

ECR Plasma CVD 淀积介质膜 折射率的神经网络模拟*

李玉鉴 谭满清 茅冬生 陆建祖

(中国科学院半导体研究所 神经网络组 光电子器件国家工程研究中心 北京 100083)

摘要 用人工神经网络方法对电子回旋共振等离子法化学气相沉积 (ECR Plasma CVD) 镀膜工艺建立了一个介质膜折射率 n 关于气流配比 $Q(N_2)/Q(SiH_4)$ 和 $Q(O_2)/Q(SiH_4)$ 的数学模型。在给定气流配比 $Q(N_2)/Q(SiH_4)$ 和 $Q(O_2)/Q(SiH_4)$ 时模型预测的成膜折射率与实验值符合得很好。

PACC: 0650, 7125T, 6590

1 引言

电子回旋共振等离子法化学气相沉积 (ECR Plasma CVD) 是一种淀积介质膜工艺, 其优点表现在^[1-3]: (1) 它是一种低温淀积技术, 淀积温度可控制在 150 °C 以下; (2) 它可通过改变气流成分和气流配比方便地实现不同的折射率膜 (SiN_x , SiO_xN_y 等); (3) 淀积的膜折射率 n 的稳定性好, 且在其它工艺条件固定的前提下, n 值可通过各种气体流量间的比例得到严格控制。虽然已经积累了大量 ECR Plasma CVD 淀积介质膜的实验数据, 但是由于工艺过程的复杂性, 目前还没有从理论上对这些实验数据建立起一个良好的数学模型。本文试图利用人工神经网络 (Artificial Neural Network, ANN) 这一新兴的建模方法, 对 ECR Plasma CVD 工艺条件 (气流配比) 和介质膜折射率的关系进行分析 (神经网络在半导体材料中的应用在国内已有报道^[4])。通过学习和训练, ANN 可以在大量的实验数据之间建立起各种复杂的函数关系。理论上已经证明, 三层 BP 网络能够以任意精度逼近任意单变量或多变量连续函数^[5-8], 因此可以作为一种有效的工具来提取各种实验数据间的内在规律性, 尤其是在这些实验数据还难于进行理论分析时, 它就能发挥其潜在的价值。本文首先用 BP 网络建立起介质膜折射率 n 关于气流配比 $Q(N_2)/Q(SiH_4)$ 和 $Q(O_2)/Q(SiH_4)$ 的数学模型, 然后利用这一模型对给定气流配比的介质膜折射率进行预测, 预测结果和实验结果符合得很好, 这将为今后对 ECR Plasma CVD 镀膜设备开发计算机辅助设计 (TCAD) 软件提供了

* 国家“863”高技术计划项目 307-06-04(01) 资助课题

李玉鉴 男, 1968 年出生, 博士生, 主要研究兴趣是人工智能和神经网络的模型、算法及其应用
1999-03-05 收到, 1999-05-22 定稿

令人满意的模型基础

2 BP 算法及其改进

Rumelhart^[9]等人于 1986 年提出反向传播算法来训练多层网络,即通常所说的BP 网络 图 1 是一个三层BP 网络,包括输入层、输出层和一个隐层,对于输入信号,要先向前传播到隐层,再把隐层的输出信息传播到输出节点,最后输出结果 神经元的作用通常选取带阈值的 S 型激活函数来描述,如:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

在图 1 中,各神经元的输出 y 由前层神经元传递给它的加权输入和它本身的激活函数所决定,即: $y = f(\sum w_i x_i - \theta)$.

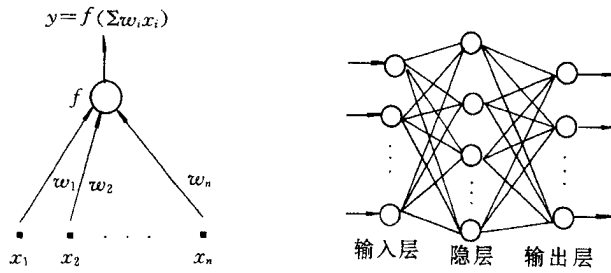


图 1 神经元的作用和三层BP 网络的结构

反向传播算法的基本精神是:从实际输出与期望输出之间的偏差出发,沿网络内信号传输的相反方向逐层推算出每一层输出的相应偏差,据此调整各层单元之间的连接权重及每个神经元的域值,使网络趋向稳定达到所要求的性能 在实际应用时,通常输入神经元的激活函数取 $f(x) = x$,隐层和输出层神经元的激活函数取(1)式表达的 Sigmoid 函数,但这种网络要求作为输出的数据必须在区间(0, 1)取值,因此就需要对实验数据进行归一化处理,从而带来一些额外的麻烦 为了克服这一缺点,本文引入了可变激活函数的概念,即给(1)式的激活函数增加几个参数,比如将其改写为:

$$f(x) = a + \frac{b}{1 + e^{-c(x-d)}} \tag{2}$$

其中 a, b, c, d 是可调参数,它们在学习训练过程中随着权值矩阵一同变化 引入可变激活函数后,网络的学习训练方法要作适当改变,但基本思想仍是梯度下降法,即将实际输出和标准输出之间的积累误差 E 看作权值 w (w 表式某两个神经元之间的连接权值)和某神经元内部参数 a, b, c, d 的函数 $E(w, a, b, c, d)$,从而容易得到:

$$w = w - \frac{\partial E}{\partial w} \eta$$

$$a = a - \frac{\partial E}{\partial a} \eta \quad b = b - \frac{\partial E}{\partial b} \eta \quad c = c - \frac{\partial E}{\partial c} \eta \quad d = d - \frac{\partial E}{\partial d} \eta$$

适当选取 η 就可保证积累误差 E 在训练时不断减小,等到 E 值的下降非常缓慢时,即可停止训练

另外, 由于传统BP 算法收敛非常缓慢, 本文还提出了下面的步长估计方法:

将误差函数 $E(w)$ 在 w (此时 w 应看作所有权值和所有内部参数构成的向量) 点沿梯度方向按步长进行一阶泰勒展开,

$$E(w - \bar{\nabla}h) = E(w) - (\bar{\nabla})^T \cdot \bar{\nabla}h + o(h)$$

其中 $\bar{\nabla}$ 是 E 在 w 点梯度 对于一个较小的正数 $\alpha(0 < \alpha < 1$, 一般可取 $\alpha = 0.05$ 或 0.1), 可用方程:

$$(1 - \alpha)E(w) = E(w) - (\bar{\nabla})^T \cdot \bar{\nabla}h \quad \text{即 } h = \frac{\alpha E(w)}{(\bar{\nabla})^T \cdot \bar{\nabla}}$$

来估算步长 h , 使BP 算法的每步迭代保证误差大约下降 $\alpha E(w)$, α 称为下降因子. 这种方法虽然非常简单, 但是却能有效地加快BP 算法收敛速度

3 气流配比 $Q(N_2)/Q(SiH_4)$ 和介质膜折射率 n 的神经网络模型

3.1 网络结构和拟合结果

本文选取两输入单输出及单隐层(包含 5 个神经元), 即 2-5-1 的BP 网络结构来建立介质膜折射率 n 关于气流配比 $Q(N_2)/Q(SiH_4)$ 和 $Q(O_2)/Q(SiH_4)$ 的函数关系 网络的输入是气流配比 $Q(N_2)/Q(SiH_4)$ 和 $Q(O_2)/Q(SiH_4)$, 输出是介质膜折射率 n , 输入神经元的激活函数取 $f(x) = x$, 隐层神经元的激活函数是 (1) 式表达的 Sigmoid 函数, 输出神经元的激活函数是 (2) 式表达的 Sigmoid 函数, 网络的权值用上节提出的改进BP 算法训练, 终止条件是拟合结果与实验结果的均方根误差 RMSE < 0.046 训练样本集和拟合结果见表 1, 模型曲面及曲线见图 2

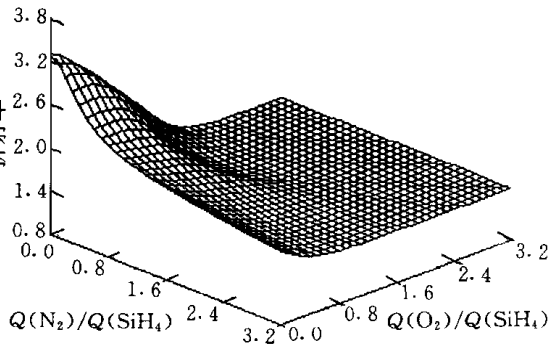


图2 折射率关于气流配比 $Q(N_2)/Q(SiH_4)$ 和 $Q(O_2)/Q(SiH_4)$ 模型曲面

3.2 预测结果和实验结果的比较

为了说明用神经网络模型建立起来的介质膜折射率 n 关于气流配比 $Q(N_2)/Q(SiH_4)$ 和 $Q(O_2)/Q(SiH_4)$ 函数关系的正确性, 有必要对模型在给定气流配比 $Q(N_2)/Q(SiH_4)$ 和 $Q(O_2)/Q(SiH_4)$ 的前提下所预测的折射率 n 值进行实验验证 表 2 给出了折射率的预测值和实验值的比较结果, 从中可以看出, 介质膜折射率 n 的神经网络预测值和实验值符合得很好.

表 1 训练样本集和网络对样本输入的计算结果

气流配比		折射率	
$Q(N_2)/Q(SiH_4)$	$Q(O_2)/Q(SiH_4)$	实验结果	拟合结果
0.000	0.000	3.277	3.281
0.727	0.000	2.342	2.368
0.545	0.000	2.597	2.532
0.364	0.000	2.784	2.780
0.909	0.000	2.267	2.262
1.636	0.000	1.994	2.036
2.545	0.000	1.884	1.880
1.091	0.000	2.228	2.186
3.273	0.000	1.830	1.798
1.455	1.455	1.589	1.577
1.818	1.091	1.639	1.631
1.818	0.909	1.673	1.661
1.818	0.727	1.708	1.696
1.818	0.545	1.740	1.742
1.818	0.364	1.780	1.806
1.818	0.182	1.820	1.891
0.909	2.545	1.500	1.501
0.000	2.000	2.989	3.003
1.455	0.000	2.068	2.078
0.000	0.727	2.600	2.649
0.455	0.000	2.580	2.645
1.182	0.000	2.221	2.155
0.909	1.710	1.553	1.522
0.509	1.710	1.550	1.561
1.091	1.310	1.608	1.618
0.727	1.310	1.633	1.644
0.455	0.909	2.191	2.178
1.091	0.545	2.071	2.057
0.509	0.509	2.482	2.521
0.236	0.509	2.694	2.792
0.236	0.236	2.963	2.961
0.000	1.636	1.636	1.710
0.236	0.909	2.344	2.340
0.800	0.400	2.293	2.273
0.800	0.727	2.123	2.144
1.455	2.000	1.532	1.502
0.000	1.000	2.368	2.377
0.000	1.364	2.044	1.964
0.000	0.545	2.820	2.825
0.545	1.309	1.677	1.711

表 2 成膜折射率 n 的预测值和实验值的比较

气流配比		折射率	
$Q(N_2)/Q(SiH_4)$	$Q(O_2)/Q(SiH_4)$	预测值	实验值
1.300	0.000	2.117	2.063
0.667	2.100	1.509	1.504
1.053	2.316	1.501	1.500
2.667	0.000	1.863	1.885
1.000	0.000	2.221	2.185
1.053	2.316	1.501	1.500
2.220	0.444	1.686	1.730
0.711	0.667	2.242	2.165
3.000	3.000	1.501	1.500

4 讨论

目前,采用电子回旋共振等离子法化学气相沉积(ECR Plasma CVD)工艺形成电介质膜的报道越来越多,这种方法已成功地在 InP 等化合物半导体衬底上淀积了 a-Si, SiN₄, SiN_x, SiO_xN_y 和 SiO₂ 介质膜。膜的折射率 n 在 3.2~ 1.48 范围内可通过 N₂/O₂/Ar 气流比率严格控制。在一定的条件下,研究各种气流配比与介质膜折射率的关系,无疑对开发 ECR Plasma CVD 镀膜工艺的计算机辅助设计软件(TCAD)具有重要意义。本文应用人工神经网络方法,成功地建立了介质膜折射率 n 关于气流配比 $Q(N_2)/Q(SiH_4)$ 和 $Q(O_2)/Q(SiH_4)$ 的数学模型,显然,这一方法可以类比地用来研究介质膜折射率与其它工艺条件的数学关系,甚至也可用来研究长膜速率与相关工艺条件的数学关系。将所有这些模型综合起来,就可能为 ECR Plasma CVD 设备开发一个功能齐全的计算机辅助设计软件,从而缩短镀膜工艺流程,并尽快生产出符合要求的镀膜材料。

5 结论

本文用人工神经网络方法对 ECR Plasma CVD 镀膜设备成功地建立了一个介质膜折射率 n 关于气流配比 $Q(N_2)/Q(SiH_4)$ 和 $Q(O_2)/Q(SiH_4)$ 的数学模型,由于此模型在给定气流配比 $Q(N_2)/Q(SiH_4)$ 和 $Q(O_2)/Q(SiH_4)$ 时所预测的介质膜折射率跟实验值符合得很好,所以它可用来作为开发 ECR Plasma CVD 镀膜设备计算机辅助设计软件的一个基本模型。

参 考 文 献

- [1] S. Dzioba and R. Rousina, J. Vac. Sci. Technol., 1994, B12(1): 433~ 440
- [2] Y. Manabe and T. Mitsuyu, J. Appl. Phys., 1989, 60(4): 2475~ 2480
- [3] 谭满清,等, ECR Plasma CVD 法淀积介质光学膜的研究, 1997 年砷化镓及有关化合物会议论文集, 310~ 312, 1997.
- [4] 张兆春,等, 半导体学报, 1998, 19(1): 49~ 53
- [5] K. Hornik, M. Stinchcombe and H. White, Neural Networks, 1989, 2: 359~ 366

- [6] K. Hornik, *Neural Networks*, 1991, 4: 251~ 257.
- [7] E. K. Blum and L. K. Li, *Neural Networks*, 1991, 4: 511~ 515.
- [8] T. Chen and H. Chen, *IEEE Trans Neural Networks*, 1995, 6(4): 911~ 917.
- [9] Rumelhart D. E., Hinton G. E. and Williams R. J., Learning internal representations by error back propagation. In D. E. Rumelhart and J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol 1: Foundations* (pp. 318~ 362). Cambridge, MA: MIT Press.

Neural-Network Simulation of Refractive Index of Dielectric Thin Film Deposition by ECR Plasma CVD

Li Yujian, Tan Manqing, Mao Dongsheng, Lu Jianzu

(*Institute of Semiconductors, The Chinese Academy of Sciences, Beijing 100083*)

Received 5 March 1999, revised manuscript 22 May 1999

Abstract By artificial neural network, a mathematical model has been established for the relation between refractive index and gas-flow ratio $Q(\text{N}_2)/Q(\text{SiH}_4)$ as well as $Q(\text{O}_2)/Q(\text{SiH}_4)$, and based on this model, a good prediction has been made for refractive index n when gas-flow ratio $Q(\text{N}_2)/Q(\text{SiH}_4)$ and $Q(\text{O}_2)/Q(\text{SiH}_4)$ are given.

PACC: 0650, 7125T, 6590