(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 111951829 A
(43)申请公布日 2020.11.17

(21)申请号 202010401597.6
(22)申请日 2020.05.13
(71)申请人 慧言科技(天津)有限公司
地址 300450 天津市滨海新区华苑产业区海泰发展六道6号海泰绿色产业基地J座210、211
申请人 深圳市康鸿泰科技有限公司
(72)发明人 关昊天 姜宇 葛檬 廖启波
(74)专利代理机构 深圳市智胜联合知识产权代理有限公司 44368
代理人 齐文剑
(51)Int.Cl.
G10L 25/51(2013.01)
G10L 21/0216(2013.01)
G10L 25/03(2013.01)
H04L 29/08(2006.01)

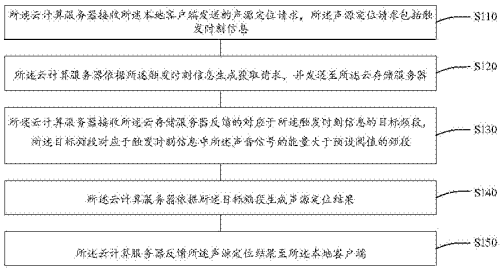
权利要求书2页 说明书11页 附图5页

(54)发明名称

基于时域单元的声源定位方法、装置及系统

(57)摘要

本发明实施例提供了一种基于时域单元的声源定位方法，应用于远场声音信号的定位，所述方法包括：所述云计算服务器接收所述本地客户端发送的声源定位请求，所述声源定位请求包括触发时刻信息；所述云计算服务器依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器；所述云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，所述目标频段对应于触发时刻信息中所述声音信号的能量大于预设阈值的频段；所述云计算服务器依据所述目标频段生成声源定位结果；所述云计算服务器反馈所述声源定位结果至所述本地客户端。通过云计算技术、大数据分析方法和机器学习算法，系统、快速的实现云端定位方法，保证数据安全。



CN 111951829 A

1.一种基于时域单元的声源定位方法，其特征在于，应用于远场声音信号的定位，所述方法涉及云计算服务器，本地客户端，以及云存储服务器；所述云存储服务器用于存储实时采集的声音信号；

所述方法包括：

所述云计算服务器接收所述本地客户端发送的声源定位请求，所述声源定位请求包括触发时刻信息；

所述云计算服务器依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器；

所述云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，所述目标频段对应于触发时刻信息中所述声音信号的能量大于预设阈值的频段；

所述云计算服务器依据所述目标频段生成声源定位结果；

所述云计算服务器反馈所述声源定位结果至所述本地客户端。

2.根据权利要求1所述的方法，其特征在于，所述接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段的步骤，包括：

接收所述云存储服务器反馈的对应于所述触发时刻的第一语音信息；

计算所述第一语音信息中每一帧的短时能量；

挑选出所述短时能量大于所述预设阈值的频段作为所述目标频段。

3.根据权利要求1所述的方法，其特征在于，所述依据所述目标频段生成声源定位结果的步骤，包括：

采用加权的可控响应功率算法计算出所述目标频段的声源定位结果。

4.一种基于时域单元的声源定位方法，其特征在于，应用于远场声音信号的定位，所述方法涉及云计算服务器，本地客户端，以及云存储服务器；所述云存储服务器用于存储实时采集的声音信号；

所述方法包括：

所述本地客户端生成声源定位请求，并发送至所述云计算服务器，所述声源定位请求包括触发时刻信息；

所述本地客户端接收所述云计算服务器反馈的所述声源定位结果，其中，所述声源定位结果为所述云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，并依据所述目标频段生成。

5.一种基于时域单元的声源定位装置，其特征在于，应用于远场声音信号的定位，所述装置涉及云计算服务器，本地客户端，以及云存储服务器；所述云存储服务器用于存储实时采集的声音信号；

所述云计算服务器具体包括：

声源定位请求接收模块，用于接收所述本地客户端发送的声源定位请求，所述声源定位请求包括触发时刻信息；

获取请求生成模块，用于依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器；

目标频段接收模块，用于接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，所述目标频段对应于触发时刻信息中所述声音信号的能量大于预设阈值的频段；

声源定位结果生成模块，用于依据所述目标频段生成声源定位结果；

声源定位结果发送模块，用于反馈所述声源定位结果至所述本地客户端。

6.一种基于时域单元的声源定位装置，其特征在于，应用于远场声音信号的定位，所述装置涉及云计算服务器，本地客户端，以及云存储服务器；所述云存储服务器用于存储实时采集的声音信号；

所述本地客户端具体包括：

声源定位请求生成模块，用于生成声源定位请求，并发送至所述云计算服务器，所述声源定位请求包括触发时刻信息；

声源定位结果接收模块，用于接收所述云计算服务器反馈的所述声源定位结果，其中，所述声源定位结果为所述云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，并依据所述目标频段生成。

7.一种基于时域单元的声源定位系统，其特征在于，应用于远场声音信号的定位，所述系统涉及云计算服务器，本地客户端，以及云存储服务器；具体包括；

所述本地客户端用于生成声源定位请求，并发送至所述云计算服务器，所述声源定位请求包括触发时刻信息；

所述云计算服务器用于依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器；

所述云存储服务器用于存储实时采集的声音信号；

所述云存储服务器还用于依据所述获取请求确定对应于所述触发时刻信息的目标频段，并发送至所述云计算服务器；

所述云计算服务器还用于依据所述目标频段生成声源定位结果；

所述本地客户端还用于接收所述云计算服务器反馈的所述声源定位结果。

8.一种电子设备，其特征在于，包括处理器、存储器及存储在所述存储器上并能够在所述处理器上运行的计算机程序，所述计算机程序被所述处理器执行时实现如权利要求1至4中任一项所述的基于时域单元的声源定位方法的步骤。

9.一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储计算机程序，所述计算机程序被处理器执行时实现如权利要求1至4中任一项所述的基于时域单元的声源定位方法的步骤。

基于时域单元的声源定位方法、装置及系统

技术领域

[0001] 本发明涉及声学信号处理技术领域，特别是涉及一种基于时域单元的声源定位方法、装置及系统。

背景技术

[0002] 基于麦克风阵列的声源定位技术被广泛应用于视频会议、语音增强、智能机器人、智能家居、车载通话设备等。

[0003] 其中，声源定位的应用场景分为室内环境和室外环境，麦克风阵列声源定位方法定义为：利用麦克风阵列去采集室内声源目标，经过对就接收到声音信号的一系列分析与与处理，找到声源目标的准确位置。

[0004] 传统的声源定位方法，如使用最小二乘法或广义互相关法获取到达时间差进而计算声源位置的基于时延估计的声源定位方法，基于高分辨率谱估计的声源定位方法，以及基于可控波束形成的声源定位方法，在低信噪比、高混响环境下，定位效果很差。

发明内容

[0005] 鉴于上述问题，提出了本发明实施例以便提供一种克服上述问题或者至少部分地解决上述问题的一种基于时域单元的声源定位方法、装置及系统。

[0006] 为了解决上述问题，本发明实施例公开了一种基于时域单元的声源定位方法，应用于远场声音信号的定位，所述方法涉及云计算服务器，本地客户端，以及云存储服务器；所述云存储服务器用于存储实时采集的声音信号；

[0007] 所述方法包括：

[0008] 所述云计算服务器接收所述本地客户端发送的声源定位请求，所述声源定位请求包括触发时刻信息；

[0009] 所述云计算服务器依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器；

[0010] 所述云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，所述目标频段对应于触发时刻信息中所述声音信号的能量大于预设阈值的频段；

[0011] 所述云计算服务器依据所述目标频段生成声源定位结果；

[0012] 所述云计算服务器反馈所述声源定位结果至所述本地客户端。

[0013] 进一步地，所述接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段的步骤，包括：

[0014] 接收所述云存储服务器反馈的对应于所述触发时刻的第一语音信息；

[0015] 计算所述第一语音信息中每一帧的短时能量；

[0016] 挑选出所述短时能量大于所述预设阈值的频段，作为所述目标频段。

[0017] 进一步地，所述依据所述目标频段生成声源定位结果的步骤，包括：

[0018] 采用加权的可控响应功率算法计算出所述目标频段的声源定位结果。

[0019] 本发明实施例公开了一种基于时域单元的声源定位方法，应用于远场声音信号的定位，所述方法涉及云计算服务器，本地客户端，以及云存储服务器；所述云存储服务器用于存储实时采集的声音信号；

[0020] 所述方法包括：

[0021] 所述本地客户端生成声源定位请求，并发送至所述云计算服务器，所述声源定位请求包括触发时刻信息；

[0022] 所述本地客户端接收所述云计算服务器反馈的所述声源定位结果，其中，所述声源定位结果为所述云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，并依据所述目标频段生成。

[0023] 本发明实施例公开了一种基于时域单元的声源定位装置，应用于远场声音信号的定位，所述装置涉及云计算服务器，本地客户端，以及云存储服务器；所述云存储服务器用于存储实时采集的声音信号；

[0024] 所述云计算服务器具体包括：

[0025] 声源定位请求接收模块，用于接收所述本地客户端发送的声源定位请求，所述声源定位请求包括触发时刻信息；

[0026] 获取请求生成模块，用于依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器；

[0027] 目标频段接收模块，用于接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，所述目标频段对应于触发时刻信息中所述声音信号的能量大于预设阈值的频段；

[0028] 声源定位结果生成模块，用于依据所述目标频段生成声源定位结果；

[0029] 声源定位结果发送模块，用于反馈所述声源定位结果至所述本地客户端。

[0030] 本发明实施例公开了一种基于时域单元的声源定位装置，应用于远场声音信号的定位，所述装置涉及云计算服务器，本地客户端，以及云存储服务器；所述云存储服务器用于存储实时采集的声音信号；

[0031] 所述本地客户端具体包括：

[0032] 声源定位请求生成模块，用于生成声源定位请求，并发送至所述云计算服务器，所述声源定位请求包括触发时刻信息；

[0033] 声源定位结果接收模块，用于接收所述云计算服务器反馈的所述声源定位结果，其中，所述声源定位结果为所述云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，并依据所述目标频段生成。

[0034] 本发明实施例公开了一种基于时域单元的声源定位系统，应用于远场声音信号的定位，所述系统涉及云计算服务器，本地客户端，以及云存储服务器；具体包括；

[0035] 所述本地客户端用于生成声源定位请求，并发送至所述云计算服务器，所述声源定位请求包括触发时刻信息；

[0036] 所述云计算服务器用于依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器；

[0037] 所述云存储服务器用于存储实时采集的声音信号；

[0038] 所述云存储服务器还用于依据所述获取请求确定对应于所述触发时刻信息的目

标频段，并发送至所述云计算服务器；

[0039] 所述云计算服务器还用于依据所述目标频段生成声源定位结果；

[0040] 所述本地客户端还用于接收所述云计算服务器反馈的所述声源定位结果。

[0041] 本发明实施例公开了一种电子设备，包括处理器、存储器及存储在所述存储器上并能够在所述处理器上运行的计算机程序，所述计算机程序被所述处理器执行时实现如上述的基于时域单元的声源定位方法的步骤。

[0042] 本发明实施例公开了一种计算机可读存储介质，所述计算机可读存储介质上存储计算机程序，所述计算机程序被处理器执行时实现如上述的基于时域单元的声源定位方法的步骤。

[0043] 本发明实施例包括以下优点：对采集到的语音检测有声段和无声段，只针对有声段进行分析，对麦克风阵列采集到的信号进行一定的语音增强处理，去除噪声和混响等干扰因素。利用波束形成原理去除对定位语音的干扰，对信号短时能量检测以选择可靠度最高的时域段，提出一种更加系统、且更加快速的云端定位方法。

附图说明

[0044] 图1是本发明的一实施例中一种基于时域单元的声源定位方法实施例的步骤流程图；

[0045] 图2是本发明的一实施例中一种基于时域单元的声源定位方法实施例的步骤流程图；

[0046] 图3是本发明的一实施例中一种基于时域单元的声源定位装置云计算服务器的结构框图；

[0047] 图4是本发明的一实施例中一种基于时域单元的声源定位装置本地客户端的结构框图；

[0048] 图5是本发明的一实施例中一种基于时域单元的声源定位方法实施例的流程图；

[0049] 图6是本发明的一实施例中一种基于时域单元的声源定位方法中语音的归一化短时能量示意图；

[0050] 图7是本发明的一实施例中一种基于时域单元的声源定位方法实施例的流程图；

[0051] 图8是本发明的一实施例中一种基于时域单元的声源定位方法实施例的90度定位结果示意图；

[0052] 图9是本发明的一实施例中一种基于时域单元的声源定位系统的结构框图；

[0053] 图10是本发明一实施例的一种计算机设备的结构示意图。

具体实施方式

[0054] 为使本发明的上述目的、特征和优点能够更加明显易懂，下面结合附图和具体实施方式对本发明作进一步详细的说明。

[0055] 本发明实施例的核心构思之一在于，提供了一种基于时域单元的声源定位方法，应用于远场声音信号的定位，所述方法包括：所述云计算服务器接收所述本地客户端发送的声源定位请求，所述声源定位请求包括触发时刻信息；所述云计算服务器依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器；所述云计算服务器接收所述云存储

服务器反馈的对应于所述触发时刻信息的目标频段，所述目标频段对应于触发时刻信息中所述声音信号的能量大于预设阈值的频段；所述云计算服务器依据所述目标频段生成声源定位结果；所述云计算服务器反馈所述声源定位结果至所述本地客户端。通过云计算技术、大数据分析方法和机器学习算法，系统、快速的实现云端定位方法，保证数据安全。

[0056] 参照图1，示出了本发明的一种基于时域单元的声源定位方法实施例的步骤流程图，应用于远场声音信号的定位，所述方法涉及云计算服务器，本地客户端，以及云存储服务器；所述云存储服务器用于存储实时采集的声音信号；具体可以包括如下步骤：

[0057] S110，所述云计算服务器接收所述本地客户端发送的声源定位请求，所述声源定位请求包括触发时刻信息；

[0058] S120，所述云计算服务器依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器；

[0059] S130所述云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，所述目标频段对应于触发时刻信息中所述声音信号的能量大于预设阈值的频段；

[0060] S140，所述云计算服务器依据所述目标频段生成声源定位结果；

[0061] S150，所述云计算服务器反馈所述声源定位结果至所述本地客户端。

[0062] 参照上述步骤S110所示，所述云计算服务器接收所述本地客户端发送的声源定位请求，所述声源定位请求包括触发时刻信息。具体的所述声源定位请求包含所述本地客户端触发的声源定位任务。

[0063] 参照上述步骤S120所示，所述云计算服务器依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器；所述云存储服务器设有指定的缓冲区，在数据采集阶段，利用本地客户端手机应用控制麦克风阵列录制多通道音频，采集到的数据暂存于云存储服务器中。在一具体实施例中，麦克风阵列为线性四个麦克风，总长度15厘米，麦克风间距5厘米。该设备属于微型麦克风阵列，对到达时间差要求较高。麦克风处于常开阶段，持续接收音频信号，保存到云存储服务器中指定的缓冲区中。

[0064] 参照上述步骤S130所述，云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，所述目标频段对应于触发时刻信息中所述声音信号的能量大于预设阈值的频段；具体地，语音短时能量是指计算较短时间内的语音能量，指的是一帧时间内的语音能量。语音的短时能量就是将语音中每一帧的短时能量都计算出来，短时能量主要用于区分浊音段和清音段、区分声母与韵母的分界和无话段与有话段分界。在本实施例中主要用于检测非静音段，并对达到预设阈值的非静音段采取切割选择等措施得到目标频段。对于一段持续输入到缓冲区的音频信号，计算一段时间内的语音能量。首先声音信号通过分帧、加窗，n时刻该语音信号的短时平均能量定义为：

[0065] $$E_n=\sum_{m=-\infty}^{+\infty}[x(m)\omega(n-m)]^2=\sum_{m=n-(N-1)}^{n}[x(m)\omega(n-m)]^2$$

[0066] 其中，N表示窗长，这里取512，ω(n)表示窗口函数，这里使用汉宁窗：

[0067] $$\omega[n]=0.5+0.5\cos(\frac{2\pi n}{N-1})$$

[0068] 为了计算整段语音信号的短时能量，需要在时域上对分帧后的信号做填补等操

作，对每一帧计算后的短时能量做归一化处理，结合绘制的短时能量图，可以清晰地看清发声时间段。完成上述操作后，将合适的语音段，利用sox(音频处理工具)操作切割得到目的频段并传输到云计算服务器，进行下一步定位操作。

[0069] 参照上述步骤S140所示，所述云计算服务器依据所述目标频段生成声源定位结果；可控响应功率算法是在双通道广义互相关算法的基础上，扩展到麦克风阵列的声源定位。通过计算每对麦克风之间的广义互相关函数并加和，在整个声源空间寻找使得函数值最大的点即为声源定位的位置。由于本系统仅涉及平面方位角定位，因此在估计时总是假设仰角为0度。

[0070] 广义互相关算法是在估计到达时间差的基础上反推方向角度的。因此麦克风之间的间距必须是已知。互相关函数和频域信号之间存在衔接关系，通过计算互功率谱，可以直接求得广义互相关函数

[0071] $$R_{x_1x_2}(\tau)=\int_0^{\pi}G_{x_1x_2}(\omega)e^{-i\omega\tau}d\omega=\int_0^{\pi}X_1(\omega)X_2^{*}(\omega)e^{-i\omega\tau}d\omega$$

[0072] 其中G表示互功率谱，通过每对麦克风接收到的信号x_1和x_2进行傅里叶变换，并对x_2进行共轭操作后卷积得到。而互相关函数数学上是通过计算两个信号的期望值得到的：

[0073] $$R_{x_1x_2}(\tau)=E[x_1(t)x_2(t-\tau)]=E[(s(t-\tau_1)+n_1(t))(s(t-\tau_2)+n_2(t))]=R_{ss}(\tau-\tau_{12})$$

[0074] 这里接收到的信号x_1和x_2实际表达如上式所示，这里假设信号和噪音，噪音和噪音之间互不相关，因此上式成立。由此可知，当$\tau_{12}=\tau_1-\tau_2$时，互相关函数取得最大值，而这恰恰是我们需要的到达时间差。

[0075] 因此只要求得在R取得最大值时的τ值即为到达时间差。而可控响应功率算法只需求每对麦克风的广义互相关函数之和即可：

[0076] $$P(\theta)=\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}R_{mn}(\tau_{mn}(\theta))$$

[0077] 其中n和m表示一对麦克风。值得一提的是，这里求n和m的广义互相关等价于求m和n的广义互相关，因此上述求和操作还可以继续缩小范围，将计算量缩减一半。

[0078] 在实际环境中，由于存在部分混响和白噪声影响，导致互相关函数的最大值不明显，到达时间差估计效果自然不佳。为了锐化互相关函数的最大值，利用波束形成的思想，对频域信号进行加权，之后再通过逆傅里叶变换得到互相关函数，达到抑制干扰的效果。加权函数根据不同的场景可以有不同的选择，为了最大效果抑制噪音，常常选择使用相位加权，即加权函数为互功率绝对值的倒数：

[0079] $$\varphi_{mn}=\frac{1}{\left|X_n(\omega)X_m^{*}(\omega)\right|}$$

[0080] 参照步骤S150所示，所述云计算服务器反馈所述声源定位结果至所述本地客户端。具体地，通过云端计算得到结果后再次返回手机端进行显示。

[0081] 本发明提出一种系统的实时声源定位的方法，采用语音能量检测确定定位语音，用加权可控响应功率对声音进行定位。本发明使用真实场景下录制的语音信号来评估系统性能，在真实环境下可以有效地对声音进行定位，其精度差控制在5°范围内。

[0082] 在一具体实施例中，本示例以真实场景录制语音为例来给出发明的实施方式。整

个系统流程如图5所示，包括数据采集、音频能量检测、声源定位这三个步骤，具体实施方式如下：真实场景下录音进行测试，房间大小5m*3m，麦克风阵列置于随机位置，高1.5m，可供选择的录制距离为1m、1.5m、2m、3m，具体的录制角度为前方0°至180°，发声数据有音乐、敲门声、说话人的声音等等。其中音乐是通过手机外放，而其他声音为真实场景下的声音。为保证鲁棒性，基于上述方案，条件随机挑选，共录制200个测试样本。为了证明系统的鲁棒性，基于上述录制场景，在加入空调声音等噪声的条件下，再次录制带噪的声音200条。

[0083] 在确定进行定位的时刻，检测语音能量是否到达阈值，选取到达阈值的语音片段进行定位。图6显示某一时刻时长为3秒的语音的短时能量。其中横坐标为时间，纵坐标为归一化的能量大小，能量的起伏反应了该时刻语音信号的强弱。

[0084] 在获取到将要定位的语音之后，调用云端算法进行定位，整个流程如图7所示。首先计算每对麦克风的广义互相关函数，在计算广义互相关函数的时候使用加权操作，目的是尽可能减少环境噪音的影响。对麦克风对的广义互相关函数求和之后，寻找使得函数取最大值的方向即为所求。图8为某次定位的结果图。

[0085] 表一列出了在没有加入噪声干扰的情况下，该方法和传统方案的精确度对比。一般来讲，远场条件下定位角度和真实角度差在5°以内，就认为该结果是准确的。测试不同声音类型下，该方案的准确率。主要测试的声音为敲门、音乐、人声、咳嗽四种类型。其中可控响应功率是传统的方法，能量检测-可控响应功率是本次提出的方案。通过表一可以看出能量检测-可控响应功率的方法在较干净的语音情况下，取得了不错的效果。

[0086]

方法 / 声音	可控响应功率	能量检测-可控响应功率
音乐	81.4%	94.8%
人声	83.5%	93.6%
敲门	77.2%	89.3%
咳嗽	78.3%	88.5%
均值	80.1%	92.0%

[0087] 表一

[0088] 为了进一步验证声源定位系统的鲁棒性，表二列出了在带噪条件下录制数据，不同信噪比下测试的结果，表二的实验结果表明，采用带噪数据，该系统的稳定性仍然高于传统方案。

[0089]

方法 / 信噪比	可控响应功率	能量检测-可控响应功率
20dB	79.7%	91.5%

[0090]

15dB	79.1%	91.2%
10dB	77.6%	90.3%
0dB	66.4%	87.6%
均值	75.7%	90.2%

[0091] 表二

[0092] 在本实施例中，所述接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段的步骤，包括：

[0093] 接收所述云存储服务器反馈的对应于所述触发时刻的第一语音信息；

[0094] 计算所述第一语音信息中每一帧的短时能量；

[0095] 挑选出所述短时能量大于所述预设阈值的频段，作为所述目标频段。

[0096] 在本实施例中，所述依据所述目标频段生成声源定位结果的步骤，包括：

[0097] 采用加权的可控响应功率算法计算出所述目标频段的声源定位结果。

[0098] 在本实施例中，还包括：

[0099] 利用人工智能模型的自学能力，建立所述声音信号与所述声源定位结果之间的对应关系。

[0100] 参照图2，本发明实施例公开了一种基于时域单元的声源定位方法，应用于远场声音信号的定位，所述方法涉及云计算服务器，本地客户端，以及云存储服务器；所述云存储服务器用于存储实时采集的声音信号；

[0101] 所述方法包括：

[0102] S210，所述本地客户端生成声源定位请求，并发送至所述云计算服务器，所述声源定位请求包括触发时刻信息；

[0103] S220，所述本地客户端接收所述云计算服务器反馈的所述声源定位结果，其中，所述声源定位结果为所述云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，并依据所述目标频段生成。

[0104] 需要说明的是，对于方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本发明实施例并不受所描述的动作顺序的限制，因为依据本发明实施例，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作并不一定是本发明实施例所必须的。

[0105] 参照图3，示出了本发明的一种基于时域单元的声源定位装置的结构框图，应用于远场声音信号的定位，所述装置涉及云计算服务器，本地客户端，以及云存储服务器；所述云存储服务器用于存储实时采集的声音信号；

[0106] 所述云计算服务器具体包括：

[0107] 声源定位请求接收模块110，用于接收所述本地客户端发送的声源定位请求，所述声源定位请求包括触发时刻信息；

[0108] 获取请求生成模块120，用于依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器；

[0109] 目标频段接收模块130，用于接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，所述目标频段对应于触发时刻信息中所述声音信号的能量大于预设阈值的频段；

[0110] 声源定位结果生成模块140，用于依据所述目标频段生成声源定位结果；

[0111] 声源定位结果发送模块150，用于反馈所述声源定位结果至所述本地客户端。

[0112] 在本实施例中，所述目标频段接收模块包括：

[0113] 接收单元，用于接收所述云存储服务器反馈的对应于所述触发时刻的第一语音信息；

[0114] 计算单元，用于计算所述第一语音信息中每一帧的短时能量；

[0115] 筛选单元，用于挑选出所述短时能量大于所述预设阈值的频段，作为所述目标频段。

[0116] 在本实施例中，所述声源定位结果生成模块包括：

[0117] 声源定位结果生成单元，用于采用加权的可控响应功率算法计算出所述目标频段的声源定位结果。

[0118] 在本实施例中，还包括：

[0119] 参照图4，本发明实施例公开了一种基于时域单元的声源定位装置，应用于远场声音信号的定位，所述装置涉及云计算服务器，本地客户端，以及云存储服务器；所述云存储服务器用于存储实时采集的声音信号；

[0120] 所述本地客户端具体包括：

[0121] 声源定位请求生成模块210，用于生成声源定位请求，并发送至所述云计算服务器，所述声源定位请求包括触发时刻信息；

[0122] 声源定位结果接收模块220，用于接收所述云计算服务器反馈的所述声源定位结果，其中，所述声源定位结果为所述云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，并依据所述目标频段生成。

[0123] 对于装置实施例而言，由于其与方法实施例基本相似，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

[0124] 本说明书中的各个实施例均采用递进的方式描述，每个实施例重点说明的都是与其他实施例的不同之处，各个实施例之间相同相似的部分互相参见即可。

[0125] 参照图9，本发明实施例公开了一种基于时域单元的声源定位系统，应用于远场声音信号的定位，所述系统涉及云计算服务器100，本地客户端，200以及云存储服务器300；具体包括；

[0126] 所述本地客户端200用于生成声源定位请求，并发送至所述云计算服务器100，所述声源定位请求包括触发时刻信息；

[0127] 所述云计算服务器100用于依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器300；

[0128] 所述云存储服务器300用于存储实时采集的声音信号；

[0129] 所述云存储服务器300还用于依据所述获取请求确定对应于所述触发时刻信息的目标频段，并发送至所述云计算服务器100；

[0130] 所述云计算服务器100还用于依据所述目标频段生成声源定位结果；

[0131] 所述本地客户端200还用于接收所述云计算服务器100反馈的所述声源定位结果。

[0132] 参照图10，示出了本发明的一种基于时域单元的声源定位方法的计算机设备，具体可以包括如下：

[0133] 上述计算机设备12以通用计算设备的形式表现，计算机设备12的组件可以包括但不限于：一个或者多个处理器或者处理单元16，系统存储器28，连接不同系统组件(包括系统存储器28和处理单元16)的总线18。

[0134] 总线18表示几类总线18结构中的一种或多种，包括存储器总线18或者存储器控制器，外围总线18，图形加速端口，处理器或者使用多种总线18结构中的任意总线18结构的局域总线18。举例来说，这些体系结构包括但不限于工业标准体系结构(ISA)总线18，微通道体系结构(MAC)总线18，增强型ISA总线18、音视频电子标准协会(VESA)局域总线18以及外围组件互连(PCI)总线18。

[0135] 计算机设备12典型地包括多种计算机系统可读介质。这些介质可以是任何能够被计算机设备12访问的可用介质，包括易失性和非易失性介质，可移动的和不可移动的介质。

[0136] 系统存储器28可以包括易失性存储器形式的计算机系统可读介质，例如随机存取存储器(RAM)30和/或高速缓存存储器32。计算机设备12可以进一步包括其他移动/不可移动的、易失性/非易失性计算机体统存储介质。仅作为举例，存储系统34可以用于读写不可移动的、非易失性磁介质(通常称为“硬盘驱动器”)。尽管图10中未示出，可以提供用于对可移动非易失性磁盘(如“软盘”)读写的磁盘驱动器，以及对可移动非易失性光盘(例如CD-ROM，DVD-ROM或者其他光介质)读写的光盘驱动器。在这些情况下，每个驱动器可以通过一个或者多个数据介质界面与总线18相连。存储器可以包括至少一个程序产品，该程序产品具有一组(例如至少一个)程序模块42，这些程序模块42被配置以执行本发明各实施例的功能。

[0137] 具有一组(至少一个)程序模块42的程序/实用工具40，可以存储在例如存储器中，这样的程序模块42包括——但不限于——操作系统、一个或者多个应用程序、其他程序模块42以及程序数据，这些示例中的每一个或某种组合中可能包括网络环境的实现。程序模块42通常执行本发明所描述的实施例中的功能和/或方法。

[0138] 计算机设备12也可以与一个或多个外部设备14(例如键盘、指向设备、显示器24、摄像头等)通信，还可与一个或者多个使得用户能与该计算机设备12交互的设备通信，和/或与使得该计算机设备12能与一个或多个其他计算设备进行通信的任何设备(例如网卡，调制解调器等等)通信。这种通信可以通过输入/输出(I/O)界面22进行。并且，计算机设备12还可以通过网络适配器20与一个或者多个网络(例如局域网(LAN))，广域网(WAN)和/或公共网络(例如因特网)通信。如图所示，网络适配器20通过总线18与计算机设备12的其他模块通信。应当明白，尽管图10中未示出，可以结合计算机设备12使用其他硬件和/或软件模块，包括但不限于：微代码、设备驱动器、冗余处理单元16、外部磁盘驱动阵列、RAID系统、磁带驱动器以及数据备份存储系统34等。

[0139] 处理单元16通过运行存储在系统存储器28中的程序，从而执行各种功能应用以及数据处理，例如实现本发明实施例所提供的基于时域单元的声源定位方法。

[0140] 也即，上述处理单元16执行上述程序时实现：接收所述本地客户端发送的声源定位请求，所述声源定位请求包括触发时刻信息；依据所述触发时刻信息生成获取请求，并发

送至所述云存储服务器;接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段,所述目标频段对应于触发时刻信息中所述声音信号的能量大于预设阈值的频段;依据所述目标频段生成声源定位结果;反馈所述声源定位结果至所述本地客户端。

[0141] 在本发明实施例中,本发明还提供一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现如本申请所有实施例提供的基于时域单元的声源定位方法:

[0142] 也即,给程序被处理器执行时实现:接收所述本地客户端发送的声源定位请求,所述声源定位请求包括触发时刻信息;依据所述触发时刻信息生成获取请求,并发送至所述云存储服务器;接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段,所述目标频段对应于触发时刻信息中所述声音信号的能量大于预设阈值的频段;依据所述目标频段生成声源定位结果;反馈所述声源定位结果至所述本地客户端。

[0143] 可以采用一个或多个计算机可读的介质的任意组合。计算机可读介质可以是计算机克顿信号介质或者计算机可读存储介质。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦可编程只读存储器(EPOM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件或者上述的任意合适的组合。在本文件中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0144] 计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括——但不限于——电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质,该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。

[0145] 可以以一种或多种程序设计语言或其组合来编写用于执行本发明操作的计算机程序代码,上述程序设计语言包括面向对象的程序设计语言——诸如Java、Smalltalk、C++,还包括常规的过程式程序设计语言——诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行或者完全在远程计算机或者服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)——连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。本说明书中的各个实施例均采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似的部分互相参见即可。

[0146] 尽管已描述了本申请实施例的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例做出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本申请实施例范围的所有变更和修改。

[0147] 最后,还需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作

之间存在任何这种实际的关系或者顺序。而且，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者终端设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者终端设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者终端设备中还存在另外的相同要素。

[0148] 以上对本申请所提供的基于时域单元的声源定位方法及装置，进行了详细介绍，本文中应用了具体个例对本申请的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本申请的方法及其核心思想；同时，对于本领域的一般技术人员，依据本申请的思想，在具体实施方式及应用范围上均会有改变之处，综上所述，本说明书内容不应理解为对本申请的限制。

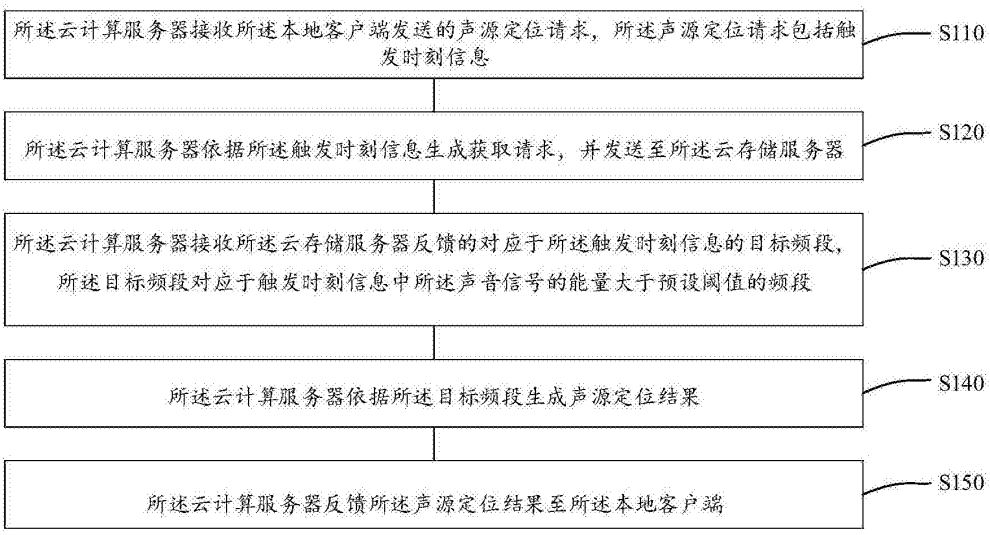

所述云计算服务器接收所述本地客户端发送的声源定位请求，所述声源定位请求包括触发时刻信息
S110
所述云计算服务器依据所述触发时刻信息生成获取请求，并发送至所述云存储服务器
S120
所述云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，所述目标频段对应于触发时刻信息中所述声音信号的能量大于预设阈值的频段
S130
所述云计算服务器依据所述目标频段生成声源定位结果
S140
所述云计算服务器反馈所述声源定位结果至所述本地客户端
S150

图1

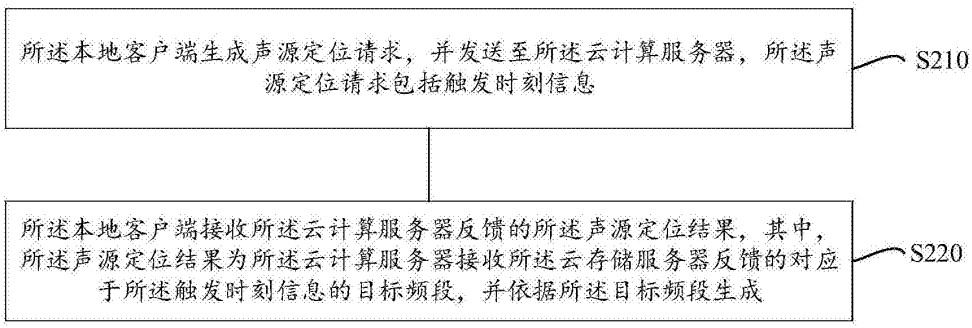

所述本地客户端生成声源定位请求，并发送至所述云计算服务器，所述声源定位请求包括触发时刻信息
S210
所述本地客户端接收所述云计算服务器反馈的所述声源定位结果，其中，所述声源定位结果为所述云计算服务器接收所述云存储服务器反馈的对应于所述触发时刻信息的目标频段，并依据所述目标频段生成
S220

图2

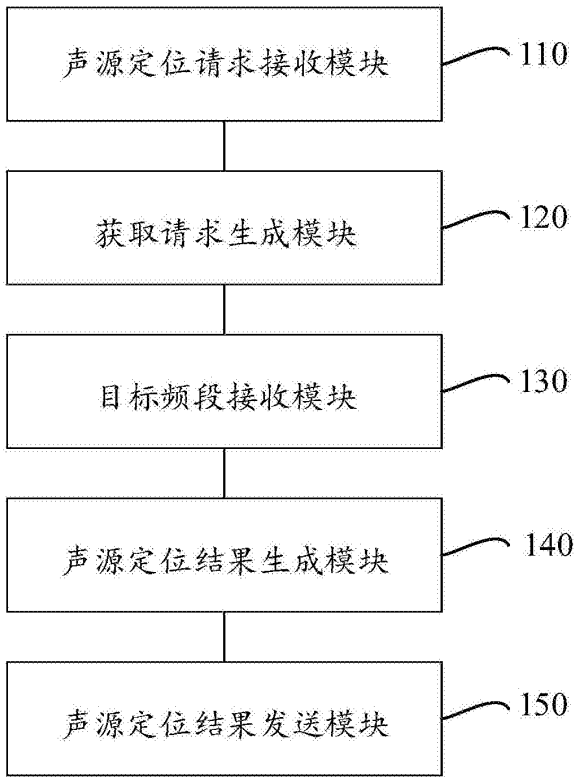

声源定位请求接收模块
110
获取请求生成模块
120
目标频段接收模块
130
声源定位结果生成模块
140
声源定位结果发送模块
150

图3

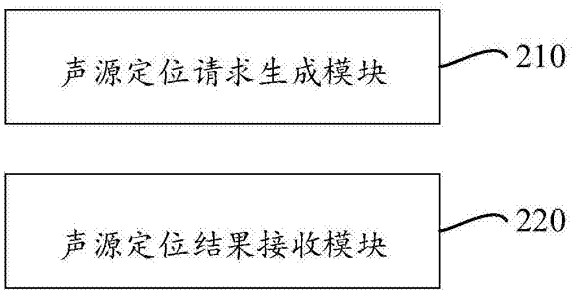

声源定位请求生成模块
210
声源定位结果接收模块
220

图4

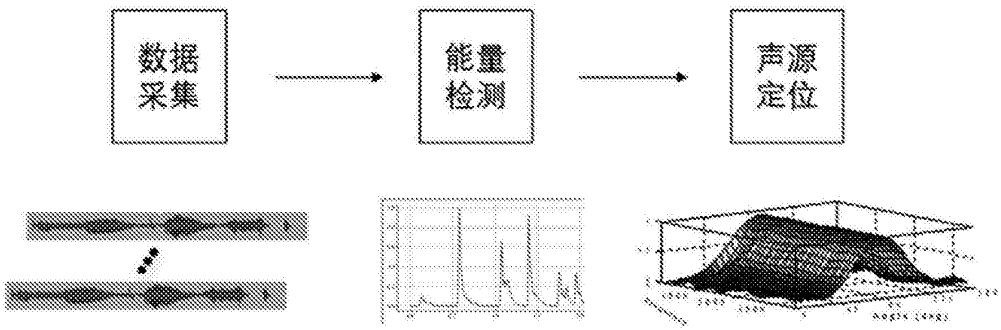

数据
采集
能量
检测
声源
定位

图5

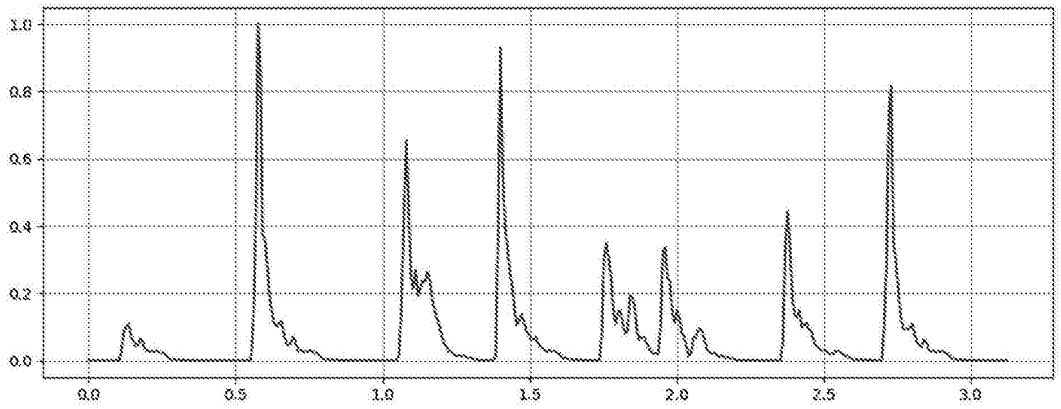

1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0

图6

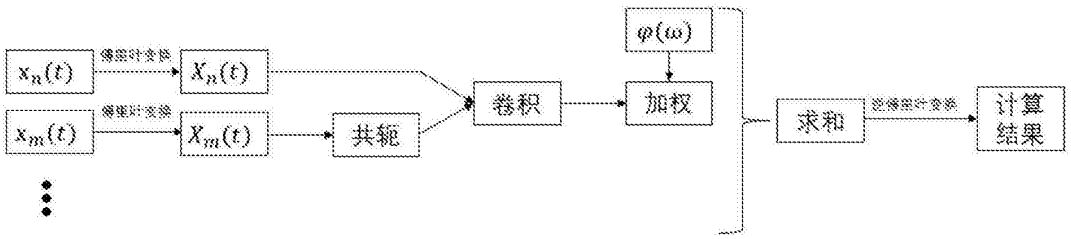

$x_n(t)$
$X_n(t)$
$x_m(t)$
$X_m(t)$
共轭
卷积
$\varphi(\omega)$
加权
求和
计算
结果

图7

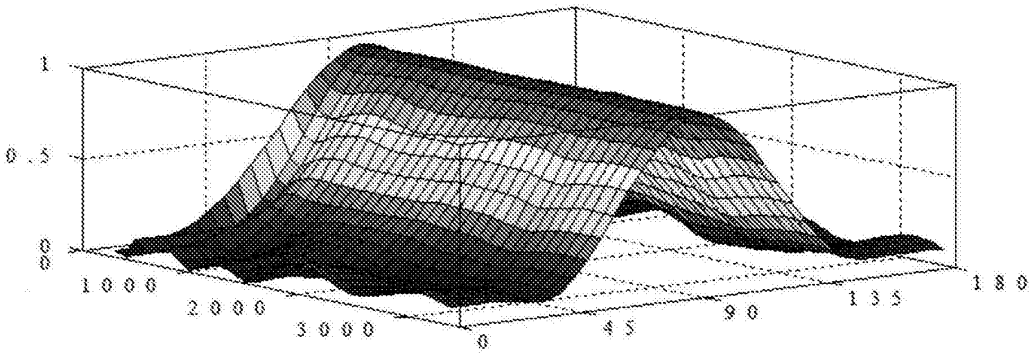

1
0.5
0
0
1000
2000
3000
0
45
90
135
180

图8

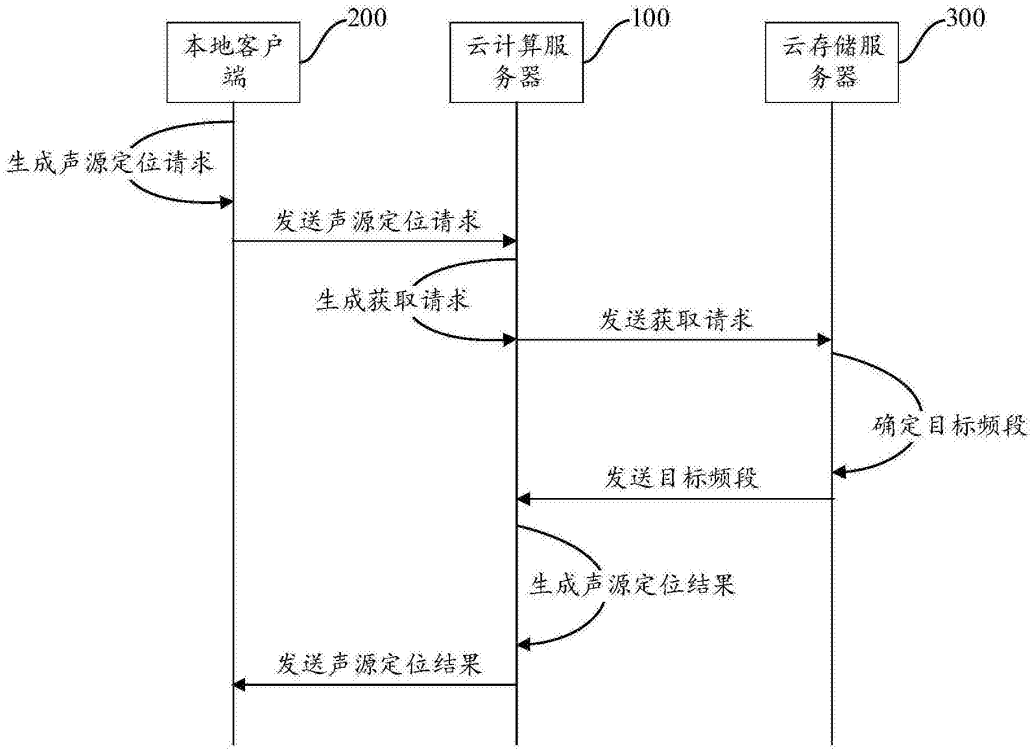

本地客户端
200
云计算服务器
100
云存储服务器
300
生成声源定位请求
发送声源定位请求
生成获取请求
发送获取请求
确定目标频段
发送目标频段
生成声源定位结果
发送声源定位结果

图9

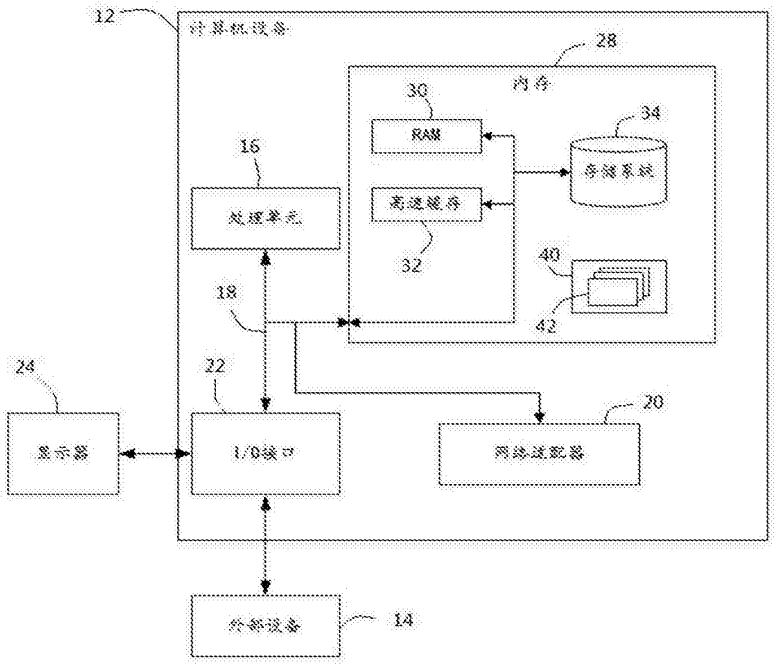

12
计算机设备
28
内存
30
RAM
34
存储系统
16
处理单元
高速缓存
32
40
42
18
24
22
20
显示器
I/O接口
网络适配器
外部设备
14

图10