Energy based Comparative Study of CBIR Techniques and a Novel approach of Image Splitting in the Frequency Domain for Efficient Retrieval

H. B. Kekre
Senior Professor
Mukesh Patel School of Technology Management and Engineering,
Mumbai-56

Archana B. Patankar
Assistant Professor
Thadomal Shahani Engineering College, Bandra (W), Mumbai-50

Aniruddh Munde
Student
Thadomal Shahani Engineering College, Bandra (W), Mumbai-50

Disha Madhrani
Student
Thadomal Shahani Engineering College, Bandra (W), Mumbai-50

Dheeraj Nagpal
Student
Thadomal Shahani Engineering College, Bandra (W), Mumbai-50

ABSTRACT
In this paper a comparative study of the existing Content Based Image Retrieval (CBIR) techniques is presented. Also a novel idea of tiling the images after transforming them into the frequency domain is proposed. The transformed images are broken down into fragments of 4X4 and feature extraction is done after taking the average of the energies of the corresponding fragments, while always selecting the highest energy coefficient. This approach provided the highest crossover point of 0.602 for Kekre Complete Transform. The compared techniques were applied to a database of 600 images spread over 12 different categories. The comparison of transforms considers 99.5% of the total energy contained in a query image and uses only those transform coefficients which contribute to this energy. The results obtained from such a novel energy based comparison show that the percentage of feature vector coefficients to be used for query execution can be as low as 12.89% as seen for Haar Column Transform.

General Terms
CBIR, Image Transforms

Keywords
CBIR, Row mean, Column mean, Image Transform, DCT, DST, Haar, Hartley, Kekre, Walsh, Image Splitting, Energy Compaction, Precision, Recall, LSRR

1. INTRODUCTION
The use of images in database technology has increased substantially in the past few years. For a long time, images were represented by textual description and observation which failed to capture the vividness of the images. This led to the rise of Content Based Image Retrieval (CBIR), which is a technique in Image Processing used for efficient retrieval and sequencing of images [1], [2], [3]. It is mainly used for efficient retrieval of images from a huge database based on some automatically extracted features. These features are automatically extracted from properties such as shape, color and texture of query image and the various images in the repository [4].

2. REVIEW OF LITERATURE
Many approaches are available in literature for content based image retrieval based on various techniques and choice of desired image components for feature extraction. Various image components like color, texture, shape, semantic image have proven useful for image retrieval methods.

Xiaohong et al. [5] introduced the basic components of content-based image retrieval system and have stressed on the semantic-based image retrieval method. Liu et al. [6] proposed an image retrieval method based on feature weight assignment operators with combined color and texture as the features. Krishnan et al. [2] proposed a method to retrieve images based on dominant colors in the foreground image.

Various transforms have also been applied for content based image retrieval system. Kekre et al. [7] presented novel image retrieval techniques based on discrete cosine transforms applied on row mean, column mean and combination for feature extraction along with image fragmentation. In [8], novel content based image retrieval techniques using row mean, column mean, energy compaction and image splitting using Kekre transform were proposed.

Kekre et al. [4] proposed Boosting Block Truncation Coding method using Kekre's LUV color space for image retrieval. In [3] novel image retrieval techniques like Kekre-Gray and Kekre-RGB using Kekre transform were presented. In [9] an image retrieval technique using the color-texture features extracted from images based on vector quantization with Kekre's fast codebook generation was proposed.

This paper discusses a number of CBIR techniques, puts forward a comparative study of all the results obtained and proposes a novel way of using image splitting for retrieval.

3. PROPOSED CBIR TECHNIQUES
Image retrieval mainly has two steps: Feature Extraction and Query Execution.

3.1 Feature Extraction
For each of the transforms, feature vectors can be extracted in three ways: applying only column transform, applying only row transform or applying the complete transform.

When only column or row transform is employed, row or column mean of the resultant is used as the feature vector for each image respectively. Absolute value of the coefficients is to be used otherwise they might cancel out each other while taking the mean. The final feature vector is obtained after taking the mean of the feature vectors of all the images in the database and performing energy compaction. This energy compaction is to be done by arranging the mean feature vector elements in descending order but maintaining their true positions and considering only the first few coefficients that contribute to 99.5% of the total energy, which is obtained by summing the square of the coefficients. The positions of such coefficients are to be used for finding Euclidean distance during query execution. Here Euclidian distance [10]-[14] is used as similarity measure. The direct Euclidian distance between an image P and query image Q can be given as below:

$$ED = \sqrt{\sum_{i=1}^{n} (Vpi - Vqi)^2} \quad (1)$$

where Vp and Vq be the feature vectors of image P and query image Q respectively with size 'n'.

Such consideration of only those major feature vector coefficients which contribute to most of the energy in the image is what distinguishes this comparative study from the ones in Kekre et al.[15], [16] which utilize reduced percentage of feature vector coefficients irrespective of their contribution to the image energy.

Each image of the database resized to 256*256 while performing row/column feature extraction is considered as a 64*64 image with 4*4 blocks taken as single units while using the complete transform. It is this novel approach proposed here in which after applying the complete transform, the energy in each of these 4*4 blocks is found and averaged over all the images. However, the highest energy coefficient is kept separate in the calculation and always selected in the feature vector. After arranging the blocks in descending order but maintaining their positions, each block's contribution to the total resultant energy is calculated and only those first few who contribute to 99.5% of the total energy are considered for feature extraction. The final feature vector, a 256*256 matrix, is obtained by assigning '1's to all the positions of such blocks otherwise a '0' is assigned. The highest energy coefficient position is always assigned a '1' since it always needs to be considered for finding Euclidean distance during query execution.

The technique of image splitting has been used earlier in Kekre et al.[8]; however, it was done before the transform was applied. The technique proposed here splits the frequency domain image i.e. after applying the complete transform for efficient retrieval.

3.2 Query Execution

Depending on the method of feature extraction, the query image is processed accordingly. In column/ row transform technique, the mean row/column feature vector of the query image as well as that of all the images in the database is obtained after applying the corresponding transform. For finding the Euclidean distance between the feature vectors of the query image and each of the images in the database, only the positions of the coefficients obtained from the final feature vector after energy compaction are considered.

When feature extraction is performed using the complete transform, the Euclidean distance is found by considering only those positions from the transformed query and database images which have a '1' in the feature vector. After finding the Euclidean distance and arranging its values in ascending order for all the images in the database, precision and recall are calculated to compare the various transforms. The following formulae [17], [18] are employed for the same:

$$Precision = \frac{Number\_of\_relevant\_images\_retrieved}{Total\_number\_of\_images\_retrieved} \quad (2)$$

$$Recall = \frac{Number\_of\_relevant\_images\_retrieved}{Total\_number\_of\_relevant\_images\_in\_database} \quad (3)$$

4. RESULTS

The CBIR techniques are tested on the image database [19], [20],[21],[22] of 600 variable sized coloured images spread across 12 categories of different flora, fauna, buildings, babies, cars, fireworks and clouds. Sample images from each category are shown in Figure 1.

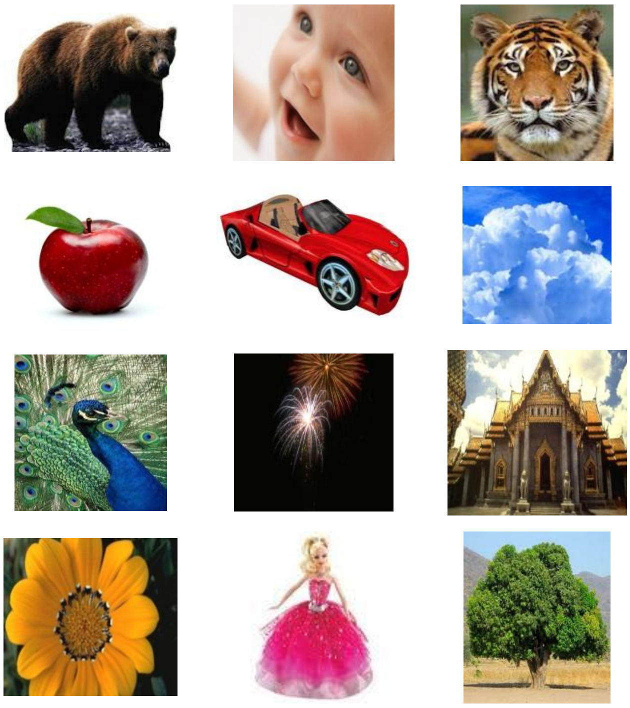

Figure 1: Sample images from the database

Total 60 queries (5 from each category of image database) are tested to get average precision, average recall and average Length of String required to Retrieve all Relevant images (LSRR) [23].

By considering 99.5% of the total energy, the percentage of feature vector coefficients to be considered for query execution for efficient computation is found out for the respective techniques.

The images in the database being color images Euclidean distance is found separately for the RGB planes. While calculating precision/recall of first n retrieved images, an m^{th} retrieved image is considered to be relevant if the m^{th} image in the list of n retrieved images of 'any one' of the RGB planes is relevant.

A cross over point defined as that value of precision when it is equal to recall is obtained by plotting precision/recall versus number of images retrieved. This cross over point is seen when the retrieved images equal the number of relevant images in the database (in this case 50).

A comparative study of the cross over points obtained for the various transforms is depicted in the bar graph shown in Figure 2. Apart from subtle differences between each of them, the similarity is the fact that their cross over point is between 0.5 and 0.6 and that the complete transform techniques have a greater cross over than their column or row counterparts.

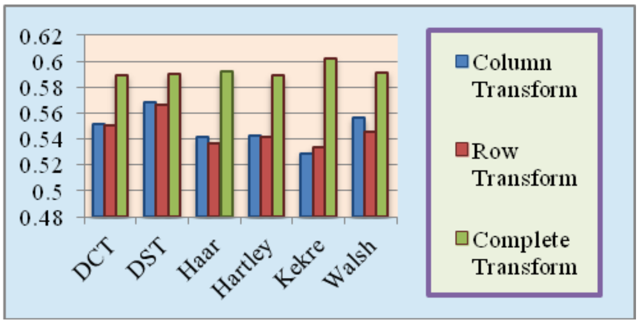

Figure 2: Comparative Study

Table 1 gives the values for the cross over point. It can be seen that the Kekre Complete Transform technique with a cross over point equal to 0.602 achieves the best results amongst all the proposed CBIR techniques.

Table 1. Cross Over Points

Transform	Column Transform	Row Transform	Total Transform
DCT	0.5520	0.5503	0.5893
DST	0.5680	0.5660	0.5900
Haar	0.5417	0.5370	0.5917
Hartley	0.5423	0.5417	0.5890
Kekre	0.5290	0.5333	0.6020
Walsh	0.5563	0.5453	0.5917

Comparing all the proposed column transform techniques, the DST Column Transform technique with a cross over point of 0.5680 is found to be the most promising. While DST Row Transform technique with a cross over point of 0.5660 gives better results than any of the other proposed row transform techniques.

Figure 3 shows the plot of precision/recall versus number of images retrieved for Kekre Complete Transform while Figures 4 and 5 show the same for DST Column Transform and DST Row Transform respectively.

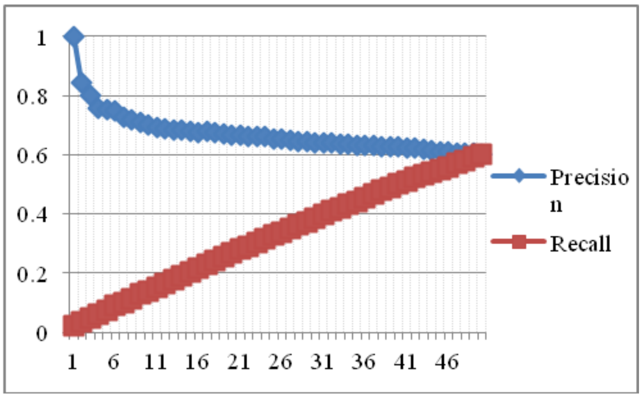

Figure 3: Kekre Complete Transform

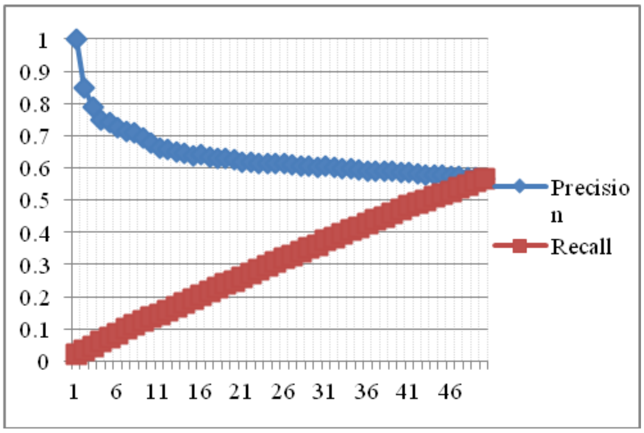

Figure 4: DST Column Transform

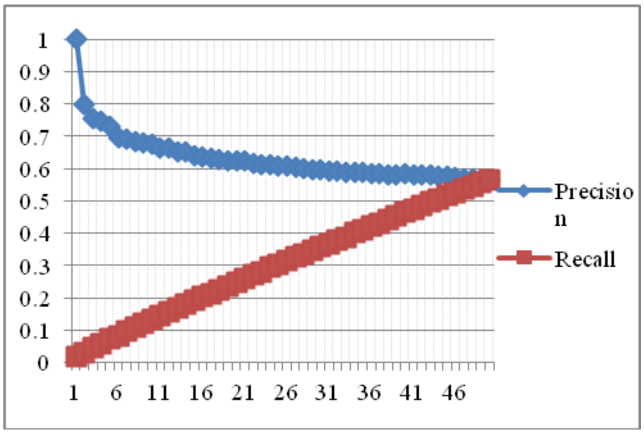

Figure 5: DST Row Transform

Another parameter comparing the CBIR techniques is the average Length of string required to Retrieve all Relevant images (LSRR) in the database. For each query image, there are 50 relevant images in the image database. LSRR is the minimum number of images required to be retrieved so that all these 50 relevant images are included in them which is true even if they appear in any one of the RGB planes' list of retrieved LSRR images. The average LSRR calculated over 5 query images of all the 12 categories of images for all the proposed CBIR techniques is listed in Table 2.

From Table 2, it is clear that DST Column transform technique retrieves all the relevant images with the smallest LSRR amongst all others. It requires on an average 386.9 i.e. 387 images to be retrieved out of the total 600 in the database so that all the 50 relevant images are included in them. With regards to all row transform techniques, DST Row Transform gives an average LSRR of 389.2333 better than any others. While Kekre Complete Transform with 396.0 average LSRR is the best when all complete transform techniques are considered.

Table 2. Average LSRR

CBIR Technique		Average LSRR	Min LSRR	Max LSRR
DCT	Column	394.6833	255	555
	Row	394.9333	208	558
	Complete	400.7500	230	570
	Column	386.9000	192	556

DST	Row	389.2333	195	562
	Complete	400.7833	230	570
Haar Transform	Column	394.3000	252	555
	Row	396.0833	216	558
	Complete	400.7000	229	570
Hartley Transform	Column	395.6333	262	556
	Row	393.6833	205	557
	Complete	400.7333	229	570
Kekre Transform	Column	393.5000	204	555
	Row	394.8000	207	556
	Complete	396.0000	233	570
Walsh Transform	Column	394.0667	239	555
	Row	394.8167	206	558
	Complete	400.6833	229	570

A final parameter comparing the CBIR techniques is the percentage of feature vector coefficients required for query execution after considering 99.5% of the total energy. The smaller the percentage the greater is the computation efficiency. Table 3 lists the percentages for all the proposed techniques.

Computation efficiency for complete transforms is obviously much worse than that for column/row transform techniques as is evident from the greater percentages the complete transform techniques require. Kekre Transform is however an exception to this observation; its complete transform technique uses lesser percentage of feature vector coefficients (69.0348%) as compared to its column/ row transform techniques (>75%). Haar Column Transform technique succeeds in being the most computationally efficient technique by using only 12.8906% of the feature vector coefficients. Closely followed by is the Haar Row Transform technique with 14.0625% DST Complete Transform technique with 46.0449% is considerably more efficient than other complete transform techniques.

Table 3. Percentage of Feature Vector Coefficients to be used

CBIR Technique		Percentage of Feature Vector Coefficients to be used
DCT	Column	15.1042
	Row	16.7969
	Complete	53.9388
DST	Column	26.1719
	Row	27.6042
	Complete	46.0449
Haar Transform	Column	12.8906
	Row	14.0625
	Complete	71.6146
Hartley Transform	Column	16.5365
	Row	17.7083
	Complete	55.5908
Kekre Transform	Column	81.7708
	Row	78.7760
	Complete	69.0348
Walsh Transform	Column	22.3958
	Row	24.6094
	Complete	69.8731

5. CONCLUSION

Amongst the myriad of CBIR techniques, a specific application would best decide which one would prove more useful. Nevertheless, as far as precision and recall at the cross over point are considered, Kekre Complete Transform is most suitable. Besides, it also has a smaller average LSRR as compared to other complete transform techniques. The percentage of feature vector coefficients to be used is lesser for Kekre Complete Transform than its column/row counterparts, which makes it even efficient to use. If better average LSRR is the need of the application, DST Column/Row Transform techniques are the catch with good precision/recall as well. Better computing efficiency can be provided by Haar Column/Row transform techniques since they require much less percentage of feature vector coefficients to be used.

6. REFERENCES

[1] H.B.Kekre, Sudeep D. Thepade, "Rendering Futuristic Image Retrieval System", In Proc. Of National Conference EC2IT-2009, KJSCOE, Mumbai, 20-21 Mar 2009.

[2] Dr.N.Krishnan, M.Sheerin Banu, C.Callins Christiyana, "Content Based Image Retrieval using Dominant Color Identification Based on Foreground Objects", In International Conference on Computational Intelligence and Multimedia Applications, 2007.

[3] H.B.Kekre, Sudeep D. Thepade, "Image Retrieval using Non-Involutional Orthogonal Kekre's Transform", International Journal of Engineering Research and Industrial Applications (IJERIA), Ascent Publication House, 2009, Volume 2, No.VII, 2009.

[4] H.B.Kekre, Sudeep D. Thepade, "Boosting Block Truncation Coding using Kekre's LUV Color Space for Image Retrieval", WASET Int. Journal of Electrical, Computer and System Engineering (IJECSE), Vol.2, Num.3, Summer 2008.

[5] Yu Xiaohong, Xu Jinhua, "The Related Techniques of Content-based Image Retrieval", In International Symposium on Computer Science and Computational Technology, 2008.

[6] Pengyu Liu, Kebin Jia, Zhuozheng Wang, Zhuoyi Lv, "A New and Effective Image Retrieval Method Based on Combined Features", In Fourth International Conference on Image and Graphics,2007.

[7] Dr. H.B.Kekre, Sudeep D. Thepade, Archana Athawale, Anant Shah, Prathamesh Verlekar, Suraj Shirke "Image retrieval using DCT on Row Mean, Column Mean and Both with Image Fragmentation", In Proc. Of the International Conference and Workshop on Emerging Trends in Technology, Mumbai, 26-27 Feb 2010.

[8] Dr.H.B.Kekre, Sudeep D. Thepade, Archana Athawale, Anant Shah, Prathamesh Verlekar, Suraj Shirke "Energy Compaction and Image Splitting for Image Retrieval using Kekre Transform over Row and Column Feature Vectors", IJCSNS Int. Journal of Computer Science and Network Security, Vol.10, No. 1, January 2010, pp. 289-298.

[9] H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, "Image Retrieval using Color-Texture Features from DCT on VQ Codevectors obtained by Kekre's Fast Codebook Generation", ICGST International Journal on Graphics,

Vision and Image Processing (GVIP), Volume 9, Issue V, Sept. 2009, pp. 1-8. http://www.icgst.com/gvip

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

[11] 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, (ICETET-08), G.H.Raisoni COE, Nagpur, INDIA.

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

[13] H.B.Kekre, Sudeep D. Thepade, "Scaling Invariant Fusion of Image Pieces in Panorama Making and Novel Image Blending Technique", Int. Journal on Imaging (IJI), Autumn 2008, Volume 1, No. A08.

[14] H.B.Kekre, Sudeep D. Thepade, "Rotation Invariant Fusion of Partial Images in Vista Creation", WASET Int. Journal of Electrical, Computer and System Engg.(IJECSE), Vol.2, No.2, Spring 2008.

[15] Dr. H. B. Kekre, Dr. Sudeep D. Thepade & Akshay Maloo, "Comprehensive Performance Comparison of Cosine, Walsh, Haar, Kekre, Sine, Slant and Hartley transforms for CBIR with Fractional Coefficients of Transformed Image", International Journal of Image Processing (IJIP), Volume (5) , Issue (3), 2011, pp. 336-351.

[16] H. B. Kekre, Sudeep D. Thepade, Archana A. Athawale, Paulami Shah, "Image Retrieval using Fractional Energy of Row Mean of Column Transformed Image with Six Orthogonal Image Transforms", International Journal of Soft Computing and Engineering (IJSCE), Volume-1, Issue-4, September 2011, pp. 168-173.

[17] H.B.Kekre, Sudeep D. Thepade, "Improving the Performance of Image Retrieval using Partial Coefficients of Transformed Image", International Journal of Information Retrieval (IJIR), Serials Publications, Volume 2, Issue 1, 2009, pp. 72-79(ISSN: 0974-6285).

[18] H.B.Kekre, Sudeep D. Thepade, "Image Retrieval using Augmented Block Truncation Coding Techniques", ACM International Conference on Advances in Computing, Communication and Control (ICAC3-2009), pp. 384-390, 23-24 Jan 2009, Fr. Conceicao Rodrigous College of Engg., Mumbai.

[19] http://www.mp-webspace4.com/media/iStock_000000466_1.jpg

[20] http://www.schroedersorchard.com/Schroedersorchard/Contact_Us_files/red%20apple_6.png

[21] http://www.popular-toys.com/wp-content/uploads/2011/01/barbieclassic.jpg

[22] Macmillan Digital Publishing, BPB 555,000 Graphics pack.

[23] Dr. H.B. Kekre, Dhirendra Mishra, Anirudh Kariwala, "A Survey of CBIR Techniques and Semantics", International Journal of Engineering Science and Technology (IJEST), Vol. 3, No. 5, May 2011, pp. 4510-4517.

AUTHOR PROFILE

Dr. H. B. Kekre has received B.E. (Hons.) in Telecomm. Engineering from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970. He has worked as Faculty of Electrical Engg. and then HOD Computer Science and Engg. at IIT Bombay. For 13 years he was working as a professor and head in the Department of Computer Engg. at Thadomal Shahani Engineering. College, Mumbai. Now he is Senior Professor at MPSTME, SVKM's NMIMS. He has guided 17 PhDs, more than 100 M.E./M.Tech and several B.E./B.Tech. projects. His areas of interest are Digital Signal processing, Image Processing and Computer Networking. He has more than 400 papers in National / International Conferences and Journals to his credit. He was Senior Member of IEEE. Presently He is Fellow of IETE and Life Member of ISTE. Recently twelve students working under his guidance have received best paper awards and six scholars have been awarded Ph.D. by NMIMS University. Currently seven research scholars are pursuing Ph.D. program under his guidance.

Dr Archana B. Patankar has received M.E. (Computer Engineering) from V.J.T.I, Mumbai University in 1999, and Ph.D. from NMIMS (Deemed to be University), Mumbai. She has more than 12 years of experience in teaching and is presently working as an Assistant Professor in Department of Computer Engineering at Thadomal Shahani Engineering College, Mumbai. Her areas of interest are Image Processing, Signal Processing and Computer Graphics. She has more than 50 papers in National /International Conferences/Journals to her credit. She has been a reviewer for International and National Journals/ Conferences. She has guided more than 20 M.E. projects and several B.E. projects. She is also a life member of ISTE and a member of CSI.

Aniruddh Munde, an NTSE scholar, is currently pursuing his B.E. degree in Computer Engineering from Thadomal Shahani Engineering College, Bandra (W), Mumbai. He has won 1st prize in National Level SJIIT Computer Quiz held in Pune in 2004. He is a member of Computer Society of India (CSI) since 2008. He was a member of IEEE Computer Society from 2009 to 2010. His areas of interest are databases, image processing, pattern recognition, operating systems, computer networks and computer programming.

Disha Madhrani is currently pursuing her B.E. degree in Computer Engineering from Thadomal Shahani Engineering College, Bandra (W), Mumbai. Her areas of interest are image processing, artificial intelligence, web designing and computer programming.

Dheeraj Nagpal is currently pursuing his B.E. degree in Computer Engineering from Thadomal Shahani Engineering College, Bandra (W), Mumbai. He has won 1st prize in National Level Technical Paper Presentation held in Mumbai. He was a member of IEEE Computer Society from 2009 to 2010. His areas of interest are image processing, pattern recognition, neural networks, artificial intelligence, web technology and computer programming.