

(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 2022/105693 A1

(43) 国际公布日
2022 年 5 月 27 日 (27.05.2022)

(51) 国际专利分类号:
G10L 13/02 (2013.01) *G10L 13/10* (2013.01)
G10L 13/08 (2013.01) *G10L 25/03* (2013.01)

(21) 国际申请号: PCT/CN2021/130459

(22) 国际申请日: 2021 年 11 月 12 日 (12.11.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202011309190.7 2020 年 11 月 20 日 (20.11.2020) CN

(71) 申请人: 北京猿力未来科技有限公司 (BEIJING YUANLI WEILAI SCIENCE AND TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市

朝阳区广顺南大街 8 号院 1 号楼 6 层 F01-03、05-10 单元, Beijing 100102 (CN)。

(72) 发明人: 王冬晓 (WANG, Dongxiao); 中国北京市朝阳区广顺南大街 8 号院 1 号楼 6 层 F01-03、05-10 单元, Beijing 100102 (CN)。 杨明祺 (YANG, Mingqi); 中国北京市朝阳区广顺南大街 8 号院 1 号楼 6 层 F01-03、05-10 单元, Beijing 100102 (CN)。 马楠 (MA, Nan); 中国北京市朝阳区广顺南大街 8 号院 1 号楼 6 层 F01-03、05-10 单元, Beijing 100102 (CN)。 夏龙 (XIA, Long); 中国北京市朝阳区广顺南大街 8 号院 1 号楼 6 层 F01-03、05-10 单元, Beijing 100102 (CN)。 郭常圳 (GUO, Changzhen); 中国北京市朝阳区广顺南大街 8 号院 1 号楼 6 层 F01-03、05-10 单元, Beijing 100102 (CN)。

(54) Title: SAMPLE GENERATION METHOD AND APPARATUS

(54) 发明名称: 样本生成方法及装置

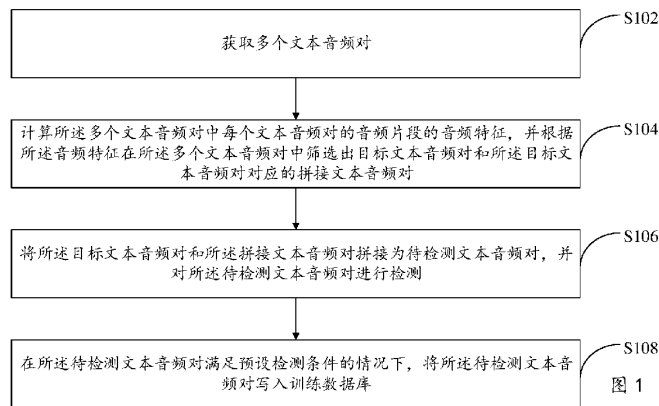


图 1

S102 Acquire a plurality of text-audio pairs
S104 Calculate an audio feature of an audio segment of each of the plurality of text-audio pairs, and select, by means of screening and according to the audio feature, a target text-audio pair and a splicing text-audio pair corresponding to the target text-audio pair from among the plurality of text-audio pairs
S106 Splice the target text-audio pair and the splicing text-audio pair into a text-audio pair to be tested, and test the text-audio pair to be tested
S108 When the text-audio pair to be tested meets a preset test condition, write the text-audio pair to be tested into a training database

(57) Abstract: Provided are a sample generation method and apparatus. The sample generation method comprises: acquiring a plurality of text-audio pairs, wherein each text-audio pair contains a text segment and an audio segment; calculating an audio feature of an audio segment of each of the plurality of text-audio pairs, and selecting, by means of screening and according to the audio feature, a target text-audio pair and a splicing text-audio pair corresponding to the target text-audio pair from among the plurality of text-audio pairs; splicing the target text-audio pair and the splicing text-audio pair into a text-audio pair to be tested, and testing the text-audio

[见续页]

(74) 代理人: 北京智信禾专利代理有限公司
(CHISPO ATTORNEYS AT LAW); 中国北京市东城区崇文门外大街8号院哈德门广场西塔9层901, Beijing 100062 (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, 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))。

pair to be tested; and when the text-audio pair to be tested meets a preset test condition, writing the text-audio pair to be tested into a training database.

(57) 摘要: 本说明书提供样本生成方法及装置, 其中所述样本生成方法包括: 获取多个文本音频对, 其中每个文本音频对中包含文本片段和音频片段; 计算所述多个文本音频对中每个文本音频对的音频片段的音频特征, 并根据所述音频特征在所述多个文本音频对中筛选出目标文本音频对和所述目标文本音频对对应的拼接文本音频对; 将所述目标文本音频对和所述拼接文本音频对拼接为待检测文本音频对, 并对所述待检测文本音频对进行检测; 在所述待检测文本音频对满足预设检测条件的情况下, 将所述待检测文本音频对写入训练数据库。