(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

(43) 国际公布日
2022 年 11 月 24 日 (24.11.2022)

WIPO | PCT

(10) 国际公布号
WO 2022/242479 A1

(51) 国际专利分类号:
G10L 19/008 (2013.01)

(21) 国际申请号: PCT/CN2022/091557

(22) 国际申请日: 2022 年 5 月 7 日 (07.05.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202110536634.9 2021年5月17日 (17.05.2021) 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: 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)。 刘帅(**LIU, Shuai**); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 王 宾 (**WANG, Bin**); 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 中 博 世 达 专 利 商 标 代 理 有 限 公 司 (**BEIJING ZBSD PATENT&TRADEMARK AGENT LTD.**); 中国北京市海淀区交大东路31号11号楼8层, Beijing 100044 (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, IQ, IR, IS,

(54) **Title:** THREE-DIMENSIONAL AUDIO SIGNAL ENCODING METHOD AND APPARATUS, AND ENCODER

(54) 发明名称: 三维音频信号编码方法、装置和编码器

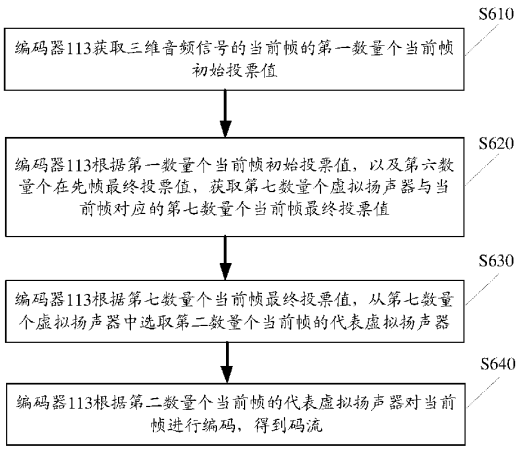


图 6

S610 An encoder 113 obtains a first number of current frame initial voting values of the current frame of a three-dimensional audio signal

S620 The encoder 113 then obtains, according to the first number of current frame initial voting values and a sixth number of previous frame final voting values, a seventh number of current frame final voting values of the current frame corresponding to a seventh number of virtual loudspeakers

S630 The encoder 113 selects representative virtual loudspeakers of a second number of current frames from the seventh number of virtual loudspeakers according to the seventh number of current frame final voting values

S640 The encoder 113 encodes the current frame according to the representative virtual loudspeakers of the second number of current frames to obtain a code stream

(57) **Abstract:** A three-dimensional audio signal encoding method and apparatus, and an encoder (113), relating to the field of multimedia. The method comprises: the encoder (113) obtains a first number of current frame initial voting values of the current frame of a three-dimensional audio signal (S610); the encoder (113) then obtains, according to the first number of current frame initial voting values and a sixth number of previous frame final voting values, a seventh number of current frame final voting values of the current frame corresponding to a seventh number of virtual loudspeakers (S620); the encoder (113) selects representative virtual loudspeakers of a second number of current frames from the seventh number of virtual loudspeakers according to the seventh number of current frame final voting values (S630); the encoder (113) encodes the current frame according to the representative virtual loudspeakers of the second number of current frames to obtain a code stream (S640). In this way, the continuity of signal azimuth between frames is enhanced, the stability of a reconstructed three-dimensional audio signal is improved, and the sound quality of the reconstructed three-dimensional audio signal is ensured.

WO 2022/242479 A1

[见续页]

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))。

(57) 摘要：一种三维音频信号编码方法、装置和编码器（113），涉及多媒体领域。该方法包括：编码器（113）获取三维音频信号的当前帧的第一数量个当前帧初始投票值（S610）后，编码器（113）根据第一数量个当前帧初始投票值，以及第六数量个在先帧最终投票值，获取第七数量个虚拟扬声器与当前帧对应的第七数量个当前帧最终投票值（S620）。进而，编码器（113）根据第七数量个当前帧最终投票值，从第七数量个虚拟扬声器中选取第二数量个当前帧的代表虚拟扬声器（S630）；编码器（113）根据第二数量个当前帧的代表虚拟扬声器对当前帧进行编码，得到码流（S640）。从而增强帧之间信号方位的连续性，提高了重建后三维音频信号的声像的稳定性，确保重建后三维音频信号的音质。