Letters

A Comparison of the von Mises and Gaussian Basis Functions for Approximating Spherical Acoustic Scatter

Rick L. Jenison and Kate Fissell

Abstract— **This paper compares the approximation accuracy of two basis functions that share a common radial basis function (RBF) neural network architecture used for approximating a known function on the unit sphere. The basis function types considered are that of a new spherical basis function, the von Mises function, and the now well-known Gaussian basis function. Gradient descent learning rules were applied to optimize (learn) the solution for both approximating basis functions. A benchmark approximation problem was used to compare the performance of the two types of basis functions, in this case the mathematical expression for the scattering of an acoustic wave striking a rigid sphere.**

I. Introduction

The use of radial basis functions (RBF's) for approximation emerged in the mathematics literature over a decade ago [14]. More recently, the technique has surfaced in the neural network literature as an alternative to nonlinear multilayer networks. A number of neural network researchers have investigated a network architecture that uses a single-layer of locally-tuned units covering a multidimensional space (e.g., [2], [9], [10], and [13]). In principle, these neural networks are designed to learn an input–output mapping from a set of examples, a procedure more generally regarded as surface approximation [13]. Multilayer sigmoidal neural networks have shown to be "universal approximators" [7], and this result has been extended to that of single-layer networks with Gaussian radial basis functions [5], [12].

Artificial neural networks and approximation techniques typically have been applied to problems conforming to a multidimensional Cartesian input space. Approximation problems that exist on a unit sphere, however, are most accurately represented by a spherical or polar coordinate system, rather than projected implicitly onto a two-dimensional Cartesian system. Spherical data arise in many areas of science such as geophysics and meteorology and projection geometry has been the focus of cartographic interest for problems involving a transformation of the spherical earth into a planar map. A well-known problem to cartographers is that area, direction and shape cannot be preserved simultaneously in these projections, hence some degree of spatial distortion must be incurred. In this paper, a spherically dependent function that models sound scattering on the surface of a rigid sphere will serve to generate a data-set for comparing basis functions since the analytic formula is available to assess the accuracy of an approximation to this known function. First, the basis functions are formally defined and then the standard RBF architecture is described. Rules for optimizing the parameters through learning are then presented. The von Mises function, which is well adapted to spherical input (see [4]) and the two-dimensional Gaussian are evaluated within the common RBF architecture with respect to accuracy and ability to generalize.

Manuscript received October 7, 1994; revised March 15, 1995.

The authors are with the Department of Psychology, University of Wisconsin, Madison, WI 53706 USA.

IEEE Log Number 9413258.

II. Gaussian Basis Function

The Gaussian basis function (GBF) is one of the most commonly used basis functions for approximation. The two-dimensional Gaussian with a center at (α, β) can be represented as

$$G(\theta, \phi, \alpha, \beta, \sigma_\alpha, \sigma_\beta) = e^{-\left(\frac{\theta-\alpha}{\sigma_\alpha}\right)^2} e^{-\left(\frac{\phi-\beta}{\sigma_\beta}\right)^2} \quad (1)$$

where the two-dimensional input space corresponds to samples of azimuth-by-elevation (θ, ϕ) mapped directly from the sphere to the plane. Any sample on the azimuth-elevation plane yields an output from each GBF proportional to the distance between the sampled and the center (α, β) of the GBF. The GBF is a member of the class of radial basis functions as defined by Micchelli [8].

III. von Mises Basis Function

The von Mises basis function (VMBF) is based on a spherical probability density function that has been used to model line directions distributed unimodally with rotational symmetry.[1] The expression for the von Mises function without the surface region coefficient $(\sin\phi)$ and a constant of proportionality is

$$VM(\theta, \phi, \alpha, \beta, \kappa) = e^{\kappa(\sin\phi \sin\beta \cos(\theta-\alpha)+\cos\phi\cos\beta)} \quad (2)$$

where the input space corresponds to samples of azimuth-by-elevation (θ, ϕ) on the sphere (see [3]). Application of the von Mises function requires an azimuthal range from zero to 2π and elevational range from zero to π. Any sample (θ, ϕ) on the sphere will induce an output from each VMBF proportional to the angle difference between the sample and the centroid of the VMBF (α, β). κ is a shape parameter called the concentration parameter, where the larger the value the narrower the function width. The VMBF resembles a bump on a sphere and can be viewed as the spherical correlate of a two-dimensional Gaussian on a plane.

IV. A Common Neural Architecture

The basis function serves as the activation function for each of the hidden layer units of the RBF architecture of the form described below. The output of the ith output node, $f_i(\theta, \phi)$, when spherical coordinates[2] are presented as input, is given by the following approximation functions

$$f_i(\theta, \phi) = \sum_j^J w_{ij} G(\theta, \phi, \alpha_j, \beta_j, \sigma_{\alpha j}, \sigma_{\beta j}) \quad (3)$$

$$f_i(\theta, \phi) = \sum_j^J w_{ij} VM(\theta, \phi, \alpha_j, \beta_j, \kappa_j) \quad (4)$$

where $G(\theta, \phi, \alpha_j, \beta_j, \sigma_{\alpha j}, \sigma_{\beta j})$ and $VM(\theta, \phi, \alpha_j, \beta_j, \kappa_j)$ represent the output of the jth Gaussian and von Mises basis function, respectively, and w_{ij} is the weight connecting the jth basis function with the ith output node. These expressions can be viewed as

[1] The probability distribution is known as the von Mises–Arnold–Fisher distribution or alternatively the Fisher distribution.

[2] The input coordinate system is spherical for both basis functions. The Gaussian, however, is considered to be localized on a two-dimensional planar projection of azimuth and elevation.

expanding the input–output mapping into a linear combination of weighted nonlinear basis functions.

V. Parameter Learning

The approximation parameters (3) and (4) for a fixed number of basis functions can be optimized by applying a gradient descent method to a particular cost function for the training input–output set. In this case, we require the sum-of-squared-error to be minimized, which is a standard technique that has been applied to GBF neural networks [10], [13], and [6]. We have derived analogous gradient equations for the von Mises basis [4]. The expression of this cost function for the pth teaching pattern is

$$E_p = \frac{1}{2}\sum_{i}^{M}(t_{ip} - f_{ip}(\theta, \phi))^2 \tag{5}$$

where t_{ip} is the ith element of the teaching pattern and $f_{ip}(\theta, \phi)$ is the ith element of the output pattern (both M-dimensional). Parameter values are learned through the successive presentation of input–output teaching pairs and application of the specific update rules for each parameter of the Gaussian or von Mises basis function. The amount of change made to each parameter following successive presentation is based on the negative derivative of the error with respect to that parameter, which requires the application of the chain rule for partial differential equations.

The specific update rules for the Gaussian basis functions have the following form

$$\Delta w_{ij} = -\frac{\partial E}{\partial w_{ij}} = (t_i - f_i(\theta, \phi))\, G(\theta, \phi, \alpha_j, \beta_j, \sigma_{\alpha j}\, \sigma_{\beta j}) \tag{6}$$

$$\Delta \alpha_j = -\frac{\partial E}{\partial \alpha_j} = \frac{2(\theta - \alpha_j)}{\sigma_{\alpha j}^2} \cdot \sum_{i}^{M}[(t_i - f_i(\theta, \phi))\, w_{ij}]\, G(\theta, \phi, \alpha_j, \beta_j, \sigma_{\alpha j}\, \sigma_{\beta j}) \tag{7}$$

$$\Delta \beta_j = -\frac{\partial E}{\partial \beta_j} = \frac{2(\phi - \beta_j)}{\sigma_{\beta j}^2} \cdot \sum_{i}^{M}[(t_i - f_i(\theta, \phi))\, w_{ij}]\, G(\theta, \phi, \alpha_j, \beta_j, \sigma_{\alpha j}, \sigma_{\beta j}) \tag{8}$$

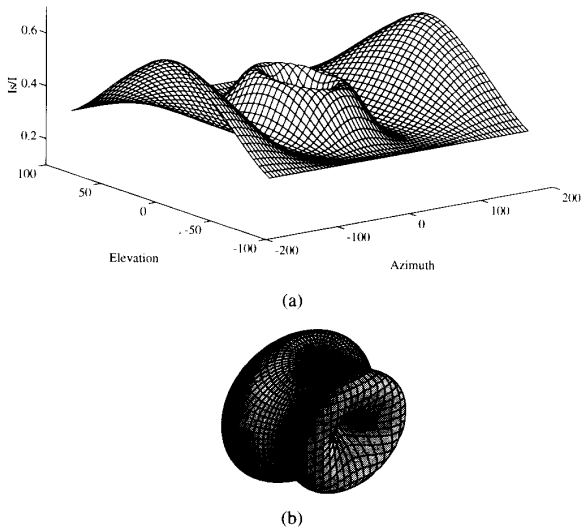
$$\Delta \sigma_{\alpha j} = -\frac{\partial E}{\partial \sigma_{\alpha j}} = \frac{2(\theta - \alpha_j)^2}{\sigma_{\alpha j}^3} \cdot \sum_{i}^{M}[(t_i - f_i(\theta, \phi))\, w_{ij}]\, G(\theta, \phi, \alpha_j, \beta_j, \sigma_{\alpha j}, \sigma_{\beta j}) \tag{9}$$

$$\Delta \sigma_{\beta j} = -\frac{\partial E}{\partial \sigma_{\beta j}} = \frac{2(\phi - \beta_j)^3}{\sigma_{\beta j}^2} \cdot \sum_{i}^{M}[(t_i - f_i(\theta, \phi))\, w_{ij}]\, G(\theta, \phi, \alpha_j, \beta_j, \sigma_{\alpha j}, \sigma_{\beta j}). \tag{10}$$

Similarly, the gradient descent update rules for the von Mises basis functions have the following form

$$\Delta w_{ij} = -\frac{\partial E}{\partial w_{ij}} = (t_i - f_i(\theta, \phi))\, VM(\theta, \phi, \alpha_j, \beta_j, \kappa_j) \tag{11}$$

$$\Delta \alpha_j = -\frac{\partial E}{\partial \alpha_j} = \kappa_j(\sin\phi \sin\beta_j \sin(\theta - \alpha_j)) \cdot \sum_{i}^{M}[(t_i - f_i(\theta, \phi))\, w_{ij}]\, VM(\theta, \phi, \alpha_j, \beta_j, \kappa_j) \tag{12}$$



(a)

(b)

Fig. 1. Directivity-pattern of scatter (15) about a rigid sphere. (a) Plotted on a planar Cartesian grid. (b) Plotted without projection distortion on the sphere.

$$\Delta \beta_j = -\frac{\partial E}{\partial \beta_j} = \kappa_j(\sin\phi \cos\beta_j \cos(\theta - \alpha_j) - \cos\phi \sin\beta_j) \cdot \sum_{i}^{M}[(t_i - f_i(\theta, \phi))\, w_{ij}]\, VM(\theta, \phi, \alpha_j, \beta_j, \kappa_j) \tag{13}$$

$$\Delta \kappa_j = -\frac{\partial E}{\partial \kappa_j} = (\sin\phi \sin\beta_j \cos(\theta - \alpha_j) + \cos\phi \cos\beta_j) \cdot \sum_{i}^{M}[(t_i - f_i(\theta, \phi))\, w_{ij}]\, VM(\theta, \phi, \alpha_j, \beta_j, \kappa_j). \tag{14}$$

VI. Approximating Spherical Acoustic Scatter

When a sound wave encounters an object, some of the wave is deflected from its original course. A scattered wave spreads out from the object, distorting and interfering with the incident plane wave. The expression for the scattered intensity derived in Morse and Ingard (1986) is

$$\frac{I_s}{I} \cong \frac{a^2}{4r^2} + \frac{a^2}{4r^2}\cot^2\left(\frac{\vartheta}{2}\right) J_m^2(ka\sin\vartheta) \tag{15}$$

where a is the radius of the sphere and r is the radial distance to the origin of a sound source. The constant k is known as the wave number and is equal to $\frac{2\pi f}{c}$ where f is the frequency of the source and c is the speed of sound. The function J_m is an mth order Bessel function of the first kind. The solid angle ϑ is computed relative to the line normal to the incident wavefront. This expression approximates intensity scatter for sounds with short wavelengths (e.g., $ka \gg 1$). Fig. 1 shows a three-dimensional rendering of the scatter that was used to compare the approximation performance of the two types of basis function. The surface was generated using (15) with the radius a of the theoretical rigid sphere equal to .09 m and the distance r of the sound source from the sphere equal to .45 m. The frequency of the sound source was three kHz. The magnitude of the scattered intensity is shown as a function of direction in azimuth and elevation. Fig. 1(a) shows the theoretical intensity surface projected on a Cartesian grid. Figl 1(b) shows the actual nondistorted surface plotted as it would appear relative to the unit sphere. The characteristics of this surface represent a

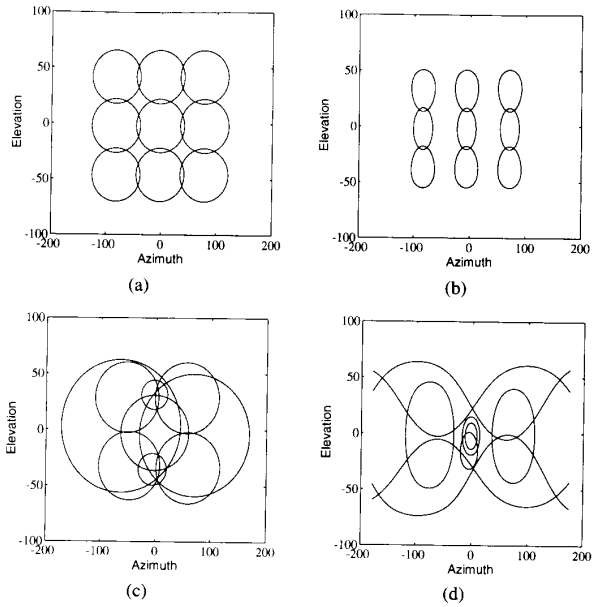


Fig. 2. Placement and shape of the Gaussian and von Mises basis functions shown as contours measured at 75% of peak. All contours are shown projected onto a plane, hence the von Mises functions, while symmetrical when projected on the sphere, will appear distorted on the plane. (a) Nine Gaussian basis functions following initialization of the RBF network. (b) Nine von Mises basis functions following initialization of the RBF network. (c) Gaussian basis functions following asymptotic learning. (d) von Mises basis functions after asymptotic learning.

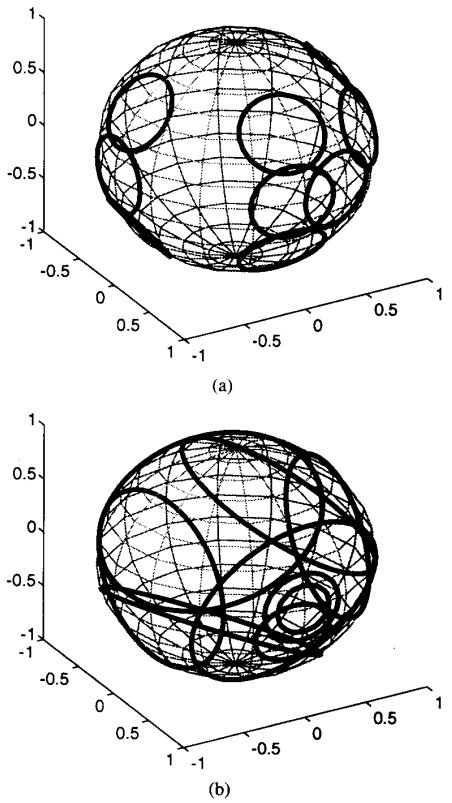


Fig. 3. Contours of nine von Mises basis functions placed on the sphere without projection distortion (a) initial placement and (b) following asymptotic learning. Parameter values are the same as those used in Fig. 2, panels a and b. The axes reflect the radius of the arbitrary unit sphere.

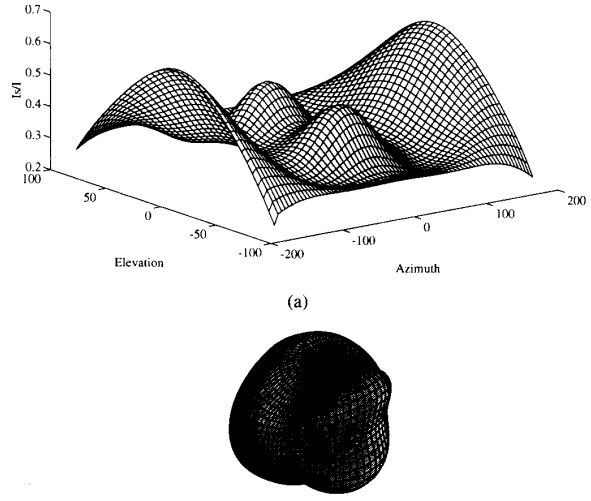
moderate challenge to functional approximation using a small number of basis functions. One challenge, in particular to the Gaussian basis function approximation, is the necessary condition that the edges (± 180 degrees azimuth) in the Cartesian input representation "match up." The second challenge is the implicit singularity at the poles; that is, at ± 90 degrees elevation all samples of azimuth are, by definition, of equal magnitude.

The two basis function were allowed to compete within a common RBF architecture by iteratively applying their respective gradient descent rules for parameter learning. A relatively small number of basis functions (nine) were used. Uniform starting placements were used for both basis functions. Fig 2(a) and (b) show the initial placement and contours of the Gaussian and von Mises basis functions respectively. Note that a Cartesian grid is being used to compare the characteristics of the two basis function types. For the case of the von Mises basis functions, the contours will appear distorted for the reasons discussed above.

A training database was obtained by uniform random sampling of 500 directions expressed in azimuth and elevation. Equation (15) was used to synthesize the corresponding teaching pattern ($t_p \equiv I_s/I$) for each of the randomly selected input directions. Gradient descent learning was allowed to proceed until asymptotic behavior was observed for both networks. This generally occurred within 200 epochs of input–output training. Fig 2 (a) and (b) show the resulting final placement and contour of the Gaussian and von Mises functions. Despite the contours of the von Mises functions appearing distorted on the Cartesian grid, the placements for both types of basis functions show a high degree of symmetry, in particular for the Gaussian basis functions. Fig. 3 is provided to further aid in the visualization of the von Mises basis function placements. Fig. 3 (a) and (b) show the actual spherical contours for the initial and final placement of the nine basis functions corresponding to the planar Cartesian projections in Fig. 2 (b) and (d). Random initial conditions have been explored in addition to the uniform initial conditions shown in Figs. 2 and 3. While all cases converged, not all final solutions show this degree of symmetry upon convergence, even though the final total sum-of-squared-error is consistently the same within basis function type.

The approximated surfaces are shown in Figs. 4 and 5 for the optimized Gaussian and von Mises basis functions, respectively. The surfaces were derived by interrogating each approximation function (3) and (4) with a dense test sampling (2664 points) of the input space (i.e., azimuth and elevation). The planar Cartesian grid is shown in panel a and the spherical surface plot is shown in panel b and should be compared to the known surface plotted in corresponding fashion in Fig. 1.

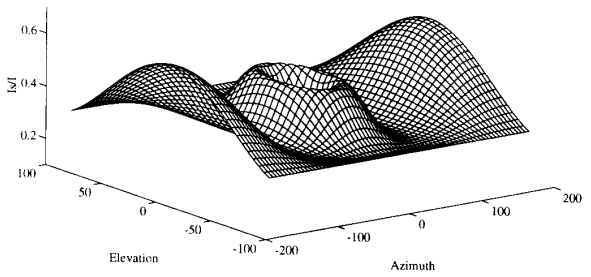
The Gaussian basis approximation in Fig. 4(b) shows evidence of the ± 180 degree "seam" corresponding to the anticipated edge effects for this basis function type. Most apparent is the loss of the cup-like scatter at zero degrees azimuth, zero degrees elevation. An additional problem is the lack of singularity at the poles (± 90 degrees), that is, equivalence across azimuth. The total sum-of-squared-error for a network of nine Gaussian basis functions was 2.117. In contrast, the network of nine von Mises basis functions closely approximates the acoustic scatter function as shown in Fig. 5. The corresponding total sum-of-squared-error was .008, nearly three orders of magnitude improvement over that of the Gaussian network with an equivalent number of basis functions. Not only was the von



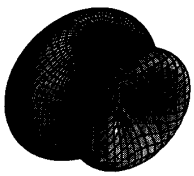
(a)

(b)

Fig. 4. Approximated directivity-pattern of scatter on a rigid sphere using nine Gaussian basis functions trained on 500 input–output pairs. (a) Plotted on a planar Cartesian map. (b) Plotted without projection distortion on the unit sphere.



(a)



(b)

Fig. 5. Approximated directivity-pattern of scatter on a rigid sphere using nine von Mises basis functions trained on 500 input–output pairs. (a) Plotted on a planar Cartesian map. (b) Plotted without projection distortion on the unit sphere.

Mises approximation more accurate, but also demonstrated better generalization as evidenced by the dense sampling of novel test data.

VII. Conclusions

The accuracy of the approximation for our benchmark acoustic scatter problem is several orders of magnitude better with the von Mises basis function network than that of the Gaussian network, when an equal number of basis functions are used and provided with equivalent opportunity to learn. We speculate that this is due to the well-matched characteristics of the spherical von Mises function to spherical problems. The advantage will likely decline as a function of increasing the number of basis functions used in the network. The benefit of using fewer basis functions, however, besides the computational efficiency, is that when the number of observations is limited, there will be less chance of overfitting and consequently better generalization [1].

References

[1] C. Bishop, "Improving the generalization properties of Radial Basis Function neural networks," *Neural Comp.*, vol. 3, pp. 579–588, 1991.
[2] D. S. Broomhead and D. Lowe, "Multivariable functional interpolation and adaptive networks," *Complex Syst.*, vol. 2, pp. 321–355, 1988.
[3] N. I. Fisher, T. Lewis, and B. J. J. Embleton, *Statistical Analysis of Spherical Data.* Cambridge, UK: Cambridge University Press, 1987.
[4] R. L. Jenison and K. Fissell, "Radial Basis Function neural network for modeling auditory space," *J. Acoust. Soc. Amer.*, vol. 95, p. 2898, 1994.
[5] E. J. Hartman, J. D. Keeler, and J. M. Kowalski, "Layered neural networks with Gaussian hidden units as universal approximations," *Neural Comp.*, vol. 2, pp. 210–215, 1990.
[6] ——, "Predicting the future: Advantages of Semilocal Units," *Neural Comp.*, vol. 3, pp. 566–578, 1991.
[7] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks,* vol. 2, pp. 359–368, 1989.
[8] C. A. Micchelli, "Interpolation of scattered data: distance matrices and conditionally positive definite functions," *Constructive Approximation,* vol. 2, pp. 11–22, 1986.
[9] J. Moody and C. Darken, "Learning with localized receptive fields," in *Proc. Connectionist Models Summer School,* 1988, pp. 1–11.
[10] ——, "Fast learning in networks of locally-tuned processing units," *Neural Comp.*, vol. 1, pp. 281–294.
[11] P. M. Morse and K. U. Ingard, *Theoretical Acoustics.* Princeton, NJ: Princeton Univ. Press, 1986.
[12] J. Park and I. W. Sandberg, "Universal approximation using radial-basis-function networks," *Neural Comp.*, vol. 3, pp. 246–257, 1991.
[13] T. Poggio and F. Girosi, "Networks for approximation and learning," *Proc. IEEE,* vol. 78, pp. 1481–1496.
[14] M. J. D. Powell, "Radial basis functions for multivariable interpolation: A review," in *Algorithms for Approximation,* J. C. Mason and M. G. Cox, Eds. Oxford, UK: Clarendon Press, 1987, pp. 143–166.

Optimization Neural Network for Solving Flow Problems

Renzo Perfetti

***Abstract*— This paper describes a neural network for solving flow problems, which are of interest in many areas of application as in fuel, hydro, and electric power scheduling. The neural network consists of two layers: a hidden layer and an output layer. The hidden units correspond to the nodes of the flow graph. The output units represent the branch variables. The network has a linear order of complexity, it is easily programmable, and it is suited for analog very large scale integration (VLSI) realization. The functionality of the proposed network is illustrated by a simulation example concerning the maximal flow problem.**

Manuscript received July 6, 1993; revised September 8, 1994.

The author is with the Istituto di Elettronica, University of Perugia I-06100 Perugia, Italy.

IEEE Log Number 9413257.