論文
9-3-2

View Point Tracking for Parallax Barrier Display Using a Low Cost 3D Imager

위성민(Sung-Min Wi)[1] 김동욱(Dong-Wook Kim)[2]

ABSTRACT

We present an eye tracking system using a low cost 3D CMOS imager for 3D displays that ensures a correct auto stereoscopic view of position- dependent stereoscopic 3D images. The tracker is capable of segmenting the foreground objects (viewer) from background objects using their relative distance from the camera. The tracker is a novel 3D CMOS Image Sensor based on Time of Flight (TOF) principle using innovating photon gating techniques. The basic feature incorporates real time depth imaging by capturing the shape of a light-pulse front as it is reflected from a three dimensional object. The basic architecture and main building blocks of a real time depth CMOS pixel are described. For this application, we use a stereoscopic type of display using parallax barrier elements that is described as well.

논문 접수 : 2008. 8. 4.
심사 완료 : 2008. 8. 14.

1) *Dept. of Electronic Materials Eng.,Kwangwoon University*
2) *Dept. of Electronic Materials Eng.,Kwangwoon University*

1. Introduction

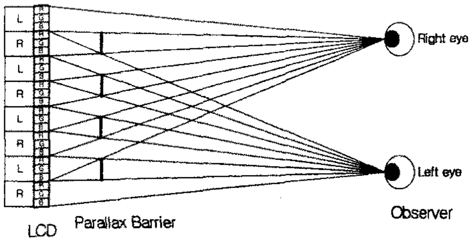
1.1. General

The European Project ATTEST (Advanced Tree-dimensional Television System Technology) that was carried out recently (2002) tried to evaluate the entire 3D Video chain for broadcast application. During this period, we at 3DV Systems developed an advanced version of the *Zcam* namely incorporating a new image shutter based on GaAs technology [1]. Recently, based on the same fundamental principles that first presented in 2001[2], a low cost version camera for gesture recognition using the GaAs was introduced [3][4]. In this paper, we will describe a 3D camera based on Si only namely using an innovative pixel approach applying the TOF measurement principle. In addition, a usage of such an imager for tracking application will be described with an emphasis on its depth dependent windowing capabilities.

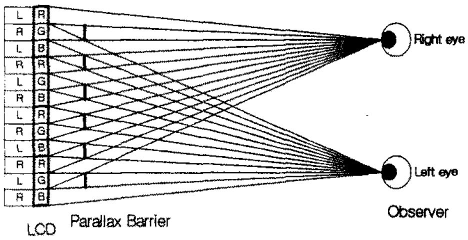
Auto stereoscopic display has a disadvantage that it tends to cause dizziness and users cannot feel a 3D effect due to reversed left and right images when the observers move over. In order to solve this problem, eye-tracking system is generally utilized [5][6]. In a method of eye tracking system the problem is solved with a technique of swapping right and left image by extracting a location of observer's eyes. In this paper, depth sensor camera that is capable of measuring the distance away from the camera in real time is utilized to implement eye-tracking system. With the depth sensor, background and face are segmented in real time and precise location is informed to PC. It enables that right and left images of the parallax barrier display are swapped to be aligned with eye's location so that the 3D effect is smoothly observed.

1.2. The Display

The tracker described in this paper is coupled to a 3D display. The input for the display are stereoscopic images, which consist of left and right images, and those images are interlaced to provide three dimensional effects using direction-selective elements such as parallax barriers or lenticular sheets [7][8]. Parallax barriers, form of fixed film or TN(Twisted Nematic) panel, have a function as a spatial de-multiplexer to separate left-eye and right-eye views from a 3D scene. Therefore, in this paper, we make the auto stereoscopic display system by the parallax barrier method using the TN panel to display stereoscopic images both in 2D and in 3D modes. The auto stereoscopic display system does not require wearing any device to separate left-eye and right-eye views, that is, the auto stereoscopic system sends those views to the corresponding eyes. Typical emissive displays radiate light equally in all directions. In order to create a twin-view auto stereoscopic display, a half of pixels must only radiate light in directions seen by the left eye and the rest of the pixels in directions seen by the right eye. The parallax barrier is the simplest way to block light using strips of the black mask.



(a)Pixel-based system



(b)Sub-pixel-based system

Fig.1. Autostereoscopic display panel geometry using parallax barrier method

The principle of the two-view parallax barrier is illustrated in Fig. 1. Fig. 1(a) is a stereoscopic display based on pixels and Fig. 1 (b) is a stereoscopic display based on sub-pixels. The sub-pixel based stereoscopic display can reduce annoying effects from strips of the black mask. The left and right images are interlaced in columns on the display. The barrier is positioned in order that left pixels of the image are blocked from the region of the right viewing windows and vice versa. The auto stereoscopic display system has the capability to switch from the 2D mode to the 3D mode electronically and vice versa [9]. Fig.2 shows the structure of the auto-stereoscopic display panel. The light is filtered at the first polarizer, modulated by liquid crystal, and filtered at the second polarizer. Here, the characteristic of the first polarizer is opposite to that of the second polarizer. In the 2D mode, since the TN panel does not operate, the modulated color light passes the TN panel and the third polarizer without any changes. Therefore, the auto stereoscopic display system operates like conventional LCD displays. In the 3D mode, however, the auto stereoscopic display system sends left-eye and right-eye images to the corresponding eye because the TN panel operates as the direction-selective element. This kind of technology has been commercialized in mobile phones and monitors [10]-[12].

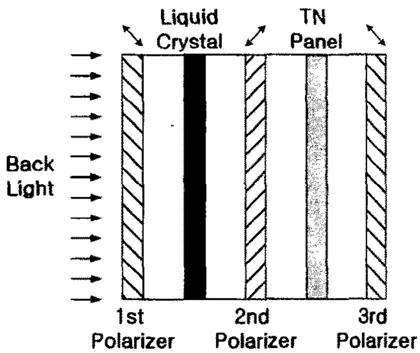


Fig. 2. Structure of autostereoscopic display panel

However, in order to get the correct position for a good stereoscopic image, scene must be rendered according to the observer's viewing angle. There is a 50 percent chance the observer will be in the correct position. The observer must stay in this position to get a good 3D perception. In the wrong position, the observer has to move his or her eye. The eye-tracking display catches the position of the observer's eye, it can display the right and left images in the appropriate zones, thus preventing pseudoscopic viewing.

2. The 3D tracker

2.1. Depth Measurement Principle

A novel small pixel 3D CMOS based image sensor is described. The concept is based on the Time-of-Flight(TOF) principle using innovating photon gating techniques The basic feature incorporates real time depth imaging by capturing the shape of a light-pulse front as it is reflected from a three dimensional object. The basic architecture and main building blocks of a real time depth CMOS pixel is described. A CMOS based image sensor capable of handling extremely short integration times then detects the reflected light. In order to provide sufficient signal to noise ratio (SNR), multiple light pulses are employed and the signals are then accumulated to a storage node within the pixel. The captured gray level image has to be normalized to account for effects such as changes in the scene's reflection coefficient, illumination/detector non-uniformity and the slope of the light edge. Capturing a second ungated image that contains all these effects without the effect of the gating performs normalization per pixel. Thus, 2 storage units are incorporated to store the gated signals and un-gated signal. If the background radiation is interfering, an additional signal should be measured and stored, namely at interval where there is no light pulse returning from the scene. The main parameters in the design include transfer time from the photodetector to the storage nodes and implementing the transfer gates with very fast clock rise/fall times. Solutions to these issues rely on the combination of careful pixel physical layout design in conjunction with special process techniques. These are incorporated in the pixel design to culminate in a high performance pixel for 3D applications.

The measurement principles are as follows: a light pulse of short duration is generated by some point light source onto a 3D scene. The light wall, emanating at the light source position is actually an expanding spherical surface of finite width as seen in Fig. 3. This wall hits the 3D objects in the scene, and it is reflected from the objects surface, carrying an imprint of the objects shape, Fig. 4. The imprint contains all the information required for the reconstruction of the scene structure. For example, each 1 mm difference in position of an object is translated to about 3.3 pico-second delay of the light wall. Since the shutter truncates the reflected light, the image sensor integrates more light for closer objects. Thus, the image shape of the returning light wall can be reconstructed to the original scene shape. The reflected light wall moves towards the camera carrying the image imprint, it enters the camera and before impinging on the imager the gate shuts off letting the front end only hit the imager. The result of the gating is a gray level image having intensity of each pixel proportional to its depth. For example, the tip of the nose will be brighter than the eye socket. The captured gray level image has to be normalized to account for effects such as changes in the scene's reflection coefficient, illumination/detector non- uniformity and the slope of the wall edge. Capturing a third un-gated image that contains all these effects without the effect of the gating performs normalization per pixel. A conceptual schematic circuit and corresponding control signals are shown in Fig. 5 and Fig. 6, respectively.

The basic operation of the pixel relies on the photoelectric effect where due to its physical properties silicon is able to detect photons of light. The photons generate electron-hole pairs in direct proportion to the intensity and wavelength of the incident illumination. The application of an appropriate bias allows users to collect the electrons and meter the charge in the form of a useful parameter such as voltage. The principle of the novel 3D CMOS Image Sensor is to mimic or perform the charge transfer similar to the CCD transfer using standard CMOS processing. A true charge coupling can be implemented in CMOS pixel designs by incorporating the Pinned Photo Diode concept. The Pinned Photodiode is the sensing element and is connected to the charge collecting nodes by transfer gate transistors that are specifically designed and fabricated. These esoteric transfer gate transistors in conjunction with the Pinned Photodiode design allow complete charge transfer from the photodiode towards the charge storage nodes (Fig. 5).

The sensor depends on a light pulse source with a known frequency. These light pulses are transmitted to illuminate a scene, which will then reflect it back to the sensor. Thus, the pulses have strong correlation with the switching time of switches f1 and f2. Figure.6 shows the related control signals that drive the 3DV sensor.

Fig.3 "Light Wall" moving towards 3D scene

Fig.4 "Light Wall" reflected from 3D scene

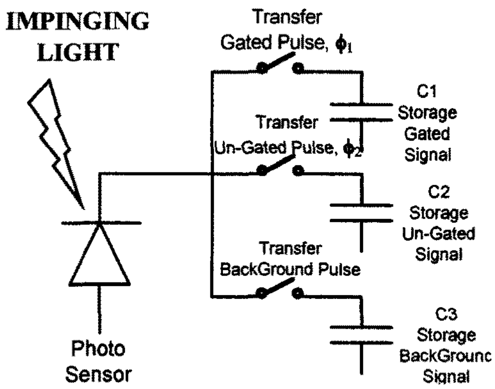


Fig. 5. 3D pixel including gated, un-gated and background storage bins

The reflected or received pulse will arrive to the sensor with a delay of TD. When pulse f1 is switched on, the accumulated charge is transferred to storage node C1. This is known as the front cut of the reflected signal. When pulse f2 is switched on, the accumulated charge is transferred to storage node C2. This is known as the back cut of the reflected signal. By knowing the front cut, back cut and background signals the range per pixel can be calculated.

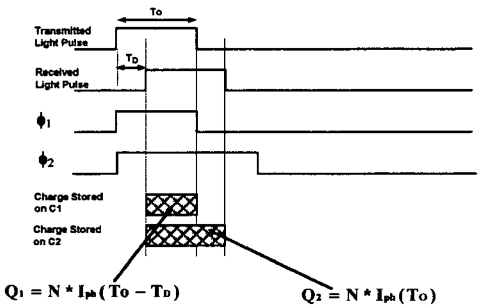


Fig. 6. Control signals for the 3D sensor

2.2. The CMOS Imager

2.2.1. Introduction

Noiseless complete charge transfer of photo-generated charge to a storage capacitor node has been performed by the well known CCD technology. This technology includes transfer of charge via double or triple level polysilicon gates. The transfer is performed by formation potential wells under the gates. By sequentially, forming and collapsing the potential wells complete charge transfer is performed.

The charge-coupling process is completely noiseless and perfectly transfers charge in silicon without loss (down to the single electron). Such fundamental qualities have yielded a technology with outstanding performance. However, the esoteric CCD technology comprises special processing techniques such as double or triple polysilicon layers, non-standard substrate doping and thermal annealing procedures. These processing techniques are incompatible with state-of-the-art standard CMOS technology.

2.2.2. The Pixel

The basic operation of the pixel relies on the photoelectric effect where due to its physical properties silicon is able to detect photons of light. The photons generate electron-hole pairs in direct proportion to the intensity and wavelength of the incident illumination. The application of an appropriate bias allows users to collect the electrons and meter the charge in the form of a useful parameter such as voltage.

The principle of the novel 3D CMOS Image Sensor is to mimic or perform the charge transfer similar to the CCD transfer using standard CMOS processing. A true charge coupling can be implemented in CMOS pixel designs by incorporating the Pinned Photo Diode concept (see Fig. 7). In the figure, signal charge is generated and collected in a pinned photodiode region that is at a fixed potential. The transfer gate, when activated high, transfers charge from the pinned region to a sense node. In that manner, the pinned region is fully depleted of majority carriers, complete charge transfer via electric field potential difference occurs. Since the photodiode is fully depleted, no kTC noise exists and noiseless charge transfer is obtained. The transfer mechanism is shown schematically in Fig.8.

The Pinned Photodiode is the sensing element and is connected to the charge collecting nodes by transfer gate transistors that are specifically designed and fabricated. These esoteric transfer gate transistors in conjunction with the Pinned Photodiode design allow complete charge transfer from the photodiode towards the charge storage nodes.

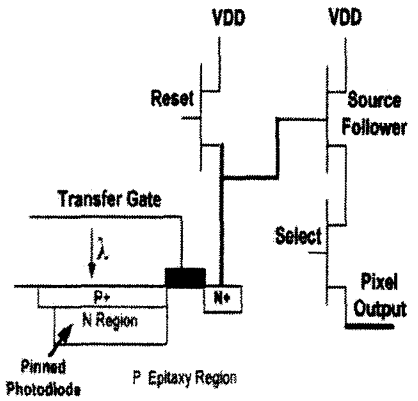


Fig. 7. CMOS pinned photodiode pixel

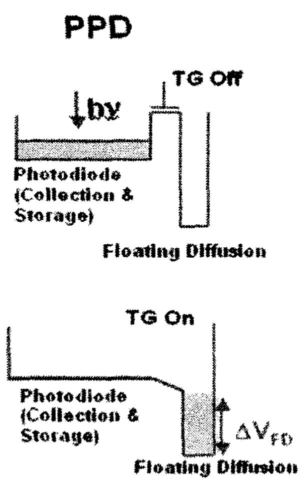


Fig. 8. PPD charge transfer

2.3. Depth Measurement Window

One of the most important features of the depth camera is the ability to change the parameters of the depth window dynamically. The depth window can be set to include certain objects, while disregarding objects that are outside the window, according to the scenario's need. In doing so, this sensor uses the entire dynamic range of the Imager, only on the measured volume. There is no "waste" of the dynamic range for measurement of the space, between the camera and the start of the depth measurement volume.

The setting of the depth measurement window width, determines the resulting depth resolution. The limit of the depth resolution will be the width of the measurement window, divided by the bit depth of the imager. For example, objects, which lie on the front of the measurement windows, will contribute maximum amount of photons to the gated image, while objects, which lies on the far end of the window, contribute almost no photos. The amount of different values that an object can generate in the gated image is equal to the bit depth of the Imager capturing the gated image, usually about 8 or 10 bits (256 or 1024 gray levels). An Imager comprising greater bit depth can be used to get even higher depth resolution at a given measurement window width. Using different measurement window widths, we can control the trade off between depth resolution and coverage area. A large depth window, allows the instantaneous capturing of a large scene, at relative course depth resolution. A narrow depth window allows the capturing of highly accurate depth maps of small scene details.

3. Experiments

Parallax barrier display has a disadvantage that it tends to cause eye fatigue and users cannot feel a 3D effect due to reversed left and right images when the viewers move. In order to solve this problem, eye-tracking system is generally utilized. In a method of eye tracking system the problem is solved with a technique of swapping right and left display image by extracting the observer's x-y-z direction. In this paper, depth sensor camera that is capable of measuring the distance away from the camera in real time is utilized to implement eye-tracking system. With the depth sensor, background and face are extracted in real time and precise location is informed to PC. It enables that right and left images of the parallax barrier display are swapped to be aligned with eye's location so that the 3D effect is smoothly observed. The method of general eye tracking is using template-matching system by extracting face region using color information. In the event of using ordinary camera for the image acquisition, the real time process is slightly delayed due to a plenty of calculation amounts in pre-processing. The pre-processing is to convert intensity of illumination on the distorted face to histogram equalization by light condition and to extract the face region through color domain conversion (RGB->HSI or RGB->YCrCv). In the event of using general web camera, over 15 frame image capture per second with the user program is difficult and as there are so much calculation amounts in pre-processing, eye tracking is, in fact, 10 frames per second. Main processing is the template matching method that is a process of separating eye's region and make coordinates.

In this paper, relatively, depth sensor algorithm is utilized in plenty of pre-processing calculations. In the event of using the depth camera, the calculation corresponding with face region extraction is processed in the sensor firmware in real time. Next, the eye recognition is the template matching method that extracts the coordinate's value on the separated face region so that the real time tracking is possible. In the experiment, the eyes coordinates value is calculated by tracking the observer's face by the depth sensor camera that is fixed on the top of the center of the monitor in indoor environment. The images transferring from the depth sensor camera to PC are the face region extracted from the firmware, from these regions, the right and left eyes coordinates are calculated on the basis of the template matching method. In the event of that the coordinates lies in the substitute eyes region, the disposition of the right and left images that will be displayed is determined by the location of the observers and monitor coordinates. The PC that controls the display converts re-deposited images to serial communication(RS-232) in alignment with 3D image output and controls 3D monitor. Accordingly, the reverse of right and left image that is the problem of auto stereoscopic display method is compensated to enable that the observers comfortably observe the 3D effect. The flowchart of the eye tracking system is depicted in Fig. 9. The depth sensor camera, 3D monitor and PC for the experimental eye tracking is depicted in Fig. 10. The captured images by the eye tracking using the depth sensor is depicted in Fig. 11. In the paper, 3DVsystem company's

DMC-100C for the depth sensor and Pavanine company's 19 inch parallax barrier display for 3D display are used on the purpose of the implementation of proposed method. Pentium 4 PC, Visual C++6.0 program and 320X240 16bit input image are utilized.

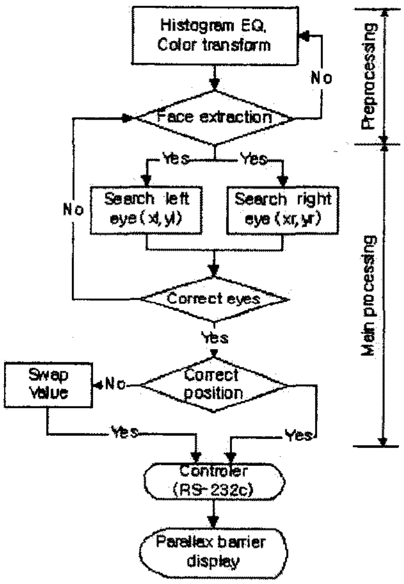


Fig. 9. Eye tracking flowchart

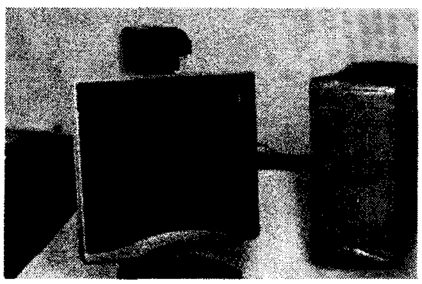


Fig. 10. Experimental system

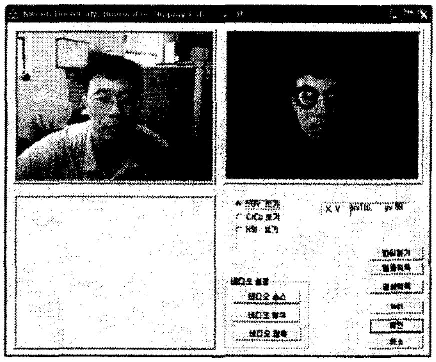


Fig. 11. Example of segmentation and eye tracking

4. Summary

We presented a novel low cost CMOS based camera that measures depth in real-time. This camera is used to generate a key signal for eye tracking application needed when using parallax barrier or similar 3D display techniques. For implementing real time eye tracking system, relatively, the depth sensor is utilized in pre-processing with the plenty of calculations. In the event of using the depth camera, the calculation corresponding with the eye recognition segmentation is processed in sensor's firmware in real time. Moreover, the eye recognition is the template matching method that enables the real time eye tracking by extracting the coordinate's value on the separated face region. The functionality of the novel camera described herein could be used as the input device for the 3D video as well and also as the gesture interface to the whole system.

< References >

[1] André Redert, Marc Op de Beeck, Christoph Fehn, Wijnand I Jsselsteijn, Marc Pollefeys, Luc Van Gool, Eyal Ofek, Ian Sexton, Philip Surman "ATTEST - Advanced Three-Dimensional Television Systems Technologies", In Proc.of. 3DPVT'02, pp.313-319, Padova, Italy, June 2002

[2] G. J. Iddan and G. Yahav, "3D imaging in the studio (and elsewhere...)," in Three-Dimensional Image Capture and Applications IV, B. D. Corner J. H. Nurre, and R. P. Pargas, eds., Proc. SPIE 4298, pp.48-55, 2001

[3] R. Gvili and A. Kaplan and E. Ofek and G. Yahav, "Depth Keying," Proc. SPIE 5006, pp.564-574, 2003

[4] G. Yahav, G.J. Iddan and D. Mandelboum, "3D Imaging Camera for Gaming Application", Consumer Electronics, 2007, ICCE 2007

[5] Y.S.Chen, C.H.Su, J.H. Chen, C.S. Chen, Y.P. Hung, and C.S. Fuh, "Video-based Eye Tracking for Autostereoscopic Displays," Optical Engineering, 40, pp.2726-2734, 2001

[6] M.Andiel, S.Hentschke, T.Elle, E.Fuchs, "Eye-Tracking for Autostereoscopic Displays using Web Cams", Proc. SPIE 4660, pp.200-206, 2002

[7] Woodgate, Graham J., Harrold Jonathan, Jacobs Adrian M., Moseley Richard R., and Ezra David, "Flat-Panel Autostereoscopic Displays: Characterization and Enhancement," Proc. SPIE 3957, pp.153-164, 2000

[8]Susumu, I., Nobuji, T., Morito, I.: Technique of Stereoscopic Image Display EP Pat., No.0354851, 1990

[9] Ezra, D., Woodgate, G.J., Omar, B.A., Holliman, N.S., Harrold, J., Shapiro, L.S. : New Autostereoscopic Display System. Sharp Technical Journal, vol. 62, pp.10-14, 1995

[10] D. H. Kim, H. J. Kang and C. H. Ahn, " A Stereoscopic Image Rendering Method for Autostereoscopic Mobile Devices", Mobile HCI 2004, vol.3160, pp. 441-445, 2004

[11] D.S.Kim, S.D.Se, K.H.Cha, and J.P.Ku, " 2D/3D Compatible display by autostereoscopy," Proc. K-IDS Three-Dimensional Display Workshop, pp.17-22, 2006

[12] O.H.Willemsen, S.T. de Zwart, M.G.H. Hiddink, D.K.G. de Boer, M.P.C. M.Krijin, " Multi-view 3D Display," Society for Information Display 2007 International Symposium Digest of Technical Paper, pp.1154-1157, 2007