(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 1669234 B
(45)授权公告日 2013.08.14

(21)申请号 03816739.5
(22)申请日 2003.07.11
(30)优先权数据
60/396,363 2002.07.15 US
10/291,320 2002.11.08 US
(85)PCT申请进入国家阶段日
2005.01.14
(86)PCT申请的申请数据
PCT/US2003/021714 2003.07.11
(87)PCT申请的公布数据
WO2004/008654 EN 2004.01.22
(73)专利权人 苹果公司
地址 美国加利福尼亚州
(72)发明人 B·G·哈斯克尔 D·W·辛格
A·杜米特拉斯 A·普里
(74)专利代理机构 中国国际贸易促进委员会专利商标事务所 11038
代理人 鲍进
(51)Int.Cl.
H04B 1/66(2006.01)
(56)对比文件
US 5436664 A,1995.07.25, 说明书第3栏第65行-第4栏第5行.
US 6400768 B1,2002.06.04, 说明书第45栏第23行-第46栏第26行、第31栏第25-46行及附图10、48.
CN 1354599 A,2002.06.19, 全文.
US 6297852 B1,2001.10.02, 全文.
审查员 阎岩

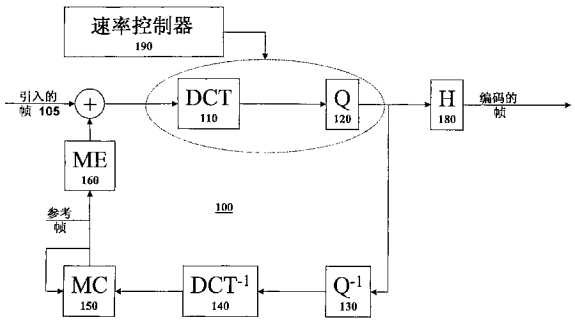
权利要求书1页 说明书6页 附图4页

(54)发明名称

数字视频编码用可变精度图像间定时规定的方法和装置

(57)摘要

本发明公开了一种数字视频编码中用于可变精度图像间定时规定的方法和装置。具体来说，本发明提供一种允许临近视频图像的相对时间可以按各种有效的方式进行编码的系统。在一个实施方式中，确定了当前视频图像(105)和临近视频图像中的显示时间差。显示时间差随后被编码(180)为视频图像的数字表达。在一个优选实施方式中，该临近视频图像为最新传输的存储图像。为有效进行编码，显示时间差可以用可变长度编码系统或算术编码进行编码。在一个可选择实施方式中，该显示时间差被作为2的幂来进行编码以减少传输比特的数目。



CN 1669234 B

1. 一种用于编码多个视频图像的方法，所述方法包括：

确定第一视频图像和临近视频图像之间的第一显示时间差；

确定第二视频图像和所述临近视频图像之间的第二显示时间差；以及

通过将所述第二视频图像的运动矢量和基于所述第一显示时间差和所述第二显示时间差的比率值相乘，为所述第一视频图像计算运动矢量。

2. 如权利要求 1 所述的方法，还包括：

传输所编码的第一视频图像和所述第一显示时间差。

3. 如权利要求 1 所述的方法，还包括：

将所述第一显示时间差编码到与所述第一视频图像相关联的切片头部中。

4. 如权利要求 1 所述的方法，其中所述第一显示时间差在所述第一视频图像中进行一次以上的编码。

5. 如权利要求 1 所述的方法，其中所述第一显示时间差包括相对时间值。

6. 如权利要求 1 所述的方法，其中所述第一显示时间差被编码为 2 的幂。

7. 如权利要求 1 所述的方法，还包括：

通过使用可变长度编码对所述第一显示时间差进行编码。

8. 如权利要求 1 所述的方法，还包括：

通过使用算术编码对所述第一显示时间差进行编码。

数字视频编码用可变精度图像间定时规定的方法和装置

技术领域

[0001] 本发明涉及多媒体压缩系统的领域。具体来说，本发明公开了一种用于规定可变精度图像间（inter-picture）定时的方法和系统。

背景技术

[0002] 基于数字的电子媒体格式最终在尖端在很大程度上取代模拟电子媒体格式。数字压缩磁盘（CD）很早以前就取代了模拟乙烯唱片。模拟盒式磁带越来越稀少。第二和第三代数字音频系统，例如 MD 盘和 MP3（MPEG 音频第三层）正在从压缩磁盘的第一代数字音频格式中分享市场份额。

[0003] 视频媒体转向数字存储和传输格式的速度比音频媒体慢。很大程度上是由于大量的数字信息需要以数字形式精确地表达视频。需要精确地表达视频的大量数字信息需要非常高容量的数字存储系统和高带宽的传输系统。

[0004] 然而，视频正在迅速地向数字存储和传输格式转变。更快的计算机处理器、高密度存储系统，和新型有效的压缩和编码算法最终使数字视频适合于消费者物价点。DVD（数字多功能光盘），即一种数字视频系统，在几年间已成为最为畅销的消费者电子产品之一。DVD 以其高视频质量、非常高的音频质量、方便性和额外特征已经迅速取代了盒式录像机（VCR）成为预录制视频回放系统的选择。老式的模拟 NTSC（国家电视标准委员会）视频传输系统正在被数字 ATSC（先进电视标准委员会）视频传输系统所取代。

[0005] 计算机系统已经使用许多年各种不同的数字编码格式。计算机系统使用的最好的数字视频压缩和编码系统是众所周知的缩写为 MPEG 的活动图像专家组（Motion Pictures Expert Croup）所支持的数字视频格式。MPEG 的最著名并被最广泛应用的三种数字媒体格式被简称为 MPEG-1，MPEG-2 和 MPEG-4。视频 CD（VCD）和早期的消费者级数字视频编辑系统使用早期的 MPEG-1 数字视频编码格式。数字多功能光盘（DVD）和 Dish Network 牌直播卫星（DBS）电视广播系统使用更高质量的 MPEG-2 数字视频压缩和编码系统。MPEG-4 编码系统迅速地被最新的基于计算机的数字视频编码器改进并与数字视频播放器相关联。

[0006] MPEG-2 和 MPEG-4 标准压缩一系列视频帧或视频场并且随后将压缩的帧或场编码为数字比特流。当对视频帧或场用 MPEG-2 和 MPEG-4 系统进行编码时，视频帧或场将被分成宏块的矩形网格。每个宏块被独立地压缩或编码。

[0007] 当压缩视频帧或场时，该 MPEG-4 标准可以将帧或场压缩成三种压缩的帧或场的一种：帧内编码帧（I-帧），单向预测帧（P 帧），或双向预测帧（B 帧）。帧内编码帧在不参照其它视频帧的情况下完全独立地将独立视频帧编码。P 帧参照单一先前显示的视频帧限定一个视频帧。B 帧参照当前帧之前显示的视频帧和在当前帧后显示的视频帧限定一个视频帧。由于对冗余视频信息的有效利用，P 帧和 B 帧通常提供最好的压缩。

发明内容

[0008] 用于数字视频编码的可变精度图像间定时规定的方法和装置被公开。具体来说，

本发明公开了一种使临近的视频图像的相对时间以相当有效的方式进行编码的系统。在一个实施方式中，确定当前视频图像和临近视频图像之间的显示时间差。该显示时间差随后被编码为视频图像的数字表示。在一个优选实施方式中，临近的视频图像是最新传输的存储图像。

[0009] 为有效地进行编码，显示时间差可以用可变长度的编码系统或算术编码进行编码。在一个可选择实施方式中，显示时间差被编码为 2 的幂来减少传输的比特的数目。

[0010] 从附图和以下的详细说明中可以使本发明的其它目的、特征和优点显而易见。

附图说明

[0011] 本领域的技术人员参照以下的详细说明将很容易地理解本发明的目的，特性和优点。

[0012] 图 1 表示一种可能的数字视频编码系统的高级框图。

[0013] 图 2 表示一系列图像应该播放的顺序排列的视频图像，其中连接不同图像的箭头代表用运动补偿创建的图像间相关性。

[0014] 图 3 表示以图像的优选传输顺序排列的图 2 中的视频图像，其中连接不同图像的箭头表示用运动补偿创建的图像间相关性。

[0015] 图 4 表示一系列视频图像，其中视频图像间互相参照的距离被选为 2 的幂。

具体实施方式

[0016] 一种多媒体压缩和编码系统中用于制定变化精度的图像间定时的方法和系统被公开。在以下的描述中，为了说明的目的，所阐述的特定术语为本发明提供全面理解。然而，对于本领域的技术人员，实施本发明并不需要这些具体细节。例如，本发明是参照 MPEG-4 多媒体压缩和编码系统来进行说明的。然而，相同的技术可以容易地应用在其它类型的压缩和编码系统中。

[0017] 多媒体压缩和编码综述

[0018] 图 1 表示技术领域中所熟知的典型的数字视频编码器 100 的高级框图。该数字视频编码器 100 接收在框图的左侧的视频帧 105 的输入视频流。每个视频帧被离散余弦变换 (DCT) 单元 110 所处理。该帧可以独立地（帧内）处理或者参照运动补偿单元（帧间）接收来自其它帧的信息来处理。然后，一个量化器 (Q) 单元 120 将来自离散余弦变换单元 110 的信息量化。最后，量化的视频帧用熵编码器 (H) 单元 180 编码来生成编码的比特流。该熵编码器 (H) 单元 180 可以用可变长度编码 (VLC) 系统。

[0019] 由于帧间编码的视频帧是参照其它临近视频帧来限定的，所以数字视频编码器 100 需要创建一个将出现在数字视频解码器中的每个帧怎样解码的复本，这样帧间就可以被编码。因此，数字视频编码器 100 的下部实际为一个数字视频解码器系统。具体来说，一个反量化 (Q^{-1}) 单元 130 将视频帧信息的量化倒置，并且反离散余弦变换 (DCT^{-1}) 单元 140 将视频帧信息的离散余弦变换进行倒置。在所有 DCT 系数从 iDCT 中重建时，该运动补偿单元将使用信息以及运动矢量来重建随后将被用作下一帧运动评估的参考帧的编码帧。

[0020] 该解码帧可随后用于对相对于解码视频帧中的信息限定的帧间 (P 帧或 B 帧）进行编码。具体来说，一种运动补偿 (MC) 单元 150 和运动评估 (ME) 单元 160 被用来确定运

动矢量并产生用于编码帧间的不同数值。

[0021] 一个速率控制器 190 接收来自数字视频编码器 100 的不同分量的信息并使用该信息为每个视频帧分配比特预算。该速率控制器 190 应按照以下方式分配比特预算，使其产生最高质量的遵循一系列指定的限制的数字视频比特流。具体来说，该速率控制器 190 在没有溢出缓冲（通过发射比能够存储的更多的信息来超过解码器中可访问的存储量）或下溢缓冲（发送视频帧不够快因此使解码器缺少进行显示的视频帧）的情况下产生最高质量的压缩视频流。

[0022] 多媒体压缩与编码综述

[0023] 在一些视频信号中，连续视频图像（帧或场）之间的时间可以不恒定。（注意：本文将用术语视频图像来一般地指视频帧或视频场）。例如，一些视频图像可能由于传输带宽的限制而丢掉。此外，该视频定时还可能由于摄像机的不规则或特殊效果，如慢动作或快动作等而改变。在一些视频流中，原始视频源可以通过设计简单地具有非一致图像间时间。例如，由于没有通过例如视频摄像系统等统一的视频捕捉系统来创建非任意视频定时，所以如计算机图像动画等合成视频可具有非一致的定时。灵活的数字视频编码系统应能够处理非一致的定时。

[0024] 许多数字视频编码系统将视频图像分成宏块的矩形网格。视频图像中的每个个体宏块被独立地压缩和编码。在某些实施方式中，使用熟知的“像素区块”的宏块中的子块。这样的像素块可以具有其本身的可内插的运动矢量。即使本发明的说明被等同地应用在宏块和像素区块中，本文将涉及宏块。

[0025] 某些视频编码标准，如，ISO MPEG 标准或 ITU H.264 标准，使用不同类型的预测宏块来将视频图像编码。在预期中，宏块可以为以下三种类型中的一种。

[0026] 1.I 宏块——一个内 (I) 宏块在其编码中不使用来自任何其它视频图像的信息（完全自定义的）；

[0027] 2.P 宏块——一种单向预测 (P) 宏块是指来自一在前视频图像的图像信息；或

[0028] 3.B 宏块——一种双向预测 (B) 宏块使用一在前图像和一在后视频图像的信息。

[0029] 如果视频图像中所有的宏块为内宏块，则视频图像为帧内编码帧。如果一个视频图像仅包括单向预测宏块或内宏块，则视频图像被已知为 P 帧。如果视频图像包括任何双向预测宏块，那么视频图像已知为 B 帧。为简单起见，本文将考虑所给图像中的所有宏块为同一类型的情况。

[0030] 将编码的视频图像的一个样本顺序可以表示为

[0031] $I_1B_2B_3B_4P_5B_6B_7B_8B_9P_{10}B_{11}P_{12}B_{13}I_{14}...$

[0032] 其中字母 (I，P 或 B) 代表如果视频图像为 I 帧，P 帧，或 B 帧，且数字代表以视频图像的顺序的视频图像的摄像顺序。该摄像的顺序为摄像机记录视频图像的顺序，因此也是视频图像应该被显示的顺序（显示顺序）。

[0033] 视频图像的在前样本系列在图 2 中图解表示。参考图 2，箭头表示来自存储图像的宏块（在此种情况下 I 帧或 P 帧）被用在其它图像的运动补偿预测中。

[0034] 在图 2 的情况中，在帧内视频图像 I_1 的编译中没有来自其它图像的信息。视频图像 P_5 是在其编码过程中使用来自在前视频图像 I_1 的视频信息的 P 帧，因此从视频图像 I_1 到视频图像 P_5 绘制了一个箭头。视频图像 B_2、视频图像 B_3、视频图图像 B_4 在其编译过程中都

使用来自视频图像 I_1 和视频图像 P_5 的信息，因此从视频图像 I_1 和视频图像 P_5 到视频图像 B_2、视频图像 B_3 和视频图像 B_4 绘制了箭头。如上所述该图像间时间通常互不相同。

[0035] 由于 B 图像使用来自在后图像（随后将被显示的图像）的信息时，传输顺序通常与显示顺序不同。具体来说，构建其它图像所需的视频图像应首先被传输。对于以上顺序，传输顺序应为

[0036] $I_1P_5B_2B_3B_4P_{10}B_6B_7B_8B_9P_{12}B_{11}I_{14}B_{13}\ldots$

[0037] 图 3 图解说明了图 2 中的视频图像的上述传输顺序。再次地，图中的箭头指示来自存储的图像的宏块（此种情况下为 I 或 P）被用在其它视频图像的运动补偿预测中。

[0038] 参照图 3，该系统首先传输并不依赖于任何其它帧的 I 帧 I_1。随后，该系统传输依赖于视频图像 I_1 的 P 帧视频图像 P_5。随后，即使视频图像 B_2 将在视频图像 P_5 前被显示，系统仍将在视频图像 P_5 后传输 B 帧视频图像 B_2。其原因是，当应对 B_2 进行解码时，解码器将已经接收并存储了将视频图像 B_2 进行解码所必须的视频图像 I_1 和 P_5 中的信息。类似地，视频图像 I_1、P_5 已准备好来将随后的视频图像 B_3 和视频图像 B_4 进行解码。该接收器 / 解码器记录用于正确显示的视频图像顺序。在本操作中，I 和 P 图像通常指存储的图像。

[0039] P 帧图像的编码过程典型地使用运动补偿，其中在图像中为每个宏块来计算运动矢量。使用计算出的运动矢量，预测宏块 (P 宏块）可以通过转变上述在前图像中的像素来进行。在 P 帧图像中的实际宏块和预测宏块之间的差别随后被编码以进行传输。

[0040] 每个运动矢量也可以通过预测编码来传输。例如，一个运动矢量预测可以用临近的运动矢量形成。在此种情况下，在实际运动矢量和运动矢量预测之间的差别被编码以进行传输。

[0041] 每个 B 宏块使用两个运动矢量 :第一运动矢量是指上述在前视频图像，第二运动矢量是指在后视频图像。从这两个运动矢量中，计算两个预测宏块。此两个预测宏块随后用某一函数结合在一起来形成最终的预测宏块。如上所述，在 B 帧图像中的实际宏块和最终预测宏块之间的差别随后被编码以进行传输。

[0042] 由于具有 P 宏块，B 宏块的每个运动矢量 (MV) 可以通过预测编码来传输。具体来说，一个预测运动矢量用临近的运动矢量来形成。随后，在实际运动矢量和预测的之间的差别被编码以进行传输。

[0043] 然而，由于 B 宏块，存在着从最近的存储图像宏块中内插运动矢量的机会。在数字视频编码和数字视频解码中都进行这样的内插。

[0044] 该运动矢量在从摄影机缓慢扫过静态背景的视频顺序的视频图像中工作得非常好。实际上，这样的运动矢量内插足够好以致可以单独使用。具体来说，这意味着对于这些使用内插的编码的 B 宏块运动矢量无需计算或传输任何差别信息。

[0045] 为进一步说明，在上述情况中，我们将图像 i 和 j 之间的图像间显示时间表示为 $D_{i,j}$，即，如果图像的显示时间分别为 T_i 和 T_j，则

[0046] $D_{i,j} = T_i - T_j$ 随后为

[0047] $D_{i,k} = D_{i,j} + D_{j,k}$

[0048] $D_{i,k} = -D_{k,i}$

[0049] 注意 $D_{i,j}$ 在某些情况下可以取负值。

[0050] 因此，如果 $MV_{5,1}$ 为参考 I_1 的 P_5 宏块的运动矢量，则对于 B_2、B_3 和 B_4 中对应的宏块，

参考 I_1 和 P_5 的运动矢量应分别地插入

[0051] $MV_{2,1} = MV_{5,1} * D_{2,1} / D_{5,1}$

[0052] $MV_{5,2} = MV_{5,1} * D_{5,2} / D_{5,1}$

[0053] $MV_{3,1} = MV_{5,1} * D_{3,1} / D_{5,1}$

[0054] $MV_{5,3} = MV_{5,1} * D_{5,3} / D_{5,1}$

[0055] $MV_{4,1} = MV_{5,1} * D_{4,1} / D_{5,1}$

[0056] $MV_{5,4} = MV_{5,1} * D_{5,4} / D_{5,1}$

[0057] 注意，由于显示时间的比率被用于运动矢量预测，所以不需要绝对显示时间。因此，相对显示时间可以用做 $D_{i,j}$ 显示时间值。

[0058] 本情况作为 H.264 标准中的实例被概括。在概括中，一个 P 或 B 图像可以使用任何先前传输的图像用于运动矢量预测。因此，在上述情况下，图像 B_3 可以在其预测中使用图像 I_1 和图像 B_2。此外，运动图像不仅可被内插还可被外推。因此，在此种情况下我们将得到：

[0059] $MV_{3,1} = MV_{2,1} * D_{3,1} / D_{2,1}$

[0060] 此种运动矢量外推（或内插）也可以应用在预测过程中来用于运动矢量的预测编码过程。

[0061] 在任何情况下，在非一致图像间时间的问题是将 $D_{i,j}$ 的相对显示时间传输到接收器上，这也是本发明的主题。在本发明的一个实施方式中，在第一图像之后对于每个图像，我们将传输当前图像和最新传输的存储图像之间的显示时间差。为了错误恢复，在图像中应将传输重复几次，例如，在 MPEG 或 H.264 标准的所谓切片头部（slice header）中。如果所有的切片头部丢失，那么推测起来，用于解码信息的依赖于丢失图像的其它图像也不能被解码。

[0062] 因此，在上述情况中，我们应传输以下：

[0063] $D_{5,1}D_{2,5}D_{3,5}D_{4,5}D_{10,5}D_{6,10}D_{7,10}D_{8,10}D_{9,10}D_{12.10}D_{11,12}D_{14,12}D_{13,14}\ldots$

[0064] 为了运动矢量的评估，用于 $D_{i,j}$ 的精度需要可以在各图像间改变。例如，如果在两个 P 帧图像 P_5 和 P_7 之间的中途只有一个单一 B 帧图像 B_6，则发送：

[0065] $D_{7,5} = 2$ 和 $D_{6,7} = -1$ 就足够了。

[0066] 其中 $D_{i,j}$ 显示时间值为相对时间值。如果，作为替换，视频图像 B_6 仅在视频图像 P_5 和视频图像 P_7 之间距离的 1/4 处，则将发送的适当的 $D_{i,j}$ 显示时间值为：

[0067] $D_{7,5} = 4$ 和 $D_{6,7} = -1$

[0068] 注意在以上两个例子中，在视频图像 B_6 和视频图像 P_7 之间的显示时间是作为显示时间“单位”来使用，并且在视频图像 P_5 和视频图像 P_7 之间的显示时间差为 4 个显示时间“单位”。

[0069] 总之，如果除数为 2 的幂则运动矢量评估的复杂性较小。如果如图 4 中所图解的，两个存储图像之间的 $D_{i,j}$（图像间时间）被选为 2 的幂，则在我们的实施方式中很容易完成。可选择地，该评估过程可以被定义为将所有除数取截取或近似为 2 的幂。

[0070] 当图像间的时间为 2 的幂时，如果只有（2 的）整数次幂被传输，而不是图像间时间的全数值被传输，数据比特的数目能够被减少。图 4 中图解说明了图像间的距离被选为 2 的幂的情况。在此种情况下，在视频图像 P_1 和视频图像 P_3 之间的 $D_{3,1}$ 显示时间值 2 传输

为 1(因为 $2^1 = 2$),并且在视频图像 P_7 和视频图像 P_3 之间的 $D_{7,3}$ 显示时间值 4 可以传输为 2(由于 $2^2 = 4$)。

[0071] 在某些情况下,可以不使用运动矢量内插。然而,仍然需要将视频图像的显示顺序传输到接收器 / 播放器系统中,这样接收器 / 播放器将按照适当的顺序显示视频图像。在此种情况下,不考虑实际显示的时间,用于 $D_{i,j}$ 的简单有符号整数就足够。在某些应用中可以只需要符号。

[0072] 图像间时间 $D_{i,j}$ 可以简单地作为简单有符号整数值传输。然而,可以使用许多方法来将 $D_{i,j}$ 数值编码来完成额外压缩。例如,一个可变长度编码数量之前的符号位相对容易实施和提供编码效率。

[0073] 可以使用的一个此种可变长度编码系统是所熟知的 UVLC(通用可变长度编码)。该 UVLC 可变长度编码系统由编码词给出 :

[0074] 1 = 1

[0075] 2 = 010

[0076] 3 = 011

[0077] 4 = 00100

[0078] 5 = 00101

[0079] 6 = 00110

[0080] 7 = 00111

[0081] 8 = 0001000...

[0082] 用于将图像间时间编码的另一种方法可以用算术编码方法。典型地,算术编码使用条件可能性来实现数据比特的非常高的压缩。

[0083] 因此,本发明介绍了一种将图像间显示时间进行编码和传输的简单但有效的方法。图像间的显示时间的编码过程可以通过使用可变长度编码或算术编码而变得非常有效。此外,所需的精度可以选为符合视频解码器的需要,而不再更多。

[0084] 上述内容说明了一种用于在多媒体压缩和编码系统中指定各种精度的图像间时间的系统。应该理解的是本领域的普通技术人员可以在不背离本发明范围的情况下对本发明元件的材料和排列进行改变和改进。

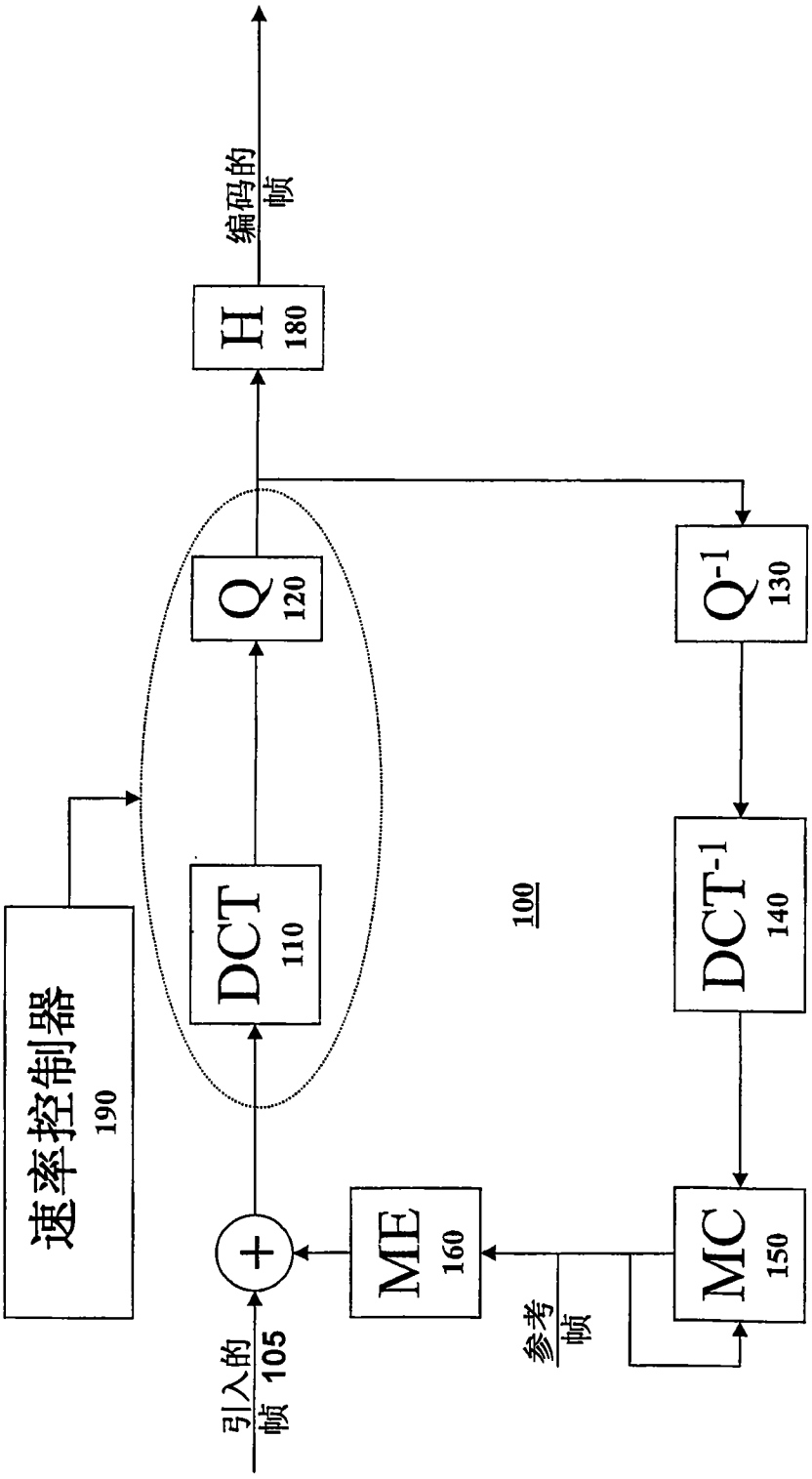

编码的帧
H
180
Q
120
Q-1
130
DCT
110
DCT-1
140
速率控制器
190
100
ME
160
MC
150
引入的帧 105
参考帧

图1

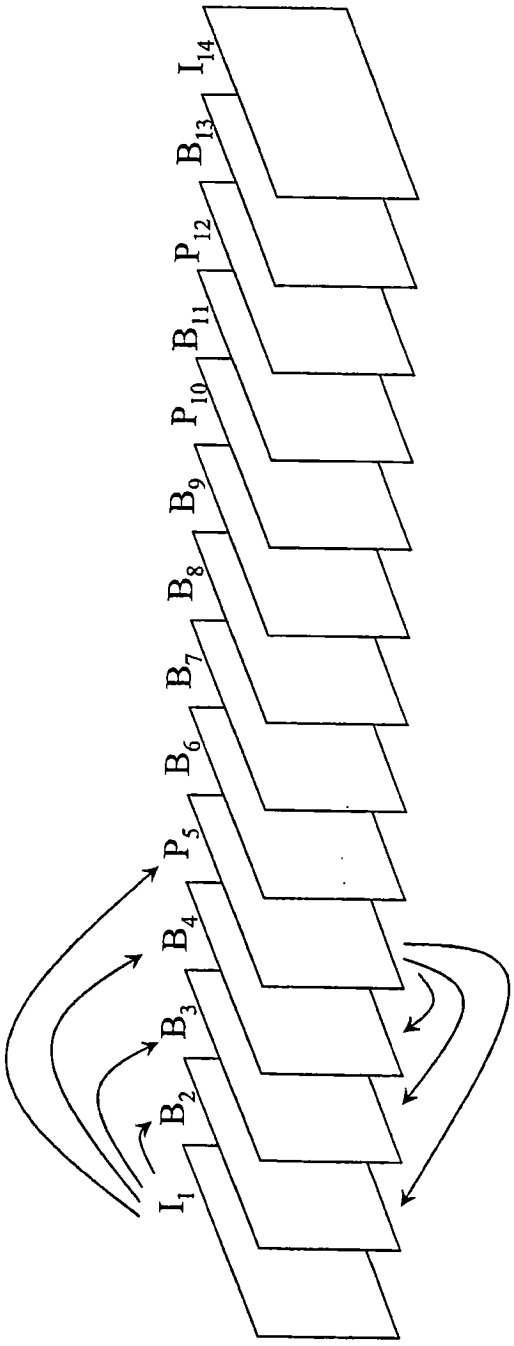


图2

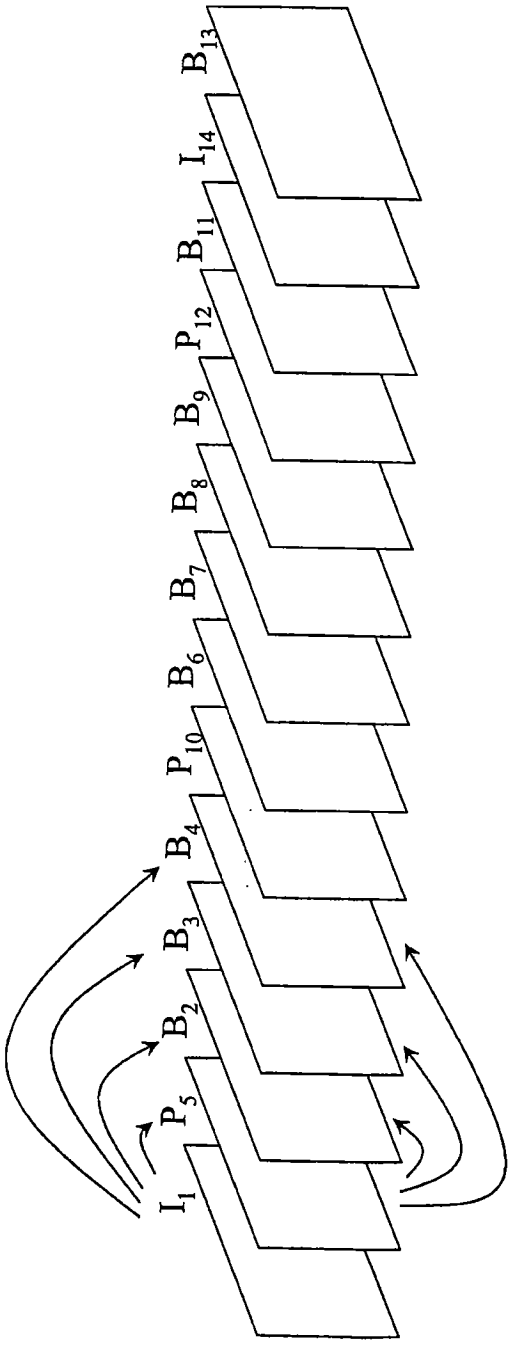


图3

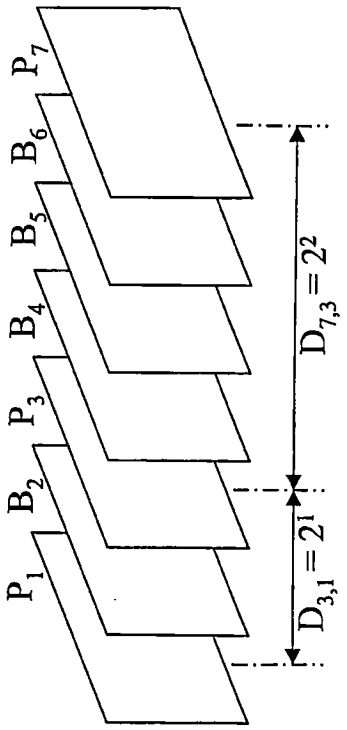


图4