

(12) 发明专利

(10) 授权公告号 CN 101395661 B

(45) 授权公告日 2013.02.06

(21) 申请号 200780007780.0

(22) 申请日 2007.03.07

(30) 优先权数据

60/743,421 2006.03.07 US

(85) PCT申请进入国家阶段日

2008.09.04

(86) PCT申请的申请数据

PCT/SE2007/050132 2007.03.07

(87) PCT申请的公布数据

W02007/102782 EN 2007.09.13

(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 A·塔莱布

(74) 专利代理机构 中国专利代理(香港)有限公司

司 72001

代理人 柯广华 王丹昕

(51) Int. Cl.

G10L 19/24 (2013.01)

G10L 19/04 (2013.01)

(56) 对比文件

US 2004/0024594 A1, 2004.02.05, 全文.

CN 1229501 A, 1999.09.22, 全文.

CN 1179848 A, 1998.04.22, 全文.

Juin-Hwey Chen et al. A Low-Delay CELP Coder for the CCITT 16 kb/s Speech Coding Standard. 《IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS》. 1992, 第 10 卷 (第 5 期), 第 830 - 849 页.

Charles D. Creusere et al. Efficient Audio Coding Using Perfect Reconstruction Noncausal I1 Filter Banks. 《IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING》. 1996, 第 4 卷 (第 2 期), 第 115 - 123 页.

Ivan V. Bajic et al. Non-causal Error Control for Video Streaming over Wireless Packet Networks. 《2005 International Conference on Wireless Networks, Communications and Mobile Computing》. 2005, 第 2 卷第 1106 - 1111 页.

审查员 张飞弦

权利要求书 6 页 说明书 13 页 附图 11 页

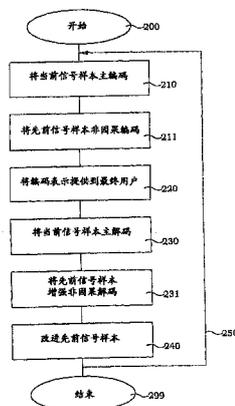
(54) 发明名称

音频编码和解码的方法和设备

(57) 摘要

一种用于音频编码和解码的方法包括将当前音频信号样本主编码 (12) 成编码表示 (T(n)) 和将第一先前音频信号样本非因果编码 (13) 成编码增强表示 (ET(n-N+))。该方法还包括向最终用户提供编码表示。在最终用户处,该方法包括将编码表示 (T*(n)) 主解码 (52) 成当前已接收音频信号样本和将编码增强表示 (ET*(n-N+)) 非因果解码 (53) 成增强的第一先前已接收音频信号样本。该方法还包括基于增强的第一先前已接收音频信号样本,改进对应于第一先前音频信号样本的第一先前已接收音频信号样本。还介绍了用于音频编码和解码的装置和系统。

CN 101395661 B



1. 一种用于音频编码和解码的方法,包括以下步骤:

将当前音频信号样本 $s(n)$ 主编码成所述当前音频信号样本的编码表示 $T(n)$;

将第一先前音频信号样本 $s(n-N^+)$ 非因果编码成所述第一先前音频信号样本的编码增强表示 $ET(n-N^+)$;

将所述当前音频信号样本的所述编码表示 $T(n)$ 和所述第一先前音频信号样本的所述编码增强表示 $ET(n-N^+)$ 提供到最终用户;

将所述当前音频信号样本的所述编码表示主解码成当前已接收音频信号样本 $\bar{s}(n)$;

将所述第一先前音频信号样本的所述编码增强表示非因果解码成增强的第一先前已接收音频信号样本 $\bar{s}(n-N^+)$;以及

基于第一先前已接收音频信号样本 $\bar{s}(n-N^+)$ 和所述增强的第一先前已接收音频信号样本 $\bar{s}(n-N^+)$, 改进对应于所述第一先前音频信号样本 $s(n-N^+)$ 的第一先前已接收音频信号样本 $\bar{s}(n-N^+)$ 。

2. 如权利要求 1 所述的方法,其中所述非因果编码是基于与第一时间瞬间后出现的时间瞬间相关联的信号样本或信号样本的表示,对与所述第一时间瞬间相关联的信号样本的编码。

3. 如权利要求 1 或 2 所述的方法,其中所述非因果编码是非因果预测编码,并且所述非因果解码是非因果预测解码。

4. 如权利要求 3 所述的方法,其中非因果预测编码的所述步骤又包括:

在开环中从第一组音频信号样本得到所述第一先前音频信号样本的第一非因果预测;

所述第一组音频信号样本包括以下样本的至少其中之一:

在所述第一先前音频信号样本后出现的至少一个先前音频信号样本;以及

所述当前音频信号样本;

计算所述第一先前音频信号样本与所述第一非因果预测之间的差作为第一差;以及

将至少所述第一差和所述第一非因果预测的参数编码成所述第一先前音频信号样本的所述编码增强表示;以及其中,非因果预测解码的所述步骤又包括:

将所述第一先前音频信号样本的所述编码增强表示解码成所述第一差和所述第一非因果预测的参数;

基于所述第一非因果预测的所述参数,从对应于所述第一组的第二组已接收音频信号样本得到所述增强的第一先前已接收音频信号样本的第二非因果预测;

计算所述第二非因果预测和所述第一差之和作为所述增强的第一先前已接收音频信号样本。

5. 如权利要求 3 所述的方法,其中非因果预测编码的所述步骤又包括:

在闭环中从第一组音频信号样本的表示得到所述第一先前音频信号样本的第一非因果预测;

所述第一组音频信号样本的表示包括以下项的至少其中之一:

与在所述第一先前音频信号样本后出现的时间相关联的先前音频信号样本的至少一

个表示 ; 以及

所述当前音频信号样本的表示 ;

计算所述第一先前音频信号样本或所述第一先前音频信号样本的表示与所述第一非因果预测之间的差作为第一差 ; 以及

将至少所述第一差和所述第一非因果预测的参数编码成所述第一先前音频信号样本的所述编码增强表示 ; 以及

其中, 非因果预测解码的所述步骤又包括 :

将所述第一先前音频信号样本的所述编码增强表示解码成所述第一差和所述第一非因果预测的参数 ;

基于所述第一非因果预测的所述参数, 从对应于所述第一组的第二组已接收音频信号样本得到所述增强的第一先前已接收音频信号样本的第二非因果预测 ;

计算所述第二非因果预测和所述第一差之和作为所述增强的第一先前已接收音频信号样本。

6. 如权利要求 4 或 5 所述的方法, 其中所述第一非因果预测和所述第二非因果预测是线性非因果预测, 由此, 所述第一非因果预测的所述参数是滤波器系数。

7. 如权利要求 1 或 2 所述的方法, 其中所述主编码是因果编码。

8. 如权利要求 1 所述的方法, 其中所述主编码是主预测编码, 并且所述主解码是主预测解码。

9. 如权利要求 8 所述的方法, 其中主预测编码的所述步骤又包括 :

在开环中从第三组先前音频信号样本得到所述当前音频信号样本的第一主预测 ;

计算所述当前音频信号样本与所述第一主预测之间的差作为第二差 ; 以及

将至少所述第二差和所述第一主预测的参数编码成所述当前音频信号样本的所述编码表示 ; 以及

其中, 主预测解码的所述步骤又包括 :

将所述当前音频信号样本的所述编码表示解码成所述第二差和所述第一主预测的所述参数 ;

基于所述第一主预测的所述参数, 从对应于所述第三组的第四组已接收音频信号样本得到所述当前已接收音频信号样本的第二主预测 ;

计算所述第二主预测和所述第二差之和作为所述当前已接收音频信号样本。

10. 如权利要求 8 所述的方法, 其中主预测编码的所述步骤又包括 :

在闭环中从第三组先前音频信号样本的表示得到所述当前音频信号样本的第一主预测 ;

计算所述当前音频信号样本与所述第一主预测之间的差作为第二差 ; 以及

将至少所述第二差和所述第一主预测的参数编码成所述当前音频信号样本的所述编码表示 ; 以及

其中, 主预测解码的所述步骤又包括 :

将所述当前音频信号样本的所述编码表示解码成所述第二差和所述第一主预测的所述参数 ;

基于所述第一主预测的所述参数, 从对应于所述第三组的第四组已接收音频信号样本

得到所述当前已接收音频信号样本的第二主预测；

计算所述第二主预测和所述第二差之和作为所述当前已接收音频信号样本。

11. 如权利要求 9 或 10 所述的方法，其中所述第一主预测和所述第二主预测是线性主预测，由此，所述第一主预测的所述参数是滤波器系数。

12. 如权利要求 11 所述的方法，其中所述第一主预测、所述第二主预测、所述第一非因果预测和所述第二非因果预测基于自适应码本范例，由此所述当前音频信号样本的所述编码表示和所述第一先前音频信号样本的所述编码增强表示包括固定和自适应码本的量化索引。

13. 如权利要求 1 或 2 所述的方法，其中提供所述当前音频信号样本的所述编码表示的所述步骤和提供所述第一先前音频信号样本的所述编码增强表示的所述步骤作为分层的编码来执行，其中，附加层包括非因果预测表示。

14. 一种用于音频编码的方法，包括以下步骤：

将当前音频信号样本主编码成所述当前音频信号样本的编码表示；

将第一先前音频信号样本非因果编码成所述第一先前音频信号样本的编码增强表示；

以及

提供所述当前音频信号样本的所述编码表示和所述第一先前音频信号样本的所述编码增强表示。

15. 一种用于音频解码的方法，包括以下步骤：

在最终用户处获得当前音频信号样本的编码表示和第一先前音频信号样本的编码增强表示；

将所述当前音频信号样本的所述编码表示主解码成当前已接收音频信号样本；

将所述第一先前音频信号样本的所述编码增强表示非因果解码成增强的第一先前已接收音频信号样本；以及

基于第一先前已接收音频信号样本和所述增强的第一先前已接收音频信号样本，改进对应于所述第一先前音频信号样本的第一先前已接收音频信号样本。

16. 一种用于音频信号样本的编码器，包括：

用于接收音频信号样本的输入；

主编码器部分，连接到所述输入并设置用于将当前音频信号样本编码成所述当前音频信号样本的编码表示；

非因果编码器部分，连接到所述输入并设置用于将第一先前音频信号样本编码成所述第一先前音频信号样本的编码增强表示；

输出，连接到所述主编码器部分和所述非因果编码器部分并设置用于提供所述当前音频信号样本的所述编码表示和所述第一先前音频信号样本的所述编码增强表示。

17. 如权利要求 16 所述的编码器，其中所述非因果编码是基于与第一时间瞬间后出现的时间瞬间相关联的信号样本或信号样本的表示，对与所述第一时间瞬间相关联的信号样本的编码。

18. 如权利要求 16 或 17 所述的编码器，其中所述非因果编码器部分是非因果预测编码器部分。

19. 如权利要求 18 所述的编码器，其中所述非因果编码器部分又包括：

非因果预测器,设置用于在开环中从第一组音频信号样本得到所述第一先前音频信号样本的非因果预测;

所述第一组音频信号样本包括以下样本的至少其中之一;

在所述第一先前音频信号样本后出现的至少一个先前音频信号样本;以及

所述当前音频信号样本;

计算部件,设置用于获得作为所述第一先前音频信号样本与所述非因果预测之间的差的第一差;以及

编码部件,设置用于将至少所述第一差和所述非因果预测的参数编码成所述第一先前音频信号样本的所述编码增强表示。

20. 如权利要求 18 所述的编码器,其中所述非因果编码器部分又包括:

非因果预测器,设置用于在闭环中从第一组音频信号样本的表示得到所述第一先前音频信号样本的非因果预测;

所述第一组音频信号样本的表示包括以下的至少其中之一;

与在所述第一先前音频信号样本后出现的时间相关联的先前音频信号样本的至少一个表示;以及

所述当前音频信号样本的表示;

计算部件,设置用于获得作为所述第一先前音频信号样本与所述非因果预测之间的差的第一差;以及

编码部件,设置用于将至少所述第一差和所述非因果预测的参数编码成所述第一先前音频信号样本的所述编码增强表示。

21. 如权利要求 16 或 17 所述的编码器,其中所述主编码器部分是因果编码器部分。

22. 如权利要求 19 所述的编码器,其中所述主编码器部分是主预测编码器部分。

23. 如权利要求 22 所述的编码器,其中所述主编码器部分又包括:

主预测器,设置用于在开环中从第二组先前音频信号样本得到所述当前音频信号样本的主预测;

计算部件,设置用于获得作为所述当前音频信号样本与所述主预测之间的差的第二差;以及

编码部件,设置用于将至少所述第二差和所述主预测的参数编码成所述当前音频信号样本的所述编码表示。

24. 如权利要求 22 所述的编码器,其中所述主编码器部分又包括:

主预测器,设置用于在闭环中从第二组先前音频信号样本的表示得到所述当前音频信号样本的主预测;

计算部件,设置用于获得作为所述当前音频信号样本与所述主预测之间的差的第二差;以及

编码部件,设置用于将至少所述第二差和所述主预测的参数编码成所述当前音频信号样本的所述编码表示。

25. 如权利要求 23 或 24 所述的编码器,其中所述主预测器和所述非因果预测器基于自适应码本范例,由此所述当前音频信号样本的所述编码表示和所述第一先前音频信号样本的所述编码增强表示包括固定和自适应码本的量化索引。

26. 如权利要求 21 所述的编码器,其中所述主编码器部分的所述编码部件和所述非因果预测器编码器部分的所述编码部件连接在一起,并设置为在所述输出处将所述当前音频信号样本的所述编码表示和所述第一先前音频信号样本的所述编码增强表示作为分层的编码信息来提供,其中,附加层包括非因果预测表示。

27. 一种用于音频信号样本的解码器,包括:

输入,设置用于接收由主编码器编码的当前音频信号样本的编码表示和由非因果编码器编码的第一先前音频信号样本的编码增强表示;

主解码器部分,连接到所述输入并设置用于将所述当前音频信号样本的所述编码表示主解码成当前已接收音频信号样本;

非因果解码器部分,连接到所述输入并设置用于将所述第一先前音频信号样本的所述编码增强表示非因果解码成增强的第一先前已接收音频信号样本;以及

信号调节器,连接到所述主解码器部分和所述非因果解码器部分并设置用于基于第一先前已接收音频信号样本和所述增强的第一先前已接收音频信号样本之间的比较,改进对应于所述第一先前音频信号样本的第一先前已接收音频信号样本。

28. 如权利要求 27 所述的解码器,其中所述非因果解码是基于与第一时间瞬间后出现的时间瞬间相关联的信号样本或信号样本的表示,对与所述第一时间瞬间相关联的信号样本的解码。

29. 如权利要求 27 或 28 所述的解码器,其中所述非因果解码器部分是非因果预测器解码器部分。

30. 如权利要求 29 所述的解码器,其中所述非因果预测器解码器部分又包括:

解码部件,设置用于将所述第一先前音频信号样本的所述编码增强表示解码成第一差和非因果预测的参数;

非因果预测器,设置用于基于所述非因果预测的所述参数,从第一组已接收音频信号样本得到所述增强的第一先前已接收音频信号样本的非因果预测;

所述第一组已接收音频信号样本包括以下样本的至少其中之一:

在所述第一先前已接收音频信号样本后出现的至少一个先前已接收音频信号样本;以及

当前已接收音频信号样本;

计算部件,设置用于获得作为所述非因果预测和所述第一差之和的所述增强的第一先前已接收音频信号样本。

31. 如权利要求 27 到 28 中任一项所述的解码器,其中所述主解码器部分是因果解码器部分。

32. 如权利要求 30 所述的解码器,其中所述主解码器部分是主预测解码器部分。

33. 如权利要求 32 所述的解码器,其中所述主解码器部分又包括:

解码部件,设置用于将所述当前音频信号样本的所述编码表示解码成第二差和主预测的参数;

主预测器,设置用于基于所述主预测的所述参数,从第二组先前已接收音频信号样本得到所述当前已接收音频信号样本的主预测;

计算部件,设置用于获得作为所述主预测和所述第二差之和的所述当前已接收音频信

号样本。

34. 如权利要求 33 所述的解码器,其中所述主预测器和所述非因果预测器基于自适应码本范例,由此所述当前音频信号样本的所述编码表示和所述第一先前音频信号样本的所述编码增强表示包括固定和自适应码本的量化索引。

35. 一种音频调解系统的终端,包括至少以下之一:如权利要求 16 到 26 中任一项所述的编码器和如权利要求 27 到 34 中任一项所述的解码器。

36. 一种音频调解系统,包括具有如权利要求 16 到 26 中任一项所述的编码器的至少一个终端和具有如权利要求 27 到 34 中任一项所述的解码器的至少一个终端。

音频编码和解码的方法和设备

技术领域

[0001] 本发明一般涉及音频信号样本的编码和解码。

背景技术

[0002] 在音频信号、且具体而言在语音信号中，在相邻样本之间存在高度相关。为了执行语音信号的高效量化和编码，可在编码前去除此类冗余。

[0003] 可通过两个缓慢时变线性预测滤波器来高效地对语音信号建模，这两个缓慢时变线性预测滤波器分别对频谱包络和频谱精细结构建模。声道的形状主要确定短时间频谱包络，而频谱精细结构主要归因于声带的周期振动。

[0004] 在现有技术中，经常使用线性模型对音频信号中的冗余建模。一种去除冗余的公知技术是通过使用预测，具体而言使用线性预测。从先前音频信号样本来预测原始当前音频信号样本，先前音频信号样本可以是原始样本或预测样本。残差定义为原始音频信号样本与预测音频信号样本之间的差。量化器搜索残差的最佳表示，例如，指向内部码本的索引。残差表示和线性预测滤波器的参数作为原始当前音频信号样本的表示提供。在解码器中，表示随后可用于重建当前音频信号样本的已接收版本。

[0005] 线性预测通常用于短期相关。理论上，可在任何阶使用 LP 滤波器。但是，由于莱文森 - 德宾算法的数值稳定性问题及在存储空间和算术运算方面所导致许多的复杂性原因，使用大阶线性预测是极不可取的。另外，将 LP 系数编码所需的比特率禁止此类使用。在实践中使用的 LP 预测器的阶数一般不超过 20 个系数。例如，宽带语音编码 AMR-WB 的标准具有 16 阶的 LPC 滤波器。

[0006] 为在保持质量的同时进一步降低所需的比特率量，需要正确地利用在有声语音段中的语音信号的周期性。为此，且由于线性预测一般将利用在不到一个基音周期中包含的相关，经常在线性预测残差上使用基音预测器。因此可利用音频信号中的长期相关性。

[0007] 虽然当前标准化的语音编解码器以极低比特率提供可接受的质量，但相信通过极少额外比特的代价，质量可得到进一步增强。现有技术的语音和音频编码算法的一个次要问题是，用于语音或音频信号的现有技术的模型虽然很有效率，但未将音频信号中存在的所有可能冗余考虑在内。在一般的音频编码中，且具体而言在语音编码中，始终存在在给定质量下降低所需比特率或在给定比特率下得到更好质量的需要。

[0008] 此外，现在经常要求使用嵌入式或分层方法以便适应质量与比特率之间的关系。但是，在给定比特率，并且对于给定编码结构，与非分层编码器相比，嵌入式或分层语音编码器经常显得存在质量损失。为了通过相同的编码结构体验相同的质量，经常要求增大比特率。

发明内容

[0009] 本发明的一个目的是进一步利用音频信号中存在的冗余。本发明的又一目的是提供一种易于在嵌入式或分层方法中应用的编码 - 解码方案。本发明还有的又一目的是提供

另外的冗余利用而不会造成太大的延迟。

[0010] 上述目的通过根据随附权利要求的方法和装置而得以实现。通常,在第一方面,用于音频编码和解码的方法包括将当前音频信号样本主编码成当前音频信号样本的编码表示和将第一先前音频信号样本非因果编码成第一先前音频信号样本的编码增强表示。该方法还包括将当前音频信号样本的编码表示和第一先前音频信号样本的编码增强表示提供到最终用户。在最终用户处,该方法包括将当前音频信号样本的编码表示主解码成当前已接收音频信号样本和将第一先前音频信号样本的编码增强表示非因果解码成增强的第一先前已接收音频信号样本。该方法还包括基于第一先前已接收音频信号样本和增强的第一先前已接收音频信号样本,改进对应于第一先前音频信号样本的第一先前已接收音频信号样本。

[0011] 在第二方面,用于音频编码的方法包括将当前音频信号样本主编码成当前音频信号样本的编码表示和将第一先前音频信号样本非因果编码成第一先前音频信号样本的编码增强表示。该方法还包括提供当前音频信号样本的编码表示和第一先前音频信号样本的编码增强表示。

[0012] 在第三方面,用于音频解码的方法包括在最终用户处获得当前音频信号样本的编码表示和第一先前音频信号样本的编码增强表示。该方法还包括将当前音频信号样本的编码表示主解码成当前已接收音频信号样本和将第一先前音频信号样本的编码增强表示非因果解码成增强的第一先前已接收音频信号样本。该方法还包括基于第一先前已接收音频信号样本和增强的第一先前已接收音频信号样本,改进对应于第一先前音频信号样本的第一先前已接收音频信号样本。

[0013] 在第四方面,用于音频信号样本的编码器包括:接收音频信号样本的输入;主编码器部分,连接到输入并设置用于将当前音频信号样本编码成当前音频信号样本的编码表示;以及非因果编码器部分,连接到输入并设置用于将第一先前音频信号样本编码成第一先前音频信号样本的编码增强表示。编码器还包括输出,该输出连接到主编码器部分和非因果编码器部分并设置用于提供当前音频信号样本的编码表示和第一先前音频信号样本的编码增强表示。

[0014] 在第五方面,用于音频信号样本的解码器包括输入,该输入设置用于接收通过主编码器编码的当前音频信号样本的编码表示和由非因果编码器编码的第一先前音频信号样本的编码增强表示。该解码器还包括:主解码器部分,连接到输入并设置用于将当前音频信号样本的编码表示主解码成当前已接收音频信号样本;以及非因果解码器部分,连接到输入并设置用于将第一先前音频信号样本的编码增强表示非因果解码成增强的第一先前已接收音频信号样本。该解码器还包括信号调节器,该调节器连接到主解码器部分和非因果解码器部分并设置用于基于第一先前已接收音频信号样本和增强的第一先前已接收音频信号样本之间的比较来改进对应于第一先前音频信号样本的第一先前已接收音频信号样本。

[0015] 在第六方面,音频调解 (mediating) 系统的终端包括根据第四方面的编码器和根据第五方面的解码器的至少之一。

[0016] 在第七方面,音频系统包括具有根据第四方面的编码器的至少一个终端和具有根据第五方面的解码器的至少一个终端。

[0017] 本发明允许有效率地使用预测原理,以便降低在语音信号中和一般在音频信号中存在的冗余。这会导致编码效率和质量的提高而不存在不可接受的延迟。通过使用通用化的预测,本发明还允许嵌入式编码。

附图说明

[0018] 通过结合附图,参照以下说明,可最好地理解本发明及其其它目的和优点,其中:

[0019] 图 1A 是因果编码的示意图;

[0020] 图 1B 是使用过去和将来信号样本编码的示意图;

[0021] 图 1C 是根据本发明的因果和非因果编码的示意图;

[0022] 图 2A 是示出开环预测编码的方框图;

[0023] 图 2B 是示出闭环预测编码的方框图;

[0024] 图 3 是示出自适应码本编码的方框图;

[0025] 图 4 是根据本发明的编码器和解码器设备实施例的方框图;

[0026] 图 5 是根据本发明的预测编码器和预测解码器的设备实施例的方框图;

[0027] 图 6 是通过使用残差参数的最优滤波和量化来增强主编码器的示意图;

[0028] 图 7 是利用非因果自适应码本范例的实施例的方框图;

[0029] 图 8 是在单个帧内使用非因果关系的示意图;

[0030] 图 9 是根据本发明的方法实施例的步骤的流程图;以及

[0031] 图 10 是估计的恶化质量的曲线图。

具体实施方式

[0032] 在本公开内容中论述了音频信号。然后,假定音频信号是在与某个时间相关联的连续信号样本中提供的。

[0033] 在使用预测模型对音频信号样本编码时,利用连续信号样本之间的关系去除冗余信息。图 1A 中示出了一个简单的草图,其示出一组信号样本 10,每个样本与某个时间相关联。当前信号样本 $s(n)$ 的编码是基于当前信号样本 $s(n)$ 及多个先前信号样本 $s(n-N)$ 、... $s(n-1)$ (原始信号样本或信号样本的表示) 而产生。此类编码称为因果编码 CE,这是因为它涉及在要将当前信号样本 $s(n)$ 编码的时间瞬间 (timeinstance) 之前可用的信息。描述信号样本 $s(n)$ 的因果编码 CE 的参数 T 随后被传送以便存储和 / 或实现最终用途。

[0034] 在当前信号样本与将来信号样本之间也存在关系。此类关系也可用于去除冗余。在图 1B 中,一个简单的草图示出了这些相关性。一般情况下,基于当前信号样本 $s(n)$ 、与时间 n 之前的时间相关联的信号样本 $s(n-1)$ 、...、 $s(n-N)$ 或其表示及与时间 n 之后的时间相关联的信号样本 $s(n+1)$ 、...、 $s(n+N)$ 或其表示,来进行时间 n 处的信号样本 $s(n)$ 的编码。涉及仅在要将信号样本编码的时间瞬间之后可用的信息的编码称为非因果编码 NCE。在其它说明中,如果应用预测编码,则也可使用术语逆测 (postdiction) 和溯测 (retrodiction)。

[0035] 由于利用了不同信号样本之间的更多关系,因此图 1B 中在时间 n 处进行的信号样本编码一般更可能优于图 1A 中提供的编码。但是,如图 1B 所示系统的主要缺点是,该编

码只在对应于 N^+ 个信号样本的某个时间延迟 D 后才可用,以便将以后信号样本中的信息也结合进去。此外,在将使用非因果编码的信号样本解码时,由于在此处也要收集“将来”的信号样本,因此,造成了附加的延迟。通常,由于为了将一个信号样本解码而需要过去和将来的已解码信号样本可用,因此,此方法不可能实现。

[0036] 根据本发明,介绍了在图 1C 中以示意图方式示出的另一非因果方法。此处,先提供基本上根据现有技术的因果编码 CE,以得到编码信号样本 $s(n)$ 的参数 P 并最终得到与其相关的解码信号。同时,为先前信号样本 $(n-N^+)$ 提供附加的非因果编码 NCE,以产生参数 NT 。如果时间和信令资源允许,可利用此附加的非因果编码 NCE 来升级或增强先前解码信号。如果此类延迟不可接受,则可忽略附加的非因果编码 NCE。如果进行已解码信号样本的升级,则确实会引入延迟。除此方法可实现外,也注意到,相对于图 1B 的编码方案,延迟减少了一半,这是因为在非因果编码到达时,在解码器处确实有全部必需的信号样本可用。将在此处下面的多个实施例中进一步描述和论述此基本设想。

[0037] 结合本设想使用的因果及非因果编码方案可以是利用连续信号样本之间的冗余的几乎任何种类方案。非排他性的示例有变换编码和 CELP 编码。因果和非因果编码的编码方案可能不一定是相同的类型,但在一些情况下,如果根据类似的方案进行这两种编码,则可能产生另外的优点。但是,在下面的实施例中,使用预测编码方案作为编码方案的模型示例。目前,预测编码方案也被视为是在本发明中使用的优选方案。

[0038] 为此,在陈述本发明的细节之前,先对现有技术的因果预测编码进行了一定深度的描述以提供科学基础。

[0039] 可区分用于去除冗余的两种类型的因果预测模型。第一种类型是所谓的开环因果预测,它基于原始音频信号样本。第二种类型是闭环因果预测,并且基于预测和重构的音频信号样本,即,原始音频信号样本的表示。

[0040] 基于利用开环因果预测的冗余去除过程的语音编解码器可如图 2A 所示大致视为基于典型预测的编码器和解码器的方框图。有关感觉加权的考虑在当前陈述中被忽略以便简化基本理解,并且因此未示出。

[0041] 作为开环预测的通用设置,通过使用以下关系,在预测器 20 中从先前的原始音频信号样本 $s(n-1)$ 、 $s(n-2)$ 、...、 $s(n-N)$ 预测提供到编码器 11 的因果预测编码器部分 16 的输入 14 的原始当前音频信号样本 $s(n)$:

$$[0042] \quad \hat{s}(n) = P(s(n-1), s(n-2), \dots, s(n-N)). \quad (1)$$

[0043] 此处, $\hat{s}(n)$ 表示用于 $s(n)$ 的开环预测,而 $P(\cdot)$ 是因果预测器,并且 N 是预测阶数。开环残差 $\tilde{e}(n)$ 在此处为减法器 22 的计算部件中定义为:

$$[0044] \quad \tilde{e}(n) = s(n) - \hat{s}(n). \quad (2)$$

[0045] 此处为量化器 30 的编码部件将搜索 $\tilde{e}(n)$ 的最佳表示 R 。通常,此类表示 R 的索引指向内部码本。表征预测器 20 的表示 R 和参数 F 提供到发射器 (TX) 40,并编码成当前音频信号样本 $s(n)$ 的编码表示 T 。编码表示 T 经存储以供将来使用,或者传送到最终用户。

[0046] 当前音频信号样本 $s(n)$ 的编码表示 T^* 的已接收版本由输入 54 接收到解码器 51 的因果预测解码器部分 56 的接收器 (RX) 41 中。在接收器 41 中,编码表示 T^* 解码成已接收残差 $\tilde{e}^*(n)$ 信号的已接收表示 R^* ,并解码成用于解码器预测器 21 的已接收参数 F^* 。理想情况下,编码表示 T^* 、已接收残差 $\tilde{e}^*(n)$ 信号的已接收表示 R^* 和已接收参数 F^* 等于编码器中

对应的量值。但是,可能存在传输误码,从而造成在已接收数据中的微小误码。因果预测解码器部分 56 的解码部件(此处为反量化器 31)提供已接收开环残差 e^* 。通常,接收内部码本索引并使用对应的码本条目。解码器预测器 21 由参数 F^* 启动,以便基于先前已接收音频信号样本 $s^*(n-1), s^*(n-2), \dots, s^*(n-N)$ 来提供预测 $\hat{s}(n)^*$:

$$[0047] \quad \hat{s}(n) = P(\bar{s}(n-1), \bar{s}(n-2), \dots, \bar{s}(n-N)). \quad (3)$$

[0048] 随后,当前已接收音频信号样本 $s^*(n)$ 在此处为加法器 23 的计算部件中计算为:

$$[0049] \quad \bar{s}(n) = \hat{s}(n) + \bar{e}(n). \quad (4)$$

[0050] 当前已接收音频信号样本 $s^*(n)$ 提供到解码器预测器 21 以供将来使用,并用作解码器 51 的输出 55 的输出信号。

[0051] 类似地,基于利用闭环因果预测的冗余去除过程的语音编解码器可如图 2B 所示大致视为基于典型预测的编码器和解码器的方框图。闭环残差信号可定义为在预测使用此处表示为 $s(n-1), s(n-2), \dots, s(n-N)$ 的重构音频信号样本、而不是原始音频信号样本时获得的信号。闭环预测在此情况下将书写为:

$$[0052] \quad \hat{s}(n) = P(\bar{s}(n-1), \bar{s}(n-2), \dots, \bar{s}(n-N)), \quad (5)$$

[0053] 并且闭环残差将书写为:

$$[0054] \quad e(n) = s(n) - \hat{s}(n). \quad (6)$$

[0055] 从 $e(n)$ 的表示 R,再次得到解码残差 $e(n)$,该残差在加法器 24 中与闭环预测 $\hat{s}(n)$ 相加,以便为预测器 20 提供重构音频信号样本 $s(n)$ 以供将来预测中使用。重构音频信号样本 $s(n)$ 因此是原始音频信号样本 $s(n)$ 的表示。

[0056] 在接收器侧,解码过程与图 2A 所示相同。

[0057] 等式 (1)、(3) 和 (5) 使用通用预测器,该预测器在一般情况下可以是非线性的。现有技术的线性预测,即使用线性预测器的估计经常用作语音和音频编解码器中去除冗余的部件。对于此类情况,预测器 $P(\cdot)$ 书写为其变元的线性函数。等式 (5) 随后变为:

$$[0058] \quad \hat{s}(n) = P(\bar{s}(n-1), \bar{s}(n-2), \dots, \bar{s}(n-N))$$

$$[0059] \quad = \sum_{i=1}^N a_i \bar{s}(n-i) \quad (7)$$

[0060] 系数 a_1, a_2, \dots, a_L 称为线性预测 (LP) 系数。大多数现代语音或音频编解码器使用时变 LP 系数以便适应音频信号的时变性质。LP 系数容易通过在自相关序列上应用例如莱文森-德宾算法估计得出,而序列是逐帧估计得出的。

[0061] 线性预测经常用于短期相关,LP 预测器的阶数一般不超过 20 个系数。例如,宽带语音编码 AMR-WB 的标准具有 16 阶的 LPC 滤波器。

[0062] 理论上,可在任何阶使用 LP 滤波器。但是,由于莱文森-德宾算法的数值稳定性及在存储空间和算术运算方面所导致许多的复杂性原因,此使用是极不可取的。另外,将 LP 系数编码所需的比特率禁止此类使用。

[0063] 为在保持质量的同时进一步降低所需的比特率量,需要正确地利用在有声语音段中的语音信号的周期性。为此,且由于线性预测一般将利用在不到一个基音周期中包含的相关,通常在线性预测残差上使用基音预测器。公知有两种不同的方法并且经常被使用以便利用语音信号中的长期相关性。

[0064] 第一种方法是基于自适应码本范例。自适应码本包含 LP 激励信号的最近过去的

重叠段。使用此方法,线性预测合成分析编码器通常将使用自适应码本贡献和固定码本贡献两者来对激励信号编码。

[0065] 就通过闭环长期预测从激励信号去除周期性和随后使用固定码本对提示信号编码而言,第二种方法更直接。

[0066] 两种方法实际上在概念上和实现方面均很类似。图 3 示出了例如由量化器 30(图 2A&B) 使用自适应 33 和固定 32 码本贡献提供的激励生成。在自适应码本方法中,在加法器 36 中作为两个分量的加权和来得到激励信号:

$$[0067] \quad \bar{e}_j(n) = g_{LTP} c'_{LTP}(n) + g_{FCB} c'_{FCB}(n) \quad (8)$$

[0068] 变量 g_{LTP} 34 和 g_{FCB} 35 分别表示自适应码本和固定码本增益。索引 j 表示固定码本 32 条目。索引 i 表示自适应码本 33 索引。此自适应码本 33 由作为最近合成的激励信号的先前段的条目组成:

$$[0069] \quad c'_{LTP}(n) = \bar{e}(n - d(i)) \quad (9)$$

[0070] 延迟函数 $d(i)$ 指定自适应码本向量的开始。由于复杂性原因,增益和索引的确定通常以顺序的方式执行。首先,找到自适应码本贡献,即,对应的索引及增益。随后,视特定实现而定,在从目标激励信号或加权语音减除后,找到固定码本的贡献。

[0071] 在优化器 19 中通过将待量化的残差信号 $e(n)$ 与 $\bar{e}(n)$ 比较,找到一组最优码本参数。残差信号的最佳表示 R 在此类情况下通常将包括 g_{LTP} 、 g_{FCB} 、 c'_{FCB} 和延迟函数 $d(i)$ 。

[0072] 自适应码本范例也具有滤波器解释,其中使用了基音预测器滤波器,并且通常书写为:

$$[0073] \quad \frac{1}{P(z)} = \frac{1}{1 - g_{LTP} z^{-d(i)}} \quad (10)$$

[0074] 也存在相同概念的几种变化,如在延迟函数不限于整数基音延迟 (pitch delay),而是也可包含分数延迟时。另外的变化是多抽头基音预测,它很类似于分数基音延迟,这是因为这两种方法都使用多抽头滤波器。另外,这两种方法产生极其类似的结果。通常,阶数为 $2q+1$ 的基音预测器由如下表达式给出:

$$[0075] \quad P(z) = 1 - \sum_{k=-q}^q b_k z^{-D+k} \quad (11)$$

[0076] 几种最新标准化的编解码器使用上述结构进行语音编码。众所周知的示例包括 3GPP AMR-NB 和 3GPP AMR-WB 编解码器。另外,AMR-WB+ 混合结构的 ACELP 部分也使用此类结构来对语音和音频进行高效编码。

[0077] 通常,在开环中估计整数基音延迟,以使得原始信号与其预测值之间的平方误差降到最低。原始信号此处是在广义方面理解,因此,也可使用加权。在允许的基音范围 (2 到 20ms) 中使用穷举搜索。

[0078] 本发明的一个重要概念是使用非因果编码,并且在优选实施例中,使用非因果预测编码作为用于去除冗余的部件和用于编码的部件。非因果预测也可称为反向时间预测。非因果预测可以为线性和非线性的。使用线性预测时,非因果预测包括例如非因果基音预测,但也可由非因果短期线性预测表示。更简单地说,将来的信号用于形成当前信号的预测。但是,由于将来通常在编码时是不可用的,因此,经常使用延迟以便访问信号的将来样本。非因果预测随后变成基于当前信号和 / 或在要预测的信号后出现的其它先前信号来预

测先前信号。

[0079] 在用于非因果预测的通用设置中,通过使用以下等式,从将来的信号样本 $s(n+1)$ 、 $s(n+2)$ 、...、 $s(n+N^+)$ 来预测原始语音信号样本 $s(n)$ 或一般为音频信号样本或者甚至任何信号样本:

$$[0080] \quad \hat{s}^+(n) = P^+(s(n+1), s(n+2), \dots, s(n+N^+)) \quad (12)$$

[0081] 此处, $\hat{s}^+(n)$ 表示用于 $s(n)$ 的非因果开环预测。上标 (+) 在此情况下用于将它与“普通”开环预测区分,并且为完整性起见,此处“普通”开环预测使用上标 (-) 重新书写为:

$$[0082] \quad \hat{s}^-(n) = P^-(s(n-1), s(n-2), \dots, s(n-N^-)) \quad (13)$$

[0083] 非因果和因果预测器表示为 $P^+(\cdot)$ 和 $P^-(\cdot)$, 并且预测器阶数分别表示为 N^+ 和 N^- 。

[0084] 同样地,开环残差可定义为:

$$[0085] \quad \tilde{e}^+(n) = s(n) - \hat{s}^+(n) \quad (14)$$

$$[0086] \quad \tilde{e}^-(n) = s(n) - \hat{s}^-(n)$$

[0087] 也可类似地定义闭环残差。对于因果预测的情况,此类定义与更上面给出的定义完全相同。但是,对于非因果预测,由于编码器实质上是因果过程,所以尽管有某个延迟,但此类定义不可能使用通过相同非因果预测、甚至通过使用另外延迟引起的预测。实际上,编码器使用非因果预测以便将样本编码,这将取决于将来的编码。因此可观察到,除非我们翻转时间之箭,否则非因果预测无法直接作用于编码或去除冗余的部件,但在该情况下,它将成为具有反向时间语音的因果预测。

[0088] 但是,可以在闭环中以间接方式有效率地使用非因果预测。一个此类实施例是主要通过因果预测器 $P^-(\cdot)$ 对信号进行编码,之后基于由因果预测器 $P^-(\cdot)$ 预测的信号,以后向闭环方式使用非因果预测器 $P^+(\cdot)$ 。

[0089] 在图 4 中,示出了应用到语音或音频编码的非因果编码的一个实施例。主编码器和非因果预测的组合作用于编码和去除冗余的部件。在本实施例中,利用了非因果预测编码,并且利用了因果预测编码为主编码。编码器 11 在输入 14 处接收信号样本 10。主编码部分,此处为因果编码部分 12,具体而言在此实施例中为因果预测编码部分 16,接收当前信号样本 10,并产生当前音频信号样本 $s(n)$ 的编码表示 T,该表示在输出 15 处提供。当前信号样本 10 也提供到非因果编码部分 13,在此实施例中是非因果预测编码部分 17。非因果预测编码部分 17 在输出 15 处提供先前音频信号样本 $s(n-N^-)$ 的编码增强表示 ET。非因果预测编码部分 17 可将其运算也基于从因果预测编码部分 16 提供的信息 18。

[0090] 在解码器 51 中,在输入 54 处接收当前音频信号样本 $s(n)$ 的编码表示 T^* 及先前音频信号样本 $s(n-N^-)$ 的编码增强表示 ET^* 。已接收编码表示 T^* 提供到主因果解码部分,此处为因果解码部分 52,并且具体而言在此实施例中为因果预测解码部分 56。因果预测解码部分 56 提供当前已接收音频信号样本 $s^-(n)$ 55⁻。编码增强表示 ET^* 提供到非因果解码部分 53,在此实施例中为非因果预测解码部分 57。非因果预测解码部分 57 提供增强的先前已接收音频信号样本。基于增强的先前已接收信号样本,先前已接收音频信号样本 $s^*(n-N^+)$ 在信号调节器 59 中增强,该调节器可以是非因果预测解码部分 57 的一部分或单独的部分。增强的先前已接收音频信号样本 $\tilde{s}^-(n-N^+)$ 在解码器 51 的输出 55⁺ 处提供。

[0091] 在图 5 中,示出了应用到音频编码的非因果闭环预测的又一详细实施例。因果预测器部分易于从图 2B 识别。但在图 5 中,示出了非因果预测器 120 如何使用主编码语音信

号 18 的将来样本。对应的样本 58 在用于非因果预测器 121 的解码器 51 中也可用。当然，要施加延迟才能访问这些样本。

[0092] 组合器 125 也引入了一个另外的“组合”函数。组合器 125 的函数由基于闭环因果预测将主编码信号（即， $s^-(n-N^+)$ ）和取决于 $s^-(n)$ 的以后样本的非因果预测器的输出组合而组成，即：

$$[0093] \quad \hat{s}^+(n-N^+) = P^+(\bar{s}^-(n-N^++1), \bar{s}^-(n-N^++2), \dots, \bar{s}^-(n)) \quad (15)$$

[0094] 此组合可以是线性或非线性的。此模块的输出可书写为

$$[0095] \quad \tilde{s}(n-N^+) = C(\hat{s}^+(n-N^+), \bar{s}^-(n-N^+)) \quad (16)$$

[0096] 优选地，组合函数 $C(\cdot)$ 经选择以将由此处为减法器 122 的计算部件提供的组合信号 $\tilde{s}(n-N^+)$ 与原始语音信号 $s(n-N^+)$ 之间的结果误差降到最低，并且该误差定义为：

$$[0097] \quad \tilde{e}(n-N^+) = s(n-N^+) - \tilde{s}(n-N^+) \quad (17)$$

[0098] 误差最小化在此处象通常一样是就某个预定的保真度标准而言的广义理解，如均方误差 (MSE) 或加权均方误差 (wMSE) 等。此结果误差残差在此处为量化器 130 的编码部件中进行量化，以便提供音频信号样本 $s(n-N^+)$ 的编码增强表示 ET。

[0099] 结果误差也可以量化，以使得结果语音信号

$$[0100] \quad \tilde{s}(n-N^+) = \tilde{e}(n-N^+) + \tilde{s}(n-N^+) \quad (18)$$

[0101] 就所述预定保真度标准而言尽可能接近原始语音信号。

[0102] 最后，应注意，预测器 $P^-(\cdot)$ 20 和 $P^+(\cdot)$ 120 及组合函数 $C(\cdot)$ 125 可以是随时间变化的并选择为符合原始语音信号的时变特征和 / 或针对保真度标准为最优的。因此，控制这些函数的时变参数也要编码并由发射器 140 传送。在解码器中接收到时，使用这些参数来使得能够解码。

[0103] 在解码器侧，非因果预测解码部分 57 在接收器 141 中接收编码增强表示 ET^* ，并且通过此处为量化器 131 的解码部件将它解码成残差样本信号。编码增强表示 ET^* 的其它参数用于非因果解码器预测器 121 以便产生预测增强信号样本。此预测增强信号样本在组合器 126 中与主预测信号样本组合，并在此处为加法器 123 的计算部件中与残差信号相加。组合器 126 和加法器 123 此处一起构成了信号调节器 59。

[0104] 与一般非线性预测相比，线性预测具有更低的复杂性，并且使用更简单。另外，一个常识是，作为用于语音信号产生的模型，线性预测是绰绰有余的。

[0105] 在前面部分中，假定预测器 $P^-(\cdot)$ 和 $P^+(\cdot)$ 及组合函数 $C(\cdot)$ 是通用的。实际上，简单的线性模型经常用于这些函数。预测器成为线性滤波器，类似于等式 (7)，而组合函数成为加权和。

[0106] 理论上，如果信号稳定，并且两个预测器使用相同的阶数，则在开环中使用相同窗口估计时的因果和非因果预测器将产生一组相同的系数。原因是，线性预测滤波器是线性相位，因此前向和后向预测误差具有相同的能量。这实际上由低延迟语音编解码器用于从过去解码语音信号得到 LPC 滤波器系数，例如，LD-CELP。

[0107] 与后向线性预测相反，非因果线性预测在一般情况下将重新估计要在一组相同的解码语音样本上应用的新“后向预测”滤波器，由此将在第一“主”编码期间发生的频谱变化考虑在内。另外，信号的非稳定性在第二遍时在增强编码器中被正确考虑在内。

[0108] 本发明很适用于分层的语音编码。首先，简短概述现有技术的分层编码。

[0109] 语音编码中的可伸缩性通过与通用音频编码相同的轴实现：带宽、信噪比和空间（信道数量）。但是，由于语音压缩主要用于多信道操作仍不是很常见的会话型通信目的，因此关于语音编码可伸缩性的最大关注集中在 SNR 和音频带宽可伸缩性上。SNR 可伸缩性始终是在总是互连到固定带宽 8kHz PSTN 的传统交换网络中的主要关注点。该 SNR 可伸缩性在处理临时拥塞情况中，例如在部署成本高和相对低带宽的大西洋通信电缆中体现其应用。近年来随着可用高端终端的出现，支持更高采样率、带宽可伸缩性已成为现实可能。

[0110] 今天使用最多的可伸缩语音压缩算法是 64kbps G. 711A/U-law 对数 PCM 编解码器。8kHz 采样的 G. 711 编解码器将 12 位或 13 位线性 PCM 样本转换为 8 位对数样本。对数样本的有序位表示允许在 G. 711 位流中窃取最低有效位 (LSB)，以使得 G. 711 编解码器实践中在 48、56 和 64kbps 之间实施 SNR 可伸缩。G. 711 编解码器的此可伸缩性在电路交换通信网络中用于实现带内控制信令的目的。使用此 G. 711 伸缩性的最近示例是 3GPP-TF0 协议，该协议使得能够通过传统 64kbps PCM 链路进行宽带语音建立和传输。原 64kbps G. 711 流的 8kbps 最初用于允许宽带语音服务的呼叫建立，而对窄带服务质量没有大的影响。在呼叫建立后，宽带语音将使用 64kbps G. 711 流的 16kbps。支持开环可伸缩性的其它更早的语音编码标准是 G. 727（嵌入式 ADPCM）和在某个程度上的 G. 722（子带 ADPCM）。

[0111] 可伸缩语音编码技术中较新的发展是 MPEG-4 标准，该标准在 SNR 域和带宽域中均为 MPEG4-CELP 提供了可伸缩性扩展。MPE 基层可通过传输附加的滤波器参数信息或另外的创新参数信息而得以增强。在 MPEG4-CELP 概念中，“BRSEL”类型的增强层是用于选定基层的 SNR 增加层，“BWSEL”层是带宽增强层，它们使得提供 16kHz 输出成为可能。结果是非常灵活的编码方案，比特率范围在不连续的阶段从 3.85 到 23.8kbps。但是，MPEG-4 语音编码器验证测试 确实显示，与固定多模式（不可伸缩）运算相比，可伸缩性所允许的额外灵活性要付出成本。

[0112] 国际电信联盟标准化部门 ITU-T 最近结束了一个名为 G. 729. EV 的新的可伸缩编解码器的认证期。这种将来的可伸缩语音编解码器的比特率范围将从 8kbps 到 32kbps。该编解码器将提供 8-12kbps 的窄带 SNR 可伸缩性、12-14kbps 的带宽可伸缩性及以 2kbps 为步长从 14kbps 到最高 32kbps 的 SNR 可伸缩性。此编解码器的主要使用情况是允许在家庭或办公室网关中有效率地共享有限的带宽资源，例如，在几个 VoIP 呼叫之间共享的 xDSL 64/128 kbps 上行链路。另外，8kbps 核将可与现有 G. 729 VoIP 终端互操作。

[0113] 图 10 示出了基于即将出现的标准的初始认证结果的估计恶化质量曲线。它示出了估计的 G. 729. EV 性能 (8(NB)/16(WB) kHz Mono)。

[0114] 除 G. 729. EV 开发外，ITU-T 在研究组 16 问题 9 中正计划开发具有 8kbps 带宽核的新的可伸缩编解码器，并且也在问题 23 中讨论一个新的工作项，即，在保持一些可伸缩性特性的同时开发完全听觉带宽编解码器。

[0115] 如果将因果、非因果和组合函数重新表示为一个运算，则可将输出表示为：

$$[0116] \quad \tilde{s}(n) = \sum_{i=-N^-}^{N^+} b_i \tilde{s}^-(n+i) \quad (19)$$

[0117] 因此可看到，使用最优因果和非因果预测器类似于将双面滤波器应用到主编码信号。双面滤波器已应用到在不同环境中的音频信号。例如，在美国专利 6738739 中介绍了一种使用平滑处理的预处理步骤，该平滑处理利用了前向和后向基音扩展。但是，整个滤波

器全部在一个相同的时刻应用,这意味着会引入时间延迟。此外,滤波器在编码器中用于平滑处理的目的,并且不涉及实际的预测过程。

[0118] 在欧洲专利申请 EP 0 532 225 中,公开了一种处理信号的方法。该方法涉及输入信号样本的优选不超过 5 毫秒的编码帧,优选以小于 每秒 16Kb 采样,编码延迟优选不超过 10 毫秒。具有相应索引信号的每个码本向量通过增益系数调整,优选通过后向自适应调整,并且应用到级联的长期和短期滤波器以生成合成候选信号。对应于最佳地逼近相关联帧的候选信号的索引和所得到的长期滤波器(例如基音)参数可供用于随后将帧解码。短期滤波器参数随后可通过后向自适应而得到。同样,在此处,整个滤波器在一个完整的过程中应用,并且应用到已经解码的信号,即,在预测编码或解码过程中不应用它。

[0119] 与此相反,在本发明中,等式 (19) 描述的运算首先在时间上划分,在该方面,第一个最初结果由主编码器在一个时间实现,并且改进或增强随后由非因果预测编码器提供。这是使该运算适用于分层音频编码的属性。此外,该运算是预测编码过程的一部分,并且因此在此“发射”侧和“接收器”侧、或更一般地说在编码和解码侧均执行。虽然 EP 0 532 225 初看起来可能与本发明有一些相似,但本文涉及完全不同的方面。

[0120] 图 6 示出了使用本发明原理的嵌入式编码结构。该图示出通过使用最优滤波的主编码器增强,由此将残差量化 (TX) 参数传送到解码器。此结构所基于的是基于主编码器的“本地合成”的输出来预测原始语音或音频信号 $s(n)$ 。这表示为 $\hat{s}_0(n)$ 。

[0121] 在通过 k 索引的每个级或增强层,得到滤波器 $W_{k-1}(z)$, 并将它应用到一个先前的层信号 $\hat{s}_{k-1}(n)$ 的“本地合成”,由此产生预测信号 $\tilde{s}_k(n)$ 。滤波器一般可以为因果、非因果的或双面的、HR 或 FIR。因此,此基本实施例对滤波器类型无限制。

[0122] 滤波器的得到使得预测误差:

$$[0123] \quad e_{k-1}(n) = s(n) - \tilde{s}_k(n) = s(n) - W_{k-1}(z)\hat{s}_{k-1}(n) \quad (20)$$

[0124] 就某个预定的保真度标准而言降到最低。还通过可以是与层相关的量化器 Q_{k-1} 将预测的残差量化和编码。这产生了量化预测误差:

$$[0125] \quad e_{k-1}(n) = Q_{k-1}(e_{k-1}(n)). \quad (21)$$

[0126] 该量化预测误差用于形成当前层的本地合成,该本地合成将用于下一层。

$$[0127] \quad \hat{s}_k(n) = \bar{e}_{k-1}(n) + W_{k-1}(z)\hat{s}_{k-1}(n) \quad (22)$$

[0128] 表示预测滤波器 $W_0(z), W_1(z), \dots, W_{k_{\max}}(z)$ 和量化器 $Q_0, Q_1, \dots, Q_{k_{\max}}$ 输出索引的参数会被编码和传送,以便在解码器侧,使用这些参数来将信号解码。

[0129] 此处应注意的是,通过剥离上部层,解码仍可进行,但是其质量低于在将所有层解码时获得的质量。

[0130] 通过每个附加层,本地合成将越来越接近原始语音信号。预测滤波器将接近于一致,而预测误差将趋向于零。

[0131] 概括而言,任何信号 $\hat{s}_0(n)$ 到 $\hat{s}_{k-1}(n)$ 都可视为通过主编码信号 $s(n)$ 而产生的信号,并且随后的信号可视为增强信号。主编码因此在一般情况下不一定只由因果分量组成,而是还可包括非因果贡献。

[0132] 滤波器与预测误差之间的这种关系可有效率地用于为预测滤波器和量化器两者统一地量化和分配位。来自主编码语音的预测用于估计原始语音。此预测的残差也可编码。此过程可重复进行,并因此提供语音信号的分层编码。

[0133] 本发明利用了此基本实施例。根据本发明,第一层包括用于提供第一近似信号的因果滤波器。此外,附加层至少之一包括非因果滤波器,它对已解码信号质量的增强给予贡献。此增强可能性由于非因果的原因而在后面的阶段才提供,并结合以后信号样本的以后因果滤波器编码来提供。根据本发明的此实施例,非因果预测用作用于嵌入式编码或分层编码的部件。因此,其中,附加层包含用于形成非因果预测的参数。

[0134] 上面进一步描述了现有技术的合成分析语音编解码器。此外,图 3 示出了在当前最新技术语音编解码器中使用的自适应码本范例背后的现有技术设想。此处下文介绍了如何通过使用称为非因果自适应码本范例的备选实现来在类似的编解码器中实施本发明。

[0135] 图 7 示出用于非因果自适应码本的当前优选实施例。此码本基于先前得到的主码本激励 $\bar{e}_y(n)$ 。索引 i 和 j 涉及每个码本的条目。

[0136] 提供利用因果自适应码本方法的主激励码本 39 作为因果预测编码部分 16 的量化器 30。不同的部分等同于前面结合图 3 所述内容。但是,不同的参数此处提供有“-”号以强调它们在因果预测中使用。

[0137] 提供利用非因果自适应码本方法的次要激励码本 139 作为非因果预测编码部分 17 的量化器 130。次要激励码本 139 的主要部分类似于主激励码本 39。自适应码本 133 和固定码本 132 分别提供具有自适应码本增益 g_{LTP}^+ 34 和固定码本增益 g_{FCB}^+ 35 的贡献。在加法器 136 中得到组成的激励信号。

[0138] 此外,如连接 37 所示,非因果自适应码本 133 还基于主激励码本 39。它使用自适应码本的将来样本作为条目,并且此非因果自适应码本 133 的输出可书写为:

$$[0139] \quad \tilde{e}_{y \rightarrow k}(n) = \bar{e}_y(n + d^+(k)) \quad (23)$$

[0140] 映射函数 $d^+(\cdot)$ 将对应的正数延迟指定到对应于后向或非因果基音预测的每个索引。该运算产生非因果 LTP 预测。

[0141] 最终的激励等于主激励和非因果自适应码本贡献及可能的次要固定码本贡献的加权线性组合。

$$[0142] \quad \tilde{e}_{y \rightarrow M}(n) = g_{LTP}^+ \bar{e}_y(n + d^+(k)) + g_{FCB}^+ c_l(n) + g_{\bar{c}} \bar{e}_y(n) \quad (24)$$

[0143] 主激励因此提供有增益 $g_{\bar{c}}$ 137 并在加法器 138 中与非因果自适应码本 133 贡献和来自次要固定码本 132 的贡献相加。增益和索引的优化和量化使得保真度标准得以优化。

[0144] 虽然只描述了码本的构成,但应注意的是,非因果基音延迟可以是分数的,由此从增大的分辨率受益并因而产生更佳的性能。情况明显与因果基音预测的情况相同。此处,也可使用多抽头基音预测器。

[0145] 非因果预测此处是在闭环中使用,因此基于原始语音信号的主编码。由于信号的主编码包括因果预测,因此可重复使用诸如基音延迟等的一些作为语音信号特征的参数,而无需额外的比特率成本,以形成非因果预测。

[0146] 具体而言,结合自适应码本范例,应注意的是,情况经常是不需要重新估计基音,而是直接重复使用为因果预测估计的相同基音延迟。这在图 7 中由虚线 38 示出。这节省了比特率,而对质量无太大的影响。

[0147] 对此过程的一个细化由只重复使用整数基音延迟、然后重新优化基音的分数部分组成。

[0148] 通常,即使重新估计基音延迟,但如果考虑到非因果基音极接近因果基音,则复杂

性及将此变量编码所需的位数量也会大大降低。因此,可有效率地应用诸如差分编码等技术。在复杂性方面,应明白的是,不必搜索所有基音范围。可只搜索因果基音周围的几个预定区域。总之,因此可以使映射函数 $d^+(\cdot)$ 自适应地与主基音变量 $d(i)$ 相关。

[0149] 仅在有个量的延迟时,才可应用非因果自适应码本的原理。实际上,需要将将来编码激励的样本以便形成增强激励。

[0150] 在逐帧基础上操作语音编解码器时,有某个提前量可用。帧通常划分成子帧。例如,在主编码信号帧之后,在第一子帧的增强编码器可访问整个帧的激励样本而无附加的延迟。如果非因果基音延迟相对较小,则通过增强编码器来完成第一子帧编码不会有额外的延迟。这也适用于如图 8 中所示的第二、第三帧,其中图 8 示出在逐帧基础上执行非因果基音预测。在此示例中,在第四子帧处,可能需要下一帧的样本,并且这将需要附加的延迟。

[0151] 如果不允许延迟,则仍可使用非因果自适应码本,但它将只对一些子帧而非所有子帧有效。因此,自适应码本使用的位数量将是可变的。活动和非活动状态的信令可以是隐含的,这是因为在收到基音延迟变量时,解码器会自动检测是否需要将来的信号样本。

[0152] 可考虑对上述实施例进行几个细化,如平滑处理预测滤波器参数的内插,使用加权误差测量和心理声学误差测量。这些细化和其它细化是本领域技术人员公知的原理,在此将不详细描述。

[0153] 图 9 示出根据本发明的方法实施例的步骤的流程图。用于音频编码和解码的方法从步骤 200 开始。在步骤 210 中,将当前音频信号样本因果编码成当前音频信号样本的编码表示。在步骤 211 中,将第一先前音频信号样本非因果编码成该第一先前音频信号样本的编码增强表示。在步骤 220 中,将当前音频信号样本的编码表示和第一先前音频信号样本的编码增强表示提供到最终用户。此步骤可视为由以下步骤组成:通过编码器提供当前音频信号样本的编码表示和第一先前音频信号样本的编码增强表示;和在最终用户处通过解码器获得当前音频信号样本的编码表示和第一先前音频信号样本的编码增强表示。在步骤 230 中,当前音频信号样本的编码增强表示被因果解码成当前接收的音频信号样本。在步骤 231 中,第一先前音频信号样本的编码增强表示被非因果解码成增强的第一先前已接收(received)音频信号样本。最后,在步骤 240 中,基于第一先前已接收音频信号样本和增强的第一先前已接收音频信号样本,改进对应于第一先前音频信号样本的第一先前已接收音频信号样本。此过程在步骤 299 中结束。如虚线箭头 250 所示,此过程基本上在音频信号的整个持续时间内重复进行。

[0154] 本公开内容介绍了一种自适应码本,其特征在于使用非因果基音贡献来形成非因果自适应码本。此外,介绍了增强激励,该激励是主编码激励和至少一个非因果自适应码本激励的组合。还示出了嵌入式语音编解码器,其特征在于,每层包含用于形成预测信号的至少一个预测滤波器、用于量化预测残差信号的量化器或编码器及用于形成本地合成增强信号的部件。对于解码器也提供了类似的部件和功能。此外,还描述了具有隐含信令的可变速率非因果自适应码本形成。

[0155] 上述实施例应理解为是本发明的少数几个说明性示例。本领域的技术人员将理解,在不脱离本发明范围的情况下,可对实施例进行不同的修改、组合和更改。具体地说,不同实施例中的不同部分解决方案可在技术上可行的情况下在其它配置中组合在一起。然而,本发明的范围由随附权利要求定义。

-
- [0156] 参考文献
- [0157] 美国专利 6738739。
- [0158] 欧洲专利申请 EP 0 532 225。

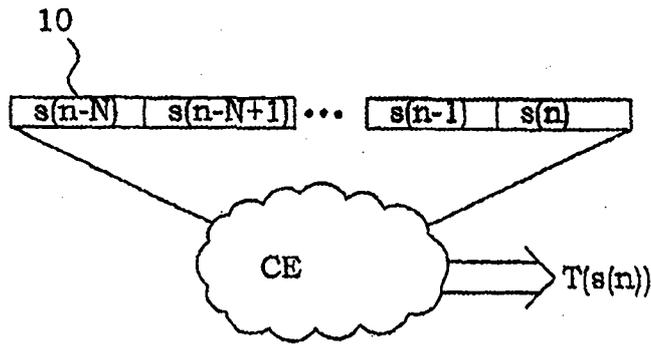


图 1A

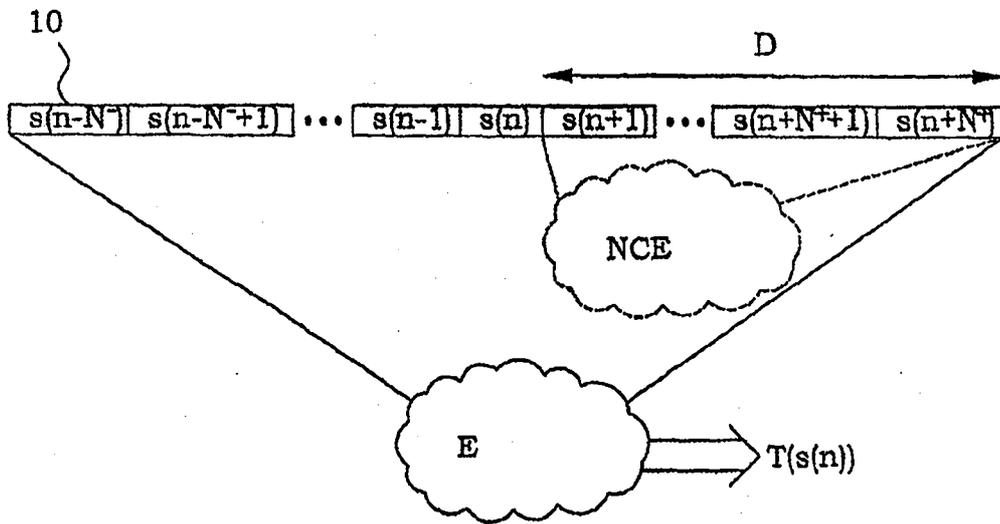


图 1B

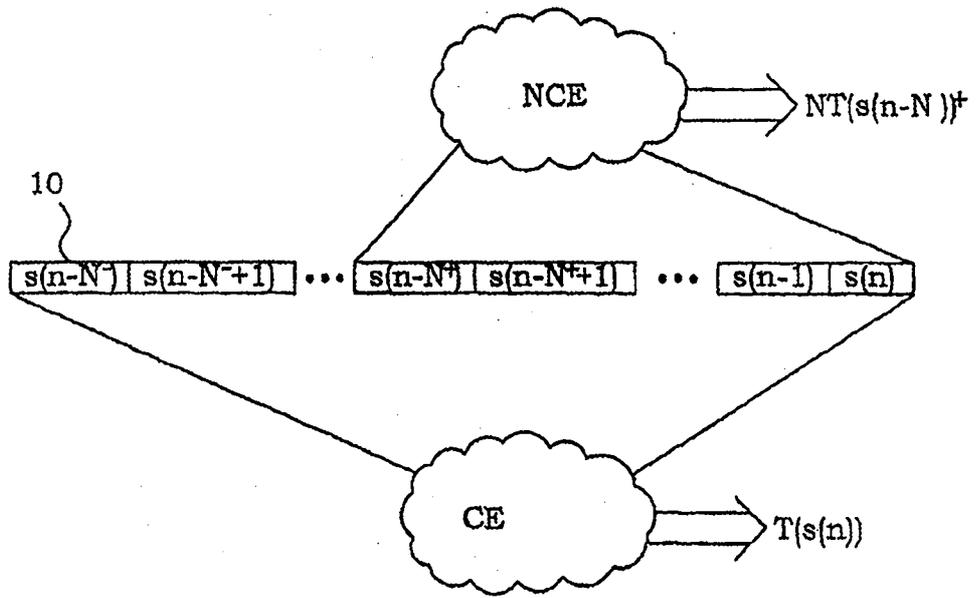


图 1C

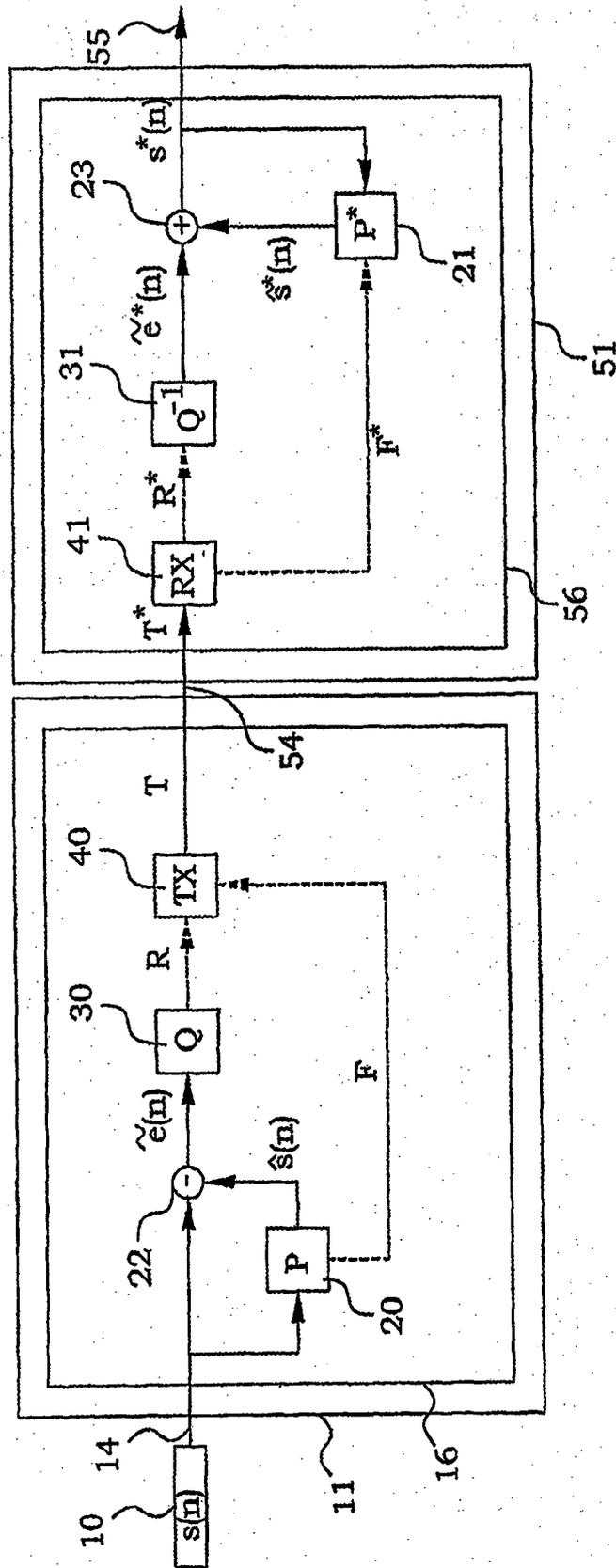


图 2A

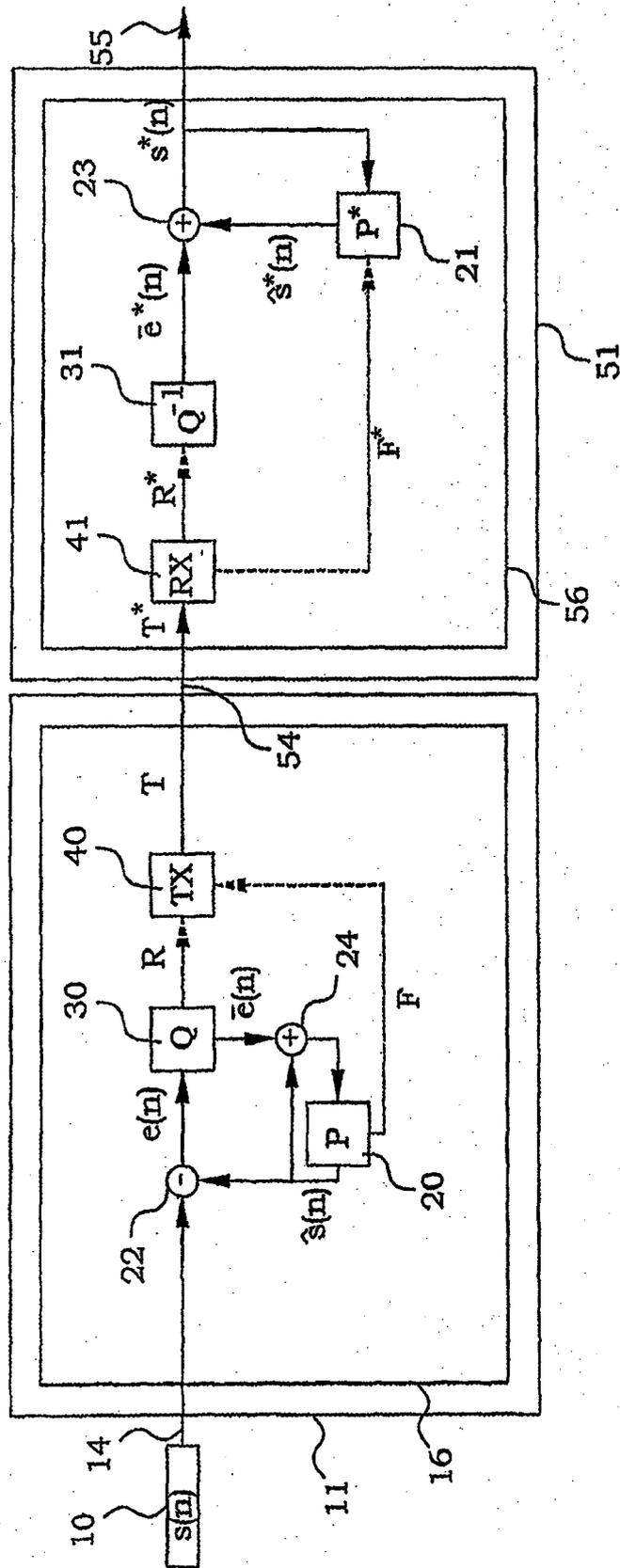


图 2B

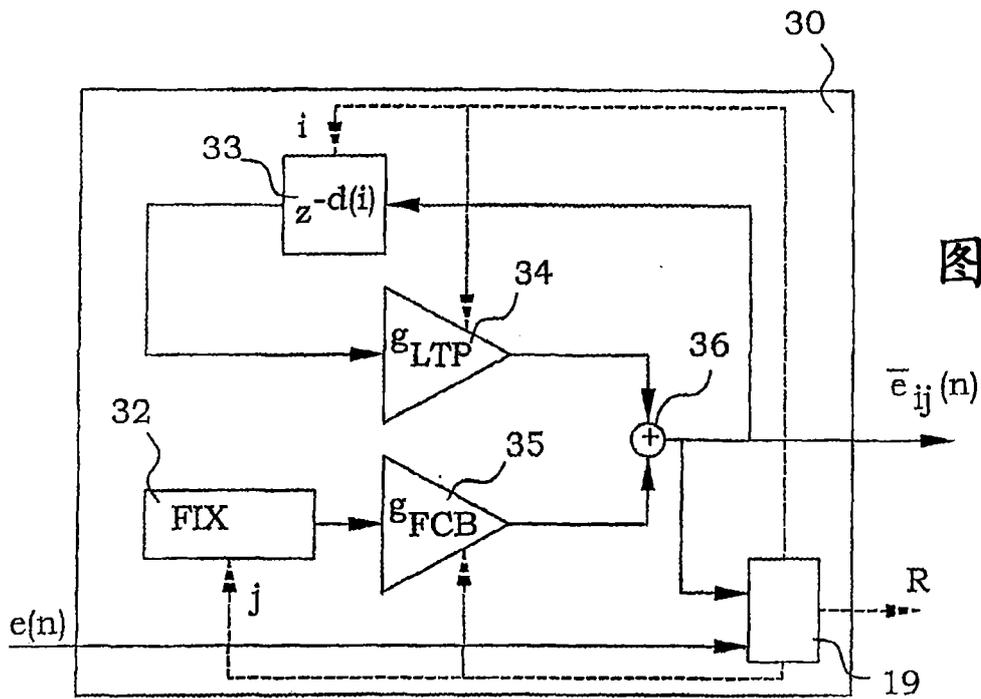


图 3

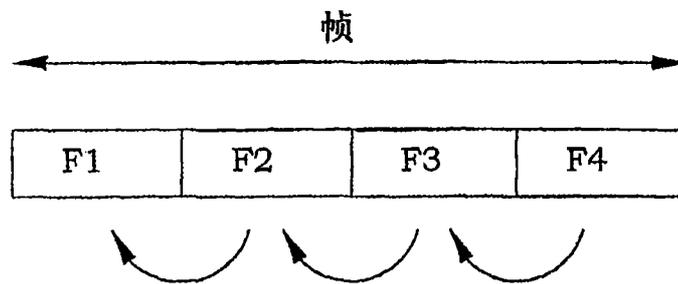


图 8

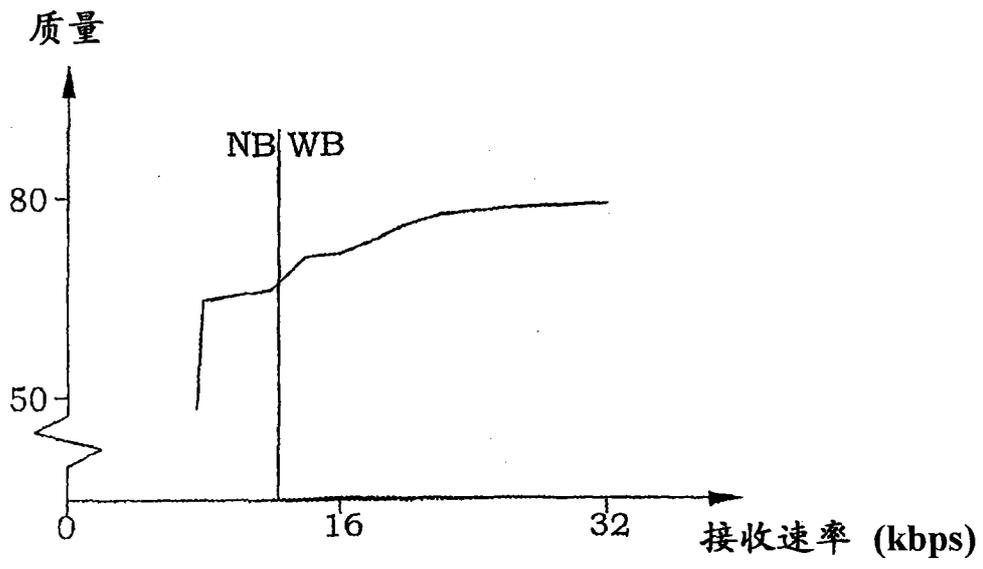


图 10

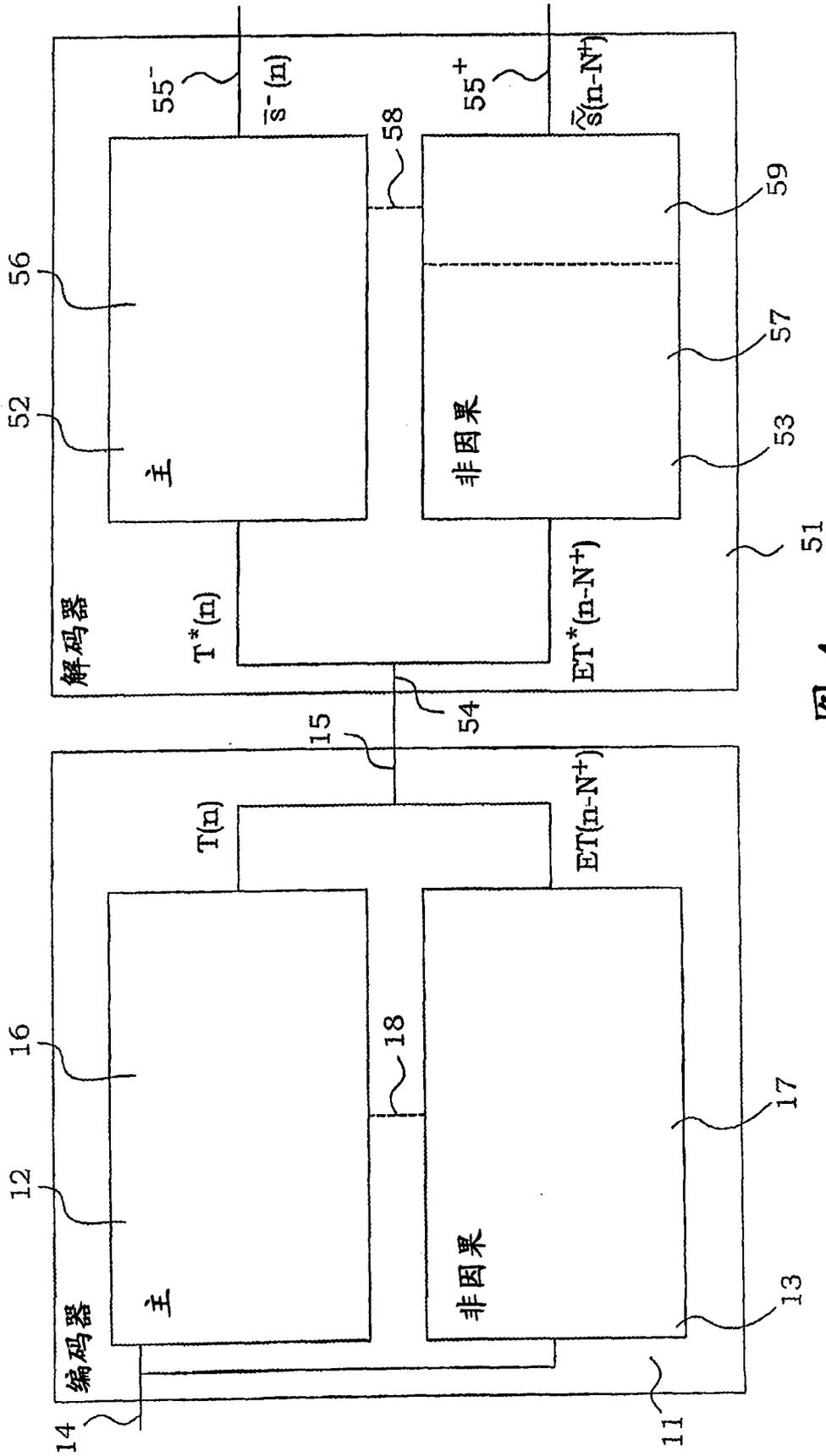


图 4

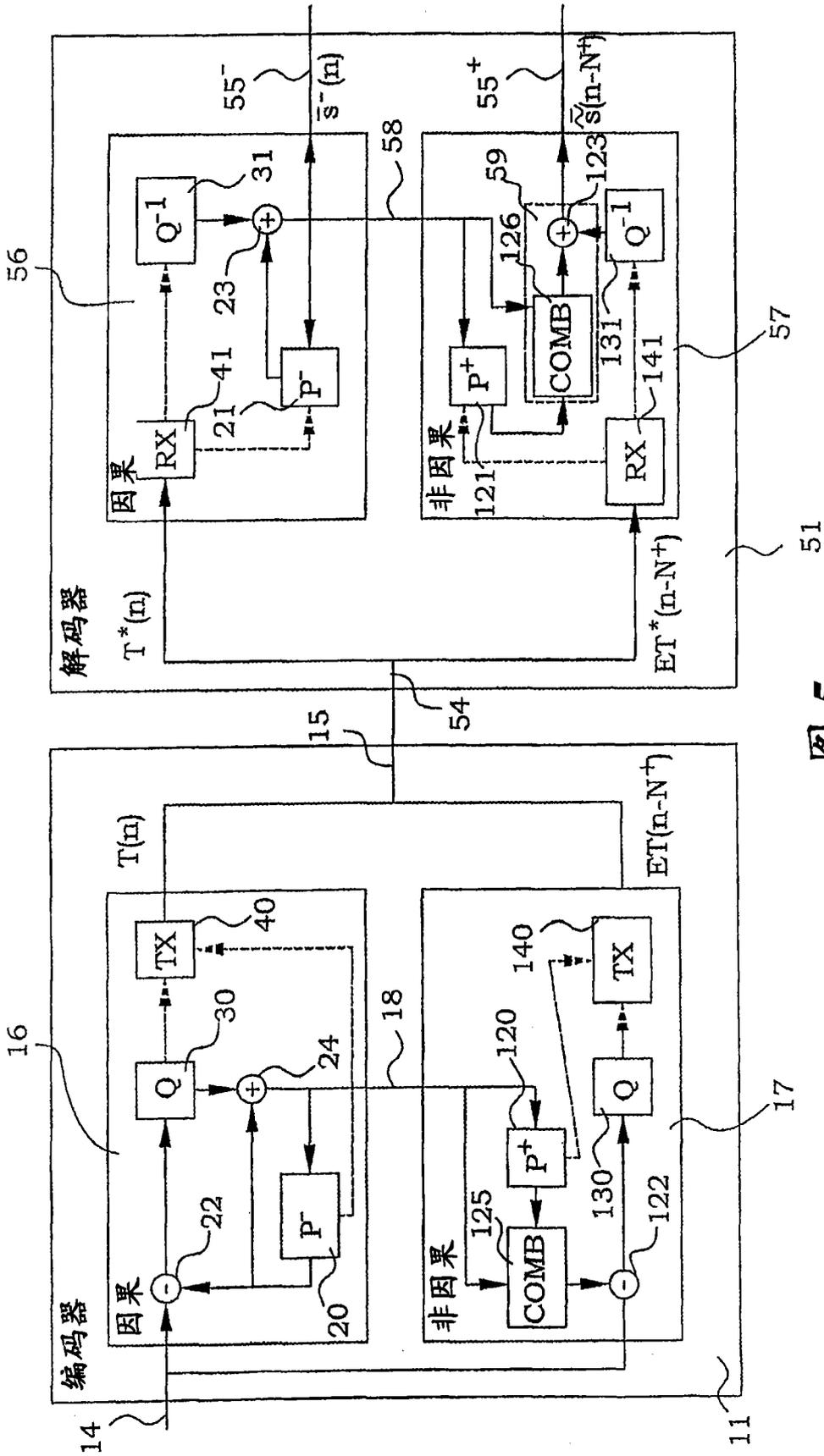


图 5

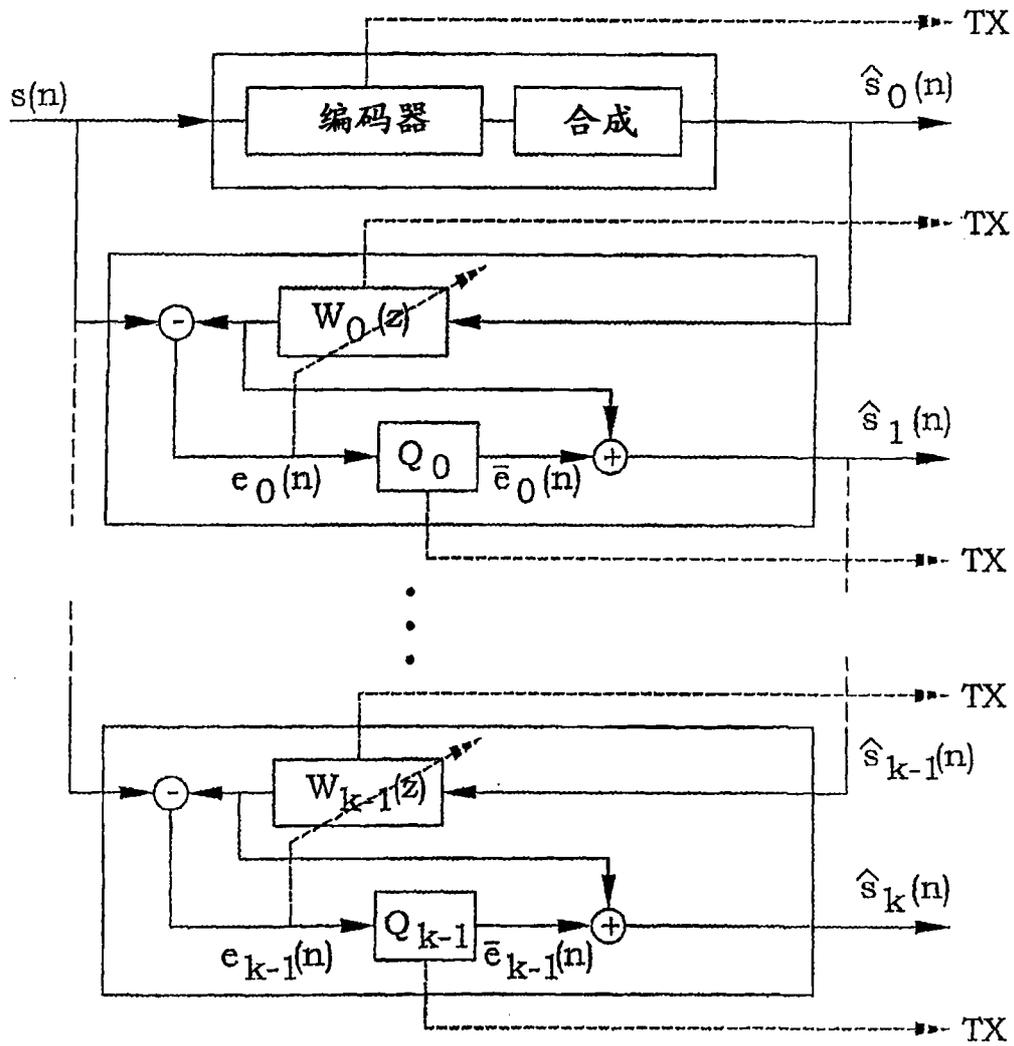


图 6

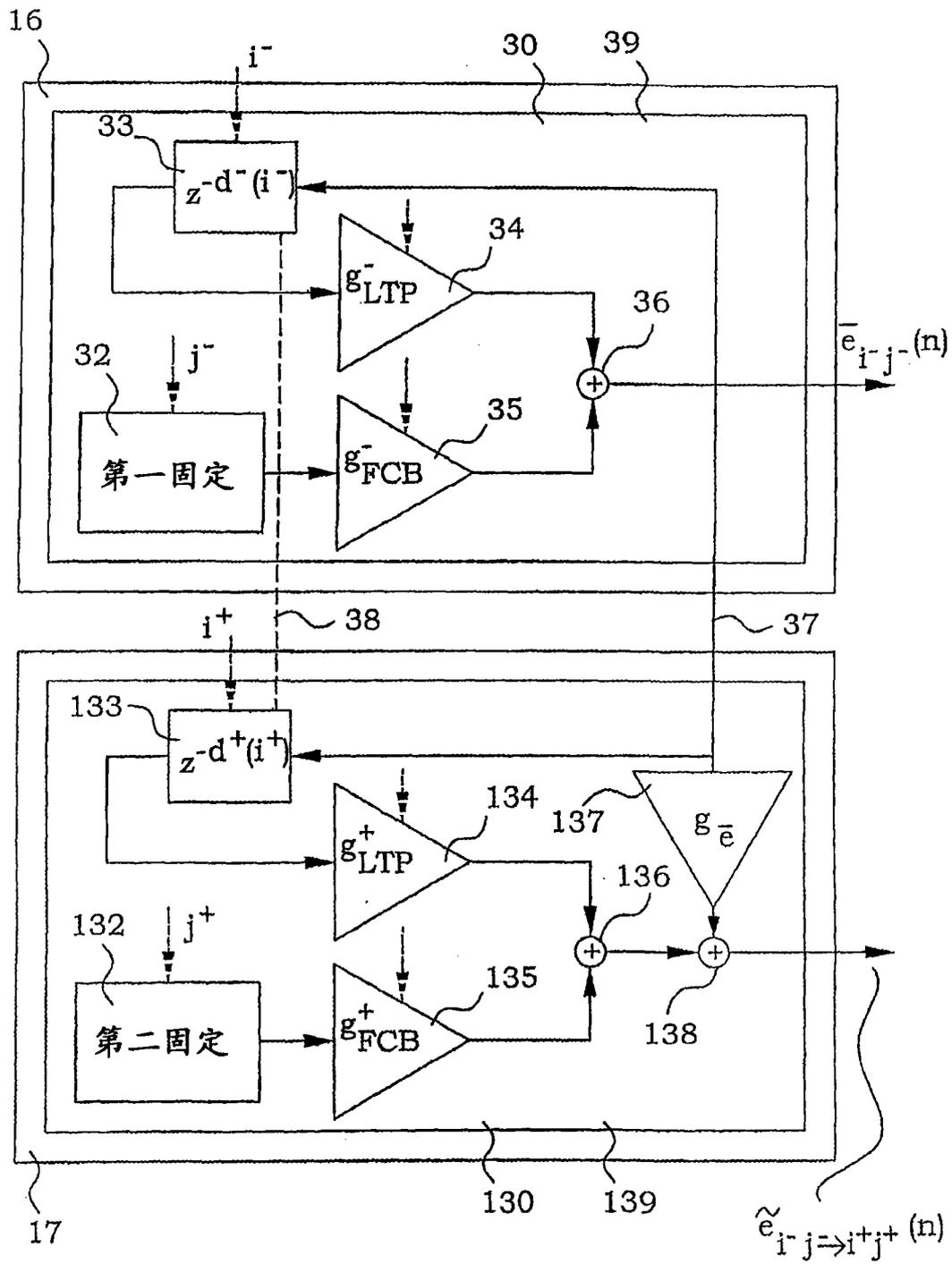


图 7

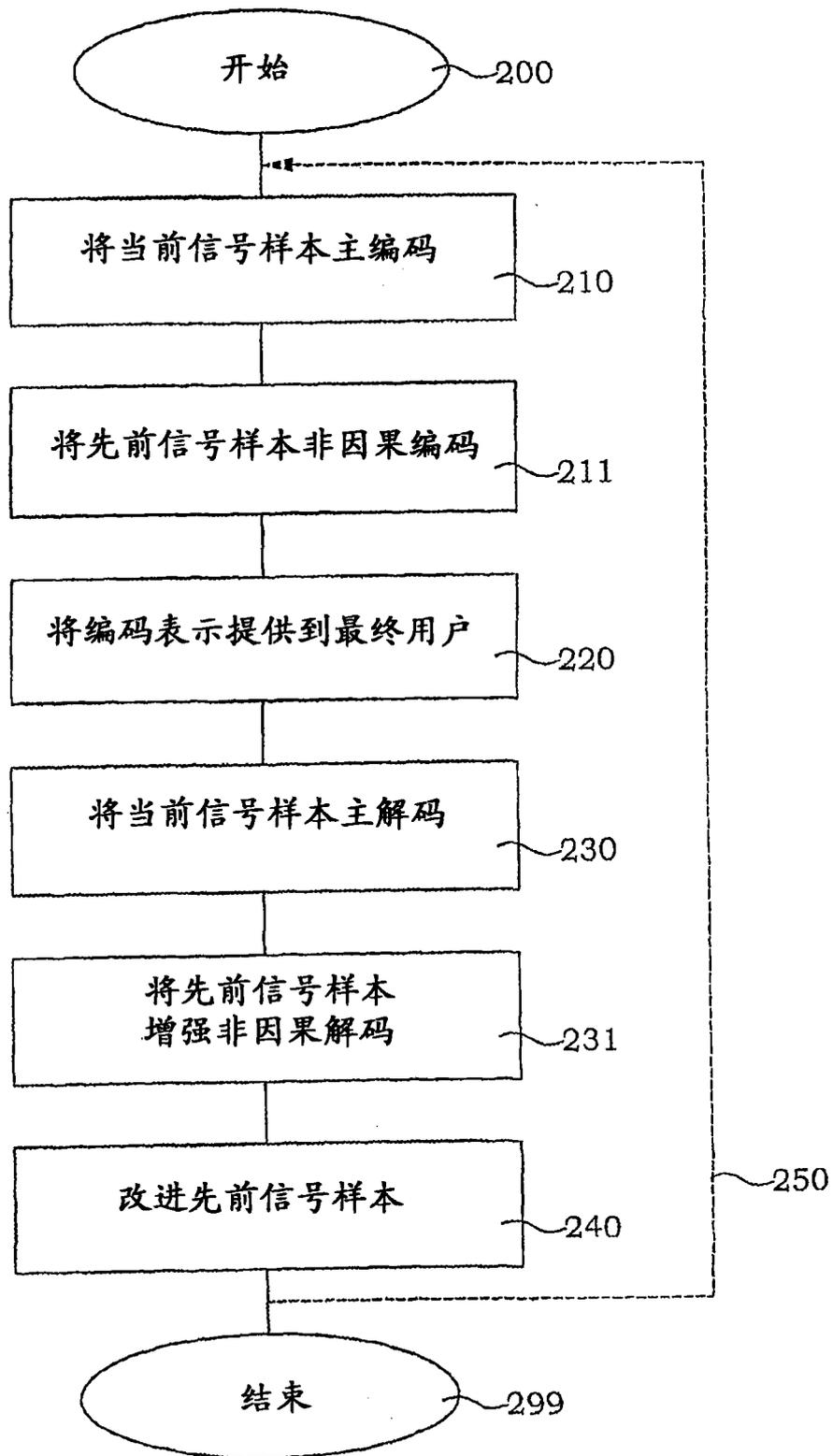


图 9