# STEREOSCOPIC ARTIFACTS ON PORTABLE AUTO-STEREOSCOPIC DISPLAYS: WHAT MATTERS?

*Atanas Boev, Atanas Gotchev, Karen Egiazarian*

Department of Signal Processing, Tampere University of Technology, Finland
firstname.lastname@tut.fi

## ABSTRACT

We aim at identifying the artifacts arising when dealing with 3D video content displayed on portable auto-stereoscopic displays. We list the artifacts related with the stages of 3D video delivery dataflow they appear at i.e. content creation, conversion to the desired format, coding/decoding, transmission, and visualization on 3D display. Eventually, we analyze which artifacts are more pronounced on portable displays and should be taken care of and which artifacts are well masked and of little effect on the stereo video perception. We suggest approaches for measuring artifacts caused by limitations of the 3D auto-stereoscopic technology employed.

## 1. INTRODUCTION

Recently, most of the building blocks of an end-to-end mobile 3DTV system have reached a maturity status. An ISO/MPEG multiview encoding standard developed as an amendment to H.264 AVC is due by the end of 2008 [1], [2]. Various algorithms have been developed for the efficient transmission of video streams over wireless networks [1], [3]. There are 3D displays optimized for a mobile use [4], [5]. While the core technologies have been developing, there is still much to be done to optimize the system to deliver the best possible visual output [6], [7]. Having a perceptually acceptable and high-quality 3D scene on a small display is a challenging task.

Estimation of the quality is the key factor in design and optimization of any visual content. All quality metrics aim at close assessment of the quality as perceived by the user. The first step towards objective quality estimation metric is to identify the artifacts, which could arise in various usage scenarios involving stereoscopic content.

In a previous paper, we have discussed the stages of 3D content delivery, which might be source of artifacts, namely content creation, format conversion, compression, transmission, and visualization [8]. Human vision is a set of separate subsystems, which transmit spatial, color and motion information to the brain using largely independent neural paths [9]. Perception of depth consists of different "layers" as well. It works by analyzing a set of separate depth cues – accommodation, binocular depth cues, pictorial cues and motion parallax [9], [10] – which are independently perceived [11] and have varying importance from observer to observer [12], [13]. We have proposed a classification of stereoscopic artifacts, which lists potential sources of artifacts, and contemplates how they would affect different "layers" of human 3D vision [14]. In this work, we focus on artifacts, characteristic for mobile 3DTV reception – broadcast related, or introduced by visualization on portable 3D display.

## 2. PORTABLE AUTO-STEREOSCOPIC DISPLAYS

Currently, there is a wide range of 3D display technologies [12], [15], but not all of them are appropriate for mobile use. For example, wearing glasses to aid the 3D perception of a mobile device is highly inconvenient. The limitations of a mobile device, such as screen size, CPU power and battery life limit the choice of a suitable 3D display technology. Another important factor is backward compatibility – a mobile 3D display should have the ability to be switched back to "2D mode" when 3D content is not available.

Autostereoscopic displays are a class of displays which create 3D effect without requiring the observer to wear special glasses. Such displays are using additional optical elements aligned on the surface of the screen, to ensure that the observer sees different images with each eye. Typically, autostereoscopic displays present multiple views to the observer, each one seen from a particular viewing angle along the horizontal direction. However, the number of views comes at the expense of resolution and brightness loss – and both are limited on a small screen, battery driven mobile device. As mobile devices are normally watched by only one observer, two independent views are sufficient for satisfactory 3D perception. At the moment, there are only a few vendors with announced prototypes of 3D displays, targeted for mobile devices [5], [16], [17]. All of them are two-view, TFT-based autostereoscopic displays.

The basic operational principle of an autostereoscopic display is to "cast" different images towards each eye of the observer. This is done by a special optical layer, additionally mounted on the screen surface which redirects the light passing through it [15], [18]. There are two common types of optical filters – lenticular sheet [18] which works by refracting the light, and parallax barrier [19] which works by blocking the light in certain directions. In both cases, the intensity of the light rays passing through the filter changes as a function of the angle, as if the light is directionally projected. These two types of optical layers are illustrated in Figure 1.
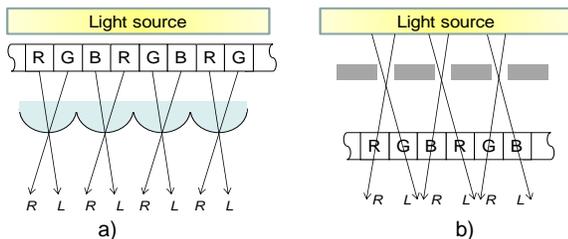


**Figure 1: Redirecting the light of auto-stereoscopic display: a) lenticular sheet, b) parallax barrier**

The main advantage of parallax barrier, besides being the cheaper technology, is the ability to be switched off, thus providing backwards compatibility with 2D content. The main disadvantage of the parallax barrier is that it blocks part of the light, resulting in a lowered brightness of the display. In order to compensate for that, one needs an extra bright backlight, which would decrease the battery life if used in a portable device.

TFT displays recreate the full color range by emitting light though red, green and blue colored components (sub-pixels). Sub-pixels are usually arranged in repetitive vertical stripes as seen in Figure 1. Since sub-pixels appear displaced in respect to the optical filter, their light is redirected towards different positions. One group will provide the image for the left eye, another – for the right. In order to be shown on a stereoscopic display, the images intended for each eye should be spatially multiplexed. This process is referred to as *interleaving* [12], or *interzigging* [20], and depends on the parameters on the optical filter used. Two topologies are most commonly used. One is interleaving on pixel level, where odd and even pixel columns belong to alternative views (Figure 2a). The other is interleaving on a sub-pixel level – where sub-pixel columns belong to alternative views (Figure 2b). In the second case, differently colored components of one pixel belong to different views.

**Figure 2: Interleaving of an image for a stereoscopic display: a) on a pixel level, b) on sub-pixel**

## 3. ARTEFACTS IN MOBILE 3D VIDEO

Not all of the stereoscopic artifacts are likely to affect a mobile 3DTV system. Some of them are not applicable to mobile device due to the technology used (e.g. LCD display, DVB-H transmission). Some of them, i.e. accommodation-convergence rivalry or puppet-theater effect are usually addressed by content providers and/or display manufacturers. In this chapter we identify the most common artifacts to be expected during the data-flow of mobile 3DTV content, how they affect the image quality, and possible ways to mitigate them.

### 3.1. Channel-related artifacts

By 'channel', we denote the technology components at the process of content creation, coding and transmission. The most common and annoying artifact introduced in the process of **capturing** a stereoscopic image is *unnatural disparity* between the images in the stereo-pair. Part of it is caused by camera geometry and position, and the remedy for it – rectification – is well known [21]. However, often a perfectly rectified stereoscopic image needs to be visualized at different size than the originally captured one. Changing the size or resolution of stereoscopic pair can also introduce unnatural disparity.
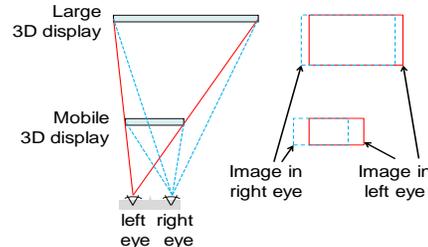


**Figure 3: Change of relative disparity while rescaling stereoscopic image pair.**

When resizing a stereoscopic pair, the relative disparity is scaled proportionally to the image size. However, as the interocular distance remains the same, observing a closely-positioned mobile 3D display would require different relative disparity range compared to when observing large 3D display placed further away. The effect is illustrated in Figure 3. Even if the mobile and large 3D displays have the same visual size, stereoscopic images on them have different disparity.

Two are the most common approaches for **encoding** 3D video data. One is multi-view coding, where video is encoded as separate streams, and temporal and inter-channel correlations are used to compress the data. This approach is being standardized as an amendment to H.264/AVC [1], [2]. Another approach is to augment the 2D video with a dense depth representation of the scene,

and then to code the 2D video and depth channel separately. MPEG defined a container format which allows H.264/AVC compression to be used for that format as well. Using block-based DCT compression is a source of *blocking artifacts*, which are thoroughly studied in 2D video, but their effect on stereoscopic quality is yet to be determined. Some authors propose that blocking might be considered as several, visually separate artifacts – *block-edge discontinuities*, *color bleeding*, *blur* and *staircase artifacts* [22]. Each of these artifacts introduces different amount of impairments to object edges and texture. The human brain has the ability to perceive single image by combining the images from left and right eye (so-called cyclopean image) [11]. As result, the same level of DCT quantization might result in different perceptual quality, based on the depth cues present in a stereo image. In Figure 4, both channels of a stereo-pair are compressed with the same quality factor. When an object appears on the same place in both frames, it is equally affected by blocking in each frame, and the perceived cyclopean image is similar to the one shown in Figure 4a. When the object has different horizontal position in each frame, the blocking artifacts will affect differently the object in each frame, which results in a cyclopean image similar to the one in Figure 4b.



**Figure 4: The impact of blocking on stereopairs with different disparity: a) q=15, disparity=0, b) q=15, disparity=4**

Artifacts due to **transmission** errors are sparse, highly variant in terms of occurrence, duration and intensity, and at very low bit rates may be masked by compression impairments. The presence of artifacts depends very much on the coding algorithms used and how the decoder copes with the channel errors. In DVB-H transmission most common are burst errors, which results in *packet losses* distributed in tight groups [23]. In MPEG-4 based encoders packet losses might result in *propagating* or *non-propagating* errors, depending on where the error occurs in respect to key frames, and the ratio between key- and predicted frames. Error patterns of the DVB-H channel can be obtained with field measurements, and then used for simulation of channel losses [23], [24]. In multiview video encoding, where one channel is predicted from the other, usually error burst is long enough to affect both

channels [25]. In that case, packet loss artifacts appear on the same absolute position in both images even though the appearance in one channel is mitigated due to the prediction (Figure 5).

If the 3D video is encoded using separate depth channel, usually depth is encoded in much lower bitrate than the video. In that case burst errors affect mainly the video channel, and the relative perceptual contribution of depth map degradation alone is very small.



**Figure 5: Packet loss artifacts affecting multi-view encoded stereoscopic video [25].**

One common artifact introduced during **receiving and decoding** of 3D video, is *temporal mismatch*, where one channel gets delayed in respect to the other. It might be caused by insufficient memory or CPU, or error concealment in one channel. The outcome is that the image from one channel do not appear with simultaneously taken image from the other channel, but with an image which is taken a few frames later. Even temporal mismatch of as low as two frames can result in a stereoscopically inadequate image pair. For comparison, two images are shown in Figure 6 - the left is done by superimposing the frame 112 from left and right channels of a movie ; the right is done by superimposing a frame 112 from the left channel and frame 115 from the right channel of the same movie.



**Figure 6: Temporal mismatch in stereo-video: a) superimposed images of a stereopair, b) superimposed images of a stereopair with 3 frames temporal mismatch**

### 3.2. Display-related artifacts

Even a perfectly captured, transmitted and received stereoscopic pair can exhibit artifacts due to various technical limitations of the stereoscopic display [26], [27], [28]. The most pronounced artifact in autostereoscopic displays is cross-talk, caused by imperfect separation of the "left" and "right" images and is perceived as *ghosting artifacts* [20]. Two factors affect the amount of crosstalk introduced by the display – position of the observer and quality of the optical filter.

Due to the size of the sub-pixels, there is a range of observation positions, from where some sub-pixels appear partially covered by the parallax barrier, or are partially in the focal field of the corresponding lenticular lens. Following that, the visibility of each channel gradually changes as a function of the observation angle, as shown in Figure 7a. This creates certain *optimal observation spots* where the two views are optimally separated (marked with "I" and "III" in Figure 7a), and transitional zone (marked with "II") where a mixture of the two is seen. However, even in the optimal observation spot one of the views is not fully suppressed – for example part of the light might "leak" through the parallax barrier as shown in Figure 7b. This determines the *minimal crosstalk* of an autostereoscopic display.

We created a test stereoscopic pair, where the "left" image contains vertical bars, and the "right" image contains horizontal bars. Then we visualized the pair on a parallax-barrier based 3D display, and photographed them from observation angles as marked with "I", "II" and "III" in Figure 7a. The resulting photos are shown in Figure 7c, d and e. Both position-dependant and minimal crosstalk effects can be seen. By knowing the observation position and the amount of crosstalk introduced by the display, the effect of crosstalk can be mitigated by pre-compensation as proposed in [29].
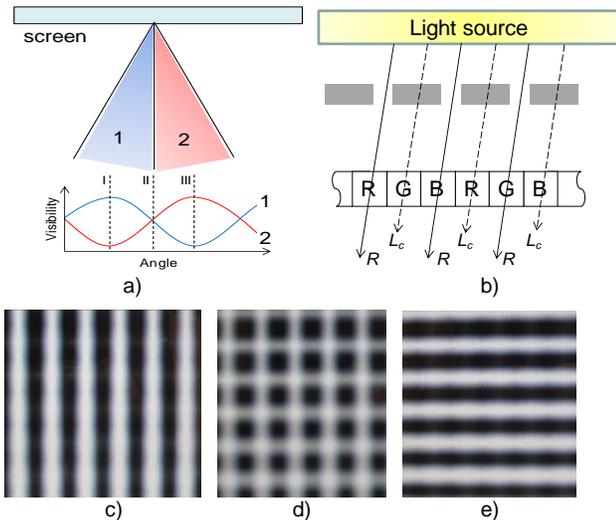


**Figure 7: Reasons for crosstalk in portable 3D displays: a) position of the observer and b) optical properties of the filter, c),d),e) photographs taken of a 3D display from positions I, II and III, as marked in a).**

There are darker gaps between sub-pixels of an autostereoscopic display. They are more visible from certain angles than from others. When an observer moves laterally in front of the screen, he perceives this as luminance changes creating brighter and darker vertical stripes over the image. Such effect is known as *banding artifacts* or *picket fence effect* and is illustrated in Figure **8**. The effect can be reduced by introducing a slant of the

optical filter in respect to the pixels on the screen [18]. Tracking of the user position in respect to the screen also can help reducing these artifacts [27], [28].

Most of 3D displays have horizontal resolution twice lower than vertical one as only half of the sub-pixels of a row form one view. This requires spatial sub-sampling of each view, before both views are multiplexed. Properly designed pre-filters should be used, in order to avoid aliasing artifacts. In 3D displays, aliasing might cause false color (shown in Figure 9a), or Moiré artifacts depending on the properties of optical filter used.



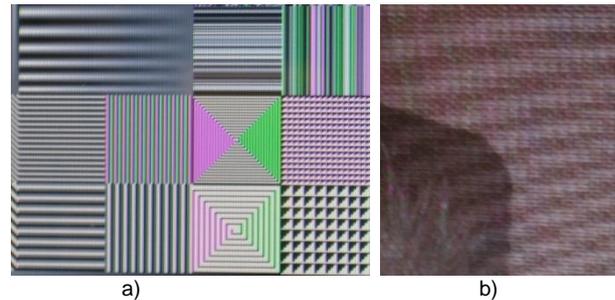**Figure 8: Banding / picked fence artifacts**



**Figure 9: Aliasing in stereoscopic displays: a) false color; b) Moiré artifacts**

Autostereoscopic displays which use parallax barrier usually have a number of interleaved "left channel" and "right channel" visibility zones, as shown in Figure 10. Such display can be used by multiple observers looking at the screen at different angles, for example positions marked with "1" and "2" in the Figure. However, an observer in position "3" will perceive *pseudoscopic* (also known as *reversed stereo*) image. For one observer, this can be avoided by using face tracking and algorithm which swaps the "left" and "right" images on the display appropriately to accommodate to the observers viewing angle.
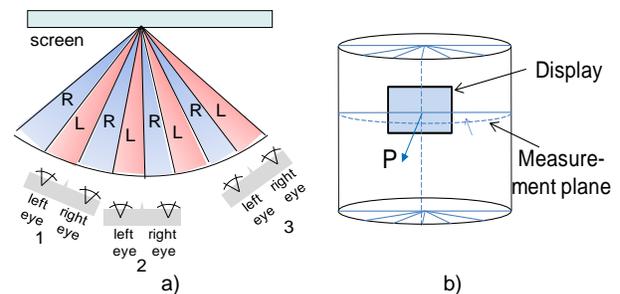
## 4. MEASUREMENT OF DISPLAY SPECIFFIC ARTIFACTS

There are two important optical properties that characterize an autostereoscopic display – crosstalk, which measures the amount of light from one view added to the other, and luminance profile, which measures the brightness of each view as it depends on the observation angle. Both can be measured based on methodology suggested in [30] using professional equipment, or closely approximated by using off-the-shelf camera as proposed in [31]. Briefly, the latter works as follows.

A number of observation points should be selected onto a plane, perpendicular to the display surface, as shown in Figure 10b. The observation points should appear along a line, parallel to the display surface as shown in Figure 11a. A number of test stereo images should be prepared, where each "view" contains all pixels set at a various brightness levels. Each test image should be photographed from each observation point.
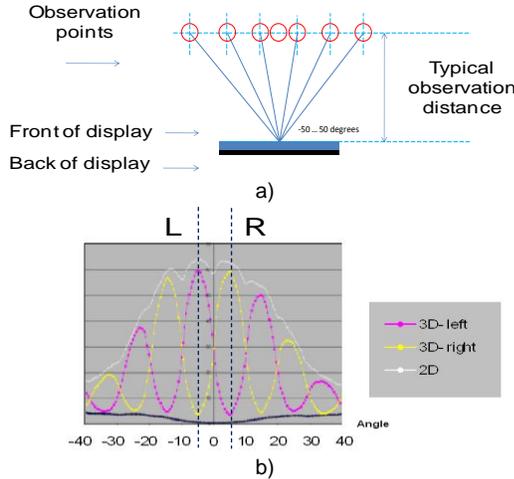


a)



b)

**Figure 11: Crosstalk measurements of autostereoscopic display**

The luminance profile of a typical parallax-barrier based display is similar to the one, shown in Figure 11b [30]. The brightness of each view changes with the angle, as shown in Figure 7a, and additionally the overall brightness of the TFT display decreases for observation angles away from the perpendicular. The brightness difference between the channels at the angle where one channel has maximal brightness (marked with "L" and "R" on Figure 11b) indicate the crosstalk in corresponding direction. The measured results can be utilized for mitigating cross-talk induced artifacts.

Banding artifacts can be measured by setting all pixels of the display to their maximum value. A number of observation points should be selected along the line

perpendicular to the display center (marked with "P" in Figure 10b). The distance, at which banding is least pronounced, should be noted as the *optimal observation distance* of the display. For other distances, the banding usually manifests itself as vertical stripes similar to the ones seen in Figure 12a. The intensity changes periodically in horizontal direction, while along vertical it is fairly similar. The span of the fluctuation can be assessed by measuring the maximum brightness in the center of a bright stripe and minimum brightness in the center of a dark stripe (as marked with $I_{max}$ and $I_{min}$ in Figure 12a). Then, the banding can be modeled as vertical bars where the intensity changes as follows:

$$I_{approx}(x, y) = \sin\left(\frac{2\pi x n}{x_{total}}\right) \cdot I_{span} + I_{mean} \qquad (1)$$

$$I_{span} = \frac{I_{max} - I_{min}}{2}; \quad I_{mean} = \frac{I_{max} + I_{min}}{2}. \qquad (2)$$

In (1) and (2) $I_{approx}$ is approximated intensity of a pixel, $x$ and $y$ are pixel coordinates in the image, $x_{total}$ is image width in pixels, and $n$ is the number of bright stripes present in the photographed image. Measured and approximated intensity of one row are compared in Figure 12b.
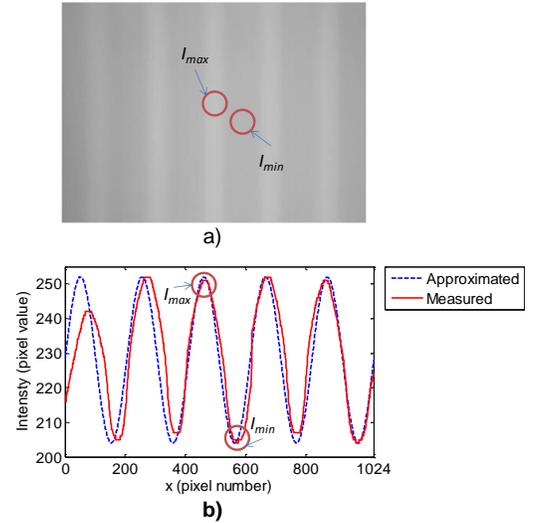


a)



b)

**Figure 12: Banding artifacts: a) measurement and b) approximation**

By counting the number of bright stripes, visible from various distances, one could approximate the parameter $n$ in (1) in the following way:

$$n = \begin{cases} \frac{q}{l}, & l > l_{optimal} \\ ql, & l < l_{optimal}, \\ 0, & l = l_{optimal} \end{cases} \qquad (3)$$

where $l$ is the distance to the display, $l_{optimal}$ is the optimal observation distance, and $q$ is experimentally derived parameter.

## 5. CONCLUSION

In this work, we have overviewed the artifacts which are most pronounced in a mobile 3DTV system involving DVB-H channel and portable parallax-barrier based 3D displays. We have commented on known '2D' artifacts such as blocking and packet losses and their appearance in 3D setting. We also have listed 'pure' stereoscopic artifacts. We have emphasized the display-specific artifacts caused by the limitations of the parallax-barrier auto-stereoscopic technology and have suggested techniques for their measurement. Next research step will be to quantify the join effect of channel-specific and display-specific artifacts by proper subjective tests.

## REFERENCES

[1] ISO/IEC JTC1/SC29/WG11, "Study Text of ISO/IEC 14496-10:2008/FPDAM 1 Multiview Video Coding", Doc. N9760, Archamps, France, May 2008.

[2] ISO/IEC JTC1/SC29/WG11, "Joint Multiview Video Model (JMVM) 8", Doc. N9762, Archamps, France, May 2008.

[3] Z. Tan and A. Zakhor, "Error control for video multicast using hierarchical FEC," in Proc. of the Int. Conf. on Image Processing, Kobe, Japan, October 1999, vol. 1, pp. 401-405.

[4] G. J. Woodgate, J. Harrold, "Autostereoscopic display technology for mobile 3DTV applications", in Proc. SPIE Vol.6490A-19, 2007

[5] S.Uehara, T.Hiroya, H. Kusanagi; K. Shigemura, H.Asada, "1-inch diagonal transflective 2D and 3D LCD with HDDP arrangement", in *Proc. SPIE-IS&T Electronic Imaging 2008, Stereoscopic Displays and Applications XIX*, Vol. 6803, San Jose, USA, January 2008

[6] S. Jumisko-Pyykkö and J. Häkkinen, "Evaluation of subjective video quality of mobile devices", Proc. *13th annual ACM int. conf. Multimedia*. New York, NY, USA: ACM Press, pp. 535–538, 2005.

[7] J. Häkkinen, M. Liinasuo, J. Takatalo, and G. Nyman, "Visual comfort with mobile stereoscopic gaming", *Proceedings of SPIE*, vol. 6055, p. 60550A, 2006.

[8] A. Gotchev, S. Jumisko Pyykkö, A. Boev, D. Strohmeier, "Mobile 3DTV System: Quality and User Perspective", European Workshop of Media Delivery (EUMOB'08), Oulu, Finland, July 2008.

[9] Wandell, B.A., *Foundations of vision,* Sinauer Associates, Inc, Sunderland, Massachusetts, USA, 1995.

[10] D. Chandler, "Visual Perception", University of Wales, Aberystwyth, http://www.aber.ac.uk/media/sections/image.html

[11] Julesz, B. Foundations of Cyclopean Perception, The University of Chicago Press, Chicago, 1971.

[12] Pastoor, "3D displays", in (Schreer, Kauff, Sikora, edts.) 3D Video Communication, Wiley, 2005.

[13] IP. Howard and BJ Rogers, "Binocular Vision and Stereopsis", in Oxford University Press, New York, Oxford, 1995

[14] A. Boev, D. Hollosi, and A. Gotchev, "Classification of stereoscopic artefacts", Mobile3DTV Project report, available online at http://mobile3dtv.eu/results/

[15] L. Onural, T. Sikora, J. Ostermann, A. Smolic, M. R. Civanlar and J. Watson: "An Assessment of 3DTV Technologies," NAB Broadcast Engineering Conference Proceedings 2006, pp. 456-467, Las Vegas, USA, April 2006.

[16] G. J. Woodgate, J. Harrold, "Autostereoscopic display technology for mobile 3DTV applications", in Proc. SPIE Vol.6490A-19 (Stereoscopic Displays and Applications XVIII), 2007

[17] S.Uehara, T.Hiroya, H. Kusanagi; K. Shigemura, H.Asada, "1-inch diagonal transflective 2D and 3D LCD with HDDP arrangement", in Proc. SPIE-IS&T Electronic Imaging 2008, Stereoscopic Displays and Applications XIX, Vol. 6803, San Jose, USA, January 2008

[18] C. Van Berkel and J. Clarke, "Characterisation and optimisation of 3D-LCD module design", in Proc. SPIE Vol. 2653, Stereoscopic Displays and Virtual Reality Systems IV, (Fisher, Merritt, Bolas, edts.), p. 179-186, May 1997

[19] W. IJzerman et al., "Design of 2d/3d switchable displays," in Proc of the SID, volume 36, Issue 1, pp. 98-101, May 2005

[20] J. Konrad and P. Angiel, "Subsampling models and anti-alias filters for 3-D automultiscopic displays", Image Processing, IEEE Transactions on , vol.15, no.1, pp. 128-140, Jan. 2006

[21] A. Gostashby, "2-D and 3-D Image Registration: for Medical, Remote Sensing, and Industrial Applications", Wiley, 2005, ISBN: 978-0-471-64954-0

[22] M. Yuen, "Coding Artefacts and Visual Distortions", in (H. Wu. K. Rao eds), Digital Video Image Quality and Perceptual Coding, ISBN 9780824727772 , CRC Press, 2005

[23] J. Poikonen, J. Paavola, "Error Models for the Transport Stream Packet Channel in the DVB-H Link Layer", Proc. ICC 2006, Istanbul, Turkey, 2006.

[24] COST207, Digital land mobile radio communications (final report), Commission of the European Communities, Directorate General Telecommunications, Information Industries and Innovation, 1989, pp. 135 147

[25] G. Akar, M. Oguz Bici, A. Aksay, A. Tikanmäki, and A. Gotchev, "Mobile stereo video broadcast", Mobile3DTV Project report, available online

[26] T. Marc, M. Lambooij, W. IJsselsteijn and I. Heynderickx, "Visual discomfort in stereoscopic displays: a review", in Prodings of the SPIE- IS&T Electronic Imaging Vol. 6490, Stereoscopic Displays and Virtual Reality Systems XIV, 5 March 2007

[27] AJ. Woods, T. Docherty and R. Koch, "Image distortions in stereoscopic video systems", in Prodings of the SPIE Vol. 1915,Stereoscopic Displays and Applications, San Jose, California, 23 September 1993

[28] MJ. Meesters, WA. IJsselsteijn and PJ.Seuntiens, "A survey of perceptual evaluations and requirements of three-dimensional TV", in Circuits and Systems for Video Technology, IEEE Trans. Vol. 14,N3,pp. 381 − 391, March 2004

[29] J. Konrad, B. Lacotte, and E. Dubois, "Cancellation of image crosstalk in time-sequential displays of stereoscopic video," Tech. Rep. 97-01, INRS-Télécommunications, Feb. 1997

[30] M. Salmimaa and T. Jarvenpaa, "3-D crosstalk and luminance uniformity from angular luminance profiles of multiview autostereoscopic 3-D displays", Journal of the SID, 16/10, pp. 1033-1040, SID, October 2008.

[31] A. Boev, A. Gotchev and K. Egiazarian, "Crosstalk measurement methodology for auto-stereoscopic screens", Proc. of 3DTV Con, Kos, Greece, 2007