(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 4 月 7 日 (07.04.2022)

WIPO | PCT

(10) 国际公布号
WO 2022/068570 A1

(51) 国际专利分类号:
G10L 19/018 (2013.01)

(21) 国际申请号: PCT/CN2021/118202

(22) 国际申请日: 2021 年 9 月 14 日 (14.09.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202011066454.0 2020年9月30日 (30.09.2020) 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 亮 (**GAO, Liang**); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 施想 (**SHI, Xiang**); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 王 满 意 (**WANG, Manyi**); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 黄 国 捷 (**HUANG, Guojie**); 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: 深圳市深佳知识产权代理事务所 (普通合伙) (**SHENPAT INTELLECTUAL PROPERTY AGENCY**); 中国广东省深圳市罗湖区南湖街道春风路庐山大厦B座18C2、18D、18E、18E2, Guangdong 518001 (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, IT, 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。

(54) **Title:** AUDIO WATERMARK ADDING METHOD AND DEVICE, AUDIO WATERMARK ANALYZING METHOD AND DEVICE, AND MEDIUM

(54) 发明名称: 一种音频水印添加、解析方法、设备及介质

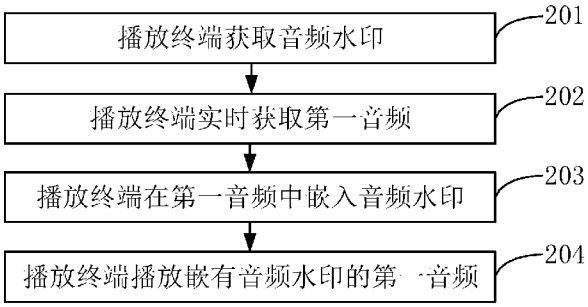


图 2

201 A playback terminal obtains an audio watermark
202 The playback terminal obtains a first audio in real time
203 The playback terminal embeds the audio watermark in the first audio
204 The playback terminal plays the first audio embedded with the audio watermark

(57) **Abstract:** An audio watermark adding method, comprising: a playback terminal obtains a first audio in real time (202); the playback terminal embeds an audio watermark in the first audio (203), the audio watermark being associated with the playback terminal; the playback terminal plays the first audio embedded with the audio watermark (204). Also provided are an audio watermark analyzing method and device, and a medium. In a real-time audio playback scenario, a playback terminal adds an audio watermark in an audio stream in real time, so that an after-treatment device can determine the playback terminal according to the audio watermark when analyzing the watermark, and the first audio can be traced after being transcribed.

(57) 摘要: 一种音频水印添加方法, 包括: 播放终端实时获取第一音频 (202); 播放终端在第一音频中嵌入音频水印 (203), 音频水印与播放终端相关联; 播放终端播放嵌有音频水印的第一音频 (204)。还提供一种音频水印解析方法、设备及介质, 在实时播放音频的场景下, 播放终端通过在音频流中实时地加入音频水印, 以使得后期设备在解析水印时能够根据该音频水印确定该播放终端, 便于在第一音频被转录后进行溯源。

WO 2022/068570 A1

[见续页]

(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))。

一种音频水印添加、解析方法、设备及介质

本申请要求于2020年9月30日提交中国专利局、申请号为“202011066454.0”、申请名称为“一种音频水印添加、解析方法、设备及介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及电子领域，尤其涉及一种音频水印添加、解析方法、设备及介质。

背景技术

水印技术是常用的图片处理方法，通过在重要的图片上标记水印，从而在图片传播的过程中能够根据该水印查询到图片的源头，使得传播者不敢轻易偷拍和传播图片，起到了震慑作用和追溯问责的能力。

当前，在音频领域也有类似的需求，例如在远程音视讯会议中，会出现会议内容外流的情况，导致商业或者个人机密信息泄露，并造成不良影响。 其主要手段是通过手机或者录音笔等设备，在某一远程终端侧，进行偷拍或者偷录，并将偷录的音视频文件，外传给其他人员，最终在互联网上传播，造成不良影响。

现有技术中的音频水印添加方法，主要是通过后期对音频进行离线处理以得到音频水印，在直播会议的场景下，对于在线的实时音频流，无法进行有效处理。

因此，现有技术中存在的上述问题还有待于改进。

发明内容

本申请实施例提供了一种音频水印添加、解析方法、设备及介质，用于解决音频水印的实时添加的问题。

有鉴于此，本申请实施例第一方面提供了一种音频水印添加方法，包括：播放终端实时获取第一音频；该播放终端在该第一音频中嵌入音频水印，该音频水印与该播放终端相关联；该播放终端播放嵌有该音频水印的该第一音频。

本实施例中，播放终端实时获取第一音频；播放终端在第一音频中嵌入音频水印，音频水印与播放终端相关联；播放终端播放嵌有音频水印的第一音频。从而在实时播放音频的场景下，通过播放终端在音频流中实时地加入音频水印，以使得后期设备在解析水印时能够根据该音频水印确定该播放终端，便于在第一音频被转录后进行溯源。

可选地，该播放终端在该第一音频中嵌入音频水印包括：该播放终端在该第一音频中确定满足第一预设条件的第一目标帧；该播放终端在该第一目标帧之后确定满足第二预设条件的第二目标帧，该第一目标帧用于标记该第二目标帧；该播放终端在该第二目标帧中嵌入该音频水印。

本实施例中，播放终端在对第一音频进行实时处理的过程中，根据第一预设条件确定第一目标帧，之后根据第二预设条件确定第二目标帧，最终在第二目标帧中嵌入音频水印。通过此种方式，使得播放终端能够在实时处理的过程中准确地找到第一音频中嵌入音频水印的合适位置。

可选地，该播放终端在该第一音频中确定满足第一预设条件的第一目标帧，包括：当

该第一音频的采样率大于或等于第一阈值时，该播放终端将低频部分最大值处于第一区间内的音频帧确定为该第一目标帧；或者，当该第一音频的采样率小于该第一阈值时，该播放终端确定包含第一特征声音的音频帧为该第一目标帧。

本实施例中，当该第一音频的采样率大于或等于第一阈值时，该播放终端将低频部分最大值处于第一区间内的音频帧确定为该第一目标帧；当第一音频的采样率小于第一阈值时，播放终端通过确定第一特征声音的方式来确定第一目标帧，不需要再嵌入同步帧标记。从而确保在第一音频中，无论第一音频的采样率大于还是小于第一阈值，均可以找到标记水印嵌入位置的第一目标帧。

可选地，当该第一音频的采样率大于或等于第一阈值时，该播放终端将低频部分最大值处于第一区间内的音频帧作为该第一目标帧之后，还包括：该播放终端在该第一目标帧中添加同步帧标记。

本实施例中，当该第一音频的采样率大于或等于第一阈值时，在第一目标帧中加入同步帧标记，以使得后续解析终端在解析时，能够根据该同步帧标记快速定位到第一目标帧。

可选地，该播放终端在该第一目标帧中添加同步帧标记，包括：该播放终端获取第一采样点，该第一采样点为中频部分的采样点；该播放终端提升该第一采样点的能量值，以使得该第一采样点的能量值与低频部分能量值的比值大于或等于第二阈值。

本实施例中，播放终端实时地判断第一音频中符合第一预设条件的目标帧作为第一目标帧或第二目标帧，之后提升第一目标帧或第二目标帧的中频部分第一采样点的能量值，以使得第一采样点的能量值与低频部分能量值的比值大于或等于预设比值，从而实现了同步帧标记的添加。

可选地，该当该第一音频的采样率小于该第一阈值时，该播放终端确定包含第一特征声音的音频帧为该第一目标帧，包括：当检测到该第一特征声音，且该第一特征声音的持续时间大于或等于预设时间时，该播放终端将包含该第一特征声音的音频帧确定为该第一目标帧。

本实施例中，第一特征声音可以为人声，或者在人声状态下，当检测到特定语句时，将该特定语句所在的目标帧确定为第一目标帧，从而确保后续的水印嵌入能够嵌入到记录有语音信息的目标帧中。

可选地，该播放终端在该第一目标帧之后确定满足第二预设条件的第二目标帧，包括：该播放终端确定中频部分能量值大于或等于第三阈值且小于第四阈值的目标帧为该第二目标帧。

本实施例中，第二目标帧是位于第一目标帧之后的目标帧，由于第二目标帧是实时加入到第一音频中的，因此无法保证第一音频中第一目标帧之后的每一帧都适合作为第二目标帧，从而需要对第二目标帧的条件进行判断，只有当目标帧满足第二预设条件时，才将该目标帧作为第二目标帧。

可选地，该播放终端在该第三目标帧中嵌入该音频水印，包括：该播放终端获取该音频水印所对应的第一数列，该第一数列中包括至少一个元素；该播放终端从该第三目标帧中获取至少一个第二采样点；该播放终端将该第一数列中的至少一个元素分别嵌入该至少

一个第二采样点中，其中，该第一数列中的一个元素对应一个第二采样点。

本实施例中，播放终端为了在第一音频中实时地嵌入音频水印，按照时序在第一目标帧之后寻找符合预设条件的第二目标帧进行音频水印的嵌入，在嵌入音频水印的过程中，通过改变第二目标帧中采样点在不同时域和/或不同频域部分能量值的能量比值，从而实现音频水印的实时嵌入，所嵌入的水印在音频转录的过程具有较强的抗干扰能力。能够通过数字信道或空气信道传播。

可选地，该播放终端将该第一数列中的至少一个元素分别加入该至少一个第二采样点中，包括：该播放终端调节该第二采样点在不同时域和/或不同频域部分能量值的能量比值，其中，一个该第二采样点的该能量比值与该第一数列中的一个元素相关联。

本实施例中，可选地，播放终端调节该第二采样点在不同时域和/或不同频域部分能量值的能量比值的具体方式可以为：提高第一子采样点前半部分的能量值，以使得该第一子采样点前半部分与后半部分能量值的比值大于或等于第五阈值，将该第一子采样点记录为1，其中，该第一子采样点为该至少一个第二采样点中的一个；提高第二子采样点后半部分的能量值，以使得该第二采样点后半部分与前半部分能量值的比值大于或等于该第五阈值，将该第二子采样点记录为0。

可选地，该播放终端将该第一数列中的至少一个元素分别加入该至少一个第二采样点之后，还包括：

当该第二采样点中高能量部分与低能量部分的比值小于该第五阈值时，提高该第二采样点中高能量部分的能量值。

可选地，该播放终端获取该音频水印所对应的第一数列之后，还包括：

播放终端在该第一数列中加入校验位，该校验位用于校验该第一数列的传输完整性。

本申请实施例第二方面提供了一种音频水印解析方法，包括：解析终端获取第一音频，该第一音频中嵌有音频水印，该音频水印与播放终端相关联，该播放终端用于将该音频水印实时嵌入该第一音频；该解析终端从该第一音频中解析该音频水印；该解析终端根据该音频水印确定该播放终端。

本申请实施例提供一种音频水印解析方法，包括：解析终端获取第一音频，第一音频中嵌有音频水印，音频水印与播放终端相关联，播放终端用于将音频水印实时嵌入第一音频；解析终端从第一音频中解析音频水印；解析终端根据音频水印确定播放终端。从而解析终端能够根据音频水印，确定将音频水印加入第一音频中的播放终端。

可选地，该解析终端从该第一音频中解析该音频水印之前，还包括：该解析终端确定该第一音频中满足第一预设条件的第一目标帧；该解析终端在该第一目标帧之后确定满足第二预设条件的第二目标帧；该解析终端从该第一音频中解析该音频水印，包括：该解析终端从该第二目标帧中解析该音频水印。

本实施例中，由于播放终端是实时将音频水印添加到第一音频中的，不能保证第一音频中的每一帧都符合水印嵌入的条件，因此在实时嵌入音频水印的方案中，不能像离线水印嵌入一样按照预设的规则进行嵌入。而是要根据第一预设条件和第二预设条件分别确定第一目标帧和第二目标帧，因此解析终端在解析时，也需要按照相同的条件解析第一目标

帧和第二目标帧。

可选地，当该第一音频的采样率小于第一阈值时，该解析终端确定该第一音频中满足第一预设条件的第一目标帧，包括：该解析终端从该第一音频中确定包含有第一特征声音，且该第一特征声音的持续时间大于或等于预设时间的目标帧作为该第一目标帧。

本实施例中，当该第一音频的采样率大于或等于第一阈值时，该解析终端将低频部分最大值处于第一区间内的音频帧确定为该第一目标帧；当第一音频的采样率小于第一阈值时，解析终端通过确定第一特征声音的方式来确定第一目标帧，不需要再解析同步帧标记。从而确保在第一音频中，无论第一音频的采样率大于还是小于第一阈值，解析终端均可以找到标记水印嵌入位置的第一目标帧。

可选地，当该第一音频的采样率大于或等于第一阈值时，该解析终端确定该第一音频中满足第一预设条件的第一目标帧；包括：该解析终端逐帧获取该第一音频中频部分与低频部分能量值的第一比值；当该解析终端获取到该第一比值大于或等于第二阈值的初始目标帧时，从该初始目标帧开始通过滑窗方式向后滑动检测该第一音频，以获取每个滑动窗口内中频部分与低频部分能量值的第二比值；该解析终端获取该第二比值最大的滑动窗口所在帧为该第一目标帧。

本实施例中，解析终端逐帧获取每一帧中频部分采样点与低频部分采样点的能量值的第一比值，当找到第一比值大于或等于第二阈值第一采样点时，确定第一采样点所在的目标帧为初始目标帧。然而，由于一帧有 2048 个采样点，第一采样点仅为其中的部分采样点，因此，当找到了符合第一比值的第一采样点后，第一采样点实际所在的第一目标帧相对于当前第一采样点所在的初始目标帧之间，可能会存在偏移。为解决此问题，需要从初始目标帧开始，通过滑窗方式向后移动，逐帧检测初始目标帧之后的第一音频，以获取每个滑动窗口内中频部分与低频部分能量值的第二比值，获取该第二比值最大的滑动窗口所在帧为该第一目标帧，从而防止解析偏差。

可选地，该第一目标帧中包括同步帧标记，该解析终端获取该第二比值最大的滑动窗口所在帧为该第一目标帧，包括：该解析终端从该第二比值最大的滑动窗口中获取中频部分能量值最高的第一采样点；该解析终端获取距离该第一采样点之前预设长度的第三采样点；该解析终端确定该第一采样点能量值与该第三采样点能量值的比值大于第七阈值的部分为该同步帧标记；该解析终端根据该同步帧标记确定该第二比值最大的滑动窗口所在帧为该第一目标帧。

本实施例中，上述同步帧标记的检测方法，是检测第一音频的中频部分与低频部分能量值的第一比值来确定的，然而在实际工作过程中，第一音频的原始内容（即非水印内容）中也可能存在中频部分与低频部分能量值比值大于第一比值的情况。从而会造成同步帧标记的误检测。对此，解析终端通过第一采样点与第一采样点之前预设长度的第三采样点的比值来确定同步帧标记，从而杜绝了上述误检测的情况发生。

可选地，该解析终端在该第一目标帧之后确定满足第二预设条件的第二目标帧，包括：该解析终端从该第一目标帧开始以帧为单位向后移动，分别获取每帧中频部分的能量大于或等于第三阈值且小于第四阈值的备选目标帧；该解析终端从该备选目标帧中获取不同时

域和/或不同频域部分能量值的能量比值大于或等于第五阈值的目标帧为该第二目标帧。

本实施例中，当解析终端检测到第一目标帧后，以第一目标帧为定位帧，即可继续寻找位于第一目标帧之后的第二目标帧，第二目标帧嵌有音频水印，且第二目标帧满足第二预设条件。解析终端可以根据该第二预设条件，在第一目标帧之后快速地找到第二目标帧。

可选地，该解析终端从该第二目标帧中解析该音频水印，包括：该解析终端从该第二目标帧中获取第二采样点，该第二采样点为该第二目标帧中该能量比值大于或等于该第五阈值的采样点；该解析终端分别获取所第二采样点中取不同时域和/或不同频域部分能量值的能量比值；该解析终端获取与该能量比值相关联的第一元素，该第一元素为该音频水印所记录的第一数列中的一个元素。

本实施例中，可选地，该第一数列中包含校验位，则该将该至少一个第二目标帧的数字按解析顺序生成第一数列之后，还包括：根据该校验位确定该第一数列是否完整；若是，则将该第一数列转化为十进制数列；若否，则忽略该第一数列。

可选地，该解析终端从该第二目标帧中解析该音频水印之前，还包括：解析终端根据该第二目标帧的时长调节该第一长度的长度值，其中，该第二目标帧的时长越长，该第一长度的长度越大。解析终端分别去除该第二目标帧头部和尾部第一长度的能量值。

可选地，该方法包括多个水印检测周期，其中，每个该水印检测周期分别解析出一个该音频水印，该方法还包括：从该多个水印检测周期所解析的音频水印中确定重复率最高的一个作为该第一音频的水印。

本实施例中，第一音频中包括多个水印解析周期，每个周期内包括一个第一目标帧和一个第二目标帧，每个水印检测周期中均嵌入有相同的音频水印。在实际解析的过程中，解析终端可能会出现一些解析错误的情况，导致不是所有水印检测周期解析得到的音频水印都是相同的数列。当解析出错时，误解析所得到的错误音频水印总是随机且不重复的，因此多个水印检测周期所解析的音频水印中重复率最高的一个，可以确定为正确的音频水印。通过这种多周期决策的方式，准确地解析出第一音频中所嵌入的正确水印，进一步防止解析终端的误解析。

本申请实施例第三方面提供一种播放终端，包括：

获取单元，用于实时获取第一音频；

执行单元，用于在所述获取单元获取的所述第一音频中嵌入音频水印，所述音频水印与所述播放终端相关联；

播放单元，用于播放由所述执行单元嵌有所述音频水印的所述第一音频。

可选地，该执行单元，还用于：

在该第一音频中确定满足第一预设条件的第一目标帧；

在该第一目标帧之后确定满足第二预设条件的第二目标帧，该第一目标帧用于标记该第二目标帧；

在该第二目标帧中嵌入该音频水印。

可选地，该执行单元，还用于：

当该第一音频的采样率大于或等于第一阈值时，将低频部分最大值处于第一区间内的

音频帧确定为该第一目标帧；或者，

当该第一音频的采样率小于该第一阈值时，确定包含第一特征声音的音频帧为该第一目标帧。

可选地，当该第一音频的采样率大于或等于第一阈值时，该执行单元，还用于：

在该第一目标帧中添加同步帧标记。

可选地，该执行单元，还用于：

获取第一采样点，该第一采样点为中频部分的采样点；

提升该第一采样点的能量值，以使得该第一采样点的能量值与低频部分能量值的比值大于或等于第二阈值。

可选地，该当该第一音频的采样率小于该第一阈值时，该执行单元，还用于：

当检测到该第一特征声音，且该第一特征声音的持续时间大于或等于预设时间时，将包含该第一特征声音的音频帧确定为该第一目标帧。

可选地，该执行单元，还用于：

确定中频部分能量值大于或等于第三阈值且小于第四阈值的目标帧为该第二目标帧。

可选地，该执行单元，还用于：

获取该音频水印所对应的第一数列，该第一数列中包括至少一个元素；

从该第三目标帧中获取至少一个第二采样点；

将该第一数列中的至少一个元素分别嵌入该至少一个第二采样点中，其中，该第一数列中的一个元素对应一个第二采样点。

可选地，该执行单元，还用于：

调节该第二采样点在不同时域和/或不同频域部分能量值的能量比值，其中，一个该第二采样点的该能量比值与该第一数列中的一个元素相关联。

本申请实施例第三方面的有益效果可参阅上述第一方面，此处不再赘述。

本申请实施例第四方面提供一种解析终端，包括：

获取单元，用于获取第一音频，该第一音频中嵌有音频水印，该音频水印与播放终端相关联，该播放终端用于将该音频水印实时嵌入该第一音频；

解析单元，用于从该获取单元获取的该第一音频中解析该音频水印；

执行单元，用于根据该解析单元解析的该音频水印确定该播放终端。

可选地，该解析单元，还用于：

确定该第一音频中满足第一预设条件的第一目标帧；

在该第一目标帧之后确定满足第二预设条件的第二目标帧；

从该第二目标帧中解析该音频水印。

可选地，当该第一音频的采样率小于第一阈值时，该解析单元，还用于：

从该第一音频中确定包含有第一特征声音，且该第一特征声音的持续时间大于或等于预设时间的目标帧作为该第一目标帧。

可选地，当该第一音频的采样率大于或等于第一阈值时，该解析单元，还用于：

逐帧获取该第一音频中频部分与低频部分能量值的第一比值；

当获取到该第一比值大于或等于第二阈值的初始目标帧时，从该初始目标帧开始通过滑窗方式向后滑动检测该第一音频，以获取每个滑动窗口内中频部分与低频部分能量值的第二比值；

获取该第二比值最大的滑动窗口所在帧为该第一目标帧。

可选地，该第一目标帧中包括同步帧标记，该解析单元，还用于：

从该第二比值最大的滑动窗口中获取中频部分能量值最高的第一采样点；

获取距离该第一采样点之前预设长度的第三采样点；

确定该第一采样点能量值与该第三采样点能量值的比值大于第七阈值的部分为该同步帧标记；

根据该同步帧标记确定该第二比值最大的滑动窗口所在帧为该第一目标帧。

可选地，该解析单元，还用于：

从该第一目标帧开始以帧为单位向后移动，分别获取每帧中频部分的能量大于或等于第三阈值且小于第四阈值的备选目标帧；

从该备选目标帧中获取不同时域和/或不同频域部分能量值的能量比值大于或等于第五阈值的目标帧为该第二目标帧。

可选地，该解析单元，还用于：

从该第二目标帧中获取第二采样点，该第二采样点为该第二目标帧中该能量比值大于或等于该第五阈值的采样点；

分别获取所第二采样点中取不同时域和/或不同频域部分能量值的能量比值；

获取与该能量比值相关联的第一元素，该第一元素为该音频水印所记录的第一数列中的一个元素。

可选地，该第一音频中包括多个水印检测周期，其中，每个该水印检测周期分别解析出一个该音频水印，该解析单元，还用于：

从该多个水印检测周期所解析的音频水印中确定重复率最高的一个作为该第一音频的水印。

本申请实施例第四方面的有益效果可参阅上述第二方面，此处不再赘述。

本申请实施例第五方面提供一种电子设备，该电子设备包括：交互装置、输入/输出（I/O）接口、处理器和存储器，该存储器中存储有程序指令；该交互装置用于获取用户输入的操作指令；该处理器用于执行存储器中存储的程序指令，执行如上述第一方面或第二方面任意一种可选的实现方式所述的方法。

本申请实施例第六方面提供一种计算机可读存储介质，包括指令，当该指令在计算机设备上运行时，使得该计算机设备执行如上述第一方面或第二方面任意一种可选的实现方式所述的方法。

附图说明

图 1 为本申请实施例所提供的音频水印添加方法的一个使用场景示意图；

图 2 为本申请实施例所提供的音频水印添加方法的一个实施例的示意图；

图 3 为本申请实施例所提供的音频水印添加方法的另一个实施例的示意图；

图 4 为本申请实施例所提供的音频水印添加方法的另一个实施例的示意图；
图 5a 为本申请实施例所提供的音频水印添加方法的另一个实施例的示意图；
图 5b 为本申请实施例所提供的音频水印添加方法的另一种实现方式的示意图；
图 5c 为本申请实施例所提供的音频水印添加方法的另一种实现方式的示意图；
图 5d 为本申请实施例所提供的音频水印添加方法的另一种实现方式的示意图；
图 5e 为本申请实施例所提供的音频水印添加方法的另一种实现方式的示意图；
图 5f 为本申请实施例所提供的音频水印添加方法的另一种实现方式的示意图；
图 5g 为本申请实施例所提供的音频水印添加方法的另一种实现方式的示意图；
图 5h 为本申请实施例所提供的音频水印添加方法的另一种实现方式的示意图；
图 5i 为本申请实施例所提供的音频水印添加方法的另一种实现方式的示意图；
图 6 为本申请实施例所提供的音频水印解析方法的一个实施例的示意图；
图 7 为本申请实施例所提供的音频水印解析方法的另一个实施例的示意图；
图 8 为本申请实施例所提供的音频水印解析方法的另一个实施例的示意图；
图 9a 为本申请实施例所提供的音频水印解析方法的另一种实现方式的示意图；
图 9b 为本申请实施例所提供的音频水印解析方法的另一种实现方式的示意图；
图 10 为本申请实施例所提供的音频水印解析方法的另一个实施例的示意图；
图 11 为本申请实施例所提供的音频水印解析方法的另一个实施例的示意图；
图 12a 为本申请实施例所提供的音频水印解析方法的另一种实现方式的示意图；
图 12b 为本申请实施例所提供的音频水印解析方法的另一种实现方式的示意图；
图 13 为本申请实施例一种使用场景的示意图；
图 14 为本申请实施例一种使用场景的示意图；
图 15 为本申请实施例所提供的一种电子设备的示意图；
图 16 为本申请实施例所提供的播放终端的示意图；
图 17 为本申请实施例所提供的解析终端的示意图。

具体实施方式

本发明实施例提供一种音频水印添加、解析方法、设备及介质，能够解决音频水印的实时添加问题。

为了使本技术领域的人员更好地理解本申请方案，下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本申请一部分的实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都应当属于本申请保护的范围。

本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”、“第三”“第四”等（如果存在）是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的实施例能够以除了在这里图示或描述的内容以外的顺序实施。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、

产品或设备固有的其它步骤或单元。

水印技术是常用的图片处理方法，通过在重要的图片上标记水印，从而在图片传播的过程中能够根据该水印查询到图片的源头，使得传播者不敢轻易偷拍和传播图片，起到了震慑作用和追溯问责的能力。

当前，在音频领域也有类似的需求，例如在远程音视讯会议中，会出现会议内容外流的情况，导致商业或者个人机密信息泄露，并造成不良影响。 其主要手段是通过手机或者录音笔等设备，在某一远程终端侧，进行偷拍或者偷录，并将偷录的音视频文件，外传给其他人员，最终在互联网上传播，造成不良影响。

当前，音频水印添加方法主要是通过后期对音频进行离线处理以得到音频水印，在直播会议的场景下，对于在线的实时音频流，无法进行有效处理。

例如，对本申请实施例所提供方法的应用场景进行说明。

请参阅图 1，如图 1 所示，本申请实施例的使用场景可以应用与会议的场景中，该场景包括会场 A101、会场 B102、会场 C103 以及在上述三个会场间调度音频的媒体中心 104。其中，会场 A101、会场 B102 和会场 C103 可以是分别位于不同地点的远程会场。在具体工作过程中，例如，会场 A101 的代表发言，会场 A101 的录音设备获取代表的发言，之后会场 A101 的通信设备将实时的音频流发送给媒体中心 104，该媒体中心 104 将该音频流分别发送给会场 B102 和会场 C103，会场 B102 和会场 C103 的通信设备获取到该音频流后通过外放设备实时播放来自会场 A101 的实时音频。从而实现了会场 A101、会场 B102 和会场 C103 之间的远程音频会议。

在上述工作过程中，会场 A、会场 B 和会场 C 的听众均有可能偷偷录制所播放的音频并传播泄露，因此在溯源问责时，需要知晓该音频具体为哪一会场的观众偷录。

为了解决上述问题，本申请实施例提供一种音频水印添加方法，能够通过对实时播放的音频添加水印，解决音频的溯源问题。

以下结合附图对本申请实施例所提供的音频水印添加方法进行详细的说明。需要说明的是，上述应用场景只是一种举例，并不构成对本申请使用场景的限定，本申请实施例所提供的方法还可以应用于其他场景中，对此本申请实施例并不进行限定。

请参阅图 2，如图 2 所示，本申请所提供音频水印添加方法的实施例包括以下步骤。

201、播放终端获取音频水印。

本实施例中，播放终端获取与该播放终端相关联的音频水印，在后续水印解析的过程中，解析终端能够通过该音频水印知晓，该音频水印是由该播放终端嵌入到音频中的。

可选地，该音频水印可以是预存在播放终端本地的，也可以是其他设备发送给播放终端的，例如服务管理中心（service management center，SMC）向多个不同的播放终端分别发送各自的音频水印。为便于说明，本实施例中以音频水印为数字 14 进行举例说明。当前播放终端获取 SMC 发送的音频水印“14”后，将该音频水印存储在本地，该数字 14 即为与当前播放终端相关联的音频水印。

202、播放终端实时获取第一音频。

本实施例中，该第一音频可以是播放终端从外部获取的，例如该播放终端可以是图 1

中的会场 B 的播放终端，也可以是图 1 中会场 C 的播放终端，该播放终端从媒体中心实时获取第一音频。可选地，在其他的应用场景中，该第一音频也可以是播放终端从自身的内存中实时获取的，对此本申请实施例并不进行限定。

203、播放终端在第一音频中嵌入音频水印。

本实施例中，播放终端的音频水印“14”与播放终端相关联，从而通过该音频水印即可唯一地确定该播放终端。在具体工作过程中，播放终端在播放第一音频的过程中实时地将该音频水印嵌入第一音频中。

可选地，本申请实施例进一步提供一种在第一音频中嵌入音频水印的具体工作方式，为便于理解，以下结合附图进行详细说明。

请参阅图 3，如图 3 所示，本申请所提供音频水印添加方法中在第一音频中嵌入音频水印的方法包括以下步骤。

301、播放终端在第一音频中确定满足第一预设条件的第一目标帧。

本实施例中，第一目标帧起到了标记的作用，在后续对音频水印进行解析的过程中，解析终端根据第一目标帧即可知晓第一目标帧之后存在音频水印，从而可以实现对音频水印的快速定位。

需要说明的是，由于第一目标帧是实时加入到第一音频中的，因此无法保证第一音频中的每一帧都适合作为第一目标帧，从而需要对第一目标帧的条件进行判断，只有当目标帧满足第一预设条件时，才将该目标帧作为第一目标帧。

可选地，放终端中可以设置一个同步帧条件检测器，该同步帧条件检测器可以为在设置在播放终端中的实体装置，也可以是存储在播放终端中的运行逻辑，该同步帧条件检测器运行上述步骤 301 所述的方法，从而确定了满足第一预设条件的第一目标帧。

需要进一步说明的是，如上所述，第一目标帧满足第一预设条件，在具体工作过程中，按照采样率大小的不同，第一预设条件并不相同。其中，当第一音频的采样率大于或等于第一阈值时，播放终端将低频部分最大值处于第一区间内的音频帧确定为第一目标帧。当第一音频的采样率小于第一阈值时，播放终端确定包含第一特征声音的音频帧为第一目标帧。为便于理解，以下分别对第一预设条件的两种情况进行详细说明。

一、当第一音频的采样率大于或等于第一阈值时，播放终端将低频部分最大值处于第一区间内的音频帧确定为第一目标帧。

本实施例中，当第一音频的采样率大于或等于第一阈值时，播放终端采用在第一目标帧中加入同步帧标记的方式来标识该第一目标帧。由于添加同步帧标记的方式是改变第一目标帧中频部分的能量值，这就要求：对于第一目标帧的原始音频而言，其低频能量不能过高或过低，在实际应用中，常常会出现原始音频低频能量过低，导致中频相对能力也过低，从而导致在录制后解析终端解析不到同步帧标记的问题。另外当原始音频低频能力过高，导致中频能量过高，会出现爆音的问题。因此要确定低频部分最大值处于第一区间内的音频帧确定为第一目标帧。

第一区间具体方式可以为（Tlow，Thigh），播放终端在确定第一目标帧时，取第一音频低频部分的能力最大值 Value，判断该 Value 是否符合公式：Tlow< Value < Thigh，若满

足条件，则将当前目标帧确定为第一目标帧。其中，Tlow 为第一区间能量值的下限，Thigh 为第一区间能量值的上限。Tlow 和 Thigh 的具体数值可以由本领域技术人员根据实际需要制定，例如，针对会议室场景、大礼堂场景及开放办公区等不同的音频录制场景设置不同的阈值范围，特别是能量过低的阈值，以保证其在不同场景下，同步帧标记在第一目标帧中可以获得良好的嵌入强度。

进一步地，当播放终端按照上述方式确定第一目标帧后，需要在目标帧中添加同步帧标记，以使得后续的解析终端能够按照该同步帧标记确定第一目标帧。同步帧标记的添加具体包括以下步骤。

1、播放终端从第一目标帧中获取第一采样点。

本实施例中，该第一采样点为第一目标帧的中频部分的采样点。

2、播放终端提升第一采样点的能量值，以使得第一采样点的能量值与低频部分能量值的比值大于或等于第二阈值。

本实施例中，第一采样点为中频部分的采样点，在第一采样点的能能量值被播放终端提高之后，第一采样点与低频部分其他采样点之间的能量值的比值出现了显著的区别，其中，第一采样点的能量值与低频部分能量值的比值大于或等于第二阈值。从而后续解析终端在获取第一目标帧时，根据中频部分采样点与低频部分采样点能量值的比值即可解析到该同步帧标记。

为便于理解，基于上述步骤1-2所述的思路，以下提供一种同步帧标记加入的更具体的实现方式。

1)、对第一目标帧的音频信号做快速傅里叶变换(FFT), 取其低频部分，求得能量的最大值 $\max_1$。计算出用于修改同步信号能量值的参考系数$\max_E_1 = \frac{max_1}{5}$。

本实施例中，如图4所示，对对第一目标帧的音频信号做快速傅里叶变换后，得到了第一目标帧的音频信号在频域和时域上的折线图。通过图4，即可知道低频部分能量的最大值。

2)、取中频部分的8个能量值$E(i)_1, 0 \leq i < 8$,计算出8个点的最大值max'_1,然后对$E(i)'_1$进行修改$E(i)'_1 = \frac{E(i) * max_E_1}{max'_1}$。

本实施例中，如图 4 所示，经过修改后，中频部分第二采样点 401 的能量值有了显著的提升，由于对中频部分能量值 E(i)'的修改公式中添加了低频部分能量最大值max_E_1，从而使得中频部分第一采样点的能量值相对低频部分的能量有了显著的提升。需要说明的是，上述取 8 个能量值只是一种优选的举例，本领域技术人员可以根据实际需要选择不同数量的能量值，对此本申请实施例并不进行限定。

3)、在中频部分取与上述 8 个能量点对称的点，进行相同的能量值提升操作。

本实施例中，由于傅里叶变换后的能量分布具有对称性，所以在与这 8 个点对称的部分取 8 个点，作相同的嵌入操作。

4)、频域信号做快速逆傅里叶变换(IFFT)。

本实施例中，在经过快速傅里叶变换之后，通过上述方式修改了第一目标帧中频部分的能量值，之后再对第一目标帧做快速逆傅里叶变换，从而得到了嵌入同步帧标记的时域信号，该时域信号即为已经嵌入同步帧标记的第一目标帧。

本实施例中，通过上述步骤1）-4）所述的方式，将第一目标帧中频部分第一采样点的能量值进行提升，以使得第一采样点的能量值与低频部分的能量值的比值超过预设的范围。以使得后续解析过程中，当解析终端获取到某一目标帧中频部分有一个采样点的能量值与低频部分能量值的比值大于预设值时，即可判断该目标帧为添加了同步帧标记的第一目标帧。从而实现了第一目标帧中同步帧标记的添加。

本实施例中，播放终端实时地判断第一音频中符合第一预设条件的目标帧作为第一目标帧，之后提升第一目标帧的中频部分第一采样点的能量值，以使得第一采样点的能量值与低频部分能量值的比值大于或等于预设比值，从而实现了同步帧标记的添加。从而在第一音频的采样率大于或等于第一阈值时，播放终端将低频部分最大值处于第一区间内的音频帧确定为第一目标帧。

需要说明的是，当第一音频的采样率小于第一阈值时，由于同步帧标记的方式提升了中频部分的能量值，对于采样率小于预设值的音频而言，同步帧标记的嵌入会影响到人耳收听第一音频的体验，因此不能再用同步帧的方式来定位第一目标帧，为了解决此问题，本申请实施例提供了第二种方案。

二、当第一音频的采样率小于第一阈值时，播放终端确定包含第一特征声音的音频帧为第一目标帧。

播放终端确定包含第一特征声音的音频帧为第一目标帧。具体方法如下。

当检测到第一特征声音，且第一特征声音的持续时间大于或等于预设时间时，播放终端将包含第一特征声音的音频帧确定为第一目标帧。

本实施例中，第一特征声音可以通过声音检测方法进行检测，其中，检测特征声音的方式可以为现有技术中的任意一种方式，本申请实施例对此并不进行限定。具体工作过程中，第一特征声音可以为人声，例如，当检测到没有人声的时间持续大于预设时间后，再次检测到人声的时刻，作为第一目标帧。这样做的好处在于能够适用于会议场景，在会议场景下，人们通过语音进行交流，为了防止语音被转录，因此需要在记录有语音的音频中加入音频水印，通过上述人声检测的方式，当检测到没有人声的时间持续大于预设时间（例如1.5s）后，将再次检测到人声的时刻作为第一目标帧，从而确保后续的水印嵌入能够嵌入到记录有语音信息的音频中。

可选地，第一特征声音还可以做更细化的判断，例如，在人声状态下，当检测到特定语句时，将该特定语句所在的目标帧确定为第一目标帧。

本实施例中，当第一音频的采样率小于第一阈值时，播放终端通过确定第一特征声音的方式来确定第一目标帧，从而通过特征声音的方式实现了水印嵌入起始位置（即第一目标帧）的确认，不需要再嵌入同步帧标记。从而确保在第一音频中，无论第一音频的采样率大于还是小于第一阈值，均可以找到标记水印嵌入位置的第一目标帧。

可选地，情况一和情况二的两种方式可由播放终端中的同步帧嵌入器来实现，该同步

帧嵌入器可以为在设置在播放终端中的实体装置，也可以是存储在播放终端中的运行逻辑，对此本申请实施例并不进行限定。

通过上述方式，播放终端在第一音频中确定了第一目标帧，此时，播放终端需要执行后续步骤，在第一目标帧之后嵌入音频水印。

302、播放终端在第一目标帧之后确定满足第二预设条件的第二目标帧。

本实施例中，第二目标帧是位于第一目标帧之后的目标帧，由于第二目标帧是实时确定的，因此无法保证第一音频中第一目标帧之后的每一帧都适合作为第二目标帧，从而需要对第二目标帧的条件进行判断，只有当目标帧满足第二预设条件时，才将该目标帧作为第二目标帧。

需要说明的是，在实际工作过程中，播放终端在第一音频中对于第一目标帧和第二目标帧的确定是周期性的。首先，由于播放终端需要在第一音频中实时添加音频水印，因此播放终端只能够按照播放时间的先后顺序对第一音频进行处理。在处理过程中，播放终端先按照第一预设条件确定第一目标帧；接着在第一目标帧之后确定满足第二预设条件的第二目标帧。此时，第一目标帧与第二目标帧构成了一个音频水印周期。在下一个音频水印周期中，播放终端依然先按照第一预设条件确定第一目标帧，接着在第一目标帧之后确定满足第二预设条件的第二目标帧。这样处理的结果是，第一音频中包括多个第一目标帧，其中，第一音频中每两个相邻的第一目标帧之间包含一个第二目标帧。

因此，作为一种可选的技术方案，播放终端在向第一音频添加水印的过程中，也可以先确定第一目标帧，再确定第二目标帧，其结果依然可以实现第一音频中每两个相邻的第一目标帧之间包含一个第二目标帧，因此本领域技术人员可以根据实际需要确定第一目标帧和第二目标帧的确定顺序，对此本申请实施例并不进行限定。为便于理解，本申请实施例仅以第二目标帧位于第一目标帧之后进行说明。

进一步地，上述确定第二目标帧的第二预设条件的具体实现方式可以为：

播放终端确定中频部分能量值大于或等于第三阈值且小于第四阈值的目标帧为第二目标帧。

本实施例中，由于在第二目标帧中嵌入水印的位置为中频区域，如果中频区域能量过低，则导致嵌入水印后，容易导致误解析或无法检测到水印；如果中频区域能量过高，则导致嵌入水印后，产生爆音。进一步地，由于连续的水印嵌入可能会导致相互干扰，因此在确定第二目标帧时，还可以进一步确保多个周期之间的第二目标帧之间保留有足够的间隔，进一步地，第三目标帧之间的间隔需要大于或等于第六阈值，该第六阈值的具体大小可由本领域技术人员根据实际情况设定，对此本申请实施例并不进行限定。

需要说明的是，第二目标帧中频部分能量值大于或等于第三阈值且小于第四阈值，该第三阈值小于该第四阈值，第三阈值与第四阈值的具体数值可以由本领域技术人员根据实际需求来确定，对此本申请实施例并不进行限定。

303、播放终端在第二目标帧中嵌入音频水印。

本实施例中，第二目标帧是位于第一目标帧之后的目标帧，在第一音频的多个水印周期中，第二目标帧位于两个第一目标帧之间，在第二目标帧中嵌入音频水印，以使得后续

水印解析的过程中，解析终端能够根据第一目标帧确定第二目标帧所在的位置，从而找到音频水印。

进一步地，本申请实施例进一步提供一种播放终端在第二目标帧中嵌入音频水印的具体实现方式，为便于理解，以下结合附图 5a，对此种情况进行详细说明，如图 5a 所示，嵌入音频水印的步骤包括。

501、播放终端获取音频水印所对应的第一数列。

本实施例中，可选地，音频水印可以通过数列的形式呈现，例如，播放终端为会场 B 所在的播放终端，如上述步骤 201 中的举例，该播放终端的标识为“14”，则该音频水印所对应的第一数列为“14”，后续步骤需要将该数列 14 作为音频水印嵌入第一音频中，从而在该第一音频中标记与之关联的播放终端。

502、播放终端从第二目标帧中获取至少一个第二采样点。

本实施例中，一个目标帧中包括 2048 个采样点，播放终端从第二目标帧中获取至少一个第二采样点，该第二采样点用于在后续工作过程中嵌入上述第一数列中的元素。

503、播放终端将第一数列中的至少一个元素分别嵌入至少一个第二采样点中。

本实施例中，播放终端将第一数列中的元素嵌入第二采样点，其中，第一数列中的一个元素嵌入到一个第二采样点中。以使得第二采样点记录了第一数列的内容，后续解析终端通过读取第二采样点中所记录的第一数列，即可解析出音频水印所记载的内容。

需要说明的是，播放终端将第一数列中的至少一个元素分别嵌入至少一个第二采样点的步骤可以通过以下方式来实现。

播放终端调节第二采样点在不同时域和/或不同频域部分能量值的能量比值。

本实施例中，上述能量比值的大小与第一数列中的数字相关联，不同的能量比值可以对应不同的数字，从而通过不同的能量比值记录了第一数列中的不同数字。由此通过改变能量比值的方式在第二采样点中记录了第一数列的内容，实现了音频水印的嵌入。

例如，播放终端需要嵌入的的音频水印为数字“14”，在将音频水印嵌入第一音频时，需要将该音频水印转化为二进制，从而将该数字“14”转化为二进制得到第一数列：“1110”，该第一数列即为需要作为音频水印嵌入第一音频中的内容。第一数列中包含 1110 四个元素，这四个元素分别嵌入第二目标帧中的四个第二采样点，从而实现音频水印的嵌入。

在具体工作过程中，通过改变第二采样点能量比值的方式，实现第一数列中每个元素在第二采样点中的嵌入。对于第二采样点中能量比值的改变可以采用以下三种不同的方案：一、改变不同时域部分能量值的比值；二、改变不同频域部分能量值的比值；三、同时改变不同时域和不同频域部分能量值的比值。为便于理解，以下结合附图对此三种方式做详细的说明。

一、改变不同时域部分能量值的比值。

本实施例中，通过改变不同时域部分能量值的比值实现时域嵌入的方式包括以下步骤。

1、降低第二子采样点前半部分的能量值，以使得第二采样点后半部分与前半部分能量值的比值大于或等于第五阈值。

本实施例中，如图 5b 所示，在第二采样点中按照时域的顺序，降低第二采样点前半部

分 5061 的能量值，以使得第二采样点后半部分 5062 的能量值显著高于前半部分 5061 的能量值。此时，可以将这种能量分布的波形预设为数字 0。

具体地，可以通过以下方式来实现。

如图 5c 所示，图 5c 为第二采样点原始帧的波形图，对图 5c 所示的原始帧进行第一次离散余弦变换 DCT 后得到如图 5d 所示的波形图，之后从图 5d 所示的波形图中选出中频部分进行第二次 DCT 变换得到如图 5e 所示的波形图，此时按照如下公式 1 对图 5e 所示波形的能量进行处理。

$$P'(j) = \frac{P(j)}{\lambda}, 1 \le j \le mid$$（公式 1）

在上述公式 1 中，j 为时域上的时间段，P(j)为如图 5e 中波形在时段 j 内的总能量，λ 为预设的系数，其中，λ 的具体数值可以根据实际需要进行调整，mid 代表图 5e 中的波形在时域上的中点。从上述公式可知，j 的取值范围为 1 到中点，即图 5e 中的前半部分，将前半部分的总能量值 P(j)除以系数 λ，从而降低了第二采样点前半部分的能量值，得到如图 5b 所示的波形图。

2、降低第二子采样点后半部分的能量值，以使得第二采样点前半部分与后半部分能量值的比值大于或等于第五阈值。

本实施例中，如图 5f 所示，在第二采样点中按照时域的顺序，降低第二采样点后半部分 5063 的能量值，以使得第二采样点前半部分 5064 的能量值显著高于后半部分 5063 的能量值。此时，可以将这种能量分布的波形预设为数字 1。

具体地，可以通过以下方式来实现。

基于上述图 5e 所示的波形图，此时按照如下公式 2 对图 5e 所示波形的能量进行处理。

$$P'(j) = \frac{P(j)}{\lambda}, mid + 1 \le j \le S' - T'$$（公式 2）

在上述公式 1 中，j 为时域上的时间段，P(j)为如图 5e 中波形在时段 j 内的总能量，λ 为预设的系数，其中，λ 的具体数值可以根据实际需要进行调整，mid 代表图 5e 中的波形在时域上的中点，S'-T'表示图 5e 中时域的上限。从上述公式可知，j 的取值范围为 mid+1 到 S'-T'，即从时域中点开始到时域的终点，即图 5e 中的后半部分，将后半部分的总能量值 P(j)除以系数 λ，从而降低了第二采样点后半部分的能量值，得到如图 5f 所示的波形图。

本实施例中，通过改变第二采样点不同时域部分能量值的比值实现水印的嵌入。具体地，调整第二采样点在时域上前后部分的能量比值，之后将前后部分不同的能量比值分别预设为 0 和 1，从而在第二采样点上实现了二进制的数字嵌入，后续可以根据需要将该二进制数列转化为十进制数列，从而实现了将第一数列加入第二采样点的工作过程。在后续解析终端对音频水印进行解析时，直接获取目标帧中前后部分能量值的比值即可实现对音频水印的解析。

需要说明的是，在上述解析的过程中，该能量比值的确定，是以目标帧在时域上的中点为界限，获取前后部分的比值后按照上述预设规则解析的。然而，在实际工作过程中，

在对第一音频进行翻录时，由于场地的回声或混响等原因，会造成高能量部分在时域上增加。例如，音频水印为 1 时，时域上以中点为界限，前半部分的能量值与后半部分能量值的比值大于或等于第五预设值。然而，由于转录过程中回声或混响的存在，导致高能量区域在时域上越过了中点的界限，从而影响了能量比值的大小，造成解析终端无法根据能量比值获取到水印。为解决上述问题，本申请实施例进一步提供一种依据频域来改变能量比值的方法，以克服转录过程中所产生的回声混响对时域上的能量分部造成的影响。

二、改变不同频域部分能量值的比值。

本实施例中，采用与上述类似的方式，改变第二采样点在频域上的的能量比值，从而实现水印的嵌入，具体的实现方式可参阅上述改变不同时域部分能量比值的方法，区别在于，本种实施方式并不是在时域上以时间作为划界，而是在频域上以频率作为划界。

通过改变不同频域部分的能量比值所得到的能量分布图如图 5g 所示，需要说明的是，前述图 5b 至图 5f 所示的图片为能量值与时域对应关系的折线图，而图 5g 所显示的能量分布图之间显示能量在时域和频域上的分布关系。在图 5g 的能量分部框中，阴影部分表示高能量部分，白色部分表示低能量部分。在第二采样点中以频域的中点划界，上半部分与下半部分的比值大于第五阈值的能量分部 5065 表示数字 1，下半部分与上半部分的比值大于第五阈值的能量分部 5066 表示数字 0，从而通过改变频域上的能量比值实现了音频水印的嵌入，克服了转录过程中所产生的回声混响对时域上的能量分部造成的影响。

三、同时改变不同时域和不同频域部分能量值的比值。

本实施例中，综合了上述方案一和方案二的方案，首先通过方案一的方法，调整第二采样点在时域上前后部分的能量比值，实现第一步的水印嵌入，例如图 5h 所示，第二采样点前半部分 5067 的能量值与后半部分 5068 的能量值的比值大于第五阈值，嵌入了音频水印为数字 1，进一步地，为了防止转录过程中所产生的回声混响在时域上对音频水印能量分部造成的影响，如图 5h 所示，对于图 5h 的后半部分（即能量较低的部分），采取上述方案二的方法，改变高频部分 50681 与低频部分 50682 的能量比值，其中，第二采样点上半部分与下半部分的比值大于第五阈值的能量分布同样表示数字 1。

同理，将上述能量分部的比值对调，即可表示数字 0。

方案三所提供的方式，在第二采样点中同时加入了一大一小两个水印。该一大一小两个水印分别记录了相同的数值， 其中，大水印为在时域上改变第二采样点前后两部分的能量比值所得到的水印，即图 5h 中第二采样点前半部分 5067 的能量值与后半部分 5068 的能量比值所形成的水印，小水印为改变第二采样点低能量部分高低频的能量比值所得到的水印，即图 5h 中后半部分 5068 的高频部分 50681 与低频部分 50682 的能量比值所形成的水印。

进一步地，对于如图 5h 所示的水印，为了防止转录过程中所产生的回声混响对时域上的能量分部造成的影响，还可以对音频水印中的低能量部分进行进一步的分割，如图 5i 所示的能量分布图用于表示数字 1。其中，阴影部分表示高能量值区域，白色部分表示低能量值的区域。如图 5i 所示，在时域上前半部分 51 为高能量部分，后半部分 52 为低能量部分。对于后半部分 52 嵌入的小水印时，将该后半部分 52 在时域上进一步分为两部分，分

别记为第一部分 521 和第二部分 522，其中，第一部分 521 为接近前半部分 51（高能量部分）的区域，第二部分 522 为远离前半部分 51（高能量部分）的区域。在嵌入小水印时，仅仅改变第二部分 522 高频部分与低频部分的能量比值，对于第一部分 521，由于靠近高能量部分，可能会由于转录过程中产生的混响导致第一部分 521 的能量相应升高，因此不再改变第一部分 521 的高低频能量比值，从而排除了转录过程中所产生的回声混响对时域上的能量分部造成的影响，仅仅在远离高能量部分的第二部分 522 通过改变高低频能量比值的方式嵌入小水印，其中，第二部分 522 的高频部分 5221 的能量值高于低频部分 5222 的能量值，同样表示数字 1。大小两个水印分别表示同样的音频水印。

同理，将图 5i 中的高能量和低能量的部分对调，即可得到表示数字 0 的音频水印。

需要说明的是，上述图 5i 所示的音频水印，对于低能量部分（即后半部分 52），在时域上分为两个部分进行处理（第一部分 521 和第二部分 522），在实际工作过程中，本领域技术人员根据实际需要，可以将该低能量部分在时域上分割为更多的部分，对此本申请实施例并不进行限定。

本实施例中，播放终端为了在第一音频中实时地嵌入音频水印，按照时序在第一目标帧之后寻找符合预设条件的第二目标帧进行音频水印的嵌入，在嵌入音频水印的过程中，通过改变第三目标帧中采样点在不同时域和/或不同频域部分能量值的能量比值，来实现音频水印的嵌入，从而实现了音频水印的实时嵌入，同时所嵌入的水印在音频转录的过程具有较强的抗干扰能力。

需要说明的是，上述方案一至方案三的方案中，播放终端通过改变能量比值实现了二进制的水印嵌入，在实际工作过程中，本领域技术人员可根据实际需要，通过改变能量比值实现其他进制的水印嵌入，例如十进制或 12 进制，对此本申请实施例并不进行限定。

需要说明的是，在上述第一数列中，为了在解析的过程中确定第一音频中所嵌入水印的完整性，在播放终端第一音频中嵌入音频水印时，需要在第一数列中加入校验位，以使得后续解析终端能够根据该校验位确定音频水印传输的完整性，防止在第一音频转录过程中，由于转录的信号丢失导致音频水印传输不完整，产生解析失败的情况。为便于理解，以下对校验位的具体添加方式进行详细说明。

本实施例中，如上所述，播放终端所获取的音频水印为数字“14”，在将音频水印嵌入第一音频时，需要将该音频水印转化为二进制，从而将该数字“14”转化为二进制，得到第一数列：“1110”，该第一数列即为需要作为音频水印嵌入第一音频中的内容。为了确保该第一数列传输的完整性，可以在第一数列中加入校验位，例如，可以采用奇偶校验的方式，例如，第一数列“1110”中包括三个数字 1，即奇数个数字“1”，此时，在第一数列的最后一位加入一个数字 1，得到新的第一数列“11101”，该第一数列中最后一位数的 1 即为校验位，该校验位的“1”用于表示当前数列中除校验位外还有奇数个数字“1”。同理，若第一数列中数字“1”的数量为偶数个，则校验位为“0”，从而通过此种方式确定了第一数列的位数，后续解析终端可以根据校验位实现对第一数列的校验，确保第一数列的传输准确。

进一步地，由于第一音频在空气录制的过程中会有部分采样点丢失的问题，当音频水印中的第一数列过长，会导致音频水印所占用的帧数过多，在检测时，用于采样点的丢失，

导致每帧的起始位和预期产生偏差，使得音频水印检测不准确。为了解决此问题，当第一数列的长度超过阈值后，进行将第一数列分割成多个子数列，将每个子数列分别按照上述方式加入校验位后再重新组合成一个大的第一数列。从而在后续解析终端对音频水印进行解析时，能够按照预设的子周期分别对第一数列进行校验，保证了第一数列的传输完整性。

本实施例中，以音频水印为数字“14”为例，在加入校验位之后，音频水印所对应的第一数列为：“11101”。播放终端获取第一个第二目标帧，按照前述步骤 506 所记载的方式，将第一数列中的第一个元素：“1”嵌入第二目标帧中的第一个第二采样点中，实现了第一数列中第一位数的嵌入。完成了水印嵌入周期中的第一子周期。之后循环执行上述步骤 506，从第二目标帧中获取第二个第二采样点，按照同样方式将第一数列中的第二个元素：“1”嵌入第二目标帧的第二个第二采样点中，实现了第一数列中第二位数的嵌入。……以此类推，播放终端通过五个水印嵌入子周期所组成的完整水印嵌入周期，将第一数列中的五位数字嵌入第二目标帧的五个第二采样点中。

可选地，在上述每个水印嵌入的子周期中，为了保证当前子周期中水印嵌入的强度足够，播放终端在每个子周期完成之后，就需要执行一次嵌入强度的检测。

需要说明的是，水印嵌入强度的具体检测方式为：检测当前子周期中第二采样点在不同时域和/或不同频域部分能量值的能量比值是否大于第五阈值。以上述步骤 506 中第一种水印嵌入方案为例，假设当前第二采样点嵌入的数字为 1，则播放终端需要判断，当前第二采样点在时域上前半部分与后半部分的能量比值是否大于第五阈值，若是，执行后续步骤，若否，重新执行水印嵌入，此时将公式 1 中的 λ 的值提升，以进一步降低第二采样点后半部分的能量值，从而提升第二采样点在时域上前半部分与后半部分的能量比值，以使得水印嵌入强度达到要求。

本实施例中，通过上述步骤 501 至 503 所述的方式，播放终端在第二目标帧中嵌入了音频水印。至此步骤 203 完成。

204、播放终端播放嵌有音频水印的第一音频。

本实施例中，播放终端将嵌有音频水印的第一音频播放出来，由于音频水印是实时嵌入第一音频中的，因此在实时播放的场景下，播放终端所播放的音频依然能够带有音频水印，从而在后续被转录的过程中，可以根据该音频水印对该第一音频实现溯源。

综上所述，本申请实施例所提供的音频水印添加方法中，播放终端实时获取第一音频；播放终端在第一音频中嵌入音频水印，音频水印与播放终端相关联；播放终端播放嵌有音频水印的第一音频。从而在实时播放音频的场景下，通过播放终端在音频流中实时地加入音频水印，以使得后期设备在解析水印时能够根据该音频水印确定该播放终端，便于在第一音频被转录后进行溯源。

通过上述音频水印添加方法加入音频水印的第一音频，无论是通过数字信道进行翻录，还是通过空气信道进行翻录，翻录所得到的第一音频中，均可以被解析终端解析出音频水印，从而实现对第一音频的溯源，解析终端可以通过音频水印确定将该音频水印加入到第一音频中的播放终端。为便于理解，以下结合附图，对本申请实施例所提供的音频水印解析方法进行详细说明。

请参阅图 6，如图 6 所示，本申请所提供的音频水印解析方法的实施例包括以下步骤。

601、解析终端获取第一音频。

本实施例中，第一音频为播放终端通过上述方法嵌入了音频水印的音频，需要说明的是，该第一音频的初始播放源为该播放终端，即：是该播放终端将该音频水印嵌入该第一音频中。当播放终端播放了该第一音频后，第一音频直接被解析终端获取，也可以经过转录，该转录可以是通过数字信道进行的传播转录，也可以是通过空气信道进行的转录，对于这两种方式的转录本申请所提供的方法均能够进行解析。

可选地，在获取到第一音频后，解析终端还需要对第一音频进行格式及采样率的转化。

本实施例中，由于在播放设备播放该第一音频后，翻录第一音频的录制设备的可能性较多，特别是不同品牌的录音设备的录制音频文件格式不尽相同，其采样率一般为 44.1K，因此需要首先进行音频文件格式的转换和采样率的变换。以得到解析终端能够处理的格式和采样率。优选地，解析终端可以将第一音频的采样率转化为 48k。

602、解析终端从第一音频中解析音频水印。

本实施例中，解析终端可以对第一音频进行实时解析，也可以进行离线的解析。对此本申请实施例并不进行限定。为便于理解，本申请实施例主要对离线解析的方法进行说明，但并不构成对本方案的限定。

请参阅图 7，如图 7 所示，本申请所提供的音频水印解析方法从第一音频中解析音频水印包括以下步骤。

701、解析终端确定第一音频中满足第一预设条件的第一目标帧。

本实施例中，第一目标帧满足第一预设条件，因此解析终端可以按照第一预设条件来获取第一音频中的第一目标帧。需要说明的是，第一音频中包括多个第一目标帧，每个第一目标帧对应一个水印解析的周期，因此每当解析终端在第一音频中确定一个满足第一预设条件的第一目标帧，则执行一次后续解析步骤。该第一预设条件可以为：将低频部分最大值处于第一区间的目标帧作为第一目标帧，关于第一区间的具体实施方式可参阅上述步骤 301 的记载，此处不再赘述。

需要说明的是，第一目标帧中包括标记信息，解析终端需要进一步根据标记信息确定第一目标帧，其中，标记信息的实现方式根据第一音频的实际情况包含两种技术方案：一、第一音频的采样率小于第一阈值时，标记信息为特征声音。二、第一音频的采样率大于或等于第一阈值时，标记信息为同步帧标记。以下分别对此两种情况进行详细说明。

一、第一音频的采样率小于第一阈值时，标记信息为特征声音。

本实施例中，当解析终端检测到第一音频的采样率小于第一阈值时，即可判定在第一音频中，第一目标帧与第二目标帧中的标记信息为特征声音，此时，解析终端检测标记信息的具体方法为：

解析终端从第一音频中分别确定包含有第一特征声音，且第一特征声音的持续时间大于或等于预设时间的目标帧作为第一目标帧。

本实施例中，第一特征声音可以通过声音检测方法进行检测，其中，检测特征声音的方式可以为现有技术中的任意一种方式，本申请实施例对此并不进行限定。具体工作过程

中，第一特征声音可以为人声，例如，当检测到没有人声的时间持续大于预设时间后，再次检测到人声的时刻所在的目标帧，确定为第一目标帧。

进一步地，播放终端与解析终端之间针对特征声音还可以约定更加细化的实施方式，例如，在人声检测的状态下，当解析终端检测到特定语句时，才将该特定语句所在的目标帧确定为第一目标帧。

本实施例中，当第一音频的采样率小于第一阈值时，解析终端通过确定第一特征声音的方式来确定第一目标帧，从而通过特征声音的方式实现了水印嵌入起始位置（即第一目标帧）的确认，第一音频中不需要再嵌入同步帧标记。从而确保在第一音频中，无论第一音频的采样率大于还是小于第一阈值，均可以找到标记水印嵌入位置的第一目标帧。

二、第一音频的采样率大于第一阈值时，标记信息为同步帧标记。

本实施例中，当解析终端确定第一音频的采样率大于或等于第一阈值时，可判定，在第一音频中，第一目标帧中的标记信息为同步帧标记。解析终端解析同步帧标记的方法具体包括以下步骤。

1、解析终端逐帧获取第一音频中频部分与低频部分能量值的第一比值。

本实施例中，由于播放终端添加同步帧的方法为：提升中频部分第一采样点的能量值，从而提高中频部分与低频部分能量值的第一比值，因此，解析终端可以通过该第一比值来确定同步帧标记。

2、当解析终端获取到第一比值大于或等于第二阈值的初始目标帧时，从初始目标帧开始通过滑窗方式向后逐帧滑动检测第一音频，以获取每个滑动窗口内中频部分与低频部分能量值的第二比值。

本实施例中，解析终端逐帧获取每一帧中频部分采样点与低频部分采样点的能量值的第一比值，当找到第一比值大于或等于第二阈值第一采样点时，确定所述第一采样点所在的目标帧为初始目标帧。然而，由于一帧有 2048 个采样点，第一采样点仅为其中的部分采样点，因此，当找到了符合第一比值的第一采样点后，第一采样点实际所在的第一目标帧相对于当前第一采样点所在的初始目标帧之间，可能会存在偏移。为解决此问题，需要从初始目标帧开始，通过滑窗方式向后移动，逐帧检测初始目标帧之后的第一音频，以获取每个滑动窗口内中频部分与低频部分能量值的第二比值。

3、解析终端获取第二比值最大的滑动窗口所在帧为第一目标帧。

本实施例中，请参阅图 8 所示，初始目标帧 801 即解析终端生成的初始的滑动窗口 801，初始目标帧 801 与第一目标帧 802 之间存在交集，第一采样点 803 位于该交集的部分，解析终端需要将滑动窗口 801 与第一目标帧 802 之间完全重合，才能确定第一目标帧所在的位置。对此，解析终端的具体工作方式为：通过滑窗方式，检测每个滑动窗口 801 内中频部分与低频部分能量值的第二比值，其中，由于播放终端主动提升了第一目标帧 802 中频部分的能量，因此，第二比值达到最大值的窗口，即为第一目标帧所在的窗口，从而通过此种方式，实现了滑动窗口 801 与第一目标帧 802 的重合。从而通过滑窗检测的方式实现了对同步帧标记的查找，且有效的防止了查找过程中产生的偏移问题，提升了后续水印检测的精度。

需要说明的是，上述同步帧标记的检测方法，是检测第一音频的中频部分与低频部分能量值的第一比值来确定的，然而在实际工作过程中，第一音频的原始内容（即非水印内容）中也可能存在中频部分与低频部分能量值比值大于第一比值的情况。从而会造成同步帧标记的误检测。事实上，同步帧标记与第一音频中的原始内容相比，由于个本质的区别在于，同步帧标记中，中频部分的第一采样点能量值相对于低频部分有一个突然的增长点，在自然录音的状态下音频是不会有这样的突增的，因此，利用这种特性，当解析终端通过上述步骤 3 的方式确定了滑动窗口所在的目标帧为第一目标帧时，可以进一步通过以下步骤确定当前目标帧中的同步帧标记是否为真正的目标帧标记，从而防止误检测的情况发生。

4、解析终端从第二比值最大的滑动窗口中获取中频部分能量值最高的第一采样点。

本实施例中，第一采样点为当前窗口中能量值最高的点。

5、解析终端获取距离第一采样点之前预设长度的第三采样点。

本实施例中，第三采样点在时域上位于第一采样点之前，第三采样点距离第一采样点的预设长度可以由本领域技术人员根据实际需要来设定，也可以由解析终端根据采样率等参数自行确定，对此本申请实施例并不进行限定。

6、解析终端确定第一采样点能量值与第三采样点能量值的比值大于第七阈值的部分为同步帧标记。

本实施例中，请参阅图 9a 和图 9b，其中，图 9a 为加入了同步帧标记的第一目标帧，图 9b 为没有加入同步帧标记的普通目标帧。图 9a 和图 9b 两幅图中，中频部分与低频部分的能量比值均满足预设条件，因此该种情况下，仅通过中频与低频部分的能量比值，无法判断哪一个是加入了同步帧标记的目标帧，从而导致误检测的发生。对此通过上述步骤 4 至 6 所示的方法，假设在图 9a 中，第一采样点 901 与第三采样点 902 之间相隔 3 个采样点，可以看到第一采样点相对第三采样点的能量发生了突增，而在图 9b 中，未加入同步帧标记的波形图中，能量值的变化是平滑的，第一采样点 903 与第三采样点 904 之间相隔 3 个采样点的情况下，第一采样点 903 的能量值相对第三采样点 904 不会有明显的变化。因此，通过此种方法，能够准确的识别同步帧标记，防止误检测的发生。

本实施例中，通过上述特征声音的方式和同步帧标记的方式，无论第一音频的采样率大小如何，解析终端均能够在第一音频中检测到播放终端确定的第一目标帧。

702、解析终端在第一目标帧之后确定满足第二预设条件的第二目标帧。

本实施例中，当解析终端检测到第一目标帧后，以第一目标帧为定位帧，即可继续寻找位于第一目标帧之后的第二目标帧，第二目标帧嵌有音频水印，且第二目标帧满足第二预设条件。解析终端可以根据该第二预设条件，在第一目标帧之后快速地找到第二目标帧。

可选地，由于第一音频是多周期水印嵌入的，每个周期中包括一个第一目标帧和一个第二目标帧，因此解析终端在离线解析时，也可以从第一目标帧开始向前获取第二目标帧，为便于理解，本申请实施例仅以从第一目标帧开始向后移动寻找第二目标帧进行说明，但并不构成对本申请实施例方案的限定。

可选地，如图 10 所示，解析终端具体可以通过以下步骤确定第二目标帧。

1001、解析终端从第一目标帧开始以帧为单位向后移动，分别获取每帧中频部分的能

量大于或等于第三阈值且小于第四阈值的备选目标帧。

本实施例中，由于播放终端在第二目标帧中嵌入水印的位置为中频区域，如果中频区域能量过低，则导致嵌入水印后，容易导致误解析和无法检测到水印；如果中频区域能量过高，则导致嵌入水印后，产生爆音。因此根据这一特性，解析终端首先解析中频部分的能量大于或等于第三阈值且小于第四阈值的目标帧作为可能存在第二目标帧的备选目标帧。

1002、解析终端从备选目标帧中获取不同时域和/或不同频域部分能量值的能量比值大于或等于第五阈值的目标帧为第二目标帧。

本实施例中，由于播放终端通过改变第二目标帧中能量比值的方式来实现音频水印的嵌入，因此，根据播放终端嵌入音频水印的具体方式，解析终端从备选目标帧中获取不同时域和/或不同频域部分能量值的能量比值大于或等于第五阈值的目标帧即为第二目标帧。

本实施例中，通过上述步骤，解析终端确定了第二目标帧，接下来可以开始从第二目标帧中解析水印。

703、解析终端从第二目标帧中解析音频水印。

本实施例中，播放终端将音频水印嵌入在第二目标帧，因此，当解析终端获取到第二目标帧时，即可从该第二目标帧中解析所嵌入的音频水印。

请参阅图 11，如图 11 所示，可选地，解析终端通过以下步骤从第二目标帧中解析音频水印。

1101、解析终端从第二目标帧中获取第二采样点。

本实施例中，第二采样点为第二目标帧中能量比值大于或等于第五阈值的采样点，解析终端从第二目标帧中获取第二采样点，该第二采样点中记录有播放终端嵌入的水印信息。

1102、解析终端分别获取第二采样点中取不同时域和/或不同频域部分能量值的能量比值。

本实施例中，播放终端在嵌入音频水印时，对于第二采样点中能量比值的改变可以采用以下三种不同的方案：一、改变不同时域部分能量值的比值；二、改变不同频域部分能量值的比值；三、同时改变不同时域和不同频域部分能量值的比值。因此，根据不同的水印嵌入方式，解析终端需要采用相应的手段进行解析。为便于理解，以下结合附图对此三种方式做详细的说明。

一、改变不同时域部分能量值的比值。

本实施例中，播放终端在嵌入音频水印时，通过改变不同时域部分能量值的比值的方式来进行，具体地，第二采样点后半部分的能量值与前半部分的能量值的比值大于第五阈值时，约定该比值所对应的数字为 0，第二采样点前半部分的能量值与后半部分的能量值的比值大于第五阈值时，约定该比值所对应的数字为 1。

基于上述水印嵌入规则，请参阅图 12a 和图 12b，解析终端所获取的第二目标帧的能量分部图分别为如图 12a 和图 12b 所示的两种情况。在能量分布图中，深色部分表示高能量区域，白色部分表示低能量区域，从而通过能量分部图，解析终端可以直观地获取到第二目标帧的能量分部。其中，对于图 12a 所示的能量分部图，第二目标帧中后半部分 1201 的能量值显著大于前半部分 1202 的能量值，且经过解析终端的计算，第二采样点后半部分

1201 的能量值与前半部分 1202 的能量值的比值大于第五阈值时，则此时解析终端将图 12a 所示的第二目标帧所嵌入的水印判定为 0。对于图 12b 所示的能量分部图，第二目标帧前半部分 1203 的能量值显著大于后半部分 1204 的能量值，且经过解析终端的计算，第二采样点前半部分 1203 的能量值与后半部分 1204 的能量值的比值大于第五阈值时，则此时解析终端将图 12b 所示的第二目标帧所嵌入的水印判定为 1。

需要说明的是，在上述解析的过程中，该能量比值的确定，是以目标帧在时域上的中点为界限，获取前后部分的比值后按照上述预设规则解析的。然而，在实际工作过程中，在对第一音频进行翻录时，由于场地的回声或混响等原因，会造成高能量部分在时域上增加。例如，音频水印为 1 时，时域上以中点为界限，前半部分的能量值与后半部分能量值的比值大于或等于第五预设值。然而，由于转录过程中回声或混响的存在，导致高能量区域在时域上越过了中点的界限，从而影响了能量比值的大小，造成解析终端无法根据能量比值获取到水印。为解决上述问题，播放终端在嵌入水印时进一步提供一种依据频域来改变能量比值的方法，以克服转录过程中所产生的回声混响对时域上的能量分部造成的影响。

二、改变不同频域部分能量值的比值。

本实施例中，播放终端在嵌入音频水印时采用与上述类似的方式，改变第二采样点在频域上的的能量比值，从而实现水印的嵌入，具体的实现方式可参阅上述改变不同时域部分能量比值的方法，区别在于，本种实施方式并不是在时域上以时间作为划界，而是在频域上以频率作为划界。

因此，解析终端在解析音频水印时，所得到第二目标帧的能量分部图如图 5g 所示，在第二采样点中以频域的中点划界，在能量分布图中，深色部分表示高能量区域，白色部分表示低能量区域，从而通过能量分部图，解析终端可以直观地获取到第二目标帧的能量分部。如图 5g 所示，上半部分与下半部分的比值大于第五阈值的表示数字 1。下半部分与上半部分的比值大于第五阈值的表示数字 0，从而通过改变频域上的能量比值实现了音频水印的嵌入，克服了转录过程中所产生的回声混响对时域上的能量分部造成的影响。

三、同时改变不同时域和不同频域部分能量值的比值。

本实施例中，播放终端在嵌入音频水印时综合了上述方案一和方案二的方案，首先通过方案一的方法，调整第二采样点在时域上前后部分的能量比值，实现第一步的水印嵌入，例如图 5h 所示，第二采样点前半部分的能量值与后半部分能量值的比值大于第五阈值，嵌入了音频水印为数字 1，进一步地，为了防止转录过程中所产生的回声混响在时域上对音频水印能量分部造成的影响，如图 5h 所示，对于图 5h 的后半部分（即能量较低的部分），采取上述方案二的方法，改变高频部分与低频部分的能量比值。

方案三所提供的方式，在第二采样点中同时加入了一大一小两个水印。该一大一小两个水印分别记录了相同的数值， 其中，大水印为在时域上改变第二采样点前后两部分的能量比值所得到的水印，小水印为改变第二采样点低能量部分高低频的能量比值所得到的水印。

进一步地，对于如图 5h 所示的水印，为了防止转录过程中所产生的回声混响对时域上的能量分部造成的影响，还可以对音频水印中的低能量部分进行进一步的分割，如图 5i 所

示，对于第二采样点中低能量部分嵌入的小水印时，将该低能量部分在时域上分为两部分，分别记为第一部分和第二部分，其中，第一部分为接近高能量部分的区域，第二部分为远离高能量部分的区域。在嵌入小水印时，仅仅改变第二部分高频部分与低频部分的能量比值，对于第一部分，由于靠近高能量部分，可能会由于转录过程中产生的混响导致第一部分的能量相应升高，因此不再改变第一部分的高低频能量比值，从而排除了转录过程中所产生的回声混响对时域上的能量分部造成的影响，仅仅在远离高能量部分的第二部分通过改变高低频能量比值的方式嵌入小水印。

需要说明的是，上述图 5i 所示的音频水印，对于低能量部分，在时域上分为两个部分进行处理，在实际工作过程中，本领域技术人员根据实际需要，可以将该低能量部分在时域上分割为更多的部分，对此本申请实施例并不进行限定。

对于播放终端按照方案三所嵌入的音频水印，解析终端在解析时首先解析大水印，大水印的解析方式可参阅上述方案一，之后解析大水印中的小水印，小水印的具体解析方法可参阅上述方案二，大水印与小水印中所解析出的数字一致时，判定水印解析正确，解析终端获取到当前第二目标帧中所嵌入的水印。

需要说明的是，上述方案一至方案三的方案中，播放终端通过改变能量比值的方式实现了二进制的水印嵌入，在实际工作过程中，本领域技术人员可根据实际需要，通过改变能量比值实现其他进制的水印嵌入，例如十进制或 12 进制，对此本申请实施例并不进行限定。

本实施例中，对于播放终端通过不同方式在第一音频中嵌入的水印，解析终端通过相应的方式进行解析，通过本申请实施例所提供的方式，音频水印具有较好的抗干扰能力，音频水印在转录的过程中不易因为空气录制产生损失，同时解析终端能够准确地解析出播放终端所嵌入的音频水印，使得音频水印的嵌入方案具有较好的稳定性和准确性。

需要说明的是，以音频水印为数字“14”为例，音频水印所对应的第一数列为：“1110”。解析终端从第二目标帧解析音频水印时，按照上述步骤 703 所记载的方式，先解析第一个第二采样点中：第一数列中的第一个元素：“1”实现了第一数列中第一位数的解析。完成了水印解析周期中的第一子周期。之后循环执行上述步骤 703，从第二目标帧中获取第二个第二采样点，按照同样方式解析第一数列中的第二个元素：“1”，实现了第一数列中第二位数的解析。……以此类推，播放终端解析四个水印嵌入子周期所组成的完整水印解析周期，实现了对音频水印的解析。

可选地，解析终端在执确定第一目标帧时，由于起始位的偏差和录制采样点丢失会带来的偏差。这些偏差使得后续的第二目标帧也会发生相应的偏差，导致第二目标帧中嵌入的音频水印受到下一帧能量的影响。因此解析终端需要对第二目标帧进行预处理，防止由于下一帧能量变化带来的影响，因此将第二目标帧两侧的能量去除，去除的长度可以由本领域技术人员根据实际需要决定，例如前后各剪裁 8 个采样点，也可以由解析终端按照预设逻辑来决定，例如由解析终端根据检测周期的长度来调节剪裁的长度。对此本申请实施例并不进行限定。最终得到的第二目标帧前后部分被减去了预设长度，仅保留中间的部分，从而排除了由于偏差带来的干扰。

本实施例中，通过上述方式，解析终端从第一音频中解析出了音频水印。

1103、解析终端对音频水印进行校验。

本实施例中，作为一种优选的实施方式，音频水印所嵌入的第一数列中包含一个校验位，解析终端根据该校验位对第一数列的完整性进行校验，以确保水印解析的准确性，例如，如上述举例，第一数列为“11101”，播放终端与解析终端之间约定采用奇偶校验的方式，第一数列中的最后一位为校验位，最后一位的数字为 1 时，代表第一数列中除校验位外包括奇数个数字 1。从而解析终端在解析出第一数列后，根据第一数列最后一位的记载，确定所解析出的第一数列中数字 1 的数量是否为奇数个，即可确定当前第一数列的解析是否完整。

1104、当解析终端完成了对整个第一音频的解析时，解析终端从多个水印检测周期所解析的音频水印中确定重复率最高的一个作为第一音频的水印。

本实施例中，第一音频中包括多组第一目标帧+第二目标帧，其中，每个第一目标帧与第二目标帧组成了一个水印检测周期，每个水印检测周期中均嵌入有相同的音频水印。然而在实际解析的过程中，解析终端可能会出现一些解析错误的情况，导致不是所有水印检测周期解析得到的音频水印都是相同的数列，通过实验观察，当解析出错时，误解析所得到的错误音频水印总是随机且不重复的，因此多个水印检测周期所解析的音频水印中重复率最高的一个，可以确定为正确的音频水印。因此通过这种多周期决策的方式，结合前述奇偶校验的校验方式，准确地解析出第一音频中所嵌入的正确水印，进一步防止解析终端的误解析。

可选地，当解析终端确定第一数列传输完整后，可以根据需要，对第一数列进行进制的转换，例如将第一数列由二进制转化为十进制，最后得到数字 14。从而最终完成了对于第一音频中音频水印的校验。

603、解析终端根据音频水印确定播放终端。

本实施例中，由于音频水印与播放终端相关联，因此根据从第一音频中解析的音频水印，即可知晓当前这段第一音频的音频水印是由哪一个播放终端添加的。例如实时会议场景中，会场 A、会场 B 和会场 C 的播放终端播放的音频内容完全相同，但是各个播放终端在播放相同的音频时各自嵌入的音频水印不同，解析终端通过解析音频水印即可知晓当前第一音频是由哪一个会场的播放终端播放的，从而实现了第一音频的溯源。

需要说明的是，本申请实施例所提供的音频水印添加方法和音频水印解析方法可用于各种不同的使用场景中，为便于理解，以下结合附图，对本申请实施例所提供方法的使用场景进行说明。

一、远程会议场景。

本实施例中的架构如图 13 所示，在如图 1 所示的架构基础上，多出了一个服务管理中心（service management center，SMC），执行以下步骤。

1301、SMC 向播放终端发送音频水印。

本实施例中，播放终端的数量可以为多个，SMC 发送的音频水印与各个播放终端一一对应，用于唯一地标记各个播放终端。

1302、播放终端将音频水印存储在本地。

本实施例中，各个播放终端在获取到音频水印后，将音频水印存储在本地，以便后续在获取到实时音频流时将音频水印嵌入音频流。

1303、媒体中心向播放终端发送第一音频。

本实施例中，媒体中心的音频流可以是由某个会场的播放终端生成后发送给媒体中心MCU的，之后媒体中心将该音频流实时发送给各个其他会场的播放终端。

1304、播放终端在第一音频中实时嵌入音频水印。

本实施例中，播放终端通过上述任意一个实施例所提供的音频水印添加方法将存储在本地的音频水印实时嵌入第一音频中。具体可参阅上述记载，此处不再赘述。

1305、播放终端播放第一音频。

本实施例中，播放终端播放的第一音频嵌入有音频水印，整个水印嵌入过程实时进行，不会影响第一音频的直播效果，同时所播放的第一音频可以根据音频水印溯源第一音频的播放终端。

1306、解析终端获取第一音频。

本实施例中，解析终端可以是通过数字信道获取到第一音频，也可以是通过空气信道获取到该第一音频，对于此两种方式所转录得到的第一音频，解析终端均能够进行解析。

1307、解析终端从第一音频中解析音频水印。

本实施例中，解析终端通过上述任意一个实施例所提供的音频解析添加方法从第一音频中解析音频水印。具体可参阅上述记载，此处不再赘述。

1308、解析终端根据音频水印确定播放终端。

本实施例中，由于音频水印与播放终端相关联，因此根据从第一音频中解析的音频水印，即可知晓当前这段第一音频的音频水印是由哪一个播放终端添加的。从而实现了第一音频的溯源。

本实施例中，该SMC分别向会场A、会场B和会场C所在的播放终端分配不同的会场标识（即音频水印），这些会场标识与会场的播放终端向关联，用于唯一地标识每个会场的播放终端。会场A、会场B和会场C的播放终端从媒体中心MCU获取到直播的音频流后执行本申请实施例所提供的音频水印添加方法，将各自获取的会场标识作为水印实时添加到播放音频流中，以使得各个会场的播放终端所播放的音频中嵌入有音频水印。解析终端根据该被转录视频中的音频水印，能够通过音频流中水印信息所记录的会场标识，确定音频具体为哪一会场的观众转录。从而实现了对音频水印的溯源。

二、云点播场景。

本实施例中，请参阅图14，如图14所示，用户通过终端向云服务器发送点播信息，点播所需要观看的音频或视频内容，云服务根据用户的点播信息向用户所在的终端实时地发送点播内容。具体步骤如下。

1401、用户终端根据用户选择点播的内容生成点播信息。

本实施例中，用户通过用户终端的交互界面选择需要点播的音频或视频内容，并生成点播信息，该点播信息用于记录用户点播的音频或视频内容。

1402、用户终端将点播信息发送给云服务器。

本实施例中，用户终端将点播信息发送给云服务器以使得云服务器知晓用户所需点播的内容。

1403、云服务器根据点播信息获取用户点播的目标内容。

本实施例中，云服务器根据用户的点播信息从数据库中获取用户所点播的目标内容。

1404、云服务器根据用户终端的终端标识生成音频水印。

本实施例中，用户终端为音频或视频的播放终端，该音频水印与用户终端相关联，用于唯一地该用户终端。从而实现对音频水印的获取。

1405、云服务器将音频水印嵌入目标内容中。

本实施例中，目标内容可以为音频或视频；若为视频，则将音频水印嵌入该视频的音频内容中，云服务器将音频水印嵌入目标内容中的具体方式可以为本申请实施例所提供的任意一种音频水印添加方法，具体可参阅前述记载，需要说明的是，在本种情况中，该水印嵌入的步骤可以由云服务器来执行，也可以由云服务器将音频水印发给用户终端，有用户终端来执行。对此本申请实施例并不进行限定。通过本申请实施例所提供的方法，云服务器能够一边向用户终端传输目标内容，一边实时地在目标内容的音频中嵌入音频水印，从而提升了工作效率。

1406、云服务器将水印内容发送给用户终端。

本实施例中，水印内容中的音频或视频内容为用户终端的用户所点播的内容，同时该水印内容中已经添加了音频水印。

1407、用户终端播放水印内容。

本实施例中是，由于用户终端播放的内容中已经添加了音频水印，因此若用户终端的用户再翻录该终端播放的内容，翻录的内容中会保留有该音频水印，从而可以溯源到翻录该内容的用户终端。

1408、解析终端获取水印内容。

本实施例中，水印内容可以是嵌入有水印的音频，或者带有音频的视频，该视频中的音频嵌入有水印，解析终端可以是通过数字信道获取到水印内容，也可以是通过空气信道获取到该水印内容，对于此两种方式所转录得到水印内容中的第一音频，解析终端均能够进行解析。

1409、解析终端从第一音频中解析音频水印。

本实施例中，解析终端通过上述任意一个实施例所提供的音频解析添加方法从第一音频中解析音频水印。具体可参阅上述记载，此处不再赘述。

1410、解析终端根据音频水印确定用户终端。

本实施例中，由于音频水印与用户终端相关联，因此根据从第一音频中解析的音频水印，即可知晓当前这段第一音频的音频水印是由哪一个用户终端添加的。从而实现了第一音频的溯源。

综上所述，本申请实施例所提供的水印添加方法和水印解析方法可以应用于各种不同的有音频水印添加和解析需求的场景中，上述两种方式只是一种举例，并不构成对本申请

实施例使用场景的限定。

从硬件结构上来描述，上述方法可以由一个实体设备实现，也可以由多个实体设备共同实现，还可以是一个实体设备内的一个逻辑功能模块，本申请实施例对此不作具体限定。

例如，上述方法可以通过图 15 中的电子设备来实现。图 15 为本申请实施例提供的一种电子设备的硬件结构示意图；该电子设备可以是本发明实施例中的播放终端或解析终端，该电子设备包括至少一个处理器 1501，通信线路 1502，存储器 1503 以及至少一个通信接口 1504。

处理器 1501 可以是一个通用中央处理器（central processing unit，CPU），微处理器，特定应用集成电路（application-specific integrated circuit，服务器 IC），或一个或多个用于控制本申请方案程序执行的集成电路。

通信线路 1502 可包括一通路，在上述组件之间传送信息。

通信接口 1504，使用任何收发器一类的装置，用于与其他设备或通信网络通信，如以太网，无线接入网（radio access network，RAN），无线局域网(wireless local area networks，WLAN)等。

存储器 1503 可以是只读存储器（read-only memory，ROM）或可存储静态信息和指令的其他类型的静态存储设备，随机存取存储器（random access memory，RAM）或者可存储信息和指令的其他类型的动态存储设备，也可以是电可擦可编程只读存储器（electrically erable programmable read-only memory，EEPROM）、只读光盘（compact disc read-only memory，CD-ROM）或其他光盘存储、光碟存储（包括压缩光碟、激光碟、光碟、数字通用光碟、蓝光光碟等）、磁盘存储介质或者其他磁存储设备、或者能够用于携带或存储具有指令或数据结构形式的期望的程序代码并能够由计算机存取的任何其他介质，但不限于此。存储器可以是独立存在，通过通信线路 1502 与处理器相连接。存储器也可以和处理器集成在一起。

其中，存储器 1503 用于存储执行本申请方案的计算机执行指令，并由处理器 1501 来控制执行。处理器 1501 用于执行存储器 1503 中存储的计算机执行指令，从而实现本申请下述实施例提供的计费管理的方法。

可选的，本申请实施例中的计算机执行指令也可以称之为应用程序代码，本申请实施例对此不作具体限定。

在具体实现中，作为一种实施例，处理器 1501 可以包括一个或多个 CPU，例如图 15 中的 CPU0 和 CPU1。

在具体实现中，作为一种实施例，电子设备可以包括多个处理器，例如图 15 中的处理器 1501 和处理器 1505。这些处理器中的每一个可以是一个单核（single-CPU）处理器，也可以是一个多核（multi-CPU）处理器。这里的处理器可以指一个或多个设备、电路、和/或用于处理数据（例如计算机程序指令）的处理核。

在具体实现中，作为一种实施例，电子设备还可以包括输出设备 1505 和输入设备 1506。输出设备 1505 和处理器 1501 通信，可以以多种方式来显示信息。例如，输出设备 1505 可以是液晶显示器（liquid crystal display，LCD），发光二级管（light emitting diode，LED)显示设备，阴极射线管（cathode ray tube，CRT）显示设备，或投影仪（projector）等。输入

设备 1506 和处理器 1501 通信，可以以多种方式接收用户的输入。例如，输入设备 1506 可以是鼠标、键盘、触摸屏设备或传感设备等。

上述的电子设备可以是一个通用设备或者是一个专用设备。在具体实现中，电子设备可以服务器、无线终端设备、嵌入式设备或有图 15 中类似结构的设备。本申请实施例不限定电子设备的类型。

本申请实施例可以根据上述方法示例对电子设备进行功能单元的划分，例如，可以对应各个功能划分各个功能单元，也可以将两个或两个以上的功能集成在一个处理单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。需要说明的是，本申请实施例中对单元的划分是示意性的，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式。

比如，以采用集成的方式划分各个功能单元的情况下，图 16 示出了本申请实施例所提供的一种播放终端的结构示意图。

如图 16 所示，本申请实施例所提供的播放终端包括。

获取单元 1601，用于实时获取第一音频；

执行单元 1602，用于在所述获取单元 1601 获取的所述第一音频中嵌入音频水印，所述音频水印与所述播放终端相关联；

播放单元 1603，用于播放由所述执行单元 1602 嵌有所述音频水印的所述第一音频。

可选地，该执行单元 1602，还用于：

在该第一音频中确定满足第一预设条件的第一目标帧；

在该第一目标帧之后确定满足第二预设条件的第二目标帧，该第一目标帧用于标记该第二目标帧；

在该第二目标帧中嵌入该音频水印。

可选地，该执行单元 1602，还用于：

当该第一音频的采样率大于或等于第一阈值时，将低频部分最大值处于第一区间内的音频帧确定为该第一目标帧；或者，

当该第一音频的采样率小于该第一阈值时，确定包含第一特征声音的音频帧为该第一目标帧。

可选地，当该第一音频的采样率大于或等于第一阈值时，该执行单元 1602，还用于：

在该第一目标帧中添加同步帧标记。

可选地，该执行单元 1602，还用于：

获取第一采样点，该第一采样点为中频部分的采样点；

提升该第一采样点的能量值，以使得该第一采样点的能量值与低频部分能量值的比值大于或等于第二阈值。

可选地，该当该第一音频的采样率小于该第一阈值时，该执行单元 1602，还用于：

当检测到该第一特征声音，且该第一特征声音的持续时间大于或等于预设时间时，将包含该第一特征声音的音频帧确定为该第一目标帧。

可选地，该执行单元 1602，还用于：

确定中频部分能量值大于或等于第三阈值且小于第四阈值的目标帧为该第二目标帧。

可选地，该执行单元 1602，还用于：

获取该音频水印所对应的第一数列，该第一数列中包括至少一个元素；

从该第三目标帧中获取至少一个第二采样点；

将该第一数列中的至少一个元素分别嵌入该至少一个第二采样点中，其中，该第一数列中的一个元素对应一个第二采样点。

可选地，该执行单元 1602，还用于：

调节该第二采样点在不同时域和/或不同频域部分能量值的能量比值，其中，一个该第二采样点的该能量比值与该第一数列中的一个元素相关联。

如图 17 所示，本申请实施例所提供的解析终端包括。

获取单元 1701，用于获取第一音频，该第一音频中嵌有音频水印，该音频水印与播放终端相关联，该播放终端用于将该音频水印实时嵌入该第一音频；

解析单元 1702，用于从该获取单元 1701 获取的该第一音频中解析该音频水印；

执行单元 1703，用于根据该解析单元 1702 解析的该音频水印确定该播放终端。

可选地，该解析单元 1702，还用于：

确定该第一音频中满足第一预设条件的第一目标帧；

在该第一目标帧之后确定满足第二预设条件的第二目标帧；

从该第二目标帧中解析该音频水印。

可选地，当该第一音频的采样率小于第一阈值时，该解析单元 1702，还用于：

从该第一音频中确定包含有第一特征声音，且该第一特征声音的持续时间大于或等于预设时间的目标帧作为该第一目标帧。

可选地，当该第一音频的采样率大于或等于第一阈值时，该解析单元 1702，还用于：

逐帧获取该第一音频中频部分与低频部分能量值的第一比值；

当获取到该第一比值大于或等于第二阈值的初始目标帧时，从该初始目标帧开始通过滑窗方式向后滑动检测该第一音频，以获取每个滑动窗口内中频部分与低频部分能量值的第二比值；

获取该第二比值最大的滑动窗口所在帧为该第一目标帧。

可选地，该第一目标帧中包括同步帧标记，该解析单元 1702，还用于：

从该第二比值最大的滑动窗口中获取中频部分能量值最高的第一采样点；

获取距离该第一采样点之前预设长度的第三采样点；

确定该第一采样点能量值与该第三采样点能量值的比值大于第七阈值的部分为该同步帧标记；

根据该同步帧标记确定该第二比值最大的滑动窗口所在帧为该第一目标帧。

可选地，该解析单元 1702，还用于：

从该第一目标帧开始以帧为单位向后移动，分别获取每帧中频部分的能量大于或等于第三阈值且小于第四阈值的备选目标帧；

从该备选目标帧中获取不同时域和/或不同频域部分能量值的能量比值大于或等于第

五阈值的目标帧为该第二目标帧。

可选地，该解析单元 1702，还用于：

从该第二目标帧中获取第二采样点，该第二采样点为该第二目标帧中该能量比值大于或等于该第五阈值的采样点；

分别获取所第二采样点中取不同时域和/或不同频域部分能量值的能量比值；

获取与该能量比值相关联的第一元素，该第一元素为该音频水印所记录的第一数列中的一个元素。

可选地，该第一音频中包括多个水印检测周期，其中，每个该水印检测周期分别解析出一个该音频水印，该解析单元 1702，还用于：

从该多个水印检测周期所解析的音频水印中确定重复率最高的一个作为该第一音频的水印。

本领域技术人员应该可以意识到，在上述一个或多个示例中，本发明所描述的功能可以用硬件、软件、固件或它们的任意组合来实现。当使用软件实现时，可以将这些功能存储在计算机可读介质中或者作为计算机可读介质上的一个或多个指令或代码进行传输。计算机可读介质包括计算机存储介质和通信介质，其中通信介质包括便于从一个地方向另一个地方传送计算机程序的任何介质。存储介质可以是通用或专用计算机能够存取的任何可用介质。

以上所述的具体实施方式，对本发明的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上所述仅为本发明的具体实施方式而已。

以上所述，以上实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的范围。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统，装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

在本申请所提供的几个实施例中，应该理解到，所揭露的通信方法、中继设备、宿主基站及计算机存储介质，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（英文全称：Read-Only Memory，英文缩写：ROM）、随机存取存储器（英文全称：Random Access Memory，英文缩写：RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

以上所述，以上实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围。

权 利 要 求

1.一种音频水印添加方法，其特征在于，包括：

播放终端实时获取第一音频；

所述播放终端在所述第一音频中嵌入音频水印，所述音频水印与所述播放终端相关联；

所述播放终端播放嵌有所述音频水印的所述第一音频。

2.根据权利要求 1 所述的方法，其特征在于，所述播放终端在所述第一音频中嵌入音频水印包括：

所述播放终端在所述第一音频中确定满足第一预设条件的第一目标帧；

所述播放终端在所述第一目标帧之后确定满足第二预设条件的第二目标帧，所述第一目标帧用于标记所述第二目标帧；

所述播放终端在所述第二目标帧中嵌入所述音频水印。

3.根据权利要求 2 所述的方法，其特征在于，所述播放终端在所述第一音频中确定满足第一预设条件的第一目标帧，包括：

当所述第一音频的采样率大于或等于第一阈值时，所述播放终端将低频部分最大值处于第一区间内的音频帧确定为所述第一目标帧；或者，

当所述第一音频的采样率小于所述第一阈值时，所述播放终端确定包含第一特征声音的音频帧为所述第一目标帧。

4.根据权利要求 3 所述的方法，其特征在于，当所述第一音频的采样率大于或等于第一阈值时，所述播放终端将低频部分最大值处于第一区间内的音频帧作为所述第一目标帧之后，还包括：

所述播放终端在所述第一目标帧中添加同步帧标记。

5.根据权利要求 3 所述的方法，其特征在于，所述播放终端在所述第一目标帧中添加同步帧标记，包括：

所述播放终端获取第一采样点，所述第一采样点为中频部分的采样点；

所述播放终端提升所述第一采样点的能量值，以使得所述第一采样点的能量值与低频部分能量值的比值大于或等于第二阈值。

6.根据权利要求 3 所述的方法，其特征在于，所述当所述第一音频的采样率小于所述第一阈值时，所述播放终端确定包含第一特征声音的音频帧为所述第一目标帧，包括：

当检测到所述第一特征声音，且所述第一特征声音的持续时间大于或等于预设时间时，所述播放终端将包含所述第一特征声音的音频帧确定为所述第一目标帧。

7.根据权利要求 2 至 6 任一所述的方法，其特征在于，所述播放终端在所述第一目标帧之后确定满足第二预设条件的第二目标帧，包括：

所述播放终端确定中频部分能量值大于或等于第三阈值且小于第四阈值的目标帧为所述第二目标帧。

8.根据权利要求 2 至 7 任一所述的方法，其特征在于，所述播放终端在所述第三目标帧中嵌入所述音频水印，包括：

所述播放终端获取所述音频水印所对应的第一数列，所述第一数列中包括至少一个元

素；

所述播放终端从所述第三目标帧中获取至少一个第二采样点；

所述播放终端将所述第一数列中的至少一个元素分别嵌入所述至少一个第二采样点中，其中，所述第一数列中的一个元素对应一个第二采样点。

9.根据权利要求 8 所述的方法，其特征在于，所述播放终端将所述第一数列中的至少一个元素分别加入所述至少一个第二采样点中，包括：

所述播放终端调节所述第二采样点在不同时域和/或不同频域部分能量值的能量比值，其中，一个所述第二采样点的所述能量比值与所述第一数列中的一个元素相关联。

10.一种音频水印解析方法，其特征在于，包括：

解析终端获取第一音频，所述第一音频中嵌有音频水印，所述音频水印与播放终端相关联，所述播放终端用于将所述音频水印实时嵌入所述第一音频；

所述解析终端从所述第一音频中解析所述音频水印；

所述解析终端根据所述音频水印确定所述播放终端。

11.根据权利要求 10 所述的方法，其特征在于，所述解析终端从所述第一音频中解析所述音频水印之前，还包括：

所述解析终端确定所述第一音频中满足第一预设条件的第一目标帧；

所述解析终端在所述第一目标帧之后确定满足第二预设条件的第二目标帧；

所述解析终端从所述第一音频中解析所述音频水印，包括：

所述解析终端从所述第二目标帧中解析所述音频水印。

12.根据权利要求 11 所述的方法，其特征在于，当所述第一音频的采样率小于第一阈值时，所述解析终端确定所述第一音频中满足第一预设条件的第一目标帧，包括：

所述解析终端从所述第一音频中确定包含有第一特征声音，且所述第一特征声音的持续时间大于或等于预设时间的目标帧作为所述第一目标帧。

13.根据权利要求 11 所述的方法，其特征在于，当所述第一音频的采样率大于或等于第一阈值时，所述解析终端确定所述第一音频中满足第一预设条件的第一目标帧；，包括：

所述解析终端逐帧获取所述第一音频中频部分与低频部分能量值的第一比值；

当所述解析终端获取到所述第一比值大于或等于第二阈值的初始目标帧时，从所述初始目标帧开始通过滑窗方式向后滑动检测所述第一音频，以获取每个滑动窗口内中频部分与低频部分能量值的第二比值；

所述解析终端获取所述第二比值最大的滑动窗口所在帧为所述第一目标帧。

14.根据权利要求 13 所述的方法，其特征在于，所述第一目标帧中包括同步帧标记，所述解析终端获取所述第二比值最大的滑动窗口所在帧为所述第一目标帧，包括：

所述解析终端从所述第二比值最大的滑动窗口中获取中频部分能量值最高的第一采样点；

所述解析终端获取距离所述第一采样点之前预设长度的第三采样点；

所述解析终端确定所述第一采样点能量值与所述第三采样点能量值的比值大于第七阈值的部分为所述同步帧标记；

所述解析终端根据所述同步帧标记确定所述第二比值最大的滑动窗口所在帧为所述第一目标帧。

15.根据权利要求 11 至 14 任一所述的方法，其特征在于，所述解析终端在所述第一目标帧之后确定满足第二预设条件的第二目标帧，包括：

所述解析终端从所述第一目标帧开始以帧为单位向后移动，分别获取每帧中频部分的能量大于或等于第三阈值且小于第四阈值的备选目标帧；

所述解析终端从所述备选目标帧中获取不同时域和/或不同频域部分能量值的能量比值大于或等于第五阈值的目标帧为所述第二目标帧。

16.根据权利要求 15 所述的方法，其特征在于，所述解析终端从所述第二目标帧中解析所述音频水印，包括：

所述解析终端从所述第二目标帧中获取第二采样点，所述第二采样点为所述第二目标帧中所述能量比值大于或等于所述第五阈值的采样点；

所述解析终端分别获取所第二采样点中取不同时域和/或不同频域部分能量值的能量比值；

所述解析终端获取与所述能量比值相关联的第一元素，所述第一元素为所述音频水印所记录的第一数列中的一个元素。

17.根据权利要求 10 至 16 任一所述的方法，其特征在于，所述方法包括多个水印检测周期，其中，每个所述水印检测周期分别解析出一个所述音频水印，所述方法还包括：

所述解析终端从所述多个水印检测周期所解析的音频水印中确定重复率最高的一个作为所述第一音频的水印。

18.一种播放终端，其特征在于，包括：

获取单元，用于实时获取第一音频；

执行单元，用于在所述获取单元获取的所述第一音频中嵌入音频水印，所述音频水印与所述播放终端相关联；

播放终端，用于播放由所述执行单元嵌有所述音频水印的所述第一音频。

19.一种解析终端，其特征在于，包括：

获取单元，用于获取第一音频，所述第一音频中嵌有音频水印，所述音频水印与播放终端相关联，所述播放终端用于将所述音频水印实时嵌入所述第一音频；

解析单元，用于从所述获取单元获取的所述第一音频中解析所述音频水印；

执行单元，用于根据所述解析单元解析的所述音频水印确定所述播放终端。

20.一种电子设备，其特征在于，所述电子设备包括：交互装置、输入/输出（I/O）接口、处理器和存储器，所述存储器中存储有程序指令；

所述交互装置用于获取用户输入的操作指令；

所述处理器用于执行存储器中存储的程序指令，执行如权利要求 1 至 9 或 10 至 17 任一所述的方法。

21.一种计算机可读存储介质，包括指令，其特征在于，当所述指令在计算机设备上运行时，使得所述计算机设备执行如权利要求 1 至 9 或 10 至 17 任一所述的方法。

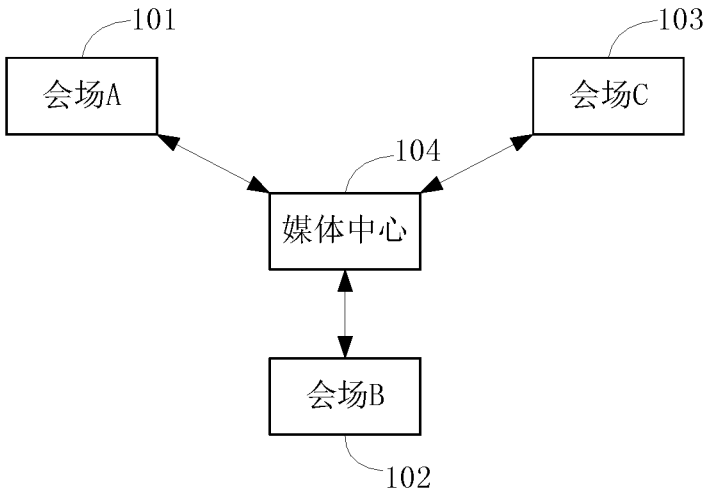


图 1

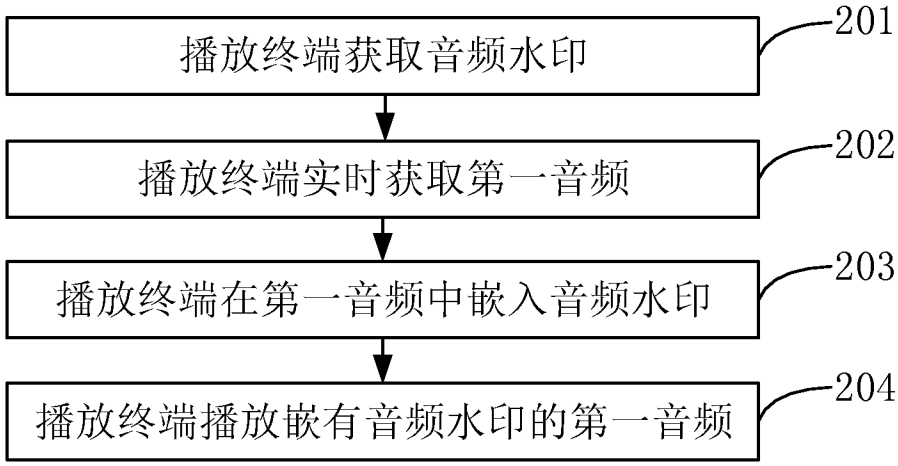

播放终端获取音频水印
201
播放终端实时获取第一音频
202
播放终端在第一音频中嵌入音频水印
203
播放终端播放嵌有音频水印的第一音频
204

图 2

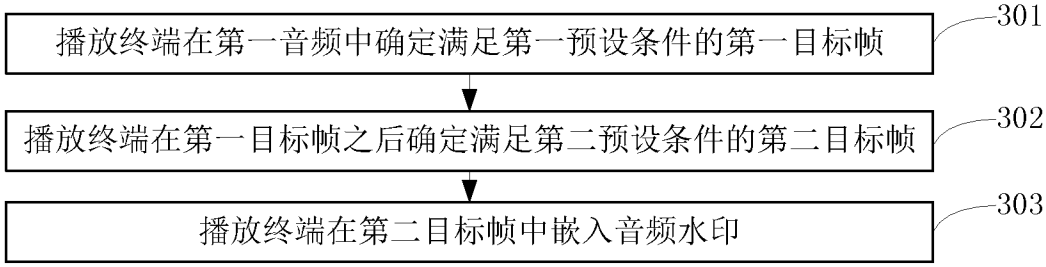

播放终端在第一音频中确定满足第一预设条件的第一目标帧
301
播放终端在第一目标帧之后确定满足第二预设条件的第二目标帧
302
播放终端在第二目标帧中嵌入音频水印
303

图 3

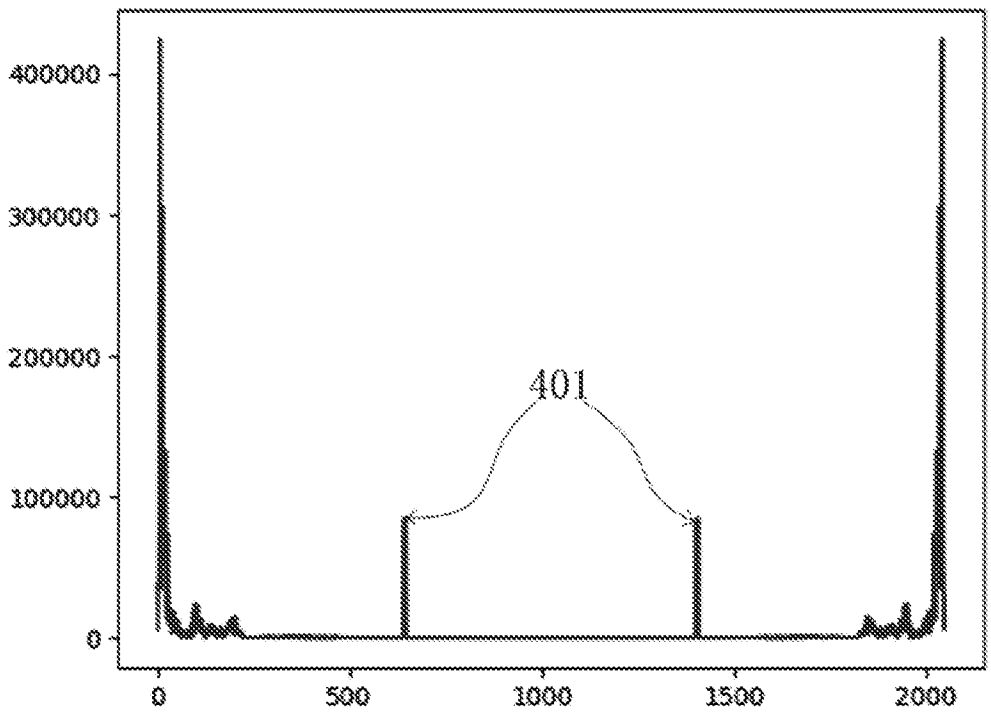

400000
300000
200000
100000
0
401
0
500
1000
1500
2000

图 4

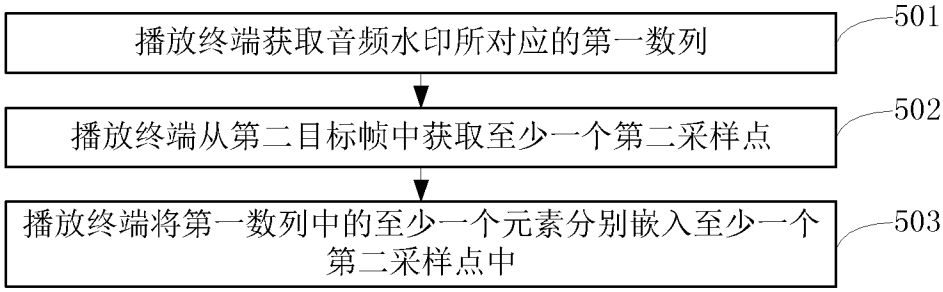

播放终端获取音频水印所对应的第一数列
501
播放终端从第二目标帧中获取至少一个第二采样点
502
播放终端将第一数列中的至少一个元素分别嵌入至少一个第二采样点中
503

图 5a

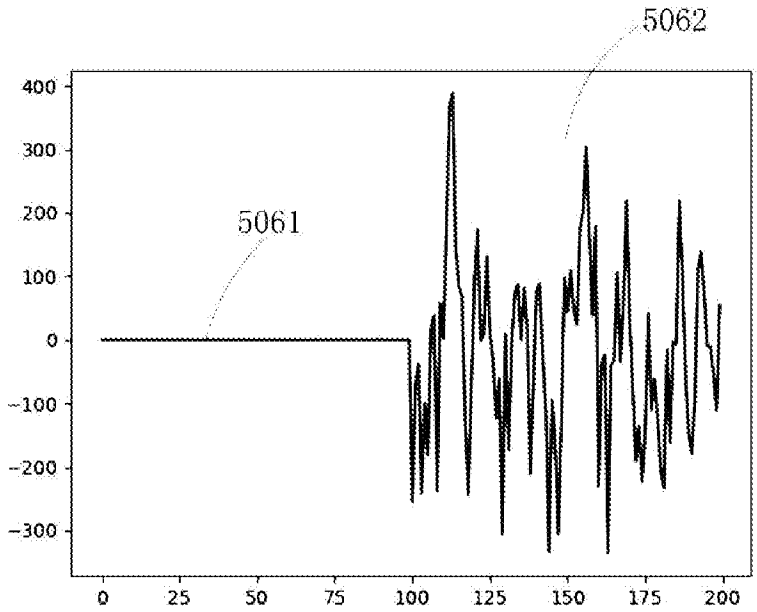

5062
5061
400
300
200
100
0
−100
−200
−300
0
25
50
75
100
125
150
175
200

图 5b

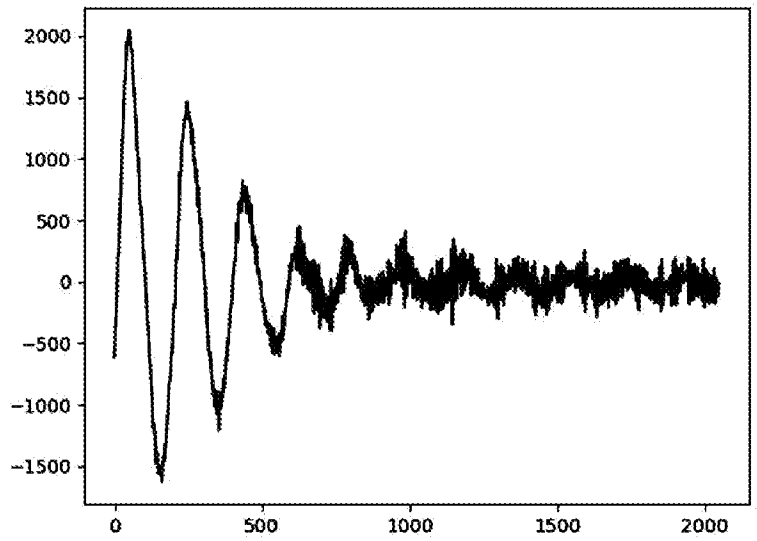

2000
1500
1000
500
0
−500
−1000
−1500
0
500
1000
1500
2000

图 5c

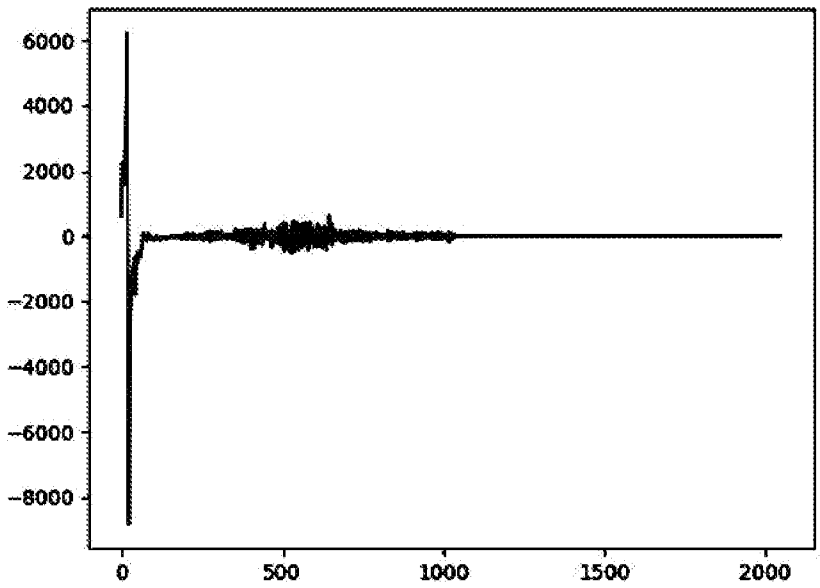

6000
4000
2000
0
−2000
−4000
−6000
−8000
0
500
1000
1500
2000

图 5d

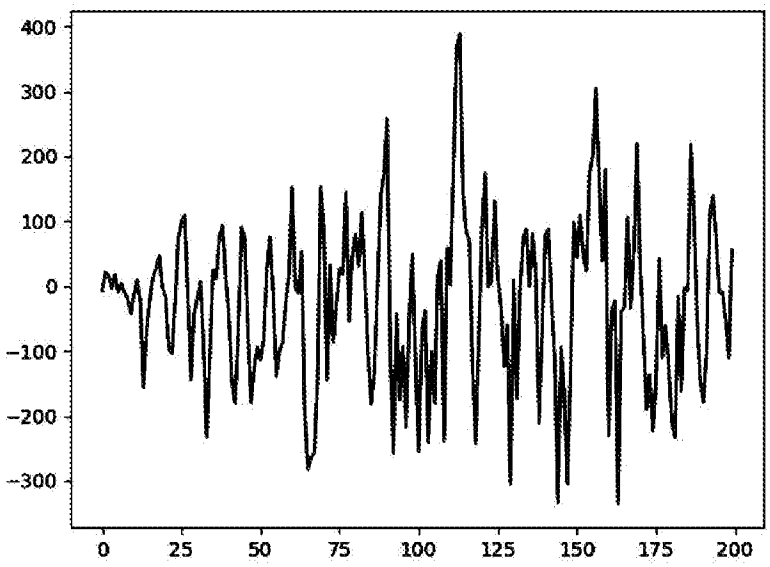

400
300
200
100
0
−100
−200
−300
0
25
50
75
100
125
150
175
200

图 5e

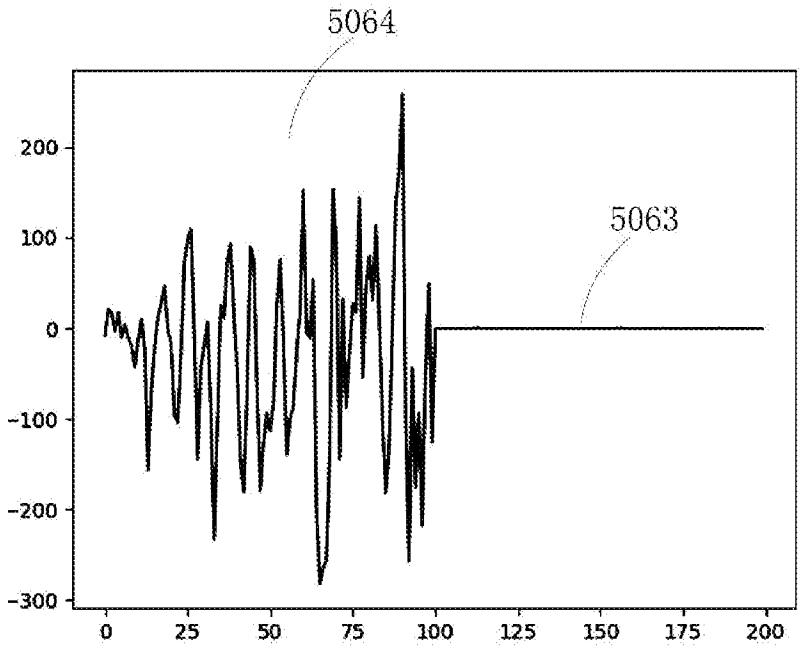

5064
5063
200
100
0
-100
-200
-300
0
25
50
75
100
125
150
175
200

图 5f

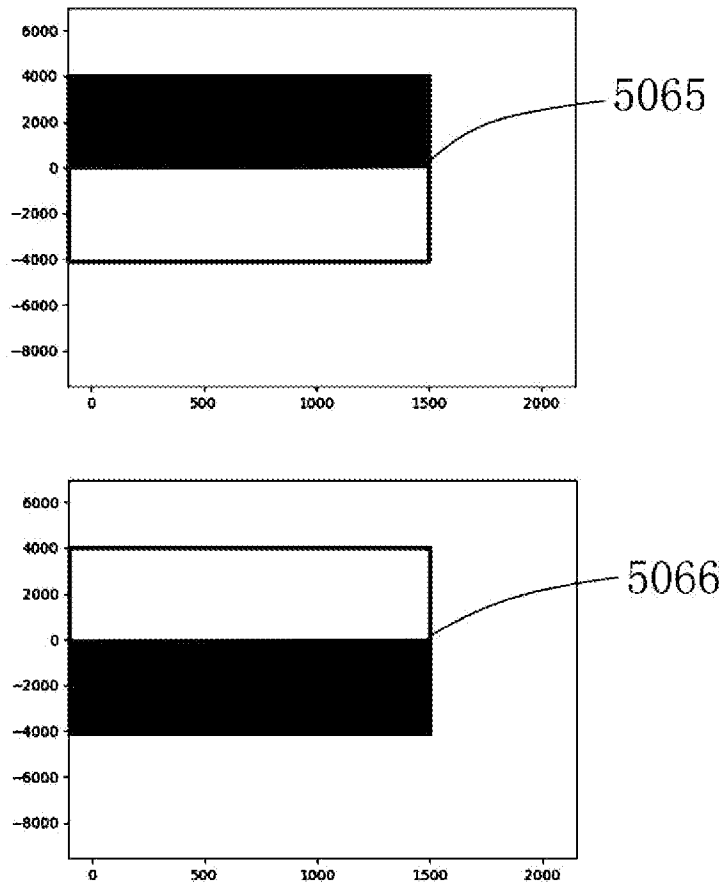

5065
5066
6000
4000
2000
0
-2000
-4000
-6000
-8000
0
500
1000
1500
2000

图 5g

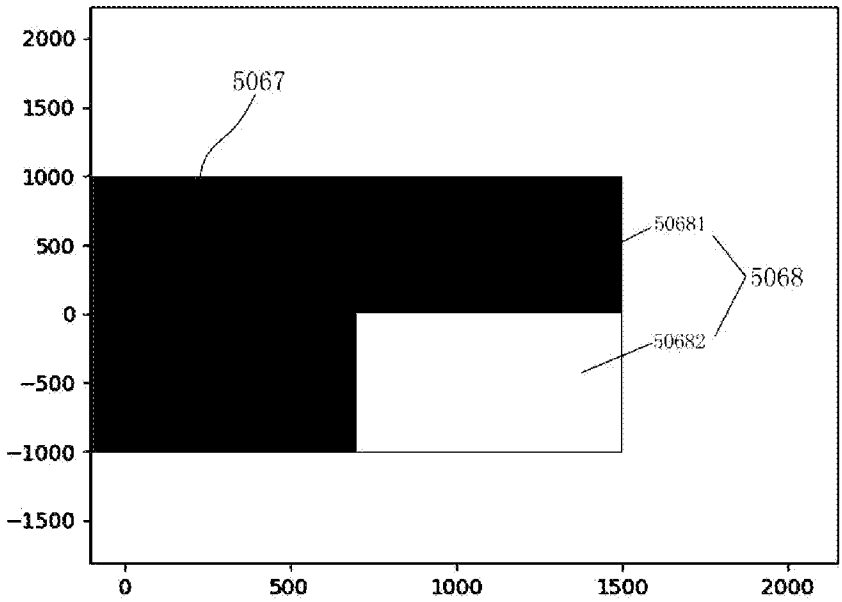

2000
1500
1000
500
0
−500
−1000
−1500
0
500
1000
1500
2000
5067
50681
5068
50682

图 5h

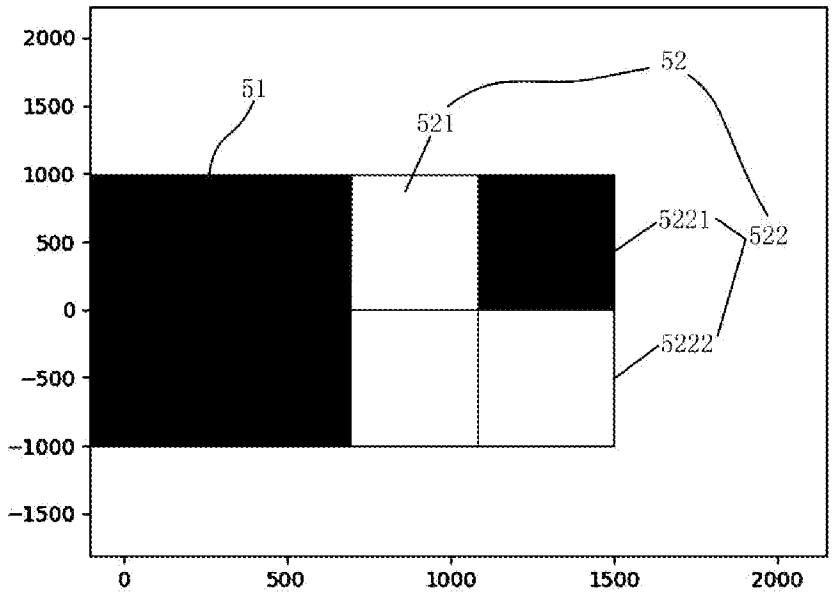

2000
1500
1000
500
0
−500
−1000
−1500
0
500
1000
1500
2000
51
52
521
5221
522
5222

图 5i

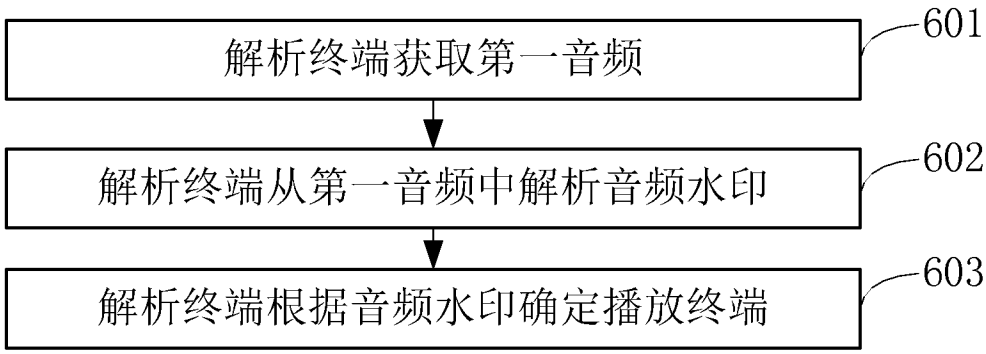

解析终端获取第一音频
601
解析终端从第一音频中解析音频水印
602
解析终端根据音频水印确定播放终端
603

图 6

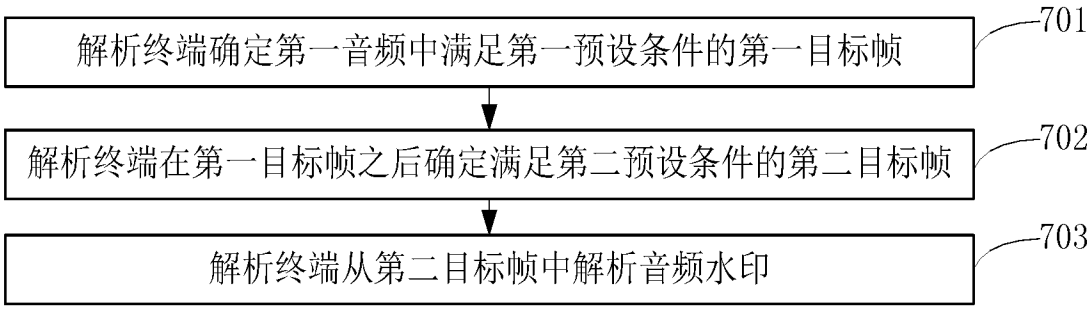

解析终端确定第一音频中满足第一预设条件的第一目标帧
701
解析终端在第一目标帧之后确定满足第二预设条件的第二目标帧
702
解析终端从第二目标帧中解析音频水印
703

图 7

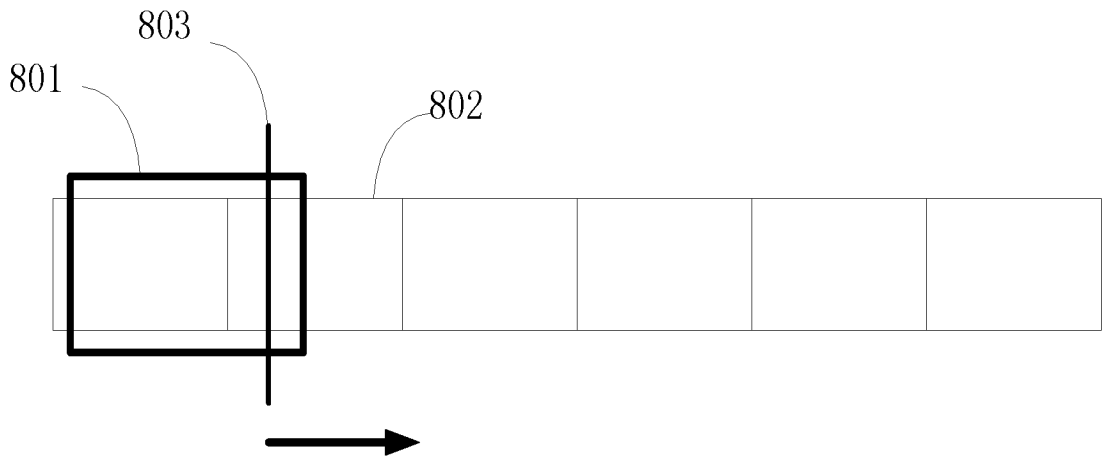

803
801
802

图 8

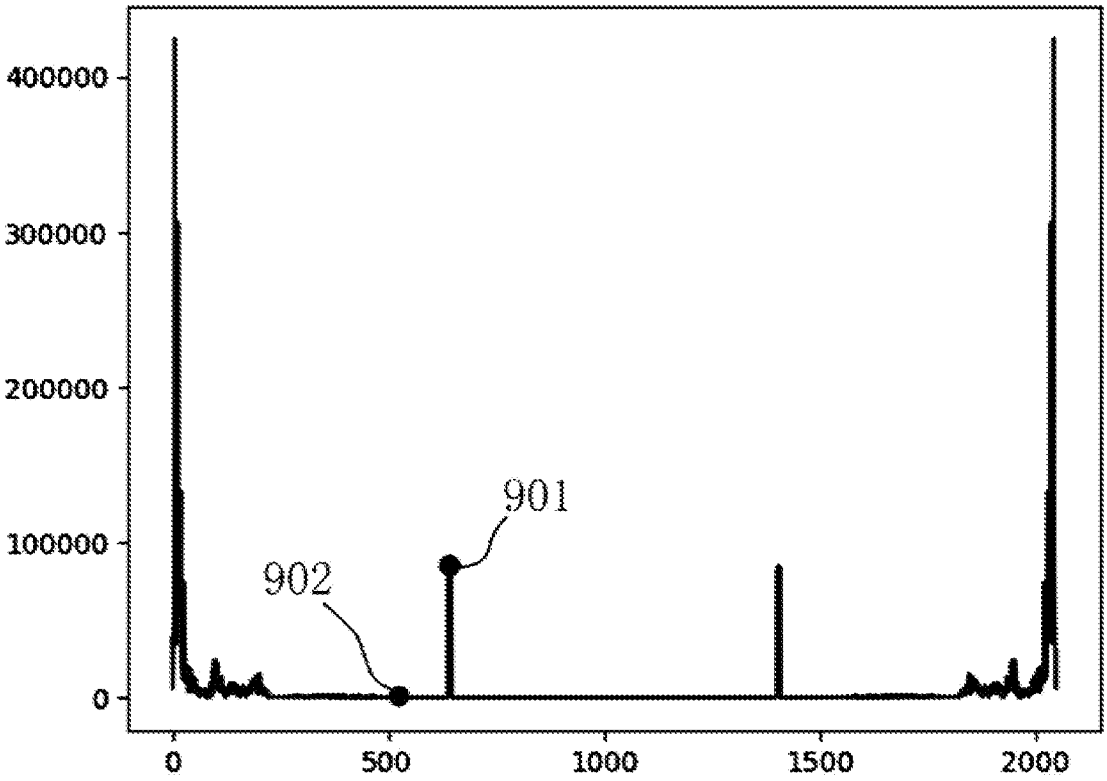

400000
300000
200000
100000
0
0
500
1000
1500
2000
901
902

图 9a

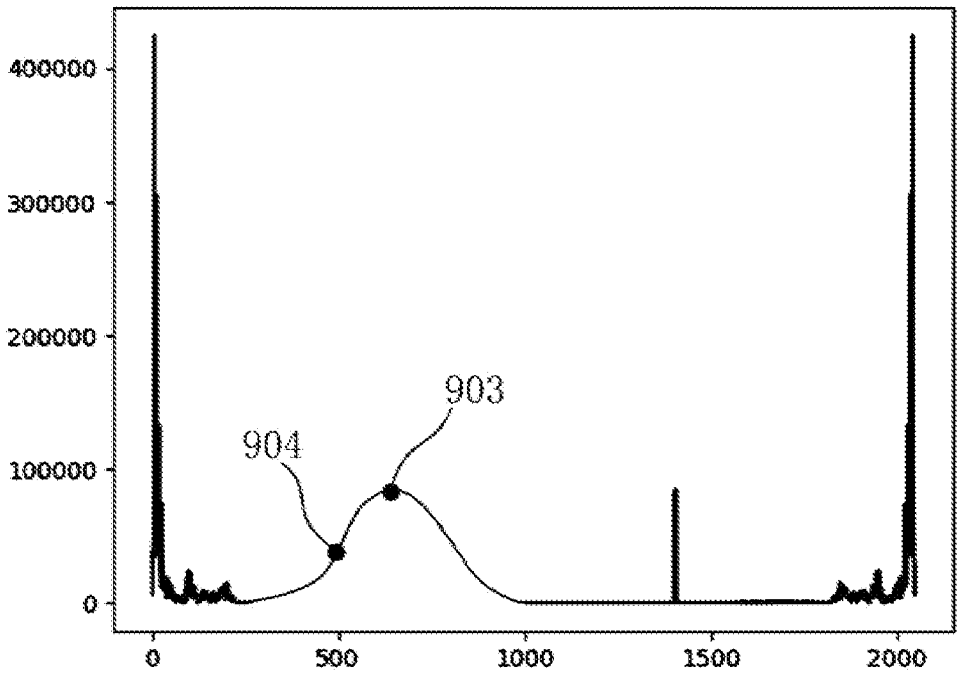

400000
300000
200000
100000
0
0
500
1000
1500
2000
903
904

图 9b

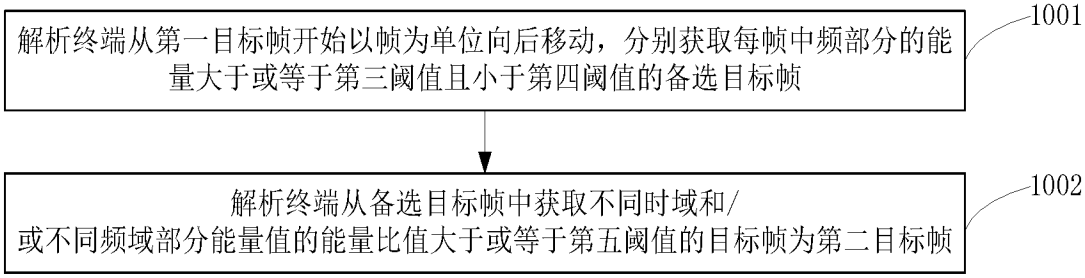

解析终端从第一目标帧开始以帧为单位向后移动，分别获取每帧中频部分的能量大于或等于第三阈值且小于第四阈值的备选目标帧
1001
解析终端从备选目标帧中获取不同时域和/或不同频域部分能量值的能量比值大于或等于第五阈值的目标帧为第二目标帧
1002

图 10

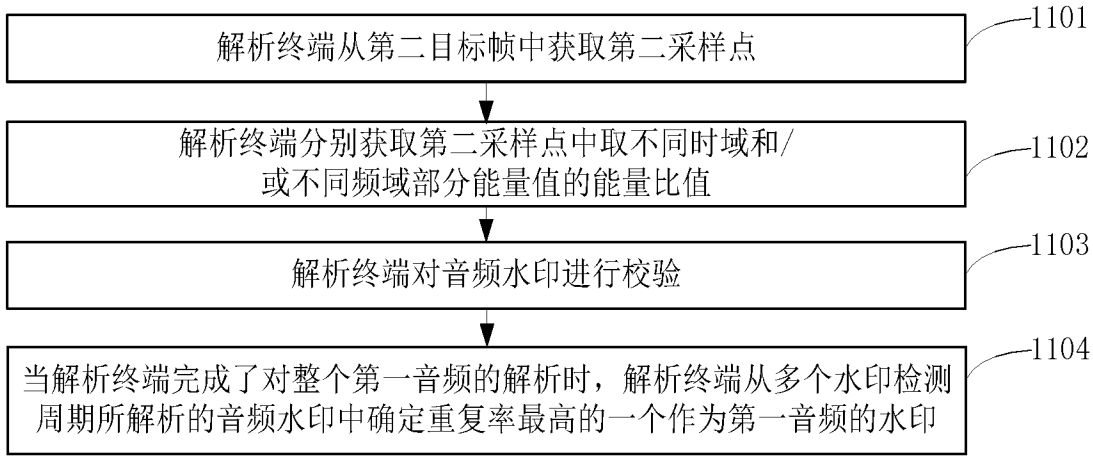

解析终端从第二目标帧中获取第二采样点
1101
解析终端分别获取第二采样点中取不同时域和/或不同频域部分能量值的能量比值
1102
解析终端对音频水印进行校验
1103
当解析终端完成了对整个第一音频的解析时，解析终端从多个水印检测周期所解析的音频水印中确定重复率最高的一个作为第一音频的水印
1104

图 11

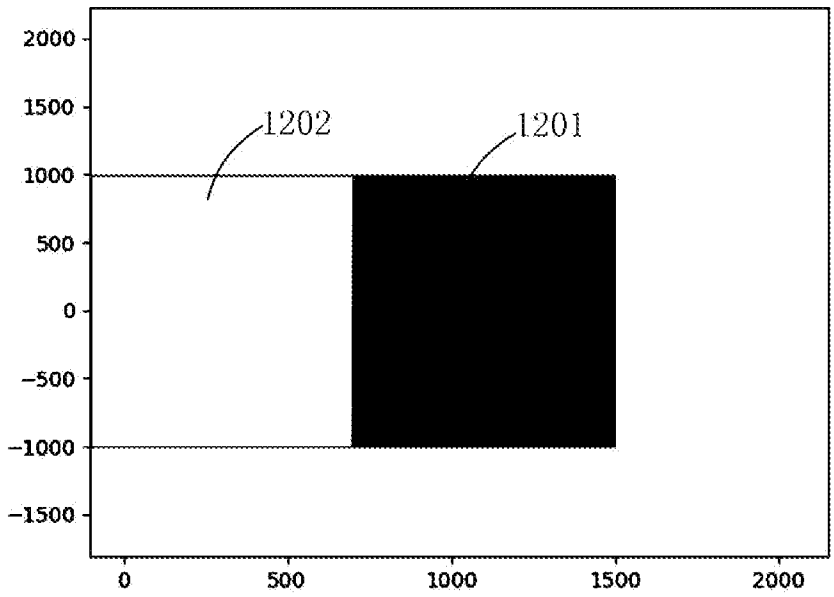

2000
1500
1000
500
0
−500
−1000
−1500
0
500
1000
1500
2000
1202
1201

图 12a

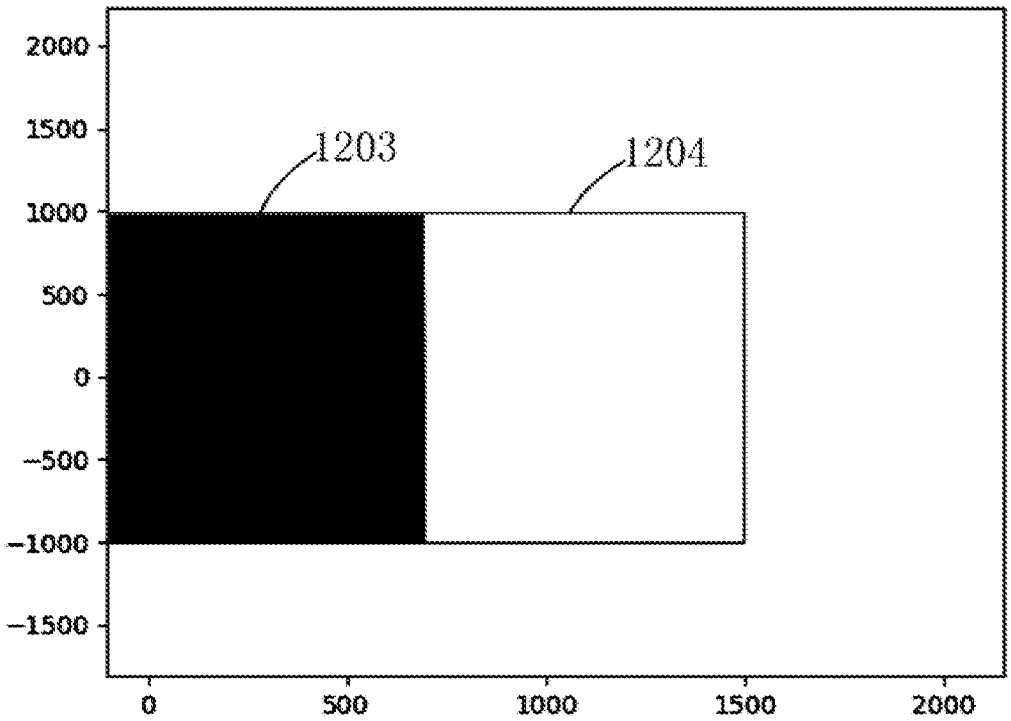

2000
1500
1000
500
0
-500
-1000
-1500
0
500
1000
1500
2000
1203
1204

图 12b

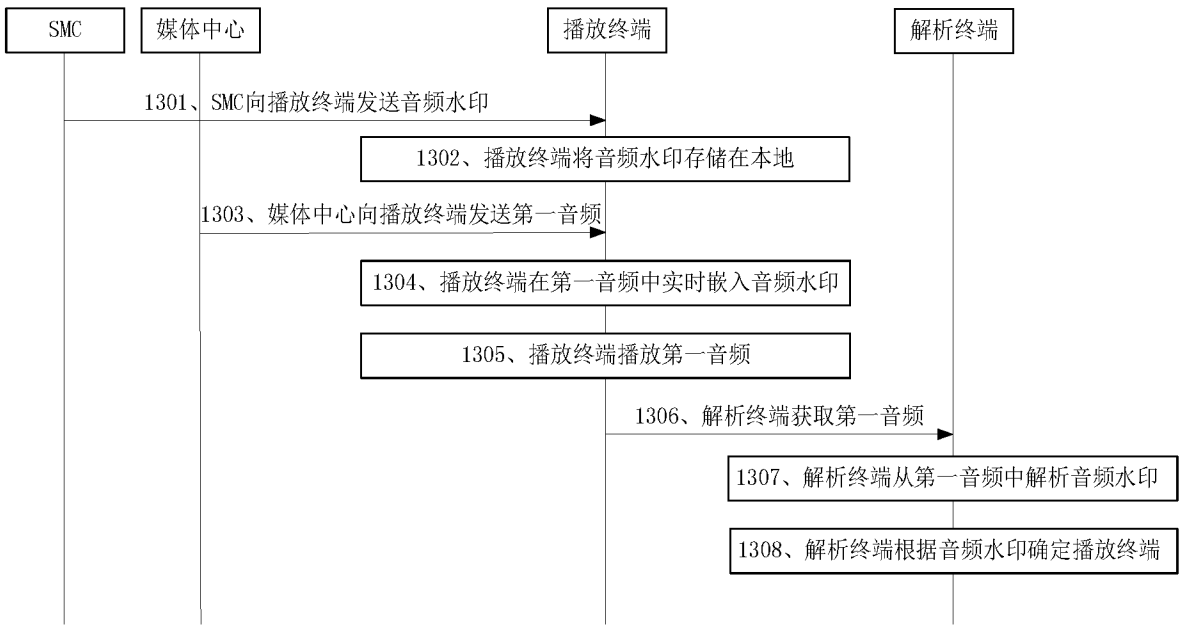

SMC
媒体中心
播放终端
解析终端
1301、SMC向播放终端发送音频水印
1302、播放终端将音频水印存储在本地
1303、媒体中心向播放终端发送第一音频
1304、播放终端在第一音频中实时嵌入音频水印
1305、播放终端播放第一音频
1306、解析终端获取第一音频
1307、解析终端从第一音频中解析音频水印
1308、解析终端根据音频水印确定播放终端

图 13

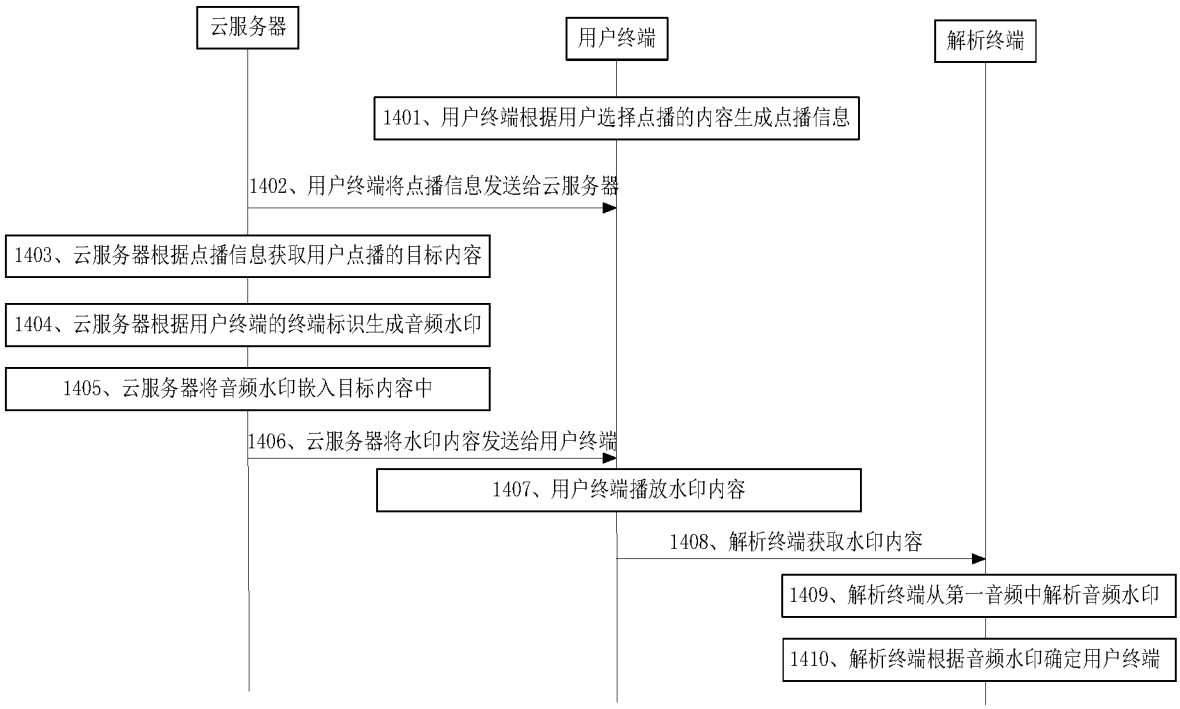

云服务器
用户终端
解析终端
1401、用户终端根据用户选择点播的内容生成点播信息
1402、用户终端将点播信息发送给云服务器
1403、云服务器根据点播信息获取用户点播的目标内容
1404、云服务器根据用户终端的终端标识生成音频水印
1405、云服务器将音频水印嵌入目标内容中
1406、云服务器将水印内容发送给用户终端
1407、用户终端播放水印内容
1408、解析终端获取水印内容
1409、解析终端从第一音频中解析音频水印
1410、解析终端根据音频水印确定用户终端

图 14

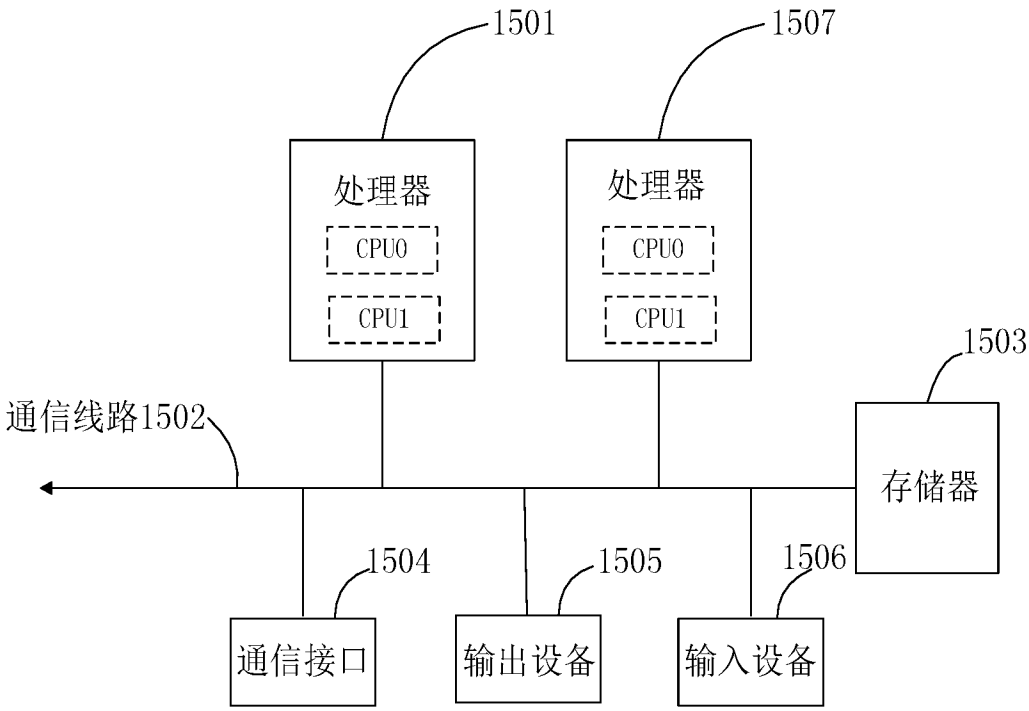

1501
1507
处理器
CPU0
CPU1
处理器
CPU0
CPU1
1503
通信线路1502
存储器
1504
1505
1506
通信接口
输出设备
输入设备

图 15

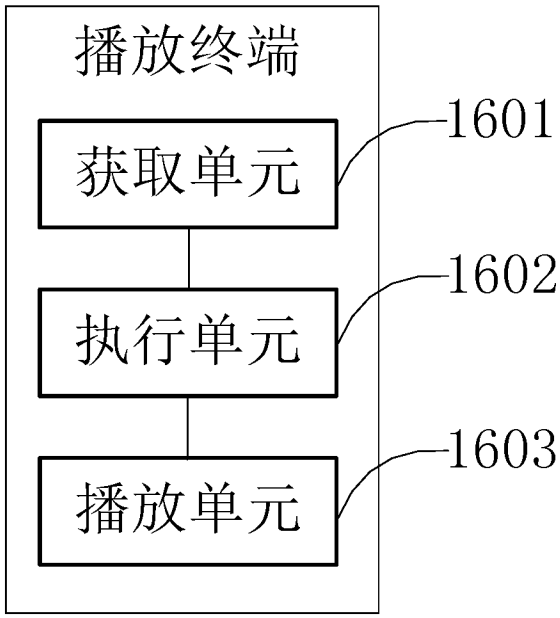

播放终端
获取单元
1601
执行单元
1602
播放单元
1603

图 16

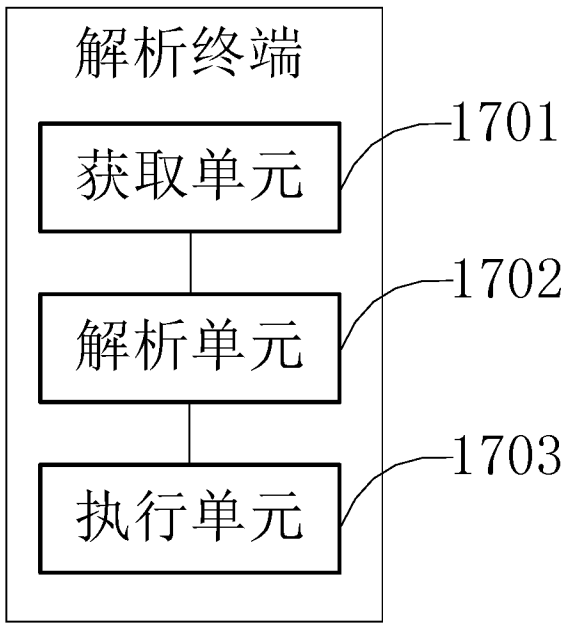

解析终端
获取单元
1701
解析单元
1702
执行单元
1703

图 17

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2021/118202

A. CLASSIFICATION OF SUBJECT MATTER

G10L 19/018(2013.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L19

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNKI, CNPAT, EPODOC, WPI: 音频, 水印, 声纹, 帧, 第一, 第二, 关联, 确定, 目标, 采样率, 阈值, 低频, 高频, 采样点, 能量, 时间, 中频, 元素, 频域, 时域, 滑动, audio, watermark, voiceprint, frame, first, second, association, determination, target, sample rate, threshold, low frequency, high frequency, sample point, energy, time, intermediate frequency, element, frequency domain, time domain, sliding

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 106921728 A (ALIBABA GROUP HOLDING LIMITED) 04 July 2017 (2017-07-04) description paragraphs [0054]-[0126]	1, 10, 18-21
A	CN 101350198 A (XIDIAN UNIVERSITY) 21 January 2009 (2009-01-21) entire document	1-21
A	CN 101266794 A (SHANGHAI JIAO TONG UNIVERSITY) 17 September 2008 (2008-09-17) entire document	1-21
A	CN 101442672 A (HUAWEI TECHNOLOGIES CO., LTD.) 27 May 2009 (2009-05-27) entire document	1-21
A	US 2003103645 A1 (LEVY, Kenneth L. et al.) 05 June 2003 (2003-06-05) entire document	1-21
A	CN 1462439 A (KONINKLIJKE PHILIPS ELECTRONICS N.V.) 17 December 2003 (2003-12-17) entire document	1-21

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 November 2021	**22 December 2021**
Name and mailing address of the ISA/CN **China National Intellectual Property Administration (ISA/CN)** **No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088, China** Facsimile No. **(86-10)62019451**	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2021/118202

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	106921728	A	04 July 2017	US	2019191276	A1	20 June 2019
				TW	I652497	B	01 March 2019
				KR	20190044101	A	29 April 2019
				JP	2019536070	A	12 December 2019
				TW	201812332	A	01 April 2018
				PH	12019500425	A1	11 November 2019
				US	10757537	B2	25 August 2020
				SG	11201901618 Q	A	28 March 2019
				EP	3509268	A1	10 July 2019
				WO	2018040947	A1	08 March 2018
CN	101350198	A	21 January 2009	CN	101350198	B	21 September 2011
CN	101266794	A	17 September 2008		None		
CN	101442672	A	27 May 2009	WO	2009065361	A1	28 May 2009
				CN	101442672	B	25 April 2012
US	2003103645	A1	05 June 2003	US	7224819	B2	29 May 2007
				US	7602978	B2	13 October 2009
				US	2009080694	A1	26 March 2009
				US	7961949	B2	14 June 2011
				US	2010027837	A1	04 February 2010
				US	2008044087	A1	21 February 2008
				US	7460726	B2	02 December 2008
CN	1462439	A	17 December 2003	US	7152161	B2	19 December 2006
				EP	1393315	B1	15 November 2006
				DE	60216085	D1	28 December 2006
				KR	20030015373	A	20 February 2003
				WO	02091376	A1	14 November 2002
				AT	345564	T	15 December 2006
				US	2003004589	A1	02 January 2003
				DE	60216085	T2	21 June 2007
				BR	0205150	A	24 June 2003
				EP	1393315	A1	03 March 2004
				JP	2004525430	A	19 August 2004
				CN	1284135	C	08 November 2006

国际检索报告

国际申请号
PCT/CN2021/118202

A. 主题的分类

G10L 19/018(2013.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G10L19

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNKI,CNPAT,EPODOC,WPI:音频，水印，声纹，帧，第一，第二，关联，确定，目标，采样率，阈值，低频，高频，采样点，能量，时间，中频，元素，频域，时域，滑动，audio, watermark, voiceprint, frame, first, second, association, determination, target, sample rate, threshold, low frequency, high frequency, sample point, energy, time, intermediate frequency, element, frequency domain, time domain, sliding

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
X	CN 106921728 A (阿里巴巴集团控股有限公司) 2017年7月4日 (2017-07-04) 说明书第[0054]-[0126]段	1，10，18-21
A	CN 101350198 A (西安电子科技大学) 2009年1月21日 (2009-01-21) 全文	1-21
A	CN 101266794 A (上海交通大学) 2008年9月17日 (2008-09-17) 全文	1-21
A	CN 101442672 A (华为技术有限公司) 2009年5月27日 (2009-05-27) 全文	1-21
A	US 2003103645 A1 (LEVY, Kenneth L. 等) 2003年6月5日 (2003-06-05) 全文	1-21
A	CN 1462439 A (皇家菲利浦电子有限公司) 2003年12月17日 (2003-12-17) 全文	1-21

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2021年11月22日	2021年12月22日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 王玥 电话号码 86-（10）- 53962545

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2021/118202

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	106921728	A	2017年7月4日	US	2019191276	A1	2019年6月20日
				TW	I652497	B	2019年3月1日
				KR	20190044101	A	2019年4月29日
				JP	2019536070	A	2019年12月12日
				TW	201812332	A	2018年4月1日
				PH	12019500425	A1	2019年11月11日
				US	10757537	B2	2020年8月25日
				SG	11201901618Q	A	2019年3月28日
				EP	3509268	A1	2019年7月10日
				WO	2018040947	A1	2018年3月8日
CN	101350198	A	2009年1月21日	CN	101350198	B	2011年9月21日
CN	101266794	A	2008年9月17日		无		
CN	101442672	A	2009年5月27日	WO	2009065361	A1	2009年5月28日
				CN	101442672	B	2012年4月25日
US	2003103645	A1	2003年6月5日	US	7224819	B2	2007年5月29日
				US	7602978	B2	2009年10月13日
				US	2009080694	A1	2009年3月26日
				US	7961949	B2	2011年6月14日
				US	2010027837	A1	2010年2月4日
				US	2008044087	A1	2008年2月21日
				US	7460726	B2	2008年12月2日
CN	1462439	A	2003年12月17日	US	7152161	B2	2006年12月19日
				EP	1393315	B1	2006年11月15日
				DE	60216085	D1	2006年12月28日
				KR	20030015373	A	2003年2月20日
				WO	02091376	A1	2002年11月14日
				AT	345564	T	2006年12月15日
				US	2003004589	A1	2003年1月2日
				DE	60216085	T2	2007年6月21日
				BR	0205150	A	2003年6月24日
				EP	1393315	A1	2004年3月3日
				JP	2004525430	A	2004年8月19日
				CN	1284135	C	2006年11月8日