Improving Generalisation in Radial Basis Function Networks for Face Recognition

A. Jonathan Howell and Hilary Buxton
School of Cognitive and Computing Sciences,
University of Sussex, Falmer, Brighton BN1 9QH, UK
{jonh,hilaryb}@cogs.susx.ac.uk

April 1996

Abstract

This paper presents experiments using an adaptive learning component based on Radial Basis Function (RBF) networks to tackle the unconstrained face recognition problem using low resolution video information. Firstly, we performed preprocessing of face images to mimic the effects of receptive field functions found at various stages of the human vision system. These were then used as input representations to RBF networks that learnt to classify and generalise over different views for a standard face recognition task.

combine the views so that a single output unit corresponds to the individual person. We have taken this idea further and have developed a 'face unit' network model, which allows rapid network training and classification of examples of views of the person to be recognised. These face units give high performance and also alleviate the problem of adding new data to an existing trained network. We are use the various views of the person to be recognised together with selected confusable views of other people as the negative evidence for the network. Our face units have just 2 outputs corresponding to 'yes' or 'no' decisions for the individual. This is in contrast with Edelman et al. (1992) who did not use such negative evidence in their study. We show that this system organisation allows flexible scaling up which could be exploited in real-life applications.

The RBF Network Model

The RBF network is a two-layer, hybrid learning network (Moody & Darken 1988, Moody & Darken 1989), with a supervised layer from the hidden to the

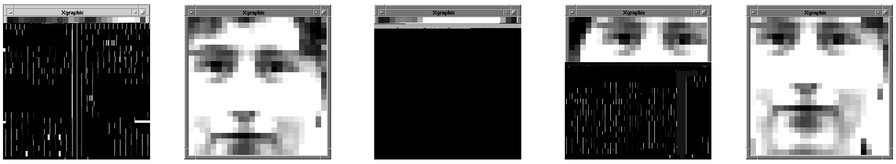
Figure 1: Entire 10-image range (rotating around the y-axis) for one person before preprocessing

matrix pseudo-inverse method (Poggio & Girosi 1990

Form of Test Data

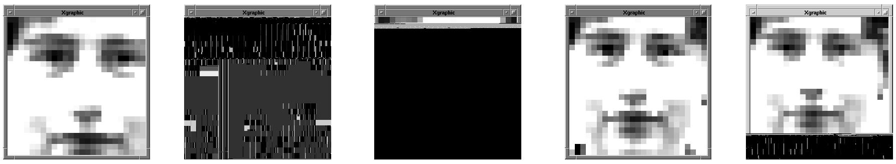
Lighting and location for the training and test face images in these initial studies has been kept fairly constant to simplify the problem. For each individual to be classified, ten images of the head and shoulders were taken in ten different positions in $10°$ steps from face-on to profile of the left side (see Figure 1), $90°$ in all. This gave a data set of 100 8-bit grey-scale 384×287 images from ten individuals.

A 100×100-pixel 'window' was located manually in each image centred on the tip of the person's nose,



(a) (b) (c) (d) (e)

Figure 2: **Shift-varying**



(a) (b) (c) (d) (e)

Figure 3: **Scale-varying** data for the 'face on' view of one individual: (a) +25% (uses 111×111 window) (b) +12.5% (107×107) (c) normal view (100×100) (b) –12.5% (94×94) (b) –25% (87×87)

Shift and Scale Invariance Properties of the RBF Network

Two further data sets were created to test the RBF network's generalisation abilities:

- A shift-varying data set with five copies of each image: one at the standard sampling 'window' position, and four others at the corners of a box where all x,y positions were ±10 pixels from the centre (see Figure 2).
- A scale-varying data set with five copies of each image: one at the standard sampling 'window' size, and four re-scaled at ±12.5% and ±25% of its surface area, ranging from 87×87 to 111×111 (see Figure 3).

Inherent Invariance   Training with Original Images Only

These experiments used only the original from each group of five for training, using all the varied ones (and the remainder of the original ones not used for training) for testing. This gives a measure of the intrinsic invariance of the network to shift and scale, *ie.*the invariance not developed during training by exposure to examples of how the data varies.

(a)

Network	Pre-processing	Initial %	% Discarded	% After Discard
100/400	DoG	14	84	21
100/400	Gabor	35	82	47
50/450	DoG	22	82	56
50/450	Gabor	37	77	53

(b)

Network	Pre-processing	Initial %	% Discarded	% After Discard
10+20	DoG	51	30	51
10+20	Gabor	57	38	52
6+12	DoG	54	32	53
6+12	Gabor	57	38	57

Table 2: Effect of pre-processing methods on **shift-varying** dataset (the original

Learnt Invariance Training with Shift and Scale Varying Images

These experiments again used a fixed selection of positions for training examples, using all five versions of each original image. This gives the network information about the shift and scale variance during training to help in learning this kind of invariance.

(a)

Pre-processing	Initial %	% Discarded	% After Discard
DoG	72	46	94
Gabor	85	35	98

(b)

Pre-processing	Initial %	% Discarded	% After Discard
DoG	84	32	93
Gabor	90	24	97

Table 4: Effect of pre-processing methods on **shift-varying** dataset (full groups of five used for training) (a) Standard 250/250 RBF Network (b) 30+60 Face Unit RBF Network

(a)

Pre-processing	Initial %	

Observations

Several points can seen from the results:

-

invariance to facial expression and refining an automated 'face-finder' routine. This is necessary for the next stage of development in which people are to be identified in natural image sequences with the usual variations in illumination as well as position, scale, view and facial expression. The statistical nature of the information successfully captured by RBF nets to do the classification task may also be effective for the face localisation task. It is clear from the work of Turk & Pentland (1991) and Bishop (1995) and others using statistically based techniques that this is the key to good performance and the RBF techniques are mathematically well-founded, which gives a clear advantage in engineering a solution to our application problems. Future work will tackle the full unconstrained recognition task by tracking faces in real-time and gathering enough information to classify them accurately with good generalisation to other image sequences containing familiar people.

References

Ahmad, S. & Tresp, V. (1993), Some solutions to the missing feature problem in vision, *in* S. J. Hanson, J. D. Cowan & C. L. Giles, eds, 'Advances in Neural Information Processing Systems', Vol. 5, Morgan Kaufmann, pp. 393–400.

Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, Oxford University Press.

Bruce, V. (1988), *Recognising Faces*, Lawrence Erlbaum Associates.

Bruce, V. & Young, A. (1986), 'Understanding face recognition', *British Journal of Psychology* **77**, 305–327.

Bruce, V., Burton, A. M. & Hancock, P. J. (1995), Missing dimensions of facial distinctiveness, *in* T. Valentine, ed., 'Cognitive and Computational Aspects of Face Recognition: Explorations in face space', Routledge, pp. 138–158.

Chen, S., Cowan, C. F. N. & Grant, P. M. (1991), 'Orthogonal least squares learning algorithm for radial basis function networks', *IEEE Transactions on Neural Networks* **2**, 302–309.

Daugman, J. G. (1988), 'Complete discrete 2-D gabor transforms by neural networks for image analysis and compression', *IEEE Transactions on Acoustics, Speech, and Signal Processing* **36**(7), 1169–1179.

Edelman, S., Reisfeld, D. & Yeshurun, Y. (1992), Learning to recognize faces from examples, *in* '2nd European Conference on Computer Vision', Genoa, Italy, pp. 787–791.

Girosi, F. (1992), 'Some extensions of radial basis functions and their applications in artifical intelligence', *Computers Math. Applic.* **24**(12), 61–80.

Hertz, J. A., Krogh, A. & Palmer, R. G. (1991), *Introduction to the Theory of Neural Computation*, Addison-Wesley.