艺术图像创作智能交互系统：挑战、方法与未来展望

邓晃煌[1]，孙傲然[2]，孟展[3]，王永威[4]，陈露阳[4]，李敦涵[2]，
李知予[4]，韩越[2]，李泽健[3]，况琨[2]，刘泽民[2]，邹宁[5]，
赵洲[2]，张克俊[2]，孙凌云[2]，赵瑜[6,4]，吴飞[2]，潘云鹤[5]
[1]浙江大学区块链与数据安全全国重点实验室，中国杭州市，310027
[2]浙江大学人工智能学院，中国杭州市，310058
[3]浙江大学软件学院，中国杭州市，310058
[4]浙江大学传媒与国际文化学院，中国杭州市，310058
[5]浙江大学计算机科学与技术学院，中国杭州市，310027
[6]浙江大学数智视听与影视艺术研究所，中国杭州市，310058

摘要：在生成式人工智能（AI）的推动下，图像视觉艺术正经历一场深刻变革。这场变革不仅超越了技术创新层面，更在重塑艺术创作的范式。本文论述了以人类为中心的图像艺术交互系统，旨在弥合技术导向与艺术导向研究之间的鸿沟。文中明确交互式生成艺术创作可控性、守正性以及创新性三大核心挑战，并深入剖析三大挑战背后潜在的制衡关系。在此基础上，系统回顾了生成模型的发展历程、艺术知识的融合方式以及协作式智能体AI系统的兴起。展望未来，本文强调需构建兼具更高可控性、保障人类创造主体性且保障合规性的人机协同框架，并探讨了该框架广泛的潜在应用场景。本研究提出连接人工智能技术、人机交互过程与艺术图像融合创新的统一分析框架，同时绘制了以任务为导向的路线图，为未来相关研究及协同创作系统开发提供指导。

关键词：跨媒体智能；人机协同；视觉艺术创作；交互系统；可控性；守正性；创新性
https://doi.org/10.1631/ENG.ITEE.2026.0124