

用于超角分辨的 Hopfield 神经网络方法

罗永光 梁甸农

(电子技术系)

摘要 Hopfield 神经网络用于方位估计可获得高的精度和超分辨力。本文综述有关研究,描述该网络模型及用于信号处理的一般步骤,用于方位估计的具体形式,讨论提高精度及分辨力的一些措施,指出进一步研究的方向。

关键词 雷达测角精度, 信号处理, 神经网络

分类号 TN953.5

方位角分辨问题是在非相关加性噪声情况下,利用传感器(天线)阵列获得的数据确定并分辨入射源的方向。普通的角分辨技术,采用直接求空间谱的方法,由于有瑞利限的限制,波束宽度不能很窄,难于分辨靠得很近的入射源。为此人们力图研究各种超分辨方法,其中模型参数法和特征结构法有较多的研究。但两者在短数据与低信噪比性能上各有所长,较难兼顾。本文另外讨论一种利用神经网络的方法,这种方法有着潜在的短数据、低信噪比性能。

八十年代, Hopfield 等人提出一种全互联神经网络,并赋予其能量函数,利用梯度法使该能量函数极小化,从而解决了一些长期未得解决的 TSP 等难题。Rastogi 等人^[1]在研究阵列信号处理时,将方位角估计问题映射成 Hopfield 神经网络的优化问题。继后又有几个学者进行了探索。

本文对有关的研究进行综述,也融合进作者的理解和看法。

1 Hopfield 神经网络模型

人工神经元可视为一多输入单输出的电子器件。图 1 是其一种典型的结构模型。其中 v_1, \dots, v_p 是神经元的输入, w_{11}, \dots, w_{ip} 是对各输入的传输权值, I_i 是偏置输入, $g(\cdot)$ 是由神经元内部状态 u_i 到其输出状态 v_i 的传输函数,在用于角分辨的 Hopfield 网络中一般取

$$v_i = g(u_i) = 1/[1 + \exp(-u_i/\lambda)] \quad (1)$$

Hopfield 神经网络如图 2 所示。图中每个圆圈代表一个神经元,每一个小黑点表示乘以一个权值。Hopfield 网络是一个全互联的对称的网络,如设第 j 个元到第 i 个元的权值为 w_{ij} , 则

$$w_{ij} = \begin{cases} w_{ij}, & i \neq j \\ 0 & i = j \end{cases} \quad (2)$$

令 $w_{ii} = 0$ 是为了避免各种神经元自反馈,以便形成自激振荡。

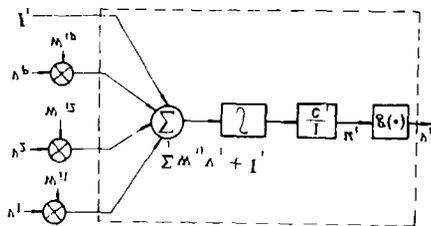


图 1 人工神经元模型

* 1991 年 2 月 28 日收稿

Hopfield 网络是一个演化式的动态网络，当演化到稳态后，各神经元的总加权输入为零，其内部状态 u_i 和输出状态 v_i 均为常数。在演化过程中则有（参考图 1）

$$C_i \frac{du_i}{dt} = \sum_{j=1}^p w_{ij} v_j + I_i \quad (3)$$

此即 Hopfield 神经网络的状态演化方程。Hopfield 为图 2 系统找到一个满足李雅普诺夫稳定性准则的能量函数：

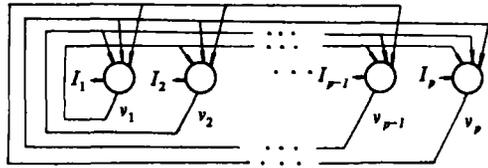


图 2 Hopfield 神经网络模型

$$E = 0.5 \sum_i v_i \sum_j w_{ij} v_j - \sum_i w_{ij} v_i I_i \quad (4)$$

由于 $w_{ij} = w_{ji}$ ，上式第一项是一个二次型。上式对 v_i 的微分为

$$\frac{dE}{dv_i} = -C_i \frac{du_i}{dt} \quad (5)$$

即 Hopfield 网络的演化方程是计算能量函数的负梯度公式。由链式微分法则可得能量函数的变化率为

$$\begin{aligned} \frac{dE}{dt} &= \sum_i \frac{dE}{dv_i} \frac{dv_i}{dt} = - \sum_i C_i \frac{du_i}{dt} \frac{dv_i}{dt} \\ &= - \sum_i C_i \frac{du_i}{dv_i} \left(\frac{dv_i}{dt} \right)^2 \end{aligned} \quad (6)$$

由于 $g(\cdot)$ 函数单调增，故 du_i/dv_i 恒大于零，故

$$\frac{dE}{dt} \leq 0, \text{ 且 } \frac{dE}{dt} = 0, \text{ 当且仅当 } \frac{dv_i}{dt} = 0, v = (v_1, v_2, \dots, v_p)^T \quad (7)$$

于是可得如下结论：Hopfield 网络演化时，其状态总是沿能量函数的负梯度方向变化，能量函数总是减小的。只有当网络达到稳态 $\left(\frac{dv_i}{dt} = 0\right)$ ，其能量函数才不再变化而停留在某一极小值上。网络的稳态总是与能量函数的极小相对应。

2 用 Hopfield 神经网络解决信号处理问题的一般步骤

利用 Hopfield 网络的演化是能量函数的极小化这一点，可以把性能函数极小化一类的信号处理问题化作 Hopfield 网络的演化来解决。为此必须：

- (1) 给问题找出一个合适的表达，使问题的解能够对应于神经元的输出；
- (2) 构造一个能量函数，使其极小值能够相应于问题的最佳解；
- (3) 由上述能量函数推导出连接权值 w_{ij} 和偏置 I_i ；
- (4) 将上述 w_{ij} 和 I_i 代入 Hopfield 网络的演化方程，使其进行状态演化（能量极小化）。

以上第 1 步是关键，即首先必须将待解决的问题映射成 Hopfield 神经网络问题。第 2 步是技术性很强的工作，选择的能量函数除其极小值能相应于最佳解之外，还要保证由其推出的 w_{ij} 和 I_i 有利于网络演化，利于避开局部极小陷阱。第 4 步工作可由计算机进行软件模拟，也可由实际电路进行硬件演化。

3 方位角估计的神经网络模型

3.1 P³ 模型

3.3.1 问题表述

以 A_i, ω_i, ϕ_i 表示第 i 个入射波源的复振幅、频率和入射角， d 为均匀线阵的阵元间距， c 为光速，则信号方位矢量可记为

$$S_i = A_i [1, r^{-j_1}, e^{-j_2}, \dots, e^{-j_{N-1}}]_i^T \quad (8)$$

$$r_i = (d/c)\omega_i \sin \phi_i \quad (9)$$

再设 n 为阵元噪声矢量, 则 M 个波源入射 N 元天线阵的数据矢量 $y = [y_1, y_2, \dots, y_N]^T$ 可记为

$$y = \sum_{i=1}^M S_i + n = [S'_1, S'_2, \dots, S'_M][1, 1, \dots, 1]^T + n \quad (10)$$

上述 M, A_i, ω_i, Φ_i 都是有待估计的未知量。为此作估计, 先将复振幅、频率、入射角分别进行 M_1, M_2, M_3 级量化, 在 $\{A, \omega, \Phi\}$ 三维空间中定出 $P = M_1 \cdot M_2 \cdot M_3$ 个刻度点, 并仿 (8)、(10) 两式构造出刻度方位矢量 S_j 和数据矢量的估计 \hat{y} 如下:

$$S_j = A_k [1, e^{-j\omega_i r_i}, e^{-j2\omega_i r_i}, \dots, e^{-j(N-1)\omega_i r_i}]^T \quad (11)$$

式中 r_i 如 (9) 式, $k=1, \dots, M_1, i=1, \dots, M_2 \cdot M_3$ 。

$$\hat{y} = [S_1, S_2, \dots, S_P][v_1, v_2, \dots, v_P] \quad (12)$$

式中, $v_i \in [0, 1], i=1, 2, \dots, P$ 。

$$\text{设法使估计方差 } Q = \|\hat{y} - y\|^2 \quad (13)$$

极小化, 即若 $S_{j'} \rightarrow S'_i$, 则 $v_{j'} \rightarrow 1$, 否则 $v_{j'} \rightarrow 0$ 。因为只有这样才有

$$\hat{y} - y \rightarrow n \quad (14)$$

从而 Q 逼近其最小值——噪声矢量 n 的方差。于是, 由哪些 $v_{j'}=1$, 就可推断有哪些 $S'_i=S_{j'}$ 的入射波。这里的 Q 和 $v_{j'}$ 可以与 Hopfield 网络的能量函数和神经元输出相比拟。

3.1.2 能量函数

为导出能量函数, 要将 Q 作些变化。令 $S = (S_1, S_2, \dots, S_P)$, 以 H 上标表示共轭转置, (13) 式可展开为

$$Q = y^H y + v^T S^H S v - 2R_c(y^H S v) \quad (15)$$

为使 v_i 收敛于 1 或收敛于零, 需在上式中加上一个惩罚项 $-\sum_{i=1}^P S_i^H S_i v_i (v_i - 1)$ 。 $y^H y$ 当数据给定后即成为常数, 与极小化无关, 可划去。再考虑到 $S^H S$ 是厄密对称阵, 最终使能量函数成为

$$Q = \text{Re} \left\{ \sum_{i=1}^P v_i \sum_{j=1}^P 2(S_i^H S_j) v_j - \sum_{i=1}^P (2y^H S_i - S_i^H S_i) v_i \right\} \quad (16)$$

3.1.3 权值与偏置

将 (16) 式与 (4) 式对比, 可得出网络的权值与偏置为

$$w_{ij} = \begin{cases} -4\text{Re}\{S_i^H S_j\}, & i \neq j \\ 0 & i = j \end{cases} \quad (17)$$

$$I_i = \text{Re}\{2y^H S_i - S_i^H S_i\} \quad (18)$$

3.1.4 状态演化

将以上 w_{ij} 和 I_i 代入 (3) 式即可进入网络的状态演化。软件演化时需将 (3) 式写成离散形式:

$$u_i(k+1) = u_i(k) + \frac{1}{C_i} \left[\sum_{j=1}^P w_{ij} v_j + I_i \right] \quad (19)$$

上述方法需大量神经元, 如 A, ω, Φ 中都作 p 级量化 ($M_1=M_2=M_3=p$), 则神经元总数 $P=p^3$, 故称为 p^3 模型。

3.2 p 模型

为使网络规模不致过大, 可将 A, ω 固定, 只考虑入射角估计, 这样可使神经元数由 p^3 压缩为 p (p 模型)。此时, 可令方位矢量 $S_j = (1, e^{-j\omega r_j}, e^{-j2\omega r_j}, \dots, e^{-j(N-1)\omega r_j})^T$

$$(20)$$

为消除复振幅与频率的失配, 应修改数据矢量的估计为

$$\hat{y} = (P_{S_1} y, P_{S_2} y, \dots, P_{S_P} y) v \quad (21)$$

其中各投影矩阵为

$$P_{S_i} = S_i (S_i^H S_i)^{-1} S_i^H \quad (22)$$

以下关于能量函数、偏置、权值的推导与前述 p^3 模型完全一样, 只需将 S_i 换成 $P_{S_i}y$ 。此时有

$$W_{ij} = \begin{cases} -4R_c \{y^H P_{S_i}^H P_{S_j} y\}, & i \neq j \\ 0, & i = j \end{cases} \quad (23)$$

$$I_i = R_c \{y^H P_{S_i} y\} \quad (24)$$

4 软件模拟

P 个节点(神经元)的 Hopfield 网络的输出状态空间是 $[0, 1]^P$, 如图 3 所示的超立方体。由于 v_i 最终取零或 1, 故 $[0, 1]^P$ 的每一个顶点都有可能是能量函数的极小值。网络总是从 $[0, 1]^P$ 中的某一初始状态点开始演化, 直到某一个顶点为止。但如该顶点不是全局极小点, 网络将陷于此点, 而达不到全局极小点, 从而造成估计误差。下面讨论避免陷于局部极小的方法。

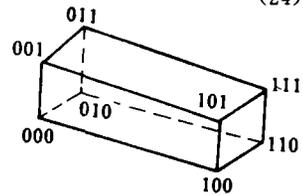


图 3 三个神经元的 Hopfield 网络的输出状态空间 $[0, 1]^3$

4.1 增益退火算法

Hopfield 网络的状态轨迹与其神经元的增益

$\left(\frac{1}{\lambda}\right)$ 有很大关系。如某一神经元的增益为 ∞ , 则其输出为数字式 (1 或 0), 网络状态轨迹将局限于状态立方体 $[0, 1]^P$ 的某相对两表面而不能驶过其内部。如果网络的全局极小顶点不在此两表面上, 则网络将陷入局部极小顶点而达不到全局极小顶点。如神经元增益低, 则状态轨迹可以驶过 $[0, 1]^P$ 内部而避开一些局部极小顶点。但如增益太低, 演化很慢, 达到稳态的时间可能会不能容忍地长。为此可在演化之初, 能量函数梯度大时降低增益, 使演化不致过快而误入陷阱。而在演化后期能量函数梯度很小时再加大增益, 使演化加速进行而尽快达到全局极小。这就是增益退火。为此有两种办法: 一是使 λ 值递减变化, 例如使^[3]

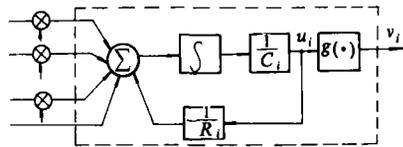


图 4 隐含增益退火的神经元模型

二是修改神经元模型为图 4 形式, 其中的负反馈起反相调节能量函数梯度的作用, 其演化方程为

$$\lambda = 10 \cdot (0.85)^k \quad (25)$$

$$C_i \frac{du_i}{dt} = \sum_j w_{ij} v_j + I_i - u_i / R_i \quad (26)$$

4.2 迭代下降算法

这种方法是当网络演化到极小点就随机地选取部分神经元, 颠倒其输出状态, 以此作为初态重新演化。颠倒状态相当于给能量函数一个增量。后续各次演化颠倒的状态越来越少, 能量函数逐渐达到全局极小。其思想是: 一, 按 Hopfield 的研究, 若

极小点与初始点的 Hamming 距离在某值以上再增加, 则收敛于该极小点的概率将下降, 颠倒状态形成新的初始点, 可增加某些极小点与初始点的距离以降低落入局部极小陷阱的概率; 二、反复独立地以随机初始点开始演化达于稳态, 按中心极限定理, 其规律是正态的, 有其均值 (全局极小值点) 和方差。若保留部分状态, 只颠倒部分状态, 则进行非独立试验, 其逼近均值的概率逐渐增加, 而试验方差越来越小, 可望得到更合理的解。

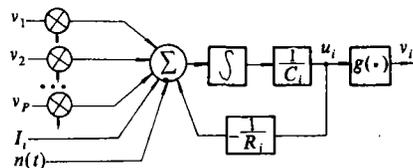


图 5 随机网络算法的神经元模型

4.3 随机网络算法

上述增益退火使状态轨迹绕过一些局部极小陷阱, 迭代下降算法是一旦落入陷阱, 就进行人工干预

使其脱离。随机网络算法则是给能量函数梯度以随机扰动,使状态轨迹在原基础上作随机摆动,即使落入陷阱也可不需人工干预而自行跳出。其演化方程为

$$C_i \frac{du_i}{dt} = -u_i/R_i + \sum_j w_{ij}v_j + I_i + n(t) \quad (27)$$

式中 $n(t)$ 是随机噪声。这又相当于图 5 所示的神经元模型。

5 多快拍形式

在上述内容只需用单次快拍数据。当输入快拍数据多时,可对各次快拍的估计误差进行时间平均以作为能量函数,重新推导演化算法。以 l 表示快拍序号,则能量函数、权值、偏置各具有如下的时间平均形式:

$$Q = \frac{1}{L} \sum_{l=1}^L R_s \left\{ \sum_{i=1}^p \sum_{\substack{j=1 \\ j \neq i}}^p y^H(l) P_{s_i, y}(l) v_i v_j - \frac{1}{2} \sum_{i=1}^p y^H(l) P_{s_i, y}(l) v_i \right\} \quad (28)$$

$$w_{ij} = -4R_s \left\{ \frac{1}{L} \sum_{l=1}^L y^H(l) P_{s_i}^H P_{s_j, y}(l) \right\} \quad (29)$$

$$I_i = R_s \left\{ \frac{1}{L} \sum_{l=1}^L y^H(l) P_{s_i, y}(l) \right\} \quad (30)$$

6 硬件实现

Hopfield 网络可以用电子电路来实现。Hopfield 本人发明了图 6 所示电路实现了网络演化。该电路对应 (26) 式的演化方程。输入电容 C_i 形成积分。输入电阻 R_i 形成负反馈,以电流作输入,电压作输出。 I_i 作偏置电流, w_{ij} 为运放间的跨导,以运放的饱和非线性作神经元的非线性特性。

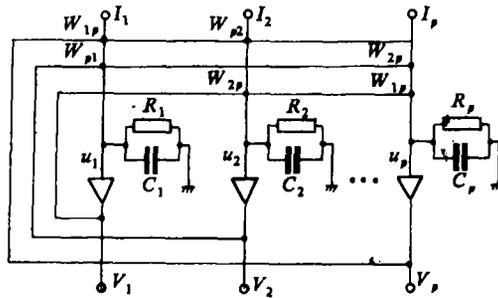


图 6 Hopfield 网络的硬件实现形式

这种电路演化稳定时间约为 $5R_i C_i$ 。如取 $R_i = 10\text{K}\Omega$, $C_i = 10\text{PF}$, 则在 500ns 时间内可完成演化。

7 精度与分辨力

Hopfield 网络是非线性网络,对其精度与分辨力进行公式化的定量分析是困难的,目前尚处于计算仿真阶段。以下是有关的一些仿真结果。

天线阵元数 $N=25$, 输入快拍数 $L=1$, 神经元数 $P=200$, 归一化频率量化值 $\Delta f = 0.00125, 0.35125 \leq f \leq 0.06$, SNR 从 10dB 降到 -5dB 时,采用普通算法的 Hopfield 网络,经 20 次独立仿真试验,对于 $f = 0.525$ 的平均估计误差 ≤ 0.002187 , 方差 ≤ 0.000504 。平均偏差不到两个神经元。在 SNR $= -10\text{dB}$ 时,误差增大到 $0.025062^{[1]}$ 。

当 $L=1$, $N=25$, $P=100$, $\Delta f = 0.01$, $0 \leq f \leq 0.99$, SNR $\approx 1\text{dB}$ 时,采用普通算法的 Hopfield 网络的分辨力仅 $0.06^{[1]}$,尚说不上超分辨。

当 $L=5$, $N=20$, $P=200$, 角度量化值 $\Delta\phi = 0.25^\circ$, $0^\circ \leq \phi \leq 50^\circ$, SNR $= 20\text{dB}$ 时,采用多快拍时间平均算法可成功地进行 33° 和 37° 两窄带信号的分辨。当 $L=5$, $N=20$, $P=50$, $\Delta\phi = 0.4^\circ$, $20^\circ \leq \phi \leq 40^\circ$

时,采用划分成 10 个小频带的频域和时域平均算法,还可在 SNR=0dB 下成功地对 28°和 32°、相对带宽为 0.67 的两宽带信号进行分辨^[4],此二例分辨力均为 4°,已突破 $N=20$ 所对应的 6.03°的瑞利限。

从以上数据可看出,Hopfield 网络有着良好的短数据和低信噪比性能,有超分辨能力。但目前仿真数据还收集得不充分,难于得出精度和分辨力与算法、网络节点数、天线阵元数、信噪比、数据快拍数等诸因素的详细关系。

8 小 结

使用 Hopfield 网络可以对入射波束进行估计与分辨,其方法是将估计误差比拟为神经网络的能量函数,将估计问题映射成神经网络在相应权值及偏置下的演化。这种方法类似于将估计问题转化为检测问题,以每一个神经元作一个方位检测器。这种方法可用于天线阵元数少和快拍数少的短数据情况,在低信噪比下也可以有较好的估计精度和超分辨性能。同时,Hopfield 网络的硬件运行有极快的实时性,因此值得对这种超角分辨方法进行更深入的研究。这种研究包括理论分析、计算机仿真及硬件试验,深入分析网络节点数、天线阵元数、数据快拍数、信噪比、算法等多种因素对精度和分辨力的影响,探索进一步提高精度和分辨力的算法,详细得出精度与分辨力对以上诸因素的依赖关系的仿真数据和硬件试验结果,使这种技术更加成熟,更合实用。

参 考 文 献

- 1 Rastogi R, et al. Array Signal Processing with Interconnected Neuron-like Elements. Proc. ICASSP, 1987; 2328~2331
- 2 Jha. S, et al. Bearing Estimation Using Neural Networks Proc ICASSP, 1988; 2156~2159
- 3 Jha. S, et al. Bearing Estimation Using Neural Optimisation Methods, Proc ICASSP, 1990; 889~892
- 4 Goryn. D, et al. Neural Networks for Narrowband and Wideband Direction Finding. Proc ICASSP, 1988; 2164~2167
- 5 Hopfield JJ, et al. Neural Computation of Decisions in Optimization Problems, Biological Cybernetics, 1985, 52; 141~152

The Hopfield's Neural Network Method for Superresolution Estimation of the Direction of Arrival

Luo Yongguang Liang Diannong
(Department of Electronic Technology)

Abstract

High accuracy and superresolution of direction of arrival (DOA) can be achieved by means of the Hopfield's neural network (HNN). In this paper a survey for related studies of this field is presented. The HNN's model, the common steps of signal processing with the HNN and the concrete form of estimation of DOA with the HNN are described. Some methods of raising accuracy and improving resolution for it are discussed. The topics for further study are pointed out.

Key words rader angle measurement accuracy, signal processing, neural networks