Attention Based Facial Symmetry Detection

Fred Stentiford

UCL Adastral Park Campus, Martlesham Heath, Ipswich, UK
f.stentiford@adastral.ucl.ac.uk

Abstract. Symmetry is a fundamental structure that is found to some extent in all images. It is thought to be an important factor in the human visual system for obtaining understanding and extracting semantics from visual material. This paper describes a method of detecting axes of reflective symmetry in faces that does not require prior assumptions about the image being analysed. The approach is derived from earlier work on visual attention that identifies salient regions and translational symmetries.

1 Introduction

Symmetries abound both in man made objects and in the structures to be found in nature itself. Symmetry is an important feature in natural scenes that attracts our attention and seems to guide the process of recognition. This has motivated many studies of symmetry and associated techniques that might be applied to image processing.

Symmetry analysis compares image regions and their transforms through translation, rotation and reflection in order to detect relevant structure. Most approaches avoid exhaustive search and reduce the enormous computational requirements by measuring intuitive features that characterise the presence of symmetrical structures. Marola [1] describes a method that can only be applied to shapes that are almost symmetric and requires the computation of the centre of mass. Sun et al [2] also make the assumption that the image is symmetric and measure the correlation between orientation histograms to detect planes of symmetry. Loy et al [3] use gradients to detect points of radial symmetry, but encounter problems of noise which are offset to some extent through the introduction of thresholds. Gradients and edges are also used by Reisfeld et al [4] who requires that symmetry transforms are local. Autocorrelation peaks are employed to determine the presence of symmetry in research by Liu et al. [5]. It was observed in this approach that significant parts of the image were overwhelmed by large expanses of background and that geometric distortions affected the results. Kiryati et al [6] develop a measure of local symmetry which is optimised using a probabilistic genetic algorithm. In the context of faces Mitra et al. [7] require an initial manual indication of the axis of symmetry, and Wu et al [8] need an alignment stage between the original and a reflected version. Symmeter [9] are able to measure the level of symmetry in faces but only if the axis is provided.

The approach taken in this paper is based upon a model of human visual attention [10] that identifies what is important in a scene. The next sections briefly outline this model and how it is modified to extract reflection symmetries. Some illustrative results on human faces are provided.

S. Singh et al. (Eds.): ICAPR 2005, LNCS 3687, pp. 112 – 119, 2005.
© Springer-Verlag Berlin Heidelberg 2005

2 Visual Attention

Salient regions in images may be detected through a process that compares small regions with others within the image. A region that does not match most other regions in the image is very likely to be anomalous and will stand out as foreground material. For example, the edges of large objects and the whole of small objects normally attract high attention scores mainly because of colour adjacencies or textures that only occur rarely in the image. Repetitive backgrounds that display a translational symmetry are assigned low attention scores. No weight is given to the presence or otherwise of reflection or rotation symmetries.

Region matching requires a few pixels (a fork) within that region to match in a translated position in another region. If the difference in colour of one pixel pair exceeds a certain threshold a mismatch is counted and the attention score is incremented.

Let a pixel $\boldsymbol{x}$ in an image correspond to a measurement $\boldsymbol{a}$ where

$$\boldsymbol{x} = (x_1, x_2) \text{ and } \boldsymbol{a} = (a_1, a_2, a_3)$$

Define a function $\boldsymbol{F}$ such that $\boldsymbol{a} = \boldsymbol{F}(\boldsymbol{x})$.

Consider a neighbourhood N of $\boldsymbol{x}$ with radius $\boldsymbol{r}$ where

$$\{\boldsymbol{x}' \in N \text{ iff } |x_i - x'_i| < r_i \ \forall\ i\}$$

Select a fork of m random points S_x in N where

$$S_x = \{\boldsymbol{x}'_1, \boldsymbol{x}'_2, \boldsymbol{x}'_3, \ldots, \boldsymbol{x}'_m\}$$

Shift S_x by a displacement $\boldsymbol{\delta}$ in the image to become S_y where

$$S_y = \{\boldsymbol{x}'_1 + \boldsymbol{\delta}, \boldsymbol{x}'_2 + \boldsymbol{\delta}, \ldots, \boldsymbol{x}'_m + \boldsymbol{\delta}\} \text{ and } \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\delta}$$

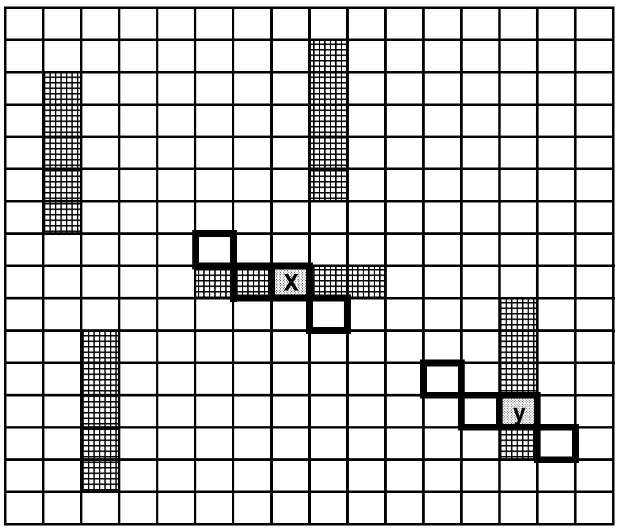

Fig. 1. Fork at $\boldsymbol{x}$ mismatching at $\boldsymbol{y}$ with $\boldsymbol{\delta} = (6,4)$

The fork S_x matches S_y if

$$|F_j(\boldsymbol{x}'_i) - F_j(\boldsymbol{x}'_i + \boldsymbol{\delta}_i)| < \varepsilon_j \quad \forall\, i,j\ .$$

In Fig. 1 a fork of m = 4 pixels x' is selected in the neighbourhood of a pixel $\boldsymbol{x}$ and is shown mismatching in the neighbourhood of pixel $\boldsymbol{y}$. The neighbourhood of the second pixel $\boldsymbol{y}$ matches the first if the colour intensities of the corresponding pixels all have values within $\boldsymbol{\varepsilon}$ of each other. The attention score $V(\boldsymbol{x})$ for each pixel $\boldsymbol{x}$ is incremented each time a mismatch occurs in the fork comparisons with a sequence of pixels $\boldsymbol{y}$. A location $\boldsymbol{x}$ will be worthy of attention if a sequence of t forks matches only a few other neighbourhoods in the space. Pixels $\boldsymbol{x}$ that achieve high mismatching scores over a range of t forks $S_{\boldsymbol{x}}$ and pixels $\boldsymbol{y}$ are thereby assigned a high estimate of visual attention. An application to image compression is described in [11].

3 Symmetry Detection

In this paper symmetries are detected using the same mechanism for measuring attention, but transforming forks through reflections *before* translation and testing for a match. Peaks in the distributions of reflection axis angles at which matches are found indicate the locations and strengths of the symmetries present in the image. Forks must include some (h) pixels that mismatch each other otherwise large background tracts of self-matching sky, for example, would appear to exhibit trivial symmetries.

A fork of m random pixels S_x is defined as a set of pixel positions where

$$S_x = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_m\}\ .$$

A series of M such forks is given by

$$S_x^k = \{\boldsymbol{x}_{1k}, \boldsymbol{x}_{2k}, \boldsymbol{x}_{3k}, \ldots, \boldsymbol{x}_{mk}\} \quad k = 1,2,\ldots, M$$

$$\text{with } |F_j(\boldsymbol{x}_{pk}) - F_j(\boldsymbol{x}_{qk})| > \varepsilon_j \text{ for at least } h \text{ values of } p. \tag{1}$$

Randomly translated and reflected forks S_y^k are generated by transforming the S_x^k as follows

$$S_y^k = \{\boldsymbol{y}_{1k}, \boldsymbol{y}_{2k}, \boldsymbol{y}_{3k}, \ldots, \boldsymbol{y}_{mk}\}$$

$$\text{with } \boldsymbol{y}_{ik} - \boldsymbol{y}_{1k} = R_\theta\left[x_{ik} - x_{1k}\right] \quad \forall\, i,k \tag{2}$$

$$\text{and } R_\theta = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix},\ \alpha_1 \le \theta \le \alpha_2$$

where α_1 and α_2 are lower and upper bounds on the random values of θ, the angle of the axis of reflection.

The fork S_y^k is now a reflected and shifted version of S_x^k and matches S_x^k indicating a possible symmetry if

$$|F_j(\boldsymbol{x}_{ik}) - F_j(\boldsymbol{y}_{ik})| < \varepsilon_j \quad \forall\, i,j\ .$$

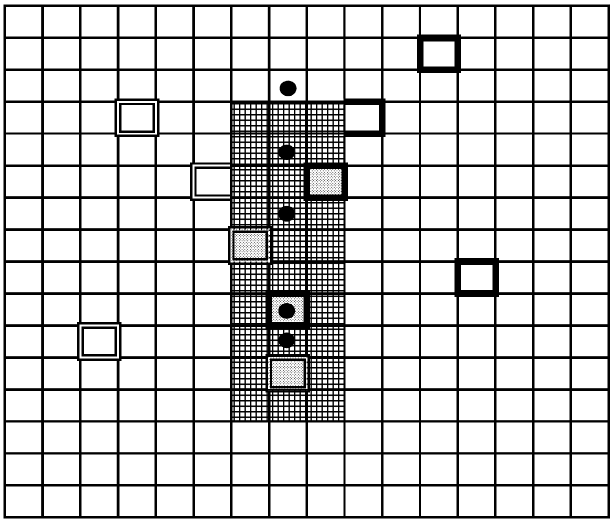

Fig. 2. Symmetric forks matching pattern with $\theta = \pi/2$

Fig. 2 shows a 5 pixel fork and its reflection about an axis at 90º with both forks fitting a vertically symmetric pattern. The mid points of lines joining corresponding fork pixels lie along the axis of symmetry of the shape as indicated by the dots. In this case reflected versions of all white or all black pixel forks would trivially match the background or totally within the shape and are excluded by (1).

Distributions of reflection or rotation symmetries are described in the following steps:

1. Set histogram of reflection axis angles to zero.
2. Generate a fork S_x^k with h pixels mismatching remaining *(m-h)* pixels
3. Reflect S_x^k about an axis at a random angle θ and apply a random shift.
4. If no match is found loop to step 3, P times else increment histogram bin at θ following a match.
5. Loop to step 2, $k = M$ times.

4 Results

Parameter values used to generate forks and symmetry distributions reported here are $m = 12$, $h = 3$, $M = 10000$, $P = 100$, $\alpha_1 = 45º$, $\alpha_2 = 135º$, $\varepsilon = 80$. The location of axes of reflection can be revealed by plotting pixels at the mid points of corresponding pixels in matching forks. The mid points will lie on the axis that was used to generate the reflected fork and a concentration of plotted points will indicate the presence of an axis of reflective symmetry. Fig. 3 shows a grey level face image (276x245) together with a display of the mid points in matching forks, a display of the optimum axis of reflection, and the distribution of reflection axis angles for matching forks. The distributions of mid points and axis angles indicates a range of spurious or less significant symmetries, but the central line of symmetry at 90º predominates.

Some grey level faces (320x243) from the Yale database B [12] are analysed in a similar fashion. Fig. 4 shows a more oval shaped face with more peaked distributions. Figs 5, 6 and 7 have slight tilts producing reflection axis peaks at 86º, 87º and 86º, respectively. In contrast Fig. 8 has a slight tilt to the left with an axis angle of 91º. In addition this face is much rounder and this is reflected in the spread of the midpoint and axis angle distributions.

To test the effectiveness of the symmetry detection on more significant deviations from the vertical, the face in Fig. 4 was rotated by 25º and analysed in the same way. The axis was located at an angle of 112º representing a 23º rotation in Fig. 9.

In addition the face in Fig. 1 was analysed with $0 \le \theta \le 180$. This revealed a secondary horizontal axis of symmetry in Fig. 10 just above the eyes that seems to balance areas of forehead against the cheeks.

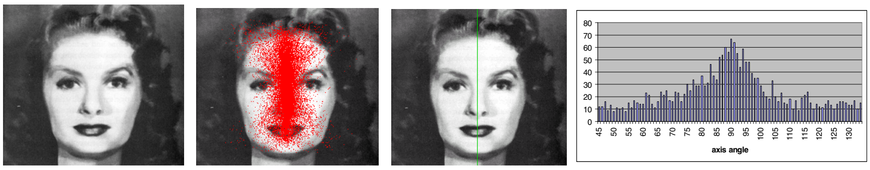

Fig. 3. Original, fork pixel midpoint locations, axis display, and axis angle distribution

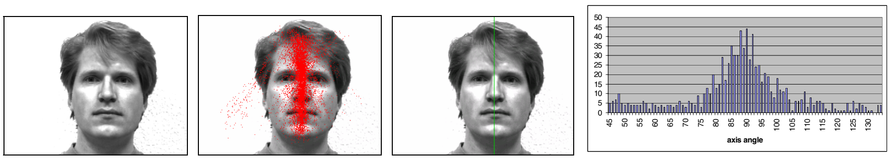

Fig. 4. Original, fork pixel midpoint locations, axis display, and axis angle distribution

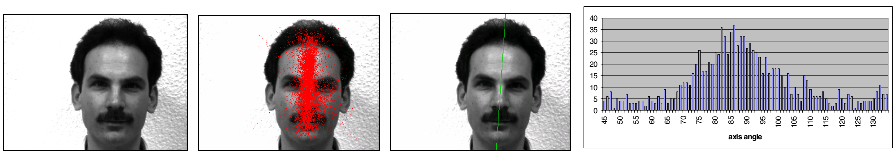

Fig. 5. Original, fork pixel midpoint locations, axis display, and axis angle distribution

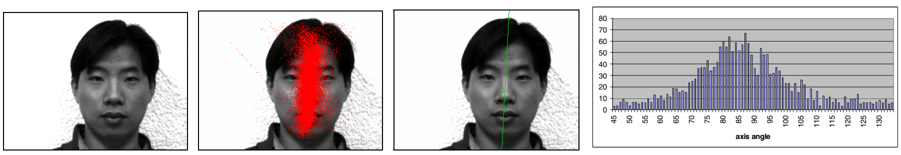

Fig. 6. Original, fork pixel midpoint locations, axis display, and axis angle distribution

Fig. 7. Original, fork pixel midpoint locations, axis display, and axis angle distribution

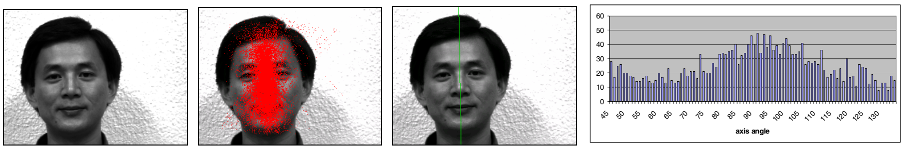

Fig. 8. Original, fork pixel midpoint locations, axis display, and axis angle distribution

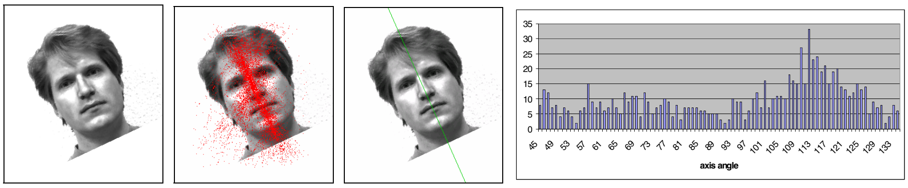

Fig. 9. Original, fork pixel midpoint locations, axis display, and axis angle distribution

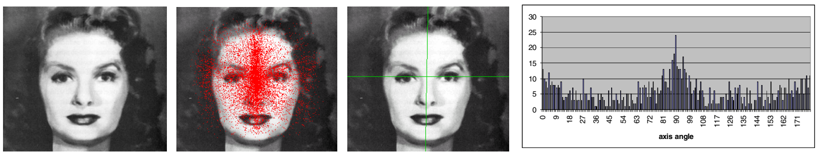

Fig. 10. Original, fork pixel midpoint locations, axis display, and axis angle distribution

5 Discussion

The results presented here only refer to grey level images, but the mechanisms apply equally to colour images. Pixel matching for colour requires that all three colour components independently have values within ε_j of each other. Preliminary experiments indicate that the choice of colour space makes little difference, but that thresholds tailored to specific images yield more informative results.

Key advantages in this approach over other techniques include the absence of any need for the specification of any a priori features that might characterise aspects of symmetrical structures. In addition no restrictions are placed on the minimum strength of any symmetry that must be present in the data for the algorithm to function effectively. Finally there is no manual intervention necessary to either initialise or guide the process. To the author's knowledge this is the first approach that is fully independent of any human involvement and therefore it would be difficult to make any fair comparisons with other methods as it is always possible to provide intuitive heuristics that gain improvements on specific sets of data. However, further work is clearly necessary to measure the errors in the estimation of angles of symmetry on much larger sets of data.

It is worth stressing that the random generation of pixel forks ensures that no solution is unwittingly precluded from the search space by the imposition of guiding heuristics. The universe of possibilities is huge but this should not be a deterrent for a simple trial and error process that is scalable and yields results. Nothing is known with logical certainty about natural image search spaces and we believe that any intuitive assumptions may only have short term benefits.

The method is not specific to the analysis of facial images but can be applied to any pattern. This necessarily means that any symmetrical form appearing in the image that does not align with the facial structure will cause errors. Asymmetrical lighting introduces shadows which do cause serious disturbance and this will be the subject of some future work on illuminant correction.

Rotation symmetries are not analysed in this work as faces do not possess this structure. However, initial experiments replacing R_θ in (2) with the rotation transform $R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ indicate similar success in extracting rotation symmetries when they are present.

The results reported in this paper have been produced with 10000 iterations of fork generation in order to yield accuracies of the order of one degree. Although the computational steps are very simple there are a large number of them and a symmetry analysis takes about 10 seconds on a 1.8GHz machine running in C++. However, the matching of forks can be carried out in parallel as each match is independent of the next and related implementations on the Texas Instruments DM642 DSP platform indicate that processing can take place at video speeds.

6 Conclusions

This paper has described a technique for extracting symmetries from 2D facial images that does not require manual intervention or the prior specification of features that characterise those symmetries. The features or forks are produced through a modified attention focussing mechanism that selects the best combination of positional and reflection transforms that maximises the matching of forks. Future work will be directed at natural colour images and illuminant correction where the objective will be to extract image relationships that can be used in Content Based Image Retrieval applications.

This research has been conducted with the support of BT and within the framework of the European Commission funded Network of Excellence "Multimedia Understanding through Semantics, Computation and Learning" (MUSCLE) [13].

References

1. Marola, G.: On the detection of the axes of symmetry of symmetric and almost symmetric planar images. IEEE Trans on PAMI. Vol. 11. (1989) 104-108
2. Sun, C., Sherrah, J.: 3D symmetry detection using the extended Gaussian image. IEEE Trans. on PAMI. Vol. 19, no. 2. (1997) 164-168
3. Loy, G., Zelinsky, A.: Fast radial symmetry for detecting points of interest. IEEE Trans. on PAMI. Vol. 25. no. 8. (2003) 959-973.
4. Reisfeld, D., Yeshurun, H.W.Y.: Context free attentional operators: the generalised symmetry transform. Int. J. Comp. Vis. Vol. 14. (1995) 119-130
5. Liu, Y., Collins, R.T.: A computational model for periodic pattern perception based on frieze and wallpaper groups. Trans. on PAMI. Vol. 26. no. 3. (2004) 354-371
6. Kiryati, N., Gofman, Y.: Detecting symmetry in grey level images: the global optimization approach. Int. J. Comp. Vis. Vol. 29. (1998) 29-45
7. Mitra, S., Liu, Y.: Local facial asymmetry for expression classification. Proc. CVPR. (2004)
8. Wu, Y., Pan, G., Wu, Z.: Face authentication based on multiple profiles extracted from range data. 4th Int. Conf. on Audio and Video based Biometric Person Authentication. Univ. Surrey, June 9-11, (2003)
9. Symmeter, http://www.symmeter.com/concept.htm
10. Stentiford, F.W.M.: Automatic identification of regions of interest with application to the quantification of DNA damage in cells. Proc. SPIE. Vol. 4662. (2002) 244-253
11. Stentiford, F.W.M.: An estimator for visual attention through competitive novelty with application to image compression. Picture Coding Symposium. (2001) Seoul
12. Georghiades, A.S., Belhumeur, P.N., Kriegman, D.J.: From few to many: illumination cone models for face recognition under variable lighting and pose. Trans. on PAMI. Vol. 23. no. 6. (2001) 643-660
13. Multimedia Understanding through Semantics, Computation and Learning, Network of Excellence. EC 6th Framework Programme. FP6-507752. http://www.muscle-noe.org/