



(12)发明专利

(10)授权公告号 CN 104246873 B

(45)授权公告日 2017.02.01

(21)申请号 201280069724.0

(22)申请日 2012.02.17

(65)同一申请的已公布的文献号
申请公布号 CN 104246873 A

(43)申请公布日 2014.12.24

(85)PCT国际申请进入国家阶段日
2014.08.15

(86)PCT国际申请的申请数据
PCT/EP2012/052734 2012.02.17

(87)PCT国际申请的公布数据
W02013/120531 EN 2013.08.22

(73)专利权人 华为技术有限公司

地址 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 郎玥 大卫·维雷特 许剑峰

(51)Int.Cl.
G10L 19/008(2013.01)
H04S 3/00(2006.01)

(56)对比文件
CN 101578658 A,2009.11.11,
CN 101460997 A,2009.06.17,
WO 2011/072729 A1,2011.06.23, (续)

审查员 陈成

权利要求书2页 说明书18页 附图4页

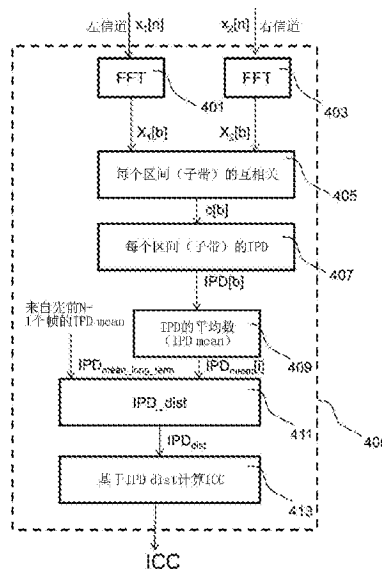
(54)发明名称

用于编码多声道音频信号的参数编码器

(57)摘要

本发明涉及一种用于产生多声道音频信号的多个音频声道信号(X₁[b]、X₂[b])中的音频声道信号(X₁[b])的编码参数(ICC)的参数音频编码器(100),每个音频声道信号(X₁[b]、X₂[b])具有音频声道信号值(X₁[k]、X₂[k]),所述参数音频编码器(100)包括参数产生器(105),所述参数产生器(105)用于-根据所述音频声道信号(X₁[b])的所述音频声道信号值(X₁[k])以及参考音频信号(X₂[b])的参考音频信号值(X₂[k])为所述多个音频声道信号中的所述音频声道信号(X₁[b])确定第一组编码参数(IPD[b]),其中所述参考音频信号是所述多个音频声道信号中的另一音频声道信号(X₂[b])或从所述多个多声道音频信号的至少两个音频声道信号中获得的下混合音频信号,-基于所述音频声道信号(X₁[b])的所述第一组编码参数(IPD[b])为所述音频声道信号(X₁[b])确定第一编码参数平均值(IPD_{mean}[i]),-基于所述音频声道信号(X₁[b])的所述第一编码参数平均值(IPD_{mean}[i])以及所述音频声道信号(X₁[b])的至少一个另一第一编码参数平均值(IPD_{mean}[i-1])为所述音频声道信号(X₁

[b])确定第二编码参数平均值(IPD_{mean_long_term}),以及-基于所述音频声道信号(X₁[b])的所述第一编码参数平均值(IPD_{mean}[i])以及所述音频声道信号(X₁[b])的所述第二编码参数平均值(IPD_{mean_long_term})确定所述编码参数(ICC)。



CN 104246873 B

[接上页]

(56)对比文件

WO 2006/000952 A1,2006.01.05,

EP 1565036 A2,2005.08.17,

1. 一种参数音频编码器(100),用于产生多声道音频信号的多个音频声道信号 $X_1[b]$ 和 $X_2[b]$ 中的音频声道信号 $X_1[b]$ 的编码参数ICC的参数音频编码器(100),每个音频声道信号 $X_1[b]$ 、 $X_2[b]$ 具有音频声道信号值 $X_1[k]$ 和 $X_2[k]$,所述参数音频编码器(100)包括参数产生器(105),所述参数产生器(105)用于

根据所述音频声道信号 $X_1[b]$ 的所述音频声道信号值 $X_1[k]$ 以及参考音频信号的参考音频信号值为所述多个音频声道信号中的所述音频声道信号 $X_1[b]$ 确定第一组编码参数IPD[b],其中所述参考音频信号是所述多个音频声道信号中的另一音频声道信号 $X_2[b]$ 或从所述多个多声道音频信号的至少两个音频声道信号中获得的下混合音频信号,

基于所述音频声道信号 $X_1[b]$ 的所述第一组编码参数IPD[b]为所述音频声道信号 $X_1[b]$ 确定第一编码参数平均值 $IPD_{mean}[i]$,

基于所述音频声道信号 $X_1[b]$ 的所述第一编码参数平均值 $IPD_{mean}[i]$ 以及所述音频声道信号 $X_1[b]$ 的至少一个另一第一编码参数平均值 $IPD_{mean}[i-1]$ 为所述音频声道信号 $X_1[b]$ 确定第二编码参数平均值 $IPD_{mean_long_term}$,以及

基于所述音频声道信号 $X_1[b]$ 的所述第一编码参数平均值 $IPD_{mean}[i]$ 以及所述音频声道信号 $X_1[b]$ 的所述第二编码参数平均值 $IPD_{mean_long_term}$ 确定所述编码参数ICC。

2. 根据权利要求1所述的参数音频编码器(100),其中所述第一组编码参数IPD[b]是以下参数中的一种:

- 声道间电平差,
- 声道间相位差,
- 声道间相干性,
- 声道间强度差,
- 子带声道间电平差,
- 子带声道间相位差,
- 子带声道间相干性,以及
- 子带声道间强度差。

3. 根据权利要求1或2所述的参数音频编码器(100),其中所述参数产生器(105)用于确定随后的音频声道信号值 $X_1[k]$ 的相位差以获得所述第一组编码参数IPD[b]。

4. 根据权利要求1或2所述的参数音频编码器(100),其中所述音频声道信号 $X_1[b]$ 以及所述参考音频信号是频域信号,并且其中所述音频声道信号值 $X_1[k]$ 以及所述参考音频信号值 $X_2[k]$ 与频率区间k或频率子带b相关联。

5. 根据权利要求1或2所述的参数音频编码器(100),进一步包括变换器FFT,用于变换频域中的多个时域音频声道信号 $x_1[n]$ 和 $x_2[n]$ 以获得所述多个音频声道信号 $X_1[b]$ 和 $X_2[b]$ 。

6. 根据权利要求1或2所述的参数音频编码器(100),其中所述参数产生器(105)用于确定所述音频声道信号 $X_1[b]$ 和 $X_2[b]$ 的每个频率区间[k]或每个频率子带[b]的所述第一组编码参数IPD[b]。

7. 根据权利要求1或2所述的参数音频编码器(100),其中所述参数产生器(105)用于确定所述音频声道信号 $X_1[b]$ 的所述第一编码参数平均值 $IPD_{mean}[i]$ 作为频率区间[k]或频率子带[b]上的所述音频声道信号 $X_1[b]$ 的所述第一组编码参数IPD[b]的平均值。

8. 根据权利要求1或2所述的参数音频编码器(100),其中所述参数产生器(105)用于确定所述音频声道信号 $X_1[b]$ 的所述第二编码参数平均值 $IPD_{mean_long_term}$ 作为所述音频声道信号 $X_1[b]$ 的多个帧上的多个第一编码参数平均值 $IPD_{mean}[i]$ 的平均值,其中每个第一编码参数平均值 $IPD_{mean}[i]$ 与所述多声道音频信号的帧 i 相关联。

9. 根据权利要求1或2所述的参数音频编码器(100),其中所述参数产生器(105)用于确定所述第二编码参数平均值 $IPD_{mean_long_term}$ 与所述第一编码参数平均值 $IPD_{mean}[i]$ 之间的差的绝对值 IPD_{dist} ;根据所确定的绝对值 IPD_{dist} 确定所述编码参数ICC。

10. 根据权利要求9所述的参数音频编码器(100),其中所述参数产生器(105)用于根据第一参数值 d 与乘以第二参数值 e 的所述所确定的绝对值 IPD_{dist} 之间的差来确定所述编码参数ICC。

11. 根据权利要求10所述的参数音频编码器(100),其中所述参数产生器(105)用于将所述第一参数值 d 设置成1并且将所述第二参数值 e 设置成1。

12. 根据权利要求1或2所述的参数音频编码器(100),进一步包括:下混合信号产生器,用于叠加所述多声道音频信号的所述音频声道信号中的至少两者,以获得下混合信号;音频编码器,具体而言单声道编码器,用于编码所述下混合信号以获得经编码的音频信号;以及组合器,用于将所述经编码的音频信号与对应的编码参数组合。

13. 一种用于产生多声道音频信号的多个音频声道信号 $X_1[b]$ 和 $X_2[b]$ 中的音频声道信号 $X_1[b]$ 的编码参数ICC的方法(400),每个音频声道信号 $X_1[b]$ 和 $X_2[b]$ 具有音频声道信号值 $X_1[k]$ 和 $X_2[k]$,所述方法(400)包括:

根据所述音频声道信号 $X_1[b]$ 的所述音频声道信号值 $X_1[k]$ 以及参考音频信号的参考音频信号值为所述多个音频声道信号中的所述音频声道信号 $X_1[b]$ 确定(407)第一组编码参数 $IPD[b]$,其中所述参考音频信号是所述多个音频声道信号中的另一音频声道信号 $X_2[b]$ 或从所述多个多声道音频信号的至少两个音频声道信号中获得的下混合音频信号,

基于所述音频声道信号 $X_1[b]$ 的所述第一组编码参数 $IPD[b]$ 为所述音频声道信号 $X_1[b]$ 确定(409)第一编码参数平均值 $IPD_{mean}[i]$,

基于所述音频声道信号 $X_1[b]$ 的所述第一编码参数平均值 $IPD_{mean}[i]$ 以及所述音频声道信号 $X_1[b]$ 的至少一个另一第一编码参数平均值 $IPD_{mean}[i-1]$ 为所述音频声道信号 $X_1[b]$ 确定(411)第二编码参数平均值 $IPD_{mean_long_term}$,以及

基于所述音频声道信号 $X_1[b]$ 的所述第一编码参数平均值 $IPD_{mean}[i]$ 以及所述音频声道信号 $X_1[b]$ 的所述第二编码参数平均值 $IPD_{mean_long_term}$ 确定(413)所述编码参数ICC。

用于编码多声道音频信号的参数编码器

技术领域

[0001] 本发明涉及音频编码。

背景技术

[0002] 例如,在电气与电子工程师协会的信号处理对音频和声学的应用的研究组会刊(Proc.IEEE Workshop on Appl.of Sig.Proc.to Audio and Acoust)中的C.法勒(C.Faller)和F.鲍姆加特(F.Baumgarte)的“使用感知参数化的空间音频的有效表示(Efficient representation of spatial audio using perceptual parametrization)”(2001年10月,第199至202页)中描述的参数立体声或多声道音频编码使用空间线索,从下混合(通常是单声道或立体声)音频信号来合成多声道音频信号,所述多声道音频信号比下混合音频信号具有更多的声道。通常,下混合音频信号由多声道音频信号(例如,立体声音频信号)的多个音频声道信号的叠加而产生。这些较少声道是经波形编码的,且将与原始信号声道关系有关的辅助信息(即,空间线索)作为编码参数添加到经编码的音频声道。解码器使用此辅助信息,基于经解码的经波形编码的音频声道而重新产生原始数目个音频声道。

[0003] 基本参数立体声编码器可以使用声道间电平差(ILD:inter-channel level difference)作为从单声道下混合音频信号产生立体声信号所需的线索。较复杂的编码器还可以使用声道间相干性(ICC:inter-channel coherence),其可以表示音频声道信号(即,音频声道)之间的相似度。此外,当编码双耳立体声信号(例如)以实现3D音频或基于头戴式耳机的环绕渲染(surround rendering)时,声道间相位差(IPD:inter-channel phase difference)也可以起到再现声道之间的相位/延迟差异的作用。

[0004] ICC线索的合成可以与大部分音频和音乐内容相关,以重新产生环境、立体声混响、声源宽度以及与以下资料中描述的空间印象有关的其他感知:J.布劳尔特(J.Blauert)的“空间听觉:人类听声辨位的心理物理学(The Psychophysics of Human Sound Localization)”,美国马萨诸塞州剑桥的麻省理工学院出版社,1997年。相干性合成可以通过使用以下资料中描述的频域中的去相关器来实施:2003年3月音频工程协会会议的第114次预印本中的E.斯凯捷(E.Schuijers)、W.欧门(W.Oomen)、B.德恩布林克尔(B.den Brinker)以及J.布雷巴特(J.Breebaart)的“高质量音频的参数编码方面的进步(Advances in parametric coding for high-quality audio)”。然而,用于估计空间线索以及合成多声道音频信号的已知合成方法的复杂度可能会增加。此外,例如,除了其他参数(例如,声道间电平差(ICLD:inter-channel level difference)和声道间相位差(ICPD:inter-channel phase difference))之外还使用ICC参数可能增加比特率开销。

发明内容

[0005] 本发明的目标是提供一种用于对表示多声道音频信号的声道之间的声道间关系的编码参数进行估计以便于有效的音频信号编码的概念。

[0006] 此目标通过独立的权利要求的特征实现。从附属权利要求、描述和图式中容易明白其他实施。

[0007] 为了详细描述本发明,将使用以下术语、缩写和符号:

[0008] BCC(Binaural cues coding):双耳线索编码,即,使用下混合和双耳线索(或空间参数)以描述声道间关系的立体声或多声道信号的编码。

[0009] 双耳线索(Binaural Cue):左耳进入信号与右耳进入信号之间的声道间线索(还参看ITD、ILD和IC)。

[0010] CLD(Channel level difference):声道电平差,与ICLD相同。

[0011] FFT(Fast Fourier Transform):DFT的快速实施方式,表示快速傅里叶变换。

[0012] STFT(Short-time Fourier transform):短时傅里叶变换。

[0013] HRTF(Head-related transfer function):头相关传递函数,即,在自由场中从源到左耳和右耳入口的声音的建模转导。

[0014] IC(Inter-aural coherence):耳间相干性,即,左耳进入信号与右耳进入信号之间的相似度。这有时也称为IAC或耳间互相关(IACC)。

[0015] ICC(Inter-channel coherence):声道间相干性,即,声道间相关性。

[0016] ICPD(Inter-channel phase difference):声道间相位差。信号对之间的平均相位差。

[0017] ICLD(Inter-channel level difference):声道间电平差。

[0018] ICTD(Inter-channel time difference):声道间时间差。

[0019] ILD(Interaural level difference):耳间电平差,即,左耳进入信号与右耳进入信号之间的电平差。这有时还称为耳间强度差(IID)。

[0020] IPD(Interaural phase difference):耳间相位差,即,左耳进入信号与右耳进入信号之间的相位差。

[0021] ITD(Interaural time difference):耳间时间差,即,左耳进入信号与右耳进入信号之间的时间差。这有时还称为耳间时间延迟。

[0022] 混合(Mixing):假设多个源信号(例如,单独录音的仪器、多声道录音),产生用于空间音频播放的立体声或多声道音频信号的过程表示混合。

[0023] 空间音频(Spatial audio):当通过适当的播放系统播放时引起听觉空间图像的音频信号。

[0024] 空间线索(Spatial cue):与空间感知相关的线索。此术语用于一对立体声或多声道音频信号的声道之间的线索(还参看ICTD、ICLD和ICC),还表示为空间参数或双耳线索。

[0025] 根据第一方面,本发明涉及一种用于产生多声道音频信号的多个音频声道信号中的音频声道信号的编码参数的参数音频编码器,每个音频声道信号具有音频声道信号值,所述参数音频编码器包括参数产生器,所述参数产生器用于

[0026] -根据音频声道信号的音频声道信号值以及参考音频信号的参考音频信号值为多个音频声道信号中的音频声道信号确定第一组编码参数,其中参考音频信号是多个音频声道信号中的另一音频声道信号,

[0027] -基于音频声道信号的第一组编码参数为音频声道信号确定第一编码参数平均值,

[0028] -基于音频声道信号的第一编码参数平均值以及音频声道信号的至少一个另一第一编码参数平均值为音频声道信号确定第二编码参数平均值,以及

[0029] -基于音频声道信号的第一编码参数平均值以及音频声道信号的第二编码参数平均值确定编码参数。

[0030] 参考音频信号可以是多声道音频信号的音频声道信号中的一者。具体而言,参考音频信号可以是形成两个声道的多声道信号的实施例的立体声信号的左音频声道信号或右音频声道信号。然而,参考音频信号可以是形成用于确定编码参数的参考的任何信号。此种参考信号可以在下混合多声道音频信号的声道之后由单声道下混合音频信号形成,或在下混合多声道音频信号的声道之后由下混合音频信号的声道中的一者形成。

[0031] 参数音频编码器可能具有较低的复杂度,因为该参数音频编码器不需要相干性或相关性计算。当ICC是利用仅需要几个步骤的粗糙量化器进行量化时,所述参数音频编码器甚至提供音频声道之间的关系精确估计。尤其针对音乐信号,还针对语言信号,使用对音频信号进行编码的编码参数是非常重要的,因为在具有正确的声音场景宽度的情况下,输出的音乐听起来更自然且不“干燥”。对于极低位率的参数立体声音频编码方案,比特预算是受限的且仅传输一个全带ICC,编码参数能够表示声道之间的全局相关性。

[0032] 在根据第一方面的参数音频编码器的第一可能实施形式中,第一组编码参数是以下参数中的一者:声道间电平差、声道间相位差、声道间相干性、声道间强度差、子带声道间电平差、子带声道间相位差、子带声道间相干性以及子带声道间强度差。

[0033] 此些参数表示音频信号之间的相似度并且因此可以由编码器使用,以便减少待传输的信息并且因此降低计算复杂度。

[0034] 在根据第一方面或根据第一方面的第一实施形式的参数音频编码器的第二可能实施形式中,参数产生器用于确定随后的音频声道信号值的相位差以获得第一组编码参数。

[0035] 需要随后的音频声道信号的相位差以用于再现声道之间的相位差和/或延迟差。当再现相位差时,语言和音乐听起来会更自然。

[0036] 在根据第一方面或根据第一方面的前述实施形式中任一者的参数音频编码器的第三可能实施形式中,音频声道信号和参考音频信号是频域信号,且音频声道信号值和参考音频信号值与频率区间或频率子带相关联。

[0037] 所使用的频率分辨率主要由听觉系统的频率分辨率激发。心理声学表明空间感知最可能基于声输入信号的临界带表示。通过使用具有子带的可逆滤波器组来考虑这种频率分辨率,所述子带具有与听觉系统的临界带宽相等或成比例的带宽。因此,参数音频编码器可以很好地适应人类感知。

[0038] 在根据第一方面或根据第一方面的前述实施形式中任一者的参数音频编码器的第四可能实施形式中,参数音频编码器进一步包括变换器,用于变换频域中的多个时域音频声道信号,以获得多个音频声道信号。

[0039] 声道脉冲响应的均衡可以在频域中有效地执行,因为时域中的卷积是频域中的倍增。因此,在频域中执行参数音频编码器的计算可以相对于计算复杂度产生更高效或产生更高精确性。

[0040] 在根据第一方面或根据第一方面的前述实施形式中任一者的参数音频编码器的

第五可能实施形式中,参数产生器用于确定音频声道信号的每个频率区间或每个频率子带的第一组编码参数。

[0041] 参数音频编码器可以将确定第一组编码参数限制到可由人耳感知且因此降低复杂度的频率区间或频率子带。

[0042] 在根据第一方面或根据第一方面的前述实施形式中任一者的参数音频编码器的第六可能实施形式中,参数产生器用于确定音频声道信号的第一编码参数平均值作为频率区间或频率子带上的音频声道信号的第一组编码参数的平均值。

[0043] 通过该种平均,参数音频编码器提供考虑所有频率分量的音频信号的短时间平均值。

[0044] 在根据第一方面或根据第一方面的前述实施形式中任一者的参数音频编码器的第七可能实施形式中,参数产生器用于确定音频声道信号的第二编码参数平均值作为音频声道信号的多个帧上的多个第一编码参数平均值的平均值,其中每个第一编码参数平均值与多声道音频信号的帧相关联。

[0045] 通过该种平均,参数音频编码器提供考虑语言信号或音乐信号的特征特性的音频信号的长时间平均值。

[0046] 在根据第一方面或根据第一方面的前述实施形式中任一者的参数音频编码器的第八可能实施形式中,参数产生器用于确定第二编码参数平均值与第一编码参数平均值之间的差的绝对值。

[0047] 通过该种差,参数音频编码器提供对长时间平均值与短时间平均值之间的差的测量并且因此能够预测语言或音乐的行为。

[0048] 在根据第一方面的第八实施形式的参数音频编码器的第九可能实施形式中,参数产生器用于根据所确定的绝对值来确定编码参数。

[0049] 当根据所确定的绝对值提供编码参数时,存在编码参数与所确定的绝对值之间的关系,所述关系可以用于有效地计算编码参数。因此降低计算复杂度。

[0050] 在根据第八实施形式或根据第一方面的第九实施形式的参数音频编码器的第十可能实施形式中,参数产生器用于根据第一参数值与乘以第二参数值的所确定的绝对值之间的差来确定编码参数。

[0051] 当编码参数被提供为第一参数值与所确定的绝对值之间的差时,存在编码参数与所确定的绝对值之间的关系,所述关系可以用于有效地计算编码参数。因此降低计算复杂度。

[0052] 在根据第一方面的第十实施形式的参数音频编码器的第十一可能实施形式中,参数产生器用于将第一参数值设置为一且将第二参数值设置为一。

[0053] 通过该种关系,参数音频编码器能够有效地计算编码参数。因此降低计算复杂度。

[0054] 在根据第一方面或根据第一方面的前述实施形式中任一者的参数音频编码器的第十二可能实施形式中,参数音频编码器进一步包括:下混合信号产生器,用于叠加多声道音频信号的音频声道信号中的至少两者,以获得下混合信号;音频编码器,具体而言单声道编码器,用于编码下混合信号以获得经编码的音频信号;以及组合器,用于将经编码的音频信号与对应的编码参数组合。

[0055] 下混合信号和经编码的音频信号可以用作参数产生器的参考信号。两个信号都包

括多个音频声道信号并且因此比用作参考信号的单声道信号提供更高的精确性。

[0056] 在根据第一方面或根据第一方面的前述实施形式中任一者的参数音频编码器的第十三实施形式中,第一编码参数平均值是指音频声道信号的当前帧,并且另一第一编码参数平均值是指音频声道信号的先前帧。

[0057] 通过使用音频声道信号的当前帧和先前帧,可以有效地执行长时间平均。

[0058] 在根据第一方面的第十三实施形式的参数音频编码器的第十四实施形式中,音频声道信号的当前帧与音频声道信号的先前帧是邻接的。

[0059] 当两个帧是连续的时,音频声道信号中的尖峰在平均值中检测到并且可以在参数音频编码器中考虑。因此编码比无法检测到尖峰的编码更精确。

[0060] 根据第二方面,本发明涉及一种用于产生多声道音频信号的多个音频声道信号中的音频声道信号的编码参数的参数音频编码器,每个音频声道信号具有音频声道信号值,所述参数音频编码器包括参数产生器,所述参数产生器用于

[0061] -根据音频声道信号的音频声道信号值以及参考音频信号的参考音频信号值为多个音频声道信号中的音频声道信号确定第一组编码参数,其中参考音频信号是从多个多声道音频信号中的至少两个音频声道信号中获得的下混合音频信号,

[0062] -基于音频声道信号的第一组编码参数为音频声道信号确定第一编码参数平均值,

[0063] -基于音频声道信号的第一编码参数平均值以及音频声道信号的至少一个另一第一编码参数平均值为音频声道信号确定第二编码参数平均值,以及

[0064] -基于音频声道信号的第一编码参数平均值以及音频声道信号的第二编码参数平均值确定编码参数。

[0065] 参考音频信号可以是多声道音频信号的音频声道信号中的一者。具体而言,参考音频信号可以是形成两个声道的多声道信号的实施例的立体声信号的左音频声道信号或右音频声道信号。然而,参考音频信号可以是形成用于确定编码参数的参考的任何信号。此参考信号可以在下混合多声道音频信号的声道之后由下混合音频信号形成,或者由单声道编码器的输出形成。

[0066] 参数音频编码器可能具有较低的复杂度,因为该参数音频编码器不需要相干性或相关性计算。当ICC是利用仅需要几个步骤的粗糙量化器进行量化时,所述参数音频编码器甚至提供音频声道之间的关系精确估计。尤其针对音乐信号,还针对语言信号,使用对音频信号进行编码的编码参数是非常重要的,因为在具有正确的声音场景宽度的情况下,输出的音乐听起来更自然且不“干燥”。对于极低位率的参数立体声音频编码方案,比特预算是受限的且仅传输一个全带ICC,编码参数能够表示声道之间的全局相关性。

[0067] 在根据第二方面的参数音频编码器的第一可能实施形式中,第一组编码参数是以下参数中的一者:声道间电平差、声道间相位差、声道间相干性、声道间强度差、子带声道间电平差、子带声道间相位差、子带声道间相干性以及子带声道间强度差。

[0068] 此些参数表示音频信号之间的相似度并且因此可以由编码器使用,以便减少待传输的信息并且因此降低计算复杂度。

[0069] 在根据第二方面或根据第二方面的第一实施形式的参数音频编码器的第二可能实施形式中,参数产生器用于确定随后的音频声道信号值的相位差以获得第一组编码参

数。

[0070] 需要随后的音频声道信号的相位差以用于再现声道之间的相位差和/或延迟差。当再现相位差时,语言和音乐听起来会更自然。

[0071] 在根据第二方面或根据第二方面的前述实施形式中任一者的参数音频编码器的第三可能实施形式中,音频声道信号和参考音频信号是频域信号,且音频声道信号值和参考音频信号值与频率区间或频率子带相关联。

[0072] 所使用的频率分辨率主要由听觉系统的频率分辨率激发。心理声学表明空间感知最可能基于声输入信号的临界带表示。通过使用具有子带的可逆滤波器组来考虑这种频率分辨率,所述子带具有与听觉系统的临界带宽相等或成比例的带宽。因此,参数音频编码器可以很好地适应人类感知。

[0073] 在根据第二方面或根据第二方面的前述实施形式中任一者的参数音频编码器的第四可能实施形式中,参数音频编码器进一步包括变换器,用于变换频域中的多个时域音频声道信号,以获得多个音频声道信号。

[0074] 声道脉冲响应的均衡可以在频域中有效地执行,因为时域中的卷积是频域中的倍增。因此,在频域中执行参数音频编码器的计算可以相对于计算复杂度产生更高效率或产生更高精确性。

[0075] 在根据第二方面或根据第二方面的前述实施形式中任一者的参数音频编码器的第五可能实施形式中,参数产生器用于确定音频声道信号的每个频率区间或每个频率子带的第一组编码参数。

[0076] 参数音频编码器可以将确定第一组编码参数限制到可由人耳感知且因此降低复杂度的频率区间或频率子带。

[0077] 在根据第二方面或根据第二方面的前述实施形式中任一者的参数音频编码器的第六可能实施形式中,参数产生器用于确定音频声道信号的第一编码参数平均值作为频率区间或频率子带上的音频声道信号的第一组编码参数的平均值。

[0078] 通过该种平均,参数音频编码器提供考虑所有频率分量的音频信号的短时间平均值。

[0079] 在根据第二方面或根据第二方面的前述实施形式中任一者的参数音频编码器的第七可能实施形式中,参数产生器用于确定音频声道信号的第二编码参数平均值作为音频声道信号的多个帧上的多个第一编码参数平均值的平均值,其中每个第一编码参数平均值与多声道音频信号的帧相关联。

[0080] 通过该种平均,参数音频编码器提供考虑语言信号或音乐信号的特征特性的音频信号的长时间平均值。

[0081] 在根据第二方面或根据第二方面的前述实施形式中任一者的参数音频编码器的第八可能实施形式中,参数产生器用于确定第二编码参数平均值与第一编码参数平均值之间的差的绝对值。

[0082] 通过该种差,参数音频编码器提供对长时间平均值与短时间平均值之间的差的测量并且因此能够预测语言或音乐的行为。

[0083] 在根据第二方面的第八实施形式的参数音频编码器的第九可能实施形式中,参数产生器用于根据所确定的绝对值来确定编码参数。

[0084] 当根据所确定的绝对值提供编码参数时,存在编码参数与所确定的绝对值之间的关系,所述关系可以用于有效地计算编码参数。因此降低计算复杂度。

[0085] 在根据第八实施形式或根据第二方面的第九实施形式的参数音频编码器的第十可能实施形式中,参数产生器用于根据第一参数值与乘以第二参数值的所确定的绝对值之间的差来确定编码参数。

[0086] 当编码参数被提供为第一参数值与所确定的绝对值之间的差时,存在编码参数与所确定的绝对值之间的关系,所述关系可以用于有效地计算编码参数。因此降低计算复杂度。

[0087] 在根据第二方面的第十实施形式的参数音频编码器的第十一可能实施形式中,参数产生器用于将第一参数值设置为一且将第二参数值设置为一。

[0088] 通过该种关系,参数音频编码器能够有效地计算编码参数。因此降低计算复杂度。

[0089] 在根据第二方面或根据第二方面的前述实施形式中任一者的参数音频编码器的第十二可能实施形式中,参数音频编码器进一步包括:下混合信号产生器,用于叠加多声道音频信号的音频声道信号中的至少两者,以获得下混合信号;音频编码器,具体而言单声道编码器,用于编码下混合信号以获得经编码的音频信号;以及组合器,用于将经编码的音频信号与对应的编码参数组合。

[0090] 下混合信号和经编码的音频信号可以用作参数产生器的参考信号。两个信号都包括多个音频声道信号并且因此比用作参考信号的单声道信号提供更高的精确性。

[0091] 在根据第二方面或根据第二方面的前述实施形式中任一者的参数音频编码器的第十三实施形式中,第一编码参数平均值是指音频声道信号的当前帧,并且另一第一编码参数平均值是指音频声道信号的先前帧。

[0092] 通过使用音频声道信号的当前帧和先前帧,可以有效地执行长时间平均。

[0093] 在根据第二方面的第十三实施形式的参数音频编码器的第十四实施形式中,音频声道信号的当前帧与音频声道信号的先前帧是邻接的。

[0094] 当两个帧是连续的时,音频声道信号中的尖峰在平均值中检测到并且可以在参数音频编码器中考虑。因此编码比无法检测到尖峰的编码更精确。

[0095] 根据第三方面,本发明涉及一种用于产生多声道音频信号的多个音频声道信号中的音频声道信号的编码参数的方法,每个音频声道信号具有音频声道信号值,所述方法包括:

[0096] -根据音频声道信号的音频声道信号值以及参考音频信号的参考音频信号值为多个音频声道信号中的音频声道信号确定第一组编码参数,其中参考音频信号是多个音频声道信号中的另一音频声道信号,

[0097] -基于音频声道信号的第一组编码参数为音频声道信号确定第一编码参数平均值,

[0098] -基于音频声道信号的第一编码参数平均值以及音频声道信号的至少一个另一第一编码参数平均值为音频声道信号确定第二编码参数平均值,以及

[0099] -基于音频声道信号的第一编码参数平均值以及音频声道信号的第二编码参数平均值确定编码参数。

[0100] 所述方法可以在处理器上有效地执行。

[0101] 参考音频信号可以是多声道音频信号的音频声道信号中的一者。具体而言,参考音频信号可以是形成两个声道的多声道信号的实施例的立体声信号的左音频声道信号或右音频声道信号。然而,参考音频信号可以是形成用于确定编码参数的参考的任何信号。此种参考信号可以在下混合多声道音频信号的声道之后由单声道下混合音频信号形成,或在下混合多声道音频信号的声道之后由下混合音频信号的声道中的一者形成。

[0102] 根据第四方面,本发明涉及一种用于产生多声道音频信号的多个音频声道信号中的音频声道信号的编码参数的方法,每个音频声道信号具有音频声道信号值,所述方法包括:

[0103] -根据音频声道信号的音频声道信号值以及参考音频信号的参考音频信号值为多个音频声道信号中的音频声道信号确定第一组编码参数,其中参考音频信号是从多个多声道音频信号中的至少两个音频声道信号中获得的下混合音频信号,

[0104] -基于音频声道信号的第一组编码参数为音频声道信号确定第一编码参数平均值,

[0105] -基于音频声道信号的第一编码参数平均值以及音频声道信号的至少一个另一第一编码参数平均值为音频声道信号确定第二编码参数平均值,以及

[0106] -基于音频声道信号的第一编码参数平均值以及音频声道信号的第二编码参数平均值确定编码参数。

[0107] 所述方法可以在处理器上有效地执行。

[0108] 参考音频信号可以是多声道音频信号的音频声道信号中的一者。具体而言,参考音频信号可以是形成两个声道的多声道信号的实施例的立体声信号的左音频声道信号或右音频声道信号。然而,参考音频信号可以是形成用于确定编码参数的参考的任何信号。此种参考信号可以在下混合多声道音频信号的声道之后由单声道下混合音频信号形成,或在下混合多声道音频信号的声道之后由下混合音频信号的声道中的一者形成。

[0109] 根据第五方面,本发明涉及一种计算机程序,当在计算机上执行时,所述计算机程序用于实施根据本发明的第三和第四方面中的一者的方法。

[0110] 所述计算机程序具有降低的复杂度并且因此可以有效地在电池寿命必须节省的移动终端中实施。当所述计算机程序在移动终端上运行时,电池寿命时间增加。

[0111] 本文中所描述的方法可以实施为数字信号处理器(DSP:Digital Signal Processor)、微控制器或任何其他辅助处理器中的软件或实施为专用集成电路(ASIC:application specific integrated circuit)内的硬件电路。

[0112] 本发明可以在数字电子电路中实施,或在计算机硬件、固件、软件或其组合中实施。

附图说明

[0113] 将相对于以下图式描述本发明的其他实施例,其中:

[0114] 图1示出根据实施形式的参数音频编码器的方块图;

[0115] 图2示出根据实施形式的参数音频解码器的方块图;

[0116] 图3示出根据实施形式的参数立体声音频编码器和解码器的方块图;以及

[0117] 图4示出根据实施形式的用于产生音频声道信号的编码参数的方法的示意图。

具体实施方式

[0118] 图1示出根据实施形式的参数音频编码器100的方块图。参数音频编码器100接收多声道音频信号101作为输入信号,并且提供比特流作为输出信号103。参数音频编码器100包括:参数产生器105,所述参数产生器耦合到多声道音频信号101上,用于产生编码参数115;下混合信号产生器107,所述下混合信号产生器耦合到多声道音频信号101上,用于产生下混合信号111或和信号;音频编码器109,所述音频信号耦合到下混合信号产生器107上,用于编码下混合信号111以提供经编码的音频信号113;以及组合器117(例如,比特流形成器),所述组合器耦合到参数产生器105和音频编码器109上以从编码参数115和经编码的信号113形成比特流103。

[0119] 参数音频编码器100实施立体声和多声道音频信号的音频编码方案,所述音频编码方案仅传输一个单音频声道,例如,下混合音频声道加上描述音频声道 $X_1[b]$ 、 $X_2[b]$ 、 \dots 、 $X_M[b]$ 之间的“可感知相关差异”的额外参数。所述编码方案是根据双耳线索编码(BCC),因为双耳线索在编码方案中起重要作用。如图中所指示,多声道音频信号101的多个(M个)输入音频声道 $X_1[b]$ 、 $X_2[b]$ 、 \dots 、 $X_M[b]$ 被下混合到一个单音频声道111中,也表示为和信号。对于立体声音频信号,M等于2。如音频声道 $X_1[b]$ 、 $X_2[b]$ 、 \dots 、 $X_M[b]$ 之间的“可感知相关差异”,编码参数115,例如,声道间时间差(ICTD)、声道间电平差(ICLD)和/或声道间相干性(ICC),是根据频率和时间进行估计,并且作为辅助信息传输到图2中所述的解码器200。

[0120] 实施BCC的参数产生器105以某一时间和频率分辨率处理多声道音频信号101。所使用的频率分辨率主要由听觉系统的频率分辨率激发。心理声学表明空间感知最可能基于声输入信号的临界带表示。通过使用具有子带的可逆滤波器组来考虑这种频率分辨率,所述子带具有与听觉系统的临界带宽相等或成比例的带宽。所传输的和信号111含有多声道音频信号101的所有信号分量是非常重要的。目标是完全保持每个信号分量。多声道音频信号101的音频输入声道 $X_1[b]$ 、 $X_2[b]$ 、 \dots 、 $X_M[b]$ 的简单求和通常会引起信号分量的放大或衰减。换句话说,“简单”总和中的信号分量的功率通常大于或小于每个声道 $X_1[b]$ 、 $X_2[b]$ 、 \dots 、 $X_M[b]$ 的对应信号分量的功率总和。因此,通过应用下混合装置107使用下混合技术,所述下混合装置使和信号111均衡化,使得和信号111中的信号分量的功率大致相同于多声道音频信号101的所有输入音频声道 $X_1[b]$ 、 $X_2[b]$ 、 \dots 、 $X_M[b]$ 中的对应功率。输入音频声道 $X_1[b]$ 、 $X_2[b]$ 、 \dots 、 $X_M[b]$ 表示子带b的声道信号。频域输入音频声道表示为 $X_1[k]$ 、 $X_2[k]$ 、 \dots 、 $X_M[k]$,其中k表示频率指数(频率区),子带b通常由若干频率区k构成。

[0121] 给定和信号111,参数产生器105合成立体声或多声道音频信号115,使得ICTD、ICLD和/或ICC接近于原始多声道音频信号101的对应线索。

[0122] 当考虑一个源的双耳房间脉冲响应(BRIR)时,存在听觉事件的宽度以及听者包围感与估计用于BRIR的早期和后期部分的IC之间的关系。然而,IC(或ICC)与普通信号(并且不仅仅是BRIR)的这些特性之间的关系并不直观。立体声和多声道音频信号通常含有同时活动的源信号的复杂混合,所述源信号由在闭合空间中录音产生的反射信号分量叠加或者由用于人工创建空间印象的录音工程师添加。不同源信号和它们的反射占据时频平面中的不同区域。这通过根据时间和频率而变的ICTD、ICLD和ICC反映。在此情况下,瞬时ICTD、ICLD和ICC与听觉事件方向以及空间印象之间的关系并不明显。参数产生器105的策略是无

目的地合成这些线索,使得这些线索接近于原始音频信号的对应线索。

[0123] 在实施形式中,参数音频编码器100使用具有带宽等于等效矩形带宽两倍的子带的滤波器组。非正式旁听揭示出,当选择更高频率分辨率时,BCC的音频质量未得到显著提高。较低频率分辨率是有利的,因为较低频率分辨率会引起需要传输到解码器的ICTD、ICLD和ICC值较小且因此位率较低。关于时间分辨率,以规则时间间隔考虑ICTD、ICLD和ICC。在实施形式中,约每4至16ms考虑ICTD、ICLD和ICC。应注意,除非在极短时间间隔内考虑线索,否则不会直接考虑优先效应。

[0124] 参考信号与合成信号之间的通常获得的可感知较小差异意味着与大范围的听觉空间图像属性有关的线索是通过以规则时间间隔合成ICTD、ICLD和ICC而隐式地进行考虑。传输这些空间线索所需的位率仅为几kb/每秒,并且因此参数音频编码器100能够以与单音频声道所需的位率接近的位率传输立体声和多声道音频信号。图4图示了ICC被估计为编码参数115的方法。

[0125] 参数音频编码器100包括:下混合信号产生器107,用于叠加多声道音频信号101的音频声道信号中的至少两者,以获得下混合信号111;音频编码器109,具体而言单声道编码器,用于编码下混合信号111以获得经编码的音频信号113;以及组合器117,用于将经编码的音频信号113与对应的编码参数115组合。

[0126] 参数音频编码器100产生多声道音频信号101的表示为 $X_1[b]$ 、 $X_2[b]$ 、 \dots 、 $X_M[b]$ 的多个音频声道信号中的一个音频声道信号的编码参数115。音频声道信号 $X_1[b]$ 、 $X_2[b]$ 、 \dots 、 $X_M[b]$ 中的每一者可以是包括频域中表示为 $X_1[k]$ 、 $X_2[k]$ 、 \dots 、 $X_M[k]$ 的数字音频声道信号值的数字信号。

[0127] 参数音频编码器100为其产生编码参数115的示例性音频声道信号是具有信号值 $X_1[k]$ 的第一音频声道信号 $X_1[b]$ 。参数产生器105根据音频声道信号 $X_1[b]$ 的音频声道信号值 $X_1[k]$ 以及参考音频信号的参考音频信号值,为音频声道信号 $X_1[b]$ 确定表示为IPD[b]的第一组编码参数。

[0128] 例如,用作参考音频信号的音频声道信号是第二音频声道信号 $X_2[b]$ 。类似地,音频声道信号 $X_1[b]$ 、 $X_2[b]$ 、 \dots 、 $X_M[b]$ 中的任何其他一者可以充当参考音频信号。根据第一方面,参考音频信号是音频声道信号中不等于产生编码参数115的音频声道信号 $X_1[b]$ 的另一音频声道信号。

[0129] 根据第二方面,参考音频信号是从多个多声道音频信号101的至少两个音频声道信号中获得(例如,从第一音频声道信号 $X_1[b]$ 和第二音频声道信号 $X_2[b]$ 中获得)的下混合音频信号。在实施形式中,参考音频信号是下混合信号111,也称为由下混合装置107产生的和信号。在实施形式中,参考音频信号是由编码器109提供的经编码的信号113。

[0130] 参数产生器105使用的示例性参考音频信号是具有信号值 $X_2[k]$ 的第二音频声道信号 $X_2[b]$ 。

[0131] 参数产生器105基于音频声道信号 $X_1[b]$ 的第一组编码参数IPD[b]为音频声道信号 $X_1[b]$ 确定表示为 $IPD_{\text{mean}}[i]$ 的第一编码参数平均值。

[0132] 参数产生器105基于音频声道信号 $X_1[b]$ 的第一编码参数平均值 $IPD_{\text{mean}}[i]$ 以及音频声道信号 $X_1[b]$ 的至少一个另一第一编码参数平均值(表示为 $IPD_{\text{mean}}[i-1]$)为音频声道信号 $X_1[b]$ 确定表示为 $IPD_{\text{mean_long_term}}$ 的第二编码参数平均值。

[0133] 在实施形式中,第一编码参数平均值 $IPD_{mean}[i]$ 是指音频声道信号 $X_1[b]$ 的当前帧 i ,并且另一第一编码参数平均值 $IPD_{mean}[i-1]$ 是指音频声道信号 $X_1[b]$ 的先前帧 $i-1$ 。在实施形式中,音频声道信号 $X_1[b]$ 的先前帧 $i-1$ 是在当前帧 i 之前接收到的帧 $i-1$,其中这两个帧之间不存在其他帧。在实施形式中,音频声道信号 $X_1[b]$ 的先前帧 $i-N$ 是在当前帧 i 之前接收到的帧 $i-N$,但是在这两个帧之间已到达多个帧。

[0134] 参数产生器105基于音频声道信号 $X_1[b]$ 的第一编码参数平均值 $IPD_{mean}[i]$ 并且基于音频声道信号 $X_1[b]$ 的第二编码参数平均值 $IPD_{mean_long_term}$ 确定表示为ICC的编码参数115。

[0135] 第一组编码参数 $IPD[b]$ 是声道间相位差、声道间电平差、声道间相干性、声道间强度差、子带声道间电平差、子带声道间相位差、子带声道间相干性、子带声道间强度差或其组合。声道间相位差(ICPD)是信号对之间的平均相位差。声道间电平差(ICLD)与耳间电平差(ILD)相同,即,左耳进入信号与右耳进入信号之间的电平差,但是更普遍地界定在任何信号对之间,例如,扩音器信号对、耳朵进入信号对等。声道间相干性或声道间相关性与耳间相干性(IC)相同,即,左耳进入信号与右耳进入信号之间的相似度,但是更普遍地界定在任何信号对之间,例如,扩音器信号对、耳朵进入信号对等。声道间时间差(ICTD)与耳间时间差(ITD)相同,有时还称为耳间时间延迟,即,左耳进入信号与右耳进入信号之间的时间差,但是更普遍地界定在任何信号对之间,例如,扩音器信号对、耳朵进入信号对等。子带声道间电平差、子带声道间相位差、子带声道间相干性以及子带声道间强度差与上文相对于子带带宽指定的参数有关。

[0136] 参数产生器101确定随后的音频声道信号值 $X_1[k]$ 的相位差,以获得第一组编码参数 $IPD[b]$ 。在实施形式中,音频声道信号 $X_1[b]$ 和参考音频信号 $X_2[b]$ 是频域信号,并且音频声道信号值 $X_1[k]$ 和参考音频信号值 $X_2[k]$ 与表示为 $[k]$ 的频率区间或表示为 $[b]$ 的频率子带相关联。在实施形式中,参数音频编码器100包括变换器,例如,用于变换频域中的多个时域音频声道信号 $X_1[n]$ 、 $X_2[n]$ 以获得多个音频声道信号 $X_1[b]$ 、 $X_2[b]$ 的FFT装置。在实施形式中,参数产生器101确定音频声道信号 $X_1[b]$ 、 $X_2[b]$ 的每个频率区间 $[k]$ 或每个频率子带 $[b]$ 的第一组编码参数 $IPD[b]$ 。

[0137] 在第一步骤中,参数产生器105将时频变换应用于时域输入声道,例如,第一输入声道 $x_1[n]$,以及时域参考声道,例如,第二输入声道 $x_2[n]$ 。在立体声的情况下,这些是左声道和右声道。在优选实施例中,时频变换是快速傅里叶变换(FFT)。在替代实施例中,时频变换是余弦调制滤波器组或复杂滤波器组。

[0138] 在第二步骤中,参数产生器105将FFT的每个频率区间 $[b]$ 的互谱计算为:

$$[0139] \quad c[b] = X_1[b]X_2^*[b],$$

[0140] 其中 $c[b]$ 是频率区间 $[b]$ 的互谱并且 $X_1[b]$ 和 $X_2[b]$ 是两个声道的FFT系数。 $*$ 表示复共轭性。对于这种情况,子带 $[b]$ 直接对应于一个频率区间 $[k]$,频率区间 $[b]$ 和 $[k]$ 确切地表示同一频率区间。

[0141] 或者,参数产生器105将每子带 $[b]$ 的互谱计算为:

$$[0142] \quad c[b] = \sum_{k=k_b}^{k_{b+1}-1} X_1[k]X_2^*[k],$$

[0143] 其中 $c[b]$ 是子带 $[b]$ 的互谱并且 $X_1[k]$ 和 $X_2[k]$ 是两个声道的FFT系数。 $*$ 表示复共轭

性。 k_b 是子带**b**的开始区间并且 k_{b+1} 是相邻子带**b+1**的开始区间。因此, k_b 与 $k_{b+1}-1$ 之间的FFT的频率区间[k]表示子带[b]。

[0144] 声道间相位差(IPD)基于互谱每子带计算为:

$$[0145] \quad \text{IPD}[b] = \angle c[b]$$

[0146] 其中运算 \angle 是用以计算 $c[b]$ 的角度的变量参数运算符。

[0147] 在实施形式中,参数产生器101确定音频声道信号 $X_1[b]$ 的第一编码参数平均值 $\text{IPD}_{\text{mean}}[i]$ 作为频率区间[b]或频率子带[b]上的音频声道信号 $X_1[b]$ 的第一组编码参数IPD[b]的平均值。

[0148] 频率区间[b]或频率子带[b]上的平均IPD(IPD_{mean})如以下方程式中定义的那样进行计算:

$$[0149] \quad \text{IPD}_{\text{mean}} = \frac{\sum_{k=1}^K \text{IPD}[k]}{K}$$

[0150] 其中K是计算平均值所考虑的频率区间或频率子带的数目。

[0151] 在实施形式中,参数产生器101确定音频声道信号 $X_1[b]$ 的第二编码参数平均值 $\text{IPD}_{\text{mean_long_term}}$ 作为音频声道信号 $X_1[b]$ 的多个帧上的多个第一编码参数平均值 $\text{IPD}_{\text{mean}}[i]$ 的平均值,其中每个第一编码参数平均值 $\text{IPD}_{\text{mean}}[i]$ 与多声道音频信号的帧[i]相关联。

[0152] 基于先前计算的 IPD_{mean} ,参数产生器105计算IPD的长期平均值。 $\text{IPD}_{\text{mean_long_term}}$ 被计算为最后N个帧(例如,N可以设置为10)上的平均值。

$$[0153] \quad \text{IPD}_{\text{mean_long_term}} = \frac{\sum_{i=1}^N \text{IPD}_{\text{mean}}[i]}{N}$$

[0154] 在实施形式中,参数产生器101确定第二编码参数平均值 $\text{IPD}_{\text{mean_long_term}}$ 与第一编码参数平均值 $\text{IPD}_{\text{mean}}[i]$ 之间的差的绝对值 IPD_{dist} 。

[0155] 为了评估IPD参数的稳定性,计算 IPD_{mean} 与 $\text{IPD}_{\text{mean_long_term}}$ (IPD_{dist})之间的距离,这示出在最后N个帧过程中的IPD的评估。在优选实施例中,局部IPD与长期IPD之间的距离被计算为局部平均值与长期平均值之间的差的绝对值:

$$[0156] \quad \text{IPD}_{\text{dist}} = \text{abs}(\text{IPD}_{\text{mean}} - \text{IPD}_{\text{mean_long_term}})$$

[0157] 可以看出,如果 IPD_{mean} 参数在先前帧上是稳定的,则距离 IPD_{dist} 变得接近0。当相位差随着时间推移稳定时,距离随后等于零。此距离对声道的相似性给出较好估计。

[0158] 在实施形式中,参数产生器101根据所确定的绝对值 IPD_{dist} 确定编码参数ICC。在实施形式中,参数产生器101根据第一参数值d与乘以第二参数值e的所确定的绝对值 IPD_{dist} 之间的差确定编码参数ICC。在实施形式中,参数产生器101将第一参数值d设置为一并且将第二参数值e设置为一。

[0159] 相干性或ICC参数被计算为 $\text{ICC} = 1 - \text{IPD}_{\text{dist}}$,因为ICC和 IPD_{dist} 具有间接的倒数关系。当声道类似时,ICC接近1,且在这种情况下 IPD_{dist} 变成等于0。

[0160] 或者,用以界定ICC与 IPD_{dist} 之间的关系的方程为被定义为 $\text{ICC} = d - e \cdot \text{IPD}_{\text{dist}}$,其中d和e经更好地选择以表示两个参数之间的倒数关系。在另一实施例中,ICC与 IPD_{dist} 之间的关系是通过在较大数据库中训练而获得并且随后被一般化为 $\text{ICC} = f(\text{IPD}_{\text{dist}})$ 。

[0161] 在音频信号(例如,对于语言信号)的相关段过程中 IPD_{dist} 较小并且在音频输入

(例如,对于音乐信号)的扩散部分过程中,此 IPD_{dist} 参数变得大得多,并且如果输入声道是去相关的,则 IPD_{dist} 参数将接近1。因此,ICC和 IPD_{dist} 具有间接的倒数关系。

[0162] 图2示出根据实施形式的参数音频解码器200的方块图。参数音频解码器200接收在通信声道上传输的比特流203作为输入信号,并且提供经解码的多声道音频信号201作为输出信号。参数音频解码器200包括:比特流解码器217,所述比特流解码器耦合到比特流203上,用于将比特流203解码成编码参数215和经编码的信号213;解码器209,所述解码器耦合到比特流解码器217上,用于根据经编码的信号213产生和信号211;参数解码器205,所述参数解码器耦合到比特流解码器217上,用于根据编码参数215解码参数221;以及合成器205,所述合成器耦合到参数解码器205和解码器209上,用于根据参数221以及和信号211合成经解码的多声道音频信号201。

[0163] 参数音频解码器200产生其多声道音频信号201的输出声道,使得声道之间的ICTD、ICLD和/或ICC接近于原始多声道音频信号的那些ICTD、ICLD和/或ICC。所描述的方案能够以比表示单声道音频信号所需的位率仅略高的位率表示多声道音频信号。由于声道对之间的所估计的ICTD、ICLD和ICC含有比音频波形少约两个数量级的信息,因此产生以上情况。所关注的不仅是低位率而且是反向兼容性方面。所传输的和信号对应于立体声或多声道信号的单声道下混合。

[0164] 图3示出根据实施形式的参数立体声音频编码器301和解码器303的方块图。参数立体声音频编码器301对应于相对于图1所述的参数音频编码器100,但是多声道音频信号101是具有左音频声道305和右音频声道307的立体声音频信号。

[0165] 参数立体声音频编码器301接收立体声音频信号305、307作为输入信号,其包括左声道音频信号305和右声道音频信号307,并且提供比特流作为输出信号309。参数立体声音频编码器301包括:参数产生器311,所述参数产生器耦合到立体声音频信号305、307上,用于产生空间参数313;下混合信号产生器315,所述下混合信号产生器耦合到立体声音频信号305、307上,用于产生下混合信号317或和信号;单声道编码器319,所述单声道编码器耦合到下混合信号产生器315上,用于编码下混合信号317以提供经编码的音频信号321;以及比特流组合器323,所述比特流组合器耦合到参数产生器311和单声道编码器319上,以将编码参数313和经编码的音频信号321组合成比特流以提供输出信号309。在参数产生器311中,抽取和量化空间参数313,随后所述空间参数在比特流中进行多路复用。

[0166] 参数立体声音频解码器303接收比特流作为输入信号,所述比特流即,在通信声道上传输的参数立体声音频编码器301的输出信号309,并且提供具有左声道325和右声道327的经解码的立体声音频信号作为输出信号。参数立体声音频解码器303包括:比特流解码器329,所述比特流解码器耦合到接收到的比特流309上,用于将比特流309解码成编码参数331和经编码的信号333;单声道解码器335,所述单声道解码器耦合到比特流解码器329上,用于根据经编码的信号333产生和信号337;空间参数解码器339,所述空间参数解码器耦合到比特流解码器329上,用于根据编码参数331解码空间参数341;以及合成器343,所述合成器耦合到空间参数解码器或解析器339以及单声道解码器335上,用于根据空间参数341以及和信号337合成经解码的立体声音频信号325、327。

[0167] 参数立体声音频编码器301中的处理能够提取延迟,并且在时间和频率中自适应地计算音频信号的级别,以产生空间参数313,例如,声道间时间差(ICTD)和声道间电平差

(ICLD)。此外,参数立体声音频编码器301对声道间相干性(ICC)合成有效地执行时间自适应滤波。在实施形式中,参数立体声编码器使用基于短时间傅里叶变换(STFT)的滤波器组,以便有效地实施具有低计算复杂度的双耳线索编码(BCC)方案。参数立体声音频编码器301中的处理具有低计算复杂度和低延迟,使得参数立体声音频编码适合于在微处理器或数字信号处理器上适用地实施以用于实时应用。

[0168] 图3中描绘的参数产生器311在功能上与相对于图1所描述的对参数产生器105相同,不同的是已添加空间线索的量化和编码用于说明。和信号317是利用常规的单声道音频编码器319进行编码。在实施形式中,参数立体声音频编码器301使用基于STFT的时频变换,以变换频域中的立体声音频声道信号305、307。STFT将离散傅里叶变换(DFT)应用于输入信号 $x(n)$ 的窗口化部分。在应用N点DFT之前,将N个样本的信号帧乘以窗口长度W。相邻窗口重叠并且移位W/2个样本的距离。窗口经选择,以使得重叠的窗口总计达恒定值1。因此,对于逆变换,不需要额外开窗口。解码器303中使用尺寸N的简单逆DFT,所述逆DFT具有W/2个样本的连续帧的时间提前。如果未修改频谱,则通过重叠/添加实现完美的重构。

[0169] 由于STFT的均匀频谱分辨率未很好地适应人类感知,因此STFT的均匀间隔开的频谱系数输出被分组成具有较好地适应感知的带宽的B非重叠分区。根据相对于图1的描述,一个分区在概念上对应于一个“子带”。在替代的实施形式中,参数立体声音频编码器301使用不均匀的滤波器组来变换频域中的立体声音频声道信号305、307。

[0170] 在实施形式中,下混合器315通过以下方程式确定均衡的和信号 $S_m(k)$ 317的一个分区b或一个子带b的频谱系数:

$$[0171] \quad S_m(k) = e_b(k) \sum_{c=1}^C X_{c,m}(k),$$

[0172] 其中 $X_{c,m}(k)$ 是输入音频声道305、307的频谱并且 $e_b(k)$ 是如下进行计算的增益系数:

$$[0173] \quad e_b(k) = \sqrt{\frac{\sum_{c=1}^C p_{\tilde{x}_{c,b}}(k)}{p_{\tilde{x}_b}(k)}},$$

[0174] 其中分区功率估计为,

$$[0175] \quad p_{\tilde{x}_{c,b}}(k) = \sum_{m=A_{b-1}}^{A_b-1} |X_{c,m}(k)|^2$$

$$[0176] \quad p_{\tilde{x}_b}(k) = \sum_{m=A_{b-1}}^{A_b-1} \left| \sum_{c=1}^C X_{c,m}(k) \right|^2。$$

[0177] 当子带信号的总和的衰减显著时,为了防止较大增益因数引起的非自然信号,增益因数 $e_b(k)$ 可以限制为6dB,即, $e_b(k) \leq 2$ 。

[0178] 在实施形式中,参数产生器311将时频变换,例如,上述STFT或FFT应用到输入声道上,例如,应用到左声道305和右声道307上。在实施形式中,时频变换是快速傅里叶变换(FFT)。在替代实施形式中,时频变换是余弦调制滤波器组或复杂滤波器组。

[0179] 参数产生器311将FFT或STFT的每个频率区间[b]的互谱计算为:

[0180] $c[b] = X_1[b]X_2^*[b]$

[0181] 对于这种情况,子带[b]直接对应于一个频率区间[k],频率区间[b]和[k]确切地表示同一频率区间。

[0182] 或者,参数产生器311将每子带[k]的互谱计算为:

[0183]
$$c[b] = \sum_{k=k_b}^{k_{b+1}-1} X_1[k]X_2^*[k]$$

[0184] 其中c[b]是区间b或子带k的互谱。 $X_1[k]$ 和 $X_2[k]$ 是左声道305和右声道307的FFT系数。运算符*表示复共轭性。 k_b 是子带k的开始区间并且 k_{b+1} 是相邻子带b+1的开始区间。因此, k_b 与 $k_{b+1}-1$ 之间的FFT或STFT的频率区间[k]表示子带[b]。

[0185] 声道间相位差(IPD)是基于互谱每子带计算为:

[0186] $IPD[b] = \angle c[b]$

[0187] 其中运算 \angle 是用以计算c[b]的角度的变量参数运算符。

[0188] 在下文中,参数产生器311如以下方程式所定义的那样计算频率区间或频率子带上的平均IPD(IPD_{mean}):

[0189]
$$IPD_{mean} = \frac{\sum_{k=1}^K IPD[k]}{K}$$

[0190] 其中K是计算平均值所考虑的频率区间或频率子带的数目。

[0191] 随后,基于先前计算的 IPD_{mean} ,参数产生器311计算IPD的长期平均值。 $IPD_{mean_long_term}$ 被计算为最后N个帧上的平均值,在实施形式中,N设置为10。

[0192]
$$IPD_{mean_long_term} = \frac{\sum_{i=1}^N IPD_{mean}[i]}{N}$$

[0193] 为了评估IPD参数的稳定性,参数产生器311计算 IPD_{mean} 与 $IPD_{mean_long_term}$ 之间的距离 IPD_{dist} ,这示出在最后N个帧过程中IPD的演进。在实施形式中,局部IPD与长期IPD之间的距离被计算为局部平均值与长期平均值之间的差的绝对值:

[0194] $IPD_{dist} = \text{abs}(IPD_{mean} - IPD_{mean_long_term})$

[0195] 可以看出,如果 IPD_{mean} 参数在先前帧上是稳定的,则距离 IPD_{dist} 变得接近0。当相位差随着时间推移稳定时,距离随后等于零。此距离对声道的相似性给出较好估计。

[0196] 在实施形式中,参数产生器311将相干性或ICC参数计算为 $ICC = 1 - IPD_{dist}$,因为ICC和 IPD_{dist} 具有间接的倒数关系。当声道类似时,ICC接近1,且在这种情况下 IPD_{dist} 变成等于0。

[0197] 或者,参数产生器311使用被定义为 $ICC = d - e \cdot IPD_{dist}$ 的ICC与 IPD_{dist} 之间的关系,其中d和e是经选择以更好地表示两个参数ICC与 IPD_{dist} 之间的倒数关系的参数。在替代实施形式中,参数产生器311通过在较大数据上训练来获得ICC与 IPD_{dist} 之间的关系,所述关系被一般化为 $ICC = f(IPD_{dist})$ 。

[0198] 在音频信号(例如,对于语言信号)的相关段过程中, IPD_{dist} 较小,并且在音频输入(例如,对于音乐信号)的扩散部分过程中,此 IPD_{dist} 参数变得大得多,并且如果输入声道是

去相关的,则 IPD_{dist} 参数将接近1。因此ICC和 IPD_{dist} 具有间接的倒数关系。

[0199] 参数产生器311使用 IPD_{dist} 来粗略地估计ICC。互谱需要低于相关性计算的复杂度。此外,在参数空间音频编码器中计算IPD的情况下,已计算此互谱并且随后总体复杂度降低。

[0200] 图4示出根据实施形式的用于产生编码参数的方法400的示意图。方法400用于产生多声道音频信号的多个音频声道信号 $x_1[n]$ 、 $x_2[n]$ 中的音频声道信号 $x_1[n]$ 的编码参数ICC。每个音频声道信号 $x_1[n]$ 、 $x_2[n]$ 具有音频声道信号值。图4描绘了其中多个音频声道信号包括左音频声道 $x_1[n]$ 和右音频声道 $x_2[n]$ 的立体声情况。方法400包括:

[0201] 将FFT变换401应用到左音频声道信号 $x_1[n]$ 并且将FFT变换403应用到右音频声道信号 $x_2[n]$ 以获得频域音频声道信号 $X_1[b]$ 和 $X_2[b]$,其中相对于频域中的频率区间 $[b]$, $X_1[b]$ 是左音频声道信号并且 $X_2[b]$ 是右音频声道信号。或者,将滤波器组变换应用于左音频声道信号 $x_1[n]$ 并且应用于右音频声道信号 $x_2[n]$ 上,以获得频率子带中的音频声道信号 $X_1[b]$ 、 $X_2[b]$,其中 $[b]$ 表示频率子带;

[0202] 确定405左音频声道信号 $X_1[b]$ 和右音频声道信号 $X_2[b]$ 的每个频率区间 $[b]$ 的互相关 $c[b]$;或者确定405左音频声道信号 $X_1[b]$ 和右音频声道信号 $X_2[b]$ 的每个频率子带 $[b]$ 的互相关 $c[b]$;

[0203] 根据音频声道信号 $X_1[b]$ 的音频声道信号值以及参考音频信号 $X_2[b]$ 的参考音频信号值为多个音频声道信号的音频声道信号 $X_1[b]$ 确定407第一组编码参数 $IPD[b]$,其中参考音频信号是多个音频声道信号中的另一音频声道信号 $X_2[b]$ 或从多个多声道音频信号的至少两个音频声道信号中获得的下混合音频信号。图4描绘了其中确定407为左音频声道信号 $X_1[b]$ 确定第一组编码参数 $IPD[b]$ 以及其中参考音频信号是右音频声道信号 $X_2[b]$ 的立体声情况;

[0204] 基于音频声道信号 $X_1[b]$ 的第一组编码参数 $IPD[b]$ 为音频声道信号 $X_1[b]$ 确定409第一编码参数平均值 $IPD_{mean}[i]$;

[0205] 基于音频声道信号 $X_1[b]$ 的第一编码参数平均值 $IPD_{mean}[i]$ 以及音频声道信号 $X_1[b]$ 的至少一个另一第一编码参数平均值 $IPD_{mean}[i-1]$ 为音频声道信号 $X_1[b]$ 确定411第二编码参数平均值 $IPD_{mean_long_term}$ 。另一第一编码参数平均值 $IPD_{mean}[i-1]$ 是根据音频声道信号 $X_1[b]$ 的先前 $N-1$ 个帧计算出;以及

[0206] 基于音频声道信号 $X_1[b]$ 的第一编码参数平均值 $IPD_{mean}[i]$ 以及音频声道信号 $X_1[b]$ 的第二编码参数平均值 $IPD_{mean_long_term}$ 确定413或计算编码参数ICC。

[0207] 在实施形式中,音频声道信号 $X_1[b]$ 的第一组编码参数 $IPD[b]$ 已可用,且方法400以上述步骤409、411和413开始。

[0208] 尽管未在图4中描绘,但是方法400适用于多声道音频信号的一般情况,参考信号随后为另一音频声道信号或相对于图1所描述的下混合音频信号。

[0209] 在实施形式中,方法400以如下方式进行处理:

[0210] 在第一步骤401、403中,将时频变换应用于输入声道上(在立体声情况下为左和右)。在优选实施例中,时频变换是快速傅里叶变换(FFT)。在替代实施例中,时频变换可以是余弦调制滤波器组或复杂滤波器组。

[0211] 在第二步骤405中,FFT的每个频率区间的互谱被计算为:

[0212] $c[b] = X_1[b]X_2^*[b]$

[0213] 其中子带[b]直接对应于一个频率区间[k], 频率区间[b]和[k]确切地表示同一频率区间。

[0214] 或者, 互谱可以每子带被计算为:

[0215]
$$c[b] = \sum_{k=k_b}^{k_{b+1}-1} X_1[k]X_2^*[k]$$

[0216] 其中c[b]是区间b或子带b的互谱。 $X_1[k]$ 和 $X_2[k]$ 是两个声道(例如, 在立体声情况下为左声道和右声道)的FFT系数。 $*$ 表示复共轭性。 k_b 是子带b的开始区间并且 k_{b+1} 是相邻子带b+1的开始区间。因此, k_b 与 $k_{b+1}-1$ 之间的FFT的频率区间[k]表示子带[b]。

[0217] 在第三步骤407中, 声道间相位差(IPD)基于互谱每子带被计算为:

[0218] $IPD[b] = \angle c[b]$

[0219] 其中运算 \angle 是用以计算c[b]的角度的变量参数运算符。

[0220] 在第四步骤409中, 频率区间(或频率子带)上的平均IPD(IPD_{mean})还如以下方程式中所定义的那样计算:

[0221]
$$IPD_{mean} = \frac{\sum_{k=1}^K IPD[k]}{K}$$

[0222] 其中K是计算平均值所考虑的频率区间或频率子带的数目。

[0223] 在第五步骤411中, 基于先前计算的 IPD_{mean} 计算IPD的长期平均值。 $IPD_{mean_long_term}$ 被计算为最后N个帧(例如, N可以设置为10)上的平均值。

[0224]
$$IPD_{mean_long_term} = \frac{\sum_{i=1}^N IPD_{mean}[i]}{N}$$

[0225] 为了评估IPD参数的稳定性, 计算 IPD_{mean} 与 $IPD_{mean_long_term}$ 之间的距离(IPD_{dist}), 这示出在最后N个帧过程中的IPD的演进。在优选实施例中, 局部IPD与长期IPD之间的距离被计算为局部平均值与长期平均值之间的差的绝对值:

[0226] $IPD_{dist} = \text{abs}(IPD_{mean} - IPD_{mean_long_term})$

[0227] 可以看出, 如果 IPD_{mean} 参数在先前帧上是稳定的, 则距离 IPD_{dist} 变得接近0。当相位差随着时间推移稳定时, 距离随后等于零。此距离对声道的相似性给出较好估计。

[0228] 在第六步骤413中, 相干性或ICC参数通过 $ICC = 1 - IPD_{dist}$ 进行计算, 因为ICC和 IPD_{dist} 具有间接的倒数关系。当声道类似时, ICC接近1, 且在这种情况下 IPD_{dist} 变成等于0。

[0229] 在第六步骤413的替代实施形式中, 用以定义ICC与 IPD_{dist} 之间的关系的方程式被定义为 $ICC = d - e \cdot IPD_{dist}$, 其中参数d和e经选择以更好地表示两个参数ICC与 IPD_{dist} 之间的倒数关系。在第六步骤413的另一实施形式中, ICC与 IPD_{dist} 之间的关系是通过在较大数据库上训练而获得, 并且可以被一般化为 $ICC = f(IPD_{dist})$ 。

[0230] 在音频信号(例如, 对于语言信号)的相关段过程中 IPD_{dist} 较小, 并且在音频输入(例如, 对于音乐信号)的扩散部分过程中, 此 IPD_{dist} 参数变得大得多, 并且如果输入声道是去相关的, 则 IPD_{dist} 参数将接近1。因此ICC和 IPD_{dist} 具有间接的倒数关系。

[0231] 根据上文,提供关于录音媒体以及类似者的各种方法、系统、计算机程序对于所属领域的技术人员而言是显而易见的。

[0232] 本发明还支持包括计算机可执行代码或者计算机可执行指令的计算机程序产品,其在执行时引起至少一个计算机执行本文中所描述的执行和计算步骤。

[0233] 本发明还支持用于执行本文中所描述的执行和计算步骤的系统。

[0234] 根据上述教示,许多替代方式、修改和变型对于所属领域的技术人员而言将是显而易见的。当然,所属领域的技术人员容易认识到,除了本文中所描述的那些应用之外,存在许多本发明的应用。尽管已参考一个或多个特定实施例描述本发明,但是所属领域的技术人员将认识到,在不脱离本发明的精神和范围的情况下可以对本发明作出许多修改。因此,应理解,在所附权利要求书及其等效物的范围内,可以不同于如本文中所具体描述的方式实践本发明。

[0235] 本发明的对应实施例可以应用于ITU-T G.722、G.722附录B、G.711.1和/或G.711.1附录D的立体声扩展的编码器中。此外,所描述的方法还可以应用于如3GGP EVS(增强语音服务)编解码器中定义的移动应用的语言和音频编码器。

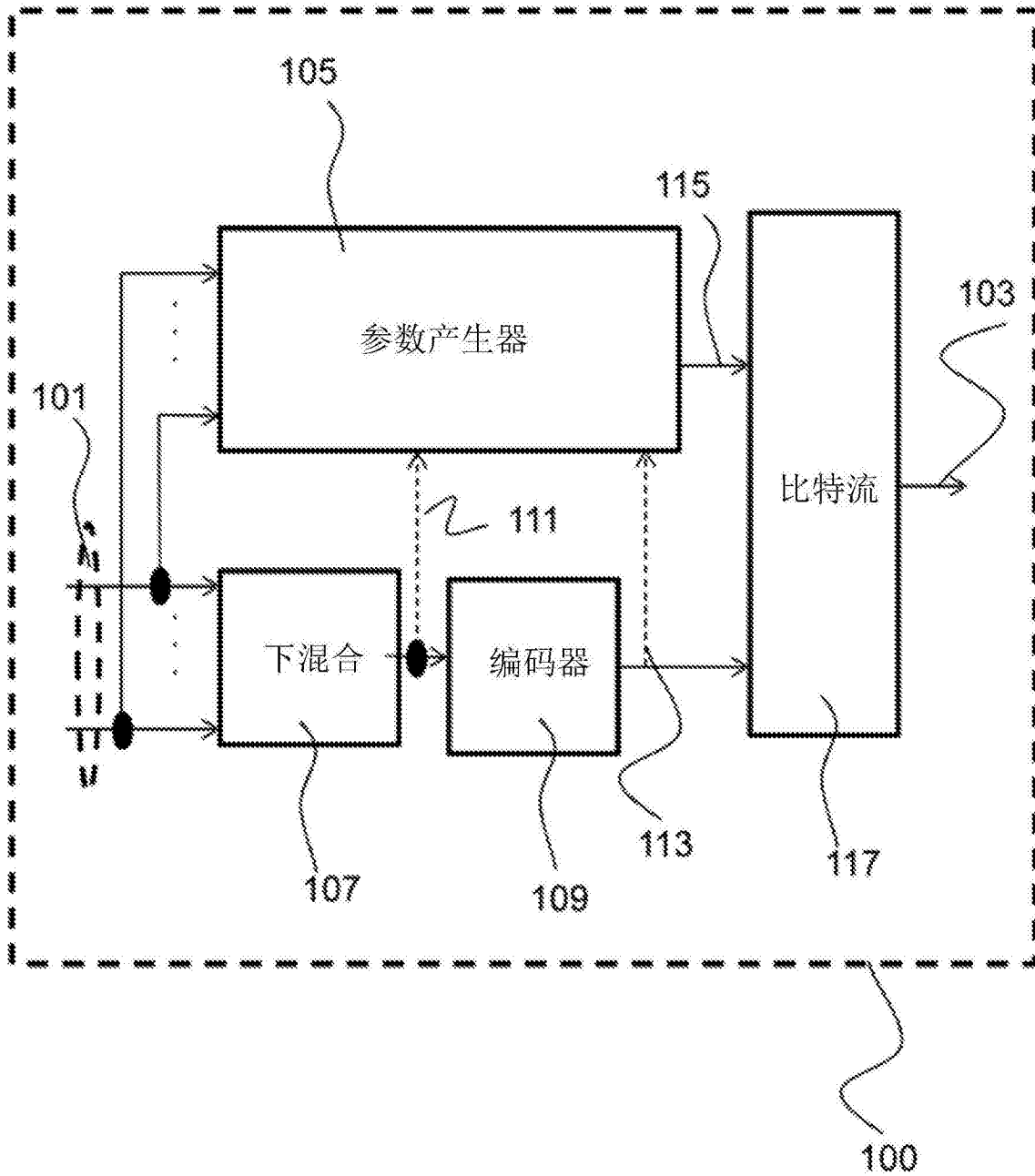


图1

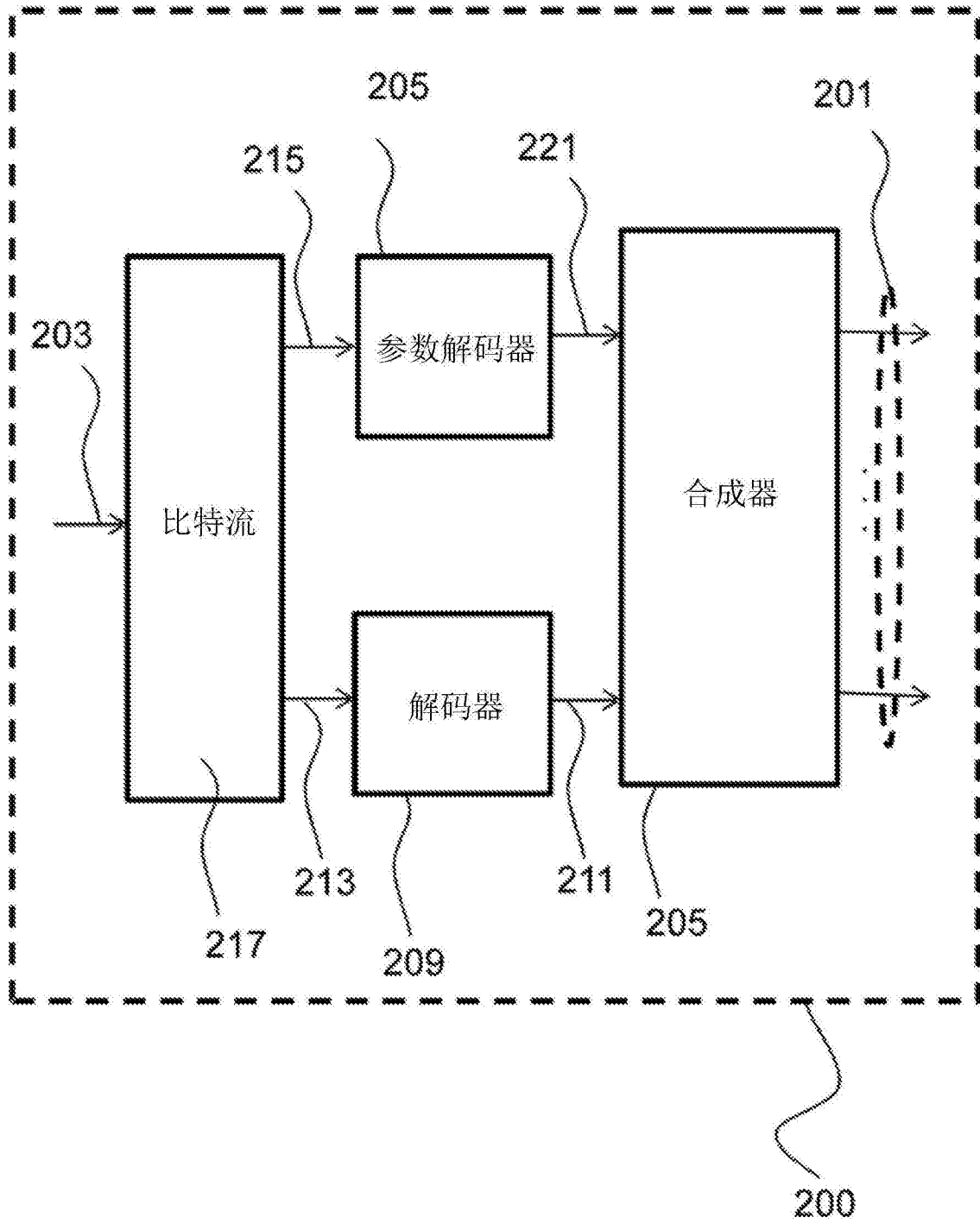


图2

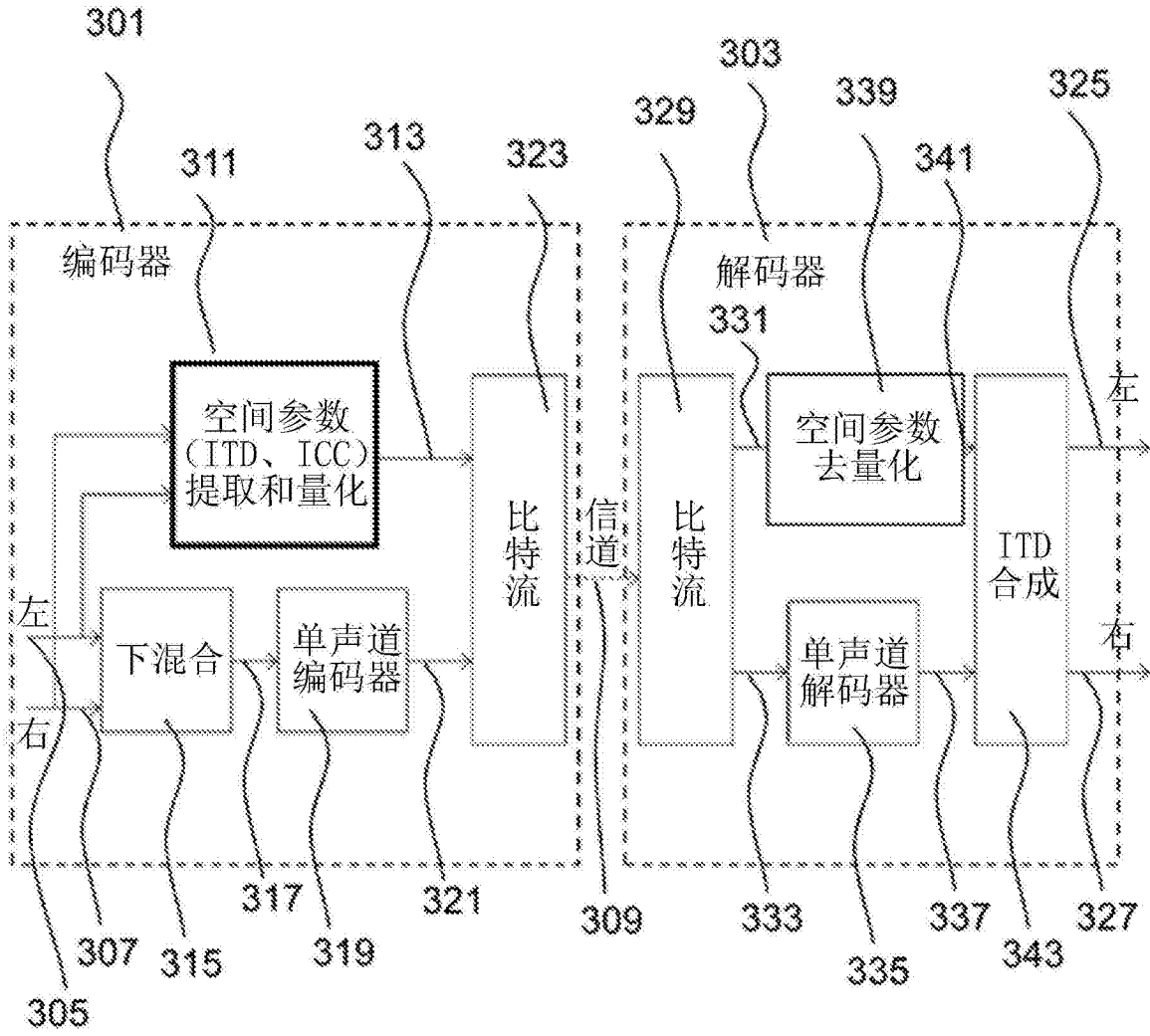


图3

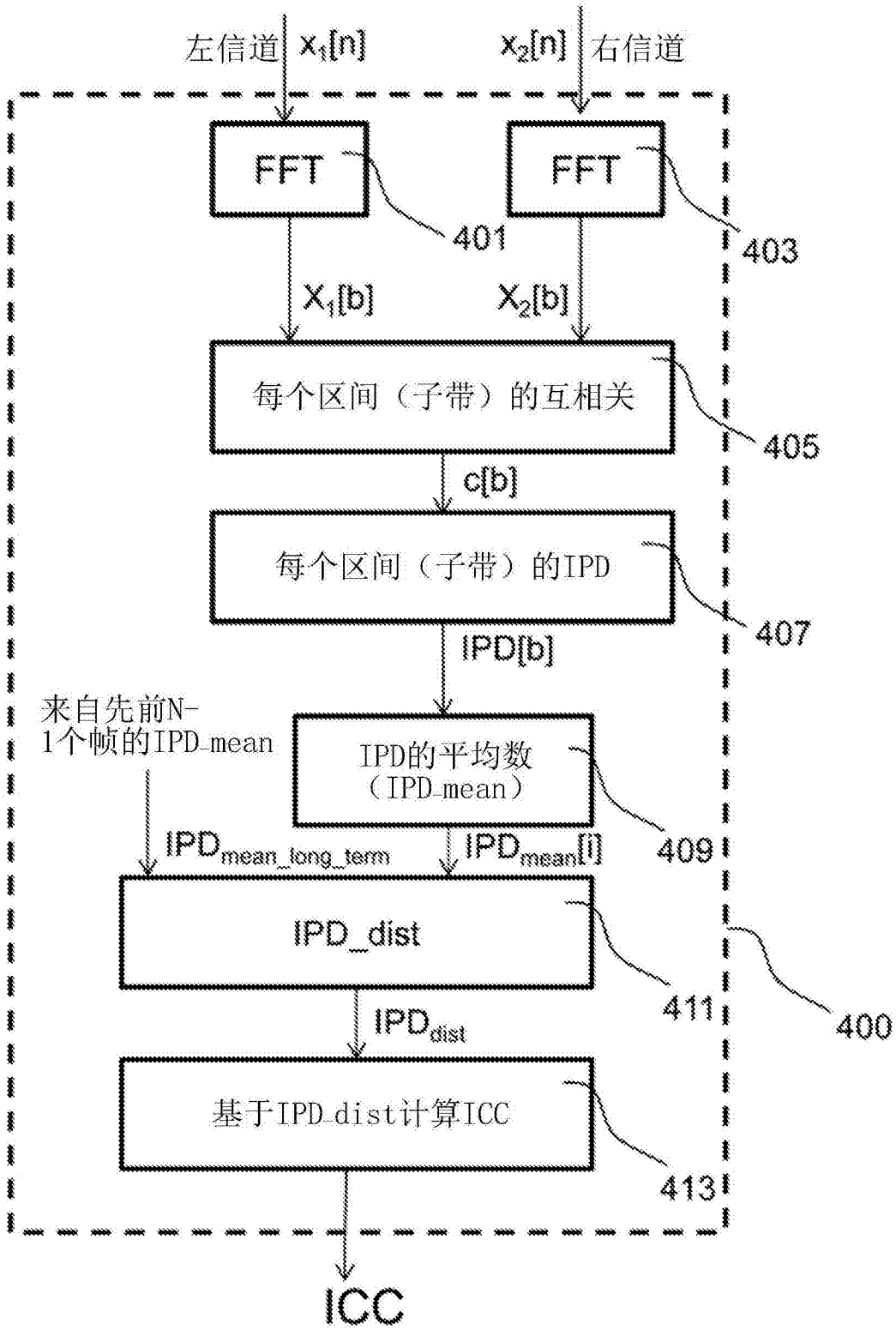


图4