ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 383-388

Face Recognition: Pre-Processing Techniques for Linear Autoassociators

E.DREGE, F.YANG, M.PAINDAVOINE, H.ABDI

LE2I - Université de Bourgogne
6 Bd. Gabriel, 21004 Dijon, France

Abstract. To improve the performance of a linear autoassociator, we explore the use of several pre-processing techniques: a Sobel operator, a Canny-Deriche operator, and a multiscale Canny-Deriche operator. The gist of our approach is to store, in addition to the original pattern, one or several pre-processed (i.e. filtered) versions of the patterns, here faces. We found that the multiscale Canny-Deriche operator gives the best performance of all models. In the framework of an automated face recognition system, we present also a module to be added prior to the neural network, based on Hough transform, for face location purposes.

1. Introduction

The problem of how to identify people from video tape is increasingly important. For instance, in the context of law enforcement usage of faces, this includes mugshot matching, bank/store security and crowd surveillance. But in most cases, security cameras capture poor quality images that have to be compared to another, usually higher quality, picture. We then have to compensate for the differences in time, and/or angle, lightening, quality, up to paraphernalia, without relying on superficial clues like clothing or hairstyle.

Linear auto-associative memories are one of the most simple and well studied neural-network models [1]. They are widely used as models for cognitive tasks as well as digital signal processing [2], or pattern, including face, recognition. Even though linear autoassociators are known to be quite robust when noise is added to the patterns to be recognized, their performance is rather bad when a *lot* of noise is added to the stimulus. One way of improving performance could be to use some pre- and post- processing of the patterns to be recognized. In this paper we explore the use of some pre-processing techniques using wavelet transform applied to multiscale edges detection. In addition, and in the framework of a more global automated face recognition system, we present a prior technique, based on the Hough transform, to locate a face in a given picture. The output can then be used to feed the neural network.

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 383-388

2. Autoassociators and wavelet transform

Description of the ANN:
The advantage of linear associators in comparison with non-linear models is that they provide for the integration of a very large number of cells in the network. Their implementation is quite easy, because they can be analyzed in terms of the singular value decomposition of a matrix [2][3].

In our description, we follow closely the formulation detailed in [3]. The patterns to be learned are represented by $L \times 1$ vectors $\mathbf{a}_k$ where k is the stimulus number. The components of $\mathbf{a}_k$ specify the values of the pattern to be applied to the L cells of the input layer for the k-th stimulus. The complete set of K stimuli is represented with a $L \times K$ matrix noted $\mathbf{A}$ (i.e., $\mathbf{a}_k$ is the k-th column of $\mathbf{A}$). The $L \times L$ synaptic weight connection matrix between the L input cells is denoted $\mathbf{W}$. In this paper learning occurs by modifying the values of the connection weight between cells using the Widrow-Hoff learning rule. The response of the model to a pattern $\mathbf{x}$ (which may or may not have been learned) is obtained as $\widehat{\mathbf{x}} = \mathbf{W}\mathbf{x}$.

Because auto-associators are generally interpreted as content addressable memories, their performance is evaluated by comparing the output of the system with a test pattern which can be a copy or a degraded version of one of the patterns previously learned by the system. This is achieved by computing similarity measures (most of a time a cosine) between input and output. The larger the similarity between input and output, the better the performance.

A pre-processing using multiscale edges:
Our objective being to improve the performance of the autoassociator when trying to*recognize* a degraded version of a learned stimulus, we decided to explore the possibility of storing one or several filtered versions of the patterns to be stored in addition to the original patterns. We refer to this technique as *pre-processing*. As we are mainly interested with image patterns, we choose filtering techniques meaningful in the this context. Because it is generally agreed that edges are essential for recognition [4], we decided to increase their importance in the image. Quite a number of algorithms have been proposed in the literature to perform edge extraction. We decided to implement three algorithms:

1. the *Sobel* operator (a differential operator) as it is considered as a standard procedure well suited for noiseless images.

2. the *Canny-Deriche* operator because it is known to be optimal for edge extraction in noisy images [5].

3. a multiscale resolution wavelet version of the Canny-Deriche operator, called the *multiscale Canny-Deriche* filter, also referred to as the *wavelet transform* [5]. It has been suggested that this technique should be more efficient than a one-scale resolution [6]. It is a separable filter when applied to 2D images. Its impulse response for a 1D signal (because

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 383-388

of its separability, the filter can be seen as two 1D filters) is given by:

$$f(x) = ksxe^{msx} + e^{msx} - e^{sx} \tag{1}$$

with $k = 0.564$, $m = 0.215$, and where $s = 2^j$ is the scale factor (with, in our case, $j \in \{0, 1, 2, 3\}$), and x being the pixel position. This method is implemented as a wavelet transform using a convolution between the image and the edge detection filter for different scales ($s = 2^j$). As a result, this filter detects edges occuring at different scale resolutions in the image [6].

In order to compare these different techniques, we implemented four models:

1. a standard auto-associator storing 40 different face images 225×151 [2],
2. an auto-associator storing the original 40 face images plus, for each face, a Sobel filtered image of the face (hence a total of 80 face images),
3. an auto-associator storing the original 40 face images plus, for each face, a Canny-Deriche filtered image of the face (hence a total of 80 face images),
4. an auto-associator storing the original 40 face images plus, for each face, four multiscale Canny-Deriche filtered face images (one face image per scale resolution, hence a total of 200 images).

All the models were tested using the same procedure. A random face (previously learned) was chosen, and random noise was added to it before presentation.

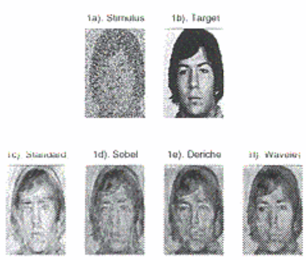

Figure 1: *Response of the models*

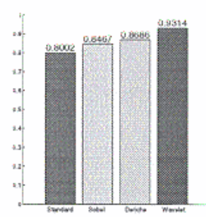

Figure 2: *Correlation of the models*

Fig.1 displays the responses of the models to the test face. The top panels present: 1a) a stimulus with additive random noise added, 1b) the original stimulus. The bottom panels show: 1c) the response of Model 1 (simple autoassociator), 1d) the response of Model 2 (Sobel operator), 1e) the response of Model 3 (Canny-Deriche filter) and 1f) the response of Model 4 (wavelet transform). The quality of recognition can be measured by computing the cosine between the vector $\widehat{\mathbf{x}}$ (i.e., the response of model) and $\mathbf{x}_k$ (i.e., the original stimulus which is also the desired response or target). Fig.2 shows the

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 383-388

correlation between response and target for the 4 models used. Clearly, the standard method is the worst of all. Pre-processing the images improves the performance of the autoassociator, with the wavelet transform giving the best result. In conclusion, the multiscale resolution (i.e., wavelet pre-processing) approach leads to the best performance for the autoassociator. Therefore, we decided, in what follows, to consider only this approach.

Pattern completion of noisy patterns:
We have applied the multiscale edge pre-processing to store a set of 40 Caucasian faces (20 males and 20 females). In order to evaluate further the improvement in performance due to pre-processing, we have decided to test the model with different levels of Gaussian random noise along with 3 different types of testing face: 1) previously learned faces, 2) new faces similar to the learned faces, and 3) new faces coming from an other race than the learned faces.

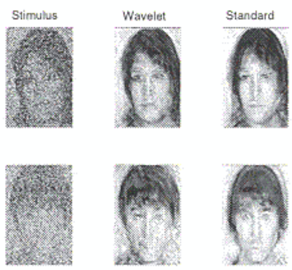

Figure 3: *Stimuli and responses.*

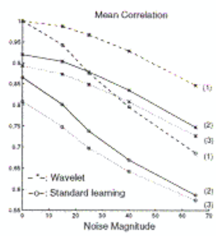

Figure 4: *Mean correlation functions.*

In all cases, pre-processing the image improves the performance of autoassociator with the improvement being more important when the noise added is larger.

3. A module based on Hough transform for face location

The Hough transform [7] is a procedure for locating analytically described shapes (straight lines, circles, parabola, ...). The search pattern is parameterised as a set of vectors from feature points (usually edges) in the pattern to a fixed reference point. With the method, our intention is to determine, the location of human faces in an image. In general, the shape of a human face cannot be described easily with an analytical curve, but can be approximated with an ellipse. If we compute coordinates of the center of the ellipse and the length of its diameters, we can determine borders of the face. This process corresponds to the first step of face identification and face categorization.

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 383-388

Edges detection and thinning:
The Shen and Castan filter [9] is used for edges detection in an image. First, we obtain the edges gradient using a convolution between the original image and the filter. Then, a test using the sign of first and second derivatives of image has been realized to thin the edges.

Hough transform for an ellipse based on an edge gradient:
If we determine a parametrization, each ellipse in the image space can be represented with a single point (the centroid) in Hough space. We exploit sign and value of edge gradient to reduce drastically the computation and the necessary memory. We present the center of an ellipse which goes through edge point (x, y) in the mathematical form using [7]:

$$x_c = x + sign(dLx)\frac{lh}{\sqrt{1+(\frac{lv}{lh})^2(\frac{dLy}{dLx})^2}} \tag{2}$$

$$y_c = y + sign(dLy)\frac{lv}{\sqrt{1+(\frac{lh}{lv})^2(\frac{dLx}{dLy})^2}} \tag{3}$$

Where x_c and y_c determine two coordinates of the center of an ellipse, and lh, lv determine two axes. dLx, dLy have been obtained from the edge pixel information. The ratio $\frac{lh}{lv}$ changes between 1.2 and 1.7 for human face [8], and we choose here $\frac{lh}{lv} = 1.5$ and $lv = 32$ as constants to simplify the computation. Since clearly an ellipse is an unperfect approximation of a face, the Fuzzy Generalized Hough Transformation (FGHT) [8][10] has been used to improve the precision of the system for our application.

Detection of the position of human faces:
To locate the ellipse, each image feature (the edges) is considered to be each of the ellipse features in turn and the corresponding locations of the reference point are calculated. An accumulator array keeps track of the frequency with which each possible reference point location is encountered. Location of human faces in an image corresponds first to search the maximum value in the accumulator array. Then, this value is set to zero to search other potential faces. (see [11] for an illustration)

4. Conclusion

In this paper, we have explored the effects of storing, in a linear auto-associator, filtered versions of face images in addition to the original images. The multiscale Canny-Deriche operator (i.e., a wavelet filter) gives the best performance for a pattern completion task involving degraded a degraded face image. It produces better generalization performance than the control with or without noise added to the image. The larger the amount of noise added, the larger the improvement in performance. Based on those performances and to start building an automatic face recognition system, we presented a module based on the Hough transform for face location. We are now exploring the possibilty of extending the location module to size invariance while working on parallel implementation to reduce the processing times [11].

References

[1] T. Kohonen: Associative memory: A system theoretic approach. Springer-Verlag, Berlin (1977)

[2] D. Valentin, H. Abdi, A.J. O'Toole: Categorization and identification of human face images by neural networks: A review of the linear autoassociative and principal component approaches. Journal of Biological Systems, 2, 413–429 (1994)

[3] H. Abdi: Les Réseaux de neurones. Presses Universitaires de Grenoble, France (1994)

[4] X. Jia, S. Nixon: Extending the feature vector for automatic face recognition. IEEE-Transactions on Patterns Analysis and Machine intelligence, 17, (1995)

[5] E. Bourennane, M. Paindavoine, F. Truchetet, F.: Amélioration du filtre de Canny-Deriche pour la détection des contours sous forme de rampe. Traitement du signal: Recherche, 10, 297–310 (1993)

[6] S. Mallat, Z. Zhong: Characterization of signal from multiscale edges. IEEE-PAMI, 14, 710–732 (1992)

[7] J.Korosec, L.Gyergyek and al.: Face contour detection based on the Hough transformation. Automatika, 31, (1990)

[8] R.Seguier: Détection et localisation de visages dans des séquences d'images vidéo. Ph.D. Thesis Rennes I, France (1995)

[9] J.Shen, S.Castan: An optimal linear operator for step edge detection. Graphical Models and Image processing ,54 (1991)

[10] J.Han, L.T.Koczy, T.Poston: Fuzzy Hough transform. Pattern Recognition Letters, 15 (1994)

[11] F.Yang, E.Drege, M.Paindavoine, H.Abdi: Parallel implementation of a face location algorithm based on the Hough transformation, Submitted.