(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 1909060 B
(45)授权公告日 2012.01.25

(21)申请号 200610108332.7
(22)申请日 2006.08.01
(30)优先权数据
10-2005-0070410 2005.08.01 KR
(73)专利权人 三星电子株式会社
地址 韩国京畿道水原市灵通区梅滩3洞416
(72)发明人 金炫秀
(74)专利代理机构 北京铭硕知识产权代理有限公司 11286
代理人 郭鸿禧 李云霞
(51)Int.Cl.
G10L 11/06(2006.01)
(56)对比文件
AHN R ET AL.HARMONIC-PLUS-NOISE DECOMPOSITION AND ITS APPLICATION IN VOICED/UNVOICED CLASSIFICATION.《IEEE》.1997,587-590.

审查员 张亚峰

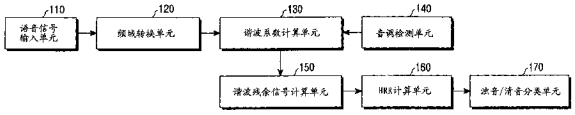
权利要求书 2 页 说明书 7 页 附图 5 页

(54)发明名称

提取浊音/清音分类信息的方法和设备

(57)摘要

本发明公开了一种从语音信号中提取精确的浊音/清音分类信息的设备和方法。该设备通过分析谐波分量对非谐波分量(或残余)的比率来提取浊音/清音分类信息。该设备使用作为基于谐波分量分析获得的特征提取方案的谐波对残余比率(HRR)、谐波对噪声比率(HNR)和子频带谐波对噪声比率(SB-HNR),从而精确地对浊音/清音进行分类。因此,在所有语音信号处理系统中,该设备和方法可用于语音编码、识别、合成、增强等。



CN 1909060 B

1. 一种使用语音信号的谐波分量提取浊音 / 清音分类信息的方法，该方法包括以下步骤：

将输入的语音信号转换为频域的语音信号；

计算有关的谐波系数，并提供计算谐波系数所需的音调；

从转换的语音信号中计算谐波信号和除了谐波信号之外的残余信号；

使用谐波信号和残余信号的计算结果来计算谐波对残余比率 HRR；以及

通过比较 HRR 和阈值对浊音 / 清音进行分类，

其中，计算有关的谐波系数，从而最小化使用谐波模型表示的语音信号中的残余信号的能量，所述谐波模型被表示为基本频率的谐波和小的残余的和；

使用计算的谐波系数来获得谐波信号，并当已经获得谐波信号时，通过从转换的语音信号中减去谐波信号来计算残余信号。

2. 如权利要求 1 所述的方法，其中，将转换的语音信号表示为：

$$S_n = a_0 + \sum_{k=1}^{L} (a_k \cos n\omega_0 k + b_k \sin n\omega_0 k) + r_n \quad (n = 0, 1, \ldots N-1)$$

$$= h_n + r_n$$

其中，“S_n”表示转换的语音信号，“r_n”表示残余信号，“h_n”表示谐波信号，“N”表示帧长度，“L”表示现有谐波的数量，“ω_0”表示音调，k 是频率槽标号，“a”和“b”是根据帧而具有不同常数值的有关的谐波系数。

3. 如权利要求 1 所述的方法，其中，以与最小平方方案相同的方式来计算谐波系数。

4. 如权利要求 1 所述的方法，其中，将残余能量表示为：

$$E = \sum_{n=0}^{N-1} r_n^2$$

。

5. 如权利要求 4 所述的方法，其中，在计算有关的谐波系数的步骤中，对残余能量关于式中的每一个“k”计算“$\partial E / \partial a_k = 0$”和“$\partial E / \partial b_k = 0$”。

6. 如权利要求 1 所述的方法，其中，计算 HRR 的步骤包括：

使用计算的谐波信号和残余信号获得谐波能量；

通过从语音信号的整个能量中减去谐波能量来计算残余能量；以及

计算计算出的谐波能量对计算出的残余能量的比率。

7. 如权利要求 1 所述的方法，其中，将 HRR 表示为：

$$HRR = 10\log_{10}(\sum h_n^2 / \sum r_n^2) dB$$

。

8. 如权利要求 1 所述的方法，其中，当使用 Parseval 定律时，将 HRR 在频域中表示为

$$HRR = 10\log_{10}\left(\sum_k |H(\omega_k)|^2 / \sum_k |R(\omega_k)|^2\right) dB$$

其中，H 表示谐波分量 h_n，R 表示残余信号 r_n，“ω”表示频率槽。

9. 如权利要求 1 所述的方法，其中，在通过比较 HRR 和阈值对浊音 / 清音进行分类的步骤中，当语音信号的 HRR 大于阈值时，将语音信号确定和分类为浊音。

10. 一种使用语音信号的谐波分量提取浊音 / 清音分类信息的设备，该设备包括：

语音信号输入单元，接收语音信号；

频域转换单元，将接收的时域语音信号转换为频域语音信号；

谐波系数计算单元，计算有关的谐波系数；

音调检测单元，提供计算谐波系数所需的音调；

谐波残余信号计算单元，从转换的语音信号中计算谐波信号和除了谐波信号之外的残余信号，其中，使用计算的谐波系数来获得谐波信号，并当已经获得谐波信号时，通过从转换的语音信号中减去谐波信号来计算残余信号；以及

谐波对残余比率 HRR 计算单元，通过使用谐波残余信号计算单元的计算结果计算谐波信号对残余信号的能量比率，

其中，所述谐波系数计算单元计算有关的谐波系数，从而最小化使用谐波模型表示的语音信号中的残余信号的能量，所述谐波模型被表示为基本频率的谐波和小的残余的和。

11. 如权利要求 10 所述的设备，其中，将 HRR 表示为

$$HRR = 10\log_{10}\left(\sum h_n^2 / \sum r_n^2\right) dB$$

其中，“h_n”表示谐波信号，“r_n”表示残余信号。

提取浊音／清音分类信息的方法和设备

技术领域

[0001] 本发明涉及一种用于提取浊音／清音分类信息的方法和设备，更具体地说，涉及一种使用语音信号的谐波分量提取浊音／清音分类信息从而将语音信号精确地分类为浊音／清音的方法和设备。

背景技术

[0002] 通常，根据其时域和频域统计特性，语音信号被分类为周期（或谐波）分量和非周期（或随机）分量（即浊音和由除了语音之外的声音或噪声产生的声音，下文中称为“清音”），从而语音信号被称为“准周期”信号。在此情况下，根据是否存在音调 (pitch) 信息来将周期分量和非周期分量确定为浊音 (voiced sound) 和清音 (unvoiced sound)，浊音具有周期特性，而清音具有非周期特性。

[0003] 如上所述，在所有语音信号处理系统中，浊音／清音分类信息是用于编码、识别、合成、增强等的最基本和关键的信息。因此，已经提出将语音信号分类为浊音／清音的各种方法。例如，一种方法是语音学编码，其中，将语音信号分类为六种类型，包括开端 (onset)、全频带稳定状态浊音、全频带瞬态浊音、低通瞬态浊音以及低通稳定状态的浊音和清音。

[0004] 具体地说，用于浊音／清音分类的特征包括低频带话音能量、过零计数、第一反射系数、预强调 (pre-emphasize) 的能量比率、第二反射系数、随意音调预测增益以及非随意音调预测增益，组合这些特征并将其用于线性鉴别器中。然而，由于还没有一种仅使用一个特征的浊音／清音分类方法，因此浊音／清音分类的性能很大程度上受到如何组合所述多个特征的影响。

[0005] 同时，在浊音期间，由于较高的功率是由发声系统（即产生语音信号的系统）输出的，因此浊音占据了语音能量的较大部分，从而语音信号中的浊音部分的失真对编码的话音的整个声音质量产生较大影响。

[0006] 在这样的浊音的话音中，由于在声门激励和声道之间的交互导致难于进行频谱估计，因此在多数语音信号处理系统中需要关于浊音的程度的测量信息。这种测量信息同样用于语音识别和语音编码。具体地说，由于测量信息是确定语音合成中的声音质量的重要参数，因此使用错误信息或错误估计值导致语音识别和合成中的性能恶化。

[0007] 然而，由于估计的现象本身按照其特性而包括一定程度的随机性，因此在预定的周期中执行这样的估计，浊音测量的输出包括随机分量。因此，可对浊音测量的估计适当地使用统计性能测量方案，并可将使用大量帧估计的混合的平均用作主索引（指示符）。

[0008] 如上所述，虽然在现有技术中使用多种特征来提取浊音／清音分类信息，但不能够通过单个特征对浊音／清音声音分类。因此，通过使用特征的组合对浊音／清音声音分类，所述特征中的任意一个特征自身不能提供可靠的信息。然而，传统方法存在所述特征之间的相关性问题以及由于噪声导致的性能恶化问题，因此需要一种能够解决这些问题的新方法。此外，传统技术没有正确地表示浊音和清音之间的实质性差别的谐波分量的存在和谐波分量的程度。因此，需要开发一种能够通过分析谐波分量对浊音／清音声音精确分类

的新的方法。

发明内容

[0009] 因此，实现本发明来满足上述要求，本发明提供一种用于通过使用语音信号的谐波分量分析从而更精确地对浊音 / 清音声音分类的提取浊音 / 清音分类信息的方法和设备。

[0010] 为此，本发明提供一种使用语音信号的谐波分量提取浊音 / 清音分类信息的方法，该方法包括：将输入的语音信号转换为频域的语音信号；从转换的语音信号中计算谐波信号和除了谐波信号之外的残余信号；使用谐波信号和残余信号的计算结果来计算谐波对残余比率 (HRR)；以及通过比较用 HRR 和阈值对浊音 / 清音进行分类。

[0011] 此外，本发明提供一种使用语音信号的谐波分量提取浊音 / 清音分类信息的方法，该方法包括：将输入的语音信号转换为频域的语音信号；从转换的语音信号中分离谐波部分和噪声部分；计算谐波部分对噪声部分的能量比率；以及使用计算结果对浊音 / 清音进行分类。

[0012] 此外，本发明提供一种使用语音信号的谐波分量提取浊音 / 清音分类信息的设备，该设备包括：语音信号输入单元，接收语音信号；频域转换单元，将接收的时域语音信号转换为频域语音信号；谐波残余信号计算单元，从转换的语音信号中计算谐波信号和除了谐波信号之外的残余信号；以及谐波对残余比率计算单元 (HRR)，通过使用谐波残余信号计算单元的计算结果来计算谐波信号对残余信号的能量比率。

[0013] 此外，本发明提供一种使用语音信号的谐波分量提取浊音 / 清音分类信息的设备，该设备包括：语音信号输入单元，接收语音信号；频域转换单元，将接收的时域语音信号转换为频域语音信号；谐波 / 噪声分离单元，从转换的语音信号中分离谐波部分和噪声部分；以及谐波对噪声能量比率计算单元，计算谐波部分对噪声部分的能量比率。

附图说明

[0014] 通过下面结合附图的详细描述，本发明的上述和其他目的、特点和优点将会变得更加清楚，其中：

[0015] 图 1 是示出根据本发明第一实施例的浊音 / 清音分类信息提取设备的结构的框图；

[0016] 图 2 是示出根据本发明第一实施例的提取浊音 / 清音分类信息的过程的流程图；

[0017] 图 3 是示出根据本发明第二实施例的浊音 / 清音分类信息提取设备的结构的框图；

[0018] 图 4 是示出根据本发明第二实施例的提取浊音 / 清音分类信息的过程的流程图；

[0019] 图 5 是示出根据本发明第二实施例的频域语音信号的示图；

[0020] 图 6 是示出根据本发明第二实施例的在分解之前的原始语音信号的波形的示图；

[0021] 图 7A 是示出根据本发明第二实施例的分解的谐波信号的示图；以及

[0022] 图 7B 是示出根据本发明第二实施例的分解的噪声信号的示图。

具体实施方式

[0023] 下文中，将参照附图描述本发明优选实施例。在以下本发明实施例的描述中，当合并到此的公知功能和配置可能使本发明的主旨模糊时，将省略对其的描述。

[0024] 本发明实现能够改善从语音信号中提取浊音 / 清音分类信息的精度的功能。为此，根据本发明，通过使用谐波对非谐波（或残余）分量比率的分析来提取浊音 / 清音分类信息。详细地说，可通过作为基于谐波分量分析获得的特征提取方法的谐波对残余比率 (HRR)、谐波对噪声分量比率 (HNR) 和子频带谐波对噪声分量比率 (SB-HNR) 精确地对浊音 / 清音声音分类。由于通过这些方案获得浊音 / 清音分类信息，因此当在所有语音信号处理系统中执行语音编码、识别、合成和增强时，可使用获得的浊音 / 清音分类信息。

[0025] 本发明测量语音或音频信号的谐波分量的强度，从而数字化地表示浊音 / 清音分类信息提取的必要属性。

[0026] 在描述本发明之前，将描述影响语音估计器的性能的因素。

[0027] 详细地说，这些因素包括对语音合成的灵敏度、对音调 (pitch) 性能的非灵敏度 (insensitivity)（例如音调为高还是低，音调是否平滑地改变、在音调间隔是否有随机性等）、对谱包络的非灵敏度、主观性能等。实际上，由于听觉系统对对语音强度的小的改变很不敏感，因此可能导致语音测量中的微小误差，但性能测量中的最重要的指标是通过聆听的主观性能。

[0028] 本发明提供一种能够仅使用单个特征不是多个不可靠的特征的组合找出浊音 / 清音分类信息（即特征）以分类浊音 / 清音声音，同时满足上述指标的分类信息提取方法。

[0029] 将描述在其中实现上述功能的浊音 / 清音分类信息提取设备的组件及其操作。为此，将参照图 1 所示的框图描述根据本发明第一实施例的浊音 / 清音分类信息提取设备。下文中，根据在本发明第一实施例中公开的结构，将整个语音信号表示为话音的谐波正弦模型，从语音信号中获得谐波系数，使用获得的谐波系数计算谐波信号和残余信号，从而获得谐波信号和残余信号之间的能量比率。在此情况下，将谐波信号和残余信号之间的能量比率定义为谐波对残余比率 (HRR)，并可通过使用 HRR 对浊音 / 清音声音分类。

[0030] 参照图 1，根据本发明第一实施例的浊音 / 清音分类信息提取设备包括语音信号输入单元 110、频域转换单元 120、谐波系数计算单元 130、音调检测单元 140、谐波残余信号计算单元 150、HRR 计算单元 160 和浊音 / 清音分类单元 170。

[0031] 首先，语音信号输入单元 110 可包括麦克风 (MIC)，并接收包括语音和声音信号的语音信号。频域转换单元 120 将输入信号从时域转换到频域。

[0032] 频域转换单元 120 使用快速傅里叶变换 (FFT) 等，从而将时域语音信号转换为频域语音信号。

[0033] 其后，当频域转换单元 120 输出信号（即整个语音信号）时，可将整个语音信号表示为话音 (speech) 的谐波正弦模型。这样能够仅以少量的计算进行高效和准确的谐波性测量。详细地说，通过使用将语音信号表示为基本频率的谐波和小的残余的和的语音信号的谐波模型，可如式 1 所示表示语音信号。也就是说，由于可将语音信号表示为余弦和正弦的组合，因此可如式 1 所示表示语音信号。

[0034]
$$\begin{aligned} S_n &= a_0 + \sum_{k=1}^{L}(a_k \cos n\omega_0 k + b_k \sin n\omega_0 k) + r_n \quad (n=0,1,\ldots N-1) \\ &= h_n + r_n \end{aligned} \qquad \ldots\ldots(1)$$

[0035] 在式 1 中，“($\alpha_k \cos n\omega_0 k + b_k + \sin n\omega_0 k$)”与谐波部分对应，“$r_n$”与除了谐波部分之外的残余部分对应。这里，“$S_n$”表示转换的语音信号，“$r_n$”表示残余信号，“$h_n$”表示谐波分量，“N”表示帧长度、“L”表示存在的谐波的数量，“ω_0”表示音调，“k”表示频率槽(frequency bin)标号，“a”和“b”表示根据帧而具有不同值的常数。在此情况下，为了使残余信号最小，执行使式 1 中的“r_n”的值最小的过程。谐波系数计算单元 130 从音调检测单元 140 接收音调值，从而将与“ω_0”对应的音调值代入到式 1 中。当接收如上所述的音调时，谐波系数计算单元 130 可根据下述方式获得使残余能量最小的“a”和“b”的值。

[0036] 首先，当关于残余部分“r_n”重新排列式 1 时，“$r_n = S_n - h_n$”，并且 $h_n = a_0 + \sum_{k=1}^{L}(a_k \cos n\omega_0 k + b_k \sin n\omega_0 k)$。同时，可将残余能量表示为式 2。

[0037] $$E = \sum_{n=0}^{N-1} r_n^2 \quad \cdots\cdots(2)$$

[0038] 这里，为了使残余能量最小，关于每一个“k”计算“$\ominus E / \ominus a_k = 0$”和“$\ominus E / \ominus b_k = 0$”。

[0039] 以与最小平方方法相同的方式获得谐波系数“a”和“b”，因为仅需少量计算，所以所述最小平方方法确保残余能量最小同时效率高。

[0040] 谐波残余信号计算单元 150 通过上述过程获得谐波系数“a”和“b”以使残余能量最小。其后，谐波残余信号计算单元 150 通过使用获得的谐波系数计算谐波信号和残余信号。详细地说，谐波残余信号计算单元 150 将计算的谐波系数和音调带入式 $h_n = a_0 + \sum_{k=1}^{L}(a_k \cos n\omega_0 k + b_k \sin n\omega_0 k)$，从而获得谐波信号。由于在获得谐波信号之后通过从转换的整个语音信号“S_n”中减去谐波信号“h_n”来计算残余信号“r_n”，因此能够计算谐波信号和残余信号。相似地，可通过从语音信号的能量中减去谐波能量的简单方式来计算残余能量。这里，残余信号是类噪声，并且在浊音的帧的情况下非常小。

[0041] 当将按上述方式获得的谐波信号和残余信号提供给 HRR 计算单元 160 时，HRR 计算单元 160 获得 HRR，所述 HRR 表示谐波对残余能量的比率。可如式 3 定义 HRR。

[0042] $$HRR = 10\log_{10}(\sum h_n^2 / \sum r_n^2) dB \quad \cdots\cdots(3)$$

[0043] 当采用 Parseval 定律时，可将式 3 在频域中表示为式 4。

[0044] $$HRR = 10\log_{10}\left(\sum_k |H(\omega_k)|^2 / \sum_k |R(\omega_k)|^2\right) dB \quad \cdots\cdots(4)$$

[0045] 在式 4 中，“ω”表示频率槽，H 表示谐波分量 h_n，R 表示残余信号 r_n。

[0046] 使用这样的测量以提取分类信息（即特征），所述分类信息表示每一帧中的信号的浊音分量的程度。通过这样的过程得到 HRR 的操作获得用于对浊音 / 清音声音分类的分类信息。

[0047] 在此情况下，采用统计分析方案以对浊音 / 清音声音分类。例如，当采用柱状图分析时，使用 95％的阈值。在此情况下，当 HRR 大于作为阈值的 2.65dB 时，可将对应的信号确定作为浊音。反之，当 HRR 小于 -2.65dB 时，可将对应的信号确定作为清音。因此，浊音 / 清音计算单元 170 通过比较获得的 HRR 和阈值来执行浊音 / 清音分类操作。

[0048] 下文中，将参照图 2 描述根据本发明第一实施例的提取浊音 / 清音分类信息的过

程。

[0049] 在步骤200，浊音/清音分类信息提取设备通过麦克风等接收语音信号。在步骤210，浊音/清音分类信息提取设备通过使用FFT等将接收的语音信号从时域转换到频域。其后，在步骤220，浊音/清音分类信息提取设备将语音信号表示为话音的谐波正弦模型，并计算对应的谐波系数。在步骤230，浊音/清音分类信息提取设备使用计算的谐波系数计算谐波信号和残余信号。在步骤240，浊音/清音分类信息提取设备通过使用步骤230的计算结果来计算谐波对残余比率(HRR)。在步骤250，浊音/清音分类信息提取设备通过使用HRR对浊音/清音声音分类。换句话说，基于谐波和非谐波（即残余）分量比率的分析来提取浊音/清音分类信息，并使用提取的浊音/清音分类信息对浊音/清音声音分类。

[0050] 根据上述本发明第一实施例，通过分析谐波区域来获得谐波和噪声之间的能量比率，所述谐波区域通常存在于比噪声区域高的电平处，从而提取使用语音和音频信号的所有系统中所需的浊音/清音分类信息。

[0051] 下文中，将描述根据本发明第二实施例的提取浊音/清音分类信息的设备和方法。

[0052] 图3是示出根据本发明第二实施例的提取浊音/清音分类信息的设备的结构的框图。

[0053] 根据本发明第二实施例的浊音/清音分类信息提取设备包括语音信号输入单元310、频域转换单元320、谐波/噪声分离单元330、谐波对噪声能量比率计算单元340和浊音/清音分类单元350。

[0054] 首先，语音信号输入单元310可包括麦克风(MIC)，并接收包括语音和声音信号的语音信号。频域转换单元320优选地使用快速傅里叶变换(FFT)等将输入信号从时域转换到频域，从而将时域语音信号转换为频域语音信号。

[0055] 谐波/噪声分离单元330从语音信号中将频域分离为谐波部分和噪声部分。在此情况下，谐波/噪声分离单元330使用音调信息从而执行分离操作。

[0056] 将参照图5更详细地描述从语音信号中分离谐波部分和噪声部分的操作。图5是示出根据本发明第二实施例的频域语音信号的示图。如图5所示，当语音信号经受谐波加噪声分解时(HND)，可将频域语音信号分离为噪声（随机）部分“B”和谐波（确知）部分“A”。由于HND是广泛公知的方案，因此将省略对其的详细描述。

[0057] 在整个HND中，将如图6所示的原始语音信号波形分离为分别如图7A和7B所示的谐波信号和噪声信号。根据本发明第二实施例，图6是示出在分解之前的原始语音信号的波形的示图。图7A是示出分解的谐波信号的示图，图7B是示出分解的噪声信号的示图。

[0058] 当输出如图7A和图7B所示的分解的信号时，谐波对噪声能量比率计算单元150计算谐波对噪声能量比率。基于整个谐波和噪声部分，可将整个谐波部分对整个噪声部分的比率定义为谐波对噪声比率(HNR)。以不同的方式根据每一预定的频带来划分谐波的整个部分和噪声部分，并可将每一频带的谐波部分对噪声部分的能量比率定义为子频带谐波对噪声比率(SB-HNR)。当谐波对噪声能量比率计算单元340已经计算HNR和SB-HNR时，浊音/清音分类单元350接收计算的HNR和SB-HNR，并执行浊音/清音分类操作。

[0059] 可如式5来定义谐波部分对噪声部分的信号能量比率HNR。将以此方式获得的HNR提供给浊音/清音分类单元350。其后，浊音/清音分类单元350通过比较接收的HNR和阈

值来执行浊音 / 清音分类操作。

[0060] $$HNR=10\log_{10}\left(\sum_k|H(\omega_k)|^2/\sum_k|N(\omega_k)|^2\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5)$$

[0061] 参照图 7A 和 7B，如式 5 所定义的 HNR 与通过将图 7A 所示的波形的较低区域除以图 7B 所示的波形的较低区域而获得的值对应。也就是说，图 7A 和 7B 所示的波形的较低区域表示能量。

[0062] 现将参照图 4 的流程图描述根据本发明第二实施例的提取浊音 / 清音分类信息的方法。在步骤 400，浊音 / 清音分类信息提取设备通过麦克风等接收语音信号。在步骤 410，浊音 / 清音分类信息提取设备通过使用 FFT 等将接收的时域语音信号转换为频域语音信号。在步骤 420，浊音 / 清音分类信息提取设备从频域语音信号分离谐波部分和噪声部分。浊音 / 清音分类信息提取设备在步骤 430 计算谐波部分对噪声部分能量比率，并进入步骤 440，在步骤 440，浊音 / 清音分类信息提取设备使用步骤 430 的计算结果对浊音 / 清音声音分类。

[0063] 同时，可重新定义本发明的特征提取方法，从而将通过比较 HNR 和阈值或比较 HRR 和阈值而获得的值包括在范围 [0,1] 中（“0”表示清音，“1”表示浊音），从而是相干的。详细地说，以 dB 为单位表示 HNR 和 HRR。然而，为了使用表示浊音程度的测量，例如，在 HNR 的情况下，可将式 5 重新定义为式 6。

[0064] $$HNR=10\log_{10}\frac{P_H}{P_N}(dB) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

[0065] 在式 6 中，“P”表示功率，其中“P_N”用于 HNR，而“P_R”用于 HRR，其可根据测量而改变。浊音的范围是无穷，而清音的范围是负无穷。此外，在式 6 中，如果$\frac{P_H}{P_N}=10^{HNR/10}$，在表示浊音程度的 [0,1] 之间测量，则式 6 可表示为式 7。

[0066] $$\delta=\frac{P_H}{P_H+P_N}=\frac{10^{HNR/10}}{10^{HNR/10}+1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7)$$

[0067] 同时，基本上，由于将残余看作过程中的噪声，与根据本发明第二实施例的浊音 / 清音分类信息对应的 HNR 可具有与 HRR 相同的概念。然而，在考虑到对根据本发明第一实施例的 HRR 使用正弦表示残余的同时，在对根据本发明第二实施例的 HNR 执行谐波加噪声分解操作之后计算噪声。

[0068] 混合的浊音显示出这样的趋势：在较低频带为周期性，在较高频带为类噪声。在此情况下，在使用分量计算 HNR 之前，可对已经通过分解操作获得的谐波和噪声分量进行低通滤波。

[0069] 同时，为了防止出现在频带之间存在很大的能量差的问题，提出根据本发明第三实施例的提取浊音 / 清音分类信息的方法。在本发明的第三实施例中，将子频带的谐波分量和噪声分量之间的能量比率定义为子频带谐波对噪声比率 (SB-HNR)。具体地说，第三方法消除了当高的能量频带主导 HNR 从而生成具有很高 HNR 值的清音的区段时可能出现的问题，并可较好地控制每一频带。

[0070] 根据第三实施例，为了计算整个比率，在对 HNR 相加之前计算每一谐波部分的 HNR，从而与其它部分相比能够更高效地归一化每一谐波部分。详细地说，参照图 7A 和 7B，

从图 7A 中由标记“c”指示的频带和图 7B 中由标记“d”指示的频带中获得 HNR。在将图 7A 和图 7B 所示的频带划分为多个频带之后，每一频带具有预定的大小，按照该方式，对每一频带计算 HNR，从而获得 SB-HNR。可如式 8 来定义 SB-HNR。

[0071] $$SB-HNR=10\sum_{n-1}^{N}\log_{10}\left(\sum_{\omega_k=\Omega_k^-}^{\Omega_k^+}|H(\omega_k)|^2 / \sum_{\omega_k=\Omega_k^-}^{\Omega_k^+}|N(\omega_k)|^2\right) \quad \text{.....(8)}$$

[0072] 在式 8 中，“Ω_k^+”表示第 n 谐波频带的上频率边界，“Ω_k^-”表示第 n 谐波频带的下频率边界，“N”表示子频带数量。在图 7A 和 7B 的情况下，如下定义 SB-HNR：

[0073] SB-HNR ＝∑每谐波频带图 7A 的区域 / 每谐波频带图 7B 的区域。

[0074] 定义一个子频带位于谐波峰值的中心并从谐波峰值向两个方向扩展半个音调。与 HNR 相比，SB-HNR 更高效地均衡谐波区域，从而每一谐波区域具有相似的加权值。此外，将 SB-HNR 看作用于时间轴的部分的 SNR 的频率轴的模拟。由于计算用于每一子频带的 HNR，因此子频带谐波对残余比率可提供用于浊音 / 清音分类的更准确的基础。这里，选择性地应用带通噪声抑止滤波器（例如具有 200kHz 的上截止频率和 3400kHz 的下截止频率的九阶 Butterworth 滤波器）。这样的滤波提供合适的高频频谱滚降 (roll-off)，并同时具有当存在噪声时弱化带外噪声的效果。

[0075] 如上所述，本发明的各种提取方法简单且实用，还对测量浊音的程度非常精确和高效。可将根据本发明的提取浊音的程度的谐波分类和分析方法容易地应用于各种语音和音频特征提取方法，当与现有方法关联时还能够更准确地执行浊音 / 清音分类操作。

[0076] 可将例如 SB-HNR 的这样的基于谐波的技术应用于诸如对用于每一子频带的浊音 / 清音声音分类所需的多频带激励声码器的各种领域。此外，由于本发明基于主要谐波区域的分析，因此本发明应该具有很强的实用性。此外，由于考虑到听觉感知现象，本发明关注在浊音 / 清音分类中实际上重要的频域，因此，本发明应该具有较优的性能。

[0077] 此外，实际上，可将本发明应用于编码、识别、增强、合成等。具体地说，由于本发明需要少量计算，并使用准确检测的谐波部分来检测浊音的分量，因此本发明可更高效地应用于各种应用（其需要移动性或快速处理，或具有计算能力和存储容量的限制，例如移动终端、通信设备、PDA、MP3 等），并还可以是用于所有语音和 / 或音频信号处理系统的信源技术。

[0078] 虽然已参照其特定优选实施例示出和描述了本发明，但本领域技术人员应理解，在不脱离由所附权利要求定义的本发明的精神和范围的情况下，可对其在形式和细节上进行各种改变。因此，本发明的范围并非由上述实施例来限定，而是由权利要求及其等同物来限定。

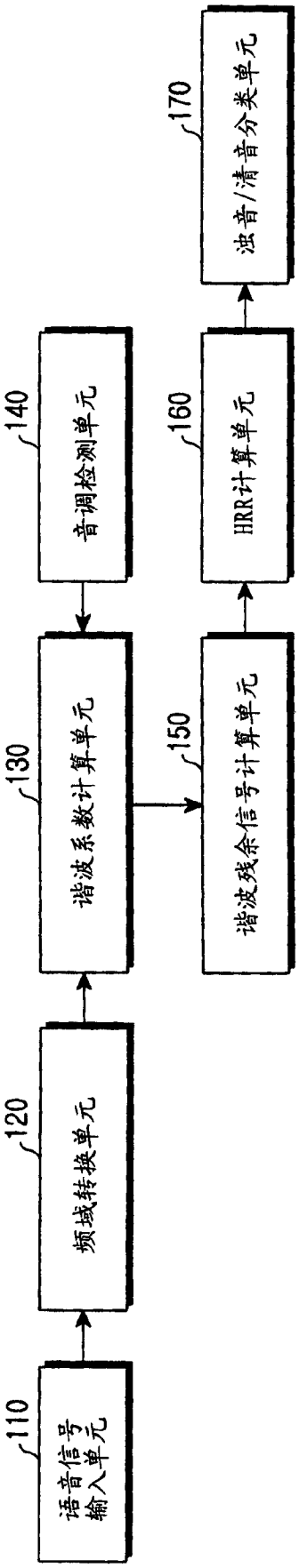

语音信号输入单元
110
频域转换单元
120
谐波系数计算单元
130
音调检测单元
140
谐波残余信号计算单元
150
HRR计算单元
160
浊音/清音分类单元
170

图 1

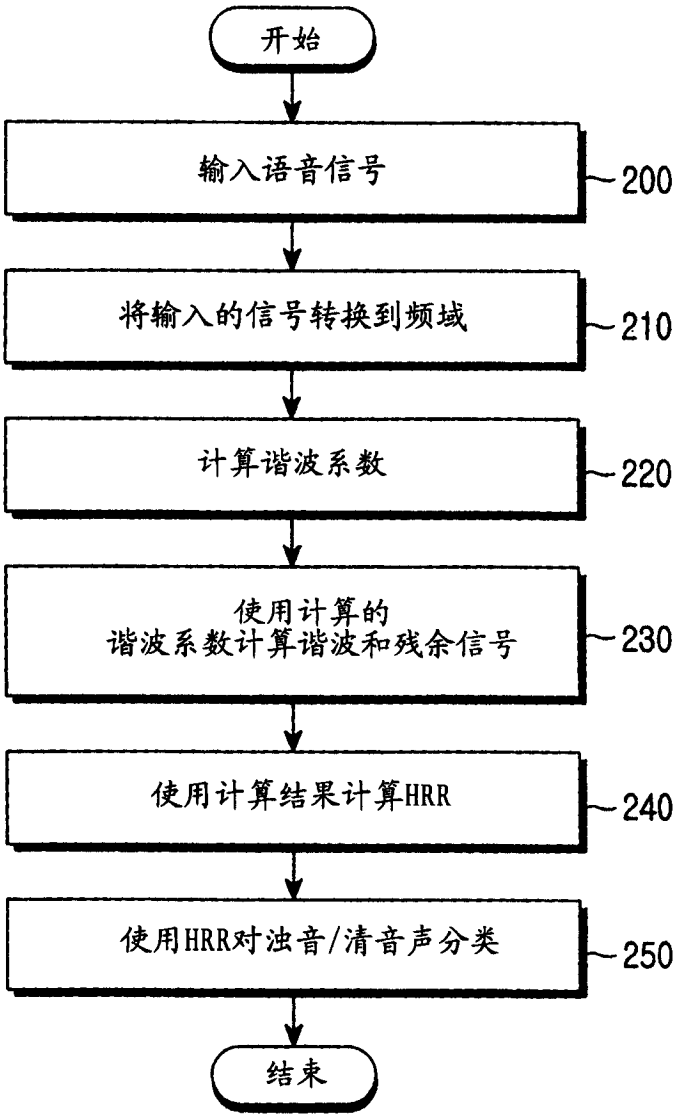

开始
输入语音信号
200
将输入的信号转换到频域
210
计算谐波系数
220
使用计算的
谐波系数计算谐波和残余信号
230
使用计算结果计算HRR
240
使用HRR对浊音/清音声分类
250
结束

图 2

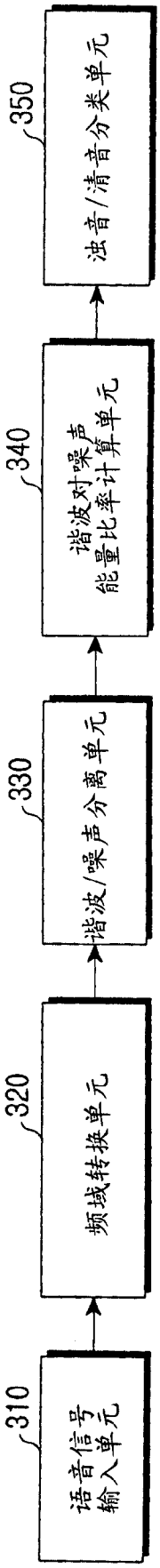

310
语音信号
输入单元
320
频域转换单元
330
谐波/噪声分离单元
340
谐波对噪声
能量比率计算单元
350
浊音/清音分类单元

图 3

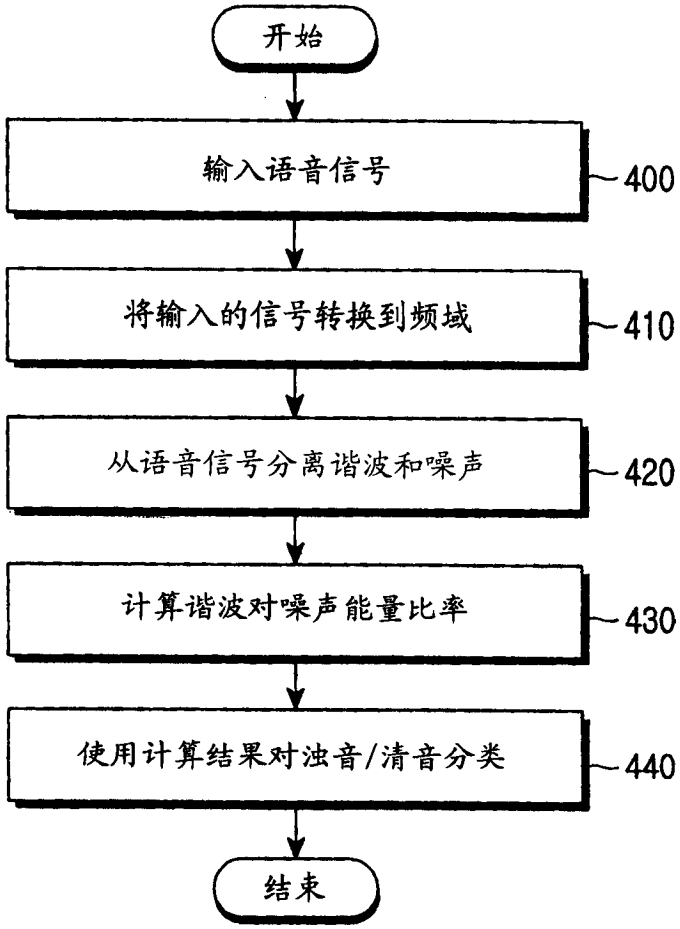

开始
输入语音信号
400
将输入的信号转换到频域
410
从语音信号分离谐波和噪声
420
计算谐波对噪声能量比率
430
使用计算结果对浊音/清音分类
440
结束

图 4

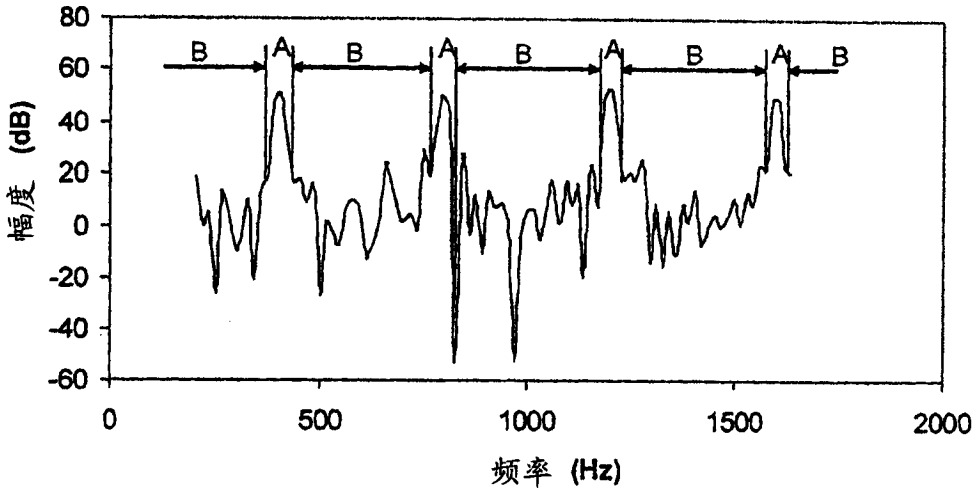

80
60
40
20
0
-20
-40
-60
幅度 (dB)
B
A
B
A
B
A
B
A
B
0
500
1000
1500
2000
频率 (Hz)

图 5

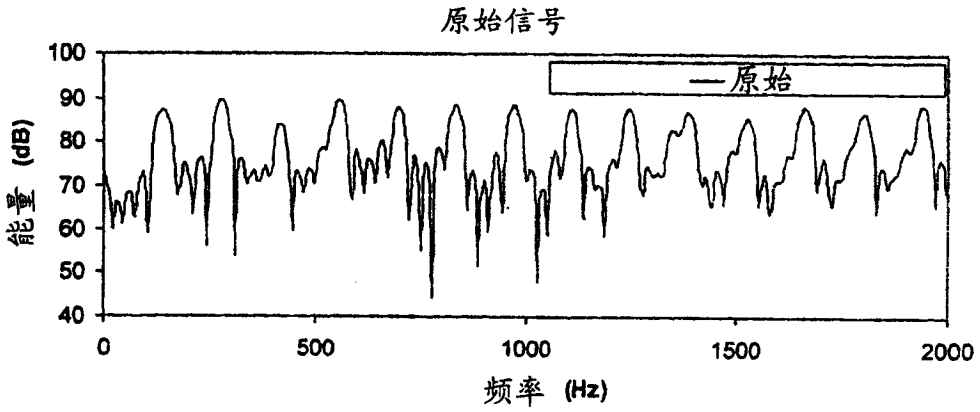

原始信号
—原始
100
90
80
70
60
50
40
能量 (dB)
0
500
1000
1500
2000
频率 (Hz)

图 6

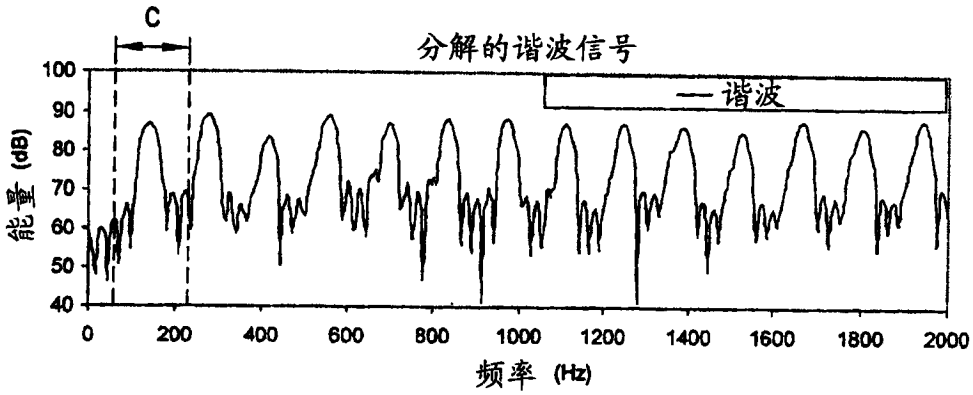

c
分解的谐波信号
—谐波
能量 (dB)
100
90
80
70
60
50
40
0
200
400
600
800
1000
1200
1400
1600
1800
2000
频率 (Hz)

图 7A

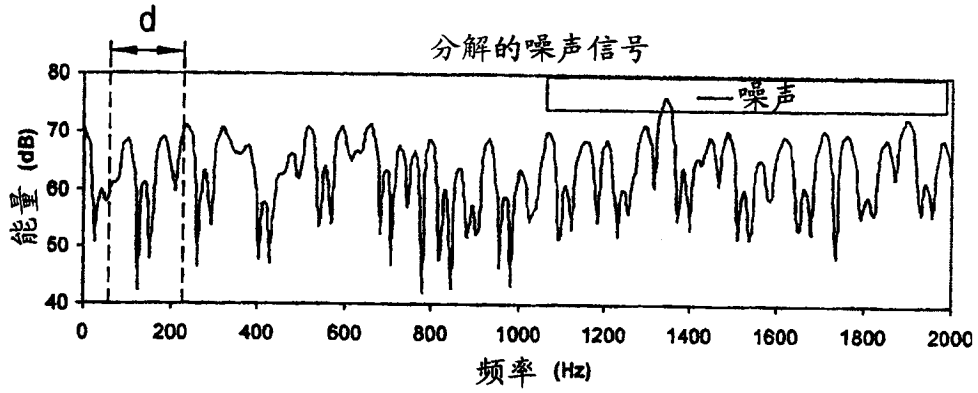

d
分解的噪声信号
—噪声
能量 (dB)
80
70
60
50
40
0
200
400
600
800
1000
1200
1400
1600
1800
2000
频率 (Hz)

图 7B