A RADIAL BASIS FUNCTION NEURAL NETWORK FOR PARTS IDENTIFICATION OF THREE DIMENSIONAL SHAPES

BORGES, D.L.; ORR, M.J.; FISHER, R.B.

DAI Research Paper No. 728
Dec 1994

Accepted for presentation at the VII Brazilian Symposium of Computer Graphics and Image Processing, SIBGRAPI, Curitiba, BrazNovember. 1994

A Radial Basis Function Neural Network for Parts Identification of Three Dimensional Shapes

DIBIO L. BORGES[1]
MARK J. ORR[2]
ROBERT B. FISHER[1]

[1]Department of Artificial Intelligence, Edinburgh University
5 Forrest Hill, Edinburgh EH1 2QL, Scotland, UK
dibio, rbf @aifh.ed.ac.uk
[2]Centre for Cognitive Science, Edinburgh University
2 Buccleuch Place, Edinburgh EH8 9LW, Scotland, UK
mjo@cns.ed.ac.uk

Abstract. The discrimination of volumetric pieces or parts of objects from range data is one key element for achieving 3-D object recognition. In this paper it is shown that previously segmented and acquired superquadrics from range data can be reliably mapped into a set of qualitative volumetric shapes (geons) by means of an RBF (Radial Basis Function) neural network classifier. We use a regularised RBF classifier and the results are shown to be both reliable and efficient in the context of range image understanding.

1 Introduction

Recognition of three dimensional articulated objects with complex curved surfaces is an important, and yet unsolved problem in Computer Vision. Volumetric modelling and recovering of articulated objects from range data is a solution to this problem, but the volumetric representations used have to offer both quantitative and qualitative information in order for the matching be flexible and efficient.

Deformable superquadrics offer a solution for representing and acquiring models from range images of articulated objects [Borges and Fisher (1993)], however because of the continuum of the superquadric parameters they do not provide a straightforward way to index shape identities into large databases of objects which is necessary for achieving efficient recognition and localisation. Geons [Biederman (19-87)] however are a set of qualitative volumetric shapes which are suitable for the purpose of indexing because of their distinctive symbolic nature.

In the context of investigating how to recognise articulated objects, the use of both deformable superquadrics and geons will strengthen the expressiveness and the indexing power of such a recognition engine. In our lab we are investigating this approach and in this paper we report on using a radial basis function (RBF) neural network to map deformable superquadrics onto a set of qualitative volumetric shapes (a subset of the geons). The network is trained using regularisation and details are given on how to tune the regularisation parameter and the basis function widths using cross-validation.

The main contributions of this work are twofold: First, it presents a new algorithm for training RBF neural networks in the context of multi-dimensional spaces, second this classification method is applied to derive indexing capabilities from fitted superquadric surfaces in order to produce suitable data for a superquadrics' model-based recognition engine.

2 Superquadrics and Geons

Volumetric primitives capture well the notion of parts in natural [Marr and Nishihara (1978)] and man-made objects [Biederman (1987)]. They are also well suited for indexing since they have a relatively small number of parameters compared to other primitives. Basically five types of volumetric primitives have been used in Computer Vision, Generalised Cylinders [Marr and Nishihara (1978), Brooks (1981)], "Sticks, plates, and blobs" [Mulgaonkar et al. (1984)], Superquadrics [Pentland (1986)], second-order topographic features [Fisher (1992)], and Geons [Raja and Jain (1992), Dickinson et al. (1992), [Bergevin and Levine (1992)]. Superquadrics are a family of parametric shapes that can be globally deformed, and thus represent a wide range of prototypical parts.

A superquadric surface can be defined by the following vector in 3D space

$$x(\eta,\omega) = \begin{bmatrix} a_1 \cos^{\epsilon_1}(\eta)\cos^{\epsilon_2}(\omega) \\ a_2 \cos^{\epsilon_1}(\eta)\sin^{\epsilon_2}(\omega) \\ a_3 \sin^{\epsilon_1}(\eta) \end{bmatrix} \quad (1)$$

where,

$$-\pi/2 \leq \eta \leq \pi/2.$$
$$-\pi \leq \omega < \pi.$$

The parameters a_1, a_2, and a_3 define the size of the superquadrics in the coordinates x, y, and z respectively. ϵ_1 and ϵ_2 are the squareness parameters. The capabilities of the superquadrics for modeling can be enhanced by introducing tapering and bending deformations.

For the purpose of recovering a superquadric surface from the data, equation (1) can be manipulated in order to get the following implicit equation

$$\left(\left(\frac{x}{a_1}\right)^{2/\epsilon_2} + \left(\frac{y}{a_2}\right)^{2/\epsilon_2}\right)^{\epsilon_2/\epsilon_1} + \left(\frac{z}{a_3}\right)^{2/\epsilon_1} = 1. \quad (2)$$

Based on this equation we can define the following inside-outside function [Solina and Bajcsy (1990)]

$$\mathrm{F}(x, y, z) = \left(\left(\left(\frac{x}{a_1}\right)^{2/\epsilon_2} + \left(\frac{y}{a_2}\right)^{2/\epsilon_2}\right)^{\epsilon_2/\epsilon_1} + \left(\frac{z}{a_3}\right)^{2/\epsilon_1}\right)^{\epsilon_1}. \quad (3)$$

Following Solina [Solina and Bajcsy (1990)] we estimate the 15 parameters (6 for position and 9 superquadric parameters, being the 5 defined above plus 4 for tapering(T_x, T_y) and bending (β_x, β_y) deformations) using the Levenberg-Marquadt method for nonlinear least squares minimisation of the expression

$$\sum_{i=1}^{N} [\mathrm{R}(x_i, y_i, z_i; a_1, ..., a_{15})]^2. \quad (4)$$

where,

$$\mathrm{R} \;=\; \sqrt{a_1 a_2 a_3}(\mathrm{F} - 1). \quad (5)$$

and N is the number of observed points. In our case, we use all of the range data points in each surface patch found by the isolating discontinuity boundaries found from segmentation [Borges and Fisher (1993)].

Using only superquadric parameters for 3-D object recognition would not be very efficient because they lack expressiveness for indexing into the object database. This problem could be solved by mapping the superquadrics parameters into a set of distinctive volumetric shapes with good potential for indexing. In the end, both qualitative and quantitative shape information would be readily available.

One set of distinctive volumetric shapes are the geons, which were proposed as part of the Recognition-by-Components theory of Biederman [Biederman (1987)]. In his theory Biederman proposes that a set of 36 fundamental part primitives (geons) can represent many complex 3D objects for the purpose of primal access.

With the set of deformations we use in the superquadrics parameters (tapering and bending) we can model 12 geon classes, the other 24 require distortions in the cross-sections to be introduced. However, for our purposes distortions in the cross-sections are not actually needed (the input of the classifier is of segmented data) and what is important is that we should have enough indexing capability (qualitative shape information) to improve our representation. Figure 3 shows the 12 geon classes used as the target mapping set.

This mapping from superquadrics to geons does not depend on all the 15 recovered superquadrics parameters. Actually in order to obtain one of the 12 geons shown on Figure 3 we need the following 5 superquadrics parameters:

- ϵ_1.
- ϵ_2.
- T_x, tapering in x.
- T_y, tapering in y.
- radius-of-bending/a_3 (derived from the bending and a_3 parameters).

The classification task is to map these 5 superquadric parameters into 12 distinct classes of geons. Reconstructing such complicated mapping functions, as efficiently as possible (*i.e.* achieving high accuracy using small amount of time and training data), is an active area of research in neural networks. The next section introduces a radial basis function network which achieves higher classification rates than other previously published algorithms for the problem at hand, making it possible in our case to deduce more accurate qualitative information for visual matching of three dimensional shapes.

3 RBF Neural Networks

Radial basis functions started as function approximation tools [Powell (1987), Jackson (1989)] and later found favour as neural network learning algorithms [Broomhead and Lowe (1988), Girosi et al. (1993)].

The basic problem is to approximate a function given a finite set of example input-output pairs (the training set). If the function is from $\Re^p$ to $\Re^q$ then the training set inputs can be written

$$\{x_k \in \Re^p; k = 1, 2, ..., n\} ,$$

and corresponding outputs

$$\{y_k \in \Re^q; k = 1, 2, ..., n\} .$$

In a straightforward regression problem the normal procedure is to assume that the data is explained by

$$y = f(x; W) + \epsilon ,$$

where $f(\cdot; \cdot)$ is a function which depends on some unknown parameters, W, and ϵ is zero mean additive noise. The unknown parameters can be estimated by some optimisation criterion such as least squares. For classification problems the same technique often works in practice. In that case the components of the training outputs $\{y_k\}_1^n$ are all zero except the component which corresponds to the correct output class, which is unity. Then the regression function $f(x; W)$ estimates the probabilities of each of the q classes given the input (features) x.

RBF networks assume that the components of the vector function $f(\cdot, \cdot)$ are given by

$$f_i(x; W) = \sum_{j=1}^{m} W_{ji}\, h(x, c_j) , \quad 1 \leq i \leq q ,$$

a sum of functions which depend on m fixed positions (centres), $c_j \in \Re^p$, in the input space and an m-by-q weight matrix whose rows are points in the output space. Such a function can be viewed as a network (see Figure 2). Certain conditions must be met by the function $h(\cdot, \cdot)$ such as being radial and monotonic for positive values. We use the Gaussian function,

$$h(x, c) = \exp\left(-\frac{\| x - c \|^2}{r^2}\right) ,$$

where the scaling parameter r is, like the $\{c_j\}_1^m$, chosen in advance. As the function is linear in the unknown parameters (W) the least squares criterion leads to a system of linear simultaneous equations to solve, a less daunting prospect than the nonlinear optimisation problems posed by some other types of neural networks.

A common problem with all regression and classification methods is the trade off between bias and variance [Geman et al. (1992)]. If the model is too flexible it will overfit (too much variance) and if it is not flexible enough it will underfit (too much bias). In RBFs there are basically three frameworks to deal with this:

1. random centre selection [Broomhead and Lowe (1988)].
2. forward centre selection [Chen et al. (1991)], and
3. regularisation [Bishop (1991)].

In this paper the regularisation framework, which stems from the work of Tikhonov and others [Tikhonov and Arsenin (1977)], is used.

A common choice for the fixed centres are the n input points in the training data, $\{x_k\}_1^n$ so that $m = n$ and $c_j = x_j$. Then the normal solution to the least squares minimisation is

$$W = (H^T H)^{-1} H^T Y ,$$

where

$$H_{kj} = h(x_k, x_j) , \quad 1 \leq k, j \leq n ,$$

and the rows of the n-by-q matrix Y are the training set outputs. The problem with this choice of centres is that the number of parameters to estimate is the same as the number of data available and any redundancy in the data will lead to overfit. Often also the matrix $H^T H$ will be ill-conditioned. The simplest regularisation technique to avoid such overfit (and also ill-conditioning) is to add a weight penalty term to the least squares criterion so that the function to be minimised (with respect to W) is

$$\sum_{k=1}^{n} \| y_k - f(x_k, W) \|^2 + \lambda \sum_{j=1}^{m} \sum_{i=1}^{q} W_{ji}^2 ,$$

(instead of just the first term). This is called zero-order regularisation [Press et al (1992)] and λ is a positive constant called the regularisation parameter. The solution to this minimisation problem is

$$W = (H^T H + \lambda I)^{-1} H^T Y .$$

The idea is that a small $\lambda > 0$ will introduce a small bias but greatly reduce the variance.

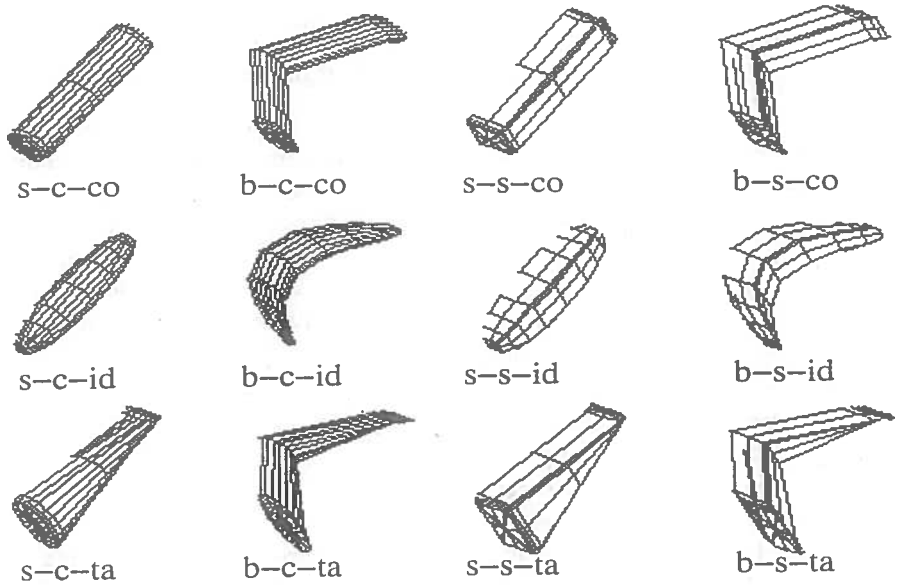


Figure 1: The set of twelve(12) geons modelled using the parameters from the deformable superquadrics.

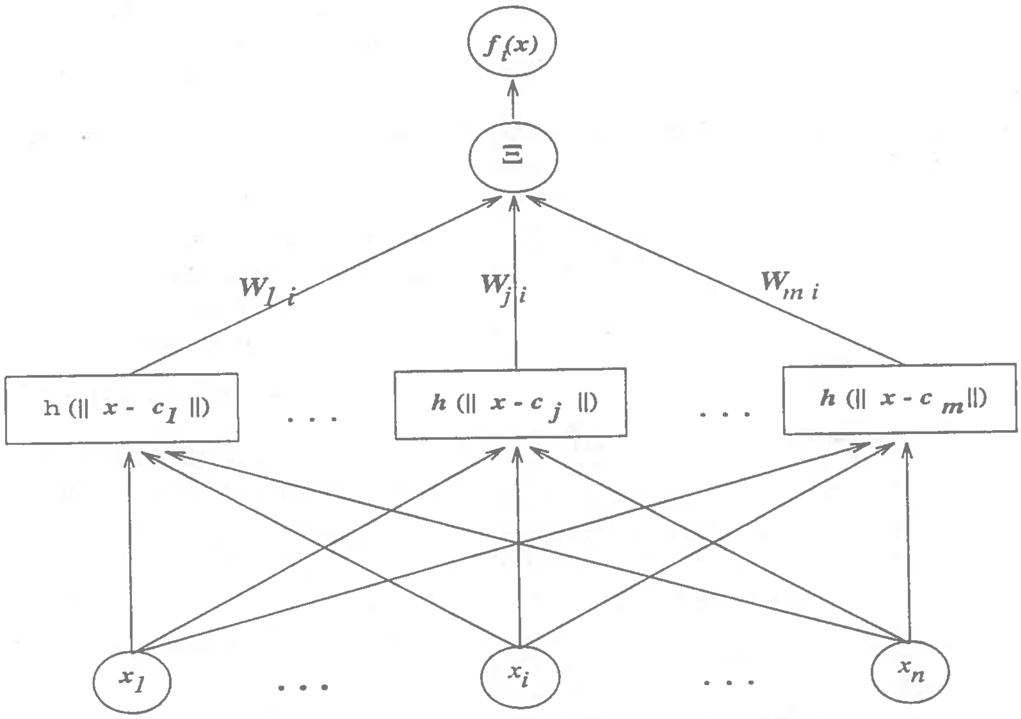


Figure 2: Radial Basis Function Neural Network.

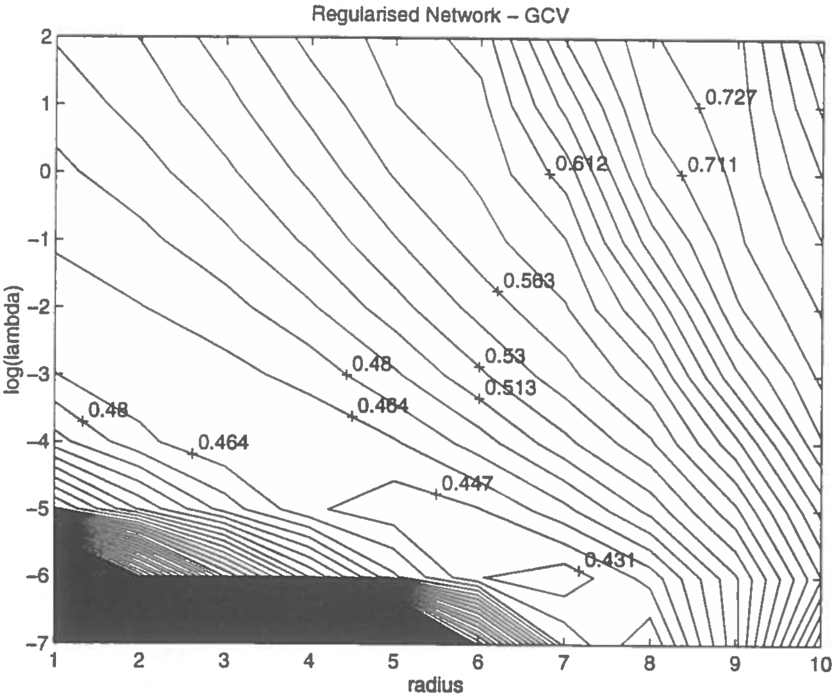
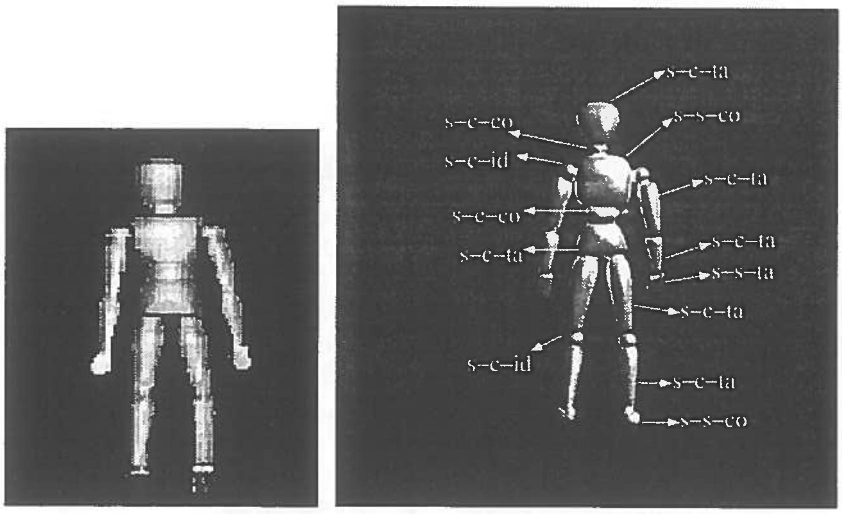


Figure 3: Regularised R.B.F. Neural Network using Cross-Validation.



(a) (b)

Figure 4: (a) Pseudo-intensity image (from range); (b) Rayshaded image of recovered superquadrics from range data and labels with RBF classification of geons. Left and right similar parts (e.g. left upper-arm, right upper-arm) are labelled the same and the figure just shows the label on one of them.

We now have two free parameters in the RBF, the Gaussian radius r and the regularisation parameter λ. The optimal values to use for these parameters depend on the amount and characteristics of the training data available for the classifier. However, there exist criteria for ranking different classifiers on the same data and one of these, generalised cross-validation (GCV) [Golub et al. (1979)], we employ here. The GCV score is given by

$$\mathrm{GCV} = \frac{\mathrm{trace}(Y^T P^2 Y)}{[(1/n)\,\mathrm{trace}(P)]^2} \ ,$$

where

$$P = I - H\,(H^T H + \lambda\, I)^{-1} H^T \ ,$$

and lower GCV score is associated with better generalisation.

Figure 3 shows a contour plot of GCV (in arbitrary units) for different choices of r and λ for the superquadrics to geon classification data (see next section). There is a clear minimum near the point $r = 7$, $\lambda = 10^{-6}$, and the classification results we describe below are based on these values.

4 Experiments

Data used for training the classifier was acquired by a segmentation procedure for range images using the algorithm of [Borges and Fisher (1993)], where wooden and plastic articulated objects were scanned by a laser striper, and then submitted to the segmentation. The segmentation works first by detecting the discontinuities on the surface, then followed by dynamic grouping of these points, and fitting of the closed regions by deformable superquadrics.

In order to achieve the results shown in Tables 1 and 2, $n = 369$ input-output training examples were used, including sometimes the same object to be segmented but in different positions. We included approximately the same number of examples from each of the 12 classes.

Table 1 gives the mean and standard deviations of all the 5 parameters used in the experiments for all the 12 geon classes. Table 2 shows the success rates for the classifier using the hold-out method (which is known to give a pessimistic estimate) indicating the individual and the overall rates for all the classes. The method involves training the classifier on some percentage of the data (we used 90%) and testing it on the rest which is held in reserve. The process is repeated a number of times to get an average success rate.

Figure 4 shows a ray-shaded image of fitted superquadrics on an original range image of a wooden manikin. Labels with the highest ranked geon class given by the RBF classifier are attached to the parts.

5 Conclusions

[Raja and Jain (1992)] presented experiments on classifying superquadrics parameters into 12 geon classes similarly to our classification problem here. They have used range images of hand-carved models of geons and ran experiments with a tree classifier and k-nearest-neighbour classifiers. They presented as the best figure achieved by a tree classifier on range data from "smooth" objects an error rate of 23.3% (i.e. 76.7% correct).

From Table 2 the success rate achieved for the RBF classifier presented here (79.0%) is better than Raja and Jain's figure. If we count the number of times the correct classification is included in the top three probabilities of the classifier then the success rate hits 94.4%. The including of more than just the highest probability value is particularly interesting for our original problem of using the classification as an additional measure for the matching of three dimensional parts, because it makes it possible to include additional hypotheses with different rankings in the recognition of the objects.

Our RBF network is being used as part of a complete system to segment and recognise (with localisation) complex articulated objects such as the recovered object shown in Figure 4. The training of the RBF net is done off-line and the final weights are then used for classification of new data in the recognition program. The algorithm could be generalised for other areas where multi-dimensional classification is necessary, and further investigation into other areas of application is being pursued.

Acknowledgements

Dibio L. Borges is supported by National Council for Research and Scientific Development (CNPq), Ministry of Science and Technology, Brazil, grant number 200389/88.2. Mark Orr is supported by a grant from the UK Joint Councils Initiative in Human Computer Interaction and Cognitive Science. This paper benefited greatly from discussions with A. Fitzgibbon.

6 References

R. Bergevin and M.D. Levine, Part decomposition of objects from single view drawings, *CVGIP: Image Understanding* **55** (1992) 73–83.

I. Biederman, Recognition-by-components: A theory of human image understanding, *Psychological Review* **94** (1987) 115–147.

Geon type	ϵ_1		ϵ_2		t_x		t_y		*rad_bend*	
	mean	st.dv.	mean	st.dv.	mean	st.dv.	mean	st.dv.	mean	st.dv.
b-c-co	0.135	0.110	0.694	0.227	0.180	0.129	0.190	0.131	3.588	1.833
b-c-id	0.771	0.192	0.776	0.174	0.388	0.305	0.310	0.251	3.607	1.606
b-c-ta	0.157	0.109	0.693	0.245	0.690	0.226	0.512	0.299	3.846	1.677
b-s-co	0.161	0.094	0.201	0.092	0.176	0.123	0.128	0.121	4.133	1.714
b-s-id	0.766	0.242	0.218	0.243	0.397	0.309	0.337	0.327	4.275	1.623
b-s-ta	0.083	0.097	0.153	0.088	0.658	0.309	0.605	0.297	4.270	1.378
s-c-co	0.139	0.107	0.722	0.210	0.140	0.122	0.158	0.160	16.037	15.714
s-c-id	0.752	0.195	0.758	0.172	0.412	0.307	0.346	0.310	15.406	9.648
s-c-ta	0.153	0.144	0.682	0.179	0.613	0.223	0.614	0.250	16.198	14.616
s-s-co	0.132	0.108	0.131	0.105	0.132	0.100	0.119	0.102	11.701	8.361
s-s-id	0.756	0.200	0.150	0.106	0.302	0.312	0.306	0.308	18.668	20.332
s-s-ta	0.089	0.090	0.160	0.083	0.627	0.304	0.631	0.326	15.983	11.877

Table 1: Mean and Standard Deviation Values for the Classification Parameters of each Class of Geon

Geon type	Hold-out success %, 40 averages, including only the highest classification	Hold-out success %, 40 averages, including the three highest classifications
b-c-co	75.0	92.4
b-c-id	71.6	96.6
b-c-ta	92.8	100.0
b-s-co	64.0	87.0
b-s-id	74.8	92.7
b-s-ta	74.3	94.5
s-c-co	88.4	97.3
s-c-id	88.6	98.9
s-c-ta	81.0	94.8
s-s-co	81.5	92.3
s-s-id	85.3	96.1
s-s-ta	78.1	94.8
Overall (all classes)	79.0	94.4

Table 2: Rates of Classification for the RBF Neural Net Classifier

C. Bishop. Improving the Generalisation Properties of Radial Basis Function Neural Networks. *Neural Computation* 3 (1991) 579–588.

D.L. Borges and R.B. Fisher, Segmentation of 3D Articulated Objects by Dynamic Grouping of Discontinuities. *Proceedings of the British Machine Vision Conference*. Guildford, (1993) 279–288.

R. Brooks. Symbolic reasoning among 3-D models and 2-D images, *Artificial Intelligence* **17** (1981) 285–348.

D.S. Broomhead and D. Lowe, Multivariate functional interpolation and adaptive networks, *Complex Systems* 2 (1988) 321–355.

S. Chen, C.F.N. Cowan, and P.M. Grant, Orthogonal Least Squares Learning for Radial Basis Function Networks, *IEEE Trans. Neural Networks* **2** (1991) 302–309.

S.J. Dickinson, A. Pentland, and A. Rosenfeld, From volumes to views: An approach to 3-D object recognition, *CVGIP: Image Understanding* **55** (1992) 130–154.

R.B. Fisher. Representation, extraction, and recognition with second-order topographic surface features, *Image and Vision Computing* **10** (1992) 156–169.

S. Geman, E. Bienenstock and R. Doursat, Neural Networks and the Bias/Variance Dilemma, *Neural Computation*, 4(1) (1992) 1–58.

F. Girosi, M. Jones, and T. Poggio, Priors, Stabilizers and Basis Functions: from Regularization to Radial, Tensor and Additive Splines, *Artificial Intelligence Laboratory, MIT, AI Memo*, **1430** (1993).

G.H. Golub, M. Heath and G. Wahba, Generalised Cross-Validation as a Method for Choosing a Good Ridge Parameter, *Technometrics*, **21** (1979) 215–223.

J.R.H. Jackson, Radial basis functions: a survey and new results, In D.C. Handscomb, editor, *Mathematics of Surfaces III*, pages 115–133. Clarendon Press, Oxford, 1989.

D. Marr and H.K. Nishihara, Representation and recognition of the spatial organisation of three-dimensional shapes, *Proc. R. Soc. London B* **200** (1978) 269–294.

P.G. Mulgaonkar, L.G. Shapiro and R.M. Haralick, Matching 'sticks, plates, and blobs' objects using geometric and relational constraints, *Image and Vision Computing* **2** (1984) 85–98.

A. Pentland, Perceptual organisation and the representation of natural form, *Artificial Intelligence* **28** (1986) 293–331.

M.J.D. Powell, Radial basis functions for multivariate approximation: a review, In J.C. Mason and M.G. Cox. editors, *Proc. of IMA Conf. on Algorithms for Approximation of Functions and Data*, pages 143–167. Oxford University Press, 1987.

W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, Numerical Recipes in C (2nd edition), *Cambridge University Press, UK*, (1992).

N.S. Raja and A.K. Jain, Recognizing geons from superquadrics fitted to range data, *Image and Vision Computing* **10** (1992) 179–190.

F. Solina and R. Bajcsy, Recovery of parametric models from range images: The case for superquadrics with global deformations, *IEEE Trans. Pattern Analysis and Machine Intelligence* **12** (1990) 133–147.

A.N. Tikhonov and V.Y. Arsenin, *Solutions of Ill-Posed Problems*, John Wiley, 1977.