Novel Keyframe Extraction for Video Content Summarization using LBG Codebook Generation Technique of Vector Quantization

Sudeep D. Thepade, Ph.D
Professor, Department of Computer Engineering
Pimpri-Chinchwad College of Engineering
Savitribai Phule Pune University, India

Pritam H. Patil
Department of Computer Engineering
Pimpri-Chinchwad College of Engineering
Savitribai Phule Pune Univesity, India

ABSTRACT
In the current era, most of the digital information in the form of multimedia with a giant share of videos. Videos do have audio and visual content where the visual content has number of frames put in a sequence. Most of the consecutive frames do have very little discriminative contents. In video summarization process, several frames containing similar information are needed to get processed. This leads to redundant slow processing speed and complexity, time consumption. Video summarization using key frames can ease the speedup of video processing. In this paper, novel key frame extraction method is proposed with Linde-Buzo-Gray (LBG) codebook generation techniques of vector quantization with ten different codebook sizes. Experimentation done with the help of the test bed of videos has shown that higher codebook sizes of LBG have given better completeness in key frame extraction for video summarization. Experimental results are also discussed to represent the validity of the proposed method for video content summarization.

Keywords
video summarization, key frame, LBG, vector quantization, codebook.

1. INTRODUCTION
In the recent era, there has been an extreme increase in assembly and storage of video data on the internet. This increase in video data specifies efficient techniques for indexing, retrieval, and summarization of this data. However, these techniques cannot be progressed at the same speed [1]. This is due to the considerably different nature of video data which is not suitable for conventional retrieval, indexing, and summarization techniques. Therefore, there is a strong need of research work to solve this problem of video summarization and retrieval. Video nature itself provides a solution for this problem, as typically every video contains a lot of unnecessary information which can be removed or neglected to make video data more suitable for retrieval as well as indexing. This approach of removing redundancies from the video and generating condensed versions of the video is called video summarization [2]. The video summary contains the a good number significant and suitable content of the video. There must not be any redundancy in the summarized video, but at the same time, the original message of the video must be preserved [3].

The video contains summary be able to generated in many different forms. However, there are two different ways to generate video contain summary are static and dynamic [4]. The static video summary is deals with the extraction of the key frames from the video. The key frames are unmoving frames extracted from the video which hold the most distiminative content of the video and thus are representative of the video [5,6]. The dynamic video summary contains small shots that are accumulated in a time ordered sequence. The mining of static summaries or key frame extraction is very much important, because it provides more flexibility. The key frame extraction can be used as a pre-processing step in video content summarization [7,8] which tackles with the problem of processing a large number of video frames.

Here novel key frame extraction method is proposed using Linde-Buzo-Gray (LBG) codebook generation technique of vector quantization (VQ) with ten different codebook sizes. Vector quantization is a three step process- with Encoding, Codebook generation and Decoding being the steps as given in Fig. 1. In a proposed method of key frame extraction for video content summarization first two steps of vector quantization are used..

2. LITERATURE SURVEY
2.1 Vector Quantization(VQ)
Vector quantization is one of the popular lossy data compression techniques which includes the process of clustering. In vector quantization, codebook is generated for each frame as signature which is calculated according to a specific VQ algorithm. In other words, Vector Quantization is mapping function that maps k dimensional vector space to finite set $CB = \{C_1, C_2 \dots C_N\}$. The set CB is called a codebook consisting of N number of code vector and each code vector $Ci= \{c_{i1},c_{i2}....c_{ik}\}$ is of K dimension [9]. Here ten different Codebooks sizes like 2, 4, 8, 16, 32, 64, 128, 256, 512 and 1024 are used.

Vector Quantization is used in many applications [10], such as speech recognition and face detection [11], [12], Image segmentation [13], speech data compression [14], Contained Based Image Retrieval (CBIR) [15], also in iris recognition [16] etc.

In process of generating codebook the video frame is separated into several k dimensional training vectors as mention above $C= \{C_{i1}, C_{i2} ..C_{ik}\}$ and codebook generation techniques are applied. In process of encoding video frame is separated into several k dimensional vectors and each vector is encoded by the index of codeword by a look-up table method. The encoded codeword results are stored in an index table. During the procedure of decoding, the receiver uses the same codebook to translate the index back to its corresponding codeword for reconstructing the image [17].

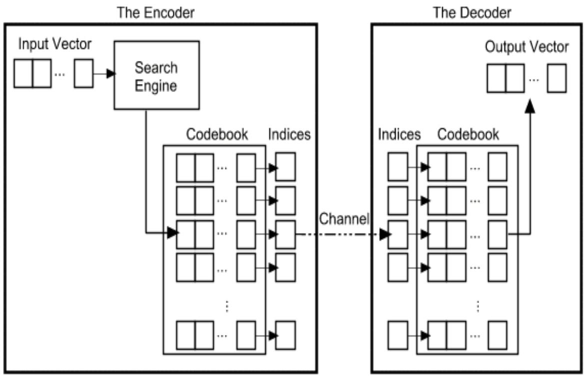

Fig. 1. The Encoder and Decoder in a vector quantization

2.2 Linde-Buzo-Gray (LBG) Codebook Generation Technique

The method one of the popular used to generate codebook is the Linde-Buzo-Gray (LBG) algorithm [18]. In this technique centroid is calculated as the first codevector for the training set. Two vectors C_1 & C_2 are then generated by adding and subtracting a constant error to from the codevector. Using Euclidean distances all the vectors are compared with vectors C_1 & C_2 and two clusters are created on the basis of the nearest of C_1 or C_2. An algorithm for the LBG codebook generation technique can be explained in following steps.

Step 1: Read an image and split the R, G and B component of image into non overlapping blocks and convert each block to vector form a training vector set.

Step 2: Calculate the centroid of a training vector set that contained we call as a codevector.

Step 3: Codevector we calculate add and subtract a constant error vector to form two codevector as C_1 and C_2.

Step 4: Using Euclidean distance we calculate distance between training vector and codevector C_1 and C_2 and split the cluster into two according to nearest basis.

Step 5: Again calculate the centroid (codevector) for clusters to obtained in the above step 4.

Step 6: Repeat step 2 to 5 till codebook of required size is obtained

Step 7: stop

2.3 Key Frame based Video Summarization

A video nothing but a continuous sequence of a number of frames, each frame being 2-D image. The basic unit of video is frame. The video can also be said as a collection of scene as well as collection of shots that have the same context of a frame.

Key frames based video content summarization can be classified into three different types. These are explained follows [19].

1) Classification based on sampling[19]:

In this method key frames choose uniformly under-sampling, without considering the video information [19]. The summary, which produced by these does not show all the video parts. So it may cause some of key frames with similar contents.

2) Classification based on scene segmentation[19]:

In this method key frame extraction is done using extraction of scenes. The scene includes all parts with a related information link in the video or in the same time [19]. The drawback of these is producing a summary, which does not take into account the temporal position of the frames.

3) Classification based on shot segmentation[19]:

In this method extract adapted key frames from video. Only first and last images of a shot are considered as key frame [19]. These methods are much more effective for steel shot and small variation in a shot. But they cannot provide a sufficient representation of shot with strong movements.

Key frame based video summarization works on frames in a three steps. In a first step 1 frames from an original video are extracted. In step 2 extracted frames are grouped together according to the respective algorithm. In step 3 selected key frames are used for video summarization. All this procedure is shown in Fig. 2.

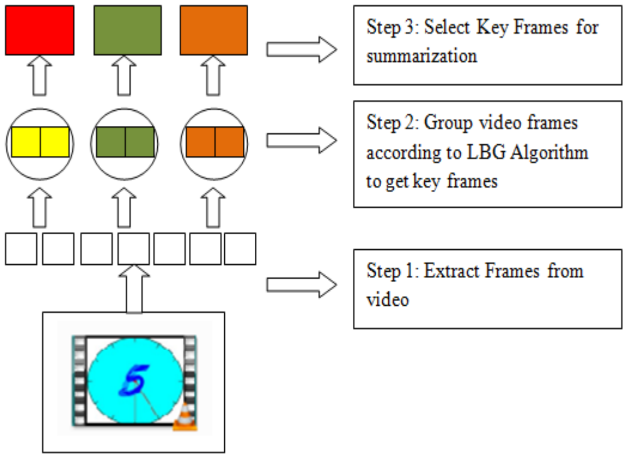

Fig. 2. Key Frame Based Video Summarization using LBG Algorithm

3. PROPOSED KEY FRAME EXTRACTION USING LBG

In his section we proposed novel key frame extraction using Linde-Buzo-Gray codebook generation technique with different codebook sizes.

3.1 Key Frame Extraction using LBG

A video consisting of 'N' frames is taken as an input and, for each pair of consecutive frames 'V_n' and 'V_{n+1}' having feature vectors respectively as 'FV_n' and 'FV_{n+1}'. Here, feature vector is nothing but the codebook for each frame calculated by LBG algorithm. The $diff(n)_i$ could be the difference between two consecutive video frame features as given in equation 1 calculated by Euclidean distance equation 1.

$$diff(n+1,n) = \sqrt{\sum_{i=1}^{n} |FV_{n+1}(i) - FV_n(i)|} \quad (1)$$

The threshold 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 2, 3 and 4.

$$Mean(M) = \frac{\sum_{i=1}^{N} diff(i)}{N-1} \qquad (2)$$

$$Standard\ Deviation(S) = \sqrt{\frac{\sum_{i=1}^{N-1}(diff(i)-M)^2}{N-1}} \qquad (3)$$

$$Threshold(T) = M + a \times S \quad Here\ a = Constant \qquad (4)$$

For getting the key frames wherever the (diff(n+i, n) > Threshold), those frames V_{n+1} will is considered as Key frame[20].

4. EXPERIMENTATION ENVIRNOMENT

The implementation of the proposed method is done in MATLAB with the basic system of an Intel core 2 duo (2.93GHz) with 4GB RAM. The test bed used for experimentation 11 videos as given in Fig. 3.

Fig. 3. Video Test Bed used in Experimentation

5. RESULTS AND DISCUSSION

The LBG codebook generation technique is used for key frame detection from the videos and key frames are extracted. For performance comparison of various codebook sizes of LBG codebooks for key frame extraction, the extracted key frames are compared with the key frames extracted manually from the same videos to find The completeness [20] of the algorithms as given in equation 5.

$$Completness = \frac{Actual\ correct\ extracted\ frames}{Total\ expected\ extraction\ of\ frames} \qquad (5)$$

Actual correct extracted frames means correct key frames extracted from video using proposed algorithms. Total expected extraction means key frames extracted manually. In Fig.3 some key frame extracted from 'clock' video by the proposed algorithm are shown in Fig. 4.

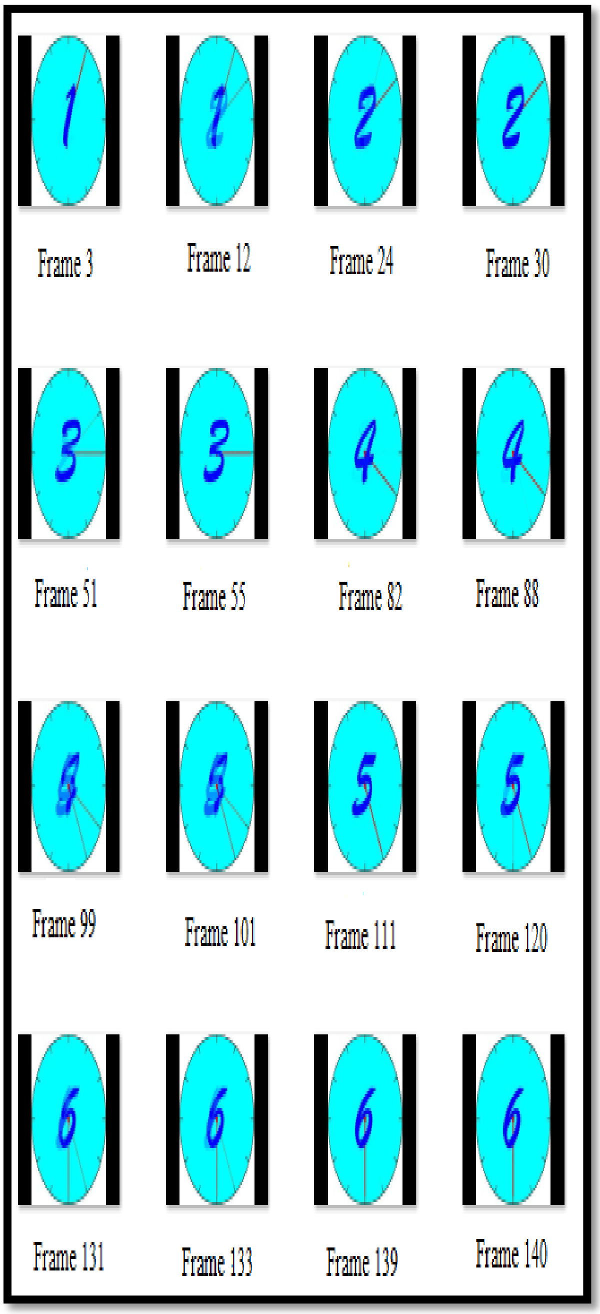

Fig. 4. Key frames Extracted using LBG Algorithm for 'clock' video

Table I shows the completeness for different codebook sizes for the videos from the video test bed. Higher codebook sizes give better Completeness values. Also in Fig. 5. The graphical representation of completeness of different codebook sizes of LBG algorithms for key frame extraction algorithms is shown

Table 1: Percentage Of Completeness For Different Codebook Sizes Of Lbg Algorithm

	LBG Codebook Sizes									
Codebook Sizes → Videos ↓	**2**	**4**	**8**	**16**	**32**	**64**	**128**	**256**	**512**	**1024**
Average Completeness	**46.68**	**53.22**	**52.03**	**51.66**	**56.83**	**59.19**	**75.78**	**70.80**	**73.44**	**81.77**
Counting	83.33	83.33	83.33	83.33	83.33	100	100	100	100	100
Bee	40.81	65.3	56.66	65	80	81.66	91.66	20	48.33	55
Clock	22.66	20	4.66	4.66	11.33	14	62.66	52	56	50
Earth	40	86.66	83	76.66	76.66	66.66	63.33	70.5	60	73.33
Zoo	5.88	23.52	29.41	26.47	32.35	35.29	35.29	88.23	97.05	85.29
Bird	8.69	8.69	8.69	8.69	8.69	8.69	95.65	88.46	88.46	100
Mobile	52.17	41.3	54.34	60.86	63.04	65.21	63.04	45.65	71.73	69.56
Cartoon	93.33	100	100	100	100	100	100	100	100	100
Number	75	79.16	70.93	66.66	87.5	100	100	100	100	100
Kid & Dog	51.66	39.56	43.33	40	38.33	41.66	40	20	18.33	88.33
Football	40	38	38	36	44	38	82	94	68	78

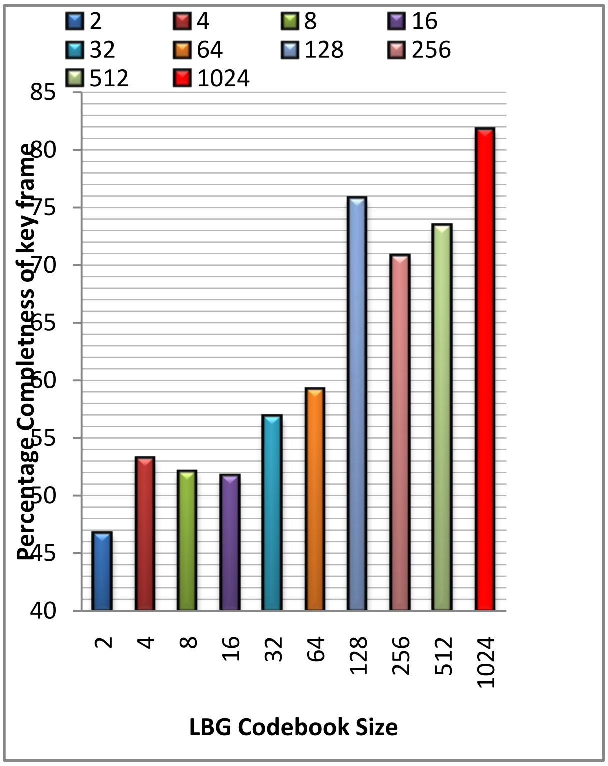

Fig.5. Percentage Completeness for Different codebook sizes of LBG algorithm for Proposed Key frame Extraction Technique

6. CONCLUSION

In this paper a novel key frame extraction technique for video content summarization using LBG codebook generation method of vector quantization is proposed and experimented for various codebook size 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 LBG codebook sizes. The LBG codebook of size 1024 give better key frames for video content summarization.

The proposed key frame extraction method is rotation and size invariant with references to the size of video frames.

7. REFERENCES

[1] A.F. Smeaton, "Techniques used and open challenges to the analysis, indexing and retrieval of digital video", Information Systems 32 (4) (2007) 545–559.

[2] A.G. Money, H. Agius, "Video summarisation: a conceptual framework and survey of the state of the art", Journal of Visual Communication and Image Representation 19 (2) (2008) 121–143.

[3] S.E.D. Avila, A.B.P. Lopes, L.J. Antonio, A.d.A. Araújo, "VSUMM: a mechanism designed to produce static video summaries and a novel evaluation method", Pattern Recognition Letters 32 (1) (2011) 56–68.

[4] Y. Li, T. Zhang, D. Tretter, "An overview of video abstraction techniques", Technical Report HP Laboratory, 2001, HP-2001-191.

[5] G. Ciocca, R. Schettini, "Innovative algorithm for key frame extraction in video summarization", Journal of Real-Time Image Processing 1 (1) (2006) 69–88.

[6] B.T. Truong, S. Venkatesh, "Video abstraction: a systematic review and classification", ACM Transactions Multimedia Computing, Communications and Applications. 3 (1) (2007).

[7] N. Ejaz, U. Manzoor, S. Nefti, S.W. Baik, A collaborative multi-agent framework for abnormal activity detection in crowded areas, International Journal of Innovative Computing, Information and Control 8 (6) (2012).

[8]] P.L. Venetianer, H. Deng, Performance evaluation of an intelligent video surveillance system- a case study, Computer Vision and Image Understanding 114 (11) (2010) 1292–1302.

[9] H. B. Kekre , Tanuja K. Sarode , Jagruti K. Save, “New Clustering Algorithm for Vector Quantization using Walsh Sequence", International Journal of Computer Applications (0975 – 8887) Volume 39– No.1,February 2012.

[10] Ahmed A. Abdelwahab, Nora S. Muharram, ‘‘A Fast Codebook Design Algorithm Based on a Fuzzy Clustering Methodology’’, International Journal of Image and Graphics, vol. 7, no. 2 pp. 291- 302, 2007.

[11] Chin-Chen Chang, Wen-Chuan Wu, ‘‘Fast Planar-Oriented Ripple Search Algorithm for Hyperspace VQ Codebook’’, IEEE Transaction on image processing, vol 16, no. 6, June 2007.

[12] C. Garcia and G. Tziritas, ‘‘Face detection using quantized skin color regions merging and wavelet packet analysis,’’ IEEE Trans. Multimedia, vol. 1, no. 3, pp. 264--277, Sep. 1999.

[13] H. B. Kekre, Tanuja K. Sarode, Bhakti Raul, ‘‘Color Image Segmentation using Kekre’s Algorithm for Vector Quantization’’,International Journal of Computer Science (IJCS), Vol. 3, No. 4, pp.287-292, Fall 2008. available: http://www.waset.org/ijcs.

[14] H. B. Kekre, Ms. 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 5, pp.: 1-8, September 2009.

[15] H. B. Kekre, Kamal Shah, Tanuja K. Sarode, Sudeep D.Thepade,’’Performance Comparison of Vector Quantization Technique KFCG with LBG, Existing Transforms and PCA for Face Recognition’’, International Journal of Information Retrieval (IJIR), Vol.02, Issue 1, pp.: 64-71, 2009.

[16] Kekre H.B, Sarode T.K. ; Bharadi V.A. ; Agrawal A.A, Arora R.J., Nair M.C.," Iris Recognition Using Vector Quantization", International Conference on Signal Acquisition and Processing, 2010. ICSAP '10. 978-1-4244-5725-0, pp. 58 - 62.

[17] Tzu -Chuen Lu, Ching-Yun Chang, “A Survey of VQ Codebook Generation”, Journal of Information Hiding and Multimedia Signal Processing, ISSN 2073-4212, Ubiquitous International Volume-1, Number 3, July 2010.

[18] Y. Linde, A. Buzo, and R. M. Gray, “An algorithm for vector quantizer design,” IEEE Trans.Commun., vol.COM-28, no. 1, pp.: 84-95, 1980.

[19] Sabbar, W.; Chergui, A.; Bekkhoucha, A., "Video summarization using shot segmentation and local motion estimation," InnovativeComputing Technology (INTECH), 2012 Second International Conference on, vol., no., pp.190, 193, 18-20 Sept. 2012.

[20] Sudeep.D.Thepade, Ashvini A. Tonge "Extraction of key frames from Video using Discrete Cosine Transform" International Conference on Control, Instrumentation, Communication and Computation technology, IEEE (1446=1449), 2014, India