World of Computer Science and Information Technology Journal (WCSIT)
ISSN: 2221-0741
Vol. 5, No. 4, 69-74, 2015

A Virtual TV System Using Head-mounted Display with Stereo Camera

Ryo HARASHIMA
Systems Information Engineering
Tokyo City University
Tokyo, Japan

Yue BAO
Faculty of Knowledge Engineering
Tokyo City University
Tokyo, Japan

Abstract— This paper propose a system to provide 3D display with unlimited parallax and display virtual TV as if real TV is put by HMD. By this system, the user is able to watch TV naturally by preventing which the moving of the virtual TV follows in the moving of head. In addition, motion parallax depending on movement is expressed without visual area restriction. The system is comprised of HMD and two Web cameras. Two Web cameras are attached to get surrounding scene. At first, an observer determines a characteristic point in the 3D image from the two cameras to set the virtual TV. The screen size of the virtual TV is set in the distance to the point calculated by triangulation using two cameras. The point is tracked by the phase correlation about anteroposterior consecutive frames. Parallax pictures are made to correspond to right side image and left side image with a size suitable for the distance. The virtual TV is displayed at the point in the input image. Two experiments were conducted to evaluate the performance of the proposal system. The experiment 1 is a experiment to evaluate image showed by virtual TV by subjects. In the experiment 2, the subjects observed two virtual cubes to confirm motion parallax. As a result, virtual TV position was fixed to the real space like the real TV and expression of the natural motion parallax was enabled by proposal system.

Keywords- Head-Mounted Display; Phase Only Correlation; Virtual TV; Mixed Reality; Motion Parallax; Stereo Camera.

I. Introduction

Recent years, stereoscopic displays have been developed and commercialized [1]. In 2010, 3D TV appeared on the market, and we have been enjoying TV and game console equipped with function of watching 3D pictures at home. We came to be able to get stronger presence than conventional 2D picture.

The stereoscopic display system includes glasses type and the naked eye type. By the glass type 3D observing the parallax images through glasses, observers can view binocular stereoscopic images. Observers can view the same stereoscopic image regardless of the point of observation. This system had been actually used at theaters or theme parks. However, it has some problems: Observers are required to wear special glasses. Because there are only two parallaxes, if observers move and view it from other points, the picture does not change, which makes observers feel less Reality.

On the other hand, about the naked eye type, observers can view the image stereoscopically at a plurality of observation positions without using special spectacles. Many stereoscopic display devices with naked eye have adopted a method called lenticular lens method. The 3D display which is put on the market by Toshiba Corporation also has adopted this method [2]. However, the lenticular lens can provide parallax only in Horizontal direction. Another problem is that the number of the displayed parallax cannot give enough motion parallax [3]. The range to view continuous stereoscopic image is limited to only main robe or side robes. In addition, a big display and large space are need when displaying a large size image. Because the screen size of TV grows big year by year in figure 1. The most suitable observation distance of TV is triple the height.

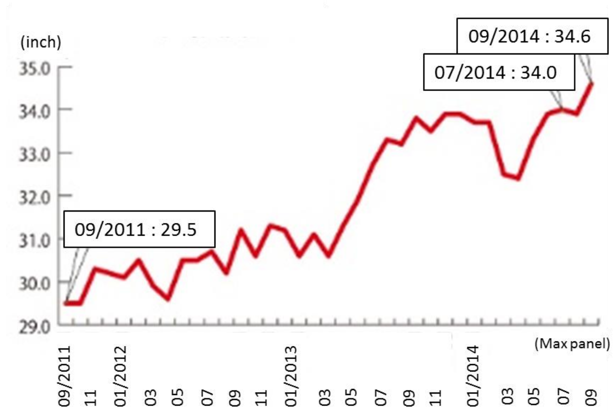


Figure 1. Average screen size of TV [4]

Head-mounted Display (HMD) is a device to enable us to enjoy big-screen 3D picture at various places. HMD is well known as a wearable display of the computers. Generally HMD is shaped as glasses or goggles. Observers can view a big screen in HMD at the same level as a movie theater. It is the characteristic of HMD that we can use it regardless of a place and time. Sony offered a HMD for home users on November 11, 2011 [5]. Afterward, a new version, HMZ-T3 was released on November 13, 2013. HMD came to be used as means to enjoy movies at home. However, its device of stereoscopic vision is the same as glasses type, and can offer only bi-ocular stereoscopic images. In addition, motion parallax is not offered even if a user moves because of original problem of HMD: the display moves with head motion of the user. To solve the problem, researchers are working on the technique to estimate the position of the user by head tracking such as the motion capture, and then switch the parallax image according to this position [6, 7, 8]. However, the technique needs installation of sensor and detector, which limits usage environment.

This paper proposes a method to provide 3D display with unlimited parallax and display virtual TV as if real TV is put by HMD. By preventing that the virtual TV follows head motion, the user becomes able to watch TV naturally. In addition, the motion parallaxes can be provided without restriction of view area.

II. PROPOSAL SYSTEM

A. Composition of System

The system is comprised of HMD and two Web cameras. Two Web cameras are attached on the front of HMD as our eyes. At first, an observer determines a characteristic point in the 3D image from the two cameras to set the virtual TV. The parallax images of the virtual TV are fitted in the points in the parallax images from the two cameras respectively. So, by observing the parallax images in HMD, the observer can view the three-dimensional virtual TV. Based on the relations of the position of an observer and the characteristic point, position and shape of the virtual TV are changed continually when the observer moves.

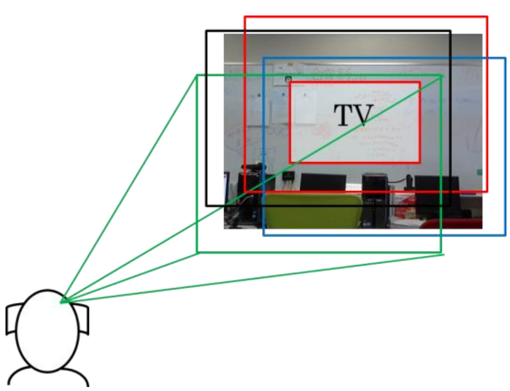


Figure 2. Viewing image of virtual TV

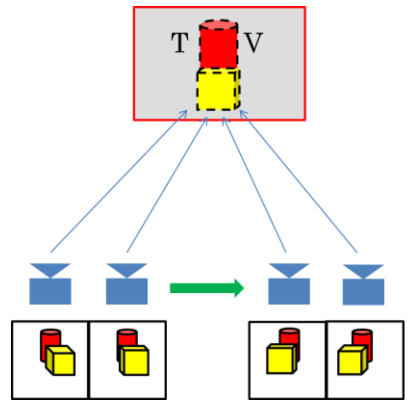


Figure 3. Viewing image of motion parallax

B. Tracking the Characteristic Point

The characteristic point is determined by distance and angle of line of sight like figure 4. The proposal method needs a tracking method that is stubborn for changes in environment. Therefore we adopted a method to use the phase-only-correlation. This method tracks the characteristic point by image matching and is not affected by the noise and the brightness in environment. At first, the cameras acquire the images of two consecutive frames as input images. Applying the phase-only-correlation to an input image, the system calculates how many pixel the characteristic point moved from a frame to the next frame. The position of the virtual TV is adjusted according to the motion range of characteristic point, so the virtual TV is displayed like a fixed TV.

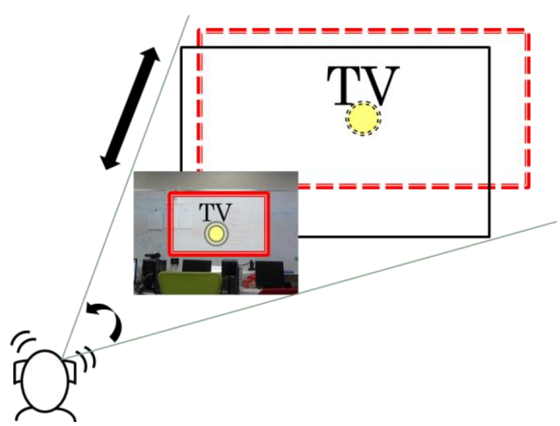


Figure 4. Characteristic point

C. Phase-Only-Correlation

The phase-only-correlation method is a method that reveals the correlation between two images using the phase information of each image [9].

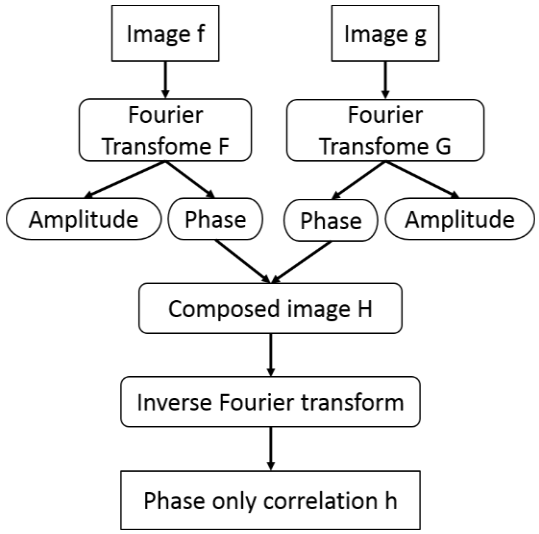


Figure 5. Flow chart of the Phase Only Correlation

The procedure of the processing is shown in figure 5. At first, performs Fourier transformation on two pieces of images. By this processing, the two images with only phase information are composed. Then normalizes the provided amplitude to leave only phase information and performs a convolution. Finally the phase-only-correlation can be searched for by performing reverse Fourier inverse transform of the image.

Two $X \times Y$ images, $f(x,y), g(x,y)$, where we assume that the index ranges are $x = 0,1,2,...,X-1$, $y = 0,1,2,...,Y-1$.

$F(u,v), G(u,v)$ donate the 2D Discrete Fourier Transforms of the two images. Two images given by

$$F(u,v) = \sum_{x=0}^{x-1}\sum_{y=0}^{y-1} f(x,y)e^{-j2\pi(xu/X+yv/Y)} = A(u,v)e^{j\theta(u,v)} \quad (1)$$

$$G(u,v) = \sum_{x=0}^{x-1}\sum_{y=0}^{y-1} g(x,y)e^{-j2\pi(xu/X+yv/Y)} = B(u,v)e^{j\phi(u,v)} \quad (2)$$

Where, $u = 0,1,2,...,X-1$, $v = 0,1,2,...,Y-1$. And $A(u,v)$, $B(u,v)$ denote amplitude spectrum, $e^{j\theta(u,v)}$ and $e^{j\phi(u,v)}$ denote phase spectrum. $F'(u,v), G'(u,v)$ which are the $A(u,v)$, $B(u,v)$ which is an amplitude spectrum were normalized to 1 given by

$$F'(u,v) = e^{j\theta(u,v)} \quad (3)$$

$$G'(u,v) = e^{j\phi(u,v)} \quad (4)$$

Figure 6. Sample of the phase limited image

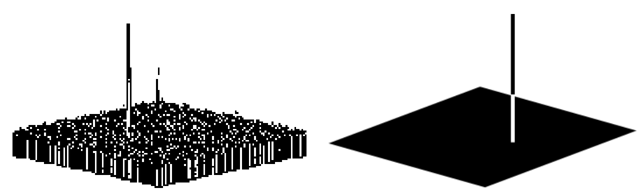
Figure 6 shows the sample of the phase limited image. Then, Synthetic image $H(u,v)$ is provided by multiplying complex conjugate of $F'(u,v)$ and $G'(u,v)$ by.

$$H(u,v) = F'(u,v)(F'(u,v))^* = e^{j(\theta-\phi)} \quad (5)$$

Inverse discrete Fourier converts $H(u,v)$, and phase only correlation $h(x,y)$ is calculated.

$$h(x,y) = \frac{1}{XY}\sum_{u=0}^{x-1}\sum_{v=0}^{y-1} H(u,v)e^{j2\pi(xu/X+yv/Y)} = \frac{1}{XY}\sum_{u=0}^{x-1}\sum_{v=0}^{y-1} (e^{j(\theta-\phi)})e^{j2\pi(xu/X+yv/Y)} \quad (6)$$

The peak of height of the correlation shows a resemblance degree of two pieces of images in figure 7(a). When the two pieces of images have completely same phase information, the value of the peak height becomes 1 in figure 7(b). In addition, the coordinate of the correlation peak is equivalent to quantity of movement of two pieces of images.



(a) (b)

Figure 7. The peak of height of the correlation

D. Distance Calculation

For giving an image with sense of perspective to the observer, the size and shape of the virtual TV screen is changed based on the observer's movement. For example, when an observer approaches the TV, the screen broadens. When an observer moves away from it, the screen becomes small. The distance from the screen to the observer can be calculated by triangulation.

In figure 8, the input images are the images at the same flame acquired from two cameras. L and R are decided as distances from right side camera and left side camera to the characteristic point. D is the real distance between two cameras, and F is the focus distance of the two cameras. By these position relations, the distance to the characteristic point is calculated. If this distance changes, the size of virtual TV screen is changed.

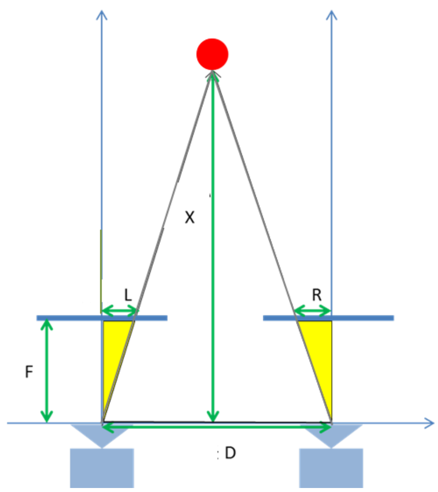


Figure 8. Distance from HMD to the point

E. Image Processing

Figure 9 shows the flowchart of the image processing. The point to fix the TV screen is set in the input images acquired from the two cameras. The point is tracked by performing the matching of the phase-only-correlation for anteroposterior consecutive frames. The screen size is decided based on the distance to the point that is for putting the virtual TV calculated by triangulation using the position of the two points in the two cameras image.

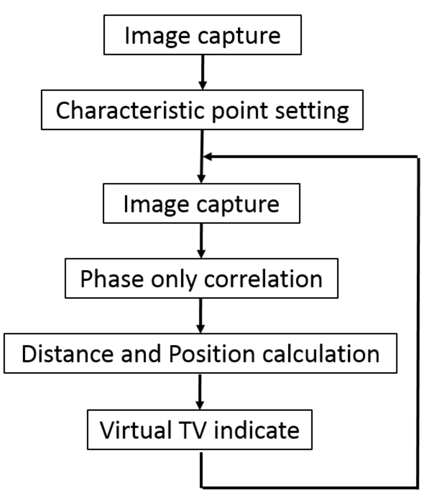


Figure 9. Image processing flowchart

III. EVALUATION EXPERIMENT

A. Experiment Method

Two experiments were conducted to evaluate the performance of the proposal system. Figure 10 is experimental conditions. The experiment 1 is an experiment to evaluate image showed by virtual TV by subjects. In the experiment, the subjects moving freely observed the virtual TV that is displayed by the HMD. At this time we gave the subjects two questions: "Was the Virtual TV fixed like a real TV?" and "Did it look like a virtual TV is set in the real world?". The subject evaluated these questions by 5-point method. Table 1 shows an evaluation criterion. The experiment was conducted with the brightness from fluorescent lamps after sunset. The subjects were 15 healthy students between the ages of 21 and 29. In the experiment 2, the subjects observed two virtual cubes to confirm motion parallax. Two cubes were displayed back and forth on a desk and were observed with stereo camera. HMD was moved in front of virtual cubes. The cubes was made using OpenGL that was API which 3DCG could draw fast. HMD and a stereo camera same as an experiment 1 were used.

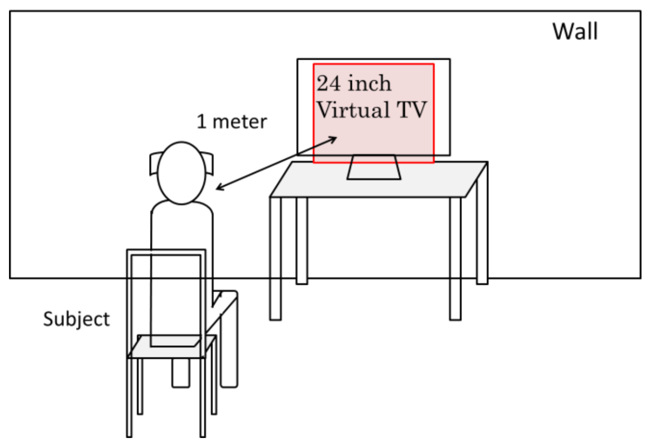


Figure 10. Experimental conditions

TABLE I. EVALUATION CRITERION

Question	Point	Evaluation criterion
"Was the Virtual TV fixed like a real TV?"	5	Extremely fixed like a real TV
	4	Slightly fixed like a real TV
	3	Neither
	2	Slightly not fixed like a real TV
	1	Extremely not fixed like a real TV
"Did it look like a virtual TV is set in the real world?"	5	Extremely set in the real world
	4	Slightly set in the real world
	3	Neither
	2	Slightly not set in the real world
	1	Extremely not set in the real world

B. Experimental Apparatus

This experiment was conducted with Oculus Rift type HMD and ovrvision. Figure 11 shows the experimental apparatus that consist of a HMD, a stereo camera and a personal computer. The specifications of personal computer are showed in table 2. The specifications of the HMD are showed in table 3 and the specifications of the stereo camera are showed in table 4.

TABLE II. SPECIFICATIONS OF COMPUTER

Model	Lenovo Windows7 PC
OS	Windows 7 Home Premium
CPU	Intel Core i7 870
RAM	6.00GB
GPU	AMD Radeon HD 7700 Series

TABLE III. SPECIFICATIONS OF HMD

Name	Oculus Rift
Resolution	1280*800
Horizontal angle	＞90°
Vertical angle	＞110°

TABLE IV. SPECIFICATIONS OF THE STEREOCAMERA

Name	Ovrvision
Resolution	640*480*2
Angle	Oculus Rift same
FPS	60fps

Figure 11. Experimental apparatus

C. *Experiment Result*

The figure 12 shows the image of experiment 1. The figure 13 and the figure 14 show the results of the experiment 1 and the result images of experiment 2 respectively. As showed in figure 12, we confirmed that the position of the virtual TV in the HMD changes as see a real TV by glasses when we move. In the experiment 1, all subjects gave a high evaluation of more than 3 points in the five-point evaluation method as shown in figure 13. As shown in figure 14, the angles of the cubes are different when the viewing point is different. It was confirmed that continuous motion parallax can be acquired.

Real TV Virtual TV

Figure 12. Result of virtual TV

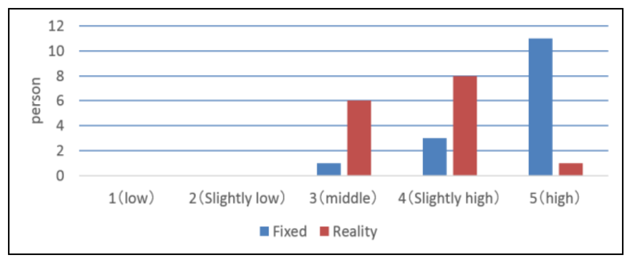
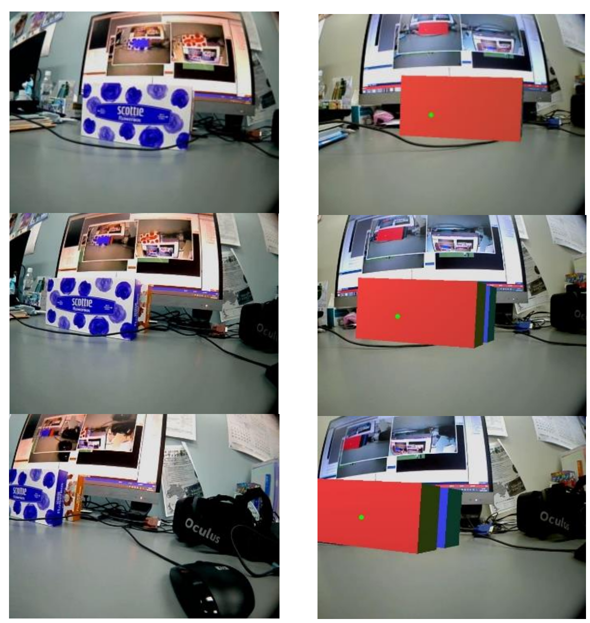


Figure 13. Result of subject experiment



Real cubes Virtual cubes

Figure 14. Result of motion parallax

IV. Conclusion

This paper proposed a method to provide 3D display with unlimited parallax and display virtual TV as if real TV is put by HMD. By preventing that a virtual TV in HMD follows head motion, the user is able to watch the virtual TV more naturally. In addition, the motion parallaxes can be offered without visual area restriction by the proposal method. In the proposal system, the TV screen in the HMD is set at the point calculated by triangulation using the two cameras. The point is tracked by the matching of anteroposterior consecutive frames by using the phase-only-correlation. So the virtual TV can be displayed at the point in the input images. Two experiments were conducted to evaluate the performance of the proposal system. The experimental results showed that the position of the virtual TV can be fixed to the real space like a real TV. When showing a 3D image, a more natural motion parallax can be gotten.

References

[1] OKUI Makoto, et al., “The Trend of Three Dimensional Image Technology”, The journal of the Institute of Image Information and Television Engineers Vol. 67, No. 1, pp. 46～52（2013）

[2] Kenichi Yamashita, et al., “Multiple-Face Tracking Technology for Glasses-Free 3D Displays”, Toshiba review Vol.68 No.2,2013

[3] Rieko Fukushima, Norihiro Nakamura,” Viewing area optimization technology for glasses-free 3D REGZA”, Toshiba review Vol.66 No.5,2011

[4] BUSINESS COMPUTER NEWS, “The mean screen size of the liquid crystal television updates a record high”, Vol.1551, 2014-10-20

[5] HMZ-T1 “Personal 3D Viewer”
http://www.sony.jp/hmd/products/HMZ-T1/

[6] Takafumi Serizawa, Yasuyuki Yanagida,” A Calibration-less Motion Tracker Installable on HMD”, No. 11−5 Preceedingsof the 2011JSME Conference on Roboticsarld Meohatronics

[7] Hiromu Miyashita, Masaki Hayashi, Kenichi Okada,” Determination of Visual Field with Motion and Gaze for Portable Display”, IPSJ SIG TechnicalReport, MBL, 2008(107), 17-24, 2008-10-30

[8] 「MREAL」 canon R&D topics,
http://web.canon.jp/technology/approach/special/mr.html

[9] Koji Kobayashi, et al, ” Principles of Phase Only Correlation and Its Applications”, ITE Technical Report, Vol.20, No.41, PP.1～6, MIP96-53, NIM96-75（July. 1996）