Feasibility of Hough-Transform-based Iris Localisation for Real-Time-Application

Klaus D. Toennies, Frank Behrens, Melanie Aurnhammer
Dept. of Computer Science, Otto-von-Guericke Universität Magdeburg, Germany
{klaus,aurnhamm}@isg.cs.uni-magdeburg.de, fbehrens@ivs.cs.uni-magdeburg.de

Abstract

We present a fast method for locating iris features in frontal face images based on the Hough transform. It consists of an initial iris detection step and a tracking step which uses iris features from initialisation for speeding up computation. The purpose of research was to evaluate the feasibility of the method for tracking at 200 frames per second or higher. Processing speed of the prototypical implementation on a 266Mhz Pentium II PC is approximately 6 seconds for initial iris detection and about 0.05 seconds for each tracking step. The algorithm was applied to images of subjects taken under normal room lighting conditions. Tests showed robustness with respect to shadowing and partial occlusion of the iris. The localisation error was below two pixels. Accuracy for tracking was within one pixel. A reduction of the number of pixels which are processed in the tracking step by 90% showed a modest degradation of the results.

1 Motivation

Eye movement analysis has been used to detect phase transitions from wakefulness to sleep in human subjects. Eye movements are controlled by the oculomotor nuclei in the brain stem. Other control centres exist adjacent to these nuclei which control vegetative functions like heart rhythm and respiration. Changes in vigilance cause changes in the function of the brainstem nuclei. Eye movements for evaluating changes in vigilance were measured using the electro-oculographic (DC-EOG) method. An analogue signal is obtained which can be sampled at frequencies of 1 kHz or higher in order to detect even the smallest saccades [1].

Although such a system offers valuable insight into physiological processes, it is inconvenient for the subject. A non-invasive, picture-based method is preferred for analysing vigilance changes in every-day-situations. Tracking the movement of the eyes with high accuracy using an infrared-based tracking system is possible (see, e.g., [2]). However, in order to impose as little constraints as possible on the experimental set-up, we want to measure saccadic motion without interference from using an infrared light beam. Such tracking is expected to be less accurate. However, it is acceptable for measuring saccadic motion as we are only interested in the relative movement of the eye. Repeatability and speed are more important than exact localisation.

Two different problems have to be solved for a tracking system:

1. Detection of iris location and determination of iris features.
2. Eye tracking based on iris features.

The initialisation phase should not be longer than a second. Tracking has to be carried out at speeds of 200-1000 fps (frames per second). Such high speed can only be achieved if the employed methods are simple. Even then it is expected that results need to be generated from updating a prediction with incomplete results.

We decided to explore the capabilities of an approach which is mostly based on the Hough transform although current methods for eye feature extraction use deformable templates (e.g., [3], [4], [5], [6]). The advantage of the Hough transform over using templates is that the final result is computed from a number of independent partial solutions. The Hough transform is therefore well-suited for the fusion of partial results with a predictions from a model.

The algorithm has been applied to test images of faces which were taken under different lighting conditions. The two irises and their features are detected during initialisation. Given the features the localisation information will be fine-tuned by evaluating a small neighbourhood of each iris. Only the last step has to be repeated during tracking.

2 Iris detection

We make the following assumptions for detecting and tracking irises:

- The image is a frontal view of the face, i.e. the two eyes are visible and have approximately the same distance from the camera.
- The iris in each eye is at least partially visible.
- The brightness of the iris is lower than that of the visible part of the sclera.
- An average ratio between iris size and distance between eyes is known.

- The line between the two centres of the irises must not deviate by more than 30° from the x-axis of the image.
- The subject may not be closer than 0.5 m to the camera and not further away than 5 m.

2.1 The generalised Hough transform for circles

We use a variant of the Hough transform for circles that has been proven to be efficient and robust [7]. After estimation of the intensity gradient, presumed edge pixels cast votes for locations of circle centres with radius r. Each edge pixel votes in direction of the gradient for a location which is at a distance of r away from the pixel. Circle centres are found at local maxima of the vote count.

A two-step matching process is used for finding the correct radii [8]. We assume that every pixel is a potential edge pixel. This makes the method insensitive to contrast variations due to shading effects. The number of votes at a location (i,j) is a function of the gradient length $\mathbf{G}(i,j)$:

$$v(i,j) = \left(\frac{|\mathbf{G}(i,j)|}{\sum_{i,j} |\mathbf{G}(i,j)|} \right)^e$$

An exponent $0<e<1$ is applied to the normalised vote in order to favour an even distribution of votes from pixels which vote for a certain circle centre.

The optimal radius r_{iris} is selected from a small number of sample radii $r_{min}<\{r_1,\ldots,r_n\}<r_{max}$. The range $\{r_{min},\ldots,r_{max}\}$ for the radius of the iris is initially computed from the focal length of the camera system, from the expected distance range of 0.5 to 5 m, and from an approximated size of the iris of 12 mm. The range is narrowed at later stages of the algorithm by using information from previous steps.

A measure is needed in order to select an optimal radius among the sample radii. A convex function would be preferred but such function is not expected to exist. There will be always more than one pair of circles in the scene and any circularity-based optimality measure will have a number of local minima.

Thus we define a circularity function which is designed to have a well-pronounced minimum for the correct iris radius and which is convex in the vicinity of this radius. Circularity is evaluated from pixels in regions $ctr(r)$ which may contain circle centres of radius r. The maximum v_{max} of all votes is computed in order to determine all pixels in $ctr(r)$. A pixel p belongs to $ctr(r)$ if $v(p)>0.33\cdot v_{max}$. Circularity $C(r)$ is computed from votes of all pixels in $ctr(r)$:

$$C(r) = \sum_{p \in ctr(r)} v(p)^{-1}$$

If the current radius is slightly higher or lower than the optimal radius it will lead to a broader distribution of votes around the centre location. This increases the circularity. We exclude pixels with $v(p)<0.33\cdot v_{max}$ in order to produce a better pronounced minimum.

2.2 Computing the iris centres

After gradients, gradient directions and votes are computed, radii, centres and average gradient length of two iris centres are determined as follows:

1. Computation of an initial estimate of the locations of the two eyes.
2. Computation of an initial estimate of the iris radii given the initial estimates on the eye positions.
3. Fine-tuning of centre position and radius of each iris.

Each of the steps use the Hough transform for iris detection. The search is carried out in three steps in order to reduce the number of pixels which are allowed to vote and in order to reduce the number of radii for which the Hough transform is applied.

During initial localisation the number of voting pixels is restricted to about 25.000 out of 1.3 million in a 1024×1280 pixel image. Pixel sites could be selected at random but for simplicity we chose pixel sites at the intersection of every 8th row with every 8th column. Radii range from 10 to 100 pixels. The size of the radius is increased by 5 pixels in each step. Two connected regions with $v(p)> 0.33\cdot v_{max}$ with highest vote count are computed for each radius r. The centres of gravity of the votes in these regions are taken as iris centres cx_1,cy_1 and cx_2,cy_2.

An optimality measure $iris(r)$ is computed from the circularity $C(r)$ of the Hough transform, the distance d_{iris} between (cx_1,cy_1) and (cx_2,cy_2) and the contrast $contr_{iris}(r)$ between presumed iris and background (see fig. 1):

$$iris(r) = C(r)\cdot\left[1+\left(10-\frac{d_{iris}}{r}\right)^2\right]\cdot contr_{iris}(r)$$

The distance measure assumes that the ideal ratio between iris radius and distance between the two iris centres is 10. It increases quadratically with deviations

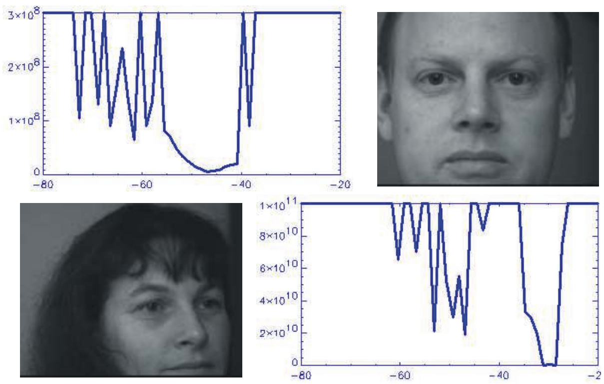


Figure 1: Examples for the optimisation measure for two images. The measure has a less pronounced minimum, if the image has shading and the view is not a frontal view.

from the ideal ratio. The ideal ratio was found experimentally. It was added into the optimality measure to reduce the likelihood of selecting circle centres which are too close or too far from each other.

The contrast measure is given by the grey level distance between the average grey level g_{out} from locations outside the iris with distances d_{out} with $r<d_{out}<1.1r$ and the average grey level g_{in} inside the iris with distances $0.9r<d_{in}<r$:

$$contr_{iris} = {g_{in}}\Big/{g_{out}} .$$

It decreases the darker the iris is with respect to its background.

A common radius r_{iris} and centre locations x_{iris1},y_{iris1} and x_{iris2},y_{iris2} are selected from the sample radii for which $iris(r)$ is minimal and the line from x_{iris1},y_{iris1} to x_{iris2},y_{iris2} deviates by less than 30° from the x-axis.

Given the initial estimate of the iris centres and sizes, a sub-image is computed in order to start a refined search. The sub-image is bounded by

$$\begin{aligned} \min(x_{iris1},x_{iris2}) - 3r_{iris} &< x < \max(x_{iris1},x_{iris2}) + 3r_{iris} \\ \min(y_{iris1},y_{iris2}) - 3r_{iris} &< y < \max(y_{iris1},y_{iris2}) + 3r_{iris} \end{aligned} .$$

The Hough transform is repeated on the sub-image for radii r_i with $r_{iris}-4 < r_i < r_{iris}+4$. The radius is increased by a pixel in each step. Every 4th pixel casts a vote in order to increase the accuracy of the result. The optimality measure $iris(r)$ is computed again for each radius. A new radius r_{iris} and new centre locations x_{iris1},y_{iris1} and x_{iris2},y_{iris2} are selected for which $iris(r)$ is minimum.

Finally, the two centres and radii from the previous steps are used for computing a final result independently for each eye. A region of interest is defined for each eye which is 10% larger than the iris size. The Hough transform is repeated for radii r with $0.9\cdot r_{iris}<r<1.1\cdot r_{iris}$. The increment of radius between two tests is one pixel. For each step, pixels are casting votes which are within a distance range d_{min} to d_{max} from the presumed iris centre. The values for d_{min} and d_{max} are $0.8\cdot r_{iris}$ and $1.2\cdot r_{iris}$.

3 Results

The method was implemented using the IDL (The Interactive Data Language) on a PC with Pentium II, 266 Mhz processor. We applied our method to a number of pictures of faces that were taken indoors under normal illumination conditions (see Fig. 2). The distance between camera and subject was about one meter which resulted in an iris radius of about 50 pixels. The resolution was 100µm per pixel. This is a realistic size for the intended use because current high speed cameras have a resolution of about 100×100 pixels if frame rates of 500 to 1000 fps are desired.

The size of the pictures at the initialisation stage was 1024×1280 pixels. The pictured region ranged from the mouth to the upper forehead.

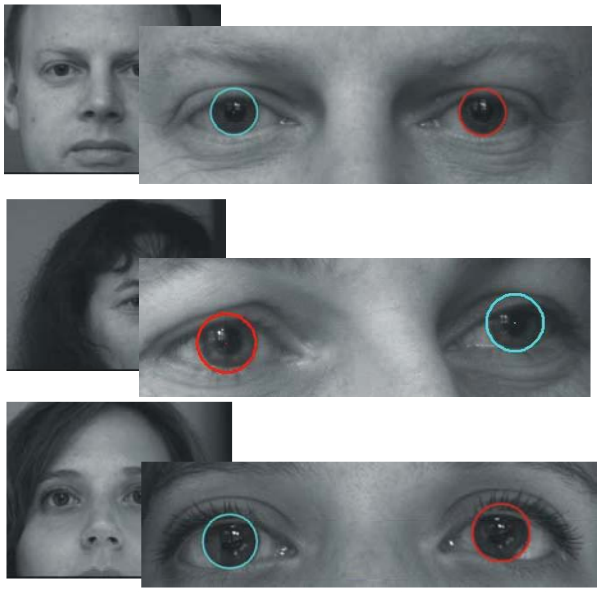


Figure 2: Results from finding the irises in three different subjects. Irises were found in all cases but accuracy depends on adhering to the assumptions made above. The red circle indicates the iris with the better circularity measure.

3.1 Initial iris localisation

An important aspect of the intialisation step – as it provides the system with estimates of iris features for subsequent real-time tracking – is its robustness. In our tests the detection succeeded if more than about 30% of the iris boundary was visible in the two eyes. In these cases the optimality function $iris(r)$ was minimal for the true radius of the iris. It was convex in the vicinity of approximately 10 pixel off this radius. Thus, the initial sampling distance of 5 pixels for two successive tests was sufficient for finding the correct radius range.

Another aspect is the robustness under shadowing. We assume that for initialisation most of the iris is visible. However, illumination may lead to a partial change of signal because some of the eye is occluded. We reduced the signal to 10% of its original brightness in experiments with images of our subjects. It resulted in a displacement error of about one pixel for the iris centre due to a radius error of 1.5 pixels (see Fig. 3). This is probably due to the less pronounced minimum for the optimisation function

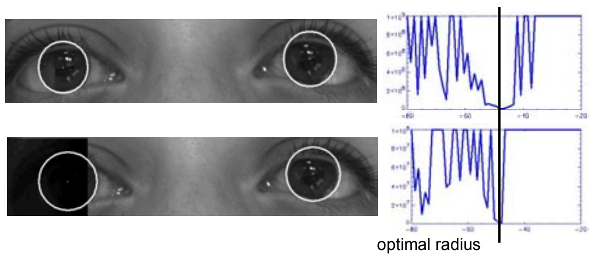
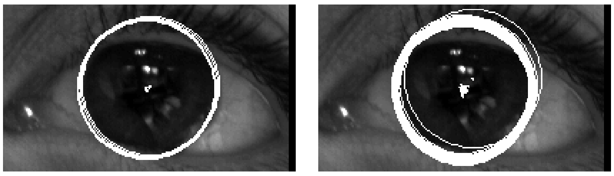


Fig. 3: Iris centres were still found if one eye was partly shadowed. However, the minimum of the optimisation function is less pronounced and the radius deviated by 1.5 pixels.



(a) (b)

Figure 4: (a) Repeatedly selecting a subset of 90% of the pixels lead to an average variation of the estimated centre of 0.87 pixel in x-direction and by 0.32 pixel in y-direction. (b) Selecting just 10% on the vote increased the variation to 1.7 pixel in x-direction and 0.9 pixel in y-direction.

(see Fig. 3). If the radius were specified, the displacement error was below 0.5 pixels. Since the latter is the case for tracking, signal decrease due to shadowing should be avoided during initial localisation or the sampling rate for estimation of the optimisation measure has to be increased.

3.2 Tracking

For tracking, the most important aspect is robustness of finding a unique location of the iris. It must not be the true centre but it needs to be unique under translations. We used the bootstrapping method [9] for a test as images from a highspeed camera are currently unavailable to us. Selecting different random subsets of 90% of the voting locations from the same image led to a standard deviation of less than one pixel for the location of the iris centre. 30% of this change was attributed to displacements along the axis between the two iris centres. This is due to the fact that most of the visible iris boundary is orthogonal to this axis (see Fig. 4a).

It will be necessary for tracking to have an indication on the change of iris size due to a change of distance between subject and camera. If the radius changes and the Hough transform is carried out with the original radius, it will lead to a dispersion of votes around the centre location. This can be measured by the change of standard deviation of votes in the vicinity of the circle centre. In our experiments, we found a non-linear relationship between standard deviation and radius change that was stable for changes of up to 3 pixels. This corresponds to a motion of 6 m/sec at sampling intervals of 5 msec.

The total computation time was about 6 seconds with 0.05 s for the final step. It makes the procedure feasible for real-time-application as only the final step has to be carried out during iris tracking.

The computation time depends largely on the number of votes being cast. They amount to about 6000 pixels per iris in the current implementation. It assumes a maximum speed of motion of the centre location of 10 pixels (= 1 mm) per frame. At an assumed frame rate of 200 fps it translates into a speed of iris motion of about 600° per second. This is in the range of the fastest saccadic motion.

Votes may be reduced by selecting a subset of pixels in the search range. The location of the detected iris centre varied by about 2 pixels on average when we repeatedly selected 10% of the 6000 candidate pixels at random and let them vote for the iris centre (see Fig. 4b). This is too high for tracking purposes but reducing the number of voting pixels for faster computation appears to be feasible.

4 Conclusions

We presented a robust and simple method for detecting and tracking the centres of the irises. The method employs the Hough transform for circles. It produces information on the centre and the radius of the eye as well as additional information on the characteristics of the gradients at the iris boundary. The features may be used for a subsequent real-time tracking of saccadic motion of the eye. The computation times suggest its applicability in a real-time environment.

Future work will concentrate on optimising the number of votes that are needed for reliable computation of the iris centres. Investigations will be undertaken to find out the use of the optimality criterion as a predictor of the result quality.

Additionally, we will work on improving the robustness of centre localisation as an accuracy of one to two pixels is not sufficient. For this purpose, we will use features from iris detection such as expected gradient length and direction for assigning a reliability measure to the votes.

We will also develop and investigate methods for using temporal continuity constraints to predict the visible portion of the iris when closing the eye lid. This will be used to estimate vote reliability at different locations on the circle.

5 References

[1] Behrens F, Schröder-Preikschat W, "An Improved Algorithm to Analyse Eye Movements under Natural Conditions as a Basis for Various Technical Applications", *Proc Intl Symp on Eye Movements and Vision in the Natural World*, Amsterdam/Rotterdam, 2000.

[2] Mulligan JB, Beuttler BR: Eye movement tracking using compressed video images, *Vision Sciences and Its Appliactions; Optical Society Technical Digest Series*, Vol.1, pp.163-166, 1995.

[3] Yuille A., Haallinan P, Cohen CS: Feature extraction from faces using deformable templates. *Intl. J. Computer Vision*, 8(2), pp.99-111, 1992.

[4] Chow G., Li X.: Towards a system for automatic facial feature detection, *Pattern Recognition*, Vol. 26(12), pp. 1739-1755, 1993.

[5] Lam K, Yan H: Locating and extracting the eye in human face images. *Pattern Recognition*, 29(5), pp.771-779, 1996.

[6] Tian Y, Kanade T, Cohn JF: Dual-state parametric eye tracking. *IEEE Intl Conf of Automatic Face and Gesture Recognition*, March 2000, pp.110-115.

[7] Kimme C., Ballard D., Slansky J: Finding circles by an array of accumulators. *Communications of the ACM*, Vol. 18, pp.120-122, 1975.

[8] Davies ER: Tradeoffs between speed and accuraccy in two-stage template matching, *Signal Processsing*, Vol. 15, pp.351-363, 1988.

[9] Pohle R, Toennies KD: A three-level evaluation process for segmentation methods in medical imaging, *Proceeding of the SPIE (Medical Imaging 2002)*, 2002, to appear.