

(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 2025/066503 A1

(43) 国际公布日
2025 年 4 月 3 日 (03.04.2025)

- (51) 国际专利分类号:
A63F 13/67 (2014.01) **G06N 3/092** (2023.01)
G06N 3/0442 (2023.01)
- (21) 国际申请号: PCT/CN2024/107801
- (22) 国际申请日: 2024 年 7 月 26 日 (26.07.2024)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202311250926.1 2023年9月26日 (26.09.2023) CN
- (71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层 518057 (CN)。

- (72) 发明人: 杨敬文 (YANG, Jingwen); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层 518057 (CN)。胡南 (HU, Nan); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层 518057 (CN)。肖一驰 (XIAO, Yichi); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层 518057 (CN)。周昊 (ZHOU, Hao); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层 518057 (CN)。
- (74) 代理人: 深圳市深佳知识产权代理事务所(普通合伙) (SHENPAT INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市罗湖区南湖街道春风路庐山大厦B座18C2、18D、18E、18E2 518001 (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,

(54) Title: REINFORCEMENT LEARNING-BASED NPC COMBAT DECISION-MAKING METHOD AND RELATED PRODUCT

(54) 发明名称: 一种基于强化学习的NPC对战决策方法及相关产品

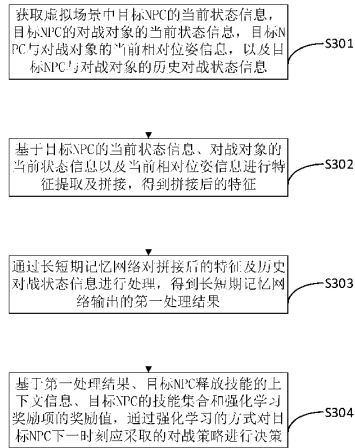


图 3

- S301 Acquire, in a virtual scene, the current state information of a target NPC, the current state information of a combat object opposing the target NPC, the current relative pose information of the target NPC and the combat object, and the historical combat state information of the target NPC versus the combat object
- S302 Perform feature extraction and concatenation on the current state information of the target NPC, the current state information of the combat object, and the current relative pose information, so as to obtain a concatenated feature
- S303 By means of a long short-term memory network, process the concatenated feature and the historical combat state information to obtain a first processing result output by the long short-term memory network
- S304 On the basis of the first processing result, contextual information of skill release of the target NPC, a skill set of the target NPC, and the bonus values of bonus items of reinforcement learning, by means of reinforcement learning, perform decision making on a combat strategy that should be taken by the target NPC at the next moment

(57) Abstract: Disclosed in the present application are a reinforcement learning-based NPC combat decision-making method and a related product. The method comprises: acquiring the current state information of a target NPC, the current state information of a combat object, the current relative pose information, and the historical combat state information of the target NPC versus the combat object, and processing the current state information of the target NPC, the current state information of the combat object and the current relative pose information, so as to obtain an concatenated feature; by means of a long short-term memory network, obtaining a first processing result output by the long short-term memory network; and, on the basis of the first processing result, contextual information of skill release of the target NPC, a skill set of the target NPC, and bonus items of reinforcement learning, by means of reinforcement learning, performing decision making on a combat strategy that should be taken by the target NPC at the next moment. According to the method

[见续页]

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, 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))。

provided in the present application, after the information required by reinforcement learning is acquired, NPC combat decisions are made by means of reinforcement learning without needing to establish a complex behavior tree, thereby reducing the labor cost.

(57) 摘要: 本申请公开一种基于强化学习的NPC对战决策方法及相关产品, 所述方法包括: 获取目标NPC的当前状态信息, 对战对象的当前状态信息, 当前相对位姿信息和目标NPC与所述对战对象的历史对战状态信息, 对目标NPC的当前状态信息、对战对象的当前状态信息和当前相对位姿信息进行处理得到拼接后的特征; 通过长短期记忆网络得到长短期记忆网络输出的第一处理结果; 基于第一处理结果、目标NPC释放技能的上下文信息、目标NPC的技能集合和强化学习的奖励项, 通过强化学习对目标NPC下一时刻采取的对战策略进行决策。本申请提供的方法在获取强化学习所需的信息后, 通过强化学习得到NPC对战决策, 无需建立复杂的行为树, 节约了人力成本。