Editorial

Special issue on advances in three-dimensional television and video: Guest editorial

Three-dimensional television (3DTV) is a new paradigm for consumer electronics, enabling viewers perceiving an extra dimension, which is not possible in conventional television systems. The 3D perception is only possible as a result of the delivery of the artificially generated light rays to the eyes of a human, appropriately by 3D displays. Hence, for 3D viewing experience, extraction of 3D scene structure as well as compression and transmission of this structure to the receiving end displays are strictly necessary.

While the scientific efforts on 3DTV are getting matured, the fundamental research on 3DTV is being replaced by end-to-end systems, considering extraction, compression, transmission and rendering issues together. Additionally, the computational complexity is also taken into account towards the consumers in the mass market. This special issue is dedicated to the recent advances in 3DTV research with particular emphasis on the 3DTV systems, which have a broader vision on the applicability of their results.

The special issue contains the extended versions of 10 selected papers from the conference 3DTV-CON 2008, which was held in İstanbul, Turkey, between 28 and 30 May 2008. The conference was organized in 2007 and 2008, and gained an international reputation in a short time. The 3DTV-CON has been initiated by the project *3DTV: Integrated Three-Dimensional Television—Capture, Transmission, and Display*. The project is a Network of Excellence project bringing together over 180 researchers from 19 European institutions and led by Bilkent University, Ankara, Turkey. The project started at September 2004 and successfully completed at the end of August 2008 by the support of European Union Sixth Framework Information Society Technologies (IST) Program.

While selecting the 10 papers for this special issue from 98 high-quality research papers that were presented in the conference, we give priority to the papers addressing an end-to-end 3DTV system, free-viewpoint video, graphics processor unit (GPU)-based approaches, and MPEG-4 3D graphics, rather than the papers addressing the specific aspects of 3DTV, such as 3D scene extraction, representation, transmission, display, holographic techniques.

Although, 3D scene extraction is a relatively mature research area, the application of this technology to 3DTV systems requires different aspects of this problem to be taken into account. One of these problems, namely the effect of depth extraction performance on arbitrary view rendering, is crucial for any multi-scopic 3D display or free-viewpoint TV (FTV) application. Obviously, defining a quality metric for rendering performance is a major step for solving this problem. Another emerged problem is the estimation of dense depth field for each view of a multi-view imaging system, noting that this new content type is expected to be standardized by International Organization for Standardization (ISO) as the representation of a 3D scene. The 3D scene extraction from uncalibrated cameras is another challenging area, considering the fact that such cameras are available in various applications of today, such as multi-camera surveillance. Finally, the implementation of these algorithms on graphics accelerator hardware is also critical for any real-time application. This issue has answers to all the aforementioned problems for 3D scene extraction.

With the assumption of the 3D scene information being obtained, the next problem for any 3DTV system is the compression and delivery of the resulting scene representation. However, in the state-of-the-art solutions for these problems, apart from conventional source coding approaches, other issues are also handled. Redundancy removal from any 3D representation should be achieved considering the quality of the rendered virtual views for 3D or FTV displays. In other words, bit-rate (of the 3D representation) and distortion (of the rendered view) should be jointly optimized for obtaining better performance. Both dense depth representations and polygonal meshes should utilize rate-distortion efficient modeling for compression. On the other hand, in case of peer-to-peer applications, the visualization of mesh representations requires the utilization of hierarchical level of detail

(LoD) approaches for efficient streaming. However, for terrestrial broadcast with limited channel utilization, depth data should be processed a priori for a smoother field. A number of papers in this issue propose novel solutions to these problems.

John Kilner and his colleagues propose an objective quality measure for free viewpoint video production. The quality metric is represented by the structural registration error in view synthesis. It can be applied as a full-reference metric to assess the quality of view synthesis with respect to a ground truth image, or as a no-reference metric to measure the error in registering scene appearance in image-based rendering.

Songkran Jarusirisawad and Hideo Saito describe a weak calibration method by using projective grid space for synthesizing free-viewpoint video from uncalibrated multiple cameras. They segment the scene into the foreground moving object and background region. The free viewpoint video is synthesized based on the reconstructed visual hull of the foreground object. The background is approximated with a set of planar segments and rendered by projecting each segment to the novel viewpoint.

Dongbo Min and his colleagues describe a method to render free viewpoint video throughout a multi-view image sequence. The method works sequentially by estimating disparity maps of a designated reference and target viewpoints. Their method provides efficient multi-view stereo matching and virtual view generation, which are very crucial for free viewpoint video.

Sammy Rogmans and his colleagues present a real-time image-based rendering framework that is capable of performing multiple intermediate view synthesis, completely on the graphics processing unit (GPU). The proposed framework seamlessly connects stereo correspondence and view synthesis. The intermediate view synthesis is implemented as a chain of consecutive processing modules. They achieve this by using a GPU-accelerated software model, where the modules are optimized implementations of recent real-time algorithms.

Yuji Mori and his colleagues propose a method to render a novel view image using pre-computed multi-view images and depth maps. They apply a post-filtering process to the depth map projected to a virtual plane to handle the problems caused by 3D warping, which is used to generate the virtual views.

Philipp Merkle and his colleagues provide a comparative study on the effect of depth compression on multi-view rendering quality. They investigate the implications of using different depth-image compression techniques on the quality of virtual view rendering. They evaluate the influence of geometry distortions resulting from coding artifacts in two ways. First, they analyze the variation of 3D surface meshes using the Hausdorff distance, and second, they evaluate the distortion for 2D view synthesis rendering.

Sehoon Yea and Anthony Vetro propose the use of view synthesis prediction for reducing rate-overhead incurred by transmitting depth-maps in free viewpoint TV applications. They describe the addition of inter-view prediction with view synthesis and novel skip and direct modes for view synthesis. Their method significantly improves the performance of multi-view video coding (MVC).

Blagica Jovanova and her colleagues describe 3D graphics compression solutions for eXtensible Markup Language (XML)-based models. The proposed software solution is composed of an Encoder, which extracts the geometry and animation parts from an XML-based scene graph to encode them by using some components of AFX tools, and a decoder to recover the original XML file. This contribution addresses the compression and interoperability issues of current on-line virtual environments.

Romain Cavagna and his colleagues describe a framework for peer-to-peer visualization of large scale 3D landscape and city models. The framework incorporates technologies, such as LoD and view-dependent streaming, which are implemented using the MPEG-4 AFX standard.

Young Kyung Park and her colleagues propose a new depth image-based rendering technique for 3DTV over T-DMB, which is backward compatible to standard monoscopic displays. They define the service architecture for depth image-based rendering based 3DTV service over T-DMB, a new depth pre-processing technique based on adaptive smoothing that is capable of reducing the bitrates required for transmission and reducing the holes, and a look-up-table-based method for fast operation.

This special issue would not be possible without the contributions of the following. First of all, we are grateful to all authors who made contributions to the special issue. The reviewers did an excellent job and their comments and suggestions improved the quality of the papers. Editor-in-Chief Murat Tekalp deserves special thanks for accepting this special issue proposal and for his support and guidance throughout the paper submission and reviewing process. The editors are also grateful to 3DTV project leader Levent Onural of Bilkent University for starting the 3DTV-CON conference series. The editors hope this special issue is useful to the researchers working in this area and open new horizons for making 3DTV possible.

Guest Editors
Uğur Güdükbay
Department of Computer Engineering,
Bilkent University, 06800 Ankara, Turkey
E-mail address: gudukbay@cs.bilkent.edu.tr

A. Aydın Alatan
Department of Electrical and Electronics Engineering,
Middle East Technical University,
06531 Ankara, Turkey
E-mail address: alatan@eee.metu.edu.tr