

(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 2022/213825 A1

(43) 国际公布日
2022 年 10 月 13 日 (13.10.2022)

(51) 国际专利分类号:
G10L 21/0224 (2013.01)

(21) 国际申请号: PCT/CN2022/083112

(22) 国际申请日: 2022 年 3 月 25 日 (25.03.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202110367186.4 2021年4月6日 (06.04.2021) CN

(71) 申请人: 京东科技控股股份有限公司
(JINGDONG TECHNOLOGY HOLDING CO., LTD.)
[CN/CN]; 中国北京市北京经济技术开发区科创十一街18号C座2层221室, Beijing 100176 (CN)。

(72) 发明人: 陈泽华(CHEN, Zehua); 中国北京市北京经济技术开发区科创十一街18号C座2层221室, Beijing 100176 (CN)。 吴俊仪(WU, Junyi); 中国北京市北京经济技术开发区科创十一街18号C座2层221室, Beijing 100176 (CN)。 蔡玉玉(CAI,

Yuyu); 中国北京市北京经济技术开发区科创十一街18号C座2层221室, Beijing 100176 (CN)。 雪巍(XUE, Wei); 中国北京市北京经济技术开发区科创十一街18号C座2层221室, Beijing 100176 (CN)。 杨帆(YANG, Fan); 中国北京市北京经济技术开发区科创十一街18号C座2层221室, Beijing 100176 (CN)。 丁国宏(DING, Guohong); 中国北京市北京经济技术开发区科创十一街18号C座2层221室, Beijing 100176 (CN)。 何晓冬(HE, Xiaodong); 中国北京市北京经济技术开发区科创十一街18号C座2层221室, Beijing 100176 (CN)。

(74) 代理人: 北京律智知识产权代理有限公司 (BEIJING INTELLEGAL INTELLECTUAL PROPERTY AGENT LTD.); 中国北京市朝阳区慧忠路5号B1605、B1606、B1607, Beijing 100101 (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: NEURAL NETWORK-BASED END-TO-END SPEECH ENHANCEMENT METHOD AND APPARATUS

(54) 发明名称: 基于神经网络的端到端语音增强方法、装置

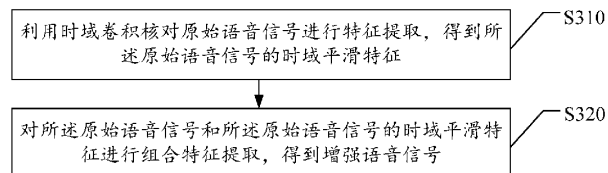


图 3

- S310 Extract a feature from an original speech signal by using a time-domain convolution kernel, so as to obtain a time-domain smoothing feature of the original speech signal
- S320 Perform combined feature extraction on the original speech signal and the time-domain smoothing feature of the original speech signal, so as to obtain an enhanced speech signal

(57) Abstract: A neural network-based end-to-end speech enhancement method and apparatus, a computer-readable storage medium, and a device. The method comprises: extracting a feature from an original speech signal by using a time-domain convolution kernel, so as to obtain a time-domain smoothing feature of the original speech signal (S310); and performing combined feature extraction on the original speech signal and the time-domain smoothing feature of the original speech signal, so as to obtain an enhanced speech signal (S320).

(57) 摘要: 一种基于神经网络的端到端语音增强方法、装置、计算机可读存储介质及设备, 该方法包括: 利用时域卷积核提取原始语音信号的特征, 得到原始语音信号的时域平滑特征 (S310); 对原始语音信号和原始语音信号的时域平滑特征进行组合特征提取, 得到增强语音信号 (S320)。

[见续页]

WO 2022/213825 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, JM, 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))。