



(12) 发明专利申请

(10) 申请公布号 CN 104575513 A

(43) 申请公布日 2015. 04. 29

(21) 申请号 201310507704. 3

H04M 1/19(2006. 01)

(22) 申请日 2013. 10. 24

(71) 申请人 展讯通信(上海)有限公司

地址 201203 上海市浦东新区张江高科技园
区祖冲之路 2288 弄展讯中心 1 号楼

(72) 发明人 蒋斌 吴晟 林福辉 徐晶明
李慧文

(74) 专利代理机构 北京集佳知识产权代理有限
公司 11227

代理人 吴靖靓 骆苏华

(51) Int. Cl.

G10L 21/0232(2013. 01)

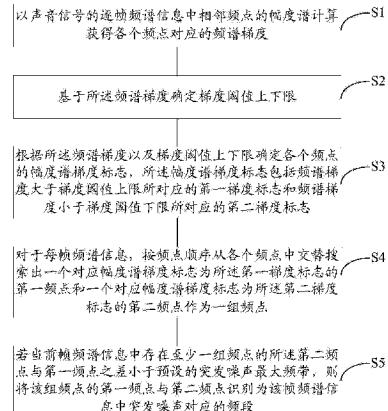
权利要求书3页 说明书11页 附图5页

(54) 发明名称

突发噪声的处理系统、突发噪声的检测及抑
制方法与装置

(57) 摘要

一种突发噪声的处理系统、突发噪声的检测
及抑制方法与装置，所述突发噪声的检测方法包
括：以声音信号的逐帧频谱信息中相邻频点的幅
度谱计算各频点对应的频谱梯度；基于频谱梯度
确定梯度阈值上下限；根据频谱梯度以及梯度阈
值上下限确定各频点的幅度谱梯度标志，所述幅
度谱梯度标志包括第一梯度标志和第二梯度标
志；按频点顺序从各频点中交替搜索出一个对应
幅度谱梯度标志为第一梯度标志的第一频点和一
个对应幅度谱梯度标志为第二梯度标志的第二频
点作为一组频点；若存在至少一组频点的第二频
点与第一频点之差小于预设的突发噪声最大频
带，则该组频点的第一频点与第二频点为突发噪
声对应的频段。本方案能有效检测和抑制背景环
境中特定突发噪声。



1. 一种突发噪声的检测方法,其特征在于,包括:

以声音信号的逐帧频谱信息中相邻频点的幅度谱计算获得各个频点对应的频谱梯度;

基于所述频谱梯度确定梯度阈值上下限;

根据所述频谱梯度以及梯度阈值上下限确定各个频点的幅度谱梯度标志,所述幅度谱梯度标志包括频谱梯度大于梯度阈值上限所对应的第一梯度标志和频谱梯度小于梯度阈值下限所对应的第二梯度标志;

对于每帧频谱信息,按频点顺序从各个频点中交替搜索出一个对应幅度谱梯度标志为所述第一梯度标志的第一频点和一个对应幅度谱梯度标志为所述第二梯度标志的第二频点作为一组频点;

若当前帧频谱信息中存在至少一组频点的所述第二频点与第一频点之差小于预设的突发噪声最大频带,则将该组频点的第一频点与第二频点识别为该帧频谱信息中突发噪声对应的频段。

2. 根据权利要求 1 所述的突发噪声的检测方法,其特征在于,所述声音信号的逐帧频谱信息通过将输入的时域声音信号进行离散傅立叶变换后得到。

3. 根据权利要求 1 所述的突发噪声的检测方法,其特征在于,交替搜索出第一频点和第二频点作为一组频点从大于频率阈值的频点开始进行,所述频率阈值高于语音信号的主要能量所在频带。

4. 根据权利要求 1 所述的突发噪声的检测方法,其特征在于,所述各个频点对应的频谱梯度通过如下方式计算获得:

$$X_D(k) = \begin{cases} 0 & k=0 \\ |X(n,k)| - |X(n,k-1)| & k=1, 2, \dots, N \end{cases}$$
, 其中 k 为频点序号, $X_D(k)$ 为序号为 k 的频点对应的频谱梯度, n 为频谱信息的帧号, $|X(n, k)|$ 为帧号 n 的频谱信息中序号为 k 的频点的幅度谱, $|X(n, k-1)|$ 为帧号 n 的频谱信息中序号为 k-1 的频点的幅度谱, N 为有效频谱点数。

5. 根据权利要求 4 所述的突发噪声的检测方法,其特征在于,所述基于所述频谱梯度确定梯度阈值上下限包括:

以 $X_D(k)$ 大于 0 的平均值与第一系数的乘积作为梯度阈值上限 Th_{up} ;

以 $X_D(k)$ 小于 0 的平均值与第二系数的乘积作为梯度阈值下限 Th_{dn} ;

所述第一系数和第二系数的取值范围均为 $(1, +\infty)$ 。

6. 根据权利要求 5 所述的突发噪声的检测方法,其特征在于,所述根据所述频谱梯度以及梯度阈值上下限确定各个频点的幅度谱梯度标志包括:

按各个频点对应的频谱梯度、所述梯度阈值上限 Th_{up} 和梯度阈值下限 Th_{dn} , 通过如下方式确定各个频点的幅度谱梯度标志 $P(k)$:

$$P(k) = \begin{cases} 1 & X_D(k) > Th_{up} \\ 0 & Th_{dn} \leq X_D(k) \leq Th_{up} \\ -1 & X_D(k) < Th_{dn} \end{cases}$$

7. 根据权利要求 6 所述的突发噪声的检测方法,其特征在于,所述对于每帧频谱信息,

按频点顺序从各个频点中交替搜索出一个对应幅度谱梯度标志为所述第一梯度标志的第一频点和一个对应幅度谱梯度标志为所述第二梯度标志的第二频点作为一组频点包括：

按频点序号由小到大的顺序先搜索出第一个幅度谱梯度标志 $P(k)=1$ 的第一频点，该第一频点的频点序号为 k_{dn} ；

在搜索出频点序号为 k_{dn} 的第一频点之后，从序号为 $k_{dn}+1$ 的频点开始继续搜索出第一个幅度谱梯度标志 $P(k)=-1$ 的第二频点，该第二频点的频点序号为 k_{up} ；

将频点序号为 k_{dn} 的第一频点和频点序号为 k_{up} 的第二频点作为搜索出的一组频点；

从序号为 $k_{up}+1$ 的频点开始重复上述步骤继续搜索，直至完成对当前帧频谱信息中序号为 N 的频点的搜索。

8. 一种突发噪声的检测装置，其特征在于，包括：

频谱梯度计算单元，适于以声音信号的逐帧频谱信息中相邻频点的幅度谱计算获得各个频点对应的频谱梯度；

第一确定单元，适于基于所述频谱梯度确定梯度阈值上下限；

第二确定单元，适于根据所述频谱梯度以及梯度阈值上下限确定各个频点的幅度谱梯度标志，所述幅度谱梯度标志包括频谱梯度大于梯度阈值上限所对应的第一梯度标志和频谱梯度小于梯度阈值下限所对应的第二梯度标志；

搜索单元，对于每帧频谱信息，按频点顺序从各个频点中交替搜索出一个对应幅度谱梯度标志为所述第一梯度标志的第一频点和一个对应幅度谱梯度标志为所述第二梯度标志的第二频点作为一组频点；

识别单元，适于在当前帧频谱信息中存在至少一组频点的所述第二频点与第一频点之差小于预设的突发噪声最大频带时，将该组频点的第一频点与第二频点识别为该帧频谱信息中突发噪声对应的频段。

9. 根据权利要求 8 所述的突发噪声的检测装置，其特征在于，所述搜索单元从大于频率阈值的频点开始搜索，所述频率阈值高于语音信号的主要能量所在频带。

10. 一种突发噪声的抑制方法，其特征在于，包括：

以权利要求 1 至 7 任一项所述的突发噪声的检测方法对声音信号的逐帧频谱信息进行检测后，对于检测出存在突发噪声的各帧频谱信息，将其中识别为该帧频谱信息中突发噪声对应的频段的第一频点和第二频点之间的所有频点对应的幅度谱进行噪声抑制处理。

11. 根据权利要求 10 所述的突发噪声的抑制方法，其特征在于，所述噪声抑制处理通过如下方式进行：

$$|Y(n, k)| = \begin{cases} \alpha * |X(n, k)| & k_{dn} - rn \leq k \leq k_{up} + rn \\ |X(n, k)| & \text{其他} \end{cases};$$

其中 n 为频谱信息的帧号，k 为频点序号， $|Y(n, k)|$ 为帧号 n 的频谱信息中序号为 k 的频点的经过所述噪声抑制处理后的幅度谱， α 的取值范围为 $[0, 1]$ ， $|X(n, k)|$ 为帧号 n 的频谱信息中序号为 k 的频点的幅度谱， k_{dn} 和 k_{up} 分别为帧号 n 的频谱信息中识别为突发噪声对应的频段的第一频点和第二频点的序号，rn 为检测出的突发噪声附近频段对应的值。

12. 一种突发噪声的抑制装置，其特征在于，包括：

第一处理单元，适于在权利要求 8 或 9 所述的突发噪声的检测装置检测出存在突发噪

声的各帧频谱信息后,将其中识别为该帧频谱信息中突发噪声对应的频段的第一频点和第二频点之间的所有频点对应的幅度谱进行突发噪声抑制处理。

13. 一种突发噪声的处理系统,其特征在于,包括:

权利要求 8 或 9 所述的突发噪声的检测装置;

权利要求 12 所述的突发噪声的抑制装置。

14. 根据权利要求 13 所述的突发噪声的处理系统,其特征在于,还包括第一变换单元,适于将输入的时域声音信号进行离散傅立叶变换以得到所述声音信号的逐帧频谱信息。

15. 根据权利要求 13 所述的突发噪声的处理系统,其特征在于,还包括第二处理单元,适于以所述突发噪声抑制处理之前声音信号的相位谱作为所述突发噪声抑制处理之后声音信号的相位谱,并将其与经过所述突发噪声抑制处理之后声音信号的幅度谱进行计算,得到经过所述突发噪声抑制处理后声音信号的逐帧频谱信息。

16. 根据权利要求 14 或 15 所述的突发噪声的处理系统,其特征在于,还包括第二变换单元,适于将经过所述突发噪声抑制处理后声音信号的逐帧频谱信息进行离散傅立叶逆变换以得到处理后的时域声音信号。

突发噪声的处理系统、突发噪声的检测及抑制方法与装置

技术领域

[0001] 本发明涉及音频处理领域,特别涉及一种突发噪声的处理系统、突发噪声的检测及抑制方法与装置。

背景技术

[0002] 移动终端(例如手机)经常在噪声环境下使用,周围噪声会极大地干扰正常的语音通信,导致语音可懂度降低和语音质量下降,不利于语音通信的舒适度;同时,录制声音信号时,噪声会干扰所需的目标声音信号。因此,降低带噪语音中的噪声可以改善语音通信的舒适度,提高录制声音信号的质量。

[0003] 通常的移动终端设备使用一个或者多个麦克风进行语音降噪处理,即单通道和多通道语音降噪技术;单通道语音降噪技术一般适用于对平稳噪声的抑制或消除,较难于去除非平稳噪声;多通道语音降噪技术可以形成指向性波束,达到抑制非平稳噪声的目的,由于移动终端设备的尺寸和阵列中麦克风个数等因素的限制,波束形成技术的性能十分有限。目前广泛使用的多通道降噪技术大多基于两个麦克风,一个作为主要麦克风拾取目标语音信号(说话人),一个作为参考麦克风拾取背景噪声,根据主要麦克风和参考麦克风的相互关系进行语音增强,该方法的性能较传统单通道降噪技术有一定提高,可以处理某些非平稳噪声。

[0004] 当移动终端在户外环境使用时,环境噪声中可能会包括一些突发噪声,比如汽车和火车的鸣笛声、工厂机器噪声等,此类噪声通常能量很强,会极大地降低通话的语音质量和语音可懂度。当前移动终端使用的单通道和双通道降噪技术无法有效地消除此类突发噪声,需要根据此类噪声的特点进行专门处理。

发明内容

[0005] 本发明解决的问题是现有技术难以有效检测和抑制背景环境中的特定突发噪声信号。

[0006] 为解决上述问题,本发明技术方案提供一种突发噪声的检测方法,包括:

[0007] 以声音信号的逐帧频谱信息中相邻频点的幅度谱计算获得各个频点对应的频谱梯度;

[0008] 基于所述频谱梯度确定梯度阈值上下限;

[0009] 根据所述频谱梯度以及梯度阈值上下限确定各个频点的幅度谱梯度标志,所述幅度谱梯度标志包括频谱梯度大于梯度阈值上限所对应的第一梯度标志和频谱梯度小于梯度阈值下限所对应的第二梯度标志;

[0010] 对于每帧频谱信息,按频点顺序从各个频点中交替搜索出一个对应幅度谱梯度标志为所述第一梯度标志的第一频点和一个对应幅度谱梯度标志为所述第二梯度标志的第二频点作为一组频点;

[0011] 若当前帧频谱信息中存在至少一组频点的所述第二频点与第一频点之差小于预

设的突发噪声最大频带，则将该组频点的第一频点与第二频点识别为该帧频谱信息中突发噪声对应的频段。

[0012] 可选的，所述声音信号的逐帧频谱信息通过将输入的时域声音信号进行离散傅立叶变换后得到。

[0013] 可选的，交替搜索出第一频点和第二频点作为一组频点从大于频率阈值的频点开始进行，所述频率阈值高于语音信号的主要能量所在频带。

[0014] 可选的，所述各个频点对应的频谱梯度通过如下方式计算获得：

$$[0015] X_D(k) = \begin{cases} 0 & k=0 \\ |X(n,k)| - |X(n,k-1)| & k=1,2,\dots,N \end{cases}, \text{其中 } k \text{ 为频点序号, } X_D(k) \text{ 为序号}$$

为 k 的频点对应的频谱梯度，n 为频谱信息的帧号，|X(n, k)| 为帧号 n 的频谱信息中序号为 k 的频点的幅度谱，|X(n, k-1)| 为帧号 n 的频谱信息中序号为 k-1 的频点的幅度谱，N 为有效频谱点数。

[0016] 可选的，所述基于所述频谱梯度确定梯度阈值上下限包括：

[0017] 以 $X_D(k)$ 大于 0 的平均值与第一系数的乘积作为梯度阈值上限 Th_{up} ；

[0018] 以 $X_D(k)$ 小于 0 的平均值与第二系数的乘积作为梯度阈值下限 Th_{dn} ；

[0019] 所述第一系数和第二系数的取值范围均为 $(1, +\infty)$ 。

[0020] 可选的，所述根据所述频谱梯度以及梯度阈值上下限确定各个频点的幅度谱梯度标志包括：

[0021] 按各个频点对应的频谱梯度、所述梯度阈值上限 Th_{up} 和梯度阈值下限 Th_{dn} ，通过如下方式确定各个频点的幅度谱梯度标志 $P(k)$ ：

$$[0022] P(k) = \begin{cases} 1 & X_D(k) > Th_{up} \\ 0 & Th_{dn} \leq X_D(k) \leq Th_{up} \\ -1 & X_D(k) < Th_{dn} \end{cases}$$

[0023] 可选的，所述对于每帧频谱信息，按频点顺序从各个频点中交替搜索出一个对应幅度谱梯度标志为所述第一梯度标志的第一频点和一个对应幅度谱梯度标志为所述第二梯度标志的第二频点作为一组频点包括：

[0024] 按频点序号由小到大的顺序先搜索出第一个幅度谱梯度标志 $P(k)=1$ 的第一频点，该第一频点的频点序号为 k_{dn} ；

[0025] 在搜索出频点序号为 k_{dn} 的第一频点之后，从序号为 $k_{dn}+1$ 的频点开始继续搜索出第一个幅度谱梯度标志 $P(k)=-1$ 的第二频点，该第二频点的频点序号为 k_{up} ；

[0026] 将频点序号为 k_{dn} 的第一频点和频点序号为 k_{up} 的第二频点作为搜索出的一组频点；

[0027] 从序号为 $k_{up}+1$ 的频点开始重复上述步骤继续搜索，直至完成对当前帧频谱信息中序号为 N 的频点的搜索。

[0028] 为解决上述问题，本发明技术方案还提供一种突发噪声的检测装置，包括：

[0029] 频谱梯度计算单元，适于以声音信号的逐帧频谱信息中相邻频点的幅度谱计算获得各个频点对应的频谱梯度；

[0030] 第一确定单元，适于基于所述频谱梯度确定梯度阈值上下限；

[0031] 第二确定单元，适于根据所述频谱梯度以及梯度阈值上下限确定各个频点的幅度谱梯度标志，所述幅度谱梯度标志包括频谱梯度大于梯度阈值上限所对应的第一梯度标志和频谱梯度小于梯度阈值下限所对应的第二梯度标志；

[0032] 搜索单元，对于每帧频谱信息，按频点顺序从各个频点中交替搜索出一个对应幅度谱梯度标志为所述第一梯度标志的第一频点和一个对应幅度谱梯度标志为所述第二梯度标志的第二频点作为一组频点；

[0033] 识别单元，适于在当前帧频谱信息中存在至少一组频点的所述第二频点与第一频点之差小于预设的突发噪声最大频带时，将该组频点的第一频点与第二频点识别为该帧频谱信息中突发噪声对应的频段。

[0034] 为解决上述问题，本发明技术方案还提供一种突发噪声的抑制方法，包括：

[0035] 以上述突发噪声的检测方法对声音信号的逐帧频谱信息进行检测后，对于检测出存在突发噪声的各帧频谱信息，将其中识别为该帧频谱信息中突发噪声对应的频段的第一频点和第二频点之间的所有频点对应的幅度谱进行噪声抑制处理。

[0036] 可选的，所述噪声抑制处理通过如下方式进行：

[0037]

$$|Y(n,k)| = \begin{cases} \alpha * |X(n,k)| & kdn - rn \leq k \leq kup + rn \\ |X(n,k)| & \text{其他} \end{cases};$$

[0038] 其中 n 为频谱信息的帧号，k 为频点序号，|Y(n, k)| 为帧号 n 的频谱信息中序号为 k 的频点的经过所述噪声抑制处理后的幅度谱， α 的取值范围为 [0, 1)，|X(n, k)| 为帧号 n 的频谱信息中序号为 k 的频点的幅度谱， kdn 和 kup 分别为帧号 n 的频谱信息中识别为突发噪声对应的频段的第一频点和第二频点的序号， rn 为检测出的突发噪声附近频段对应的值。

[0039] 为解决上述问题，本发明技术方案还提供一种突发噪声的抑制装置，包括：第一处理单元，适于在上述突发噪声的检测装置检测出存在突发噪声的各帧频谱信息后，将其中识别为该帧频谱信息中突发噪声对应的频段的第一频点和第二频点之间的所有频点对应的幅度谱进行突发噪声抑制处理。

[0040] 为解决上述问题，本发明技术方案还提供一种突发噪声的处理系统，包括：上述突发噪声的检测装置以及上述突发噪声的抑制装置。

[0041] 与现有技术相比，本发明的技术方案至少具有以下优点：

[0042] 通过声音信号的逐帧频谱信息中相邻频点的幅度谱计算出各个频点对应的频谱梯度，并基于各频点的频谱梯度对应的幅度谱梯度标志，检测各帧频谱信息中是否存在突发噪声，将搜索出的每一组频点中两者之差小于预设的突发噪声最大频带的第一频点与第二频点识别为该帧频谱信息中突发噪声谱峰对应的频段，由此能够实现针对特定突发噪声的特点有效检测出背景环境中的特定突发噪声信号。

[0043] 对于检测出存在突发噪声的各帧频谱信息，将其中识别为该帧频谱信息中突发噪声对应的频段的第一频点和第二频点之间的所有频点对应的幅度谱进行噪声抑制处理，以实现对特定突发噪声进行针对性的噪声抑制，从而能够有效降低噪声，改善语音可懂度和语音质量，提高语音通信的舒适度。

附图说明

- [0044] 图 1 是本发明实施方式提供的突发噪声的检测方法的流程示意图；
- [0045] 图 2 是本发明实施例一的突发噪声的检测方法的流程示意图；
- [0046] 图 3 是本发明实施例一的突发噪声频谱搜索的流程示意图；
- [0047] 图 4 是本发明实施例一的确定突发噪声谱峰的示意图；
- [0048] 图 5 是本发明实施例一的突发噪声的检测装置的结构示意图；
- [0049] 图 6 是本发明实施例三的突发噪声的处理系统的结构示意图。

具体实施方式

[0050] 如背景技术所述，现有的语音降噪技术无法有效地消除环境噪声中可能会包括的一些如汽车和火车的鸣笛声、工厂机器噪声等突发噪声。本发明实施方式提出一种突发噪声的检测及抑制方法，根据背景环境中特定突发噪声的特点进行专门处理，通过分析当前的声音信号频谱中各频点（频率点）的频谱梯度相对应的幅度谱梯度标志，对是否存在特定突发噪声进行检测，进而抑制此类噪声，从而能够降低噪声，改善语音可懂度和语音质量，提高语音通信的舒适度。

- [0051] 参阅图 1，本发明实施方式提供的突发噪声的检测方法包括如下步骤：
 - [0052] 步骤 S1，以声音信号的逐帧频谱信息中相邻频点的幅度谱计算获得各个频点对应的频谱梯度；
 - [0053] 步骤 S2，基于所述频谱梯度确定梯度阈值上下限；
 - [0054] 步骤 S3，根据所述频谱梯度以及梯度阈值上下限确定各个频点的幅度谱梯度标志，所述幅度谱梯度标志包括频谱梯度大于梯度阈值上限所对应的第一梯度标志和频谱梯度小于梯度阈值下限所对应的第二梯度标志；
 - [0055] 步骤 S4，对于每帧频谱信息，按频点顺序从各个频点中交替搜索出一个对应幅度谱梯度标志为所述第一梯度标志的第一频点和一个对应幅度谱梯度标志为所述第二梯度标志的第二频点作为一组频点；
 - [0056] 步骤 S5，若当前帧频谱信息中存在至少一组频点的所述第二频点与第一频点之差小于预设的突发噪声最大频带，则将该组频点的第一频点与第二频点识别为该帧频谱信息中突发噪声对应的频段。
 - [0057] 基于上述突发噪声的检测方法，本发明实施方式还提供一种突发噪声的抑制方法，包括：以上述突发噪声的检测方法对声音信号的逐帧频谱信息进行检测后，对于检测出存在突发噪声的各帧频谱信息，将其中识别为该帧频谱信息中突发噪声对应的频段的第一频点和第二频点之间的所有频点对应的幅度谱进行噪声抑制处理。
 - [0058] 本发明实施方式提供的突发噪声的检测方法以及抑制方法只针对特定突发噪声进行处理，对目标声音信号造成的失真很小，保证目标声音信号的同时尽可能地降低突发噪声，它可以单独使用，也可以结合现有的单通道和多通道语音降噪技术使用，可以用在语音降噪处理之前作为语音降噪处理的前处理部分，也可以用在语音降噪处理之后作为语音降噪处理的后处理部分，两者相结合可以消除目标语音外的其他噪声信号。
 - [0059] 为使本发明的上述目的、特征和优点能够更为明显易懂，下面结合附图对本发明

的具体实施例做详细的说明。

[0060] 实施例一

[0061] 本实施例中,以对移动终端(例如手机)采集的声音信号中突发噪声的检测为例,对所述突发噪声的检测方法的具体实施进行说明。

[0062] 如图 2 所示,首先执行步骤 S201,将输入的时域声音信号进行短时傅里叶变换,得到声音信号的逐帧频谱信息。

[0063] 步骤 S201 是生成输入时域声音信号的频谱信息的过程。假设移动终端有 L 个麦克风,则当前的第 j 通道麦克风采集到的离散数字声音信号分别表示为 $x_j(m)$, $j=1, \dots, L$, m 表示时间的离散时刻,通过短时傅立叶变换便能够得到逐帧的频谱信息。

[0064] 具体地,对采集的时域声音信号通过重叠分帧得到逐帧信号,对每帧信号进行加窗函数处理,对加窗后的信号进行离散傅立叶变换,分别得到第 n 帧的频谱表示为 $X_j(n, k)$,其中 n 为频谱信息的帧号, k 表示离散傅立叶变换后的频点序号,计算公式如下:

[0065]

$$X_j(n, k) = \sum_{m=-\infty}^{+\infty} x_j(m) w(nR - m) e^{-i2\pi km/NFFT} \quad (1)$$

[0066] 其中 R 是逐帧信号的采样点更新数, NFFT 是离散傅立叶变换点数, $w(m)$ 是窗函数。窗函数类型可以选择汉明窗、汉宁窗、矩形窗等语音信号处理常用的窗函数,帧长选择为 10 ~ 40 毫秒(ms),典型值为 20ms;前后两帧之间有重叠,重叠范围根据实际情况选择。

[0067] 本实施例中,离散傅立叶变换可以通过快速傅立叶变换(FFT, Fast Fourier Transform)来实现,对加窗后的信号直接进行 FFT,或者为了提高频率分辨率和满足 FFT 的长度是 2 的指数倍条件,进行补零操作,如进行 NFFT 点 FFT 变换,可以得到 NFFT 点频谱值。比如,若采样率 F_s 为 16kHz,选择汉明窗,帧长为 20ms,帧间重叠为 50%,则每帧信号共有 320 个采样点,采样点更新数 $R=160$,通过在信号末尾补零得到 512 个采样点,进行 $NFFT=512$ 点的 FFT 可得到 512 个频点值。

[0068] 需要说明的是,步骤 S201 中声音信号的频谱信息处理过程可以和其他语音信号处理过程共享,比如在其他语音降噪技术中,同样需要生成声音信号的频谱信息。因此,本发明实施方式提供的突发噪声的检测方法并非必然包括步骤 S201,在其他实施例中,也可以直接对输入的声音信号的逐帧频谱信息进行后续处理。

[0069] 步骤 S201 之后,执行步骤 S202,计算所述逐帧频谱信息中各频点的幅度谱。

[0070] 本实施例中,根据短时频谱信息计算得到的幅度谱,可以记为 $|X(n, k)|$,其中 $|\cdot|$ 为取模符号,以幅度谱作为处理的主要特征。

[0071] 步骤 S202 之后,执行步骤 S203,以声音信号的逐帧频谱信息中相邻频点的幅度谱计算获得各个频点对应的频谱梯度。

[0072] 具体地,根据逐帧频谱信息中各频点的幅度谱 $|X(n, k)|$ 得到各频点对应的频谱梯度可以通过如下方式进行计算:

[0073]

$$X_D(k) = \begin{cases} 0 & k=0 \\ |X(n, k)| - |X(n, k-1)| & k=1, 2, \dots, N \end{cases} \quad (2)$$

[0074] 其中 k 为频点序号, $X_D(k)$ 为序号为 k 的频点对应的频谱梯度, n 为频谱信息的帧号, $|X(n, k)|$ 为帧号 n 的频谱信息中序号为 k 的频点的幅度谱, $|X(n, k-1)|$ 为帧号 n 的频谱信息中序号为 k-1 的频点的幅度谱, N 为有效频谱点数, 通常可以为 $N=NFFT/2+1$ 。

[0075] 由公式(2)可知, 各频点对应的频谱梯度能够反映出各频点的幅度变化量, 后续步骤搜索突发噪声谱峰正是基于各频点的幅度变化情况进行的, 至于所述突发噪声谱峰的确定过程将在下面详细说明。

[0076] 获得各个频点对应的频谱梯度后, 执行步骤 S204, 基于所述频谱梯度确定梯度阈值上限和梯度阈值下限。

[0077] 本实施例中, 步骤 S204 具体可以采用以下方式确定梯度阈值上下限: 以 $X_D(k)$ 大于 0 的平均值与第一系数 μ_{up} 的乘积作为梯度阈值上限 Th_{up} ; 以 $X_D(k)$ 小于 0 的平均值与第二系数 μ_{dn} 的乘积作为梯度阈值下限 Th_{dn} ; 所述第一系数 μ_{up} 和第二系数 μ_{dn} 的取值范围均为 $(1, +\infty)$ 。

[0078] 具体地, 在计算所述梯度阈值上限 Th_{up} 时, 以 $X_D(k)>0$ 的平均值为基准, 可以通过如下方式表示:

[0079]

$$Th_{up} = \underset{X_D(k)>0}{\text{mean}} \{X_D(k)\} * \mu_{up} \quad (3)$$

[0080] 其中 μ_{up} 的取值范围为 $\mu_{up}>1$, 典型值为 5, $\text{mean}\{\cdot\}$ 表示计算满足给定条件的向量的均值。

[0081] 同时, 在计算梯度阈值下限 Th_{dn} 时, 以 $X_D(k)<0$ 的平均值为基准, 可以通过如下方式表示:

[0082]

$$Th_{dn} = \underset{X_D(k)<0}{\text{mean}} \{X_D(k)\} * \mu_{dn} \quad (4)$$

[0083] 其中 μ_{dn} 的取值范围为 $\mu_{dn}>1$, 典型值为 4。

[0084] 确定梯度阈值上下限之后, 执行步骤 S205, 根据所述频谱梯度以及梯度阈值上下限确定各个频点的幅度谱梯度标志, 所述幅度谱梯度标志包括频谱梯度大于梯度阈值上限所对应的第一梯度标志和频谱梯度小于梯度阈值下限所对应的第二梯度标志。

[0085] 本实施例中, 根据各频点对应的频谱梯度以及梯度阈值上下限计算各个频点的幅度谱梯度标志, 可以通过如下方式进行:

[0086]

$$P(k) = \begin{cases} 1 & X_D(k) > Th_{up} \\ 0 & Th_{dn} \leq X_D(k) \leq Th_{up} \\ -1 & X_D(k) < Th_{dn} \end{cases} \quad (5)$$

[0087] 由公式(5)可以看出, 当任一频点对应的频谱梯度大于所述梯度阈值上限 Th_{up} , 则该频点对应的幅度谱梯度标志 $P(k)=1$, 即所述幅度谱梯度标志 $P(k)$ 包括的第一梯度标志为“1”; 当任一频点对应的频谱梯度小于所述梯度阈值下限 Th_{dn} , 则该频点对应的幅度谱梯度标志 $P(k)=-1$, 即所述幅度谱梯度标志 $P(k)$ 包括的第二梯度标志为“-1”; 当任一频点对应的频谱梯度介于所述梯度阈值下限 Th_{dn} 和梯度阈值上限 Th_{up} 之间, 则该频点对应的幅度

谱梯度标志 $P(k)=0$, 可以将“0”作为所述幅度谱梯度标志 $P(k)$ 包括的第三梯度标志。

[0088] 需要说明的是, 本实施例中分别以“1”和“-1”作为所述第一梯度标志和第二梯度标志, 主要是为了便于后续搜索和计算, 提高处理效率, 在其他实施例也可以采用其他符号或数字作为第一梯度标志和第二梯度标志。

[0089] 在确定各频点对应的幅度谱梯度标志之后, 执行步骤 S206, 对于每帧频谱信息, 按频点顺序从各个频点中交替搜索出一个对应幅度谱梯度标志为所述第一梯度标志的第一频点和一个对应幅度谱梯度标志为所述第二梯度标志的第二频点作为一组频点。

[0090] 根据得到的各频点对应的幅度谱梯度标志, 便可以开始对逐帧频谱信息中的突发噪声进行检测。本实施例中, 根据各频点对应的幅度谱梯度标志对逐帧频谱信息中的突发噪声进行检测是通过先搜索出有可能确定为突发噪声谱峰的一组或一组以上的频点, 然后判断每一组频点中的第一频点与第二频点是否满足可识别为突发噪声的条件。

[0091] 本实施例中, 步骤 S206 对于每帧频谱信息按频点顺序从各个频点中交替搜索出第一频点和第二频点作为一组频点的具体实现过程如下:

[0092] 按频点序号由小到大的顺序先搜索出第一个幅度谱梯度标志 $P(k)=1$ 的第一频点, 该第一频点的频点序号为 k_{dn} ; 在搜索出频点序号为 k_{dn} 的第一频点之后, 从序号为 $k_{dn}+1$ 的频点开始继续搜索出第一个幅度谱梯度标志 $P(k)=-1$ 的第二频点, 该第二频点的频点序号为 k_{up} ; 将频点序号为 k_{dn} 的第一频点和频点序号为 k_{up} 的第二频点作为搜索出的一组频点; 从序号为 $k_{up}+1$ 的频点开始重复上述步骤继续搜索, 直至完成对当前帧频谱信息中序号为 N 的频点的搜索。

[0093] 在通常情况下, 考虑到突发噪声(比如鸣笛声)的主要能量一般集中在频率较高处, 因此, 在本实施例中, 步骤 S206 中交替搜索出第二频点和第一频点作为一组频点是从大于频率阈值的频点开始进行的, 所述频率阈值高于语音信号的主要能量所在频带。由于在语音信号的频带中也有可能搜索出满足识别为突发噪声对应频段的第一频点和第二频点, 因此, 从大于所述频率阈值的频点开始进行搜索的方式, 能够防止后续识别是否为突发噪声时产生误判的情况, 进而避免对相应语音信号的误处理, 从而能降低语音失真。

[0094] 在实际实施时, 由于语音信号的主要能量频带通常在 1500Hz 以下, 为了尽可能降低语音失真, 一般要求从 1500Hz 以上的频率点开始搜索(所述频率阈值为 1500Hz), 即满足:

[0095]

$$fdn_k > \text{int}\left(\frac{1500}{F_s/\text{NFFT}}\right) \quad (6)$$

[0096] 其中 $\text{int}()$ 表示取整数, fdn_k 为搜索的起始频率点所对应的频点序号。

[0097] 搜索出当前帧频谱信息中的一组或一组以上频点后, 执行步骤 S207, 判断当前帧频谱信息中是否存在至少一组频点的所述第二频点与第一频点之差小于预设的突发噪声最大频带, 是则执行步骤 S208, 将该组频点的第一频点与第二频点识别为该帧频谱信息中突发噪声对应的频段, 否则执行步骤 S209, 识别出当前帧频谱信息中不存在突发噪声, 并跳转至步骤 S206 搜索下一帧频谱信息中的各组频点。

[0098] 通过步骤 S206 能够搜索出有可能确定为突发噪声谱峰的各组频点, 然后通过步

骤 S207 进一步判断根据搜索出的各组频点能否确定为突发噪声谱峰,从而实现对当前帧频谱信息中突发噪声的检测。

[0099] 在实际实施时,所述突发噪声最大频带可以根据实际突发噪声的特点进行有针对性的设定。由于鸣笛声等突发噪声往往是窄带噪声,因此在搜索出一组频点后,还需要接着判断该组频点中的第二频点与第一频点之差是否小于预先设定的窄带噪声最大频带,即判断 $k_{up}-k_{dn} < M_k$,其中 M_k 为预先设定的窄带噪声最大频带所对应的值, M_k 的值还与采样率 F_s 以及 FFT 点数 NFFT 有关,比如设定突发噪声最大频带为 400Hz,在采样率 F_s 为 16kHz 和 NFFT 为 512 点时,则: $M_k = \text{int}\left(\frac{400}{F_s/NFFT}\right) = 12$ 。

[0100] 如果搜索到的任何一组 k_{up} 和 k_{dn} 均无法满足 $k_{up}-k_{dn} < M_k$,则说明当前帧频谱信息不存在突发噪声,如果搜索到一组或一组以上 k_{up} 和 k_{dn} 满足 $k_{up}-k_{dn} < M_k$,则满足该条件的各组 k_{up} 和 k_{dn} 即是该帧中突发噪声谱峰对应的频段,后续可以用于进行噪声抑制处理。

[0101] 突发噪声检测和突发噪声频谱搜索过程如图 3 所示,下面予以详细介绍。

[0102] 按频点序号由小到大的顺序,从 $k=f_{dn_k}$ 开始搜索,首先找到第一个满足 $P(k)=1$ 的频点,该频点即为搜索出的首个第一频点,通过“ $k_{dn}=k$ ”实现赋值后所述第一频点的频点序号为 k_{dn} ,然后从序号为 $k_{dn}+1$ 的频点开始搜索出第一个满足 $P(k)=-1$ 的频点,该频点即为搜索出的首个第二频点,通过“ $k_{up}=k$ ”实现赋值后所述第二频点的频点序号为 k_{up} ,如此就得到对应的第一组频点序号 k_{dn} 和 k_{up} 。

[0103] 需要说明的是,本实施例中,在每次搜索出一组频点后,还需要根据该组频点中的第一频点和第二频点进一步识别该组频点是否为能够确定突发噪声谱峰对应频段的一组频点,即判断该组频点中序号为 k_{up} 的第二频点与序号为 k_{dn} 的第一频点是否满足 $k_{up}-k_{dn} < M_k$,若是则表明该组频点是能够确定为突发噪声谱峰的一组 k_{up} 和 k_{dn} ,并搜索当前帧频谱信息中的下一组频点,否则直接开始搜索下一组频点。在其他实施例中,也可以在当前帧频谱信息中先搜索出各组频点,进而判断各组频点是否满足确定为突发噪声谱峰的条件。

[0104] 在找到第一组频点并判断是否满足确定为突发噪声谱峰的条件后,以上述同样的方式从序号为 $k_{up}+1$ 的频点开始继续搜索出第二个满足 $P(k)=1$ 的频点以及第二个满足 $P(k)=-1$ 的频点、第三个满足 $P(k)=1$ 的频点以及第三个满足 $P(k)=-1$ 的频点,……,以此类推,直至完成对当前帧频谱信息中序号为 N 的频点的搜索,并在搜索出之后一一判断是否满足确定为突发噪声谱峰的条件。

[0105] 在整个搜索过程中,通过“ $k=k+1$ ”实现频点序号 k 的不断递增,直至 $k=N$ (不再满足“ $k < N$ ”),即表示完成对当前帧频谱信息中所有有效频点的搜索。

[0106] 为了便于更为直观地理解上述确定突发噪声谱峰的过程,下面结合图 4 进行说明。

[0107] 图 4 示出的坐标轴中,横轴表示频谱信息中的各频点,频率沿横轴箭头方向由小到大依次递增,纵轴表示各频点对应的幅度谱梯度标志 $P(k)$,可以理解的是,将各频点对应的频谱梯度进一步转化为各自对应的幅度谱梯度标志后,各频点的幅度变化情况就更容易被体现出来了,考察幅度变化情况存在短时较大突变的频点将有可能寻找出突发噪声谱峰。

[0108] 如图 4 所示,沿横轴向右方向依次搜索,假设频点 B 是搜索过程中找到的第一个满足 $P(k)=1$ 的频点,频点 B 可以认为是幅度变化情况存在短时较大突变的一个起始频点,在频点 B 之前,会出现 $P(k)=0$ 的频点,例如频点 A,在频点 B 之后,通常还会出现 $P(k)=1$ 和 $P(k)=0$ 的频点,例如频点 C 和频点 D,由频点 B 开始构成频谱峰值的上升沿,假设频点 E 是找到的第一个满足 $P(k)=-1$ 的频点,频点 E 可以认为是幅度变化情况存在短时较大突变的一个终止频点,至频点 E 构成频谱峰值的下降沿。由此可知,频点 B、频点 E 及其中间的频点构成的频谱则可能为突发噪声谱峰。若频点 B 的频点序号为 k_{dn} ,频点 E 的频点序号为 k_{up} ,则判断 $k_{up}-k_{dn} < M_k$ 是否满足就可以确定频点 B、频点 E 及其中间的频点构成的频谱是否为突发噪声谱峰。当然,在频点 E 之后,需要按照上述方式继续去搜索第二个满足 $P(k)=1$ 的频点,在搜索过程中通常先会出现 $P(k)=-1$ 和 $P(k)=0$ 的频点,例如频点 F 和频点 G。

[0109] 对应于上述突发噪声的检测方法,本实施例还提供一种突发噪声的检测装置。如图 5 所示,所述突发噪声的检测装置 10 包括:频谱梯度计算单元 101,适于以声音信号的逐帧频谱信息中相邻频点的幅度谱计算获得各个频点对应的频谱梯度;第一确定单元 102,适于基于所述频谱梯度确定梯度阈值上下限;第二确定单元 103,适于根据所述频谱梯度以及梯度阈值上下限确定各个频点的幅度谱梯度标志,所述幅度谱梯度标志包括频谱梯度大于梯度阈值上限所对应的第一梯度标志和频谱梯度小于梯度阈值下限所对应的第二梯度标志;搜索单元 104,对于每帧频谱信息,按频点顺序从各个频点中交替搜索出一个对应幅度谱梯度标志为所述第一梯度标志的第一频点和一个对应幅度谱梯度标志为所述第二梯度标志的第二频点作为一组频点;识别单元 105,适于在当前帧频谱信息中存在至少一组频点的所述第二频点与第一频点之差小于预设的突发噪声最大频带时,将该组频点的第一频点与第二频点识别为该帧频谱信息中突发噪声对应的频段。

[0110] 本实施例中,所述搜索单元 104 从大于频率阈值的频点开始搜索,所述频率阈值对应于语音信号的频带的最高频率。

[0111] 所述突发噪声的检测装置的具体实施可以参考本实施例所述突发噪声的检测方法的实施,此处不再赘述。

[0112] 实施例二

[0113] 在实施例一的基础上,本实施例对所述突发噪声的抑制方法的具体实施进行说明。

[0114] 以实施例一提供的突发噪声的检测方法对声音信号的逐帧频谱信息进行检测后,对于检测出存在突发噪声的各帧频谱信息,将其中识别为该帧频谱信息中突发噪声对应的第一频点和第二频点之间的所有频点对应的幅度谱进行噪声抑制处理。

[0115] 具体地,对于各帧频谱信息,如果检测出不存在突发噪声,则不对声音信号进行处理,如果检测出存在突发噪声,则对搜索得到的一组或一组以上满足 $k_{up}-k_{dn} < M_k$ 的 k_{up} 和 k_{dn} 进行处理。本实施例中,所述噪声抑制处理通过对突发噪声及其附近频带给一较小增益,得到处理后幅度谱 $|Y(n, k)|$ 如下:

[0116]

$$|Y(n, k)| = \begin{cases} \alpha * |X(n, k)| & kdn - rn \leq k \leq kup + rn \\ |X(n, k)| & \text{其他} \end{cases} \quad (6)$$

[0117] 其中 n 为频谱信息的帧号, k 为频点序号, $|Y(n, k)|$ 为帧号 n 的频谱信息中序号为 k 的频点的经过所述噪声抑制处理后的幅度谱, α 为一较小值, α 的取值范围为 $[0, 1)$, 其典型值为 0, $|X(n, k)|$ 为帧号 n 的频谱信息中序号为 k 的频点的幅度谱, kdn 和 kup 分别为帧号 n 的频谱信息中识别为突发噪声对应的频段的第一频点和第二频点的序号, rn 为检测出的突发噪声附近频段对应的值, rn 的取值范围为 $0 \sim 10$, 典型值为 4, 该值与采样率 F_s 、FFT 点数 $NFFT$ 有关。

[0118] 本实施例中, 对于经过上述噪声抑制处理后的频谱信息, 可以进一步通过短时傅立叶逆变换重新得到时域声音信号。具体地, 使用未处理声音信号的相位谱 $\angle X(n, k)$ 作为处理后声音信号的相位谱, 并将该相位谱与经过所述噪声抑制处理之后声音信号的幅度谱进行计算, 得到处理后声音信号的逐帧频谱信息 $Y(n, k) = |Y(n, k)| * \exp[i * \angle X(n, k)]$, 接着通过短时傅里叶逆变换得到逐帧的时域声音信号, 重叠相加得到处理后的声音信号。

[0119] 在实际实施时, 实施例一的突发噪声的检测方法以及实施例二的突发噪声的抑制方法可以单独使用, 也可以和其他处理过程结合使用, 例如可以单独对输入声音信号进行处理, 也可以作为声音信号处理的其中一个环节。此外, 所述突发噪声的检测方法以及突发噪声的抑制方法可以在其他语音降噪技术之前作为前处理, 也可以在其他语音降噪技术之后作为后处理, 当作为前处理时, 对各个通道的声音信号分别进行处理, 处理之后的各个通道信号作为其他语音降噪技术的输入信号, 当作为后处理时, 对其他语音降噪技术输出的增强后信号进行处理。

[0120] 对应于上述突发噪声的抑制方法, 本实施例还提供一种突发噪声的抑制装置, 所述突发噪声的抑制装置包括: 第一处理单元, 适于在上述突发噪声的检测装置检测出存在突发噪声的各帧频谱信息后, 将其中识别为该帧频谱信息中突发噪声对应的频段的第一频点和第二频点之间的所有频点对应的幅度谱进行突发噪声抑制处理。

[0121] 所述突发噪声的抑制装置的具体实施可以参考本实施例所述突发噪声的抑制方法的实施, 此处不再赘述。

[0122] 实施例三

[0123] 在实施例一提供的突发噪声的检测装置以及实施例二提供的突发噪声的抑制装置, 本实施例提供一种突发噪声的处理系统。

[0124] 如图 6 所示, 本实施例提供的突发噪声的处理系统包括: 实施例一提供的突发噪声的检测装置 10 以及实施例二提供的突发噪声的抑制装置 20, 所述突发噪声的抑制装置 20 包括第一处理单元 201。

[0125] 本实施例中, 所述突发噪声的处理系统还包括第一变换单元 30, 适于将输入的时域声音信号进行离散傅立叶变换以得到所述声音信号的逐帧频谱信息。

[0126] 所述突发噪声的处理系统还包括第二处理单元 40, 适于以所述突发噪声抑制处理之前声音信号的相位谱作为所述突发噪声抑制处理之后声音信号的相位谱, 并将其与经过所述突发噪声抑制处理之后声音信号的幅度谱进行计算, 得到经过所述突发噪声抑制处理后声音信号的逐帧频谱信息。

[0127] 所述突发噪声的处理系统还包括第二变换单元 50，适于将经过所述突发噪声抑制处理后声音信号的逐帧频谱信息进行离散傅立叶逆变换以得到处理后的时域声音信号。

[0128] 需要说明的是，在实际实施时，可以将所述第一变换单元 30、第二处理单元 40、第二变换单元 50 与所述突发噪声的检测装置 10 和突发噪声的抑制装置 20 集成在一起作为单独模块使用，即输入和输出都为时域信号，有独立的短时傅立叶变换和逆变换，也可以将所述突发噪声的检测装置 10 和突发噪声的抑制装置 20 作为一个模块集成在其他语音信号处理模块之内，此时输入和输出都为短时频谱信息。

[0129] 本领域技术人员可以理解，实现上述实施例中突发噪声的处理系统及其突发噪声的抑制装置的全部或部分是可以通过程序来指令相关的硬件来完成，所述的程序可以存储于计算机可读存储介质中，所述存储介质可以是 ROM、RAM、磁碟、光盘等。

[0130] 虽然本发明披露如上，但本发明并非限定于此。任何本领域技术人员，在不脱离本发明的精神和范围内，均可作各种更动与修改，因此本发明的保护范围应当以权利要求所限定的范围为准。

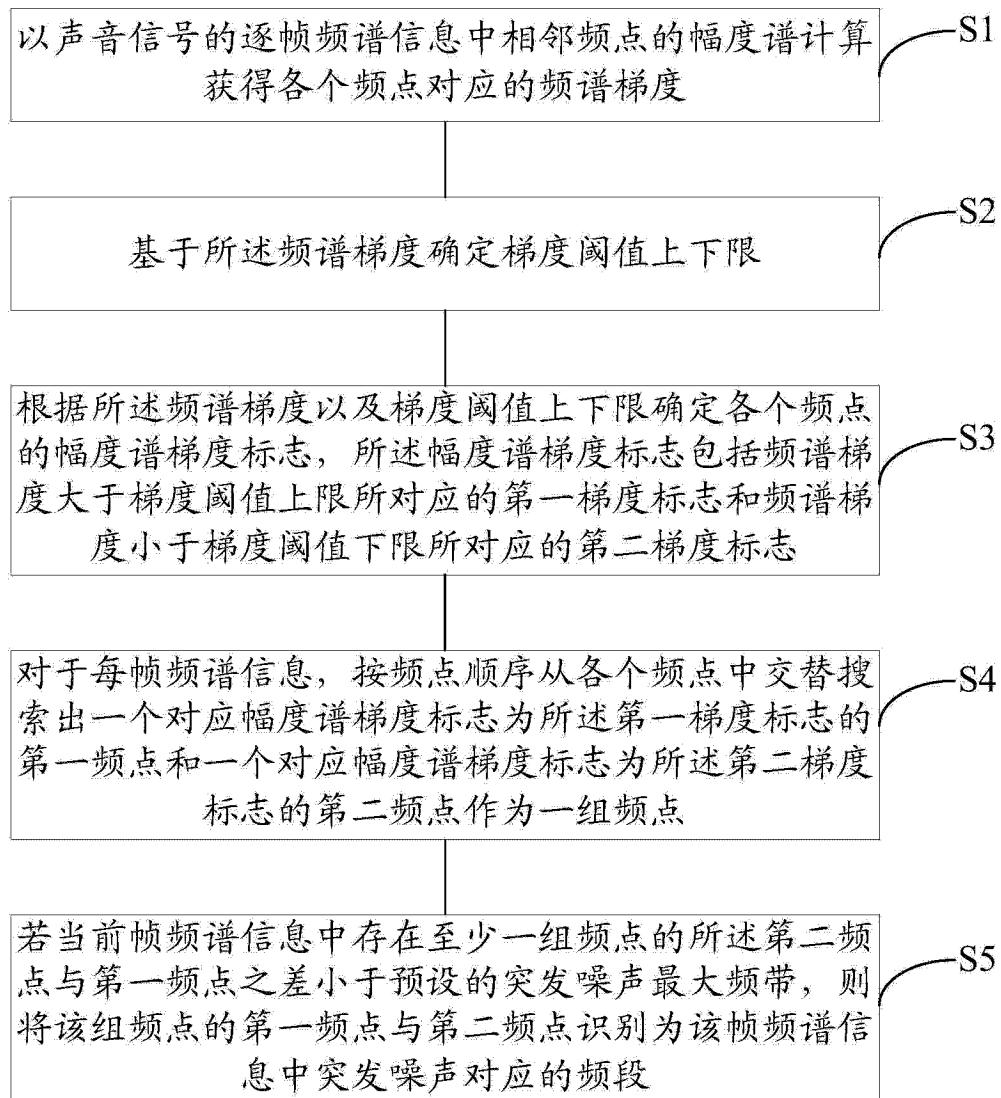


图 1

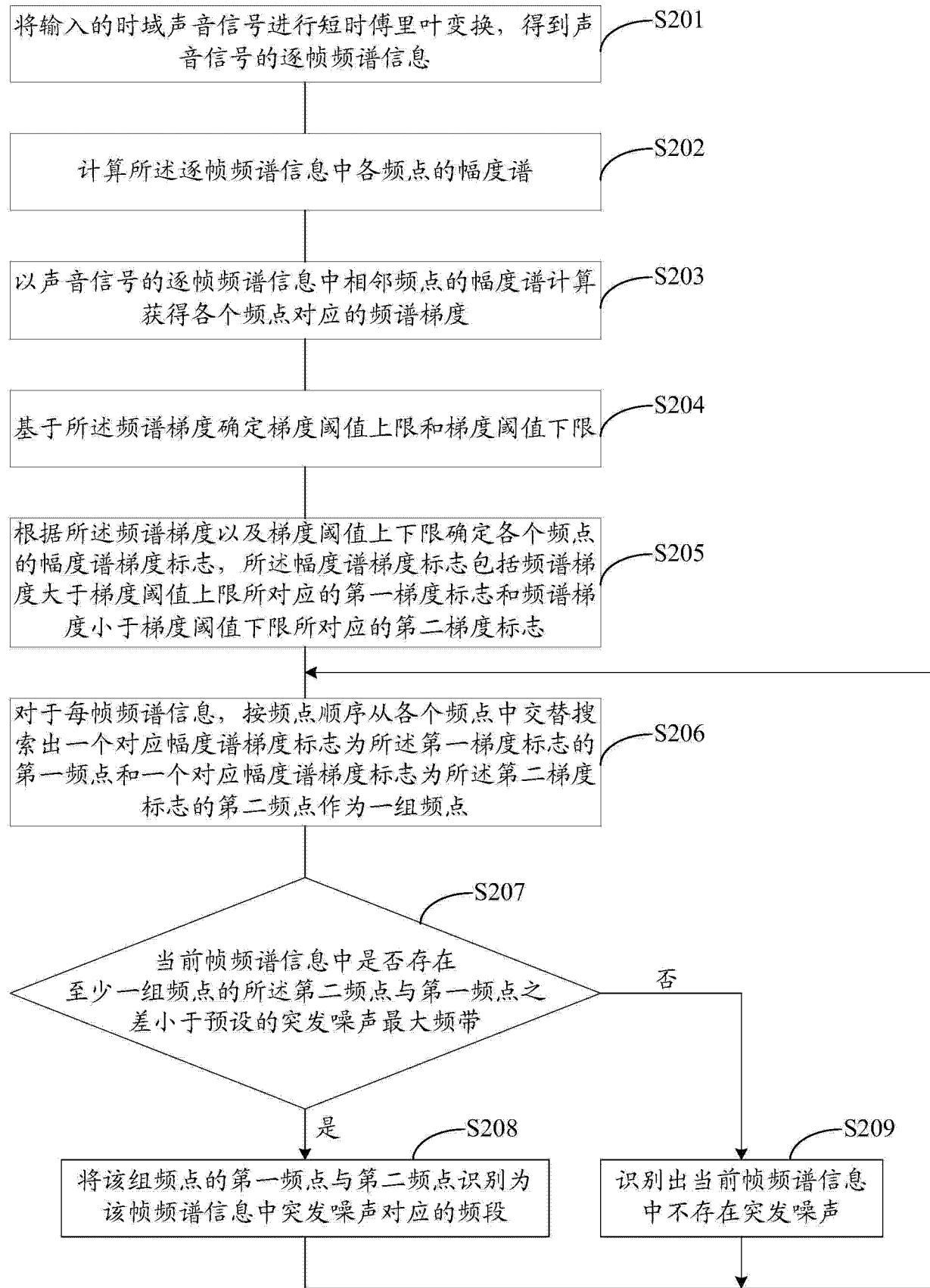


图 2

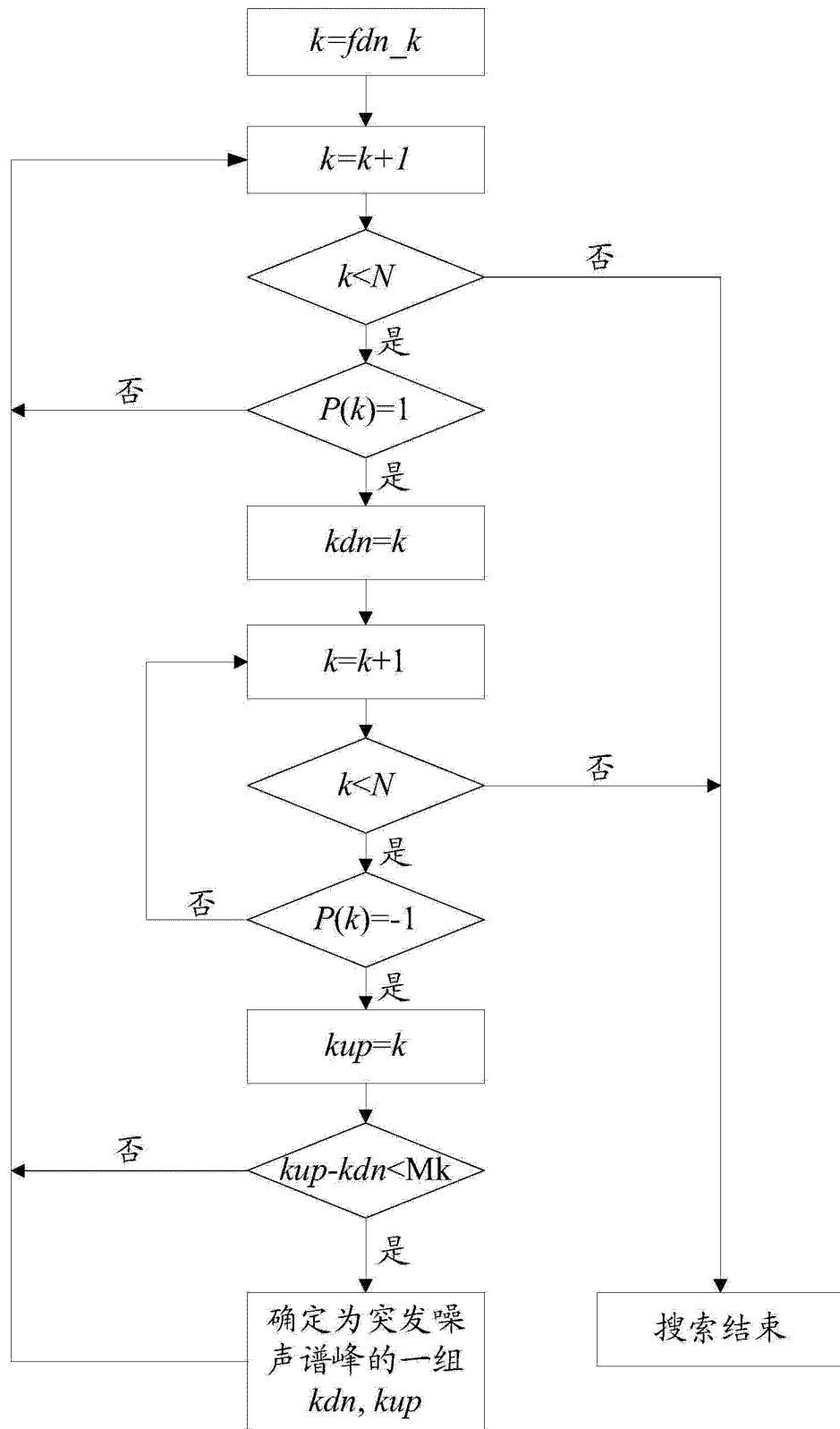


图 3

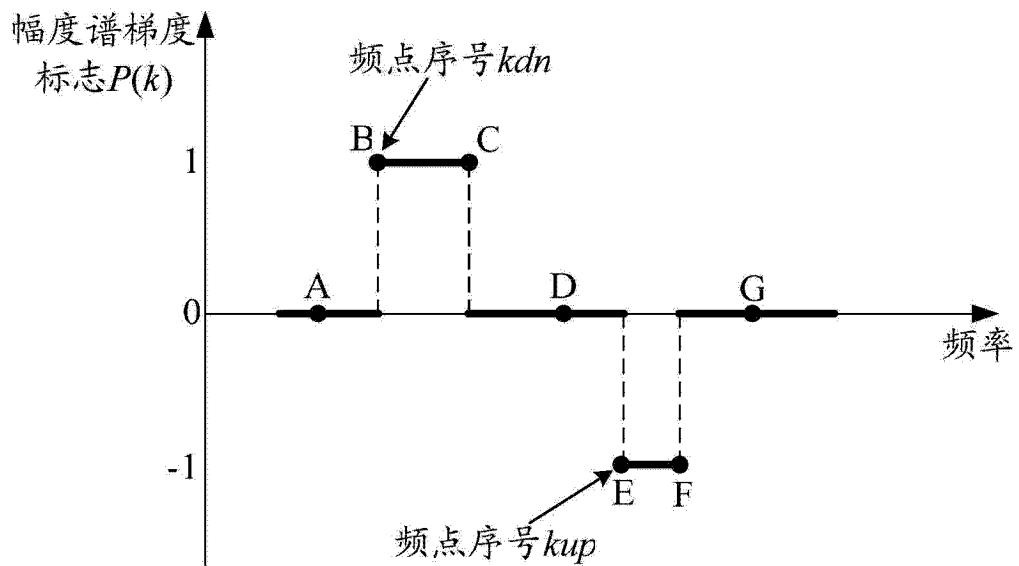


图 4

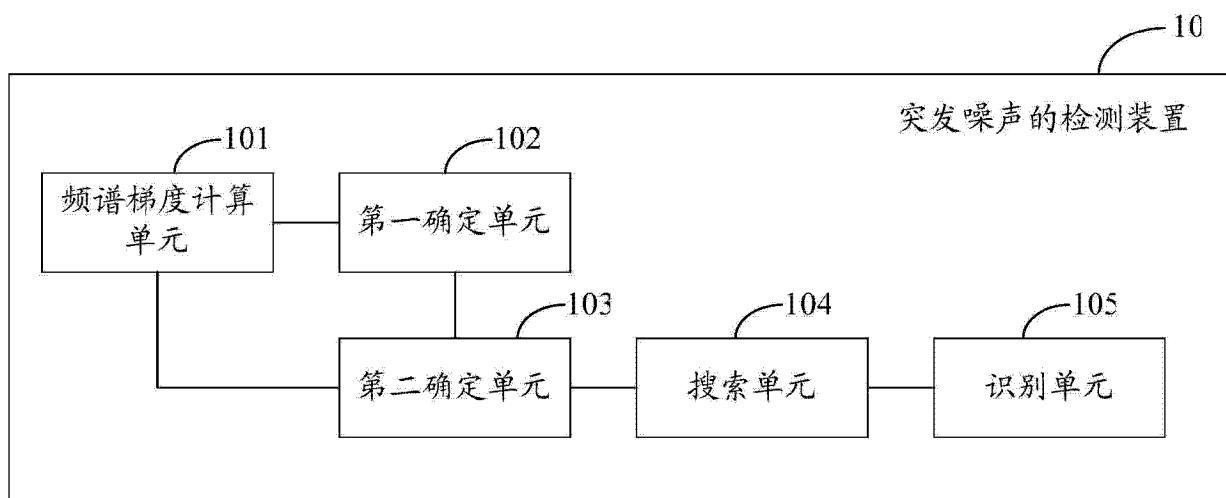


图 5

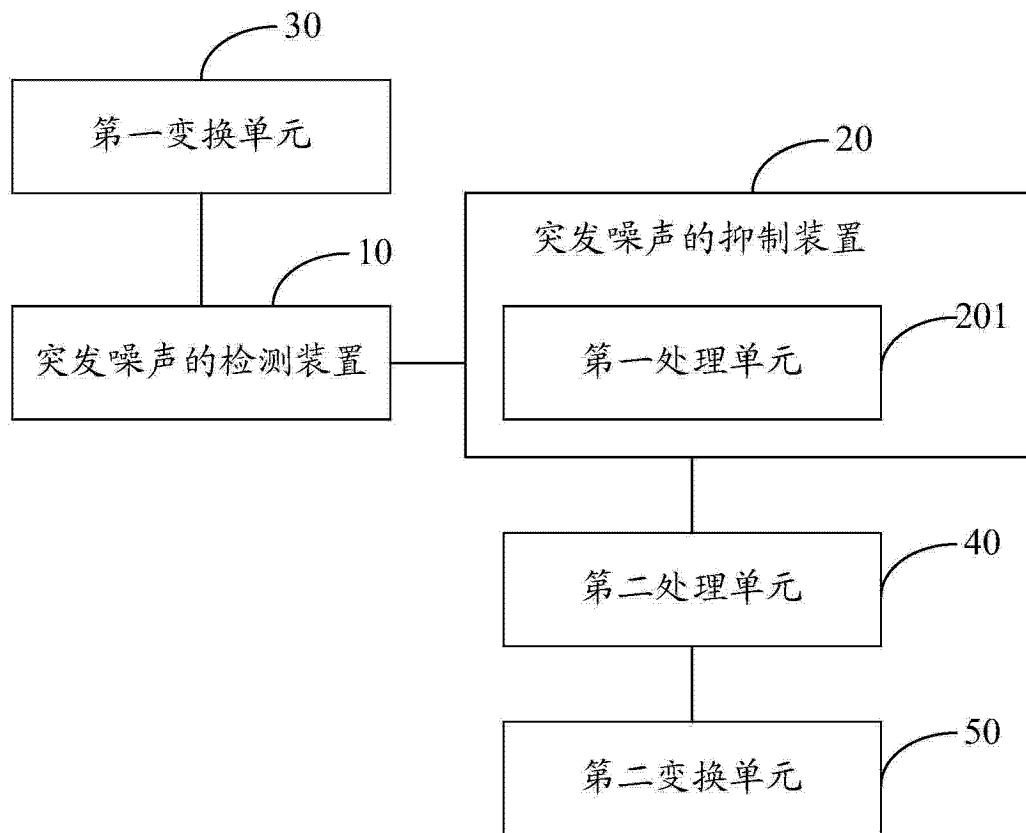


图 6