Efficient and Robust Detection of Users Wearing Stereoscopic Glasses for Automatic 2D/3D Display Mode Selection

Zoran Zivkovic and Erwin Bellers
Trident Microsystems (Europe) B.V., The Netherlands

***Abstract*--A system for automatic 2D/3D display mode selection is presented. Most 3D displays require special glasses to create an illusion of observing a 3D scene. The proposed system is using a camera to automatically detect if there are users in front of the display wearing the special 3D viewing glasses. An efficient algorithm is described to detect faces wearing the glasses in the camera images. Based on the detection, the display automatically switches between the 2D and 3D viewing modes.**

I. INTRODUCTION

The glasses-free 3D displays are currently mainly used for small handheld devices where only narrow viewing angles need to be considered. Most of the larger 3D capable displays require the special glasses. The glasses are used to filter the displayed content such that different images are shown to the user left and right eye, creating an illusion of observing a 3D scene.

A 3D capable display usually operates either in the 2D or 3D mode. The 3D video material can be presented as 2D, for example showing just the image intended for one eye. A 2D video can be converted to 3D using some 2D to 3D conversion technique. Watching the display in 3D mode without the special 3D viewing glasses is unpleasant since the left and right eye images intended for 3D are observed by both eyes. A remote control is usually used to switch between the 2D and 3D viewing modes.

A 3D display system is presented here that can perform automatic switching between the 2D and 3D viewing modes. The users wearing the special 3D viewing glasses are detected using a camera and this information is used to switch between the 2D and 3D modes, see Figure 1. Practical realization of the system is discussed here.

II. EFFICIENT DETECTION OF USERS WEARING 3D GLASSES

Detecting the users wearing the 3D glasses is related to the widely studied problem of face detection. The Viola-Jones algorithm [1] is the most popular face detection algorithm that inspired many consumer electronic implementations. This algorithm will serve as the starting point here.

The Viola-Jones detector is based on a statistical classification procedure that is trained to detect faces using a large set of example face images. A set of example images of people wearing glasses can be made and the detector can be trained to detect users wearing the 3D glasses. However, there are two important practical constraints that will be considered here. First, the detection of the users wearing the 3D glasses needs to be very robust and reliable. Second, the detection should be very efficient; such that it could, for example, run on the embedded processor already present in the display.

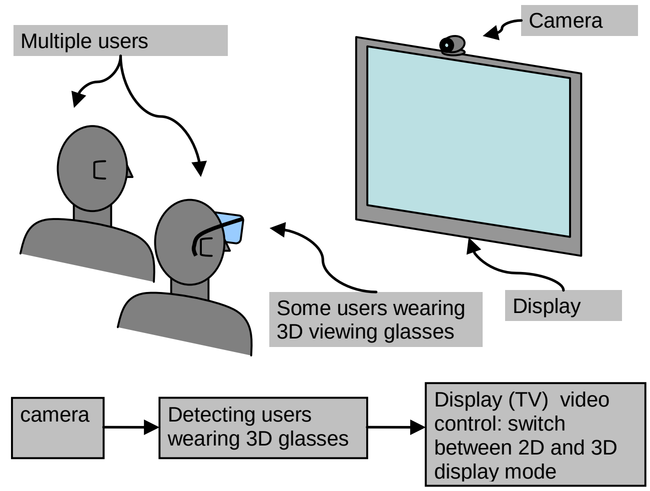

Fig. 1. User friendly automatic 2D/3D display mode selection.

Adding special markers to the glasses is proposed for easier detection and highly efficient detection algorithm is presented.

A. *3D glasses with special markers for easier detection*

If the glasses have special markers, the efficiency and robustness of detection can be greatly improved. The special markers make it also easier to distinguish the 3D glasses from other dark glasses. Furthermore, if the markers are made of reflective material, it will be possible to detect the glasses even in low light conditions typical for a living room.

There are many ways to design the special markers. The markers should be large enough to be possible to detect them when the user is at 4-5 meters from the display. A practical design proposed here is to make the whole upper part of the glasses of a reflective material, see Figure 2. Since the human skin has no blue pigment, the upper part of the glasses is made form a blue reflective material for easier detection.

B. *Efficient detection*

The Viola-Jones detector [1] can be reasonably efficient when only up-front faces are detected. Even then the storage and memory requirements are large. For detecting also in-plane rotated faces [5][6] the costs increase significantly. Many hardware and software solutions are proposed to improve the efficiency of the detector and limit the memory bandwidth costs, e.g. [2][3][4].

When the glasses have special, easy to detect, markers, it will be typically possible to detect the glasses using other techniques much more efficiently than using the Viola-Jones detector directly. Once the glasses are detected the Viola-Jones detector trained of faces with glasses can be used to decide if it is a person wearing the glasses. Huge gain in efficiency is obtained since the detector is applied only to a

Digest of ICCE 2012

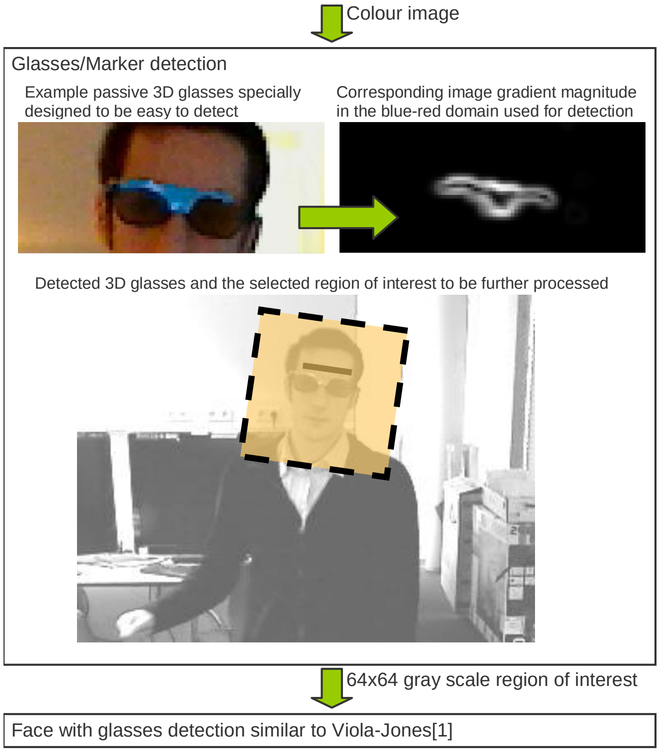

Fig. 2. Efficient detection of users wearing specially designed 3D viewing glasses.

small region of interest where the glasses are detected, see Figure 2.

For the glasses marker design proposed here, line detector can be used for fast detection. Since the glasses are of blue reflective material and human skin contains usually red pigment, the line detection is performed by first detecting gradient in image obtained by subtracting for each pixel red color from the blue color. Figure 2, shows an example part of an image and the computed gradient using a 5x5 Gaussian kernel gradient detector. The reflective part of the glasses becomes clearly visible. The line segments are detected by applying a threshold to the gradient image to extract only the strong gradient points and using random sampling approach RANSAC for the fast line detection [7].

In the second processing stage, for each detected line a region of interest is extracted and scaled to a small gray scale image. This small image is then used to apply the Viola-Jones detector and decide if indeed the detected line corresponds to a face with the 3D glasses, see Figure 2. The second stage of the algorithm is evaluated for each detected line. As result, the face detector is executed for a few times but on very small images. The computation costs of the second stage will depend on the number of detected lines. In our tests on average there were 5 lines detected per image.

C. Evaluation

The OpenCV c-code [8] of the Viola-Jones detector is used to train the face-with-glasses detector and to execute it, see Figure 3 for example images used for the training. Testing the processing speed was done by measuring the average execution time on a 2GHz PC. The standard face detector required on average 50ms for a VGA image when only up-front faces are detected. The presented approach required on average only 5ms. Since the orientation of the detected glasses marker give the estimate of the face in-plane rotation, the rotated faces are also detected.

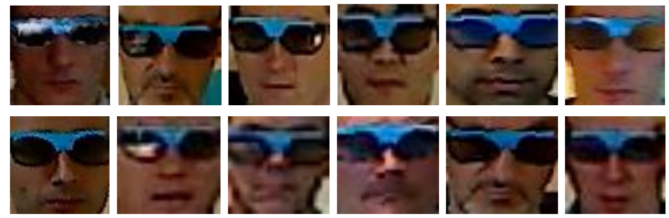

Fig. 3. Example faces with the specially designed 3D glasses used for training of the classification algorithm.

To evaluate the proposed scheme, a set of videos is made containing people wearing the special glasses. Some of the videos were recorded in very low light conditions to simulate common situation in a living room. For 1000 extracted test frames of up-front faces the detection rate was 99% and equal to running the Viola-Jones detector directly. For another 1000 frame set of rotated faces the recognition rate remained the same, while the Viola-Jones detector applied directly detected only 30% of the faces.

The algorithm is also compiled for a small 24K MIPS embedded processor. The algorithm required just around 15 million cycles per frame on average, making it suitable for running directly on the embedded processor already present in many displays.

III. CONCLUSIONS

Automatic 2D/3D mode selection is described that can improve user experience when watching the standard glasses-based 3D displays. A camera is used to detect if there are users in front of the display wearing the special 3D viewing glasses. Based on the detection, the display automatically switches between the 2D and 3D viewing modes. An efficient detection algorithm is described suitable for embedded processor implementation. Special markers are added to the glasses to make the efficient implementation possible and also increase the robustness of the detection.

REFERENCES

[1] P. Viola, M. Jones, "Robust Real-Time Face Detection", International Journal of Computer Vision 57(2), 137–154, 2004.

[2] J. Cho, B. Benson, R. Kastner, "Hardware acceleration of multi-view face detection,", *SASP '09. IEEE 7th Symposium on Application Specific Processors*, pp.66-69, 27-28 July 2009.

[3] C. He, A. Papakonstantinou, D. Chen, "A novel SoC architecture on FPGA for ultra fast face detection," *IEEE International Conference on Computer Design*, pp.412-418, 4-7 Oct. 2009.

[4] H.J.W. Belt, "Storage Size Reduction for the Integral Image", Philips Research Technical Note TN-2007/00784. Issued: 12/2007.

[5] H. Rowley, S. Baluja, and T. Kanade, "Rotation Invariant Neural Network-Based Face Detection, " IEEE Conference on Computer Vision and Pattern Recognition, pp. 38-44, 1998.

[6] B. Li, A. Yang, J. Yang, "Rotated Face Detection Using AdaBoost," *2nd International Conference on Information Engineering and Computer Science (ICIECS)*., pp.1-4, 25-26 Dec. 2010.

[7] M. Aly, "Real time Detection of Lane Markers in Urban Streets". IEEE Intelligent Vehicles Symposium, The Netherlands, June 2008.

[8] OpenCV Library, http://opencv.willowgarage.com