Fraunhofer Heinrich-Hertz-Institut

Real-Time Video Processing for Autostereoscopic 3D Displays

White Paper

Title: Real-Time Video Processing for Autostereoscopic 3D Displays

Authors: R. de la Barré, S. Jurk, P. Kauff

Date: January 8, 2018

Version: 1.0

Table of Contents

Table of Figures

1 Introduction

At autostereoscopic 3D displays, the ergonomics and user-friendliness of the viewing situation must be taken into account. In addition to the size and resolution of the display and its ability to show the depth, this also includes the user (s)'s viewing distance and the depth he/she accepts individually. Such displays are called adaptive, intelligent and sometimes mathematical. In the case of commercial mass-produced products, this adaptability must be a feature of the display. It must therefore be integrated into the display interface. Researchers at the HHI developed a number of required modules such as depth map estimators, automatic stereo analysis, eye-tracking algorithms and a number of rendering solutions for special 3D presentation modes. These components can preferably be implemented on display controller hardware, but they can also be run on a computer's graphics controller unit to implement the functions and display modes. This whitepaper reports on them.

2 Real-time processing chain

The presented system covers the whole real-time processing chain from a given stereo stream to its visualization at arbitrary autostereoscopic (i.e. glasses-free) 3D displays. Figure 1 shows a block diagram of the existing processing chain with multifold output options. The 3D content can be shown on either tracked single-user displays or multi-user multi-view displays as well as on upcoming integral imaging displays. For this purpose the modules can flexibly be configured such that the same processing and system architecture supports a multitude of different 3D display formats and operating modes. A further outstanding feature is its future-proof performance enabling perfect quality at auto-stereoscopic 3D display panels with very high resolution (4k and beyond) and extremely small pixel pitches. In summary, our solutions allow highest quality for a wide range of 3D applications by only one common and cost-effective processing system.

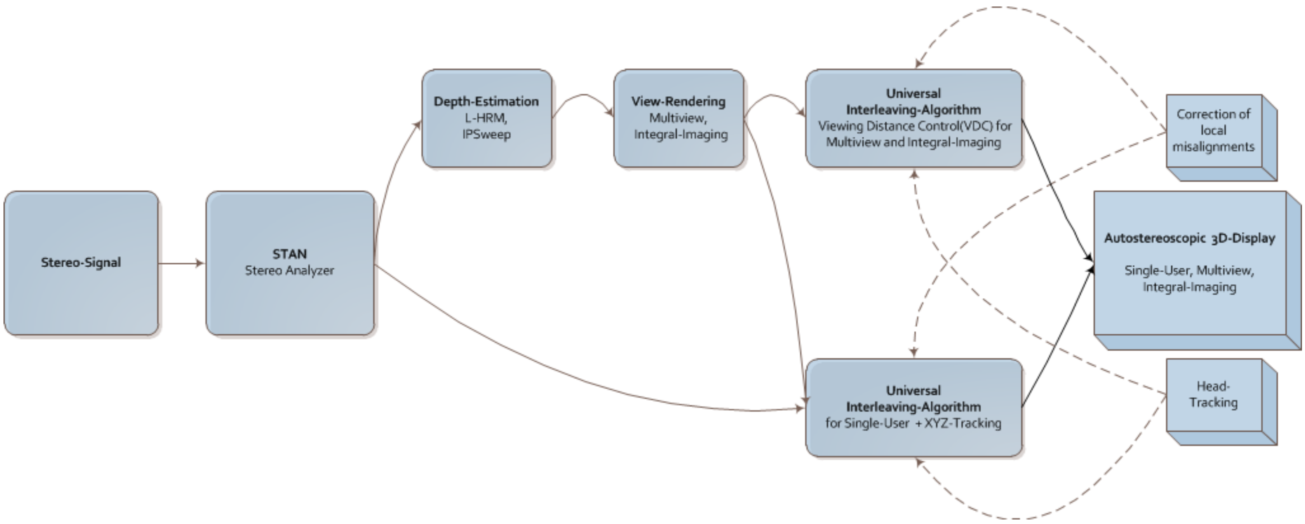

Figure 1. Block diagram of the existing process chain with multifold output options.

Additional input components allow the decompression of video streams encoded in HEVC, for example, and enable 3D display systems compatible with many transmission paths.

3 Enabling image processing modules

3.1 STAN – Stereoscopic Analyzer

The Stereoscopic Analyzer (STAN) is used as the initial processing step. It ensures that the stereo stream respects basic quality requirements such as line alignment, color matching, avoidance of perception conflicts and compliance with comfortable viewing conditions. It controls the stereo quality and corrects the stereo input in real-time if necessary. Figure 2 shows the basic idea of this software tool, which is also implemented in hardware.

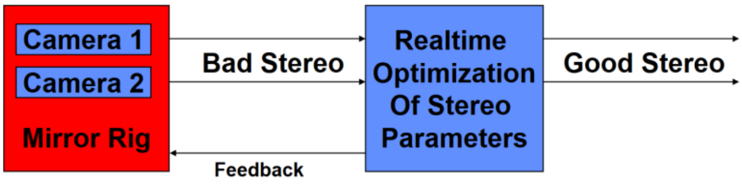

Figure 2. Basic idea of the STAN.

In summary, the STAN is a system for producing perfect stereo by combining real-time image analysis with intelligent automated tools. [1-4, 21]

3.2 Depth estimation

Based on the corrected stereo input from STAN output, depth information is extracted from the stereoscopic video by using a hybrid recursive matching (HRM) technique [5]. The current system of the processing chain uses a fast, robust and efficient version called L-HRM. It enables an estimation of disparity maps in real time [6]. Furthermore texture-adaptive multi-lateral post-filtering detects and corrects mismatches, resulting in reliable disparity maps of highest quality [7-9, 22].

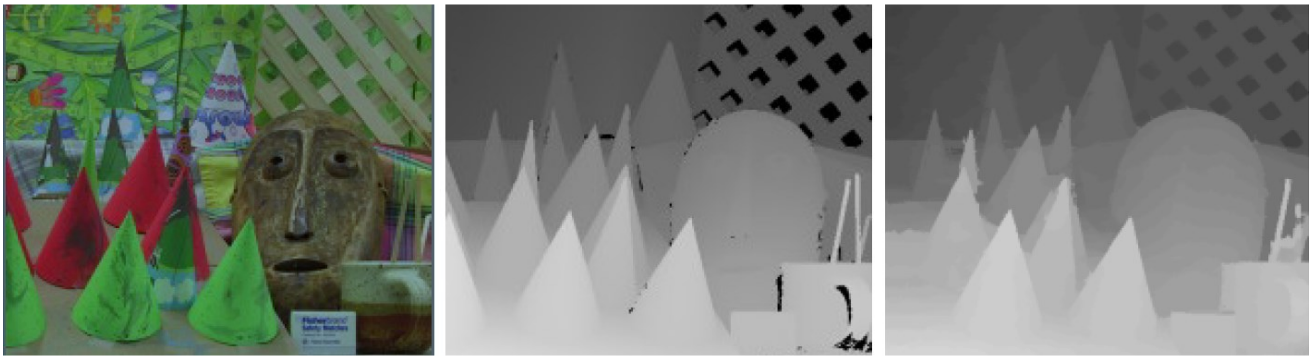

Figure 3. Resulting disparity map of a Middlebury test image. (left image / ground truth / disparity map estimated by HHI algorithm)

Currently, a new enhanced depth estimator is under development and will be integrated soon. The most significant advantage of this new so-called IDSweep algorithm is the furtherly improved real-time performance based on a high degree of parallelization especially suitable for GPU implementations [10].

3.3 View rendering

Using the depth information from L-HRM, a depth-image based renderer (DIBR) calculates all virtual views that are needed for the specific 3D display to be supported [11]. The view renderer is optimized for autostereoscopic 3D displays and runs in real time. Number and position of the virtual views (see Figure 5) depends on the display design and its operating mode.

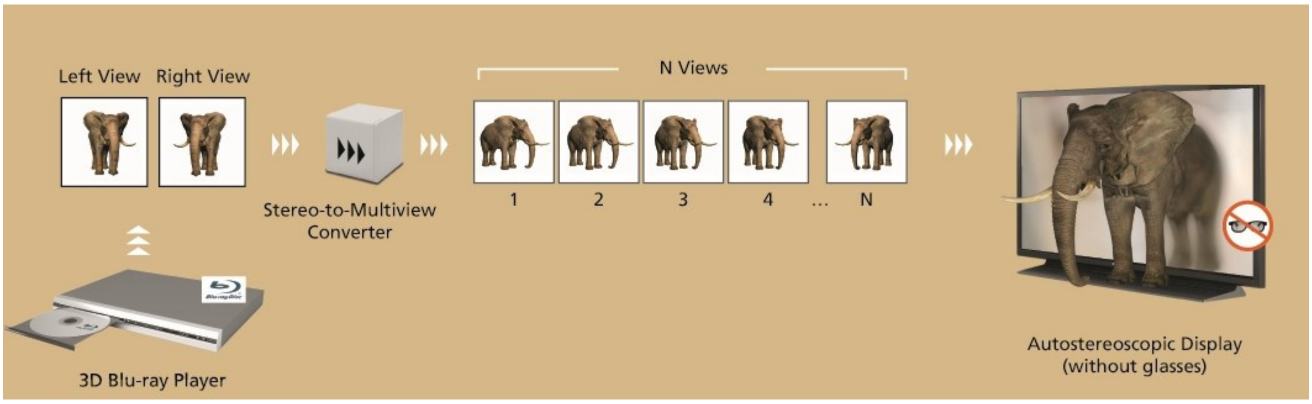

Figure 4. Conversion of generic Stereoscopic 3D Video to Multiview by generating virtual camera images.

The renderer has a variety of options to be configured accordingly, e. g. as shown in Figure 4. Thus, it represents the core component of the entire system allowing the high flexibility to support any kind of display type and to be able to switch on-the-fly between different operating modes (e.g. tracked single-user, multi-view or integral imaging). Figure 5 illustrates a section of the processing chain.

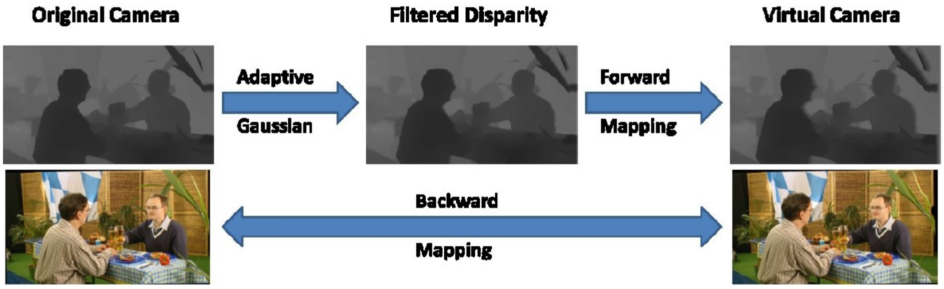

Figure 5. View interpolation using image and disparity for rendering.

Moreover, it allows to calibrate the alignment of display panel and optical lens precisely to each other, just by electronical processing – an important feature, especially in case of high resolution panels with extremely small pixel or sub-pixel pitch. [15]

Specifications:

- Highly real-time optimized core algorithms like L-HRM, SKB, STAN
- All core algorithms are designed for a maximum of parallelization
- Real-time implementation on 2 Intel Hexa-Core CPUs and Nvidia GeForce GTX 590 GPU
- Support for all commercially available autostereoscopic 3D displays

3.4 Head tracking

Our head tracking technology detects the eye positions of several users and calculates their positions. The tracking algorithm is able to work with both video and stereoscopic video. In a first step, the faces of potential users are identified by software algorithms. In the second step, when faces were found, a quick pattern tracking is performed. Depending on the camera resolution, typical facial features such as eyebrows, iris and nostrils can be detected and stored to provide reference points for tracking. Robust tracking performance is based on recording high-quality video images.

The tracking software can be flexibly implemented for different numbers of video cameras and recording frequencies. Figure 6 shows such an example. At stereoscopic camera images, similar shots are necessary. For this reason, an exposure process changes camera parameters when the pattern quality is poor. [13]

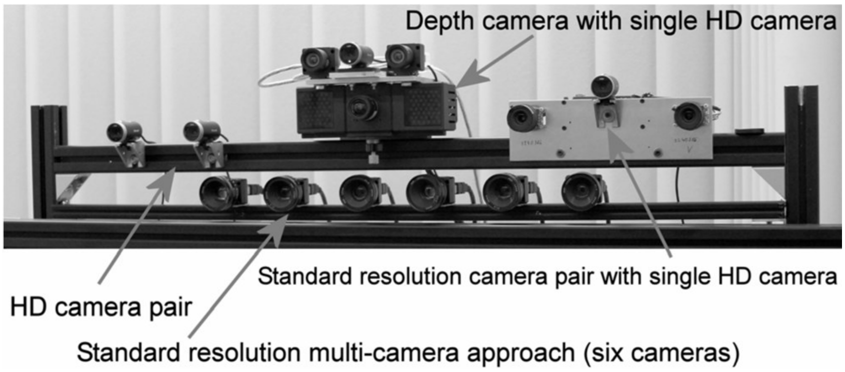

Figure 6. Various implementations of the HHI tracking software.

The tracking provides the position coordinates of the eyes to the view rendering and, if necessary, to the software application. This allows both adaptation of user position and a dynamic perspective of the presented scene.

3.5 Correction of misalignments

Autostereoscopic displays require significantly higher panel resolution than other display devices. In this way, they compensate for the resolution losses resulting from the directed emission of several views. The following applies: The higher the panel resolution, the higher the pixel density, the smaller the size of the pixel structures. The latter causes that the distance between display and image distributor reduces. This also increases the influence of alignment errors of the optical image splitter and the pixel structure. The visible effects of such dislocations and distance fluctuations can be detected and corrected in real time by our image processing system (see also Figure 7). [19]

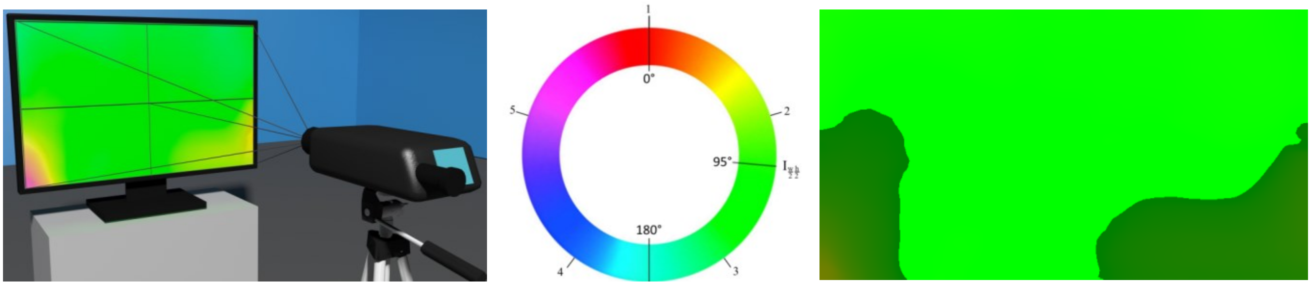

Figure 7. Left: capturing by color spectro-meter using a colored test pattern. **Center: HSV color circle split for calculation of alignment adjustment.** **Right: final correction map coded with shift values and shift directions.**

Main features are:

- Flexible method for the correction of allocation errors between the display panel and the image splitter (e.g. lenticular)
- Applicable to single-user, multi-view and integral imaging displays
- Errors caused by inaccurate bonding can be compensated
- Reduced fabrication costs due to electrical adaptation instead of mechanical calibration

4 Presentation modes for autostereoscopic 3D displays

Based on the idea of an almost continuous content shift, electronic x-y-z adaptation methods were developed at the HHI. They were described in [18]. Their further development [24] serves as a basis for adapting the image representation to the position of the user. In addition, image processing algorithms have been developed to combine several display formats and operating modes such as Single User [23, 24], Multi-View [25-30] and Integral Imaging [31, 32]. They share the same image reproduction process by using different on-board processing configurations on the same 3D display without changing the properties of the optical lenses. The scheme in Figure 8 shows the interplay of the associated HHI patent literature and its interaction with different types of autostereoscopic displays.

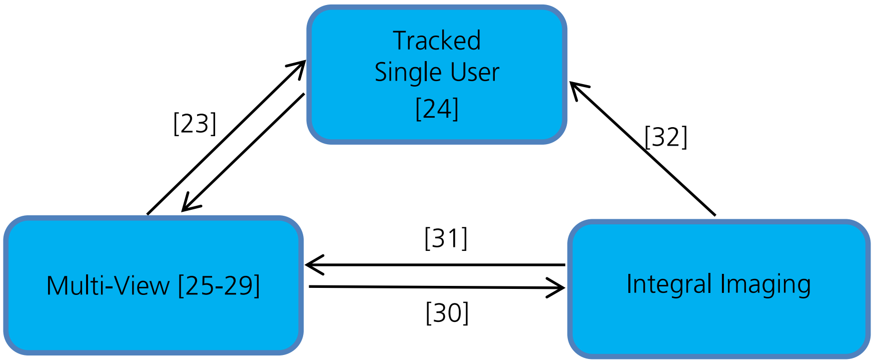

Figure 8. Options for transforming the properties of glasses-free 3D displays by electronic image processing.

Autostereoscopic 3D displays use arbitrary horizontal or slanted optical image splitter. For instance, a multi-view 3D display with lenticular lenses can be converted to an integral imaging device, just by electronical means without modifying the optics themselves and vice versa. 3D displays with tracking units can be switched on-the-fly from tracked single-user mode to a multi-view multi-user display and vice versa. In addition, the technique allows for adapting the viewing zones to a tracked user position or desired viewing distances in case of multi-view multi-user displays or integral imaging displays [20].

4.1 Universal interleaving-algorithm for adaptive two-view mode

This adaptive basic mode is a generic method for interleaving of two stereo images for each given autostereoscopic 3D display design and allows a precise electronical calibration between display panel and optical lens. It can be used for autostereoscopic single workstations, multi-view and integral screens with any inclination and division of the image splitter elements [12, 14, 23, 24, 32]. A line by line processing of the continuous subpixel shift illustrates Figure 9, which shows the treatment of image data for a line segment of a spatially multiplexed, eyeglass-free 3D display with several subpixels per left and right pixel within the line.

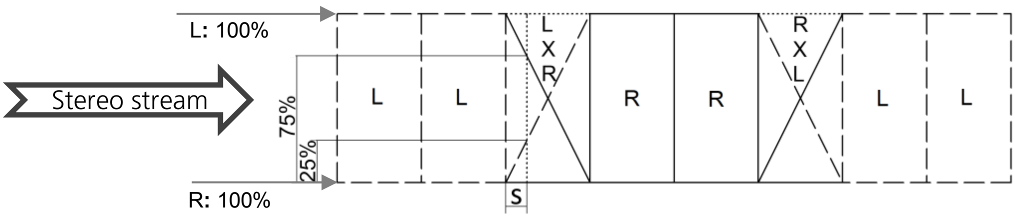

Figure 9. Situation at a two-channel playback displayed at a linear shift S by 25% of a subpixel width. The shift parameter S determines the mixing ratio between the active left and right subpixel.

Main features are:

- Universal interleaving algorithm for autostereoscopic 3D displays with subpixel reserve
- Adjustment of the viewing zone in the room before the 3D display according to a tracking signal
- Local correction of misalignments between display-panel and lens grid
- Hardware implementation on FPGA board for 5K Display

4.2 Discrete and continuous multiview via Viewing-Distance-Control (VDC)

The Viewing Distance Control (VCD) shown in Figure 10 is a generic method that adapts the interleaving pattern to the desired viewing distance depending on the optical design of a given autostereoscopic 3D display [25, 26]. For each subpixel position, it uses newly calculated render angles resulting from the desired viewing distance and the 3D display properties. The division into discrete and continuous multiview is carried out, as both types of multiview representation can actually be performed with this method, irrespective of how many input images are present and how many beam directions the display operates [15-17].

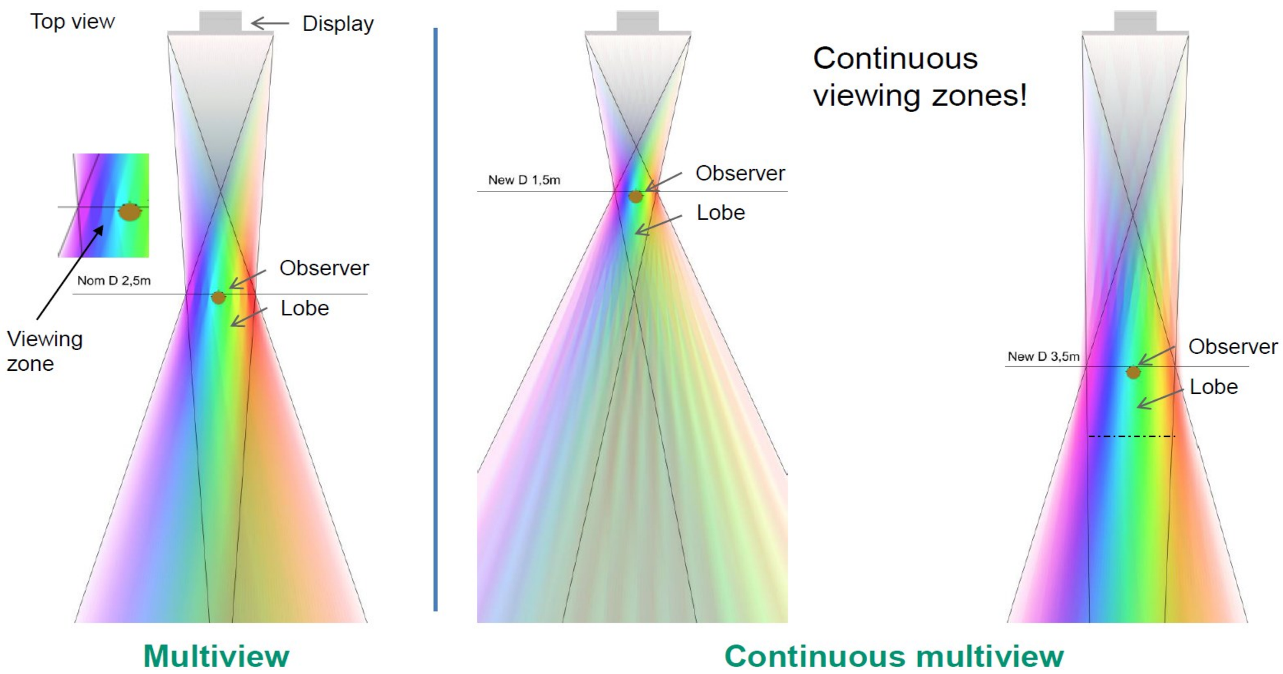

Figure 10. Multiview display left with separate stereo zones (discrete multiview) and right with continuous adaptation of viewing zones to changing viewing distances by VDC (center, right)

The possibilities in a glance:

- Flexible technology for all autostereoscopic 3D displays with more than two views
- Adaptation of optimum viewing distance in front of or behind the nominal distance
- Continuous lateral displacement of the viewing zone possible
- Lateral adjustment of the position of the viewing zones is possible by additional views.
- Continuous change of perspective at a the lateral head movement
- Changing the number of visible zones by modulating the period width of switchable lenses [32]
- Image content can be produced by depth-image based rendering (DIBR) or computer generated imagery (CGI)

4.3 Switching an integral display to multi-view display mode via VDC

This mode enables the conversion of a 3D display with perspective projection (integral imaging) into a multi-view 3D display with central perspective projection. This is done using a similar processing as described in section 4.2 [27]. Figure 11 shows the effect of both modes on the stereo lobe and indicates the flexibility of this design approach.

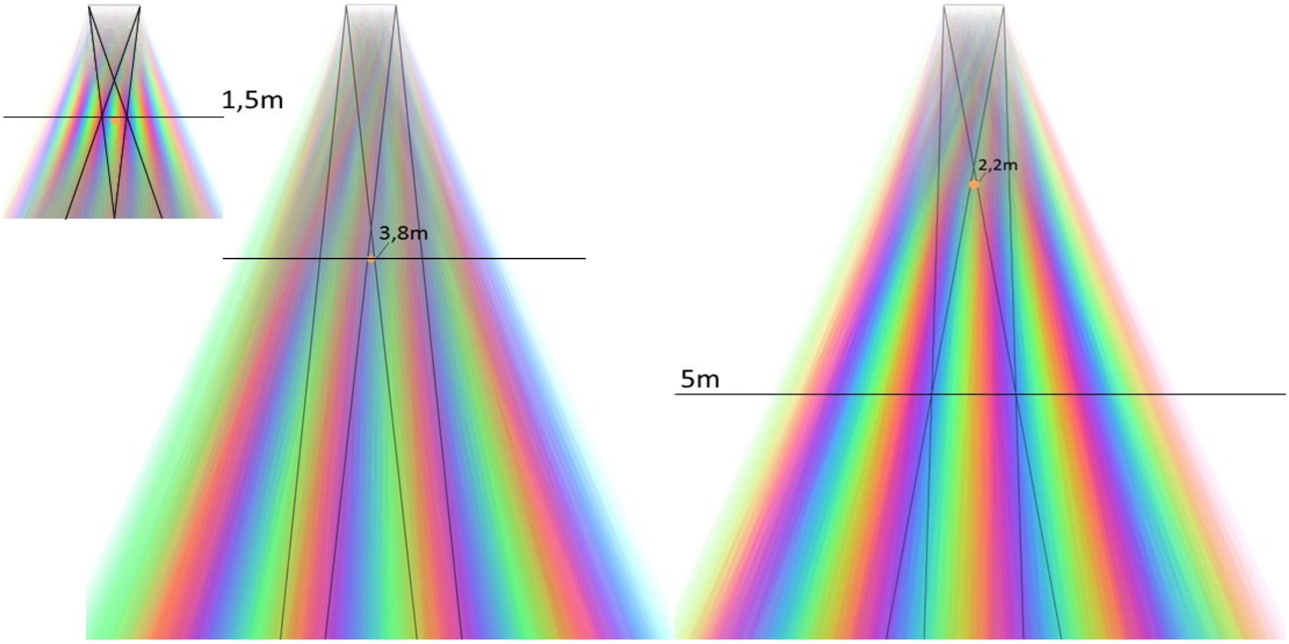

Figure 11. Example for the viewing space at same 3D display: Multi-view mode (left), conventional integral imaging (center) and optimal trade-off (right) for the desired viewing condition

4.4 Depth enhanced integral imaging display

In contrast to section 4.3, the opposite conversion direction causes a transformation of a multi-view 3D display with perspective projection into an integral-imaging 3D display with parallel projection. This special application has the advantage of finding an optimal compromise between both features, high depth of field and wide viewing angle for a desired viewing condition [31].

The advantages of this design approach are obvious:

- Increased depth of field compared to conventional integral imaging displays
- Reduction of the blurring effect
- No adaptation of the viewing area in Z-direction necessary
- Allows a reduction of the minimum viewing distance
- Allows the correct 3D perception at different viewing distances

4.5 Multi-modal Super Hi-Vision 3D representation

Super High-Vision panels with resolutions of 8k or 16k or even beyond will be available in near future. Due to these high resolutions autostereoscopic 3D display (multi-view or integral imaging) will be able to offer completely new features. In particular, it might be feasible to create a high number of separated stereo channels with wide viewing angles. Hence, in combination with tracking technologies, several persons at different viewing positions can watch individual stereo views under optimal conditions, even with different content. Such viewing formats are known as multimodal representations. [14, 18, 20]

Our universal processing chain is also able to support these future applications. One of the unique possibilities of these multimodal presentation modes is shown in Figure 12, where each viewer uses his or her own individually selected program with a different display mode.

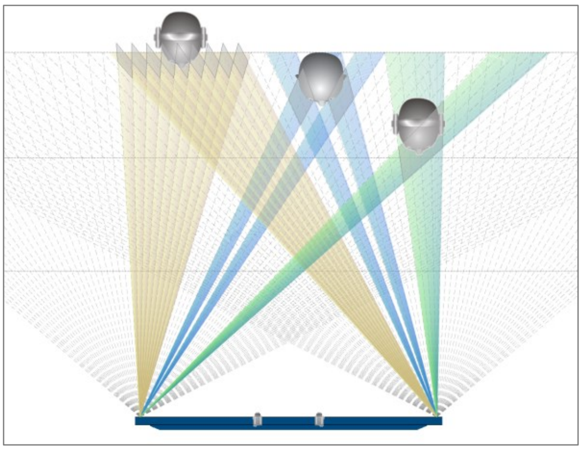

Figure 12. Multiple display modes with different contents, played back simultaneously on a 3DTV-set with 32 views (left: 3D multiview mode, center: 3D stereo mode, right: 2D mode).

Here are a few examples:

- Multiview 3D displays with mixed viewing zones, e.g. a tracked stereo zone and a wide angle 2D viewing zones for special application like surgery [23]
- Multiview 3D displays with a few tracked stereo zones and minimized crosstalk [28]
- Multi-Content representation where several people can watch with tracked stereo views of different content [29]
- Transforming an integral image display with parallel radiation geometry into a tracked multi-user display with two matching views each [30]

References

[1] Frederik Zilly, Peter Eisert, and Peter Kauff, "Real-Time Analysis and Correction of Stereoscopic HDTV Sequences", 6th European Conference on Visual Media Production (CVMP 2009), London, UK, November 2009, Short Paper.

[2] Frederik Zilly, Marcus Müller, Peter Eisert, and Peter Kauff, "The Stereoscopic Analyzer - An Image-Based Assistance Tool for Stereo Shooting and 3D Production", Proceedings of International Conference on Image Processing (ICIP10), Hong Kong, pp. 4029-4032, 2010.

[3] Frederik Zilly, Marcus Müller, Peter Kauff, and Ralf Schäfer, "STAN — An assistance system for 3D productions: From bad stereo to good stereo", Proceedings of 14th ITG Conference on Electronic Media Technology (Dortmunder Fernsehseminar), Dortmund, Germany, 2011.

[4] Frederik Zilly, Christian Riechert, Peter Eisert, and Peter Kauff, "Semantic Kernels Binarized – A Feature Descriptor for Fast and Robust Matching", Proceedings of European Conference on Visual Media Production (CVMP), London, UK, 2011.

[5] Nicole Atzpadin, Peter Kauff, and Oliver Schreer, "Stereo Analysis by Hybrid Recursive Matching for Real-Time Immersive Video Conferencing", IEEE Transactions on Circuits and Systems for Video Technology, special Issue on Immersive Telecommunications, vol. 14, no. 3, pp. 321-334, 2004.

[6] Christian Riechert, Frederik Zilly, Marcus Müller, and Peter Kauff, "Real-Time Disparity Estimation Using Line-Wise Hybrid Recursive Matching and Cross-Bilateral Median Up-Sampling", International Conference on Pattern Recognition (ICPR 2012), Japan, 2012.

[7] Christian Riechert, Frederik Zilly, Marcus Müller, and Peter Kauff, "Advanced Interpolation Filters for Depth Image Based Rendering", 3D TV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), Zürich, Switzerland, 2012.

[8] Marcus Müller, Frederik Zilly, Christian Riechert, and Peter Kauff, "Spatio-temporal consistent depth maps from multi-view video", 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), pp. 1-4, 2011.

[9] Christian Riechert, Frederik Zilly, Marcus Müller, and Peter Kauff, "Real-Time Disparity Estimation Using Line-Wise Hybrid Recursive Matching and Cross-Bilateral Median Up-Sampling", International Conference on Pattern Recognition (ICPR 2012), Tsukuba Science City, Japan, November 2012.

[10] W. Waizenegger, I. Feldmann, O. Schreer, P. Kauff and P. Eisert, "REAL-TIME 3D BODY RECONSTRUCTION FOR IMMERSIVE TV", Int. Conf. on Image Processing, ICIP 2016, Phoenix, Arizona, USA, September 2016.

[11] Christian Riechert, Frederik Zilly, Peter Kauff, Jens Güther, and Ralf Schäfer, "Fully Automatic Stereo-to-Multiview Conversion in autostereoscopic Displays", IBC 2012, Amsterdam, The Netherlands, September 2012.

[12] Silvio Jurk and René de la Barré, "A new Tracking Approach for XYZ-User-Adaptation of Stereoscopic Content", Electronic Displays 2014, Electronic Proceedings ISBN 978-3-645-50132-3, Nuremberg, 2014.

[13] Klaus Hopf, Frank Neumann and David Przewozny, „Multi-user eye tracking suitable for 3D display applications" 3DTV-CON 2011, DOI: 10.1109/3DTV.2011.5877222, 2011.

[14] Silvio Jurk and Rene de la Barré, "Improvements of aid free 3D Presentation", Electronic Displays Conference 2014, Electronic Proceedings ISBN 978-3-7723-0862-8, Nuremberg, 2011.

[15] Silvio Jurk, Mathias Kuhlmey and René de la Barré, "A new Type of Multiview Display", IS&T/SPIE Electronic Imaging 2015, Stereoscopic Displays and Applications XXVI, DOI: 10.1117/12.2081005, San Francisco, California, USA, Volume: 9391, 2015.

[16] Silvio Jurk, Bernd Duckstein, Sylvain Renault, Mathias Kuhlmey, René de la Barré and Thomas Ebner, "Enhancing multi-view autostereoscopic displays by viewing distance control (VDC)", Proceedings Volume 9011, Stereoscopic Displays and Applications XXV; 90111C (2014); DOI: 10.1117/12.2040755, IS&T/SPIE Electronic Imaging 2014.

[17] René de la Barré, "Viewpoint Adaption for Multiview 3D Displays", Web-page, Fraunhofer HHI, 2017. https://www.hhi.fraunhofer.de/en/departments/vit/research-groups/capture-display-systems/research-topics/multiview-and-integral-imaging-displays/viewpoint-adaption-for-multiview-3d-displays.html

[18] René de la Barré, „Adapting of Electronic Image Content for Autostereoscopic Presentation", Whitepaper, Fraunhofer HHI, Berlin, 2010.

[19] Silvio Jurk, Mathias Kuhlmey, Bernd Duckstein and René de la Barré, "Electronical correction of misalignments between optical grid and pixel panel on autostereoscopic displays", IS&T International Symposium on Electronic Imaging 2016, DOI: 10.2352/ISSN.2470-1173.2016.5.SDA-445, 2016.

[20] René de la Barré, Klaus Hopf, Silvio Jurk and Ulrich Leiner, "TRANSFORMERS - Autostereoscopic displays running in different 3D operating modes", SID Symposium Digest of Technical Papers 42(1):452-455, DOI: 10.1889/1.3621351, 2011.

[21] Peter Kauff, Christian Riechert and Frederik Zilly, "Feature descriptor for image sections", Patent App. WO 2012/136642 A1,

[22] Peter Kauff, Marcus Müller, Christian Riechert, Oliver Schreer and Frederik Zilly, "Apparatus and method for real- time capable disparity estimation for virtual view rendering suitable for multi-threaded execution", US Patent 9,158,994 B2,

[23] René de la Barré and Silvio Jurk, "Method for displaying image information on auto-stereoscopic display screen", US Patent 8,633,972 B2,

[24] René de la Barré and Silvio Jurk, "Autostereoscopic screen and method for reproducing image information", EP Patent 2708029 B1,

[25] René de la Barré and Silvio Jurk, "Autostereoscopic display and method of displaying a 3d image", US Patent 9,628,784 B2,

[26] René de la Barré and Silvio Jurk, "An autostereoscopic screen and method for reproducing a 3D image", US Patent 9,398,291 B2,

[27] René de la Barré and Silvio Jurk, "Method for displaying image information and autostereoscopic screen", US Patent 8,358,335 B2,

[28] René de la Barré and Silvio Jurk, "Screen and method for representing picture information" US Patent 8,687,051 B2,

[29] René de la Barré and Silvio Jurk, "An autostereoscopic screen and method for the simultaneous reproduction of more than two different pictures", CN Patent 104541504 B,
Originally published as Patent Appl. WO 2014/023321 A1

[30] René de la Barré and Silvio Jurk, "Method for reproducing image information, and autostereoscopic screen", Patent Appl. WO 2014/183774 A1,

[31] René de la Barré and Silvio Jurk, "Method for reproducing image information and autostereoscopic screen", Patent Appl. WO 2014/048512 A1

[32] René de la Barré, "Image display using matrix screen and periodic grating having a fixed period", US Patent 9,319,663 B2,

Fraunhofer
Heinrich-Hertz-Institut

White Paper

Fraunhofer
Heinrich Hertz Institute
Einsteinufer 37
10587 Berlin
Germany
Phone +49 (0) 30 310 02 - 0