

(12) 发明专利

(10) 授权公告号 CN 102057423 B

(45) 授权公告日 2013. 04. 03

(21) 申请号 200980121577. 5

US 6614370 B2, 2003. 09. 02,

(22) 申请日 2009. 06. 09

CN 1679082 A, 2005. 10. 05,

(30) 优先权数据

61/060, 342 2008. 06. 10 US

Colin Perkins et al. A Survey of Packet loss Recovery Techniques for Streaming Audio. 《IEEE Network》. 1998, 第 12 卷 (第 5 期),

(85) PCT 申请进入国家阶段日

2010. 12. 10

审查员 王馨宁

(86) PCT 申请的申请数据

PCT/US2009/046692 2009. 06. 09

(87) PCT 申请的公布数据

W02009/152124 EN 2009. 12. 17

(73) 专利权人 杜比实验室特许公司

地址 美国加利福尼亚

(72) 发明人 H·米施

(74) 专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 杨小明

(51) Int. Cl.

G10L 19/005 (2013. 01)

(56) 对比文件

US 6421802 B1, 2002. 07. 16,

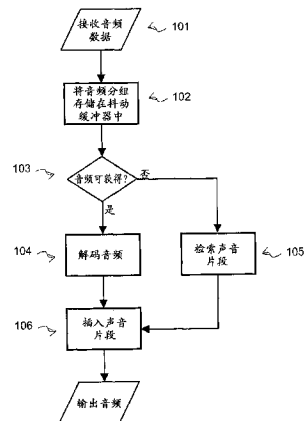
权利要求书 3 页 说明书 11 页 附图 5 页

(54) 发明名称

用于隐藏音频伪迹的方法、系统、计算机系统用途

(57) 摘要

识别音频信号中的至少一个段。所述音频段与音频信号内的伪迹相关联,并具有持续时间。检索至少一个存储的具有超过与音频段相关联的持续时间的持续时间的声音片段。将检索的声音片段与所述音频信号混合,并且检索的声音片段可听地补偿音频伪迹。



1. 一种用于隐藏音频伪迹的方法,包括以下步骤:

识别音频信号中的至少一个音频段,所述至少一个音频段具有与其相应的持续时间,其中,所述音频段与所述音频信号内的伪迹相关联;

检索至少一个存储的具有等于或者超过与所述至少一个音频段相关联的持续时间的持续时间的声音片段;

其中,所述检索步骤包括以下步骤:

检测识别的至少一个音频段中的音频伪迹;

基于所述音频伪迹的特性查询存储的声音片段的储存库;和

基于所述声音片段和所述特性之间的匹配响应于查询步骤返回所述声音片段,

其中,所述特性包括:

与识别的至少一个音频段相应的持续时间;和

与音频伪迹相应的至少一个音频属性;

以及,

将检索的至少一个声音片段与所述音频信号混合;

其中,将所述至少一个检索的声音片段与所述音频信号混合使得所述音频伪迹不能被感觉到。

2. 根据权利要求1所述的方法,其中,所述音频伪迹包括所述音频段的丢失或损坏部分中的一个或多个;和

其中,所述方法还包括以下步骤:

接收音频流,其中,所述音频流包括编码的音频数据的多个分组;和

从接收的音频分组收集音频信号。

3. 根据权利要求1所述的方法,其中,当检测到音频伪迹时,所述检索步骤还包括以下步骤:

确定所述音频伪迹的特性;和

其中,响应于检测步骤或确定步骤中的至少一个执行查询步骤。

4. 根据权利要求3所述的方法,其中,所述音频伪迹的特性是频率相关的;

其中,所述确定步骤包括以下步骤:

预测对应于与频率相关的特性的频谱;和

其中,所述查询步骤包括以下步骤:

将预测的频谱与和存储的声音片段相关联的频谱特性进行比较;

其中,所述匹配包括预测的音频伪迹频谱和声音片段频谱特性之间的显著相似性。

5. 根据权利要求4所述的方法,还包括以下步骤:

至少部分地基于预测的频谱与和存储的声音片段相关联的频谱特性的比较,确定与存储的声音片段相关联的级别;和

调整存储的声音片段级别;

其中,所述混合步骤包括以下步骤:

将经级别调整的声音片段与所述音频段混合;

其中,经过混合步骤,经级别调整的声音片段显著地掩蔽所述音频伪迹。

6. 根据权利要求5所述的方法,其中,经过混合步骤,经级别调整的声音片段基本上掩

蔽所述音频伪迹。

7. 根据权利要求 1 所述的方法,还包括以下步骤:

将所述至少一个声音片段存储在声音片段储存库中;

监视与存储的声音片段相关的上下文信息;

其中,存储步骤包括基于上下文信息更新存储的声音片段中的一个或多个的步骤。

8. 根据权利要求 7 所述的方法,其中,所述音频信号与基于网络的游戏相关;和

其中,所述上下文信息与和游戏相关联的虚拟环境相关。

9. 根据权利要求 1 所述的方法,其中,所述方法是由计算机系统执行的。

10. 一种用于隐藏音频伪迹的方法,包括以下步骤:

识别音频信号中的至少一个音频段,所述至少一个音频段具有与其相应的持续时间,其中,所述音频段与所述音频信号内的伪迹相关联;

检索至少一个存储的具有等于或者超过与所述至少一个音频段相关联的持续时间的持续时间的声音片段;

其中,所述检索步骤包括以下步骤:

检测识别的至少一个音频段中的音频伪迹;

基于所述音频伪迹的特性查询存储的声音片段的储存库;和

基于所述声音片段和所述特性之间的匹配响应于查询步骤返回所述声音片段,

其中,所述特性包括:

与识别的至少一个音频段相应的持续时间;和

与音频伪迹相应的至少一个音频属性;

以及,

将检索的至少一个声音片段与所述音频信号混合;

其中,将所述至少一个检索的声音片段与所述音频信号混合使得所述音频伪迹不能被感觉到;

其中,所述音频伪迹包括所述音频段的丢失或损坏部分中的一个或多个。

11. 根据权利要求 10 所述的方法,还包括以下步骤:

接收音频流,其中,所述音频流包括编码的音频数据的多个分组;和

从接收的音频分组收集音频信号;

其中,与丢失的或损坏的音频段相关联的时间位置被完全包含在所述音频片段的时间位置中。

12. 一种用于隐藏音频伪迹的系统,包括:

用于识别音频信号中的至少一个音频段的装置,所述至少一个音频段具有与其相应的持续时间,其中,所述音频段与所述音频信号内的伪迹相关联;

用于检索至少一个存储的具有等于或者超过与所述至少一个音频段相关联的持续时间的持续时间的声音片段的装置;

其中,所述用于检索的装置包括以下装置:

用于检测识别的至少一个音频段中的音频伪迹的装置;

用于基于所述音频伪迹的特性查询存储的声音片段的储存库的装置;和

用于基于所述声音片段和所述特性之间的匹配响应于用于查询的装置返回所述声音

片段的装置，

其中，所述特性包括：

与识别的至少一个音频段相应的持续时间；和

与音频伪迹相应的至少一个音频属性；

以及，

用于将检索的至少一个声音片段与所述音频信号混合的装置；

其中，将所述至少一个检索的声音片段与所述音频信号混合使得所述音频伪迹不能被感觉到。

用于隐藏音频伪迹的方法、系统、计算机系统用途

[0001] 本申请要求由 Hannes Muesch 于 2008 年 6 月 10 日提交的相关的共同未决的题目为“Concealing Audio Artifacts”的美国临时专利申请 No. 61/060,342 的权益,包括优先权,该临时专利申请已转让给本申请(杜比实验室文号 D07046US01)的受让人。

[0002] 技术领域

[0003] 本发明涉及音频信号处理。更具体地讲,本发明的实施例涉及隐藏音频伪迹(artifact)。

[0004] 背景技术

[0005] 现代音频通信可涉及音频信息通过分组(packet)交换网络(例如因特网)的传输。通过分组交换网络的音频通信可以是电话、在线计算机游戏、视频和远程会议(teleconferencing)及其它应用的特征。

[0006] 例如,多玩家在线计算机游戏可涉及各个游戏玩家之间的现场(live)语音通信。在该环境中,语音通信路径可包括语音编码器,语音编码器的输出被分包(packetize),并通过分组交换网络被转发(relay)给其他游戏玩家。

[0007] 在这个部分中所描述的应用、情况或问题可被追踪(persue),但不必然是以前已构想或追踪的。除非另外指明,不应该仅由于在这个部分中所描述的任何方案被包括在这里,就将这些方法假设为现有技术。类似地,除非另外指明,不应该基于这个部分就假设针对一个或多个应用或情况而发现的问题已在任何现有技术中被认识到。

[0008] 附图说明

[0009] 在附图的图中以示例方式,而不是以限制方式,对本发明进行说明,在附图中,相同的标号表示类似的元件,其中:

[0010] 图 1 描绘根据本发明的实施例的第一示例性过程的流程图;

[0011] 图 2 描绘根据本发明的实施例的第二示例性过程的流程图;

[0012] 图 3 描绘根据本发明的实施例的第三示例性过程的流程图;

[0013] 图 4 描绘通过其可实现本发明的实施例的示例性计算机系统平台;和

[0014] 图 5 描绘通过其可实现本发明的实施例的示例性集成电路装置平台。

具体实施方式

[0015] 这里描述涉及隐藏音频伪迹的示例性实施例。在以下描述中,为了解释的目的,对许多特定细节进行了阐述,以提供对本发明的透彻理解。然而,应该清楚,可在没有这些特定细节的情况下实施本发明。在其它实例中,为了避免封闭、模糊或混淆本发明,没有对公知的结构和装置进行详尽描述。

[0016] 概述

[0017] 本发明的实施例涉及隐藏音频伪迹。在音频信号中识别至少一个段。音频段与音频信号内的伪迹相关联,并具有持续时间。检索至少一个存储的具有持续时间的声音片段,所述持续时间与和所述音频段相关联的持续时间匹配或者超过和所述音频段相关联的持续时间。将检索的声音片段与所述音频信号混合,并且检索的声音片段可听地(audibly)

补偿音频伪迹。

[0018] 本发明的实施例利用已知为连续错觉 (continuity illusion) 或时间诱导 (temporal induction) 的心理现象。为了便于理解本发明的实施例,现在对这种现象进行解释:

[0019] 如这里所使用的,术语连续错觉和时间诱导与听觉错觉有关,在听觉错觉中,如果第二声音防止收听者获得第一声音中的中断出现的迹象 (evidence),则收听者感知中断的第一声音是连续的。

[0020] 例如,如果连续音调被一系列间隙 (gap) 周期性地中断,则收听者将停止听到连续音调,并且作为替代将感知到一系列脉动离散音调 (pulsating discrete tone)。如果引入第二声音,例如引入在音调中断时间期间出现的一系列噪声短脉冲串 (noise burst),并且如果噪声的频谱和水平使得如果音调没有被中断则该噪声将会掩蔽 (mask) 音调,那么收听者将停止听到音调为中断的。替代地,收听者将感知到伴随一系列噪声短脉冲串的未中断 (比如,连续) 的音调。第二声音 (噪声短脉冲串) 的添加创建了第一声音 (被中断的音调) 连续的错觉。

[0021] 从此,第一声音将被称为“目标声音”,第二声音将被称为“掩蔽物 (masker)”或“掩蔽声音 (masking sound)”。

[0022] 对于连续错觉的发生,必须满足两个条件:第一,收听者必须具有目标信号连续的合理预期。连续的预期从上下文得到:例如,听到句子的开头的短语,收听者预期还听到该句子的最后的单词。第二,掩蔽物必须防止收听者获得目标声音中断的任何迹象。当掩蔽物的听觉表示与收听者预期在中断时间段期间听到的目标声音的听觉表示完全重叠时,掩蔽声音防止收听者获得中断的迹象。关于听觉表示的时间位置和幅度,重叠必须完全。

[0023] 合适的听觉表示的示例为听觉神经中的基膜的激励和放电模式 (firing pattern),或者它们的数学模型。

[0024] 可用简易信号 (例如音调) 和复合信号 (例如音乐或讲话) 引起连续错觉。将适当放置的掩蔽声音添加到中断的讲话信号不仅给予连续的错觉,而且未中断的讲话还使得大脑中的语言中心能够使用上下文信息来“填充”丢失的讲话段,从而帮助讲话理解。

[0025] 本发明的实施例用于通过添加战略 (strategically) 放置的掩蔽声音引起连续错觉来隐藏从错误音频传输引起的简短的音频伪迹。所描述的实施例提供用于选择或者产生对于引起连续错觉有效的且适合于收听环境的掩蔽信号的方法。

[0026] 示例性过程

[0027] 图 1 描绘根据本发明的实施例的第一示例性过程 100 的流程图。在步骤 101 中,接收音频信号中的数据分组 (比如,通过音频接收器)。音频信号可包括一系列音频数据分组。在步骤 102 中,对接收的音频数据分组进行缓冲 (比如,临时存储在与音频接收器相关联的抖动 (jitter) 缓冲器中)。与接收音频数据分组的音频接收器相关联的音频解码器可达到或者呈现这样的状态,即,解码器准备接收构成音频信号的分组系列中的下一个音频分组以用于顺序解码。

[0028] 在步骤 103 中,关于缓冲的音频分组,查询抖动缓冲器。如果可在抖动缓冲器中或者从抖动缓冲器获得音频分组,则在步骤 104 中,将缓冲的音频分组传递给解码器。然而,如果请求的音频分组不可获得,则解码器或者产生丢失音频信号的预测,或者将具有与丢

失分组的持续时间相应的持续时间的间隙插入到解码的音频流中。

[0029] 如这里所使用的,术语“掩蔽”可涉及通过提供其听觉表示与正被掩蔽的音频信号的听觉表示完全重叠的“掩蔽声音”或者“掩蔽物”来使得音频信号不能被听见。与其它音频信息相似,可对掩蔽声音进行分类、编纂 (codify)、建索引、存储、从储存器检索、以及/或者呈现 (render)。可在存储介质存储和从存储介质检索掩蔽声音,所述存储介质包括,但不限于,计算机存储器、存储盘或静态驱动器、或者音频储存库或数据库。

[0030] 在步骤 105 中,从存储介质检索用作与间隙(或预测的信号部分)相关的掩蔽声音的声音片段。在步骤 106 中,与音频信号中的间隙(或失真)基本上时间对应地将检索的掩蔽声音片段混合(比如,插入)到解码的音频信号中。

[0031] 在本论述的上下文中,“掩蔽间隙”的概念可表示提供掩蔽声音,该掩蔽声音为收听者将会在间隙出现的时间合理地预期听到的信号的有效掩蔽物。

[0032] 实施例提供涉及连续错觉的功能,在该功能中,掩蔽声音基本上(比如,完全)掩蔽与丢失或损坏的信号部分明显地类似(比如,相同、基本相同、紧密相似)的声音。因而,实施例的功能是使掩蔽物的水平及其频谱特性与掩蔽间隙或预测的信号部分所需的水平和频谱特性匹配。

[0033] 例如,实施例的功能是调整掩蔽物的水平,以使得在接收的音频信号的剩余部分的上下文中掩蔽物水平足以掩蔽间隙或缺陷。还例如,实施例的功能是调整掩蔽物的频率成分,以使得在接收的音频信号的剩余部分的上下文中该频率成分适合于掩蔽间隙或缺陷。过程 100 可通过相对高水平的、宽带掩蔽声音工作,所述掩蔽声音可足以掩蔽可能接收或者遇到的音频信号中的预期持续时间的间隙或者预期失真。

[0034] 图 2 描绘根据本发明的实施例的第二示例性过程 200 的流程图。在实施例中,过程 200 用过程 100 的一个或多个步骤或者步骤序列执行(图 1)。因此,过程 200 可从步骤 101 开始,在步骤 101 中,接收音频数据分组。在步骤 102 中,将接收的音频分组(比如,临时)存储在抖动缓冲器中。当音频解码器处于要接收音频流中的后面的(比如,下一个)音频分组以用于解码的状况时,在步骤 103 中对抖动缓冲器进行查询。如果存储的音频分组可获得,则在步骤 104 中,将该分组传递给解码器。然而,如果请求的音频分组不可获得,则解码器将间隙或丢失音频的预测插入到解码的音频中。

[0035] 当将间隙或预测音频插入到解码的音频中时,在步骤 202 中从储存器检索第一掩蔽声音。在步骤 203 中,计算与第一掩蔽声音相应的听觉表示(比如,听觉掩蔽模式)。

[0036] 在步骤 201 中预测丢失(或损坏)音频数据的特性。例如,可通过重复丢失部分之前的音频部分来得到丢失音频数据的一种或多种特性。

[0037] 在步骤 204 中,计算通过预测信号生成的听觉表示(比如,激励模式)。在步骤 205 中,将计算的预测信号的听觉表示与第一检索的掩蔽物的听觉表示进行比较。如果比较表明掩蔽物不完全掩蔽预测音频信号,则在步骤 206 中将小的固定增益施加于掩蔽物,并重复掩蔽计算。这个迭代过程可持续到掩蔽物本质上完全掩蔽预测音频信号为止。

[0038] 预测音频信号和掩蔽物的频谱之间的显著不匹配可能要求增益增大以掩蔽预测音频信号。所要求的增益水平可变得比所期望的大,比如,为了似真性 (plausibility) 或者舒适。实施例可选择至少一个可供选择的掩蔽声音,并通过该可供选择的掩蔽声音重复掩蔽预测。可选地,在步骤 207 中,作为选择,可与可供选择的掩蔽预测相关地选择增益。

[0039] 在步骤 208 中根据决定规则选择掩蔽物候选之一。实施例可至少部分地基于一个或多个标准选择掩蔽物。例如,与步骤 208 相关的决定函数可从多个候选掩蔽物之中选择要求最小增益的掩蔽物。在步骤 106 中,将所选择的掩蔽声音插入到音频流中以掩蔽间隙或缺陷。

[0040] 时间诱导在广范围的收听情况下作用。然而,时间诱导作为隐藏音频信号中的遗漏(dropout)的手段并不总是实用的。例如,将噪声短脉冲串插入到电话会话中以引起连续错觉可能造成比对隐藏遗漏不采取任何措施更差的用户体验。时间诱导仅在用于诱导连续错觉的掩蔽者适合于其的应用中实用。

[0041] 例如,实施例可用于带现场聊天的在线游戏应用。在带现场聊天的在线游戏中,用户接收从两组源发出的音频。第一组音频源包括通过分组交换数据网络实时接收的编码的语音信号。通过分组交换网络实时发送的音频源可能遭受语音信号中的丢失数据分组和伴随的(比如,伴发的)遗漏。

[0042] 第二组音频源包括由游戏引擎创建的多种环境声音(以及与用户和游戏引擎部署或定位于其中的物理环境相关联的可能的周围环境噪声或其它声音)。典型的游戏声音场景包括几种声音的叠加,多个这些声音(可能许多)具有短的持续时间。示例包括雷鸣声、枪击、爆炸等。

[0043] 周围环境声音通常可存储在物理上靠近用户的位置中,例如,存储在用户本地的数据存储装置处。因此,可至少部分地基于游戏运动的进展,可能显著地基于游戏运动的进展动态地发起对本地存储的声音的回放。在一些实例中,播放环境声音的定时可极大不同,而对声音场景的似真性没有显著的负面影响。在这样的应用中,具有提供遗漏隐藏的时间诱导功能的实施例是有用且实用的。

[0044] 图 3 描绘根据本发明的实施例的第三示例性过程 300 的流程图。过程 300 可用于例如游戏引擎的应用和/或与例如游戏引擎的应用集成。在步骤 301 中,决定是否发生了听觉场景的变化。如果发生了场景变化,则在步骤 302 中,识别场景相关的音频资源(比如,所有可访问的音频资源)。

[0045] 并不是所有的场景相关的音频资源可适合于遗漏隐藏。例如,过长的、具有不合适地窄的频率范围的、或者如果以掩蔽典型的讲话信号所需的级别播放则将会不象真实(implausible)的音频资源可能不足以满足实际的遗漏隐藏。因此,在步骤 303 中,从场景相关的音频资源之中选择适合于遗漏隐藏的音频资源子集。在步骤 304 中,使得所选择的音频资源的子集可用于(比如,提供给)根据过程 100 和/或 200(图 1、图 2)的遗漏隐藏。

[0046] 示例性计算机系统实现平台

[0047] 可用计算机系统、用电子电路和组件配置的系统、例如微控制器的集成电路(IC)器件、现场可编程门阵列(FPGA)或者应用特定 IC(ASIC)和/或包括这样的系统、装置或组件中的一个或多个的设备来实现本发明的实施例,例如过程 100、200 和 300(图 1、图 2、图 3)的一部分。

[0048] 图 4 描绘通过其可实现本发明的实施例的示例性计算机系统平台 400。计算机系统 400 包括用于传送信息的总线 402 或其它通信机制、以及与总线 402 耦接的用于处理信息的处理器 404。计算机系统 400 还包括与总线 402 耦接的用于存储信息和将由处理器 404 执行的指令的主存储器 406,例如随机存取存储器(RAM)或其它动态存储装置。主存储器

406 还用于在将由处理器 404 执行的指令的执行期间存储临时变量或其它中间信息。

[0049] 计算机系统 400 还包括与总线 402 耦接的用于存储静态信息和用于处理器 404 的指令的只读存储器 (ROM) 408 或其它静态存储装置。提供例如磁盘或光盘的存储装置 410, 存储装置 410 与总线 402 耦接, 用于存储信息和指令。处理器 404 可执行一种或多种数字信号处理 (DSP) 功能。另外或者作为选择, DSP 功能可由另一个处理器或实体 (这里用处理器 404 表示) 来执行。

[0050] 计算机系统 400 通过总线 402 与用于向计算机用户显示信息的显示器 412 耦接, 显示器 412 例如液晶显示器 (LCD)、阴极射线管 (CRT) 等。包括字母数字键和其它键的输入装置 414 与总线 402 耦接, 用于将信息和命令选择传送给处理器 404。另一种类型的用户输入装置为光标控制器 416, 例如, 用于将方向信息和命令选择传送给处理器 404 和用于控制显示器 412 上的光标移动的鼠标、跟踪球或者光标方向键。这种输入装置通常具有允许装置指定平面中的位置的两个轴 (第一轴 (比如, x) 和第二轴 (比如, y)) 的两个自由度。

[0051] 本发明涉及使用计算机系统 400 隐藏音频伪迹。根据本发明的一个实施例, 计算机系统 400 响应于处理器 404 执行主存储器 406 中所包含的一个或多个指令的一个或多个序列来提供音频伪迹的隐藏。可将这样的指令从另一个计算机可读介质 (例如存储装置 410) 读取到主存储器 406 中。主存储器 406 中所包含的指令序列的执行使处理器 404 执行这里所述的处理步骤。多处理布置中的一个或多个处理器也可用于执行主存储器 406 中所包含的指令序列。在可供选择的实施例中, 可使用硬接线电路来代替实现本发明的软件指令, 或者可与实现本发明的软件指令组合地使用硬接线电路。因此, 本发明的实施例不限于硬件电路和软件的任何特定组合。

[0052] 这里所使用的术语“计算机可读介质”可表示参与将指令提供给处理器 404 以用于执行的任何介质。这样的介质可采取许多形式, 包括但不限于, 非易失性介质、易失性介质和传输介质。非易失性介质包括, 例如, 光盘或磁盘, 例如存储装置 410。易失性介质包括动态存储器, 例如主存储器 406。传输介质包括同轴电缆、铜线和其它导体、以及光纤, 包括包含总线 402 的配线。传输介质还可采取例如在无线电波和红外数据通信期间产生的那些形式的声波或电磁 (比如, 光) 波的形式。

[0053] 计算机可读介质的共同形式包括, 例如, 软盘、柔性盘、硬盘、磁带或任何其它磁介质、CD-ROM、任何其它光介质、打孔卡、纸带、具有孔图案的任何其它旧式的或其它的物理介质、RAM、PROM 和 EPROM、FLASH-EPROM、任何其它存储芯片或盒、如下所述的载波、或者计算机可从其读取的任何其它介质。

[0054] 各种形式的计算机可读介质可涉及将一个或多个指令的一个或多个序列承载到处理器 404 以用于执行。例如, 指令一开始可承载在远程计算机的磁盘上。远程计算机可将指令加载到其动态存储器中, 并使用调制解调器通过电话线发送所述指令。计算机系统 400 本地的调制解调器可接收电话线上的数据, 并使用红外发送器将数据转换为红外信号。与总线 402 耦接的红外检测器可接收红外信号中承载的数据, 并将该数据放置在总线 402 上。总线 402 将数据承载到主存储器 406, 处理器 404 从主存储器 406 检索指令并执行指令。可选地, 可在处理器 404 执行之前或者之后将主存储器 406 接收的指令存储在存储装置 410 上。

[0055] 计算机系统 400 还包括与总线 402 耦接的通信接口 418。通信接口 418 提供与网

络链路 420 耦接的双路数据通信,网络链路 420 与本地网络 422 连接。例如,通信接口 418 可以是综合服务数字网 (ISDN) 卡或数字用户线 (DSL)、将数据通信连接提供给相应类型的电话线的电缆或其它调制解调器。作为另一个示例,通信接口 418 可以是提供与可兼容局域网 (LAN) 的数据通信连接的 LAN 卡。还可实现无线链接。在任何这样的实现中,通信接口 418 发送和接收承载表示各种类型的信息的数字数据流的电、电磁或光学信号。

[0056] 网络链路 420 通常通过一个或多个网络提供与其它数据装置的数据通信。例如,网络链路 420 可通过局域网 422 提供与主机 424 或者与由因特网服务提供商 (ISP) 426 操作的数据设备的连接。ISP 426 进而通过全球分组数据通信网络提供数据通信服务,全球分组数据通信网络现在普遍称为“因特网”428。局域网 422 和因特网 428 都使用承载数字数据流的电、电磁或光学信号。将数字数据承载到计算机系统 400 或者从计算机系统 400 承载数字数据的、通过各种网络的信号以及网络链路 402 上且通过通信接口 418 的信号为运载信息的示例性形式的载波。

[0057] 计算机系统 400 可通过网络、网络链路 420 和通信接口 418 发送消息和接收包括程序代码的数据。在因特网示例中,服务器 430 可通过因特网 428、ISP 426、局域网 422 和通信接口 418 发送请求的应用代码。根据本发明,一个这样的下载应用提供如这里所述的实现可靠地与媒体内容一致的媒体指纹。

[0058] 当接收的代码被接收、和 / 或被存储在存储装置 410 或其它非易失性存储器中以用于以后执行时,所述代码可被处理器 404 执行。以这种方式,计算机系统 400 可获得载波形式的应用代码。

[0059] 示例性平台

[0060] 图 5 描绘通过其可实现本发明的实施例的示例性 IC 装置 500。IC 装置 500 可具有输入 / 输出 (I/O) 特征 501。I/O 特征 501 接收输入信号,并通过路由结构 (routing fabric) 510 将它们路由到与存储器 503 一起工作的中央处理单元 (CPU) 502。I/O 特征 501 还从 IC 装置 500 的其它组件特征接收输出信号,并可通过路由结构 510 控制信号流的一部分。数字信号处理 (DSP) 特征至少执行与数字信号处理相关的功能。接口 505 访问外部信号,并将它们路由到 I/O 特征 501,并允许 IC 装置 500 输出信号。路由结构 510 在 IC 装置 500 的各种组件特征之间路由信号和电力 (power)。

[0061] 可配置的和 / 或可编程的处理元件 (CPPE) 511,例如逻辑门阵列,可执行 IC 装置 500 的专用功能,在实施例中,这些专用功能可涉及提取和处理可靠地与媒体介质一致的媒体指纹。存储器 512 贡献足够的存储单元以供 CPPE 511 有效率地工作。CPPE 可包括一个或多个专用 DSP 特征 514。

[0062] 本发明的实施例涉及隐藏音频伪迹。在音频信号中识别至少一个段。音频段与音频信号内的伪迹相关联,并具有持续时间。检索至少一个存储的具有持续时间的声音片段,所述持续时间与和音频段相关联的持续时间匹配或者超过和音频段相关联的持续时间。将检索的声音片段与音频信号混合,并且检索的声音片段可听地补偿音频伪迹。音频伪迹可包括音频段的部分丢失或数据组成的损坏。可接收包括编码的音频数据的多个分组的音频流。从接收的音频分组收集 (assemble) 音频信号。

[0063] 可将声音片段存储在储存库中。检索声音片段可包括检测识别的至少一个音频段中的音频伪迹、基于音频伪迹的特性查询储存库、以及基于声音片段和伪迹特性之间的匹

配响应于所述查询返回声音片段。伪迹特性可包括与识别的段相应的持续时间和与音频伪迹相应的至少一个音频属性。

[0064] 当检测到音频伪迹时,检索声音片段可包括确定音频伪迹的特性,在确定音频伪迹的特性时,响应于检测伪迹或确定伪迹的特性来执行查询。音频伪迹的特性是频率相关的。因此,确定伪迹的特性可包括预测对应于与频率相关的特性的频谱。

[0065] 执行查询可包括将预测的频谱与和存储的声音片段相关联的频谱特性进行比较。因此,匹配可包括预测的音频伪迹频谱和声音片段频谱特性之间的显著相似性。显著相似性可包括预测的音频伪迹频谱和声音片段频谱特性之间的基本相同的对应性。

[0066] 至少部分地基于预测的频谱与和存储的声音片段相关联的频谱特性的比较,确定与存储的声音片段相关联的级别。可相应地调整存储的声音片段级别。因此,将声音片段和音频信号混合可包括将经级别调整的声音片段与音频段混合。当将经级别调整的声音片段与音频段混合时,经级别调整的声音片段显著地、可能基本上(或者甚至本质上完全地)掩蔽音频伪迹。

[0067] 可对与存储的声音片段相关的上下文信息进行监视。因此,存储声音片段可包括基于上下文信息更新存储的声音片段中的一个或多个。音频信号可与基于网络的游戏相关。因此,上下文信息可与和游戏相关联的虚拟环境相关。音频信号还可与电话、视频或音频会议、或者相关应用相关联。

[0068] 示例性实施例

[0069] 本发明的实施例可涉及以下枚举的示例中的一个或多个。

[0070] 1、一种方法,包括以下步骤:

[0071] 识别音频信号中的至少一个音频段,所述至少一个音频段具有与其相应的持续时间,其中,所述音频段与所述音频信号内的伪迹相关联;

[0072] 检索至少一个存储的具有等于或者超过与所述至少一个段相关联的持续时间的持续时间的声音片段;和

[0073] 将检索的至少一个声音片段与所述音频信号混合;

[0074] 其中,将所述至少一个检索的声音片段与所述音频信号混合使得所述音频伪迹不能被感觉到。

[0075] 2、根据枚举的示例性实施例 1 所述的方法,其中,所述音频伪迹包括所述音频段的丢失或损坏部分中的一个或多个。

[0076] 3、根据枚举的示例性实施例 2 所述的方法,还包括以下步骤:

[0077] 接收音频流,其中,所述音频流包括编码的音频数据的多个分组;和

[0078] 从接收的音频分组收集音频信号。

[0079] 4、根据枚举的示例性实施例 2 所述的方法,与丢失的或损坏的音频段相关联的时间位置被完全包含在所述音频片段的时间位置中。

[0080] 5、根据枚举的示例性实施例 1 所述的方法,还包括以下步骤:

[0081] 将所述至少一个声音片段存储在声音片段储存库中。

[0082] 6、根据枚举的示例性实施例 5 所述的方法,其中,所述检索步骤包括以下步骤:

[0083] 检测识别的至少一个音频段中的音频伪迹;

[0084] 基于音频伪迹的特性查询储存库;和

- [0085] 基于所述声音片段和所述特性之间的匹配响应于查询步骤返回所述声音片段。
- [0086] 7、根据枚举的示例性实施例 6 所述的方法,其中,所述特性包括:
- [0087] 与识别的至少一个段相应的持续时间;和
- [0088] 与音频伪迹相应的至少一个音频属性。
- [0089] 8、根据枚举的示例性实施例 6 所述的方法,其中,当检测到音频伪迹时,所述检索步骤还包括以下步骤:
- [0090] 确定音频伪迹的特性;和
- [0091] 其中,响应于检测步骤或确定步骤中的至少一个执行查询步骤。
- [0092] 9、根据枚举的示例性实施例 8 所述的方法,其中,所述音频伪迹的特性是频率相关的;
- [0093] 其中,所述确定步骤包括以下步骤:
- [0094] 预测对应于与频率相关的特性的频谱。
- [0095] 10、根据枚举的示例性实施例 9 所述的方法,其中,所述查询步骤包括以下步骤:
- [0096] 将预测的频谱与和存储的声音片段相关联的频谱特性进行比较;
- [0097] 其中,所述匹配包括预测的音频伪迹频谱和声音片段频谱特性之间的显著相似性。
- [0098] 11、根据枚举的示例性实施例 10 所述的方法,其中,所述显著相似性包括预测的音频伪迹频谱和声音片段频谱特性之间的基本相同的对应性。
- [0099] 12、根据枚举的示例性实施例 10 所述的方法,还包括以下步骤:
- [0100] 至少部分地基于预测的频谱与和存储的声音片段相关联的频谱特性的比较,确定与存储的声音片段相关联的级别;和
- [0101] 调整存储的声音片段级别;
- [0102] 其中,所述混合步骤包括以下步骤:
- [0103] 将经级别调整的声音片段与音频段混合;
- [0104] 其中,经过混合步骤,经级别调整的声音片段显著地掩蔽音频伪迹。
- [0105] 13、根据枚举的示例性实施例 12 所述的方法,其中,经过混合步骤,经级别调整的声音片段基本上掩蔽音频伪迹。
- [0106] 14、根据枚举的示例性实施例 5 所述的方法,还包括以下步骤:
- [0107] 监视与存储的声音片段相关的上下文信息;
- [0108] 其中,所述存储步骤包括以下步骤:
- [0109] 基于所述上下文信息更新存储的声音片段中的一个或多个。
- [0110] 15、根据枚举的示例性实施例 14 所述的方法,其中,所述音频信号与基于网络的游戏相关;和
- [0111] 其中,所述上下文信息与和游戏相关联的虚拟环境相关。
- [0112] 16、一种方法,包括以下步骤:
- [0113] 识别音频信号中的至少一个音频段,所述至少一个音频段具有与其相应的持续时间,其中,所述音频段与所述音频信号内的伪迹相关联;
- [0114] 检索至少一个存储的具有等于或者超过与所述至少一个段相关联的持续时间的持续时间的声音片段;和

- [0115] 将检索的至少一个声音片段与所述音频信号混合；
- [0116] 其中,将所述至少一个检索的声音片段与所述音频信号混合使得所述音频伪迹不能被感觉到。
- [0117] 17、根据枚举的示例性实施例 16 所述的方法,其中,所述音频伪迹包括所述音频段的丢失或损坏部分中的一个或多个。
- [0118] 18、根据枚举的示例性实施例 17 所述的方法,还包括以下步骤:
- [0119] 接收音频流,其中,所述音频流包括编码的音频数据的多个分组;和
- [0120] 从接收的音频分组收集音频信号。
- [0121] 19、根据枚举的示例性实施例 17 所述的方法,其中,与丢失的或损坏的音频段相关联的时间位置被完全包含在音频片段的时间位置中。
- [0122] 20、根据枚举的示例性实施例 16 所述的方法,还包括以下步骤:
- [0123] 将所述至少一个声音片段存储在声音片段储存库中。
- [0124] 21、根据枚举的示例性实施例 20 所述的方法,其中,所述检索步骤包括以下步骤:
- [0125] 检测识别的至少一个音频段中的音频伪迹;
- [0126] 基于所述音频伪迹的特性查询储存库;和
- [0127] 基于所述声音片段和所述特性之间的匹配响应于查询步骤返回所述声音片段。
- [0128] 22、根据枚举的示例性实施例 21 所述的方法,其中,所述特性包括:
- [0129] 与识别的至少一个段相应的持续时间;和
- [0130] 与音频伪迹相应的至少一个音频属性。
- [0131] 23、根据枚举的示例性实施例 21 所述的方法,当检测到音频伪迹时,所述检索步骤还包括以下步骤:
- [0132] 确定音频伪迹的特性;和
- [0133] 其中,响应于检测步骤或确定步骤中的至少一个执行所述查询步骤。
- [0134] 24、根据枚举的示例性实施例 23 所述的方法,其中,所述确定步骤包括以下步骤:
- [0135] 预测所述特性的听觉表示。
- [0136] 25、根据枚举的示例性实施例 24 所述的方法,其中,所述查询步骤包括以下步骤:
- [0137] 将所述特性的听觉表示与存储的声音片段的听觉表示进行比较;
- [0138] 其中,所述匹配包括与所述特性的听觉表示完全重叠的存储的声音片段的听觉表示。
- [0139] 26、根据枚举的示例性实施例 24 所述的方法,其中,所述查询步骤包括一系列迭代步骤:
- [0140] 将所述特性的听觉表示与存储的声音片段的听觉表示进行比较,以确定存储的声音片段的听觉表示是否与所述特性的听觉表示完全重叠;和
- [0141] 以比较结果为条件,调整存储的声音片段的级别,并重复该比较,直到存储的声音片段的听觉表示与所述特性的听觉表示完全重叠;和
- [0142] 其中,所述混合步骤包括以下步骤:
- [0143] 将经级别调整的声音片段与音频段混合。
- [0144] 27、一种系统,包括:
- [0145] 用于执行枚举的示例性实施例 1-26 中的一个或多个所述的方法的一个或多个步

骤的装置。

[0146] 28、一种基于计算机的设备,包括:

[0147] 至少一个处理器;和

[0148] 包括指令的计算机可读存储介质,当在所述至少一个处理器上执行所述指令时,所述指令控制计算机执行枚举的示例性实施例 1-26 中的一个或多个所述的方法的一个或多个步骤。

[0149] 29、一种装置,包括:

[0150] 路由结构;

[0151] 与所述路由结构耦接的多个活动组件,其被配置为执行处理或逻辑相关功能中的至少一个;和

[0152] 与路由功能耦接的存储介质,其包括指令,当在活动组件上执行所述指令时,所述指令控制所述装置执行以下步骤中的一个或多个:

[0153] 枚举的示例性实施例 1-26 中的一个或多个所述的方法的各步骤:

[0154] 配置所述活动组件;或者

[0155] 执行与以下中的一个或多个相关的功能:

[0156] 根据枚举的示例性实施例 27 所述的系统;或者

[0157] 根据枚举的示例性实施例 28 所述的设备。

[0158] 30、根据枚举的示例性实施例 29 所述的装置,其中,所述装置包括集成电路。

[0159] 31、根据枚举的示例性实施例 30 所述的装置,其中,所述集成电路包括应用特定集成电路。

[0160] 32、根据枚举的示例性实施例 30 所述的装置,其中,所述路由结构、活动组件或存储介质中的一个或多个是可编程的或者可配置的。

[0161] 33、根据枚举的示例性实施例 32 所述的装置,其中,所述集成电路包括以下中的一个或多个:

[0162] 可编程逻辑装置;

[0163] 微控制器;或者

[0164] 现场可编程门阵列。

[0165] 34、一种计算机可读存储介质,包括:

[0166] 指令,当在一个或多个处理器上执行所述指令时,所述指令控制执行根据枚举的示例性实施例 1-26 中的一个或多个所述的方法的步骤。

[0167] 35、一种计算机可读存储介质,包括指令,当在一个或多个处理器上执行所述指令时,所述指令控制执行根据枚举的示例性实施例 1-26 中的一个或多个所述的方法的一个或多个步骤。

[0168] 36、一种计算机可读存储介质,包括指令,当在一个或多个处理器上执行所述指令时,所述指令执行一种或多种功能,包括:

[0169] 控制以下中的一个或多个的步骤的一种或多种功能:

[0170] 根据枚举的示例性实施例 27 中所述的系统;

[0171] 根据枚举的示例性实施例 28 中所述的设备;或者

[0172] 根据枚举的示例性实施例 29-33 中的一个或多个所述的装置;

[0173] 或者

[0174] 配置以下中的一个或多个或者对以下中的一个或多个进行编程：

[0175] 根据枚举的示例性实施例 27 所述的装置；

[0176] 根据枚举的示例性实施例 28 所述的处理器或介质；或者

[0177] 根据枚举的示例性实施例 29-33 中的一个或多个所述的路由结构、活动组件或介质中的一个或多个。

[0178] 等同、扩展、可选和其它

[0179] 如此描述了涉及隐藏音频伪迹的示例性实施例。在此前述说明书中，参照可在实现与实现之间变化的许多特定细节对本发明的示例性实施例进行了描述。因此，关于什么是本发明以及申请人要将什么作为发明的唯一和排他的指示是以发布权利要求的特定的形式从本申请发布的这些权利要求的集合，包含任何随后的校正。这里对于包含于这些权利要求中的术语表达的任何定义应支配在权利要求中使用的这些术语的含义。由此，没有在权利要求中明确表述的限制、要素、性能、特征、优点或属性不应以任何的方式限制这种权利要求的范围。因此，说明书和附图应被视为是示例性的而不是限制性的。

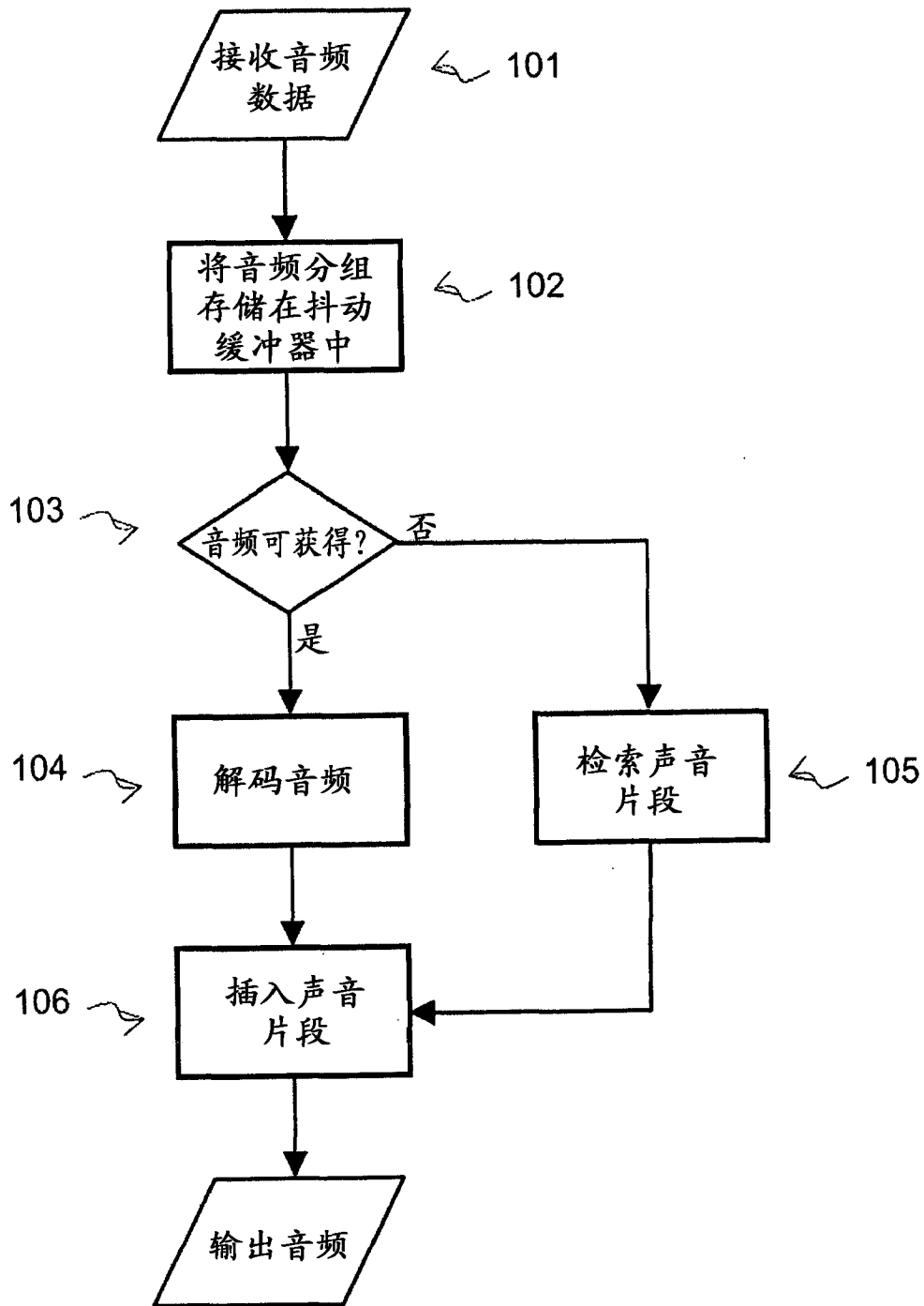


图 1

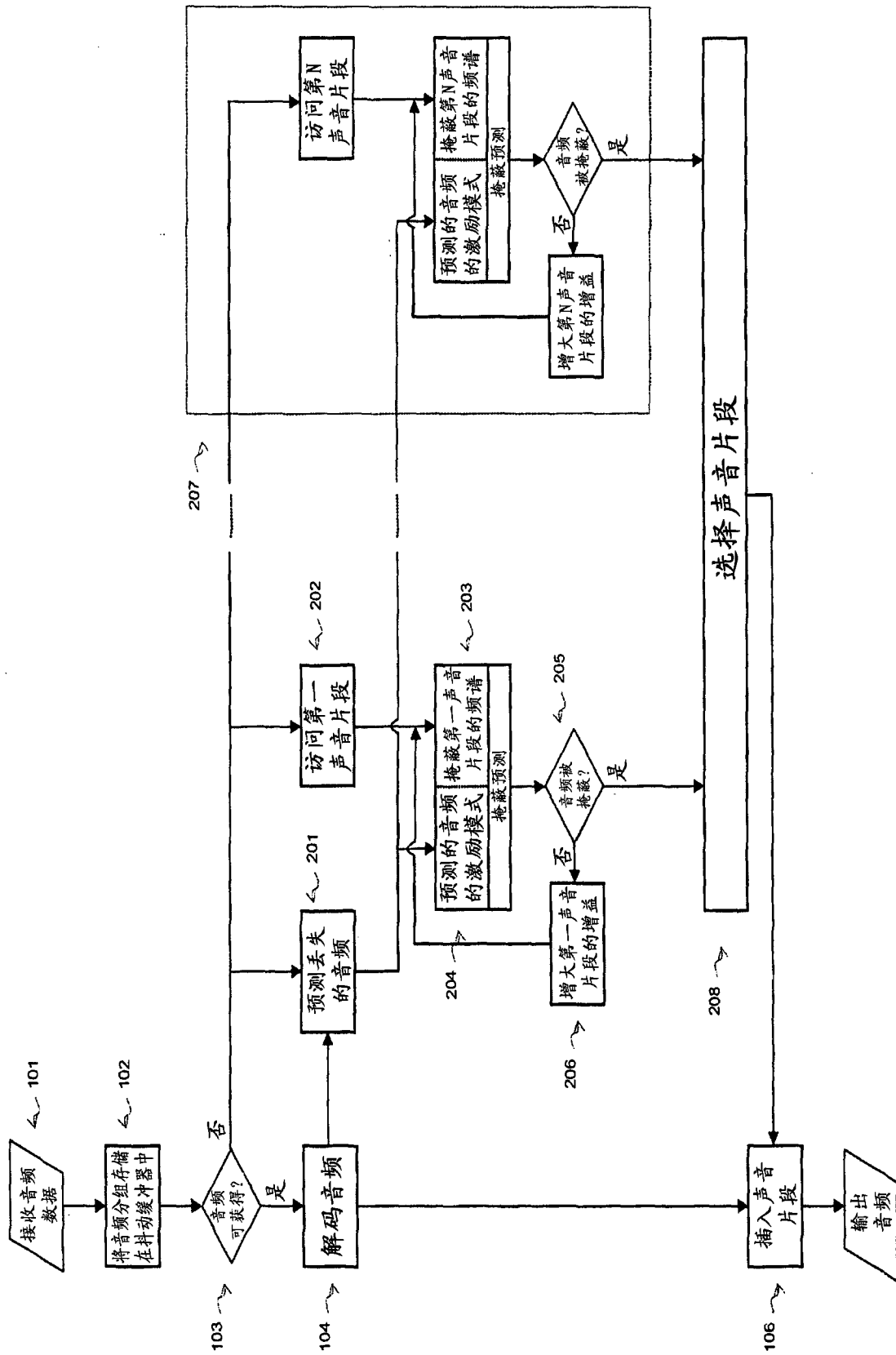


图 2

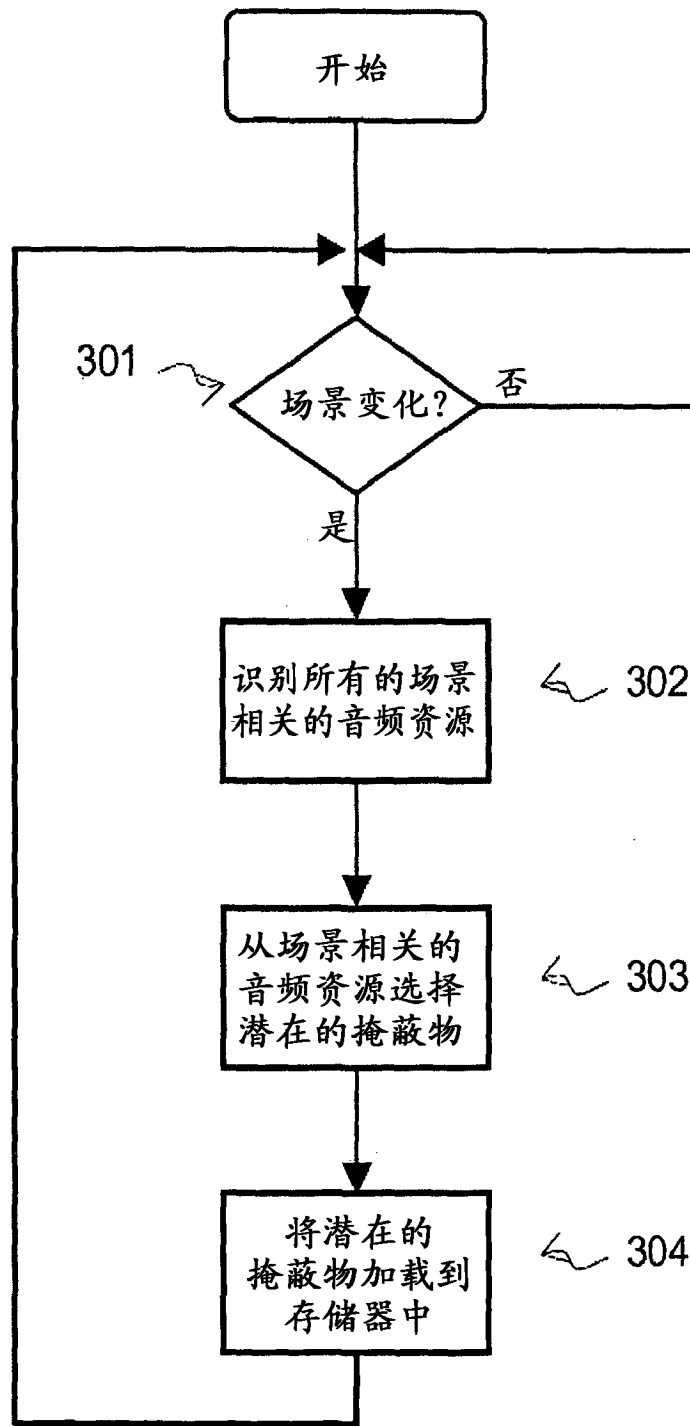


图 3

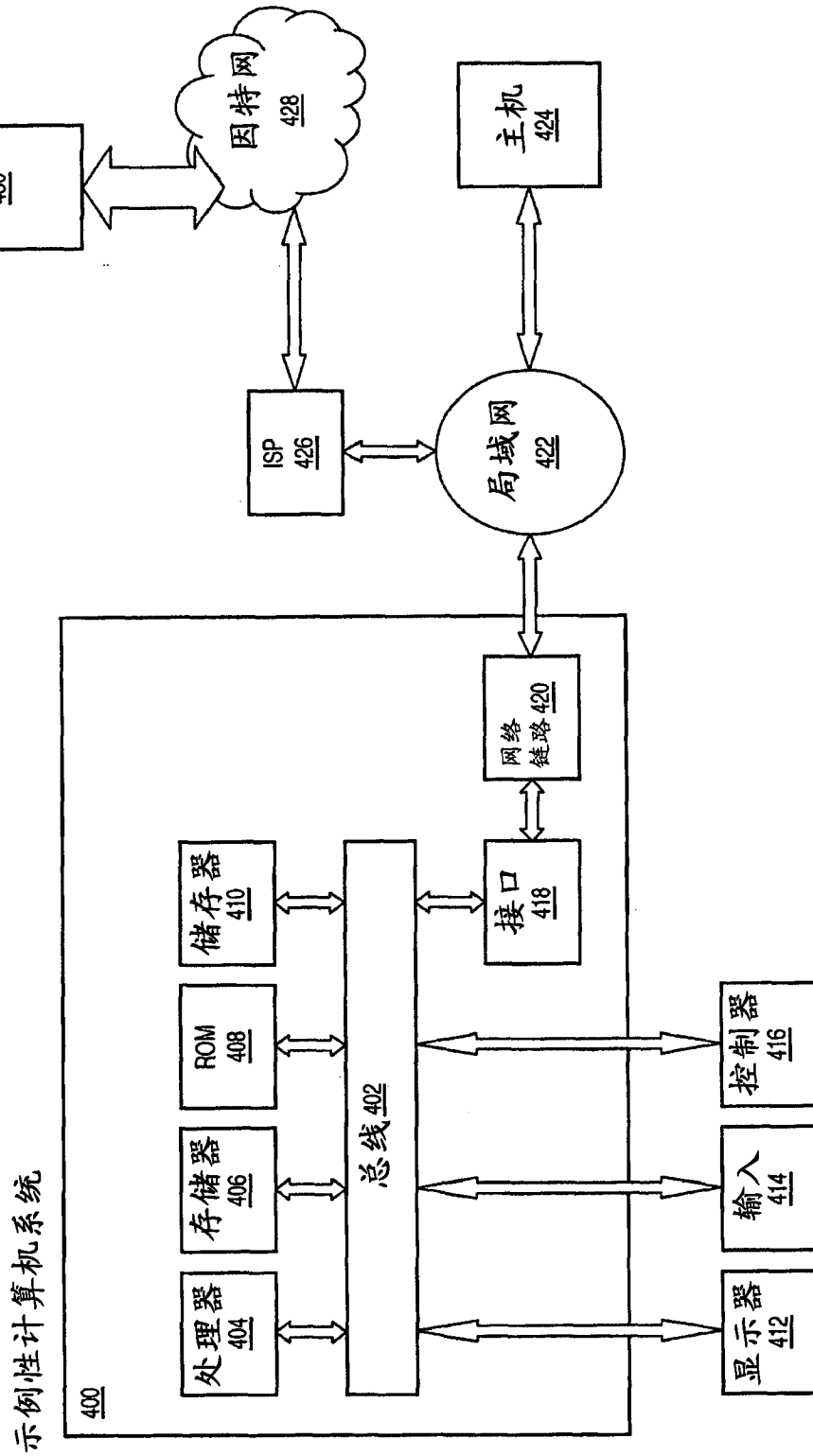


图 4

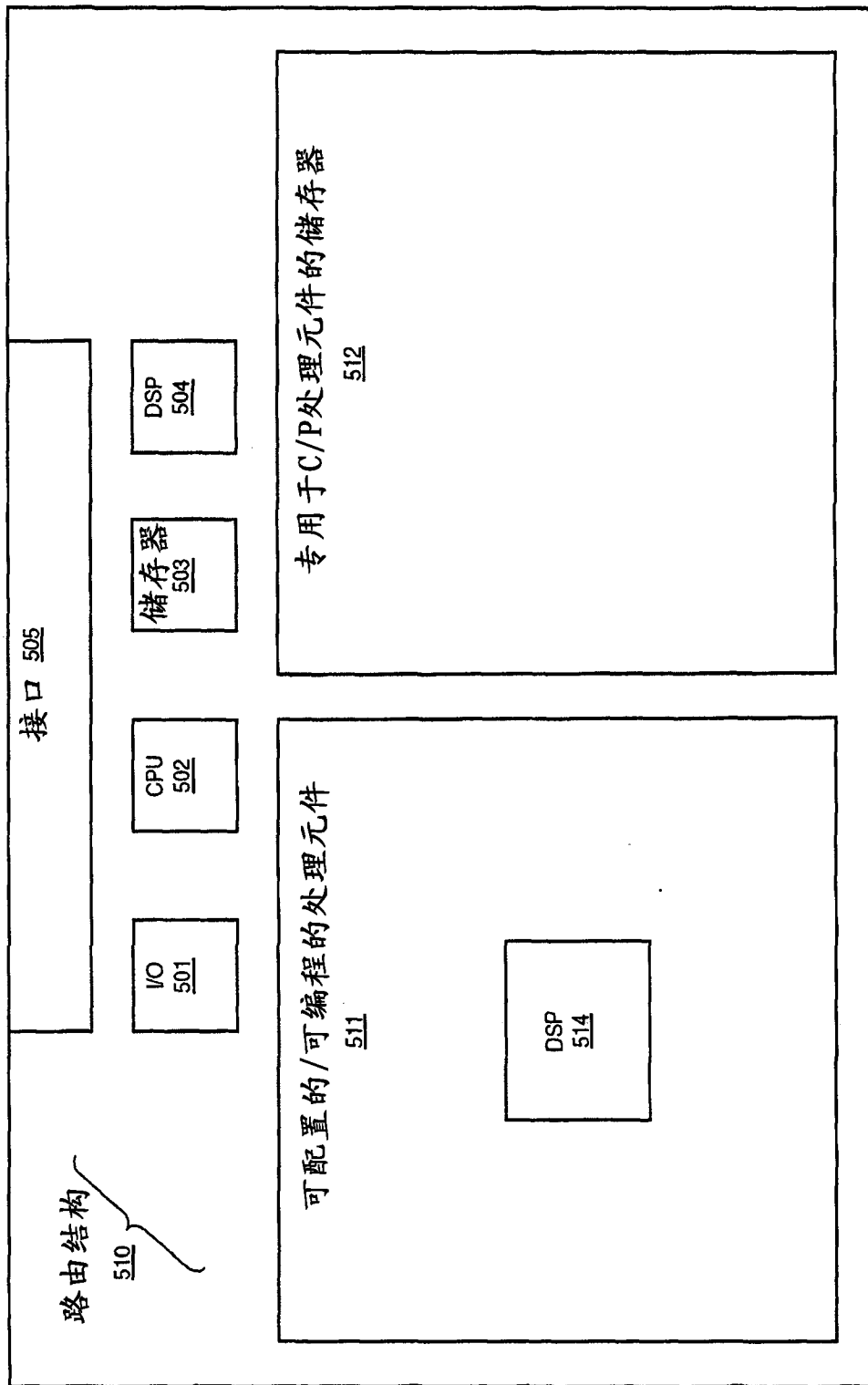


图 5