A Graph Representation Based Fuzzy C Means Approach to Video Summarization

Shanmukhappa Angadi
Department of Computer Science and Engineering, Visvesvaraya Technological University, Belgaum, India
Email: vinay_angadi@yahoo.com

Vilas Naik
Department of Computer Science and Engineering, Basveshwar Engineering College, Bagalkot, India
Email: vilasnaik_h@rediffmail.com

Abstract—**The Video summarization normally includes shot boundary detection, key frame extraction from each shot and generation of summary. The proposed video summarization approach eliminates shot boundary detection by employing segmentation based key frame extraction scheme and the segmentation is achieved by fuzzy C Means (FCM) clustering. The video summarization algorithms basically perform structural analysis of the video. Recently algorithms utilizing graph theoretic approaches for structure analysis are found in literature. The proposed method models a video segment formed by FCM clustering as undirected weighted graph with frames in cluster as nodes of graph and Euclidean feature distance as edge weight. The eccentricity of each graph node is used to determine its connectivity to other nodes of the graph. If the eccentricity of node is greater than the mean eccentricity of graph that node is one of the important node and the frame represented by that node is selected as key frame. This is done over all the clusters and the summary is created using key frames and merging them on the basis of their timeline. This method ensures that video summary represents the most unique frames of the input video and gives equal attention to preserving continuity of the summarized video. The performance of the algorithm is evaluated for compactness, fidelity and informativeness measures.**

Index Terms—**fuzzy C means clustering, graph modeling, graph eccentricity, keyframe extraction, video summarization**

I. Introduction

Nowadays, the ever increasing number of videos has necessitated automatic analysis, synopsis generation and extraction of visual information. Video summarization appears as a compact representation of a video sequence useful for various video applications such as video browsing, indexing and retrieval. A video summary can be a preview of a sequence combining a limited number of video segments or a collection of key frames properly chosen from the video sequence. Although a key frame-based video summarization may lose the spatial-temporal properties and audio content of the original video sequence, it is clearly the simplest method. The use of key frames reduces the amount of information required in video indexing and provides the framework for dealing with the video content for retrieval purposes.

The reported research in the field of video analysis and video summarization relies on the detection of shot boundaries. A shot is detected when the difference in a metric between consecutive frames exceeds a threshold. This measure can be computed by using either features containing global information, such as color histograms or more complex features, such as image edges, motion vectors and probability density. After the shot detection stage, key frames can be selected by techniques which explore the visual and motion structure of the shot. The purpose of these methods was to maintain the temporal continuity of the extracted key frames providing fast indexing or browsing ability.

In this work, a unified method for shot clustering and video summarization in the uncompressed video domain utilizing the video structure representation in the feature space is proposed. The technique is based on the Fuzzy C Means clustering methodology. The experiments have revealed that clustering group frames of similar visual features. The clustering process is followed by an efficient structural description in the form of undirected graph and further analysis of graph in terms of graph parameters to detect representative frames. The graph parameter employed in the proposed work is eccentricity of vertices which is defined as maximum of shortest distances from a vertex to rest of the vertices.

The remaining part of the paper is organized in to 4 sections. Section II presents the related work and background of the algorithms used. Section III describes the new proposed graph theoretic algorithm. In Section IV experimentation and results are presented. Section V brings up the conclusion

II. Related Work

Previous methods in frame-based video summarization mainly relied in shot boundary detection [1]. A shot is detected when difference of a certain measure between consecutive frames exceeds a threshold. This measure can be computed by using either features containing global information (colour histograms) or more complex

Manuscript received July 28, 2014; revised July 2, 2015.

doi: 10.18178/ijsps.4.4.304-310

features, such as image edges, motion vectors and probability densities. From a shot, key frames are to be extracted. In a very simple approach the first frame [2] or the first and the last frames [3] in each shot are selected as keyframes. Other methods use a certain frame rate in order to extract a down-sample version of video. All these approaches do not consider the dynamics of the visual content or the motion analysis and the type of the shot boundary and they often extract a fixed number of key frames per shot. Such approaches are basically called as Sampling based methods for key frame extraction for summarization.

There are other types of key frame extraction methods like shot based [4], segmentation based [5] are also available. Shot based approaches detect the shots using shot boundary detection method and selects key frames from each shot based on some logic. The segmentation based approaches segment the videos in to groups frames of similar visual content by clustering like technique and extracts key frames from each segment.

Clustering has been successfully applied for both key frame selection and video summarization. The mechanism in [6] cluster the frames using connectivity clustering algorithm and selects frames closest to the centroids of each cluster as key frames. As the content of the shot may change significantly due to camera operations and object motion, in [7] the clustering based algorithm is modified to extract more than one key frames. Two key frames are extracted for clusters with high intercluster variance (the closest and the farthest to the centroid). Frames with large deviations from the average luminance of the shot are selected as key frames too.

In [8] an agglomerative hierarchical clustering is applied to extract a predefined number of key frames and used them to represent videotaped meetings and presentations. A key point based method in which global features are utilized to identify scenes through clustering due to the visual similarity among video frames of the same scene, and local features to summarize each scene is presented in [9]. Video key-frame extraction for summarization using unsupervised clustering using multi-features is presented in [10]. To deal with problems existed in the traditional clustering algorithms, an improved key-frame extraction algorithm based on fuzzy C-means clustering using color feature is presented in [9]. The method clusters shots into several sub-shots. As maximum image entropy corresponds to the maximum amount of information in the information theory, the maximum entropy frame is extracted as the key-frame from each sub shot.

The literature also divulges the employment of graph theoretic approaches for pattern recognition problems, where in pair-wise similarities between all data objects are used to construct a weighted graph as an *adjacency matrix* (*weight matrix* or *similarity matrix*) that contains all necessary information. Representing the data set in the form of an edge-weighted graph converts the data processing problem into a graph analysis problem. Researchers have proposed graph theoretic solutions to issues in video summarization. The solutions are proposed for shot boundary detection, Key frame extraction, event retrieval and video indexing. The graph theoretic Shot Boundary Detection (SBD) algorithm implemented is explained in [11]. Graph theoretic segmentation algorithms are getting more popular in the field of pattern recognition and computer vision *Dominant sets*, which is a very novel concept, is a remarkably important contribution to the graph theory domain and has already found applications in image segmentation. This novel concept is applied to video scene detection problem in [12]. The idea of key frame extraction from single shots in video sequences is presented in [13]. The method is implemented by an efficient two-step algorithm. In the first step, a graph is constructed, where each node is associated to a single frame of the shot. The second step, extracts key frames based on the principle of their maximum spread. The number of the selected key frames is controlled by an adaptively defined threshold, while the validity of the results is evaluated by the fidelity measure.

In [14] a novel approach for video summarization based on graph optimization is presented. The approach emphasizes both a comprehensive visual-temporal content coverage and visual coherence of the video summary. The approach has three stages. First, the source video is segmented into video shots, and a candidate shot set is selected from the video shots according to some video features. Second, a dissimilarity function is defined between the video shots to describe their spatial-temporal relation, and the candidate video shot set is modeled into a directional graph. Third, a dynamic programming approach to search the longest path in the graph which will represent final video summary. The work in [15] proposes a new approach for event detection and summarization of news videos. The approach is mainly based on two graph algorithms Optimal Matching (OM) and Normalized Cut (NC). Initially, OM is employed to measure the visual similarity between all pairs of events under the one-to-one mapping constraint among video shots Then, news events are represented as a complete weighted graph and NC is carried out to globally and optimally partition the graph into event clusters.

The algorithm in [16] presents a method for summarizing multi-view videos by constructing a spatio-temporal shot graph and to formulate the summarization problem as a graph labeling task. The spatio-temporal shot graph is derived from a hypergraph, which encodes the correlations with different attributes among multi-view video shots in hyperedges. The shot graph identifies clusters of event-centered shots, and pick out the candidates for summarization by random walks. The summarization result is generated through solving a multi-objective optimization problem based on shot importance evaluated using a Gaussian entropy fusion scheme.

The paper [17] proposes a unified approach for video summarization based on the analysis of the video structure. The method originates from a data learning technique that uses the membership values produced by an over-partitioning mode of the FCM algorithm to find the connection strength between the resulting set of prototype centers. The final clustering stage is implemented by using the minimal spanning tree produced by the connectivity matrix. Based on the MST edge weights value, the clusters are derived without supervision. The algorithm is finalized by the detection of video shots and the selection of key frames from each one. The method is evaluated by using objective and subjective criteria and its applicability to elongated video data set structures is very satisfactory.

Various Key frame extraction approaches have been proposed in literature including cluster based approaches. Most of the clustering based approaches perform clustering of similar frames within a shot after shots have been separated using SBD algorithm. Then some frames are extracted, generally one frame from each cluster. The work presented in this paper employs FCM clustering of similar frames of video sequence in feature space. The RGB entropy features are used to represent every frame. The second step of method represents each cluster as weighted undirected graph with each frame as node, every edge weight is the Euclidean distance between feature vectors representing frame nodes connected by the edge. The frames representing nodes whose eccentricity is greater than average eccentricity of all the nodes in graph are selected as key frames.

III. Proposed Clustering Based Graph Theoretic Algorithm for Key Frame Extraction

In this work, a unified approach for video summarization based on the analysis of the video structure is proposed. The method originates from a data learning and grouping technique that uses the membership values produced by the FCM (Fuzzy C Means) algorithm to group the similar frames of video into clusters. The connection strength within the resulting set of cluster is represented by a connectivity weighted graph. The frames in a cluster are arranged in time line and weighted undirected graph is constructed where in each node represents a frame in that cluster and edges have weight equivalent to Euclidean distance between nodes connected by edges in feature space.

Finally key frames key frames are selected from each cluster represented as weighted undirected graph. The method calculates eccentricity of each node which is maximum of the Euclidean distances calculated between a node and rest of the nodes. The frames which have eccentricity value greater than threshold are selected as key frames.

The proposed key frame extraction algorithm works in three stages, extracting features for cluster as a undirected graph and extracting key frames from each cluster. The three stages of proposed algorithm are presented in subsequent sections.

A. Feature Extraction

The color, texture, shape and motion estimation have been commonly used features in the literature. It is observed that color and motion features play a dominant role in the extraction of characteristics from the videos and hence the color information is adopted in this work. Normally color distribution is estimated in the form of histogram, the proposed work estimates the information of color content at different color intensity planes of images in terms of entropy values. The entropy computed for a color plane in an image gives the average information conveyed by an image. In the proposed work the entropy of three layers of RGB are computed and feature vector is written as in (1).

$$E_{f_i} = [\, E_r \quad E_g \quad E_b] \tag{1}$$

where:

E_{f_i}: Feature vector of i^{th} frame from video sequence

Er: Entropy of Red plane

Eg: Entropy of green plane

Eb: Entropy of blue plane of an RGB frame from video sequence

For all the N frames in the video sequence feature vectors are determined and these feature vectors are clustered by clustering algorithm.

B. Fuzzy C-Means (FCM) Clustering

The purpose of clustering based video segmentation is to segment video sequence into groups where each group represents a sequence of frames having similar contents, and then select key frames from each group for summarization. The video segmentation methods can be classified into two groups: the Shot Change Detection (SCD) approach for which thresholds have to be pre-assigned, and the clustering approach for which a prior knowledge of the number of clusters is required. In this work a video segmentation method using an entropy features-based fuzzy c-means (EBFCM) clustering algorithm is proposed. The EBFCM clustering algorithm is composed of two phases: the feature extraction phase, the clustering phase. The Image entropy values of R, G and B color channels which reflects the uniformity of the image color distribution and it represents the amount of information contained in the image color channels and are employed as features for representing image frame in feature space and perform FCM clustering.

Fuzzy C Means (FCM) is a method of clustering which allows one piece of data to be in the right position to two or more clusters. FCM starts with an initial guess for the cluster centers, which are proposed to mark every data point a membership rank for every cluster. By iteratively updating the cluster centers and the membership grades for each data point, FCM iteratively moves the cluster centers to the correct place within a dataset. This iteration is based on minimizing an objective function that symbolizes the distance from any given data point to a cluster center weighted by that data point's membership rank.

Let f_v=$f_1\, f_2 \ldots\ldots f_N$ indicates a video with N frames to be partitioned into c clusters, where E_{f_i} represents feature vector containing three entropy values of each frame f_i as in (1). The algorithm is an iterative optimization that minimizes the cost function defined as follows:

$$J = \sum_{j=1}^{N} \sum_{i=1}^{C} U_{ij}^{m} \left\| E_{f_i} - V_i \right\|^2 \tag{2}$$

where U_{ij} represents the membership of frame f_j in the i^{th} cluster, V_i is the i^{th} cluster center. The parameter m controls the fuzziness of the resulting partition. The cost function is minimized when frames close to the centroid of their clusters and are assigned high membership values, and low membership values are assigned to frames with data far from the centroid. The membership function represents the probability that a frame belongs to a specific cluster. In the FCM algorithm, the probability is dependent solely on the distance between the frame feature vector and each individual cluster center in the feature domain. The membership functions and cluster centers are updated by the following (3) and (4).

$$U_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{E_j - V_i}{E_j - V_k} \right)^{2/(m-1)}} \tag{3}$$

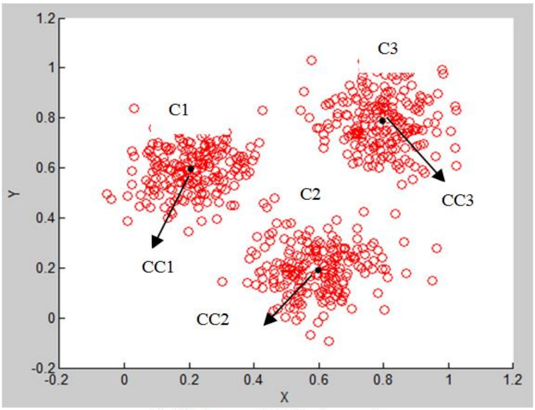
$$V_i = \frac{\sum_{j=1}^{N} U_{ij}^{m} \; E_{f_j}}{\sum_{j=1}^{N} U_{ij}^{m}} \tag{4}$$

U_{ij}: membership of frame f_j in the i^{th} cluster

V_i: i^{th} cluster center

E_{f_i}: feature vector containing three entropy values of each frame f_i

Starting with an initial guess for each cluster center, the FCM converges to a solution for V_i representing the local minimum or a saddle point of the cost function.



C: Cluster CC Cluster centers

Figure 1. Result of fuzzy c-means clustering of frames.

Fuzzy C-Means (FCM) is a data clustering technique in which a dataset is grouped into n clusters with every datapoint in the dataset belonging to every cluster to a certain degree. A certain data point that lies close to the center of a cluster will have a high degree of membership to that cluster. One advantage of FCM clustering from other techniques like k-means is the type of clustering output that is a membership matrix. This matrix specifies the most representative frames of each cluster directly. The Experiments in this works shows that most of the clusters contain maximum frames from single semantic shot and very few frames from other part of video. The color information representation is used as feature for clustering and RGB entropy values are extracted from each frame to construct a feature vector. The key frame extraction algorithm selects key frames as explained in the next section. The Fig. 1 depicts the typical clustering in two dimensional feature space. The work proposed work performs clustering of 3 dimensional feature vectors.

The key frame extraction algorithm selects key frames as explained in the next section.

C. Graph Based Approach for Key Frame Extraction

The proposed method cluster the video clip under consideration in to clusters of similar frames by employing Fuzzy C Means (FCM) clustering as explained in previous section are represented as undirected weighted graph. Each cluster basically contain frames from one or more shots with similar visual content. Based on these shots, a complete undirected weighted graph, with frames as its nodes and with frame similarities as weights of its edges is constructed to model the similarity among all pairs of frame in a video. The edges of the graph are represented by the Euclidean distance between pair of frames represented as two vertices connected by the edge. The metric space formed by vertex set and distance function is used to define the individual vertex connectivity to rest of the graph. The eccentricity ∈ of a vertex *υ* is the greatest geodesic distance between *υ* and any other vertex. It can be thought of as how far a node is from the node most distant from it in the graph and as well as from centre of the graph. The nodes of the graph in this work are frames clustered using FCM algorithm. So the nodes exhibit lot of similarity, but nodes with higher eccentricity are on the diameter of the graph and typically different from rest of the nodes. Such frames representing these nodes are typically different in content from rest of the frames in that cluster so these frames are considered as Key frames. To attain comprehensive coverage and select informative Key frames all those frames represented by nodes having eccentricity greater than 90% of maximum eccentricity values are considered as key frames. The node of graph model is represented as feature vector as

$$D_i = E_{f_i} = [\, E_r \quad E_g \quad E_b] \tag{5}$$

$$W_{ij} = \sqrt{(\, D_i^{\,2} - D_j^{\,2})} \text{ (Euclidean Distance)}$$

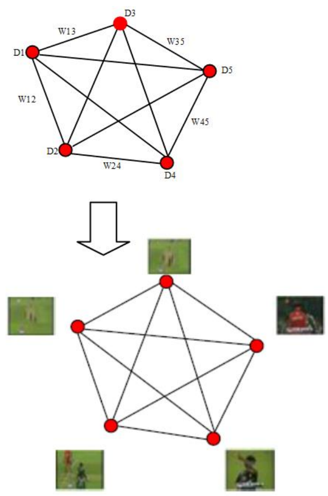


Figure 2. Graph model representation of frames clustered by fuzzy C means clustering.

The Fig. 2 depicts the graph representation of a cluster with vertices and edges.

IV. Experimental Results

In the experiments conducted, the video sequences used are MPEG and AVI compressed and downloaded from the You Tube and soccer video repository [18]. They are first decompressed using the official MPEGcodec from MPEG Software Simulation Group [19]. The proposed method was applied on 10 randomly selected video clips. The performance evaluation of the proposed approach is based on the parameters; compactness, Informativeness and fidelity.

In the proposed work first entropy values of all the frames are computed as explained in Section 3.1. On these feature vectors representing video frames Fuzzy C Means (FCM) clustering is performed and the plot of feature vectors and clusters formed are shown in Fig. 3.

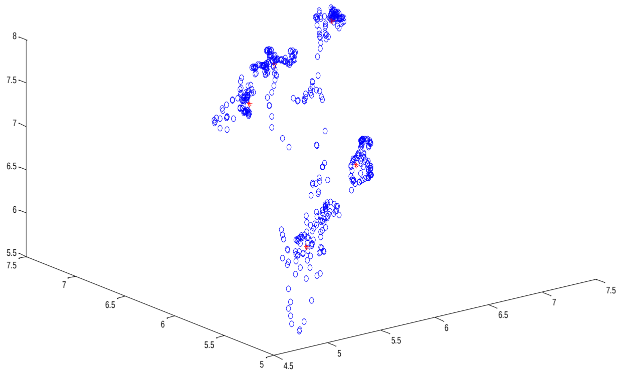


Figure 3(a). Plot of feature vectors of frames and cluster centers.

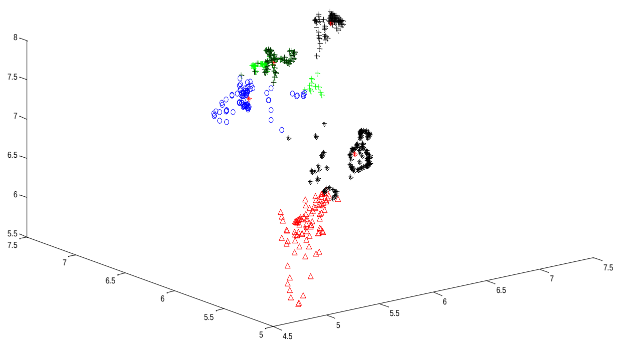
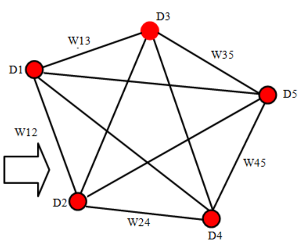


Figure 3(b). Plot of feature vectors of different clusters detected and cluster centers.

w11	w12	w13	w14	w15
w21	w22	w23	w24	w25
w31	w32	w33	w34	w35
w41	w42	w43	w44	w45
w51	w52	w53	w54	w55



Di: Nodes of the graph
W_{ij}: Edge between node *i* and *j*

Figure 4. Similarity matrix and graph model of cluster.

The clusters formed by FCM are modeled as undirected weighted graph whose vertices are represented by frames within cluster and edge weights by similarity values by constructing similarity matrix. The similarity between frames is computed as Euclidian distance as shown in (5).

An example similarity matrix and its graph model representation is as shown in Fig. 4.

From the graph model representation of each cluster formed key frames are selected as explained in Section III(c).

A. *Performance Evaluation of Algorithm*

A difficult issue of the key-frame extraction problem is related to the evaluation of the extracted key-frames, since it is rather subjective which frames are the best representatives of the content of a shot. There are several quality measures that can be used to evaluate the efficiency of the algorithms. Two quality measures are used in this work. The first is compactness or compression ratio second is the Fidelity measure, and finally Informativeness measure to access the capability of content coverage is used.

A video summary should not contain too many key frames since the aim of the summarization process is to allow users to quickly grasp the content of a video sequence. For this reason, algorithm is also evaluated with the compactness (compression ratio) of the summary that can be generated by extracted key frames. The compression ratio is computed by dividing the number of key frames in the summary by the length of video sequence. For a given video sequence “St”, the compression rate is thus defined as in (6). This metric gives an indication of the size of the summary with respect to the size of the original video.

$$CRatio(S_t) = 1 - {}^{\gamma_{NKF}}/_{\gamma_{NF}} \qquad (6)$$

Fidelity is used as a measure of the comprehensive coverage of a video. It is based on the metric of the semi-Hausdorff distance. The Fidelity measure is computed as the maximum of the minimum distances between the key frame set and the frames in the original video sequence. The high eccentricity value which is considered as parameter for selecting key frames will pick up high fidelity frames to attain comprehensive coverage.

Suppose that the key frame set *R* consists of *K* frames, $=\{KFj|j=1, 2, ..., k\}$, while the shot frame set *S* consists of *N* frames, $S=\{Fi|i=1, 2, \dots, k\}$ Let the distance between any two frames *KFj* and *Fi* be $d(KFj, Fi)$. Define *di* for each frame *Fi* as:

$$di = min(d(KFj, Fi)), j = 1, 2, ..., k$$

Then the Semi-Hausdorff distance between *S* and *R* is given as:

$$d_{sh} = max(d_i), i = 1.2.3 \dots \dots N$$

The fidelity measure is defined as in (7):

$$fidelity = 1 - \frac{d_{sh}}{\max(\max(d_{ij}))} \qquad (7)$$

where d_{ij} denotes the dissimilarity matrix of the shot set *S*. The bigger the fidelity is, the more accurate the global scan of key frames over the original video is.

Informativeness: To evaluate the informativeness, human observers are required to watch both the source and the automatic summary, and evaluate the percentage of information of original source video contained in the summary. The values given by the observers for particular pair of video and summary are indicative of amount of information conveyed by summary from source video to individual observer. The results of experimentation on 5 randomly selected videos are presented in Table I. Finally the average of these values is taken and the same is depicted in the Table I.

The proposed algorithm is different from other clustering based algorithms for key frame extraction that pick up some fixed frames from cluster, for example Centroid and frames near to it. The proposed method is extension to regular cluster based methods. Instead of picking some key frames from cluster on sampling bases the algorithm models each cluster into undirected weighted graph and selects key frames depending on frames connectivity with rest of the frames in the cluster in terms of eccentricity. The eccentricity of node representing frame is used as measure to estimate its connectivity with rest of the graph. All those frames represented by nodes very near to maximum eccentricity value are considered key frames. The nodes with higher value of eccentricity always exhibit dissimilarity from similar frames clustered in to segment. The example cluster and key frames selected from it are shown in the Fig. 5.

Figure 5(a). Example cluster containing 91 frames formed using fuzzy c means clustering.

Figure 5(b). Set of 09 key frames selected by proposed algorithm from cluster of 91 frames.

Figure 5(c). Set of 07 key frames selected from cluster of 60 frames by proposed algorithm.

TABLE I. PERFORMANCE OF PROPOSED ALGORITHM

Video	Number of Frames	Number of Key frames	Compression Ratio in %	Average Fidelity	Informativeness
1	500	52	90	93%	90%
2	1342	128	91	94%	92%
3	1300	120	91	95%	90%
4	2600	202	92	96%	95%
5	800	70	91	95%	95%
6	1000	85	91	94%	93%
Average			91	94.5%	92.5%

V. CONCLUSION

In this paper a key frame extraction algorithm for sports video summarization is presented, It selects the variable number of key frames from undirected weighted graph modeled from frames clustered using Fuzzy C Means clustering. The entropy features are used to represent video frames in to feature space. The proposed key frame selection algorithm picks the key frames based on Euclidean distance of frame feature from the rest of the frames in the cluster modeled as undirected graph. All those frames which are on diameter of graph are picked as key frames. The results have shown that the algorithm is able to summarize the video capturing all salient events in the video sequence. The compression ratio which is ratio of number of key frames utilized to build the summary to total number of frames in video for the algorithm is 91%. The value of performance parameter reveals that the algorithm can preserve over all information while compressing video by over 91%. The average fidelity of 94.5% indicates the comprehensive coverage of video by selected key frames. The informativeness of 92.5 is measure of information the summary conveys to individual users. The results in the form of performance parameters reveal that FCM is suitable for temporal segmentation of video and suitability of graph theoretical approach with graph eccentricity for key frame extraction for video summary generation, is brought out by the results.

REFERENCES

[1] J. S. Boreczky and L. A. Rowe, "Comparison of video shot boundary detection techniques," in *Proc. Int. Conf. Storage Retr. Still Image Video Databases*, 1996, pp. 170-179.

[2] S. Behzard and D. S. Gibbon, "Automatic generation of pictorial transcripts of video programs," in *Proc. SPIE Multimedia Computer Networking*, 1995, pp. 512-518.

[3] C. Panagiotakis, A. Doulamis, and G. Tziritas, "Equivalent key frames selection based on iso-content principles," *IEEE Trans. on Circuits Systems for Video Technology*, vol. 19, pp. 447-451, 2008

[4] G. S. Zhao, "A novel approach for shot boundary detection and key frames extraction," in *Proc. International Conference on Multi Media and Information Technology*, 2008.

[5] Y. Chen, Y. Deng, Y. Guo, and W. Wang, "A temporal video segmentation and summary generation method based on shots' abrupt and gradual transition boundary detecting," in *Proc. Second International Conference on Communication Software and Networks*, 2010.

[6] Y. Xiao and L. Xia, "Key frame extraction based on connectivity clustering," in *Proc. Second International Workshop on Education Technology and Computer Science*, 2010, pp. 174-177.

[7] A. M. Ferman and A. M. Tekalp, "Efficient filtering and clustering methods for temporal video representation and visual summarization," *J. Visual Com. & Image Rep.*, vol. 9, pp. 336-351, 1998.

[8] I. Koprinska, J. Clark, and S. Carrato, "Video GCS - A clustering-based system for video summarization and browsing," in *Proc. 6th COST 276 Workshop on information and Knowledge Management for Integrated Media Communication*, Thessaloniki, Greece, 2004.

[9] A. Girgensohn and J. Boreczky, "Time-Constrained keyframe selection technique," *Multimedia Tools & Appl.*, vol. 11, pp. 347-358, 2000.

[10] Z. Wang, K. Yu, S. Mei, M. He, and D. Feng, "Video summarization with global and local features," in *Proc. IEEE International Conference on Multimedia and Expo Workshops*, July 2012, pp. 570-575.

[11] R. Pan, Y. Tian, and Z. Wang, "Key-Frame extraction based on clustering," in *Proc. IEEE International Conference on Progress in Informatics and Computing*, 2010, pp. 867-871.

[12] J. Yuan, H. Wang, and L. Xiao, "A formal study of shot boundary detection," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 168-186, 2007.

[13] U. Sakarya and Z. Telatar, "A graph-based approach for video scene detection," in *Proc. 16 IEEE Conference in Signal Processing, Communication and Applications Conference*, 2008.

[14] S. Lu, I. King, and M. R. Lyu, "Key frame extraction in video sequences: a vantage points approach," in *Proc. IEEE 9th Workshop on Multimedia Signal Processing*, 2007, pp. 434-437.

[15] S. Lu and I. King, "Video summarization by spatial-temporal graph optimization," in *Proc. International Symposium on Circuits and Systems*, May 2004.

[16] Y. Peng and C. W. Ng, "Hot event detection and summarization by graph modeling and matching," *Image and Video Retrieval Lecture Notes in Computer Science*, vol. 3568, pp. 257-266, 2005.

[17] Y. Fu, Y. Guo, Y. Zhu, F. Liu, C. Song, and Z. H. Zhou, "Multi-View video summarization," *IEEE Transactions on Multimedia*, vol. 12, no. 7, pp. 717-729, 2010.

[18] Videosourse. [Online]. Available: www.soccer.net/koolwap.in/videos/soccer_mp4.htm

[19] The most recent version of the codec source code is available by ftp. [Online]. Available: ftp://ftp.mpeg.org/pub/mpeg/mssg/mpeg2vid.zip

Shanmukhapp Angadi is a Professor in the PG Centre, Department of Computer Science, Visvesvaraya Technological University, Belgaum, India. He earned Bachelor Degree in Electronics and Communication Engineering from Karnataka University, Master Degree in Computer Engineering from the Sri Jayachamarajendra College of Engineering, Mysore and PhD degree in Computer Science from the Department of Studies in Computer Science University of Mysore India. His research areas include image processing Pattern Recognition, Character Recognition, Fuzzy Systems, Neural Networks and Genetic Algorithms, Optimization and Graph Theoretic techniques, Embedded Systems, Intelligent Systems, Web Technology and Internet of things. He is a co-author of a book on C-programming language. He is life member of professional bodies like IEEE, ISTE and IETE.

Vilas Naik received BE (Electronics and Communication) from Karnataka University Dharwad and Master of Engineering in Computer Technology from Shri Guru Govind Singh College of Engineering Nanded under Sri Ramanand Teerth Marathawada University Nanded. He is currently a research scholar registered to Visvesvaraya Technological University, Belgaum in the area of Image and Video processing working on the issues of Multimodal Video summarization and selection. Currently he is working as Assistant Professor in the Department of Computer science and Engineering, Basaveshwar Engineering College Bagalkot. His subjects of interest are Image and Video processing, Data Communications and Computer Networks, Computer Architectures and Multimedia computation and Communication.