MOBILE 3DTV TECHNOLOGY DEMONSTRATOR BASED ON OMAP 3430

Atanas Gotchev[1], *Antti Tikanmäki*[1], *Atanas Boev*[1], *Karen Egiazarian*[1]
Ivan Pushkarov[2], *Nikolai Daskalov*[2]

[1]Department of Signal Processing, Tampere University of Technology, Finland
[2]MM Solutions Ltd., Sofia, Bulgaria

ABSTRACT

We present receiver-side components of the mobile 3DTV technology being developed. We focus on the DVB-T/H receiving module, the 3D video decoder and player, and the auto-stereoscopic display. These three components have been integrated within the OMAP 3430 EVM. We have developed software decapsulator which decapsulates MPE-FEC tables out of MPEG2 transform streams into RTP to feed the H.264 based video decoder modified to decode side-by-side stereo video. An interface card has been developed to interface a parallax barrier auto-stereoscopic display to the system. We provide details about the implementation of these modules and focus on some image processing techniques aimed at correct display of the stereo video content.

Index Terms— Mobile 3DTV, stereo video, OMAP 3430, DVB-H, MPE-FEC, H.264

1. INTRODUCTION

Achieving higher realism of visual scenes trough providing the 3rd dimension has been long-lasting aim of researchers and technology developers. Recently, 3D video applications have gained a renewed interest due mainly to the achievements of the 3D display technology. Novel auto-stereoscopic displays have been developed thus eliminating the need for the user to wear special glasses to experience the stereoscopic 3D effect [1], [2]. Novel representation formats and compression techniques have been developed to handle the higher amount of data [3]. Use scenarios have been discussed so to find out where the new technology will be accepted first [4].

3DTV is mainly expected to target the users' living rooms by big 3D displays and immersive 3D content [5]. In turn, mobile 3DTV is expected to appeal to the mobile users by bringing 3D content to 'cool' and fashionable handhelds [6]. While it is expected to come first as it targets the more dynamic and open for new technologies mobile market, it brings specific and challenging problems. The new techniques for packet decapsulation, stereo video decoding and view-rendering have to be (jointly) optimized so to work efficiently on power-constrained handheld device, equipped with an auto-stereoscopic display This paper addresses the current development of receiver-side technology demonstration of Mobile 3DTV. Techniques for channel decoding, video decomporession and display are implemented on the OMAP 3430 EVM.

2. SYSTEM OVERVIEW

The transmission part of a mobile 3DTV system is illustrated in Fig. 1. We consider 3D video represented in the form of two-channel stereo video content. The advantage of this presentation is that such video can be directly displayed on an auto-stereoscopic display with proper intezigging of the two channels [7]. A straight forward approach for compression of two-channel video is to align the left and right views one to the other, i.e. side-by-side, and to compress the composite frames by an existing, e.g. H.264 based encoder. This approach, known as simulcast, is inferior in terms of computational efficiency however, allows for direct use of existing coding tools and standards. An alternative is provided by so-called Multi-view coding (MVC), which utilizes also the inter-channel correlations. This approach has been recently standardized as an amendment of H.264 AVC [8]. Its application might lead to up to 35% compression gain compared with simulcast. Inter-channel prediction is similar to the motion prediction applied to temporal frames in H.264 and does not increase the complexity substantially. However, MVC is based and extends the H.264/AVC High Profile thus imposing a requirement for the decoder to support that profile.

The compressed content is then to be transmitted over a broadcast channel. Our system builds on the DVB-H channel considered to be the broadcast media of future mobile 3DTV. As a standard especially developed to serve handheld devices, DVB-H provides specific error-protection and power-saving tools [9].

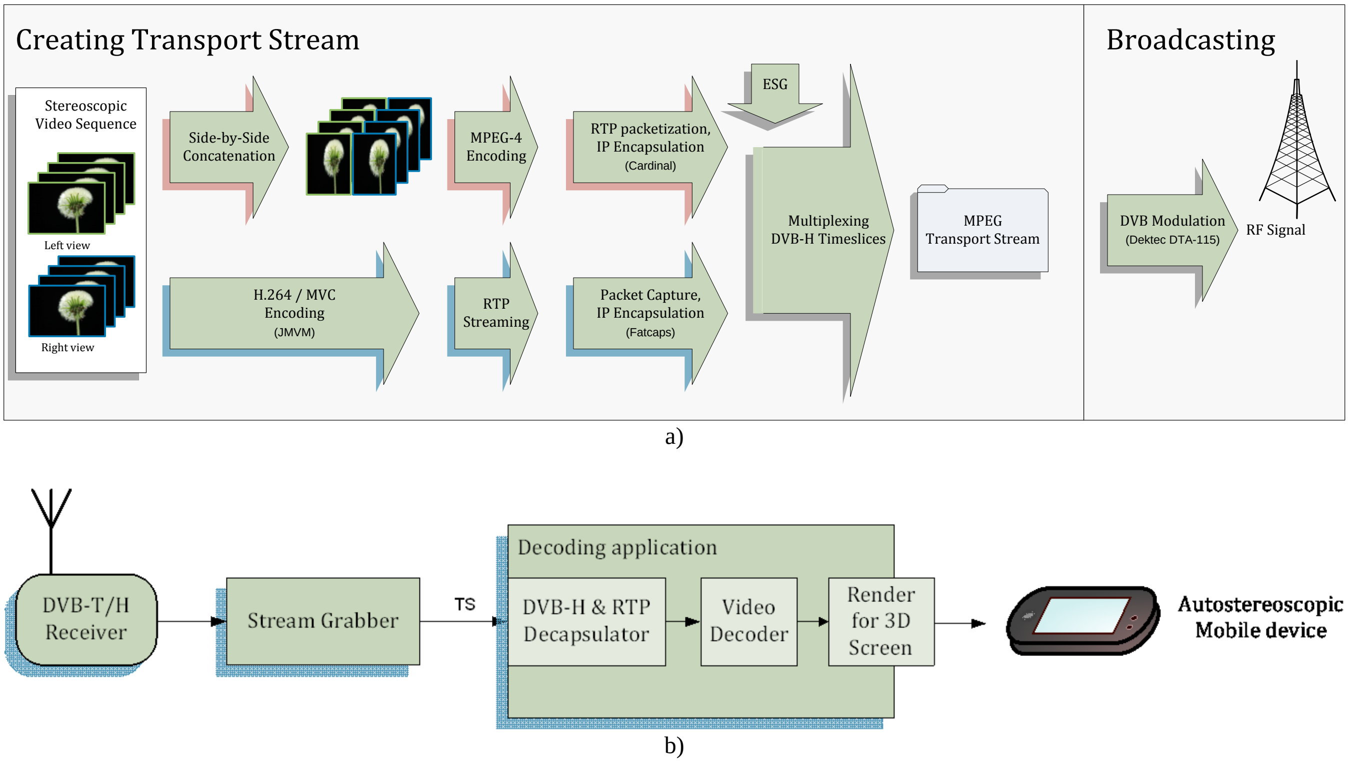

Fig. 1. Mobile 3DTV system over DVB-H chanel.a) Transmitter part; b) Receiver part.

For error protection it uses forward error correction, implemented by Reed-Salomon codes through multiprotocol encapsulation (MPE-FEC). For power efficiency it uses time-slicing to allow data transmission in bursts. In our system, the compressed video in the form of IP datagrams is passed to the link network layer where it is encapsulated into MPE-FEC frames to be packetized in MPEG-2 transport streams (TS) and then transmitted over the channel [10].

At the receiver side, the reverse operations are performed. A DVB-T receiver demodulates the radio signal and delivers TS packed to be then decapsulated to MPE-FEC tables and RTPs. The RTPs are delivered to a streaming client and decoder. The decoded two-channel video is displayed then on an auto-stereoscopic display.

In this paper we focus on the implementation issues related with the receiver device. We provide details on the implementation of the DVB-H stack, video decoder and the display renderer.

3. DVB-H FRONT END

In our current application, the DVB-H/T front end is based on a DVB-T receiver connected to the EVM trough USB interface. A stream grabber within the Linux operating system collects the TS, which are then decoded using the *Decaps* software [11], which was recompiled for the Linux OS running on the OMAP EVM. Compared to some Linux built-in tools, such as *dvbnet*, the *Decaps* software adds several features such as MPE-FEC error correction and collection of error statistics. The IP decapsulator performs the reverse operations of IP encapsulator. It extracts MPE and MPE-FEC sections from the TS. Errors or erasures in the TS are reliably detected by the CRC-32 code included in each section. The data and RS tables of an MPE-FEC frame are filled with the correctly received sections, and RS decoding is performed to recover the lost data. Integrity of the recovered data is finally verified using the UDP checksum. After the error correction, the IP datagrams are fed to the stereo video client over a network connection. In our system, we implemented the IP decapsulator according to the specifications given in [12], [13]. Video players with RTP support can be used to view the decapsulated broadcasts. The software components of the Linux receiver setup are illustrated in Figure 2.

4. STEREO VIDEO DECODING

The OMAP 3430 SDP is bundled with video decoder test application, capable to decode MPEG4, H.264, MPEG2 and WM9 video streams. The application is based on OpenMAX (OMX) and implements optimized decoder running on the DSP and using some vendor-specific hardware accelerators. The block diagram of the decoder application and its components is shown in Fig. 3. It decodes the video-streams coming through the IP datagrams and renders the decoded channels on the LCD display.

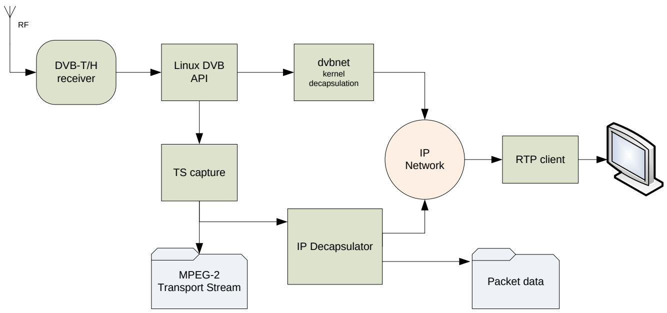

Fig. 2. Linux-based DVB-H receiver.

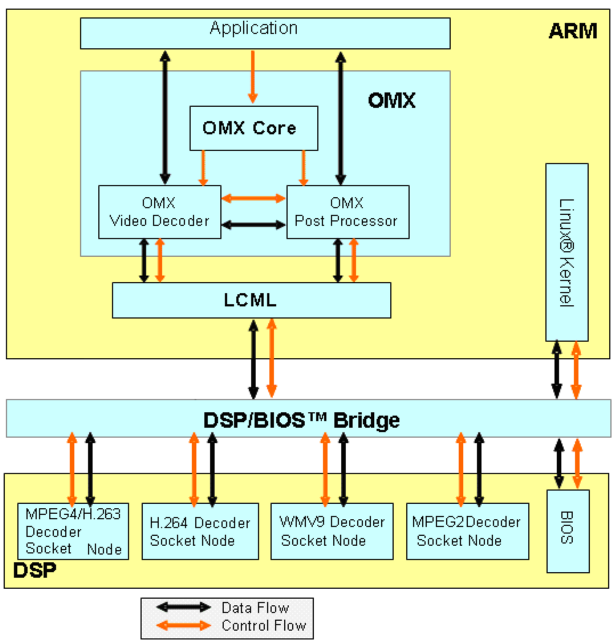

Fig. 3. Block diagram of video decoding architecture in OMAP 3430 SDP.

At that stage, the simulcast encoding approach has been adopted and respectively its decoding part has been implemented. The result of the simulcast, where left and right parts of the stereo image are put side- by-side, is H.264 stream with double frame size. Decoding this stream is equivalent to decoding a stream with double frame size.

Though inferior, the simulcast approach has been chosen in this first version of the demonstrator as it allowed for a direct utilization of the H-264 decoding framework. Furthermore, this first version of the demonstrator has been developed to serve as a platform of studying the perceived quality of 3D video content [14]. With this respect, simulcast allows manipulating and experimenting with videos of different quality (varying quality factors such as bitrate, framerate, transmission modes, etc.) – simply code and decode the streams and put the left and right channels side by side to be played further for the need of the subjective tests.

5. AUTO-STEREOSCOPIC DISPLAY INTEGRATION

5.1. Overview of auto-stereoscopic displays

Auto-stereoscopic displays require no special glasses to create the 3D illusion for the observer. Instead, such displays use dedicated optical elements aligned on the surface of the screen so to ensure that the observer sees different images with each eye. Auto-stereoscopic displays are capable of presenting multiple views to the observer, each one seen from a particular viewing angle along the horizontal direction. However, the number of views comes as a trade-off. More views lower the spatial resolution and, depending on the optical elements, might lead to brightness loss. Therefore, for small-screen, battery-driven mobile devices, two-view displays are considered optimal. This is in line with the use scenario, where a handheld is normally used by single user, and correspondingly its display is watched by a single observer as well [6]. Thus, two 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 [1], [2]. All of them are two-view, TFT-based autostereoscopic displays [15].

The basic operational principle of an auto-stereoscopic 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. There are two common types of optical filters – lenticular sheet [16] which works by refracting the light, and parallax barrier [17] 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 technologies are shown in Fig. 4.

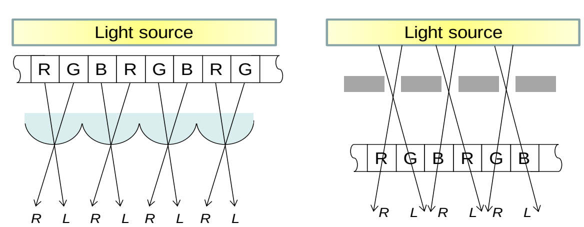

Fig. 4. Principles of operation of lenticular sheet and parallax-barrier auto-stereoscopic displays.

Comparing the cost of developing and manufacturing parallax barrier and lenticular sheet, the former is much cheaper than the latter. In our current implementation we have used a display developed by the parallax barrier technology [18]. The barrier is in fact an additional LCD precisely mounted on top of the main LCD. While switched on, it blocks the light to certain directions depending on the angle, thus creating two different views for the two eyes. The barrier can be switched on and off, so the display works either in 3D or in 2D mode, thus providing backwards compatibility with 2D content. While in 3D mode, the effect of blocking part of the light, lowers also the display brightness. To compensate this effect, an extra bright backlight can be used, which however, would decrease the battery life.

5.2. Display characteristics and interface

We have procured and integrated a commercially available parallax barrier display produced by MasterImage. It is a 4.3”WVGA (800 pixels by 480 pixels) transmissive LCD display. Additional feature of this display is the ability to switch the barrier between horizontal and vertical mode, which allows landscape and portrait mode of 3D operation. Due to the operation principle, a parallax barrier lowers the horizontal resolution twice when operating in 3D mode. For the MarterImage display this results in the following set of resolutions: 800x480 in 2D mode, 400x480 in landscape 3D mode, and 240x800 in portrait 3D mode.

In order to provide proper functionality of the 3D LCD, we have designed a dedicated daughter card connecting the display to the main development platform has been designed. It provides proper voltage to the requested components, generates the supply voltage for the back-light, provides proper power on/off sequence of the 3D LCD, and level-shifts the signals from the platform to the main display. Fig. 5 shows the daughter-card hosting the 3D LCD as connected to the OMAP3430SDP.

Fig. 5. Auto-stereoscopic display interface card for the OMAP 3430 SDP

Regarding the Linux OS, changes on kernel display drivers and frame buffer have been accomplished. These changes have been implemented for the L12.20 baseline release of the platform software.

5.2. View rendering and compensation of 3D artifacts

The left and right view images obtained as a result of video decoding have to be correctly displayed. Recall that in the auto-stereoscopic displays there are group of pixels responsible for one view and group of pixels responsible for the other view. 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 *interzigging* [19], and depends on the parameters on the optical filter used. For the particular display in use, pixels on every other column are responsible for the same view, as illustrated in Fig. 6.

	Pixel 1			Pixel 2			Pixel 3		
Row	R	G	B	R	G	B	R	G	B
1	L	L	L	R	R	R	L	L	L
2	L	L	L	R	R	R	L	L	L
3	L	L	L	R	R	R	L	L	L

Fig. 6. View interzigging for a two-view autostereoscopic display

The view rendering has been implemented by a direct DMA operation on order to safe CPU computational time and power. Two DMA channels were used in order to simultaneously transfer the two halves of the original picture into a final interlaced image as shown in Fig. 7. By using two DMA channels, the two halves are automatically transferred to the correct places with no complex loops. The DMA channels use double-indexing (i.e. indexing within an element frame and an indexing of frames).

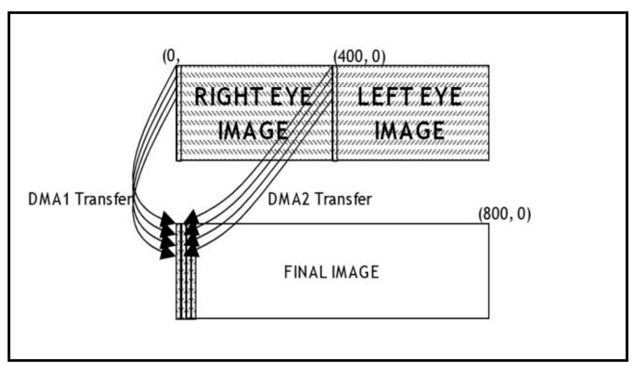

Fig. 7. DMA data transfer from side-by-side format to interlaced frame.

Once rendered, the stereo images suffer from display-specific artifacts. The most pronounced artifact is the cross-talk, caused by imperfect separation of the “left” and “right”

images and perceived as ‘ghosting’ [19]. The perceived amount of cross-talk is determined by the position of the observer and the quality of the optical filter.
There is a range of observation positions, from where some sub-pixels appear partially covered by the parallax barrier causing an angle-dependent perceived quality of each channel. The effect is illustrated in Fig. 8 (left). For certain positions, the two channels are best separated. Such positions are referred to as sweet spots (marked with “I” and “III” in Fig. 8. There are also transitional zones (marked with “II”) where the mixture of the two channels reaches a maximum. Due to ‘leakage’ of light through the parallax barrier, even in the optimal observation spot one of the views might not be fully suppressed. The effect is illustrated in Fig. 8 (right).

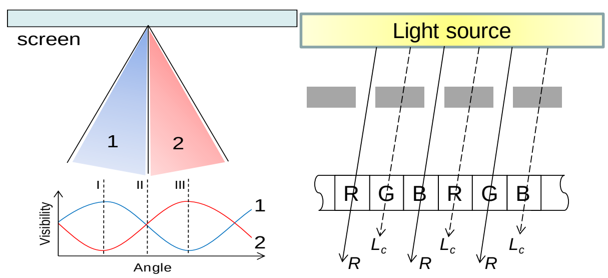

Fig. 6. Cross-talk in an auto-stereoscopic display as result of the viewer position and optical mask.

The discrete structure of the barrier is manifested as darker gaps between sub-pixels of the auto-stereoscopic display. Their appearance is also angle dependent. While moving laterally in front of the screen, one observes brighter and darker vertical stripes over the image. Such effect is known as *banding artifacts* or *picket fence effect*. To compensate this effect, we have developed a camera-based application allowing for determining the relative user position with respect to the screen. First, the observer’s face is detected and then the positions of the two eyes are detected as well. Based on this information, the views are slightly shifted to adjust them to the user’s position. The face and eye detection system is described in more details in a companion paper [20].

The same system compensates an artifact known as pseudoscopy or reverse stereo. As illustrated in Fig. 7, parallax-barrier type of autostereoscopic displays create a number of interleaved “left channel” and “right channel” visibility zones. Based on this, there are multiple positions, from which the stereo effect can be perceived. Equivalently, multiple observers can also experience 3D if looking at the screen from the proper angle. Such proper positions are marked in Fig. 7 by “1” and “2”. However, an observer in position “3” will perceive pseudoscopic image. For single observer use, the effert is effectively avoided by the eye position detection system, i.e. after the eyes are properly detected, the “left” and “right” views are swapped as needed [20].

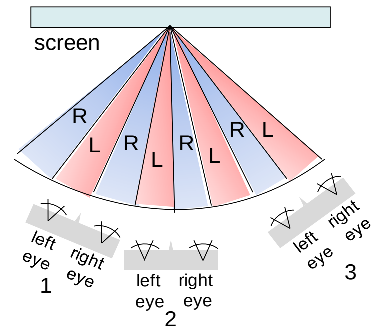

Fig. 7. True (1 and 2) and false (pseudo) – 2 stereoscopic positions.

6. CONCLUSIONS

In this paper, we have presented the first version of a technology demonstrator of a mobile device capable of receiving DVB-H signals, decapsulating the transport packets to IP datacasts containing compressed stereo video and their decoding and playing on auto-stereoscopic display. Our implementation strongly relies on the OMAP 3430 SDP. More specifically, we developed a software decapsulator capable of interpreting DVB-T TS as carrying DVB-H data, i.e. decapsulating the MPE-FEC tables out of the TS. This application runs on the ARM core of the platform. The H.264 video decoding application runs mainly on the DSP core based on the OMX specifications. The third component, the auto-stereoscopic display support was built as a daughter card interfaced to the platform with the corresponding Linux driver support. A DMA based implementation provides the proper interleaving (spatial multiplexing) of the left and right views so to create the stereo effect. Based on the knowledge of display-specific 3D artifacts, we have developed an application for their mitigation, based on the information about the position of the viewer’s eyes.

7. ACKNOWLEDGMENTS

The work was supported by the Academy of Finland, project No. 213462 (Finnish Centre of Excellence program (2006 - 2011) and by the EC within FP7 (Grant 216503 with the acronym MOBILE3DTV). The OMAP 35xx EVM at TUT site was provided by Texas Instruments through the TI University Programme.

8. REFERENCES

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

[2] 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.

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

[4] 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.

[5] Discussion forum "3D in the Home: How Close are We?", Stereoscopic Displays and Applications XVIII, San Jose, California, USA, 2007, www.stereoscopic.org.

[6] Jumisko-Pyykkö, S., Weitzel, M., Strohmeier, D. "Designing for User Experience: What to Expect from Mobile 3D TV and Video? " . Proceedings of the First International Conference on Designing Interactive User Experiences for TV and Video. October 22 - 24, 2008, Silicon Valley, California, USA.

[7] A. Boev, A. Gotchev, K. Egiazarian, 'Stereoscopic artifacts on portable auto-stereoscopic displays: what matters?', in Proc. VPQM 2009, Scottsdale, Arizona, Jan 2009.

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

[9] G. Faria , J. A. Henriksson , E. Stare and P. Talmola "DVB-H: digital broadcast services to handheld devices," Proc. IEEE, vol. 94, pp. 194, Jan. 2006.

[10] M. Oguz Bici, A. Aksay, A. Tikanmäki, A. Gotchev and G. Bozdagi Akar, "Stereo Video Broadcasting Simulation for DVB-H", NEM Summit, Oct. 2008, France.

[11] Decaps software [Online] Available. http://www.mobile3dtv.eu/download/

[12] ETSI, TR 102 377, "Digital Video Broadcasting (DVB); DVB-H Implementation Guidelines," V1.2.1 (2005-11).

[13] ETSI, EN 301 192, "Digital Video Broadcasting (DVB); DVB Specification for Data Broadcasting." V1.4.1 (2004-11).

[14] Jumisko-Pyykkö, S., Strohmeier, D. Report on research methodologies for the experiments, Technical report of the MOBILE3DTV project, December 2008, available at www.mobile3dtv.eu/results/

[15] I. Sexton, P. Surman, "Stereoscopic and autostereoscopic display systems.", in IEEE Signal Processing Magazine, pp. 85-99, 1999.

[16] 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.

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

[18] MasterImage web site, www.masterimage.co.kr

[19] 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.

[20] M. Georgiev, A. Boev, A. Gotchev, N. Daskalov, 'A system for face and eye position detection based on OMAP 3430', submitted to EUSIPCO 2009.