Monocular 3D Vision with Correct Depth Using Polypyrrole Film Actuator for Light-weight Display Unit

Koji Ohmori, Kunio Sakamoto, Shusaku Nomura,
Tetsuya Hirotomi, Kuninori Shiwaku and Masahito Hirakawa

Abstract— **The human vision system has visual functions for viewing 3D images with a correct depth. These functions are called accommodation, vergence and binocular stereopsis. Most 3D display system utilizes binocular stereopsis. The authors have developed a monocular 3D vision system with accommodation mechanism, which is useful function for perceiving depth. This vision unit needs an image shift optics for generating monocular parallax images. But conventional image shift mechanism is heavy because of its linear actuator system. To improve this problem, we developed a light-weight 3D vision unit for presenting monocular stereoscopic images using a polypyrrole linear actuator.**

Index Terms—**head mounted display, monocular stereoscopic display, real-time stereogram, 3-D display**

I. Introduction

A study of virtual-reality system has been popular and its technology has been applied to medical engineering, educational engineering, a CAD/CAM system and so on. The 3D imaging display system has two types in the presentation method; one is a 3D display system using a special glasses and the other is the monitor system requiring no special glasses. A display system requiring no special glasses is useful for a 3D TV monitor, but this system has demerit such that the size of a monitor restricts the visual field for displaying images. The 3D display system using special glasses can display virtual images over a wide area.

The human vision system has visual functions for viewing 3D images with a correct depth. These functions are called accommodation, vergence and binocular stereopsis. Accommodation is a useful function for perceiving a depth by the monocular vision system. Binocular stereopsis is one of the most important processes for the human 3D perception. Then most 3D display system utilizes this binocular stereopsis. The binocular vision mostly utilizes next two functions. One is simultaneous perception, which is an ability to perceive dichoptically presented images simultaneously and in the correct position. The other is binocular fusion, which is an ability to perceive two dichoptic images in left and right eyes as one image.

K. Ohmori and K. Sakamoto are with Konan University, 8-9-1 Okamoto, Higashinada, Kobe 658-8501, Japan (e-mail: kunio@konan-u.ac.jp).

S. Nomura is with Nagaoka University of Technology, 1603-1 Kamitomioka, Nagaoka 940-2188, Japan

T. Hirotomi, K. Shiwaku and M. Hirakawa are with Shimane University, 1060 Nishikawatsu-cho, Matsue, Shimane 690-8504, Japan

Meanwhile accommodation is a useful function for perceiving a depth by the monocular vision system. And humans have the function called the simultaneous perception. Using this function, humans perceive dichoptically presented images simultaneously and in the correct position. So the authors have developed a monocular 3D vision system which enables us to provide correct and natural 3D images with a depth perceiving by the accommodation.

II. Vision System

A. HMD System

A study of virtual-reality system has been popular and its technology has been applied to medical engineering, educational engineering, a CAD/CAM system and so on. The 3D imaging display system has two types in the presentation method; one is a 3D display system using a special glasses and the other is the monitor system requiring no special glasses. A display system requiring no special glasses is useful for a 3D TV monitor, but this system has demerit such that the size of a monitor restricts the visual field for displaying images. The 3D display system using a special glasses can display virtual images over a wide area.

In general, a 3D display system using a special glasses utilizes the binocular parallax, and many methods have been proposed. Recently an imaging device and a lens-optical system have been miniaturized and lightened of its weight. Then a 3D head mounted display (HMD), which is composed of these imaging and optical devices, has been researched and developed.

The available characteristics of this HMD system are as follows [1]:

1) The glasses of an HMD is constituted of light and small devices, so that the system can construct a virtual large screen without requiring an extensive space.
2) The system can display virtual images with the feeling of being at a real space because observers can watch the large screen right before their eyes.
3) The HMD system can display 3D images utilizing a binocular stereoscope with two monocular display systems.
4) The HMD system can display the combination of real and virtual images using a see-through mechanism.

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

B. Human vision system and perceiving depth

The human vision system has visual functions for viewing 3D images with a correct depth. These functions are called accommodation, vergence and binocular stereopsis. Accommodation is a useful function for perceiving a depth by the monocular vision system. Binocular stereopsis is one of the most important processes for the human 3D perception.

The binocular vision system has many functions as follows [2]:

1) Simultaneous perception: The ability to perceive dichoptically presented images simultaneously and in the correct position.
2) Binocular fusion: The ability to perceive two dichoptic images in left and right eyes as one image.
3) Coarse stereopsis: The ability to perceive dichoptically presented images with a parallax as one image with a coarse depth.
4) Fine stereopsis: The ability to perceive dichoptically presented images with a parallax as one image with a fine depth.
5) Crossed fusion: The ability to perceive dichoptically presented images with crossed disparities as one image.
6) Uncrossed fusion: The ability to perceive dichoptically presented images with uncrossed disparities as one image.
7) Dynamic stereopsis: The ability to perceive depth in moving random dot stereogram images.

C. Monocular stereoscopy

Accommodation is a useful function for perceiving a depth by the monocular vision system. And humans have the function called the simultaneous perception. Using this function, humans perceive dichoptically presented images simultaneously and in the correct position. So the 3D HMD system can display a spatial image with merit that lightness of real and virtual images is not lacking, when humans watch a real 3D image with the left eye and a virtual 3D image with the right eye, for example.

Now, the following is devised how to display monocular 3D images.

1. Electoro-holography

A hologram can display an ideal 3D image containing all information that the human vision system requires. So humans perceive a 3D image of the electoro-holography with just one eye. However, note that the size and weight of an electoro-holography can be reduced within technical limits.

2. Super multi-view stereo

This new stereo approach called Super Multi-View is proposed by members of the 3D project [3]. This super multi-view means the following condition; more than two parallax images pass through a pupil of eye. At the super multi-viewing, these parallax images passed through a pupil are focused the same position on a retina. So humans perceive a 3D image with a correct accommodation due to the monocular parallax.

III. Concept

3D stereo imaging, in which the left and right eyes see slightly different images, is a mature concept that is very popular, and the many techniques of presenting stereo images are proposed. These methods require that one eye not see the view presented to the other eye. In the 3D stereo image viewer, two separate images are presented and a physical barrier blocks the unwanted view. The authors have ever researched 3D imaging display systems using the polarized glasses and the liquid crystal shutter glasses, the image splitter such as a parallax barrier or a lenticular screen and the holographic optical elements [4][5][6]. But this paper describes another technology for displaying 3D image.

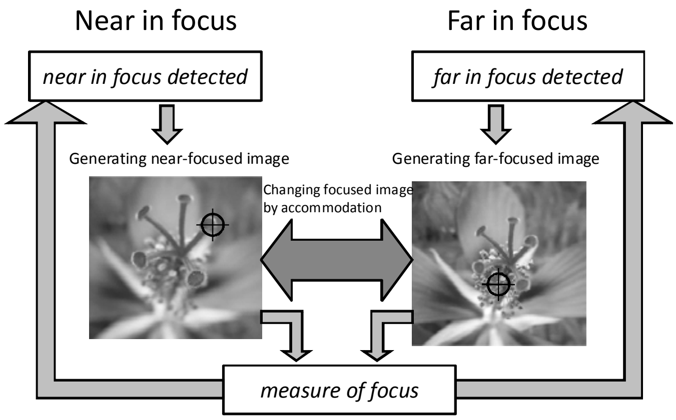


Fig. 1 Flow of generating focused image

To realize natural 3D viewing, we have developed the monocular vision system, which can directly project stereoscopic image on a retina, and a 3D image generation system, which can make 3D computer graphics in accordance with accommodation. Assume that an actual object is in the real world. When you perceive this object, a part of the projected image on a retina might be a blur by the lens of an eye. It is a physiological response called as accommodation. In case of virtual 3D image viewing as shown in Fig. 1, you can watch a correct 3D image as the actual object is in there if the projected retina image has appropriate blur in compliance with focus adjustment of your eye. Then you might perceive virtual images with same accommodation as you watch real objects. Thus a monocular 3D vision system can provide correct 3D viewing with accommodation, vergence and binocular stereopsis and without a tired feeling at long time watching when the retina image is directly projected and external stimulation induces the focus adjustment by changing the thickness of an eye lens.

IV. Display System

Fig. 2 shows the principle of a reconstruction of the 3D image by the stereo-viewing. Fig. 2 (a) shows the optical configuration of a conventional stereogram. To display a point object P, observers turn on a point P_L for the left eye and a point P_R for the right eye as shown in this figure. The observer perceives that a point object exists on a spatial position P due to the binocular parallax. Fig. 2 (b) shows the optical configuration of a monocular multi-view stereogram.

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

To display a point object P, observers turn on points P_{R1} and P_{R2} for the right eye as shown in this figure. At the monocular multi-viewing, the observer adjusts the focal length of an eye to match with the spatial position P, then projected images of pixels P_{R1} and P_{R2} are focused to the same position on the retina. So the observer naturally perceives that a point object exists on a spatial position P due to the monocular parallax.

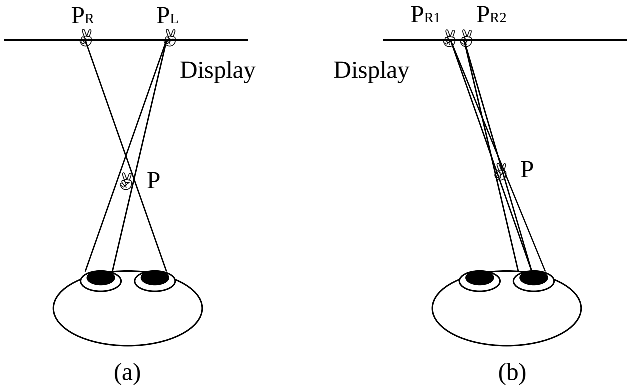


Fig. 2 The principle of a stereogram

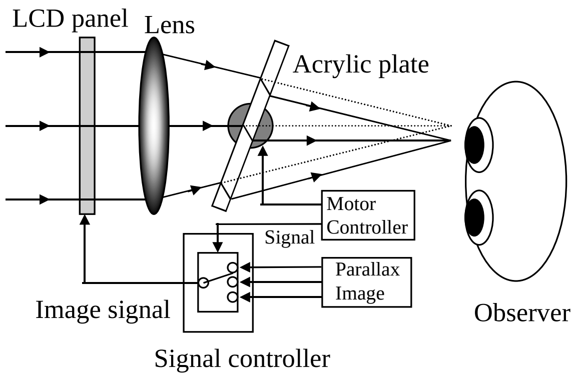


Fig. 3 Optical layout of the monocular 3D display

Fig. 3 shows the principle of the 3D vision system using monocular stereoscopy. This display system consists of an LCD panel, an acrylic plate and an optical lens. The observers perceive parallax images at the just point, which the optical lens converges the light on. To perceive multiple parallax images with just one eye, the image shifting optics consists of a parallel plane acrylic plate, whose inclination causes the image to shift as shown in Fig. 4. In Fig. 4, the image shift Δx and the angle of inclination θ have the relation as follows

$$\Delta x = d \cdot \sin\theta(1-\cos\theta/\sqrt{n^2-\sin^2\theta})$$

where d is the thickness of a glass plate and n is the refractive index of a glass. An LCD panel is used as the displaying plane of parallax images. The signal controller sends an image signal to the LCD panel to the tune of a control signal. Then the displaying plane of parallax images creates monocular multi-viewing images. Fig. 5 shows the appearance of the LCD panel and controller.

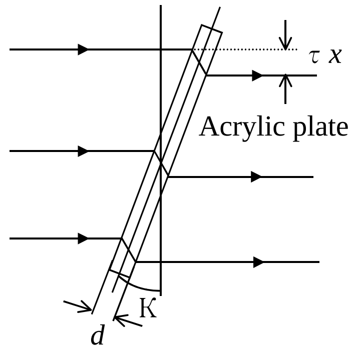


Fig. 4 Image shifting optics

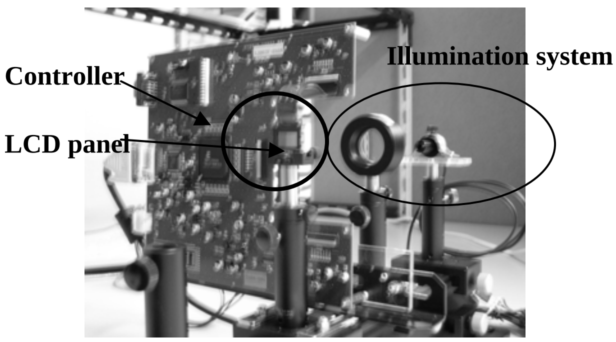


Fig. 5 Appearance of LCD panel and controller

V. Eye Position Spatial Measurement

Fig. 6 shows the result of detecting eyes from the captured face image. Using this result of detecting eyes, it is possible to measure the distance from the camera to the eye and to determine the direction using the camera image captured by the single camera.

Fig. 7 shows the illustration of capturing the face of a user using the pin-hole camera. The face in the real world and the perspective image of a user's face on the plane of a film are indicated in this figure. In Fig. 7, let (x, y, z) be three space coordinate and parameter z shows the depth. The S_X-S_Y coordinate system shows the image plane of a camera. Assume that the points A, B, C and D show the interest points of the eyes and the points A', B', C' and D' on the image plane show the corresponding points of the eyes. Moreover, the positions of eye E_R and E_L are the midpoints of the side AB and CD respectively. The followings are assumed at capturing:

1) The interval of eyes (the length of the side E_LE_R) is R.
2) AB : BC : CD = k : 1 : k .
3) $\overrightarrow{OA} = t_A\overrightarrow{OA'}, \overrightarrow{OB} = t_B\overrightarrow{OB'}, \overrightarrow{OC} = t_C\overrightarrow{OC'}$ (here, t_A, t_B and t_C are variables).

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Fig. 6 Detected eyes

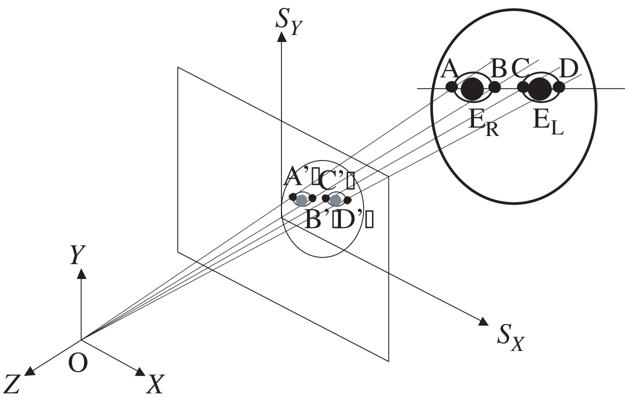


Fig. 7 Measurement of the 3D position

In Fig. 7, the relation of the vectors OA, OB and OC is given by

$$(k+1)\overrightarrow{OB} = \overrightarrow{OA} + k\overrightarrow{OC}.$$

Hence, the following equations are given.

$$t_B(k+1)\overrightarrow{OB'} = t_A\overrightarrow{OA'} + t_C k\overrightarrow{OC'} \quad (1)$$

$$\left|\overrightarrow{E_L E_R}\right| = \left|\overrightarrow{AC}\right| = R \quad (2)$$

Assume that A'B' : B'C' = k' : 1, the following equation is given.

$$(k'+1)\overrightarrow{OB'} = \overrightarrow{OA'} + k'\overrightarrow{OC'}. \quad (3)$$

Comparing the coefficients of equations (1) and (3), the parameters t_B and t_C are given by

$$t_B = \beta t_A, \ t_C = \gamma t_A.$$

Here, $\beta = \dfrac{k'+1}{k+1}, \gamma = \dfrac{k'}{k}$.

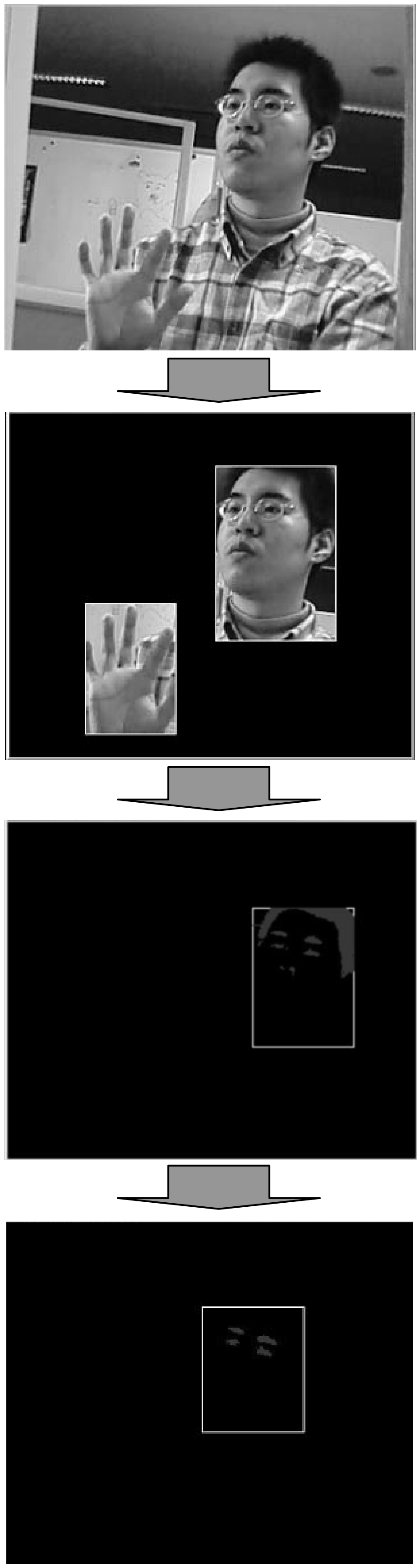


Fig. 8 Image processing for detecting observer's both eyes

Using the relation $\left|\overrightarrow{AC}\right| = R$, the following equation is given.

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$$\left|\overrightarrow{AC}\right|^2 = \left|\overrightarrow{OC} - \overrightarrow{OA}\right|^2$$

$$= {t_C}^2\left|\overrightarrow{OC'}\right|^2 - 2t_A t_C \overrightarrow{OA'} \cdot \overrightarrow{OC'} + {t_A}^2\left|\overrightarrow{OA'}\right|^2 = R^2$$

The z parameter of the point A is always minus, the parameter t_A is given by

$$t_A = \frac{R}{\sqrt{\gamma^2\left|\overrightarrow{OC'}\right|^2 - 2\gamma\overrightarrow{OA'} \cdot \overrightarrow{OC'} + \left|\overrightarrow{OA'}\right|^2}} (> 0)$$

Thus the parameters t_A, t_B and t_C are given and the spatial position of eyes is uniquely determined.

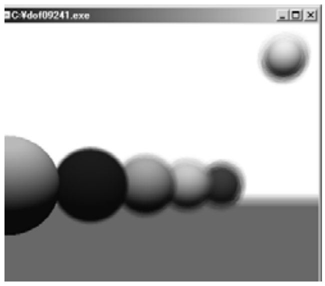
Consider the image in Fig. 8 below for demonstration of detecting eyes. The detecting process and image processing go along the step as follows. The image processing software firstly determines the regions of the observer's face and hands from the video frame image of the single camera using the skin color detection. At the skin detection, we utilized the image processing using the HSV space. Secondly it needs to choose only the face area from detected regions. This process is performed by the hair-color detection using the RGB space and the histogram approach. The part of hair is attached to the face area and thus it enables the system to choose the region of the face. And the parts of black colors are extracted from the selected face region.

Next, parts of the hair are eliminated. This operation is the image processing using the RGB space. Countering the numbers of connected pixels with black colors, the eyes and eyebrows are only extracted. Finally, it chooses the eyes from this image after analyzing the alignments of parts on the face and the intervals between the parts.

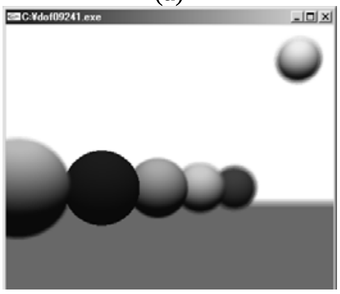
We incorporated this measuring method into the eye tracking system. In this system, the location of the human face is detected and its head pose is estimated continuously. To estimate the head pose, the image processing software recognized the positions of performer's eyes by detecting the regions of eyes from the extracted face area. By analyzing the detected human face region and finding the positions of eyes, we can estimate the head pose; the slant angle, *i.e.*, how he turns his head right or left.

VI. Real-time Generating Focused Image

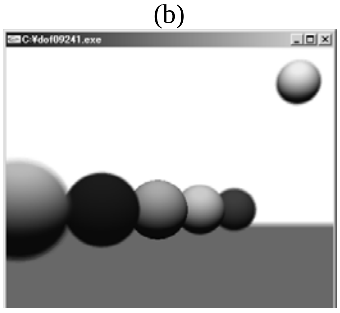
To induce accommodation, the display system needs to generate an appropriate focused image in accordance with a result of the focus point detection by a computer graphics. We used the OpenGL developed by Silicon Graphics Inc for real-time generating of 3D objects. Fig. 9 shows computer graphics images generated on the Windows XP using the OpenGL graphics library. In this figure, five balls are on the same plane and arrangements become farther and farther from left to right. At near in focus, an outline of the farthest ball is a blur (Fig. 9(a)). At far in focus the nearest ball is a blur (Fig. 9(e)). When focus at center ball, other balls are out of focus (Fig. 9(c)). Thus we can confirm that the OpenGL produce virtual 3D images in accordance with the focus depth by computer graphics.



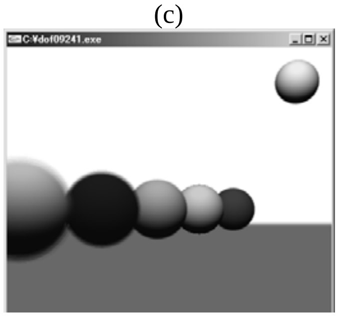
(a)



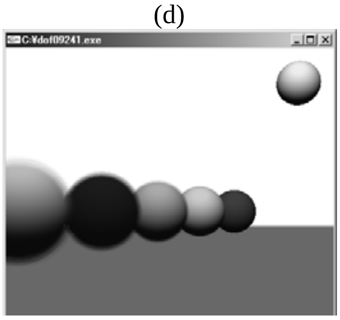
(b)



(c)



(d)



(e)

Fig. 9 Real-time generated focused images

VII. Improvement Of Image Shift Optics

As shown in Fig. 10, the image shift Δx is generated by an inclined acrylic plate. To incline the plate, former vision unit used a motor. Therefore its weight is heavy so that observers wear it as glasses. To improve this problem, the authors developed new image shift mechanism for monocular 3D vision. This image shifting optics consists of an acrylic plate and a polypyrrole linear actuator as shown in Fig. 10. Fig. 11

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

shows the structure of the polypyrrole actuator. This plastic film actuator is made of the polypyrrole. The thin layer of polypyrrole is attached on an acetate film as a support film. The polypyrrole can expand when it absorbs moisture in the air. When it discharges moisture, it can contract. Meanwhile the polypyrrole can also contract when the voltage is applied to it. The principle lies in the desorption of water vapor caused by Joule heating, where the electric field controls the sorption and the film absorbs and desorbs moisture reversibly in response to the applied voltage. The plastic film actuator is bent because the polypyrrole layer contacts but the acetate film can not contract as shown in Fig. 12.

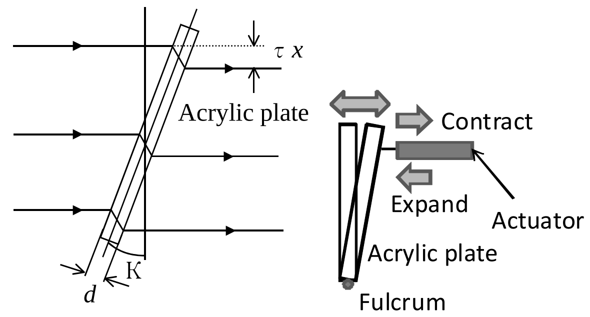


Fig. 10 Image shifting optics using linear actuator

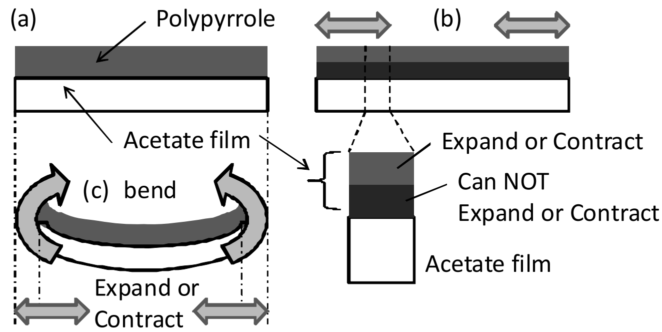


Fig. 11 Polypyrrole linear actuator

(a) off state (b) on state

Fig. 12 Appearance of polypyrrole linear actuator

The polypyrrole film containing perchlorate is electrochemically synthesized by the anodic oxidation of pyrrole. The plastic film actuator can incline the acrylic plate when this film actuator is bent (or contracts) by controlling the applied voltage as shown in Fig. 10. The film actuator is thin and light. So the authors could develop the light-weight monocular display unit for 3D viewing with a correct and natural depth using this polypyrrole actuator.

ACKNOWLEDGMENT

This research is partially supported by "Grant-in-Aid for Young Scientists(B)" #20700112 and "Scientific Research (C) (General)" #20500481 from Ministry of Education, Culture, Sports, Science and Technology Japan(MEXT) and also by a grant from the Hyogo Science and Technology Association. A part of this work is done while the author is partially supported by the Hirao Taro Foundation of the Konan University Association for Academic Research, Japan.

REFERENCES

[1] K. Fukai, "3D Head mounted Display (*in Japanese*)", Proc. of the 1998 ITE Annual Convention, pp. 462-465, (1998).

[2] M. Emoto et al., "Stereoscopic Test Materials – Vision Tests - (*in Japanese*)", Proc. of the 1998 ITE Annual Convention, pp. 218-219, (1998).

[3] Y. Kajiki et al., "Hologram-like Video Images by 45-view Stereoscopic ", SPIE Proc. Vol. **3012A-19**, (1997).

[4] K. Sakamoto, H. Nakayama, S. Taneji, "Field-lens Display: Headtracking enables 3D image viewing at any position", Advances in intelligent IT, Active Media Technology 2006, pp. 277-280, (2006).

[5] K. Sakamoto, M. Takaki, M. Nishida, "Parallax Barrier 3D Reflection Display Using Holographic Screen", Proc. of 12th International Display Workshops, pp.1769-1772, (2005).

[6] K. Sakamoto, R. Kimura, M. Takaki, "Elimination of Pseudoscopic Region of Parallax Barrier 3D Display", Proc. of 11th International Display Workshops, pp. 1497-1498, (2004).