



(12) 发明专利

(10) 授权公告号 CN 1922657 B

(45) 授权公告日 2012.04.25

(21) 申请号 200580005256.0

(22) 申请日 2005.02.02

(30) 优先权数据

04100664.4 2004.02.19 EP

(85) PCT申请进入国家阶段日

2006.08.17

(86) PCT申请的申请数据

PCT/IB2005/050437 2005.02.02

(87) PCT申请的公布数据

W02005/083684 EN 2005.09.09

(73) 专利权人 NXP 股份有限公司

地址 荷兰艾恩德霍芬

(72) 发明人 阿夫尼什·马和施瓦里

维德·克鲁伊吉泽

(74) 专利代理机构 中科专利商标代理有限责任公司
11021

代理人 王波波

(51) Int. Cl.

G10L 19/14 (2006.01)

(56) 对比文件

US 5638069 A, 1997.06.10, 全文.

US 6081783 A, 2000.06.27, 摘要, 第2栏第26-33行, 第7栏第52-第10栏第13行, 第70栏第45行-第71栏第5行, 第72栏第59行-第73栏第8行及图1, 2, 6-15.

STEVE VERNON. DOLBY DIGITAL: AUDIO CODING FOR DIGITAL TELEVISION AND STORAGE APPLICATIONS. AES 17TH INTERNATIONAL CONFERENCE. HIGH QUALITY AUDIO CODING. . 1999, 40-57.

VERNON S. DESIGN AND IMPLEMENTATION OF AC-3 CODERS. IEEE TRANSCATION ON CONSUMER ELECTRONCS, IEEE INC. NEW YORK, US413. 1995, 41(3), 754-759.

SAKAMOTO H ET AL. A DOLBY AC-3/MPEG1 AUDIO DECODER CORESUITABLE FOR AUDIO/VISUAL SYSTEM INTEGRATION. PROCEEDINGS OF THE IEEE 1997 CUSTOM INTEGRATED CIRCUITS SYMPOSIUM. 1997, 241-244.

审查员 杨艳兰

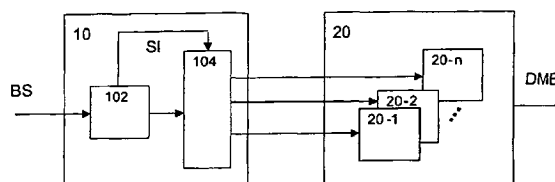
权利要求书 1 页 说明书 5 页 附图 3 页

(54) 发明名称

用于可变块尺寸信号的解码方案

(57) 摘要

本发明涉及一种两步解码方法, 其中首先基于来自比特流的信息的子集, 计算或确定媒体块的尺寸。该尺寸信息限定该媒体块的字节数或长度。然后利用该尺寸信息, 从随后的第二媒体块和比特流的剩余部分砍掉或抽取第一媒体块。该步骤与目前的解码步骤相比, 计算或处理量更少。然后可以进行第一媒体块的正常解码, 同时该并行结构的处理元件可以利用第一步中获得的尺寸信息跳至第二媒体块, 而不需要等待第一媒体块的处理结束。这样, 由于该基本结构能够通过同时解码多个块而利用并行性, 所以解码时间减少。



1. 一种解码装置,用于解码包括多个数据块的数据流,所述多个数据块是媒体应用帧的音频块,每一个音频块具有块参数、打包的指数信息、比特分配信息和尾数块,所述装置包括:

- a. 尺寸确定模块 (102),用于处理所述数据流的信息的子集,以确定要被解码的第一音频块的长度;
- b. 分离模块 (104),基于所述确定的长度,从所述数据流中分离所述第一音频块;以及
- c. 并行处理模块 (20),用于当所述第一音频块被解码的同时,处理随后的第二音频块;

其中,所述尺寸确定模块 (102) 用于确定尾数块的尺寸,并且,其中,所述分离模块 (104) 安排用于对第一音频块的副信息进行解包,然后解析指数信息并将其发送至所述并行处理模块 (20) 的第一处理单元,解析比特分配信息并将其发送至所述并行处理模块 (20) 的第二处理单元,以及解析尾数块并将其发送至所述并行处理模块 (20) 的第三处理单元,然后跳至第二音频块。

2. 如权利要求 1 所述的装置,其中所述尺寸确定模块 (102) 适用于产生尺寸信息,并将所述尺寸信息 ($f_sz_of_blk$) 提供至所述分离模块 (104)。

3. 如权利要求 2 所述的装置,其中所述分离模块 (104) 使用所述尺寸信息,以从所述数据流中分离出所述第一音频块。

4. 如前述任一权利要求所述的装置,其中所述尺寸确定模块 (102) 的所述处理是累加处理,用于累加所述第一音频块的尾数块的确定的比特数。

5. 如权利要求 4 所述的装置,其中所述确定的比特数是从比特分配处理中获得的。

6. 如权利要求 5 所述的装置,其中所述比特分配处理基于至少一个心理声学模型,其中将功率谱密度与掩模曲线比较,以揭示所述比特数。

7. 如权利要求 5 或 6 所述的装置,其中所述并行处理模块 (20) 安排用于解析所述数据流的第一帧的比特流信息,然后无需等待所述第一帧中提供的音频块的副信息的解析结束,跳至随后的第二帧的开始。

8. 一种解码数据流的方法,其中该数据流包括多个音频块,所述多个数据块是媒体应用帧的音频块,每一个音频块具有块参数、打包的指数信息、比特分配信息和尾数块,所述方法包括步骤:

处理所述数据流的信息的子集,以确定要被解码的第一音频块的长度;

基于所述确定的长度,从所述数据流中分离出所述第一音频块;

当解码所述第一音频块的同时,处理随后的第二音频块;

安排用于对第一音频块的副信息进行解包;

然后解析指数信息并将其发送至并行处理模块 (20) 的第一处理单元;

解析比特分配信息并将其发送至所述并行处理模块 (20) 的第二处理单元;以及

解析尾数块并将其发送至所述并行处理模块 (20) 的第三处理单元,然后跳至第二音频块。

用于可变块尺寸信号的解码方案

[0001] 本发明涉及一种解码方法和解码装置,用于解码包括多个数据块的数据流。具体地,本发明涉及用于具有可变块尺寸的媒体数据流的音频和/或视频解码方案。

[0002] 数字音频的普及度在稳定增长。越来越多的人在使用压缩数字音频,用于在因特网上交换音乐和音频文件。数字通用磁盘(DVD)、音乐CD、电视和无线广播工业,所有这些都认识到传送具有良好品质的压缩音频的优点。DVD和HDTV(高清晰电视)工业已经承诺提供给他们用户多声道、剧场品质的声音体验。杜比(Dolby)数字编码系统,也公知为杜比AC-3,它是DVD和HDTV广播的音频压缩标准,它严重降低了声道节目的数据速度,例如,从6Mb/s(6声道,20比特,48kHz),降低至384kb/s,这对应于15比1的降低。

[0003] 为了这样的媒体应用,比特流格式由帧结构组成,其中一帧由若干媒体块组成。这些媒体块又包括他们自己的参数和数据。在体系结构世界中,趋势是走向并行处理结构。在这些结构中,目的是从比特流中分离和取得媒体块,并将它们并行馈入该结构的各处理元件。为了达到这个目的,需要识别媒体块的结束端,以至于它们可以彼此分离。为了识别媒体块之间的分离,当前使用两种方法:

[0004] 1、每个媒体块具有明确的分隔符字段,其中在每个媒体块的结束端增加该分隔符字段。这有助于识别一个媒体块的结束端和另一个媒体块的起始端。

[0005] 2、每个媒体块的字节尺寸限制为固定的。由于此时每个媒体块都具有固定的尺寸,所以它可以跳过该固定字节数,从而识别出下一个媒体块的起始端。

[0006] 但是,存在这样的标准,在这些标准中,这些媒体块没有固定的尺寸和任何分隔符字段。这样的标准例如是上述用于DVD和HDTV广播的杜比AC-3标准。在类似的标准中,上述两个方法无效。

[0007] 本发明的目的是提供一种解码方法和装置,由此对于具有可变块尺寸而无分隔符字段的媒体应用,可以实现并行处理结构。

[0008] 通过如权利要求1所述的解码装置和权利要求10所述的方法实现该目的。

[0009] 相应地,由于下述事实解码需要更少的计算或处理:进行第一数据块解码的同时,该并行结构的处理元件可以利用从尺寸确定中获得的块尺寸已经跳到第二数据块,而不需要等待第一数据块的处理结束。这样,由于可以同时解码多个块,使得该基本结构可以开发或利用并行性,所以解码时间减少。

[0010] 尺寸确定模块适用于产生尺寸信息和向分离模块提供尺寸信息。然后分离模块使用该尺寸信息,以从数据流中分离出第一数据块。因此,当根据从同时运行的解码模块之一中获得的反馈信息产生尺寸信息时,可以提供优先的块分离,以通过分离模块跳至第二数据块。

[0011] 尺寸确定模块的处理可以是累加处理,用于累加第一数据块的各预定部分的确定的比特数。

[0012] 具体地,多个数据块可以是诸如AC-3帧的媒体应用帧的音频块,预定部分可以是尾数部分。这样,可以在数据流的预先解析或解码运行期间顺序获得各数据块的长度。可以从比特分配处理中获得所确定的比特数目。该比特分配处理可以基于至少一个心理声学

模型,其中将功率谱密度和掩模曲线(mask curve)作比较,以揭示所述比特数目。

[0013] 并且,该并行处理模块可以安排用于解析数据流的第一帧的比特流信息,然后跳至随后的第二帧的开始,而不需要等待第一帧中提供的音频块的副信息的解析结束。这样,可以在该音频块的解析结束之前,开始第二帧的比特流信息的解析和解码,因此,增加了并行性。

[0014] 此外,该分离模块可以安排用于对第一音频块的副信息进行解包,然后解析指数信息并将其发送至该并行处理模块的第一处理单元,解析比特分配信息并将其发送至该并行处理模块的第二处理单元,解析尾数块并将其发送至该并行处理模块的第三处理单元,然后跳至第二音频块。因此,只是将信息进行解析并发送至各个过程,不需要在跳至块序列的下一音频块之前,等待这些过程完成。

[0015] 进一步有利的修改在从属权利要求中定义。

[0016] 本发明现在将在优选实施例的基础上,参照附图进行描述,其中:

[0017] 图1表示可以应用本发明的媒体应用的帧的典型比特流结构;

[0018] 图2表示根据本发明的两步解码方案的示意性结构图;

[0019] 图3表示典型的杜比数字解码方案的示意性流程图;以及

[0020] 图4表示根据优选实施例的杜比数字解码过程的示意性功能图。

[0021] 现在基于杜比数字解码器,即杜比 AC-3 音频解码器描述优选实施例。

[0022] 在过去的几年中,数字音频数据压缩已经变成音频工业中重要的技术。杜比 AC-3 是灵活的音频数据压缩技术,可以将一定范围的音频声道格式编码为低速率比特流。AC-3 技术的起源来自,希望提供用于高清晰电视(HDTV)的高级多声道声音的局域化。该目标是得到可以被尽可能广泛的观众使用的编码音频。潜在的观众范围从希望拥有完全的声音体验的商业电影院的顾客或家庭剧院的爱好者,至安静的旅馆房间中以很低的音量收听单声道电视但希望听到所有节目内容的居住者。

[0023] 杜比 AC-3 标准接受 PCM(脉冲编码调制)音频作为它的输入,并产生编码比特流。编码过程的第一步是将音频的表示从 PCM 时间采样序列变换至频率系数块序列。将 512 个时间采样的重叠块乘以时间窗,并变换至频域。由于重叠块的原因,每个 PCM 输入采样在两个连续变换块中表示。然后可以因数 2 对该频域表示进行抽样,使得每个块包括 256 个频率系数。在瞬态信号的情况下,通过使用块切换技术实现了性能的提高,该块切换技术计算两个 256 点的变换,代替 512 点的变换。浮点转换过程将变换系数分为指数/尾数对。然后基于参量的比特分配模型,将尾数量化为可变的比特数。将 6 个音频块(1536 个音频采样)的频谱包络(指数)和粗量化尾数格式化为 AC-3 帧。

[0024] 图1表示诸如 AC-3 的媒体应用的典型帧 F 的示意性结构。比特流是这样的帧的序列。如图所示,在该帧结构图中,每个帧包括多个媒体块 MB0-MBn,例如在 AC-3 帧的情况下的音频块。每个媒体块又包括媒体块参数 MBP 和媒体块数据 MBD。并且,每个帧 F 可以包括同步字或模式 SYNC、误差校正码(循环冗余码)CRC#1、比特流信息 BSI,和辅助信息 AUX。

[0025] 在 AC-3 帧的特定情况下,媒体块数据 MBD 包括打包的指数和尾数块。为了提高解码过程的并行性,希望提供一种解析例程或解码例程,适用于跳过该解码计算量繁重的尾数块,并开始解析或解码下一个音频块。为此,解码过程或方案应当能够识别音频或媒体块之间的“分离点”。如上所述,这在传统的媒体标准中,通常通过在这样的媒体块之间加入唯

一可识别的“分隔符字段”或通过具有固定尺寸的媒体块来实现。但是,上述解决方案没有一个可以在没有分离信息的情况下应用于特定的可变尺寸的媒体应用,诸如 AC-3 比特流。

[0026] 根据优选实施例,提出下面的两步或两级解码方法。

[0027] 图 2 表示用于指示根据优选实施例的解码过程或方案的示意性结构图。在第一步或第一级 10 中,通过尺寸确定函数或单元 102,从输入比特流 BS 计算或确定媒体块例如尾数块的尺寸,其中输入比特流 BS 包括例如 PCM 数据。产生相应的尺寸信息 SI,并将它传递至分离函数或单元 104。在分离单元 104 中,然后使用尺寸信息 SI,从剩余的比特流中分割出媒体块,并将分离的媒体块提供至第二级 20 中多个解码处理函数或单元 20-1 至 20-n 中所选择的单元。然后,在选择了解码处理单元 20-1 至 20-n 中,至少执行所提取媒体块的部分并行解码。然后将解码的媒体块 DMB 组合成一个数据流或直接并行提供至第二级 20 的输出端。

[0028] 下面,基于 AC-3 解码过程,给出优选实施例的更详细的描述。

[0029] 图 3 表示典型的 AC-3 解码流程的示意性流程图。在步骤 1 中,典型地,从发送或存储系统输入比特流。下一步骤 2 用于建立帧定位。这包括找到 AC-3 同步字 SYNC,然后确认 CRC 误差检测字指示没有误差。

[0030] 在步骤 3 中,对副信息进行解包,诸如采样速率、帧尺寸、比特速率、声道数、与例如语言码的音频相关的信息、版权等,其中比特流信息 BSI 每帧出现一次,音频块的副信息在每个音频块出现一次,例如每帧 6 次。然后,在步骤 4 中,指数是以编码形式在比特流中传递的。利用来自比特流的副信息对指数解码,并发送至步骤 5 中执行的比特分配例程。该比特分配步骤包括基于心理声学模型的计算,将音频的功率谱密度与掩模曲线作比较。这些计算显示出每个尾数被分配多少个比特。

[0031] 如后面结合优选实施例所解释的,使用获得的比特分配数确定或计算尾数块的尺寸。

[0032] 粗量化的尾数组成了大量 AC-3 数据流。在步骤 6 中,通过分离或抽取由比特分配例程所指示的比特组,而解包尾数数据。然后分组的尾数被解分组。单独编码的尾数值被转换成解量化值。当使用耦合时,在步骤 7 中,利用用于各单独声道的公共耦合声道和耦合坐标,重新构造耦合声道的高频分量。对于每个音频模块,编码器规定其动态范围,并且基于该值,解码器利用该动态范围字改变指数和尾数的幅值。

[0033] 在二声道模式中,如果编码器使用如步骤 8 所示的重新矩阵化,那么在步骤 8 中使用和值以及差值,来抽取左声道和右声道。在步骤 9 的动态范围压缩后,利用步骤 10 中的逆变换,将频域系数转换回时域。在步骤 11 中,对各个时间采样块进行窗口操作,将相邻的块交叠相加在一起,以重新构造最终连续的时域 PCM 音频信号。

[0034] 但是,流中声道的数目可能与用户房屋中扬声器的数目不匹配。在这种情况下,需要如步骤 12 中所示的缩混(downmixing),以混合流中的声道,使得它们能够基于用户房屋中扬声器的数目而再现出来。

[0035] 最后,在步骤 13 中,典型地,以采样速率或以适合与数模转换器(DAC)互连的形式或其他任意形式,将 PCM 输出写入缓冲器。

[0036] 注意,图 3 中所示的一系列步骤只是多种解码 AC-3 音频流的方式中的一种。例如,由于步骤 12 中的缩混是线性操作,它可以在时域或在频域中完成。

[0037] 并且,可以理解,图 3 的流程图隐藏了其中的循环。步骤 1、2、11、12 和 13 在帧的基础上工作,而步骤 3-10 在音频块的基础上迭代。因此,帧 F 的典型解码顺序将意味着对于每一帧 F 执行步骤 1 和 2 一次,然后重复步骤 3-10 的次数为一帧中媒体块 MB 的数目,例如,AC-3 帧中的 6 个音频块,然后在帧的基础上执行步骤 11-13。也意味着顺次执行步骤 3-10。换言之,在对第一个音频块进行步骤 6 时,不能对随后的第二个音频块开始步骤 3。

[0038] 在优选实施例中,提出了一种解决方案,使得在一个过程网络中,能够独立并同时执行对应于步骤 1-13 的过程。

[0039] 图 4 表示根据优选实施例,AC-3 解码器方案的功能过程模型。该模型基于经由先入先出存储器 (fifo)、移位寄存器存储器等而彼此相连的过程的集合。这些过程和 fifo 经由过程的端口连接。

[0040] 在图 4 的功能图中,提出一种技术,用于从图 3 所示的固有串行的 AC-3 解码算法中提取出并行性。在图 4 的功能图中,椭圆表示过程,箭头表示 fifo。应当注意,图 4 没有示出这些过程的所有细节。例如,它没有示出端口过程和分叉过程。如上所述,端口用于连接各过程和 fifo。需要分叉过程来复制标记。这发生在这种情况下:对于一个标记,有一个生产者和多个使用者。那么,分叉过程从它的输入 fifo 读出标记,并将它的副本写入多个输出 fifo。当经由 fifo 在过程之间进行标记通信时,标记代表值,而不是值的引用。这意味着如果两个过程必须共用数据,它们则通过从 fifo 写入和读出数据来显式地共用数据,而不是写入和读出数据的指针。

[0041] 并且,图 4 没有示出代表过程之间 fifo 的箭头的完全列表。为了该图的可读性,在图 4 中描述所有 fifo 箭头是不可行的。

[0042] 在图 4 中,过程 1、2 和 8 分别对应于图 3 的步骤 1、2 和 8,而过程 5、6 和 9 至 12 对应于图 3 的步骤 4、5 和 10 至 13。图 3 的“解包 BSI,副信息”步骤(步骤 3)分为两个过程 3 和 4,即“unp_bsi_info”(过程 3)和“unp_audio_info”(过程 4)。并且,图 3 的步骤 6、7 和 9 即“解包、解分组、解量化、抖动尾数(dither mantissa)”、“解耦合”和“动态范围压缩”合并为一个过程 7,即过程“decode_mants”。下面解释这些原因。

[0043] 取代图 4 的两个过程 3 和 4,在图 3 中有一个步骤 3,该步骤覆盖了“解包 BSI,副信息过程”的功能。在图 3 的传统方案中,以下述方式解析 AC-3 帧。首先,比特流信息(BSI)被解码或解析,这在每个帧的帧头处出现一次。然后,第一个音频块的副信息被解析或解码。为了开始解析下一帧的 BSI,步骤 3 首先必须完成所有音频块的副信息的解析或解码。为了增加并行性,提出了解析或解码第一帧的 BSI,然后跳至下一帧的开始。这样,不需要等待第一帧的音频块解析结束,就可以开始第二帧的 BSI 的解析或解码。

[0044] 根据图 4 的优选实施例,图 3 的步骤 3 分为两个过程 3 和 4,其中在过程 3 中对帧的 BSI 信息进行解包,在过程 4 中对音频块的音频副信息进行解包。这样,过程 3 在帧的基础上工作,并只解析或解码每个帧的 BSI 信息,而帧的其余部分传递至过程 4,其中过程 4 在音频块的基础上工作,并解析包含在每个音频块中的副信息。

[0045] 根据 AC-3 帧结构,每个 AC-3 帧包括六个音频块。每个音频块又包括参数、打包指数和一个尾数块。因此,如上所述,希望跳过该尾数块并开始下一个音频块的解析。为此,必须在尾数块之间识别“分离点”。为了解决该问题,使用图 2 的两步解码方法。换言之,图 4 的过程 4(解包音频副信息)首先对第一音频块的副信息进行解包,然后解析编码的指数

并将其发送至过程 5(解码指数),解析比特分配数据并将其发送至过程 6(比特分配),解析压缩的尾数块并将其发送至过程 7(解码尾数),然后对于第二个音频块重复该流程。关键点是:只是解析信息并将其发送至相应的过程,然后不需要等待由其他过程 5 至 7 完成第一个音频块的处理,就跳至随后的第二个音频块。

[0046] 上述并行流程需要已知压缩尾数块的尺寸。为了克服该运算障碍,提出了利用过程 6。利用心理声学模型,对于第一个尾数,过程 6 确定应当从尾数块中剥离出多少个比特。它将该信息存储在可变的、所谓的比特分配指针(BAP)中。然后过程 7 使用该 BAP 为第一个尾数从压缩尾数块中剥离出比特。该尾数被解码,并储存在数组中,用于进一步的处理。接下来,计算第二个尾数的 BAP,过程 7 使用该 BAP,以从比特流的压缩尾数块中剥离出比特。对于该第一音频块中存在的所有声道的所有尾数,找到或获得 BAP,然后使用该 BAP 从比特流中剥离出比特的过程被重复进行。当从比特流中剥离出第一音频块的所有尾数时,可以顺次进行第二音频块或下一音频块的解析或解码。

[0047] 但是,如果将第一音频块的所有 BAP 加在一起,那么该总和将代表第一音频块的压缩尾数块的尺寸。因此,技巧是将该确定的或计算出的 BAP 的总和经由 fifo “f_size_of_blk”(图 4 中的虚线箭头)发送至过程 4,使得过程 4 能够从比特流中“砍掉”对应于该 BAP 总和的比特数,并将该压缩尾数块发送至过程 7。这样,过程 4 无需等待第一音频块的处理结束,就可以开始解析第二音频块。

[0048] 在上述方法中,每个过程只等待其完成计算所需要的充分必要信息。附带地,这也是在如 YAPI 的抽象层次上的算法操作如何节省大量周期的很好示例。再参照图 2,可以理解,图 2 的第一级 10 对应于图 4 的过程 4,而图 2 的第二级 20 基本对应于过程 5 至 7。

[0049] 总之,提出了两步解码方法,其中,首先基于比特流中信息的子集,计算或确定媒体块的尺寸。该尺寸信息限定了媒体块的字节数或尺寸。然后利用该尺寸信息,从随后的第二个媒体块和比特流的剩余部分砍掉或抽取出第一个媒体块。该步骤与目前的解码步骤相比,计算或处理量更少。然后进行第一媒体块的正常解码,同时该并行结构的处理元件可以利用第一步中获得的尺寸信息跳至第二媒体块,而不需要等待第一媒体块的处理结束。这样,由于该基本结构能够通过同时解码多个块而利用并行性,所以解码时间减少。

[0050] 注意,本发明不限制于上述优选的 AC-3 实施例,它可以在处理可变尺寸块的任意解码装置或方法中实现。具体地,可以使用比特流信息的任何适合子集来计算或导出任意类型块的尺寸,使得能够至少部分地同时或并行处理在随后的块中提供的信息。这样,各种优选的实施例可以在后附的权利要求保护范围内变化。

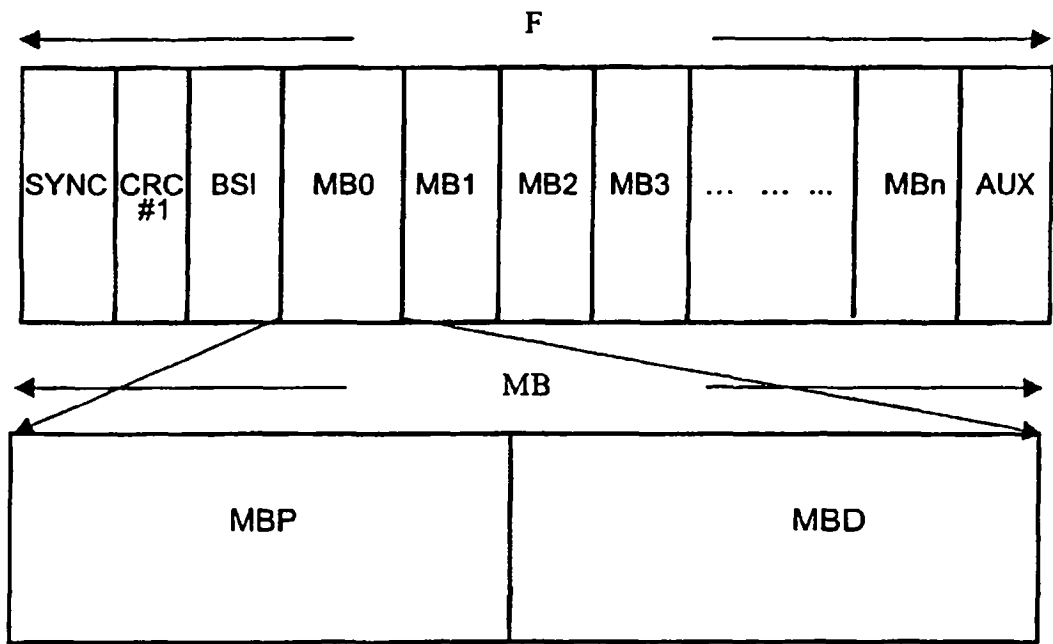


图1

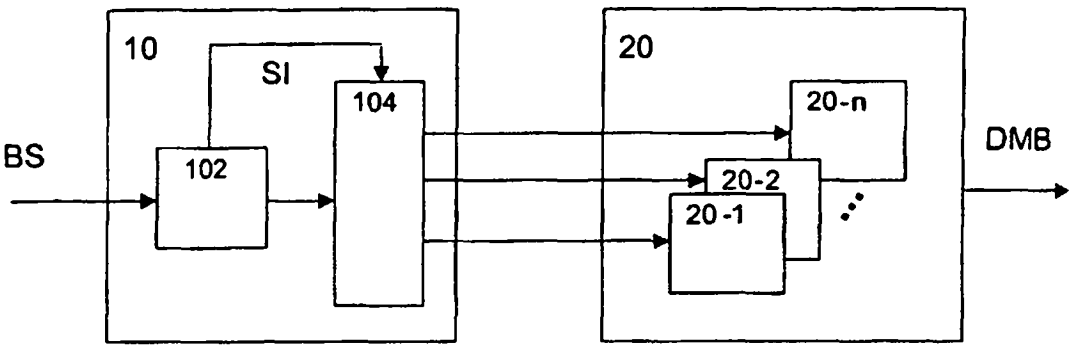


图2

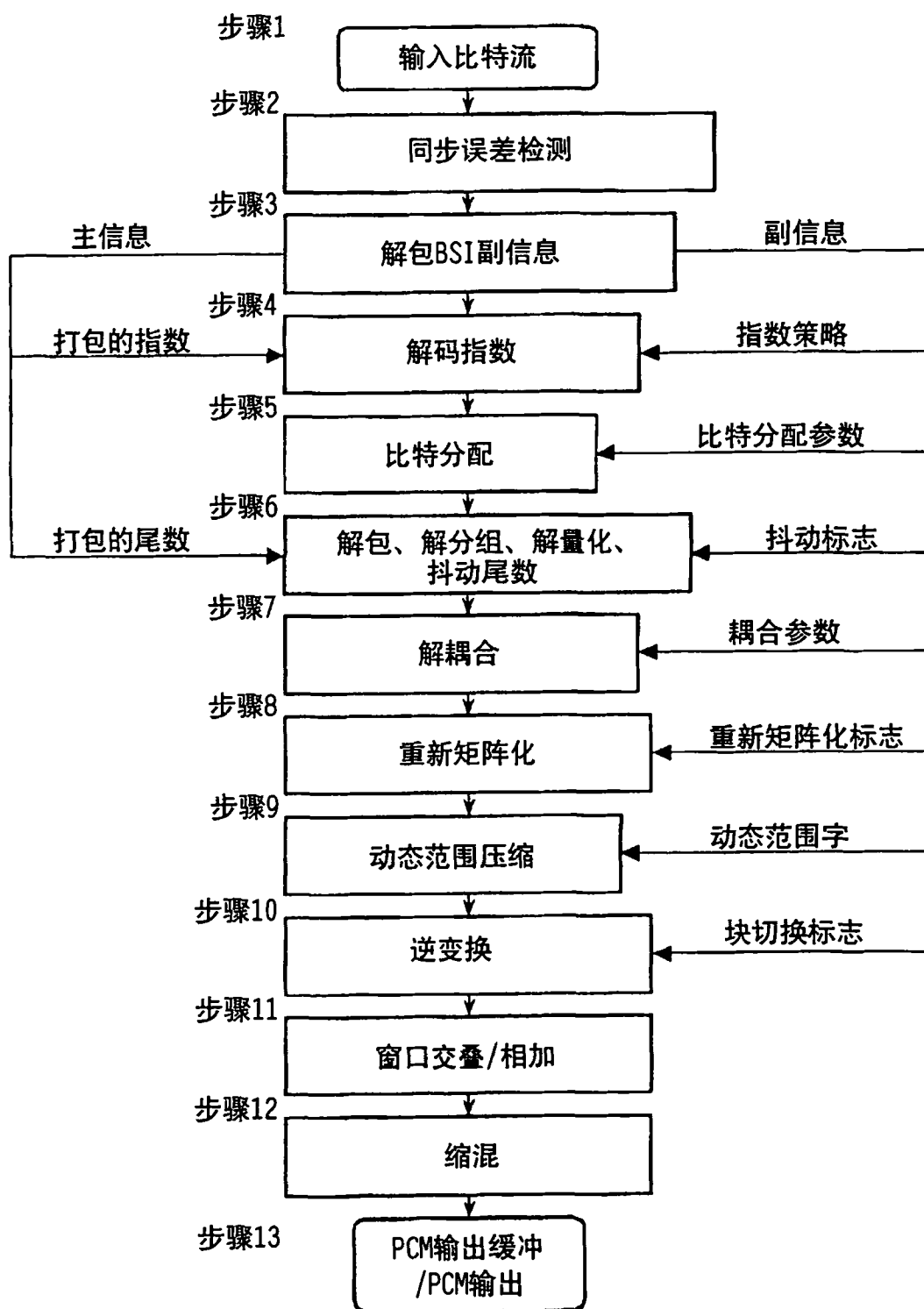


图3

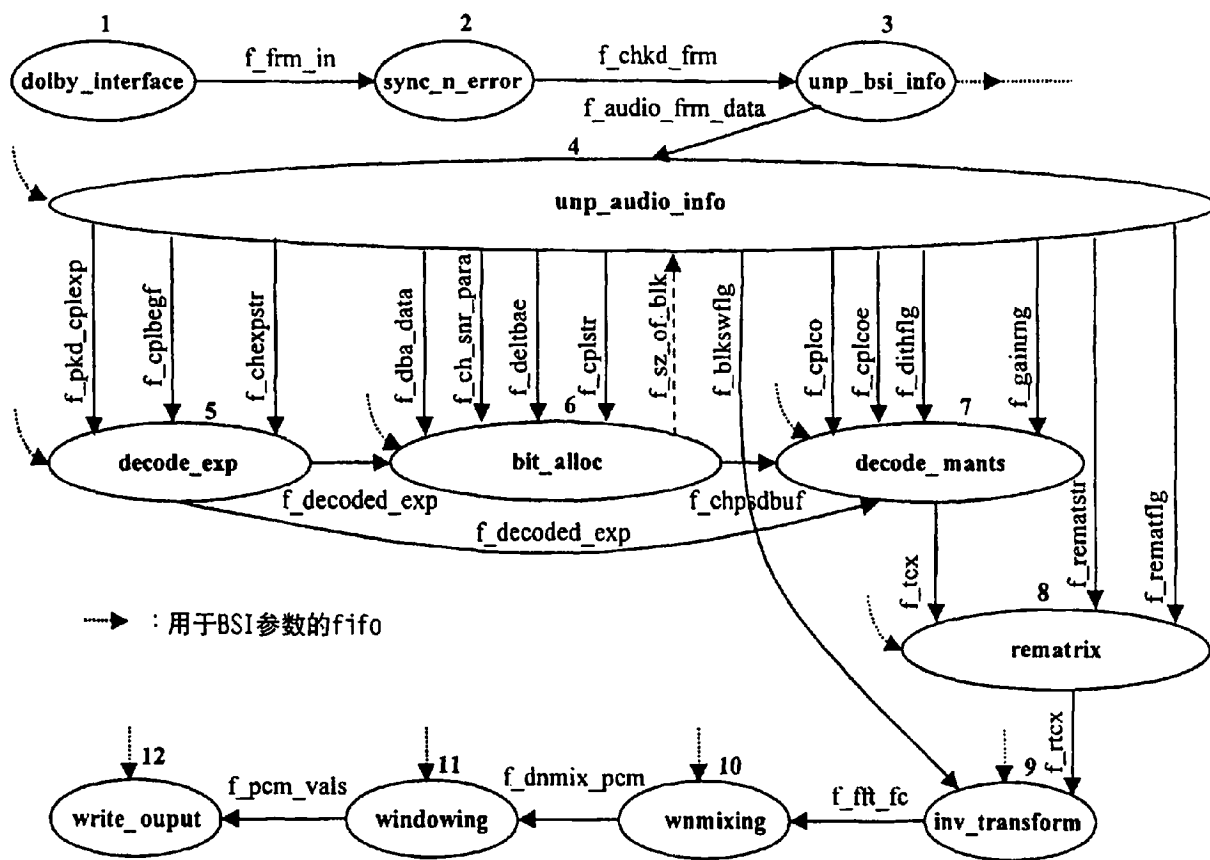


图4