

人机物 CPSS 智能融合的平行创作架构与关键技术研究

郭超¹, 鲁越^{1,2}, 王晓^{1,3}, 易达¹, 王琥⁴, 王飞跃^{1,5}

(1. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室, 北京 100190;

2. 中国科学院大学人工智能学院, 北京 100049;

3. 青岛智能产业技术研究院平行智能创新中心, 山东 青岛 266114;

4. 普渡大学, 美国 西拉法叶 IN 47907-2050;

5. 中国国家博物馆科学与艺术研究所, 北京 100006)

摘要: 随着人工智能探索领域的不断拓展, 艺术创作成为人工智能发展和应用的重要研究热点。基于平行系统理论与 ACP 方法构建风格多样、内容逼真、笔触灵活和描述精准的平行艺术创作元宇宙, 为提升人工智能的创造能力提供了一种可行的实现途径, 并提供了应用案例。通过 AI 算法创作、人类筛选和评估、机器人执行, 构建了人机物 CPSS 智能融合的平行创作架构, 阐述了基于计算实验的绘画风格迁移、内容组合、笔触生成和图像描述等关键技术, 并对所构建的平行创作系统进行了实验验证。平行创作系统融合了人、AI 创作算法、机器人的优势, 提升了人工智能艺术创作系统在虚拟和物理空间中的创作水平, 促进了人机物协同艺术创作的发展。

关键词: 平行系统; ACP; 艺术创作; 元宇宙; 人机物智能

中图分类号: TP391

文献标志码: A

doi: 10.11959/j.issn.2096-6652.202246

Architecture and key techniques of parallel creation through the fusion of human-cyber-physical intelligence in CPSS

GUO Chao¹, LU Yue^{1,2}, WANG Xiao^{1,3}, YI Da¹, WANG Xiao⁴, WANG Fei-Yue^{1,5}

1. The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

2. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China

3. Parallel Intelligence Innovation Center, Qingdao Academy of Intelligent Industries, Qingdao 266114, China

4. Purdue University, West Lafayette IN 47907-2050, USA

5. Institute of Science and Art, National Museum of China, Beijing 100006, China

Abstract: With the expansion of the fields covered by AI, artistic creation will become the next hot spot for AI research and applications. Building a metaverse with diverse styles, realistic contents, flexible strokes and accurate descriptions based on parallel system theory and ACP approach will provide a feasible way to improve AI creation capability. The intelligence of human, AI, and robots were fused to develop a parallel creation architecture through the creation by AI, the evaluation by humans, and the execution by robots. The parallel creation with the key methods of style transfer, content combination, stroke generation and image captioning in computational experiments were explained. The parallel creation system was validated through the painting experiments. The parallel creation system will improve the creation capability of artificial intelligence in cyberspace and physical space, and promote the human-cyber-physical collaborative creation through the fusion of them.

Key words: parallel system, ACP, artistic creation, metaverse, human-cyber-physical intelligence

收稿日期: 2022-08-01; 修回日期: 2022-08-08

通信作者: 王晓, x.wang@ia.ac.cn

基金项目: 国家自然科学基金资助项目 (No.61533019)

Foundation Item: The National Natural Science Foundation of China (No.61533019)

0 引言

1955 年达特茅斯提案 (Dartmouth Proposal) 提出了人工智能 (artificial intelligence, AI) 的定义, 其中将创造力作为人工智能研究的一个重要目标^[1]。几十年来, 国内外众多研究者围绕这一目标, 致力于让人工智能向着增强创造力的方向不断发展, 取得了显著的进步^[2-6]。其中, AlphaGo^[3]针对围棋这一博弈竞技场景, 在围棋攻防策略和博弈方法等方面表现出了强大的创造力^[4]。随着在人工智能创造力方面的研究不断深入, 人工智能的创造任务从常规的模式识别任务扩展到了艺术创作任务^[5]。近年来, OpenAI 探索了人工智能生成艺术图像的创造力, 其设计的 DALL-E 2 图像生成 AI 创作算法^[6]能够根据语言描述生成逼真的艺术图像。同时, 国内外学者利用人工神经网络和深度学习围绕艺术作品的检测^[7]、分类^[8]、评估^[9]和生成^[10]等方面展开研究, 得到了广泛的关注^[11]。上述人工智能在艺术创造中的初步探索, 使人工智能与艺术的交叉领域成为当下人工智能技术与系统研究的一个热点。

现有的 AI 创作算法大多是从人类的艺术创作过程中得到启发的。人类的艺术创作过程往往是先构思, 在头脑中进行模拟, 然后在不断地尝试中完成实际创作。与此类似, AI 创作算法通过组合已有的绘画元素、变换绘画风格等方法进行绘画实验和探索。然而, 当前的 AI 创作算法具有显著的局限性: 一方面缺乏对因果关系和时空约束的探索, 另一方面仅可应对数字绘画图像, 而无法完成实际的绘画过程。为了解决上述问题, 提升人工智能的创造能力, 需要融合人 (人)、AI 创作算法 (机)、机器人 (物) 各自的优势。在 AI 创作算法的基础上融合人和机器人的角色, 人提供创作示范样例, 并对绘画进行评估, 机器人完成物理空间中的创作。由此, 人机物 (human-cyber-physical) 的协同将进一步释放人工智能进行艺术创作的巨大潜力。

早在 1994 年, 学术界就出现了借助虚拟计算的力量拓展人类计算力量的想法^[12]。中国科学院自动化研究所王飞跃研究员于 2004 年提出平行系统理论^[13-16], 用于解决复杂系统建模与控制问题。该理论利用人工系统、计算实验和平行执行来构造闭环系统, 拓展人类思维计算的空间, 成为解决复杂社会经济及复杂系统问题的有效计算研究理论和方法体系。平行系统理论^[17]为实现人机物的协同创作提供了理论

支撑。可借助数字孪生^[18]和人工智能等技术方法, 构造平行的人工智能艺术创作系统, 以 AI 算法创作、人类筛选和评估、机器人执行的方式, 构建一种人机物智能融合平行创作的范式。

本文以人工系统、计算实验、平行执行 (artificial systems, computational experiments, parallel execution, ACP) 平行智能理论为支撑, 构造了一种融合人机物各自优势的平行创作系统, 通过对人类绘画过程的建模建立人工系统; 基于计算实验进行风格迁移、内容组合、笔触生成和图像描述等对人类艺术创作过程的模仿和推演; 根据探索结果引导绘画机器人的绘画过程; 通过人工绘画系统与实际绘画机器人的平行执行, 实现 AI 学习并创作、人类示教和评估、机器人执行绘画过程的人机物智能融合的平行创作。在现有 AI 创作算法的基础上, 平行创作系统构建起包含人、AI 创作算法、机器人的艺术创作闭环, 实现了人机物的协同创作, 为进一步提升机器的创造能力奠定了基础。

本文着重描述了如何通过计算实验来融合人机物智能的优势, 构建风格多样、内容逼真、笔触灵活和描述精准的平行艺术创作元宇宙。其中, AI 创作算法完成计算实验中的风格迁移、内容组合、笔触生成和图像描述等模仿推演过程, 是人机物智能融合的平行创作中的重要实现技术。第 1 节介绍人工智能艺术创作、元宇宙与平行系统等相关工作; 第 2 节介绍基于平行系统的人机物 CPSS 智能融合的平行创作架构; 第 3 节介绍平行创作系统的设计与实现, 包括风格迁移、内容组合、笔触生成和图像描述等关键技术; 第 4 节进行总结和展望。

1 相关工作

现有的人工智能创作方法大多关注采用 AI 创作算法进行数字绘画图像生成的探索。本文提出一种人机物 CPSS 智能融合的平行创作架构, 并据此构建平行创作系统, 在 AI 创作算法的基础上, 融合人和机器人在艺术创作中的优势, 完成对人类绘画过程的模仿和基于机器人的绘画的物理实现。其中, AI 创作算法是人机物协同创作的核心环节, 本节介绍相关的主流算法及其代表性工作, 并阐述基于平行系统的人机物智能融合创作及其与数字孪生、元宇宙之间的关系。

1.1 人工智能艺术创作

本节围绕绘画风格、绘画内容、绘画笔触和绘

画描述这4个绘画创作中的关键研究内容,分别对风格迁移、内容组合、笔触生成和图像描述方法进行分析 and 总结。其中,风格迁移和内容组合聚焦于如何产生具有特定绘画风格和绘画内容的数字图像,笔触生成方法通过笔触序列实现绘画的序列化生成和绘制,图像描述则建立绘画与文字描述之间的多模态关联,并实现对绘画的赏析。

(1) 绘画的风格迁移

风格迁移任务旨在分别提取两张图像的内容和风格,并将它们组合起来,生成新的图像^[19]。早期的风格迁移方法大致分为笔触风格渲染^[20]、基于区域渲染^[21]、基于实例类比^[22]和基于图像滤波^[23]。上述方法一般只适用于某种特定的风格,在绘画风格上的可扩展性较差,同时在提取图像的高层次结构特征方面存在不足。随着深度学习的发展,基于卷积神经网络的神经风格迁移算法有效解决了上述问题,提高了图像风格迁移的质量,同时实现了单模型单风格^[24]、单模型多风格^[25]与单模型任意风格^[26]等多种类型的风格迁移模型。风格迁移实现了对图像风格信息的提取与组合,使人工智能艺术创作能够在保持绘画内容不变的情况下对多种绘画风格进行探索。

(2) 绘画的内容组合

内容组合是人类在绘画中常用的重要创造形式之一^[27],可用于生成全新的绘画内容。根据内容组合的实现方式,现有绘画内容组合方法主要包括两种,即基于生成模型^[28-29]和基于绘画元素组合^[30-31]。前者基于绘画的文字描述等给定条件凭空生成绘画内容,高度依赖模型的生成能力,探索空间较大;后者将现有内容拆分、组合,从而形成新的绘画内容,探索空间相对较小,且需要提前采集绘画,并对其进行分割。与基于绘画元素组合的方法相比,基于生成模型的方法难以生成笔触序列以完成绘画过程。

(3) 绘画的笔触生成

笔触生成方法主要包括基于笔触参数优化的方法^[32]、启发式方法^[33]和基于强化学习^[34]的方法。目前基于笔触参数优化和基于强化学习的方法可以实现较好的图像渲染效果且较为灵活,但这两种方法通常采用复杂的笔触模型,不适用于机器人绘画过程。启发式方法则依赖人工设计的绘画规则,笔触形状往往采用简单的线条,能够直接应用于机器人绘画过程。因此,目前用于数字绘画图像渲染的笔触生成多采用基于笔触参数优化的方

法,而涉及实际绘画过程的机器人绘画等场景则多采用启发式的方法。

(4) 绘画的图像描述

图像描述的目的在于分析图像数据,并以自然语言的形式表达其语义信息^[35]。绘画图像描述在人工智能艺术创作的语义理解等方面具有重要作用。例如,将大规模的绘画图像转换为结构化的自然语言,便于人工智能在艺术创作中根据自然语言来筛选合适语义的绘画图像,并将其作为绘画学习的样本。根据建模方式,现有图像描述方法可以分为基于文字搜索和基于文字生成两种^[36]。基于文字搜索的方法在文字数据库中选择与输入图像最匹配的描述句子,基于文字生成的方法根据图像特征直接生成描述文字。基于文字搜索的方法能够得到较为流畅的描述文字,适用于数据集规模较小的场景,但在生成文字的多样性方面存在不足。基于文字生成的方法能够输出表达多样的描述文字,有利于人工智能艺术创作中的绘画语义理解^[37]。

1.2 基于平行系统的平行创作

艺术创作活动体现了人类独特的复杂性和创造性,对艺术创作的复杂过程进行模拟与研究需要有效的方法,平行系统理论与 ACP 方法提供了一种可行方案。平行系统理论已在智能控制^[38-40]、智能交通^[41-43]和自动驾驶^[44-47]等场景取得了广泛应用。图1展示了基于 ACP 的平行系统架构体系,其由实际系统和与之对应的人工系统组成,可有效解决有人参与的控制和学习问题,并提高复杂系统的管理与机器学习的效率。其中,计算实验是体现平行系统的智能和功能的核心,其将复杂系统理论与计算机仿真技术相结合,构建与实际系统对应的人工系统,并通过模拟实际系统的运行规则、参数、推演过程进行大量的反复实验,以对实际系统进行分析与解释,为人机物智能融合的平行创作提供理论支撑。

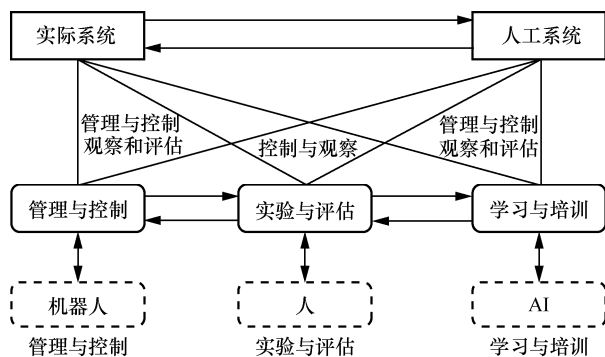


图1 基于 ACP 的平行系统架构体系

通过 ACP 方法实现人的实验与评估、AI 的学习与培训、机器人的管理与控制，从而构建起人机器 CPSS 智能融合的平行创作架构和系统。在该架构中，人提供示范数据并进行评估；AI 创作算法则通过计算实验在人工系统中对创作方案进行探索，直接决定了创作的结果，是实现融合智能平行创作的关键环节；机器人完成对人工系统中探索的创作方案的实际部署和控制。具体地，计算实验由绘画风格迁移、内容组合、笔触生成和图像描述等 AI 创作算法实现，用于模拟绘画创作过程，对绘画方案进行充分探索，最终形成优化的绘画方案并提炼绘画知识，从而引导实际绘画创作系统完成艺术创作，提升创作效果。

1.3 数字孪生、元宇宙、平行系统与人工智能艺术创作

近年来，随着数字孪生^[17, 48]和元宇宙^[49-50]概念的提出，平行系统的含义和特点得到了更清晰的体现。数字孪生通过对物理系统构建一对一的数字系统，实现对物理系统的实时、全面的信息呈现。元宇宙通过构建数字替身实现参与者在虚拟数字空间中的互动和演化。数字孪生和元宇宙都针对实际物理系统构建了虚拟的数字系统，但其在系统复杂性和社会属性方面的建模仍不充分，无法实现对实际物理系统的引导和迭代优化，不能很好地发挥虚

拟数字系统在人工智能艺术创作中的作用。平行系统有效融合了数字孪生与元宇宙的优势，通过数字孪生对物理世界进行数字化，同时实现元宇宙对人类的想象力和创造力的表达^[51]。本质上，平行系统是数字孪生与元宇宙的更高维度的融合形态，能够根据应用场景的需求变化，实现数字孪生特性和元宇宙特性的切换或组合。平行系统同时支持对物理世界的虚拟扩展和对人类想象空间的虚拟扩展，可为元宇宙及人工智能艺术创作元宇宙提供基础科学理论和关键技术体系支撑。

2 人机器智能融合的平行创作

在艺术创作过程中，人类艺术家首先在大脑中进行构思，再进行真实的艺术创作，并实时地根据大脑中的构思修正真实创作效果。上述人类艺术家的艺术创作过程可等价于平行系统中的人工系统与实际系统之间的相互作用过程。因此，本文根据平行系统理论设计了人机器 CPSS 智能融合的艺术创作系统架构。首先，通过人工系统对创作过程进行模拟，再通过计算实验进行创作构思，对比和筛选大量创作思路，最后通过虚拟创作和实际创作的平行执行与交互，实现人、AI 创作算法和机器人互相配合的平行艺术创作。图 2 展示了人机器 CPSS 智能融合的平行创作架构。所构建的平行创作系统基于 ACP 方法与

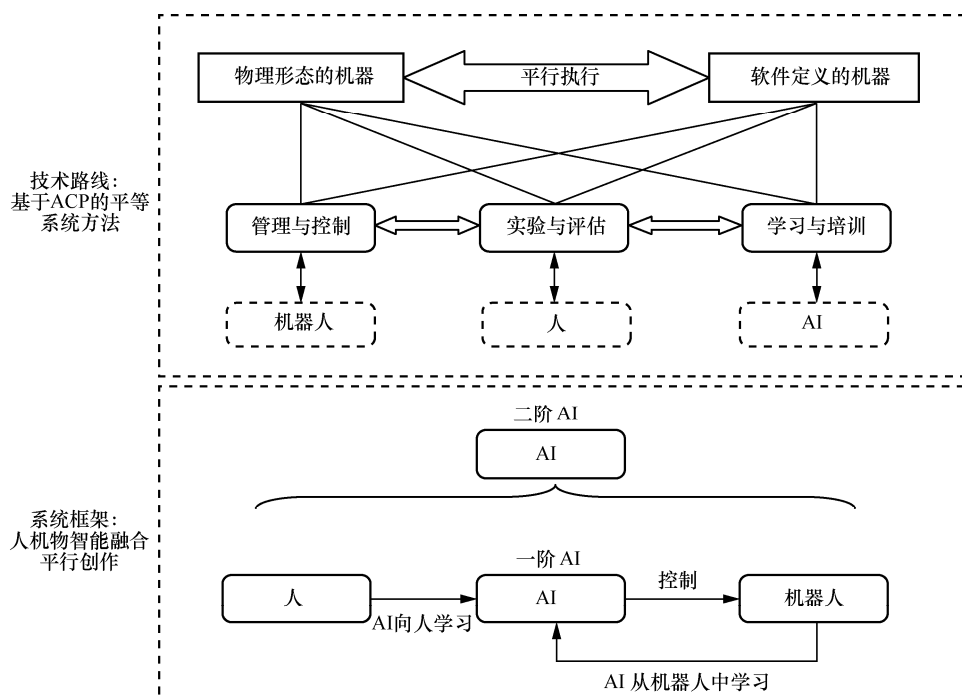


图 2 人机器 CPSS 智能融合的平行创作架构

平行系统理论,通过人类进行实验与评估,对AI进行学习培训,对机器人进行管理与控制。平行创作系统通过一阶AI和二阶AI两个层次的AI来实现对创作过程的优化。一阶AI在机器人绘画执行中对产生的新笔触和画法进行学习,从而产生新的创作技术;二阶AI对人机物智能融合的整体创作过程进行监测,控制三者的执行和交互顺序。上述两个层次的AI能够实现局部绘画策略的控制和对整体绘画流程的把控,构建多层闭环的平行创作过程。

2.1 人工系统

人工系统是人机物CPSS智能融合的平行创作架构的基础组成部分,提供后续计算实验和平行执行的实验环境。本质上,人工系统是人类艺术家在创作环境中构思过程的反映。人类艺术家在进行艺术创作时,首先在脑海中形成作品的概念和印象,在虚拟的脑海世界里想象和绘制出艺术作品,并将其作为后续艺术创作的基础。本文模仿人类艺术家的创作过程,构建人工系统。

首先,构建艺术创作人工系统的软件环境,包括绘画工具建模、绘画环境模拟,为后续绘画实验提供条件。在绘画工具建模方面,对画板、颜料和画笔等绘画工具进行建模,还进行不同绘画工具之间的交互建模。在绘画环境模拟方面,对绘画工具间的相互作用进行模拟,对绘画环境进行设定,使人工系统的绘画条件尽量接近实际的绘画条件。同时,考虑绘画规模、绘画复杂度、绘画重复性等方面的需求,进行人工系统中的参数设计。相比实际系统,人工系统具备实验便捷、操作可逆等优势。例如,对某一绘画因素进行大量的参数尝试,或撤销某一不当的绘画操作带来的影响等。

然后,利用人工系统的软件环境生成大量的实验数据。基于软件环境,人工系统能够吸纳真实艺术家的绘画轨迹等艺术创作数据,并较为方便地对数据进行分割、重组和优化等再加工,进而增强数据的多样性和规模。例如,通过生成式对抗网络和对比学习等技术进行新数据的生成,利用生成的大规模的虚拟数据增大可用数据的规模;通过在生成模型中引入随机的控制变量,在生成的大规模的虚拟数据中增加随机且符合数据分布的数据。上述人工系统的数据生成和推演过程均能够大规模和长时间地运行,解决了实际系统中由材料消耗和人力限制带来的实验资源不足等问题,为艺术创作系统提供了大规模和长时间推演的实验条件。

最后,人工系统对平行艺术创作中各种艺术创作研究内容分别提供相应的实验环境与数据,包括风格迁移、内容组合、笔触生成和图像描述等。针对风格迁移任务,整合多种绘画风格的训练和测试数据,为平行创作中的风格生成提供充足数据;针对内容组合任务,人工系统模拟大规模的、多样的内容组合形式,为平行艺术创作的艺术构图等方面提供更多的选择;针对笔触生成任务,人工系统构建能够进行笔触交互和渲染的笔触实验环境,为平行创作中的笔触策略学习和执行提供基础;针对图像描述任务,人工系统提供大规模的、高质量的训练数据,增强平行创作中的绘画语义理解模型的鲁棒性。

2.2 计算实验

计算实验是平行系统中的核心环节。通过虚拟空间中的计算实验可以进一步扩展对状态和行动的探索空间,提高探索效率。经过十几年的发展,计算实验已成为从小数据生成大数据、从大数据提炼深度智能或精确知识的常规手段,是分析复杂系统的重要方法之一。

艺术创作过程具有复杂度高的特点,人类艺术创作需要观察、构思、创作3个环节的大量尝试、交互与循环,机器艺术创作同样需要通过大量的探索来获得创作技能与审美能力。在人工系统中,AI可以在绘画的认知和表达等方面进行计算实验,例如空间与语义关系认知、笔触色彩与情感表达、笔触序列生成等,从而获得绘画知识和经验。

计算实验中AI艺术创作的案例如图3所示。在绘画风格迁移方面,如图3(a)所示,把一幅图像中风格、纹理等特征迁移到另一幅目标图像的内容中。风格迁移的难点在于把图像的风格特征和内容特征区分开,从而生成具有风格图像的纹理和参考图像的内容的目标图像。在绘画内容组合方面,如图3(b)所示,根据绘画需求选定绘画背景和绘画元素,并进行组合优化生成全新的绘画内容。在绘画笔触生成方面,如图3(c)所示,首先对图像进行语义分割等内容分析,然后在不同区域根据不同规则计算笔触排布和笔触颜色,从而得到单个笔触,最终将笔触整合排序形成笔触序列,实现机器人绘画过程。在绘画图像描述方面,如图3(d)所示,通过文字形式的自然语言对绘画的语义信息进行描述,实现绘画语义的理解,为平行创作提供精准的结构化的语义理解,提高风格迁移和笔触生成等任务中的语义理解精准度。

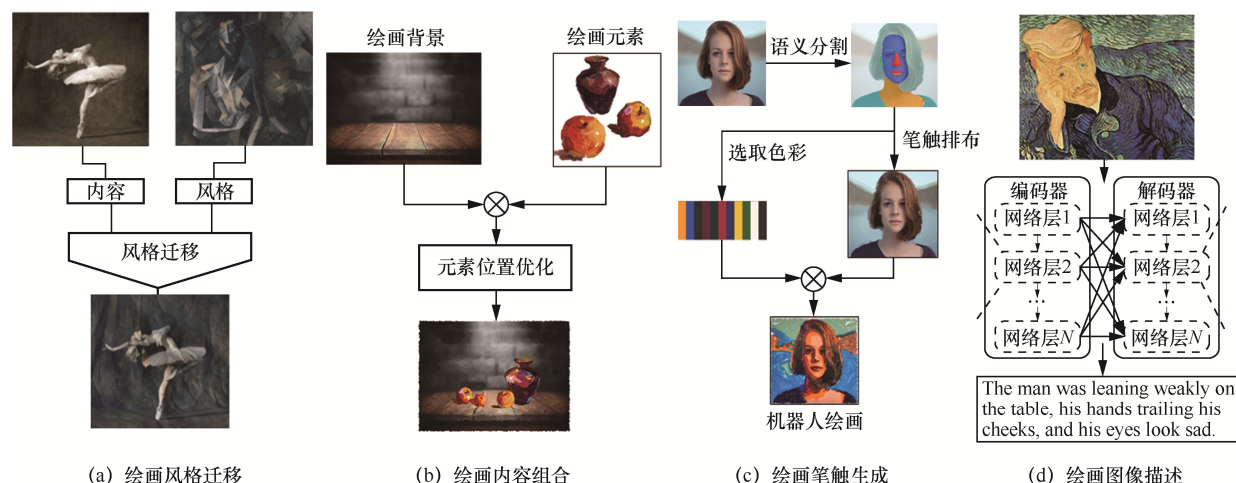


图3 计算实验中 AI 艺术创作的实际案例

2.3 平行执行

在人机物 CPSS 智能融合的平行创作架构中, 平行执行连接人工系统和实际系统, 实现人工系统对实际系统执行过程的交互和引导。人工系统中进行大规模绘画计算实验所获得的创作方案和知识将引导改善实际创作系统, 提升创作效果和效率。该环节直接作用于创作过程, 对创作结果起到决定性作用。平行执行的实现有正向、反向两个过程, 正向过程由人工系统引导实际系统, 根据计算实验结果推荐创作方案, 并将该推荐创作方案部署到实际系统中产生作用。反向过程由实际系统更新人工系统, 通过反馈迭代优化人工系统参数。

在由人工系统引导实际系统的环节, 根据计算实验的创作方案调整实际创作方案, 实现对实际系统运行的改善。实际创作系统和创作计算实验同步执行, 迭代优化, 从而发挥平行创作系统在拓展探索空间和提升探索效率上的优势。在由实际系统更新人工系统的环节, 一方面获取实际创作过程中的状态, 在计算实验中为形成动作方案提供决策依据; 另一方面, 由实际系统数据扩充人工系统, 更新人工系统参数, 在实际系统的演变中使人工系统与实际系统之间拥有较小的差异, 保证人工系统中计算实验对实际系统迁移的有效性。

本节基于平行系统理论构建了平行创作系统, 通过人工系统、计算实验和平行执行分别实现了虚拟创作的构思、创作构思的迭代优化、虚拟与实际创作系统的平行执行。其中, 人工系统、计算实验和平行执行互为基础, 构建了一个不断迭代优化的平行创作闭环系统。通过平行系统理论, 本节对人类艺术家从脑海构思到艺术创作的流程进

行了模拟, 通过类人化的建模方式, 构建了艺术创作人工系统, 为平行创作系统的设计与实现提供了框架基础。

3 平行创作系统的设计与实现

为了验证平行创作系统的有效性, 本文设计了绘画机器人系统^[52-53], 并以此为案例进行分析。绘画机器人系统主要包括绘画过程采集、绘画计算实验和机器人绘画执行 3 个子系统。绘画过程采集子系统的作用是采集动态绘画过程, 一方面收集人类绘画数据, 另一方面为机器人绘画执行子系统提供视觉反馈。机器人绘画执行子系统的作用是完成实际的绘画创作过程。在绘画机器人系统中, 最关键的是绘画计算实验子系统, 其作用是生成绘画方案, 这直接决定了最终绘画结果。因此后文将重点对计算实验子系统风格迁移、内容组合、笔触生成与图像描述等绘画计算实验方法进行详细阐述, 并呈现创作效果, 分析其在水机物智能融合的平行创作中的作用。

3.1 绘画的风格迁移

绘画计算实验子系统中采用风格迁移方法实现对绘画纹理风格的探索。风格迁移的基本思想是分别从内容图像和风格图像中提取内容特征和风格特征, 并将这两个特征重新组合成目标图像, 之后根据生成图像与内容图像和风格图像之间的差异对生成图像进行迭代优化。在实验中, 本文将内容损失函数定义为两者通过 VGG 网络提取的特征之间的欧氏距离, 将风格损失函数定义为两者通过 VGG 网络提取的特征之间的格拉姆 (Gram) 矩阵的欧氏距离。与现有方法对目标图像的像素进行优

化的做法不同,本文采用二次贝塞尔曲线表示笔触模型,通过优化风格损失与内容损失来迭代更新笔触参数,使得采用优化后笔触参数渲染的目标绘画图像同时具备风格图像的风格和内容图像的内容。对于一个深层网络来说,浅层特征更加关注色彩等低维信息,深层特征更加关注内容语义等高维信息。因此,本文使用 VGG-19 网络提取的浅层网络特征计算得到风格损失,深层网络特征计算得到内容损失^[54]。相比直接对目标图像像素进行优化,本文案例中生成的风格化绘画结果具有明显的笔触形态。绘画的风格迁移结果样例如图 4 所示。



(a) 内容图像



(b) 绘画风格图像



(c) 风格迁移结果

图4 绘画的风格迁移结果样例

3.2 绘画的内容组合

内容组合是人工系统中绘画计算实验的重要方式,体现了平行创作系统对绘画对象的创造能力。由于机器人绘画关注笔触序列的叠加过程,因此内容组合在机器人绘画中的表现形式为笔触序列集合的组合。一幅绘画可以认为由若干的绘画内容元素构成,内容组合的目标则是优化这些绘画内容元素

的参数,使构成的最终绘画图像审美评价指标最大化。本文将优化的组合参数定义为每个绘画内容元素的平面位置和大小尺度,采用遗传算法对位置和尺度参数的集合进行优化。在计算得到绘画元素的组合参数之后,将其对应的笔触序列在优化后的画布位置按时序执行,从而完成完整的绘画过程。基于内容组合的绘画生成结果样例如图 5 所示。



(a) 绘画背景



(b) 绘画元素



(c) 组合之后的绘画渲染结果

图5 基于内容组合的绘画生成结果样例

3.3 绘画的笔触生成

基于绘画内容和纹理风格,进一步通过笔触生成方法生成完成绘画所需的笔触序列。笔触序列生成通过绘画笔触叠加的动态过程实现数字绘画构建,其采用特定参数组合的笔触将画布填满,使得其与目标绘画图像的误差最小化。笔触参数通常包括笔触模型、笔触排布和笔触颜色。在笔触模型方面,机器人需要形状规整的笔触形状从而适应其执行过程,因此本文基于笔尖截面形状按照特定轨迹拖动的方法^[20]生成单个绘画笔触。基于上述定义的笔触模型、笔触排布和笔触颜色决定了机器人绘画的结果。

在笔触排布方面,需要考虑机器人对生成笔触轨迹的可执行性,采用基于人工设计绘画规则的启发式方法。笔触的排布主要考虑笔触位置和笔触方向两个因素:笔触的位置从与目标绘画差异较大的位置开始选择;笔触方向则由绘画内容图像的纹理方向确定,纹理方向由像素点处的法向表示。根据笔触位置确定该笔触的起始点,基于该起始点处的纹理方向前进笔尖半径大小的步长,得到笔触轨迹的下一个关键点,依此类推,迭代更新得到笔触轨迹的骨架点。笔尖沿着骨架点表示的笔触轨迹运动得到单个笔触排布结果。在笔触色彩方面,考虑机器人绘画中基颜料和所能画出的色彩是有限的,因此针对整幅绘画范围,提取数目有限的主导颜色,并将其分配到对应位置的笔触。其中,主导色通过对不同语义区域进行像素的色彩聚类获得。针对目标绘画图像的笔触生成渲染和机器人绘画结果样例如图6所示。



(a) 目标绘画图像



(b) 笔触渲染结果



(c) 机器人绘画结果

图6 针对目标绘画图像的笔触生成渲染和机器人绘画结果样例

3.4 绘画的图像描述

绘画图像描述实现将绘画语义从视觉形式转换为文本语言形式,能够为机器绘画的语义理解提供支撑。在平行创作系统的数据获取和绘画创作过程中,图像描述能够以自然语言输出的形式辅助AI对绘画内容进行筛选和控制。本文针对绘画图像描述任务中训练数据稀缺的现状,对模型的结构、训练方法、测试方法进行了设计。在绘画图像描述模型的结构上,采用卷积神经网络和语言生成模型,先对绘画图像提取图像特征,再根据马尔可夫决策过程逐个单词地生成描述文字。在训练方面,在人工系统中使用风格迁移技术生成大规模高质量的虚拟训练数据。在测试方面,采用真实的绘画图像和本文平行创作系统创作的绘画图像进行测试。

具体来说,绘画图像描述模型的构建包括数据准备和模型训练两方面。在数据准备方面,考虑到现有的带标签数据不能很好地满足视觉和语言的联合建模需求,采用人工系统的方式生成大规模的虚拟标注数据用于训练。其中,采用风格迁移技术将现有的自然图像描述数据集中的图像迁移为绘画风格,同时保留其语义信息和标注信息,形成具有绘画风格的虚拟绘画图像标注数据集。在模型训练方面,为了充分利用图像中全局和局部的视觉特征,通过目标检测方法来提取物体级别特征,再通过网状记忆 Transformer 模型^[55]生成文字描述。目标检测提取多个物体的特征,每个物体的特征由1024维度的向量表示,将特征作为网状记忆 Transformer 模型的输入,再计算生成句子中每个位置上的单词的预测概率值。因此,上述过程通过风格迁移进行了人工系统的构建和虚拟数据的生成与扩充,再经过语言生成模型生成绘画图像的描述文字。绘画图像描述模型的输出结果样例^[56-57]如图7所示,通过对模型输出文字、标签文字和绘画图像的对比,能够看出绘画图像描述模型在绘画语义信息提取上的作用。

本文基于平行创作架构设计了绘画机器人系统,通过采集和模拟人类的动态绘画过程建立绘画人工系统,基于风格迁移、内容组合、笔触生成和图像描述进行绘画计算实验,并引导机器人完成实际的绘画创作过程。与现有AI创作算法相比,该绘画创作系统模拟人类动态绘画过程,同时实现物理意义上的绘画创作,提升了绘画创作效果。更重

要的是,该机器人绘画系统构建起人、AI创作算法、机器人之间协同创作的闭环以及通用的机器绘画创作算法流程,为进一步释放机器的创造能力奠定了基础。

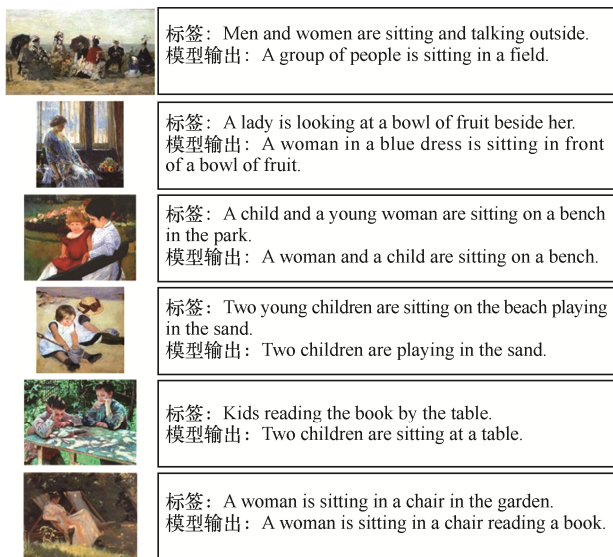


图7 绘画图像描述模型的输出结果样例

4 结束语

本文针对人工智能艺术创作,对人机物的智能进行了融合利用,基于平行系统理论与ACP方法设计了平行创作系统。通过上述平行创作系统,探索了平行创作中的风格迁移、内容组合、笔触生成和图像描述等绘画计算实验关键方法,取得了一系列的阶段成果。在平行创作系统中,人类艺术家和机器人将相互受益,机器人一方面从人类的示范中学习绘画技巧,另一方面为人类提供辅助绘画创作服务,形成了闭环迭代的创作优化框架。人机物智能融合的平行创作发挥了人、AI创作算法和机器人的优势,实现人机融合学习和融合创作,体现了平行系统与其结合应用的必然趋势和重要作用。

未来,可以将风格迁移、笔触生成、图像描述等内容作为平行创作中的交互“小游戏”,让元宇宙中更多的人参与到平行创作中,进一步增强平行创作对人工智能艺术创作的作用。未来,还可以对各种艺术创作研究进行更加深入的优化。例如,在绘画风格迁移方面,现有基于笔触的风格迁移方法容易忽视笔触中的语义信息,使不同的语义区域间笔触存在同质化的问题,未来

可尝试在考虑图像语义的同时进行笔触风格的迁移。在绘画图像描述方面,可挖掘绘画中更多的信息,生成更加全面、丰富的绘画语义信息描述,为平行艺术创作元宇宙提供更加精准全面的语义信息解读,例如画面的风格和情感等。同时,可以对更加广泛的研究内容进行实验和探讨,不断丰富平行创作的组成部分,例如,对艺术品美学质量评估进行研究,使平行创作系统像人类一样对图像的美感进行感知、分析和决策。

参考文献:

- [1] MCCARTHY J, MINSKY M, ROCHESTER N, et al. A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955[J]. AI Mag, 2006, 27: 12-14.
- [2] BODEN M A. Creativity and artificial intelligence[J]. Artificial Intelligence, 1998, 103(1/2): 347-356.
- [3] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. Nature, 2016, 529(7587): 484-489.
- [4] WANG F Y, ZHANG J J, ZHENG X H, et al. Where does AlphaGo go: from church-Turing thesis to AlphaGo thesis and beyond[J]. IEEE/CAA Journal of Automatica Sinica, 2016, 3(2): 113-120.
- [5] 郭超, 鲁越, 林懿伦, 等. 平行艺术: 人机协作的艺术创作[J]. 智能科学与技术学报, 2019, 1(4): 335-341.
- [6] GUO C, LU Y, LIN Y L, et al. Parallel art: artistic creation under human-machine collaboration[J]. Chinese Journal of Intelligent Science and Technology, 2019, 1(4): 335-341.
- [7] RAMESH A, DHARIWAL P, NICHOL A, et al. Hierarchical text-conditional image generation with CLIP latents[J]. arXiv preprint, 2022, arXiv:2204.06125.
- [8] GONTHIER N, GOUSSEAU Y, LADJAL S, et al. Weakly supervised object detection in artworks[M]//Lecture notes in computer science. Cham: Springer International Publishing, 2019: 692-709.
- [9] RODRIGUEZ C S, LECH M, PIROGOVA E. Classification of style in fine-art paintings using transfer learning and weighted image patches[C]//Proceedings of 2018 12th International Conference on Signal Processing and Communication Systems. Piscataway: IEEE Press, 2018: 1-7.
- [10] CETINIC E, LIPIC T, GRGIC S. A deep learning perspective on beauty, sentiment, and remembrance of art[J]. IEEE Access, 2019, 7: 73694-73710.
- [11] DA YI, GUO C, BAI T X. Exploring painting synthesis with diffusion models[C]//Proceedings of 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence. Piscataway: IEEE Press, 2021: 332-335.
- [12] CETINIC E, SHE J. Understanding and creating art with AI: review and outlook[J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 2022, 18(2): 1-22.
- [13] WANG F Y. Shadow systems: a new concept for nested and embedded co-simulation for intelligent systems[R]. 1994.

- [13] 王飞跃. 关于复杂系统研究的计算理论与方法[J]. 中国基础科学, 2004, 6(5): 3-10.
WANG F Y. Computational theory and method on complex system[J]. China Basic Science, 2004, 6(5): 3-10.
- [14] 王飞跃. 人工社会、计算实验、平行系统: 关于复杂社会经济系统计算研究的讨论[J]. 复杂系统与复杂性科学, 2004, 1(4): 25-35.
WANG F Y. Artificial societies, computational experiments, and parallel systems: a discussion on computational theory of complex social-economic systems[J]. Complex Systems and Complexity Science, 2004, 1(4): 25-35.
- [15] 王飞跃. 平行系统方法与复杂系统的管理和控制[J]. 控制与决策, 2004, 19(5): 485-489, 514.
WANG F Y. Parallel system methods for management and control of complex systems[J]. Control and Decision, 2004, 19(5): 485-489, 514.
- [16] 王飞跃, 史帝夫·兰森. 从人工生命到人工社会: 复杂社会系统研究的现状和展望[J]. 复杂系统与复杂性科学, 2004, 1(1): 33-41.
WANG F Y, LANSING J S. From artificial life to artificial societies—new methods for studies of complex social systems[J]. Complex Systems and Complexity Science, 2004, 1(1): 33-41.
- [17] WANG F Y, WANG X, LI L X, et al. Steps toward parallel intelligence[J]. IEEE/CAA Journal of Automatica Sinica, 2016, 3(4): 345-348.
- [18] 杨林瑶, 陈思远, 王晓, 等. 数字孪生与平行系统: 发展现状、对比及展望[J]. 自动化学报, 2019, 45(11): 2001-2031.
YANG L Y, CHEN S Y, WANG X, et al. Digital twins and parallel systems: state of the art, comparisons and prospect[J]. Acta Automatica Sinica, 2019, 45(11): 2001-2031.
- [19] GATYS L A, ECKER A S, BETHGE M. Image style transfer using convolutional neural networks[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 2414-2423.
- [20] HERTZMANN A. Painterly rendering with curved brush strokes of multiple sizes[C]//Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques. [S.l.:s.n.], 1998: 453-460.
- [21] GOOCH B, COOMBE G, SHIRLEY P. Artistic vision: painterly rendering using computer vision techniques[C]//Proceedings of the 2nd International Symposium on Non-photorealistic Animation and Rendering. [S.l.:s.n.], 2002: 83.
- [22] HERTZMANN A, JACOBS C E, OLIVER N, et al. Image analogies[C]//Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques. [S.l.:s.n.], 2001: 327-340.
- [23] TOMASI C, MANDUCHI R. Bilateral filtering for gray and color images[C]//Proceedings of the 6th International Conference on Computer Vision. Piscataway: IEEE Press, 1998: 839-846.
- [24] JOHNSON J, ALAHI A, LI F F. Perceptual losses for real-time style transfer and super-resolution[C]//Computer Vision – ECCV 2016. [S.l.:s.n.], 2016: 694-711.
- [25] CHEN D W, YUAN L, LIAO J, et al. Stylebank: an explicit representation for neural image style transfer[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 1897-1906.
- [26] HUANG X, BELONGIE S. Arbitrary style transfer in real-time with adaptive instance normalization[C]//Proceedings of 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 1510-1519.
- [27] COLTON S, WIGGINS G A. Computational creativity: the final frontier? [C]//Proceedings of the 20th European Conference on Artificial Intelligence. [S.l.:s.n.], 2012: 21-26.
- [28] ESSER P, ROMBACH R, OMMER B. Taming transformers for high-resolution image synthesis[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 12868-12878.
- [29] RAMESH A, PAVLOV M, GOH G, et al. Zero-shot text-to-image generation[J]. arXiv preprint, 2021, arXiv:2102.12092.
- [30] ZHENG Q Y, LI Z R, BARGTEIL A. Learning aesthetic layouts via visual guidance[J]. arXiv preprint, 2021, arXiv:2107.06262.
- [31] LI J N, YANG J M, HERTZMANN A, et al. LayoutGAN: synthesizing graphic layouts with vector-wireframe adversarial networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(7): 2388-2399.
- [32] ZOU Z X, SHI T Y, QIU S, et al. Stylized neural painting[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 15684-15693.
- [33] HEGDE S, GATZIDIS C, TIAN F. Painterly rendering techniques: a state-of-the-art review of current approaches[J]. Computer Animation and Virtual Worlds, 2013, 24(1): 43-64.
- [34] HUANG Z W, ZHOU S C, HENG W. Learning to paint with model-based deep reinforcement learning[C]//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 8708-8717.
- [35] VINYALS O, TOSHEV A, BENGIO S, et al. Show and tell: a neural image caption generator[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2015: 3156-3164.
- [36] HOSSAIN M Z, SOHEL F, SHIRATUDDIN M F, et al. A comprehensive survey of deep learning for image captioning[J]. ACM Computing Surveys, 2019, 51(6): 1-36.
- [37] BAI S, AN S. A survey on automatic image caption generation[J]. Neurocomputing, 2018, 311: 291-304.
- [38] WANG W S, NA X X, CAO D P, et al. Decision-making in driver-automation shared control: a review and perspectives[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 7(5): 1289-1307.
- [39] LIU T, TIAN B, AI Y F, et al. Parallel reinforcement learning-based energy efficiency improvement for a cyber-physical system[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 7(2): 617-626.
- [40] WEI Q L, LI H Y, WANG F Y. Parallel control for continuous-time linear systems: a case study[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 7(4): 919-928.
- [41] LI X S, LIU Y T, WANG K F, et al. A recurrent attention and interaction model for pedestrian trajectory prediction[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 7(5): 1361-1370.
- [42] LIU K H, YE Z H, GUO H Y, et al. FISS GAN: a generative adversarial network for foggy image semantic segmentation[J]. IEEE/CAA Journal of Automatica Sinica, 2021, 8(8): 1428-1439.
- [43] LU J W, WEI Q L, WANG F Y. Parallel control for optimal tracking via adaptive dynamic programming[J]. IEEE/CAA Journal of Automatica Sinica, 2021, 8(8): 1428-1439.

tica Sinica, 2020, 7(6): 1662-1674.

- [44] SUN C, VIANNEY J M U, LI Y, et al. Proximity based automatic data annotation for autonomous driving[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 7(2): 395-404.
- [45] WANG S Y, HOUSDEN J, BAI T X, et al. Robotic intra-operative ultrasound: virtual environments and parallel systems[J]. IEEE/CAA Journal of Automatica Sinica, 2021, 8(5): 1095-1106.
- [46] ZU C Y, YANG C, WANG J, et al. Simulation and field testing of multiple vehicles collision avoidance algorithms[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 7(4): 1045-1063.
- [47] TAN J Y, XU C L, LI L, et al. Guidance control for parallel parking tasks[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 7(1): 301-306.
- [48] 张俊, 许沛东, 王飞跃. 并行系统和数字孪生的一种数据驱动形式表示及计算框架[J]. 自动化学报, 2020, 46(7): 1346-1356.
ZHANG J, XU P D, WANG F Y. Parallel systems and digital twins: a data-driven mathematical representation and computational framework[J]. Acta Automatica Sinica, 2020, 46(7): 1346-1356.
- [49] JAYNES C, SEALES W B, CALVERT K, et al. The metaverse: a networked collection of inexpensive, self-configuring, immersive environments[C]//Proceedings of the Workshop on Virtual Environments 2003 – EGVE'03. New York: ACM Press, 2003: 115-124.
- [50] DUAN H H, LI J Y, FAN S Z, et al. Metaverse for social good: a university campus prototype[C]//Proceedings of the 29th ACM International Conference on Multimedia. New York: ACM Press, 2021: 153-161.
- [51] WANG F Y. Parallel intelligence in metaverses: welcome to Hanoi![J]. IEEE Intelligent Systems, 2022, 37(1): 16-20.
- [52] GUO C, BAI T X, LU Y, et al. Skywork-daVinci: a novel CPSS-based painting support system[C]//Proceedings of 2020 IEEE 16th International Conference on Automation Science and Engineering. Piscataway: IEEE Press, 2020: 673-678.
- [53] GUO C, BAI T X, WANG X, et al. ShadowPainter: active learning enabled robotic painting through visual measurement and reproduction of the artistic creation process[J]. Journal of Intelligent & Robotic Systems, 2022, 105(3): 1-17.
- [54] KOTOVENKO D, WRIGHT M, HEIMBRECHT A, et al. Rethinking style transfer: from pixels to parameterized brushstrokes[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 12191-12200.
- [55] CORNIA M, STEFANINI M, BARALDI L, et al. Meshed-memory transformer for image captioning[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 10575-10584.
- [56] LU Y, GUO C, DAI X Y, et al. Image captioning on fine art paintings via virtual paintings[C]//Proceedings of 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence. Piscataway: IEEE Press, 2021: 156-159.
- [57] LU Y, GUO C, DAI X Y, et al. Data-efficient image captioning of fine art paintings via virtual-real semantic alignment training[J]. Neurocomputing, 2022, 490: 163-180.

[作者简介]



郭超 (1992–), 男, 中国科学院自动化研究所复杂系统管理与控制国家重点实验室助理研究员, 主要研究方向为机器学习、强化学习、机器艺术创作、图像生成、机器人绘画等。



鲁越 (1994–), 男, 中国科学院自动化研究所复杂系统管理与控制国家重点实验室博士生, 主要研究方向为机器学习、小样本学习、风格迁移等。



王晓 (1988–), 女, 博士, 中国科学院自动化研究所复杂系统管理与控制国家重点实验室副研究员, 主要研究方向为群体行为的激发与汇聚激励、群体智能和社交网络挖掘与分析。



易达 (1996–), 男, 中国科学院自动化研究所复杂系统管理与控制国家重点实验室硕士生, 主要研究方向为机器学习、深度学习、绘画图像生成、艺术风格迁移等。



王琥 (1996–), 男, 普渡大学计算机系博士生, 主要研究方向为自监督学习、计算生物学和智能交通系统等。



王飞跃 (1961–), 男, 中国科学院自动化研究所复杂系统管理与控制国家重点实验室研究员, 国防科技大学军事计算实验与并行系统技术研究中心教授, 中国科学院大学中国经济与社会安全研究中心主任, 主要研究方向为智能系统和复杂系统建模、分析与控制。