

# AI生成时代的图像作品： 平台语境中的审美样板化与知识生产问题

丁依然 董晨宇 颜晓鹏

**【摘要】**生成式人工智能的兴起正深刻改变着社交平台中人像生产的逻辑。研究以小红书平台为场域,通过对306条AI人像教程文本及其生成的790张图像的实证考察,探讨AI图像作品的特征及相关的图像知识生产方式。研究发现,在图像层面,人像生产由传统摄影的现实记录转向算法驱动的参数化生成,其核心在于复刻当下既存的审美公约数、打造差异化的审美样板;提示词作为一种审美-技术中介,将感性经验简化为功能指令,预设了视觉表达范畴,使个体的审美判断深度受制于算法体系。研究进一步揭示了AI赋权下的图像民主化悖论,并对图像和审美知识的窄化现象进行了省思。

**【关键词】**生成式人工智能;提示词;审美样板化;知识生产

**【中图分类号】**G206 **【文献标识码】**A

## 一、研究背景与研究问题

提示词:

"prompt": "专业高清摄影写真,18岁东亚美女,短黑波波头刘海遮额,蓝色美瞳超大无辜眼睛(kawaii可爱眼神),俏皮甜美微笑嘴角上扬带小酒窝,头部微微侧倾向上看镜头(head tilt upward playful),眼神灵动带点害羞俏皮;冷调白瓷肌,柔雾野生眉,暖桃橙眼影+卧蚕,狗狗眼细棕眼线,簇状睫毛,高饱和珊瑚粉晒伤腮红大面积鼻梁眼下,珠光鼻尖点高光,咬唇暖红珊瑚唇模糊晕染;湿透白短T恤+深蓝牛仔裤,站在蓝调时刻浅水区,水到小腿;可爱小动作:一只手轻触水面玩水溅起小水花,或另一手微微比心/手指轻卷刘海;水面精细涟漪+闪耀反射+月光光路,背景广阔宁静夜色水域;柔和冷调漫射光,无硬

影,头发轮廓光、湿衣水珠折射、皮肤自然光影过渡;轻朦胧梦幻柔焦氛围,电影感,杰作,8K级高清,极致细节,可爱俏皮表情,甜美灵动",

……

上述这类在社交平台上随处可见的提示词教程,正成为当下视觉生产的通用脚本。随着生成式人工智能(AIGC)渗入日常传播,人像生产正经历着从具身摄影向程序化指令的范式转变。图像不再仅是物理现实的索引,而是演变为由参数驱动的、可计算的样板化产品。如何理解这种技术中介下的视觉文化产物,构成了本文的核心关切。

在社交平台的视觉场域中,人工智能工具的普及表面上铲平了专业技艺的壁垒,实现了某种形式的图像民主化。实际上,这种技术赋能的过程伴随着更为隐蔽的规训逻辑。当个体的感性审美意图必须经由提示词(Prompt)这一功能性中介进

行转译时,原本多维的审美经验被降维为离散的代码指令,独属于人的审美主权在追求效率与确定性的过程中发生让渡。这使得关注算法时代的视觉政体与知识转型问题变得尤为重要。

在此背景下,本文聚焦于两个相互关联的问题:人工智能提示词教程引导下的平台图像在审美上呈现出怎样的特征?提示词作为审美—技术中介,如何转化感官维度的审美经验?为此,研究以社交平台小红书为研究场域,通过对300余条提示词教程文本及其对应的790张生成图像进行实证考察,试图剖析审美经验向程序化指令转译的过程,并探讨图像知识所经历的本体性变迁。

## 二、文献综述

### (一) 图像生产的逻辑:从索引到生成

提及图像,早期的社会大众往往将其视作现实世界的镜像反映或是创作者情感意志的艺术表达。在摄影技术主导的漫长周期内,图像主要被认为是物理世界在感光材料上留下的真实痕迹。这种直觉式的认知将图像看作是中立的、客观的记录工具,其审美价值往往取决于创作者捕捉瞬间的敏感度以及对器材的掌控力。然而,随着生成式人工智能的兴起,这种联系正在被离散的数据计算所取代,致使图像的生产逻辑从基于现实的索引性记录转向基于算法的模拟性生成。

瓦尔特·本雅明(Walter Benjamin)在探讨摄影技术的早期社会影响时,提出“灵韵”(aura)这一核心概念。在本雅明看来,传统艺术品的价值源于其在特定时空中的唯一性,而机械复制技术通过大规模生产剥离了这种“此时此地”的独特性<sup>①</sup>。即便如此,本雅明时代的影像技术依然保持着与物理世界的紧密联系。无论图像如何被复制,它在最初生成的时刻,必须依赖光线在感光材料上留下的痕迹。这种物理上的关联构成了传统影像的索引性(indexicality)特征。

尽管本雅明在其著作中并未直接使用这一术语,但他通过“痕迹”(Spur)这一概念描述了影像与现实之间的物理关联。在《摄影小史》中,本雅明观察到早期摄影由于曝光时间极长,“拍照对象势必得在镜头前久久静止不动,照片的成影便因而综合了人物在这段时间所出现的表情”<sup>②</sup>。这样,早期银版照片的“像”本身便与现实世界形成了一种不可分割的物理纽带。这种关系在符号学中被查尔斯·桑德斯·皮尔斯(C. S. Peirce)定义为“索引符号”,即符号与其对象之间存在因果性的物理联系<sup>③</sup>。此后,罗兰·巴特(Roland Barthes)将其感性地表述为“此曾在”(That-has-been),以此确认拍摄对象的曾在场<sup>④</sup>。

苏珊·桑塔格(Susan Sontag)在《论摄影》中进一步强化了图像作为“现实痕迹”的讨论。她认为摄影本质上是一种对世界的采集和记录,照片被视为现实世界的某种延伸或战利品。桑塔格指出,图像的生产建立在拍摄者对现实的“占有”之上<sup>⑤</sup>,这种占有确认了拍摄对象曾经在场。在此视角下,图像不仅是审美对象,更是证明现实存在的物理证据。这种对现实的高度依赖,确立了传统摄影时代图像本体论的基石:即没有物质现实的在场(或被指涉物的物理呈现),就没有图像的生成。

如果说Photoshop等图像处理软件在2000年代的普及以及Instagram等视觉平台于2010年左右的兴起在不断挑战这一基础,那么人工智能生成图像等模型的日益成熟则标志着这种索引性逻辑的彻底断裂。图像的产生不再依赖于某个具体人脸对光线的反射,而是源于算法对海量人像数据中视觉特征的提取与概率合成。这意味着图像本体从现实痕迹的捕捉转变为数据的重组。图像开始作为一种纯粹的、脱离了物理索引的视觉参数集合而存在。亚历山大·R·加洛韦(Alexander R. Galloway)生动地描述二者的区别:“如果说摄影之眼是凸出的(convex),像船首从中间扎入世界一样的话,

那么计算之眼便是凹陷的(concave),从世界边缘包抄其侧方并包围着世界”。<sup>⑥</sup>

列夫·马诺维奇(Lev Manovich)在《新媒体的语言》中提出的“数据库逻辑”为上述这种参数化生产提供了理论解释。马诺维奇认为,新媒体作品不再是传统的叙事结构,而表现为数据库中各种元素的排列组合<sup>⑦</sup>。他提出的数值化呈现与模块化法则,可以用于理解生成式图像的构造方式。前者是指,“所有的新媒体对象,无论是完全由计算机创建出来的,还是从其他的模拟媒体资源转化而来的,都是由数字符码构成的……这一事实带来了两个重要结果:(1)一个新媒体对象是可以利用特定的形式(数学形式)来描述的……(2)新媒体对象受算法操控。”<sup>⑧</sup>在生成式图像中,无论是人像的皮肤质感、照明的光影角度,还是面部的骨骼比例,这些图像的视觉属性都可被转化为可计算的离散数值。模块化则是指,“所有的媒体元素都表现为离散采样(像素、多边形、立体像素、字符、脚本)的集合。”<sup>⑨</sup>它使得审美体验发生了显著的知识化转向。综述第三部分将对此展开详述。

值得注意的是,媒介的数值编码与模块化结构共同赋予了新媒体对象操作层面的自动化潜质,并在此基础上促成了其形态上无限版本增殖的可变性<sup>⑩</sup>。而在现实传播中,社交平台等视觉场域中充斥的图像却常常表现出高度的特征收敛。那么,技术赋予的自由度为何在现实中呈现为特定的雷同样貌?这促使我们超越图像生产逻辑的探讨,转向考察社交平台这一图像的重要流通场域。

## (二)平台上的视觉政体:孕育类同审美

视觉生产并非发生在孤立的技术真空中,它始终受到特定场域逻辑的引导。当图像作为社会资本在平台中流通时<sup>⑪</sup>,其生成的自由度往往让位于对可见性的追逐,进而受制于一套预设的审美规范与分发法则。因此,有必要关注决定“何为美、如何看、何者可见”的深层结构。

马丁·杰(Martin Jay)提出的“视觉政体”(scopic regimes)概念具有启发性。杰认为,在不同的历史时期,社会对于“如何看”有着不同的规则和共识。例如,笛卡尔透视主义强调单眼的、静态的、理性的观察视角。这种政体将世界看作是处于几何规律约束下的自然客体,观者与对象之间保持着冷静的距离。在透视模式下,人们习惯于像旁观者一样冷静,试图从画面中找出中心的逻辑和因果<sup>⑫</sup>。除了笛卡尔透视主义,杰还总结了另外两种视觉政体:发源于17世纪荷兰的描绘艺术以及巴洛克视觉。

杰所讨论的宏大视觉范式显然难以解释当下新技术语境中的视觉秩序。过去那种涉及哲学观和空间感的视觉政体,在一定程度上,正在被简化和压缩为平台上的具体的视觉风格。而且,作为视觉对象的主要环境和支撑条件,平台成为影响事物可见状态以及可见标准的重要因素。托马斯·波尔(Thomas Poell)与大卫·尼堡(David Nieborg)提出,文化生产正被纳入由数据化、商品化和选择性三个维度构成的平台底层逻辑中。他们将这一过程概念化为“文化生产的平台化”(platformization of cultural production)<sup>⑬</sup>。平台之上的视觉文化生产当然在此讨论范围中。

数据化是这一过程的起点。它呼应了前文马诺维奇的观点,指出平台将原本主观的审美体验转化为可度量的参数,如提示词组合、光影数值以及用户的互动数据。这种转化使得美感变得可计算,平台得以通过后台监测识别出具有传播潜力的视觉模式。审美因此从一种感性的表现,转变为一组可以被追踪、对比和优化的量化指标。

商品化逻辑将图像视为一种标准化的社交资产。为提高审美产品在平台上的流通效率,平台通过提供预制模板和同款功能,将视觉生产转变为可快速复制的流程。这种标准化过程显著降低了用户的创作门槛,同时也鼓励用户产出符合

流量逻辑的视觉产品。在这种机制下,个体的个性化尝试往往被效率更高的样板化生产所取代。

选择性则是通过算法实施的审美准入与过滤机制。算法基于数据反馈,对符合平台偏好的图像或内容进行优先分发,从而决定了其可见程度<sup>⑧</sup>。研究已证实,社交媒体博主常通过模仿和规范化生产来换取潜在的平台资本<sup>⑨</sup>。例如,为争取可见性,创作者会迎合平台假定的受众聚类逻辑,采用高度标准化的音频、视觉结构与标签。研究者将此现象概念化为“回声式连结”(echoic affiliation)<sup>⑩</sup>。在此机制下,样板成为一种应对分发的生存策略,复杂的个人表达被简化为易于机器识别的数据点。这种在“顺从”与“创造”间的权衡<sup>⑪</sup>,最终在不觉中固化了主流的视觉表现形式。

对于不期待依靠可见性获利的普通大众,主动使用样板则很可能是出于强烈的靠近“美”的欲望。本雅明在《机械复制时代的艺术作品》中揭示了大众的这种深层心理动机。他指出,“现在的大众对于‘拉近’自己与艺术品的空间距离和情感距离所投注的关切,就和本身那种以接受复制品来压制艺术原作的唯一性的倾向同样强烈”。曾经,“人们通过持有周遭事物的影响……来掌握这些事物本身……”;如今,通过调用算法指令,人们就能占有别致的审美品味、神秘的AI知识,甚至让自身的像进入那个曾经具有唯一性的世界。这也成为样板得以在平台迅速风靡的原因<sup>⑫</sup>。

样板文化(template culture)不仅体现了技术架构对内容生产的深度规训,还揭示了平台治理的微观逻辑。平台并不直接要求用户生成内容的形式,而是通过流量分配和界面指引,使特定的视觉特征成为共识。这种共识便可被视为一种隐形的视觉政体。它在一定程度上解释了为什么社交平台上会形成标准相对固化的、集体性的审美认同。

不过,即便平台划定了明确的视觉样板边界,对于绝大多数缺乏专业艺术训练的普通用户

而言,从理解样板到产出样板之间仍存在巨大的技能鸿沟。这种集体性的“去技能化”生产究竟通过何种媒介得以实现?

(三)图像的知识论转型:从审美技艺到程序化指令

根据波兰尼对隐性知识(tacit knowledge)的阐述,知识的默会性在很大程度上源于身体感官的不可替代性<sup>⑬</sup>。这一概念常被后来学者延伸至审美创作领域,用以解释艺术实践中那些难以言明的创造性直觉与身体技艺。具体说来,摄影师对光影细微动态的捕捉,或画师对色彩微妙差值的感知,都是基于长期感官训练形成的身体直觉。本雅明如是说:“实际上,摄影师与他们所掌握的摄影技术的关系,已对摄影产生最具决定性的影响。”<sup>⑭</sup>这些知识与个体经验深度绑定,难以通过结构化的语言进行完整拆解与传递。而生成式技术正在使审美能力与传统的身体技艺发生根本性的剥离。在此过程中,审美直觉被重构为两类显性知识:关于图像的知识,以及作为知识的图像。

### 1. 关于图像的知识

在算法生产的视野下,关于图像的知识不再聚焦“感知”,其重要性迅速被参数知识、结构知识、模型特征知识等智能技术维度的知识取代,正是这些知识高度影响着算法的生成路径、生成逻辑和生成结果。

提示词是此类知识中相对特殊的部分。它可被视为桥接审美和技术世界的关窍,因此本文将视其作为一种审美—技术中介。但必须指出,这种桥接必须依靠语言文字的转译才能达成。因此它实际上是一种依靠文字中介才能被调用的中介。而关于审美的默会知识在被转化为文字时很可能已经遭到第一次损耗。

研究指出,使用提示词设定艺术风格本质上是一种“可供性实现(affordance actualization)”实践。技术修饰词(Modifiers)与技术参数

(如`--style raw, --sref`)协助个体实现其审美偏好,原本抽象的艺术风格由此转化为与特定词汇相关联的预测值<sup>⑥</sup>。这样,原本需要通过数年训练才能掌握的审美技艺,便被简化为只要记住并输入特定指令即可完成的操作。用户不再需要掌握如何绘画或摄影,也无需理解算法和模型的工作原理,而只需要掌握关于这些行为的语言描述,便可进行“美”的生产。

随着人们更加娴熟地进行“逆向工程”——即使用各种技术进行将人类摄影作品反向解析为文本提示词,具备同等审美吸引力的“合成孪生图像”被批量生产出来<sup>⑦</sup>。早已受到重创的“灵韵”、本真性和“此时此地”在此消解无踪。

不过必须承认,程序化知识的建构显著提升了视觉内容的传播效率。在新闻摄影等领域,提示词能够将感性认知自动转化为符合专业构图与光影标准的视觉参数<sup>⑧</sup>。在商业应用中,基于提示词生成的图像在点击率等量化指标上甚至超过了人类专业摄影作品<sup>⑨</sup>。但是这种高效的符号传播,也进一步推动同质化视觉风格的蔓延。

## 2. 作为知识的图像

在本研究的语境下,“作为知识的图像”即根据提示词最终生成的图像本身。这种依托计算生成的视觉产物在多大程度上能被当作知识?波兰尼将知识定义为一种认知行为,而非单纯的客观信息。他认为知识是认知主体将零散的细节整合为有意义的整体的过程<sup>⑩</sup>。如上所述,传统摄影中,图像凝结了摄影师与相机、光影之间难以言传的具身经验,它对于相关行动者而言具备意义,而其所承载的社会价值(如记忆)再次确证了其意义。这便能解释诸如中国“希望工程”的标志图像“大眼睛姑娘”、叙利亚难民危机中的“海滩上的男孩”等图像可被视为知识的原因。

而经由算法生成的图像所整合的细节来自不同层级的数据和参数,其生产过程几乎不涉及认

知主体(即生成图像者)的有意识的认知整合动作。但是,这并不意味着这类图像无法成为知识。实际上,传统摄影图像意义的起点始于拍摄过程,而生成图像意义的起点则是将其传播之后。

研究指出,基于提示词生成的图像在点击率等量化指标上表现优异,根据平台的流量分配规律,它们便有可能在算法分发中占据统治地位<sup>⑪</sup>。形成平台上可辨识的图像样板之后,这些图像不仅成为审美对象,更成为衡量美学价值的标准答案,一种确定性的知识,其所承载的意义是一条关于美的宣称——这就是美的。最终,AI生成的图像及其样板化审美甚至可能演变为一种权威知识。

这种转型也使得人们在社交平台上的学习行为发生了转向:他们不再学习如何审美,而是学习如何调用关于审美的指令。这时,发现现实不再是人们生产图像的理由。个体在创作中的角色从感知的发现者转变为指令的编排者。

需要警惕的是,当图像样板被视为一种确定性的审美知识时,它同时也可能内化并放大技术底层的结构性偏见。研究发现,AI在视觉生成中存在系统性的窄化偏向,其产出的图像往往固化了特定的身份权力结构<sup>⑫</sup>。这种通过提示词激活的刻板印象,使得既有的视觉霸权得以通过标准化知识的形式,在用户交互中实现自动化复制,甚至能被未成年用户直观感知<sup>⑬</sup>。

综上,生成技术对于图像知识的影响可能在于,它不仅改变了关于审美的知识形式,更重塑了人们理解美、占有美的方式。为了更有针对性地剖析这一微观过程,本文开展了如下实证研究。

## 三、研究方法

随着智能技术深度介入现实,中国视觉分析的研究对象已拓展至人工智能生成的媒介产品。视觉分析的研究取向也逐渐从对表征层面的描述,转向对算法偏见与隐性权力等深层机制的剖

析<sup>⑧</sup>。类似实证研究中呈现出显著的计算转向,但计算逻辑倾向于将复杂的视觉修辞降维为离散的物理变量,这种还原主义视角容易剥离图像生成的具体社会语境<sup>⑨</sup>。

本研究采取内容分析与案例分析相结合的方法。内容分析旨在对AI人像教程进行系统化的类目建构,通过对视觉特征、生成指令与参数逻辑等内容的分类归纳,揭示平台场域内审美生产的共性特征。案例分析则聚焦于对典型样本的深描,旨在透视图像生成过程中技术逻辑与个体经验的交互细节,阐释审美生产向程序化知识转型的微观过程。

#### (一) 研究设计与样本筛选

本研究选择小红书作为研究平台,主要基于该平台在AI绘画、AI摄影等视觉领域具有较高的内容集中度。研究聚焦于Nano Banana系列模型生成的图像,该模型是基于Gemini大语言模型架构开发的AI图像生成与编辑工具。因其生成的人像效果高度契合社交媒体流行审美,在AI摄影圈层中具有广泛的影响力。

数据抓取阶段,研究者以“Banana 写真”“Banana 摄影”“Banana 一致性”为核心关键词进行检索。前两个关键词偏向审美维度,第三个关键词偏向技术维度。样本量设定为300个帖子。抓取数据时,研究综合了“最新”与“最热”两种排序逻辑:80%的“最新”样本用于捕捉AI技术快速迭代下的审美动态与个体实践的多样性;20%的“最热”样本用于观察在平台算法中获得高频互动的审美样板。

#### (二) 数据采集与预处理

研究最终取得306条帖子的详细元数据。数据内容涵盖:帖子标题、正文文本/提示词文本、标签、互动数据(点赞、收藏、评论)、帖子链接以及帖子包含的所有原始图片(共1000余张)。经整理,共获得可供编码的图像样本790张。数据

分析前,研究者对文本内容进行结构化预处理。提示词部分被细分为技术参数与感性描述语,便于观察审美意图如何转化为具体的算法指令。

#### (三) 编码类目构建与过程

进行图片编码前,研究构建了包含视觉表征、摄影知识、本真性表演及深层质性阐释四个维度的编码表。视觉表征维度探讨图像如何建立数字美学标准,如完美合成感、超真实质感等;摄影技术维度考察人类摄影专业知识(如构图、光影布局)在AI生成中的体现;本真性表演维度探讨AI人像的凝视关系、情感表露与所处的空间语境等;深层质性阐释维度旨在挖掘图像背后的社会心理、审美异质性及情感投射。

编码过程遵循“视觉先行”的序贯原则。第一阶段,编码员仅观察图像,在不接触文字的情况下完成物理指标编码,以保证审美直觉不受干扰。第二阶段,编码员开启提示词与文案,探讨文字意图与图像结果的偏离程度。两阶段都允许复选(如照片的人像主体可能既像处在自然光照条件下,又像处于算法生成的光照条件下,此时保留两个编码),以便保留质性阐释的开放性和丰富度。

编码由两名有经验的编码员使用NVivo15软件完成。首先随机抽取20组样本进行预编码与一致性校准。在针对歧义点进行辩论并达成标准对齐后,计算得出信度指标Cohen's Kappa值大于0.8,表明信度达标。随后,编码员完成全样本编码。在软件中,每个帖子被设定为一个案例,使图像编码、文本属性与互动数据实现关联。

## 四、研究发现

(一) 差异化的审美样板:个体差异下的结构趋同

根据曲凯音的研究,自辛亥革命以来的百年间,中国女性审美标准经历了从觉醒释放到多元并存的变迁。1911年至1949年,伴随妇女解放运

动,女性美获得空前释放,呈现出以胡蝶、阮玲玉为代表的温婉时尚与以林徽因、杨步伟为代表的知性气质并存的多元特征;1950年至1976年间,审美标准趋于政治化与男性化,以邢燕子“铁姑娘”形象为典型,强调劳动奉献,而将女性温婉气质异化与遮蔽;进入1977年至2000年的改革开放时期,审美回归理性,女性通过“幸子衫”、红裙子等流行服饰大胆追美,气质得以释放张扬;步入21世纪后,审美标准走向全面多元化,性感美、知性美与中性美并行崛起,女性审美从被动迎合男性转向主动的主体表达<sup>⑥</sup>。

本研究对790张人像图像样本的物理特征进行内容分析发现,人工智能人像生产呈现出一种“差异化审美样板”的景观。

首先,图像空间布局呈现出极强的稳定性,确保观察者与被摄主体之间维持较为固定的辨认距离。在790张样本中,采用“平视”视角的图像占比达87.8%,采用“中心构图”的比例为81.2%。由于样本多为社交媒体写真,景别高度集中于近景与中景。这种空间关系的标准化,使得不同生成图像在视觉心理层面产生了一种相似的社交接近感。

其次,在面部细节与皮肤质感的处理上,82.2%的样本呈现出明显的“奶油肌、陶瓷感、无毛孔”特征,皮肤纹理呈现非生物性的平滑质感。89.7%的图像被修正为“无死角脸型”,91.1%的样本对毛发、睫毛等细节进行了超越物理相机极限的根根分明式刻画。这些数据不仅体现了算法对人脸结构的理想化重塑,也呼应了加洛韦的论述,即“计算的视觉理所当然地认为,对象与世间存有可以被全方位观看”<sup>⑦</sup>。这种对人脸面貌的全方位修正,呼应了大众人像审美从追求自然美到偏向量化美、精准美的转型趋势<sup>⑧</sup>。高频出现的暖橙调与多粉调妆面色调则在很大程度上统一了画面的主要氛围,从而在感官层面构建起高度一致且易于识别的视觉模态。

再次,在图像的光影布局上,“物理指向光”

(52.5%)与“算法全方位发光”(54.3%)的分布数量基本持平。在实际编码中,大量样本同时具备这两个特征。这意味着AI在保留明确光源方向的同时,普遍对阴影区域进行了算法层面的补偿,使人像面部(如鼻翼两侧、眼窝等部位)几乎不留任何深重阴影。这种光影处理方式弱化了现实环境的复杂性,使人物看起来更像是被置于一个受控的、理想化的照明体系中。算法微光的广泛存在,进一步强化了图像的拟像属性。它使得图像在感官上呈现出一种奇怪的雷同感:几乎每个人物都处于某种最优化的光场中。这种过于精致的处理在抹除瑕疵的同时,也消解了影像应有的鲜活性与不可预期性,并从根本上抹平了图像本应承载的时空差异,使其演变为脱离具体现实语境的视觉符号。其呈现结果正如马诺维奇谈及照相现实主义三维计算机动画时所指出的,这之间存在一个悖论:“其图像并不逊色于传统摄影,它们是完美的真实,但又过于真实了”<sup>⑨</sup>。

本研究将实证样本呈现出的种种特征描述为“差异化的审美样板”。这里的样板化并非指简单重复,而是意在指出,虽然图像保留了用户原始照片的个体辨识度,但标准化的美学加工体系又将这些辨识度抹平。即便存在不同的审美风格(如纯欲风、日杂感),但风格间的差异只能从着装风格、布景等较表面的元素捕捉。在这些看似迥异的风格中,几乎始终存在着一些共性的审美结构。

在这种模式下,主体的独特性在很大程度上经历了去语境化处理,转化为一种符合平台审美公约数的符号资产。这种个体面孔相异、但视觉体验趋同的现象,反映了个体审美直觉向算法统计逻辑的收敛,一定程度上标志着审美生产从感性的经验表达转向了对预设样板的参数化拟合。

(二)审美的参数化转译:提示词作为审美—技术中介

1.语义的模块化拆解:从感性词汇到技术参数



风写真操作指南:Step 1:打开你的「香蕉模型」; Step 2:上传一张面部和姿态清晰的高清半身照; Step 3:直接复制粘贴下方提示词!”

这种叙事结构取消了创作者在拍摄现场的具身性思考,审美产出的成败被简化为对既有步骤的执行力。特征词表中,“生成”(279次)、“指令”(64次)、“操作”(22次)与“速成”(16次)的高频共现,进一步佐证了这种生产模式对效率的追求优于对意境的推敲。

有趣的是,部分帖文尝试在标准化的结构中嵌入感性话语,为枯燥的技术指令披上一层“情感伪装”,以中和教程文本过于厚重的工具属性。以标题为“转运锦鲤2026新年写真提示词分享”的帖子为例,在全篇607字的文本中,核心提示词占据了466字,构成了绝对的内容主体。但在结构安排上,帖子开头利用“金沙摇曳,红妆点缀,是独属于中国新年的浪漫”这一视觉化描述建立了特定的审美语境,结尾则辅以“新的一年,要明媚,要张扬,要光芒万丈”等呼吁性、情感化的祝愿。这实质上是试图利用文学性话语来对冲程序化流程的机械感,却因感性表述与技术逻辑之间存在极大差异,导致话语层面显示出严重失调的感觉。我们从中能够感受到创作者在试图弥合算法逻辑与人文体验,但这种刻意的、流于形式的拼贴反而凸显了技术指令与文学修辞之间的巨大差异。

其次,教程文本通过建构“审美公式”,将复杂的视觉风格转化为标准化的技术指令。在这种生产逻辑下,离散的提示词与参数组合替代了创作者的感性直觉。例如,部分博主在文案中用类似导引性的文字封装风格、情绪、场景等具体描述:“【抄作业时间——AI通用提示词】(这是封面图的咒语,可以直接复制使用……)。”“抄作业”可以代表不同类型用户间在此类平台上形成的某种话语契约。风格描述则居于其中,随时等待被调用。由于用户能够直观地对封面图的风格

下模糊判断,即决定“我”是不是喜欢或想要这种风格,他们并不需要细致分析或揣摩提示词中的具体风格描述内容或参数,也不再需要理解“电影感”或“高级感”背后的光影原理。

再次,教程文本在话语体系上表现出对校验指标的过度依赖。大量样本显示,原本抽象的审美评价,正逐渐变成对图像各项技术参数的“达标检查”。如某帖文在结尾处强调:“今天不绕弯子,直接上一个核心结论:AIGC数字模特进入3.0时代的唯一标志,就是‘绝对的人物一致性’被彻底打通了。这意味着,你可以像导演指挥真人演员一样,去精准调度一个‘数字生命体’。”这种话语将“美”降维为“精准还原”与“像素级对齐”。由此,图像的优劣不再取决于其传递的情感深度或艺术意蕴,而在于其技术参数是否符合预设的标准,个体的审美直觉在标准化的参数评估面前被边缘化。这种逻辑将审美评价从感性的艺术体验,简化为关于算法达成度的技术核验。

尤其值得注意的是,上述这类高度标准化的叙事特征揭示了人机交互逻辑对社会传播话语的深层重塑。即创作者在面向受众传递审美经验时,不自觉地借用了与算法对话的指令思维,使得原本用于人际沟通的文字,在逻辑与语法上正无形中向人机指令深度靠拢。

### 3.典型案例:

本节选取一个典型案例(见图3),旨在通过互文性解析,还原个体美学意图如何被收编进指令体系。

该案例生成的图像序列展现出极高的技术完成度与视觉工业水准。从图像质感上看,AI模型对物理世界的模拟已达到肉眼难辨虚实的程度,特别是皮肤纹理的细腻程度、毛孔的微小细节以及发丝在自然光下的光晕表现,均符合专业人像摄影的高标准。无论是“日系居家风”(见下图3左上角图片)中温暖、通透且带有轻微胶片

让 Nano banana 帮你拍百变写真! (附提示词)

最近在测怎么用精准的提示词搞图  
我的逻辑一般是, 找到你想要参考图片  
然后用 AI 反推提示词, 以上用的是千问 or Gemini  
上传一张自己的正面照  
生成的模型用了 Nanobanana 和 seedream4.0  
这两个比较好用, 其他我用起来傻傻的...  
不得不说不成片的人物质感和皮肤纹理真的真假难辨

日系居家风 (我个人比较喜欢的):  
风格锚点: 日系胶片=暖调柔光+颗粒感+自然过曝 (区别于韩系高饱和、美式复古做旧);  
场景符号: 居家感 (沙发/地毯/书架)、自然光 (纱帘透光、窗光影子)、生活道具 (老相机/书堆/水果);  
人物状态: 松弛互动 (自拍/躺卧/搞怪)+ 日常穿搭 (白T+牛仔)+ 细节妆容 (雀斑、小耳饰);  
光影色彩: 柔和暖光、低对比度、浅景深 (突出主体)、轻微颗粒 (模拟胶片质感)。

红色背景那一套我也觉得很酷:  
神态: 眼神清澈、专注, 带着一丝慵懒和疏离感, 嘴角微抿, 整体传递出一种沉浸在自己世界里的的情绪。动作: 她以蹲姿坐在椅子上, 一只手把耳机贴近耳边, 另一只手拿着播放器, 动作自然放松, 强化了“听音乐”的沉浸感。  
服饰: 她穿了一件宽松的黑色长西装, 搭配黑色百褶裙, 脚上是黑色乐福鞋配白色蕾丝袜, 形成强烈的黑白对比。  
颜色: 她的妆面用了偏暖的色调, 腮红和口红让整体有一点复古感, 与黑色服装形成鲜明对比。色彩与构图: 色彩: 背景是大面积的红色, 与人物的黑色服装形成高对比度的视觉冲击, 让主体非常突出。构图: 她的位置偏左, 画面留有一定的呼吸空间; 整体为竖构图, 让人物更有张力; 红色背景简洁干净, 突出人物的情绪和风格。比例 3:4。

真的是0成本预制, 但是一般第一次生成出来的都傻傻的, 可以多次调教 2-3 次就正常了。感恩大模型让我又当清纯老妹又当酷拽学姐又当摩登女郎

#ai 写真 #AI 绘画有点东西 #创意写真 #日杂写真 #我和 AI 有话说 #Nanobanana #seedream4 #ai 关键词 #画报写真 #时尚写真



图3 案例展示(已做面部处理)

颗粒感的柔和氛围, 还是“红色背景”系列中(见图3右上角图片)高饱和度、高对比度的现代时尚感, 图像在光影分布和色彩平衡上都表现得较为成熟。这种高质量的视觉产出并非由于创作者具备深厚的摄影修养, 而是算法对大众审美公约数进行优化后的结果, 展现出一种数字原生的、理想化的“真实感”。

即便如此, 这一生成实践的深层逻辑仍是对特定审美样板的复刻与收编。博主通过明确的“风格锚点”和“场景符号”, 将复杂的艺术风格还原为一组可替换的离散标签。例如, “日系胶片”被降维为“暖调柔光+颗粒感+自然过曝”的等式, 这使得审美表达从一种整体性的生命体验收缩为参数的精确对齐。即便图像中的人物面孔保留了用户的身份特征, 但支撑人物神态的疏离感或松弛感实际上是由算法预设的情绪模组驱动的。

从生产过程看, 该案例清晰地揭示了审美生产如何从感性探索演变为指令执行。博主的创作逻辑并非源于具身的观察, 而是基于反推提示词这一技术化的逆向工程。在此路径下, 原本抽象的审美意图被转译为由道具、光位、服装颜色等构成的清单。博主提到的“调教”过程, 实质上是

用户利用指令对算法生成的随机性进行规训, 直至产出结果符合其预设的“正常”样板。这种“0成本预制”的便捷性, 实际上是以放弃个体审美探索的主体性为代价, 换取对算法权力体系下标准化美学的调用权。美感在此处不再是创造的终点, 而变成了对特定代码路径成功调用的产物。

五、结论与讨论

通过对社交平台小红书内帖子文本及生成图像的实证考查, 研究发现生成式人工智能背景下的人像生产已由对现实世界的索引记录转向了基于参数对齐的样板化生成逻辑。尽管技术手段保留了用户的面部识别特征, 但个体元素已被系统性地纳入高度标准化的视觉框架与算法模组之中。视觉生产的核心目标不再仅是捕获独特的物质瞬间, 而是通过对美学公约数的复刻来确保产出结果的稳定性。这表明, 个体在获得即时美感体验的同时, 其审美判断的自主过程正受到生成算法体系的深度介入。

在知识生产维度, 研究揭示了提示词作为审美—技术知识转译中介的关键作用。作为集成了美学意向、摄影技艺与算法逻辑的综合性中介,

提示词将原本依赖主体感官经验的审美探索,简化为一套功能导向且路径明确的指令系统。提示词不仅作为生产工具存在,更作为人机协作中负责审美翻译的媒介,通过将复杂的直觉经验降维来预设视觉表达的范畴与方式。上述发现为进一步讨论社交媒体环境下的图像民主化和知识生产问题提供了实证基础。

### (一)新视觉政体下的图像民主化错觉

在剖析生成式人像的生产逻辑时,可以观察到一种新型视觉政体正在社交平台中成型。人工智能时代的新视觉政体,其核心特征是将传统的感性审美感知,转化为对一系列等待平台、图像生产者、图像观看者等多方行动者进行达标核验的技术指标。在此政体下,图像不再作为对现实的见证,而是表现为对预设美学标准的拟合程度。观看也不再是与物质世界的偶然遭遇,因为新视觉政体不仅在生产实践中为视觉权力的重新分配提供了技术背景,更在认知层面重塑了“如何看”的社会规则<sup>⑧</sup>。

随着图像生产的门槛降低,传统的视觉霸权在技术赋能下表现出了消解迹象。传统上,视觉生产是由专业摄影师和媒体机构垄断的特权领域<sup>⑨</sup>。如 Midjourney 等生成式模型的出现赋予了缺乏专业训练的普通用户一种“生成性视觉公民身份”,使其能够通过文本提示参与视觉文化的构建与公共表达<sup>⑩</sup>。这在一定程度上促成了大众层面的生产力释放。然而,由于生产过程深度依赖平台基础设施及算法模组,视觉权力的转移往往伴随着更为隐蔽的规训逻辑。生产门槛的降低与技术控制的集中,使图像民主化呈现出一种矛盾:大众获得了生产力,但审美主导权却被算法和平台收编了。

一方面,从政治经济学维度审视,人工智能工具的应用加深了视觉生产对平台基础设施的依附<sup>⑪</sup>。随着互联网接口向人工智能代理演进,人

机交互模式发生了结构性转变。用户由主动的视觉探索者转化为面对算法输出的接收者<sup>⑫</sup>。由于算法黑箱替代了创作者深层的视觉决策,用户无法直接掌控视觉标准。另一方面,在垂直整合的平台架构中,少数科技巨头通过控制大模型等基础设施占据枢纽位置,统摄着数据与视觉内容的流转<sup>⑬</sup>。更微观地,平台算法通过预设参数对大众的视觉表达进行规范统一<sup>⑭</sup>。因此,AI 技术专业壁垒的消解看似赋予了公众基础生产力,但在决定图像生成结果的核心审美规则制定上,用户仍缺乏实质性的参与渠道<sup>⑮</sup>。

由此可见,当下的图像民主化在很大程度上只是一种“错觉”。表面上,技术赋予了用户随心所欲生成图像的自由。实际上,该自由始终处于算法和平台预设的路径之内,而且它被包装得更隐蔽、介入得更深层。正如研究发现所示,规训实则在图像生成的算法运作过程中就已完成。用户在调用技术接口时,其审美直觉早已被收编。因此,这种图像民主化的本质是在技术划定的边界内进行的有序生产,技术的深度介入和便利掩盖了主体让渡审美权力的过程。

### (二)图像知识的再省思

在生成式人像的生产逻辑中,技术的干预已不仅限于视觉产出的表象,更触及了审美知识的深层问题。

首先是审美知识的窄化问题。在社交平台语境中,“氛围感”等词汇本应高度依赖于个体的具身感知,其内涵因人、因时而异,具有极强的流动性。但在人工智能的生成路径中,这些抽象意境被固化为一组特定的提示词后缀与算法参数。这种处理方式可能导致审美话语的语义萎缩。当大众习惯于以此类技术定义的标准来锚定审美时,语言对复杂意象的阐释力便随之削弱。感性语言经过长时间的标准化规训后,可能最终导致审美想象力的贫瘠。

更重要的是,传统的审美生产涉及跨越物理三维世界与图像二维美学的复杂互动,包括对光线的物理感知、镜头的透视控制以及创作者具身的感官参与。然而,在人工智能介入后,这一充满偶然性的探索过程被降格为单一的、文本层面的交互。二者在意义丰富度上极不对称。其结果之一便是机器拟合的审美与人类真实感性之间仍存在错位。研究亦证实,由于人工智能图像常带有感知上的非流利性与视觉不确定性,人类在潜意识层面依然偏好能唤起深层共情(如怀旧、自我映射)的人类画作<sup>⑤</sup>。这表明审美维度中某些感性经验仍抗拒被完全计算化。

尽管有学者试图提出“半灵韵”概念<sup>⑥</sup>,主张在人机协作中寻找本真性的残余,但这一折中方案在理论上存在张力。本雅明的“灵韵”严格依托于对象在特定时空中的独一无二性,而人工智能生成的图像是去中心化、去语境化且可无限变体的统计推断。二者之间恐怕很难达成真正的折中。

此外,审美样板的流行也可能造成新的知识危机。由于生成式人工智能的模型参数倾向于收敛至训练数据的统计平均值,即一种平滑、符合大众刻板印象的“完美”图像,用户为了追求生成效率与视觉快感,选择主动“从主体性中解放”,接纳模型提供的最优解<sup>⑦</sup>。在此过程中,用户对美的感知和描述能力便被限制在提示词库的特定维度内。这种现象构建了一种“虚假的接近性”<sup>⑧</sup>,用户误以为通过简单的提示词掌控了复杂的艺术创作,但其实既未接近真正的审美知识,也未在实质上理解其所调用的技术知识。

综上,在新的技术背景下,图像或许正从一种承载长久记忆与认知价值的知识载体,悄然质变为仅在当下生效的瞬时信息。不同于自然美所蕴含的“遥远”,“数字美的时间性是没有未来也没有过往的当下,它就在眼前”<sup>⑨</sup>。原本具备时间属性的视觉符号,在指尖划过屏幕的刹那间便完

成了它的全部使命,不再具有向未来延展的生命力。与此同时,那些作为审美中介的提示词,虽在形式上构成了某种新兴知识,却因其纯粹的功能向度而必将陷入“用后即忘”的宿命。这种知识缺乏温度,无法转化为可供反复思索的文化经验,最终使关于美的体悟仅仅剩下一串串冰冷的指令本身。当图像失去时间属性、指令取代意义的沉淀,人类的审美世界或许正面临着一种系统性的空心化。它就隐匿在每一次便捷的生成与消费之中,正无声地重塑着我们感知世界的方式。

本研究仍存在一定局限:首先,样本局限于大众消费审美场域,未来可转向专业艺术或先锋美学领域,探讨不同文化语境下的逻辑差异。其次,研究多聚焦于产出结果,对用户的主观创作体验挖掘不足。后续研究可通过访谈或田野调查,观察用户如何利用对抗性提示词或技术偶发性,在算法规训中重塑主体性。最后,受限於技术的高速迭代,本研究仅记录了特定阶段的视觉切片。关于生成式技术对现实身体伦理的深层塑造,仍需开展长期的跨周期追踪。

(丁依然:北京师范大学计算传播学研究中心、北京师范大学新闻传播学院讲师;董晨宇:中国人民大学新闻学院副教授;颜晓鹏:北京师范大学计算传播学研究中心、北京师范大学新闻传播学院博士研究生)

#### 注释:

①②⑩[德]瓦尔特·本雅明:《机械复制时代的艺术作品:本雅明精选集》,庄仲黎译,杭州:浙江大学出版社,2025年,第7-10页、第149页、第11-12页、第154页。

③Peirce, C. S. Collected papers of Charles Sanders Peirce (C. Hartshorne & P. Weiss, Eds.; Vols. 1 - 8). Cambridge: Harvard University Press, 1931 - 1958.

④Barthes, R. Camera lucida: Reflections on photography (R. Howard, Trans.). New York: Hill and Wang, 1981.

⑤[美]苏珊·桑塔格:《论摄影》,黄灿然译,上海:上海译文出版社,2012年,第2页。

⑥[美]亚历山大·R·加洛韦:《不可计算:漫长数字时

代的的游戏与政治》，李逸帆译，南京：南京大学出版社，2025年，第59页。

⑦⑧⑨⑩⑪[俄]列夫·马诺维奇：《新媒体的语言》，车琳译，贵阳：贵州人民出版社，2020年，第222-240页、第27-32页、第30页、第32、36页、第207页。

⑫Jurgenson, Nathan. *The Social Photo: On Photography and Social Media*, New York: Verso Books, 2019, pp.1-45.

⑬马丁·杰：《现代性的视觉政体(1988)》，唐宏峰主编：《现代性的视觉政体：视觉现代性读本》，郑州：河南大学出版社，2018年，第76-100页。

⑭⑮Nieborg, D. B. & Poell, T. "The platformization of cultural production: Theorizing the contingent cultural commodity." *New Media & Society*, vol. 20, no. 11, 2018, pp.4275-4292.

⑯Gillespie, T., 2014. *The relevance of algorithms. Media Technologies: Essays on Communication, Materiality, and Society*, no.167, 2014.

⑰Duffy, B. E. (Not) Getting Paid to Do What You Love, New Haven, CT: Yale University Press, 2017, pp.1-45.

⑱Trillò, Tommaso. "PoV: You are reading an academic article." *The memetic performance of affiliation in TikTok's platform vernacular.* *New Media & Society*, vol. 28, no. 1, 2026, pp.380-400.

⑲Trillò, T., Hallinan, B. & Shifman, L. "A typology of social media rituals." *Journal of Computer-Mediated Communication*, vol. 27, no. 4 (2022): zmac011.

⑳①Polanyi, M. *Personal knowledge: Towards a post-critical philosophy*. Chicago: University of Chicago Press, 1958.

㉑Laba, N. "Beyond magic: Prompting for style as affordance actualization in visual generative media", *New Media & Society*, vol.28, no.1, 2026, pp.148-168.

㉒⑳⑳ Hartmann, J., Exner, Y. & Domdey, S. "The power of generative marketing: Can generative AI create superhuman visual marketing content?", *International Journal of Research in Marketing*, vol.42, no.1, 2025, pp.13-31.

㉓Papa, V. & Theodosiou, Z. "Visual news values in the age of AI: Exploring newsworthiness in generative image creation", *Journalism*, 2025, 14648849261418348.

㉔党明辉、马林、李翼帆：《AI生成战争报道图像的视觉框架偏向——基于人工智能驱动平台NewsGPT中新闻报道图片的计算视觉分析》，《当代传播》2025年第6期，第34-41+48页。

㉕Vartiainen, H., Kahila, J., Tedre, M. et al. "Enhancing children's understanding of algorithmic biases in and with text-to-image generative AI", *New Media & Society*, vol.27, no.9, 2025, pp.5342-5368.

㉖王彦博、任吴炯、姚俊臣：《符号权力的“智能再生”：中文AI大模型生成记者图像分析》，《新闻界》2025

年第9期，第74-86页。

㉗周葆华、王风范：《计算传播研究中的计算机视觉分析：基于B站政务视频的研究案例与讨论》，《青年记者》2022年第20期，第25-28页。

㉘曲凯音：《百年中国女性审美价值标准变迁的文化评析》，《山西师大学报(社会科学版)》2012年第39卷第2期，第121-125页。

㉙[美]亚历山大·R·加洛韦：《不可计算：漫长数字时代的的游戏与政治》，李逸帆译，南京：南京大学出版社，2025年，第59-60页。

㉚张晓梅：《中国美：中国第一本美女标准粉皮书》，北京：新华出版社，2005年。

㉛秦兰珺：《编码日常：大众软件批判》，北京：文化艺术出版社，2023年，第188-189页。

㉜Alper, M. "War on Instagram: Framing conflict photojournalism with mobile photography apps", *New Media & Society*, vol.16, no.8, 2014, pp.1233-1248.

㉝Schwalbe, C. B., Silcock, B. W. & Candello, E. "Gatecheckers at the visual news stream: A new model for classic gatekeeping theory" *Journalism Practice*, vol.9, no.4, 2015, pp.465-483.

㉞⑬ Bak Herrie, M., Maleve, N. R., Philipsen, L. et al. "Democratization and generative AI image creation: aesthetics, citizenship, and practices", *AI & SOCIETY*, vol.40, no.5, 2025, pp.3495-3507.

㉟王金鹏、曼纽尔·卡斯特：《从搜索引擎到生成式AI：互联网接口权力的演变与社会重构——访谈曼纽尔·卡斯特教授》，《全球传媒学刊》2025年第12卷第4期，第213-224页。

㊱van Dijck, J. "Seeing the forest for the trees: Visualizing platformization and its governance", *New Media & Society*, vol.23, no.9, 2021, pp.2801-2819.

㊲Popiel, P. & Vasudevan, K. "Platform frictions, platform power, and the politics of platformization", *Information, Communication & Society*, vol.27, no.10, 2024, pp.1867-1883.

㊳Samo, A. & Highhouse, S. "Artificial intelligence and art: Identifying the aesthetic judgment factors that distinguish human- and machine-generated artwork", *Psychology of Aesthetics, Creativity, and the Arts*, 2023.

㊴Salas Espasa, D. & Camacho, M. "From aura to semi-aura: reframing authenticity in AI-generated art—a systematic literature review", *AI & Society*, vol.40, no.8, 2025, pp.6727-6759.

㊵Park, S. "The work of art in the age of generative AI: aura, liberation, and democratization", *AI & Society*, vol.40, no.3, 2025, pp.1807-1816.

㊶Bar-Gil, O. "The transformation of artistic creation: from Benjamin's reproduction to AI generation", *AI & SOCIETY*, 2025, pp.1-15.

㊷[德]韩炳哲：《美的救赎》，关红玉译，北京：中信出版社，2019年，第34页。

(责任编辑：王珏)