

DOI:10.19344/j.cnki.issn1671-5276.2026.03.052

基于增量式 Q 学习的硅材料电火花线切割控制方法研究

熊峰, 罗福源, 苏圣霖, 王开阳

(南京航空航天大学 机电学院, 江苏 南京 210016)

摘 要: 硅材料等半导体材料的加工制造是国家电子信息制造业的基石, 半导体材料的电火花加工是其中的热点之一。以硅材料工件为研究对象, 针对电火花加工过程提出一种增量式 Q 学习的控制方法, 通过与环境交互获得数据从而实现控制器的训练与学习, 使加工过程中放电间隙维持稳定。通过神经网络建立进给速度与电流脉冲概率相互关系的数学模型。仿真和实验表明: 该算法相对于传统依赖于经验公式的控制方法有更好的加工效果。

关键词: Q 学习; 神经网络; 数据驱动; 硅材料; 电火花线切割

中图分类号: TP391.9 **文献标志码:** B **文章编号:** 1671-5276(2026)03-0266-05

Research on Control Method of WEDM of Silicon Material Based on Incremental Q -Learning

XIONG Feng, LUO Fuyuan, SU Shenglin, WANG Kaiyang

(College of Mechanical and Electronic Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: As the processing and manufacturing of semiconductor materials such as silicon is the cornerstone of the national electronic information manufacturing industry, and EDM for semiconductor materials is one of the hot spots, this paper, taking silicon material workpiece as the research object, proposes an incremental Q -learning control method for the EDM process. Data is obtained by interacting with the environment to realize the training and learning of the controller, maintaining the stability of the discharge gap in the machining process. The mathematical model of the relationship between feed rate and current pulse probability is established by neural network. Simulation and experiments show that the proposed algorithm has better processing effect than the traditional control method dependent on empirical formula.

Keywords: Q -learning; neural network; data-driven; silicon materials; WEDM

0 引言

在科技高速发展的今天, 半导体材料在集成电路、航空航天、新能源等领域已经得到广泛应用, 因此如何更高效地加工半导体材料以及提高其加工质量成为机械加工领域所面临的问题之一^[1]。常见的加工方式存在机械应力, 硅晶体易崩碎、断裂^[2]等问题。而电火花线切割 (wire electrical discharge machining, WEDM) 凭借其高加工精度、低切割损耗、低制造成本等特点, 在机械加工生产中得到了广泛的应用^[3]。但由于硅晶体半导体不同于金属的放电性质, 传统电火花加工伺服控制系统应用于硅材料加工时易出现过跟踪及欠跟踪现象, 硅晶体放电加工持续性受到影响^[4]。

在电火花线切割加工过程中最重要的是保证

合适的放电间隙。电源参数 (脉宽、脉间、极性) 等)、伺服控制参数 (进给速度、伺服电压) 等)、工艺参数 (工件材料、切削液浓度和种类) 等) 对电火花线切割过程的间隙状态有着复杂的影响^[5-7]。在硅材料电火花线切割加工过程中, 由于半导体的体电阻作用, 在短路状态和正常加工状态电压变化不明显, 难以用常规方法进行加工状态的判断^[8]。通常情况下, 对硅材料进行电火花加工的参数调节都是依赖人工或经验公式, 例如构造电流放电概率与进给速度的关系式来实现加工速度的伺服控制。然而, 电火花线切割过程复杂, 实际上难以搭建精准的数学模型, 这就导致传统控制方法适应性差。随着人工智能时代的来临, 强化学习作为无监督的智能算法被学者越来越重视。事实证明可在面对无模型或复杂模型的场景下, 强化学习算法, 展现出极强的适应性, 因此被广泛

第一作者简介: 熊峰 (2000—), 男, 安徽六安人, 硕士研究生, 研究方向为数控技术与装备自动化控制, 18551717262@163.com。

通信作者简介: 罗福源 (1977—), 男, 广西桂林人, 副教授, 硕士生导师, 研究方向为数控技术与装备自动化控制, fuyuanluo@126.com。

运用于机器人控制、路径规划等场景。王立群等^[9]针对传统柔性作业车间调度依赖人工的情况,将车间调度转化为序列决策的问题,将深度学习与强化学习相结合对决策序列进行优化,并通过仿真验证该算法能够降低完工时间并均衡机器负载。

本文采用基于数据驱动的增量 Q 学习方法对硅材料电火花线切割过程进行伺服控制,以期获得更好的控制适应性与更优的加工质量。

1 电流脉冲概率检测原理

在电火花线切割加工硅锭的过程中,根据其电流电压脉冲的不同大致可以分为3种状态,即开路、正常放电和短路^[10-12],在硅材料切割过程中,短路状态仍可以进行放电加工,其电压信号与正常放电相比没有较大区别。加工过程中电压和电流波形如图1所示。

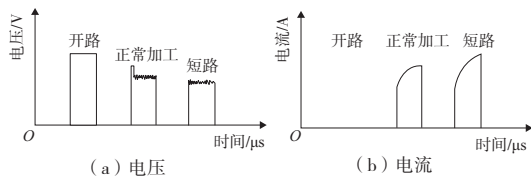


图1 硅材料电加工电特性

电流脉冲概率就是在一定周期时间内参与放电的电流脉冲个数与电压脉冲个数的比值,电流脉冲概率反应了一段时间内放电间隙的大小,是后续对进给速度控制的依据。单次采样的电流脉冲如图2所示。

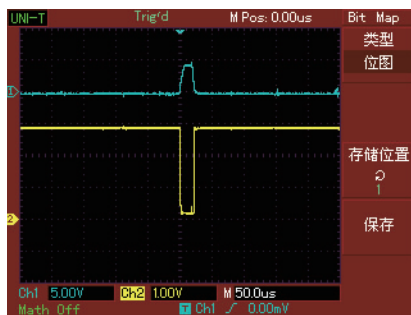


图2 电流脉冲波形

2 Q -Learning 控制算法

强化学习的智能体(agent)一般由环境感知器、学习器和动作选择器3个组成部分。智能体通过环境感知器得到当前所处状态 S ,然后在动作选择器的作用下输出相应的动作 a ,之后环境

在智能体做出 a 动作后会变为状态 S' ,此时智能体会根据新状态对 a 动作进行评估,并更新学习器,循环该过程直至达到期望的状态^[13-15]。

智能体在环境作用下,逐步形成对刺激的评价,产生能最终获得最大收益回报的学习性行为。马尔可夫决策过程 M 定义如下

$$M = (S, A, P, r, \gamma) \quad (1)$$

式中: S 为智能体所处环境离散化的状态空间集; A 为智能体所能执行的离散化动作集; P 为 S 状态转移到 S' 状态的转移函数; r 为各状态下给予智能体的奖励值; γ 为各时间序列的衰减因子并且 $\gamma \in [0, 1]$ 。

在对硅材料进行电火花加工时,要根据当前状态的放电脉冲概率来调整进给速度,保持加工的稳定。本文采用增量式 Q 学习方法对进给速度进行控制。相比传统 Q 学习方法,每次选择的动作空间覆盖整个可选空间范围,增量式 Q 学习每次的动作输出选择为增量动作 Δa 。增量式 Q 学习一次迭代学习的过程如下:

步骤1:获取当前电流脉冲概率大小,确定当前状态 S_p ,初始化绝对动作值 A_{basic} 及整个 Q 表,以及优化回合结束标志。

步骤2:确定增量动作值 Δa 的阈值范围。

步骤3:构建动作策略 T ,策略 T 映射在状态空间中各状态所对应的最大奖励动作,记为 $T_\Delta(S_i)$ 。

步骤4:初始化整个 Q 表,并在 $t+1$ 时刻获取当前状态,对 $Q(S_t, A_t)$ 进行更新和 A_{basic} 。

步骤5:判断是否达到该回合结束标志,若满足则结束该迭代过程,不满足则继续跳转至步骤3。

Q 表结构如表1所示。

表1 Q 表结构

状态 S_p	动作 A_n				
	A_1	...	A_y	...	A_n
S_{p1}	$Q(S_{p1}, A_1)$	...	$Q(S_{p1}, A_y)$	...	$Q(S_{p1}, A_n)$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
S_{px}	$Q(S_{px}, A_1)$	...	$Q(S_{px}, A_y)$	...	$Q(S_{px}, A_n)$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
S_{pm}	$Q(S_{pm}, A_1)$	...	$Q(S_{pm}, A_y)$	...	$Q(S_{pm}, A_n)$

在步骤1中,一开始需对状态空间进行离散,即将整个状态空间以一定步长均匀分成 n 个区间,对当前放电状态用电流脉冲概率衡量且额外规定正常和异常状态的集合 S_n 和 S_u 。

步骤2设置了增量动作的阈值,阈值的大小影响控制系统精度以及平稳性,本文将增量动作集设置成: $\Delta = \{-32, -16, -8, -4, -2, -1, 0, 1, 2, 4, 8, 16, 32\}$,单位 $\mu\text{m/s}$ 。

在步骤3中,运用 ε -greedy 选择动作增量

$$a_{\Delta} = \begin{cases} A_{Q_{\max}}(S_p), & \lambda > \varepsilon \\ A_{\Delta}, & \lambda \leq \varepsilon \end{cases} \quad (2)$$

式中: $A_{Q_{\max}}(S_p)$ 为根据策略 π 选择当前状态下能得到最大奖励的动作; A_{Δ} 为在步骤2中设置的增量动作集; λ 为 $0 \sim 1$ 的随机数。

步骤4中,初始化整个 Q 表,令

$$Q(S_{pm}, A_n) = 0, \quad S_{pm} \in S, A_n \in A \quad (3)$$

在每个迭代代数中,需要对 Q 表进行更新,具体的更新方式如下

$$Q(S, A) \leftarrow Q(S, A) + \alpha [R + \gamma \max_{A'} Q(S', A') - Q(S, A)] \quad (4)$$

式中: α 为学习率; γ 为折扣因子,表示对未来奖励值的衰减,主要用于对计算累计奖励回报的调和^[16-17]; $\max_{A'} Q(S', A')$ 为在 S' 状态下采取动作能获得的最大的 Q 值; R 为状态 S 到状态 S' 的奖励,本文将奖励 R 定义为与目标电流脉冲概率的偏差有关的分段函数,即

$$R = \begin{cases} -k \cdot |I_c - I_e|, & I_c \in S_n \\ -100, & I_c \in S_u \\ 1, & I_c \in S_o \end{cases} \quad (5)$$

式中: k 为相关系数; I_c 为当前电流脉冲概率大小; I_e 为目标电流脉冲大小; S_o 为达到控制目标状态的集合。另外,对绝对动作值 A_{basic} 进行如下的更新

$$A_{\text{basic}} = A_{\text{basic}} + a_{\Delta} \quad (6)$$

3 仿真与实验

3.1 建立仿真模型

本文采用 BP 神经网络对系统进行建模,研究对象是进给速度和电流脉冲概率的关系,神经网络模型如图3所示^[18-19]。

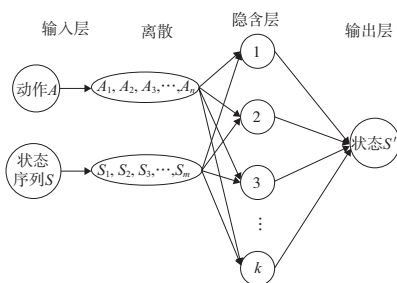


图3 2-N-1 BP神经网络模型

本文在上述搭建的实验平台进行实验,实验过程中相关加工参数设定如表2所示。

表2 电火花线切割实验加工参数

项目	内容	项目	内容
工件材料	P型单晶硅	电极	钨丝
伺服电压/V	115	进电方式	正极性
占空比	1:10	脉宽/ μs	30

对线切割加工过程中的原始数据采集并整理,筛选出1000组进行训练,并对数据集进行分割,其中20%作为测试集,80%作为训练集。训练误差曲线如图4所示。可以看到在约160轮时该网络收敛。

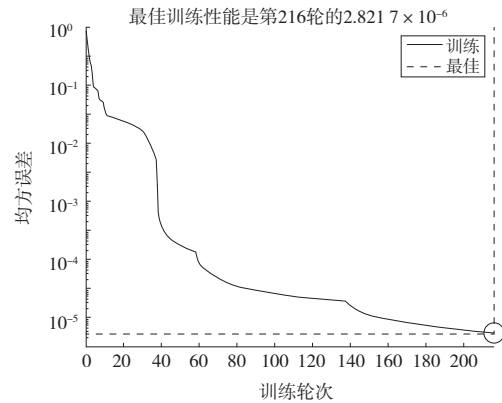


图4 训练误差曲线

测试集部分数据的验证结果如图5所示。可以看到模型所预测的输出和期望输出误差较小,因此该模型可以作为强化学习的训练环境进行迭代学习。

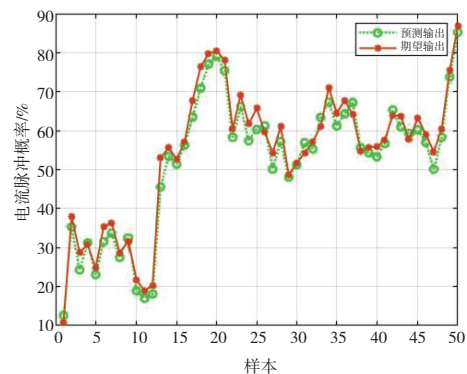


图5 BP2-5-1神经网络输入输出对比

3.2 仿真结果与分析

将上文所建立的神经网络模型作为强化学习训练环境,算法中相关学习参数设置如表3所示。

表3 学习参数设置

参数	数值	参数	数值
学习率	0.50	最大失败步数	50
折扣因子	0.99	最大探索率	0.50
最大步数	6 000	最小探索率	0.01

在训练过程中,不同训练次数下的训练结果如图6所示,从图6中可以看出,随着训练次数的增多,智能体不断修正和更新 Q 表,最终使整个系统朝着期望的状态逼近。

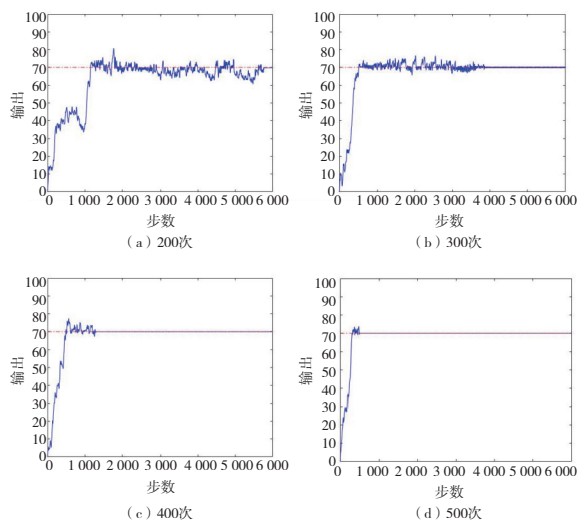


图6 不同训练次数系统响应输出

训练过程中智能体获得的累积奖励如图7所示,累计奖励值在约第500次训练达到最大值并收敛。

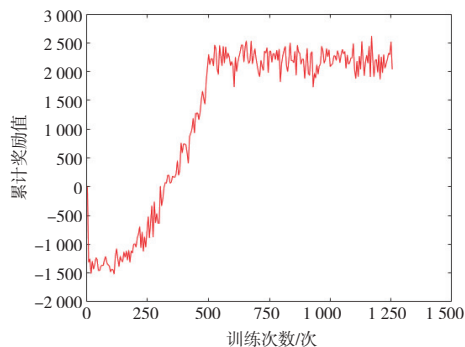


图7 训练过程的累计奖励值变化

因此对硅材料进行电火花线切割加工来说该控制方法是可行的,并且该控制方法以其灵活性和适应性强的特点而非常适合在随机变化的电火花加工环境中使用。

3.3 实验结果与分析

在上述搭建的实验平台基础上,应用基于增量 Q 学习的控制算法对线切割机床步进电机进

行控制。基于上文中设定的工艺参数进行实验,加工现场如图8所示。

加工过程中电流脉冲概率的变化曲线如图9所示。从图9中可以看出虽然加工过程中切割厚度在不断变化,但电流脉冲概率仍能维持在设定目标范围内,这表明该控制算法能适应厚度加工。



图8 硅材料电火花线切割现场

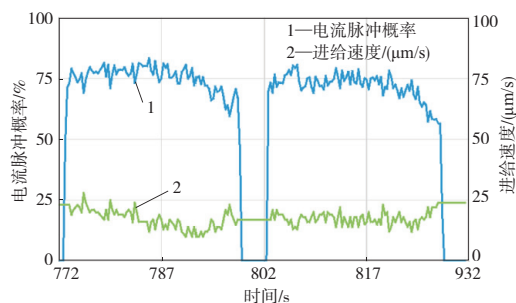


图9 电流脉冲概率变化曲线

基于增量 Q 学习控制方法下的电极丝进给加工距离变化曲线,如图10所示。

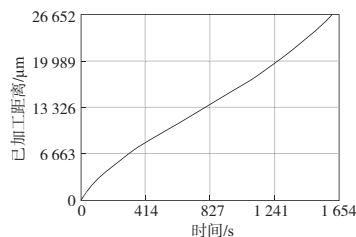


图10 加工距离变化曲线

将训练好的控制器用于硅棒电火花线切割加工,记录各切割次数下的硅片表面粗糙度,同时统计加工时间并计算其加工效率。不同切割次数后硅片表面粗糙度和切割效率的变化曲线,如图11所示。

可以看出随着切割次数的增加,切割效率变化不大,但加工出来的工件表面粗糙度越来越小,这说明智能体在不断调整进给速度使表面加工质量逐渐提高。

目前传统控制算法多基于经验公式对进给速度进行控制^[20],保持相关工艺参数不变,使用经验公式对硅棒进行电火花线切割,对加工后的硅

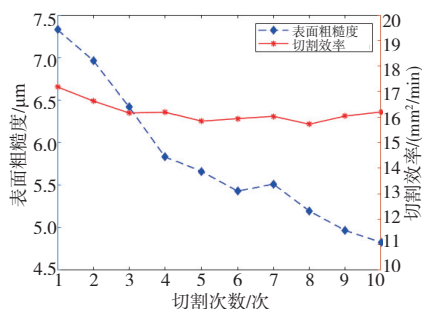


图 11 不同切割次数下数据对比

片的表面粗糙度 Ra 和切割效率 v 进行比较。基于增量式 Q 学习控制方法下加工的硅片表面质量更加均匀。

不同控制方法测得的表面粗糙度与加工效率如表 4 所示。采用增量式 Q 学习控制方法获得的硅片表面粗糙度要优于依赖经验公式方法获得的,同时切割效率也略高于经验公式方法控制的,并且在该控制方法下随着切割次数的增加,其加工精度会随之进一步得到提高。

表 4 不同控制方法对比

项目	经验公式控制	增量式 Q 学习控制
单次测量	5.768	4.812
表面粗糙度 $Ra_n/\mu\text{m}$	5.532	4.932
	5.764	4.743
平均表面粗糙度 $Ra/\mu\text{m}$	5.688	4.829
加工效率 $v/(\text{mm}^2/\text{min})$	19.41	18.35

4 结语

本文将基于增量 Q 学习的自适应控制方法应用于硅材料电火花线切割的伺服控制。首先建立了用于强化学习的 BP 神经网络数学模型,通过实验对该模型进行反复训练,使其能较好地模拟真实环境的状态变化。然后将其用于硅材料电火花线切割进给速度的自适应控制。仿真和实验表明:该控制方法能通过不断试错学习而有效防止切割过程中发生断丝的危险,使硅材料电火花线切割维持合适的放电间隙。本文提出的控制算法在加工精度和加工效率上都优于使用经验公式的控制方法。

参考文献:

- [1] 王旭. 半导体材料的应用现状及发展趋势[J]. 光源与照明,2022(1):67-69.
- [2] 侯志坚,索海生. 硅晶体切片设备的研究现状与进

展[J]. 机床与液压,2011,39(2):37-39.

- [3] XAKALASHE B S, TANGSTAD M. Silicon processing: from quartz to crystalline silicon solar cells[J]. Southern African Pyrometallurgy International Conference, 2011: 1-18.
- [4] 邵程杰. 基于闭环脉冲电源的硅晶体放电加工试验研究[D]. 南京:南京航空航天大学,2020.
- [5] 罗福源,张琪. 基于往复渐进运动控制的硅材料电火花线切割[J]. 南京航空航天大学学报,2013,45(6): 830-836.
- [6] 陈良树. 基于 DOE 方法的线切割加工效率优化及应用[J]. 中国新技术新产品,2024(21):64-66.
- [7] 施宏宇,陈南光,杨润林,等. PtIr 高温合金电火花线切割工艺参数优化研究[J]. 贵金属,2023,44(增刊1):23-27.
- [8] 潘慧君,刘志东,黄赛娟,等. 半导体电火花线切割加工电流脉冲概率特性及伺服控制研究[J]. 电加工与模具,2013(4):28-31,35.
- [9] 王立群,唐敦兵,朱海华,等. 基于强化学习的柔性作业车间动态调度算法研究[J]. 机械制造与自动化,2023,52(3):1-3,10.
- [10] 王锋,李佳旺,张粤东,等. 基于时域-频域双路网络的深度学习电火花线切割放电状态识别[J]. 电加工与模具,2022(5):23-29,40.
- [11] 陈晟,奚学程,李子伦,等. 电火花线切割小圆弧加工自适应控制[J]. 电加工与模具,2022(1):24-28,50.
- [12] 高阳,马虎亮,王燕青,等. 电火花线切割伺服控制系统特性实验研究[J]. 机床与液压,2021,49(13): 51-55,126.
- [13] 丁世飞,杜威,张健,等. 多智能体深度强化学习研究进展[J]. 计算机学报,2024,47(7):1547-1567.
- [14] 曹宏业,刘潇,董绍康,等. 面向强化学习的可解释性研究综述[J]. 计算机学报,2024,47(8):1853-1882.
- [15] 唐蕾,牛园园,王瑞杰,等. 强化学习的可解释方法分类研究[J]. 计算机应用研究,2024,41(6): 1601-1609.
- [16] 褚晶,邓旭辉,岳硕. 基于 Q -learning 的搜救机器人自主路径规划[J]. 南京航空航天大学学报,2024, 56(2):364-374.
- [17] 宋宇航,陈宇帆,魏延岭,等. 基于强化学习环境设计策略的电动汽车充电路径规划[J]. 电力系统自动化,2024,48(11):184-196.
- [18] 高杨,曹仰杰,段鹏松. 神经网络模型轻量化方法综述[J]. 计算机科学,2024,51(增刊1):23-33.
- [19] 严如强,商佐港,王志颖,等. 可解释人工智能在工业智能诊断中的挑战和机遇:先验赋能[J]. 机械工程学报,2024,60(12):1-20.
- [20] 刘璐,王娟. 基于电火花线切割加工工艺指标对加工质量的影响分析[J]. 内燃机与配件,2021(24): 94-96.

收稿日期:2023-11-30