Visualization of Radial Basis Function Networks

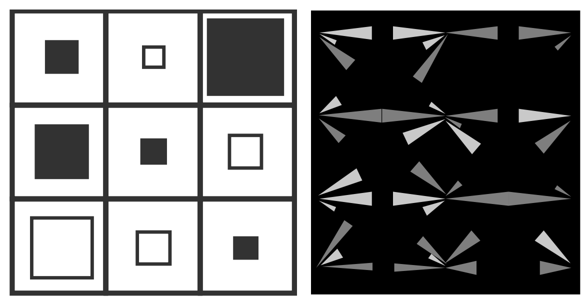
Adrian Agogino (agogino@ece.utexas.edu)
Joydeep Ghosh (ghosh@ece.utexas.edu)
Laboratory for Artificial Neural Systems
http://www.lans.ece.utexas.edu

Cheryl Martin (cheryl@cadlips.ece.utexas.edu)
Laboratory for Intelligent Processes and Systems
http://www.lips.utexas.edu

Department of Electrical and Computer Engineering
University of Texas at Austin
Austin TX 78712
USA

Abstract

This paper presents a method for the 3D visualization of the structure of radial basis function networks. This method allows the visualization of basis function characteristics (centers and widths) along with second level weights. Network properties can be displayed simultaneously with the training data or test data in the same input space. Principal component analysis is used to transform the input data so that its most salient dimensions can be visualized. This method also allows changes made while graphically editing the network structure, in transformed space, to be projected back into the original input space.



a. b.

Figure 1. Conventional Forms of Visualization. A. Hinton Diagram. B. Bond Diagram.

I. Introduction

Neural networks are traditionally used as a "black box." In general, the user has little knowledge about how a network is forming a solution to a particular problem. The impact of specific internal network parameters is often difficult to comprehend. Examining the matrix of numbers representing a trained network's weights generally provides little insight, and current visualization methods to help with this task are still in the primitive stages.

In general, visualization techniques developed specifically for neural networks have focused on the representation and interpretation of network weights [7]. These techniques include Hinton diagrams [5] (Figure 1a), bond diagrams (Figure 1b), WV-diagrams [1], WV-curves, and weight trajectory diagrams [6]. However, network weights may not be the most important thing to visualize in neural networks. As summarized by [4], "Attempts to visualize the internal states of neural networks usually concentrate on visualization of the connection weights. The value of these weights however, has no direct significance that can be related to the input data or to the output classification. When using neural networks as statistical classifiers, it makes more sense to visualize the internal network state in terms of decision boundaries drawn in feature space." In this paper we will focus on visualizing the network structure in ways that are meaningful in terms of the inputs and the outputs.

One reason that so much previous work has focused on weight-space visualization is that the type of network generally visualized is the multi-layer perceptrion (MLP). The visualization of radial basis function (RBF) networks is practically unheard of in the literature and in simulation descriptions. This is unfortunate because the nature of RBF networks helps move the focus away from the weight space to the input-output mapping of the network. Building a visualization tool that concentrates on this mapping should allow a deeper understanding of RBF networks and the problems to which they are applied.

The visualization of RBF networks must draw from previous visualization work in general and from previous work on visualization of neural networks in particular. Due to the nature of RBF networks, the visualization method proposed by this paper is grounded in the network's input space and can therefore benefit from previous work in visualization of high-dimensional data sets. Visualization of the second layer weights and network outputs can leverage previous work in neural network visualization.

II. Visualization Method

The visualization tool presented in this paper attempts to visualize all the parameters of a Gaussian radial basis function network, including the center, width, and associated weights for each basis function. Since the RBF network topology is derived from the training data, the tool also gives the user the important ability to visualize the data sets. Due to the unique topology of RBF networks, the data points can be visualized in the same space as the basis functions, giving the user immediate visual feedback into the relationship between the network and the data. The interactive visualization environment allows the user to zoom, rotate, add or remove information from the display, and select viewing dimensions if more than three are available. In addition to visualizing the network topology and data, this visualization tool also lets the user interactively modify or create the network topology and view the performance of the resulting network.

A. Visualizing Basis Functions

This paper assumes that the basis functions being visualized are Gaussian in the form of

$$\phi_j(x) = \exp\left(-\frac{\left\|x-\mu_j\right\|^2}{2\sigma_j^2}\right) \tag{1}$$

To visualize the basis function, they are drawn as spheres with a radius of σ. Since their centers are in the same space as the data points they are be drawn intermixed in the data (Figure 2). The basis functions are visualized either as translucent spheres or wire-frame spheres so that the data they surround can be seen. This allows by inspection the user to see the relation between the data and the basis function. It also gives an idea of the sphere of influence of a given basis.

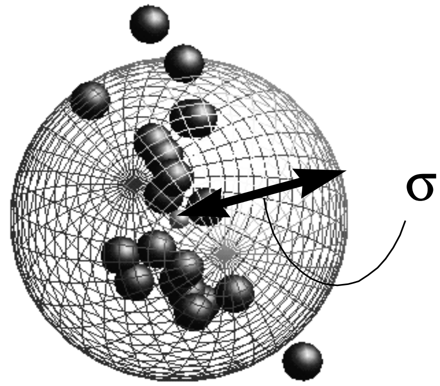


Figure 2. Each radial basis function is visualized as a sphere centered at the basis center. The radius is equal to the basis width. Note that the sphere is drawn in wire-frame mode so that the cluster of data that the basis surrounds can be seen. (Actual program uses color rendering. A demo is located at *www.lans.ece.utexas.edu/projects/rbfvis*)

B. Visualizing Weights

While the visualization of the radial basis functions gives important information about how the network is processing the clusters of data, the visualization of the second layer weights is needed to see how the network is finding a solution. To visualize the weights, we used a method similar to bond diagrams. In a classification problem the number of weights used for each radial basis function is equal to the number of classes. Each weight for a particular radial basis is visualized as a cone coming out of the basis center whose color corresponds to the class that the weight is going to (Figure 3). In regression problems, only one output is visualized at a time, and there is only one weight/cone per basis center. The strength of the weight is visualized by the size of the cone. Positive weights are represented as solid cones and negative weights are represented as banded cones. In terms of a bond diagram, this can be seen as bonds being drawn from the basis center to output nodes that are located at infinity.

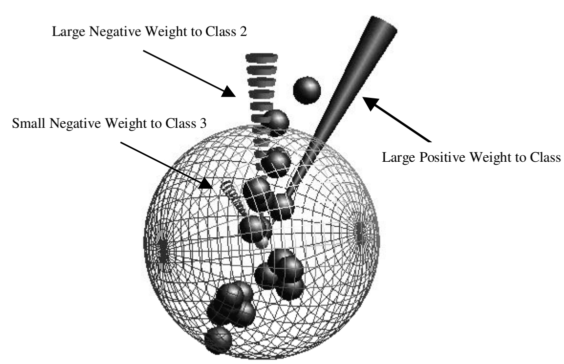


Figure 3. Visualization of weights for a single radial basis function is done by drawing them as cones coming out of the center of the basis. Each cone is a weight connected to a different class output. The magnitude of the weight is represented as the size of the cone. Negative weights are banded.

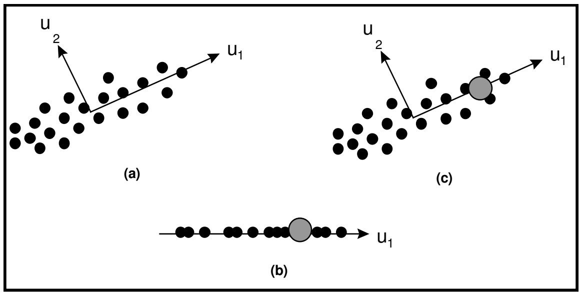


Figure 4. (a) Original data in two dimensions. (b) New data point added along first principal component. (c) New point is transformed back to original dimensions.

C. PCA

Since the dimensionality of input data will generally be much higher than three dimensions, dimensionality reduction is used so that the most important information in the data can be visualized in the three dimensional space. This paper uses principal component analysis (PCA) to perform the dimensionality reduction. This technique is used, because it is computationally efficient and it is an optimal form of linear dimensionality reduction (in the least squares sense) [2]. PCA works by forming a square matrix using the eigenvectors of the covariance matrix of the data. This minimizes the error in equation 2 when the *M* most important dimensions are used.

$$E_M = \frac{1}{2}\sum_{i=M+1}^{d}\sum_{n}\left\{\mu_i^T\left(x^n - \bar{x}\right)\right\}^2 \qquad (2)$$

Each data point and basis center is transform by multiplying it by the eigenvector matrix U.

$$p^{'} = pU \qquad (3)$$

The three dimensions with the highest associated eigen-values are then displayed. While using non-linear methods such as principal curves [3] could do better dimensionality reduction, the linear method is used since it has the important property of being invertable. This property allows us to transform any of the editing done in the transformed space to be converted back to the original space (Figure 4) by applying the inverse of the matrix U.

$$p = p'U^{-1} \qquad (4)$$

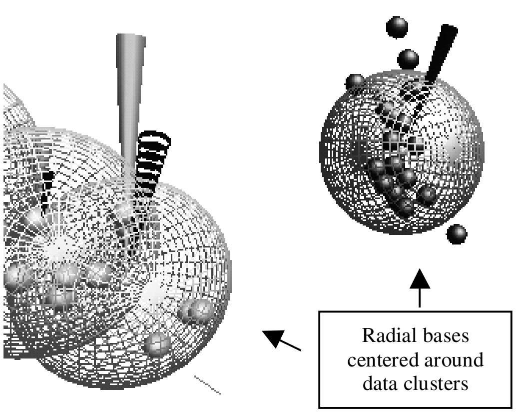


Figure 5: Visualization of RBFN with Iris data set

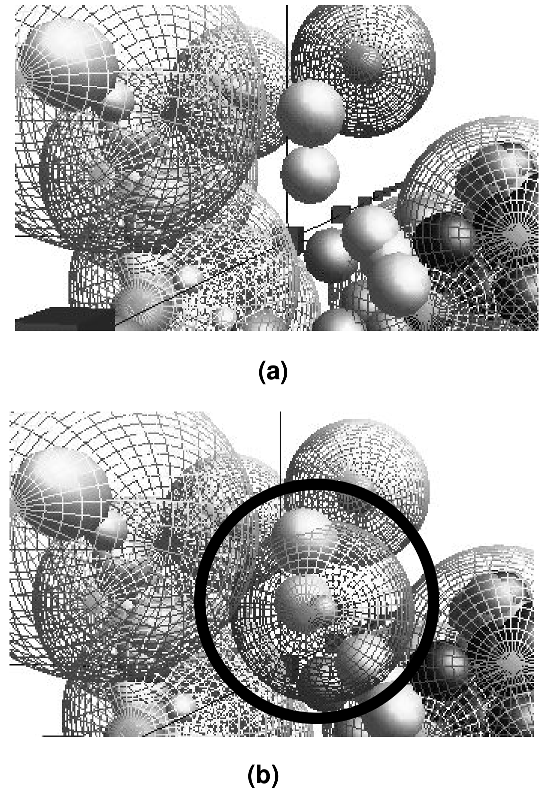


Figure 6. RBFN of Glass data set. The user can add new basis functions around data clusters for modest performance gains. In this example on the "glass" data set, the misclassification rate on the test set went from 56% to 54% when this basis function was added.

III. Results

The visualization package was tested on 11 data sets of varying dimensionality, taken from the UCI depository. With data sets that had low intrinsic dimensionality it was easy to see how the radial basis function network was operating. The geometric relations between of the basis functions and the data clusters, along with how each basis affected the classification was clear. Also the basis centers could be manually edited in the three-dimensional space to improve the performance of the network.

A. Iris data set

This visualization proved that it could be effective in visualizing RBFN with simple data sets such as the Iris data set (Figure 5). The clusters in the data set can easily be seen as well as the influence of each basis function with respect to the data. Looking at the weights it was

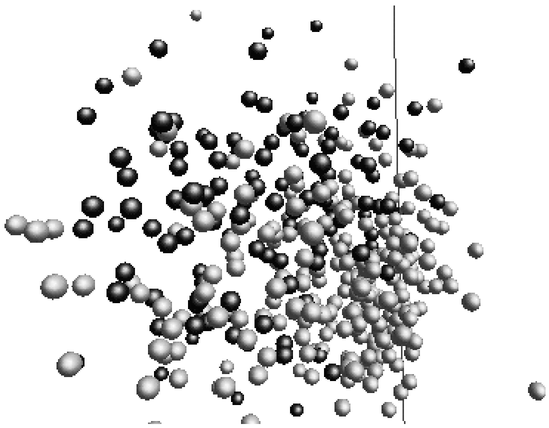


Figure 7: Visualization of Diabetes data set.

also easy to interpret by inspection how the network was discriminating between different classes.

B. Glass data set

This data set had much higher dimensionality, but our visualization method was still able to convey a lot of information about the RBFN built around this data set, such as clustering. Also through visual inspection the network could be edited to improve performance (Figure 6).

C. Diabetes data set

This data set had eight dimensions, with all dimensions having about an equal variance. Since each dimension was about the same, PCA could do little to effectively reduce the data dimensionality. As a result it was hard to see patterns in the data when visualized (Figure 7). The visualization does give insight though on why it is difficult to create a good classifier with this data set.

IV. Future Work

As previously mentioned, our visualization tool shows its limitations in trying to visualize high dimensional data that cannot be reduced with PCA. Other linear data reduction methods can be applied such as of Fisher's linear discriminant, which looks at the target values of the input data to create a more informed transform. Also non-linear methods could be used, but then the graphical edits made to the RBFN in transformed space could not be easily transformed back to input space.

There are many other improvements in progress for this visualization tool. We are in the process of modifying the tool for visualization and manipulation of elliptical basis functions. In addition, we are adding more options for visualization of network outputs including confusion matrices and decision surfaces.

We are also in the process of looking at methods that would allow the user to perform more advanced "surgery" on the network that goes beyond the simple changing basis centers and size. These methods would involve allowing the user to specify how the outputs of the network should behave, and what areas in the data space are of special importance. Changes could then be made automatically to the network, and the user could immediately view the results.

V. Conclusion

Visualization is a powerful mechanism for representing data. It allows humans to discover complex patterns in otherwise incomprehensible data sets. Combining the results of previous work with novel visualization methods, we built a unique tool that visualizes radial basis function neural networks. The tool allows the visualization of data sets, basis functions, second-layer weights, and performance measures. It also allows the user to modify existing networks and provided feedback on these modifications. This tool has proved helpful in understanding RBF networks and also in improving their performance. With future additions, this tool could become an integral part of the development process for RBF networks and other similar networks. Many different types of researchers and developers could benefit from its visualization and network manipulation capabilities.

[1] H. Bischof, A. Pinz, and W. G. Kropatsch. (1992) "Visualization Methods for Neural Networks," In Proceedings of the *11th IARP International Conference on Pattern Recognition*, The Hague, Netherlands, 581-585.

[2] C. M. Bishop. (1995) *Neural Networks for Pattern Recognition*. New York: Oxford University Press.

[3] K.-Y. Chang and J. Ghosh, "Principal Curve Classifier - A Nonlinear Approach to Pattern Classification", 1998 *IEEE Intl. Conf. on Neural Networks*, Anchorage, May 1998, pp. 695-700.

[4] F. Fierens, G. G. Wilkinson, and I. Kanellopoulos. (1994) "Studying the Behaviour of Neural and Statistical Classifiers by Interaction in Feature Space," In Proceedings of the *Image and Signal Processing for Remote Sensing*, Rome, Italy, 483-493.

[5] G. Hinton. (1986) "Connectionist Learning Procedures," *Artificial Intelligence* 40: 185-234.

[6] J. Wejchert and G. Tesauro. (1991) "Visualizing Processes in Neural Networks," *IBM Journal of Research and Development* 35(1-2): 244-253.

[7] G. Whittington and T. Spracklen. (1992) "Applying Visualisation Techniques to the Development of Real-World Artificial Neural Network Applications," In Proceedings of the *Applications of Artificial Neural Networks III*, Orlando, FL, 1024-1033.