

# 受艺术家创造性思维启发的扩散模型多阶段引导

齐旺<sup>1</sup>, 邓晁煌<sup>2</sup>, 李太豪<sup>1</sup>

<sup>1</sup>之江实验室跨媒体智能研究中心, 中国杭州市, 311500

<sup>2</sup>浙江大学计算机科学与技术学院, 中国杭州市, 310027

**摘要:** 目前文本生成图像的研究已显示出与普通画家类似的水平, 但与艺术家绘画水平相比仍有很大改进空间; 艺术家水平的绘画通常将多个意象的特征融合到一个意象中, 以表示多层次语义信息。在预实验中, 我们证实了这一点, 并咨询了3个具有不同艺术欣赏能力的群体的意见, 以确定画家和艺术家之间绘画水平的区别。之后, 利用这些观点帮助人工智能绘画系统从普通画家水平的图像生成改进为艺术家水平的图像生成。具体来说, 提出一种无需任何进一步预训练的、基于文本的多阶段引导方法, 帮助扩散模型在生成的图像中向多层次语义表示迈进。实验中的机器和人工评估都验证了所提方法的有效性。此外, 与之前单阶段引导方法不同, 该方法能够通过控制不同阶段之间的指导步数来控制各个意象特征在绘画中的表现程度。

**关键词:** 文本生成图像; 扩散模型; 多层次语义; 多阶段引导

<https://doi.org/10.1631/FITEE.2300313>