[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06K 9/36 (2006.01)
G06K 9/46 (2006.01)

[12] 发明专利申请公布说明书

[21] 申请号 200480025588.0

[43] 公开日 2007 年 4 月 18 日

[11] 公开号 CN 1950832A

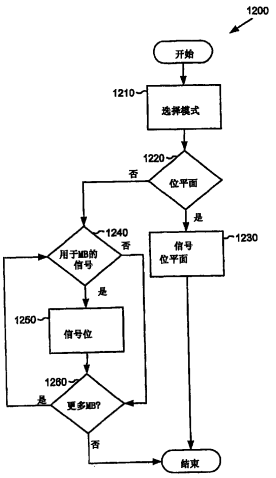
[22] 申请日 2004.9.3
[21] 申请号 200480025588.0
[30] 优先权
[32] 2003. 9. 7 [33] US [31] 60/501,081
[32] 2004. 9. 2 [33] US [31] 10/933,959
[86] 国际申请 PCT/US2004/029032 2004.9.3
[87] 国际公布 WO2005/027494 英 2005.3.24
[85] 进入国家阶段日期 2006.3.6
[71] 申请人 微软公司
地址 美国华盛顿州
[72] 发明人 P·苏 S·斯里尼瓦杉
T·W·赫尔科比 K·慕克吉
林志隆
[74] 专利代理机构 上海专利商标事务所有限公司
代理人 张政权
权利要求书 5 页 说明书 33 页 附图 22 页

[54] 发明名称

用于 AC 预测状态以及宏块半帧/帧编码类型信息的位平面编码和解码

[57] 摘要

一方面，编码器/解码器从多个可用位平面模式组中选择一种位平面模式，并根据选定的位平面模式处理位平面，其中该位平面指示视频图片的多个宏块的 AC 预测状态信息。 另一方面，对于视频序列中的第一隔行扫描视频帧，解码器解码第一隔行扫描视频帧的帧层处用信号表示的位平面。 位平面表示第一隔行扫描视频帧的多个宏块的半帧/帧变换类型。 对于视频序列中的第二隔行扫描视频帧，对于第二隔行扫描视频帧的多个宏块的至少一个但非全部之中的每一个，解码器处理宏块层处用信号表示的每宏块的半帧/帧变换类型比特。



1.一种方法，包括：

对于视频序列中的第一隔行扫描视频帧，解码在第一隔行扫描视频帧的帧层处用信号表示的位平面，其中所述位平面表示用于第一隔行扫描视频帧的多个宏块的半帧/帧变换类型；以及

对于视频序列中的第二隔行扫描视频帧，对于所述第二隔行扫描视频帧的多个宏块的至少一个但非全部之中的每一个，处理在宏块层处用信号表示的每宏块的半帧/帧变换类型比特。

2.如权利要求 1 所述的方法，其特征在于，所述第二隔行扫描视频帧包括一个或多个帧内宏块以及一个或多个帧间宏块，其中所述每宏块的半帧/帧变换类型比特对所述一个或多个帧内宏块中的每一个而不对所述一个或多个帧间宏块用信号表示。

3.如权利要求 2 所述的方法，其特征在于，不同的比特流元素表示用于一个或多个帧间宏块中的每一个的半帧/帧变换类型。

4.如权利要求 1 所述的方法，其特征在于，所述第一隔行扫描视频帧是隔行扫描 I 帧，而所述第二隔行扫描视频帧是隔行扫描 P 帧。

5.如权利要求 1 所述的方法，其特征在于，解码所述位平面包括从多个可用解码模式中选择一个解码模式。

6.如权利要求 5 所述的方法，其特征在于，所述多个可用解码模式包括行跳过、列跳过、成对 VLC、六个一组 VLC 和一个或多个差分模式。

7.如权利要求 1 所述的方法，其特征在于，还包括：

对于第一隔行扫描视频帧，在解码位平面后，对于第一隔行扫描视频帧的多个宏块中的每一个，至少部分基于为位平面中的宏块用信号表示的半帧/帧变换类型来处理用于宏块的宏块层信息。

8.一种其上存储有计算机可执行指令的计算机可读媒体，所述指令用于使得通过其编程的计算机系统执行如权利要求 1 所述的方法。

9.一种方法，包括：

对于视频序列中的第一隔行扫描视频帧，编码位平面并在第一隔行扫描视

频帧的帧层处用信号表示位平面，其中所述位平面表示用于第一隔行扫描视频帧的多个宏块的半帧/帧变换类型；以及

对于视频序列中的第二隔行扫描视频帧，对于第二隔行扫描视频帧的多个宏块的至少一个而非全部之中的每一个，在宏块层处用信号表示每宏块的半帧/帧变换类型比特。

10.如权利要求 9 所述的方法，其特征在于，所述第二隔行扫描视频帧包括一个或多个帧内宏块以及一个或多个帧间宏块，其中所述每宏块的半帧/帧变换类型比特对所述一个或多个帧内宏块中的每一个而不对所述一个或多个帧间宏块用信号表示。

11.如权利要求 10 所述的方法，其特征在于，不同的比特流元素表示用于所述一个或多个帧间宏块中的每一个的半帧/帧变换类型。

12.如权利要求 9 所述的方法，其特征在于，所述第一隔行扫描视频帧是隔行扫描 I 帧，而所述第二隔行扫描视频帧是隔行扫描 P 帧。

13.如权利要求 9 所述的方法，其特征在于，编码位平面包括从多个可用编码模式中选择一个编码模式。

14.如权利要求 13 所述的方法，其特征在于，所述多个可用编码模式包括行跳过、列跳过、成对 VLC、六个一组 VLC 以及一个或多个差分模式。

15.一种其上存储有计算机可执行指令的计算机可读媒体，所述指令用于使得通过其编程的计算机系统执行如权利要求 9 所述的方法。

16.一种解码器，包括：

用于解码在帧层处用信号表示的位平面的装置，其中所述位平面包括用于第一隔行扫描视频帧的第一位平面，其中所述第一位平面表示用于第一隔行扫描视频帧的多个宏块的半帧/帧变换类型；以及

用于处理在宏块层处用信号表示的每宏块的半帧/帧变换类型比特的装置，其中所述每宏块的半帧/帧变换类型比特包括用于第二隔行扫描视频帧的多个宏块的一个或多个而非全部之中的每一个的每宏块的半帧/帧变换类型比特。

17.如权利要求 16 所述的解码器，其特征在于，所述第二隔行扫描视频帧是运动补偿帧，其中用于处理每宏块的半帧/帧变换类型比特的装置处理用于第二隔行扫描视频帧的帧内宏块而非帧间宏块的宏块半帧/帧变换类型比特。

18.如权利要求 16 所述的解码器，其特征在于，所述第一隔行扫描视频帧是隔行扫描 I 帧，而所述第二隔行扫描视频帧是隔行扫描 P 帧。

19.如权利要求 16 所述的解码器，其特征在于，还包括：

用于从多个可用位平面解码模式中选择一个位平面解码模式的装置。

20.如权利要求 19 所述的解码器，其特征在于，所述多个可用编码模式包括行跳过、列跳过、成对 VLC、六个一组 VLC 和一个或多个差分模式。

21.一种方法，包括：

从多个可用位平面模式的组中选择一种位平面模式；以及

根据选定的位平面模式处理位平面，其中所述位平面指示用于视频图片的多个宏块的 AC 预测状态信息。

22.如权利要求 21 所述的方法，其特征在于，所述 AC 预测状态信息为所述多个宏块中的每一个指示是否用 AC 预测来编码所述宏块。

23.如权利要求 21 所述的方法，其特征在于，所述视频图片是逐行扫描帧内帧、隔行扫描帧内帧或隔行扫描帧内半帧。

24.如权利要求 21 所述的方法，其特征在于，还包括：

对于第二视频图片，对于所述第二视频图片的多个宏块的至少一个而非全部之中的每一个，处理在宏块层处用信号表示的 AC 预测状态比特。

25.如权利要求 21 所述的方法，其特征在于，所述第二视频图片包括一个或多个帧内宏块和一个或多个帧间宏块，其中每宏块的 AC 预测状态比特对所述一个或多个帧内宏块中的每一个而不用于一个或多个帧间宏块中的任一个用信号表示。

26.如权利要求 21 所述的方法，其特征在于，所述多个可用位平面编码/解码模式包括行跳过、列跳过、成对 VLC、六个一组 VLC 和一个或多个差分模式。

27.如权利要求 21 所述的方法，其特征在于，所述处理包括编码。

28.如权利要求 21 所述的方法，其特征在于，所述处理包括解码。

29.一种其上存储有计算机可执行指令的计算机可读媒体，所述指令用于使得通过其编程的计算机系统执行如权利要求 21 所述的方法。

30.一种方法，包括：

编码一位平面，所述位平面指示用于视频图片的多个宏块的 AC 预测状态信息；以及

用信号表示所述被编码的位平面。

31.如权利要求 30 所述的方法，其特征在于，所述位平面包括用于所述多个宏块中的每一个的一个 AC 预测状态比特，以指示是否使用 AC 预测来编码该宏块。

32.如权利要求 30 所述的方法，其特征在于，在比特流中在半帧或帧层处用信号表示所述编码的位平面。

33.如权利要求 30 所述的方法，其特征在于，还包括：

对于第二视频图片的一个或多个帧内宏块中的每一个，在宏块层处用信号表示每宏块的 AC 预测状态比特。

34.如权利要求 33 所述的方法，其特征在于，所述第二视频图片是运动补偿图片。

35.如权利要求 33 所述的方法，其特征在于，所述第二视频图片是帧内编码图片。

36.一种其上存储有计算机可执行指令的计算机可读媒体，所述指令用于使得通过其编程的计算机系统执行如权利要求 30 所述的方法。

37.一种方法，包括：

接收编码的位平面；以及

解码所述位平面，其中所述位平面指示用于视频图片的多个宏块的 AC 预测状态信息。

38.如权利要求 37 所述的方法，其特征在于，所述位平面包括用于所述多个宏块中的每一个的一个 AC 预测状态比特，以指示所述宏块是否用 AC 预测进行编码。

39.如权利要求 37 所述的方法，其特征在于，在比特流中在半帧或帧层处用信号表示所述编码的位平面。

40.如权利要求 37 所述的方法，其特征在于，还包括：

对于第二视频图片的一个或多个帧内宏块中的每一个，在宏块层处接收每宏块的 AC 预测状态比特。

41.如权利要求 40 所述的方法，其特征在于，所述第二视频图片是运动补偿图片。

42.如权利要求 40 所述的方法，其特征在于，所述第二视频图片是帧内编码图片。

43.一种其上存储有计算机可执行指令的计算机可读媒体，所述指令用于使得通过其编程的计算机系统执行如权利要求 37 所述的方法。

用于 AC 预测状态以及宏块半帧/帧编码类型信息的位平面编码和解码

版权许可

本专利文献的一部分揭示内容包含受版权保护的材料。版权所有人不反对任一专利揭示内容的复制，如其在专利和商标局的专利文件或记录中所显现的，但其他情况则保留全部版权。

技术领域

描述了用于视频编码和解码的技术和工具。例如，编码器位平面编码 AC 预测状态信息。作为另一示例，一编码器用信号表示用于隔行扫描帧编码图片中的宏块的半帧/帧变换类型。一解码器执行相应的解码。

背景

数字视频消耗大量存储和传输容量。典型的原始数字视频序列包括每秒 15 或 30 个图片。每张图片可包括几万或几十万个像素（也称象素）。每个像素都表示图片的微小元素。按原始形式，计算机一般用 24 个比特或以上来表示像素。因此，典型原始数字视频序列的每秒比特数或比特率可以是 5 百万比特/秒或以上。

多数计算机或计算机网络缺乏资源来处理原始数字视频。为此，工程师使用压缩（也称作译码或编码）来降低数字视频的比特率。压缩可以是无损的，其中不损害视频质量而是通过视频的复杂性限制比特率的减小。或者，压缩可以是有损的，其中损害视频质量而比特率的减小更惊人。解压缩和压缩正好相反。

一般，视频压缩技术包括“帧内”压缩和“帧间”或预测压缩。对于视频帧，帧内压缩技术压缩个别帧，通常称作 I 帧或关键帧。帧间压缩技术参考先前的和/或后面的帧对帧进行压缩，且帧间压缩的帧通常被称作预测帧、P 帧或 B 帧。

I.Windows Media Video 版本 8 和 9 中的帧间和帧内压缩

微软公司的 Windows Media Video 版本 8[WMV8]包括一视频编码器和一视频

解码器。WMV8 编码器使用帧内和帧间压缩，且 WMV8 解码器使用帧内和帧间解压缩。Windows Media Video 版本 9[WMV9]使用类似的架构用于许多操作。

A.帧内压缩

图 1A 示出了 WMV8 编码器内关键帧中的像素块 105 的基于块的帧内压缩 100。块是一组像素，例如一 8×8 像素排列。WMV8 编码器将关键视频帧分成 8×8 的像素块并将 8×8 离散余弦变换[DCT]110 应用于诸如块 105 的各块。DCT 是一种频率变换类型，它将 8×8 的像素块（空间信息）转换成作为频率信息的 DCT 系数 115 的 8×8 块。DCT 操作自身是无损或接近无损的。但与原始像素值相比，DCT 系数使编码器更有效地压缩，因为多数有效信息被集中于低频系数（常规地，块 115 的左上部分）且许多高频系数（常规地，块 115 的右下部分）具有零值或接近于零。

随后，编码器量化 DCT 系数（120），形成量化 DCT 系数 125 的 8×8 块。例如，编码器向每个系数应用均匀的标量量化步长。量化是有损的。由于低频 DCT 系数趋于具有较高的值，量化造成精度损失但不造成用于系数的信息的损失。另一方面，由于高频 DCT 系数趋于具有零值或接近于零，所以高频系数的量化通常形成零值的连续区域。此外，在一些情况下，与低频 DCT 系数相比，更粗糙地量化高频 DCT 系数，对高频 DCT 系数造成精度/信息的更大损失。

随后，编码器为熵编码准备量化 DCT 系数 125 的 8×8 块，熵编码是一种无损压缩形式。熵编码的准确类型可根据系数是 DC 系数（最低频率）、上行或左列中的 AC 系数（其它频率）还是其它 AC 系数而变化。

编码器编码 DC 系数 126，作为与邻近的 8×8 块的 DC 系数的差，它是正被编码的块的先前编码的近邻（例如，上面或左面）。（图 1A 示出了位于帧中正被编码的块的左面的近邻块 135）。编码器熵编码（140）该差。

熵编码器可编码 AC 系数的左列或上行，作为与邻近 8×8 块的相应左列或上行的差。这是 AC 系数预测的示例。图 1A 示出了作为与邻近（实际上，左面）块 135 的左列 137 的差 147 被编码的 AC 系数的左列 127。差分编码增加了差分系数具有零值的机会。其余 AC 系数来自于量化 DCT 系数的块 125。

图 1B 示出了用于 I 帧内的 8×8 块的 AC 预测候选。对于上预测，上邻近块 175 的 AC 系数的上行 177 被用作量化 DCT 系数的块 125 中的 AC 系数的上行 129

的预测值。对于左预测，左邻近块 135 中 AC 系数的最左列 137 被用作为块 125 中 AC 系数的最左列的预测值。

在一些模式中，在差分值的计算或与差分值组合之前，AC 系数预测值被缩放或另外处理。

如果邻近块不存在于规定的预测方向中，则将最左列或上行中的所有七个 AC 系数的预测值设定为零。例如，如果预测方向是向上且当前块处于顶行中，则当前块的顶行中的每个预测 AC 系数都被设定为零，因为在向上方向中不存在邻近的块。将预测行或列中的 AC 系数添加到当前块中的相应解码 AC 系数（它们是差分值），以形成完全重建的量化变换系数块。

编码器将量化 AC DCT 系数的 8×8 块 145 扫描（150）成一维数组 155，随后利用行程编码 160 的变化熵编码所扫描的 AC 系数。编码器从一个或多个运行/级别/最后（run/level/last）表 165 中选择一熵码并输出该熵码。

B.帧间压缩

WMV8 编码器中的帧间压缩使用基于块的运动补偿预测编码，继之以残差的变换编码。图 2 和 3 示出了 WMV8 编码器中预测帧的基于块的帧间压缩。特别是，图 2 示出了对预测帧 210 的运动估计且图 3 示出了用于预测帧的运动补偿块的预测残差的压缩。

例如，图 2 中，WMV8 编码器计算用于预测帧 210 中的宏块 215 的运动矢量。为了计算该运动矢量，编码器在参考帧 230 的查找区域 235 中进行查找。在该查找区域 235 内，编码器比较来自预测帧 210 的宏块 215 和各种候选宏块，以找到作为良好匹配的候选宏块。编码器输出指定用于匹配宏块的运动矢量（熵编码的）的信息。运动矢量相对于运动矢量预测值被差分编码。

在通过将所述差添加到运动矢量预测值来重建运动矢量后，解码器使用该运动矢量通过利用来自参考帧 230 的信息计算宏块 215 的预测宏块，该参考帧是编码器和解码器处可用的先前重建的帧。预测很少是理想的，所以编码器通常编码预测宏块和宏块 215 本身之间的像素差的块（也称作误差或残差块）。

图 3 示出了 WMV8 编码器中误差块 335 的计算和编码示例。误差块 335 是预测的块 315 和原始当前块 325 之间的差。编码器将 DCT 340 应用于误差块 335，形成量化 DCT 系数 355 的 8×8 块。编码器将该 8×8 块 355 扫描（360）成一维数组

365，以便从最低频率到最高频率地将系数一般排序。编码器用行程编码 370 的变化来熵编码所扫描的系数。编码器从一个或多个运行/级别/最后表 375 中选择一熵码并输出该熵码。

图 4 示出了用于帧间编码块的相应解码过程 400 的示例。在图 4 的概要中，解码器解码（410、420）表示预测残差的熵编码的信息，这使用具有一个或多个运行/级别/最后表 415 的可变长度解码 410 和行程解码 420。解码器将存储熵解码信息的一维数组 425 反扫描（430）成二维块 435。解码器反量化和反 DCT（一起，440）数据，形成重建的误差块 445。在分开的运动补偿路径中，解码器使用相对参考帧的位移的运动矢量信息 455 计算预测块 465。解码器将预测块 465 和重建的误差块 445 组合（470）以形成重建块 475。

在用于先前 WMV 编码器的软件和用于先前 WMV 解码器的软件中，在比特流中的宏块级别以每宏块一个比特为基础来用信号表示 AC 预测信息。

ACPRED 半帧是一个比特的宏块级别比特流元素，它指定 AC 预测是否用于解码用于宏块中的所有块的 AC 系数。ACPRED 在 I 帧中并在预测帧的 IMV 帧内宏块中出现。ACPRED=0 一般指示宏块中不使用 AC 预测，而 ACPRED=1 一般指示宏块中使用 AC 预测。预测值块是当前块的紧接上面或左面的块。然而，在预测帧（例如，P 帧或 B 帧）中，如果上预测值块和左预测值块不是帧内编码的，则即便 ACPRED＝1 也不使用 AC 预测。

编码器和解码器也使用用于隔行扫描帧的 AC 预测的信令。ACPREDMB 标记是帧编码的宏块的宏块级别处出现的一个比特的值，它指定 AC 预测是否被用于宏块中的所有块。ACPREDTFIELD 和 ACPREDBFIELD 标记是半帧编码的宏块的宏块级别处出现的一个比特的值，它们分别指定 AC 预测是否被用于当前宏块的上和下半帧中的块。

II.隔行扫描视频和逐行扫描视频

视频帧包含视频信号的空间信息的线。对于逐行扫描视频，这些线包含从一个时刻开始持续通过逐行扫描线到达帧底部的采样。逐行扫描 I 帧是帧内编码的逐行扫描视频帧。逐行扫描 P 帧是用正向预测编码的逐行扫描视频帧，且逐行扫描 B 帧是用双向预测编码的逐行扫描视频帧。

典型的隔行扫描视频帧由在不同时间处开始扫描的两个半帧构成。例如，参考图 5，隔行扫描视频帧 500 包括上半帧 510 和下半帧 520。通常，偶数线（上半帧）在一个时间（例如，时间 t）开始被扫描，而奇数线（下半帧）在不同（通常稍晚）时间（例如，时间 t+1）处开始被扫描。该定时会形成隔行扫描视频帧区域内的锯齿状特性，在该区域中当在不同时间处开始扫描这两个半帧时呈现出运动。为此，可以根据半帧结构重排隔行扫描视频帧，其中将奇数线在一个半帧中组合在一起并将偶数线在另一半帧中组合在一起。被称作半帧编码的这种排列在用于在高运动图片中减少这种锯齿边缘人工效应是很有用的。另一方面，在静止区域中，在没有这种重排的情况下，可以更有效地保留隔行扫描视频帧的图像细节。因此，帧编码常在静止或较低运动的隔行扫描视频帧中使用，其中保留了原始的交替半帧线排列。

用于先前 WMV 编码器的软件和用于先前解码器的软件使用根据隔行扫描视频帧中的半帧结构（半帧编码宏块）或帧结构（帧编码宏块）排列的宏块。图 6 示出了编码器和解码器中的半帧编码宏块的结构。置换隔行扫描宏块 610 以使所有的上半帧线（例如，偶数线 0、2、...14）被置于半帧编码宏块 620 的上半部分，且所有下半帧线（例如，奇数线 1、3、...15）被置于半帧编码宏块的下半部分。对于帧编码的宏块，上半帧线和下半帧线遍及该宏块交替，如同在隔行扫描宏块 610 中那样。

先前的编码器和解码器在隔行扫描帧中使用 4：1：1 宏块格式。4：1：1 宏块由四个 8×8 亮度块和每个色度通道的两个 4×8 块构成。在半帧编码的 4：1：1 宏块中，细分置换的宏块，以使上部两个 8×8 亮度块和每个色度通道中的上部 4×8 色度块仅包含上半帧线，而下部两个 8×8 亮度块和每个色度通道中的下部 4×8 色度块仅包含下半帧线。

典型的逐行扫描视频帧由具有非交替线的内容的一个帧构成。对照隔行扫描视频，逐行扫描视频不将视频帧划分成分开的半帧，且在单个时间处开始，从左到右、从上到下地扫描整个帧。

III.用信号表示隔行扫描宏块的帧/半帧模式

在用于先前的 WMV 编码器和解码器的软件中，INTRLCF 半帧是一个比特的

帧层元素，用于用信号表示宏块是仅按帧模式还是按半帧或帧模式被编码。如果INTRLCF＝0，帧中的所有宏块都按帧模式编码。如果INTRLCF＝1，则帧中的宏块可按半帧或帧模式编码，且INTRLCMB字段在比特流中跟随以指示每个宏块的帧/半帧编码状态。INTRLCMB是存在于逐行扫描I帧、隔行扫描I帧、隔行扫描P帧和隔行扫描B帧中的位平面。解码的INTRLCMB位平面按从左上方到右下方的光栅扫描顺序将每个宏块的隔行扫描状态表示为一个比特值的半帧。0值指示相应的宏块按帧模式编码。1值指示相应的宏块按半帧模式编码。

在逐行扫描I帧、隔行扫描I帧、隔行扫描P帧和隔行扫描B帧中为每个宏块用信号表示半帧/帧编码模式，且该半帧/帧编码模式仅通过位平面在帧级别被用信号表示。没有宏块层信令选项可用于用信号表示半帧/帧编码模式，这限制了信令的灵活性。

IV.用于先前的WMV编码器和解码器的软件中的位平面编码

在用于先前WMV编码器和解码器的软件中，用于帧中宏块的某些二进制信息按七种位平面编码模式之一被编码为二维数组，并在帧头中发送。

编码器和解码器使用位平面编码以便在用于帧中宏块的帧级别用信号表示四个不同种类的二进制信息：（1）跳过/不跳过宏块，（2）隔行扫描图片中的半帧或帧编码模式，（3）一个运动矢量[1MV]或四个运动矢量[4MV]编码模式；以及（4）B帧中的直接/非直接预测模式。在位平面编码方案中使用以下的语法元素。

INVERT

INVERT字段是一个比特的代码，它指示位平面有更多比特等于0还是更多比特等于1。根据INVERT和位平面编码模式，解码器可逆转所解码的位平面以重新形成原始的。

IMODE

IMODE字段是表示位平面编码模式的可变长度码[VLC]。一般，更短的码用于编码更频繁出现的编码模式。

DATABITS

DATABITS字段是基于IMODE字段中用信号表示的编码模式的符号的熵编

码流。每个二维数组的大小是 rowMB×colMB，其中 rowMB 和 colMB 分别是帧中宏块行和列的数量。在比特流内，每个数组都按七种位平面编码模式之一被编码为一组连续比特。以下描述这七种位平面编码模式。

1.原始模式

在原始模式中，位平面被编码为按自然扫描顺序扫描的每像素一个比特。DATABITS 在长度方面是 rowMB×colMB 比特。

2.行跳过模式

在行跳过模式中，ROWSKIP 字段指示对于位平面中的每个行是否存在 ROWBITS 字段。如果位平面中的整行值是零，则 ROWSKIP＝0 且 ROWBITS 被跳过。如果行中的至少一个值是非零，则 ROWSKIP＝1 且 ROWBITS 包含用于行中每个值的一个比特。从帧的上部到下部扫描行。

3.列跳过模式

在列跳过模式中，COLUMNSKIP 字段指示对于位平面中的每个列是否存在 COLUMNBITS 字段。如果位平面中的整列值是零，则 COLUMNSKIP＝0 且 COLUMNBITS 被跳过。如果列中的至少一个值是非零，则 COLUMNSKIP＝1 且 COLUMNBITS 包含用于行中每个值的一个比特。从帧的左面到右面扫描列。

4.正常－2 模式

在正常－2 模式中，如果 rowMB×colMB 是奇数，则第一个符号用匹配其值的一个比特简单地表示，且后续符号用二进制 VLC 表按自然扫描顺序成对地进行编码。

5.正常－6 模式

在正常－6 模式中，以六个像素的组编码位平面。这些像素被组成为 2×3 或 3×2 的块。位平面使用一组平铺规则最大地被平铺，且其余像素用行跳过和列跳过模式的变型进行编码。如果且仅如果 rowMB 是 3 的倍数而 colMB 不是，则使用 3×2“垂直”块。否则，使用 2×3“水平”块。

首先编码六个元素块，接着是列跳过和行跳过编码的线性块。如果数组大小是 3×2 的倍数或者 2×3 的倍数，则后者的线性块不存在且仅用六个元素的矩形块平铺该位平面。

6，7.Diff-2 和 Diff-6 模式

如果使用任一差分模式（Diff-2 或 Diff-6），则用相应的正常模式（分别是正常－2 或正常－6）来解码“差分比特”的位平面。差分比特被用于重新生成原始位平面。

对于关于位平面编码的更多信息，参见 2002 年 12 月 16 日提交的美国专利申请 No.10/321415，标题为“Skip Macroblock Coding（跳过宏块编码）”。

V.用于视频压缩和解压缩的标准

几个国际标准涉及视频压缩和解压缩。这些标准包括运动图像专家组[MPEG]1、2 和 4 标准以及来自国际电信联盟（ITU）的 H.261,H.262（MPEG-2 的另一名称）,H.263 和 H.264（也称作 JVT/AVC）标准。这些标准规定视频解码器的各方面以及用于压缩的视频信息的格式。直接地或通过暗示，它们也规定某些编码器细节，但未规定其它编码器细节。这些标准使用（或支持其使用）帧内和帧间解压缩和压缩的不同组合。

A.用信号表示标准中的半帧或帧编码的宏块

一些国际标准描述了用于隔行扫描图片中的宏块的半帧编码或帧编码的信令。

JVT/AVC 标准的草案 JVT-d157 描述了 mb_field_decoding_flag 语法元素，它用于用信号表示在隔行扫描 P 帧中是按帧模式还是半帧模式来解码宏块对。章节 7.3.4 描述了比特流语法，其中在序列参数（mb_frame_field_adaptive_flag）指示宏块中帧和半帧解码之间的切换且片头元素（pic_structure）将图片结构标识为隔行扫描帧图片的情况下，mb_field_decoding_flag 作为片数据的元素发送。

1998 年 5 月 28 日，MPEG-4 的委员会草案描述了 dct_type 语法元素，它用于用信号表示宏块是帧 DCT 编码的还是半帧 DCT 编码的。根据章节 6.2.7.3 和 6.3.7.3，dct_type 是仅存在于隔行扫描内容中的 MPEG-4 比特流中的宏块层元素，其中宏块具有非零编码的块图案或者是帧内编码的。

在 MPEG-2 中，dct_type 元素也是宏块层元素，它指示宏块是帧 DCT 编码的还是半帧 DCT 编码的。MPEG-2 还描述图片编码扩展元素 frame_pred_frame_dct。当 frame_pred_frame_dct 被设定为‘1’时，在隔行扫描帧中仅使用帧 DCT 编码。当 frame_pred_frame_dct＝1 且 dct_type 元素不存在于比特流中时，“导出”条件 dct_type＝0。

B.用信号表示标准中的 AC 系数预测

一些国际标准描述了用于宏块的不同空间 AC 系数预测模式的信令。

1998 年 5 月 28 日，MPEG-4 委员会草案描述了 ac_pred_flag 语法元素，它是用于用信号表示是否差分编码帧内宏块的第一行或列中的 AC 系数的一个比特的标记。在 MPEG-4 比特流中，ac_pred_flag 以每宏块一个比特为基础在视频对象平面的数据划分数据结构（例如，data_partitioned_I_VOP()，data_partitioned_P_VOP()）或者在宏块层数据结构（macroblock()）中发送。

在 H.263 标准中，附录 1 描述了任选地使用 AC 预测的高级帧内编码模式。宏块层元素 INTRA_MODE 是可变长度码，它用信号表示是否以使用 AC 预测的模式编码宏块。

C.标准的限制

这些国际标准在几个重要方面受到限制。例如，尽管这些标准提供用于用信号表示半帧/帧类型信息和 AC 预测，但信号表示通常是以每宏块一个比特为基础进行的。

由于视频压缩和解压缩对数字视频的重要性，不难理解视频压缩和解压缩是丰富发展的领域。但无论先前的视频压缩和解压缩技术的好处是什么，它们都不具有以下技术和工具的优点。

发明内容

总之，详细说明针对用于视频编码和解码的各种技术和工具。例如，编码器位平面编码 AC 预测状态信息。作为另一示例，编码器用信号表示用于隔行扫描帧编码图片中的宏块的半帧/帧变换类型。解码器执行相应的解码。所述实施例实现了所述技术和工具中的一个或多个，包括但不限于以下：

一方面，编码器/解码器从多个可用位平面模式组中选择一个位平面模式，并根据选定的位平面模式来处理位平面，其中该位平面指示视频图片的多个宏块的 AC 预测状态信息。

另一方面，编码器编码指示视频图片的多个宏块的 AC 预测状态信息的一个位平面并用信号表示编码的位平面。

另一方面，解码器接收编码的位平面并解码该位平面，其中该位平面指示视

频图片的多个宏块的 AC 预测状态信息。

另一方面，对于视频序列中的第一隔行扫描视频帧，解码器解码在第一隔行扫描视频帧的帧层处用信号表示的位平面。位平面表示第一隔行扫描视频帧的多个宏块的半帧/帧变换类型。对于视频序列中的第二隔行扫描视频帧，对于第二隔行扫描视频帧的多个宏块的至少一个但非全部之中的每一个，解码器处理宏块层处用信号表示的每宏块的半帧/帧变换类型比特。

另一方面，对于视频序列中的第一隔行扫描视频帧，编码器编码位平面并在第一隔行扫描视频帧的帧层处用信号表示该位平面。位平面表示第一隔行扫描视频帧的多个宏块的半帧/帧变换类型。对于视频序列中的第二隔行扫描视频帧，对于第二隔行扫描视频帧的多个宏块的至少一个但非全部之中的每一个，编码器在宏块层处用信号表示每宏块的半帧/帧变换类型比特。

另一方面，解码器包括用于解码帧层处用信号表示的位平面的装置，其中该位平面包括用于第一隔行扫描视频帧的第一位平面，其中第一位平面表示第一隔行扫描视频帧的多个宏块的半帧/帧变换类型，以及用于处理在宏块层处用信号表示的每宏块的半帧/帧变换类型比特的装置，其中每宏块的半帧/帧变换类型比特包括用于第二隔行扫描视频帧的多个宏块的一个或多个但非全部之中的每一个的每宏块的半帧/帧变换类型比特。

各种技术和工具可组合或独立使用。

附加的特点和优点将通过以下不同实施例的详细描述并通过参考附图而变得显而易见。

附图说明

图 1A 是根据现有技术的 8×8 像素块的基于块的帧内压缩的示图。图 1B 示出了根据现有技术的用于 8×8 像素块的 AC 预测候选。

图 2 是示出根据现有技术的视频编码器中的运动估计的示图。

图 3 是示出根据现有技术的视频编码器中预测残差的 8×8 块的基于块的压缩的示图。

图 4 是示出根据现有技术的视频编码器中预测残差的 8×8 块的基于块的解压缩的示图。

图 5 是示出根据现有技术的隔行扫描帧的示图。

图 6 是示出根据现有技术的隔行扫描宏块的半帧编码的示图。

图 7 是可结合用于实现几个所述实施例的合适计算环境的框图。

图 8 是可结合用于实现几个所述实施例的一般化的视频编码器系统的框图。

图 9 是可结合用于实现几个所述实施例的一般化的视频解码器系统的框图。

图 10 是几个所述实施例中使用的宏块格式的示图。

图 11A 是隔行扫描视频帧的一部分的示图，示出了上半帧和下半帧的交替线。图 11B 是为编码/解码被组织为一帧的隔行扫描视频帧的示图，其图 11C 是为编码/解码组织为半帧的隔行扫描视频帧的示图。

图 12 是示出用于利用帧级别位平面编码或宏块级别编码来用信号表示用于宏块的半帧/帧类型信息的技术的流程图。

图 13 是示出 AC 预测信息的位平面编码的技术的流程图。

图 14 是示出在组合实现中用于逐行扫描 I 帧的帧层比特流语法的示图。

图 15 是示出在组合实现中用于逐行扫描 I 帧的宏块层比特流语法的示图。

图 16 是示出在组合实现中用于隔行扫描 I 半帧或 BI 半帧的帧层比特流语法的示图。

图 17 是示出在组合实现中用于隔行扫描 I 半帧的半帧层比特流语法的示图。

图 18 是示出在组合实现中用于隔行扫描 I 帧的帧层比特流语法的示图。

图 19 是示出在组合实现中用于隔行扫描 P 帧的帧层比特流语法的示图。

图 20 是示出在组合实现中用于隔行扫描 B 帧的帧层比特流语法的示图。

图 21 是示出在组合实现中用于隔行扫描 I 帧的宏块的宏块层比特流语法的示图。

图 22 是示出在组合实现中用于隔行扫描 P 帧的宏块的宏块层比特流语法的示图。

图 23 是示出在组合实现中用于隔行扫描 B 帧的宏块的宏块层比特流语法的示图。

图 24A－24C 是示出用于根据现有技术的正常－6 和 diff－6 位平面编码模式的块的示图。

具体实施方式

本申请涉及用于逐行扫描和隔行扫描视频的有效压缩和解压缩的技术和工具。在各种描述的实施例中，视频编码器和解码器结合了用于编码和解码逐行扫描及隔行扫描视频的技术，以及用于与包括不同层或级别（例如，序列级别、帧级别、半帧级别、宏块级别和/或块级别）的比特流格式或语法一起使用的相应信令技术。

这里所述实现的各种可选方案也是可能的。例如，通过改变流程图中示出的阶段的排序，通过重复或省去某些阶段等，可改变参考流程图描述的技术。作为另一示例，虽然参考特定的宏块格式描述了一些实现，但也可以使用其它格式。此外，参考正向预测描述的技术和工具也可应用于其它类型的预测。

各种技术和工具可组合或独立使用。不同的实施例实现了一个或多个所述技术和工具。这里描述的一些技术和工具可以在视频编码器或解码器中使用，或者在不特别限于视频编码或解码的其它系统中使用。

I.计算环境

图 7 示出了合适计算环境 700 的一般化示例，其中可以实现几种所述实施例。计算环境 700 不旨在暗示对使用范围或功能的任何限制，因为这些技术和工具可在不同的通用或专用计算环境中实现。

参考图 7，计算环境 700 包括至少一个处理单元 710 和存储器 720。图 7 中，该最基本配置 730 包含在虚线内。处理单元 710 执行计算机可执行指令并且可以是实际的或虚拟的处理器。在多处理系统中，多个处理单元执行计算机可执行指令以提升处理能力。存储器 720 可以是易失性存储器（例如，寄存器、高速缓存、RAM），非易失性存储器（例如，ROM、EEPROM、闪存等）或者这两者的组合。存储器 720 存储实现具有一个或多个上述技术和工具的视频编码器或解码器的软件 780。

计算环境可具有附加的特性。例如，计算环境 700 包括存储装置 740、一个或多个输入装置 750、一个或多个输出装置 760 以及一个或多个通信连接 770。诸如总线、控制器或网络的互连机制（未示出）互连计算环境 700 的组件。通常，操作系统软件（未示出）为计算环境 700 中执行的其它软件提供操作环境，并协调计算环境 700 的组件的活动。

存储装置 740 可以是可移动或不可移动的，并包括磁盘、磁带或盒式磁带、CD-ROM、DVD 或可用于存储信息并可以在计算环境 700 内访问的任何其它介质。

存储装置 740 存储实现视频编码器或解码器的软件 780 的指令。

输入装置 750 可以是诸如键盘、鼠标、笔或轨迹球的触摸输入装置，语音输入装置、扫描装置或向计算环境 700 提供输入的其它装置。对于音频或视频编码，输入装置 750 可以是声卡、视频卡、TV 调谐器或以模拟或数字形式接收音频或视频输入的类似装置，或者将音频或视频采样读入计算环境 700 的 CD-ROM 或 CD-RW。输出装置 760 可以是显示器、打印机、扬声器、CD 刻录机或提供来自计算环境 700 的输出的其它装置。

通信连接 770 支持通信媒体上到另一计算实体的通信。通信媒体在调制的数据信号中传送信息，诸如计算机可执行指令、音频或视频输入或输出或者其它数据。调制的数据信号是按编码信号中的信息的方式设定或变化其一个或多个特征的信号。作为示例而非限制，通信媒体包括用电、光、RF、红外线、声或其它载体实现的有线或无线技术。

这些技术和工具可在计算机可读媒体的一般上下文中进行描述。计算机可读媒体是可以在计算环境内访问的任何可用媒体。作为示例而非限制，在使用计算环境 700 的情况下，计算机可读媒体包括存储器 720、存储装置 740、通信媒体和以上的任何组合。

这些技术和工具可在计算环境下在目标实际或虚拟处理器上执行的诸如程序模块中包含的那些计算机可执行指令的一般上下文中进行描述。一般，程序模块包括例程、程序、库、对象、组件、数据结构等，它们执行特殊任务或者实现特殊的抽象数据类型。如各种实施例中所期望的，程序模块的功能可以在程序模块之间组合或分离。用于程序模块的计算机可执行指令可在本地或分布式计算环境内执行。

为便于表达，详细描述使用如“估计”、“补偿”、“预测”和“应用”的术语来描述计算环境中的计算机操作。这些术语是计算机执行的操作的高级抽象，且不应与人执行的动作相混淆。对应于这些术语的实际计算机操作根据实现情况而变化。

II.一般化的视频编码器和解码器

图 8 是结合用于实现一些所述实施例的一般化视频编码器 800 的框图。图 9 是结合用于实现一些所述实施例的一般化视频解码器 900 的框图。

所示出的编码器 800 和解码器 900 内模块之间的关系表示编码器和解码器中

信息的一般流程，出于简化的原因而未示出其它关系。特别是，图 8 和 9 通常不示出指示用于视频序列、图片、宏块、块等的编码器设定、模式、表等的辅助信息。这种辅助信息通常在辅助信息的熵编码后在输出比特流中发送。输出比特流的格式可以是 Windows Media Video 版本 9 格式或其它格式。

编码器 800 和解码器 900 处理视频图片，它们可以是视频帧、视频半帧或帧和半帧的组合。图片和宏块级别的比特流语法和语义可取决于是使用帧还是半帧。可存在对宏块组织和总定时的变化。编码器 800 和解码器 900 是基于块的，并对帧使用 4：2：0 宏块格式，其中每个宏块包括四个 8×8 亮度块（有时处理为一个 16×16 宏块）和两个 8×8 色度块。对于半帧，可以使用相同或不同的宏块组织和格式。在不同阶段处可进一步细分这些 8×8 块，例如在频率变换和熵编码阶段。以下更详细地描述示例视频帧组织。

根据所期望的实现和压缩类型，可以将编码器或解码器的模块添加、省略、分成多个模块、与其它模块组合和/或用类似模块代替。在可选实施例中，具有不同模块和/或其它模块配置的编码器或解码器执行一个或多个所述技术。

A.视频帧组织

在一些实现中，编码器 800 和解码器 900 处理如下组织的视频帧。帧包含视频信号的空间信息的线。对于逐行扫描视频，这些线包含从一个时刻开始并继续通过连续的线到达帧底部的采样。逐行扫描视频帧被分成宏块，诸如图 10 所示的宏块 1000。宏块 1000 包括四个 8×8 亮度块（Y1 到 Y4）以及同这四个亮度块共同定位但只有一半的水平和垂直分辨率的两个 8×8 色度块，遵循常规的 4：2：0 宏块格式。8×8 块可以在不同的阶段被进一步细分，例如在频率变换（例如，8×4，4×8 或 4×4 DCT）和熵编码阶段。逐行扫描 I 帧是帧内编码的逐行扫描视频帧。逐行扫描 P 帧是用正向预测编码的逐行扫描视频帧，且逐行扫描 B 帧是用双向预测编码的逐行扫描视频帧。逐行扫描 P 和 B 帧可包括帧内编码的宏块以及不同类型的预测宏块。

隔行扫描视频帧由对帧的两次扫描构成——一次包含帧的偶数线（上半帧）且另一次包括帧的奇数线（下半帧）。这两个半帧可表示两个不同的时间周期或者它们可以来自于同一时间周期。图 11A 示出了隔行扫描视频帧 1100 的一部分，包括隔行扫描视频帧 1100 的左上部分处上半帧和下半帧的交替线。

图 11B 示出了为编码/解码而被组织为帧 1130 的图 11A 的隔行扫描视频帧 1100。隔行扫描视频帧 1100 已被划分成宏块，诸如宏块 1131 和 1132，如图 10 所示使用 4：2：0 格式。在亮度平面内，每个宏块 1131 和 1132 包括与来自下半帧的 8 条线交替的来自上半帧的 8 条线，总共 16 条线，且每条线都是 16 个像素长。（未示出宏块 1131、1132 内亮度块和色度块的实际组织和放置，且实际上可针对不同的编码决策而变化）。在给定的宏块内，上半帧信息和下半帧信息可在各阶段中的任一个处共同或分开地被编码。隔行扫描 I 帧是隔行扫描视频帧的两个帧内编码半帧，其中宏块包括用于这两个半帧的信息。隔行扫描 P 帧是使用正向预测编码的隔行扫描视频帧的两个半帧，而隔行扫描 B 帧是用双向预测编码的隔行扫描视频帧的两个半帧，其中宏块包括用于这两个半帧的信息。隔行扫描 P 和 B 帧可包括帧内编码的宏块以及不同类型的预测宏块。隔行扫描 BI 帧是隔行扫描 I 帧和隔行扫描 B 帧的混合物；它是帧内编码的但不用作其它帧的锚。

图 11C 示出了为编码/解码而组织为半帧 1160 的图 11A 的隔行扫描视频帧 1100。将隔行扫描视频帧 1100 的两个半帧中的每一个分成宏块。上半帧被分成诸如宏块 1161 的宏块，且下半帧被分成诸如宏块 1162 的宏块。（再次，宏块使用如图 10 所示的 4：2：0 格式，且未示出宏块内亮度块和色度块的组织和放置。）在亮度平面内，宏块 1161 包括来自上半帧的 16 条线且宏块 1162 包括来自下半帧的 16 条线，且每条线都是 16 像素长。隔行扫描 I 半帧是隔行扫描视频帧的单个个别表示的半帧。隔行扫描 P 半帧是用正向预测编码的隔行扫描视频帧的单个个别表示的半帧，且隔行扫描 B 半帧是用双向预测编码的隔行扫描视频帧的单个个别表示的半帧。隔行扫描 P 和 B 半帧可包括帧内编码的宏块以及不同类型的预测宏块。隔行扫描 BI 半帧是隔行扫描 I 半帧和隔行扫描 B 半帧的混合物；它们是帧内编码的，但不用作其它半帧的锚。

为编码/解码组织为半帧的隔行扫描视频帧可包括不同半帧类型的各种组合。例如，这种帧可在上和下半帧中具有相同的半帧类型或者在每个半帧中具有不同的半帧类型。在一个实现中，半帧类型的可能组合包括 I/I,I/P,P/I,P/P,B/B,B/BI,BI/B 和 BI/BI。

术语图片一般表示源、编码的或重建的图像数据。对于逐行扫描视频，图片是逐行扫描视频帧。对于隔行扫描视频，图片可表示隔行扫描视频帧、帧的上半帧

或帧的下半帧，这取决于上下文。

或者，编码器 800 和解码器 900 是基于对象的，使用不同的宏块或块格式，或者对与 8×8 块和 16×16 宏块不同大小或配置的像素组进行操作。

B.视频编码器

图 8 是一般化的视频编码器系统 800 的框图。编码器系统 800 接收包括当前图片 805 的视频图片序列（例如，逐行扫描视频帧、隔行扫描视频帧或隔行扫描视频帧的半帧），并产生压缩的视频信息 895 作为输出。视频编码器的特殊实施例通常使用一般化编码器 800 的变型或补充版本。

编码器系统 800 压缩预测图片和关键图片。为便于表达，图 8 示出了关键图片通过编码器系统 800 的路径和用于预测图片的路径。编码器系统 800 的许多组件被用于压缩关键图片和预测图片。这些组件执行的准确操作可根据正被压缩的信息的类型而变化。

根据从一个或多个其它图片（它们通常被称作参考图片或锚）的预测（或差异）来表示预测图片（例如，逐行扫描 P 帧或 B 帧，隔行扫描 P 半帧或 B 半帧，或者隔行扫描 P 帧或 B 帧）。预测残差是被预测的图片和原始图片之间的差。相反，在不参考其它图片的情况下压缩关键图片（例如，逐行扫描 I 帧、隔行扫描 I 半帧或隔行扫描 I 帧）。

如果当前图片 805 是正向预测的图片，则运动估计器 810 相对于一个或多个参考图片（例如图片存储 820 中缓冲的重建的先前图片 825）估计当前图片 805 的宏块或其它像素组的运动。如果当前图片 805 是双向预测的图片，则运动估计器 810 相对于达 4 个重建参考图片（例如对于隔行扫描 B 半帧）估计当前图片 805 中的运动。通常，运动估计器相对于一个或多个时间上先前的参考图片和一个或多个时间上将来的参考图片估计 B 图片中的运动。因此，编码器系统 800 可使用分开的存储装置 820 和 822 用于多个参考图片。对于关于逐行扫描 B 帧以及隔行扫描 B 帧和 B 半帧的更多信息，参见 2003 年 7 月 18 日提交的美国专利申请序号 No.10/622378，标题为“Advanced Bi-Directional Predictive Coding of Video Frames（视频帧的高级双向预测编码）”，以及 2004 年 6 月 29 日提交的美国专利申请序号 No.10/882135，标题为“Advanced Bi-Directional Predictive Coding of Interlaced Video（隔行扫描视频的高级双向预测编码）”，在此并入全文以供参考。

运动估计器 810 可按像素、1/2 像素、1/4 像素或其它增量估计运动，并可逐图片或以其它为基础切换运动估计的精度。运动估计器 810（和补偿器 830）也可以每帧或其它为基础在参考图片像素插值类型之间切换（例如，在双三次和双线性之间）。运动估计的精度可以水平和垂直地相同或不同。运动估计器 810 输出诸如差分运动矢量信息的运动信息 815 作为辅助信息。

运动补偿器 830 将运动矢量应用于重建图片 825 以形成运动补偿的当前图片 835。但预测很少是完美的，且运动补偿的当前图片 835 和原始当前图片 805 之间的差是预测残差 845。在稍后的图片重建期间，将预测残差 845 添加到运动补偿的当前图片 835 中以获得更接近于原始当前图片 805 的重建图片。但在有损的压缩中，原始当前图片 805 中仍损失一些信息。或者，运动估计器和运动补偿器应用另一类型的运动估计/补偿。

频率变换器 860 将空间域视频信息转换成频域（即频谱）数据。对于基于块的视频图片，频率变换器 860 将 DCT、DCT 的变型或其它块变换应用于像素数据或预测残差数据的块，产生频率变换系数块。或者，频率变换器 860 应用诸如傅立叶变换的另一常规频率变换或使用小波或子频带分析。频率变换器 860 可应用 8×8、8×4、4×8、4×4 或其它大小的频率变换。

随后，量化器 870 量化频谱数据系数的块。量化器将均匀的标量量化应用于具有逐图片或以其它为基础变化的步长的频谱数据。或者，量化器将另一种量化应用于频谱数据系数，例如不均匀矢量或不自适应的量化，或者在不使用频率变换的编码器系统中直接量化空间域数据。除了自适应量化，编码器 800 可使用丢帧、自适应滤波或其它技术用于速率控制。

编码器 800 可使用特殊的信令用于跳过的宏块，它是没有特定类型信息的宏块。

当后续运动估计/补偿需要重建的当前图片时，反量化器 876 对被量化的频谱数据系数进行反量化。随后，反频率变换器 866 进行频率变换器 860 的反操作，形成重建的预测残差（用于预测图片）或重建的关键图片。如果当前图片 805 是关键图片，则将重建的关键图片作为重建的当前图片（未示出）。如果当前图片 805 是预测图片，则将重建的预测残差加到运动补偿的当前图片 835 以形成重建的当前图片。图片存储 820 和 822 之一或两者缓冲重建的当前图片，用于在运动补偿预测

中使用。在一些实施例中，编码器将解块滤波器应用于重建的帧以自适应地平滑图片中的不连续和其它人工效应。

熵编码器 880 压缩量化器 870 的输出以及某些辅助信息（例如运动信息 815、量化步长）。典型的熵编码技术包括算术编码、差分编码、Huffman 编码、行程编码、LZ 编码、字典编码和以上的组合。熵编码器 880 通常使用不同的编码技术用于不同种类的信息（例如 DC 系数、AC 系数、不同种类的辅助信息），并可以从特殊的编码技术内多个码表中进行选择。

熵编码器 880 向多路复用器[MUX]890 提供压缩的视频信息 895。MUX890 可包括一缓冲器，且可将缓冲器级指示器反馈给比特率自适应模块，用于速率控制。在 MUX890 之前或之后，压缩的视频信息 895 可以被信道编码，以便在网络上传输。信道编码可将错误检测和纠正数据应用于压缩的视频信息 895。

C.视频解码器

图 9 是一般视频解码器系统 900 的框图。解码器系统 900 接收用于压缩的视频图片序列的信息 995 并产生包括重建图片 905 的输出（例如，逐行扫描视频帧、隔行扫描视频帧或隔行扫描视频帧的半帧）。视频解码器的特殊实施例通常使用一般化解码器 900 的变型或补充型式。

解码器系统 900 解压缩预测图片和关键图片。为便于呈现，图 9 示出了用于通过解码器系统 900 的关键图片的路径以及用于正向预测图片的路径。解码器系统 900 的许多组件被用于解压缩关键图片和预测图片两者。这些组件执行的准确操作可根据解压缩的信息类型而变化。

DEMUX 990 接收用于压缩视频序列的信息 995 并使得接收到的信息对熵解码器 980 可用。DEMUX 990 可包括抖动缓冲器以及其它缓冲器。在 DEMUX 990 之前或之后，可将压缩的视频信息信道解码并处理，用于错误检测和纠正。

熵解码器 980 熵解码被熵编码的量化数据以及被熵编码的辅助信息（例如，运动信息 915、量化步长），通常应用编码器中执行的熵编码的逆。熵解码技术包括算术解码、差分解码、Huffman 解码、行程解码、LZ 解码、字典解码和以上的组合。熵解码器 980 通常使用不同的解码技术用于不同种类的信息（例如，DC 系数、AC 系数、不同种类的辅助信息），并可在特殊的解码技术内的多个码表中进行选择。

解码器 900 例如通过计算用于运动矢量的一个或多个预测值、熵解码差分运动矢量以及组合解码的差分运动矢量和预测值来重建运动矢量而解码运动信息 915。

运动补偿器 930 将运动信息 915 应用于一个或多个参考图片 925 以形成正重建的图片 905 的预测 935。例如，运动补偿器 930 使用一个或多个宏块运动矢量来找到参考图片 925 中的宏块。一个或多个图片存储（例如，图片存储 920、922）存储用作参考图片的先前重建图片。通常，B-图片具有超过一个参考图片（例如，至少一个时间上先前的参考图片和至少一个时间上将来的参考图片）。因此，解码器系统 900 可使用分开的图片存储 920 和 922 用于多个参考图片。运动补偿器 930 可按像素、1/2 像素、1/4 像素或其它增量补偿运动，并可逐图片或以其它为基础切换运动补偿的精度。运动补偿器 930 也可以每帧或其它为基础在参考图片像素插值类型之间切换（例如，在双三次和双线性之间）。运动补偿的精度可以水平和垂直地相同或不同。或者，运动补偿器应用另一类型的运动补偿。运动补偿器进行的预测很少是完美的，因此解码器 900 也重建预测残差。

反量化器 970 反量化被熵解码的数据。一般，反量化器将均匀的标量反量化应用于具有逐图片或以其它为基础变化的步长的熵解码数据。或者，反量化器将另一类型的反量化应用于数据，例如以便在不均匀矢量或非自适应量化后重建，或者在不使用反频率变换的解码器系统中直接反量化空间域数据。

反频率变换器 960 将被量化的频域数据转换成空间域视频信息。对于基于块的视频图片，反频率变换器 960 将反 DCT[IDCT]、IDCT 的变型或其它反向块变换应用于频率变换系数块，分别产生用于关键图片或预测图片的像素数据或预测残差数据。或者，反频率变换器 960 应用诸如反傅立叶变换的另一常规反频率变换或者使用小波或子频带合成。反频率变换器 960 可应用 8×8、8×4、4×8、4×4 或其它大小的反频率变换。

对于预测图片，解码器 900 将重建的预测残差 945 与运动补偿预测 935 组合以形成重建的图片 905。当解码器需要重建图片 905 用于后续运动补偿时，图片存储（例如，图片存储 920）之一或两者缓冲重建的图片 905 用于在预测下一图片时使用。在一些实施例中，解码器 900 将解块滤波应用于重建图片，以自适应地平滑图片中的不连续和其它人工效应。

III.位平面编码

在一些实现中，宏块特有的二进制信息可被编码为位平面并以高于宏块级别的级别在比特流中传送（例如，帧级别、半帧级别或一些其它级别）。编码器例如可基于用于编码某种数据的不同编码模式的相对效率在几种编码模式中进行选择。位平面编码技术利用帧或半帧中宏块间的某些种类的二进制信息的空间相关。在与宏块级别信令组合时，所述技术和工具提供灵活和有效的宏块特有信息的信令。

在一些实现中，如下所述，编码器将 INVERT、IMODE 和 DATABITS 语法元素用于位平面编码。

INVERT 元素是一比特代码，它指示位平面具有更多比特等于 0 还是更多比特等于 1。根据 INVERT 和位平面编码模式，解码器可逆转解码的位平面以重新形成原始的位平面。IMODE 元素是表示位平面编码模式的 VLC。一般，较短的代码被用于编码更频繁出现的编码模式。DATABITS 元素是基于 IMODE 元素中用信号表示的编码模式的符号的熵编码流。例如，可将用于帧或半帧中宏块的二进制信息编码为二维数组并在帧或半帧的头中传送。每个数组的大小是 rowMB×colMB，其中 rowMB 和 colMB 分别是帧或半帧中宏块行和列的数量。

在一些实施例中，在七种位平面编码模式中，每个数组都被编码为一组连续比特（例如，表示按光栅扫描顺序的宏块）。这七种位平面编码模式如下所述。

行跳过模式

在行跳过模式中，ROWSKIP 元素指示 ROWBITS 元素是否对于位平面中的每行都存在。如果位平面中的整行值都是零，ROWSKIP=0 且跳过 ROWBITS。如果行中的至少一个值非零，则 ROWSKIP＝1 且 ROWBITS 包含用于该行中每个值的一个比特。

列跳过模式

在列跳过模式中，COLUMNSKIP 元素指示 COLUMNBITS 元素是否对于位平面中的每列都存在。如果位平面中的整列值是零，则 COLUMNSKIP＝0 且跳过 COLUMNBITS。如果该列中至少一个值非零，则 COLUMNSKIP＝1 且 COLUMNBITS 包含用于该列中每个值的一个比特。

正常－2 模式

在正常－2 模式中，如果 rowMB×colMB 是奇数，则用匹配第一个符号值的

一个比特来表示第一个符号，且使用 VLC 表成对地编码后续符号。

正常－6 模式

在正常－6 模式中，以 2×3 或 3×2 块的六个像素的组编码位平面。位平面使用一组平铺规则被组成为 2×3 和 3×2 块，且其余像素（如果有）用行跳过和列跳过模式的变型进行编码。如果数组大小是 3×2 或 2×3 的倍数，则仅用 6 个元素的矩形块平铺位平面。如果且仅如果 rowMB 是 3 的倍数而 colMB 不是，则使用 3×2“垂直”块。否则，使用 2×3“水平”块。利用可变长度和固定长度码的组合来编码六个元素的矩形块。

Diff-2 和 Diff-6 模式

如果使用任一差分模式（Diff-2 或 Diff-6），则用相应的正常模式（分别是正常－2 或正常－6）来解码“差分比特”的位平面。差分比特被用于重新生成原始位平面。

重新生成过程是二进制字母表上的 2-D DPCM。为了在位置（i,j）处重新生成比特，预测值 b_p（i,j）如下地生成（从位置（i,j）处的比特 b(i,j)）起）：

$$b_p(i,j)=\begin{cases} A & i=j=0,\ 或\ b(i,j-1)\neq b(i-1,j) \\ b(0,j-1) & i=0 \\ b(i-1,j) & ot\ 其它 \end{cases}$$

对于差分编码模式，不执行基于 INVERT 的逐位逆过程。然而，以不同的容量使用 INVERT 标记，来指示用于以上示出的预测值导出的符号 A 的值。更具体地，如果 INVERT=0 则 A＝0，而如果 INVERT＝1 则 A＝1。通过用解码的差分比特值对预测值求异或来获得位平面的实际值。在以上等式中，b(i,j)是最终解码后（即在进行正常－2/正常－6 之后，接着是用其预测值差分异或）第 i,j 位置处的比特。

原始模式

原始模式用每个二进制符号一个比特来编码宏块特有的二进制信息。虽然原始模式比特可在未压缩位平面中以帧或半帧级别进行编码，但对于比特流中与用于宏块的其余宏块级别信息相同层处的每个宏块，也可每次一个地编码原始模式比特。

其它模式在帧或半帧级别处编码位平面并在编码期间使用经过帧或半帧的第二次通过。例如，编码器在第一次通过中收集宏块级别处用于每个宏块的二进制信

息（例如，AC 预测状态），并在第二次通过中以帧或半帧级别编码宏块特有信息。但是，在低延时情形，避免第二次通过会是有利的。因此，将二进制信息的编码从帧或半帧级别切换到宏块级别的能力提供了附加的灵活性。

在一些实现中，编码器使用原始模式用于片编码（slice-coded）帧。片代表宏块的一个或多个邻接行。相对于空间预测，片中宏块的第一行一般被处理为图片中宏块的第一行。在使用片时，通常在图片层处的位平面中表示的信息在宏块层处以原始模式被用信号表示，使得每个宏块承载其自己的局部信息。

除了上述位平面编码模式，也可使用其它位平面编码模式。例如，一编码器可以任一大小的像素组编码位平面。在正常－6 模式的变型中，编码器可采用“最大矩形”模式，诸如当位平面中的二进制信息按可预测模式排列时。编码器可使用不同的平铺规则组将位平面组成为任一大小的块。随后，可用可变长度和固定长度码的组合或一些其它编码结构来编码这些块。

关于一些实现中位平面编码的更多信息，参见以下的章节 IV、V 和 VI。

IV.用于隔行扫描帧编码图片的半帧/帧类型信令的改进

所述实施例包括用于用信号表示隔行扫描帧编码图片（例如，隔行扫描 I 帧、隔行扫描 P 帧、隔行扫描 B 帧等）中的半帧/帧类型信息的技术和工具。例如，所述技术和工具包括能使用帧级别位平面编码或宏块级别编码来编码半帧/帧类型信息的编码器。解码器执行相应的解码。所述技术和工具可相互结合使用或者与其它技术和工具结合使用，或者可用单独使用。

一般，当存在较高半帧间运动时，半帧编码宏块是更有效的，且在存在较少运动时帧编码的宏块是更有效的。编码器可选择用于给定宏块的半帧/帧类型并指示它在比特流中被帧编码还是半帧编码。宏块的半帧/帧类型指示宏块的内部组织。再次参考图 6，在半帧编码中，置换隔行扫描宏块 610，以使亮度块中的所有上半帧线（例如，偶数线 0、2、…14）置于半帧编码的宏块 620 的上半部分中，并使亮度块中的所有下半帧线（例如，奇数线 1、3、…15）置于半帧编码的宏块的下半部分中。对于帧编码的宏块，上半帧线和下半帧线遍及该宏块交替，如同在隔行扫描宏块 610 中那样。对于半帧编码的宏块和帧编码的宏块两者，色度块保持隔行扫描。

因此，在一些实现中，编码器在几种编码模式中选择每宏块模式一个比特并

以宏块级别发送用于帧中的每个宏块的单个比特（例如，FIELDTX 标记），对其进行模式判断以明确地向解码器通知宏块是被半帧编码还是被帧编码。在编码这种信息时，选择宏块级别、每个宏块一个比特模式或帧级别、压缩位平面模式的能力提供了附加的灵活性。

特别是，当为帧的一些而非全部宏块用信号表示半帧/帧编码模式判定比特时，宏块级别信令是有效的。例如，在隔行扫描 P 帧中，为帧内宏块而非帧间宏块（否则为它用信号表示半帧/帧判定比特）用信号表示半帧/帧判定比特。这种上下文中的位平面编码导致无关信息的信令。另一方面，在一些帧中，每个宏块都具有半帧/帧判定比特。将半帧/帧类型比特组合在一起并在帧级别位平面中将它们编码可以利用帧中的宏块间存在的半帧/帧类型中的空间相关。

例如，图 12 示出了用于使用位平面编码或宏块级别编码用信号表示用于宏块的半帧/帧类型信息的技术 1200。在 1210，编码器选择编码模式。在 1220，如果编码器模式是位平面编码模式，则 1230 处编码器用信号表示半帧/帧类型信息作为位平面。否则，编码器在 1240 处确定是否发送信号比特（在 1250）以指示用于该宏块的半帧/帧判定。在 1260，如果存在将对其进行信令判断的更多宏块，则编码器确定是否为这些宏块发送信号比特。解码器执行相应的解码。

例如，在一个实现中，FIELDTX 是所有隔行扫描 I 帧宏块和隔行扫描 BI 帧宏块中以及隔行扫描 P 和 B 帧中的帧内编码宏块中存在的 1 比特语法元素。FIELDTX＝1 指示宏块被半帧编码，且 FIELDTX＝0 指示宏块被帧编码。FIELDTX 比特可以在数据流中以帧级别被位平面编码，或者对于一些或全部宏块以每宏块一个比特为基础以宏块级别被用信号表示。例如，在隔行扫描 I 帧和隔行扫描 BI 帧中，FIELDTX 比特在比特流中以帧级别或者以每宏块一个比特为基础以宏块级别被位平面编码。在隔行扫描 P 帧和隔行扫描 B 帧中，以每宏块一个比特为基础以宏块级别为帧内宏块发送 FIELDTX 比特。

或者，不同地用信号表示半帧/帧类型。例如，可以用不同大小或结构的比特流元素用信号表示（例如以宏块级别）半帧/帧类型。作为另一可选方案，可以以宏块级别以外的低于帧级别的级别（例如，片级别或一些其它级别）来用信号表示半帧/帧类型。作为另一可选方案，对于帧内编码帧中的不到全部宏块，可以宏块级别（或者一些其它级别）选择性地用信号表示半帧/帧类型。

V.用于逐行扫描和隔行扫描图片的 AC 预测信令中的改进

所述实施例包括用于在逐行扫描和隔行扫描图片（例如，逐行扫描 I 帧、隔行扫描 I 帧等）中用信号表示 AC 预测信息的技术和工具。例如，所述技术和工具包括用于利用位平面编码用信号表示帧或半帧中宏块的 AC 预测的使用的技术和工具。所述技术和工具可以相互结合使用或者与其它技术和工具结合使用，或者可以单独使用。

例如，编码器/解码器在熵编码之前预测 AC 系数。编码器从当前块中的 AC 系数值中减去相邻预测值块中的相应 AC 系数值以获得 AC 系数差。随后将该差熵编码。根据预测方向，预测值块是当前块的紧邻上方的块或者是当前块的左边的块。再次参考图 1B，对于顶部预测，当前块紧邻上方的块 175 中 AC 系数的顶行 177 被用作为当前块 125 中 AC 系数的顶行 129 的预测值。对于左预测，当前块的紧邻左边的块 135 中 AC 系数的最左面列 137 被用作为当前块 125 中 AC 系数的最左列的预测值。在一个实现中，为 DC 系数预测选择的预测方向被用作为用于 AC 系数预测的预测方向。或者，还用信号表示 AC 预测的方向。

AC 预测的具体机制可以针对不同的图片类型而不同。在其最简单形式中，AC 预测包括简单地使用相邻块的顶行或左列 AC 系数作为系数预测值。在更复杂的 AC 预测中，系数预测值可通过与当前块和相邻块中的量化级别相关的因素而被缩放。或者，编码器和解码器使用具有其它机制的 AC 预测。

如果特定预测方向中不存在块，则将用于最左列或顶行中的所有七个 AC 系数的预测值设定为零。例如，如果预测方向向上且当前块位于顶行中，则由于在向上方向中没有相邻块，所以将当前块的顶行中的每个预测 AC 系数设定为零。将预测行或列中的 AC 系数添加到当前块中的相应解码 AC 系数（它们是差）以形成完全重建的量化变换系数块。

为帧或半帧中的所有宏块进行 AC 系数预测会是低效的；在一些宏块中，预测将是无效的。因此，编码器/解码器使用信令来指示对于帧或半帧中的个别宏块启用还是禁用 AC 预测。可以在宏块级别为每个宏块发送信息的单个比特以明确向解码器通知该宏块中是否使用 AC 预测。然而，仅通过每宏块一个比特地发送来用信号表示 AC 预测的编码器/解码器不能利用半帧或帧中用于宏块的 AC 预测状态中的可能空间相关。

因此，在一些实现中，通过将用于单个宏块的 AC 预测信息组合在一起并将该信息编码为位平面(例如，以帧级别、半帧级别或宏块级别以上的一些其它级别)，编码器/解码器利用帧或半帧中的这种空间相关。在一种实现中，位平面是帧级别/半帧级别位平面 ACPRED。以下在章节 VII 中进一步详细描述该 ACPRED 位平面。

图 13 示出了用于 AC 预测信息的位平面编码或解码的技术 1300。在 1310，编码器/解码器从一组多个可用编码模式中选择一种编码模式。在 1320，编码器/解码器根据所选编码模式处理位平面，所述位平面包括表示视频帧中的宏块是否用 AC 预测进行编码的二进制信息。

所述位平面编码技术还可结合宏块级别的每宏块一个比特编码使用。例如，在一种实现中，在宏块层中被用信号表示时 ACPRED 语法元素指示 AC 预测是否用于该宏块中的块。当 ACPRED＝1 时，基于预测值块中的系数将宏块的解码块中的 AC 系数视为差分值。以下在章节 VII 中进一步详细描述宏块级别 ACPRED 比特流元素。

或者，不同地用信号表示 AC 预测信息。例如，可以用不同大小或结构的比特流元素用信号表示 AC 预测信息。作为另一可选方案，可以以帧级别或半帧级别以外的宏块级别以上的级别（例如，片级别或一些其它级别）在位平面中编码 AC 预测信息。作为另一可选方案，对于帧内编码帧中的不到全部宏块，可以宏块级别（或一些其它级别）选择性地用信号表示 AC 预测信息，其中在帧级别（或一些其它级别）可能使用位平面编码（或者其它编码技术）用于用信号表示不在宏块级别处用信号表示的宏块的半帧/帧类型。

VI.组合实现

除了与主要组合实现的较小差异的可选组合实现，现在描述比特流语法、语义和解码器的详细组合实现。

A.比特流语法

在各种组合实现中，数据按具有多层的比特流形式呈现（例如，序列、入口点、帧、半帧、片、宏块、块和/或子块层）。

在语法图中，箭头路径示出了语法元素的可能流程。用方角边界示出的语法元素指示固定长度的语法元素；用圆角边界示出的那些表示可变长度的语法元素且用外圆边界内的圆角边界示出的那些表示由更简单的语法元素构成的语法元素(例

如，位平面）。固定长度的语法元素被定义为其长度不取决于语法元素本身中的数据的语法元素；固定长度语法元素的参数是常数或者由语法流程中的先前数据确定。层图中的较低层（例如，帧层图中的宏块层）由矩形内的矩形表示。

序列级别元素用于解码压缩视频图片的序列。序列级别数据可影响其它级（例如，入口点级别、帧级别、宏块级别等）处元素的解释或出现。一般，入口点标记比特流中解码器可以开始解码的位置（例如，I帧或其它关键帧）。换句话说，不需要比特流中入口点之前的图片来解码入口点之后的图片。入口点头可用于用信号表示编码控制参数中的变化（例如，启用或禁用用于入口点以后的帧的压缩工具）。

对于逐行扫描I帧，图14示出了帧级别比特流元素。（用于逐行扫描BI帧的帧级别比特流元素与用于逐行扫描I帧的那些相同。）用于每个帧的数据由帧头继之以用于宏块层（用于帧内宏块）的数据构成。图15示出了用于逐行扫描I帧的宏块级别比特流元素。

对于具有隔行扫描I半帧和/或BI半帧的隔行扫描视频帧，图16示出了帧级别比特流元素。用于每个帧的数据由帧头继之以用于半帧层的数据构成（示作每个半帧的重复“FieldPicLayer”元素）。构成用于隔行扫描I半帧的半帧头的比特流元素在图17中示出。（用于隔行扫描BI半帧的半帧级别比特流元素与用于隔行扫描I半帧的那些相同。）构成用于隔行扫描I半帧和隔行扫描BI半帧的宏块层的比特流元素与逐行扫描I帧中的宏块的那些相同。）

对于隔行扫描I帧、P帧和B帧，帧级别比特流元素分别在图18、19和20中示出。（用于隔行扫描BI帧的帧级别比特流元素与用于隔行扫描I帧的那些相同。）在隔行扫描I帧、P帧和B帧中，用于每个帧的数据由帧头继之以用于宏块层（无论是用于帧内或各种帧间型宏块）的数据构成。构成用于隔行扫描I帧、P帧和B帧的宏块层的比特流元素分别在图21、22和23中示出。（用于隔行扫描BI帧中的宏块的比特流元素与用于隔行扫描I帧中的宏块的那些相同。）

以下部分描述帧、半帧和宏块层中的选定比特流元素。尽管在特殊层的上下文中描述选定比特流元素，但一些比特流元素可以在超过一个层中使用。

1.选定的帧和半帧层元素

图14、16、18、19和20是分别示出用于逐行扫描I帧/逐行扫描BI帧、具有隔行扫描I半帧或BI半帧的帧、隔行扫描I帧/隔行扫描BI帧、隔行扫描P帧和隔

行扫描 B 帧的帧级别比特流语法的示图。图 17 是示出用于隔行扫描 I 半帧的半帧级别比特流语法的示图。一些元素（例如，ACPRED 位平面）还出现于用于隔行扫描 BI 半帧的半帧级别比特流语法中，以及潜在地存在于其它图片类型的语法中。以下描述特定比特流元素。

帧编码模式（FCM）（可变大小）

FCM 是用于指示图片编码类型的可变长度代码字[VLC]。FCM 采用用于帧编码模式的值，如以下的表 1 所示：

表 1：帧编码模式 VLC

FCM 值	帧编码模式
0	逐行扫描
10	帧隔行扫描
11	半帧隔行扫描

半帧图片类型（FPTYPE）（3 比特）

FPTYPE 是包含隔行扫描 I 半帧和/或隔行扫描 BI 半帧以及潜在地包括其它种类半帧的帧的帧头中出现的三比特语法元素。根据以下的表 2，FPTYPE 采取隔行扫描视频帧中半帧类型的不同组合的值。

表 2：半帧图片类型 FLC

FPTYPE FLC	第一半帧类型	第二半帧类型
000	I	I
001	I	P
010	P	I
011	P	P
100	B	B
101	B	BI
110	BI	B
111	BI	BI

图片类型（PTYPE）(可变大小)

PTYPE 是用于隔行扫描 P 帧和隔行扫描 B 帧（或者诸如隔行扫描 I 帧的其它种类的隔行扫描帧）的帧头中存在的可变大小的语法元素。根据以下的表 3，PTYPE 采取用于不同帧类型的值。

表 3：图片类型 VLC

PTYPE VLC	图片类型
110	I
0	P
10	B
1110	BI
1111	跳过

如果 PTYPE 指示帧被跳过，则该帧被处理为与其参考帧相同的 P 帧。被跳过的帧的重建在概念上等效于复制参考帧。被跳过的帧意味着不为该帧传送进一步的数据。

半帧变换位平面（FIELDTX）（可变大小）

在帧级别或半帧级别，FIELDTX 是指示隔行扫描 I 帧内的宏块被帧编码还是半帧编码的位平面。在以下以及以上的章节 IV 中进一步详细说明了 FIELDTX。

AC 预测位平面（ACPRED）（可变大小）

对于逐行扫描 I 帧和 BI 帧且对于隔行扫描 I 帧和 BI 帧，使用位平面编码的语法元素来共同地编码将在所有宏块中存在的 1 比特 ACPRED 语法元素，该位平面编码的语法元素指示帧中每个宏块的 AC 预测状态。解码的位平面将用于每个宏块的 AC 预测状态表示为 1 比特值。以下以及在以上的章节 V 中进一步详细描述了 ACPRED 位平面。

3.选定的宏块层元素

图 15、21、22 和 23 分别是示出组合实现中用于逐行扫描 I 帧/隔行扫描 I 半帧/隔行扫描 BI 半帧、隔行扫描 I 帧/隔行扫描 BI 帧、隔行扫描 P 帧和隔行扫描 B

帧中的宏块的宏块级别比特流语法的示图。以下描述具体的比特流元素。用于宏块的数据由宏块头继之以块层数据构成。宏块层中的比特流元素（例如，ACPRED，FIELDTX 等）也可为其它图片类型的宏块而提供。

AC 预测标记（ACPRED）（1 比特）

在组合实现中，ACPRED 语法元素存在于所有逐行扫描 I 帧、隔行扫描 I 帧、隔行扫描 BI 帧、隔行扫描 I 半帧和隔行扫描 BI 半帧宏块，以及隔行扫描 P 半帧中的帧内宏块，隔行扫描 B 半帧，隔行扫描 P 帧，和隔行扫描 B 帧中。ACPRED 也存在于逐行扫描 P 帧和 B 帧中的宏块中。在宏块级别，ACPRED 是 1 比特语法元素，它指明宏块中的块是否用 AC 预测进行编码。ACPRED＝0 表示不使用 AC 预测。ACPRED＝1 表示使用 AC 预测。ACPRED 也可编码作为帧级别位平面，如以上章节 V 中进一步详细说明的。

半帧变换标记（FIELDTX）（1 比特）

FIELDTX 是隔行扫描 I 帧、隔行扫描 BI 帧、隔行扫描 P 帧和隔行扫描 B 帧中的帧内编码宏块中存在的 1 比特语法。该语法元素指示宏块是被帧编码还是被半帧编码（基本上，宏块的内部组织）。FIELDTX＝1 表示宏块被半帧编码。否则，宏块被帧编码。在帧间编码宏块中，该语法元素可从宏块级别比特流元素 MBMODE 中推断出。FIELDTX 也可以被编码为帧级别位平面，如以上章节 IV 中进一步详细说明的。

B.位平面编码

可以按每宏块一个二进制符号来编码诸如跳过比特的宏块特有的二进制信息。例如，可以用一个比特用信号表示一宏块是否被跳过。在这些情况下，用于半帧或帧中所有宏块的状态可以编码为一位平面并在半帧或帧的头中传送。该规则的一个例外在于如果位平面编码模式被设定为原始模式，在这种情况中每个宏块的状态被编码为每符号一个比特并与其它宏块级别语法元素一起在宏块级别处传送。

半帧/帧级别位平面编码被用于编码二维二进制数组。每个数组的大小是 rowMB×colMB，其中 rowMB 和 colMB 分别是所述半帧或帧中宏块行和列的数量。在比特流内，每个数组都被编码为一组连续比特。将七种模式之一用于编码每个数组。七种模式是：

1.原始模式－信息被编码为每符号一个比特并作为 MB 级语法的一部分传送；

2.正常－2 模式－两个符号被共同编码；

3.差分－2 模式－位平面的差分编码，继之以共同编码两个残差符号；

4.正常－6 模式－共同编码六个符号；

5.差分－6 模式－位平面的差分编码，继之以共同编码六个残差符号；

6.行跳过模式－－一个跳过比特以用信号表示没有设定比特的行；以及

7.列跳过模式－－一个跳过比特以用信号表示没有设定比特的列。

半帧或帧级别的位平面的语法元素位于以下序列中：INVERT、IMODE 和 DATABITS。

逆转标记（INVERT）

INVERT 语法元素是 1 比特值，它如果被设定就指示位平面的设定比特多于零比特。根据 INVERT 和模式，解码器将逆转所译码的位平面以重新形成原始的位平面。注意，在使用原始模式时将忽略该比特的值。以下提供在解码位平面时如何使用 INVERT 值的描述。

编码模式（IMODE）

IMODE 语法元素是可变长度值，它指示用于编码位平面的编码模式。表 4 示出了用于编码 IMODE 语法元素的代码表。以下提供解码位平面时如何使用 IMODE 值的描述。

表 4：IMODE VLC 代码表

IMODE VLC	编码模式
10	正常－2
11	正常－6
010	行跳过
011	列跳过
001	差分－2
0001	差分－6
0000	原始

位平面编码比特（DATABITS）

DATABITS 语法元素是编码用于位平面的符号流的可变大小的语法元素。通过 IMODE 的值来确定用于编码位平面的方法。以下部分中描述了七种编码模式。

原始模式

在该模式中，位平面被编码为每符号一个比特（例如，以光栅扫描顺序），并作为宏块层的一部分发送。或者，以半帧或帧级别按原始模式编码信息且 DATABITS 在长度上是 rowMB×colMB。

正常－2 模式

如果 rowMB×colMB 是奇数，则第一符号被原始编码。后续符号被成对编码，按自然扫描顺序。表 5 中的二进制 VLC 表用于编码符号对。

表 5：正常－2/差分－2 代码表

符号 2n	符号 2n+1	代码字
0	0	0
1	0	100
0	1	101
1	1	11

差分－2 模式

如上所述，正常－2 方法用于产生位平面，且随后将 $Diff^1$ 运算应用于该位平面，如下所述。

正常－6 模式

在正常－6 和差分－6 模式中，以六个像素一组编码位平面。这些像素被组成为 2×3 或者 3×2 块。使用一组规则最大地平铺位平面，且其余像素用行跳过和列跳过模式的变型进行编码。如果且仅如果 rowMB 是 3 的倍数而 colMB 不是，则使用 2×3“垂直”块。否则，使用 3×2“水平”块。图 24A 示出了 2×3“垂直”块的简化示例。图 24B 和 24C 示出了 3×2“水平”块的简化示例，其中伸长的暗矩形是 1 像素宽并用行跳过和列跳过编码进行编码。对于如图 24C 所示地平铺的平面，采用沿图片的顶边和左边的线性平铺，块的编码顺序遵循以下模式。首先编码 6 元素块，继之以列跳过和行跳过编码的线性块。如果数组大小是 2×3 或 3×2 的

倍数，则后者线性块不存在并完美地平铺位平面。

用可变长度和固定长度代码的组合来编码 6 元素矩形块。设 N 为块中设定比特的数量，即 0≤N≤6。对于 N<3，VLC 被用于编码块。对于 N=3，固定长度转义码继之以 5 比特固定长度代码。对于 N>3，另一固定长度转义码继之以 VLC。对于 N>3，跟随转义码的 VLC 与在 N<3 情况下用于编码该块的补码的 VLC 相同。用于 N>3 情况的固定长度转义码不同于用于 N=3 情况的固定长度转义码。矩形块包含 6 比特的信息。使得 k 作为与块相关联的代码，其中 $k=\sum_i b_i 2^i$，b_i 是块内按自然扫描顺序的第 i 个比特的二进制值。因此，0≤k<64。VLC 和转义码的组合加上固定长度代码被用于用信号表示 k。

差分－6 模式

如上所述，正常－6 模式被用于产生位平面，随后将 $Diff^1$ 运算应用于该位平面，如下所述。

行跳过模式

在行跳过编码模式中，用一个比特的开销跳过全零行。该语法如下：对于每个行，单个 ROWSKIP 比特指示是否跳过该行；如果跳过该行，则接着处理用于下一个行的 ROWSKIP 比特；否则（不跳过该行），接着处理 ROWBITS 比特（行中用于每个宏块的比特）。因此，如果整行是零，则发送零比特作为 ROWSKIP 符号，且跳过 ROWBITS。如果行中存在一个设定比特，则将 ROWSKIP 设为 1，且原始地发送整行（ROWBITS）。在半帧或帧中从上到下地扫描行。

列跳过模式

列跳过是行跳过的转置。在半帧或帧中从左到右地扫描列。

$Diff^1$：：反差分解码

如果使用任一差分模式（差分－2 或差分－6），则首先用相应的正常模式（分别是正常－2 或正常－6）来解码“差分比特”的位平面。差分比特被用于重新生成原始位平面。重新生成过程是二进制字母表上的 2-D DPCM。为了在位置（i,j）处重新生成比特，如下地生成预测值 b_p（i,j）（在位置（i,j）处从比特 b（i,j））：

$$b_p(i,j)=\begin{cases} A & i=j=0,\ 或 b(i,j-1)\neq b(i-1,j) \\ b(0,j-1) & i=0 \\ b(i-1,j) & 其它 \end{cases}$$

对于差分编码模式，不进行基于 INVERT 的按比特逆转过程。但是，以不同的容量使用 INVERT 标记以指示用于以上示出的预测值的导出的符号 A 的值。更具体地，如果 INVERT=0 则 A＝0，且如果 INVERT=1 则 A＝1。通过用解码的差分比特值对预测值求异或而获得位平面的实际值。在以上公式中，b(i,j)是最终解码后（即在进行正常－2/正常－6 继之以用其预测值的差分异或之后）第 i,j 位置处的比特。

已参考各种实施例描述和说明了本发明的原理，可以理解，可以在结构和细节方面修改各种实施例而不背离这些原理。应理解，除非另外指出，这里的程序、过程或方法不涉及或不限于任何特殊类型的计算环境。根据这里所述的教导使用各种类型的通用或专用计算环境或进行操作。以软件示出的实施例元素可在硬件中实现，反之亦然。

考虑到可应用本发明原理的许多可能的实施例，我们声明所有在所附权利要求书及其等效物范围和精神内的实施例都是我们的发明。

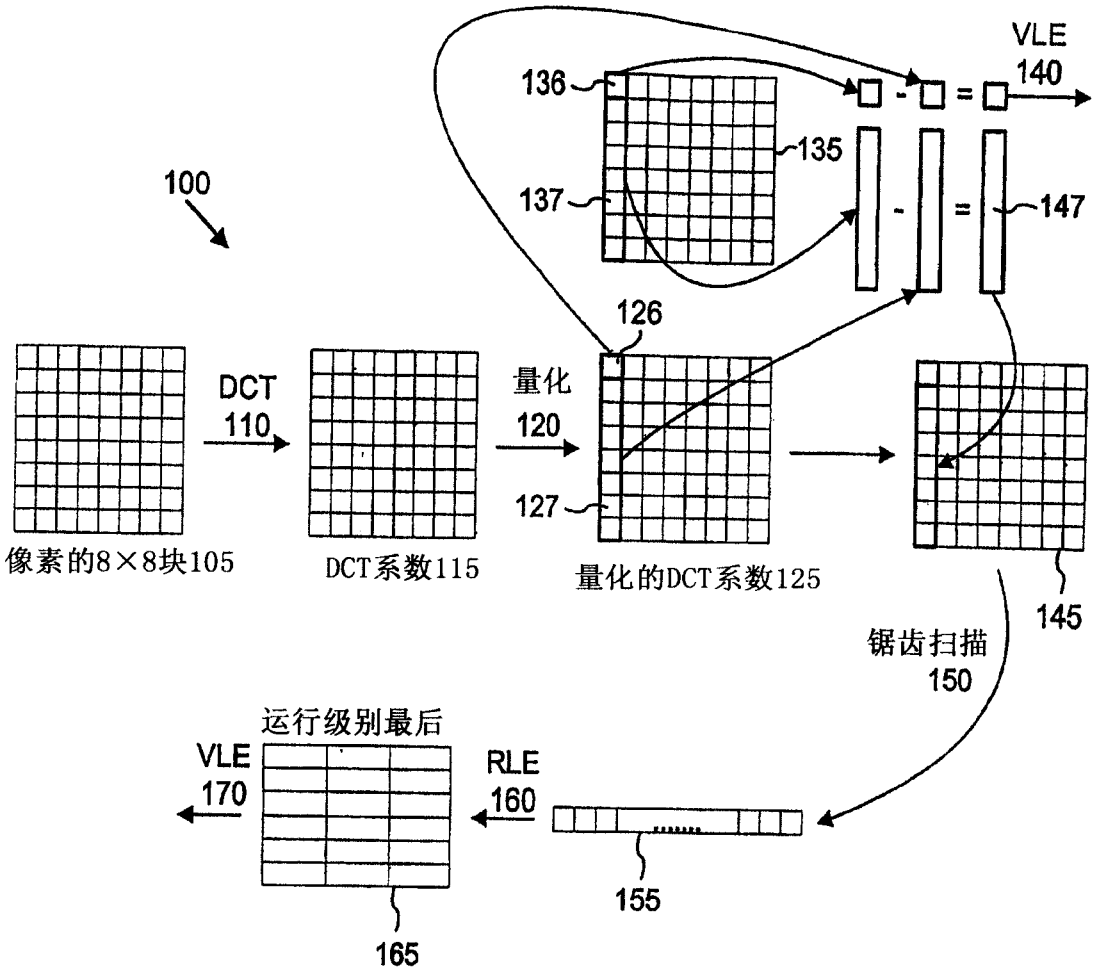

100
DCT
110
量化
120
像素的8×8块105
DCT系数115
量化的DCT系数125
126
127
136
135
137
147
VLE
140
145
锯齿扫描
150
155
RLE
160
运行级别最后
VLE
170
165

图 1A

(现有技术)

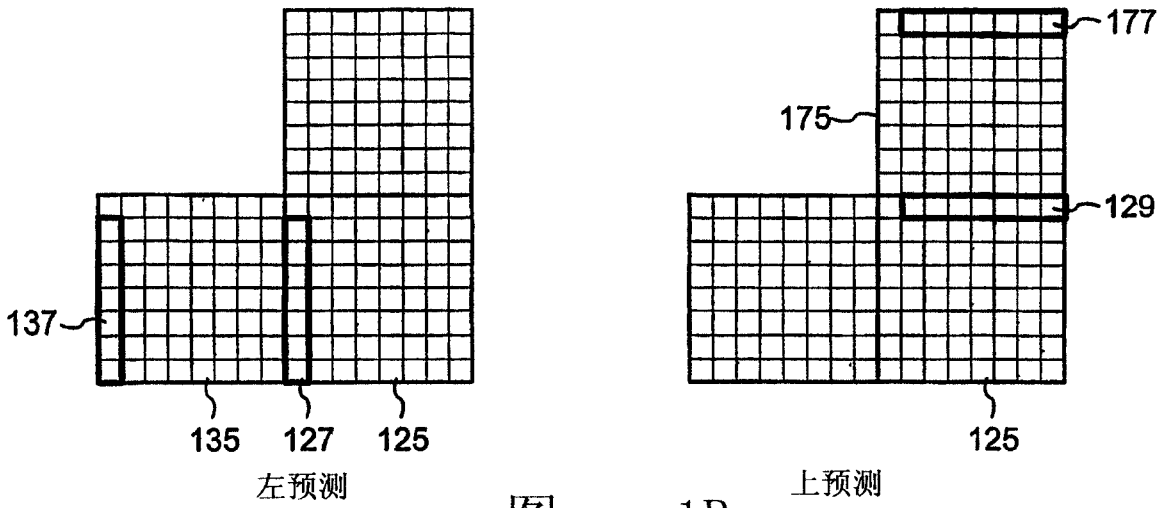


图 1B

(现有技术)

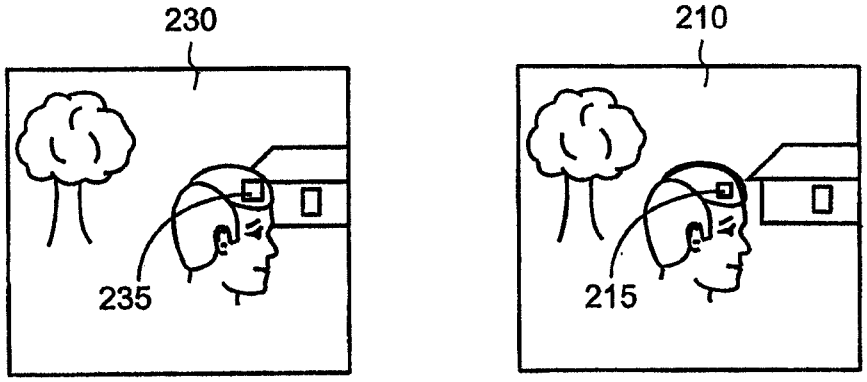


图 2

(现有技术)

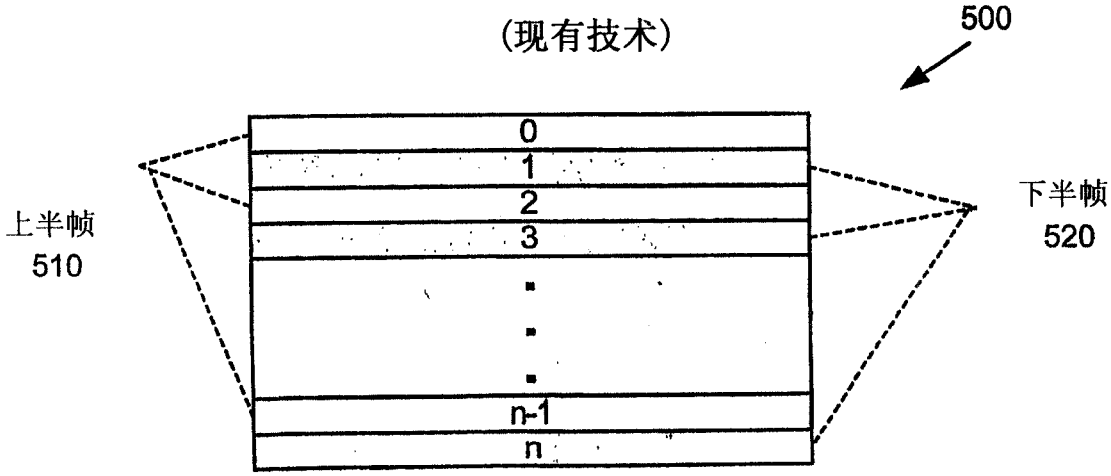
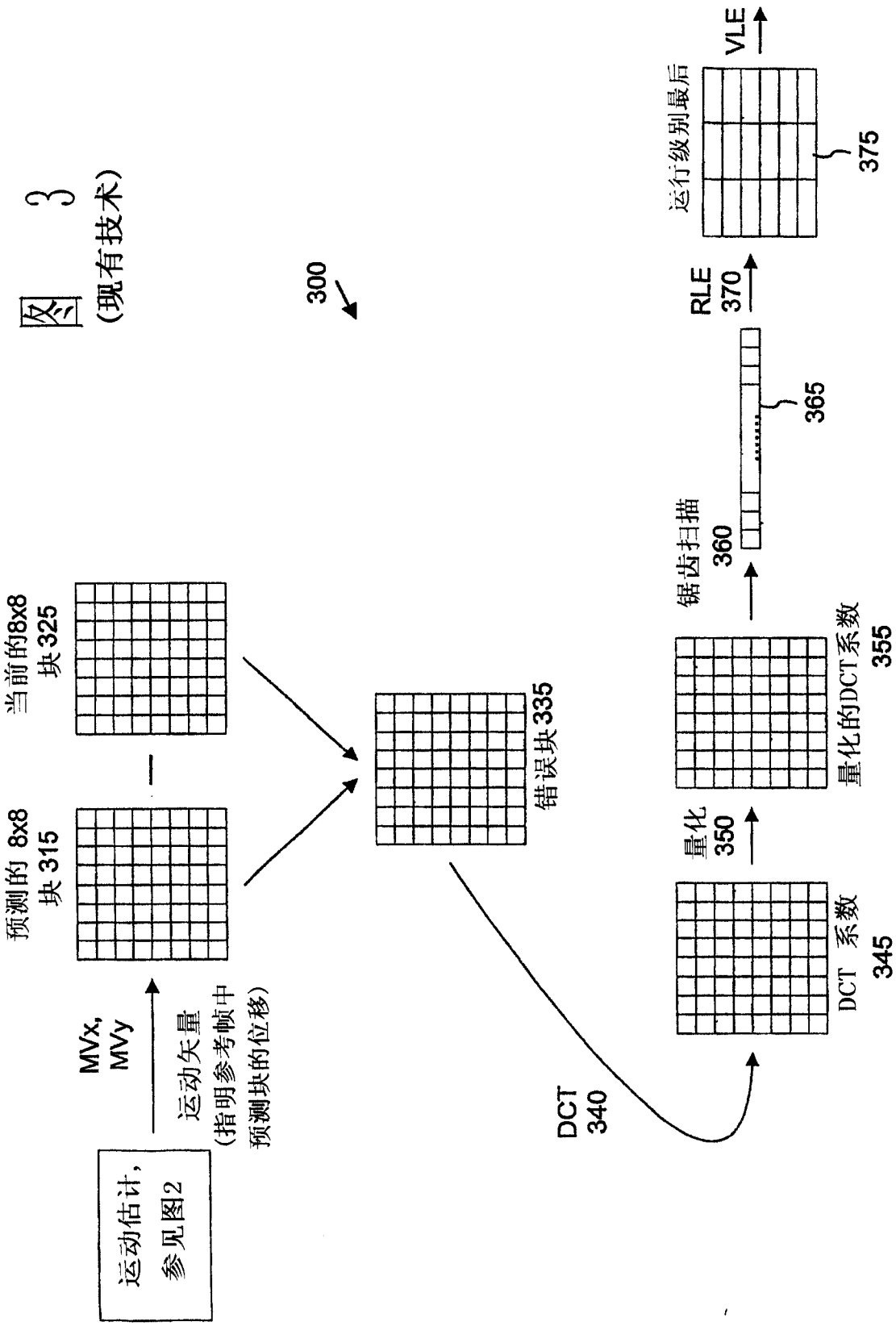


图 5

(现有技术)

图 3
(现有技术)


300
当前的8x8
块 325
预测的 8x8
块 315
—
MVx,
MVy
运动矢量
(指明参考帧中
预测块的位移)
运动估计,
参见图2
错误块335
DCT
340
DCT 系数
345
量化
350
量化的DCT系数
355
锯齿扫描
360
365
RLE
370
运行级别最后
375
VLE

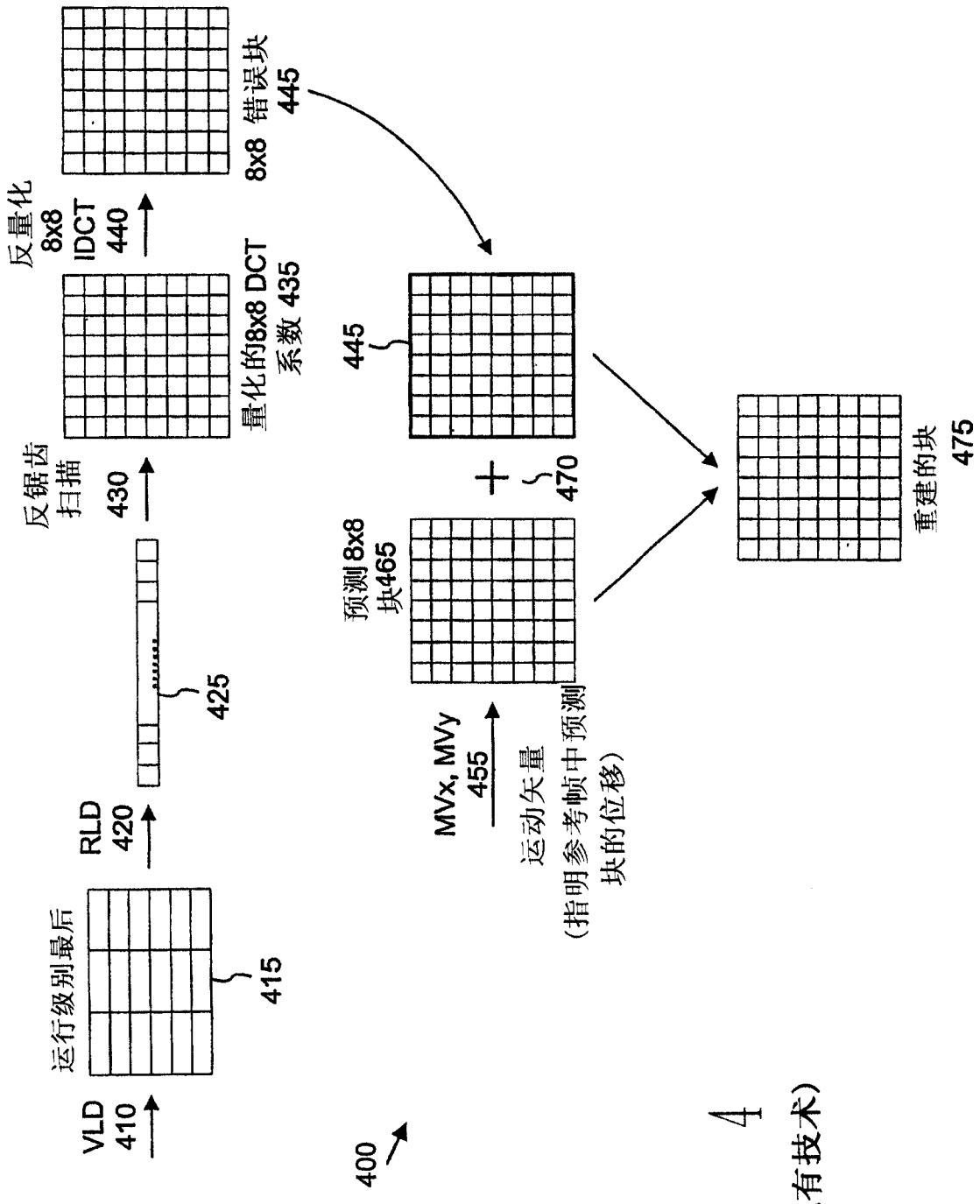

VLD
410
运行级别最后
415
RLD
420
425
反锯齿
扫描
430
量化的8x8 DCT
系数 435
反量化
8x8
IDCT
440
8x8 错误块
445
400
MVx, MVy
455
运动矢量
(指明参考帧中预测
块的位移)
预测 8x8
块465
470
445
重建的块
475

图 4
(现有技术)

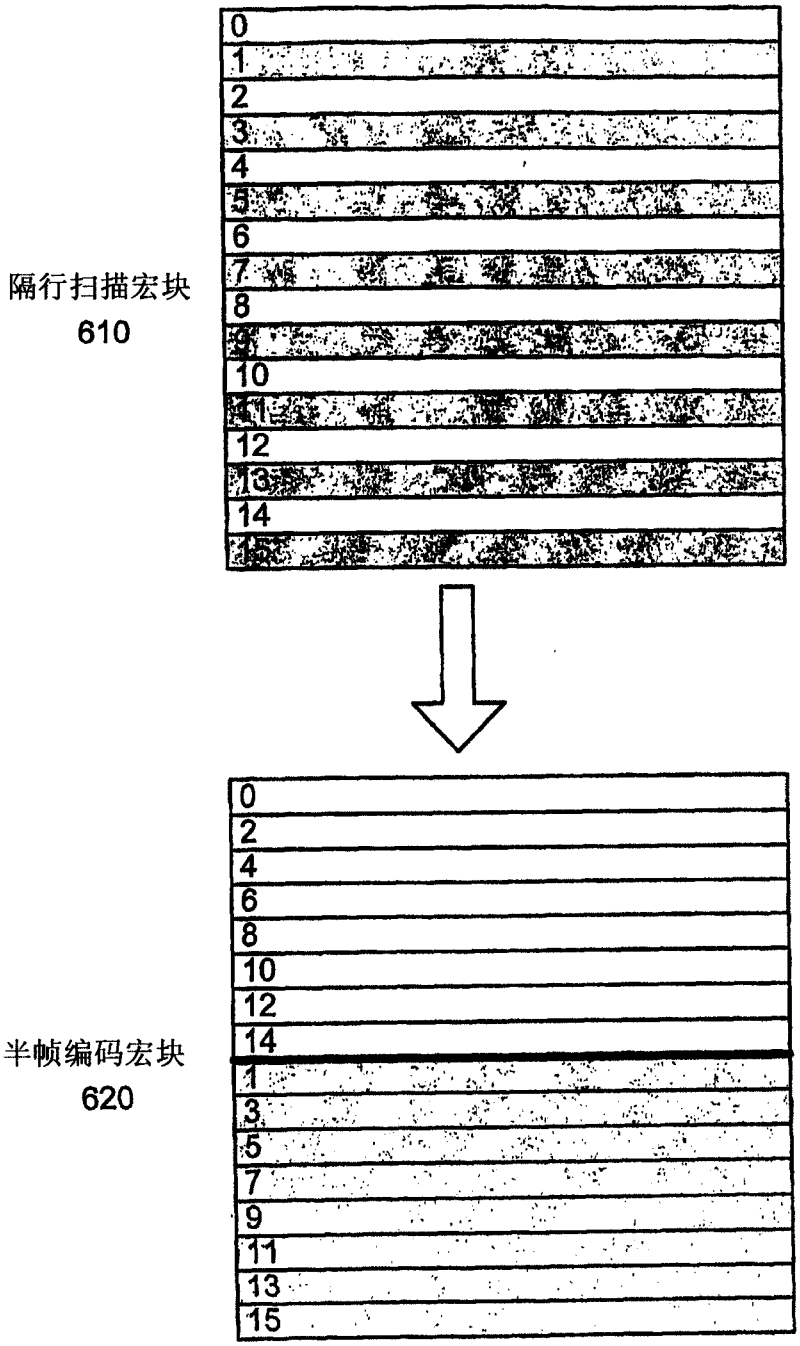

隔行扫描宏块
610
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
半帧编码宏块
620
0
2
4
6
8
10
12
14
1
3
5
7
9
11
13
15

图 6

(现有技术)

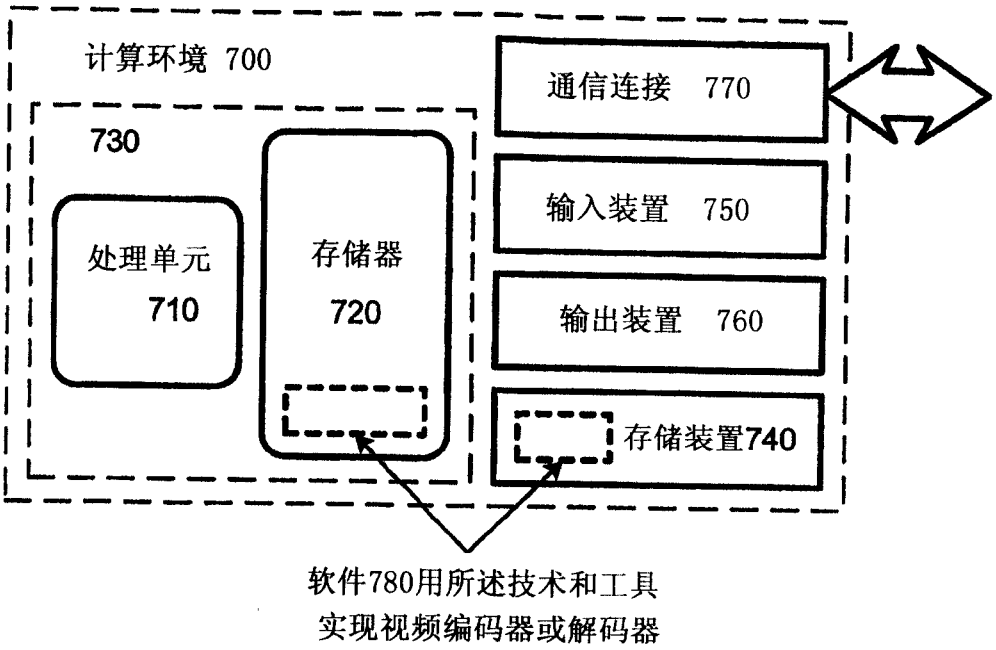

计算环境 700
730
处理单元
710
存储器
720
通信连接 770
输入装置 750
输出装置 760
存储装置740
软件780用所述技术和工具
实现视频编码器或解码器

图 7

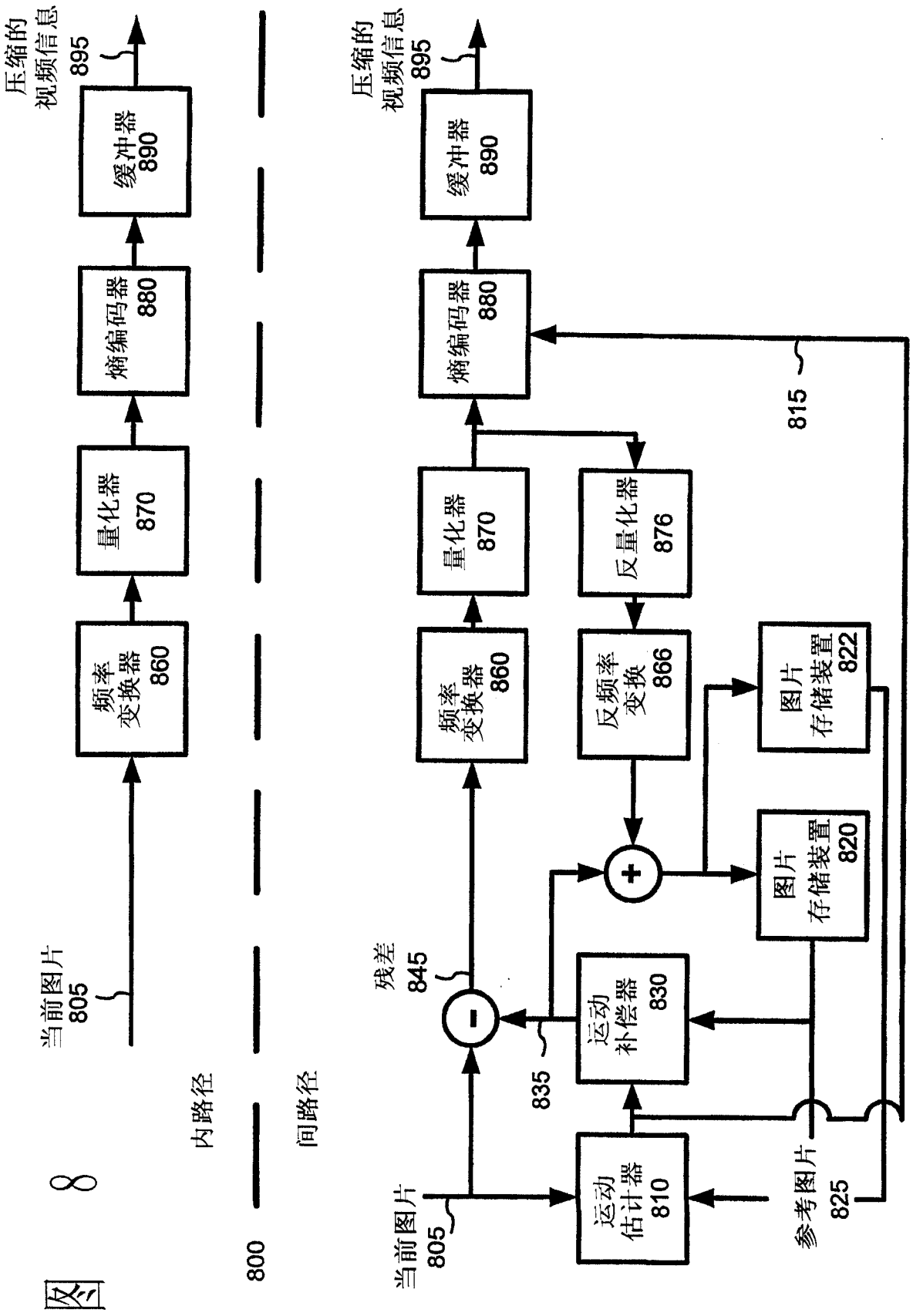

压缩的视频信息 895
缓冲器 890
熵编码器 880
量化器 870
频率变换器 860
当前图片 805
内路径
间路径
800
压缩的视频信息 895
缓冲器 890
熵编码器 880
815
量化器 870
反量化器 876
频率变换器 860
反频率变换 866
图片存储装置 822
图片存储装置 820
残差 845
运动补偿器 830
835
当前图片 805
运动估计器 810
参考图片 825

图 8

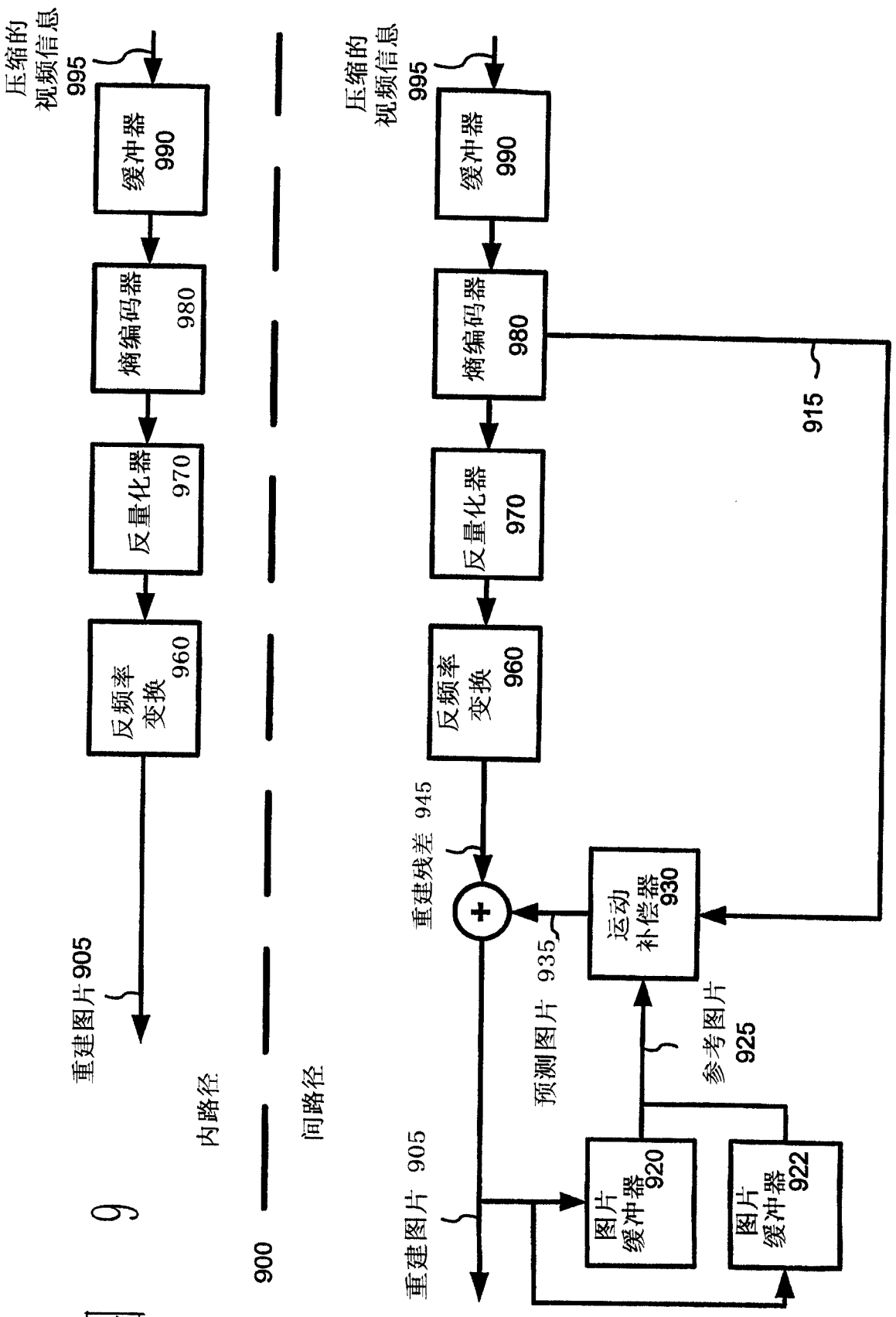

压缩的视频信息 995
缓冲器 990
熵编码器 980
反量化器 970
反频率变换 960
重建图片 905
内路径
间路径
900
压缩的视频信息 995
缓冲器 990
熵编码器 980
915
反量化器 970
反频率变换 960
重建残差 945
运动补偿器 930
预测图片 935
参考图片 925
重建图片 905
图片缓冲器 920
图片缓冲器 922

图 9

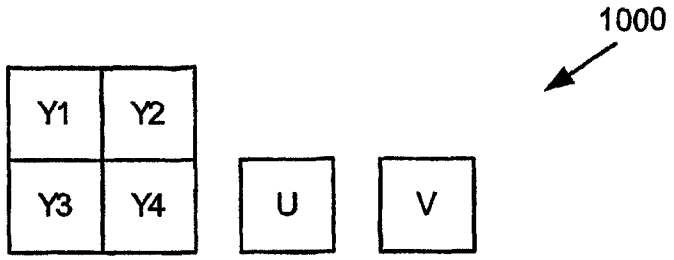


图 10

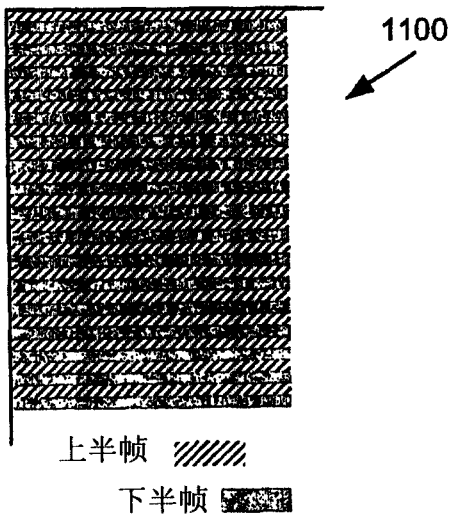


图 11A

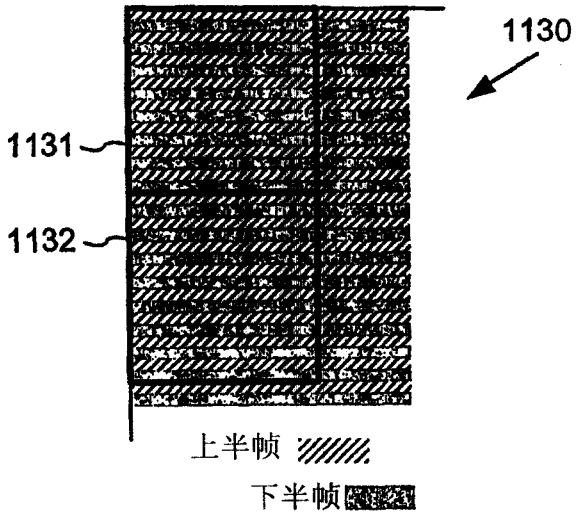


图 11B

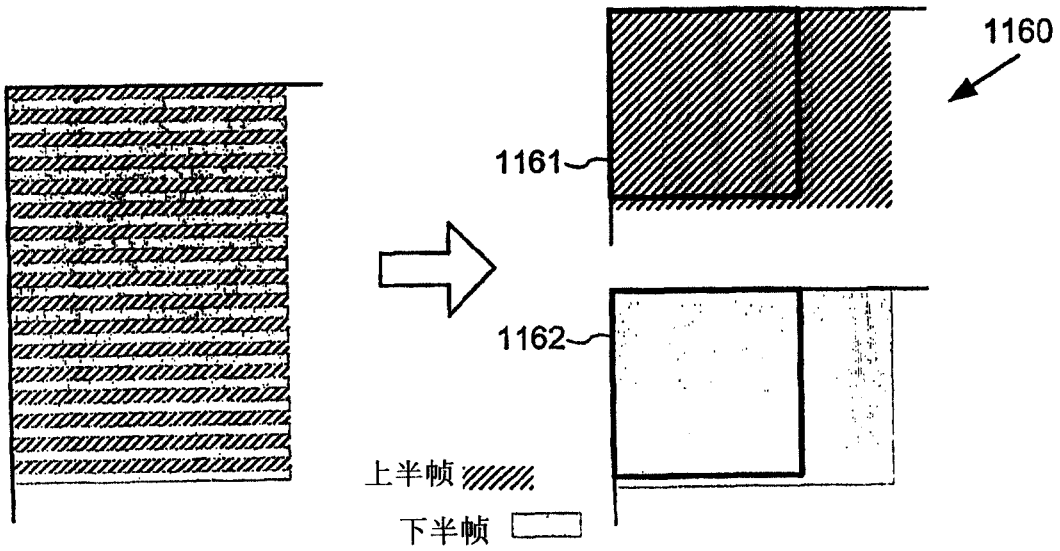


图 11C

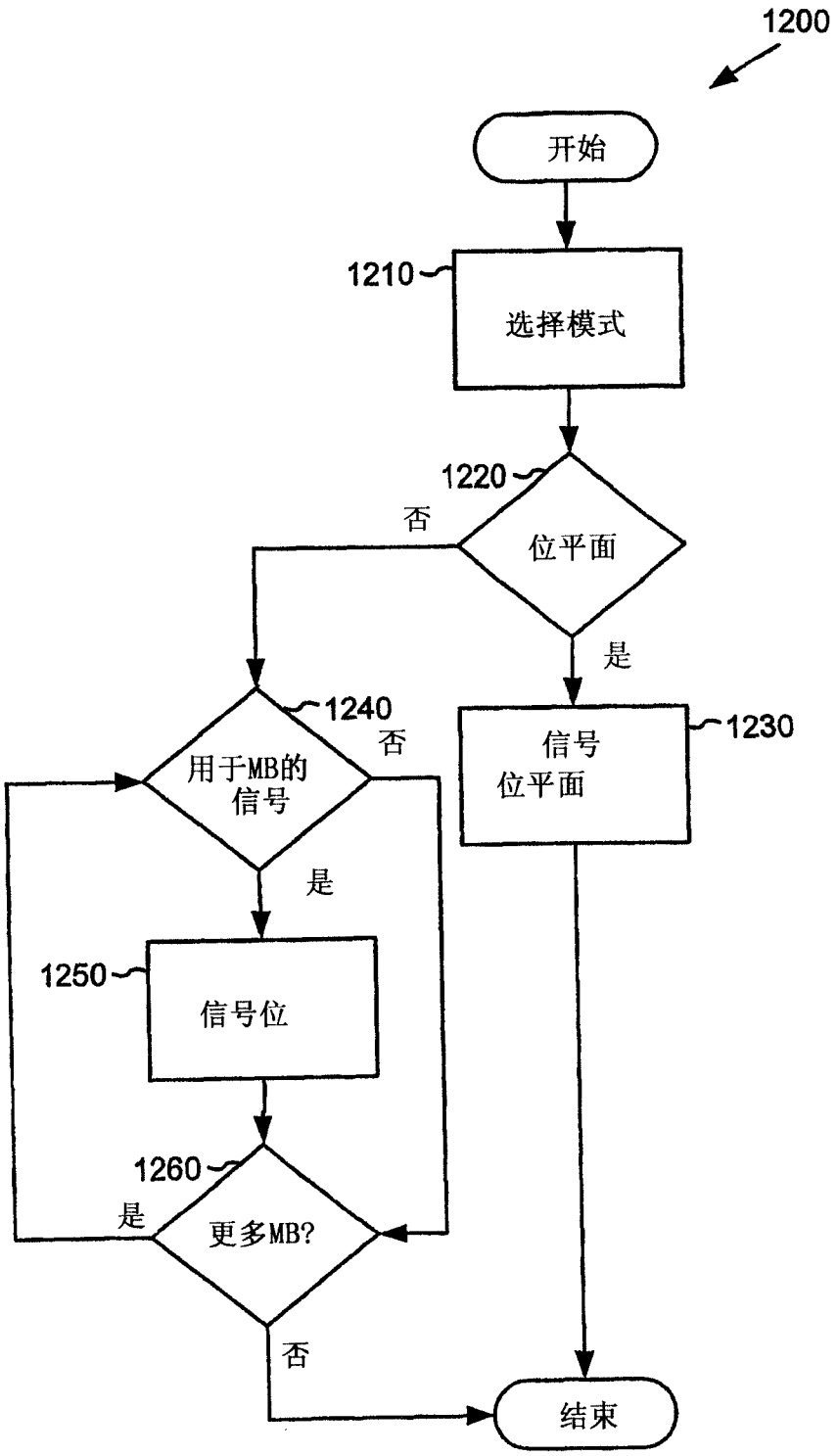

1200
开始
1210
选择模式
1220
位平面
否
是
1240
用于MB的信号
否
是
1230
信号位平面
1250
信号位
1260
更多MB?
是
否
结束

图 12

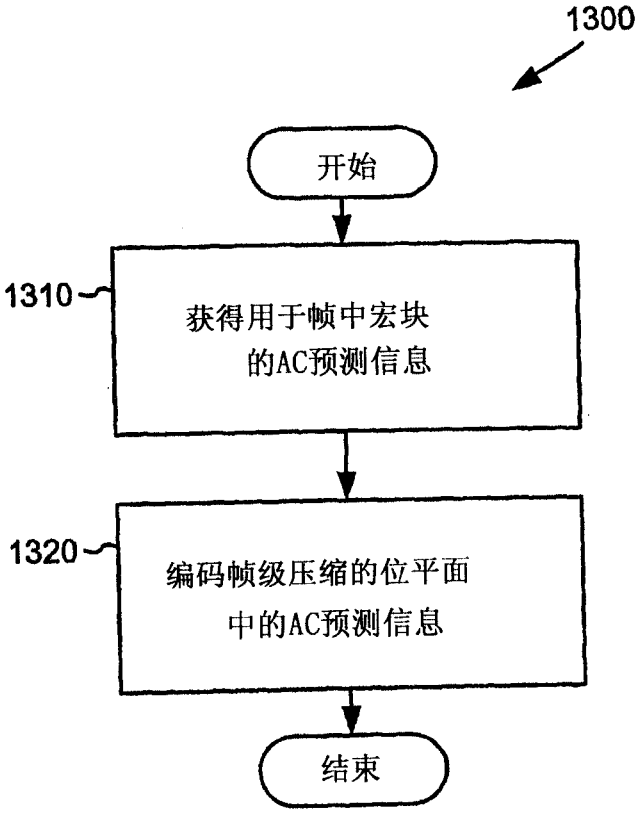


图 13

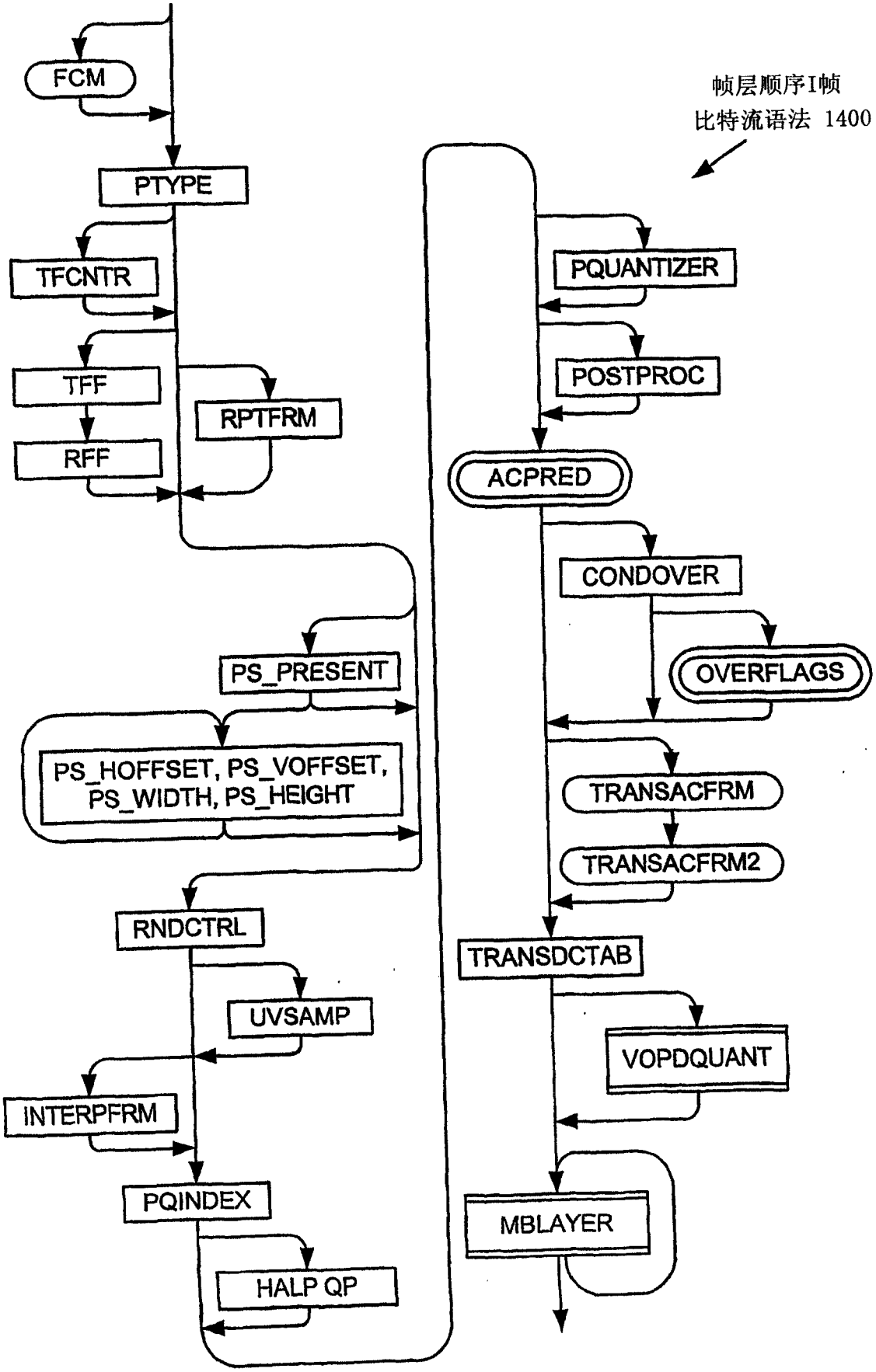

帧层顺序I帧
比特流语法 1400
FCM
PTYPE
TFCNTR
TFF
RPTFRM
RFF
PS_PRESENT
PS_HOFFSET, PS_VOFFSET,
PS_WIDTH, PS_HEIGHT
RNDCTRL
UVSAMP
INTERPFRM
PQINDEX
HALP QP
PQUANTIZER
POSTPROC
ACPRED
CONDOVER
OVERFLAGS
TRANSACFRM
TRANSACFRM2
TRANSDCTAB
VOPDQUANT
MBLAYER

图 14

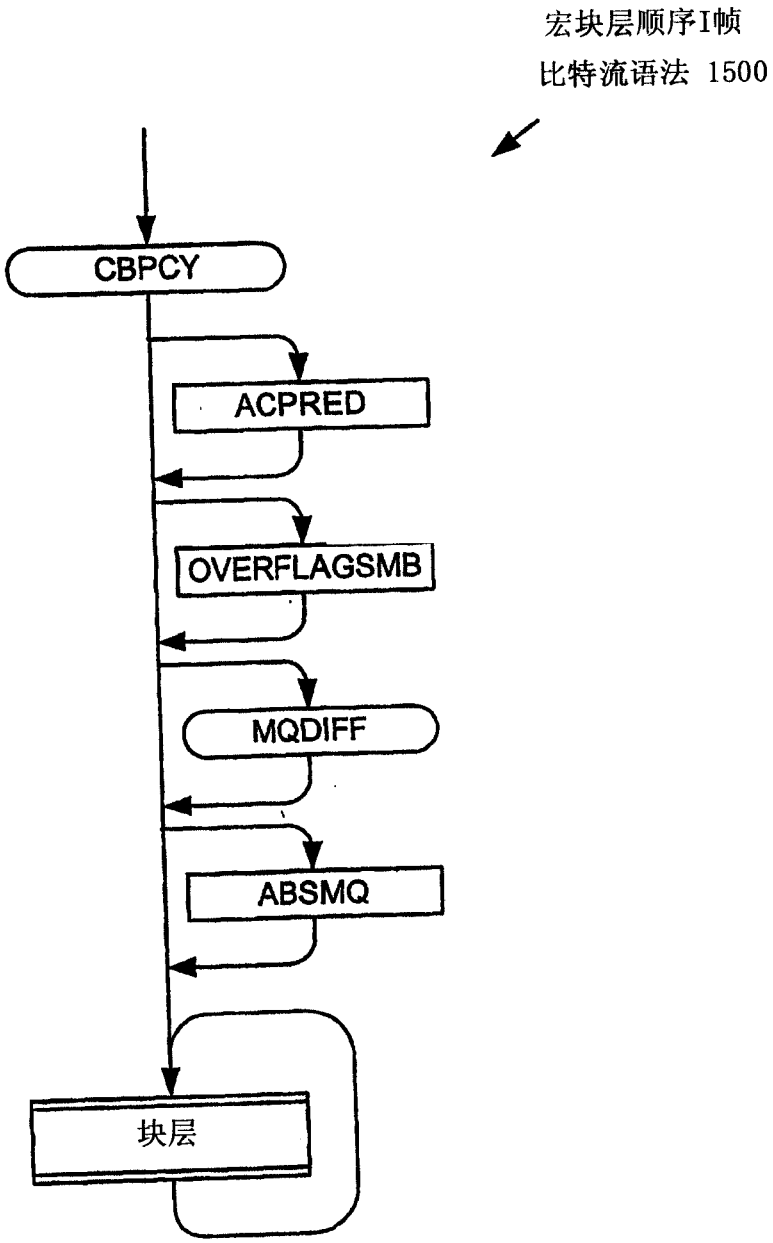

宏块层顺序I帧
比特流语法 1500
CBPCY
ACPRED
OVERFLAGSMB
MQDIFF
ABSMQ
块层

图 15

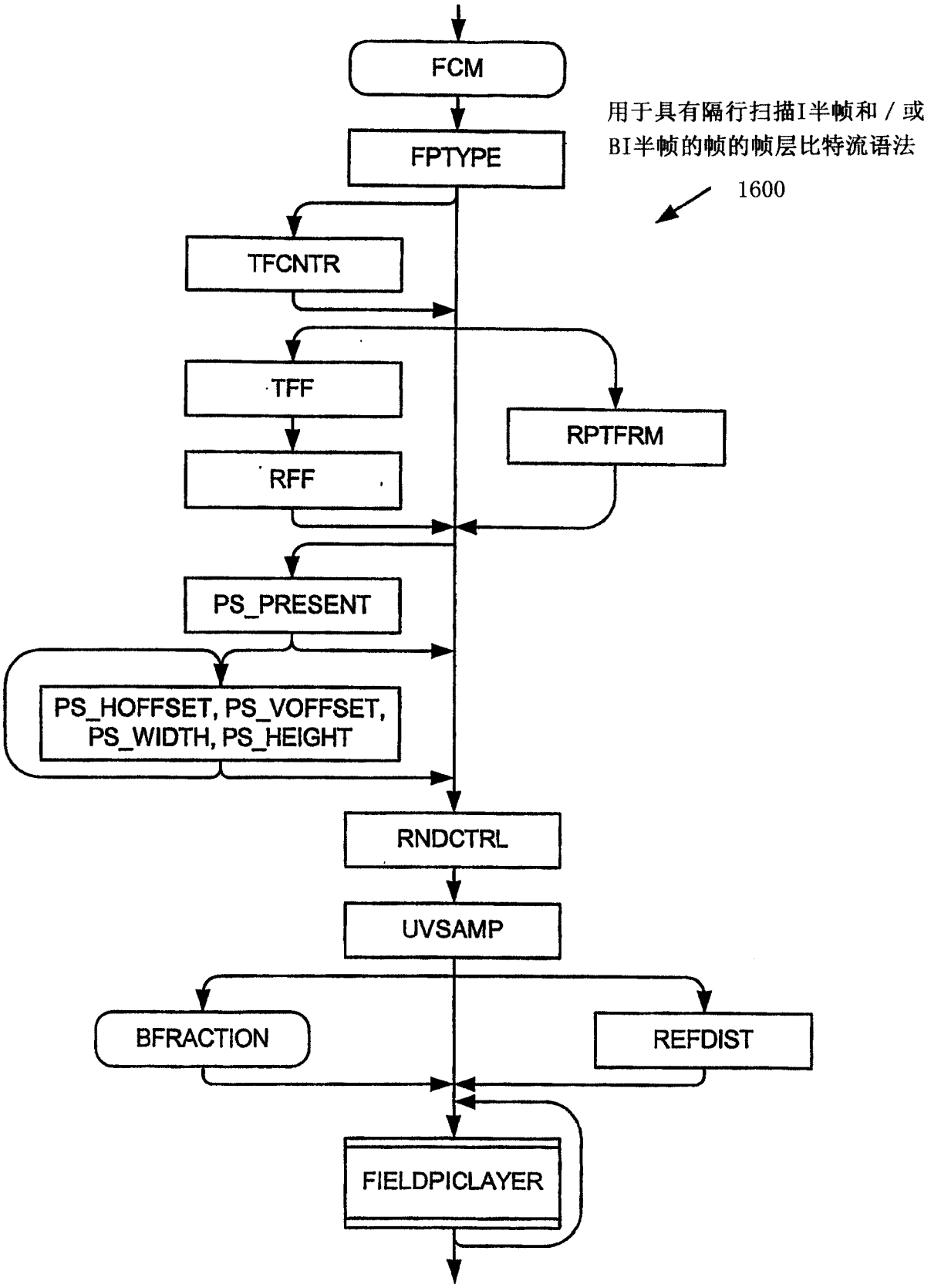

FCM
FPTYPE
用于具有隔行扫描I半帧和/或BI半帧的帧的帧层比特流语法
1600
TFCNTR
TFF
RFF
RPTFRM
PS_PRESENT
PS_HOFFSET, PS_VOFFSET, PS_WIDTH, PS_HEIGHT
RNDCTRL
UVSAMP
BFRACTION
REFDIST
FIELDPICLAYER

图 16

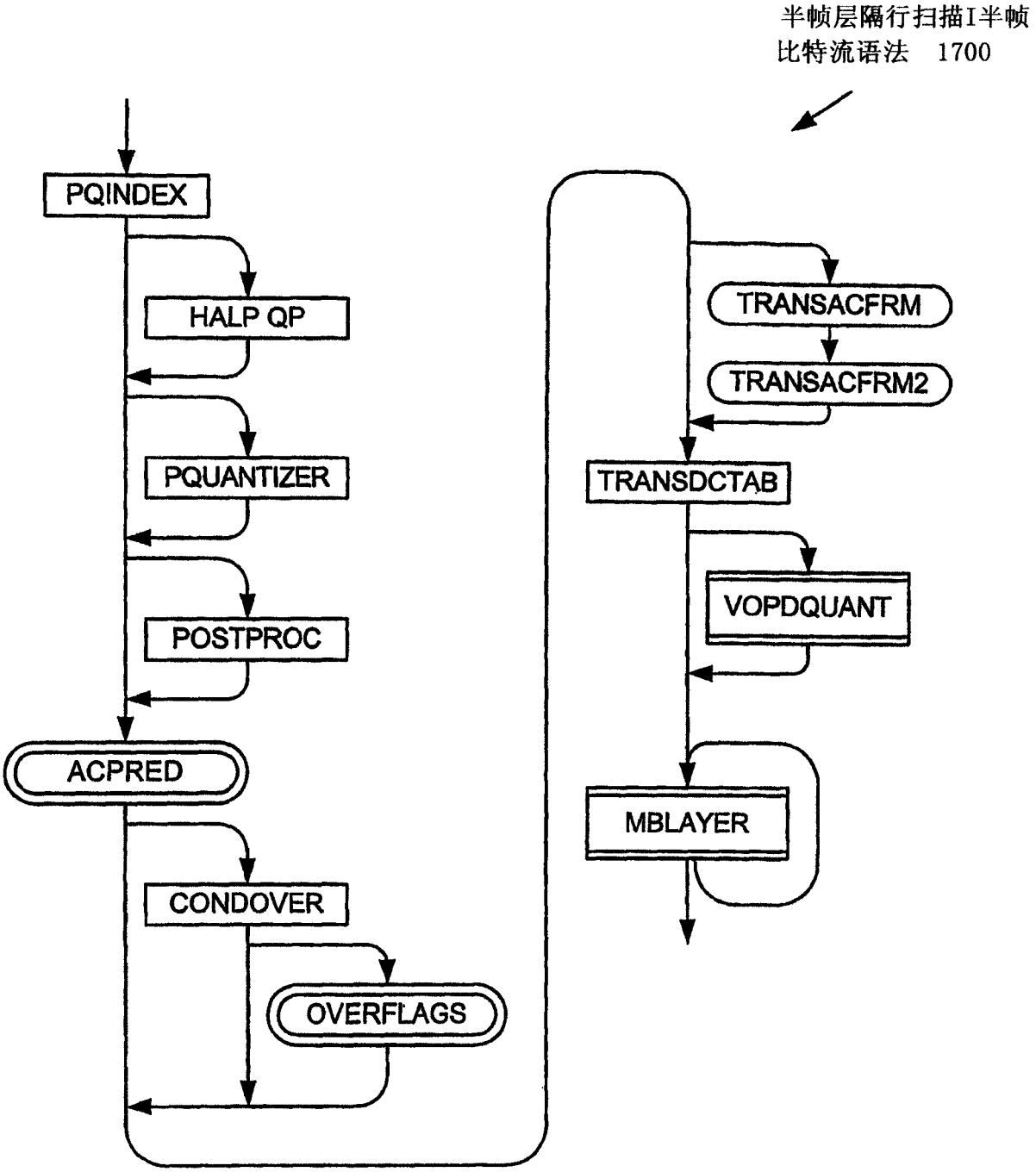

半帧层隔行扫描I半帧
比特流语法 1700
PQINDEX
HALP QP
PQUANTIZER
POSTPROC
ACPRED
CONDOVER
OVERFLAGS
TRANSACFRM
TRANSACFRM2
TRANSDCTAB
VOPDQUANT
MBLAYER

图 17

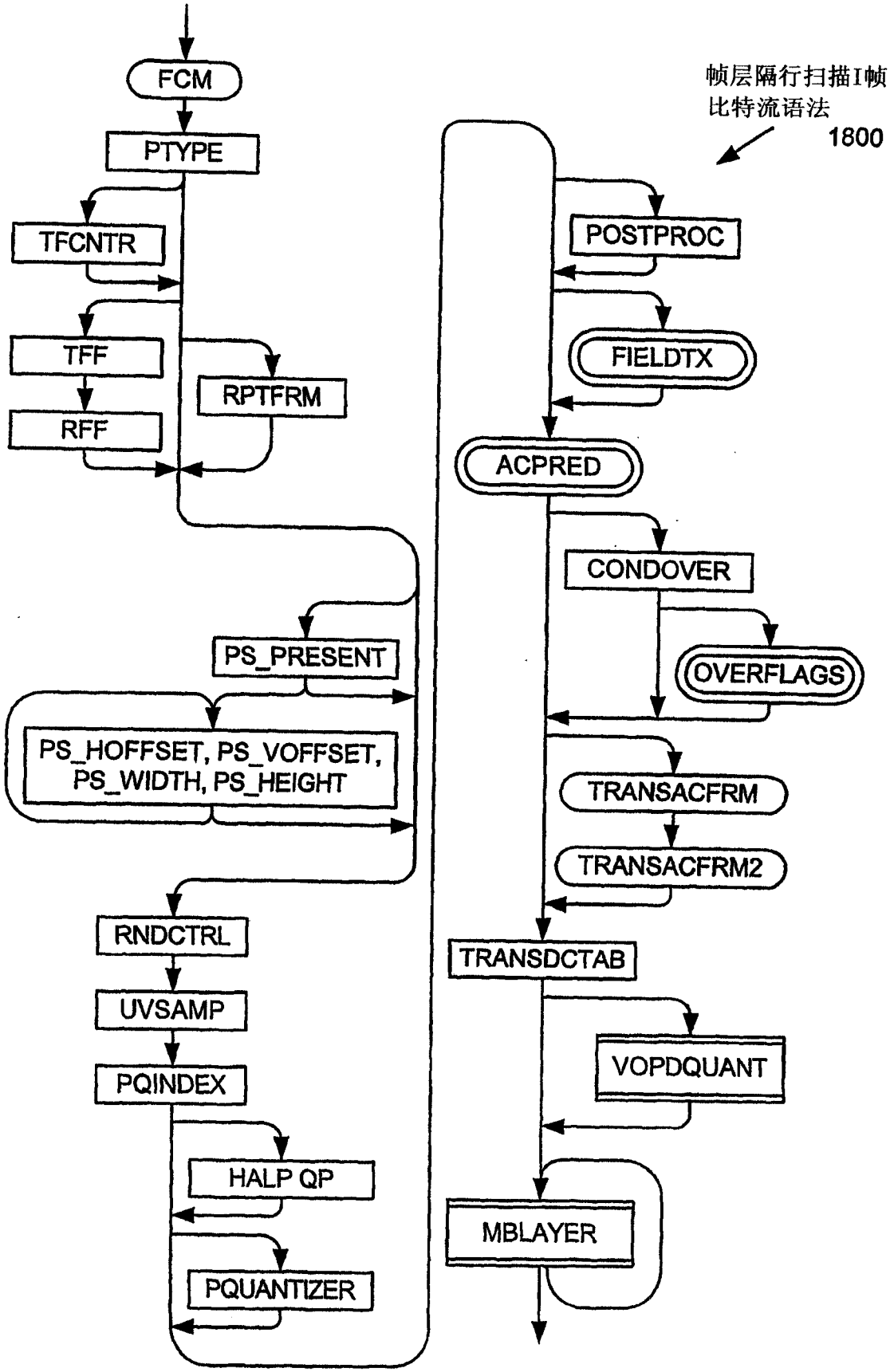

帧层隔行扫描I帧
比特流语法
1800
FCM
PTYPE
TFCNTR
TFF
RFF
RPTFRM
PS_PRESENT
PS_HOFFSET, PS_VOFFSET,
PS_WIDTH, PS_HEIGHT
RNDCTRL
UVSAMP
PQINDEX
HALP QP
PQUANTIZER
POSTPROC
FIELDTX
ACPRED
CONDOVER
OVERFLAGS
TRANSACFRM
TRANSACFRM2
TRANSDCTAB
VOPDQUANT
MBLAYER

图 18

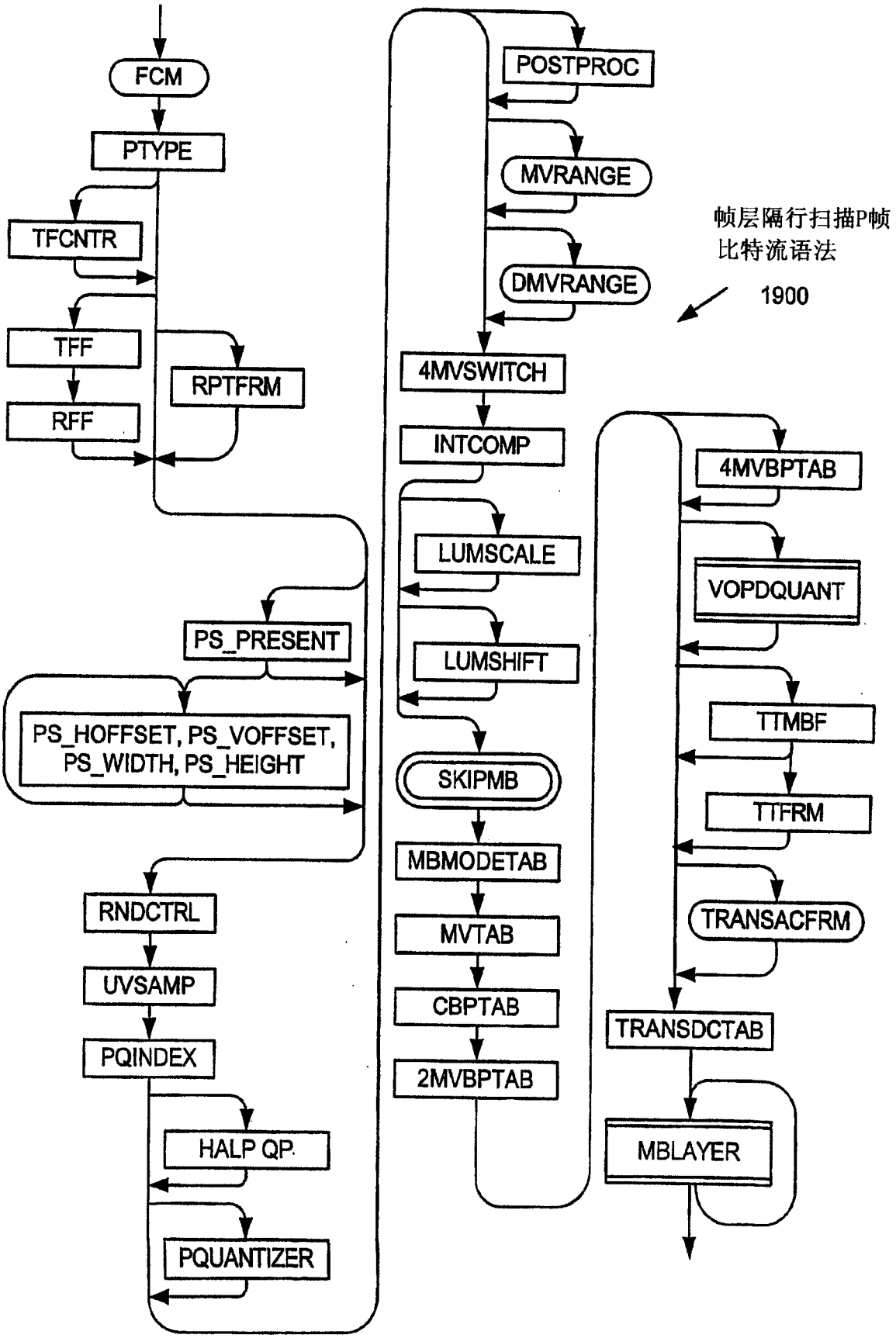

帧层隔行扫描P帧
比特流语法
1900
FCM
PTYPE
TFCNTR
TFF
RFF
RPTFRM
PS_PRESENT
PS_HOFFSET, PS_VOFFSET,
PS_WIDTH, PS_HEIGHT
RNDCTRL
UVSAMP
PQINDEX
HALP QP
PQUANTIZER
POSTPROC
MVRANGE
DMVRANGE
4MVSWITCH
INTCOMP
LUMSCALE
LUMSHIFT
SKIPMB
MBMODETAB
MVTAB
CBPTAB
2MVBPTAB
4MVBPTAB
VOPDQUANT
TTMBF
TTFRM
TRANSACFRM
TRANSDCTAB
MBLAYER

图 19

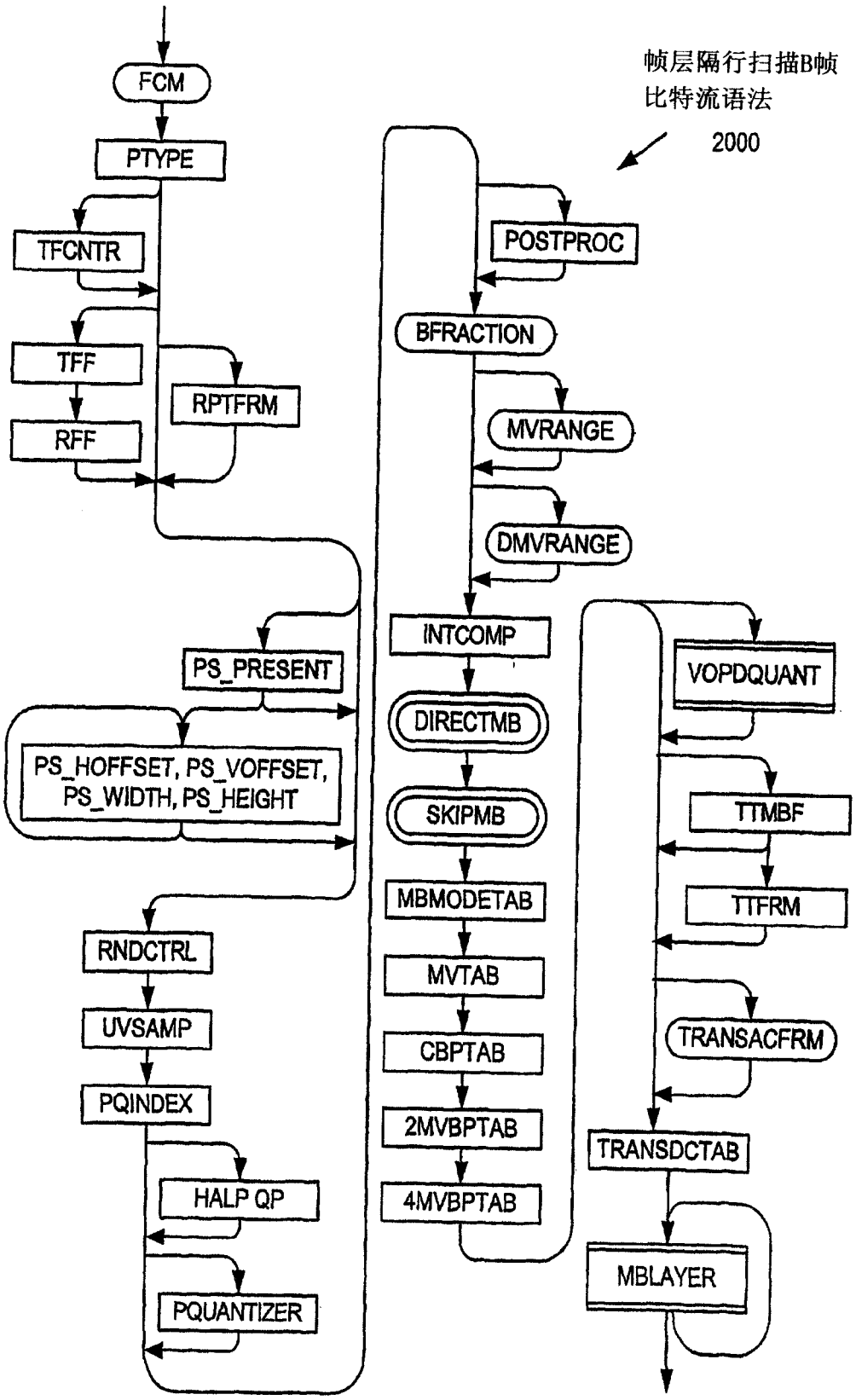

帧层隔行扫描B帧
比特流语法
2000
FCM
PTYPE
TFCNTR
TFF
RFF
RPTFRM
PS_PRESENT
PS_HOFFSET, PS_VOFFSET,
PS_WIDTH, PS_HEIGHT
RNDCTRL
UVSAMP
PQINDEX
HALP QP
PQUANTIZER
POSTPROC
BFRACTION
MVRANGE
DMVRANGE
INTCOMP
DIRECTMB
SKIPMB
MBMODETAB
MVTAB
CBPTAB
2MVBPTAB
4MVBPTAB
VOPDQUANT
TTMBF
TTFRM
TRANSACFRM
TRANSDCTAB
MBLAYER

图 20

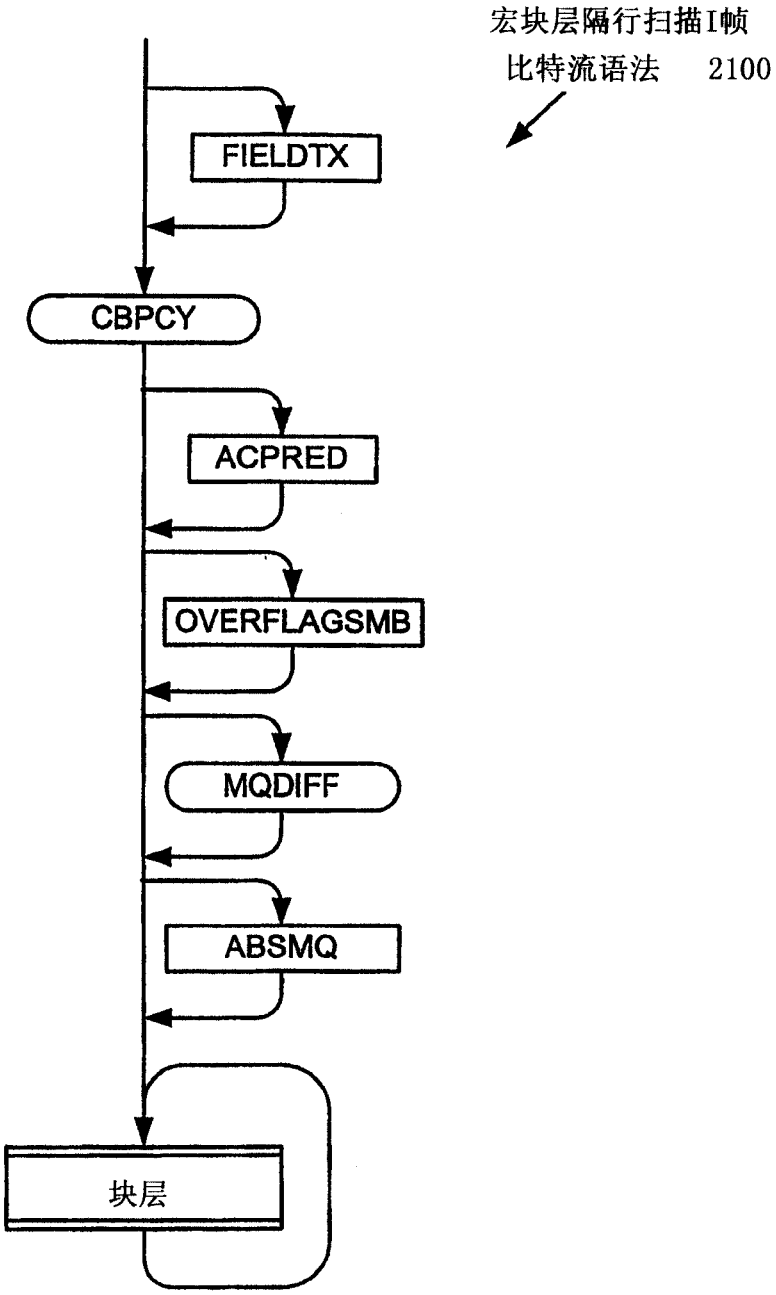

宏块层隔行扫描I帧
比特流语法 2100
FIELDTX
CBPCY
ACPRED
OVERFLAGSMB
MQDIFF
ABSMQ
块层

图 21

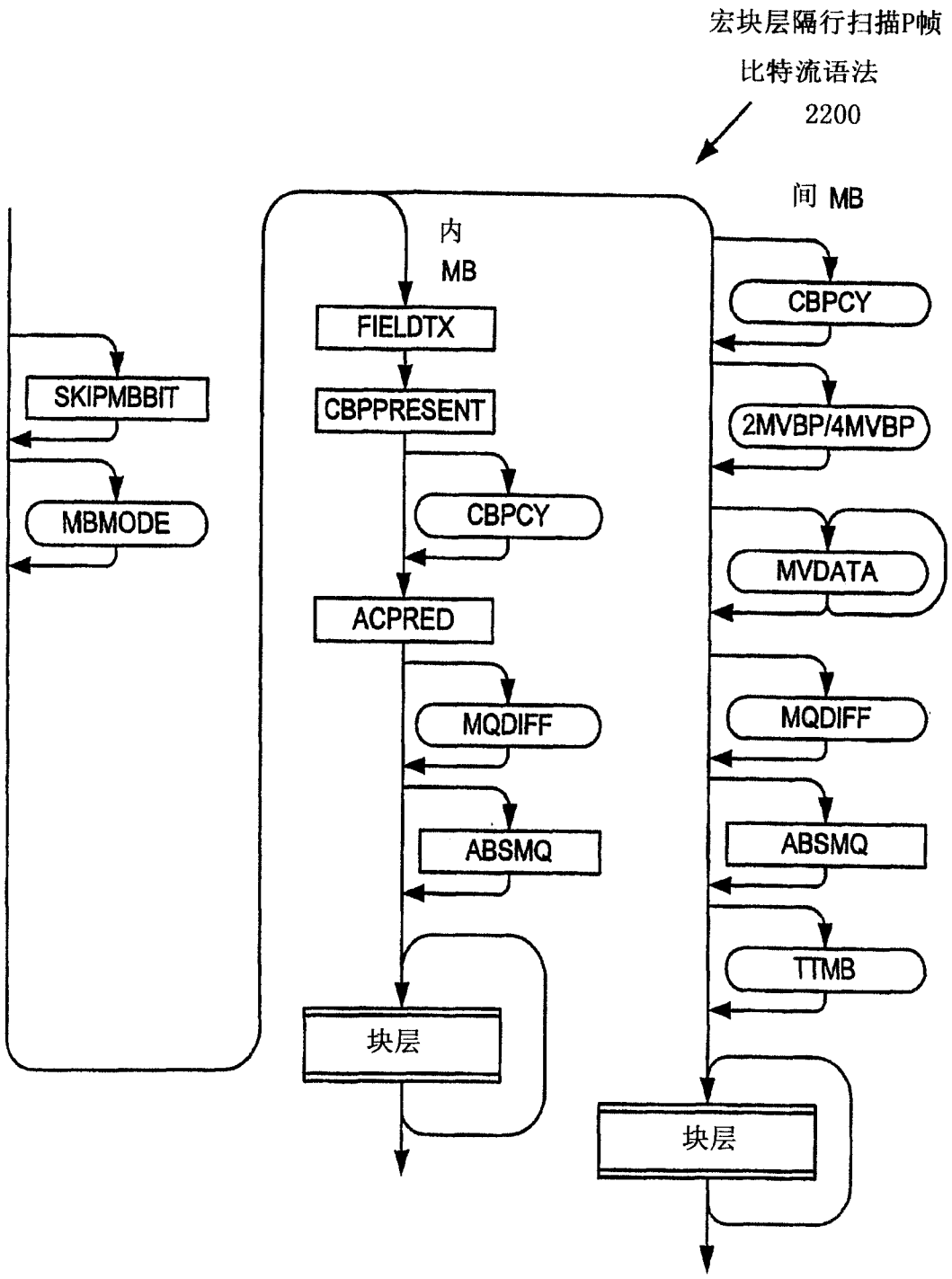

宏块层隔行扫描P帧
比特流语法
2200
间 MB
内
MB
SKIPMBBIT
MBMODE
FIELDTX
CBPPRESENT
CBPCY
ACPRED
MQDIFF
ABSMQ
块层
CBPCY
2MVBP/4MVBP
MVDATA
MQDIFF
ABSMQ
TTMB
块层

图 22

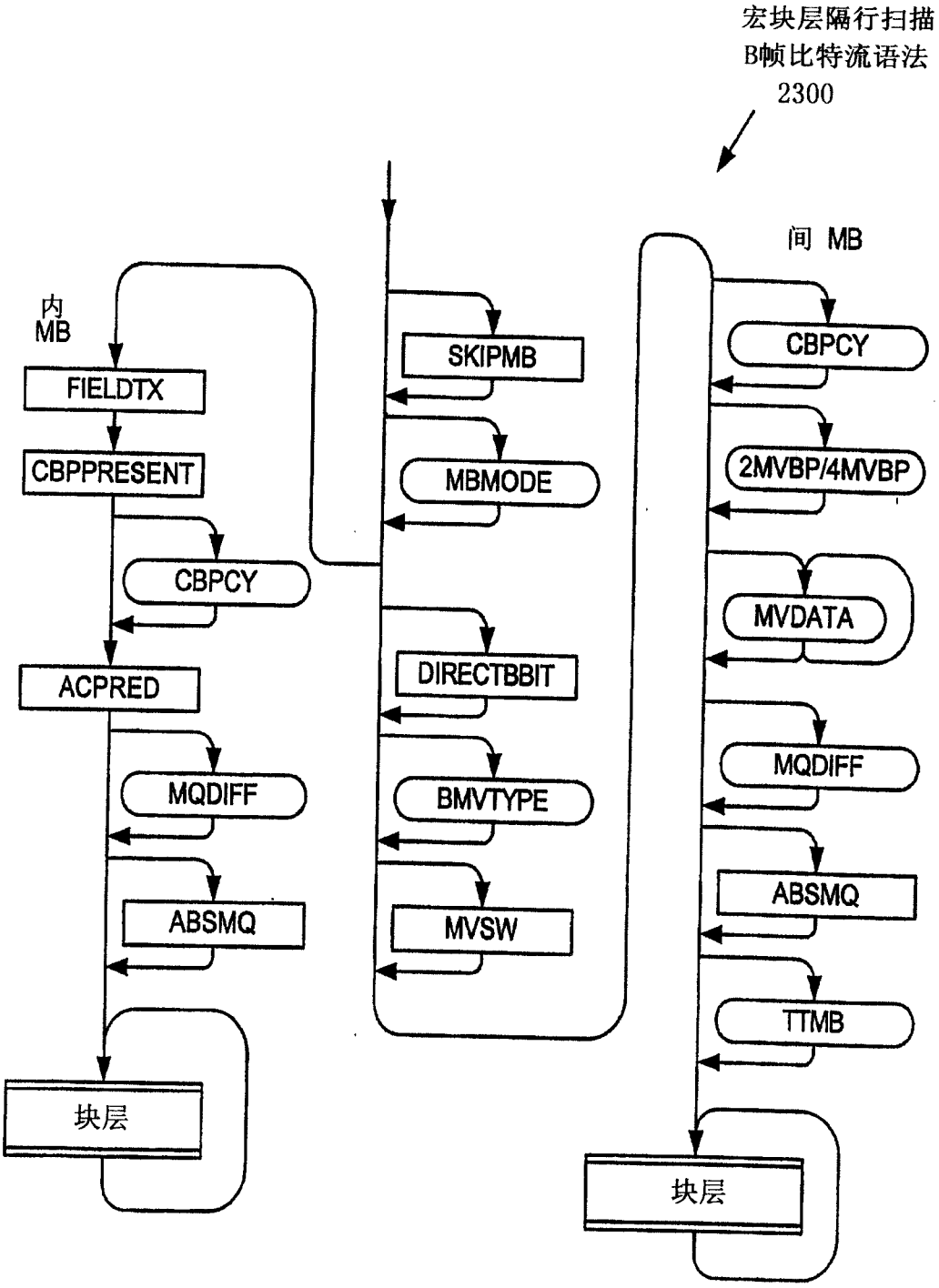

宏块层隔行扫描
B帧比特流语法
2300
内
MB
FIELDTX
CBPPRESENT
CBPCY
ACPRED
MQDIFF
ABSMQ
块层
SKIPMB
MBMODE
DIRECTBBIT
BMVTYPE
MVSW
间 MB
CBPCY
2MVBP/4MVBP
MVDATA
MQDIFF
ABSMQ
TTMB
块层

图 23

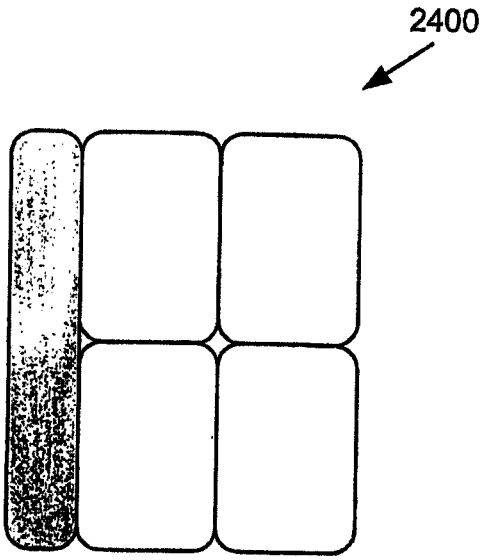


图 24A

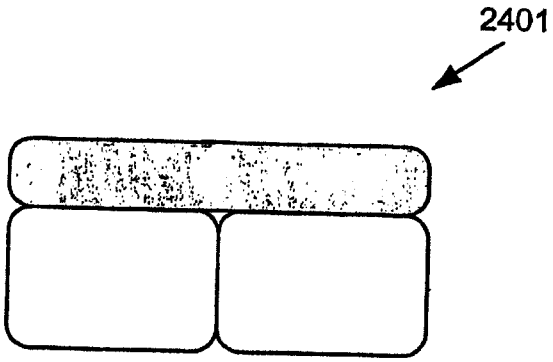


图 24B

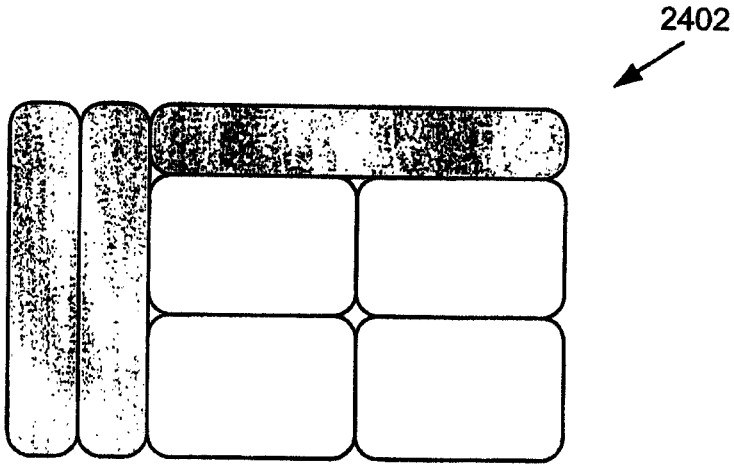


图 24C