A robotic vision system emulating fixational eye movements and retinal sampling

著者	Yotsumoto Takanori, Hayashida Yuki, Yasukawa Shinsuke
journal or publication title	Proceedings of International Conference on Artificial Life & Robotics (ICAROB2022)
volume	27
page range	519-523
year	2022-01-22
URL	http://hdl.handle.net/10228/00008971

doi: https://doi.org/10.5954/ICAROB.2022.OS6-4

A Robotic vision system emulating fixational eye movements and retinal sampling

Takanori Yotsumoto
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan

Yuki Hayashida
Dept. Inform. Eng., Grad. Eng., Mie University,
1577 Kurima-machiya, Tsu, Mie 514-8507, *Japan*

Shinsuke Yasukawa
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan

E-mail: yotsumoto.takanori573@mail.kyutech.jp, hayashida@ai.info.mie-u.ac.jp, s-yasukawa@brain.kyutech.ac.jp

Abstract

Recent studies on visual physiology have suggested that fixational eye movements contribute to the information processing function of the retina. In this study, we built a robotic vision system that emulates the characteristics of fixational eye movements and retinal sampling as an assistant tool for clarifying information processing through the collaboration of eye movements and neural circuits. The proposed robotic vision system consists of two galvanometers, a field-programmable gate array (FPGA) board, and a high-speed, high-resolution camera. We measured the frequency response of this robotic vision system, and the results showed that the system satisfied the requirements for emulating the frequency characteristics of biological eye movements. In addition, we generated command signals in the FPGA to emulate biological eye movements and confirmed the satisfactory operation of the system.

Keywords: Fixational eye movement, Robot vision, Galvano mirror, Bio-inspired vision system

1. Introduction

The visual nervous system contains a large number of neurons and synapses, and it is difficult to monitor each of them simultaneously with sufficient resolution in physiological experiments. Therefore, it is useful to develop a neural-circuit model that describes the characteristics of neurons and their interactions based on physiological knowledge, and to use a simulation approach to reproduce the activity of a group of neurons. In simulation experiments, we can systematically investigate the functions and properties of each element of the visual network by freely controlling the parameters that are difficult to control in physiological experiments.

Eye movements, body movements, and adaptations that occur at all levels of the visual system can be viewed as active actions of the biological visual system that constantly seeks to acquire new information. If the biological visual system is considered as a system that performs visual computation while interacting with the external world, it is essential to perform the biological vision simulation in real time.

A methodology that attempts to understand the intricate essence of a system that undergoes complex interactions with the external environment by constructing an artificial system with a similar structure as the hardware and verifying its operation is called a constructivist methodology. Since Mead et al. demonstrated an analog circuit that mimics the structure of the retina [1], research has been extensively conducted to reproduce the structure of retinal neurons and their network circuits as faithfully as possible on silicon. These studies are significant from the standpoint of understanding biological vision and engineering applications (reviewed in [2]). A real-time simulation of the visual system has also been realized [3].

Although there have been attempts at hardware emulation of the visual nervous system, there are few examples of hardware implementation of eye movements, which are the entry points for light input to the visual system, for simulation purposes. As mentioned earlier, if we consider the visual system as an active process, it is important to realize the hardware that can mimic eye movements using a bio-visual simulation method. In addition, findings from visual physiology and visual psychology experiments suggest that fixational eye movement contributes to the realization of retinal neural circuit functions, such as high spatial resolution [4], foreground–background separation [5], and target prediction [6].

In this study, we developed a robotic vision system that can receive visual input considering eye movements, for simulation of biological vision and evaluated the frequency characteristics of the visual input during fidgety microtremors in the robotic vision system.

2. Fixational eye movements and retinal sampling

Human fixation eye movement is composed of three components: tremor, drift, and microsaccade.

Tremor is an eye movement of small amplitude occurring at a frequency of 90 Hz or less, whereas microsaccade is an eye movement with a velocity of approximately 120 °/s occurring at a frequency of 2.0 Hz or less. Drift is a low-speed eye movement that occurs simultaneously with tremor and in between microsaccades [7]. To reproduce these movements using an engineering approach, it is necessary for a robot vision system to have a cutoff frequency of 90 Hz or higher and to move at a speed of 120 °/s or higher.

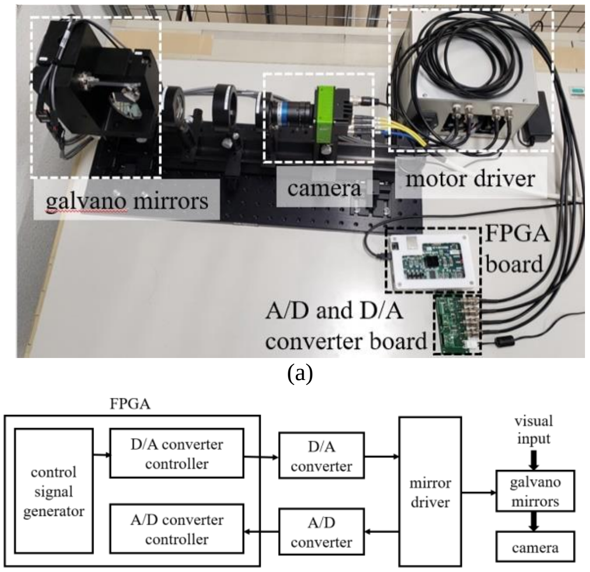
The spatio-temporal characteristics of retinal sampling are used as a reference for selecting a camera module. The density of cone cells in the photoreceptor layer of the central fossa of the retina is approximately 150000 cells/mm^2 [8] and the relationship between the angle and length of the retina is approximately 300 μm/° [9]. Therefore, the spatial-resolution requirement of the robotic vision system is estimated to be 116 pixels/°. In this study, we assumed that one pixel of the image sensor corresponds to one cell of the retinal photoreceptor layer. Next, the temporal characteristics of retinal ganglion cells are considered to determine the frame rate of the camera module. Because the incidence of spike response of ganglion cells in the retina is once every 5 ms (200 Hz) on average, the frame rate requirement of the robotic vision system is 200 Hz or higher.

3. System integration

3.1. *Hardware*

Fig. 1(a) and Fig.1 (b) show the hardware setup and block diagram of the proposed system, respectively. The proposed robotic vision system consists of an optical pan-tilt mechanism with two rotating mirrors, its driver module, digital circuits for control signal generation, A/D and D/A converters, and a high-speed, high-resolution camera.

In general, in a pan/tilt camera, the line of sight is mechanically controlled using a rotary actuator. However, the mechanical control of gaze direction using a rotary actuator does not provide sufficient high-speed performance. In this study, we optically controlled the line-of-sight direction at high speed using a rotating mirror with two axes (pan and tilt), based on the study presented in [10]. In [10], a pupil-transmission system was installed between the camera and the mirror to resolve the trade-off between light intensity and angle of view. Signals for pan-tilt control are generated by digital circuits in the field-programmable gate array (FPGA); these signals are then D/A converted and input to the driver module for the motor driving the mirror. Signals representing the rotation angle of the mirror are output from the encoders installed on each axis and input to the FPGA via A/D converters. The mirror speed



(a)

(b)

Fig. 1 (a) Hardware setup of the proposed system. (b) Block diagram of the proposed system.

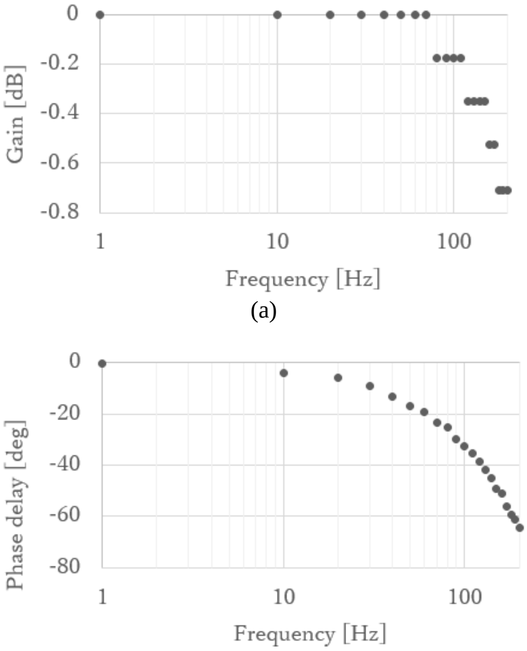
of the system, operating frequency, operating range, and lens diagonal angle of view of the pupil transfer system are set to 280 °/s, 200 Hz, ±20 °, and 38.6°, respectively.

Fig. 2 depicts the measured frequency response of the optical pan-tilt system. Sinusoidal signals of various frequencies with an amplitude of 0.5 V (±1.26°) are input to the optical pan-tilt system, and the angle signal of the mirror is measured by the encoder. The gain and phase characteristics of the optical pan-tilt system are measured on the basis of the relationship between the input and output signals. The cutoff frequency is higher than the target value of 200 Hz, which meets the requirements of the system.

The camera module used is a JAI SP-25000C-CXP4A, which has a resolution of 5120 × 3832 pixels when operated at a resolution of 200 fps. The spatial resolution of this camera is 140 pixel/°, which satisfies the requirement (116 pixel/°) for the robotic vision systems described in Section 2.

3.2. *Control signals for drift motion*

In this study, we focused on the drift motion even during fixational eye movements. Considering the temporal characteristics of the drift motion, we Implemented a digital circuit to generate control signals for the optical pan-tilt system based on the study by Hasegawa et al. [11]. To reproduce the motion, the displacement vector from the gazing center position was used as the control signal for the mirror control motor.

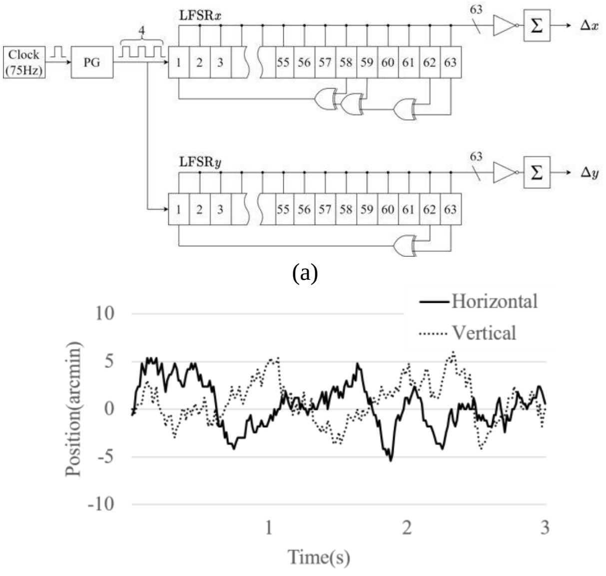


(a)

(b)

Fig. 2 Frequency response of optical pan-tilt system. (a) Gain characteristics. (b) Phase characteristics.

Fig. 3(a) shows a schematic of the circuit configuration used to generate the displacement vector. The circuit consists of a pulse generator (PG), a pair of 63-bit linear feedback shift registers (LFSRx and LFSRy), and a pair of summing modules (Σ). The pulse generator sends a sequence of four pulses as a clock signal to the linear feedback shift registers; this shifts the information in the registers by four bits. LFSRx and LFSRy generate M-sequences that are independent of each other. The summing module outputs the number of bit 0's in each connected linear-feedback shift register. The displacement vectors Δx and Δy are obtained by rounding the number of bits to integer values. Figure 3(b) shows a typical eye position transition pattern, where one pixel corresponds to the size of a cone cell in the central part of the human retina (visual angle of 0.6 arcmin). The width and velocity distribution of the fixational eye movement are reasonable [12].



(a)

(b)

Fig. 3 (a) Digital calculation circuit for displacement vector of fixational eye movements. (b) Variation of the displacement vector generated by the digital circuitry.

4. Experiments and Results

Experiments and analyses were performed using the following procedure: First, the two images shown in Fig.4 were input to the robotic vision system in two cases viz. drifting motion of the mirror and stationary motion of the mirror. We applied a three-dimensional Fourier transform to each of the obtained movies to transform them into a spatio-temporal frequency space. Next, at each time frequency, the data was two-dimensionalized by averaging over each of the same spatial frequencies (center images of Fig. 4). Finally, by integrating in the time direction at each spatial frequency, a graph was created to represent the spatial frequency characteristics of the system.

The results demonstrate that the spatial frequency response of the visual input with the mirror stationary is that of a natural image, with the amplitude spectrum decaying to 1/f. However, the low spatial frequency response of the visual input with the mirror in motion was flatter than that with the mirror stationary. This is similar to the effect of eye movement on the visual input of a biological vision system.

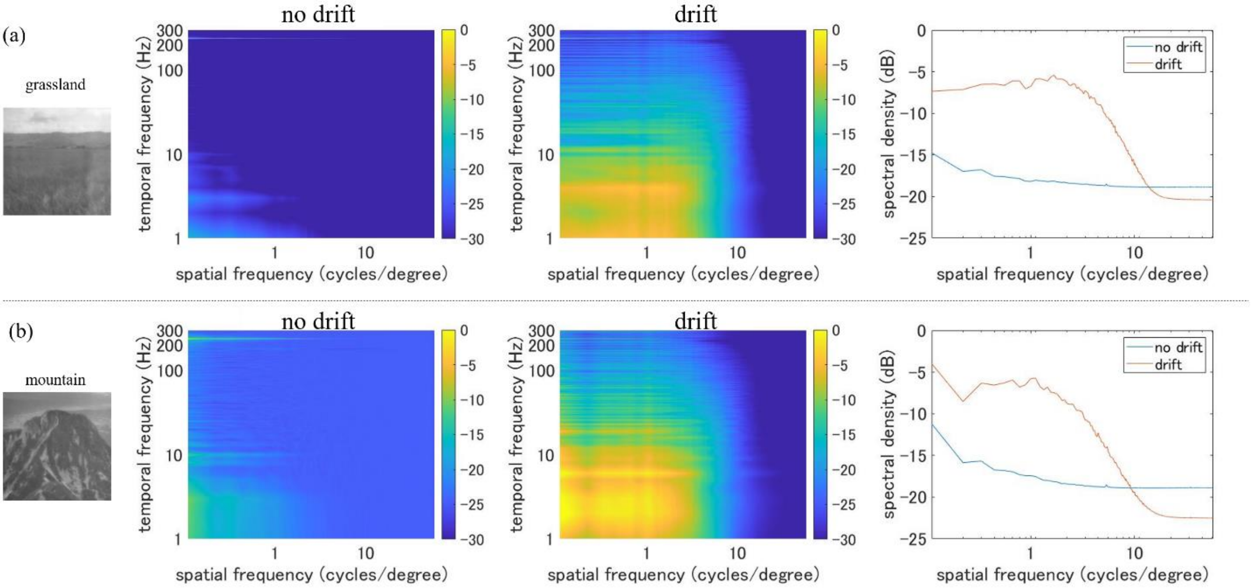


Fig.4 Spatio-temporal frequency characteristics of visual stimuli when the system is presented with images of (a) a meadow and (b) mountain, as input images. The images in the center left and center right show the spatio-temporal frequency characteristics when the mirror is stationary and when the mirror is moving, respectively. The graphs on the right show the spatial frequency response when the amplitude component is integrated in the time frequency direction.

5. Conclusion

We developed a robotic vision system that simulates the characteristics of human fixational eye movements and retinal sampling. To achieve fast response characteristics, such as those of fixational eye movements, we integrated an optical pan-tilt system using galvanometer mirrors and a digital circuit that generates control signals simulating eye movements and emulated dynamic characteristics similar to the drift component of actual human eye movements. In addition, we measured the spatio-temporal characteristics of the retinal input with and without fixational eye movement using the proposed system. From the results, we confirmed that the proposed system can simulate the spatio-temporal frequency characteristics of the visual input during human fixational eye movements. This system is expected to be a useful tool for understanding the relationship between eye movements and visual information processing in the biological visual system.

Acknowledgements

This research was partly supported by JSPS KAKENHI Grant Number 20K19895.

References

1. C. Mead, “Neuromorphic electronic systems.” Proceedings of the IEEE, Vol. 78, No. 10, pp. 1629-1636, 1990.
2. S. C. Liu, et al.(Eds.), “Event-based Neuromorphic Systems.” John Wiley & Sons, 2014.
3. Y. Hayashida, et al., ”Retinal Circuit Emulator With Spatiotemporal Spike Outputs at Millisecond Resolution in Response to Visual Events.”, IEEE transactions on biomedical circuits and systems, Vol. 11, No. 3, pp. 597-611, 2017.
4. M. Rucci, & J. D. Victor, “The unsteady eye: an information-processing stage, not a bug.” Trends in Neurosciences, Vol. 38, No. 4, pp. 195-206, 2015.
5. S. A. Baccus, et al. “A retinal circuit that computes object motion.” Journal of Neuroscience, Vol. 28, No. 27, pp. 6807-6817, 2008.
6. A. Matsumoto, M. Tachibana, “Rapid and coordinated processing of global motion images by local clusters of retinal ganglion cells.” Proceedings of the Japan Academy, Series B, Vol. 93, No. 7, pp. 524-524, 2017.
7. S. Martinez-Conde, S. L. Macknik, & D. H. Hubel, “The role of fixational eye movements in visual perception.” Nature reviews neuroscience, Vol. 5, No. 3, pp. 229-240, 2004.
8. R. W. Rodieck, “The First Steps in Seeing.” Sunderland, MA: Sinauer Associates, Vol. 1, 1998.
9. N. Drasdo, & C. W. Fowler, “Non-linear projection of the retinal image in a wide-angle schematic eye.” The British Journal of Ophthalmology, Vol. 58, No. 8, pp. 709-714, 1974.
10. K. Okumura, H. Oku, & M. Ishikawa, “High-speed gaze controller for millisecond-order pan/tilt camera.” In 2011 IEEE International Conference on Robotics and Automation, pp. 6186-6191, May. 2011.
11. J. Hasegawa, & T. Yagi, “Emulation of retinal cell responses during fixational eye movements.” IEICE Electronics Express, Vol. 7, No. 3, pp. 184-189, 2010.
12. I. Murakami, “Correlations between fixation stability and visual motion sensitivity.” Vision Research, Vo. 44, No. 8, pp.751-761, 2004.
13. X. Kuang, et al., “Temporal encoding of spatial information during active visual fixation.” Current Biology, Vol. 22, No.6, pp. 510-514, 2012.

Authors Introduction

Mr. Takanori Yotsumoto

He received his B.S. degree from the Department of Biorobotics, Kyushu Sangyo University, Japan in 2020. He is currently a Master’s course student in Kyushu Institute of Technology, Japan.

Dr. Yuki Hayashida

He received his Ph.D. in information engineering from Kyushu Institute of Technology, Kitakyushu, Japan, in 1998. He is currently a Professor with the Graduate School of Engineering, Mie University, Japan. His research interests include physiological and engineering studies on biological nervous and cardiac systems, neural prostheses, and nanomaterial-based neural electrodes.

Dr. ShinsukeYasukawa

He received his Ph.D. degree in electrical, electronic, and information engineering from Osaka University, Japan, in January 2017. Currently, he is an Associate Professor in the Department of Human Intelligence Systems, Kyushu Institute of Technology, Japan. His research interests include information processing in biological systems and their applications in robotics.