

(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 2025/156808 A1

(43) 国际公布日
2025 年 7 月 31 日 (31.07.2025)

(51) 国际专利分类号:
G10L 13/08 (2013.01) **G10L 21/013** (2013.01)
(21) 国际申请号: PCT/CN2024/134607
(22) 国际申请日: 2024 年 11 月 26 日 (26.11.2024)
(25) 申请语言: 中文
(26) 公布语言: 中文
(30) 优先权:
202410096064.X 2024 年 1 月 23 日 (23.01.2024) CN

(71) 申请人: 北京字跳网络技术有限公司
(**BEIJING ZITIAO NETWORK TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市海淀区紫金数码园4号楼2层0207 100190 (CN)。

(72) 发明人: 楚书廷(**CHU, Shuting**); 中国北京市朝阳区七圣中街12号院融中心B1小邮局 100028 (CN)。
陈漫佳(**CHEN, Manjia**); 中国北京市朝阳区七圣中街12号院融中心B1小邮局 100028 (CN)。

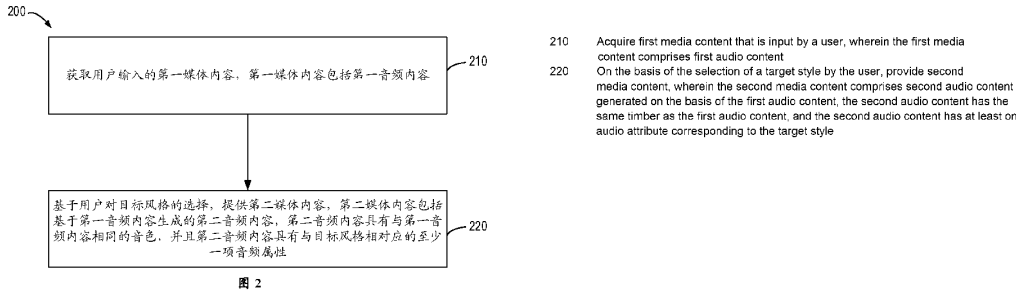
(74) 代理人: 北京世辉律师事务所(**SHIHUI PARTNERS**); 中国北京市朝阳区建国门外大街2号北京银泰中心C座42层 100022 (CN)。

(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。

(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,

(54) **Title:** AUDIO PROCESSING METHOD AND APPARATUS, AND DEVICE AND STORAGE MEDIUM

(54) 发明名称: 音频处理的方法、装置、设备和存储介质



(57) **Abstract:** An audio processing method and apparatus, and a device and a storage medium. The method comprises: acquiring first media content that is input by a user, wherein the first media content comprises first audio content (210); and on the basis of the selection of a target style by the user, providing second media content, wherein the second media content comprises second audio content generated on the basis of the first audio content, the second audio content has the same timber as the first audio content, and the second audio content has at least one audio attribute corresponding to the target style (220). The method can improve the voice changing effect while preserving the timber.

(57) 摘要: 一种音频处理的方法、装置、设备和存储介质, 方法包括: 获取用户输入的第一媒体内容, 第一媒体内容包括第一音频内容 (210); 基于用户对目标风格的选择, 提供第二媒体内容, 第二媒体内容包括基于第一音频内容生成的第二音频内容, 第二音频内容具有与第一音频内容相同的音色, 并且第二音频内容具有与目标风格相对应的至少一项音频属性 (220)。该方法能够在保留音色的基础上, 提高变声效果。

[见续页]

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))。