
北 京 科 技 大 学

硕士学位研究生 选题报告及文献综述

论文题目：利用机器学习构建材料成分-结构-性能
关系模型



指导教师：金莹（研究员）、时鹏（副研究员）

单 位：国家材料服役安全科学中心

学 号：B20170427

作 者：赵朝阳

专业名称：材料科学与工程

入学时间：2017 年 9 月

2018 年 07 月 20 日

目录

1 选题依据.....	1
1.1 课题来源.....	1
1.2 研究目的及意义.....	1
2 文献综述.....	3
2.1 材料基因组计划.....	3
2.1.1 材料基因组计划概述.....	3
2.1.2 高通量材料实验.....	4
2.1.3 高通量材料计算设计.....	11
2.1.4 数字化数据平台.....	12
2.2 机器学习.....	13
2.2.1 机器学习概述.....	15
2.2.2 监督学习.....	15
2.2.3 非监督学习.....	18
2.2.4 强化学习.....	21
2.3 构建材料成分相图.....	23
2.3.1 相图.....	23
2.3.2 相图的构建方法.....	24
2.3.3 目前存在的不足与改进方案.....	25
2.4 预测碳钢的大气腐蚀模型.....	26
2.4.1 碳钢的大气腐蚀.....	26
2.4.2 碳钢大气腐蚀的预测模型.....	26
2.4.3 目前存在的不足与改进方案.....	27
2.5 探究不锈钢焊接接头成分、硬度与腐蚀性能的联系.....	27
2.5.1 不锈钢焊接接头研究的必要性.....	27
2.5.2 不锈钢的焊接性和焊接特点.....	28
2.5.3 不锈钢焊接接头腐蚀性能的研究现状.....	29
2.5.4 目前存在的不足与改进方案.....	30
3 研究方案.....	31
3.1 研究内容.....	31
3.2 研究方法.....	33
3.3 创新点.....	34
3.4 预期目标.....	34
3.5 进度安排.....	35
参考文献.....	36

1 选题依据

1.1 课题来源

国家重点研发计划：基于同步辐射光源和先进中子源的高通量材料表征技术与装置(项目号：2017YFB0701900)。

1.2 研究目的及意义

材料基因组技术是近年来兴起的材料研究新理念和新方法，是当今世界材料科学与工程领域的前沿。材料基因组技术的实质是通过融合高通量材料计算设计、高通量材料实验和材料数据库三大组成要素构建材料设计研发的协同创新网络，加速新材料从发现到应用的全过程。其中，高通量材料实验经过 20 多年的发展，目前面向多种形态材料和多种服役性能形成了一系列成功案例，高通量材料计算设计和材料数据库建设近几年也取得了较大进展。

机器学习在 1970 年就陆续出现了相关的算法。因为近年计算机计算能力的爆炸式增长，使得我们可以使用机器学习去解决日益复杂的问题。当能捕获和存储大量的数据时，机器学习能通过在数据自身中寻找解决问题的“答案”。

因为高通量材料实验产生的海量实验数据，传统的人工分析实验数据分析的速度远不能跟上实验结果的输出进度。当面临大量数据时，人工同样难以有效发现、提取实验数据中的潜在规律和联系。此时，机器学习的诸多算法便能帮助科研人员高效地辨识、分析海量的实验数据。

构建相图的传统方法是每次获取单个样品的合成与表征，这种成本较高、费时较长且不系统的方法在科技加速的当今远不能满足科研工作者的需求。自 19 世纪 60 年代，就有科研人员尝试寻找更有效建立成分-相关关系的方法。组合材料芯片技术、组合材料库的高通量合成与快速表征结合机器学习的相关算法对于加速材料筛选和优化显示出巨大的潜力。

在金属遭受的腐蚀中，大气腐蚀是最普遍和严重的，有资料显示，大气腐蚀造成的金属损失约占腐蚀总损失量的 50% 以上。研究碳钢在大气环境中的腐蚀性能，了解腐蚀特性和规律，对于合理选用材料并提供相应的防腐蚀措施，延长设备和构件的使用寿命，减少腐蚀造成的经济损失，具有重要意义。过去使用的碳钢大气腐蚀预测模型多是预测值与期望值平方差之和最小的回归模型。这些模型仅在非常有限的范围内有效，难以捕捉腐蚀过程的非线性变化。因此，科研人员一直在寻找能够在更广范围的气候和污染水平状况下预测大气腐蚀速率的数学模型。机器学习中的神经网络模型，因在合理

选择网络结构和参数的情况下具备模拟复杂模型的能力，所以可作为碳钢大气腐蚀的预测模型。

不锈钢由于具备良好的物理、化学性能，它的生产和应用在深度和广度上不断发展和扩大。不锈钢的使用范围十几年来向轻工和家庭生活方面迅速扩展。由于不锈钢的使用过程中，很大一部分都要通过焊接的方式将构件进行连接，因此焊接质量的好坏直接影响着不锈钢构件的性能。在不锈钢构件的腐蚀失效分析中，焊接接头所受到的腐蚀是整个构件腐蚀当中最严重的。因此深入研究焊接接头的耐蚀性能对不锈钢构件的安全使用和寿命评估起着相当重要的作用。但是不锈钢焊接接头的区域一般较小，有限的实验区域必然会限制实验数据的体量，即小样本数据。机器学习中的深度强化学习往往能充分适应小样本的数据，并提供较理想的预测模型。

2 文献综述

2.1 材料基因组计划

2.1.1 材料基因组计划概述

美国总统巴拉克·奥巴马于 2011 年 6 月 24 日在卡耐基·梅隆大学作的以“先进制造业伙伴关系”为主题的演讲中明确地指出材料基因组计划的总目标是“将先进材料的发现、开发、制造和使用的速度提高一倍”。白宫科技政策办公室在 2011 年发布相应的白皮书《具有全球竞争力的材料基因组计划》中阐述了材料创新基础设施的三个平台：计算工具平台、实验工具平台和数字化数据（数据库及信息学）平台，如图 1 所示^[1]。材料基因组计划提倡的并不只是开发和应用计算材料学，而是要集成计算工具、实验工具和数据库来加快材料的设计与应用。白皮书中强调材料创新基础设施是一个材料设计与制造的“加速器”。

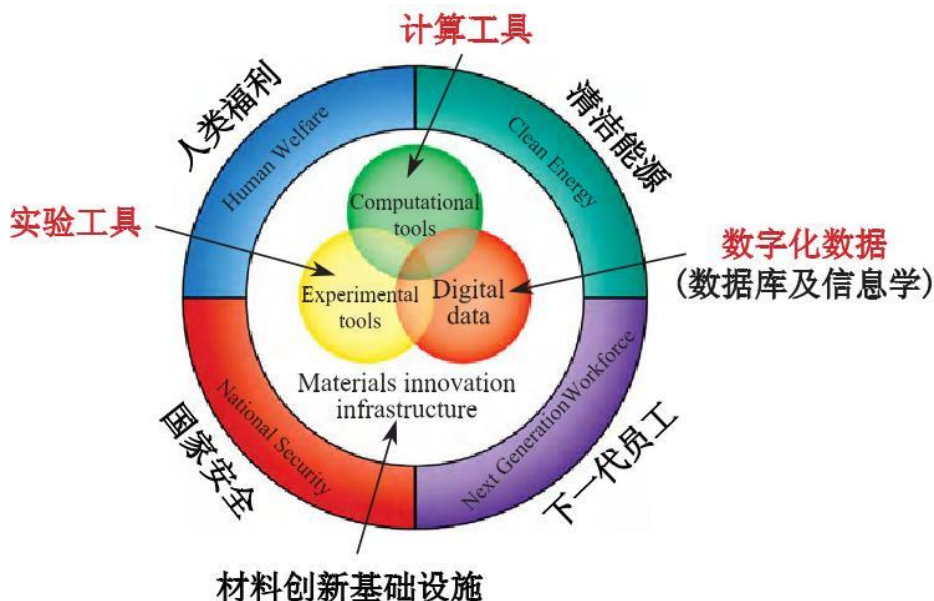


图 1 材料基因组计划中的材料创新基础设施的内涵^[1]

材料基因组计划中的材料基因组技术包括高通量计算模拟、高通量材料试验和材料数据库三大组成要素；其中材料计算模拟是实现“材料按需设计”的基础，可以帮助缩小高通量材料实验范围，提供实验理论依据；高通量材料试验起着承上启下的角色，既可以为材料模拟计算提供海量的基础数据和实验验证，也可以充实材料数据库，并为材料信息学提供分析素材，同时还

可以针对具体应用需求，直接快速筛选目标材料^[2]；材料数据库可以为材料计算模拟提供计算基础数据，为高通量材料实验提供实验设计的依据，同时计算和实验所得的材料数据亦可丰富材料数据库的建设。

2.1.2 高通量材料实验

（1）高通量材料实验概述

高通量在物理层面上的含义指单位时间内流经某单位面积的某属性量；高通量是指提高材料制备及性能测试效率，加速材料研发进程的一种方法。材料高通量实验是指在短时间内完成大量样品的制备与表征，其核心思想是将传统材料研究中的顺序迭代方法改为并行处理，从而以量变引起材料研究效率的质变^[2]。

材料基因组计划中的实验平台的主要实验内容是材料高通量实验。经过 20 多年的发展，高通量组合材料实验技术已取得较大的进展，目前已形成覆盖薄膜、块体、粉体等多种材料形态高通量制备和满足热学、电学、光学、电磁学、电化学、物相等各类材料性能高通量表征的完整实验技术体系。尤其是近年来，随着电子信息技术的发展，一批具有更高实验通量、更快实验速度和更高时间/空间分辨率的制备与表征装备得到发展，并在催化剂、半导体和合金等领域取得较好示范验证^[3]。

（2）高通量材料实验的制备技术

高通量组合材料样品按照维度从低到高可分为粉体材料、薄膜材料和块体材料，其制备过程通常可概括为“组合”与“成相”两步，其中前者实现了高通量样品的成分可控分布，后者实现了高通量样品的物相结构可控分布。因本人课题的实验材料为材料组合芯片，属于薄膜材料，下文将着重介绍基于薄膜材料的高通量制备和表征技术。

基于薄膜形态的组合材料芯片是目前发展最为成熟的高通量材料制备技术，该技术可基于磁控溅射、电子束蒸发等传统薄膜制备方法实现。其“组合”步骤按照分布可控程度由低至高可分为共沉积法和物理掩模法，物理掩模法又可细分为分立掩模法和连续掩模法。

1) 共沉积法

共沉积法是指利用多个沉积源同时工作，共同完成单一样品的制备，所形成的样品材料成分渐变连续分布，且通过改变沉积源的出射角度或相对位置等制备参数，可调整高通量样品的成分分布。如图 2 所示，文献^[4]利用三靶磁控共溅射装置在并行纳米量热器件阵列上单次实验完成 25 种不同成分

的 Cu-Au-Si 玻璃态合金材料样品库制备；并且通过改变溅射靶相对垂直方向的倾斜角度和溅射功率，可进一步改变所制备样品的成分分布。该方法无需额外的物理掩模即可获得不同成分连续分布，并可实现不同材料原子级的均匀混合，而无需额外扩散热处理，可直接对样品进行高温结晶成相，材料制备工艺和制备装备相对简单，但该方法无法实现多元材料组合的精确可控分布和多元成分空间的完整覆盖，限制了该方法的应用范围。

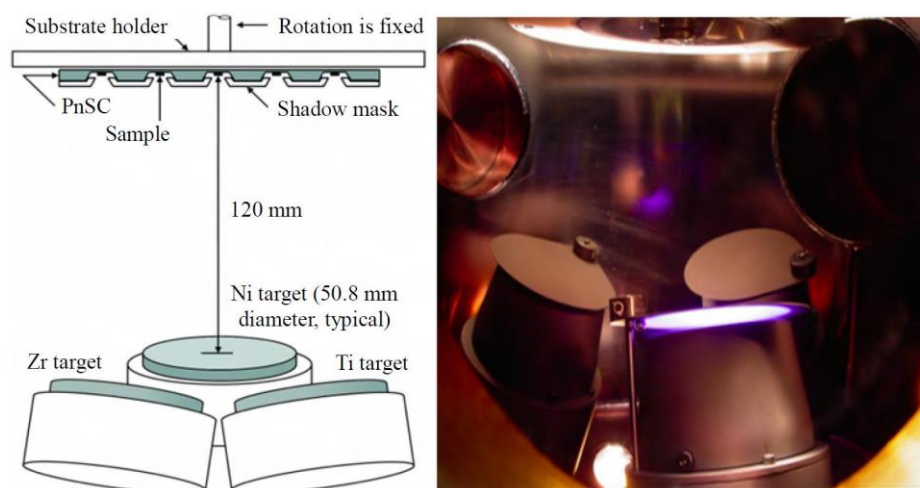
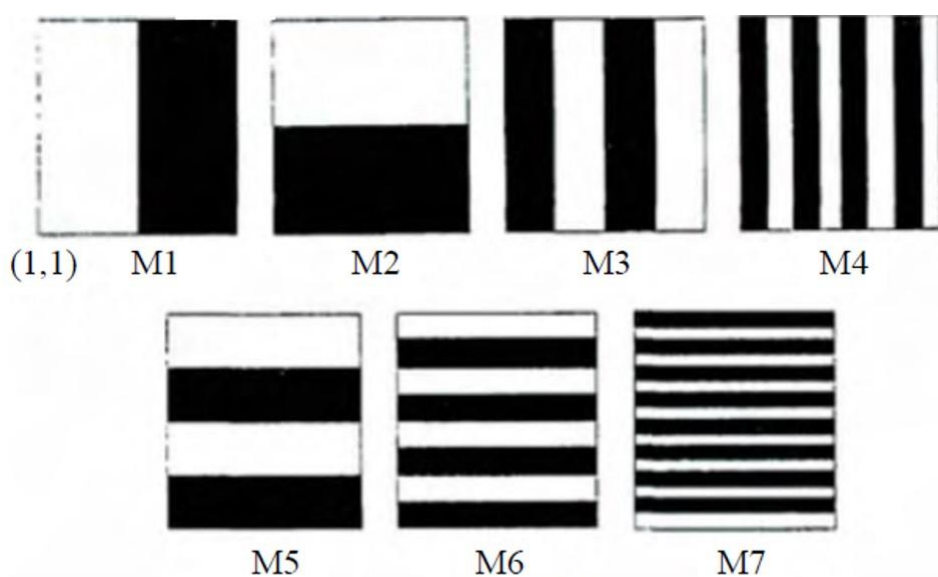


图 2 共沉积法示意图

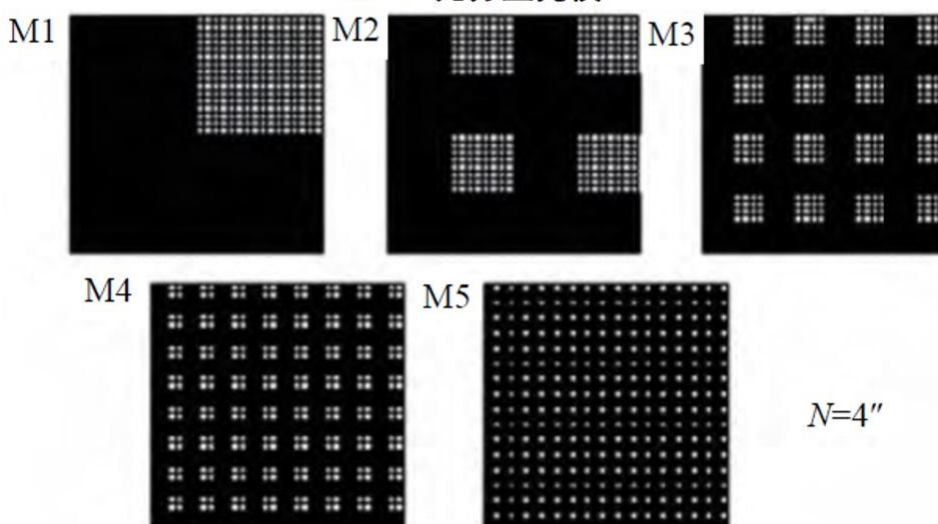
2) 物理掩模法

物理掩模法是实现高通量样品可控“组合”的方法，分再为分立掩模法和连续掩模法。

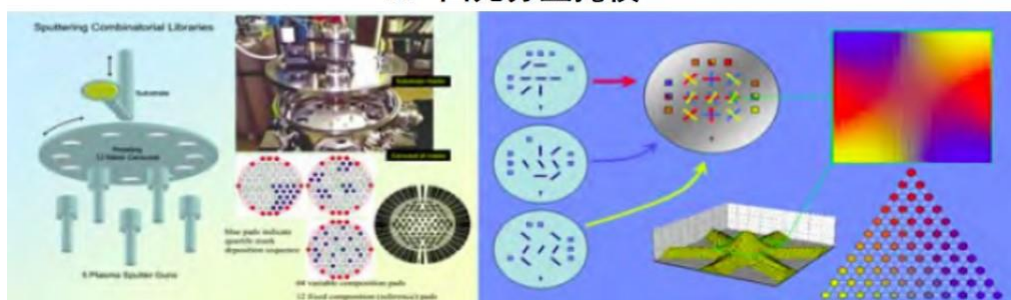
分立掩模法将物理掩模技术和薄膜材料沉积技术相结合，单层材料沉积使用一块掩模板和一种沉积源，并多次组合和更换不同的掩模板和沉积源，在薄膜均匀沉淀的前提下，实现叠层薄膜的依次沉积、多元材料的组合和样品单元的空间可控分布。常用的分立掩模包括二元掩模、四元掩模和多元掩模，如图 3 所示。



a. 二元分立掩模



b. 四元分立掩模



c. 多元分立掩模

图3 分立掩模法示意图

其中文献^[5]利用脉冲激光沉积、离子束溅射等制备技术，将二元和四元分立掩模应用于荧光材料、超导材料和介电材料等的高通量制备，单个基片

最多可制备 1024 个不同成分样品单元，极大地提高了材料研究的效率。文献^[6]设计了由 12 种不同掩模组成的多元分立掩模组，可实现 2 英寸基片上多达 64 个不同样品的制备，结合预制的电极阵列，应用于燃料电池电极材料的电化学性能高通量表征。分立掩模法可用于组成元素多、成分空间跨度大的材料样品库制备，适合电学、电化学等需要对分立样品进行表征的技术要求，尤其可用于器件的高通量研究，但由于其实验通量相对有限，依然无法实现多元材料体系的系统、完整研究。

连续掩模法是指在薄膜沉积速率均匀可控的前提下，利用有电机控制下可连续移动的掩模板实现薄膜厚度梯度沉积的方法，如图 4 所示为连续掩模制备三元相图的工作流程^[2]，结合旋转角度精确控制的基片台，可用于三元连续相图组合材料芯片的制备。

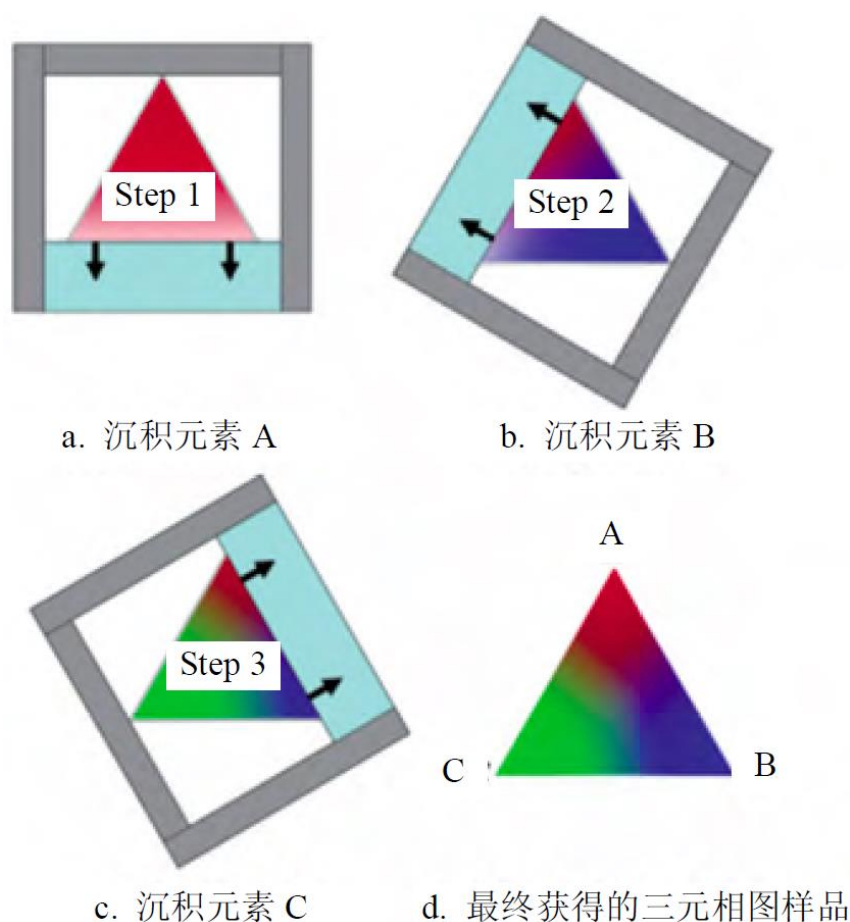


图 4 连续掩模法示意图

有文献将该技术用于 Ge-Sb-Te 相变存储材料样品库的制备，通过步进电机控制连续掩模步长 1mm/步精确控制，在 1inch 基片上实现了样品密度 2 个/mm² 的样品制备。

上述基于物理掩模技术制备的叠层薄膜只实现了材料堆叠，还需要低温扩散热处理过程方可实现多元材料的均匀“扩散”，此外还需“原位”或“离位”热处理“成相”过程才可实现组合材料芯片的制备。其中，低温扩散热处理是组合材料芯片制备的关键，倘若未实现完全均匀扩散就对进行高温结晶，则会由于存在复相而影响实验结论，这是因为多层膜界面上会发生“扩散”和“结晶”竞争热处理过程，界面一旦结晶则阻碍进一步进行。

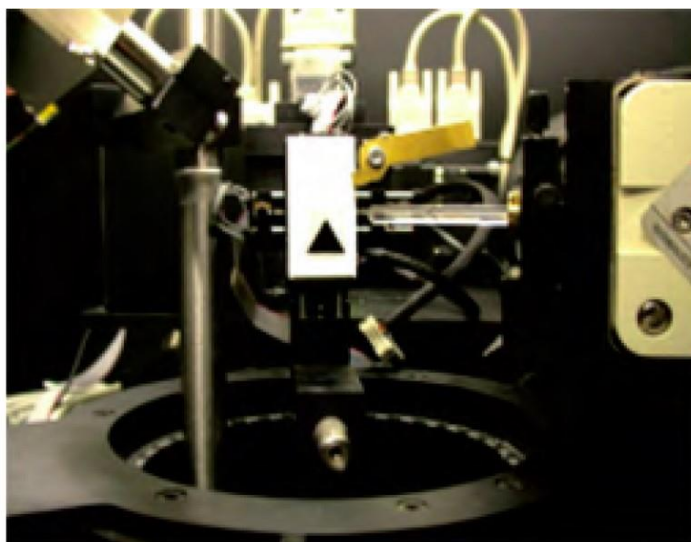
(3) 高通量材料实验的表征技术

高通量材料表征是高通量材料实验技术的重要组成部分。针对材料不同的应用需求，近年来研究人员已陆续发展了面向材料围观基本单元、介观材料等不同尺度的高通量表征技术，涉及电学、磁学、电化学、电磁学和热力学等多种材料性能表征。

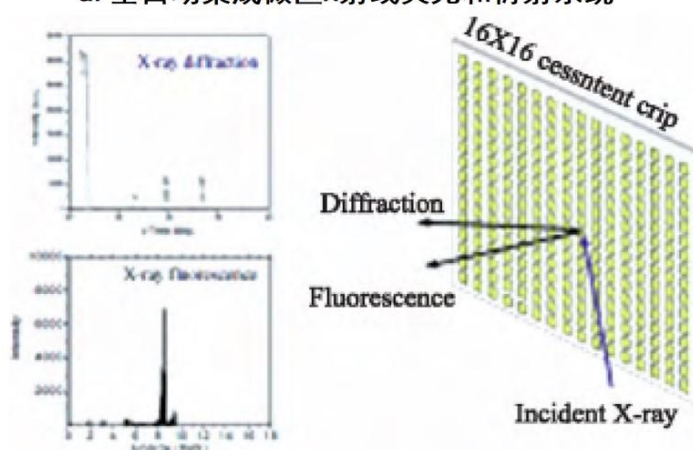
1) 高通量微区成分、结构表征

光学方法是研究材料成分、结构最直接有效的表征方法，常用的光学检测方法包括 X 射线衍射/散射、X 射线荧光光谱分析、X 射线能谱仪、紫外/可见/红外分光光度计等。但是常规的光学检测仪器由于其光通量密度较低，影响了空间分辨率的进一步提升，从而限制了常规光学仪器在高通量微区表征中的应用。同步辐射光源克服了光通量低下的缺点，从红外至硬 X 射线全光谱范围内均能实现高亮度微聚焦，提升了表征的空间分辨率，因此成为理想的材料成分、结构高通量表征手段。

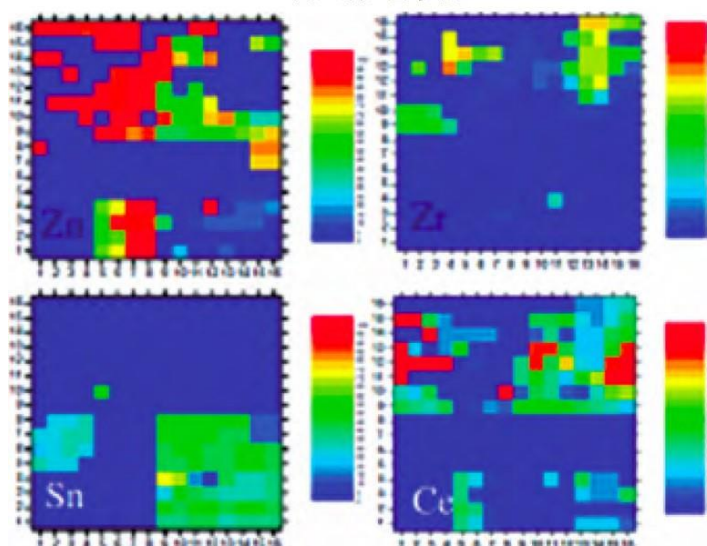
由于同步辐射光源等大型科学装置资源有限，在使用机时方面并不能得到充分保证。在此背景下，美国劳伦斯伯克利国家实验室开发了全自动集成微区 X 射线荧光和衍射系统，该系统的微束聚焦 X 射线的亮度是普通 X 光源的 20 多倍，空间分辨率最高可达 $10\ \mu\text{m}$ ，可以同时高通量实验样品进行成分和结构的快速表征。图 5 分别展示了全自动集成微区 X 射线荧光和衍射系统的仪器照片、工作原理、以及用该仪器对四元(Zn, Zr, Sn, Ce)组合材料芯片快速表征所得的组分分布图。此外，布鲁克公司生产的 D8 discover 系列 X 射线衍射仪也可以实现最小 $50\ \mu\text{m}$ 束斑的微区物相表征。



a. 全自动集成微区X射线荧光和衍射系统



b. 测量原理示意



c. 四元组合材料芯片组分分布彩图

图 5 全自动集成微区 X 射线荧光和衍射系统

2) 高通量微区电磁学表征

电磁学特性是科学研究中最基本的物理特性,分为电学特性和磁学特性,其中:电学特性包括超导性质、电导率、介电常数、铁电常数、磁阻效应、电子迁移率、扩散长度、腐蚀、接触电阻、界面参数、能级对准等,磁学特性包括磁化率、自旋共振等。衰减微波探针显微镜是研究以上电磁学特性最有效的高通量研究工具。衰减微波探针显微镜利用衰减波的原理,将微波信号控制在谐振腔内针尖上极小的微区,通过针尖与样品间发生相互作用,进而改变谐振腔的频率和品质因子,结合相关物理模型,最终获得材料的电磁学性质。图6是美国劳伦斯伯克利国家实验室开发的衰减微波探针显微镜的工作原理和仪器照片。该系统拥有极高的微区分辨率,结合样品台的自动化控制和数据采集处理功能,能够应用于材料芯片电磁学特性的高通量、快速和自动化表征。

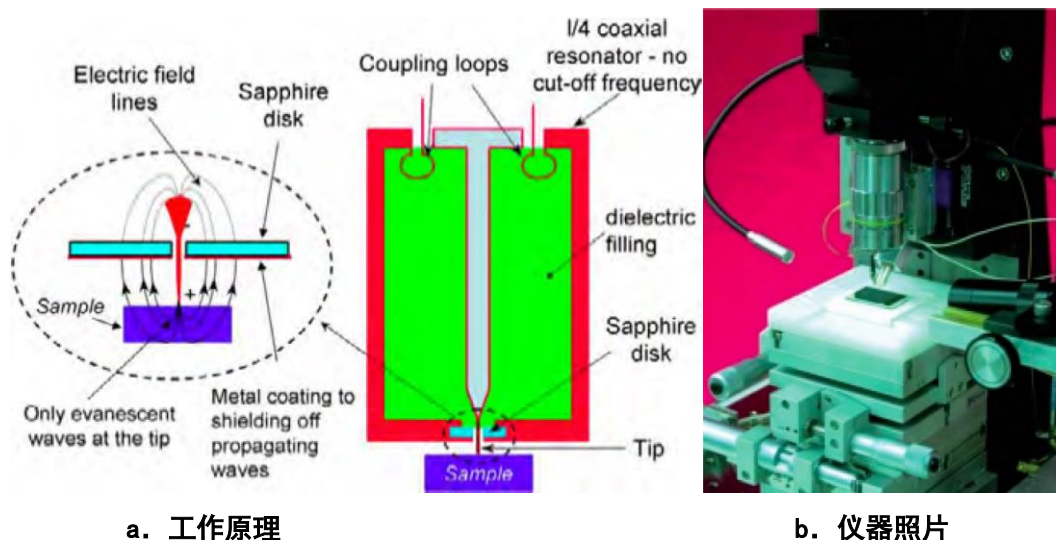


图6 衰减微波探针显微镜^[7]

3) 高通量微区电化学表征

对于电极、电解质等电池、电容材料或器件而言,其电化学性质具有重要的研究意义。作为材料基因芯片高通量表征的基本要求,电化学性质的表征仪器必须具有高分辨率和自动化特性。

目前,广泛应用于锂电池正负极、薄膜电解质、半导体等重要材料的高通量组合电化学表征仪器是由美国 Princeton Applied Research、AMETEK, Inc. 开发的 VersaSCAN 微区电化学扫描系统^[8]。该仪器的特点是样品定位精度高,平台空间分辨率可达 50nm,样品测试区域为 100×100mm,且满足高密度组合材料样品的全自动编程测试需要。该高通量微区电化学测试平台基于电化

学过程和材料电化学特性，可提供扫描电化学显微镜、扫描开尔文探针、扫描振动电极测试、微区电化学阻抗测试、扫描电解液微滴测试、非接触式微区形貌测试等电化学测试。

2.1.3 高通量材料计算设计

（1）高通量计算材料设计概述

高通量材料计算设计是材料科学中一片新兴的领域。通过将先进的热力学、电子结构与智能的数据挖掘、数据库构建以及利用如今超级计算机的架构，科学家们生产、管理和分析海量的材料数据仓库来寻找优质的材料。

所有技术都与某种特定材料密切相关。钢铁制成的蒸汽发动机推动了 18 世纪的工业革命，硅为信息通讯技术打下来基础。一旦某种材料被一项技术选中，这种组合就因为建立大规模生产线的投资难以变动而基本固定了。这意味着改变一项已经建立起来的技术中所使用的材料的情况是很少见的，新材料的使用往往会引发该项技术的革命。为了让技术持久有效，技术人员对材料的初始选择是极其苛刻的。更重要的是，最近出现一些新型的每项技术都需要不同的材料。因此，开发新材料的压力变得不堪重负。材料的筛选要根据的技术的基本属性、要求来确定，比如要求不包含有毒元素、当需要制备大量材料时，原材料因偏重价格低廉的选项。因此，寻找材料是一个需要从多个维度同时考虑的问题。

尽管对于材料的需求在持续的增长，实验发现仍受限于高昂、耗时的合成工艺。有其它的方法解决材料设计的问题吗？有，那就是计算材料新兴的一个领域“高通量”计算材料设计。量子力学热力学方法、大量基于数据库结构的技术以及智能数据挖掘的结合造就了高通量计算材料设计。概念简单但有效：创建一个大型的数据库，数据库中包含对已存在的材料及假想的材料进行计算得来的热力学信息及电子性质；之后在数据库中查找带有理想性质的材料。显然，这个工作需要真实的结果来验证其有效性，比如，能正确预测已存在材料的性能，假设的材料最终也能被制备合成出来。这种真实性检验反馈到理论上能更好的构建数据库并提升预测能力。

高通量实验方式在 100 多年前被 Edison 和 Ciamician 率先提出，但直到高效、精确的理论工具及廉价的计算机出现，才使得其具备解决材料设计问题的能力。因此，过去 10 年来，高通量计算材料研究的发展紧随着高通量实验方法高歌猛进的发展势头。

（2）高通量材料设计的应用

高通量计算的应用受到三个紧密联系步骤的约束：①虚拟材料的生长，材料热力学和电子结构的计算；②合理材料存储，数据仓库中信息的系统化存储；③材料表征与筛选，致力于筛选合适的材料或获取新的物理观点的数据分析。

高通量通常因自身产生的大型数据库而被人们所熟知（例如，AFLOWLIB.org 和 Materials Project）。这里我们假设三个高通量阶段都是高度必须的，但是最后一个约束是最具挑战性也是最重要的。材料的表征与筛选是指从数据库中提取有效的信息，但是要求实验人员对手中问题的物理背景有相对透彻的理解。数据库的智能搜索通过“描述器”的方式来执行。描述器可能是定性的经验、不必要的观测，材料的微观计算参数到宏观属性。换言之，描述器是研究者与数据库对话的语言。

一旦定义了一个好的描述器，在数据库中筛选更佳材料的的操作可以分为内源性执行或外源性执行，分别对应于最优解是否已经包括在计算结果的集合里。内源性搜索只包括上述约束的第三步，只需要快速的描述器，也可以采用多种信息检索技术。内源性检索的案例有搜寻更佳的阴极材料、揭示未知的化合物、了解优质的拓扑绝缘体或者是热电材料。因为需要通过迭代寻找最优解会导致数据库的扩大，外源性搜索包含上述的三个步骤。

外源性高通量计算搜索的一个重要成分是范式。范式能够利用数据库中已有数据上的描述器的评估来指导数据库中尚不存在的数据的计算。已发表文献中的范式有进化算法、遗传算法、基于特征分解和贝叶斯概率的数据挖掘和结构映射分析等。神经网络和支持向量机在一些案例中被结合使用。这些方法有时可能跳过高通量分析的第三步，即描述器需要具备物理含义。这样即使对手中问题的物理背景不甚了解仍可进行搜索。

2.1.4 数字化数据平台

材料数据分为计算数据和实验数据。长期以来，材料数据研究处于单打独斗和小规模的“数据制造-简单处理”模式，往往采用图表和统计方法等传统低通量人工数据处理方法，针对单次或数次计算、实验得出的少量数据进行分析，并对其规律进行猜想和提出经验公式，无法严谨预测和深度挖掘材料本质科学规律，造成材料研究经验结论多于理论的现状，无法完成从“试错”材料研究向材料理性设计的转变，同时也使得相同工作盲目重复进行，极大地浪费了有限的科研资源。为解决上述问题，目前国际上已有多个国家建立了跨机构的材料基因数据库，将计算数据与实验数据实时全面地搜集储

存和共享，以便深度发掘有用信息和规律。

美国麻省理工学院建立的 Materials Project 数据库，主要集中在无机固体上，尤其以锂离子电池材料为主。Materials Project 利用密度泛函理论(density functional theory)收集的巨型数据库来预测模拟物质模型的实际属性。目前该数据库里保存了大约 10 万种可能存在的材料。为了充分发挥这些数据在新材料研发中的作用，研究人员用人工筛选结合机器学习的方式来探索这些数据间蕴含的材料本质性能规律。Materials Project 采用分布式计算的原理，使用者可以通过在电脑上下载一个程序来进行运算并返还结果。美国哈佛大学清洁能源计划 建立起来的 Molecular Space 数据库也是基于密度泛函理论，采用人工加机器学习的方式来挖掘数据库的潜力。目前，Molecular Space 数据库在网上发布了 230 万种元素组合供研究人员使用。日本国立材料科学研究所建立的材料数据库是在其原有的 11 个材料数据库基础上整合建立的，涵盖了聚合物、无机非金属材料、金属材料、超导材料、复合材料以及扩散等内容，是目前世界上最大的、最全的材料数据库系统。目前，其含有数据库及应用系统已达到 20 个，包括 8 个材料基本性能数据库，3 个工程应用数据库，5 个在线结构材料数据库以及 4 个数据库应用系统。目前注册用户超过 80000 名，分别来自 149 个国家的 21228 个组织机构。

2.2 机器学习

通过对已有计算数据和实验数据的挖掘来发现一些模式，基于这些模式进而获得对材料性质的定量或者定性描述，是加快材料研发的重要方法，这种基于数据挖掘的方法可被称为数据驱动的方法（以下又称数据方法）。数据驱动方法的有别于那种有着因果关系和物理内涵的模型方法，它更强调基于大量的数据，寻找数据之间的关联^[9]。

2016 年 5 月 5 日，Raccuglia 等关于机器学习算法改变材料发现方式的文章，被《Nature》以封面论文的形式刊出，其中提出“从失败中学习”：哈佛大学研究者利用机器学习算法，用失败或不成功的实验数据预测了新材料的合成，“我们的机器学习模型获得了比传统人类合成策略更好的效果，并成功预测了有机模板合成的无机物形成条件，成功率达到 89%”^[11]，这意味着机器学习也能帮助加快材料研发，发明新材料的可能性也大幅提高。这种算法的好处在于，失败的实验数据也能用作下一轮的输入，继而不断完善算法。随着材料科学的发展，数据积累越来越庞大，如何从数以亿万计纷繁复杂的数据中提取有用信息，分析并梳理材料成分-工艺-结构-组织-性能的关系，必然成为材料研究的核心和关键。机器学习方法带领研究人员进入材料空间更

深入、更复杂的认知新模式，可以更加科学高效地推动材料设计的发展。

通过数据驱动方法帮助新材料研发已有很多成功案例。1992 年，Makishima 介绍了利用玻璃材料数据库和知识库开发的一个材料设计专家系统：它能从数据库中选出满足一定要求的玻璃组份，再利用知识库中有关玻璃生成规则，预报这些组份生成玻璃的可能性，并利用玻璃组份的原子半径和分解能的数据计算该玻璃材料的物性。1992 年，Yasui 和 Futagami 利用回归处理方法对数据库中的数据进行挖掘，建立起玻璃材料的组份与性能之间的关系，并构造了一个专家系统来预测钙钛矿的生成。1999 年，Ashby 提出了“材料选择器的概念”（material selector），即利用已知物质的数据库帮助特定用途材料的选择。

此外，通过数学方程建立结构和材料性质之间的函数关系，即通过定量构效关系（quantitative structure activity relationships, QSAR）也能帮助发现材料的一些性质。Rajan 在 2011 年提到，他的团队没有通过第一性原理计算，而是借用了 QSAR 的理念，通过数据驱动的方法，识别与已知晶体结构和电子结构相关的离散标量描述符和观察到的材料性能之间的关系，从而找到了决定高温压电钙钛矿关键性能的“结构-性能”描述符。

高通量材料计算和机器学习的集成，即 QM/ML 的方法（quantum mechanics/machine learning）更是引起目前业界的普遍关注。QM/ML 主要强调通过量子力学计算，产生大量的数据，然后从该数据中学习一些模式，利用该模式来预测材料的性质。2006 年，Ceder 课题组利用结构之间的关联性发展了数据挖掘结构预测（data mining structure predictor, DMSP）的方法，并成功实现了银镁合金的基态结构预测^[11]。2012 年，Saad 和 Bobbitt 等结合计算机与材料化学领域的优势共同研究了数据挖掘在材料结构预测的几种常用算法，利用监督学习方法以平均 95% 的准确率实现了二元合金的结构预测。此外，该研究组还利用机器学习的统计回归方法基于第一性原理计算结果进行了材料熔点的预测，平均相对误差小于 12.8%^[12]。Wolverton 研究组于 2013 年就开始在已经建立的 OQMD 中进行数据挖掘，发展了巨正则线性规划（grand canonical linear programming GCLP）的机器学习方法，通过组分就可以实现对材料稳定结构的预测^[13]，并将该方法成功应用于锂离子电池阳极材料^[14]和镁基三元长周期堆垛有序（long-period stacking ordered LPSO）结构^[15]的预测。陈冠华研究组是开展 QM/ML 方法较早也是做得比较成功的团队之一，该研究组在 2003 年便提出了利用第一性原理计算与神经网络（机器学习方法）相耦合的方法提高材料计算的精度，其关键技术在于利用神经网络挖掘实验数据与计算结果之间的定量关系，从而对第一性原理计算结果进行校

正^[16]。Bligaard 利用含有 64000 有序金属合金 (ordered metallic alloys) 的数据库, 利用经济学里的帕累托优化方法 (Pareto optimal), 寻找到了低压缩性, 高稳定性, 并且低成本的合金优化方法。他们首先利用高通量第一性原理计算, 计算了 64149 种多达 4 个原子晶胞结构的面心立方和体心立方结构的合金状态方程, 建立数据库^[17], 然后利用该数据库并结合帕累托优化法进行多目标优化, 寻找到满足特定应用需要的优化合金方案。

2.2.1 机器学习概述

机器学习属于计算科学领域人工智能的一个分支, 它不用清晰的编码规则而是用数据通过统计学技术赋予计算机“学习”的能力。“机器学习”一词是 Arthur Samuel 在 1959 年首次提出来的。机器学习起源于模式识别和人工智能中的计算学习理论, 它研究和构建能够基于数据学习和制定预测的算法。这些算法通过简单的输入制定数据驱动的预测/决策来克服随之而来的严格静态编码指令。机器学习被用于解决难以用清晰的指令获得良好效果的计算任务。

机器学习主要包括监督学习、非监督学习、强化学习等方面, 下文将分别进行详细介绍。

2.2.2 监督学习

(1) 监督学习概念

监督学习 (Supervised Learning) 是机器学习中的一类方法, 可以由训练资料中学到或建立一个模式, 并依次模式推测新的实例。训练资料是由输入样本和样本标签所组成。当模型的输出是一系列连续的值时, 称为回归分析; 当模型的输出是一系列分类标签时, 称为分类预测。

监督学习包含众多的分类/回归算法, 如决策树、逻辑斯谛回归、支持向量机、遗传算法、神经网络等算法。基于本文的工作所用到的监督学习方法, 下文将着重介绍神经网络算法。

(2) 神经网络

人工神经网络 (Artificial Neural Networks, ANN) 简称神经网络 (NN), 是对人脑或自然神经网络若干基本特性的抽象和模拟, 是一种基于连接学说构造的智能仿生模型, 是由大量神经元组成的非线性动力系统。

1) 神经元模型

人工神经元是对生物神经元的模拟与抽象，是神经网络的基本处理单元。目前人工神经网络模型大多采用由心理学家 W•McCulloch 和数理逻辑学家 W•H•Pitts 共同提出的 M—P 模型，图 7 表示一个人工神经元模型。

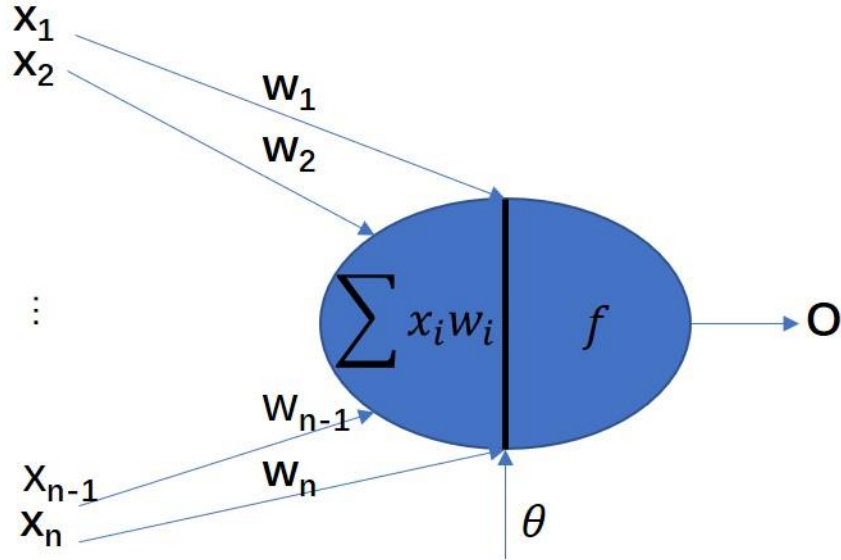


图 7 人工神经元模型

图中 $X = (x_1, x_2, \dots, x_n)$ 为该神经元的 n 个输入，来自外部或其他神经元的输出。 $W = (w_1, w_2, \dots, w_n)$ 表示与该神经元相连的 n 个神经元之间的连接强度，称为权值； $\sum wx$ 称为激活值，表示这个人工神经元的输入总和； O 表示这个神经元的输出； θ 表示这个人工神经元的阈值。这个输入信号的加权和超过 C ，则人工神经元被激活。这样，人工神经元的输出可描述为：

$$O = F(\sum WX - \theta) \quad (1)$$

其中 $F(\cdot)$ 表示神经元输入-输出关系函数，称为激活函数或输出函数。阈值 θ 一般不是一个常数，它是随着神经元的兴奋程度而变化的。

2) 激活函数

在神经网络中，网络解决问题的能力与效率除了与网络结构有关外，在很大程度上取决于网络所采用的激活函数。激活函数的选择对网络的收敛速度有较大的影响，针对不同的实际问题，激活函数的选择也应不同。常用的激活函数有以下几种形式^[18]：

① 阈值函数(Threshold)，其表达式为：

$$p(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)$$

该函数通常也称为阶跃函数。当激活函数采用阶跃函数时，人工神经网络

络即为 MP(McCulloch-Pitts)模型。此时神经元的输出取 1 或 0，反映了神经元的兴奋或抑制。

②线性函数(Linear)，其表达式为：

$$y = kx + b \quad (3)$$

该函数可以在输出结果为任意值时作为输出神经元的激活函数，但是当网络复杂时，线性激活函数会大幅降低网络的收敛性^[19]，故实际中较少采用。

③对数 S 型函数(Sigmoid)，其表达式为：

$$f(x) = \frac{1}{1+e^{-x}} \quad (4)$$

对数 S 型函数的输出介于 0~1 之间，常被要求为输出在 0~1 范围的信号选用。它是神经元中使用最为广泛的激活函数。

④双曲正切 S 型函数(Tanh)，其表达式为：

$$f(x) = \tanh(x) = \frac{1-e^{-x}}{1+e^{-x}} \quad (5)$$

双曲正切 S 型函数类似于被平滑的阶跃函数，形状与对数 S 型函数相同，以原点对称，其输出介于-1~1 之间，常被要求为输出范围在-1~1 之间的信号选用。

3) 人工神经网络的特点

①固有的并行结构和并行处理^[20]

人工神经网络和人类的大脑类似，不但结构上是并行的，它的处理顺序也是并行和同时的。在同一层内的处理单元都是同时操作的，即神经网络的计算功能分布在多个处理单元上，而一般计算机通常有一个处理单元，其处理顺序是串行的。

②知识的分布存储

在神经网络中，知识不是存储在特定的存储单元中，而是分布在整个系统中，要存储多个知识就需要很多链接。在计算机中，只要给定一个地址就可得到一个或一组数据。在神经网络中要获得存储的知识则采用“联想”的办法，这类似人类和动物的联想记忆。人类善于根据联想正确识别图形，人工神经网络也是这样。

③容错性

人工神经网络具有很强的容错性。它可以从不完善的数据和图形中进行学习并做出决定。由于知识存在于整个系统中，而不只是一个存储单元中，预订比例的结点不参与运算，对整个系统的性能不会产生重大的影响。能够处理那些有噪声或不完全的数据，具有泛化功能和很强的容错能力。

④自适应性

自适应性根据所提供的数据,通过学习和训练,找出输入和输出之间的内在关系,从而求取问题的解,而不是依据对问题的经验知识和规则,因而具有自适应功能,这对于弱化权重确定人为因素是十分有益的。

⑤模式识别能力

目前有各种各样的神经网络模型,其中有很多网络模型善于模式识别。模式识别是 ANN 最重要的特征之一。它不但能识别静态信息,对实时处理复杂的动态信息(随时间和空间变化的)也具有巨大潜力^[21]。模式识别往往是非常复杂的,各个因素之间相互影响,呈现出复杂的非线性关系,人工神经网络为处理这类非线性问题提供了强有力的工具。

2.2.3 非监督学习

在机器学习,非监督学习(Unsupervised Learning)的问题是在未加标签的数据中,试图找到隐藏的结构。因为提供给学习者的实例是未标记的,因此没有错误或报酬信号来评估潜在的解决方案。非监督学习与统计数据密度估计的问题密切相关。然而无监督学习还包括寻求,总结和解释数据的主要特点等诸多技术。在非监督学习中使用的许多方法是基于处理数据的数据挖掘方法。

非监督学习方法可以分成两大类,一类为基于概率密度函数估计的直接方法,指设法找到各类别在特征空间的分布参数在进行分类;另一类称为基于样本间相似性度量的间接聚类方法,其原理是设法定出不同类别的核心或初始类核,然后依据样本与这些核心之间的相似性度量将样本聚集成不同类别。

非监督学习同样包含众多算法,如 K-means、Apriori 算法、FP-growth 算法以及众多聚类方法。基于本文的工作所用到的非监督学习方法,下文将着重介绍聚类算法。

(1) 聚类算法

目前聚类算法有很多种,算法的选择取决于数据的类型、聚类的目的和应用。传统的聚类算法可以被分为五类层次方法、划分方法、基于密度方法、基于网格方法和基于模型方法^[22,23,24]。

1) 层次的方法(Hierarchical Method)

层次的方法对给定的数据对象集合进行层次的分解。按照层次的形成方式,层次方法可以分为凝聚的方法和分裂的方法^[25]。凝聚的方法,也称为自底向上的方法,一开始将每个对象都作为单独的一个类,然后相继地合并相近

的对象或类，直到所有的类合并成一个层次的最上层，或者达到一个终止条件。分裂的方法，也称为自顶向下的方法，一开始将所有的对象都置于同一个类中，然后通过不断的迭代，在迭代的每一步中，一个类被分裂为更小的类，最终直到每个对象被归入某个单独的类中，或者达到某个终止条件。

考虑了互连 CURE^[26]和 Chameleon^[27,28]是层次方法的两个有效的算法。解决了偏好球形和相似大小的问题，在处理孤立点上也更加健壮。它采用了一种新颖的层次聚类算法，该算法选择基于质心和基于代表对象方法之间的中间策略。它不用单个质心或对象来代表一个簇，而是选择数据空间中固定数目的具有代表性的点，并将这些点乘以一个适当的收缩因子，使它们更靠近簇的中心。选择多个代表使该算法可以适合非球状的几何形状，簇的收缩或凝聚可以有助于控制噪声的影响。同时，该算法采用了随机抽样与划分相结合来提高效率。对于大型数据库，它 also 具有良好的伸缩性，没有牺牲聚类质量，并且速度很快，对于容量为 n 的样本，时间复杂度是 $O(n^2)$ ，空间复杂度为 $O(n)$ 。CURE 不处理分类属性，收缩因子等参数设置对聚类结果有显著影响。

Chameleon 是一个采用动态模型的层次聚类算法。对于基于动态模型的算法，只要定义了相似度函数就可应用于所有类型的数据。其他算法往往忽略了簇间的互连性（互连性是指两个簇间交叉链的数目）或近似度，而 Chameleon 算法解决了这一问题。首先通过一个图划分算法将数据对象聚类为大量相对较小的子聚类，然后用一个凝聚的层次聚类算法，通过反复地合并子类来找到真正的结果簇。可以看出该算法从全局上是分裂的，而从局部上是凝聚的。它考虑了簇间的近似度，特别是用簇内部的特征，来确定最相似的子簇，这样它不依赖于一个静态的用户提供的模型，能够自动地适应被合并的簇的内部特征。Chameleon 比 CURE 在发现高质量的任意形状的聚类方面有更强的能力。但是，在最坏的情况下，其对含有 n 个对象的高维数据的复杂度是 $O(n^2)$ 。

2) 划分方法(Partitioning Method)

典型的划分方法有 K-means 方法和 K-中心点方法以及它们的改进算法。K-means 方法采用簇中对象的平均值作为参照点，而 K-中心点方法不采用簇中对象的平均值作为参照点，而是选用簇中位置最中心的对象作为参照点。因此，“噪声”和孤立点数据对 K-中心点方法的影响比方法小，但复杂度较高。典型的 K-中心点方法有 PAM 算法，CLARA 算法，CLARANS 算法，K-中心点方法在入侵检测等数据挖掘领域得到了广泛的应用。这两种方法都要求用户指定结果簇的数目 k 。

PAM 算法对“噪声”和孤立点数据不敏感，且由它所发现的簇与聚类数

据的输入顺序无关，能够处理不同类型的数据点。为了找到 k 个簇，PAM 算法为每一个簇定义一个代表对象，这个代表对象被称为中心点，它是簇内位置最中心的对象。当 k 个中心点选定以后，剩余的 $n-k$ (n 为对象的个数) 个非选中对象被归入离它最近的选中对象所代表的簇。为了找出 k 个中心点，PAM 算法首先随机地选择了 k 个对象。然后在每一步中，用一个非选中对象 o_h 替换一个选中对象 o_i ，只要这样的替换能够提高聚类质量，聚类质量（选择的中心点的总体质量）由对象和它所属簇的中心点之间的平均相异度量，相异度一般采用欧几里得公式计算。为了估算 o_h 与 o_i 之间替换的效果，PAM 算法为每一个非选中对象 o_j 计算代价 C_{jih} 。PAM 只适合处理小数据集，因为在 PAM 算法的每一次迭代，在计算每一个代表对象 $o_i (i=1, 2, \dots, k)$ 与非代表对象 $o_h (h=k+1, k+2, \dots, n)$ （共有 $k(n-k)$ 对）的替换代价时需要为每一个非选中对象 $o_j (j=k+1, k+2, \dots, n, j \neq h)$ 的计算代价 C_{jih} ，即 PAM 算法的每一步的时间复杂度 $O(k \times (n-k)^2)$ ，一般算法需要迭代若干步 t ，则 PAM 算法时间复杂度为 $O(t \times k \times (n-k)^2)$ ，很明显，当 n, k 很大时”算法的运行时间会大大增加。

CLARA 算法不是从整个数据集中发现代表对象，而是从数据集中抽取一部分样本，然后用替换代价的计算方法从样本中选择代表点。如果样本是以随机方式选取的，则样本的代表对象能近似代表整个数据集的代表对象。为了更好地达到近似，CLARA 算法对数据集进行多次取样，并将最好的样本聚类所得的代表对象作为算法的竞争结果。CLARA 算法每次随机抽取数目为 $40+2k$ 的对象数据进行聚类，可随机抽取样本 m 次，并将最好的聚类结果作为竞争。算法的每一次迭代，对取样对象选取代表对象的时间复杂度为 $O(k \times (40+k)^2)$ ，然后将数据集中的每个对象 o_j ，将它归入最近的代表对象，CLARA 算法的每一次迭代的时间复杂度为 $O(k \times (40+k)^2 + k \times (n-k))$ 。这种方法适合大数据集，但因为该方法是基于取样，聚类结果很容易陷入局部最小。如果最佳的中心点没有被包括在抽样集合中，CLARA 算法就无法得到最佳的聚类结果。

3) 基于密度的方法(Density-based Method)

绝大多数划分方法基于对象之间的距离进行聚类，这样的方法只能发现球状的类，而在发现任意形状的类上遇到了困难。基于密度的聚类方法的主要思想是只要临近区域的密度（对象或数据点的数目）超过某个闭值就继续聚类。也就是说，对给定类中的每个数据点，在一个给定范围的区域中必须至少包含某个数目的点。这样的方法可以用来过滤噪声孤立点数据，发现任

意形状的种类。

4) 基于网格的方法(Grid-based Method)

基于网格的方法把对象空间量化为有限数目的单元,形成一个网络结构。有的聚类都是在这个网络结构(即量化的空间)上进行。这种方法的主要优点是它的处理速度很快,其处理时间独立于数据对象的数目,只与量化空间中每一维的单元数目有关。基于网格的代表性的算法包括:

①STING 算法,它利用了存储在网格单元中的统计信息。

②WaveCluster 算法,它利用小波变换方法进行聚类。

③CLIQUE 算法,它是在高维数据空间中基于网格和密度的聚类方法。

这些算法都具有处理高维数据如空间数据库的能力,计算复杂度低,具有良好的伸缩性等优点。尤其是算法,它不仅能有效的处理大数据集,而且能够发现任意形状的种类,成功地处理孤立点,对输入的顺序不敏感,不要求指定聚类数目或邻域半径等输入参数。

5) 基于模型的方法(Model-based Method)

基于模型的方法为每个类假定了一个模型,算法主要是寻找数据对给定模型的最佳拟合。一个基于模型的算法可能通过构建反映数据点空间分布的密度函数来定位聚类,它基于标准的统计数字自动决定聚类的数目,考虑噪声数据和孤立点,从而产生健壮的聚类算法。

除上述五大类以外,还存在大量的聚类方法,如基于遗传算法的聚类方法,处理高维数据的聚类方法,处理动态数据的聚类方法,以及将基本聚类方法与各种新技术相结合的聚类方法等。

2.2.4 强化学习

强化学习(Reinforcement Learning, RL)作为近来机器学习领域一个研究热点,已经广泛应用于工业制造、仿真模拟、机器人控制、优化与调度、游戏博弈等领域。RL 的基本思想是通过最大化智能体(agent)从环境中获得的累计奖赏值,以学习到完成目标的最优策略。因此 RL 方法更加侧重于学习解决问题的策略。

RL 是一种从环境状态映射到动作的学习,目标是使 agent 在与环境的交互过程中获得最大的累计奖赏。马尔可夫决策过程(Markov Decision Process, MDP)可以用来对 RL 问题进行建模,通常将 MDP 定义为一个四元组 (S, A, ρ, f) , 其中:

①S为所有环境状态的集合, $s_t \in S$ 表示 agent 在 t 时刻所处的状态;

② A 为 agent 可执行动作的集合。 $a_t \in A$ 表示 agent 在 t 时刻所采取的动作；

③ $\rho: S \times A \rightarrow R$ 为奖赏函数。 $r_t \sim \rho(s_t, a_t)$ 表示 agent 在状态 s_t 执行动作 a_t 获得的立即奖赏值；

④ $f: S \times A \times S \rightarrow [0,1]$ 为状态转移概率分布函数。 $s_{t+1} \sim f(s_t, a_t)$ 表示 agent 在状态 s_t 执行动作 a_t 转移到下一状态 s_{t+1} 的概率。

在 RL 中，策略 $\pi: S \rightarrow A$ 是状态空间到动作空间的一个映射。表示为 agent 在状态 s_t 选择动作 a_t ，执行该动作并以概率 $f(s_t, a_t)$ 转移到下一状态 s_{t+1} ，同时接受来自环境反馈的奖赏 r_t 。假设未来每个时间步所获的立即奖赏都必须乘以一个折扣因子 γ ，则从 t 时刻开始到 T 时刻情节结束时，奖赏之和定义为

$$R_t = \sum_{t'=t}^T \gamma^{t'-t} r_{t'} \quad (6)$$

其中， $\gamma \in [0,1]$ ，用来权衡未来奖赏对累计奖赏的影响。

状态动作值函数 $Q^\pi(s, a)$ 指的是在当前状态 s 下执行动作 a ，并一直遵循策略 π 到情节结束，这一过程中 agent 所获得的累计回报表示为

$$Q^\pi(s, a) = E[R_t | s_t = s, a_t = a, \pi] \quad (7)$$

对于所有的状态动作对，如果一个策略 π^* 的期望回报大于或等于其它所有策略的期望回报，那么称策略 π^* 为最优策略。最优策略可能不止一个，但它们共享一个状态动作值函数：

$$Q^*(s, a) = \max_{\pi} E[R_t | s_t = s, a_t = a, \pi] \quad (8)$$

上式被称为最优状态动作值函数，且最优状态动作值函数遵循贝尔曼最优方程。即

$$Q^*(s, a) = E_{s' \sim s} \left[r + \gamma \max_{a'} Q(s', a') | s, a \right] \quad (9)$$

在传统的 RL 中，一般通过迭代贝尔曼方程求解 Q 值函数：

$$Q_{i+1}(s, a) = E_{s' \sim s} \left[r + \gamma \max_{a'} Q(s', a') | s, a \right] \quad (10)$$

其中，当 $i \rightarrow \infty$ 时， $Q_i \rightarrow Q^*$ 。即通过不断地迭代会使状态动作指函数最终收敛，从而得到最优策略： $\pi^* = \operatorname{argmax}_{a \in A} Q^*(s, a)$ ，然而对于实际问题来讲，

通过迭代式求解最优策略显然是不可行的，因为在大状态空间下，用迭代贝尔曼方程求解 Q 值函数的方法计算代价太大，针对此问题，在 RL 算法中，通常使用线性函数逼近器近似表示状态动作值函数，即 $Q(s, a | \theta) \approx Q^*(s, a)$ 。

随着人类社会的飞速发展，在越来越多复杂的现实场景任务中，需要利用深度学习(Deep Learning, DL)来自动学习大规模输入数据的抽象表征，并以

此表征为依据进行自我激励的 RL，优化解决问题的策略。由此，谷歌的人工智能团队 DeepMind 创新性地具有感知能力的 DL 和具有决策能力的 RL 相组合，形成了人工智能领域新的研究热点，即深度强化学习 (Deep Reinforcement Learning, DRL)。此后，在很多挑战性领域中，DeepMind 团队构造并实现了人类专家级别的 agent。这些 agent 对自身知识的构建和学习都直接来自原始输入信号，无需任何的人工编码和领域知识。因此，DRL 是一种端对端(End-to-End)的感知和控制系统，具有很强的通用性。其学习过程可以描述为：

①在每个时刻 agent 与环境交互得到一个高纬度的观察，并利用 DL 方法来感知观察，以得到具体的状态特征表示；

②基于预期回报来评价各动作的价值函数，并通过某种策略将当前状态映射为相应的动作；

③环境对此动作做出反应，并得到下一个观察。通过不断循环以上过程，最终可以得到实现目标的最优策略。DRL 的原理框架如图 8 所示：

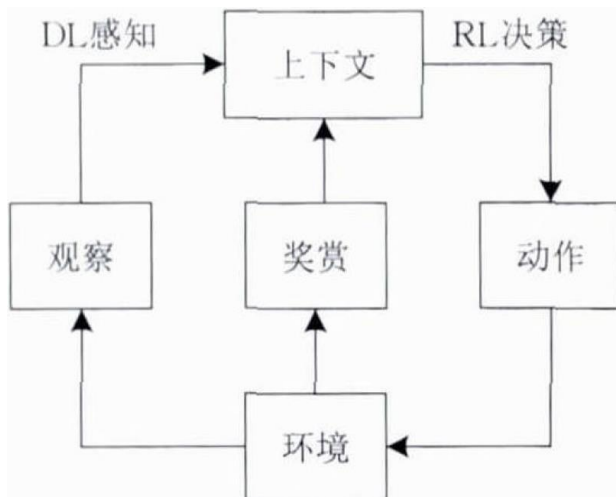


图 8 DRL 原理框架图

目前 DRL 技术在游戏、机器人控制、参数优化、机器视觉等领域得到了广泛的应用，并被认为是迈向通用人工智能(Artificial General Intelligence, AGI)的重要途径。

2.3 构建材料成分相图

2.3.1 相图

(1) 相图的基本概念

相图亦称平衡图或状态图，在金属学中，它是表示金属或合金系中各种相的平衡存在条件以及各相之平衡共存关系的一种简明图解。

(2) 相图的作用

相图可以帮助人们系统地了解金属和合金在不同条件下可能出现的各种组态，以及条件改变时，各种组态可能发生的转变的方向和限度。相图的直接作用仅限于此。但是，如果能把相图与相变的机理和动力学结合起来，那么，相图便可成为分析组织形成和变化的有力工具，它可将有关组织变化的极其复杂的现象抽象化、系统化。相图因此而成为金属材料的研制、生产和加工的重要参考资料。相图可用于指导热加工和制造参数的选择、预测材料性能、指导热处理工艺参数的选择、解决如材料失效分析材料性能问题以及其他多种用途。

2.3.2 相图的构建方法

(1) 相图传统的构建方法

构建相图的传统方法是每次获取单个样品的合成与表征，这种成本较高、费时较长且不系统的方法在科技加速的当今远不能满足科研工作者的需求。

(2) 相图构建方法的现状

自 19 世纪 60 年代，就有科研人员尝试寻找更有效建立成分-相关系的方法。组合材料芯片技术、组合材料库的高通量合成与快速表征对于加速材料筛选和优化显示出巨大的潜力。

Yoo^[29]等指出在传统的体合金相图中，每次只能制备和表征单个离散的点，因此每个样品点具有不同的热力学背景。连续相图(Continuous Phase Diagram, CPD)为把薄膜相图的高通量构建与一个能应用于所有成分点的通用热力学背景联系起来提供了一个独一无二的机会。

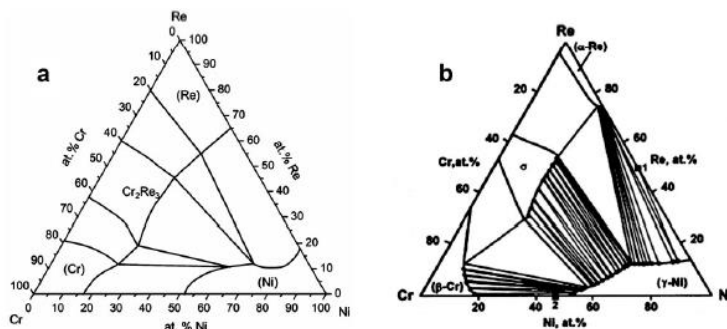


图 9 左图 1100°C 下 Cr-Ni-Re 薄膜相图，右图为文献中 1150°C 下体合金相图

Janghorban^[30]等为了建立 Cr-Ni-Re 三元合金薄膜材料相图联合使用组合制备技术与高通量表征技术，薄膜的成分覆盖了全部的成分范围。在 940 和 1100℃ 下退火后，相平衡达到三元相图中富 Ni 区域，这一结果与已发表文献中的体相图一致。940℃ 下退火的材料库不足以达到三元相图中的富 Re 区域，但是 1100℃ 下退火的材料库形成了预期中的相（能观测到三个固溶体和一个紧密堆积的化合物）。其文章中完整的建立了 1100℃ 下 Cr-Ni-Re 薄膜相图，且与体相图的结果一致。这意味着组合薄膜相图技术对于昂贵或高熔点的成分来说是极其值得尝试的方法。

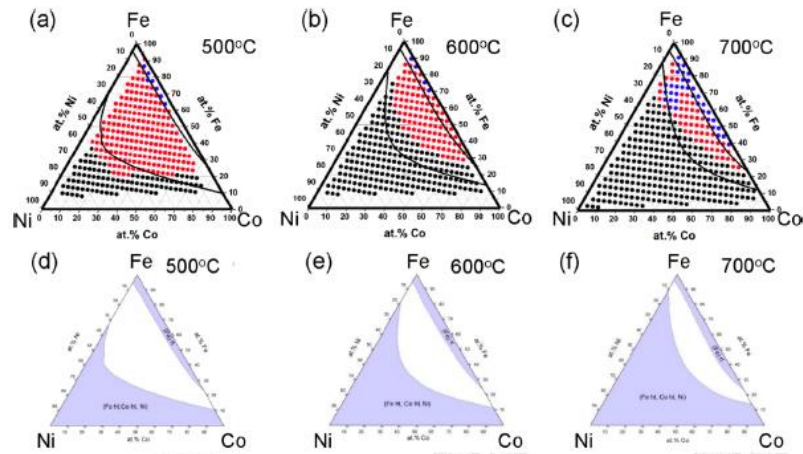


图 10 由 500、600、700℃ 下热处理的材料组合芯片构建的相图与文献中结果对比

Xing^[31]等通过磁控共溅射制备 100nm 厚的 Fe-Co-Ni 材料芯片并在 500, 600, 700℃ 下绝热退火。逐点的成分和结构匹配由同步光源的微束 X 射线表征。衍射谱的图像按照每秒一张的速率被记录下来。XRD 衍射图谱被自动处理，相的识别与归类由层次聚类的算法完成，并用来构建组合成分相图。所构建的相图与 ASM 合金相图数据库记载的绝热章节内容一致，验证了其构建相图方法的有效性。

2.3.3 目前存在的不足与改进方案

(1) XRD 曲线特征峰峰位偏移

解决方案：采用统计学的方法统计峰位偏移的距离 d ，将原有曲线平移 d 个单位距离；

采用自适应的峰位调节；

(2) 算法模型使用范围较小

解决方案：使用表达能力更强、复杂度更高的模型去映射材料成分与晶

体结构之间的关系；

尝试使用多种模型，比较各种模型的适用范围；

2.4 预测碳钢的大气腐蚀模型

2.4.1 碳钢的大气腐蚀

大气腐蚀是指金属材料在大气环境下，干湿交替过程中发生的腐蚀现象。在金属遭受的腐蚀中，大气腐蚀是最普遍和严重的，有资料显示，大气腐蚀造成的金属损失约占腐蚀总损失量的 50%以上。因此，研究钢材表面锈层的生长规律对钢材的寿命预测具有重要意义。

2.4.2 碳钢大气腐蚀的预测模型

目前常见的大气腐蚀模型有如下几种：

① $D = AT^n$ 模型^[32]

式中，D 表示年平均腐蚀深度(um/a)；

T 表示暴露时间(a)，特别地当 $T=1a$ 时， $D=A$ ，即 A 具有第一年腐蚀量的意义。

② $D = AT^B e^{C/T}$ 模型^[33]

式中，D 表示年平均腐蚀深度(um/a)；

T 表示暴露时间(a)；

A, B, C 表示待定系数。

③ GM(1,1)模型^[34]

灰色系统的 GM(1,1)模型能对一系列数据建立一个或一群数学函数，模型的数学表达式为：

$$F(k+1) = Ae^{-\eta k} + B$$

式中，F 表示 k-1, k-2 个对应时间的年平均腐蚀深度之和(um/a)；

A, B, n 为待定系数。将上式进行变换，可得：

$$D = A(1 - e^n)e^{(1-T)}$$

GM(1,1)模型要求各处理点的实验时间间隔相等，以实现能够预测非处理点的功能。

④ 线性模型^[35]

$$E = A + B \times Cl^- + C \times SO_2 + D \times TOW$$

式中， D 表示年平均腐蚀深度($\mu\text{m/a}$)；

A , B , C , D 表示待定系数。

Cl^- : 沉积速率($\text{mg/m}^2\text{d}$)

SO_2 : 沉积速率($\text{mg/m}^2\text{d}$)

TOW : 润湿时间

⑤ 人工神经网络模型^[36]

详情请见本章 2.2.2 节

传统的预测模型多是预测值与期望值平方差之和最小的回归模型。这些模型仅在非常有限的范围内有效，难以捕捉腐蚀过程的非线性变化。因此，科研人员一直在寻找能够在更广范围的气候和污染水平状况下，预测大气腐蚀速率的数学模型。人工神经网络模型因在合理选择网络结构和参数的情况下具备模拟任何复杂模型的能力，最近在腐蚀研究领域腐蚀模型预测方面引起了一番热潮，具体原理可见本章 2.2.2 节。

2.4.3 目前存在的不足与改进方案

(1) 模型仅在小数据范围内有效

解决方案：扩大碳钢大气腐蚀模型训练的数据集。

(2) 模型难以捕捉腐蚀过程的非线性变化，表达能力有限

解决方案：使用包含更多变量、规模更大的模型去拟合碳钢大气腐蚀的实际过程。例如使用深度神经网络。

2.5 探究不锈钢焊接接头成分、硬度与腐蚀性能的联系

2.5.1 不锈钢焊接接头研究的必要性

不锈钢自上世纪初问世，到现在已有近 100 年的历史。由于石油、化工、能源及原子能、宇航、海洋开发等尖端技术的迅速发展，对不锈钢提出了更高的综合性能要求，不仅要求有良好的耐蚀性，还要求有高强度、耐高温高压、防辐射、耐低温等性能，使不锈钢的品种类型得到进一步的开拓。特别是从上世纪 70 年代以来，不锈钢的工业精炼技术普遍采用 AOD(氩氧脱碳)和 VOD(真空氧脱碳)法，为新型钢种的开发提供了良好的工艺条件。在过去的 10 年，中国不锈钢消费量以年平均超过 20% 的速度增长。然而，我国人均不锈钢消费量仅有 4.1kg，仍然非常低，这表明我国不锈钢消费量仍有很大的增长空间。目前，中国已经成为世界不锈钢市场发展的中心。

不锈钢由于具备良好的物理、化学性能，它的生产和应用在深度和广度上不断发展和扩大。不锈钢的使用范围十几年来向轻工和家庭生活方面迅速扩展。由于不锈钢的使用过程中，很大一部分都要通过焊接的方式将构件进行连接，因此焊接质量的好坏直接影响着不锈钢构件的性能。在不锈钢构件的腐蚀失效分析中，焊接接头所受到的腐蚀是整个构件腐蚀当中最严重的。因此深入研究焊接接头的耐蚀性能对不锈钢构件的安全使用和寿命评估起着相当重要的作用。当两种不同金属焊接时，因电焊造成的高温会导致焊缝区和热影响区材料的成分和组织发生较大的变化，从而为高通量的材料成分分布提供了温床。

2.5.2 不锈钢的焊接性和焊接特点

由于不锈钢在使用过程中，很大一部分都要焊接，因此，焊接质量的好坏直接影响着不锈钢件的性能。焊接性是指同种金属材料或异种金属材料在焊接加工条件下，能够形成具备一定使用性能的焊接接头特性。焊接性受到材料、焊接方法、构件类型及使用要求等因素的影响。焊接性包括两个方面的意义：一是结合性，即一定的金属材料在指定的焊接工艺下，对焊接缺陷的敏感性，即工艺焊接性；二是使用性能，是金属材料在指定的焊接条件下所形成的焊接接头适用使用条件的程度，也称使用焊接性。不锈钢焊接接头有焊缝金属区、热影响区和母材（基本金属）区三部分组成，图 11 给出了不锈钢焊接接头的组织和性能变化示意图。

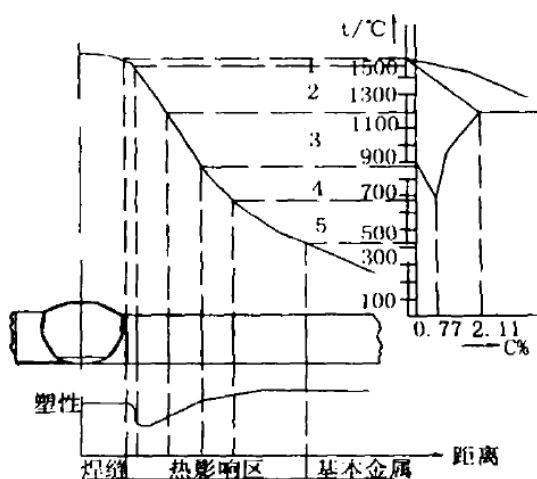


图 11 不锈钢焊接接头的组织和性能变化示意图

1——半熔化区；2——过热区；3——正火区；4——部分相变区；5——

再结晶区。

①焊缝金属区(WM): 焊接加热时, 焊缝处的温度在液相线以上, 基本金属和填充金属共同形成熔池, 冷凝后成为铸态组织。焊缝金属区由于冷却速度较快, 晶粒较细小, 再加上焊条药皮的保护作用和合金化作用, 其强度一般不低于母材区。

②热影响区(HAZ): 是指焊缝两侧因局部加热而产生组织和性能变化的区域, 以下着重讨论半熔化区和过热区对焊接接头的不良影响。半熔化区, 位于焊缝和基本金属的交界处, 加热温度约为 $1490\sim 1530^{\circ}\text{C}$ 。部分金属加热到融化状态, 结晶后成为铸态组织; 另一部分金属为严重过热的奥氏体, 冷却后成为粗晶粒, 故该区域的塑形和冲击韧性很差。过热区, 紧靠半熔化区, 加热温度约为 $1100\sim 1490^{\circ}\text{C}$, 在如此高温作用下, 奥氏体晶粒急剧长大, 形成过热组织, 使塑形和韧性大大降低。

③母材区(BM): 由于焊接接头试样取样长度有限, 母材区在焊接过程中也不同程度的受到焊接热输入的影响。因此, 母材区的晶粒与原热轧板材有些区别, 等轴化比较明显、组织相对均匀一些, 晶粒尺寸有所长大。

2.5.3 不锈钢焊接接头腐蚀性能的研究现状

不锈钢在各个行业越来越广泛的应用, 促使人们对不锈钢以及不锈钢焊接接头的耐腐蚀性能进行更深入的研究。近些年来, 对不锈钢以及不锈钢焊接接头的研究主要有以下四个方面: ①焊接工艺对不锈钢焊接接头的性能的影响; ②不锈钢焊接接头抗晶间腐蚀和点蚀的研究; ③不锈钢焊接接头的腐蚀微区电化学研究; ④不锈钢焊接接头的微观结构和组织与性能的关系。通过查阅文献, 国内外的相关研究者所做的与不锈钢焊接接头的腐蚀微区电化学研究相关的工作如下:

骆鸿, 魏丹, 董超芳^[37]介绍了电化学扫描探针显微镜的基本原理, 重点对微区电化学技术在金属腐蚀研究中的应用进行了阐述, 研究发现: 在金属-电解质溶液界面, 阳极区和阴极区可能会表现出不同的特征, 常规的电化学方法无法深入研究其腐蚀行为和机理, 微区电化学技术正好弥补了这方面的不足。

陈志林, 陈国强^[38], 李军等通过电化学方法研究了钎料为 $60\%\text{BNi}_2+40\%\text{BNi}_5$ (质量分数) 和 BNi_2 的 316L 不锈钢钎焊接头在人造海水中的耐腐蚀性能。结果表明: 母材、 $60\%\text{BNi}_2+40\%\text{BNi}_5$ 钎料和 BNi_2 钎料这

三者的钝化电流相近,进入钝化的电位依次增加,表明母材的耐腐蚀性最好,60%BNi₂+40%BNi₅ 钎料次之,BNi₂ 钎料最差。Cr 含量的不同是造成三者耐腐蚀性存在差异的主要原因。

C.Garcia 等^[39]运用动态交流阻抗谱方法和动电位阳极极化方法对奥氏体不锈钢焊接构件在海水中的抗点蚀性能进行了评估。结果表明:焊接热影响区是不锈钢构件发生点蚀的最严重的部分,这与其成分和微观组织特征有着直接的关系。

金维松,郎宇平,荣凡等^[40]利用电化学动电位再活化(即 EPR)法检测了奥氏体不锈钢晶间腐蚀敏感性的程度.研究了不同条件对敏化度的影响,特别是不同敏化时间、敏化温度、扫描速度、溶液温度对奥氏体不锈钢晶间腐蚀性能的影响。结果表明:双环 EPR 法能定量、定性的检测不锈钢的晶间腐蚀敏感性。对 304L 不锈钢而言,敏化时间越长,温度越高,其晶间腐蚀敏感性越强。扫描速度越快,溶液温度越低,同一材料的 Ra 值越小。

林晓云等^[41]介绍了奥氏体不锈钢焊接接头抗晶间腐蚀能力的检验和评定方法,对焊接接头晶间腐蚀产生的原因进行了分析,阐述了在焊接过程中实施的有效控制方法,以提高不锈钢焊件抗晶间腐蚀的能力。结果表明:不锈钢焊接接头的晶间腐蚀是由于在 450~850℃这一危险温度范围内,碳化物在晶界沉淀,因而在晶界及邻近区域形成贫铬区,从而大大的降低了晶界的耐蚀性能。

2.5.4 目前存在的不足与改进方案

(1) 焊接接头实验区域较小,少有科研人员进行大量的实验用于定量模型的构建

解决方案:在焊接接头有限的实验区域内进行微区实验(如:微区电化学实验),可进行较多组实验。

(2) 关于焊接接头成分、硬度与腐蚀性能之间的定性关系的描述较多,定量关系或相关预测模型的工作较少

解决方案:尝试构建新的模型来映射焊接接头成分、硬度与腐蚀性能之间的关系。

3 研究方案

3.1 研究内容

(1) 通过层次聚类快速构建 Fe-Cr-Ni 成分相图

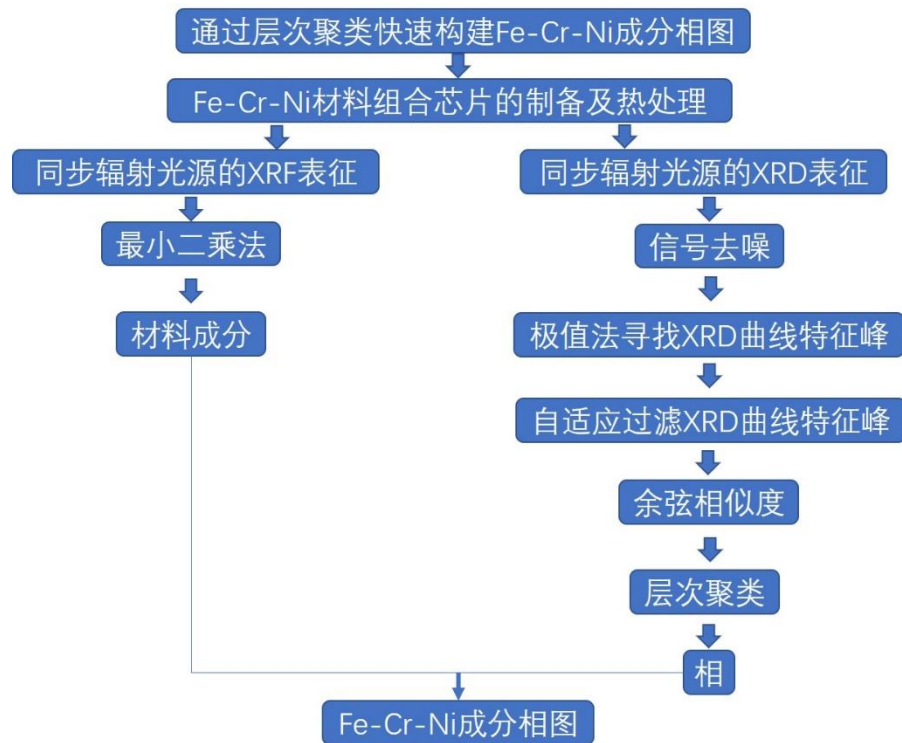


图 12 通过层次聚类快速构建 Fe-Cr-Ni 成分相图技术路线

①利用降噪技术去除实验数据噪声，利用最小二乘法计算材料成分；
 ②利用聚类方法识别材料晶体结构与成分之间的联系，以构建相图；
 ③将实验构建的相图与已发表文献中的相图进行对比，验证实验结果的正确性。

(2) 利用多种神经网络预测碳钢的大气腐蚀

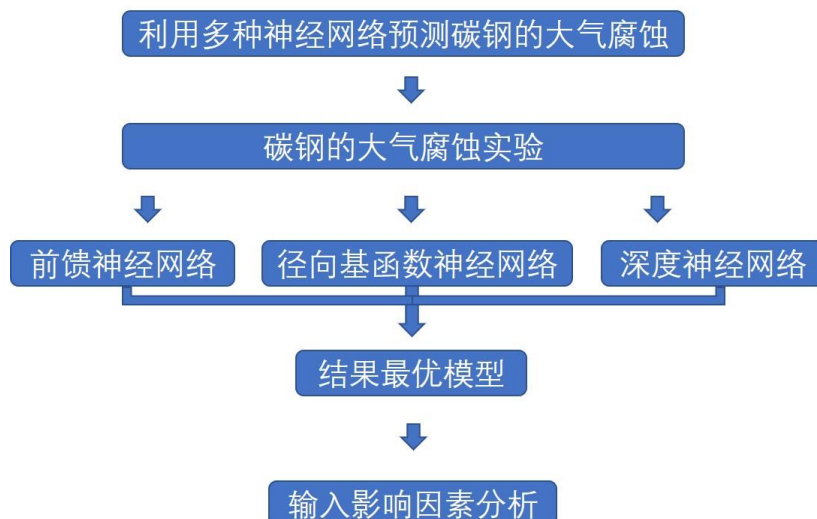


图 13 利用多种神经网络预测碳钢的大气腐蚀技术路线

- ①搜集包含碳钢大气腐蚀数据的文献，并提取有效腐蚀数据；
- ②选择合适的神经网络构建大气腐蚀模型；
- ③将数据带入模型训练，观察模型的输出结果；
- ④对输出结果最佳的模型进行输入因素（腐蚀条件）影响分析。
- ⑤确定最终的大气腐蚀模型。

（3）利用深度强化学习建立不锈钢焊接接头成分、硬度与腐蚀性能的联系

系

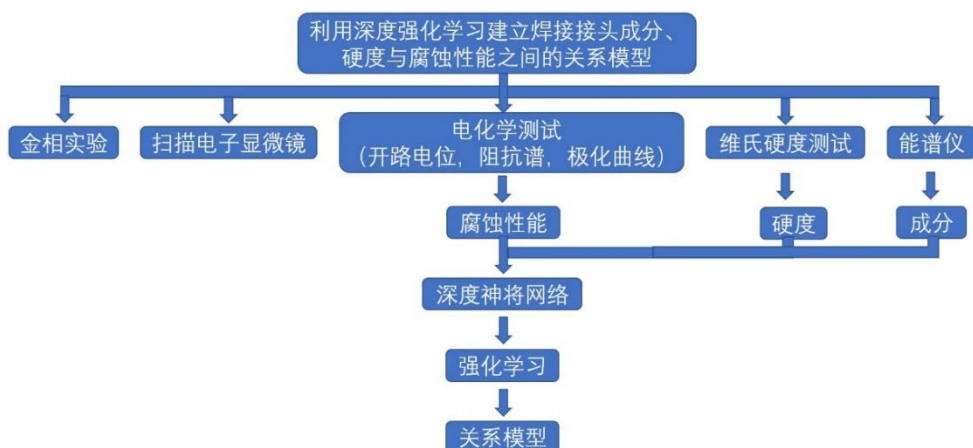
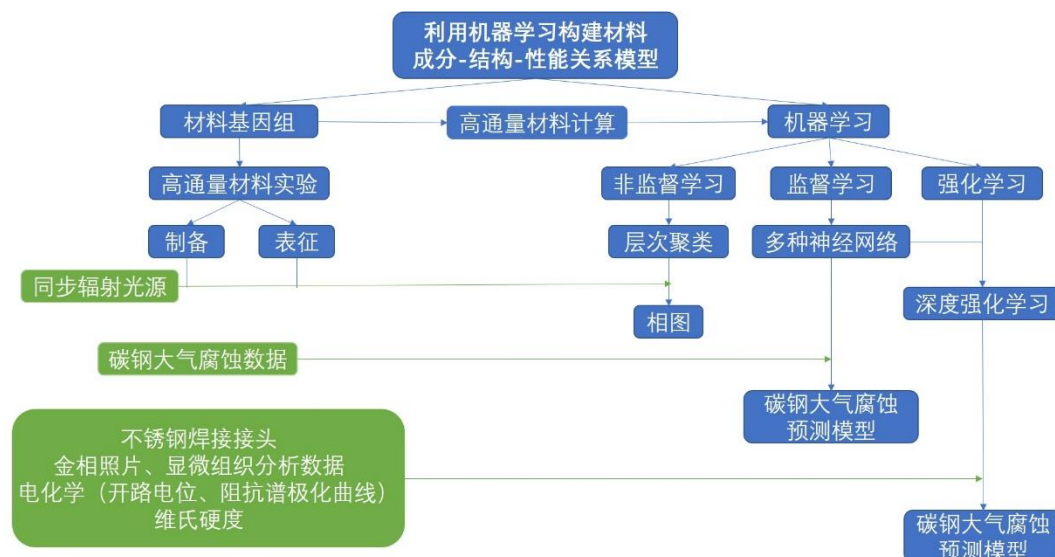


图 14 利用深度强化学习建立不锈钢焊接接头成分、硬度与腐蚀性能的联系技术路线

- ①通过实验获取不锈钢焊接接头的成分、结构信息；
- ②选择深度学习和强化学习的具体结合方式；
- ③将数据输入深度强化学习模型，对模型进行训练；

④根据模型结果得到焊接接头成分、结构与腐蚀性能的联系。

3.2 研究方法



(1) 通过层次聚类快速构建 Fe-Cr-Ni 成分相图

①利用 X 射线荧光、衍射分析技术分析实验数据, 得到材料成分、相;

②利用机器学习聚类方法中的层次聚类根据材料实验数据的相似性进行

③根据相图原则和相图绘制规定绘制相图。

(2) 利用多种神经网络预测碳钢的大气腐蚀

①根据搜集的文献确定对碳钢大气腐蚀的主要影响因素;

②利用神经网络构建腐蚀预测模型:

③利用输入因素分析技术（例：Fuzzy Curve Analysis）分析各项腐蚀条件对碳钢大气腐蚀的影响。

(3) 利用深度强化学习建立不锈钢焊接接头成分、硬度与腐蚀性能的关系模型

①利用 EBSD、EPMA 等实验确定焊接接头的成分、结构;

②利用电化学实验确定焊接接头的腐蚀性能:

③利用深度强化学习构建预测模型:

④将焊接接头的成分、结构信息作为预测模型的输入，将模型的预测输出的腐蚀性能与实验得到的腐蚀性能进行对比。

3.3 创新点

(1) 通过层次聚类快速构建 Fe-Cr-Ni 成分相图

①对 XRD 数据采用去噪处理, 避免遗漏材料在 XRD 曲线中的特征峰;

②采用自适应的策略解决 XRD 曲线峰位偏移的问题, 减少人工干预影响;

③可能尝试不同相似度的度量方法(皮尔逊相关系数、Jensen-Shanon);

④在材料 XRD 曲线的相似度计算中引入成分信息。

(2) 利用多种神经网络预测碳钢的大气腐蚀

①收集得到的腐蚀数据量较大(2000~3000 条);

②腐蚀预测模型(径向基函数神经网络, 深度神经网络)较为新颖;

(3) 利用深度强化学习建立不锈钢焊接接头成分、硬度与腐蚀性能的关系模型

①利用焊接接头的成分、结构信息预测其腐蚀性能的模型较少

②深度强化学习对线性/非线性的模型都具备强大的表征能力, 因技术出现在近几年, 目前较少在材料研究领域中出现应用, 值得尝试;

3.4 预期目标

(1) 通过层次聚类快速构建 Fe-Cr-Ni 成分相图

①获得 Fe-Cr-Ni 成分相图;

②构建完整的自动化实验数据处理、分析流程。

(2) 利用多种神经网络预测碳钢的大气腐蚀

①获得能根据腐蚀环境、材料成分精准预测碳钢大气腐蚀速率的预测模型;

②分析得到腐蚀环境和材料成分对于碳钢大气腐蚀的影响。

(3) 利用深度强化学习建立不锈钢焊接接头成分、硬度与腐蚀性能的关系模型

①获得能根据不锈钢焊接接头成分、结构精准预测不锈钢焊接接头腐蚀性能的预测模型;

②分析得到不锈钢焊接接头成分、结构对于不锈钢焊接接头腐蚀性能的影响。

3.5 进度安排

(1) 通过材料组合芯片技术层次聚类快速构建 Fe-Cr-Ni 成分相图

2018/08-2018/10

去除 XRD 曲线噪声, 提取 XRD 曲线有效特征峰的峰位和强度。将 XRD 曲线转换成对应的特征向量, 计算特征向量的相似度 (余弦相似度)。

2018/10-2019/02

用层次聚类根据相似性将 XRD 曲线分为 11 类, 与标准相图上的 11 中相对应。

2019/03-2019/08

结合聚类结果和成分信息构建相图。

2019/09-2019/12

针对可能存在的问题进行补充研究, 撰写论文。

(2) 利用多种神经网络预测碳钢的大气腐蚀

2018/08-2018/09

搜集包含碳钢大气腐蚀的文献, 统一腐蚀数据格式, 搭建三种神经网络。

2019/01-2019/03

将实验数据带入神经网络进行训练、调试神经网络参数。

2019/04-2019/08

确定最优神经网络模型, 分析输入因素对碳钢腐蚀性能的影响。

2019/09-2019/12

针对可能存在的问题进行补充研究, 撰写论文。

(3) 利用深度强化学习建立不锈钢焊接接头成分、硬度与腐蚀性能的关系模型

2018/08-2018/10

学习深度强化学习的具体方法, 查阅相关论文。从原始实验数据中提取有效信息。

2019/01-2019/04

搭建深度强化学习模型。

2019/05-2019/08

将材料成分、硬度数据带入模型进行训练, 调整模型参数。

2019/09-2019/12

确定模型结构, 针对可能存在的问题进行补充研究, 撰写论文。

参考文献

- [1] 黄乾尧, 李汉康, 陈国良. 高温合金[M]. 北京: 冶金工业出版社, 2000
- [2] 王海舟, 汪洪, 丁洪, 等. 高通量材料实验与表征[J]. 科技导报, 2015, 33(10): 31-49. WANG Hai-zhou, WANG Hong, DING Hong, et al. Highthroughput experimentation for materials genome[J]. Science & Technology Review, 2015, 33(10): 31-49.
- [3] 向勇, 闫宗楷, 朱焱麟, 等. 材料基因组技术前沿进展[J]. 电子科技大学学报, 2016, 45(4): 634-649.
- [4] MCCLUSKEY P J, VLASSAK J J. Glass transition and crystallization of amorphous Ni-Ti-Zr thin films by combinatorial nano-calorimetry[J]. Scripta Materialia, 2011, 64(3): 264-267. MCCLUSKEY P J, VLASSAK J J. Combinatorial nanocalorimetry[J]. Journal of Materials Research, 2011, 25(11): 2086-2100.
- [5] XIANG X D, SUN X, BRICENO G, et al. A combinatorial approach to materials discovery[J]. Science, 1995, 268(5218): 1738-1740.
- [6] COOPER J S, ZHANG G H, MCGINN P J. Plasma sputtering system for deposition of thin film combinatorial libraries[J]. Review of Scientific Instruments, 2005, 76: 062221
- [7] GAO C, DUEWAR F, XIANG X D. Quantitative microwave evanescent microscopy[J]. Applied Physics Letters, 1999, 75(19): 3005-300.
- [8] VersaSCAN. Electrochemical scanning system[EB/OL]. [2016-01-28]. <http://www.princetonappliedresearch.com/Our-products/Electrochemical-Scanning-System/index.aspx>.
- [9] 杨小渝, 任杰, 王娟, 等. 基于材料基因组计划的计算和数据方法[J]. 科技导报, 2016, 34(24): 62-67.
- [10] Raccuglia P, Elbert K C, Adler P D F, et al. Machine-learning-assisted materials discovery using failed experiments[J]. Nature, 2016, 533(7601): 73.
- [11] Fischer C C, Tibbetts K J, Morgan D, et al. Predicting crystal structure by merging data mining with quantum mechanics[J]. Nature materials, 2006, 5(8): 641.
- [12] Hautier G, Jain A, Ong S P, et al. Phosphates as Lithium-Ion Battery Cathodes: An Evaluation Based on High-Throughput Ab Initio[C]//Lithium Battery Discussion Electrode Materials. 2011.

- [13] Saal J E, Kirklin S, Aykol M, et al. Materials design and discovery with high-throughput density functional theory: the open quantum materials database (OQMD)[J]. Jom, 2013, 65(11): 1501-1509.
- [14] Kirklin S, Meredig B, Wolverton C. High - Throughput Computational Screening of New Li - Ion Battery Anode Materials[J]. Advanced Energy Materials, 2013, 3(2): 252-262.
- [15] Saal J E, Wolverton C. Thermodynamic stability of Mg-based ternary long-period stacking ordered structures[J]. Acta Materialia, 2014, 68: 325-338.
- [16] Hu L H, Wang X J, Wong L H, et al. Combined first-principles calculation and neural-network correction approach for heat of formation[J]. The Journal of Chemical Physics, 2003, 119(22): 11501-11507.
- [17] Bligaard T, Jóhannesson G H, Ruban A V, et al. Pareto-optimal alloys[J]. Applied physics letters, 2003, 83(22): 4527-4529.
- [18] 何华灿. 人工智能基础理论研究的重大进展--评钟义信的专著《高等人工智能原理》[J]. 智能系统学报, 2015(1):163-166.
- [19] 杨淑莹, 张桦. 模式识别与智能计算:MATLAB 技术实现.第 3 版[M]. 电子工业出版社, 2015.
- [20] 高隽. 人工神经网络原理及仿真实例.第 2 版[M]. 机械工业出版社, 2007.
- [21] 叶晨洲, 廖金周. 车辆牌照字符识别系统[J]. 计算机系统应用, 1999, 8(5):10-13.
- [22] JiaweiHan, MichelineKamber, JianPei,等. 数据挖掘:概念与技术[M]. 机械工业出版社, 2012.
- [23] 汤效琴, 戴汝源. 数据挖掘中聚类分析的技术方法[J]. 微计算机信息, 2003(1):3-4.
- [24] 唐勇智. 基于聚类的 RBF-LBF 串联神经网络的学习算法及其应用[D]. 江南大学, 2009.
- [25] Voorhees E M. Implementing agglomerative hierarchic clustering algorithms for use in document retrieval ☆[J]. Information Processing & Management, 1986, 22(6):465-476.
- [26] Guha S, Rastogi R, Shim K, et al. CURE : An Efficient Clustering Algorithm for Large Databases[C]// Elsevier Ltd, 1998:73-84.
- [27] Karypis G, Han E H, Kumar V. CHAMELEON A hierarchical clustering algorithm using dynamic modeling[J]. Computer, 2008, 32(8):68-75.
- [28] Ng R T, Han J. Efficient and Effective Clustering Methods for Spatial Data

- Mining[J]. In: Proceedings of the 20th international, 1994, 88(9):144--155.
- [29] Yoo Y K, Xue Q, Chu Y S, et al. Identification of amorphous phases in the Fe–Ni–Co ternary alloy system using continuous phase diagram material chips[J]. Intermetallics, 2006, 14(3):241-247.
- [30] Janghorban A, Pfetting-Micklich J, Frenzel J, et al. Investigation of the Thin - Film Phase Diagram of the Cr - Ni - Re System by High - Throughput Experimentation[J]. Advanced Engineering Materials, 2014, 16(5):588-593.
- [31] Hui X, Zhao B, Wang Y, et al. Rapid Construction of Fe–Co–Ni Composition-Phase Map by Combinatorial Materials Chip Approach[J]. Acs Combinatorial Science, 2018.
- [32] Panchenko Y M, Marshakov A I. Prediction of First-Year Corrosion Losses of Carbon Steel and Zinc in Continental Regions.[J]. Materials, 2017, 10(4):422.
- [33] Institution B S. Corrosion of metals and alloys. Corrosivity of atmospheres. Guiding values for the corrosivity categories[J].
- [34] 唐其环. 用 GM(1 1)模型拟合大气腐蚀数据[J]. 腐蚀与防护, 1993(1).
- [35] Chico B, Fuente D D L, Díaz I, et al. Annual Atmospheric Corrosion of Carbon Steel Worldwide. An Integration of ISOCORRAG, ICP/UNECE and MICAT Databases[J]. Materials, 2017, 10(6).
- [36] Pintos S, Queipo N V, Rincón O T D, et al. Artificial neural network modeling of atmospheric corrosion in the MICAT project[J]. Corrosion Science, 2000, 42(1):35-52.
- [37] 骆鸿, 魏丹, 董超芳,等. 金属腐蚀微区电化学研究进展(1)扫描电化学显微技术[J]. 腐蚀与防护, 2009, 30(7):437-441.
- [38] 陈志林, 陈国强, 李军,等. 不锈钢钎焊接头电化学行为的研究[J]. 热加工工艺, 2008, 37(5):16-17.
- [39] Lee D J, Jung K H, Sung J H, et al. Pitting corrosion behavior on crack property in AISI 304L weld metals with varying Cr/Ni equivalent ratio[J]. Materials & Design, 2009, 30(8):3269-3273.
- [40] 金维松, 郎宇平, 荣凡,等. EPR 法评价奥氏体不锈钢晶间腐蚀敏感性的研究[J]. 中国腐蚀与防护学报, 2007, 27(1):54-59.
- [41] 林晓云. 不锈钢焊接接头晶间腐蚀性能的控制[J]. 中国修船, 2007, 20(1):37-39.

