International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 3 | ISSN : 2456-3307

Survey on Summarization of Multiple User-Generated Videos

Binju Bentex*[1], Shandry K . K[2]
[1]PG Student, Department of Computer Science, College Of Engineering, Kidangoor, Kottayam, Kerala, India
[2]Assist. Professor, Department of Computer Science, College Of Engineering, Kidangoor, Kottayam, Kerala, India

ABSTRACT

The rapid developments in camera technology and mobile devices bring a flourish of user-generated videos with rich metadata, which can be great information resources for prospective tourists to preview a place of interest. In a video retrieval system, a simple query will return many videos. To provide users with a convenient way to explore videos, generate summarization from multiple user generated videos, which is composed of segments from different input videos. Apparently, summarizing a video is an important process. A video nothing but a synchronous sequence of a number of frames, each frame being a 2-D image. So the basic unit in a video is a frame. The video can also be thought of as a collection of many scenes, where a scene is a collection of shots that have the same context.

Keywords: Video Summarization, Multimedia, Indexing, Feature Extraction, Static Video Summarization, Dynamic Video Summarization.

I. INTRODUCTION

In recent years, because of the rapid growth in multimedia information, the advance in internet communication and digital video technologies, multimedia information indexing and retrieval has become more and more important and lots of research efforts have been devoted to the video retrieval and video analysis based on audio or visual features. This analysis shows that, when developing retrieval applications and video indexing, we first have to consider the issue of structuring the huge and rich amount of heterogeneous information related to video content. In addition, to retrieve information from the audio or visual content is a very challenging since it requires the extraction of high-level semantic information from low-level audio or visual data. Video summarization is an important process that facilitates faster browsing of large video collections and also more efficient content indexing and access.

The user may not have enough time to watch the entire video. In such cases, the user may just want to view the abstract of the video instead of watching the whole video which provides more information about the occurrence of various incidents in the video. As the name implies, video summarization is a mechanism for generating a short summary of a video, which can either be a sequence of stationary images or moving images.

II. LITERATURE SURVEY

A. Summarizing Wearable Video

Reference[1], Personal experiences are usually maintained by such media as diaries, pictures and movies. For example, we memorize our daily experiences in a diary, and we use photos and movies to keep our special events such as travels etc. However, those existing media, so far, keeps only a small part of our life. Even though we make use of

photos and movies, we always miss the best moments because the camera is not always ready.

"We want to record our entire life by video" is the motivation of this research. They propose an approach to automatic structuring and summarization of wearable video. (Wearable video is our abbreviation of "video captured by a wearable camera".) In this approach, they make use of a wearable camera and a sensor of brain waves. The video is firstly structured by objective features of video, and the shots are rated by subjective measures based on brain waves.

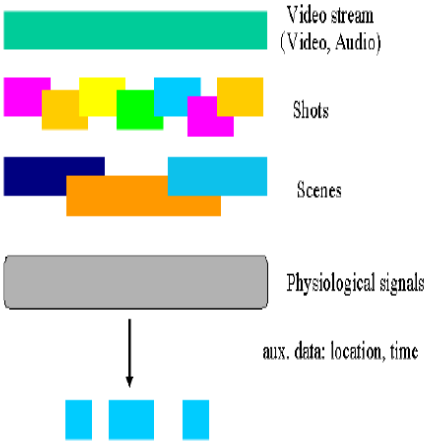

Figure1. The framework of summarization of wearable video.

Constantly recording will be feasible from hardware points of view. The most difficult issue is how to extract and reproduce what we want to see. The data size is too huge, then its automatic summarization is the most critical. The framework of the approach is automatic summarization illustrated in figure 1.
Two major steps:

1) The video stream is analyzed and segmented based on the objective visual features,
2) Video segment is evaluated based on the subjective signals.

B. A Study on Video Summarization Techniques

Reference [2] covers study of various techniques for key frames based video summarization. There have been tremendous needs of video processing applications to deal with abundantly available & accessible videos. One of the research areas of interest is Video Summarization that aims creating summary of video to enable a quick browsing of a collection of large video database.

Video summarization is a mechanism for generating a short summary of a video, which can either be a sequence of stationary images (key frames) or moving images (video skims).

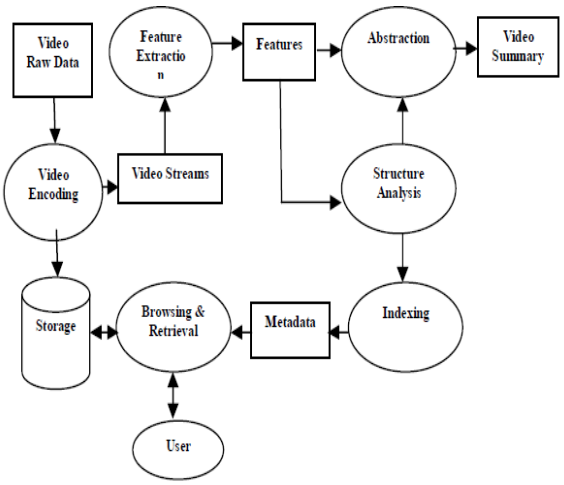

Figure2. General application of the video analysis and indexing tasks

Video can be summarized by two different ways which are as follows.

1) Key Frame Based Video Summarization:

These are also called representative frames, R-frames, still-image abstracts or static storyboard, and a set consists of a collection of salient images extracted from the underlying video source

2)Video Skim Based Video Summarization:

This is also called a moving-image abstract, moving story board, or summary sequence. The original video is segmented into various parts which is a video clip with shorter duration. Each segment is joined by either a cut or a gradual effect. The trailer of movie is the best example for video skimming.

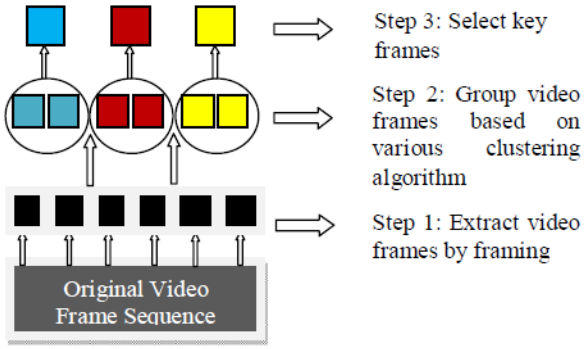

Figure 3. Key frames based video summarization

key frames based video summarization works on frames so first step is to extract frames from original video frame sequence. In step two extracted video frames are cluster that have redundant content obviating the need for shot detection. Selection of key frames is proceeding in step three. The entire procedure is shown in figure 3.

C. Automatic Soccer Video Analysis and Summarization

Reference[3] propose a fully automatic and computationally efficient framework for analysis and summarization of soccer videos using cinematic and object-based features. The proposed framework includes some novel low-level soccer video processing algorithms as well as some higher-level algorithms.

Semantic analysis of sports video generally involves use of cinematic and object-based features. Cinematic features refer to those that result from common video composition and production rules, such as shot types and replays. Objects are described by their spatial, e.g. color, texture, and shape, and spatio-temporal features, such as object motions and interactions. Object-based features enable high-level domain analysis, but their extraction may be computationally costly for real-time implementation. Cinematic features, on the other hand, offer a good trade-off between the computational requirements and the resulting semantics.

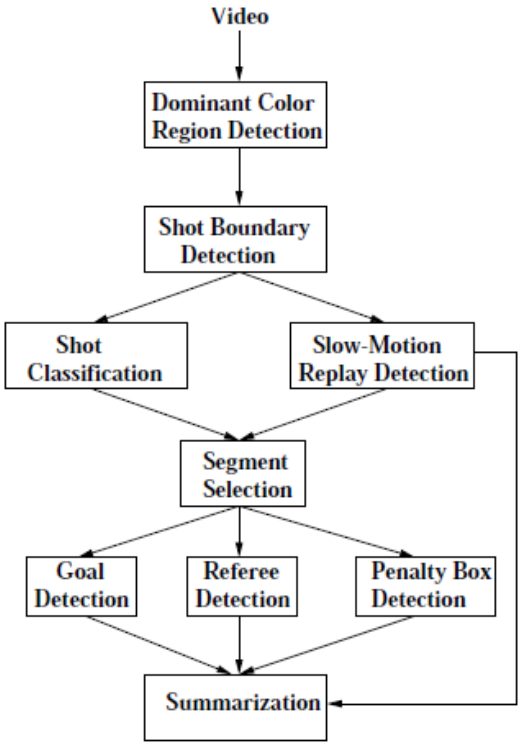

Figure 4. The flowchart of the system

A flowchart of the proposed framework is shown in Figure 4. The main contributions are:

1) Propose new dominant color region and shot boundary detection algorithms that are robust to variations in the dominant color.
2) Propose two novel features for shot classification in soccer video. They provide robustness to variations in cinematic features, which is due to slightly different cinematic styles used by different production crews.
3) Introduce new algorithms for automatic detection of i) goal events, ii) referee, and iii) penalty box in soccer videos.
4) Finally, propose an efficient and effective framework for soccer video analysis and summarization that combines these algorithms in a scalable fashion.

D. Video Summarization: Techniques and Applications

Video summarization has been proposed to improve faster browsing of large video collections and more efficient content indexing and access. Reference[4], focus on approaches to video summarization. The

video summaries can be generated in many different forms. However, two fundamentals ways to generate summaries are static and dynamic.

In one hand, static video summary represents a video sequence in a static imagery form (one or more selected representative frames from the original video, or a synthesized image generated from the selected keyframes). According to different sampling mechanisms, a set of keyframes are extracted from shots of the original video. Then, the selected keyframes are arranged or blended in a two-dimensional space. On the other hand, dynamic summarization consists in selecting the most relevant small dynamic portions (video skims) of audio and video in order to generate the video summary.

1) Static Video Summarization:

Keyframe extraction is fundamental process in video content management. It involves selecting one or multiple frames that will represent the content of the video and used for generating video summaries. Fig. 5 shows hierarchical structure in a video sequence in the extraction of such keyframes.

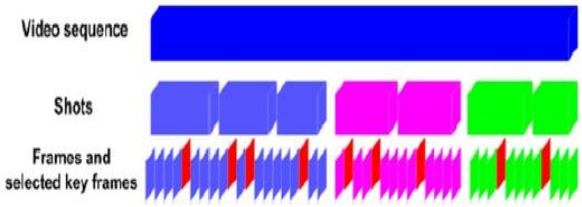

Figure 5. Hierarchical structure of a video sequence

2) Dynamic video Summarization

The fundamental idea of video skim which is a short video composed of informative scenes from the original video presented to the user to be able to receive an abstract of the video story, but in video format. For dynamic summarization (skimming), most mechanisms extract and segment video clips from the original video. Compared with static storyboard summary, there are relatively few works being addressed for dynamic video skimming. Techniques for dynamic video skimming include applying SVD (Singular Value Decomposition), motion model and semantic analysis. Most techniques are based mainly on visual information and some others approaches where audio and linguistic information are also incorporated in order to derive semantic meaning.

E. Video Summarization Based on Balanced AV-MMR

Among the techniques of video processing, video summarization is a promising approach to process the multimedia content. Reference[5], present a novel summarization algorithm, Balanced Audio Video Maximal Marginal Relevance (Balanced AV-MMR or BAV-MMR), for multi-video summarization based on both audio and visual information. Balanced AVMMR exploits the balance between audio information and visual information, and the balance of temporal information in different videos. AV-MMR is a simple extension of Video- MMR, and does not consider the characteristics of audio and video track. Therefore, in this paper propose a novel algorithm, Balanced AV-MMR, to improve AVMMR by:

1) Considering the balance between audio information and visual information in a short time
2) Analyzing and using the influence of audio genres
3) Exploiting audio changes from one genre to another
4) Analyzing and utilizing the information brought by the face
5) Using the temporal distance of video frames in a set
6) Finally designing a novel mechanism to combine these features

Assume that in a short time the audio attracted more attention from the user, the user would pay less attention to video content and vice versa, because the attention of a person in a short time is limited. In an audio segment, the duration is usually short. Therefore, there is a balance between audio

information and visual information in an audio segment. Consequently they give the name “balance”. Balanced AV-MMR exploits the information from audio genre, the face and the time to improve the balance information and similarities of frames in semantic level.

III. CONCLUSION

This paper presents a literature survey on various techniques for video summarization. There are different methods through which we can summarize videos in an efficient manner. Regardless of the methods used, that are static or dynamic forms, the evaluation process showed that the techniques proposed produces video summaries of high visual quality, and some approaches are suitable for real-time video processing. There is not any best technique for abstracting a video sequence, as video abstraction is still in the research phase largely. Also, practical applications are still limited. So, there is a scope of research in many fields such as personalized videos, consumer videos, and movie videos as well.

IV. REFERENCES

[1]. Kiyoharu Aizawa, Ken-Ichiro Ishijima, Makoto Shiina, "Summarizing Wearable Video," DOI: 10.1109/ICIP.2001.958135 • Source: IEEE Xplore.

[2]. RaviKansagara, DarshakThakore, MahaswetaJoshi," A Study on Video Summarization Techniques" ISSN(Online): 2320-9801, Vol. 2, Issue 2, February 2014.

[3]. Ahmet Ekin, A. Murat Tekalp, and Rajiv Mehrotra, "Automatic Soccer Video Analysis and Summarization" IIS-9820721.

[4]. Zaynab El khattabi, Youness Tabii, Abdelhamid Benkaddour," Video Summarization: Techniques and Applications" International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:9, No:4, 2015.

[5]. Yingbo Li and Bernard Merialdo," Video Summarization Based on Balanced AV-MMR" EURECOM, Sophia Antipolis, France {Yingbo.Li, Bernard.Merialdo}@eurecom.fr.

[6]. Ying Zhang and Roger Zimmermann," Efficient Summarization From Multiple Georeferenced User-Generated Videos" IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 18, NO. 3, MARCH 2016.