CrossMark

Fuzzy-based estimation of continuous Z-distances and discrete directions of home appliances for NIR camera-based gaze tracking system

Jae Woong Jang[1] · Hwan Heo[1] · Jae Won Bang[1] ·
Hyung Gil Hong[1] · Rizwan Ali Naqvi[1] ·
Phong Ha Nguyen[1] · Dat Tien Nguyen[1] ·
Min Beom Lee[1] · Kang Ryoung Park[1]

Abstract With the development of eye gaze tracking technology, much research has been performed to adopt this technology for interfacing with home appliances by severely disabled and wheelchair-bound users. For this purpose, two cameras are usually required, one for calculating the gaze position of the user, and the other for detecting and recognizing the home

✉ Kang Ryoung Park
   parkgr@dgu.edu

   Jae Woong Jang
   toshinya@dongguk.edu

   Hwan Heo
   gjghks@dongguk.edu

   Jae Won Bang
   bangjw@dgu.edu

   Hyung Gil Hong
   hell@dongguk.edu

   Rizwan Ali Naqvi
   rizwanali@dongguk.edu

   Phong Ha Nguyen
   phongnhhn92@gmail.com

   Dat Tien Nguyen
   nguyentiendat@dongguk.edu

   Min Beom Lee
   mblee@dongguk.edu

[1] Division of Electronics and Electrical Engineering, Dongguk University, 30, Pildong-ro 1-gil, Jung-gu, Seoul 100-715, South Korea

🜨 Springer

appliance. In order to accurately calculate the gaze position on the home appliance that the user looks at, the Z-distance and direction of the home appliance from the user should be correctly measured. Therefore, stereo cameras or depth-measuring devices such as Kinect are necessary, but they have limitations such as the need for additional camera calibration, and low acquisition speed for two cameras or a large-size of Kinect device. To overcome this problem, we propose a new method for estimating the continuous Z-distances and discrete directions of home appliances using one (small-sized) near-infrared (NIR) web camera and a fuzzy system. Experimental results show that the proposed method can accurately estimate the Z-distances and directions to home appliances.

1 Introduction

With the rapid development of human–computer interaction, the quality of life of severely disabled people has been considerably improved. Previous research proposed computer assistive technology for users to interact with a computer using joysticks, switches, head pointers, trackballs, neural interfaces, and eye tracking devices, etc. [38, 44, 51]. The first four devices have been widely used as computer-assistive devices because they can be executed by hands, feet, chin, mouth (blowing), thumbs, palm, and head. In addition, they can be used for controlling a wheelchair or a mouse pointer [26, 38].

In previous studies [21, 22 38, 53], the head and face were tracked by computer vision technology, which enables the control of mouse cursors. However, these methods have disadvantages in that they cannot be used by the severely disabled, such as quadriplegics. That is because it is often the case that the severely disabled, such as quadriplegics cannot move even the head and face. To overcome this problem, the method based on various physiological signals such has been researched.

In previous research [43], they showed that various physiological signals such as electro-oculograms (EOGs), electroencephalograms (EEGs), and electromyograms (EMGs) can be used as the interface with patients with motor disabilities, e.g. people with high-level spinal cord injury or with locked-up syndrome. The EOG signal is the electrical one produced by the difference of potential between the retina and cornea, from 2 to 20 mV. However, it is usually represented in the range from 15 to 200 uV.

Sometimes people with severe motor disability, e.g. locked-in syndrome cannot control the activity of muscle for communication. In this case, brain computer interface (BCI) based on EEG signal can be used because even the people with severe motor disability can maintain cerebral and sensory functions intact. Although the accuracies by invasive method of measuring EEG signal, e.g. electrocorticogram (ECoG) are higher than those by non-invasive method, the former has the disadvantages of involving clinical risks. Therefore, non-invasive method has been highlighted in recent BCI researches.

In case that the brain commands a muscle to contract, signals are known as being sent to motor neurons which control some fibers. As the membrane fiber is depolarized, it is reported that the electrical potential is produced in the neighborhood of the muscle fibers. The summation of the action potentials through the fibers

generates the motor unit action potential (MUAP) of a motor unit (MU). The ensemble of MUAP train is usually named as the EMG signal.

In previous study [8], they proposed a method of using EMG signals based on muscle responses as a computer interface for the upper limb disabled. They extracted the EMG signals from the muscles of lower arm, and multi-layered neural network (MLP) was trained in order to classify six patterns for controlling the movement and clicking a cursor. In addition, they designed on-screen keyboard in order to input Korean and Roman letters on the computer.

In other researches [3, 11], they proposed the interface systems using EOG signals, which can provide information of rough gaze direction based on eye movements. They designed the hardware circuits for measuring EOG signal including the eye-movement coils generating the voltages from 50 uV to 3500 uV with a frequency ranging from dc-100 Hz [11]. With this device, they also designed the graphic user interface (GUI) for controlling TV or game. In previous research [3], Barea et al. proposed the method of solving the problem that the detection accuracy of EOG-based eye movement degrades due to user's tiredness caused by using the system for long period. For that, they remove the noises and obtain more accurate result from EOG data based on wavelet transform and radial basis function (RBF) neural network.

However, the cost of system for measuring these kinds of physiological signals is not cheap, and the user's convenience is low because of the attachment of sensors on user's skin for the measurement of long period. In addition, the attachment of sensors can cause the errors in the measured signals in case of the movement of user's head, face, or body.

In order to solve these problems, camera-based gaze tracking methods have been widely researched as an alternative [45]. In previous research [12], authors addressed the dynamic and active nature of biometrics, in general, and gaze analysis, in particular, including background and motivation. They introduces a critical survey of conventional gaze analysis methods, challenges caused by uncontrolled environments and lack of standards, and explains promising R&D directions. In addition, they propose criteria for performance evaluation, and compares state-of-the art gaze analysis methods with the same database. In [2], they introduce a state-of-the art survey for head movement detection and eye tracking methods proposed in the literature. They also explain various examples of different fields of applications for these technologies, such as driving assistance systems, human computer interaction, and assistive technologies.

Most previous studies on gaze tracking were performed in the environment of 2D monitors [17, 23, 36]. In [36], the authors introduced the previous researches on eye tracking-based human computer interface system including its features, functionality and methods in addition to the challenges included in using the eye as an input modality, and real life applications. Lankford proposed the eye-based typing method based on the eye-gaze response interface computer aid (ERICA) system made by the University of Virginia [23]. In order to enhance the accuracy of gaze clicking, they introduced the method of gaze clicking based on dwell time (including collapse stages) and zooming methodology of the screen where a user is looking. In addition, they showed the method of enhancing the performance of eye-typing based on collapse on the key of visual keyboard. In other study, they introduced various previous studies on eye movements as a user interface (UI) for analyzing the UI in terms of usability and control in human-computer dialogue [17]. However, most of these studies of gaze tracking in 2D monitors have disadvantages in that the change of Z-distance and the direction of the monitor from the user are limited after user calibration. In addition, in the case of a gaze tracking-based interface for the disabled, the user can look at a home appliance in the 3D world instead of a 2D monitor, where the change of Z-distance and direction of the home appliance from the user can be large after user calibration, which can degrade the accuracy of gaze tracking system.

Some studies examined gaze tracking for controlling in 3D space [14, 37, 41, 47, 54, 57]. Shi et al. proposed a gaze tracking method using two cameras, one for eye tracking and the other for frontal object recognition. However, they did not consider the change of Z-distance and direction of the frontal object from the user after user calibration, which can cause the degradation of accuracy of gaze detection [47]. Other studies proposed gaze tracking using a head-mounted display (HMD) or wearable device [37, 41, 54]. In [41], authors used an HMD system with one camera for gaze tacking and one for frontal viewing, in an augmented reality application. Mardanbegi et al. also proposed a wearable device including two cameras for eye tracking and frontal viewing [37]. Turner et al. proposed a wearable device including one camera for gaze tracking and the two cameras for frontal viewing [54]. Using two frontal viewing cameras can increase the viewing range of the user irrespective of the natural movements of user's head. In [57], authors proposed a wearable eye-gaze tracking system for observational purposes on flight decks. In their research, because a pilot uses this system while sitting in the cockpit, the viewing angle and the Z-distance range from user to object are limited. In other research [14], the authors proposed a wearable device including two cameras, one for eye gaze detection, and the other for the recognition of frontal objects and gestures. However, all of these methods require the user to wear an HMD or wearable device, but wearing these devices for a long time can cause much inconvenience or cybersickness to the user, especially the disabled. In addition, these studies did not consider the change of Z-distance and direction of the frontal object from the user after user calibration, which can cause the degradation of accuracy of gaze detection.

In order to estimate the Z-distance of the object [55], stereo cameras based on homography matrix [7, 16, 25, 39, 40] or depth-measuring devices, such as ultrasonic sensors or Kinect devices [19] have been widely used. However, they have limitations, such as the need for additional calibration between two cameras and low acquisition speed of images by two stereo cameras [7, 25, 39, 40]. Because one camera for eye tracking should be used in our research, if two stereo cameras are used for frontal viewing, the total number of cameras becomes three. Therefore, the image capturing speed can be much reduced. In addition, in our research, each people can see the TV or heater at the Z distance of maximum 3 m from people to TV or heater. In order to measure the accurate Z distance in this case, the inter-distance between two cameras should be large, which can increase the size of system. The measurement by ultrasonic sensors has the limitation that inaccurate measurement can be caused by other obstacles which exists between the people and TV or heater. In case of using Kinect devices, the size of our system increases due to the large size of the Kinect device. Although Joglekar et al. proposed a method of estimating Z-distance based on single camera, their method requires additional information about the road geometry, including vanishing point and the point of contact on the road [18], which cannot be obtained indoors where our interface system for the disabled is applied.

Although they did not deal with the method of gaze detection, our fuzzy system-based method can be compared to previous machine learning-based approaches. In previous research [30, 31], they proposed the method for forecasting the water quality of a station over the next few hours, based on multitask multi-view learning method to combine multiple datasets from different domains. In details, in [30], their learning model is composed of two alignments such as the spatio-temporal view alignment, which fuse local spatial and temporal information of each station, and the prediction alignment among stations, which acquires their spatial correlations and performs co-predictions by incorporating these correlations. In [31], they also proposed the method of forecasting the water quality of a station from a data-driven perspective, using the water quality data and water hydraulic data reported by existing monitor stations and various data

sources observed in the city, such as pipe networks, meteorology, point of interests (POIs), and structure of road networks. Then, they identify the influential factors that have effect on the urban water quality through extensive experiments, and used a multi-task multi-view learning method to fuse those multiple datasets from different domains into an unified learning model. In other research [32], they proposed the method of scientifically and systematically studying the feasibility of career path prediction from social network data by seamlessly combining information from multiple social networks to comprehensively describe a user and characterize progressive properties of his or her career path. For that, they used a multi-source learning framework with fused lasso penalty, which jointly regularizes the career-stage relatedness and source.

By comparing gaze tracking of our research, in previous research [28], they proposed the method for generic human motion tracking by using both low- and high-dimensional trackers. Annealed Particle Filtering (APF) shows its advantage in the tracking without limitation on motion types. In addition, Gaussian Process Dynamical Model (GPDM) is trained to learn the motion model of low-dimensional tracker, which performs better on specific types of motion. They combined these two methods and automatically sample trackers according to the motion types. In [10], they also proposed a fusion formulation which combines low- and high-dimensional tracking methods into one framework. The low-dimensional method successfully overcomes the high-dimensional problem of tracking the motions with available training data by learning motion models, but it only works with specific kinds of motion types. On the other hand, although the high-dimensional approach can recover the motions without learned models by sampling directly in the pose space, it cannot guarantee robustness and efficiency. Therefore, they used within the framework, the two parallel approaches, low- and high-dimensional, which are fused via a probabilistic approach at each time step. All of these previous researches show high performance, but they require the procedure of learning with training data whereas our fuzzy system-based estimation (of Z-distance and direction of home appliance) and gaze tracking method do not require any training procedure. This is the difference between our research and previous ones [10, 28, 30–32].

By comparing NIR patterns-based method for measuring the Z-distance and direction in our research, in previous research [29], they proposed the method for complicated activity recognition being composed of two components such as temporal pattern mining, which supplies a mid-level feature representation for activities, encodes temporal relatedness among actions, and acquires the intrinsic properties of activities, and adaptive multi-task learning, which acquires relatedness. In [33], they also proposed the method to identify temporal patterns among actions and utilize the identified patterns to represent activities for automated recognition based on machine learning method using support vector machine (SVM) or k-nearest neighborhood (kNN) classifier. In [34], they defined an atomic activity-based probabilistic framework that uses Allen's interval relations to show local temporal dependencies. The framework introduces a latent variable from the Chinese Restaurant Process to explicitly characterize these unique internal configurations of a particular complex activity as a variable number of tables. However, their method also performs the training procedure to learn the parameters. In [27], they proposed a new method to retrieve the Kid's activities from video sequences captured in smart home environment. They showed the retrieval method with two steps, first using spatio-temporal matching to get a coarse result, second using the method to learn dominant child-adult interactive behaviors from a sequence of home videos based on K-means clustering and scale-invariant feature transform (SIFT) matching. In [35], authors proposed the unsupervised method to recognize physical activities by smartphone accelerometers. In their method, features extraction is performed with the raw acceleration data gathered

by smartphone. With these features, physical activity is recognized by unsupervised classification method called molecular complex detection (MCODE) method. Different from these researches [27, 29, 33–35], our NIR patterns-based method for measuring the Z-distance and direction does not require any training procedure. This is the difference between our research and previous ones [27, 29, 33–35].

Considering these factors, we propose a new method of estimation of Z-distances and directions of home appliances using one (small-sized) web camera based on fuzzy system. Our research is novel in the following four ways compared to previous works:

– First, the NIR patterns (attached on the boundary of the home appliance) captured by the frontal viewing camera are used for measuring the Z-distance and direction of the home appliance from the user.
– Second, the size and average of four lengths of sides of a quadrangle defined by the NIR patterns are used as two features for a fuzzy system to estimate the Z-distance.
– Third, the difference between the sum of two obtuse angles and that of two acute angles of the quadrangle, and the difference between the two y positions of left- and right-top NIR patterns are used as two features for the fuzzy system to estimate the direction.
– Fourth, the accuracy of user's gaze detection on the home appliance is enhanced by compensating the size and symmetrical shape of the quadrangle area (defined by the NIR patterns) by offset according to the Z-distance and direction measured by the fuzzy system.

The structure of this article is as follows. The proposed system and our methodology are presented in Section 2. Section 3 describes the experimental setup and the results, and the conclusions and some ideas for future work are presented in Section 4.

2 Proposed estimation method of Z-distances and directions

2.1 Proposed system and method

Figure 1 shows examples of using the proposed system. NIR-LEDs attached to the corners and edges of a home appliance can be captured by the frontal viewing camera where the NIR cutting filter is replaced with a NIR passing filter. The NIR-LED wavelength of 850 nm is not dazzling to the user's eye. According to various positions of NIR-LEDs, various kinds of home appliances, such as a TV, heater, and switch, can be recognized by the frontal viewing camera. The additional eye tracking camera, including a NIR passing filter with an 8×8 NIR-LED illuminator, is used for measuring the user's gaze position in the 3D world. These two cameras and illuminator are attached to the wheelchair. By using NIR cameras with NIR-LEDs and a NIR-LED illuminator, the performance of gaze tracking and recognizing the home appliance is not affected by changes in environmental visible light.

Figure 2 shows the summarized steps of our method. As the first and second steps, one image is captured by frontal viewing camera, and one by the eye tracking camera. Then, the four features are extracted from the NIR-LED pattern in the image by frontal viewing camera. With these features and a fuzzy system, the Z-distance and direction from the user to the home appliance are estimated. With the image from the eye tracking camera, the user's gaze position is calculated, and the system determines whether the user is looking at the home appliance on the basis of the Z-distance and direction with the user's gaze position.
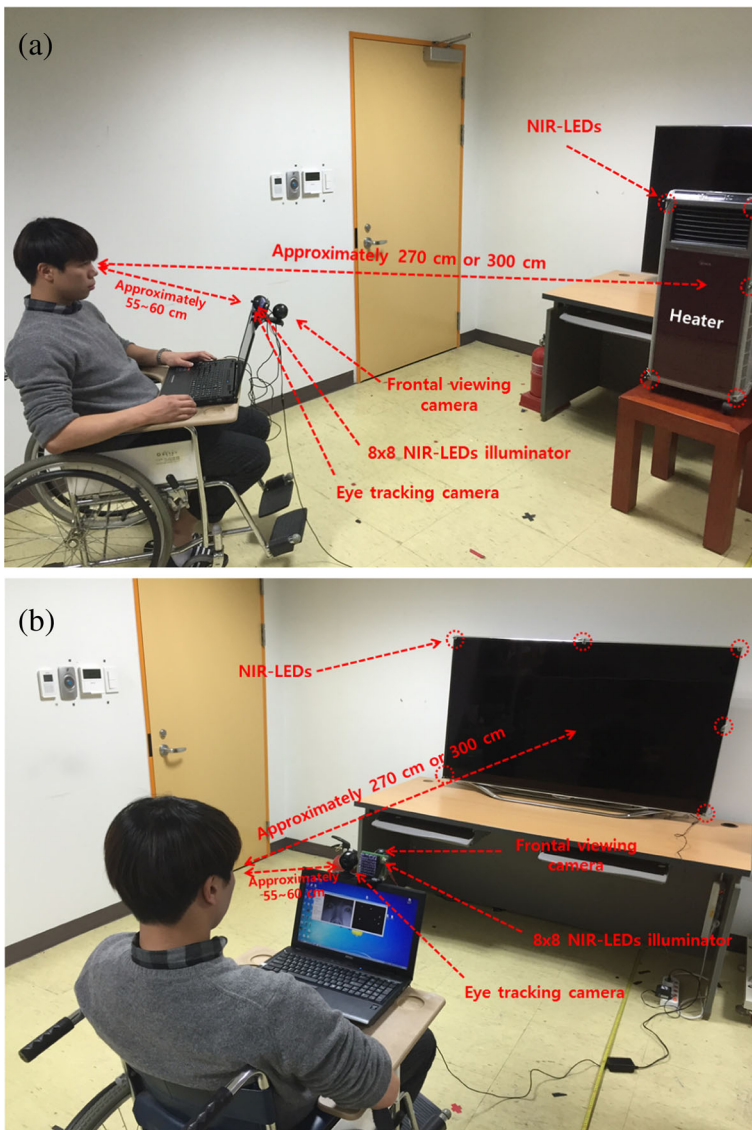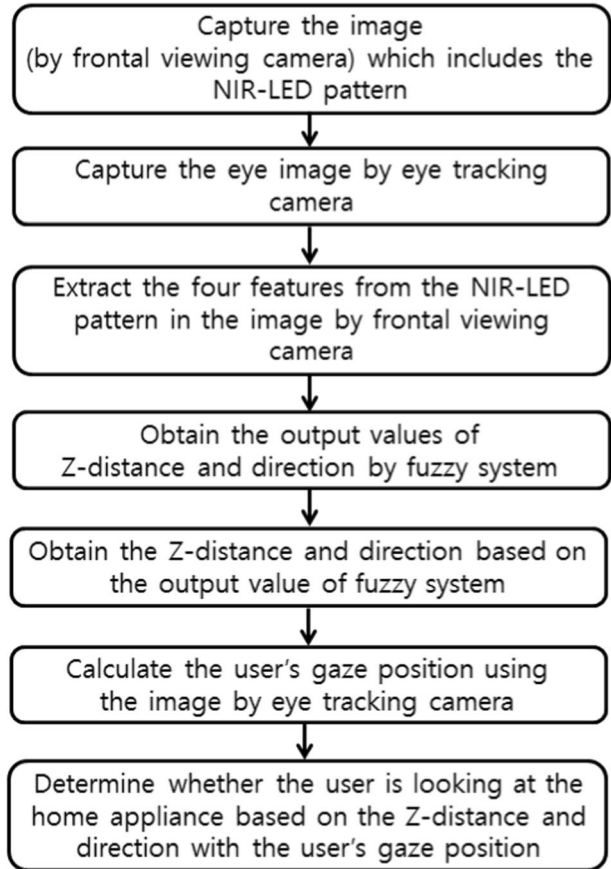
Fig. 1 Example of using the proposed system: cases of (**a**) looking at a heater, (**b**) looking at a TV

2.2 Capturing the images of the NIR-LED patterns of home appliances by the frontal viewing camera

The experimental setup of our research is shown in Fig. 3. The Z-distances of the TV and heater from the user are 270 cm and 300 cm, respectively, whereas those of the switch from user are 170 cm and 200 cm, respectively. The reason the Z-distances from the switch are smaller than those from the TV and heater is that the switch is smaller than the TV or heater. We performed the experiments in the three directions of right, center, and left of the user's position relative to the TV, heater, and switch, as shown in Fig. 3. That is, the experiments

Fig. 2 Summarized steps of the proposed method

Capture the image (by frontal viewing camera) which includes the NIR-LED pattern

↓

Capture the eye image by eye tracking camera

↓

Extract the four features from the NIR-LED pattern in the image by frontal viewing camera

↓

Obtain the output values of Z-distance and direction by fuzzy system

↓

Obtain the Z-distance and direction based on the output value of fuzzy system

↓

Calculate the user's gaze position using the image by eye tracking camera

↓

Determine whether the user is looking at the home appliance based on the Z-distance and direction with the user's gaze position

were performed at three viewpoints of the frontal viewing camera as shown in Fig. 3. In our research, we define the case that user is at right position based on home appliance as shown in Fig. 3(c) as left viewpoint because the home appliance is positioned at left side based on the frontal viewing camera of our system.

As explained in Section 2.1, each home appliance has different patterns of NIR-LEDs. Examples of images of the home appliances, including the NIR-LED patterns captured by the frontal viewing camera, are shown in Fig. 4.

With the captured images of Fig. 4, the features for estimating the direction (viewpoint) and Z-distance of the user from the home appliance are extracted. Detailed explanations are as follows.

The quadrangular shape of the NIR-LED patterns becomes smaller according to the increase in Z-distance of the user from the home appliance, because of the principle of camera perspective projection. On the basis of these characteristics, we define the features for estimating the Z-distance of the user from the home appliance. As the two features for estimating the Z-distance, we use the size and average of four lengths (L1 ~ L4 of Fig. 5(b)) of the sides of the quadrangle defined by the NIR patterns. The farther the Z-distance is, the smaller these two features become.
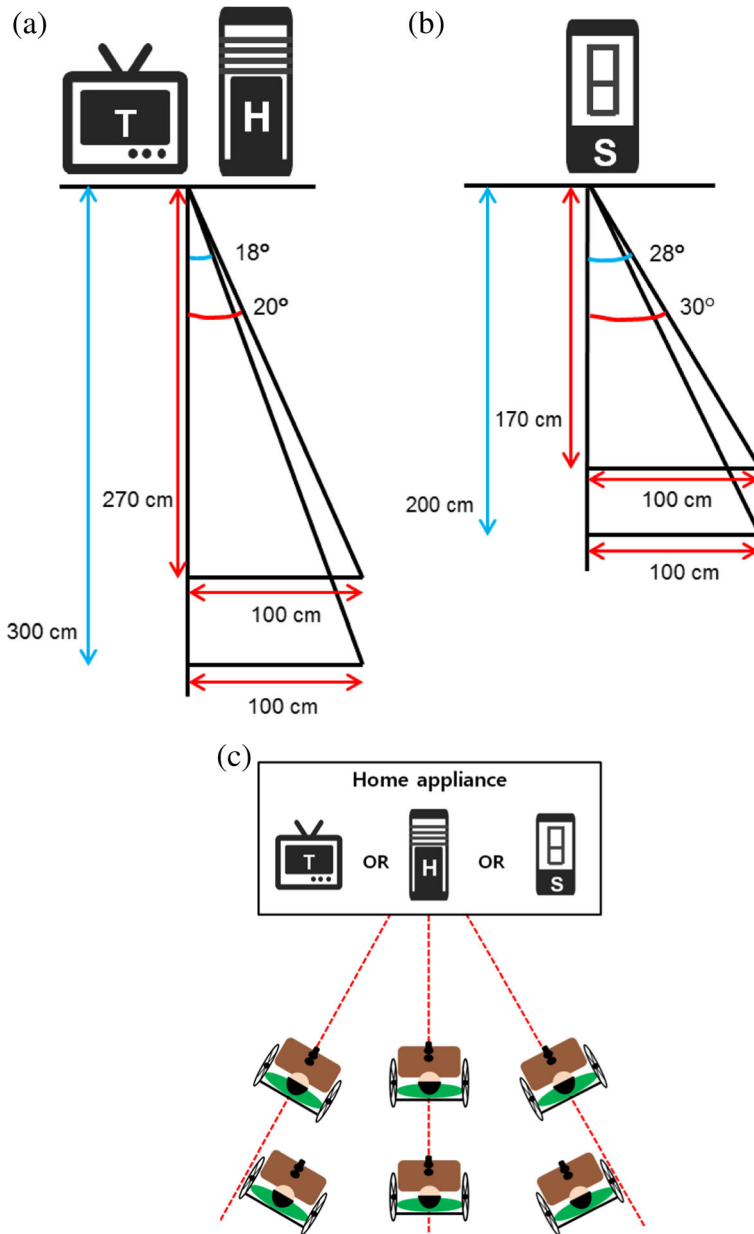
Fig. 3 Experimental setup of (a) the TV and heater, and (b) switch. (c) Top view of experimental setup

In addition, as shown in Fig. 5, the symmetrical quadrangular shapes of the NIR-LED patterns are different according to three different viewpoints of the frontal viewing camera. For example, as shown in the left image of Fig. 5(a) in the case of right viewpoint of frontal viewing camera (the user's position in the right direction), the upper-left corner point of quadrangular shape is higher than the upper-right corner point. In addition, there are differences between the sum of the two obtuse angles and that of the two acute angles of the
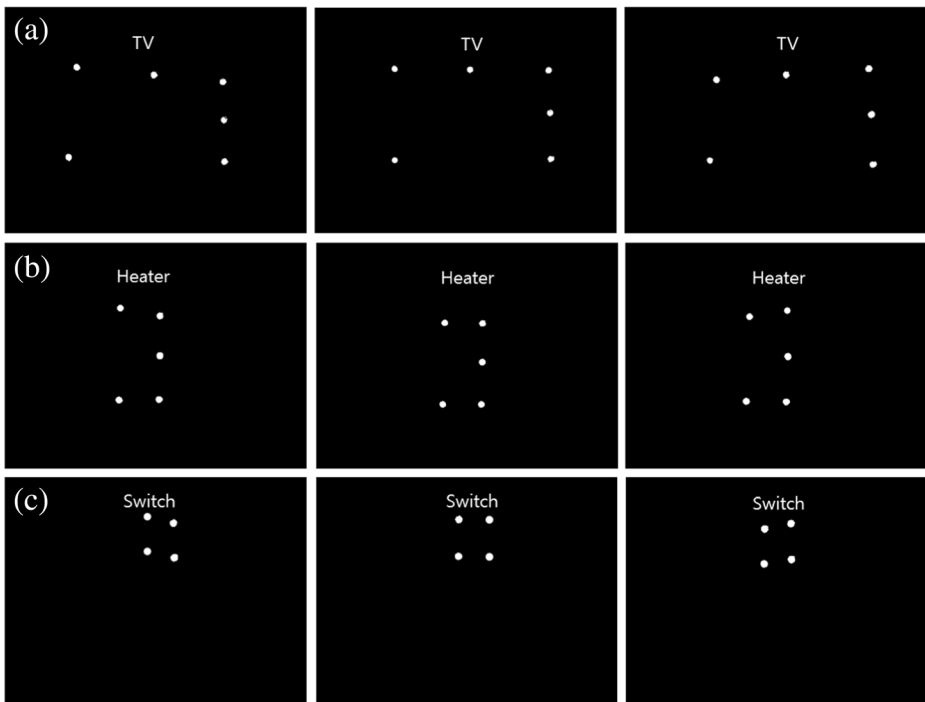
Fig. 4 Captured images of (**a**) TV, (**b**) heater, and (**c**) switch at three different viewpoints of the frontal viewing camera. Left, center, and right images represent the images captured from the right, center, and left viewpoints, respectively

symmetrical quadrangle. For example, in case of the left image of Fig. 5(a), the sum of angles of R1 and R4 is smaller than that of R2 and R3. On the basis of these, we use two features for estimating the direction of the user (viewpoint of the frontal viewing camera) as follows. The first feature is the difference between the two y positions of left- and right-top NIR-LED patterns. The second feature is the difference between the sum of the two obtuse angles and that of the two acute angles of the quadrangle.

Using these features, we estimate the Z-distance and direction using a fuzzy system. Detailed explanations are presented in the next section.

2.3 Fuzzy-based estimation of Z-distances and directions

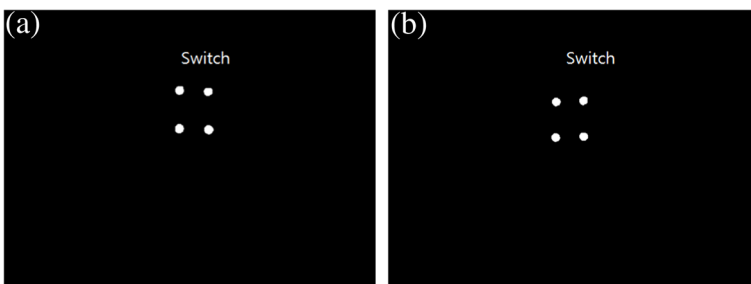#### 2.3.1 Definition of fuzzy membership functions

In some cases, it is difficult to accurately estimate the Z-distance and direction of a home appliance from the user only with the NIR-LED pattern of the image by the frontal viewing camera. The NIR-LED pattern in the case of the left position of the user (Fig. 6(a)) looks similar to that of the right position of the user (Fig. 6(b)), which makes it difficult to distinguish the direction of the home appliance. To overcome this problem, we adopt a fuzzy logic system. As explained in Section 2.2, we use two features for estimating the Z-distance, including the size (F_1_dist) and average (F_2_dist) of the four lengths (L1 ~ L4 of Fig. 5(b)) of the quadrangle sides defined by the NIR patterns. In addition, we use two features for estimating the direction
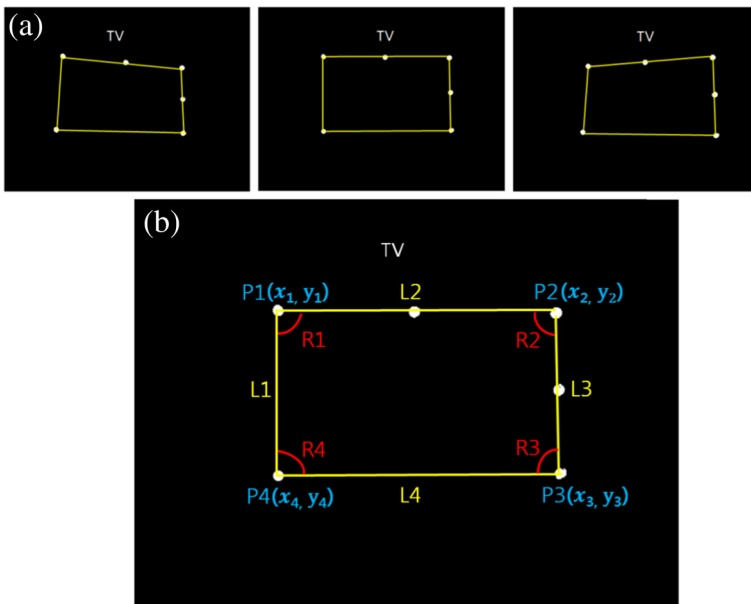
Fig. 5 Symmetrical quadrangular shapes of NIR-LED patterns of the TV, and features for estimating the direction (viewpoint) and Z-distance of the user from the home appliance: **a** Shapes of NIR-LED patterns at three viewpoints (right, center, and left, respectively) at the Z-distance of 270 cm. **b** Extracted lengths of L1, … L4 and angles of R1, … R4 from the captured image

of the user, including the difference (F_1_direct) between the two y positions of the left- and right-top NIR-LED patterns, and the difference (F_2_direct) between the sum of the two obtuse angles and that of the two acute angles of the quadrangle. Because the ranges of the four features are different from each other, each feature value is represented in the range from 0 to 1 by Z-score normalization [49], and they are used as the inputs to the fuzzy logic system [1, 20, 52].

Figure 7 shows the fuzzy membership functions used in this study. We use the linear (triangular) function because it has been widely adopted in fuzzy systems considering its fast processing speed and the problem complexity [4, 5, 58]. In Fig. 7(a), Features 1 and 2 represent F_1_dist and F_2_dist in the case of using the fuzzy system for estimating the Z-distance of the home appliance from the user. Features 1 and 2 represent F_1_direct and F_2_direct in the case of using the fuzzy system for estimating the direction of the home appliance from the user. As shown in Fig. 7, the low (L), middle (M), and high (H) areas are

Fig. 6 Example case where the direction of the home appliance from user is difficult to distinguish: **a** right direction and **b** left direction
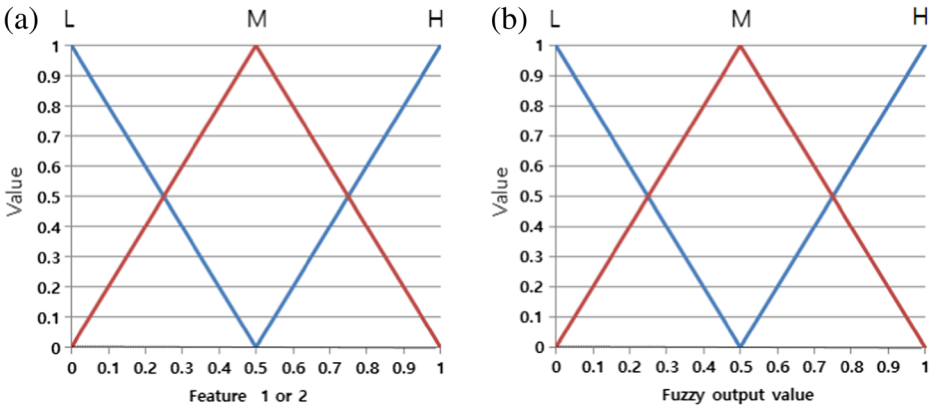
Fig. 7 Illustrations showing the input and output fuzzy membership functions: **a** input fuzzy membership function and **b** output fuzzy membership function

usually designed to be overlapped. We use these membership functions and fuzzy rules table (Table 1) to obtain an output value of the fuzzy system through the defuzzification method.

2.3.2 Designing the fuzzy tables and obtaining the output by the defuzzification method
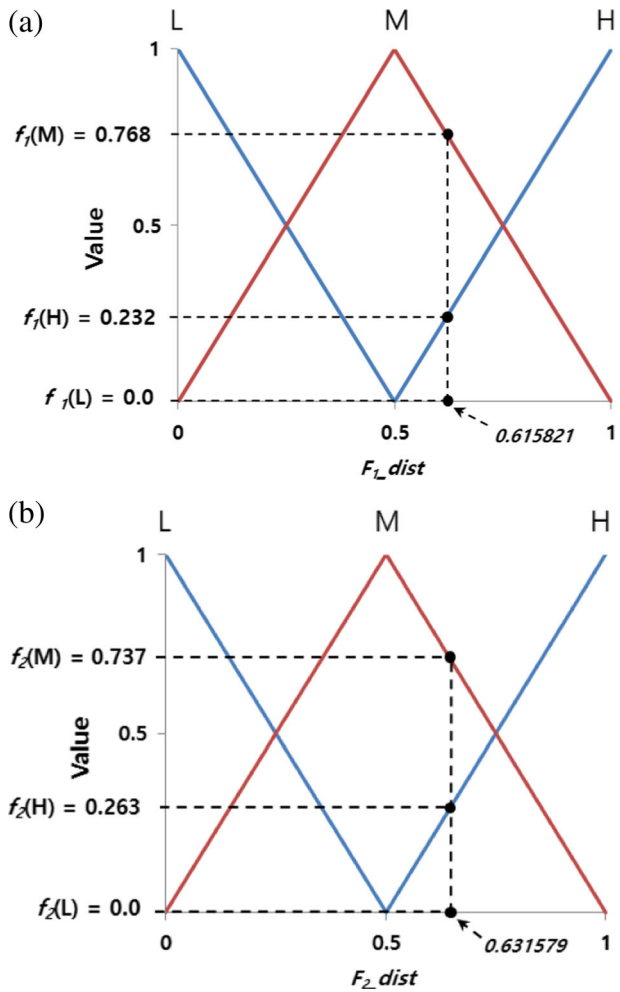
In Table 1, we design the fuzzy rule tables. Table 1(a) is used for estimating the Z-distance, and Table 1(b) is for estimating the direction of the home appliance from the user. Based on the 5th rule (if "M" and "M", then "M"), the 1st ~ 4th rules are symmetrical to the 6th ~ 9th ones if exchanging L and H, each other as shown in Tables 1(a) and (b).

Lower values of F_1_dist and F_2_dist represent a farther Z-distance of the home appliance from the user. Therefore, we design the output of the fuzzy system as H in cases of low values

Table 1 The fuzzy rule tables (Low (L), Middle (M), High (H)). Rule tables for estimating (a) the Z-distance and (b) direction

Input 1 (F_1_dist)	Input 2 (F_2_dist)	Output
L	L	H
	M	H
	H	M
M	L	H
	M	M
	H	L
H	L	M
	M	L
	H	L
(a)		
L	L	M
	M	H
	H	H
M	L	L
	M	M
	H	H
H	L	L
	M	L
	H	M
(b)		

Fig. 8 Illustrations of obtaining the output from the input membership function: **a** outputs by F_1_dist and **b** outputs by F_2_dist

of F_1_dist and F_2_dist as shown in Table 1(a). Higher values of F_1_dist and F_2_dist represent a nearer Z-distance of the home appliance from the user. Therefore, we design the output of the fuzzy system as L in cases of high values of F_1_dist and F_2_dist as shown in Table 1(a). As with this method, the fuzzy rules of Table 1(b) are also designed.

Then, we introduce the procedure of obtaining output value of fuzzy system as follows. Figure 8 shows the conventional procedure of obtaining the fuzzy inference value from input feature based on input membership function whereas Fig. 9 shows the conventional procedure (defuzzification) of obtaining the final output value of fuzzy system from the fuzzy inference value. These has been widely used in conventional fuzzy system [4–6, 24, 46, 48, 58].

Three outputs of $f_1(L), f_1(M)$, and $f_1(H)$ are obtained from one input feature value (F_1_dist) by three input membership functions of L, M and H, respectively, as shown in Fig. 8(a). In addition, three outputs of $f_2(L), f_2(M)$, and $f_2(H)$ are obtained from other input feature value (F_2_dist) by three input membership functions of L, M, and H, respectively, as shown in Fig. 8(b). For example, if the F_1_dist is 0.615821, the $f_1(L), f_1(M)$, and $f_1(H)$ are 0.0, 0.768,

Fig. 9 Illustrations of obtaining the final output by the defuzzification method: **a** FOM, LOM, MOM, and MeOM; **b** COG
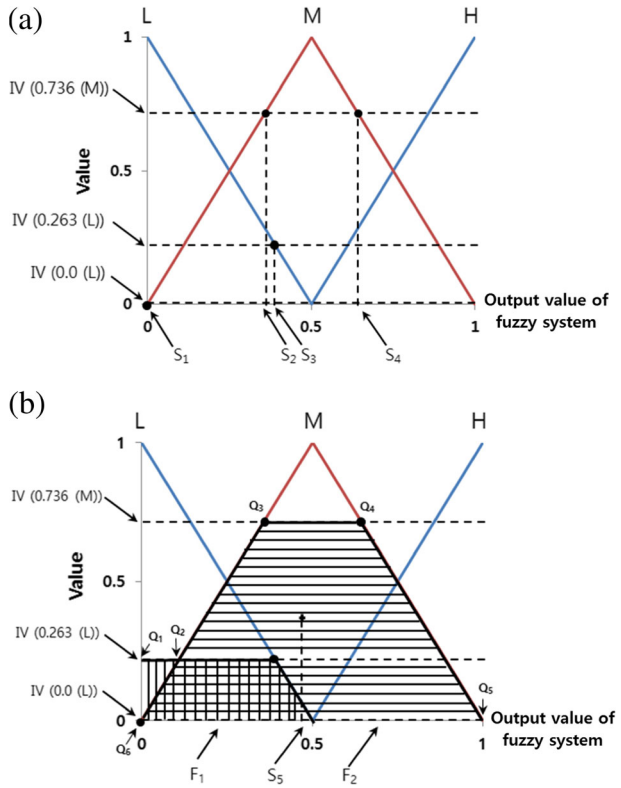
and 0.232, respectively, as shown in Fig. 8(a). If the F_2_dist is 0.631579, the f_2(L), f_2(M), and f_2(H) are 0.0, 0.737, and 0.263, respectively, as shown in Fig. 8(b).

With these two pairs of three outputs, we can obtain the nine combinations of output ((f_1(L), f_2(L)), (f_1(L), f_2(M)), (f_1(L), f_2(H)), (f_1(M), f_2(L)), (f_1(M), f_2(M)), (f_1(M), f_2(H)), (f_1(H), f_2(L)), (f_1(H), f_2(M)), (f_1(H), f_2(H))) as shown in Table 2. Then, based on the Min or Max rule with the fuzzy rule table of Table 1, we can obtain the output values as shown in Table 2. For example, in the 6th index of pairs, in the case of 0.768(M) and 0.263(H), from Table 1 we can obtain L as the output in the case of M and H of input. In addition, if using the Min rule, the smaller value of 0.263 is selected, whereas the larger value of 0.768 is selected in the case of using the

Table 2 Illustration of the nine combinations of output values of Fig. 8, and the selected ones by the Min or Max rule

Index of pairs	Output of $f_1(\cdot)$	Output of $f_2(\cdot)$	Min rule	Max rule
1	0.0(L)	0.0(L)	0.0(H)	0.0(H)
2	0.0(L)	0.737(M)	0.0(H)	0.737(H)
3	0.0(L)	0.263(H)	0.0(M)	0.263(M)
4	0.768(M)	0.0(L)	0.0(H)	0.768(H)
5	0.768(M)	0.737(M)	0.737(M)	0.768(M)
6	0.768(M)	0.263(H)	0.263(L)	0.768(L)
7	0.232(H)	0.0(L)	0.0(M)	0.232(M)
8	0.232(H)	0.737(M)	0.232(L)	0.737(L)
9	0.232(H)	0.263(H)	0.232(L)	0.263(L)

Max rule. Therefore, in the case of using the Min and Max rules, we can obtain 0.263(L) and 0.768(L), respectively, as shown in Table 2. We call these values of 0.263(L), and 0.768(L) as inference values (IVs). Then, through the defuzzification method with these IVs, we can obtain the output of the fuzzy system.

The detailed explanation of the defuzzification method is as follows. For example, with Fig. 9, three IVs are assumed to be 0.0(L), 0.263(L), 0.736(M). Using these three IVs, we can obtain the final output of the fuzzy system by using the defuzzification method. There are various kinds of defuzzification methods, such as first of maxima (FOM), last of maxima (LOM), middle of maxima (MOM), mean of maxima (MeOM), and center of gravity (COG) [6, 24]. The FOM method selects the first output value calculated by the maximum IV. In Fig. 9, we can obtain output (S_2) calculated by the maximum IV (0.736(M)) by the FOM method. LOM selects the last output value calculated by the maximum IV. We can obtain output (S_4) calculated by the maximum IV (0.736(M)) by the LOM method. MOM selects the middle of the first and last outputs calculated by the maximum IV. We can obtain the output (($S_2 + S_4$)/2) calculated by maximum IV (0.736(M)) by the MOM method. MeOM selects the mean of all the outputs calculated by the maximum IV. In Fig. 9, the output (($S_2 + S_4$)/2) obtained by MeOM is same to that by MOM because all the outputs (S_2 and S_4) calculated by the maximum IV are same as the first and last outputs (S_2 and S_4) calculated by the maximum IV. COG select S_5 (the geometrical center of the polygonal region defined by $Q_1 \sim Q_6$) as the output as shown in eq. (1) [6, 46, 48] and Fig. 9(b). In eq. (1), $\tilde{F}$ is the combined area of two regions (F_1 and F_2 of Fig. 9(b)). V and S are the variables of the values of vertical and horizontal axes, respectively.

$$S5 = \frac{\int V \, \tilde{F}(S) \cdot S \, dS}{\int V \, \tilde{F}(S) \, dS} \tag{1}$$

On the basis of the fuzzy system output value, the Z-distance and direction of the home appliance from the user are obtained. For example, in the case of the TV and heater, the output value in the range of 0–1 is linearly mapped into that in the range of 270–300 cm, whereas the value is linearly mapped into that of 170–200 cm in the case of the switch.

2.4 Detecting the user's gaze position based on the home appliance

On the basis of the measured Z-distance and direction, we can obtain accurate gaze position relative to the home appliance. At one Z-distance and direction (center) with each home appliance, each user performs a user-dependent calibration. For the calibration, the user gazes at the four (upper left, upper right, lower right, and lower left) corners of the home appliances. From this, four images are acquired. Within each image, the pupil center is detected by sub-block-based template matching, binarization, canny edge detection, and ellipse fitting. In addition, from the detected pupil center, the center of corneal specular reflection (SR) is located by binarization and calculation of geometric center position. Figure 10 shows examples of detected pupil centers and centers of SR.

With these four data of detected pupil centers and centers of SR, the position of the pupil center is compensated on the basis of the center of SR, which can reduce the variation of gaze position caused by head movement [15]. Because these four data of compensated pupil centers are obtained when a user looks at the four corners of the home appliance, we can obtain the
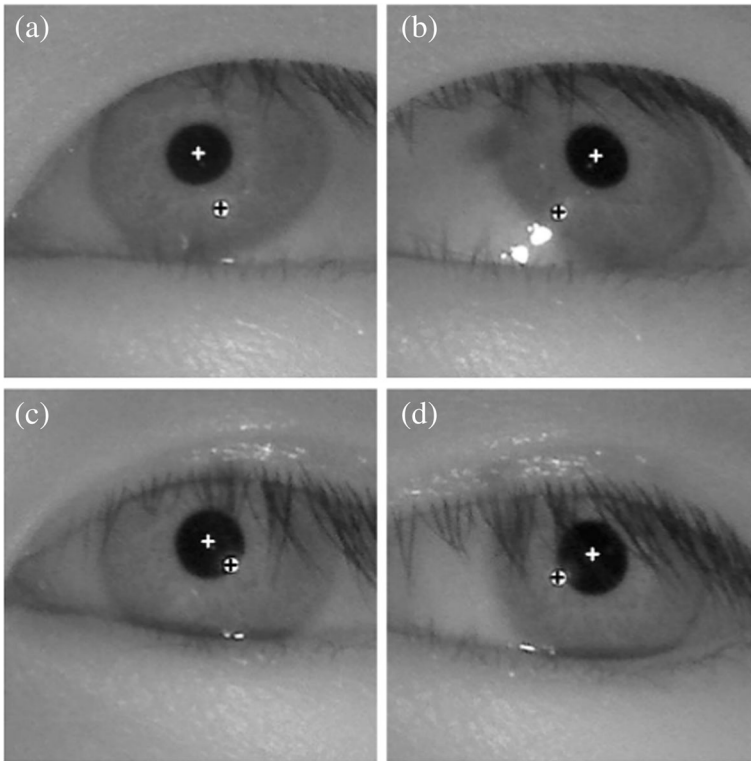
Fig. 10 Four images including the detected centers of pupil and corneal SR when a user is looking at the (**a**) upper left, (**b**) upper right, (**c**) lower left, and (**d**) lower right corners of a TV

geometric transform matrix, which defines the relationship between the pupil movable area and the region of the home appliance as shown in Fig. 11.

Then, we can obtain a geometric transform matrix like eq. (2), and the user's gaze position (GP_x, GP_y) can be obtained by eq. (3).

$$\begin{bmatrix} HA_{x0} \ HA_{x1} \ HA_{x2} \ HA_{x3} \\ HA_{y0} \ HA_{y1} \ HA_{y2} \ HA_{y3} \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \end{bmatrix} \begin{bmatrix} PC_{x0} & PC_{x1} & PC_{x2} & PC_{x3} \\ PC_{y0} & PC_{y1} & PC_{y2} & PC_{y3} \\ PC_{x0}PC_{y0} & PC_{x1}PC_{y1} & PC_{x2}PC_{y2} & PC_{x3}PC_{y3} \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (2)$$

$$\begin{bmatrix} GP_x \\ GP_y \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \end{bmatrix} \begin{bmatrix} PC'_x \\ PC'_y \\ PC'_x PC'_y \\ 1 \end{bmatrix} \quad (3)$$

However, the geometric transform matrix including the coefficients of a, b, … h of eqs. (2) and (3) is obtained when a user performs calibration at one Z-distance and direction. Therefore, if the Z-distance and direction are changed from those of initial calibration, the accuracy of the user's gaze position calculated by the geometric transform matrix of eqs. (2) and (3) becomes lower. To solve this problem, we use the Z-distance and direction measured by the fuzzy
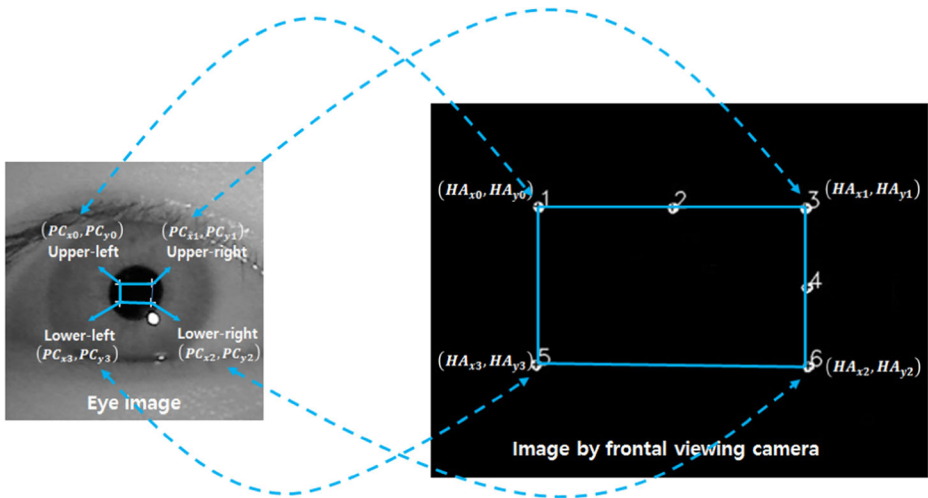
Fig. 11 Relationship between the pupil movable region (the quadrangle defined by (PC_{x0}, PC_{y0}), (PC_{x1}, PC_{y1}), (PC_{x2}, PC_{y2}), and (PC_{x3}, PC_{y3})) on the eye image and the recognized area of the home appliance (the quadrangle defined by (HA_{x0}, HA_{y0}), (HA_{x1}, HA_{y1}), (HA_{x2}, HA_{y2}), and (HA_{x3}, HA_{y3})) on the image by the frontal viewing camera

system (explained in Section 2.3). The area of the home appliance for user's gaze position (GP_x, GP_y) is compensated by offset according to the measured Z-distance and direction. For example, if a user gazes at the TV at 300 cm although the calibration was performed at 270 cm, the TV size seen by the user inevitably becomes smaller. Therefore, we redefine the TV area (the right image of Fig. 11) by using the smaller area when our system determines whether the user gazes at the TV. For example, if a user gazes at the TV at the right position although the calibration is performed at the center, the TV shape seen by the user inevitably becomes distorted as shown in the left image of Fig. 5(a). Therefore, we redefine the TV area (the right image of Fig. 11) by using the distorted area like the left image of Fig. 5(a) when our system determines whether the user gazes at the TV.

3 Experimental results

Two web cameras (Logitech C600 [56]) with the illuminator of 8×8 NIR-LEDs are used for implementing our system. Our algorithm is implemented using Microsoft visual studio 2012 C++, and OpenCV (version 2.3.9) library [42], which is executed on a laptop computer (Intel Core i5 at 2.8 GHz with 8 GB of memory). As the home appliance, a 60-in. TV, heater, and switch are used. The sizes of the images captured by the scene and eye tracking cameras are 640×480 pixels and 1600×1200 pixels, respectively. Experiments were performed by 20 people with three home appliances from three viewpoints (left, center, and right) and two Z-distances (270 and 300 cm in case of the TV and heater, 170 and 200 cm in case of the switch) as shown in Fig. 3. Each subject performed the experiments five times, and we captured 1800 scene images for the experiments. The percentage of scene images of users looking at the target is 50%, and that of users looking outside the target boundaries is also 50%. Among 20 participants, 18 were male and 2 were female, and none of them wore glasses

or contact lens. The average age (standard deviation) of the 20 participants is 27.4 (1.85), and all of them are able-bodied.

As explained in Section 2.3.1, we use two features for estimating the Z-distance, including the size (F_1_dist) and average (F_2_dist) of four lengths (L1 ~ L4 of Fig. 5(b)) of the sides of the quadrangles defined by the NIR patterns. In addition, we use two features for estimating the direction of the user, including the difference (F_1_direct) between the two y positions of left- and right-top NIR-LED patterns, and the difference (F_2_direct) between the sum of the two obtuse angles and that of the two acute angles of the quadrangle. Table 3 shows the calculated feature values according to the Z-distance and directions of each home appliance. From the Table 3, we can find that there are differences among the feature values according to the Z-distance and directions, and these feature values can be effective as inputs to fuzzy system.

With the feature values of Table 3, the output values by the fuzzy system are obtained as shown in Fig. 12. Figure 12 shows the examples of the outputs by fuzzy system with the COG method and Min rule. As shown in Fig. 12, we can find that all the outputs by the fuzzy system can be completely separated according to Z-distances and directions.

For the next experiments, according to various defuzzification methods such as FOM, LOM, MOM, MeOM, and COG with the Min and Max rules, we measured the correct recognition rate of the Z-distance and direction of the TV, heater, and switch. For example,

Table 3 Calculated feature values according to the Z-distance and the direction of each home appliance: (a) F_1_dist, (b) F_2_dist, (c) F_1_direct, and (d) F_2_direct

| Home appliance	Z-distance	Directions		
		Right	Center	Left
TV	270 cm	65,197	64,649	64,846
	300 cm	49,760	48,750	47,593
Heater	270 cm	14,824	13,621	14,189
	300 cm	11,466	11,112	10,632
Switch	170 cm	4936	4597	3814
	200 cm	3217	2976	2673
(a)				
TV	270 cm	267	267	266
	300 cm	233	231	227
Heater	270 cm	130	125	128
	300 cm	115	113	111
Switch	170 cm	71	68	63
	200 cm	57	55	52
(b)				
TV	270 cm	−34	−5	31
	300 cm	−29	−1	24
Heater	270 cm	−4	1	11
	300 cm	-4	1	8
Switch	170 cm	−5	2	8
	200 cm	−3	1	7
(c)				
TV	270 cm	6	−4	−15
	300 cm	6	-4	−14
Heater	270 cm	10	0	−10
	300 cm	6	0	−8
Switch	170 cm	4	0	−2
	200 cm	5	0	−5
(d)				
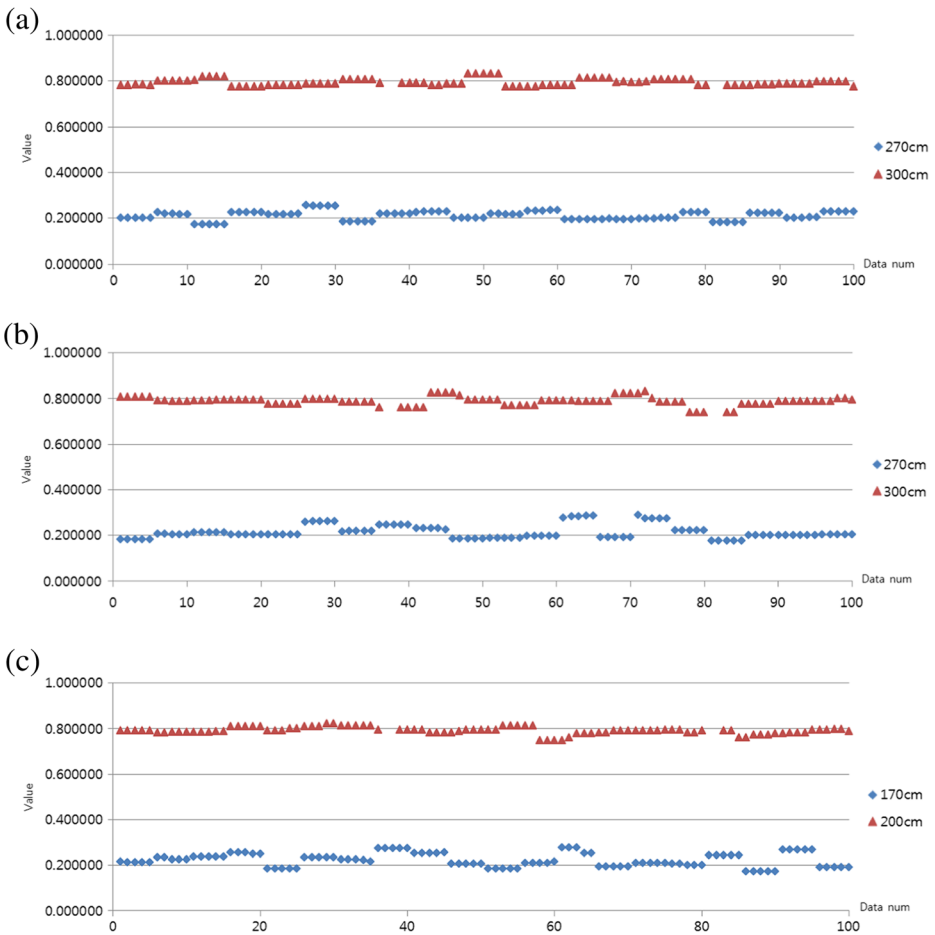
Fig. 12 Examples of outputs by the fuzzy system with the COG method and Min rule: **a–c** are the cases of using F_1_dist and F_2_dist; **d–f** are the cases of using F_1_direct and F_2_direct; **a** TV, **b** heater, **c** switch, **d** TV, **e** heater, and **f** switch

with the TV at the far Z-distance (300 cm) in the right direction, if our fuzzy-based system can correctly recognize this as a TV at the far Z-distance (300 cm) in the right direction, this case is counted as correct recognition. Experimental results showed that the average recognition rate by COG_MIN (COG with Min rule) is the best (100%) and that by FOM_MIN (FOM with Min rule) is the second-best (99.5%).

For the next tests, we measured the root mean square (RMS) errors of the Z-distance and direction estimated by our fuzzy-based system. From the fuzzy output of 0–1, we obtain the actual Z-distance and direction by linear mapping, and calculate the RMS error between the estimated and ground-truth values. Because the correct recognition rates of COG_MIN and FOM_MIN are the first- and second-best, respectively, we show the RMS errors by the COG_MIN and FOM_MIN as shown in Table 4. As shown in Table 4, we can find that the RMS error of estimating the Z-distance is less than approximately 2 cm and that of estimating direction is less than approximately 4°.

(d)

(e)

(f)
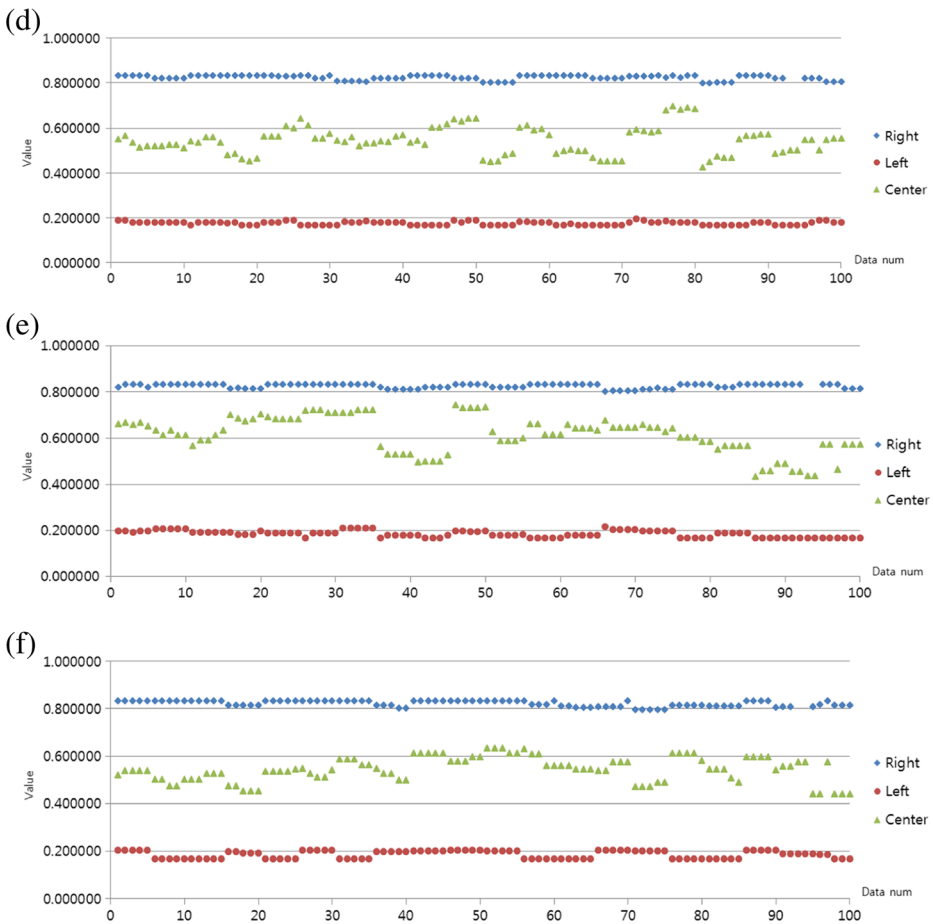
Fig. 12 continued.

For the next experiment, we measured the accuracies of gaze detection on home appliances in terms of true positive rate (TPR) and true negative rate (TNR) by proposed method and previous one [15]. TPR means the rate at which the calculated gaze position belongs to the area of the home appliance boundary when the user is actually looking at the home appliance. The TNR is that at which the calculated gaze position exists outside of the area of the home appliance boundary when the user is not actually looking at the home appliance. The same dataset made up of 1800 scene images (explained at the beginning of Section 3) is used for this experiment. The details of the employed dataset are explained at the beginning of Section 3.

The gaze detection with our method is performed by compensating the size and shape of the quadrangle area (defined by the NIR patterns as shown in the right image of Fig. 11) by offset according to the Z-distance and direction measured by fuzzy system as explained in Section 2.4. The gaze detection by previous method [15] is performed with geometric transform matrix for gaze detection without the estimation of Z-distance and direction. As shown in Table 5, the TPR and TNR with our method are higher than those by previous method, which shows the effectiveness of our method in gaze detection on home appliances.

Table 4 The RMS errors of the Z-distance and direction estimated by the fuzzy-based system: (a) RMS errors of the Z-distance, (b) RMS errors of the direction at the near Z-distance (TV and heater (270 cm), switch (170 cm)), (c) RMS errors of the direction at the far Z-distance (TV and heater (300 cm), switch (200 cm))

(unit: cm)

Home appliance	270 cm		300 cm			
	FOM_MIN	COG_MIN	FOM_MIN	COG_MIN		
TV	2.28	1.72	4.42	1.76		
Heater	3.10	1.73	4.24	2.04		
Switch	2.08	1.89	3.79	1.46		
(170, 200 cm)						

(a)

(unit: °)

	Left		Center		RIGHT	
	FOM_MIN	COG_MIN	FOM_MIN	COG_MIN	FOM_MIN	COG_MIN
TV	0.79	0.22	3.05	2.71	0.14	0.00
Heater	1.00	0.00	5.68	3.31	0.45	0.00
Switch	3.01	0.00	5.69	3.92	2.83	0.00

(b)

(unit: °)

	Left		Center		RIGHT	
	FOM_MIN	COG_MIN	FOM_MIN	COG_MIN	FOM_MIN	COG_MIN
TV	2.81	0.85	2.41	2.25	2.46	0.92
Heater	2.85	0.78	4.58	3.66	5.13	1.96
Switch	8.28	1.90	6.80	3.75	6.57	3.14

(c)

In order to prove that the average TPR and TNR with our method are statistically higher than that by previous method, a t-test is performed. The t-test is a statistical hypothesis test which is frequently used in statistical analysis [50]. When the t-test is performed using two independent samples, the average TPR by previous method and that with our method, the calculated p-value is 0.006, which is smaller than the 99% (0.01) significance level. When the t-test is performed using two independent samples, the average TNR by previous method and that with our method, the calculated p-value is 0.015, which is smaller than the 95% (0.05) significance level. The null-hypothesis for the t-test is determined that there is no difference between two TPRs or two TNRs, and this null-hypothesis is rejected on the basis of the p-value. Therefore, we can conclude that there is a significant difference in the average TPR and TNR by previous method [15] and with ours.

For the next analysis, Cohen's d analysis is performed, through which the size of the difference between the two groups can be shown on the basis of effect size [9]. Cohen's d is usually determined as small at approximately 0.2 ~ 0.3, medium at approximately 0.5, and large at or higher than 0.8. In Table 6, the calculated Cohen's d is 1.728 in the case of TPR, and

Table 5 TPR and TNR with our method and by previous method (unit: %)

Home appliance	TPR/TNR	Previous method [15]	Proposed method
TV	TPR	97.22	99.29
	TNR	83.33	99.38
Heater	TPR	62.50	99.75
	TNR	84.72	96.50
Switch	TPR	81.94	98.92
	TNR	88.89	99.17

Table 6 *P*-value, Cohen's d, and effect size calculated between the average TPR and TNR by previous method and with our method

	P-value	Cohen's d	Effect size
TPR with our method vs. by previous method [15]	0.006 (smaller than the 99% (0.01) significance level)	1.728	Large
TNR with our method vs. by previous method [15]	0.015 (smaller than the 95% (0.05) significance level)	1.455	Large

we can conclude that the difference between the average TPR by previous method and with our method is large. In addition, in Table 6, the calculated Cohen's d is 1.455 in the case of TNR, and we can conclude that the difference between the average TNR by previous method and with our method is large. From the *p*-value and Cohen's d, we can conclude that the average TPR and TNR with our method are higher than those by previous method [15] with a significant difference.

As the last experiment, we measured the processing speed of our method. As shown in Table 7, we found that our system can be operated at a fast speed (about 67.2 (1000/14.887) frames/s).

In our system, the home appliance in the image by frontal viewing camera can be recognized based on the bright spots of NIR pattern. Therefore, even in the case that multiple home appliances are close to each other, all of them can be correctly recognized and the positions of each quadrangle defined by NIR patterns can be detected as shown in Fig. 14.

With the image, the positions of all the bright spots can be located by image binarization and component labeling [13]. In order to group the bright spots according to the kind of home appliance, the following method is used. Two bright spots which are the closest to the left-upper corner (0, 0) and left-lower corner (0, 479) of the image is firstly located, respectively. The image size by frontal viewing camera is 640 × 480 pixels in our system. In Fig. 14, these two bright spots are the left-upper and left-lower corners of heater, respectively. Then, based on these two bright spots, two searching regions are respectively defined at the right positions of these bright spots, and the positions of third and fourth bright spots are also located. From that, the first quadrangle defined by four bright spots is located. Then, except for the four bright spots of the first quadrangle, same procedure is repeated in order to detect the second or third quadrangle. After detecting all the quadrangles in the image, with the four bright spots of each quadrangle, our system checks whether there exists additional bright spot between two corner bright spots of the quadrangle. Based on this information, our system can recognize the kind of home appliance.

Through this procedure, our system can correctly recognize multiple home appliances close to each other, and detect the positions of each quadrangle defined by NIR patterns as shown in Fig. 14. Then, based on the methods of Sections 2.2 and 2.3, the Z-distance and direction from user to the home appliances are estimated. All the Z-distances and directions of Fig. 14 are correctly recognized by our method.

Table 7 Processing speed of our method (unit: ms)

Estimating the Z-distance	Estimating the direction	Calculating gaze position	Total
1.012	1.313	12.562	14.887

Fig. 13 T-test of average TPR and TNR by previous and proposed methods. **a** TPR. **b** TNR
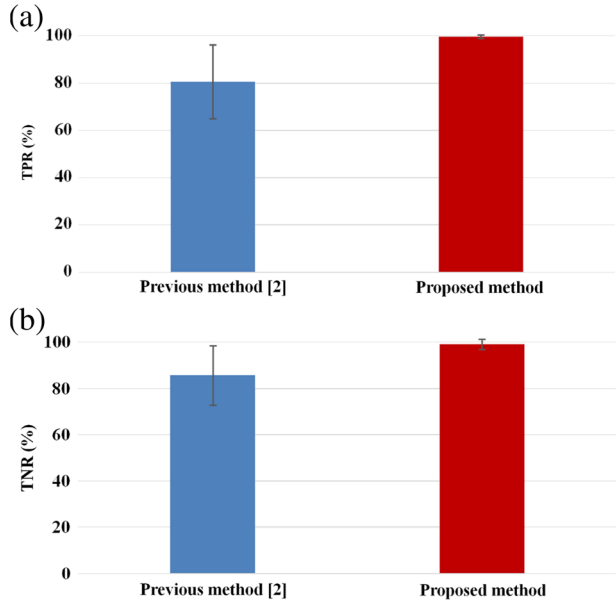
Our fuzzy system-based method (for estimating Z-distance and direction of home appliance) and gaze tracking algorithm do not require any training procedure, and the performance of our method is not affected by the training data. As shown in Section 3 with Tables 4, 5, 6, and 7, and Figs. 12, 13, and 14, we showed the experimental results with analyses for showing the efficiency of our method. Our fuzzy system can produce the continuous values of Z-distance (i.e. 171 cm, 195 cm, 283 cm, etc) and directions (left direction of -3° from the center, right direction of +21° from the center, etc) as shown in Fig. 12 and Table 4 (not discrete value). Therefore, our system can be used in
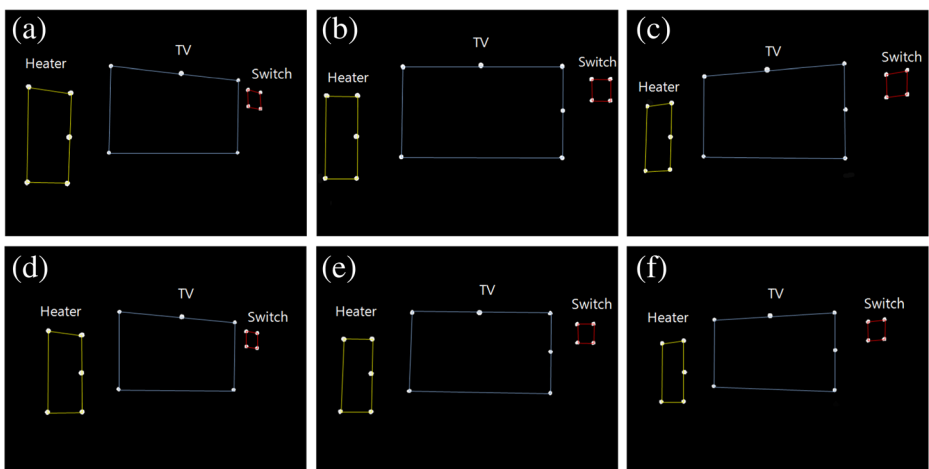
Fig. 14 The examples of recognizing multiple home appliances close to each other. The Z-distances of (**a**) ~ (**c**) and (**d**) ~ (**f**) are correctly estimated as 270 cm and 300 cm, respectively. The directions from user to home appliances are correctly estimated as (**a**), (**d**) right viewpoint, (**b**), (**e**) center viewpoint, (**c**), (**f**) left viewpoint

various positions (not in the same position). Based on these continuous values of Z-distance and directions, our system compensates the size and shape of the quadrangle area (defined by the NIR patterns as shown in the right image of Fig. 11) by offset. For example, if a user gazes at the TV at 295 cm although the calibration was performed at 270 cm, the TV size seen by the user inevitably becomes smaller. Therefore, we redefine the TV area (the right image of Fig. 11) by using the smaller area (whose size is decreased by the offset based on the measured (continuous) Z-distance) when our system determines whether the user gazes at the TV. For example, if a user gazes at the TV at the right position of +4° from the center although the calibration is performed at the center, the TV shape seen by the user inevitably becomes distorted (whose shape is distorted by the offset based on the measured angle) as shown in the left image of Fig. 5(a). Therefore, we redefine the TV area (the right image of Fig. 11) by using the distorted area like the left image of Fig. 5(a) when our system determines whether the user gazes at the TV.

A direct trigonometric computation can be considered for estimating Z-distance and direction. However, only the Z-distance and direction to home appliance (coordinate) based on frontal-viewing camera (coordinate) can be obtained by this method. In our research, there exists another gaze tracking camera (coordinate), and the relationship between the frontal viewing camera and gaze tracking camera coordinates should be calibrated also in case of using the trigonometric computation. In addition, user's eye (coordinate) based on the gaze tracking camera (coordinate) should be considered, also. However, through our fuzzy system-based method, we can easily obtain the gaze position on home appliance (coordinate) directly without these complicated calibration procedures among different cameras, home appliance, and user's eye coordinates. This is the advantage of our method.

4 Conclusion

In this research, we propose a new method of estimation of Z-distances and directions of home appliances using one (small-sized) web camera and a fuzzy system. Two feature values are extracted for estimating the Z-distance and another two feature values are obtained for estimating the direction. These feature values are used as the input to the fuzzy system, and we can obtain the Z-distance and direction of the home appliance based on the output of the fuzzy system. Experimental results show that the correct recognition rate of Z-distance and direction is 100%. The RMS errors of estimating Z-distance and direction are less than approximately 2 cm and 4°, respectively. In addition, the TPR and TNR with our method are higher than those by previous method, which shows the effectiveness of our method in gaze detection on home appliances.

For future work, we would test the performance of our system with more home appliances, and enhance the accuracy of our system by combining the fuzzy system with a training-based neural network.

References

1. Adeli H, Sarma KC (2006) Cost optimization of structures: fuzzy logic, genetic algorithms, and parallel computing. Willey, Hoboken
2. Al-Rahayfeh A, Faezipour M (2013) Eye tracking and head movement detection: a state-of-art survey. IEEE J Translational Eng Health Medicine 1:2100212–2100212
3. Barea R, Boquete L, Ortega S, López E, Rodríguez-Ascariz JM (2012) EOG-based eye movements codification for human computer interaction. Expert Syst Appl 39:2677–2683
4. Barua A, Mudunuri LS, Kosheleva O (2014) Why trapezoidal and triangular membership functions work so well: towards a theoretical explanation. J Uncertain Syst 8:164–168
5. Bayu BS, Miura J (2013) Fuzzy-based illumination normalization for face recognition. In: Proceedings of the IEEE Workshop on advanced robotics and its social impacts, Tokyo, Japan, 7–9 November 2013, pp. 131–136
6. Broekhoven EV, Baets BD (2006) Fast and accurate center of gravity defuzzification of fuzzy system outputs defined on trapezoidal fuzzy partitions. Fuzzy Sets Syst 157:904–918
7. Chao C-H, Hsueh B-Y, Hsiao M-Y, Tsai S-H, Li T-HS (2009) Fuzzy target tracking and obstacle avoidance of mobile robots with a stereo vision system. Int J Fuzzy Syst 11:183–191
8. Choi C, Kim J (2007) A real-time EMG-based assistive computer interface for the upper limb disabled. In: Proceedings of the 10th International Conference on rehabilitation robotics, Noordwijk, Netherlands, 13–15 June 2007; pp. 459–462
9. Cohen J (1992) A power primer. Psychol Bull 112:155–159
10. Cui J, Liu Y, Xu Y, Zhao H, Zha H (2013) Tracking generic human motion via fusion of low- and high-dimensional approaches. IEEE Trans Syst Man Cybern Part A Syst Hum 43:996–1002
11. Deng LY, Hsu C-L, Lin T-C, Tuan J-S, Chang S-M (2010) EOG-based human-computer Interface system development. Expert Syst Appl 37:3337–3343
12. Galdi C, Nappi M, Riccio D, Wechsler H (2016) Eye movement analysis for human authentication: a critical survey. Pattern Recogn Lett 84:272–283
13. Gonzalez RC, Woods RE (1992) Digital Image Processing, 1st edn. Addison-Wesley, Boston
14. Hales J, Rozado D, Mardanbegi D (2013) Interacting with objects in the environment by gaze and hand gestures. In: Proceedings of the 3rd International Workshop on pervasive eye tracking and mobile eye-based interaction, Lund, Sweden, 13 august 2013, pp. 1–9
15. Heo H, Lee JM, Jung D, Lee JW, Park KR (2014) Nonwearable gaze tracking system for controlling home appliance. Sci World J 2014:1–20
16. Homography (computer vision). Available online: https://en.wikipedia.org/wiki/Homography_(computer_vision). Accessed on 4 Feb 2016
17. Jacob RJK, Karn KS (2003) Eye tracking in human–computer interaction and usability research: ready to deliver the promises. In: The mind's eye: cognitive and applied aspects of eye movement research, 1st edn. Elsevier, Oxford, pp 573–605
18. Joglekar A, Joshi D, Khemani R, Nair S, Sahare S (2011) Depth estimation using monocular camera. Int J Comput Sci Inf Technol 2:1758–1763
19. Kinect for Windows. Available online: http://www.microsoft.com/en-us/kinectforwindows/meetkinect/default.aspx. Accessed on 7 July 2015
20. Klir GJ, Yuan B (1995) Fuzzy Sets and fuzzy logic-theory and applications. Prentice-Hall, NJ
21. Ko E, Ju JS, Kim EY, Goo NS (2009) An intelligent wheelchair to enable mobility of severely disabled and elder people. In: Proceedings of the International Conference on consumer electronics, Las Vegas, USA, 10-14 January 2009, pp. 1-2
22. Kocejko T, Bujnowski A, Wtorek J (2008) Eye mouse for disabled. In: Proceedings of the Conference on human system interactions, Krakow, Poland, 25–27 May 2008, pp. 199–202
23. Lankford C (2000) Effective eye-gaze input into WindowsTM. In: Proceedings of the symposium on eye tracking research & applications, Palm Beach Gardens, FL, USA, 6–8 November 2000, pp. 23–27
24. Leekwijck WV, Kerre EE (1999) Defuzzification: Criteria and classification. Fuzzy Sets Syst 108:159–178
25. Lim Y-C, Lee C-H, Kwon S, Jung W-Y (2008) Distance estimation algorithm for both long and short ranges based on stereo vision system. In: Proceedings of IEEE intelligent vehicles symposium, Eindhoven, Netherlands, 4–6 June 2008, pp. 841–846
26. Lin C-S, Ho C-W, Chen W-C, Chiu C-C, Yeh M-S (2006) Powered wheelchair controlled by eye-tracking system. Opt Appl 36:401–412
27. Liu Y, Zhang X, Cui J, Wu C, Aghajan H, Zha H (2010) Visual analysis of child-adult interactive behaviors in video sequences. In: Proceedings of the 16th International Conference on virtual systems and multimedia, Seoul, Republic of Korea, 20-23 October 2010, pp. 26-33

28. Liu Y, Cui J, Zhao H, Zha H (2012) Fusion of low-and high-dimensional approaches by trackers sampling for generic human motion tracking. In: Proceedings of the 21st International Conference on pattern recognition, Tsukuba, Japan, 11-15 November 2012, pp. 898-901
29. Liu Y, Nie L, Han L, Zhang L, Rosenblum DS (2015) Action2Activity: recognizing complex activities from sensor data. In: Proceedings of the 24th International joint Conference on Artificial Intelligence, Buenos Aires, Argentina, 25-31 July 2015, pp. 1617-1623
30. Liu Y, Zheng Y, Liang Y, Liu S, Rosenblum DS (2016) Urban water quality prediction based on multi-task multi-view learning. In: Proceedings of the 25th International joint Conference on Artificial Intelligence, New York, USA, 9-15 July 2016, pp. 2576-2582
31. Liu Y, Liang Y, Liu S, Rosenblum DS, Zheng Y (2016) Predicting urban water quality with ubiquitous data. CoRR, pp. 1–14
32. Liu Y, Zhang L, Nie L, Yan Y, Rosenblum DS (2016) Fortune teller: predicting your career path. In: Proceedings of the 13th AAAI Conference on Artificial Intelligence, phoenix, USA, 12-17 February 2016, pp. 201-207
33. Liu Y, Nie L, Liu L, Rosenblum DS (2016) From action to activity: sensor-based activity recognition. Neurocomputing 181:108–115
34. Liu L, Cheng L, Liu Y, Jia Y, Rosenblum DS (2016) Recognizing complex activities by a probabilistic interval-based model. In: Proceedings of the 13th AAAI Conference on Artificial Intelligence, Phoenix, USA, 12-17 February 2016, pp. 1266-1272
35. Lu Y, Wei Y, Liu L, Zhong J, Sun L, Liu Y (2017) Towards unsupervised physical activity recognition using smartphone accelerometers. Multimed Tools Appl in press
36. Majaranta P, Bulling A (2014) Eye tracking and eye-based human-computer interaction. In: Advances in physiological computing. Springer, London, pp 39–65
37. Mardanbegi D, Hansen DW (2011) Mobile gaze-based screen interaction in 3D environments. In: Proceedings of the Conference on novel gaze-controlled applications, Karlskrona, Sweden, 25–26 May 2011
38. Mauri C, Granollers T, Lorés J, García M (2006) Computer vision interaction for people with severe movement restrictions. Hum Technol 2:38–54
39. Medeiros MD, Gonçalves LMG, Frery AC (2010) Using fuzzy logic to enhance stereo matching in multiresolution images. Sensors 10:1093–1118
40. Mrovlje J, Vrančić D (2008) Distance measuring based on stereoscopic pictures. In: Proceedings of the 9th International PhD Workshop on systems and control: young generation viewpoint, Izola, Slovenia, 1–3 October 2008, pp. 1–6
41. Nilsson S, Gustafsson T, Carleberg P (2009) Hands free interaction with virtual information in a real environment: eye gaze as an interaction tool in an augmented reality system. PsychNology J 7:175–196
42. OpenCV. Available online: http://opencv.org/. Accessed on 6 Feb 2017
43. Pinheiro CG, Naves ELM, Pino P, Losson E, Andrade AO, Bourhis G (2011) Alternative Communication Systems for People with Severe Motor Disabilities: a Survey. Biomed Eng Online 10:1–28
44. Rebsamen B, Teo CL, Zeng Q, Ang MH, Burdet E, Guan C, Zhang H, Laugier C (2007) Controlling a wheelchair indoors using thought. IEEE Intell Syst 22:18–24
45. Ren Y-Y, Li X-S, Zheng X-L, Li Z, Zhao Q-C (2015) Analysis of drivers' eye-movement characteristics when driving around curves. Discrete Dyn Nat Soc 462792:1–10
46. Ross T, Fuzzy J (2010) Logic with engineering applications. Hoboken, Willey
47. Shi F, Gale A, Purdy K (2007) A new gaze-based Interface for environmental control. Lect Notes Comput Sci 4555:996–1005
48. Siddique N, Adeli H (2013) Computational Intelligence: synergies of fuzzy logic, neural networks and evolutionary computing. Willey, Hoboken
49. Standard score. Available online: https://en.wikipedia.org/wiki/Standard_score. Accessed on 31 July 2015
50. Student's T-Test. Available online: http://en.wikipedia.org/wiki/Student's_t-test. Accessed on 31 July 2015
51. Su M-C, Wang K-C, Chen G-D (2006) An eye tracking system and its application in aids for people with severe disabilities. Biomed Eng: Appl Basis Commun 18:319–327
52. Suh IH, Kim TW (2000) A visual servoing algorithm using fuzzy logics and fuzzy-neural networks. Mechatronics 10:1–18
53. Tu J, Tao H, Huang T (2007) Face as mouse through visual face tracking. Comput Vis Image Underst 108:35–40
54. Turner J, Bulling A, Gellersen H (2012) Extending the visual field of a head-mounted eye tracker for pervasive eye-based interaction. In: Proceedings of the symposium on eye tracking research and applications, Santa Barbara, USA, 28–30 March 2012, pp. 269–272

55. Vernon D, Tistarelli M (1990) Using camera motion to estimate range for robotic parts manipulation. IEEE Trans Robot Autom 6:509–521
56. Webcam C600. Available online: https://support.logitech.com/en_us/product/5869. Accessed on 7 July 2015
57. Weibel N, Fouse A, Emmenegger C, Kimmich S, Hutchins E (2012) Let's look at the cockpit: exploring mobile eye-tracking for observational research on the flight deck. In: Proceedings of the symposium on eye tracking research and applications, Santa Barbara, USA, 28–30 March 2012, pp. 107–114
58. Zhao J, Bose BK (2002) Evaluation of membership functions for fuzzy logic controlled induction motor drive. In: Proceedings of the IEEE Annual Conference of the Industrial Electronics Society, Sevilla, Spain, 5–8 November 2002, pp. 229–234

Jae Woong Jang received a BS and Master degrees in Computer Science from Sangmyung University, Seoul, South Korea in 2014 and 2016, respectively. His research interests include pattern recognition, biometrics, and image processing.

Hwan Heo received a Ph.D degree in Electronics and Electrical Engineering at Dongguk University in 2015. His research interests include pattern recognition, biometrics, and image processing.

Jae Won Bang received a PhD degree in Electronics and Electrical Engineering at Dongguk University in 2015. He is working in KIST since 2016. His research interests include biometrics and pattern recognition.


Hyung Gil Hong received a BS and Master degree in electronics engineering from Dongguk University, Seoul, South Korea, in 2012 and 2014, respectively. He is pursuing PhD degree in Electronics and Electrical Engineering at Dongguk University. His research interests include biometrics and pattern recognition.

Rizwan Ali Naqvi received his BS degree in computer engineering from Comsats Institute of Technology, Pakistan, in 2008. He received his MS degree in electrical engineering from Karlstad University, Sweden, in 2011. He was a lecturer at Superior University, Pakistan, until 2014. Currently, he is a PhD candidate in the Department of Electronics and Electrical Engineering at Dongguk University. His research interests include image processing, computer vision, and HCI.

Phong Ha Nguyen received a BS degree in Mechatronics, Hanoi University of Science and Technology (HUST), in 2015. He is pursuing a Master degree in Electronics and Electrical Engineering at Dongguk University. His research interests include biometrics and pattern recognition.
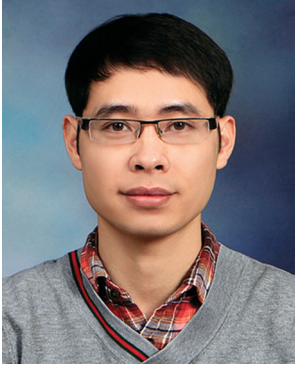
Dat Tien Nguyen received a BS degree in Electronics and Telecommunication, Hanoi University of Science and Technology (HUST), in 2009. He received PhD degree in Electronics and Electrical Engineering at Dongguk University in 2015. He has been an assistant professor in Electronics and Electrical Engineering at Dongguk University since 2015. His research interests include biometrics and pattern recognition.

Min Beom Lee received a BS degree in Computer Engineering, Dongyang University, in 2016. He is pursuing the combined course of Master and Ph.D degree in Electronics and Electrical Engineering at Dongguk University since 2016. His research interests include biometrics and pattern recognition.

Kang Ryoung Park received his BS and Master degrees in electronic engineering from Yonsei University, Seoul, South Korea, in 1994 and 1996, respectively. He received his PhD degree in electrical and computer engineering from Yonsei University in 2000. He has been a professor in the division of electronics and electrical engineering at Dongguk University since March 2013. His research interests include image processing and biometrics.