

(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 2022/017451 A1

(43) 国际公布日
2022 年 1 月 27 日 (27.01.2022)

(51) 国际专利分类号:
H04N 5/262 (2006.01) *H04N 21/845* (2011.01)
H04N 5/76 (2006.01)

(21) 国际申请号: PCT/CN2021/107817

(22) 国际申请日: 2021 年 7 月 22 日 (22.07.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010728035.2 2020 年 7 月 23 日 (23.07.2020) 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: 曾衍(ZENG, Yan); 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)。 赵晨(ZHAO, Chen); 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)。 郑起凡(ZHENG, Qifan); 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)。 付平非(FU, Pingfei); 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: 北京市柳沈律师事务所(LIU, SHEN & ASSOCIATES); 中国北京市海淀区彩和坊路 10 号 1 号楼 10 层, Beijing 100080 (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: VIDEO DUBBING METHOD, DEVICE, APPARATUS, AND STORAGE MEDIUM

(54) 发明名称: 视频配音方法、装置、设备及存储介质

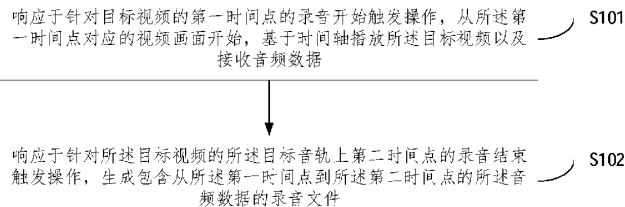


图 1

- S101 In response to receiving an audio recording start trigger operation for a first time point of a target video, play, starting from a video frame corresponding to the first time point, the target video on the basis of a timeline, and receive audio data
- S102 In response to receiving an audio recording end trigger operation for a second time point on a target audio track of the target video, generate an audio recording file containing the audio data from the first time point to the second time point

(57) Abstract: The present disclosure provides a video dubbing method, a device, an apparatus, and a storage medium. The method includes: upon receiving an audio recording start trigger operation for a first time point of a target video, playing, starting from a video frame corresponding to the first time point, the target video on the basis of a timeline, and receiving audio data; and upon receiving an audio recording end trigger operation for a second time point, generating an audio recording file, wherein the audio recording file has a linked relationship with a timeline of a video segment having a video frame corresponding to the first time point as a start frame, and a video frame corresponding to the second time point as an end frame. The present disclosure enables timeline-based audio recording to be performed while playing a target video, and in turn generates an audio recording file having a linked relationship with the timeline of a corresponding video segment, such that timelines of the audio recording file and the video segment do not need to be re-aligned for future operations, thereby facilitating an accurate video dubbing result, and preventing tedious operations of manual timeline alignment from causing inaccurate dubbing results.

[见续页]

WO 2022/017451 A1

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, 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: 本公开提供了一种视频配音方法、装置、设备及存储介质, 所述方法包括: 在接收针对目标视频的第一时间点的录音开始触发操作时, 从第一时间点对应的视频画面开始, 基于时间轴播放目标视频以及接收音频数据。当接收到针对第二时间点的录音结束触发操作时, 生成录音文件, 该录音文件与以第一时间点对应的视频画面为起始帧且以与第二时间点对应的视频画面为结束帧的视频片段的时间轴具有联动关系。本公开在播放目标视频的同时, 基于时间轴录音, 生成的录音文件与对应的视频片段的时间轴具有联动关系, 后续无需对录音文件与视频片段的时间轴重新匹配, 即可得到较准确的视频配音结果, 避免了因人工匹配时间轴导致的操作繁杂且配音结果不准确的问题。