

(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 2019/184517 A1

(43) 国际公布日
2019 年 10 月 3 日 (03.10.2019)

(51) 国际专利分类号:
G10L 17/02 (2013.01) *G10L 25/24* (2013.01)
G10L 25/18 (2013.01)

(21) 国际申请号: PCT/CN2018/125491

(22) 国际申请日: 2018 年 12 月 29 日 (29.12.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810273669.6 2018年3月29日 (29.03.2018) CN

(71) 申请人: 北京字节跳动网络技术有限公司 (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市

石景山区实兴大街30号院3号楼2层B-0035
房间, Beijing 100041 (CN)。

(72) 发明人: 李根(LI, Gen); 中国北京市海淀区知春路
63号中国卫星通信大厦今日头条小邮局, Beijing
100190 (CN)。 李磊(LI, Lei); 中国北京市海淀区
知春路63号中国卫星通信大厦今日头条小
邮局, Beijing 100190 (CN)。 何轶(HE, Yi); 中国
北京市海淀区知春路63号中国卫星通信大厦
今日头条小邮局, Beijing 100190 (CN)。

(74) 代理人: 北京中原华和知识产权代理有限
责任公司 (HUAHE IP LIMITED); 中国北京市
朝阳区北辰东路8号汇宾大厦A座909室徐
民、寿宁, 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,

(54) Title: AUDIO FINGERPRINT EXTRACTION METHOD AND DEVICE

(54) 发明名称: 一种音频指纹提取方法及装置

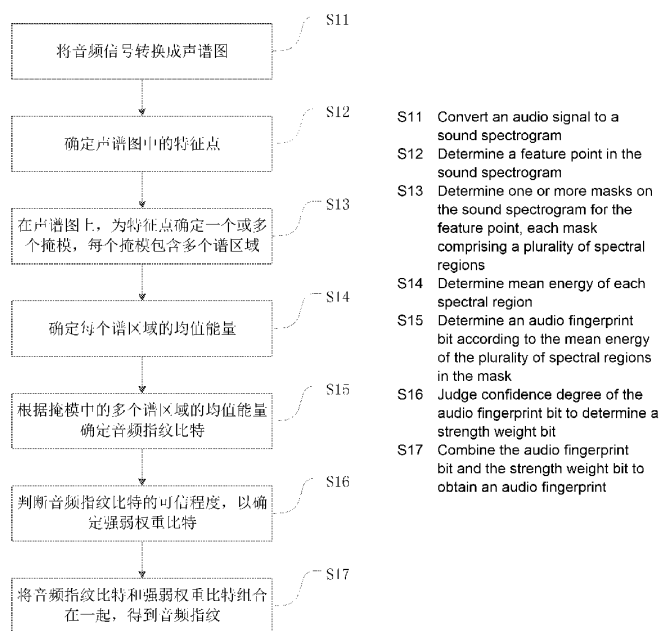


图 1

(57) Abstract: An audio fingerprint extraction method and device. The method comprises: converting an audio signal to a sound spectrogram; determining a feature point in the sound spectrogram; determining one or more masks on the sound spectrogram for the feature point, each mask comprising a plurality of spectral regions; determining mean energy of each spectral region; determining an audio fingerprint bit according to the mean energy of the plurality of spectral regions in the mask; judging confidence degree of the audio fingerprint bit to determine a strength weight bit; and combining the audio fingerprint bit and the strength weight bit to obtain

[见续页]

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, 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)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

an audio fingerprint.

(57) 摘要: 一种音频指纹提取方法及装置, 该方法包括: 将音频信号转换成声谱图; 确定该声谱图中的特征点; 在该声谱图上, 为该特征点确定一个或多个掩模, 每个掩模包含多个谱区域; 确定每个谱区域的均值能量; 根据该掩模中的多个谱区域的均值能量确定音频指纹比特; 判断该音频指纹比特的可信程度以确定强弱权重比特; 将该音频指纹比特和该强弱权重比特进行组合, 得到音频指纹。