

(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 2022/116825 A1

(43) 国际公布日
2022 年 6 月 9 日 (09.06.2022)

- (51) 国际专利分类号:
G10L 21/0208 (2013.01) *G10L 25/30* (2013.01)
G10L 25/03 (2013.01) *G10L 25/51* (2013.01)
- (21) 国际申请号: PCT/CN2021/131187
- (22) 国际申请日: 2021 年 11 月 17 日 (17.11.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202011410814.4 2020 年 12 月 3 日 (03.12.2020) 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, Guangdong 518057 (CN)。
- (72) 发明人: 吴雯(WU, Wen); 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, Guangdong 518057 (CN)。 夏咸军(XIA, Xianjun); 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, Guangdong 518057 (CN)。
- (74) 代理人: 北京派特恩知识产权代理有限公司(CHINA PAT INTELLECTUAL PROPERTY OFFICE); 中国北京市海淀区海淀南路 21 号中关村知识产权大厦 B 座 2 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, 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,

(54) **Title:** ARTIFICIAL INTELLIGENCE-BASED AUDIO PROCESSING METHOD AND APPARATUS, ELECTRONIC DEVICE, COMPUTER READABLE STORAGE MEDIUM, AND COMPUTER PROGRAM PRODUCT

(54) 发明名称: 基于人工智能的音频处理方法、装置、电子设备、计算机可读存储介质及计算机程序产品

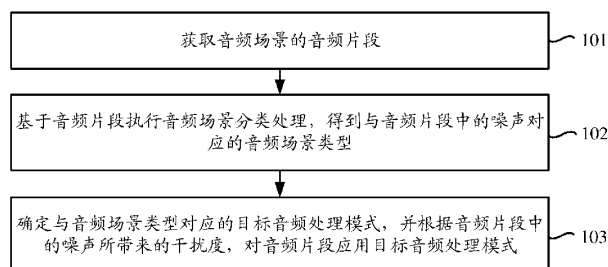


图 3

- 101 OBTAIN AN AUDIO CLIP OF AN AUDIO SCENE
- 102 PERFORM AUDIO SCENE CLASSIFICATION PROCESSING ON THE BASIS OF THE AUDIO CLIP TO OBTAIN AN AUDIO SCENE TYPE CORRESPONDING TO NOISE IN THE AUDIO CLIP
- 103 DETERMINE A TARGET AUDIO PROCESSING MODE CORRESPONDING TO THE AUDIO SCENE TYPE, AND APPLY THE TARGET AUDIO PROCESSING MODE TO THE AUDIO CLIP ACCORDING TO THE INTERFERENCE DEGREE BROUGHT BY THE NOISE IN THE AUDIO CLIP

(57) **Abstract:** The present application relates to cloud technology and artificial intelligence technology, and provides an artificial intelligence-based audio processing method and apparatus, an electronic device, a computer readable storage medium, and a computer program product. The method comprises: obtaining an audio clip of an audio scene, the audio clip comprising noise; performing audio scene classification processing on the basis of the audio clip to obtain an audio scene type corresponding to the noise in the audio clip; determining a target audio processing mode corresponding to the audio scene type, and applying the target audio processing mode to the audio clip of the audio scene according to the interference degree brought by the noise in the audio clip.

[见续页]

WO 2022/116825 A1

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: 本申请提供了一种基于人工智能的音频处理方法、装置、电子设备、计算机可读存储介质及计算机程序产品; 涉及云技术和人工智能技术; 方法包括: 获取音频场景的音频片段, 其中, 音频片段中包括噪声; 基于音频片段执行音频场景分类处理, 得到与音频片段中的噪声对应的音频场景类型; 确定与音频场景类型对应的目标音频处理模式, 并根据音频片段中的噪声所带来的干扰度, 对音频场景的音频片段应用目标音频处理模式。