HIGH-RESOLUTION BEARING ESTIMATION VIA UNITARY DECOMPOSITION ARTIFICIAL NEURAL NETWORK (UNIDANN)

Shun-Hsyung Chang, Tong-Yao Lee

Department of Electrical Engineering
National Taiwan Ocean University
Keelung, Taiwan, R.O.C.
E-mail:b0091@notu66.notu.edu.tw

Wen-Hsien Fang

Department of Electronic Engineering
National Taiwan Institute of Technology
43 Keelung Rd., Taipei, Taiwan, R.O.C.
E-mail:whf@et.ntit.edu.tw

ABSTRACT

In this paper, a novel artificial neural network (ANN) called the UNItary Decomposition ANN (UNIDANN), which can perform the unitary (Schur) decomposition of the synaptic weight matrix, is presented. It is shown both analytically and quantitatively that if the synaptic weight matrix is positive definite and normal, the dynamic equation involved will converge to a unitary matrix which can transform the weight matrix into an upper triangular one via the Schur decomposition. In particular, if the synaptic weight matrix is also Hermitian (symmetric for real case), the UNIDANN will perform the eigendecomposition. Compared with other existing ANN's, the proposed one possesses several attractive features such as more versatile in the sense that it is capable of performing the Schur decomposition, low computation time and no synchronization problem due to the application of the structure of analog circuit, and faster convergent speed. Some simulations with particular emphasis at the MUSIC bearing estimation algorithm have been provided to justify the validity of the proposed ANN.

1. INTRODUCTION

The Direction of Arrival (DOA) problem, which occurs in many other areas of signal processing such as sonar, radar, and seismology etc., has been an active research area in recent years. The traditional methods like linear prediction (maximum entropy) and minimum variance methods can not produce satisfactory performance when two sources are closely located or when the signal-to-noise ratio (SNR) is not high enough. To overcome this difficulty, many high-resolution methods such as Pisarenko, MUSIC, ESPRIT, and their variations [1] have been proposed. These high-resolution methods are all based on the subspace concept and require the eigenstructure decomposition (ED) of the input correlation matrix. This ED procedure calls for a lot of computations, thus hindering these subspace-based methods from real-time applications.

Recently, there has been a resurgence of interest in the investigation of the Artificial Neural Network (ANN). An ANN contains volumnious connected simple processing elements (neurons) and it can offer massively parallel computing capability. Hence, it becomes an attractive approach to alleviate the high computations involved in the ED. The utilization of neural networks for high-resolution bearing estimation was first considered by Rastogi *et al* [2], and later improved by Jha *et al* [3] to achieve more likely the global minima. Both methods utilize the Hopfield model by associating the minimization problem with the corresponding Liapunov's energy function. This type of ANNs, however, suffers the drawbacks of requiring a large number of neurons and is unable to achieve the global minima every time. Another approach is based on the hebbian rule. For example, a digital neural network called APEX and its complex extension for extracting the principle components of the input correlation matrix has been recently reported in [4] and [5], respectively. On the other hand, Luo and Li [6] considered an analog neural network by using the generalized hebbian algorithm (GHA) addressed by Sanger [7].

In this paper, we extend the GHA and come up with a novel neural network called UNItary Decomposition ANN (UNIDANN) which can perform the unitary (Schur) decomposition [8] of the synaptic weight matrix. It is shown both *analytically* and *quantitively* that if the synaptic weight matrix is positive definite and normal, the dynamic equation involved will converge to a unitary matrix which can transform the weight matrix into an upper triangular one via the Schur decomposition. Moreover, the transformed triangular matrix has diagonal elements corresponding to the eigenvalues of the synaptic weight matrix in a descending order. In particular, if the synaptic weight matrix is also Hermitian (symmetric for real case), the UNIANN will perform the ED and the associated principle components will correspond to the front portion of the resulting unitary matrix.

Compared with other existing ANNs, the UNIDANN possesses the following attractive features: (1) It can perform the Schur decomposition so that it's more versatile than other existing ones like [2, 3, 4, 5, 6], which can perform the ED for symmetric positive definite matrix only, (2) Compared with the discrete counterparts of [4, 5], it obeys the continuous-time dynamics, thus possessing the merits

This work was supported by National Science Council of R.O.C. under contract NSC 83-0404-E019-004 and 82-0404-E011-197.

0-7803-2431-5/95 $4.00 © 1995 IEEE

of the analog circuits such as low computation time and no synchronization problem, (3) It offers faster convergent speed and more accurate final results than the analog ANN recently reported in [6].

2. THE MUSIC BEARING ESTIMATION ALGORITHM

Consider a Uniform Linear Array (ULA) of M omnidirectional sensors impinged by $N(N < M)$ narrowband zero-mean signals, $s_k(t)$, $k = 1, 2, \ldots, N$. Assume that the signals (of the central frequency ω_0) are stationary and far-field to allow for a planar wavefront approximation. Then the output of the sensor array can be described by

$$\mathbf{Y}(t) = \mathbf{AS}(t) + \mathbf{N}(t), \qquad t = 1, 2, \ldots, L \qquad (1)$$

where

$$\begin{aligned}
\mathbf{Y}(t) &= [y_1(t), y_2(t), \ldots, y_M(t)]^T, \ t = 1, 2, \ldots, L \\
\mathbf{A}(\theta) &= [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N], \\
\mathbf{a}_k &= [1, \ e^{j\tau_k}, \cdots, e^{j(M-1)\tau_k}]^T, \ k = 1, 2, \cdots, N \\
\tau_k &= \frac{\omega_0 d \sin\theta_k}{c}, \ k = 1, 2, \cdots, N \\
\mathbf{S}(t) &= [s_1(t), s_2(t), \ldots, s_N(t)]^T, \ t = 1, 2, \ldots, L \\
\mathbf{N}(t) &= [n_1(t), n_2(t), \ldots, n_M(t)]^T, \ t = 1, 2, \ldots, L
\end{aligned}$$

the superscript ^{T} denotes matrix transposition, matrix $\mathbf{A}$ stands for the steering matrix, $s_k(t)$ is the k^{th} narrowband signal arriving from angle θ_k, d is the distance between two adjacent sensors, and $n_i(t)$ are the additive zero mean white noise.

Assume that the signal sources are noncoherent, then the correlation matrix of the input sequence becomes

$$\mathbf{R} = E[\mathbf{YY}^H] = \mathbf{ADA}^H + \sigma^2\mathbf{I}, \qquad (2)$$

where the superscript ^{H} denotes the Hermitian transposition,

$$\begin{aligned}
\sigma^2 &= E[|n_i(t)|^2] = \text{noise power}, \quad i = 1, 2, \ldots, M, \\
\mathbf{D} &= \text{diagonal matrix of } P_1, P_2, \ldots P_N
\end{aligned}$$

and P_i, $i = 1, \cdots, N$, denotes the power of the i^{th} signal. Let $\mathbf{v}_i$ be the orthonormal eigenvector of $\mathbf{R}$ corresponding to eigenvalue λ_i ($\lambda_1 \geq \ldots \geq \lambda_N \geq \lambda_{N+1} \geq \cdots \geq \lambda_M$). Since eigenvectors $\{\mathbf{v}_{N+1}, \mathbf{v}_{N+2}, \ldots, \mathbf{v}_M\}$ all correspond to eigenvalue σ^2 (noise component), the subspace spanned by $\mathbf{v}_{N+1}, \mathbf{v}_{N+2}, \ldots, \mathbf{v}_M$ is interpreted as the noise subspace by Schmidt [9], while the subspace spanned by $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$ is interpreted as the signal subspace. The MUSIC algorithm proposed by Schmidt is to use the orthogonality of these two subspaces to locate the arriving angles θ_i.

The pseudospectrum of MUSIC algorithm is defined as

$$P_{MUSIC}(\theta) = \frac{1}{abs\{\mathbf{a}^H(\theta)\mathbf{E}_N\mathbf{E}_N^H\mathbf{a}(\theta)\}} \qquad (3)$$

where $\mathbf{E}_N = [\mathbf{v}_{N+1}, \ldots, \mathbf{v}_M]$, and abs$\{\cdot\}$ stands for the absolute value. In the MUSIC algorithm, when the arriving angle $\theta = \theta_i (i = 1, 2, \ldots, N)$, the denominator of $P_{MUSIC}(\theta)$ goes to zero since $\mathbf{a}(\theta_i)$ lies in the signal subspace. As a result, the peak values of $P_{MUSIC}(\theta)$ correspond to the directions of signals [9]. Unfortunately, the intensive computation load required for the MUSIC algorithm in the ED of the correlation matrix precludes it from real-time implementation. In the next section, we consider a new neural structure which possesses massive computing capability to overcome this difficulty.

3. UNITARY DECOMPOSITION ARTIFICIAL NEURAL NETWORK

In this section, we consider the characteristics of the performance of the proposed UNIDANN which is dictated by the following continuous-time dynamic equation:

$$\begin{aligned}
\frac{d}{dt}\mathbf{V}(t) &= c[\mathbf{WV}(t) - \mathbf{V}(t) \times TRI\{\mathbf{V}^H(t)\mathbf{WV}(t)\}] \\
&\quad \times [DIAG\{\mathbf{V}^H(t)\mathbf{WV}(t)\}]^{-1}. \qquad (4)
\end{aligned}$$

where $\mathbf{V}(t)$ is an $M \times N$ matrix corresponding to the output of the UNIDANN at time t, $\mathbf{W}$ is an $M \times M$ matrix standing for the synaptic weight matrix, and c is a positive constant related to the capacity and the resistance of the circuit. $TRI\{\cdot\}$ sets all elements above (or below) the diagonal to zero, while $DIAG\{\cdot\}$ sets elements in both areas to zero. A neural structure for the UNIDANN is shown in Fig. 1.

Based on the dynamics of (4), we have the following theorem:

Theorem: If the synaptic weight matrix $\mathbf{W}$ is positive definite and normal, and $N = M$, the UNIDANN will accomplish the Schur decomposition. More specifically, $\mathbf{V}(t)$ will converge to a unitary matrix $\mathbf{V}_{final}$ with probability 1 so that $\mathbf{S} = \mathbf{V}_{final}^H\mathbf{WV}_{final}$ is upper triangular. Moreover, the diagonal elements of $\mathbf{S}$ correspond to the eigenvalues of $\mathbf{W}$ in a descending order.

Proof: Here we only consider the case for $N = 1$ (the proof for general N can be found it [10]) and, without loss of generality, assume $c = 1$. Since $\mathbf{W}$ is normal, there esisting M linearly independent unit eigenvectors, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_M$ with corresponding eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_M$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$ which has been arranged in a descending order. Multiplying both sides of (4) by $\mathbf{v}_k$ yields

$$\frac{d}{dt}(\mathbf{v}_k\mathbf{V}) = \mathbf{v}_k^T\mathbf{V}[\lambda_k(\mathbf{V}^T\mathbf{WV})^{-1} - 1] \qquad (5)$$

where we have used the fact that $\mathbf{Vv}_k = \lambda_k\mathbf{v}_k, k = 1, \cdots, M$. Define $\theta_{k1} = \frac{\mathbf{v}_k^T\mathbf{V}}{\mathbf{v}_1^T\mathbf{V}}$ (assume $\mathbf{v}_k^T\mathbf{V}(0) \neq 0$), then we have

$$\frac{d}{dt}\theta_{k1} = (\lambda_k - \lambda_1)(\mathbf{V}^T\mathbf{WV})^{-1}\theta_{k1} \qquad (6)$$

where we have used (5). Since $\mathbf{W}$ is positive definite and $\lambda_1 \geq \cdots \geq \lambda_M$, $(\lambda_k - \lambda_1)(\mathbf{V}^T\mathbf{WV})^{-1}\theta_{k1} < 0$ if $k > 1$. Hence, $\theta_{k1} \to 0$ as $t \to \infty$ if $k > 1$. This implies that $\mathbf{V}_{final}$ lis in the direction of $\mathbf{v}_1$, i.e. $\mathbf{V}_{final} = c_1\mathbf{v}_1$. Substituting this into (5) with $k = 1$ and using the fact that $\mathbf{v}_1$ has unit length yields

$$\frac{d}{dt}c_1 = \frac{1}{c_1}(1 - c_1^2) \qquad (7)$$

If we define a corresponding Lyapunov function $U = (c_1^2 - 1)^2$, it can be readily verify that $\frac{d}{dt}U = 4(c_1^2-1)(1-c_1^2) < 0$. Therefore, $c_1 \to \pm 1$ and $\mathbf{V}_{final} \to \pm \mathbf{v}_1$, thus completing the proof. ■

In most of the practical applications, $\mathbf{W}$ is not only positive definite but also Hermitian (or symmetric if $\mathbf{W}$ is real). In such cases, the above theorem can be simplified to the following Corollary.

Corollary: When $\mathbf{W}$ is Hemitian positive definite, the output of the UNIDANN will approach the principle N eigenvectors of $\mathbf{W}$; that is,

$$\mathbf{V}_{final} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N], \tag{8}$$

where $\mathbf{v}_i$ is the eigenvector related to the eigenvalue λ_i of $\mathbf{W}$ with $\lambda_1 \geq \cdots \geq \lambda_N \geq \cdots \geq \lambda_M$.

To apply the proposed UNIDANN to the high-resolution bearing estimation algorithm like the MUSIC algorithm, we need to estimate the input spatial correlation matrix $\hat{\mathbf{R}}$ first by

$$\hat{\mathbf{R}} = \frac{1}{\ell}\sum_{t=1}^{\ell} \mathbf{Y}(t)\mathbf{Y}^H(t), \tag{9}$$

where $\mathbf{Y}(t)$ is the input data received at time t by the array and ℓ is the number of snapshots. Then we set the interconnection strength by using the computed correlation matrix which is obviously Hermitian positive definite. Therefore, the output of the UNIDANN will converge to the desired results by applying the Corollary above. The overall steps can be summarized in the following algorithm:

Algorithm: (UNIDANN for computing the signal subspace of the MUSIC algorithm)

1) Compute the input spatial correlation matrix according to (9).
2) Set the synaptic weight matrix $\mathbf{W}$ equal to $\hat{\mathbf{R}}$.
3) Select an appropriate N and initialize $\mathbf{V}(0)$ to some random matrix.
4) Update $\mathbf{V}(t)$ via the equation (4) until a stopping criterion is satisfied. Note that the number of the signals $\hat{N}$ can be obtained by checking the diagonal elements of $\mathbf{V}_{final}^H\mathbf{W}\mathbf{V}_{final}$. The first $\hat{N}$ columns of $\mathbf{V}_{final}$ will correspond to the eigenvectors of the signal subspace.

4. SIMULATION RESULTS AND DISCUSSION

In this section, some simulations are performed to justify the validity of the proposed UNIDANN.

Example 1

Consider an arbitrary positive definite matrix W as

$$\mathbf{W} = \begin{bmatrix} 1.2441 & -0.0970 & 0.4037 \\ 0.1782 & 1.8466 & -0.1567 \\ 0.4197 & -0.2576 & 1.7650 \end{bmatrix}$$

which has eigenvalues 2.0616, 1.7871, 1.0070. To obtain the Schur decomposition of $\mathbf{W}$, we applying the UNIDANN by setting the synaptic weight matrix as $\mathbf{W}$ and choosing the stopping criteria as $\|\mathbf{V}(n+1) - \mathbf{V}(n)\| < 10^{-10}$ with $\|\cdot\|$ denoting the 2-norm. After it converges, we can get

$$\mathbf{V}_{final} = \begin{bmatrix} -0.4524 & 0.1571 & -0.8779 \\ 0.2489 & 0.9675 & 0.0449 \\ -0.8564 & 0.1982 & 0.4768 \end{bmatrix}.$$

It can be readily verified that

$$\mathbf{V}_{final}^T\mathbf{W}\mathbf{V}_{final} = \begin{bmatrix} 2.0616 & 0.2191 & -0.0232 \\ 0 & 1.7871 & -0.1940 \\ 0 & 0 & 1.0070 \end{bmatrix}$$

and $\mathbf{V}_{final}^T\mathbf{V}_{final} = \mathbf{I}$.

where $\mathbf{I}$ is a 3×3 identity matrix. Therefore, it justifies that the UNIDANN can indeed accomplish the Schur decomposition.

Example 2

Consider a uniform linear array of 5 sensors and two sources at 42°, 46° with SNR=20 dB. The interelement sensor spacing is set as $\frac{\lambda}{2}$, and sensors are omnidirectional. The pseudospectra of the MUSIC algorithm by using the UNIDANN and the one addressed in [6] are shown in Figs. 2 and 3, respectively. A vivid description of the results can also observed by the corresponding contour plots as shown in Figs 4 and 5, respectively. We can observe that the UNIDANN converges faster under the stopping criteria $\|\mathbf{V}(n+1) - \mathbf{V}(n)\| < 10^{-8}$. This can be explained by comparing the dynamic equations of the UNIDANN and that of [6]. For the UNIDANN, there is an extra term $[DIAG\{\mathbf{V}^H(t)\mathbf{W}\mathbf{V}(t)\}]^{-1}$, which stands for a normalization term for each column of $[\mathbf{W}\mathbf{V}(t) - \mathbf{V}(t)TRI\{\mathbf{V}^H(t)\mathbf{W}\, V(t)\}]$ in each iteration. Thus, it is not as sensitive as that of [6] in choosing c. Therefore, for the UNIDANN, it is possible to use a larger c, which in turn offers faster convergent speed.

5. CONCLUSION

In this paper, we advance a novel neural network,UNIDANN, which can perform the unitary decomposition of the synaptic weight matrix. This new structure exhibits several advantageous features when compared with other existing ones. Since many high-resolution bearing estimation methods are based on the subspace concept and require the ED of the input correlation matrix, the developed UNIDANN becomes an attractive alternative to rendering real-time implementation of these methods.

6. REFERENCES

[1] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*. Prentice Hall, Inc.. Englewood Cliff, New Jersey, 1992.

[2] R. Rastogi, P. K. Gupta and R. Kumerason, "Array signal processing with interconnected neuron-like element," *Proc. ICASSP*, pp. 54.8.1–54.8.3, 1987.

[3] S. Jha and T. S. Durrani, "Direction of arrival estimation using artificial neural network," *IEEE Trans. Syst., Man, and Cybernetic*, vol. 21, no. 5, pp. 1192–1201, 1991.

[4] S. Y. Kung and K. I. Diamantaras, "A network learning algorithm for adaptive principal component extraction (APEX)," *Proc. ICASSP*, pp. 861–864, 1990.

[5] Y. Chen and C. Hou, "High resolution adaptive bearing estimation using a complex-weighted neural network," *Proc. ICASSP*, pp. II-317–II-320, 1992.

[6] F. L. Luo and Y. D. Li, "Real-time neural computation of the noise subspace for the MUSIC algorithm," *Proc. ICASSP*, pp. I485–I487, 1993.

[7] T. D. Sanger, "Optimal unsupervised learning in a single-layer linear feedforward neural network," *Neural Networks*, vlo.2, no. 6, pp. 459–473, 1989.

[8] G. H. Golub, and C. F. Van Loan, *Matrix Computations*, 2nd ed., Johns Hopkins University Press, Baltimore, Martland, 1989.

[9] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *Proc. RADC Spectrum Estimation Workshop*, Rome, NY, 1979.

[10] S.-H. Chang, T.-Y. Lee, and W.-H. Fang, "High-resolution bearing estimation via unitary decomposition artificial neural network (UNIDANN)," submitted to IEEE Trans. Syst., Man, and Cybernetic.

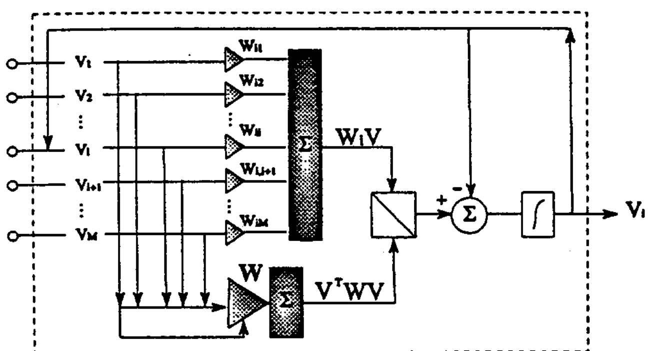

Figure 1: The neural structure of the UNIDANN.

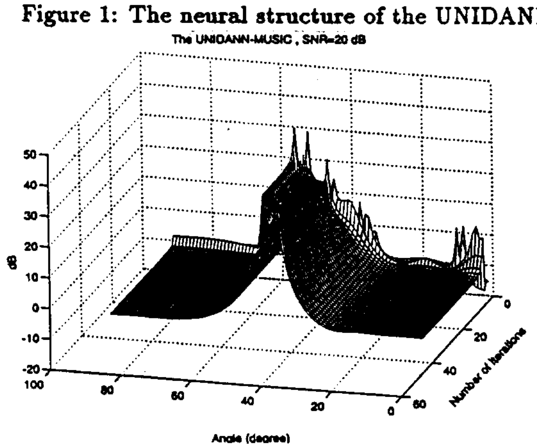

Figure 2: The pseudospectra of the MUSIC algorithm by using the UNIDANN.

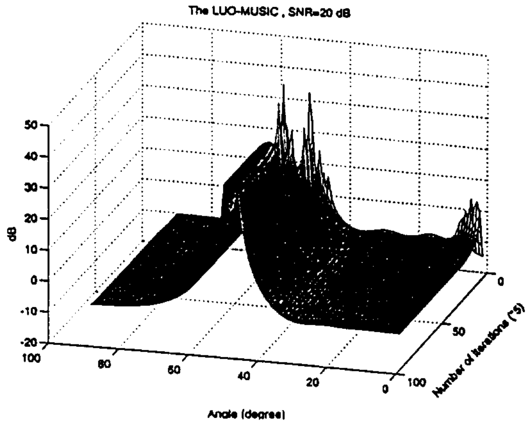

Figure 3: The pseudospectra of the MUSIC algorithm by using the one of [6].

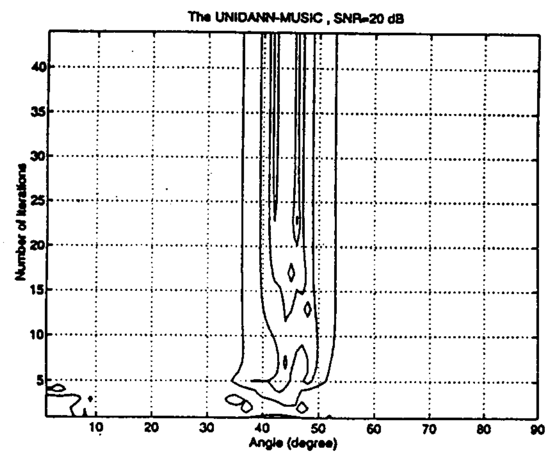

Figure 4: The corresponding contour plot by using the UNIDANN in Example 2.

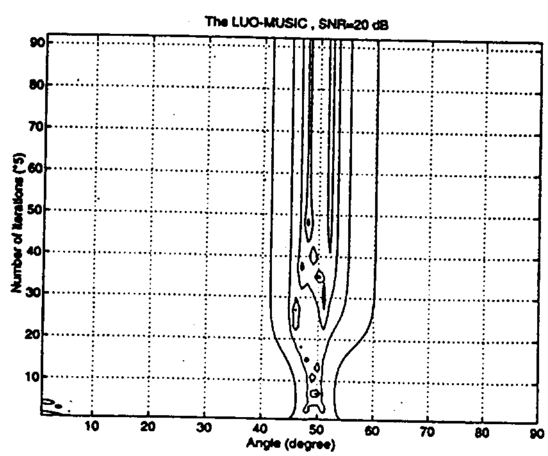

Figure 5: The corresponding contour plot by using the algorithm of [6] in Example 2.