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

Hesham Farouk *, Kamal ElDahshan**, Amr Abozeid **, Mayada Khairy*

** Computers and Systems Dept., Electronics Research Institute, Cairo, Egypt.*

Hesham@eri.sci.eg, Mayada@eri.sci.eg

*** Dept. of Mathematics, Computer Science Division,*

Faculty of Science, Al-Azhar University, Cairo, Egypt.

Dahshan@gmail.com, Amrapozaid@gmail.com

***Abstract*— With the fast advancement 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 at the mean while tools for retrieving the archived video with fast techniques.**

These will be achieved across 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 be across 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, which 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 differs 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 innovation of this project 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, Key Frame Extraction, Video skimming, Home video, Mobile computing. Video indexing and retrieval**

I. Introduction

Nowadays, multimedia communications has 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, MSN, 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 dramatically increases traffic of videos 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, variety of relay equipment, different terminal decoding speeds, various display screen resolution, and 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 supply such heterogeneous demands [4]. Another problem is that, the increasing amount of content and the intensive nature of video data make the management and browsing of stored video collections, as well as their search and retrieval, this increases the system handling difficulties [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 and etc.) for mobile devices is a key challenge. The proposed system aims to solve some of these challenges.

This paper is organized as follows: Section II introduces the research approach and methodology. Section III presents the MDLSS design architecture and discusses its modules. Finally, in section IV we conclude the paper and suggest a future work.

II. Research Approach and Methodology

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 offers a solution for meeting such heterogeneous requirements [6].

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 scalabilities are [7, 8]:

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 AVC. 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. 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 Fig. 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 times 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 [9].

Video summarization is a mechanism for generating compact representation of a video sequence, which includes only the important parts in the original video [10]. 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 [11]. 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 [9].

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

There are two fundamental types of video summaries [12]: 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 more good 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 [13].

C. The Third motivation of this system

The Digital library for media file under Media Asset Management (MAM) system which 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 scientific research academy 2 years ago.

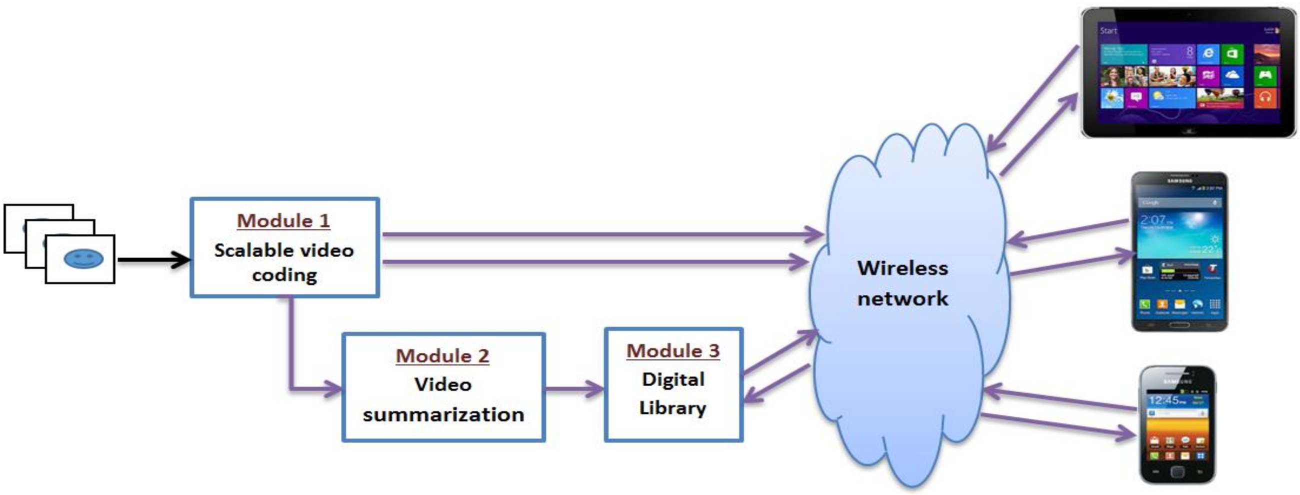

Fig.1 MDLSS architecture

III. 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 Fig.1. MDLSS aims to provide 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 scalable video coding and video summarization techniques to enhancing digital video library. This goal can be further specified in the following:-

- Design and develop Scalable Video Coding (SVC) algorithm to meet the requirements of applications and devices heterogeneities.
- Design and develop an automatic video summarization algorithm, which engage 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:
General adapted methods for SVC

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

For module 2:
A general adapted method for video summarization module is shown in Fig. 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 split the video stream into frames. The video sequence is decoded and each frame is extracted and treated separately [14].

2- Feature extraction

Digital video contains many features like color, motion, and voice etc. Color feature is considered an important aspect of video. That's why it has been used quite often for video summarization. Color based summarization techniques are very simple and easy to use. However, their accuracy is not reliable, as color based techniques may consider noise as part of the summary [15].

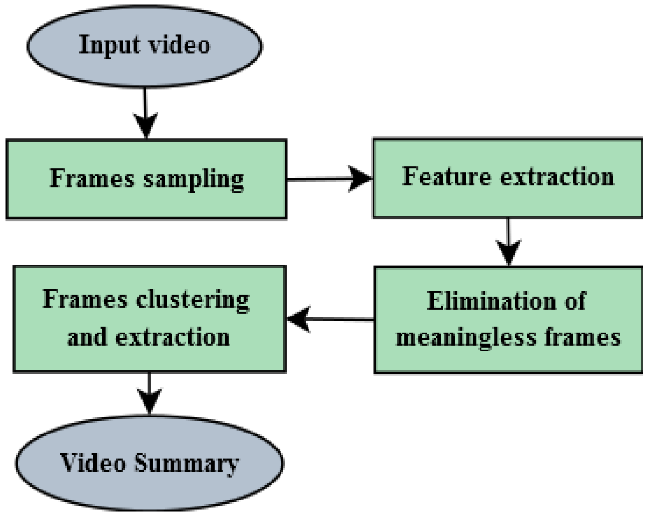

Fig. 2 Flowchart of video summarization module

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 [13].

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 [16].

Video summarization is hot research filed 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 [9, 17-19]. Moreover, the authors in [15] introduce 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.

IV. Conclusions and Future Work

The scalable video coding as well as video summarization plays an important role in many video services. So, in this paper we present an efficient media digital library design of summarized

ISBN: 978-1-61804-271-2

video based on scalable video coding for H.264 (MDLSS). The proposed design will utilize the conjunction between scalable video coding and video summarization techniques to enhance the digital video library.

In the future we arrange to develop this proposed project (MDLSS). Our implementation activities will be organized as follows**:**

- Literature survey of related works and highlight our goals**.** Which we did a lot of this by publishing comparative paper [15]
- Analysing the system requirement for each module and for integration
- Develop a system prototype.
- Testing the system and updates.

References

[1] YouTube Statistics, last access date is 10-12-2014; http://www.youtube.com/yt/press/statistics.html.

[2] V. N. I. (VNI). Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018, last access date is 10-12-2014, 2014; http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html.

[3] N. V. Uti, and R. Fox, The Challenges of Compressing and Streaming Real Time Video Originating from Mobile Devices, Multimedia Services and Streaming for Mobile Devices: Challenges and Innovation, pages. 1, 2011.

[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, volume(32), issue(4), pages. 577-587, 2010.

[5] L. Herranz, and J. M. Martínez, Combining MPEG Tools to Generate Video Summaries Adapted to the Terminal and Network, The Computer Journal, volume(56), issue(5), pages. 529-553, 2013.

[6] M. Ransburg, E. M. Graciá, T. Sutinen, J. O. Murillo, M. Sablatschan, and H. Hellwagner, Scalable video coding impact on networks, Mobile Multimedia Communications, pages. 571-581, 2012.

[7] S. Ibrahim, A. H. Zahran, and M. H. Ismail, SVC-DASH-M: Scalable video coding dynamic adaptive streaming over HTTP using multiple connections, Telecommunications (ICT), 21st International Conference on, pages. 400-404, 2014.

[8] D. Zhang, H. Li, and C. Chen, Robust Transmission of Scalable Video Coding Bitstream over Heterogeneous Networks, Circuits and Systems for Video Technology, IEEE Transactions on, 2014.

[9] B. T. Truong, and S. Venkatesh, Video abstraction: A systematic review and classification, ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP), volume(3), issue(1), pages. 37, 2007.

[10] G. Guan, Z. Wang, S. Mei, M. Ott, M. He, and D. D. Feng, A top-down approach for video summarization, ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), volume(11), issue(1), pages. 4, 2014.

[11] 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, 2008.

[12] 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, volume(32), issue(1), pages. 56-68, 2011.

[13] J. Almeida, N. J. Leite, and R. d. S. Torres, Online video summarization on compressed domain, Journal of Visual Communication and Image Representation, volume(24), issue(6), pages. 729-738, 2013.

[14] J. Niu, D. Huo, K. Wang, and C. Tong, Real-time generation of personalized home video summaries on mobile devices, Neurocomputing, ScienceDirect, volume(120), pages. 404-414, 2013.

[15] 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, volume(14), issue(2), pages. 37-50, 2014.

[16] 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. pages. 1260-1264.

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

[18] [18] R. Pal, A. Ghosh, and S. K. Pal, Video Summarization and Significance of Content: A Review, Handbook on Soft Computing for Video Surveillance, pages. 79, 2012.

[19] [19] M. Ajmal, M. H. Ashraf, M. Shakir, Y. Abbas, and F. A. Shah, Video summarization: techniques and classification, Computer Vision and Graphics, Springer, pages. 1-13, 2012

Acknowledgment

The work give agreat acknowledgement to Dr Alaa Hefnawy, Dr. Alaa Hamdy and Dr. Amr Mohamed for their great support and honest consultation in the core technical specilaization for each of them

Biographies

Assoc. Prof. Farouk is an associate Prof. since 2012. He joined the Electronics Research Institute, Egypt, in 1993. His fields of research are signal processing, mobile systems, Neural Networks, image compression, video processing, video compression, video indexing and retrieval, video on demand, pattern recognition and machine vision. Dr. Farouk received his Ph.D. at 2001 from Electronics & Communications Dept., Faculty of Engineering, Cairo Univ. and his M.Sc. at 1996 from Electronics & Communications Dept., Faculty of Engineering, Cairo Univ. Dr. Hesham participated in many national projects in MCIT developed based on portals and digital libraries. He also participated in some strategic studies as mobile for development.. Then he is a Research and innovation dept acting manager in ITI. Meanwhile, In ERI he is managing Technology Transfer office since June 2013 and the ERI technical office.

Prof. Kamal Abdelraouf ElDahshan is a professor of Computer Science and Information Systems at Al-Azhar University in Cairo, Egypt.
An Egyptian national and graduate of Cairo University, he obtained his doctoral degree from the Université de Technologie de Compiègne in France, where he also taught for several years.
During his extended stay in France, he also worked at the prestigious Institute National de Télécommunications in Paris.
Professor ElDahshan's extensive international research, teaching, and consulting experiences have spanned four continents and include academic institutions as well as government and private organizations. He taught at Virginia Tech as a visiting professor; he was a Consultant to the Egyptian Cabinet Information and Decision Support Centre (IDSC); and he was a senior advisor to the Ministry of Education and Deputy Director of the National Technology Development Centre. Prof. ElDahshan has taught graduate and undergraduate courses in information resources and centers, information systems, systems analysis and design, and expert systems.

Professor ElDahshan is a professional Fellow on Open Educational Resources as recognized by the United States Department of State.

Prof. Eldahshan wants to work in collaboration with the Ministry of Education to develop educational material for K-12 levels. Prof. Eldahshan is interested in training instructors to be able to use OER in their teaching and hopes to make his university a center of excellence in OER and offer services to other universities in the country.

Amr Abozeid received B.Sc. degree in computer science and mathematics from Al-Azhar University, Cairo, Egypt, in 2005. He received the M.Sc. in computer science from department of mathematics and computer science, faculty of science, Ain Shams University, in 2012. He is now a Ph.D. student in department of mathematics and computer science, faculty of science, Al-Azhar University. His main research topic is focused on video processing and computer vision. He is working as lecturer assistant of computer science and has eight years' experience in the field of teaching computer science topics, tools and technologies, supervising students' projects and leading software development projects.

Mayada Khairy Shehata graduated from the Telecommunications and Electronics Department, Faculty of Engineering, Helwan University, Cairo, Egypt in 2004. She received the M.Sc. in video processing from Faculty of Engineering, Helwan University, in 2010 ,Registered PhD in video processing, Faculty of Engineering, Helwan University in 2014 .Mayada was working on the field of broadcast (TV& Radio) as Projects Manger in Systems Design company from 2004 to 2013 dealing with Egyptian Radio Television Union(ERTU) , Egyptian Media Predication City (EMPC) ,Nielsat, and Private satellite TV Channels. She is dealing with international companies in the field such as; Miranda , Grass Vally Harris, Avid, DMTsys, etc… .Attending international Radio and Television exhibitions, Technical seminars and training for the mentioned companies in England, Dubai, Lebanon, ,Oman ,Jordan. Also working as Assistant Lecturer in Faculty of Engineering Thebes Academy (Part Time).Now working as Assistant Researcher in Computers and Systems department Electronics Research Institute (ERI) from July 2013 up till now.

ISBN: 978-1-61804-271-2