



(12) 发明专利

(10) 授权公告号 CN 102982810 B

(45) 授权公告日 2016. 01. 13

(21) 申请号 201210478344. 4

(22) 申请日 2008. 11. 06

(30) 优先权数据

61/033, 992 2008. 03. 05 US

(62) 分案原申请数据

200880128616. X 2008. 11. 06

(73) 专利权人 尼尔森(美国)有限公司

地址 美国伊利诺伊州

(72) 发明人 亚历山大·帕夫洛维奇·托普奇

阿伦·拉马斯瓦米

韦努戈帕尔·斯里尼瓦桑

(74) 专利代理机构 北京三友知识产权代理有限公司

11127

代理人 吕俊刚 孙海龙

(51) Int. Cl.

G10L 25/48(2013. 01)

H04H 60/58(2008. 01)

(56) 对比文件

WO 02/065782 A1, 2002. 08. 22, 全文.

WO 03/009277 A2, 2003. 01. 30, 全文.

US 2006/0020958 A1, 2006. 01. 26, 全文.

US 5437050 A, 1995. 07. 25, 全文.

CN 101077014 A, 2007. 11. 21, 全文.

审查员 康丹丹

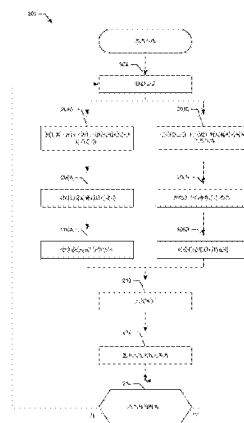
权利要求书1页 说明书12页 附图11页

(54) 发明名称

生成签名的方法和装置

(57) 摘要

本发明公开了一种生成签名的方法和装置。该生成签名的方法包括以下步骤:将第一函数应用于数字数据的一部分,以生成第一加窗块;将与所述第一函数不同的第二函数应用于所述数字数据的所述一部分,以生成第二加窗块;确定所述第一加窗块中频带的第一特征;确定所述第二加窗块中所述频带的第二特征;将所述第一特征与所述第二特征进行比较;以及基于所述第一特征和所述第二特征的比较指派表示所述数字数据的所述一部分的签名位。



1. 一种生成表示数字数据的签名的方法,所述方法包括以下步骤:  
将第一函数应用于数字数据的一部分,以生成第一加窗块;  
将与所述第一函数不同的第二函数应用于所述数字数据的所述一部分,以生成第二加窗块;  
确定所述第一加窗块中频带的第一特征;  
确定所述第二加窗块中所述频带的第二特征;  
将所述第一特征与所述第二特征进行比较;以及  
基于所述第一特征和所述第二特征的比较指派表示所述数字数据的所述一部分的签名位,其中,所述数字数据包括视频信息、网页或静止图像。
2. 根据权利要求 1 所述的方法,其中,所述第一函数和第二函数包括互补窗口函数。
3. 根据权利要求 1 所述的方法,其中,从窗口函数集中选择所述第一函数和所述第二函数。
4. 根据权利要求 1 所述的方法,其中,所述第一特征和所述第二特征包括第一加窗块和第二加窗块的频谱平坦度。
5. 一种生成表示数字数据的签名的装置,所述装置包括:  
加窗器,用于将第一窗口函数应用于数字数据的一部分,以生成第一加窗块并将第二窗口函数应用于所述数字数据的所述一部分,以生成第二加窗块;  
特征确定器,用于确定所述第一加窗块中的频带的第一特征并确定所述第二加窗块中的频带的第二特征;  
比较器,用于将所述第一特征与所述第二特征进行比较;以及  
签名确定器,用于基于所述第一特征和所述第二特征的比较指派表示所述数字数据的所述一部分的签名位,其中,所述数字数据包括视频信息、网页或静止图像。
6. 根据权利要求 5 所述的装置,其中,所述第一窗口函数和第二窗口函数包括互补窗口函数。
7. 根据权利要求 5 所述的装置,其中,从窗口函数集中选择所述第一窗口函数和所述第二窗口函数。
8. 根据权利要求 5 所述的装置,其中,所述第一特征和所述第二特征包括第一加窗块和第二加窗块的频谱平坦度。

## 生成签名的方法和装置

[0001] 本申请根据专利法实施细则第四十二条提出,是国际申请日为 2008 年 11 月 6 日、国际申请号为 PCT/US2008/082657、国家申请号为 200880128616. X、发明名称为“生成签名的方法和装置”中国专利申请的分案申请。

### 技术领域

[0002] 本公开一般地涉及媒体监视、多媒体内容搜索和检索,更具体地说,涉及生成用于标识媒体信息的签名的方法和装置。

### 背景技术

[0003] 利用签名匹配技术标识媒体信息(更具体地说是音频信号(例如,音频流中的信息))的方法已经很好地建立了。签名同样还被公知为并且经常称为指纹。签名匹配技术常用于电视观众和无线广播听众计量应用,并利用生成和匹配签名的几种方法实现。例如,在电视观众计量应用中,在监测地(例如,被监测家庭)和基准地生成签名。监测地通常包括诸如监测听众/观众成员的媒体消费量的家庭的位置。例如,在监测地,可以基于与所选择的频道、无线电台等相关联的音频流生成被监测签名。然后可以将被监测签名发送给中心数据收集设备用于分析。在基准地,基于广播区域内提供的已知程序生成签名,通常称为基准签名。基准签名可以存储在基准地和/或中心数据收集设备处并与监测地生成的被监测签名进行比较。可能发现被监测签名与基准签名匹配,并且与匹配基准签名相应的已知程序可以标识为在监测地出现过的程序。

### 发明内容

[0004] 根据本发明的一个方面,提供了一种生成表示数字数据的签名的方法,该方法包括以下步骤:将第一函数应用于数字数据的一部分,以生成第一加窗块;将与所述第一函数不同的第二函数应用于所述数字数据的所述一部分,以生成第二加窗块;确定所述第一加窗块中频带的第一特征;确定所述第二加窗块中所述频带的第二特征;将所述第一特征与所述第二特征进行比较;以及基于所述第一特征和所述第二特征的比较指派表示所述数字数据的所述一部分的签名位,其中,所述数字数据包括信号、时间序列数据、或媒体信息。

[0005] 根据本发明的另一方面,提供了一种生成表示数字数据的签名的装置,该装置包括:加窗器,用于将第一窗口函数应用于数字数据的一部分,以生成第一加窗块并将第二窗口函数应用于所述数字数据的所述一部分,以生成第二加窗块;特征确定器,用于确定所述第一加窗块中的频带的第一特征并确定所述第二加窗块中的频带的第二特征;比较器,用于将所述第一特征与所述第二特征进行比较;以及签名确定器,用于基于所述第一特征和所述第二特征的比较指派表示所述数字数据的所述一部分的签名位,其中,所述数字数据包括信号、时间序列数据、或媒体信息。

### 附图说明

- [0006] 图 1A 和 1B 例示了用于生成签名和标识音频流的示例性音频流标识系统。
- [0007] 图 2 是例示了示例性签名生成过程的流程图。
- [0008] 图 3 是示例性的被监视音频流的时域表示。
- [0009] 图 4 是一部分被监测音频流（即，音频块）的示例的曲线图，其为正弦曲线。
- [0010] 图 5 是可以应用于图 4 的音频块的示例性窗口的曲线图。
- [0011] 图 6 是将图 5 的窗口应用于图 4 的音频块得到的加窗后音频块的曲线图。
- [0012] 图 7 是可以应用于图 4 的音频块的第二示例性窗口的曲线图。
- [0013] 图 8 是将图 7 的窗口应用于图 4 的音频块得到的加窗后音频块的曲线图。
- [0014] 图 9 是图 5 的窗口的曲线图，其示出了所述窗口的相应频率响应。
- [0015] 图 10 是图 7 的窗口的曲线图，其示出了所述窗口的相应频率响应。
- [0016] 图 11 是第二另选示例性窗口的曲线图及其相应频率响应。
- [0017] 图 12 是第三另选示例性窗口的曲线图及其相应频率响应。
- [0018] 图 13 是示例性签名匹配过程的流程图。
- [0019] 图 14 是示出了可以如何根据图 13 的流程图来比较签名的图。
- [0020] 图 15 是基于音频流或音频块生成签名的示例性签名生成系统的框图。
- [0021] 图 16 是用于对签名进行比较的示例性签名比较系统的框图。
- [0022] 图 17 是可以用于实现本文所述的方法和装置的示例性处理器系统的框图。

## 具体实施方式

[0023] 尽管下面公开了在其他组件当中利用硬件上执行的软件实现的示例性系统，但应当注意的是，这些系统仅仅是说明性的，而不应当看作是限制。例如，可以预料，这些硬件和软件组件全部或部分都能够单独以硬件、单独以软件或者以硬件和软件的任何组合来实施。因此，尽管下面描述了示例性实施方式，本领域的技术人员很容易理解，所提供的示例不是实现这些系统的仅有的方式。

[0024] 本文描述的方法和装置一般地涉及生成可以用于标识媒体信息的数字签名。数字签名，或数字指纹，是准确表征用于匹配、索引或数据库搜索和检索的信号的信号描述符。具体地说，从基于音频流或音频块（例如，音频信息）生成数字音频签名的方面描述了所公开的方法和装置。然而，本文描述的方法和装置还可以用于基于任何其它类型的信号、时间序列数据和媒体信息（例如，视频信息、网页、静态图像、计算机数据等）生成数字签名。此外，媒体信息可以与广播信息（例如，电视信息、无线电信息等）、从任何存储介质（例如，压缩盘（CD）、数字多功能盘（DVD）等）再现的信息或者与音频流、视频流相关联的任何其它信息或者为之生成数字签名的任何其它媒体信息相关联。在一个具体示例中，基于包括监测地（例如，被监测家庭）处生成的被监测数字签名和在基准地和 / 或中心数据收集设备处生成和 / 或存储的基准数字签名标识音频流。

[0025] 如下文详细所述的，本文描述的方法和装置基于数字签名标识媒体信息（包括音频流或任何其它媒体）。本文描述的示例技术例如利用音频样本的单个音频块在特定时刻计算签名，但是利用两个或更多个窗口函数来处理该音频块以得到两个或更多个加窗后音频块。对加窗后音频块的进一步处理使得能够检测对该音频块的音频频谱的窗口效应。根据两个或更多个窗口函数对该音频块的效应得出该音频块唯一或基本唯一的签名值。即，

本文描述的示例性技术使得能够在不使用时间置换音频块的情况下计算或确定音频签名。当然,可以调节窗口函数的选择,也可以选择用于确定签名的变换类型、参数和 / 或分辨率。

[0026] 如下文详细描述,在对音频样本块应用窗口函数之后,通过例如利用离散傅里叶变换 (DFT) 或不管是否基于傅里叶变换的任何其它合适的变换 (例如,离散余弦变换 (DCT)、修正离散余弦变换 (MDCT)、哈尔 (Haar) 变换、沃尔什 (Walsh) 变换等) 将加窗后音频块从时域变换到频域,来生成加窗后音频块的频率成份。可以使用该变换来分析加窗后音频块中的频率成份并标识各频率成份的频谱功率。然后可以利用频谱功率生成数字签名。

[0027] 可以在将窗口函数应用于音频块之后使用其他技术。例如,可以利用将音频数据从时域变换到小波域的小波变换处理加窗后音频块。通常,小波变换可以用于将数据块或帧 (例如,时域音频样本) 分解成多个子波段,由此使得能够以不同比例和 / 或分辨率分析数据集。通过将数据分成多个子波段,小波变换可以用于以期望比例或分辨率分析数据的各个时间间隔。

[0028] 另选地,除向音频样本的时域块应用时域窗口函数外,也能够频域中实现加窗,其中,可以将与时域窗口对应的频率响应与音频块的频谱进行卷积。如果使用包括卷积的频域处理,则可以利用傅里叶变换执行音频块到频域的转换,其中,在音频块之间进行调节,以解决不连续性。此外,如果在频域中实现窗口的处理和应用,则可以选择频率特征具有多个非零元素 (例如,3 到 5 个非零元素) 的时域窗口。

[0029] 可以利用上述技术基于与观众 / 听众消费的或呈现给观众 / 听众的媒体信息 (例如,被监测音频流) 相关联的音频流在监测地生成被监测签名。例如,可以基于在监测地呈现的电视节目的音轨或任何其他媒体生成被监测签名。然后可以将被监测签名传送给中心数据收集设备,用于与一个或更多个基准签名进行比较。

[0030] 利用上述技术对与已知媒体信息相关联的音频流在基准地和 / 或中心数据收集设备处生成基准签名。该已知媒体信息可以包括在一个区域内广播的媒体、在家庭内再现的媒体、通过互联网接收到的媒体等。各基准签名与诸如歌曲标题、电影标题等的媒体标识信息一起存储在存储器中。当在中心数据收集设备处接收到被监测签名时,将被监测签名与一个或更多个基准签名相比较,直到找到匹配为止。然后可以将该匹配信息用于标识从中生成被监测签名的媒体信息 (例如,被监测音频流)。例如,可以查阅查询表或数据库,以检索与从中生成被监测签名的媒体信息相应的媒体标题、节目标号、幕号等。

[0031] 在一个示例中,生成被监测签名和基准签名的速率可以不同。例如,为了处理或其他方面,被监测签名可以是基准签名的数据率的 25%。例如,可以每 0.032 秒生成 48 位基准签名,这样得到基准数据率为 48 位 \*31.25/秒或 187.5 字节 / 秒。在该构造中,可以每 0.128 秒生成 48 位基准签名,这样得到基准数据率为 48 位 \*7.8125/秒或 46.875 字节 / 秒。当然,在被监测签名和基准签名的数据率不同的构造下,在将被监测签名与基准签名进行比较时必需考虑这种差别。例如,如果监测速率是基准速率的 25%,则各连续被监测签名将对应于每第四个基准签名。

[0032] 图 1A 和 1B 例示了用于生成数字频谱签名和标识音频流的示例性音频流标识系统 100 和 150。示例性音频流标识系统 100 和 150 可以分别实现为电视广播信息标识系统

和无线电广播信息标识系统。示例性音频流标识系统 100 包括监测地 102 (例如,被监测家庭)、基准地 104 和中心数据收集设备 106。

[0033] 监测电视广播信息涉及基于电视广播信息的音频数据在监测地 102 生成被监测签名并将该被监测签名通过网络 108 传送给中心数据收集设备 106。基准签名可以在基准地 104 生成,也可以通过网络 108 传送给中心数据收集设备 106。在监测地 102 处生成的被监测签名表示的音频内容可以通过将被监测签名与一个或多个基准签名进行比较直到找到匹配为止来在中心数据收集设备 106 处进行标识。另选地,可以将被监测签名从监测地 102 传递到基准地 104,并且在基准地 104 处与一个或多个基准签名进行比较。在另一示例中,可以将基准签名传送到监测地 102 并与监测地 102 处的被监测签名进行比较。

[0034] 例如,监测地 102 可以是监测观众/听众的媒体消费的家庭。通常,监测地 102 可以包括多个媒体投放装置 110、多个媒体呈现装置 112 和用于生成与监测地 102 处呈现的媒体相关联的被监测签名的签名生成器 114。

[0035] 该多个媒体投放装置 110 可以包括例如机顶盒调谐器(例如,电缆调制器、卫星调制器等)、PVR 装置、DVD 播放器、CD 播放、收音机等。例如机顶盒调谐器的一些或所有媒体投放装置 110 可以可通信地连接到一个或多个广播信息接收装置 116,广播信息接收装置 116 可以包括电缆、圆盘式卫星电视天线(satellite dish)、天线和/或用于接收广播信息的任何其他合适装置。媒体投放装置 110 可以被配置成例如基于广播信息和/或存储的信息再现媒体信息(例如,音频信息、视频信息、网页、静态图像等)。可以从广播信息接收装置 116 获得广播信息,并可以从任何信息存储介质(例如,DVD、CD、磁带等)获得存储的信息。媒体投放装置 110 可通信地连接到媒体呈现装置 112 并且可被配置成向媒体呈现装置 112 传送媒体信息以用于呈现。媒体呈现装置 112 可以包括具有显示装置和/或一组扬声器的电视,例如,观众/听众由此消费广播电视信息、音乐、电影等。

[0036] 如下文更详细描述,签名生成器 114 可以用于基于音频信息生成被监测数字签名。具体地说,在监测地 102 处,签名生成器 114 可以被配置成基于由媒体投放装置 110 再现和/或由媒体呈现装置 112 呈现的被监测音频流生成被监测签名。签名生成器 114 可以通过音频监测接口 118 可通信地连接到媒体投放装置 110 和/或媒体呈现装置 112。这样,签名生成器 114 可以获得与媒体投放装置 110 再现和/或由媒体呈现装置 112 呈现的媒体信息相关联的音频流。附加地或另选地,签名生成器 114 可以可通信地连接到置于媒体呈现装置 112 附近的麦克风(未示出)以检测音频流。签名生成器 114 还可以通过网络 108 可通信地连接到中心数据收集设备 106。

[0037] 网络 108 可以用于在监测地 102、基准地 104 和中心数据收集设备 106 之间交流签名(例如,数字频谱签名)、控制信息和/或配置信息。诸如广播电缆网络、DSL 网络、蜂窝电话网络、卫星网络和/或任何其它通信网络的任何有线或无线通信系统都可以用于实现网络 108。

[0038] 如图 1A 所示,基准地 104 可以包括多个广播信息调谐器 120、基准签名生成器 122、发射器 124、数据库或存储器 126 和广播信息接收装置 128。基准签名生成器 122 和发射器 124 可以可通信地连接到存储器 126 以将基准签名存储在存储器 126 中和/或从存储器 126 检索存储的基准签名。

[0039] 广播信息调谐器 120 可以可通信地连接到广播信息接收装置 128,广播信息接收

装置 128 可以包括电缆、天线、圆盘式卫星天线和 / 或用于接收广播信息的任何其他合适装置。各广播信息调谐器 120 可以被配置成调谐到特定广播频道。通常,基准地 104 处的调谐器的数量等于特定广播区域中可用的频道数量。这样,可以针对广播区域内所有频道上发送的所有媒体信息生成基准签名。可以将调谐的媒体信息的音频部分从广播信息调谐器 120 传送给基准签名生成器 122。

[0040] 基准签名生成器 122 可以被配置成获取特定广播区域中可用的所有媒体信息的音频部分。然后基准签名生成器 122 可以基于该音频信息生成多个基准签名(例如,利用下面更详细描述的处理)并在存储器 126 中存储所述基准签名。尽管图 1 中示出了一个基准签名生成器,但是在基准地 104 中可以使用多个基准签名生成器。例如,所述多个签名生成器中的各签名生成器可以可通信地连接到各自的广播信息调谐器 120。

[0041] 发射器 124 可以可通信地连接到存储器 126 并被配置成从存储器 126 检索签名并将基准签名通过网络 108 传送给中心数据收集设备 106。

[0042] 中心数据收集设备 106 可以被配置成将从监测地 102 接收到的被监测签名与从基准地 104 接收到的基准签名相比较。此外,中心数据收集设备 106 可以被配置成通过将被监测签名与基准签名相匹配并利用该匹配信息来从数据库检索电视目标标识信息(例如,节目标题、广播时间、广播频道等)来标识被监测音频流。中心数据收集设备 106 包括接收器 130、签名分析器 132 和存储器 134,所有这些都可以通过通信地如图 1 所示连接起来。

[0043] 接收器 130 可以被配置成通过网络 108 接收被监测签名和基准签名。接收器 130 可通信地连接到存储器 134 并被配置成将被监测签名和基准签名存储在存储器 134 中。

[0044] 签名分析器 132 可以用于将基准签名与被监测签名进行比较。签名分析器 132 可通信地连接到存储器 134 上并被配置成从该存储器 134 中检索被监测签名和基准签名。签名分析器 132 可以被配置成从存储器 134 中检索被监测签名和基准签名,并将被监测签名与基准签名进行比较,直到找到匹配。可以利用任何机器可访问信息存储介质(例如,一个或多个硬盘、一个或多个光存储装置等)来实现存储器 134。

[0045] 尽管在图 1A 中,签名分析器 132 位于中心数据收集设备 106 处,但是签名分析器 132 还可以位于基准地 104 处。在这样的结构中,可以将被监测签名从监测地 102 通过网络 108 传送到基准地 104。另选地,存储器 134 可以位于监测地 102 处,并且可以由发射器 124 通过网络 108 将基准签名周期性地添加到存储器 134 中。另外,尽管签名分析器 132 示出为与签名生成器 114 和 122 是分立的装置,但签名分析器 132 也可以与基准签名生成器 122 和 / 或签名生成器 114 集成在一起。另外,尽管图 1 描绘了单个监测地(即,监测地 102)和单个基准地(即,基准地 104),但是可以通过网络 108 将多个这种场地连接到中心数据收集设备 106。

[0046] 图 1B 的音频流标识系统 150 可以被配置成监测并标识与无线电广播信息相关的音频流或任何其它媒体。通常,音频流标识系统 150 用于监测在特定广播区域中多个无线电台广播的内容。与用于监测观众消费的电视内容的音频流标识系统 100 不同,音频流标识系统 150 可以用于监测广播区域内广播的音乐、歌曲等以及广播它们的次数。该类型的媒体轨道可以用于确定与各音频组成的使用费支付,版权的正确使用等。音频流标识系统 150 包括监测地 152、中心数据收集设备 154 和网络 108。

[0047] 监测地 152 被配置成接收特定广播区域可用的所有广播信息并基于无线电广播

信息生成被监测签名。监测地 152 包括多个广播信息调谐器 120、发射器 124、存储器 126 和广播信息接收装置 128, 所有这些都结合图 1A 在上面进行了描述。此外, 监测地 152 包括签名生成器 156。当在音频流标识系统 150 中使用时, 广播信息接收装置 128 被配置成接收无线电广播信息, 而广播信息调谐器 120 被配置成调谐到无线电广播台。监测地 152 处的广播信息调谐器 120 的数量可以等于特定广播区域中的无线电广播台的数量。

[0048] 签名生成器 156 被配置成从各广播信息调谐器 120 接收所调谐到的音频信息并对其生成被监测签名。尽管仅示出了一个签名生成器 (即, 签名生成器 156), 但监测地 152 可以包括多个签名生成器, 各签名生成器可以可通信地连接到广播信息调谐器 120 中的一个上。签名生成器 156 可以将被监测签名存储在存储器 126 中。发射器 124 可以从存储器 126 中检索被监测签名, 并将它们通过网络 108 传送给中心数据收集设备 154。

[0049] 中心数据收集设备 154 被配置成从监测地 152 接收被监测签名, 基于基准音频流生成基准签名, 并将被监测签名与基准签名进行比较。中心数据收集设备 154 包括接收器 130、签名分析器 132 和存储器 134, 所有这些都结合图 1A 在上面进行了更详细的描述。此外, 中心数据收集设备 154 包括基准签名生成器 158。

[0050] 基准信号生成器 158 被配置成基于音频流生成基准签名。基准音频流可以存储在任意类型的机器可访问介质, 例如, CD、DVD、数字音带 (DAT) 等上。通常, 艺术家和 / 或录制制作公司将其音频作品 (即, 音乐、歌曲等) 发送到中心数据收集设备 154 以便添加到基准库中。基准签名生成器 158 可以从机器可访问介质读取音频数据并基于各音频作品 (例如, 图 3 的所捕获的音频 300) 生成多个基准签名。然后基准签名生成器 158 可以将基准签名存储在存储器 134 中, 以用于签名分析器 132 随后的检索。与各基准音频流相关联的标识信息 (例如, 歌曲标题、艺术家姓名、音轨号码等) 可以存储在数据库中, 并且可以基于基准签名索引。这样, 中心数据收集设备 154 包括数据库, 所述数据库包括基准签名以及对应于所有已知和可获得歌曲标题的标识信息。

[0051] 接收器 130 被配置成从网络 108 接收被监测签名并将被监测签名存储在存储器 134 中。签名分析器 132 从存储器 134 检索被监测签名和基准签名, 以用于标识在广播区域内广播的被监测音频流。签名分析器 132 可以首先使被监测签名与基准签名匹配来标识被监测音频流。然后利用匹配信息和 / 或匹配的基准签名来从存储器 134 中存储的数据库检索标识信息 (例如, 歌曲标题、歌曲音轨、艺术家等)。

[0052] 尽管图 1B 中示出了一个监测地 (例如, 监测地 152), 但是多个监测地可以可通信地连接到网络 108 并被配置成生成被监测签名。具体地说, 各监测地可以位于各自的广播区域中并且被配置成监测各自广播区域内的广播台的内容。

[0053] 图 2 是示出了示例性签名生成过程 200 的流程图。如图 2 所示, 签名生成过程 200 首先捕获要由签名表征的音频块 (框 202)。图 3 中附图标记 300 处示出了可能捕获到的示例性音频时域曲线。例如可以通过与音频源的有线连接或通过与音频源的无线连接 (例如, 音频传感器) 从音频源捕获音频。如果音频源是模拟的, 则该捕获包括利用例如模数转换器对模拟音频源进行抽样。在一个示例中, 音频源可以以 8 千赫兹 (kHz) 速率 (称为抽样速率 ( $F_s$ )) 进行抽样, 这意味着由其数字抽样表示模拟音频, 该抽样是以每秒 8000 个样本的速率或每 125 微秒 1 个样本的速率所取的。可以由单音 (monoaural) 16 位分辨率表示各音频样本。



[0054] 在一个示例中,对应于 8192 个样本的音频块被捕获以用于处理。以前面的抽样率 8kHz,这对应于 1.024 秒的音频。然而,这只是一个示例,所收集的样本数量可以对应于范围在大约 1 秒到 2 秒持续时间内的任何音频片段。通常,这里用变量  $N$  来表示音频块中捕获的样本数量。因此,在上述示例中, $N = 8192$  并且所捕获的音频的时间范围对应于  $t \dots t + N/F_s$ 。图 4 中附图标记 402 处示出了音频块的表示,其中,为了例示的目的,音频块对应于正弦曲线。

[0055] 在捕获到音频块(框 202)之后,过程 200 对音频块应用第一窗口函数(称为  $W_1$ ) (框 204A),以生成第一加窗后音频块。此外,过程 200 利用第二窗口函数(称为  $W_2$ ) 对音频块进行加窗(框 204B),以产生第二加窗后音频块。例如,例如如图 5 中附图标记 502 所示,窗口可以是高斯(Gaussian)或贝尔(bell)型函数,其中, $W_1$ 502 的高端和低端具有零值,而窗口 502 的中心的值为 1。在一个示例中,加窗是窗口函数的值与音频块的各样本之间的样本方面的乘法(sample-wise multiplication)。例如,利用窗口 502 对音频块 402 进行加窗,得到加窗的音频块 602,如图 6 所示,其中,在窗口 502 的两端处加窗后音频块 602 的幅度为零,并且加窗后音频块 602 的中心处具有与音频块 402 相同的幅度。

[0056] 另选地,除向音频块应用利用窗口函数的样本方面的乘法在时域应用窗口函数外,也能够在此域中实现加窗,其中,可以将与时域窗口对应的频率响应与音频块的频谱进行卷积。如上所述,如果使用包括卷积的频域处理,则可以利用傅里叶变换执行音频块到频域的转换,其中,在音频块之间进行调节,以解决不连续性。此外,如果在频域中实现窗口的处理和应用,则可以选择频率特征具有多个非零元素(例如,3 到 5 个非零元素)的时域窗口。

[0057] 对  $W_1$  和  $W_2$  选择的窗口可以在实质上是互补的。例如,如果对  $W_1$  选择了图 5 中所示的窗口 502,则可以对  $W_2$  选择图 7 所示的窗口 702。如图 7 所示,窗口 702 是窗口  $W_1$  的反转形式,即, $W_2(k) = 1 - W_1(k)$ ,其中, $k$  是窗口域内的抽样索引。窗口  $W_2$  在窗口 702 的高端和低端达到统一值,并且在窗口 702 的中心具有零值。因此,当窗口 702 应用于音频块 402 时,如图 8 所示,得到加窗后音频块 802。如图 8 所示,加窗后音频块 802 在其中心具有零幅度,但在加窗后音频块 802 的两端具有基本与音频块 402 的幅度匹配的幅度。

[0058] 如图 9 和图 10 所示,窗口 502 和 702 具有各自的频率响应 902 和 1002。因此,将窗口 502 和 702 应用于音频块(例如,音频块 402)影响了音频块的频谱。如下所述,被检测的是不同窗口对音频块的不同影响,以确定表示音频块的签名。

[0059] 尽管为以上说明选择的窗口 502、702 分别类似于 Hann 窗口和反转 Hann 窗口,但也可以使用其它窗口形状。例如,如图 11 和图 12 所示,可以选择两个非对称窗口 1102、1202,其中,第一窗口 1102 占据加窗空间的上半,而第二窗口 1202 占据加窗空间的下一半。非对称窗口 1102、1202 的频率响应相同,如图 11 和图 12 附图标记 1104 和 1204 处所示,但是因为窗口作用于音频块的大部分不同的部分上,因此加窗的结果对于不是正弦曲线的音频信号具有不同的频谱特性。

[0060] 尽管本文描述了窗口形状的特定示例,但也可以使用其它窗口形状。例如,可以对第一窗口和第二窗口(例如, $W_1$  和  $W_2$ ) 两者任意选择窗口形状,其中,从一组窗口函数中作出选择。当然,可以在不同时刻使用不同窗口,只要监测地和基准地使用同样的即可。此外,可以使用多于两个窗口。

[0061] 返回到图 2,在完成加窗(框 204A 和 204B)之后,分别对加窗后音频块进行变换(框 206A 和 206B)。在一个示例中,该变换可以是从时域到频域的变换。例如,已经加窗的捕获的音频的  $N$  个样本可以转换成由  $N/2$  复合 DFT 系数表示的音频频谱。另选地,可以使用任何合适的变换,例如,小波变换、DCT、MDCT、哈尔变换、沃尔什变换等。

[0062] 在完成变换(框 206A 和 206B)之后,过程 200 对各变换的结果进行特征化(框 208A 和 208B)。例如,该过程可以确定各变换结果的  $K+1$  个不同波段的各波段内的能量。即,利用窗口  $W_1$  获得的加窗后音频的变换的结果(框 206A)可以例如分成 16 不同波段,并且可以确定 16 个不同波段的各波段中的能量。这可以表示为  $E_j(w_1)$ ,其中,  $j$  范围从 0 到 15,并且  $w_1$  表示该能量与将窗口  $W_1$  应用于抽样后的音频(即,音频块)得到的频谱相关联。类似地,利用窗口  $W_2$  得到的加窗后音频块的变换结果(框 206B)可以例如分成 16 个不同波段,其能量可以确定并表示为  $E_j(w_2)$ ,其中,  $j$  范围为从 0 到 15,  $w_2$  表示该能量与应用窗口  $W_2$  得到的频谱相关联。另选地,可以使用能量以外的不同频谱特征来表征这些结果。例如,可以使用能量分布的频谱平坦度。

[0063] 在对各组变换结果进行特征化(框 208A 和 208B)之后,过程 210 比较这些特征化的结果。例如,可以相互减去各波段的这些特征化的结果。在一个示例中,可以将中间值计算为  $d_j = E_j(w_2) - E_j(w_1)$ ,其中,  $j$  范围从 0 到  $K$ 。对于  $K = 15$  的上述特定示例,可以计算中间值  $d_j$ ,其中,  $d_j = E_j(w_2) - E_j(w_1)$ ,并且  $j$  范围为从 0 到 15。因此,这种比较结果得到 16 种不同的中间值(例如,  $d_0, d_1, d_2 \dots d_{15}$ ),其中,各中间值是例如对加窗后音频块进行变换得到的频谱的相似频段的特征的差。

[0064] 在计算出中间值以表示多个特征的比较(框 210)之后,过程 200 基于该比较确定签名(框 212)。例如,如果中间值  $d_j > 0$ ,则签名位  $S_j$  可以指派值 1,否则可以指派值 0,其中,  $j$  范围从 0 到  $K$ 。更具体地说,这在上述  $K = 15$  的示例中提到的,因此存在中间值与值 0 的 16 种比较,并且基于这些比较,将生成 16 位签名,来表示图 2 的框 202 处捕获的音频块。在确定了签名(框 212)之后,过程 200 循环进行(框 214)并捕获附加音频(框 202),以得到附加签名。

[0065] 尽管前面描述了选择第一窗口( $W_1$ )和第二窗口( $W_2$ )并且利用所选择的窗口来确定捕获的音频块的所有签名,但其它方式也是可能的。例如,可以利用第一窗口对(例如, $W_1$ 和 $W_2$ )确定表示捕获的音频块的签名的一些位,而可以利用不同的窗口对(例如, $W_3$ 和 $W_4$ )来确定签名的其它位。此外,可以使用第三窗口对(例如, $W_1$ 和 $W_3$ )来确定附加签名位。在一些情况下,可以以预定方式或任意方式选择唯一的窗口对来确定各签名位的值,只要在基准地选择了相同的窗口对来作用于同样的窗口块即可。

[0066] 前面描述了签名技术,其可以用来确定表示捕获的音频的一部分的签名。图 13 示出了一个示例性签名匹配过程 1300,其可以执行用来将基准签名(即,基准地确定的签名)与被监测签名(即,监测地确定的签名)进行比较。签名匹配的最终目的是要找出询问的音频签名(例如,被监测音频)和数据库中的签名(例如,基于基准音频所取的签名)之间最接近的匹配。可以在基准地、监测地或能够访问被监测签名和包含基准签名的数据库的任何其它数据处理地执行该比较。

[0067] 现在来详细描述图 13 的示例性方法,示例性过程 1300 涉及获取被监测签名及其相关的定时(框 1302)。如图 14 所示,签名集可以包括多个被监测签名,在图 14 中附图标记

802、804 和 806 处示出了其中三个。各签名由  $\sigma$  (sigma) 表示。各被监测签名 1402、1404、1406 可以包括定时信息 1408、1410、1412, 无论该定时信息是隐式的还是显式的。

[0068] 然后查询包含基准签名的数据库 (框 1304), 以标识数据库中具有最接近匹配的签名。在一种实现方式中, 签名之间的相似指标 (接近度) 可以取为汉明 (Hamming) 距离, 即, 查询的值和基准位串不同的位置的号码。在图 14 中, 在附图标记 1416 处示出了签名和定时信息的数据库。当然, 数据库 1406 可以包括来自不同媒体表示的任何数量的不同签名。然后在与匹配的基准签名相关联的节目和未知签名之间进行关联 (框 1306)。

[0069] 可选地, 过程 1300 可以之后确定被监测签名和基准签名之间的偏差 (框 1308)。需要偏差值, 以便更好更确信地决定查询签名块是否良好地匹配了基准签名。由于监测 (观测) 的原因, 通常将短查询块中所有的签名的偏差值保持为相对于各基准签名几乎恒定。

[0070] 在多于一个基准签名的所有描述符都与预定汉明距离阈值以下的汉明距离相关联的情况下, 可能需要多于一个的被监测签名与可能的匹配基准音频流的各基准签名相匹配。相对来说不可能的是, 基于被监测音频流生成的所有被监测签名会与多于一个基准音频流的所有基准签名相匹配, 因此, 可以防止将多于一个的基准音频流与被监测音频流进行错误地匹配。

[0071] 可以通过硬件、软件和 / 或它们的任何组合来实现上述示例性方法、过程和 / 或技术。更具体地说, 可以以图 15 和图 16 的框图定义的硬件来执行这些示例性方法。也可以通过诸如图 16 的处理器系统 1610 的处理器系统上执行的软件来实现这些示例性方法、过程和 / 或技术。

[0072] 图 15 是用于生成数字频谱签名的示例性签名生成系统 1500 的框图。具体地说, 示例性签名生成系统 1500 可以用于基于上述的对音频块进行加窗、变换、特征化以及比较来生成被监测签名和 / 或基准签名。例如, 示例性签名生成系统 1500 可以用于实现图 1A 的签名生成器 114 和 122 或者图 1B 的签名生成器 156 和 158。另外, 示例性签名生成系统 1500 可以用于实现图 2 的示例性方法。

[0073] 如图 15 所示, 示例性签名生成系统 1500 包括样本生成器 1502、定时装置 1503、基准时间生成器 1504、加窗器 1506、变换器 1508、特征确定器 1510、比较器 1512、签名确定器 1514、存储器 1516 和数据通信接口 1518, 如图所示, 所有这都可以可通信地连接。示例性签名生成系统 1500 可以被配置成获得示例性音频流、从示例性音频流获取多个音频样本形成音频块, 并从该单个音频块生成代表它的签名。

[0074] 样本生成器 1502 可以被配置成获取示例性音频流, 例如, 得到图 3 的捕获音频的流 300。流 300 可以是任何模拟或数字音频流。如果示例性音频流是模拟音频流, 则样本生成器 1502 实现为利用模数转换器。如果示例性音频流是数字音频流, 则样本生成器 1502 可以实现为利用数字信号处理器。此外, 样本生成器 1502 可以被配置成以任何期望抽样频率  $F_s$  获取和 / 或提取音频样本。例如, 如上所述, 样本生成器可以被配置成以 8kHz 获取  $N$  个样本, 并且利用 16 位来表示各样本。在这种安排下,  $N$  可以是任何数量的样本, 例如,  $N = 8192$ 。样本生成器 1502 还可以通知基准时刻生成器 1504 何时开始音频样本采集过程。样本生成器 1502 将样本传送给加窗器 1506。

[0075] 定时装置 1503 可以被配置成生成时间数据和 / 或时间戳信息, 并且可以由时钟、定时器、计数器和 / 或任何其它合适的装置来实现。定时装置 1503 可以可通信地连接到

基准时刻生成器 1504, 并且可以被配置成将时间数据和 / 或时间戳传送给基准时刻生成器 1504。定时装置 1503 还可以可通信地连接到样本生成器 1502, 并且可以表明起始信号或中断, 以指示样本生成器 1502 开始收集或采集音频样本数据。在一个示例中, 定时装置 1503 可以通过具有以毫秒为分辨率、记录时间的 24 小时周期制的实时时钟实现。在这种情况下, 定时装置 1503 可以被配置成在午夜时刻复位到零, 并相对于午夜时刻以毫秒记录时间。然而, 通常时间戳可以将完整的年、月、日、时、分、秒信息表示为从过去的预定时刻经过的秒数, 例如, 2005 年 1 月 1 日 00:00AM。可以根据确定的收集的音频签名的采集速率进行推导来增加到次秒 (subsecond) 级分辨率。

[0076] 基准时刻生成器 1504 可以在从样本生成器 1502 接收到通知时初始化基准时刻  $t_0$ 。基准时刻  $t_0$  可以用于指示音频流内生成签名的时刻。具体地说, 基准时刻生成器 1504 可以被配置成在被样本生成器 1502 通知开始样本采集过程时从定时装置 1503 读取时间数据和 / 或时间戳值。然后基准时刻生成器 1504 可以将该时间戳值存储为基准时刻  $t_0$ 。

[0077] 例如, 加窗器 1506 将例如两个窗口应用于样本生成器 1502 输出的音频块。因此, 加窗器 1506 得到的是两个加窗后的音频块。如上所述, 窗口可以是任何窗口集。然而, 互补窗口是优选的, 因为它们很容易保证在平均上说能量的两个值相同, 从而导致准相等的位分布。

[0078] 变换器 1508 可以被配置成对各加窗的音频块执行 N 点 DFT, 其中, N 是由样本生成器 1502 获得的样本数量。例如, 如果样本生成器获取 8192 个样本, 则变换器将根据这些样本生成频谱, 其中, 该频谱由 4096 个复值傅立叶系数表示。

[0079] 特征确定器 1510 可以被配置成标识变换器 1508 生成的 DFT 频谱特征内的几个频段 (例如, 16 个频段)。所选择的频段可以但最好不相互交叠。可以根据任何技术选择这些频段。当然, 可以选择任何数量的合适频段 (例如, 48 个)。特征确定器 1510 然后确定各频段的特征。例如, 特征确定器 1510 可以确定各频段的能量。因此, 特征确定器 1510 得到的是针对例如 16 个频段中各频段的两个特征集。例如, 如果选择 16 个频段, 则特征确定器 1510 的输出将是 32 个能量指标, 各 DFT 中每个频段一个能量指标。可以由  $E_j(w1)$  和  $E_j(w2)$  表示这些特征,  $j$  范围为从 0 到 K (例如, 0 到 15), 并且  $w1$  和  $w2$  分别表示窗口 1 和窗口 2。

[0080] 比较器 1512 比较各频段的特征, 以确定中间值。例如, 比较器 1512 可以根据  $d_j = E_j(w2) - E_j(w1)$  生成中间值, 使得相互减去 DFT 的各频段中的能量。

[0081] 签名确定器 1514 对从比较器 1512 得到的值进行运算, 以针对各中间值生成一个签名位。该运算可以与结合图 2 在上面描述的过程 212 非常类似或相同。即, 签名位值可以基于中间值与零的比较。将签名位输出到存储器 1516。

[0082] 存储器可以是用于容纳签名存储器的任何合适介质。例如, 存储器 1516 可以是诸如随机存取存储器 (RAM)、闪存等的存储器。附加地或另选地, 存储器 1516 可以是诸如硬盘驱动器、光学存储介质、带式驱动器等的大容量存储器。

[0083] 存储器 1516 连接到数据通信接口 1518。例如, 如果系统 1500 是监测地 (例如, 在个人家里), 则可以利用数据通信接口 1518 将存储器 1516 中的签名信息传送给收集设备、基准地等。

[0084] 图 16 是用于比较数字频谱签名的示例性签名比较系统 1600 的框图。具体地说,

示例性签名比较系统 1600 可以用于将被监测签名与基准签名进行比较。例如,示例性签名比较系统 1600 可以用于实现图 1A 的签名分析器 132,以将被监测签名与基准签名进行比较。另外,示例性签名比较系统 1600 可以用于实现图 13 的示例性过程。

[0085] 示例性签名比较系统 1600 包括被监测签名接收器 1602、基准签名接收器 1604、比较器 1606、汉明距离滤波器 1608、媒体标识器 1610 和媒体标识查询表接口 1612,如图所示,所有这些都可以可通信地连接。

[0086] 被监测签名接收器 1602 可以被配置成通过网络 108(图 1) 获取被监测签名并将被监测签名传送给比较器 1606。基准签名接收器 1604 可以被配置成从存储器 134(图 1A 和 1B) 获取基准签名并将基准签名传送给比较器 1606。

[0087] 比较器 1606 和汉明距离滤波器 1608 可以被配置成利用汉明距离将基准签名与被监测签名进行比较。具体地说,比较器 1606 可以被配置成将被监测签名的描述符与来自多个基准签名的描述符进行比较,并针对各比较生成汉明距离值。然后汉明距离滤波器 1608 可以从比较器 1606 获取汉明距离值,并基于该汉明距离值滤除不匹配的基准签名。

[0088] 在找出匹配基准签名后,媒体标识器 1610 可以获取匹配的基准签名,并结合媒体标识查询表接口 1612,可以标识与未标识的音频流(例如,图 3 的示例性被监测音频流 300) 相关联的媒体信息。例如,媒体标识查询表接口 1612 可以可通信地连接到媒体标识查询表或用于基于基准签名交叉引用媒体标识信息(例如,电影标题、演出标题、歌曲标题、艺术家姓名、幕号等)的数据库。这样,媒体标识器 1610 可以基于匹配的基准签名从媒体标识数据库检索媒体标识信息。

[0089] 图 17 是可以用于实现本文所述的装置和方法的示例性处理器系统 1710 的框图。如图 17 所示,处理器系统 1710 包括连接到互连总线或网络 1714 的处理器 1712。处理器 1712 包括寄存器组或寄存器空间 116,其在图 17 中描述为整体在片上(on-chip),但是另选地也可以整体或部分离片(off-chip) 并通过专用电连接和/或通过互连网络或总线 1714 直接连接到处理器 1712。处理器 1712 可以是任何合适的处理器、处理单元或微处理器。尽管图 17 中未示出,但系统 1710 可以是多处理器系统,并且因此可以包括与处理器 1712 相同或相似并且可通信地连接到互连总线或网络 1714 的一个或更多个附加处理器。

[0090] 图 17 的处理器 1712 连接到芯片组 1718,芯片组 1718 包括存储器控制器 1720 和输入/输出(I/O) 控制器 1722。众所周知,芯片组通常提供 I/O 和存储器管理功能以及可被连接到该芯片组的一个或更多个处理器访问或使用的多个通用和/或专用寄存器、定时器等。存储器控制器 1720 执行使处理器 1712(有多个处理器时使所述多个处理器) 能够访问系统存储器 1724 和大容量存储器 1725 的功能。

[0091] 系统存储器 1724 可以包括任何期望类型的易失性和/或非易失性存储器,例如,静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、闪存、只读存储器(ROM) 等。大容量存储器 1725 可以包括任何期望类型的大容量存储装置,包括硬盘驱动器、光驱、带式存储装置等。

[0092] I/O 控制器 1722 执行使得处理器 1712 能够通过 I/O 总线 1730 与外围输入/输出(I/O) 装置 1726 和 1728 通信的功能。I/O 装置 1726 和 1728 可以是任何期望类型的 I/O 装置,例如键盘、视频显示器或监视器、鼠标等。尽管存储器控制器 1720 和 I/O 控制器 1722 在图 17 中描述为芯片组 1718 内的独立功能块,但由这些块执行的功能可以集成在单个半

导体电路内或者可以利用两个或更多个独立的集成电路实现。

[0093] 本文描述的方法可以利用计算机可读介质上存储的由处理器 1712 执行的指令实现。计算机可读介质可以包括利用大容量存储装置（例如，盘驱动器）、可移动存储装置（例如，软盘、存储卡或棒等）和 / 或集成存储装置（例如，随机存取存储器、闪存等）的任何期望组合实现的固态、磁和 / 或光媒体的任何期望组合。

[0094] 尽管本文描述了特定的方法、装置和产品，但本专利的覆盖范围不限于此。

[0095] 相关申请的交叉引用

[0096] 本申请要求 2008 年 3 月 5 日提交的美国临时申请 61/033, 992 的优先权，通过引用将其全部内容并入本文中。

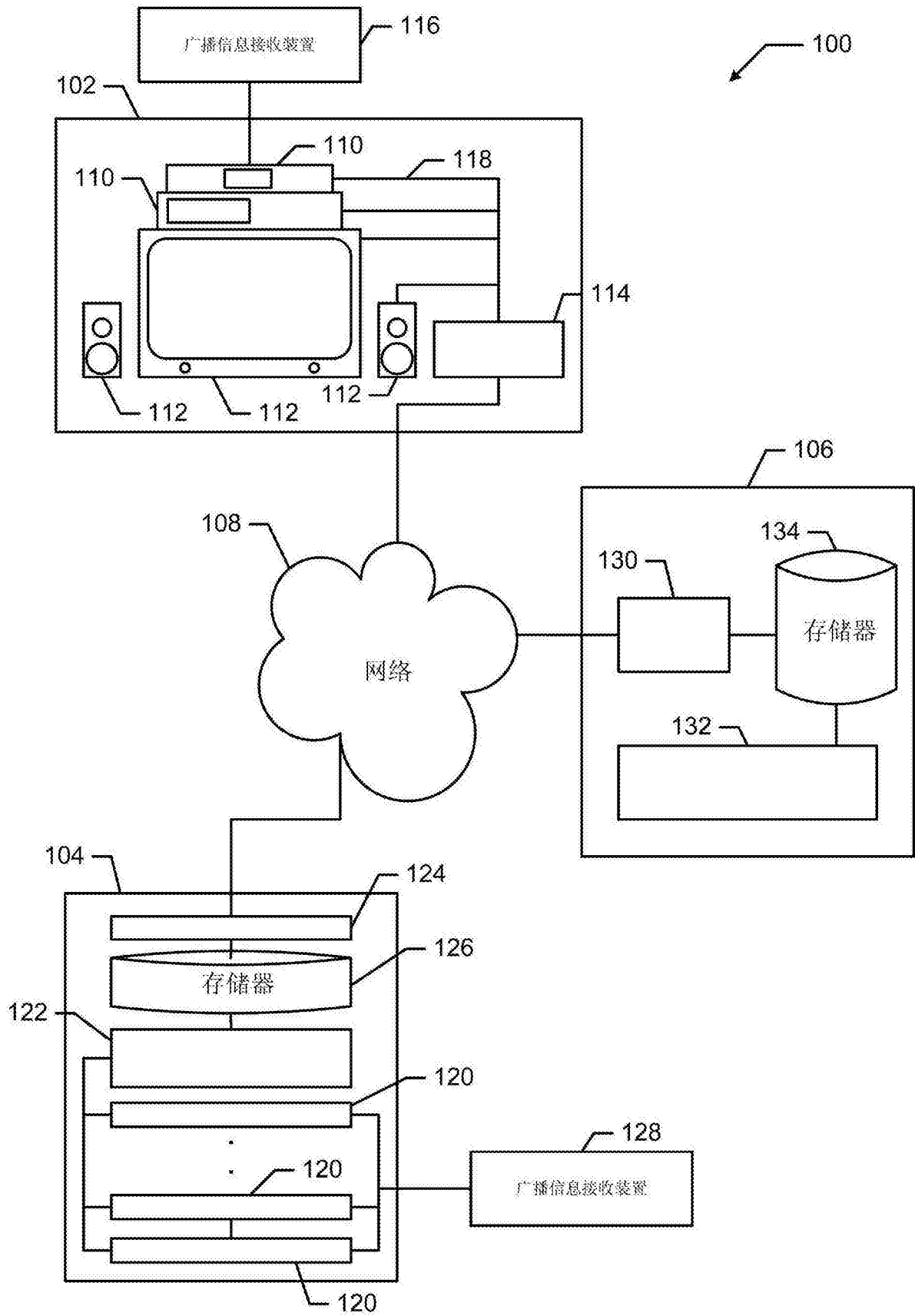


图 1A

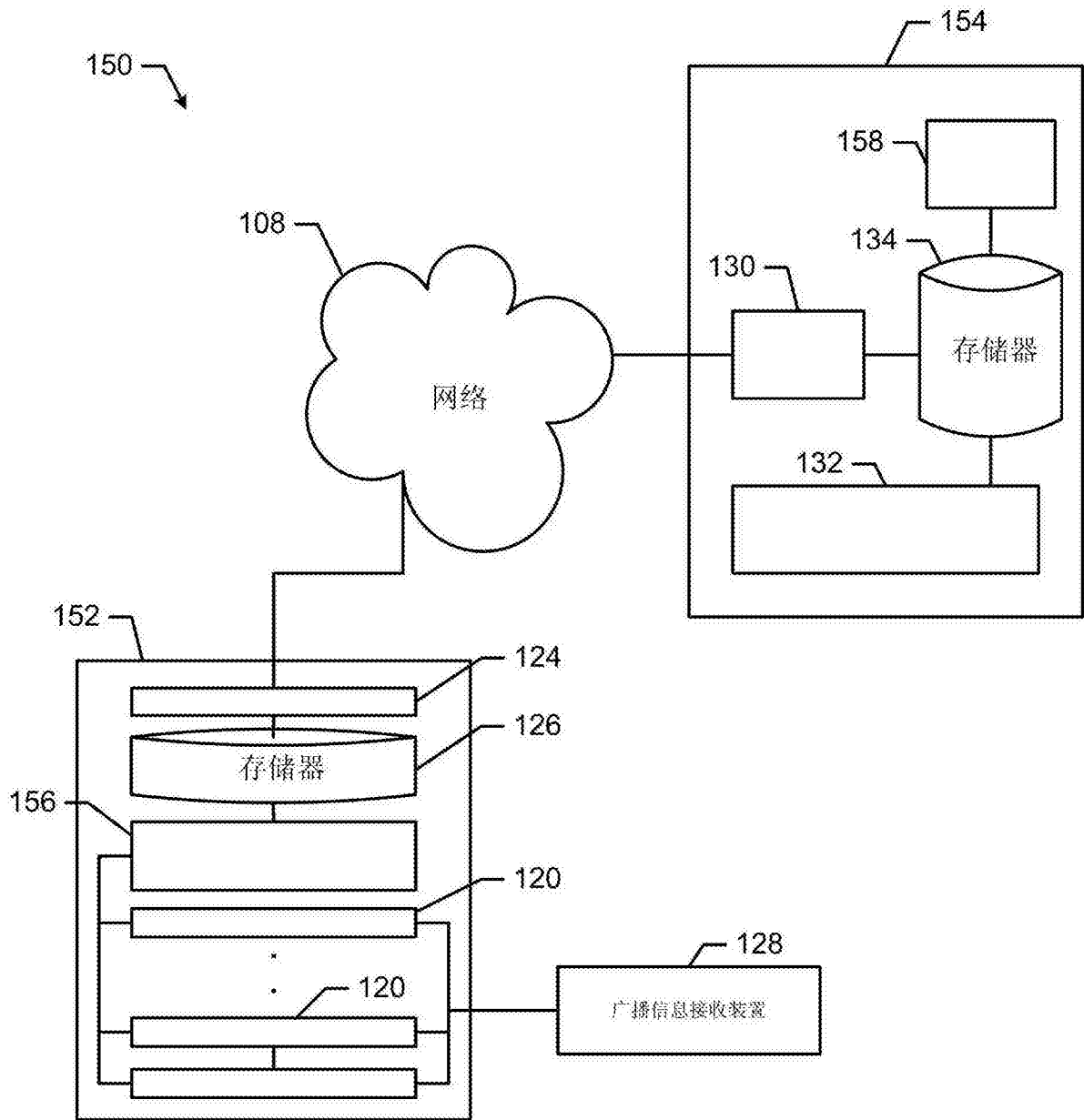


图 1B



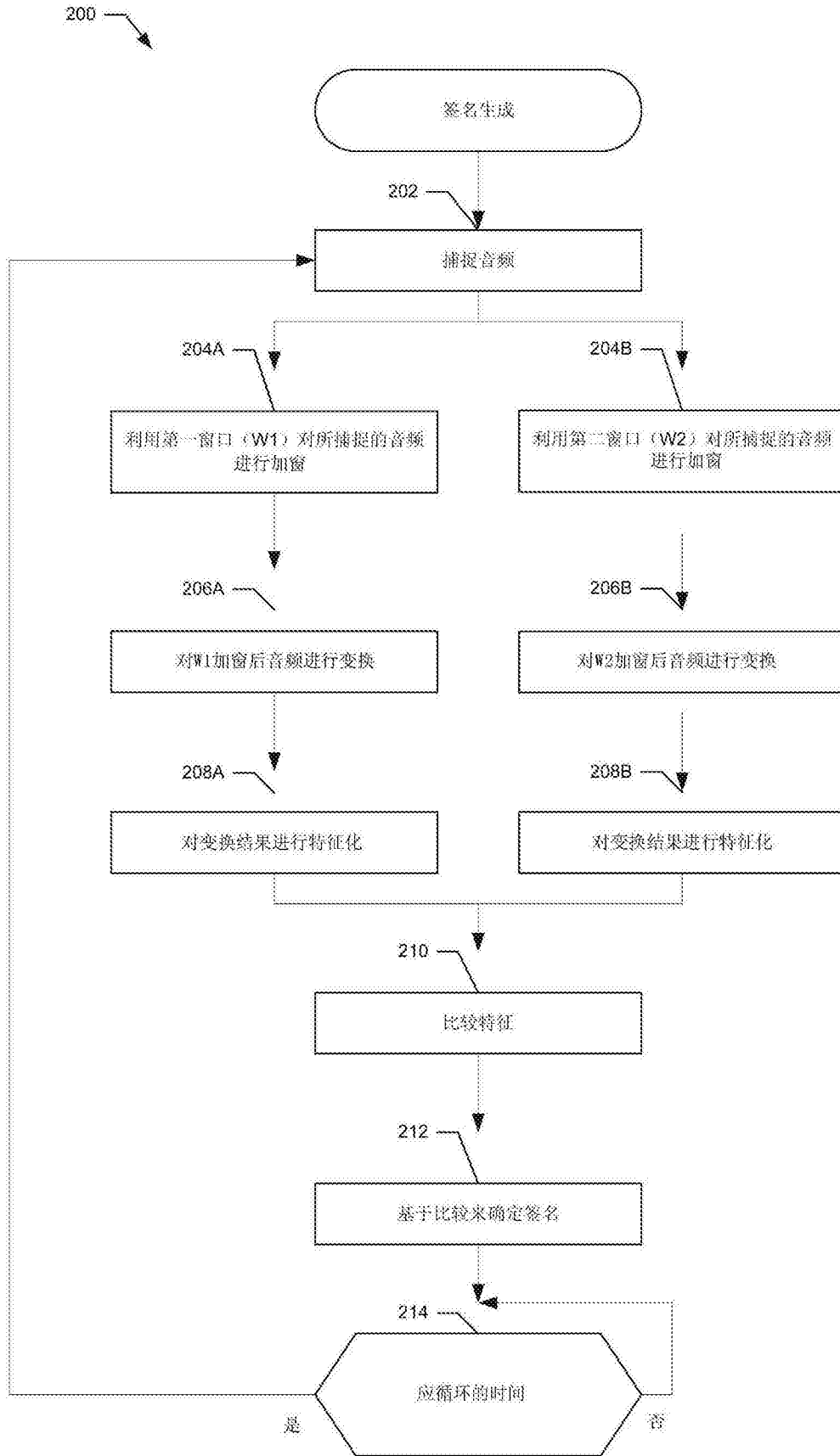


图 2

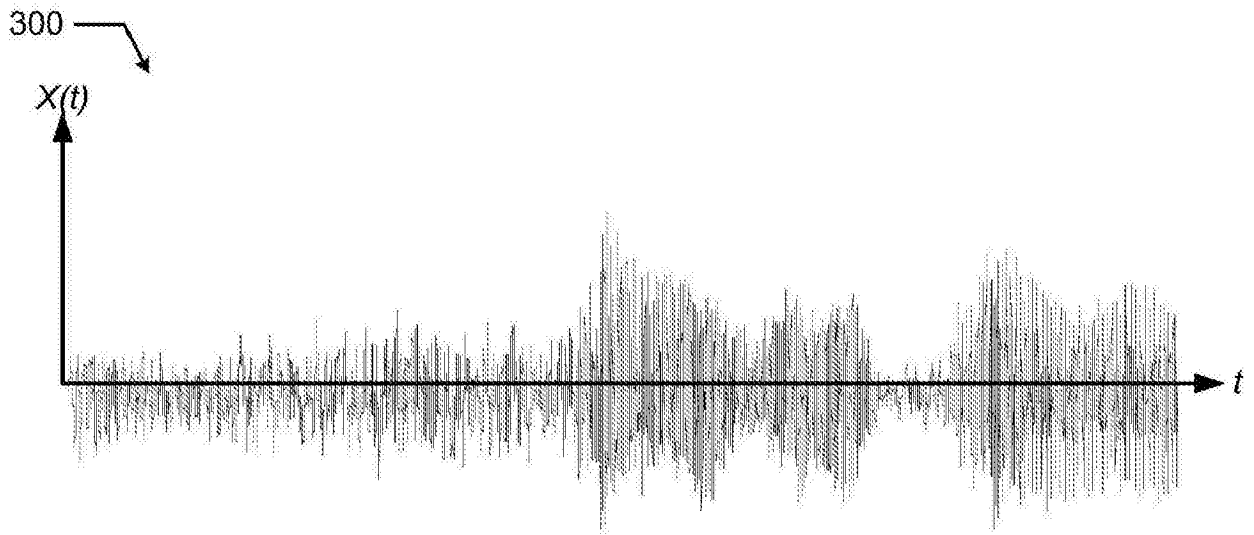


图 3

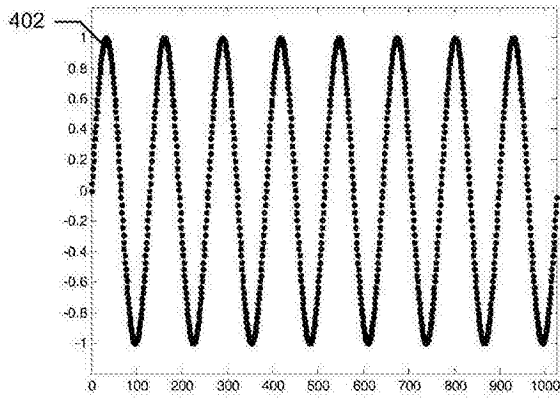


图 4

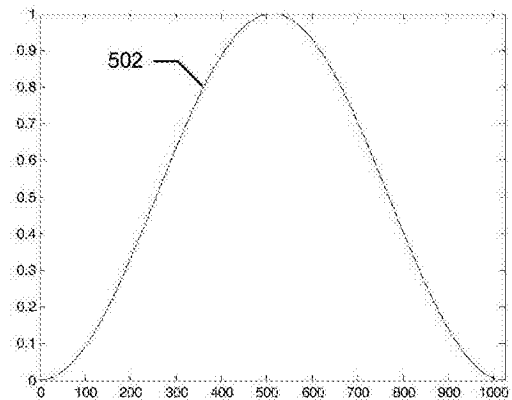


图 5

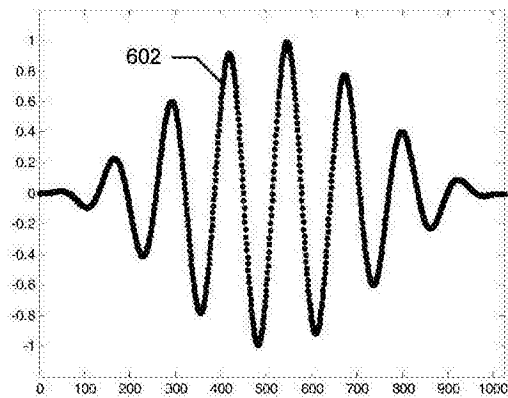


图 6

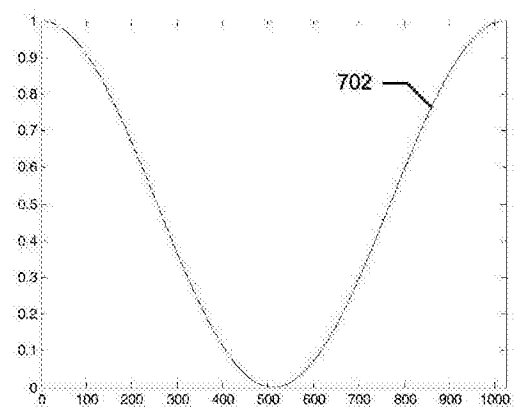


图 7

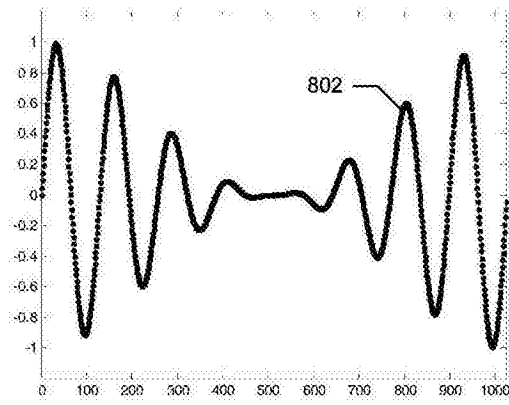


图 8

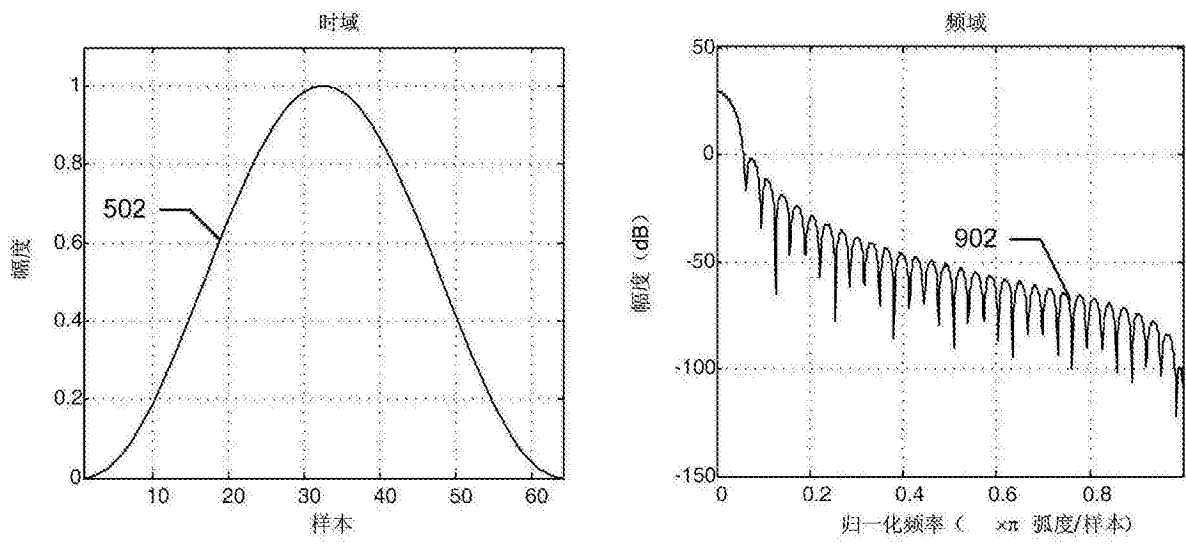


图 9

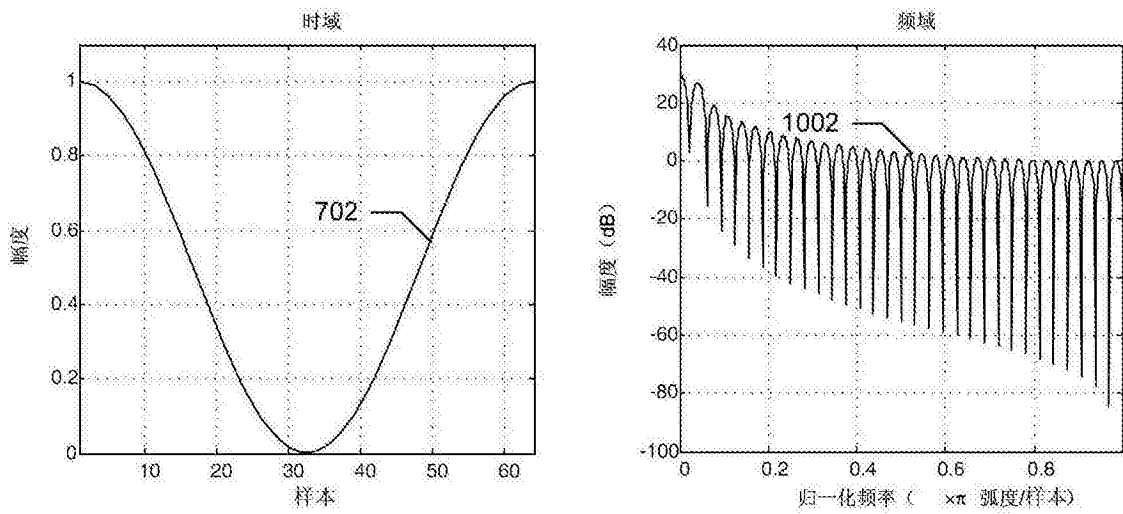


图 10

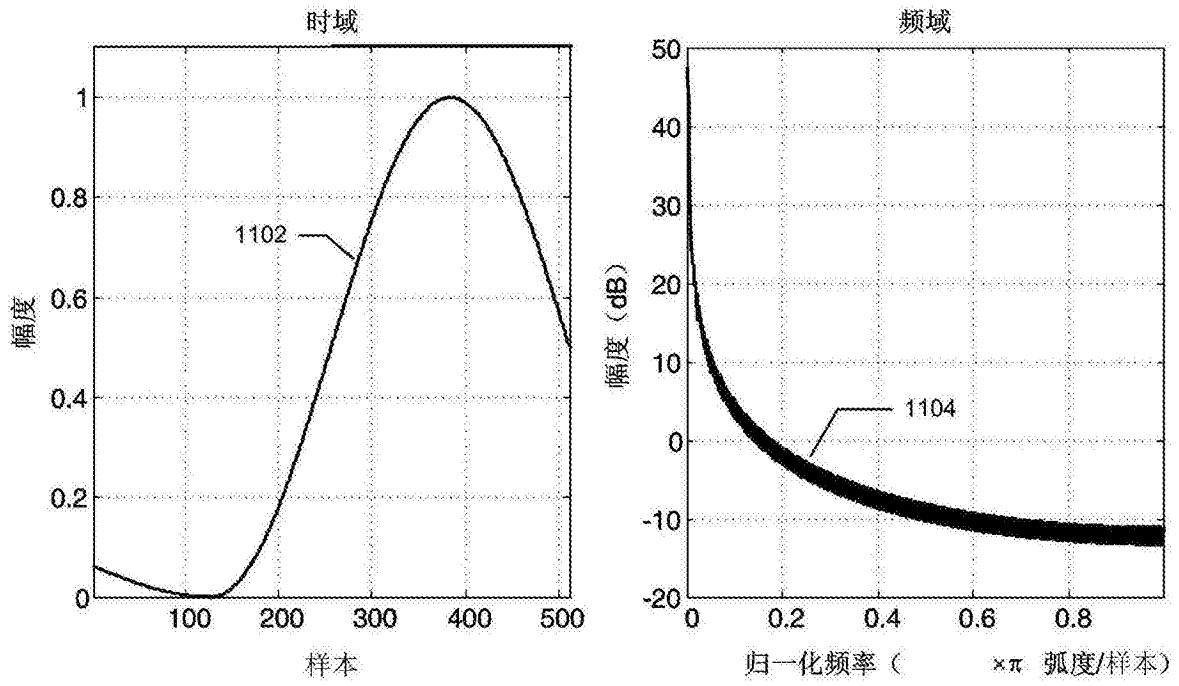


图 11

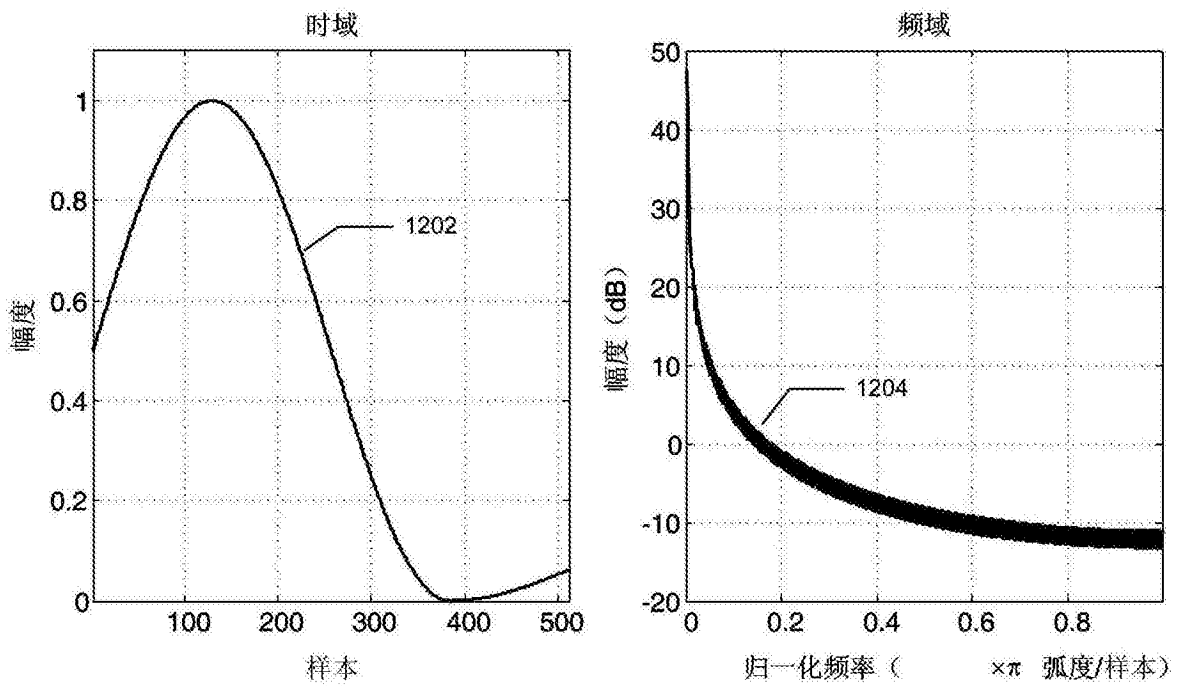


图 12

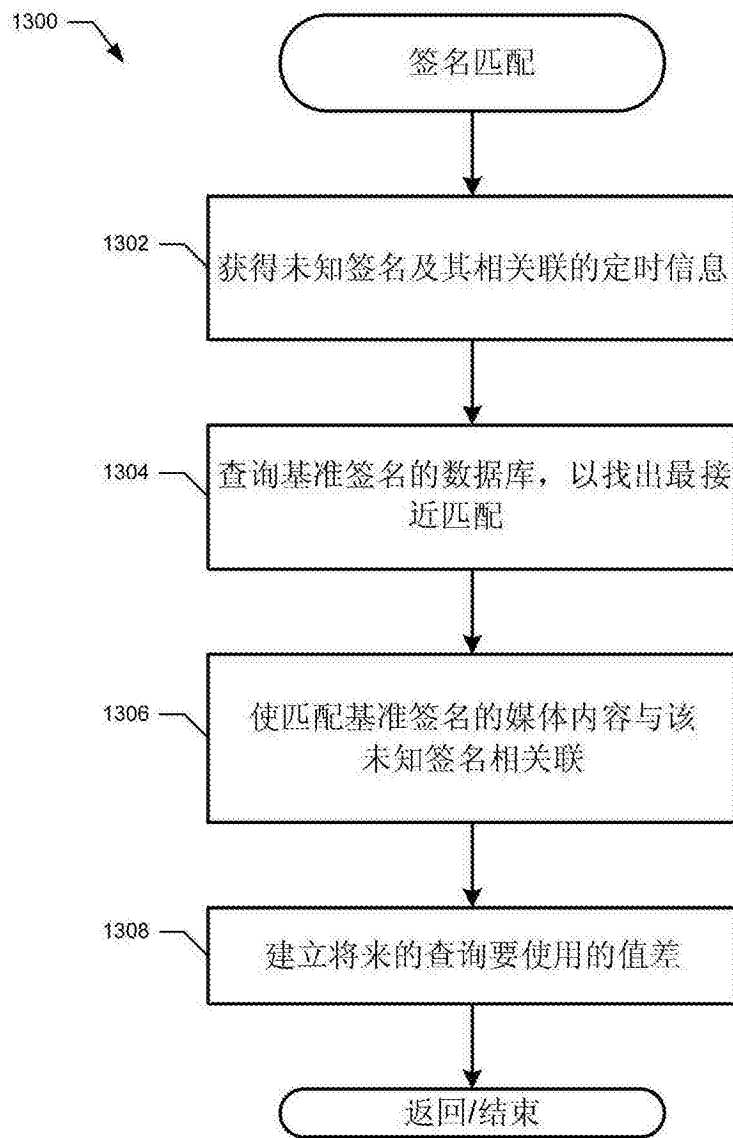


图 13

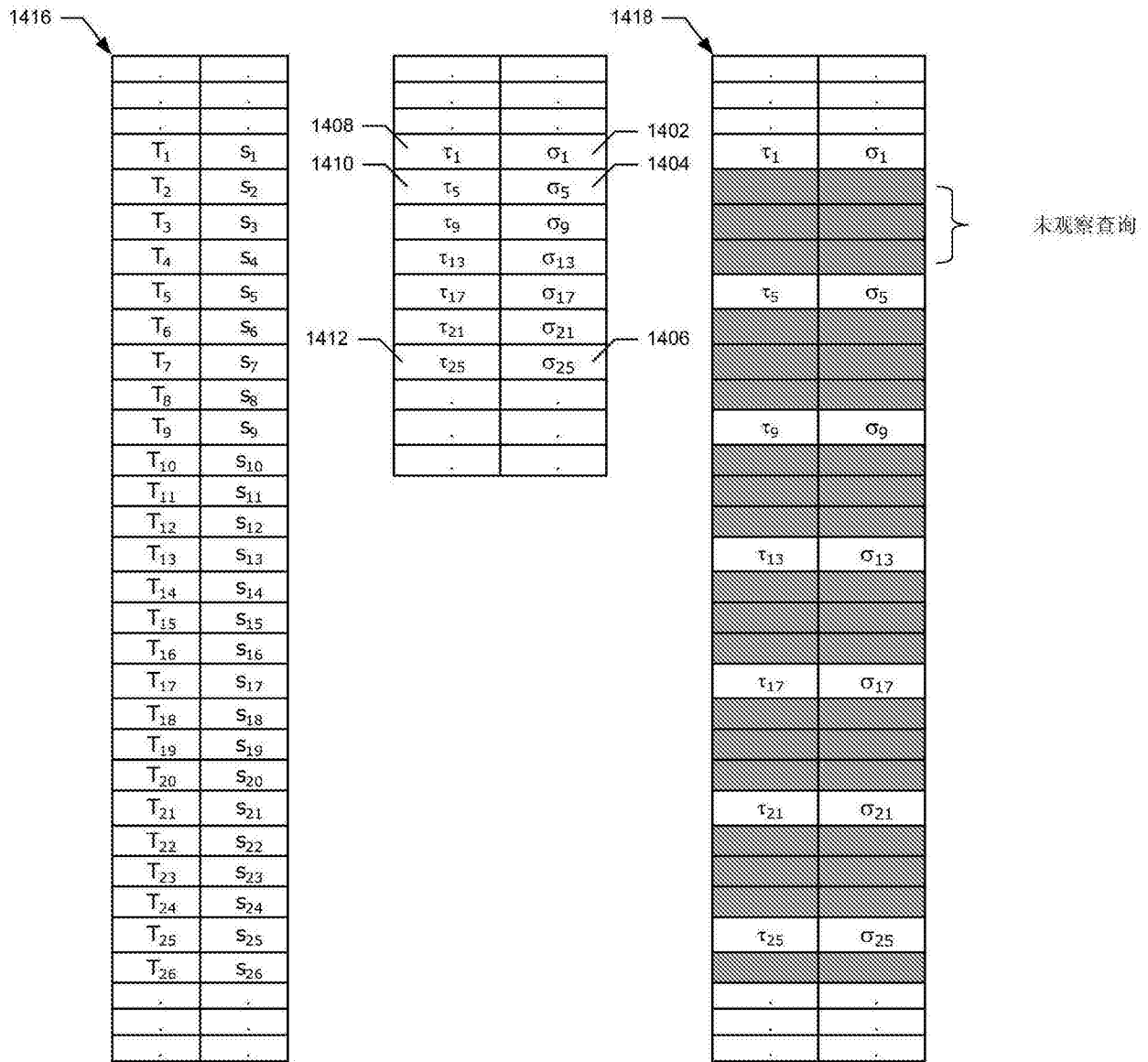


图 14

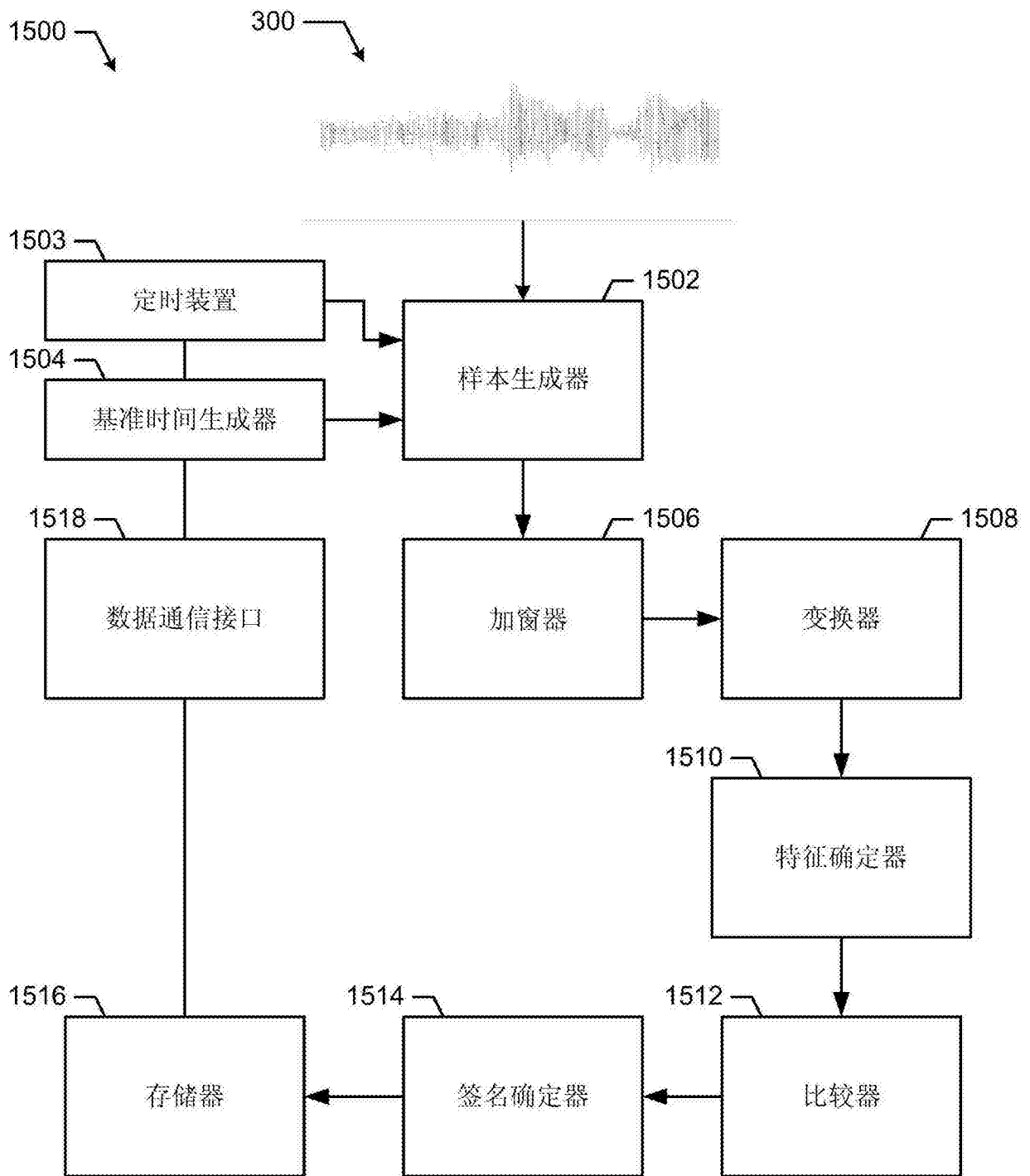


图 15

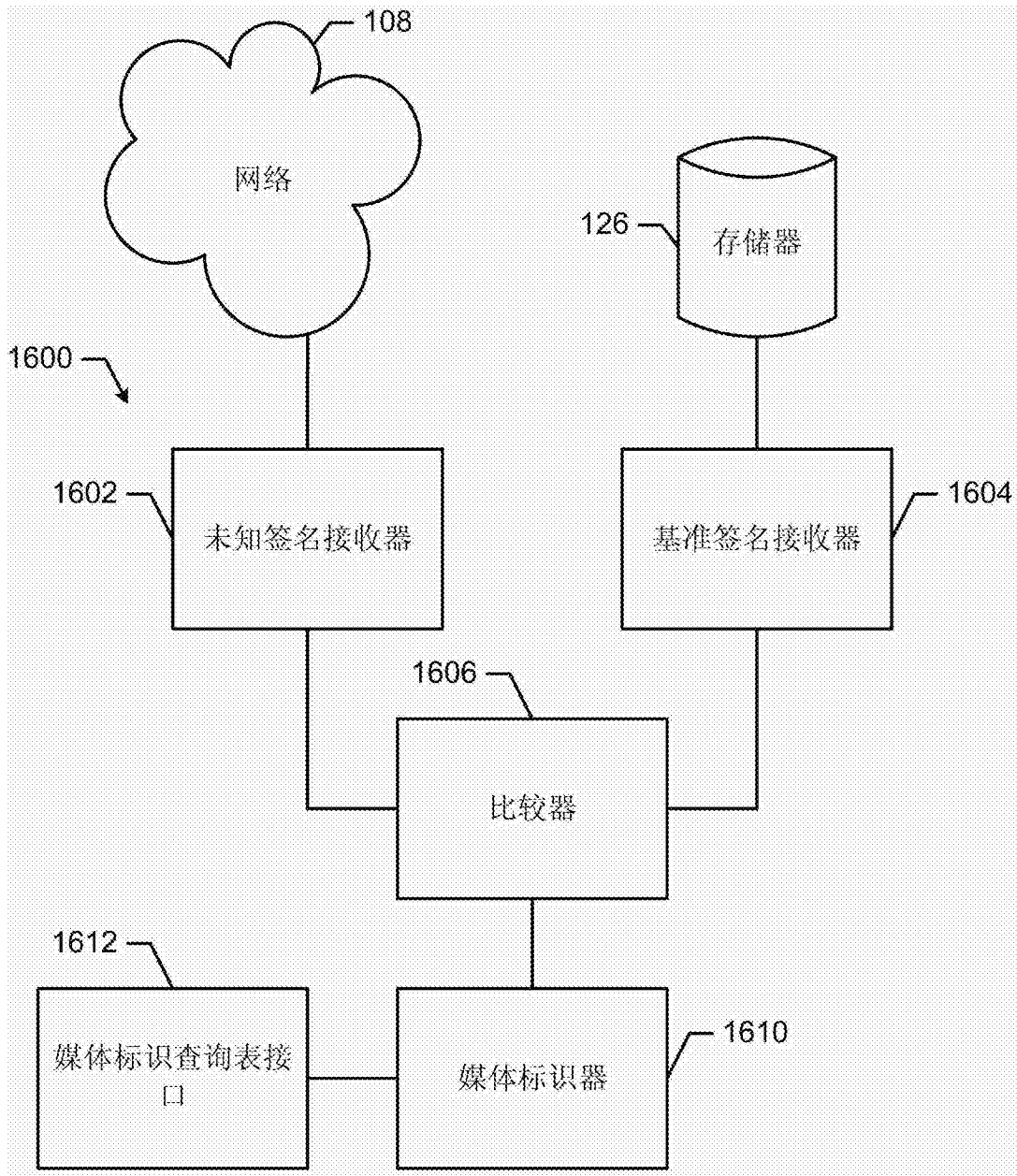


图 16



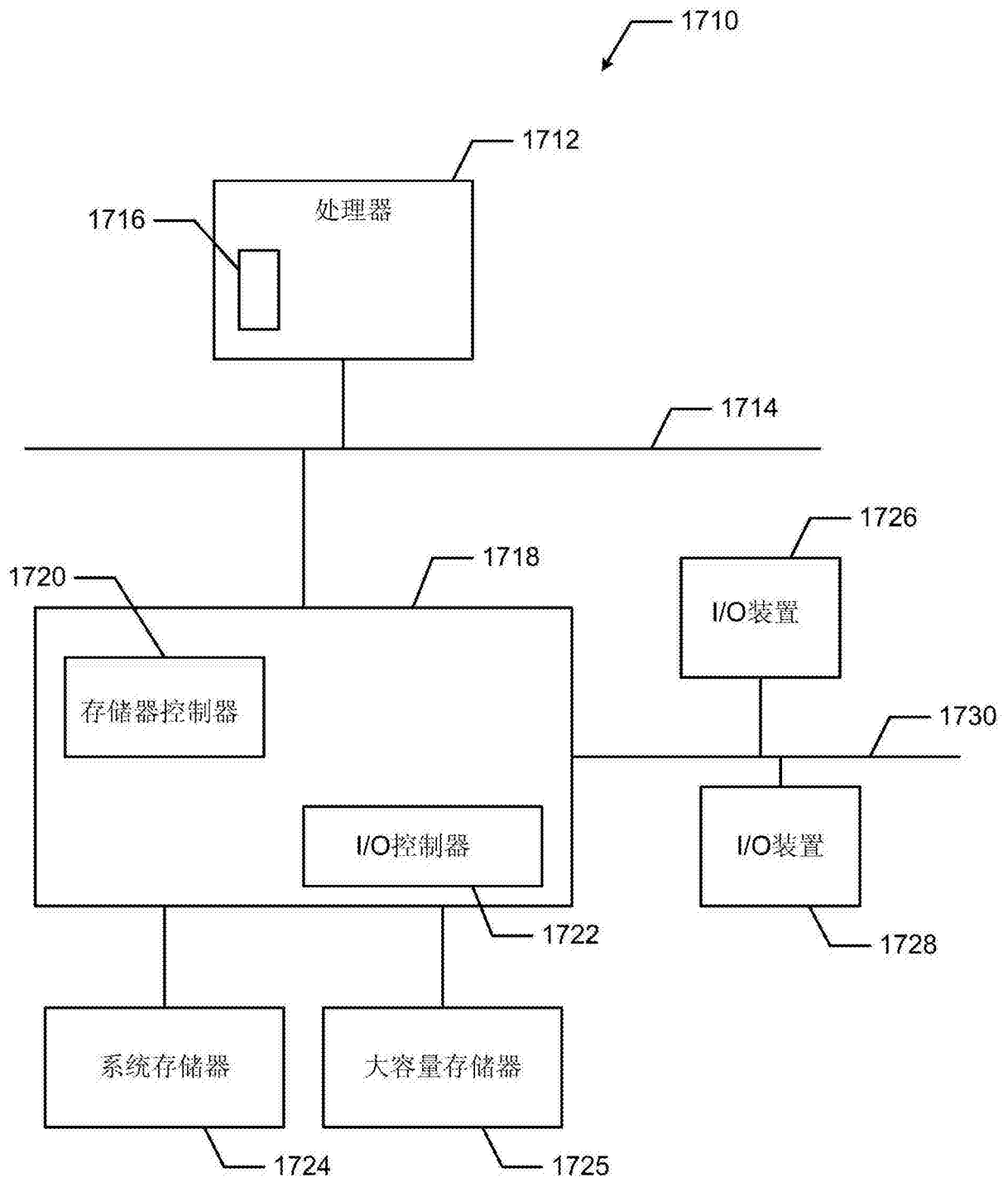


图 17