Improved Video-Based Eye-Gaze Detection Method

Yoshinobu Ebisawa, *Member, IEEE*

***Abstract*—Recently, some video-based eye-gaze detection methods used in eye-slaved support systems for the severely disabled have been studied. In these methods, infrared light was irradiated to an eye, two feature areas (the corneal reflection light and pupil) were detected in the image obtained from a video camera and then the eye-gaze direction was determined by the relative positions between the two. However, there were problems concerning stable pupil detection under various room light conditions. In this paper, methods for precisely detecting the two feature areas are consistently mentioned. First, a pupil detection technique using two light sources and the image difference method is proposed. Second, for users wearing eye glasses, a method for eliminating the images of the light sources reflected in the glass lens is proposed. The effectiveness of these proposed methods is demonstrated by using an imaging board. Finally, the feasibility of implementing hardware for the proposed methods in real time is discussed.**

***Index Terms*— Biomedical communication, biomedical equipment, biomedical measurements, biomedical signal detection, computer interfaces, handicapped aids, image processing.**

I. Introduction

MANY of the physically disadvantaged, such as amyotrophic lateral sclerosis patients, who cannot use basic body movements accurately, do indeed have eye movement ability. Support systems controlled by eye-gaze input can easily be projected as the most useful and desirable way for such people to independently aid themselves [1], [2]. In such a system, menu options would appear on a computer display, and the user would make selections by staring at the desired option for a short time. The user is therefore able to control peripheral assistant devices or can communicate electronically with others. However, there still has not been any eye-gaze detection method developed which is truly suitable for support systems.

To achieve such a system would necessitate equipment for accurately detecting where on the display the user is looking. Many conventional eye-gaze detection methods already exist (e.g., the infrared limbus reflection method [3]), but these require eye glass frame type or goggle type detectors, which must be attached to the user's face. Moreover, recalibration becomes necessary each time detectors are attached. Of course, this is an extra burden not only for the patient but also for the attendant. On the other hand, Hutchinson *et al.* [1] extracted the corneal reflection light (glint) and pupil image (bright eye) from the eye's image obtained from a video camera located immediately below the center of a computer display, while irradiating the eye with invisible infrared light using a light emitting diode (LED) positioned in front of the camera lens. Subsequently, the eye-gaze position was determined from the relative positions of the glint and pupil centers. The method is dramatically suitable for support systems because it is not only a noncontact method (neither need a sensor nor a marker attached to a user's head) but also allows for slight head movement. Moreover, although most eye-monitors measure eye direction relative to head direction, fundamentally in this system the eye direction can be determined from only the video images. So head movement measurement (especially rotation) is not necessary. Furthermore, unlike the other video-based methods with a light source located at largely eccentric positions (e.g., 20 deg) from the camera axis [4], the light source is attached to the camera. By this arrangement the light source and camera are simultaneously directed to the eye. This will facilitate the daily setup of the system and lead to its betterment in the future: i.e., automatic head (eye)-pursuing system under almost complete free head conditions. However, consistent and reliable pupil detection is difficult, as described by Hutchingson *et al.* in their paper on this method, because the pupil does not differ sufficiently in brightness from its surroundings.

To solve this problem, in this paper, a pupil detection technique using two light sources and the image difference method will be proposed, and it will be shown that the method facilitates pupil detection under various room light conditions. Furthermore, for eye glass users a simple method which enables eye-gaze detection will be proposed, and its effectiveness will be demonstrated.

II. Methods

A. Pupil Detection Technique Using Two Light Sources and the Image Difference Method

When the eye (face) is illuminated by a near-infrared light source, the light enters the pupil and is reflected off the retina and comes back to *the light source* again through the pupil because of the property of the optical system of the eyeball. Therefore, if the light source is set in coaxial with the camera (bright eye condition), the pupil image appears as a larger half-lighted disc against a darker background, called the bright eye (bright pupil) [1], [5], as shown in Fig. 1(a) and (d). In contrast, if the eye is illuminated by a light source uncoaxial with the camera (dark eye condition), the pupil image appears as a darker area against the background, called the dark eye (dark pupil) [4], as shown in Fig. 1(b) and (e).

Under both bright and dark eye conditions, a fraction of the infrared light is reflected off the front corneal surface and appears as a small intense area in the image, called the glint [1] or corneal reflection light, which is the first Purkinje image of the light source [3], as shown in Fig. 1(a), (b), (d), and (e).

Manuscript received June 14, 1994. This work was supported in part by the Nissan Science Foundation.

The author is with the Faculty of Engineering, Shizuoka University, Johoku 3-5-1, Hamamatsu, 432-8561 Japan.

Publisher Item Identifier S 0018-9456(98)09922-7.

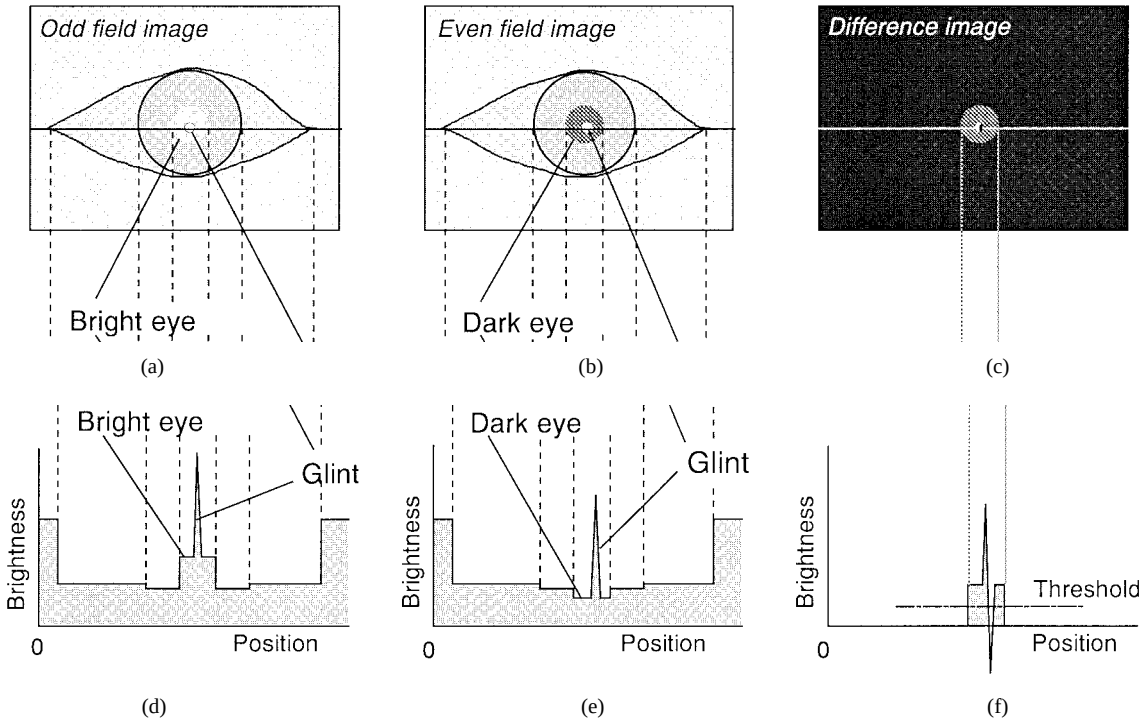

Fig. 1. Explanation of the pupil detection technique using two light sources and the image difference method. (a) and (b) indicate the images under the bright and dark eye conditions, respectively. (d) and (e) indicate the brightness distributions on the lines shown in (a) and (b), respectively. (c) Difference image obtained by subtracting (b) from (a). (f) Brightness distribution on the line shown in (c).

Although the glint position remains almost fixed in the image field as long as the user's head remains stationary, the position of the pupil (the bright or dark eye) moves, following the rotation of the eyeball. For example, if the subject looks at the camera, glint appears in the center of the pupil image. If the subject looks above the camera, the pupil center moves upward against the glint position. So under both bright and dark eye conditions, the direction of gaze against the camera can be determined from the relative positions of the pupil and glint center.

In the pupil detection technique using two light sources and the image difference method, one light source set in coaxial with the camera is switched on during the odd fields and another one set uncoaxial with the camera is switched on during the even fields (one frame of video signal consists of one odd field and one even field). The difference images are obtained by subtracting the even field images from the consecutive odd field images. As a result, the background almost vanishes as shown in Fig. 1(c) and (f). Here, it is necessary to make the background brightness levels in the odd and even field images equal by controlling the light sources' current. After binarizing the obtained difference image, the pupil area is detected. Here, the positional difference between the two light sources generates a small hole-like image on the pupil [see Fig. 3(c) right]. This hole-like image will be described in the next section and Section IV-A.

B. Elimination of Eye Glass Reflection Light (GRL)

The eye glass lens reflects the infrared light and yields the intense area on it (GRL). Fig. 2(a) and (b) schematically indicate the brightness distributions in the neighborhood of the pupil and GRL in an image, respectively. In these figures the brightness distributions of the odd and even fields are presented together. Fig. 2(b) and (h) indicate the corresponding brightness distributions in the difference image. Here, the negative range of the brightness level is not presented. The GRL usually saturates in the odd and even field images as shown in Fig. 2(g). In addition, the GRL profiles in both field images show a large difference in their positions. Therefore, in the difference image, the GRL area remains as shown in Fig. 2(h). The area causes a large error in pupil center detection.

The right and left figure lines in Fig. 2 explain the GRL elimination and pupil remaining processes, respectively. Both the GRL and pupil images are identically processed. Since in the difference image the GRL is usually brighter than the pupil [compare Fig. 2(h) with (b) and refer to Section IV-B], most of it can be eliminated by nullification of the parts brighter than the threshold as shown in Fig. 2(i). However, this high brightness elimination process widens the hole in the pupil image caused by the glint [compare Fig. 2(c) with (b)]. It is necessary for precise pupil center detection to eliminate the hole in the pupil image and to exclude the remaining GRL image. For this compensation, mathematical morphological operations (dilation and erosion [6]) were applied after binarization. First, the binary image [Fig. 2(d) and (j)] is two-dimensionally dilated with a smaller structuring element to eliminate the hole in the pupil image [compare Fig. 2e with (d)], and then is eroded with a larger structuring element to exclude the GRL [compare Fig. 2l with (k)]. These operations reduce the area of the pupil image and shift its center. However, these changes do not influence the results of

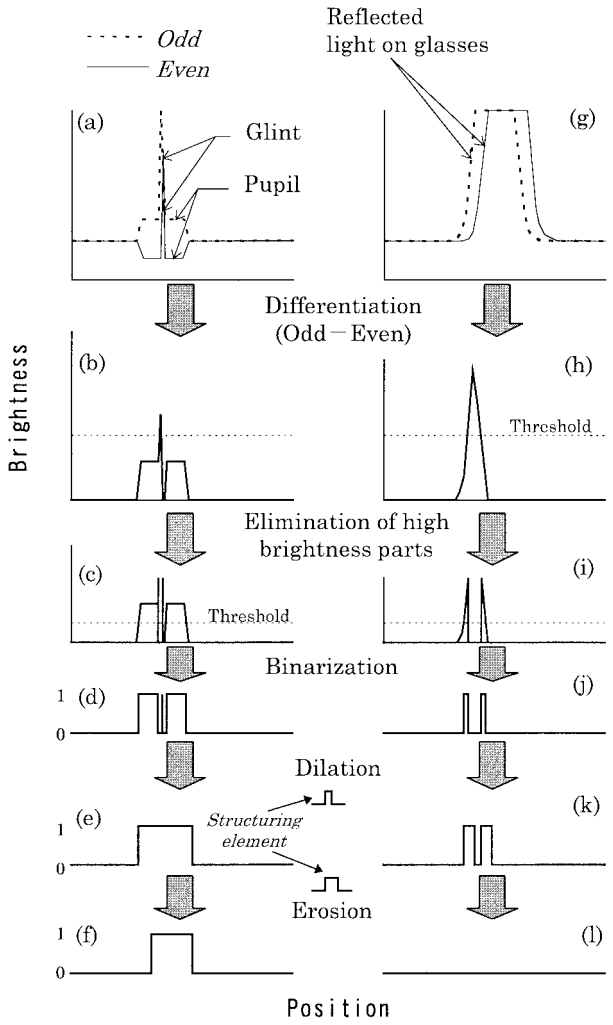

Fig. 2. The proposed method for eliminating the reflected light in eye glass lens. Left figures, (a)–(f), indicate the glass reflection light elimination process. Right figures, (g)–(l), indicate the pupil remaining process.

eye-gaze position detection after calibrating the relationship between the system outputs and actual eye-gaze positions if the sizes of the structuring elements are constant.

However, if the GRL can not be perfectly removed at the last stage in Fig. 2, it causes an error in determining the pupil center. A circular moving window would solve this problem. The reduction of the GRL area by the mathematical morphological operations brings the centroid (*the center of gravity*) close to the true pupil center. Therefore, if the next difference image in real time processing were windowed so that the center of the unmasked area (window) coincided with the centroid, it would be possible to perfectly eliminate the GRL from the beginning.

This window processing is also very important for glint detection, as mentioned in Section IV-C.

C. Experimental Methods

A standard interlaced scanning CCD camera (NTSC, 30 frames per second, Konica KS-30) having near-infrared sensitivity was used after removing the installed infrared-cutoff filter. The camera was fixed on a pedestal which manually rotates two-dimensionally. It was placed just below a 15 in computer display.

A motorized zoom lens (COSMICAR C6Z1218MESP-3, $f = 12.575$ mm, F1.8, $\times$ 6) was attached to the camera, with an extension ring (COSMICAR 2-EX, $\times$ 2) between them. The ring doubled the image, while making the focal length longer. To shorten the focal length, a close-up lens (COSMICAR CU1-49, +1D, 49 mm dia.) was attached to the front of the zoom lens. Consequently, the total focal range became about 50–100 cm. On the front of the close-up lens, an infrared pass filter (COSMICAR IR-80, 800 nm) was attached to reduce the amount of ambient visible light reaching the CCD. Finally a lens-protecting filter (58-mm dia.) was installed in the front of the infrared pass filter, with an adapter-ring which converts the diameters (49 mm(58 mm) between them. In the lens-protecting filter, two 4.7-mm-dia. circle holes were made; one was located at the center and the other was located 23 mm eccentrically. Two infrared LED's (880 nm, lens-can type, TOSHIBA TLN201) were put in the holes, respectively, and then glued toward the camera's subject. The LED's had strong directivity with a half-power angle of $\pm 7°$. The center LED, in coaxial with the camera, and the off-center LED, uncoaxial with the camera, were alternately switched by using an even/odd signal derived from an imaging board (CYBERNETICS TECHNOLOGY CT-9800B). The center LED was switched on during the odd fields, and the off-center LED was switched on during the even fields. The images were obtained in a room having windows toward the north, under fluorescent light during the day (1300 $l\times$) and at night (700 $l\times$). The subject sat, with his left side toward the windows. Images of his right eye were obtained. His head was held in place facing the computer display by a chin-and-forehead rest stand. The distance between the eye and the center LED was 50 cm. The center LED current when lit up was set at 40 mA, and then the off-center LED current when lit up was controlled and the brightness in the surroundings of the pupil in the even field images was adjusted to that of the odd field images on a black and white video monitor for camera-image feedback. The total infrared light power at the position of the eye when the sensor was directed to the camera (the center LED) was 88 and 41 μW/cm^2 in the day and night experiments, respectively. In contrast, that produced when the sensor was directed to the windows in the daytime experiment was about 1.6×10^3 μW/cm^2. These results indicate that the infrared light power irradiated to the eye by the LED's (about 40 μW/cm^2) was much weaker than that included in the sunlight.

When the subject wore eye glasses, the light that reflected off the front and rear surfaces of the lens appeared in the image fields as large, bright spots. In the experiments with eye glasses, the camera located just below the center of the computer display was moved to the right (about 15 deg) so that the GRL would not be superimposed on the pupil in the video images [see Fig. 4(a) and (b)].

Either with or without eye glasses, consecutive odd and even field images were taken by using the imaging board (aspect ratio 1 : 1) with a resolution of 512H $\times$ 256V pixels in each field, and were stored on a hard disc. The obtained images were processed later by a personal computer (NEC PC9801DA2)

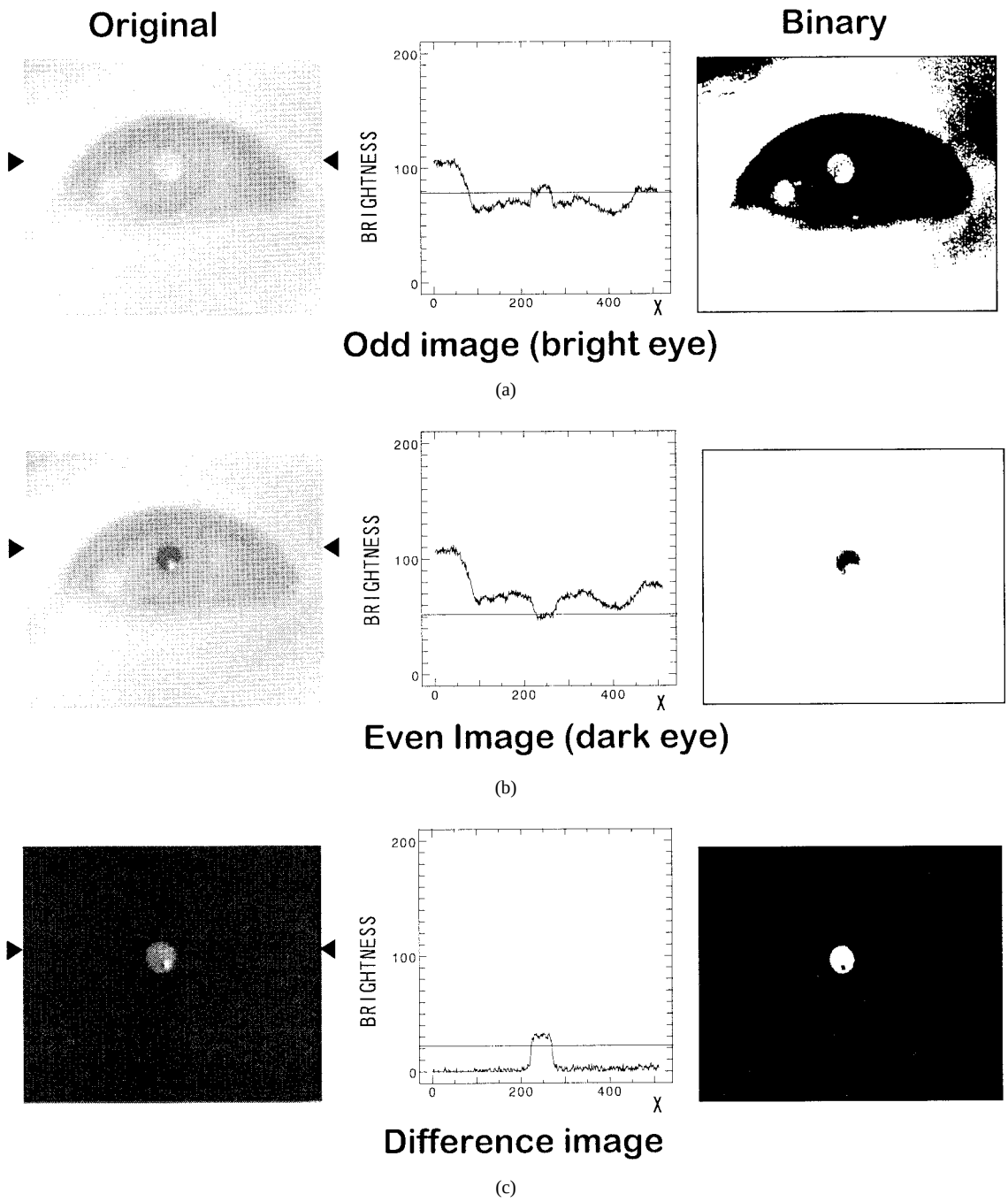

Fig. 3. Comparison in pupil detection between the (a) bright eye, (b) dark eye, and (c) proposed methods.

through the imaging board. The image differences (resolution 512H $\times$ 256V) were obtained by subtracting the even images from the odd images. Here, when the differences indicated negative values, the results were calculated as zero (black).

III. Results

A. Effectiveness of the Proposed Pupil Detection Method

Pupil detection was not easier during the daytime than at night. So only the results obtained from the daytime experiments will be shown.

Fig. 3 compares the difference image with the odd and even field images, which were obtained from the experiments without eye glasses. The vertically arranged figures in the center show the brightness distributions, on the lines located between the two triangles (▶ ◀), in the corresponding original images on the left. The right figures show the binary images obtained from the corresponding original images by using each threshold (horizontal line) shown in the central figures.

Although the pupil brightens in the odd field image, the surroundings of the eye are brighter than the pupil, because the sunlight includes a strong infrared light. Therefore, it is impossible to detect the pupil by using simple image

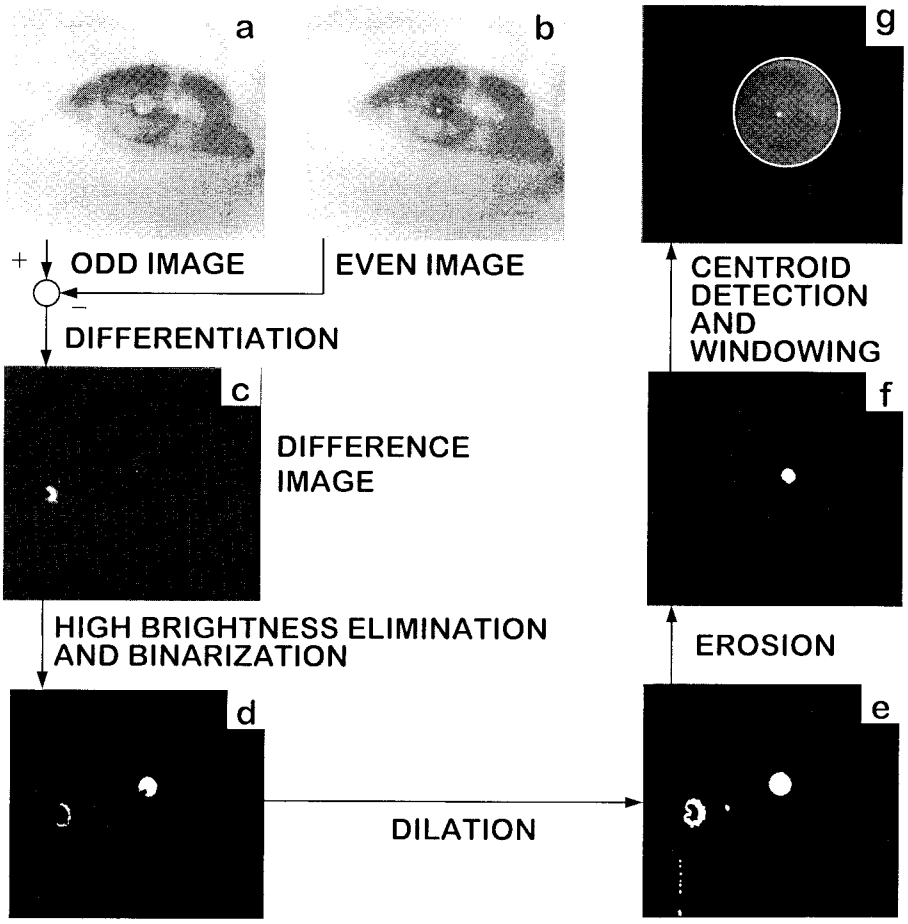

Fig. 4. The result of a pupil and glint detection while wearing eye glasses.

processing algorithms. In the even field image, the pupil is a little darker than the surroundings. However, the threshold determination is very difficult. In contrast, in the difference image, the surroundings of the pupil indicate the almost zero level in brightness. Consequently, although the pupil brightness changes under various conditions, the pupil can be easily detected by setting the threshold as low as possible.

B. Pupil Detection While Wearing Eye Glasses

Fig. 4 shows the effectiveness of the proposed algorithm for eliminating the GRL and detecting the pupil. Fig. 4(a) and (b) show the consecutive odd and even field images obtained from the experiment with eye glasses. In these images, a large, bright area on the left indicates the GRL. In addition, the surrounding objects in the room reflect in the glass lens. They also disturb pupil detection in the odd image (bright eye method). However, in the difference image [Fig. 4(c)], the objects perfectly disappear, though the GRL cannot been removed completely. Here, the pupil can not be seen because of its low brightness: about 20 in the range of 0–255. The difference image was binarized by the level of 10 after nullifying the pixels brighter than 25. In the binarized image [Fig. 4(d)], the chipped pupil and slight GRL can be seen. Here, a higher threshold for high brightness elimination made the GRL image larger. The image (Fig. 4d) was dilated with a rectangular structuring element of 6×3 pixels (it appeared a square; note the resolution of the difference image). This operation compensated the chipped pupil image though emphasizing the GRL image in size [Fig. 4(e)]. Then it was eroded with the element of 16×8 pixels. The resulting image [Fig. 4(f)] shows that the pupil can be successfully detected while eliminating the GRL.

From the image the centroid was calculated, and then was applied to the center of the circular window that should be given to the next difference image. Here, however, the window is given to the image shown in Fig. 4(a) (odd field image). The boundary of the window is indicated by the white circle [Fig. 4(g)]. The window position indicates that the pupil can be more easily detected in the next difference image. Here, although there is a discrepancy between the window and true pupil centers due to the dilation and erosion operations, it can be easily compensated.

IV. Discussion

A. Noise Removal for Precise Pupil Detection

Conventionally, the following two types of pupil center determination methods have been used: the centroid detection method and the edge detection method. The former is advantageous to noisy images. Although the latter does not allow any background noise, it is possible to estimate the

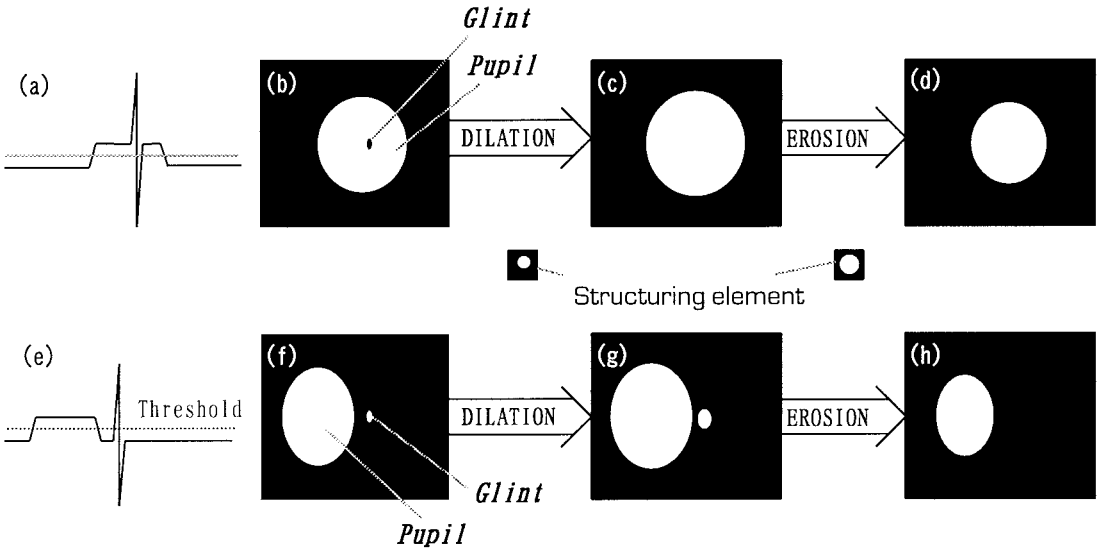

Fig. 5. Schematic diagrams of the glint noise elimination process for precise pupil detection. (a) and (e) indicate the brightness distributions in the difference images when the eye is directed toward the camera and when the eye is directed toward the right of the camera, respectively. (b) and (f) indicate the binary images of (a) and (e), respectively. They are processed in the same way as the eye glass reflection light elimination process. However, these diagrams do not display the stage of high brightness elimination. Consequently, the glint noise can be eliminated as shown in (d) and (h). Here, for simplicity the structuring elements are expressed with circles.

pupil shape. It could be available to quick menu selection by using a blink. In the pupil detection technique using two light sources and the difference image method, the glints in the odd and even field images partially cancel each other out. Therefore, the glint appears as a small area, compared with the bright or dark eye method. This is an additional advantage of the proposed method in pupil center determination as described below.

When the eye is directed toward the camera, the glint exists within the pupil area and it causes a small hole-like image as shown in Fig. 5(a) and (b) [see Fig. 3(c) right]. In contrast, when the eye is directed to largely eccentric positions from the camera axis, the glint comes out of the pupil image, especially during the daytime. The glint easily comes out of the pupil image due to pupil constriction and yields a small dot as shown in Fig. 5(e) and (f). These noises (hole and dot) cause an error in pupil center detection even while using the centroid detection method, especially when the pupil becomes small. However, because the noises are small in size, they can be easily eliminated by the mathematical morphological operations used in the GRL elimination process, as shown in Fig. 5(b)–(d) or (f)–(h). (Indeed, in Fig. 4 the sizes of the structuring elements were empirically determined so that the dot coming out of the pupil image could be eliminated.) By the operations, furthermore, precise eye-gaze detection becomes possible in a wide range of eye direction (over $\pm 30°$) even by using the edge detection method.

The operations can also compensate defective binarized pupil images. Under bad conditions (e.g., a low LED current), the pupil becomes darker. Even when the threshold is adjusted to an optimal level, the binarized images show overshadowed pupil and background noise. The operations can compensate the pupil image completely while eliminating the background noise perfectly. This not only allows stabler pupil detection but also enables restraining the LED light as low as possible, for the safety of the eye and the extension of the LEDs' life span.

B. Effect of High Brightness Elimination (HBE) and Pupil Brightness Control

When the GRL is not brighter than the pupil image in the difference image, the HBE process is ineffective. However, because in such a situation the GRL is smaller than the pupil, it can be easily removed by mathematical morphological operations. In contrast, when the GRL is brighter and larger than the pupil, it cannot be completely removed by the operations without HBE. In such a case, the centroid obtained from the image after applying the operations does not approach the pupil center. Accordingly, the window cannot be applied to the proper positions even after several images in real-time processing. Obviously, the HBE process plays an important role in pupil detection when wearing eye glasses, especially when the GRL is brighter than the pupil.

The pupil brightness in the difference images fluctuates largely in accordance with the pupil area (especially at night). To increase the effectiveness of HBE, the threshold should be set as low as possible but a little higher than the pupil brightness level. The pupil brightness fluctuation makes it difficult to set the threshold effectively. Since the pupil brightness in the difference images has linear relationships with the pupil area and with the LEDs' current, it is possible to keep the pupil brightness low. In fact, controlling in real time the current of the LED's in accordance with the pupil area obtained from a pupil pixel counter, we have succeeded in keeping the pupil brightness almost constant at a low level. This process enables effective HBE.

C. Glint Detection and Window Processing

Glint is another important feature area in eye-gaze position determination. When wearing eye glasses, glint detection by

simple binarization is difficult because it resembles the GRL in brightness [compare Fig. 2(a) with (g) or (b) with (h)]. Mathematical morphological operations cannot exclude the GRL image without eliminating the tiny glint. However, since the glint always exists near the pupil, it is possible to detect the glint after applying the window identical with that used for pupil detection [see Fig. 4(g)]. The window processing is indispensable for glint detection when wearing eye glasses.

When using the image difference method, there are four images from which glint would could be detected.

1) Odd field image.
2) Even field image.
3) Difference image (Odd minus even field image).
4) Difference image (Even minus odd field image).

Applying the window to the even minus odd field image may appear to facilitate the detection of glint because its surrounding image, including the pupil, disappears. However, the head and/or eye movement sometimes causes the glints in the odd and even fields to cancel out each other almost completely. As a result, stable glint detection from the difference images, including the odd minus even field image, is impractical. In contrast, it is easy in the odd or even field images because the glint is much brighter than its surroundings inside the window, when keeping the pupil brightness low as mentioned in Section IV-B. (However, to detect the glint in the even field image, the position of the off-center LED should be fixed. In the current system, the LED moves with the zooming level.) Accordingly, glint should be detected in the odd and/or even field images.

Without eye glasses, the window processing is not so important for glint detection. However, during the daytime an ambient infrared light raises the brightness level of the background (e.g., eyelid) in both odd and even images, as indicated in Fig. 1(d) and (e). This makes it difficult to set an appropriate threshold for glint detection. As a result, even without eye glasses, the employment of the window is desirable.

D. Feasibility of Hardware Implementation

We have already developed an image differentiator and glint and pupil detectors.

The image differentiator, based on first-in-first-out (FIFO) memories, can output the difference image obtained from consecutive odd and even fields, every 1/60 s as a video signal. The time delay from the input of the first field to the output of the first difference image is 1/30 s. The resolution is 640H $\times$ 256V.

The pupil detector consists of a binarizer, a noise eliminator, an edge detector and a pupil counter. The threshold necessary for the binarizer is given by the personal computer. In the noise eliminator, the dilation and erosion operations are applied for the binarized signal. The conventional methods for the operations need much time for calculation or a large scale circuit for large structuring elements, as used in the proposed method. Our noise eliminator [7], which is based on one-dimensional dilation circuits, image field memories and scanning direction changers, implements two-dimensional dilation and erosion operations for large structuring elements in spite of its small circuit scale. It needs 1/30 s for both operations, outputs every 1/60 s, and has a resolution of 256 $\times$ 256 pixels in an aspect ratio 1 : 1. The edge detector, in the binary image after noise elimination, calculates the x coordinates of the right and left extremes and the y coordinates of the top and bottom extremes of the pupil, and then sends them to a personal computer within the corresponding video field time. From these coordinates, the computer calculates the center coordinates of the pupil image. This pupil center detection method (edge detection) was almost equivalent to the centroid detection method in precision, because of our successful noise elimination.

The glint center coordinates are calculated in the same way as pupil, by using the glint detector consisting of a binarizer and an edge detector. (The glint detector does not include the noise eliminator.)

Without eye glasses, the current system can easily detect the pupil center from the difference images every 1/60 s, and can independently detect the glint center from the odd and even fields at the same rate. (The input terminals of the glint and pupil detectors are common in the current system.) For the simultaneous detection of the glint and pupil, a minor change of the system is necessary.

We are now developing two moving circular window processors (size changeable) based on a microprocessor and a video switch. We are also developing a high brightness eliminator based on an analog comparator and a video switch. Next, we plan to develop a microprocessor-based centroid detector. Since high precision is not necessary for determining the moving window position, centroid calculation with high sampling rate is not necessary. Accordingly, centroid detection using a microprocessor within the corresponding field time is easy.

As mentioned above, the implementation of the proposed methods is not difficult.

V. Conclusion

In this paper, a new video-based eye-gaze detection method overcoming the weaknesses of the conventional methods was proposed for the development of eye-slaved support systems for the physically disabled. First, a pupil detection technique using two light sources and the image difference method was proposed to facilitate the detection of the pupil, which is one feature area for eye-gaze position determination, under various room light conditions. Second, for users wearing eye glasses, a method was proposed to precisely detect not only the pupil but also the glint, which is another feature area to determine the eye-gaze position. The effectiveness of the proposed methods was demonstrated by some simulations using the imaging board. Finally, the feasibility of hardware implementation for the proposed methods was discussed.

References

[1] T. E. Hutchinson, K. P. White, Jr., K. C. Reichert, and L. A. Frey, "Human-computer interaction using eye-gaze input," *IEEE Trans. Syst., Man, Cybern.*, vol. 19, pp. 1527–1533, Nov./Dec. 1989.

[2] L. A. Frey, K. P. White, Jr., and T. E. Hutchinson, "Eye-gaze word processing," *IEEE Trans. Syst., Man, Cybern.*, vol. 20, pp. 944–950, July/Aug. 1990.
[3] L. R. Young and D. Sheena, "Methods and designs: Survey of eye movement recording methods," *Behavior Res. Methods Instrum.*, vol. 7, no. 5, pp. 397–429, 1975.
[4] A. M. Meyers, K. R. Sherman, and L. Stark, "Eye monitor, microcomputer-based instrument uses an internal model to track the eye," *Computer*, Mar. 1991, pp. 14–21.
[5] J. R. Charlier and J. C. Hache, "New instrument for monitoring eye fixation and pupil size during the visual field examination," *Med. Biol. Eng. Comput.*, vol. 20, pp. 23–28, Jan. 1982.
[6] R. M. Haralick, S. R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, pp. 532–550, July 1987.
[7] S. Kojima, Y. Ebisawa, and T. Miyakawa, "Fast morphology hardware using large size structuring element," *Syst. Comput. Japan*, vol. 25, no. 6, pp. 41–49, 1994.

Yoshinobu Ebisawa (M'95) was born in Hiratsuka, Japan, in February 1961. He received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1984, 1986, and 1989, respectively.

He is currently an Associate Professor in the Department of Systems Engineering, Faculty of Engineering, Shizuoka University, Hamamatsu, Japan. His current research interests include visual and oculomotor system and man–machine interface.