

(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 2022/156655 A1

(43) 国际公布日
2022 年 7 月 28 日 (28.07.2022)

- (51) 国际专利分类号:
G10L 21/0272 (2013.01) *H04M 3/56* (2006.01)
- (21) 国际申请号: PCT/CN2022/072460
- (22) 国际申请日: 2022 年 1 月 18 日 (18.01.2022)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202110083388.6 2021 年 1 月 21 日 (21.01.2021) 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: 李娟娟(LI, Juanjuan); 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: 北京德琦知识产权代理有限公司 (DEQI INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区知春路 1 号学院国际大厦 7 层, Beijing 100083 (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: VOICE CALL CONTROL METHOD AND APPARATUS, AND COMPUTER-READABLE MEDIUM AND ELECTRONIC DEVICE

(54) 发明名称: 语音通话的控制方法、装置、计算机可读介质及电子设备

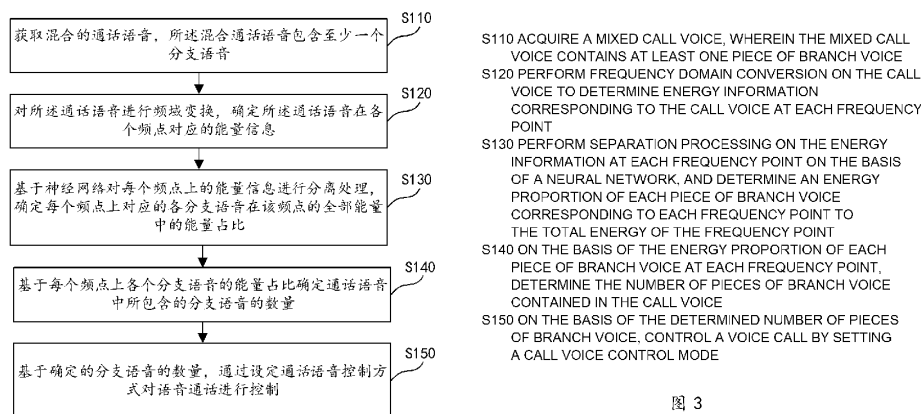


图 3

(57) Abstract: A voice call control method and apparatus (1400), and a computer-readable medium and an electronic device. The voice call control method comprises: acquiring a mixed call voice in real time during a cloud conference call process, wherein the mixed call voice contains at least one piece of branch voice (S110); performing frequency domain conversion on the call voice to determine energy information corresponding to the call voice at each frequency point in a frequency domain (S120); then performing separation processing on the energy information at each frequency point on the basis of a neural network, and determining an energy proportion of each piece of branch voice corresponding to each frequency point to the total energy of the frequency point (S130); on the basis of the energy proportion of each piece of branch voice at each frequency point, determining the number of pieces of branch voice contained in the call voice (S140); and finally, on the basis of the number of pieces of branch voice, controlling a voice call by setting a call voice control mode (S150).

[见续页]

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: 一种语音通话的控制方法、装置(1400)、计算机可读介质及电子设备。语音通话的控制方法包括: 在云会议通话进行过程中实时获取混合的通话语音, 混合通话语音包含至少一个分支语音(S110), 并对通话语音进行频域变换以确定通话语音在频域中各个频点对应的能量信息(S120), 之后基于神经网络对每个频点上的能量信息进行分离处理, 确定每个频点上对应的各个分支语音在频点的全部能量中的能量占比(S130), 以基于每个频点上各分支语音的能量占比确定通话语音中所包含的分支语音的数量(S140), 最后基于分支语音的数量, 通过设定通话语音控制方式对语音通话进行控制(S150)。