

(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 2020/143263 A1

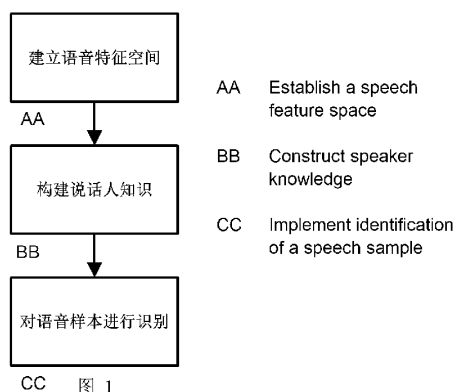
(43) 国际公布日
2020 年 7 月 16 日 (16.07.2020)

- (51) 国际专利分类号:
G10L 17/04 (2013.01) *G10L 17/02* (2013.01)
G10L 17/08 (2013.01)
- (21) 国际申请号: PCT/CN2019/111530
- (22) 国际申请日: 2019 年 10 月 16 日 (16.10.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910027145.3 2019 年 1 月 11 日 (11.01.2019) CN
- (71) 申请人: 华南理工大学 (SOUTH CHINA UNIVERSITY OF TECHNOLOGY) [CN/CN]; 中国广东省广州市天河区五山路 381 号, Guangdong 510640 (CN)。
- (72) 发明人: 贺前华 (HE, Qianhua); 中国广东省广州市天河区五山路 381 号, Guangdong 510640 (CN)。
吴克乾 (WU, Keqian); 中国广东省广州市天河

- 区五山路 381 号, Guangdong 510640 (CN)。 谢伟 (XIE, Wei); 中国广东省广州市天河区五山路 381 号, Guangdong 510640 (CN)。 庞文丰 (PANG, Wenfeng); 中国广东省广州市天河区五山路 381 号, Guangdong 510640 (CN)。
- (74) 代理人: 广州市华学知识产权代理有限公司 (GUANGZHOU HUAXUE INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国广东省广州市天河区五山路 381 号物资大楼首层, Guangdong 510640 (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, 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,

(54) Title: SPEAKER IDENTIFICATION METHOD BASED ON SPEECH SAMPLE FEATURE SPACE TRAJECTORY

(54) 发明名称: 一种基于语音样本特征空间轨迹的说话人识别方法



(57) Abstract: A speaker identification method based on a speech sample feature space trajectory, the method comprising: clustering unmarked speech data features to obtain a speech feature space representing an identifier subset; using a marked speech sample to implement speaker registration and obtaining motion trajectory information and distribution information of the speaker in the speech feature space; and using the distribution information of the speaker in the speech feature space and motion trajectory information of the speech sample to implement identification of a speech sample to be identified. Using the idea of speaker speech feature spatial positioning, the complexity of identifying the speaker is low, solving the problem of the high degree of complexity of GMM-UBM calculation; the speech feature space of a speaker in one language can serve as the speech feature space for identification of the speaker in another language, implementing data sharing.

[见续页]

PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, 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)。

根据细则4.17的声明:

— 发明人资格 (细则4.17(iv))

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

(57) 摘要: 一种基于语音样本特征空间轨迹的说话人识别方法, 方法包括对无标注语音数据特征进行聚类、得到语音特征空间表示: 标识子集合; 使用标注语音样本进行说话人注册, 得到说话人在语音特征空间中的分布信息和运动轨迹信息; 对待识别语音样本利用说话人语音特征空间分布信息以及语音样本的运动轨迹信息进行识别。采用说话人语音特征空间定位的思路, 说话人识别计算复杂度低, 解决了GMM-UBM计算复杂度高的问题; 并且一个语种的说话人语音特征空间可作为另一语种的说话人识别的语音特征空间, 实现了数据的共享。