Hadamard based Video Key Frame Extraction using Thepade's Transform Error Vector Rotation with Assorted Similarity Measures

Pritam H. Patil
Dept. of Computer Engg.
PCCOE, SPPU
Pune, India

Sudeep D. Thepade
Ph.D. Computer Engineering
Dept. of Computer Engg.
PCCOE, SPPU
Pune, India

Babita Sonare
Asst. Professor IT Dept.
Dept. of Computer Engg.
PCCOE, SPPU
Pune, India

ABSTRACT
In Video summarization is a method to reduce redundancy and generate succinct representation of the video data. In video summarization process, several frames containing similar information need to get processed, this leads to slower processing speed and higher complexity, consuming. More time Video summarization using key frames can ease the speed up of video processing. One of the mechanisms to generate video summaries is to extract key frames which represent the most important content of the video by identifying neard duplicate frames in video. In this paper, novel key frames extraction method is proposed with Thepade's Walsh Hademard Error Vector Rotation (THdEVR) with ten different codebook sizes and and assorted similarity measures. Experimentation done with help of the test bed of videos has shown that higher codebook sizes give better completeness in key frame extraction for video summarization. Experimental results are discussed for video content summarization with five assorted similarity measures like Euclidean Distance, Canberra Distance, Square-Chord Distance, Mean Square Error, Sorensen Distance with proposed THadVR.

Keywords
Key frame, video summarization, vector quantization, hademard

1. INTRODUCTION
Now a day's increase in video data have imposed new challenges in managing such an enormous information. Due to which Video summarization is attracting more attention as the applications like information browsing and video retrieval [1], [2] [13] need summary of the videos. A video summary is a small part of the complete video sequence and aims to give a meaningful visual outline of a entire video with fewer distinct video frames.

The goal of a video summarization is to provide a crisp video visualization so that the user can understand overall content of video [3] and remove redundant information from the video. Generally, video summarization include the steps like video segmentation, feature extraction, after that redundancy detection based on features and finally video summarization with the non redundant features (key frame) [4].

In [5] and [6], the video summary is represented in two ways, first is a static video summary (e.g. storyboard) and a dynamic video skimming (e.g. movies trailer). In dynamic video skimming most relevant small video skims (shots) are selected or generating a video summary. But, in static video summary are identified most distinct nearby frames (keyframes) of a video sequence [7] are identified. A static video summary is more appropriate for browsing, retrieval and indexing. Video is made up of still frames among that , Fig. 1. gives detailed structure of video.

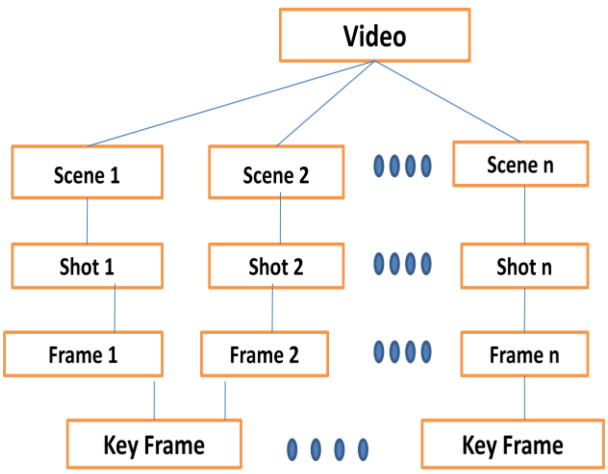

Fig 1: Video Structure in terms of scene, shot. frame and keyframe

This paper proposes a keyframe extraction using a Thepade's Transform Error Vector Rotation using Hademard Transform [10] method of vector quantization with five similarity measures like Euclidean distance, Mean Square Error, Canberra Distance, Square Chord distance Sorensen Distance with ten different codebook sizes resulting into 50 variation of the proposed technique are compared using completeness as performance measure.

The rest of the paper is organized as follows: in section II literature survey on VQ is given. Section III provides the details of our proposed framework. In section IV the experimentation environment is elaborated. Section V gives Results and observations of experimentation. Section VI concludes the paper

2. VECTOR QUANTIZATION
Vector quantization is defined as mapping function that maps k dimensional vector space to a finite set of codebook CB={C_1, C_2, C_3, C_4, ..., C_N}. The codebook (CB) consisting N numbers of codevector and every codevector represent as C_i={c_{i1}, c_{i2}, c_{i3}, c_{i4}, ..., c_{ik}}is of dimension k. Vector

quantization is using three steps: Codebook design, Encoding and Decoding. For key frame extraction method only first two steps of vector quantization are used. There are various codebook generation algorithms available [8]. The method most commonly used to generate codebooks is the Linde-Buzo-Gray (LBG) algorithm [9].

VQ is suitable for lossy data compression. Vector quantization is used in many application like face detection, iris recognition biometric identification image and retrieval etc.

2.1 Thepade's Hademard Transform Error Vector Rotation(TH_dEVR) Codebook Generation Algorithm

Please The In LBG, KPE algorithm some drawbacks like cluster elongation and constant error and proportionate error vector addition are observed. To overcome this drawback Thepare et. al. [10] have desired codebook generated in spatial domain by clustering algorithms.

In the Hadamard transformation matrix the positive and the negative values are replaced respectively with 1's and -1's to obtain the Thepade's Hadamard error vector matrix to be used in proposed key frame extraction technique using Thepade's Hadamard Transform Error Vector Rotation (THdEVR) codebook generation algorithm. In this algorithm each row of Hadamard error matrix e_i for i^{th} row will be an error vector to be added and subtracted from the centroid of cluster for dividing into two cluster. Fig. 2 shows Hademard sequence consist for size 8×8.

1	1	1	1	1	1	1	1
1	-1	1	-1	1	1	1	-1
1	-1	-1	1	-1	1	1	1
1	1	-1	-1	1	-1	1	1
1	-1	1	-1	-1	1	-1	1
1	-1	-1	1	-1	-1	1	-1
1	-1	-1	-1	1	-1	-1	1
1	1	-1	-1	-1	1	-1	-1

Fig 2: Hademard Error Vector of size 8×8

Steps THdEVR are explained as follow.

Step 1: Separate R, G, B components of the image are split into non overlapping blocks and convert each block to vector thus forming a training vector set Initialize i=1.

Step 2: Take a column wise mean of training vector and compute a centroid (codevector).

Step 3: Use Hademard transform matrix and form a Thepade's Hademard error vector e.

Step 4: Add and subtract error 'i' in codevector and form a two codevectors v1 and v2.

Step 5: Compute Euclidean distance between all the training vectors belonging to this cluster and the vectors v_1 and v_2 and split the cluster into two.

Step 6: Compute a centroid (codevector) for respective clusters obtained in the above Step 7.

Step 7: Increment 'i' by one and repeat Step 2 to Step 6 for each codevector and for desired codebook size is obtained

step 8: Stop.

3. PROPOED KEY FRAME EXTRACTION USING TH_DEVR WITH HADEMARD TRANSFORM

3.1 Key Frame Extraction

Please A video consisting of 'N' number of frames is taken as an input with consecutive frames be 'Vn' and 'Vn+1' having feature vectors respectively for n^{th} and $n+1^{th}$ frame as 'FVn' and 'FVn+1'. The diff(n) be the difference between two consecutive video frame feature vectors. The feature vectors taken m proposed techniques are codebooks generated using THdEVR with Hadamard transform. Here five similarity measures are used to calculate the diff(n) in proposed method as mean square error(MSE), Euclidean Distance, Canberra Distance, Squared-Chord Distance, and Sorensen Distance. Calculating diff(n) Mean(M) and standard deviation(SD) of all the video frames are calculated based on the five assorted similarity measures considered. The Threshold(T) value is computed by adding mean and multiplying constant 'a' with standard deviation. By comparing the threshold with the difference of consecutive video frames the key frames are computed as given in equation (1), (2) and (3).

$$Mean(M) = \frac{\sum_{n=1}^{N} diff(n)}{N-1} \quad (1)$$

$$SD = \sqrt{\frac{\sum_{n=1}^{N} (diff(n) - M)^2}{N-1}} \quad (2)$$

$$T = M + (a \times SD) \quad (3)$$

Here 'a' as a constant and after this calculate a key frame using If (diff (n) > T) output of n^{th} frame set as a key frame [9].

3.2 Key frame Extraction Using TH_dEVR with Hademard Transform and Assorted Similarity Measures

Please Video V, consisting of 'N' frames taken as an input. Consecutive frames are read as Vc and Vc+1 and then THdEVR with Hadamard applied on two consecutive frames is applied to calculate feature vectors. Then, following similarity measures are used for calculate difference:

1) *Mean Square Error (MSE)-*

The Mean Square Error can be represented mathematically by the equation (4).

$$MSE = \frac{1}{N} \sum_{c=1}^{N} [V_c - V_{c+1}] \quad (4)$$

2) Euclidean Distance (ED)-

The Euclidean distance between two frame defined as an equation (5).

$$ED = \sqrt{\sum_{c=1}^{N} |V_c - V_{c+1}|^2} \quad (5)$$

3) Square-Chord Distance (SC)-

The equation (6) gives a representation of SC.

$$SC = \sum_{C=1}^{N} \left[\left[\sqrt{V_c} - \sqrt{V_{c+1}} \right] \right]^2 \qquad (6)$$

4) Canberra Distance (CAB)-

Representation of Canberra distance is given in equation (7).

$$CAB = \sum_{c=1}^{N} \frac{|V_c - V_{c+1}|}{V_c + V_{c+1}} \qquad (7)$$

5) Sorensen Distance (SD)-

The Sorensen distance between consecutive frames are calculated by equation (8).

$$Sorensen = \frac{\sum_{c=1}^{N} |V_c - V_{c+1}|}{\sum_{c=1}^{N} V_c - V_{c+1}} \qquad (8)$$

Mean square error, Euclidean Distance from La family[12], Sorensen and Canberra are from L_1 Family[12] Square chord belong to the Fidelity family[12] and after using this follow the same process explain in key frame extraction.

Table 1. Table captions should be placed above the table

4. EXPERIMETATION ENVIRONOMENT

The key frame extraction using Thepade's Hadamard Transform Error Vector Rotation (THdEVR) is implemented with Matlab. The experiment is performed with Intel i5 with 4 GB RAM. In all 15 videos are taken with 150 initial frames per video for experimentation. The test bed used for experimentation of 15 videos is as given in Fig. 3.

Fig 3: Sample Videos of the Data Set Used

5. RESULTS AND DISCUSSIONS

The proposed algorithm is used for key frame detection from videos and key frames are extracted. These frames are compared with the key frames extracted manually from the same videos. The completeness [9] [11] of the algorithms is calculated by equation (9).

$$Completness = \frac{Actual\ correct\ extracted\ frames}{Total\ expected\ extraction\ of\ frames} \quad (9)$$

Actual correct extracted frames means correct key frames extracted from video using proposed algorithms. Total expected extraction means key frames extracted manually.

In Fig.4 some key frame extracted of 'counting number' video by the proposed algorithm are shown.

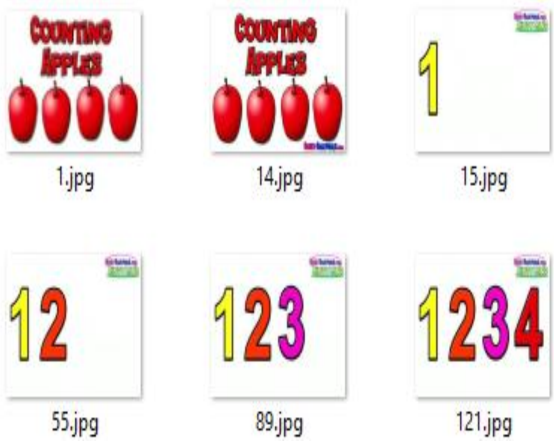

Fig 4: Key Frame Extracted for Proposed Algorithm of Counting Video

Table I, shows the completeness for different codebook sizes for the videos from the video test bed. Higher codebook sizes give better Completeness values. It also shows that the Euclidean distance gives better performance than that of other similarity measures used here. Sorensen also gives better completeness followed by that of Euclidean distance. Also in Fig. 5. the graphical representation of completeness of different codebook sizes of THdEVR algorithms for key frame extraction algorithms is shown with different similarity measures. Earlier key frame extraction methods are proposed with vector quantization algorithm like LBG [9], KPE[16], TCEVR[15] and THEVR [17]. The Comparison of THdEVR based key frame extraction with these existing is done in Fig. 6 here. The Fig. 6 shows comparison between THdEVR with TCEVR, THEVR, LBG [9] and KPE based video key frame extraction methods using Euclidean Distance. Fig. 6. Clearly shows that TCEVR given better completeness than that of LBG, KPE and THdEVR, THEVR for higher codebook sizes. THdEVR is faster in computations than of LBG and KPE.

Table 2. Percentage of Completeness for Different Codebook Sizes of Th_devr Different Similarity Measures used in Video Key Frame Extraction

	TH_DEVR with different Codebook Sizes									
Codebook Sizes → Similarity Measures ↓	**2**	**4**	**8**	**16**	**32**	**64**	**128**	**256**	**512**	**1024**

Canberra	56.04	57.49	57.26	55.10	51.40	54.77	54.00	40.82	33.57	18.63
Euclidean Distance	56.67	59.88	59.43	64.20	63.69	70.28	81.82	72.65	79.85	85.55
MSE	35.23	37.39	39.64	44.91	49.04	60.44	73.76	67.18	71.85	77.84
Square Chord	35.13	36.20	37.70	42.07	46.28	58.06	72.98	67.74	72.60	78.72
Sorensen	54.54	57.13	58.08	60.91	58.93	67.15	76.96	69.67	72.69	72.82

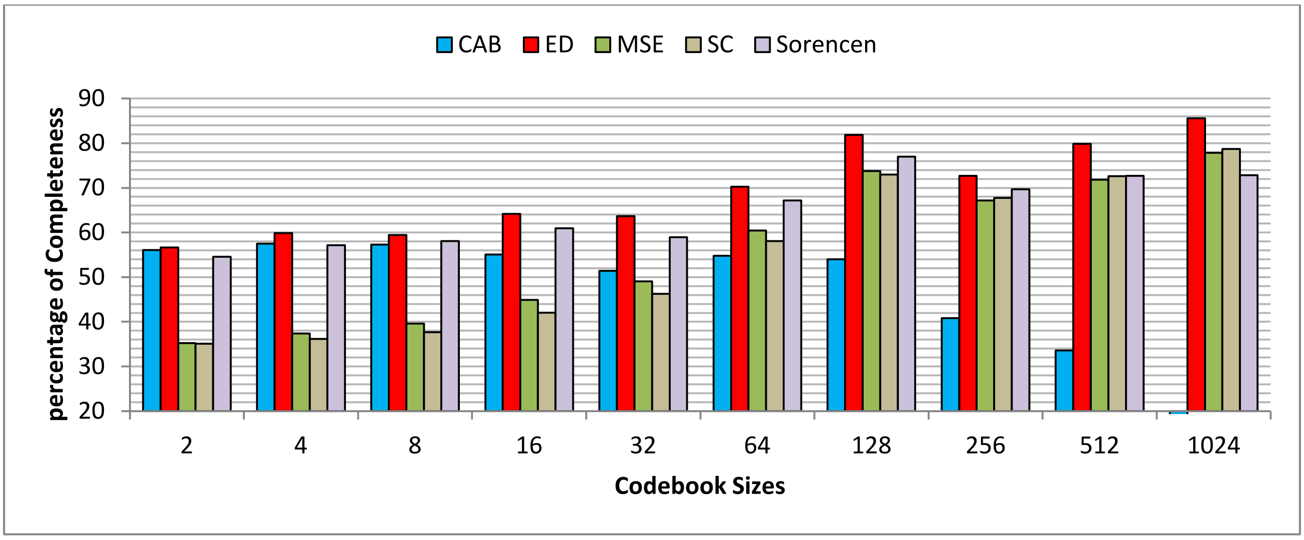

Fig 6: Percentage Completeness for Different codebook sizes for Proposed Key frame Extraction Technique using TH_DEVR

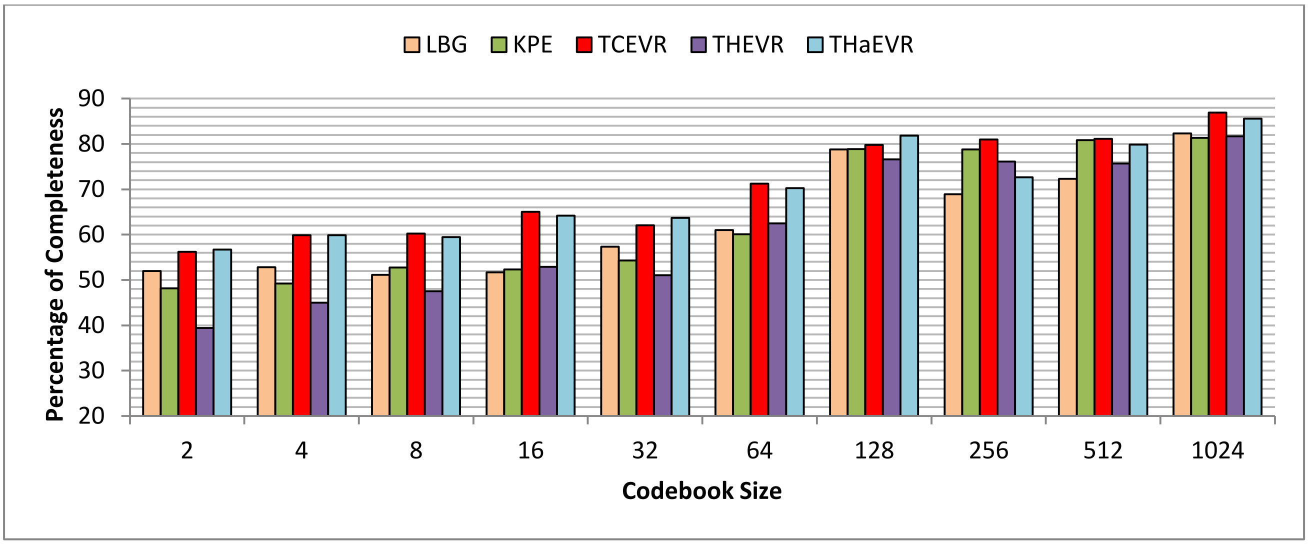

Fig 6: Percentage Completeness Performance Comparison of Completeness For LBG, KPE, TCEVR, THEVR and THdEVR With Euclidean Distance used in video Keyframe Extraction

6. CONCLUSION

In this paper a novel key frame extraction technique for video content summarization using Thepade's Hadamard Error Vector Rotation (THdEVR) codebook generation method of vector quantization is proposed and experimented for various codebook sizes and assorted similarity measures with a video test bed. The performance of the variations of proposed key frame extraction method is done using percentage completeness. The experimentation results have given better completeness percentage with higher THdEVR codebook sizes. For THdEVR codebook of size 1024 Euclidean distance give better key frames for video content summarization followed by the Sorensen Distance.

Thepade's Cosine Error Vector Rotation (TCEVR) gives better completeness than that other vector quantization algorithm i.e. Linde-Buzzo-Gray (LBG) ,Kekre Proportionate Error (KPE) and Thepade's Haar Error vector Rotation (THEVR) with Haar. The Thepade's Hademard Error Vector

Rotation (THdEVR) is computationally lighter than Thepade's Cosine Transform Error vector Rotation (TCEVR) codebook generation algorithm.

7. REFERENCES

[1] Z. Xiong, X. S. Zhou, Q. Tian, Y. Rui, and H. TS, "Semantic retrieval of video - review of research on video retrieval in meetings, movies and broadcast news, and sports," IEEE Signal Processing Magazine, vol. 23, no. 2, pp. 18–27, march 2006.

[2] V. Valdes and J. Martinez, "Efficient video summarization and retrieval tools," in 2011 9th International Workshop on Content-Based Multimedia Indexing (CBMI), june 2011, pp. 43–48.

[3] M. Furini, F. Geraci, M. Montangero, M. Pellegrini, a1, a2, and a3, "Stimo: Still and moving video storyboard for the web scenario," Multimedia Tools Appl., vol. 46, no. 1, pp. 47–69, Jan. 2010.

[4] Edward J. Y. Cayllahua Cahuina, Guillermo Camara Chavez," A New Method for Static Video Summarization Using Local Descriptors and Video Temporal Segmentation", IEEE Conference on Graphics, Patterns and Images (SIBGRAPI)XXVI. 2013, pages 226 - 233.

[5] B. T. Truong and S. Venkatesh, "Video abstraction: A systematic review and classification," ACM Transactions on Multimedia Computing, Communications and Applications, 2007.

[6] A. G. Money and H. Agius, "Video summarisation: A conceptual framework and survey of the state of the art," Journal of Visual Communication and Image Representation, pp. 121–143, 2008.

[7] Y. Gao, W.-B. Wang, J.-H. Yong, and H.-J. Gu, "Dynamic video summarization using two-level redundancy detection," Multimedia Tools and Applications, pp. 233–250, 2009.

[8] Y. Linde, A. Buzo, R. M. Gray. "An algorithm for vector quantizer design". IEEE Trans. Commun., COM-28(1):84-95, 1980.

[9] Sudeep D. Thepade, Pritam H. Patil, "Novel Keyframe Extraction for Video Content Summarization using LBG Codebook Generation Technique of Vector Quantization", International Journal of Computer Applications, Volume 111, Number 9, pp.-49-53, ISBN: 973-93-80885-22-9, Feb-2015

[10] Dr. Sudeep Thepade, Vandana Mhaske," New Clustering Algorithm for Vector Quantization using Haar Sequence", International and Communication Technologies (ICT), pp-1144-1149, April 2013.

[11] Sudeep D. Thepade, Pritam H. Patil, ''Novel Visual Content summarization in Videos using Keyframe Extraction with Thepade's Sorted Ternary Block Truncation Coding and Assorted Similarity Measures'', International Conference on Communication, Information & Computing Technology (ICCICT), Jan. 16-17, Mumbai, India pp. 1-5, 2015.

[12] Sung-Hyuk Cha, "Comprehensive Survey on Distance/Similarity Measures betwwen Probability Density Functions", International Journal of Mathematical Models and methods in Applied Sciences, Issue 4, Volume 1, 2007(300-307).

[13] 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), Volume (7) : Issue (3) : 2013

[14] Dr. H. B. Kekre, Sudeep D. Thepade,Varun K. Banura," Performance Comparison of Gradient Mask Texture Based Image Retrieval Techniques using Walsh, Haar and Kekre Transforms with Image Maps", International Conference on Technology Systems and Management,No. 3, 2011

[15] Sudeep D. Thepade, Pritam H. Patil, ''Novel Video Keyframe Extraction Technique using Thepade's Transform Error Vector Rotation Algorithm Cosine Transform and Assorted Similarity Measures'' Futuristic Trends in Computational analysis and Knowledge management 25th-27yh Feb 2015.

[16] Sudeep D. Thepade, Pritam H. Patil, ''Novel Video Keyframe Extraction using KPE Vector Quantization with Assorted Similarity Measures in RGB and LUV Color Spaces'', International Conference on Industrial Instrumentation and Control (ICIC) 28th -30th may 2015.

[17] Sudeep D. Thepade, Pritam H. Patil, ''Video Key Frame Identification using Thepade's Transform Error Vector Rotation Algorithm with Haar Transform and Assorted Similarity Measures'', International Conference on Communication and Signal Processing 6th-8th April 2015.