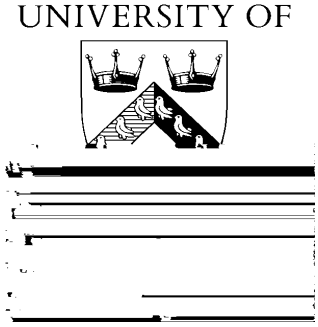
Receptive Field Functions for Face Recognition

A. Jonathan Howell and Hilary Buxton

CSRP 391

December 1995

ISSN 1350-3162



Cognitive Science Research Papers

Receptive Field Functions 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

December 1995

Abstract

Preprocessing of face images was performed 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 Radial Basis Function (RBF) networks that learnt to classify and generalise over different views for a standard face recognition task. Two main organisations of the RBF networks (standard and face unit) and two main types of preprocessing (Difference of Gaussian filtering and Gabor wavelet analysis) were compared. Quantitative and qualitative differences in these schemes are described and conclusions drawn about the best approach for our face recognition problem using low resolution images.

1 Introduction

Face recognition has been the subject of a great deal of research in computer vision and work on biologically-motivated approaches has begun to deliver real solutions. One of the main problems is dimensionality reduction to remove much of the redundant information in the original images. There are many possibilities for effectively representing this data, including principal component analysis, Gabor filters and various isodensity map or feature extraction schemes.

generalise over a wide range of conditions to capture the essential similarities of a given face. In this paper, we are concentrating on the issues of finding an effective input representation for our networks. In particular, we contrast the use of Difference of Gaussian filtering and Gabor wavelet analysis at a range of scales. One way of thinking about these input representations and mapping them onto our RBF networks is to use the analogy with visual neurons. The receptive field of such a neuron is the area of the visual field (image) where the stimulus can influence its response. For the different classes of these neurons, a receptive field function $f(x, y)$ can be defined. For example, retinal ganglion cellas and lateral geniculate cells early in the visual processing have receptive fields which can be implemented as Difference of Gaussian filters (Marr & Hildreth 1980). Later, the receptive fields of the simple cells in the primary visual cortex are oriented and have characteristic spatial frequencies. Daugman (1988) proposed that these could be modelled as complex 2-D Gabor filters. Petkov et al. (1993) successfully implemented a face recognition scheme based on Gabor wavelet input representations to imitate the human vision system. The question we want to ask here is whether these later stages of processing make more information explicit than the earlier DoG filters for our face recognition task.

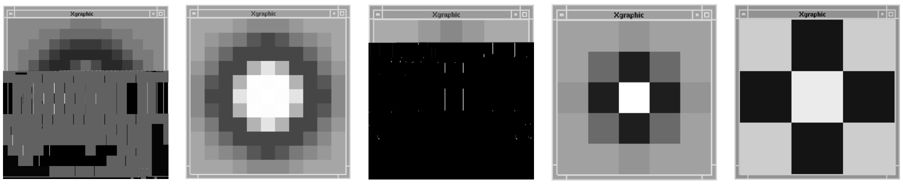
2 The RBF Network Model

The RBF network is a

Grey-Levels	Fixed %	Fixed % with 1.5 Discard	Epochs
Full	88	95	1148381
Reduced	92	100	335524

(a)

Grey-Levels	Ave. %	Ave. % with 1.5 Discard	Ave. Epochs
Full	94		



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

Figure 1: Masks created from various DoG scales (with mask sizes): (a) 1.6 (15×15) (b) 1.2 (11×11) (c) 0.8 (7×7) (d) 0.4 (5×5) (e) 0.15 (3×3)

In summary, good generalisation performance was obtained, although the training times were unacceptably long. The reduction of the range of grey-levels gave very much shorter convergence times, though still very slow.

5 Difference of Gaussians (DoG) Pre-Processing

Where there is a change of intensity in an image, peaks or troughs are found in the first derivative of the intensity, and zero-crossings in the second derivative. To isolate the latter, Marr & Hildreth (1980) suggested the $\nabla^2 G$, or *Laplacian of the Gaussian*, operator, which can be closely apprimated by a *Difference of Gaussians* (DoG) operator, constructed from two Gaussians G of the form:

$$G(x, y) = \frac{1}{\sigma^2} \exp(-\frac{x^2 + y^2}{2\sigma^2}), \quad (1)$$

where the space constants σ have a ratio of 1:1.6. The DoG masks were constructed using the POPVISION CONVOLVE_DOG_2D routines. Figure 5 shows these masks at various scale values, whilst Figure 2 shows the result of their convolution with an image at a fixed resolution.

5.1 DoG Gradients *vs.* 'Zero Crossings'

With a typical, grey level image, such as Figure 3(a), DoG convolution will give continuously-valued[2] gradient information, as shown in Figure 3(b). Where these values change from one sign to the other is the 'zero-crossing' point; if the values are thresholded at 0 into either 0 (for negative) and 1 (for positive), the boundaries between black and white are the zero-crossings for the image,

wforconstan as shown in Figure 3(c). Thoei20Td1(scale)Tj2(as)TJ20.16020Td(obtained,)T020Tdigurelicit990Td(2000.

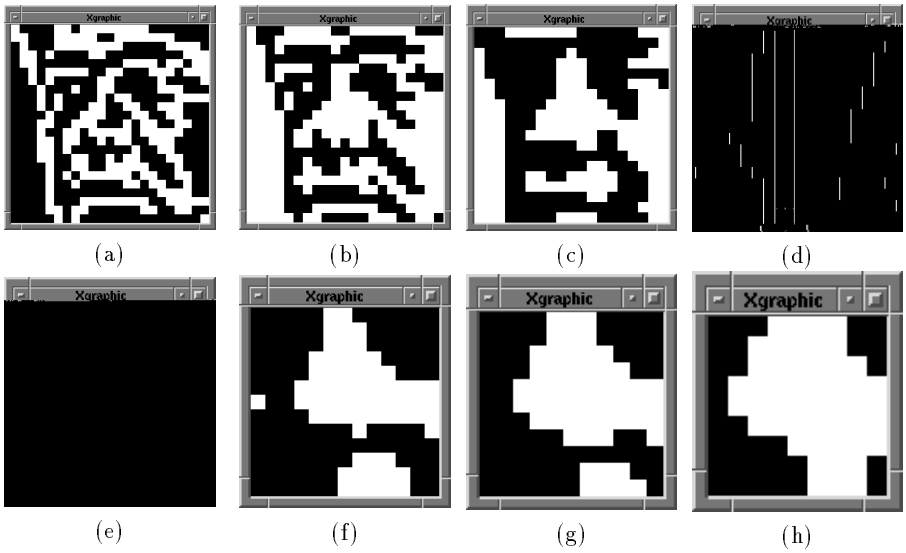


Figure 2: DoG scales applied to 25×25 image (with convolved image sizes): (a) 0.15 (23×23) (b) 0.4 (21×21) (c) 0.8 (19×19) (d) 1.0 (17×17) (e) 1.2 (15x15) (f) 1.4 (13x13) (g) 1.6 (11x11) (h) 1.9 (9x9)

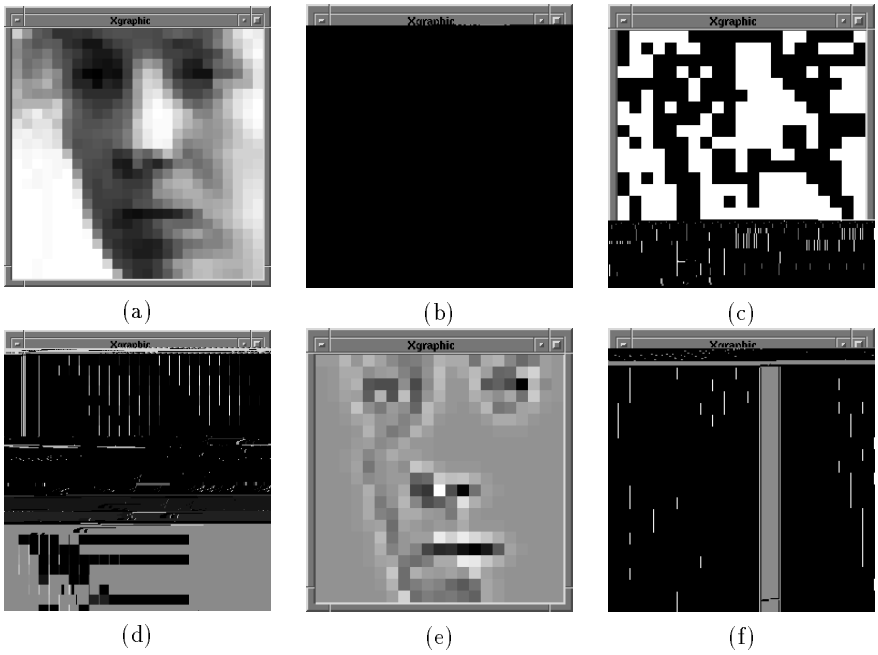
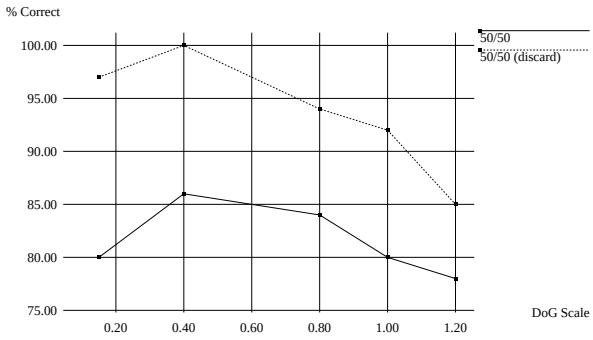


Figure 3: Effect of reducing range of grey-levels on 25×25 image (a) full range of grey-levels (b) after non-thresholded DoG (c) after thresholded DoG (d) reduced range of grey-levels (e) after non-thresholded DoG (f) after thresholded DoG

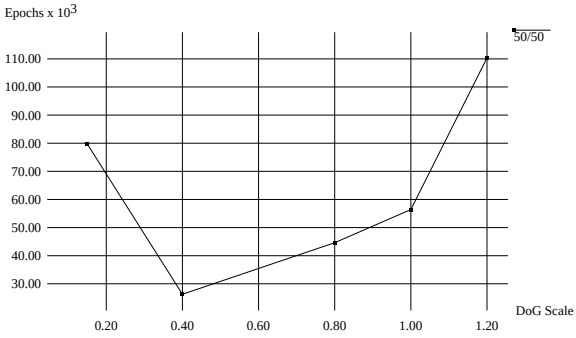
Scale	Thres-holding	Grey-Levels	Fixed %	Fixed % with 1.5 Discard	Epochs
0.4	No	Full	50	59	54515
0.4	No	Reduced	68	86	35801
0.4	Yes	Full	72	90	13311
0.4	Yes	Reduced	86	100	27463
0.15, 0.4, 0.8, 1.6	Yes	Reduced	78	90	19119

(a)

Scale	Thres-holding	Grey-Levels	Ave. %	Ave. % with 1.5 Discard	Ave. Epochs

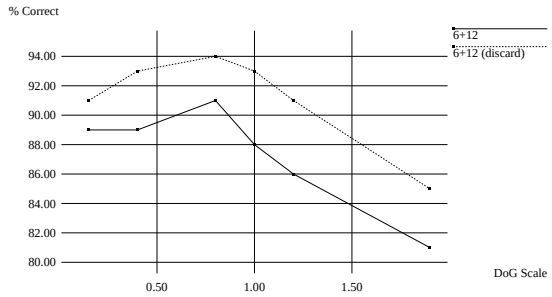


(a)

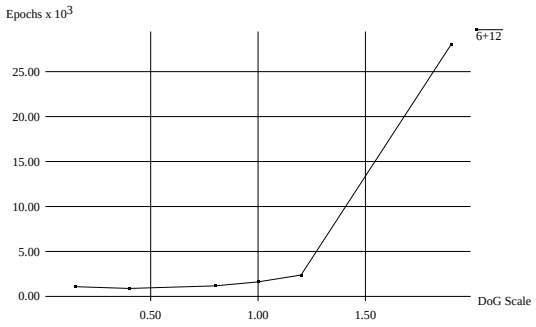


(b)

Figure 4: Effect of varying the scale in DoG pre-processing (a) on test generalisation (b) on training epochs with 50/50 RBF networks

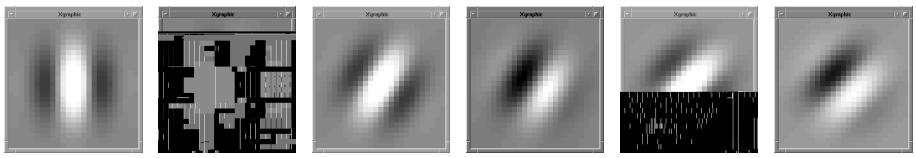


(a)



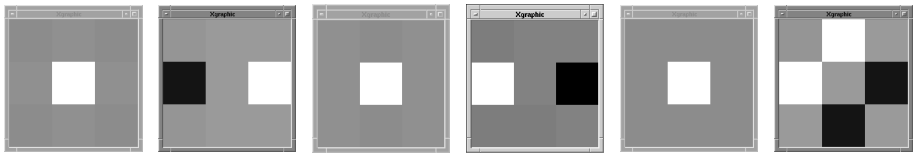
(b)

Figure 5: Effect of varying the scale in DOG pre-processing (a) on test generalisation (b) on training epochs with 6+12 RBF 'face unit' networks



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

Figure 6: 25×25 masks created from Gabor filter of period 13: (a) 0° real (b) 0° imaginary (c) 30° real (d) 30° imaginary (e) 45° real (f) 45° imaginary



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

Figure 7: 25

Scheme	Orientations (degrees)	Scales	Overlapping	Matrix	Coefficients Per Image
A1	0	4	No	Square	170
A2	0, 180	4	No	Square	340
A3	0, 120, 240	4	No	Square	510
A3X	60, 180, 300	4	No	Square	510
A3S	30, 150, 270	4	No	Square	510
A4	0, 90 180, 270	4	No	Square	680
A6	30, 90, 150 210, 270, 330	4	No	Square	1020
B3	0, 120, 240	4	Most	Square	510
C3	0, 120, 240	4	Less	Square	510
D3	0, 120, 240	3	No	Circular	420

Table 3: Types of Gabor sampling schemes tested, with filter orientations and number of coefficients sampled per image

6.2 Gabor Sam ling Schemes

In order to reduce the number of coefficients calculated for each image, sparse sampling schemes were constructed, with a range of scales. The 'A' square matrix sampling scheme which had the least amount of overlap on sampling points, proved to be the most successful arrangement. Others were tested which used large amounts of overlap on the sampling receptive fields, or circular sets of sampling points; Table 3 summarises the different sampling schemes used. Tables 4(a) and (b) show the sampling arrangements for the 'A' and 'B' square matrix sampling schemes, with Figures 8(a) and (b) showing how these masks were positioned to cover the image area. Note that the 'A' scheme only covers 24×24 at the 8×8 scale and the some overlap was needed to fit the 2×2 and 4×4 scales.

The 'C' square matrix sampling scheme (Table 4(c) and Figure 8(c)) was devised after the 'B' scheme performed poorly. The scales used were intended to retain fine detail from the original image.

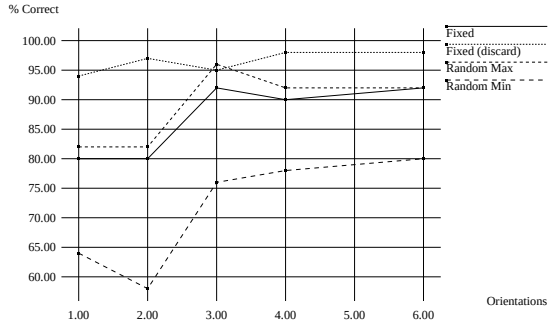
Table 4(d) and Figure 8(d) show similar details for the 'D' circular matrix sampling scheme. Note that due to the fairly coarse alignment to pixel boundaries in the low resolution 25×25 image area, some masks placements do not coincide with the exact mathematical position.

Scheme	Coefficients per Image	Fixed %	Fixed % with 1.5 Discard	Epochs
A3	510	92	95	35752
A3R	510	92	100	40016
Non-Thresholded A3	510	82	90	330035
A3 (Sine mask only)	255	86	95	30806
A3 (Cosine mask only)	255	46	72	5555197
B3	510	86	97	28832
C3	510	84	92	24122
D3	420	82	94	40532

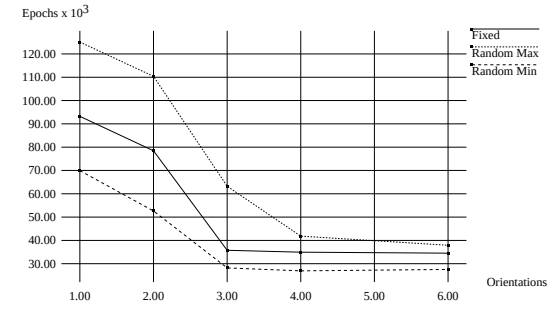
Table 5:

Scheme	Coefficients per Image	Ave. %	Ave. % with 1.5 Discard	Ave. Epochs
A3	510	96	98	654
A3R	510	95	98	755
Non-Thresholded A3	510	91	95	8288
A3 (Sine mask only)	255	94	99	368
A3 (Cosine mask only)	255	83	88	18980
B3	510	88	100	698
C3	510	92	96	556
D3	420	89	93	881

Table 6: Gabor Preprocessing for 6+12 Face Unit RBF Network (all schemes used full range of grey-levels except A3R)

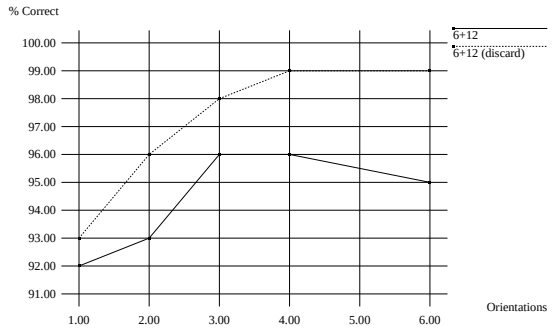


(a)

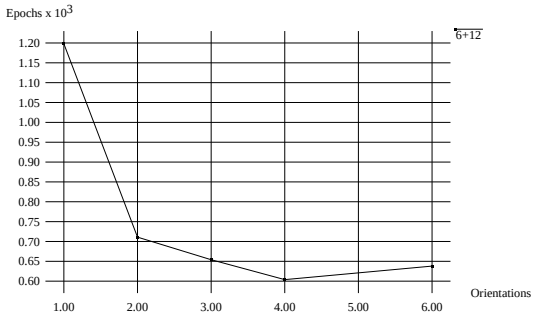


(b)

Figure 9: Effect of varying the number of orientations in Gabor pre-processing (a) on test generalisation (b) on training epochs for 50/50 RBF networks

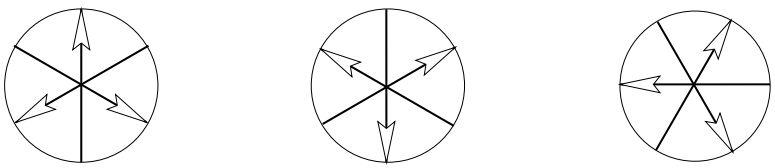


(a)



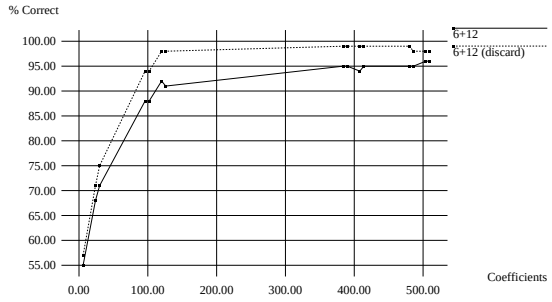
(b)

Figure 10: Effect of varying the number of orientations in Gabor pre-processing (a) on test generalisation (b) on training epochs for 6+12 RBF 'face unit' networks

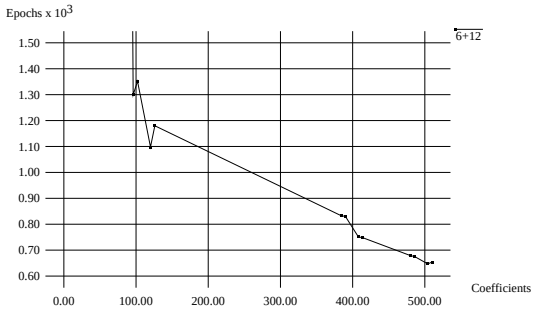


(a) (b) (c)

Figure 11: Angles used for three-orientation tests: (a) A3 (b) A3X (c) A3S

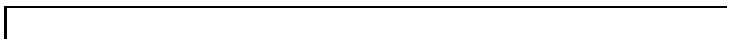


(a)



(b)

Figure 12: Effect of specific arrangements of scales in Gabor pre-processing (a) on test generalisation (b) on training epochs for 6+12 RBF 'face unit' networks, with total number of coefficients for the combination of scales



different scales and orientations to be closely tailored to the task at hand. The DoG preprocessing, on the other hand, results in an image-like representation with as many coeeficients as there are pixels. In our future work, we will extend the face unit RBF scheme and look at the problem of tracking faces in image sequences. We will also extend the Gabor preprocessing scheme to the space-time case.

Acknowledgments

We would like to thank David Young at COGS at the University of Sussex for his valuable help in providing POPLOG library functions for DoG and Gabor mask convolution.

References

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.

Howell, A. J. & Buxton, H. (1995*a*), 'Invariance in radial basis function neural networks in human face classification', *Neural Processing Letters* **2**(3), 26–30.

Howell, A. J. & Buxton, H. (1995*b*), A scaleable approach to face identification, *in* 'Proceedings of International Conference on Artificial Neural Networks (ICANN'95)', Vol. 2, EC2 & Cie, Paris, pp. 257–262.

Marr, D. & Hildreth, E. (1980), 'Theory of edge detection', *Proc. R. Soc. London* **B207**, 187–217.

Moody, J. & Darken, C. (1988), Learning with localized receptive fields, *in* D. Touretzky, G. Hinton & T. Sejnowski, eds, 'Proceedings of the 1988 Connectionist Models Summer School', Morgan Kaufmann, pp. 133–143.

Pentland, A., Moghaddam, B. & Starner, T. (1994), View-based and modular eigenspaces for face recognition, *in* 'IEEE Conference on Computer Vision and Pattern Recognition', pp. 84–91.

Petkov, N., Kruizinga, P. & Lourens, T. (1993), Biologically motivated approach to face recognition, *in* 'Proceeding of International Workshop on Artificial Neural Networks', pp. 68–77.

Rao, R. P. N. & Ballard, D. H. (1995), Natural basis functions and topographic memory for face recognition, *in* 'Proceeding of International Joint Conference on Articial Intelligence (IJCAI'95)', Montréal, Canada, pp. 10–17.

Turk, M. & Pentland, A. (1991), 'Eigenfaces for recognition', *Journal of Cognitive Neuroscience* **3**(1), 71–86.