

文章编号: 1004-0609(2005)01-0105-05

基于分子动力学结合神经网络的 Au 表面能计算方法^①

汤 伟, 朱定一, 陈丽娟, 关翔锋
(福州大学 材料科学与工程学院, 福州 350002)

摘 要: 利用嵌入原子模型, 采用分子动力学方法计算了贵金属 Au 低指数晶面及部分简单高指数晶面的表面能。同时, 采用 Levenberg-Marquardt 算法, 建立了 Au 表面能的 BP 神经网络模型; 结合分子动力学模型的计算数据, 通过大量数据的自学习训练, 完成神经网络模型对 Au 高指数晶面表面能的预测。计算结果表明: 该方法具有较高的预测精度, 能正确预言低指数晶面表面能的排序, Au 各晶面的表面能随其晶面与(111)密排面夹角的增大呈现先增大后减小的特点。

关键词: 表面能; 嵌入原子势; 人工神经网络; Levenberg-Marquardt 算法

中图分类号: TG 146.31; TP 183

文献标识码: A

Calculation of Au surface energy by molecular dynamics combined with neural networks

TANG Wei, ZHU Ding-yi, CHEN Li-juan, GUAN Xiang-feng

(School of Materials Science and Engineering, Fuzhou University, Fuzhou 350002, China)

Abstract: Via embedded-atom model and molecular dynamics simulation, the surface energies of three low-index and some high-index planes were calculated for precious metal Au, and the error back-propagation network (BP) developed by Levenberg-Marquardt algorithm was adopted. Combining the data calculated with the molecular dynamics model, a great deal of data were trained many times and compared with the calculated data, and the prediction of high-index surface energy was performed. The results show that the method has high predicting accuracy. The order of the three low-index planes was predicted exactly. The surface energies on the other planes show a tendency that first increasing and then decreasing with angle between the planes and (111) plane increasing.

Key words: surface energy; embedded-atom potential; artificial neural network; Levenberg-Marquardt algorithm

表面能是表征材料表面特性的重要参量, 在表面、强度与断裂、相变等很多材料问题的研究中有广泛应用。由于晶体在结构上的各向异性, 不同的晶体表面具有不同的表面能, 其物理属性也有很大差异。Au 在电子工业和尖端科学技术领域是极其重要的功能材料, 在电子工业中 90% 用于表面镀层。镀金可提高材料的耐腐蚀性, 镀金的另一个作用是对红外有强烈的反射作用, 广泛用于红外探测仪反导弹装置上, Au 已成为现代卫星通讯仪器中

的关键材料^[1]。近几年, Au 的表面催化作用得到广泛的应用^[2]。因此研究 Au 的表面能具有重要意义。目前很难通过实验精确测定固态材料的表面能, 即使通过实验能获得一些简单结构晶体的表面能, 但误差一般都比较大大。现在已能够由第一原理方法在超级计算机上直接计算一些简单结构晶体的表面能, 但需要花费大量机时, 同时也存在一定的误差^[3]。若想通过该方法计算实际条件下各种复杂结构晶体的表面能, 目前计算机的运算速度和存储

① 基金项目: 国家自然科学基金资助项目(50471007); 福建省自然科学基金资助项目(A0210008)

收稿日期: 2004-05-25; 修订日期: 2004-11-07

作者简介: 汤 伟(1981-), 男, 硕士研究生。

通讯作者: 朱定一, 副教授; 电话: 0591-87893540; E-mail: zdy7081@163.com

能力仍远远不够。在这种情况下,利用近年来迅速发展的各种多体理论,并结合分子动力学方法进行计算机模拟计算材料的表面能,可能是目前确定固态材料表面能的有效方法,它既有一定的准确性,又对计算机要求不太高,本文作者采用嵌入原子模型对贵金属 Au 的低指数及部分高指数晶面的表面能进行分子动力学模拟计算,但由于复杂高指数晶面表面能用分子动力学方法计算困难,目前鲜有报道,本文作者探索性地采用人工神经网络方法,结合分子动力学模拟的计算数据,通过大量数据的自学习训练,完成神经网络模型对高指数晶面表面能的预测,并与实验结果进行比较。

1 嵌入原子模型及参数的确定

EAM 理论^[4]是在密度泛函理论基础上,运用有效介质或准原子近似推导出来的,它包括两个主要的假设:一是原子电子密度分布的球面平均假设;二是基体电子密度是原子的电子密度的线性迭加。根据 EAM 理论,一个原子集团的总能量可表示为

$$E_{\text{tot}} = \sum_i F_i(\rho) + \frac{1}{2} \sum_{i,j} \phi_{i,j}(r_{i,j}) \quad (1)$$

$$\rho = \sum_{j \neq i} f(r_{ij}) \quad (2)$$

式中 $F_i(\rho)$ 为嵌入能; $\phi_{i,j}(r_{i,j})$ 为对势项,根据需要可以取不同的形式; ρ 为除第 i 个原子以外的所有其他原子的核外电子在第 i 个原子处产生的电子云密度之和; $f(r_{ij})$ 为原子电子密度分布函数; r_{ij} 为两原子间的间距。

要得到一个嵌入原子模型,须确定上述函数的形式,根据 Johnson^[5] 提出模拟计算方法:

$$F(\rho) = -E_c(1 - \ln x)x - 6\phi_e y \quad (3)$$

$$f(r) = f_e \exp[-\beta(r/r_e - 1)] \quad (4)$$

$$\phi(r) = \phi_e \exp[-\gamma(r/r_e - 1)] \quad (5)$$

式中 $x = (\rho/\rho_e)^{\alpha\beta}$; $y = (\rho/\rho_e)^{\gamma\beta}$; $\alpha = 3(\Omega B/E_c)^{1/2}$; $f_e = E_c/\Omega$; $\phi_e = E_c/6$; 下标 e 表示平衡状态时的值。

通过拟合晶格常数 a 、原子体积 Ω 、内聚能 E_c 、单空位形成能 E_{UF} 、体积模量 B 和剪切模量 G 等物性参数和通过拟合获得的模型参数列于表 1。

2 人工神经网络原理及模型

2.1 人工神经网络原理

人工神经网络(ANN)是目前国际上迅速发展

表 1 贵金属 Au 的物理参数和模型参数^[5]

Table 1 Physical parameters and model parameters of precious metal Au

Physical parameters					
a/nm	Ω/nm^3	E_c/eV	E_{UF}/eV	B/GPa	G/GPa
0.408	0.016 98	3.93	0.90	171	31.0
Model parameters					
f_e	ϕ_e/eV	α	β	γ	
0.23	0.65	6.37	6.67	8.20	

的前沿交叉学科,其本质是建立一种映射关系,在工程应用中属于典型的数据挖掘技术。人工神经网络以样本数据为基础,按照某种算法经过有限次迭代获得一个反映样本数据内在规律的数学模型,因此具有适合于研究复杂非线性系统和不确定过程的优点,具有广泛的应用前景^[6~8]。由于误差反向传播神经网络(BP 网络)具有较强的联想记忆和泛化能力,本文选用 BP 网络。BP 网络通过输入数据正向传播和输出误差反向传播两个过程,修正各层联接权值和阈值,直至输出达到或逼近所要求的响应为止。具有一个输入层、一个隐含层、一个输出层的 3 层 BP 神经网络如图 1 所示(图中, P 代表输入层矩阵; w_1 、 w_2 分别代表隐含层和输出层的权值矩阵; b_1 、 b_2 分别代表隐含层和输出层的偏差向量(阈值); a_1 、 a_2 分别代表隐含层和输出层的输出矩阵; n_1 、 n_2 分别代表隐含层和输出层神经元的输出矩阵)。经典 BP 算法是采用最速下降静态寻优的一阶算法,其在目标函数的极值点附近收敛速度慢,而经典二阶算法(如高斯-牛顿法)在远离目标函数的极值点时收敛速度很慢^[8]。本研究采用一种拟二阶算法——Levenberg-Marquardt 算法^[9],其基本迭代公式为

$$\Delta w = (J^T J + \mu I)^{-1} J^T e \quad (6)$$

式中 J 为误差函数对权值向量微分的 Jacobian 矩阵; e 为误差向量, μ 是某个可以进行自适应调整的非负数。该修正方法依赖于 μ 的幅值的改变,光滑地在两种极端之间变化:即二阶的高斯-牛顿法(μ

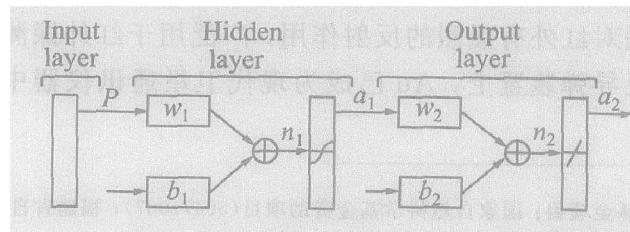


图 1 三层 BP 神经网络

Fig. 1 Three layers BP neural network

→0) 和经典的梯度法 ($\mu \rightarrow \infty$), 充分发挥了两种算法的优势^[10]。

2.2 表面能预测神经网络模型

2.2.1 BP 神经网络模型的建立

将各晶面与 3 个低指数晶面 ((100)、(110) 和 (111)) 的夹角作为神经网络训练的输入矢量, 根据立方晶系两个晶面 ($h_1k_1l_1$) 和 ($h_2k_2l_2$) 间的夹角公式:

$$\cos \theta = \frac{h_1h_2 + k_1k_2 + l_1l_2}{\sqrt{(h_1^2 + k_1^2 + l_1^2) + (h_2^2 + k_2^2 + l_2^2)}} \quad (7)$$

可求得各晶面与 (100)、(110) 和 (111) 3 个低指数晶面的夹角, 将所对应的根据分子动力学模拟计算的表面能值作为输出矢量, 其结果均列入表 2 中。

由于 BP 神经网络选用 Sigmoid 函数作为网络神经元的传递函数。根据 S 函数的特征, 为了保证网络神经元的非线性作用, 对原始数据要进行标准化处理, 标准化处理有助于提高网络的学习速度。为了将样本输入和输出标准化, 使它们的均值为 0, 而标准差为 1, 作者对样本进行如下标准化处理:

$$x_p = \frac{x_p - \bar{x}}{\sigma} \quad (8)$$

式中 x_p 为样本数据, 平均值 $\bar{x} = \frac{1}{p} \sum_{p=1}^n x_p$, 标准

$$\sigma = \sqrt{\frac{1}{p-1} \sum_{p=1}^n (x_p - \bar{x})^2}。$$

Hecht-Nielson^[11] 证明了对于任何在闭区间内的一个连续函数都可以用一个单隐层的 BP 网络来逼近, 一个三层的 BP 网络可以完成任意的 n 维到 m 维的映射。网络结构选用具有一个隐含层的 BP 网络, 隐含层采用 S 型激活函数, 输出层采用线性激活函数。

2.2.2 BP 网络神经元数目的选择

隐含层神经元的数目是决定神经网络有效性和准确性的关键因素之一, 一般根据问题的复杂程度和目标的精度选择最适合的神经元数目。因此, 首先要设定隐含层节点数和误差限, 在训练神经网络的过程中不断调节隐含层节点数目、训练次数和训练误差限, 直至找到最佳预测结果。通过多次训练和调节, 综合考虑网络的学习速度和泛化能力, 最终确定 6 个节点的隐含层, 网络的拓扑结构为 3-6-1。

3 表面能的计算与讨论

将上述网络模型在 MatLab6.1 中编制成可运

行程序, 并将标准化后的夹角数据作为输入矢量, 表面能作为输出矢量分别代入网络模型进行训练。用训练所得到的权值, 对表面能进行了预测, 同时将 BP 神经网络预测相对误差值列入表 2 中。由表 2 可知, 所训练的网络对表面能预测的误差范围是比较稳定和合理的。图 2 所示为表面能的神经网络预测结果与 MD 计算结果的比较, 并给出预测表面能和计算表面能的线性相关曲线。从图可以看出, 预测值和 MD 计算值有较高的线性相关性, 线性相关系数 $R = 0.999\ 65$ 。由此可知, 该预测模型具有高的预测精度。平均相对误差 $\bar{\varepsilon}$ 可以用以下公式计算:

$$\bar{\varepsilon} = \frac{1}{n} \sum_{i=1}^n \frac{100 |d_i - o_i|}{d_i} \quad (9)$$

式中 d_i 为目标输出值; o_i 为预测输出值, n 为输出数据数目。经计算可得: $\bar{\varepsilon} = 0.1489\%$, 表明表面能的计算值和预测值非常接近, 这说明上面建立的模型具有较高的精度。需要指出的是, 由于本文作者对数据进行了标准化处理, 程序所控制的目标误差是对归一化后的数据而言, 而上面计算的相对误差是对原始数据而言的。对于一个神经网络, 关键在于网络结构、权值和阈值。一旦网络训练完成, 这些数据就可以得到, 将它们保存起来, 编写成预测函数, 再次进行训练, 就无需再对神经网络进行训练, 预测的速度大大加快。

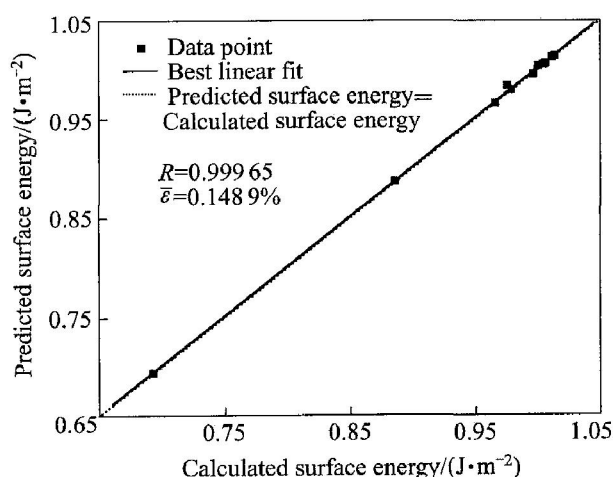


图 2 表面能的神经网络预测结果与 MD 计算结果的比较

Fig. 2 Comparison of ANN-predicted and calculated surface energy

利用上述模型, 作者对部分未知表面的晶面能进行了预测, 其结果列于表 2 中。

分析训练结束后的权值矩阵可以发现, 与输入矢量 θ 相关的权值明显高于其他两个输入矢量,

表 2 Au 不同晶面与 3 个低指数晶面的夹角及各晶面的表面能

Table 2 Angle between planes and three low-index planes and surface energies						
(hkl)	$\gamma_s^{(hkl)} / (\text{J} \cdot \text{m}^{-2})$		Predicted related error/ %	Angle between planes and plane (100)	Angle between planes and plane (110)	Angle between planes and plane (111)
	MD calculated	Predicted		$\theta_1 / (^\circ)$	$\theta_2 / (^\circ)$	$\theta_3 / (^\circ)$
100	0.886	0.887 0	0.112 87	0	45.00	54.74
110	0.976	0.983 9	0.809 43	45.00	0	35.26
111	0.693	0.693 6	0.086 58	54.74	35.26	0
112		0.851 7		65.91	54.73	19.47
121		0.829 1		65.91	30.00	19.47
210	1.014	1.012 6	- 0.138 07	26.57	18.43	39.23
211		0.830 6		35.26	30.00	19.47
221		0.859 9		48.19	19.47	15.78
310	0.997	0.996 0	- 0.100 30	18.43	26.87	43.09
311		0.982 1		25.24	31.48	29.49
312		0.807 8		36.70	40.90	22.21
320	1.011	1.012 0	0.098 91	33.69	11.30	36.81
321		0.872 9		36.70	19.11	22.21
331		0.907 8		46.51	13.27	22.00
332		0.805 8		50.24	25.24	10.04
410	0.979	0.980 0	0.102 15	14.04	30.96	45.56
411		0.999 8		19.47	33.56	35.26
412		0.925 8		29.20	39.51	28.13
413		0.851 8		38.33	46.10	25.07
430	1.005	1.005 1	0.009 95	39.66	8.15	36.07
510	0.966	0.966 0	0	11.31	33.68	47.20
520	1.006	1.005 8	- 0.019 88	21.80	23.21	41.37
530	1.014	1.013 7	- 0.029 59	30.96	14.05	37.62
540	1.001	1.003 8	0.279 72	38.66	6.33	35.76

表明各晶面的表面能大小主要与密排面(111)的夹角有关。表面能 γ 随各晶面与(111)面夹角 θ 的变化如图 3 所示。可以看出,面心立方的(111)密排面的表面能最小,和预期结果一致。其他各晶面表面能随其与(111)面夹角的增大,在角度较小时,呈现明显的增大趋势,并在 $\theta \approx 40^\circ$ 附近时达到最大值, θ 继续增大,表面能快速下降。对本文计算获得的各晶面表面能值研究发现,表面能值随 θ 增大快速下降的区域,是同属于 $\langle 001 \rangle$ 晶带轴的高密勒指数晶面(如(410)和(510)等)。本文计算获得的 Au 各晶面表面能随其与(111)晶面夹角 θ 的变化关系与张建民等^[12]采用改进嵌入原子法计算 Cu 晶体的表面能得到的表面能面密度随各晶面与(111)晶面夹角 θ 的变化趋势基本一致。可以预知具有面心立方特征的金属各晶面表面能随其与(111)晶面夹角 θ 的变化趋势基本上呈现先增大后减小的特点。对比图 3 和表 2 还可以发现,当不同晶面和密排面夹角相同时,就必须考虑与其他二个低指数晶

面的夹角关系。
经计算可得,预测的所有晶面表面能的平均值为 $0.920\ 2\ \text{J}/\text{m}^2$,同实验测得的 Au 多晶体的表面能

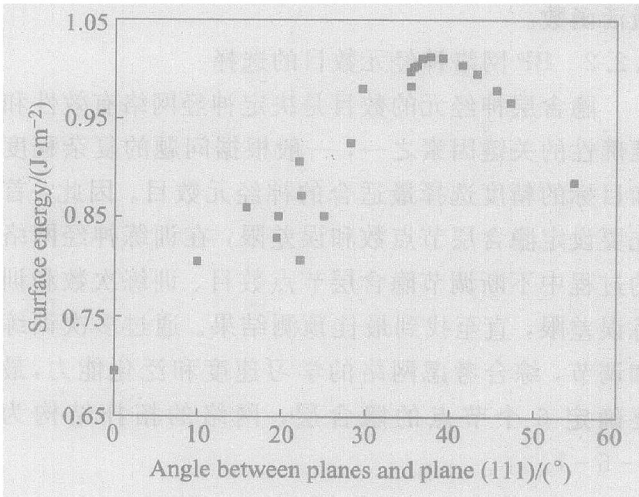


图 3 表面能随各晶面与(111)晶面夹角的变化
Fig. 3 Variation of surface energy with angle between planes and plane (111)

面密度平均值 1.50 J/m^2 ^[13]、Skriver 的第一性原理计算值^[3]和 Foiles 的 EAM 计算值^[14]比较发现, EAM 理论模型计算表面能及本方法预测的表面能均存在结果偏小的问题。这主要是: 1) 实验测定表面能 1.50 J/m^2 是所有晶面的表面能的平均值, 而本文计算的晶面较少; 2) 对通过分子动力学方法计算的几个晶面没有考虑到表面原子的松弛。因此仅考虑 3 个输入变量将会使预测结果存在一定的误差。而 Skriver 第一性原理计算获得的最密排面 (111) 晶面的表面能值为 1.61 J/m^2 , 比实验获得的平均表面能还要高, 同样显示了 Skriver 方法的不足。目前尚没有一种模拟计算方法可以完全解决计算晶面表面能同实验数据存在误差的问题。由表 2 中表面能预测数据及图 2 可知所建立的联合模型可以正确预言低指数晶面表面能 E_s 的排序: $E_{S(111)} < E_{S(100)} < E_{S(110)}$, 即表面的开放度越大, 其表面能越大。(111) 是最密排面, 该面的表面能最低。这表明, 表面原子排列越紧密, 表面能越低。根据固体表面物理学的有关理论^[15], 表面原子排列越紧密时, 原子间的相互作用越强, 从而使材料表面能降低, 这与计算结果是一致的。

4 结论

1) 利用神经网络对晶面夹角和表面能样本集的学习, 采用改进的 BP 网络算法——Levenberg-Marquardt 算法建立了表面能人工神经网络模型; 计算表明该神经网络模型具有较高的精度和很好的泛化能力。该方法可以直接用于其他面心立方金属表面能的计算, 也可以类推到其他的晶体结构。

2) Au 晶体其他晶面的表面能随其与 (111) 面夹角的增大呈现出先增大后减小的趋势, 并在 $\theta \approx 40^\circ$ 附近时达到最大值。表面能值随 θ 增大快速下降的区域, 是同属于 $\langle 001 \rangle$ 晶带轴的高密勒指数晶面(如 (410) 和 (510) 等)。

3) 分子动力学计算的表面能数据的整体精度及所考虑影响因素的全面性, 决定了对 Au 各晶面表面能的预测精度。

REFERENCES

- [1] 阮德水, 李卫萍. 金的化学[J]. 高等函授学报(自然科学版), 2000, 13(1): 25-29.
RUAN De-shui, LI Wei-ping. Gold chemistry[J]. Journal of Higher Correspondence Education(Natural Science), 2000, 13(1): 25-29.
- [2] 董守安. 纳米技术中的金元素[J]. 贵金属, 2003, 24(1): 54-61.
- [3] DONG Shou-an. The role of gold in the nanotechnology[J]. Precious Metals, 2003, 24(1): 54-61.
- [4] Skriver H L, Rosengard N M. Surface energy and work function of elemental metals[J]. Phys Rev B, 1992, 46(11): 7157-7168.
- [5] Daw M S, Baskes M I. Embedded atom method: derivation and application to impurities, surface, and other defects in metals[J]. Phys Rev B, 1984, 29(12): 6443-6453.
- [6] Johnson R A. Alloy models with the embedded atom method[J]. Phys Rev B, 1989, 39(17): 12554-12559.
- [7] QI Le-hua. Research on prediction of the processing parameters of liquid extrusion by BP network[J]. Journal of Materials Processing Technology, 1999, 95: 232-237.
- [8] Basheer L A. Artificial neural network: fundamentals, computing, design, and application[J]. Journal of Microbiological Methods, 2000, 43: 3-31.
- [9] Joines J A, White M W. Improved generalization using robust cost functions[A]. IEEE/INNS Int Joint Conference of Neural Networks[C]. New York: IEEE Press, 1992. 911-918.
- [10] Hagan M T, Menhaj M. Training feedforward networks with Marquardt algorithm[J]. IEEE Transaction on Neural Networks, 1994, 5(6): 989-993.
- [11] 张智星, 孙春在, 水谷英二, 等. 神经-模糊和软计算[M]. 西安: 西安交通大学出版社, 1998. 156-215.
ZHANG Zhixing, SUN Chunzai, Mizutani E, et al. Neuro-Fuzzy and Soft Computing[M]. Xi'an: Xi'an Jiaotong University Press, 1998. 156-215.
- [12] Hecht-Nielsen R. Neurocomputing[M]. Massachusetts: Addison Wesley Publishing Company, 1991.
- [13] 张建民, 徐可为, 马 飞. 用改进嵌入原子法计算 Cu 晶体的表面能[J]. 物理学报, 2003, 52(8): 1993-1999.
ZHANG Jianmin, XU Ke-wei, MA Fei. Calculation of surface energy of Cu crystal with modified embedded atom method[J]. Acta Physica Sinica, 2003, 52(6): 1993-1999.
- [14] De Boer F R, Boom R, Mattens W C M, et al. Cohesion in metals[M]. Amsterdam, North-Holland: Elsevier Science, 1988. 716.
- [15] Foiles S M, Baskes M I, Daw M S. Embedded atom method functions for the FCC metals Cu, Ag, Au, Ni, Pd, Pt, and their alloys[J]. Phys Rev B, 1986, 33(12): 7983-7990.
- [16] 闻立时. 固体材料界面研究的物理基础[M]. 北京: 科学出版社, 1991. 21.
WEN Lishi. Physics Basis of Solid Materials Interface Research[M]. Beijing: Science Press, 1991. 21.

(编辑 陈爱华)