Robust Detection of Facial Features by Generalized Symmetry

Daniel Reisfeld and Yehezkel Yeshurun

Department of Computer Science, Tel Aviv University, 69978, Tel Aviv, Israel

Abstract

Locating facial features is crucial for various face recognition schemes. We suggest a robust facial feature detector based on a generalized symmetry interest operator. No special tuning is required if the face occupies 15-60% of the image. The operator was tested on a large face data base with a success rate of over 95%.

1   Introduction

Recognition of human faces is a problem that appears time and again in the field of computer vision (see [2] for a review). This task, which seems effortless for humans, does not lend itself easily to computational approaches. These approaches can be generally classified into two categories – *feature based* recognition and *Gestalt* based recognition. The feature based approaches focus on detection of individual features such as eyes, nose, mouth, and head outline; and define a face model by the relative position of the features and their size. The approaches inspired by the Gestalt school of perception [11, 8] process faces as entities, where the parts are not the main cues.

The importance of developing an automatic method for locating facial features is self evident in the feature based schemes. It is equally important in the Gestalt approaches; in a typical system faces are stored in at least 128×128 pixel resolution and the problem is transformed to a classification task in $R^{10,000}$ or more. Kirby and Sirovich [6] suggested to use the Karhunen-Loève procedure to reduce the dimensionality of the problem. This was suggested also in [10]. In Figure 1 we demonstrate that this method can be substantially improved by normalizing faces using the eyes and mouth location as anchor point for an affine transformation. Recently [4], we have obtained good recognition rates on normalized faces.

An early system for locating facial features is described in [9, 5]. This system applies a threshold over the laplacian of a wide gaussian convolved with the original image. This operation transform the pictures into a line drawing and then a simple heuristic is applied to find the facial feature. Although the results reported are impressive, this system requires fine tuning for every image. In a more recent attempt, the Sobel operator is applied with a simple contour following heuristic [3]. Again, this cannot be used on real data. Baron [1] locates the eyes by cross-correlating the image with a series of eye templates and looking for sufficiently high levels of correlation. This technique is inadequate for real images with varying sizes, angles, and considering the fact that the eyes are plastic objects.

An attempt to achieve a more complete solution is to use *deformable templates*, which are parameterized models (resembling snakes) of each feature in which the parameter values are determined by interaction with the image [13]. The deformable template success relaies on descending to a local minima of an energy function, whose final parameters characterize the feature we look for. In order for the function to converge to a local minima, which is actually the feature we look for, the initial guess should be close enough. We suggest an operator which enables estimation of these locations. Finding the exact location, can then be performed by simpler and more robust methods.

We have developed an operator that detects the eyes and mouth in images of human faces and can be used as a preprocessor for virtually any face recognition system. No special tuning is required if the face occupies 15-60% of the image. We have tested it on a face image data base (including [10]) with a success rate of over 95%. Some examples are shown in Figure 2.

2   The Generalized Symmetry Operator

Biological and machine vision tasks involve the processing of an enormous amount of information. In order to analyze it under plausible time and space constraints primates direct their computational resources
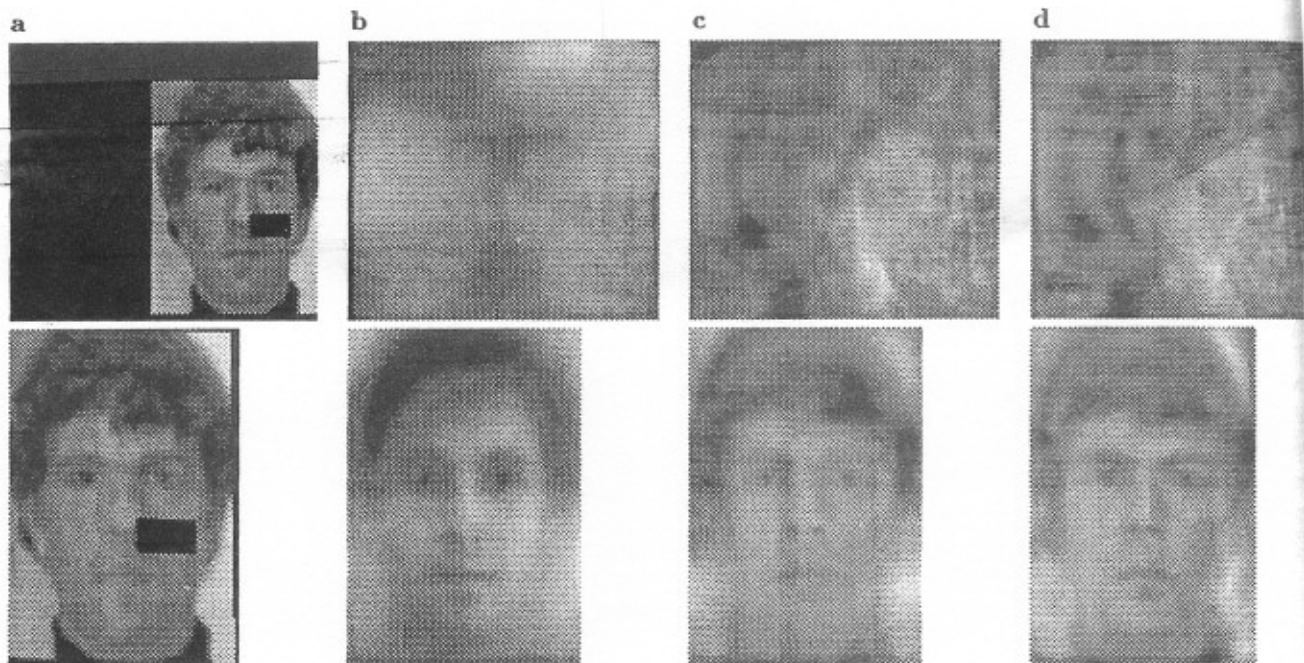
117

Figure 1: Projection of a new face on eigenvectors obtained using Karhunen-Loève procedure (principle component analysis).
Top: (a) A new image presented to the system. (b) The expectancy of the original data base obtained from [10] (432 images). (c) Projection of a on the first 25 eigenvectors. (d) Projection of a on the first 40 eigenvectors.
Bottom: Same procedure applied to a normalized by the location of the eyes and the mouth; the image data base is normalized as well.
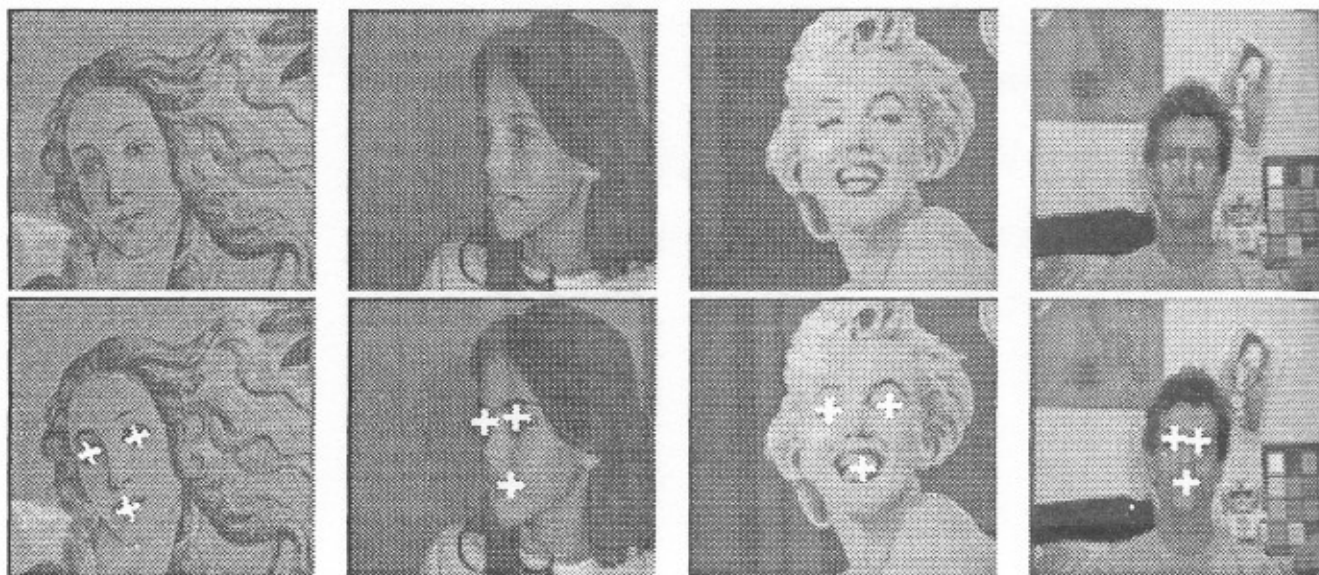

Figure 2: **Top:** Original intensity images. **Bottom:** Crosses mark centers of facial features detected. Notice the variations in size, contrast, posture and orientation of the faces.

toward *attention points* and *fixation points*. We have suggested an *interest operator* based on the intuitive (rather than the formal mathematical) notion of symmetry that can provide a basis for a more thorough investigation by higher processes.

The motivation and formalization of the operator can be found elsewhere [7, 12]. We give here a slightly different formulation of the operator. This formulation was actually used as the first step for facial features detection.

We first define a symmetry measure for each point. Let $p_k = (x_k, y_k)$ be any point ($k = 1, \ldots, K$), and denote by $\nabla p_k = \left(\frac{\partial}{\partial x} p_k, \frac{\partial}{\partial y} p_k\right)$ the gradient of the intensity at point p_k. We assume that a vector $v_k = (r_k, \theta_k)$ is associated with each p_k such that $r_k = \log(1 + \|\nabla p_k\|)$ and $\theta_k = \arctan\left(\frac{\partial}{\partial y} p_k / \frac{\partial}{\partial x} p_k\right)$. For each two points p_i and p_j, we denote by l the line passing through them, and by α_{ij} the angle counterclockwise between l and the horizon. We define the set $\Gamma(p)$, a distance weight function $D_\sigma(i,j)$, and a phase weight function $P(i,j)$ as

$$\Gamma(p) = \left\{(i,j) \,\middle|\, \frac{p_i + p_j}{2} = p\right\}$$
$$D_\sigma(i,j) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\|p_i - p_j\|}{2\sigma}\right)$$
$$P(i,j) = (1 - \cos(\theta_i + \theta_j - 2\alpha_{ij}))(1 - \cos(\theta_i - \theta_j))$$

We define the contribution of the points p_i and p_j as

$$C(i,j) = D_\sigma(i,j)P(i,j)r_i r_j$$

The symmetry measure $M_\sigma(p)$ of each point p is defined as

$$M_\sigma(p) = \sum_{(i,j)\in\Gamma(p)} C(i,j)$$

We define the direction of the symmetry $\beta(p) = \frac{p_i + p_j}{2}$ such that $M(i,j)$ is maximal for $(i,j) \in \Gamma(p)$. Thus, the symmetry of the point p is defined as

$$S_\sigma(p) = (M_\sigma(p), \beta(p))$$

We have used a similar operator to demonstrate the detection of interest points in general [7]. Recently, another symmetry measure was suggested [14].

3 Locating Facial Features

The symmetry operator is a powerful tool for finding interest points in arbitrary natural scenes without a priori knowledge of the world. It can be turned
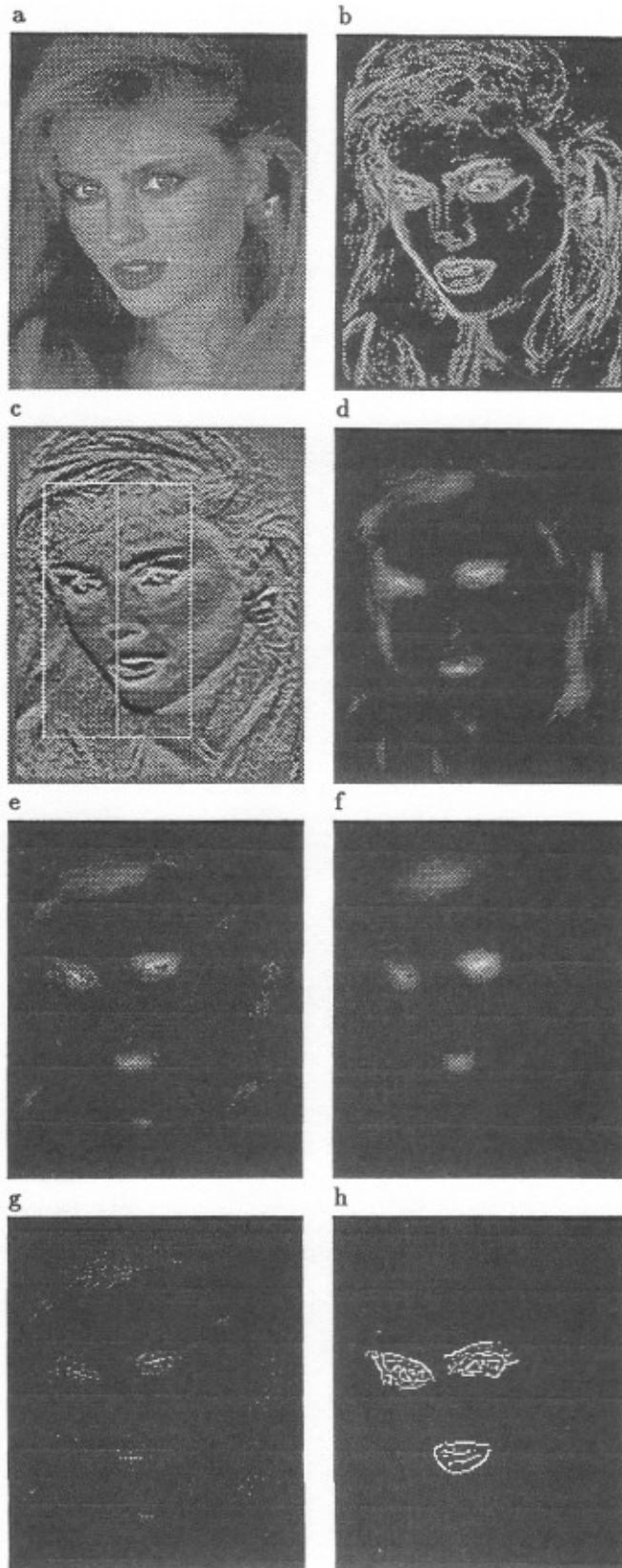

Figure 3: Various stages of the operator. (see text).

into a feature detector by considering domain specific knowledge as demonstrate in Figure 3. (a) is an original image and (b) is its edges result by a convolution with derivatives of gaussians in two directions. We start by detecting the vertical midline of the face, by employing global correlation of directional intensity gradients of left and right half images (c). This can be done in various orientation and the midline is determined according to the best correlation found. We, then expect the eyes to be on the two sides of the midline, the mouth to cross the midline and to have certain geometrical relations between the various locations.

Applying the symmetry operator on an image result in a *symmetry map* which assigns a symmetry magnitude (d) and a symmetry orientation to each point. This kind of representation is similar to edge representation and therefore it can be processed by similar techniques such as edge linking. (e) is the projection of the symmetry map along the normal to the midline.

(f) is the *accumulated symmetry* evaluated by linking the symmetry values based on dynamic programming. Each symmetry point adds to its accumulated value a fraction of the accumulated value of its left neighbours in a recursive manner. The same is done with the right neighbors and the two accumulated values (left and right) are multiplied. In this way the center of each symmetry cluster is singled out as being a local maxima in the accumulated symmetry map.

(g) is another way to reduce the information in the symmetry map by suppressing symmetry values which are not maximal in the direction perpendicular to their symmetry orientation.

The edge values that contributed to the high symmetry value can be selected for segmentation purposes (h).

Combining these processes yields good localization as demonstrated in Figure 2. However, this process is based on a geometrical relations among the features assuming the existence of a single face in the image. It can be made more robust if a pre-trained classifier is applied to validate the existence of the facial features detected; the resulted normalized image can be in turn examined by an additional classifier for the existence of a face.

4 Conclusions

In this work we have demonstrated that the notion of generalized symmetry can be used to define a robust operator, that can detect the centers of the eyes and the mouth for a wide range of sizes, rotations and lighting conditions.

References

[1] R. J. Baron. Mechanisms of human facial recognition. *International Journal of Man-Machine Studies*, 15:137–178, 1981.

[2] V. Bruce and M. Burton. Computer recognition of faces. In A. W. Young and H. D. Ellis, editors, *Handbook of Research on Face Processing*, pages 487–506. Elsevier, North-Holland, 1989.

[3] I. G. Craw, H. D. Ellis, and J. R. Lishman. Automatic extraction of face-features. *Pattern Recognition Letters*, 5:183–187, 1983.

[4] S. Edelman, D. Reisfeld, and Y.Yeshurun. Learning to recognize faces from examples. In *Second European Conference on Computer Vision*, S. Margherita, Ligure - Italy, May 1992.

[5] T. Kanade. *Picture processing system by computer complex and recognition of human faces*. PhD thesis, Dept. of Information Science, Kyoto University, November 1973.

[6] M. Kirby and L. Sirovich. Application of the karhunen-loève procedure for characterization of human faces. *IEEE Trans. Pattern Anal. Machine Intell.*, 12(1):103–108, 1990.

[7] D. Reisfeld, H. Wolfson, and Y. Yeshurun. Detection of interest point using symmmetry. In *Third International Conference on Computer Vision*, pages 62–65, Osaka, Japan, December 1990.

[8] I. Rock and S. Palmer. The legacy of gestalt psychology. *Scientific American*, 263(6):48–61, December 1990.

[9] T. Sakai, M. Nagao, and T. Kanade. Computer analysis and classification of photographs of human faces. In *Proceedings of the first USA-Japan computer conference*, pages 55–62, 1972.

[10] M. Turk and A. Pentland. Eigenfaces for recognition. *J. of Cognitive Neuroscience*, 3:71–86, 1991.

[11] W. Köhler. *Gestalt psychology*. Liveright, New York, 1929.

[12] Y. Yeshurun, D. Reisfeld, and H. Wolfson. Symmetry: A context free cue for foveated vision. In Harry Wechsler, editor, *Neural Networks for Perception, Vol. 1, Human and Machine Perception*, pages 477–490. Academic Press, Inc., San Diego, CA, 1992.

[13] A. L. Yuille, D. S. Cohen, and P. W. Halilinman. Feature extraction from face using deformable templates. In *Proc. CVPR*, San Diego, CA, Jun 1989.

[14] H. Zabrodsky, S. Peleg, and D. Avnir. A measure of symmetry based on shape similarity. In *Proc. CVPR*, Champaign, IL, June 1992.