



Exploring Gaze Movement Gesture Recognition Method for Eye-Based Interaction Using Eyewear with Infrared Distance Sensor Array

Futami, Kyosuke

Tabuchi, Yuki

Murao, Kazuya

Terada, Tsutomu

(Citation)

Electronics, 11(10):1637

(Issue Date)

2022-05

(Resource Type)

journal article

(Version)

Version of Record

(Rights)

© 2022 by the authors. Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

(URL)

<https://hdl.handle.net/20.500.14094/0100482014>



Article

Exploring Gaze Movement Gesture Recognition Method for Eye-Based Interaction Using Eyewear with Infrared Distance Sensor Array [†]

Kyosuke Futami ^{1,2,*} , Yuki Tabuchi ¹, Kazuya Murao ^{1,3}  and Tsutomu Terada ⁴ 

¹ Graduate School of Information Science and Engineering, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu 525-8577, Shiga, Japan; is0417hs@ed.ritsumei.ac.jp (Y.T.); murao@cs.ritsumei.ac.jp (K.M.)

² Digital Spirits Teck, Kusatsu 525-8577, Shiga, Japan

³ Strategic Creation Research Promotion Project (PRESTO), Japan Science and Technology Agency (JST), 4-1-8 Honmachi, Kawaguchi 332-0012, Saitama, Japan

⁴ Graduate School of Engineering, Kobe University, 1-1 Rokkodaicho, Nada, Kobe 657-8501, Hyogo, Japan; tsutomu@eedept.kobe-u.ac.jp

* Correspondence: futami@fc.ritsumei.ac.jp

[†] This paper is an extended version of our paper published in (1) Futami, K. A Method to Recognize Eye Movements Based on Uplift Movement of Skin. In Proceedings of the Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers; London, United Kingdom, 9–13 September 2019; pp. 624–627. (2) Futami, K.; Tabuchi, Y.; Murao, K.; Terada, T. A Method to Recognize Eyeball Movement Gesture Using Infrared Distance Sensor Array on Eyewear. In Proceedings of the 23rd International Conference on Information Integration and Web Intelligence, Linz, Austria, 29 November–1 December 2021; pp. 645–649.



Citation: Futami, K.; Tabuchi, Y.; Murao, K.; Terada, T. Exploring Gaze Movement Gesture Recognition Method for Eye-Based Interaction Using Eyewear with Infrared Distance Sensor Array. *Electronics* **2022**, *11*, 1637. <https://doi.org/10.3390/electronics11101637>

Academic Editor: Hsiang-Chieh Chen

Received: 21 April 2022

Accepted: 14 May 2022

Published: 20 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: With the spread of eyewear devices, people are increasingly using information devices in various everyday situations. In these situations, it is important for eyewear devices to have eye-based interaction functions for simple hands-free input at a low cost. This paper proposes a gaze movement recognition method for simple hands-free interaction that uses eyewear equipped with an infrared distance sensor. The proposed method measures eyelid skin movement using an infrared distance sensor inside the eyewear and applies machine learning to the time-series sensor data to recognize gaze movements (e.g., up, down, left, and right). We implemented a prototype system and conducted evaluations with gaze movements including factors such as movement directions at 45-degree intervals and the movement distance difference in the same direction. The results showed the feasibility of the proposed method. The proposed method recognized 5 to 20 types of gaze movements with an F-value of 0.96 to 1.0. In addition, the proposed method was available with a limited number of sensors, such as two or three, and robust against disturbance in some usage conditions (e.g., body vibration, facial expression change). This paper provides helpful findings for the design of gaze movement recognition methods for simple hands-free interaction using eyewear devices at a low cost.

Keywords: glasses; eyewear; wearable device; infrared distance sensor; photo reflector; skin movement; eye-based interaction; gaze movement; hands-free input

1. Introduction

In recent years, eyewear-type information devices have become popular with information devices able to be used anytime, anywhere. For example, there are now smart glasses/optical see-through head-mounted displays (e.g., Epson MOVERIO, google glass), AR/XR Glasses (e.g., HoloLens, Nreal Air), and smart audio glasses (e.g., Bose Frames). It has become essential to provide simple hands-free input methods because these devices are used in various situations in which it is difficult to use the hands for input operations.

In connection with this, the gaze movement input function can be one of the helpful options for hands-free input using eyewear devices. This is because gaze movement directly reflects the user's intention and has been used as a hands-free gesture input method [1–5]. In addition, it is easy to apply gaze movement recognition sensors to eyewear devices. Furthermore, gaze movement input is useful for situations where other promising input methods are not suitable for a variety of reasons, including psychological pressure in terms of privacy and embarrassment when inputting (e.g., voice input [6,7], facial expression input [8]), physical fatigue (e.g., hand input in the air [9]), and requirements to carry and wear additional equipment (e.g., ear accessory devices [10], finger ring devices [11]).

Therefore, a gaze movement recognition method suitable for hands-free input on an eyewear device in continuous daily use is required. Some methods are promising. For example, the method of mounting a small camera on eyewear can recognize eye movements with high accuracy; however, the methods are not suitable for continuous daily wear for the general public. This is because the price of such consumer products is high [12,13] and power consumption and processor performance are high due to real-time processing of the image data [14–16]. Although methods of mounting the EOG sensor on the nasal bridge of eyewear are currently the least expensive, these methods have issues in terms of the recognition accuracy [16,17]. There is still a need to explore methods which can enable continuous daily use at a low cost.

One existing method uses an infrared distance sensor on eyewear to constantly sense eyelid skin movements. For example, this method is used for blink recognition (e.g., Google Glass [18], Dual Blink [19]) and facial expression recognition [20,21]. Infrared distance sensors are low-cost and sufficiently small to be installed in various eyewear devices (e.g., vision correction glasses and AR/VR glasses). In addition, previous studies have demonstrated that infrared distance sensors can effectively recognize skin movements and require low processor and power consumption for data processing [11,19,22]. Since it has been found that gaze movements may appear through the skin of the eyelids (e.g., gaze vibrations in sleep appear through the eyelids [23]), it is assumed that gaze movements can be recognized using this method. Although the recognition accuracy of this approach is assumed to be less than that of methods that use camera images, it is important to know the recognition accuracy of this approach and what applications it can be used for. If this approach is effective, this approach could help add simple gaze movement interaction functions to many eyewear devices for daily use at a low cost.

Therefore, we propose a gaze movement recognition method for simple hands-free interaction that uses eyewear equipped with an infrared distance sensor. The proposed method measures eyelid skin movement using an infrared distance sensor inside the eyewear, applies machine learning to the time-series sensor data, and recognizes gaze movement (e.g., up, down, left, and right). We implemented a prototype system of the proposed method. We conducted three types of evaluations. In Evaluation 1, we evaluated 20 types of gaze movement with 14 subjects to verify the feasibility of the proposed method as a simple hands-free input interface. In Evaluation 2, we investigated the necessary sensor positions when the number of sensors is small, such as two or three. The results of this evaluation were intended to help improve the proposed method so that it could be used with fewer sensors. In Evaluation 3, we evaluated whether the recognition accuracy of the proposed method changed under certain conditions, such as reattachment of the device, body vibration, and facial expression change. The results of this evaluation were important to aid understanding of the robustness of the proposed method for various situations typical of the use of the wearable devices.

We published the concept of the proposed method in work-in-progress papers in 2019 [24] and 2021 [25]. This paper improves Evaluation 1 and adds Evaluation 2 and Evaluation 3. Although Evaluation 1 was described in a previous paper, this paper improved Evaluation 1 by increasing the number of subjects from 5 to 14 and by adding an analysis of the recognition accuracy for each gesture. This paper also adds Section 2.2, Section 2.3 and Section 7.

2. Related Work

2.1. Eye Activity Sensing Technology Using Eyewear

In recent years, various methods have been proposed to sense eye activity using eyewear in different situations.

One existing method uses a camera attached to eyewear to recognize eye activity by processing the acquired images of the eye. Representative consumer devices include the Tobii Pro Glasses [12] (Tobii), Pupil Core [13] (Pupil Labs), and the SMI eye-tracking glasses (SMI). These methods can perform highly accurate recognition; however, the accuracy of such devices can be reduced by various factors, such as lighting conditions [14,15]. There are also issues related to cost and hardware. The data are acquired high frequently; thus, processor performance and power consumption to process the camera data are high. Since high processor processing power is required for data processing, a computer must process the data after data collection; thus, it is difficult to realize real-time applications, e.g., live notifications. Therefore, this method is mostly used for expert purposes, such as research, and is not used regularly by general consumers. Although a camera is not used, a method exists for an infrared corneal limbus tracker that uses a light source (infrared LED) and an optical sensor (phototransistor) to recognize the contour of the cornea. This method is also used with eyewear to sense gaze [26].

Another method uses an electrooculogram (EOG) sensor mounted on eyewear. This method senses the change in electric potential when the eye moves by placing electrodes on the skin near the eye, which is effective because the human eye is an electric dipole. Such EOG methods are used to estimate fatigue by sensing eyelid and eye movements [27,28]. Recently, some studies have attempted to reduce the size and number of electrodes assuming constant use in daily life [29]. Although it does not involve eyewear, an existing method recognizes the gaze movement direction using an EOG sensor attached to other devices, such as earphones and headphones [2,30]. In addition, the JINS MEME [31] is a consumer device that uses an EOG sensor on the nasal bridge and can be used at all times [32,33]. However, this method is not good at recognizing vertical gaze movement and distinguishing gaze and eyelid movements, and there is a limit to its ability to recognize eye activity with high resolution [16,17].

Another method uses an infrared distance sensor on eyewear, similar to our study. This method recognizes eye activity using an infrared distance sensor installed in front of the eye to recognize blinks. In addition, Google Glass has a function to recognize winks (blinks) that the user exaggerates as a gesture input [34]. The Dual Blink application [19] has shown that this method can recognize blinks and that it is suitable for constant use from various perspectives, including power consumption. Dual Blink can also induce blinks by providing a stimulus, e.g., blowing air onto the cornea. Futami [24] showed examples of the use of three sensors to recognize four types of gaze movements (up, down, left, right) and the gazing point of nine segments of the field of view. Masai et al. [35] also provided examples of using 16 sensors to recognize 7 types of gestures, including blink and gaze movement with movement directions at 90° intervals, and the gazing point of 25 segments of the field of view. In addition, this method has been used to recognize facial expressions by sensing eyelid and cheek movements [21,22,36]. Based on these previous studies, it is expected that eyewear equipped with infrared distance sensors will increase in the future.

2.2. Recognition Method of Skin Movement Using Infrared Distance Sensor

An infrared distance sensor attached to the wearer has been used to recognize the skin's movement. For example, many studies have applied infrared distance sensors to eyeglass-type wearing as outlined below.

Fukumoto et al. [21] proposed a method that recognizes a smile by sensing the movements around the cheeks and outer corners of the eyes. Masai et al. [22] proposed a method that recognizes facial expressions, such as smiles and expressions of surprise, in daily life by sensing the movements of the eyelids and cheeks. Masai et al. [37] proposed a method that recognizes gesture input of rubbing the cheeks with the hands by sensing

the movements of the cheeks. Regarding blink detection, Google Glass [18] recognizes intentional blink gestures and Dual Blink [19] recognizes natural blinks. Dual Blink [19] also has functions to physically induce user's blinks by hitting the eyelids with an air cannon. Futami et al. [24,25] and Masai et al. [35] proposed a method that recognizes gaze movements. In addition, some studies have shown that an infrared distance sensor is suitable for wearable devices that can be continuously used from multiple perspectives (e.g., power consumption) [19,22]. For improving the VR experience, an infrared distance sensor is used in a head-mounted display (HMD), with mapping of the movement of the skin of the user's face to the facial expression of an avatar in the virtual space [36]. An infrared distance sensor is used for the earphone to recognize input gestures such as methods of pulling the ear [38] and moving the tongue [39]. Other examples include ear accessory devices to recognize facial gestures [10], a wristband to recognize hand-shape gestures [40], a ring to recognize finger movement gestures [11], and a mouthpiece to recognize tongue gestures [41].

These studies have shown that an infrared distance sensor can recognize skin movements with high accuracy and robustness. In addition, infrared distance sensors are inexpensive, lightweight, compact, have low power consumption, and use little data. This reduces the cost and size of the battery and processor, making them suitable for wearable devices that are continuously used. Based on previous studies, an infrared distance sensor is considered to be suitable for the proposed method.

2.3. Simple Hands-Free Input Method

This paper investigates the feasibility of simple hands-free input for the proposed method. A hands-free input method is useful in situations and for people where a general input interface is not available, such as for people with a disability or in situations where both hands cannot be used for input. Previous studies have shown that various hands-free input methods expand the usage scenarios of information devices and make the use of information devices comfortable. Similar to these previous studies, our method is predicted to expand the usage scenarios of information devices and make the use of information devices comfortable.

Many methods for hands-free input have been proposed. Recognition methods of face or body movements (e.g., face [8,10], ear [38], tongue [39], finger [11]) are often used for simple hands-free input. Postures of the wrist [42] and torso [43] are also used for navigation input. The input method of speech recognition with a microphone is used for applications such as text input [6] and navigation [7]. One of the disadvantages of this method is that the recognition accuracy of the method decreases in situations where there is significant environmental sound [44]. Gaze movement is also used as an input method [2,3,24,25,35] because gaze can reflect the user's intention. There is also a method that uses movement of the head, for purposes such as turning pages when browsing [45] and operating the cursor (e.g., desktop devices [46], mobile devices [47]). There are methods that use a combination of head movement and gaze [48] and a combination of brain and gaze [49].

3. Method

The proposed method recognizes gaze movements (e.g., up, down, left, and right) based on the movement of the skin around the eyes that accompanies gaze movements. A flowchart of the proposed method is shown in Figure 1. The proposed method comprises three main steps: (1) First, for the sensing mechanism, multiple infrared distance sensors measure the skin movements of the eyelids. This skin movement is sensed based on the change in distance from the skin and the infrared distance sensor installed in front of the eyes (i.e., the inner circumference of the glasses). The infrared distance sensor uses infrared light to measure distance. (2) Second, the feature amounts of time series data are extracted. (3) Finally, in the recognition step, machine learning is applied to the data to recognize gaze movement.

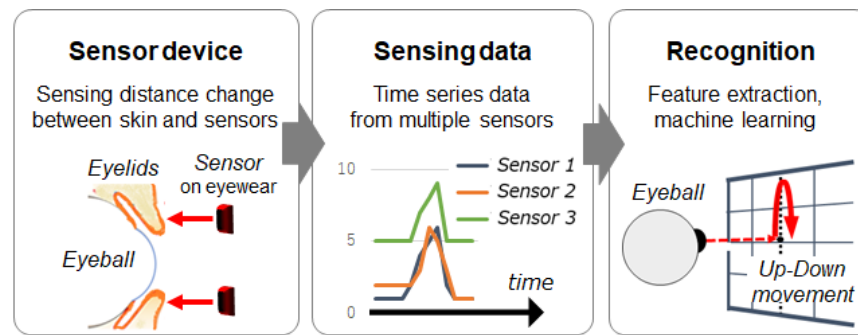


Figure 1. Flowchart of the proposed method. Reprinted with permission from Ref. [25]. Copyright 2021 ACM.

3.1. Recognition Mechanism

To recognize gaze movement, DTW and kNN were used for time series data. DTW (dynamic time warping) is an algorithm for calculating the similarity value of time series.

The details are as follows: (1) First, the similarity value between the acquired data and the training data is calculated by DTW. The training data includes all the gesture data that were prepared in advance. The similarity value is calculated for each sensor. (2) From the training data, the data with high similarity to the acquired data is selected by kNN. From the percentage of gesture labels of the selected data, we calculate the affiliation probability of which gesture label the acquired data is. For example, if kNN ($k = 3$) selected three training data of gesture label 1, the affiliation probability of a gesture label 1 is 100%. The affiliation probability is calculated for each sensor. (3) Then, the gesture label with the highest sum of the affiliation probability of all sensors is judged as the recognition result of the acquired data. For example, if the total number of sensors is two and the affiliation probability of gesture label 1 of sensor one and sensor two are 0.3 and 0.4, the sum of the affiliation probability of gesture label 1 of the acquired data is 0.7 (i.e., the total value of all sensors).

3.2. Implementation

We implemented a prototype system of the proposed method. The entire prototype system consisted of a sensor device, microprocessor (Arduino), laptop, and software. The software was implemented with Processing and Python. Figure 2 shows the system configuration. The prototype device is shown in Figure 3B and consisted of 16 infrared distance sensors (TRP-105, SANYO Electric Co., Ltd., Moriguchi, Japan) that were installed inside the spectacle frame. The sampling rate was 200 [Hz]. k ($k = 7$) of kNN was set to the optimum parameter based on the data of the experimenter although recognition accuracy does not change significantly when the value of k is changed. We prepared two types of feature amounts to determine which one had higher recognition accuracy. The first one was a 16-dimensional pattern, which was the value of the 16 infrared distance sensors. The second one was a 40-dimensional pattern, which was the sum of three factors, i.e., the 16 values of the 16 infrared distance sensors, 16 values of the difference between the 16 adjacent sensors, and eight values of the difference between the 16 diagonal sensors.

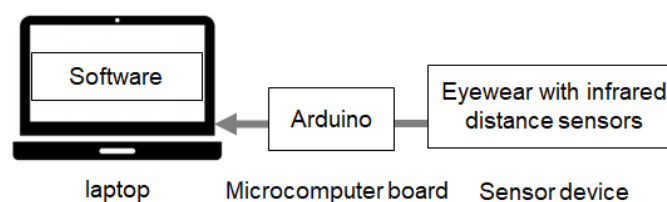


Figure 2. System configuration.

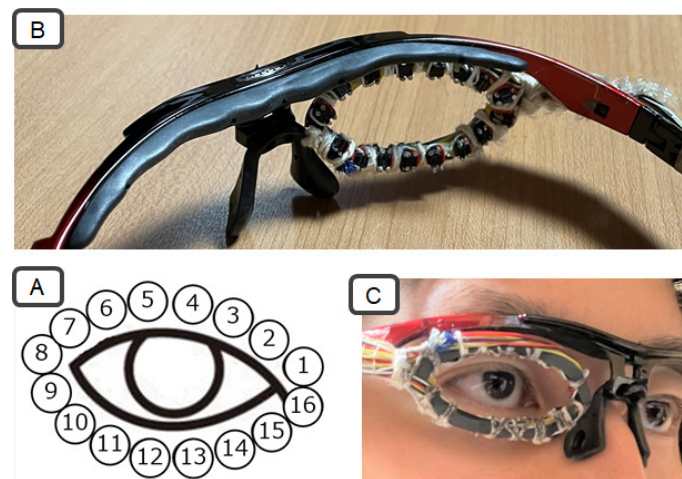


Figure 3. (A) Sensor positions, (B) sensor device, and (C) wearing the sensor device. Reprinted with permission from Ref. [25]. Copyright 2021 ACM.

4. Evaluation 1: Gaze Movement Recognition

In this experiment, we evaluated the eyeball movement recognition accuracy of the proposed method. This experiment focused on two points: (1) First, we evaluated the proposed method's accuracy and limitations in terms of gaze movement recognition. Here, multiple gaze movement patterns were prepared, and the recognition accuracy of each pattern was evaluated. (2) Second, we evaluated the feasibility of a gaze input interface based on gaze movement.

The subjects included 14 college students (average age: 21 years, maximum age: 32 years, minimum age: 20 years; 10 males, 4 females). This study was approved by the research ethics committee of Kobe University (Permission number: 03-19) and was carried out according to the guidelines of the committee.

4.1. Types of Gaze Movement

We prepared 20 types of gaze movement gestures as shown in Figure 4.

The detailed movement involved the following: There were essentially 10 types of gaze movement gestures (G1 to G10) (i.e., Gesture 1 to Gesture 10), where each gesture involved small and large movement patterns, e.g., G1S and G1L (i.e., G1 Small and G1 Large). The 10 main types of movement patterns are summarized as follows: G1 (up and down movement), G2 (up and down movement in the right diagonal direction), G3 (right and left movement), G4 (down and up movement in the left diagonal direction), G5 (down and up movement), G6 (down and up movement in the right diagonal direction), G7 (left and right movement), G8 (up and down movement in the left diagonal direction), G9 (hourglass-shaped movement), and G10 (square movement).

The detailed sizes of each gesture were as follows: Figure 5A shows the distance interval of the marks used when moving the gaze. The marks were placed at a transparent shield, and the transparent shield was positioned in front of the subject's face with a visor, as shown in Figure 5B. Note that the letter P in the following explanation indicates the point P in Figure 5A. For movement patterns G1 to G8, a small movement pattern was between the start point of P13 and one next point (e.g., P12, P7, or P8). For example, pattern G1S involved a movement order of P13, P12, and P13. Pattern G2S involved a movement order of P13, P7, and P13. A large movement pattern was between the start point of P13 and the two next points (e.g., P11, P1, and P3). For example, pattern G1L involved a movement order of P13, P11, and P13, and G2L involved a movement order of P13, P1, and P13. For movement patterns G9 and G10, a small movement pattern was between the start point (P17) and the two next points (e.g., P7). For example, G9S involved a movement order of P17, P7, P19, P9, and P17, and G10S involved a movement order of P17, P7, P9, P19, and P17. Then, a large movement pattern was between the start point (P21) and two next points

(e.g., P1). For example, G9L involved a movement order of P21, P1, P25, P5, and P21, and G10L involved a movement order of P21, P1, P5, P25, and P21. Note that the movement speed was set to 0.5 s when moving between points. For example, about 0.5 s were required to move from P13 to P8 and to move from P13 to P3; thus, gestures G1 to G8 required 1 s, and gestures G9 to G10 required 2 s. How to perform each eye movement gesture and the speed of movement were instructed using a video and the experimenter's explanation.

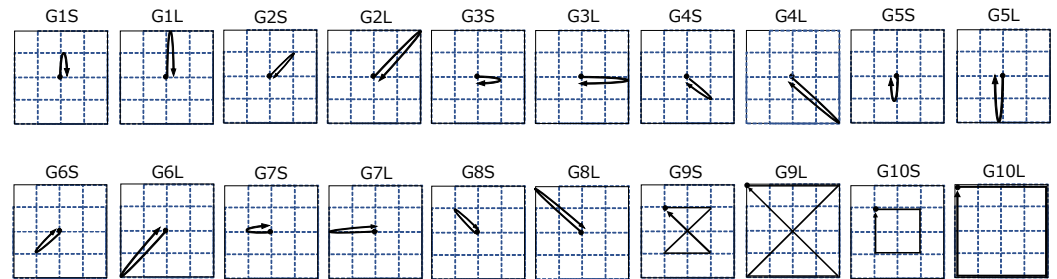


Figure 4. Types of gaze movement. Reprinted with permission from Ref. [25]. Copyright 2021 ACM.

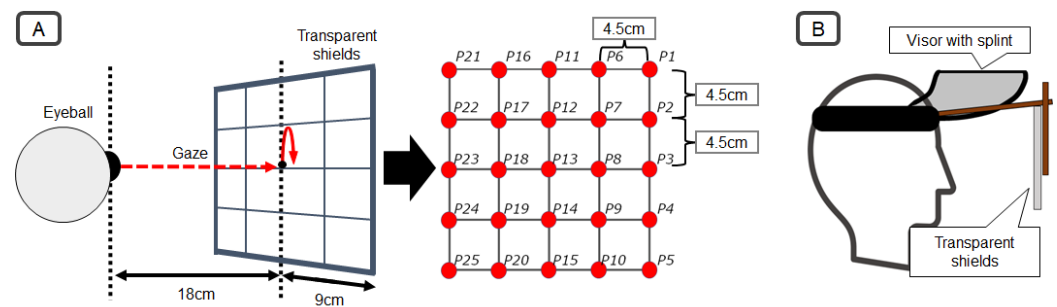


Figure 5. (A) Distance interval of marks used for the eye movement task, and (B) wearing equipment to present marks in eye movement task. Reprinted with permission from Ref. [25]. Copyright 2021 ACM.

4.2. Experimental Procedure

The experimental gaze movement task involved performing a designated gaze movement. Here, the subject wore the prototype device and sat on a chair. In addition, a shield (Figure 5B) was attached to the subject's head. Visible points were arranged in front of the shield to guide the subject's gaze movements. Based on the points on the shield, the subject was instructed on how to perform each eye movement gesture and the speed of movement using a video and the experimenter's explanation. One trial involved performing the 20 types of gaze movements. The order of gaze movements performed was random. 10 trials were performed in this experiment. Therefore, data per person consisted of 10 trials for each of 20 different gestures. For gaze movements G1S to G8S and G1L to G8L, the time-series data were recorded for 1 s (i.e., 200 samples) because the time required for a single movement was approximately 1 s. For gaze movements G9S, G10S, G9L, and G10L, the time-series data were recorded for 2 s (i.e., 400 samples) because the time required for a single movement was approximately 2 s. This experiment assumed an intentional gaze input gesture; thus, the subjects did not blink while performing the gaze movements.

10-fold cross-validation was performed on the data acquired in the 10 trials. Recognition accuracy was evaluated for the following three patterns.

- (1) The 5-movement pattern comprised G1S, G3S, G5S, G7S, and G9S. This was set to evaluate the feasibility of a gaze input interface using the proposed method. An example of an application with a hands-free input interface is a media player (e.g., music, video, and still images). To operate such applications, it is sufficient to use approximately five types of commands, e.g., play, stop, forward, and back. In fact,

the effectiveness of the hands-free input method has been evaluated previously using approximately five gestures [8]. In addition, if five gestures can be recognized, dozens of diverse inputs can be produced by combining those gestures. This movement pattern was considered to evaluate whether the proposed method could recognize differences in moving direction at every 90°s interval.

- (2) The 10-movement pattern comprised G1S to G10S. This pattern was set to evaluate whether the proposed method could recognize differences in moving direction every 45°s interval.
- (3) The 20-movement pattern comprised all movements containing the small and large movement patterns. This pattern was set to evaluate whether the difference in the degree of movement in the same direction could be recognized.

4.3. Result

Figure 6 shows the F-value results. Table 1 shows the results for F-value, precision, and recall. The results are shown for each type of feature amount. Figure 7 shows the confusion matrix for each gaze movement. The value shown is the F-value. This result is the 16-dimensional pattern. This figure helps us understand the tendency toward misrecognition of each gaze direction due to the increased number of gaze movements. Figure 8 and Table 2 show the results for each individual. The value shown is the F-value. This result is the 16-dimensional pattern. This helps us understand individual differences in the recognition accuracy of the proposed method.

The 5-movement pattern gave an average F-value of 1.0. The 5-movement pattern consisted of gaze movements with vertical and horizontal movement directions at 90° intervals. Therefore, this result indicates that the proposed method was able to recognize the moving direction at 90° intervals with high accuracy.

The 10-movement pattern gave an average F-value of approximately 0.99. In addition, the proportion of F-values of 0.9 or more was 100% (14 subjects). The 10-movement pattern consisted of gaze movements with movement directions at 45° intervals. Therefore, this result indicates that the proposed method was effective for all subjects and could recognize the moving direction at 45° intervals with high accuracy.

The result of the 20-movement pattern gave an average F-value of approximately 0.96. In addition, the proportion of F-values of 0.9 or more was 86% (12 subjects). The 20-movement pattern consisted of gaze movements with movement directions at 45° intervals containing small and large movement patterns. These results indicate that our method could recognize the change in gaze movement, which was about twice as large, in the same direction with high accuracy.

The tendency for erroneous recognition due to increased gaze movements was as follows: Erroneous recognition between small and large movement patterns (e.g., between G5S and G5L) tended to increase, as demonstrated by the 20-movement pattern. Therefore, erroneous recognition was assumed to increase as the difference in the degree of movement in the same direction decreased. There was almost no difference between the results for the 16-dimensional pattern and the 40-dimensional pattern. Therefore, it is appropriate to adopt the 16-dimensional pattern considering the calculation cost.

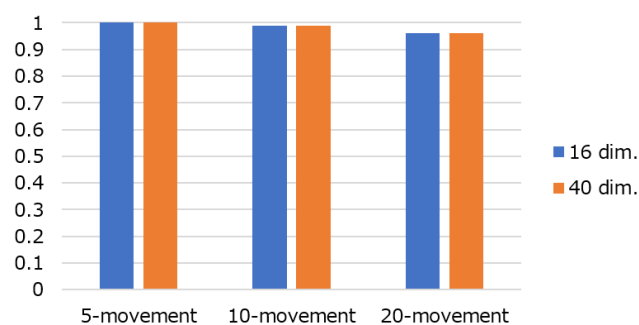


Figure 6. Results for gaze movement pattern in Evaluation 1. The value is the F-value.

Table 1. Results for gaze movement pattern in Evaluation 1. (R: Recall, P: Precision, F: F-value).

	5-Movement Pattern			10-Movement Pattern			20-Movement Pattern		
	F	P	R	F	P	R	F	P	R
16 dim.	1.00	1.00	1.00	0.99	0.99	0.99	0.96	0.96	0.97
40 dim.	1.00	1.00	1.00	0.99	0.99	0.99	0.96	0.96	0.97

5-movement pattern

G1S	1	0	0	0	0
G3S	0	1	0	0	0
G5S	0	0	1	0	0
G7S	0	0	0	1	0
G10S	0	0	0	0	1

10-movement pattern

G1S	0.98	0.02	0	0	0	0	0	0	0
G2S	0.01	0.99	0	0	0	0	0	0	0
G3S	0	0	1	0	0	0	0	0	0
G4S	0	0	0	0.99	0.01	0	0	0	0
G5S	0	0	0	0.01	0.99	0	0	0	0
G6S	0	0	0	0	0	1	0	0	0
G7S	0	0	0	0	0	0	0.99	0.01	0
G8S	0	0	0	0	0	0	0	1	0
G9S	0	0	0	0	0	0	0	0	0.99
G10S	0	0	0	0	0	0	0	0	1

20-movement pattern

G1L	0.99	0.01	0	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G1S	0.02	0.96	0	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G2L	0.01	0	0.98	0	0	0	0	0	0	0	0	0	0	0	0	0	0.01	0	0
G2S	0	0.01	0.01	0.99	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G3L	0	0	0	0	0.98	0.02	0	0	0	0	0	0	0	0	0	0	0	0	0
G3S	0	0	0	0	0.01	0.99	0	0	0	0	0	0	0	0	0	0	0	0	0
G4L	0	0	0	0	0	0	0.96	0.03	0	0.01	0	0	0	0	0	0	0	0	0
G4S	0	0	0	0	0	0	0	0.99	0	0.01	0	0	0	0	0	0	0	0	0
G5L	0	0	0	0	0	0	0	0	0.96	0.04	0	0.01	0	0	0	0	0	0	0
G5S	0	0	0	0	0	0.01	0	0.02	0.96	0	0.01	0	0	0	0	0	0	0	0
G6L	0	0	0	0	0	0	0	0	0	0.9	0.09	0	0.01	0	0	0	0	0	0
G6S	0	0	0	0	0	0	0	0	0	0.05	0.95	0	0	0	0	0	0	0	0
G7L	0	0	0	0.01	0	0	0	0	0	0	0	0	0.96	0.03	0	0	0	0	0
G7S	0	0	0	0	0	0	0	0	0	0	0	0	0.01	0.98	0	0.01	0	0	0
G8L	0	0.01	0	0	0	0	0	0	0	0	0	0.01	0	0.98	0.01	0	0	0	0
G8S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
G9L	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.93	0	0.06	0.01
G9S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.98	0	0.02
G10L	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
G10S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.01	0.01	0.99

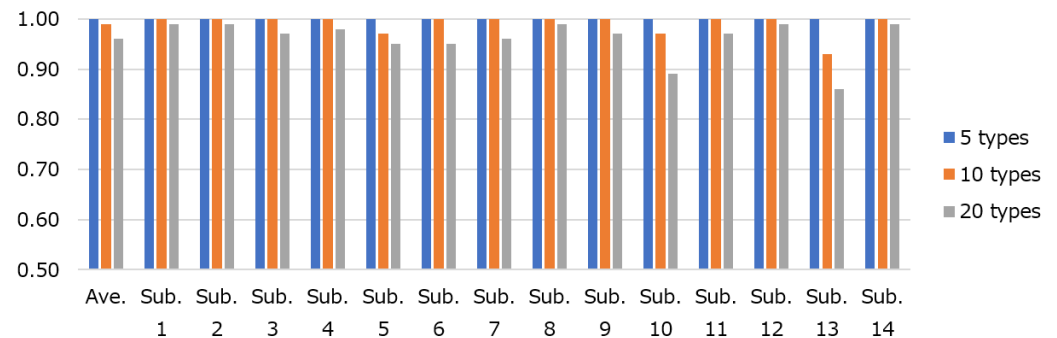
Figure 7. Results of each gaze movement in Evaluation 1. The value is the F-value.**Figure 8.** Results of each individual in Evaluation 1. The value is the F-value.

Table 2. Results of each individual in Evaluation 1. The value is the F-value.

	5-Movement	10-Movement	20-Movement
Sub. 1	1.00	1.00	0.99
Sub. 2	1.00	1.00	0.99
Sub. 3	1.00	1.00	0.99
Sub. 4	1.00	1.00	0.97
Sub. 5	1.00	1.00	0.98
Sub. 6	1.00	0.97	0.95
Sub. 7	1.00	1.00	0.95
Sub. 8	1.00	1.00	0.96
Sub. 9	1.00	1.00	0.99
Sub. 10	1.00	1.00	0.97
Sub. 11	1.00	0.97	0.89
Sub. 12	1.00	1.00	0.97
Sub. 13	1.00	1.00	0.99
Sub. 14	1.00	0.93	0.86
Ave.	1.00	0.99	0.96

4.4. Discussion

The results showed the feasibility of the proposed method for a simple hands-free input method. The gaze movement appears on the eyelid skin, and the proposed method recognized the pattern of gaze movement. The proposed method can recognize 5 to 20 types of gaze movement patterns with an F-value of 0.9 or higher. Movement directions at 90° intervals were recognized with high accuracy, and movement directions at 45° intervals and the movement distance difference in the same direction were also recognized. A previous study showed that about five types of command recognition are necessary and sufficient for simple hands-free input [8]. For example, to operate a media player (e.g., music, video, still image), it was sufficient to use approximately five types of commands, e.g., play, stop, forward, and back. In addition, previous studies investigating a simple hands-free input method showed that five to seven types of input gestures of the face or gaze are recognized with F-values of 0.85 to 0.9 [8,35,50], although these studies were not based on the same experiments. Based on these, the proposed method seems to have the same level of recognition accuracy as the previous study, and the proposed method can be utilized as a simple hands-free input method.

5. Evaluation 2: Recognition Accuracy with a Small Number of Sensors

In this experiment, we investigated whether the proposed method could recognize gaze movement gestures with a small number of sensors, such as two or three sensors. In addition, we investigated the necessary sensor positions for recognition when the number of sensors was small. It was feasible to implement the proposed method using the minimum number of sensors; thus, this evaluation provided an example of its application.

Here, the proposed method was evaluated with five gestures that were the same as in the previous section. We investigated the combination of two or three sensors with high recognition accuracy. We narrowed down the number of sensors to be used to eight on the upper side of the eyeglass frame. This was because the sensors on the lower side of the eyeglass frame could not be used for gaze movement recognition when the lower side of the eyeglass frame was wide. This setting was enough to assess whether the proposed method could recognize a small number of sensors. With two sensors, the recognition accuracy was calculated for 28 patterns of the combination of two sensors. With three sensors, the recognition accuracy was calculated for 56 patterns of the combination of three sensors. Since this evaluation aimed to assess whether the proposed method could recognize a small number of sensors, we only used a single pattern of feature amounts of the 16-dimensional pattern.

5.1. Result

10-fold cross-validation was performed on the data acquired in the 10 trials. Tables 3 and 4 show the combinations of the top five patterns. Regarding the two-sensor pattern, the highest F-value was 0.918, and the combination of sensors with the highest F-value included the 1st and 8th sensors. Regarding the three-sensor pattern, the highest F-value was 0.966, and the combination of sensors with the highest F-value was the 4th, 7th, and 8th sensors. Figure 9 shows the examples of the proposed method using two or three sensors.

Table 3. Result of combination of two sensors with high accuracy in Evaluation 2. (R: Recall, P: Precision, F: F-value).

Rank	Combination	F	P	R
1st	1, 8	0.918	0.919	0.928
2nd	4, 8	0.902	0.903	0.916
3rd	5, 8	0.888	0.890	0.905
4th	1, 5	0.865	0.866	0.882
5th	2, 8	0.864	0.875	0.881

Table 4. Result of combination of three sensors with high accuracy in Evaluation 2. (R: Recall, P: Precision, F: F-value).

Rank	Combination	F	P	R
1st	4, 7, 8	0.966	0.962	0.974
2nd	1, 4, 7	0.965	0.961	0.973
3rd	3, 4, 8	0.961	0.957	0.970
4th	2, 4, 6	0.960	0.955	0.970
5th	1, 4, 6	0.959	0.955	0.969

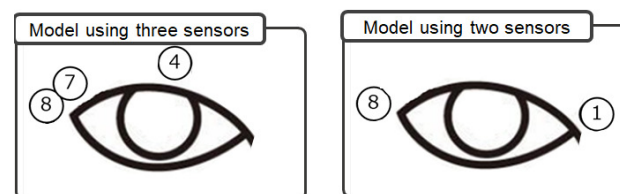


Figure 9. The sensor position with the highest recognition accuracy when the number of sensors was two or three in Evaluation 2.

5.2. Discussion

The results showed that the proposed method could be used with a small number of sensors, such as two or three sensors. Sensor positions where gaze movement was likely to be sensed were also identified. The proposed method with a small number of sensors was assumed to be adequate for a simple hands-free input method since it had an F-value of 0.9 or more. Although reducing the number of sensors can decrease the recognition accuracy, reducing the number of sensors is an appropriate design for implementing the proposed method for reasons such as lower cost and lighter weight of the whole system, if the necessary recognition accuracy can be obtained.

6. Evaluation 3: Robustness in Gaze Movement Recognition

This experiment evaluated whether the recognition accuracy of the proposed method changed under various conditions assuming the usage scenario of wearable devices. The conditions were reattachment, body vibration, and facial expression change. Each condition was as follows:

Condition 1. Reattachment condition

This condition evaluated the accuracy of the proposed method after the sensor device was reattached. This condition was selected to assess whether it is necessary to reacquire the learning data when using the proposed method after the sensor device is reattached. Thirteen subjects participated in this experiment. Firstly, each subject performed ten trials for the 5-movement pattern task. This task was the same as for Evaluation 1 and under the same conditions as Evaluation 1. Subjects who had already participated in Evaluation 1 did not perform this data acquisition exercise. The data for these ten trials were used as training data. Then, each subject reattached the sensor device and performed five trials of the 5-movement pattern task. The data for these five trials were used as test data.

Condition 2. Body vibration condition

This condition evaluated the accuracy of the proposed method when body vibration occurs (e.g., walking). Twelve subjects participated in this experiment. In this evaluation, walking vibration was considered. Strong body vibration (e.g., a vibration that occurs when dashing or dancing) was not assessed because such vibration is considered to clearly decrease the accuracy of the proposed method. Firstly, each subject performed ten trials of 5-movement pattern task. This task was the same as for Evaluation 1 and performed under the same conditions as Evaluation 1. Subjects who had already participated in Evaluation 1 did not perform this data acquisition. The data for these ten trials were used as training data. Then, each subject performed five trials of the 5-movement pattern task while walking on the spot. The walking speed reflected the natural speed of each individual. The data for these five trials were used as test data.

Condition 3. Facial expression change condition

Although the same value for the infrared distance sensors can be obtained with the same facial expression, we should verify whether it is necessary to reacquire the learning data when the facial expression changes. Therefore, this condition evaluated whether the learning data acquired when normal facial expressions can be used for different facial expressions. A smile was adopted for the facial expression. This was because the previous study [19], which involved detection of blinks using an infrared distance sensor, reported an example in which the accuracy of blink recognition was reduced due to the facial expression of a smile. The number of subjects was thirteen. Firstly, each subject performed ten trials of the 5-movement pattern task. This task was the same as for Evaluation 1 and under the same conditions as Evaluation 1. Subjects who had already participated in Evaluation 1 did not perform this data acquisition exercise. The data for these ten trials were used as training data. Each subject performed five trials of the 5-movement pattern task while smiling. The smiling expression involved the intentional raising of both corners of the mouth so that the teeth could be seen. The data of these five trials were used as test data.

6.1. Result

The recognition results for each condition are shown in Figure 10 and Table 5. Values indicate F-value, precision, and recall. This result relates to the 16-dimensional pattern.

The results of the reattachment conditions were as follows: As an overall tendency, the recognition accuracy after reattachment slightly decreased when using the learning data before reattachment compared with Evaluation 1. The average F-value was 0.95 for the 5-movement pattern, which was lower than that in Evaluation 1. From the results for each individual, the proportion of F-values of 0.9 or more was 77% (10 subjects of 13 subjects). These results showed that the 5-movement pattern could be recognized with high recognition accuracy even after reattachment.

The results of the body vibration condition were as follows. As an overall tendency, the recognition accuracy decreased. The average F-value was 0.89 for the 5-movement pattern. From the results for each individual, the proportion of F-values of 0.9 or more was 58% (7 of 12 subjects). These results indicate that the accuracy of our method is assumed to decrease in situations where the body vibrates. However, more than half of the people had high recognition accuracy, which indicates that there were individual differences in the effect of body vibration on accuracy.

The results of the facial expression change conditions were as follows. As an overall tendency, the recognition accuracy slightly decreased. The average F-value was 0.91 for the 5-movement pattern. From the results for each individual, the proportion of F-values of 0.9 or more was 77% (10 of 13 subjects). Although these results indicate that the accuracy of our method is assumed to decrease when facial expression changes, more than half of the participants had high recognition accuracy.

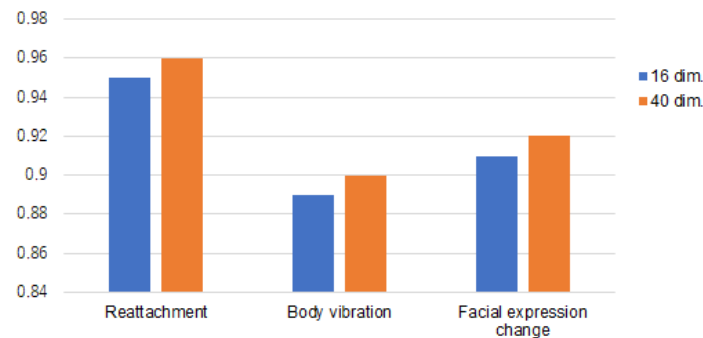


Figure 10. Results for gaze movement pattern in Evaluation 3. The value is the F-value.

Table 5. Result of each condition for each pattern in Evaluation 3. (R: Recall, P: Precision, F: F-measure).

		5-Movement Pattern		
		F	P	R
Reattachment	16 dim.	0.95	0.96	0.95
	40 dim.	0.96	0.97	0.96
Body vibration	16 dim.	0.89	0.92	0.90
	40 dim.	0.90	0.93	0.91
Facial expression change	16 dim.	0.91	0.93	0.92
	40 dim.	0.92	0.95	0.92

6.2. Discussion

The results showed that the proposed method could recognize the gaze movement pattern, although the recognition accuracy decreased due to the disturbance in each condition. The recognition accuracy did not decrease significantly after reattachment. Therefore, it can be assumed that the proposed method can obtain high recognition accuracy if the mounting positions of the eyewear devices are almost the same. However, if the mounting position shifts significantly after remounting and the recognition accuracy decreases, it is necessary to correct the mounting position. If the recognition accuracy does not improve after correcting the mounting position, it is necessary to acquire the learning data again. Since the learning data can be acquired in a short time (e.g., about 1 min is required for five types of gestures and five trials), the burden on the user is not large. Under the body vibration condition and facial expression change condition, the recognition accuracy decreased to about 0.9 of the F-value. This F-value seems to be sufficient for the proposed method. However, in order to obtain high recognition accuracy under such conditions, the user should apply ingenuity when using hands-free input, such as returning the facial expression to neutral (i.e., returning to the same facial expression as when acquiring the learning data) and reducing body vibration (e.g., slowing down walking speed).

7. Limitations and Future Work

This paper showed the following. Evaluation 1 showed that the proposed method could recognize patterns of gaze movements and be feasible for a simple hands-free input

method. Evaluation 2 provided an example in which the proposed method could recognize the eye movement pattern even with a small number of sensors, such as two to three, leading to reducing the cost and weight of the entire system. Evaluation 3 showed that the proposed method could recognize the gaze movement pattern, although the recognition accuracy decreased due to disturbance in the usage scene of the wearable device. These results are helpful for designing gaze movement recognition methods using eyewear with infrared distance sensors. This section describes future work.

Individual differences that may affect recognition accuracy: We plan to investigate individual characteristics that reduce the recognition accuracy of the proposed method. For example, Sub.14 in Evaluation 1 had lower recognition accuracy than other subjects. Some factors may affect the recognition accuracy of the proposed method. For example, recognition accuracy will decrease if the eyelid skin does not move much when moving the gaze. In addition, recognition accuracy may decrease for people with very dark skin color since the infrared distance sensor often has difficulty responding to black color. Since the subjects were mainly young Asians, we plan to evaluate subjects with various attributes in the future.

Evaluation in the natural environment: We plan to evaluate the proposed method in the natural environment. For example, the intensity of ambient light is a factor that affects the recognition accuracy of the proposed method. The value of the infrared distance sensor differs between outdoors and indoors because the intensity of ambient light differs between outdoors and indoors. Although the same value of infrared distance sensors can be obtained with the same lighting intensity, we should verify whether it is necessary to reacquire the learning data when entering an environment with different lighting intensity.

Investigation of the limits of recognizable gaze movement patterns and expanding the number of input commands while maintaining recognition accuracy: This paper examined about 20 types of gaze movement patterns. Although these patterns are assumed to be enough in many hands-free input scenes, we plan to investigate how much difference in movement patterns can be recognized in the future (e.g., the difference in moving direction at 10, 20, and 30 degrees intervals). In addition, although one gesture was performed independently in this experiment, the number of input commands can be increased while maintaining high recognition accuracy by combining gestures with high recognition accuracy. For example, if users want 10 input commands with higher recognition accuracy, instead of using the gaze movements with 45° intervals used in Evaluation 1, combining multiple gaze movements with 90° intervals is assumed to be useful. Command examples consisting of two gestures include continuous inputs of G1 and G3 or repeating G1 twice.

Application of the proposed method: We plan to verify whether the gaze movement recognition of the proposed method can be applied to applications other than hands-free input interfaces. Recognizing and clarifying the characteristics of humans is one of the applications. For example, medically important findings of individual characteristics have been clarified from eye activities, such as ADHD [51], autism [52], Williams syndrome [53], schizophrenia, and Parkinson's disease. In addition, the characteristics of the gaze movement pattern of a highly skilled player (e.g., gaze movement of a highly skilled basketball player before the shot is small [54]) may be investigated. Although the recognition accuracy is assumed to be lower if gaze movement contains randomized disturbances, the proposed method can recognize them if a similar pattern of gaze movements occurs consistently.

8. Conclusions

In this study, we proposed a method for recognizing gaze movements using eyewear equipped with multiple infrared distance sensors as a simple method to add gaze interaction functions to eyewear. We implemented a prototype system and conducted evaluations of the gaze movements, including movement directions at 45° intervals and the movement distance difference in the same direction. The results showed that the proposed method could recognize gaze movement patterns and be feasible as a simple hands-free input method. The proposed method recognized five types of movement with an F-value of

1.0, 10 types of movement with an F-value of 0.99, and 20 types of movement with an F-value of 0.96. In addition, the proposed method recognized the gaze movement pattern in conditions of reattachment, body vibration, and facial expression change, although the recognition accuracy decreased. The results also showed that the proposed method recognized gaze movements even with a small number of sensors, such as two to three. Since the previous studies for a simple hands-free input method showed that five to seven types of input gestures of the face or gaze are recognized with an F-value of 0.85 to 0.9, the proposed method seems to have the same level of recognition accuracy as the previous study. These results are helpful for designing gaze movement recognition methods for continuous daily use, using eyewear with infrared distance sensors.

Author Contributions: Conceptualization, K.F., Y.T., K.M. and T.T.; methodology, K.F. and Y.T.; software, Y.T.; validation, K.F. and Y.T.; formal analysis, K.F. and Y.T.; investigation, K.F. and Y.T.; resources, K.F. and Y.T.; data curation, K.F. and Y.T.; writing—original draft preparation, K.F. and Y.T.; writing—review and editing, K.F. and Y.T.; visualization, K.F. and Y.T.; supervision, K.F., K.M. and T.T.; project administration, K.F., K.M. and T.T.; funding acquisition, K.F. and T.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by JSPS (Japan Society for the Promotion of Science) KAKENHI Grant Number JP19K20330, and JST, CREST Grant Number JPMJCR18A3, Japan.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Commission of Kobe University (n. 03-19, 1 November 2021).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hutchinson, T.E.; White, K.P.; Martin, W.N.; Reichert, K.C.; Frey, L.A. Human-Computer Interaction Using Eye-Gaze Input. *IEEE Trans. Syst. Man Cybern.* **1989**, *19*, 1527–1534. [\[CrossRef\]](#)
2. Manabe, H.; Fukumoto, M.; Yagi, T. Conductive Rubber Electrodes for Earphone-Based Eye Gesture Input Interface. *Pers. Ubiquitous Comput.* **2015**, *19*, 143–154. [\[CrossRef\]](#)
3. Jacob, R.; Stellmach, S. What You Look at Is What You Get: Gaze-Based User Interfaces. *Interactions* **2016**, *23*, 62–65. [\[CrossRef\]](#)
4. Salvucci, D.D.; Anderson, J.R. Intelligent Gaze-Added Interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, The Hague, The Netherlands, 1–6 April 2000; pp. 273–280.
5. Menges, R.; Kumar, C.; Müller, D.; Sengupta, K. GazeTheWeb: A Gaze-Controlled Web Browser. In Proceedings of the 14th International Web for All Conference, Perth, Australia, 2–4 April 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 1–2.
6. He, J.; Chaparro, A.; Nguyen, B.; Burge, R.; Crandall, J.; Chaparro, B.; Ni, R.; Cao, S. Texting While Driving: Is Speech-Based Texting Less Risky than Handheld Texting? In Proceedings of the Fifth International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Eindhoven, The Netherlands, 28–30 October 2013; pp. 124–130.
7. Feng, J.; Sears, A. Using Confidence Scores to Improve Hands-Free Speech Based Navigation in Continuous Dictation Systems. *ACM Trans. Comput.-Hum. Interact. (TOCHI)* **2004**, *11*, 329–356. [\[CrossRef\]](#)
8. Amesaka, T.; Watanabe, H.; Sugimoto, M. Facial Expression Recognition Using Ear Canal Transfer Function. In Proceedings of the 23rd International Symposium on Wearable Computers, London, UK, 9–13 September 2019; pp. 1–9.
9. Fejtová, M.; Figueiredo, L.; Novák, P.; Štěpánková, O.; Gomes, A. Hands-Free Interaction with a Computer and Other Technologies. *Univers. Access Inf. Soc.* **2009**, *8*, 277–295. [\[CrossRef\]](#)
10. Futami, K.; Oyama, K.; Murao, K. Augmenting Ear Accessories for Facial Gesture Input Using Infrared Distance Sensor Array. *Electronics* **2022**, *11*, 1480. [\[CrossRef\]](#)
11. Ogata, M.; Sugiura, Y.; Osawa, H.; Imai, M. IRing: Intelligent Ring Using Infrared Reflection. In Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology, Cambridge, MA, USA, 7–10 October 2012; pp. 131–136.
12. Niehorster, D.C.; Hessels, R.S.; Benjamins, J.S. GlassesViewer: Open-Source Software for Viewing and Analyzing Data from the Tobii Pro Glasses 2 Eye Tracker. *Behav. Res. Methods* **2020**, *52*, 1244–1253. [\[CrossRef\]](#)
13. Kassner, M.; Patera, W.; Bulling, A. Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-Based Interaction. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, Seattle, WA, USA, 13–17 September 2014; pp. 1151–1160.
14. Javadi, A.H.; Hakimi, Z.; Barati, M.; Walsh, V.; Tcheang, L. SET: A Pupil Detection Method Using Sinusoidal Approximation. *Front. Neuroeng.* **2015**, *8*, 4. [\[CrossRef\]](#)

15. Tonsen, M.; Zhang, X.; Sugano, Y.; Bulling, A. Labelled Pupils in the Wild: A Dataset for Studying Pupil Detection in Unconstrained Environments. In Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research and Applications, Charleston, SC, USA, 14–17 March 2016; pp. 139–142.
16. Rostamina, S.; Mayberry, A.; Ganesan, D.; Marlin, B.; Gummesson, J. IliD: Low-Power Sensing of Fatigue and Drowsiness Measures on a Computational Eyeglass. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2017**, *1*, 1–26. [CrossRef]
17. Ahmad, R. Understanding the Language of the Eye: Detecting and Identifying Eye Events in Real Time via Electrooculography. Ph.D. Thesis, University of California, San Diego, CA, USA, 2016.
18. The Google Glass Wink Feature Is Real | TechCrunch. Available online: <https://techcrunch.com/2013/05/09/the-google-glass-wink-feature-is-real/> (accessed on 29 December 2021).
19. Dementyev, A.; Holz, C. DualBlink: A Wearable Device to Continuously Detect, Track, and Actuate Blinking for Alleviating Dry Eyes and Computer Vision Syndrome. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2017**, *1*, 1–19. [CrossRef]
20. Masai, K.; Sugiura, Y.; Ogata, M.; Suzuki, K.; Nakamura, F.; Shimamura, S.; Kunze, K.; Inami, M.; Sugimoto, M. AffectiveWear: Toward Recognizing Facial Expression. In Proceedings of the ACM SIGGRAPH 2015 Emerging Technologies, Los Angeles, CA, USA, 9–13 August 2015; p. 1.
21. Fukumoto, K.; Terada, T.; Tsukamoto, M. A Smile/Laughter Recognition Mechanism for Smile-Based Life Logging. In Proceedings of the fourth Augmented Human International Conference, Stuttgart, Germany, 7–8 March 2013; pp. 213–220.
22. Masai, K.; Sugiura, Y.; Ogata, M.; Kunze, K.; Inami, M.; Sugimoto, M. Facial Expression Recognition in Daily Life by Embedded Photo Reflective Sensors on Smart Eyewear. In Proceedings of the 21st International Conference on Intelligent User Interfaces, Sonoma, CA, USA, 7–10 March 2016; pp. 317–326.
23. Matsui, S.; Terada, T.; Tsukamoto, M. Smart Eye Mask: Sleep Sensing System Using Infrared Sensors. In Proceedings of the 2017 ACM International Symposium on Wearable Computers, Maui, HI, USA, 11–15 September 2017; pp. 58–61.
24. Futami, K. A Method to Recognize Eye Movements Based on Uplift Movement of Skin. In Proceedings of the Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers, London, UK, 9–13 September 2019; pp. 624–627.
25. Futami, K.; Tabuchi, Y.; Murao, K.; Terada, T. A Method to Recognize Eyeball Movement Gesture Using Infrared Distance Sensor Array on Eyewear. In Proceedings of the 23rd International Conference on Information Integration and Web Intelligence, Linz, Austria, 29 November–1 December 2021; pp. 645–649.
26. Ishiguro, Y.; Mujibiyah, A.; Miyaki, T.; Rekimoto, J. Aided Eyes: Eye Activity Sensing for Daily Life. In Proceedings of the 1st Augmented Human International Conference, Megève, France, 2–3 April 2010; pp. 1–7.
27. Hsieh, C.-S.; Tai, C.-C. An Improved and Portable Eye-Blink Duration Detection System to Warn of Driver Fatigue. *Instrum. Sci. Technol.* **2013**, *41*, 429–444. [CrossRef]
28. Bulling, A.; Roggen, D.; Tröster, G. Wearable EOG Goggles: Seamless Sensing and Context-Awareness in Everyday Environments. *J. Ambient. Intell. Smart Environ.* **2009**, *1*, 157–171. [CrossRef]
29. Picot, A.; Caplier, A.; Charbonnier, S. Comparison between EOG and High Frame Rate Camera for Drowsiness Detection. In Proceedings of the 2009 Workshop on Applications of Computer Vision (WACV), Snowbird, UT, USA, 7–9 December 2009; pp. 1–6.
30. Manabe, H.; Fukumoto, M. Full-Time Wearable Headphone-Type Gaze Detector. In Proceedings of the CHI'06 Extended Abstracts on Human Factors in Computing Systems, Montréal, QC, Canada, 22–27 April 2006; pp. 1073–1078.
31. Uema, Y.; Inoue, K. JINS MEME Algorithm for Estimation and Tracking of Concentration of Users. In Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers, Maui, HI, USA, 11–15 September 2017; pp. 297–300.
32. Bulling, A.; Roggen, D.; Tröster, G. Eyemote—towards Context-Aware Gaming Using Eye Movements Recorded from Wearable Electrooculography. In Proceedings of the International Conference on Fun and Games, Eindhoven, The Netherlands, 20–21 October 2008; Springer: Berlin, Germany, 2008; pp. 33–45.
33. Bulling, A.; Ward, J.A.; Gellersen, H.; Tröster, G. Eye Movement Analysis for Activity Recognition Using Electrooculography. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 741–753. [CrossRef]
34. Ishimaru, S.; Kunze, K.; Kise, K.; Weppner, J.; Dengel, A.; Lukowicz, P.; Bulling, A. In the Blink of an Eye: Combining Head Motion and Eye Blink Frequency for Activity Recognition with Google Glass. In Proceedings of the 5th Augmented Human International Conference, Kobe, Japan, 7–8 March 2014; pp. 1–4.
35. Masai, K.; Kunze, K.; Sugimoto, M. Eye-Based Interaction Using Embedded Optical Sensors on an Eyewear Device for Facial Expression Recognition. In Proceedings of the Augmented Humans International Conference, Kaiserslautern, Germany, 16–17 March 2020; pp. 1–10.
36. Suzuki, K.; Nakamura, F.; Otsuka, J.; Masai, K.; Itoh, Y.; Sugiura, Y.; Sugimoto, M. Recognition and Mapping of Facial Expressions to Avatar by Embedded Photo Reflective Sensors in Head Mounted Display. In Proceedings of the 2017 IEEE Virtual Reality (VR), Los Angeles, CA, USA, 18–22 March 2017; pp. 177–185.
37. Masai, K.; Sugiura, Y.; Sugimoto, M. Facerubbing: Input Technique by Rubbing Face Using Optical Sensors on Smart Eyewear for Facial Expression Recognition. In Proceedings of the 9th Augmented Human International Conference, Seoul, Korea, 7–9 February 2018; pp. 1–5.

38. Kikuchi, T.; Sugiura, Y.; Masai, K.; Sugimoto, M.; Thomas, B.H. EarTouch: Turning the Ear into an Input Surface. In Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, Vienna, Austria, 4–7 September 2017; pp. 1–6.
39. Taniguchi, K.; Kondo, H.; Kurosawa, M.; Nishikawa, A. Earable TEMPO: A Novel, Hands-Free Input Device That Uses the Movement of the Tongue Measured with a Wearable Ear Sensor. *Sensors* **2018**, *18*, 733. [[CrossRef](#)] [[PubMed](#)]
40. Fukui, R.; Watanabe, M.; Gyota, T.; Shimosaka, M.; Sato, T. Hand Shape Classification with a Wrist Contour Sensor: Development of a Prototype Device. In Proceedings of the 13th International Conference on Ubiquitous Computing, Beijing, China, 17–21 September 2011; pp. 311–314.
41. Hashimoto, T.; Low, S.; Fujita, K.; Usumi, R.; Yanagihara, H.; Takahashi, C.; Sugimoto, M.; Sugiura, Y. TongueInput: Input Method by Tongue Gestures Using Optical Sensors Embedded in Mouthpiece. In Proceedings of the 2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), Nara, Japan, 11–14 September 2018; pp. 1219–1224.
42. Crossan, A.; Williamson, J.; Brewster, S.; Murray-Smith, R. Wrist Rotation for Interaction in Mobile Contexts. In Proceedings of the Tenth International Conference on Human Computer Interaction with Mobile Devices and Services, Amsterdam, The Netherlands, 2–5 September 2008; pp. 435–438.
43. Probst, K.; Lindlbauer, D.; Haller, M.; Schwartz, B.; Schrempf, A. A Chair as Ubiquitous Input Device: Exploring Semaphoric Chair Gestures for Focused and Peripheral Interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, ON, Canada, 26 April–1 May 2014; pp. 4097–4106.
44. Hirsch, H.-G.; Pearce, D. The Aurora Experimental Framework for the Performance Evaluation of Speech Recognition Systems under Noisy Conditions. In Proceedings of the ASR2000-Automatic Speech Recognition: Challenges for the New Millenium ISCA Tutorial and Research Workshop (ITRW), Paris, France, 18–20 September 2000.
45. Tang, Z.; Yan, C.; Ren, S.; Wan, H. HeadPager: Page Turning with Computer Vision Based Head Interaction. In Proceedings of the Asian Conference on Computer Vision, Taipei, Taiwan, 20–24 November 2016; pp. 249–257.
46. Gorodnichy, D.O.; Roth, G. Nouse ‘Use Your Nose as a Mouse’ Perceptual Vision Technology for Hands-Free Games and Interfaces. *Image Vis. Comput.* **2004**, *22*, 931–942. [[CrossRef](#)]
47. Crossan, A.; McGill, M.; Brewster, S.; Murray-Smith, R. Head Tilting for Interaction in Mobile Contexts. In Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services, Bonn, Germany, 15–18 September 2009; pp. 1–10.
48. Jalaliniya, S.; Mardanbeigi, D.; Pederson, T.; Hansen, D.W. Head and Eye Movement as Pointing Modalities for Eyewear Computers. In Proceedings of the 2014 11th International Conference on Wearable and Implantable Body Sensor Networks Workshops, Washington, DC, USA, 16–19 June 2014; pp. 50–53.
49. Zander, T.O.; Gaertner, M.; Kothe, C.; Vilimek, R. Combining Eye Gaze Input with a Brain–Computer Interface for Touchless Human–Computer Interaction. *Int. J. Hum.-Comput. Interact.* **2010**, *27*, 38–51. [[CrossRef](#)]
50. Matthies, D.J.; Strecker, B.A.; Urban, B. Earfieldsensing: A Novel in-Ear Electric Field Sensing to Enrich Wearable Gesture Input through Facial Expressions. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 1911–1922.
51. Fried, M.; Tsitsiashvili, E.; Bonne, Y.S.; Sterkin, A.; Wagnanski-Jaffe, T.; Epstein, T.; Polat, U. ADHD Subjects Fail to Suppress Eye Blinks and Microsaccades While Anticipating Visual Stimuli but Recover with Medication. *Vis. Res.* **2014**, *101*, 62–72. [[CrossRef](#)] [[PubMed](#)]
52. Schmitt, L.M.; Cook, E.H.; Sweeney, J.A.; Mosconi, M.W. Saccadic Eye Movement Abnormalities in Autism Spectrum Disorder Indicate Dysfunctions in Cerebellum and Brainstem. *Mol. Autism* **2014**, *5*, 1–13. [[CrossRef](#)] [[PubMed](#)]
53. Seiple, W.; Rosen, R.B.; Garcia, P.M. Abnormal Fixation in Individuals with Age-Related Macular Degeneration When Viewing an Image of a Face. *Optom. Vis. Sci.* **2013**, *90*, 45–56. [[CrossRef](#)] [[PubMed](#)]
54. Vine, S.J.; Wilson, M.R. The Influence of Quiet Eye Training and Pressure on Attention and Visuo-Motor Control. *Acta Psychol.* **2011**, *136*, 340–346. [[CrossRef](#)] [[PubMed](#)]