

(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



(10) 国际公布号
WO 2025/026043 A1

(43) 国际公布日
2025 年 2 月 6 日 (06.02.2025)

(51) 国际专利分类号:
G06T 13/40 (2011.01)

(21) 国际申请号: PCT/CN2024/105266

(22) 国际申请日: 2024 年 7 月 12 日 (12.07.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202310957070.5 2023 年 7 月 31 日 (31.07.2023) CN

(71) 申请人: 支付宝 (杭州) 信息技术有限公司 (ALIPAY (HANGZHOU) INFORMATION TECHNOLOGY CO., LTD.) [CN/CN]; 中国浙江省杭州市西湖区西溪路 556 号 8 层 B 段 801-11, Zhejiang 310000 (CN)。

(72) 发明人: 朱帅 (ZHU, Shuai); 中国浙江省杭州市西湖区西溪路 556 号 8 层 B 段 801-11, Zhejiang 310000 (CN)。 杨明晖 (YANG, Minghui); 中国浙江省杭州市西湖区西溪路 556 号 8 层 B 段

801-11, Zhejiang 310000 (CN)。 曹雨(CAO, Yu); 中国浙江省杭州市西湖区西溪路 556 号 8 层 B 段 801-11, Zhejiang 310000 (CN)。

(74) 代理人: 北京亿腾知识产权代理事务所 (普通合伙) (E-TONE INTELLECTUAL PROPERTY FIRM (GENERAL PARTNERSHIP)); 中国北京市海淀区中关村紫金数码园 3 号楼 707, Beijing 100190 (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。

(54) Title: DIGITAL HUMAN ACTION ORCHESTRATION METHOD AND DEVICE

(54) 发明名称: 一种数字人动作编排方法及装置

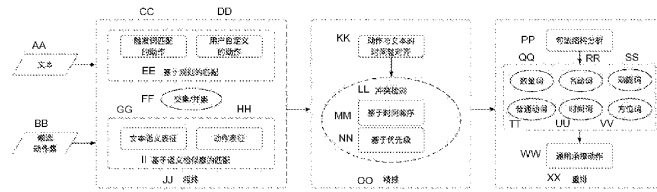


图 1

AA	Text	MM	Based on a time sequence
BB	Candidate action set	NN	Based on a priority
CC	An action matching a trigger word	OO	Fine orchestration
DD	User-defined action	PP	Syntactic structure analysis
EE	Rule-based matching	QQ	Quantifier
FF	Intersection/union	RR	Nominal verb
GG	Text semantic representation	SS	Converb
HH	Action representation	TT	Ordinary verb
II	Semantic similarity-based matching	UU	Temporal word
JJ	Coarse orchestration	VV	Noun of locality
KK	Time axis alignment between the action and the text	WW	General connecting action
LL	Conflict detection	XX	Re-orchestration

(57) Abstract: Embodiments of the present description relate to a digital human action orchestration method and device. The method comprises: inputting a first word sequence corresponding to a target text, and a first action set, wherein the first action set comprises a plurality of actions, and any action has a preset trigger word and a corresponding action duration; then, on the basis of the trigger words corresponding to the actions, matching words in the first word sequence with the actions in the first action set to obtain a first matching result, wherein the first matching result comprises alternative action sets corresponding to the words; and on the basis of the positions of the words in the target text and the action durations of the actions, determining, from the first matching result, a group of action sequences which do not overlap in a time dimension so as to obtain a first action sequence, wherein the first action sequence is used for instructing a digital human to make corresponding actions when broadcasting the target text.

[见续页]

(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, 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: 本说明书实施例涉及一种数字人动作编排方法及装置, 所述方法的输入包括目标文本对应的第一词语序列和第一动作集合, 第一动作集合中包含若干动作, 任一动作具有预设的触发词以及对应的动作持续时长; 然后, 根据动作对应的触发词, 将所述第一词语序列中的各个词语与所述第一动作集合的动作进行匹配, 得到第一匹配结果, 第一匹配结果中包含各个词语对应的备选动作集合; 然后, 根据各个词语在所述目标文本中的位置, 以及动作的动作持续时长, 从所述第一匹配结果中确定一组在时间维度上没有重叠的动作序列, 得到第一动作序列, 所述第一动作序列用于指示所述数字人在播报所述目标文本时做出相应的动作。