

AN IMPROVED PUPIL LOCALIZATION TECHNIQUE FOR REAL-TIME VIDEO-OCULOGRAPHY UNDER EXTREME EYELID OCCLUSION

SUNU WIBIRAMA*, IGI ARDIYANTO, THORIQ SATRIYA, TEGUH BHARATA ADJI
NOOR AKHMAD SETIAWAN AND MUHAMMAD TAUFIQ SETIAWAN

Intelligent Systems Research Group
Department of Electrical Engineering and Information Technology
Faculty of Engineering
Universitas Gadjah Mada
Bulaksumur, Yogyakarta 55281, Indonesia
*Corresponding author: sunu@ugm.ac.id

Received December 2018; revised April 2019

ABSTRACT. *Video-oculography (VOG) is a useful tool for detecting eye movement abnormalities as well as for human-computer interaction. Performance of a real-time VOG system relies on how the center of pupil is estimated under various occlusions. However, the accuracy of the state-of-the-art algorithms decreases significantly with extreme eyelid occlusion, although they can estimate pupil center under different illumination conditions and irregularity of pupil contour. To handle this problem, we propose a novel method to localize pupil in a real-time VOG during extreme eyelid occlusion. We have improved the geometric ellipse fitting method by incorporating a cumulative histogram processing for robust binarization and the random sample consensus (RANSAC) for removing outliers. Experimental results with 90% eyelid occlusion show that the proposed algorithm increases the accuracy of pupil localization up to 51% compared with a state-of-the-art method. During real-time pupil tracking, the proposed algorithm achieves significant accuracy improvement ($N = 9000$ frames; $p < .05$) with less than 10 ms of computational time per frame. The proposed method shows promising results for accurate measurement of horizontal and vertical eye movements.*

Keywords: Eye tracking, Pupil localization, Video-oculography, Ellipse fitting, Random sample consensus

1. Introduction. Eye tracking is a method to detect and to measure human eye movements. Eye tracking plays essential roles in various fields, such as human-computer interaction [1], applied application in an embedded system [2], multimodal sensory systems [3], medical robotics [4], behavioral research for industrial security [5], user experience in manufacturing systems [6], and computer-aided diagnosis [11]. Measurement of human eye movements is generally conducted using three kinds of eye tracking methods: scleral search coil (SSC), electro-oculography (EOG), and video-oculography (VOG) [7].

SSC has been regarded as the most accurate eye tracking technique with rigorous temporal and spatial resolution. This technique measures human eye movements based on changes in electromagnetic waves. The wave is based on voltage changes in a coil induced in a magnetic field. The coil is inserted inside a contact lens placed on the surface of the eye. However, the usage of SSC shall be accompanied by a medical expert as this

technique requires a complex configuration [8, 9]. Some researchers opt to use electro-oculography (EOG) because this method does not need direct placement of a contact lens during data collection.

EOG is based on the difference of corneo-retinal potential between the back and the front of the human eye [7]. EOG is useful for measuring eye movements when the eye is totally covered with eyelid – such as during sleeping activity. However, EOG is an obtrusive method because the electrodes have to be placed around the eyes. This may result in uncomfortable feeling when the device is used for more than 30 minutes [10]. Due to various constraints in SSC and EOG, eye movements researchers commonly use video-oculography (VOG) in practical studies.

VOG estimates eye movements from a captured video stream. VOG has gained popularity due to its non-invasiveness as it implements optical image processing algorithms instead of a contact lens or electrodes placement. Some VOG systems measure eye movements not only in horizontal and vertical but also in torsional direction [11, 12, 15]. Additionally, advanced image processing techniques may be incorporated in VOG systems to convert eye movements data into gaze information in horizontal, vertical, and depth direction [13, 14].

Additionally, VOG has been widely used as a computer-aided diagnosis system during a clinical investigation of eye movements disorders – such as nystagmus and saccadic dyskinesia [24, 37]. Nystagmus is involuntary oscillations of the eyes initiated by slow eye movements. Saccadic dyskinesia is an abnormal eye movement that drives the eyes away from their visual target and disrupts visual fixation. Both nystagmus and saccadic dyskinesia are used to observe various balance disorders [38], upper respiratory infection [39], migraine, head injury [40], and brain tumor [41].

As a medical diagnosis system, VOG implements three types of image processing-based pupil localization technique: center of gravity, curvature algorithm, and ellipse fitting (see Table 1). The center of gravity method is based on moment estimation of the pupil contour. A moment is a gross characteristic of the pupil contour computed by integrating or summing over all pixels of the contour. During the early stage of VOG research, the center of gravity was introduced in several research works [15, 17, 18, 19]. Unfortunately, this method is prone to error due to the irregularity of pupil shape, poor contrast, and unwanted corneal reflection resulting from light sources.

The curvature algorithm was introduced to solve these drawbacks [16, 20]. Initially, the algorithm defines a pupil blob – all points related to the pupil. The largest blob is then isolated, and all others are eliminated. Next, the algorithm identifies visible pupil boundary and separates it from edge points of eyelids, eyelashes, corneal reflections, and shadows using a heuristic curvature algorithm. The heuristic algorithm is mainly based on observations of the curvature characteristics of pupil boundaries during various occlusions by considering their arc length and Euclidean distance. Finally, circle fitting is conducted on the edges found by the curvature algorithm. The curvature algorithm yields high accuracy, but this method may cause false results due to curvature transition from the eyelid to the visible pupil.

On the other hand, pupil localization based on an ellipse fitting technique is performed by fitting an ellipse model to the detected pupil contour. This method allows fast pupil localization in a real-time VOG system. Therefore, most research works in VOG have implemented and combined this method with various techniques of contour detection. For instance, Roig et al. [21] combined morphological operations and edge detection with ellipse fitting. In 2013, Swart et al. [22] applied histogram processing, Sobel edge detection, and ellipse fitting in a 125 Hz VOG system. Recent research in 2018 by Ding et al. [27] implemented ellipse fitting based on circle difference method in a 50 Hz VOG system.

TABLE 1. Comparison of various state-of-the-art pupil localization methods in video-oculography systems for medical diagnosis

Authors	Sampling rate	Methods of pupil localization	Addressing extreme eyelid occlusion*
Moore et al. (1991) [15]	25 Hz	Center of gravity	No
Zhu et al. (1999) [16]	N/A	Curvature algorithm	No
Iijima et al. (2001) [17]	250 Hz	Center of gravity with histogram processing	No
Yagi (2008) [18]	240 Hz	Center of gravity	No
MacDougall et al. (2009) [19]	250 Hz	Center of gravity	No
Ong and Haslwanter (2010) [20]	130 Hz	Curvature algorithm	No
Roig et al. (2012) [21]	63 Hz	Ellipse fitting	No
Swart et al. (2013) [22]	125 Hz	Ellipse fitting with histogram processing	No
Wibirama et al. (2013) [11]	25 Hz	Center of gravity with gaussian smoothing	No
Wibirama and Hamamoto (2014) [13]	30 Hz	Ellipse fitting with convex hull	No
Barbosa and James (2014) [23]	3 Hz	Ellipse fitting with integro-differential operator	No
Charoenpong et al. (2015) [25]	25 Hz	Ellipse fitting with blob detection	No
Slama et al. (2017) [26]	30 Hz	Ellipse fitting with geodesic active contour	No
Ding et al. (2018) [27]	50 Hz	Ellipse fitting based on circle difference	No
Kong et al. (2018) [28]	30 Hz	Ellipse fitting	No
This research (2019)	30 Hz	Improved ellipse fitting with cumulative histogram and outliers removal	Yes

*) The pupil is partially occluded up to 70% of its vertical diameter (see Figure 2).

Despite of its wide use, previous works assumed that the pupil was not under extreme occlusion (i.e., the pupil was not partially occluded up to 70% of its vertical diameter). In reality, however, severe eyelid occlusion affects the estimation of the pupil center. Both Nyström et al. [42] and Choe et al. [43] have argued that accurate identification of the pupil center is important towards accuracy and precision of a VOG system.

To deal with this research gap, we propose a novel pupil localization technique based on an improved ellipse fitting algorithm with outliers removal. The proposed method aims to achieve better accuracy of pupil localization under extreme eyelid occlusion in a real-time VOG system. A new adaptive image binarization based on cumulative histogram is proposed to deal with varying brightness. The proposed method has been validated using static eye images and video streams. Additionally, we have compared the accuracy and the computational time of the proposed method with a state-of-the-art pupil localization

technique that achieved a remarkable performance in eye movements data consisting of nystagmus and random components [22].

2. Materials and Methods. Figure 1 shows the proposed pupil localization technique. The proposed algorithm consists of four main stages: masking region of interest, contour extraction, improved ellipse fitting with outliers removal, and moving average filter.

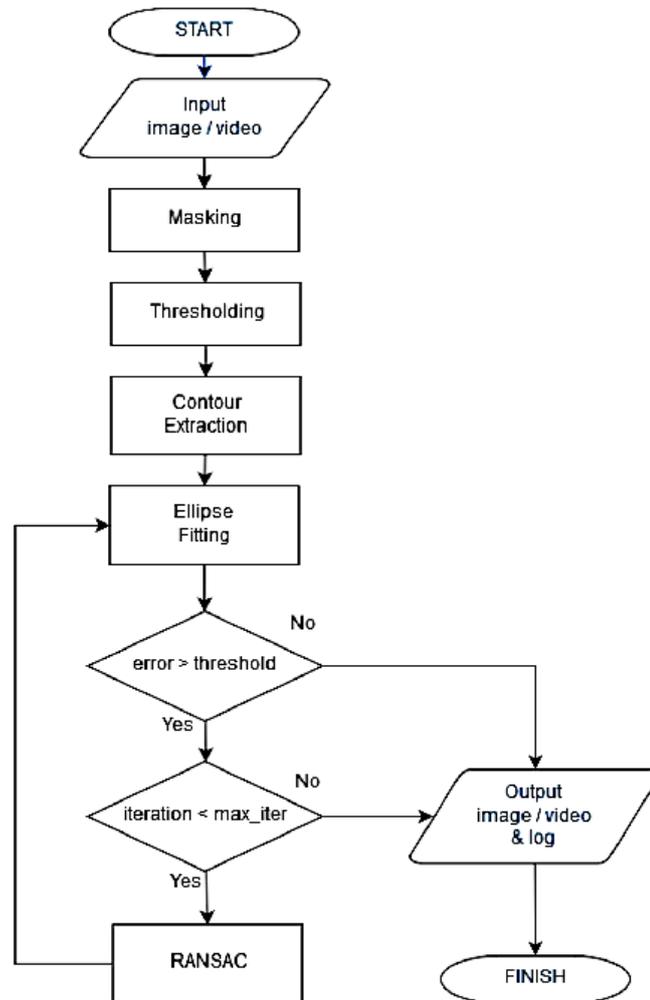


FIGURE 1. Pipeline of the proposed pupil localization technique

2.1. Masking region of interest (ROI). Pupil localization is performed by first converting an RGB image to a gray image. It is worth mentioning that capturing a clear eye image from a close-range RGB sensor or an infrared camera is challenging. In our study, the captured raw image data come with darker area at the corner, yielding noise in contour extraction process. To cope with this problem, image masking is introduced by comparing the captured eye image and a binary mask image. The region of interest (ROI) that contains pupil area is defined through an empirical observation. Additionally, various gray pixels outside the ROI are eliminated. Once the ROI is configured, the computational time of the following image processing stages becomes faster because the amount of the processed pixels decreases significantly. The masking process can be defined as:

$$dst(x, y) = \begin{cases} src(x, y) & \text{if } mask(x, y) = 1 \\ 0 & \text{if } mask(x, y) = 0 \end{cases} \quad (1)$$

Here, $dst(x, y)$, $src(x, y)$, and $mask(x, y)$ are pixels intensity value of the masked eye image, raw eye image, and mask image in x and y position, respectively.

2.2. Adaptive image binarization based on cumulative histogram and contour detection. Before extracting pupil contour, the masked eye image is converted to binary image. To automate binarization process, we propose an image binarization technique based on cumulative histogram. In Algorithm 1, $hist$, $cmlHist$, and $ratioDarkPix$ represent the image histogram, the cumulative histogram, and the ratio of pixels that composes dark area, respectively. The threshold value is determined by analyzing histogram of the masked eye image. In this case, pixels with intensity 0 (perfect black) and 255 (perfect white) are excluded from the computation since these values fill the largest part of the histogram.

Algorithm 1 Adaptive image binarization based on cumulative histogram

```

compute  $hist$ 
normalize  $hist$ 
for  $i = 2$  to  $i = 255$  do
     $cmlHist[i] += hist[i - 1]$ 
    if  $cmlHist[i] > ratioDarkPix$  then
        return  $i$ 
    end if
end for

```

To define the threshold value, the cumulative histogram is compared with the ratio of pixels that composes dark area in the image, without considering perfect black. Based on empirical analysis, we define the ratio of pixels that composes dark area around 3% out of all pixels. Compared with the conventional thresholding commonly used in biometric research [30], this binarization process is more robust to various eye images captured under near-infrared illumination. Thus, the proposed binarization technique may be used not only in static images but also in video processing with various brightness levels.

Following the binarization process, a simple contour detection algorithm is implemented to obtain the border of the pupil [29, 31]. The contour detection algorithm works by tracing each pixel of pupil edge, so that the boundaries between the pupil and its background are extracted. For each pixel, we combine 4- and 8-connectivity on 3×3 neighboring pixels.

2.3. Improved ellipse fitting with outliers removal. After the contours are extracted, an ellipse model is created based on these contours. The main idea of ellipse fitting is to find minimum geometric distance between the ellipse model and the contour points [44]. Let $\mathbf{p}_1 \dots \mathbf{p}_N$ be a set of N contour points, $\mathbf{p}_i = [x_i, y_i]^T$. Let $\mathbf{x} = [x^2, xy, y^2, x, y, 1]^T$, $\mathbf{p} = [x, y]^T$, and

$$f(\mathbf{p}, \mathbf{a}) = \mathbf{x}^T \mathbf{a} = ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (2)$$

as the equation of generic ellipse with the parameter vector $\mathbf{a} = [a, b, c, d, e, f]^T$. Next, we find the parameter vector \mathbf{a}_0 associated with the ellipse that matches $\mathbf{p}_1 \dots \mathbf{p}_N$ as the solution of:

$$\min_a \sum_{i=1}^n \|\mathbf{p} - \mathbf{p}_i\|^2 \quad (3)$$

with the constraint that \mathbf{p} belongs to $f(\mathbf{p}, \mathbf{a}) = 0$. To solve Equation (3), we can use Lagrange multipliers by defining an objective function:

$$L = \sum_{i=1}^n \|\mathbf{p} - \mathbf{p}_i\|^2 - 2\lambda f(\mathbf{p}, \mathbf{a}) \quad (4)$$

If we set $\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} = 0$, we obtain

$$\mathbf{p} - \mathbf{p}_i = \lambda \nabla f(\mathbf{p}, \mathbf{a}) \quad (5)$$

As \mathbf{p} is yet to be known, we introduce two approximations:

$$0 = f(\mathbf{p}, \mathbf{a}) \approx f(\mathbf{p}_i, \mathbf{a}) + [\mathbf{p} - \mathbf{p}_i]^T \nabla f(\mathbf{p}_i, \mathbf{a}) \quad (6)$$

and

$$\nabla f(\mathbf{p}) \approx \nabla f(\mathbf{p}_i) \quad (7)$$

Considering Equation (7), we can rewrite Equation (5) as

$$\mathbf{p} - \mathbf{p}_i = \lambda \nabla f(\mathbf{p}_i, \mathbf{a}) \quad (8)$$

Substituting Equation (8) into Equation (6), we obtain

$$\lambda = \frac{-f(\mathbf{p}_i, \mathbf{a})}{\|\nabla f(\mathbf{p}_i, \mathbf{a})\|^2} \quad (9)$$

Substituting Equation (5) into Equation (9), we finally obtain

$$\|\mathbf{p} - \mathbf{p}_i\| = \frac{|f(\mathbf{p}_i, \mathbf{a})|}{\|\nabla f(\mathbf{p}_i, \mathbf{a})\|^2} \quad (10)$$

We estimate goodness of fit for the ellipse model by minimizing the geometric distance between the model and the contour of pupil. However, some contour points are not useful to find the best ellipse model. These points are outliers, without which the ellipse fitting process is more accurate. If the pupil is under extreme occlusion – the pupil area is partially occluded up to 70% or more – eyelashes and eyelid will produce erroneous contour points. In this case, the ellipse fitting procedure incorporates the random sample consensus (RANSAC) [36].

Thus, when $\|\mathbf{p} - \mathbf{p}_i\|$ exceeds an empirical threshold, RANSAC is implemented (see Algorithm 2). The acceptable deviation of the model (ϵ) is set to be one standard deviation from the mean of the geometric distance error between the contour points and the ellipse

Algorithm 2 RANSAC for improved ellipse fitting

```

given  $n$  contour data  $X = \{x_1, x_2, x_3, \dots, x_n\}$ 
set  $m = 5$ 
set  $t = 20$ 
set iteration counter  $k = 1$ 
set iteration limit  $K = 10$ 
for  $k < K$  do
  choose  $m$  random items from  $X$  and fit ellipse model
  determine how many data from  $X$  within  $\epsilon$ 
  if data within  $\epsilon > t$  then
    return fit ellipse model over the data within  $\epsilon$ 
  end if
  set  $k = k + 1$ 
end for

```

model. As expressed by Equation (2), one can observe that $(a, b, c, d, e, f) \in \mathbb{P}^5(\mathbb{R})$. Hence, at least five points are needed ($m = 5$) to recover the coefficients of an ellipse.

The parameter t represents enough data points to confirm an acceptable model [36] and enough data points to refine the hypothesis to the final best estimate [48]. Fischler and Bolles have suggested to set t so that $t - m > 5$ [36]. This provides a higher than 95% probability to avoid compatibility with an incorrect model. In our case, the size of the consensus set t is set empirically to be 20 to ensure we get the best ellipse model over the data within ϵ .

2.4. Moving average filter. The moving average filter is incorporated in real-time video processing to suppress the spikes that occur due to pupil occlusion. This filtering process is performed by calculating the average positions of tracked pupil from previous frames:

$$\mathbf{P}_b[i] = \frac{1}{M} \sum_{i=0}^{M-1} \mathbf{P}_a[i+1] \quad (11)$$

Here, \mathbf{P}_a is the pupil coordinates before filtering and M is the length of the filter window.

2.5. Experimental evaluation of the proposed algorithm. The algorithm was implemented using OpenCV 3.1.0 in Microsoft Visual C# 2013 on a personal computer with an Intel Core i5 2.3 GHz processor, a 4 GB memory, and a Microsoft Windows 10 operating system. We conducted three experiments to measure the robustness of the proposed algorithm.

2.5.1. 1st experiment: evaluation on CASIA dataset. In the first experiment, we validated the proposed pupil localization technique on static eye images from Chinese Academy of Sciences (CASIA) dataset as used in the previous works [32, 33, 34]. CASIA dataset was collected from 1,000 participants with 10 times eye recording for each participant. The resolution of each eye image was 640×480 pixels. We developed a simple batch script to randomly generate 35 test images from 10,000 images of the dataset. Each image in CASIA dataset was formatted with filename S6***So*.jpg, where * represented a digit of 0 to 9. We iterated 35 times to pick a test image image by randomizing those 4 digits. The chosen test image was then copied to an experiment folder.

Figures 2 and 3 illustrate a method to simulate the eyelid occlusion. On each test image, we measured the distance of upper and bottom part of pupil in pixels unit. The distance was then divided into 10 parts. Each part represented 10% eyelid occlusion.

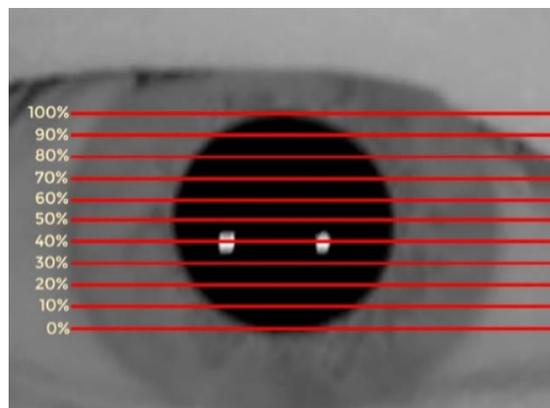


FIGURE 2. Range of pupil area that can be captured by infrared camera. 100% represents the whole pupil area is visible while 0% represents the whole pupil area is occluded by eyelid.

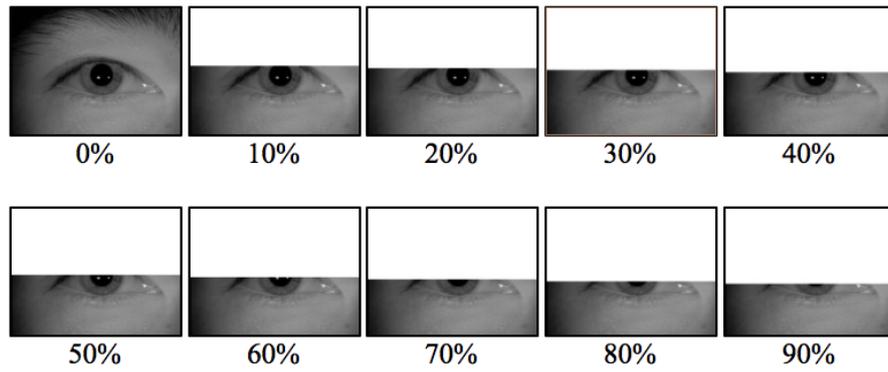


FIGURE 3. Eye images from CASIA dataset are modified to simulate various eyelid occlusion ranging from no occlusion (0%) to high occlusion (90%).

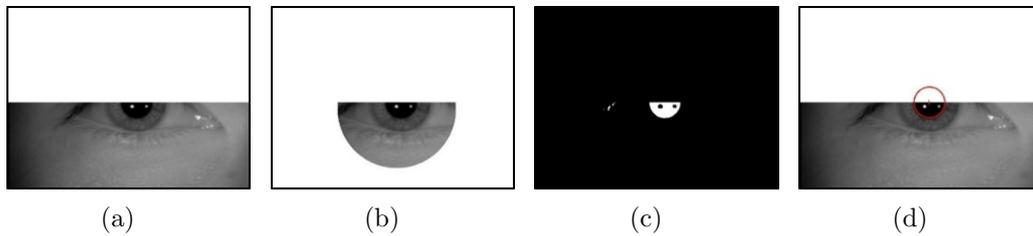


FIGURE 4. Validation on an image of CASIA dataset: (a) simulation of eyelid occlusion; (b) implementation of masking an ROI; (c) image binarization with automatic thresholding; (d) pupil localization using the proposed algorithm

Each testing image was modified by manually adding a white space simulating the eyelid occlusion. Note that images with no occlusion (0% occlusion) were used as the ground truth to measure the accuracy of the proposed algorithm.

Figure 4 shows implementation of the proposed algorithm on a test image. We measured the accuracy of the proposed algorithm by computing errors of pupil center estimation. Errors were computed from the difference of pupil center position between a benchmark image (an eye image without eyelid occlusion) and a test image (an eye image with occlusion):

$$err = \sqrt{(x - x')^2 + (y - y')^2} \quad (12)$$

where err denotes the absolute error in pixel unit. x and y represent the horizontal and vertical coordinates of pupil center in benchmark image, respectively. x' and y' represent the horizontal and vertical coordinates of pupil center in test image, respectively. The measurement was conducted five times and errors were averaged for each test image. The accuracy of algorithm (in %) was then computed with the following equation:

$$acc = \left(1 - \frac{err}{rad}\right) * 100\% \quad (13)$$

Here, acc denotes algorithm accuracy in % and rad represents pupil radius in a benchmark image.

We compared the accuracy of the proposed algorithm with a state-of-the-art pupil localization method proposed by Swart et al. [22]. Swart et al. implemented their algorithm on a medical VOG system with infrared illumination in a controlled experimental environment. Their setting of experiment – more or less – was similar to what we were performed

in our study. Additionally, we disregarded several other algorithms commonly used in a gaze tracking system for interactive applications in real-world environment since those algorithms were developed under different constraints with uncontrolled environment [45].

Swart et al. implemented a color transformation through histogram processing and a feature identification using ellipse fitting [22]. They evaluated their algorithm on signals of eye movements with pure nystagmus as well as signals with nystagmus and random components. They reported a remarkable performance with an average specificity of 96.47%. However, the performance of their algorithm during extreme eyelid occlusion is yet to be known.

In this experiment, the proposed algorithm and the state-of-the-art algorithm were implemented on 35 test images of the CASIA dataset, yielding an average accuracy under various eyelid occlusion. We then conducted a one-tailed statistical hypothesis testing with 5% of significance level. Since the sample size was more than 30, we used Z -test for hypothesis testing. The null hypothesis was the proposed pupil localization technique did not significantly improve the accuracy of pupil localization.

2.5.2. *2nd experiment: evaluation on a captured video from a Sony PS3 Eye camera.* We also evaluated the proposed adaptive binarization method using eye images captured from a Sony PS3 Eye camera (Sony Interactive Entertainment LLC, California). The Sony PS3 Eye camera was used in this study because it provided a wide range of configuration available for computer vision research. To capture the raw data from the camera, we installed CL Eye Driver version 5.3.0.0341 and CL Eye Platform SDK version 1.6.4.0028 [35]. The resolution of the camera was 640×480 pixels with a frame rate of 75 Hz. As depicted in panel (a) of Figure 5, the camera was positioned 17.5 cm from the ground and 23 cm in front of a printed image of human eye. The size of the eye image was 14×14 cm, attached on a white background paper with size of 35×35 cm. Panel (b) of Figure 5 shows: (1)

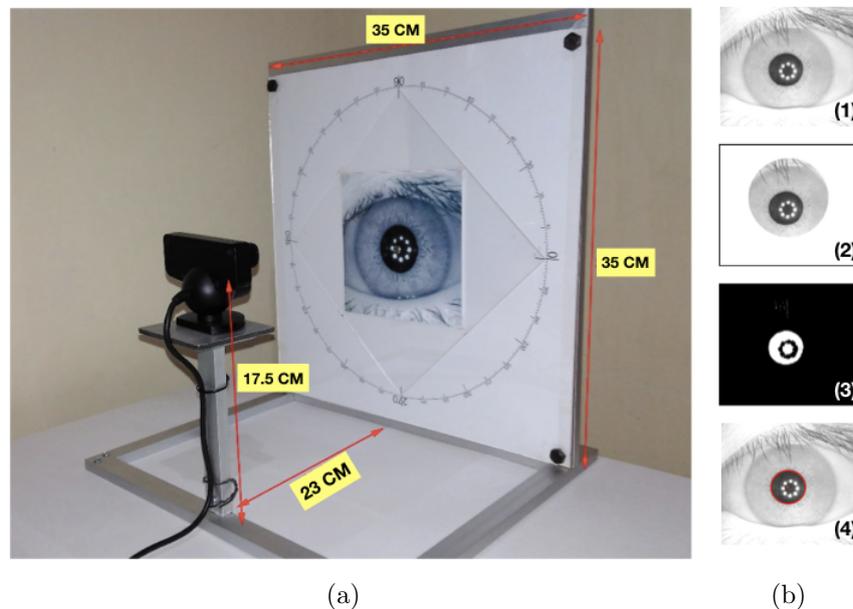


FIGURE 5. The experimental setting to evaluate the proposed adaptive thresholding method: (a) a Sony PS3 Eye camera was located in front of a printed eye image; (b) stages of pupil localization: (1) the captured gray scale image, (2) the eye image with masking, (3) result of adaptive thresholding, (4) the eye image with localized pupil

the captured gray scale image before image processing; (2) the eye image with masking; (3) the result of adaptive thresholding; and (4) the localized pupil. The experiment was performed by adjusting gain values of the camera to simulate various brightness levels. The gain values were set from 0 to 79 – the minimum and maximum values were based on the default configuration provided by CL Eye Platform SDK. We compared the proposed adaptive binarization method and a conventional binarization method [30]. The threshold value of the conventional method was defined to be 84. This predefined value was the closest value from an average threshold value resulted by the proposed adaptive binarization method. Comparative analysis of both binarization methods is presented in the experimental results.

2.5.3. 3rd experiment: evaluation on captured video streams from a VOG system. In the third experiment, the proposed algorithm and the state-of-the-art pupil localization technique were implemented in video streaming. The video data were captured using a Mobile Eye VOG system manufactured by Applied Science Laboratories (Bedford, Massachusetts). Three male participants (average of age: 24.7 years old) joined the experiment on voluntary basis. All participants were healthy with normal eyes. The participants wore a VOG goggle during data recording. The video data were recorded with a resolution of 768×480 pixels at a frame rate about 30 Hz. The detailed procedure of data collection was similar to a prior work by McMurrough et al. [46]. Errors of pupil localization were obtained by measuring absolute difference between the hand-labeled ground truth and the computed pupil positions. We processed 9,000 frames in real time using the proposed algorithm and the state-of-the-art pupil localization algorithm.

3. Experimental Results.

3.1. 1st experiment: evaluation on CASIA dataset. Figure 6 shows the accuracy of the proposed method and the state-of-the-art pupil localization algorithm proposed by Swart et al. [22] during various pupil occlusion. We validated the performance of our algorithm and the algorithm proposed by Swart et al. on eye images with 10%-90% range of pupil occlusion. Experimental results show that the proposed algorithm outperformed the state-of-the-art algorithm. The proposed method yielded 75.79% of accuracy even when the pupil was 70% occluded. On the other side, the state-of-the-art algorithm [22] yielded only 37.23% of accuracy. Statistical analysis showed that the performance of the proposed algorithm was superior to the state-of-the-art algorithm up to 70% eyelid occlusion ($Z = 2.23$, $p < .05$). However, the performance of the proposed method decreased when the pupil was 80% occluded (accuracy = 51.42%, $Z = 1.03$, $p < .05$). In this case, the proposed algorithm failed to estimate pupil center correctly because it could not find the best model that fitted the intended pupil curvature.

3.2. 2nd experiment: evaluation on a captured video from a Sony PS3 Eye camera. Figure 7 shows a comparative analysis of accuracy of pupil localization using the proposed adaptive binarization and a conventional binarization method [30]. We computed the average accuracy in overall gain values. Experimental results show that the proposed adaptive binarization method yielded higher accuracy (i.e., 99.36%) compared with the conventional binarization method (i.e., 60.8%). This results show that the proposed method can handle various brightness levels. This ability plays an important role when the proposed pupil localization algorithm is implemented in real-time video processing – a condition where unwanted brightness level may be incorporated in the recorded frames.

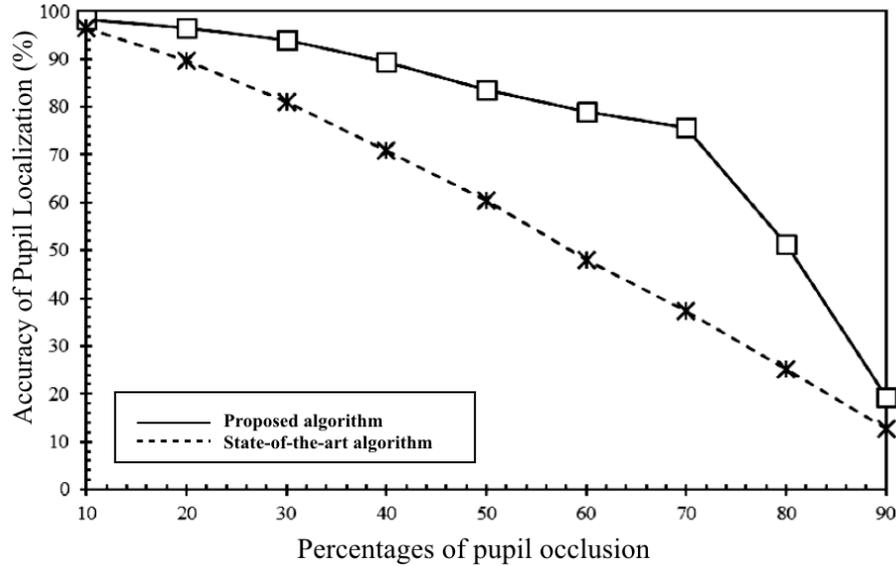


FIGURE 6. Comparative analysis of the proposed algorithm and a state-of-the-art algorithm [22] on estimating pupil center during various percentages of occlusion. Solid and dashed lines represent the proposed and the state-of-the-art algorithm, respectively. Both algorithms were implemented on 35 images randomly taken from CASIA dataset.

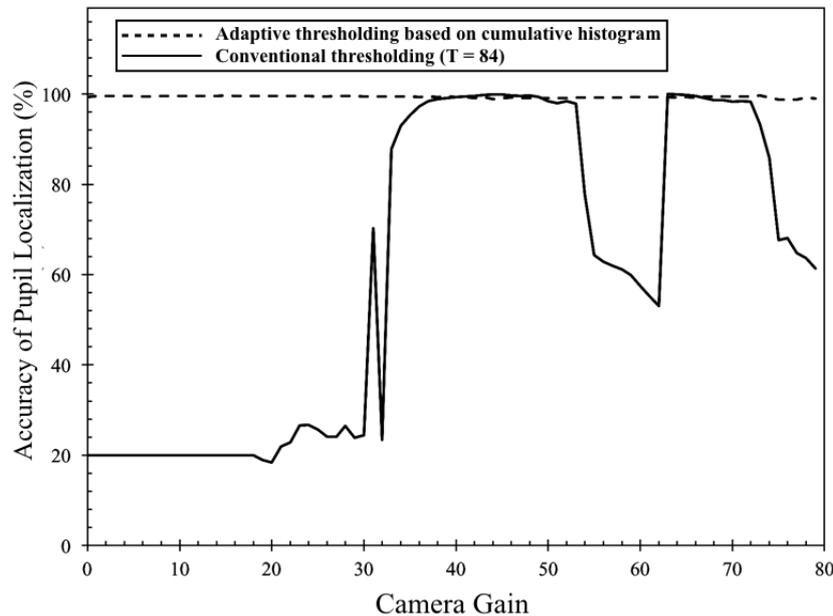


FIGURE 7. Comparative analysis of accuracy of pupil localization using the proposed adaptive thresholding and an existing thresholding method ($T = 84$) [30]

3.3. 3rd experiment: evaluation on captured video streams from a VOG system. Figures 8 and 9 show the recorded horizontal and vertical eye movement of the first participant. The proposed algorithm and the state-of-the-art algorithm are shown in panel (a) and panel (b), respectively. Figure 8 shows that the proposed algorithm minimized tracking error by removing spikes caused by eyelid occlusion – such as at 4 seconds,

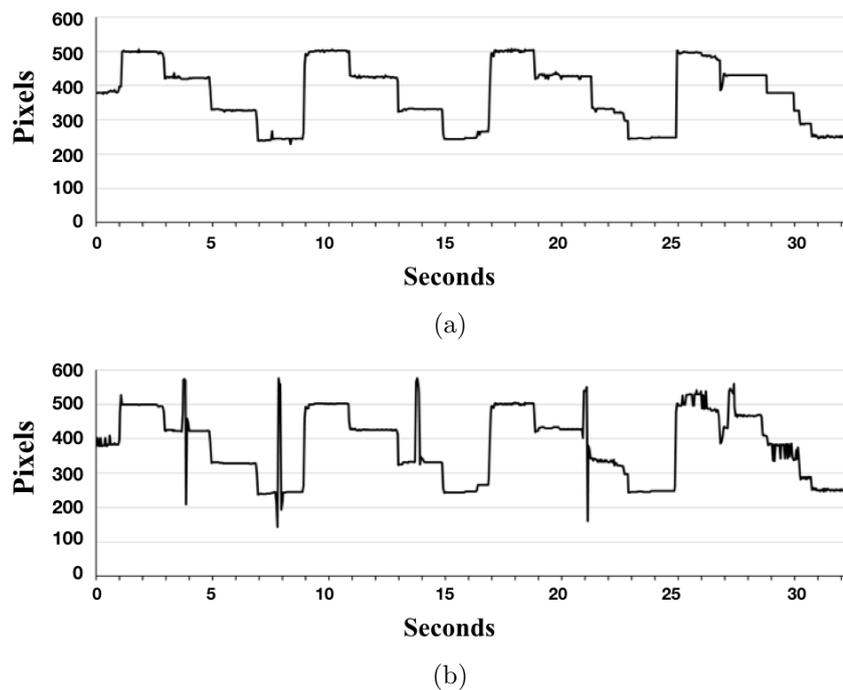


FIGURE 8. Sample of eye movement graph from the first participant during real-time video streaming. The graphs show comparison of pupil position (in pixels) in horizontal direction using (a) the proposed algorithm; (b) a state-of-the-art pupil localization algorithm [22]. The vertical and horizontal axes are coordinate (in pixels) and timestamp (in seconds), respectively.

8 seconds, 14 seconds, and 21 seconds of horizontal eye movements. Similar trends can be observed in Figure 9. The proposed algorithm removed sharp spikes at 4 seconds, 8 seconds, 14 seconds, and 21 seconds of vertical eye movements.

Figure 10 depicts several conditions during which the proposed algorithm overcame extreme eyelid occlusion. The proposed algorithm was able to locate the pupil position, albeit there was a slight error on estimating coordinates of pupil center (see panel (e) of Figure 10). On the contrary, the state-of-the-art pupil localization technique failed to detect the pupil while producing false coordinates of pupil center (panel (f) of Figure 10).

Table 2 shows comparison of errors of the proposed algorithm and the state-of-the-art pupil localization method. Experimental results show that the proposed algorithm yielded better performance during real-time eye movements tracking. In this case, the z-scores resulted from all participants were more than 1.645 – a critical value for one-tail statistical testing with 5% significance level ($p < .05$). We also compared the computational time between the proposed algorithm and the state-of-the-art algorithm. Although there was an overhead in contour extraction and ellipse fitting stage (0.27 ms), the proposed algorithm was still appropriate for real-time processing. Overall comparison of processing time for a video frame is shown in Table 3.

Our experimental results show that the proposed algorithm achieves remarkable accuracy despite of extreme eyelid occlusion. Significant improvement has been achieved due to implementation of cumulative histogram in the adaptive binarization and removal of unwanted outliers during ellipse fitting. Compared with the state-of-the-art pupil localization method [22], our algorithm spent 20.3% longer computational time on processing a video frame (see Table 3). This longer computational time is due to implementation

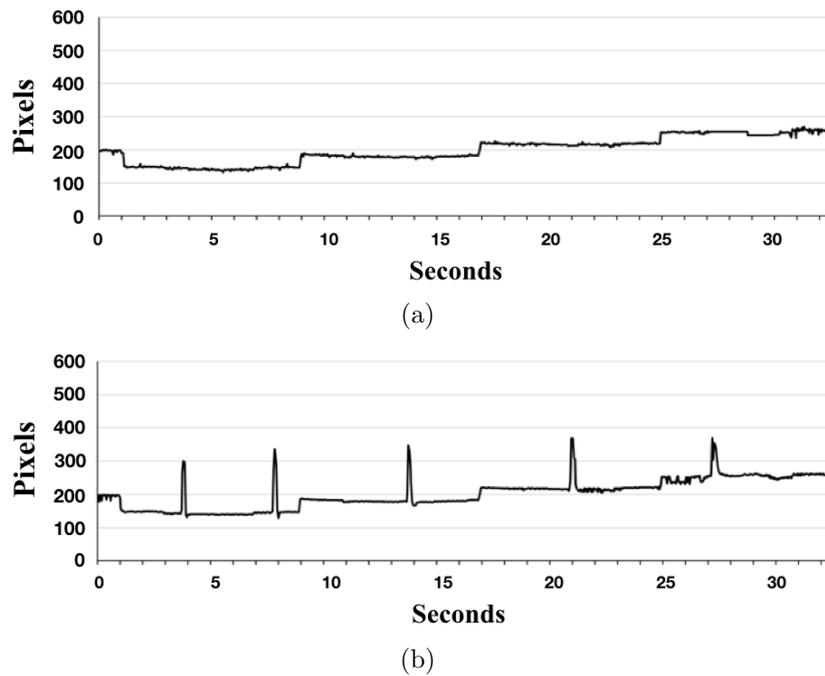


FIGURE 9. Sample of eye movement graph from the first participant during real-time video streaming. The graphs show comparison of pupil position (in pixels) in vertical direction using (a) the proposed algorithm; (b) a state-of-the-art pupil localization algorithm [22]. The vertical and horizontal axes are coordinate (in pixels) and timestamp (in seconds), respectively.

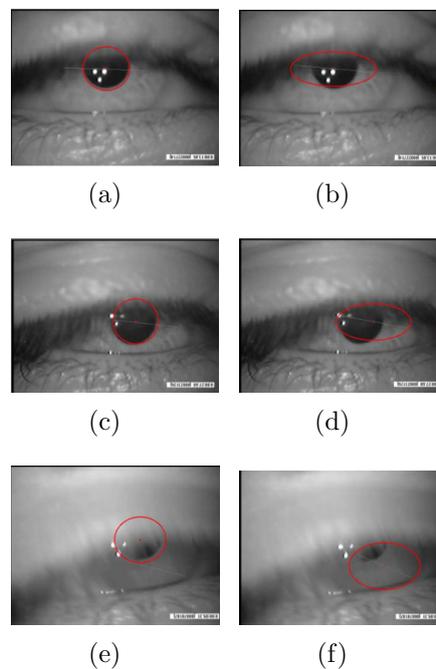


FIGURE 10. Comparative performance of pupil localization on a participant during real-time video streaming using the proposed algorithm (left panel) and a state-of-the-art algorithm [22] (right panel) at 14 s (panel (a) and (b)), 27 s (panel (c) and (d)), and at 21 s (panel (e) and (f)).

TABLE 2. Errors of pupil localization algorithms during real-time streaming (in pixels)

Participant	The proposed algorithm	The state-of-the-art algorithm [22]	Statistical z-score	Significance
1	12.33	19.80	5.09	$p < .05$
2	10.60	13.02	1.75	$p < .05$
3	19.87	21.87	2.95	$p < .05$

$N = 9000$ frames

TABLE 3. Comparative analysis of computational time on processing a video frame (in ms)

Process	The proposed algorithm	The state-of-the-art algorithm [22]
Masking of ROI	0.22	0.22
Binarization	0.10	0.11
Pupil center estimation	0.27	0.18
TOTAL	3.09	2.18

of RANSAC as outliers removal during the ellipse fitting (see Figure 1). As presented in Algorithm 2, RANSAC iteratively estimates its parameter until the data within ϵ exceed a certain threshold t . Despite of this longer computational time, our algorithm is still appropriate to be used in real-time processing.

Although the proposed algorithm yields better performance than the state-of-the-art pupil localization method, automatic binarization process incorporated in the pupil localization algorithm requires an empirical value of the ratio of pixels that composes dark area in the image. In this case, assumption has to be made prior to defining a threshold as a baseline of acceptable cumulative histogram. Based on our empirical analysis, we define the ratio of pixels that composes dark area around 3% out of the whole pixels. Furthermore, we found that outliers removal based on RANSAC is prone to inconsistency as RANSAC is a non-deterministic algorithm. In future, we plan to develop a better outliers removal technique to obtain more consistent results.

Another constraint of this study is an assumption made before implementing the proposed method to human participants. The participants should be healthy and wearing no prescription glasses nor contact lenses. Our algorithm has been designed to deal with extreme eyelid occlusion. On the other hand, occlusions caused by prescription glasses or contact lenses are yet to be considered. In reality, a particular clinical trial may require the participant to wear prescription glasses or contact lenses to see the presented stimulus clearly. In this case, the accuracy of the proposed algorithm may decrease significantly because the prescription glasses and the contact lenses produce unwanted light reflection [47]. In addition to this constraint, our proposed method requires the pupil to be captured on its most common shape – the camera of the VOG system is assumed to be positioned in front of the eye, with or without the usage of infrared reflector [11, 49, 50]. In this ideal case, the optical axis of the camera coincides with the optical axis of the eyeball. Our method is not designed to deal with off-axial camera position, in which the shape of the pupil may resemble an ellipse with various irregularities.

4. Conclusions. Pupil localization is an important part of a video-oculography (VOG) system. The accuracy of eye movements measurement likely depends on how the pupil

center is estimated in real time. Although previous works were able to estimate pupil center with various contrasts and various image resolutions, the accuracy of the state-of-the-art pupil localization algorithms decreases significantly when the pupil is extremely occluded by the eyelid. In this research, we propose a novel method to improve the accuracy of pupil localization in real-time VOG during extreme eyelid occlusion. We implement an adaptive image binarization based on cumulative histogram and an improved ellipse fitting algorithm with outliers removal. Experimental results with 90% eyelid occlusion show that the proposed algorithm increases the accuracy of pupil localization up to 51% compared with a state-of-the-art pupil localization technique. Additionally, implementation of the proposed algorithm in video streaming demonstrates appropriateness of the proposed method in a real-time VOG with significant accuracy improvement ($p < .05$). Our future research will be conducted in a larger clinical trial. Although the proposed algorithm yielded a good accuracy during real-time video streaming, a real clinical trial is needed to investigate the usefulness of the proposed algorithm during detection of various eye movements disorders. In this case, an experimental investigation involving subjects suffering from a balance disorder or a head injury may reveal how the proposed algorithm gives a significant impact in medical diagnosis.

REFERENCES

- [1] W. Zhang, M. L. Smith, L. N. Smith and A. Farooq, Eye center localization and gaze gesture recognition for human-computer interaction, *Journal of the Optical Society of America A*, vol.33, no.3, pp.314-325, 2016.
- [2] A. Soetedjo and I. K. Somawirata, Implementation of eye detection using dual camera on the embedded system, *International Journal of Innovative Computing, Information and Control*, vol.13, no.2, pp.397-409, 2017.
- [3] J. Chen, D. Chen, X. Li and K. Zhang, Towards improving social communication skills with multi-modal sensory information, *IEEE Trans. Industrial Informatics*, vol.10, no.1, pp.323-330, 2014.
- [4] Y. Cao, S. Miura, Y. Kobayashi, K. Kawamura, S. Sugano and M. G. Fujie, Pupil variation applied to the eye tracking control of an endoscopic manipulator, *IEEE Robotics and Automation Letters*, vol.1, no.1, pp.531-538, 2016.
- [5] B. B. Anderson, J. L. Jenkins, A. Vance, C. B. Kirwan and D. Eargle, Your memory is working against you: How eye tracking and memory explain habituation to security warnings, *Decision Support System*, vol.92, pp.3-13, 2016.
- [6] L. Wu, Z. Zhu, H. Cao and B. Li, Influence of information overload on operator's user experience of human-machine interface in LED manufacturing systems, *Cognition, Technology and Work*, vol.18, no.1, pp.161-173, 2016.
- [7] A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, 2nd Edition, Springer, 2007.
- [8] J. N. van der Geest and M. A. Frens, Recording eye movements with video-oculography and scleral search coils: A direct comparison of two methods, *Journal of Neuroscience Methods*, vol.114, no.2, pp.185-195, 2002.
- [9] K. Eibenberger, B. Eibenberger, D. C. Roberts, T. Haslwanter and J. P. Carey, A novel and inexpensive digital system for eye movement recordings using magnetic scleral search coils, *Medical and Biological Engineering and Computing*, vol.54, no.2, pp.421-430, 2016.
- [10] T. Haslwanter and A. H. Clarke, Eye movement measurement: Electro-oculography and video-oculography, *Handbook of Clinical Neurophysiology*, vol.9, pp.61-79, 2010.
- [11] S. Wibirama, S. Tungjitkusolmun and C. Pintavirooj, Dual-camera acquisition for accurate measurement of three-dimensional eye movements, *IEEEJ Trans. Electrical and Electronics Engineering*, vol.8, no.3, pp.238-246, 2013.
- [12] J. Otero-Millan, D. C. Roberts, A. Lasker, D. S. Zee and A. Kheradmand, Knowing what the brain is seeing in three dimensions: A novel, noninvasive, sensitive, accurate, and low-noise technique for measuring ocular torsion, *Journal of Vision*, vol.15, no.14, pp.1-15, 2015.
- [13] S. Wibirama and K. Hamamoto, 3D gaze tracking on stereoscopic display using optimized geometric method, *IEEEJ Trans. Electronics Information and Systems*, vol.134, no.3, pp.345-352, 2014.
- [14] R. I. Wang, B. Pelfrey, A. T. Duchowski and D. H. House, Online 3D gaze localization on stereoscopic displays, *ACM Trans. Applied Perception*, vol.11, no.1, pp.1-21, 2014.

- [15] S. T. Moore, I. S. Curthoys and S. G. McCoy, VTM – An image processing system for measuring ocular torsion, *Computer Methods and Programs in Biomedicine*, vol.35, pp.219-230, 1991.
- [16] D. Zhu, S. T. Moore and T. Raphan, Robust pupil center detection using a curvature algorithm, *Computer Methods and Programs in Biomedicine*, vol.59, no.3, pp.145-157, 1999.
- [17] A. Iijima, H. Minamitani and N. Ishikawa, Image analysis of quick phase eye movements in nystagmus with high-speed video system, *Medical and Biological Engineering and Computing*, vol.39, no.1, pp.2-7, 2001.
- [18] T. Yagi, Nystagmus as a sign of labyrinthine disorders – Three-dimensional analysis of nystagmus, *Clinical and Experimental Otorhinolaryngology*, vol.1, no.2, pp.63-74, 2008.
- [19] H. G. MacDougall, K. P. Weber, L. A. McGarvie, G. M. Halmagyi and I. S. Curthoys, The video head impulse test diagnostic accuracy in peripheral vestibulopathy, *Neurology*, vol.73, no.14, pp.1134-1141, 2009.
- [20] J. K. Ong and T. Haslwanter, Measuring torsional eye movements by tracking stable iris features, *Journal of Neuroscience Methods*, vol.192, no.2, pp.261-267, 2010.
- [21] A. B. Roig, M. M. Vidal, J. Espinosa, J. Perez, D. Mas and C. Illueca, Pupil detection and tracking for analysis of fixational eye micromovements, *Optik – International Journal for Light and Electron Optics*, vol.123, no.1, pp.11-15, 2012.
- [22] W. Swart, C. Scheffer and K. Schreve, A video-oculography based telemedicine system for automated nystagmus identification, *Journal of Medical Devices*, vol.7, no.3, 2013.
- [23] M. Barbosa and A. C. James, Joint iris boundary detection and fit: A real-time method for accurate pupil tracking, *Biomedical Optics Express*, vol.5, no.8, pp.2458-2470, 2014.
- [24] J. Espinosa, A. B. Roig, J. Pérez and D. Mas, A high-resolution binocular video-oculography system: Assessment of pupillary light reflex and detection of an early incomplete blink and an upward eye movement, *Biomedical Engineering Online*, vol.14, no.1, pp.1-12, 2015.
- [25] T. Charoenpong, P. Pattapisetwong and V. Mahasitthiwat, A new method to detect nystagmus for vertigo diagnosis system by eye movement velocity, *Proc. of the 14th IAPR International Conference on Machine Vision Applications (MVA)*, pp.174-177, 2015.
- [26] A. B. Slama, A. Mouelhi, H. Sahli, S. Manoubi, C. Mbarek, H. Trabelsi, F. Fnaiech and M. Sayadi, A new preprocessing parameter estimation based on geodesic active contour model for automatic vestibular neuritis diagnosis, *Artificial Intelligence in Medicine*, vol.80, pp.48-62, 2017.
- [27] Z. Ding, J. Luo and H. Deng, Accelerated exhaustive eye glints localization method for infrared video oculography, *Proc. of the 33rd Annual ACM Symposium on Applied Computing*, pp.620-627, 2018.
- [28] Y. Kong, S. Lee, J. Lee and Y. Nam, A head-mounted google-type video-oculography system for vestibular function testing, *EURASIP Journal on Image and Video Processing*, vol.2018, no.1, pp.1-10, 2018.
- [29] R. Krishnamoorthi and G. Annapoorani, A simple boundary extraction technique for irregular pupil localization with orthogonal polynomials, *Computer Vision and Image Understanding*, vol.116, no.2, pp.262-273, 2012.
- [30] S. A. H. Nair and P. Aruna, Comparison of DCT, SVD and BFOA based multimodal biometric watermarking systems, *Alexandria Engineering Journal*, vol.54, no.4, pp.1161-1174, 2015.
- [31] K. A. S. Suzuki, Topological structural analysis of digital binary image by border following, *Computer Vision, Graphics, and Image Processing*, vol.30, no.1, pp.32-46, 1985.
- [32] CASIA, *CASIA Iris Database*, <http://biometrics.idealtest.org/>.
- [33] T. Satriya, S. Wibirama and I. Ardiyanto, Robust pupil tracking algorithm based on ellipse fitting, *Proc. of International Symposium on Electronics and Smart Devices (ISESD)*, pp.253-257, 2017.
- [34] M. T. Setiawan, S. Wibirama and N. A. Setiawan, Robust pupil localization algorithm based on circular hough transform for extreme pupil occlusion, *Proc. of the 4th International Conference on Science and Technology (ICST)*, vol.1, pp.1-5, 2018.
- [35] CodeLaboratories, *CL-Eye Platform SDK*, <http://codelaboratories.com/products/eye/sdk/>.
- [36] M. A. Fischler and R. C. Bolles, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM*, vol.24, no.6, pp.381-395, 1981.
- [37] A. M. F. Wong, *Eye Movement Disorders*, 1st Edition, Oxford University Press, 2007.
- [38] L. S. Parnes, S. K. Agrawal and J. Atlas, Diagnosis and management of benign paroxysmal positional vertigo (BPPV), *Canadian Medical Association Journal*, vol.169, no.7, pp.681-693, 2003.

- [39] H. H. Elsanadiky and Y. A. Nour, Diagnosis and management of posterior semicircular canal benign paroxysmal positional vertigo: A practical approach, *Egyptian Journal of Ear, Nose, Throat and Allied Sciences*, vol.16, no.2, pp.161-166, 2015.
- [40] K. Jahn, T. Langhagen, A. S. Schroeder and F. Heinen, Vertigo and dizziness in childhood – Update on diagnosis and treatment, *Neuropediatrics*, vol.42, no.4, pp.129-134, 2011.
- [41] H. J. Kim, S. H. Park, J. S. Kim, J. W. Koo, C. Y. Kim, Y. H. Kim and J. H. Han, Bilaterally abnormal head impulse tests indicate a large cerebellopontine angle tumor, *Journal of Clinical Neurology*, vol.12, no.1, pp.65-74, 2016.
- [42] M. Nyström, I. Hooge and R. Andersson, Pupil size influences the eye-tracker signal during saccades, *Vision Research*, vol.121, pp.95-103, 2016.
- [43] K. W. Choe, R. Blake and S. H. Lee, Pupil size dynamics during fixation impact the accuracy and precision of video-based gaze estimation, *Vision Research*, vol.118, pp.48-59, 2016.
- [44] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*, Prentice Hall, 1998.
- [45] W. Fuhl, M. Tonsen, A. Bulling and E. Kasneci, Pupil detection for head-mounted eye tracking in the wild: An evaluation of the state of the art, *Machine Vision and Applications*, vol.27, no.8, pp.1275-1288, 2016.
- [46] C. D. McMurrough, V. Metsis, J. Rich and F. Makedon, An eye tracking dataset for point of gaze detection, *Proc. of Symposium on Eye Tracking Research and Applications (ETRA'12)*, pp.305-308, 2012.
- [47] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka and J. van de Weijer, *Eye Tracking: A Comprehensive Guide to Methods and Measures*, OUP Oxford, 2011.
- [48] H. W. Sorenson, Least-squares estimation: From Gauss to Kalman, *IEEE Spectrum*, vol.7, no.7, pp.63-68, 1970.
- [49] S. Wibirama and K. Hamamoto, Design and implementation of gaze tracking headgear for Nvidia 3D Vision®, *Proc. of the 2013 International Conference on Information Technology and Electrical Engineering (ICITEE)*, pp.84-87, 2013.
- [50] S. Wibirama, H. A. Nugroho and K. Hamamoto, Evaluating 3D gaze tracking in virtual space: A computer graphics approach, *Entertainment Computing*, vol.21, pp.11-17, 2017.