3D TV Using MPEG-2 and H.264 View Coding and Autostereoscopic Displays

Lakis Christodoulou, Liam M. Mayron, Hari Kalva, Oge Marques, and Borko Furht

Dept. of Computer Science and Engineering, Florida Atlantic University, Boca Raton, FL 33431

ABSTRACT

There is a renewed interest in the 3DTV research primarily due to the advances in low cost 3D display technologies. The two views required for 3DTV can be compressed using standard video compression techniques. MPEG-2 is widely used in digital TV applications today and H.264/MPEG-4 AVC is expected to be the leading video technology standard for digital video in the near future. The compression gains and quality of 3DTV will vary depending on the video coding standard used. While inter-view prediction will likely improve the compression efficiency, new approaches such as asymmetric view coding are necessary to greatly reduce bandwidth requirements for 3DTV. This demonstration will show the quality of 3D video experience on autostereoscopic displays using H.264 and MPEG-2. We will also show the benefits of using asymmetric view coding as well as the role of eye dominance on 3DTV experience.

Categories and Subject Descriptors

I.4.2 [Image Processing and Computer Vision]: Compression (Coding) – Approximate methods

General Terms

Algorithms, Performance, Human Factors, Experimentation

Keywords

H.264, 3DTV, Eye dominance, asymmetric view coding

1. INTRODUCTION

Latest advances in video compression and 3D displays have necessitated a further understanding and development of video coding algorithms for 3DTV applications. The emergence of low cost autostereoscopic displays is expected to drive the growth of 3DTV services [1]. As video compression is a key issue in traditional video communications and the use of two video streams makes it even more important for 3DTV [2][3]. New approaches beyond the traditional predictive coding are necessary to reduce the bandwidth requirements for 3DTV.

One of the first applications of asymmetric view coding for MPEG-2 based 3DTV applications was reported in [4]. Our work uses a similar approach but considers the human visual system as well as the autostereoscopic displays which do not require glasses for experiencing 3D video. Our work on the use of asymmetric view coding based on H.264 was reported in [5]. The key challenges in 3D TV development are presented in [6]. While H.264 is expected to replace MPEG-2 video over the next several years, these two technologies will co-exist during this migration and 3DTV systems based on both MPEG-2 and H.264 will likely be used. A good understanding of 3DTV quality and compression

when using MPEG-2 as well as H.264 is necessary to plan and build 3DTV services.

We are currently studying the characteristics of the human visual system that can be exploited to compress individual stereo views at different qualities without affecting the perceptual quality of the 3D video. In the particular case, efforts have been focused on the human factors, such as the eye dominance of the viewer, in order to investigate the impact on the 3D video quality, and so exploit the asymmetric video coding. We examine the bounds of asymmetric stereo view compression and its relationship to eye-dominance based on a user study. This demonstration also presents the design and development of a modular video player with stereoscopic and multi-view capabilities, including a discussion of useful tools for accelerating the development and enhancing flexibility. The experimental results indicate that eye-dominance influences 3D perception, and as a result will impact the coding efficiency of 3D video. Preliminary results show that the deblocking filters in H.264 have beneficial impact on the quality of 3D video, whereas, the blocky artifacts that are visible in MPEG-2 at lower quality degrade the 3D experience. Avoiding blocky artifacts in MPEG-2 requires higher bitrates and hence higher bandwidth requirements for 3DTV.

This demonstration will show the application of asymmetric view coding utilizing MPEG-2 and H.264/MPEG-4 AVC to compare and evaluate compression efficiency, savings in bit rate, and picture quality on the experience of digital 3D TV. The key elements of this demonstration are:

- 3D TV using MPEG-2
- 3D TV using H.264/MPEG-4 AVC
- The impact of the compression algorithm on the experience of the 3D TV visual quality
- The impact of the asymmetric view coding on the experience of the 3D TV visual quality
- The impact of the eye dominance on the experience of the 3D TV visual quality

2. 3D SYSTEM OVERVIEW

Most of the research groups have paid little attention to the scalability of video surveillance systems, considering sufficient availability of network bandwidth, computational power, and bit rate. Our research motivation is to determine new algorithms that can support 3D video surveillance infrastructures incorporating computer vision algorithms with the most economical approach in bit rate, and simultaneously succeeding in compression efficiency. Under these considerations, we are currently developing a 3D/multi-view video coding system with applications in 3D TV, security, and surveillance. The goal of this project is to develop technologies and tools for efficient compression, communication, and playback of multi-view and 3D video.

Figure 1 shows the general architecture of a 3D video system. The stereo views are encoded at the sender by exploiting the large amount of redundancies among the views. This asymmetric view

coding provides the opportunity to exploit and achieve a range of video quality, in which the quality of each view can be dropped, so that can reduce the video bit rate without significantly affecting the 3D observation's quality. We use H.264/MPEG-4 AVC standard as the core compression engine with inter-view prediction to increase compression efficiency. The coded views are communicated to the receiver while the decoded views are rendered on an appropriate display. The 3D displays use a pair of coded views (left and right view) to display 3D video with depth perception.
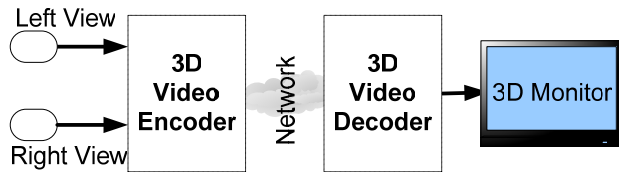

Figure 1. 3D/Multiview Video System

We use the Sharp LL-151-3D autostereoscopic display to render the stereoscopic videos. The display is 15-inches, XGA resolution (1024 by 768 pixels). This display which uses lenticular imaging techniques and renders depth very accurately gives a true 3D experience. The perception of depth is achieved by a parallax barrier that diverts different patterns of light to the left and right eye.

3. STEREOSCOPIC & MV VIDEO PLAYER

This section presents the architecture of a modular video player, Figure 2, with stereoscopic and multi-view capabilities. The designed architecture supports the playback of monoscopic and stereoscopic multi-view video. Here the stereoscopic video is generated by the binocular fusion (combined visual perception of the scene) composed by the inter-pixel-column interpolation of the independent left and right sequences, or as a single video formatted with the left and right views side-by-side or top-to-bottom. The software architecture of the MV video player was implemented and tested on the Microsoft Windows XP platform. The Microsoft DirectShow framework was used for the capture and transform functions. MFC was used to implement the interface. An open source project, AviSynth Tool was used for the video playback of our 3D research experiments.
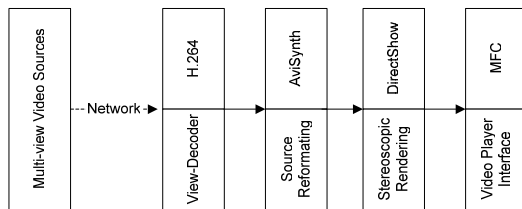

Figure 2. 3D/MV player architecture

Figure 3 shows the snap shots of the left and the right views coded using H.264 with asymmetric video quality. The left view was coded at a very low quality (PSNR of 26.2 dB) and the right view was coded at broadcast TV quality (PSNR of 38.4 dB). The left view has a very low quality (with completely blurred facial features) and is not acceptable by itself but results in a good 3D experience when used with a higher quality right view.

5. References

[1] N.A. Dodgson, "Autostereoscopic 3D Displays," *Computer*, Volume 38, Issue 8, Aug. 2005 pp. 31 – 36.

[2] B.L. Tseng and D. Anastassiou, "Multi-Viewpoint Video Coding with MPEG-2 Compatibility," *IEEE Trans. Circuits and Systems for Video Tech.*, Vol. 6, No. 4, pp. 414-419.

[3] A. Puri, R. V. Kollarits and B. G. Haskell, "Stereoscopic Video Compression Using Temporal Scalability," *Proc. SPIE Visual Communications and Image Processing'95*, Taiwan, May 1995.

[4] L B. Stelmach, W. J. Tam, "Stereoscopic image coding: Effect of disparate image-quality in left- and right-eye views", *Signal Processing: Image Communication*, Vol. 14, pp.111-117, 1998.

[5] H. Kalva, L. Christodoulou, L. Mayron, O. Marques, and B. Furht, "Design and Evaluation of 3D Video System Based on H.264 View Coding", *Proceedings of the NOSSDAV 2006*, Newport, Rhode Island, May 22–23, 2006, pp. 68-73.

[6] H. Kalva, L. Christodoulou, L. Mayron, O. Marques, and B. Furht, "Challenges and opportunities in video coding for 3D TV ", *IEEE International Conference on Multimedia & Expo (ICME) 2006*, July 9-12, 2006, Toronto, Canada.

Figure 3. Snapshots of the right view coded at a low quality (above) and left view coded at a high quality (below)