



(12) 发明专利

(10) 授权公告号 CN 101617363 B

(45) 授权公告日 2012. 09. 05

(21) 申请号 200780051682. 7

(22) 申请日 2007. 02. 21

(85) PCT申请进入国家阶段日
2009. 08. 21

(86) PCT申请的申请数据
PCT/SE2007/050100 2007. 02. 21

(87) PCT申请的公布数据
W02008/103087 EN 2008. 08. 28

(73) 专利权人 艾利森电话股份有限公司
地址 瑞典斯德哥尔摩

(72) 发明人 T·特朗普

(74) 专利代理机构 中国专利代理(香港)有限公司
72001

代理人 王岳 李家麟

(51) Int. Cl.
G10L 21/02(2006. 01)
H04M 9/08(2006. 01)

(56) 对比文件

EP 0847180 A1, 1998. 06. 10, 全文.

US 6138092 A, 2000. 10. 24, 全文.

US 20051047235 A1, 2005. 07. 07, 全文.

审查员 刘红梅

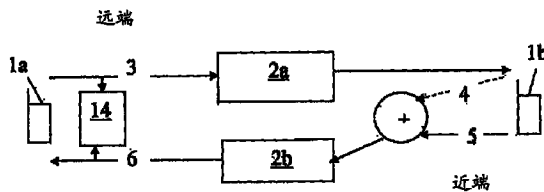
权利要求书 3 页 说明书 9 页 附图 4 页

(54) 发明名称

双端话音检测器

(57) 摘要

双端话音检测器 (14), 用于通过指示何时所接收编码语音信号 (6) 由非回声信号 (5) 占支配地位即存在所谓的双端话音而控制在电信系统中的回声路径估计。这是通过当信号功率超过第一阈值时从所接收编码语音信号的编码语音帧中提取 LSP、把每个所述提取的 LSP 转换成 LSP(线频谱频率) Q_i 和计算在每两个相邻的 LSF 之间的距离而被确定的。对于小于第二阈值的每个距离, 在两个 LSF 之间定位频谱峰, 以及确定所述频谱峰是否为回声。当预定数目的非回声频谱峰位于接收语音信号 (6) 中时, 将指示双端话音, 以及可以停止回声路径估计。



1. 一种检测所接收编码语音信号 (6) 是由非回声信号 (5) 占支配地位的方法,其特征在于以下步骤:

- 如果所接收语音信号 (6) 的信号功率超过第一阈值,则从所述所接收语音信号 (6) 的编码语音帧中提取 LSP(线频谱对);

- 把每个所述提取的 LSP 转换成 LSF(线频谱频率) ω_i , 以及计算在每两个相邻的 LSF 之间的距离;

- 对于小于第二阈值的每个所述距离,计算由所述 LSF 围绕的频谱峰 ω_c 的频率,以及确定所述频谱峰是否为回声。

2. 按照权利要求 1 的方法,其特征在于进一步的步骤:

- 对于不是回声的每个所定位的频谱峰,对计数器进行增量;

- 当计数器达到预定的阈值时指示双端话音。

3. 按照权利要求 1 的方法,其特征在于确定频谱峰是否为回声包括以下步骤:

- 从第一语音信号 (3) 的编码语音帧中提取 LSP(线频谱对),并把所述 LSP 转换成对应的 LSF(线频谱频率) ω_i , 第一语音信号从第一移动电话发送;

- 如果在第一语音信号 (3) 中围绕所述频谱峰 ω_c 的相邻 LSF 之间的距离小于第三阈值,则确定所述频谱峰是回声。

4. 按照权利要求 1-3 的任一项的方法,其特征在于只有在所述频谱峰 ω_c 的频率低于第四阈值时才确定频谱峰 ω_c 是否为回声。

5. 按照权利要求 3 的方法,其特征在于如果所述频谱峰不是回声则通过以下附加步骤搜索在第一语音信号中接近地定位的第二频谱峰:

- 从两个相邻的 LSF 计算在第一语音信号 (3) 中的第二频谱峰;

- 如果在计算的第二频谱峰与中心频率 ω_c 之间的距离小于第五阈值,则确定所述第二频谱峰是回声。

6. 按照权利要求 1-3 的任一项的方法,其中由 ω_i 表示的 LSF(线频谱频率)通过使用关系 $q_i = \cos(\omega_i)$ 把每个提取的 LSP(线频谱对) q_i 转换成对应的 LSF 而得到。

7. 按照权利要求 1-3 的任一项的方法,其中频谱峰的中心频率 ω_c 通过定位两个相邻的 LSF, $\omega_{i+1} - \omega_i$, 和计算 $\omega_c = (\omega_{i+1} - \omega_i) / 2$ 而被确定。

8. 按照权利要求 2 的方法,其特征在于双端话音的指示停止电信系统中的回声路径估计。

9. 一种估计电信系统中的回声路径的方法,其特征在于回声路径估计通过由按照权利要求 2 的方法得到的双端话音指示而被停止。

10. 一种估计电信系统中的回声路径的方法,其中当所接收编码语音信号 (6) 是由非回声信号 (5) 占支配地位时停止回声路径估计,该方法的特征在于以下步骤:

- 如果所接收语音信号 (6) 的信号功率超过第一阈值,则从所述所接收语音信号 (6) 的编码语音帧中提取 LSP(线频谱对);

- 把每个所述提取的 LSP 转换成对应的 LSF(线频谱频率) ω_i , 以及计算在每两个相邻的 LSF 之间的距离;

- 对于小于第二阈值的每个所述距离,计算由所述 LSF 围绕的频谱峰 ω_c 的频率,以及确定所述频谱峰是否为回声;

当预定数目的频谱非回声峰被定位时,停止回声路径估计。

11. 按照权利要求 10 的方法,其特征在于确定频谱峰是否为回声包括以下步骤:

- 从第一语音信号 (3) 的编码语音帧中提取 LSP(线频谱对),并把每个提取的 LSP 转换成对应的 LSF(线频谱频率) ω_i ,第一语音信号从第一移动电话发送;

- 如果在第一语音信号 (3) 中围绕所述频谱峰 ω_c 的 LSF 之间的距离小于第三阈值,则确定所述频谱峰是回声。

12. 按照权利要求 10 的方法,其特征在于只有在所述频谱峰 ω_c 的频率低于第四阈值时才确定频谱峰 ω_c 是否为回声。

13. 按照权利要求 10-12 的任一项的方法,其中语音编码是基于 AMR-CELP- 编码的。

14. 一种在电信系统中被布置成连接到编码的所接收语音信号 (6) 和编码的所发送第一语音信号 (3) 的双端话音检测器 (14),用于确定何时所述所接收语音信号 (6) 由非回声信号占支配地位,双端话音检测器的特征在于:

- 信号功率监视器 (16),被布置成确定何时所接收语音信号 (6) 的信号功率超过第一阈值;

频谱峰定位器 (18),被布置成:

- 从所述所接收语音信号的编码语音帧中提取 LSP(线频谱对),并把每个所述提取的 LSP 转换成对应的 LSF(线频谱频率) ω_i ;

- 计算在每两个相邻的 LSF 之间的距离和计算由小于第二阈值的距离所隔开的两个 LSF 所围绕的频谱峰 ω_c ;

- 回声定位器 (20),被布置成确定所定位的频谱峰 ω_c 是否为回声。

15. 按照权利要求 14 的双端话音检测器,还被布置成如果预定数目的频谱非回声峰位于接收信号 (6) 中则指示所述所接收语音信号 (6) 由非回声信号占支配地位。

16. 按照权利要求 14 的双端话音检测器,还包括计数器 (22),用于计数在所接收语音信号中频谱非回声峰的数目。

17. 按照权利要求 14-16 的任一项的双端话音检测器,其特征在于,回声定位器 (20) 被布置成:

- 从第一语音信号 (3) 的编码语音帧中提取 LSP(线频谱对),并把所述提取的 LSP 转换成对应的 LSF;

- 计算在围绕所述频谱峰 ω_c 的相邻 LSF 之间的距离,以及如果所述计算的距离小于第三阈值,则确定所述频谱峰是回声。

18. 按照权利要求 14-16 的任一项的双端话音检测器,其中所述回声定位器 (20) 被布置成只有在所述频谱峰 ω_c 的频率低于第四阈值时才确定频谱峰 ω_c 是否为回声。

19. 按照权利要求 17 的双端话音检测器,其中所述回声定位器 (20) 被进一步布置成如果所述频谱峰不是回声则通过以下步骤来搜索在第一语音信号中接近地定位的第二频谱峰:

- 从所述 LSF 计算第二频谱峰,以及如果到中心频率 ω_c 的距离小于第五阈值则确定所述第二频谱峰是回声。

20. 按照权利要求 14-16 的任一项的双端话音检测器,其中由 ω_i 表示的 LSF(线频谱频率)通过使用关系 $q_i = \cos(\omega_i)$ 把每个提取的 LSP(线频谱对) q_i 转换成对应的 LSF 而

得到。

21. 按照权利要求 14-16 的任一项的双端话音检测器,其中频谱峰的中心频率 ω_c 通过定位两个相邻的 LSF, $\omega_{i+1}-\omega_i$, 和计算 $\omega_c = (\omega_{i+1}-\omega_i)/2$ 而被确定。

22. 一种用于电信系统的回声路径估计装置 (15), 被布置成停止估计由非回声信号占支配地位的回声路径, 其特征在于包括按照权利要求 20-21 的任一项的、用于编码的语音信号的双端话音检测器 (14)。

23. 一种用于电信系统的回声路径估计装置 (15), 被布置成停止估计由非回声信号占支配地位的回声路径, 回声路径估计装置包括被布置成连接到在所述电信系统中编码的所接收语音信号 (6) 和编码的所发送第一语音信号 (3) 的双端话音检测器 (14), 其特征不在于所述双端话音检测器包括:

- 信号功率监视器 (16), 被布置成确定何时所接收语音信号 (6) 的信号功率超过第一阈值;

- 频谱峰定位器 (18), 被布置成:

- 从所述所接收语音信号的编码语音帧中提取 LSP (线频谱对), 并把每个所述提取的 LSP 转换成对应的 LSF (线频谱频率) ω_i ;

- 计算在每两个相邻的 LSF 之间的距离和计算由小于第二阈值的距离间隔开的两个 LSF 围绕的频谱峰 ω_c ;

- 回声定位器 (20), 被布置成确定所定位的频谱峰 ω_c 是否为回声。

24. 按照权利要求 23 的回声路径估计装置, 其中双端话音检测器还被布置成如果预定数目的频谱非回声峰位于接收信号 (6) 中, 则指示所述所接收语音信号 (6) 由非回声信号占支配地位。

25. 按照权利要求 24 的回声路径估计装置, 其中双端话音检测器包括计数器 (22), 用于计数在所接收信号中频谱非回声峰的数目。

26. 按照权利要求 23-25 的任一项的回声路径估计装置 (15), 其特征不在于回声定位器 (20) 还被布置成:

- 从第一语音信号 (3) 的编码语音帧中提取 LSP (线频谱对), 并把所述提取的 LSP 转换成对应的 LSF;

- 计算在围绕所述频谱峰 ω_c 的相邻 LSF 之间的距离, 以及如果所述计算的距离小于第三阈值则确定所述频谱峰是回声。

27. 按照权利要求 23-25 的任一项的回声路径估计装置, 其中回声定位器 (20) 被布置成只有在所述频谱峰 ω_c 的频率低于第四阈值时才确定频谱峰 ω_c 是否为回声。

28. 按照权利要求 26 的回声路径估计装置, 其中所述回声定位器 (20) 还被布置成如果所述频谱峰不是回声则通过以下步骤来搜索在第一语音信号中接近地定位的第二频谱峰:

- 从两个相邻的 LSF 计算第二频谱峰, 以及如果到中心频率 ω_c 的距离小于第五阈值, 则确定所述第二频谱峰是回声。

双端话音检测器

技术领域

[0001] 本发明涉及一种检测所接收编码语音信号是由非回声信号占支配地位的方法和一种估计在电信系统中的回声路径的方法。而且,本发明涉及一种双端话音(double talk)检测器以及一种包括双端话音检测器的回声路径估计装置。

背景技术

[0002] 例如由移动电话和 IP 电话进行的语音传输通常牵涉到语音编码,它是将语音压缩成一种代码以用语音编解码器进行传输。CELP(码激励线性预测)编码是通常使用的语音编码方法,其包括两个阶段(stage)即,对频谱包络(spectral envelope)进行建模的线性预测阶段和对线性预测阶段的剩余量进行建模的码本(code-book)阶段。

[0003] 除了对信号的实际语音编码以外,信道编码也可以被用于信号的传输以便避免由于传输误差造成的损失,并且在语音数据流中最重要的比特常常通过较鲁棒的信道编码进行编码以便得到最佳的总编码结果。

[0004] 重要的是减小噪声和干扰以便提高移动电话中的语音质量。回声(即话音信号反射回到讲话方)是主要的干扰,并且电话网中的主回声源是由 PSTN(公共交换电话网)的本地交换机中 4 线到 2 线转换的阻抗失配造成的、所谓的混合电路中的电反射。通常,这个电回声由在电话系统中-例如在作为在例如使用 IP(互联网协议)那样的分组交换网络与例如 PSTN 那样的电路交换网络之间的接口的媒体网关中或者在作为在移动网络与 PSTN 之间的接口的移动业务交换中心中-靠近回声源安装的网络回声消除器而去除。在国际交换机(international exchange)中也需要网络回声消除器,并且在具有大的端对端传输延时的国家电话交换机中可能需要网络回声消除器。而且,如果在靠近回声源的电话交换机中没有回声消除器,则在另一个国家的国际运营商可能想要通过检测和去除在长途电话交换机中生成的回声来减小回声。

[0005] 在移动通信网内另一个回声源是由话筒与扬声器之间的声音耦合引起的、在移动电话或 IP 电话内出现的声音串扰。为了按照标准要求减小声音耦合,移动电话通常提供回声衰减。然而,虽然移动电话提供按照要求的回声衰减,但源自声音串扰的回声仍旧可能发生,例如由于移动电话位置的很大变化或线路电平离标称电平的偏差而可能发生。

[0006] 虽然传统的网络回声消除器能够控制电回声,但源自声音串扰的回声需要不同的回声消除器。由于移动通信网中的信号在语音编码器中进行编码并且然后通过引入误码的无线信道进行发送,所以回声路径将是非线性和非稳定的并引入未知的延时。由此,传统的网络回声消除器不能处理从移动电话返回的声音回声。

[0007] 传统上,回声控制包括确定所接收语音信号是由源自接收机附近即源自所谓的近端的分量还是由源自远处即源自所谓的远端的已知语音信号的反射(回声)占支配地位。从远端反射的已知语音信号即回声将被延时、变换、并与语音信号和源自近端的噪声混合。这示意性示于图 1 中,图 1 示出了第一移动电话 1a 和第二移动电话 1b。第一语音信号 3 从第一移动电话 1a 发送,并且在到达第二移动电话 1b 之前在第一网络路径 2a 中被延时和被

变换。然而,这个语音信号的反射部分 4 将被反射并且通过第二网络路径 2b 返回以被第一移动电话 1a 接收作为已知第一语音信号 3 的回声。因此,由第一移动电话接收的这个回声信号即远端信号源自第一语音信号,且通过了两个网络路径 2a, 2b。

[0008] 从第二移动电话 1b 发送的第二语音信号 5 将被加到源自第一语音信号 3 的回声信号 4 上。因此,到达第一移动电话 1a 的接收语音信号 6 将包括回声信号分量 4(即远端信号)和这个第二语音信号分量 5(即近端信号),后者对于第一移动电话 1a 是未知的。由近端信号 5 而不是由回声信号 4 占支配地位的所接收语音信号 6 可被称为双端话音,而关于语音信号由近端信号占支配地位的确定,此后被称为双端话音检测。作为第一语音信号 3 的反射的、接收信号 6 的远端分量可以被回声控制装置抑制以便减小干扰和噪声。

[0009] 回声控制装置通常估计回声路径的特性,这个估计将被源自近端的未知语音信号所打扰。所以,传统的回声控制装置避免在有源自近端的语音的情况下估计回声路径的特性。作为替代,回声控制装置将通过上述的双端话音检测来检测近端语音的存在,并且将在所接收信号是由近端话音占支配地位时的时段内停止(inactivate)或禁止估计回声路径特性。

[0010] 双端话音检测例如可以通过比较近端分量与远端分量的信号电平以便检测双端话音而被执行,诸如例如由如例如在以下文献中描述的 Geigel 检测器 :D. L. Duttweiler, "A twelve-channel digital echo canceller", IEEE Transactions on Communications, Vol. COM-26, No. 5, May 1978。然而,这种双端话音检测的准确性是相当低的,因为它假设回声信号功率总是低于恒定倍数的远端信号功率,以及如果从近端返回的信号具有高于恒定倍数的远端信号功率的短期功率,则声称是双端话音。由此,检测器将遗漏由于线路电平的差异或由于近端扬声器以低于远端扬声器的话音进行谈话所引起的任何弱的双端话音状况。另外,这个常数可能很难确定,特别是对于声音回声,由于在回声路径中的放大,它可以强于引起它的远端信号。

[0011] 可替代地,双端话音检测包括计算近端分量和远端分量的互相关、协方差或相干函数,如例如在美国专利 No. 6, 035, 034 和 No. 6, 766, 019 中描述的。这导致改进的检测性能,但需要更高的计算复杂性。

[0012] 如上所述,在移动通信网中的语音信号通常以编码的格式被传输,AMR(自适应多速率)是针对语音编码所优化的音频数据压缩方案的一个例子。AMR 通常被使用来对 GSM(全球移动通信系统)和 UMTS(通用移动通信系统)网络中的语音信号进行编码,并且它涉及基于链路条件从八个不同的比特速率之一中选择的链路适配。AMR 可以使用不同的技术,诸如例如上述的 CELP、或 DTX(非连续传输)、VAD(Voice Activity Detection, 语音活动检测)或 CNG(Comfort Noise Generation, 舒适噪声生成),并且链路适配可以选择最好的编解码模式来满足本地无线电信道和容量要求。在很差的无线电传输的情形下,信道编码将增加,这将改进网络连接的质量和鲁棒性,但将导致恶化的话音信号。

[0013] 同样地,IP 电话语音信号通常在发送的移动电话中被编码并且通过网络被传输到另一个移动终端/电话,而不用网络中的任何解码。

[0014] 因此,网络回声控制将必须被施加到编码的信号,优选地,通过直接修改在编码的比特流中的参数,而不用解码信号且不用在去除回声后执行第二编码,因为解码随后编码可能破坏 TFO(Tandem Free Operation, 免级联运行)和 TrFO(Transcoder Free

Operation, 无转码器运行) 的肯定的语音质量效果, TFO 和 TrFO 通常在现代电信网中被引用以便增强语音质量。

[0015] 传统的双端话音检测中的附加缺点在于, 为了计算检测变量需要信号波形, 要求在检测之前解码语音信号。然而, 由于为了增强语音质量而使用 TrFO(无转码器运行) 和 TFO(免级联运行), 对编码的比特流直接进行操作的能力变得越来越重要, 因为解码随后编码会降低 TFO(免级联运行) 和 TrFO(无转码器运行) 的肯定的语音质量效果。

[0016] 而且, 由于网络回声控制通常牵涉到双端话音检测, 即确定所接收语音信号由近端信号占支配地位, 所以改进的双端话音检测将改进网络回声控制。

[0017] 所以, 为实现可应用于编码的语音信号的改进且准确的双端话音检测, 仍旧存在问题。

发明内容

[0018] 本发明的目的是解决上述的问题, 这个目的和其它目的是通过按照所附权利要求的方法和设备实现的。

[0019] 按照第一方面, 本发明提供一种检测所接收编码语音信号是由非回声信号占支配地位的方法, 该方法包括以下步骤:

[0020] - 如果所接收语音信号的信号功率超过第一阈值, 则从所述所接收语音信号的编码语音帧中提取 LSP (Line Spectral Pair, 线频谱对);

[0021] - 把每个所述提取的 LSP 转换成 LSF (线频谱频率) ω_i , 以及计算在每两个相邻的 LSF 之间的距离;

[0022] - 对于小于第二阈值的每个所述距离, 计算由所述 LSF 围绕的频谱峰 (spectral peak) ω_c 的频率, 以及确定所述频谱峰是否是回声。

[0023] 由此, 完成可靠的双端话音检测, 即所接收语音信号由非回声信号占支配地位。由于为了得到所需要的参数只需要部分解码, 例如 LSF (线频谱频率) 通过对从编码的语音信号中提取的 LSP (线频谱对) 进行转换而得到, 所以本发明可直接应用于编码的语音和非线性回声路径。

[0024] 方法可包括以下进一步的步骤:

[0025] - 对于不是回声的每个定位的频谱峰, 对计数器进行增量;

[0026] - 当计数器达到预定的阈值时指示双端话音。

[0027] 确定频谱峰是否为回声可包括以下步骤:

[0028] - 从第一语音信号的编码语音帧中提取 LSP (线频谱对), 并把所述 LSP 转换成 LSF (线频谱频率) ω_i ;

[0029] - 如果在第一语音信号中围绕所述频谱峰 ω_c 的相邻 LSF 之间的距离小于第三阈值, 则确定所述频谱峰是回声。

[0030] 而且, 该方法可以只有在所述频谱峰 ω_c 的频率低于第四阈值时才确定频谱峰 ω_c 是否为回声, 以便改进噪声容限 (tolerance)。

[0031] 为了提高准确性, 如果频谱峰不是回声, 则可以通过以下步骤在第一语音信号中搜索接近地 (closely) 定位的第二频谱峰:

[0032] - 从两个相邻的 LSF 计算在第一语音信号中的第二频谱峰;

[0033] - 如果在计算的第二频谱峰与中心频率 ω_c 之间的距离小于第五阈值, 则确定所述第二频谱峰是回声。

[0034] 由 ω_i 表示的 LSF (线频谱频率) 可以通过使用关系 $q_i = \cos(\omega_i)$ 把每个提取的 LSP (线频谱对) q_i 转换成对应的 LSF 而从 LSP (线频谱对) q_i 而得到, 并且频谱峰的中心频率 ω_c 可以通过定位两个相邻的 LSF, $\omega_{i+1} - \omega_i$, 和计算 $\omega_c = (\omega_{i+1} - \omega_i)/2$ 而被确定。而且, 频谱编码可以是基于 AMR-CELP- 编码 (AMR-CELP-coding)。

[0035] 按照第二方面, 本发明提供一种在电信系统中回声路径估计的方法, 以及如由按照第一方面的方法检测的, 当所接收编码语音信号由非回声信号占支配地位时停止 (deactivate) 回声路径估计。

[0036] 按照第三方面, 本发明提供一种双端话音检测器, 其在电信系统中被布置成连接到编码的所接收语音信号和编码的所发送第一语音信号以便确定何时所述所接收语音信号由非回声信号占支配地位。双端话音检测器包括:

[0037] - 信号功率监视器, 被布置成确定何时所接收语音信号的信号功率超过第一阈值;

[0038] - 频谱峰定位器, 被布置成:

[0039] - 从所述所接收语音信号的编码语音帧中提取 LSP (线频谱对), 并把每个所述提取的 LSP 转换成对应的 LSF (线频谱频率) ω_i ;

[0040] - 计算在每两个相邻的 LSF 之间的距离和计算由小于第二阈值的距离所隔开的两个 LSF 所围绕的频谱峰 ω_c ;

[0041] - 回声定位器, 被布置成确定所定位的频谱峰 ω_c 是否为回声。

[0042] 双端话音检测器还可被布置成如果预定数目的频谱非回声峰位于接收信号中则指示所述所接收语音信号由非回声信号占支配地位, 以及可包括用于计数在所接收语音信号中频谱非回声峰的数目的计数器。

[0043] 所述回声定位器可被布置成:

[0044] - 从第一语音信号的编码语音帧中提取 LSP (线频谱对), 并把所述提取的 LSP 转换成对应的 LSF;

[0045] - 计算在围绕所述频谱峰 ω_c 的相邻 LSF 之间的距离, 以及如果所述计算的距离小于第三阈值, 则确定所述频谱峰是回声。

[0046] 所述回声定位器还可被布置成: 只有在所述频谱峰 ω_c 的频率低于第四阈值时才确定频谱峰 ω_c 是否为回声; 以及如果所述频谱峰不是回声, 则通过从所述 LSF 计算第二频谱峰, 来搜索在第一语音信号中接近地定位的第二频谱峰, 以及如果到中心频率 ω_c 的距离小于第五阈值则确定所述第二频谱峰是回声。

[0047] 按照第四方面, 本发明提供用于电信系统的回声路径估计装置, 被布置成停止估计由非回声信号占支配地位的回声路径, 回声路径估计装置包括按照第三方面的双端话音检测器, 该双端话音检测器被布置成连接到在所述电信系统中编码的所接收语音信号和编码的所发送第一语音信号。

[0048] 因此, 按照本发明的双端话音检测将通过当双端话音检测器指示所接收信号由不是回声的近端信号占支配地位时使得能实现停止在编码语音信号中回声路径估计而提供改进的回声控制。

附图说明

- [0049] 现在将参照附图更详细地描述本发明,其中:
- [0050] - 图 1 是示意地说明在两个移动电话之间的通信时的回声路径的框图;
- [0051] - 图 2 说明传统的 AMR-CELP-合成 (AMR-CELP-synthesis);
- [0052] - 图 3 图形说明 $1/A(z)$ 的频率特性和 $A(z)$ 的根的位置;
- [0053] - 图 4 图形说明 $A(z)$ 的复数根 (complex root) 和对应的 LSF;
- [0054] - 图 5 是说明按照本发明的实施例的双端话音检测的流程图;
- [0055] - 图 6 是按照本发明的、配备有双端话音检测器的、按照图 1 的框图;以及
- [0056] - 图 7 是说明按照本发明的双端话音检测器以及包括这样的双端话音检测器的回声路径估计的框图。

具体实施方式

[0057] 在以下的说明中,阐述具体的细节,诸如特定的体系结构和步骤序列,以便提供对本发明的透彻了解。然而,本领域技术人员可以看到,本发明可以以可能不同于这些具体细节的其它实施例来实践。

[0058] 而且,可以看到,所描述的功能可以使用结合编程的微处理器或通用计算机起作用的软件和 / 或使用专用集成电路来实施。在本发明以方法的形式被描述的情况下,本发明也可以被体现在计算机程序产品中以及在包括计算机处理器和存储器的系统中,其中存储器用可以执行所描述的功能的一个或多个程序来编码。

[0059] 参照图 1,本发明的概念是通过检查通过转换从编码语音帧中提取的 LSP (线频谱对) 而得到的 LSF (线频谱频率),从而检测所接收语音信号 6 由非回声信号 (即由来自近端的未知语音 / 噪声分量 5) 占支配地位。LSF 在包括所述未知近端分量 5 和来自远端的回声分量 4 的接收信号 6 中和直接在所述第一语音信号 3 中被检查,所述回声分量 4 是从第一移动电话 1a 所发送的第一语音信号的反射。按照本发明,当在所接收语音信号 6 的信号功率是较大时,在所接收语音信号 6 中检测到预定数目的一个或多个频谱非回声峰时,确定所接收语音信号 6 由非回声信号占支配地位,即存在双端话音。更具体地,如果在所接收语音信号 6 中 LSP 的检查揭示在两个相邻的 LSF 之间的距离小于阈值以及如果在第一语音信号 3 中没有检测到对应的频谱峰而同时所接收语音信号 6 的功率超过功率阈值,则检测到频谱非回声峰。

[0060] 按照本发明,包括双端话音检测器的回声路径估计装置被布置成当双端话音检测器指示存在双端话音时禁止估计回声路径的特性,以避免由双端话音引起的干扰。

[0061] 因此,为了区分未知近端信号 5 的存在与作为已知信号的反射的远端回声信号 4,本发明确定在第一语音信号 3 与所接收信号 6 之间的非相似性。所接收信号 6 包括源自第一移动电话 1a 发送的已知第一语音信号 3 的反射回声分量 4 和来自近端的即源自第二移动电话 1b 的未知第二语音 / 噪声信号 5。非相似性通过定位在所接收信号 6 中和在第一语音信号 3 中的频谱峰以及比较在各个信号中峰的位置而被确定。语音信号中的频谱峰通过提取在编码的语音信号中以 q_i 表示的 LSP (线频谱对) 以及使用关系 $q_i = \cos(\omega_i)$ 把 LSP 转换成以 ω_i 表示的 LSF (线频谱频率) 而得到。从包括第一语音信号 3 的反射分量 4

和第二语音 / 噪声信号 5 的接收信号 6 中以及直接在第一语音信号 3 中提取 LSP, 然后提取的 LSP 被转换成 LSF。如果关于 LSF 和关于接收信号 6 的信号功率的某些条件被满足, 则将指示双端语音。

[0062] 图 2 是示意地说明按照示例性实施例的、本发明可应用于其的、基于 CELP (码激励线性预测编码) 合成模型的 AMR (自适应多速率) 语音解码器的框图。然而, 本发明的另外实施例可应用于其它语音编解码器, 特别是基于 CELP 合成模型的那些编解码器。在图 2 中, 通过将来自自适应码本 7 (adaptive codebook) 的、以音高增益 (pitch gain) 进行缩放 (在 12 中) 的自适应码向量和来自固定码本 8 的、以革新增益 (innovative gain) 进行缩放 (在 11 中) 的固定码向量相加, 并通过在 LP (线性预测) 合成滤波器 9 中滤波激励信号以重建语音, 来构建激励信号。最后, 重建的语音信号传送通过自适应后置滤波器 10。

[0063] AMR 编解码器通常被应用到具有 20ms 长度、包括 160 个样本的帧, 并且每个帧被划分成具有 5ms 长度和 40 个样本的子帧, 在比特流中可得到的参数是 LSP (线频谱对) 向量、分数音高滞后 (fractional pitch lag) (pitch period, 音高周期)、革新的码向量、以及音高增益和革新增益。LSP 向量承载关于语音的 AR (自回归) 模型的信息, 并且代表以“量化友好 (quantization friendly)”形式的线性预测 (LP) 合成多项式系数。LSP 向量精确地承载语音信号的、与以下的 LP 多项式表达相同的频谱信息:

$$[0064] \quad A(z) = 1 + \sum_{k=1}^K a_k z^{-k}$$

[0065] 在这个多项式中, a_k 表示 LP (线性预测) 系数, 以及 z 是复变量。出于量化和内插目的, LP 滤波器系数 ($a_k, k = 1, \dots, 10$) 可转换到 LSP (线频谱对) 表示形式, 以及对于 10 阶 LP 滤波器, LSP 被定义为分别是以下的求和多项式 (sum polynomial) 与差值 (difference) 多项式的根:

$$[0066] \quad F'_1(z) = A(z) + z^{-11}A(z^{-1}) \text{ 以及 } F'_2(z) = A(z) - z^{-11}A(z^{-1})$$

[0067] 以及多项式 $F'_1(z)$ 和 $F'_2(z)$ 分别是对称和反对称的。这些多项式的每个根都位于单位圆上并且是交替的, 以使得 $F'_1(z)$ 具有根 $z = -1$ ($\omega = \pi$) 而 $F'_2(z)$ 具有根 $z = 1$ ($\omega = 0$)。为了消去 (eliminate) 这两个根, 定义两个新的多项式, 即:

$$[0068] \quad F_1(z) = F'_1(z) / (1+z^{-1}) \text{ 以及 } F_2(z) = F'_2(z) / (1-z^{-1})$$

[0069] 每个多项式具有在单位圆 ($e^{+j\omega_i}$) 上的五个共轭根, 所以多项式可被写为:

$$[0070] \quad F_1(z) = \prod_{i=1,3,\dots,9} (1 - 2q_i z^{-1} + z^{-2}) \text{ 以及}$$

$$[0071] \quad F_2(z) = \prod_{i=2,4,\dots,10} (1 - 2q_i z^{-1} + z^{-2})$$

[0072] 所谓的线频谱频率 LSF, ω_i 可被用来定位语音帧的频谱峰以及在存在于编码的语音信号中的以 q_i 表示的 LSP (线频谱对) 之间的关系, LSF 用 $q_i = \cos(\omega_i)$ 描述, 以及 ω_i 满足排序特性 $0 < \omega_1 < \omega_2 < \dots < \omega_{10} < \pi$ 。

[0073] LP 滤波器把声道建模为自回归 (all-pole, 全极点) 滤波器并且能够表示多达五个频谱峰。语音信号通常具有有峰的短时频谱 (peaky short-time spectrum), 图 3 中的有峰线图说明在典型的 AMR 语音帧中 $1/A(z)$ 在从 0 到 0.9 的归一化频率上的频率特性。

LSF, ω_i 的频率位置用 x 符号表示, 接近地间隔开的 LSF 围绕 $A(z)$ 的对应根和频谱峰的每个位置。所述根在说明它们的角位置的图上用小圆表示, $A(z)$ 的每个根对应于由两个接近地间隔开的 LSF 围绕的所示频谱峰。应当指出, $A(z)$ 的根将以复共轭对 (complex conjugate pair) 出现, 但在图 3 中仅仅说明了对应于正频率的根。

[0074] 图 4 图形说明被表示为小圆的、 $A(z)$ 的所述复数根的实部和虚部。LSF 在图中用小的 x 符号表示, 并且两个接近地间隔开的 LSF 围绕单位圆上前四个根中的每个根, 这四个根对应于图 3 所示的四个根。图 4 揭示了在 LP 滤波器频率响应中的频谱峰的狭窄度 (narrowness) 确定单位圆上对应的 $A(z)$ 根的接近度 (closeness)。图 3 中的第一个根 (对应于图 4 中的根 $z \approx 0.8 + j0.33$) 定位得更接近于图 4 的单位圆并且与图 3 中的第四个根 (对应于图 4 中的根 $z \approx -0.75 + j0.39$) 相比具有更接近地间隔开的 LSF。因此, LSF 常常处在单位圆上并且定位得围绕 $A(z)$ 的根。另外, 在围绕接近于单位圆的 $A(z)$ 根的两个 LSF 之间的距离小于在围绕更远离单位圆的 $A(z)$ 根的两个 LSF 之间的距离。本发明在检测双端话音时使用如在图 3 和 4 中表示的、在两个相邻的 LSF 之间的距离和对应的根 / 频谱峰的关系, 即所接收语音信号由来自近端的未知语音 / 噪声分量占支配地位, 而不是由来自远端的回声分量占支配地位。

[0075] 因此, 本发明通过使用在频谱峰的高度与在围绕峰的 LSF 之间的距离之间的关系来检测双端话音, 并且通过提取编码的 LSP, q_i , 和根据关系 $q_i = \cos(\omega_i)$ 把 LSP 转换成对应的 LSF 来检查在接收的语音帧中的 LSF, ω_i 。如果在所接收信号 6 中的信号功率很大并且如果相邻的且接近地间隔开的 LSF 指示在所接收信号 6 中存在一个或多个频谱峰且不对应于在第一语音信号 3 中的任何频谱峰, 则将声称双端话音。

[0076] 按照本发明的第一示例性实施例, 双端话音的检测, 即所接收编码语音信号由非回声信号占支配地位, 是通过监视所接收信号 6 的功率而执行的, 并且当这个功率超过第一阈值 (例如 -45dBm0) 时, 在所接收信号 6 中的编码语音帧的 LSP 被提取并被转换成对应的 LSF。计算在每两个相邻的 LSF 之间的距离 $\Delta i = \omega_{i+1} - \omega_i$, 并且如果这个距离小于第二阈值 (例如 0.03π), 则在所述两个 LSF 之间定位的频谱峰的中心频率 ω_c 被计算为 $\omega_c = (\omega_{i+1} - \omega_i) / 2$ 。此后, 通过定位在第一语音信号 3 中的任何对应的频谱峰而确定在所接收语音信号 6 中所定位的频谱峰是否为回声。按照这个第一示例性实施例, 通过从第一语音信号 3 的编码语音帧中提取 LSP、把 LSP 转换成对应的 LSF 和定位围绕接收信号的所述频谱峰的中心频率 ω_c 的 LSF 对, 来检查第一语音信号 3。如果在这个 LSF 对之间的距离大于第三阈值 (例如 $4 \cdot \Delta i$), 则确定在第一语音信号 3 中即在已知的远端信号中没有存在对应的频谱峰, 以及在接收信号中在 ω_c 处的频谱峰不是回声信号。

[0077] 当例如通过计数器达到预定值而发现在接收信号 6 的语音帧中预定数目的非回声频谱峰时, 确定所接收信号 6 由未知近端信号 5 占支配地位, 并将声称双端话音。

[0078] 按照本发明的第二实施例, 通过比较在接收信号 6 中所定位的频谱峰的频率 ω_c 与第四阈值 (例如 0.8π) 以及仅在所定位的频谱峰的频率 ω_c 低于这个阈值时才检查第一语音信号 3, 来达到更好的噪声容限。否则, 在频率 ω_c 处的频谱峰被假设为由噪声引起的。

[0079] 按照本发明的第三实施例, 在第一语音信号 3 中没有发现精确对应的频谱峰的情形下, 通过定位在第一语音信号 3 中具有稍微高于或低于接收信号 6 的频谱峰 ω_c 的频率的、任何存在的频谱峰, 来达到更高的准确性。如果能找到这样的频谱峰, 并且在这

个频谱峰的频率与接收信号的频谱峰 ω_c 之间的差值小于第五阈值,则在接收信号中所定位的频谱峰 ω_c 将被假设为是回声,并且不会对双端话音计数器进行增量。

[0080] 按照本发明的回声路径估计装置包括按照本发明的实施例的双端话音检测器,并且该回声路径估计装置被布置成在双端话音检测器指示存在双端话音时,即接收信号由源自近端的非回声信号占支配地位,而不是由源自远端的回声信号占支配地位时,禁止回声路径估计。

[0081] 图 5 是说明按照本发明的上述第一实施例的、检测双端话音的方法的流程图。首先,在步骤 51,测量接收信号 6 的功率,并且在步骤 53,如果它超过第一阈值,则将从编码的语音帧中提取 LSP。LSF 将通过把每个 LSP 转换成对应的 LSF 而获得,并且将在步骤 54 中计算在每两个相邻的 LSF 之间的距离。对于小于第二阈值的每个距离,在步骤 55,计算位于在 LSF 之间的频谱峰的中心频率。在步骤 56,从第一语音信号 3 中提取 LSP,并把它转换成 LSF,随后确定围绕位于接收信号 6 中的每个频谱峰频率的、第一语音信号 3 的相邻 LSF 以及计算它们之间的距离。在步骤 57,确定每个计算的距离是否大于第三阈值,由此指示没有对应的频谱峰定位在第一语音信号 3 中,以及在接收信号中的频谱峰不是回声。对于不是回声的每个频谱峰,在步骤 59,将对双端话音计数器进行增量。如果双端话音计数器达到预定的阈值,则将指示为双端话音,这就可以禁止在包括这样的双端话音检测器的回声控制装置中的任何回声路径估计。

[0082] 图 6 是说明双端话音检测器 14 的位置和从第一移动电话 1a 发送的、通过在 2a 中延时和变换、到达第二移动电话 1b 的第一语音信号 3 的路径的框图。这个第一语音信号 3 的回声 4 连同源自第二移动电话 1b 附近的第二未知语音 / 噪声信号 5 一起被发回到第一移动电话 1a,组合的接收信号 6 在到达第一移动电话 1a 之前通过 2b 中的延时和变换。按照本发明的示例性实施例的双端话音检测器 14 被布置成监视接收信号 6 的信号功率,以及当信号功率超过功率阈值时提取在第一语音信号 3 中和在接收信号 6 中的 LSP (线频谱对) q_i 。通过使用关系 $q_i = \cos(\omega_i)$ 把 LSP 转换成 LSF (线频谱频率) ω_i ,并且将检查 LSF 以便检测在接收信号 6 中的、不是源自第一语音信号 3 的回声的频谱峰。如果在接收信号 6 的语音帧中检测到预定数目的非回声频谱峰,则确定所接收信号由源自近端的非回声信号占支配地位,并且将指示双端话音。在指示双端话音时,可以禁止在配备有双端话音检测器 14 的回声路径估计装置 15 中的回声路径估计。

[0083] 图 7 是说明被连接到能够估计电信系统中的回声路径的回声路径估计装置 15 的、按照本发明的示例性实施例的双端话音检测器 14 的功能特征的框图。

[0084] 按照本发明的这个实施例的双端话音检测器 14 包括信号功率监视器 16,该信号功率监视器 16 被布置成当测量的信号功率超过第一阈值时开始定位在所接收语音信号 6 中的频谱峰。双端话音检测器 14 包括频谱峰定位器 18,该频谱峰定位器 18 被布置成通过提取 LSP, q_i 和使用关系 $q_i = \cos(\omega_i)$ 把提取的 LSP 转换成对应的 LSF, ω_i 而定位在所接收语音信号中的任何频谱峰。此后,计算在每两个相邻的 LSF 之间的距离并与第二阈值进行比较,并且对于小于所述第二阈值的每个距离,计算频谱峰 ω_c 的频率。而且,双端话音检测器包括回声定位器 20,用于确定在所接收语音信号中所定位的每个频谱峰 ω_c 是否为回声。当在所接收语音信号中定位预定数目的非回声频谱峰时,将指示双端话音并且将停止回声路径估计。

[0085] 按照本发明的这个示例性实施例,回声定位器 20 被布置成通过从第一语音信号的编码的语音帧中提取 LSP、把 LSP 转换成对应的 LSF 和定位围绕 ω_c 的两个 LSF,而确定所接收语音信号的频谱峰 ω_c 是否为回声。此后,计算这两个 LSF 之间的距离,并且如果这个距离大于第三阈值,则确定在第一语音信号中不存在对应的频谱峰以及在所接收语音信号中所定位的频谱峰 ω_c 是非回声频谱峰而不是回声。

[0086] 按照本发明的、检测所接收语音信号由非回声信号占支配地位的方法、以及双端话音检测器和回声估计装置允许可靠地检测双端话音,并可应用于非线性回声路径以及编码的语音,因为为了得到需要的参数(例如来自从编码的语音信号提取的 LSP(线频谱对)的 LSF(线频谱频率)),只需要部分解码。因此,按照本发明的双端话音检测器通过当双端话音检测器指示所接收信号由不是回声的近端信号占支配地位时使能停止回声路径估计而将提供改进的回声控制。

[0087] 虽然本发明是参照具体的示例性实施例描述的,但该描述通常只打算说明本发明的概念,而不应当看作为限制本发明的范围。

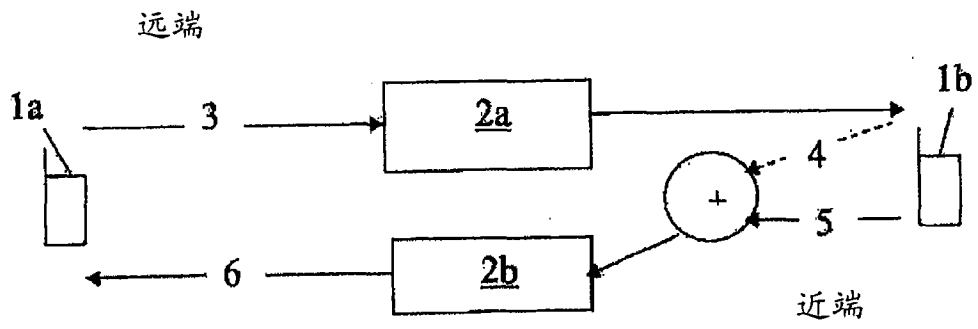


图 1

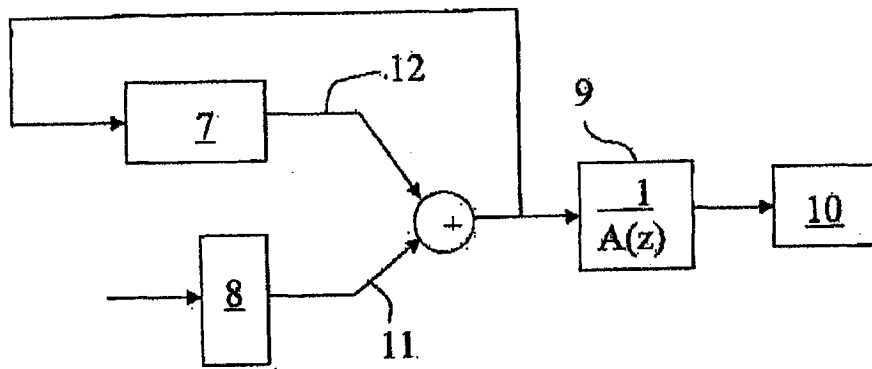


图 2

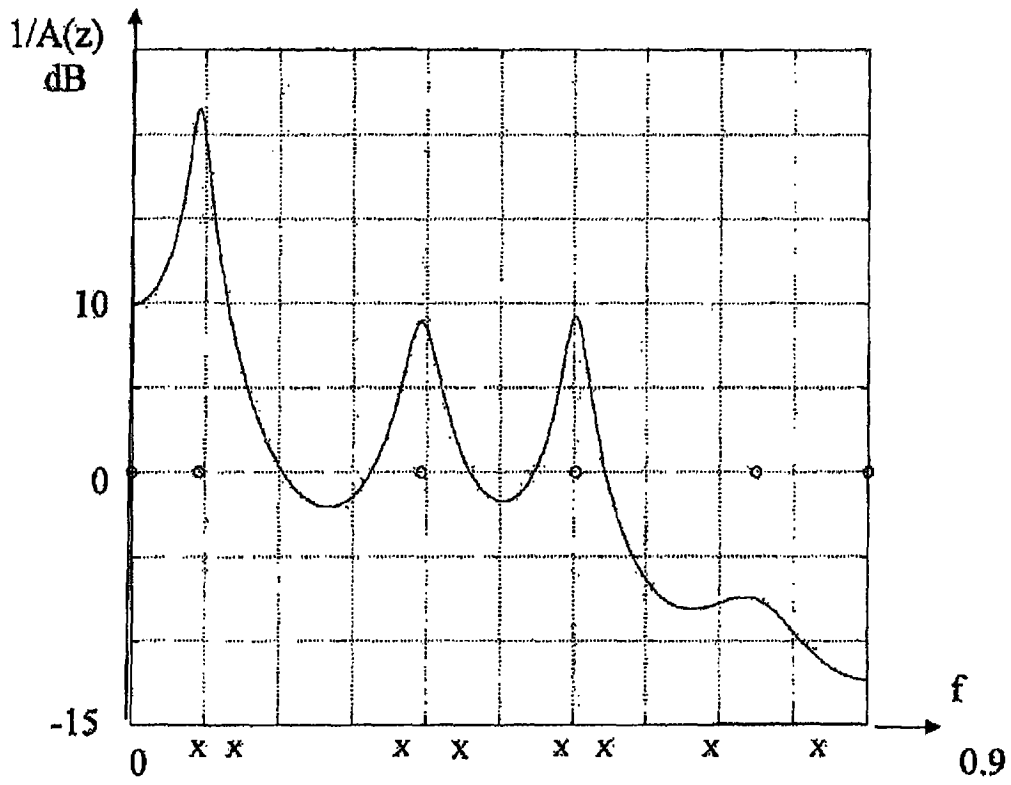


图 3

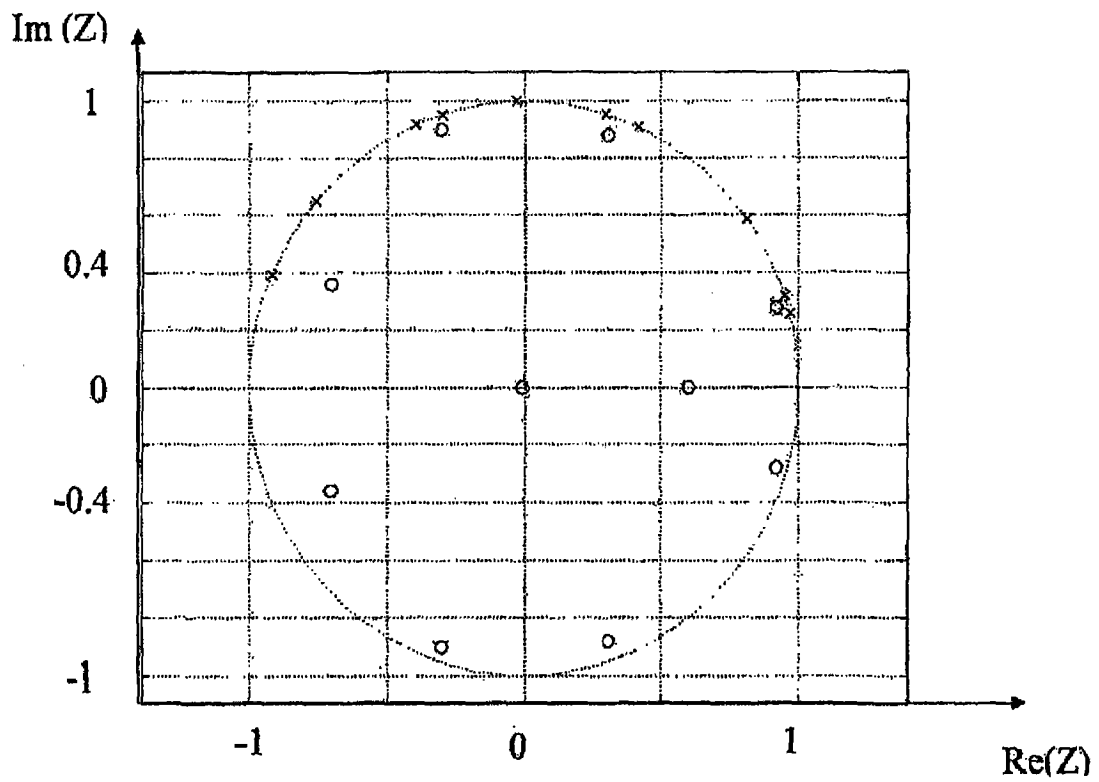


图 4

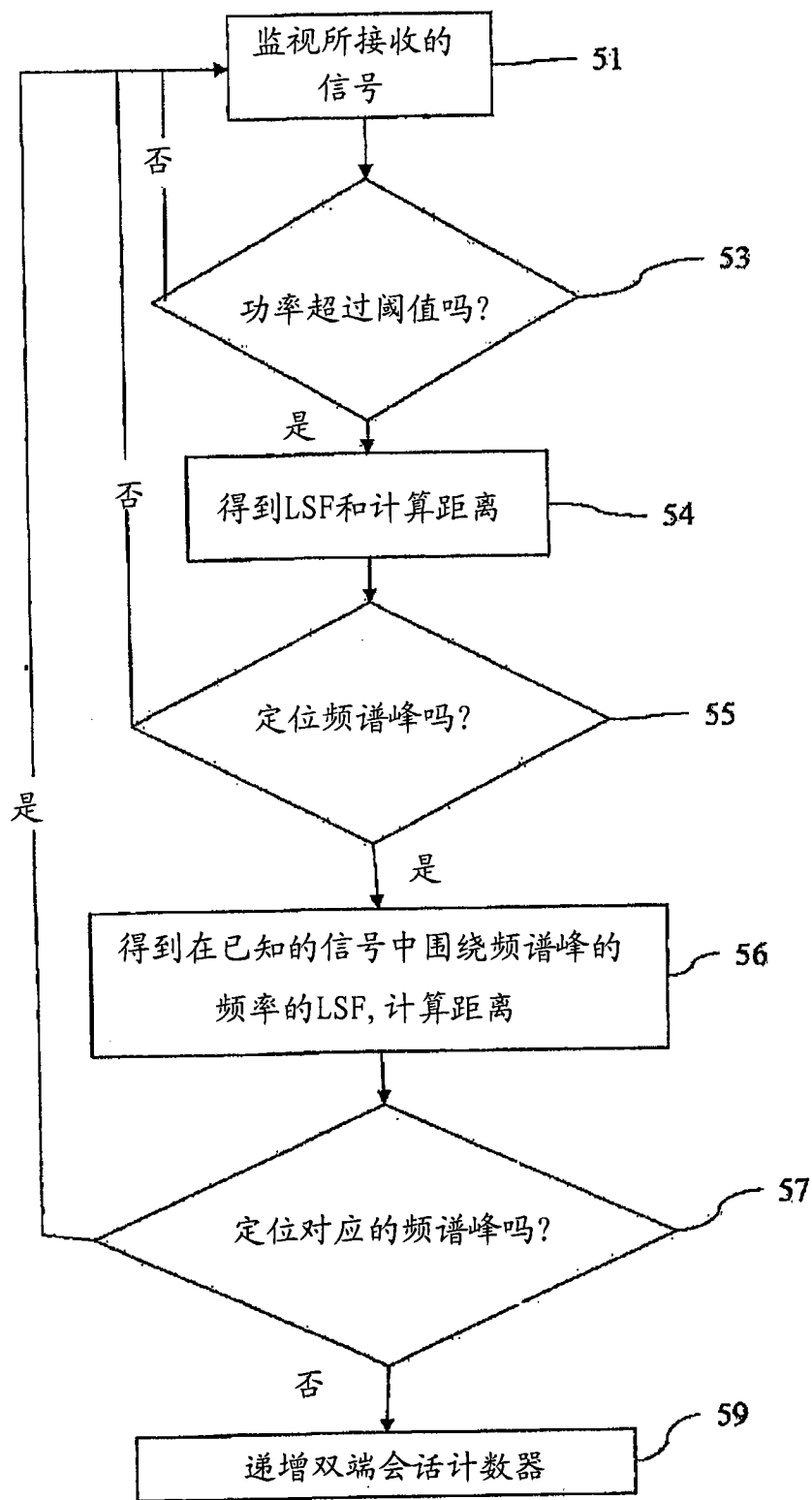


图 5

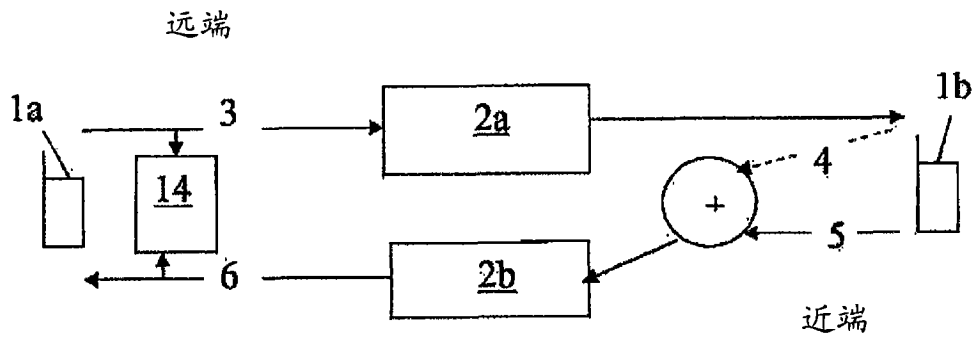


图 6

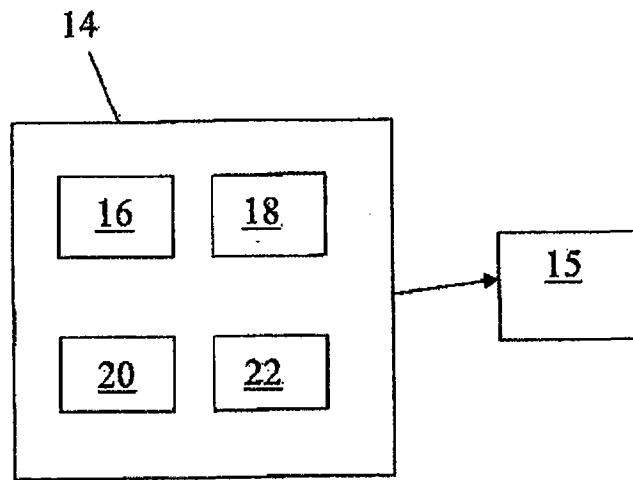


图 7