Resolution of a Neural Network Direction Finding Array

Hugh L. Southall
USAF Rome Laboratory
Electromagnetics and Reliability Directorate
Hanscom AFB, MA, 01731-3010

1 Introduction

We have developed neural network based signal processing techniques for array antennas [1]. The objective is to train the network at angles within the antenna field of view to perform the desired antenna function with arrays which cost less to manufacture. In this paper, we show that the direction finding (DF) resolution increases with larger array apertures the same way it does for a conventional rf beampformer.

2 The Neural Network DF Array

The neural beamformer with its associated preprocessing and postprocessing has been described in detail in [1]. The system consists of antenna measurement preprocessing, an artificial neural network, and postprocessing. It is important to note that we preprocess raw antenna measurements; there is no calibration or traditional compensation look-up tables.

The neural network architecture was described in [1]. Network input nodes receive preprocessed antenna data and broadcast this data as input vectors to hidden layer processing nodes. From a set of \( L \) measurements at different angles of arrival (AOAs) within the antenna field of view (\(-60^\circ\) to \(+60^\circ\)), we obtain \( L \) input vectors \( \mathbf{z}_i = (z_{i1}, z_{i2}, \ldots, z_{i(N-1)}) \), where \( i = 1, 2, \ldots, L \). The components of each input vector are the \( N-1 \) phase differences between the \( N \) array elements. Measured input vectors include the effects of array element nonuniformities, degradations, etc., and, possibly, the effects of the local environment, such as near field scattering. When we develop a new neural net application, we select a subset of the \( L \) input vectors to train the network, then we use the rest of them to test the network to see how well it performs. The network contains a hidden layer of processing nodes with nonlinear activation functions. Processing node \( i \) has an output which depends on the radial distance between the input vector and node center \( m_i \), and on \( \sigma \), the spread parameter. This type of network is called a radial basis function (RBF) network. We use periodic basis functions (PBFs) to overcome network training difficulties associated with \( 2\pi \) periodic phase discontinuities introduced by the phase receiver and arbitrary initial phases [2].

When we present the network with input vector \( \mathbf{z}_i \), the activation, or output, of processing node \( i \) is given by the PBF [3]

\[
\phi_{ii} = \exp \left( -\frac{\sum_{k=1}^{N-1} \cos(z_{ik} - m_{ik})}{\sigma^2} \right),
\]

where \( i = 1, 2, \ldots, q \), for \( q \) hidden layer nodes. Center \( m_{ik} \) is equal to the input vector at the \( k \)th training angle.

*This research was sponsored by the US Air Force Office of Scientific Research (AFOSR), Bolling AFB, DC.

U.S. Government Work Not Protected by U.S. Copyright

1988
The response at each output node due to input \( z_j \) is a weighted sum of processing node values given by

\[
y_j = \sum_{i=1}^{q} w_{ji} \phi_i
\]

where \( j = 1, 2, \ldots, r \) for \( r \) output nodes. We use the same number of processing nodes, \( q \), as output nodes, \( r \). The number of output nodes is equal to the number of training angles. The weights, \( w_{ji} \), are computed by training the network.

In Section 3, we use ideal input vectors, not measured data, to theoretically analyze the resolution of the neural network DF array. For ideal input vectors, each of the \( N-1 \) phase differences is equal to

\[
\pi(\theta) = \frac{2\pi}{\lambda} d \sin \theta = \pi \sin \theta = \pi u,
\]

for element separation \( d = \lambda/2 \), and \( u = \sin \theta \). Therefore, in Equation 1, the input vector and center do not depend on index \( k \), and the PBFs become

\[
\phi_i(u - u_i) = \exp \left( \frac{(N - 1) \cos(\pi(u - u_i)) - (N - 1)}{\sigma^2} \right),
\]

where \( u_i = \sin \theta_i \) corresponds to the \( i \)th PBF center for training angle \( \theta_i \), \( i = 1, 2, \ldots, r \). To simplify, we analyze a two node/two element system (\( N=2 \)), where the PBFs are

\[
\phi_i(u - u_i) = \exp \left( \frac{\cos(\pi(u - u_i)) - 1}{\sigma^2} \right).
\]

The response of the first output node is

\[
y_1(u) = w_{11} \phi_1(u - u_i) + w_{12} \phi_2(u - u_j).
\]

In Section 3, we analyze the resolution of a neural network DF array by comparing it with a Butler matrix DF beamformer [3]. Like a Butler matrix beam, the response of output node one is normalized to unity at \( u_1 \) and forced to be zero at \( u_2 \). This can be expressed as

\[
y_1(u_1) = w_{11} \exp \left( \frac{\cos(\pi(u_1 - u_j)) - 1}{\sigma^2} \right) = 1
\]

\[
y_1(u_2) = w_{11} \exp \left( \frac{\cos(\pi(u_2 - u_i)) - 1}{\sigma^2} \right) + w_{12} = 0,
\]

which can be solved for the two unknown weights below.

\[
w_{11} = \frac{1}{1 - \exp \left( \frac{\cos(\pi(u_1 - u_j)) - 1}{\sigma^2} \right)}
\]

\[
w_{12} = -\frac{\exp \left( \frac{\cos(\pi(u_2 - u_i)) - 1}{\sigma^2} \right)}{1 - \exp \left( \frac{\cos(\pi(u_1 - u_j)) - 1}{\sigma^2} \right)}
\]

Solving for the weights constitutes network training.

3 Neural Network Resolution

The resolution of the Butler matrix increases (DF angle error becomes smaller) as \( N \) increases. Classic rf beamformers, such as the Butler matrix and monopulse [3] coherently sum \( N \) element rf signals and one expects greater resolution for larger \( N \) (larger antenna aperture). The neural network DF array, unlike classic beamformers, does not pass the entire rf signal through the network. Without coherent signal combination, it is unclear how neural network resolution should depend on aperture size, especially for ideal (simulated) data where all inputs (phase differences) are equal to the same number for a given AOA.

Recall that each neural network output node response is a weighted sum of basis functions. For ideal input vectors at the training angles and \( N=2 \), the network weights are given in Equations 9 and 10. For larger \( N \), analytical solutions are impractical, and we must use a
Figure 1: Log of the DF RMS angle error in degrees as a function of the number of network output nodes (dashed) and number of Butler matrix beams (solid). The circles indicate 2, 4, 8 and 16 beam systems.

Figure 2: Neural net output node response (dashed) for a two node/two element system compared with the corresponding modified sinc beam of a two beam/two element Butler matrix (solid).

computer to solve for the weights. Additionally, we use the Butler matrix orthogonal beam locations for network training points, so \( u_1 = -5 \) and \( u_2 = 1.5 \). We can calculate \( y_1(u) \) using Equation 6, again using ideal input vectors from Equation 3. In a similar manner, we can calculate \( y_2(u) \). We use values of these adjacent output node responses to interpolate and obtain an estimate of the AOA [1]. This estimate is used to calculate the RMS angle error, which is the RMS difference between estimated and actual AOs over the field of view.

In Figure 1, we plot the log of the RMS angle error for \( N=2,4,8 \) and 16 for both the network and the Butler matrix. For the network, \( \sigma \) has been optimized for minimum angle error. The neural network resolution increases with \( N \) as does for the classic beamformer. The Butler matrix resolution is better, which is expected for the ideal case with perfectly matched elements. The fundamental reason that the Butler matrix resolution increases with \( N \) is that the beams get narrower and are more closely spaced. Therefore, interpolation between beam centers to obtain an AOA estimate becomes more accurate. The \( j \)th Butler matrix beam is a modified sinc function, \( \sin[N \frac{\pi}{4}(u - u_j)]/N \sin[\frac{\pi}{4}(u - u_j)] \), for \( d = \frac{\lambda}{4} \) [3], and network output node responses are closely related to these beams, as is resolution.

From Equations 5 and 6, note that \( y_1(u) \) depends on the internal network parameter \( \sigma \).
for any $u$. For example, $y_1(0) = \frac{1}{\sqrt{2}} \sinh \frac{1}{\sqrt{2}}$, which quickly converges to .5 for $\sigma$ larger than about 2. In fact, $y_1(u)$ is independent of $\sigma$ when $\sigma$ is larger than 2. We plot $y_1(u)$ using $\sigma = 3$ in Figure 2 and compare it with a corresponding Butler matrix beam. The agreement is poor, which is why the neural net RMS angle error is larger than the Butler matrix error for $N=2$ in Figure 1. For $N=4$, we plot a similar node response in Figure 3, where the agreement is better, particularly within the main beam region between zero crossings on either side of the beam center. Even though the neural network node responses converge to the Butler matrix modified sinc beams for larger $N$, there is still a small difference between them. As shown in Figure 1, this difference results in a larger angle error for the neural network. The absolute value of the difference obviously decreases rapidly with $N$.

Acknowledgements

Dr. Bob Mailloux of Rome Laboratory provided valuable insight into the analogy between the neural network DF array and the Butler matrix. Thanks to Drs. Fran Narcowich and Joe Ward of Texas A&M for helpful discussions on Section 2 and PBFs in particular.

References


Figure 3: Neural net output node response (dashed) for a four node/four element system compared with the corresponding modified sinc beam of a four beam/four element Butler matrix (solid).