
This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Visual focus of attention estimation using eye center localization

PLEASE CITE THE PUBLISHED VERSION

PUBLISHER

IEEE

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

LICENCE

All Rights Reserved

REPOSITORY RECORD

Cai, Haibin, Bangli Liu, Jianhua Zhang, Shengyong Chen, and Honghai Liu. 2019. "Visual Focus of Attention Estimation Using Eye Center Localization". figshare. <https://hdl.handle.net/2134/9955508.v1>.

Visual Focus of Attention Estimation Using Eye Center Localization

Haibin Cai, Bangli Liu, Jianhua Zhang, *Member, IEEE*, Shengyong Chen, *Senior Member, IEEE*, and
Honghai Liu, *Senior Member, IEEE*

Abstract—Estimating people visual focus of attention (VFOA) plays a crucial role in various practical systems such as human–robot interaction. It is challenging to extract the cue of the VFOA of a person due to the difficulty of recognizing gaze directionality. In this paper, we propose an improved integrodifferential approach to represent gaze via efficiently and accurately localizing the eye center in lower resolution image. The proposed method takes advantage of the drastic intensity changes between the iris and the sclera and the grayscale of the eye center as well. The number of kernels is optimized to convolute the original eye region image, and the eye center is located via searching the maximum ratio derivative of the neighbor curve magnitudes in the convolution image. Experimental results confirm that the algorithm outperforms the state-of-the-art methods in terms of computational cost, accuracy, and robustness to illumination changes.

Index Terms—Convolution, eye center localization, human–robot interaction, integrodifferential, visual focus of attention (VFOA).

I. INTRODUCTION

IT is evident that the visual focus of attention (VFOA) of a person is a strong indicator for assisting communication in human–robot interactions and cognitive science. Developing computational attention models has many potential applications such as surveillance, automatic target detection, navigational aids, and robotic control [1]. VFOA, as a perceptual process, has been embodied in the polyscheme cognitive architecture which integrates different computational mechanisms to achieve human-level intelligence [2]. Moreover, the polyscheme model is adapted to construct models of human–

robot interaction [3]. Moreover, the gaze of people is also one of the key issues for humanoid robots to acquire human-like intelligence [4]. The goal of this research is to develop methods for VFOA. We see this work as taking steps toward developing robotics systems with human-like intelligence.

Various VFOA systems have been developed in the literature, such as Stiefelwagen *et al.* [5] (where VFOA during face-to-face communication in a meeting is estimated by using head pose), Danninger *et al.* [6] (where VFOA in office environment is estimated through head orientation data), and Katzenmaier *et al.* [7] (where the addressee in human–human–robot interaction is identified by combining head pose and acoustic). However, accurate VFOA is determined by the eye gaze [8]. The VFOA has inherent relation with gaze directionality that is usually decoded from eye movement. On the other hand, information on eye movement has connection to the emotional, cognitive, and psychological states of a human. Due to the difficulty of locating eye movement in a 3-D space, hence human-like intelligence is required to interpret the information behind eye movement with continuous trajectories of eye center position [9].

This paper gives priority to eye center location in that it has become an active research topic owing to its wide applications in face recognition, human–robot interaction, human–machine interfaces, psychology, and cognitive linguistics [10]. Although high-accuracy eye center location can be obtained through high-quality eye-gaze tracking systems, the intrusive or expensive devices make these existing methods unattractive [11].

Generally speaking, the reported eye center detection methods can be classified into three categories, namely, appearance-based, shape-based, and hybrid methods. Appearance-based methods, known as holistic methods or image template methods, use an entire eye image as input. Jesorsky *et al.* present a method of detecting face in an image using Hausdorff distance as a similarity measure and then locate the pupil by a multilayer perceptron trained with pupil centered images [12]. Cristinacce *et al.* employ a coarse-to-fine method which combines a face detector, feature detectors, and pairwise reinforcement of feature responses to find the features on the human face [13]. Hamouz *et al.* identify ten feature points on the face through Gabor filters and an appearance model [14]. Niu *et al.* use two cascade classifiers which are designed by bootstrapping positive samples and negative samples, respectively, to locate the eye [15]. Kroon *et al.* [16] locate the eye position by searching for the maximum response of a trained fisher linear discriminant classifier. Markus *et al.*, with an assumption of the pupil localization as regression based on low-level image

H. Cai is with the Institution of Intelligent Systems, College of Computer Science, Zhejiang University of Technology, Hangzhou 310014, China, and also with the Intelligent Systems and Biomedical Robotics Group, The School of Computing, University of Portsmouth, Portsmouth PO1 3HE, U.K. (e-mail: haibin.cai@port.ac.uk).

B. Liu is with the School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China (e-mail: jlliubangli@gmail.com).

J. Zhang and S. Chen are with the Institution of Intelligent Systems, College of Computer Science, Zhejiang University of Technology, Hangzhou 310014, China (e-mail: zjh@zjut.edu.cn; sy@ieee.org).

H. Liu is with the Intelligent Systems and Biomedical Robotics Group, The School of Computing, University of Portsmouth, Portsmouth PO1 3HE, U.K., and also with the State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: honghai.liu@icloud.com).

features, apply an ensemble of randomized regression trees to achieve high accuracy [17]. Kim *et al.* propose an eye location method based on multiscale Gabor feature vectors [18].

Unlike appearance-based methods, shape-based methods make use of the geometric characteristics of the eye or its surroundings. The shapes of the eye, limbus, and eye corner are identified as commonly used features. Asteriadis *et al.* employ the length and slope information of pixel vectors which point to the closest pixel of an edge eye map to locate the eye center [19]. Bai *et al.* present a radial symmetry measure derived from generalized symmetry transform proposed by Reisfeld *et al.* to locate the eye center [20], [21]. Timm *et al.* propose a gradient-based method utilizing the dot product of image gradient vector and displacement vector for eye center location [22]. Valenti *et al.* take advantage of the isophote curvature where curves connecting points are of equal intensity to design a voting scheme for pupil localization [23]. Asadifard *et al.* calculate the histogram cumulative density function of the eye region, followed by a minimum intensity pixel filter to locate the eye center [24]. Daugman employs the intensity change between the iris and the sclera and localize the iris boundaries using an integrodifferential operator, i.e., IDO [25]. Due to its efficiency, IDO-based methods has become one of the most popular iris location algorithms in the area of iris biometrics [26]. However, IDO-based iris localization methods are computationally more expensive than the other methods and require high resolution with a minimum of 50 pixels in iris radius [27], [28].

Hybrid methods combine the advantages of appearance-based methods and shape-based methods to achieve better performance. Turkan *et al.* employ edge projections and support vector machine based classifier to estimate the most probable location of each eye [29]. Campadelli *et al.* locate the eye position by using two support vector machines trained on properly selected Haar wavelet coefficients [30]. Valenti *et al.* add scale invariance to their former isophote-based method using a scale space pyramid and finally determine the eye location by matching the SIFT vector of each candidate with a database [11].

This paper presents an improved integrodifferential solution to localize the eye center. The proposed method is computationally much cheaper than the original integrodifferential method and also achieves higher accuracy in relatively lower resolution images. The rest of this paper is organized as follows: a brief introduction of the integrodifferential operator is presented in Section II. Section III describes the detail of the proposed method. Section IV presents the data set used in this study along with the accuracy measure method and the experimental results. Finally, this paper is concluded with discussions in Section V.

II. INTEGRODIFFERENTIAL OPERATOR

It is evident that Daugman's method has been highly cited, especially in iris recognition area [25]. It is assumed that an iris is with a circular form and the integrodifferential operator is applied as defined

$$\max_{(r, x_0, y_0)} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r, x_0, y_0} \frac{I(x, y)}{2\pi r} ds \right| \quad (1)$$

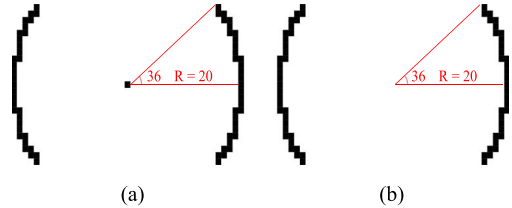


Fig. 1. (a) Kernel with center weight at a radius size of 20. (b) Kernel without center weight at a radius size of 20.

where $G_\sigma(r)$ is a smoothing Gaussian with a scale of σ , $*$ represents convolution, I is the image of an eye, and ds is the contour of the r radius circle with the center point of (x_0, y_0) . The operator searches for the maximum along the circle path in the blurred image via the Gaussian kernel, partial derivative with respect to increasing radius r . In order to deal with the obscure of the upper and lower limbi by the eyelids, the angular arc of contour integration s is restricted in range to two opposing 90 cones centered on the horizon. Moreover, the computation complexity of this method is $O(X, Y, R)$ [31], where X and Y are the number of rows and columns of the center coordinate and R is the counts of increasing radius.

In discrete implementation of the integrodifferential operator, the order of convolution and differentiation is interchanged and concatenated to improve the speed. After replacing the convolution and contour integrals with sums, the equation is derived as follows:

$$\max_{(n\Delta r, x_0, y_0)} \left| \frac{1}{\Delta r} \sum_k \left\{ (G_\sigma((n-k)\Delta r) - G_\sigma((n-k-1)\Delta r)) \right. \right. \\ \left. \left. \times \sum_m I[(k\Delta r \cos(m\Delta\theta) + x_0), (k\Delta r \sin(m\Delta\theta) + y_0)] \right\} \right| \quad (2)$$

where Δr means small increment in radius and $\Delta\theta$ is the angular sampling interval along the circular arcs.

III. PROPOSED ALGORITHM

Although an angular sampling interval $\Delta\theta$ is used to find points along the circular arc ds in (2), it makes it challenging in choosing the appropriate value of $\Delta\theta$. If $\Delta\theta$ is too small, the computational cost can be very high; otherwise, the accuracy decreases. It should be noted that the original integrodifferential operator only uses the optimization of the circle curve integral of the gradient magnitudes, and the center point of the pupil, another important feature information, is not taken into account. The proposed method utilizes all of the pixels along the circle by convoluting different sizes of circle kernels with the eye region image. The grayscale of the eye center is considered by designing the kernel with a weight in the center point. The size of the kernel is $2r + 1$, where r stands for the radius of the circle. Furthermore, the pixels along the circle are assigned a normalized value. Fig. 1 shows the reverse value of the kernel, where the black squares represent pixels with a weight and the white part pixels mean a value of zero. In order to cope with the

obscure of eye lids, the upper and lower parts are not assigned, and the angle of the arc is set to be 36.

Instead of using a differential method at the integral of circle intensity, this paper calculates a ratio derivative between neighbor curve magnitudes. The ratio derivative is formulated as follows:

$$\begin{cases} I_r = K_r * I \\ I'_{r+1} = K'_{r+1} * I \\ D_r = \frac{I'_{r+1}}{I_r} \\ \arg \max_{(r,x,y)}(D_r) \\ r \in [r_{\min}, r_{\max}] \end{cases} \quad (3)$$

where K_r is the kernel with a center weight and r stands for the radius of the circle inside the kernel. The kernel without a center weight is represented as K'_{r+1} , whose radius is $r + 1$. I_r and I'_{r+1} are the results of the convolution of the different kernels with eye image I . D_r stands for the ratio derivative calculated by the division of the convolution result image. r_{\min} and r_{\max} are set according to the size of the eye image representing the minimum and maximum of the radius r . The weights of the points around the circular arcs are of equal value and normalized to 1, and the weight of the center point is settled to a valid value. In order to locate the eye center and radius, the proposed method searches the maximum of different radii of D_r , and the smoothing function in the original integrodifferential operator is not employed. Employing FFT in the realization of convolution, the computation complexity can be reduced, where only 2-cycle DFT and 1-cycle IDFT are required. The computational complexity of this method is $O(P \log_2(P)R)$, where counts P need to satisfy $P \leq X + Y + C$. X, Y are the number of rows and columns of the center coordinate, C is a constant number, and R is the number of increasing radius.

IV. EXPERIMENTAL RESULTS

The boosted cascade face detector is employed with default parameters in order to obtain the approximate location of the face [32], and then, rough eye regions are extracted through anthropometric relations with the face as stated in [11] and [17].

A. Data Sets

The proposed method is validated on the BioID face database [33] which consists of 1521 grayscale images of 23 different people with a resolution of 384×288 pixels [27]. The images in the database are head and shoulder frontal view images with a large variety of illumination, background, scale, and pose. Some people in the database are wearing glasses. In some images, the eyes are closed or completely hidden by reflections on the glasses. Because of these conditions, the BioID database is considered to be challenging and realistic. Some snapshots are provided in Figs. 2 and 3, and the number of detected images is 1469.

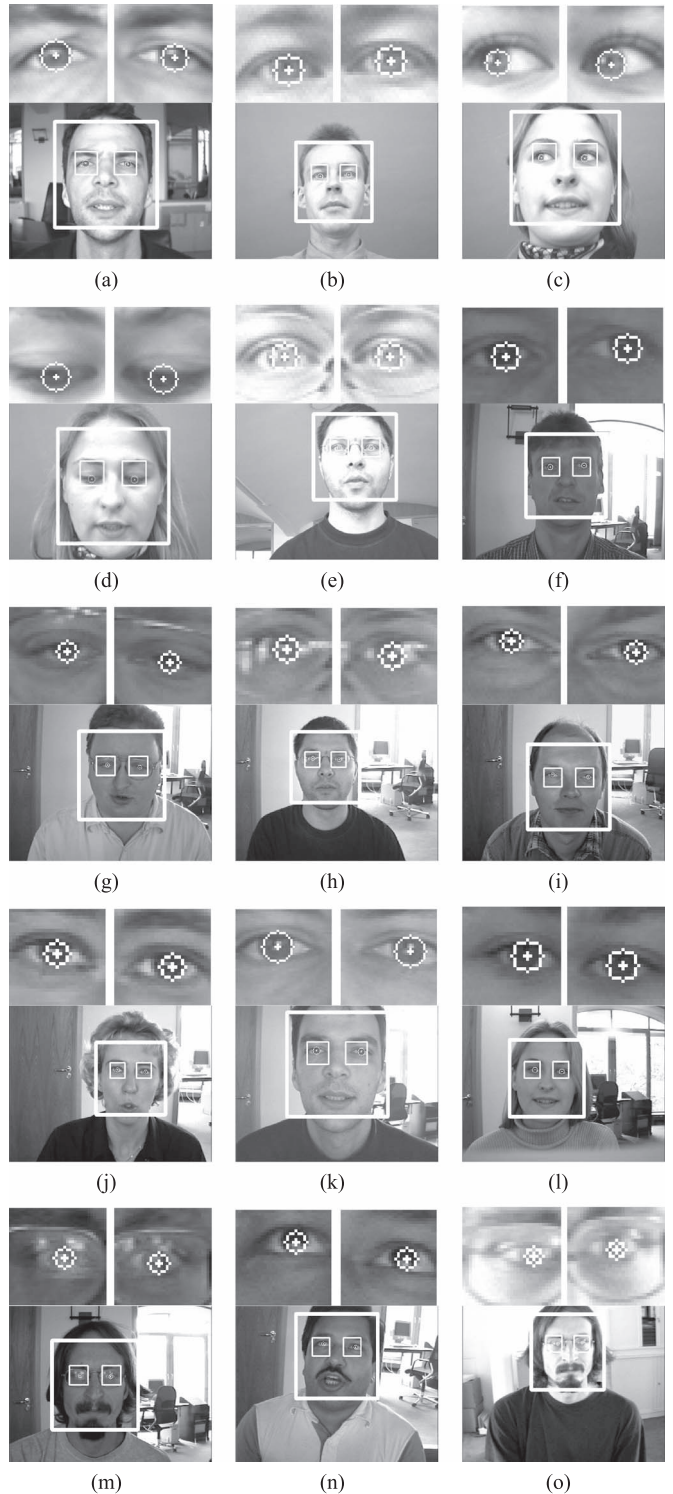


Fig. 2. Snapshots with accurate eye center estimation.

B. Localization Accuracy

The accuracy measure of eye location is calculated in normalized error which records the maximum error of both eye points. The measure was introduced by Jesorsky *et al.* [12], and it is defined as follows:

$$e = \frac{\max(d_l, d_r)}{d} \quad (4)$$

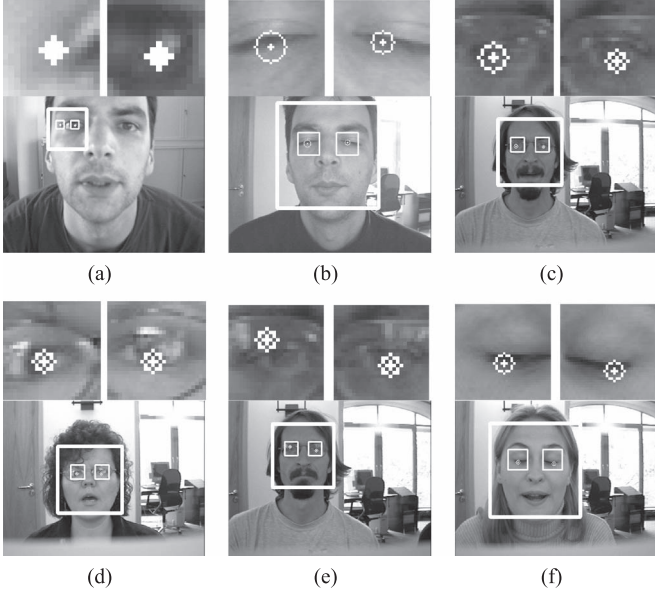


Fig. 3. Snapshots with inaccurate eye estimation.

where d_l and d_r are the Euclidean distances between the detected left and right eye centers and the ones in the ground truth and d is the Euclidean distance between the left and right eyes in the ground truth. Herein, a relative error of $e \leq 0.25$ is equal to a distance of half an eye width, $e \leq 0.1$ means the diameter of the iris, and $e \leq 0.05$ corresponds to the length of the pupil. In order to compare the proposed method with other published methods which also employ the normalized error measure, the minimum normalized error which is calculated by the minimum of d_l and d_r , and the average normalized error which is obtained by the average of d_l and d_r , are also calculated.

C. Results

Figs. 2 and 3 present some snapshots with accurate eye center locations and inaccurate eye locations, respectively, using the proposed method. In Fig. 2, the proposed method is able to deal with some slight head pose changes, changing illumination, shadows, squint, reflection of the light from the iris, and presence of glasses. On the other hand, the proposed method failed to localize the eye center where the face is wrongly localized, eyes are closed, or strong highlights exist because of the glasses, as shown in Fig. 3.

D. Comparison With the Original Method

The performances of the proposed method were compared with the original integrodifferential method in terms of computation and accuracy factors. Figs. 4 and 5 show the quantitative accuracy results of the original integrodifferential method and the proposed method tested on the BioID database. It can be observed that the modified method significantly improves the accuracy for both $e \leq 0.05$ and $e \leq 0.1$ regardless of the minimum or maximum normalized error standard. The impact of the size of $\Delta\theta$, where the circle is divided into n parts

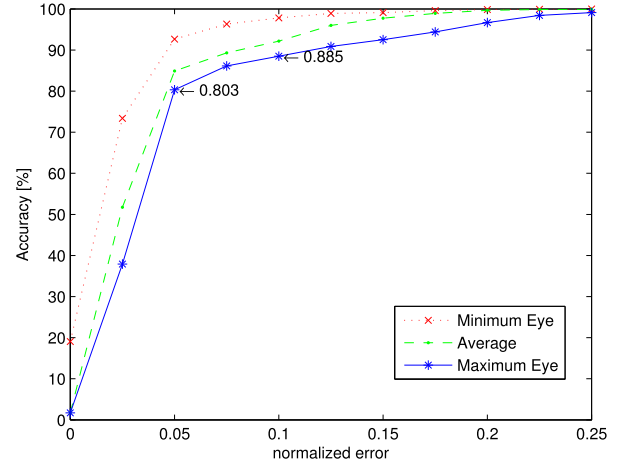


Fig. 4. Accuracy of the original integrodifferential method ($n = 60$) on the BioID database. The three lines correspond to the minimum, average, and maximum normalized errors from the top to the bottom, respectively.

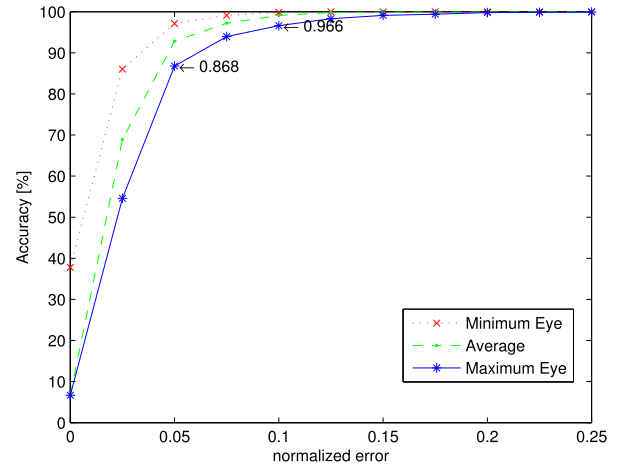


Fig. 5. Accuracy of the proposed method on the BioID database. The three lines correspond to the minimum, average, and maximum normalized errors from the top to the bottom, respectively.

to acquire different sizes of $\Delta\theta$, i.e., $\Delta\theta = 2\pi/n$, over the accuracy on the original method is also validated. Fig. 6 shows the qualitative maximum normalized error of the different n 's of the original method, together with the accuracy of the proposed method. It is clear that the accuracy declines drastically if n is too small; on the other hand, the increase of n will lead to less influence on the accuracy if n is already sufficiently big. The performance of the original method is kept stable when n is equal to 60, where n is the parameter adopted in Fig. 4. It should be noted that the proposed method outperforms the original one even when n is very large.

In addition, the proposed method has a low computational cost in comparison with the original method. In order to compare with the computational load, both methods have been implemented by C++ in Ubuntu (2.4-GHZ Intel Core TM i5 CPU). Both of the methods deal with the same rectangle area of one eye region of the BioID database, and the average processing time is shown in Table I. The proposed method is almost

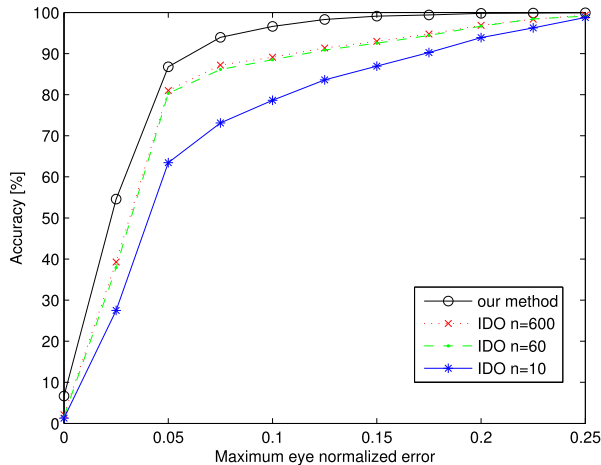


Fig. 6. Increase in accuracy when using different numbers of n and the accuracy of the proposed method of the maximum normalized error.

TABLE I
AVERAGE PROCESSING TIME OF ONE EYE REGION

Method	Time(s)
integro-differential operator(IDO) $n = 10$	0.04595
integro-differential operator(IDO) $n = 60$	0.100073
integro-differential operator(IDO) $n = 600$	0.615077
The proposed method	0.00115615

TABLE II
COMPARISON OF MAXIMUM NORMALIZED ERROR IN THE BioID DATABASE. * MEANS THE VALUE ASSUMED BY [11]

Method	$e \leq 0.05$	$e \leq 0.10$	$e \leq 0.25$
Bai et al.[20]	37.0%*	64.0%	96.0%
Timm et al.[22]	82.5%	93.4%	98.0%
Kim et al.[18]	n/a	96.4%	98.8%
Turkan et al.[29]	19.0%*	73.7%	99.5%
Valenti et al.[23]	84.1%	90.9%	98.5%
Valenti et al.[11]	86.1%	91.7%	97.9%
Markus et al.[17], $p = 7$	85.7%	95.3%	99.7%
Markus et al.[17], $p = 31$	89.9%	97.1%	99.7%
Hamouz et al.[14]	59.0%	77.0%	93.0%
Jesorsky et al.[12]	40.0%	79.0%	91.8%
Cristinacce et al.[13]	56.0%*	96.0%	98.0%
Niu et al.[15]	75.0%*	93.0%	98.0%
Kroon et al.[16]	65.0%*	87.0%	98.8%
Asteriadis et al.[19]	74.0%*	81.7%	97.4%
Asadifard et al.[24]	47.0%	86.0%	96.0%
Campadelli et al.[30]	62.0%	85.2%	96.1%
integro-differential operator(IDO)	80.3%	88.5%	99.1%
The proposed method	86.8%	96.6%	99.9%

40 times faster than the original integrodifferential method when $n = 10$. Thus, the proposed method greatly cuts down the computational cost in low-resolution images and has potential of being integrated into a real-time systems.

E. Comparison With the State-of-the-Art Methods

It is reported that many eye center detection methods have employed the BioID database to evaluate their performances. Table II presents the comparison of the proposed method with the other state-of-the-art methods using the same database and same accuracy measure. Although the highest accuracy

$e \leq 0.05$ is obtained by Markus *et al.* [17] when the number of sampled rectangles p is 31, their method needs to train plenty of images, and the images need to be annotated, which require extra work. On the other hand, the proposed method does not need to be trained and can be easily combined into several existing applications without significant changes. It is evident that the proposed method ranks the first position in comparison with the selected methods requiring no training process.

V. CONCLUSION AND FUTURE DIRECTIONS

This paper has made a significant contribution to the estimation of people's VFOA via localizing human eye center. It is an essential subsystem that should be integrated to achieve human-level intelligence, especially in the area of cognitive systems, human-level AI, and human-robot interaction.

We have proposed a novel integrodifferential method; not only does it offer better performance in comparison with the original integrodifferential method in terms of accuracy and computational cost for low-resolution images, but it also outperforms the state-of-the-art methods. The experimental results also confirm that the proposed method can be easily integrated into real-time systems with acceptable recognition accuracy. This paves the way to resolving bottleneck problems in VFOA required applications such as human-machine interaction.

The proposed method has a significant benefit in that it does not require high-quality images and the hardware can be built using only a Web camera without requiring an IR light or specialized environment. The robustness of eye tracking, in practical, can be further improved by combining head pose direction to address the influence of pose changes. By integrating a 3-D eye model to our model, the gaze direction can be further determined. Future work is targeted to interpret information behind eye movement with cognitive, psychological, and human-like knowledge in the context of human-robot interaction and human-robot skill transfer [34], [35].

REFERENCES

- [1] L. Itti and C. Koch, "Computational modelling of visual attention," *Nat. Rev. Neurosci.*, vol. 2, no. 3, pp. 194–203, Mar. 2001.
- [2] N. L. Cassimatis, "A cognitive substrate for achieving human-level intelligence," *Artif. Intell. Mag.*, vol. 27, no. 2, pp. 45–56, Jul. 2006.
- [3] N. L. Cassimatis, J. G. Trafton, M. D. Bugajska, and A. C. Schultz, "Integrating cognition, perception and action through mental simulation in robots," *Robot. Auton. Syst.*, vol. 49, no. 1, pp. 13–23, 2004.
- [4] R. A. Brooks, "Prospects for human level intelligence for humanoid robots," in *Proc. 1st Int. Symp. Humanoid Robot.*, 1996, pp. 17–24.
- [5] R. Stiefelhagen, M. Finke, J. Yang, and A. Waibel, "From gaze to focus of attention," *Vis. Inf. Inf. Syst.*, pp. 765–772, 1999.
- [6] M. Danninger, R. Vertegaal, D. P. Siewiorek, and A. Mamuji, "Using social geometry to manage interruptions and coworker attention in office environments," in *Proc. Conf. Graph. Interface*, 2005, pp. 211–218.
- [7] M. Katzenmaier, R. Stiefelhagen, and T. Schultz, "Identifying the addressee in human-human-robot interactions based on head pose and speech," in *Proc. 6th Int. Conf. Multimodal Interfaces*, 2004, pp. 144–151.
- [8] K. Smith, S. O. Ba, J.-M. Odobez, and D. Gatica-Perez, "Tracking the visual focus of attention for a varying number of wandering people," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 7, pp. 1212–1229, Jul. 2008.
- [9] A. Metallinou *et al.*, "Context-sensitive learning for enhanced audiovisual emotion classification," *IEEE Trans. Affective Comput.*, vol. 3, no. 2, pp. 184–198, Apr./Jun. 2012.

- [10] D. W. Hansen and Q. Ji, "In the eye of the beholder: A survey of models for eyes and gaze," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 478–500, Mar. 2010.
- [11] R. Valenti and T. Gevers, "Accurate eye center location through invariant isocentric patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1785–1798, Sep. 2012.
- [12] O. Jesorsky, K. J. Kirchberg, and R. W. Frischholz, "Robust face detection using the Hausdorff distance," in *Proc. 3rd Int. Conf. Audio-Video-based Biometric Person Authent.*, 2001, pp. 90–95.
- [13] D. Cristinacce, T. F. Cootes, and I. M. Scott, "A multi-stage approach to facial feature detection," in *Proc. British Mach. Vis. Conf.*, 2004, pp. 1–10.
- [14] M. Hamouz *et al.*, "Feature-based affine-invariant localization of faces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 9, pp. 1490–1495, Sep. 2005.
- [15] Z. Niu, S. Shan, S. Yan, X. Chen, and W. Gao, "2D cascaded adaboost for eye localization," in *Proc. 18th Int. Conf. Pattern Recognit.*, 2006, vol. 2, pp. 1216–1219.
- [16] B. Kroon, A. Hanjalic, and S. M. Maas, "Eye localization for face matching: Is it always useful and under what conditions?" in *Proc. Int. Conf. Content-based Image Video Retrieval*, 2008, pp. 379–388.
- [17] N. Markuš, M. Frljak, I. S. Pandžić, J. Ahlberg, and R. Forchheimer, "Eye pupil localization with an ensemble of randomized trees," *Pattern Recognit.*, vol. 47, no. 2, pp. 578–587, Feb. 2014.
- [18] S. Kim *et al.*, "Multi-scale Gabor feature based eye localization," in *Proc. World Acad. Sci., Eng. Technol.*, 2007, vol. 21, pp. 483–487.
- [19] S. Asteriadis, N. Nikolaidis, A. Hajdu, and I. Pitas, "An eye detection algorithm using pixel to edge information," in *Proc. Int. Symp. Control, Commun. Signal Process.*, 2006, pp. 1–4.
- [20] L. Bai, L. Shen, and Y. Wang, "A novel eye location algorithm based on radial symmetry transform," in *Proc. 18th Int. Conf. Pattern Recognit.*, 2006, vol. 3, pp. 511–514.
- [21] D. Reisfeld, H. Wolfson, and Y. Yeshurun, "Context-free attentional operators: The generalized symmetry transform," *Int. J. Comput. Vis.*, vol. 14, no. 2, pp. 119–130, Mar. 1995.
- [22] F. Timm and E. Barth, "Accurate eye centre localisation by means of gradients," in *Proc. 6th Int. Conf. Comput. Vis. Theory Appl.*, 2011, pp. 125–130.
- [23] R. Valenti and T. Gevers, "Accurate eye center location and tracking using isophote curvature," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8.
- [24] M. Asadifard and J. Shanbezaheh, "Automatic adaptive center of pupil detection using face detection and CDF analysis," in *Proc. Int. MultiConf. Eng. Comput. Sci.*, 2010, vol. 1, pp. 130–133.
- [25] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 11, pp. 1148–1161, Nov. 1993.
- [26] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, "Image understanding for iris biometrics: A survey," *Comput. Vis. Image Understand.*, vol. 110, no. 2, pp. 281–307, May 2008.
- [27] F. Jan, I. Usman, and S. Agha, "A non-circular iris localization algorithm using image projection function and gray level statistics," *Optik-Int. J. Light Electron Opt.*, vol. 124, no. 18, pp. 3187–3193, Sep. 2013.
- [28] J. Daugman, "The importance of being random: Statistical principles of iris recognition," *Pattern Recognit.*, vol. 36, no. 2, pp. 279–291, 2003.
- [29] M. Türkan, M. Pardas, and A. E. Cetin, "Human eye localization using edge projections," in *Proc. Conf. Comput. Vis. Theory Appl.*, 2007, pp. 410–415.
- [30] P. Campadelli, R. Lanzarotti, and G. Lipori, "Precise eye localization through a general-to-specific model definition," in *Proc. British Mach. Vis. Conf.*, 2006, pp. 187–196.
- [31] N. B. Puhana, N. Sudha, and A. S. Kaushalram, "Efficient segmentation technique for noisy frontal view iris images using Fourier spectral density," *Signal Image Video Process.*, vol. 5, no. 1, pp. 105–119, Mar. 2011.
- [32] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, May 2004.
- [33] The BioID Face Database. [Online]. Available: <http://www.bioid.com/downloads/facedb/index.php>
- [34] H. Liu, "Exploring human hand capabilities into embedded multifingered object manipulation," *IEEE Trans. Ind. Informat.*, vol. 7, no. 3, pp. 389–398, Aug. 2011.
- [35] Z. Ju and H. Liu, "Human hand motion analysis with multisensory information," *IEEE/ASME Trans. Mechatron.*, vol. 19, no. 2, pp. 456–466, Apr. 2014.



Haibin Cai received the B.Eng. degree from China Jiliang University, Hangzhou, China, in 2012 and the M.Sc. degree from Zhejiang University of Technology, Hangzhou, in 2015. He is currently working toward the Ph.D. degree at the Intelligent Systems and Biomedical Robotics Group, University of Portsmouth, Portsmouth, U.K.

His research interests include computer vision, gaze estimation, image segmentation, and object tracking.



Bangli Liu received the B.Eng. degree from China Jiliang University, Hangzhou, China, in 2012 and the M.Sc. degree from East China University of Science and Technology, Shanghai, China, in 2015. She is currently working toward the Ph.D. degree at the Intelligent Systems and Biomedical Robotics Group, University of Portsmouth, Portsmouth, U.K.

Her research interests include computer vision, pattern recognition, image segmentation, and human motion analysis.



Jianhua Zhang (M'11) received the M.Sc. degree from Zhejiang University of Technology, Hangzhou, China, in 2009 and the Ph.D. degree from the University of Hamburg, Hamburg, Germany, in 2012.

He was a Research Assistant with the City University of Hong Kong, Hong Kong, in 2008. He is currently working with the College of Computer Science, Zhejiang University of Technology. His research interests include category discovery, object detection, image segmentation, and medical image analysis.



Shengyong Chen (M'01–SM'10) received the Ph.D. degree in computer vision from the City University of Hong Kong, Hong Kong, in 2003.

He joined Zhejiang University of Technology, Hangzhou, China, in February 2004, where he is currently a Professor with the Department of Computer Science. He received a fellowship from the Alexander von Humboldt Foundation of Germany and worked at the University of Hamburg, Hamburg, Germany, in 2006–2007. He has published over 100 scientific papers in international journals and

conferences. His research interests include computer vision, robotics, 3-D object modeling, and image analysis.

Dr. Chen is a Fellow of IET and a committee member of the IET Shanghai Branch.



Honghai Liu (M'02–SM'06) received the Ph.D. degree in intelligent robotics from King's College London, London, U.K., in 2003.

He is currently a Chair Professor of intelligent systems and robotics. His research interests include wearable sensing, biomechanics, intelligent video analytics, intelligent robotics, and their practical applications, with an emphasis on approaches that could make contribution to the intelligent connection of perception to action using contextual information.

Dr. Liu is an Associate Editor of the *IEEE TRANSACTIONS ON HUMAN MACHINE SYSTEMS*, *IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS*, and *IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS*.