Self-Mutation of Hybrid Wavelet Transform with Cosine-Kekre, Cosine-Haar, Cosine-Walsh, Walsh-Cosine, Haar-Cosine and Kekre-Cosine for Content Based Video Retrieval

Nalini Yadav
Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Savitribai Phule Pune University
Pune, India
yadav.nalini@gmail.com

Sudeep D. Thepade
Ph.D., Professor Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Savitribai Phule Pune University
Pune, India
sudeepthepade@gmail.com

Abstract – Efficient access of multimedia data is a big requirement of current decade. With advancement of storage and network technologies, there is huge generation of multimedia data. To store and access this multimedia data efficiently and accurately is a challenge. Videos are part of multimedia data. Important activities in video retrieval are efficient storage and accurate retrieval of relevant videos. This paper proposes a novel Content based Video Retrieval method using energy contents of the video. Proposed energy contents extraction method includes self-mutation of Hybrid Wavelet Transform. Further in proposed approach the efficient method to reduce the feature vector size using partial energies of the video content is also enlightened. Paper also list downs the performance comparison of four similarity measures alias Euclidean Distance, City Block Metric, Sorensen Distance, Kulczynski Distance. This paper aims at using a novel method for generation of orthogonal transform called Self Mutation of Hybrid Wavelet Transform and judging the efficiency of same for Content Based Video Retrieval with energy compaction using partial energies.

Keywords— Self Mutation; Hybrid Wavelet Transform; Energy Compaction; Similarity measures; Partial Energy; Cosine Transform; Haar Transform; Walsh Transform;

I. INTRODUCTION

Content based video retrieval is needed for efficient indexing and retrieval of relevant videos after storage in large databases. Due to swift development of image processing techniques and digital equipments, a huge amount of videos are getting generated. To access these videos efficiently and accurately is the biggest challenge for the research. Due to uncertain length and format of the video, it is becoming more difficult to index and access them. Thus the well-structured analysis of the video is required. The efficient index and access of the video decides the accuracy of the search engines.

In earlier days, there were a text based video retrieval systems. But the performance of annotation or textual based video retrieval system depends heavily upon metadata (in form of text) attached with video which are subjective and also the process is time consuming and manual. Problems with traditional method of indexing and retrieval of video generated a need of new generation retrieval system based on automatically derived features of videos. These features can be Color, Motion or Speech which are actual content of the video. Content or features of the video describes the video precisely than the annotations attached with it. Content based retrieval system can be best described as a system which searches the video based on the content rather than the metadata of the video. Input to the Content based video retrieval system is a query video and output is relevant retrieved videos from the database; which are matched with query video.

The goal of every Content based video retrieval system is to index the videos in dataset with efficiency in both aspects of storage, accuracy and retrieval in faster manner. Efficient video retrieval system provides the way to access, update and retrieve the videos in flexible and less complex manner. Video is moving sequence of images. Frame image is base of every video and sequence of those images is matter of interest. Videos are described in form of key frames. Key frames are image frames having distinguishable content than other subsequent image frames.

II. LITERATURE SURVEY

Video consist of multiple images. Image can be described with contents like color, edge, shape and intensity. Video is rich in contents hence high storage space required to store video. The aim of Content Based Video Retrieval system is to minimize the signature size required for the video and index it with efficiency. To minimize the space required to store the video, it is necessary to consider only discriminating features of the video. Orthogonal transforms are widely used to extract the discriminating features of the image frame [4]. Multiple orthogonal transforms are available out of them Haar and Walsh with partial coefficients have shown better performance for Content Based Video Retrieval [2, 3]. The Wavelet transforms generated from orthogonal transforms give higher energy compaction which results in reduction in feature vector size in image retrieval [1].

As compared to orthogonal transforms, respective wavelet transforms are proven better in many applications [5]. Hybrid Wavelet Transform is generated from two constituent orthogonal transforms. In the application of video retrieval the hybrid wavelet transform proved better compared to their constituent transforms [6, 9]. Moreover from combinations of multiple orthogonal transforms, hybridization of Cosine and Walsh showed better performance in Content Based Video Retrieval [6] for partial energies. As compared to orthogonal transforms, respective wavelet transforms and hybrid wavelet transforms are proven better in many applications [7, 8].

A. *Generation of Hybrid Wavelet Transform*

The hybrid wavelet transform $W_{ab\times ab}$ will be generated as follows: Consider two matrices I of size a×a & J of size b×b as given in figure 1.

I_{11}	I_{12}	…	I_{1a}
I_{21}	I_{22}	…	I_{2a}
…	…	…	…
I_{a1}	I_{a2}	…	I_{aa}

J_{11}	J_{12}	…	J_{1b}
J_{21}	J_{22}	…	J_{2b}
…	…	…	…
J_{b1}	J_{b2}	…	J_{bb}

Fig.1. Two Orthogonal Transforms $I_{a\times a}$ and $J_{b\times b}$

$I_{11}=\begin{pmatrix}J_{11}\\J_{21}\\\vdots\\J_{b1}\end{pmatrix}$	$I_{12}=\begin{pmatrix}J_{11}\\J_{21}\\\vdots\\J_{b1}\end{pmatrix}$	- - -	$I_{1a}=\begin{pmatrix}J_{11}\\J_{21}\\\vdots\\J_{b1}\end{pmatrix}$	$I_{21}=\begin{pmatrix}J_{11}\\J_{21}\\\vdots\\J_{b1}\end{pmatrix}$	$I_{22}=\begin{pmatrix}J_{11}\\J_{21}\\\vdots\\J_{b1}\end{pmatrix}$	- - -
I_{21}	I_{22}	-	I_{2a}	0	0	-
0	0	-	0	I_{21}	I_{22}	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	0	0	0	0	0	0
I_{31}	I_{32}		I_{3a}	0	0	0
0	0	0	0	I_{31}	I_{32}	-
⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	0	0	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
I_{a1}	I_{a1}		I_{aa}	0	0	-
0	0	0	0	I_{a1}	I_{a2}	-
⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	0	0	0	0	0	0

Fig.2. Generation of Hybrid Wavelet Transform from two orthogonal transforms $I_{a\times a}$ and $J_{b\times b}$

In the above matrix the Hybrid Wavelet Transform (W_{abxab}) is generated from mixture of matrix I & matrix J as 1st a number of rows generated from product of each element from 1st row of matrix I with every column of matrix J. Secondly remaining rows will be filled with circular shift with every row of matrix I [5].

B. *Partial Energy*

Orthogonal Transforms have the property that when they are applied on image then they discriminate the high energy and low energy regions from each other. High energy contents of the image contain the most distinguishing features of the image. Thus high energy contents can only be considered as features representing the image which can distinguish one image from another image. When all the coefficients of the transformed image are considered then the feature vector size becomes very large, but when few of the low frequency coefficients are considered to generate the feature vector then the size of feature vector reduces drastically and this leads to lesser number of computations in comparing feature vectors in retrieval process. Extracting partial energy components for feature vector has proved efficient for image retrieval [10].

Extraction of partial energies from the all energy coefficients is a 3 step process [11]. These steps are summarized in figure 3.

a. Generation of average energy matrix

b. Building a summed energy matrix

c. Extracting Partial energy coefficient table

Steps are summarized in a Figure 3.

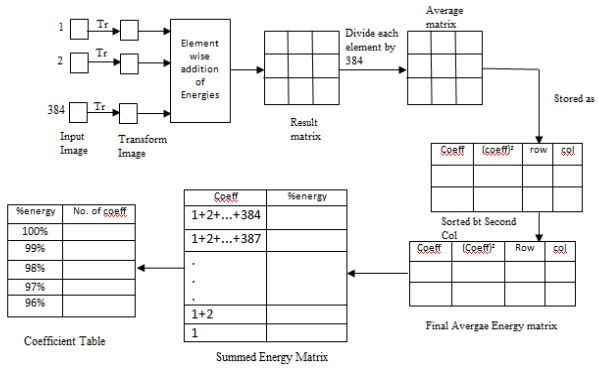

Fig.3. Energy Compaction using partial energy

III. PROPOSED CONTENT BASED VIDEO RETRIEVAL TECHNIQUE

Proposed Content Based Video Retrieval system emphasizes upon energy contents of the videos. A novel method of generation of orthogonal transform is induced [13] called as Self Mutation of Hybrid Wavelet Transform (SMHWT). In this method a hybrid wavelet transform evolves itself for mutation, gives rise to a new orthogonal transforms. This paper aims at evaluating the efficiency of the self mutated hybrid wavelet transform with Content Based Video Retrieval. Energy of video is derived using SMHWT method. These energy contents are considered partially in deriving the features of the particular video.

I_{11}	I_{12}	…	I_{1a}
I_{21}	I_{22}	…	I_{2a}
…	…	…	…
I_{a1}	I_{a2}	…	I_{aa}

Fig.4. Base hybrid Wavelet Transform for generation of Self Mutated Hybrid Wavelet Transform

Figure 4 shows the base hybrid wavelet transform which is generated by using the steps given in figure 2. This base hybrid wavelet transform is used for creating the self mutated hybrid wavelet transform.

Figure 5, details about the self mutation of hybrid wavelet transform.

$I_{11}=\begin{pmatrix}I_{11}\\I_{21}\\\vdots\\I_{b1}\end{pmatrix}$	$I_{12}=\begin{pmatrix}I_{11}\\I_{21}\\\vdots\\I_{b1}\end{pmatrix}$	-	$I_{1a}=\begin{pmatrix}I_{11}\\I_{21}\\\vdots\\I_{b1}\end{pmatrix}$	$I_{21}=\begin{pmatrix}I_{11}\\I_{21}\\\vdots\\I_{b1}\end{pmatrix}$	$I_{22}=\begin{pmatrix}I_{11}\\I_{21}\\\vdots\\I_{b1}\end{pmatrix}$	-
I_{21}	I_{22}	-	I_{2a}	0	0	-
0	0	-	0	I_{21}	I_{22}	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	0	0	0	0	0	0
I_{31}	I_{32}		I_{3a}	0	0	0
0	0	0	0	I_{31}	I_{32}	-
⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	0	0	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
I_{a1}	I_{a1}		I_{aa}	0	0	-
0	0	0	0	I_{a1}	I_{a2}	-
⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	0	0	0	0	0	0

Fig.5. Novel method of Self Mutation of Hybrid Wavelet Transform (SMHWT)

In the self mutation of hybrid wavelet transform (SMHWT), a hybrid wavelet transform gets mutated with itself to evolve a new orthogonal transform. A resultant self mutated hybrid wavelet transform is transformed on the videos to extract the energy contents of the video. These energy contents are used partially like 99%, 97%, 95%, 90%, 85% and 80% for feature vector generation. Accuracy of the video retrieval system is evaluated for proposed partial energy contents of the video.

The proposed technique also explores the different similarity measures for getting higher accuracy of retrieval. In proposed technique query feature vector is compared with feature vectors of database videos with four different similarity measures alias Euclidean Distance, City Block Metric, Sorensen Distance and Kulczynski Distance to evaluate suitability of similarity measure for proposed Content Based Video Retrieval [12].

The proposed research paper aims at increasing the performance of content based retrieval system by reducing the feature vector size with partial energy compaction and increase in accuracy by using efficient similarity measure for matching the query feature vector with database of feature vectors.

To assess the retrieval effectiveness, proposed system has used the average accuracy as statistical comparison parameter. Higher accuracy indicates more accurate method for feature extraction. Accuracy is fraction of retrieved videos that are relevant as given in equation 1.

$$Accuracy = \frac{\{Number\ of\ relevant\ videos\ \ retrived\}}{\{Total\ number\ of\ videos\ retrieved} \quad (1)$$

IV. Experimentation Environment

Here the platform used for experimentation is MATLAB with processor CORE i3.

The experimentation test bed has 500 Videos across 10 categories, each category has 50 videos. Fig. 6 shows the sample from collection of videos considered in data set.

Fig.6. Samples from the Categories of Video Dataset

V. Results and Discussion

The aim of experimentation of the proposed method is to find out the impact of self mutation of Hybrid Wavelet Transform on Content based Video Retrieval using energy compaction.

The purpose of research work is to generate the self mutated Hybrid Wavelet Transform and taking out the partial energies of the transformed content of the video. Different Hybrid Wavelet Transform viz. Cosine-Haar, Cosine-Kekre, Cosine-Walsh, Haar-Cosine, Kekre-Cosine and Walsh-Cosine are self mutated in proposed experiments to generate new orthogonal Transform. Energy compaction of transformed visual contents is achieved with partial coefficients viz. 99%, 97%, 95%, 90%, 85% and 80% of total energy content. The similarity measures used for comparison of feature vector are Euclidean Distance, City Block Metric, Sorensen Distance and Kulczynski Distance [12].

The experimentation is conducted using video dataset of 500 videos spread across 10 video categories. On the considered database total 500 queries are fired. The average accuracy values for four similarity measures - Euclidean Distance, City Block Metric, Sorensen Distance and Kulczynski Distance for six orthogonal self mutated hybrid wavelet transforms - Cosine-Haar, Cosine-Kekre, Cosine-Walsh, Haar-Cosine, Kekre-Cosine and Walsh-Cosine with 99%, 97%, 95%, 90%, 85% and 80% partial energy coefficients is shown in figure 7.

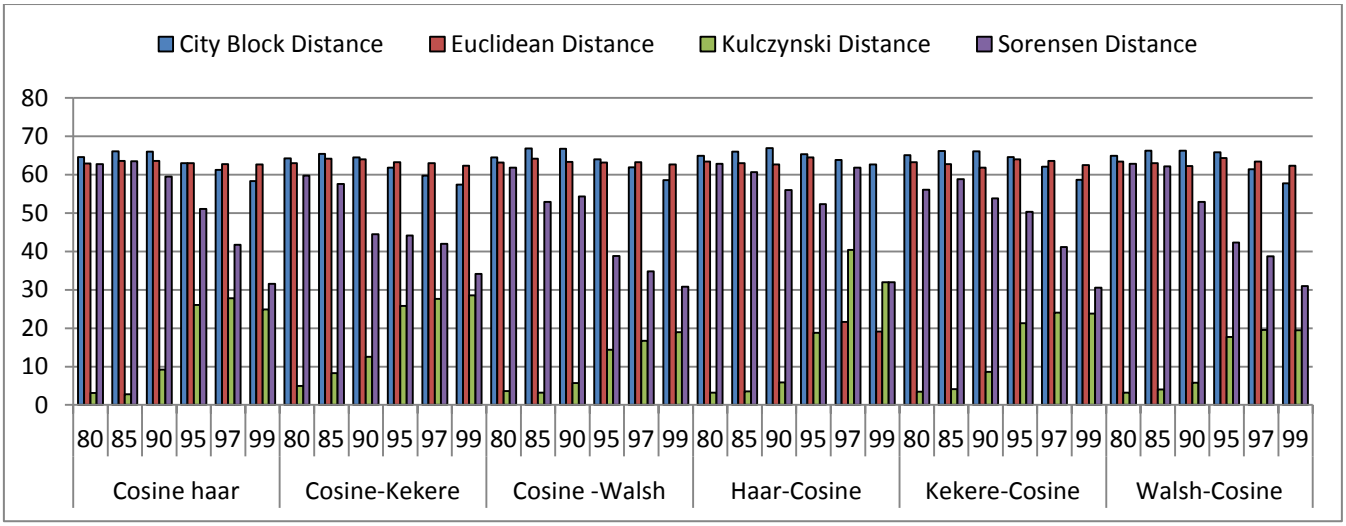

Fig.7. Average accuracy for Cosine-Haar, Cosine-Kekre, Cosine-Walsh, Haar-Cosine, Kekre-Cosine and Walsh-Cosine Self Mutated Hybrid Wavelet transform for four similarity measures and 100%, 99%, 97%, 95%, 90%, 85% and 80% partial energy coefficients

It is observed from the figure 7, that out of four similarity measures City Block Metric is performing best with all six self mutated hybrid wavelet transform. The goal of increasing the accuracy of retrieval is achieved with use of best performing similarity measure viz. City Block Metric which gives the highest accuracy.

TABLE I AVERAGE ACCURACY USING CITY BLOCK METRIC OF SELF MUTATED HYBRID WAVELET TRANSFORMS WITH PARTIAL ENERGY

% Parial Energy	Cosine-Haar	Cosine-Kekere	Cosine - Walsh	Haar-Cosine	Kekere-Cosine	Walsh-Cosine
80	64.608	64.268	64.524	64.936	65.06	64.936
85	66.052	65.388	66.816	66.012	66.184	66.232
90	66.012	64.536	66.736	66.956	66.084	66.244
95	62.988	61.868	64	65.36	64.548	65.828
97	61.244	59.792	61.904	63.832	62.052	61.408
99	58.328	57.412	58.6	62.692	58.644	57.732

Table I and figure 8 enlightens the best energy compaction percentage of Self mutated Hybrid Wavelet Transform with partial energy contents with highest accuracy. Out of 99%, 97%, 95%, 90%, 85% and 80%, 85% is performing best for all six Self Mutated Hybrid Wavelet Transforms with City Block Metric as similarity measure. From the proposed experimentation it is observed that 85% partial energy coefficient gives highest accuracy indicating better Retrieval.

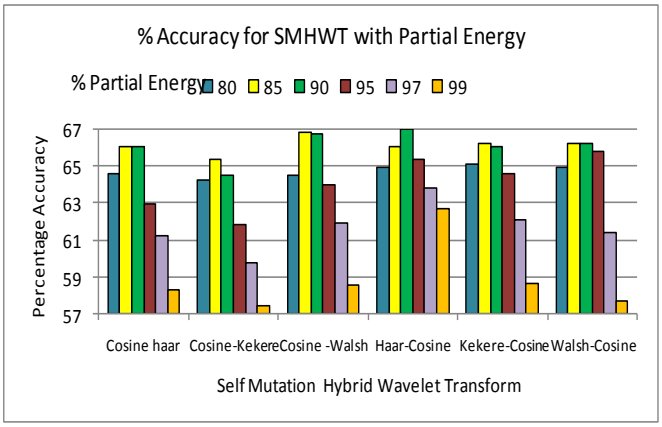

Fig.8. Summary of average accuracy of experimented Self Mutated orthogonal Hybrid Wavelet transforms for 100%, 99%, 97%, 95%, 90%, 85% and 80% of partial energy coefficients with City Block Metric similarity measure

Figure 9 compares all six self mutated hybrid wavelet transforms for 85% of energy compaction. It is observed that at 85% of energy coefficients Cosine-Walsh self mutated hybrid wavelet transform gives 66.82% of highest accuracy.

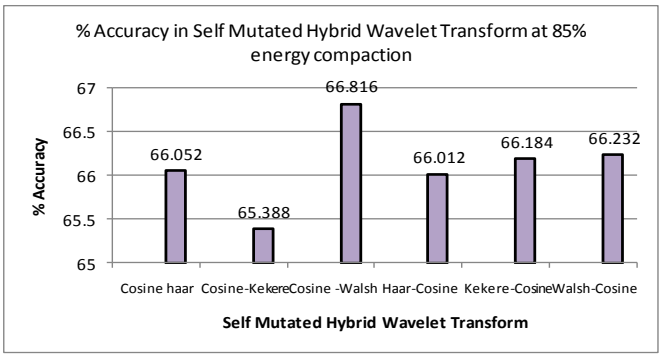

Fig.9. Summary of average accuracy of experimented six Self Mutated

orthogonal Hybrid Wavelet transforms for 85% of partial energy coefficients with City Block Metric similarity measure

Table II summarizes the performance given by five orthogonal transforms for different energy percentages considered in proposed transformed content based video retrieval.

TABLE II SUMMARY OF AVERAGE ACCURACY AND REDUCTION IN FEATURE VECTOR SIZE FOR DIFFERENT PERCENTAGE OF PARTIAL ENERGY COEFFICIENTS

Self Mutated Hybrid Wavelet Transform	Energy considered to form feature vector	Number of coefficients considered to form Feature vector	Percentage Reduction in size of Feature vector (In %)	Accuracy (In %)
Cosine-Haar	85	24	99.963	66.052
Cosine-Walsh	85	20	99.969	66.816
Cosine-Kekare	85	25	99.962	65.388
Haar-Cosine	85	19	99.971	66.012
Walsh-Cosine	85	19	99.971	66.232
Kekare-Cosine	85	23	99.965	66.184

From table II, it is can be observed that for the best performing self mutated Cosine-Walsh Hybrid wavelet transform at 85% of partial energy only 20 coefficients are required thus there is huge reduction of feature vector size. Hence the goal of reduction in size complexity is achieved through reduction in feature vector size with 85% of partial energy coefficients and highest accuracy of retrieval. When there is reduction in feature vector size then there is less number of comparisons required for matching the query video with database videos which indirectly improves the speed of retrieval. Hence the third aim of improving the speed of video retrieval is also achieved with partial energy coefficients.

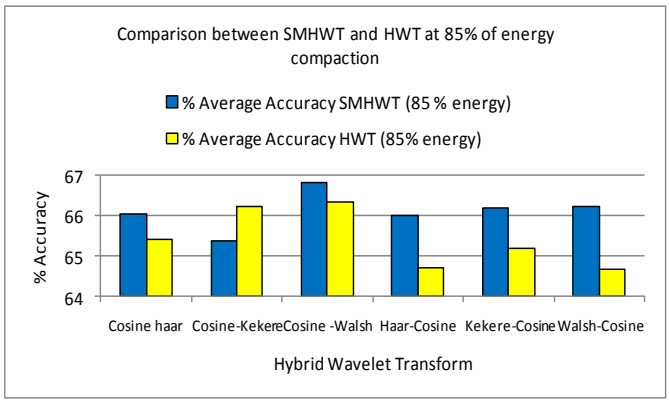

Fig.10. Comparison between SMHWT and HWT at 85% of energy compaction

Figure 10 shows the performance comparison for Hybrid Wavelet Transform and Self mutated hybrid wavelet transform for 85% of partial energy coefficients. The self mutated hybrid wavelet transform performs best in all transforms except Cosine-Kekre Self mutated hybrid wavelet transform. Figure 11, details about the comparison of HWT and SMHWT. It clearly compares the HWT and proposed SMHWT technique with accuracy, number of coefficients required and percentage reduction in feature vector size. It is concluded from figure 11 that in all three aspects of efficiency proposed SMHWT method supersedes the existing HWT method.

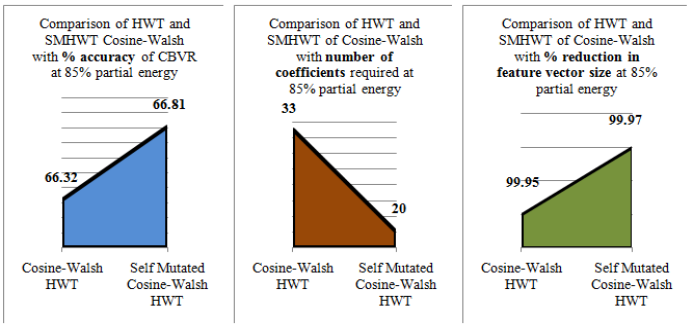

Fig.11. Comparison between SMHWT And HWT

VI. CONCLUSION

Goals of Content Based Video Retrieval system are to increase the efficiency of retrieval system with reduction in feature vector size, increase in retrieval speed and improvising accuracy. The proposed research is focused around these goals. Proposed research evaluated impact of energy compaction with 99%, 97%, 95%, 90%, 85% and 90% partial energy coefficients on retrieval accuracy and reduction in feature vector size. The novel method of generating an orthogonal transform named self mutated hybrid wavelet transform (SMHWT) is proposed in this paper. This paper aims at evaluating the efficiency of the self mutated hybrid wavelet transform with Content Based Video Retrieval

Self Mutated Hybrid Wavelet Transforms viz. Cosine-Haar, Haar-Cosine, Cosine-Walsh, Walsh-Cosine, Cosine-Kekare and Kekare-Cosine are used for Content Based Video Retrieval with partial energy coefficients as 99%, 97%, 95%, 90% 85% and 80% and four similarity measures viz. Euclidean Distance, City Block Metric, Sorensen Distance and Kulczynski Distance. After experimentation with dataset of 500 videos, it is observed that Self Mutated Cosine-Walsh Hybrid Wavelet Transform is performing best with 85% partial energy coefficients i.e. 99.38% in reduction in feature vector size. At 85% of partial energy coefficients Cosine-Walsh gives 66.82% average accuracy with city block metric similarity measure. With comparison to existing HWT method, proposed SMHWT method performs better. Reduction in feature vector size improvises speed thus results into efficient retrieval system.

References

[1] Dr. H. B. Kekre, Dr. Tanuja K. Sarode, Dr. SudeepD.Thepade, Ms.Sonal Shroff, "Instigation of OrthogonalWavelet Transforms using Walsh, Cosine, Hartley, KekreTransforms and their use in ImageCompression", International Journal of Computer Science and Information Security (IJCSIS), Vol 9, No 6, pp.125-133, June2011

[2] Nalini Yadav, Dr. Sudeep Thepade, "Comprehensive Performance Comparison of Fourier, Walsh, Haar, Sine and Cosine Transforms for Video Retrieval with Partial Coefficients of Transformed Video", International Journal of Computer Applications (IJCA), Volume 120, No 19, June 2015.

[3] Dr. H.B. Kekre, Dr. Sudeep D. Thepade & Saurabh Gupta, "Content Based Video Retrieval in Transformed Domain using Fractional Coefficients", International Journal of Image Processing (IJIP), Vol 7, Issue 3, pp 238,274, 2013.

[4] H.B.Kekre, Sudeep D. Thepade, "Improving the Performance of Image Retrieval using PartialCoefficients of Transformed Image", International Journal of Information Retrieval (IJIR), Vol 2, Issue 1, pp. 72-79, 2009.

[5] Dr. H. B. Kekre, Dr. Tanuja K. Sarode, Dr. SudeepD.Thepade, Ms.Sonal Shroff, "Instigation of OrthogonalWavelet Transforms using Walsh, Cosine, Hartley, KekreTransforms and their use in ImageCompression", International Journal of Computer Science and Information Security (IJCSIS), Vol 9, No 6, pp.125-133, June2011

[6] Dr. Sudeep Thepade, Nalini Yadav, "Partial Energy of Hybrid Wavelet Transformed Videos for Content Based Video Retrieval with various Similarity measures using Cosine, Haar and Walsh Transforms", Global Conference on Communication Technology (GCCT), IEEE Conference, April 2015.

[7] Dr. Sudeep Thepade, Pushpa R. Mandal, "Performance Comparison of Cosine, Walsh, Haar, Kekre and Hartley Transforms for Iris Recognition using Fractional Energies of the Transformed Iris Images", International Journal of Emerging Technologies in Computational and Applied Sciences, 8(1), March-May, 2014.

[8] Dr. H. B. Kekre, Dr. Sudeep D. Thepade, Varun K. Banura, Ankit Khandelwal, "Augmentation of Image Retrieval using Fractional Coefficients of Hybrid Wavelet Transformed Images with Seven Image Transforms", IJCST Vol. 3, Iss ue 1, Jan. - March 2012.

[9] Dr. Sudeep D. Thepade, Nalini Yadav, "Novel Efficient Content Based Video Retrieval Method using Cosine-Haar Hybrid Wavelet Transform with Energy Compaction", IEEE Conference, ICCUBA, Pune, India, Feb 2015.

[10] 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).

[11] Dr. Sudeep Thepade, Pushpa R. Mandal, "Energy compaction based novel Iris recognition techniques using partial energies of transformed iris images with Cosine, Walsh, Haar, Kekre, Hartley Transforms and their Wavelet Transforms.", India Conference (INDICON), 2014 Annual IEEE Conference, Pune, India, pp 1-6, 11-13 Dec. 2014.

[12] Sudeep D. Thepade, Nalini Yadav, "Assessment of Similarity Measurement Criteria in Thepade's Sorted Ternary Block Truncation Coding (TSTBTC) for Content Based Video Retrieval", 2015 International Conference on Communication, Information & Computing Technology (ICCICT), Jan. 16-17, Mumbai, India.

[13] Dr. Sudeep D. Thepade, Jaya H. Dewan, "Proposing Self Mutation of Hybrid Wavelet Transform with Cosine-Kekre, Cosine-Sine and Cosine-Walsh for image compression", IEEE International Conference on Futuristic Trends in Computational Analysis and Knowledge Management, 26-17 Feb. 2015, Amity University, Noida Delhi.