OPTIMIZED VISUALIZATION OF STEREO IMAGES ON AN OMAP PLATFORM WITH INTEGRATED PARALLAX BARRIER AUTO-STEREOSCOPIC DISPLAY

Atanas Boev[1]*, Mihail Georgiev*[1]*, Atanas Gotchev*[1]*, Nikolay Daskalov*[2]*, Karen Egiazarian*[1]

[1]Department of Signal Processing, Tampere University of Technology
P. O. Box 553, FI-33101, Tampere, Finland
phone: + 358 3 3115 4349, fax: + 358 3 3115 3857, email: firstname.lastname@tut.fi
[2]MM Solutions Ltd.
Izgrev, 15 Tintiava Str, Sofia 1113, Bulgaria
phone: +359 2 868 8162, fax: +359 2 962 4404, email: ndaskalov@mm-sol.fi

ABSTRACT

In this paper, we describe a system for optimized visualization of stereo images on a mobile platform. The system utilizes a front camera, and face and eye tracking to find the position of the observer's eyes. Depending on this position, the left and right views targeting the corresponding eyes are maintained properly based on measured optical characteristics of the used parallax-barrier 3D display.

An efficient implementation on the OMAP 3430 platform is targeted by splitting the processes of face and eye detection between the ARM and DSP cores.

The final system allows for dynamic switching the display between 2D and 3D mode and swapping the left and right views so to avoid high cross-talk between view channels and reverse stereo effect.

1. INTRODUCTION

Stereoscopic video content has became more and more popular and available in the form of 3D movies for 3D movie theatres [1], [2], and through 3D display solutions for home [3], [4] and mobile entertainment [5], [6], [7].

The mobile use of 3D video content is especially challenging since it requires creating an immersive 3D effect on a small display and processing big amount of data in a power-constrained handheld device. Autostereoscopic displays requiring no special glasses to deliver the 3D effect have been considered attractive for mobile 3D devices. Such displays however, suffer from 3D artefacts usually related with the position and angle the display is observed from. Special image processing methods are needed to prepare the images for such displays and to mitigate the corresponding artefacts. In this paper, we propose a system, which optimizes the 3D imaging to adapt it to the observation angle of the user. The system is based on OMAP 3430 and employs front-facing camera and face and eye-detection, to find the user's position and angle with respect to the screen in order to create the 3D image accordingly. The paper is organized as follows: First, we briefly present the mobile auto-stereoscopic displays, the respective artefacts and the particular display we deal with. Then, we present the suggested system for optimized visualization. Section 4 presents the implementation details concerning the face and eye detection module on the OMAP platform.

2. MOBILE 3D DISPLAYS

There are two important requirements for mobile 3D display: to create 3D effect without the need of special glasses, and to be able to switch back to "2D mode" when 3D content is not available. Autostereoscopic displays create glasses-free 3D effect by emitting different images towards each eye of the observer. In such displays, a standard portable TFT display is used to generate the images and an additional optical filter is used to redirect the light from the pixels. Thus, groups of pixels (denoted as *view*) are seen only from a specific angle. For mobile devices, normally watched by single observer, two independent views are sufficient for satisfactory 3D perception. In order to be shown on a stereoscopic display, the images intended for each eye should be spatially multiplexed. This process is known as *interleaving* [8], and depends on the parameters of the optical filter.

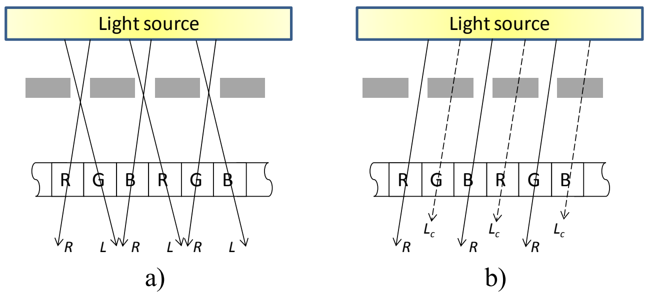
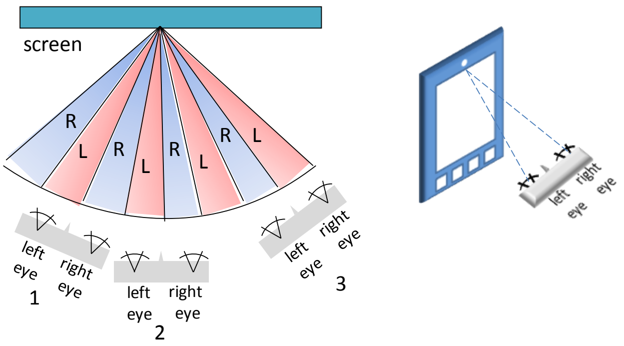


Figure 1. Parallax barrier: a) operation principle and b) crosstalk towards the optimal observation angle.

The most common approach for separating the images intended for each eye utilizes a layer called *parallax barrier*. The barrier blocks the light in certain directions as shown in Figure 1a. Two separate views are formed as a result. Due to the repetitive structure of the parallax barrier, each view is seen from a number of observation angles, as illustrated in Figure 2. In order to perceive proper 3D image, the observer should be properly positioned with respect to the display (e.g. positions "1" or "2" but not "3" in Figure 2).



a) b)

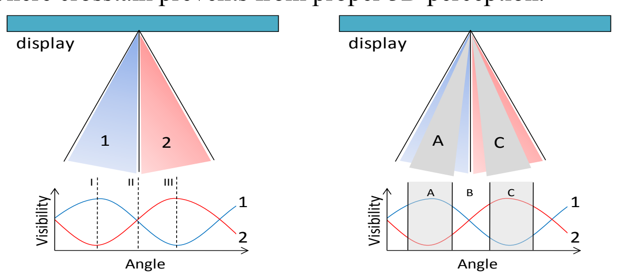
Figure 2. Position of the observer in respect to the display: a) visibility zones of views and b) position of the user as detected by the front facing camera

The parallax barrier is a cheap technology providing 2D backward compatibility through switching off the barrier.

2.1 Visual artefacts in parallax barrier-based displays

Crosstalk

Crosstalk is the effect of mixing the views. It is caused by imperfect optical separation of the views. The visual manifestation of crosstalk is a double-contoured, "ghost" images which significantly reduce the perceived 3D quality. There are two causes of crosstalk in parallax barrier-based displays. *First*, it arises when the display is observed from a position between two observation zones. The visibility of each view gradually changes as a function of the observation angle, as exemplified in Figure 3. At a certain angle, pixels of one view are fully visible, while the pixels of the other are fully covered by the parallax barrier. Such *optimal observation angle* for view 1 is marked by "I" in Figure 3, while the optimal observation angle of view "2" is marked by "III" in the same figure. At angle "II", both views are only partially covered by the barrier causing what we call *inter-zone crosstalk*. It reaches minimum at the optimal observation angle of a view, and maximum on the bisection between two neighbouring optimal observation angles. With respect to the inter-zone crosstalk, there are "high quality" areas, with no noticeable crosstalk (areas "A" and "C" in Figure 3b) and "low quality" areas ("B" in Figure 3b), where crosstalk prevents from proper 3D perception.



a) b)

Figure 3. Crosstalk versus observation angle: a) visibility of a view as a function of the observation angle and b) "high-quality" zones with low inter-zone crosstalk

The second cause for crosstalk is the transparency of the parallax barrier. It is usually implemented as a second, not fully opaque, LCD layer. Even at an optimal angle, part of the light passes through the barrier as illustrated in Figure 1b. This amount of *minimum crosstalk* is always presented in view (cf. Figure 3).

Pseudoscopy

Regarding Figure 2a, positions "1" and "2" are proper for perceiving 3D effect. However, at position "3" in the same figure, an observer will see the "left" image with the right eye and vice versa, thus perceiving a *pseudoscopic image* (aka *reverse stereo*). Both the observation zones of the two views and the correct and pseudoscopic positions alternate. In between each correct or pseudoscopic zone an inter-zone crosstalk is perceived as exemplified in Figure 4. Moving away from a correct observation (e.g."C" in Figure 4), the observer passes through a zone where high crosstalk is visible, and then falls into zone with low amount of crosstalk, but incorrect (pseudo) stereo ("P" zone). This effect causes what is perhaps the biggest inconvenience with parallax barrier-based displays. Most observers have the instinctive ability to move away from zones where low-quality, ghosting impaired image is seen. However, pseudoscopic stereo is not immediately perceived as "bad", which might cause a user to stay at an angle maintaining low crosstalk but wrong stereo.

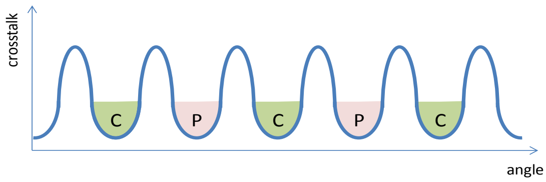


Figure 4. "Correct stereo" and "Pseudoscopic stereo" zones as function of the observation angle

2.2 Stereoscopic 3D LCD in our system

Technical data

The display model used in our system is Stereoscopic 3D LCD MB403M0117135, produced by masterImage [7]. It is 4.3" WVGA autostereoscopic display with switchable parallax barrier, which can operate in 2D or 3D mode. Additionally, the parallax barrier of the 3D LCD module can be switched between "3D horizontal" and "3D vertical" mode, thus operating in landscape 3D or portrait 3D mode. Two signals, "Chip Select" and "Mode Select" determine the mode of the display. The combinations are given in **Table 1**.

Table 1 – display modes of 3D LCD module

CS (Chip Select)	MS (Mode Select)	Display mode
Low	Low	3D Horizontal
Low	High	3D Vertical
High	Low	2D
High	High	2D

Artefacts quantified

We have measured the crosstalk and the angles between the observation zones using the methodology in [9], [10]. We found the optimal observation distance of the display to be 42 cm. At the optimal observation distance, the optimal observation points of each view found are shown in Figure 5.

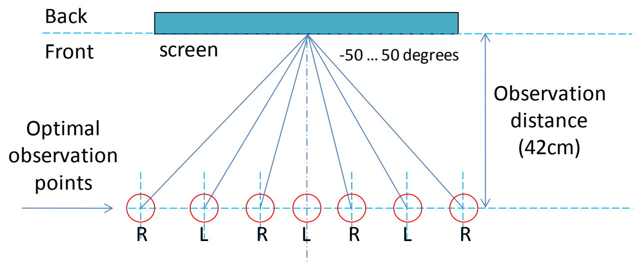
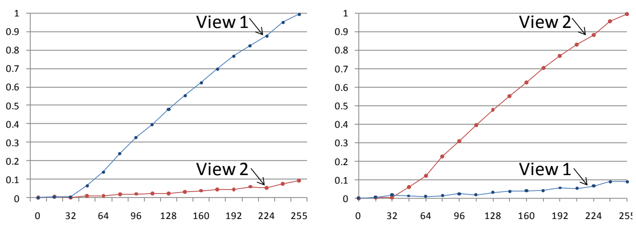


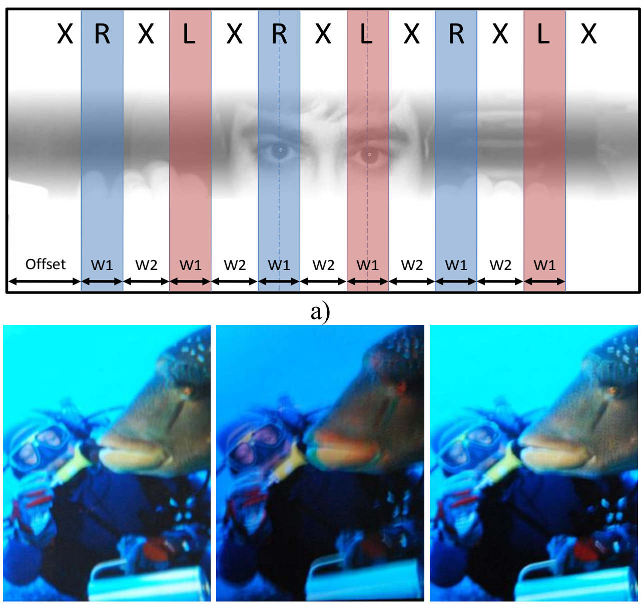
Figure 5. Observation points used in measurements

The angle between two neighbouring optimal observation vectors is 9.4°. The minimum crosstalk is 9%, and it is symmetrical with respect to the channels. The results of the crosstalk measurements at the optimal observation points of both views are shown in Figure 6.



a) b)

Figure 6. Minimum crosstalk: a) view 2 introduced in view 1 and b) view 1 introduced in view 2



a)

b) c) d)

Figure 7. a) Map of "high quality" observation zones with low inter-zone crosstalk, b) image seen from area "L", c) image seen from area "X" and d) image seen from area "R" of the map

By subjective evaluation, we measured the width of the "high quality", ghost artefact-free areas ("A" and "C in Figure 2b). A test 3D image was shown on the display, and a front facing camera was mounted on the device as shown in Figure 2b. An observer assessed the image looking with one eye from the optimal observation point of one of the views. Then he started moving to the left, till noticeable ghosting was appeared in the image. The position of the pupil was recorded by taking a snapshot of the observer's eye with the camera. The right border of the ghost-free was found in a similar manner. The process was repeated for all optimal observation points, which resulted in a map of areas where the pupil of the observer must reside in order to perceive image with no hosting artefacts. The measured map is shown in Figure 7a. When the camera operates in VGA resolution, the width of ghost-free zones (marked with "W1" in Figure 7) is 20px, separated by crosstalk-impaired zones (marked with "W2") of 31px each. Images in Figure 7b, c and d show photos taken from zones "L", "X", and "R" correspondingly, and give example of the inter-zone crosstalk observed in between the "high quality" zones.

3. VISUAL OPTIMIZATION FOR 3D LCD DISPLAY

We propose a system for visual optimization of stereo imagery for an autostereoscopic display. The system is based on OMAP 3430 SDP and integrates a 3D LCD module and front-mounted camera with VGA resolution. The system tracks the position of the observer's eyes, and adapts the system to avoid three cases of visual discomfort.

Reverse stereo is avoided by simply flipping the left and right channel based on eye detected being at the opposite view zone, see Figure 8a. The pseudoscopic regions (marked with "P" in Figure 4) are replaced with zones where both channels are flipped (marked with "F" in Figure 8a), thus allowing correct stereo image to be perceived.

Ghosting artefacts are avoided by switching the display into "2D" mode, in cases when the observer's eyes falls into area with pronounced crosstalk, where the 3D perception is anyway impossible (see in Figure 7a, areas marked by 'X' and Figure 8b).

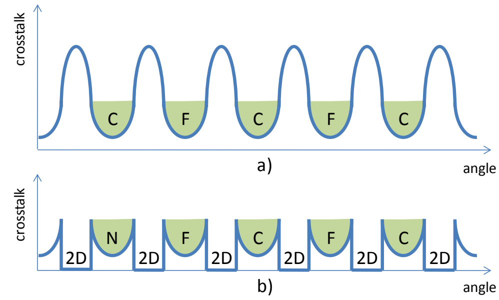


Figure 8. 3D image correction following the position of the observer's eyes: a) correction for the pseudoscopic regions, and b) correction for the regions with inter-zone crosstalk

Finally, making use of the 3D LCD module ability to switch between horizontal 3D and vertical 3D modes, our system selects 3D mode and scene orientation according to the orientation of the observer's eyes, as illustrated in Figure 9. When the face of the observer is not in horizontal of vertical direction in respect to the display, 3D effect is not possible, and thus the system switches the display into 2D mode.

The block diagram of the algorithm is shown in Figure 10. It goes through the following stages:

1. Face detection is attempted four times, each time rotating the camera image at a right angle. If face is not detected, it is possible that the face of the observer is at a wrong angle or too far away from the centre of the display. In both cases 3D perception is not possible, and the system switches the display into 2D mode.

2. If face detection is successful, its direction is stored, and eye tracking is performed according to the direction.

3. The position of the eyes is matched against the map of “high quality” observation regions. The map in use is selected to match the direction of the face.

4. If both eyes are found in the corresponding regions, the system switches into 3D mode. If both eyes appear in the regions of the opposite view, the system flips the channels and switched into 3D mode. If both eyes fall into the observation zone of the same view, or at least one eye falls in an inter-zone crosstalk area, the system switches into 2D mode.

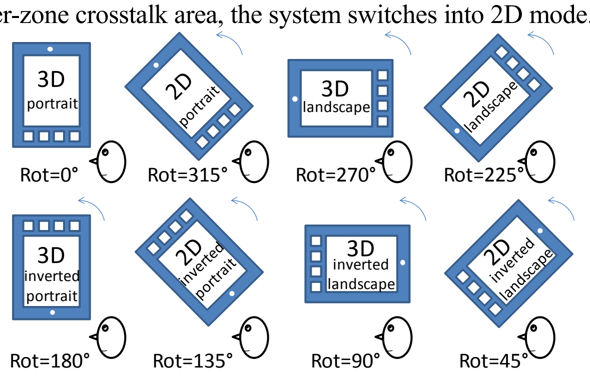


Figure 9. Selection of 3D mode and scene orientation according to the orientation of the observer’s head

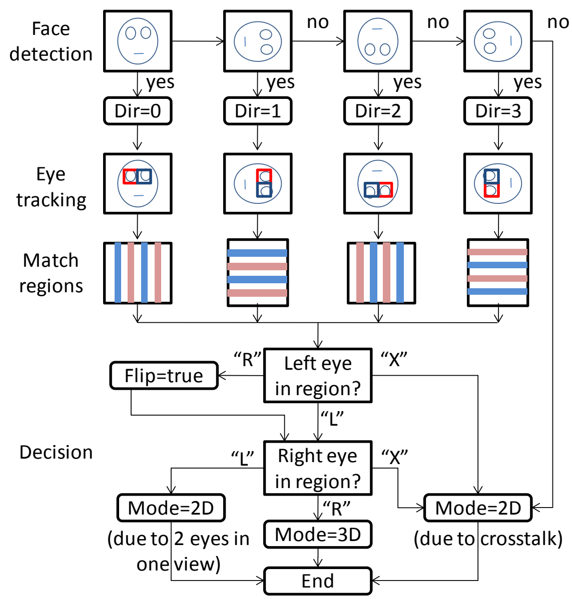


Figure 10. Block diagram of the proposed algorithm for visual optimization

4. IMPLEMENTATION

The system is implemented on the OMAP 3430 SDP running Linux OS (L12.20 baseline release). A parallax-barrier auto-stereoscopic display has been integrated to the platform [11]. OMAP 3430 is a dual-core processor, which includes general purpose ARM core and a TMS320 compatible DSP core. The ARM side provides access to C compiler and Linux environment, which allows code from existing open source libraries to be reused. The ARM side has been used for code prototyping while time-critical functions has been ported to the DSP in a block-by-block fashion. The dual-core architecture allows the output of both implementations to be compared and simplifies the debugging process.

The application processing modules have been distributed between the ARM and DSP processors as shown in Figure 11, aiming at an efficient implementation. The ARM side is engaged by the Linux OS. It is also responsible for maintaining the camera images, detecting the face and generating the stereo views. The DSP is engaged by the computationally-intensive eye detection algorithm. An effective inter-processor communication protocol is used through a queued mailbox-interrupt mechanism [16].

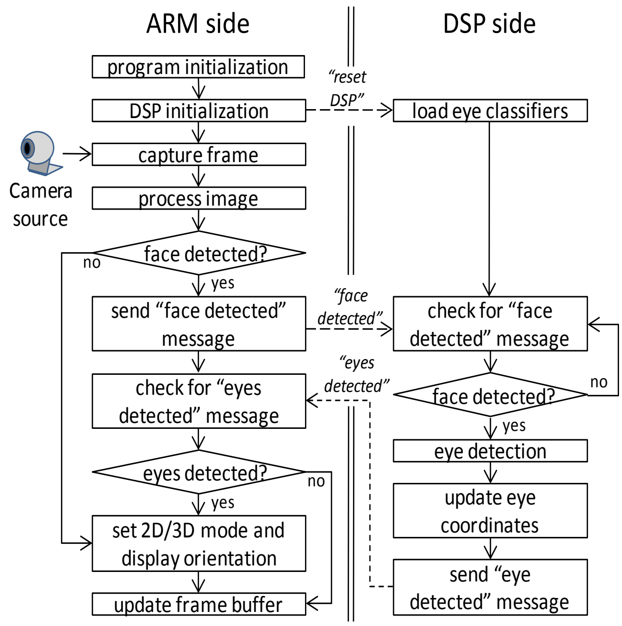


Figure 11. Application flow diagram of DSP and ARM.

4.1 Face and eye detection

We have ported an OpenCV realization of face detection algorithm by Viola and Jones [17] by modifying the classifiers to use fixed point arithmetic. Our own face detection algorithm is being ported to the OMAP as well. It is based on a two-stage hybrid technique, combing skin detection with feature-based face detection [12], [13]. In our face detection implementation, the search for faces is done for a subset of face sizes – limited by the expected facial size of an observer within the visual comfort zone for the 3D display. It applies a large-to-small scale search strategy, and the search is satisfied by the first face found, thus ensuring that the display mode is set appropriately for the *closest* observer.

The eye detection is implemented on the DSP core. It detects the two pupils by a Bayesian classifier working on Dual-Tree Complex Wavelet Transform (DT-CWT) features. The DT-CWT has been chosen as a low-cost alternative to Gabor transform for real-time feature extraction implementation [14], [15]. The DT-CWT features are formed by a four-scale DT-CWT applied on a spatial area of 16x16 pixels around a landmark., with six differently-oriented sub-bands per scale. The resulting twenty four matrices as shown in Fig. 5 form a landmark jet [14], [15].

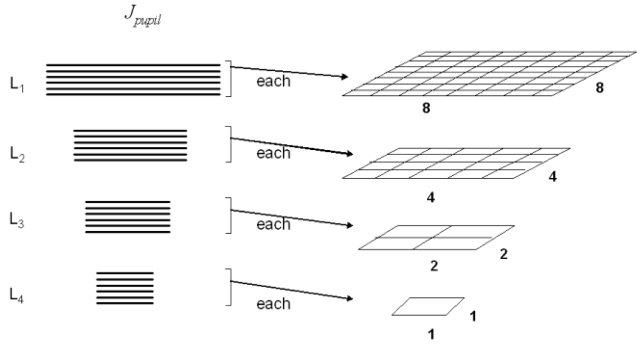


Figure 12. DT-CWT based feature jet.

Two landmark classes are modeled: pupil and non-pupil respectively. For modeling a particular landmark class, we have trained Gaussian mixture model (GMM) for each sub-band in the jet, thus leading to 24 models. We have used utmost 5 Gaussian components for each slice.

The positions of the detected eyes are returned back to the ARM by the by means of the shared SDRAM. Based on the position of the eyes, the stereo images are properly manipulated, as described in the algorithm of Section 3. The view rendering is implemented by a direct DMA operation [11].

5. CONCLUSIONS

We presented a system for optimizing the image on a 3D LCD display module by using eye-tracking and adapting to the position of the observer's eyes.

Instead of presenting an improper, low quality stereo-pair, our system switches to 3D mode only if the 3D effect is guaranteed. The resulting system will deliver 3D image only when looked at from a set of observation angles, but will avoid confusing the user by showing him low-quality image when 3D perception would not be otherwise possible.

An efficient implementation on the OMAP 3430 platform has been targeted by splitting the processes between the ARM and DSP cores. The system is part of bigger system for playing stereo video content on 3D mobile device.

6. ACKNOWLEDGMENT

This work is supported by the European Commission within the ICT Programme of FP7 under Grant 216503 with the acronym MOBILE3DTV.

REFERENCES

[1] B. Schiffman, "Movie Industry Doubles Down on 3-D", Wired magazine, April 2008, available at http://www.wired.com/techbiz/media/news/2008/04/3d_movies

[2] The Illustrated 3D Movie List, available online at http://www.3dmovielist.com/list.html

[3] "Mitsubishi Digital Electronics America Showcases Large-Screen 3D-Ready HDTV at CES 2008", press release by Mitsubishi Electric, Dec 2008, available at: www.mitsubishi-tv.com/pressreleases.html

[4] 3D DLP HDTV white paper, Texas Instruments, 2009, available online at http://www.dlp.com/hdtv/3-d_dlp_hdtv.aspx

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

[6] 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

[7] "Stereoscopic 3D LCD Display module", Product Brochure, masterImage, 2009, available online at http://www.masterimage.co.kr/new_eng/product/module.htm

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

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

[10] A. Boev, A. Gotchev and K. Egiazarian "Stereoscopic Artifacts on portable auto-stereoscopic displays: what matters", In Proc. of VPQM 2009, Scottsdale, AZ, USA.

[11] A. Gotchev, A. Tikanmäki, A. Boev, K. Egiazarian, I. Pushkarov, N. Daskalov, 'Mobile 3DTV technology demonstrator based on OMAP 3430', submitted to DSP 2009 conference.

[12] A. Boev, M. Georgiev, A. Gotchev and K. Egiazarian, 'Optimized single viewer mode of multiview autostereoscopic display', in Proc. EUSIPCO 2008, August, Lausanne, Switzerland.

[13] V. Uzunov, A. Gotchev, K. Egiazarian, J.Astola "Face Detection by Optimal Atomic Decomposition", Proceedings of the SPIE, Volume 5916, pp. 160-171 (2005).

[14] N G Kingsbury, Complex wavelets for shift invariant analysis and filtering of signals, *Journal of Applied and Computational Harmonic Analysis*, vol. 10, no 3, May 2001, pp. 234-253.

[15] H. Essaky Sankaran, A. Gotchev, K. Egiazarian, and J. Astola, "Complex wavelets versus Gabor wavelets for facial feature extraction: a comparative study", Proceedings of SPIE, vol. 5672, pp. 407-415, 2005.

[16] OMAP35x Applications Processor Interprocessor Communication (IPC) Module, Technical Reference Manual, TI.

[17] P. Viola, M. Jones, "Robust Real-time Object Detection," In proc. J. Of computer vision, 2001