

(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 2022/033534 A1

(43) 国际公布日
2022 年 2 月 17 日 (17.02.2022)

- (51) 国际专利分类号:
H04L 29/06 (2006.01) *G10L 25/27* (2013.01)
H04N 5/262 (2006.01)
- (21) 国际申请号: PCT/CN2021/112140
- (22) 国际申请日: 2021 年 8 月 11 日 (11.08.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202010806612.5 2020 年 8 月 12 日 (12.08.2020) CN
- (71) 申请人: 北京字节跳动网络技术有限公司 (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市石景山区实兴大街 30 号院 3 号楼 2 层 B-0035 房间, Beijing 100041 (CN)。
- (72) 发明人: 兰枫 (LAN, Feng); 中国北京市海淀区知春路 63 号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。

- (74) 代理人: 上海光栅知识产权代理有限公司 (LINKER IP LLC); 中国上海市长宁区延安西路 889 号 702 室, Shanghai 201109 (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, MK, MT, NL, NO, PL, PT,

(54) Title: METHOD FOR GENERATING TARGET VIDEO, APPARATUS, SERVER, AND MEDIUM

(54) 发明名称: 用于生成目标视频的方法、装置、服务器和介质

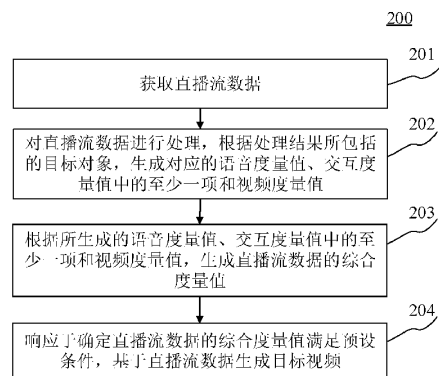


图 2

- 201 Obtain live broadcast stream data
- 202 Perform processing on live broadcast stream data, and generate at least one among corresponding voice metric value and interaction metric value, and corresponding video metric value according to target object included in processing result
- 203 Generate overall metric value for live broadcast stream data according to generated at least one among voice metric value and interaction metric value, and generated video metric value
- 204 In response to determining that comprehensive metric value for live broadcast stream data satisfies preset condition, generate target video on basis of live broadcast stream data

(57) Abstract: Disclosed in embodiments of the present application are a method for generating a target video, an apparatus, a server, and a medium. A specific embodiment of the method comprises: obtaining live broadcast stream data, wherein the live broadcast stream data comprises at least one among voice data and live broadcast interaction data, and video data; performing processing on the live broadcast stream data, and generating at least one among a corresponding voice metric value and interaction metric value, and a corresponding video metric value according to a target object included in a processing result; generating an overall metric value for the live broadcast stream data according to the generated metric values; and in response to determining that the comprehensive metric value for the live broadcast stream data satisfies a preset condition, generating a target video on the basis of the live broadcast stream data. The present embodiment implements automatic generation of a target video, performs an overall selection for a basis for generation of the target video from many aspects, such as video and at least one among voice and live broadcast interaction, and improves quality of the generated target video and generation efficiency.

[见续页]

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: 本申请实施例公开了用于生成目标视频的方法、装置、服务器和介质。该方法的一具体实施方式包括: 获取直播流数据, 其中, 该直播流数据包括语音数据、直播交互数据中的至少一项和视频数据; 对该直播流数据进行处理, 根据处理结果所包括的目标对象, 生成对应的语音度量值、交互度量值中的至少一项和视频度量值; 根据所生成的各种度量值, 生成该直播流数据的综合度量值; 响应于确定该直播流数据的综合度量值满足预设条件, 基于该直播流数据生成目标视频。该实施方式实现了目标视频的自动生成, 并且从语音、直播交互中的至少一项和视频等多方面综合选定目标视频的生成基础, 提高了所生成的目标视频的质量和生成效率。