IGLANCE: TRANSMISSION TO MEDICAL HIGH DEFINITION AUTOSTEREOSCOPIC DISPLAYS

Daniel Ruijters[1] and Svitlana Zinger[2]

[1]Philips Healthcare, Cardio/Vascular Innovation, Best, the Netherlands
[2]Eindhoven University of Technology, Video Coding and Architectures, Eindhoven, the Netherlands

ABSTRACT

The healthcare branch of the iGLANCE project aims at making high quality high definition autostereoscopic displays available in the clinical operating room. Displaying medical images on an autostereoscopic display poses different requirements than consumer usage would. For entertainment it is sufficient when the perceived image is convincing, even when deviating from the actual imaged scene. For medical usage it is essential that the perceived image represents the actual clinical data. The challenge that the iGLANCE project intends to address is the transmission of the autostereoscopic data through a bandwidth limited channel, while maintaining an image that does not contain significant image artifacts, like e.g. visible disocclusions.

Index Terms— Three-dimensional displays, Stereo vision, Biomedical imaging, HDTV, Interpolation

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 healthcare branch of the project. Displaying medical images in a clinical context, e.g. for diagnosis or treatment, imposes several restrictions on the image transmission chain. The error that is introduced (e.g. due to discretization, compression, transfer function, dynamic range limitations, etc.) has to be quantified and kept below a stringent threshold. The reason for this is obvious: medical decisions are taken based on these images, and flaws in the image might lead to misinterpretations.

The development of high resolution LCD grids (such as QuadHD grids) has brought high resolution autostereoscopic screens within reach. These screens, however, introduce a new challenge, since the amount of the visualized data becomes enormous, while the images have to be rendered and transmitted to the display in real-time. To cope with this, the iGLANCE project proposes the usage of view interpolation, using several 2D+depth views. This approach alleviates the rendering system, since fewer views need to be generated. Further, it reduces the strain on the transmission channel, because less data has to be transmitted, which is cost efficient and allows existing infrastructures to be used. In this paper we intend to describe the iGLANCE objectives regarding the usage of high definition autostereoscopic displays in a clinical environment.

2. THE AUTOSTEREOSCOPIC DISPLAY

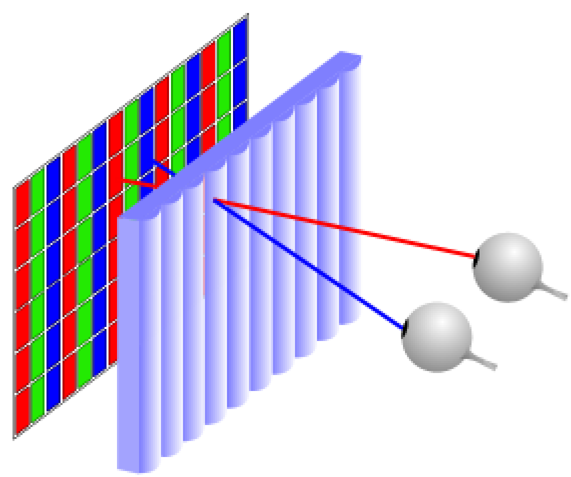

Figure 1. The autostereoscopic lenticular screen. The various sub-pixels are refracted to different angles by the sheet with the lenticular cylindrical lenses. In this way the left and the right eye are presented with different views.

Autostereoscopic displays allow a stereoscopic view of a 3D scene, without the use of any external aid, such as goggles. The additional depth impression that a stereoscopic image offers, allows a natural interpretation of 3D data. For the usage of such displays during medical interventions, the absence of goggles is a significant benefit, since there is no compromise of sterility by any external attributes and the goggles might be considered to be disturbing when the clinician is not looking at the stereoscopic display (which typically will be the major part of the clinical procedure) [1].

The multi-view autostereoscopic lenticular display consists of a cover sheet of cylindrical lenses (lenticulars) placed on top of an LCD, in such a way that the LCD image plane is positioned at the focal plane of the lenses [2], see Figure 1. As a consequence of this arrangement, different LCD pixels underneath the lenses become visible when viewed from various directions. Provided that these pixels are loaded with suitable stereo information, a 3D stereo effect is obtained, in which the left and right eye see different, but corresponding, information, see Figure 2 [3, 4]. A commercially available display line offers nine distinct views[1], but our technology will be applicable to any number of views.

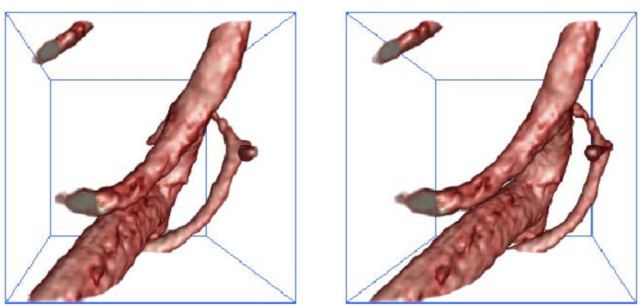

Figure 2. The same scene rendered from the most left and most right viewpoint.

The fact that mutually exclusive subsets of LCD pixels are assigned to different views (spatial multiplex), leads to a lower effective resolution per view than the intrinsic resolution of the LCD grid [5]. In order to distribute this loss of resolution over the horizontal and vertical axis, the lenticular cylindrical lenses are not placed vertically and parallel to the LCD column, but slanted at a small angle.

The resulting assignment of a set of interweaved LCD pixels is illustrated in Figure 3. Note that the red, green and blue color channels of a single pixel are depicted in different views, and that 9 different views are covered within a column of 3 pixels.

[1] Commercially available as the Philips WOWvx 3D display line

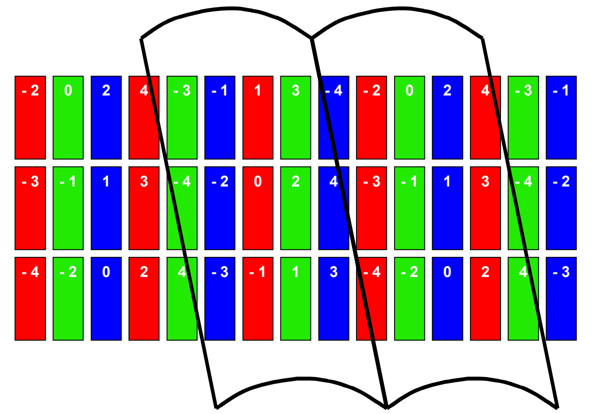

Figure 3. The cylindrical lenses depict every sub-pixel in a different view. The numbers in the sub-pixels indicate in which view they are visible.

3. VIEW INTERPOLATION

In order to reduce the load of the 3D rendering on the workstation and the transmission channel, we create fewer views than are displayed on the lenticular screen. The missing views are interpolated after decoding the video stream at the receiver side, see Figure 4. The interpolation method that is used is known from free viewpoint interpolation [6], see Figure 5. This method takes into account the camera parameters, and the depth information per pixel, which quantifies the distance between the screen and the object displayed at a particular pixel. The camera focus point and the pixel location define a ray in a virtual ray space [7]. By using this information more accurate interpolated views can be created than using a naïve interpolation method, such as linear interpolation.

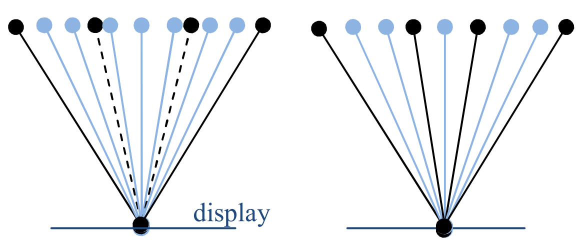

Figure 4. Two configurations for 4 transmitted views, and 9 displayed views. Solid black: transmitted views that can be mapped directly on an output view. Dashed: transmitted views that cannot be mapped on an output view. Light blue: interpolated view.

A QuadHD LCD grid consists of 3840 * 2160 pixels. Assuming that a nine view QuadHD autostereoscopic display is used, it would make sense to build up a single view in a resolution of 1280 * 720 pixels [8]. The views that are used for the interpolation algorithm can consist of 32 bits per pixel; 24 bits for RGB and 8 bits for depth information. Usage in clinical interventions requires a minimum frame rate of 24 frames per second (fps). When for example four views are transmitted at 24 fps, and the others are interpolated (see Figure 4), that would require a bandwidth of 4 views * 1280 * 720 pixels * 24 fps * 32 bits = 2.6 Gbit/s (for uncompressed video data) versus 4.4 Gbit/s for nine views without depth information.

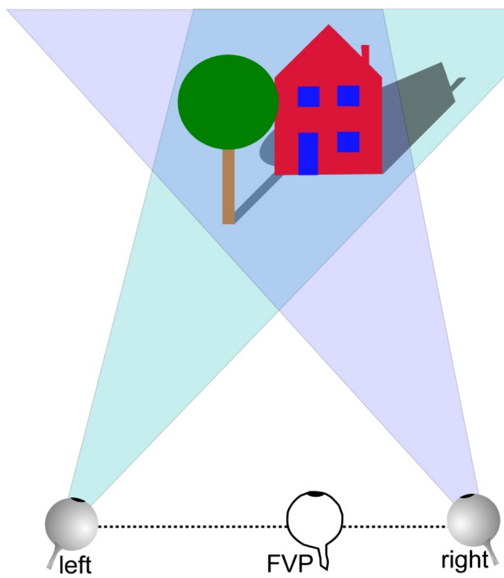

Figure 5. Free viewpoint (FVP) interpolation.

There is a particular type of artifact that may appear when using free viewpoint interpolation, which is the occurrence of disocclusions [9,10]. When this happens, a part of the scene becomes visible that has been hidden in any of the transmitted views, see Figure 6. Consequently, there is no proper information available that should be filled in at the affected pixels. Fortunately, the impact of this effect is very limited for our application, since the transmitted views and the interpolated views are very close to each other. Disocclusions mainly occur for views that are rather far apart, which is not the case for our setup.

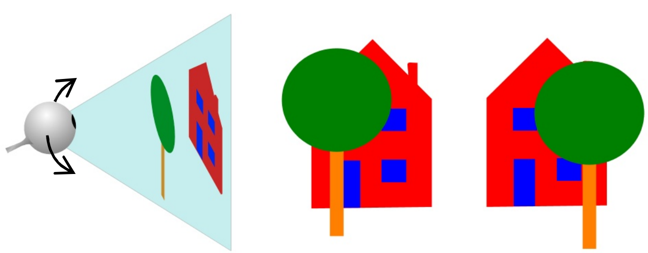

Figure 6. Depending on the viewing position different parts of a background object are occluded.

Another free viewpoint interpolation related artifact concerns semi-transparent parts of the depicted scene. The free viewpoint algorithm expects a single depth datum per pixel. However, when the object that is being depicted is semi-transparent, a pixel can contain visible information that is composed of the light that is reflected by several objects. Once the reflected light has been blended into a single pixel color, it is impossible to dissect it. In the rendering of 3D medical data this effect is even amplified, since the depicted data is often the result of a volume rendering process [3], whereby a ray of light traverses through a continuous range of semi-transparent material.

4. THE IGLANCE PIPELINE

The iGLANCE image rendering, transmission and displaying pipeline for the healthcare application is built up as follows (see Figure 7):

- Rendering multiple 2D+depth views
- Encoding the views into a video stream
- Transmission of the video stream
- Decoding the views
- View interpolation to generate the missing views
- Interweaving of the output signal
- Stereoscopic displaying

It should be noted that the iGLANCE project focuses mainly on the decoding, view interpolation and interweaving steps. It aims at developing dedicated hardware for these steps.

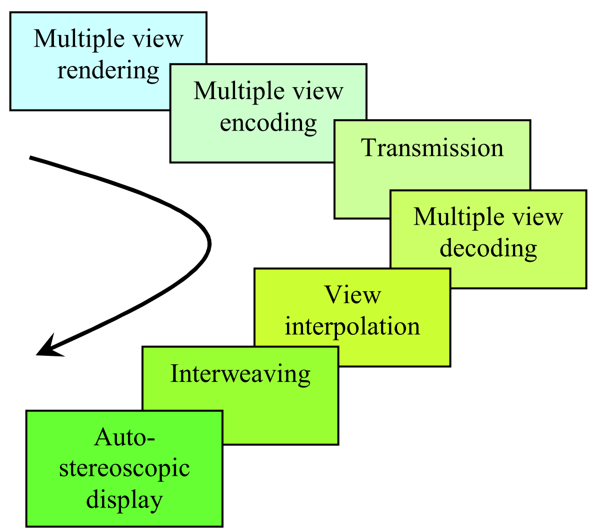

Figure 7. The iGLANCE pipeline for the healthcare branch.

It is the intention that the encoding, transmission and decoding steps use the H.264 codec, which involves lossy compression. The depth maps are then encoded as grayscale images, which puts all information in the hue components. Further efficiency can be reached by employing the multi-view extension of the H.264 codec [11]. The literature suggests that lossy compression ratios of 15:1 to 20:1 are still acceptable for medical images [12]. While it is necessary to use an efficient encoding strategy to fit the large amount of streaming data in a bandwidth limited channel, the usage of on-the-fly lossy compression in medical imaging should be evaluated carefully.

5. DISCUSSION

In this article we have presented the objectives of the iGLANCE project to achieve high definition autostereoscopic displaying of medical images during clinical interventions. The iGLANCE project focuses at creating algorithms, hardware and software design and a reference hardware implementation of the decoding and free viewpoint interpolation blocks of the proposed pipeline.

Free viewpoint interpolation algorithms are employed to 1) generate and 2) transmit fewer views. Rendering fewer views relieves the processing on the workstation, and therefore significantly aids in offering an interactive station that can visualize the requested data on-the-fly. Interactive and real-time 3D rendering is of the greatest importance for e.g. advanced minimally invasive clinical procedures. The bandwidth is further reduced by using H.264 compression. The reduction of the bandwidth requirements for the transmission channel will furthermore help to reduce the cost price.

The free viewpoint interpolations, as well as the H.264 compression may lead to image artifacts. Further investigations and quantification is needed to evaluate whether these artifacts are acceptable for medical usage.

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 and Olivier Muller from TIMA, Edmond Boyer and Radu Horaud from INRIA and all others who contributed.

7. REFERENCES

[1] D. Ruijters, "Integrating Autostereoscopic Multi-View Lenticular Displays in Minimally Invasive Angiography," *Proc. MICCAI AMI-ARCS workshop,* New York, USA, pp. 87-94, September 10, 2008.

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

[3] D. Ruijters, "GPU-Accelerated Volume Rendering to 3D Lenticular Displays," *IEEE Workshop on Content Generation and Coding for 3D-Television Vol. 1*, Eindhoven, the Netherlands, 2006.

[4] S. Zinger, D. Ruijters, P.H.N. de With, "iGLANCE project: free-viewpoint 3D video", *Proc. WSCG,* February, 2009.

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

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

[7] M. Tanimoto, "FTV (Free viewpoint TV) and creation of ray-based image engineering," *ECTI Trans. Electrical Eng., Electronics, Communications, Vol 6, No. 1,* pp. 3-14, Feb. 2008

[8] 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. doi:10.1155/2009/843753, 2009.

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

[10] S. Ince, and J. Konrad, "Geometry-based estimation of occlusions from video frame pairs," *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05)*, Philadelphia, USA, Vol. 2, pp. ii/933-ii/936, 2005.

[11] E. Martinian, A. Behrens, J. Xin, A. Vetro, H. Sun, "Extensions of H.264/AVC for Multiview Video Compression," *IEEE International Conference on Image Processing (ICIP),* pp. 2981-2984, October, 2006

[12] Y.-H. Shiao, T.-J. Chen, K.-S. Chuang, C.-H. Lin, and C.-C. Chuang, "Quality of Compressed Medical Images," *Journal of Digital Imaging, Vol. 20, No. 2,* pp. 149-159, June 2007.