IGLANCE: INTERACTIVE FREE VIEWPOINT FOR 3D TV

Svitlana Zinger[1], *Luat Do*[1], *Daniel Ruijters*[2], *Peter H. N. de With*[1]

[1]Eindhoven University of Technology, Video Coding & Architectures Research Group
[2]Philips Healthcare, Cardio/Vascular Innovation

ABSTRACT

The iGLANCE project aims at making interactive free viewpoint selection possible in 3D TV broadcasted media. This means that the viewer can select and interactively change the viewpoint of a stereoscopic streamed video. The interactivity is enabled by broadcasting a number of video streams from several viewpoints, consisting of a traditional 2D video and additionally depth information for each frame. Any desired view location in-between is generated by free viewpoint interpolation, using the depth information. The interpolated images are then displayed on a stereoscopic screen, giving a 3D impression to the audience. This article focuses on the broadcast chain, the decoding aspects and the view interpolation algorithm.

Index Terms — 3D video, view interpolation, autostereoscopic display, interactive free viewpoint.

1. INTRODUCTION

The iGLANCE project is part of the European MEDEA+ program. MEDEA+ is an industry-initiated pan-European program for advanced co-operative research and development in microelectronics. It has been set up and labeled within the framework of EUREKA (E! 2365) to ensure Europe's continued technological and industrial competitiveness in this sector. The iGLANCE consortium consists of STMicroelectronics, Logica, TIMA, INRIA and 4D View Solutions in France, and Philips Healthcare, the University of Technology Eindhoven, TASK24, Silicon Hive, Verum and Prodrive in the Netherlands. The iGLANCE project targets the healthcare and the consumer market. This article deals with the consumer branch of the project.

After high-definition television (HDTV), stereoscopic broadcasted and displayed content is expected to be the next innovation in the television market. Many movies are already recorded in a stereoscopic format today, and commercially available stereoscopic displays are emerging. The iGLANCE project aims at taking this innovation one step further, by introducing interactivity to the stereoscopic content, and allowing the viewer to freely select the viewing (camera) position of choice in a continuous fashion (as opposed to selecting from a discrete number of camera positions).

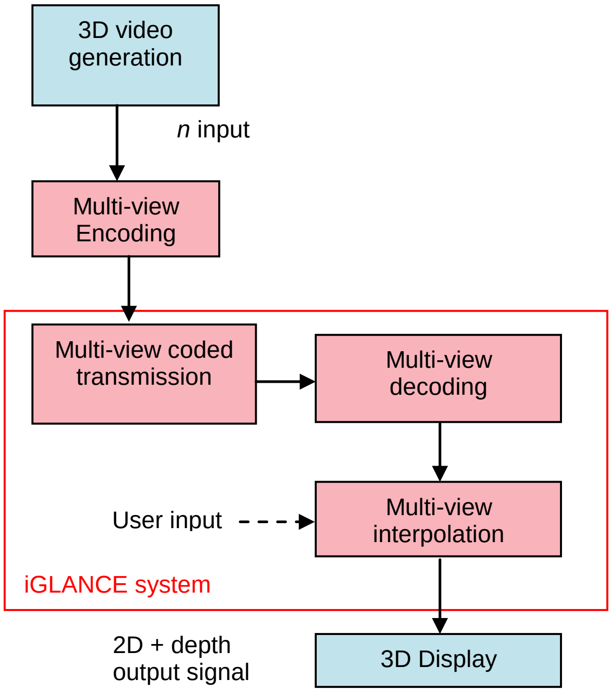

Figure 1. The iGLANCE transmission chain components.

It is the objective of the iGLANCE project to design a system that can make use of the infrastructure available today or in the nearby future (see Figure 1), to develop the algorithms necessary to perform this free viewpoint selection, and to design and implement a hardware solution for the real-time decoding of the video streams and processing of the free viewpoint algorithms.

2. STEREOSCOPIC DISPLAYS

A stereoscopic display presents the viewer with different images for the left and the right eye. Provided that these images contain proper stereoscopic information, the viewer will have the sensation of seeing depth. Principally there are two kinds of stereoscopic displays: the first type requires the viewer to wear goggles or glasses, and the second type, called auto-stereoscopic display, allows stereoscopic viewing without any external aid. While the iGLANCE project aims at supporting both kinds of displays, the primary focus is on the auto-stereoscopic category.

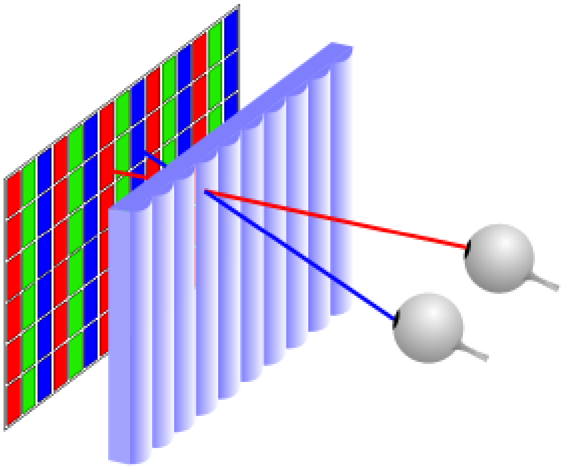

Figure 2. The auto-stereoscopic lenticular screen. The various sub-pixels are refracted to different angles by the sheet with the lenticular cylindrical lenses. In this way the left and the right eye are presented with different views.

The auto-stereoscopic displays typically use lenticular lenses (see Figure 2) or parallax barriers to achieve the stereoscopic effect. Modern auto-stereoscopic displays emit between 8 and 47 views in order to achieve a smooth transition when the viewer moves his head [1-3,10]. We currently target a lenticular display available on the market, which accepts an input signal that consists of a 2D image and a grey scale depth image which indicates the distance perpendicular to the screen for each pixel, see Figure 3. The onboard processing of the display then generates itself the individual stereoscopic views from this 2D + depth input signal. The advantage of this format is the fact that it decouples the input signal from the physical display properties, like lens configuration, number of views, angular interval between the views, etc [8].

3. FREE VIEWPOINT INTERPOLATION

Free viewpoint algorithms use smart interpolation routines to generate images for virtual camera locations, given a number of images depicting the same scene taken from different camera positions. Free viewpoint video applies this principle to video streams, whereby at least two synchronized video stream of the same scene are provided. The virtual camera location can be interactively chosen by the viewer (see Figure 4), e.g. by using a joystick as input device.

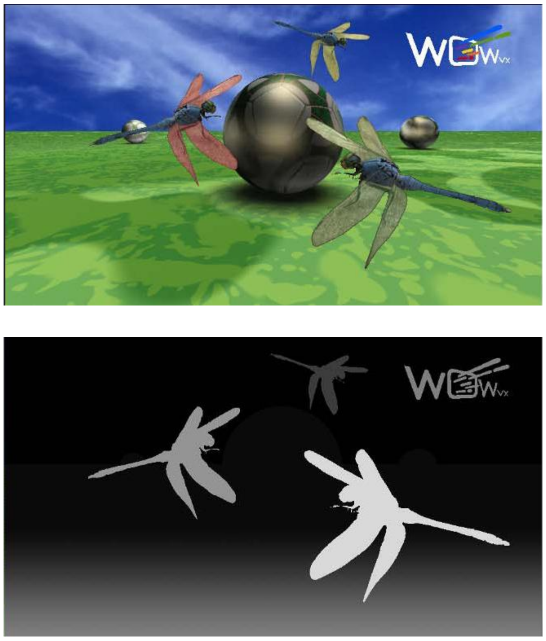

Figure 3. A 2D image (above) and its corresponding depth image (below) [8].

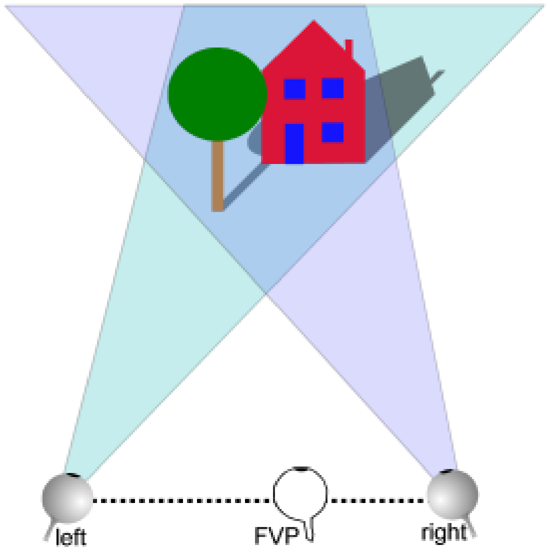

Figure 4. The free viewpoint can lie anywhere between two broadcasted views.

The input video streams are supposed to encode conventional video (texture) and a depth or disparity map for each frame in the video streams. Obtaining such depth images for real-world video sequences [9] is not an explicit objective of the iGLANCE project; the availability of such information is assumed as a precondition and can be obtained with off-line computations. For full 3D data, such as in virtual reality, computer games or medical volumetric data, the generation of the depth information is more or less trivial.

The challenge in applying free viewpoint video in interactive consumer applications lies in its real-time nature. The interpolated view has to be obtained with low-latency in order to provide proper interactivity. Furthermore low-latency is also of importance when viewing video from live events, like e.g. sports. In order to achieve this low-latency, the iGLANCE project aims both at using effective algorithms and developing dedicated hardware for the decoding and interpolation of the video streams. The project strives to stay within a maximum latency of 250 ms.

To make the output signal suitable for auto-stereoscopic displays, the interpolation algorithm generates a depth map along with the interpolated texture view.

4. ALGORITHM

Free viewpoint algorithms include multi-view image interpolation, image and depth map warping and their interpolation [6]. The fundamental challenge that remains in free viewpoint interpolation is displaying occluded parts of a scene [11]; in an interpolated view disocclusions may occur, which means that a part of the scene becomes visible that has been hidden in any of the transmitted views, see Figure 5.

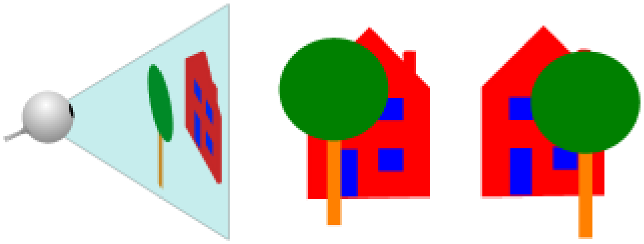

Figure 5. Depending on the viewing position different parts of a background object are occluded.

Our free viewpoint interpolation algorithm is based on calculating corresponding points from the neighboring transmitted views to the virtual viewing plane associated with the user defined free viewpoint [4]. The depth maps that accompany the video frames are back-projected to the virtual viewing plane [12], see Figure 6. For each pixel in the virtual plane, the original view that is least likely to be occluded is determined by taking the maximum from the corresponding pixels in the back-projected depth maps (higher depth values correspond to pixel content that is closer to the viewer).

The remaining pixels in the virtual depth map that are occluded in both views are filled in by taking the highest neighboring depth value.

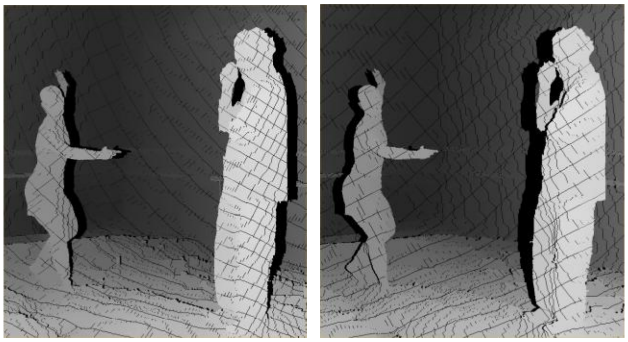

Figure 6. Back-projected depth maps from the original left and right views, using the Ballet Sequence from Microsoft Research [5]. Black pixels are due to rounding errors and disocclusions.

Since the edges of contours in the original textures and depth map are usually not sharp, ghost contours may occur. Especially when there is a large discontinuity between foreground and background depth, parts of the foreground can be projected to background locations. This effect is overcome by dilating the foreground objects in the joint depth map. As a result the foreground textures are not projected on the background anymore, see Figure 7, at the cost of some extra disocclusions.

Figure 7. Left: interpolated image with ghost effect. Right: without ghost effect, after the foreground depth has been dilated.

5. DISCUSSION AND CONCLUSIONS

The iGLANCE project aims for interactive free viewpoint video, displayed on a stereoscopic screen. To realize this objective, the project concentrates on the decoding and rendering of multi-view video with an arbitrary in-between viewpoint and considers also the hardware design of such a receiving platform.

Though stereoscopy is already an old technology (Sir Charles Wheatstone already developed a stereoscopic viewing device in 1838), the development and availability of stereoscopic displays for the mass market is currently enjoying an increasing momentum. Also many Hollywood productions are already being prepared for stereoscopic viewing.

The iGLANCE initiative nicely fits in these developments. The project does not rely on a single display technology, but focuses rather on the transmission and decoding chain and introduces interactivity through the freedom in choice of viewpoint. The central role of the interactivity distinguishes the iGLANCE project from existing research projects [7,9] and current commercial 3DTV solutions.

Working in close collaboration with internationally renowned industrial partners and research institutions, while developing software and hardware simultaneously, the iGLANCE project intends to achieve not only scientific but also relevant results for the industry.

6. ACKNOWLEDGEMENTS

This project is sponsored in the framework of the EUREKA (E! 2365) MEDEA+ program by the French and Dutch national research programs.

We would also like to acknowledge all the partners in the iGLANCE project, especially Michel Imbert, Regis Miquel and Emiliano Piccinelli from STMicroelectronics, Luc Verhaegh from Task24, Menno Lindwer and Aleksandar Beric from Silicon Hive, Egor Bondarev from TU/e, Nick de Koning and Dion Cornelissen from Prodrive, Henk Katerberg and Ladislau Posta from Verum, Sander Denissen, Willemien van der Linden and Paul Merkus from Philips, Florian Geffray from 4D View Solutions, Matthieu Millequant from Logica, Frédéric Pétrot from TIMA, Edmond Boyer and Radu Horaud from INRIA and all others who contributed.

7. REFERENCES

[1] C. van Berkel. Image Preparation for 3D-LCD. In *Proc. SPIE, Stereoscopic Displays and Virtual Reality Systems VI*, Vol. 3639, pages 84-91. 1999.

[2] N.A. Dodgson. Autostereo displays: 3D without glasses. *EID: Electronic Information Displays*. 1997

[3] D. Maupu., M.H. van Horn, S. Weeks, E. Bullit. 3D Stereo Interactive Medical Visualization. *IEEE Computer Graphics and Applications*, 25(5):67-71. IEEE, 2005.

[4] L. McMillan. An Image-Based Approach on Three-Dimensional Computer Graphics, PhD thesis, University of North Carolina at Chapel Hill. 1997.

[5] Microsoft Research, Video sequences with depth maps. http://research.microsoft.com/en-us/um/people/sbkang/3dvideodownload/

[6] Y. Morvan, D. Farin, P.H.N. de With. System architecture for free-viewpoint video and 3D-TV. *IEEE Transactions on Consumer Electronics*, 54:925-932. IEEE, 2008.

[7] 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 *Proc. SPIE, Vol 6187, Photon Management II*, Strasbourg, France, pages 61870T1-61870T10. 2006.

[8] Philips 3D Solutions. 2008. 3D Interface Specifications White Paper.

[9] A. Redert, M. op de Beeck, C. Fehn., W. IJsselsteijn, M. Pollefeys, L. van Gool, E. Ofek., I. Sexton, P. Surman. Advanced three-dimensional television system technologies. In *Proceedings of First International Symposium on 3D Data Processing Visualization and Transmission*, pages 313-319. 2002.

[10] D. Ruijters. Dynamic Resolution in GPU-Accelerated Volume Rendering to Autostereoscopic Multiview Lenticular Displays. *EURASIP Journal on Advances in Signal Processing*, vol. 2009, Article ID 843753, 8 pages. 2009.

[11] C. Vázquez, W.J. Tam. 3D-TV: Coding of Disocclusions for 2D+Depth Representation of MultiView Images. In *Proceedings of Tenth IASTED International Conference on Computer Graphics and Imaging (CGIM)*, Innsbruck, Austria, pages 26-32. 2008.

[12] P.H.N. de With, S. Zinger. Free-viewpoint rendering algorithm for 3D TV. In *Proceedings of the2nd International Workshop of Advances in Communication*, Boppard, Germany, pages 19-23. 2009.