



(12) 发明专利申请

(10) 申请公布号 CN 101894565 A

(43) 申请公布日 2010. 11. 24

(21) 申请号 200910140488. 7

(22) 申请日 2009. 05. 19

(71) 申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为
总部办公楼

申请人 北京邮电大学

(72) 发明人 武穆清 李默嘉 吴大鹏 魏璐璐

甄岩 苗磊 许剑峰

(74) 专利代理机构 深圳市深佳知识产权代理事

务所(普通合伙) 44285

代理人 彭愿洁 李文红

(51) Int. Cl.

G10L 21/04 (2006. 01)

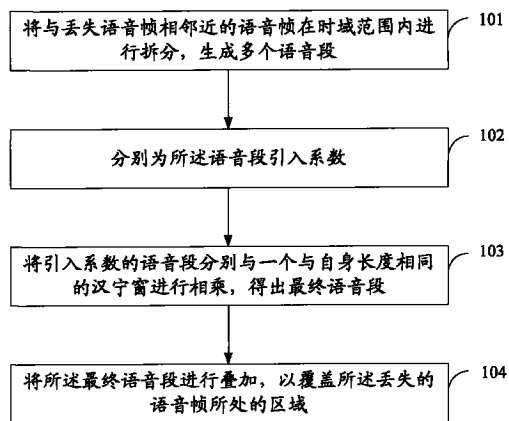
权利要求书 2 页 说明书 9 页 附图 6 页

(54) 发明名称

语音信号修复方法和装置

(57) 摘要

本发明实施例公开了一种语音信号修复方法,包括步骤:将与丢失语音帧相邻近的语音帧在时域范围内进行拆分,生成多个语音段;分别为所述语音段引入系数;将引入系数的语音段分别与一个与自身长度相同的汉宁窗进行相乘,得出最终语音段;将所述最终语音段进行叠加,以覆盖所述丢失语音帧所处的区域。同时,本发明实施例还公开了一种语音信号修复装置。通过本发明,使在采用语音拉伸法进行语音修复时,能使叠加后的波形更大程度地恢复原语音信号的幅值,避免新生成的语音信号幅度趋势同原语音信号差距过大,由此改善语音质量。



1. 一种语音信号修复方法,其特征在于,包括:
将与丢失语音帧相邻近的语音帧在时域范围内进行拆分,生成多个语音段;
分别为所述语音段引入系数;
将引入系数的语音段分别与一个与自身长度相同的汉宁窗进行相乘,得出最终语音段;
将所述最终语音段进行叠加,以覆盖所述丢失语音帧所处的区域。
2. 根据权利要求1所述的方法,其特征在于,所述将与丢失语音帧相邻近的语音帧在时域范围内进行拆分,生成多个语音段包括:
确定所述语音段的长度和放置位置;
根据所述语音段的长度、放置位置和叠加部分相关性最大化原则在所述相邻近的语音帧的波形上选取所述语音段。
3. 如权利要求1所述的方法,其特征在于,所述分别为所述语音段引入系数包括:分别为所述语音段引入增益因子。
4. 如权利要求3所述的方法,其特征在于,所述增益因子包括:所述语音段将要叠加的位置的原语音波形的平均幅值和所述语音段本身平均幅值的比值。
5. 如权利要求1至4任意一项所述的方法,其特征在于,还包括:
判断所述语音段是否处于语音异常期;
当判断为否时,所述分别为所述语音段引入系数包括,分别为所述语音段引入增益因子;
当判断为是时,所述分别为所述语音段引入系数包括,分别为所述语音段引入预先设定的因数,所述因数根据统计分析或网络传输情况产生,并为一个小于或等于1的正数。
6. 如权利要求5所述的方法,其特征在于,所述判断所述语音段是否处于语音异常期包括:
确定所述语音段将要叠加位置的原语音波形的能量和所述语音段本身能量的比值,若所述比值近似等于1,则所述语音段未处于语音异常期;否则,所述语音段处于语音异常期;或者,
确定所述语音段进行叠加时叠加部分的相关性,若所述相关性大于或等于预先设定的阈值,则所述语音段未处于语音异常期;否则,所述语音段处于语音异常期。
7. 一种语音信号修复装置,其特征在于,包括:
语音段生成单元,用于将与丢失语音帧相邻近的语音帧在时域范围内进行拆分,生成多个语音段;
系数引入单元,用于分别为所述语音段生成单元中生成的所述语音段引入系数;
汉宁窗引入单元,用于将引入系数的语音段分别与一个与自身长度相同的汉宁窗进行相乘,得出最终语音段;
语音段叠加单元,用于将所述最终语音段进行叠加,以覆盖所述丢失语音帧所处的区域。
8. 如权利要求7所述的装置,其特征在于,还包括:语音异常期判断单元,用于判断所述语音段是否处于语音异常期。
9. 如权利要求8所述的装置,其特征在于,所述语音异常期判断单元包括:

能量比值计算子单元,用于计算所述语音段将要叠加位置的原语音波形的能量和所述语音段本身能量的比值;

第一比较子单元,用于判断能量比值计算子单元计算出的能量比值是否近似等于1;或者,

相关性计算子单元,用于计算所述语音段进行叠加时叠加部分的相关性;

第二比较子单元,用于将相关性计算子单元计算所得的相关性与设定阈值进行比较。

10. 如权利要求8或9所述的装置,其特征在于,所述系数引入单元包括:

增益因子计算子单元,用于计算用于所述语音段的增益因子,所述增益因子为语音段将要叠加位置的原语音波形的平均幅值和所述语音段本身平均幅值的比值;

第一相乘子单元:用于将计算得出的所述增益因子与所述语音段进行相乘;或者,

因数生成子单元,用于根据统计分析或网络传输情况生成用于所述语音段的因数;

第二相乘子单元,用于将所述生成的因数与所述语音段进行相乘。

语音信号修复方法和装置

技术领域

[0001] 本发明涉及通信技术领域,更具体地说,涉及一种语音信号修复方法和装置。

背景技术

[0002] 随着无线网络技术的飞速发展,以及网络传输质量的不断提高,相比传统的有线网络,无线网络在便捷性和移动性方面已经表现出了相当巨大的优势。同时,基于无线网络的各种应用也迅速发展起来,而基于无线网络的VoIP(Voice over IP)技术便是其中之一。VoIP指的是利用IP网络进行话音传输,由于在分组网络中语音传输可以很容易地与其它业务结合,实现多媒体通信,并且以分组形式传输的语音信息利用了互联网低成本的特定,使其费用通常比传统的电话网传输要低,因此,受到了广大用户的欢迎。

[0003] 但由于无线网络本身的不稳定性,导致基于无线网络的VoIP语音包的传输面临着大量的丢包情况,而当VoIP业务的丢包率超过了5%,就会对语音通信质量产生比较明显的影响,而在前向纠错已无法产生作用的时候,就需要依靠接收端通过一系列的丢包恢复技术来抵消无线网络大量丢包对语音通信质量造成的不良影响。

[0004] 其中,丢包恢复技术属于丢包处理技术的一种,它指的是在已经发生了丢包的情况下,采用隐藏丢包技术,使人主观上产生一种并没有丢包的感觉的技术。对于语音信号来说,丢包恢复技术主要是利用了人类在听到不完整波形的时候的一种下意识的修复能力,在对收到的波形进行一定的改动之后,可以在相当大的程度上减轻丢包对人产生的主要影响,让接收端的人耳感官上认为并没有丢包或者丢包并不是特别严重。

[0005] 在现有技术中,通常采用波形相似叠加(WSOLA)方法来进行语音信号的丢包恢复。WSOLA方法是一种在语音处理领域常用的时域拉伸方法,它是在基于语音波形相似性的前提下工作,可以保证主观质量的前提下改变语音信号的长度。其实现过程为:当接收端探测到有一个语音帧由于传输环境影响被丢弃以后,就可以利用WSOLA方法将丢失的帧之前接收到的几个完好的语音帧进行时域拉伸,使拉伸后的语音数据长度覆盖过丢失了的语音帧的位置,使得接收端的人耳听起来就好像没有丢包一样。

[0006] 在实现本发明创造的过程中,发明人发现,上述方法至少存在以下问题:传统的WSOLA方法可能导致拉伸生成的语音信号幅度趋势同原音信号差距较大,而且容易在新生成的信号中造成幅度突变,从而降低了语音的质量。

发明内容

[0007] 本发明实施例提供一种语音信号修复方法和装置,使在对语音信号进行恢复时,新生成的语音信号幅度趋势更加接近于原语音信号,相应提高了语音质量。

[0008] 本发明实施例提供一种语音信号修复方法,包括:

[0009] 将与丢失语音帧相邻近的语音帧在时域范围内进行拆分,生成多个语音段;

[0010] 分别为所述语音段引入系数;

[0011] 将引入系数的语音段分别与一个与自身长度相同的汉宁窗进行相乘,得出最终语

音段；

[0012] 将所述最终语音段进行叠加，以覆盖所述丢失语音帧所处的区域。

[0013] 本发明实施例提供了一种语音信号修复装置，包括：

[0014] 语音段生成单元，用于将与丢失语音帧相邻近的语音帧在时域范围内进行拆分，生成多个语音段；

[0015] 系数引入单元，用于分别为所述语音段生成单元中生成的所述语音段引入系数；

[0016] 汉宁窗引入单元，用于将引入系数的语音段分别与一个与自身长度相同的汉宁窗进行相乘，得出最终语音段；

[0017] 语音段叠加单元，用于将所述最终语音段进行叠加，以覆盖所述丢失语音帧所处的区域。

[0018] 本发明实施例通过对原语音帧进行拆分，生成语音段，并为新生成的语音段引入一个系数，将引入系数的语音段与汉宁窗相乘得出最终语音段，将所述最终语音段进行叠加以覆盖丢失语音帧所处的区域的技术手段，使叠加后的波形能够更大程度上地恢复原语音信号的幅值，从而提高语音质量。

附图说明

[0019] 为了更清楚地说明本发明实施例中的技术方案，下面将对实施例中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

[0020] 图 1 为本发明实施例所涉及的一种语音信号修复方法的流程图；

[0021] 图 2 为本发明实施例所涉及的另一种语音信号修复方法的流程图；

[0022] 图 3 为本发明实施例所涉及的第三种语音信号修复方法的流程图；

[0023] 图 4 为本发明实施例所涉及的一种语音信号修复装置的结构示意图；

[0024] 图 5 为本发明实施例所涉及的另一种语音信号修复装置的结构示意图；

[0025] 图 6 为本发明实施例所涉及的一种异常期判断单元的结构示意图；

[0026] 图 7 为本发明实施例所涉及的另一种异常期判断单元的结构示意图；

[0027] 图 8 为本发明实施例所涉及的一种系数引入单元的结构示意图；

[0028] 图 9 为本发明实施例所涉及的另一种系数引入单元的结构示意图；

[0029] 图 10 为本发明实施例所涉及的一种语音信号修复的方法的流程图。

具体实施方式

[0030] 下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

[0031] 在现有技术中，在对已损坏或丢失的语音帧进行修复时，是在保证主观质量的前提下改变语音信号的长度，但由于在这个过程中，由于只考虑到了保持语音基音频率的稳定以及重叠语音的相位一致，未考虑生成语音波形与原波形在幅度上的一致，因此会导致

修复后的语音质量较低。

[0032] 本发明实施例提供了一种语音信号修复方法,具体流程如图 1 所示:

[0033] 步骤 101:将与丢失语音帧相邻近的语音帧在时域范围内进行拆分,生成多个语音段;

[0034] 步骤 102:分别为所述语音段引入系数;

[0035] 步骤 103:将引入系数的语音段分别与一个与自身长度相同的汉宁窗进行相乘,得出最终语音段;

[0036] 步骤 104:将所述最终语音段进行叠加,以覆盖所述丢失的语音帧所处的区域。

[0037] 本实施例所提供的一种语音信号修复方法,通过对原语音帧进行拆分,生成语音段,并为新生成的语音段引入一个系数,使叠加后的波形能够更大程度上地恢复原语音信号的幅值,从而提高语音质量。

[0038] 同时,本发明实施例还提供了另外一种语音信号修复方法,具体流程如图 2 所示:

[0039] 步骤 201:将与丢失语音帧相邻近的语音帧在时域范围内进行拆分,生成多个语音段;

[0040] 在步骤 201 中,将该丢失的语音帧邻近的几个完整语音帧进行拆分,生成语音段,在这个过程中,首先要确定所要使用的语音帧和丢失语音帧的总长度,此处将该长度称为语音帧总长度,该长度决定了语音段进行叠加放置后所形成波形的总长度。在进行拆分所生成语音段的长度和该语音段将要放置的位置的确定过程中,可以有多种方式,而这些方式需要满足的条件是相邻语音段之间必须进行叠加,其目的是为了保证在将语音段进行放置后,各波段之间能有一个平滑过度。为便于实际应用技术方案实现,可以预先设定语音段的个数,并将语音段的长度取为相同,同时使相邻两个语音段之间相互重叠一半,从而可以根据上述几个条件求得生成语音段的长度。

[0041] 在语音段的个数、长度以及相互之间的叠加关系均确定以后,需要将语音段从原语音帧上进行拆分,该过程可以通过如下方式进行:

[0042] 从原语音帧的起始处取与语音段长度相同的一段作为第 1 个语音段,并将该语音段放置于语音帧总长度的起始处。

[0043] 在选取第 2 个语音段时,首先为该语音段的起始位置选取一个范围,使该语音段在该范围内进行选取时与第 1 个语音段进行叠加时能满足相关性最大,即与第 1 个语音段进行叠加时能够尽可能地相位保持一致。

[0044] 同理,可以进行后面所有语音段的选取。

[0045] 步骤 202:分别为所述语音段引入增益因子;

[0046] 步骤 202 的目的是为了使语音段进行叠加后所形成的新的波形与原语音波形在幅度上尽量相同。其中,该处引入的增益因子可以为:该语音段将要叠加位置的原语音波形的平均幅值和该语音段本身的平均幅值的比值。这样,在语音段与该增益因子相乘后,可以在叠加时尽量在幅度和原语音波形保持一致。

[0047] 步骤 203:将引入增益因子的语音段分别与一个与自身长度相同的汉宁窗进行相乘,得出最终语音段;

[0048] 由于在进行语音段的叠加过程中,语音段的重叠必然导致叠加后的语音幅度的增加,因此,需要对参与叠加的每个语音段施加一个汉宁窗,即参与叠加的每个语音段分别与

一个与自身长度相同的汉宁窗相乘,这样,在语音段的叠加部分会有一个变化的衰减,并且该衰减可以确保叠加部分的最终幅度不至于过大。

[0049] 步骤 204:将所述最终语音段进行叠加,以覆盖所述丢失的语音帧所处的区域。

[0050] 在得到最终的语音段之后,根据之前所确定的各个语音段的放置位置,将最终得到的语音段放置于对应位置,以覆盖所述丢失的语音帧所处的区域。

[0051] 本实施例所提供的一种语音信号修复方法,通过对原语音帧进行拆分,生成多个语音段,并分别为新生成的语音段引入相应的增益因子,使叠加后的波形能够更大程度上地恢复原语音信号的幅值,从而提高语音质量。

[0052] 相应地,本发明实施例还提供了第三种语音信号修复方法,具体流程如图 3 所示:

[0053] 步骤 301:将与丢失语音帧相邻近的语音帧在时域范围内进行拆分,生成多个语音段;

[0054] 在步骤 301 中,将该丢失的语音帧邻近的几个完整语音帧进行拆分,生成语音段,在这个过程中,首先要确定所要使用的语音帧和丢失语音帧的总长度,此处将该长度称为语音帧总长度,该长度决定了语音段进行叠加放置后所形成波形的总长度。在进行拆分所生成语音段的长度和该语音段将要放置的位置的确定过程中,可以有多种方式,而这些方式需要满足的条件是相邻语音段之间必须进行叠加,其目的是为了保证在将语音段进行放置后,各波段之间能有一个平滑过度。为便于实际应用技术方案实现,可以预先设定语音段的个数,并将语音段的长度取为相同,同时使相邻两个语音段之间相互重叠一半,从而可以根据上述几个条件求得生成语音段的长度。

[0055] 在语音段的个数、长度以及相互之间的叠加关系均确定以后,需要将语音段从原语音帧上进行拆分,该过程可以通过如下方式进行:

[0056] 从原语音帧的起始处取与语音段长度相同的一段作为第 1 个语音段,并将该语音段放置于语音帧总长度的起始处。

[0057] 在选取第 2 个语音段时,首先为该语音段的起始位置选取一个范围,使该语音段在该范围内进行选取时与第 1 个语音段进行叠加时能满足相关性最大,即与第 1 个语音段进行叠加时能够尽可能地相位保持一致。

[0058] 同理,可以进行后面所有语音段的选取。

[0059] 步骤 302:判断语音段是否处于语音异常期;

[0060] 在步骤 302 中,由于生成的多个语音段有可能会处于语音异常期,如语音转换期或白噪声期等,其中,语音转换期可以理解为:当任意长度的一段语音幅度变化较频繁,并且有很多的语音幅度为零值。从而需要对所生成的语音段分别判断其是否处于语音异常期。在本实施例中,可以采用如下两种方法实现对语音段是否处于语音异常期的判断:

[0061] 方法一:计算语音段将要叠加位置的原语音波形的能量和语音段本身能量的,如果两者相差过大,则可以认为该语音段处于语音异常期,换一种说法为:如果语音段将要叠加位置的原语音波形的能量和语音段本身能量的该比值近似等于 1,认为该语音段未处于语音异常期;否则,认为处于语音异常期。

[0062] 方法二:在语音段与其他语音段进行叠加放置时,计算该语音段的叠加部分的相关性,当该相关性大于预先设定的阈值时,表示该语音段未处于语音异常期;否则,表示新语音段处于语音异常期。在本方法中,如果计算得出的相关性小于设定的阈值,表明该语

音段很难与其他语音段进行叠加时实现相位上的一致,则可以认为该语音段处于语音异常期。

[0063] 步骤 303:根据判断结果分别为语音段引入相应的系数;

[0064] 步骤 303 的目的是为了防止在新语音段进行叠加放置后所形成的新的波形与原语音波形在幅度上存在过大差距。

[0065] 在步骤 302 中,分别对生成的语音段进行语音异常期的判断后,根据判断结果为语音段引入相应系数:对于未处于语音异常期的语音段,为其引入增益因子;而对于处于语音异常期的语音段,相应为其引入一个预先设定的因数。

[0066] 其中,增益因子的计算方法前面已做过介绍,在此不做赘述;而预先设定的因数则可以根据统计结果和当前网络的传输状况来得到,例如,将该传输网络之前长时间的传输状况进行统计分析,依据以往数据设定一个值,也可以只考虑当前网络的传输状况设定一个值,通常情况下,网络传输状况较差时,易发生语音处于异常期的情况,则相应地,语音段需要一个较大的衰减,则所设定的因数要较小。基本上,该处所设定的因数都为一个小于或等于 1 的正数。

[0067] 需要说明的是,也可以先对所有语音段进行增益因子的计算,再进行语音异常期的判断,并根据语音异常期的判断结果,决定所计算出的增益因子是否被采用。在这里,这两步的具体顺序不做特殊要求。

[0068] 步骤 304:将引入系数的语音段分别与一个与自身长度相同的汉宁窗进行相乘,得出最终语音段;

[0069] 由于在进行语音段的叠加过程中,语音段的重叠必然导致叠加后的语音幅度的增加,因此,需要对参与叠加的每个语音段施加一个汉宁窗,即参与叠加的每个语音段分别与一个与自身长度相同的汉宁窗相乘,这样,在语音段的叠加部分会有一个变化的衰减,并且该衰减可以确保叠加部分的最终幅度不至于过大。

[0070] 步骤 305:将所述最终语音段进行叠加,以覆盖所述丢失的语音帧所处的区域。

[0071] 在得到最终的语音段之后,根据之前所确定的各个语音段的放置位置,将最终得到的语音段放置于对应位置,以覆盖所述丢失的语音帧所处的区域。

[0072] 本实施例所提供的一种语音信号修复方法,通过对原语音帧进行拆分,生成多个语音段,对所生成的语音段分别进行语音异常期的判断,并根据判断结果分别为新生成的语音段引入相应的系数,使叠加后的波形能够更大程度上地恢复原语音信号的幅值,从而提高语音质量。

[0073] 本发明实施例还相应提供了一种语音信号修复装置,如图 4 所示,该装置包括:

[0074] 语音段生成单元 401,用于将与丢失语音帧向邻近的语音帧在时域范围内进行拆分,生成多个语音段;

[0075] 系数引入单元 402,用于分别为所述语音段生成单元中生成的所述语音段引入系数;

[0076] 汉宁窗引入单元 403,用于将引入系数的语音段分别与一个与自身长度相同的汉宁窗进行相乘,得出最终语音段;

[0077] 语音段叠加单元 404,用于将所述最终语音段进行叠加,以覆盖所述丢失的语音帧所处的区域。

[0078] 结合以上装置,对语音信号进行恢复包括:

[0079] 语音段生成单元 401 将与丢失语音帧相邻近的语音帧在时域范围内进行拆分,生成多个语音段;为了使语音段在叠加后能够尽量与原语音的波形保持一致,需要通过系数引入单元 402 根据每个语音段的不同情况分别引入不同的系数;由于语音段在进行叠加时会导致叠加后的语音幅度的增加,因此,需要汉宁窗引入单元 403 将引入系数的语音段分别与一个与自身长度相同的汉宁窗进行相乘,并得出最终语音段;之后,语音段叠加单元 404 将生成的最终语音段进行叠加,以覆盖所述丢失的语音帧所处的区域。

[0080] 本实施例所提供的一种语音修复装置,通过对原语音帧进行拆分,生成多个语音段,并分别为新生成的语音段引入相应的增益因子,使叠加后的波形能够更大程度上地恢复原语音信号的幅值,从而提高语音质量。

[0081] 本发明实施例还相应提供了另外一种语音信号修复装置,如图 5 所示,该装置包括:

[0082] 语音段生成单元 501,用于将与丢失语音帧向邻近的语音帧在时域范围内进行拆分,生成多个语音段;

[0083] 语音异常期判断单元 502,用于判断所述语音段是否处于语音异常期;

[0084] 系数引入单元 503,用于分别为所述语音段生成单元中生成的所述语音段引入系数;

[0085] 汉宁窗引入单元 504,用于将引入系数的语音段分别与一个与自身长度相同的汉宁窗进行相乘,得出最终语音段;

[0086] 语音段叠加单元 505,用于将所述最终语音段进行叠加,以覆盖所述丢失的语音帧所处的区域。

[0087] 其中,语音异常期判断单元可以进一步包括如图 6 所示子单元:

[0088] 能量比值计算子单元 601,用于计算所述语音段将要叠加位置的原语音波形的能量和所述语音段本身能量的比值;

[0089] 第一比较子单元 602,用于判断能量比值计算子单元计算出的能量比值是否近似等于 1,如果是,确定所述语音段未处于语音异常期;否则,确定所述语音段处于语音异常期;

[0090] 另外,语音异常期判断单元还可以进一步包括如图 7 所示子单元:

[0091] 相关性计算子单元 701,用于计算所述语音段进行叠加时叠加部分的相关性;

[0092] 第二比较子单元 702,用于将相关性计算子单元计算所得的相关性与设定阈值进行比较,如果所述相关性大于预先设定的阈值,确定所述语音段未处于语音异常期;否则,确定所述语音段处于语音异常期。

[0093] 另外,根据判断结果的不同,语音段所引入的系数也存在不同,当判断语音段未处于语音异常期时,为该语音段引入一个增益因子;否则,为该语音段引入一个预先设定的因数。

[0094] 因此,系数引入单元也相应包括以下两种结构:

[0095] 一种如图 8 所示,包括:

[0096] 增益因子计算子单元 801,用于计算用于所述语音段的增益因子,所述增益因子为语音段将要叠加位置的原语音波形的平均幅值和所述语音段本身平均幅值的比值;

[0097] 第一相乘子单元 802 :用于将计算得出的所述增益因子与所述语音段进行相乘。

[0098] 另外一种如图 9 所示,包括 :

[0099] 因数生成子单元 901,用于根据统计分析或网络传输情况生成用于所述语音段的因数 ;

[0100] 第二相乘子单元 902,用于将所述生成的因数与所述语音段进行相乘。

[0101] 结合以上装置,对语音信号进行恢复具体为 :

[0102] 语音段生成单元 501 将与丢失语音帧相邻近的语音帧在时域范围内进行拆分,生成多个语音段 ;由于生成的语音段有可能处于语音异常期,而影响修复的效果,需要由语音异常期判断单元 502 判断语音段是否处于语音异常期 ;根据判断结果,如果该语音段未处于语音异常期,则由增益因子计算子单元 801 为该语音段计算增益因子,并通过第一相乘子单元 802 将计算得出的增益因子与该语音段进行相乘 ;而如果该语音段处于语音异常期,则由因数生成子单元 901 生成的语音段的因数,并通过第二相乘子单元 902 将计算得出的增益因子与该语音段进行相乘 ;由于语音段在进行叠加时会导致叠加后的语音幅度的增加,因此,需要汉宁窗引入单元 504 将引入系数的语音段分别与一个与自身长度相同的汉宁窗进行相乘,并得出最终语音段 ;之后,语音段叠加单元 505 将生成的最终语音段进行叠加,以覆盖所述丢失的语音帧所处的区域。

[0103] 本实施例所提供的一种语音信号修复装置,通过对原语音帧进行拆分,生成多个语音段,对所生成的语音段分别进行语音异常期的判断,并根据判断结果分别为新生成的语音段引入相应的系数,使叠加后的波形能够更大程度上地恢复原语音信号的幅值,从而提高语音质量。

[0104] 结合上述方法和具体应用情况,本实施例对本发明的技术方案做进一步介绍。

[0105] 假设发送端发送 3 帧语音信号,但由于网络原因,第 3 帧信号在传输过程中丢失,接收端需要对前面两个完好的语音帧进行拉伸,使其覆盖过第 3 个语音帧的位置。具体步骤如图 10 所示 :

[0106] 步骤 1001 :将接收到的 2 个完整的语音帧拆分为 3 段长度相同的语音段。

[0107] 在步骤 1001 中,假设接收到的 2 个完好的语音帧的长度分别为 20ms,在 8000Hz 的采样频率之下,该 2 个语音帧分别包括 160 个采样点。由于要满足相邻两个语音段之间相互重叠一半,而重叠之后的语音段刚好覆盖 3 个语音帧,即 480 个样点长度的数据,由此可以得出,拆分后的语音段长度应为 240 个样点。

[0108] 下面,对如何进行语音帧的拆分做详细介绍 :由于要将现有的两个 160 个样点长度的语音帧拆分为 3 个 240 个样点长度的语音段,因此,要进行如下操作 :

[0109] 通常情况下,将输入的两帧语音的开始处作为第 1 个语音段的起始位置,则第 1 个语音段应该是从第 1 个样点到第 240 个样点,在对第 2 个语音段进行选取的过程中,为便于实现,该语音段的起始位置可以从第 1 至第 41 个样点中进行选取,并根据选取的起始位置,向后数 240 个样点,形成第 2 个语音段 ;同理,第 3 个语音段的起始位置则在第 41 至第 81 个样点中进行选取,并根据选取好的起始位置向后数 240 个样点,形成第 3 个语音段。需要注意的是,在选取语音段的起始位置的时候,要考虑到在 3 个语音段进行叠加时,尽量使相互叠加的语音段的相位保持一致,即让两段语音信号的波峰和波峰相叠加,波谷和波谷相叠加,因此,在选取语音段的起始位置的时候,通常要首先计算每个语音段叠加部分的相关

性,最大相关性所对应的样点,即为该语音段的起始位置。

[0110] 步骤 1002:判断语音段是否处于语音异常期,例如语音转换期或白噪声期等,如果是,则进入步骤 1003;否则,进入步骤 1004。

[0111] 在步骤 1002 中,在进行语音转换期或白噪声期的判别时,可以采用如下方式:

[0112] 以第 2 语音段为例,采用如下公式进行判断:

$$[0113] \quad g1 = \frac{XY}{X^2} = \frac{\sum xy}{\sum x^2} \quad g2 = \frac{Y^2}{X^2} = \frac{\sum y^2}{\sum x^2}$$

[0114] 其中, X 为第 2 个语音段在原语音帧中所处位置的采样点值, Y 为第 2 个语音段在将要叠加的位置的采样点值。将计算得出的 g1 和 g2 进行比较,如果 g1 约等于 g2,也就是说该语音段将要叠加位置的原语音波形的能量和该语音段本身能量的比值近似等于 1 则说明该语音段不处于语音转换期或白噪声期,否则,说明该语音段处于语音转换期或白噪声期。

[0115] 同理,可进行对第 3 语音段的判断。

[0116] 除了利用上述方法进行语音转换期或白噪声期的判断外,还可采用以下方法进行判断:

[0117] 仍然以第 2 个语音段为例,前面已做过介绍,在进行第 2 个语音段的起始位置的选择时,选择范围为原语音帧的第 1 至第 41 个样点,而由于进行叠加后,第 2 个语音段的前 120 个样点将与第 1 个语音段的后 120 个样点产生重叠,将该范围内的每一个样点都假设为第 2 个语音段的起点,并依次与第一个语音段的后 120 个样点进行相关性计算,计算所得最大值所对应的样点即为第 2 个语音段的起始位置,而如果计算所得相关性的最大值大于预先设定的阈值,说明该语音段未处于语音转换期或白噪声期;否则,说明该语音段处于语音转换期或白噪声期。

[0118] 同理,可进行对第 3 语音段的判断。

[0119] 在该步骤中,通常情况下,将阈值设在 0.5 ~ 2 之间,该方法的优点是不需要额外复杂的计算,在进行语音段拆分的过程中已将每个数据段的相关性做了计算。

[0120] 步骤 1003:对语音段使用一个预定义的衰减。

[0121] 在步骤 1003 中,对第 2 和第 3 语音段使用一个预定义的衰减,可以是将第 2 和第 3 语音段分别与一个预定的小于 1 的系数相乘,从而实现第 2 和第 3 语音段的衰减,而对于第 1 语音段,可以不对其幅值进行改变。

[0122] 步骤 1004:计算分别用于 3 个语音段的增益因子。

[0123] 步骤 1004 是该实施例的一个关键步骤,目的是使叠加后生成的波形包络幅度能很好地和原波形相匹配。

[0124] 其中,用于第 2 个语音段的增益因子可以通过如下公式进行计算:

$$[0125] \quad C2 = \frac{\sum_{n=121}^{360} x(n)}{\sum_{n=s}^{s+239} x(n)} = \frac{\sum_{n=121}^{360} x(n)}{\sum_{n=s}^{s+239} x(n)} \cdot \frac{240}{240}$$

[0126] 其中,C2 表示第 2 语音段将要叠加的位置的原语音波形的平均幅值和第 2 语音段

本身平均幅值的比值；S 指代第 2 个语音段的起始位置。

[0127] 而用于第 3 个语音段的增益因子通过如下公式进行计算：

$$[0128] \quad C3 = \frac{3 \sum_{n=241}^{320} x(n)}{\sum_{n=s'} x(n)} = \frac{3 \sum_{n=241}^{320} x(n) / 240}{\sum_{n=s'} x(n) / 240}$$

[0129] 其中，C3 表示第 3 语音段将要叠加的位置的原语音波形的平均幅值和第 3 语音段本身平均幅值的比值；S' 指代第 3 个语音段的起始位置。

[0130] 而由于第 1 语音段在原语音帧中的位置与叠加放置后的位置相同，因此，其增益因子为 1，可以认为不用对第 1 语音段引入增益因子。

[0131] 步骤 1005：将步骤 1004 中计算得出的增益因子分别与第 2 个语音段和第 3 个语音段进行相乘。

[0132] 步骤 1006：利用同每个语音段长度相同的汉宁窗分别与每个语音段相乘。

[0133] 由于要进行 3 个语音段的叠加，而在对语音段进行重叠后，重叠的部分必然导致语音幅度的增加，因此，对参与叠加的语音段施加一个汉宁窗，可以在叠加部分受到一个变化的衰减，使叠加部分的幅度增加不至于过大。

[0134] 步骤 1007：将拆分所得的 3 个语音段在一个 480 个样点的语音区间内进行叠加。

[0135] 在步骤 1007 中，由于每个语音段均具有 240 个样点，在进行语音段的重叠过程中，需要满足相邻两个语音段之间相互重叠一半，因此，将第 2 个语音段的前 120 个样点与第 1 个语音段的后 120 个样点进行重叠，将第 3 个语音段的前 120 个样点与第 2 个语音段的后 120 个样点进行重叠，从而实现 3 个 240 样点的语音段在 480 个样点的区域内完成重叠。

[0136] 本领域普通技术人员可以理解实现上述方法实施方式中的全部或部分步骤是可以通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取存储介质中，这里所称的存储介质，如：ROM/RAM、磁碟、光盘等。

[0137] 对所公开的实施例的上述说明，使本领域专业技术人员能够实现或使用本发明。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的，本文中所定义的一般原理可以在不脱离本发明的精神或范围的情况下，在其它实施例中实现。因此，本发明将不会被限制于本文所示的这些实施例，而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。

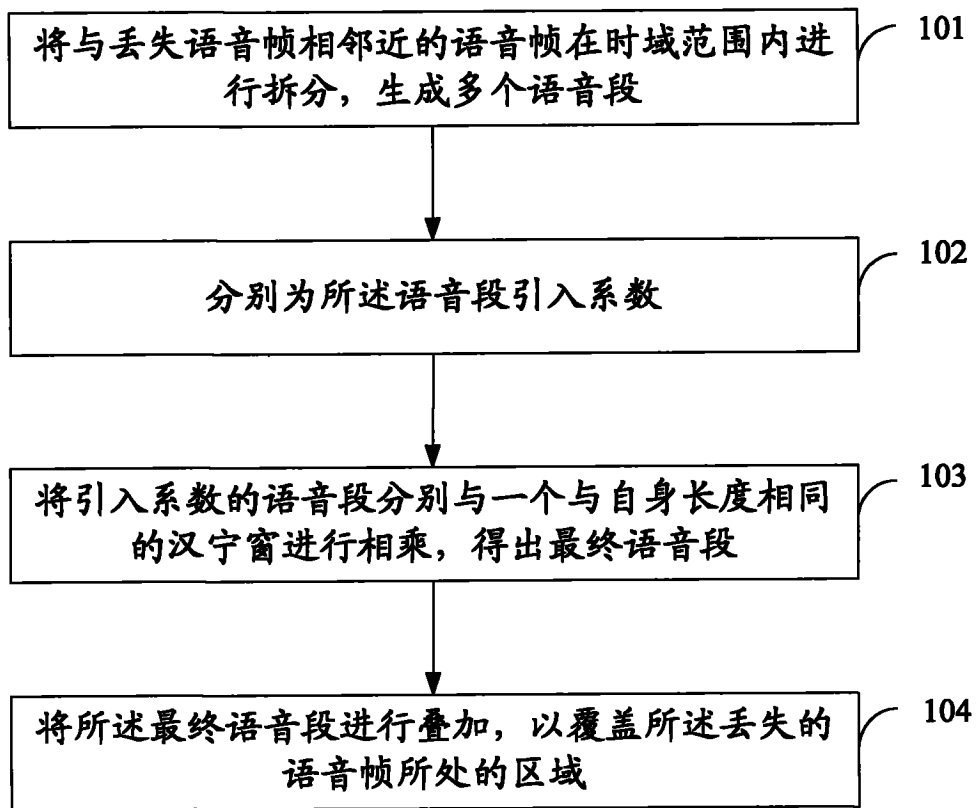


图 1

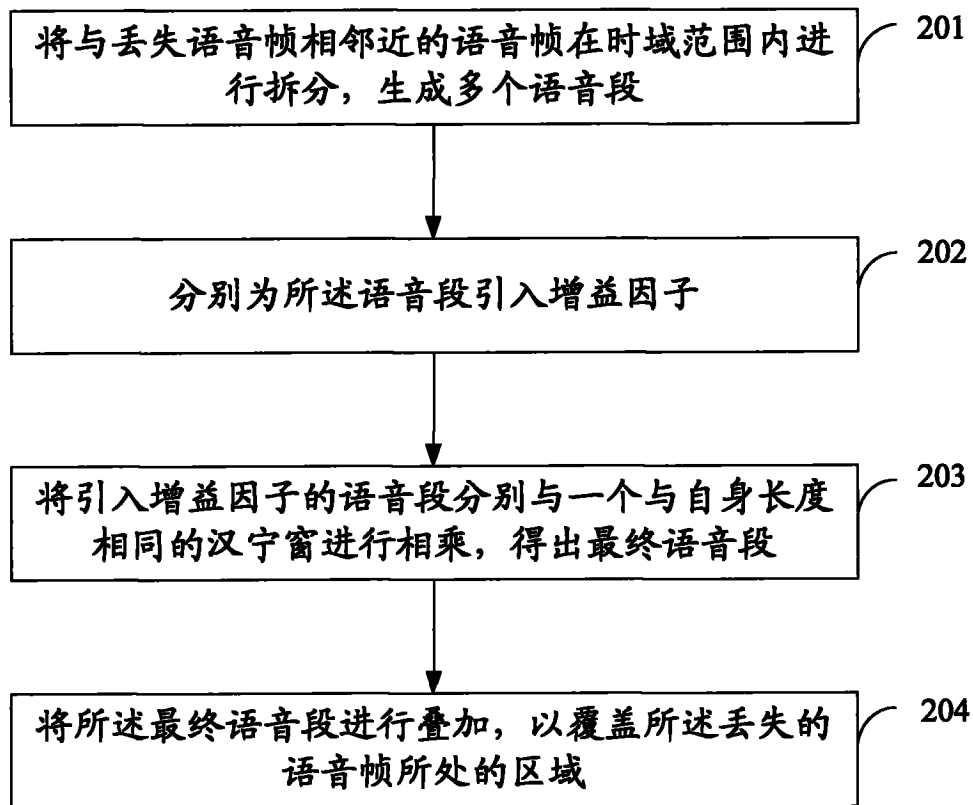


图 2

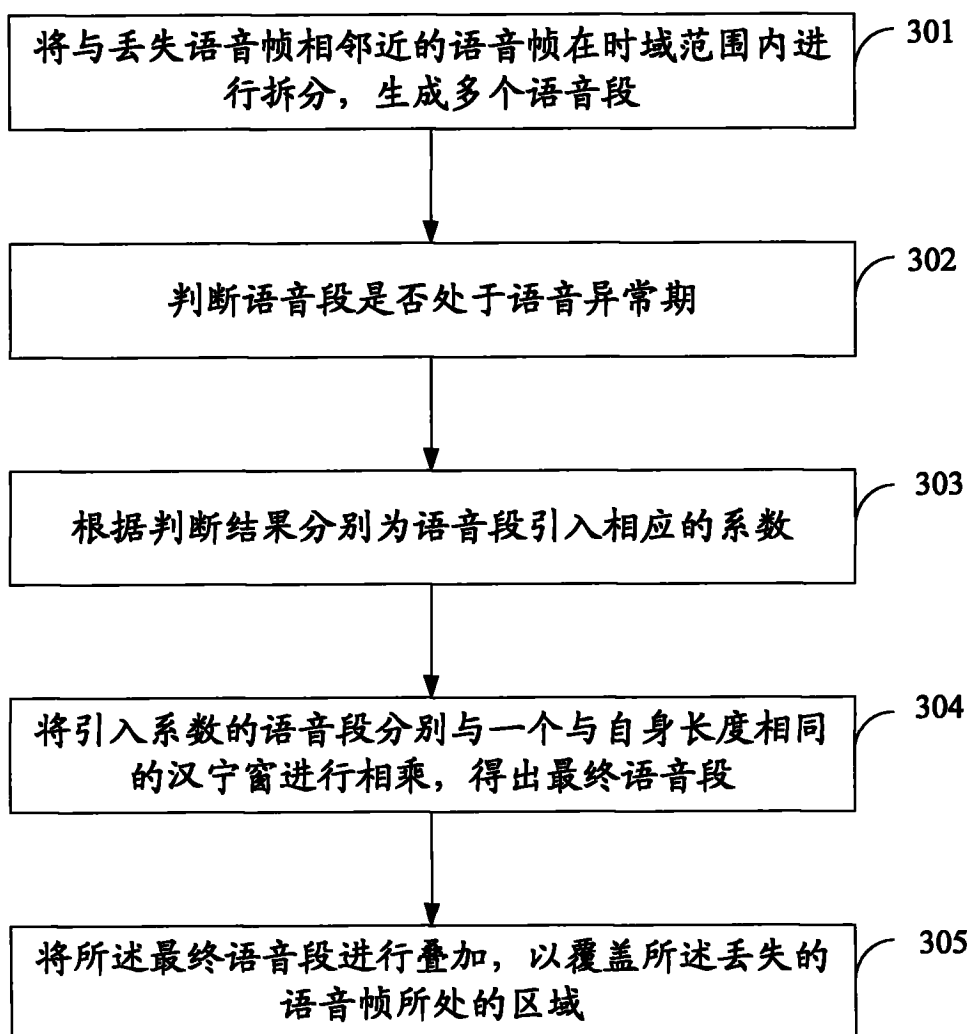


图 3

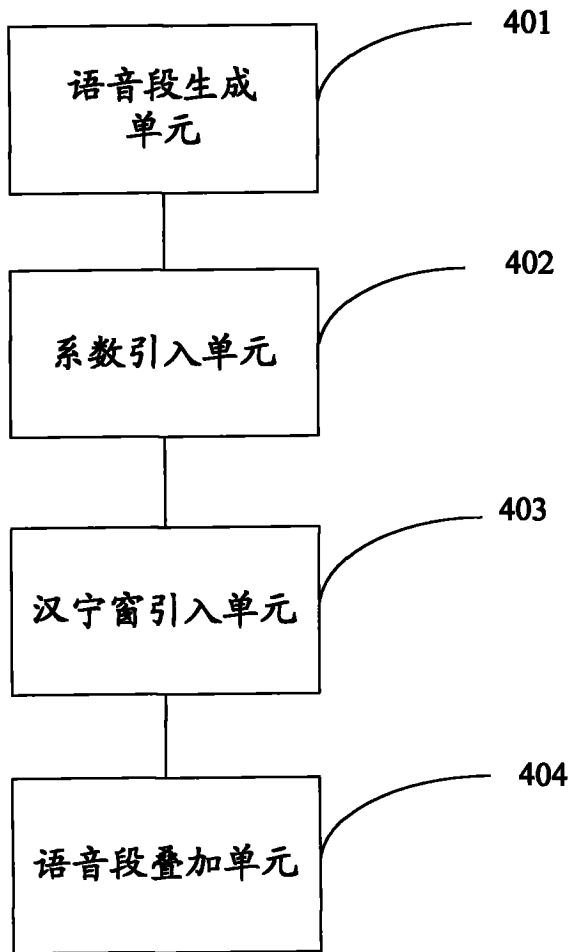


图 4

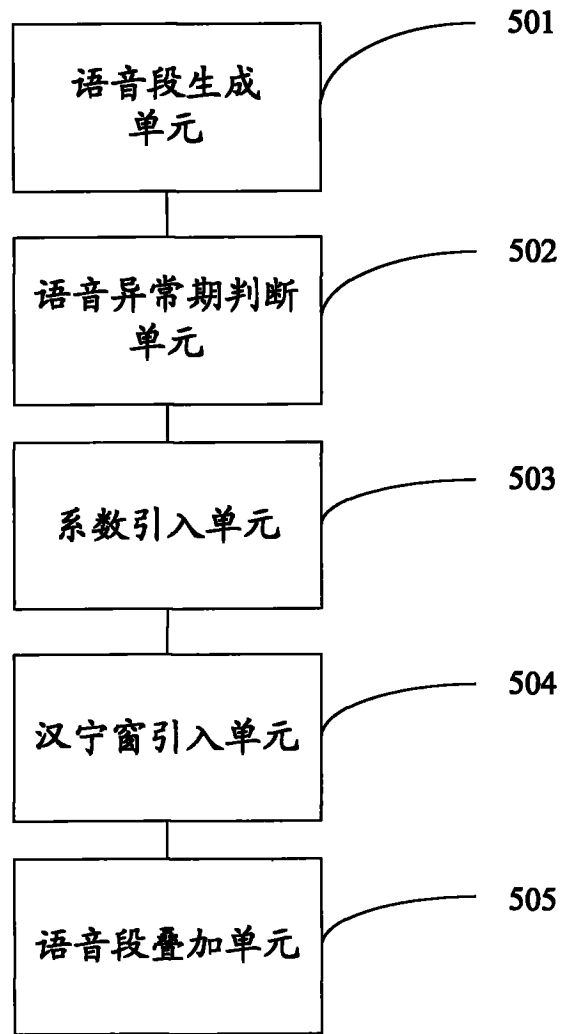


图 5

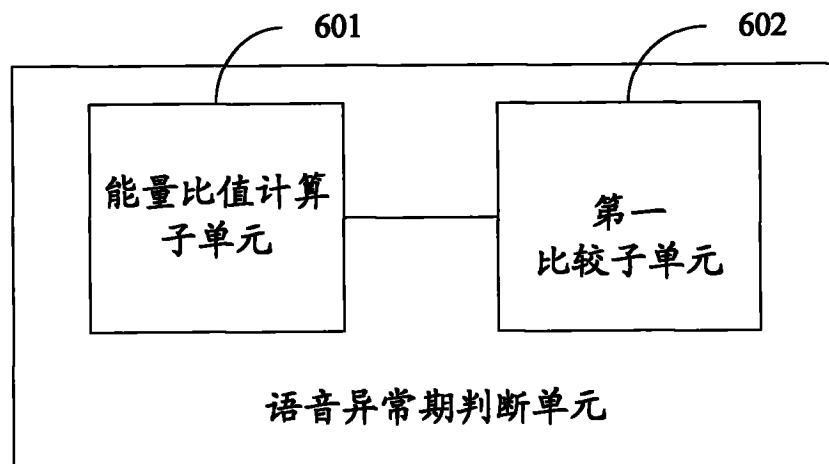


图 6

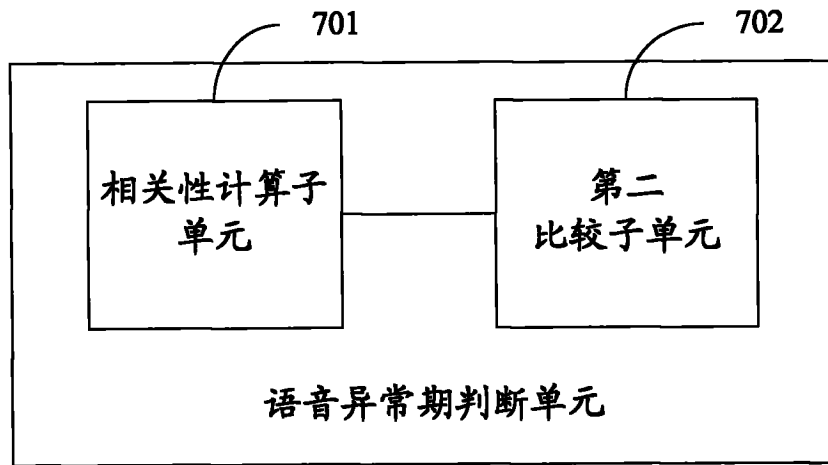


图 7

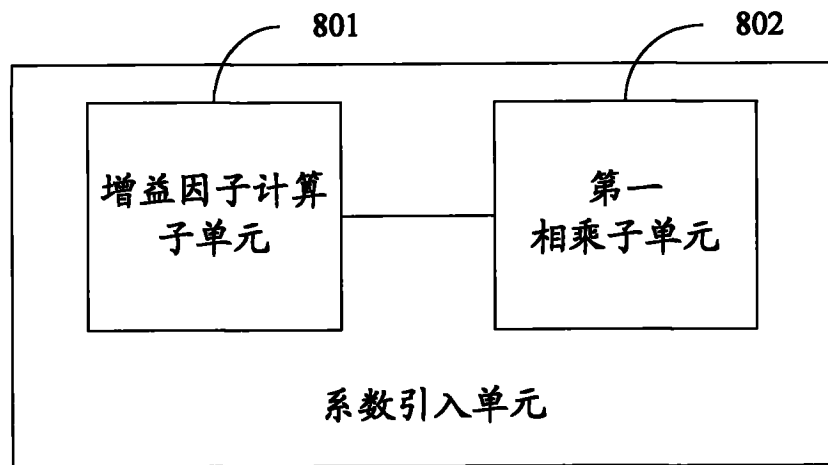


图 8

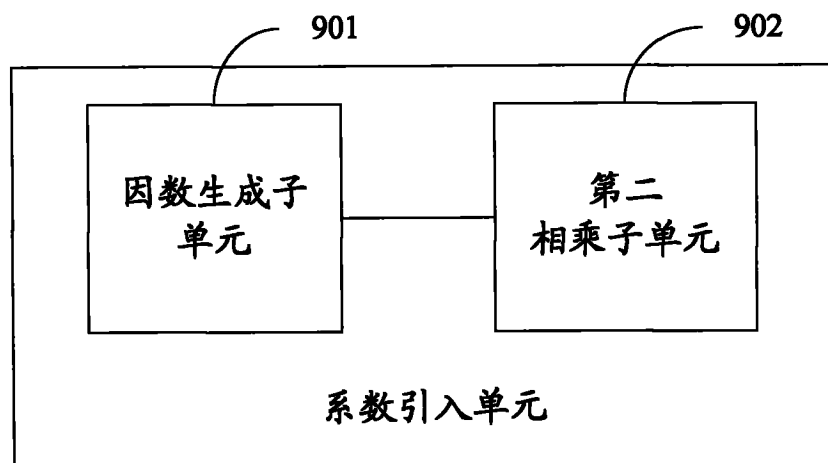


图 9

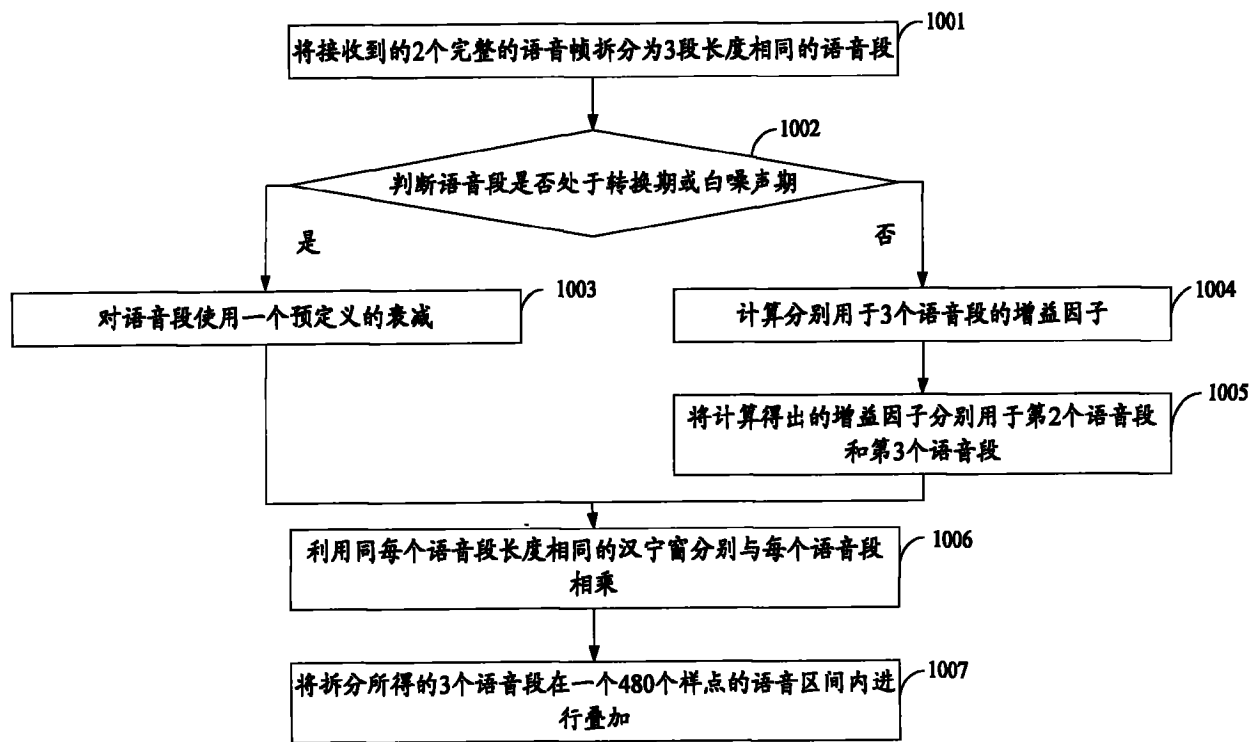


图 10