckily, a lot of research has been conducted over the last two decades leading to various approaches for content-based image retrieval. Content-based image retrieval (CBIR) does not rely on textual attributes but allows search based on features that are directly extracted from the images. This however is, not surprisingly, rather challenging and often relies on the notion of visual similarity between images or image regions. While humans are capable of effortlessly matching similar images or objects, machine vision research still has a long way to go before it will reach a similar performance for computers. Currently, many retrieval approaches are based on low-level features such as color, texture, and shape features, leaving a ‘semantic gap’ to the high-level understanding of users. Several approaches for bridging this gap have been introduced, such as relevance feedback or automatic image annotation, but much work still remains to be done for CBIR to become truly useful. Content-based image retrieval uses the visual contents of an image such as color, shape, texture, and spatial layout to represent and index the image. In typical content-based image retrieval systems, the visual contents of the images in the database are extracted and described by multi-dimensional feature vectors. The feature vectors of the images in the database form a feature database. To retrieve images, users provide the retrieval system with example images or sketched figures. The system then changes these examples into its internal representation of feature vectors. The similarities /distances between the feature vectors of the query example or sketch and those of the images in the database are then calculated and retrieval is performed with the aid of an indexing scheme. The indexing scheme provides an efficient way to search for the image database. Recent retrieval systems have incorporated users' relevance feedback to modify the retrieval process in order to generate perceptually and semantically more meaningful retrieval results. Recent works on color image coding using vector quantization has established that color and pattern information can be used as image indices for classification and retrieval purposes. The diagram indicating the process is shown below.

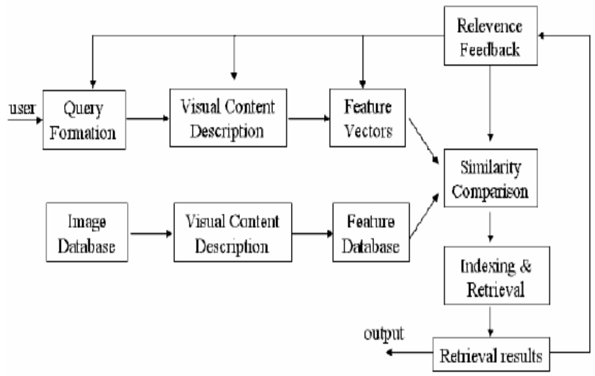

Fig 1 Process of Content Based Image Retrieval

B.Artificial neural networks

An artificial neural network, often just named a neural network, is a mathematical model inspired by biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In most cases a neural network is an adaptive system changing its structure during a learning phase. Neural networks are used for modeling complex relationships between inputs and outputs or to find patterns in data. The inspiration for neural networks came from examination of central nervous systems. In an artificial neural network, simple artificial nodes, called "neurons", "neurodes", "processing elements" or "units", are connected together to form a network which mimics a biological neural network.

Generally, it involves a network of simple processing elements exhibiting complex global behavior determined by the connections between the processing elements and element parameters. Artificial neural networks are used with algorithms designed to alter the strength of the connections in the network to produce a desired signal flow.

Neural networks are also similar to biological neural networks in performing functions collectively and in parallel by the units, rather than there being a clear delineation of subtasks to which various units are assigned. The term "neural network" usually refers to models employed in statistics, cognitive psychology and artificial intelligence. Neural network models which emulate the central nervous system are part of theoretical neuroscience and computational neuroscience.

In modern software implementations of artificial neural networks, the approach inspired by biology has been largely abandoned for a more practical approach based on statistics and signal processing. What has attracted the most interest in neural networks is the possibility of learning.

The cost function C is an important concept in learning, as it is a measure of how far away a particular solution is from an optimal solution to the problem to be solved. Learning algorithms search through the solution space to find a function that has the smallest possible cost.

For applications where the solution is dependent on some data, the cost must necessarily be a function of the observations; otherwise we would not be modeling anything related to the data. It is frequently defined as a statistic to which only approximations can be made.

Choice of model will depend on the data representation and the application. Overly complex models tend to lead to problems with learning.

Learning algorithm: There is numerous trades-offs between learning algorithms. Almost any algorithm will work well with the correct hyper parameters for training on a particular fixed data set. However selecting and tuning an algorithm for training on unseen data requires a significant amount of experimentation.
Robustness: If the model, cost function and learning algorithm are selected appropriately the resulting ANN can be extremely robust.

The utility of artificial neural network models lies in the fact that they can be used to infer a function from observations. This is particularly useful in applications where the complexity of the data or task makes the design of such a function by hand impractical.

The neural network classifier consists of three layers with an input layer, a hidden layer, and an output layer. The input layer has input nodes, the hidden layer has hidden nodes, and the output layer has output nodes. The neural network is trained and changed weights until the minimum error reduces to 0.1.

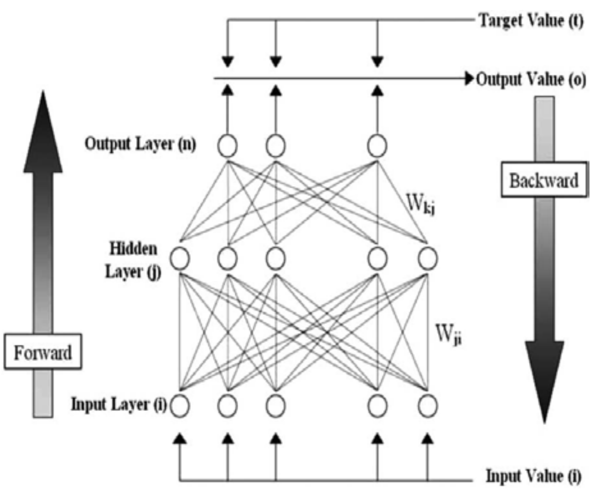

Fig 2 Various layers of an artificial neural network

C.Similarity Measure

Finding good similarity measures between images based on some feature set is a challenging task. On the one hand, the ultimate goal is to define similarity functions that match with human perception, but how humans judge the similarity between images is a topic of ongoing research. Many Current Retrieval systems take a simple approach by using typically norm-based distances (e.g., Euclidean distance) on the extracted feature set as a similarity function. The main premise behind these CBIR systems is that given a "good set" of features extracted from the images in the database (the ones that significantly capture the content of images.) then for two images to be "similar" their extracted features have to be "close" to each other. The Direct Euclidian Distance between an image P and query image Q can be given as the equation below.

$$ED = \sqrt{\sum_{i=0}^{n} (v_{pi} - v_{qi})^2}$$

Where, and v_{pi} and v_{qi} are the feature vectors of image P and Query image Q respectively with size 'n'. Beside the Euclidean Distance, there are many ways to measure feature distance between two images for example: Manhattan distance; the Mahalanobis Distance; Earth Mover's Distance (EMD) and the chord distance.

III. FEATURE VECTOR EXTRACTION

In pattern recognition and in image processing, extraction is a special form of dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant then the input data will be transformed into a reduced representation set of features (also named features vector). Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.

Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately.
Gabor Magnitude

The Gabor transform, named after Dennis Gabor, is a special case of the short-time Fourier transform. It is used to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time. The function to be transformed is first multiplied by a Gaussian function, which can be regarded as a window function, and the resulting function is then transformed with a Fourier transform to derive the time-frequency analysis. The window function means that the signal near the time being analyzed will have higher weight.

In image processing, a Gabor filter, named after Dennis Gabor, is a linear filter used for edge detection. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation and discrimination. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. The Gabor filters are self-similar: all filters can be generated from one mother wavelet by dilation and rotation. J. G. Daugman discovered that simple cells in the visual cortex of mammalian brains can be modeled by Gabor functions. Thus, image analysis by the Gabor functions is similar to perception in the human visual system.

Gabor filters are directly related to Gabor wavelets, since they can be designed for a number of dilations and rotations. However, in general, expansion is not applied for Gabor

ISSN 2250-3153

wavelets, since this requires computation of bi-orthogonal wavelets, which may be very time-consuming. Therefore, usually, a filter bank consisting of Gabor filters with various scales and rotations is created. The filters are convolved with the signal, resulting in a so-called Gabor space. This process is closely related to processes in the primary visual cortex. The Gabor space is very useful in image processing applications such as optical character recognition, iris recognition. Relations between activations for a specific spatial location are very distinctive between objects in an image. Furthermore, important activations can be extracted from the Gabor space in order to create a sparse object representation.

The filters of a Gabor filter bank are designed to detect different frequencies and orientations. We use them to extract features on key points detected by interest operators. From each filtered image, Gabor features can be calculated and used to retrieve images. The systematic steps for extracting the Gabor feature vector are shown in Figure 4. The algorithm (figure 5) and the equations related to extraction of Gabor features are also mentioned below.

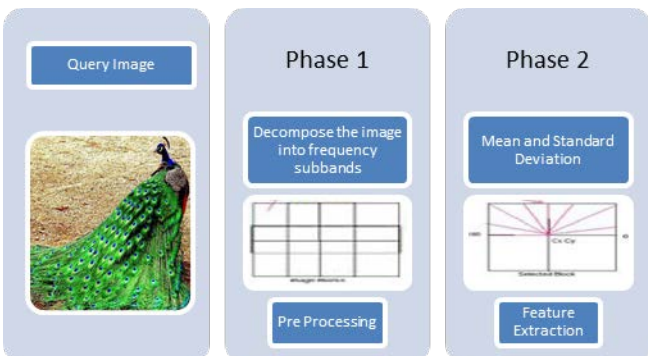

Fig 3 Gabor Filter Feature Extraction

For a given image I(x, y), the discrete Gabor wavelet transform is given by a convolution:

$$W_{mn} = \sum \sum I(x_1, y_1) g_{mn} * (x - x_1, y - y_1)$$

Where * indicates complex conjugate and where m, n specify the scale and orientations of wavelet respectively.

After applying Gabor filters on the image with different orientation different scale, we obtain an array of magnitudes:

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

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.

The standard deviation of the magnitude of the transformed coefficients is:

$$\sigma_{mn} = \sqrt{\frac{\sum_x \sum_y (W_{mn}(x,y) - \mu_{mn})^2}{P \times Q}}$$

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

A feature vector f (texture representation) is created using mn and mn as the feature components. M scales and N orientations are used and the feature vector is given by:

$$f = [\sigma_{00}, \sigma_{01}, \dots \dots \dots, \sigma_{(m-1)(n-1)}]$$

$$f_{gabor} = \frac{1 - \mu}{\sigma}$$

Where µ is the mean and is the standard deviation . The basic algorithm for Gabor feature extraction is shown below:

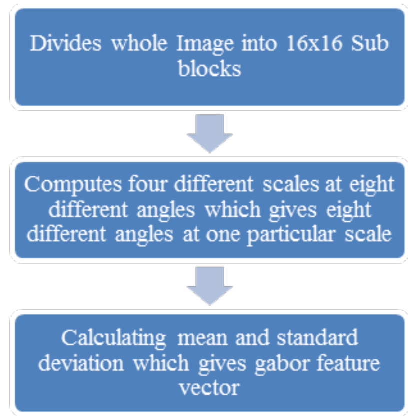

Fig 4 Gabor Filter Algorithm

IV. PROPOSED CBIR TECHNIQUE

1. We consider a database of images transferred in the hard disk. These images will be transferred to the Gabor filter.
2. Gabor filter will first scale the original image or query image into the size of 256 x 256 pixels. After scaling, the image is divided into equal blocks each of the size 16x16 pixels.
3. Applying Gabor Filter with different orientations and different scales on all the blocks of an image, we obtain a set of magnitudes for each block.
4. We compare the standard deviation for each orientation and obtain Gabor feature for that block and subsequently a feature vector is obtained for the entire image. Thus, features are calculated for each and every image and the resulted feature vectors are stored in feature vector database.
5. Label the feature vectors of images from all the classes (training).
6. Train ANN according to the label attached to the feature vectors.
7. Label the query image with the class it belongs.
8. Extract variables of the trained ANN for required class.
9. Perform classification of the images.
10. The indexed images are stored in the result folder.

IMPLEMENTATION

In this approach Gabor feature vector is obtained by calculating Standard Deviation of Gabor. The ANN is trained by labeling the feature vectors of the known images. The variables generated while training the ANN are then used for the purpose of classification and hence the images are retrieved.

V. RESULTS AND DISCUSSION

The methods ANN-Gabor Magnitude and only Gabor Magnitude Features were applied to the image database having 900 images spread across 10 categories. The query and database image matching is done using ANN classifier. The average precision and average recall are calculated by grouping the number of retrieved images sorted according to classification of database images.

Fig 5 Sample of Image Database

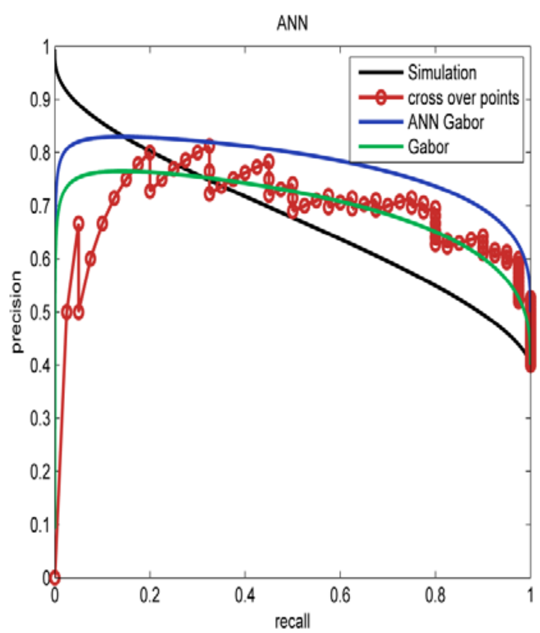

Fig 6 Precision Recall and Crossover Plot with ANN

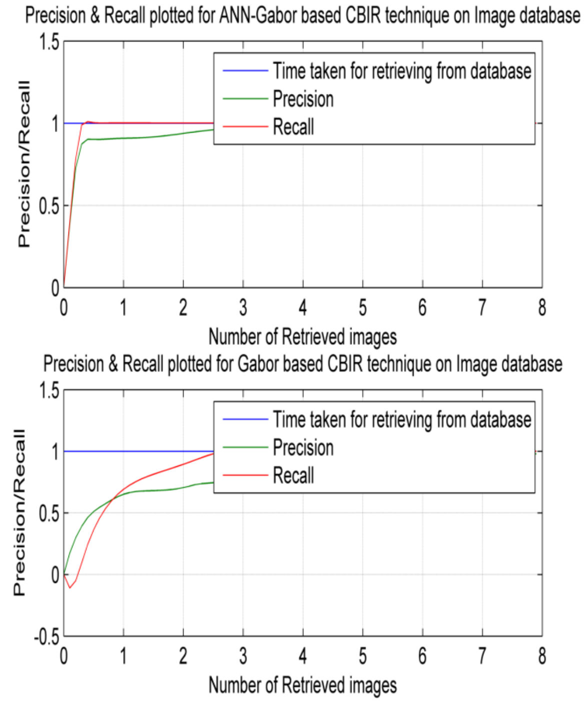

Fig 7 Precision Recall Comparative Plot for Gabor with ANN and without ANN

Figure 6 shows the graph of precision/recall values plotted for proposed image retrieval techniques. It can be seen that ANN-Gabor based image retrieval technique gives the highest precision/recall crossover values specifying the best performance. The crossover point varies for different image category. Figure 8 shows results obtained using Gabor Filtered image features based CBIR technique and results obtained using ANN-Gabor based CBIR technique.

There is a significant improvement in results using ANN based technique which can be seen clearly from the plots shown in Figure 6 and 7. The precision/recall values and crossover points from the plot proves that the discrimination capability of ANN - Gabor Magnitude based CBIR technique is better than Gabor Filtered image features based CBIR technique. However the distinction in the performance of all these techniques is not very clear. The height of crossover point of precision and recall curves plays very important role in performance comparison of CBIR methods. Ideally this crossover point height should be one. Higher the value of this crossover point better the performance is. A sample of the classified and retrieved images of 2 general categories is also shown below.

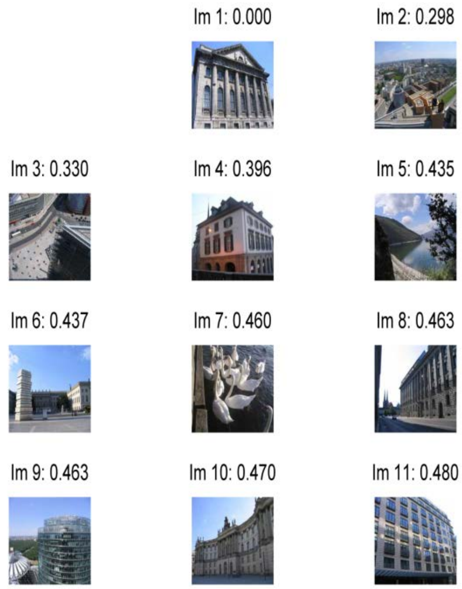

Fig 8 Sample of Retrieved Image Class

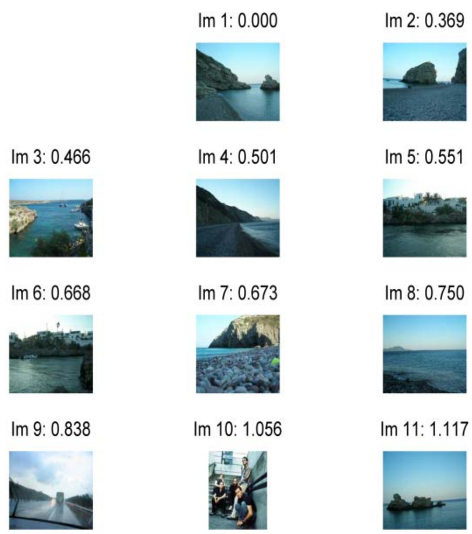

Fig 9 Sample of Retrieved Image class

VI. CONCLUSION

In this paper we have proposed a Content Based Image Retrieval System using Artificial Neural Network based on Gabor Filter's Response. The proposed system is giving higher Precision and Recall as compared to the CBIR technique when only Gabor magnitude features are used. The superiority of the system is because of the Gabor feature gives good response to texture of the image and makes it very clear and simple for ANN to classify and retrieve the required image.

REFERENCES

[1] H.B.Kekre, Sudeep D. Thepade, "Image Retrieval using Augmented Block Truncation Coding Techniques", ACM International Conference on Advances in Computing, Communication and Control (JCAC3-2009), pp. 384-390, 23-24 Jan 2009, Fr.ConceicaoRodrigous College of Engg., Mumbai. Is uploaded on online ACM portal.

[2] Minh N. Do, Martin Vetterli, "Wavelet-Based Texture Retrieval Using Generalized Gaussian Density and Kullback-Leibler Distance", IEEE Transactions On Image Processing, Volume II, Number 2, pp.146-158, February 2002.

[3] B.G.Prasad, K.K. Biswas, and S. K. Gupta, "Region –based image retrieval using integrated color, shape, and location index", International Journal on Computer Vision and Image Understanding Special Issue: Colour for Image Indexing and Retrieval,Volume 94, Issues 1-3, April-June 2004, pp.I93-233.

[4] H.B.Kekre, Tanuja Sarode, Sudeep D. Thepade, "OCT Applied to Row Mean and Column Vectors in Fingerprint Identification", In Proceedings of International Conference on Computer Networks and Security (ICCNS), 27-28 Sept. 2008, VIT, Pune .

[5] Zhibin Pan, Kotani K., Ohmi T., "Enhanced fast encoding method for vector quantization by finding an optimally-ordered Walsh transform kernel", ICIP 2005, IEEE International Conference, Volume I, pp I - 573-6, Sept. 2005.

[6] H.B.kekre, Sudeep D. Thepade, "Improving 'Color to Gray and Back' using Kekre's LUV Color Space", IEEE International Advanced Computing Conference 2009 (IACC'09), Thapar University, Patiala, INDIA, 6-7 March 2009. Is uploaded and available online at IEEE Xplore.

[7] H.B.Kekre, Sudeep D. Thepade, "Color Traits Transfer to Grayscale Images", In Proc.of IEEE First International Conference on Emerging Trends in Engg. & Technology, (JCETET-08), G.H.Raisoni COE, Nagpur, INDIA. Uploaded on online IEEE Xplore.

[8] H.B.Kekre, Sudeep D. Thepade, "Scaling Invariant Fusion of Image Pieces in Panorama Making and Novel Image Blending Technique", International Journal on Imaging ([JI), www.ceser.res.inliji.html. Volume I. No. A08. pp. 31-46, Autumn 2008.

[9] H.B.Kekre, Sudeep D. Thepade, "Creating the Color Panoramic View using Medley of Grayscale and Color Partial Images ", W ASET International Journal of Electrical, Computer and System Engineering (IJECSE), Volume 2, No. 3, Summer 2008.

[10] H.B.Kekre, Sudeep D. Thepade, "Image Blending in Vista Creation using Kekre's LUV Color Space", SPIT –IEEE Colloquium and International Conference, Sardar Patel Institute of Technology, Andheri, Mumbai, 04-05 Feb 2008.

[11] H.B.Kekre, Sudeep D. Thepade, "Rendering Futuristic Image Retrieval System", National Conference on Enhancements in Computer, Communication and Information Technology, EC2IT-2009, 20-21 Mar 2009, K.J.Somaiya College of Engineering, Vidyavihar, Mumbai-77.

[12] H.B.Kekre, Sudeep D. Thepade, "Image Retrieval using Non Involutional Orthogonal Kekre's Transform", International Journal of Multidisciplinary Research and Advances in Engineering (TJMRAE), Ascent Publication House, 2009, Volume I, No.1, pp 189-203, 2009. Abstract available online at www.ascent-journals.com (ISSN: 0975-7074)

[13] H.B.Kekre, Sudeep D. Thepade, ArchanaAthawale, Anant Shah, PrathmeshVerlekar, SurajShirke,"Walsh Transform over Row Mean and Column Mean using Image Fragmentation and Energy Compaction for Image Retrieval", International Journal on Computer Science and Engineering (UCSE), Volume 2S, Issuel, January 2010, (ISSN: 0975-3397).

[14] H.B.Kekre, Tanuja Sarode, Sudeep D. Thepade,"Color-Texture Feature based Image Retrieval using DCT applied on Kekre's Median Codebook", International Journal on Imaging (UI), Volume 2, Number A09, Autumn 2009,pp. 55-65. Available online at www.ceser.res.in/iji.html (ISSN: 0974-0627).

[15] H.B.Kekre, Sudeep D. Thepade, "Boosting Block Truncation Coding using Kekre's LUV Color Space for Image Retrieval", W ASET International Journal of Electrical, Computer and System Engineering (TJECSE), Volume 2, Number 3, pp. 172- 180, Summer 2008.

AUTHORS

First Author – Swathi Rao G, M.tech, Dehradun Institute of Technology, swaumababu@gmail.com
Second Author – Manipriya Singh, Mtech , SRM University, manipriyasingh01@gmail.com
Third Author – Ramnath Panwar Singh, B.tech, Veera College of Engineering, rammani.pr21@gmail.com