Efficient Media Digital Library of Summarized Video based on Scalable Video Coding for H.264 (MDLSS): Design and Implementation

Hesham Farouk, Mayada Khairy
Dept. of Computers and Systems
Electronics Research Institute, ERI
Cairo, Egypt
{hesham, mayada}@eri.sci.eg

Kamal A. ElDahshan, Amr Abozeid
Dept. of Math., CS Division
Faculty of Science, Al-Azhar University
Cairo, Egypt
{dahshan, amrapozaid}@gmail.com

Alaa Hamdy, Amr Elsayed
Dept. Telecomm. And Electronics.
Faculty of Engineering, Helwan Univ.
Cairo, Egypt
{alaa.hamdy, amr18188}@gmail.com

***Abstract*—With the fast progress of wireless networks bandwidth and mobile devices, large scale digital video library systems are growing rapidly. However, the huge increasing of content and the data intensive nature of video make the management and browsing of video collections, as well as their search and retrieval, increasingly difficult.**

The need of having a media digital library is essential these days with intelligent tools for indexing the video with allocating the suitable metadata that describe the content of such videos and appropriate tools for retrieving the archived video with fast techniques. These will be achieved in 3 steps, working on the stream coding with multi bit rates and methods of handling, representing the video with summarized stream carrying the same information of the full stream and deriving a media digital library for indexing and retrieval process. The first step, stream handling, will address implementing scalable video techniques which set the bit rate according to the required application and the delivery devices. Scalable video coding offers a solution for meeting such heterogeneous requirements. The second step, video summarization, plays an important role in this context; it makes navigation easier and provides the user with a quick idea about the content. Another issue is that the same video content can be accessed from a wide variety of terminal devices which differ with respect to bandwidth limitation, decoding complexity, power constraints and screen size. The third step is implementing a media digital library for storing the code and/ or the summarized video based on Media Asset management system.

The main contribution of this paper is to explore the use of scalable video coding and video summarization techniques to enhancing a digital video library and the integration between these 3 modules.

Keywords-Video processing; Scalable video coding; Video summarization; Digital library.

I. Introduction

Nowadays, multimedia communications have significantly facilitated and enriched people's daily life. People have witnessed the fast development of various wireless multimedia applications, such as video content distribution (e.g., YouTube) and live video communications (e.g., Skype, Microsoft Network (MSN) messenger, etc.). As a result, the volume of video data is rapidly increasing; over 6 billion hours of video are watched each month on YouTube and more than 100 hours of video are uploaded to YouTube every minute [1]. Moreover, the increased popularity of mobile devices and wireless networks and their ubiquitous use for video recording and streaming leads to a dramatic increase traffic of videos traffic on such devices. Cisco stated that "Mobile Video will generate over 69 Percent of Mobile Data Traffic by 2018". Mobile makes up almost 40% of YouTube's global watch time [2].

A. Problem identification

Video delivery, especially, via mobile wireless networks, faces diverse challenges, including limited bandwidth, dynamic network conditions with low stability, a variety of relay equipment, different terminal decoding speeds, various display screen resolutions, limited battery capacity, etc. [3]. Therefore, the video coding system must encode the video sequence in different frame sizes, frame rates, and bit rates to meet such heterogeneous demands [4]. Another problem is that, the increasing amount of content and the intensive nature of video data make very difficult the management and browsing of stored video collections, as well as their search and retrieval [5].

B. Need for the system

Video is increasingly becoming one of the most pervasive technologies in terms of everyday usage, both for entertainment and in the enterprise environments. Mobile video is responsible for a majority of the growth seen in mobile broadband data volume. The demand for better video services (streaming, storing, retrieving, browsing, etc.) for mobile devices is a key challenge. The proposed system aims to solve some of these challenges.

The Middle East region, especially Egypt, has a huge amount of audio and visual heritage. There is a real need for developing a system to help in archiving digitized content and operate these rich video libraries with up to date video technologies, such as scalable video and video summarization.

This paper is organized as follows: Section II introduces the related work. Section III explains the research approach and methodology. Section IV presents the proposed system, Media Digital Library of Summarized Video based on Scalable Video Coding, MDLSS, design architecture and discusses its modules. Finally, in Section V, we conclude the paper and hint to future work.

 ISBN: 978-1-61208-408-4

II. REALTED WORK

This section will look at the state of the art of Module 1 and 2 from the proposed system (Module 3 will be the host of the deliverables from Module 1 and 2).

A. Scalable video coding (SVC)

A number of fast algorithms of fast mode decision schemes have been proposed for SVC [4][6][7][8]. For spatial scalability, a fast mode decision algorithm based on distribution relationship between the base layer and enhancement layers is used [4][9].

The scheme in Li et al.[8] represents the mode distribution relationship between base layer and enhancement layer. It is employed to reduce the candidate mode set at enhancement layers. The experimental results show that the proposed scheme provides significant reduction in computational complexity without any noticeable coding loss. Kim et al. [7] proposed a fast mode spatial, temporal, quality, and combined scalability. This algorithm is based on Coded Block Pattern (CBP) of 16×16 mode in current frame and CBP of coded best mode in selected reference frame.

B. Video summarization

Video summarization is the process of extracting the most important information and reduces the amount of redundant information from the video. The input video must be well processed in order to extract only the most useful contents [10]. But, to generate a good video summary, a full understanding of the video is required, which is still a research challenge.

In literature, many video summarization approaches have been introduced [11]. Farouk et al. [12] the authors have analyzed and compared between various techniques of mobile video summarization according to criteria. Some examples of these criteria include: content structure, final summary representation, features based, targeted devices, summarization speed, summarization purposes, adaptability and complexity. These criteria have been extracted from the reading and the analysis of literature and works in the video summarization field. Here is a summary of the observations derived from the current literature:

- The final representation of video summarization techniques is a static summary.
- Color feature was widely used in literature to summarize the video contents
- In recent years (from 2010), there is a new research direction to summarize home videos originating from the mobile camera as in [13][14].
- The goal is to reduce the technique complexity and use it to generate an online video summary.
- The purpose of mobile video summarization approaches is usually for browsing or/and streaming to another device.

III. RESEARCH APPROACH AND METHODOLOGY

The proposed research in this paper will address 3 research tracks:

1. The scalable video coding will target to enhance the bit rate transmission for spatial video streams.
2. The video summarization will target to develop an effective summarization approach comparable to the state of the art approaches.
3. The Digital library technology will work on developing a smart indexing and retrieving technique based on Media Asset Management concepts already installed in Electronics Research Institute, ERI.
4. The Integration phase, which will work on implementing the work flow described in Figure 1.

A. The first motivation of this system

Today, there is a wide range of different devices available for viewing video content, including smartphones, tablets, laptops and televisions. Every client's requirement differs with respect to bandwidth limitation, decoding complexity, power constraints and screen size. Scalable Video Coding (SVC) offers a solution for meeting such heterogeneous requirements [15].

A video bit stream is called scalable if a part of the stream can be removed in such a way that the resulting bit stream is still decodable. The three types of scalability are [16]:

1. Temporal (frame rate) scalability: the motion compensation dependencies are structured so that complete pictures (i.e., their associated packets) can be dropped from the bit stream. Temporal scalability is already enabled by H.264/MPEG-4 Advanced Video Coding (AVC) [17]. The SVC has only provided supplemental enhancement information to improve its usage.
2. Spatial (picture size) scalability: video is coded at multiple spatial resolutions. The data and decoded samples of lower resolutions can be used to predict data or samples of higher resolutions in order to reduce the bit rate to code the higher resolutions.
3. Signal to Noise Ratio (SNR)/Quality/Fidelity scalability: video is coded at a single spatial resolution, but at different qualities. The data and decoded samples of lower qualities can be used to predict data or samples of higher qualities in order to reduce the bit rate to code the higher qualities.

This work is represented as Module 1 of the proposed work given in Figure 1.

B. The second motivation of this system

The content of video may be huge and crowded with much redundant information, so that it often takes a long time to browse the content from the beginning to the end. Also, the user may not have sufficient time to watch the entire video or the video content, as a whole, may not be of interest to the user. In such cases, the user may just want to view the summary of the video instead of watching the whole video [18].

Video summarization is a mechanism for generating compact representation of a video sequence, which includes only the important parts in the original video [19]. Video summarization is useful when a system is operating under tight constraints (e.g., limited bandwidth, watching time, or memory size). For example, in surveillance applications the video may be recorded nearly for 24 hours per day, a summary version of the original video may be useful to watch the important events only in such case. Also, video summarization is useful when we need to transmit an important video segment to another device in real time [20]. Video summarization techniques target different domains of video data, such as sports, news, movies, documentaries, e-learning, surveillance, home videos, etc., and discuss various assumptions and viewpoints to produce an optimal or good video summary [18].

This work is represented as module 2 of the proposed work given in Figure 1.

There are two fundamental types of video summaries [21]: static video summary (also called representative frames, still-image abstracts or static storyboard) and dynamic video skimming (also called video skim, moving image abstract or moving storyboard). The static video summary is a collection of video frames extracted from the original video. The dynamic video summary is a set of short video clips, joined in a sequence, and played as a short video clip. Usually, from the users viewpoint, a dynamic video summary may provide a better choice since it contains both audio and motion information that makes the summarization more interesting and natural, while static video summary may provide a glance of video contents in a more concise way. In addition, once video frames are extracted, there are further possibilities of organizing them for browsing and retrieving purposes [22].

C. *The third motivation of this system*

The Digital library for media file under Media Asset Management (MAM) system will be the main storage system for the processed video by Module 1 and Module 2 and will be based on the MAM purchased by ERI through the 'EQUIPME' initiative issued by research academy 2 years ago [23].

IV. THE PROPOSED SYSTEM

A. *The proposed system architecture*

The architecture of MDLSS consists of three modules; scalable video coding, video summarization and digital library, as shown in Figure 1. MDLSS aims at providing better video services (streaming, storing, retrieving and browsing) for mobile devices which increase the interactions and activities between users and digital libraries. In other words, the main goal of MDLSS is to explore the use of SVC and video summarization techniques for enhancing digital video library. This goal can be further specified in the following.

- Design and develop SVC algorithm to meet the requirements of applications and devices heterogeneities.
- Design and develop an automatic video summarization algorithm, which engages in providing concise and informative video summaries to help in browsing and managing video files efficiently.

B. *The system design methods and procedures*

For Module 1:

The scalable video coding has three types, namely, spatial scalability, quality scalability and temporal scalability. This paper focuses on spatial scalability.

The output of SVC stage (Module 1 in the proposed system) is one bit stream of compressed video in H.264/SVC format. This bit stream has two levels of spatial scalability. The first one is the base-layer Quarter Common Intermediate Format, QCIF, and the second one is the enhancement-layer Common Intermediate Format, CIF. There are seven modes for inter prediction (SKIP, 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, and 4x4) and there are two intra prediction modes (INTRA 4x4 and INTRA 16x16) based on the minimum rate cost equation.

General adapted methods for SVC:

1. Determine number of layers for scalable video.
2. Determine number of bitrates available.
3. Analyze the video stream.
4. Select the type of scalability according to 3 steps before.
5. Implement the scalable video types according to previous steps.

The proposed algorithm for Module 1 achieved saving in time of encoding with up to 50%, and saving in time of decoding with up to 30% compared to Joint Scalable Video Coding, JSVC. This is done by selecting best macroblock mode.

For Module 2:

A general adapted method for video summarization module is shown in Figure 2. Each step is described as follows:

1. Frames sampling

The first step towards automatic video summarization is splitting the video stream into a set of meaningful and manageable basic elements (e.g., shots, frames) that are used as basic elements for summarization. Most of existing methods for automatic video summarization have focused on splitting the video stream into frames. The video sequence is decoded and each frame is extracted and treated separately [13].

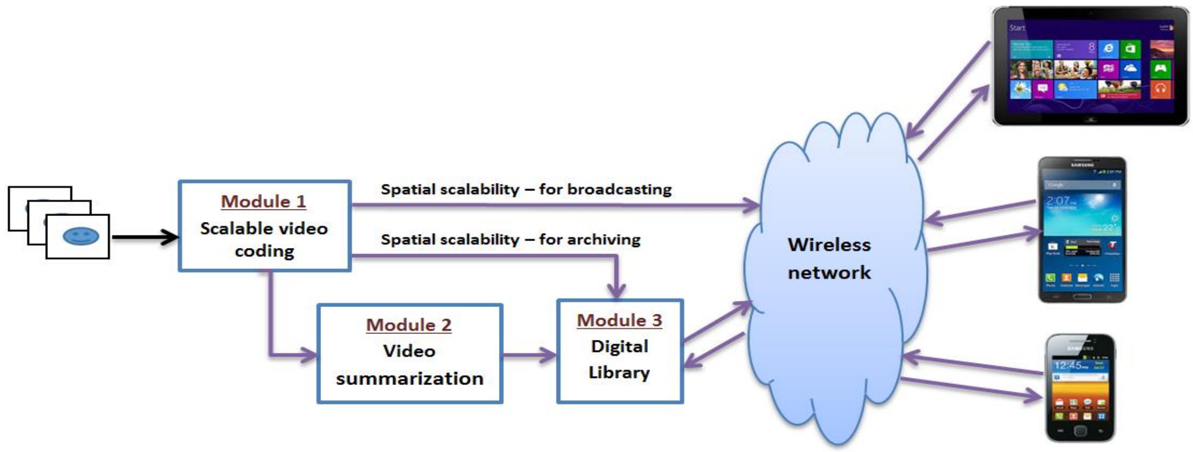

Figure 1. MDLSS architecture

2. Feature extraction

Digital video contains many features, like color, motion, voice, etc. Color feature is considered an important aspect of video. That is why it has been used quite often for video summarization. Color based summarization techniques are very simple and easy to use. However, color-based methods accuracy is not reliable, as color based techniques may consider noise as part of the summary [12].

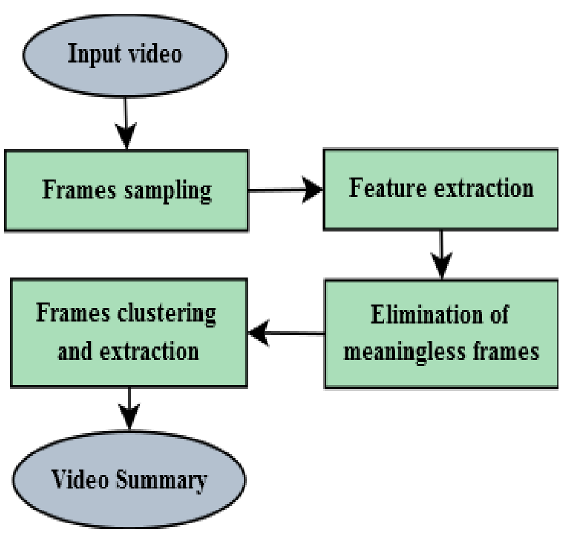

Figure 2. Flowchart of video summarization module

3. Elimination of meaningless frames

The goal of this step is to avoid possible meaningless frames in a video summary. It has been generally observed that a video, usually, has some meaningless frames, such as totally black frames, totally white frames (a monochromatic frame) and faded frames [22].

4. Frames clustering and extraction

The goal of this step is to group similar video frames together and to select a representative frame per each group, to produce the video summary. The effectiveness of grouping similar frames depends on the suitable choice of a similarity metric used for comparing two frames [24].

Video summarization is has been a very active research field in recent years due to its important role in many video services (e.g., browsing, indexing and streaming). The reader can find a comprehensive review of video summarization techniques in [11][25]. Moreover, Farouk [12] produced an analysis and comparative study between various techniques proposed in literature for the summarization of video content, which can be useful for mobile applications.

For Module 3:

In this step, we will manage video storage after editing and applying Meta data through the Media Asset Management we already have in our Digital Signal Processing Lab in Electronics Research Institute.

V. A CASE STUDY OF VIDEO SUMMARIZATION

In this case study, we show how can generate a static video summary from the SVC originated from Module 1.

A. Frames sampling

The input of video summarization (Module 2) is a QCIF video in H.264/SVC format. This video is partially decoded to select sample frames from it based on a predefined sampling rate. In this case study, the sampling rate is set to be two frames per second.

B. Feature extraction

In this case study, a color histogram is applied to describe the visual content of video. There are two key issues in applying the color histogram technique which are the selection of a suitable color space and the quantization of that color space [21]. Since the Human Visual System (HVS) is more sensitive to luminance than color [17], we choose the YUV 4:2:0 color space and compute histogram only for Y (luminance) component and discard the chrominance

 ISBN: 978-1-61208-408-4

components U and V. The quantization of the color histogram is set to 16 color bins that are normally used for hue component aiming at reducing the amount of data significantly without losing the important information.

C. Elimination of meaningless frames

In this step, for each candidates frame the standard deviation of pixels is computed, as in [21][26]. If the standard deviation is very low (close to zero) then this frame is considered as a meaningless frame and is discarded. We apply this step before the clustering step (as preprocessing step) like [21], which saves the computation cost.

D. Frames clustering and extraction

In this case study, we use the adopted Zero-mean Normalized Cross Correlation (ZNCC) [22] as the similarity metric between two frames. ZNCC is widely used in template matching, motion analysis, stereo vision, and industrial inspection. Let H_{t_1} and H_{t_2} be the color histograms extracted from the video frames F_{t_1} and F_{t_2} taken at the times $t_1$andt_2, respectively. The ZNCC between $H_{t_1}$and H_{t_2} is defined by (1).

$$ZNCC(H_{t_1}, H_{t_2}) = \frac{\sum_i\left(\left(H_{t_1}^i - \bar{H}_{t_1}\right) \times \left(H_{t_2}^i - \bar{H}_{t_2}\right)\right)}{\sqrt{\sum_i\left(H_{t_1}^i - \bar{H}_{t_1}\right)^2 \times \sum_i\left(H_{t_2}^i - \bar{H}_{t_2}\right)^2}}. \quad (1)$$

Where H^i be the ith bin of the color histogram H and $\bar{H}$ is the mean value of all entries of H. The ZNCC function returns a real value from -1 to 1. The value of -1 is returned for situations in which those histograms are not similar at all, and the value of +1 is returned for situations in which they are identical [22].

In order to group similar video frames together, we apply the cluster algorithm which is described as Algorithm 1.

Where ε is a threshold for the similarity between frames and through the experiment, we have found that the values of ε between 0.3 and 0.7 are best choices. Finally, the middle frame is selected from each cluster to form the video summary.

Algorithm 1: the cluster algorithm

Input: F_{t_k}, $k = 1,2,\dots,n$ // the set of candidates frames
Output: C_j, $j = 1,2,\dots,m$; $m < $ n // a set of
Start
1. Initialize j=1
2. Loop for each $k: 1 \rightarrow N$
3. If ($ZNCC\left(H_{t_k}, H_{t_{k+1}}\right) > \varepsilon$) then
4. Add $F_{t_k}, F_{t_{k+1}}$ to the cluster C_j // without duplicate
5. k=k+1
6. Else
7. Add F_{t_k} to the cluster C_j // without duplicate
8. k=k+1, j=j+1
9. End loop

End

.

E. Experiments and Results

The setup environment for testing the proposed modules is based on a network has bandwidth 100 Mb/s. We assume 70% of the bandwidth is dedicated for transferring videos.

1) Data set and testing device

This experiment is carried out on 4 videos from the standard data set available at the Video Summarization (VSUMM) web site [27]. Table I contains the descriptions of these videos. Each video format is MPEG-1 with resolution of 352×240 pixels, 30 frames per second and in color and with sound. Because of the input format to our approach is H.264/AVC, each video is firstly transcoded to match the input format.

TABLE I. THE TEST VIDEOS DESCRIPTION

#	Video name	Duration	#Frames	Genre
1	Exotic Terrane, segment 01 of 12	00:01:38	2,940	Documentary
2	A New Horizon, segment 05 of 13	00:01:59	3,561	Documentary
3	Senses And Sensitivity, Introduction to Lecture 3 presenter	00:02:32	4,566	Lecture
4	Digital Jewelry: Wearable Technology for Every Day Life	00:03:00	4,204	Educational

We implemented a prototype to test this case study based on Java platform. The JCodec library is used to partially decode the input video [28]. All the experiments were performed on a PC device equipped with an 8 GB of DDR3-memory and Intel Core i7 processor.

2) Evaluation method

The Mean Opinion Score (MOS) method proposed in [21][29] is used in this evaluation. In this method, the quality of the automatically generated summary is compared to the users (Mostly, five users) generated summary. Then, we compute the Accuracy Rate (AR) and Error Rate (ER) metrics as in (2) and (3) respectively.

$$AR = \frac{N_{TM}}{N_{US}}. \quad (2)$$

$$ER = \frac{N_{FM}}{N_{US}}. \quad (3)$$

Where N_{TM} denotes the total number of the frames that exists as key frame in both the user and the automatic

ISBN: 978-1-61208-408-4

summary. The symbol N_{FM} denotes the total number of the frames that is exists in the automatic summary and not exists in the user summary. Finally, N_{US} denotes the total frames number in the user summary. Both accuracy rate and error rate are complementary metrics and the highest quality of summary was achieved when AR = 1 and ER = 0.

3) Comparison with other techniques

In order to evaluate the quality of the proposed approach, we compare it with other static video summary approaches found in the literature. The compared approaches include VSUMM [21], Delaunay Triangulation (DT) [30], STIll and MOving (STIMO) video storyboard for the web scenario [31] and Open Video Project (OV) [32].

The comparative results are provided in Table II. Also, an example is shown in Figure 3. The results demonstrate that the proposed approach achieved an average AR of 0.84 and an average ER of 0.24 with respect to the users' generated summary. These results indicated that, the proposed approach has a balance between a high accuracy rate and a low error rate.

TABLE II. THE COMPARISON OF DIFFERENT TECHNIQUES

	OV		***DT***		***STIMO***		***VSUMM***		***Proposed***	
NO.	***AR***	***ER***	***AR***	***ER***	***AR***	***ER***	***AR***	***ER***	***AR***	***ER***
1	0.98	1.39	0.48	0.32	0.66	0.53	0.82	0.53	0.85	0.23
2	0.63	0.12	0.24	0.08	0.43	0.15	0.82	0.19	0.81	0.12
3	0.6	0.3	0.38	0.34	0.79	0.57	0.84	0.43	0.82	0.32
4	1	0.7	0.6	0.1	0.94	0.46	0.9	0	0.89	0.27
Average	**0.8025**	**0.6275**	**0.425**	**0.21**	**0.705**	**0.4275**	**0.845**	**0.2875**	**0.8425**	**0.235**

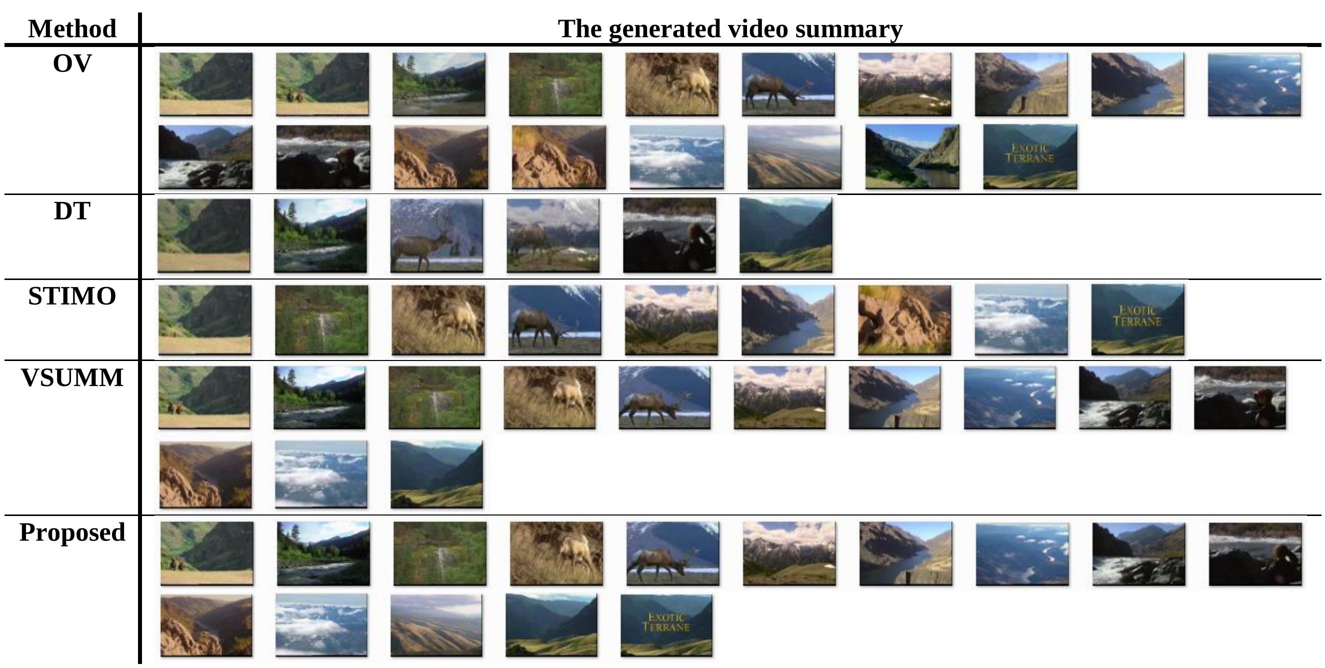

Figure 3. Comparison of video summary extraction for "Exotic Terrane, segment 01 of 12" video

VI. CONCLUSION AND FUTURE WORK

The SVC, as well as the video summarization, plays an important role in many video services. So, in this paper, we presented an efficient media digital library framework design of summarized video based on SVC for H.264 (MDLSS) with a test case study as a proof of concept. The proposed design will utilize the conjunction between SVC and video summarization techniques to enhance the digital video library.

 ISBN: 978-1-61208-408-4

In the future, we plan to do a full implementation to this proposed system (MDLSS). Our implementation activities will be organized as follows:

- Analyzing the system requirement for each module and for integration
- Developing a system prototype.
- Testing the system and updates.

REFERENCES

1. Youtube statistics. Available from: http://www.youtube.com/yt/press/statistics.html, [retrieved: 5, 2015].
2. V. N. I. (VNI). Cisco visual networking index: Global mobile data traffic forecast update, 2013–2018. Available from: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html, [retrieved: 5, 2015].
3. N. V. Uti and R. Fox, The challenges of compressing and streaming real time video originating from mobile devices, in Multimedia services and streaming for mobile devices: Challenges and innovation. 2011, IGI Global. p. 1-24.
4. P.-C. Wang, G.-L. Li, S.-F. Huang, M.-J. Chen, and S.-C. Lin,"Efficient mode decision algorithm based on spatial, temporal, and inter-layer rate-distortion correlation coefficients for scalable video coding," ETRI journal, vol. 32, 2010, pp. 577-587, doi:org/10.4218/etrij.10.0109.0622.
5. L. Herranz and J. M. Martínez,"Combining mpeg tools to generate video summaries adapted to the terminal and network," The Computer Journal, vol. 56, 2013, pp. 529-553, doi:10.1093/comjnl/bxs104.
6. L. Shen, Z. Liu, P. An, R. Ma, and Z. Zhang,"Fast mode decision for scalable video coding utilizing spatial and interlayer correlation," Journal of Electronic Imaging, vol. 19, 2010, pp. 033010-033010-8.
7. T.-J. Kim, J.-J. Yoo, J.-W. Hong, and J.-W. Suh,"Fast mode decision algorithm for scalable video coding based on luminance coded block pattern," Optical Engineering, vol. 52, 2013, doi:10.1117/1.OE.52.1.017401.
8. H. Li, Z. Li, C. Wen, and L.-P. Chau,"Fast mode decision for spatial scalable video coding," Proc. IEEE International Symposium on Circuits and Systems (ISCAS 2006). . IEEE, May 2006, pp. 4, doi:10.1109/ISCAS.2006.1693257.
9. H. Li, Z. Li, and C. Wen,"Fast mode decision algorithm for inter-frame coding in fully scalable video coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 16, 2006, pp. 889-895, doi:10.1109/TCSVT.2006.877404.
10. H. Karray, M. Ellouze, and A. Alimi, Indexing video summaries for quick video browsing, in Pervasive computing. 2010, Springer. p. 77-95.
11. M. Ajmal, M. H. Ashraf, M. Shakir, Y. Abbas, and F. A. Shah,"Video summarization: Techniques and classification," Computer Vision and Graphics, Springer, vol. 7594, 2012, pp. 1-13, doi:10.1007/978-3-642-33564-8_1.
12. H. Farouk, K. ElDahshan, and A. Abozeid,"The state of the art of video summarization for mobile devices: Review article," Graphics, Vision and Image Processing GVIP, vol. 14, 2014, pp. 37-50.
13. J. Niu, D. Huo, K. Wang, and C. Tong,"Real-time generation of personalized home video summaries on mobile devices," Neurocomputing, ScienceDirect, vol. 120, 2013, pp. 404-414, doi:10.1016/j.neucom.2012.06.056.
14. G. Abdollahian, C. M. Taskiran, Z. Pizlo, and E. J. Delp,"Camera motion-based analysis of user generated video," IEEE Transactions on Multimedia, vol. 12, 2010, pp. 28-41, doi:10.1109/TMM.2009.2036286.
15. M. Ransburg, et al.,"Scalable video coding impact on networks," Mobile Multimedia Communications, vol., 2012, pp. 571-581.
16. S. Ibrahim, A. H. Zahran, and M. H. Ismail,"Svc-dash-m: Scalable video coding dynamic adaptive streaming over http using multiple connections," Proc. 21st International Conference on Telecommunications (ICT). IEEE, May 2014, pp. 400-404, doi:10.1109/ICT.2014.6845147.
17. I. E. Richardson, The h. 264 advanced video compression standard. John Wiley & Sons, 2011,
18. B. T. Truong and S. Venkatesh,"Video abstraction: A systematic review and classification," ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP), vol. 3, 2007, pp. 37, doi:10.1145/1198302.1198305.
19. G. Guan, et al.,"A top-down approach for video summarization," ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 11, 2014, doi:10.1145/2632267.
20. L. Zhu, Z. Fan, and K. Aggelos K,"Joint video summarization and transmission adaptation for energy-efficient wireless video streaming," EURASIP Journal on Advances in Signal Processing, vol., 2008, doi:10.1155/2008/657032.
21. S. E. F. de Avila and A. P. B. Lopes,"Vsumm: A mechanism designed to produce static video summaries and a novel evaluation method," Pattern Recognition Letters, Elsevier vol. 32, 2011, pp. 56-68, doi:10.1016/j.patrec.2010.08.004.
22. J. Almeida, N. J. Leite, and R. d. S. Torres,"Online video summarization on compressed domain," Journal of Visual Communication and Image Representation, vol. 24, 2013, pp. 729-738, doi:10.1016/j.jvcir.2012.01.009.
23. Ncpower. Available from: http://www.norcom.de/en/features-ncpower, [retrieved: 5,2015].
24. S.-H. Ou, C.-H. Lee, V. S. Somayazulu, Y.-K. Chen, and S.-Y. Chien,"Low complexity on-line video summarization with gaussian mixture model based clustering," Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) IEEE, May 2014, pp. 1260-1264, doi:10.1109/ICASSP.2014.6853799.
25. R. Pal, A. Ghosh, and S. K. Pal, Video summarization and significance of content: A review, in Handbook on soft computing for video surveillance. 2012, CRC Press. p. 79-102.
26. N. Ejaz, T. B. Tariq, and S. W. Baik,"Adaptive key frame extraction for video summarization using an aggregation mechanism," Journal of Visual Communication and Image Representation, vol. 23, 2012, pp. 1031-1040.
27. Vsumm (video summarization). Available from: http://www.npdi.dcc.ufmg.br/VSUMM [retrieved: 4, 2015].
28. Jcodec. Available from: http://jcodec.org/, [retrieved: 5, 2015].
29. N. Ejaz, I. Mehmood, and S. W. Baik,"Feature aggregation based visual attention model for video summarization," Computers & Electrical Engineering, vol. 40, 2014, pp. 993-1005.
30. P. Mundur, Y. Rao, and Y. Yesha,"Keyframe-based video summarization using delaunay clustering," International Journal on Digital Libraries, vol. 6, 2006, pp. 219-232.
31. M. Furini, F. Geraci, M. Montangero, and M. Pellegrini,"Stimo: Still and moving video storyboard for the web scenario," Multimedia Tools and Applications, vol. 46, 2010, pp. 47-69.
32. The open video project. Available from: http://www.open-video.org/index.php, [retrieved: 5, 2015].

 ISBN: 978-1-61208-408-4