Error resilient scalable video transmission system using multiple description coding

Yujie Zhang[1] ,Yao Zhao[2] ,Huihui Bai[3] ,Chunyu Lin[4]
Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China
yzhao@bjtu.edu.cn

Koh Hang[5],Anhong Wang[6]
A*STAR Institute for Infocomm Research(12R) [5]
Taiyuan University of Science and Technology,Taiyuan 030024,China[6]

Abstract—**In this paper a practical real-time video transmission system is proposed by combining SVC (Scalable Video Coding) and MDC (Multiple Description Coding). Each description of MDC can be designed as scalable bit stream which is helpful for UDP/RTP based transmission over practical packet loss network, such as Internet. To provide different spatial resolution / fidelity videos and flexible transmission over error-prone network, a new coding architecture needs to be revised. In the proposed coding scheme, each MDC description includes independent copy of raw video with lower quality, and each copy can be compressed using SVC. The proposed system achieved better experimental results when compared to schemes without MDC, especially when packet loss occurs.**

Keywords- Scalable video coding; H.264/AVC; RTP; video capture

I. INTRODUCTION

Recently, video applications have become widely available with the prosperity and development of Internet. Video compression today is used in a wide range of applications ranging from multimedia messaging, video telephony and video conferencing over mobile TV, wireless and Internet video streaming, to standard- and high-definition TV broadcasting. Furthermore, video transmission is often exposed to variable transmission conditions, which needs to be dealt with using scalability features. At the same time, video content is delivered to a variety of decoding devices with heterogeneous display and computational capabilities.

The Scalable Video Coding (SVC) amendment [1][2] of the H.264/AVC standard (H.264/AVC) provides network-friendly scalability at a bit stream level with a moderate increase in decoder complexity relative to single-layer H.264/AVC. It supports various functionalities such as bit rate, format, and power adaptation, graceful degradation in lossy transmission environments，which provide enhancements to transmission and storage applications. The target of SVC is that within a very wide range of bit rates obtaining high transmission efficiency and decoding quality, within a lower complexity providing temporal, spatial and quality scalability coding, offering network, terminal and a seamless adaptive release of streaming media.

Multiple Description Coding (MDC) has emerged as a promising technology for robust transmission over error-prone channels, which has been attracting more and more researchers [3]. MDC are proposed based on the assumption that multiple channels exist between the source and destination and it is impossible that all the channels fail at the same time. Therefore, using MDC the source can generate multiple bit streams (descriptions) with equal priority at the encoder, which then can be transmitted over multiple channels. At the decoder, each description can be decoded independently to produce a minimum fidelity which is measured by side distortion. With the increasing of the received descriptions, the reconstructed quality can be enhanced further. In a simple architecture of two channels, the distortion generated by two received descriptions is called central distortion [4].

In this paper a practical real-time video transmission system is proposed by combining SVC and MDC, which has two advantages. One is the basic codec of SVC is not modified to keep the compatibility with the current video standard. And the other is each description of MDC can be designed as scalable bit stream which is helpful for UDP/RTP based transmission over practical packet loss network. It is noted that in this paper we mainly focus on the combination of SVC and MDC. Therefore, the rate-distortion performance may be improved by using better MDC schemes.

II. THE PROPOSED SYSTEM

A. Overview of the system

Fig. 1 shows the diagram of the proposed system. We have developed the compliant-standard scalable encoder to realize encoding-once for heterogeneous decoding according to the users' settings and the devices' computation capacity. Then the network transport protocol RTP/UDP/IP is utilized for scalable streaming. This system has achieved reliable transmission of video streams on wired and wireless network, and supports spatial scalability at different endpoints and devices.

Our real-time video transmission system is based on Client/Server model, and its architecture is shown in Fig. 2. Our system includes three parts: Server, Internet and Client. The Internet part is using the existing network in our University. Server contains three subparts of video capturing, video encoding and data sending. Client also includes three subparts: data receiving, video decoding and video display. It should be noted that the encoder and decoder are the combination of SVC and MDC.

978-0-7695-4738-1/12 $26.00 © 2012 IEEE
DOI 10.1109/CMCSN.2012.6

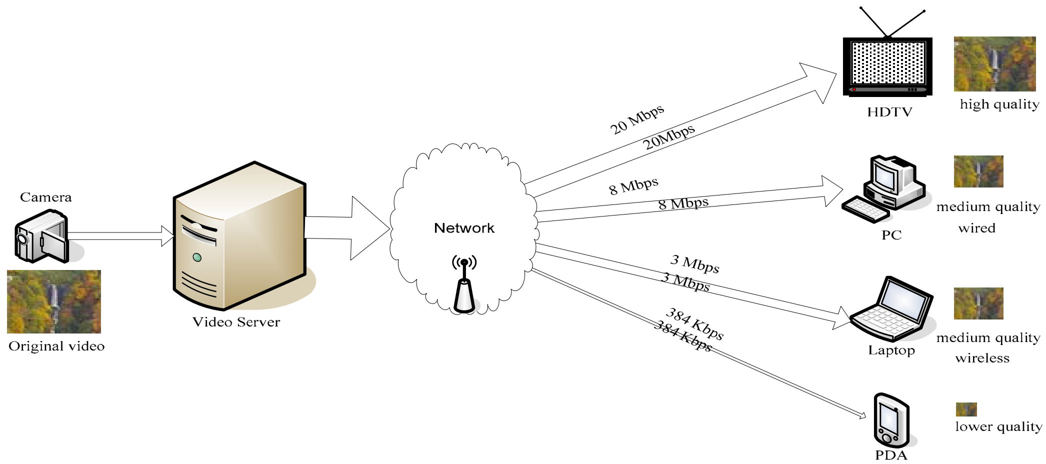

Figure 1. The diagram of the proposed system

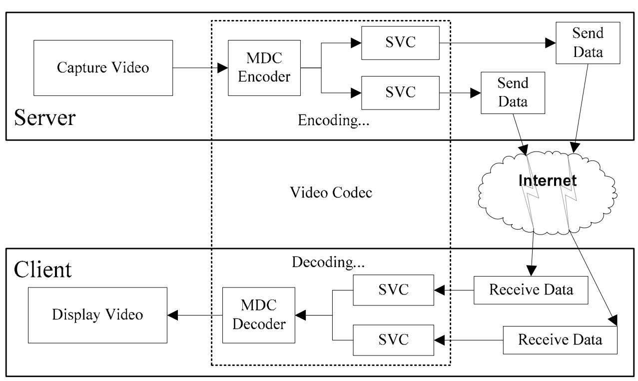

Figure 2. Framework of video transmitting

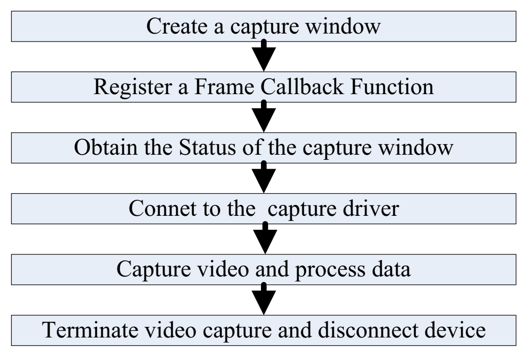

Figure 3. The process of VFW video capture

B. Video Capturing and Display

There are typically three video capturing technologies: VFW, DirectShow and SDK. Among them, VFW (Video for Windows) is Microsoft's digital video package which provides a set of library functions, to implement features such as video capture, image compression and image playback. The usage of VFW has the advantage that it followed the installed Windows operating system, thus the executable file can run without additional library files. VFW can capture digital video signals from a video source and store to a file or directly handle the video cache, which makes real-time video processing possible. This is exactly in line with our current hardware facilities. Fig. 3 gives the flow chart of VFW video capture [8].

We can use function DrawDibDraw in VFW to display the video. Function DrawDibDraw requires RGB video format, however the decoding video format is YUV, thus we need to change video format from YUV to RGB using the standard algorithm available.

C. Video Codec

This part contains the combination of MDC and SVC. MDC can provide robustness of our system over error-prone internet. We adopt an easy scheme of MDC because of the real-time requirement, including two descriptions: the original video sequence (original description) and the rotation description (clockwise rotate 180°), each of which contains many slices (at least one). In the encoder, we encode both the original description and rotation description. In the decoder, there are many cases according to the loss of two descriptions as shown in Fig. 4.

After SVC decoding, we achieved different reconstructed frames according to the received descriptions. If we only received one description, for example the original slice 2, but lost the rotated slice 4, we can use the original slice to replace the rotated slice. If we received both descriptions, we can average the two descriptions for reconstruction.

D. Network data transmission

After the video encoding, the compressed data is streamed from the Server to the Client, over the Internet. One simple way is P2P based on UDP, one Server one Client. However, there may be multi clients in our system, such as different PC Client (different network, different platform,

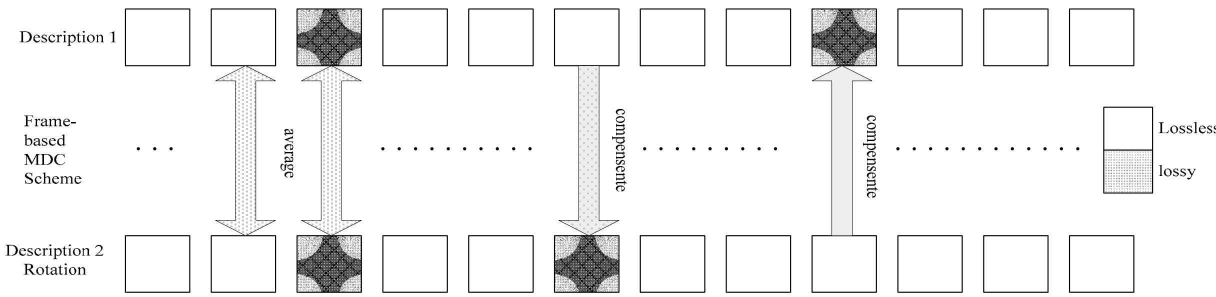

Figure 4. Figure 1 Schemes of MDC decodin

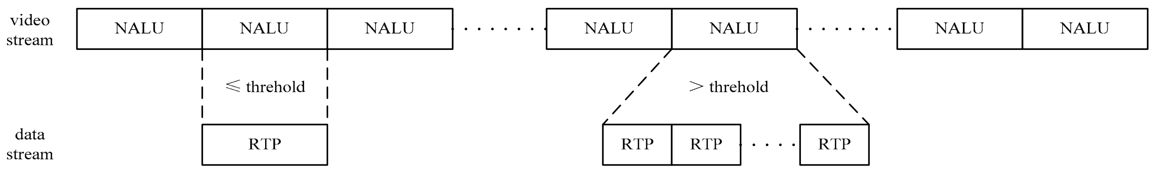

Figure 5. NALU package method

and different operation system). So the most appropriate network communication model is multicast, which can achieve one-to-many data transmission.

As to the packetisation format of the compressed video stream, we follow the RFC 3984 [7] and RFC 6190 [5]. In summary, our system uses UDP/RTP protocol and socket to transmit data (Fig. 5 gives a simple note of how to packetize the video stream after encoding to sending data in the internet in which we need a threshold to decide whether to divide NALU or not). In order to receive data and display video at the same time, multi-thread programming is used.

Our MFC-based system is developed under Visual Studio 2005 in Windows and only feasible in LAN. We have tested it in Windows XP and Win7. Fig. 6 shows the interface of our system.

Figure 6. System Interface

III. TESTS AND PERFORMANCE

We use ordinary USB camera, which provide 5 video resolutions: 640x480, 352x288 (CIF), 320x160, 176x144 (QCIF), 160x120. The Server provides the preview of captured video. The Client can set the IP of Server and set different display layer. For example, if the video source is 704x576 (4CIF) with three layers encoding, the client can choose three different display resolutions freely (4CIF, CIF, QCIF) as shown in Fig. 7.

Figure 7. Results of different display settings

It can run smoothly in wired and wireless networks. We build a LAN using a gigabit router which supports wired and wireless networks. After we start the Server, the Client can join the multicast network at anytime.

Table 1 shows the efficiency of encoder and decoder. From table 1, we know that the codec library can meet the real-time requirement. The sequence is the standard test sequence of JSVM (Joint Scalable Video Model). FPS denotes number of frames per second.

TABLE I. THE EFFICIENCY OF CODEC LIBRARY

Sequence	**Frame Number**	**Encoder**		**Decoder**	
		Time	***FPS***	***Time***	***FPS***
CITY_4CIF	600	11.402	52.6	15.397	38.9
ICE_4CIF	240	3.108	77.2	5.990	40.0
SOCCER_4CIF	300	5.923	50.6	7.472	40.1

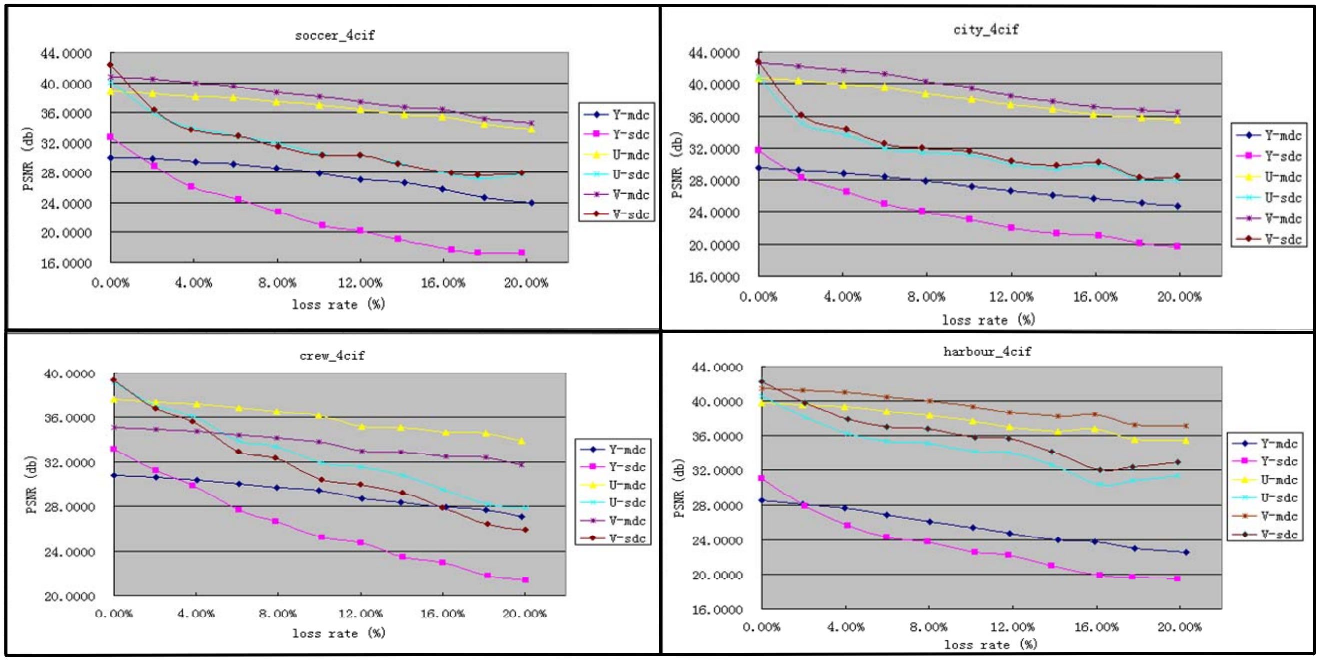

Figure 8. Results of MDC and SDC

As to our MDC schemes, we compare it with the original SVC (SDC, Single Description Coding) which uses smaller QP to ensure that the two schemes are in the same bit-rate. The results in Fig. 8 show that the MDC outperforms the SDC.

As described above, there are problems to be solved. First, we'd better provide more encoding settings in the Server and display parameters in the Client. Second, the codec library only supports spatial scalability so far, it will be better if it also supports temporal and quality scalability. Third, we can increase the diversity of endpoints, such as we can use PC, laptop, PAD, phone and so on.

IV. Conclusion

As described above, our real-time video transmission system based on SVC, using UDP/RTP/IP as network transport protocol, has achieved reliable transmission of video streams on wired and wireless network, and supports spatial scalability in different endpoints and devices. Those all meet our original designation and have certain universality in the applications of H.264/AVC SVC. Our MDC schemes also achieve a good performance in our test result.

Reference

[1] Advanced Video Coding for Generic Audiovisual Services, ITU-T Rec. H.264 and ISO/IEC 14496-10 (MPEG-4 AVC), ITU-T and ISO/IEC JTC 1, Sept.2009.

[2] Heiko Schwarz, Detlev Marpe, Thomas Wiegand. Overview of the Scalable Video Coding Extension of the H. 264 /AVC Standard [J]. IEEE Transactions on circuits and systems for video technology, 2007, 17 (9): 1103 – 1120

[3] Y. Wang, A. R. Reibman, and S. Lin, "Multiple description coding for video delivery," Proceedings of the IEEE, vol. 93, no. 1, pp. 57–69, 2005.

[4] V. K. Goyal, "Multiple description coding: compression meets the network," IEEE Signal Processing Magazine, vol. 18, no. 5, pp. 74–93, 2001.

[5] S. Wenger, Y.-K. Wang, T. Schierl, A. Eleftheriadis. "RTP Payload Format for Scalable Video Coding", RFC6190, May 2011.

[6] H. Schulzrinne,S. Casner, R. Frederick, V. Jacobson. "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.

[7] S. Wenger, M.M. Hannuksela, T. Stockhammer, M. Westerlund, D. Singer. "RTP Payload Format for H.264 Video", RFC3984, Feb 2005.

[8] Microsoft MSDN. The Introduction of Using Video Capture. http://msdn.microsoft.com/zh-cn/library/ms713463(v=vs.85).aspx, Oct 2011/Feb 2012.

[9] Heiko Schwarz, Tobias Hinz, Karsten Suehring. JSVM Software Manual, HHI. Mar 2011.

[10] Thomas Schierl, Thomas Stockhammer, Thomas Wiegand. Mobile Video Transmission Using Scalable Video Coding. IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Scalable Video Coding, Vol. 17, No. 9, pp. 1204-1217, Sep 2007.

[11] Stephan Wenger, Ye-Kui Wang, Thomas Schierl. Transport and Signaling of SVC in IP Networks. IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Scalable Video Coding, Vol. 17, No. 9, pp. 1164-1173, Sep 2007.