

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2025 年 9 月 4 日 (04.09.2025)



(10) 国际公布号
WO 2025/179758 A1

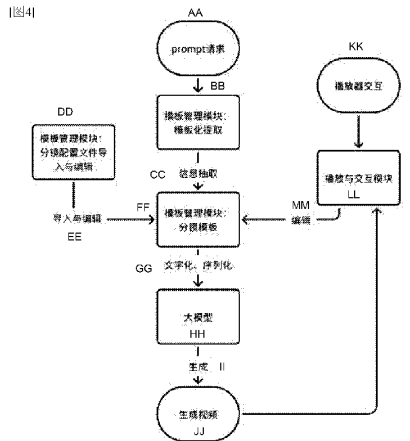
- (51) 国际专利分类号:
G06T 11/60 (2006.01) **H04N 21/81** (2011.01)
G06F 16/738 (2019.01)
- (21) 国际申请号: PCT/CN2024/106325
- (22) 国际申请日: 2024 年 7 月 19 日 (19.07.2024)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202410233999.8 2024年3月1日 (01.03.2024) CN
- (72) 发明人: 及
- (71) 申请人: 余音(YU, Yin) [CN/CN]; 中国福建省福州市鼓楼区铜盘路219号7座302 350000 (CN)。
- (72) 发明人: 杨子言(YANG, Ziyang); 中国福建省福州市鼓楼区八一七北路189号安泰中心A区4座707单元 350000 (CN)。
- (74) 代理人: 福州市京华专利代理事务所(普通合伙) (FU ZHOU JING HUA PATENT AGENCY

(GP.)); 中国福建省福州市鼓楼区软件大道89号福州软件园F区2号楼大鹏智慧中心楼层标号7层701单元-1 350003 (CN)。

- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO,

(54) Title: INTELLIGENT PICTURE OR VIDEO GENERATION METHOD AND SYSTEM FOR LARGE MODEL INTERACTION

(54) 发明名称: 一种大模型交互的智能图片或视频的生成方法和系统



AA Prompt request
BB Template management module: template-based extraction
CC Information extraction
DD Template management module: storyboard configuration file import and editing
EE Import and edit
FF Template management module: storyboard template
GG Text conversion and serialization
HH Large model
II Generate
JJ Generate a video
KK Player interaction
LL Playing and interaction module
MM Edit

(57) Abstract: Provided in the present invention are an intelligent picture or video generation method and system for large model interaction. The method comprises a picture or video generation process: S11, receiving an original prompt request from a user, extracting description information from the original prompt request, and filling same into a storyboard template inside a system; alternatively, directly importing the original prompt request and a storyboard template outside the system, and mapping the storyboard template outside the system into the storyboard template inside the system; S12, serializing and textualizing the content in the storyboard template, so as to obtain at least one prompt request segment, wherein each prompt request segment corresponds to one picture or one storyboard video; and S13, submitting to a large model the prompt request segments one by one or in batches, and the large model generating picture or video content and returning same. In the present invention, a systematic storyboard template is constructed and provided for a user, thereby realizing systematic and structured editing and management; moreover, batch generation can be supported, and interaction modes are rich and diverse, thereby improving the interaction experience of the user.

[见续页]

PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 本发明提供一种大模型交互的智能图片或视频的生成方法和系统, 方法包括图片或视频的生成过程: **S11**、接收用户的原始prompt请求, 抽取所述原始prompt请求中的描述信息并填充到系统内部的分镜模板中; 或直接导入原始prompt请求和系统外部的分镜模板, 将系统外部的分镜模板映射到系统内部的分镜模板中; **S12**、序列化、文字化所述分镜模板中的内容, 得到至少一个prompt请求分段, 每个prompt请求分段对应一图片或分镜视频; **S13**、将prompt请求分段逐条或批量提交给大模型, 由大模型生成图片或视频内容后返回。本发明为用户构建提供系统化分镜模板, 从而实现系统化、结构化编辑与管理, 同时可支持批量化生成, 交互方式丰富多样, 提升用户交互体验。