DOI 10.34658/jacs.2018.26.2.201-212

JOURNAL OF APPLIED
COMPUTER SCIENCE
Vol. 26 No. 2 (2018), pp. 201-211

Pupil Detection Methods for Eye Tracking

Filip Rynkiewicz [1], Marcin Daszuta [1], Piotr Napieralski[1]

[1]*Lodz University of Technology*
Institute of Information Technology
ul. Wólczańska 215, 90-924 Lodz, Poland
piotr.napieralski@p.lodz.pl

Abstract. *Abstract. The theory that pupillary changes can reflect human psychophysical character was confirmed in several conducted studies. Authors of those studies have used devices called eye trackers in order to measure pupillary changes, blinking rate, eye movement patterns or even a gaze position. The main goal of those apparatuses is to locate the pupil on a given image using implemented pupil detection algorithms. In this paper the main focus will be set on several states of art pupil detection methods. All represent different approaches to performing the same tasks in difficult situations for ex.: low/high condition, low image resolutions or low contrast.*
Keywords: *pupil, eye tracking,.*

1. Introduction

Biometrics, the measuring of biological data in order to get detailed information about a person's condition, can be used in multiple areas. For human eye biometrics, that is, the anatomy of human eye 1, vastly different features can be extrapolated. Blink rate measurement can provide information about possible fatigue [1] or even detect preliminary Parkinson's disease [2]. Iris scans are used in identity authentication [3] or predict obstructive lung disease [4]. In the case of saccades, a dysfunction in the area of rapid movement of the eye between phases of

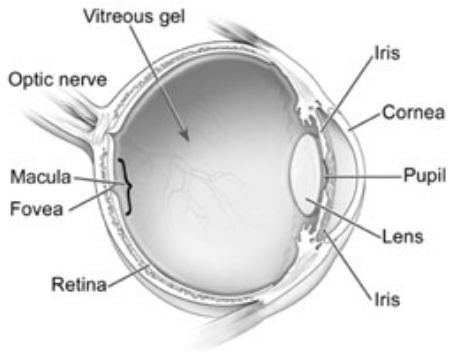
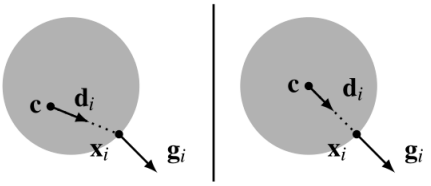


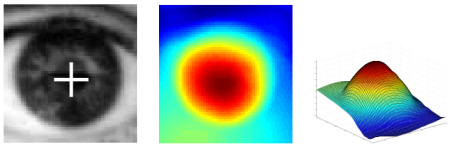
Figure 1: The anatomy of the human eye. Source `http://www.afb.org/image.ashx?ImageID=3484`

fixation [5],it can be correlated with Huntington's [6] or Alzheimer's disease [7]. In 1628, Harvey's discourse on circulation of the blood noted: "*In anger the eyes are fiery, and the pupils contracted" ("Irâ rubent oculi, constringitur pupilla)*". From that time. pupil dilatation was strongly correlated with the emotional state [8, 9, 10, 11].

In order to observe these changes, biometric devices called *eye trackers* can be used. Two types of those apparatuses can be specified. The first, is standalone, which can be placed far from the examinee and consist of camera that is not pointed directly at human eye, but rather directed at the face of the examinee. The second one comes in the form of glasses that have cameras pointed directly at each eye with an additional global camera to record everything that is in front of examined. Most types of eye trackers use *near infra red cameras*, in order to take advantage of *dark pupil eye tracking*. This technique causes the pupil to appear darkest in the picture with a clear separation from the iris to speed up further calculations. This apparatus provides a wide range of possible measurements of the human eye, such as blink rates, pupil dilatation, point of gaze, eye position or eye movement. The primary aim of this article is to present the most common state of art algorithms used in eye trackers to localize the pupil that can be used on both types of the device. For standalone devices, a few additional steps must be implemented, because the camera can record the whole upper body of a person. The first step is



(a) Example of displacement vector d_i and the gradient vector g_i calculation. On right have the same orientation, on left not. Source [12]



(b) Calculation of maximum Equ. 1 for possible center (white cross). Source [12]

Figure 2: Gradient approach

to recognize the face on this image, and then locate the position of the eyes. For the second type of device, when the camera is centered on the eyes, those steps can be omitted.

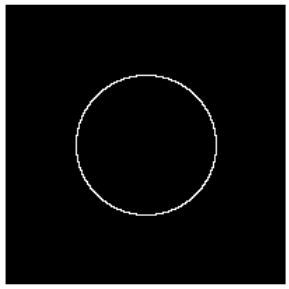
Input for all presented algorithms are infrared images in 8bit gray scale, and it is assumed that on this image a human eye is presented. The first step for all methods is to locate the eye on image, which is not the topic of this article, and then in each specific way, to locate the pupil.

2. Gradient approach

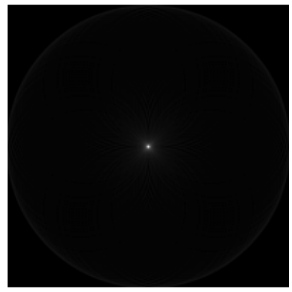
The input image for this algorithm [12] must include only the eye image, thus it is suitable for both types of *eye trackers*, and it is assumed that the pupil on image must have an almost circular shape. The basic idea behind this is gradient displacement and can be described in mathematical terms as follows.

Let c be a possible center and g_i the gradient vector at position x_i .The normalized displacement vector d_i must have the same orientation, except for the sign, as the gradient g_i, (Fig. 2a). If the vector field of gradients will be taken into consideration as a dot product between the normalized displacement vector and the gradient vectors g_i, the optimal center c of a circular object in an image with pixel positions x_i ,$i \in \{1..., N\}$ is given by

$$c^* = \underset{c}{argmax} \left\{ \frac{1}{N} \sum_{i=1}^{N} (d_i^T g_i)^2 \right\} \tag{1}$$



(a) Input image for CHT.



(b) Hough accumulator. Intensity of white color is corresponded to vote value.

Figure 3: Voting approach

$$d_i = \frac{x_i - c}{\|x_i - c\|_2}, \forall_i : \|g_i\|_2 = 1 \tag{2}$$

Figure 2b presents the plot of Equation 1. There is an assumption that the maximum response will be threated as the center of circular object, in this case, the pupil.

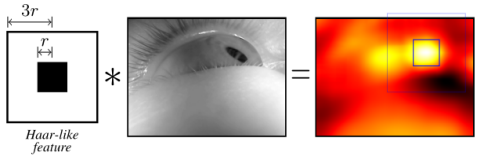
3. Voting approach

Starting with the work of Paul Hough in 1962 about the detection features of a particular shape, multiple approaches have diverged into line detection, circle detection, curve detection or even 3D objects detection. As in the previous algorithm, these calculations are suitable for both types of eye trackers, and they only detect circular objects. Advantage of this approach over the previous one is that can give more information on the output. It detects all the features of a circular shape, the position of the center and the radius.

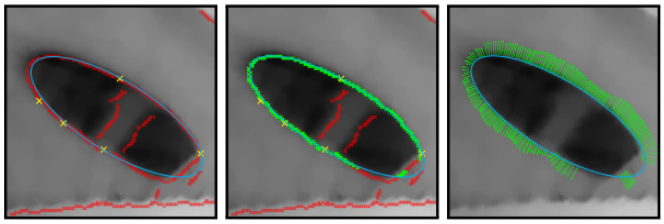
Basic Circular Hought Transform (CHT) can be represented as 3 parameter space over N dimensional image. The circle is described by radius r and pair x y, corresponded to the center position, Equation 3.

$$(x - a)^2 + (y - b)^2 - r^2 = 0 \tag{3}$$

All points of circle C on the image are transformed as the center of new circle C_n with radius r. For each r a two-dimensional array with the same dimensions as the input image is created, a so-called *Hough accumulator* (HA). In the same HA for



(a) Eye image is convolved with Haar-like feature for radius r. Maximal response over all pixels and radii is considered as pupil. Source [13]



(b) Rasnac method.

Figure 4: Convolution and fitting approach. Source [13]

each intersection of circles C_n, indexes corresponding to intersection position is incremented by one, this action is called *voting* procedure. After all intersection for all r are checked, indexes with the highest voting point is sought. Position of indexes with the highest value of vote is treated as the center of the circle (Fig. 3b). When we know what HA gives its highest vote, the r can be extracted.

4. Convolution and fitting approach

Based on Haar-like feature [14] (Fig. 4a), and RASNAC [15] fitting algorithms (Fig.4b), *Swirski* [13] presented a new approach for pupil localization in real-life scenarios. It is strictly designed for glasses-like eye trackers and provides an advantage over other ideas presented herein. It more suitable for a mobile device in various scenarios like sports, automotive or marketing, when the mobility of the device is a must. And it focuses only on the localization of the ellipse with no information about its center.

First step of this algorithm is to use, as previously mentioned, a Haar-like feature convolved with the input image, to tighten the boundaries of the input image as much as possible in order to minimize further calculations. Output of this convolution is squared region P on the image (Fig. 4b). P is segmented using k-means

clusters to get two regions, dark and light, correspondent to pupil pixels and background intensity values. The dark sector is assumed to be the rough estimation of the pupil position.

For the removal of small bright gaps in the black pupil region, *morphological open* technique is used. This step is important for further optimization and correctness of the algorithm; this action removes pixels that cannot be part of the pupil contour.

Instead of a voting system, like the previously described CHT, this approach uses a search-based method. Based on extrapolated pixels the eclipse can be fit, using *RASNAC* . This technique can sort out so called *inliers* and *outliners* (Fig. 4b). The first one can be described by data that is correlated somehow with model parameters, the second one consists of data that does not fit into the model. On Figure 4b first picture represents ellipse fitted with 5 initial points. Next, the search of inliers for this ellipse is executed and the ellipse is then refit with new inliers. When the magnitude of the image gradient at each inlier is orthogonal to the ellipse, it is assumed that this is the best fitted shape.

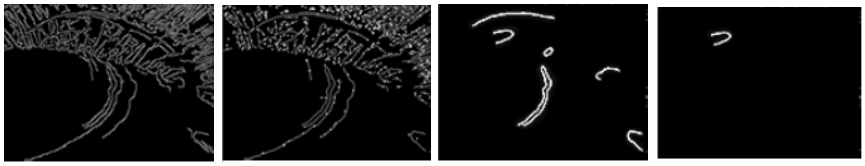
5. The divided condition approach

The last algorithms [16] called ExCuSe focus on extracting specific points that can be treated as the edge of the pupil. It has the highest advantage over all the methodology presented herein. It provides special steps for the most challenging input images such as big bright spots called *glints*. To detect them, the histogram of the input image has to be calculated first. These values are analyzed to search for the highest peaks, valuing above the threshold $th_1 = 200$, throughout the histogram. After this step, the algorithm diverges into two different paths.

5.1. An image with bright area peak

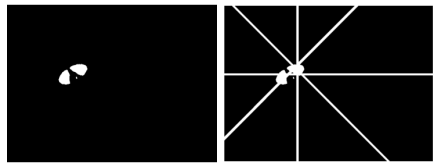
For this case, a Canny filter is applied to extrapolate all the possible edges on the image. Then, the 4-step filtering of unwanted edges is done:

- All neighboring pixels are removed, which (considered vectors) have angels greater than 90° between each other.
- Distinguishing straight and curved lines, thinning them using morphological operations.



(a) (b) (c) (d)

Figure 5: *(a)* contain image after Canny filter was applied. All thin lines are removed and centroids calulated on *(b)*. All straight lines removed on *(c)*. Final curve *(d)*.



(a) (b)

Figure 6: *(a)* thresholded input image, *(b)* white lines represents calculated AIPF.

- Removing all straight lines calculating their centroid and comparing that to line segments. If the distance between centroid and line segment is smaller then $di_1 = 3$, the line is assumed to be straight.

- For each curved line left, the values of gray intensity are calculated along it. The longest curve in the darkest area is picked as most probable to be part of the pupil edge.

The last step in this branch is to used to *direct least squares method* to fit the ellipse into gathered points on the calculated edge.

5.2. Image without bright area peek

First step in this case is to extract coarse pupil location by thresholding the image by th_2. This value is equal of half the standard deviation of the image intensity. This part removes 90% of unwanted pixels on image. For 0° 45° 90° 135° the Angular Integral Projection Function [17] is calculated, described by Equation 4

$$AIPF(\Theta, p, h) = \frac{1}{h+1} * \int_{j=-\frac{h}{2}}^{\frac{h}{2}} I\Big((x_0 + p\cos\Theta) + (j\cos(\Theta + 90^\circ)), (y_0 + p\sin\Theta) + (j\sin(\Theta + 90^\circ))\Big)\mathrm{d}j \quad (4)$$

Output of this function can be represented as a histogram and the maximum value of each one is searched (Fig. 4(b)). When those white lines intersect, the points between the intersection are treated as a rough estimation of the area where the pupil can be found. The image is then truncated to the size of a rough pupil area, and again resized by 10% or the original input image. This step is crucial because the rough estimation and resize can cut an important part of image where pupil might be. These steps gave the output of possible pupil area P on the input image. For P, the previous Canny filter values are thresholded by the value of the full standard deviation of image. The edges of the edge image are preselected by overlay with the threshold image. The next step is to get only the edges that are close to the border of the threshold region. Irrelevant edges are removed by the previously described morphological operations, which in output provides a curve. The assumption at this point is that the center of P is contained inside the pupil. With this information, rays are shot from the center of P. When they intersect, any remaining edges intersection points are treated as input to the eclipse fitting algorithm.

6. Results

All algorithms were compared on the data sets provided by ExCuSe which can be obtained at `http://www.ti.uni-tuebingen.de/Pupil-detection.1827.0.html`.

As it can be seen in Table 1 the ExCuSe algorithm has the highest detection rate value which provides two ways of the localization of the pupil. That separation gives a high advantage in pictures with very bright spots. Following this, the Swirski can handle highly o axis images and small obstacles like eyelashes but cannot handle high value of illuminations or reflection. For the gradient approach and Hough Circular Transform, the detection rate for pupil localization is the lowest one, and it mostly depends on light condition and the position of the eye relative

Data Set	Images	Challenges	Swirski	ExCuSe	Gradient	Hough
Swirski	600	Highly off axis, eyelashes	74%	86%	63%	11%
I	6.554	Reflections	3%	62%	34%	8%
II	505	Reflections, bad illumination	28%	31%	7%	13%
III	9.799	Reflections, recording errors, bad illumination	11%	31%	4%	7%

Table 1: Discussion of the presented method

to the camera. If it is assumed that pupil is elliptical in shape, circle detection will always have a smaller precision rate.

7. Conclusion

The trend of the newest and more precise algorithms are directed to the use of glass-like eye trackers, rather than standalone ones. This is dictated by the usage of eye trackers as mobile devices for real-life scenarios, and with the position of cameras directed directly on the eyes. This configuration simplifies calculation and the step of face recognition can be omitted.

Algorithms presented in this paper provide information about how many approaches can be taken to complete the same task. From basing the approach on gradient information, through a vote-based system, such as Hough Transform, to search-based approaches like ExCuSe or Swirski. Each one bases their solution and ideas on information provided by previous researchers. For more real-life scenarios when the pupil must be consider more like an ellipse than a circle, ExCuSe and Swirski are more relevant. In the case of high glints, reflection and more complicated real-life scenarios, the best algorithm is the ExCuSe.

References

[1] Haq, Z. A. and Hasan, Z., *Eye-blink rate detection for fatigue determination*, In: 2016 1st India International Conference on Information Processing (IICIP), Aug 2016, pp. 1–5.

[2] Fitzpatrick, E., Hohl, N., Silburn, P., O'Gorman, C., and Broadley, S. A., *Case–control study of blink rate in Parkinson's disease under different conditions*, Journal of Neurology, Vol. 259, No. 4, Apr 2012, pp. 739–744.

[3] Thomas, T., George, A., and Indira Devi, K., *Effective Iris Recognition System*, Procedia Technology, Vol. 25, 12 2016, pp. 464–472.

[4] Bansal, A., Agarwal, R., and K. Sharma, R., *Pre-Diagnostic Tool to Predict Obstructive Lung Diseases Using Iris Recognition System*, 01 2019, pp. 71–79.

[5] Cassin, B. and Rubin, M., *Dictionary of Eye Terminology*, Triad Pub., 2006.

[6] García-Ruiz, P., Fontán, A., Cenjor, C., Ulmer, E., Gómez-Tortosa, E., Sanchez Pernaute, R., Rojo, A., and Yébenes, J., *[Disorders in gaze saccades in Huntington disease. Clinical correlations]*, Der Nervenarzt, Vol. 72, 06 2001, pp. 437–40.

[7] Molitor, R., Ko, P., and Ally, B., *Eye Movements in Alzheimer's Disease*, Journal of Alzheimer's disease : JAD, Vol. 44, 09 2014.

[8] Bradley, M. M., Miccoli, L., Escrig, M. A., and Lang, P. J., *The pupil as a measure of emotional arousal and autonomic activation*, Psychophysiology, Vol. 45, No. 4, Jul 2008, pp. 602–607, 18282202[pmid].

[9] Leknes, S., Wessberg, J., Ellingsen, D.-M., Chelnokova, O., Olausson, H., and Laeng, B., *Oxytocin enhances pupil dilation and sensitivity to 'hidden' emotional expressions*, Social Cognitive and Affective Neuroscience, Vol. 8, No. 7, 2013, pp. 741–749.

[10] Silk, J. S., Dahl, R. E., Ryan, N. D., Forbes, E. E., Axelson, D. A., Birmaher, B., and Siegle, G. J., *Pupillary reactivity to emotional information in child and adolescent depression: links to clinical and ecological measures*, Am J Psychiatry, Vol. 164, No. 12, Dec 2007, pp. 1873–1880, 18056243[pmid].

[11] Vasquez-Rosati, A., Brunetti, E. P., Cordero, C., and Maldonado, P. E., *Pupillary Response to Negative Emotional Stimuli Is Differentially Affected in Meditation Practitioners*, Front Hum Neurosci, Vol. 11, May 2017, pp. 209–209, 28515685[pmid].

[12] Timm, F. and Barth, E., *Accurate Eye Centre Localisation by Means of Gradients*, In: VISAPP, 2011.

[13] Świrski, L., Bulling, A., and Dodgson, N. A., *Robust real-time pupil tracking in highly off-axis images*, In: Proceedings of ETRA, March 2012.

[14] Viola, P. and Jones, M., *Rapid object detection using a boosted cascade of simple features*, In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Vol. 1, Dec 2001, pp. I–I.

[15] Fischler, M. A. and Bolles, R. C., *Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography*, Commun. ACM, Vol. 24, No. 6, June 1981, pp. 381–395.

[16] Fuhl, W., Kübler, T., Sippel, K., Rosenstiel, W., and Kasneci, E., *ExCuSe: Robust Pupil Detection in Real-World Scenarios*, Vol. 9256, 09 2015.

[17] Mohammed, G., Hong, B., and Alkazzaz, A., *Accurate Pupil Features Extraction Based on New Projection Function.* Computing and Informatics, Vol. 29, 01 2010, pp. 663–680.