Multi-view Video Coding for 3DTV

Nurulfajar Abd Manap & John Soraghan
Department of Electronic and Electrical Engineering
University of Strathclyde, UK / Universiti Teknikal Malaysia Melaka (UTeM)

Abstract

With the advent in display technology, the 3DTV will provide a new viewing experience without the need of wearing special glasses to watch the 3D scenes. One of the key elements in 3DTV is the multi-view video coding; obtained from a set of synchronized cameras, capture the same scene from different viewpoints. The video streams are synchronized and subsequently used to exploit the redundancy contained among video sources. A multi-view video consists of components for data acquisition, compression, transmission and display. This paper outlines the design and implementation of a multi-view video system for transmission over a wireless channel. Synchronized video sequences acquired from four separate cameras and coded with H.264/AVC.

Background

The demand for multi-view video coding is driven by the development in new 3D display technologies and the growing use of multi-camera arrays. A variety of companies are starting to produce 3D display technologies that do not require special glasses and can be viewed by multiple people simultaneously. This technology provides a good platform for new applications to emerge such as 3D scene communication [1]. Even with 2D displays, multi-camera arrays are increasingly being used to capture a scene from many angles. The resulting multi-view data sets allow the viewer to observe a scene from any viewpoint and serve as another application of multi-view video compression.

Multiple camera views of the same scene require a large amount of data to be stored or transmitted to the user. Furthermore for real-time multi-view video processing it demand extensive processing capabilities and predicted to consume a large portion of the bandwidth available in the future [1]. Therefore, efficient compression techniques are essential. The simplest solution for this would be to encode all the video signals independently using a state-of-the-art video codec such as H.264/AVC [2, 3]. However, this is inefficient, as it does not exploit the correlation or inter-view statistical dependencies that exist in the multi-views. These redundancies can be exploited, where images are not only predicted from temporal neighboring images but also from corresponding images in adjacent views, referred to as multi-view video coding (MVC).

In this paper, a multi-view video system for wireless applications will be presented. The system consists of components for data acquisition, compression, transmission and display. The main features of the system include wireless video transmission system for up to four cameras, by which videos can be acquired, encoded and transmitted wirelessly to a receiving station. The video streams can be displayed on a single 3D or on multiple 2D displays. The encoding for the multi-view video through inter-view and temporal redundancies increased the compression rates. The H.264/AVC multi-view compression techniques has been exploited and tested during the implementation process. One of the highlights in this paper is the low cost implementation of a multi-view video system, which using only typical web cameras attached to a single PC.

Multi-View Video Compression

Many 3DTV systems are based on scenarios, where a 3D scene is captured by a number of N cameras [4]. The simplest case is classical stereo video with two cameras. More advanced systems apply 8, 16 and more cameras. Some systems traditionally apply per sample depth data that can also treated as video signals. An overview of compression algorithms and standards can be found in [5], which includes the conventional stereo video coding, video plus depth data and multi-view video coding. Depending on the degree of common content shared by a subset of the cameras, a coding gain can be achieved in comparison to single-view coding.

In multi-view coding, correlations between adjacent cameras are exploited in addition to temporal correlations within each sequence. The multi-view video coding adds another compression dimension on the top of single-view coding: the inter-view direction. Exploiting redundancies among the multi-view video images is the key to efficient compression. Therefore, for this research, the multi-view video coding has been selected due to the fact that this technique provides a high compression rates compared to the simulcast coding. The video data will be compressed with H.264/AVC algorithms before transmitted over the wireless channel.

System Architecture

The proposed multi-view video system shown in Figure 1 mainly consists of four video cameras, one acquisition PC, multi-view codec with error protection and correction, transmission, reception and display. These components can be classified into four modules: acquisition, data encoding and decoding, error protection and correction, and lastly the display.

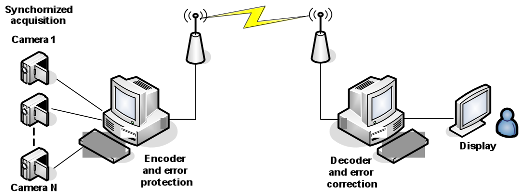

Figure 1 - General scheme of the proposed system

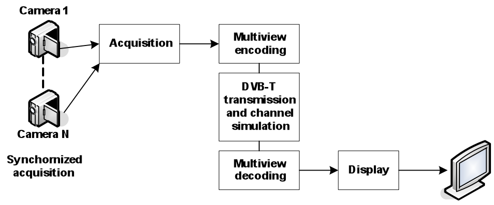

Figure 2 – Block diagram of system design

The whole system can be simplified as shown in Figure 2, which consist all the modules. The system divided into several modules for some practical reasons. Such modules match naturally the functions provided by real devices used in a real implementation of the system. The division of a major problem into a set of smaller problems reduces design and implementation complexity of the original problem. By defining self-contained modules, it helped the debugging process and allows reusability. In addition, it also simplifies code maintenance and modification.

Results

The first goal of the test is to ensure that the H.264/AVC reference software could handle the multi-view video streams. Every view of the cameras contains 50 frames in YUV format and CIF size captured at 15 fps. The simulcast coding is achieved by coding each view sequence separately using H.264/AVC standard. The quality of the encoded sequences was measured by the average PSNR of their frames.

The parameters that have been set for the encoding process were 352x288 image format, 16 search range, IPPP sequence type and full Motion Estimation scheme search. A set of multi-view video sequences, called 'Book' was captured. It contains of four views of 50 frames each at 15 fps. The sequences were set to produce CIF size sequences. To illustrate the nature of the captured data sets, frames 25 of the four cameras from 'Book' sequences are shown in Figure 3. With H.264 Analyzer released by MMRG team [6], the output of the coded multi-view video (macroblocks) labeled with selected colors to distinguish which camera sequence they are referred from.

Conclusion

A multi-view video coding simulation based on H.264/AVC for wireless channel has been presented. The coding scheme processed the frames of sequences captured by multiple cameras from a scene. The codec is based on the JM H.264/AVC software version 10. Five modes of operation are simulated based on the MMRG H.264 Multi-view Extension. The acquisition stage consists of an array of synchronized cameras that are connected to a single PC through the USB connection. One of the highlights in this research was that the implementation cost for this system is quite low since it used a typical web camera attached to the PC. The system can be upgraded to higher state with better specification of the equipment from acquisition to display.

References

[1] M. Flierl and B. Girod, "Multiview Video Compression," *Signal Processing Magazine, IEEE*, vol. 24, pp. 66-76, 2007.

[2] P. Merkle, A. Smolic, K. Muller, and T. Weigand, "Efficient Prediction Structures for Multi-View Video Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, 2007.

[3] G. Li and Y. He, "A Novel Multi-View Video Coding Scheme Based On H.264," *Proceedings of the Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing,* vol. 1, pp. 15-18, 2003.

[4] I. Ahmad, "Multi-View Video: Get Ready for Next-Generation Television," *Distributed Systems Online, IEEE*, vol. 8, pp. 6-6, 2007.

[5] H. Y. Shum, S. B. Kang, and S. C. Chan, "Survey of Image-Based Representations and Compression Techniques," *IEEE Trans. Circuits Systems Video Technology*, vol. 13, pp. 1020-1037, Nov. 2003.

[6] C. Bilen, A. Aksay, and G. B. Akar, "A Multi-view Video Codec Based on H.264," *IEEE ICIP 2006*, Oct. 2006.