Performance Comparison of Image Classifier using Discrete Cosine Transform and Walsh Transform

H.B. Kekre
Senior Professor, Computer Engineering, MP'STME, SVKM'S NMIMS University, Mumbai, India

Tanuja K. Sarode
Asst.Professor, Thadomal Shahani Engineering College, Mumbai, India

Meena S. Ugale
Lecturer, Xavier Institute of Engineering, Mumbai, India

ABSTRACT
In recent years, the accelerated growth of digital media collections and in particular still image collections, both proprietary and on the Web, has established the need for the development of human-centered tools for the efficient access and retrieval of visual information. The need to manage these images and locate target images in response to user queries has become a significant problem. Image categorization is an important step for efficiently handling large image databases and enables the implementation of efficient retrieval algorithms.

Image Classification refers to grouping of a digital image into different classes within a particular dataset, based on attribute values. It is done to replace visual analysis of the image data with quantitative techniques.

This paper presents the image classification techniques based on feature vectors of transformed images using Discrete Cosine Transform and Walsh Transform. The various sizes of feature vectors are generated such as 8X8, 16X16, 32X32, 64X64 and 128X128. The proposed algorithm is worked over database of 1000 images spread over 10 different classes. The Euclidean distance is used as similarity measure. A threshold value is set to determine to which category the query image belongs to.

General Terms
Algorithm, Experimentation

Keywords
Discrete Cosine Transform (DCT), Walsh Transform, Image Database, Transform Domain, Feature Vector

1. INTRODUCTION
Image Classification refers to grouping of a digital image into different classes within a particular dataset, based on attribute values. It is done to replace visual analysis of the image data with quantitative techniques.

The need to manage these images and locate target images in response to user queries has become a significant problem [26]. Image classification is an important task for many aspects of global change studies and environmental applications.

In recent years, the accelerated growth of digital media collections and in particular still image collections, both proprietary and on the Web, has established the need for the development of human-centered tools for the efficient access and retrieval of visual information. As the amount of information available in the form of still images continuously increases, the necessity of efficient methods for the retrieval of the visual information becomes evident [30].

Image categorization is an important step for efficiently handling large image databases and enables the implementation of efficient retrieval algorithms. It can help users to organize and to browse images. Although this is usually not a very difficult task for humans, it has been proved to be an extremely difficult problem for computer programs.

Digital image processing is a collection of techniques for the manipulation of digital images by computers. Classification generally comprises four steps [27]:

1. Pre-processing: E.g. atmospheric correction, noise suppression, and finding the band ratio, principal component analysis, etc.
2. Training: Selection of the particular feature which best describes the pattern.
3. Decision: Choice of suitable method for comparing the image patterns with the target patterns.
4. Assessing the accuracy of the classification.

Image classification refers to the labeling of images into one of predefined semantic categories.

The most important step in an Image Classification system is the image description. Indeed, features extraction gives a feature vector per image which is a reduced representation of the image visual content, because images are too big to be used directly for indexing and retrieval [30].

In this paper the use of Discrete Cosine Transform (DCT) and Walsh Transform is investigated. This paper presents the image classification techniques based on feature vectors of transformed images using Discrete Cosine Transform and Walsh Transform. A feature vector is extracted for an image of size N X N using DCT or Walsh Transform. The similarity measurement (SM), where a distance (e.g., Euclidean distance) between the query image and each image in the database using their feature vectors is computed so that the top ―closest images can be retrieved [7, 14, 17].

2. RELATED WORK

Many image classification systems have been developed since the early 1990s. Various image representations and classification techniques are adopted in these systems: the images are represented by global features, block-based features, region-based local features, or bag-of-words features[8], and various machine learning techniques are adopted for the classification tasks, such as K-nearest neighbor (KNN)[24], Support Vector Machines (SVM)[24], Hidden Markov Model(HMM)[21], Diverse Density(DD)[29], DD-SVM[28] and so on.

Recently, a popular technique for representing image content for image category recognition is the bag of visual word model [10, 6].

In the indexing phase, each image of the database is represented using a set of image attribute, such as color [25], shape [9, 1], texture [2] and layout [26]. Extracted features are stored in a visual feature database. In the searching phase, when a user makes a query, a feature vector for the query is computed. Using a similarity criterion, this vector is compared to the vectors in the feature database.

A heterogeneous image recognition system based on content description and classification is used in which for image database several features extraction methods are used and applied to better describes the images content. The features relevance is tested and improved through Support Vectors Machines (SVMs) classifier of the consequent images index database [26].

In literature there are various Image classification methods. Some of these methods use wavelets transform and support vector machine [33]; some methods use effective algorithm for building codebooks for visual recognition [14]; some advanced image classification techniques use Artificial Neural Networks, Support Vector Machines, Fuzzy measures and Genetic Algorithms [23] whereas some methods are proposed for classifying images, which integrates several sets of Support Vector Machines (SVM) on multiple low level image features [32].

3. DISCRETE COSINE TRANSFORM (DCT)

In general, neighboring pixels within an image tend to be highly correlated. As such, it is desired to use an invertible transform to concentrate randomness into fewer, decorrelated parameters [13].The Discrete Cosine Transform (DCT) has been shown to be near optimal for a large class of images in energy concentration and decorrelating. It has been adopted in the JPEG and MPEG coding standards [12][3]. The DCT decomposes the signal into underlying spatial frequencies, which then allow further processing techniques to reduce the precision of the DCT coefficients consistent with the Human Visual System (HVS) model. The DCT coefficients of an image tend themselves as a new feature, which have the ability to represent the regularity, complexity and some texture features of an image and it can be directly applied to image data in the compressed domain [31]. This may be a way to solve the large storage space problem and the computational complexity of the existing methods.

For the full 2-Dimensional Walsh transform applied to image of size NxN, the number of additions required are $2N^2(N-1)$ and absolutely no multiplications are needed in Walsh transform [18].

The two dimensional DCT can be written in terms of pixel values *f(i, j)* for *i,j* = 0,1,…,*N*-1 and the frequency-domain transform coefficients *F(u,v)*:

$$F(u,v) = \frac{1}{\sqrt{2N}} C(u)C(v) \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} f(i,j) \times \cos\left[\frac{(2i+1)u\pi}{2N}\right]\cos\left[\frac{(2j+1)v\pi}{2N}\right]$$

$$for\ u,v = 0,1,\ldots,N-1$$

$$where \quad C(x) = \begin{cases} \frac{1}{\sqrt{2}} & for\ x = 0 \\ 1 & otherwise \end{cases}$$

The inverse DCT transform is given by

$$f(i,j) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} C(u)C(v)F(u,v) \times \cos\left[\frac{(2i+1)u\pi}{2N}\right]\cos\left[\frac{(2j+1)v\pi}{2N}\right]$$

$$for\ i,j = 0,1,\ldots,N-1$$

The DCT tends to concentrate information, making it useful for image compression applications and also helping in minimizing feature vector size in CBIR [23]. For full 2-Dimensional DCT for an NxN image the number of multiplications required are $N^2(2N)$ and number of additions required are $N^2(2N-2)$.

4. WALSH TRANSFORM

Walsh transform matrix [18] is defined as a set of N rows, denoted Wj, for j = 0, 1, ... , N - 1, which have the following properties:

- Wj takes on the values +1 and -1.
- Wj[0] = 1 for all j.
- $W_j \times W_K^T = 0$, for $j \neq k$ and $W_j \times W_K^T = N$, for j=k.
- Wj has exactly j zero crossings, for j = 0, 1, ..., N-1.
- Each row Wj is even or odd with respect to its midpoint.

Walsh transform matrix is defined using a Hadamard matrix of order N. The Walsh transform matrix row is the row of the Hadamard matrix specified by the Walsh code index, which must be an integer in the range [0... N -1]. For the Walsh code index equal to an integer j, the respective Hadamard output code has exactly j zero crossings, for j = 0, 1... N - 1.

Table 1. Computational Complexity for applying transforms to image of size NxN [18]

	DCT	Walsh
Number of Additions	$2N^2(N-1)$	$2N^2(N-1)$
Number of Multiplications	$N^2(2N)$	0
Total Additions for transform of 128 x128 image	37715968	4161536

[Here one multiplication is considered as eight additions for last row computations]

5. PROPOSED ALGORITHM

The proposed algorithm makes use of well known Discrete Cosine Transform (DCT) and Walsh Transform to generate the feature vectors for the purpose of search and retrieval of database images.

For the DCT or Walsh transform, we convert an RGB image into gray level image. For spatial localization, we then use the DCT or Walsh transformation. Each image is resized to *N*N* size. DCT or Walsh Transform is applied on the image to generate a feature vector as shown in figure 1.

5.1 Algorithm for Image Classification

1. Feature vector of the query image is generated as shown in figure 1.
2. Feature vector of the query image is compared with the feature vectors of all the images in the database. Euclidean distance measure is used to check the closeness of the query image and the database images.
3. Euclidean distance values are sorted w.r.t. ascending order sequence to find first 100 closest matches with query image.
4. The closest matches with query image for all 10 categories are calculated.
5. A threshold value is set to determine to which category the query image belongs to.
6. Display the category of the query image.

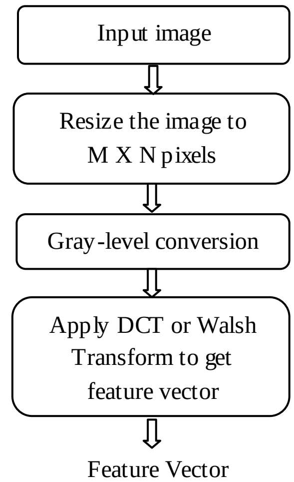

Figure 1: Flowchart for feature extraction

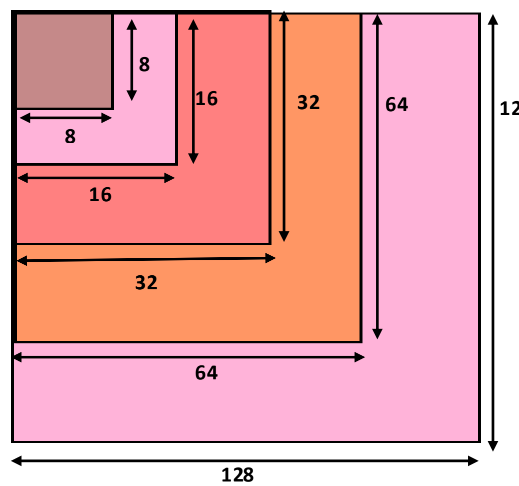

Figure 2: Selection of varying size portion from feature

6. RESULTS AND DISCUSSION

The implementation of the proposed algorithm is done in MATLAB 7.0 using a computer with Intel Core2 Duo Processor E4500 (2.20GHz) and 2 GB RAM. The DCT and Walsh Transform algorithm is tested on the image database of 1000 variable size images collected from Corel Collection [23] and Caltech-256 dataset [11]. These images are arranged in 10 semantic groups: Elephants, Horses, Roses, Coins, Mountains, Birds, Buses, Rainbows, Dinosaurs and Seashores. It includes 100 images from each semantic group. The images are in JPEG format.

The algorithm is executed with 10 query images selected randomly for each category in the image database. The results obtained are shown in the following tables. Here as per the set threshold, it is seen that algorithm classifies the images in the image database to the different categories such as Dinosaur, Roses, Horses, Elephant, Rainbow, Mountains, Coins, Seashores and Birds.

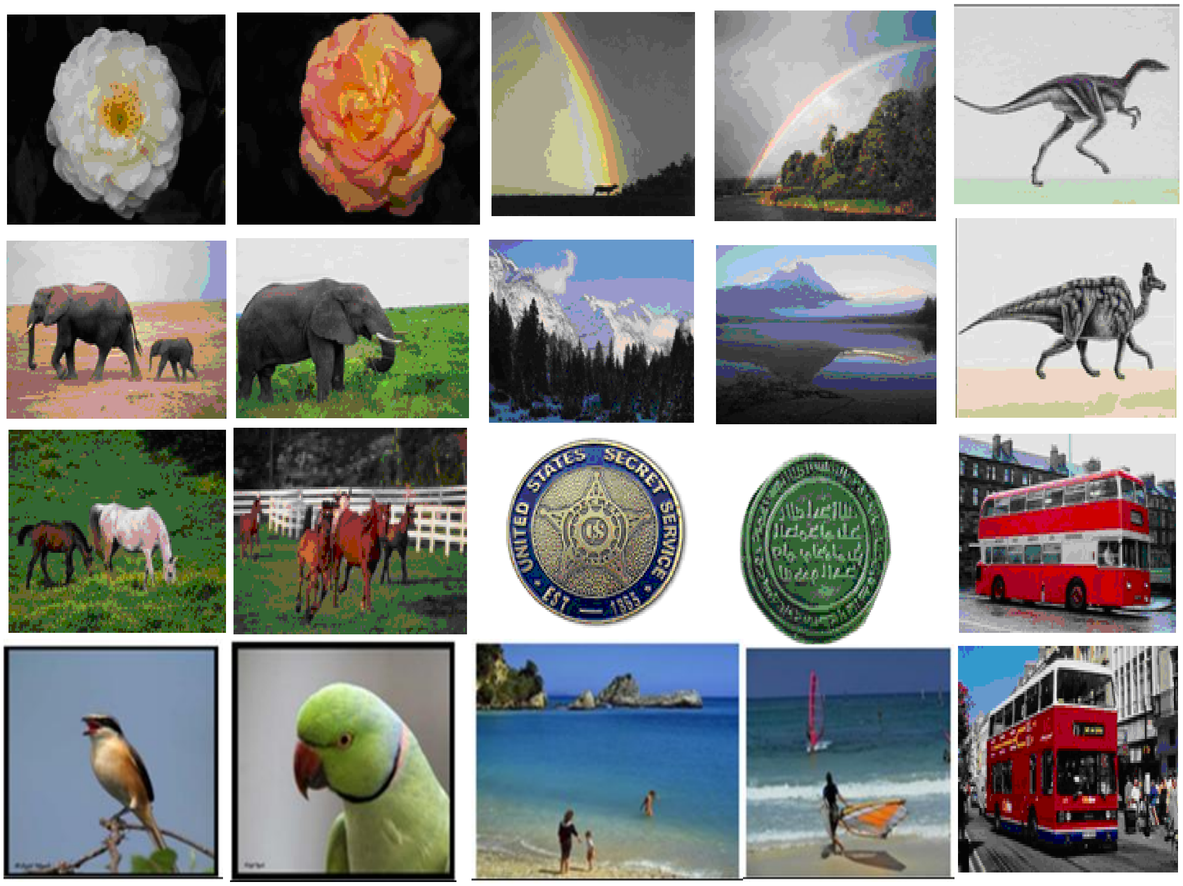

Figure 3: Sample Database Images

[Image database contains total 1000 images with 10 categories]

Table 2: Average Identification rate for 10 query images for all categories for different Feature Vector size using DCT

No	Feature Vector Size	Rainbow	Mountains	Horse	Rose	Elephant	Dinosaur	Bus	Sea	Coins	Bird
1	128 X 128	48.8	31.7	33.9	69.2	35.3	89.1	4.2	34.1	46.8	48.8
2	64 X 64	46.5	31.5	37.5	67.5	35.9	88.6	5.3	33.7	49.6	49.7
3	32 X 32	44.7	30.7	37.4	65.9	37.1	87.9	7.9	33.4	51.4	51.5
4	16 X 16	40.3	29.5	39.9	65.1	38.4	88.6	11.8	32.3	51.9	47.5
5	8 X 8	35.6	28.4	38.8	64.4	39.5	89.7	18.3	30.3	53.6	41.8

Table 2 show that the DCT algorithm gives 87.9% to 89.7% images from Dinosaur category, 64.4% to 69.2% images from Rose category, 35.3% to 39.5% images from Elephant category., 33.9% to 39.9% images from Horse category, 28.4% to 31.7% images from Mountains category, 30.3% to 34.1% images from Seashores category, 4.2% to 18.3% images from Bus category., 35.6% to 48.8% images from Rainbow category, 41.8% to 51.5% images from Birds category and 46.8% to 51.9% images from Coins category.

Table 3: Average Identification rate for 10 query images for all categories for different Feature Vector size using Walsh Transform

No	Feature Vector Size	Rainbow	Mountains	Horse	Rose	Elephant	Dinosaur	Bus	Sea	Coins	Bird
1	128 X 128	47.3	29.7	36.8	64.6	31.4	88	5.1	32.2	48.5	48.9
2	64 X 64	45.8	30.1	37.6	63.6	32	87.6	6.6	31.9	49.7	50
3	32 X 32	42.3	29.6	38.2	62.6	33	87.8	9	31	50.9	49.5
4	16 X 16	38.4	27.7	37.5	61.7	33.6	88.4	13.3	29.7	52.2	42.7
5	8 X 8	33.9	27.5	36.7	61.1	34.6	89.5	17.8	28.1	51.5	37.4

Table 3 show that the Walsh Transform algorithm gives 87.6% to 89.5% images from Dinosaur category, 61.1% to 64.6% images from Rose category, 35.3% to 39.5% images from Elephant category., 33.9% to 39.9% images from Horse category, 28.4% to 31.7% images from Mountains category, 30.3% to 34.1% images from Seashores category, 4.2% to 18.3% images from Bus category., 35.6% to 48.8% images from Rainbow category, 41.8% to 51.5% images from Birds category and 46.8% to 51.9% images from Coins category.

Table 4: Average accuracy of Image Classification

Results of Image Classification										
No	Feature Vector Size	Rainbow	Mountains	Horse	Rose	Elephant	Dinosaur	Sea	Coins	Bird
1	128 X 128	100	100	100	100	100	100	100	100	100
2	64 X 64	100	100	100	100	100	100	100	100	100
3	32 X 32	100	100	100	100	100	100	100	100	100
4	16 X 16	100	100	100	100	100	100	100	100	100
5	8 X 8	100	100	100	100	100	100	100	100	100

The threshold of 28% is set for DCT and 27% is set for Walsh Transform. Here as per the set threshold, it is seen that algorithm classifies the images in the image database to the different categories such as Dinosaur, Roses, Horses, Elephant, Rainbow, Mountains, Coins, Seashores and Birds. The average accuracy of image classifier is 90%; only the Bus images are misclassified. If "Bus" category is excluded, the classifier accuracy increases up to about 100% as shown in Table 4.

7. CONCLUSION

In recent years, thousands of images are generated everyday, which implies the necessity to classify, organize and access them by easy and faster way. The need for image classification is becoming increasingly important. Classifying the images into semantic categories is a challenging problem.

In this paper, a simple but effective algorithm of Image Classification which uses Discrete Cosine Transform (DCT) or Walsh Transform is presented. To evaluate this algorithm, a heterogeneous image database is used.

Experimental results on a database of 1000 images from 10 semantic groups in image classification were reported for all categories using DCT and Walsh Transform.

The complexity comparison of DCT and Walsh Transform shows that the complexity of DCT is more by 9.063 times than the complexity of Walsh Transform.

The average accuracy of image classifier is 90%; only the Bus images are misclassified. If "Bus" category is excluded, the classifier accuracy increases up to about 100%.

8. REFERENCES

[1] A.K.Jain and A.Vailaya, "Image retrieval using color and shape", Pattern recognition, vol.29, no.8, pp.1233-1244, 1996.

[2] B.S.Manjunath and W.Y.Ma, "Texture feature for browsing and retrieval of image data", IEEE PAMI, no. 18, vol. 8, pp. 837- 842, 1996.

[3] D. J. Le Gall, "The MPEG Video Compression Algorithm: A review," SPIE 1452 (1991) 444-457.

[4] Dr. H.B.Kekre, Sudeep D. Thepade, Akshay Maloo, "Image Retrieval using Fractional Coefficients of Transformed Image using DCT and Walsh Transform", International Journal of Engineering Science and Technology , Vol. 2(4), pp.362-371, 2010.

[5] Dr. H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, "Image Retrieval using Color-Texture Features from DCT on VQ Code vectors obtained by Kekre's Fast Codebook Generation", ICGST-GVIP Journal, Volume 9, Issue 5, September 2009.

[6] E. Nowak, F. Jurie, and B. Triggs. Sampling strategies for bag-of-features image classification. In ECCV,Part IV, LNCS 3954, pp. 490–503, 2006.

[7] Emma Newham, "The biometric report," SJB Services, 1995.

[8] F. Li and P. Perona, "A Bayesian Hierarchical Model for Learning Natural Scene Categories," Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, pp. 524-531, 2005.

[9] F.Mokhtarian and S.Abbasi, "Shape similarity retrieval under affine transforms", Pattern Recognition, 2002, vol. 35, pp. 31-41.

[10] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray. "Visual categorization with bags of key points", In. Proc. ECCV'04 workshop on Statistical Learning in Computer. Vision , pp. 59–74,2004

[11] G. Griffin, A. Holub, and P. Perona. Caltech 256 objects category dataset. Technical Report UCB/CSD-04-1366, California Institute of Technology, 2007.

[12] G. K. Wallace, "Overview of the JPEG still Image Compression standard," SPIE 1244 (1990) 220-233.

[13] Golam sorwar, Ajith abraham, "DCT based texture classification using soft computing approach", Malaysian Journal of Computer Science, vol. 17,2004

[14] H. B. Kekre, Dhirendra Mishra, "Digital Image Search & Retrieval using FFT Sectors" published in proceedings of National/Asia pacific conference on Information communication and technology(NCICT 10) 5TH & 6TH March 2010..

[15] H. B. Kekre, Sudeep Thepade, Akshay Maloo, "Performance Comparison of Image Retrieval Using Fractional Coefficients of Transformed Image Using DCT, Walsh, Haar and Kekre's Transform", CSC-International Journal of Image processing (IJIP), Vol.. 4, No.2, pp.:142-155, May 2010.

[16] H. B. Kekre, Tanuja Sarode, Shachi Natu, Prachi Natu, "Performance Comparison Of 2-D DCT On Full/Block Spectrogram And 1-D DCT On Row Mean Of Spectrogram For Speaker Identification", CSC International Journal of Biometrics and Bioinformatics (IJBB), Volume (4): Issue (3).

[17] H.B.Kekre, Dhirendra Mishra, "Content Based Image Retrieval using Weighted Hamming Distance Image hash Value" published in the proceedings of international conference on contours of computing technology pp. 305-309 (Thinkquest2010) 13th & 14th March 2010.

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

[19] H.B.Kekre, Sudeep D. Thepade, Archana Athawale, Anant Shah, Prathmesh Verlekar, Suraj Shirke,"Energy Compaction and Image Splitting for Image Retrieval using Kekre Transform over Row and Column Feature Vectors", International Journal of Computer Science and Network Security (IJCSNS),Volume:10, Number 1, January 2010, (ISSN: 1738-7906) Available at www.IJCSNS.org.

[20] H.B.Kekre, Sudeep D. Thepade, Archana Athawale, Anant Shah, Prathmesh Verlekar, Suraj Shirke, "Performance Evaluation of Image Retrieval using Energy Compaction and Image Tiling over DCT Row Mean and DCT Column Mean", Springer-International Conference on Contours of Computing Technology (Thinkquest-2010), Babasaheb Gawde Institute of Technology, Mumbai, 13-14 March 2010, The paper will be uploaded on online Springerlink.

[21] J. Li and J. Wang, "Automatic Linguistic Indexing of Pictures by a statistical modeling approach", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, pp. 1075-1088, 2003.

[22] J.R.Smith and C.S.Li, "Image classification and quering using composite region templates", Academic Press, Computer Vision and Understanding, 1999, vol.75, pp.165-174.

[23] J.Z.Wang, J.Li and G.Wiederhold, "SIMPLIcity: semantic sensitive integrated matching for picture libraries", IEEE Trans. on Pattern Analysis and Machine Intelligence, 2001, vol.23, no.9, pp.947-963.

[24] M. Szummer and R.W. Picard, "Indoor-Outdoor Image Classification," IEEE International Workshop on Content-based Access of Image and Video Databases, in conjunction with ICCV'98, pp. 42-51, 1998.

[25] M.J.Swain and D.H.Ballard, "Color indexing", International Journal of Computer Vision, vol.7, no.1, pp.11-32, 1991.

[26] M.R. Naphade and T.S. Huang, "Extracting semantics from audio-visual content: the final frontier in multimedia retrieval", IEEE Trans. on Neural Networks, vol. 13, no. 4, pp. 793–810, July 2002.

[27] M.Seetha, I.V.MuraliKrishna, B.L. Deekshatulu, "Comparison of Advanced Techniques of Image Classification", Map World Forum Hyderabad, India.

[28] O. Chapelle, P. Haffner, and V. Vapnik, "Support vector machines for histogram-based image classification", IEEE Transactions on Neural Networks, vol. 10, pp. 1055-1064, 1999.

[29] O. Maron and A.L. Ratan, "Multiple-Instance Learning for Natural Scene Classification", Proceedings of the Fifteenth International Conference on Machine Learning, pp. 341-349, 1998.

[30] Rostom Kachouri, Khalifa Djemal and Hichem Maaref, Dorra Sellami Masmoudi and Nabil Derbel ,"Content description and classification for Image recognition system", Information and Communication Technologies: From Theory to Applications, ICTTA ,3rd International Conference, April 2008.

[31] Sang-Mi Lee, Hee_Jung Bae, and Sung-Hwan Jung, "Efficient Content-Based Image Retrieval Methods Using Color and Texture", *ETRI Journal* 20 (1998) 272-283.

[32] Y. Chen and J.Z. Wang, "Image Categorization by Learning and Reasoning with Regions", Journal of Machine Learning Research, vol. 5, pp. 913-939, 2004.

[33] Zhu Xiangbin ," Cartoon Image Classification Based on Wavelet Transform", Asia-Pacific Conference on Information Processing, pp. 80-83,July 2009.