



(12) 发明专利

(10) 授权公告号 CN 103189916 B

(45) 授权公告日 2015. 11. 25

(21) 申请号 201180054354. 9

(22) 申请日 2011. 11. 07

(30) 优先权数据

10190709. 5 2010. 11. 10 EP

(85) PCT国际申请进入国家阶段日

2013. 05. 10

(86) PCT国际申请的申请数据

PCT/IB2011/054951 2011. 11. 07

(87) PCT国际申请的公布数据

W02012/063185 EN 2012. 05. 18

(73) 专利权人 皇家飞利浦电子股份有限公司

地址 荷兰艾恩德霍芬

(72) 发明人 E. F. 奇奇

(74) 专利代理机构 中国专利代理(香港)有限公

司 72001

代理人 董宁 汪扬

(51) Int. Cl.

G10L 25/90(2013. 01)

(56) 对比文件

CN 101149924 A , 2008. 03. 26,

CN 101599272 A , 2009. 12. 09,

CN 101853240 A , 2010. 10. 06,

EP 0153787 A2 , 1985. 09. 04,

EP 1944754 A1 , 2008. 07. 16,

US 6018706 A , 2000. 01. 25,

审查员 张鑫

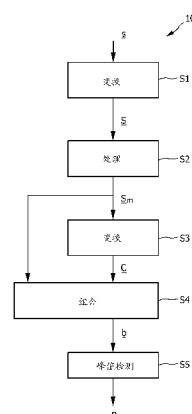
权利要求书1页 说明书9页 附图8页

(54) 发明名称

估计信号模式的方法和设备

(57) 摘要

本发明涉及一种用于估计模式,具体地说,估计具有周期分量、准周期分量或实际周期分量的信号中的基音和/或基频的方法,其中所述信号从时域变换为频域以获取所述信号的频谱,对所述频谱进行处理以获取所述信号的零相位频谱,所述信号的频谱被变换为所述时域以获取相关信号,所述频谱和所述相关信号被组合为组合频谱,以及根据所述组合频谱估计所述模式。



1. 用于估计具有周期分量、准周期分量或实际周期分量的信号($\underline{s}$)中的基音和 / 或基频的方法(10 ;30 ;50), 包括以下步骤:

- 将所述信号($\underline{s}$)从时域变换(S1 ;S8)为频域以获取信号($\underline{s}$)的频谱($\underline{S}$),
- 处理(S2 ;S9)所述频谱($\underline{S}$)以获取所述信号($\underline{s}$)的零相位频谱($\underline{S}_m$),
- 将所述信号($\underline{s}$)的所述零相位频谱($\underline{S}_m$)变换(S3 ;S12)为所述时域以获取相关信号($\underline{c}$),
- 将所述频谱($\underline{S}$)和所述相关信号($\underline{c}$)组合(S4 ;S14)为组合频谱($\underline{b}$), 以及
- 根据所述组合频谱($\underline{b}$)估计(S5 ;S15)所述基音和 / 或基频。

2. 根据权利要求1的方法, 其中所述将信号($\underline{s}$)从时域变换(S1 ;S8)为频域的步骤包括傅里叶变换(S8)。

3. 根据权利要求1或2的方法, 其中所述信号通过DC陷波滤波器(54)进行处理(S6)。

4. 根据权利要求3的方法, 其中DC滤波的信号($\underline{S}_f$)与窗函数(32)进行相乘(S7)。

5. 根据权利要求1的方法, 其中所述零相位频谱($\underline{S}_m$)是所述信号($\underline{s}$)的幅度谱($\underline{S}_m$)。

6. 根据权利要求5的方法, 其中将所述信号($\underline{s}$)的幅度谱($\underline{S}_m$)压缩(S10)为压缩幅度谱($\underline{S}_c$)。

7. 根据权利要求1的方法, 其中通过窗函数(34)对信号($\underline{s}$)的频谱($\underline{S}$)执行加窗处理(S11)。

8. 根据权利要求1的方法, 其中将所述信号($\underline{s}$)的所述零相位频谱($\underline{S}_m$)变换(S13 ;S12)为所述时域包括傅里叶反变换(S12)。

9. 根据权利要求1的方法, 其中所述相关信号($\underline{c}$)通过窗函数(36)进行衰减(S13)。

10. 根据权利要求1的方法, 其中所述组合(S4 ;S14)所述频谱($\underline{S}$)和所述相关信号($\underline{c}$)包括重新采样所述频谱($\underline{S}$)或所述相关信号($\underline{c}$)中的至少一项。

11. 根据权利要求1的方法, 其中所述估计(S5 ;S15)所述基音和 / 或基频包括搜索所述组合信号($\underline{b}$)的绝对最大值。

12. 根据权利要求1的方法, 其中所述信号具体通过全波整流函数进行整流(S16)。

13. 根据权利要求12的方法, 其中整流的信号($\underline{r}$)的零相位频谱($\underline{R}_m$)与非整流的信号($\underline{s}$)的零相位频谱($\underline{S}_m$)进行比较, 并且其中这些信号的最大值与所述相关信号($\underline{c}$)进行组合以形成所述组合信号($\underline{b}$)。

14. 用于估计具有周期分量、准周期分量或实际周期分量的信号($\underline{s}$)中的基音和 / 或基频的设备(26), 包括:

- 第一变换构件(80), 用于将信号($\underline{s}$)从时域变换为频域以获取信号($\underline{s}$)的频谱($\underline{S}$),
- 处理构件(82), 用于处理所述频谱($\underline{S}$)以获取所述信号($\underline{s}$)的零相位频谱($\underline{S}_m$),
- 第二变换构件(84), 用于将所述信号($\underline{s}$)的所述频谱($\underline{S}$)变换为所述时域以获取相关信号($\underline{c}$),
- 组合构件(86), 用于将所述频谱($\underline{S}$)和所述相关信号($\underline{c}$)组合为组合频谱($\underline{b}$), 以及
- 估计构件(88), 用于根据所述组合频谱($\underline{b}$)估计所述基音和 / 或基频。

估计信号模式的方法和设备

技术领域

[0001] 本发明涉及一种用于估计模式,更具体地说,估计具有周期分量、准周期分量或实际周期分量的信号中的基音(pitch)和 / 或基频的方法、对应的设备以及对应的计算机程序。

背景技术

[0002] 基音检测可用于诸如声音改变、文本到语音转换、语音编码、音乐信息检索、音乐演奏系统、生物识别测量、天体物理测量等之类的不同的应用。对于基音检测,时域和频域方法是众所周知的。时域方法的实现成本低且简单,其实现方式例如包括测量过零率(如 C. H. Chen 所著的“Signal Processing Handbook, Dekker, 纽约, 1988 年, 第 531 页”中描述的那样)或利用后续基音周期的相似度的自相关变差(如 R. Bracewell 所著的“The Autocorrelation Function, in The Fourier Transform and Its Applications, MacGraw-Hill, 纽约, 1965 年, 第 40-45 页”中描述的那样)。频域方法通常更为复杂并包括以下步骤:将时域信号变换为频域信号的快速傅里叶变换(FFT),仅通过考虑频率分量的功率来消除相位影响,压缩值以降低频谱包络的影响,通过基础谐波的关联(例如分谐波求和)和选择最高峰值来查找候选产生基音候选。这些方法在 D. J. Hermes 所著的“Measurement of pitch by subharmonic summation, 发表于 Journal of the Acoustic Society of America, 第 83 卷, 1988 年, 第 257-264 页”中进行描述。另一可能获取基音候选的方法是通过傅里叶反变换(IFFT)将频域信号变换回时域。例如,在“B. E. Bongart 等人所著的“The Frequency Analysis of Time Series for Echos: Cepstrum, Pseudoautocovariants, Cross-Cepstrum and Saphe Cracking, 发表于 Proceedings of the Symposium on Time Series Analysis, Wiley, 纽约, 1963 年, 第 15 章, 第 209-243 页)”中已知的基音检测算法基于频谱分析并使用对数函数进行压缩。如果将幅度(magnitude)用作压缩运算,则最终的反向变换结果为零相位信号。在此方面可使用自相关,前提是不对功率谱应用任何压缩。

[0003] 诸如对数函数之类的强压缩放大了噪声影响并形成错误的基因候选。诸如幅度运算之类的小型压缩太低,而不能抑制频谱包络的影响,因此,从较高谐波产生错误候选。一种折衷的方式是对 R. Taori 等人所著的“Harmony-1: A Versatile Low Bit Rate Speech Coding System, Nat. Lab. Technical Note 157/97”中已知的协调语音编码器中使用的幅度执行平方根运算。提供基音检测方法是為了从多个候选中判定正确的候选,但是,如果候选相互接近,则可能选择错误候选。进一步地,如果突出表示较高和 / 或较低的音高八度,则现有技术中已知的基音检测方法有可能选择错误候选。

发明内容

[0004] 本发明的目标是提供一种改进的用于估计模式,具体是指更可靠地估计信号中的基音和 / 或基频的方法、设备和计算机程序。

[0005] 在本发明的第一方面,用于估计模式,具体地说,估计具有周期分量、准周期分量或实际周期分量的信号中的基音和 / 或基频的方法包括:

[0006] - 将所述信号从时域变换为频域以获取信号的频谱,

[0007] - 处理所述频谱以获取所述信号的零相位频谱,

[0008] - 将所述信号的所述零相位频谱变换为所述时域以获取相关信号,

[0009] - 将所述频谱和所述相关信号组合为组合频谱,以及

[0010] - 根据所述组合频谱估计所述模式。

[0011] 在本发明的进一步的方面,提供对应的设备,例如包括用于执行上述方法的步骤的处理单元。

[0012] 在本发明的进一步的方面,提供对应计算机程序,其中包括当所述计算机程序在计算机上执行时,用于执行所述方法的步骤的程序代码构件。

[0013] 本发明的优选实施例在从属权利要求中定义。应该理解,所声明的设备和所声明的计算机程序具有与所声明的方法类似和 / 或相同的优选实施例并且所述优选实施例在从属权利要求中定义。

[0014] 本发明基于这样一种理念:即在附加步骤中,所述频域频谱与其时域变换相结合,以便形成的频谱在基音位置具有明显的峰值,在较高和较低的八度上具有明显衰减。此方法可用于估计信号的基音和 / 或基频。由于所形成的频谱在基音位置和 / 或所述基频上具有明显的峰值,因此可非常可靠地轻松检测基音和 / 或基频。

[0015] 根据优选实施例,将信号从时域变换为频域的步骤包括傅里叶变换,具体指快速傅里叶变换。这提供以低付出实现从时域到频域变换的可能性。

[0016] 根据进一步的实施例,通过 DC 陷波滤波器处理所述信号。所述 DC 陷波滤波器消除低频信号以防止错误检测。

[0017] 所述 DC 滤波信号优选地与窗函数相乘。该加窗运算将频谱限于包含至少两个基音周期的区域。

[0018] 根据进一步的实施例,处理所述信号的频谱以获取所述信号的幅度谱(magnitude spectrum)。所述信号幅度的计算提供了压缩运算,此运算可轻松实现并在反向变换之后形成零相位信号。

[0019] 根据进一步的实施例,将所述信号的所述频谱压缩为压缩频谱,具体指通过平方根运算。可替换地,一般而言,压缩函数可以是使用例如 0.6 作为指数的根一函数。该运算强调基音的谐波并且减弱频谱包络的影响。

[0020] 根据进一步的实施例,通过窗函数对所述信号的所述频谱进行加窗处理,具体指使用汉宁窗口的右半部分或具有类似作用的其他窗函数。该加窗运算减弱嘈杂的高频分量。

[0021] 根据进一步的实施例,所述零相位频谱,具体是所述信号的压缩幅度谱,变换为时域包括傅里叶反变换。由于所述频谱,具体是压缩频谱的相位为零,因此只需计算所述频谱实部的正轴。这提供获取在多个基音周期上具有峰值的相关信号的可能性。

[0022] 根据进一步的优选实施例,所述相关信号通过窗函数进行衰减。该加窗运算减弱所述频谱包络对所述相关信号的影响。

[0023] 根据优选实施例,所述频谱和所述相关信号的组合包括重新采样所述频谱或所述

相关信号中的至少一项。所述重新采样提供组合具有反比例轴的所述相关信号和所述频谱的可能性。具体而言,优选地使用对数尺。这提供组合在不同域的高低频分辨率方面具有巨大差别的频谱和信号的可能性。

[0024] 根据优选实施例,所述估计所述模式包括搜索所述组合信号的绝对最大值。这提供查找所述信号的所述基音和 / 或所述基频的可靠、简单的可能性。

[0025] 根据优选实施例,具体通过全波整流函数对所述信号进行整流。这提供当所述基频丢失时,在不降低非滤波信号性能的情况下判定信号的所述基音和 / 或所述基频的可能性。

[0026] 根据优选实施例,所述整流信号的所述零相位频谱与所述非整流信号的所述零相位频谱进行比较,其中选择这些信号的最大值并将其与所述相关信号进行组合以形成所述组合信号。采取最大频谱是因为在纯正弦信号的情况下,所述整流消除了所述基频并且仅产生较高谐波。为降低失真,所述整流和所述非整流的频谱通过选择这些频谱中的最大值进行组合。

附图说明

[0027] 本发明的这些和其他方面通过下面描述的(多个)实施例将变得显而易见,并参考这些实施例进行阐述。在下面的附图中:

[0028] 图 1 示出根据本发明的基音检测方法的示意性流程图,

[0029] 图 2 示出待处理的源信号以及基音检测方法根据所述源信号导出的压缩频谱、相关信号、组合频谱和测量基音的图表,

[0030] 图 3 示出根据本发明执行基音检测的设备的示意图,

[0031] 图 4 示出基音检测方法的实施例的流程图,

[0032] 图 5 示出基音检测方法的进一步实施例的流程图,

[0033] 图 6 示出执行根据图 4 的方法的处理单元的示意性框图,

[0034] 图 7 示出执行根据图 5 的方法的处理单元的示意性框图,以及

[0035] 图 8 示出执行根据图 1 的方法的处理单元的示意性框图。

具体实施方式

[0036] 图 1 示出一般由 10 表示的检测具有周期分量、准周期分量或实际周期分量的信号中的基音和 / 或基频的方法的流程图。这些信号的实例包括浊语音记录、乐器的乐音、诸如心跳之类的身体信号、来自星球的无线电信号、活动监视信号。在步骤 S1 将输入信号 s (诸如语音信号之类的准周期或实际周期信号) 从时域信号变换为频域频谱。所述变换优选地包括快速傅里叶变换(FFT)。步骤 S1 提供信号 s 的频谱 S 。频谱 S 在步骤 S2 进行处理以消除频谱的相位信息来获取零相位频谱(S_m)。所述处理包括计算频谱 S 的幅度,并且选择性地对频谱 S 进行频谱压缩,例如通过平方根运算。所述处理和压缩步骤 S2 强调基音的谐波并减弱频谱包络的影响。步骤 S2 提供零相位频谱 S_m 。

[0037] 零相位频谱 S_m 在步骤 S3 优选地使用傅里叶反变换从频域变换为时域。所述变换步骤 S3 提供相关信号 c , 该信号在多个基音周期上包括峰值。

[0038] 零相位频谱 S_m 和相关信号 c 在步骤 S4 被组合为组合频谱 b 。组合频谱 b 在基音

上包括明显的峰值,其中频谱和多个基音周期中的较高谐波被衰减,留下基音和 / 或基频作为主要峰值。所述组合 S4 通过用零相位频谱 S_m 乘以相关信号 c 来执行。

[0039] 根据组合频谱 b , 执行峰值检测 S5 以估计信号中的基音和 / 或基频。峰值检测 S5 包括搜索组合频谱 b 的最大值并提供输出信号 p , 该输出信号对应于源信号 s 中的基音和 / 或基频。

[0040] 组合零相位频谱 S_m 与其时域变换 c 的步骤 S4 导致组合频谱 b , 该频谱在基音位置和 / 或基频上具有明显的峰值, 在较高和较低的八度上具有明显衰减。因此, 峰值检测是可靠的, 因为基音位置和 / 或基频对应于组合频谱 b 中的最高峰值。

[0041] 图 2 示出五个图表图 2A-E, 这五个图表示出源信号 s 的振幅、压缩频谱 S_c 的频率、相关信号 c 的频率、组合频谱 b 的频率, 以及输出信号, 源信号 s 的基音 p 对时间。

[0042] 图 2A 所示的源信号 s 是英文语句“do they take the car when they go aboard”的时域。通过所述变换步骤 S1 和所述处理和压缩步骤 S2 从源信号 s 导出的压缩信号 S_c 在图 2B 中示出。

[0043] 通过所述变换步骤 S3 从压缩频谱 S_c 导出的相关信号 c 的频率在图 2C 中示出。

[0044] 通过步骤 S4 从压缩频谱 S_c 和相关信号 c 的组合导出的组合频谱 b 的频率在图 2D 中示出。

[0045] 通过步骤 S5 的峰值检测从组合频谱 b 导出的基音 p 对时间在图 2E 中示出。

[0046] 因此, 图 2 示出特定方法步骤 S1 至 S5 提供的信号或频谱对时间。

[0047] 图 3 示出一般由 20 表示的执行基音检测的装置的示意性框图。

[0048] 装置 20 包括信号输入端 22 和信号输出端 24, 分别用于接收源信号 s 以及提供输出信号 p 。装置 20 包括处理单元 26, 用于处理输入信号 s 以及估计输入信号 s 中的基音和 / 或基频。处理单元 26 将输出信号 p 提供给装置 20 的输出端 24。处理单元 26 包括存储器 28, 用于存储导致处理单元 26 执行方法步骤来处理输入信号 s 的程序代码。

[0049] 处理单元 26 可以通过集成电路或计算机实现, 也可以通过执行必要处理步骤的离散元件和 / 或设备实现。

[0050] 图 4 示出一般由 30 表示的基音检测方法的流程图以及特定方法步骤提供的对应信号或频谱。

[0051] 源信号 s 在第一步骤 S6 优选地通过 DC 陷波滤波器进行滤波。输入信号 s 的低频率可能导致基音检测过程失真, 这是因为在执行从时域到频域的傅里叶变换之前, 执行加窗步骤。该加窗处理步骤将主要 DC 信号的能量加到较高频率, 并且可以强调源信号 s 的弱低频。为了防止错误检测, 需要在下面的加窗处理过程之前消除源信号 s 的低频。步骤 S6 的 DC 陷波滤波器用于消除源信号 s 的低频。根据 S6 的 DC 陷波滤波器包括转换函数:

[0052]

$$H_d(z) = \frac{1 - z^{-1}}{1 - \alpha \cdot z^{-1}}$$

[0053] 其中

[0054]

$$\alpha = \frac{f_s}{f_s + f_c}$$

[0055] 并且 f_s 是采样频率, f_c 是以 Hz 为单位的截止频率, 在该截止频率上, DC 陷波滤波器的输出功率降为输入功率 (-3dB) 的 50%。

[0056] 时域中的滤波实现为:

$$s_f[n] = s[n] - s[n-1] + \alpha \cdot s_f[n-1]$$

[0058] 包括源信号 $\underline{s}$, DC 滤波信号 $\underline{s}_f$ 作为步骤 S6 的输出信号, 并且包括 n 作为第 n 个输入样本。对于语音信号, 8 kHz 的采样频率和 500 Hz 的截止频率, α 接近 0.94。DC 陷波滤波器的输出信号 $\underline{s}_f$ 不包括低频分量, 如图 4 所示。

[0059] 下面的步骤 S7 为窗函数。DC 滤波信号 $\underline{s}_f$ 与窗函数 32 相乘。窗函数 32 减弱边缘上可能的不连续并将信号限于包括至少两个基音周期的区域。例如, 如果预期最低基音为 40 Hz, 则窗口持续时间需要为至少 50 msec。优选地, 使用汉宁窗函数:

[0060]

$$w[n] = 0.5 - 0.5 \cdot \cos\left(\frac{2\pi(n+0.5)}{L}\right)$$

$$0 \leq n < L$$

[0062] 可替换地, 可使用汉宁窗函数或其他任何具有类似特征的窗函数。 L 取决于采样频率, 其中对于 8 kHz 的采样频率和 50 msec 持续时间, L 为 400。

[0063] 加窗运算定义如下:

$$s_w[n] = s_f[n] \cdot w[n], \quad 0 \leq n < L$$

[0065] 其中 s_w 是步骤 S7 的加窗函数的输出信号。

[0066] 信号 $\underline{s}_w$ 在步骤 S8 从时域变换为频域。该变换包括离散傅里叶变换 (DFT) 以提供信号 $\underline{s}_w$ 的频谱 $\underline{S}$ 。离散傅里叶变换的变换函数由以下公式给出:

$$S[k] = \sum_{n=0}^{L-1} s_w[n] \cdot e^{-j \frac{2\pi kn}{M}},$$

$$0 \leq k < M$$

[0069] 为了提高效率, 优选地使用 radix-2 FET。在这种情况下, DFT 的大小 M 的幂为 2, 并且接近但不小于 L 。例如, 对于 L 等于 400, M 被设为 512。

[0070] 在步骤 S9, 计算频谱 $\underline{S}$ 的幅度谱。由于 $\underline{s}_w$ 是实数值信号, 并且 $\underline{S}$ 在零周围对称, 因此仅使用正轴计算幅度。这样, 上述傅里叶变换的公式可写为:

[0071]

$$S_R[k] = \sum_{n=0}^{L-1} s_w[n] \cdot \cos\left(\frac{2\pi kn}{M}\right)$$

$$0 \leq k < \frac{M}{2}$$

[0073]

$$S_I[k] = \sum_{n=0}^{L-1} s_w[n] \cdot \sin\left(-\frac{2\pi kn}{M}\right)$$

[0074] 其中 $\underline{S}_R$ 是频谱的实部, $\underline{S}_I$ 是频谱的虚部。在步骤 S9, 通过以下公式计算幅度:

[0075]

$$\underline{S}_m[k] = |\underline{S}[k]| = \sqrt{(\underline{S}_R[k])^2 + (\underline{S}_I[k])^2}$$

[0076] $0 \leq k \leq \frac{M}{2}$

[0077] 其中 $\underline{S}_m$ 是步骤 S9 的输出频谱。在下面的步骤 S10, 通过平方根运算压缩幅度谱 $\underline{S}_m$ 。

[0078] $\underline{S}_c[k] = \sqrt{\underline{S}_m[k]},$

[0079] $0 \leq k \leq \frac{M}{2}$

[0080] 该平方根运算强调基音的谐波并且减弱频谱包络的影响 (例如, 语音信号中的共振峰)。所述压缩 S10 的输出信号为压缩幅度谱 $\underline{S}_c$ 。

[0081] 在步骤 S11, 在频域中对压缩幅度谱 $\underline{S}_c$ 执行加窗, 以优选地使用汉宁窗的右半部分减弱嘈杂的高频分量:

[0082] $\underline{S}_w[k] = \underline{S}_c[k] \cdot W[k],$

[0083] $0 \leq k \leq \frac{M}{2}$

[0084] 其中

[0085]

$$W[k] = \begin{cases} 0,5 + 0,5 \cdot \cos\left(\frac{\pi(k+0,5)}{N}\right), & 0 \leq k < N \\ 0, & \text{否则} \end{cases}$$

[0086] N 确定通带的大小。对于采样频率为 8 kHz 以及通带为 2 kHz 的语音信号, $N = \frac{M}{4}$

。S10 的窗函数在 34 示出。步骤 S11 的输出信号为加窗压缩幅度谱 $\underline{S}_w$, 如图 4 所示。

[0087] 加窗压缩幅度谱 $\underline{S}_w$ 在步骤 S12 使用傅里叶反变换 (IFT) 变换为时域。FFT 大小保持为上面示出的那样:

[0088] $c[n] = \sum_{k=0}^{N-1} \underline{S}_w[k] \cdot e^{j \frac{2\pi nk}{M}},$

[0089] $0 \leq n < M$

[0090] 由于加窗压缩幅度谱 $\underline{S}_w$ 的相位为零, 因此仅需频谱实部的正轴用于反变换:

[0091] $c[n] = \sum_{k=0}^{N-1} \underline{S}_w[k] \cdot \cos \frac{2\pi nk}{M},$

[0092] $0 \leq n \leq \frac{M}{2}$

[0093] 这个到时域的变换用于获取相关信号 $\underline{c}$, 该相关信号在多个基音周期上包括峰值, 如图 4 所示。

[0094] 在步骤 S13, 对相关信号 $\underline{c}$ 进行加窗以进一步减弱频谱包络的影响。优选地, 使用

简单窗函数 36 执行此衰减步骤：

$$[0095] \quad c_w[n] = c[n] \cdot n,$$

$$[0096] \quad 0 \leq n \leq \frac{M}{2}.$$

[0097] 步骤 S13 的输出信号为加窗相关信号 c_w 。

[0098] 在步骤 14, 用压缩幅度谱 S_c 乘以衰减的相关信号 c_w 形成组合频谱 b 。该组合频谱 b 在基频上具有明显的峰值。通过这些频谱相乘, 频谱中的较高谐波和多个基音周期发生衰减, 其中基频和 / 或基音保持为主要峰值。在组合频谱之前, 使用对至少一个频谱的重新采样, 因为轴成反比, 其中：

$$[0099] \quad n = \frac{M}{k}.$$

[0100] 由于不同域之间的低高频的分辨率具有差别, 因此优选地使用对数尺执行组合：

$$[0101] \quad b[i] = c_w[n_i] \cdot S_c[k_i],$$

$$[0102] \quad 0 \leq i \leq R$$

$$[0103] \quad i = R \frac{\log k - \log k_{\min}}{\log k_{\max} - \log k_{\min}},$$

$$[0104] \quad k_{\min} \leq k \leq k_{\max}$$

[0105]

$$k_i = \frac{\exp(\frac{i}{R}(\log k_{\max} - \log k_{\min}) + \log k_{\min})}{M}$$

$$[0106] \quad n_i = \frac{M}{k_i},$$

[0107] 其中 $k_{\min}$ 和 $k_{\max}$ 对应于有效基音范围。例如, 对于语音, 常见的基音范围基于 40 到 600 Hz 之间。R 确定输出阵列大小。针对 R 使用输入窗口长度便足以, 其中 $L=R$ 。

[0108] 优选地使用样条插值执行重新采样运算：

[0109]

$$\begin{aligned} c_w[n_i] = & c_w[n_i' - 1] \left(\left(\left(\frac{4}{5} - \frac{1}{3} n_i'' \right) n_i'' - \frac{7}{15} \right) n_i'' \right) + c_w[n_i'] \left(\left(\left(n_i'' - \frac{9}{5} \right) n_i'' - \frac{1}{5} \right) n_i'' + 1 \right) \\ & + c_w[n_i' + 1] \left(\left(\left(\frac{6}{5} - n_i'' \right) n_i'' + \frac{4}{5} \right) n_i'' \right) \\ & + c_w[n_i' + 2] \left(\left(\left(\frac{1}{3} n_i'' - \frac{1}{5} \right) n_i'' + \frac{2}{15} \right) n_i'' \right) \end{aligned}$$

[0110] 其中 $n_i' = [n_i]$, $n_i'' = n_i - n_i'$ 和 $[n_i]$ 表示消除分数部分的量化运算。相同的插值还应用于 S_w , 其中 k_i' 是 k_i 的量化索引。

[0111] 所述量化索引以及样条系数可以预先计算并存储在阵列中, 从而避免冗长的复杂对数和指数运算计算。在 S14 组合的重新采样频谱在图 4 中示出并由 38、40 表示。

[0112] 最后一步 S15 的峰值位置检测包括搜索组合频谱 $\underline{b}$ 的最大值：

$$[0113] \quad m_i = \max_i \{b[i]\},$$

$$[0114] \quad 0 \leq i \leq R$$

[0115]

$$p_i = i \mid b[i] = m_i$$

[0116] 其中 m_i 是最大值, p_i 是带刻度对数域中的最大值位置。以 Hz 为单位的线性域中的基音通过以下公式判定：

[0117]

$$p = f_s \frac{\exp(\frac{p_i}{R}(\log k_{\max} - \log k_{\min}) + \log k_{\min})}{M}$$

[0118] 在图 5 中, 基音检测方法的进一步的实施例一般由 50 表示。方法 50 与图 4 中的方法 30 类似。相同的步骤和信号由相同的参考标号表示, 其中仅详细解释不同之处。

[0119] 方法 50 优选地用于在基频丢失时, 查找源信号 $\underline{s}$ 的基音。在基音检测之前对信号应用高通滤波器的情况下(例如, 就像电话语音), 基频丢失。提供方法 50 是为了在不降低非滤波信号性能的情况下取回基频。

[0120] 方法 50 包括单独的路径 52 来提供 DC 滤波信号 $\underline{s}_f$ 的整流频谱。

[0121] DC 滤波信号 $\underline{s}_f$ 在步骤 S16 进行整流以提供整流信号 $\underline{r}$ 。优选地, DC 滤波信号 $\underline{s}_f$ 通过全波整流器进行全波整流。全波整流器的公式如下所示：

$$[0122] \quad r[n] = |s_f[n]|.$$

[0123] 在整流步骤 S16 之后, 执行步骤 S6' 至 S10' 以提供整流信号的整流压缩幅度谱 $\underline{R}_c$ 。步骤 S6' 至 S10' 与上述步骤 S6 至 S10 相同。在步骤 S17, 组合非整流信号 $\underline{s}$ 的压缩幅度谱 $\underline{S}_c$ 和整流压缩幅度谱 $\underline{R}_c$ 。为降低失真以及对于整流消除基频并只产生较高谐波的情况, 组合整流信号 $\underline{r}$ 的整流压缩幅度谱 $\underline{R}_c$ 和非整流信号 $\underline{s}$, 其中这些谱的最大值根据以下的公式选择：

$$[0124] \quad R'_c[k] = \max\{dR_c[k], S_c[k]\},$$

$$[0125] \quad 0 \leq k \leq \frac{M}{2}$$

[0126] 其中, d 是缩放系数, 优选地被设为 2。S17 的输出信号为 $\underline{R}'_c$, 它是整流信号和非整流信号的压缩幅度谱的最大值。

[0127] S17 的输出信号与衰减的相关信号 $\underline{c}_w$ 在步骤 S14 进行组合, 如上所述。

[0128] 图 6 示出图 3 所示的处理单元 26 的实施例的示意性框图。根据图 6 的处理单元 26 包括特定离散元件或设备, 提供这些元件或设备是为了执行根据图 4 的方法步骤。

[0129] 输入端 22 与执行步骤 S6 的 DC 陷波滤波器 54 相连。DC 陷波滤波器 54 与执行步骤 S7 的加窗元件 56 相连。加窗元件 56 与执行步骤 S8 的傅里叶变换元件 58 相连。傅里叶变换元件 58 与被提供用于根据步骤 S9 计算幅度的幅度元件 60 相连。幅度元件 60 与执

行步骤 S10 的方根运算元件 62 相连。方根运算元件 62 与被提供用于执行步骤 S11 的加窗元件 64 相连。加窗元件 64 与被提供用于执行步骤 S12 的傅里叶反变换元件 66 相连。傅里叶反变换元件与被提供用于执行 S13 的加窗元件 68 相连。加窗元件 68 与被提供用于执行 S14 的组合元件 70 相连。方根运算元件 62 还与组合元件 70 相连以将压缩幅度谱 $\underline{s}$ 提供给组合元件 70。组合元件 70 与被提供用于执行步骤 S15 的峰值位置检测元件 72 相连。峰值位置检测元件 72 与处理单元 26 的输出端相连以将基音 p 提供给输出端 24。

[0130] 图 7 示出图 6 所示的处理单元 26 的实施例的示意性框图。参考图 6, 其中相同的步骤、元件和信号由相同的参考标号表示, 并且仅详细解释不同之处。根据图 7 的处理单元 26 包括特定离散元件或设备, 提供这些元件或设备是为了执行根据图 5 的方法步骤。

[0131] 根据该实施例, 图 7 的处理单元 26 包括附加并行路径 74 以提供源信号 $\underline{s}$ 的整流压缩幅度谱。路径 74 执行图 5 所示的路径 52 的步骤。路径 74 包括整流器 76, 该整流器与 DC 陷波滤波器 54 相连以执行步骤 S16。整流器 76 与串联的元件 54'、56'、58'、60' 和 62' 相连, 这些元件分别与元件 54、56、58、60 和 62 相同, 用于执行步骤 S6'、S7'、S8'、S9' 和 S10'。

[0132] 方根运算元件 62 和 62' 与执行步骤 S17 的最大值判定元件 78 相连。最大值判定元件 78 与执行步骤 S14 的组合元件 70 相连。

[0133] 图 8 示出图 3 所示的用于执行根据图 1 的方法的处理单元 26 的实施例的示意性框图。一般而言, 处理单元 26 也称为“设备”或“系统”。

[0134] 处理单元 26 包括执行步骤 S1 的第一变换单元 80、执行步骤 S2 的处理单元 82、执行步骤 S3 的第二变换单元 84、执行步骤 S4 的组合单元 86 以及执行步骤 S5 的估计单元 88。

[0135] 因此, 方法 10、30 和 50 的步骤可以通过上述处理单元 26 中的离散元件执行。在可替换实施例中, 方法 10、30 和 50 的步骤可由处理单元 26 执行, 所述处理单元可通过集成电路(例如 FPGA 或 ASIC 或类似的电路)实现, 也可通过计算机或控制单元上运行的软件实现。

[0136] 尽管在附图和上述描述中详细示出和描述了本发明, 但是此类图示和描述只能视为示意性的或示例性的, 并非旨在进行限制; 本发明不限于所公开的实施例。所公开实施例的其他变形可被实施本发明的所属领域技术人员通过对附图, 说明书和所附权利要求的研究而理解和实现。

[0137] 在权利要求中, 单词“包括”不排除其他元素或步骤, 不定冠词“一”或“一个”不排除多个。单个元件或其他单元可以实现权利要求中所述的多个项的功能。某些措施被记载在相互不同的从属权利要求中并不表示这些措施的组合不能被有利地使用。

[0138] 计算机程序可以存储/分布在适当的介质上, 例如光存储介质或与其他硬件一起提供或作为其他硬件的一部分提供的固态介质, 但是也可以通过其他形式分布, 例如通过因特网或其他有线或无线电信系统。

[0139] 权利要求中的任何参考标号都不应被视为对范围的限制。

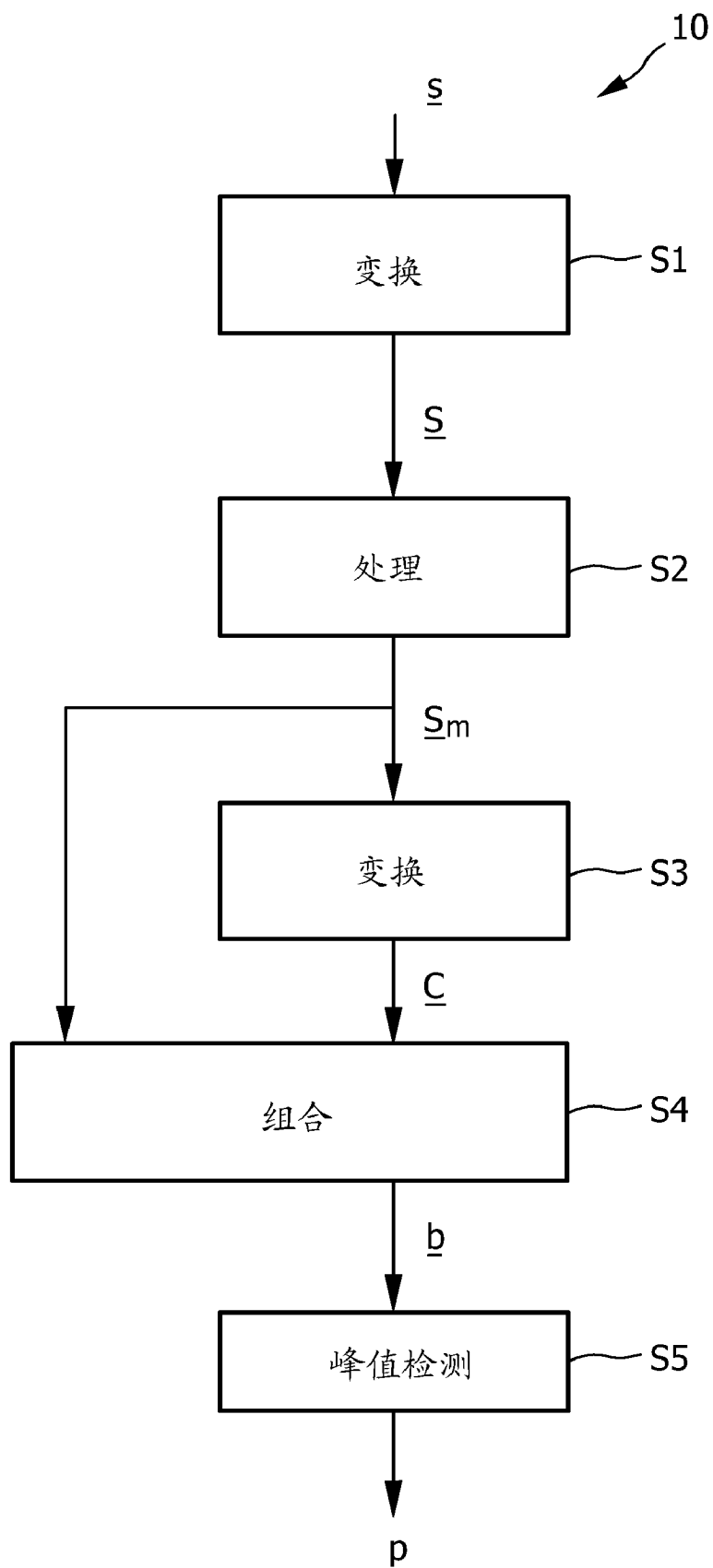


图 1

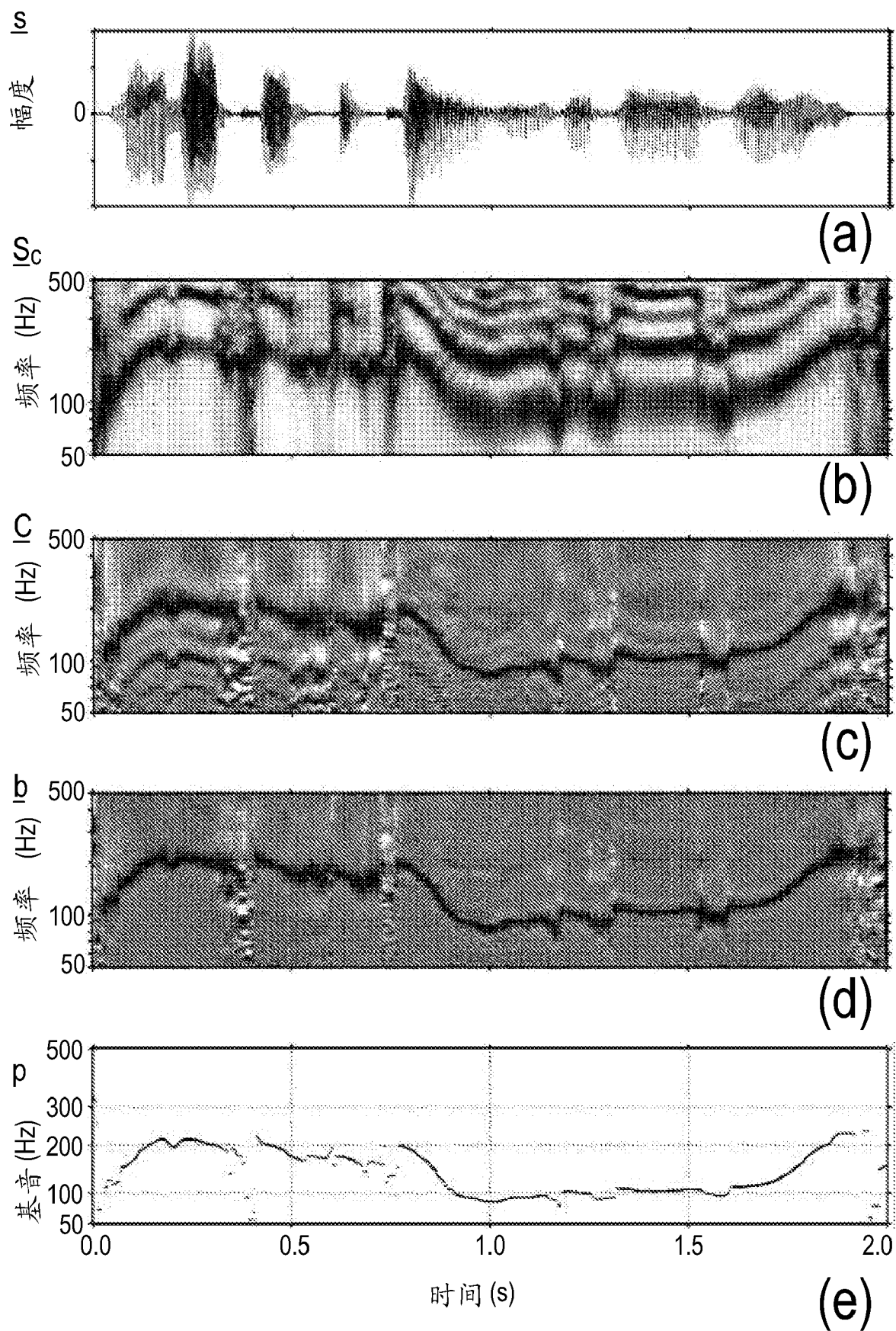


图 2

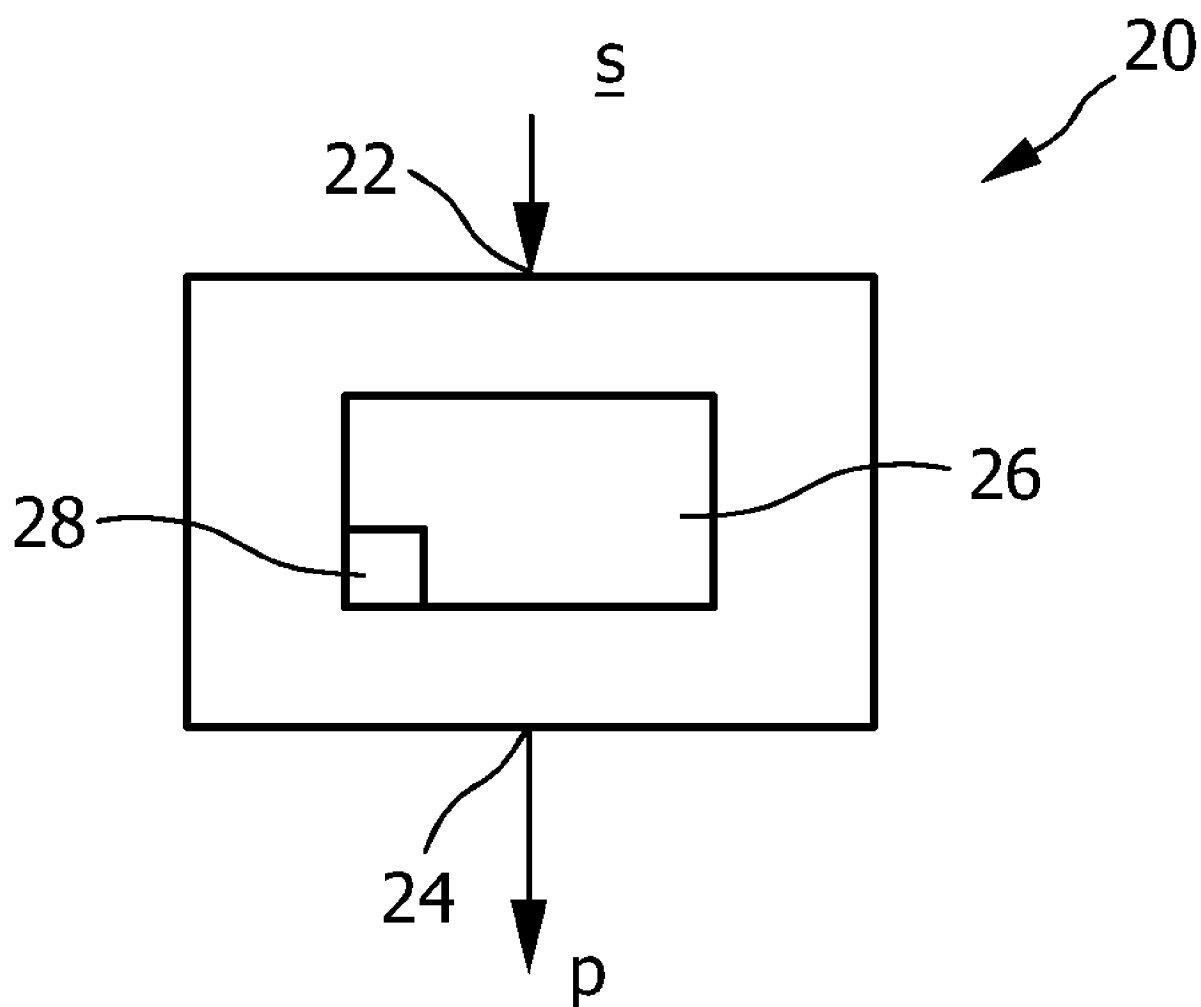


图 3

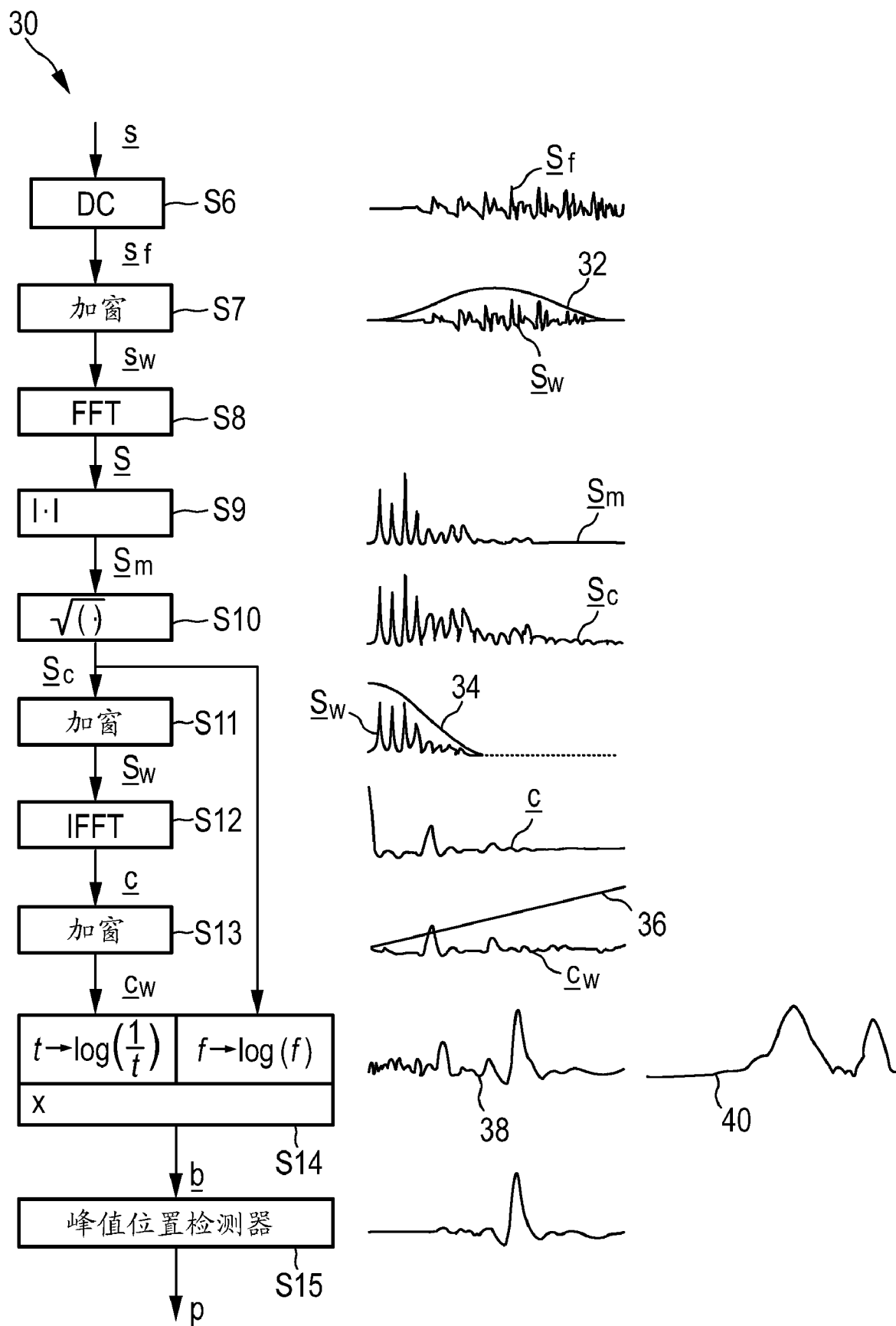


图 4

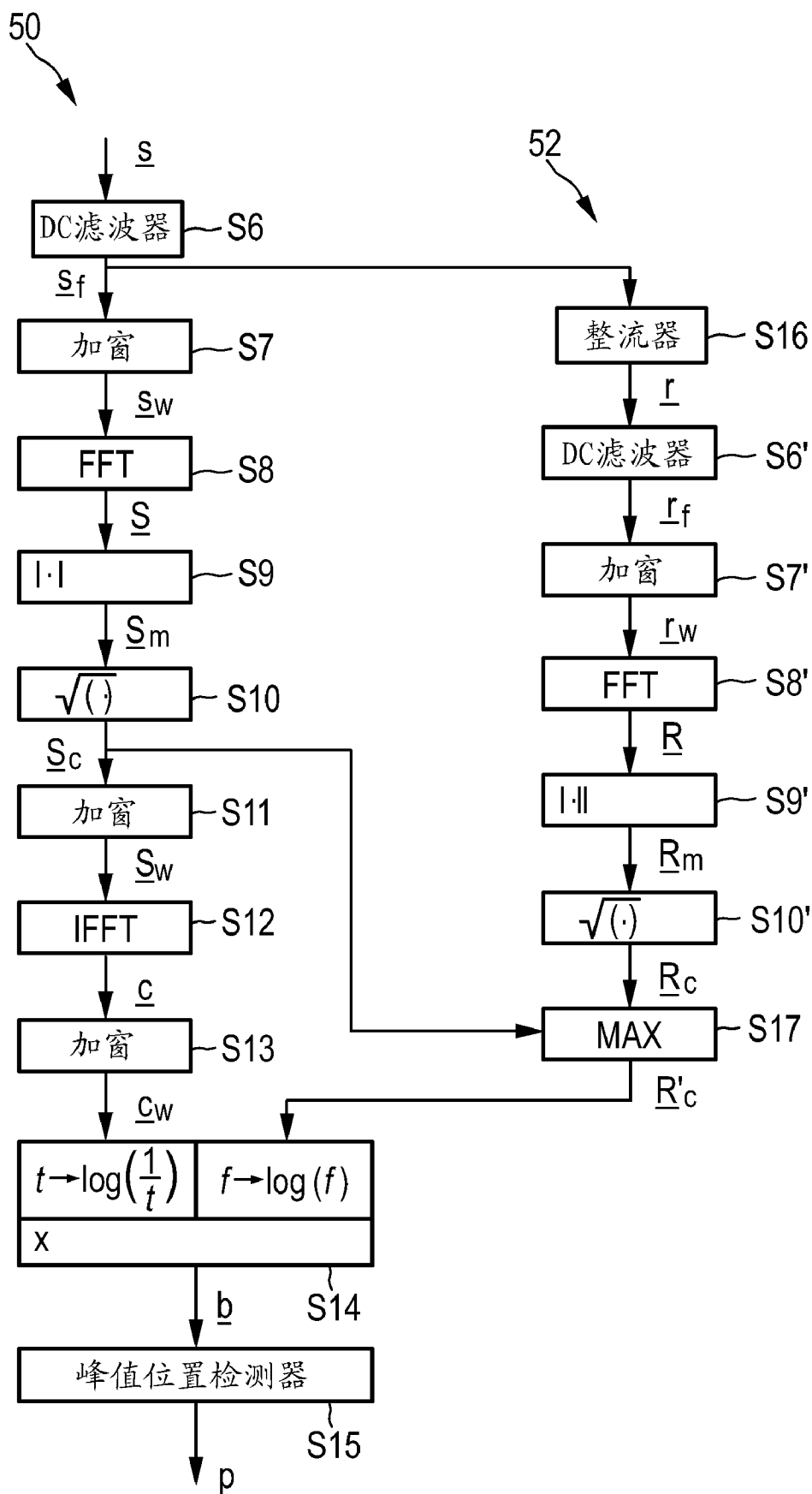


图 5

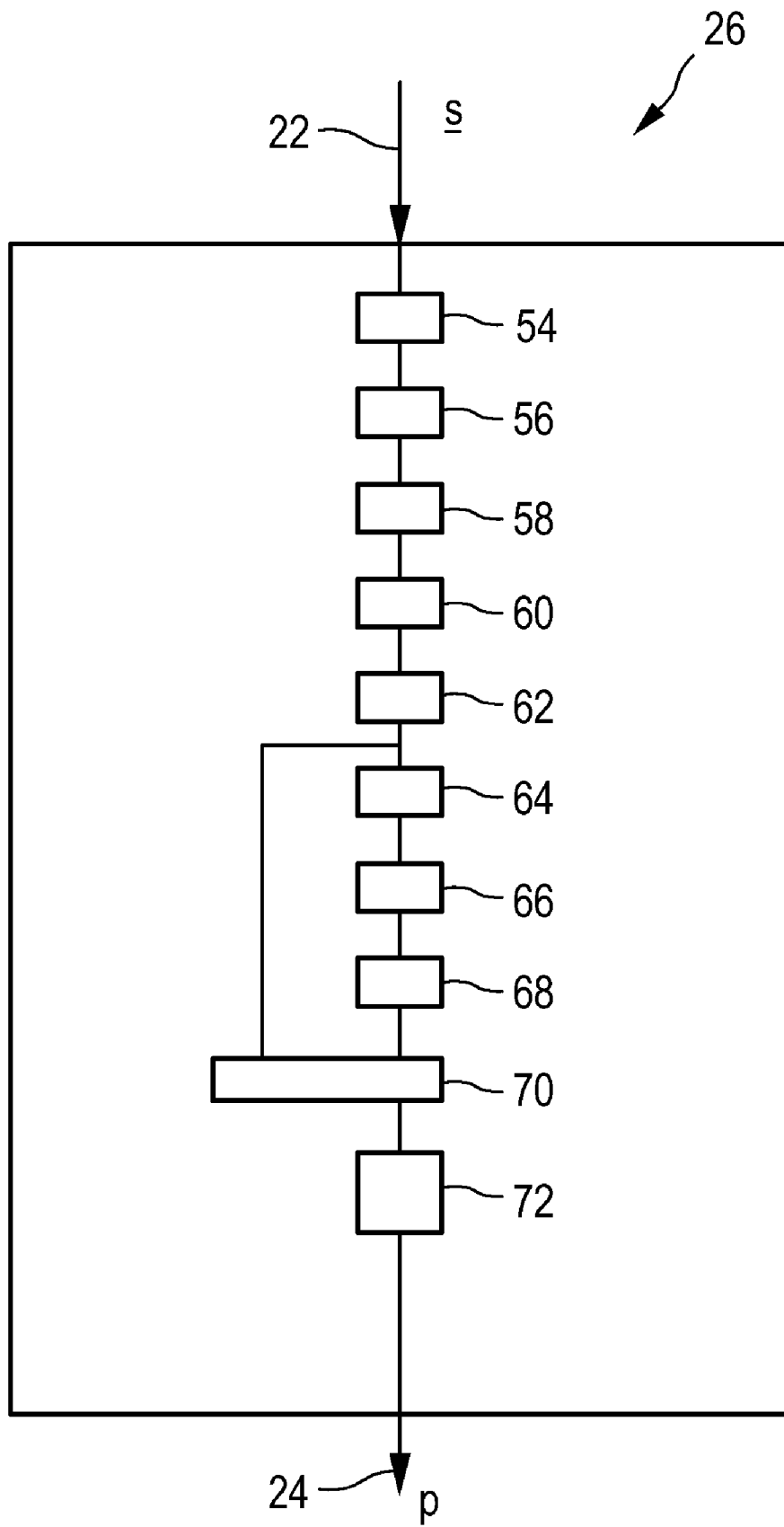


图 6

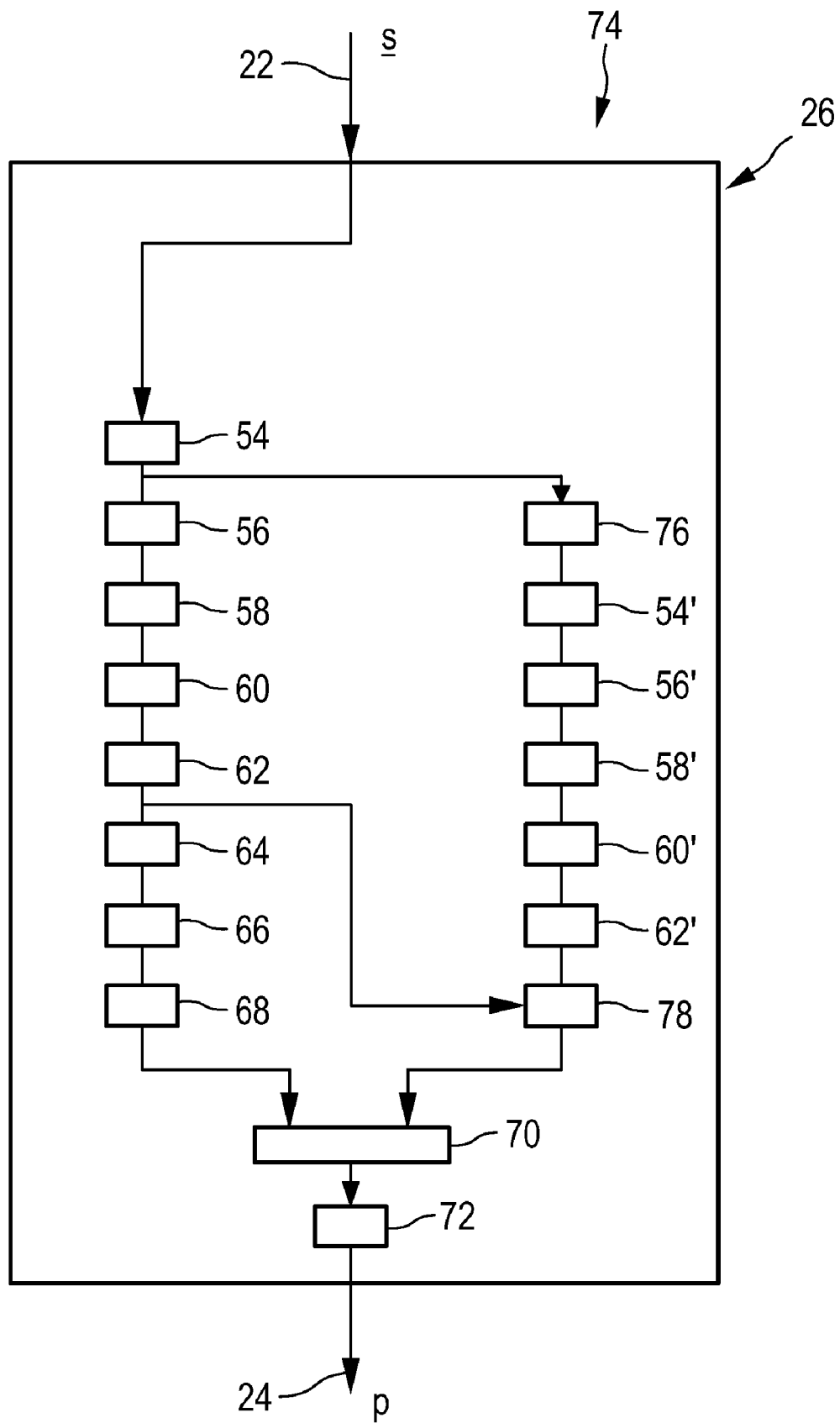


图 7

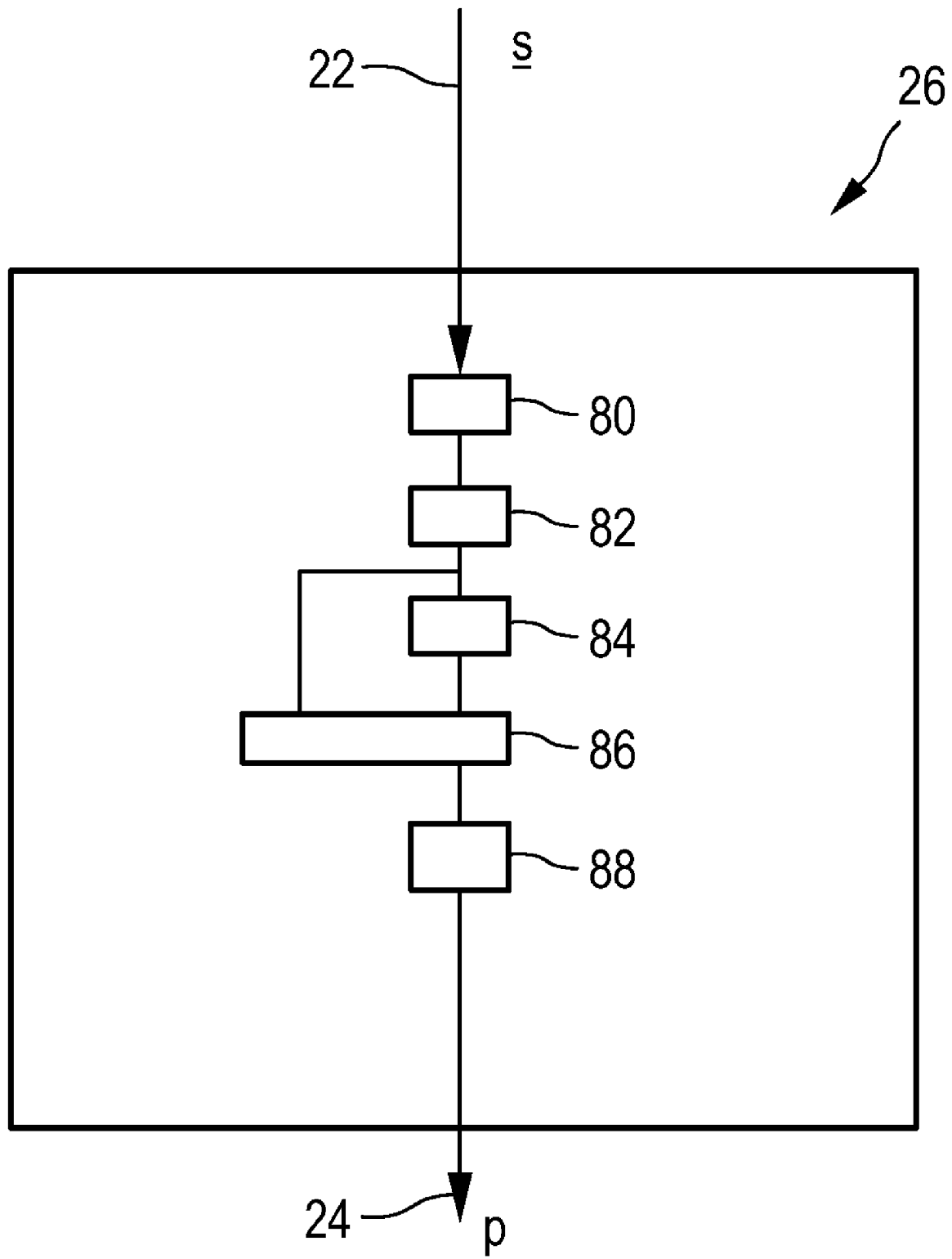


图 8