Development of Algorithms for Optimal Viewing Distance Control for Multi-User Application

Version 1.3

White Paper

Fraunhofer HHI

Title:	Development of Algorithms for Optimal Viewing Distance Control for Multi-User Application
Authors:	René de la Barré, Silvio Jurk
Date:	April 30, 2013

Table of Contents

Table of Figures

1 Introduction

The early attempts at autostereoscopic multiview imaging were made with vertically arranged barriers or lenticular plates. Each vertical lens covered multiple columns of pixel stripes, providing perspective views and multiple viewing zones. This traditional approach had two main disadvantages. Firstly, it ensures an unbalanced distribution of spatial resolution. Typically, the horizontal spatial resolution is greatly reduced without affecting the vertical resolution. The horizontal resolution for each eye is defined by dividing the number of horizontal LCD pixels by the number of views offered. Second, these approaches show disturbing moiré effects caused by the gaps between the pixels and the perception of dark bands on the screen.

Cees van Berkel (Philips) proposed a method to avoid these drawbacks. He introduced a technology that uses an oblique alignment of the image separation component. This method dissolves moiré-like black bands and supports a more symmetrical loss of spatial resolution in horizontal and vertical alignment of the displayed views.

However, the current multi-view approaches based on lenticular plates only maintain the stereo effect for limited changes in the user's viewing distance. Fraunhofer HHI proposes a new technology that supports distance movements of observers with good stereo quality. This method calculates a large number of perspective intermediate views in a fragmented arrangement that is dynamically adapted to the individual viewing distances. The general concept of this new approach describes the next chapter. Detailed information on the algorithms used can you find in the following chapters.

2 General Concept

Conventional autostereoscopic multi-view displays support sufficient stereo quality only within a limited viewing distance around a fixed nominal distance. The stereo image quality decreases with increasing deviation from the nominal viewing distance. In the case of larger deviations, the perceived left and right image of a stereo pair is not displayed in the correct stereo view. The left and right images are spatially composed of different views, which in the worst case result in pseudo-stereoscopic image perception. Figure 1 shows the perceived image information according to different viewing positions in a common multi-view display. The viewing position (A) leads to correct binocular depth perception, as the left and right image of a stereo pair is correctly displayed in the entire image area. Observing from position (B) provides blurred images with limited depth as overlapping images are displayed. Position (C) creates a stereopsis, although each stereoscopic image consists of three adjacent stereoscopic views. The viewing position (D) cannot provide stereopsis. Positions (E) and (F) produce pseudo-stereoscopic perception.

Correcting incorrect view alignments requires a method that allows continuous variation of multiple views. The discrete arrangement of the pixels on the display makes a simple rearrangement of the image distribution impossible. For this reason, a new method is recommended that allows continuous view adjustment based on the calculation of intermediate views and their column-by-column playback. This technology realizes a shift of the nominal viewing distance to a new distance supporting stereopsis. In combination with appropriate eye-tracking technology, a dynamically adapted multiple view is realized, which supports a large change in the viewing distance.

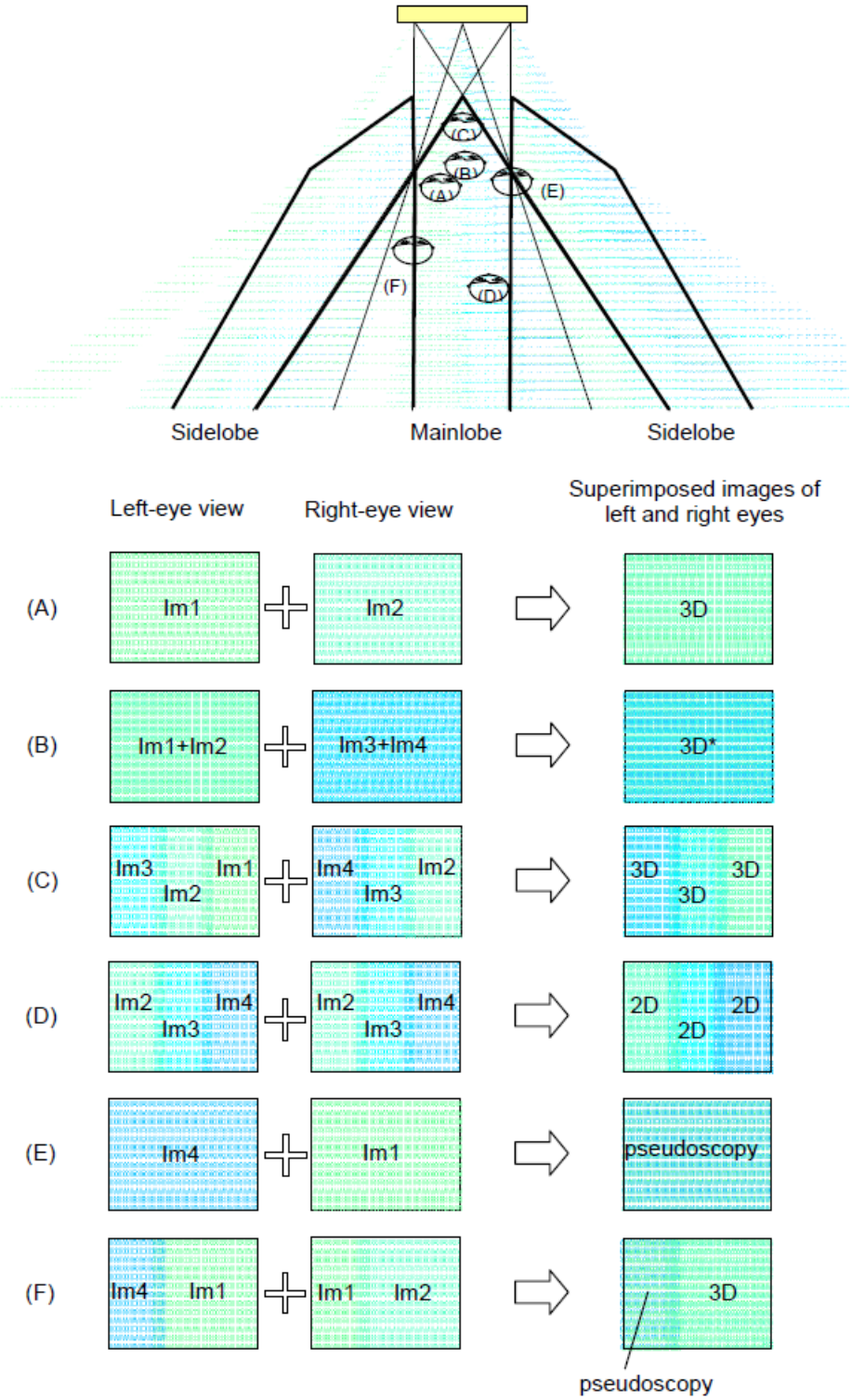

Figure 1: Observed views depending on observers viewing positions [1]

3 Description of the Applied Algorithm

3.1 General Method

A section of a 9-view representation is shown in Figure 2 as an example. For reasons of clarity, this figure shows only a few selected optical paths that refer to 'view 5'. From the nominal viewing distance (nom D), 'view 5' with all its subpixels can be correctly perceived across the entire display width. A shift of the viewing distance to shift D results in a shift of the perceived views with non-vertical axes. This diagram is not to scale, but shows the trends in view shifts.

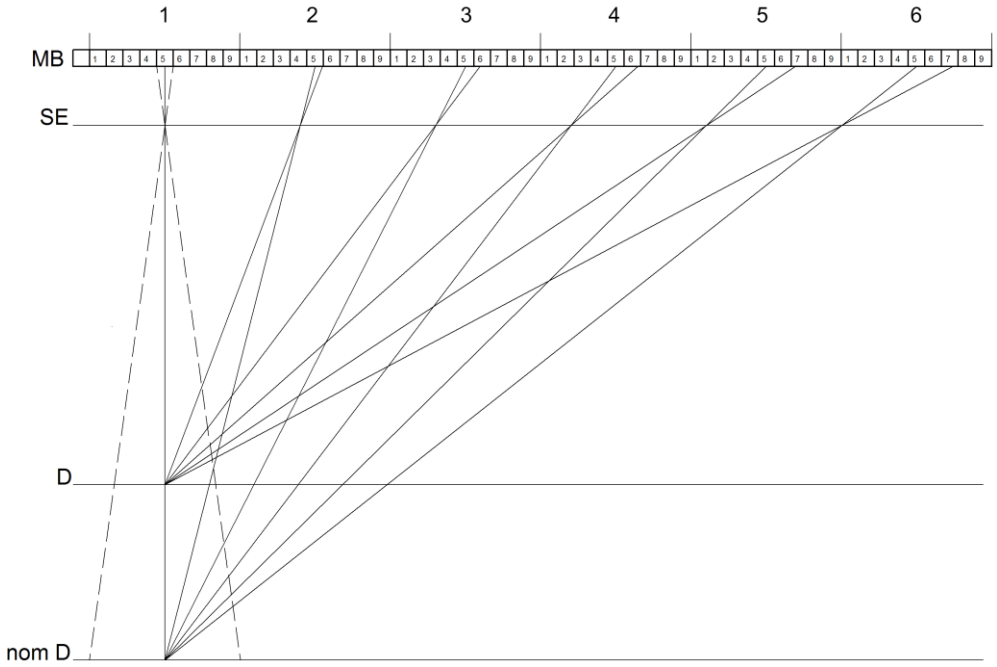

Figure 2: Illustration of misalignments of corresponding views depending on the distance

Figure 3 illustrates a further example showing the situation for one viewing eye. It can be seen that the viewer's impression results from partly one view and partly two superimposed views. This blending of different views leads to blurred image perception.

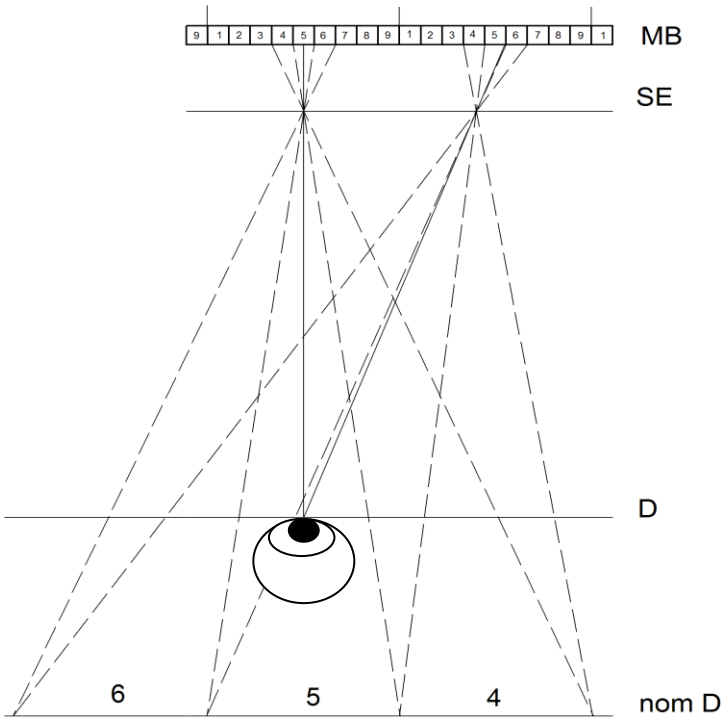

Figure 3: Fusion of views are evoking blurred image perception

Incorrect image alignment and blurring are reduced by the following method of image reproduction. This technique requires image processing technology that is capable of providing multiple specific intermediate views. The method does not fully reflect some views within the entire image plane. The nested views are displayed line by line, taking into account the alignment of the separating grid. Figure 4 shows virtual camera positions depending on different viewing distances (D). The occurring changes of the render directions are calculated and used for the continuous shifting of the image contents.

Figure 4 shows six image strips corresponding to different viewing distances in a 9-view display implementation. The vertical line under each strip represents the deviation from the nominal viewing distance. The strip no.1 shows the view assignment for the vertical axis and the middle area of the display (no deviation from ND). The following strips (no.2 - no.6) indicate the views for progressive reduction of the viewing distance.

This adapted image content ensures a correct and sharper image presentation. With increasing deviation from the nominal distance (nom D), the stereoscopic representation changes from "constantly changing perspectives" arranged in diamonds. The remaining blur can be reduced by a display design that uses more than one view between the interpupillary distances and minimizes crosstalk between views.

With this technology, the nominal distance can be shifted in the direction of the display. The continuously adjusted perspectives lead to a continuous change of perspective for the viewer moving sideways. This includes an adaptation to the individual needs of the observation (interpupillary eye distance, offered perspective). This method of controlling the optimal viewing distance (OVDC) extends the capabilities of traditional multi-view technology to systems that support greater changes in viewing distance.

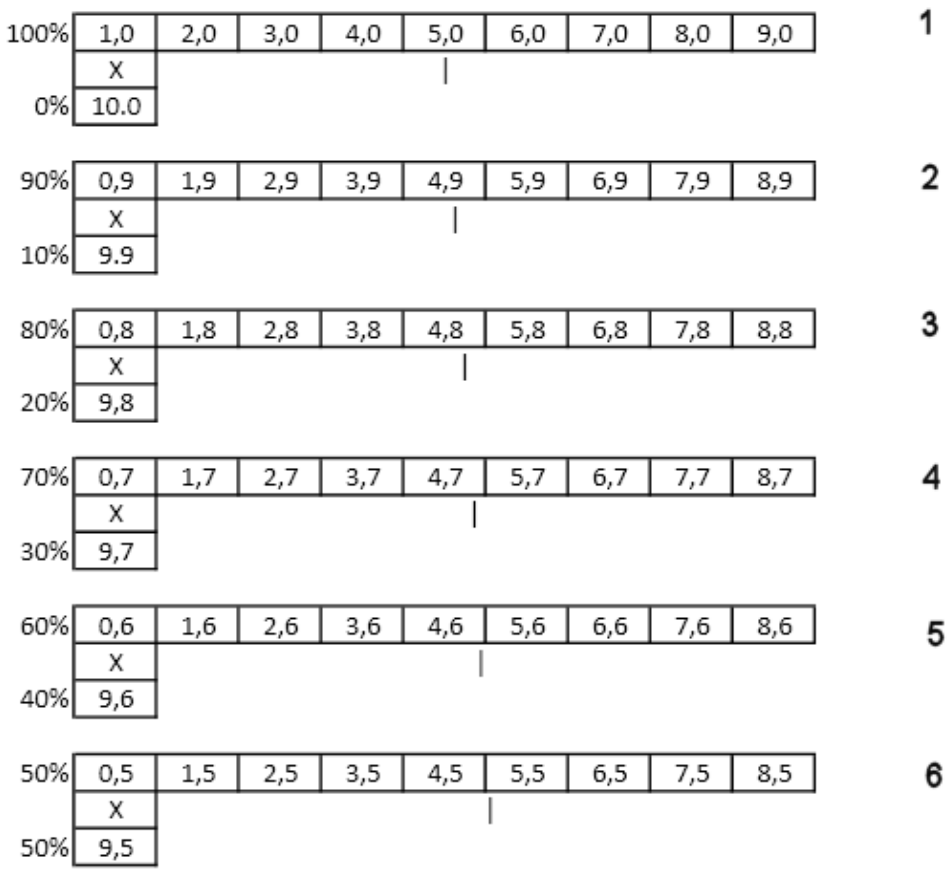

Figure 4: Continuously shifting of image content based on the calculation of specific views

3.2 OVDC Content Generation

Suitable content should be available for OVDC calculations. This content can be based on image processing technology that uses stereo to multi-view conversions such as depth image-based rendering (DIBR). Computer Generated Content (CGI) is also a possible basis for OVDC technology (see Figure 5).

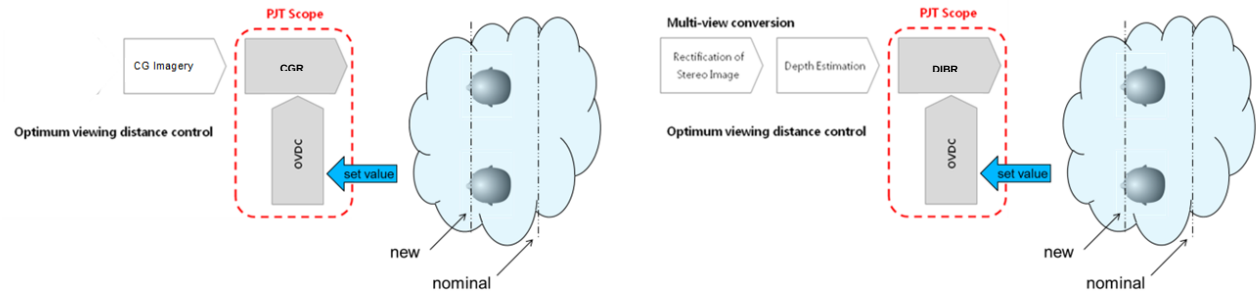

Figure 5: OVDC based DIBR (left) and CGR (right)

3.3 Block Diagram

Figure 6 shows the information flow for the implementation of OVDC technology in a block diagram. The OVDC contains two basic modules, the OVDC block with camera control and a spatial filter that modifies the content on the outer subpixels.

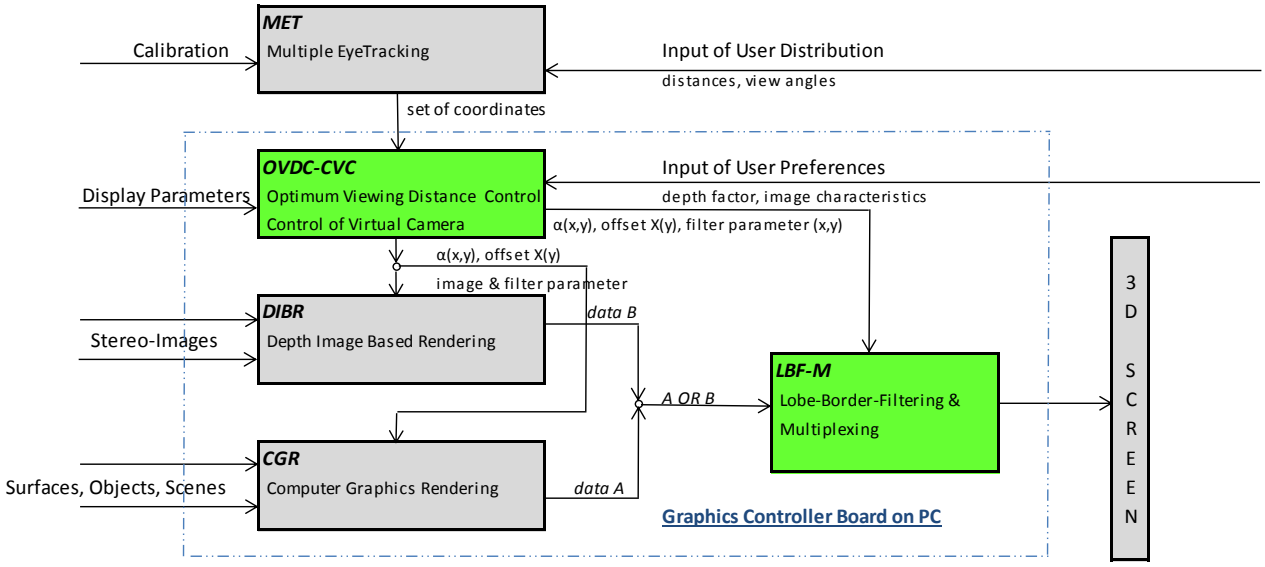

Figure 6: Information flow for OVDC technology

3.4 OVDC Implementation

The OVDC rendering method requires a calculation of numerous intermediate image perspectives. Image strips of specific views are selected and combined to represent the entire image. In general, the image consists of image strips in sub-pixel width. Each stripe contains the corresponding view for a correct stereo display in the entire display area. The selection of the corresponding viewing strips is controlled by a so-called "target value", which represents a certain eye position. The set-point can be derived from a real-time eye tracker to provide a dynamic OVDC.

If an observer leaves the set distance and approaches the viewing distances, the integer values identifying the initial views must be stretched to intermediate values (Figure 7, left side). If an observer chooses to wide viewing distances, the initial values must be placed in a compressed order (Figure 7, right side).

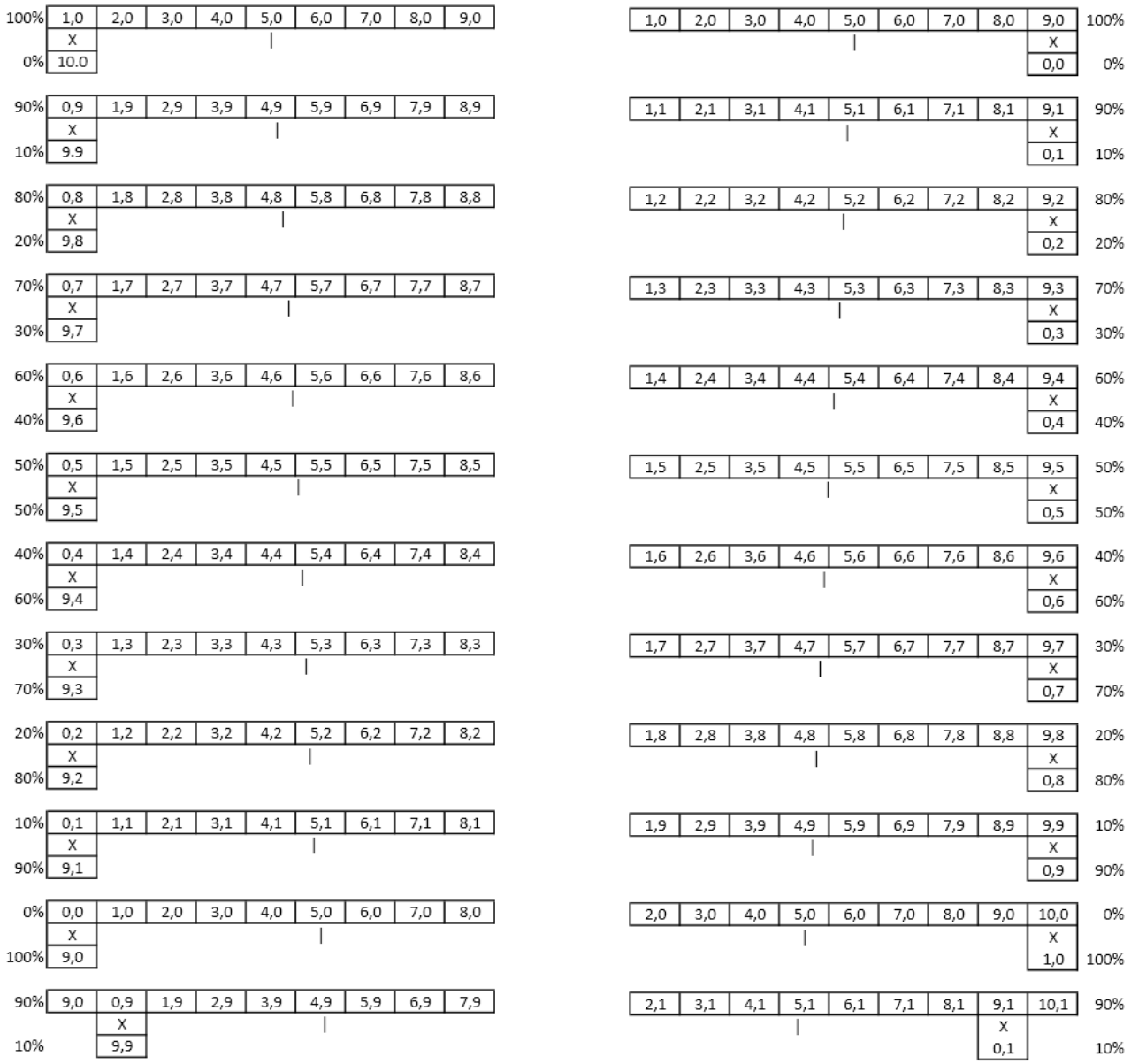

Figure 7: Stretching (on the left) and compression (on the right) of views depending on different viewing distances

3.4.1 Transformation of Eye Location Data into Global Shift Values

In general, a head tracking system provides eye position data based on a Cartesian coordinate system. In a first step, the eye position data provided is transferred to global displacement values in the XY plane and YZ plane. The global displacement values thus cover a two-dimensional plane. A method based on quadratic regression is used to map the three-dimensional eye tracking data to a two-dimensional coordinate system.

3.4.2 Transformation of Global Shift Values into Local Shift Values

The global shift values of the XY-plane and the YZ-plane have to be transformed into local shift values to provide smooth adaptation of the image content when an observer deviates from the nominal viewing distance. The shift value XY represents a uniform shift of all sub-pixels in a specific extent. A value of 1.0 corresponds to a content shift in the extent of one sub-pixel. The shift value sign identifies the shift direction. The shift value YZ is a factor representing the extent of compression and stretching. For each sub-pixel group the global shift values have to be transformed individually into local shift values. The

formula shown in (Figure 8) defines a factor S representing a periodical process of a pattern shift in an extent of the whole width of a sub-pixel group.

$$S = frac\left(\left(x - frac\left(\frac{y}{\frac{W}{H} * vp}\right) - 0.5 - frac\left(\frac{0.5}{\frac{W}{H} * vp}\right)\right) * YZ + XY\right)$$

Figure 8: Definition of the factor S

Explanation:

- x: normalized horizontal display coordinates
- y: normalized vertical display coordinates
- W: horizontal display resolution
- H: vertical display resolution
- vp: slant of the interlacing pattern (specified in lines)
- XY: global shift value in XY plane
- YZ: global compress/stretch value in YZ-plane

Figure 10 shows a visualization of the factor S for subpixel groups in different gray levels. The factor vp defines the inclination of each period. This is necessary to obtain correct displacement values (s) and displacement directions (d). These must follow the inclination of the nesting pattern. The left part of Figure 10 shows a compression of the shift value. The shift values drift from the display edge to the display center. The right part of Figure 10 shows an elongation of the displacement values, with the displacement gradually decreasing from the display center to the display edges.

The following formula can be used for a version with vertically arranged raster plate.

$$S = frac\left(\left(x - 0.5 - frac\left(\frac{0.5}{\frac{W}{H} * vp}\right)\right) * YZ\right)$$

Figure 9: Definition of factor S for vertically arranged image splitter

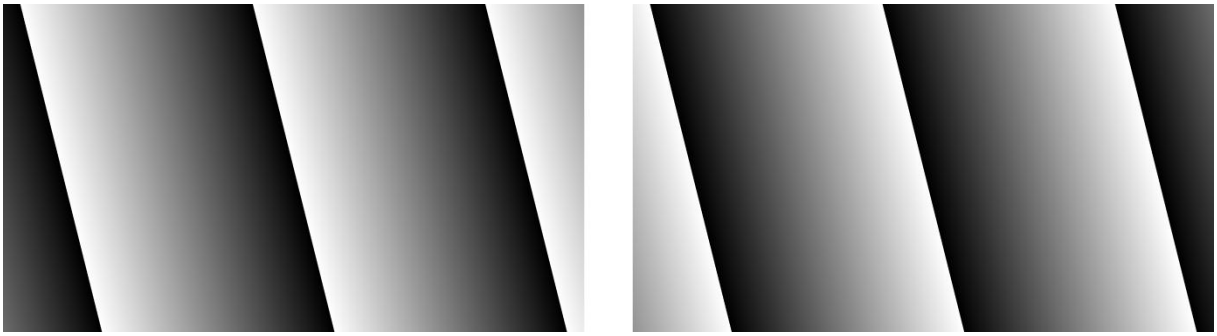

Figure 10: Visualization of the factor S for compression (left) and stretching (right) of image content

3.4.3 Calculation of Local Shift Values

The determination of the normalized local displacement value s and the shift direction d results from the formula that defines the normalized factor S (Figure 8). This segmentation is described in the following algorithm. To determine the switching direction d, the factor S must first be divided into two equal parts (*if-then-else* condition). If factor S is less than 0.5, the displacement drifts to the right. For S-values greater than 0.5, the shift to the left drifts. With a left shift, the number of subpixels within a subpixel group V multiplies the local shift value s. The left part of the displacement is divided into V/2 segments, since the local shift value s must not be greater than 0.5. Since the shift value s could reduce the range of 1 when multiplied by V, this value is further separated. The integer part of this value is stored in the factor *offset*, which represents a switching of the interlacing pattern in the range of n subpixels. The fractal part is stored as shift value s. The factor *offset* depends on the shift direction, which determines the shift direction of the interlacing pattern.

$$
\begin{aligned}
&if(S < 0.5)\\
&\{\\
&\qquad d = 1.0\\
&\qquad s = S * V\\
&\qquad offset = floor(-s)\\
&\}\\
&else\\
&\{\\
&\qquad d = 0.5\\
&\qquad s = (1.0 - S) * V\\
&\qquad offset = floor(s)\\
&\}\\
&s = frac(s)
\end{aligned}
$$

Figure 11 shows a color-coded visualization of the segmentation of factor S for each sub-pixel group. The red channel represents the shift value s and the green channel represents the shift direction. In the dark-green areas the shift occurs to the left direction and in the light green areas the shift occurs to the right. The left part of Figure 11 visualizes a compression and the right part a stretching of the image content. The separation into an integer part, which is stored as factor *offset* and a fractile part s, results in a further fragmentation of the period of factor S in V sub-cycles. V specifies the number of sub-pixels in a sub-pixel group.

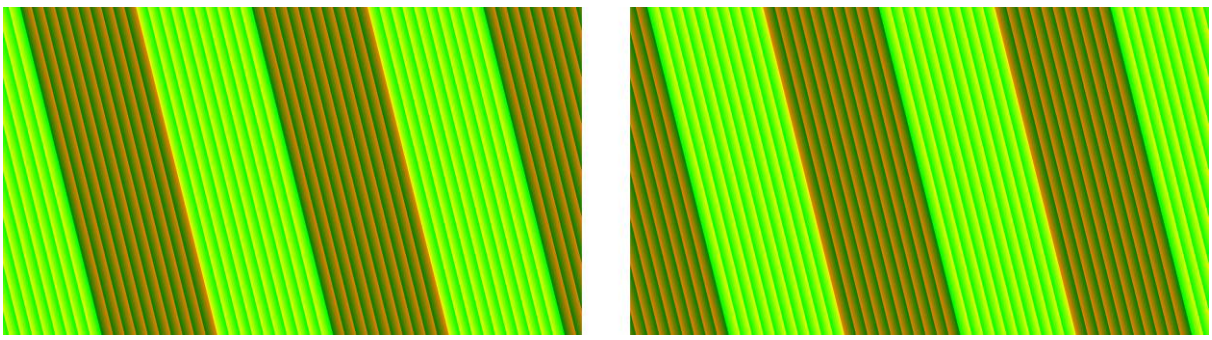

Figure 11: Segmentation of factor s in shift values s and shift directions d for compression (left) and stretching (right) of image content

3.5 OVDC Mode of Operation

The number of views required depends on the quality and the crosstalk parameters to be achieved. Evaluations with a 42" Philips 9-View WoW Display (70 mm IPD, 350 cm ND) showed that the use of 30 views offers sufficient image quality. Figure 12 shows selected optical paths in a conventional multi-view display that provides diamond-shaped viewing zones in nominal size for the "single row view distribution" mode.

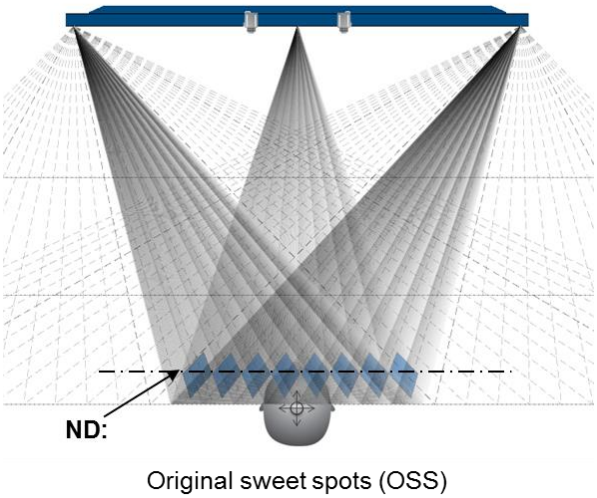

Figure 12: Selected optical paths in a common multi-view dispay providing rhomb-shabed viewing zones in nominal viewing distance

The perception of the individual perspectives in the sweetspots changes to an even motion parallax at reduced and increased nominal distances. (see Figure 13 and Figure 14)

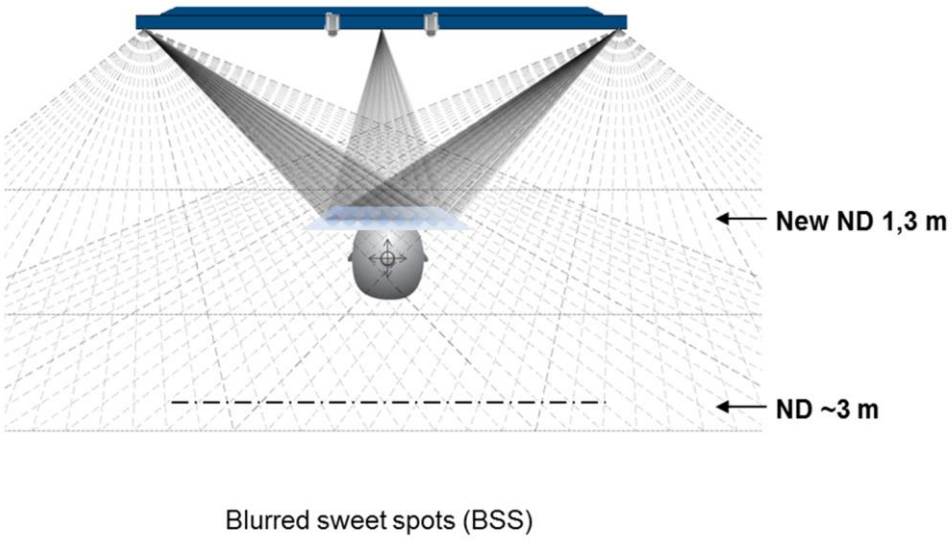

Figure 13: The perception of separate perspectives changes to smooth simulated motion parallax with reduced nominal distance

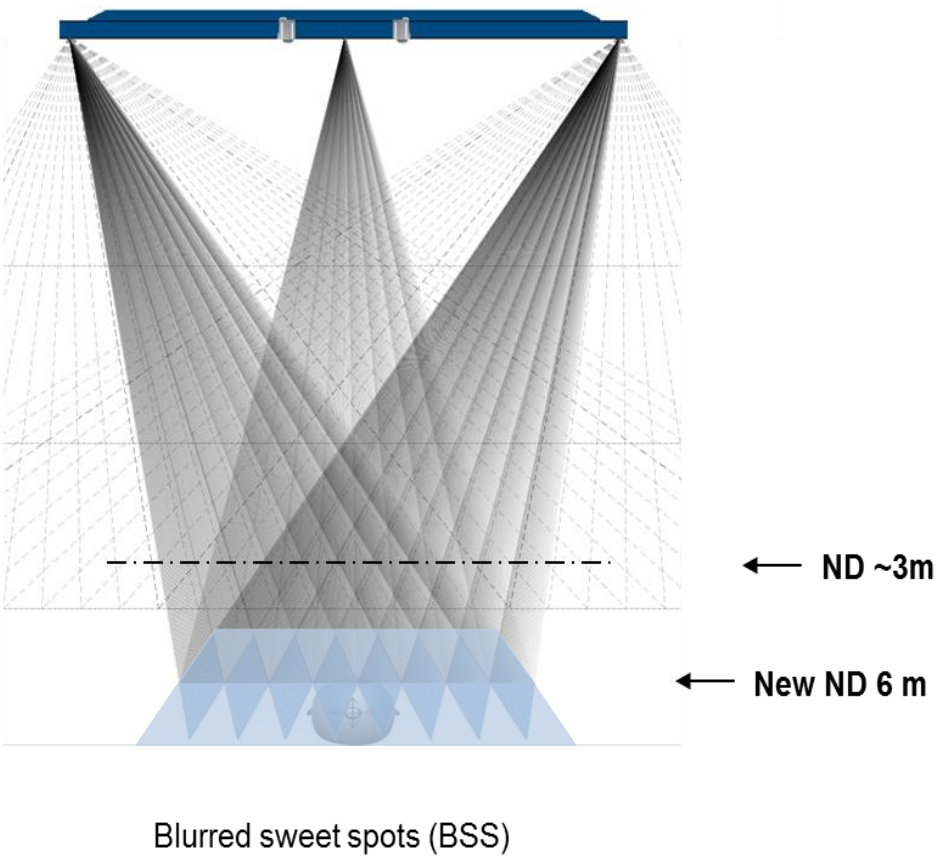

Figure 14: The perception of separate perspectives changes to smooth simulated motion parallax with enlarged nominal distance

4 References

[1] "ISO TR 9241-331".

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