Iris Detection and Tracking for Micro-Movement Analysis in Video Sequences

Alister Giles
Dept. of Mechatronics Engineering
University of Canterbury
Christchurch, New Zealand
agi73@uclive.ac.nz

Richard Green
Dept. of Computer Science & Software Engineering
University of Canterbury
Christchurch, New Zealand
richard.green@canterbury.ac.nz

***Abstract*—This paper presents a robust method for detecting and tracking micro-movements of the iris in video sequences. The proposed method leverages the Circular Hough Transform (CHT) to localise the iris region, followed by circularity filtering to mitigate false detections caused by specular reflections. A Kalman filter is employed to predict and correct eye positions over time, facilitating smoother tracking and reducing noise-induced jitter. The method was evaluated on a dataset comprising videos from four subjects, with a total duration of 10:03 minutes and 7:39 minutes of eyes-in-frame. The results demonstrate an F1 score of 0.97, outperforming similar studies in the literature, with a precision of 0.98 and a recall of 0.95. The proposed technique exhibits high accuracy and robustness in detecting and tracking iris movements, making it suitable for applications in biometrics, human-computer interaction, and medical diagnostics.**

I. Introduction

Eye detection is a widespread technique commonly utilised in computer-vision applications **??**. Typically, algorithms initially detect the face, marking landmarks throughout and use these landmarks to locate the eye region. While this is a logical process to undertake, various use cases require the entire face to be in frame. This caveat provides reason to invest further in the area.

II. Background and Literature Review

A. *Iris Detection and Tracking Techniques*

Iris detection and tracking methodologies represent a cornerstone for advancements across biometrics, human-computer interaction, and medical diagnostics. Initial approaches typically detected the entire face, subsequently locating eyes within the defined facial structure [1]. This technique leveraged relative positioning of features, enhancing accuracy and efficiency of eye detection algorithms. Appearance-based methods extract visual features from the eye region to detect and localise the iris. Haar cascade algorithms and Adaboost classifiers [2]–[4] identify the iris based on distinctive appearance patterns. While computationally efficient and robust to illumination changes, these methods exhibit sensitivity to occlusions and limited generalisation capabilities. Shape-based techniques [5], [6] exploit geometric iris properties, assuming circular or elliptical shapes. The Circular Hough Transform (CHT) [7]–[11] accumulates votes in a parameter space, identifying peaks corresponding to potential circle centres and radii. Robust to noise and effective for circular iris boundaries, the CHT struggles with non-circular deformations and varying iris sizes. Hybrid methods combine appearance and shape cues. Gwon et al. [12] performed eye detection using appearance-based techniques, subsequently applying circular Hough transforms for accurate pupil localisation within detected eye regions. Deep learning architectures, such as Convolutional Neural Networks (CNNs), learn hierarchical feature representations directly from data, potentially capturing complex patterns and generalizing better to diverse scenarios. However, these methods require large labeled training datasets and are computationally intensive [13].

B. *Challenges and Limitations*

Despite progress, existing techniques face several challenges:
Occlusions: Eyelids, eyelashes, glasses, and other obstructions significantly impact accuracy [8].
Deformations: Many methods assume circular or elliptical iris shapes, which may not hold true for non-circular deformations due to perspective distortions or physiological factors [10].
Illumination: Shadows, specular reflections, or low-contrast environments affect performance [9].
Real-time Performance: Complex computations or deep learning models may struggle to achieve real-time performance, especially on resource-constrained devices or high-resolution videos [12].
Generalisation: Methods developed and evaluated on specific datasets may degrade when applied to diverse populations or environments [11].

C. *Motivation*

Given existing challenges and limitations, robust and efficient iris detection and tracking methods are needed.

The proposed approach aims to: Combine shape-based techniques, such as the Circular Hough Transform, with additional filtering and tracking mechanisms to mitigate false detections and improve robustness. Incorporate circularity filtering to handle non-circular iris deformations and specular reflections within the iris region [Section III.D]. Employ Kalman filtering techniques to predict and correct eye positions over time, facilitating smoother tracking and reducing noise-induced jitter [Section III.E]. Explore computational optimisations and potential improvements, such as multi-scale or hierarchical approaches, to enhance real-time performance and handle a larger range of iris sizes [Section VI.B]. By addressing these challenges, the proposed method aims to achieve improved accuracy, robustness, and computational efficiency in detecting and tracking micro-movements of the iris, enabling applications in biometrics, human-computer interaction, and medical diagnostics.

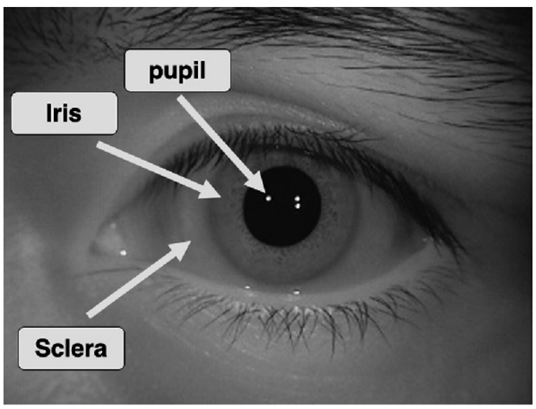

Fig. 1. A Labelled Periocular Image [14]

D. Circular Hough Transform

One method used commonly in the field is the Circle-Hough-Transform (CHT) [8] [9] [10] [11] [15]. The Circle-Hough-Transform (CHT) is a method used to locate circles within a still frame. Equation 1 shows the representation of a circle in 2-D space.

$$(x - x_{centre})^2 + (y - y_{centre})^2 = r^2 \tag{1}$$

In Equation 1, x_{centre} and y_{centre} represent the coordinates of the centre of a potential circle, x and y denote the location of all given circumference points, and r denotes its radius.

For each edge within the frame, a predefined number of circle centres and radii are located [16]. During the CHT process, each edge location in the frame, casts votes into a 3-D array representing potential circle centres and radii. These votes accumulate to form peaks in the array, indicating the most probable circle parameters. A threshold is set to define the minimum number of votes required to define the presence of a circle.

An illustration can be seen in Figure 2. The red circle represents detected edges of a potential circle in an image, the blue dots indicate specific edge points selected for voting, and the dotted circles illustrate the voting process. Each blue dot votes for all possible circle centres that could pass through it, based on different radii.

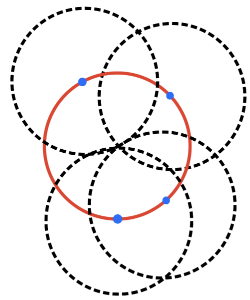

Fig. 2. CHT Diagram

III. Proposed Method

For the sake of consistency, specific computer details can be seen outlined in Table I, and software components in Table II.

TABLE I
HARDWARE SPECIFICATIONS

Device	Specification
Computer	M2 MacBook Air
OS	macOS
Processor	4 cores @ 3.49 GHz - 4 cores @ 2.42 GHz
RAM	8 GB
Camera	30fps

TABLE II
SOFTWARE SPECIFICATIONS

Component/Package	Version/Specification
IDE	Visual Studio Code: 1.89.0
Language	Python: 3.10.8
Opencv-python	4.9.0.80
numpy	1.26.4

The proposed method involves separating the input video into individual frames. Each frame is then subjected to the process outlined in Figure 3.

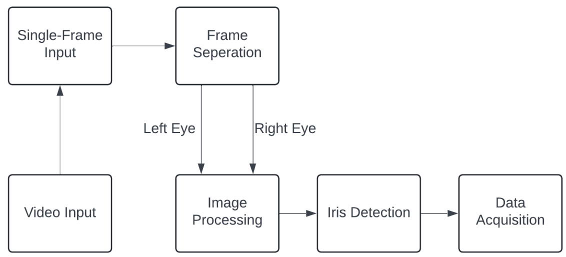

Fig. 3. System Flow Diagram

A. Video Separation

Video separation serves as the initial step in the proposed methodology, aimed at breaking down the continuous stream of video data into individual frames. The input video, captured at a frame rate of 30 frames per second (fps), undergoes this process to facilitate frame-level analysis. By isolating each frame, the subsequent algorithms can operate independently on distinct instances of visual data, enabling precise detection and tracking of eye movements. While majority of processing was completed on a frame-by-frame basis, various data are saved throughout, including an implementation of a Kalman Filter.

B. Frame Separation

Following video separation, each frame undergoes frame-specific analysis to identify regions of interest (ROIs). The frame is then divided into left and right hemispheres. This process assumes all inputs are centered, requiring at maximum one eye in each of the hemispheres. This frame-by-frame approach allows for dynamic adaptation to changes in lighting conditions, occlusions, and other environmental factors. Frame separation lays the groundwork for subsequent pre-processing and analysis steps, ensuring that eye detection algorithms operate effectively on singular image data.

C. Pre-processing Methods

Pre-processing techniques are applied to individual frames to enhance the visibility of eye features and mitigate noise artifacts. Gaussian blurring is employed to reduce high-frequency noise, while adaptive thresholding enhances local contrast to delineate eye boundaries. Additionally, morphological operations such as erosion and dilation help refine the detected regions, ensuring accurate localisation of eye structures. These pre-processing steps prepare the input frames for subsequent analysis by improving feature contrast and reducing noisy data.

D. CHT Implementation

The Circular Hough Transform (CHT) is a key component of the proposed method, utilized for the detection of circular objects within the identified hemispheres. By employing the CHT, the algorithm searches for circular patterns indicative of pupils, leveraging parameters such as minimum and maximum radii to constrain the search space. This detection technique enables the accurate localisation of pupils within varying contexts, including changes in scale, orientation, and illumination. The implementation of the CHT enhances the algorithm's capability to identify and track eye movements with high reliability.

One common issue encountered when applying the CHT for iris detection is the localisation of circles within the iris itself due to light reflections. These reflections can create circular shapes within the iris region, leading to false positives in the detection process. To address this issue, a technique known as circularity filtering was employed as a post-processing step.

To address the issue of false detections caused by light reflections within the iris, a circularity index was computed for each detected circle using the following algorithm. Equation 3 presents the variables used in calculations.

$$A = Contour\ Area, \tag{2}$$
$$P = Arc\ Length \tag{3}$$

If the perimeter P of the contour was zero, indicating a degenerate case, the circularity index was set to zero. Otherwise, the circularity index was found using equation 4

$$circularity = \frac{4\pi A}{P^2} \tag{4}$$

This formula quantifies the roundness of the detected contour, where A represents the area enclosed by the contour, and P denotes the perimeter of the contour. By comparing the circularity index of each detected circle to a predefined threshold, false detections caused by specular reflections were effectively mitigated, enhancing the robustness of the iris localisation algorithm.

E. Data Acquisition

Detected circles representing pupils are subjected to further analysis to extract relevant information, including current and mean position. Kalman filtering techniques are utilized to predict and correct eye positions over time, facilitating smoother tracking and reducing noise-generated jitter in eye movement estimation. By continuously updating and refining the tracked positions, the data acquisition process ensures accurate representation of eye movements within the video sequence. This approach enables the extraction of valuable insights into dynamic eye behaviors.

IV. Results

The proposed method was evaluated with a dataset comprising of videos from four subjects. These videos have a total length of 10:03 minutes with a total time of eyes-in-frame of 7:39 minutes. The testing was completed within the following framework.

A. Determining Threshold Values

Threshold values for circularity and spectrum intensity tests were determined empirically in order to optimise the performance of the method. Various thresholds were tested to find the values that maximize the F1 Score. Balancing precision and recall effectively resulted in an adequate circularity index of 0.97 and spectrum intensity of 190. These values were used in the methods previously described.

B. F1 Score

The F1 Score is a commonly used metric in classification tasks, and provides a comprehensive evaluation of a method's performance by considering both precision and recall. Precision can be seen in equation 5, and measures the accuracy of positive predictions, indicating the proportion of correctly predicted positive cases out of all cases predicted as positive.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5}$$

Recall quantifies the model's ability to capture all positive cases, revealing the proportion of correctly predicted positive cases out of all true positive cases. This can be seen in equation 6

$$\text{Recall} = \frac{TP}{TP + FN} \tag{6}$$

The F1 Score, shown in equation 7, can be defined as the harmonic mean of precision and recall. Combining these two metrics into a single value provides an assessment of a model's effectiveness in handling both false positives and false negatives. By combining precision and recall, the F1 Score obtains an overarching view of the method's performance.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

Figures 4 and 5 present examples of True Positives, with a small margin of error with difference between the true centre and this method's centre being 1.6% and 2.9% for Figures 4 and 5, respectively. In these figures, the green circle is generated by the proposed method, and the red-dotted circle is a mark of ground-truth.

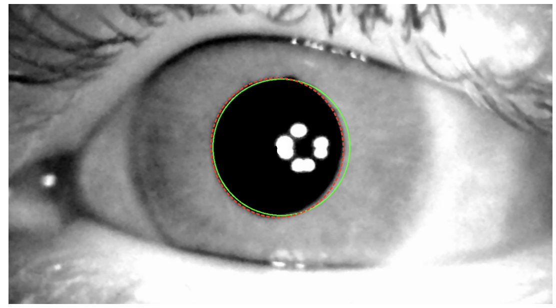

Fig. 4. True Positive: Example 1

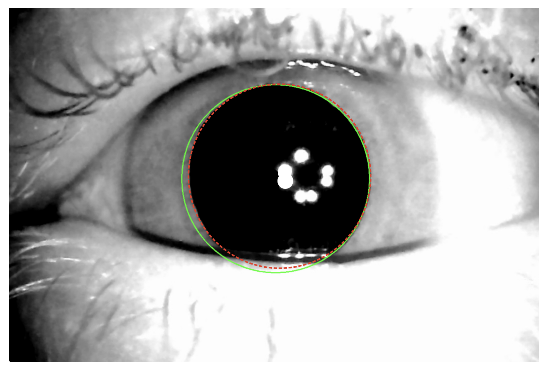

Fig. 5. True Positive: Example 2

Figures 6 and 7 present examples of a False Positive and a False Negative, respectively.

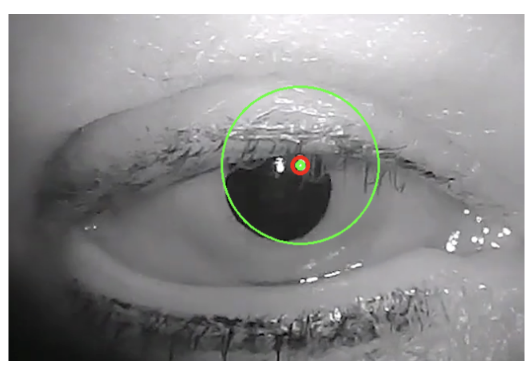

Fig. 6. False Positive

TABLE IV
ACCURACY COMPARISON WITH SIMILAR STUDIES

Source	Accuracy Score
Biswas et al.	92%
Kaudki and Bhurchandi	97%
Athinarayanan	95%
Proposed Method	**97%**

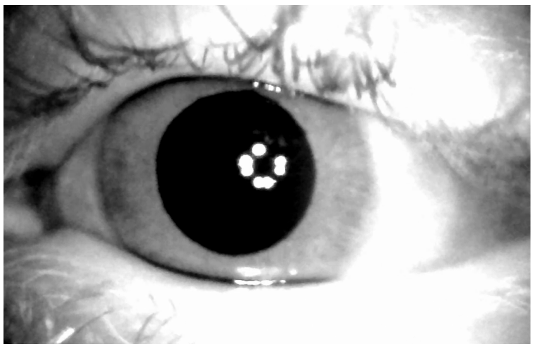

Fig. 7. False Negative

Table III presents the results given utilisation of the aforementioned methods.

TABLE III
RAW RESULTS

Input Video	#1	#2	#3	#4	Avg (%)
Frames Analysed	990	2640	5190	4950	-
True Positives (TP)	962	2390	4645	4828	93.7
False Positives (FP)	3	57	233	13	1.8
False Negatives (FN)	25	193	312	109	4.5

While F1 scores are typically reserved for testing machine learning methods, the score presents a valuable comparison between different methods of achieving the same goal.

The proposed method achieved a precision score of 0.98, a recall of 0.95, and an F1 Score of 0.97.

In comparison to similar studies, Biswas et al. [17], Kaudki and Bhurchandi [18], and Athinarayanan [19] this method provides a high level of accuracy represented by F1 score.

Table IV presents the comparisons between similar studies.

V. CONCLUSIONS

The proposed method for detecting and tracking micro-movements of the iris in video sequences demonstrates promising results. By combining the Circular Hough Transform, circularity filtering, and Kalman filtering techniques, the method achieves high accuracy and robustness in localising and tracking iris movements. The evaluation on a diverse dataset showcases the method's effectiveness, with an F1 score of 0.97, outperforming similar studies in the literature. The low false positive and false negative rates further highlight the method's reliability in handling challenging scenarios. While the current approach assumes perfectly circular iris boundaries and relies on grayscale images, potential improvements and future directions include exploring elliptical iris modeling, incorporating colour information, and evaluating the method on larger and more diverse datasets to enhance its robustness further. Overall, the proposed method demonstrates promising capabilities for applications in biometrics, human-computer interaction, and medical diagnostics, where accurate detection and tracking of iris movements are crucial.

VI. METHOD LIMITATIONS & FUTURE WORK

A. Limitations

- The current method assumes perfectly circular iris boundaries, which may lead to false positives when the iris appears elliptical from the camera's perspective.
- The Circle Hough Transform used in this work has difficulty handling a large range of potential iris sizes efficiently.
- The method relies on grayscale images, which could limit its performance in low-contrast or challenging lighting conditions where colour information may be beneficial.
- The computational complexity of the Circle Hough Transform may hinder real-time performance, especially on resource-constrained devices or for high-resolution videos.

B. Potential Improvements and Future Directions

- The issue of a large potential iris size range could be mitigated by employing multi-scale or hierarchical approaches within the Circle Hough Transform algorithm.
- To handle non-circular iris boundaries more effectively, future work could explore elliptical iris modeling techniques or other geometric transformations.
- Incorporate colour information or explore deep learning-based methods that can make use of the colour-spectrum for improved performance in challenging lighting conditions.
- Evaluate the method on larger and more diverse datasets, including variations in ethnicity, age, occlusions, and light conditions, to improve robustness.

REFERENCES

[1] C. Morimoto, D. Koons, A. Amir, and M. Flickner, "Real-time detection of eyes and faces," pp. 117–120, 1998.

[2] R. Lienhart and J. Maydt, "An extended set of haar-like features for rapid object detection. in proceedings of international conference on image processing," *IEEE. doi*, vol. 10, 2002.
[3] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, pp. 137–154, 2004.
[4] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," in *Computational Learning Theory: Eurocolt 95*, pp. 23–37, Springer-Verlag, 1995.
[5] Yap and R. Green, "Real time face detection and recognition," tech. rep., Computer Vision Lab, Department of Computer Science and Software Engineering, University of Canterbury, Christchurch, 2017.
[6] B. Goulter and R. Green, "Eye gaze tracking," tech. rep., Computer Vision Lab, Department of Computer Science and Software Engineering, University of Canterbury, Christchurch, 2013.
[7] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.
[8] N. Cherabit, F. Z. Chelali, and A. Djeradi, "Circular hough transform for iris localization," *Science and Technology*, vol. 2, no. 5, pp. 114–121, 2012.
[9] W. Kong and D. Zhang, "Accurate iris segmentation based on novel reflection and eyelash detection model," in *Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing. ISIMP 2001 (IEEE Cat. No.01EX489)*, pp. 263–266, 2001.
[10] L. Ma, Y. Wang, and T. Tan, "Iris recognition using circular symmetric filters," in *2002 International Conference on Pattern Recognition*, vol. 2, pp. 414–417 vol.2, 2002.
[11] C.-l. Tisse, L. Martin, L. Torres, M. Robert, *et al.*, "Person identification technique using human iris recognition," in *Proc. Vision Interface*, vol. 294, pp. 294–299, 2002.
[12] S. Y. Gwon, C. W. Cho, H. C. Lee, W. O. Lee, and K. R. Park, "Robust eye and pupil detection method for gaze tracking," *International Journal of Advanced Robotic Systems*, vol. 10, no. 2, p. 98, 2013.
[13] P. Jain and R. Green, "Automatic face recognition and tracking," tech. rep., Computer Vision Lab, Department of Computer Science and Software Engineering, University of Canterbury, Christchurch, 2016.
[14] K. Ogawa and K. Kameyama, "Adaptive selection of classifiers for person recognition by iris pattern and periocular image," in *Neural Information Processing. ICONIP 2021* (Taufik, M. Lee, M. Ayu, K. W. Wong, and A. N. Hidayanto, eds.), vol. 13111 of *Lecture Notes in Computer Science*, Springer, 2021.
[15] R. Wildes, "Iris recognition: an emerging biometric technology," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1348–1363, 1997.
[16] OpenCV, "Image processing in opencv: Hough circle transform." https://docs.opencv.org/4.x/da/d53/tutorial_py_houghcircles.html.
[17] R. Biswas, J. Uddin, and M. J. Hasan, "A new approach of iris detection and recognition," *International Journal of Electrical and Computer Engineering*, vol. 7, no. 5, p. 2530, 2017.
[18] O. Kaudki and K. Bhurchandi, "A robust iris recognition approach using fuzzy edge processing technique," in *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, (Bengaluru, India), pp. 1–6, 2018.
[19] S. Athinarayanan, K. Navaz, R. L. Priya, and V. Ebenezer, "Effective image processing techniques based iris attendance system," *ICTACT Journal on Image and Video Processing*, pp. 2071–2075, 2019.