

Iris Localisation for a Head-Mounted Eye Tracker

Hilary Tunley and David Young
School of Cognitive and Computing Sciences
University of Sussex
Brighton, England BN1 9QH.

Email: hilaryt@cogs.susx.ac.uk davidy@cogs.susx.ac.uk

Abstract

This paper discusses the image processing techniques needed to localise the iris outline in images of the eye produced by a novel, portable eye tracker. The tracker allows gaze direction to be measured by supplying simultaneous views of a subject's eye and of the world from a head-mounted camera. Finding gaze direction relative to the head requires accurate and robust measurement of the iris outline under a wide range of lighting conditions, in the presence of highlights, and when the iris is in an extreme position. We describe a reliable method to solve this problem, which achieves a gaze direction accuracy of under 2 degrees.

1 Introduction

There has always been a great deal of interest in the role of eye movements in visual perception. The human visual system, with its non-uniform retinal sampling, uses sophisticated processes for directing gaze and maintaining fixation, which merit study. However, measuring eye movements and gaze direction is difficult, particularly if the head is not fixed in position. Head fixation, which has been used in many experiment on eye movements, cannot, of course, be used for studies of real-world activities such as driving. Various methods of eye movement monitoring using head-mounted apparatus have been devised, and amongst these one of the most straightforward, and potentially most accurate, is an innovative tracker developed by M. F. Land at Sussex, which uses a miniature head-mounted camera to give images both of the scene and of the subject's eye simultaneously. This paper discusses aspects of the automation of the iris position measurement needed to obtain gaze direction using this device.

Land's eye tracker [1] consists of a light head-frame with a small CCD camera mounted above the left eye and pointing downward (Figure 1). A half-silvered mirror below the camera allows it a view forwards (placing a virtual camera, fixed relative to the head, into the orbit) and a concave mirror mounted just below the eye gives the camera an inverted view of the eye in the lower part of the field (Figure 2).



Figure 1: The head-mounted eye tracker

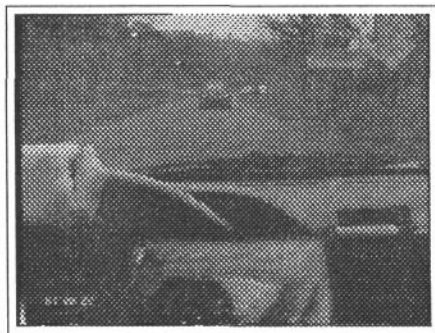


Figure 2: Example frame from tracker

Analysis of each frame requires finding the iris and/or pupil in the image of the eye and, after calibration, mapping this centre to the point in the image of the scene being fixated by the subject. Since the tracker is used in natural conditions, variations in lighting are inevitable, extreme eye positions occur, and bright reflections from the corneal surface are common. These factors make automatic measurement of the eye position much more challenging than for laboratory-based equipment. In our system, gaze direction is determined from the position of the iris outline. This is done by first finding candidate contours for the iris outline using a variety of heuristics, then by fitting ellipses, constrained by the eye model used, to outline contours. It would be possible to incorporate tracking from frame to frame during smooth pursuit eye movements, but the system needs to be able to cope with very fast saccades which may include blinks (and thus redundant video frames), so this has not been implemented for this initial study. Iris detection is highly reliable and the accuracy approaches that achieved by hand fitting the iris outline.

2 The Eye Model

A simple model of the eye (Figure 3), similar to that used by Robertson et al. [2], allows accurate ellipse fitting. The eye is treated as a sphere rotating about its centre, and the iris outline is taken to be a circle on the surface of the sphere. Scaled parallel projection is assumed. The parameters of this model are the scaled eyeball radius R_e , the scaled iris radius R_i , both measured in image pixels, and the coordinates of the projection of the centre of the eye, $(x_c y_c)$. These parameters have to be found for each experimental session, since they vary between subjects and also depend on the exact positioning of the apparatus on the head. Gaze direction relative to the head is specified by two variables, corresponding to the two degrees of freedom of the eye, pan and tilt (cyclotorsional movements are ignored). The variables give the position of the projection of the centre of the iris relative to the eye centre and can be combined with the eye model parameters to calculate the elliptical projection of the iris outline in the image. At present, the eye parameters are found manually, by adjusting them until the projected model iris outline can be made to fit the image by varying the model gaze direction, for a variety of different subject gaze directions. As a starting point, the iris radius

can be easily determined from an image in which the subject is looking straight ahead and the iris projection is circular.

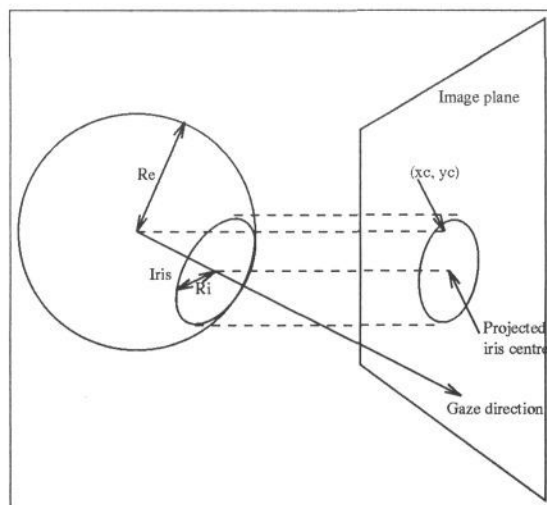


Figure 3: Eye Model

3 Iris Contour Detection

Iris contour detection encounters a number of problems, namely: dominating highlights; spatial changes in contrast at the iris margin due to oblique lighting; strong contours from the eye margins. This section discusses these problems and the contour-based measures used to overcome them.

3.1 Highlights and Changes in Contrast and Lighting

As the surface of the eye is smooth and moist there is a tendency for highlights to be present on the tracker images (Figure 4). The lack of lighting restrictions further increases the possibility of highlights, whilst the eye's convexity also increases the number of high contrast curved highlights which are more likely to be interpreted as potential iris contours. Smoothing at the edge detection stage cannot overcome this as the contours of interest become grossly inaccurate.

Though highlight interference cannot be completely removed it can be significantly reduced using the assumptions that highlights are often spatially small or 'noisy' and/or contained within the area of larger contours (Section 3.2.1). Combining these observations with an individual contour-based analysis (Section 4) ensures that any highlights which do not quite 'qualify' for pruning do not exert an unduly large influence upon the iris detection process.

A further consequence of allowing the wearer free rein in choice of orientation and environment is that of lighting variability, both in space and time, and its

inevitable influence upon the contrast of the iris against the sclera. Extreme sun and shade can reduce this contrast, whilst oblique side lighting may cause variable contrast (Figure 5). Pre-processing therefore takes advantage of hysteresis thresholding in the style of Canny [3]. Any contour point which exceeds a threshold of 30% of the maximum determined contrast is used to “gate” neighbouring pixels for which the gradient is a local maximum along the direction perpendicular to the local equi-gradient contour, but which have a gradient value below the 30% threshold. Figure 6 shows the result of simple thresholding and the resulting reduction in contour information, whilst Figure 7 illustrates the advantage of gate-based thresholding in which even the weak region of the iris is detected. A linking procedure (Section 3.2.2) further enhances coherent contour segments by ‘filling in’ gaps where there is no local gradient evidence.

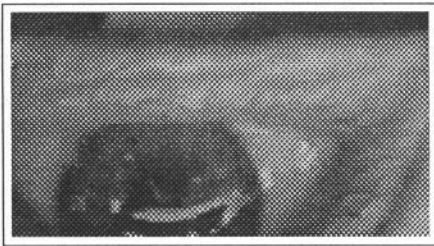


Figure 4: Strongly-highlighted iris

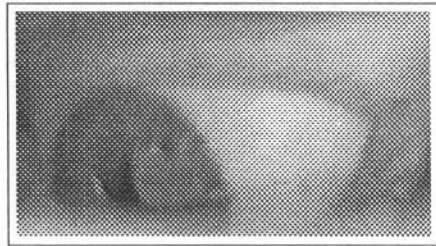


Figure 5: Spatially-varying iris contrast

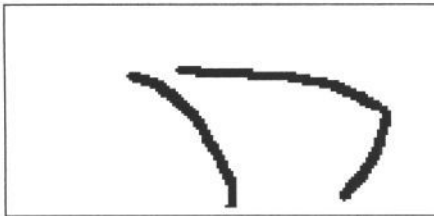


Figure 6: Simple Thresholding

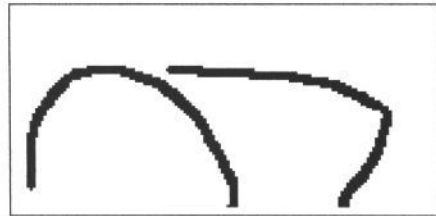


Figure 7: Gate-based Thresholding

3.2 Extracting Contours

The approach used to detect the iris starts with a Canny-based image [3] thresholded in the manner discussed above. This thresholded image includes the iris (or part of the iris) contour in approximately 98% of cases without introducing excessive noise. In the unsuccessful cases the lack of iris is detected by a low Hough Transform score (Section 4) and the process repeated with a reduced threshold. The next stage is to determine the best way to emphasise iris-based edge points and group them into contours whilst restricting spurious edge segments.

3.2.1 Length, Shape and Position-Based Restrictions

A sensible restriction on valid contours is a minimum length in terms of edge pixels. Very short edge segments are more likely to be noise (the low gradient threshold

ensures that iris contours are rarely broken up). An initial minimum length for contour validation is set at $2R_i$. In cases in which this setting does not pick up strong iris contours the value is reduced to R_i pixels.

Once a contour is detected various factors concerning its structure can be calculated, including (all in pixels): length; horizontal extent; vertical extent; area; region centre. This information can be used to determine the validity of each contour. Figure 8 shows an example of a complete iris, in which the contour length approaches 80% of the circumference of the contour's area, whilst Figure 9 illustrates a less ideal case of an iris section, highlight and eye edge.

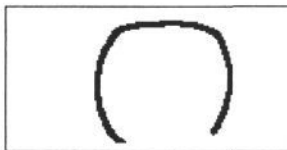


Figure 8: $M_c=0.78$

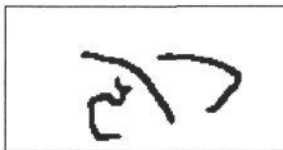


Figure 9: Weak contour

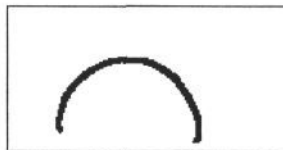


Figure 10: $M_c=0.58$

A simple scale-independent measure of compactness is the ratio of contour length, l , to the perimeter of its bounding box (dimensions l_x and l_y):

$$M_c = \frac{l}{2(l_x + l_y)} \quad (1)$$

Most unambiguous iris contours have values of M_c in the range 0.5 to 0.8, with the lower value occurring when the iris is occluded by the eyelid, eye boundary, or image border (Figure 10). Values lower than this correspond to contours that are too elongated, such as those produced by the lower eye boundary, whilst higher values are mainly from noisy highlights.

Whilst the contour measure, M_c , provides a reliable indication of contour validity, non-valid contours, such as highlights, are often still present. Position-based pruning uses the region centre and area measures to determine whether a contour has a bounding box which is a complete subset of a larger contour's bounding box. Such 'subset' contours can represent highlights and, when combined with M_c measures near the top of the valid range, indicate potential pruning candidates. This method thus favours bigger contours without explicitly defining an area-based threshold which would be too restrictive and tend to remove small irises too (Section 4.2).

3.2.2 Linking Contours

Segmentation based on continuous contours is applied at a later stage, but contours close to the minimum allowable length often represent only a small part of the potentially visible iris outline. Therefore contours which are spatially aligned (i.e. with closely connected end points and compatible edge gradients near those end points) are joined into single longer contours. This is particularly important as, although the pruning process does not actively discriminate against short contours (i.e. those $\approx 2R_i$ – or R_i if threshold lowering has occurred), short contours do

not make good candidates for the ellipse-fitting stage, which works with each individual contour (Section 4), and so relies upon well-connected ellipse data to produce strong unambiguous validity judgements.

3.3 Strong Eye Edges

An approach which uses raw grey-level gradient information as input to the ellipse-fitting stage creates stronger support for examples with eye margins which are more dominant than the iris edge (Figure 11) *even though the iris supplies potentially more non-zero ellipse points overall*. Figure 12 shows the resulting spurious fit from using gradient information. This result, which also occurs for some strong highlights, emphasises the need to consider the *shapes* of the contours as well as their contrast intensity, to avoid spurious matches.

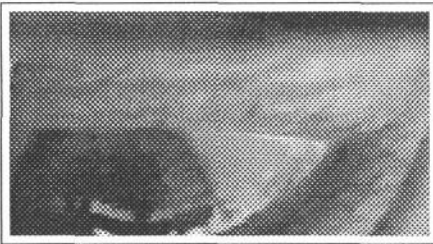


Figure 11: Dominant eye edge

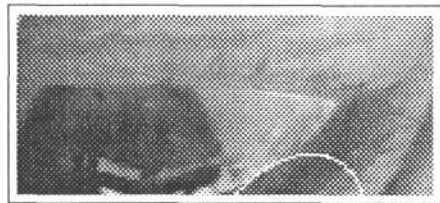


Figure 12: Intensity-based Result

4 Ellipse Fitting

A Hough Transform (HT) was used for the ellipse-fitting stage with a 2D space corresponding to the two degrees of freedom of the eye. Each edge point voted only for those ellipses passing through it which could be generated by the eye model using polar coordinates for the iris centre [5].

4.1 Individual Contour-Based HT

Since *contour segmentation* has already been carried out, the HT is applied to each contour fragment separately. Although slightly less computationally efficient than doing a global transform, it avoids problems of ‘cross-contamination’ [4], so giving very clean peaks. Figures 13 and 14 illustrate typical accumulator arrays for a gradient approach and one using pre-processing and individual contour assessment. Since each edge point is only processed once, the additional computational cost is not excessive, and the preprocessing reduces the number of edge points considered. The largest response from the individual Hough arrays supplies the correct iris position in 99.9% of cases. In order to achieve this performance, the following problems need to be overcome: small iris contours due to extreme iris positions; blinking and excessive iris occlusion; iris fuzziness.

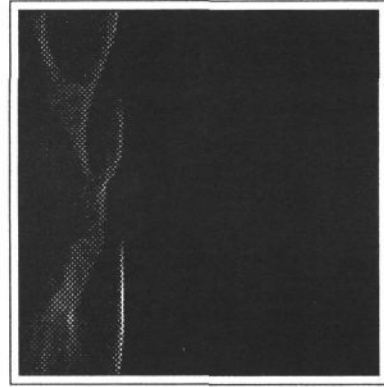
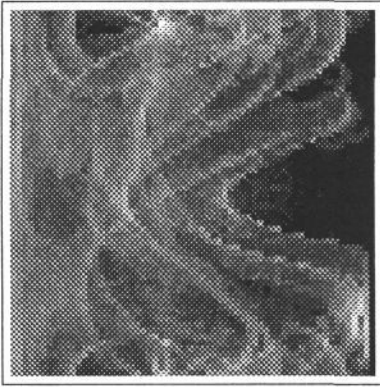


Figure 13: Typical gradient HT accumulator Figure 14: Typical contour HT accumulator

4.2 Iris Occlusion and Blinking

Extreme eye positions, containing very limited iris data, though relatively rare with total head motion freedom [1], still require accurate detection. In order to do this, a weighting scheme is used in choosing the correct HT peak by compensating for loss of edge points in extreme iris positions.

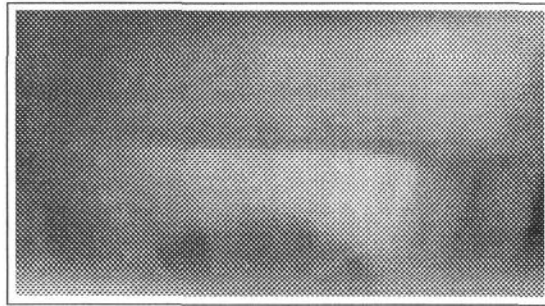


Figure 15: Extreme iris position (note the dominant eye edge rather than the iris)

The problem is that central iris positions can be favoured simply because longer contours are available to support them, even if that contour taken as a whole is not as good a fit to the iris ellipse as a more peripheral contour. The measure used is therefore the ratio of the HT support to the number of ellipse points present in the image – a percentage measure of support. In this situation more centrally-placed ellipses require greater HT support than more peripheral examples. When no contour score exceeds 40% it is assumed that no iris can be detected and the process repeated with a lower edge detection threshold. Further failure concludes that the iris is too small to detect – either due to extreme occlusion or blinking.

4.3 Iris Outline Blurring and Non-elliptical Iris Contours

Though at first glance most iris edges to be clearly defined, close inspection often reveals a large degree of boundary fuzziness, especially in subjects with light irises.

Furthermore the iris is often not perfectly elliptical, causing curved iris boundaries to appear straight (Figure 16). In order to allow for inaccuracy in edge determination, the edge data input to the Hough Transform is blurred slightly, using a 3x3 Gaussian mask of standard deviation 0.5 pixels (Figure 17). This allows pixels neighbouring image gradient peaks to contribute to the HT, supplying the added advantage that the eye model parameters do not need to be precisely accurate – errors of a few pixels are acceptable.

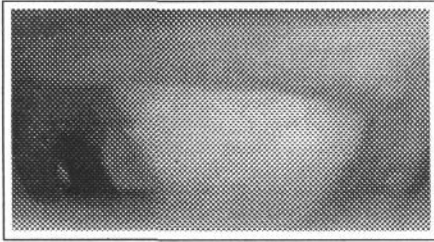


Figure 16: Original iris image

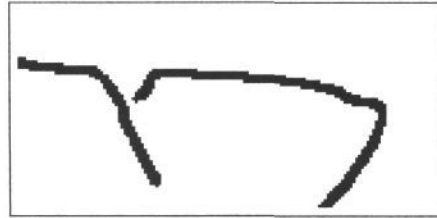


Figure 17: Squared-off iris contour

5 Ellipse-Fitting Results

The algorithm was tested on three available image sequences (two whilst car driving under very variable lighting conditions, the other under artificial indoor lighting). Between them, these sequences contained 1100 frames of irises at almost all allowable locations, of which 1099 were successfully detected (under non-blinking conditions). Results for the irises in this paper, currently achieved using off-line processing, are shown in Figures 18 to 21.

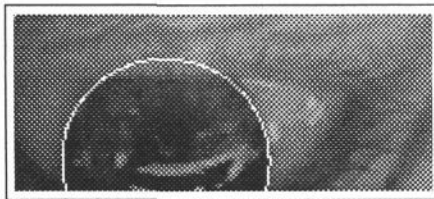


Figure 18: Highlighted iris ellipse-fitting

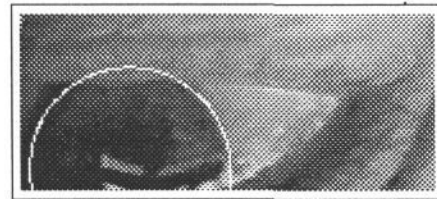


Figure 19: Dominant edge iris ellipse-fitting

6 Gaze Position Determination

From the iris centre position, and the eye model, the direction of gaze relative to a head-fixed coordinate system can easily be obtained, using equations such as those of [2]. This direction can then be related to position in the image of the world by means of a calibration step in which the subject looks at a variety of easily discernable features, such as marks on a plain surface, using predominantly eye rather than head movements. The absolute positions of these features are not

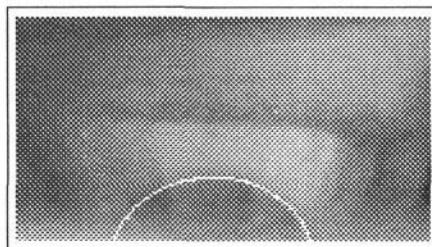
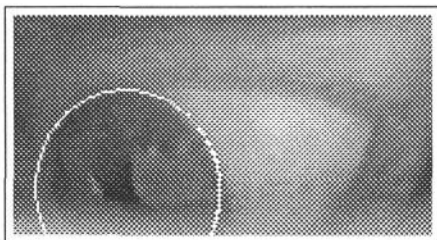


Figure 20: Contrast-varying iris ellipse-fitting

Figure 21: Small iris ellipse-fitting

needed, but it must be possible to localise their positions in the image.

To calibrate the system, we assume that the subject is fixating the target mark and determine the iris centre and the position of the intersection of the line of sight with the image plane. We model the relationship between the camera's and the eye's views of the world by assuming an affine transformation between the coordinates of the line of sight intersection and the coordinates of the image of the target. Given a reasonably large number of targets, the parameters of the transformation can be fitted by least squares, and the RMS residual then gives an estimate of the accuracy of the system in terms of image pixels (or, with camera calibration, in terms of gaze angle).

The performance, which approached that achieved by hand-fitting the ellipse, was consistently better than 2 deg – RMS errors of 0.85 deg horizontally, 1.49 deg vertically. Figure 22 shows a vector representation of these errors for each fixation point chalked in a rectangular grid on a vertical frontoparallel blackboard. The discrepancy between horizontal and vertical errors would appear to be directly related to the degree of iris contour *structure* present. Fixation errors in the central field of view are small due to the large portion of iris contour available for fitting. As the iris moves horizontally to the left or right there is always at least one whole side of contour available (dependent, obviously on eyelid occlusion during blinking). Similarly, movements directly above and below the centre also supply good constraining information (both sides of the iris are available – albeit in small amounts). The main problems occur for large diagonal fixations in which two 'sides' of each contour are effectively occluded and the constraints available are severely reduced. In such cases hand-fitting does not reduce the error as it is due to insufficient information rather than inaccuracies in the model or image processing. The use of pupil information is a possible method of improvement – though the pupil may also be occluded for extreme diagonal-peripheral positions – or manually drawing back the eyelids [2], though this is more suited to a laboratory-based system in which the user has limited mobility.

7 Conclusions

The present method involves considerably more processing than using raw edge data and, although the simpler approach works in a large percentage of cases, it is

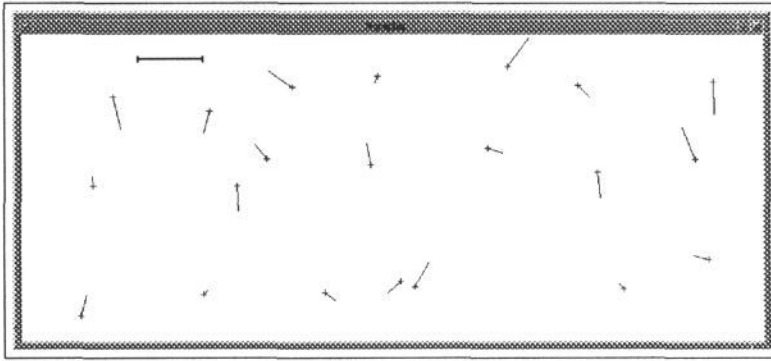


Figure 22: Vector Fixation Errors. The fixation point positions are at the base of each vector and the horizontal line measure represents 5 deg. Note that the 'grid' of points is not regular as the vectors are plotted at the image point which they represented for that particular frame. The distortions are therefore due to head motion during calibration.

susceptible to gross errors caused by highlights and eyelid boundary contours. The use of the pre-processing techniques described in this work reduces the error rate from around 10% to 0.1% and is therefore considerably more robust. The accuracy of the system approaches that obtainable by laborious hand-fitting of the iris ellipse and thus supplies a useful, potentially real-time, automation method.

Acknowledgements

Supported by the ESRC/MRC/SERC Initiative in Cognitive Science/HCI and the SERC Image Interpretation Initiative. The authors would like to thank Mike Land and his technicians for the design and construction of the eye-tracker.

References

- [1] Land ML, "Eye-head coordination during driving", Proceedings of the Systems, Man and Cybernetics Conference, **3**, 1993, pp 490.
- [2] Robertson G, Craw I and Donaldson B, "Human eye location for quantifying Eye Muscle Palsy", Proceedings of the British Machine Vision Conference, 1994, pp 357-366.
- [3] Canny J, "A computational approach to edge detection", Pattern Analysis and Machine Intelligence **8**, 1986.
- [4] Yuen K, "Connective Hough Transform", Proceedings of the British Machine Vision Conference, 1991, pp 127-133.
- [5] Young, DS, Tunley, H and Samuels, R, "Specialised Hough transform and adaptive contour methods for real-time eye tracking", 1995, (in preparation).