

Facial Features Localisation

Robust Eye Centre Extraction Using the Hough Transform

David E. Benn, Mark S. Nixon and John N. Carter

Department of Electronics and Computer Science, University of Southampton,
Highfield, Southampton SO17 1BJ, England.

Email: deb95r@ecs.soton.ac.uk and msn@ecs.soton.ac.uk

Abstract. Finding the eyes is an important stage of feature extraction in automatic face recognition. Current approaches include standard feature extraction techniques using heuristic methods specifically developed for human eyes. We present a new method for finding eye centres using a gradient decomposed Hough Transform (HT) which embodies the natural concentricity of the eye region in a peak reinforcement scheme to improve accuracy and robustness. This enhances a standard feature extraction technique with an analytic approach, which can be applied to the whole face without priming of estimates of eye position and size. In a database of 54 eyes this new method is shown to be less constrained, more robust and resulted in a three-fold improvement in accuracy over using the standard HT.

1 Introduction

Automatic face recognition research has many potential commercial applications such as for access control or face reconstruction. An attractive benefit of automatic face recognition is that, unlike finger print analysis, it requires no contact. Chellappa [2] reviews how face feature extraction can be performed using a variety of techniques including statistical moments, eigen-faces and singular value decomposition. Geometric measurement of face components such as the distance between the eyes, width of mouth, length of nose etc. provides a simple method of effective feature extraction. The eyes are spatially the most well defined structures on the face and the distance between them varies little with different facial expressions. Thus, measurements of eye spacing have been used to provide invariance to distance from the camera and normalise other face measurements [4] [9].

Riesfield and Yeshurun [7] [8] used a generalised symmetry operator to locate the eyes. A face usually exhibits a high degree of symmetry about a line drawn downwards through axis of the nose. For example, the eyes are either side of the nose axis as well being locally symmetric. Thus the symmetry operator can be used to locate the position of the eyes. One apparent advantage of the approach over correlation based method such as [4] is that it is independent of scale or orientation. However since no *a priori* knowledge of the face is used in the search for symmetry, the process is computationally intensive. The authors mention a success rate of 95% providing the faces occupy between 15-60% of the image. The method does not use the explicit circular shape of the iris, which offers the potential to improve the success rate.

A deformable template was used in [12] [10] to locate the eyes. The eyes can be modelled mathematically as a circular iris which is enclosed or occluded by two parabolae, the eyelids, containing two white regions, the sclera [11] [12]. In order to determine the parameters of the iris, sclera and eyelids, appropriate energy functions

are defined using the information about the valleys, edges, peaks and intensity of the face image. The eye template interacts with the face image by adjusting its parameters to minimise a composite energy functional. Thus, finding the eyes reduces to numerical optimisation of minimising the energy function. The template is sufficiently flexible to locate the eyes despite variations in size, orientation and lighting conditions. More shape information is used, however sometimes the algorithm not converge to the desired result. The algorithm was computationally costly, involving sequential change in the values of up to 11 parameters, followed by numerical optimisation of the energy term. The success of the process was very sensitive to initial starting conditions. When the template was initially positioned above the eyebrows the algorithm failed to distinguish between the eyebrows and the eyes. In [11] Xie et al used a Kalman filter to track a deformable template's results [10]. It was claimed that by using a better choice parameters and optimisation techniques, the problems resulting in non-convergence suffered in [12] were alleviated, however the size of their database is not reported.

In [9] Stringa uses the bright spot of light reflected from an illuminating light source to locate the pupil. Even though it is possible for the light spot be absent, Stringa relies on this and then postulates that the horizontal intensity variation across the eye will be symmetric about this bright spot, with a change in intensity between sclera and iris and another sharp change in intensity between bright spot and iris. The shape of the eye is used only indirectly, via the intensity variations. Stringa reported 100% success on a database of 333 faces with a fast processing time. However, the technique is intrusive and may not be so successful in cases where an artificial light source is not be available.

A standard Hough transform (SHT) has been used [6] to detect the instance of a circular shape and an ellipsoidal shape which approximate the perimeter of the iris and sclera respectively. Both the measured iris and sclera centre were used to provide eye spacing measurements. The iris centre measurements were within ± 2 pixels of the subjective estimate of the eye centre. A recent approach has employed a non explicit form of concentricity [5] and achieved good results. However, the databases used in [6] and [5] were relatively small comprising of only 6 pairs of eye measurements.

In section 2, we describe the new HT formulation; in section 3 we describe its implementation. In section 4 we illustrate that a simple peak count in the SHT as used in [6] may not correspond to pixels of the whole iris, instead sometimes the arch of eyebrows or eyelids. Attempts to locate the iris based on a count per unit radius tend to favour small circles in the corner of the sclera. As such, none of the approaches hitherto have used one particularly strong feature of the eye, namely that it is the centre of a set of concentric shapes. This allows extension of a proven technique (the HT provides a result equivalent to matched filtering) in a non-heuristic manner. The resulting technique requires few parameters, and none specific to eyes. Also, it offers potential for finding the eyes without priming by prior knowledge. Since the eyes are the only feature with this concentricity a technique formulated to use it will deliver the eyes automatically, when applied to a whole face image. The eyebrows enclose the eye lids which enclose the iris which contains the pupil. Thus, concentricity affords the basis for a new method of eye location.

2 HT for Circles using Gradient Information

The standard HT has excessive computational requirements in complexity and memory. To alleviate this we have used a gradient decomposed HT [1]. A circle can be described as

$$(x - x_o)^2 + (y - y_o)^2 = r^2 \quad (1)$$

where (x, y) is a point the circle with centre (x_o, y_o) and radius r . Differentiating with respect to x yields

$$\frac{dy}{dx} = (x - x_o) / (y - y_o) \quad (2)$$

where dy and dx are the vertical and horizontal components of the gradient intensity at the point (x, y) . By substitution,

$$x_o = x \pm \frac{r}{\sqrt{1 + dx^2/dy^2}} \quad (3)$$

$$y_o = y \pm \frac{r}{\sqrt{1 + dy^2/dx^2}} \quad (4)$$

Thus, the problem of finding the centre of the circle can be reduced to finding the peak in two 2D accumulators. The centre of the concentric circles is located at the peak of the summation of x_o and y_o over the radii of interest. Defining

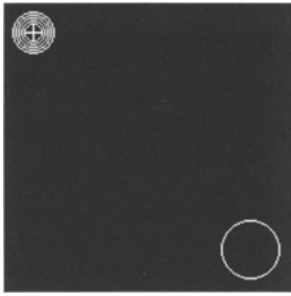
$$X_o = \sum_{r \min}^{r \max} x_o \quad (5)$$

$$Y_o = \sum_{r \min}^{r \max} y_o \quad (6)$$

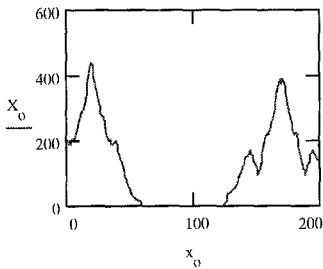
where $rmin$ and $rmax$ are the minimum and maximum likely radii of the iris. Then the centre of the circle is given by the co-ordinates

$$(\max(X_o), \max(Y_o)) \quad (7)$$

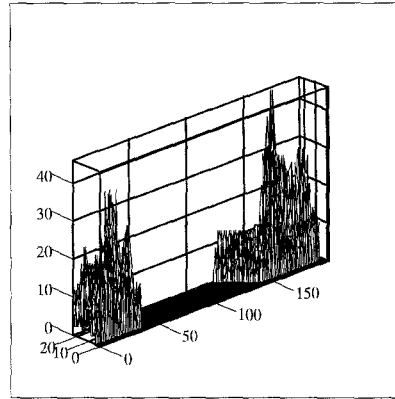
Applying this technique to the synthesised image of Fig. 1(a), the cross shows that the group of concentric circles (top left) yields a higher count than the single larger circle (bottom left). Fig. 1(b) shows the accumulator space corresponding to Fig. 1(a). Note that the larger circle yields the highest count in the accumulator space. However, Fig. 1(c) show the X_o histogram calculated by summing the counts over the radii of interest, thus finding the centre of concentricity. Note that the smaller concentric circles combine to produce a higher peak than the single larger circle. If we now extend the analogy to the eye region, a large number of small concentric circles such as pixels within the iris, sclera and pupil, should produce a better result than a single large count from say, an eyebrow.



(a). Synthetic Circles



(c). Summed Accumulator



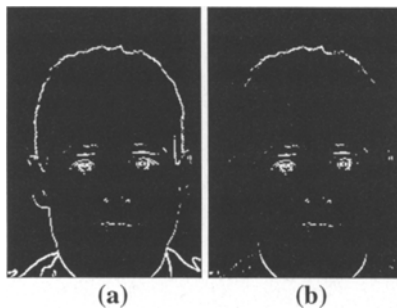
Accumulator

(b). Accumulator for Synthetic Circles

Fig. 1 Extracting Synthetic Eye Centres.

3 Implementation

In theory, the new approach can be applied to the whole face and the two peaks in the X_0 histogram provide the x co-ordinate of the eyes. In practice it is necessary to pre-filter the edge map to remove edge points that lie on vertical or horizontal straight lines which can introduce valid but, undesirable solutions to equations (3) and (4) and a peak in the accumulator which corresponds to a straight line, instead of a circle. The sides of the face can contribute to this undesirable effect, Consequently, straight lines consisting of more than 40 pixels and $\pm 10^\circ$ of vertical or horizontal were removed.

**Fig. 2.** Edge Map Before and After Removing Straight Lines.

The new method was then applied to the resultant edge map. The face images measured approximately 300×400 pixels. Fig. 2 shows a face edge map produced using a 3×3 Sobel edge detector implementing dx and dy from equation (2) followed by uniform thresholding above 160, removing the sides of the face as described above, see Fig. 2(b). Fig. 4 shows a processed face image with a cross to mark the detected eyes. First x co-ordinates of the eyes corresponding to the two prominent peaks in the X_o histogram are found using the whole face (see Fig. 3(a)). The face can then be split into two halves, the centre line being midway between the eyes found using the X_o histogram. The y co-ordinates of the eyes are found using separate accumulators for the left and right side of the face, thus compensating for any non vertical orientation of the face (see Figs 3(b) and 3(c)). The SHT requires a 3-dimensional accumulator in r , x , y . In order to minimise memory requirements of the SHT, both methods were re-applied in a window measuring 110×110 pixels centred on the extracted eye centres.

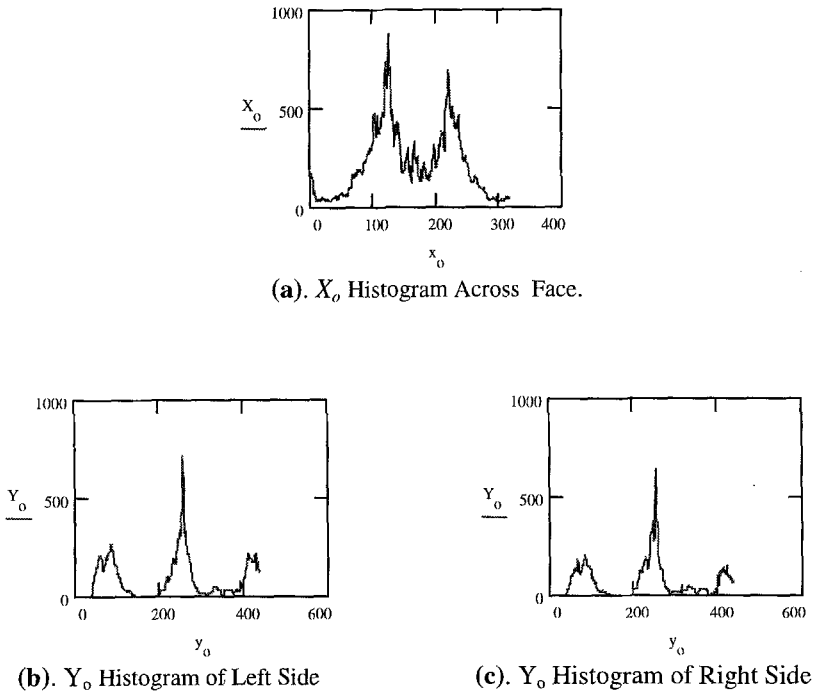


Fig. 3. Concentricity Histograms for the Whole Face.

4 Results

Fig. 5 clearly illustrates the advantage of the new method. The extracted eye centre using the new method is marked with a "+" while an "x" is used to mark the eye centre extracted via the SHT. Using the concentricity property of the eye, a cluster of pixels, centred around the pupil are extracted. In contrast the SHT located a large arch of an eyelid which is not centred on the pupil. Table 1 shows the statistics of the difference

between the estimates of the eye co-ordinates provided by the standard HT and the new formulation, both from the values of eye co-ordinates estimated manually. Table 1 gives the statistics for a large radius variation (as to be expected in application) and for a smaller radius variation. (A small radius variation can be justified in applications where face contour extraction [3] precedes eye location and primes a small expected radius variation.) The difference between the eye co-ordinates is less between the new formulation and manual estimates, compared with the standard HT. For wide variations, the mean difference for the new formulation on the left eye is 1.8 pixels, whereas for the same eye the mean difference for the SHT is 6.8 pixels. For the right eye, the mean difference is 2.5 and 6.4 pixels for the new and the SHT methods respectively. Reducing the potential radius variation improves the mean difference for the new concentricity formulation which is still considerably lower than for the SHT, being nearly half its value. Overall, for the 54 eye database, the use of concentricity in the new technique resulted in a mean difference of 3 pixels, whereas the mean difference for the SHT was 10 pixels, a three-fold improvement in accuracy.



Fig. 4. Extracted Eye Centres

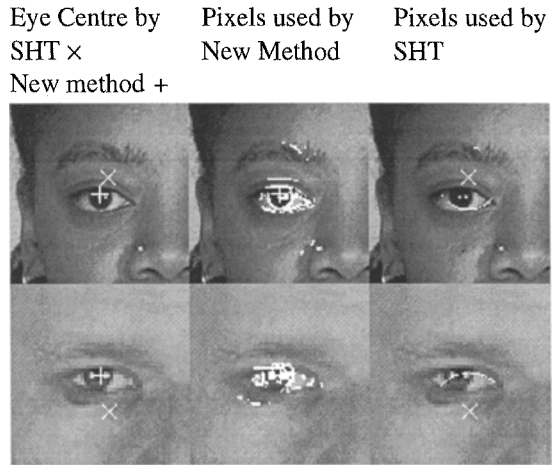


Fig. 5. Results and Data for SHT and New Method

method		SHT (left)			new (left)			SHT (right)			new (right)		
actual - measured		x	y	d	x	y	d	x	y	d	x	y	d
Mean error	radius=4..25	3.6	10	11	2.4	1.3	3.1	3.8	9	10	3.3	1.7	4.1
Sdev error	radius=4..25	2.5	5.3	5.4	3.6	2	3.9	3.1	4.9	5.2	2.6	2.5	3
Mean error	radius=4..10	2.5	2.7	4.3	1.4	0.7	1.9	1.7	1.9	3.1	1.8	1.7	2.8
Sdev error	radius=4..10	2.8	3.8	4.1	1.6	1	1.6	3.5	3.2	4.4	2.4	3.2	3.7
x=measured error in x_0 y=measured error in y_0 d= Euclidean error													

Table 1. Comparing SHT and new method for finding eye centres

5 Conclusions and Further Work

Concentricity affords a new basis to extend the HT for eye extraction in a manner which is well suited to the normal appearance of eyes in a face image. The standard HT (SHT) can provide good results when a good prior estimate of iris radius is known, but even under these conditions the new method can perform better than the SHT. Under wider ranges of radius, the new method performs much better than the SHT and thus appears to be a more robust and accurate method for extracting eye centres in face images. We have applied the new method over the *whole* face without the need for an intermediate eye window location stage or *a priori* knowledge of the face. Our results show that the eyes are the centres of concentricity on the face and this can be used as the basis for simple but, very effective eye feature extraction. Furthermore, the accuracy of our eye model, by concentric circles, may improved by considering the eye as the centre of concentric ellipses.

6 Acknowledgements

We acknowledge gratefully financial support provided by the Home Office (UK) and the interest shown by Tony Kitson and Bob Nicholls in the research.

7 References

- [1] A. S. Aguado, M. E. Montiel and M. S. Nixon, "On Using Directional Information for Parameter Space Decomposition in Ellipse Detection", *Patt. Recog.*, **29**(3) pp. 369-381, (1996).
- [2] R. Chellappa, C. Wilson and S. Sirohey, "Human and Machine Recognition of Faces: A survey". *Proc of the IEEE* **83** (5) pp. 704-741, (1995).
- [3] S. R. Gunn and M. S. Nixon, "Snake Head Boundary Extraction Using Global and Local Energy Minimisation", *Proc. 13 ICPR*, **2**, pp. 581-585, (1996).
- [4] X. Jia and M. S. Nixon, "Extending the Feature Vector for Automatic Face Recognition". *IEEE Trans. PAMI.*, **17** (12) pp. 1167-1176, (1995).
- [5] R. Kothari and J. L. Mitchell, "Detection of Eye Locations in Unconstrained Visual Images", *Proc. ICIIP '96*, **III**, pp. 519-523, (1996).
- [6] M. Nixon, "Eye Spacing Measurement for Facial Recognition". *SPIE Proc.*, **575**, pp. 279-285, (1985).
- [7] D. Riesfield, H. Wolfson and Y. Yeshuran, "Context-free Attentional Operators: The Generalised Symmetry Transform". *Int. J. of Comp. Vision* **14** pp. 119-130, (1995).
- [8] D. Riesfield and Y. Yeshuran, "Robust Detection of Facial Features by Generalised Symmetry". *Proc. 11th Int Conf. on Patt. Recog.*, pp. 117-120, (1992).
- [9] L. Stringa, "Eyes Detection For Face Recognition". *Applied Artificial Intelligence* **7** pp. 365-382, (1993).
- [10] X. Xie, R. Sudhakar and H. Zhuang, "On Improving Eye Feature Extraction using Deformable Templates". *Patt. Recog.*, **27**(6) pp. 791-799, (1994).
- [11] X. Xie, R. Sudhakar and H. Zhuang, "Real-Time Eye Feature Tracking from a Video Sequence Using Kalman Filter". *IEEE Trans. on SMC*, **25** (12) pp. 1568-577, (1995).
- [12] A. Yuille, D. Cohen and P. Hallinan, "Feature Extraction from Faces using Deformable Templates". *Int. J. Comp. Vision* **8**(20) pp. 99-111, (1989)