

玉米深加工副产物混菌固态发酵饲料工艺优化

满云^{1,2*}, 蒋光乐¹, 蒋舒晴¹, 刘庆宇¹

(1. 蚌埠学院 食品与生物工程学院, 安徽 蚌埠 233030; 2. 安徽省生物乙醇重点实验室, 安徽 蚌埠 233030)

摘要:为提升玉米深加工副产物的饲用价值,以乳酸杆菌、枯草芽孢杆菌、酵母菌复合菌种进行混菌固态发酵,采用单因素试验及 $L_9(3^4)$ 正交试验,探究发酵温度、水分含量及发酵时间对发酵饲料的感官评价、pH值、益生菌数量及可溶性蛋白含量的影响。结果显示,各因素影响主次顺序为:发酵时间>发酵温度>水分含量。最优工艺参数为:发酵温度30℃,水分40%,发酵时间4d。在此条件下,发酵饲料呈浓郁酒香与酱香味,质地松散湿润,无结块,综合评分84分。

关键词:玉米深加工副产物;固态发酵;混菌发酵;正交试验;蛋白饲料;可溶性蛋白

中图分类号:S816.6

文献标识码:A

文章编号:2095-297X(2026)05-0001-07

Optimization of Mixed-Strain Solid-State Fermentation Process for Feed Using Corn Deep-Processing By-Products

MAN Yun^{1,2*}, JIANG Guangle¹, JIANG Shuqing¹, LIU Qingyu¹

(1. School of Food and Biological Engineering, Bengbu University, Bengbu, 233030, Anhui;

2. Anhui Provincial Key Laboratory of Bioethanol, Bengbu, 233030, Anhui)

Abstract: To improve the feeding value of by-products from corn deep processing, it adopted mixed strain solid-state fermentation with *Lactobacillus*, *Bacillus subtilis* and *Saccharomyces* as composite strains. Single-factor tests and $L_9(3^4)$ orthogonal test were used to investigate the effects of fermentation temperature, moisture content and fermentation time on sensory quality, pH value, probiotic count and soluble protein content of fermented feed. The results showed that the order of influencing factors was fermentation time > fermentation temperature > moisture content. The optimal parameters were determined as fermentation temperature 30℃, moisture content 40%, and fermentation time 4 d. Under these conditions, the fermented feed had rich wine and soy sauce aroma, loose and moist texture without caking, with a comprehensive score of 84 points.

Key words: corn deep processing by-products; solid-state fermentation; mixed strain fermentation; orthogonal test; protein feed; soluble protein

蛋白饲料是支撑畜牧业与水产养殖业高效发展的核心原料,传统植物蛋白原料普遍存在大分子蛋白消化率低、抗营养因子残留、功能性成分不足等问题,限制了饲用效率提升。混菌固态发酵依托多菌种协同代谢,可定向降解抗营养因子、将大分子蛋白转化为小分子肽与游离氨基酸,同时富集益生菌、有

机酸及风味物质,显著改善饲料营养价值与适口性,已成为非常规饲料资源化利用的关键技术。

当前,玉米淀粉、酒精深加工产业规模化生产过程中副产物产出量大,玉米酒精糟(简称“DDGS”)、玉米蛋白粉等副产物粗蛋白含量偏高、原料易得、生产成本低廉,但普遍存在适口性差、抗营养因子残

收稿日期:2026-04-13 * 通讯联系人

基金项目:安徽省重点实验室建设项目(S202305a12020040);安徽省大学生创新创业项目(S202511305125)。

作者简介:满云(1976—),女,安徽寿县人,副教授,博士。E-mail:manyun760124@163.com

留、动物消化利用率偏低等缺陷,资源化利用受限^[1-3]。国内外围绕农副产品固态发酵制备发酵饲料已开展大量研究;王立等^[1]证实发酵改性大豆蛋白可显著提升肉牛采食率与日增重;Chen等^[2]利用乳酸菌单一菌种固态发酵果蔬副产物,但单一菌株代谢途径单一,难以同步实现蛋白转化与多风味物质富集;叶丙奎^[3]研究证实混菌发酵可有效提升饲料消化吸收率,但试验原料以豆粕、秸秆为主,极少聚焦玉米深加工系列副产物复配底物;冯娜等^[4]、洪婷婷等^[5]分别证实单一菌种发酵饲料可改善羊、仔猪生长性能与机体免疫,但工艺仅针对单一发酵参数优化,缺乏温度、水分、发酵时间多因素耦合探究;侯永帅等^[6]、姚志芳等^[7]分别探究了枯草芽孢杆菌、酵母菌单独发酵特性,结果表明单一菌种发酵产物营养提升幅度有限;国外 Abbas等^[8]采用枯草芽孢杆菌复合菌株发酵桑叶制备抗氧化肽,验证了混菌在蛋白分解方面的优势,但底物类型与玉米加工副产物理化特性差异较大。现有相关研究多选用单一菌种或仅优化单个工艺参数^[4-5],针对乳酸菌、枯草芽孢杆菌、酵母菌三菌复配协同发酵玉米深加工副产物的系统性工艺优化研究偏少,关键工艺参数交互作用规律尚不明确,难以支撑工业化稳定生产^[9]。本研究以乳酸杆菌、枯草芽孢杆菌、酵母菌为复合菌剂,以 DDGS、玉米蛋白粉等玉米深加工副产物为主要底物,通过单因素试验明确关键参数取值范围,结合正交试验优化发酵工艺,旨在确定稳定高效的混菌固态发酵方案,为非常规蛋白资源高值化利用提供科学依据^[10-15]。

1 材料与方法

1.1 材料与菌种

原料:DDGS、豆粕、玉米粉、麦麸、玉米蛋白粉、鱼粉、红糖,中粮生物科技提供;

菌种(饲料级):乳酸菌(1×10^{11} CFU · g⁻¹)、枯草芽孢杆菌(1×10^{11} CFU · g⁻¹)、酵母菌(2×10^{10} CFU · g⁻¹),仙农生物科技(上海)公司提供;

试剂与培养基:平板计数琼脂(PCA)、0.1 mg · mL⁻¹牛血清蛋白(BSA)、0.1 mg · mL⁻¹考马斯亮蓝染液,实验室自制。

1.2 主要仪器

FA214A 电子天平,上海海康电子厂;KW-1000DA 恒温水浴锅,金坛市文华科教实验仪器厂;PWT-P42B 恒温培养箱,华德利(合肥)科学器材;YX-24HDD 手提式高压灭菌锅,江阴滨江医疗设备有限公司;SW-CJ-1G 净化工作台,苏州安泰空气技

术有限公司,pH818 pH 计,万创电子制品有限公司;TU-1901 分光光度计,北京普析通用仪器有限责任公司。

1.3 试验流程

试验流程图如图 1 所示。

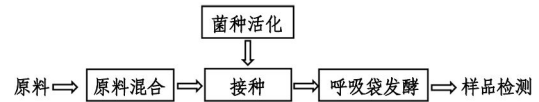


图 1 试验流程图

(1)原料配比:玉米蛋白粉 30.00 g、豆粕 10.00 g、玉米粉 10.00 g、DDGS 6.67 g、鱼粉 6.67 g、麦麸 3.33 g、红糖 2 g、水 31.33 g,充分混匀备用。

(2)菌种活化:称取 1 g 乳酸杆菌、5 g 红糖、100 mL 纯净水充分混合溶解,35 ℃ 活化 2 h。酵母菌和枯草芽孢杆菌活化同上。

(3)微生物接种:2.5% 枯草芽孢杆菌 + 1.5% 酵母菌 + 0.75% 乳酸菌。

(4)发酵:12 cm × 17 cm 呼吸袋密闭发酵。

1.4 试验方法

1.4.1 单因素试验设计

发酵基础条件:发酵温度 33 ℃、发酵时间 3 d、发酵水分 40%。

固定其他参数,考察:发酵温度:30、33、36、39 ℃;发酵水分:30%、35%、40%、45%;发酵时间:2、3、4、5、6 d。以感官评定和理化指标进行综合评分。

1.4.2 正交试验设计

根据单因素结果,采用 L₉(3⁴) 正交表,以发酵温度、水分含量、发酵时间为考察因素,以综合评分为评价指标,优化最佳工艺组合。正交试验 L₉(3⁴) 设计表如表 1 所示。

表 1 正交试验 L₉(3⁴) 设计表

组别	A(发酵温度) /℃	B(水分) /%	C(发酵时间) /d
试验 1	1	1	1
试验 2	1	2	2
试验 3	1	3	3
试验 4	2	1	2
试验 5	2	2	3
试验 6	2	3	1
试验 7	3	1	3
试验 8	3	2	1
试验 9	3	3	2

1.5 检测方法

指标测定包括产气情况、气味、组织形态、产酸、益生菌总量、可溶性蛋白含量等。

选择 10 名具有感官评价经验的专业人员根据感官评定表(见表 2)^[13]对饲料进行评价打分,取平均值作为感官评价分值。

1.5.1 感官评价标准

表 2 感官评定表

等级	产气量(35 分)	气味(35 分)	组织形态(30 分)
一级产品 (90 - 100 分)	大量产气 (严重胀袋)	浓郁的芳香酸味,伴有酒香味,酱香味。	质地松散湿润,有颗粒感,色泽均匀,无结块现象。
二级产品 (80 - 89 分)	中等产气 (中度胀袋)	有芳香酸味,但并不浓郁,酒香味和酱香味较浅。	有轻微的结块现象,易松散,色泽较均匀,饲料整体质感没有一级产品蓬松柔软。
三级产品 (70 - 79 分)	少量产气 (轻微胀袋)	酸味较重,带有不太愉悦的气味(可能刺鼻),没有明显腐败和霉味。	手感粗糙、坚硬,有明显硬块、结团或有部分发稠现象。
四级产品 (< 70 分)	无或微量产气 (无胀袋)	腐败,发霉或难闻的气味如刺鼻的酸臭味。	严重结块、硬化明显霉变、变色,或浓稠腐败。

1.5.2 pH 值的测定

称取 5 g 发酵饲料样品放入锥形瓶中,加入 50 mL 去离子水(1:10 左右),摇匀,室温 5 min 后,过滤,取上清液待测。

1.5.3 益生菌数量的测定

样品制备:称取 5 g 发酵饲料,加入 50 mL 无菌生理盐水(0.9% NaCl)于锥形瓶中,此时料液比为 1:10。得到稀释倍数为 10^{-1} 初始液,取 1 mL 菌悬液加 9 mL 无菌生理盐水于试管中,依次得到 10^{-2} 、

10^{-3} 、 10^{-4} 、 10^{-5} 、 10^{-6} 等系列稀释倍数菌悬液,取 0.1 mL 进行平板涂布,培养箱 37 ℃ 培养 24 h,直接计数法确定。

1.5.4 可溶性蛋白含量测定

标准曲线制备:采用考马斯亮蓝法,根据表 3 准备试管,室温放置 2 min,595 nm 波长测吸光度并绘制标准曲线。标准曲线试验设计表如表 3 所示。牛血清蛋白标准曲线图如图 2 所示。

表 3 标准曲线试验设计表

项目	0	1	2	3	4	5
牛血清蛋白(0.1 mg · mL ⁻¹)	0	100	200	300	400	500
ddH ₂ O	1000	900	800	700	600	500
考马斯亮蓝 G-250 溶液	5000	5000	5000	5000	5000	5000
各试管总溶液量	总量为 6000					uL

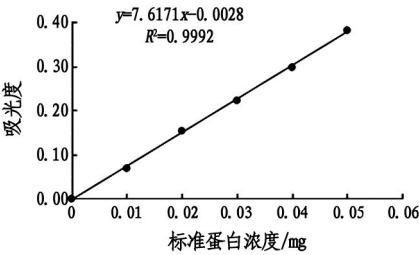


图 2 牛血清蛋白标准曲线图

样品测定:取发酵饲料 0.5 g 加入 50 mL ddH₂O 于锥形瓶中摇匀,离心后按标准曲线方法测定吸光值。代入标准曲线方程,计算总可溶性蛋白质含量。

1.5.5 综合评价表设计

根据感官评定分数分析,对其中占比为 40% 和理化指标分数占比为 60% 进行加权,计算综合评价分数。其中感官评定分数按表 2 确定,理化指标分数按表 4 确定。

表4 理化指标设计表

理化指标	等级	分数/分
pH 值(35 分)	3.80 – 4.00	30 – 35
	4.01 – 4.40	10 – 29
	>4.40	0 – 9
益生菌数量(35 分)	2.00 – 2.60($\times 10^{10}$)	30 – 35
	1.00 – 1.99($\times 10^{10}$)	10 – 29
	<1.00($\times 10^{10}$)	0 – 9
可溶性蛋白含量(30 分)	4.00% – 5.00%	25 – 30
	2.00% – 3.99%	10 – 24
	<2.00%	0 – 9

2 结果与分析

2.1 单因素试验结果分析

2.1.1 发酵温度对混菌固态发酵蛋白饲料的影响

不同发酵温度条件发酵饲料的理化指标数据及综合评价得分分别如表5、表6所示。

表5 不同发酵温度条件发酵饲料的理化指标数据表

理化指标	30	33	36	39
pH 值	4.38	4.20	4.24	4.35
益生菌数(10^{10})	1.22	1.54	1.15	1.05
可溶性蛋白质含量/%	2.98	3.00	3.57	2.95

表6 不同发酵温度下的发酵饲料综合评价得分表

项目	30	33	36	39
感官评价(占比40%)	产气量	29	33	31
	气味	28	30	30
	组织形态	30	30	30
	感官评价总分	87	93	91
理化指标(占比60%)	pH 值	20	28	26
	益生菌数量	21	24	19
	可溶性蛋白质含量	15	16	22
	理化指标总分	56	68	67
综合评定得分(感官 $\times 40\%$ + 理化 $\times 60\%$)		68.4	78.0	76.6

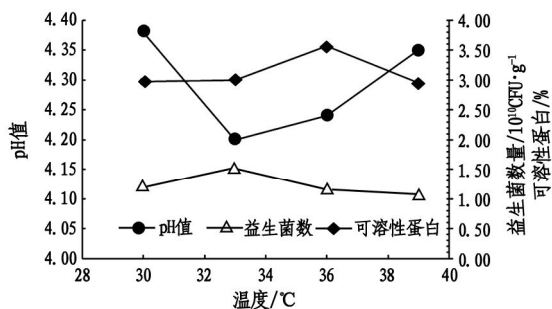


图3 理化指标随发酵温度变化图

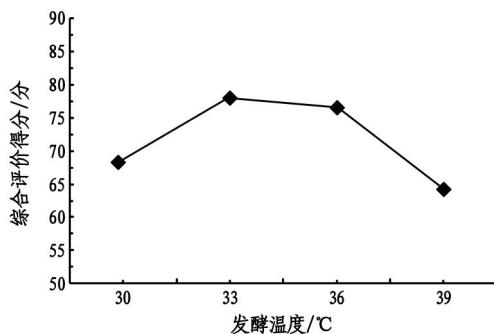


图4 综合评价得分随发酵温度变化图

由图3、图4可知,随着发酵温度增加,感官评分先上升后下降,pH先下降后上升,益生菌数量和可溶性蛋白质先上升后下降,加权综合评分先上升后下降。原因是:当发酵温度较低时,微生物活性减弱,代谢较慢,利用营养物质能力较弱,生成乳酸等风味物质较少;温度增至适宜发酵温度,微生物活性增加,代谢增强,可溶性蛋白质增多,同时代谢生成乳酸等风味物质较多;当发酵温度超过适宜发酵温度时,微生物活性受到抑制,代谢减弱,生长增殖能力减弱,利用营养物质转化为可溶性蛋白质和生成乳酸等风味物质能力减弱。

2.1.2 水分对混菌固态发酵蛋白饲料的影响

不同水分条件发酵饲料理化指标数据及综合评价得分分别如表7、表8所示。

表7 不同水分条件发酵饲料理化指标数据表

理化指标	30	35	40	45
pH 值	4.37	4.06	3.94	4.19
益生菌数(10^{10})	1.54	1.73	2.14	1.96
可溶性蛋白质含量/%	2.78	3.17	4.71	4.43

表 8 不同水分条件的发酵饲料综合评价得分表

项目		30	35	40	45
感官评价(占比 40%)	产气量	30	30	31	29
	气味	28	30	32	31
	组织形态	25	27	29	26
	感官评价总分	83	87	92	86
理化指标(占比 60%)	pH 值	15	25	31	20
	益生菌数量	20	22	26	24
	可溶性蛋白质含量	18	22	27	25
综合评定得分(加权)		65	76.2	87.2	75.8

由图 5、图 6 可知,随着发酵水分增加,感官评总分先上升后下降,pH 值先下降后上升,益生菌数量和可溶性蛋白先上升后下降,加权综合评分先上升后下降。原因是:水分较低时,微生物生长、增殖、代谢较慢,利用营养物质转化为可溶性蛋白质和生成乳酸等风味物质较少;随着水分增加,微生物快速生长、增殖,代谢旺盛,可溶性蛋白质和乳酸等风味物质含量增加;当水分超过 40% 时易受杂菌污染,菌种生长、增殖、代谢受到抑制。

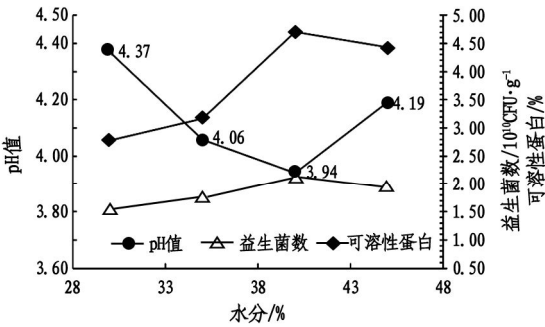


图 5 理化指标随水分变化图

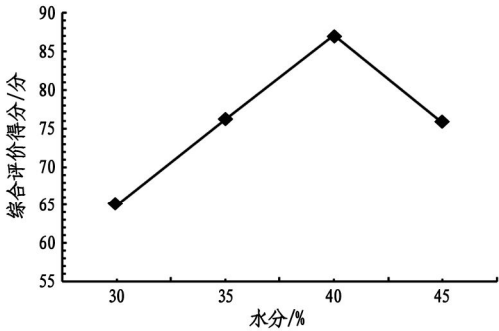


图 6 综合评分随水分变化图

2.1.3 发酵时间对混菌固态发酵蛋白饲料的影响

不同发酵时间下发酵饲料理化指标数据及综合评价得分分别如表 9、表 10 所示。

表 9 不同发酵时间饲料理化指标数据表

理化指标	2	3	4	5	6
pH 值	4.38	4.28	4.21	4.15	4.10
益生菌数量(10^{10})	0.62	1.32	1.58	1.66	1.54
可溶性蛋白质含量/%	2.10	2.56	3.25	3.14	2.85

表 10 不同发酵时间饲料的综合评价得分表

项目		2	3	4	5	6
感官评价(占比 40%)	产气量	28	30	31	30	30
	气味	28	30	30	27	23
	组织形态	30	29	30	26	22
	感官评价总分	86	89	91	83	75
理化指标(占比 60%)	pH 值	14	18	21	24	25
	益生菌数量(10^{10})	8	15	18	20	16
	可溶性蛋白质含量/%	15	18	23	21	19
综合评定得分(加权)		56.6	66.2	73.6	72.2	66

由图 7、图 8 可知,随着发酵时间增加,感官评定分数、益生菌数量和可溶性蛋白均呈先上升后下降,而 pH 值持续降低。可能原因是:随着发酵时间持续增加,微生物生长增殖不断利用营养物质转化为可溶性蛋白质和生成乳酸等风味物质,而发酵时间过长,有机酸、醇类等过度积累、蛋白质过度降解,饲料产生刺鼻、酸败气味,影响饲料的品质。

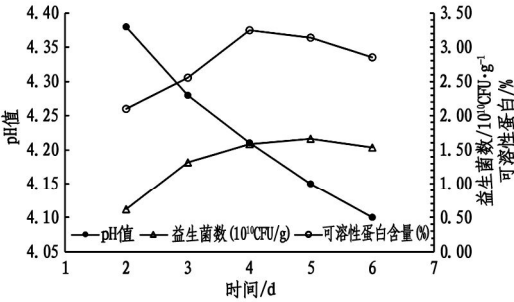


图 7 理化指标随发酵时间变化图

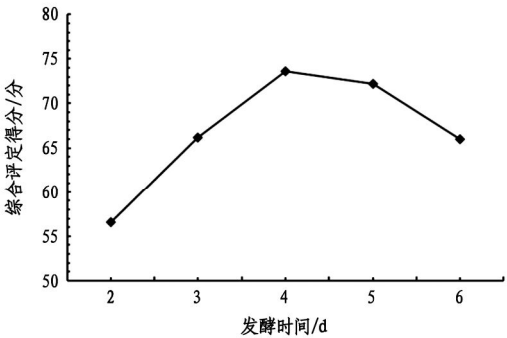


图 8 综合评定得分随发酵时间变化图

2.2 正交试验结果分析

正交试验理化指标、综合评价以及结果分析分别如表 11、表 12、表 13 所示。

由表 13 可知,三个影响因素主次顺序是 C > A > B。正交试验结果得出,混菌固态发酵蛋白饲料的最优工艺条件为 A₁B₂C₁。即发酵温度为 30 ℃,发酵饲料水分 40%,发酵时间为 4 d。

表 11 正交试验理化指标呈现表

理化指标	1	2	3	4	5	6	7	8	9
pH 值	4.02	4.12	4.21	4.29	4.28	4.23	4.36	4.08	4.10
益生菌数量(10 ¹⁰)	2.54	2.04	1.43	1.86	2.11	2.18	1.39	1.69	1.84
可溶性蛋白质含量/%	3.41	2.82	2.50	2.45	2.54	3.04	2.12	2.86	2.64

表 12 正交试验综合评价得分表

项目		1	2	3	4	5	6	7	8	9
感官评价(占比 40%)	产气量	32	29	28	34	28	26	25	32	32
	气味	34	35	34	34	33	34	33	34	33
	组织形态	27	29	27	28	27	28	27	26	28
	感官评价总分	93	89	89	95	88	88	85	92	93
理化指标(占比 60%)	pH 值	29	26	23	20	21	22	17	28	27
	益生菌数量(10 ¹⁰)	32	31	24	26	33	35	23	25	26
	可溶性蛋白质含量(%)	18	15	13	12	14	17	14	16	14
综合评定得分(加权)		84.6	78.8	71.6	72.8	76	79.6	66.4	78.2	77.4

表 13 正交试验结果分析表

分组	A(温度/℃)	B(水分/%)	C(发酵时间/d)	综合评价得分/分
1	1(30)	1(35)	1(4)	84.6
2	1	2(40)	2(5)	78.8
3	1	3(45)	3(6)	71.6
4	2(33)	1	2	72.8

续表
表 13 正交试验结果分析表

分组	A(温度/℃)	B(水分/%)	C(发酵时间/d)	综合评价得分/分
5	2	2	3	76
6	2	3	1	79.6
7	3(36)	1	3	66.4
8	3	2	1	78.2
9	3	3	2	77.4
均值 1	78.333	74.600	80.800	
均值 2	76.133	77.667	76.333	
均值 3	74.000	76.200	71.333	
极差	4.333	3.067	9.467	

2.3 验证试验

最优工艺下三次重复试验平均综合评分为 84 分,与正交试验最高评分接近,工艺稳定可靠。

3 结论

以玉米深加工副产物为主要原料,通过单因素和正交试验确定混菌固态发酵最优工艺:发酵温度 30℃,发酵水分 40%,发酵时间 4 d。发酵饲料以感官评价、pH、益生菌数量、可溶性蛋白含量进行综合评价,优化后发酵饲料具有浓郁酒香与酱香味,质地松散湿润,无结块,综合评分达到 84 分。

利用乳酸菌、枯草芽孢杆菌、酵母菌复配发酵玉米深加工副产物,有效改善原料适口性与蛋白利用率,降低饲料生产成本,契合当下非常规饲料资源化利用发展方向^[9,15]。然而试验仅优化温度、水分、时间三项关键参数,后续可探究菌种配比、接种量、原料复配比例等条件,为酒精副产物资源化利用与高品质发酵蛋白饲料工业化生产提供更全面的理论依据与技术参考^[14]。

参考文献:

- [1] 王立,张翠翠.大豆蛋白饲料在肉牛饲养中的应用效果[J].畜牧业环境,2024(9):16-18.
- [2] Wang Chen, Wang Yuan, Shan Weiting, et al. Solid-state fermentation with lactic acid bacteria of citric acid degrading for hawthorn fruit improvement: sensory, flavor, antioxidant properties[J]. Food Bioscience, 2024, 62: 105402.
- [3] 叶丙奎.混菌固态发酵对饲料营养价值及生长育肥猪生产性能的影响[D].咸阳:西北农林科技大学,2009.
- [4] 冯娜,刘小东,高瀚洋,等.发酵饲料在羊养殖生产中的

应用[J].畜牧兽医杂志,2025,44(1):84-89.

- [5] 洪婷婷,李小丽,翁卿甫,等.乳酸杆菌发酵饲料对仔猪生长性能、免疫力及血清生化指标的影响[J].今日养猪业,2024(3):15-18.
- [6] 侯永帅,高亮,蒋松,等.枯草芽孢杆菌的生理功能及其在反刍动物生产中的应用[J].上海畜牧兽医通讯,2025(1):47-51.
- [7] 姚志芳,冯宇哲,张晓卫,等.酵母菌在湿鲜生物发酵饲料中的应用研究进展[J].青海畜牧兽医杂志,2021,51(3):46-51,61.
- [8] Abbas Zaheer, Zhang Jing, Tong Yucui, et al. Production and characterization of novel antioxidant peptides from mulberry leaf ferment using *B. subtilis* H4 and *B. amyloliquefaciens* LFB112[J]. Food Chemistry, 2025, 482: 144022.
- [9] 张朋,姚有莉,罗金婷,等.发酵非常规饲料在生猪养殖中的应用研究进展[J].青海畜牧兽医杂志,2025,55(2):56-60,72.
- [10] 刘志国.秸秆前处理固态发酵制备生物肥料及促生效果的研究[D].呼和浩特:内蒙古大学,2024.
- [11] 姚竹青,代二朋,纪春暖.蛋鸡饲料中添加混合型饲料添加剂对其后期生产性能和蛋品质的影响[J].现代畜牧科技,2025(5):80-82.
- [12] 李凤玲,郑志然.生物酶发酵饲料对海兰灰蛋鸡生产性能和鸡蛋品质的影响[J].新农民,2024(29):112-114.
- [13] 胡敏洁,展丽梅,李吉琴,等.昆明犬幼犬全价饲料固态发酵工艺研究[J].饲料工业,2025,46(1):138-147.
- [14] 魏文江,周彦兵.甘肃省生物发酵饲料产业发展现状与趋势[J].甘肃畜牧兽医,2024,54(3):45-49.
- [15] 林秀蔚,刘文,李青莹,等.非常规饲料在反刍动物生产中的研究进展[J].现代畜牧科技,2025(4):114-117.

责任编辑:汪张贵

复方中药酵素液的制备工艺优化及体外降糖降脂作用

徐世晶*, 熊 柯, 李正阳, 王书胜, 刘丙辉

(蚌埠学院 食品与生物工程学院, 安徽 蚌埠 233030)

摘 要:通过响应面优化试验,研究了七种药食同源原料(山药、茯苓、桑叶、山楂、白果燕麦、杭白菊)经复合乳酸菌(保加利亚乳杆菌、嗜酸乳杆菌、嗜热链球菌、乳双歧杆菌、植物乳杆菌)发酵,以 α -淀粉酶抑制率为主要评价标准,对菌种添加量、原料配比、发酵时长及发酵温度等关键工艺参数进行优化,并测定发酵液对 α -淀粉酶、 α -葡萄糖苷酶及猪胰脂肪酶的抑制率以及对甘氨酸脱氧胆酸钠、牛磺脱氧胆酸钠的结合率。结果表明,当山药:茯苓:桑叶:山楂:白果:燕麦:杭白菊:水=10:9:8:7:2:3:1:250(质量比)时,接种1.5%复合乳酸菌,40℃条件下发酵2 d,酵素液含量10%,其对 α -淀粉酶的抑制率达到94.2%,与发酵前相比提高3.64倍,差异极显著($P<0.01$)。酵素液含量50%时,对 α -葡萄糖苷酶及胰脂肪酶的抑制率最大,分别为66.47%和69.57%,与发酵前相比分别提高2.00倍和1.16倍,差异极显著($P<0.01$)。发酵液与甘氨酸脱氧胆酸钠、牛磺脱氧胆酸钠的结合率分别达到65.24%和67.82%,比发酵初分别提高2.20倍和1.39倍,且均显著高于对照组考来烯胺(60.42%,63.85%),差异显著($P<0.05$)。可见,乳酸复合菌发酵液有利于抑制肠道糖脂代谢相关的上述几种酶的活力,对胆酸盐也有较好的吸附能力,这为研发控制食源性血糖血脂升高的功能性食品提供参考。

关键词:酵素液;工艺优化; α -淀粉酶;降糖降脂

中图分类号:TS201.4

文献标识码:A

文章编号:2095-297X(2026)05-0008-07

Optimization of Preparation Process for Compound Traditional Chinese Medicine Yeast Extract and Study on Its In Vitro Hypoglycemic and Lipid-Lowering Effects

XU Shijing*, XIONG Ke, LI Zhengyang, WANG Shusheng, LIU Binghui

(School of Food and Biological Engineering, Bengbu University, Bengbu, 233030, Anhui)

Abstract:Through response surface optimization experiments, the fermentation of seven medicinal and edible raw materials (yam, Poria cocos, mulberry leaf hawthorn, ginkgo, oat, and Hangbai chrysanthemum) by a composite lactic acid bacteria (*Lactobacillus bulgaricus*, *Lactobacillus acidophilus*, *Streptococcus thermophilus*, *Bifobacterium lactis*, and *Lactobacillus plantarum*) was studied. Using the inhibition rate of α -amylase as the main evaluation criterion, key process parameters such as the amount of bacterial inoculation raw material ratio, fermentation duration, and fermentation temperature were optimized. The inhibition rates of the fermentation liquid on α -amylase, α -glucosidase, and porcine pancreatic lipase, well as its binding rates for sodium glycocholate and sodium taurocholate, were also determined. The results showed that with a mass ratio of yam: Poria coc: mulberry leaf: hawthorn: ginkgo: oat: Hangbai chrysanthemum: water = 10: 9: 8: 7: 2: 3: 1: 250, inoculation of 1.5% composite lactic acid bacteria, fermentation at 40℃ for 2 days, and an enzyme liquid content fraction of 1%, the inhibition rate of α -amylase reached 94.2%, which was 3.64 times higher than before fermentation, with a highly significant difference ($P<0.01$). When the enzyme liquid content fraction was 50%, the inhibition rates of α -glucosidase and pancreatic li-

收稿日期:2026-03-13 * 通讯联系人

基金项目:蚌埠学院国家级大学生创新创业项目(202411305053)。

作者简介:徐世晶(1977—),女,湖北罗田人,讲师,硕士。E-mail:122699730@qq.com

pase were the highest, at 6647% and 69.57%, respectively, which were 2.00 and 1.16 times higher than before fermentation, with highly significant differences ($P < 0.01$). The binding rates of the fermentation liquid with sodium glycocholate and sodium taurocholate reached 65.24% and 7.82%, respectively, which were 2.20 and 1.39 times higher than at the beginning of fermentation, and were significantly higher than those of control group cholestyramine (60.42%, 63.85%), with significant differences ($P < 0.05$). The study indicated that the lactic acid bacteria fermentation liquid is beneficial for inhibiting the activities of the aforementioned enzymes related to intestinal carbohydrate and lipid metabolism, and also has a good adsorption capacity for bile salts. This provides theoretical basis for the development of functional foods to control diet-induced elevations in blood glucose and blood lipids.

Key words: enzyme solution; process optimization; α -amylase; hypoglycemic and lipid-lowering

植物酵素作为一种富含多种生理活性成分的功能性食品,包括黄酮化合物、多酚类物质、氨基酸、有机酸、多糖以及多种功能性酶系。现有研究证实,这些活性成分普遍表现出抗氧化特性、解酒保肝、免疫调节、糖脂代谢调控以及促进消化吸收等多种功效^[1-5]。随着人们健康意识的提升,消费者对植物酵素产品的需求不断增长,科研人员也更倾向于运用同源性的药食两用资源进行酵素研发^[4]。不过,目前国内植物酵素的生产工艺与标准模式尚不完善,不同酵素产品所需的发酵条件、品质指标、菌种选择存在较大差异,且发酵产物的功能活性与安全评估仍需开展系统的科学研究^[6]。

药理学研究证实,茯苓多糖与三萜类物质能降低血清总胆固醇、甘油三酯及低密度脂蛋白的含量,提升高密度脂蛋白浓度^[7],茯苓酸等活性成分可以减轻肝脏、血管壁的脂质沉积^[8],具有较好的血脂调节能力。经常食用山药,发生糖尿病相关并发症的概率明显下降^[9]。山楂黄酮与三萜成分能够阻断胆固醇合成关键酶的活性,而其中所含有的有机酸、脂肪酶则能促进脂质分解^[10]。被大家普遍认可的功能性食品,燕麦 β -葡聚糖能结合肠道胆固醇,干扰其吸收过程,并且促进胆酸排出,进而降低低密度脂蛋白含量^[11]。临床研究表明,规律摄入适量的燕麦能够对血脂代谢紊乱起到有效的改善作用^[12]。杭白菊黄酮、菊甙类成分一方面抑制胆固醇沉积,另一方面改善血液流变学特性。糖尿病患者的临床监测显示,合理饮用杭白菊茶不仅可改善脂代谢紊乱,还能缓解血管并发症,并表现出一定的血糖调控效果^[13]。白果含有的不饱和脂肪酸、植物甾醇、黄酮类物质,通过抑制胆固醇合成关键酶,降低低密度脂蛋白和甘油三酯水平^[14]。

本研究拟采用山药、茯苓、桑叶、山楂、燕麦、白果及杭白菊作为原料,通过复合乳酸菌发酵工艺,促进新型降糖降脂活性成分的生成或富集,初步探讨

发酵产物对 α -淀粉酶、 α -葡萄糖苷酶、胰脂肪酶的抑制效应以及对脱氧胆酸盐的结合能力,为研发具备调节血糖血脂功能的保健食品提供依据。

1 材料与方法

1.1 材料与设备

山药、茯苓、桑叶、山楂、燕麦、白果、杭白菊(中药饮片,均购自同仁堂)、复合乳酸发酵菌粉(保加利亚乳杆菌、嗜酸乳杆菌、嗜热链球菌、乳双歧杆菌、植物乳杆菌混合菌粉,食品级)、牛磺脱氧胆酸钠($AR \geq 95\%$)、甘氨酸脱氧胆酸钠($AR \geq 98\%$)、考来烯胺(BR, 麦克林)、胰蛋白酶(1:250)、胃蛋白酶($3000 \text{ U} \cdot \text{mg}^{-1}$)、 α -淀粉酶($2000 \text{ U} \cdot \text{g}^{-1}$)、 α -葡萄糖苷酶($20 \text{ U} \cdot \text{mL}^{-1}$)、猪胰脂肪酶(10 万 U/g , BR, 麦克林)、碳酸钠、盐酸、硫酸、氯化钠、磷酸缓冲液、氢氧化钠、阿卡波糖(均为分析纯,国药集团)、3,5-二硝基水杨酸(3,5-Dinitrosalicylic acid, DNS, 上海创赛科)、对硝基苯基- β -D-半乳糖苷(4-Nitrophenyl β -D-galactopyranoside, PNPG)、4-硝基苯丁酸酯 4-Nitrophenyl butyrate, PNPB)。

DHG 电热恒温干燥箱(上海一恒)、756PC 型紫外可见分光光度计(上海恒平)、HH-8 型数显恒温水浴锅(常州天瑞仪器)、LCE 高速冷冻离心机和高压灭菌锅(均为力辰科技)、pH 计(METTLER TOLEDO)、超净工作台、中药粉碎机、SNJ-17C 型可调酸奶机。

1.2 实验方法

1.2.1 复方中药配方对酵素液 α -淀粉酶抑制率的影响

将山药、茯苓、桑叶、山楂、白果、燕麦及杭白菊分别研磨,红外灭菌处理后按表 1 配制 6 组实验配方(编号 1-6)。按 4:25 的固液比加入蒸馏水,混匀制成发酵基质,煮沸 15 min,冷却到 25 $^{\circ}\text{C}$,移取 5 mL 上清液,测定 pH 值并作为发酵初待测样品,剩

余体系发酵使用。各组接入 1% 复合乳酸菌, 无菌密封条件下 42 ℃ 连续培养 3 d, 测定不同中药配方发酵液 pH 值及对 α -淀粉酶活性的抑制率, 并与发酵前液比较分析。

表 1 复合中药酵素配方表(质量比/%)

配方	山药	茯苓	桑叶	山楂	白果	燕麦	杭白菊
1	10	9	8	7	1	3	2
2	10	9	8	7	2	3	1
3	9	10	9	6	1	3	2
4	9	10	9	6	2	2	2
5	8	8	10	9	1	2	2
6	8	8	10	9	2	2	1

1.2.2 复合菌接种量对 α -淀粉酶抑制率的影响

通过 1.2.1 试验结果来固定中药配方, 制备发酵基质, 分别接种菌粉 0、0.5%、1%、1.5%、2%、2.5%, 于 42 ℃ 静置发酵 3 d, 测定酵素液 pH 值及对 α -淀粉酶的抑制率。

1.2.3 发酵时间对 α -淀粉酶抑制率的影响

采用恒定配方与接种菌粉量配制发酵底物, 42 ℃ 恒温发酵, 分别在发酵第 0、1、2、3、4、5 d, 每日定时定量取样, 测定酵素液 pH 值以及对 α -淀粉酶的抑制率。

1.2.4 发酵温度对 α -淀粉酶抑制率的影响

固定发酵配方、菌接种量, 制备发酵基质, 分别在 36、38、40、42、44 ℃ 条件下进行一定时间的发酵, 测定酵素液 pH 值以及对 α -淀粉酶的抑制率。

1.2.5 响应面优化发酵工艺参数

以 α -淀粉酶抑制率为评价指标, 设计三因素三水平响应面法优化乳酸复合菌接种量、发酵时间及发酵温度等关键工艺参数, 并对预测结果进行验证。

1.2.6 酵素液对 α -淀粉酶抑制率的测定

将酵素液稀释到体积分数为 2%、4%、6%、8%、10% 的不同浓度梯度, 分别取 500 μ L 样品与 pH 6.8 的 α -淀粉酶溶液 ($6.5 \text{ U} \cdot \text{mL}^{-1}$) 等体积 37 ℃ 下混合 10 min, 加 1% 可溶性淀粉溶液 500 μ L 混匀, 37 ℃ 保温 10 min 后, 加 1 mL DNS 试剂, 沸水浴 5 min, 冰水浴冷却至 25 ℃, 加去离子水至 15 mL, 混合后测定 540 nm 波长处吸收值^[16], 发酵初始液及阿卡波糖液(浓度 $0.5 \text{ mg} \cdot \text{mL}^{-1}$) 的测定同样处理。

$$\text{抑制率}\% = \frac{(A_c - A_{bc}) - (A_s - A_{sc})}{A_c - A_{bc}} \times 100 \quad (1)$$

式中 A_s 为样品组吸光值, A_{sc} 为样品对照组吸光

值, A_c 为阴性对照组吸光值, A_{bc} 为阴性空白对照组吸光值。

1.2.7 酵素液对 α -葡萄糖苷酶抑制率的测定

将酵素液稀释到体积分数为 10%、20%、30%、40%、50% 的不同浓度梯度, 分别取 50 μ L 样品与 100 μ L α -葡萄糖苷酶 ($20 \text{ U} \cdot \text{mL}^{-1}$, pH 6.8) 37 ℃ 混合 10 min, 加 $20 \text{ mmol} \cdot \text{L}^{-1}$ 底物 PNPB 50 μ L。37 ℃ 保温 1 h, 加入 50 μ L 1 M Na_2CO_3 溶液终止反应, 400 nm 处测吸收值。按照公式(1) 计算对 α -葡萄糖苷酶的抑制率^[15], 发酵初始液及阿卡波糖液(浓度 $0.5 \text{ mg} \cdot \text{mL}^{-1}$) 的测定同样处理。

1.2.8 酵素液对胰脂肪酶抑制率的测定

将酵素液稀释到体积分数为 10%、20%、30%、40%、50% 的不同浓度梯度, 分别取 50 μ L 样品, 加入 100 μ L 胰脂肪酶溶液 ($0.125 \text{ mg} \cdot \text{mL}^{-1}$, pH 7.4), 混合均匀, 37 ℃ 下保温 10 min, 加 50 μ L 11.2 mM PNPB 溶液, 37 ℃ 下保温反应 20 min, 在 405 nm 处测吸光度, 按照公式(1) 计算脂肪酶的抑制率^[16], 发酵前液及奥利司他(浓度 $8 \text{ } \mu\text{g} \cdot \text{mL}^{-1}$) 的测定同样处理。

1.2.9 酵素液对脱氧胆酸盐结合率的测定

配制甘氨酸脱氧胆酸钠 (0.03 、 0.06 、 0.12 、 0.18 、 0.24 、 $0.3 \text{ mmol} \cdot \text{L}^{-1}$) 和牛磺脱氧胆酸钠 (0.05 、 0.1 、 0.15 、 0.2 、 0.25 、 $0.3 \text{ mmol} \cdot \text{mL}^{-1}$) 标准溶液, 分别移入 1 mL 胆酸盐溶液、3 mL 60% H_2SO_4 溶液, 混匀, 70 ℃ 水浴保温 40 min, 再冰浴 5 min, 387 nm 处测吸光度^[17]。

甘氨酸脱氧胆酸钠标准曲线为 $y = 0.128x - 0.0279$, $R^2 = 0.9826$, 牛磺脱氧胆酸钠标准曲线为 $y = 0.24x - 0.169$, $R^2 = 0.9782$ (其中 x 为胆酸盐浓度, y 为吸光度值)。

将酵素液稀释到体积分数为 2%、4%、6%、8%、10% 的不同浓度梯度, 分别取 200 μ L 样品, 加等体积 $10 \text{ mg} \cdot \text{mL}^{-1}$ 胃蛋白酶, 加 0.01 M 盐酸溶液 600 μ L, 37 ℃ 振荡消化 1 h; 用 1 mM NaOH 溶液缓慢调节 pH 至 6.3, 加入 $10 \text{ mg} \cdot \text{mL}^{-1}$ 胰蛋白酶 800 μ L, 37 ℃ 振荡消化 1 h; 结束后分别加入 800 μ L 0.4 mM 甘氨酸脱氧胆酸钠、800 μ L 0.5 mM 牛磺脱氧胆酸钠, 37 ℃ 振荡 1 h。结束后 8000 $\text{r} \cdot \text{min}^{-1}$ 离心 10 min, 取 1 mL 上清液, 加 3 mL 60% H_2SO_4 混匀, 70 ℃ 水浴 40 min, 冰浴 5 min, 在 387 nm 处测吸光度, 计算胆酸盐含量, 按公式(2)、公式(3) 计算甘氨酸脱氧胆酸钠和牛磺脱氧胆酸钠结合率^[17]。标准品、样品及酶液的配制均以 pH6.3 的 PBS 为溶剂, 发酵前液及考来烯胺 ($1 \text{ mg} \cdot \text{mL}^{-1}$) 溶液的测定同样处理。

$$\text{甘氨酸脱氧胆酸钠结合率}\% = \frac{C_1 - C_2}{C_1} \times 100 \quad (2)$$

式中: C_1 为甘氨酸脱氧胆酸钠加入量, μmol ; C_2 为甘氨酸脱氧胆酸钠剩余量, μmol 。

$$\text{牛磺脱氧胆酸钠结合率}\% = \frac{C_3 - C_4}{C_3} \times 100 \quad (3)$$

式中: C_3 为牛磺脱氧胆酸钠加入量, μmol ; C_4 为牛磺脱氧胆酸钠剩余量, μmol 。

1.3 数据分析

工艺优化过程各试验重复 3 次,均以未接种菌作空白对照。采用 WPS office 软件对数据进行图表处理,Design-Expert 软件对星点设计进行响应面分析,以 * 或#表示显著差异($P < 0.05$),* * 或##表示显著极差异($P < 0.01$)。

2 结果与分析

2.1 单因素试验结果分析

2.1.1 原料配方对 α -淀粉酶抑制率及 pH 值的影响

如图 1 所示,各组配方发酵后,对 α -淀粉酶抑制率均极显著增加($P < 0.01$)。其中 2 号配方对 α -淀粉酶抑制率最高,达到 91.31%,此时发酵液 pH 值最低 3.2。这表明发酵乳酸复合菌的生长状态良好,故配方 2 作为最佳选择。

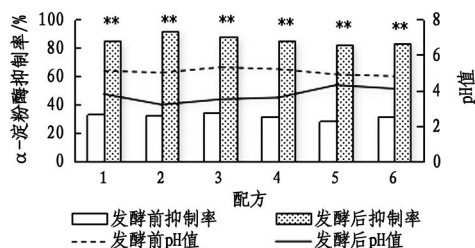


图 1 不同配方发酵液对 α -淀粉酶抑制率及 pH 值的影响

2.1.2 菌接种量对 α -淀粉酶的抑制率及 pH 值的影响

如图 2 所示,配方 2 发酵后,各组对 α -淀粉酶抑制率均极显著增加($P < 0.01$)。当复合乳酸菌添加比例处于 0.5% - 1.5% 范围时,发酵液对 α -淀粉酶的抑制率逐渐增强,最高 93.01%。超过 1.5% 的接种量,抑制率稍有减弱。随着菌种添加量的提升,培养体系的 pH 值显著降低,表明微生物处于理想的增殖状态。基于上述实验结果,确定复合乳酸菌适宜接种量为 1.5%。

2.1.3 发酵时间对 α -淀粉酶抑制率及 pH 值的影响

如图 3 所示,从发酵开始至第 2 d,发酵液 pH

值逐渐下降,表明乳酸菌处于高活性代谢状态,有机酸生成效率较高,此阶段发酵液对 α -淀粉酶抑制率持续上升至 93.27%,第 3 d 稍有下降,当发酵进入第 4 - 5 d 阶段,酶活性抑制效果显著下降,表明 2 d 发酵周期可获得最佳效果。

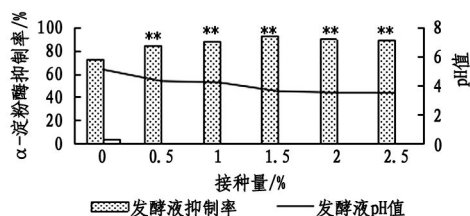


图 2 接种量对 α -淀粉酶抑制率及 pH 值的影响

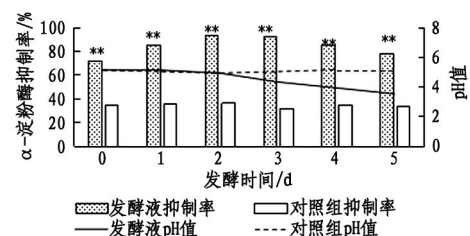


图 3 发酵时间对 α -淀粉酶抑制率及 pH 值的影响

2.1.4 发酵温度对 α -淀粉酶抑制率及 pH 值的影响

如图 4 所示,随着发酵温度提高,发酵液 pH 值逐渐下降,表明相对较高温度有利于乳酸复合菌的增殖。pH 值在低于 40 °C 范围内下降较快,超过 40 °C 时则轻微下降。发酵液对 α -淀粉酶的抑制率恰好在 40 °C 时达到最大,达到 94.03%,故初步确定 40 °C 为最佳发酵温度。

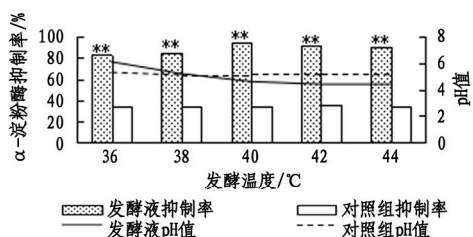


图 4 发酵温度对 α -淀粉酶抑制率及 pH 值的影响

2.2 响应面试验结果分析

采用 Design-expert 11 软件对应的 Central Composite 制定优化方案。在配方 2 的基础上,以复合菌接种量(A)、发酵时间(B)及发酵温度(C)三个因素为响应面变量,以 α -淀粉酶抑制率(Y)为响应值优化发酵工艺,结果如表 2 和表 3 所示。

表 2 响应面因素水平及试验结果表

序号	A/%	B/d	C/℃	Y/%
1	1.5	3	38	78.23
2	2	1	40	81.15
3	1.5	1	42	81.5
4	1.5	2	40	94.65
5	1	2	42	80.06
6	1.5	3	42	88.34
7	1.5	2	40	95.98
8	2	3	40	86.37
9	1	1	40	78.63
10	1.5	2	40	93.72
11	1	3	40	81.69
12	1.5	2	40	94.56
13	1	2	38	79.35
14	1.5	2	40	94.79
15	2	2	38	80.1
16	1.5	1	38	77.47
17	2	2	42	86.65

表 3 响应面分析结果表

方差来源	平方和	自由度	均方	F 值	P 值	显著性
模型	741.6	9	82.4	67.08	<0.0001	*

A	26.43	1	26.43	21.51	0.0024	*
B	31.52	1	31.52	25.66	0.0015	*
C	57.25	1	57.25	46.61	0.0002	**
AB	1.17	1	1.17	0.95	0.3623	
AC	8.53	1	8.53	6.94	0.0337	*
BC	9.24	1	9.24	7.52	0.0288	*
A ²	167.78	1	167.78	136.6	<0.0001	**
B ²	176.12	1	176.12	143.39	<0.0001	**
C ²	199.74	1	199.74	162.61	<0.0001	**
残差	8.6	7	1.23			
失拟项	5.98	3	1.99	3.04	0.1554	不显著
纯误差	2.62	4	0.66			
总离差平方和	750.2	16				

建立的二次多项式回归模型为 $Y = 94.74 + 1.82A + 1.99B + 2.68C + 0.54AB + 1.46AC + 1.52BC - 6.31A^2 - 6.47B^2 - 6.89C^2$ 。决定系数 $R^2 = 0.9885$, 调整决定系数 $R^2_{adj} = 0.9738$, 表明该模型的预测较为可靠, 预测采用配方 2, 接种复合乳酸菌 1.5%, 40℃ 发酵 2 d, 可获得对 α-淀粉酶 94.74% 的抑制率。

由图 5 所示, 乳酸复合菌接种量与发酵温度互作效应显著, 发酵时间与发酵温度互作效应显著。

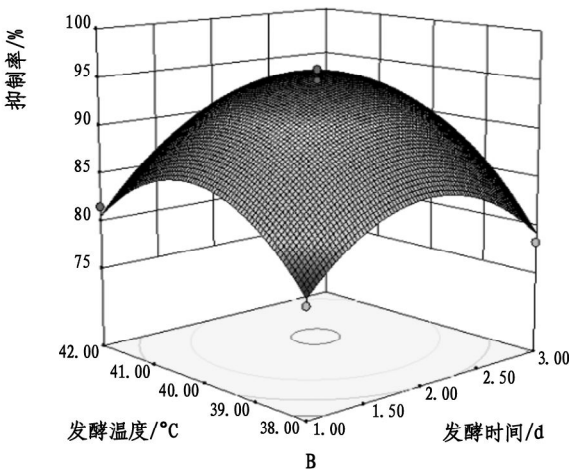
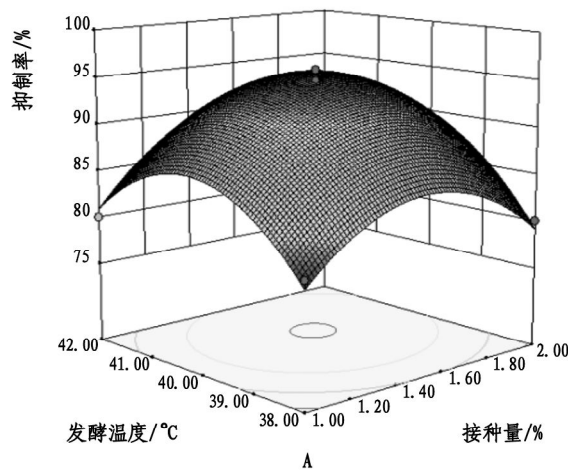


图 5 发酵过程中的互作效应

2.3 酵素液对三种酶活性的体外抑制率

由图 6 可知, 随着发酵产物浓度的增加, 发酵液对 α-淀粉酶、α-葡萄糖苷酶及脂肪酶的抑制率均逐渐增大, 呈剂量依赖关系。

由图 6A 可知, 酵素液含量分数为 4% 时与阿卡波糖对 α-淀粉酶抑制率 (63.62%) 相当, 当含量分数达到 10% 时, 对 α-淀粉酶的抑制率最大 (94.2%), 与预测值较为一致, 与发酵前 20.32% 相

比提高 3.64 倍,差异极显著($P < 0.01$)。

由图 6B 可知,酵素液对 α -葡萄糖苷酶抑制率随着含量分数的增加呈现递增的趋势,但与阿卡波糖液(84.42%)相比要偏低,当酵素液含量分数为 50% 时,对 α -葡萄糖苷酶的抑制率最大,达到 66.47%,与发酵前 22.13% 相比提高 2.00 倍,差异极显著($P < 0.01$)。

由图 6C 可知,酵素液含量分数为 40% 时与奥利司他对脂肪酶的抑制率(50.63%)相当,当含量分数达到 50% 时,对脂肪酶的抑制率最大,达到 69.57%,与发酵前 31.25% 相比提高 1.16 倍,差异极显著($P < 0.01$)。这表明发酵液具有抑制肠道消化酶的能力,一定程度上能够减少食源性糖脂的摄入。

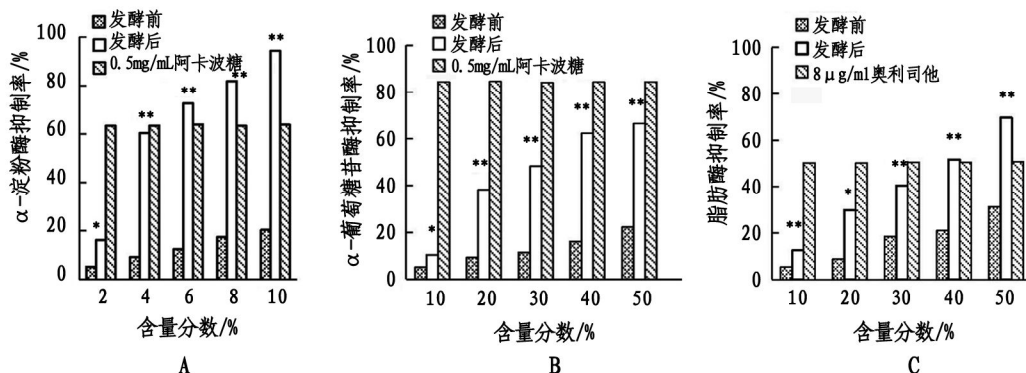
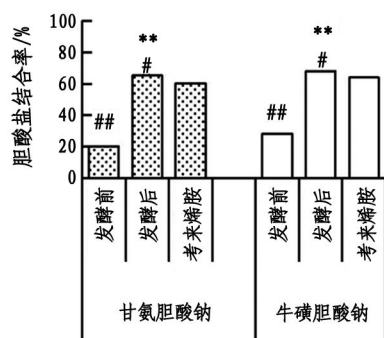


图 6 酵素液对 α -淀粉酶、 α -葡萄糖苷酶、胰脂肪酶的抑制率的影响

2.4 中药酵素液与胆酸盐的结合率

如图 7 所示,复方中药发酵前液对甘氨酸脱氧胆酸钠的吸附率为 20.37%,经发酵处理后提升至 65.24%,吸附效率上升 2.20 倍,差异极显著($P < 0.01$),考来烯胺对照组吸附率为 60.42%,差异显著($P < 0.05$)。在牛磺脱氧胆酸钠体系中,发酵前液结合率为 28.37%,发酵后增至 67.82%,提升幅度达 1.39 倍,差异极显著($P < 0.01$),考来烯胺对照组吸附率为 63.85%,差异显著($P < 0.05$)。实验表明,发酵后样品对脱氧胆酸盐的吸附能力显著增强,提示发酵过程中可能生成或富集了具有生物活性组分,这些组分在胆酸盐的吸附过程中可能发挥了作用。



注: #表示与考来烯胺相比差异显著($P < 0.05$);
##表示与考来烯胺相比差异极显著($P < 0.01$)。

图 7 复方中药发酵液与胆酸盐的结合率

3 结论与讨论

α -淀粉酶作为糖代谢关键酶之一,分解淀粉分子内的 α -1,4-糖苷键,生成短链糊精、麦芽糖和少量葡萄糖。 α -葡萄糖苷酶则作用于双糖或寡糖的末端 α -1,4-糖苷键,将其分解为葡萄糖为主的单糖分子,这两类酶加速食物消化后的血糖上升。反之,抑制这两类酶的催化活性,能有效控制餐后血糖水平。研究发现,植物乳杆菌发酵的蓝莓汁^[18]以及乳酸发酵接骨木果汁^[19]均显著提升对 α -淀粉酶和 α -葡萄糖苷酶的抑制率。曹英等^[20]从驼乳中分离出 6 种抑制 α -淀粉酶活性的乳酸菌株,其中最高抑制率达到 88.59%。本研究采用 5 种复合乳酸菌对山药、桑叶、茯苓、山楂、燕麦、白果及杭白菊等中药材进行共发酵,也获得了药食同源酵素液对 α -淀粉酶、 α -葡萄糖苷酶抑制率显著提升的结果。

此外,胰脂肪酶将食物中的甘油三酯分解为甘油、脂肪酸和甘油一酯,在胆汁酸盐的辅助下吸收入血,抑制该酶活性可以减少机体对脂类物质的消化与吸收。再者,外源物质与胆酸盐的结合,可以有效减少肠道胆汁酸的重吸收,促进肝脏加速胆固醇代谢以补充胆酸储备。本试验中发酵产物与未发酵前相比,对脂肪酶的抑制率显著提升,与甘氨酸脱氧胆酸钠和牛磺脱氧胆酸的结合力显著增强。潘胜国等^[21]发现乳酸发酵枇杷汁能显著抑制脂肪酶活力,且对胆汁酸盐的吸附能力也大大提升。党娅等^[22]研制的蓝莓酵素液证实对高血脂小鼠有良好的降脂

效果。大量研究表明^[23-24],选择合适的原料、发酵菌群及发酵工艺可以减少脂类的吸收,间接促进体内胆固醇的分解代谢,从而调节血脂。可见,通过微生物发酵并控制发酵条件以优化目标产物的合成可能是降糖降脂的重要策略之一。

相较于化学合成制剂,天然来源的功能成分具有更好的安全性和经济性,这使得基于天然产物的血糖血脂调节剂成为糖尿病干预领域的研究热点。发酵体系的 pH 变化是监控微生物代谢活动的主要参数之一,能够反映微生物的生长状态。本研究未对菌体生长、发酵液功能组分与糖脂代谢关键酶等关联性进行定量测定,后续可深入解析其组分构成及作用机制,以促进药用植物资源在健康食品产业中的开发与利用。

参考文献:

- [1] 陈雨欣,杨峰山,付海燕,等. 食用植物酵素发酵过程中主要成分与功能研究进展[J]. 中国农学通报,2024,40(24):143-150.
- [2] 李艳杰. 食用植物酵素菌种选择及功能作用研究进展[J]. 辽宁林业科技,2023(6):56-59.
- [3] 王桂丹. 植物酵素发酵过程中几种功效酶的动态变化及减肥功效评价[J]. 食品安全导刊,2021(29):112-113.
- [4] 周亚楠,郑舒怡,柳嘉雯,等. 植物酵素代谢产物形成机制研究进展[J]. 食品工业科技,2024,45(15):380-391.
- [5] 饶智,陈彦坤,刘斌,等. “药食同源”植物酵素研究进展[J]. 食品与发酵工业,2020,46(9):290-294.
- [6] 郭小舟,刘昱辛,吴越,等. 药食同源发酵中药产业发展现状、问题与对策[J]. 药食同源杂志,2025,1(4):1-8.
- [7] 赵辉. 基于 UPLC-HDMS 的代谢组学研究茯苓和茯苓皮抗高脂血症大鼠的生物化学作用机制[D]. 西安:西北大学,2020.
- [8] 关彦豪. 茯苓酸改善非酒精性脂肪肝脂质堆积的机制研究[D]. 广州:广州中医药大学,2023.
- [9] 龚凌霄,池静雯,王静,等. 山药中主要功能性成分及其作用机制研究进展[J]. 食品工业科技,2019,40(16):312-319.
- [10] 张亮亮. 山楂不同提取物的成分分析、生物活性及抑菌机理研究[D]. 临汾:山西师范大学,2020.
- [11] 于永华,徐飞飞,林琳,等. 燕麦 β -葡聚糖降血脂作用的研究进展[J]. 临床与病理杂志,2022,42(2):486-491.
- [12] 高晓晗,王利国,张诗文,等. 燕麦青稞粉对 II 型糖尿病小鼠糖脂代谢及体内氧化应激的影响[J]. 中国食品学报,2022,22(6):190-197.
- [13] 陈明珠. 杭白菊中黄酮成分的抗糖尿病活性及其机制研究[D]. 天津:天津科技大学,2019.
- [14] 夏梦雨,张雪,王云,等. 白果的炮制方法、化学成分、药理活性及临床应用的研究进展[J]. 中国药房,2020,31(1):123-128.
- [15] 李云姣,李琪,杜佳峰,等. 水果酵素体外抗氧化及抑制 α -淀粉酶和 α -葡萄糖苷酶活性的研究[J]. 中国食品学报,2019,19(4):79-84.
- [16] 史方丹,周际松,李敏,等. 核桃谷蛋白的营养价值及其降脂活性[J]. 食品工业科技,1-17[2026-04-19].
- [17] 余金毅,王红波,方云龙,等. 白扁豆花及其发酵品粗多糖的结构特征与抗氧化、降糖、降脂活性[J]. 中国食品学报,2025,25(9):61-74.
- [18] Zhang Yu, Liu Weipeng, Wei Zehua, et al. Enhancement of functional characteristics of blueberry juice fermented by *Lactobacillus plantarum*[J]. LWT-Food Science and Technology, 2021, 139:110590.
- [19] 封弦,翁佩芳,吴祖芳,等. 乳酸菌发酵接骨木果汁降血糖与抗氧化活性机理[J]. 食品与生物技术学报,2022,41(8):95-103.
- [20] 曹英,侯敏,易光平,等. 驼乳制品中抑制 α -淀粉酶和 α -葡萄糖苷酶活性乳酸菌的筛选及益生特性研究[J]. 食品工业科技,2022,43(19):191-201.
- [21] 潘胜国,王群,刘制杰,等. 乳酸菌发酵对枇杷汁体外降血糖、降血脂活性的影响[J]. 食品工业科技,2025,46(7):161-168.
- [22] 党娅,乔雨雨,雷格格,等. 蓝莓酵素对高脂血症小鼠的降血脂及肠道菌群调节作用[J]. 食品工业科技,2025,46(10):360-370.
- [23] 张田宇,周晨露,范晓源,等. 食用酵素的降血脂作用研究进展[J]. 食品安全导刊,2022(10):128-133.
- [24] 颜梦秋. 食药两用菌液体发酵及功能活性成分研究现状与展望[J]. 微生物学报,2023,50(7):3211-3231.

责任编辑:汪张贵

复合低聚糖固体饮料制备及其抗氧化活性研究

郝梦媛,张雨晨,杜传来,徐静静*

(安徽科技工程大学 食品科学与工程学院,安徽 滁州 239000)

摘要:参考膳食宝塔中不同膳食成分的摄入量和人体每日膳食纤维推荐摄入量,开发了一款主要成分为低聚木糖、低聚果糖、壳寡糖和低聚半乳糖的复合低聚糖固体饮料。通过单因素和正交试验优化复合低聚糖固体饮料的配方,并测定饮料的理化指标和抗氧化活性。结果表明,该饮料的配方是纯净水250 mL,低聚果糖12 g,低聚木糖6 g,壳寡糖2 g,低聚半乳糖9 g;该饮料的还原糖含量 (3.31 ± 0.04) g/100 g,总酸含量 (0.256 ± 0.003) g/100 g,pH值4.28,可溶性固形物含量10.00%,饮料口感酸甜适中;体外抗氧化活性分析表明复合低聚糖固体饮料的自由基清除率和铁离子还原能力显著大于市售低聚果糖固体饮料。该研究可为功能性低聚糖产品的开发提供参考。

关键词:膳食宝塔;复合低聚糖;固体饮料;抗氧化活性;功能性低聚糖

中图分类号:TS278

文献标识码:A

文章编号:2095-297X(2026)05-0015-07

Preparation and Antioxidant Activity of Compound Oligosaccharide Solid Beverage

HAO Mengyuan,ZHANG Yuchen,DU Chuanlai,XU Jingjing*

(School of Food Science and Engineering,Anhui Science and Technology University,Chuzhou,239000,Anhui)

Abstract:A novel compound oligosaccharide solid beverage was developed in the paper,with main components including xylo-oligosaccharides,fructo-oligosaccharides,chitosan oligosaccharides,and galacto-oligosaccharides,based on the intake of different dietary components from Chinese Food Guide Pagoda and the recommended daily intake of dietary fiber for humans. The formula of the compound oligosaccharide solid beverage was optimized through single-factor and orthogonal experiments,and the physicochemical properties,and the antioxidant activity of this beverage were determined. The results showed that the formula of compound oligosaccharide solid beverage was 250 mL of purified water,12 g of fructo-oligosaccharides,6 g of xylo-oligosaccharides,2 g of chitosan oligosaccharides,and 9 g of human milk oligosaccharides. This beverage had a reducing sugar content of (3.31 ± 0.04) g/100g,a total acid content of (0.256 ± 0.003) g/100g,a pH value of 4.28,and a soluble solid content of 10.00%. This beverage was a sweet and sour taste. The free radical scavenging rate and ferric reducing antioxidant power of the compound oligosaccharide solid beverage were significantly higher than those of commercially fructo-oligosaccharides solid beverages by in vitro antioxidant activity analysis. This research will provide a reference for the development of functional oligosaccharide products.

Key words:Chinese Food Guide Pagoda;compound oligosaccharides;solid beverage;antioxidant activity;functional oligosaccharides

低聚糖是一类由2-10个单糖分子通过糖苷键 连接形成的碳水化合物,具有益生元^[1]特性、调节

收稿日期:2025-11-25 * 通讯联系人

基金项目:广东省基础与应用基础研究基金(2022A1515110247);净化环境的活性益生菌产品开发项目(882176);安徽省应用型高峰培育学科(840122)。

作者简介:郝梦媛(1996—),女,山西阳泉人,硕士研究生。E-mail:1499880931@qq.com

糖脂代谢等多种生理活性。在传统的食品加工^[2]领域,低聚糖的应用主要集中在大豆制品^[3]、果蔬制品等,来源单一,难以发挥各组分间的功能互补。随着消费者意识的增强与需求的多样化,功能性饮料的研发从“单一功能”向“复合功能”转变。李文强等^[4]将琼脂低聚糖添加到乳酸菌饮料中,丰富了低聚糖在饮料中的应用。

低聚糖固体饮料^[5]是一种以低聚糖为主要成分的固体饮料,具有多种功效。低聚糖固体饮料中的益生元特性,可以促进肠道内有益菌群^[6]的生长。与传统的固体饮料相比,低聚糖固体饮料具有易消化,溶解速度快,携带方便的特点,且呈现多种保健功能。然而目前低聚糖固体饮料成分单一^[7],复合低聚糖功能饮料亟待开发与创新。

研究参考膳食宝塔^[8]中不同膳食成分的摄入量 and 人体每日膳食纤维推荐摄入量(25–30 g),开发出一款复合低聚糖固体饮料^[9]。该固体饮料以低聚木糖^[10]、低聚果糖^[11]、壳寡糖^[12]和低聚半乳糖^[13]为主要功能成分,低聚木糖的添加量参考每日膳食纤维推荐摄入量 and 主食类食物推荐摄入比例,低聚果糖的添加量参考每日膳食纤维推荐摄入量 and 果蔬类食物推荐摄入比例,壳寡糖参考每日膳食纤维推荐摄入量 and 动物类食物推荐摄入比例,低聚半乳糖参考每日膳食纤维推荐摄入量 and 奶制品类食物推荐摄入比例,然后检测固体饮料的理化指标和抗氧化活性,为功能性固体饮料的研制提供参考。

1 材料与方法

1.1 实验材料

低聚果糖(食品级)、低聚半乳糖(食品级),山东科源生化有限公司;低聚木糖(食品级),山东龙力生物科技有限公司;壳寡糖(食品级),山东英才生物科技有限公司;1,1-二苯基-2-苦肼基(DPPH)、2-联氮-二(3-乙基-苯并噻唑-6-磺酸)二铵盐(ABTS),上海麦克林生化科技股份有限公司;乙酸锌、亚铁氰化钾、葡萄糖、氢氧化钠、酚酞等,上海阿拉丁生化科技股份有限公司;0.1 mol·L⁻¹醋酸钠缓冲液(pH3.6)、10 mmol·L⁻¹氯化三苯四氮唑(TPTZ),福州飞净生物科技有限公司;H₂O₂,天津市永大化学试剂有限公司;FeCl₃·6H₂O、水杨酸、

FeSO₄·7H₂O,国药集团化学试剂有限公司。

1.2 实验仪器

FA224 电子分析天平,上海力辰邦西仪器科技有限公司;SFG-02B 恒温鼓风干燥箱,黄石市恒丰医疗器械有限公司;YCD-EL300 冰箱,中科美菱低温科技股份有限公司;LGJ-10 冷冻干燥机,北京松源华兴科技发展有限公司;多功能粉碎机,东莞市华太电器有限公司;FB2200PH 计,奥豪斯仪器常州有限公司;WZS-32 手持式折光仪,上海仪电物理光学仪器有限公司;DSX-18L-I 高压灭菌锅,上海申安医疗器械厂;SHZ-C 恒温水浴振荡器,上海蚁霖科学有限公司。

1.3 实验方法

1.3.1 复合低聚糖固体饮料的制备

通过超微粉碎机对低聚果糖、低聚木糖、壳寡糖和低聚半乳糖分别进行粉碎,利用 120 目筛进行干法筛分以分别去除以上四个低聚糖中的杂质及粒径较大的颗粒,保证原料粒径均一性。将过筛后的低聚果糖、低聚木糖、壳寡糖、低聚半乳糖分别按照一定的比例添加到小型三维混合机中进行充分混匀,得到复合低聚糖固体粉末。称取 29 g 复合低聚糖固体粉末加入到 250 mL 纯净水中,搅拌溶解,然后通过 100 °C 蒸汽灭菌 30 min,冷却至室温后,在无菌条件下分装至食品级瓶子,旋紧密封即得复合低聚糖固体饮料。

1.3.2 单因素实验

本实验控制纯净水的添加量为 250 mL,分别对低聚果糖的添加量(10、11、12、13、14 g),低聚木糖的添加量(4、5、6、7、8 g),壳寡糖的添加量(2、2.5、3、3.5、4 g),低聚半乳糖的添加量(5、6、7、8、9 g),进行单因素^[14]实验,观察主要因素对饮料品质的影响,进行感官评分。

1.3.3 正交试验

为了获得最佳配方的复合低聚糖固体饮料,在单因素试验的基础上,观察低聚果糖、低聚木糖、壳寡糖和低聚半乳糖的添加量对固体饮料感官评分的影响,采用四因素三水平 L₉(3⁴)的正交试验设计,以感官评分为主要指标,对生产配方进行优化,如表 1 所示,确定复合低聚糖固体饮料的最佳配方^[15]。

表 1 正交试验因素水平表

水平	因素			
	A 低聚果糖添加量	B 低聚木糖添加量	C 壳寡糖添加量	D 低聚半乳糖添加量
1	11	5	2	7
2	12	6	2.5	8
3	13	7	3	9

1.3.4 感官评分

选择 15 人作为评分小组,对复合低聚糖固体饮料^[16]的感官进行评价。感官评价包括色泽与形态(30 分)、风味(30 分)、口感(40 分),总计 100 分,评分标准如表 2 所示,对复合低聚糖固体饮料的各项指标进行打分,得到最终得分^[17]。

表 2 复合低聚糖固体饮料的感官评价标准表

感官特征	评分标准	分数/分
色泽与形态 (30 分)	淡黄色,透明有光泽,质地均一	25-30
	淡黄色,透明,质地基本均一	15-25
	黄色,不透明,质地不均	1-15
风味 (30 分)	酸甜适中,无异味	20-30
	酸甜较适中,轻微异味	10-20
	酸甜失调,味道寡淡,异味较明显	1-10
口感 (40 分)	口感顺滑,清爽,无粗糙感	30-40
	口感较顺滑,较清爽,微有粗糙感	20-30
	口感较差,不清爽,粗糙感明显	10-20

1.3.5 理化指标的测定

还原糖含量的测定。取 20 mL 复合低聚糖固体饮料,加入 50 mL 水,依次加入乙酸锌溶液和亚铁氰化钾溶液各 5 mL,混匀后静置 30 min,备用。分别吸取碱性酒石酸铜甲液和乙液各 5 mL 于锥形瓶中,加入 10 mL 水和 2-4 粒玻璃珠,加热至沸腾后,用葡萄糖标准溶液滴定,至溶液蓝色完全褪尽且 15 s 内不恢复为终点,平行测定 3 次,取平均值计算碱性酒石酸铜溶液当量。随后,按同样方法分别吸取碱性酒石酸铜甲液和乙液各 5 mL,加入 10 mL 水和 2-4 粒玻璃珠,加热至沸腾后,用上述样品溶液进行滴定,至蓝色完全褪尽为终点。平行测定 3 次,计算还原糖含量。还原糖的含量可根据 GB 5009.7-2016《食品安全国家标准 食品中还原糖的测定》的方法,按照公式(1)计算:

$$X = \frac{m_1}{m \times F \times V/250 \times 1000} \times 100 \quad (1)$$

式中: X 为试样中还原糖的含量(g/100 g); m_1 为碱

性酒石酸铜溶液(甲、乙液各半)相当于某种还原糖的质量(mg); m 为试样的质量(g); F 为换算系数,含淀粉的食品为 0.8,其他食品为 1; V 为滴定时所消耗试样溶液的体积(mL);250 为定容体积(mL);100 为换算系数。

总酸的测定。取 20 mL 复合低聚糖固体饮料,滴加 2 滴酚酞指示液,用 0.1 mol·L⁻¹ 氢氧化钠溶液进行滴定,30 s 内不褪色,即为滴定终点,每个样品平行测定三次,取平均值。根据 GB 12456-2021《食品安全国家标准 食品中总酸的测定》的方法,按照公式(2)计算:

$$X = \frac{[c \times (V_1 - V_2)] \times k \times F}{m} \times 1000 \quad (2)$$

式中: X 为总酸的含量; c 为氢氧化钠的浓度(mol·L⁻¹); V_1 为滴定时消耗氢氧化钠的体积(mL); V_2 为空白试验时消耗氢氧化钠的体积(mL); k 为酸的换算系数; F 为试液的稀释倍数; m 为试样的质量。

pH 值的测定。pH 值的测定参考 GB 5009.237-2016《食品安全国家标准 食品 pH 值的测定》。

可溶性固形物的测定。可溶性固形物含量测定参考 GB/T 12143-2008《饮料通用分析方法》。

DPPH 自由基清除率的测定。吸取 0.1 mL 复合低聚糖固体饮料与 3.9 mL 新鲜配制的 DPPH 乙醇溶液混合,室温避光放置 60 min,测定 517 nm 处的吸光度,空白对照组是去离子水,DPPH 自由基清除率^[19]按公式(3)计算:

$$Y = \frac{A_c - A_s}{A_c} \times 100 \quad (3)$$

式中: Y 为样品的 DPPH 阳离子自由基清除率(%); A_c 为空白组的吸光度; A_s 为样品组的吸光度。

ABTS 阳离子自由基清除率的测定。吸取 20 uL 复合低聚糖固体饮料与 180 uL 的 ABTS 混合,测定 734 nm 处吸光值,去离子水作为空白对照组。ABTS 阳离子自由基清除率按公式(4)计算:

$$Y = \frac{A_c - A_s}{A_c} \times 100 \quad (4)$$

式中: Y 为样品的 ABTS 阳离子自由基清除率(%); A_c 为空白组的吸光度; A_s 为样品组的吸光度。

羟基自由基清除率测定。复合低聚糖固体饮料羟自由基清除方法如表3所示。取8 mL固体饮料样液、稀释5倍、稀释10倍的样品,标记为样品1、2、3,分别加入1 mL $9 \text{ mmol} \cdot \text{L}^{-1} \text{FeSO}_4$ 、1 mL $9 \text{ mmol} \cdot \text{L}^{-1}$ 水杨酸、1 mL $8.8 \text{ mmol} \cdot \text{L}^{-1} \text{H}_2\text{O}_2$,混匀之后置于37 ℃水浴锅加热30 min,在波长510 nm处测定吸光度 A_i ;用1 mL蒸馏水代替 $8.8 \text{ mmol} \cdot \text{L}^{-1} \text{H}_2\text{O}_2$ 测定吸光度 A_j ,去离子水作为空白对照组,测定吸光度 A_0 ,所有样品重复测试3次。按公式(5)计算羟自由基清除能力。

$$\text{羟自由基清除率} = \frac{A_0 - (A_i - A_j)}{A_0} \times 100\% \quad (5)$$

式中: A_0 为1 mL FeSO_4 + 1 mL 水杨酸 + 1 mL H_2O_2 + 8 mL 蒸馏水的吸光度; A_i 为1 mL FeSO_4 + 1 mL 水杨酸 + 1 mL H_2O_2 + 8 mL 固体饮料样液的吸光度; A_j 为1 mL FeSO_4 + 1 mL 水杨酸 + 1 mL 蒸馏水 + 8 mL 固体饮料样液的吸光度。

表3 复合低聚糖固体饮料羟自由基清除方法

试剂	组成/mL		
	A_0	A_i	A_j
$9 \text{ mmol} \cdot \text{L}^{-1} \text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	1	1	1
$9 \text{ mmol} \cdot \text{L}^{-1}$ 水杨酸-乙醇溶液	1	1	1
样品	0	8	8
蒸馏水	8	0	1
$8.8 \text{ mmol} \cdot \text{L}^{-1} \text{H}_2\text{O}_2$	1	1	0

铁离子还原能力测定。将醋酸钠缓冲液($300 \text{ mmol} \cdot \text{L}^{-1}$ 、 $\text{pH} = 3.6$)、TPTZ溶液($10 \text{ mmol} \cdot \text{L}^{-1}$)和 FeCl_3 溶液($20 \text{ mmol} \cdot \text{L}^{-1}$)按照10:1:1的体积比混匀,于37 ℃预热后得到FRAP工作液。分别吸取不同浓度的 FeSO_4 标准溶液($0.15 - 1.5 \text{ mmol} \cdot \text{L}^{-1}$)与900 μL FRAP工作液混匀,静置反应10 min,于593 nm处测定其吸光度,绘制 FeSO_4 溶液标准曲线。取50 μL 样品或去离子水(空白对照)按上述条件与FRAP工作液反应后测定吸光度值,根据标准曲线计算 FeSO_4 浓度($\text{mmol} \cdot \text{L}^{-1}$),即得到该样品的FRAP值,其大小可表明样品的铁离子还原能力。

1.4 数据处理

所有实验均重复测定3次,实验采用Origin 2024软件和SPSS 27软件进行作图和统计分析,实验结果均采用平均值 $\pm$ 和标准差表示。

2 结果与分析

2.1 单因素实验结果分析

2.1.1 低聚果糖添加量对复合低聚糖固体饮料感官评分的影响

低聚果糖可以改善饮料风味,其添加量会对饮料的感官评价有显著的影响^[20]。如图1所示,低聚果糖的添加量为10 g时,低聚果糖含量偏低,饮料甜味不足,导致整体风味偏淡,饮料口感略显寡淡。低聚果糖添加量在12 g以上时,甜度特征逐渐增强,与体系中酸味成分的协同作用减弱,造成酸甜比例失衡,从而使感官评分下降。低聚果糖添加量从10 g增加至14 g的过程中,感官评价总分呈现出先升高后降低的变化趋势,低聚果糖添加量为12 g时,饮料的感官评分最高,复合低聚糖固体饮料的甜度适中,与酸味形成优势互补,饮料的综合感官评分最高。由此可见,11 g、12 g和13 g的低聚果糖添加量可分别作为正交试验水平,有助于在后续实验中对配方进行优化,并分析与其他成分的相互关系。

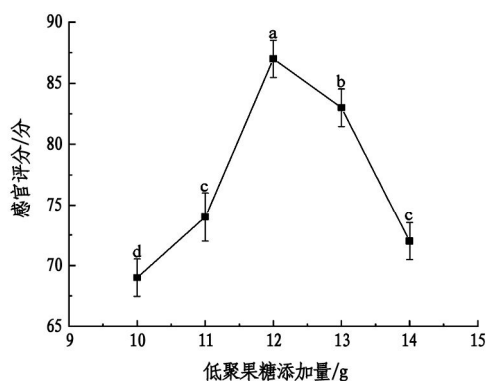


图1 低聚果糖添加量对饮料感官评分的影响

注:折线上小写字母不同表示差异显著($P < 0.05$),下同。

2.1.2 低聚木糖添加量对复合低聚糖固体饮料感官评分的影响

低聚木糖具有优良的热稳定性和酸稳定性,因而可作为饮料中的理想功能成分^[21]。如图2所示,低聚木糖添加量为4 g时,低聚木糖含量较少,浓度低于风味感知阈值,未能为饮料提供足够的风味强度与醇厚感,导致饮料整体口感单薄,风味强度不足;低聚木糖添加量在6 g以上时,低聚木糖带有轻微涩感和木质味,过量添加会打破酸、甜及其他风味成分之间的平衡,导致感官评分显著降低。低聚木糖添加量从4 g增加至8 g的过程中,感官评价总分呈现先升高后降低的趋势,低聚木糖添加量为6 g时,饮料的感官评分最高,复合低聚糖固体饮料酸甜适中,品质最佳。由此可见,5 g、6 g和7 g的低聚木糖添加量可分别作为正交试验因素水平,有助于确定其在复合固体饮料中的最优添加量。

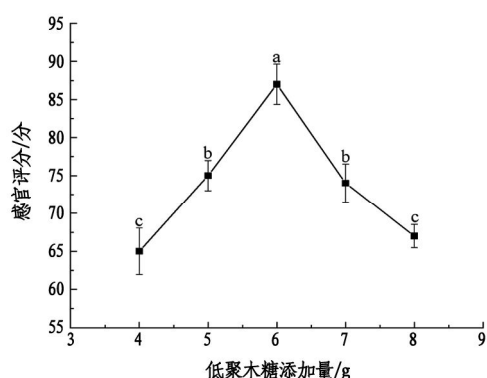


图2 低聚木糖添加量对饮料感官评分的影响

2.1.3 壳寡糖添加量对复合低聚糖固体饮料感官评分的影响

壳寡糖是壳聚糖经酶法降解等方式得到的低聚糖,具有良好的生物活性^[22]与很高的生物利用度,在功能型固体饮料的开发中展现出广阔的前景。如图3所示,壳寡糖添加量为2 g时,壳寡糖甜度较低,添加量较少,导致饮料风味不够,口感偏淡。壳寡糖添加量在2.5 g以上时,壳寡糖含量升高,不仅影响饮料色泽,使其透明度下降,而且壳寡糖含有微量的甲壳素,导致饮料异味感较重,感官评价总分反而降低。壳寡糖添加量从2 g增加至4 g的过程中,感官评价总分呈现先升高后降低的趋势,壳寡糖添加量为2.5 g时,饮料的感官评分最高,复合低聚糖固体饮料的口感酸甜适中,无异味,品质最佳。由此可见,2 g、2.5 g和3 g的壳寡糖添加量可分别作为正交试验的因素水平,进一步优化配方。

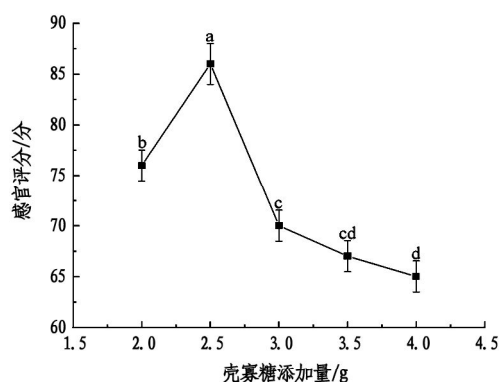


图3 壳寡糖添加量对饮料感官评分的影响

2.1.4 低聚半乳糖添加量对复合低聚糖固体饮料感官评分的影响

低聚半乳糖是3-10个半乳糖和葡萄糖经 β -半乳糖苷酶催化形成的低聚糖,具有典型的益生元特性,广泛用于固体饮料,可以促进肠道健康、提高免疫力。如图4所示,低聚半乳糖添加量为5 g时,低聚半乳糖含量较少,甜度较低,酸味过于突出,口

感协调较差。低聚半乳糖添加量在8 g以上时,低聚半乳糖含量较多,甜味过重,感官评价总分降低。低聚半乳糖添加量从5 g增加至9 g的过程中,感官评价总分呈现先升高后降低的趋势,低聚半乳糖添加量为8 g时,饮料的感官评分最高,酸甜适中,水溶液清澈透明,品质最佳。由此可见,7 g、8 g和9 g的低聚半乳糖添加量可分别作为正交试验的因素水平,进一步优化配方。

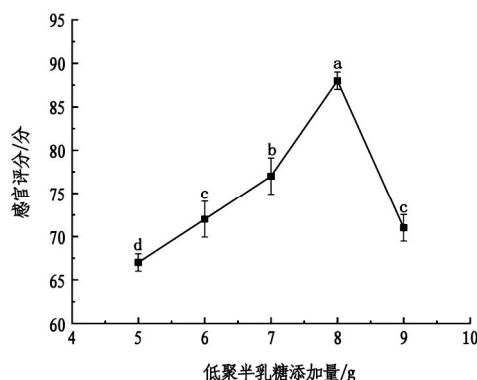


图4 低聚半乳糖添加量对饮料感官评分的影响

2.2 正交试验结果分析

研究采用 $L_9(3^4)$ 的正交试验设计,以低聚果糖、低聚木糖、壳寡糖、低聚半乳糖的添加量为4因素,每个因素设置3个水平,以感官评分作为评价标准进行优化。结果如表4所示。根据极差分析结果可知,各因素对复合低聚糖固体饮料感官品质的影响程度依次为:低聚果糖添加量>低聚半乳糖添加量>低聚木糖添加量>壳寡糖添加量。其中,低聚果糖的添加量对复合低聚糖固体饮料的感官品质影响最显著,从感官评价的总分数来看最优组合是 $A_2B_2C_1D_3$ 。

2.3 验证试验

在以上试验结果中,单因素实验的最佳组合是 $A_2B_1C_2D_3$,正交试验的最优组合是 $A_2B_2C_1D_3$,进一步通过感官评价试验来确定和验证最优组合。由表5试验结果得出,组合 $A_2B_2C_1D_3$ 的感官评分是 93 ± 0.2 ,大于 $A_2B_1C_2D_3$ 。因此组合 $A_2B_2C_1D_3$ 为复合低聚糖固体饮料的最佳配方,该配方条件下的复合低聚糖固体饮料颜色呈淡黄色,清亮透明有光泽,质地均匀,口感酸甜适中,无异味,整体品质最佳。因此,复合低聚糖固体饮料的最佳工艺配方是:纯净水250 mL,低聚果糖12 g,低聚木糖6 g,壳寡糖2 g,低聚半乳糖9 g。

表 4 正交试验设计及结果

试验号	A 低聚果糖添加量/g	B 低聚木糖添加量/g	C 壳寡糖添加量/g	D 低聚半乳糖添加量/g	感官评分/分
1	1	1	1	1	86
2	1	2	2	2	88
3	1	3	3	3	87
4	2	1	2	3	91
5	2	2	3	1	88
6	2	3	1	2	89
7	3	1	3	2	85
8	3	2	1	3	87
9	3	3	2	1	78
K_1	261	262	262	252	
K_2	268	263	257	262	
K_3	250	254	260	265	
k_1	87	87	87	84	
k_2	89	88	85	87	
k_3	83	85	86	88	
R	6	3	2	4	
因素影响等级			$A > D > B > C$		
最佳方案			$A_2B_2C_1D_3$		

表 5 固体复合饮料验证实验结果

组合	感官评分
$A_2B_1C_2D_3$	91 ± 0.3
$A_2B_2C_1D_3$	93 ± 0.2

2.4 理化性质分析

根据 GB5009.7-2016 中食品安全国家标准中还原糖的测定方法,测得复合低聚糖固体饮料的还原糖含量(3.31 ± 0.04)g/100 g。

根据 GB12456-2021 食品安全国家标准食品中总酸的测定方法,测得复合低聚糖固体饮料的总酸含量(0.256 ± 0.003)g/100 g。

根据 GB5009.237-2016 食品安全国家标准中 pH 值的测定方法,测得复合低聚糖固体饮料的 pH 值 4.28。

根据 GB/T12143-2008 饮料通用分析方法中可溶性固形物的测定方法,测得复合低聚糖固体饮

料的可溶性固形物含量是 10.00%。

复合低聚糖固体饮料和佳福蒙益生元低聚果糖固体饮料的抗氧化活性检测结果如表 6 所示。复合低聚糖固体饮料的 DPPH 自由基清除率、ABTS 阳离子自由基清除率和羟基自由基清除率分别是(36.3 ± 0.38)%、(41.3 ± 0.53)%和(62.1 ± 0.79)%,分别显著大于佳福蒙益生元低聚果糖固体饮料的(15.3 ± 0.20)%、(20.2 ± 0.35)%、(35.2 ± 0.43)%。铁离子还原能力的检测结果表明,复合低聚糖固体饮料的 FRAP 值为(0.69 ± 0.08)mmol $\cdot$ L $^{-1}$,高于佳福蒙益生元低聚果糖固体饮料的 FRAP 值。综上所述,复合低聚糖固体饮料的 DPPH 自由基清除能力、ABTS 阳离子自由基清除能力、羟自由基清除能力和铁离子还原能力均高于佳福蒙益生元低聚果糖固体饮料,表明复合低聚糖固体饮料具有良好的体外抗氧化活性水平。

表 6 固体复合饮料抗氧化活性测定结果

样品种类	DPPH 自由基清除能力/%	ABTS 阳离子自由基清除能力/%	羟自由基清除能力/%	FRAP 值 /mmol $\cdot$ L $^{-1}$
------	----------------	-------------------	------------	--------------------------------

佳福蒙益生元 低聚果糖固体饮料	15.3 ± 0.20	20.2 ± 0.35	35.2 ± 0.43	0.51 ± 0.10
复合低聚糖固体饮料	36.3 ± 0.38	41.3 ± 0.53	62.1 ± 0.79	0.69 ± 0.08

3 结论

通过单因素与正交试验确定复合低聚糖固体饮料最优配方是纯净水 250 mL,低聚果糖 12 g,低聚木糖 6 g,壳寡糖 2 g,低聚半乳糖 9 g。理化指标结果显示,该饮料的还原糖含量(3.31 ± 0.04) g/100 g,总酸含量(0.256 ± 0.003) g/100 g,pH 值 4.28,可溶性固形物含量 10.00%;抗氧化活性试验结果显示,该饮料 DPPH 自由基清除率(36.3 ± 0.38)%,ABTS 阳离子自由基清除率(41.3 ± 0.53)%,羟基自由基清除率(62.1 ± 0.79)%,铁离子还原能力(0.69 ± 0.08) mmol · L⁻¹ ($P < 0.05$)。

参考文献:

- [1] 王猛,胡金梅,游凤珍,等. 益智仁低聚糖的结构特征、体外活性及其对肠道菌群的调控机制[J]. 食品工业科技,2026,47(10):139-149.
- [2] 陈文彬,石会玲. 茺菁功能成分及其功能性食品研究进展[J]. 食品安全质量检测学报,2025,16(7):148-156.
- [3] 侯越,李振江,王立东. 大豆低聚糖的分离纯化研究进展[J]. 农产品加工,2023(16):79-83.
- [4] 李文强,胡佳麒,赵越,等. 琼脂低聚糖性质及在乳酸菌饮料中的应用研究[J]. 食品与发酵工业,2022,48(22):235-240.
- [5] 唐志峰,杨林. 红枣粉固体饮料加工研究进展[J]. 食品安全导刊,2025,19(28):138-140.
- [6] 赵人杰. 大蒜低聚糖的制备、消化特性及其益生作用[D]. 泰安:山东农业大学,2022.
- [7] 李伟. 麦芽低聚糖在食品生产中的应用分析[J]. 食品安全刊,2021(26):157,163.
- [8] 臧雪妍,关宁,伏彦慧. “中国居民平衡膳食宝塔”及其信息化应用[J]. 食品安全导刊,2023(36):97-100.
- [9] 王磊,黄燕巧,杨瑞利,等. 一种益生元复合固体饮料配方及通便功能研究[J]. 食品科技,2023,48(11):114-122.
- [10] Forsan F C,úvila F P, Masarin F, et al. Cellooligosaccharides and xylooligos-accharides: production processes, potential prebiotics, and metabolism routes by Lactobacillus and bifidobacterium [J]. Food Research International, 2025, 222(P1):117661.
- [11] Kaewsulud T, Liony M J, Tongdonyod S, et al. Development of fructoolig-osaccharide-rich sugarcane juice by enzymatic method and enhancement of its microbialsafety using high-pressure processing[J]. Foods (Basel, Switzerland), 2025, 14(19):3417.
- [12] 宋雪娜,李莉,赵奎,等. 功能性壳寡糖在食品领域的应用研究进展[J]. 食品研究与开发,2023,44(22):218-224.
- [13] Zhao Jingyu, Cao Yuan, Miao Yujiao, et al. Structure, enzymatic production, biological activities, and food applications of galacto-oligosaccharides: a review [J]. The Journal of Nutrition, 2025, 155(12):4113-4122.
- [14] 岑燕霞,银倍侨,梁玉才,等. 药食同源黄精复合饮料的配方优化[J]. 食品安全导刊,2024(4):127-129.
- [15] 崔梦冉,王延艳,常学东. 五味子沙棘复合饮料的研制及抗氧化活性评价[J]. 食品研究与开发,2025,46(2):144-153.
- [16] 任国秋. 熟化方式和花色苷强化对紫玉米粉固体饮料品质的影响[D]. 沈阳:沈阳农业大学,2024.
- [17] 许冰洁,徐娜,张杰. 蒲公英-猕猴桃复合饮料的研制[J]. 饮料工业,2025,28(4):62-68.
- [18] 胡晓婷,黄建芳,文明,等. 乳酸菌发酵百香果豆清饮料的动态品质分析:营养价值、抗氧化活性和感官风味[J]. 食品与发酵工业,2026,52(5):250-260.
- [19] 左亚锋,代永康,徐启兰,等. 山药杏仁粟米复合饮料制备及抗氧化活性[J]. 湘南学院学报,2023,44(2):22-27.
- [20] 陈又铭,李宁,袁卫涛,等. 低聚果糖的功能性质及其在食品中的应用[J]. 中国食品添加剂,2022,33(1):11-15.
- [21] 韦帮偶,覃丽媛,黄丽琼,等. 甘蔗渣木聚糖酶解液中低聚木糖的脱盐脱色工艺研究[J]. 轻工科技,2025,41(2):1-5.
- [22] Lin Conghao, Yuan Yuting, Tan Wenqiang, et al. Preparation of cationic chitooligos-accharide derivatives bearing N-halogenated benzyl pyridinium and assessment of their antimicrobial activities [J]. Polymer Bulletin, 2022, 80(11):12193-12210.

责任编辑:汪张贵

玻璃纤维粉目数对二氧化硅气凝胶性质影响的研究

高燕*, 王悦悦, 钱紫薇, 张丽娜, 张家嘉

(蚌埠学院 材料与化学工程学院, 安徽 蚌埠 233030)

摘要:以正硅酸四乙酯为硅源,采用溶胶-凝胶法,成功制备了不同目数和不同含量的玻璃纤维粉二氧化硅气凝胶复合材料。结果表明:玻璃纤维粉与二氧化硅气凝胶进行了复合,形成了疏水性良好的新复合材料;复合材料随着玻璃纤维粉含量的增加,其粒径及 BET 比表面积均呈现减小的趋势,与玻璃纤维粉目数的选择无关;相较于相同含量的 3000 目和 8000 目玻璃纤维粉,目数更小的 800 目玻璃纤维粉更易与二氧化硅气凝胶形成比表面积大且热稳定性好的复合材料。复合材料的高效合成为隔热、降噪、防火涂层等方向提供新的实践依据。

关键词:二氧化硅气凝胶复合材料;溶胶-凝胶法;正硅酸四乙酯;玻璃纤维粉

中图分类号:O613.72

文献标识码:A

文章编号:2095-297X(2026)05-0022-07

Effects of Glass Fiber Powder Mesh Size on the Properties of Silica Aerogel

GAO Yan*, WANG Yueyue, QIAN Ziwei, ZHANG Lina, ZHANG Jiajia

(School of Materials and Chemical Engineering, Bengbu University, Bengbu, 233030, Anhui)

Abstract: Silica aerogel composites with glass fiber powder of varying mesh sizes and contents were successfully prepared via the sol-gel method by using tetraethyl orthosilicate as the silicon source. The results indicated that the glass fiber powder was incorporated with the silica aerogel, forming new composites with good hydrophobicity. As the content of glass fiber powder increased, both the particle size and BET specific surface area of the composite showed a decreasing trend, irrespective of the selected mesh size of the glass fiber powder. Compared to composites with the same content of 3000-mesh and 8000-mesh glass fiber powder, the smaller 800-mesh glass fiber powder more readily formed composites with silica aerogel, exhibiting a larger specific surface area and improved thermal stability. The efficient synthesis of composites offers practical insights for thermal insulation, noise reduction, and fire-resistant coatings.

Key words: Silica aerogel composites; sol-gel method; tetraethyl orthosilicate; glass fiber powder

二氧化硅气凝胶是一类孔隙中充满气态分散介质的三维纳米材料,具有丰富的孔道结构(80% - 99.8%)、超大的比表面积($500 - 1200 \text{ m}^2 \cdot \text{g}^{-1}$)、极低的热导率($0.01 - 0.03 \text{ W} \cdot \text{m}^{-1} \cdot \text{K}^{-1}$)以及质轻、低密度($0.003 - 0.5 \text{ g} \cdot \text{cm}^{-3}$)等特点,在隔热、保温、航天航空、催化等领域有着广泛的应用^[1-4]。溶胶-凝胶法是二氧化硅气凝胶的常用制备方法,一般包括溶胶的生成、凝胶的制备、凝胶的老化以及凝胶的干燥^[5]。其中凝胶的老化和干燥是十分重

要的步骤,决定制备二氧化硅气凝胶的品质,是否具有完整的三维结构^[6]。表面改性剂溶液老化法以及常压干燥法有助于二氧化硅气凝胶形成完整的三维网络结构和节约成本。在常规的溶胶-凝胶法中引入价格低廉的无机表面增强剂玻璃纤维粉,玻璃纤维粉作为刚性骨架弥补了纯二氧化硅气凝胶结构脆弱、易坍塌、力学强度低的缺陷^[7-8]。

近年来二氧化硅气凝胶与纤维类形成复合材料得到广泛研究。崔连超以隔热性能稳定的碳纤维毡

收稿日期:2025-10-25

* 通讯联系人

基金项目:安徽省大学生创新创业训练项目(202411305015, S202411305002);安徽省质量工程教育研究项目(2024jyxm0485)。

作者简介:高燕(1987—),女,安徽宿州人,副教授,博士。E-mail:2857965284@qq.com

为基体材料,探究气凝胶粒径、填充质量分数对复合材料隔热性能的影响^[9]。邢士翠通过种子诱导结合溶剂热的方法制备二氧化钛纳米针纤维增强二氧化硅气凝胶复合材料,探究了溶剂热的反应温度和时间对复合材料形成的影响,并对制备复合材料的耐温性和隔热性能进行了研究^[10]。路林等制备了莫来石纤维与二氧化硅气凝胶复合材料,并考察了二氧化硅气凝胶及莫来石纤维的添加量对复合材料防火性能的影响^[11]。Zhou 等以甲基三甲氧基硅烷和水玻璃为共前驱体制备了二氧化硅气凝胶,采用冷冻干燥法制备了玻璃纤维增强二氧化硅气凝胶复合材料^[12]。各种研究表明采用有机纤维或者无机纤维有助于增强二氧化硅气凝胶的物理化学性能,提高其应用场景的多样化^[13]。

采用无机纤维增强法改善二氧化硅气凝胶性能,以气凝胶多孔结构中融入适当含量和目数的玻璃纤维作为支撑骨架,增加其三维结构稳定性^[14]。粒径约为 0.018 mm 的 800 目玻璃纤维粉机械强度高,粒径约为 0.005 mm 的 3000 目玻璃纤维粉具有一定的平衡强度与细腻度,粒径约为 0.0016 mm 的 8000 目玻璃纤维粉是化学活性强的超细粉体。将不同目数的玻璃纤维粉简单地溶于二氧化硅气凝胶溶胶中,充分混合进行凝胶化形成复合材料,采用接触角、激光粒径等多种表征方法研究不同含量和目数的玻璃纤维粉对二氧化硅气凝胶复合材料结构及物理、化学性质的影响。研究制备的比表面积大、热稳定性良好的玻璃纤维粉二氧化硅气凝胶复合材料,可通过涂料隔热用于外墙保温,显著降低建筑能耗,进一步拓展其在建筑隔热节能领域的潜在应用价值。

1 材料与方法

1.1 药品和试剂

玻璃纤维粉(800 目、3000 目和 8000 目),富华新材料有限公司;六甲基二硅氧烷(99%),正硅酸四乙酯(98%),上海麦克林生化科技有限公司;无水乙醇($\geq 99.7\%$),国药集团化学试剂有限公司;浓硫酸(95%–98%)、氨水(25.0%),天津市永大化学试剂有限公司;浓盐酸(36.0%–38.0%),江苏彤晟化学试剂有限公司;去离子水($\geq 182\text{ M}\Omega$),实验室纯净水系统。

1.2 实验仪器

电热鼓风恒温干燥箱(DNG-9076A),绍兴市苏珀仪器有限公司;分析天平(FA1004B),上海越平科学仪器制造有限公司;超声波清洗器(KQ-100),昆

山市超声仪器有限公司;集热式磁力搅拌器(KSW-40-15),东莞三量量具有限公司;激光粒度分析仪(Nicom380Z3000),美国 PSS 粒度仪公司;接触角测量仪(JCY 系列),上海方瑞仪器有限公司;X 射线衍射仪(Smart Lab SE),日本岛津公司;傅里叶变换红外光谱仪(Nicolet IS 10),美国赛默飞世尔有限公司;比表面及孔径分析仪(TriStarII3020),美国麦克默瑞提克仪器有限公司;热重分析仪(STA 2500),德国 ZEISS 公司。

1.3 前驱体的制备

前驱体 P900 的制备^[15]。具体制备过程:取正硅酸乙酯 19.7061 g(理论 19.6480 g)边搅拌边加入乙醇 3.5719 g(理论 3.5340 g)于烧杯中得溶液 A;取乙醇 3.6627 g(理论 3.5340 g)边搅拌边加入浓硫酸 0.0243 g(理论 0.0164 g)于烧杯中得溶液 B。将溶液 A 置于磁力搅拌器中进行室温搅拌,磁力搅拌器的转数控制在 350–450 r/min,在半小时内缓慢将溶液 B 滴入溶液 A 中,得到 A 和 B 的混合溶液 C,继续搅拌半小时使溶液充分混合。称取 3.0724 g 蒸馏水在两小时内缓慢滴加到溶液 C 中,蒸馏水的滴加速率要适当,以避免局部过量的水导致硅溶胶的过度水解。最终制备得到前驱体 P900 溶液于密封容器中保存,并置于阴凉处以防止前驱体污染。

1.4 玻璃纤维粉二氧化硅气凝胶复合材料的制备

取 4.0180 g 左右的 800 目玻璃纤维粉溶于 10 mL 无水乙醇中,将混合液超声振动震荡 10 min,充分混合后静置 5 min,得到玻璃纤维粉乙醇混合液。分别向 6 个烧杯中加入 0、0.500、1.000、1.500、2.000、2.500 mL 的玻璃纤维粉乙醇混合液,再依次加入 3.500、3.000、2.500、2.000、1.500、1.000 mL 的无水乙醇,然后向每个烧杯中加入 0.280 mL 蒸馏水和 3.000 mL P900 溶液,搅拌均匀后置于聚苯乙烯盒中,向聚苯乙烯盒迅速加入 3.500 mL 六甲基二硅氧烷和 0.250 mL $2\text{ mol}\cdot\text{L}^{-1}$ 的氨的乙醇溶液,玻璃棒快速搅拌,直到凝胶出现停止搅拌,将盖好盖子的聚苯乙烯盒置于 75 °C 的烘箱中 2 h,然后切成小块置于 150 °C 的烘箱中 1 h,取出得到二氧化硅气凝胶,以及不同 800 目玻璃纤维粉含量(3%,6%,9%,12%,15%)的二氧化硅气凝胶复合材料。不同含量的 3000 目和 8000 目玻璃纤维二氧化硅气凝胶复合材料的制备方法与上述方法类似。

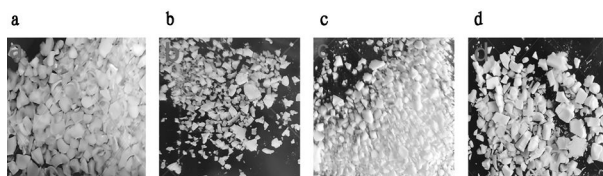
2 结果与分析

2.1 复合材料的疏水性分析

如图 1 所示,制备的二氧化硅气凝胶呈现类似

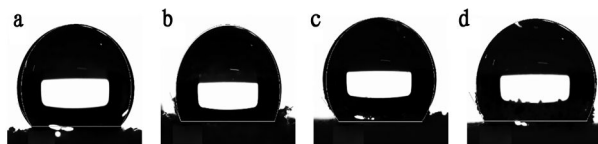
凝胶的半透明淡蓝色状态,与 3% 含量不同目数玻璃纤维粉复合后的二氧化硅气凝胶呈现不透明的均匀乳白色质地,形貌上看出复合前后二氧化硅气凝胶的状态发生明显改变。采用接触角测试来研究材料的疏水性,图 2 展示了二氧化硅气凝胶与 3% 含量不同目数玻璃纤维粉复合前后接触角的变化。此外,将不同含量和不同目数玻璃纤维粉制备二氧化硅气凝胶复合材料的接触角汇总于表 1。这里可以看出,不同含量和不同目数玻璃纤维粉制备二氧化

硅气凝胶复合材料均呈现较好的疏水性,接触角均大于 115° ,表明制备材料具有良好的疏水性。当玻璃纤维粉含量是 9% 时,对于 800 目、3000 目和 8000 目玻璃纤维粉二氧化硅气凝胶复合材料的接触角分别是 126.28° 、 122.25° 和 130.52° 。对于不同目数玻璃纤维粉复合的二氧化硅气凝胶均在玻璃纤维粉含量为 9% 时呈现较大的接触角。适宜的玻璃纤维粉添加量有助于增加二氧化硅气凝胶的疏水性。



(a 不含玻璃纤维粉, b、c 和 d 分别含 3% 的 800 目、3000 目和 8000 目玻璃纤维粉)

图 1 不同目数玻璃纤维粉制备复合材料实物图



(a 不含玻璃纤维粉, b、c 和 d 分别含 3% 的 800 目、3000 目和 8000 目玻璃纤维粉)

图 2 不同目数玻璃纤维粉制备复合材料接触角示意图

表 1 不同含量和不同目数玻璃纤维粉制备二氧化硅气凝胶复合材料的接触角汇总表

($^{\circ}$)

目数	0%	3%	6%	9%	12%	15%
800 目	125.85	116.09	126.87	126.28	119.31	119.12
3000 目	125.85	129.73	122.48	122.25	121.21	115.84
8000 目	125.85	120.65	126.72	130.52	124.55	124.17

2.2 复合材料的粒径分析

激光粒径仪分析测试用于了解二氧化硅气凝胶发生复合反应前后粒径的变化。图 3 是不同含量和不同目数玻璃纤维粉制备二氧化硅气凝胶复合材料的强度粒径分布图,将测试得到的体积粒径、强度粒径和数量粒径汇总于表 2。对于不同目数的玻璃纤

维粉,二氧化硅气凝胶复合材料的粒径会随着玻璃纤维粉含量的增加逐渐变小。以体积粒径为例,800 目玻璃纤维粉二氧化硅气凝胶复合材料随着玻璃纤维粉含量从 3% 增加到 15%,体积粒径的变化是 44.70、40.70、29.20、29.50 nm 和 24.20 nm,呈现依次减小的趋势。

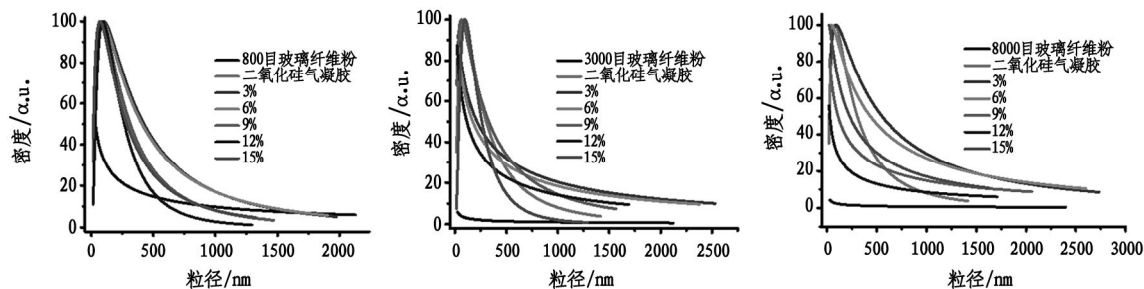


图 3 不同含量和不同目数玻璃纤维粉制备二氧化硅气凝胶复合材料的强度粒径分布图

表 2 不同含量和不同目数玻璃纤维粉制备二氧化硅气凝胶复合材料的粒径汇总表

项目	平均体积粒径分布/nm	平均强度粒径分布/nm	平均数量粒径分布/nm
800 目玻璃纤维粉	42.50	229.80	25.70
二氧化硅气凝胶	27.20	152.70	18.00
3% (800 目)	44.70	213.20	24.80
6% (800 目)	40.70	205.40	23.80
9% (800 目)	29.20	158.60	18.70
12% (800 目)	29.50	139.80	17.30
15% (800 目)	24.20	143.50	16.70
3000 目玻璃纤维粉	41.60	229.20	25.50
3% (3000 目)	65.60	273.40	30.70
6% (3000 目)	55.60	256.70	28.80
9% (3000 目)	28.30	169.00	19.40
12% (3000 目)	28.50	182.40	20.40
15% (3000 目)	31.40	139.40	17.60
8000 目玻璃纤维粉	54.60	258.60	28.80
3% (8000 目)	90.30	295.10	33.70
6% (8000 目)	73.70	280.60	31.80
9% (8000 目)	40.10	222.30	24.90
12% (8000 目)	28.60	184.60	20.60
15% (8000 目)	27.60	176.00	19.80

3000 目玻璃纤维粉二氧化硅气凝胶复合材料随着玻璃纤维粉含量的增加,体积粒径呈现依次减小的趋势。不同含量 8000 目玻璃纤维粉二氧化硅气凝胶复合材料,粒径依次减小,表明增加玻璃纤维粉在复合材料中的含量有助于得到粒径更小的二氧化硅气凝胶复合材料,说明在高性能二氧化硅气凝胶复合材料的制备过程中控制玻璃纤维粉的含量是十分重要的。

2.3 复合材料的 X 射线衍射分析

通过 X 衍射技术对二氧化硅气凝胶复合材料进行分析,对复合材料是否是晶体结构进行判断。

图 4 可以看出不同目数的玻璃纤维粉与二氧化硅气凝胶进行了复合,800 目玻璃纤维粉和无掺杂的二氧化硅气凝胶均为无定型结构,不同含量的 800 目玻璃纤维粉复合后的二氧化硅气凝胶仍呈现无定型结构。而 3000 目和 8000 目玻璃纤维粉本身存在晶体结构,与无定型的二氧化硅气凝胶复合后得到的复合材料仍然存在与 3000 目和 8000 目玻璃纤维粉相近的晶体结构,但是强度要弱。X 射线衍射分析表明玻璃纤维粉与二氧化硅气凝胶进行了充分的复合,形成了新的复合材料。

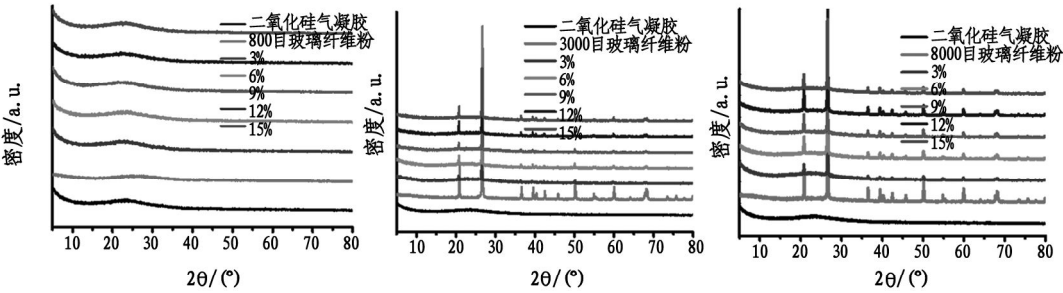


图 4 不同含量和不同目数玻璃纤维粉制备二氧化硅气凝胶复合材料的 X 射线衍射图

2.4 复合材料的红外光谱分析

在红外光谱分析中,一般将 3300 cm^{-1} – 3900 cm^{-1} 处的吸收峰归属于水中 O-H 的伸缩振动峰, 1600 cm^{-1} – 1700 cm^{-1} 处的吸收峰归属于水中 O-H 的弯曲振动峰,图 5 展示出二氧化硅气凝胶及不同含量和不同目数玻璃纤维粉复合得到的二氧化硅气凝胶在这两处具有较弱的吸收峰,说明制备的二氧化硅气凝胶及其复合物具有较少的-OH 基团存在,表现出显著的疏水性,与前面接触角的测试结果一致。此外,从图中可以看出无掺杂的二氧化硅气凝胶在 1080 、 843 、 452 cm^{-1} 处都有明显的吸收峰,这些峰为二氧化硅气凝胶的典型吸收峰,其中 1080 cm^{-1} 可归属为 Si-O-Si 键的非对称伸缩振动峰; 452 cm^{-1} 可归属为 Si-O-Si 键的弯曲振动峰; 843 cm^{-1} 的吸收峰是 Si-C 伸缩振动峰^[16];无其他明显杂质峰,

说明以 Si-O-Si 键为主要构成的二氧化硅气凝胶及其玻璃纤维粉复合物的形成。对于 800 目玻璃纤维粉本来的特征吸收峰 701 cm^{-1} 在复合材料中不再显示;对于 800 目玻璃纤维粉 2364 cm^{-1} 的特征吸收峰,当复合材料中玻璃纤维粉含量较高时($\geq 9\%$),才会显示出来特征吸收峰。对于 3000 目玻璃纤维粉 695 cm^{-1} 的特征吸收峰在复合材料中不再显示;3000 目玻璃纤维粉 797 cm^{-1} 的特征吸收峰在复合材料中可以清晰地显示出来。对于 8000 目玻璃纤维粉 702 cm^{-1} 的特征吸收峰在复合材料中不显示;3000 目玻璃纤维粉 805 cm^{-1} 的特征吸收峰在复合材料中可以清晰地显示出来。这些均说明不同目数的玻璃纤维粉与二氧化硅气凝胶进行了复合,形成了新的复合材料。

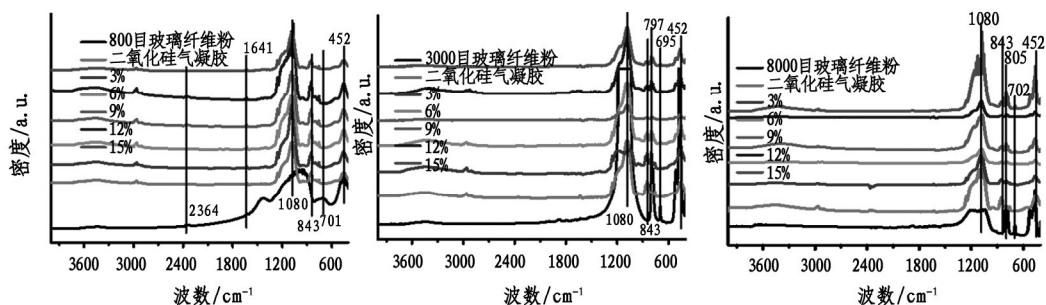


图 5 不同含量和不同目数玻璃纤维粉制备二氧化硅气凝胶复合材料的红外光谱图

2.5 复合材料的热重分析

如图 6 所示,利用热重分析来了解二氧化硅气凝胶复合材料的热稳定性。800 目和 3000 目玻璃纤维热稳定性好,在室温到 $850\text{ }^{\circ}\text{C}$ 的温度范围内几乎没有热量损失。800 目玻璃纤维复合得到的二氧化硅气凝胶在 800 目玻璃纤维含量在 12% 和 15% 的热稳定性大于未复合前的二氧化硅气凝胶,而当 800 目玻璃纤维含量为 9% 时,复合材料的热稳定性低于二氧化硅气凝胶本身。3000 目玻璃纤维复合得到的二氧化硅气凝胶在 3000 目玻璃纤维含量在

9% 时具有较好的热稳定性,在温度 $420\text{ }^{\circ}\text{C}$ 大于未复合前的二氧化硅气凝胶,而当 3000 目玻璃纤维含量为 12% 和 15% 时,复合材料的热稳定性均低于二氧化硅气凝胶本身。8000 目玻璃纤维具有相对较差的热稳定性,在与二氧化硅气凝胶进行复合得到的复合材料的热稳定性均低于二氧化硅气凝胶本身的热稳定性。研究表明玻璃纤维的目数和含量对复合材料的热稳定性产生较大的影响,需要严格控制玻璃纤维的目数和含量,以得到热稳定性更好的复合材料。

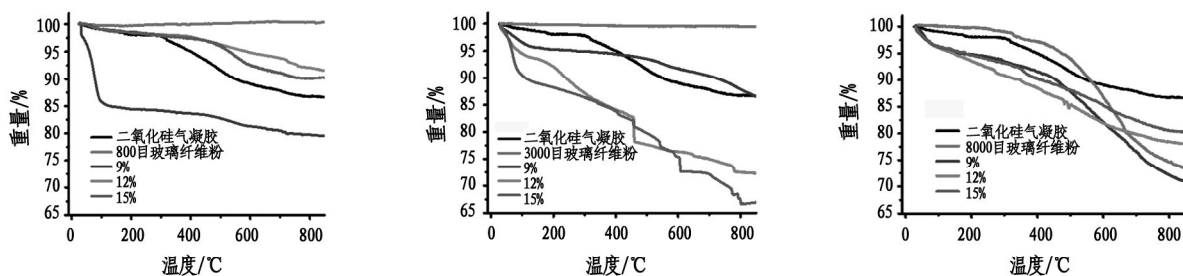
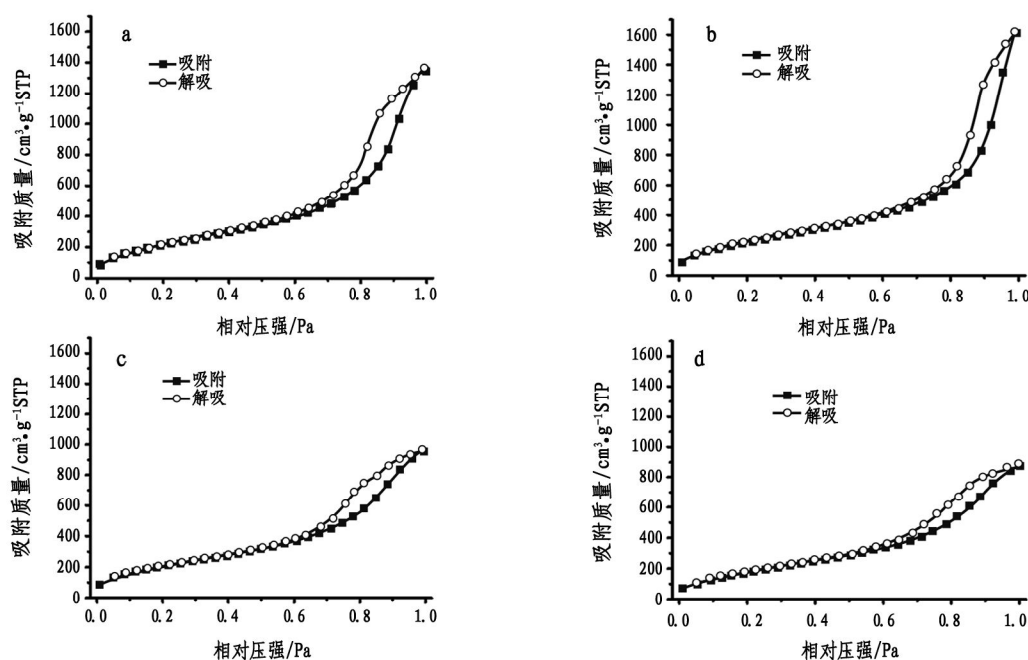


图 6 不同含量和不同目数玻璃纤维粉制备二氧化硅气凝胶复合材料的热重分析曲线

2.6 复合材料的比表面积和孔径分析

二氧化硅气凝胶复合材料具有较大的比表面积和丰富的孔道结构,如图7所示,对玻璃纤维粉二氧化硅气凝胶复合材料(玻璃纤维粉含量为3%)进行了氮气等温吸附脱附曲线分析。表3汇总了从氮气等温吸附脱附曲线得到的不同含量和不同目数玻璃纤维粉制备二氧化硅气凝胶复合材料的 BET 比表

面积和 BET 平均吸附孔径。含量为3%的800目玻璃纤维粉形成的二氧化硅气凝胶复合材料的 BET 比表面积($872.90 \text{ m}^2 \cdot \text{g}^{-1}$)大于含量为3%的3000目玻璃纤维粉形成的二氧化硅气凝胶复合材料的 BET 比表面积($774.37 \text{ m}^2 \cdot \text{g}^{-1}$)、含量为3%的8000目玻璃纤维粉二氧化硅气凝胶复合材料的 BET 比表面积($707.65 \text{ m}^2 \cdot \text{g}^{-1}$)以及未掺杂。



(a: 不含玻璃纤维粉; b: 含800目玻璃纤维粉; c: 含3000目玻璃纤维粉; d: 含8000目玻璃纤维粉)

图7 不同目数玻璃纤维粉制备二氧化硅气凝胶复合材料的氮气吸附-脱附等温线

表3 不同含量和不同目数玻璃纤维粉制备二氧化硅气凝胶复合材料的 BET 比表面积及吸附、脱附平均孔径汇总表

项目	BET 比表面积/ $\text{m}^2 \cdot \text{g}^{-1}$	BJH 吸附平均孔径/nm	BJH 脱附平均孔径/nm
800 目玻璃纤维粉	0.03	— — — —	— — — —
二氧化硅气凝胶	846.81	5.67	7.04
3% (800 目)	872.90	5.59	7.81
6% (800 目)	863.21	5.26	7.44
9% (800 目)	845.50	5.55	7.34
12% (800 目)	769.35	5.41	7.70
15% (800 目)	754.56	5.40	7.19
3000 目玻璃纤维粉	3.89	6.83	5.29
3% (3000 目)	774.37	5.50	5.91
6% (3000 目)	667.22	5.23	5.14
9% (3000 目)	631.22	5.01	5.24
12% (3000 目)	507.71	4.72	5.53
15% (3000 目)	421.88	4.73	5.21
8000 目玻璃纤维粉	4.95	5.57	4.78

续表

表 3 不同含量和不同目数玻璃纤维粉制备二氧化硅气凝胶复合材料的 BET 比表面积及吸附、脱附平均孔径汇总表

项目	BET 比表面积/ $\text{m}^2 \cdot \text{g}^{-1}$	BJH 吸附平均孔径/nm	BJH 脱附平均孔径/nm
3% (8000 目)	707.65	5.35	5.89
6% (8000 目)	590.48	5.49	6.96
9% (8000 目)	514.87	5.22	5.40
12% (8000 目)	449.08	5.24	6.03
15% (8000 目)	400.96	5.32	5.83

同样的对于其他含量的玻璃纤维粉二氧化硅气凝胶复合材料也存在类似规律。也就是当玻璃纤维粉的含量相同时,二氧化硅气凝胶复合材料随着玻璃纤维粉目数的增加呈现 BET 表面积逐渐减小的趋势。不同含量和不同目数玻璃纤维粉二氧化硅气凝胶复合材料的 BJH 吸附、解析平均吸附孔径均小于 10 nm,呈现介孔结构。结果表明:含 3% 800 目玻璃纤维粉的二氧化硅气凝胶复合材料具有最大的 BET 比表面积,可以通过控制玻璃纤维粉的含量来得到不同比表面积的二氧化硅气凝胶复合材料。

3 结论

二氧化硅气凝胶复合材料因其独特的物理和化学性质在多个领域有着广泛的应用前景。通过不同目数和不同含量的玻璃纤维粉成功制备了一系列具有良好疏水性及热稳定性的二氧化硅气凝胶复合材料,研究表明:它们的比表面积及粒径均随玻璃纤维粉含量及目数的变化而呈现规律性变化。研究为玻璃纤维粉二氧化硅气凝胶复合材料的高效制备提供了参考,并拓展了其在建筑隔热领域的应用空间。

参考文献:

[1] Li Hao,Xu Weidong,Zhu Liyan,et al. Study on thermal insulation performance of silica aerogel thermal insulation blankets[J]. Gels,2024,10(11):707.

[2] 刘杰,明贺娜,朱平,等. 疏水 SiO_2 气凝胶涂层涤纶织物的制备及隔热性能[J]. 印染,2023,49(12):63-66.

[3] Xi Shuang,Wang Yijun,Zhang Xiaoxue,et al. Fire-resistant polyimidesilica aerogel composite aerogels with low shrinkage,low density and high hydrophobicity for aerospace applications [J]. Polymer Testing, 2023, 129: 108259.

[4] 徐婉婷. SiO_2 气凝胶的增强及其在建筑材料领域的应用研究概况[J]. 建材技术与应用,2024,5:33-39.

[5] Soleimani D A, Abbasi M H. Silica aerogel: synthesis, properties and characterization [J]. Journal of Materials

Processing Technology,2008,199(1-3):10-26.

[6] Brinker C,Scherer G W. Sol-Gel science:the physics and chemistry of Sol-Gel processing[M]. New York:Academic Press,1990.

[7] 明贺娜,赵鹏,张凯迪,等. 高强度二氧化硅气凝胶的研究进展[J]. 棉纺织技术,2026,54(3):85-90.

[8] 李思盈. 海泡石纳米纤维增强二氧化硅气凝胶的制备及性能研究[D]. 重庆:重庆交通大学,2023.

[9] 崔连超. 二氧化硅气凝胶/碳纤维毡复合材料制备及隔热性能研究[D]. 哈尔滨:哈尔滨工程大学,2024.

[10] 邢士翠. 纤维增强二氧化硅气凝胶复合材料的设计合成及性能研究[D]. 济南:山东大学,2023.

[11] 路林,张增辉,贺优优. 莫来石纤维与二氧化硅气凝胶对防火涂料性能的影响[J]. 电镀与涂饰,2025,44(1):30-36.

[12] Zhou Ting, Cheng Xudong, Pan Yuelei, et al. Mechanical performance and thermal stability of glass fiber reinforced silica aerogel composites based on co-precursor method by freeze drying[J]. Applied Surface Science,2018,437:321-328.

[13] Berthon F S,Hildenbrand C,Ilbizian P,et al. Evaluation of lightweight and flexible insulating aerogel blankets based on Resorcinol-Formaldehyde-Silica for space applications [J]. European Polymer Journal,2017,93:403-416.

[14] Bardy E R,Mollendorf J C,Pendergast D R. Thermal conductivity and compressive strain of aerogel insulation blankets under applied hydrostatic pressure[J]. Journal of Heat Transfer-Transactions of the Asme, 2007, 129 (2): 232-235.

[15] Pajonk G M,Elaloui E,Achard P,et al. Physical properties of silica gels and aerogels prepared with new polymeric precursors[J]. Journal of Non-Crystalline Solids,1995,186:1-8.

[16] Rami A O,Houssam E R. Synthesis and characterization by FTIR spectroscopy of silica aerogels prepared using several $\text{Si}(\text{OR})_4$ and $\text{R}^n\text{Si}(\text{OR}')_3$ precursors[J]. Journal of Molecular Structure,2009,919:140-145. 责任编辑:李云

沸石咪唑酯骨架材料作为缓蚀剂在海洋环境铜防护中的应用

曹娇娇*, 柯顺安, 葛雨童

(蚌埠学院 材料与化学工程学院, 安徽 蚌埠 233030)

摘要:探讨了一种新型沸石咪唑酯骨架材料(ZIFs)在中性介质中对铜的缓蚀作用机制。通过扫描电子显微镜、X-单晶衍射及红外光谱表征,证实所合成材料具有高度规整的晶体结构。电化学测试结果表明,该材料在适宜浓度下对铜的缓蚀效率可达90%以上。机理分析表明,其缓蚀作用主要归因于有机配体在铜表面形成的吸附保护层,能有效阻隔腐蚀介质。研究证实了ZIFs作为高效铜缓蚀剂的可行性,并为功能导向设计金属有机框架材料用于腐蚀防护提供了新思路。

关键词:ZIFs;铜;电化学;缓蚀剂;功能导向设计

中图分类号:TG174.42;TB34

文献标识码:A

文章编号:2095-297X(2026)05-0029-05

Research on Metal-Organic Framework Materials as Corrosion Inhibitors in the Protection of Copper in Marine Environment

CAO Jiaojiao*, KE Shunan, GE Yutong

(School of Materials and Chemical Engineering, Bengbu University, Bengbu, 233030, Anhui)

Abstract: It systematically investigated the inhibition mechanism of a newly synthesized zeolite imidazolate framework (ZIF) material on copper in a neutral medium. The synthesized material was confirmed to possess a highly ordered crystalline structure through characterization by scanning electron microscopy, X-ray single-crystal diffraction, and infrared spectroscopy. Electrochemical test results indicated that the inhibition efficiency of this material on copper could exceed 90% at an optimal concentration. Mechanistic analysis revealed that the corrosion inhibition was primarily attributed to an adsorbed protective layer formed by the organic ligands on the copper surface, which effectively blocks corrosive media. It confirmed the feasibility of ZIFs as highly efficient corrosion inhibitors for copper and provided a new research direction for the functional design of metal-organic framework materials for corrosion protection.

Key words: zeolite imidazolate frameworks; copper; electrochemistry; corrosion inhibitor; function-oriented design

铜合金管道因其卓越的导热性、机械加工性与抗生物污损能力,被广泛应用于船舶制造、滨海电站、海水淡化系统以及海洋平台的各种换热器与管道网络中。然而,严酷的海洋环境对铜金属构成了极其严重的腐蚀威胁^[1-3]。氯离子会破坏铜表面自然形成的保护性氧化膜,引发点蚀、缝隙腐蚀等局部破坏,显著缩短设备服役寿命。据统计,中国每年因

各类腐蚀问题导致的经济损失高达3至5万亿元人民币,约占国民生产总值的3%至5%^[4-5]。针对金属腐蚀的防护策略主要涵盖三大类:表面涂层技术、电化学保护以及缓蚀剂应用^[6-8]。相较而言,缓蚀剂技术因其添加量小、成本效益高、操作简便、适应性广且对基体金属几乎无附加损伤等突出优点,被公认为是减缓铜腐蚀速率最为经济有效的技术手段

收稿日期:2025-12-15 * 通讯联系人

基金项目:蚌埠学院应用项目(2024YYX26QD);安徽省大创项目(S202511305145)。

作者简介:曹娇娇(1993—),女,安徽阜阳人,讲师,博士。E-mail:caojiaojiao@bbc.edu.cn

之一^[9-11]。传统的铜缓蚀剂以苯并三唑(BTA)及其衍生物为代表,它们能在铜表面形成配位聚合物膜,有效抑制腐蚀。近年来,金属有机框架(MOFs)材料作为一种新兴的缓蚀剂也展现出巨大潜力^[12-14]。其中,沸石咪唑酯骨架材料(ZIFs)是其中一类 MOFs。传统缓蚀剂,其合成与应用过程存在明显的环境隐患,它们通常含有或在合成中使用有毒、难生物降解的组分(如磷酸盐、铬酸盐及某些有机溶剂),这些物质在海洋环境中持续累积,将对水生生态系统和人类健康构成长期威胁,这与当前全球倡导的绿色、可持续发展理念背道而驰。在此背景下,兼具纳米尺寸效应、结构可设计性、高化学稳定性及良好生物相容性的新型 MOFs 材料,迅速成为环境友好型缓蚀剂研发的前沿焦点^[15-16]。MOFs 材料具有巨大的比表面积和高度有序的孔道结构,这赋予了其优异的物理吸附能力;更重要的是,通过精准选择有机配体和金属节点,能设计出具有特定防护功能的“智能”缓蚀剂。基于此,以锌离子为金属节点,咪唑衍生物为有机配体,合成了一种沸石咪唑酯骨架材料。性能表征表明,该材料是一种高效的铜腐蚀抑制剂。

1 材料与方法

1.1 试验材料

实验仪器:水热釜、超声波清洗机、分析天平、滴定管、移液管、烧杯、扫描电子显微镜(SEM)、傅里叶变换红外光谱仪(FTIR)、离心机。

化学试剂:六水合硝酸锌(分析纯)、4-硝基咪唑(分析纯)、甲醇(分析纯)、无水乙醇。

1.2 样品微观形貌观测

通过以下实验方法对合成材料进行表征,并系统评估其缓蚀性能。

扫描电子显微镜(SEM)方法:在 15 kV 的加速电压下,对经喷金处理的样品进行形貌观测。该分析旨在清晰揭示 MOF 晶体的微观形貌、表面纹理、颗粒尺寸分布及其均匀性,为理解其结构与性能的关系提供直观的形貌学依据。

X 射线衍射(XRD)方法:在 $0^{\circ} - 60^{\circ}$ 的扫描范围内,使用 Cu K α 辐射($\lambda = 1.5418 \text{ \AA}$)对粉末样品进行物相分析。将测得的衍射谱图与依据单晶结构模拟的理论谱图进行比对,以确定所合成材料的晶体结构、相纯度与结晶度,此为 MOF 材料成功合成的决定性证据。

傅里叶变换红外光谱(FTIR)方法:采用 KBr 压片法,在 $4000 - 400 \text{ cm}^{-1}$ 波长范围内对样品进行扫

描。通过分析特征吸收峰的位移与强弱,用以确认 MOF 框架中关键官能团的存在、化学键类型,进而验证其化学结构是否与预期设计相符,为材料的成功合成提供关键证据。

电化学性能测试:采用标准三电极体系,在 3.5 wt% 氯化钠(NaCl)溶液中进行。通过测量塔菲尔(Tafel)极化曲线,定量评估 MOF 缓蚀剂的性能。Tafel 曲线用于计算腐蚀电流密度与缓蚀效率,揭示其缓蚀机理。

1.3 试验方法

首先,将六水合硝酸锌与 4-硝基咪唑以 1:1 的摩尔比溶于甲醇中,并在 25°C 下恒温超声处理 10 分钟,使其充分混合。随后,将混合溶液转移至聚四氟乙烯内衬的不锈钢水热釜中,于 120°C 的烘箱内反应 24 小时。反应结束后,待反应釜自然冷却至室温,将所得产物进行离心分离,并用去离子水和乙醇洗涤数次,最后在 60°C 的真空干燥箱中干燥,所得固体粉末即为沸石咪唑酯骨架材料,即 N-MOF,保存备用。

2 结果与分析

2.1 沸石咪唑酯骨架材料 N-MOF 的形貌表征

通过对合成产物进行 SEM、XRD 与 FT-IR 表征,其结果证实我们成功制备出了结构规整、化学组成符合预期的目标材料,如图 1 所示。

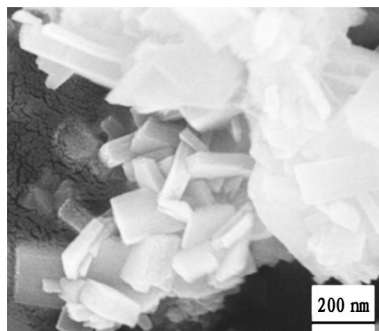


图1 N-MOF 的扫描电镜形貌图

如图 1 所示,该材料的微观形貌呈现高度均一的长方体棒状结构,其形状规整,表面光滑,尺寸分布集中。进一步的统计测量表明,该长方体的截面边长约为 100 nm ,长度主要分布在 $300 - 400 \text{ nm}$ 之间。这种均一的纳米级结构不仅证明了合成路线的有效性,也为其作为缓蚀剂在金属表面形成致密、均匀的吸附层奠定了形貌基础。

对所得材料进行的 X 射线衍射(XRD)分析结果如图 2 所示。

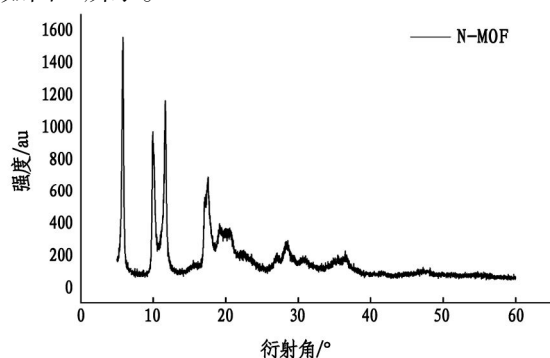


图 2 N-MOF 的 XRD 图

由图 2 可见,样品在 2θ 为 5° 、 10° 、 11° 、 17° 、 22° 及 27° 等处出现了一系列尖锐且清晰的衍射峰。该衍射图谱与文献中报道的 ZIF-8 实验 XRD 谱图高度吻合^[17-18]。各特征峰位一一对应,这表明所合成的材料具有与目标晶体结构一致的有序性。特别值得注意的是,位于低角度区域(如 5° 、 10°)的强衍射峰,这是具有孔径结构的 MOF 材料的典型特征,证明了其多孔框架的成功构建。此外,所有衍射峰均无杂峰出现,表明合成产物具有高的结晶度和相纯度。综上所述,XRD 结果有力地证实了我们成功合成了具有预期晶体结构的高纯度 MOF 材料。

此外,通过图 3 的红外光谱分析,可以进一步确认材料的化学键合与官能团信息。

在 421 cm^{-1} 附近出现的明显特征吸收峰,可归属为锌离子与咪唑氮原子之间配位键(Zn-N)的伸缩振动,这是形成 ZIFs 类金属有机框架结构的关键证据^[19-21]。同时,在 $1300\text{—}1550\text{ cm}^{-1}$ 的波数范围内出现的一组特征峰,分别对应于硝基($-\text{NO}_2$)的不对称伸缩振动和对称伸缩振动^[22],这清晰地表明来

自 4-硝基咪唑配体的硝基官能团被成功引入到最终的 N-MOF 骨架中。

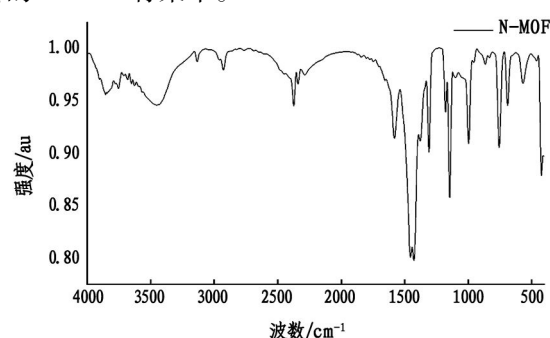


图 3 N-MOF 的红外谱图

2.2 电化学实验结果

为深入探究 N-MOF 缓蚀剂对铜电极阳极与阴极反应过程的抑制机制,塔菲尔(Tafel)极化曲线被用于进行分析。不同浓度 N-MOF 溶液中的 Tafel 曲线如图 4 所示。曲线的塔菲尔线性区通过经典外推法进行拟合,腐蚀电流密度(i_{corr})等关键电化学参数得以准确估算,相关结果汇总于表 1 中。

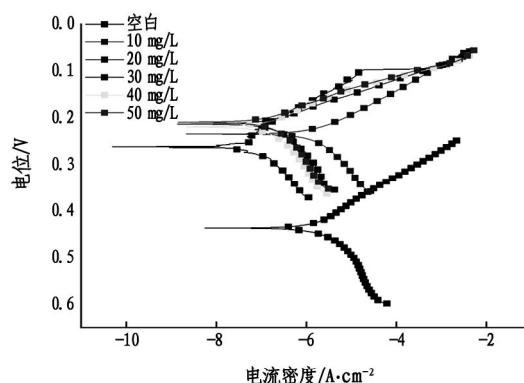


图 4 不同浓度的 N-MOF 对铜的缓蚀作用极化曲线

表 1 空白组和不同 N-MOF 组的极化参数

Samples/ $\text{mg} \cdot \text{L}^{-1}$	$b_a/\text{mV} \cdot \text{dec}^{-1}$	$b_c/\text{mV} \cdot \text{dec}^{-1}$	$i_{\text{corr}}/\text{A} \cdot \text{cm}^{-2}$	$\eta/\%$
空白	59.00	-117.20	4.00×10^{-6}	/
10	53.70	-111.30	2.77×10^{-6}	30.75
20	37.98	-116.20	4.97×10^{-7}	87.58
30	40.51	-98.10	2.54×10^{-7}	93.65
40	43.30	-114.00	4.21×10^{-7}	89.48
50	37.70	-115.60	4.65×10^{-7}	88.38

由数据可知,空白溶液中铜的腐蚀电流密度最高,这主要归因于氯离子(Cl^-)对铜表面的直接侵

蚀,促使 Cu 被氧化为 Cu^+ 。当腐蚀介质中加入 N-MOF 缓蚀剂后,腐蚀电流密度显著下降。例如,在

$10 \text{ mg} \cdot \text{L}^{-1}$ 的 N-MOF 溶液中, i_{corr} 值为 $2.77 \mu\text{A} \cdot \text{cm}^{-2}$; 而当浓度增至 $30 \text{ mg} \cdot \text{L}^{-1}$ 时, i_{corr} 进一步降低至 $0.254 \mu\text{A} \cdot \text{cm}^{-2}$ 。这一变化趋势明确表明, 随着 N-MOF 浓度的增加, 其对铜电极的腐蚀抑制作用持续增强^[23]。

计算得到的缓蚀效率(η)变化与此吻合: 从 $10 \text{ mg} \cdot \text{L}^{-1}$ 时的 30.75%, 显著提升至 $30 \text{ mg} \cdot \text{L}^{-1}$ 时的 93.65%, 达到最大值。这表明在铜表面形成的 N-MOF 吸附膜能有效阻隔腐蚀介质, 同时抑制阴极还原和阳极溶解反应, 因此该 N-MOF 属于混合型缓蚀剂。值得注意的是, 当浓度继续升高至 $40 \text{ mg} \cdot \text{L}^{-1}$ 和 $50 \text{ mg} \cdot \text{L}^{-1}$ 时, 缓蚀效率出现小幅回落, 这可能源于过高浓度下分子间相互作用导致的吸附层局部脱附。此外, 对比各曲线的阴极分支可见, 它们基本呈现平行关系, 说明 N-MOF 对阳极反应的抑制程度总体大于对阴极反应的影响^[24-25]。

2.3 SEM 结果分析

通过腐蚀浸泡实验, 我们对比观察了铜在含及不含 N-MOF 缓蚀剂的 3.5 wt% NaCl 溶液中浸泡 24 小时后的表面形貌, 结果如图 5 所示。在未添加 N-MOF 的空白溶液中, 铜电极表面呈现出典型的氯离子诱导腐蚀特征。其表面被大量疏松、不均匀的腐蚀产物所覆盖, 这主要是由于溶液中的 Cl^- 与溶解氧协同作用, 持续侵蚀铜基体并促进了 Cu_2O 等产物的生成与堆积^[26-27]。相比之下, 在添加了 N-MOF 缓蚀剂的实验组中, 电极表面变得平整光滑, 仅出现零星且轻微的腐蚀迹象。这一形貌改善直观地证明, N-MOF 分子能够有效地在铜/溶液界面发生吸附, 形成一层致密且稳定的保护膜, 从而物理性阻隔腐蚀介质的侵袭。该宏观形貌观察结果与前述电化学测试中腐蚀电流密度显著降低、缓蚀效率显著提升的趋势高度吻合, 从不同维度共同证实了所合成 N-MOF 材料对铜在氯化钠介质中优异的缓蚀性能。

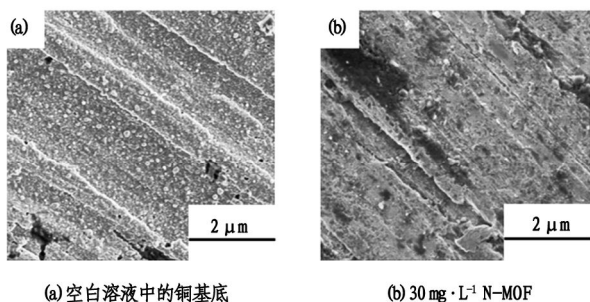


图 5 不同腐蚀环境中铜表面形貌

2.4 机理分析

表面分析与腐蚀电化学测试结果表明, 将该 N-MOF 材料添加至模拟海洋环境的中性氯化钠溶液中, 能对铜合金起到极其显著的缓蚀作用。其卓越的性能源于一种协同防护机制。首先, N-MOF 纳米颗粒凭借其高表面能, 能够物理吸附并致密地覆盖在铜合金表面, 形成一道初始物理屏障。其次, 更为关键的是, 4-硝基咪唑配体中的氮、氧杂原子能够与铜表面发生强烈的化学配位, 同时 N-MOF 结构中开放的锌位点也可能与铜基底相互作用, 从而共同构建起一层牢固、稳定的杂化保护膜。这层膜能有效阻隔腐蚀性氯离子、水和氧气的渗透, 同时抑制铜的阳极溶解和氧的阴极还原这两个电化学过程。

3 结论

在水热合成条件下, 研究成功制备出一种具有规整长方体棒状结构的 ZIF 衍生材料。

电化学测试结果表明, 当该 N-MOF 作为缓蚀剂在 3.5 wt% NaCl 溶液中的浓度达到 $30 \text{ mg} \cdot \text{L}^{-1}$ 时, 对铜金属可实现高达 93% 以上的缓蚀效率, 展现出卓越的防护性能。

研究不仅为应对海洋环境中的铜腐蚀问题提供了一种高效、环境友好的新材料解决方案, 更重要的是, 它完整展示了从分子结构设计、绿色合成实践到微观作用机制剖析的系统研究路径。

参考文献:

- [1] 刘群, 魏宏, 许冲, 等. 低碱条件下组合抑制剂在铜硫分离中的试验研究[J]. 铜业工程, 2025(5): 97-103.
- [2] 郑富凯, 韩光达, 张思雨, 等. 铜合金在增材制造领域的研究进展[J]. 铜业工程, 2025(2): 110-129.
- [3] ZHANG H, BIAN H, ZHANG X, et al. Preparation and characterization of hollow ceria based smart anti-corrosive coatings on copper[J]. Surfaces and Interfaces, 2024, 45: 103930.
- [4] 陈磊. 典型三元电偶体系腐蚀行为研究及有限元模拟[D]. 合肥: 中国科学技术大学, 2025.
- [5] Devadig R, Handa V, Jkalmithu J, et al. Gadhamshetty, indirect impact of quorum sensing on corrosion of graphene-coated copper via sulfate-reducing bacteria biofilm enhancement[J]. Corrosion Science, 2024: 112988.
- [6] Mzioud K, Habsaoui A, Rached S, et al. Experimental investigation and theoretical modeling of copper corrosion inhibition by urguinea maritima essential oil[J]. Materials Today Sustainability, 2024, 27: 100906.
- [7] Qiao Yanxin, Wang Tianyu, Chen Zhilin, et al. Corrosion techniques and strategies for used fuel containers with cop-

- per corrosion barriers under deep geological disposal conditions; a review[J]. *International Journal of Minerals, Metallurgy and Materials*, 2024, 31: 2582-2606.
- [8] She Xiaomeng, Jian Peng, Qiang Yujie, et al. Recent advances in protective technologies against copper corrosion[J]. *Journal of Materials Science & Technology*, 2024, 201: 75-94.
- [9] Li Yuanxing, Chen Yiwen, Wang Caoxu, et al. Exploring the potential of plant extracts as corrosion inhibitors; a comprehensive review[J]. *Progress in Organic Coatings*, 2025, 198: 108915.
- [10] Puzikova D, Khussurova G, Leontyeva X, et al. Review of organic corrosion inhibitors; application with respect to the main functional group[J]. *Journal of Saudi Chemical Society*, 2025, 29: 20.
- [11] Verma C, Barsoum I, Rhee K, et al. Physicochemical bonding and corrosion protection; non-covalent and coordination bonds in transport and adsorption of corrosion inhibitors[J]. *Coordination Chemistry Reviews*, 2026, 508: 217154.
- [12] Chen Chao, Yu Mei, Xiao Jian, et al. One-step in-situ growth by one-pot method for uniform BTA@ ZIF-8 layer to improve the corrosion protection of Sol-Gel coatings[J]. *Corrosion Science*, 2025, 250: 112887.
- [13] Guo Chengbin, Lian Yubo, Huang Chen, et al. Sustained-release system based on BTA @ MOF-5 for self-healing coating application[J]. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 2022, 654: 130139.
- [14] Zhao Chenyang, Huang Yi, Li Yue, et al. Construction of BTA-inbuilt ZIF-8 coated with molybdate-intercalated ZnAl-LDH as Core-Shell structured nanocontainers toward Ph-responsive smart corrosion protection of waterborne epoxy coating[J]. *Progress in Organic Coatings*, 2024, 194: 108623.
- [15] 傅伟强, 马心语, 杨家乐, 等. 金属有机框架材料在防腐涂层中的研究进展[J]. *聚酯工业*, 2025, 38(2): 77-80.
- [16] 谢峰, 岳浩, 赵蕊, 等. 电沉积时间对铝合金表面锰基 MOF 膜层的电化学性能影响[J]. *中国表面工程*, 2026, 39(3): 239-248.
- [17] Alipanah N, Dehghani A, Abdolmaleki M, et al. Designing environmentally-friendly Ph-responsive self-redox polyaniline grafted graphene oxide nano-platform decorated by zeolite imidazole ZIF-9 MOF for achieving smart functional epoxy-based anti-corrosion coating[J]. *Journal of Environmental Chemical Engineering*, 2023, 11: 109048.
- [18] Chen Junnan, Jinhua Feng, Liang Chengke, et al. Application of ZIF nanosheets for construction of epoxy composite coatings with enhanced corrosion protection performance[J]. *Progress in Organic Coatings*, 2023, 175: 107335.
- [19] Anadebe V C, Chukwuike V L, Chidiebere M A, et al. Synthesis, characterization, and evaluation of Co-MOF based ZIF-67 for CO₂ corrosion inhibition of X65 steel: insights from electrochemical studies and a machine learning algorithm[J]. *The Journal of Physical Chemistry C*, 2023, 127: 9871-9886.
- [20] Zhang Dan, Xiao Linming, Xiong Gaodian, et al. Recent progress of zeolitic imidazolate frameworks (ZIFs) in superhydrophobic and anticorrosive coatings for metals and their alloys[J]. *Journal of Coatings Technology and Research*, 2023, 20: 1157-1177.
- [21] Zhang Xuejun, Zhang Yang, Wang Yanling, et al. Preparation and corrosion resistance of hydrophobic zeolitic imidazolate framework (ZIF-90) film@ Zn-Al alloy in NaCl solution[J]. *Progress in Organic Coatings*, 2018, 115: 94-99.
- [22] Li Jizhen, Fan Xuezhong, Liu Xiaogang, et al. Thermal decomposition of 4-nitroimidazole catalyzed by Pb(NO₃)₂[J]. *Chinese Journal of Chemistry*, 2008, 26: 127-129.
- [23] 李瑾. 工业水处理中的铜材料点蚀组装保护技术研究[D]. 上海: 东华大学, 2020.
- [24] Solmaz R, Şahin E A, Kardas G L, et al. Bingo? 1 pollen self-assembled natural thin films: fabrication, characterization, and corrosion inhibition performance for copper protection in a NaCl environment[J]. *ACS omega*, 2025, 22: 23281-23296.
- [25] ZHENG Z, LU J, WANG Z, et al. Improving corrosion resistance of copper-nickel alloys by the microalloying-facilitated formation of protective corrosion product films[J]. *Corrosion Science*, 2025, 245: 112693.
- [26] Hu Yong, Liao Bokai, Chen Lijuan, et al. Mitigating sulfate-reducing bacteria-induced corrosion of pure copper in simulated oilfield-produced water using cetylpyridinium chloride[J]. *Coatings*, 2025, 15: 308.
- [27] Wei Boxin, Cai Zheng, Niu Mengchao, et al. Synergistic microbial interactions and electrochemical mechanisms driving microbiologically influenced corrosion in offshore platform produced seawater at 60°C C[J]. *Corrosion Science*, 2025, 248: 112797.

责任编辑: 刘娟

渗碳淬火对螺旋锥齿轮齿面误差影响分析与修正

顾井飞*, 陈兴强

(蚌埠学院 机械与车辆工程学院, 安徽 蚌埠 233030)

摘要:为补偿渗碳淬火导致的热处理变形,提升螺旋锥齿轮的齿面精度,通过有限元仿真与实验验证相结合的方法开展研究。首先,基于 ABAQUS 软件模拟螺旋锥齿轮的渗碳淬火过程,揭示了其齿面变形的定性规律。随后,以一对模数 2.5 mm、齿数比 33:8 的螺旋锥齿轮为例,依据仿真预测的变形趋势,对机床加工参数进行反向预调整,使齿面形状在热处理后变形至理论设计位置。最后,通过三坐标测量与齿面接触区检验,验证了该预变形修正方法的有效性与工程实用性。

关键词:渗碳淬火;螺旋锥齿轮;齿面误差;预变形修正

中图分类号:TH132.421;TG156.81 **文献标识码:**A

文章编号:2095-297X(2026)05-0034-06

Influence Analysis and Modification of Surface Error of Spiral Bevel Gear for Carburizing and Quenching

GU Jingfei*, CHEN Xingqiang

(School of Mechanical and Vehicle Engineering, Bengbu University, Bengbu, 233030, Anhui)

Abstract: To compensate for heat treatment deformation induced by carburizing and quenching and improve the tooth surface accuracy of spiral bevel gears, an investigation was conducted through a combined approach of finite element simulation and experimental verification. First, the carburizing and quenching process of a spiral bevel gear was simulated using ABAQUS software, and the qualitative deformation pattern of the tooth surface was revealed. Subsequently, taking a pair of spiral bevel gears with a module of 2.5 mm and a tooth number ratio of 33:8 as an example, the machining parameters of the machine tool were reversely pre-adjusted based on the deformation trend predicted by the simulation, so that the tooth surface geometry could be deformed to the theoretical design position after heat treatment. Finally, the effectiveness and engineering practicality of this pre-deformation compensation method were verified through coordinate measuring machine inspection and tooth contact pattern tests.

Key words: carburizing and quenching; spiral bevel gear; tooth surface error; pre-deformation compensation

螺旋锥齿轮广泛应用在摩托车、汽车、飞机等传动装置中,其齿面是复杂空间曲面,齿面啮合质量直接影响传动平稳性和承载能力,传动平稳性主要体现在振动和噪声,齿面接触应力对齿轮的使用寿命有直接影响。齿面的啮合质量除了装配误差外主要受齿面误差影响,其主要由加工误差和热处理变形

等因素造成^[1-2]。为了获得理想的啮合质量,减小齿面误差,一些学者在如何减少齿面加工误差方面做了一些研究。张宇等^[3-4]建立双重螺旋法的切齿数学模型,推导机床调整参数的计算过程,并建立了刀具参数与轮齿接触特征之间的作用规律,为优化双重螺旋法的切齿加工参数提供了理论依据。吴顺

收稿日期:2025-12-31 * 通讯联系人

基金项目:安徽省高校省级自然科学基金项目(KJ2020A0741);蚌埠学院自然科学基金项目(2024YYX05pj,2024YYX03pj)。

作者简介:顾井飞(1977—),男,安徽宿州人,实验师。E-mail:366298149@qq.com

兴等^[5-6]建立齿面误差修正模型,并采用含信赖域策略的最小二乘法和 Levenberg-Marquard 算法对齿面误差进行修正。唐进元、宋碧芸^[7-10]通过建立机床、齿轮加工的数学模型,运用解析法和非线性迭代法求解,确定了准双曲面齿轮加工过程中对齿面误差影响较大的调整参数误差项,并在加工中基于三坐标测量机的螺旋锥齿轮加工参数反调修正,减小齿面误差。由于传动强度和寿命需要,常采用热处理对锥齿轮进行强化处理。根据齿轮传动的特点,常选用低碳合金钢和渗碳淬火工艺,热处理中温度变化所产生的组织应力和热应力会引起齿面变形,影响齿轮的传动质量。

1 齿面误差及热处理对其影响

1.1 螺旋锥齿轮齿面误差的定义

螺旋锥齿轮齿面误差定义为实际齿面与理论齿面在理论齿面法向方向上的偏差,包括了齿圈跳动、齿距误差、齿形及齿向等误差的综合误差^[11]。锥齿轮的齿面误差评定可以通过数值分析,如图 1 所示,根据本研究齿轮参数及结构尺寸,将齿面按齿长和齿高的方向取 5 行 9 列共 45 个点,由啮合原理可以获得齿面方程,进而获得 45 个点的理论坐标值,通过三坐标测量仪测量实际齿面对应 45 个点的实际坐标,其理论值与测量值的差值即为误差值。

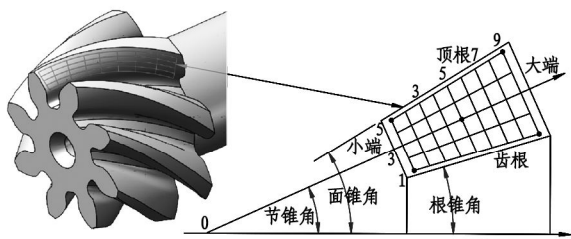


图 1 齿面点离散示意图

生产中,锥齿轮的齿面误差也可以通过着色法并以齿轮副相互啮合的“接触区”做定性分析。齿面接触区主要表达两个啮合齿面接触部分的形状、大小和在齿面的位置,且接触区检查方法具有直观、操作简单等优点,在生产过程中大量采用,如图 2 所示。在 ISO17485-2009 锥齿轮精度标准中,接触区质量未根据精度等级进行关联,而是依据加工实践经验和客户的要求以满足最终使用作为判定接触区质量的依据。美国 GLEASON 公司对锥齿轮副理想的接触区位置定义为接触区位于齿面中部附近略偏大端,占齿面 60% 左右,形状为椭圆形。

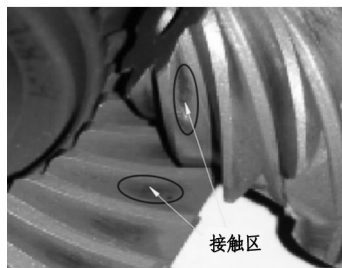


图 2 接触区检查示意图

1.2 热处理对齿面误差的影响因素

热处理对齿形误差的影响因素主要包括齿轮材料、结构尺寸、热处理工艺参数等,一些学者对此进行了研究。螺旋锥齿轮由于空间曲面的复杂性,当锥齿轮处于淬冷阶段的时候,会发生变形,这主要是因为两类应力在起作用。其一是热应力,在锥齿轮进行淬冷之时,其各个部位所发生的温度变化是存在差异的,且这些部位的冷却速度也并非一致,使得不同部位所产生的热应力出现不一致的情形,进而致使变形呈现出不一致的状态。其二是组织应力,即锥齿轮的金相组织从奥氏体转变为马氏体时,由于芯部与表面的冷却速度存在差别,所以组织从奥氏体向马氏体转变的顺序以及转变量也都不相同,而马氏体相变这一过程会导致体积出现膨胀的情况,这种组织转变过程的不一致性会催生出组织应力,最终使得变形呈现不一致的状况^[12]。就渗碳锥齿轮来讲,因为齿轮的含碳量是从表层开始一直往里不断发生变化的,并且从表层一直到芯部,马氏体相变的温度是逐步升高的,因此其组织应力所起到的作用就更为复杂,对其做出精准分析更为困难。所以,对于渗碳锥齿轮变形展开有限元分析,了解渗碳以及马氏体相变的具体情况是必要的^[13]。基于多场耦合软件 DEFORM-HT 的扩散与相变模型,采用数值分析方法反馈介质流速对弧齿锥齿轮冷却特性的影响,给出了齿轮在不同介质流速下温度、马氏体体积分数、硬度及应力的演化规律^[14-15]。由于影响热处理变形的因素多且复杂,前面的研究分析大都是针对具体情况和定性的分析。

金属材料在经过热处理之后所产生的变形,一般被划分成两种类型。其中一种是内应力而引发的塑性形变,主要是因为在对工件进行不断的加热、保温以及冷却的操作过程当中,工件的不同部位受热的程度并不均匀,冷却的速率也有不小的差异,材料所具有的热胀冷缩效应在各个不同的位置上所呈现出的效果也是不一样的,因此产生了应力塑性形变。另一种形变是由于组织转变带来的形变,比如马氏

体为体心立方晶格,致密度为 0.68,而奥氏体是面心立方晶格,其致密度为 0.74。因此奥氏体在转化为马氏体时体积会膨胀,由此产生组织应力,再加上淬火冷却快,工件芯部温度比表面高,巨大温差会引起热应力。凡是对上述两种变形有影响的因素都会影响齿面误差,其主要包括以下一些因素。

工件材料自身特性。工件材料自身特性主要指材料淬透性、材料成分、材料组织均匀性和晶粒度。材料的淬透性决定着材料淬硬深度和硬度分布,影响材料热处理后的力学性能,材料的淬透性主要受奥氏体的化学成分和奥氏体化条件影响。材料中元素的不同含量,会影响奥氏体的转变曲线,从而会影响工件的变形。

工件的结构、尺寸。在相同材料及热处理的条件下,模数较小,齿数较少,也就是尺寸较小的工件变形相对较大,齿轮螺旋角越大,变形相对越大。工件的结构越复杂、壁厚差异越大,则变形相对越大。

毛坯类型和组织情况。齿轮毛坯的类型主要有锻造、铸造和圆钢加工,锻造可以减少偏析、变形,有利于均匀组织。为了消除锻造应力,锻造毛坯一般需要做正火处理,正火的毛坯硬度高低会影响后续加工的变形,硬度过高或过低都不好,一般控制在 HB170-220。同时组织状态也影响热处理变形,一般组织越均匀,晶粒越细小,变形越小。

冷加工工艺。工件在冷加工中的工艺不同,工件表面的残余应力大小和性质就不同,也会影响热处理变形。应力越小越均匀,热处理变形越小。

热加工工艺。渗碳淬火齿轮的装料方法、淬火温度、预冷时间、淬火介质等热处理工艺方法和工艺参数都会对变形产生影响。

2 热处理对齿面误差的影响机理及仿真

2.1 热处理过程的温度场

温度场的变化是研究热处理过程中其他场变化规律的基础,齿轮热处理过程是一个复杂的传热过程,具有内、外热源相互作用的三维非稳态问题^[16]。根据傅里叶能量守恒方程和传热定律,推导出瞬态传热方程为:

$$\rho c \frac{\partial T}{\partial t} = \lambda \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \right) + Q_x \quad (1)$$

式中: ρ 为材料的密度, kg/m^3 ; c 为材料的比热容, $\text{J} \cdot \text{kg}^{-1} \cdot ^\circ\text{C}^{-1}$; T 为工件的温度, $^\circ\text{C}$; t 为进行时间, s ; λ 为材料的导热系数, $\text{W} \cdot \text{mm}^{-1} \cdot ^\circ\text{C}^{-1}$; x 为横向坐标, mm ; y 为纵向坐标, mm ; z 为轴向坐标, mm ; Q_x

为相变热量, J , 其表示物体温度升高需要的热量,来自于两部分,等于内热源和外部流入热量的热量之和,当内在和外部热源的热流密度相对稳定时,温度达到稳定,此时瞬态方程可以转化为稳态方程。

$$\lambda \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \right) + Q_x = 0 \quad (2)$$

2.2 热处理过程的组织转变

热处理过程中,铁素体和贝氏体属于扩散型相变,组织转变量的数学计算式表示为:

$$V = 1 - \exp(-bt^n) \quad (3)$$

式中: V 为组织转变量; t 为等温时间, s ; b 为新相形核系数; n 为新相长大系数。其中 b, n 可从等温转变曲线中对应某一温度 T 的两个等温时间 t_1 和 t_2 的转变量 V_1 和 V_2 求得

$$b = - \frac{\ln(1 - V_1)}{t_1^n} \quad (4)$$

$$n = \frac{\ln \left[\frac{\ln(1 - V_1)}{\ln(1 - V_2)} \right]}{\ln \left(\frac{t_1}{t_2} \right)} \quad (5)$$

而马氏体转变属于非扩散型转变,由过冷度决定转变量而与时间无关,转变量的计算公式为:

$$V = V_{\max} \{ 1 - \exp[c(M_s - T)] \} \quad (6)$$

式中: $V_{\max}$ 为马氏体最大转变量; c 为应力影响参数; M_s 为马氏体相变温度, $^\circ\text{C}$; T 为过冷度, $^\circ\text{C}$ 。

2.3 热处理过程的仿真

利用 3D 建模软件 UG 建立锥齿轮模型,将模型导入有限元软件 ABAQUS 中,仿真模拟锥齿轮在给定热处理工艺条件下的淬火过程。首先要确定模型的材料种类,本研究齿轮选用的材料为 20CrMnTi,其和仿真模型的材料设定一致;选取合适的单元类型并进行合理的网格划分,设定正确的边界条件和合理的约束条件。依据传热方程和热处理工艺参数,选取温度场模拟所需的性能参数,仿真模拟温度场的变化,研究齿轮各部位的冷速情况。

模拟结果表明齿轮各处不同冷却速度和内外温差是造成变形的主要原因。根据应力场模拟理论方程,对锥齿轮淬火应力场进行仿真模拟,其结果表明淬火过程初期阶段,造成应力变化和变形的主要原因是热应力,而在发生相变阶段组织应力又起主要作用。同时热应力和组织应力又相互影响,最终作用的结果是齿轮表面呈现为压应力,心部呈现为拉应力的状态。

3 对热处理造成齿面误差的修正方法

渗碳淬火造成的变形误差直接影响齿轮的啮合质量,消除变形造成的影响主要有三种方法,根据不同实际情况选择。

3.1 磨齿消除热处理齿面变形

如果对于齿轮啮合质量要求高,齿轮的结构、尺寸又满足磨齿的条件,通常采用铣齿时留有磨削余量,热处理后在磨齿机上用砂轮切除余量消除热处理带来的变形,但其成本较高,本研究不再阐述。

3.2 研磨修正法

如果对于齿轮啮合质量要求不是很高,或者齿轮的结构、尺寸不适合磨齿,可以采用研磨修正法,如图3所示。配对齿轮以装配的位置状态进行对研,在研磨膏的作用下,“凸”的变形被部分修正。研齿可以消除部分齿轮加工或热处理后产生的齿形误差,改善齿面接触质量;研磨使齿面接触区域分布更合理,避免局部应力集中,提高承载能力;研齿还可以消除齿面微观不平度,从而降低噪声。另外研

磨修正法工艺简单,效率高,但会对齿面造成一定的“磨损”,影响寿命。而且,对于配对研磨的齿轮最好配对使用。



图3 研齿加工示意图

3.3 预变形修正法

3.3.1 实例齿轮参数

本研究以模数为 2.5 mm,大小齿轮齿数比为 33:8 的一对螺旋锥齿轮为例验证该方法的有效性。螺旋锥齿轮几何参数和齿轮铣齿加工参数分别如表1和表2所示。

表1 螺旋锥齿轮几何参数

齿轮	齿数	节圆直径 /mm	外端大径 /mm	中点螺旋角 /(°)	面锥角 /(°)	节锥角 /(°)	根锥角 /(°)	中点锥距 /mm	冠顶距 /mm
大轮	33	82.5	83.081	35	79.297	76.373	71.071	29.945	8.801
小轮	8	20	25.965	35	18.929	13.627	10.793	29.945	40.507

表2 齿轮铣齿加工参数

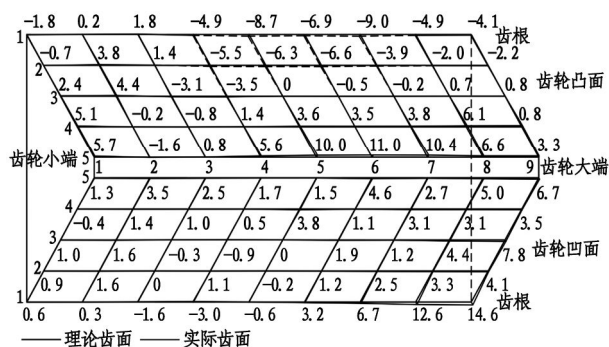
齿轮	刀盘直径 /英寸	刀号	错刀距 /mm	螺旋方向	摇台角 /(°)	摇台摆角 /(°)	根锥角 /(°)	偏心角 /(°)	滚比
大轮	2.5	14	0.9	左旋	320	85.46	71.071	51.146	1.029
小轮	2.5	14	0.9	右旋	91.14	67.69	10.793	51.146	4.244

齿轮按照铣齿加工参数加工并进行接触区检查,如图4所示,接触区位于齿面中部附近略偏大端,占齿面高度和长度都在60%左右,形状近似为椭圆形,符合要求。

利用三坐标测量所得齿面误差的数值如图5所示,由图可知凸面最大误差为-11 μm,误差的平方和为1041 μm²,绝对平均误差为3.2 μm,凹面最大误差为14.6 μm,误差的平方和为723 μm,绝对平均误差为2.7 μm。这些误差主要是刀具的制造误差、零件装夹误差和测量误差等。



图4 理论铣齿齿面接触区

图5 理论铣齿齿面小锥齿网格数据(单位 μm)

3.3.2 实例齿轮热处理

该对齿轮材料为 20CrMnTi, 热处理采用渗碳淬火, 渗层厚度 0.5 – 0.8 mm, 小齿轮表面硬度换算成洛氏硬度为 57 – 62 HRC, 大齿轮硬度为 55 – 60 HRC。齿轮热处理工艺流程如图 6 所示, 包含升温、渗碳、强渗、扩散、保温、淬火等流程。首先炉温加热到 880 $^{\circ}\text{C}$, 从升温开始 3000 s 后, 然后向炉中通入渗碳气体, 控制碳势为 1%, 温度升高到 900 $^{\circ}\text{C}$, 持续时间为 7200 s。由热处理经验表明, 如果采用一种渗碳条件不变, 表层碳元素含量不均匀。为此将对工件进行“强渗 + 扩散处理”, 工艺过程为炉温由 900 $^{\circ}\text{C}$ 升高到 915 $^{\circ}\text{C}$ 、碳势升高到 1.2% 的环境中保温 6000 s, 这是“强渗阶段”; 然后再将碳势与温度分别降至 0.9% 与 900 $^{\circ}\text{C}$, 持续时间 4800 s, 这是“扩散处理”。为了减小变形与应力集中, 采用预冷淬火工艺路线, 首先将炉温由 900 $^{\circ}\text{C}$ 降至 830 $^{\circ}\text{C}$, 保温 600 s 结束后, 将料架沉入温度为 60 $^{\circ}\text{C}$ 淬火液内淬火。

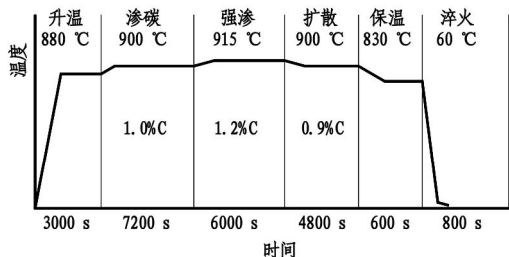


图6 热处理渗碳淬火工艺流程

通过金相分析, 轮齿表面马氏体含量较高, 测量硬度为 63.5 HRC, 具有很高的耐磨性。芯部组织主要由铁素体、珠光体、马氏体组成, 体积分数分别为 0.15、0.29 和 0.46。由于芯部碳含量较低, 而且轮齿芯部主要为板条状马氏体组织, 测量芯部硬度为

40.7 HRC, 硬度和耐磨性较低表面低, 但是具有较强韧性和抗疲劳强度, 低温回火后满足大小齿轮热处理的硬度和金相组织要求。

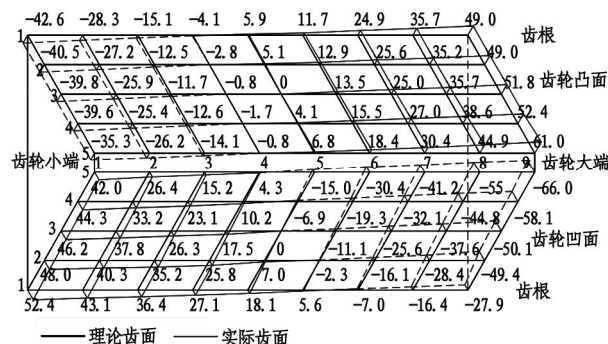
3.3.3 热处理变形

分别热处理后进行接触区检查和三坐标数值检查, 结果如图 7 和图 8 所示。从接触区看小齿轮凹面偏小端, 凸面偏大端, 体现了齿形被“拉直”的趋势, 但由于齿面形状的复杂性, 并不是均匀变形。



图7 理论铣齿热处理后齿轮接触区

从图 8 可以看出, 热处理变形比较明显, 凸面最大误差为 61 μm , 误差的平方和为 34 748 μm^2 , 绝对平均误差由 3.2 μm 变为 24.2 μm , 凹面最大误差为 -66 μm , 误差的平方和为 43539 μm^2 , 绝对平均误差由 2.7 μm 变为 29 μm 。

图8 理论铣齿热处理后齿面小锥齿网格数据(单位 μm)

3.3.4 预变形修正

预变形法就是根据热处理对齿面变形的影响规律, 在热处理前加工齿轮时调整相关加工参数, 使齿面处于“非正确”的位置, 然后通过热处理变形“变”到较理想的位置。通过多次实验调整, 把小齿轮加工参数中的偏心角减小 20 分, 摇台角增大 15 分时, 加工参数见表 3, 热处理后的齿面质量理想。热处理前接触区图和小锥齿的齿面网格数据分别如图 9 和图 10 所示。从接触区看小齿轮凹面偏大端, 凸面偏小端。

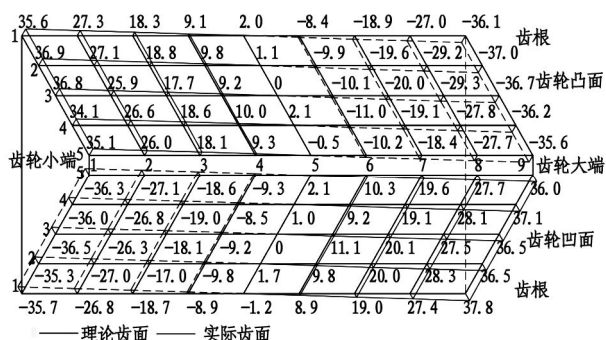
表3 调整后的齿轮铣齿加工参数

齿轮	刀盘直径 /英寸	刀号	错刀距 /mm	螺旋方向	摇台角 /(°)	摇台摆角 /(°)	根锥角 /(°)	偏心角 /(°)	滚比
大轮	2.5	14	0.9	左旋	320	85.46	71.071	51.146	1.029
小轮	2.5	14	0.9	右旋	91.39	67.69	10.793	50.813	4.244



图9 预调整铣齿热处理前齿轮接触区

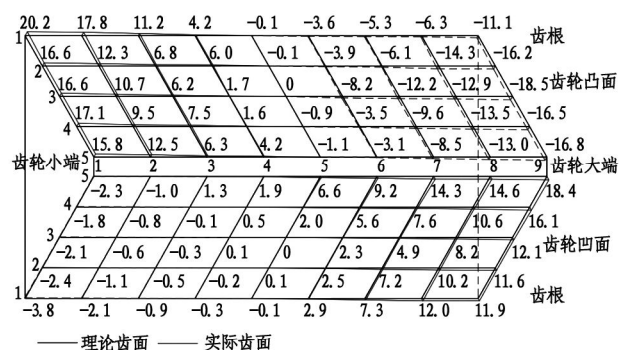
从图10可以看出,凸面最大误差为 $36.9 \mu\text{m}$,误差的平方和为 $21\,373 \mu\text{m}^2$,绝对平均误差为 $20.5 \mu\text{m}$,凹面最大误差为 $37.8 \mu\text{m}$,误差的平方和为 $21\,601 \mu\text{m}^2$,绝对平均误差为 $20.6 \mu\text{m}$ 。

图10 预调整铣齿热处理前齿面小锥齿网格数据(单位 μm)

热处理后接触区图和小锥齿的齿面网格数据分别如图11和图12所示,从图11看出热处理后接触区位置处于较理想状态。从图12可以看出,凸面最大误差为 $20.2 \mu\text{m}$,误差的平方和为 $4426 \mu\text{m}^2$,绝对平均误差为 $9.1 \mu\text{m}$,凹面最大误差为 $18.4 \mu\text{m}$,误差的平方和为 $2265 \mu\text{m}^2$,绝对平均误差为 $4.9 \mu\text{m}$ 。



图11 预调整铣齿热处理后前齿轮接触区

图12 预调整铣齿热处理后齿面小锥齿网格数据(单位 μm)

3.3.5 两种方法的比较

采用研磨方法修正齿形和预先调整参数修正齿形的两种方法均能有效修正齿面接触区,提高接触质量,但后者明显使得齿面形状在热处理后更接近理论值,齿面接触质量更好,经实际验证,后者的噪音明显要低 $2 \sim 4 \text{ dB}$ 。

4 结论

本研究分析了产生齿面误差的原因、齿面误差对传动质量的影响,以及热处理对齿形误差的影响因素,其主要有齿轮材料、结构尺寸、热处理工艺参数等,即使利用大型有限元软件 ABAQUS,也很难确定热处理引起的变形规律,只能做定性分析。消除热处理对齿面变形的影响通常有三种方法。本研究以模数为 2.5 mm ,大小轮齿数比为 $33:8$ 的一对螺旋锥齿轮为例,采用研齿和预变形法两种工艺减少热处理对齿面的影响,并通过利用三坐标测量机上进行测量和齿面接触器检验,都验证了该方法的正确性和工程实用性。两种方法各有特点,研齿工艺简单,会提高齿面接触质量,但只能部分修正误差,还会对齿面造成一定的“磨损”,影响寿命;研磨的齿轮配对使用效果最好,但增加了生产管理难度。预变形法可以有效修正误差,而且是一次性批量完成,但由于产品的结构、尺寸、材料等因素影响,对修正量不好确定,需要在仿真和三坐标测量机测量的基础上实验验证。另外,对那些需要热处理后磨齿的齿轮,也可以用预变形减少变形量,从而减少磨齿的加工余量。

(下转第62页)

基于 Kriging 模型的龙门加工中心主轴箱多目标优化

陈宇*, 吴紫鹏, 彭兆春, 赵林昊, 李奥迪

(巢湖学院 机械工程学院, 安徽 巢湖 238024)

摘要:为加强龙门加工中心主轴箱综合性能,并进一步实现轻量化与动态性能,选取某高速龙门加工中心主轴箱为研究对象,利用 ANSYS 对主轴箱部件进行静态分析得出了基础结构性能、最大形变量与最大等效应力值,以龙门加工中心横加强筋间距、竖加强筋间距、主轴箱前后壁厚、主轴箱左右壁厚、横加强筋厚度、竖加强筋厚度的相关尺寸参数为优化设计变量,以龙门加工中心主轴箱的几何结构质量、几何结构、最大等效应力、最大总变形为优化目标,对龙门加工中心主轴箱进行响应面优化分析,与优化前相比,龙门加工中心主轴箱最大变形量增加了 6.33%,最大应力减小了 12.47%,质量减轻了 7.25%,实现减重释放结构潜力,为龙门加工中心主轴箱的性能优化提供了理论指导。

关键词:主轴箱;静态特性分析;Kriging 模型;多目标优化

中图分类号:TG659;TH122

文献标识码:A

文章编号:2095-297X(2026)05-0040-06

Multi-objective Optimization of Headstock of Gantry Machining Center Based on Kriging Model

CHEN Yu*, WU Zipeng, PENG Zhaochun, ZHAO Linhao, LI Aodi

(School of Mechanical Engineering, Chaohu University, Chaohu, 238024, Anhui)

Abstract: In order to enhance the comprehensive performance of the spindle box of the portal machining center and further achieve lightweighting and dynamic performance, a certain high-speed portal machining center spindle box was selected as the research object. Using ANSYS, the static analysis of the spindle box components yielded the basic structural performance, the maximum deformation and maximum equivalent stress values. The relevant dimension parameters of the portal machining center, such as the horizontal spacing of the reinforcing ribs, the vertical spacing of the reinforcing ribs, the thickness of the front and rear walls of the spindle box, the thickness of the left and right walls of the spindle box, and the thickness of the horizontal reinforcing ribs and the vertical reinforcing ribs, were used as the optimization design variables. The geometric structural quality, geometric structure, maximum equivalent stress, and maximum total deformation of the portal machining center spindle box were used as the optimization objectives. The response surface optimization analysis showed that compared with the original design, the maximum deformation of the portal machining center's spindle box increased by 6.33%, the maximum stress decreased by 12.47%, and the weight was reduced by 7.25%. This minor weight reduction released structural potential, providing theoretical guidance for the performance optimization of the portal machining center's spindle box.

Key words: main spindle box; static characteristics analysis; Kriging model; multi-objective optimization

龙门加工机床主轴箱是机床核心部件之一,它通过电机驱动主轴带动刀具工作,靠轴承和齿轮系统保持稳定、减少振动,保证主轴旋转精度以避免加工偏差,直接影响着机床的加工精度、稳定性和工作效率。对龙门加工中心主轴箱做出静态分析与响应面的多目标优化,提高了机床加工精度与稳定性,同时质量与体积这两个关键指标均呈现出显著降低趋势。由此可见,该优化流程实现了机械结构的轻量化设计。

国内研究多集中在龙门加工中心关键部件,如横梁和滑枕。蒋莉^[1]等利用 Kriging 模型和遗传算法相结合的方法,避免了在方案比选优化上的不确定性和烦琐性,显著提升了优化设计的效率,比传统的优化比选有着更优的优化效果。杜贝^[2]通过建立龙门加工中心,并对各零部件做静态特性分析和模态分析,有效提升了主轴箱的固有频率,降低了低频率共振的可能性,使得主轴箱稳定性上升,满足了实验设计。张玥^[3]对龙门加工中心主轴箱做了静态特性分析和目标优化,优化后的加工中心主轴箱外形未发生改变,但最大等效应力减少了 15%,主轴箱重量减少了 7.8%,主轴箱优化前后的振动特性对比试验也验证了优化后的主轴箱结构的动态性能有了大幅的提高,满足了实验设计。张子琦^[4]通过 Kriging 代理模型对风机做多目标优化,优化后的风机在全压和效率方面分别提升了 12.7% 和 7.1%,这证明能量损失减少,风机性能提升,满足了实验设计。陈鹏^[5]通过建立龙门加工中心主轴箱模型,对其进行静态特性分析和模态分析,并作出目标优化,优化后的主轴箱变形幅度下降了 2%,优化后的最大等效应力下降了 15%,优化效果明显。陈国雄等^[6]针对龙门架结构存在质量偏大和隔振能力较弱的问题,提出了一种基于有限元法-近似模型-遗传算法三者结合的多目标优化方法。优化后龙门架质量减少了 7.34%,最大静变形量增加了 0.0434 mm,最大等效应力增加了 4.16 MPa,1 阶模态频率提高了 1.22 Hz,可满足焊接平台的设计要求。李有堂等^[7]采用拓扑优化与二维截面优化相结合的方法确定材料分布情况,基于 MOGA 遗传算法,以质量、最大静变形量与第 1 阶固有频率为优化目标,进行参数优化。结果表明,最大静变形量减小了 34.61%,质量减小了 7.64%,第 1 阶固有频率提高了 6.54%,验证了龙门架的优化效果。彭伊丽^[8]针对切削激励下机床工作空间振动特性,提出了随机激励下基于响应的频响函数正则化方法、基于模态质量分布和聚类算法的工作空间分区方法,以及

基于模态状态空间模型和卡尔曼滤波的主振模态在线识别方法,并在龙门加工中心上验证了这些方法的有效性,最终实现了以主振模态对应薄弱环节动柔度最小为目标、通过合理选择加工位置与切削参数来减小机床强迫振动。

本研究基于 ANSYS 仿真软件对龙门加工中心主轴箱进行静态分析和模态分析,获取加工状态下的结构静态特性与动态性能特征;采用非参数化建模方法,基于 Kriging 模型构建响应面,以前后壁厚、左右壁厚、加强筋厚度和增大筋间距为主要参数,通过 MOGA 算法进行多目标优化;最后通过实例验证优化效果,以提升其结构性能并实现轻量化。

1 静态分析

1.1 主轴箱的建模

利用 UG 软件对主轴箱进行三维建模,为保证仿真结果的精确性,对一些细节做出了简化,去除了对一些对模型静态特征分析影响较小的特征,如倒角,圆角,螺钉孔等。图 1 为某公司设计的一款高速龙门加工设备,产品型号为 PM1220HA。

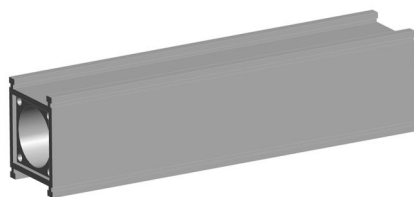


图 1 主轴箱三维建模

1.2 材料类型与网格属性划分

选取 HT200 为主轴箱,材料参数如表 1 所示。

表 1 材料参数

参数	密度/ $\text{kg} \cdot \text{m}^{-3}$	弹性模量/GPa	泊松比
数值	7.2	110	0.28

网格划分精度直接影响求解结果的真实性,研究采用全局-局部耦合的网格优化算法,通过设定 12 mm 特征尺寸阈值,实现了单元密度梯度控制。建立的模型共划分为 243 414 个节点和 140 329 个单元,网格划分结果如图 2 所示。

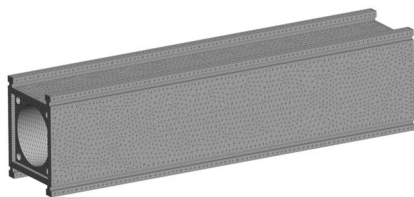


图 2 主轴箱网格划分

1.3 工况设置

龙门加工中心运行过程中,主轴箱承受的载荷状况呈现多样性特征。主要对应钻削、铣削、镗削等典型加工工况,其中铣削工况载荷最具代表性,本研究以铣削工况为核心开展受力分析。

铣削力分析遵循金属切削原理,铣削进给力的公式如下

$$F = C_F \times a_p^x \times a_e^y \times f_z^m \times Z \times K_F \quad (1)$$

式中: F 为铣削力; C_F 为与刀具工件材料有关的系数; a_p 为铣削进给深度; a_e 为铣削进给宽度; f_z 为每齿进给量; Z 为铣刀齿数; K_F 为进给力修正系数。

$$F_x = k_x \cdot F, k_x \in [0.3, 0.4] \quad (2)$$

$$F_y = k_y \cdot F, k_y \in [0.5, 0.55] \quad (3)$$

$$F_z = \sqrt{F^2 - F_x^2 - F_y^2} \quad (4)$$

式中: F_x 为横向切削力; F_y 为纵向切削力; F_z 为垂直切削力; k_x 、 k_y 分别为横、纵向切削力比例系数。

龙门加工中心铣削加工参数如下:铣削刀具是硬质合金钢材料,铣削刀具的尺寸为 50 mm,铣削刀具的齿数为 8 齿,被铣削工件属性为 45 钢,工件属性系数 120,铣削深度 18 mm,铣削宽度 12 mm,每齿进给量 0.25 mm/z,指数 x 、 y 、 m 分别为 0.9、0.7、0.15。将参数带入公式(1)至公式(4)可得:

$$F = 554 \text{ N}$$

$$F_x = 2217.6 \text{ N}$$

$$F_y = 3049.2 \text{ N}$$

$$F_z = 4064.6 \text{ N}$$

1.4 结果分析

主轴箱静态特性分析:先算所需受力大小,再施加对应载荷力及固定约束,求解时插入总变形和等效应力,因力为分量,需对各分量做定向变形,得到总变形及等效应力变化图,分别如图 3、图 4 所示。

由上可知主轴箱最大总变形量为 0.000 375 42 mm,变形量较小,满足静态特性分析要求,其最大等效应力为 0.659 63 MPa,对比材料的屈服极限应力,变化量微小,本课题主轴箱满足实验要求。

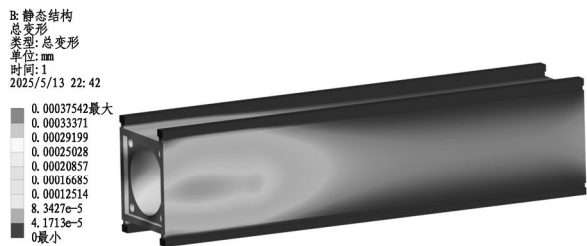


图 3 总变形变化图



图 4 等效应力变化图

2 龙门加工中心主轴箱响应面优化

2.1 响应面优化设计变量

本研究选取横加强筋间距 $P3$ 、竖加强筋间距 $P7$ 、主轴箱前后壁厚 $P4$ 、主轴箱左右壁厚 $P8$ 、横加强筋厚度 $P5$ 、竖加强筋厚度 $P6$ 共六个参数作为设计变量;将主轴箱几何结构质量 $P1$ 、主轴箱体积 $P2$ 、最大等效应力 $P9$ 、最大总变形 $P10$ 设置为响应指标,组建优化参数体系。通过调整设计变量的取值,实现主轴箱静态特性、动态性能及质量的协同优化,完成轻量化设计目标。

龙门加工中心主轴箱输入参数横加强筋间距 $P3$ 、竖加强筋间距 $P7$ 、主轴箱前后壁厚 $P4$ 、主轴箱左右壁厚 $P8$ 、横加强筋厚度 $P5$ 、竖加强筋厚度 $P6$ 的设计如表 2 所示。

表 2 参数设计

输入参数/mm	下限值	初始值	上限值
$P3$	270	300	330
$P4$	27	30	33
$P5$	18	20	22
$P6$	18	20	22
$P7$	72	80	88
$P8$	27	30	33

2.2 实验结果及分析

本次实验采用 Box-Behnken 设计,无试验点镜像。以横加强筋间距($P3$)、竖加强筋间距($P7$),主轴箱前后壁厚($P4$)、主轴箱左右壁厚($P8$),横加强筋厚度($P5$)、竖加强筋厚度($P6$)为输入变量;以主轴箱几何结构质量($P1$)、主轴箱体积($P2$)、最大等效应力($P9$)、最大总变形($P10$)为输出变量。通过优化设计得到 25 组实验参数及变化范围,如表 3 所示。各个参数之间的影响的响应面分析变化如图 5 至图 7 所示。

表 3 实验设计

名称	P3/mm	P4/mm	P5/mm	P6/mm	P7/mm	P8/mm	P1/kg	P9/MPa	P10/mm
1	300	30	20	20	80	30	935.3	0.30	0.00030
2	270	27	20	18	80	30	895.5	0.32	0.00030
3	330	27	20	8	80	30	895.5	0.32	0.00030
4	270	33	20	18	80	30	964.6	0.30	0.00029
5	330	33	20	18	80	30	964.6	0.29	0.00029
6	270	27	20	22	80	30	905.9	0.31	0.00030
7	330	27	20	22	80	30	905.9	0.32	0.00030
8	270	33	20	22	80	30	975.0	0.30	0.00029
9	330	33	20	22	80	30	975.0	0.27	0.00029
10	300	27	18	20	72	30	899.5	0.33	0.00030
11	300	33	18	20	72	30	968.6	0.29	0.00029
12	300	27	22	20	72	30	901.9	0.31	0.00030
13	300	33	22	20	72	30	971.0	0.30	0.00029
14	300	27	18	20	88	30	899.5	0.32	0.00030
15	300	33	18	20	88	30	968.6	0.30	0.00029
16	300	27	22	20	88	30	901.9	0.32	0.00030
17	300	33	22	20	88	30	971.0	0.30	0.00029
18	300	30	18	18	80	27	894.3	0.31	0.00031
19	300	30	22	18	80	27	896.7	0.31	0.00031
20	300	30	18	22	80	27	904.7	0.25	0.00031
21	300	30	22	22	80	27	907.0	0.31	0.00031
22	300	30	18	18	80	33	963.4	0.31	0.00028
23	300	30	22	18	80	33	965.9	0.31	0.00028
24	300	30	18	22	80	33	973.8	0.30	0.00028
25	300	30	22	22	80	33	976.2	0.30	0.00028

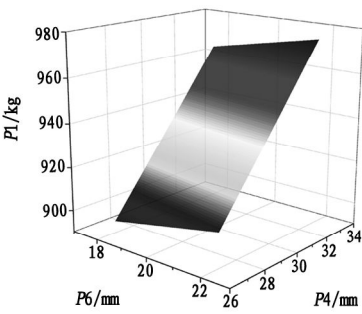


图 5 P4、P6 对质量 P1 响应图

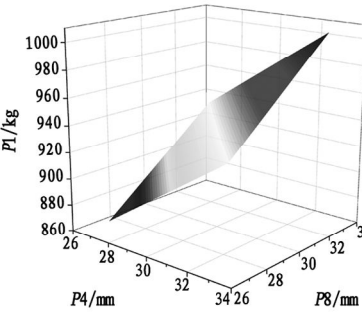


图 6 P4、P8 对质量 P1 响应图

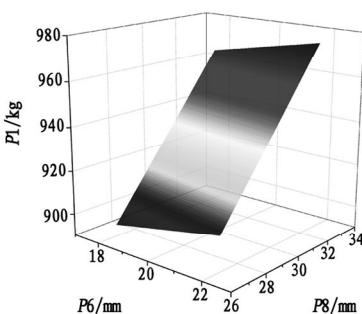


图 7 P6、P8 对质量 P1 响应图

由图 5 至图 7 可知, P_1 、 P_4 呈线性关系, P_4 越小,主轴箱质量 P_1 越小, P_6 相对关小较小。 P_1 对 P_4 和 P_8 呈线性关系, P_4 和 P_8 越小, P_1 越小。主轴箱质量 P_1 对主轴箱左右壁厚 P_8 呈线性关系, P_8 越小, P_1 越小, P_6 相对关小较小。

对表 3 数据进行分析可得主轴箱几何结构质量 P_1 和输入参数之间的关系为线性正相关,各输入参数对最大等效应力 P_9 和最大总变形 P_{10} 的影响如图 8 至图 13 所示。

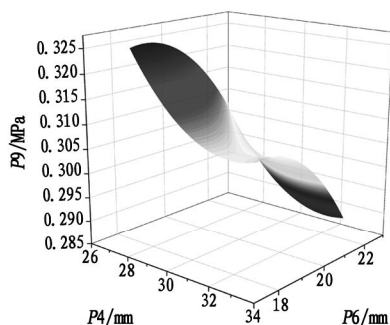


图 8 P_4 、 P_6 对最大等效应力 P_9 响应图

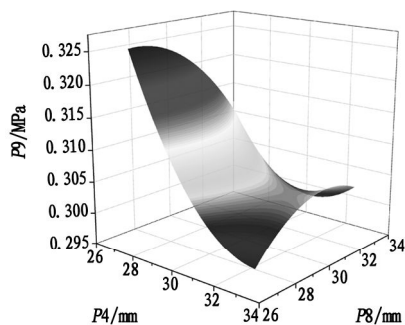


图 9 P_4 、 P_8 对最大等效应力 P_9 响应图

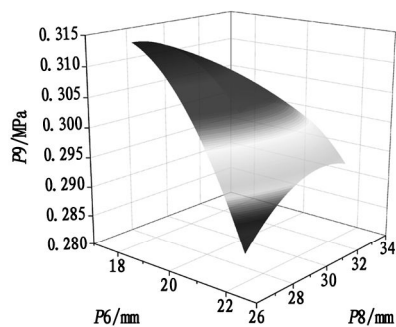


图 10 P_6 、 P_8 对最大等效应力 P_9 响应图

由图 8 至图 10 可知,随着 P_4 减小, P_9 增大, P_6 减小,那么 P_9 会急剧增大。随着 P_4 减小, P_9 先减小后增大, P_6 减小, P_9 会先增大然后再缓慢减小。随着 P_6 减小,那么 P_9 缓慢变大, P_8 减小,那么 P_9 会先增大后减小。

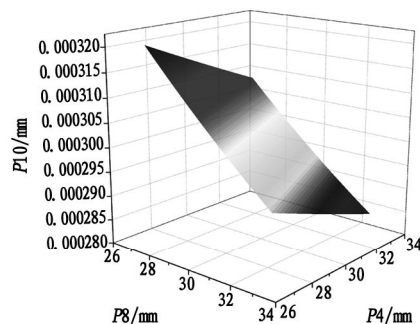


图 11 P_4 、 P_8 对最大总变形 P_{10} 响应图

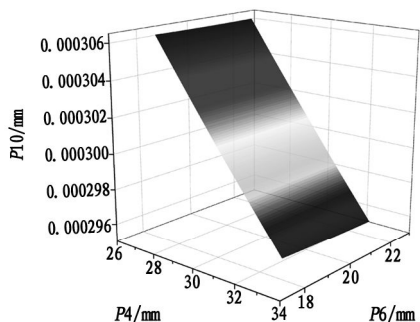


图 12 P_4 、 P_6 对最大总变形 P_{10} 响应图

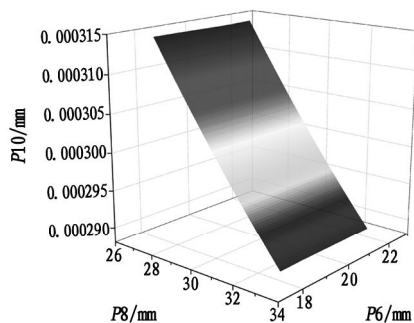


图 13 P_6 、 P_8 对最大总变形 P_{10} 响应图

由图 11 至图 13 可知,随着主轴箱前后壁厚 P_4 减小,主轴箱左右壁厚 P_8 减小,主轴箱最大总变形变大。随着主轴箱前后壁厚 P_4 减小,主轴箱最大总变形变大,竖加强筋厚度 P_6 影响不大。随着主轴箱左右壁厚 P_8 减小,主轴箱最大总变形变大,竖加强筋厚度 P_6 影响不大。

3 响应面优化多目标优化结果

以龙门加工中心主轴箱几何结构质量 P_1 最小化为本次实验项目的设计目标,以横加强筋间距 P_3 、主轴箱前后壁厚 P_4 、横加强筋厚度 P_5 、竖加强筋厚度 P_6 、竖加强筋间距 P_7 、主轴箱左右壁厚 P_8 为设计变量,结合静态分析结果设定约束条件,建立主轴箱优化数学模型如下

$$\begin{cases} \min M(P) \\ \delta_{\max} \leq [\delta] = 0.000301 \\ \sigma_{\text{eq,max}} \leq [\sigma] = 0.306666 \\ P = [P_3, P_4, P_5, P_6, P_7, P_8] \\ P_{i,\text{low}} \leq P_i \leq P_{i,\text{up}}, i = 3, 4, \dots, 8 \end{cases} \quad (5)$$

式中： M 为主轴箱质量； P 为结构设计变量； $\delta_{\max}$ 为主轴箱最大总变形； $[\delta]$ 为许用最大变形量（mm）；

$\sigma_{\text{eq,max}}$ 为主轴箱最大等效应力； $[\sigma]$ 为材料许用等效应力（MPa）； $P_{i,\text{low}}$ 、 $P_{i,\text{up}}$ 分别为各设计变量的下限、上限取值。

结合实验设计要求，选取多目标遗传算法（MOGA）进行优化求解，该算法能有效处理多约束、多目标优化问题，收敛性好、优化效率高。求解得到 3 组最优候选点，候选点参数如表 4 所示。

表 4 候选点参数表

名称	P3/mm	P4/mm	P5/mm	P6/mm	P7/mm	P8/mm	P1/kg	P9/MPa	P10/mm
1	329.54	27.452	18.62	21.985	87.841	27.017	876.3	0.2646	0.00031
2	328.37	27.304	18.57	21.984	87.941	27.032	874.8	0.2660	0.00031
3	329.59	27.077	18.61	21.972	87.857	27.017	872.0	0.2684	0.00032

通过对 3 组候选点的对比，选取龙门加工中心主轴箱质量最轻的候选点，对响应面优化前后的参

数作对比，如表 5 所示。

表 5 响应面优化前后模型数据对比

模型信息	原模型	响应面优化模型	与原模型对比（%）
最大变形量/mm	0.00037542	0.00032008	-14.74
最大应力值/MPa	0.6596	0.2684	-59.31
质量/kg	935.3	872.0	-6.77

由表 5 可知，优化后的模型比原模型性能得以提升，优化后最大变形量为 0.000 320 08 mm，与原模型比减小了 14.74%。优化后的最大应力值为 0.2648 MPa，与原模型相比减小了 59.31%。优化后的质量为 872 kg，与原模型相比质量减少了 6.77%。综合来看，优化后主轴箱的最大变形量控制在允许范围内，最大等效应力和质量显著降低，符合优化设计目标，实现了主轴箱轻量化，节约了材料成本。

4 结论

通过 UG 建立主轴箱的三维立体图形，再进入与 UG 关联的 Ansys WorkBench 中主轴箱材料类型并设定参数，对主轴箱作网格划分，随后施加力做工况分析，完成主轴箱静态特性分析的前期准备。

添加静态分析模块，在此模块中添加总变形和应力，算出主轴箱最大变形量和最大等效应力。然后在静态特性分析的基础上添加响应面优化模块，选择 6 组参数对主轴箱作响应面优化。实验设计为 Box-Behnken 设计，响应面类型为 Kriging 类型，随后带入参数进行优化。主轴箱的质量从 935.3 kg 变

为了 872 kg，质量减轻了 6.77%，提升了静态特性，减小了主轴箱的质量，实现了轻量化设计。

参考文献：

[1] 蒋莉,马增辉,许新勇,等. 基于 Kriging 模型与遗传算法的月牙肋钢岔管结构优化设计[J]. 华北水利水电大学学报(自然科学版),2026,47(3):123-132.

[2] 杜贝. 龙门加工中心结构与静态特性分析[D]. 西安:陕西科技大学,2016.

[3] 张玥. 高速龙门加工中心电主轴用主轴箱静态特性分析及优化研究[D]. 苏州:苏州大学,2018.

[4] 张子琦. 基于敏感度分析与 Kriging 代理模型的离心风机优化设计[D]. 郑州:郑州轻工业大学,2024.

[5] 陈鹏. 基于 ANSYS 的龙门加工中心主轴箱模态特性分析及优化研究[D]. 成都:西华大学,2019.

[6] 陈国雄,曹阳,吴家雄,等. 焊接平台龙门架结构轻量化设计[J]. 机械设计与制造,2024(4):194-199.

[7] 李有堂,候宗仁. 立式加工中心龙门架的多步复合优化设计[J]. 现代制造工程,2024(11):108-114,123.

[8] 彭伊丽. 切削激励下机床工作空间振动特性分析方法研究[D]. 武汉:华中科技大学,2021. 责任编辑:李云

基于带状宽度约束策略的 Smith-Waterman 算法优化

甘秋云^{1*}, 黄 玮^{2*}, 李华亮^{1,3}, 侯翔宇¹, 赵 佳¹

(1. 福州理工学院 计算与信息科学学院, 福建 福州 350014; 2. 福州工商学院 人工智能学院, 福建 福州 350700;
3. 鼋龙(厦门)生物医药研究院有限公司, 福建 厦门 361000)

摘 要:针对传统 Smith-Waterman 序列比对算法存在的时间和空间复杂度较高的问题,提出一种改进算法。一方面引入带状宽度约束策略,通过减少动态规划矩阵的计算范围,有效降低打分时间,提高比对效率;另一方面,设计动态存储结构,优化得分矩阵,降低空间复杂度。以金黄色葡萄球菌(*Staphylococcus aureus*)、银葡萄球菌(*Staphylococcus argenteus*)序列为测试数据集,设计在不同序列长度、不同带状宽度 k 值条件下的多组比对实验,提出基于序列特征因子的自适应 k 值选择方法,并采用多尺度验证策略确定最优参数 k 。选取 2019 新型冠状病毒(2019-nCoV)、中东呼吸综合征冠状病毒(MERS-CoV)、SARS 冠状病毒(SARS-CoV)及蝙蝠 SARS 样冠状病毒(bat-SL-CoV ZC45)的全基因组序列作为测试数据集,在全序列范围内进行局部比对。实验结果表明:在保证最佳匹配得分和序列相似度与传统算法一致的前提下,改进算法显著降低了比对时间,提升了运算效率。

关键词:Smith-Waterman 算法;序列比对; k 值;带状宽度;打分矩阵;改进算法

中图分类号:TP301.6

文献标识码:A

文章编号:2095-297X(2026)05-0046-11

Optimization of the Smith-Waterman Algorithm Based on a Banded Width Constraint Strategy

GAN Qiuyun^{1*}, HUANG Wei^{2*}, LI Hualiang^{1,3}, HOU Xiangyu¹, ZHAO Jia¹

(1. School of Computer and Information Science, Fuzhou Institute of Technology, Fuzhou, 350014, Fujian;

2. School of Artificial Intelligence, Fuzhou Technology and Business University, Fuzhou, 350700, Fujian;

3. TuoLong (Xiamen) Biomedical Research Institute Co., Ltd, Xiamen, 361000, Fujian)

Abstract: The conventional Smith-Waterman algorithm for local sequence alignment suffers from high computational and memory complexity, limiting its utility in large-scale genomic analyses. To address these limitations, we proposed an optimized approach that integrates two key innovations: a banded alignment strategy that constrains the dynamic programming matrix to a defined diagonal region, significantly reducing computation time, and a dynamic storage structure that minimizes memory requirements by efficiently handling the scoring matrix. We systematically evaluated our method using bacterial genomes from *Staphylococcus aureus* and *Staphylococcus argenteus* under varying sequence lengths and band widths (k -values). This led to the development of an adaptive k -selection method based on sequence characteristics, with optimal parameters determined through multi-scale validation. Further validation was performed using complete genomes of four coronaviruses: 2019-nCoV, MERS-CoV, SARS-CoV and bat-SL-CoV ZC45. Experimental results demonstrated that our optimized algorithm maintains identical alignment accuracy and similarity scores while achieving substantially reduced execution times compared to the standard implementation. These improvements enable more efficient large-scale sequence comparisons without compromi-

收稿日期:2026-01-20 * 通讯联系人

基金项目:福州理工学院校级科研培育计划项目(FTKY2025004);2025 年度福州理工学院校级教学团队建设项目(LGJC2024028)。

作者简介:甘秋云(1986—),女,福建宁德人,副教授,硕士。E-mail:gqy5344@gmiot.com

sing result quality.

Key words: Smith-Waterman algorithm; sequence alignment; k value; band width; scoring matrices; improved algorithm

生物序列比对是生物信息学的重要研究领域之一,它是多个基因序列之间相似性和同源性分析的基础,对揭示生物序列中的结构、功能和进化信息的发现具有重要的意义^[1]。基于动态规划算法的序列比对方法是由 Saul Needleman 和 Christian Wunsch^[2]于 1970 年提出,Needleman-Wunsch 算法适合于序列的全局比对,而在两条较长的序列中寻找最佳匹配的子序列往往更有意义^[3]。因此,在此基础上,1980 年 Temple Smith 和 Michael Waterman 给出了 Smith-Waterman 算法,该算法是基于动态规划算法下实现局部序列相似度的比较,具有较高的灵敏度,奠定了局部序列比对算法的基础。此后出现了多种启发式算法,其中应用广泛的有 FATSA 算法^[4]和 BLAST 算法^[5]。

然而,在实际问题中,Smith-Waterman 局部比对算法虽然结果准确,但是比对时间相对较长,同时面对呈指数级增长的生物数据,序列得分矩阵存储空间也随之增大^[6]。因此,提出一种改进的 Smith-Waterman 算法策略,降低序列比对时间和空间存储成本具有一定的现实意义。

1 基本原理

1.1 算法简介

Smith-Waterman 算法旨在核苷酸序列或蛋白质序列之间发现相似或匹配的局部区域的一种局部序列比对算法^[7]。依据特定打分规则,计算两个序列的最优局部比对得分(打分方法使用了置换矩阵和空位罚分),从而判断序列之间的最高相似性区域。该算法在计算得分时,回溯从打分最高的矩阵单元开始,直到遇到单元打分为零的位置停止,即算法不存在负分值(负分值均以“0”表示),这也是它与 Needleman-Wunsch 算法的主要区别^[8]。

1.2 算法过程

构建一个 $(m+1) \times (n+1)$ 阶的序列矩阵 H ,初始化 H 的第 0 行和第 0 列元素均设为 0。即 $H_{i,0} = H_{0,j} = 0 (i=0,1,\dots,m; j=0,1,\dots,n)$ 。

递推计算 $H_{i,j}$,打分规则如下:

$$H_{i,j} = \max \begin{cases} 0 \\ H_{i-1,j-1} + S(s_i, t_j) \\ H_{i-1,j} + S(s_i, -) \\ H_{i,j-1} + S(-, t_j) \end{cases} \quad (1)$$

其中 0 表示重新启动一个新的比对区域,也是区别于全局比对的部分。

$S(s_i, t_j)$ 是序列碱基匹配得分, W_1 和 W_2 分别是得分值,即:

$$S(s_i, t_j) = \begin{cases} W_1, s_i = t_j; \\ -W_2, s_i \neq t_j; \\ -W_2, s_i = - \text{ 或 } t_j = - \end{cases} \quad (2)$$

其中空位罚分使用连续罚分,单个空位罚分为 d ,空位的延伸罚分为 e ,则 n 个空位的罚分 $W_2 = d + (n-1) \times e$ ^[9]。

回溯寻找最佳比对路径。从矩阵 H 中得分最高的单元开始,根据得分的来源回溯至上一位置,如此反复直至遇到得分为 0 的单元,终止该条路径。通过回溯方法,寻找出两条序列的最优比对结果^[10]。

若两条序列分别为 $s = \text{GGTTGACTA}$, $t = \text{TGT-TACGG}$ 。使用局部比对算法进行打分,得到的打分矩阵及回溯路径如图 1(a) 所示,比对结果如图 1(b) 所示。

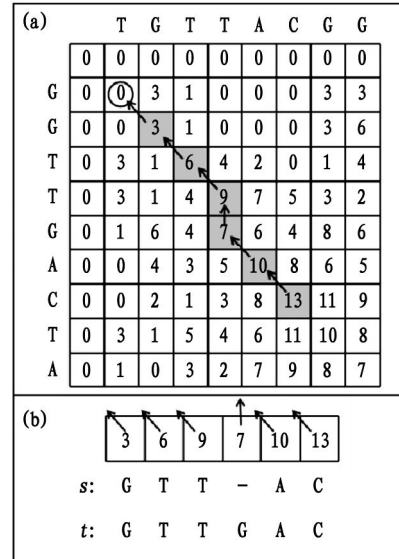


图1 打分矩阵及比对结果

2 算法改进

虽然 Smith-Waterman 算法可以有效求解两条序列中的高相似区域,但时间开销和空间开销过大。根据 Smith-Waterman 算法的实现过程,可以发现,

计算得分以及填充矩阵所需要的时间开销和空间开销均为 $O(m \times n)^{[11]}$ 。因此,研究基于带状宽度约束策略和动态存储结构设计对算法时间开销和空间开销进行优化^[12]。

2.1 基于带状宽度约束策略

在原有的算法中,针对两条长序列需要将序列中的每个碱基依次使用打分方式进行计算得分,算法时间相对较长。而在实际中发现,高度同源的两条序列在比对时最佳匹配的路径是非常靠近主对角线区域的,因此,基于主对角线的带状比对方法就是依据该事实进行优化改进^[13]。

假设两条 DNA 序列分别为 s 和 t , 长度分别为 m 和 n , 带状比对方法是以主对角线为主轴中心, 设置宽度为 k 的区域进行打分计算, 而区域外的元素则不参与打分直接设置为 0 即可。其中 k 满足 $m - n - k \leq i - j \leq k$, 因此, 在宽度为 k 的带状区域内计算的矩阵单元数量较所有单元数量明显下降, 时间复杂度从传统的 $O(m \times n)$, 变为 $O(k \times \min(m, n))$ 。图 2 是两条序列在阴影区域宽度 $k = 2$ 时的带状比对示意图, 实线表示跨出带状区域的比对路径。

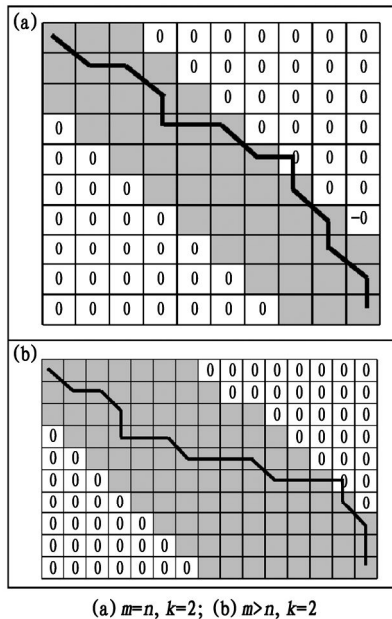


图 2 两条序列的带状比对示意图

由于带状宽度 k 值对序列比对结果具有一定的影响, 研究中的改进算法基于序列特征因子分析提供推荐的 k 值, 当序列较大时, 改进对 k 值的调整, 在推荐 k 值的基础上进行多尺度验证, 最终选择合适的带状宽度。

(1) 基于序列特征因子自适应 k 值。为适应不同序列的生物学特性, 提出一种基于序列特征因子的自适应 k 值选择方法。基础 k 值设置为两条序列中较短序列的长度 $\min_len$, 同时通过序列相似度因子、GC 含量因子、序列复杂因子和保守性因子进行调整, 调整后的 k 值范围为 $5 < k \leq 0.5 \times \min_len$ 。

序列相似度因子 (F_{sim}): 反映两条序列在全局范围内的匹配程度。通过传统 Smith-Waterman 算法对序列进行初步比对, 计算匹配碱基数占总长度的比例。若匹配比例 > 0.8 , 说明序列高度相似, 带状区域可适当缩小, 因子取 0.6; 若匹配比例 < 0.4 , 说明序列差异较大, 需扩大搜索范围, 因子取 1.2; 其余情况可取 0.8 - 1.0 范围值。

GC 含量因子 (F_{GC}): 衡量两条序列中 G 和 C 碱基所占比重的差异。设两条序列的 GC 含量分别为 GC_1 和 GC_2 , 差异度 $\Delta GC = |GC_1 - GC_2|$ 。若 $\Delta GC < 0.05$, 说明序列碱基组成相似, 因子取 0.9; 若 $\Delta GC > 0.15$, 说明组成存在显著差异, 因子取 1.1; 其余情况取 1.0。

序列复杂因子 (F_{comp}): 反映序列中短片段组成的多样性, 通过评估序列中短片段 (本次实验使用 3 个碱基的片段, 即 3-mer) 的分布来量化序列的复杂性。若 $H > 0.7$, 说明序列复杂度高, 可能包含较多变异区域, 需扩大搜索范围, 因子取 1.1; 若 $H < 0.3$, 说明序列重复性高, 因子取 0.9。

保守性因子 (F_{cons}): 综合 GC 含量差异与序列复杂度, 估计序列的保守性。若两条序列 GC 含量相近且复杂度较低, 说明序列保守性高, 因子取 0.8; 若 GC 含量差异大且复杂度高, 说明保守性低, 因子取 1.2。

最终推荐的初始 k 值计算如下:

$$k = \min(\min_len \times \frac{F_{sim} \times F_{GC} \times F_{comp} \times F_{cons}}{z}, 0.5 \times \min_len)$$

其中 z 为缩放因子, 一般取 5 - 10, 用于控制 k 的数量级。各因子设置如表 1 所示。

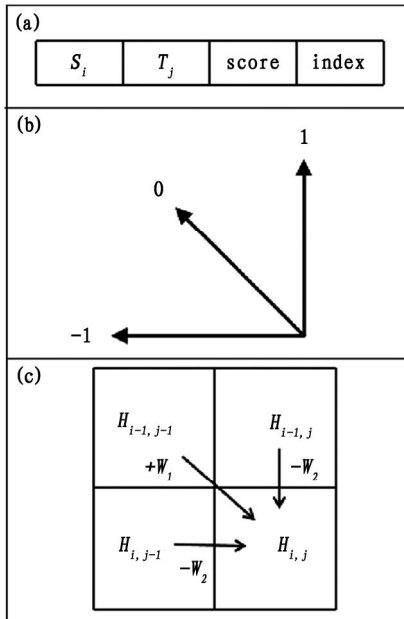
表 1 多种序列统计特征与因子关系

统计特征	相似度				GC 含量		序列复杂性		保守性	
	> 0.8	$0.6 - 0.8$	$0.4 - 0.6$	< 0.4	< 0.05	> 0.15	> 0.7	< 0.3	> 0.7	< 0.3
因子	0.6	0.8	1.0	1.2	0.9	1.1	1.1	0.9	0.8	1.2

(2)多尺度验证 k 值。基于方法(1)得到的两条序列的特征计算初始推荐的 k 值,根据推荐值,选取不同大小的 k 值进行多次测试验证,本实验中采用 $0.5k$ 、 k (由方法(1)推荐得到的 k 值)、 $1.5k$ 和 $2k$ 值进行测试。以传统算法得到的结果作为参考,分别使用候选 k 值进行改进后算法实现,将结果与参考结果比较并计算得分比例,将得分比例最高的 k 值做为最终推荐的 k 值。同时,在计算过程中,根据序列长度差异自动调整带状区域的中心,对选择得到的 k 值进行精度验证,若 k 值得分比例达到 98% 以上,则提前终止测试,节省程序运行时间。为了确保改进算法的可靠性,比对最佳匹配得分和相似度误差控制在 2% 以内。

2.2 动态存储结构设计

打分矩阵填充结束后,需要从矩阵中找到得分最大的单元进行回溯,直至遇到得分为 0 的即可终止。为了节省回溯的时间,在计算填充的过程中,实时记录当前最大的得分单元,每个单元不仅包含得分值,同时包含得分单元位置、分值源等信息。打分单元 $H_{i,j}$ 的结构示意图如图 3(a) 所示,其中, S_i 和 T_j 分别表示该得分单元在两条序列上的碱基位置, $score$ 记录 $H_{i,j}$ 的得分值, $index$ 表示得分源,判断单元得分是否来自竖直向上、对角线向左和水平向左三个方向,分别以 1, 0, -1 表示,如图 3(b) 所示,得分示意图如图 3(c) 所示。



(a) 打分单元存储结构示意图;
(b) 得分源位置标记 $index$; (c) 得分示意图

图 3 得分单元结构示意图

在图 1 中, $H_{7,6}$ 单元表示计算来自两条序列中 S_7 和 T_6 位置 (此时碱基元素均为“C”) 的两个碱基匹配得分, $score$ 得分为 13, $index$ 为 0, 表示当前得分来自对角线方向, 即得分来自 $H_{6,5}$ 单元。

原有的算法需要的存储空间为 $O(m \times n)$, 随着序列长度的增加, 对计算机内存空间要求越高^[14]。由于带状比对方法只要考虑主对角线附近宽度为 k 的区域, 区域外的无需保存, 从而大大降低存储开销。因此, 在存储各个碱基结点和打分单元时, 本研究设计为动态存储结构。

动态存储得分矩阵主要对两条序列进行存储, 每个碱基单元包含碱基符号和所在序列下标位置, 此时需要开辟 $O(2m + 2n)$ 的空间开销。而各个打分单元结构包含碱基位置, 得分值和得分来源 (见图 3(a)), 在计算第 i 行 (或第 j 列) 元素时, 只需要提供第 $i-1$ 行 (或第 $j-1$ 列) 值即可, 其中 S_i 和 T_j 分别链接序列中碱基位置, 打分矩阵动态存储结构示意图如图 4 所示。从理论上, 空间存储容量为两条序列存储空间 $2m + 2n$ ^[15], 各个打分单元在设置带状宽度为 k 的条件下所需的存储容量为 kn , 总的算法空间复杂度为 $O(2m + 2n + kn)$, 当两条序列长度较为接近时, 算法空间复杂度为 $O(n)$ 。

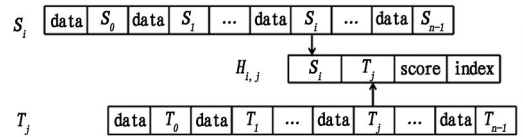


图 4 打分矩阵动态存储结构示意图

3 实验结果与分析

3.1 序列比对结果测试

为了确保改进后的算法在降低比对时间的同时, 得到最优的匹配得分且比对结果准确, 实验中随机选取序列长度为 500 bp 的两条序列进行算法测试, 传统的 Smith-Waterman 算法需要填充整个动态规划矩阵, 改进后的算法设置带状宽度 $k = \lceil L/10 \rceil$ 。本实验所采用的替换矩阵参数为: $match = 5$, $mis-match = -4$, 空位的延伸罚分为 0.5^[16]。

测试结果如图 5 所示, 结果中可以看到算法改进前后局部比对最佳匹配得分均为 414, 序列相似度为 56.12%, 改进前后得分一致, 相似度差异为 0, 验证了改进后算法的准确性。同时, 传统的算法运行时间为 5 ms, 改进后的算法时间为 1 ms, 说明改进后的 Smith-Waterman 算法可以降低比对时间, 提高效率。

```

=== Smith-Waterman序列比对算法测试系统 ===
请选择序列输入方式:
1. 手动输入序列
2. 随机生成序列
3. 从文件读取序列(文件要和程序在同一个目录下)
请选择 (1-3): 2
请输入第一条序列的长度 (bp): 500
请输入第二条序列的长度 (bp): 500
正在生成两条随机DNA序列...
序列1长度: 500 bp
序列2长度: 500 bp
生成的序列1 (前50个碱基): TTCCAATTGCTTACTATTAACTTAAAGACTCGAGGTGATGGGGCGCAT
生成的序列2 (前50个碱基): CCTGAATAGACGAGCTGCTTCTACTGCTAGGATGCTGAAATCCCGGCA
请输入改进算法的带状宽度k值 (推荐值: 50): 50
改进算法带状宽度: 50
开始序列比对...
正在运行传统Smith-Waterman算法... 完成
正在运行改进的Smith-Waterman算法... 完成
=== 传统Smith-Waterman算法结果 ===
最佳匹配得分: 414
序列相似度: 56.12%
比对长度: 531
执行时间: 5.00 ms
=== 改进的Smith-Waterman算法结果 ===
最佳匹配得分: 414
序列相似度: 56.12%
比对长度: 531
执行时间: 1.00 ms
=== 算法性能比较 ===


| 算法类型   | 执行时间    | 最佳匹配得分 | 序列相似度  |
|--------|---------|--------|--------|
| 传统SW算法 | 5.00 ms | 414    | 56.12% |
| 改进SW算法 | 1.00 ms | 414    | 56.12% |


性能改进:
- 时间减少: 80.00%
- 带状宽度: 50
- 得分一致性: 是
- 相似度差异: 0.00%

```

图5 改进前后算法测试结果 ($L=500\text{ bp}$)表2 改进前后算法关于5组长度范围不同的序列比对结果 ($k=L/10$)

实验组	subject	query	序列一致性/%		比对时间/ms		最佳匹配得分/分	
			改进前	改进后	改进前	改进后	改进前	改进后
1	S-400	Q-416	55.30	55.30	4	1	314	314
2	S-500	Q-527	54.69	54.69	5	1	386	386
3	S-800	Q-792	55.74	55.74	12	2	669	669
4	S-1000	Q-984	54.09	54.09	17	4	694	694
5	S-1500	Q-1485	55.40	55.40	35	10	1207	1207

从表2中可以发现,在带状宽度为 $[L/10]$ 的情况下,5组实验结果中,在最佳匹配得分不变的情况下,改进后的 Smith-Waterman 算法在比对时间上普遍低于改进前的时间,5组实验的序列一致性分别为 55.30%、54.69%、55.74%、54.09% 和 55.40%,改进前后完全一致,平均差异为 0%。图6是5组长度不相同的序列,在算法改进前后序列比对所耗用的时间。从图中可以发现,使用改进后的算法进行序列比对,时间消耗低于改进前的算法,随着序列长度的增加,比对消耗的时间差距越明显,传统算法的执行时间从 4 ms 到 35 ms,呈近似线性增长趋势,而改进算法执行时间从 1 ms 到 10 ms,增长幅度更为平缓,5组实验的平均加速比约为 3.67 倍,表明改进的算法具有更好的可扩展性。

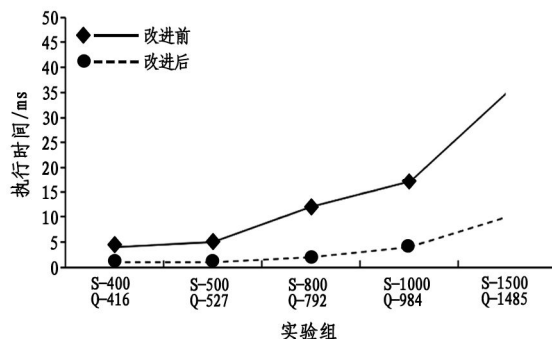
在上述实验中,选取的带状宽度 k 为 $[L/10]$,随着序列长度 L 的增大,带状宽度 k 值有可能影响实际比对区域,若主对角线带状区域过小,将导致实际非 0 的得分被过滤而终止路径。因此,在保证带

3.2 算法时间对比

为了验证在保证比对得分最优,比对结果准确的情况下,随着序列长度增加,改进后算法可以有效地降低算法运行时间,本研究设计了序列长度范围在 400 – 1500 bp 的 5 组不同的对比实验。

为了仿真序列中碱基的插入缺失情况,设计程序随机生成 5 条目标序列,长度 L 分别为 400、500、800、1000、1500 bp。再随机生成 5 条长度在“目标序列 $\pm 50\text{ bp}$ ”不等的待比对序列。同时,在本次实验中,设置带状宽度 $k = [L/10]$ (k 值为待测序列长度 $\div 10$ 后的取整结果)。为了计算两条序列的一致性、比对时间以及最佳匹配得分,分别采用改进前和改进后的算法,依次进行 5 组实验,结果如表 2 所示,其中目标序列和比对序列分别使用 subject 和 query 表示。如 S-400 表示目标序列长度为 400 bp, Q-416 表示待比对的序列长度为 416 bp。

状比对方法可以有效减小算法比对时间的情况下,继续测试本实验中不同的带状宽度 k 对序列比对结果是否存在实质性影响。

图6 算法改进前后5组长度不同的序列比对消耗的时间 ($k=L/10$)

3.3 带状宽度 k 值对对比结果的影响

实验中,从 NCBI 下载具有较高同源性的 *Staphylococcus aureus* (金黄色葡萄球菌,序列大小约 2.82

Mb, *GCF_000013425.1*) 和 *Staphylococcus argenteus* (银葡萄球菌, 序列大小约为 2.76 *Mb*, *GCF_000236925.1*) 分别为参考序列和待比对序列。设计程序, 从序列文件中分别读取 5 组不同长度范围的序列, 每组序列碱基起始位点默认都从 1 开始, 序列长度范围分别在 (500、1000、1500、2500、5000) *bp* ±

50 *bp* 的序列, 分别计算在不同的带状宽度 k 值下, 算法改进前后比对得分、用时、序列相似性等情况, 结果如表 3 所示。其中 S-456 表示 *Staphylococcus aureus* (金黄色葡萄球菌) 序列长度为 456 *bp*, Q-482 表示 *Staphylococcus argenteus* (银葡萄球菌) 序列长度为 482 *bp*。

表 3 不同长度范围、不同带状宽度 k 值下序列比对结果

实验组	subject	query	带状宽度	传统 SW 算法			改进 SW 算法		
				执行时间 /ms	最佳匹配 得分/分	序列相 似度/%	执行时间 /ms	最佳匹配 得分/分	序列相 似度/%
1	S-445	Q-451	$k = L/5$	5	1663	85.25	2	1663	85.25
			$k = L/10$	5	1663	85.25	1	1663	85.25
			$k = L/15$	4	1663	85.25	1	1663	85.25
			$k = L/20$	4	1663	85.25	0	1663	85.25
2	S-995	Q-1012	$k = L/5$	17	4161	90.50	6	4161	90.50
			$k = L/10$	14	4161	90.50	4	4161	90.50
			$k = L/15$	14	4161	90.50	2	4161	90.50
			$k = L/20$	15	4161	90.50	1	4161	90.50
3	S-1510	Q-1495	$k = L/5$	33	6507	92.58	15	6507	92.58
			$k = L/10$	33	6507	92.58	8	6507	92.58
			$k = L/15$	31	6507	92.58	5	6507	92.58
			$k = L/20$	32	6507	92.58	4	6507	92.58
4	S-2510	Q-2535	$k = L/5$	87	10939	92.56	43	10939	92.56
			$k = L/10$	89	10939	92.56	22	10939	92.56
			$k = L/15$	88	10939	92.56	16	10939	92.56
			$k = L/20$	88	10939	92.56	12	10939	92.56
5	S-4980	Q-5035	$k = L/5$	339	21640	92.46	171	21640	92.46
			$k = L/10$	340	21640	92.46	87	21640	92.46
			$k = L/15$	338	21640	92.46	60	21640	92.46
			$k = L/20$	341	21640	92.46	43	21640	92.46

改进前的序列比对得分考虑的是 $k = L$ 的情况 (即每个碱基都参与打分), 因此能够保证当前两条序列比对得到的是最佳匹配结果得分。从表 3 中可以发现, 5 组不同长度范围的两条序列在不同带状宽度 k 值下, 改进算法在最佳匹配得分和序列相似度与传统算法一致, 说明基于带状宽度的改进策略能在降低比对时间的同时, 保持比对结果的准确性。

从结果中可以发现, 改进算法的执行时间在不同序列长度与带状宽度下均显著低于传统算法。在第 4 组实验 (S-2510、Q-2535) 中, 传统算法耗时为 87-89 ms, 改进算法在 $k = L/20$ 时仅需 12 ms,

时间降低率达 86.2%。在同一组实验中, 随着带状宽度 k 值的减小, 执行时间进一步降低, 在第 3 组实验 (S-1510、Q-1495) 中, 当 $k = L/5$ 时, 改进时间为 15 ms, 时间降低 54.5%, 当 $k = L/20$ 时, 改进时间为 4 ms, 时间降低 87.9%, 可以得到带状宽度与执行时间呈现负相关性, 表明更小的带状宽度具有更高的计算效率, 同时也验证了算法复杂度从 $O(m \times n)$ 到 $O(k \times \min(m, n))$ 理论的合理性。

范围在 500 *bp* 的第 1 组实验中, 当 $k = L/20$ 时, 改进算法的执行时间为 0 ms, 主要是由于序列长度过短, 算法执行时间低于系统计算精度而导致

的,并非实际耗时为 0 ms。在较短序列长度中,改进算法时间降低平均为 75% - 93.3%,在较长序列中,改进算法时间降低平均为 74.3% - 87.3%。图 7 是不同序列长度范围、不同带状宽度 k 值下的序

列比对结果图,从图中可以明显看到,随着序列长度从 500 bp 增长至 5000 bp,传统算法的执行时间从 5 ms 急剧增加到 340 ms 左右,增长了约 68 倍。

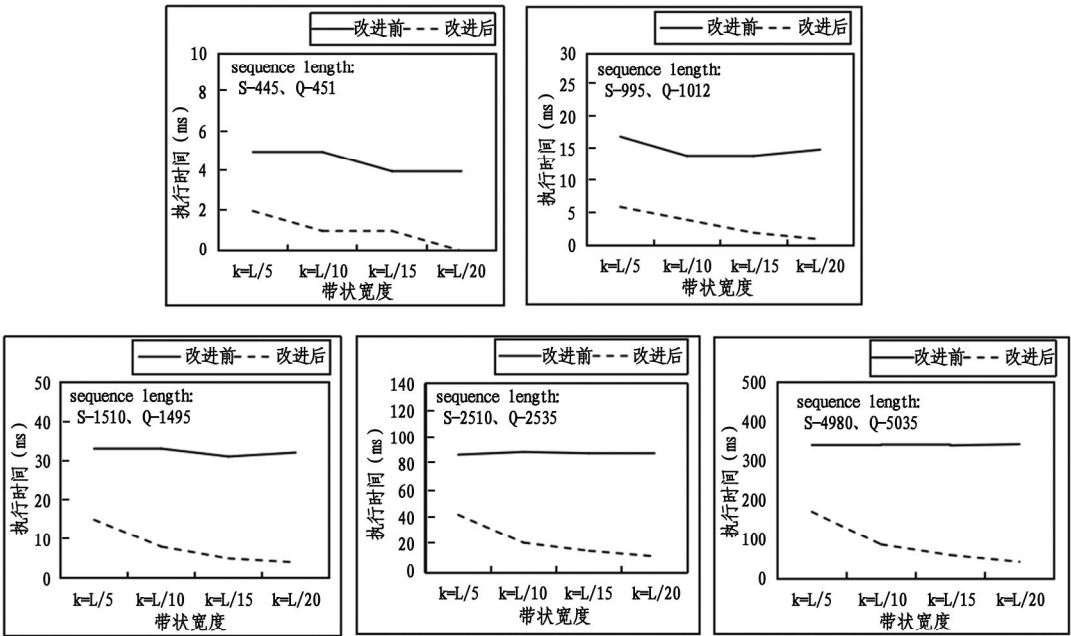


图 7 不同长度范围、不同带状宽度 k 值下序列比对结果图

为了进一步测试带状宽度 k 值对比对结果的影响,设计 5 组实验,设置读取序列的起始位点不同,分别从第 1000、3000、5000、10 000、20 000 个碱基位置开始。序列长度均在 5000 bp $\pm$ 100 bp 范围,在不

同的 k 值条件下,实现序列比对,根据实验结果分析算法改进前后的得分、用时及局部最优相似度,结果如表 4 所示,测试结果如图 8 所示(以 S - 5029、Q - 5302 在 $k = L/20$ 情况下的运行结果为例)。

表 4 不同起始位点、不同宽度 k 值下序列比对结果

实验组	subject	query	起始点	带状宽度	传统 SW 算法			改进 SW 算法		
					执行时间 /ms	最佳匹配得分/分	序列相似度/%	执行时间 /ms	最佳匹配得分/分	序列相似度/%
1	S - 4980	Q - 5110	1000	$k = L/5$	345	21953	93.24	171	21953	93.24
				$k = L/10$	344	21953	93.24	89	21953	93.24
				$k = L/15$	344	21953	93.24	57	21953	93.24
				$k = L/20$	346	21953	93.24	44	21953	93.24
2	S - 5007	Q - 5355	3000	$k = L/5$	364	21793	92.67	179	21793	92.67
				$k = L/10$	361	21793	92.67	88	21793	92.67
				$k = L/15$	363	21793	92.67	57	21793	92.67
				$k = L/20$	358	21793	92.67	47	21793	92.67
3	S - 4850	Q - 5130	5000	$k = L/5$	334	21388	93.32	168	21388	93.32%
				$k = L/10$	337	21388	93.32	84	21388	93.32
				$k = L/15$	337	21388	93.32	58	21388	93.32
				$k = L/20$	337	21388	93.32	42	21388	93.32

续表
表 4 不同起始位点、不同宽度 k 值下序列比对结果

实验组	subject	query	起始 起点	带状 宽度	传统 SW 算法			改进 SW 算法		
					执行时间 /ms	最佳匹配 得分/分	序列相 似度/%	执行时间 /ms	最佳匹配 得分/分	序列相 似度/%
4	S-5011	Q-5410	10000	$k = L/5$	365	19366	87.29	178	19366	87.29
				$k = L/10$	369	19366	87.29	89	19366	87.29
				$k = L/15$	372	19366	87.29	62	19366	87.29
				$k = L/20$	371	19366	87.29	47	19366	87.29
5	S-5029	Q-5302	20000	$k = L/5$	360	21115	91.40	172	21115	91.40
				$k = L/10$	364	21115	91.40	92	21115	91.40
				$k = L/15$	362	21115	91.40	61	21115	91.40
				$k = L/20$	363	21115	91.40	47	21115	91.40

```

=== 传统Smith-Waterman算法结果 ===
最佳匹配得分: 21115
序列相似度: 91.40%
比对长度: 5025
执行时间: 363.00 ms

=== 改进的Smith-Waterman算法结果 ===
最佳匹配得分: 21115
序列相似度: 91.40%
比对长度: 5025
执行时间: 47.00 ms

=== 算法性能比较 ===

```

算法类型	执行时间	最佳匹配得分	序列相似度
传统SW算法	363.00 ms	21115	91.40%
改进SW算法	47.00 ms	21115	91.40%

图 8 不同带状宽度、不同起始位点下
SW 算法改进前后测试结果

从表 4 中可以发现,虽然每组实验的读取序列的起始位置不同,但是同组实验中,不同的带状宽度 k 值下,改进前后的算法得分、相似度仍然一致,差

异性为 0,表明改进后的算法精度是可靠的。在同组实验中,算法执行时间随着 k 值的减小而减小,这是由于 k 值设置越小,主对角线附近区域最小,计算区域越小,运行时间越短。

为了验证算法适用长序列且算法具有良好的扩展性,将参考序列和待比对序列从任意起始位置开始,设计 A ~ F 共 6 组实验,序列长度从 4500 - 30000 bp 依次递增,同一组的序列长度相差 20 - 100 bp,根据计算自适应 k 值和多尺度 k 值验证,推荐最优的带状宽度 k 值,每组实验 3 次取平均结果,实验结果如表 5 所示。其中 S(345;4980)表示参考序列起始位置从第 345 个碱基开始,序列长度为 4980 bp。

表 5 任意起始位置、不同长度下序列比对结果

实验组	subject	query	算法类型	k 值	执行时间/ms	最佳匹配得分/分	序列相似度/%
A	S(345;4980)	Q(350;5024)	传统 SW 算法	-	306	22042	93.47
			改进 SW 算法	709	157	22042	93.47
B	S(1335;4850)	Q(1420;5112)	传统 SW 算法	-	304	20897	92.96
			改进 SW 算法	970	149	20897	92.96
C	S(513;7805)	Q(450;7740)	传统 SW 算法	-	745	33747	93.19
			改进 SW 算法	1032	263	33747	93.19
D	S(645;8980)	Q(718;9024)	传统 SW 算法	-	995	39126	93.12
			改进 SW 算法	1196	350	39126	93.12
E	S(7806;20980)	Q(8018;21016)	传统 SW 算法	-	5324	85806	90.17
			改进 SW 算法	2098	1464	85806	90.17
F	S(4507;34210)	Q(4485;34176)	传统 SW 算法	-	14283	130482	86.38
			改进 SW 算法	3416	3956	130482	86.38

从表 5 中可以发现,任意位置起始,序列长度设置一定范围,使用自适应推荐 k 值做进一步的多尺度验证,计算获取最终使用的带状宽度 k 值,可以提高序列比对的可靠性和准确性,在同一条件下(k 值使用自适应和多尺度验证推荐得到),算法执行时间随着序列长度的增加而增加,不同性能环境下执行时间有所差异,但是结果分析中可以得出改进算法的执行时间增长曲线更为平缓,表明改进算法能有效地适用于序列长度的增长,具有良好的扩展性,如图 9 所示。

以 B 组实验(S(1335;4850)、Q(1420;5112))为例,测试结果如图 10 所示,由分析结果可得,序列分析结果中 GC 含量比例为 33.08%,自适应推荐 k 值为 691,为了保证 k 值选取的合理性与最优,进行多尺度 k 值验证,最终选择的带状宽度 k 值为 970,从测试结果可知,当前比对时间减少 52.96%,最佳匹配得分和相似度一致。多次测试发现,序列长度 < 10000 bp,推荐使用自适应的 k 值,若序列长度 > 10000 bp,建议在推荐 k 值基础上进行多尺度验证计算 k 值。

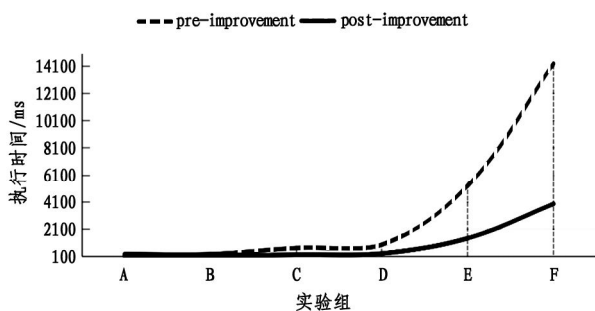


图9 基于动态调整 k 值下改进前后 SW 算法执行时间对比

由上可得改进后的 Smith-Waterman 算法在保证最佳比对结果下,可以有效地降低序列比对时间。

3.4 算法测试

3.4.1 相似度序列(>60%)的算法性能测试

为了评估本研究所提改进算法的实际应用的有效性,实验选取了 2019-nCov(2019 新型冠状病毒)的基因组序列^[17],分别与 MERS-CoV(中东呼吸综合征冠状病毒)、SARS-CoV(SARS 病毒),以及 bat-

SL-CoV ZC45(蝙蝠 SARS 样冠状病毒)进行两两局部序列比对。每组实验重复 5 次并取平均值,实验结果记录了比对结果的局部相似性、最佳匹配得分以及算法的运行时间,并与改进前的算法进行对比,实验结果如表 6 所示。其中,2019-nCov 新型冠状病毒、MERS-CoV 病毒、SARS 病毒和蝙蝠 SARS 样冠状病毒从 NCBI 下载(<https://www.ncbi.nlm.nih.gov/>)。



图 10 基于动态调整 k 值下任意起始位置、不同长度下序列比对测试结果图

表 6 改进前后 Smith-Waterman 算法关于不同病毒序列比对结果

实验组		序列局部相似性/%		最佳匹配得分/分		比对时间/ms	
		改进前	改进后	改进前	改进后	改进前	改进后
2019-nCov	MERS-CoV	61.19	61.19	43572	43572	11157	3196
2019-nCov	SARS-CoV	80.48	80.48	97758	97758	10827	3911
2019-nCov	bat-SL-CoV ZC45	88.15	88.15	118043	118043	10803	3037

从表 6 中可知,在 2019-nCoV 与 SARS-CoV、2019-nCoV 与 bat-SL-CoV ZC45 的比对中,改进后算法所得到的序列局部相似性、匹配最佳得分均与改进前完全一致,算法执行时间降低了 64% - 71.9%,这一结果证实了,改进算法在实际应用中是可信的,特别在长序列比对中,时间效率更为显著,未因性能提升而存在较大误差。

从表 6 结果分析可以看到,三组比对结果呈现的相似性梯度约为 61%、80%、88% 左右,2019-nCoV 与 MERS-CoV 差异最大,仅为 61.19%。有研究表明,RdRp 作为病毒复制关键酶,在 β 冠状病毒属内保持较高序列保守性,即使在远缘冠状病毒间也维持 60% - 65% 的相似度,特别是催化活性中心的 GDD 基序及其周边区域^[18-19]。因此,可以推断该局部最优相似区域可能位于 RNA 依赖性 RNA 聚合酶(RdRp)的保守核心区域。

2019-nCoV 与 SARS-CoV 亲缘关系较近,局部相似比为 80.48%,最优局部区域可能涵盖多个高保守功能区,包括膜蛋白跨膜区域,刺突蛋白非受体结合区域等,有研究结果显示,2019-nCoV 与 SARS-CoV

在复制酶复合体基因的相似度达 86%,与本实验结果吻合^[20]。

局部相似性最高的是蝙蝠冠状病毒 bat-SL-CoV ZC45,表明二者可能存在直接的进化关联,局部最优区域可能覆盖基因组的大片段连续区域,可能包括结构蛋白的编码基因、基因组 5' 和 3' 端的保守非翻译区等,这为新冠病毒的动物起源假说提供了有力的数据支持^[21]。

3.4.2 相似度序列(<30%)的算法性能测试

为了检验算法在序列相似性极低情况下的准确性和可靠性,实验选取系统发育距离较远的物种基因组序列片段作为测试数据。实验所用序列均下载自 NCBI 数据库,分别来自人类线粒体基因组(*Homo sapiens*, NC_012920.1)、斑马鱼线粒体基因组(*Danio rerio*, NC_002333.2)以及拟南芥叶绿体基因组(*Arabidopsis thaliana*, NC_000932.1)。通过从上述基因组中提取不同长度的序列片段,构建了三组相似度均低于 30% 的比对实验组,每组实验重复 5 次,记录最佳匹配得分、序列相似度及平均执行时间,结果如表 7 所示。

表 7 低相似度序列(<30%)的算法比对结果

实验组		序列局部相似/%		最佳匹配得分/分		比对时间/ms	
		改进前	改进后	改进前	改进后	改进前	改进后
<i>Homo sapiens</i> (1020 bp)	<i>Danio rerio</i> (1045 bp)	27.2	27.2	412	412	21	11
<i>Homo sapiens</i> (2015 bp)	<i>Arabidopsis thaliana</i> (1988 bp)	21.8	21.8	561	561	64	29
<i>Danio rerio</i> (3050 bp)	<i>Arabidopsis thaliana</i> (3012 bp)	24.3	24.3	935	935	164	67

从表 7 的结果可以看出,在序列相似度低于 30% 的极端情况下,改进后的 Smith-Waterman 算法仍能准确确定最佳的局部比对区域。在三组实验中,改进前后的最佳匹配得分和序列相似度完全一致。这说明,尽管带状宽度约束策略限制了计算区域,但基于自适应与多尺度验证的 k 值计算方法,确保了关键高分区域未被遗漏,从而维持了结果的准确性。

在时间效率上,对于长度约为 1000 bp、2000 bp 和 3000 bp 的低相似度序列,改进后的算法执行时间比传统算法分别降低了约 47.6%、54.7% 和 59.1%。结果表明,即使序列间同源性较低,带状约束策略通过自适应 k 值计算,仍然能够有效聚焦于

潜在的局部保守区域,从而大幅减少了无效计算,提升了计算效率。实验结果进一步验证了改进后的算法在比对低相似度序列时具有一定的普适性。

4 结论

在序列比对问题中,寻找序列最佳匹配的子序列具有实际的生物学意义,针对传统的 Smith-Waterman 算法运行时间和空间存储成本高等问题,提出了一种基于带状宽度为约束条件的优化方法。根据不同的序列长度,探讨不同带状宽度 k 值下的序列匹配结果,分析序列特征和多尺度验证,引入自适应的带状宽度参数 k ,在保证比对准确性的前提下显著降低了算法的时间复杂度。实验结果表明,改进

后的算法在不同序列长度条件下均能有效提升比对效率,较传统算法具有显著的时间性能。

在存储结构方面,提出了一种构建动态存储结构的方法,通过仅保存必要的计算状态,理论上可实现空间复杂度的优化。研究结果发现,带状宽度参数 k 的选取对算法精度及时间性能均具有重要影响,随着序列长度的增大,带状宽度不宜过小,以免过滤实际得分值导致比对结果非最优。当序列相似度较高时,可选择小 k 值,当序列相似度较低时,可选择大 k 值。然而,在实际应用中,并不能完全预估序列的相似度。因此,本研究提出的基于序列特征的多尺度 k 值验证,可以保证效率的同时提升比对的准确性。

虽然改进后的 Smith-Waterman 算法在时间效率和空间存储成本上均实现了优化,但是随着序列长度的增加,存储需求仍呈指数级增长。后续研究将重点对空间复杂度进行深入探究和优化,探索更高效的数据存储设计与计算方法,进一步提高算法处理大规模序列比对的性能。

参考文献:

- [1] 甘秋云. 基于动态规划的 Needleman-Wunsch 双序列比对算法的分析与研究[J]. 计算机工程与科学, 2021, 43(2): 340-346.
- [2] Smith T F, Waterman M S. Identification of common molecular subsequences[J]. Journal of Molecular Biology, 1981, 147(1): 195-197.
- [3] 罗志兵. 基于动态规划的基因双序列比对研究[J]. 现代计算机(专业版), 2017(32): 28-33, 37.
- [4] Lipman D, J Person W R. Rapid and sensitive protein similarity searches[J]. Science, 1985, 227: 1435-1441.
- [5] Altschul S F, Gish W, Miller W, et al. Basic local alignment search tool[J]. Journal of Molecular Biology, 1990, 215(3): 403-410.
- [6] 罗静初. 双序列比对基础和应用实例[J]. 生物信息学, 2023, 21(1): 1-19.
- [7] 陈俊涛, 邹权. 基于 Profile 比对的改进星比对算法[J]. 电子科技大学学报, 2022, 51(3): 353-356.
- [8] 高佳, 徐云. 基于组合 minimizer 种子的泛基因组图序列比对算法[J]. 计算机工程, 2025, 51(8): 53-61.
- [9] 杨波, 陈洋广, 徐胜超. 基于 Spark 云计算的生物基因多序列比对方法[J]. 计算机测量与控制, 2024, 32(7): 274-279, 287.
- [10] 罗静初. EMBOSS 软件包序列分析程序应用实例[J]. 生物信息学, 2021, 19(1): 1-25.
- [11] 张晓玲, 杨健. 动态规划法进行基因双序列比对的探索 and 实现[J]. 大理大学学报, 2019, 4(12): 6-11.
- [12] 袁若兰. 计算机算法在生物信息学中运用对策探究[J]. 数字通信世界, 2024(3): 121-123.
- [13] 石海鹤, 周卫星. 基于动态规划的双序列比对算法构件设计与实现[J]. 计算机研究与发展, 2019, 56(9): 1907-1917.
- [14] 王勇献, 王正华. 生物信息学导论——面向高性能计算的算法与应用[M]. 北京: 清华大学出版社, 2011: 64-68.
- [15] Notredame C, Higgins D G, Heringa J. T-coffee: a novel method for fast and accurate multiple sequence alignment[J]. Journal of Molecular Biology, 2000, 1: 205-217.
- [16] Lassmann T. Kalign 3: multiple sequence alignment of large datasets[J]. Bioinform, 2019, 36(6): 1928-1929.
- [17] Lu Roujian, Zhao Xiang, Li Juan, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding[J]. Lancet (London, England), 2020, 395: 565-574.
- [18] Snijder E J, Bredenbeek P J, Dobbe J C, et al. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage[J]. Journal of Molecular Biology, 2003, 331(5): 991-1004.
- [19] Gorbalenya A E, Enjuanes L, Ziebuhr J, et al. Nidovirales: evolving the largest RNA virus genome[J]. Virus Research, 2006, 117(1): 17-37.
- [20] Zhou Peng, Yang Xinglou, Wang Xianguang, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin[J]. Nature, 2020, 579: 270-273.
- [21] Andersen K G, Rambaut A, Lipkin W I, et al. The proximal origin of SARS-CoV-2[J]. Nature Medicine, 2020, 26(4): 450-452.

责任编辑: 葛圣志

基于 STM32 和 OV2640 的嵌入式图像采集系统设计

侯静云*, 闻国才, 杨 晟, 冯光双

(淮南师范学院 计算机学院, 安徽 淮南 232038)

摘 要:针对传统图像采集系统成本高、体积大、不易移植等问题,设计了一款基于 STM32 和 OV2640 的嵌入式图像采集系统。该系统以 STM32F407ZGT6 芯片为核心,利用图像传感器 OV2640 进行图像采集,并通过 LCD 实时显示。同时,还能够将温湿度异常的图像通过 WIFI 传输模块发送至 PC 机。经测试,该系统能够实时、准确地完成图像的采集、显示以及传输,在移动终端、工业生产等领域具有一定的应用价值。

关键词:STM32 系统;OV2640 系统;图像采集

中图分类号:TP274

文献标识码:A

文章编号:2095-297X(2026)05-0057-06

Design of Embedded Image Acquisition System Based on STM32 and OV2640

HOU Jingyun*, WEN Guocai, YANG Sheng, FENG Guangshuang

(School of Computer Science, Huainan Normal University, Huainan, 232038, Anhui)

Abstract: Aiming at the problems of high cost, large volume and difficult transplantation of traditional image acquisition system, an embedded image acquisition system based on STM32 and OV2640 was designed. The system takes STM32F407ZGT6 chip as the core, uses image sensor OV2640 for image acquisition, and displays it in real time through LCD. At the same time, it can also send the image of abnormal temperature and humidity to the PC through the WIFI transmission module. After testing, the system can complete image acquisition, display and transmission in real time and accurately, and has certain application value in mobile terminals, industrial production and other fields.

Key words: STM32 system; OV2640 system; image acquisition

随着人工智能、图像处理等技术的不断发展,以嵌入式为基础的图像采集装置得到了广泛的应用。目前,常用的图像采集系统大多以高性能的 CPU、FPGA^[1-2] 为主控芯片设计。蒋俊伦采用 FPGA 结合具有强大运算能力的 GPU 设计了一款实时图像采集存储速度高达 437.5 Mbps 的图像采集处理系统,但 FPGA 操作起来流程复杂,应用门槛较高^[3]。此外,还有使用 USB 接口摄像头如 OMAP3530 或 S3C2410(ARM9)等采集图像数据^[4-5]。这些系统虽然能够获得超高分辨率的图像数据,但其单位时间数据传输量较小,无法满足本设计的视频流量的需求。因此,本系统以 STM32F407ZGT6 芯片为核

心,选择 CMOS 图像传感器 OV2640 代替 USB 摄像头进行图像采集来保证 LCD 视频图像的实时性。此外,该系统还可将温湿度异常的图像以 JPG 格式通过 WIFI 传输模块发送到 PC 机,降低了传输过程的功耗,提高了视频监控环境的安全性。

1 系统硬件设计

本研究设计的嵌入式图像采集系统主要包括 STM32F407ZGT6 单片机模块、采集模块、IS62WV51216 存储模块、LCD 显示模块以及 WIFI 传输模块五个模块,总体结构如图 1 所示。采集模块作为本系统的最前端,主要利用模块中的 OV2640 图像传感器和温湿

收稿日期:2024-10-16 * 通讯联系人

基金项目:淮南师范学院重点科学研究项目(2025XJZD008);淮南师范学院智慧课程建设项目(2025hske06)。

作者简介:侯静云(1990—),女,安徽淮南人,讲师,硕士。E-mail:779100671@qq.com

DHT11 度传感器完成图像和温湿度的采集工作。然后,将采集的图像和数据经由 STM32F407ZGT6 单片机模块处理后传送到 LCD 模块显示^[6]。其中,STM32F407ZGT6 存储容量有限,IS62WV51216 芯片作为扩展存储模块用来暂存图像数据。而 WIFI 传输模块主要是实现与 PC 机的通信,将温湿度异常的图像传送给 PC 机进行显示和存储。系统硬件图如图 2 所示。

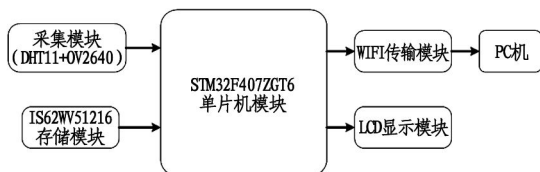
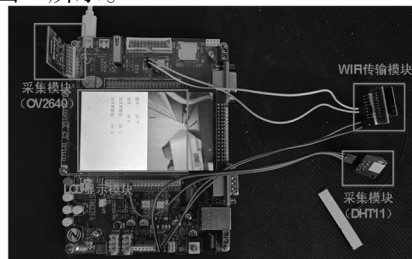


图 1 系统总体结构图

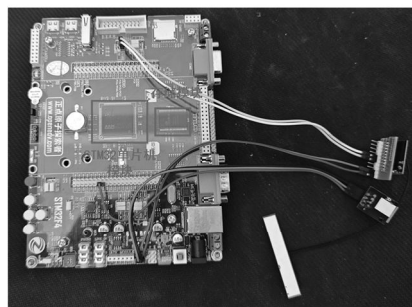
1.1 STM32F407ZGT6 单片机模块

STM32F407ZGT6 是基于 ARM Cortex-M4 处理器的 32 位单片机(MCU),不仅内部拥有 1 MB 的闪存(Flash)、192 kB 的 SRAM、168 MHz 的运行频率以及 12 位的数模转换器,还提供 I²C、SPI、CAN 等

多种接口能够与传感器实现多通道双向数据传输功能,满足本系统对实时数据传输的应用需求,其外围电路如图 3 所示。



(a) 正面



(b) 背面

图 2 系统硬件图

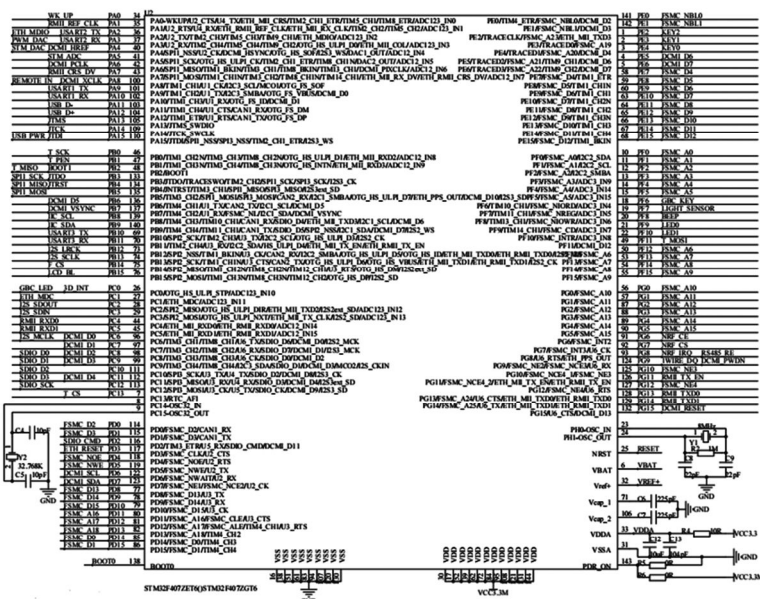


图 3 STM32 芯片电路图

1.2 采集模块

1.2.1 DHT11 传感器模块

本研究采用 DHT11 温湿度传感器监测环境中的温湿度参数。该芯片能够测量 0℃ - 50℃ 范围内的温度和 20% 至 90% RH 范围内的湿度,且误差分别能够控制在 $\pm 1^\circ\text{C}$ 和 $\pm 5\%$ RH,满足了大部分的应用场景。同时,它还具有低功耗、响应时间短、模块化设计方便安装、成本低等一系列优点,在

嵌入式系统设计中被广泛应用。研究中,DHT11 采用单总线通信协议与 STM32 的 PF7 接口相连实现通信,其电路如图 4 所示。

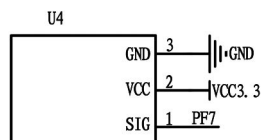


图 4 DHT11 传感器电路图

1.2.2 OV2640 图像采集模块

图像采集模块选用的是 OV (OmniVision) 公司生产的 CMOS 图像传感器 OV2640, 其体积仅为 1/4 寸, 还具有 1.7 ~ 3.3 V 的较低 IO 电压, UXGA (1632 × 1232) 的最大输出格式, 可通过 SCCB 总线以 8/10 位的方式输出 UXGA 格式 200 W 像素的影像数据^[7]。OV2640 图像采集模块主要负责图像捕获任务, 它首先利用内部的光学元件和电子元件将传感器接收到的外部光学信号转换为电信号, 并通过模数转换器(ADC)将其转换为数字图像数据; 然后利用模块内置的图像信号处理器 (ISP) 对原始图像数据进行色彩校正、白平衡调整、曝光控制以及噪声降低等优化处理, 从而提高图像的清晰度; 最后, 处理后的图像数据通过 DCMI 接口传输至 STM32F407ZGT6 主控芯片上, 其电路如图 5 所示。

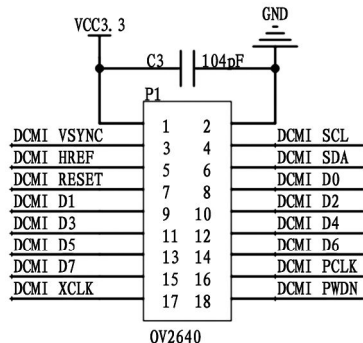


图 5 OV2640 传感器电路图

1.3 WIFI 传输模块

WIFI 传输模块选用的是 USR-C215B 芯片, 该芯片支持 802.11b/g/n 无线通信标准, 为系统提供稳定的无线连接。同时提供 UART、SPI 等串行通信接口, 方便与微控制器或其他控制单元连接。此外, 还支持 TCP/IP 网络协议栈, 可以进行 IP、TCP、UDP、DHCP、DNS 等网络通信。研究中将 WIFI 模块作为 AP 模式使用, 通过 AT 指令将 UART 串口通信信号转换成 WIFI 信号, 实现与上位机的图像数据传输, 其电路如图 6 所示。

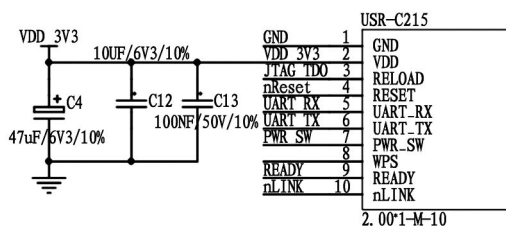


图 6 USR-C215B 芯片电路图

1.4 LCD 显示模块

LCD 显示模块选用的是 ALIENTEK 公司的 4.3

寸 TFT_LCD, 其主要利用 TFT 技术控制液晶分子的排列, 实现对每个像素点的精确控制, 从而使得图像显示更加清晰。微控制器 STM32F407ZGT6 通过 FSMC (可变静态存储控制器) 接口来控制 TFTLCD 的显示, LCD 接口原理图如图 7 所示。在本系统中, TFT-LCD 显示模块主要负责显示 OV2640 图像采集模块捕获到的实时视频图像、DHT11 温湿度传感器模块检测的环境温湿度信息以及用户自定义的温湿度异常报警阈值。

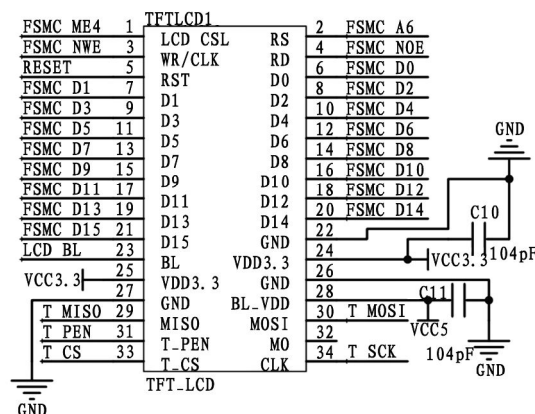


图 7 TFT-LCD 芯片电路图

1.5 IS62WV51216 存储模块

研究所设计的嵌入式图像采集系统需要处理大量的视频数据, 但 STM32F407ZGT6 主控模块自带的存储容量仅有 192 kB, 因此, 扩展了一个拥有 8 M 存储容量的 SRAM 芯片 IS62WV51216。研究中 IS62WV51216 通过 FSMC 总线与微控制器连接, 其接口电路如图 8 所示。

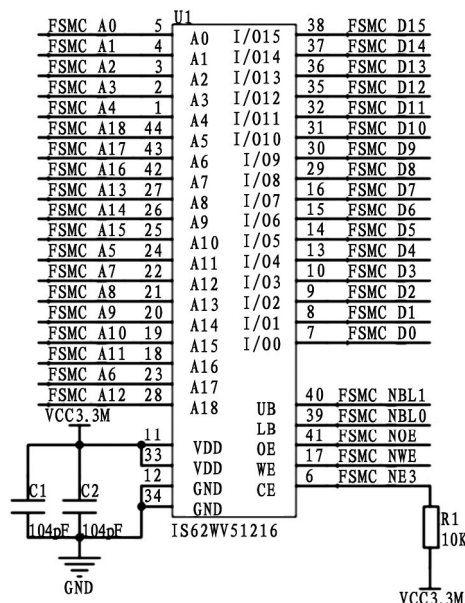


图 8 存储模块电路图

2 系统程序设计

系统程序设计是指利用 Keil uVision5 集成开发环境,用 C 语言对图像采集、图像处理以及图像传输等功能进行程序设计,其流程如图 9 所示。

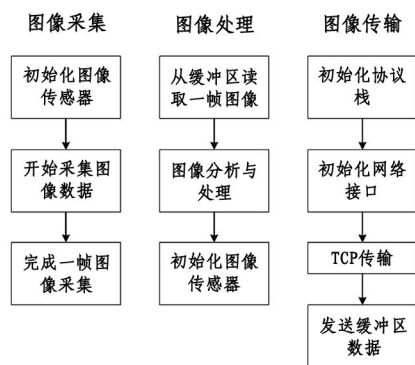


图 9 系统软件设计框图

2.1 OV2640 图像采集程序设计

通过 OV2640_Init() 函数对 OV2640 引脚、时钟的使能、SCCB 总线以及寄存器等进行初始化。如果初始化未完成,在 LCD 屏幕上输出“OV2640 错误!”的信息,否则,在 LCD 屏幕上显示 2 秒的“OV2640 正常”的信息。然后,利用 DCMI 数字摄像接口来驱动图像传感器 OV2640 来采集图像数据,其流程如图 10 所示。OV2640 图像采集的主要代码如下:

```

while( OV2640_Init() )
{
    Show_Str( 30, 190, 240, 16, “OV2640 错误!”,
    16, 0);
    delay_ms( 200 );
    LCD_Fill( 30, 190, 239, 206, WHITE );
    delay_ms( 200 );
}
Show_Str( 30, 190, 200, 24, “OV2640 正常!”,
24, 0);
delay_ms( 2000 );
OV2640_RGB565_Mode();
My_DCMI_Init();
DCMI_DMA_Init( ( u32 ) &LCD - > LCD_RAM,
0, 1, DMA_MemoryDataSize_HalfWord, DMA_Memory-
Inc_Disable );
OV2640_mode = 1;
OV2640_OutSize_Set( lcddev. width, lcddev.
height( 2 ); OV2640_jpg_photo( 0 ) );
DCMI_Start();

```

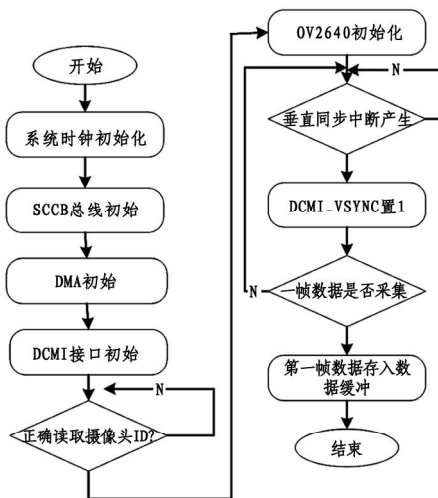


图 10 OV2640 图像采集程序设计流程图

2.2 图像处理软件设计

将图像传感器 OV2640 采集的图像转换为分辨率为 1600×1200 的图像,需要经过以下设置,其过程简图如图 11 所示。

(1)首先,需要通过 OV2640_Window_Set 函数配置寄存器 0X03/0X19/0X1A/0X07/0X17/0X18 来实现传感器开窗大小的设置。

(2)然后,利用 OV2640_ImageSize_Set 函数实现图像尺寸设置,即设置所选格式的输出分辨率 1600×1200 ;根据图像尺寸小于传感器开窗大小的要求,通过 OV2640_ImageWin_Set 函数设置图像开窗大小^[8]。

(3)最后,通过函数 OV2640_OutSize_Set 实现输出到外部设备图像大小的设置,即设置最终在 LCD 和 PC 机上显示的图像分辨率。

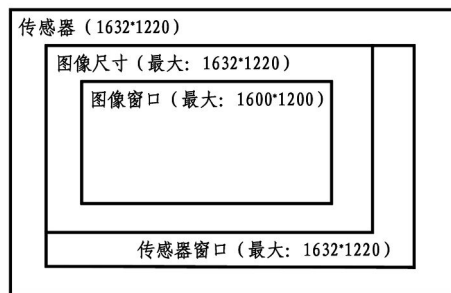


图 11 OV2640 图像窗口设置过程简图

2.3 图像传输软件设计

LCD 显示模块通过 DCMI_Start 函数实现图像传输显示。首先,利用 LCD_SetCursor 函数设置图像在 LCD 屏幕上显示的起始位置,然后调用 LCD_WriteRAM_Prepare 函数让 LCD 做好接收图像数据的准备。接着,调用 DMA_Cmd 函数启用 DMA 传输,允许 DCMI 捕获 OV2640 采集的数据,并直接传

输给 LCD 显示屏,其主要代码如下:

```
Void DCMI_Start( void)
{
    LCD_SetCursor(0,0);
    LCD_WriteRAM_Prepare( );
    DMA_Cmd( DMA2_Stream1, ENABLE);
    DCMI_CaptureCmd( ENABLE);
}
```

当温度或湿度大于设定的阈值时,通过 DMA (直接存储器访问) 传输数据,并通过 USR-C215BWIFI 无线传输模块传输将接收到的数据传输到 PC 机上进行显示和存储,主要代码如下:

```
for(j=0;j<jpeg_data_len*4-i;j++)
{
    While( USART1 -> SR&0X40) == RESET);
    USART1 -> DR = pbuf[j];
    Key = KEY_Scan(0);
    If( key) break;
}
```

3 系统测试

3.1 硬件测试

打开系统后,检测 OV2640 图像采集模块是否连接,并测试 OV2640 图像采集模块是否能够正确采集图像信息并传输到 LCD 屏幕上实时显示,如图 12 所示。

启动嵌入式图像采集系统后,系统会自动执行其预设的监测任务,检测当前环境温度与湿度数据并在 LCD 屏幕上显示。当用户手指轻轻放置在与系统相连的 DHT11 温湿度检测模块的感应区域上,系统测得当前湿度为 77%,大于系统预设的湿度报警值 70%,如图 13 所示。

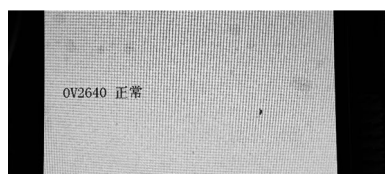


图 12 OV2640 图像采集模块连接检测

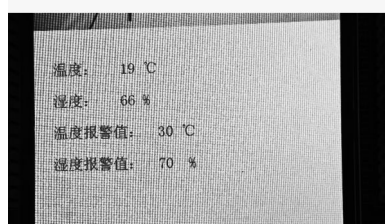


图 13 OV2640 图像采集模块实时显示

当采用上述用手指模拟温度异常情况时,系统就会触发 UART 中断,通过 USR-C215BWIFI 传输模块将当前图像数据传输到上位机中的 XCAM 软件中显示出来,如图 14 所示。同时 XCAM 软件将图像(分辨率 640×480)数据统一保存到事先在电脑上设置的指定的文件夹中。经测试,输出的图像数量每秒可达 23 帧,过程比较流畅。



图 14 温湿度出现异常传输图像

3.2 数据采集模块功能测试

将数据采集模块采集的温度、湿度与手持式温湿度仪进行对比,如表 1 所示。可以看出,误差基本符合设计要求,能够完成对温湿度准确测量的任务。

表 1 温湿度测试对比

数据采集模块		温湿度仪	
温度/°C	湿度/% RH	温度/°C	湿度/% RH
36	45	36	45
30	40	31	41
32	40	31	42
27	35	25	35

3.3 图像显示测试

将图像采集终端置于不同距离(5、10、20、30 m)对同一画面进行录像,通过记录画面帧数来判断图像采集的延迟时间。同一分辨率(640×480)下延迟测试对比结果如表 2 所示。

表 2 延迟测试对比

距离/m	延迟/s
5	<0.6
10	<0.85
20	<1
30	<1.5

由表 2 可知,在不同的传输距离条件下,系统传输的时间延迟都很小,可以很好地实现图像采集。

4 结论

以 STM32 为核心处理器设计了一款嵌入式图像采集系统,该系统不仅可以通过 CMOS 图像传感器 OV2640 使图像输出较为流畅,还将温湿度异常的图像数据以 JPG 格式传输来降低传输功耗。此外,还采用 WIFI 无线传输方式实现与 PC 端的数据传输,减轻布线的工作量。本系统整体成本低、功耗低,在家庭安防、无人或高危工业生产等领域具有一定的应用价值。

参考文献:

- [1] 余国良. 嵌入式图像处理系统中 FPGA 重构技术研究[D]. 武汉:华中科技大学,2015.
- [2] 吴铭,黄国宁,汪保祥,等. 基于 FPGA 的多路视频采集及 AI 加速[J]. 南昌大学学报(工科版),2024,46(3):386-394.

(上接第 39 页)

参考文献:

- [1] 曾韬. 螺旋锥齿轮设计与加工[M]. 哈尔滨:哈尔滨工业大学出版社,1989:11-15.
- [2] 郑昌启. 弧齿锥齿轮和准双曲面齿轮:啮合原理,齿坯设计,加工调整和齿面分析计算原理[M]. 北京:机械工业出版社,1988:26-28.
- [3] 张宇,严宏志,曾韬. 弧齿锥齿轮双重螺旋法切齿原理及齿面接触分析研究[J]. 机械工程学报,2015,51(21):15-23.
- [4] 张宇,严宏志,曾韬. 螺旋运动及刀具参数对双重螺旋法加工齿面特征的影响规律[J]. 机械工程学报,2018,54(5):53-61.
- [5] 吴顺兴. 双重螺旋法加工重载螺旋锥齿轮齿面修正与修形方法研究[D]. 长沙:中南大学,2023:34-37.
- [6] WU S X, YAN H Z, WANG Z Y. Tooth surface error correction of hypoid gears machined by duplex helical method[J]. Journal of Central South University, 2021, 28(5):1402-1411.
- [7] 唐进元,曹康,李国顺. 机床调整参数误差对小轮齿面误差影响规律的理论研究[J]. 机械工程学报,2010,46(17):179-185.
- [8] 唐进元,刘艳平,蒲太平. 基于五轴联动数控铣齿机的

- [3] 蒋俊伦,丰大强,徐新瑞,等. 基于 FPGA + GPU 的图像采集处理系统设计[J]. 计算机测量与控制,2023,31(8):273-279,305.
- [4] Liu Yonglin, Liang Ying, Cheng Yaoyu. Design and implementation of image acquisition system based on ARM and Linux[C]//Proceedings of 2011 International Conference on Electronics and Optoelectronics, IEEE, 2011, 3:49-52.
- [5] 袁姜红,陈二阳,蒋毅. 基于嵌入式的车载视频监控防盗系统的设计与实现[J]. 安阳师范学院学报,2019,21(5):50-52.
- [6] 敖利丞,孟令军,谢宇斌,等. 基于 CH565W 的图像采集系统设计[J]. 集成电路与嵌入式系统,2024,24(4):88-92.
- [7] 付东,张国平,汪馨童,等. 基于 STM32 的低功耗图像采集系统设计[J]. 电子测量技术,2020,43(19):138-143.
- [8] 秦伟,余朝刚. 基于 STM32 的数显轨距尺图像数据采集系统设计[J]. 计算机测量与控制,2018,26(1):173-176.

责任编辑:葛圣志

螺旋锥齿轮刀倾法制造仿真[J]. 北京工业大学学报,2011,37(4):487-493.

- [9] TANG J Y, HU Z H, WU L J. Effect of static transmission error on dynamic responses of spiral bevel gears[J]. Journal of Central South University, 2013, 20(3):640-647.
- [10] 宋碧芸,唐进元,容铠彬,等. 减小螺旋锥齿轮齿面加工误差的参数修正方法[J]. 西安交通大学学报,2022,56(2):101-109.
- [11] 齿轮手册编委会编. 齿轮手册(第二版)(下册)[M]. 北京:机械工业出版社,2001:356-358.
- [12] 谢玲珍,李倩,王忠. 四种渗碳齿轮钢热处理畸变特点研究[J]. 热处理技术与装备,2018,39(2):33-36.
- [13] 张国强. 重载渗碳齿轮钢的疲劳性能研究[D]. 昆明:昆明理工大学,2009.
- [14] 刘赣华,钱锦年,邹洋,等. 淬火冷却介质流速对弧齿锥齿轮渗碳淬火热处理的影响[J]. 材料热处理学报,2018,39(2):152-160.
- [15] KIM N K, BAE K Y. Analysis of deformation in the carburizing-quenching heat treatment of helical gears made of SCM415H steel[J]. International Journal of Precision Engineering and Manufacturing, 2015, 16(1):73-79.
- [16] 侯兴隆. 锥齿轮淬火变形控制的数值模拟研究[D]. 哈尔滨:哈尔滨工程大学,2017.

责任编辑:李壮

基于模态差异感知与双向注意力融合的 RGBT 目标跟踪方法

李鹏程

(蚌埠学院 计算机与信息工程学院,安徽 蚌埠 233030)

摘要:RGBT 目标跟踪因跨模态成像差异显著,以及强光、雾霾、极暗等复杂退化场景下互补信息的有效挖掘不足,导致跟踪系统鲁棒性显著降低。针对上述问题,提出一种融合模态差异感知与双向注意力机制的 RGBT 跟踪算法。在特征提取阶段,构建模态感知增强(MAE)模块,通过通道注意力与空间注意力的协同作用,实现模态特定特征的强化与噪声抑制。在特征融合阶段,设计跨模态双向注意力融合(CMBAF)模块,基于 Transformer 交叉注意力机制建模模态间长距离语义依赖,实现纹理引导的热红外信息补偿与显著性引导的可见光去噪。进一步地,引入基于信息熵的不确定性感知动态门控机制,根据环境可靠性自适应调整模态融合权重,增强模型在模态退化场景下的泛化能力。GTOT、RGBT234 和 LasHeR 三个公开数据集上的实验表明,所提方法在 LasHeR 数据集上取得 67.4% 的精确率和 53.6% 的成功率,较主流方法提升约 2.3%,验证了所提方法的有效性。

关键词:RGBT 目标跟踪;模态感知增强;双向注意力融合;Transformer 交叉注意力机制;不确定性建模

中图分类号:TP391.41

文献标识码:A

文章编号:2095-297X(2026)05-0063-07

Modality-Aware Perception and Bidirectional Attention Fusion for RGBT Object Tracking

LI Pengcheng

(School of Computer and Information Engineering, Bengbu University, Bengbu, 233030, Anhui)

Abstract: RGBT object tracking faces persistent challenges arising from inherent cross-modal imaging discrepancies and insufficient exploitation of complementary information under degraded environmental conditions, such as strong illumination, haze, and extreme low-light scenarios, all of which substantially compromise tracking robustness. To address these limitations, it proposed an RGBT tracking framework that integrates modality-aware perception with bidirectional attention fusion. At the feature extraction stage, a modality-aware enhancement(MAE) module was constructed, leveraging the synergy between channel attention and spatial attention to strengthen modality-specific representations while suppressing irrelevant noise. At the feature fusion stage, a cross-modal bidirectional attention fusion(CMBAF) module was devised based on Transformer cross-attention mechanisms, enabling the modeling of long-range semantic dependencies across modalities for texture-guided thermal infrared compensation and saliency-guided visible-light denoising. Moreover, an uncertainty-aware dynamic gating mechanism grounded in information entropy was introduced to adaptively calibrate modality fusion weights in response to varying environmental reliability, thereby enhancing model generalization under modality degradation. Extensive experiments on three public benchmarks GTOT, RGBT234, and LasHeR demonstrated that the proposed method achieves a precision rate of 67.4% and a success rate of 53.6% on LasHeR, surpassing state-of-the-art methods by approximately 2.3% in precision, which validates the effectiveness of the proposed approach.

Key words: RGBT object tracking; modal-aware enhancement; bidirectional attention fusion; Transformer cross-attention mechanism; uncertainty modeling

视觉目标跟踪是计算机视觉领域的基础研究问题之一,在视频监控、自动驾驶及人机交互等应用场景中具有重要意义^[1]。传统基于单一可见光(RGB)模态的跟踪方法在光照剧烈变化、夜间环境以及复杂背景干扰等情况下,往往难以维持稳定的跟踪性能。热红外(TIR)传感器通过感知物体的热辐射信息进行成像,具有不依赖光照条件的优势,能够在低照度甚至全黑环境下稳定工作。因此,将RGB与TIR两种模态进行融合的可见光-热红外(RGBT)目标跟踪技术,因其在信息互补性方面的显著优势,逐渐成为研究热点。

尽管RGBT跟踪能够有效融合多模态信息,但仍面临诸多挑战。首先,不同模态之间存在显著的异构性,导致特征分布差异较大。RGB图像包含丰富的颜色与纹理信息,但在弱光条件下易受噪声干扰;TIR图像能够突出目标轮廓,却缺乏细粒度纹理信息,并容易受到热交叉现象的影响。现有多数方法^[2]通常采用共享权重的骨干网络直接提取特征,缺乏针对模态特性的差异化建模,使得模态特有噪声在融合前未得到有效抑制。其次,融合机制的设计仍存在局限。早期方法^[3-4]多采用特征拼接或加权求和等简单策略,难以刻画复杂的跨模态关系。当某一模态受到严重干扰甚至失效时,静态融合策略可能引入噪声信息,进而导致跟踪性能下降。近年来,深度学习方法通过引入自适应融合机制在一定程度上缓解了上述问题。Zhu等^[5]通过多层特征聚合及通道权重预测实现自适应融合。为进一步提升模型对复杂场景的适应能力,部分研究开始从挑战属性建模角度出发。Li等^[6]将跟踪任务中的挑战划分为模态特定问题与模态通用问题,并通过多分支结构与动态权重分配机制提升鲁棒性;Xiao等^[7]提出基于属性解耦的渐进式融合框架,通过分阶段融合策略提升模型在不同场景下的泛化能力。近年来,基于Transformer^[8]的模型在RGBT跟踪任务中表现出显著优势。Hui等^[9]通过构建模板与搜索区域之间的交互机制增强信息互补性。He等^[10]利用时空交互机制提升定位精度。随着大模型的发展,基于提示学习(visual prompt)和状态空间模型(mamba)的新方法^[11-12]进一步增强了跨模态特征对齐能力。然而,在极端环境下(如目标与背景温度接近),现有方法仍难以有效应对模态退化问题。

针对上述问题,提出一种模态差异感知与双向注意力融合的RGBT目标跟踪方法(MDBA-Track),通过增强模态特征表达能力、强化跨模态语义交互以及引入动态融合机制,有效提升模型在复杂环境

下的鲁棒性与泛化能力。主要工作内容如下:

特征增强机制:提出MAE模块,通过通道与空间注意力机制对RGB和TIR模态进行特征强化。该模块在特征提取阶段实现了特征解耦与预提纯,有效抑制了模态的成像噪声。

双向语义交互:设计CMBAF模块,利用Transformer建立模态间的长距离语义依赖,实现信息互补,提高特征的判别能力。

动态决策调节:引入不确定性度量,构建动态门控决策机制。该机制根据实时环境信噪比自适应调整融合权重,增强了系统在模态失效等极端场景下的鲁棒性。

1 模态差异感知与双向注意力融合网络

1.1 网络整体架构

MDBA-Track采用双流网络架构,主要由特征提取模块、模态感知增强模块、跨模态双向注意力融合模块以及预测头组成,如图1所示。

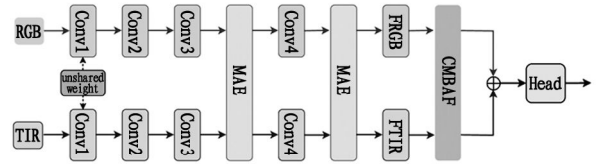


图1 网络框架图

具体而言,模型以模板帧与搜索帧的RGB和TIR图像对作为输入。采用ResNet50作为骨干网络,并去除其最后的全连接层及部分残差块。同时,两个模态分支采用不共享权重的设计,以更好地适应模态间的异构性。在特征提取阶段,从ResNet50的中高层(Conv3与Conv4)输出特征后,引入MAE模块对模态特征进行差异化增强与噪声抑制。随后,将处理后的特征展平并输入至CMBAF模块,通过跨模态双向注意力机制建模长距离语义依赖,实现模态间的深层信息交互。最终,融合特征被送入无锚点预测头,用于输出目标分类得分与边界框回归结果。

1.2 模态感知增强模块

为缓解不同模态在成像过程中引入的噪声干扰,并强化模态特定的判别性特征,设计了模态感知增强模块(MAE)。该模块通过联合建模通道注意力与空间注意力,实现对模态特征的差异化增强。模态感知增强模块如图2所示。

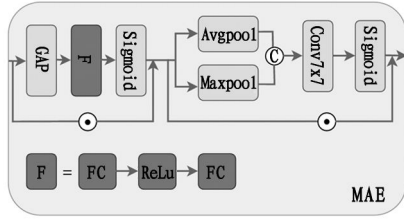


图 2 模态感知增强模块

对于可见光分支,考虑到其在光照充足时纹理信息丰富的特点,而热红外分支,鉴于其主要反映物体的热辐射分布且纹理较弱,因此采用通道注意力机制(channel attention)可以分别筛选两个模态包含特定目标的关键通道,抑制由于光照变化产生的噪声通道和抑制背景中的冷噪声与热弥散。具体而言,以可见光分支为例,假设输入特征为 $X_{rgb} \in R^{C \times H \times W}$,先执行全局平均池化(GAP),将 $C \times H \times W$ 的空间信息压缩为 $C \times 1 \times 1$ 的通道描述向量 V_c ;再将 V_c 向量输入至一个共享权重的多层感知机(MLP),MLP 包含两层全连接层,并引入一个通道压缩比 $Y=16$,以减少参数量并建模通道间的非线性关系。最后通过 Sigmoid 函数(记为 σ)生成通道权重向量 $M_c \in R^{C \times 1 \times 1}$ 。计算公式为:

$$M_c(X_{rgb}) = \sigma(FC(GAP(X_{rgb})))$$

而增强后的特征为: $X_{rgbfinal} = X_{rgb} \odot M_c(X_{rgb})$

考虑到 RGB 和 TIR 模态间可能存在微小空间非对齐问题,研究从空间维度引导网络专注于目标区域,忽略特征图中的杂乱背景。具体做法为:对通道增强后的特征 $X_{rgbfinal}$ 分别执行通道维度的平均池化和最大池化,得到两个 $H \times W \times 1$ 的空间信息描述图。将两张空间描述图在通道维度上拼接,接着使用一个 7×7 的卷积核对拼接后的特征图进行处理,并通过 Sigmoid 函数生成空间权重图 $W_c \in R^{H \times W}$ 。计算为: $W_c = \sigma(\text{Conv}_{7 \times 7}(\text{Concat}[\text{Avgpool}(X_{rgbfinal}), \text{Maxpool}(X_{rgbfinal})]))$,最后将空间权重特征 W_c 与通道增强后的特征 $X_{rgbfinal}$ 进行乘法操作得到最终的模态感知特征: $F_{MAE_{rgb}} = W_c \odot X_{rgbfinal}$

该模块能够在特征提取阶段显式增强模态特定信息,减少噪声对后续融合过程的影响,为跨模态交互提供更高质量的输入特征。

1.3 跨模态双向注意力融合模块

经过增强后的双模态特征 $F_{MAE_{rgb}}$ 和 $F_{MAE_{tir}}$ 仍处于相对独立的表征空间,缺乏深层语义关联。为此,构建了跨模态双向注意力融合模块(CMBAF),利用 Transformer 的多头交叉注意力机制进行双向增强。

具体来说,首先将特征图展平并添加位置编码,

得到 CMBAF 模块包含两个并行的交叉注意力支路,通过交换 Query(Q)实现双向增强。

RGB-to-TIR 交互流:利用可见光特征查询红外特征中的互补信息,旨在利用红外辐射信息对可见光分支进行语义补偿与鲁棒性增强。查询矩阵 Q 来自 F_{rgb} ,键矩阵 K 和值矩阵 V 来自 F_{tir} 。交叉注意力计算公式为: $F_{R \rightarrow T} = \text{Softmax}(\frac{Q_{rgb} K_{tir}^T}{\sqrt{d_k}}) V_{tir}$,其中, d_k

为缩放因子。此过程将红外图像中的显著性热辐射信息注入到可见光特征中。

TIR-to-RGB 交互流:利用红外特征查询可见光特征。 Q 来自 F_{tir} , K 和 V 来自 F_{rgb} 。计算公式为:

$$F_{T \rightarrow R} = \text{Softmax}(\frac{Q_{tir} K_{rgb}^T}{\sqrt{d_k}}) V_{rgb}$$

该过程利用可见光的高分辨率纹理信息修补红外图像中因热交叉导致的目标模糊。

最后,交互后的特征通过逐元素相加实现进一步融合。该融合过程不仅实现了模态间信息的双向补偿,还有效增强了目标特征的判别能力。尽管 CMBAF 模块能够在特征层面实现深度语义融合,但其融合过程仍属于静态建模机制,即对不同模态的贡献缺乏动态适应能力。在复杂场景中,当某一模态受到严重干扰甚至完全失效时,固定融合策略可能引入噪声信息,从而影响最终跟踪性能。

1.4 不确定性感知动态决策

为弥补跨模态双向注意力融合模块在动态适应能力方面的不足,进一步引入不确定性感知动态决策机制。该机制从特征统计分布角度出发,对不同模态的可靠性进行显式建模,从而实现融合权重的自适应调节。研究基于信息论中的香农熵原理,构建轻量化不确定性评估函数。通过提取融合前各模态特征的全局统计信息,并将其映射为概率分布,进而计算对应的不确定性分数: $U_m = \sigma(\psi F_m)$ 。其中, σ 为 Sigmoid 激活函数, ψ 代表感知机参数。该指标能够实时感知环境可靠度:当某一模态包含大量噪声或语义缺失时,其对应的不确定性得分 U_m 将显著升高。为了保证特征维度的一致性,将之前通过逐元素相加的特征经过一个轻量级的 1×1 卷积网络 Φ ,基于此来构建动态融合权重 g_m 。同时引入平滑系数 k 防止数值不稳定:

$$g_m = \text{Softmax}(\frac{k \cdot \Phi(F_{R \rightarrow T} \oplus F_{T \rightarrow R})}{\sqrt{d_k} \cdot (U_{rgb} + U_{tir}) + \varepsilon})$$

其中 ε 为极小值,防止分母为 0。最终融合特征为: $F_{fused} = g_{rgb} \cdot F_{R \rightarrow T} + g_{tir} \cdot F_{T \rightarrow R}$ 。图 3 展示了基于不确定性感知的跨模态双向注意力融合模块。

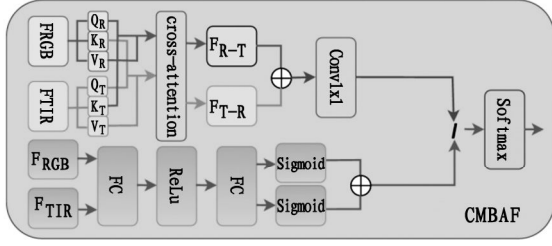


图3 基于不确定性感知的跨模态双向注意力融合模块

1.5 损失函数及跟踪头设计原理

在视觉目标跟踪任务中,跟踪头(tracking head)作为整个网络结构的决策终端,其核心职责是对主干网络提取的深层特征进行解码与判别,输出包括目标位置、尺度等边界框参数以及目标存在的置信度评分。这些输出结果不仅直接决定当前帧的跟踪性能,还将作为下一帧搜索区域生成与候选区域筛选的依据,因而其设计合理性与预测精度直接影响整个跟踪系统的性能上限。而跟踪任务被建模为分类和回归的联合优化问题。总损失函数 L 由分类损失 L_{cls} 和回归损失 L_{reg} 组成: $L = \lambda_1 L_{cls} + \lambda_2 L_{reg}$ 。其中, λ_1 和 λ_2 为平衡超参数。分类损失采用 Focal Loss 以解决正负样本不平衡问题;回归损失采用 GIoU Loss 以提升边界框的定位精度。

2 实验结果与分析

为验证本研究所提出模块 MAE、CMBAF 及不确定性机制的有效性,设计了一系列实验,从整体性能、挑战属性以及模块贡献三个方面进行系统评估,并建立方法设计与实验结果之间的对应关系。

2.1 实验设置

2.1.1 数据集

实验在 GTOT^[13]、RGBT234^[14] 和 LasHeR^[15] 3 个主流数据集上进行。LasHeR 是目前最大的 RGBT 跟踪数据集,包含 1224 个高度对齐的视频序列,涵盖了极为丰富的场景挑战,划分为训练集和测试集。RGBT234 包含 234 个视频对,总帧数约 23.4 万帧,标注了 12 种典型属性(如光照变化、热交叉、部分遮挡等),用于评估算法的泛化能力。GTOT 包含 50 个视频序列,作为早期基准用于验证算法的基本性能。

2.1.2 实施细节

实验利用 PyTorch 和 Python 框架进行模型实现、数据处理及构建实验环境。服务器硬件配置及软件版本如表 1 所示。

算法基于 PyTorch 深度学习框架实现。骨干网

络 ResNet50 使用 ImageNet^[16] 预训练权重初始化,其余参数采用 Xavier 初始化。训练在 NVIDIA RTX 3080 GPU 上进行,批量大小设为 32。训练周期共 10 轮,前 5 个训练周期冻结骨干网络参数进行预热,初始学习率设为 4×10^{-5} ,并在之后每个 epoch 衰减 0.1 倍。优化器采用随机梯度下降(SGD),动量设为 0.9,权重衰减为 0.0005。

表 1 实验硬件配置及软件版本

属性	配置
操作系统	Windows 11
CPU	12 th Gen Intel i7 - 12700 10
GPU	NVIDIA RTX 3080
内存	16 G
Python	v3.12.12
PyTorch	v2.9.1

2.1.3 评价指标

采用精确率(precision rate, PR)和成功率(success rate, SR)作为主要评价指标。PR 计算中心位置误差(center location error, CLE)小于阈值(通常为 20 像素)的帧数占总帧数的百分比。SR 通过计算预测框与真实框的重叠率(intersection over union, IoU),绘制成功率曲线,并取曲线下面积(area under curve, AUC)作为最终得分。

2.2 对比实验与分析

为了验证 MDBA-Track 的整体性能与鲁棒性,选取 10 种主流 RGBT 跟踪算法(传统深度学习、孪生网络、Transformer 类方法)开展对比实验,结合研究中模态差异感知、双向注意力融合、不确定性动态决策三大创新点,从跨模态融合效率、复杂场景适应性、模态退化鲁棒性三个核心维度分析结果,并提炼研究结论。具体如表 2 所示。

2.2.1 LasHeR 数据集实验结论

MDBA-Track 在 LasHeR 数据集上取得 67.4% 精确率、53.6% 成功率,较孪生网络类 CAT 方法成功率提升 21.2%,与 ViPT、OneTracker 等先进 Transformer 方法保持同等竞争力,推理速度达 32FPS。

核心研究结论:研究提出的不确定性感知动态门控机制,可根据模态可靠性自适应分配融合权重,彻底解决了传统静态融合在模态退化场景下引入噪声的缺陷;MAE 模态感知增强模块在特征提取阶段完成模态专属特征强化与噪声抑制,从源头提升特征质量。二者协同使模型在大规模、高复杂度的 LasHeR 数据集上实现性能突破。

表 2 实验结果与其他方法的比较

%

方法	主干网络	GTOT PR ↑/SR ↑	RGBT234 PR ↑/SR ↑	LasHeR PR ↑/SR ↑	FPS ↑
CAT ^[6]	VGG	88.9/71.7	80.4/56.1	45.0/31.4	—
APFNet ^[7]	VGG	90.5/73.7	82.7/57.9	50.0/36.2	1.9
CMPP ^[17]	VGG	92.6/73.8	82.3/57.5	—/—	—
DMCNet ^[18]	VGG	—/—	83.9/59.3	49.0/35.5	—
ViPT ^[19]	ViT	—/—	83.5/61.7	65.1/52.5	—
OneTracker ^[20]	ViT	—/—	85.7/64.2	67.2/53.8	—
DAPNet ^[21]	VGG	88.2/70.7	76.6/53.7	43.1/31.4	—
MANet ^[22]	VGG	89.4/72.4	77.7/53.9	45.5/32.6	2.1
MaCNet ^[23]	VGG	—/—	79.0/55.4	48.2/35.0	1.6
DAFNet ^[24]	VGG	89.1/71.6	79.6/54.4	44.8/31.1	20.5
MDBA-Track	ResNet-50	90.1/75.0	86.1/65.6	67.4/53.6	32

2.2.2 RGBT234 数据集实验结论

在 RGBT234 数据集上,MDBA-Track 成功率超越主流算法约 1.4%,在严重遮挡(HO)、热交叉

(TC)等 12 种挑战属性子集上均取得最优性能,HO、TC 场景成功率较 APFNet 分别提升 7.0%、5.9%。具体挑战属性对比如图 4 所示。

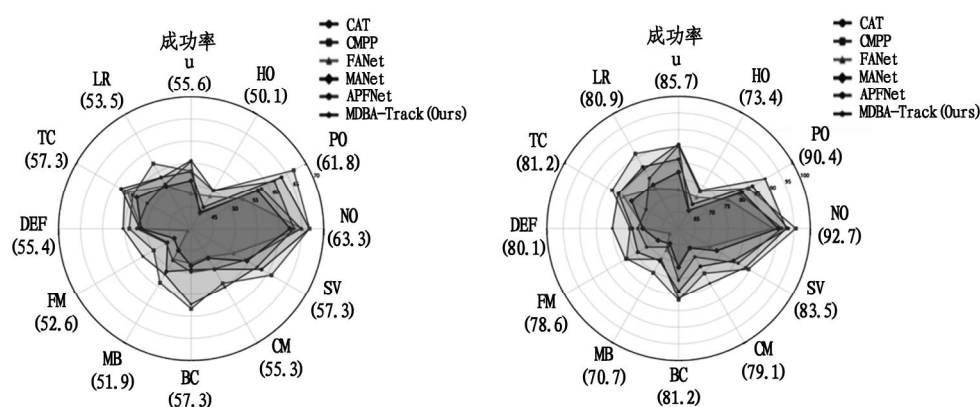


图 4 各方法在 RGBT234 数据集上的挑战属性对比

核心研究结论:MAE 模块可针对性抑制可见光光照噪声、红外热弥散噪声,适配遮挡、光照突变等场景;CMBAF 双向注意力融合模块实现 RGB 与 TIR 模态的长距离语义互补,热交叉导致红外失效时,可依托可见光纹理精准判别目标;不确定性机制实时切换模态权重,进一步强化极端场景鲁棒性。充分证明本方法的多模块协同设计,可全面解决 RGBT 跟踪的核心场景挑战。

2.2.3 GTOT 数据集实验结论

MDBA-Track 在 GTOT 数据集上取得 90.1% 精确率、75.0% 成功率,整体性能优于所对比的多数算法,仅小目标跟踪精度略低于 CMPP。

核心研究结论:研究方法在不同规模、不同类型

的 RGBT 数据集上均具备良好泛化能力,验证了模态差异感知+双向注意力融合策略的通用性。小目标性能差距源于 CMPP 的特征金字塔专属优化,也为后续小目标场景的算法改进指明方向。

综合三大数据集对比结果可证:MDBA-Track 通过模态感知增强、跨模态双向注意力融合、不确定性动态决策三大创新,有效解决了 RGBT 跟踪中模态异构差异大、互补信息挖掘不足、模态退化鲁棒性差的关键问题。方法在复杂退化场景下的精度与鲁棒性显著优于传统方法,与先进 Transformer 方法相比具备性能优势,充分验证了本研究算法的有效性与先进性。

2.3 消融分析

为了系统评估研究所提出的关键模块(MAE、CMBAF及不确定性感知动态决策机制)的有效性,本节设计了逐步引入模块的消融实验,通过对比不同组合下的性能变化,深入分析各模块在模态特征增强、跨模态交互及动态融合中的具体作用。表3详细展示了模块的性能对比结果。

表3 模块有效性验证消融实验结果 %

算法	RGBT234	LasHeR
	PR ↑/SR ↑	PR ↑/SR ↑
基线	72.2/51.4	41.3/30.2
+ RGB-to-TIR	79.6/54.3	42.6/30.7
+ TIR-to-RGB	80.1/55.4	43.4/31.1
+ MAE	81.7/56.8	45.4/32.5
+ CMBAF	81.3/56.1	44.8/32.2
+ MAE + CMBAF (Ours)	86.1/65.6	67.4/53.6

具体消融实验设置:(1)基线:仅使用双流 Res-Net50 提取特征,并采用简单的通道拼接(concatenation)进行融合;(2)在基线方法上仅利用可见光引导红外增强(+RGB-to-TIR):在该设置中,仅保留从可见光到热红外的单向注意力分支。实验结果显示,在可见光照充足、红外热交叉严重的场景下,SR 较基线提升明显,但在低光照或全黑环境下,由于可见光信号失效,该分支无法提供有效的语义补偿,性能提升受限;(3)在基线方法上利用红外引导可见光抗噪(+TIR-to-RGB):仅保留从热红外到可见光的单向注意力分支。实验发现,该分支在夜间场景下表现稳定,能利用红外显著性引导网络关注目标,但在目标与背景温度接近(热交叉)时,由于缺乏可见光纹理的补充,容易发生跟踪漂移;(4)在基线方法基础上引入模态感知增强模块(+MAE):结果显示 SR 分别提升了 5.4% 和 2.3%。这表明在特征提取阶段进行模态解耦与去噪,能够有效抑制特有噪声对后续融合过程的负面影响;(5)在基线方法上加入基于感知动态决策的双向注意力融合模块(+CMBAF):相比简单的拼接和静态加权,SR 分别提升了 4.7% 和 2.0%,表明引入不确定性度量因子作为门控约束以及基于 Transformer 的双向交互能够挖掘更深层的语义关联的有效性;(6)同时引入 MAE、CMBAF 及不确定性决策机制(MAE + CMBAF + 不确定性感知):SR 和 PR 均达到最优,证明各模块间存在良性的耦合效应,共同提升了跟踪性能。

2.4 可视化分析

为了直观展示算法性能,图5给出了在典型挑战序列上的跟踪可视化结果。在低光照场景下,可见光图像几乎全黑,单一 RGB 跟踪器完全丢失目标。而 MDBA-Track 通过不确定性感知,自动降低了 RGB 分支权重,主要依赖 TIR 分支准确锁定了目标。在热交叉场景下(目标与背景温度接近),TIR 图像对比度极低。此时,算法利用 CMBAF 模块中的 TIR-to-RGB 交互流,借助可见光纹理信息成功区分了目标与背景,避免了跟踪漂移。在部分遮挡场景下,MDBA-Track 预测的边界框相比对比算法更加紧致,证明了双向注意力机制有助于在遮挡发生时利用局部特征补全整体语义。

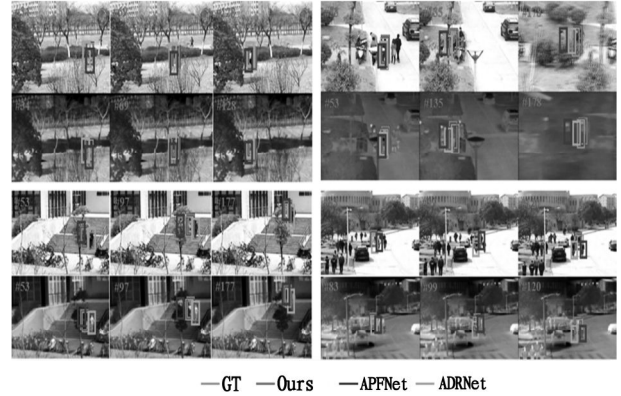


图5 可视化对比分析

3 结论

针对 RGBT 目标跟踪中模态差异显著及模态退化带来的挑战,提出了一种基于模态差异感知与双向注意力融合的跟踪方法。通过引入 MAE 模块,实现了模态特征的差异化增强与噪声抑制;通过 CMBAF 模块,实现了跨模态的深层语义交互;通过不确定性感知机制,实现了融合权重的动态调节。实验结果表明,该方法在多个公开数据集上均取得了优异表现,验证了其在复杂场景下的有效性与鲁棒性。未来工作将从以下三个方向持续推进:探索轻量化 Transformer 骨干结构,以提升模型在资源受限嵌入式平台上的部署能力;引入时序记忆机制,增强模型对长时遮挡与目标重现场景的适应能力;研究模型压缩与量化策略,进一步降低推理延迟,推动算法向实时应用场景落地。

参考文献:

- [1] 卢湖川,李佩霞,王栋. 目标跟踪算法综述[J]. 模式识别与人工智能,2018,31(1):61-76.
- [2] Li Chenglong, Wu Xiaohao, Zhao Nan, et al. Fusing two-stream convolutional neural networks for RGB-T object

- tracking[J]. Neurocomputing, 2018, 281:78-85.
- [3] Lan Xiangyuan, Ye Mang, Zhang Shengping, et al. Modality-correlation-aware sparse representation for RGB-infrared object tracking[J]. Pattern Recognition Letters, 2020, 130:12-20.
- [4] Li Chenglong, Xiang Zhiqiang, Tang Jin, et al. RGBT tracking via noise-robust cross-modal ranking[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(9):5019-5031.
- [5] Zhu Yabin, Li Chenglong, Tang Jin, et al. Quality-aware feature aggregation network for robust RGBT tracking[J]. IEEE Transactions on Intelligent Vehicles, 2020, 6(1):121-130.
- [6] Li Chenglong, Liu Lei, Lu Andong, et al. Challenge-aware RGBT tracking[C]//Computer VisionECCV2020. Chan; Springer international Publishing, 2020:222-237.
- [7] Xiao Yun, Yang Mengmeng, Li Chenglong, et al. Attribute-based progressive fusion network for RGBT tracking[C]. Proceedings of the 36th AAAI Conference on Artificial Intelligence, 2022, 36(3):2831-2838.
- [8] Vaswani Ashish, Shazeer Noam, Parmar Niki, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30.
- [9] Hui Tianrui, Xun Zizheng, Peng Fengguang, et al. Bridging search region interaction with template for RGB-T tracking[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023:13630-13639.
- [10] He Kaijie, Zhang Canlong, Xie Sheng, et al. Target-aware tracking with long-term context attention[C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(1):773-780.
- [11] 周姣. 基于跨模态 Mamba 交互融合网络的 RGBT 目标跟踪[J]. 电脑编程技巧与维护, 2026(3):174-176.
- [12] Chen Jia, Xu Rui, Chen Si, et al. ProMoT: progressive prompting of modality and temporal dynamics for RGB-T tracking[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2026, 36(6):8092-8107.
- [13] Li Chenglong, Cheng Hui, Hu Shiyi, et al. Learning collaborative sparse representation for grayscale-thermal tracking[J]. IEEE Transactions on Image Processing, 2016, 25(12):5743-5756.
- [14] Li Chenglong, Liang Xinyan, Lu Yijuan, et al. RGB-T object tracking: benchmark and baseline[J]. Pattern Recognition, 2019, 96:106977.
- [15] Li Chenglong, Xue Wanlin, Jia Yaqing, et al. LasHeR: a large-scale high-diversity benchmark for RGBT tracking[J]. IEEE Transactions on Image Processing, 2021, 31:392-404.
- [16] Deng Jia, Dong Wei, Socher Richard, et al. ImageNet: a large-scale hierarchical image database[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009:248-255.
- [17] Wang Chaoqun, Xu Chunyan, Cui Zhen, et al. Cross-modal pattern-propagation for RGB-T tracking[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020:7064-7073.
- [18] Lu Andong, Qian Cun, Li Chenglong, et al. Duality-gated mutual condition network for RGBT tracking[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 36(3):4118-4131.
- [19] Zhu Jiawen, Lai Simiao, Chen Xin, et al. Visual prompt multi-modal tracking[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023:9516-9526.
- [20] Hong Lingyi, Yan Shilin, Zhang Renrui, et al. Onetracker: unifying visual object tracking with foundation models and efficient tuning[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024:19079-19091.
- [21] Zhu Yabin, Li Chenglong, Bin Luo, et al. Dense feature aggregation and pruning for RGBT tracking[C]. Proceedings of the 27th ACM International Conference on Multimedia, 2019:465-472.
- [22] Li Chenglong, Lu Andong, Zheng Aihua, et al. Multi-adapt-er RGBT tracking[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2019:2262-2270.
- [23] Zhang Hui, Zhang Lei, Zhuo Li, et al. Object tracking in RGB-T videos using modal-aware attention network and competitive learning[J]. Sensors, 2020, 20(2):393.
- [24] Gao Yuan, Li Chenglong, Zhu Yabin, et al. Deep adaptive fusion network for high performance RGBT tracking[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2019:91-99. 责任编辑:李云

基于改进 YOLOv12 的中药饮片识别标注研究

郇悦*, 朱锦龙, 刘儒香

(亳州学院 电子与信息工程系, 安徽 亳州 236800)

摘要:针对中药饮片识别标注问题,提出一种改进的 YOLOv12 目标检测模型。首先,在骨干网络设计上引入 EDFFN 和多级辅助骨干网络结构,以强化样本特征捕捉能力和骨干网络特征融合能力,实现跨路径、跨尺度的深度特征融合与增强,提取中药饮片的潜在语义信息。其次,在颈部网络中集成 EDFFN 和三重注意力机制用以提炼三种尺度的特征信息,获取中药饮片空间特性和语义特征。最后,结合所提出的骨干网络和颈部网络构建一种 MF-YOLO 模型。实验结果表明, MF-YOLO 在精确率、召回率、mAP@0.5、mAP@0.75 和 mAP@0.5:0.95 上相比于 YOLOv12 分别提升 4.2%、11.94%、11.02%、11% 和 11.01%, 各项指标均优于 YOLOv12, 在中药饮片识别标注上具有一定有效性。

关键词: 中药饮片; EDFFN; 注意力机制; 目标检测

中图分类号: TP391.41

文献标识码: A

文章编号: 2095-297X(2026)05-0070-08

Research on Identification and Labeling of Chinese Medicine Decoction Pieces Based on Improved YOLOv12

TAI Yue*, ZHU Jinlong, LIU Ruxiang

(Department of Electronic and Information Engineering, Bozhou University, Bozhou, 236800, Anhui)

Abstract: An improved YOLOv12 object detection model was proposed to address the issue of recognition and annotation for Chinese herbal decoction pieces. Firstly, in the design of the backbone network, the enhanced discriminative frequency feedforward network (EDFFN) and a multi-level auxiliary backbone structure were introduced. This design aims to enhance the sample feature capture capability and the feature fusion ability of the backbone network, thereby achieving cross-path and cross-scale deep feature fusion and enhancement, as well as extracting the potential semantic information of Chinese herbal decoction pieces. Secondly, in the neck network, the EDFFN and a triple attention mechanism were integrated to refine feature information at three scales, which enables the acquisition of both spatial characteristics and semantic features of Chinese herbal decoction pieces. Finally, the proposed backbone network and neck network were combined to construct the MF-YOLO model. Experimental results demonstrated that compared with YOLOv12, the MF-YOLO model achieves improvements of 4.2%, 11.94%, 11.02%, 11% and 11.01% in precision, recall, mAP@0.5, mAP@0.75, and mAP@0.5:0.95, respectively. All metrics of MF-YOLO outperform those of YOLOv12, indicating that the model exhibits certain effectiveness in the recognition and annotation of Chinese herbal decoction pieces.

Key words: Chinese herbal decoction pieces; enhanced discriminative frequency feedforward network; attention mechanism; object detection

收稿日期: 2025-08-17 * 通讯联系人

基金项目: 安徽省高等学校科学研究项目(2024AH051305); 安徽省高等学校省级质量工程项目(2024aijy441)。

作者简介: 郇悦(1995-), 男, 安徽淮北人, 助教, 硕士。E-mail: taiyue@bzuu.edu.cn

随着深度学习技术在传统医药领域的深度融合,中药饮片的智能化识别标注成为研究热点。传统中药饮片鉴别主要依赖人工经验,存在效率低、主观性强等局限性。在实际应用中,中药饮片形态多样性与背景复杂性导致特征提取困难,其次中药饮片原始样态和炮制加工后的药材颜色纹理变化影响传统算法鲁棒性,在应用场景还需要兼顾检测速度与精确率。

传统中药饮片检测方法通过结合传统机器学习方法对中药饮片进行识别分析^[1-2]。机器学习方法在检测过程中存在特征提取困难,泛化能力不足,识别效率低,漏检和误检的情况时有发生^[3]。目前深度学习在中药饮片检测领域已经取得一定进展。王健庆等^[4]采用 GoogleNet 的卷积神经网络模型对自建的 100 类 8886 张图片进行训练,结果显示卷积神经网络相比传统机器学习算法具有优势。胡晓东^[5]在 MobileNetV2 的基础上压缩了模型,采用多尺度特征提取层,提出一种基于 CNN 的多尺度轻量网络模型用于中药饮片识别,该模型在精确率、内存使用量等方面均有提升。谷瑞等^[6]在多尺度特征和重要特征上进行改进,提出一种 BIM-Res2Net-ICA 模型,该模型在 Res2Net 基础上采用双向融合方法并嵌入坐标注意力机制,提升模型对中药饮片的特征提取,增强了在空间和通道上的信息捕获能力。周丽媛^[7]通过引入多尺度卷积、并联空洞卷积和注意力机制来构建中药饮片识别模型,并采用数据增强方式对 18 种 3200 张草本藏药材进行识别实验。高爽等^[8]采用 YOLOv3 对不同场景下的 152 种 19 692 张图片进行实验,该数据集存在遮挡、残缺等复杂样本,采用数据增强方式在中药饮片数据集上有较好的识别效果。谭代庆等^[9]基于残差网络和 Transformer 对 12 类 5 组 8720 张中药饮片数据集进行实验,该方法结合集成学习思想形成一种强分类器,使用投票法获得识别结果,可以有效分类易混淆的目标。为提升小尺寸数据检测性能并降低模型复杂度,孙兴等^[10]提出了一种轻量化的改进 YOLOv8 模型,其核心改进包括使用 GhostC2f 模块和 DySnakeC2f 模块替换主干网络模块,并引入 CA 注意力机制,减小模型体量的同时提高了识别准确性。

1 基于 YOLO 的目标检测

YOLO 作为目标检测的典型代表,以实时性强、精度高的特点,为中药饮片快速识别提供了新的技术路径。早期 YOLO 主要从模型设计的角度建立了 YOLO 系列框架。YOLO 提出一种新颖的检测方法,

通过将目标检测问题转化为回归问题,直接预测空间分离的边界框和相关的类别概率^[11]。后续 YOLOv5 的推出在目标检测方面取得了一定进步,至今还仍在许多领域发挥重要作用。随着不断地迭代,YOLO 系列取得了重要进展。Tian 等^[12]提出了 YOLOv12 模型,YOLOv12 是对经典 YOLO 进行创新性改进,其通过简化注意力机制和融入残差高效层聚合网络(residual efficient layer aggregation networks, R-ELAN) 进而提高检测效率。在检测模型中,如何降低模型复杂性是一个重要问题,其中一种方法是采用局部注意力机制,传统方式包括窗口注意力、十字交叉注意力、轴向注意力等。YOLOv12 则采用一种简单而有效的区域注意力机制,如图 1 所示。

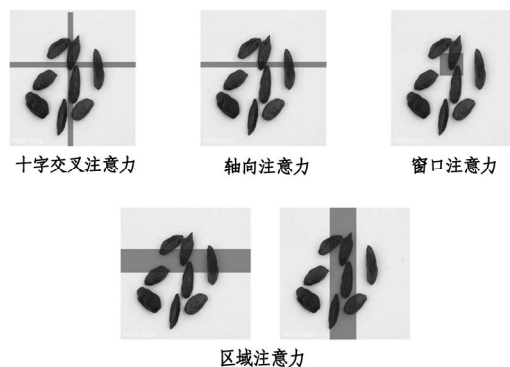


图 1 传统注意力与区域注意力

区域注意力机制是将尺寸为 (H, W) 的特征图进行重新划分,可以划分为 $(\frac{H}{l}, W)$ 或者 $(H, \frac{W}{l})$, 这种操作可以消除显示窗口,提高计算效率。另外高效层聚合网络虽能增强特征聚合能力,却存在梯度阻塞风险及输入输出残差连接的缺失,增加优化复杂性^[13]。为了解决该问题,学者们提出了 R-ELAN 网络结构,采用带缩放因子的块级残差,建立输入至输出的直接残差路径,融合层缩放技术并强化注意力机制^[14]。在 YOLOv12 中,通过引入 R-ELAN 和可分离卷积,并利用残差增强机制和卷积操作提取判别性特征,再进行跨层级特征拼接,最终实现了高效的目标检测效果。

YOLOv12 作为新一代的 YOLO 系列模型,集成了近年来目标检测领域的多项核心模块,其计算量约为 6# GFLOPs,具备更高的计算效率,即更优的 FLOPs/参数比,计算和参数带来的“收益”潜力更大,体现其良好的性能及模块的轻量化,同时具备一定跨领域泛化能力和拓展能力。此外,在中药饮片识别任务中,对检测精度和计算资源都有一定要求,

YOLOv12 相对于传统模型采用更优的架构,在相似的计算量下,其通过架构先进性、计算效率通常能在实际硬件上获得更好的效率。尽管 YOLOv12 在公开数据集上展现了较优的性能,但其在特定垂直领域(如中药饮片识别)的表现未达预期,这揭示了模型在应对中药饮片这类任务时存在的固有局限,包括对低分辨率目标的检测能力不足,对相似特征区分度不高,领域数据分布不均导致的泛化能力弱。正是基于这些不足,本研究对 YOLOv12 模型进行改进,通过增强特征融合能力与特征提取能力,将改进的 YOLOv12 应用于中药饮片识别领域,为实际中药

饮片识别提供技术支撑。

2 网络结构设计

2.1 EDFFN 模块

EDFFN(enhanced discriminative frequency feed-forward network)是一种基于频域的 FFN,主要目标是在频域中高效地、有选择性地处理图像特征,以提升图像恢复的质量^[15]。EDFFN 通过在频域中筛选特征,尤其是在图像去模糊任务中能够自适应地决定保留有效频率信息,提高图像恢复的质量。EDFFN 模块如图 2 所示。

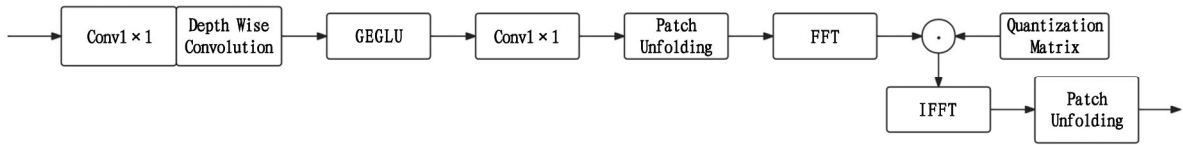


图 2 EDFFN 模块

EDFFN 首先对输入数据 EF_{in} 经过 1×1 卷积和深度可分离卷积进行特征提取,也可处理经标准化处理后的特征,再经过 GEGLU 激活函数,增加非线性特性,提升模型表达能力。处理后的数据再次通过 1×1 卷积,将得到的特征进行频域变换处理,通

过特征图填充和分块处理,将处理后的数据经过量化矩阵和 FFT 进行频域变换得到 Φ ,随后经过逆变换将数据恢复和重组,得到最终的输出 EF_{out} ,EDFFN 计算方式如式(1)、式(2):

$$\Phi = \text{FFT}(\text{PatchUnfolding}(\text{Conv}(\text{GEGLU}(\text{DConv}(\text{Conv}(EF_{in})))))) \quad (1)$$

$$EF_{out} = \text{PatchUnfolding}(\text{IFFT}(\Phi \odot W)) \quad (2)$$

2.2 MLAE 骨干网络

在骨干网络设计上,本研究提出一种基于 EDFFN 的多级辅助(Multi Level Auxiliary based on EDFFN,MLAE)网络结构,以 YOLOv12 骨干网络为基础融入 EDFFN 模块和多级辅助信息结构。这一设计将 EDFFN 用于提取中药饮片模糊特征,多级辅助信息结构用于强化语义,进而提高检测效果。MLAE 骨干网络如图 3 所示。

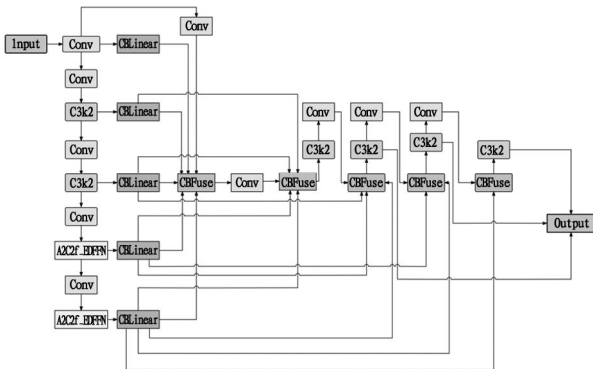


图 3 MLAE 骨干网络

MLAE 采用多个卷积层和模块化设计,通过堆

叠多种模块以增强特征表示能力。在处理输入过程中,依旧采用多个卷积层和 C3k2 进行特征提取,获得多个不同尺寸的特征图。在第四层级(P4)和第五层级(P5)下采样时引入 EDFFN,即 A2C2f_EDFFN 模块。通过 A2C2f_EDFFN 和卷积操作获得新的特征图,再经过一轮 A2C2f_EDFFN 模块得到骨干网络这一阶段输出。

根据多层级卷积处理和多尺度特征融合思想,本研究提出采用多级辅助骨干网络用于增强原有特征的提取。由图 3 可以看出,在原骨干网络的基础上,将经过 P1 至 P5 层级下采样特征送入 CBLinear 模块中,该输出再与原始输入经过 2 次卷积特征所产生的特征图均输入到 CBFuse 中进行特征融合。经过组合不同层级后的特征图,可以有效提高模型的特征表达能力。在 MLAE 中,融合了 5 层特征的 CBFuse 再进行卷积操作,得到的输出再与浅层模块特征融合,再卷积,经过多次 C3k2 与卷积操作及多尺度特征融合,得到多组最终的输出。

2.3 FETN 网络结构

在 YOLOv12 基础上提出一种基于 EDFFN 和三重注意力机制的特征增强网络 (Feature Enhancement based on EDFFN and Triple Attention Mechanism Network, FETN) 来构成 Neck (颈部网络) 部分, 其中三重注意力模块用来增强上文语义, 消除骨干网络输入的背景噪声。三重注意力模块由三个分支组成, 其中两个分支负责捕获空间维度 H 或 W 与通道维度 C 之间的交互特征^[16]。其通过三种不同分支对输入张量进行分区计算, 分别为左分支、中分支和右分支, 三重注意力模块如图 4 所示。

在左分支中, 首先对原始张量 t_{in} 进行池化操作, 将输入张量 $C \times H \times W$ 降至 $2 \times H \times W$, 经过 7×7 卷积层进行特征提取, 再经过 Batch Norm 和 Sigmoid 函数获得注意力权重值, 最终与输入 t_{in} 逐元素相乘得到输出特征 bt_1 , 如式(3)所示:

$$bt_1 = \text{Sigmoid}(\text{BatchNorm}(\text{Conv}_{7 \times 7}(\text{ChannelPool}(t_{in})))) \odot t_{in} \quad (3)$$

在中分支中, 首先对原始张量 t_{in} 改变其维度顺序, 即从 $C \times H \times W$ 沿着高度轴逆时针旋转 90° 得到 $W \times H \times C$, 记为张量 pt_2 , 通过 Z-pool 池化对宽度下采样至 $2 \times H \times C$ 。再通过 7×7 卷积层进行特征提取得到 $1 \times H \times C$, 再经过 Batch Norm 和 Sigmoid 函数, 生成权重参数, 再与旋转后张量 pt_2 逐元素相

$$bt_2 = \text{Permute}(\text{Sigmoid}(\text{BatchNorm}(\text{Conv}_{7 \times 7}(\text{Z-Pool}(pt_2)))) \odot pt_2) \quad (5)$$

同理在右分支中, 对原始张量 t_{in} 进行 W 轴逆时针旋转改变其维度顺序, 获得输入张量 $H \times C \times W$, 再下采样至 $2 \times C \times W$, 再通过 7×7 卷积层降至 $1 \times C \times W$, 再经过 Batch Norm 和 Sigmoid 函数生成注意

$$bt_3 = \text{Permute}(\text{Sigmoid}(\text{BatchNorm}(\text{Conv}_{7 \times 7}(\text{Z-Pool}(pt_3)))) \odot pt_3) \quad (7)$$

综合 bt_1, bt_2, bt_3 三个输出, 获得三种不同维度上的特征信息, 再将三个输出取均值获得三重注意力模块的最终输出 t_{out} , 如式(8)所示:

$$t_{out} = \text{Avg}(bt_1, bt_2, bt_3) \quad (8)$$

这种设计使得模型更有效地捕捉特征的全局语义信息, 实现通道和空间之间的跨维度交互。FETN 网络结构如图 5 所示。

FETN 网络接收骨干网络的特征输入后, 首先执行 SPFF、C2PSA 及上采样操作。将原始输入特征与上采样所得特征进行拼接, 拼接结果输入到 EDFFN 模块。在后续处理中, FETN 网络进行多次特征拼接, 产生的特征经 EDFFN 模块处理后, 传递至三重注意力模块。FETN 通过堆叠多个 EDFFN 模块并结合多尺度融合策略处理的特征, 最终输入三

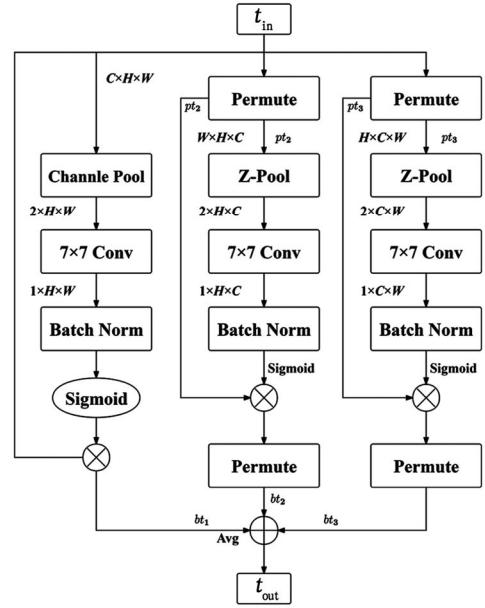


图 4 三重注意力模块

乘, 输出结果沿高度轴顺时针旋转 90° 以恢复原始输入张量的形状, 记为 bt_2 。中分支主要用于计算通道维度 C 和空间维度 H 的注意力权重, 如式(4)、式(5)所示:

$$pt_2 = \text{Permute}(t_{in}) \quad (4)$$

力权重参数, 再与 pt_3 逐元素相乘后顺时针旋转恢复原始形状, 记为 bt_3 。如式(6)、式(7)所示:

$$pt_3 = \text{Permute}(t_{in}) \quad (6)$$

重注意力模块, 产生网络的最终输出。

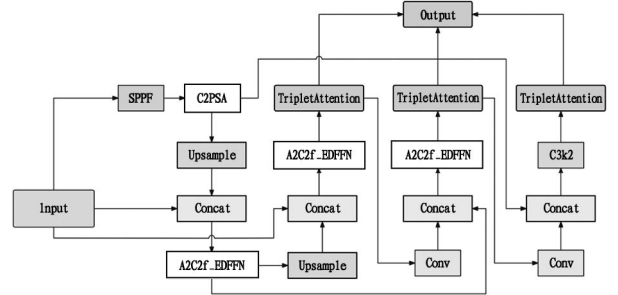


图 5 FETN 网络结构

2.4 MF-YOLO 模型

中药饮片检测识别存在数据稀缺性、环境复杂性及算法泛化性的叠加挑战。为提高识别效果, 结

合本研究所提出的 MLAE 和 FETN 网络结构,提出一种 MF-YOLO(MLAE and FETN-YOLO)模型用于

中药饮片识别,MF-YOLO 模型如图 6 所示。

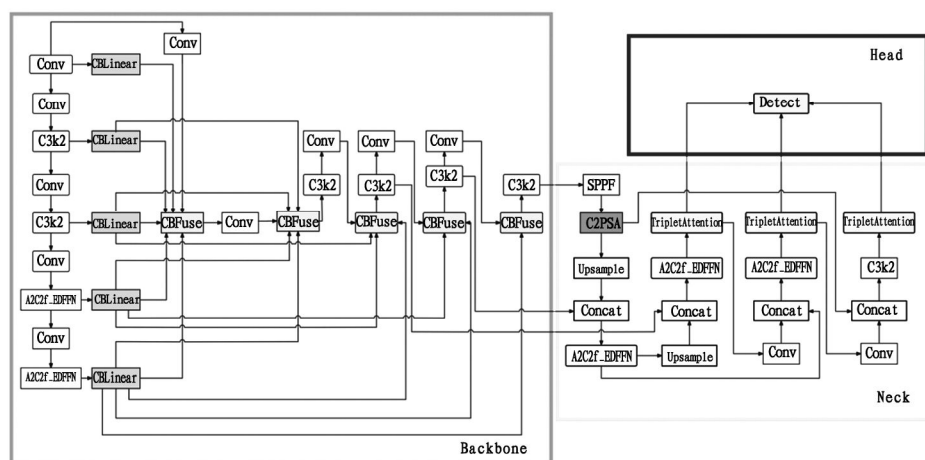


图 6 MF-YOLO 模型

MF-YOLO 骨干网络采用本研究所提出的 MLAE 网络结构,旨在将复杂的中药饮片模糊特征进行削弱,提高关键特征的捕获能力。MLAE 将原始图像经过主分支和多分支获得的特征融合到 CB-Fuse 模块中,经过多级辅助信息结构获得多组特征输出。Neck 则采用本研究提出的 FETN,其在每个尺度上添加 EDFFN 和三重注意力机制用于捕获空间特征,经过 EDFFN 和三重注意力分别提炼到三种尺度的特征图并送入检测头中得到最终的输出。

3 实验结果与分析

3.1 实验数据集与参数配置

为验证模型的有效性,实验分别在中药饮片数据集和公开数据集 VisDrone2019 上进行模型评估。其中所用中药饮片数据集来源于 5117 张尺寸为 416×416 的中药饮片图像,根据中药饮片类别分为枸杞子、人参、蜈蚣等 50 类,中药饮片数据集统计信息如表 1 所示。

表 1 中药饮片数据集统计信息

数据集信息	数量	类别	目标数	类别	目标数	类别	目标数
总图片数	5117	人参	150	菊花	140	蜈蚣	108
总目标数	10543	水蛭	152	白术	148	桑椹	289
平均每张图片目标数量	2.06	大枣	418	核桃仁	263	川贝母	128
平均每个类别目标数量	210.86	百合	138	天麻	195	山药	201
总目标框对数量	17826	黄连	109	鳖甲	156	蝉蜕	229
IOU $\geq 30\%$ 的重叠对	169	乌梅	360	当归	167	葶澄茄	174
IOU $\geq 50\%$ 的重叠对	18	厚朴	310	山楂	192	柴胡	108
IOU $\geq 75\%$ 的重叠对	2	牡蛎	127	藏红花	126	羚羊角	193
IOU $\geq 95\%$ 的重叠对	0	海马	206	蛇蜕	95	胖大海	384
单张图片包含最多目标个数	30	罗汉果	131	杜仲	249	莲子	229
包含 1 个目标的图片数量	3284	甘草	146	白茅根	114	蛤壳	341
包含 2 个目标的图片数量	683	三七	121	金银花	218	川芎	284
包含 3 个目标的图片数量	363	通草	117	栀子	553	大蒜	321
包含 30 个目标的图片数量	1	薏苡仁	127	全蝎	119	银杏叶	331
类别中包含最多的目标数	553	枸杞子	310	鱼腥草	102		
类别中包含最少的目标数	95	八角	342	土鳖虫	212		

由表 1 可以看出,数据集分辨率较低,类别丰富,目标重叠较少,标注质量较好。本次实验中药饮片数据集按 8:2 比例划分为训练集和验证集,训练过程中 batch 设置为 32,使用 SGD 优化器,epochs 为 300 次。公共数据集 VisDrone2019 包括训练集 6471 张图像,验证集 548 张图像,训练过程中 batch 设置为 4,使用 SGD 优化器,epochs 为 300 次。

3.2 评价指标

实验采用目标检测领域常用的精确率 P ,召回率 R 和 mAP 作为评价指标,计算见式(9)、式(10)。

$$P = \frac{TP}{TP + FP}, R = \frac{TP}{TP + FN} \quad (9)$$

$$AP = \int_0^1 P(r) dr, mAP = \sum_{i=1}^n \frac{AP(i)}{n} \quad (10)$$

为了更好衡量模型的综合性能,细化模型的识别能力,其中 mAP 分为 $mAP@0.5$, $mAP@0.75$ 和 $mAP@0.5:0.95$ 。

3.3 消融实验

为验证 MF-YOLO 模型改进的有效性,在中药饮片数据集上进行消融实验来验证模型性能的提升

效果,消融实验结果如表 2 所示。MF-YOLO 包含两种模块和两种网络结构,模型 1 和模型 2 分别评估两个核心模块的独立作用,即 EDFFN 与三重注意力。模型 3 为两个核心模块的结合效果,MLAE 与 FETN 分别为模型 4,模型 5。由于 MF-YOLO 是 MLAE 网络与 FETN 网络相结合,即最终模型通过整合各个模块和网络结构形成模型 6。由表 2 可以看出,与基线相比,单独引入三重注意力模块(模型 2)或 MLAE 网络(模型 4)、FETN 网络(模型 5)均在不同程度上提升了召回率与 mAP ,体现了这些模块或网络的独立有效性。从模型 3 可以看出,EDFFN 与三重注意力进行结合出现性能下降,表明这两个模块的简单叠加未产生预期协同效应,需要通过整体网络结构进行优化整合。完整的 MF-YOLO(模型 6)取得所有模型中的最优性能,其 mAP 显著优于基线及其他消融模型,说明 MF-YOLO 整体架构设计具有一定的合理性,各组件经过系统整合后能够有效互补,共同带来显著的性能提升。

表 2 消融实验对比结果 %

模型	EDFFN	三重 注意力	MLAE	FETN	P	R	$mAP@$ 0.5	$mAP@$ 0.75	$mAP@$ 0.5:0.95
YOLOv12					55.47	48.49	50.17	23.10	25.79
1	√				54.39	45.75	47.55	21.29	23.91
2		√			54.76	51.50	52.55	25.00	26.74
3	√	√			49.57	45.70	45.64	20.19	22.62
4			√		56.09	52.94	53.56	24.37	27.16
5				√	56.12	50.90	53.12	23.55	26.64
6	√	√	√	√	57.80	54.28	55.70	25.64	28.63

3.4 对比实验

为验证 MF-YOLO 模型改进的有效性,将 MF-YOLO 和 YOLOv12 在公共数据集 VisDrone2019 上进行对比实验,如表 3 所示。可以看出,MF-YOLO 模型在各项评估指标上均优于基线模型 YOLOv12,

验证了其综合性能的提升,并非针对特定领域的过拟合优化,具备较好的跨领域泛化能力,其中核心指标 $mAP@0.5$ 提升约 9.93%, $mAP@0.75$ 提升约 13.18%, $mAP@0.5:0.95$ 提升约 11.65%。

表 3 MF-YOLO 和 YOLOv12 在 VisDrone2019 上的对比结果 %

模型	P	R	$mAP@0.5$	$mAP@0.75$	$mAP@0.5:0.95$
MF-YOLO	55.71	43.22	45.38	28.94	27.99
YOLOv12	51.70	39.74	41.28	25.57	25.07

为验证 MF-YOLO 在中药饮片领域的目标检测能力,将 MF-YOLO 与当前主流的目标检测模型进行综合对比,即将 MF-YOLO 和 YOLOv12^[12], YOLOv10n^[17], YOLOv9s^[18], YOLOv9t^[18], YOLOv8, YOLOv5, RT-DETR^[19], SSD^[20] 在中药饮片数据集上进行对比实验。

首先在中药饮片训练集中,各个模型分别在 mAP@0.5 和 mAP@0.5:0.95 上进行了对比,如表 4 所示。可以看出 MF-YOLO 在 300 次迭代过程中, MF-YOLO 的 mAP@0.5:0.95 相对于 YOLOv12 提升约为 11.13%。MF-YOLO 的 mAP@0.5 明显优于其他各个模型,其中相对于基线模型 YOLOv12 提升约 10.76%。

为更好地验证 MF-YOLO 模型性能的有效性,各模型在中药饮片验证集中进行对比实验,如表 5 所示。由表 5 可以看出在各项指标上, MF-YOLO 显著高于基线模型 YOLOv12 和其他模型,在精确率上相对于基线 YOLOv12 提升约 4.2%,在召回率上提升约 11.94%。在核心性能指标 mAP 上, MF-YOLO 在 mAP@0.5 上提升约 11.02%,在 mAP@0.75 上提升约 11%,在 mAP@0.5:0.95 上提升约 11.01%。

同时也可以看出 MF-YOLO 模型性能与模型体

量之间较为均衡,尽管 MF-YOLO 的参数量和计算量高于部分轻量模型,但其性能提升幅度超过其复杂度增长。相较于参数量较少的 YOLOv9t, mAP@0.5:0.95 指标提升约 17.1%,体现其代价可控且收益显著。与参数量相近的 YOLOv9s 相比, MF-YOLO 模型在各项指标更高的同时,计算量降低了约 50% 并取得了 mAP@0.5:0.95 指标 5.41% 的提升,凸显了其更高的计算效率,与参数量和计算量较大的模型相比,在核心综合性能指标 mAP 上全面优于 RT-DETR。

表 4 不同模型在中药饮片训练集下的对比结果 %

模型	mAP@0.5	mAP@0.5:0.95
MF-YOLO	55.81	28.66
YOLOv12	50.39	25.79
YOLOv10n	51.90	26.76
YOLOv9t	48.46	24.46
YOLOv9s	53.93	27.15
YOLOv8	55.31	27.91
YOLOv5	53.93	27.18
RT-DETR	45.78	23.62
SSD	52.94	25.30

表 5 不同模型在中药饮片验证集下的对比结果

模型	P/%	R/%	mAP@0.5/%	mAP@0.75/%	mAP@0.5:0.95/%	Parameters/M	FLOPs/G
MF-YOLO	57.80	54.28	55.70	25.64	28.63	4.84	11.2
YOLOv12	55.47	48.49	50.17	23.10	25.79	2.52	5.9
YOLOv10n	56.62	48.51	51.88	24.89	26.74	2.27	6.6
YOLOv9s	56.03	52.68	53.82	23.95	27.16	6.21	22.2
YOLOv9t	54.57	45.90	48.19	21.18	24.46	1.74	6.5
YOLOv8	56.14	54.03	55.07	25.22	27.91	2.69	6.9
YOLOv5	56.64	51.66	53.93	23.65	27.19	2.20	6.0
RT-DETR	56.49	47.77	45.72	21.75	23.56	32.09	103.7
SSD	56.55	52.72	52.40	21.20	25.20	19.16	34.3

4 结论

YOLOv12 在应对中药饮片这一特殊任务场景时存在一定局限性,包括对低分辨率数据和相似特征区分能力不足,中药饮片数据分布不均导致的泛化能力弱等情况,但是在计算效率与架构设计上具有优势,相对其他模型更为轻量。基于此,本研究展开了针对性的改进,提出了基于 YOLOv12 改进的

MF-YOLO 模型。MF-YOLO 模型不仅超越了基准模型 YOLOv12,也优于其他对比模型,验证了本次改进工作的有效性和必要性。

MF-YOLO 针对骨干网络和颈部网络进行了改进,骨干网络通过集成 EDDFN 模块和多级辅助信息分支,实现多尺度特征的高效提取与融合,提高了小尺寸、小样本和复杂环境下数据的特征表达。Neck 则采用 EDDFN 和三重注意力相结合的方式,增强融

合骨干网络提供的多尺度特征,有助于网络在深层特征中捕获更具判别性的信息。MF-YOLO 整体采用模块化设计优化特征提取效率,经实验验证了模型的有效性,提高了中药饮片的检测效果。未来工作将围绕数据集容量、模型网络优化等方面展开,探索中药饮片智能化分拣、溯源等应用场景的技术支撑,进一步推动智慧中药转型。

参考文献:

- [1] 李梦, 张小波, 刘绍波, 等. 部分可解释机器学习方法的高光谱人参产地识别和分析[J]. 光谱学与光谱分析, 2022, 42(4): 1217 – 1221.
- [2] WANG J F, CAI C Z, KONG C Y, et al. A computer method for validating traditional Chinese medicine herbal prescriptions[J]. The American Journal of Chinese Medicine, 2005, 33(2): 281 – 297.
- [3] 傅弘, 池哲儒, 常杰, 等. 基于人工神经网络的叶脉信息提取——植物活体机器识别研究 I [J]. 植物学通报, 2004(4): 429 – 436.
- [4] 王健庆, 戴恺, 李子柔. 基于深度学习的中药饮片图像识别研究[J]. 时珍国医国药, 2020, 31(12): 2930 – 2933.
- [5] 胡晓东. 基于改进深度学习算法的中药饮片图像识别研究[D]. 长春: 吉林农业大学, 2023.
- [6] 谷瑞, 宋翠玲, 李元昊. 基于 BIM-Res2Net-ICA 的中药饮片图像识别模型[J]. 时珍国医国药, 2025, 36(6): 1192 – 1200.
- [7] 周丽媛. 基于深度学习的草本藏药材识别方法研究[D]. 拉萨: 西藏大学, 2023.
- [8] 高爽, 周志强, 黄显章, 等. 基于 YOLOv3 算法的中药饮片检测研究[J]. 中药材, 2023, 46(1): 57 – 61.
- [9] 谭代庆, 肖志鸿, 吴浩忠, 等. 融合改进残差网络和 Transformer 的易混淆中药饮片识别研究[J]. 中国数字医学, 2023, 18(6): 42 – 50.
- [10] 孙兴, 李杨, 毛天驰, 等. 改进 YOLOv8 在中药饮片自动识别标注中的研究与应用[J]. 中草药, 2025, 56(1): 216 – 226.
- [11] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779–788.
- [12] TIAN Y J, YE Q X, DOERMANN D. YOLOv12: Attention-centric real-time object detectors[J]. arXiv preprint arXiv: 2502.12524, 2025.
- [13] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 7464 – 7475.
- [14] TOUVRON H, CORD M, SABLAYROLLES A, et al. Going deeper with image transformers [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 32 – 42.
- [15] KONG L, DONG J, TANG J, et al. Efficient visual state space model for image deblurring[C] // Proceedings of the Computer Vision and Pattern Recognition Conference. 2025: 12710 – 12719.
- [16] MISRA D, NALAMADA T, ARASANIPALAI A U, et al. Rotate to attend: Convolutional triplet attention module [C] // Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2021: 3139 – 3148.
- [17] WANG A, CHEN H, LIU L, et al. YOLOv10: Real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107984 – 108011.
- [18] WANG C Y, YE H I, LIAO H Y M. YOLOv9: Learning what you want to learn using programmable gradient information[C] // European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 1 – 21.
- [19] ZHAO Y, LYU W, XU S, et al. Detsr beat Yolovs on real-time object detection[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 16965 – 16974.
- [20] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector[C] // European Conference on Computer Vision. Cham: Springer International Publishing, 2016: 21 – 37.

责任编辑: 葛圣志

双孔平行盾构掘进地表及管线沉降对比分析

王 鹏*, 王 芸, 李永康, 辛海洋, 张晶晶, 戴光明

(马鞍山学院 建筑工程学院, 安徽 马鞍山 243000)

摘 要:通过某地铁盾构实测、理论和模拟对比研究双线平行盾构掘进引起的地表和管线变形规律,结果表明:双线平行隧道线间距为14 m时,单双线盾构掘进引起的地表和管线横向沉降均近似呈V形变化;盾构机掘进引起地表纵向沉降的范围约为35 m,主要经历了缓慢发展、迅速发展和逐步稳定三个阶段,地面行车荷载增加了地表纵向沉降;管片顶部和两侧分别承受较大的压应力和拉应力,存在应力集中现象;隧道横断面由设计“圆形”变形为“椭圆形”,竖向呈现挤压变形,水平呈现拉伸变形。

关键词:双孔平行盾构;地表沉降;管线沉降;数值模拟

中图分类号:U455.43

文献标识码:A

文章编号:2095-297X(2026)05-0078-06

Comparative Analysis of Ground and Pipeline Settlement During Double-Tube Parallel Shield Tunneling

WANG Peng*, WANG Yun, LI Yongkang, XIN Haiyang, ZHANG Jingjing, DAI Guangming

(School of Architecture Engineering, Ma' anshan University, Ma' anshan, 243000, Anhui)

Abstract:Through a comparative analysis of field measurements, theoretical studies, and simulations conducted on a subway shield tunneling project, it explored the deformation patterns of ground surface and pipelines induced by double-tube parallel shield tunneling, and showed that when spacing between double parallel tunnels is 14 meters, lateral settlements of ground surface and pipelines caused by single and double-tube shield tunneling both exhibit an approximate V-shaped variation; range of longitudinal surface settlement caused by shield tunneling was approximately 35 meters, which mainly goes through three stages: slow development, rapid development and gradual stabilization, load of ground vehicles increases longitudinal surface settlement; top and both sides of segment are subjected to significant compressive and tensile stresses with stress-concentration phenomena; cross-section of tunnel deforms from a “circular shape” to an “elliptical shape”, with compressive deformation in the vertical direction and tensile deformation in the horizontal direction.

Key words:double-tube parallel shield; ground settlement; pipeline settlement; numerical simulation

交通运输部已公布全国共有54个城市开通运营城市轨道交通线路343条,里程11 710.3 km,其中43个城市开通运营地铁、轻轨线路282条。因地铁工程多位于地下,不占用地上空间资源,不受地面交通影响,运行效率较高,速度较快(部分线路速度

为160 km/h),可有效缓解城市交通拥挤等问题。地铁隧道施工一般采用盾构机掘进施工,难免会下穿建(构)筑物、给排水等管道,盾构机掘进施工时会切削地下岩土体,局部会扰动岩土体。切削岩土体使得地下应力部分释放,会引起地表建(构)筑物

收稿日期:2026-03-06 * 通讯联系人

基金项目:安徽省高等学校自然科学基金资助项目(KJ2020A0845);马鞍山学院科研基金重点项目(QS2022005);马鞍山学院大学生创新创业训练项目(x202513614256)。

作者简介:王鹏(1990—),男,安徽亳州人,讲师,硕士。E-mail:992917071@qq.com

及管线的变形,甚至威胁其稳定性、安全性,对地下结构产生附加应力等,因此有必要研究地铁盾构下穿管道的受力变形影响^[1-4]。

早在 1969 年 Peck 就总结出单线盾构施工引起的地表沉降经验公式^[5]。目前国内各高校、研究所等主要通过理论分析、室内模型试验、工程实测数据和数值模拟展开研究。刘波等^[4]针对全国各地地铁盾构穿越建(构)筑物、地下管道等分析总结,归纳其引起结构变形问题的控制难点,给出控制结构变形的可行方案和加固措施。冯国辉等^[6]基于已有下穿管线影响的理论分析,采用力学平衡法总结出管线竖向受力变形的方程,并通过数值模拟验证分析其可靠性。史江伟等^[7]通过模型试验研究隧道施工对上覆管线的影响,提出管线-开挖面净距是控制管线变形的关键参数,建立了地下管线最大沉降预测方法。任旭东等^[8]基于遗传算法优化的 BP 神经网络研究盾构掘进地表沉降及掘进参数优化,得出降低泥浆黏度和提高泥浆密度可控制地表沉降,提高推进速度可降低对地表沉降的影响。席培胜^[9]、王忠昶等^[10-11]利用 FLAC3D 软件分析地铁盾构施工对地下管线产生的变形影响,总结出管线与盾构中心线距离是重要因素,地表横向沉降影响距离距隧道中心约为 30 m,隧道横断面最终呈“椭圆形”。周绍宾等^[12]利用 Midas-GTS 软件研究盾构隧道下穿排水箱涵与地下管线引起的变形,提出了控制沉降的技术措施和时间节点。双线盾构施工引起的地表沉降并不是简单的两条线沉降叠加,魏纲等^[13]、马可栓^[14]、宫亚峰等^[15]基于 Peck 经验公式总结出双线盾构引起的地表沉降公式。潘茁等^[16]基于实测数据分析了沉降槽宽度与沉降槽宽度系数之间的关系,沉降主要集中在距隧道轴线水平距离 2 倍沉降槽宽度系数范围内。

部分研究盾构掘进引起管线沉降时并未考虑到地面行车荷载的影响,当道路较宽、车流量较多时,车辆行车荷载对盾构掘进引起的地面和管线变形不

应该被忽视。本研究以某地铁区间盾构掘进下穿给排水管道为研究背景,通过现场实测、经验理论公式和数值模拟对比验证分析单、双线盾构掘进对地表及管线沉降影响并考虑地面行车荷载对其的影响。

1 工程概况

1.1 工程简介

某地铁 1 号线盾构区间沿科苑横路前行,下穿石大路、日慎精工电子厂、松山湖法院和华勤地块中间区域进入新城路至松山湖站,长约 3.1 km。起点里程右线 CK33 + 637.390,终点里程右线 CK36 + 748.147。区间双线隧道采用土压平衡盾构机施工,左右线盾构中心点距离为 14 m,埋深为 15.88 ~ 34.11 m,隧道内径为 6 m,管片厚度为 0.35 m,管片宽度为 1.5 m,错缝拼接,采用同步注浆方式对管片进行充填,先施工左线,后施工右线。其盾构区间平面图如图 1 所示。

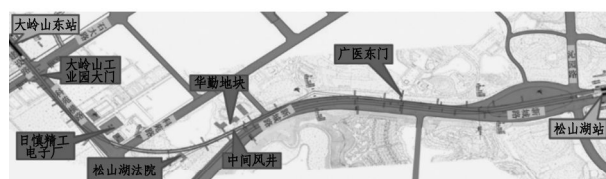


图 1 大岭山站—松山湖站区间平面图

1.2 工程地质与水文地质

根据勘察报告,该区间勘探范围内揭露的岩土层如下:黏性土素填土 1-1,黄褐色,含少量粉砂、碎石等,平均厚度为 2.85 m;淤泥质黏土,平均厚度为 4.30 m;砂质黏性土 6-5/6-6,平均厚度为 13.85 m;全风化混合花岗岩 8-1,棕红色、褐黄色,坚硬土状,平均厚度为 4.76 m;强风化碎块状混合花岗岩 8-2-1,灰黄色、棕红色,平均厚度为 3.15 m;中风化混合花岗岩 8-3,灰色、灰白色。

地下水属于水文地质 II 单元,主要是第四系松散岩类孔隙水(一)和基岩裂隙水(二)。

各岩土层分布及物理力学参数如表 1 所示。

表 1 岩土物理力学参数

岩土层名称	密度 $\rho/\text{g} \cdot \text{cm}^{-3}$	黏聚力 c/kPa	内摩擦角 $\varphi/(\circ)$	泊松比 ν	弹性模量 E_s/MPa
黏性土素填土 1-1	1.82	10	8	0.25	7
淤泥质黏土 3-4-1	1.75	8	6	0.27	9
砂质黏土 6-5/6-6	1.80	18	17	0.27	8
全风化花岗岩 8-1	1.85	19	20	0.28	10
强风化花岗岩 8-2-1	1.96	30	28	0.23	25
中风化花岗岩 8-3	2.50	400	35	0.20	49.5

1.3 监测内容及结果分析

盾构机始发不久即下穿石大路两条钢筋混凝土结构给水管,直径为 600 mm,埋深约为 3 m,一条电子厂排水管,直径为 200 mm,埋深约为 2 m。石大路为东西主干道,交通车流量较大,如果盾构下穿管线时引起管线过大沉降甚至破坏,可能造成严重后果。为此沿地表横向和纵向方向布置监测点,监测盾构机掘进过程对地表横向和纵向沉降、管线沉降的影响,横断面沿纵向间距为 30 ~ 50 m,地表和管线纵向监测点间距为 10 m,监测频率为 1 次/1 天,变形较大或出现报警时增加监测频率,并分析原因,对施工过程实施全面监控和有效控制管理。监测点平面布置如图 2 所示。

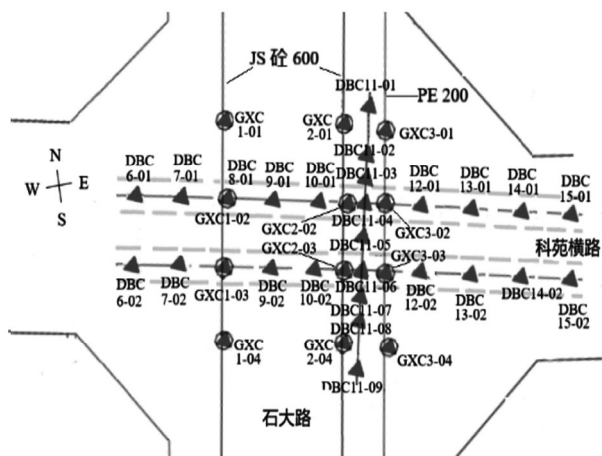


图 2 监测点布置图

地表纵向沉降监测点为 DBC6 至 DBC15,地表横向监测点为 DBC11-01 至 DBC11-09,由南向北方向第一、二、三条管线监测点分别为 GXC1-01 至 GXC1-04、GXC2-01 至 GXC2-04、GXC3-01 至 GXC3-04,采用天宝 DINI03 仪器采集监测数据。地表沉降设计限值为 $(-30) \sim (+10)$ mm,管线设计限值为 ± 30 mm。

2 双线隧道数值模型及模拟

2.1 双线隧道数值模型

FLAC3D 软件是将连续快速拉格朗日分析法应用于岩土工程问题上,使用手册中也有隧道相关的建模知识,因此可用于地铁盾构施工模拟。三维模

型尺寸为横向 72 m × 纵向 90 m × 竖向 46 m,靠近隧道附近的网格较密。输入“print information”获得“List Information”清单,显示“Number of Gps: 334 019”,表明模型共有节点 334 019 个,显示“Number of Zones: 323 688”,表明模型共有 323 688 个网格单元。岩土体采用莫尔库伦弹性模型,利用 Null 命令模拟盾构掘进时岩土体开挖,盾构管片安装采用 liner 衬砌单元。固定约束模型四周边界和底边界。数值模拟时盾构管片弹性模量为 30.0 GPa,密度为 $2.5 \text{ g} \cdot \text{cm}^{-3}$,泊松比为 0.2^[2-3]。地铁盾构和地下管线三维模型如图 3 所示。

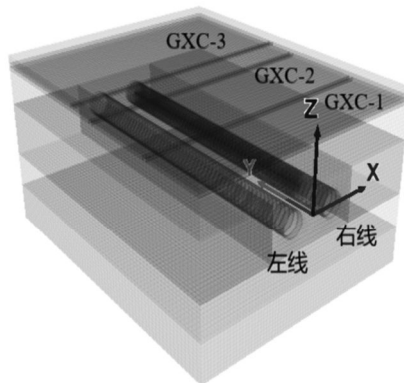


图 3 地铁盾构和地下管线三维模型

2.2 数值模拟结果

数值模拟前在 GXC1 和 GXC2 内部及两侧 3 m 范围施加 20 kPa 的均布荷载,在左线($x = -12 \sim -2$ m 和 $x = 7 \sim 17$ m)范围施加 10 kPa 的均布荷载,模拟石大路和科苑横路交通车流量产生的荷载。数值模拟时先平衡初始地应力,使岩土层在自重作用下完成固结沉降,并将位移值和位移速率重新设置为 0。为保持数值模拟和地铁盾构施工的一致性,先掘进 5 个管片进深,然后在盾尾第一个管片建立一次衬砌,之后每掘进一个管片建立一次衬砌,依次循环至先行左线隧道贯通,然后掘进右线直至双线隧道贯通。在模型纵向($y = 50$ m)设置地表横向沉降监测点 73 个,3 条管线横断面均布置管线横向沉降监测点 73 个,左线($x = -7$ m)和右线($x = 7$ m)隧道中心线地表纵向沉降均布置 60 个监测点。先行左线隧道贯通、后行右线隧道贯通变形模拟云图如图 4 所示。

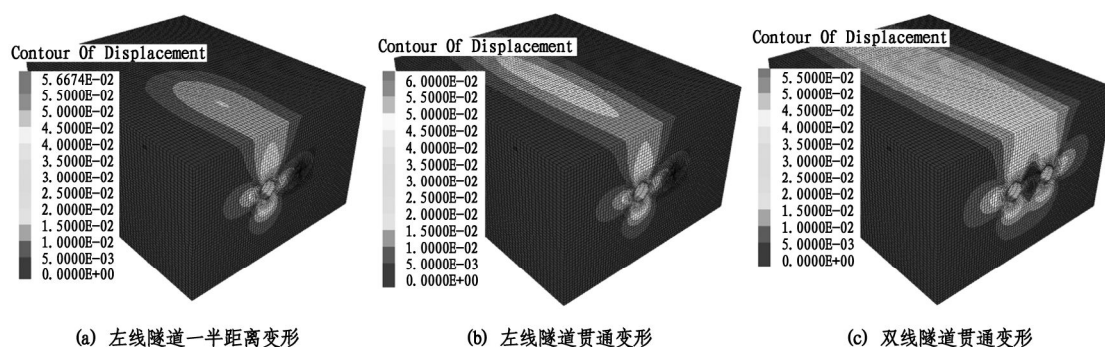


图4 双线隧道模拟云图

3 理论、实测和模拟对比分析

3.1 地表横向沉降分析

Peck^[5]认为地表沉降是在不排水条件下发生的,假定地层损失率在长度上均匀分布,沉降槽的体积等于地层损失的体积,单线盾构掘进地表横向沉降分布为正态分布曲线,Peck 公式为:

$$S(x) = \frac{\pi R^2 \eta}{i \sqrt{2\pi}} \exp^{-\frac{x^2}{2i^2}} \quad (1)$$

$$i = \frac{H}{\sqrt{2\pi} \tan(45^\circ - \varphi/2)} \quad (2)$$

式中 i_f, i_l 分别为先行、后行盾构引起的地表沉降槽宽度系数; η_f, η_l 分别为先行、后行盾构引起的地层损失率, L 为双线隧道中心点间距。

结合本工程实测数据分析,横断面 DBC11 处后行与先行盾构的土体损失比约为 1.25 ~ 1.35。先行左线和双线贯通时横断面 DBC11 处地表横向沉降实测值、理论值和模拟值变化规律如图 5 所示。

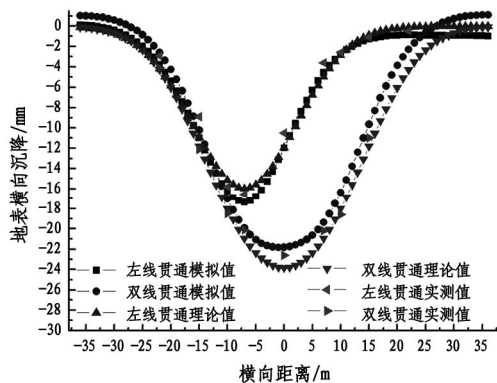


图5 地表横向沉降对比

由图 5 可知,先行左线和双线隧道贯通时横断面 DBC11 处地表横向沉降实测、理论和模拟值随地表横向距离的变化规律基本一致,均以隧道中心线 ($y = -7$ m) 和两隧道中心 ($y = 0$ m) 呈 V 形。

式中 $S(x)$ 为距隧道中心线 x 处的地表沉降, i 为沉降曲线反弯点至盾构中心距离,即地面沉降槽的宽度系数, η 为地层损失率, R 为隧道半径, H 为隧道埋深, φ 为岩土体内摩擦角。

因后行盾构掘进时岩土体受到二次扰动的效应,后行盾构对先行成形隧道产生影响,马可栓^[14]在 Peck 公式基础上总结双线平行隧道先、后行盾构掘进引起的地表变形公式:

$$S(x) = \frac{\pi R^2 \eta_f}{i_f \sqrt{2\pi}} \exp\left[-\frac{(x-0.5L)^2}{2i_f^2}\right] + \frac{\pi R^2 \eta_l}{i_l \sqrt{2\pi}} \exp\left[-\frac{(x+0.5L)^2}{2i_l^2}\right] \quad (3)$$

3.2 地表纵向沉降分析

盾构掘进时对掌子面前后地表沉降均有不同程度的影响,先行左线和后行右线对左右盾构中心线地表纵向沉降也有不同程度的影响,通过实测数据和模拟值总结出盾构机掘进时掌子面前后地表纵向沉降的规律如图 6 所示。

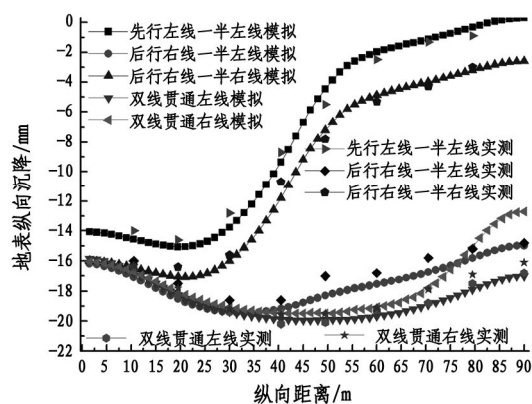


图6 地表纵向沉降对比

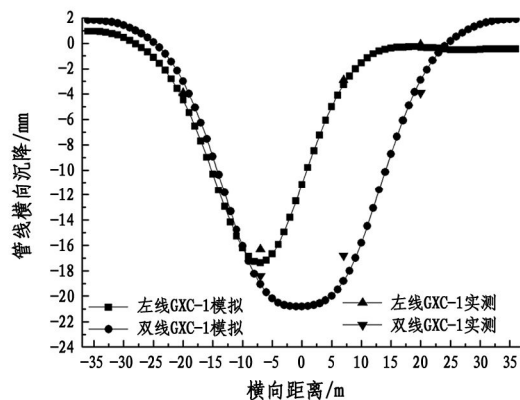
由图 6 分析可知,分别提取左线盾构机和右线盾构机均位于模型纵轴一半位置 ($y = 45$ m) 的纵向沉降数据,先行左线和后行右线盾构引起的左右线盾构中心线上方的地表纵向沉降规律模拟值和实测

值基本一致,即便受地面行车荷载的影响,盾构机前方地表纵向沉降仍然较小,沉降影响范围约为 15 m,发展较为缓慢,甚至在前方 35 m 以外地面出现隆起的情况。盾构机继续掘进,地表纵向沉降逐渐增加至趋于稳定,盾构机后方沉降影响范围约为 20 m,发展较为迅速,因此盾构机掘进对前后地表纵向沉降影响范围约为 35 m,距离盾尾较远位置的沉降已趋于稳定。地表纵向沉降主要经历了缓慢发展、迅速发展和逐步稳定三个阶段。在地面行车荷载作用下,双线盾构均贯通后,地表纵向沉降略有增加,最终稳定在 19.91 mm,小于 30 mm 设计要求值,满足安全要求。

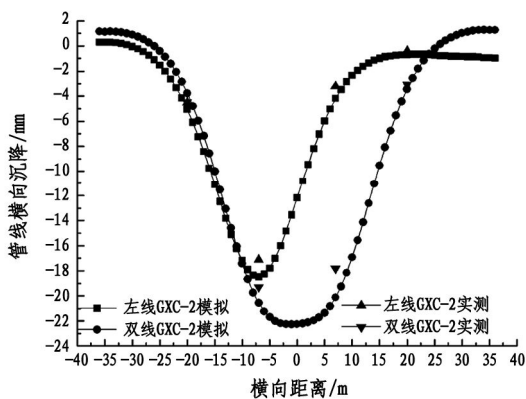
3.3 管线沉降分析

盾构机掘进引起的 GXC-1、GXC-2、GXC-3 的沉降模拟值和实测值变化规律如图 7 所示。

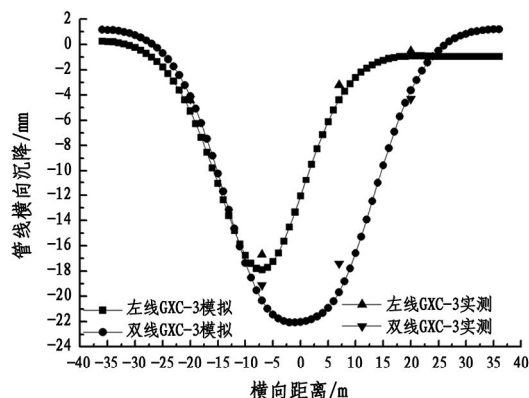
由图 7 可知,盾构机掘进引起的 GXC-1、GXC-2、GXC-3 的沉降模拟值和实测值变化规律基本一致,均以隧道中心线($y = -7$ m)和两隧道中心($y = 0$ m)近似呈 V 形。GXC-2 所在车道受石大路和科苑横路交叉影响,其沉降值是最大的。管线沉降模拟值和实测值均小于 30 mm,满足要求。



(a) GXC-1沉降对比



(b) GXC-2沉降对比

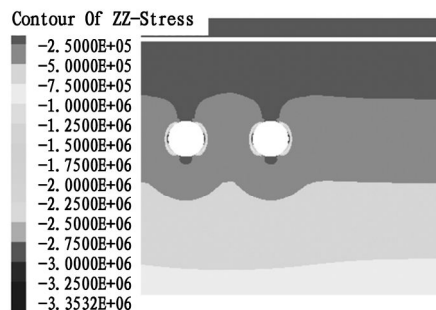


(c) GXC-3沉降对比

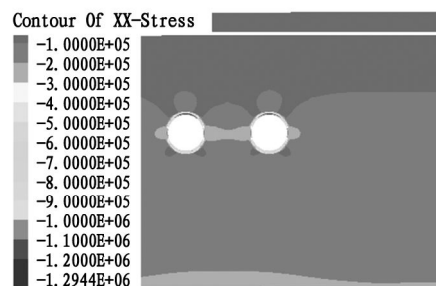
图 7 管线沉降对比

3.4 隧道管片应力分析

GXC-2 所在横断面沉降量最大,选择该断面的竖向 ZZ 和水平 XX 应力图如图 8 所示。由图可知在管片水平和竖直方向存在应力集中现象,且水平应力(3.35 MPa)大于竖向应力(1.29 MPa)。管片顶部承受较大的压应力,管片水平左右两侧承受拉应力,均满足规范和设计要求。



(a) ZZ应力云图



(b) XX应力云图

图 8 隧道管片应力云图

分别提取盾构管片应力和变形模拟值,对比现场实测数据绘制隧道最终变形断面图如图9所示。由图9可知,盾构管片拼装后受地层沉降、应力等影响,由设计“圆形”断面近似变形为“椭圆形”断面,竖向呈现挤压变形,水平呈现拉伸变形。

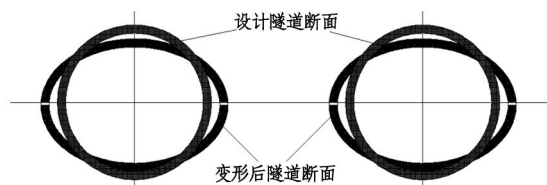


图9 隧道管片变形图

4 结论

地铁左右线隧道间距相对较小时,单线和双线盾构掘进引起的地表横向沉降均近似呈V形变化,由于管线近似垂直隧道中心线,管线横向沉降规律和地表横向沉降规律保持一致。

盾构机前方地表纵向沉降较小,且较远位置会出现隆起,沉降影响范围约为15 m,盾构机后方地表纵向沉降逐渐增加至趋于稳定,盾构机后方沉降影响范围约为20 m。地表纵向沉降主要经历了缓慢发展、迅速发展和逐步稳定三个阶段,地面行车荷载使得地表纵向沉降量略有增加。

管片水平和竖直方向存在应力集中现象,且水平应力大于竖向应力,管片顶部承受较大的压应力,管片水平左右两侧承受拉应力。

盾构管片拼装后受地层沉降、应力等影响,由设计“圆形”断面近似变形为“椭圆形”断面,竖向呈现挤压变形,水平呈现拉伸变形。

参考文献:

- [1] 邓皇适,傅鹤林,史越,等. 盾构隧道曲线段掘进引发临近地下管线变形分析[J]. 中南大学学报(自然科学版),2022,53(8):3008-3020.
- [2] 王鹏,谭庆振,潘随伟,等. 上软下硬地层地铁盾构施工地表沉降模拟分析[J]. 河南城建学院学报,2024,33(1):49-54.
- [3] 王鹏,高铭,骆方圆,等. 冻结管偏斜对煤矿井冻结温度场发展影响[J]. 蚌埠学院学报,2024,13(5):90-94.
- [4] 刘波,李烜,韩彦辉. 城市地下空间盾构隧道穿越工程研究综述[J]. 施工技术(中英文),2024,53(17):35-46.
- [5] Peck R B. Deep excavations and tunneling in soft ground [C]. Proceedings of the 7th International Conference on Soil Mechanics and Foundation Engineering, Mexico, 1999: 225-290.
- [6] 冯国辉,肖明清,倪建中,等. 盾构下穿引起上覆管线变形的简化计算方法[J]. 湖南大学学报(自然科学版),2024,51(5):68-75.
- [7] 史江伟,王金朴. 盾构开挖面失稳下既有管线三维变形特性研究[J]. 地下空间与工程学报,2025,21(3):1057-1065.
- [8] 任旭东,张凤凯,丁万涛,等. 基于神经网络的泥水盾构地表沉降预测与掘进参数优化[J]. 科学技术与工程,2025,25(5):2090-2099.
- [9] 席培胜,周冠宇,张军. 盾构隧道伴行对高压管线变形的影响分析[J]. 四川轻化工大学学报(自然科学版),2024,37(1):79-85.
- [10] 王忠昶,魏新平,王川,等. 地铁盾构隧道施工对临近管线的影响分析[J]. 大连交通大学学报,2018,39(4):84-87.
- [11] 王忠昶,常龙,夏洪春. 地铁盾构双线隧道施工地层变形及衬砌结构应力数值分析[J]. 大连交通大学学报,2017,38(1):81-85.
- [12] 周绍宾,郭佳奇,徐冲. 盾构隧道下穿交叉市政管线地表沉降分析与控制[J]. 城市轨道交通研究,2021(8):41-46.
- [13] 魏纲,庞思远. 双线平行盾构隧道施工引起的三维土体变形研究[J]. 岩土力学,2014(9):2562-2568.
- [14] 马可栓. 盾构施工引起地基移动与近邻建筑保护研究[D]. 武汉:华中科技大学,2008.
- [15] 宫亚峰,王博,魏海斌,等. 基于Peck公式的双线盾构隧道地表沉降规律[J]. 吉林大学学报(工学版),2018,48(5):1411-1417.
- [16] 潘苗,江玉生,葛振义. 盾构施工引起横向沉降槽宽度分析[J]. 科学技术与工程,2015,15(28):196-199.

责任编辑:汤庆国

光纤法布里-珀罗传感器解调算法发展

赵小明¹, 吴凯鹏¹, 周程^{1*}, 楼亮亮¹, 田佳峻²

(1. 台州学院 人工智能学院, 浙江 台州 318000;
2. 哈尔滨工业大学(深圳) 电子与信息工程学院, 广东 深圳 518055)

摘要:为梳理光纤法布里-珀罗传感器解调算法的发展现状,从信息利用方式和建模范式角度,综述了强度解调、频域解调、谱峰追踪解调和深度学习解调方法。分析表明,传统解调方法具有物理模型清晰、实现方式成熟等特点,但在低信噪比、宽动态范围、多腔串扰和复杂环境条件下易受有限特征提取与模型假设限制。深度学习解调方法能够利用全谱信息建立非线性映射,在串扰抑制、多参数解耦和复杂工况适应方面具有潜力,但仍存在样本依赖、泛化能力不足和物理一致性欠缺等问题。未来研究可重点围绕物理模型——数据驱动融合鲁棒解调、有限样本与稀疏采样智能解调,以及采集硬件与解调算法协同优化展开。

关键词:光纤法布里-珀罗传感器;解调算法;深度学习;数据驱动;智能解调

中图分类号:TP212.14;TN253

文献标识码:A

文章编号:2095-297X(2026)05-0084-12

Development of Demodulation Algorithms for Fiber Fabry-Perot Sensors

ZHAO Xiaoming¹, WU Kaipeng¹, ZHOU Cheng^{1*}, LOU Liangliang¹, TIAN Jiajun²

(1. School of Artificial Intelligence, Taizhou University, Taizhou, 318000, Zhejiang;
2. School of Electronic Information and Engineering, Harbin Institute of Technology, Shenzhen, 518055, Guangdong)

Abstract: To review the development of demodulation algorithms for fiber-optic Fabry-Perot sensors, intensity-based demodulation, frequency-domain demodulation, spectral-peak tracking demodulation, and deep-learning-based demodulation were summarized from the perspectives of information utilization and modeling paradigm. It is shown that traditional methods are characterized by clear physical models and mature implementation, but are limited by finite-feature extraction and model assumptions under low signal-to-noise ratio, wide dynamic range, multi-cavity crosstalk, and complex environments. Nonlinear mappings can be established by deep-learning-based demodulation using full-spectrum information, showing potential in crosstalk suppression, multi-parameter decoupling, and adaptation to complex conditions. However, limitations related to sample dependence, insufficient generalization, and limited physical consistency remain. Future research should be focused on robust demodulation based on physics-model and data-driven fusion, intelligent demodulation with limited samples and sparse sampling, and co-optimization of acquisition hardware and demodulation algorithms.

Key words: fiber-optic Fabry-Perot sensor; demodulation algorithm; deep learning; data-driven; intelligent demodulation

法布里-珀罗传感器(Fabry-Perot, F-P)因体积小、抗电磁干扰能力强、耐腐蚀、耐高温等优势,已广

泛应用于恶劣环境下温度、加速度、应变等多参量检测^[1-3],并在医疗、气体检测、水声探测、电力监测等

收稿日期:2026-04-10 * 通讯联系人

基金项目:浙江省科技计划项目(2024C01022);台州市科技计划项目(24nya19)。

作者简介:赵小明(1964—),男,浙江台州人,教授,硕士。E-mail:229191720@qq.com

领域受到持续关注^[4-9]。受结构设计、材料特性及加工工艺等因素制约^[10-12],F-P传感器的检测性能仍受到灵敏度和测量精度受限、干涉条纹可辨识度不足等问题影响。面对高精度、多参数及复杂环境测量需求,光纤F-P传感器在抗噪性能、多物理量协同感知等方面仍存在不足,其工程应用范围在一定程度上受到限制。深度学习在信号处理与特征提取领域展现出较强的建模能力,尤其在复杂非线性关系表征和自适应学习方面具有显著优势,对于提高解调精度具有潜在价值。将深度学习引入F-P传感器解调过程,有助于提升解调精度与环境鲁棒性,增强复杂工况下的适应能力。本研究综述传统光纤F-P传感器解调算法的发展现状,分析深度学习在解调中的应用进展与演化趋势,为F-P传感器解调技术的发展提供参考。

1 传感器工作原理

光纤F-P传感器的典型结构如图1所示。通常包含两反射面,并由一定长度的介质隔开^[13]。当一束光入射,光线在两个反射面之间进行多次反射,最终反射回的干涉光强会受到反射面反射率、腔长以及光波波长等因素的调制。其工作原理的相应表达式^[14]为:

$$I \approx [1 - \cos(\frac{4\pi}{\lambda}nL)] I_0 \quad (1)$$

式中 I_0 为入射光光强, λ 为光波波长, L 为F-P腔的腔长, n 为F-P介质的折射率。

从公式(1)可看出输出光强中蕴含着F-P腔的有效光程信息,后续解调方法的介绍均是围绕着如何去解调出有效光程展开。由于参数的微小变化会对光强产生显著影响以及干涉的周期性多值性。因此,需采用曲线拟合等处理方法并结合相位解调等相关解调算法,来获取精准的腔长信息。

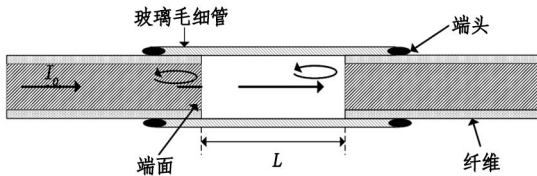


图1 F-P传感器结构示意图

2 强度解调算法

2.1 单波长解调算法

由图2可知,波峰与波谷附近存在光强变化灵敏度低、变化特征不清晰等问题。单波长解调算法的核心在于将工作点设置于光强变化最敏感的位置,

通常选取光谱中斜率最大 Q 点^[15]作为初始工作点。在 Q 点附近 $\lambda/4$ 的范围内,腔长 L 与光强 I 可近似视为一一对应的线性关系

$$I = \alpha I_0 L \quad (2)$$

式中, α 为灵敏度系数。

该算法具有灵敏度高、响应速度快以及硬件实现相对简便等优点^[16]。该算法属于一种“以点代面”的测量策略,干涉信息被显著压缩,存在若干内在局限。其一,其线性工作范围受到约束。公式(2)成立的前提是相位变化小于 $\pi/2$,对应的腔长变化不得超过 $\lambda/4$,这表明该方法在获得较高灵敏度的同时,测量范围受到明显约束。其二,待测腔长信息与光源强度波动共同耦合于测量值 L 中,实际应用对光源稳定性和光路稳定性具有较强依赖,复杂环境或恶劣工况下的测量可靠性因而受到限制。其三,外界环境变化可引起 L 的漂移,并导致干涉光谱基准位置发生偏移。图2所示 Q 点位置随之变化,表明该方法缺乏自校准能力^[17]。

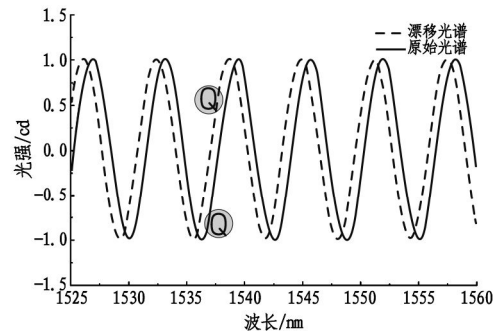


图2 波长漂移对 Q 点的影响

2.2 双波长解调算法

双波长解调算法的核心思想在于利用两个不同波长(λ_1 、 λ_2)的干涉信号构建差分机制,以补偿单波长方法的内在缺陷。其基本原理为:在 Q 点附近,两波长对应光强变化对腔长响应的灵敏度符号相反,通过构造两路信号的相位差,并采用DCM^[18]、atan^[19]及椭圆拟合方法^[20]进行相位提取,实现共模分量抑制。该方法降低了系统对光强绝对稳定性的依赖,对光强波动具有一定抑制作用,同时扩展了测量动态范围^[21],突破了 $\lambda/4$ 的限制,实现了对腔长变化的连续跟踪,理论上可覆盖整个干涉周期。

双波长解调虽在一定程度上弥补了单波长算法的不足,但仍引入了新的误差来源。如图3所示,实验过程中通常需要配置两个光源,二者功率波动差异及响应灵敏度偏差均可能被误判为相位变化。波长间隔的选取同样构成关键约束:间隔过小,两路信

号相关性过高,有效区分难度增加;间隔过大,干涉条纹相关性下降,相位差提取精度受损。针对上述问题,后续研究开始引入第三个波长,以改善双波长方法的性能局限。

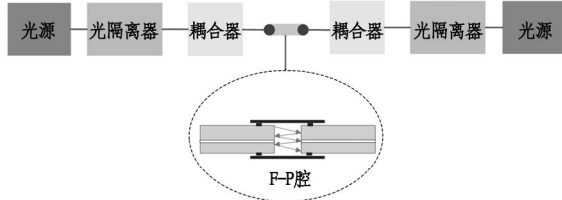


图3 双波长解调算法实验结构图

2.3 三波长解调算法

三波长解调算法的提出,主要针对双波长算法对两路探测器响应特性严格匹配的高度依赖。该条件在工程实践中通常难以满足。其核心思路在于引入第三个波长^[22],增加系统信息冗余,使硬件在难以实现精确匹配时,仍可通过算法完成待测信号的解耦提取^[23]。实验装置如图4所示。三路干涉信号经窄带滤波后,各路光强可表示为:

$$I_k = B \sin\left(\frac{4\pi nL}{\lambda_k}\right) \frac{4\pi n s(t)}{\lambda_k} \quad (3)$$

式中, I_k 为第 k 路信号光强, B 为常数, λ_k 为第 k 个滤波器的工作波长, $s(t)$ 为待测信号。

依据公式(3),三路信号可构成三个线性方程,并用于联合求解同一未知量。该结构提高了系统自由度,可对三路探测器之间的响应差异进行自治校正,降低了系统对硬件一致性的依赖。但该算法仍具有明确的适用边界。信号间的近似关系建立在公式(3)中 $\sin x$ 的泰勒展开,当 x 较小时忽略高阶项,可得到 $\sin x \approx x$ 。当信号幅值 $s(t) < 25$ nm 时,近似的相对误差小于 1%;但超过该阈值,近似条件不再满足,算法性能将明显下降,误差过大时甚至失效^[24]。这一边界表明,三波长方法本质上以数学近似换取硬件匹配要求的放宽,但有效性仍依赖于模型精度及适用条件的严格满足。

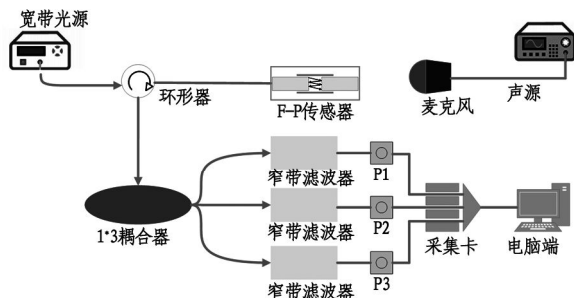


图4 三波长小信号解调实验结构

3 频域解调算法

3.1 快速傅里叶变换法

快速傅里叶变换(Fast Fourier Transform, FFT)不同于强度法提取少数特征点,而是利用整个波长域的干涉光谱,并通过傅里叶变换将其映射至频率域^[25],从而将腔长信息的处理由相位测量转化为频率估计。实际处理中,通常先对光谱进行三次样条插值^[26]或FFT重采样^[27],并结合小波分解^[28]提取低频包络,变换后的主峰位置 f_{peak} 与 L 满足如下关系:

$$L = \frac{f_{\text{peak}} c}{2\nu} \quad (4)$$

该映射关系拓展了干涉光谱的处理路径。其一,各波长点所含微弱信号可在频谱主峰处实现能量累积,有利于提高信噪比并提升光谱信息利用率。其二,测量对象由光强绝对值转变为频谱峰值频率,频率量对光源功率波动、光纤损耗等共模扰动具有较强抑制能力,因而方法整体鲁棒性较强。但该算法仍存在一定问题。首先,受有限光谱扫描范围限制,FFT解调的频率分辨率有限,制约了腔长分辨率 ΔL 的提升。通常, ΔL 与光谱扫描范围之间满足如下关系:

$$\Delta L = \frac{\lambda_1 \lambda_2}{2n(\lambda_2 - \lambda_1)} \quad (5)$$

以 1525 ~ 1575 nm 波段,即 50 nm 带宽为例,受有限光谱扫描范围限制,解调的腔长分辨率约为 24 μm 。经波数域重采样、零填充和频谱细化处理后,频谱主峰位置的估计精度可进一步提高,在较高信噪比条件下可实现亚微米量级的腔长估计。需要指出的是,零填充主要提高频谱采样密度,并不改变由光谱扫描范围决定的真实频率分辨率。若需进一步提高真实分辨率,通常需要扩大光谱扫描范围;而增加有效采样点数和采用频谱细化处理则有助于改善峰值定位精度,但会增加采集或计算时间,并降低系统动态响应能力^[29]。此外,有限采样长度、非整周期截断以及由波长域等间隔采样转换至波数域时的非均匀采样问题,均易引起频谱泄漏,使变换后出现旁瓣、主峰展宽和峰值漂移等现象。加窗法可在一定程度上抑制频谱泄漏^[30],但通常伴随频率分辨率下降。已有研究表明,当信噪比高于 18 dB 时,可在一定程度上避免干涉级次跳变。当误差小于 0.05% 时,通常要求信噪比超过 20 dB。当存在多个 F-P 腔串扰,频谱中可能出现多个幅值接近的峰值,导致主峰识别困难并引起腔长误判^[31]。

3.2 稀疏傅里叶变换法

干涉光信号频谱通常具有显著稀疏性^[32],有效信息主要集中于少数频率分量,而多数频谱分量能量接近于零^[33-34]。这一特性表明,FFT对全频谱进行统一计算时存在较多冗余操作。稀疏傅里叶变换(Sparse Fourier Transform, SFT)的核心思想在于利用频谱重排、窗函数滤波、频域降采样以及循环定位与估值等步骤,快速定位频谱中幅值最大的 k 个傅里叶系数,并据此提取与腔长对应的频率信息,实现信号解调。SFT的主要优势在于减少了全频谱计算带来的时间开销,从而提高了解调效率^[35]。当信号长度较大且频谱稀疏性更强时,该优势更为突出。该方法的适用性仍以信号满足稀疏性假设为前提。SFT实现过程相对复杂,通常需要针对具体问题设计采样策略及信号恢复算法^[36]。因此,该方法更适用于频谱稀疏特征明显的应用场景,其优势主要体现在较高的计算时效性。

4 谱峰解调算法

4.1 单峰解调算法

单峰法的核心思想是跟踪反射光谱中某一个特定波峰的漂移量,通过检测该特征点的波长变化来反推出 F-P 腔长的相对变化,该方法的表达式为:

$$L_d = \frac{m}{2} \Delta \lambda_m, m = 1, 2, \dots \quad (6)$$

式中, L_d 为相对腔长变化量, m 为该点对应的干涉级次, $\Delta \lambda_m$ 为所选波峰的漂移量。

该方法实质上是将整个干涉光谱压缩为单一特征点的波长信息,并通过跟踪该特征点的位置变化实现测量。从方法本质上看,该算法与单波长法的核心思想一致,区别仅在于其测量对象由光强转变为波长。

由于光谱峰值位置具有较高的波长定位精度,且特征点跟踪算法实现较为简便,单峰法能够实现腔长变化的快速、高灵敏度检测。在高信噪比和高质量干涉光谱条件下,其测量精度通常优于强度解调方法。然而,该方法对目标波峰的连续可辨识性具有较强依赖。当腔长变化引起的波峰漂移量超过相邻波峰间距的一半时,目标波峰与相邻波峰可能发生混淆,进而导致干涉级次判定歧义。以 1440 ~ 1675 nm 光谱范围为例,受波峰可追踪范围限制,单峰法的动态测量范围上限约为 0.8375 μm ^[37]。此外,该方法高度依赖峰值定位准确性,噪声干扰、光谱畸变或多腔串扰均可能引起峰值偏移,并向腔长解调结果引入附加误差^[38]。因此,单峰法更适用于

信噪比较高、谱峰结构稳定且腔长变化范围较小的静态或准静态测量场景。为突破单一特征点跟踪限制,需进一步引入多特征点信息进行联合约束。

4.2 双峰解调算法

该算法通过利用两个波峰 λ_m 与 λ_{m+q} 之间的波长间隔直接计算绝对腔长^[39-40],从而摆脱了对干涉级次的依赖。在计算过程中,无需预先获得 m 的具体数值,因此相较于单峰法,该方法能够实现绝对腔长测量,其动态范围在理论上也不再受干涉周期限制,从而突破了单峰法的瓶颈。然而,该算法仍存在相应缺陷。如公式(7)所示:

$$\frac{L_1}{L} \approx \frac{\Delta \lambda}{\lambda_{m+q} - \lambda_m} \cdot \sqrt{1 + \left(\frac{q \lambda_m}{\lambda_{m+q} - \lambda_q} \right)} \quad (7)$$

腔长测量误差 L_1 与两个波峰的定位误差 $\Delta \lambda$ 满足公式(7)。由此可知,双峰法的解调精度主要受波峰间距和峰值定位误差共同影响。当两个波峰在波长域上距离较近时,波长差值较小,峰值定位误差对腔长计算结果的放大作用将显著增强。因此,若为提高局部跟踪灵敏度而选取相邻波峰,可能导致更明显的误差放大效应。与单峰法相比,双峰法虽然能够实现绝对腔长测量,但其解调误差会受到波峰间距项的进一步放大,导致整体测量精度下降^[41]。为在保留绝对测量能力的同时抑制误差放大,有必要进一步引入干涉级次信息作为约束。

4.3 双峰级次腔长解调算法

双峰级次腔长解调算法是针对双峰法中误差放大问题提出的改进方法,其核心思想是在双峰法的基础上进一步引入干涉级次信息作为约束,通过干涉级次与波峰位置的联合求解,在保留绝对测量能力的同时抑制误差放大效应。该方法在光谱中选取一个待跟踪的谱峰,并结合另一个与其相隔 q 个周期的谱峰,联合计算对应的干涉级次 m ,从而实现腔长解调。干涉级次关系为:

$$m = \frac{q \lambda_{m+q}}{\lambda_m - \lambda_{m+q}} \quad (8)$$

获得干涉级次后,可进一步通过公式(7)求得绝对腔长。该方法实现了单峰法高精度特性与双峰法绝对测量的有机结合。相关实验结果表明,在 $L = 100 \mu\text{m}$ 、中心波长为 1568 nm、峰值定位误差为 30 pm 的条件下,该方法的解调误差约为 2 nm,而双峰法在相同条件下的误差约为 0.345 μm ^[42]。对比可见,该方法有效体现了干涉级次信息在抑制误差放大方面的关键作用。然而,该方法也引入了新的局限性。由于干涉级次信息 m 的获取依赖于相关公式中的取整操作,因此对两个波峰的定位误差提出

了较高要求,必须保证其足够小,才能避免干涉级次发生跳变。

4.4 传统算法总结与局限性

表 1 从算法本质、核心贡献、固有代价、关键参数(定量指标)和适用场景五个维度,对传统解调方法进行了比较。从整体上看,传统方法的发展并不是单一性能指标的线性提升,而是在测量精度、动态范围、抗干扰能力、实时性和实现复杂度之间不断权

衡的过程。算法的演进较为典型地体现了由单一特征向多特征约束发展的逻辑:测量参考系由固定特征点依赖,逐步过渡到特征点间关系约束,再到峰值信息与干涉级次的联合解算。与之对应,信息利用方式也由单一特征点波长信息,扩展到双特征点间距信息,并进一步发展为特征点信息与级次信息的联合求解。信息维度的增加在一定程度上缓解了前序方法的约束。

表 1 传统方法对比表

方法	算法本质	核心贡献	固有代价	关键参数(定量指标)	适用场景
单波长法	固定点测量	建立了光强 - 腔长的直接线性映射,实现了高灵敏度快速测量	线性区硬约束,对光源波动敏感,外部环境易干扰待测信号	线性区上限: $\lambda/4$, Q 点取灵敏度最高的点	静态、高稳定性环境下的微位移测量
双波长法	差分测量	摆脱对 Q 点的依赖,消除共模干扰	硬件抑制性要求高,通道失配存在误差	动态范围大于一个干涉周期	中等动态范围、能够接受硬件复杂度的场合
三波长法	冗余测量	通过信息冗余降低对硬件匹配的依赖	小信号具有 25 nm 的幅度上限,超过后误差增大	小信号有效范围 < 25 nm,超过失效	微弱振动信号的高灵敏度测量
FFT	全谱分析	利用全光谱信息,实现信噪比提升	分辨率受限于扫描范围,存在频谱泄露	信噪比要求 > 18 dB,避免级次跳变, > 22 dB 保证 0.05% 精度	单腔传感器、信噪比 > 20 dB 的通用测量场景
SFT	稀疏采样	计算效率显著提升	以信号稀疏性为前提,实现复杂	运算时间低	实时性要求高,频谱稀疏的大规模多通道解调
单峰法	特征点追踪	高分辨率、直观快速,实现高精度的相对测量	动态范围受限,对噪声敏感	动态范围上限: $0.8375\text{ }\mu\text{m}$	静态、小范围腔长变化测量
双峰法	多特征点解算	摆脱对先验级次的依赖	误差放大机制,精度相比单峰法低	误差约为 $0.345\text{ }\mu\text{m}$	需要绝对腔长测量,接受较低精度的场景
双峰级次腔长法	联合解算	融合单峰精度与双峰绝对测量能力,有效抑制了误差放大	级次取整可能会引入 ± 1 的整数误差,对峰值定位精度要求高	解调误差: 仿真 < 0.7 nm , 实验 < 1.4 nm 。	高精度绝对腔长测量、对鲁棒性要求较高的场景

从上述比较可以看出,传统方法的改进并未从根本上消除误差,而是改变了误差来源及其表现形式。方法的每一次演进,往往是在缓解已有约束的同时引入新的适用条件或误差来源。多数传统方法仍依赖人工选取的有限特征或显式构造的中间量,属于有限信息驱动的测量模式。在这一模式下,干

涉光谱中的大量信息尚未被充分纳入解调过程。信息利用的不充分,使得相关方法在测量精度^[43]、抗干扰能力^[44]、动态范围^[45]和适用场景^[46]等方面存在较为明确的性能边界。当信噪比降低或光谱畸变削弱特征点的可辨识度时,依赖有限特征提取的解调过程容易受到影响,测量稳定性与结果可靠性也

会下降。

上述局限促使研究者探索不同于有限特征驱动的解调路径。面向高精度、高频动态测量^[47]和多参数感知^[48]等应用需求,突破离散特征驱动的测量模式逐渐成为方法演进的趋势之一。以深度学习为代表的智能解调方法,不再依赖少数特征点的提取与追踪,而是尝试利用全谱信息进行建模,并学习光谱与待测物理量之间的复杂非线性关系,从而展现出弥补传统方法短板的可能。同时传统算法仍是理解 F-P 干涉与解调机理的重要基础,其理论价值和可解释性也为后续智能解调方法的设计提供了重要支撑。

5 结合神经网络的解调算法

5.1 从物理模型到数据模型:深度学习解调的范式转变

基于上述分析,传统解调方法的局限并不只是某一具体算法精度不足,而是与其有限特征驱动和显式物理建模的基本框架有关。强度解调依赖少数波长点处的光强变化,频域解调虽利用更完整的频谱信息,但仍受到频谱分辨率和信号模型假设的制约;谱峰追踪方法则始终依赖特征峰的稳定可辨。当光谱受到复杂噪声污染、多参数耦合或剧烈环境变化影响时,人工构造特征往往难以充分表征信号中的有效信息,基于理想假设建立的物理模型也难以准确刻画真实场景中的输入-输出映射关系。由此,深度学习方法逐渐被引入 F-P 传感器解调领域。

深度学习解调体现出一种范式层面的转变。其核心在于构建多层神经网络,直接以原始干涉光谱为输入、以待测物理量为输出,使模型能够自动学习光谱特征与物理量之间的非线性映射关系。针对传统解调中的若干核心瓶颈,这一方法提供了不同的解决思路。在信息利用方面,深度学习无需依赖人工特征筛选,能够从全光谱数据中自适应提取对解调任务具有判别性的特征;在建模方式方面,深度学习通过数据驱动方式直接拟合输入与输出关系,在一定程度上减少对线性化或小信号近似的依赖;在抗干扰方面,深度学习可借助覆盖不同工况的样本训练,学习待测信号与环境干扰之间的统计差异,从而提高复杂环境下的解调鲁棒性。

5.2 前馈神经网络的应用

前馈神经网络(Feedforward Neural Network, FNN)是深度学习基础模型之一,具有结构简单、训练过程相对成熟等特点,适用于学习输入到输出之间的静态映射关系。Wu 等将 FNN 应用于复用 F-P

传感器解调,针对多传感器串扰问题提出了数据驱动的解决方案^[49]。复用 F-P 传感器的解调本质上属于由混叠光谱到多个传感器应变量的多元回归问题。由于多个 F-P 腔的干涉信号叠加于同一光谱中,传统 FFT 方法在腔长接近时易出现频谱峰重叠,导致各传感器信号难以有效分离。FNN 的多输入多输出结构与该任务具有良好适配性:输入层接收离散化光谱数据,输出层同时给出多个传感器的应变值,隐藏层则负责学习从混叠光谱中分离各传感器贡献的非线性映射关系。该端到端学习方式绕过了传统方法中先分离后解调的处理流程,从原理上减少了串扰误差的累积。

该研究利用已知应变组合条件下采集的光谱数据构建训练集,避免了传统方法中对全组合工况进行逐一校验的繁琐过程。实验结果表明,当训练数据能够充分覆盖应变范围及其组合方式时,模型对采样间隔变化表现出较低敏感性,并具有较好的鲁棒性。然而,模型性能显著依赖于训练数据分布:当测试样本超出训练范围时,其输出结果可能与真实值产生较大偏差。与 FFT 方法的实验对比结果显示,FNN 在解调过程中表现出更优的线性度和更低的串扰水平。在传感器间串扰较为严重的情况下,FFT 可能无法识别被掩盖的传感器峰值,而 FNN 仍能够实现有效解调,表明数据驱动方法在复杂光谱场景下具有较大应用潜力。

5.3 DBN 与集成学习

深度信念网络(Deep Belief Network, DBN)由多个限制玻尔兹曼机堆叠而成,结构如图 5 所示,其核心优势在于无监督预训练与有监督微调相结合的训练机制,使其在训练数据有限的情况下仍能有效提取光谱中的深层特征。Wang 等将 DBN 引入 F-P 温度传感器解调,并进一步融合集成学习思想以提升解调性能^[50]。F-P 传感器的光谱数据具有高维、强相关、含噪声的特点,而 DBN 的逐层无监督预训练能够从原始光谱中自动学习波长漂移等关键特征,无需人工设计特征提取器;其深层结构能够建模光谱演变与温度变化之间的复杂非线性关系。这种自动特征学习能力使 DBN 在处理多参数耦合或噪声干扰时具有天然优势。

然而,单一 DBN 模型可能因初始化、训练样本选择等因素产生偏差和方差。为克服这一局限,该研究采用 Stacking 架构构建多模型融合方法:由多个采用不同激活函数的 DBN 作为基学习器,各自独立对光谱数据进行解调,将其输出作为元特征输入至 BP 神经网络进行融合。不同激活函数所形成的

模型差异性有效降低了单模型的偏差与方差,元学习器进一步整合各 DBN 提取的互补特征,从而获得更接近真实值的解调结果。实验结果表明,基于 Stacking 的多模型融合方法在解调精度和鲁棒性上显著优于单一 DBN 模型,展现了集成学习思想在 F-P 解调中的成功应用。这一策略的本质是用多个“弱学习器”的集体智慧弥补单一模型的局限性,为提升深度学习解调的可靠性提供了有效途径。

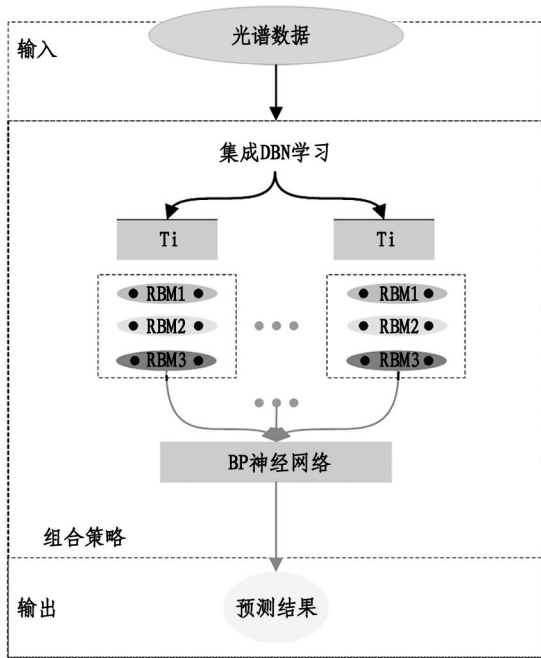


图5 DBN处理模型

5.4 CNN + LSTM 模型解调方法

卷积神经网络与长短期记忆网络的混合模型 (Convolutional Neural Network and Long Short-Term Memory, CNN-LSTM) 结合了卷积神经网络在局部特征提取方面的优势与长短期记忆网络在序列长时依

赖建模方面的能力,适用于处理具有局部相关性与长距离依赖关系的一维光谱序列。Wei 等将 CNN-LSTM 应用于 F-P 传感器的光谱解调^[51],模型架构如图 6 所示。模型绕过了传统方法中对峰值波长或腔长等中间量的显式求解,实现了从归一化光谱到外部物理量的端到端回归。由于光谱中各干涉特征点之间存在长距离依赖关系,传统峰值追踪算法在离散采样或噪声干扰下易丢失关键信息,傅里叶变换等方法则需先计算腔长再转换为应变,存在多步误差累积。CNN-LSTM 的级联结构与该任务具有较好的适配性:前端卷积层提取光谱中的局部波形特征并进行降维压缩,后端 LSTM 层捕捉光谱随应变变化的全局趋势,从而在不引入中间物理量的前提下建立光谱与应变之间的映射关系。该端到端学习方式简化解调流程,有效减少了传统串行处理带来的误差传递。

该研究利用悬臂梁应变平台采集覆盖一定应变范围的光谱数据构建训练集,减少了对手动寻峰和显式腔长计算等中间步骤的依赖。实验结果表明, CNN-LSTM 模型在测试集上具有较高的拟合优度与较低的测试误差,其解调精度相较于峰值追踪、快速傅里叶变换及互相关等传统方法有明显提升。然而,深度学习方法依赖于训练数据对应变范围的充分覆盖,模型性能受限于数据分布;当测试样本超出训练应变区间时,其外推能力尚需进一步验证。与单独使用卷积神经网络或长短期记忆网络相比, CNN-LSTM 在解调精度与训练成本之间取得了较好的综合平衡。在低分辨率光谱条件下,该模型仍能保持较好的性能,表明其在低成本、高精度光纤传感解调系统中具有一定的应用潜力。

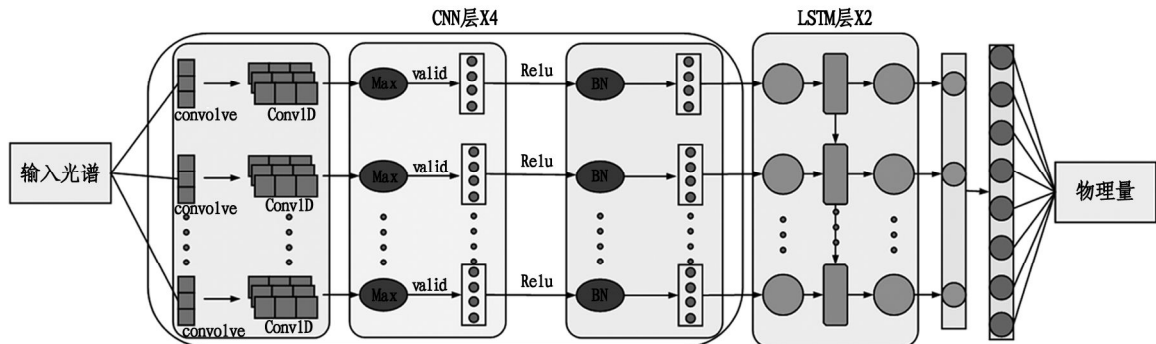


图6 CNN + LSTM 处理模型

5.5 物理信息神经网络解调方法

物理信息神经网络 (Physics-Informed Neural Network, PINN) 是一类将物理规律引入数据驱动模型的深度学习方法,能够在标注数据有限的条件下,通过物理约束提升模型的泛化能力和预测合理性。对于存在多参数耦合和非线性交叉敏感的 F-P 传感器, PINN 的优势在于能够将传感机理作为约束嵌入模型训练过程,从而减弱纯数据驱动模型对大规模标注样本的依赖。Li 等将动态物理权重分配与双分支解耦架构引入 PINN,提出 DPWA-DualPINN 模型,并将其应用于光纤布拉格光栅-法布里珀罗 (FBG-FP) 级联传感器的温度-压力双参数同步解耦^[52],模型架构如图 7 所示。通过物理约束与网络结构设计相结合,改善了传统方法在非线性交叉敏感和串扰抑制方面的不足。双参数解调属于由波长

漂移量到温度、压力变化量的多输出回归问题。在该结构中,FBG 主要用于表征温度变化,FP 则同时受到温度和压力影响,且两类参量对 FP 波长漂移的影响可能存在方向差异。因此,传统灵敏度矩阵方法容易受到线性假设、矩阵病态以及参数串扰等问题的限制。DPWA-DualPINN 与该任务的适配性主要体现在两个方面。其一,动态物理权重分配机制将物理方程中的相关系数设置为可学习参数,并引入可学习的全局权重,以自适应平衡数据损失和物理损失,从而增强网络对非线性关系的刻画能力。其二,双分支解耦架构分别构建温度和压力对应的物理子空间,通过特征隔离降低梯度冲突和特征串扰,并在输出层前引入参数共享的注意力机制进一步提高解耦效果。

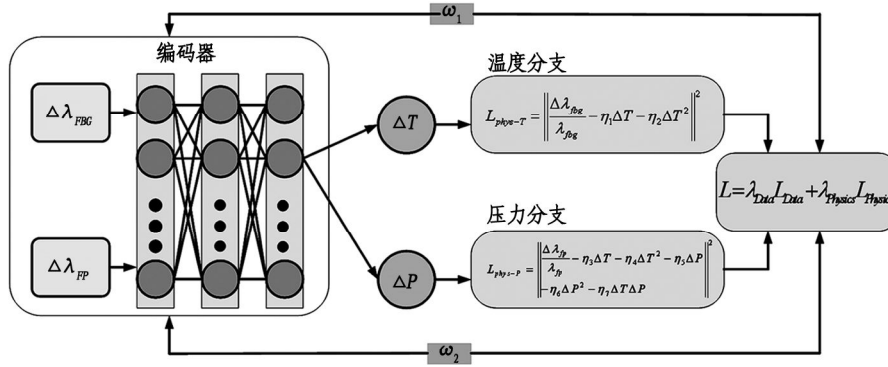


图 7 PINN 处理模型

该研究在恒温水槽与深度可调平台采集了不同温度与压力组合下的光谱数据,并结合仿真数据构建训练集,避免了传统方法中对全工况组合进行逐一标定的繁琐过程。与灵敏度矩阵方法相比,DPWA-DualPINN 使温度测量的最大残差降低了约 49%,预测精度提升了 3.59 倍;压力测量的最大残差降低了约 74.4%,预测精度提升了 5.23 倍。然而,深度学习方法的性能仍受训练数据分布范围的影响,当测试工况超出训练覆盖区间时模型的外推能力尚需进一步验证。与传统方法相比, PINN 在处理非线性耦合与交叉串扰方面表现出显著优势,且在数据量有限的条件下仍能保持较好的解耦精度。该方法在海洋温度-深度原位监测等复杂环境中展现出良好的应用潜力。

5.6 基于深度学习解调的共性优势与待解决问题

综合上述研究,深度学习解调相较于传统方法主要体现出以下几方面特点。在特征提取方面,传统方法的性能往往受限于人工特征设计的完备程度,而深度学习能够从原始数据中学习任务相关

的特征表示,捕捉人工方式难以显式构造的复杂光谱特征。在建模流程方面,传统解调通常将复杂任务拆分为若干独立步骤,步骤间误差可能逐级传播并累积;深度学习通常可将多个处理环节纳入统一的优化框架,有助于减少中间环节的误差传递。在非线性映射方面,传统方法在建立可计算模型时,通常需要对物理过程做线性化处理或小信号近似,深度学习则可直接拟合复杂的非线性关系,在一定程度上减少对线性化或小信号假设的依赖。在多参数协同解调方面,传统方法处理多参数耦合时往往需要设计专门的解耦算法,深度学习则可通过多输出结构同时预测多个物理量,并在训练过程中自动捕捉参数间的耦合关系。

尽管深度学习在上述方面表现出一定潜力,但在其实际应用上仍面临若干挑战。数据依赖是其中较为突出的问题。网络性能在很大程度上取决于训练数据集的代表性与完备性,而在 F-P 传感场景下,获取覆盖全量程、多参数耦合及复杂工况条件的高质量数据集,工程成本较高,实现难度也较大。其

次,模型在测试条件超出训练数据分布范围时,解调结果可能出现明显偏差甚至失真,外推能力存在一定局限。可解释性不足是另一项制约因素。传统算法的各处理步骤通常对应明确的物理含义,而神经网络内部表征机制较为隐含,这一特征使其在高可靠性要求场景下的应用受到限制。此外,物理一致性问题同样值得关注。纯数据驱动模型若缺少必要的物理约束,可能输出与基本干涉机理不一致的结果,影响模型的稳定性与可用性。计算资源需求与实时性之间的平衡也是需要考虑的因素。深度学习解调通常依赖较高的算力支持,在边缘计算或嵌入式部署条件下实现仍存在一定挑战。

深度学习与传统方法之间更适理解解为两种具有互补性的解调范式,而非简单的替代关系。传统方法以明确的物理模型为基础,具有较好的可解释性和相对清晰的性能边界;深度学习的优势则主要体现在复杂映射拟合和自动特征提取方面。未来较为值得关注的方向,是推动二者的深度融合——以物理模型提供基础框架,以深度学习提升模型表达能力。

6 展望

研究围绕强度解调、频域解调、谱峰追踪解调和深度学习解调等方法,梳理了 F-P 传感器解调算法的发展脉络。从这一发展过程来看,F-P 传感器解调技术是在有限光谱观测信息条件下,对测量精度、动态范围、抗干扰能力、实时性与实现复杂度进行权衡。传统解调方法建立在明确的物理模型之上,具有较好的可解释性,但在低信噪比、谱形畸变、宽动态范围以及长期漂移等条件下,其性能容易受到模型假设和工作边界的制约。深度学习等智能解调方法在复杂非线性映射建模方面表现出优势,但仍需面对样本依赖、泛化能力、可解释性和工程部署成本等问题。在这一背景下,F-P 传感器解调技术的后续发展不宜仅着眼于单一算法精度或运算速度的提升,还应面向复杂工况和工程应用需求,围绕鲁棒解调、低成本智能解调以及采集与算法协同优化等方向进一步展开。

面向复杂扰动的鲁棒解调仍是 F-P 传感器高可靠应用需要解决的基础问题。强度解调、频域解调和谱峰追踪等方法在稳定工作点、高信噪比、清晰谱峰结构或充分频谱分辨率条件下可以获得较好性能,但当传感器处于高动态、强噪声或多腔干涉环境时,容易出现谱峰误判、相位非线性、频谱泄漏和解调不连续等问题。这表明,传统解调方法的性能在

很大程度上仍依赖理想化的信号条件。围绕这一问题,近期研究开始在保留 F-P 干涉物理模型主体框架的基础上,引入特征增强、参数优化和误差校正等机制,以提高复杂环境下的解调可靠性。在低信噪比光谱解调方面,Yang 等^[53]针对光谱信号质量退化导致的腔长解调误差问题,采用多级离散小波变换提取原始光谱的近似系数,并将其作为降维特征输入由支持向量机和 K 近邻回归器构成的集成模型。该方法的核心在于利用小波变换的噪声抑制能力和集成学习的非线性映射能力,对低信噪比条件下的光谱退化进行补偿。在相位解调方面,Liu 等^[54]提出五步相移与多波长平均相结合的方法,通过在多个连续工作波长点上分别获取相位估计值并进行平均处理,引入冗余波长信息以抑制随机噪声,从而提升解调的稳定性和动态范围。针对高灵敏度、大动态范围测量场景,Yang 等^[55]通过优化相位生成载波(Phase Generated Carrier,PGC)解调中的关键参数,提高了相位提取精度。Li 等^[56]则针对 PGC 解调中载波相位延迟引起的正交分量衰减和信噪比退化问题,提出移位采样预标定方法,在测量前获取载波相位延迟标定值,并将其用于后续相位补偿,从而减小正交分量幅值不均衡带来的解调误差。上述研究分别从光谱域特征增强、多波长冗余平均、PGC 参数优化和相位延迟预标定等角度展开,处理对象和技术路径并不相同,但其共同特征在于:并未简单以数据模型替代干涉物理模型,而是在保留物理模型主体结构的基础上,对噪声、参数失配、相位延迟和信号退化等非理想因素进行针对性补偿。这一思路表明,提升复杂环境下 F-P 传感器解调鲁棒性的关键,可能并不在于传统物理方法与数据驱动方法之间的取舍,而在于二者如何结合。后续研究可进一步将腔长、相位、光程差和谱峰位置等物理中间量嵌入学习模型的训练或推理过程,同时引入信号质量评估、异常状态识别和误差补偿,构建一种混合解调框架,在保持物理可解释性的同时提升复杂环境下的适应能力。

面向小样本、稀疏采样和低成本硬件的智能解调,为缓解数据依赖与系统成本约束提供了一条可行思路。现有深度学习解调方法对标注样本的规模、质量和分布覆盖程度通常有较高要求,而 F-P 传感实验数据的获取受限于标定设备、加载平台、极端环境复现条件和传感器批次差异等因素。与此同时,高分辨率光谱仪或密集波长扫描系统也会增加系统体积、硬件成本以及实时处理压力。因此,如何在有限采样点和有限标注样本条件下实现可靠解

调,是智能解调方法从高分辨率实验条件走向低成本应用时需要面对的问题。针对光谱重构层面的尝试, Ren 等^[57]提出基于无监督表征学习的光谱重构解调框架,利用无监督学习从光谱数据中提取表征,并在较少采样点条件下恢复干涉光谱,为降低对高分辨率光谱仪和密集采样过程的依赖提供了一种可能方案。Xu 等^[58]提出资源受限环境下的高分辨率光谱重构框架,通过深度学习融合超低分辨率光谱样本与阵列波导光栅的投影强度信息,实现有限采样条件下的高分辨率光谱重建。两项工作均未单纯追求网络规模扩大,而是在采样条件或硬件资源受限的前提下,提高有限观测条件下的光谱表达能力。这类研究表明,当前相关工作已开始将数据端和硬件端的约束纳入方法设计,关注样本效率、采样效率与硬件成本之间的协同。然而,目前相关研究更多集中于光谱恢复或重构环节,与下游解调任务之间的端到端联合优化仍有进一步拓展空间。弥合这一差距的关键,在于使模型在有限标注条件下学习兼具通用性与任务相关性的光谱表征。后续研究可进一步结合物理仿真数据增强、迁移学习和自监督预训练等方法,使模型从无标签或少标签光谱数据中提取干涉条纹周期性、谱峰迁移和光谱包络变化等特征,再通过少量标定数据完成特定器件和工况下的快速适配。

采集硬件与解调算法的协同优化,则从系统层面为提升 F-P 传感系统实时性与工程可部署性提供了新的思路。传统解调过程多采用固定采集与后端处理分离的顺序式框架,即传感系统按预设扫描范围、采样点数和采集频率获取光谱,再由后端算法完成参数反演。在这一模式下,光源调谐方式、采样策略、光谱探测结构和解调模型之间相对独立,难以根据待测信号频率、实时信噪比、动态范围以及系统资源约束进行联合调整。近年来,部分研究开始从系统层面探索采集与解调的一体化设计。Xu 等^[59]提出基于扫描激光的高速光谱解调方法,利用分布式布拉格反射(Distributed Bragg Reflector, DBR)激光器进行快速波长扫描,在不依赖光谱仪的条件下实现了宽范围 F-P 腔长解调,其腔长解调范围可达 0.3 ~ 13 mm。Li 等^[60]构建了基于光纤阵列并行光谱探测的超高速相位解调系统,通过阵列波导光栅分光与 48 通道光电探测并行采集,将 F-P 传感器相位解调速率提升至 200 kHz,并实现 60 ~ 700 μm 范围内的腔长解调。Bai 等^[61]进一步利用稀疏光谱的时空特征,将密集波分复用器、多通道光电探测器与神经网络模型相结合,实现最高 500 kHz 的解调速

率,并支持 15 ~ 950 μm 范围内的绝对腔长解调。Chen 等^[62]将阵列波导光栅与神经网络结合,通过光谱编码硬件将波长变化映射为多通道强度信息,再由学习模型完成非线性反演,展示了光谱编码硬件与数据驱动模型联合设计的应用潜力。从上述工作可以看出, F-P 解调技术正由单一后端算法优化,逐步转向光源扫描、光谱采集、特征编码与参数反演等环节的系统级协同设计。面向长期连续监测和高频动态测量应用,后续仍需进一步引入在线校准、漂移监测与模型更新机制,以提升系统在复杂工况下的长期稳定性和工程可部署性。

综上所述, F-P 传感器解调技术的未来发展需要进一步聚焦复杂工况和工程应用中的关键问题,并围绕三类具有现实基础的路径展开:一是面向低信噪比、宽动态范围和复杂谱形扰动的物理模型——数据驱动融合鲁棒解调;二是面向有限标注样本、稀疏光谱采样和低成本硬件条件的智能解调;三是面向高速、实时和长期部署需求的采集硬件——解调算法协同优化。随着光谱采集器件、传感器封装工艺、智能学习模型和在线校准技术的协同发展, F-P 传感器解调研究将不再局限于单一参数反演,而是进一步拓展为融合物理建模、数据学习、系统优化和可靠性评估的综合测量问题。

7 结论

围绕光纤 F-P 传感器解调算法,对强度解调、频域解调、谱峰追踪解调和深度学习解调方法进行了系统梳理。强度解调方法结构简单、响应速度快,但受线性工作范围、光强波动和环境稳定性限制;频域解调方法能够利用更完整的光谱信息,提高信噪比和抗干扰能力,但其性能受光谱扫描范围、频率分辨率、频谱泄漏和多腔串扰等因素制约;谱峰追踪类方法具有较高测量精度,并在绝对腔长解调方面形成了多种改进策略,但对特征峰的稳定识别具有较强依赖,在低信噪比、谱形畸变和宽动态范围条件下存在峰值误判与级次模糊风险。随着应用需求向高精度、多参数和复杂环境测量拓展,传统解调方法中有限特征提取和理想物理模型假设的局限逐渐显现。深度学习方法为 F-P 传感器解调提供了新的建模思路。前馈神经网络、深度信念网络、CNN-LSTM 和物理信息神经网络等模型能够从光谱数据中自动提取特征,并建立光谱与待测物理量之间的非线性映射关系,在串扰抑制、多参数解耦、低分辨率光谱解调和复杂工况适应等方面展现出应用潜力。然而,深度学习解调仍面临训练样本覆盖不足、模型泛化能

力有限、物理一致性欠缺、可解释性不足和工程部署成本较高等问题。

综合来看,F-P 传感器解调算法的发展体现出两条演进主线:一是信息利用方式由单点或少数特征点提取逐渐转向全谱信息融合;二是建模方式由显式物理模型驱动逐步转向物理约束与数据驱动协同。传统方法具有明确的物理基础和可解释边界,智能方法增强了复杂非线性关系建模和自适应特征提取能力,二者在复杂应用场景中具有互补关系。未来研究应在明确各类方法适用条件和失效机制的基础上,推动物理模型、数据学习、光谱采集硬件和在线校准机制的协同优化,以提升 F-P 传感系统在复杂环境下的精度、鲁棒性、实时性和长期稳定性。

参考文献:

- [1] 李文豪,贾平岗,王军,等. 基于硅 MEMS 技术的高灵敏度微型光纤法布里-珀罗压力传感器[J]. 光子学报, 2024,53(5):0553110.
- [2] 孙萌萌,葛益娴,沈令闻,等. 基于电弧放电优化的光纤法珀湿度传感器[J]. 光学学报, 2022, 42(10): 1006002.
- [3] 盛天宇,李健,李鸿昌,等. 光纤法珀式 SiC 耐高温压力传感器的制造与测试[J]. 中国机械工程, 2022, 33(15):1803-1809.
- [4] Chen Ling, Tian Jiajun, Wu Qiang, et al. Hybrid fabry-pérot interferometer: polymer-induced microcavity for temperature and gas pressure measurement[J]. Optics Communications, 2024,563:130578.
- [5] Perez-Herrera R A, Bravo M, Leandro D, et al. Multiparameter sensor based on a multi-interferometric serial configuration for temperature and strain measurements[J]. IEEE Journal of Selected Topics in Quantum Electronics, 2021, 27(6):5600704.
- [6] Liu Haojie, Peng Nian, Meng Wei, et al. Optical fiber temperature sensing probe based on F-P cavity for human body temperature monitoring[J]. IEEE Sensors Journal, 2024, 24(10):16188-16195.
- [7] Hou Liangtao, Li Yan, Liu Yi, et al. High-integration optical fiber sensor with Vernier effect based on spatial beam splitting[J]. Optics & Laser Technology, 2024, 171: 110415.
- [8] Zhao Xinyu, Ren Xiwang, Bai Jiandong, et al. Enhanced high-temperature gas pressure sensor based on a fiber optic fabry-pérot interferometer probe[J]. Journal of Lightwave Technology, 2025, 43(2):977-982.
- [9] Zhu Chen, Zheng Hongkun, Ma Lingmei, et al. Advances in fiber-optic extrinsic fabry-perot interferometric physical and mechanical sensors: a review[J]. IEEE Sensors Journal, 2023, 23(7):6406-6426.
- [10] Kilic O, Digonnet M J F, Kino G S, et al. Miniature photonic-crystal hydrophone optimized for ocean acoustics[J]. The Journal of the Acoustical Society of America, 2011, 129(4):1837-1850.
- [11] Yu Bing, Kim Dae Woong, Deng Jiangdong, et al. Fiber fabry-perot sensors for detection of partial discharges in power transformers[J]. Applied Optics, 2003, 42(16): 3241-3250.
- [12] Dong Bo, Han Ming, Sun Liquan, et al. Sulfur hexafluoride-filled extrinsic Fabry-Pérot interferometric fiber-optic sensors for partial discharge detection in transformers[J]. IEEE Photonics Technology Letters, 2008, 20(18): 1566-1568.
- [13] 谭心,李自鹏,杨桥,等. 一种双环形凹面镜结构的光纤 F-P 腔位移传感器设计[J]. 激光杂志, 2025, 46(2):29-33.
- [14] Santos J L, Leite A P, Jackson D A. Optical fiber sensing with a low-finesse Fabry-Perot cavity[J]. Applied Optics, 1992, 31(34):7361-7366.
- [15] 王东平,王伟,张军英,等. 光纤法布里-珀罗传感器双峰-干涉级次定位联合解调算法[J]. 光学学报, 2022, 42(16):213-219.
- [16] Feng Fei, Chen Baojie, Alshehri H M, et al. The demodulation algorithm of optical fiber Fabry-Perot sensor using the nonlinear dynamic characteristics of the optical fiber broadband light source[J]. Fractals, 2022, 30(2):2240072.
- [17] 李春成,王鸣,夏巍,等. 基于 F-P 腔强度解调的微小位移传感器[J]. 光学学报, 2014, 34(6):0628001.
- [18] 江覃晴,吕桂贤,赵文彬,等. 基于 DFB 激光器的振动传感系统改进 PGC 解调算法[J]. 半导体光电, 2024, 45(2):319-326.
- [19] 胡雨润,王目光,孙春然,等. 光纤干涉传感器相位生成载波解调算法研究[J]. 激光技术, 2022, 46(2):213-219.
- [20] 朱森荣,刘杰徽. 基于最小二乘法椭圆拟合的改进型快速算法[J]. 舰船电子工程, 2022, 42(1):33-35.
- [21] Ren Qianyu, Jia Pinggang, An Guowen, et al. Dual-wave-length demodulation technique for interrogating a shortest cavity in multi-cavity fiber-optic Fabry-Pérot sensors[J]. Optics Express, 2021, 29(20):32658-32669.
- [22] 任乾钰,贾平岗,钱江,等. 多腔型光纤法布里-珀罗传感器三波长动态解调技术[J]. 光子学报, 2022, 51(9):0906005.
- [23] Islam M R, Ali M M, Lai M H, et al. Chronology of Fabry-Perot interferometer fiber-optic sensors and their applications: a review[J]. Sensors, 2014, 14(4):7451-7488.
- [24] 黄宇浩,任乾钰,王军,等. 基于 Lissajous 曲线拟合的 Fabry-Perot 传感器三波长小信号标定方法[J]. 光子学

- 报,2025,54(6):0606001.
- [25] 付亚平, 闻旭, 王巧云. 基于 FFT-COSS 的光纤 F-P 传感器解调算法[J]. 东北大学学报(自然科学版), 2022, 43(11): 1529-1535.
- [26] 雷小华, 陈伟民, 章鹏, 等. 基于三次样条插值的光纤 F-P 传感器傅里叶变换解调研究[J]. 光子学报, 2008, 37(4): 705-708.
- [27] Chinaev A, Enzner G, Schmalenstroer J. Fast and accurate audio resampling for acoustic sensor networks by poly-phase-Farrow filters with FFT realization[C]//2018 13th ITG-Symposium on Speech Communication, Oldenburg, Germany. Oldenburg: VDE, 2018: 1-5.
- [28] 王晶, 张子豪, 余鑫宇, 等. 法布里-珀罗干涉信号的小波变换相位补偿算法[J]. 光子学报, 2021, 50(7): 39-45.
- [29] Liu Zheng, Lou Ruizhe, Zhang Lingling, et al. High precision and low noise demodulation scheme using adaptive fast Fourier transform for fiber-optic interferometric sensors[J]. Optics Express, 2025, 33(1): 817-828.
- [30] 姚琼, 刘政, 刘刚, 等. 光纤法布里-珀罗传感器阵列串扰抑制方法[J]. 光学学报, 2023, 43(23): 2306004.
- [31] 江俊峰, 邹盛亮, 王双, 等. 空间扫描型光纤法布里-珀罗传感解调中信噪比的影响研究[J]. 光学学报, 2015, 35(11): 1106003.
- [32] Chen Shengchao, Ren Sufen, Yang Jianli, et al. Reconstruction of Fabry-Perot interferometric sensor spectrum from extremely sparse sampling points using dense neural network[J]. IEEE Photonics Technology Letters, 2022, 34(24): 1337-1340.
- [33] Guo Zilong, Chen Cang, Zhang Junying, et al. Demodulation of fiber-optic Fabry-Perot sensors through a double-frequency self-correlation method[J]. Optical Fiber Technology, 2021, 64: 102563.
- [34] 陶珺, 陈杨黎, 卢景琦. 基于稀疏快速傅里叶变换的光纤 F-P 传感器腔长解调方法[J]. 中国激光, 2018, 45(5): 0510001.
- [35] 那美丽, 周志刚, 李霏霏. 基于稀疏傅里叶变换的低采样率宽带频谱感知[J]. 电子技术应用, 2015, 41(11): 85-88.
- [36] Hassanieh H, Indyk P, Katabi D, et al. Nearly optimal sparse Fourier transform[C]//Proceedings of the 44th Annual ACM Symposium on Theory of Computing, New York, USA. New York: ACM, 2012: 563-578.
- [37] 席婷苇, 温嘉琳, 董玉明, 等. 光纤法布里珀罗超声传感器解调方法进展[J]. 激光与光电子学进展, 2025, 62(5): 0500001.
- [38] Li Chenxi, Zhao Xinyu, Qi Hongchao, et al. Integrated fiber-optic Fabry-Perot vibration/acoustic sensing system based on high-speed phase demodulation[J]. Optics & Laser Technology, 2024, 169: 110131.
- [39] 葛吉瑶, 赵庆超, 孙晓, 等. 基于 F-P 腔的高灵敏光纤振动传感器[J]. 光电子·激光, 2025, 36(3): 231-238.
- [40] Zhang Weiguang, Yu Jia, Zhang Xiongxing, et al. High-precision cavity length demodulation method for Fiber-Optic Fabry-Perot sensors based on dual superluminescent diodes[J]. Sensors, 2022, 22(15): 5898.
- [41] Li Zhen, Wu Guoxiu, Meng Lingjie, et al. Multimodal demodulation algorithm for fiber optic MEMS Fabry-Perot sensors[J]. Journal of Optics, 2024, 54: 1625-1633.
- [42] 陈斯彤, 王伟, 周立凡, 等. FP 传感器双峰级次腔长解调算法[J]. 光电子·激光, 2024, 35(2): 128-134.
- [43] Qi Wang, Li Zhiqi, Fan Yijie, et al. High precision and speed demodulation algorithm for optical fiber Fabry-Perot temperature sensor[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 9516709.
- [44] Liu Yang, Li Cheng, Lu Shanshan, et al. Temperature-compensated high-sensitive graphene Fabry-Perot acoustic sensor[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 9522607.
- [45] Li Shiyu, Jie Ruimin, Alsalman Osamah, et al. Large-range and high-sensitivity displacement sensing based on extrinsic Fabry-Perot interferometer assisted microwave photonic filter[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 7008511.
- [46] Ma Chaoqun, Chen Yanxia, Qu Wenwen, et al. High-sensitivity fiber temperature and pressure sensor based on Fabry-Perot interferometry and vernier effect[J]. Optics & Laser Technology, 2025, 181: 111999.
- [47] Daponte P, De Vito L, Picariello F, et al. Assessment of Sparse Fourier Transform for Spectral Measurements[C]//2023 IEEE International Instrumentation and Measurement Technology Conference (I2MTC). IEEE, 2023: 1-6.
- [48] 王志峰. 基于 FBG 集成 F-P 干涉的温度-应变传感器研究[J]. 光学仪器, 2024, 46(6): 49-54.
- [49] Wu Ying, Xia Li, Wu Nishan, et al. Optimized feedforward neural network for multiplexed extrinsic Fabry-Perot sensors demodulation[J]. Journal of Lightwave Technology, 2021, 39(13): 4564-4569.
- [50] Wang Lixiong, Liu Hanjie, Pan Zhen, et al. Temperature demodulation for optical fiber F-P sensor based on DBNs with ensemble learning[J]. Optics & Laser Technology, 2023, 162: 109275.
- [51] Wei Chuanhao, Liu Qiang, Wang Yiping, et al. Accurate strain sensing with fiber-optic Fabry-Pérot sensors based on CNN-LSTM model[J]. IEEE Sensors Journal, 2024, 24(11): 17725-17732.

基于 MacTempas 软件模拟 DO_{24} 结构相的选区电子衍射

程宏伟*, 孟维利, 周家峰, 屈苏平, 董 超

(蚌埠学院 数理学院, 安徽 蚌埠 233030)

摘 要:以 DO_{24} 结构相为对象, 基于 Wyckoff 位点构筑不同原子占位模型, 利用 MacTempas 软件模拟不同晶带轴下的选区电子衍射花样。将模拟获得的清晰衍射图谱与实际实验衍射图片进行比对分析, 从而确定 DO_{24} 结构相的原子组态。该方法实现了选区电子衍射实验过程的可视化与结果推演的无限试错, 构建了一个高度拟真且完全开放的“虚拟电镜实验室”。研究表明, 模拟与实验相结合的电子晶体学方法可有效解析复杂合金相的原子尺度结构, 为材料科学中的相分析提供新思路。

关键词: Wyckoff 位点; MacTempas 模拟; 选区电子衍射; 结构组态

中图分类号: TG115.23

文献标识码: A

文章编号: 2095-297X(2026)05-0096-04

Selected Area Electron Diffraction Simulation of DO_{24} Phases via MacTempas

CHENG Hongwei*, MENG Weili, ZHOU Jiafeng, QU Suping, DONG Chao

(School of Physics and Mathematics, Bengbu University, Bengbu, 233030, Anhui)

Abstract: The DO_{24} structured phase was investigated by constructing atomic occupancy models based on Wyckoff positions. Selected area electron diffraction (SAED) patterns along different zone axes were simulated using MacTempas software and were compared with experimental results, by which the atomic configuration of the DO_{24} phase was determined. Visualization of the SAED experimental process and infinite trial-and-error in result interpretation were enabled through this approach, establishing a highly realistic “virtual electron microscopy laboratory”. It has been demonstrated that complex alloy phases can be effectively resolved at the atomic scale by combining simulation with experimental electron crystallography, and new insights are provided for phase analysis in materials science.

Key words: Wyckoff site; MacTempas simulation; selected area electron diffraction; configuration

长期以来,材料科学的研究一直围绕着材料的结构—性能的关系展开。对于绝大多数材料,晶体结构及各类缺陷决定了其性能和使役行为^[1]。因此,分析表征材料的晶体结构是材料研究的核心内容。自从德国电气工程师 Emst Ruska 与 Max Knoll 发明了电子显微镜后,经过百年的不断发展,电子显微镜已成为研究材料晶体结构及缺陷表征最常用、最有力的工具之一,是材料研究不可或缺的重要手段^[2]。电子显微术的发展和应用极大地拓展了人们对材料结构的认知,推动了材料科学的迅猛发展。

然而,在利用透射电子显微学分析材料的结构时,研究者往往被阻隔在昂贵精密的仪器与深奥难懂的衍射图谱之间。如何将电子束与晶体原子相互作用产生的抽象信息,转化为可感知、可探索的直观经验,成为深化材料科学必须破解的关键命题。

透射电子显微分析长期徘徊于两种模式之间:一是依赖二维图片,难以构建原子排布与衍射斑点之间动态、立体的联系;二是受制于设备稀缺、成本高昂、操作复杂,真实实验往往沦为短暂观摩。对 Wyckoff 位点、晶带轴、倒易空间等核心概念的理解

收稿日期: 2026-01-21 * 通讯联系人

基金项目: 安徽省自然科学基金项目(2025AHGXZK31133);安徽省质量工程项目(2024xjzlt061)。

作者简介: 程宏伟(1985—),男,安徽怀远人,讲师,博士。E-mail: hwcheng@mail.ustc.edu.cn

容易流于表面,更难以自主地将衍射图谱的几何特征与样品台下真实的原子组态一一对应。这种认知隔阂,不仅削弱了人们探究微观世界的兴趣,也制约了利用电镜技术解决实际材料科学问题的创新能力。

正是在这一背景下,探讨了 718Plus 合金中析出相的晶体学结构信息,结合 Wyckoff 位点构筑不同原子占位的 DO₂₄ 结构相。利用 MacTempas 软件模拟其“选区电子衍射”,与实验中得到的选区电子衍射结果对比。其核心思路:将抽象的晶体学原理(Wyckoff 位点)与具体的材料结构模型(DO₂₄ 结构相)相结合。通过专业的模拟软件(MacTempas)将“晶体结构→衍射条件→衍射图谱”这一核心认知链条进行了拆解、重组,实现过程的可视化与结果的无限试错,从而获得对 DO₂₄ 结构相原子组态的深刻理解。

1 选区电子衍射

晶体具有周期性排列的原子结构,可视为由一系列晶面间距相等、相互平行的原子面堆垛而成。图 1(a)展示了空间群为 Fm-3m 的晶体结构,该图呈现了其 2×2×2 晶胞范围内的原子排布示意图。透射电子显微学中,电子衍射的本质:电子束穿过晶体时,会于晶体的周期性晶格发生散射,其中满足 Bragg 条件的晶面,才能产生衍射电子,衍射电子被探测器捕捉,就形成了衍射图(见图 1(b))。选区电子衍射上的单晶斑点,可看作厄瓦尔德球和倒易点阵相截时在二维底片平面上的投影^[3]。图 1(c)中的网格点代表倒易点阵的阵点,其中每个点(如 G)对应真实晶体中的一族晶面(hkl)。倒易矢量 $\mathbf{g}$ 连接倒易原点 O^* 和倒易点 G。其长度 $|\mathbf{g}| = 1/d$,

其方向垂直于真实空间中的(hkl)晶面。以入射波矢起 C 为球心,以 $1/\lambda$ 为半径作球面。入射波 $\mathbf{k}_0$: 从球心 C 指向倒易原点 O^* ,其长度 $|\mathbf{k}_0| = 1/\lambda$ 。球面意义:在几何上,从球心 C 到球面上任意一点的连线构成的矢量 $\mathbf{k}$,其长度都等于 $1/\lambda$,代表一个可能存在的与入射波等波长的散射波方向。衍射的几何条件:当某个倒易点 G 恰好落在厄瓦尔德球的球面上时,连接球心 C 到 G 的矢量 $\mathbf{k}$,就代表了一束衍射束的方向。此时,矢量关系 $\mathbf{k} - \mathbf{k}_0 = \mathbf{g}$ 成立,这完全等价于布拉格定律 $2d\sin\theta = \lambda$ (当 $n = 1$ 时)。因此,“倒易点落在球面上”是布拉格定律在倒易空间中的几何表达^[4]。单晶电子衍射的实际情形,在 TEM 实验中,电子波长极短(如:加速电压为 200 kV 时, $\lambda \approx 0.00251$ nm),因此球半径 $1/\lambda$ 非常大(约 400 nm^{-1}),而低指数晶面的 $|\mathbf{g}|$ 很小(如 $d = 0.2$ nm 时, $|\mathbf{g}| = 5 \text{ nm}^{-1}$)。这使得在倒易原点 O^* 附近,巨大的球面可以近似视为一个平面。同时,由于样品非常薄(通常 < 100 nm),其倒易点在垂直于样品薄片的方向会被拉长成倒易杆。这两个效应(大球面和倒易杆)使得球面非常容易与一系列倒易杆相交。这些相交点在垂直于入射束方向(即 O^* 点法向平面)上的投影,就形成了在屏幕上看到的规则排列的二维衍射斑点阵列。不同组晶面的衍射电子,会在电子探测器上形成分散的衍射斑点。每个斑点对应一组满足衍射条件的晶面的衍射束方向,且这些斑点空间排列规律与晶体的晶格对称性严格一致。通过单晶衍射斑点的分布和间距,计算晶面间距 d 值,结合布拉格方程($2d\sin\theta = n\lambda$)标定晶格常数,确定材料的晶体结构。

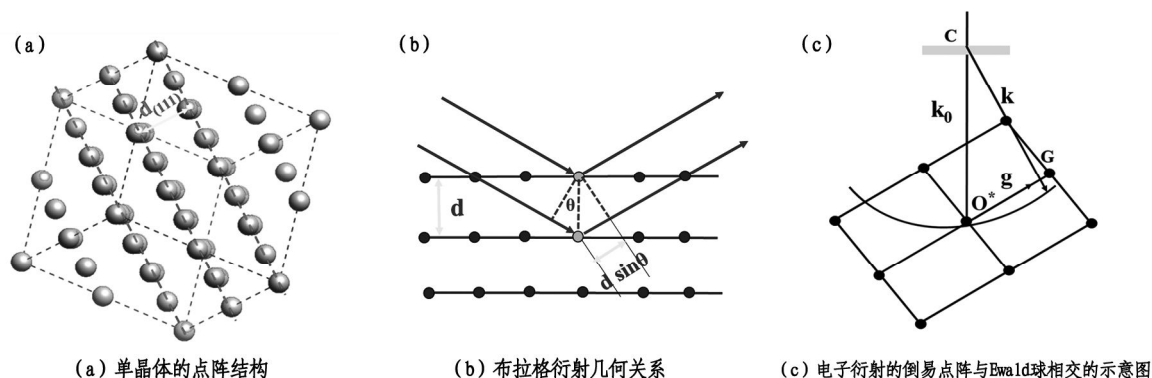


图 1 电子衍射的本质示意图

2 实验方法

以相对较难理解和掌握的六方晶系相 η -Ni₃Ti 相作为研究对象。研究表明: Ni₃Ti 相的空间群为 194, 对称性为 P6₃/mmc, 晶格常数为 $a = 5.096 \text{ \AA}$, $c = 8.304 \text{ \AA}$; $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$ ^[5]。根据 Wyckoff No. 得到晶体中原子占位情况如表 1 所示。

表 1 η -Ni₃Ti 结构数据

原子	Wyckoff No.	x	y	z
Ti1	2a	0	0	0
Ti2	2c	1/3	2/3	0
Ni1	6g	1/2	0	0
Ni2	6h	-1/6	1/6	1/4

注:表中数据由 Laves 等^[6]结果得出。

利用 VESTE 软件构建不同原子占位 η -Ni₃Ti 相的晶体结构,如图 2 所示。首先设定 Nb, Ti 和 Al 原子随机占据 2a/2c 原子位置(图 2(a)中 Ti 原子占据的点阵位置), Ni 原子占据 6g 和 6h 位,构建 η -Ni₃Ti 相无序晶体结构如图 2(a)所示。设定 Nb 占据 2a 原子位, Ti 和 Al 原子随机占据 2c 原子位, Ni 原子依然占据 6g 和 6h 位。构建出有序的 η -Ni₃Ti 晶体结构如图 2(b)所示,分别导出保存为 CIF 格式的文件。

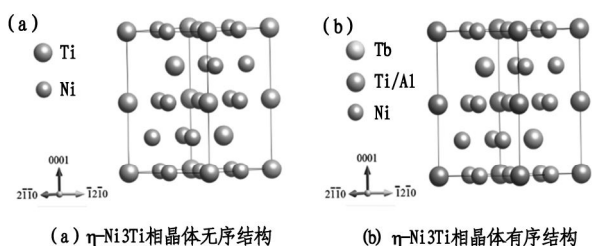


图 2 DO₂₄ 结构的原子组态

3 模拟结果

打开 TEMPER, 通过 File→Import 导入 η -Ni₃Ti 相晶体无序结构与有序结构的 CIF 文件, 选择 SAED 模式, 分别输入 $[10\bar{1}0]$ 、 $[2\bar{1}\bar{1}0]$ 和 $[0001]$ 晶带轴, 设置加速电压 200 kV ($\lambda \approx 0.00251 \text{ nm}$), 相机常数 $15 \text{ mm} \cdot \text{\AA}$, 选区大小 $3 \mu\text{m}$, 点击 Simulate→SAED 模式, 自动计算生成衍射花样。用 Indexing 工具点击自动匹配晶面指数, 晶面间距, 验证晶带轴, 勾选标注信息, File→Save 保存为 TIFF/PNG。图 3 为利用 MacTempas 软件模拟 DO₂₄ 结构无序相

与有序相结构在 $[10\bar{1}0]$ 、 $[2\bar{1}\bar{1}0]$ 和 $[0001]$ 晶带轴下的电子衍射图。图 3(a)至图 3(c)为 MacTempas 软件模拟无序结构的选区电子衍射图谱。图 3(d)至图 3(f)为 MacTempas 软件模拟有序结构的选区电子衍射图谱。对比相同晶带轴下有序与无序结构的电子衍射图谱, 证实晶体的结构组态对电子衍射花样有很大的影响。电子衍射图谱可以直观地反映晶体的结构。

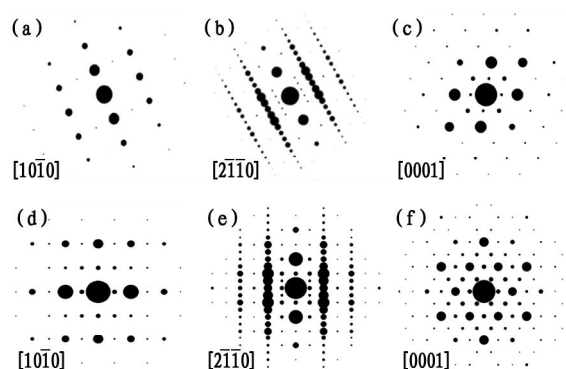


图 3 DO₂₄ 结构相在 $[10\bar{1}0]$ 、 $[2\bar{1}\bar{1}0]$ 和 $[0001]$ 晶带轴下的模拟电子衍射图

图 4 为 718Plus 合金中 η 结构相分别在 $[10\bar{1}0]$ 、 $[2\bar{1}\bar{1}0]$ 和 $[0001]$ 晶带轴下的电子衍射图。分别对衍射花样进行标定, 具体对应的晶面如图 4(a)至图 4(c)所示。经计算得到合金中的 η 相, 晶格常数为 $a = 0.52 \text{ nm}$, $c = 0.84 \text{ nm}$; $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$, 六方结构。合金中的析出相的电子衍射花样与 MacTempas 软件模拟得到的有序结构相的衍射花样(图 3(d)至图 3(f))一致, 证实合金中 η 相的原子组态为 Nb 占据 2a 原子位, Ti 和 Al 原子随机占据 2c 原子位, Ni 原子占据 6g 和 6h 位的 DO₂₄ 有序相结构^[7-8]。可以看出利用 MacTempas 软件模拟晶体的选区电子衍射实验, 能提供很好的实验佐证。同时有效填补昂贵精密的仪器与深奥难懂的衍射图谱之间的鸿沟。

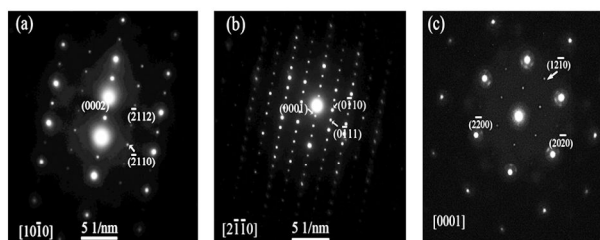


图 4 718Plus 合金中 η 结构相分别在 $[10\bar{1}0]$ 、 $[2\bar{1}\bar{1}0]$ 和 $[0001]$ 晶带轴下的选区电子衍射图

图 5 为 [0001] 晶带轴下有序相电子衍射图的晶面间距的模拟结果。图中箭头所指的衍射斑点, 对应图 4(c) 中 (2020) 晶面。模拟结果显示 d_{20-20} 晶面间距为 2.21 Å。根据六方晶系的晶面间距公式

$$d_{hkl} = \frac{1}{\sqrt{\frac{4}{3} \frac{h^2 + hk + k^2}{a^2} + \frac{l^2}{c^2}}} \quad (a, c \text{ 为晶格常数}, h, k, l$$

为晶面指数), 代入实验测得的数值, 得到 $d_{20-20} = \frac{\sqrt{3}}{4}a = 2.25 \text{ Å}$, 与模拟结果得到的晶面间距一致。证实模拟数据可以很好地还原真实晶体的电子衍射结果。

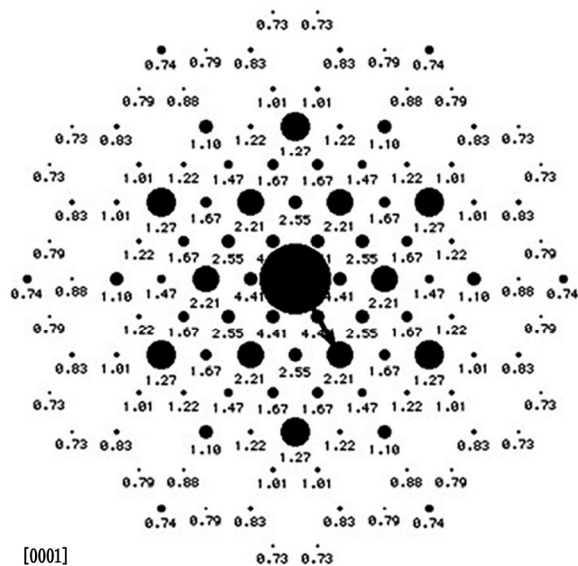


图 5 [0001] 晶带轴下, 有序相结构晶面间距模拟结果

4 结论

根据析出相的实际晶体结构特征, 确定其原子在晶胞中的 Wyckoff 位置, 明确不同元素的原子占位情况。利用专业的模拟软件 (MacTempas) 将“晶

体结构 → 衍射条件 → 衍射图谱”的物理过程进行拆解、重构, 构建出具备过程的可视化与结果的无限试错功能, 且高度拟真、完全开放的“虚拟电镜实验室”。同时模拟结果能够有效指导实验操作与实验结果分析, 深入理解 DO24 结构相的原子组态特征。该模式进一步增强了对电子显微分析原理的掌握和实际应用能力。

参考文献:

- [1] 栗慧, 王晓军, 卢裕臻, 等. 材料科学基础课程的思政探索——以“晶体缺陷”为例[J]. 科学咨询, 2025(5): 116-120.
- [2] 彭海琳. α -Cu₂Se 有序起伏结构的原子级分辨率表征[J]. 物理化学学报, 2019(2): 137-138.
- [3] 黄孝瑛. 材料微观结构的电子显微学分析[M]. 北京: 冶金工业出版社, 2008: 69-70.
- [4] 章晓中. 电子显微分析[M]. 北京: 清华大学出版社, 2006: 74-77.
- [5] Pickering E J, Mathur H, Bhowmik A, et al. Grain-boundary precipitation in Allvac 718Plus[J]. Acta Materialia, 2012, 60(6/7): 2757-2769.
- [6] Laves F, Wallbaum H J. On the crystal structure of Laves phases[J]. Zeitschrift für Kristallographie, 1939, 101: 78-93.
- [7] Ma Shuan, Wang Mingqing, Ju Yiwei, et al. Atomic-scale observation of the interfaces between η and γ matrix in ATI718Plus[J]. Materials Characterization, 2024, 215: 114196.
- [8] Cheng Hongwei, Tang Yuling, Wang Mingqin. Growth mechanisms of eta phase in 718Plus alloy[J]. Materials Communications, 2024, 38: 108452. 责任编辑: 刘娟

租赁设备两阶段维护策略优化

黄显兵*, 姚梦月, 董克, 袁绪潘

(蚌埠学院 经济与管理学院, 安徽 蚌埠 233030)

摘要: 租赁设备在全生命周期内呈现非平稳退化特征, 且失效速率随服役时间延长而加速提升, 传统单一维护策略难以兼顾维护精准性与管理经济性。为此, 提出一种两阶段混合预防性维护模型: 第一阶段采用基于可靠度阈值的非等周期预防维护策略, 第二阶段执行等周期预防维护策略。该模型综合考虑小修成本、预防维护成本以及惩罚成本, 构建以期望总成本最小为目标的两阶段优化模型, 并设计了求解算法。算例结果显示, 该策略的期望总成本优于单一维护策略, 为租赁企业制定维护方案提供了可选方案。

关键词: 租赁设备; 两阶段维护; 可靠度阈值; 等周期维护

中图分类号: F273.4

文献标识码: A

文章编号: 2095-297X(2026)05-0100-06

A Two-Stage Maintenance Strategy Optimization for Rental Equipment

HUANG Xianbing*, YAO Mengyue, DONG Ke, YUAN Xupan

(School of Economics and Management, Bengbu University, Bengbu, 233030, Anhui)

Abstract: Leased equipment exhibits non-stationary degradation characteristics throughout its lifecycle, with its failure rate accelerating as service time increases. Traditional single maintenance strategies struggle to balance maintenance precision with economic efficiency. To address this issue, it proposed a two-stage hybrid preventive maintenance strategy. In the first stage, a non-periodic preventive maintenance strategy based on a reliability threshold was adopted, while a periodic preventive maintenance strategy was implemented in the second stage. This model comprehensively considered factors such as minimal repair costs, preventive maintenance costs, and penalty costs. A two-stage optimization model was developed with the objective of minimizing the expected total cost, and a corresponding solution algorithm was designed. The results of a numerical example showed that the expected total cost of the proposed strategy was superior to that of single strategies, providing an alternative solution for leasing companies in developing maintenance plans.

Key words: leased equipment; two-stage maintenance; reliability threshold; periodic maintenance

近年来,随着设备技术迭代速度不断加快,越来越多的行业更倾向于选择租赁设备,而非购置设备从事生产经营活动^[1]。设备租赁作为一种提供使用权的服务模式,已在建筑设备、工业设备、医疗设备、交通运输等^[2-3]领域得到广泛应用,且成为企业降低成本、提升运营灵活性的重要手段。然而,随着

设备向大型化、智能化、高价值方向发展,设备故障造成的停机损失与维修成本显著增加。同时,设备在全生命周期中呈现初期性能稳定、后期故障概率攀升的非均匀退化特征,进一步加剧了维护管理的复杂性与经济风险。在此背景下,如何从供应链管理视角出发,结合租赁设备运营特点,设计能够协同

收稿日期: 2026-03-31 * 通讯联系人

基金项目: 教育部重点项目(DKA220345); 蚌埠学院科研启动基金项目(2025GQD057); 蚌埠学院校级重点项目(2023SK01zd)。

作者简介: 黄显兵(1972—),男,安徽六安人,讲师,硕士。E-mail: 56692843@qq.com

维护决策与供应链响应的优化策略,以提升设备利用率、降低综合运营成本,具有重要的研究意义。

根据预防维护活动的不同时间间隔,预防维护一般可分为周期性和顺序性两类^[4]。针对周期预防维护,张思源等^[5]以流水线车间为研究对象,在周期预防性维护的约束条件下,以最小化最大完工时间作为优化目标,针对置换车间与非置换车间两种典型情形,分别构建了相应的优化模型。文献[6]基于经典EPQ模型,结合周期性不完全预防维护,以单位产品总成本最小为目标,引入役龄回退因子与等效役龄概念,构建了包含动态维护与次品处理成本的生产批量模型。李雯静等^[7]为了降低露天矿卡车维护成本,基于役龄回退模型与威布尔分布,建立了周期预防性维护与小修相结合的预防维护总成本模型,并设计求解算法以获取最优维护策略。但在实际生产中,设备的失效率会随着使用年限的增长而上升^[8]。

为降低后期高故障风险带来的停机损失,预防维护时间间隔应随着设备使用年限的增加而逐步缩短,应随工龄增加逐步缩短间隔,而非坚持采用等周期维护方式^[9-10]。文献[11-14]均为基于可靠度的顺序维护方法,该方法能够依据设备的实际运行状况调整预防性维护时间,从而适应设备非均匀运行的特性。此外,有学者对多阶段顺序维护进行了研究。曲玉祥等^[15]基于不完全维护理论,提出分阶段顺序维护概念,以设备维护周期内成本率最低为目标,建立一般顺序与分阶段顺序两种预防维护模型。Zhou等^[16]融合周期维护与顺序维护的优势,从出租方角度提出一种多阶段预防性维护策略:将租赁周期划分为多个阶段,各阶段内维护周期相等,阶段间周期逐阶段递减,以应对不完全维护的影响,实现维护总成本最小化。然而,该研究未考虑可靠度对设备维护的影响。邓超等^[17]针对传统顺序维护操作复杂的问题,提出一种分阶段顺序维护模型,将维修过程划分为三个固定间隔阶段,以可靠度为约束构建最小总维护费率优化模型。但基于可靠度的顺序维护在后期设备可靠度较低时,可能导致维护过于频繁,不利于操作与管理。

综上,提出一种顺序预防维护与等周期维护相结合的两阶段租赁设备维护策略模型。该模型将设备的维护分为两个阶段: L_1 和 L_2 阶段。在 L_1 阶段,实行顺序预防维护,即非等周期预防维护,依据给定的可靠度阈值 R_{th} ,对设备进行维护;在 L_2 的初始阶

段,先进行一次预防维护,后续维护将 L_2 分为 m 个等长周期,维护周期时间为 L_2/m ,在周期 L_2 结束时,不进行维护。从出租方的视角出发建立数学模型,以期总成本最小化为优化目标,通过给出的优化算法,最终获得期望总成本最优相对应的维护策略,算例验证了所提模型的有效性,可为出租方实施设备维护提供理论依据。

1 问题描述与基本假设

1.1 问题描述

出租方购置生产设备用于出租,进而获取租金利润。然而,依据租赁合同,出租方需承担设备的日常维护工作。在租赁期间,若设备发生故障,则由出租方对其进行小修,小修将产生小修成本。当设备发生故障的次数超过合同约定次数时,将给出租方带来额外的惩罚成本。为减少设备故障次数,出租方对设备进行预防维护,由此将产生预防维护成本。如何在小修成本、预防维护成本以及惩罚成本之间进行权衡,从而使得设备在租赁期间的维护成本最小化,是出租方亟待解决的问题。

1.2 基本假设

(1) 假设设备的故障率函数随时间单调递增,即形状参数 $\beta > 1$ 。

(2) 在相邻两次预防维护之间,若设备发生故障,则采取小修处理。小修仅修复当前故障,使设备“恢复如旧”,不改变其故障强度。

(3) 假设预防维护为不完美维护,即维护后设备状态介于“全新”与“如旧”之间,无法恢复至初始水平。

(4) 相较于整个租赁周期,小修及预防维护所占用的时间可忽略不计。

2 数学模型

2.1 租赁设备两阶段预防维护策略模型

假定租赁的设备寿命服从二参数的Weibull分布,其尺度参数和形状参数分别记为 α 和 β ,则故障率函数可表示为 $\lambda(t) = (\frac{\beta}{\alpha})(\frac{t}{\alpha})^{\beta-1}$,累积故障率函数为 $\Lambda(t) = (\frac{t}{\alpha})^\beta$ 。设出租方对设备进行 $n+m$ 次预防维护,第 k 次预防维护周期为 T_k ($k=1,2,\dots,n+m$),则系统在第 k 次预防维护后故障率函数 $\lambda_{k+1}(t)$ 为:

$$\lambda_{k+1}(t) = b_k \lambda_k(a_k T_k + t), t \in [0, T_{k+1}], k=1,2,\dots,n+m \quad (1)$$

式中, a_k, b_k 分别为年龄减少因子和故障率增加系数, 且 a_k 和 b_k 满足如下约束: $0 \leq a_1 < \dots < a_n < a_{n+1} < \dots < a_{n+m} \leq 1$ 和 $1 \leq b_1 < \dots < b_n < b_{n+1} < \dots < b_{n+m}$ 。已知 $\lambda(t)$ 为设备在初始状态区间 $[0, T_1]$ 上的故障率函数, 结合公式(1), 可得 $[0, T_k]$ 上的故障率函数 $\lambda_k(t)$ 为:

$$\lambda_k(t) = B_k \lambda(A_k + t) \quad (2)$$

式中, $A_k = \sum_{i=1}^{k-1} a_i T_i, B_k = \prod_{i=1}^{k-1} b_i$, 且 $A_1 = 0, B_1 = 1$ 。

2.2 L_1 阶段维护策略模型

根据设备可靠度与故障率函数之间的关系, 利用递归方程(2), 可以得到第 k 个预防维护周期的设备可靠度 $R_k(t)$ 为:

$$R_k(t) = \exp[-B_k \int_{A_k}^{A_k+t} \lambda(k) dt] \quad (3)$$

若 $T_k (k=1, 2, \dots, n+m)$ 对应于设备的可靠度阈值水平 R_{th} , 则:

$$R_{th} = R_k(T_k) = \exp[-B_k \int_{A_k}^{A_k+T_k} \lambda(k) dt] \quad (4)$$

已知设备的 $\lambda(t)$ 、 $A(t)$, 则设备在第 k 个维护

周期的可靠度函数 $R_k(t)$ 为:

$$R_k(t) = \exp[B_k[(\frac{A_k}{\alpha})^\beta - (\frac{A_k + T_k}{\alpha})^\beta]] \quad (5)$$

由此解得, 第 k 个预防维护周期 T_k 的表达式为:

$$T_k = \alpha[(\frac{A_k}{\alpha})^\beta - (\frac{\ln R_{th}}{B_k})^\frac{1}{\beta}] - A_k \quad (6)$$

2.2.1 L_1 阶段预防维护成本

假设 L_1 内的预防维护成本为 c_p /次, 维护次数为 n 次, 则维护总成本 C_{PM} 为:

$$C_{PM}(L_1) = nc_p \quad (7)$$

2.2.2 L_1 阶段期望小修成本

由文献[11]可知, 第 k 个预防维护区间内小修故障次数的期望值为:

$$E[N_1(T_k)] = B_k \int_{A_k}^{A_k+T_k} \lambda(k) dt \quad (8)$$

从而, 整个 L_1 阶段内的期望小修故障总次数表示为:

$$E[N_1(L_1)] = \sum_{k=1}^n E[N_1(T_k)] + B_n b_n \int_{A_n}^{A_n+\alpha_n T_n+x} \lambda(t) dt \quad (9)$$

式中, $x = L_1 - \sum_{k=1}^n T_k$ 为第 n 次维护与 L_1 阶段结束的剩余时长。

若单次小修成本为 c_r , 则在 L_1 内的期望小修成本为:

$$C_R = c_r E[N(L_1)] \quad (11)$$

2.2.3 L_1 阶段期望惩罚成本

根据约定, 在 L_1 内若设备故障次数超过给定的阈值 N_{max} , 出租方会受到惩罚。引入二元变量 ψ , 将其定义为:

$$\psi = \begin{cases} 1, E[N(L_1)] - N_{max} > 0 \\ 0, \text{其它} \end{cases} \quad (12)$$

假定当故障次数超过给定的阈值 N_{max} 时的惩罚成本为 c_f /次, 则在 L_1 内的期望惩罚成本表达式为:

$$C_R = \psi c_f (E[N(L_1)] - N_{max}) \quad (13)$$

2.2.4 L_1 阶段期望总成本

由式(7)、(11)和(13)可得, 在 L_1 阶段内的期望总成本 $C(L_1, n)$ 为:

$$C(L_1, n) = nc_p + c_r E[N(L_1)] + \psi c_f (E[N(L_1)] - N_{max}) \quad (14)$$

将式(10)代入式(14)并化简, 可得:

$$C(L_1, n) = nc_p + (c_r + \psi c_f) (\sum_{k=1}^n B_k \int_{A_k}^{A_k+T_k} \lambda(t) dt + B_n b_n \int_{A_n}^{A_n+\alpha_n T_n+x} \lambda(t) dt) - \psi c_f N_{max} \quad (15)$$

2.3 L_2 阶段维护策略模型

在租赁期 L_2 的起始阶段, 先进行一次预防维护, 后续的维护是将 L_2 分为 m 个等周期, 维护周期长为 L_2/m , 在 L_2 周期结束时, 不进行维护, 则在 L_2 阶段的维护成本主要由三部分组成: 小修成本、预防维护成本以及惩罚成本。

2.3.1 L_2 阶段预防维护成本

$$E[N_2(T_k)] = B_{n+k+1} \int_{A_{n+k+1}}^{A_{n+k+1}+\alpha_{n+k+1}x+C_k(L_2/m)+L_2/m} \lambda(t) dt \quad (17)$$

整个 L_2 阶段内的期望小修故障总次数表达式为:

$$E[N_2(L_2)] = \sum_{k=1}^m B_{n+k+1} \int_{A_{n+k+1}}^{A_{n+k+1}+\alpha_{n+k+1}x+C_k(L_2/m)+L_2/m} \lambda(t) dt \quad (18)$$

若单次小修成本为 c_r , 则 L_2 内的期望小修成本为:

$$C_R(L_2) = c_r E[N_2(L_2)] \quad (19)$$

假定 L_2 内的预防维护成本为 c_p /次, 维护次数为 m , 则 L_2 内的维护总成本 $C_{PM}(L_2)$ 为:

$$C_{PM}(L_2) = mc_p \quad (16)$$

2.3.2 L_2 阶段期望小修成本

令 $C_m = \frac{L_2}{m} \sum_{i=1}^{m-1} \alpha_{n+1+i}, C_1 = 0$, 则第 k 个预防维护

区间的期望小修故障数为:

2.3.3 L_2 阶段期望惩罚成本

在整个租赁期 L 内,若二手设备的故障次数超过给定的阈值 $N_{\max}$,依据合同约定,出租方则会受到惩罚。引入二元变量 ψ ,定义如下:

$$\psi = \begin{cases} 1, E[N_1(L_1) + N_2(L_2)] - N_{\max} > 0 \\ 0, \text{其它} \end{cases} \quad (20)$$

则在整个租赁期 L 内的期望惩罚成本表达式

$$C(L_2, m) = mc_p + c_r E[N_2(L_2)] + \psi c_f (E[N_1(L_1)] + E[N_2(L_2)] - N_{\max}) \quad (22)$$

将式(18)代入式(22)并化简,可得:

$$C(L_2, m) = \begin{cases} mc_p + c_r \sum_{k=1}^m B_{n+1+k} \int_{A_{n+1} + \alpha_{n+1}x + C_m(L_2/m)}^{A_{n+1} + \alpha_{n+1}x + C_m(L_2/m) + L_2/m} \lambda(t) dt + \\ \left(\sum_{k=1}^n B_k \int_{A_k}^{A_k + T_k} \lambda(t) dt + B_n \int_{A_n + \alpha_n T_n}^{A_n + \alpha_n T_n + x} \lambda(t) dt \right) - \psi c_f N_{\max} \\ + \sum_{k=1}^m B_{n+1+k} \int_{A_{n+1} + \alpha_{n+1}x + C_m(L_2/m)}^{A_{n+1} + \alpha_{n+1}x + C_m(L_2/m) + L_2/m} \lambda(t) dt \end{cases} \quad (23)$$

2.4 目标函数

由上可知,整个租赁期 L 的期望总成本函数 $C(L_1, L_2, n, m)$ 由 L_1 和 L_2 两个阶段的小修成本、预防维护成本以及惩罚成本构成,假定在第一、二阶段的最优期望成本分别为 $C(L_1, n^*)$ 和 $C(L_2, m^*)$,则在整个租赁期的期望总成本函数最优值 $C(L_1, L_2, n^*, m^*)$ 的表达式为:

$$C(L_1, L_2, n^*, m^*) = C(L_1, n^*) + C(L_2, m^*) \quad (24)$$

基于上述推导,给出计算最优期望总成本的算法,具体步骤如下:

Step1:参数初始化,输入参数 $\alpha, \beta, c_p, c_r, c_f, L_1, L_2, L, N_{\max}, n_{\max}, m_{\max}$ 以及 a_k, b_k 。

Step2: R_{th} 从 0.5 到 1.0,步长为 0.01,对每个 R_{th} 执行 step3 和 step4。

Step3:第一阶段优化(非等周期维护),对 $n = 1$ 到 $n_{\max}$ 满足以下过程。

(1)根据公式(6)计算第 k 个维护时间间隔 T_k 。

(2)计算期望故障数、惩罚成本。

(3)计算总成本 $C(L_1, n)$,记录最优 n^* 和期望最小成本 $C(L_1, n^*)$ 。

Step4:第二阶段优化(等周期维护),对 $m = 1$

为:

$$C_p(L_2) = \psi c_f (E[N_1(L_1)] + E[N_2(L_2)] - N_{\max}) \quad (21)$$

2.3.4 L_2 阶段期望总成本

由式(16)、(19)和(21)可得,在 L_2 内的期望总成本 $C(L_2, m)$ 为:

到 $m_{\max}$ 满足以下过程。

(1)计算等周期长度 $T = L_2/m$ 。

(2)计算期望故障数、惩罚成本(考虑两阶段的期望总故障数)。

(3)计算总成本 $C(L_2, m)$,记录最优 m^* 和期望最小成本 $C(L_2, m^*)$ 。

Step5:计算租赁期 L 的期望总成本。

$C(L_1, L_2, n^*, m^*) = C(L_1, n^*) + C(L_2, m^*)$,并记录对应 R_{th} 下的 n^* 和 m^* 。

3 算例分析

假设出租方购置某型号设备用于租赁业务,依据租赁合同,租赁期限 L 设定为 60 个月。若设备关键部件的故障发生概率服从两参数 Weibull 分布:

$$\lambda(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha} \right)^{\beta-1} \quad (25)$$

其中, α, β 分别为两参数 Weibull 分布的尺度参数和形状参数,则累积故障率函数为:

$$\Lambda(t) = \left(\frac{t}{\alpha} \right)^{\beta} \quad (26)$$

假设设备的相关参数取值如表 1 所示。

表 1 相关参数取值

β	α	$c_r/\text{次}$	$c_p/\text{次}$	$c_f/\text{次}$	$N_{\max}$	$n_{\max}$	$m_{\max}$	a_k	b_k	$L_1/\text{月}$	$L_2/\text{月}$	$L/\text{月}$
3.5	21.5	650 \$	50 \$	350 \$	1	15	15	$k/(3k+7)$	$(12k+1)/(11k+1)$	30	30	60

3.1 结果分析

将可靠度阈值的寻优范围设为 $[0.5, 1]$,步长取 0.01,该精度可满足实际维护需求。选定的 $n_{\max}$ 与 $m_{\max}$ 的取值均设得足够大,确保覆盖所有可能的

预防性维护方案。采用表 1 数据进行计算,通过变动 R_{th} 的取值得到不同组合,表 2 给出了取不同值时两阶段维护的最优期望总成本及相应最优预防维护次数。

表 2 R_{th} 取不同值时的两阶段维护最优维护成本及维护次数取值

R_{th}	n^*	m^*	$C(L_1, L_2, n^*, m^*)$
0.92	3	5	1842.3
0.91	2	5	1440.6
0.90	2	5	1466.0
0.89	2	5	1479.8
0.87	2	5	1494.5
0.85	2	5	1628.7
0.83	2	5	1703.8
0.81	2	5	1782.4
0.79	2	5	1894.8

表 2 为当可靠度阈值取不同值时,两阶段维护最优期望总成本与最优维护次数的取值情况,由表 2 可知:当可靠度阈值 $R_{th} = 0.91$ 时,所得两阶段维护期望总成本的值最小,此时 $C(L_1, L_2, n^*, m^*)$ 为 1440.6 \$,对应地,第一阶段最优维护次数 n^* 为 2 次,第二阶段最优维护次数 m^* 为 5 次,最优维护次数共计为 7 次,具体取值如图 1 所示。与文献[11]相比,当可靠度阈值 R_{th} 取值为 0.83 时,其最优值 $C(T_1^*, \dots, T_n^*)$ 为 1509.0 \$,最优维护次数亦为 7 次。文中所提出的两阶段维护方法,使得出租方获得更优的维护成本,而维护总次数相同。

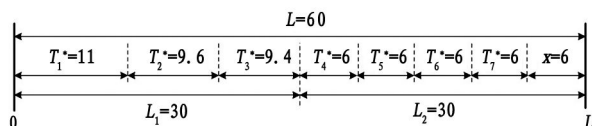
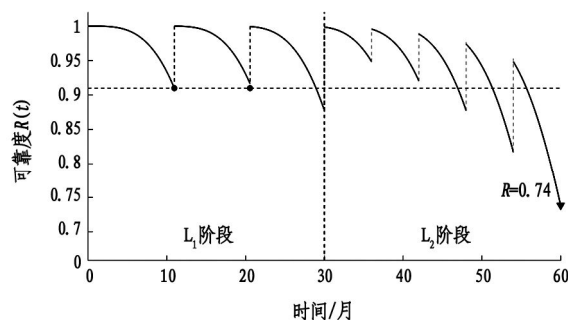


图 1 最优维护时间间隔示意图(单位:月)

由图 1 可以看出,当可靠度阈值 $R_{th} = 0.91$ 时,执行两阶段预防维护策略。根据优化结果可知,在维护的第一阶段 L_1 ,执行非等周期维护,维护时间节点为: $T_1 = 11.0, T_2 = 9.6$,在第一阶段结束节点进行的维护算作下一阶段的开始。所以,在 L_2 阶段的首次维护间隔为距离 L_1 阶段最后一次维护的时间 $T_3 = 9.4$,随后执行本阶段内等周期预防维护,维护时间节点为 $T_4 = T_5 = T_6 = T_7 = 6$ 。

图 2 为 L_1 阶段的可靠度阈值 R_{th} 取 0.91 时,设备的可靠度变化趋势图,利用算法优化可知,在 L_1 阶段执行非等间隔期预防维护,当维护次数 $n^* = 2$ 时,该阶段的期望维护总成本值最优。在 L_2 阶段,执行等间隔期预防维护,当维护次数 $m^* = 5$ 时,该

阶段的期望维护总成本值最优。经计算得出,在租赁期结束时,设备的可靠度阈值为 0.74。在租赁合同未对可靠度阈值作出明确限制的情况下,本方法优于文献[11]中的方法。

图 2 可靠度阈值 $R_{th} = 0.91$ 时的设备可靠度曲线变化图

3.2 灵敏度分析

当其它参数取值不变时,随着可靠度阈值 R_{th} 在 $[0.50, 0.97]$ 内以步长 $k = 0.01$ 变化,期望总成本最优值的变化情况如图 3 所示。由图 3 可知,期望总成本的最优值随可靠度阈值的提高呈“U”型曲线,存在唯一的最优阈值 $R_{th}^* = 0.91$ 。

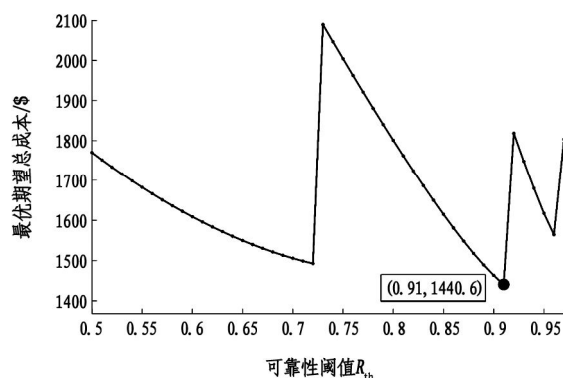


图 3 期望总成本随可靠度阈值变化曲线示意图

当其它参数取值不变,可靠度阈值 R_{th} 变化时,期望维护总次数 $n^* + m^*$ 的变化情况如图 4 所示。从图 4 中可以看出,期望维护总次数随 R_{th} 的增大呈阶梯式增长,即当可靠度阈值变大时,设备需要更多的预防维护,从而阻止故障的发生。

当其它参数取值不变, $R_{th} = 0.91$, 第一阶段的长度 L_1 在 $[15, 55]$ 个月内以步长 $k = 1$ 变化时,期望总成本的变化情况如图 5 所示。由图 5 可得结论:随着第一阶段 L_1 的长度增大,期望总成本 $C(L_1, L_2, n^*, m^*)$ 呈波动下降趋势,在 $L_1 = 42$ 时,期望总成本的值达到最小, $C(L_1, L_2, n^*, m^*) = 1297.7$ \$;当 $L_1 > 42$ 时,期望总成本 $C(L_1, L_2, n^*, m^*)$ 呈波动上升趋势。可见,两阶段长度不相等时可获得更

优的维护经济性。尽管如此,本研究为与文献[11]进行对比,并充分展示等周期与非等周期维护策略的差异,仍选取两阶段长度相等的情形进行分析,即 $L_1 = L_2 = 30$ 个月。

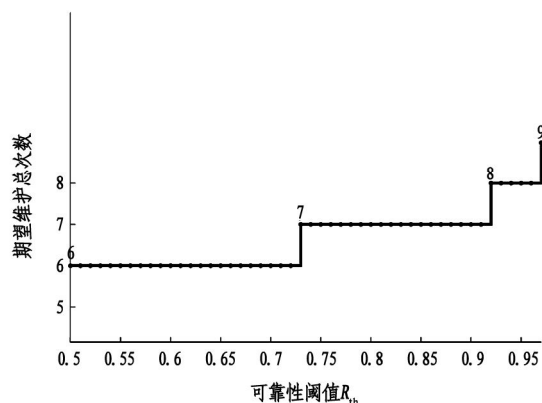


图4 期望维护总次数随可靠性阈值变化曲线示意图

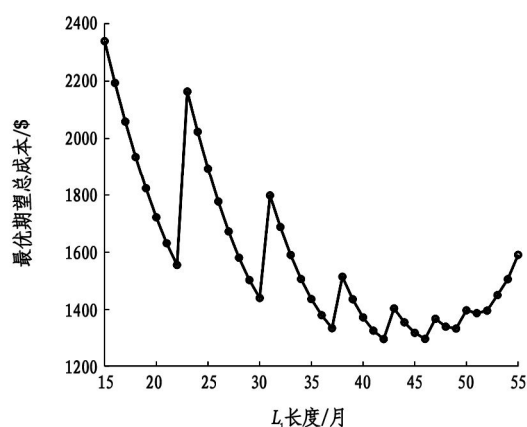


图5 期望总成本随 L_1 长度变化的示意图

4 结论

将基于可靠性阈值的非等周期维护与等周期维护相结合,避免了因设备不完全维修导致其退化速度不均而引发的“过维修”或“欠维修”问题。通过具体数值分析,验证了所提模型的可靠性,并获得了使维护总成本最小的最优维护策略,可为出租方的设备维护提供理论指导。在更广泛场景下与多种策略的对比,是下一步的研究方向。

参考文献:

- [1] Zhang Wenyi, Chen Bi, Yang Feng. Purchasing or leasing? Equipment upgrades for a capital-constrained manufacturer [J]. Journal of Environmental Management, 2025, 373: 123635.
- [2] 张云正,曾建潮,张晓红,等. 多设备租赁条件下基于服务质量评估的预防性维修策略优化[J]. 计算机集成制

造系统,2018,24(5):1162-1170.

- [3] Ben Mabrouk A, Chelbi A, Radhoui M. Optimal imperfect preventive maintenance policy for equipment leased during successive periods [J]. International Journal of Production Research, 2016, 54(17): 5095-5110.
- [4] 席启超,曹继平,陈桂明,等. 一种基于可靠度约束的分阶段等周期预防性维修模型研究[J]. 兵工学报, 2017, 38(11): 2251-2258.
- [5] 张思源,陆志强,崔维伟. 考虑设备周期性维护的流水车间生产调度优化算法[J]. 计算机集成制造系统, 2014, 20(6): 1379-1387.
- [6] 卢震,郭巧顺,徐健. 基于周期性不完全预防性维修的最优经济生产批量决策[J]. 系统工程理论与实践, 2017, 37(10): 2621-2629.
- [7] 李雯静,邱莉,林志勇,等. 基于役龄回退模型的露天矿卡车周期预防性维护策略[J]. 煤炭科学技术, 2023, 51(4): 209-214.
- [8] 黄尧,刘勤明,倪静然,等. 考虑质量控制的设备维护与批量生产联合优化研究[J]. 机械设计, 2026, 43(2): 102-109.
- [9] Zhou Xiaojun, Xi Lifeng, Lee Jay. Reliability-centered predictive maintenance scheduling for a continuously monitored system subject to degradation [J]. Reliability Engineering & System Safety, 2007, 92(4): 530-534.
- [10] 彭聪,上官伟,王宗耀,等. 考虑单位成本效益的列控车载设备预防性维修决策研究[J]. 铁道学报, 2024, 46(12): 77-85.
- [11] Ben Mabrouk A, Chelbi A, Radhoui M. Optimal hybrid imperfect preventive maintenance policy for leased equipment [J]. IFAC-PapersOnLine, 2017, 50(1): 13698-13703.
- [12] 李建华,徐家生,任丽娜. 可靠度约束下的不完全预防更换维修模型[J]. 太阳能学报, 2022, 43(4): 446-452.
- [13] 王智明,王兆旭. 可靠度约束下考虑剩余寿命的单元部件可修复系统不完全预防维护[J]. 机械设计与研究, 2025, 41(4): 100-104, 112.
- [14] 姚梦月,董克,李少付,等. 二手设备租赁最优混合预防维护策略优化[J]. 蚌埠学院学报, 2025, 14(2): 99-105.
- [15] 曲玉祥,吴甦. 基于不完全维护的劣化系统分阶段顺序预防维护策略[J]. 机械工程学报, 2011, 47(10): 164-170.
- [16] Zhou Xiaojun, Li Yanting, Xi Lifeng, et al. Multi-phase preventive maintenance policy for leased equipment [J]. International Journal of Production Research, 2015, 53(15): 4528-4537.
- [17] 邓超,段超群,吴军. 基于设备可靠度的分阶段顺序维修经济优化模型[J]. 计算机集成制造系统, 2016, 22(2): 568-575.

责任编辑:刘娟

基于决策树分析的车间生产工位竞争研究

杨露*, 石林

(合肥大学 设计学院, 安徽 合肥 230061)

摘要: 车间生产调度作为车间生产的关键环节,其效率直接影响车间生产的效益。以DCT车间为研究对象,通过分析车间生产工位竞争的影响因素,采用CHAID决策树算法进行数据分析,构建工位竞争决策树分析模型并用决策树算法获取工位竞争决策规则,对车间生产进行优化分析,确定出车间生产调度新策略。相对于原策略,生产工艺基本各工位效率均得到提高,进而提高了车间生产调度的效率,对提升车间生产调度决策,改进车间生产效率具有重要的意义。

关键词: 工位竞争;因素选取;CHAID决策树算法;决策规则;优化分析

中图分类号:F406.2

文献标识码:A

文章编号:2095-297X(2026)05-0106-07

Research on Workshop Production Workstation Competition Based on Decision Tree Analysis

YANG Lu*, SHI Lin

(School of Design, Hefei University, Hefei, 230061, Anhui)

Abstract: As the key link in the workshop production, the efficiency of workshop production scheduling is directly associated with the economic benefits of workshop production. The DCT workshop was selected as the research object, and the research was conducted to analyze the influencing factors of workstation competition in workshop production, with the CHAID decision tree algorithm adopted for data analysis. A decision tree analysis model for workstation competition was constructed, and the decision rules for workstation competition were obtained by using the decision tree algorithm, so that an optimization analysis of workshop production can be conducted. It has been confirmed that compared with the original strategy, the efficiency of virtually all workstations in the production process has been improved under the new workshop production scheduling strategy, and thus the efficiency of workshop production scheduling has been enhanced. This is of great significance for the improvement of decision-making in workshop production scheduling and the promotion of the overall production efficiency of the workshop.

Key words: workstation competition; factor selection; CHAID decision tree algorithm; decision rules; optimization analysis

车间生产调度作为车间生产的关键环节,在制造业企业中占有重要地位,其效率直接影响车间生产的效益。工位竞争作为车间生产调度中的重要问题,具有很大的研究价值,国内外学者对此已进行大量研究。唐艺军等^[1]建立高维多目标柔性车间调

度模型,利用改进的NSGA-II算法进行求解,验证改进算法的有效性。王豪渊等^[2]对传统遗传算法进行自适应改进,建立数学模型,更快求得柔性作业车间调度问题的最优解。钟宏扬等^[3]研究加工与装配的双阶段调度问题,构建基于遗传算法与优先分

收稿日期:2025-12-24 * 通讯联系人

基金项目:安徽省高校哲学社会科学研究重点项目(2023AH052157);安徽省高等学校省级本科质量工程建设类项目(2023zygzs044)。

作者简介:杨露(1989—),女,安徽六安人,助理实验师,硕士。E-mail:969411365@qq.com

派规则的混合求解算法,验证调度的有效性。王亮等^[4]应用近端策略优化强化学习算法与虚拟仿真技术研究作业车间集成调度问题。陈雯雯^[5]提出基于数据挖掘和调度规则的遗传优化算法(GA-DDR)来解决作业车间调度问题。姚层林^[6]构建AHP-DEMATEL-ISM综合分析模型,开展离散智能制造车间生产效率优化研究。ZAHMANI等^[7]实时解决动态作业车间调度问题,采用遗传算法与数据挖掘技术,为待加工工件选择最优调度规则,实现实时调度。陈睿奇等^[8]将Job shop的加工工序构建为工序智能体,并基于工序特征和最早加工时间设计拟合动作值函数的深度神经网络,调度模型根据系统状态和工序智能体特征即可生成调度策略。MUHAMMAD等^[9]研究基于遗传算法求解柔性作业车间调度问题。PALOMBARINI等^[10]训练深度Q网络来选择智能体的动作以修复调度计划。

目前国际上对车间生产调度的研究日益增多,而工位竞争是车间生产调度中的重要问题,同时数据挖掘方法也是目前数据分析领域中的重要方法,但是在分析此类问题中的应用相对较少。基于此,笔者利用车间生产调度的大量历史数据研究车间生产调度中的工位竞争问题,获取车间生产工位竞争的决策规则,改进车间生产工位竞争策略,提高车间的生产调度效率。

1 问题描述及工艺分析

1.1 问题描述

车间作为制造型企业的重要组成部分,车间生产在企业生产中占据重要地位。车间生产调度作为车间生产的关键环节,主要包含工件的调度、机械的调度、运输的调度和仓储的调度等。工位竞争作为车间生产调度中的重要问题,具有很大的研究价值。

针对双离合变速器(dual clutch transmission, DCT)车间生产调度中的工位竞争问题,工件遵循怎样的决策规则将直接影响工位竞争处的工件生产调度顺序。因此分析影响车间生产工位竞争因素,利用决策树算法进行数据分析,获取有效的车间生产工位竞争决策规则,并应用此车间生产工位竞争决策规则来解决有限资源下的工件排序问题,得到一个可靠的车间生产调度策略是十分必要的。

1.2 工艺分析

DCT车间生产工艺流程包括变速器三种自制零件(轴承自制零件、齿轮自制零件、主壳体自制零件)的加工及这三种自制零件进入整个产品的装配,这三种自制零件只是最终产品装配中的三个零

件,属于整个产品装配的一部分,最终生产得到产品是DCT变速器。

整个生产线上的加工采取先进先出(FIFO)规则,分析DCT车间整个生产工艺流程,发现三种自制零件(轴承自制零件、齿轮自制零件、主壳体自制零件)的单独加工与变速器装配工艺过程中存在同一个工序中同一台设备上工件竞争加工的情况,即存在车间生产工位竞争加工冲突问题。存在竞争的工序主要有6道,分别是工序OP60(精珩内孔)、工序OP90(研磨另一端面、锥面)、工序OP120(打印标记)、工序OP230(安装双离合器堵头)、工序OP260(拧紧联接螺栓)、工序OP270(安装阀体一)。以往采取的先进先出调度规则在竞争工位处往往不会得到最佳的结果,考虑应用规则调度法提取更加行之有效的规则,重新分配工艺的加工次序,可以解决车间生产调度中的工位竞争问题,更好地起到了改进车间生产调度的作用。

2 因素选取及数据获取

2.1 因素选取

研究以满足产品订单交货期和生产周期为目标的车间生产工位竞争问题,基于逆序排序的方法提取车间工位竞争决策规则,即先加工的工序后排序,后加工的工序先排序。基于上述提取原则和因素选取实施方法,选取以下因素作为车间生产工位竞争问题的影响因素。其中绝对剩余排序时间(absolute remaining sorting time, ARST)和工件交货期(delivery time, DT)可通过宏观信息直接选取而得,相对剩余排序时间(relative remaining sorting time, RRST)和宽松比率(loose rate, LR)可通过相关计算得到。

$$ARST = \sum_{i=1}^k t_{\text{behind}} \quad (1)$$

$$RRST = \frac{ARST}{T} \quad (2)$$

$$LR = \frac{DT}{ARST} \quad (3)$$

绝对剩余排序时间用ARST表示,相对剩余排序时间用RRST表示,工件交货期用DT表示,宽松比率用LR表示,工件总加工时间用T表示。相对剩余排序时间RRST为绝对剩余排序时间ARST与工件总加工时间T的比值,宽松比率LR为工件交货期DT与绝对剩余排序时间ARST的比值。

2.2 数据获取

2.2.1 数据准备

选取DCT车间一段时间(1万套1个月排班计划)内的车间生产数据作为实际研究案例,以此段时

期内的加工节拍、工序顺序、设备加工零件种类等数据作为研究对象,数据来源于该车间的数据库。数据库中的信息主要包括工件信息、排班信息、设备信息等。共收集了与本次数据挖掘任务相关的一个数据库中三个子数据库的信息数据,时间跨度为 1 个月,合计为 9873 条数据。

根据选取的影响因素进行数据获取工作,对多表进行连接,得到了与此次挖掘密切相关的信息数据表,其内容包括如下:工序号、加工节拍、加工顺序、开工时间、完工时间、设备上前一道工序的结束时间、工件交货期。通过得到的信息数据表获得最终的数据挖掘表,其内容包括工序号、绝对剩余时间(ARST)、相对剩余时间(RRST)、工件交货期(DT)、宽松比率(LR)。具体各数据表之间的关系如图 1 所示。

2.2.2 数据预处理

当车间生产调度中工序存在工位竞争情况时,分别比较上述四项影响因素(ARST、RRST、DT、LR)相关数据的大小,此四项数据的初始值需要通过数据预处理得到。以工序 OP60 处的工位竞争部分数

据为例,说明决策树算法是如何应用于车间生产工位竞争数据预处理的,工序 OP60 工位竞争部分初始数据如表 1 所示。

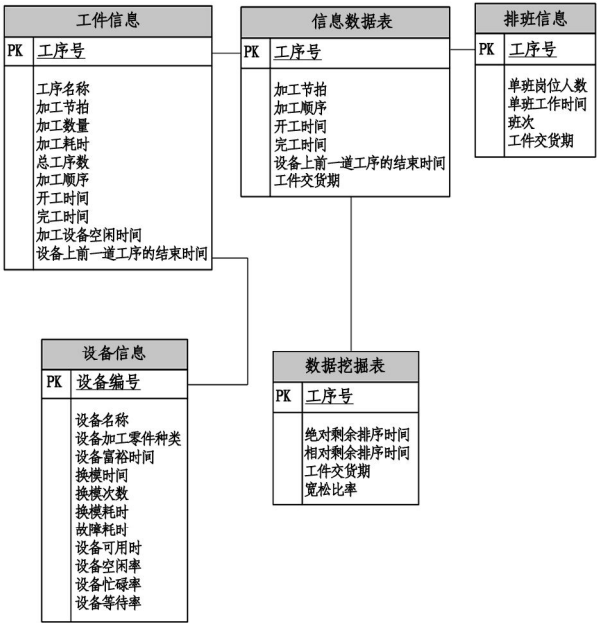


图 1 数据表关系图

表 1 工序 OP60 工位竞争部分初始数据

齿轮到 OP60 时间	主壳体到 OP60 时间	齿轮工件完成时间	主壳体工件完成时间	齿轮交货期	主壳体交货期
0: 03: 10	0: 18: 00	0: 13: 56	0: 28: 15	0: 26: 07	0: 38: 14
0: 04: 20	0: 22: 22	0: 15: 18	0: 38: 14	0: 28: 26	0: 40: 33
0: 05: 30	0: 26: 44	0: 26: 07	0: 41: 27	0: 30: 47	0: 42: 53
0: 06: 40	0: 31: 06	0: 26: 56	0: 44: 40	0: 33: 06	0: 45: 39
0: 07: 50	0: 35: 28	0: 28: 26	0: 47: 53	0: 35: 27	0: 48: 52
0: 09: 00	0: 39: 50	0: 29: 15	0: 50: 29	0: 37: 46	0: 51: 29
0: 10: 10	0: 40: 38	0: 30: 47	0: 54: 08	0: 40: 07	0: 55: 08
0: 11: 20	0: 47: 46	0: 31: 35	0: 58: 30	0: 42: 27	0: 59: 30
0: 12: 30	0: 52: 08	0: 33: 06	1: 02: 52	0: 44: 46	1: 03: 52
0: 13: 40	0: 56: 30	0: 33: 55	1: 07: 14	0: 47: 07	1: 08: 14

DCT 车间生产工位竞争问题中,记齿轮工件到 OP60 的时间为 ARSTB,齿轮工件完成时间为 TB,齿轮工件交货期为 DTB;主壳体工件到 OP60 的时间为 ARSTC,主壳体工件完成时间为 TC,主壳体工件

交货期为 DTC,进而算出 RRSTB、LRB、RRSTC、LRC。

上述数据集成和转换后的工序 OP60 工位竞争部分数据如表 2 所示。

表 2 数据集成和转换后工序 OP60 工位竞争部分数据

NO	ARSTB	RRSTB	DTB	LRB	ARSTC	RRSTC	DTC	LRC
1	0: 03: 10	0. 23	0: 26: 07	8. 25	0: 18: 00	0. 64	0: 38: 14	2. 12
2	0: 04: 20	0. 28	0: 28: 26	6. 56	0: 22: 22	0. 59	0: 40: 33	1. 81
3	0: 05: 30	0. 21	0: 30: 47	5. 60	0: 26: 44	0. 65	0: 42: 53	1. 60

续表

表2 数据集成和转换后工序 OP60 工位竞争部分数据

NO	ARSTB	RRSTB	DTB	LRB	ARSTC	RRSTC	DTC	LRC
4	0:06:40	0.25	0:33:06	4.97	0:31:06	0.70	0:45:39	1.47
5	0:07:50	0.28	0:35:27	4.52	0:35:28	0.74	0:48:52	1.38
6	0:09:00	0.31	0:37:46	4.20	0:39:50	0.79	0:51:29	1.29
7	0:10:10	0.33	0:40:07	3.95	0:40:38	0.75	0:55:08	1.36
8	0:11:20	0.36	0:42:27	3.74	0:47:46	0.82	0:59:30	1.25
9	0:12:30	0.38	0:44:46	3.58	0:52:08	0.83	1:03:52	1.23
10	0:13:40	0.40	0:47:07	3.45	0:56:30	0.84	1:08:14	1.21

综合上述数据,得到需要的车间生产工位竞争数据挖掘表部分数据如表3所示。

表3 车间生产工位竞争数据挖掘表部分数据

NO	ARST	RRST	DT	LR
1	0:17:10	0.45	0:54:07	3.15
2	0:18:00	0.64	0:38:14	2.12
3	0:18:48	0.49	0:56:27	3.00
4	0:22:22	0.59	0:40:33	1.81
5	0:23:10	0.58	0:58:47	2.54
6	0:26:44	0.65	0:42:53	1.60
7	0:27:32	0.67	1:01:07	2.22
8	0:31:06	0.70	0:45:39	1.47
9	0:31:54	0.75	1:03:27	1.99
10	0:35:28	0.74	0:48:52	1.38

对于车间生产调度中工位竞争的6道工序(工序 OP60、工序 OP90、工序 OP120、工序 OP230、工序 OP260、工序 OP270),分别比较 ARST、RRST、DT、LR 的大小,并将比较结果作为该样本的属性取值,参与比较的工序数即为该样本属性的取值数。数值较小的记为 S,数值较大的记为 L,若出现数据相等的情况,记为 S。在排序结果中,将该工序的加工顺序作为该样本的类别值,先加工的工序记为类别 A,后加工的工序记为类别 B。现截取车间生产工位竞争数据挖掘转换表部分数据如表4所示。

经过数据准备及数据预处理,数据得到了有效缩小,有用的数据挖掘对象全部集中到表4中。此表为即将进行数据挖掘的研究对象,表4仅为该表中的部分数据。

表4 车间生产工位竞争数据挖掘转换表部分数据

NO	ARST	RRST	DT	LR	TYPE
1	S	S	L	L	A
2	L	L	S	S	B
3	S	L	S	S	A
4	L	S	L	L	B
5	S	S	L	L	A
6	L	L	S	S	B
7	S	L	S	S	A
8	L	S	L	L	B
9	S	S	L	L	A
10	L	L	S	S	B

3 基于决策树分析的工位竞争决策

3.1 CHAID 决策树算法设计

以表4中的数据作为训练样本,首先选定表中的类别值(TYPE)作为目标变量,然后选定属性值(ARST、RRST、DT、LR)与类别值(TYPE)进行交叉分类,产生一系列二维频数表;分别计算所生成二维分类表的卡方统计量或似然估计统计量,比较统计量的大小,以最大统计量的二维表作为最佳初始分类表,在最佳二维分类的基础上继续使用属性值(ARST、RRST、DT、LR)对类别值(TYPE)进行分类,重复上述过程直到分类条件满足为止。

令 y 为目标变量 TYPE, $x_1, x_2, \dots, x_m$ 为输入变量 ARST、RRST、DT、LR。

CHAID 算法具体如下:

(1) 建立交叉分类表。将变量 $x_k (1 \leq k \leq m)$ 与 y 进行交叉分类,形成一个二维交叉分类表。所有的 x_k 都与 y 进行交叉分类,得到 m 个交叉分类表。

(2) 计算每个交叉分类表的卡方统计量或极大似然估计值

$$\text{卡方统计量: } \chi^2 = \sum_j \sum_i \frac{(f_{ij} - \hat{F}_{ij})^2}{\hat{F}_{ij}} \quad (4)$$

$$\text{极大似然估计: } L^2 = 2 \times \sum_i \sum_j f_{ij} \ln \frac{f_{ij}}{\hat{F}_{ij}} \quad (5)$$

其中 $f_{ij} = y_{ij}/n$, 为第 k 个交叉分类表 x_k 第 i 类第 j 类的实际分布频率, y_{ij} 为分布频数, n 为分类表样本总数。为 $\hat{F}_{ij}$ 与 f_{ij} 对应的期望分布频率的估计值, $\hat{F}_{ij} = (y_{i.}/n) \times (y_{.j}/n)$, $y_{i.}$ 和 $y_{.j}$ 分别为二维交叉分类表中第 j 列元素总和与第 i 行元素总和。

(3) 选择分类变量。比较 m 个交叉分类表 χ^2 或 L^2 相伴概率 P 的大小。假如 x_k 与 y 的 P 值最大, 则选 x_k 为最佳交互分类方法, 即 x_k 与 y 的交叉分类最能体现 y 的分布差异。

(4) 确定分类方向。对已分好的最优二维表继续对 y 进行交叉分类(排除已使用过的最优变量)形成三维交叉表。重复步骤(1)至(5), 就得到不同阶段多维交互表并找到针对 y 的最优分类。

(5) 分类停止条件确定。一般可以用以下三种方法确定停止分类的条件: 根据预先给定的显著性水平 α , 如果 χ^2 或 L^2 统计量相伴概率 $P < \alpha$, 则停止; 设置树的深度, 如果决策树的层数超过给定的最大值, 则停止; 设置最小样本数, 如果交叉分类的样本数小于给定的最少样本数, 则停止。

3.2 CHAID 决策树构建

应用 CHAID 决策树算法解决车间生产工位竞争问题, 主要在于根据确定的因变量确定决策树模型的树根节点, 即所有的车间生产数据作为树根节点的全部数据, 进一步根据目标变量 TYPE 和输入变量 ARST、RRST、DT、LR 之间的相关程度进行分组, 将相关性强的分为一组, 据此实现相关性程度的不同分组情况。按照卡方显著性不断划分或组合为树状结构, 实现最终的决策树分类过程。根据决策树生成的调整 P 值及卡方值得出决策树模型拟合度, 确定工位竞争决策树模型是否倾向于被接受, 据此被接受的决策树模型得到车间生产工位竞争决策规则。

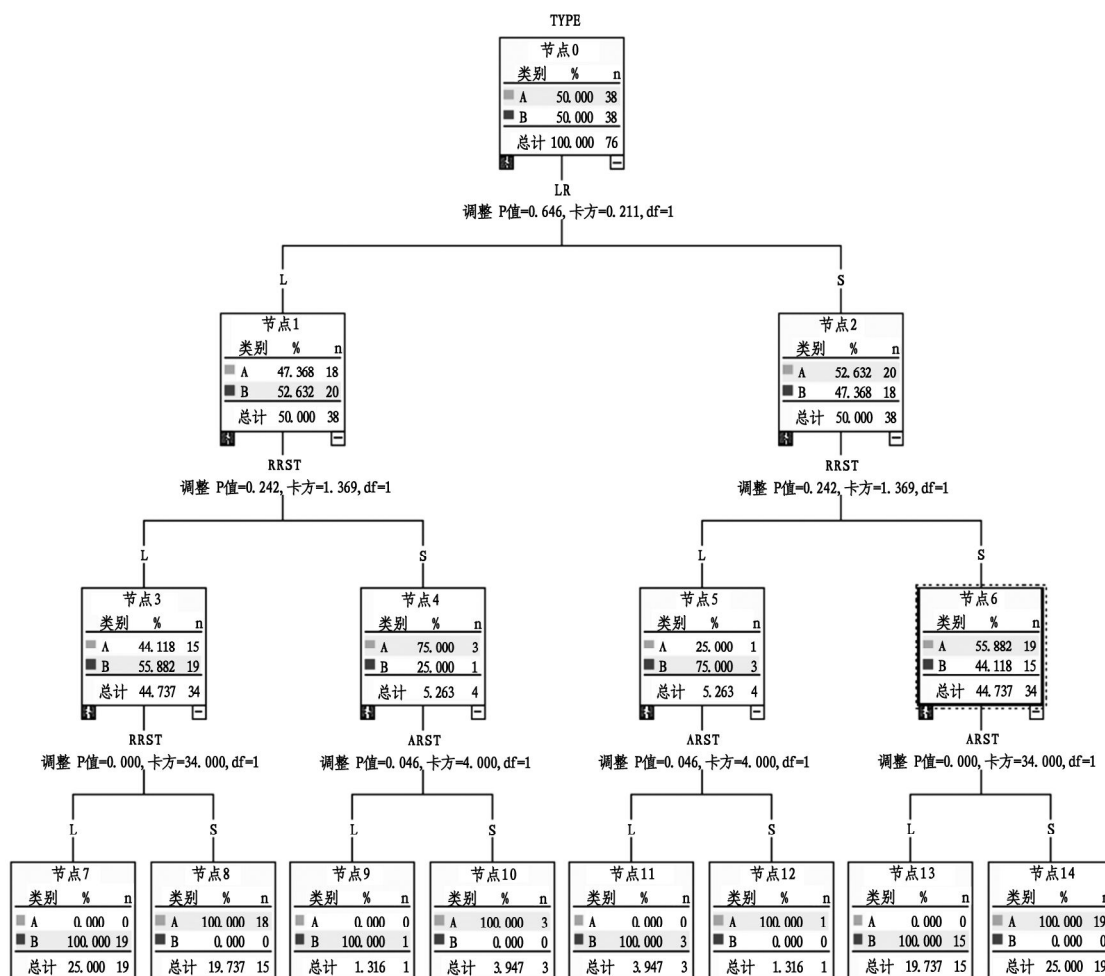


图2 工序 OP90 决策树模型

用表 4 的样本数据作为 CHAID 决策树构建的训练样本数据,以工位竞争处工件加工顺序类别 TYPE 作为目标变量,记类别 A 为 1,类别 B 为 0,以 ARST、RRST、DT、LR 为输入变量,取显著性水平 α 为 0.05,使用 Bonfermni 法调整 P 值,建立 CHAID 决策树模型。

以工序 OP90 处两个工件竞争同一台设备的问题为例,构建工位竞争决策树模型,从各竞争工位相关加工冲突信息和加工顺序结果等数据中提取车间生产工位竞争决策规则,用以说明利用决策树算法解决车间生产工位竞争的问题。依据决策树算法过程,应用于表 4,最终得到如图 2 所示的工序 OP90 决策树模型。

观察到决策树在分割生长分支时是通过调整 P 值及卡方值控制的。最终调整 P 值等于 0,即该决策树模型达到了统计性的显著。在这种情况下,卡方值显著, P 值越小越好,即车间生产数据挖掘表中的数据之间以模型中的关系进行关联,要远好于数据之间没有任何关系。因此,得出结论:模型拟合度较高,倾向于接受此工序 OP90 决策树模型。

至此,应用决策树算法得到 DCT 车间生产 6 个竞争工序处(工序 OP60、工序 OP90、工序 OP120、工序 OP230、工序 OP260、工序 OP270)的决策树模型,通过对决策树模型的分析获取车间生产工位竞争决策规则,将该车间生产工位竞争决策规则应用于车间生产调度中,指导 DCT 车间实际车间生产调度。

3.3 决策规则获取

图 2 工序 OP90 决策树模型共 4 层 15 个节点,其中,节点 0 为根节点;节点 1、2、3、4、5、6 为中间节点;节点 7、8、9、10、11、12、13、14 为叶子节点,每一个叶子节点到根节点的路径构成一个分类规则,8 个叶子节点共对应 8 条分类规则。由该模型得到的对应车间生产工位竞争决策规则如表 5 所示。

表 5 工序 OP90 决策树模型对应工位竞争决策规则

NO	LR	RRST	ARST	RESULT
1	L	L	L	B
2	L	L	S	A
3	L	S	L	B
4	L	S	S	A
5	S	L	L	B
6	S	L	S	A
7	S	S	L	B
8	S	S	S	A

根据上述工序 OP90 决策树模型,从根结点到叶结点的每一条路径都可以是一条车间生产工位竞争决策规则,从表 5 可以得到如下 8 条车间生产工位竞争决策规则。

- (1) If LR = L, RRST = L, ARST = L Then RESULT = B
- (2) If LR = L, RRST = L, ARST = S Then RESULT = A
- (3) If LR = L, RRST = S, ARST = L Then RESULT = B
- (4) If LR = L, RRST = S, ARST = S Then RESULT = A
- (5) If LR = S, RRST = L, ARST = L Then RESULT = B
- (6) If LR = S, RRST = L, ARST = S Then RESULT = A
- (7) If LR = S, RRST = S, ARST = L Then RESULT = B
- (8) If LR = S, RRST = S, ARST = S Then RESULT = A

以规则(8)为例对上述车间生产工位竞争决策规则的意义进行说明:在某一时刻,多道工序竞争同一台设备时,在逆序的情况下,宽松比率小、相对剩余排序时间小、绝对剩余排序时间小的优先排序。

同理,可得到另外 5 个存在竞争的工序处(工序 OP60、工序 OP120、工序 OP230、工序 OP260、工序 OP270)的车间生产工位竞争决策规则。将此方法生成的车间生产工位竞争决策规则应用于车间生产中,获得数据挖掘决策规则,得出车间生产调度新策略,并建立车间生产仿真系统评估模块,可以用于指导 DCT 车间实际生产调度。

4 工位竞争优化分析

基于上述 CHAID 决策树算法形成决策树模型得到的车间生产工位竞争决策规则,利用仿真技术,验证车间生产工位竞争决策规则应用于车间生产得到的新策略相对于原策略的有效性,据此分析 DCT 车间生产调度新策略是否起到优化的作用。对比 DCT 车间生产调度原策略与根据数据挖掘得到的新策略,得出结论,通过对车间生产工艺流程中各工位效率进行比较,为提升车间生产调度决策,改进车间生产效率提供依据。轴承工艺、齿轮工艺、主壳体工艺及装配工艺在各工位处的效率,如图 3 至图 6 所示。

由图 3 至图 6 可以看出使用决策树算法得到的车间生产工位竞争决策规则集成于车间生产调度新策略后基本各工位效率均得到提高。车间生产调度中加工工艺(轴承工艺、齿轮工艺及主壳体工艺)和装配工艺工序的各工位效率,将影响车间生产调度的效率。以轴承工艺为例,对其进行分析,得到在轴承工艺加工过程中,新策略下各工位的效率相比原策略均有提高,其中 OP200、OP230、OP260、OP270 工位效率提高尤为显著,分别提高 12%、14%、14%、16%,其余 OP210、OP225、OP250、OP280 工位

效率也分别提高 3%、2%、2%、5%，证明新策略起到了改进竞争工位效率的作用。

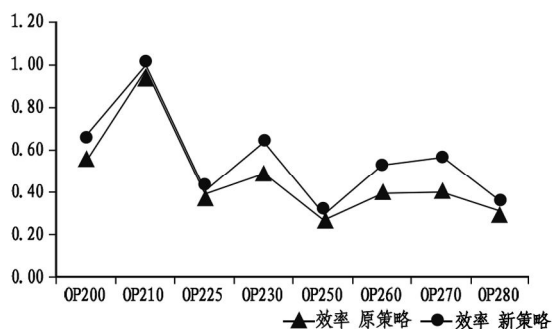


图 3 各工位的轴承工艺效率

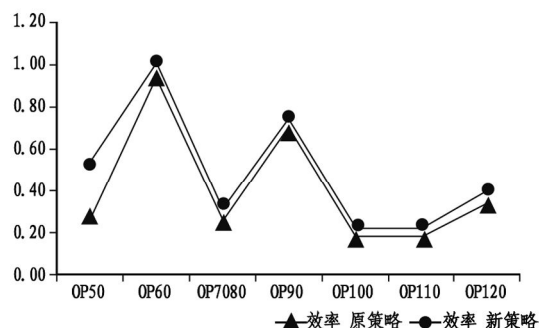


图 4 各工位的齿轮工艺效率

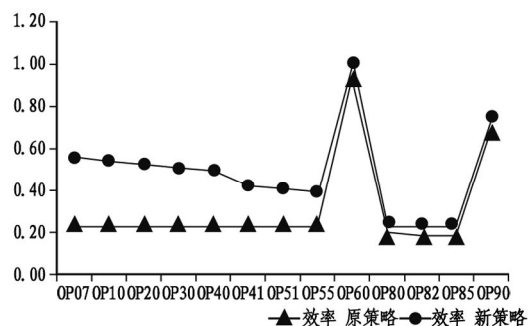


图 5 各工位的主壳体工艺效率

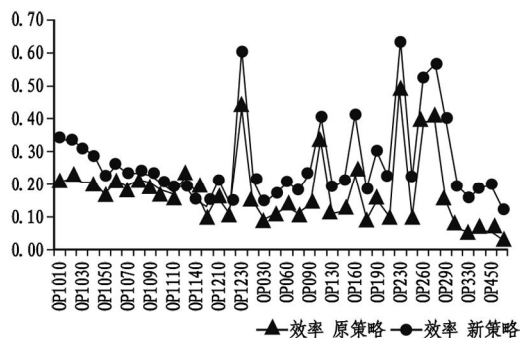


图 6 各工位的装配工艺效率

综上所述,决策树算法得到的车间生产工位竞争决策规则使各工位效率均得到提高。证明本次数据挖掘实现了原有目的,通过改进车间工位的效率,改善了车间生产调度的效率,提升了 DCT 车间生产调度决策能力,进而改善了整体生产调度效果,实现了生产效率的提升。

5 结论

将 CHAID 决策树算法应用于 DCT 车间工位竞争问题中,结合车间生产工位竞争影响因素(绝对剩余排序时间、相对剩余排序时间、工件交货期、宽松比率),分析了车间生产工位竞争历史数据,构建了工位竞争决策树模型,并获取对应的车间生产工位竞争决策规则,确定车间生产调度新策略。新策略相对于原策略,各工位效率均得到提高,进而来改善车间生产调度的效率,达到了优化 DCT 车间生产调度决策,改进车间生产效率的目的。

参考文献:

- [1] 唐艺军,王艳灵,孙浩强. 基于改进 NSGA-II 算法的高维多目标柔性冲压车间生产调度优化[J]. 锻压技术, 2023,48(11):204-211.
- [2] 王豪渊,文涛. 改进自适应遗传算法在柔性车间调度的应用[J]. 自动化应用,2023(17):178-181.
- [3] 钟宏扬,刘建军,黎英杰,等. 多层级装配作业车间分批调度算法设计与分析[J]. 工业工程,2023,26(4):124-134.
- [4] 王亮,李吴谦,唐堂,等. 基于虚拟仿真与深度强化学习的作业车间集成调度[J]. 实验室研究与探索,2023,42(5):133-136,178.
- [5] 陈雯雯. 基于数据挖掘的作业车间调度问题研究[D]. 沈阳:沈阳工业大学,2022.
- [6] 姚层林. 离散智能制造车间生产效率优化研究[D]. 武汉:武汉科技大学,2025.
- [7] Zahmani M H, Atmani B. Multiple dispatching rules allocation in real time using data mining, genetic algorithms, and simulation[J]. Journal of Scheduling, 2020, 24:1-22.
- [8] 陈睿奇,黎雯馨,王传洋,等. 基于深度强化学习的工序交互式智能体 Job shop 调度方法[J]. 机械工程学报, 2023,59(12):78-88.
- [9] Muhammad K A, Shahid I B, Rubeena K, et al. Recent research trends in genetic algorithm based flexible job shop scheduling problem[J]. Mathematical Problems in Engineering, 2018(8):1-32.
- [10] Palombarini J A, Martinez E C. Closed-loop rescheduling using deep reinforcement learning[J]. IFAC-PapersOn-Line, 2019, 52(1):231-236.

责任编辑:刘娟

技术工具扩张对高校思想政治教育价值实现的冲击与应对策略

梁祥凤

(安徽中医药大学 马克思主义学院,安徽 合肥 230012)

摘要:随着技术不断革新,高校思想政治教育为了适应时代发展和提高教育成效,多样化技术工具的使用必不可少,但是技术工具的过度使用,必将导致高校思想政治教育价值失衡。针对技术工具过度扩张导致高校思想政治教育面临价值虚化、主体淡化和思维弱化三重困境,提出“制度-技术-队伍-评价”四位一体的约束机制,为平衡技术工具使用与高校思政教育价值的实现提供系统性解决方案。

关键词:技术工具扩张;思想政治教育;价值实现

中图分类号:G641

文献标识码:A

文章编号:2095-297X(2026)05-0113-05

Impact of Technological Tool Expansion on the Realization of Educational Value in University Ideological and Political Education and Coping Strategies

LIANG Xiangfeng

(School of Marxism, Anhui University of Chinese Medicine, Hefei, 230012, Anhui)

Abstract: With continuous technological innovation, the use of diverse technical tools has become essential for universities adapting to improve the effectiveness of ideological and political education. However, excessive use of these tools may lead to a value imbalance. To address the triple challenges of the devaluation of value, the weakening of subjectivity and the erosion of thought caused by the overexpansion of technology, a “system-technology-workforce-evaluation” integrated constraint mechanism was proposed in this study, which will offer a systematic solution to balance technological application and the realization of educational values in higher education’s ideological and political work.

Key words: technological tool expansion; ideological and political education; value realization

在数智技术全方位渗透社会肌理的当下,以人工智能、大数据、云计算为代表的技术工具正以前所未有的速度重塑着教育生态^[1]。高校思想政治教育作为培育时代新人、传播主流价值的重要阵地,在技术浪潮的裹挟下,既迎来了创新发展的历史机遇,也面临着价值实现被冲击、被消解的严峻挑战。如何在技术工具扩张的背景下,坚守高校思想政治教育的价值内核,破解其价值实现过程中的重重困境,成为新时代思想政治工作者必须直面的重大课题。

1 技术工具过度扩张造成思政教育价值实现的三重困境

技术工具过度扩张,指的是高校思想政治教育领域中,一种过度推崇技术手段、效率管控与实用目标的管理倾向^[2]。这种倾向借助量化考核、智能平台、数据画像等技术载体,将复杂且充满思想价值引领的育人过程,粗暴简化为一组组冰冷可测的指标:用线上资源的点击率衡量学生的思考深度,靠算法

收稿日期:2026-05-06

基金项目:安徽省教育厅高校质量工程思想政治理论课教研重大项目(2022sxzz045);安徽省高校质量工程高等教育重大决策部署研究项目(2022jchs030)。

作者简介:梁祥凤(1972—),女,安徽巢湖人,教授,博士。E-mail:404262203@qq.com

的精准推送取代师生间的思想交流,以标准化的教学流程限制学生的个性发展。最终,教育异化为片面追求效率的“绩效管理”,不仅严重弱化了高校思想政治教育的思想内核,更让其价值导向被技术工具所裹挟,陷入“重器轻道”的发展泥沼。具体而言,主要体现在以下三个方面。

1.1 主流价值被虚化

技术工具的过度扩张,推动高校思想政治教育陷入对技术化、量化评估方法过度依赖怪圈。当社会主义核心价值观的深刻阐释,被简化为标准化题库里的选择题;当理想信念教育,异化为VR场景中走形式的打卡任务;当智能助教的机械问答,取代了师生间触及心灵的深度对话;当算法的精准推荐,消解了情感共鸣的培育空间……这种以技术量化评估成效的方式,一定程度弱化了对社会主义核心价值观、理想信念等核心内容的深层灌输与内化。

1.2 主体性被淡化

技术工具过度扩张强调单向控制和信息输出逻辑,导致教育者与受教育者之间的关系异化为机械的“施教-受教”模式^[3]。传统高校思想政治教育中强调的双向沟通、情感互动和主体间平等对话被削弱,教育主体被视为被动接收信息的“客体”,而非具有自主性和反思能力的参与者。例如,某高校推行标准化思政课程设计,将45分钟的课堂机械切割为“10分钟理论灌输+25分钟案例播放+10分钟随堂测试”的流水线操作。这种用标准化课件与计时器构建的知识输送流水线,让教师彻底沦为教学进度的执行者,而非思想的引导者,师生间“教学相长”的双向奔赴,被异化为单向的任务完成。

1.3 思维批判力被弱化

技术工具的实用导向使高校思想政治教育聚焦于短期技能训练和知识灌输,忽略了对学生批判性思维、独立判断能力的培养^[4]。在技术治理框架下,受教育者的思维过程被简化为信息处理。例如,智能教学系统通过知识点拆解、答题模式训练等标准化设计,将社会议题转化为可量化的测评指标。例如政治理论课中的辩证分析常被简化为立场选择,学生通过高频练习形成条件反射式应答,而非基于证据链的自主论证。同时,大数据推荐系统构建的信息茧房,使受教育者习惯接受符合预设观点的信息推送。当社会热点事件发生时,学生更倾向于检索既有结论而非构建分析框架。

以上三重困境源于技术工具的使用与高校思想政治教育价值实现的分离,后者强调正确的政治立场、理想信念、人文关怀和道德导向的回归。技术工

具应用虽能提升效率,但若缺乏边界约束,无序扩张,则会加剧高校思想政治教育主客体关系的隐蔽化与交互失衡。因此,需通过强化主流价值教育、优化沟通机制和完善监督反馈,以实现技术工具与高校思想政治教育价值实现的辩证统一。

2 技术工具过度扩张阻滞高校思想政治教育价值实现的发生机制

2.1 认知测量异化导致认知评价系统偏差

在技术工具深度介入高校思想政治教育的当下,量化考核正以近乎“霸权”的姿态重塑着教育评价的逻辑,将高校思想政治教育推向“指标化生存”的困境,最终引发高校思想政治教育价值的根本性异化。教育学家哈蒂在《可见的学习》中曾尖锐地指出:“过度测量反损教育本质。”这一论断在当前高校思政教育实践中不断得到现实印证。当“抬头率”“作业完成度”“考试通过率”等冷冰冰的量化指标^[5],取代了对学生思想深度、价值认同与道德践行等核心维度的考量,成为衡量高校思想政治理论课成效的主要标准时,一种“可测量性崇拜”便悄然兴起。这种崇拜迫使教学实践向标准化流程严重倾斜——教师为追求亮眼的数据表现,不得不将原本充满思辨性与人文温度的高校思想政治教育课堂,异化为围绕指标展开的“精准生产”。

更值得警惕的是,这种以量化为核心的评价导向,正在潜移默化地侵蚀学生的心理结构与价值体系。当学生长期处于单一指标主导的评价环境中,其自我认知也逐渐被数据所绑架。考试分数、作业完成情况等指标,成为衡量自我价值的唯一尺度。一旦遭遇学业挫折或职业发展的不确定性,那些长期受量化标准规训的学生,极易因抗挫能力薄弱而陷入“习得性无助”。部分学生可能开始质疑思想政治教育所倡导的理想信念与价值追求,甚至滑向价值虚无主义的深渊。他们不再相信崇高理想能够指引方向,也不再认同道德规范具有行为约束力,转而将功利主义与实用主义奉为人生信条,最终背离了“立德树人”的根本目标。

2.2 制度运行机制梗阻

高校思想政治教育的制度体系是保障育人目标实现的骨架与脉络,然而在技术工具过度扩张的背景下,这一体系正面临着“碎片化”执行与“数字化”异化的双重梗阻,使得制度的育人效能难以有效释放,甚至在一定程度上成为技术工具消解高校思想政治教育价值的“帮凶”。

这种困境的根源在于管理架构的分散性与协同

机制的缺失。高校思想政治教育的实施主体被分散配置于教务处、学工部、团委等多个职能部门,每个部门都承担着部分思想政治教育职责:教务处负责思想政治课程的教学安排与考核,学工部聚焦学生日常思想政治管理,团委则负责各类思想政治主题实践活动。但各部门之间既缺乏跨部门的资源整合平台,又未建立统一的指挥调度机制,各自为政的现象十分突出。比如教务处在设计思想政治课程时,往往只关注课程的学时安排、考试通过率等量化指标,很少与学工部沟通学生在日常管理中呈现的思想动态;学工部在开展学生思想政治教育活动时,也难以借助教务处的课程资源进行深度拓展;团委组织的实践活动,更是常常与课堂教学内容脱节,无法形成理论与实践的有效衔接。

这种“多头管理”模式直接导致责任边界模糊,形成“人人有责却无人负责”的制度性空转。由于缺乏合力意识,各部门在开展思想政治教育工作时,往往只从自身利益出发,追求各自的考核指标,而忽视了思想政治教育的整体性与系统性。比如为了完成各自的量化任务,教务处不断增加思想政治课程的考试难度,学工部频繁开展各类具有形式主义的思想汇报活动,团委则一味追求实践活动的数量。这些工作之间缺乏有机衔接,不仅无法形成育人合力,反而让学生疲于应付,极易产生抵触情绪,严重制约了育人效能的释放。

2.3 教育技术赋能异化风险

数智媒体造成知识过量,引发学生选择焦虑与专注障碍,反而导致信息接收效能递减,形成“高供给-低内化”的价值传导现象。例如,在数智媒体环境中,学生日均接收信息量已达到传统教育场景的数十倍。这种爆炸式供给本应提升知识获取效率,却因学生认知的生理限制,导致选择焦虑与专注障碍。某高校调研显示,使用智能学习平台的学生中,多数学生存在“标签收藏癖”(持续标记未阅读内容),实际知识转化率远远低于标记量。其次,技术赋能异化造成情感联结弱化。技术工具大规模使用削弱师生共情场域,当教育停留于“机械说理”而缺乏“以情润心”,理论灌输难以触发情感共鸣和行为认同。例如,当前教育技术普遍采用的视频会议、VR 课堂等虚拟交互模式,剥离了传统教室中的微表情识别(如皱眉迟疑)、肢体语言反馈(如前倾倾听)等关键共情要素,直接导致情感共鸣能力下降,影响高校思想政治教育效果。

2.4 技术工具扩张助长实践转化链条断裂

技术工具的过度使用导致了高校学子能力培养

的偏离与替代。技术工具原本是辅助学生提升效率和能力的手段,但在实际应用中,却逐渐演变成替代学生思考和行动的工具。实践转化链条是连接知识认知与行动落地的核心纽带,其断裂的本质是知行转化的系统性脱节,成为制约高校人才培养质量的关键问题。实践转化链条的断裂首先体现在教育供给端的偏差——部分课程在技术工具的加持下陷入“理论至上”的误区,过度侧重知识的快速传递,却弱化了实践能力的系统性培养。技术工具的便利性让知识传授效率呈指数级提升:教师可通过在线课件、虚拟仿真平台在短时间内完成海量理论知识的输出,学生也能借助搜索引擎、AI 答疑工具轻松获取标准化答案。然而,这种“轻实践、重理论”的模式,正在制造“认知接受但行动滞后”的矛盾:学生虽能在考试中精准复述概念、背诵规范,却缺乏将理论转化为解决实际问题的能力及其场景训练。

综上,认知维度的偏差、制度层面的缺陷、技术异化现象与实践脱耦危机共同构成了一个复杂的阻滞系统。这些要素并非孤立存在,其深层根源在于教育现代化进程中,技术工具过度扩张与高校思想政治教育之间的结构性矛盾。因此,亟须构建系统性解决方案,针对不同层面的要素实施精准干预。

3 建立技术工具的约束机制,助力高校思想政治教育价值实现

技术工具对教育过程的过度渗透,导致的师生关系疏离、情感联结弱化以及价值传导链条断裂,这一困境的破解需构建“制度-技术-队伍-评价”四位一体的约束机制,通过系统性协同实现价值重塑。具体而言,需以制度整合为抓手强化组织协同效能,依托数字技术重构师生情感联结网络,在虚实融合的教育实践中完成高校思想政治教育的价值重建。同时,要突破认知灌输局限,实现从知识传递向价值内化的根本转向。

3.1 加强制度整合,发挥协同效应

制度保障的核心在于价值锚定、边界厘清、动态调适的三维联动,通过刚性制度防止技术异化,确保范式创新始终服务于立德树人根本目标^[6]。

校级统筹机制的规范化建设。优化组织架构,成立实体化运作的高校思想政治教育协同中心,采用“双主任”负责制:由分管学生工作的校党委副书记任政治主任,分管教学的副校长任业务主任,全面统筹协调全校思政教育工作。

创新协同机制。推行联席会议制度,每月固定召开教务处、学工部、马克思主义学院三方核心会

议;每季度扩展至二级学院党总支、专业教研室、学生社团等多元主体参与的扩大会议。会议实施“三单管理”(议题清单、责任清单、时限清单)模式,确保决策有效落实。

明晰权责边界。重点规定正面清单,聚焦课程思政建设、实践育人等核心协同事项,明确牵头部门的统筹权限与配合部门的执行义务。例如,课程思政建设由教务处主责,马克思主义学院协同开发教学案例库,形成“一事项一专班”的落实机制;形成负面清单,界定禁止性协作障碍(如信息壁垒、责任推诿等)。建立负面清单技术防控开发协同工作数字化平台,强制要求主责部门在立项时上传《权责确认书》,设立预警机制,防患于未然。

创新制度保障体系。完善双轨审批制度,构建“党委政治审查委员会+技术专家评审组”双轨并行审批体系,形成价值判断与技术评估的闭环管理^[7]。党委政治审查主要范围:政治立场(审查是否符合国家网络安全法及社会主义核心价值观)、意识形态安全(内容生成类技术需通过主流价值观对齐检测)、技术伦理(建立AI伦理审查委员会,对算法偏见、隐私侵犯等风险进行分级评估)。技术专家主要审查系统兼容性、数据安全等级、可持续运维成本、《技术准入负面清单》动态更新,明确禁止情感识别、算法推荐等12类高风险应用场景等。

3.2 重置技术工具范式,强化高校思政教育实效

推动高校思想政治课教学内容生产智能化。首先,在智能推送系统中部署“破茧算法模块”,采用动态加权策略确保异质内容占比在一定的比例,通过强制曝光打破信息同质化循环。开发“具身化”数字思政产品设计原则(反对界面交互,强调情境沉浸)。其次,重塑知识供给范式,构建智能知识图谱,破解碎片化困局的系统工程^[8]。以“农村包围城市”战略为例,系统可自动关联脱贫攻坚典型案例库中的区域发展数据,实现马克思主义理论知识的动态化演绎与可视化呈现,形成“历史经验-当代实践-理论创新”的闭环论证体系,将知识串联起来,有助于学生学习。再次,开发时政转化引擎,建立重大事件的12小时响应机制。部署基于DeepSeek的实时分析系统,在嫦娥成功落月、9·3大阅兵、“两会”等重大热点议题发生时,系统可自动生成包含背景知识、历史对比、理论解读的复合型知识包,并通过多终端实现分钟级触达。

重构场景,虚实融合深化价值传导。首先,落实实体场域主导原则。构建“实体主导-虚拟补充”的混合式教学框架,严格将虚拟技术应用比例

控制在30%以内,确保工厂实训、社会实践等实体教学场景占比不低于70%。通过思想政治理论指导下的亲身体验,实现知识从感官输入到认知内化的完整闭环。其次,构建时空穿越式体验。系统采用多光谱扫描技术,构建毫米级精度的三维历史场景模型。以延安杨家岭革命旧址数字化项目为例,学习者通过惯性动作捕捉套装参与“大生产运动”情景模拟,在时空交错的沉浸式体验中完成革命精神的具象化传承。最后,完善交互式思辨培养体系。基于生成对抗网络构建动态价值观训练沙盘,针对技术伦理等争议性话题,通过人机协同的认知调节机制,实时生成包含观点演变路径、共识达成曲线等维度的可视化认知图谱,实现辩证思维能力的量化培养。

3.3 加强教师队伍建设,构建新时代育人主体

伴随技术工具的广泛使用,教师已从单一的知识传授者转向“育人工程师”角色。数字化浪潮带来的不仅是智能黑板等技术工具革新,更要求教师平衡技术工具使用和思想政治教育价值引领^[9]。

强化教师数字素养。年长的教师多依赖传统板书、口授等成熟教学方式,对智能教学工具的操作复杂性存在预设性焦虑,以及现有考核制度未将数字技能纳入职称评定标准等原因,造成“部分年长教师排斥技术”的现象。同时,部分年轻教师仅掌握基础PPT制作、视频播放等简单功能,难以将虚拟仿真、大数据分析等技术与价值观引导有机结合,教学融合能力偏弱。因此,有必要建立思政教师技术培训体系。可以建立面向全体教师开展“数字教学工具工作坊”,按年龄分层设计课程:55岁以上教师侧重智能板书系统、课堂互动平台等轻量化工具实操培训;45岁以下教师增加虚拟仿真实验设计、教育大数据分析等进阶内容。建立“技术辅导员”制度,遴选年轻教师与老教师结对,通过“1对1”现场指导解决教师智能设备操作问题,每月开展“数字教学案例分享会”展示融合创新成果。

预防教学技术过度扩张,特别是资本的渗透。一方面,预防教师线上教学替代线下教学,弱化教师言传身教的价值引领。实践证明,当直播授课、AI助教等技术手段过度替代线下课堂时,教师肢体语言传递的共情力、课堂即时生成的思想火花等不可量化要素被压缩。例如,某高校思政课采用“慕课+线上答疑”模式后,学生价值观困惑的发现率大幅下降,印证了技术载体对教育情境的扁平化改造。另一方面,社会资本通过技术开始大规模对高校思政教育进行渗透,主要体现在知识取代价值观教育。

教育科技企业通过“智能批改系统”“知识点图谱”等产品,将思政教育简化为可量化的知识传递。某平台数据显示,其“红色题库”知识点使用率高,但涉及“理想信念”“家国情怀”等深层价值讨论的互动模块却参与度不足,揭示了技术工具对教育内涵的窄化。

3.4 聚焦知行合一,重构评估体系

在高校思想政治教育的宏大叙事中,评估体系始终是指挥棒与风向标。基于技术工具扩张形成的评估模式,过度倚重量化指标,催生了“重形式、轻实效”“重纸面、轻实践”的教育异化现象,使得学生陷入“刷时长、凑材料”的功利化泥沼,思想政治教育的价值引领功能被严重弱化。为此,必须以“知行合一”为核心导向,系统性重构高校思想政治教育评估体系,让评估真正成为推动高校思想政治教育入脑入心、落地生根的核心动力。

构建三维评价体系,替代单纯数据竞赛。构建“认知内化-情感认同-行为践行”三维评价体系,全面衡量思政教育对学生思想观念与行为模式的深层影响。认知内化维度:通过课程论文答辩、思想政治主题辩论赛、马克思主义理论知识应用案例分析等方式,考查学生对高校思想政治理论的理解深度与迁移应用能力,重点关注学生能否用理论知识解释现实问题、分析社会现象。情感认同维度:依托思想政治主题演讲比赛、红色文化研学心得分享、榜样人物访谈等活动,评估学生对主流价值观的情感共鸣程度,通过学生的语言表达、情感流露、价值判断等细节,捕捉其思想观念的内化轨迹。行为践行维度:建立“思想政治行为观察档案”,通过日常行为表现记录、班级民主评议、社会实践跟踪等方式,考查学生在人际交往、集体活动、社会服务等场景中的道德行为,重点关注学生能否将思想政治理论转化为自觉行动。

构建复合型价值体系,打造中国特色思想政治教育话语。传统思想政治教育评估往往陷入某种误区,将思想政治教育视为实现某种功利目标的手段,忽视了其价值引领的核心功能。为此,我们必须推动评估体系范式转换,构建以马克思主义理论及其中国化的理论成果为指导、以中华优秀传统文化为根基、人类文明成果为补充的复合型价值体系。首先,确立以马克思主义理论及其中国化的理论成果为指导的价值体系^[10]。将辩证唯物主义和历史唯物主义的立场、观点、方法贯穿于思想政治教育的全过程,帮助学生正确分析纷繁复杂的当今世界,引导学生树立正确的世界观、人生观、价值观,培养学生

的辩证思维能力与历史分析能力。其次,夯实以中华优秀传统文化为根基的价值体系。深入挖掘中华优秀传统文化中的思想政治教育资源,将“仁爱、诚信、正义、和合”等传统美德与现代思想政治教育相结合,打造具有中国特色、中国风格、中国气派的思想政治教育话语体系,增强思想政治教育的文化底蕴与民族认同。再次,拓展以人类文明成果为补充的价值体系。积极吸收借鉴世界各国的优秀文明成果,引导学生以开放包容的心态看待不同文化,培养学生的全球视野与国际理解能力,推动思想政治教育的国际化发展。

总之,技术工具在高校思想政治教育中的应用是一把双刃剑,既为高校思想政治教育带来新的机遇,也引发了价值失衡等问题。我们必须深刻认识技术工具过度扩张对高校思想政治教育价值实现的阻滞效应,通过构建“制度-技术-队伍-评价”四位一体的约束机制,平衡技术工具与思想政治教育价值实现的关系。在技术革新的背景下,坚守高校思想政治教育的初心和使命,充分发挥技术工具的赋能作用,推动高校思想政治教育高质量发展,培养更多担当民族复兴大任的时代新人。

参考文献:

- [1]程诚,李玉.突破技术工具论:重审数字文明时代的历史学范式[J].南京社会科学,2025(8):11-23.
- [2]王易,朱惠羽.数字技术赋能高校思想政治理论课教学论析[J].马克思主义理论学科研究,2024,10(3):29-38.
- [3]栾淳钰.思想政治教育“可视化”:价值、困境与举措[J].马克思主义与现实,2023(6):141-148.
- [4]吴满意.困境与纾解:高校思想政治教育数字化进程中的风险防范研究[J].思想教育研究,2024(6):70-75.
- [5]约翰·哈蒂.可见的学习:对800多项关于学业成就的元分析的综合报告[M].彭正梅,邓莉,高原,译.北京:高等教育出版社,2009:16-21.
- [6]李娜,史柳.新时代高校思政课理想信念教育价值实现的三重逻辑[J].现代教育科学,2025(4):69-74.
- [7]金梭,陈首珠.技术治理推动高校思想政治教育效能提升的现实困境与应对路径[J].哈尔滨学院学报,2024,45(4):128-132.
- [8]孙其昂.思想政治教育基础理论研究反思及再出发[J].思想政治教育研究,2024,40(3):69-75.
- [9]王易.数字技术赋能思政课教学[J].马克思主义研究,2024(3):29-38.
- [10]教育部.新时代学校思想政治理论课改革创新实施方案[Z].2025.

责任编辑:朱家席

基于工业机器人的摩擦力测量实验平台设计及其教学应用

罗少轩*, 乔爱民

(蚌埠学院 电子与电气工程学院, 安徽 蚌埠 233030)

摘要:针对机器人工程专业工业机器人系统集成、机器人传感技术、电气控制与PLC应用等核心课程现有实验设备功能单一、跨课程知识点难以贯通的实验教学痛点,课程组将摩擦力自动测量作为实验应用场景,整合六自由度工业机器人、三轴力传感器、传送带、信号变送器、PLC等分立设备,搭配自研的上位机测控软件,研发了一套基于工业机器人的摩擦力测量实验平台。该实验平台可精准测定不同材料间的滑动摩擦力并计算摩擦系数,支持学生自主设计实验流程、规划机器人运动轨迹和更换摩擦副材料,能充分发挥学生的主观能动性。依托该平台,课程组打通课程壁垒,将工业机器人、传感器技术、电气控制的相关知识点有效融合,并设计了6个创新实验实训项目。教学实践表明,该实验平台能有效激发学生求知欲与学习兴趣,弥补传统实验设备的功能短板,显著提升实验教学质量,并为同类型机器人实验设备的升级改造与创新研发提供实践依据。

关键词:工业机器人;摩擦力测量;三轴力传感器;实验设备研发;教学研究

中图分类号:TP242.2;G642

文献标识码:A

文章编号:2095-297X(2026)05-0118-06

Friction Force Measurement Experimental Platform Based on Industrial Robot and Teaching Application

LUO Shaoxuan*, QIAO Aimin

(School of Electronics and Electrical Engineering, Bengbu University, Bengbu, 233030, Anhui)

Abstract: Aiming at the experimental teaching deficiencies of core courses for robotics engineering, including Industrial robot system integration, robot sensing technology and electrical control and PLC application, where existing experimental devices possessed only single functions and failed to connect knowledge points among different courses, automatic friction measurement was adopted as the experimental scenario by the teaching team. A friction measurement platform based on industrial robots was developed by integrating discrete hardware including a six-degree-of-freedom robot, a triaxial force sensor, a conveyor belt, a signal transmitter and a PLC, and was matched with self-designed upper computer measurement and control software. The sliding friction of different materials was accurately detected and the friction coefficients were calculated by the platform. Experimental schemes were designed, robot trajectories were planned, and friction pairs were replaced independently by students to give full play to their subjective initiative. Curriculum barriers were broken, multidisciplinary knowledge involving robots, sensing and electrical control was integrated, and six innovative experimental projects were constructed by the teaching team. Teaching practice verified that students' learning enthusiasm was inspired, the functional defects of traditional equipment were remedied, the quality of experimental teaching was improved, and practical references were provided for the upgrading and innovative development of similar robot experimental equipment.

收稿日期:2026-06-17 * 通讯联系人

基金项目:安徽省高等学校质量工程项目(2025cjk074);蚌埠学院校级质量工程项目(2025jyxm7)。

作者简介:罗少轩(1980—),男,湖南岳阳人,教授,硕士。E-mail:lsx@bbc.edu.cn

Key words: industrial robot; friction force measurement; three-axis force sensor; experimental equipment design; teaching research

工业机器人系统集成、机器人传感技术、电气控制与 PLC 应用等课程是机器人工程专业核心课程^[1],要求学生重点掌握机器人轨迹控制、传感检测、PLC 程序设计等核心知识与软硬件结合的实操能力^[2-3]。目前高校相关实验教学存在显著缺陷,现有设备功能单一、各课程实验体系相互独立,仅满足单课程验证性教学需求,难以实现跨课程知识融合^[4-5],导致学生知识掌握碎片化,无法建立系统化工程思维,综合实践能力培养受限^[6]。此外,传统实验平台流程与参数固化,学生自主设计、调试与创新空间不足,难以满足新工科复合型工程人才的培养要求^[7-8]。

针对上述教学问题,契合新工科实践教学改革需求,本文依托课程组在工业机器人领域的科研积累,自主研发了工业机器人摩擦力测量跨课程实验平台。该平台整合六自由度工业机器人、高精度三轴力传感器、信号变送器、PLC 控制系统及自研上位机软件,支持学生自主设计实验方案、规划机器人运动轨迹、更换摩擦副材料,可完成机械运动控制、传感器信号采集、PLC 电气控制、上位机数据处理的全流程实操训练,可同时支撑三门专业核心课程的创新实验教学,实现多课程知识点有机融合。该平台紧密贴合专业应用型人才培养目标,可开发工程性、综合性较强的创新实验项目,有效提升设备利用率、降低教学设备投入成本^[9],打破了传统单课程验证式教学模式,实现设备、实验项目与多门核心课程的深度融合^[10],有效解决了知识割裂、工程实践性薄弱等教学痛点^[11-12]。同时依托科研成果反哺课堂,丰富了专业实践教学内容与手段,为新工科背景下机器人工程专业跨课程综合实验教学改革与创新实践人才培养提供了有效的实践思路与应用范例。

1 实验平台的组成及工作原理

本次研制的实验平台主要由六轴工业机器人、夹爪、传送带、三轴力传感器、信号变送器、通信线缆、PC 机和上位机软件组成,如图 1(a)所示。

摩擦力测量采用方形且底部带凸台的高精度三轴力传感器,如图 1(b)所示,传感器尺寸为 40 mm × 40 mm × 17 mm,其底部带有 18 mm × 18 mm 的方形凸台,凸台高度 1 mm, Z 轴量程 20 N, X 轴和 Y 轴量程为 10 N,传感器综合测量误差低于 0.5% F. S。试样(铜箔)固定在三轴力传感器的凸台上,对偶件

(铜箔)固定在传送带上。机器人通过夹爪抓取三轴力传感器,调整夹爪姿态,使得传感器 X 轴方向与传送带运动方向保持一致、传感器 Z 轴方向与传送带平面垂直,传感器凸台与传送带上的对偶件轻微接触。三轴力传感器的输出信号通过 3 根信号线与信号变送器的输入端连接,信号变送器通过 RS485-USB 通信线与 PC 机通信,PC 机中的上位机软件对信号变送器上传的数据进行分析和处理,计算并显示正向压力、摩擦力和摩擦系数,上位机软件主界面如图 2 所示。

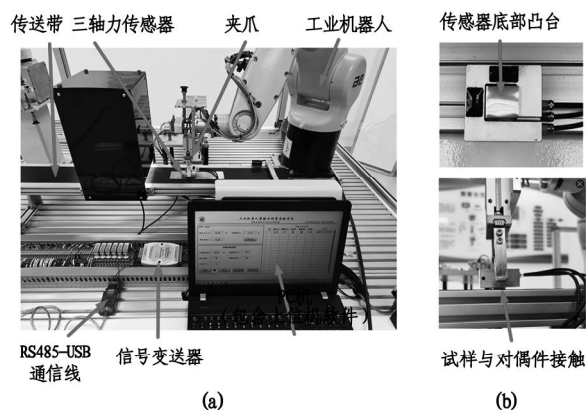


图 1 实验平台实物图



图 2 上位机软件主界面

利用实验系统测量摩擦力并计算摩擦系数 μ , 需要以下 6 个步骤:

第 1 步:安装待测试样。将试样安装在三轴力传感器底部的凸台上;将对偶件固定在传送带上。

第 2 步:控制六轴工业机器人抓取三轴力传感器。学生可以通过示教器手动控制机器人和夹爪去抓取三轴力传感器,也可以通过示教器编程让机器人自动抓取传感器。

第 3 步:传感器零点校准。三轴力传感器被夹爪抓取后, Y 轴、Z 轴输出会分别受夹持力和自身重力影响,而多轴力传感器本身也存在一定的交感误差,因此这时需要对传感器各轴零点进行零点校准操作,降低测量误差。在上位机软件中点击“零点

校准”按钮,软件自动根据当前各轴受力情况,清除皮重值。

第 4 步:调整传感器方向及位置。对机器人各关节姿态进行调整,使得夹爪上的三轴力传感器 Z 轴与传送带垂直; X 轴与传送带运动方向平行;并且传感器底部凸台上的试样与传送带上的对偶件轻微接触。

第 5 步:测量摩擦力。控制机器人沿 Z 轴方向缓慢下压,直至传感器 Z 轴压力达到设定阈值后停止,同时控制传送带沿 X 轴方向匀速运动,测量并记录传感器 X 轴输出值,即为滑动摩擦力。另一种测量方式是,传感器 Z 轴压力达到设定阈值后,传送带不动,通过示教器编程,让机器人自动沿 X 方向匀速平移一段距离,同样可以测量 X 方向的滑动摩擦力。学生在此步骤中,可以充分发挥主观能动性,设计测量方式和运动轨迹。

第 6 步:计算滑动摩擦系数 μ 。在上位机软件中点击“计算系数 μ ”按钮,软件自动计算摩擦系数,并将测量结果保存在界面右侧的实验记录表格中。

实验系统的主要参数指标如表 1 所示。

表 1 实验系统的主要参数指标

技术指标	技术参数
工业机器人载荷	8 kg
三轴力传感器量程	$Z:20\text{ N}; X,Y:10\text{ N}$
传感器测量精度	0.5% F.S
摩擦副材质	可替换
系统综合测量误差	$\leq 10\%$
传送带速度	0 ~ 4 m/min 可调
供电电压	220 VAC

2 摩擦力测量平台的设计方法

2.1 系统结构设计

本次设计的摩擦力测量系统,包括硬件设计和软件设计两部分。其中,硬件部分主要包括六轴工业机器人、夹爪、三轴力传感器、信号变送器、传送带、PLC、RS485-USB 通信线和 PC 机;软件部分主要包括控制机器人运行轨迹的示教程序、控制传送带的 PLC 程序和 Labview 编写上位机软件。

系统结构图如图 3 所示。三轴力传感器的输出信号是 mV 级电压信号,经过变送器处理后,变成数字信号,通过 RS485-USB 通信线,发送给 PC 机分析处理。

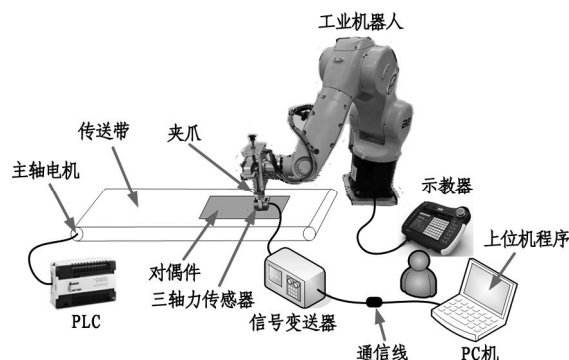


图 3 系统结构图

2.2 实验平台工作原理

本研究实验平台,可用于精准测定材料接触界面的滑动摩擦力与摩擦系数,模拟工业零部件、耐磨构件的摩擦特性测试,可实现恒定压力、恒定滑动速度条件下的精准摩擦性能检测。在实验过程中,也可以向学生说明,各类机械传动、接触运动构件的摩擦系数与摩擦力稳定性,直接影响设备运行功耗、使用寿命与运动精度,若摩擦参数异常波动,极易造成构件磨损加剧、传动卡顿、失效故障等问题。

实验开展时,设备对位与压力控制直接决定测量精度。依托配天 AIR8 工业机器人高精度重复定位、恒力按压的作业优势,可精准控制三轴力传感器与试样的下压位置、接触角度与正压力大小,有效提升实验数据的稳定性与准确性。同时可根据不同材质、不同尺寸的待测试样,合理调节机器人按压行程与按压压力,适配多种摩擦测试工况,规避压力过载损伤传感器、压力不足导致接触不充分的实验误差问题。试样硬度较高、接触面积较大,可匹配适中压力保证贴合紧密;试样材质偏软、接触面积较小,则可适当降低按压压力,防止试样挤压变形影响测试结果。

实验系统中,三轴力传感器底部凸台固定待测试样,传送带表面固定对偶件,实验时由 PLC 程序控制传送带保持匀速稳定运行,模拟均匀滑动摩擦工况。由配天 AIR8 工业机器人抓取三轴力传感器,精准带动传感器及底部试样垂直按压在传送带的对偶件表面,通过三轴力传感器实时采集接触过程中的正压力(Z 轴受力)与滑动摩擦力(X 轴受力)数据,进而完成摩擦力的定量测量并计算摩擦系数 μ 。若传感器实时采集的沿 Z 轴法向正压力为 F_N ,沿 X 轴滑动方向实测摩擦力为 F_f ,则可通过二者的比值计算出待测试样与对偶件接触面的摩擦系数 μ 。

2.3 三轴力传感器信号采集与解耦

本次设计的实验系统搭载自研高精度三轴力传感器,量程参数配置为Z轴20 N、X、Y轴均为10 N。传感器各轴原始输出为微弱mV级电压信号,信号易受现场电气噪声干扰,因此配套多通道信号变送器完成统一处理:依次实现信号滤波放大、模数转换与三轴力耦合解耦,处理后的有效力值数据经由RS485总线上传至上位机完成存储与运算。

其中解耦算法采用标定矩阵求逆法实现三轴力耦合误差消除。三轴力传感器弹性体受力时,X、Y、Z三个方向载荷会相互产生交叉干扰,各轴输出电压 U_x 、 U_y 、 U_z 与各轴受力 F_x 、 F_y 、 F_z 的关系如公式(1)所示,其中 C_{xy} 为Y向作用力对应的影响系数,表征Y向作用力 F_y 对X方向输出电压 U_x 的耦合影响^[13]。

$$\begin{bmatrix} U_x \\ U_y \\ U_z \end{bmatrix} = \begin{bmatrix} C_{xx} & C_{xy} & C_{xz} \\ C_{yx} & C_{yy} & C_{yz} \\ C_{zx} & C_{zy} & C_{zz} \end{bmatrix} \begin{bmatrix} F_x \\ F_y \\ F_z \end{bmatrix} \quad (1)$$

单轴受力下其余通道仍会输出电压信号,该耦合关系可通过公式(2)描述:

$$U = C \cdot F \quad (2)$$

式中 U 为各通道电压输出向量, F 为三轴实际受力向量, C 为系统耦合标定矩阵。实验前通过标准砝码对传感器完成多组单向加载标定,采集多组受力-电压样本,拟合得到完整3阶耦合标定矩阵 C ,实际测量时对矩阵 C 求逆得到解耦矩阵,如公式(3)所示。

$$F = C^{-1}U \quad (3)$$

由变送器对采集的三路电压信号进行矩阵运算分离出X、Y、Z轴真实独立受力,剔除三轴力学交叉耦合造成的测量偏差,能大幅提升法向压力、摩擦力的检测精度与数据稳定性,解耦后三轴力传感器的综合测量误差 $\leq 0.5\% F.S.$ 。

2.4 软件设计

实验平台软件程序分为三个主要部分:(1)驱动传送带的PLC程序;(2)控制工业机器人运动的示教程序;(3)PC端的上位机数据采集处理程序。其中(1)和(2)两项程序需要学生在课堂实验环节独立构思、全程自主设计编写调试;第(3)项上位机程序由课程组团队教师预先开发完善,实验课上直接交付学生操作、调取测量数据使用。

上位机软件基于LabVIEW开发,软件主界面见图2,主要实现五项功能:(1)配置传感器量程、摩擦副配对材质等系统参数;(2)建立与信号变送器的数据通信链路,实时采集三轴力传感器的各轴受力数据;(3)待工业机器人夹持传感器就位后,执行传感器零点校准,消除空载皮重带来的测量偏差;(4)依据采集的受力数值运算求解摩擦系数;(5)保存实验数据记录。系统运行流程如图4所示。

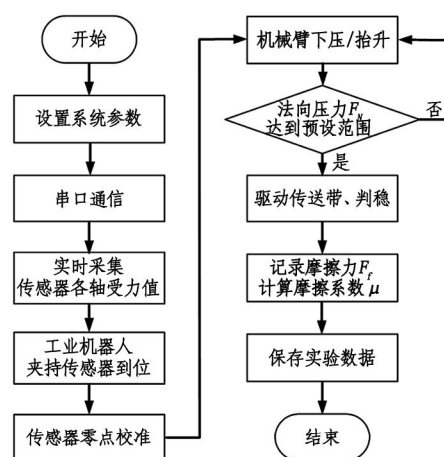


图4 系统运行流程

2.5 摩擦系数测量实验数据

实验选用厚度0.5 mm铜箔作为摩擦副材料,设置Z轴法向压力阈值区间7.50~8.50 N;待法向载荷达到设定值后,驱动传送带匀速拖动摩擦副运动,实时采集X方向摩擦力并计算摩擦系数 μ ,每组连续重复测量10次,实验数据如表2所示。

表2 摩擦系数测量实验数据

测量次数	1	2	3	4	5	6	7	8	9	10
法向压力 F_N/N	7.89	8.32	8.16	8.37	7.92	8.05	8.19	7.73	8.25	8.11
摩擦力 F_f/N	3.23	4.08	3.84	4.27	3.01	3.38	3.69	3.32	4.04	3.57
摩擦系数 μ	0.41	0.49	0.47	0.51	0.38	0.42	0.45	0.43	0.49	0.44

经测算,实验测量相对重复性误差稳定控制在10%以内,平台测试稳定性优异。相较于弹簧测力

计手动拖拽的传统测试方法,人工操作难以保持匀速,施力方向易偏移,加之弹簧测力计精度低、信号

采集滞后,测量薄铜箔摩擦副时数据离散严重,仅可做定性演示。本平台借助六自由度机器人和传送带实现匀速闭环运动,搭配高精度三轴力传感器同步采集多维力信号,精准维持恒定法向载荷,消除人为操作带来的随机误差。平台可完成摩擦系数精准定量测试,支撑多变量综合创新实验,精度完全满足机器人工程专业跨课程实训教学要求,弥补了传统手动设备精度差、数据不稳定的短板。

3 教学应用及教学效果

本次研制的摩擦力测量实验平台,既契合教育部鼓励本科院校自研教学实验器材的导向要求,也

为机器人工程、自动化专业的实践教学革新开辟了可行路径。相较于市场上的商业化实验仪器,自行研制的实验设备更贴合课堂教学内容,有效解决传统标准化实验台功能单一、模拟实际工况效果差的问题;同时便于教师把机器人工作站搭建、三轴力传感信号解耦等科研成果转化为教学资源,推动科研、教研深度融合,提升实践教学质量。结合“工业机器人系统集成”“机器人传感技术”“电气控制与 PLC 应用”三门主干课程的教学大纲要求,课题组团队依托此实验平台设计了 6 个配套的创新实验项目,如表 3 所示。

表 3 基于摩擦力测量实验平台开发的 6 个实验项目

实验名称	实验类型	课程名称
传送带启停与调速控制实验	基础型实验	电气控制与 PLC 应用
PLC、工业机器人、PC 机多设备通信实验	基础型实验	电气控制与 PLC 应用
多通道数字信号变送器设计	综合性实验	机器人传感技术
三轴力传感器信号采集与解耦实验	综合性实验	机器人传感技术
摩擦力测量工作站设计与仿真	课程设计	工业机器人系统集成
摩擦力测量系统上位机软件设计	课程设计	工业机器人系统集成

将自研机器人摩擦力检测平台投入对应课程实验教学后,收获了显著的实践教学成效:首先,学生不但了解了滑动摩擦力、摩擦系数的自动化检测方法,还掌握了机器人工作站的设计、仿真和调试流程,充分调动学习主动性与探索欲,为今后其他类型的自动化检测场景和机器人工作站设计打下良好基础,切实锻炼了学生的工程实践能力。其次,开发的全新实验实训项目扩充了原有实验课教学体系,直观展示了机器人控制、传感器信号采集、PLC 逻辑控制等知识点在工业检测场景里的落地形式,夯实了学生工程应用思维,提升了实验教学质量。再次,检测平台由教师团队依托自身相关科研成果自主搭建,拓展性强,可依据不同课程的各种实践任务灵活调整硬件结构、控制流程与上位机程序;平台亦可充当课程设计、实习实训、毕业设计的实践载体,为机器人工程、自动化专业的实践教学改革提供可行方案。最后,设备研制阶段吸纳多名实操能力突出的学生全程参与整机装配布线、软件程序编写以及整机联调、运行测试等工作。在平台搭建与优化的过程中,学生得以把课堂理论知识转化为动手实操能力,科研创新能力得到充分锻炼。不少学生围绕本实验平台的技术内容开展深度研究,不仅成功申报并获得多项实用新型专利授权,还以此为支撑立项

大学生创新创业训练项目,为工科实践教学、产学研协同育人提供了优质示范载体。

4 结论

本次研发的基于工业机器人的摩擦力测量实验平台可精准测量各种材料之间的滑动摩擦力与摩擦系数,并针对机器人工程专业的“工业机器人系统集成”“机器人传感技术”“电气控制与 PLC 应用”三门主干课程设计了 6 项创新实验项目,充实了原有实践教学资源,也为本校及同类本科院校自研机电一体化教学装备提供参考范例。整套实验平台高度还原工业检测现场工况,能够充分调动学生学习积极性,有效锻炼学生工程实践与创新思考能力,提升了机器人工程专业实践课程的教学质量。

参考文献:

- [1] 钟秋波,邵千钧,谢行恒,等. 地方应用型高校工科项目制教学模式构建——基于宁波工程学院机器人学院的探索[J]. 高等工程教育研究, 2025(3): 61-66, 78.
- [2] 李康康,王智慧,王成军. AI 赋能机器人工程专业课程教学改革研究[J]. 中国现代教育装备, 2026(9): 56-59.
- [3] 何巍,高鸿隽,孙星,等. 新工科视域下“产-教-研-

- 学”四位一体的实验教学改革与育人模式探索[J/OL]. 力学与实践, 1-7. <https://link.cnki.net/urlid/11.2064.03.20260203.1217.004>.
- [4] 左文娟, 宁萌, 王琨, 等. 基于数字孪生的机器人工程专业虚拟仿真实践平台设计[J]. 实验室研究与探索, 2024, 43(3): 6-11.
- [5] 王嘉, 冯佳亮, 刘海龙, 等. 工业机器人关节精密减速器加速寿命实验装置研制及教学应用[J]. 实验室研究与探索, 2024, 43(12): 43-46, 78.
- [6] 杨波, 王晨辰. 基于科研驱动的角度测量实验教学研究[J]. 黄山学院学报, 2025, 27(5): 96-99.
- [7] 赵鹏飞, 赵聪, 温洋, 等. 工业机器人柔性执行器实验装置设计[J]. 实验室研究与探索, 2025, 44(3): 61-66.
- [8] 杨亮, 谢金玲, 沈慧, 等. 基于视觉反馈的机械臂教学实验装置开发[J]. 实验室研究与探索, 2025, 44(5): 54-58, 63.
- [9] 王海欣, 黄海宏, 汪海宁, 等. 新型“电力电子技术”实验教学装置[J]. 电气电子教学学报, 2025, 47(4): 210-213.
- [10] 杨照坤, 房立存, 李昀书, 等. 自动化专业实验仪器开发的探索与实践[J]. 中国现代教育装备, 2021(19): 56-57, 63.
- [11] 郭丰, 林琳, 何俊, 等. 教育元宇宙视域下的新工科实验教学改革初探[J]. 机械设计, 2026, 43(2): 199-206.
- [12] 李贺, 刘金库, 刘新丽, 等. 人工智能赋能虚拟仿真实验课程升级的路径探析[J]. 中国电化教育, 2025(12): 20-24.
- [13] Luo Shaoxuan, Qiao Aimin, Wang Yanchun, et al. A three-axis force sensor for stress monitoring of foundation buildings[C]//2025 5th International Conference on Sensors and Information Technology (ICSI), Nanjing, China, 2025: 43-47.

责任编辑: 汤庆国

(上接第 95 页)

- [52] Li Bin, Liang Shun, Luo Jingpeng, et al. A physics-informed neural network for dual-parameter decoupling and crosstalk suppression based on dynamic physical weight assignment and dual-branch decoupling[C]//Eighth Global Intelligent Industry Conference (GHIC 2025). SPIE, 2025, 13685: 243-251.
- [53] Yang Yiguang, Geng Dahe, Chen Hao, et al. Low SNR multimirror Fabry-Perot pressure sensor optic spectrum signal analysis and demodulation via SVM-KNN regressors[J]. Applied Optics, 2024, 63(6): A16-A23.
- [54] Liu Zheng, Yao Qiong, Xia Ji, et al. Five-step phase-shift-based multiwavelength averaging for extrinsic Fabry-Perot interferometric sensors[J]. Photonics, 2023, 10(10): 1076.
- [55] Yang Yang, Zhao Xinyu, Zheng Yongqiu, et al. Optimized phase-generated carrier demodulation algorithm for membrane-free Fabry-Pérot acoustic sensor with high sensitivity[J]. Micromachines, 2025, 16(2): 196.
- [56] Li Qun, Shao Jian, Wu Peng, et al. Research on high-precision PGC demodulation method for Fabry-Perot sensors based on shifted sampling pre-calibration[J]. Sensors, 2025, 25(19): 5990.
- [57] Ren Sufen, Chen Shengchao, Xu Haoyang, et al. Unsuper-
- vised representation learning-based spectrum reconstruction for demodulation of Fabry-Pérot interferometer sensor[J]. IEEE Sensors Journal, 2023, 23(12): 13810-13816.
- [58] Xu Haoyang, Chen Shengchao, Ren Sufen, et al. High-resolution spectrum reconstruction of cascaded FBG-FPI sensor in resource-constrained environments[J]. IEEE Sensors Journal, 2024, 24(7): 11878-11885.
- [59] Xu Yufu, Qi Hongchao, Zhao Xinyu, et al. High-speed spectrum demodulation of fiber-optic Fabry-Perot sensor based on scanning laser[J]. Optics and Lasers in Engineering, 2024(c), 178: 108192.
- [60] Li Chenxi, Qi Hongchao, Han Xiao, et al. Ultrahigh-speed phase demodulation of a Fabry-Perot sensor based on fiber array parallel spectral detection[J]. Optics Letters, 2024, 49(3): 714-717.
- [61] Bai Jiandong, Cao Shengjie, Chai Haoyan, et al. Ultrahigh-speed demodulation for Fabry-Pérot sensors based on spatio-temporal features of sparse spectra[J]. Journal of Lightwave Technology, 2025, 43(14): 6957-6968.
- [62] Chen Yuanfeng, Wang Tianhao, Yang Pengcheng, et al. Optical vernier effect-based demodulation system for fiber-optic sensors: an AWG and a neural network-based perspective[J]. Applied Optics, 2025, 64(21): 6077-6084.

责任编辑: 汤庆国

AI 赋能医学生创新创业能力提升的教学改革及实践探索

张亨国

(安徽医科大学口腔医学院,安徽医科大学附属口腔医院,安徽省口腔疾病研究重点实验室,安徽 合肥 230032)

摘要:针对医科类高校创新创业教育中专业融合不足、实践场景不深和评价支持不精准等问题,构建“全过程融入、多模块协同、AI 精准赋能”的教学改革模式,将生成式 AI 融入课前导学、课中实践、课后迭代和成果孵化全过程,支持个性化学习、项目设计、方案优化和多元评价。结果显示,AI 赋能教学有助于提升学生学习主动性、创新思维、项目设计能力、团队协作能力和路演表达能力,促进医学专业教育与创新创业教育深度融合。此研究为医科类高校 AI 辅助创新创业教育改革和双创人才培养提供了实践参考。

关键词:人工智能;大学生创新创业;教学改革;医学生;实践教学

中图分类号:G642.0

文献标识码:A

文章编号:2095-297X(2026)05-0124-05

Teaching Reform and Practical Exploration of AI-Empowered Enhancement of Innovation and Entrepreneurship Abilities among Medical Students

ZHANG Hengguo

(College & Hospital of Stomatology, Anhui Medical University, Key Lab. of Oral Diseases Research of Anhui Province, Hefei, 230032, Anhui)

Abstract: To address problems in innovation and entrepreneurship education in medical universities, such as insufficient integration with professional education, limited depth of practical scenarios, and imprecise evaluation support, a teaching reform model was constructed in this study characterized by “whole-process integration, multi-module collaboration, and AI-enabled precision support.” Generative AI was integrated throughout the entire process of pre-class guidance, in-class practice, post-class iteration, and achievement incubation to support personalized learning, project design, proposal optimization, and diversified evaluation. The results showed that AI-enabled teaching helped improve students’ learning initiative, innovative thinking, project design ability, teamwork ability, and roadshow presentation skills, while promoting the deep integration of medical professional education and innovation and entrepreneurship education. It provided practical reference for AI-assisted reform of innovation and entrepreneurship education and the cultivation of innovation-oriented entrepreneurial talents in medical universities.

Key words: artificial intelligence, AI; college students’ innovation and entrepreneurship; teaching reform; medical student; practical teaching

大学生创新创业能力培育工作,是高校践行立德树人、完善高质量人才培养体系、助力国家战略布局的关键组成部分。为了推动高校完善创新创业教学体系,相继出台的多项政策明确提出把创新创业

能力提升贯穿人才培养各个阶段,实现专业授课与创新创业培育深度交融^[1-2]。立足于“新医科”建设大潮,医学院校除了完成基础医学人才培育工作之外,还需助力健康中国战略实施、落实医学技术创

收稿日期:2025-05-26

基金项目:安徽省优秀青年教师培育项目(YQYB2024013)。

作者简介:张亨国(1988—),男,安徽合肥人,副教授,博士。E-mail:zhanghengguo@ahmu.edu.cn

新研究、加快医药科研成果转化。由此可见,培育医学生创新思维、创业认知、多学科融合能力,是高水平医学类院校建设过程中人才培养质量提升与学科建设体系完善的核心工作^[3]。

随着科技革新与产业升级进程持续提速,人工智能、大数据、云平台、智能医疗等前沿技术持续重塑行业发展模式与就业市场格局,也对当代学生创新思维、创业认知、跨学科统筹能力、复杂难题处理能力设立了全新标准。现阶段高校双创教学不能只停留在创新创业基础理论讲解、各类创业赛事指导层面,应当转型为依托真实行业难题、搭建实景实践平台、侧重综合实操素养提升的一体化育人体系。

当前,许多高校已开设创新创业基础课程,并推出大学生创新创业竞赛和校企协同实践平台,在激发学生创新意识、拓展实践机会方面发挥了积极作用。相关研究认为,创新创业教育应从单一知识传授转向能力培养和价值塑造,并进一步融入高校人才培养体系^[4-5]。然而,在实际教学过程中,仍普遍存在课程内容泛化、案例更新滞后、学生参与深度不足、项目指导个性化不够、评价方式单一等问题。尤其在医科类高校中,学生专业学习任务较重,创新创业教育与医学专业知识、临床需求、健康产业发展和科技成果转化结合不够紧密,导致学生虽然具备一定专业知识基础,却缺乏专业知识转化成创新项目、创业方案或社会服务产品的能力。

人工智能(artificial intelligence, AI)技术的发展为创新创业教学改革提供了新的工具和路径。近年来,生成式 AI 在个性化学习、智能答疑、学习行为分析、自动化评价和教学资源生成等方面的应用不断深化,为高校教学模式升级提供了重要支撑^[6-9]。落实到医学院校的教学场景中, AI 技术的作用早已超越普通课堂辅助软件范畴。立足临床一线真实存在的各类医疗难题,该技术可协助师生梳理临床服务缺口、拆解各类医学研究数据、构思智能医疗相关产品、搭建完整健康服务场景,同时完成科研成果落地转化相关实操训练,给医学生双创素养培育提供高度贴合医学专业特色的配套支撑。依托生成式 AI,学生可完成市场行情摸排、行业矛盾拆解、同类产品优劣比较、商业运营框架搭建、答辩稿件润色打磨等一系列工作;生成式 AI 使用的全过程数据记录与分析,便于授课教师精准定位学生在创新思考、课题规划、小组协同配合等环节存在的各类薄弱点。

因此,本研究借鉴 AI 赋能课程教学改革的相关研究思路^[10],以大学生创新创业能力提升为目标,以医科类高校学生为实证对象,立足新医科建设和

医学教育高质量发展需求,聚焦医学生创新创业能力培养中的专业融合不足、实践场景不深和评价支持不精准等问题,围绕生成式 AI 赋能教学资源建设、学习过程支持、项目实践训练和综合评价优化等方面展开研究,构建“全过程融入、多模块协同、AI 精准赋能”的教学改革模式。研究设计上,将通用创新创业教育理论与医科类高校人才培养特点相结合,以医学真实问题和健康产业应用场景为牵引,探索 AI 技术支持下的课程资源重构、项目训练优化、过程指导强化和多元评价改进路径,旨在形成适用于医科类高校、具有一定推广价值的创新创业教育改革实践范式。

1 医学生创新创业能力培养的教学现状与核心挑战

当前,高校创新创业教育已形成课程教学、项目训练、学科竞赛、创业孵化和校企合作相结合的基本体系,并依托大学生创新创业训练计划、“互联网+”大学生创新创业大赛、挑战杯竞赛、大学科技园和校企合作平台等,为学生提供实践训练与成果转化机会。近年来,线上资源、翻转课堂、项目式教学和混合式教学逐步应用于课程改革,在提升学生自主学习和课堂参与方面发挥了一定作用^[11-12]。

在医科类高校中,创新创业教育应服务新医科建设、健康中国战略和医学科技成果转化需求。但现有教学仍存在专业融合不深、真实场景不足、指导不精准和评价方式单一等问题。部分课程偏重通用创业理论、案例分析和商业计划书训练,与临床诊疗、公共卫生、口腔健康、康复护理、医学检验、智慧医疗等医学真实场景结合不够。已有研究指出,创新创业教育应强化与专业教育、产业需求和真实项目的结合,避免停留在形式化和活动化层面^[4-5]。医学生虽具备专业知识基础,也能从临床实践、科研训练和社会服务中发现问题,但在用户调研、产品设计、商业模式、知识产权、医学伦理、注册合规和成果转化等方面训练不足,导致项目难以从“医学问题”转化为“创新方案”或“应用产品”。

此外,医学生创新创业能力还涉及问题意识、创新思维、团队协作、技术转化、伦理合规和健康产业理解等多维能力。随着 AI、大数据、智能制造和智慧医疗等技术快速发展,医学生亟须提升数字素养、技术应用能力和跨学科创新能力^[13]。现阶段双创课程考核手段,大多依靠期末作业完成质量、创业计划书撰写情况、最终路演得分作为评判依据,很少关注学生全程学习状态、小组内实际贡献、医学项目落

地价值等。学生课堂投入程度直接决定最终学习成效。由此可见,长期收集、分析学生全周期学习数据,对教学优化具有很高的参考意义^[14]。

2 AI 赋能医学生双创综合素养培育的内在逻辑

AI 辅助教学模式的普及,给医学类院校创新与创业相关课程开辟了全新实施载体,同时丰富线上学习资源库,革新传统课程考核评判的整套实施逻辑。AI 引入课堂的核心意义,绝不只是单纯压缩授课耗时、提升课堂运转效率,其核心优势体现在依托海量学习数据做量化研判、实时同步学习反馈、搭建虚拟模拟实训场景等多重路径,贯穿医学生创新思维塑造、创业实操能力养成的全周期。当下数字化教学改革已是各大高校发展的必然方向,智能工具深度介入课堂,会从教师授课逻辑、学生自主学习路径、期末考核评判规则三个维度产生颠覆性改变,相关观点已有诸多研究佐证^[6-9]。

目前医学生的专业发展侧重点、参与科研项目积累、临床实操阅历、数字软件操作熟练度存在差异,依托生成式 AI 评估学生阶段性测试得分、课后作业完成质量、创新课题推进情况等多元信息,就能生成专属个人学习画像,据此推送适配其现阶段能力的学习材料,搭建因人而异的成长提升路径,真正做到分层式针对性培养^[6,15]。生成式 AI 还可全程协助医学生开展创新课题全链条准备工作:发掘行业现存问题、检索相关文献资料、解读医疗卫生领域相关扶持政策、调研行业市场供需现状、对标市面上同类医疗产品优劣、搭建整套创新落地规划,从源头优化医学创新课题选题方向,完善整套设计方案框架^[9,16]。实操过程中,学生可借助智能系统深挖不同医疗细分赛道的现实需求,涵盖口腔常态化健康养护、老年群体康养配套服务、慢性疾病长效管控方案、医学影像智能诊断设备、康复配套器械研发、全民医疗卫生知识普及、乡镇基层医疗配套升级、数字化智慧医院搭建等方向;同步梳理目标服务人群特征、提炼行业现存核心短板、规划产品核心使用功能、预判项目落地潜在隐患、整理科研成果商业化落地整套流程。借助生成式 AI,授课教师能够分析课题完成进度、小组协作状态、方案修改痕迹、答辩综合表现等多维度原始数据,形成完整记录学生双创能力提升的成长台账。

3 AI 赋能医学生创新创业能力提升的教学改革模式构建

基于医学生创新创业能力培养目标和 AI 技术应用特点,本研究构建“全过程融入、多模块协同、AI 精准赋能”的教学改革模式(见图 1)。该模式以医学生创新创业能力提升为主线,将教学过程整合为课前启发、课中实践、课后迭代和成果孵化四个阶段,将教学内容整合为创新意识、创业知识、项目实践、技术应用、团队协作和成果转化六个模块,实现 AI 技术与医学创新创业教育的深度融合。该设计与当前高校混合式教学改革和 AI 赋能教学模式升级趋势具有一致性^[6-12]。

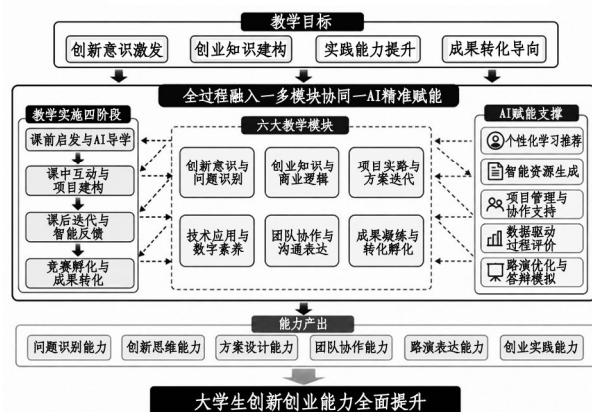


图 1 AI 赋能大学生创新创业能力提升的教学模式

本课程构建了“课前—课中—课后—孵化”一体化医学双创育人模式。课前结合学生专业背景,依托生成式 AI 梳理医学创新案例、政策文件和行业报告,引导学生围绕临床诊疗、口腔健康、护理康复、公共卫生、健康管理和智慧医疗等场景开展问题识别与需求分析。课中教师基于学情数据开展精准教学,通过案例研讨、角色扮演、项目路演、小组答辩等形式,指导学生完成可行性论证、用户画像、技术路线、产品设计和商业模式构建;生成式 AI 承担资料汇总、竞品分析、商业画布生成、风险梳理和文稿润色等辅助工作,教师重点把关项目落地性、医学严谨性、伦理合规性和临床应用潜力。课后学生依据教师点评与生成式 AI 持续优化项目方案,对创业计划书、申报书和路演 PPT 进行系统排查,提升项目逻辑、数据支撑、表达质量和风险预判水平。课后进一步衔接大创项目、学科竞赛、科研实验室、合作医院和医疗企业资源,形成项目孵化培育闭环。课程最终目标是提升学生创新创业能力提升,包括问题识别能力、创新思维能力、方案设计能力、团队协作能力、路演表达能力、创业实践能力,推动医学双创教育走向以临床真实问题为导向、项目实践为核心、智能技术精准赋能的新型育人模式。

4 教学实施效果评估

为检验 AI 赋能创新创业教学改革实施效果,可采用对照研究、问卷调查、项目成果分析和过程性数据评价相结合的方法开展评估。已有课程教学改革研究常采用学习成绩、学习行为、满意度和能力发展等维度评价教学效果^[10,14],本文在此基础上构建创新创业课程教学改革评价框架。

4.1 方案设计

本次调研选取安徽医科大学 22 级口腔医学专业选修《创新教育与创业基础》课程的 64 名本科生作为观测样本。研究采用随机均分分组方式,32 人划入试验组,全程运用生成式 AI 辅助开展创新创业教学;剩余 32 人为对照组,沿用院校传统课堂模式完成课程学习。两组学生入学年级一致,前期专业课程积累水平无显著差距,课程授课师资完全相同。

对试验组,课程全周期融入生成式 AI,覆盖课前资料引导、课堂项目设计、课后方案修改、路演孵化梳理等全部环节。对照组授课形式仅保留理论知识点讲解、行业案例剖析、小组课后作业、期末统一路演答辩四类传统形式。

4.2 效果评价

教学实践结果显示,生成式 AI 有效提升了课程学习成效与学生双创综合素养。试验组课程平均分为 86.2 分,高于对照组的 77.8 分,整体提升 10.8%。相较于传统教学模式,试验组学生在课题材料完整度、方案设计逻辑、路演表现等方面表现更优。学生能够围绕临床诊疗、口腔健康管理、老年慢性病随访、康复设备和智慧护理等真实医学场景开

展课题设计,形成较为完整的项目方案、申报材料和路演表达。结果表明,AI 辅助教学能够增强学生主动学习意识,促进医学专业知识与创新创业实践深度融合,提升课程育人质量和项目孵化潜力。

4.3 问卷调研

课程教学环节结束后,面向试验组全体学生发放匿名调查问卷,调研内容覆盖教学资源供给质量、生成式 AI 实操感受、创新项目实训效果、小组协同作业体验、路演实践效果、课程整体育人收获六大方向。问卷细分考察指标包含案例资料、项目模板的丰富程度;个性化学习推荐、智能答疑、文本诊断;选题、调研、方案设计、项目迭代;任务分工、进度管理、沟通协同;路演稿优化、答辩模拟、PPT 结构调整;对创新创业能力提升的总体评价等细分内容。依托学生主观体验调研数据,可客观评判智能教学模式适配程度,同步梳理后续教学体系优化调整的具体方向。

AI 赋能大学生创新创业能力提升的教学改革满意度调查结果如表 1 所示。由问卷统计数据显示,绝大多数受试学生认可生成式 AI 具备教学辅助价值,能够拓宽创新思考角度、压缩文稿撰写耗时、优化答辩成品整体质量。但一部分学生也客观指出,智能程序自动生成文本普遍存在内容高度趋同、文字表述宽泛空洞、医疗细分专业细节缺失等突出短板。该反馈内容提示授课教师,必须完善 AI 使用管理规范,正向引导学生把智能工具定位为辅助载体,杜绝单纯依靠生成文本替代自主思考、线下实地调研的惰性学习行为。

表 1 AI 赋能大学生创新创业能力提升的教学改革满意度调查结果

评估维度	调查内容	调查结果/人			主要反馈
		不满意	一般	满意	
教学资源	案例资料、项目模板的丰富程度	1	5	26	学生认为课程资源更加丰富,案例更贴近实际
AI 学习支持	个性化学习推荐、智能答疑、文本诊断	1	6	25	有助于拓展思路,但部分内容仍需教师把关
项目实践	选题、调研、方案设计、项目迭代	2	6	24	项目推进更有方向,方案修改效率提高
团队协作	任务分工、进度管理、沟通协同	1	8	23	有助于提升协作效率,但团队主动性仍需加强
路演表达	路演稿优化、答辩模拟、PPT 结构调整	1	5	26	能够提升表达逻辑和答辩准备质量
综合效果	对创新创业能力提升的总体评价	1	4	27	多数学生认可 AI 赋能教学模式

4.4 教改现存短板与长期优化实施路径

AI 与医学双创实训课程相互融合落地,现阶段仍处在起步探索阶段,完整配套育人体系暂未成型,教学落地推进过程中,逐步显现多项亟待调整优化

的现实问题。第一,部分学生形成重度工具依赖习惯,直接照搬智能生成内容填充课题申报材料,最终成品虽格式规范完整,但缺少线下实地调研原始数据、实景实践过程作为支撑,整体内容空洞,落地可

行性严重不足。第二,大量一线授课教师智能教学实操能力存在明显短板,医学专属提示词撰写、智能评价数据深度解读等核心工作能力欠缺。第三,未来双创教育智能体的搭建,离不开信息化平台、细分专业教学数据库、配套实训资源三重支撑,对医学院校硬件配套条件、常态化教学管理机制均提出更高建设标准。

针对本次教学改革过程中暴露的各类短板,后续双创教学体系优化可同步推进制度规范、师资培育、考核机制升级三项工作:第一,院校应落地标准化智能工具使用管理细则,清晰划定学生完成课后作业、申报各级创新项目过程中智能工具的使用范围与约束边界;第二,定期组织校内教师开展智能教学能力专项培训,全方位提升教师将人工智能融入课程设计、一对一课题指导的综合授课能力;第三,迭代升级全过程多元化综合考核体系,把线下调研原始资料、课题多轮修改留存记录、团队成员分工贡献台账、智能工具规范使用全部纳入综合考核指标,持续推动医学院校双创教学朝着标准化、精细化、实效化方向升级。

5 结论

人工智能技术快速迭代,为高校培育学生创新创业能力提供全新教学支撑,同时拓宽传统课堂改革空间。本文针对国内高校双创教学普遍存在的实践场景不足、学生实操能力偏弱、个性化指导缺失、考核方式单一且主观等短板,搭建“全流程融入、多模块协同、AI 精准赋能”的课堂改革方案。整套改革框架把生成式人工智能嵌入课前预习、课堂实操、课后方案优化、项目孵化的完整教学链条,搭建以学生能力成长为核心、创新项目实训为载体、智能数据反馈为支撑的闭环双创教学体系。

对照教学实践结果表明,AI 赋能的教学改革能提高主动学习的热情,提升创新思维、项目规划、团队协作、路演答辩等多项能力,方便教师开展分层精准教学、全过程动态考核。对于医科类院校而言,融合智能工具的双创教学模式能够打通专业授课与创业培养之间的隔阂,引导学生围绕医疗卫生行业真实难题开展创新实训,提升学生适配健康产业发展、满足社会医疗服务需求的综合实践能力。

未来,高校应进一步完善 AI 赋能创新创业教育的课程体系、平台建设、师资培训和评价机制,推动人工智能从辅助性工具转变为教学改革的重要支撑

力量,不断提升大学生创新创业教育质量,为培养具有创新精神、实践能力和社会责任感的高素质复合型人才提供有力保障。

参考文献:

- [1] 国务院办公厅. 关于深化高等学校创新创业教育改革的实施意见[Z]. 国办发[2015]36号,2015-05-04.
- [2] 教育部办公厅. 关于做好深化创新创业教育改革示范高校 2019 年度建设工作的通知[Z]. 教高厅函[2019]22号,2019-03-28.
- [3] 陈培,陈维贤,黄勇,等. 新医科时代背景下医学教育高质量发展路径研究[J]. 继续医学教育,2025,39(7): 116-120.
- [4] 王占仁. “广谱式”创新创业教育的体系架构与理论价值[J]. 教育研究,2015,36(5):56-63.
- [5] 李家华,卢旭东. 把创新创业教育融入高校人才培养体系[J]. 中国高等教育,2010(12):9-11.
- [6] 翟雪松,许家奇,童兆平,等. 人工智能赋能高校韧性教学生态的路径研究[J]. 中国远程教育,2023,43(1): 49-58.
- [7] 喻国明,李钊,滕文强. AI+教育:人工智能时代的教学模式升维与转型[J]. 宁夏社会科学,2024(2):191-198.
- [8] 郑旭东,周冰. 人工智能融入“三个课堂”的总体目标、理论基础与具体路径[J]. 电化教育研究,2025,46(11):85-91.
- [9] 石磊,方海光. 生成式人工智能重塑开放教育教学场景:样态、价值及实践路径[J]. 成人教育,2025,45(9): 47-54.
- [10] 王翠,周宇. 基于 AI 赋能的供应链管理课程教学改革与实践探索[J]. 蚌埠学院学报,2026,15(2):122-128.
- [11] 王维国,徐健,盖印. 经管类专业课程体系数智化升级与教学方法创新[J]. 中国大学教学,2022(3):31-36.
- [12] 孙慧,李玉蓉,谢萌. 未来教育视域下高校新型混合式教学研究[J]. 黑龙江高教研究,2024,42(1):17-22.
- [13] 赵芮,单莹莹. 教育数字化背景下高校学生创新创业教育研究[J]. 创新教育研究,2024,12(9):510-516.
- [14] 梁存良. 在线学习行为投入对学习绩效的影响——基于国内外 30 篇实证研究文献的元分析[J]. 河南师范大学学报(哲学社会科学版),2024,51(5):151-156.
- [15] 赵晶莹,吴坚. 美国 K-12 教育领域人工智能赋能个性化学习的实践路径[J]. 比较教育学报,2025(5):153-166.
- [16] 余辉,夏文蕾,程钰,等. 基于生成式人工智能的课堂教学实验设计与实践[J]. 实验室研究与探索,2025,44(11):121-125,221.

责任编辑:汤庆国

蚌埠学院学报

(综合类双月刊)

编辑委员会

主 任:陈国龙

副 主 任:张亚新

委 员:(以姓氏笔画为序)

王艳春	韦文联	吕 部	吕效华	朱兰保	朱家席
邬旭东	刘 娟	许 晖	孙兰萍	孙西超	李大胜
李文瑛	李如林	李宗群	吴长法	何 蕾	张 斌
张永芳	张亚新	张茂林	陈 荣	陈发来	陈国龙
陈恩红	周仁战	郑晓奋	胡 飞	赵生学	赵东勇
袁 飞	郭有强	黄迎辉	崔执树	葛金龙	魏天路

主 编:张亚新

副 主 编:郭有强 邬旭东 朱家席

目 次

食品工程·生命科学

- 玉米深加工副产物混菌固态发酵饲料工艺优化 满云,蒋光乐,蒋舒晴,等(1)
- 复方中药酵素液的制备工艺优化及体外降糖降脂作用 徐世晶,熊柯,李正阳,等(8)
- 复合低聚糖固体饮料制备及其抗氧化活性研究 郝梦嫒,张雨晨,杜传来,等(15)

硅基新材料·化学工程

- 玻璃纤维粉目数对二氧化硅气凝胶性质影响的研究 高燕,王悦悦,钱紫薇,等(22)
- 沸石咪唑酯骨架材料作为缓蚀剂在海洋环境铜防护中的应用..... 曹娇娇,柯顺安,葛雨童(29)

机械工程

- 渗碳淬火对螺旋锥齿轮齿面误差影响分析与修正 顾井飞,陈兴强(34)
- 基于 Kriging 模型的龙门加工中心主轴箱多目标优化 陈宇,吴紫鹏,彭兆春,等(40)

电子与电气工程·人工智能及计算机技术

- 基于带状宽度约束策略的 Smith-Waterman 算法优化 甘秋云,黄玮,李华亮,等(46)
- 基于 STM32 和 OV2640 的嵌入式图像采集系统设计 侯静云,闻国才,杨晟,等(57)
- 基于模态差异感知与双向注意力融合的 RGBT 目标跟踪方法 李鹏程(63)
- 基于改进 YOLOv12 的中药饮片识别标注研究 邵悦,朱锦龙,刘儒香(70)

水利科学·土木建筑工程

- 双孔平行盾构掘进地表及管线沉降对比分析 王鹏,王芸,李永康,等(78)

基础科学研究

- 光纤法布里-珀罗传感器解调算法发展 赵小明,吴凯鹏,周程,等(84)
- 基于 MacTempas 软件模拟 DO_{24} 结构相的选区电子衍射 程宏伟,孟维利,周家峰,等(96)
- 租赁设备两阶段维护策略优化 黄显兵,姚梦月,董克,等(100)
- 基于决策树分析的车间生产工位竞争研究 杨露,石林(106)

教育与教学研究

- 技术工具扩张对高校思想政治教育价值实现的冲击与应对策略 梁祥凤(113)
- 基于工业机器人的摩擦力测量实验平台设计及其教学应用 罗少轩,乔爱民(118)
- AI 赋能医学生创新创业能力提升的教学改革及实践探索 张亨国(124)

Journal of Bengbu University

Vol. 15, No. 5, Sum. No. 89, 2026

Main Contents

- Optimization of Mixed-Strain Solid-State Fermentation Process for Feed Using Corn Deep-Processing By-Products MAN Yun, JIANG Guangle, JIANG Shuqing, et al. (1)
- Optimization of Preparation Process for Compound Traditional Chinese Medicine Yeast Extract and Study on Its In Vitro Hypoglycemic and Lipid-Lowering Effects
..... XU Shijing, XIONG Ke, LI Zhengyang, et al. (8)
- Effects of Glass Fiber Powder Mesh Size on the Properties of Silica Aerogel
..... GAO Yan, WANG Yueyue, QIAN Ziwei, et al. (22)
- Research on Metal-Organic Framework Materials as Corrosion Inhibitors in the Protection of Copper in Marine Environment CAO Jiaojiao, KE Shunan, GE Yutong (29)
- Influence Analysis and Modification of Surface Error of Spiral Bevel Gear for Carburizing and Quenching GU Jingfei, CHEN Xingqiang (34)
- Optimization of the Smith-Waterman Algorithm Based on a Banded Width Constraint Strategy
..... GAN Qiuyun, HUANG Wei, LI Hualiang, et al. (46)
- Design of Embedded Image Acquisition System Based on STM32 and OV2640
..... HOU Jingyun, WEN Guocai, YANG Sheng, et al. (57)
- Modality-Aware Perception and Bidirectional Attention Fusion for RGBT Object Tracking
..... LI Pengcheng (63)
- Comparative Analysis of Ground and Pipeline Settlement During Double-Tube Parallel Shield Tunneling WANG Peng, WANG Yun, LI Yongkang, et al. (78)
- Development of Demodulation Algorithms for Fiber Fabry-Perot Sensors
..... ZHAO Xiaoming, WU Kaipeng, ZHOU Cheng, et al. (84)
- Selected Area Electron Diffraction Simulation of DO_{24} Phases via MacTempas
..... CHENG Hongwei, MENG Weili, ZHOU Jiafeng, et al. (96)
- A Two-Stage Maintenance Strategy Optimization for Rental Equipment
..... HUANG Xianbing, YAO Mengyue, DONG Ke, et al. (100)
- Research on Workshop Production Workstation Competition Based on Decision Tree Analysis
..... YANG Lu, SHI Lin (106)
- Impact of Technological Tool Expansion on the Realization of Educational Value in University Ideological and Political Education and Coping Strategies LIANG Xiangfeng (113)
- Friction Force Measurement Experimental Platform Based on Industrial Robot and Teaching Application LUO Shaoxuan, QIAO Aimin (118)
- Teaching Reform and Practical Exploration of AI-Empowered Enhancement of Innovation and Entrepreneurship Abilities among Medical Students ZHANG Hengguo (124)