AN OVERVIEW OF A NEW EUROPEAN CONSORTIUM: INTEGRATED THREE-DIMENSIONAL TELEVISION – CAPTURE, TRANSMISSION AND DISPLAY (3DTV)*

L. ONURAL

Electronics and Electrical Engineering Department
Bilkent University
TR-06800 Ankara, Turkey
onural@bilkent.edu.tr

T. SIKORA

Communications Department
Technical University of Berlin
Berlin, Germany
sikora@nue.tu-berlin.de

A. SMOLIC

Heinrich Hertz Institute
Fraunofer Institute
Berlin, Germany
Aljoscha.Smolic@hhi.fraunhofer.de

A new European consortium is formed as a Network of Excellence to integrate the research works of 19 institutions in the field of 3DTV. The consortium is funded by EC under the FP6 thematic area Information Society Technologies within the strategic objective Cross-media Content for Leisure and Entertainment. The project will last 48 months, but the collaboration among the partners is expected to be longer. The technical focus of the consortium is 3DTV with all its aspects except audio. Various techniques of 3D scene capture will be investigated and compared. Representation of captured 3D content in abstract form using mainly computer graphics approaches is the key feature which decouples user from the input.. Compression of 3D scene information, and forming the bitstream structure for effective streaming are parts of the project. The user may interact with the captured scene and get a visual display based on the choice of display technology. A rich variety of different display techniques, including stereoscopy and holographic displays are among the main focus of the consortium. The plan covers various integration and dissemination activities.

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

1. Objectives of the 3DTV NoE

The objectives of the 3DTV NoE are centered around the common theme: the recording, processing, interacting, and displaying 3D visual information. Capturing three-dimensional visual information of a real-life scene and creating an exact (except the scale) optical duplicate of it at a remote site instantaneously, or at a later time, are ultimate goals in visual communications. We identify the major components of the integrated system as: a) scene capture, b) scene representation, c) compression and coding of the scene, d) transmission (streaming), and e) display. Therefore, the topic is inherently multidisciplinary, and the researchers in the network reflect this diversity. The ultimate technical objective is a 3DTV system in its entirety, with high-end displays providing true 3D views. Figure 1 shows the artists impression of the ultimate 3D viewing experience targeted by the research.

Figure 1. Artists impression of a true 3D video display. (Courtesy of Erdem Yücel)

The primary objective of this project is to align the research works of participating European researchers along the 3DTV focus. The diverse experiences and activities of the researchers in distinct, yet complementary, areas encompass signal processing, with emphasis on multi-dimensional signals and video; computer graphics, communication and information theory; optics and physics. Naturally, the interest is all sorts of 3D visual data, and not only TV in its restrictive sense. That includes transmission modes through the internet, as well. The team targets a long term colaboration well beyond the life-time of the project. The project will create the highly needed synergy among the European partners at a critical time since 3DTV related research has been significantly accelerating throughout the world, and therefore will boost the European competitiveness in this field. Potential application areas, like haptics, telepresence, medical imaging, dentistry, museums, instruments, entertainment, video games will be examined.

Another important related item is the social impact of the 3DTV. It is believed that the 3DTV concept is well understood by a large majority of the public. Since simple stereo is a centuries old technique, we believe that most of the people had some experience with it. However, it is also a fact that the 3D mode is not the popular choice of the video technologies, today. While the 3D display is perceived as a highly futuristic desirable mode of real-time visual communications as implied in some popular motion pictures, there seems to be a lack of enthusiasm to incorporate 3D video in daily life. Reasons behind this seemingly contradictory situation are interesting. More than that, understanding those reasons are essential in overcoming the resistance toward 3D. There may be technical reasons like the lack of true 3D experience as a consequence of unnatural parallax associated with most of the simple stereoscopy that creates uncomfortable viewing conditions. But, there may be other psychological and social reasons, as well. Such issues will also be investigated within the NoE activities.

The essential efforts of the consortium will be naturally targeted to solve the technical problems associated with the general 3DTV topic explained above. It is not too difficult to identify some of the missing technological building blocks; for example, lack of practical fully electronic means of 3D scene capture and 3D scene display units is probably the main complicating factor. A more careful analysis would reveal that there are many other technological missing components. The group believes that the technological infrastructure to conduct the aimed research is ripe, and therefore, the jointly conducted research work will be productive with many innovative solutions that in turn might bring the 3DTV into common daily usage.

2. The work plan

The planned 48 months of work is rich with a number of activities aimed at different purposes.

As expected from an NoE, there will be major efforts for successful integration of the research works of partners. These include various small and large scale meetings, both technical and non-technical. The joint technical work is supported by an exchange program which will mobilize both graduate students and researchers. A number of research collaboration infrastructure activities are under way. For example, we will set-up and maintain software libraries to share developed software. A key effort will be on establishing a shared work environment with all its aspects, to facilitate collaboration.

The consortium believes that the success is based primarily on a strong organizational structure with carefully designed operational and administrative units equipped with the right amount of decision-making powers to cater the needs of the participants. Initial phase of the activities is concentrated to form the functional consortium bodies. At the time of writing of this paper, the formation of such bodies is almost complete. Furthermore, a strong management infrastructure is established with a staffed Project Management Office at Bilkent University who coordinates the 3DTV activities.

Dissemination efforts encompass a wide range of planned activities. In addition to expected publicity efforts, and technical paper publication in reputable conferences and journals, the consortium will publish an edited book on 3DTV, and design a graduate course for the same purpose. A web-site for public will be maintained throughout the project.

Naturally, the bulk of the efforts will be devoted to jointly conducted research work. Though a strong communication and effective collaboration among all partners dealing with different technical aspects of the problem are planned, there will be seven major workpackages each dealing with a different technical aspect of the 3DTV theme. The workpackages are further grouped into five Technical Committees whose number may change as the needs arise. TC1 will cover all techniques needed to capture moving 3D live scenes (a workpackage), and converting the captured scene into an abstract moving 3D scene representation (another workpackage). Fundamentally different approaches will be comparatively examined within the scope of the capture technologies. For example, video camera based techniques (single or multiple camera), as well as totally different holographic cameras, LIDAR based techniques, and object-mounted sensor arrays are all included. Some combinations of these different technologies may be very promising for practical purposes. Other techniques like photometric procedures will also be investigated.

It is targeted to achieve an abstract representation of the captured/constructed moving 3D scenery. The planned representation phase plays a fundamental role in the

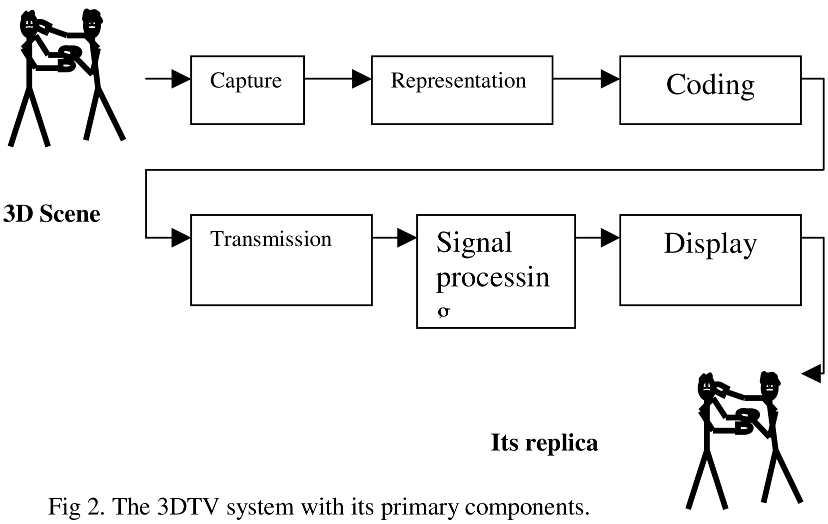

Fig 2. The 3DTV system with its primary components.

functionality of the overall system shown in Figure 2: The purpose of the "capture" part is to generate an input which is not necessarily transmitted to the user. Instead, the captured scene is used essentially to generate data to complete the representation. The user interacts with the representation only. Naturally, it is not necessary to have the "capture" and the "representation" physically in the same place; therefore, another transmission block might be added between these two blocks. The representation issues include mesh structures and generation, motion representation and animation, dense depth fields, volumetric techniques, etc. Human face and body specific techniques will get special attention.

TC2 will be dealing with 3D specific compression, coding, bitstream syntax, compatibility issues, intellectual property rights issues and related topics. TC3 will specifically target streaming protocols adapted to 3D video signals as represented. Information theoretic issues, as well as protocols are the main focus.

It is envisioned to have a system that can operate with a full range of display modes. Those users with a lower-end display (like the current 2D monitors) can still interact with the "3D representation" and chose an angle of view, zoom parameters, etc. to see it. However, those users with an ultimate high-end display, like the full holographic display, can view true 3D video. There is a wide range of different types of stereoscopic (or multiple-view) display techniques, in between. Therefore TC5 will be investigating different display technologies, together with applications of the 3DTV technique, since the applications will essentially be centered around the display. The research team has scientists who can also contribute to the very design of micro-mechanical devices, and optical structures to support novel display technologies.

Finally TC4 will be dealing with application specific signal processing techniques which operate on the abstract representation of captured motion 3D scenery, and convert that to those specific signals needed to drive the designated display. Furthermore, considering the propagating optical waves as a multi-dimensional signal to be generated by a holographic display device, the TC4 will also concentrate on the effects of various approximations at different stages to the subsequent optical quality of the generated true 3D scene by a given type of holographic display.

The consortium will consider standardization issues, as well as other commercialization issues right from the beginning of the work.

3. Conclusion

A European consortium, in the form of a Network of Excellence is formed and operational under the EC FP6. The team is strong and well motivated. A long-lasting collaboration with productive partnership and successful results are targeted.

4. References

1. L. Onural, G. Bozdağı and A. Atalar, ``New High-resolution Display Device for Holographic 3-D Video: Principles and Simulations'', {Optical Engineering}, vol 33, no 3, pp. 835 - 844, March 1994.

2. G. Bozdağı, A. M. Tekalp and L. Onural, ``3-D Motion Estimation and Wireframe Adaptation Including Photometric Effects for Model-Based Coding of Facial Image Sequences'', {IEEE Trans. On Circuits and Systems for Video Technology}, vol 4, no 3, pp. 246 - 256, June 1994.

3. A. A. Alatan and L. Onural, ``Estimation of Depth Fields Suitable for Video Compression using 3-D Structure and Motion of Objects" {IEEE Trans. on Image Processing}, vol 7, no 6, pp. 904 - 908, June 1998.

4. G. Bozdağı, A. M. Tekalp and L. Onural, ``3-D Motion Estimation and Wireframe Adaptation Including Photometric Effects for Model-Based Coding of Facial Image Sequences", {IEEE Trans. On Circuits and Systems for Video Technology}, vol 4, no 3, pp. 246 - 256, June 1994.

5. A. A. Alatan and L. Onural, ``Estimation of Depth Fields Suitable for Video Compression using 3-D Structure and Motion of Objects" {IEEE Trans. on Image Processing}, vol 7, no 6, pp. 904 - 908, June 1998.

6. G. Bozdağı, A. M. Tekalp and L. Onural, ``3-D Motion Estimation and Wireframe Adaptation Including Photometric Effects for Model-Based Coding of Facial Image Sequences", {IEEE Trans. On Circuits and Systems for Video Technology}, vol 4, no 3, pp. 246 - 256, June 1994.

7. A. A. Alatan and L. Onural, ``Estimation of Depth Fields Suitable for Video Compression using 3-D Structure and Motion of Objects" {IEEE Trans. on Image Processing}, vol 7, no 6, pp. 904 - 908, June 1998.

8. D.Tzovaras, N.Grammalidis and M.G.Strintzis: "Object-Based Coding of Stereo Image Sequences using Joint 3-D Motion/ Disparity Compensation", IEEE Trans. on Video Technology, Vol.7, No.2, pp.312-328, April 1997.

9. N.Ploskas, D. Simitopoulos, D.Tzovaras, G.A.Triantafyllidis and M.G.Strintzis: "*Rigid and Non-Rigid 3D Motion Estimation from Multiview Image Sequences*", Elsevier Signal Processing: Image Communication Journal, Vol. 18, Issue 3, pp.185-202, March 2003.

10. D.Tzovaras, N.Grammalidis and M.G.Strintzis: "3-D Camera Motion Estimation and Foreground / Background Separation for Stereoscopic Image Sequences", Optical Engineering, Vol.36, No.2, pp.574-580, February 1997.

11. M.G.Strintzis and S.Malassiotis: "*Object-Based Coding of Stereoscopic and 3D Image Sequences: A Review*", IEEE Signal Processing Magazine, Special Issue on Stereo and 3D Imaging (invited paper), Vol.16, No.3, pp.14-28, May 1999.

12. M.G.Strintzis and S.Malassiotis: "Optimal Biorthogonal Wavelet Decomposition of Wireframe Surfaces Using Box Splines, and its Application to the Hierarchical Coding of 3D Objects", IEEE Trans. on Image Processing, Vol.8, No.1, pp.41-58, January 1999.

13. D.Tzovaras, N.Grammalidis, S.Malassiotis and M.G.Strintzis: "*Coding for the Storage and Communication of 3-D Medical Data*", Image Communication, Vol.13, No.1, pp.65-87, July 1998.

14. S.Malassiotis and M.G.Strintzis: "Object-Based Coding of Stereo Image Sequences Using 3D Models", IEEE Trans. on Circuits and Systems for Video Technology, Vol.7, No.6, pp.892-906, December 1997.

15. I. Kompatsiaris, D.Tzovaras and M.G.Strintzis: "*3D Model-Based Segmentation of Videoconference Image Sequences*", IEEE Trans. On Circuits and Systems for Video Technology, Vol. 8, No. 5, September 1998.

16. Sexton I, Surman P, "Stereoscopic and Autostereoscopic Displays", IEEE Signal Processing, Vol. 16, No.3, May 99.

17. D. Dudley, W. Duncan, J. Slaughter, "Emerging Digital Micromirror Device (DMD) Applications", SPIE Proceedings vol. 4985, 2003.

18. L.J. Hornbeck, "Digital Light Processing and MEMS: Timely Convergence for Bright Future", Plenary Session, SPIE Micromachining and Microfacrication 95, Austin, Texas, October 1995.