

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
G10L 21/02 (2006.01)



# [12] 发明专利申请公布说明书

[21] 申请号 200680000799.8

[43] 公开日 2007年10月24日

[11] 公开号 CN 101061535A

[22] 申请日 2006.6.30

[21] 申请号 200680000799.8

[30] 优先权

[32] 2005.7.13 [33] DE [31] 102005032724.9

[86] 国际申请 PCT/EP2006/063742 2006.6.30

[87] 国际公布 WO2007/073949 德 2007.7.5

[85] 进入国家阶段日期 2007.3.12

[71] 申请人 西门子公司

地址 德国慕尼黑

[72] 发明人 B·盖瑟 P·贾克斯 S·尚德尔

H·塔德伊 A·特勒 P·瓦里

[74] 专利代理机构 中国专利代理(香港)有限公司  
代理人 刘春元 魏 军

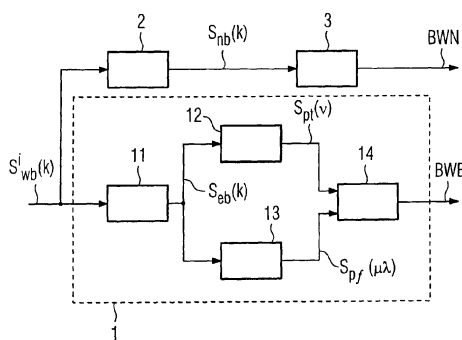
权利要求书 3 页 说明书 14 页 附图 1 页

## [54] 发明名称

用于人工扩展语音信号的带宽的方法和装置

## [57] 摘要

一种用于人工扩展语音信号的带宽的方法，其具有以下步骤：a) 提供宽带的输入语音信号 ( $s_{wb}^i(k)$ )；b) 从宽带输入语音信号 ( $s_{wb}^i(k)$ ) 的扩展频带中确定扩展带宽所需要的宽带输入语音信号 ( $s_{wb}^i(k)$ ) 的信号分量 ( $s_{eb}(k)$ )；c) 确定用于扩展带宽的信号分量 ( $s_{eb}(k)$ ) 的时间包络；d) 确定用于扩展带宽的信号分量 ( $s_{eb}(k)$ ) 的频谱包络；e) 对时间包络和频谱包络的信息进行编码，并提供经过编码的信息来用于扩展带宽；f) 对经过编码的信息进行解码，并从经过编码的信息中产生时间包络和频谱包络以用于产生扩展了带宽的输出语音信号 ( $s_{wb}^o(k)$ )。本发明还涉及用于人工扩展语音信号的带宽的装置。



1. 一种用于人工扩展语音信号的带宽的方法，其特征在于以下步骤：

a) 提供宽带的输入语音信号 ( $s_{wb}^i(k)$ )；

b) 从宽带输入语音信号 ( $s_{wb}^i(k)$ ) 的扩展频带中确定扩展带宽所需要的宽带输入语音信号 ( $s_{wb}^i(k)$ ) 的信号分量 ( $s_{eb}(k)$ )；

c) 确定用于扩展带宽的信号分量 ( $s_{eb}(k)$ ) 的时间包络；

d) 确定用于扩展带宽的信号分量 ( $s_{eb}(k)$ ) 的频谱包络；

e) 对时间包络和频谱包络的信息进行编码，并提供经过编码的信息来用于扩展带宽；

f) 对经过编码的信息进行解码，并从经过编码的信息中产生时间包络和频谱包络以用于产生扩展了带宽的输出语音信号 ( $s_{wb}^o(k)$ )。

2. 根据权利要求1所述的方法，其特征在于，所述扩展带宽所需要的信号分量 ( $s_{eb}(k)$ ) 通过滤波、尤其是带通滤波从宽带输入语音信号 ( $s_{wb}^i(k)$ ) 中确定。

3. 根据权利要求1或2所述的方法，其特征在于，在步骤c) 中对时间包络的确定与在步骤d) 中对频谱包络的确定无关地进行。

4. 根据上述权利要求之一所述的方法，其特征在于，在步骤e) 中对时间包络和频谱包络编码之前对时间包络和频谱包络进行量化。

5. 根据上述权利要求之一所述的方法，其特征在于，在用于确定频谱包络的步骤d) 中确定用于扩展带宽的信号分量 ( $s_{eb}(k)$ ) 的频谱子带的信号功率 ( $P_f(\mu, \lambda)$ )。

6. 根据权利要求5所述的方法，其特征在于，为了确定所述频谱子带的信号功率 ( $P_f(\mu, \lambda)$ )，产生用于扩展带宽的信号分量 ( $s_{eb}(k)$ )，其中尤其对该信号分量进行特殊的变换，尤其是FF变换。

7. 根据上述权利要求之一所述的方法，其特征在于，在用于确定时间包络的步骤c) 中确定用于扩展带宽的信号分量的时间信号段的信号功率 ( $P_t(v)$ )。

8. 根据上述权利要求之一所述的方法，其特征在于，在步骤f)

中对已编码的信息进行解码以重构地形成时间包络和频谱包络。

9. 根据上述权利要求之一所述的方法, 其特征在于, 激励信号 ( $s_{exc}(k)$ ) 在解码器 (5) 中从传送给该解码器 (5) 的信号 ( $s_{si}(k)$ ) 中产生, 其中所传送的信号 ( $s_{si}(k)$ ) 在对应于宽带输入语音信号 ( $s_{wb}^i(k)$ ) 的扩展频带频率范围的频率范围内具有这样的信号强度, 即该信号强度使得可以产生激励信号 ( $s_{exc}(k)$ )。

10. 根据权利要求 9 所述的方法, 其特征在于, 向所述解码器 (5) 传送经过调制的窄带信号以产生激励信号 ( $s_{exc}(k)$ ), 该窄带信号具有在宽带输入语音信号的扩展频带之下的频带范围。

11. 根据权利要求 9 或 10 所述的方法, 其特征在于, 所述激励信号 ( $s_{exc}(k)$ ) 具有传送给所述解码器 (5) 的信号 ( $s_{si}(k)$ ) 的基频的谐波。

12. 根据权利要求 8 和 11 所述的方法, 其特征在于, 从经过解码的时间包络和激励信号 ( $s_{exc}(k)$ ) 的信息中确定第一校正系数 ( $g_1(k)$ )。

13. 根据权利要求 12 所述的方法, 其特征在于, 从第一校正系数 ( $g_1(k)$ ) 和激励信号 ( $s_{exc}(k)$ ) 中重构地形成时间包络, 尤其是通过将第一校正系数 ( $g_1(k)$ ) 与激励信号 ( $s_{exc}(k)$ ) 相乘。

14. 根据权利要求 13 所述的方法, 其特征在于, 对时间包络的重构形式进行滤波, 并在滤波器中产生脉冲响应 ( $h(k)$ )。

15. 根据权利要求 14 所述的方法, 其特征在于, 从所述脉冲响应 ( $h(k)$ ) 和时间包络的重构形式中重构地形成频谱包络。

16. 根据权利要求 15 所述的方法, 其特征在于, 从频谱包络的重构形式中重构出宽带输入语音信号 ( $s_{wb}^i(k)$ ) 的扩展频带的信号分量 ( $s_{eb}(k)$ )。

17. 根据上述权利要求之一所述的方法, 其特征在于, 向解码器 (5) 传送窄带信号 ( $s_{nb}(k)$ ), 其具有在宽带输入语音信号 ( $s_{wb}^i(k)$ ) 的扩展频带之下的频带范围。

18. 根据权利要求 16 或 17 所述的方法, 其特征在于, 从传送给解码器 (5) 的窄带信号 ( $s_{nb}(k)$ ) 和频谱包络的重构形式中、尤其是从这两个信号的和中确定扩展了带宽的输出语音信号 ( $s_{wb}^o(k)$ ), 并作为解码器 (5) 的输出信号提供出去。

19. 根据上述权利要求之一所述的方法, 其特征在于, 步骤 a) 至 e) 在编码器 (1) 中执行, 在步骤 d) 中产生的已编码的信息作为数字信号 (BWE) 传送给解码器。

20. 根据上述权利要求之一所述的方法, 其特征在于, 所述宽带输入语音信号 ( $s_{wb}^i(k)$ ) 包括在大约 50Hz 至大约 7kHz 之间的带宽。

21. 根据上述权利要求之一所述的方法, 其特征在于, 所述宽带输入语音信号 ( $s_{wb}^i(k)$ ) 的扩展频带包括从大约 3.4kHz 到大约 7kHz 的频率范围。

22. 根据权利要求 17 所述的方法, 其特征在于, 所述窄带信号 ( $s_{nb}(k)$ ) 包括宽带输入语音信号 ( $s_{wb}^i(k)$ ) 从大约 50Hz 到大约 3.4kHz 的信号范围。

23. 一种用于人工扩展可被施加宽带输入语音信号 ( $s_{wb}^i(k)$ ) 的语音信号的带宽的装置, 其特征在于,

a) 用于从宽带输入语音信号 ( $s_{wb}^i(k)$ ) 的扩展频带中确定扩展带宽所需要的宽带输入语音信号 ( $s_{wb}^i(k)$ ) 的信号分量 ( $s_{eb}(k)$ ) 的装置;

b) 用于确定用于扩展带宽的信号分量 ( $s_{eb}(k)$ ) 的时间包络的装置;

c) 用于确定用于扩展带宽的信号分量 ( $s_{eb}(k)$ ) 的频谱包络的装置;

d) 用于对时间包络和频谱包络进行编码并提供经过编码的信息来用于扩展带宽的编码器 (1); 以及

e) 用于对经过编码的信息进行解码并从经过编码的信息中产生时间包络和频谱包络以产生扩展了带宽的输出语音信号 ( $s_{wb}^o(k)$ ) 的解码器 (5)。

24. 根据权利要求 23 所述的装置, 其特征在于, a) 至 d) 中的装置实施为编码器 (1)。

## 用于人工扩展语音信号的带宽的方法和装置

本发明涉及用于人工扩展语音信号的带宽的方法以及装置。

语音信号覆盖很宽的频率范围,该频率范围大约从与说话者有关而位于 80 至 160Hz 范围内的语音基频到超过 10kHz 的频率。但是在通过特定传输介质如电话而进行的语音通信中,由于带宽有效性的原因只能传输有限片段,其中保证大约 98% 的单句清晰度。

对应于特定于电话系统的最低带宽 300Hz 至 3.4kHz,语音信号基本上可以分为 3 个频率范围。每个频率范围在此都表征特定的语音特征以及主观感受。从而大约在 300Hz 以下的更低的频率基本上出现在有声的语音段期间,例如对于元音而言。在这种情况下,该频率范围包含音调分量,尤其是语音基频以及与音高有关的可能若干谐波。

这些低音频率对于主观感受语音信号的音量和动态性很重要。相应地,人类收听者基于虚拟音调高度的心理声学特性即使在缺乏低音频率时也能从更高频率范围内的谐波结构中感受到语音基频。从而在语音活动中从大约 300Hz 到大约 3.4kHz 范围内的平均频率基本上存在于语音信号中。该平均频率通过多个共振峰的随时间变化的频谱音调色彩以及时间和频率的微观结构表征说出的各个声音或音素。通过这种方式,平均频率传达了对理解语言很重要的信息的主要部分。

另一方面,在无声的音素中,尤其是对于尖锐的音素如“s”或“f”更是特别强烈地出现位于约 3.4kHz 以上的高频分量。所谓的爆破音如“k”或“t”具有含有强高频分量的宽频谱。因此该信号在该上频率范围中更多为噪声特性而不是音调特性。在该范围中存在的共振峰的结构相对而言不随时间变化,但是对不同的说话者有所不同。高频分量对语音信号的清晰度、精确度以及自然程度而言具有重要意义,因为没有高频分量语音就显得很沉闷。此外通过这种高频分量可以更好地区分摩擦音和辅音,该高频分量由此也保证增强对该语音的理解。

在通过具有有限带宽的传输信道的语音通信系统来传输语音信

号时，原则上希望而且也一直以此作为目标：能够以最可能的高质量从发送者向接收者传送待传输的语音信号。但是在此该语音质量是具有多个部分的主观参数，其中语音信号的理解度对这种语音通信系统最重要。

在现代数字传输系统中已经可以达到比较高的语音理解度。其中公知通过为电话带宽增加高频(大于 3.4kHz)以及低频(小于 300Hz)可以改善对该语音信号的主观判断。因此在主观质量改善的意义下力求在用于语音通信的系统中实现比常见电话带宽更大的带宽。在此可能的措施在于，修正该传输并借助编码方法促使传输带宽加宽或者可替换地执行人工带宽扩展。通过这种带宽扩展在接收端将频率带宽加宽到 50Hz 至 7kHz 的范围。借助合适的信号处理算法从窄带语音信号的短片段中利用模式识别的方法确定宽带模型的参数，接着将该参数用于估计该语音所缺乏的信号分量。在这种方法中从窄带语音信号中产生频率分量在 50Hz 至 7kHz 范围内的宽带对应物，并引起对主观感受的语音质量的改善。

在当前的语音信号和音频信号编码算法中更多地采用人工带宽扩展的技术。例如在带宽范围(声学带宽 50Hz 至 7kHz)内采用诸如 AMR-WB(适应性多比率宽带)编码解码算法的语音编码标准。在这种 AMR-WB 标准中从低频分量外推出上面的子频带(大约 6.4 至 7kHz 的频率范围)。在这种编码解码方法中通常通过比较小数量的辅助信息进行带宽扩展。该辅助信息例如可以是滤波器系数或放大系数，其中滤波器系数例如可以通过 LPC(线性预测滤波器)方法产生。该辅助信息以编码的位流传送给接收器。基于扩展带宽技术的其它标准目前可以在标准 AMR-WB+和扩展的 aac+语音/音频编码解码方法中找到。用于对信息进行编码和解码的方法称为 Codec(编解码器)，既包括编码器又包括解码器。每个数字电话，不管是为固网建立的还是为移动通信网络建立的，都包含这种将模拟信号转换为数字信号并将数字信号转换为模拟信号的 Codec。这种 Codec 可以用硬件或软件来实现。

在语音/音频信号编码算法的当前实现中采用了带宽扩展的技术，其中借助已经提到的 LPC 编码技术对扩展频带如 6.4 至 7kHz 的频率范围中的分量进行编码和解码。在此在编码器中对输入信号的

扩展频带进行 LPC 分析，并对剩余信号的子帧的 LPC 系数以及放大系数进行编码。在解码器中产生扩展频带的剩余信号，将传送的放大系数和 LPC 合成滤波器用于产生输出信号。上述过程可以直接应用于宽带的输入信号，也可以应用于在极限范围或临界范围中的具有扩展频带的下采样子带信号。

在经过扩展的 aac+语音/音频编码解码标准中采用 SBR（频谱带复制）技术。其中借助 64 信道 QMF 滤波器组将宽带音频信号划分为频率子带。对于高频滤波器带信道来说，对信号分量的子带采用经过推敲和技术上高度发展的参数编码，为此需要采用大量的检测器和估计器来检查位流内容。虽然在公知的标准和编码解码方法中已经可以改善语音信号的语音质量，但还是力求进一步提高语音质量。此外上述标准和编码解码方法耗费很大并且具有非常复杂的结构。

因此本发明要解决的技术问题是提供一种用于人工扩展语音信号的带宽的方法和装置，利用它们可以改善语音质量和提高语音理解度。此外该方法和装置还可以比较简单和花费少的方式实现。

该技术问题是通过具有按照权利要求 1 的特征的方法以及具有按照权利要求 23 的特征的装置来解决的。

在本发明的用于人工扩展语音信号的带宽的方法中执行以下步骤：

- a) 提供宽带的输入语音信号；
- b) 从宽带输入语音信号的扩展频带中确定扩展带宽所需要的宽带输入语音信号的信号分量；
- c) 确定用于扩展带宽的信号分量的时间包络；
- d) 确定用于扩展带宽的信号分量的频谱包络；
- e) 对时间包络和频谱包络的信息进行编码，并提供经过编码的信息来用于扩展带宽；以及
- f) 对经过编码的信息进行解码，并从经过编码的信息中产生时间包络和频谱包络以用于产生扩展了带宽的输出语音信号。

通过本发明的方法可以改善语言理解度和提高语音信号传输过程中的语音质量，其中语音信号也理解为声频信号。此外本发明的方法还对传输过程中的干扰具有很强的抵抗性。

优选的，扩展带宽所需要的信号分量通过滤波、尤其是带通滤波从宽带输入语音信号中确定，由此可以对需要的信号分量进行简单和不太费事的选择。

在步骤 c) 中对时间包络的确定优选与在步骤 d) 中对频谱包络的确定无关地进行。由此精确地确定包络，由此可以避免相互影响。

优选的，在步骤 e) 中对时间包络和频谱包络编码之前对时间包络和频谱包络进行量化。优选的，在用于确定频谱包络的步骤 d) 中确定用于扩展带宽的信号分量的频谱子带的信号功率。由此可以非常精确地确定用于表征时间包络和频谱包络的参数。

为了确定频谱子带的信号功率，优选产生用于扩展带宽的信号分量，其中对该信号分量进行特殊的变换，尤其是 FF（快速傅立叶）变换。此外，优选在用于确定时间包络的步骤 c) 中确定用于扩展带宽的信号分量的时间信号段的信号功率。由此以不费事的方式确定所需要的参数。

优选的，在步骤 f) 中对已编码的信息进行解码以重构地形成时间包络和频谱包络。

激励信号优选在解码器中从传送给该解码器的信号中产生，其中所传送的信号在对应于宽带输入语音信号的扩展频带频率范围的频率范围内具有这样的信号功率，即该信号功率使得可以产生激励信号。优选向解码器传送经过调制的窄带信号以产生激励信号，该窄带信号具有频率低于宽带输入语音信号的扩展频带的频带范围的频率的频带范围。该激励信号优选具有传送给该解码器的信号的基频的谐波。

优选的，从经过解码的时间包络和激励信号的信息中确定第一校正系数。此外从第一校正系数和激励信号中重构地形成时间包络，尤其是通过将第一校正系数与激励信号相乘。此外，优选对时间包络的重构形式进行滤波，并在滤波器中产生脉冲响应。从该脉冲响应和时间包络的重构形式中重构地形成频谱包络。此外从频谱包络的重构形式中重构出宽带输入语音信号的扩展频带的信号分量。由此非常可靠和非常精确地执行时间包络和频谱包络的重构。

在优选实施方式中向解码器传送窄带信号，其具有频率低于宽带输入语音信号的扩展频带的频率的频带范围。



优选的，从传送给解码器的窄带信号和频谱包络的重构形式中、尤其是从这两个信号的和中确定扩展了带宽的输出语音信号，并作为解码器的输出信号提供出去。由此可以产生和提供保证高语音理解度和高语音质量的输出信号。

优选的，步骤 a) 至 e) 在编码器中执行，该编码器优选设置在发射器中。优选的，在步骤 e) 中产生的已编码的信息作为数字信号传送给解码器。优选的，至少步骤 f) 在接收器中执行，其中解码器设置在该接收器中。还可以将本发明方法的所有步骤 a) 至 f) 都在接收器中执行。在这种情况下将接收器中的步骤 a) 至 e) 都替换成（不同实现的）估计方法。步骤 a) 至 e) 还可以分离地在发射器中执行。

宽带输入语音信号优选包括在大约 50Hz 至大约 7kHz 之间的带宽。宽带输入语音信号的扩展频带优选包括从大约 3.4kHz 到大约 7kHz 的频率范围。此外，窄带信号包括宽带输入语音信号从大约 50Hz 到大约 3.4kHz 的信号范围。

本发明的用于人工扩展可被施加宽带输入语音信号的语音信号的带宽的装置至少包括以下部件：

- a) 用于从宽带输入语音信号的扩展频带中确定扩展带宽所需要的宽带输入语音信号的信号分量的装置；
- b) 用于确定用于扩展带宽的信号分量的时间包络的装置；
- c) 用于确定用于扩展带宽的信号分量的频谱包络的装置；
- d) 用于对时间包络和频谱包络进行编码并提供经过编码的信息来用于扩展带宽的编码器；
- e) 用于对经过编码的信息进行解码并从经过编码的信息中产生时间包络和频谱包络以产生扩展了带宽的输出语音信号的解码器。

本发明的装置使得可以在通信设备中改善在语音信号传输过程中的语音质量和提高语言理解力，该通信设备例如是移动通信设备或 ISDN 设备。

a) 至 d) 中的装置优选实施为编码器。该编码器可以设置在发射器或接收器中，其中解码器设置在接收器中。

本发明方法的优选实施方式只要可以转换就也作为本发明装置的优选实施方式。

下面借助示意性的附图详细解释本发明的实施例。

图 1 示出本发明装置的编码器；以及

图 2 示出本发明装置的解码器。

在下面详细解释的发明中，语音信号的概念也包括音频信号。在图 1 和图 2 中相同或功能相同的元件具有相同的附图标记。

在图 1 中示出用于人工扩展语音信号的带宽的本发明装置的编码器 1 的示意电路连接图。编码器 1 既可以实现为硬件又可以作为算法实现为软件。编码器 1 在该实施例中包括用于对宽带输入语音信号  $s_{wb}^i(k)$  进行带通滤波的块 11。此外，编码器 1 包括与块 11 连接的块 12 和块 13。在此块 12 用于确定用于扩展带宽的信号分量的时间包络，这些信号分量是从宽带输入语音信号的扩展频带中确定出来的。按照相应的方式，块 13 用于确定用于扩展带宽的信号分量的频谱包络，这些信号分量是从宽带输入语音信号的扩展频带中确定出来的。

此外从图 1 中可以看出块 12 和块 13 与块 14 连接，其中块 14 用于量化通过块 12 和 13 产生的时间包络和频谱包络。

在图 1 中还示出实施为带通滤波器的块 2，在块 2 上施加宽带的输入语音信号  $s_{wb}^i(k)$ 。块 2 还与另一个块 3 连接，其中块 3 实施为另一个编码器。

在该实施例中编码器 1 以及块 2 和块 3 都设置在第一电话设备中。宽带输入语音信号在本实施例中具有从大约 50Hz 至大约 7kHz 的带宽。按照本发明，该宽带输入语音信号  $s_{wb}^i(k)$  施加在编码器 1 的带通滤波器或块 11 上。借助块 11 从在本实施例中包括从大约 3.4kHz 至大约 7kHz 的带宽的扩展频带中确定出扩展带宽所需要的信号分量。扩展带宽所需要的信号分量通过信号  $s_{eb}(k)$  来表征并作为块 11 的输出信号传送给两个块 12 和 13。在此在块 12 中，从信号  $s_{eb}(k)$  中确定出时间包络。按照相应的方式在块 13 中确定通过信号  $s_{eb}(k)$  表征的信号分量的频谱包络。

下面详细解释如何确定时间包络和频谱包络。在此，首先对表征扩展带宽所需要的信号分量的信号  $s_{eb}(k)$  进行分段，并对该窗口化的信号段进行变换。信号  $s_{eb}(k)$  的分段在各  $k$  扫描值的长度的帧内进行。全部面向帧地执行下面的所有步骤和子算法。每个语音帧（例

如具有 10ms 或 20ms 或 30ms 的持续时间) 可以有利地分为多个子帧 (持续时间例如为 2.5 或 5ms)

然后对窗口化的信号段进行变换。在该实施例中借助 FFT (快速傅立叶变换) 变换到频域中。经过 FFT 变换的信号段在此按照以下公式 1) 确定:

$$S_{wf}(i) = \sum_{\kappa=0}^{N_f-1} s_{eb}(\mu \cdot M_f + \kappa) \cdot w_f(\kappa) \cdot e^{-j\kappa \frac{2\pi}{N_f} i}$$

在该公式 1) 中,  $N_f$  表示 FFT 长度或帧长度,  $\mu$  表示帧下标,  $M_f$  表示窗口化的信号段的帧的重叠。此外  $w_f(\kappa)$  表示窗口函数。下面接着在频域中计算扩展频带的频率范围的子带中的信号功率。信号强度或信号功率的计算按照以下公式 2) 进行:

$$P_f(\mu, \lambda) = \sum_{i \in EB_\lambda} w_\lambda(i) \cdot |S_{wf}(i)|^2$$

在该公式 2) 中  $\lambda$  表示相应子带的下标, 其中  $EB_\lambda$  表征在第  $\lambda$  个频域窗口  $w_\lambda(i)$  中包含所有具有非零系数的 FFT 间隔区域  $i$  的集合。按照公式 2) 的子带的信号功率  $P_f(\mu, \lambda)$  表征传送给解码器的频谱包络的信息。

按照类似于确定频谱包络的方式在时域中确定时间包络, 并以经过带通滤波的宽带输入信号  $s_{wb}^i(k)$  的短暂的窗口化片段为基础。由此在确定时间包络的时候也考虑信号的信号段  $s_{eb}(k)$ 。对于每个窗口化段按照以下公式 3) 计算信号功率:

$$P_t(v) = \sum_{\kappa=0}^{N_t-1} (s_{eb}(v \cdot M_t + \kappa) \cdot w_t(\kappa))^2$$

在公式 3) 中,  $N_t$  表示帧长度,  $v$  表示帧下标,  $M_t$  表示信号段的帧的重叠。要注意一般用于提取时间包络的帧长度  $N_t$  和帧的重叠  $M_t$  远小于用于确定频谱包络的对应参数  $N_f$  和  $M_f$ 。

从信号  $s_{eb}(k)$  中提取时间包络的参数的替换方式在于, 对该信号  $s_{eb}(k)$  执行希尔伯特变换 (90°相移滤波)。经过滤波的部分和原始

部分的短片段信号功率的和给出了短暂的时间包络，对该时间包络下采样以确定信号功率  $P_t(v)$ 。这些信号段的信号功率  $P_t(v)$  就表征时间包络的信息。

表征时间包络和频谱包络的信号  $S_{Dt}(v)$  和  $S_{Df}(\mu, \lambda)$ ，在块 14 中量化和编码，这些信号分别表征按照公式 2) 和公式 3) 提取的信号功率的参数。块 14 的输出信号是数字信号 BWE，其表征按照编码方式包含时间包络和频谱包络的信息的位流。

将该数字信号 BWE 传送给解码器，下面将对该解码器详细解释。要注意在根据公式 2) 和 3) 提取的信号强度的参数之间存在冗余时可以执行同一种或关联的编码，该编码例如可以通过向量量化来实现。

此外从图 1 可以看出，宽带输入语音信号还传送给块 2。借助实施为带通滤波器的块 2 对该宽带输入语音信号  $s_{wb}^i(k)$  的窄带范围的信号分量进行滤波。在本实施例中，该窄带范围位于 50Hz 与 3.4kHz 之间。块 2 的输出信号是窄带信号  $s_{nb}(k)$  并传送给在本实施例中实施为另一个编码器的块 3。在块 3 中对窄带信号  $s_{nb}(k)$  进行编码，并作为数字信号 BWN 的位流传送给下面解释的解码器。

在图 2 中示出用于人工扩展语音信号带宽的本发明装置的这种解码器 5 的示意电路连接图。从图 2 可以看出，数字信号 BWN 首先传送给另一个解码器 4，该解码器 4 对包含在数字信号 BWN 中的信息解码并从中又产生窄带信号  $s_{nb}(k)$ 。此外解码器 4 产生另一个包含辅助信息的信号  $s_{si}(k)$ 。该辅助信息例如可以是放大系数或滤波器系数。该信号  $s_{si}(k)$  传送给解码器 5 的块 51。块 51 在该实施例中用于产生处于扩展频带的频率范围中的激励信号，为此考虑信号  $s_{si}(k)$  的信息。

此外在本实施例中设置在接收器中的解码器 5 具有块 52，该块 52 用于对通过编码器 1 和解码器 2 之间的传输段传输的信号 BWE 进行解码。要注意数字信号 BWN 也通过编码器 1 和解码器 2 之间的传输段传输。从图 2 可以看出，块 51 和块 52 都与解码器区域 53 至 55 连接。下面详细解释解码器 5 和在解码器 5 中执行的本发明方法的分步骤的功能原理。

如上所述，包含在编码后的数字信号 BWE 中的信息在块 52 中解

码，并重构出根据公式 2) 和 3) 计算并表征时间包络和频谱包络的信号功率。从图 2 中可以看出，在块 51 中产生的激励信号  $s_{exc}(k)$  是用于重构地形成时间包络和频谱包络的输入信号。该激励信号  $s_{exc}(k)$  在此基本上是任意信号，其中作为该信号的重要前提是，该信号必须具有在宽带输入频谱信号  $s_{wb}^i(k)$  的扩展频带的频率范围中足够的信号功率。例如，作为激励信号  $s_{exc}(k)$  采用经过调制的窄带信号  $s_{nb}(k)$  或任意的噪声。如上所述，该激励信号负责精确建立在宽带输出语音信号  $s_{wb}^o(k)$  的扩展频带的信号分量中的频谱包络和时间包络。因此有利的是，按照这样的方式产生该激励信号  $s_{exc}(k)$ ，使得其具有窄带信号  $s_{nb}(k)$  的基频的谐波。

在分级式语音编码的情况下，实现这一点的一种可能性在于，使用其它解码器 4 的参数。如果例如  $\Delta_k$  为基频的分数或实数值的偏差， $b$  为 CELP 窄带解码器内的自适应码本的 LTB 放大因子，那么例如可以利用谐波频率在当前基频的整数倍时通过带通滤波器对任意信号  $n_{eb}(k)$  的 LTP 合成滤波（扩展频带的频率范围）来进行激励。

这里根据下式 (4) 来产生激励信号：

$$s_{exc}(k) = n_{eb}(k) + f(b) \cdot s_{exc}(k - \Delta_k)$$

这里 LTP 放大因子可以通过函数  $f(b)$  来降低或限制，以便能够防止所产生的扩展频带的信号分量胜出。需要指出，可以实现多个其它的替代方案，以便借助于窄带编解码器的参数执行合成的宽带激励。

产生激励信号的另一种可能性在于，用固定频率的正弦函数来调制窄带信号  $s_{nb}(k)$ ，或通过直接采用任意的信号  $n_{eb}(k)$ ，这在上面已经被定义过。需要强调，用于产生激励信号  $s_{exc}(k)$  的方法完全取决于数字信号 BWE 的生成以及该数字信号 BWE 的格式以及该数字信号 BWE 的解码。因此就此进行独立的调整。

下面详细解释时间包络的重构式成型。数字信号 BWE 如上所述在块 52 中解码，并根据信号  $s_{pt}(\nu)$  和  $s_{pf}(\mu, \lambda)$  提供根据公式 2) 和 3) 计算的信号功率表征时间包络和频谱包络的参数。为此从图 2 中看出，在本实施例中首先重构地形成时间包络。这在解码区域 53 中执行。为此将激励信号  $s_{exc}(k)$  以及信号  $s_{pt}(\nu)$  传送给解码区域 53。如图 2 所

示,激励信号  $s_{exc}(k)$  既传送给块 531 又传送给乘法器 532。还将信号  $s_{pt(v)}$  传送给块 531。从传送给块 531 的信号中产生比例校正系数  $g_1(k)$ 。该比例校正系数  $g_1(k)$  由块 531 传送给乘法器 532。然后在乘法器 532 中将激励信号  $s_{exc}(k)$  与该比例校正系数  $g_1(k)$  相乘,从而产生输出信号  $s'_{exc}(k)$ , 该输出信号表征对时间包络的重构式成型。输出信号  $s'_{exc}(k)$  具有接近正确的时间包络,但是就正确的频率而言还不是很精确,由此在下面的步骤中需要重构地形成频谱包络,从而能够将不精确的频率与需要的频率相匹配。

在图 2 中可以看出,输出信号  $s'_{exc}(k)$  传送给解码器 5 的第二解码区域 54, 信号  $s_{pf(\mu, \lambda)}$  也传送给第二解码区域 54。第二解码区域 54 具有块 541 和块 542, 其中块 541 用于对输出信号  $s'_{exc}(k)$  进行滤波。从输出信号  $s'_{exc}(k)$  和信号  $s_{pf(\mu, \lambda)}$  中产生脉冲响应  $h(k)$ , 该脉冲响应从块 541 传送给块 542。然后在块 542 中由输出信号  $s'_{exc}(k)$  和脉冲响应  $h(k)$  来重构形成频谱包络。然后通过块 542 的输出信号  $s''_{exc}(k)$  表征重构的频谱包络。

在按照图 2 示出的实施例,在产生第二解码区域 54 的输出信号  $s''_{exc}(k)$  之后在解码器 5 的第三解码区域 55 中再次重构地形成时间包络。时间包络的重构形成按照类似于在第一解码区域 53 的方式进行。在此在第三解码区域 5 中从输出信号  $s''_{exc}(k)$  和信号  $s_{pt(v)}$  中通过块 551 产生第二比例校正系数  $g_2(k)$ , 将该系数传送给乘法器 552。然后提供表征扩展带宽所需要的信号分量的信号  $s_{eb}(k)$  作为解码器 5 的第三解码区域 55 的输出信号。将该信号  $s_{eb}(k)$  传送给求和器 56, 窄带信号  $s_{eb}(k)$  也传送给求和器 56。通过窄带信号  $s_{eb}(k)$  和信号  $s_{eb}(k)$  的求和, 产生扩展了带宽的输出信号  $s^{\circ}_{wb}(k)$ , 并作为解码器 5 的输出信号提供。

要注意图 2 所示的实施例只是示例性的,对于本发明来说像在第一解码区域 53 中进行的那样重构地形成时间包络一次以及像在第二解码区域 54 中进行的那样重构地形成频谱包络一次就足够了。同样要注意还可以在于第一解码区域 53 中重构地形成时间包络之前在第二解码区域 54 中重构地形成频谱包络。这意味着在该实施例中第二解码区域 54 设置在在第一解码器区域 53 之前。还可以再次继续交替地执行时间包络的重构形成和频谱包络的重构形成, 并且例如在

图 2 所示的实施例中在第三解码区域 55 之后接着设置另一个解码区域，在该另一个解码区域中重新重构地形成频谱包络。

如上所述，本发明在该实施例中以有利方式用于具有大约 50Hz 至 7kHz 频率范围的宽带输入语音信号。同样，在该实施例中本发明可用于人工扩展语音信号的带宽，其中在此扩展频带通过大约 3.4kHz 至大约 7kHz 的频率范围来预定。还可以将本发明用于设置在低频频率范围中的扩展频带。例如，该扩展频带在此可以包括大约 50Hz 或更低的频率至大约 3,4kHz 的频率范围。要着重说明，本发明的方法可以按照以下方式用于人工扩展语音信号的带宽，即使扩展频带包括至少部分在大约 7kHz 频率以上并例如达到 8kHz、尤其是 10kHz 或更高频率的频率范围。

如上所述，时间包络的重构形成在按照图 2 的第一解码区域 53 中通过将第一比例校正系数  $g_1(k)$  和激励信号  $s_{exc}(k)$  相乘来产生。在此要注意，在时域中的乘法对应于频域中的卷积运算，由此给出以下公式 (5)：

$$s'_{exc}(k) = g(k) \bullet s_{exc}(k) ;$$

$$S'_{exc}(z) = G(z) * S_{exc}(z)$$

只要频谱包络在原理上没有被第一解码区域 53 改变，则第一比例校正系数或放大系数  $g_1(k)$  就应当具有严格的低通频率特性。

为了计算放大系数或第一校正系数  $g_1(k)$ ，通过在上面已经用于分段和分析对时间包络的提取或在编码器 1 中借助块 12 从信号  $s_{eb}(k)$  中产生信号  $s_{pt}(v)$  的方式来分段和分析激励信号  $s_{exc}(k)$ 。通过公式 3) 计算的经过解码的信号功率和经过分析的信号强度的结果  $P^{exc}_i(v)$  之间的比例产生了第  $v$  个信号段的期望放大系数  $\gamma(v)$ 。第  $v$  个信号段的该放大系数根据以下公式 6) 计算：

$$\gamma(v) = \sqrt{\frac{P_i(v)}{P_i^{exc}(v)}}$$

从该放大系数  $\gamma(v)$  中通过内插和低通滤波计算放大系数或第一

校正系数  $g_1(k)$ 。为了限制该放大系数或第一校正系数  $g_1(k)$  对频谱包络的影响，低通滤波在此具有很重要的意义。

扩展频带所需要的信号分量的频谱包络的重构形式通过对表征时间包络的重构形式的输出信号  $s'_{exc}(k)$  进行滤波来确定。在此该滤波操作在时域或在频率中进行。为了避免脉冲响应  $h(k)$  具有较大的时间散射或时间扩展幅度，分析第一解码区域 53 的输出信号  $s'_{exc}(k)$ ，以便能构找到信号功率  $P_f^{exc}(\mu, \lambda)$ 。扩展频带的频率范围的对应子带的期望放大系数  $\Phi(\mu, \lambda)$  根据以下公式 7) 计算：

$$\Phi(\mu, \lambda) = \sqrt{\frac{P_f(\mu, \lambda)}{P_f^{exc}(\mu, \lambda)}}$$

频谱包络的形状滤波器的频率特性  $H(\mu, i)$  可以通过对放大系数  $\Phi(\mu, \lambda)$  进行内插并在考虑频率的情况下进行平滑来计算。如果频谱包络的形状滤波器应当用在时域中，例如通过线性相位 FIR 滤波器，则滤波器系数可以通过对频率特性  $H(\mu, i)$  和后面的窗口化的反 FFT 变换来计算。

如通过上面的实施例解释和展示的，时间包络的重构形成影响频谱包络的重构形成，反之亦然。因此有利的是，如在该实施例中解释和在图 2 中示出的那样，在迭代过程中交替地执行时间包络的重构形成和频谱包络的重构形成。由此可以明显改善扩展频带的信号分量的时间包络和频谱包络的一致性，该时间包络和频谱包络在解码器中重构，并且能达到在编码器中相应产生的时间包络和频谱包络。

在按照图 2 的上述实施例中，执行一个半迭代（重构时间包络、重构频谱包络和再次重构时间包络）。通过本发明实现的带宽扩展使得很容易产生具有处于正确频率下的谐波的激励信号，该正确频率例如是瞬时音素的基频的整数倍。要注意，本发明还可以用于宽带输入信号的被下采样的子带信号分量。这在要求极少的计算成本时是很有利的。

优选的，编码器 1 以及块 2 和块 3 都设置在发射器中，其中按逻辑在块 2 和块 3 以及编码器 1 中执行的方法步骤也在该发射器中执行。块 4 以及解码器 5 优选可以设置在接收器中，由此也很清楚在



解码器 5 和块 4 中执行的前面的步骤要在接收器中处理。要注意，本发明还可以这样实现，即在编码器 1 中执行的方法步骤在解码器 5 中执行，由此只在接收器中执行。在此可以在解码器 5 中估计按照公式 2) 和 3) 计算的信号功率。尤其是块 52 用于估计信号功率的参数。该实施例使得可以消除在数字信号 BWE 中传送的辅助信息的潜在传送错误。通过预先估计包络例如由于数据丢失而失去的参数，可以防止麻烦地转换信号带宽。

与用于人工扩展语音信号的带宽的公知方法不同，在本发明中不向解码器传送已经采用的放大系数和滤波器系数作为辅助信息，而只是传送期望的时间包络和频谱包络作为辅助信息。在设置在接收器内的解码器中才计算放大系数和滤波器系数。由此可以成本低的方式在接收器中分析带宽的人工扩展，并在必要时进行校正。此外按照本发明的方法和装置可以非常稳定地抵抗激励信号的干扰，例如所接收的窄带信号的这种干扰可能通过传输错误而引起。

通过分开执行对时间包络和频谱包络的分析、传送和重构成形，可以在时域和频域中都达到非常好的分辨率或分隔。这导致对静止音素和音调以及临时或短时信号的非常好的再现性。对于语音信号，尤其是停止辅音和爆破音的再现得到了明显改善的时间分辨率。

与传统的带宽扩展不同，通过本发明可以通过线性相位 FIR 滤波器而不是 LPC 合成滤波器来进行频率成型。由此还可以降低典型的伪影（滤波器环）。此外本发明还可以非常灵活和模块化的结构实现，此外该结构还使得可以简单方式更换或调节在接收器和解码器 5 中各个块。优选的，这种更换或调节不需要改变发射器和编码器 1 或传输信号的格式一经过编码的信息就以该格式传送给解码器 5 或接收器。此外利用本发明的方法可以运行不同的解码器，由此可以根据可提供的计算功率以不同的精度再次产生宽带输入信号。

要注意所接收的表征频谱包络和时间包络的参数不仅可用于扩展带宽，还可用于支持后面的信号处理块如后滤波，或者附加的编码组件如变换编码器。

所产生的窄带语音信号  $s_{nb}(k)$ ，如向用于扩展带宽的算法提供的，例如可以在减小扫描频率一半之后以 8kHz 的扫描速率给出。

利用本发明和带宽扩展所基于的原理可以产生 G.729+标准信息的宽带激励。在数字信号 BWE 中传送的辅助信息的数据率大约是 2kbit/s。此外在本发明中需要小于 3WMOPS 的不太复杂的计算系统或不太复杂的计算花费。此外，本发明的方法和装置能非常稳定地抵抗 G.729+标准的基带干扰。本发明还可以优选用于在通过 IP 的语音中的使用。此外本发明的方法以及装置与 TDAC 包络兼容。另外本发明还具有极度模块化和灵活的结构以及模块化和灵活的概念。

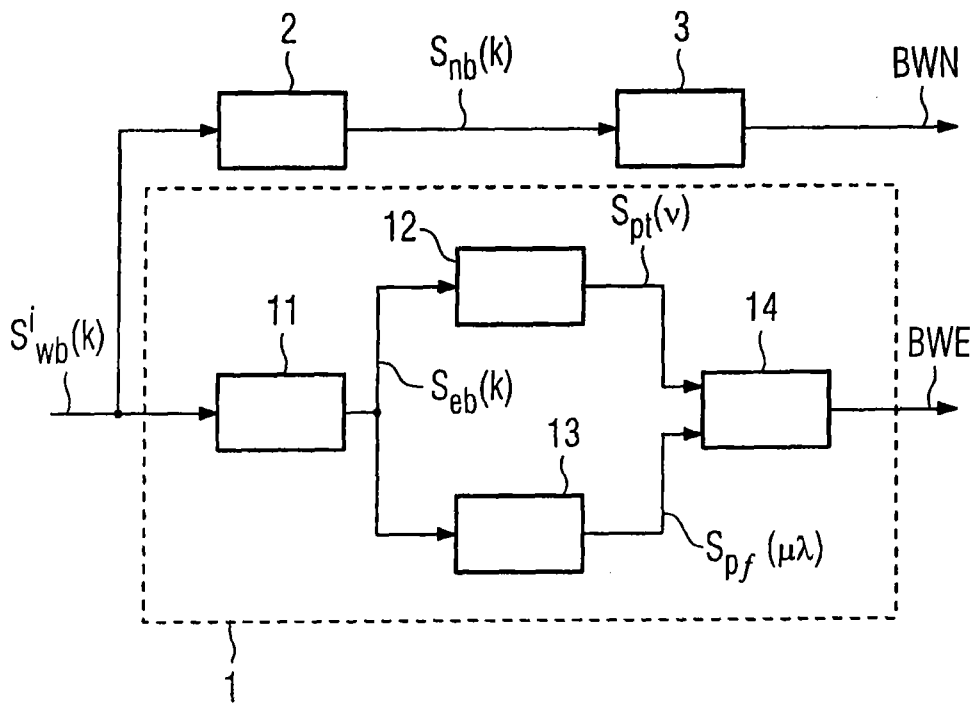


图 1

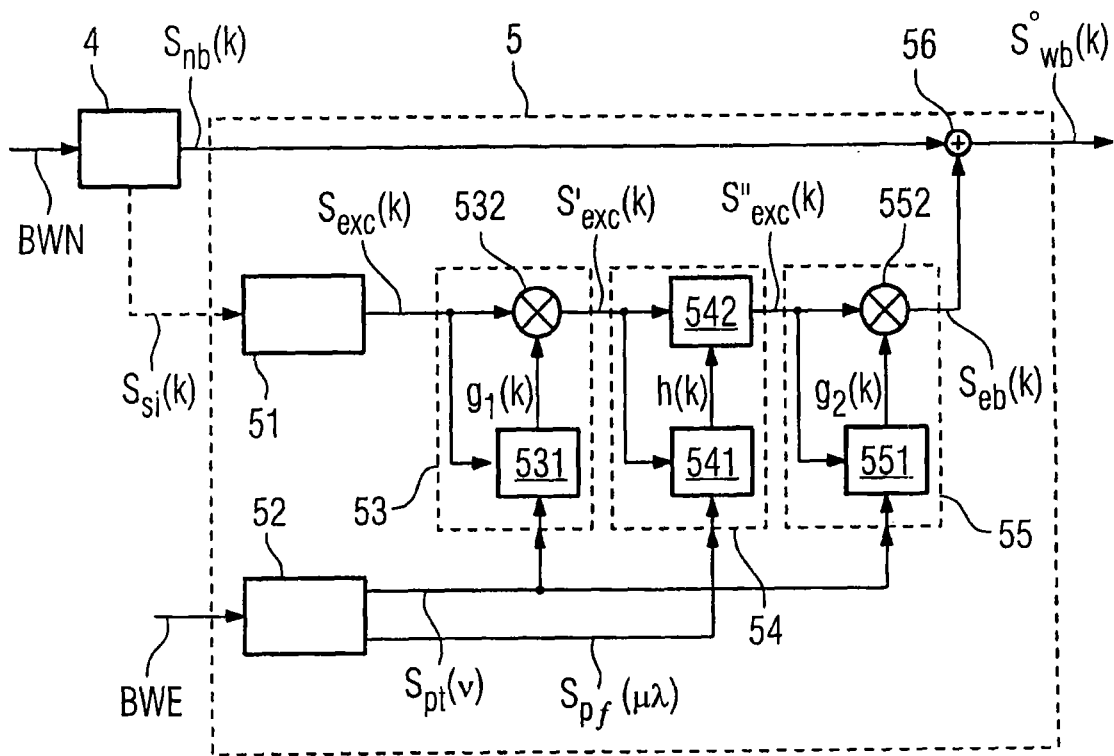


图 2