Available online at www.sciencedirect.com
ScienceDirect
Signal Processing: *Image Communication* 22 (2007) 103–107
SIGNAL PROCESSING: IMAGE COMMUNICATION
www.elsevier.com/locate/image

Guest Editorial

Special issue on three-dimensional video and television

The term three-dimensional television (3DTV) is understood as a natural extension of two-dimensional television—which produces a flat image on a screen—into the third dimension. This means that the impression of the viewer also involves the perception of depth. Our two eyes view the world from slightly different positions and the two slightly different images registered by our eyes play a primary role in depth perception. The brain merges this information and creates the perception of a single 3D view. In the absence of the original scene or object, the same perception can be created by employing systems exploiting this mechanism. Two cameras capture scenery from viewpoints corresponding to the human eye positions; such a system is referred to as a stereo camera. If a display system can then ensure that each eye sees only the one corresponding view, the impression of depth will be created. Such so-called stereo video systems have been known for decades. Most of these systems employ special glasses in order to ensure that each eye perceives only the one corresponding view it is supposed to perceive. In an anaglyph system, the images are overlaid and glasses with colour filters (for instance, red and green) are used. Some other systems use glasses with shutters and temporally interleaved video (as in IMAX theatres).

The necessity to wear glasses while viewing has often been considered as a major obstacle in front of wide acceptance of 3DTV technology. Despite this, 3DTV has found acceptance and is growing in niche markets including, for instance, computer games, which are in most cases produced in 3D anyway. Graphics drivers that produce stereo video output are freely available. By purchasing appropriate glasses (available below 50 USD), an ordinary personal computer system can be converted into a 3D display system, allowing 3D games to be played. Natural stereo video is also becoming more and more available on the Internet and on DVDs and similar media. Such content is either originally captured in stereo (such as in some IMAX movies) or is converted from two-dimensional (2D) video. 2D to 3D conversion is possible with user-assisted production systems, which are of great interest for content owners and producers. But perhaps more significantly, newer 3D displays are now available, where glasses are not necessary to obtain the impression of 3D. For instance, autostereoscopic displays use lenticular screens to display two images at the same time, such that each eye can see only one of them.

Another drawback of classical stereo video systems that capture and display two fixed views to the user is that, head motion parallax effects cannot be supported. A human watching a 3D scene expects occlusion and disocclusion effects when moving with respect to the scene: certain parts of objects should appear and disappear as one moves around. It is desirable to support at least to some extent looking around objects. This is not possible with two fixed views. Therefore, classical stereo video, where the viewpoint is fixed to the position of the stereo camera, produces an unnatural result if the observer is moving.

Adding a limited free viewpoint video (FVV) functionality may enable viewers to move. FVV offers the same functionality that is known from 3D computer graphics. The user can choose a viewpoint and viewing direction within a visual scene; that is, interactive free navigation is possible. In contrast to pure computer graphics applications, FVV targets real world scenes as captured by real cameras. As in computer graphics, FVV relies on a certain 3D representation of the scene. If from that 3D representation a virtual view (not an available

0923-5965/$ - see front matter
doi:10.1016/j.image.2006.12.002

camera view) corresponding to an arbitrary viewpoint and viewing direction can be rendered, FVV functionality will have been achieved. In most cases, it will be necessary to restrict to some practical limits the navigation range (the allowed virtual viewpoints and viewing directions). Rendering a stereo pair corresponding to the human eyes from the 3D representation will not only provide 3D perception, but also support natural head motion parallax.

In contrast to stereoscopy, holographic approaches to 3DTV are less mature. However, advances in optical technology and computing power have now brought us to a point where holographic approaches can also be seriously considered. These are based on recording and reconstructing the whole optical wavefield emanating from an object or scene and thus aim to provide the truest 3D image possible. Research on holography has been continuing for several decades, but only with recently available computing power has it been possible to contemplate moving 3D images.

Although the final form 3DTV will take is not certain, current visions involve ghost-like but crisp moving video images floating in space or standing on a tabletop-like display. Viewers will be able to peek or walk around the images to see them from different angles or maybe even from behind. Undoubtedly, this will be of immense interest for consumer TV, games, virtual reality, and entertainment in general. Non-consumer applications are also diverse. In medicine, 3DTV images may aid diagnosis as well as surgery. In industry, they may aid design and prototyping of machines or products involving moving parts. In education and science, they may allow unmatched visualization capability. The possibilities are endless.

In recent years, tremendous efforts have been invested worldwide to develop convincing 3DTV systems, algorithms, and applications. This includes improvements over the whole processing chain, including image acquisition, 3D representation, compression, transmission, signal processing, interactive rendering, and display, as illustrated in Fig. 1. The overall design has to take into account the strong interrelations between the various subsystems. For instance, an interactive display that requires random access to 3D data will affect the performance of a coding scheme that is based on data prediction.

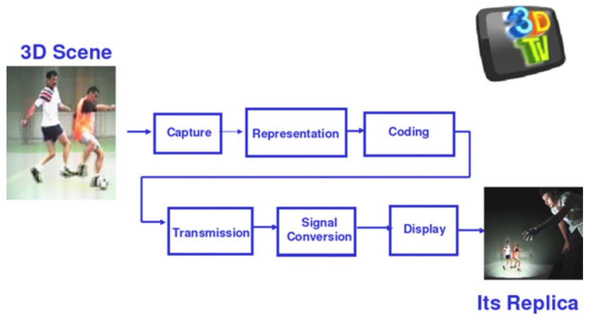

Fig. 1. Basic components of a 3DTV system. Reprinted with permission from L. Onural, H.M. Ozaktas, E. Stoykova, A. Gotchev, J. Watson, An overview of the holographic display related tasks within the European 3DTV project. In: Photon Management II: SPIE Proceedings 6187, SPIE, Bellingham, Washington, 2006, pp. 0T-1–0T-10.

The choice of a certain 3D scene representation format is of central importance for the design of any 3DTV system. On the one hand, it sets the requirements for acquisition and signal processing. On the other hand, it determines the rendering algorithms, degree of and mode of interactivity, as well as the need for and means of compression and transmission. Various 3D scene representations are already known from computer graphics and may be applied to 3DTV systems as well. These include different types of data representations, such as 3D mesh models, multiview video, per-pixel depth, or holographic data representations. Different capturing systems which may be considered include multi-camera systems, stereo cameras, lidar (depth) systems, or holographic cameras. Different advanced signal-processing algorithms may be involved on the sender side, including 3D geometry reconstruction, depth estimation, or segmentation, in order to transform the captured data into the selected 3D scene representation.

Specific compression algorithms need to be applied for the different data types. Transmission over different channels requires different strategies. The vast amount of data and user interaction for FVV functionality essential to many systems complicate this task even more. On the receiver side, the data needs to be decoded, rendered, and displayed. In many cases this may require specific signal conversion and display adaptation. Interactivity needs to be taken care of. Finally, the images need to be displayed. Autostereoscopic displays have already been mentioned, but there are also other types of displays, which aim to go even further. Such displays include volumetric displays, immersive displays and, of course, holographic displays. For many who are involved in the ambitious applications of 3D imaging, the fully interactive, full parallax, high-resolution holographic display is the ultimate goal. Whether or not this is achievable depends very much on the ability to efficiently handle the vast amounts of raw data required by a full holographic display and the ability to exploit the rapid developments in optical technologies.

The diversity of technologies necessary to make 3DTV a reality is very great. Successful realization of such products will require significant interdisciplinary work. Few, if any, researchers can claim to be experts in all of the areas involved. The purpose of this Special Issue is to provide a forum for some of the latest high-quality research results on various aspects of 3DTV. Of the 25 papers submitted to the issue, 9 were selected for inclusion.

Two of the papers deal with sender side signal processing or extraction. The paper "Towards 3D Scene Reconstruction from Broadcast Video" presents an efficient technique for generating 3D views using regular broadcast video, when the camera is stationary relative to the objects in the scene. Advanced algorithms for stereo reconstruction are presented in the paper "Graph-Cut-Based Stereo Matching Using Image Segmentation with Symmetrical Treatment of Occlusions".

Two papers target compression and transmission of 3DTV data. The first one, "3D Scanning-Based Compression Technique for Digital Hologram Video" falls within a rather pioneering field of research, contributing to the groundwork for future holographic video systems. The second one, "Stereoscopic Video Streaming with Content-Adaptive Rate and Format Control" targets efficient compression and transmission of classical stereo video.

Receiver side signal-processing issues are dealt with in three papers. The paper "Some Signal Processing Issues in Diffraction and Holographic 3DTV" reviews several key signal processing issues and tools, which are of importance in dealing with optical signals, including sampling and efficient numerical approaches. The paper "Diffraction Field Computation from Arbitrarily Distributed Data Points in Space" deals with forward and inverse problems in diffraction towards the goal of computing the optical field which should be presented to the observer in order to achieve 3D visualization by holographic means of the transmitted 3D object or scene. Efficient intermediate view synthesis is dealt with by the paper "Real Time View Synthesis from a Sparse Set of Views".

Finally, two papers elaborate on specific aspects of complete 3DTV systems. The paper "Towards High-Quality 3D Video of Dynamic Scenes" presents a system for active dynamic 3D scene reconstruction. A dynamic 3D point cloud is automatically captured and can be rendered from arbitrary viewpoints. A multiview video plus multiview depth system is presented in "Depth Map Creation and Image Based Rendering for Advanced 3DTV Services Providing Interoperability and Scalability", including capturing, data extraction, efficient 3D scene representation, and interactive rendering.

The guest editors would like to thank all authors who contributed their excellent research results to this Special Issue, and are grateful to all reviewers for their help and their fruitful suggestions, which have improved the quality of the manuscripts. We would also like to thank the Editor-in-Chief Murat Tekalp, for suggesting this Special Issue and for his support and guidance throughout the process. We hope that this Special Issue will provide interested readers a window into the state-of-the-art in 3DTV research and that it will inspire further research that will help make 3DTV a reality in the near future.

This Special Issue was motivated by, and some of the papers derived from work conducted within, the Integrated 3DTV—Capture, Transmission, and Display project, which is a Network of Excellence (NoE) funded by the European Commission 6th Framework Information Society Technologies Programme and led by Bilkent University, Ankara. The project involves 19 partner institutions from 7 countries and over 180 researchers throughout Europe and extends over the period from September 2004 to August 2008. The project web site is www.3dtv-research.org. We would like to take this opportunity to also thank project leader Levent Onural of Bilkent University for his encouragement of this Special Issue.

This work is supported by EC within FP6 under Grant 511568 with the acronym 3DTV.

Dr. M. Reha Civanlar is a Vice President and Director of the Media Lab in DoCoMo USA Labs. He was a visiting professor of Computer Engineering at Koc University in Istanbul for 4 years starting at 2002. He was also leading a multi-national European Research Project on 3DTV transport, and participating in numerous Turkish Industrial Boards. He is serving on the advisory boards of Argela Technologies Inc., on 3G multimedia systems and Layered Media Inc., on multipoint videoconferencing. Before Koc University, he was the head of the Visual Communications Research Department at AT&T Labs-Research, where he worked since 1991. In the same department, he also held Technology Consultant and Technology Leader positions before heading the group. Prior to that, Dr. Civanlar was at Pixel Machines Department of Bell Laboratories, where he worked on parallel architectures and algorithms for image and volume processing and scientific visualization. His career started as a researcher in the Center for Communications and Signal Processing of NCSU, where he worked on image processing, upon receiving his ECE PhD in 1984 from the same university. He received his BS and MS degrees in electrical engineering from METU, Turkey. He has numerous publications, several key contributions to the international multimedia communications standards, and over 40 patents either granted or pending. Dr. Civanlar is a recipient of 1985 Senior Award of the ASSP society of IEEE. His current research interests include packet video systems, networked video and multimedia applications with particular emphasis on the Internet and wireless systems, video coding, 3DTV and digital data transmission. Dr. Civanlar is a Fulbright scholar and a member of Sigma Xi. He served as an editor for IEEE Transactions on Communications and IEEE Transactions on Multimedia and JASP. He is currently an editor for EURASIP Image Communications. He served as a member of MMSP and MDSP technical committees of the Signal Processing Society of IEEE. Dr. Civanlar is a fellow of the IEEE.

Joern Ostermann studied Electrical Engineering and Communications Engineering at the University of Hannover and Imperial College London, respectively. He received Dipl.-Ing. and Dr.-Ing. degrees from the University of Hannover in 1988 and 1994, respectively. From 1988 till 1994, he worked as a Research Assistant at the Institut fuer Theoretische Nachrichtentechnik conducting research in low bit-rate and object-based analysis-synthesis video coding. In 1994 and 1995 he worked in the Visual Communications Research Department at AT&T Bell Labs on video coding. He was a member of Image Processing and Technology Research within AT&T Labs-Research from 1996 to 2003. Since 2003 he is Full Professor and Head of the Institut fuer Informationsverarbeitung at the Universitaet Hannover, Germany. From 1993 to 1994, he chaired the European COST 211 sim group coordinating research in low bitrate video coding. Within MPEG-4, he organized the evaluation of video tools to start defining the standard. He chaired the Adhoc Group on Coding of Arbitrarily shaped Objects in MPEG-4 Video. Joern was a scholar of the German National Foundation. In 1998, he received the AT&T Standards Recognition Award and the ISO award. He is a Fellow of the IEEE and member of the IEEE Technical Committee on Multimedia Signal Processing and past chair of the IEEE CAS Visual Signal Processing and Communications (VSPC) Technical Committee. Joern served as a Distinguished Lecturer of the IEEE CAS Society. He published more than 50 research papers and book chapters. He is coauthor of a graduate level textbook on video communications. He holds 22 patents. His current research interests are video coding and streaming, 3D modeling, face animation, and computer–human interfaces.

Haldun M. Ozaktas received a BS degree from Middle East Technical University, Ankara in 1987, and a PhD degree from Stanford University, California in 1991. He joined Bilkent University, Ankara in 1991, where he is presently Professor of Electrical Engineering. In 1992, he was at the University of Erlangen-Nurnberg, Bavaria as an Alexander von Humboldt Foundation Postdoctoral Fellow. Over the summer of 1994 he worked as a Consultant for Bell Laboratories, New Jersey. He is the author of over 80 refereed journal articles, one book, many book chapters, and over 80 conference presentations and papers, 25 of which have been invited. A total of over 2500 citations to his work are recorded in the Science Citation Index (ISI). He is the recipient of the 1998 ICO International Prize in Optics and one of the youngest recipients of the Scientific and Technical Research Council of Turkey Science Award (1999), among other awards and prizes. Haldun M. Ozaktas is also one of the youngest members of the Turkish Academy of Sciences and a Fellow of the Optical Society of America. His academic interests include optical information processing, signal and image processing, and optoelectronic and optically interconnected computing systems.

Aljoscha Smolic received the Dipl.-Ing. degree in electrical engineering from the Technical University of Berlin, Germany in 1996, and the Dr.-Ing. degree in electrical and information engineering from Aachen University of Technology (RWTH), Germany in 2001. He joined the Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institute (HHI), Berlin, Germany, as a student in 1994. In 1996, he became a Research Assistant and is employed as Project Manager since 2001. He has been involved in several national and international research projects, where he conducted research in various fields including video processing, video coding, computer vision and computer graphics and published more than 60 refereed papers in these fields. In this context he has been involved in ISO standardization activities where he contributed to the development of the multimedia standards MPEG-4 and MPEG-7. In current projects he is responsible for research in free viewpoint and 3D video processing and coding, augmented reality, 3D reconstruction and video-based rendering. Since 2003 he lectures at the

Technical University of Berlin in Multimedia Communications and Statistical Communications Theory. Dr. Smolic received the Rudolf-Urtlel-Award of the German Society for Technology in TV and Cinema (FKTG) for his dissertation in 2002. Frequently he serves as reviewer for various IEEE and other journals and conferences. He chairs the MPEG ad hoc group on 3DAV.

John Watson obtained a GradInstP degree in 1973 and a PhD from The University of St. Andrews in 1978. After 5 years with the UK Atomic Energy Authority he took up a lecturing post in Electrical and Electronic Engineering at Robert Gordons Institute of Technology, Aberdeen in 1981. He moved to the Department of Engineering at The University of Aberdeen in 1984 and (from Senior Lecturer and Reader) was appointed to a Professorial Chair in Optical Engineering in 2004. Professor Watson has published extensively in the scientific literature in his field of laser applications and optical engineering. His main research interests are in holography, underwater holography, laser welding and laser-induced spectroscopy. He has participated in and led many European research projects. He is a Fellow of the (UK) Institute of Physics, a Fellow of the (UK) Institution of Engineering and Technology (formerly IEE) and is a Senior Member of IEEE.

M. Reha Civanlar,
Media Labs, DoCoMo Labs USA,
Palo Alto, CA 94304, USA

Joern Ostermann,
Leibniz Universitaet Hannover,
Appelstr. 9A, 30167 Hannover, Germany

Haldun M. Ozaktas*,
Bilkent University, 06800 Bilkent, Ankara, Turkey
E-mail address: haldun@ee.bilkent.edu.tr

AljoschaSmolic,
Fraunhofer Institute for Telecommunications/
Heinrich-Hertz-Institut,
Einsteinufer 37, Berlin 10587, Germany

JohnWatson,
The University of Aberdeen, King's College,
Aberdeen AB24 3FX, UK

*Corresponding author.