

(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 2025/011361 A1

(43) 国际公布日
2025 年 1 月 16 日 (16.01.2025)

(51) 国际专利分类号:
G10L 19/008 (2013.01)

(21) 国际申请号: PCT/CN2024/102282

(22) 国际申请日: 2024 年 6 月 28 日 (28.06.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202310843541.X 2023 年 7 月 10 日 (10.07.2023) CN

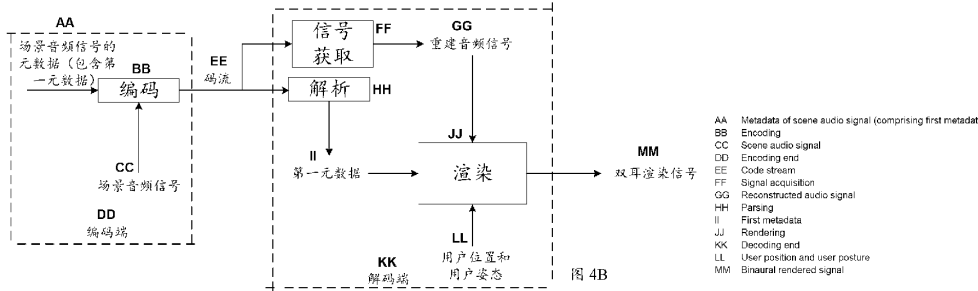
(71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(72) 发明人: 恩格尔·阿隆索·马丁内斯胡安·艾萨克 (ENGEL ALONSO MARTINEZ, Juan Isaac); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。徐德著 (XU, Dezhu); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。高原 (GAO, Yuan); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。王喆 (WANG, Zhe); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(74) 代理人: 北京润泽恒知识产权代理有限公司 (BEIJING RUN ZEHENG INTELLECTUAL PROPERTY LAW FIRM); 中国北京市海淀区

(54) Title: DECODING METHOD AND ELECTRONIC DEVICE

(54) 发明名称: 解码方法及电子设备



(57) Abstract: A decoding method and an electronic device. The method comprises: receiving a code stream; then parsing the code stream to obtain metadata of a scene audio signal; and acquiring a reconstructed audio signal of the scene audio signal on the basis of the code stream, wherein the scene audio signal is used for describing a sound field of a sound source in a scene, the scene comprises a plurality of microphones, the metadata comprises first metadata, the first metadata is used for determining a virtual loudspeaker radius on the basis of a distance between two microphones among the plurality of microphones, the virtual loudspeaker radius and a virtual loudspeaker signal are rendered to obtain a rendered audio signal, and the virtual loudspeaker signal is generated on the basis of the reconstructed audio signal. In this way, a decoding end does not need to acquire metadata of a scene audio signal from other electronic devices, and does not need to locally generate metadata of a scene audio signal, thereby reducing the calculation amount for generating the metadata of the scene audio signal, and further saving the time for performing rendering to obtain a binaural rendered signal and improving the user experience.

(57) 摘要: 一种解码方法及电子设备。方法包括: 首先, 接收码流; 接着, 从码流中解析出场景音频信号的元数据; 以及, 根据码流, 获取场景音频信号的重建音频信号; 场景音频信号用于描述场景中声源的声场, 场景包括多个麦克风, 元数据包括第一元数据, 第一元数据用于与多个麦克风中的两个麦克风之间的距离确定虚拟扬声器半径, 虚拟扬声器半径和虚拟扬声器信号用于渲染以得到渲染后的音频信号, 虚拟扬声器信号是根据重建音频信号生成的。这样, 解码端无需从其他电子设备获取场景音频信号的元数据, 也无需在本地生成场景音频信号的元数据, 降低了生成场景音频信号的元数据的计算量; 从而可以节省渲染得到双耳渲染信号的时间, 提高用户体验。

WO 2025/011361 A1

[见续页]

中关村南大街甲 18 号北京国际 C 座 6 层606, Beijing 100081 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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))。