



[12] 发明专利说明书

专利号 ZL 02805857.7

[45] 授权公告日 2006 年 11 月 8 日

[11] 授权公告号 CN 1284355C

[22] 申请日 2002.3.1 [21] 申请号 02805857.7

[30] 优先权

[32] 2001. 3. 2 [33] US [31] 09/798,346

[86] 国际申请 PCT/US2002/006078 2002.3.1

[87] 国际公布 WO2002/071735 英 2002.9.12

[85] 进入国家阶段日期 2003.9.2

[71] 专利权人 杜比实验室许可股份有限公司

地址 美国加利福尼亚州

[72] 发明人 G·A·迪莫斯 D·鲁霍夫

审查员 杨 娟

[74] 专利代理机构 上海专利商标事务所有限公司

代理人 李家麟

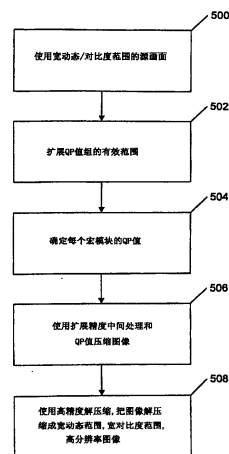
权利要求书 1 页 说明书 16 页 附图 5 页

[54] 发明名称

视频图象的高精度编码和解码

[57] 摘要

用于提高视频压缩质量的方法、系统和计算机程序。通过在中间编码和解码处理步骤期间保存较多数目的位，可以使来自 MPEG 型视频编码的图象质量提高。通过使解码器的反离散余弦变换 (IDCT) 功能数值算法与编码器的解码部分所使用的 IDCT 功能数值算法确切匹配，可以消除反离散余弦变换失配的问题。还包括通过扩展“量化参数”或“QP”的范围而把高精度压缩施加于宽动态范围图象。通过直接增加 QP 的范围或间接地通过非线性变换可以完成 QP 的扩展。还包括把经扩展的中间处理精度和经扩展的 QP 范围施加于图象对比度降低的区域来扩展精度，用该扩展精度对低一对比度部分进行压缩编码。



1. 一种用于在把包括帧序列的数字化视频图象的序列压缩成包括 I 帧的经压缩的位流的期间增加平均压缩比的方法，所述方法包括：

转换所述帧序列的颜色空间以产生第一输出；

将离散余弦变换施加于所述第一输出以产生第二输出；

量化所述第二输出以产生第三输出；以及

将与后继的解码过程期间施加的反离散余弦变换数值算法匹配的反离散余弦变换数值算法对帧序列施加于所述第三输出，从而产生经压缩位流，所述经压缩位流具有增加的 I 帧间间隔。

2. 如权利要求 1 所述的方法，其特征在于，进一步包括：

去量化所述经压缩位流以产生第四输出；

将与编码期间施加的反离散余弦变换数值算法匹配的反离散余弦变换数值算法施加于所述第四输出，以产生包括 I 帧的有符号输出。

3. 一种用于在把包括帧序列的数字化视频图象的序列压缩成包括 I 帧的经压缩的位流的期间增加平均压缩比的系统，所述系统包括：

用于转换所述帧序列的颜色空间以产生第一输出的装置；

用于将离散余弦变换施加于所述第一输出以产生第二输出的装置；

用于量化所述第二输出以产生第三输出的装置；以及

用于将与后继的解码过程期间施加的反离散余弦变换数值算法匹配的反离散余弦变换数值算法对帧序列施加于所述第三输出，从而产生经压缩位流的装置，所述经压缩位流具有增加的 I 帧间间隔。

4. 如权利要求 3 所述的系统，其特征在于，进一步包括：

用于去量化所述经压缩位流以产生第四输出的装置；

用于将与编码期间施加的反离散余弦变换数值算法匹配的反离散余弦变换数值算法施加于所述第四输出，以产生包括 I 帧的有符号输出的装置。

视频图象的高精度编码和解码

技术领域

本发明涉及视频压缩，尤其，涉及基于 MPEG 类的编码和解码系统的新颖改进的提高了质量的视频压缩。

背景技术

MPEG 背景

MPEG-2 和 MPEG-4 是定义视频语法的国际视频压缩标准，所述语法提供一种有效的方法，以多个压缩编码数据的形式来表示视频图象序列。编码位的语言是“语法”。例如，几个标记可以表示整个取样块(例如，MPEG-2 的 64 个取样)。两种 MPEG 标准还描述解码(重构)过程，其中，把编码位从压缩表示映射到图象序列原始格式的近似。例如，在编码位流中的一个标志给出信号表示，跟随在后的一些位是否在用离散余弦变换(DCT)算法进行解码之前是否前面有一个预测算法。通过由这些 MPEG 标准定义的语义学来调整包括解码过程的所述算法。可以把这个语法应用于开发公共视频特征，诸如空间冗余、时间冗余、一致运动、空间掩模等。事实上，这些 MPEG 标准定义编程语言以及数据格式。MPEG 解码器必须能够对输入数据流进行分析和解码，但是只要数据流符合相应的 MPEG 语法，就可以使用广泛范围的多种可能的数据结构以及压缩技术。还有可能在另外的语法中携带所需要的语义学。

这些 MPEG 标准使用多种压缩方法，包括帧内方法以及帧间方法。在大多数视频情景中，背景保持相当稳定而在前景中发生动作。背景可以移动，但是大量情景是冗余的。这些 MPEG 标准通过创建被称为“内部”帧或“I 帧”的参考帧来开始压缩。压缩 I 帧而无需参考其它帧，因此包含视频信息的整个帧。I 帧提供用于随机访问的数据位流的进入点，但是只可以进行适度的压缩。一般，在位流中每 12 到 15 帧放置表示 I 帧的数据。此后，由于只有小部分落在参考 I 帧之间的帧与括号 I 帧不同，所以只捕获、压缩和存储图象差异。对于这种差异使用两类帧—预测帧或 P 帧以及双向内插帧或 B 帧。

一般参考过去帧(I 帧或以前的 P 帧)对 P 帧进行编码，并且一般使用 P 帧

作为后继 P 帧的参考。P 帧接收十分高的压缩量。B 帧提供最高的压缩量，但是需要过去和将来两种参考帧，以便进行编码。从来不把双向帧用作为参考帧。

宏模块是图象像素的区域。对于 MPEG-2，宏模块是四个 8×8 DCT(离散余弦变换)模块组成的 16×16 像素，以及用于 P 帧的一个运动矢量和用于 B 帧的一个或两个运动矢量。可以使用内部帧编码或帧间(预测)编码对 P 帧中的宏模块进行单独编码。可以使用内部帧编码、正向预测编码、反向预测编码、或正向和反向(即，双向内插)两种预测编码对 B 帧中的宏模块进行单独编码。

在编码之后，MPEG 数据位流包括 I、P 和 B 帧的一个序列。一个序列可以包括几乎 I、P 和 B 帧的任何模式(在它们的放置上有几个次要的语义的限制)。然而，在工业实践中较普通的是具有固定模式(例如，IBBPBBPBBPBBPBB)。

已经知道了有一些时候，当通过利用分层运动搜索来确定运动矢量时，减少了计算。例如，MPEG 算法试图寻找“宏模块”区域之间的匹配。MPEG 型以及其它经运动补偿的 DCT(离散余弦变换)编码器试图使当前帧中的每个宏模块区域与以前帧(P 帧)中的一个位置或以前和后继帧(B 帧)中的一个位置匹配。然而，不是始终需要寻找优良匹配，由于在这种情况下 MPEG 可以对新的宏模块进行编码作为刷新的独立(“内部”)宏模块而无需使用以前或后继帧。在这种运动补偿 DCT 系统中对于 MPEG-2 的每个宏模块区域需要一个宏模块运动矢量。在 MPEG-4 中，相应于每个 8×8 区域有一个矢量的 4 个运动矢量的一个组(即，每个宏模块 4 个矢量)也是一种任选的编码模式。

MPEG 精度

参考 MPEG-2 和 MPEG-4 视频编解码器实施利用下列编码方法：

- a) 当从 RGB 转换到 YUV 颜色空间时，只保留要编码的位数(例如，在编码中把 MPEG-2 限制到 8 位，因此 YUV 值也限制到 8 位)。
- b) 当编码和解码时，只保存已经编码的位数，小心地施加舍入以减少伪象。
- c) 当转换回 RGB 时，由于保存位数的限制(诸如对于 MPEG-2，最多为 8 位)而限制了精度。

图 1 是现有技术 MPEG-2 参考视频编码方法的方框图。把按每种颜色 8 位/像素进行编码的 RGB 输入帧 102 施加于 RGB 到 YUV 转换器 104，它有意地在它的输出端限制为每种颜色 8 位的精度。把结果施加于 DCT 功能 106，然后施加

于量化器功能 108，然后到反 DCT 功能 110，按相同的精度存储最终输出作为输出数据。

用相同方法实施 MPEG-4 的参考视频编码器，虽然中间精度可以扩展到 12 位(虽然 VLC-可变长度编码-表不支持全范围的使用)。

利用随机地抖动受限制的精度值(在 MPEG-2 中，最大为每种颜色分量 8 位)的技术来减少步长改变时的明显的可见性。然而，由于这种抖动，还由于使用了有限的中间处理精度，产生了编码中的噪声和伪象。

除了有限的中间处理精度之外，MPEG-2 和 MPEG-4 允许在编码期间使用的反 DCT(IDCT)算法(经常以高精度浮点表示来实施)与解码期间使用的 IDCT 算法略有不同。这是众知的“IDCT”失配。IDCT 失配导致信号离开指定解码值的不可预测的逐渐漂移。传统上通过使用 ICDT 最高频率中的低阶位(MPEG-2 和 MPEG-4 中使用的 8×8 DCT 模块尺寸的 7 次谐波)的随机抖动而使这种漂移减少。这种抖动在信号上增加附加的噪声和伪象。

图 2 是现有技术 MPEG-2 参考视频解码方法的方框图。把经编码的输入位流 202 施加于具有与输入位流的精度(一般对于 MPEG-2 为 8 位)匹配的有限精度的去量化器功能 204。把结果施加于 IDCT 功能 206(它可能和编码器的 IDCT 功能 110 不匹配)，它输出有符号的 8-比特值 208。这个输出包括 I 帧 210，或与来自先前帧 212 或后继帧 214(两者都具有相同的精度)的数据组合而产生的新帧 216。因此，MPEG-2 解码过程限制了中间处理精度到最大为 8 位。相似地，还把 MPEG-4 视频解码的中间处理精度限制到在编码中使用的位数(最大为 12 位，但是经常设置成 8 位)。

在 MPEG-2 和 MPEG-4 中的有限精度还限制了动态范围(即，可以显示的图象亮度等级数)以及对比度范围(即，分配给相似对比度的图象区域的不同级别数)。相应地，与原始输入图象相比，在 MPEG-2 和 MPEG-4 中使用的编码和解码方法降低了输出的、经压缩的图象的潜在质量。本发明致力于解决这些限制。

发明内容

根据对于 MPEG 类编码和解码系统的新颖改进，本发明针对提高质量的视频压缩。在一个方面，本发明提供一种方法，用于通过在中间编码和解码处理步骤期间保存较多位数而显著地提高来自 MPEG 型的视频编码的图象质量。令

人惊讶地，这种质量的提高并不导致对图象序列进行编码所需要的总位数成比例地增加。此外，通过使解码器的 IDCT 功能数值算法与编码器的解码部分所使用的 IDCT 功能数值算法正确地匹配，可以排除 IDCT 失配的问题。排除 IDCT 失配允许通过减少所需要的 I 帧数来增加压缩比。

在另一个方面中，本发明包括通过扩展“量化参数”或“QP”的范围而把高精度压缩施加于宽动态范围图象。动态范围扩展把低 QP 用于暗区以及把高 QP 用于亮区。可以通过直接增加 QP 的范围或间接地通过非线性变换(诸如一个函数或查找表)来完成 QP 的扩展，所述非线性变换把 QP 值的小范围映射到用于除(压缩)或乘(解压缩)的输出 QP 值的宽广范围。

在另一个方面中，本发明包括应用经扩展的中间处理精度和经扩展的 QP 范围，以减小图象的对比度范围，以扩展对低对比度部分进行压缩编码所具有的精度。可以与低对比度(常常是远的)区域一起利用低 QP，而对于高对比度(常常是前景)区域利用高 QP。如此，例如，如果摄像机传感器具有足够的比特分辨率，则当解压缩时，完全可以辨别在模糊日子的空中或地上的远目标。

尤其，在本发明的一个方面中包括一种方法，用于对包括第一颜色空间中按第一精度表示的帧序列的数字化视频图象序列进行压缩，所述方法包括把帧序列变换成第二颜色空间中按比第一精度高的第二精度的显示，并按第二精度执行后继的编码步骤，以产生经压缩的输出。

本发明的另一个方面是一种方法，用于对包括第一颜色空间中按第一精度表示的帧序列的数字化视频图象序列进行压缩，所述方法包括把帧序列变换成第二颜色空间中按比第一精度高的第二精度的显示；按第二精度在帧序列上执行运动补偿离散余弦变换，以产生第一编码输出；按第二精度对第一编码输出进行量化，以产生经量化输出；在经量化输出上按第二精度执行反离散余弦变换，以产生按第二精度的经压缩输出。

在另一个方面中，本发明包括一种方法，用于对包括第一颜色空间中按第一精度表示的帧序列的数字化视频图象序列进行压缩和解压缩，所述方法包括对帧序列进行编码，以产生经压缩位流，并按高于第一精度的第二精度在经压缩位流上执行解码步骤。

本发明的另一个方面是一种方法，用于对包括第一颜色空间中按第一精度表示的帧序列的数字化视频图象序列进行压缩和解压缩，所述方法包括把帧序列变换成第二颜色空间中按比第一精度高的第二精度的表示；按第二精度执行

后继的编码步骤，以产生经压缩位流；以及通过对经压缩位流按第二精度进行去量化而对经压缩位流进行解码，以产生经去量化的输出，在经去量化输出上按第二精度施加反离散余弦变换，以产生经解压缩输出，以及按第二精度从经解压缩输出产生图象帧。

本发明的再另一个方面是一种方法，用于对包括第一颜色空间中按第一精度表示的帧序列的数字化视频图象序列进行压缩和解压缩，所述方法包括把帧序列变换成第二颜色空间中按比第一精度高的第二精度的表示；按第二精度执行后继的编码步骤以产生经压缩位流包括施加反离散余弦变换数值算法；按第二精度在经压缩位流上执行后继的解码步骤包括施加匹配的反离散余弦变换数值算法。

本发明的另一个方面是一种方法，用于在把包括帧序列的数字化视频图象序列压缩成包括 I 帧的经压缩位流期间增加平均压缩比，所述方法包括通过施加与后继的解码过程期间施加的反离散余弦变换数值算法匹配的反离散余弦变换数值算法对帧序列进行编码，从而产生在 I 帧之间具有增加空间的经压缩位流。

本发明的另一个方面是一种方法，用于对具有宽动态范围或宽对比度范围中至少一种范围的数字化视频图象序列进行压缩，视频图象包括按第一精度表示的帧，每个帧包括多个定义区域的宏模块，所述方法包括确定量化参数，QP，所述量化参数具有表示可能值的直接对应或表示可能值的扩展范围的代码；确定每个帧的每个宏模块的 QP 值；按比第一精度高的第二精度对每个帧进行压缩以产生经压缩的帧，这种压缩包括施加这种帧的确定的 QP 值以减少对这种帧进行编码所需要的位数；使 QP 代码与压缩期间使用的确定 QP 值相关联；以及输出每个经压缩帧以及相关联的 QP 代码。

本发明包括相应的计算机程序实施以及设备实施。

在下面的附图和说明书中陈述本发明的一个或多个实施例的详细说明。从说明书和附图以及从权利要求书。本发明的其它特征、目的和优点将变得明白。

附图说明

图 1 是现有技术 MPEG-2 参考视频编码方法的方框图；

图 2 是现有技术 MPEG-2 参考视频解码方法的方框图；

图 3 是根据本发明的 MPEG 类编码方法的方框图；

图 4 是根据本发明的 MPEG 类解码方法的方框图；

图 5 是概括在图象压缩期间扩展动态范围和/或对比度的较佳方法的流程图。

在各附图中的相同参考符号表示相同元件。

具体实施方式

较高精度中间图象处理

设计 MPEG-2 和 MPEG-4 中的有限精度假设和技术，以使存储 I、B 和 P 帧需要的存储器的量最少。然而，帧存储器是当前十分买得起的。本发明部分地基于一个发现，即，通过保存中间编码和解码处理步骤期间的较多数目的位数，可以从 MPEG 风格的视频编码得到显著提高的图象质量。令人惊讶地，这种质量的提高并不导致对图象序列进行解码所需要的总位数成比例地增加。事实上，使用本发明通常使位数减少。

图 3 是根据本发明的 MPEG 类编码方法的方框图。把 RGB 输入帧 302 施加于 RGB 到 YUV 转换器 304。因为调制解调器帧产生装置(例如，视频摄像机和高分辨率胶片扫描器)可以输出较大的颜色范围(例如，每种颜色 10 位/象素)的图象，这种装置作为输入源是较佳的。然而，输入源可以是传统的 MPEG-2 或 MPEG-4 的 8-位装置。把结果施加于 DCT 功能 306，它保存比原始输入信号中存在的精度位数更多的精度位数(例如，16 位)。把 DCT 功能 306 的输出施加于量化器功能 308，然后到 IDCT 功能 310，它再保存比原始输入信号中存在的精度位数(诸如 16 位，对于 P 和 B 帧是有符号的，对于 I 帧是无符号的，16 位是数字硬件和软件系统的方便表示)更多的精度位数(例如，16 位)。一般，最终输出信号 312 是 YUV 信号按相同精度存储作为中间处理精度。在输出中的这种增加精度的重要特征在于它允许改进后继的 P 和 B 帧的预测。

在 MPEG-2 和 MPEG-4 视频编码中体现的 IDCT 失配的概念是基于一种假设，即，解码所使用的计算可能与编码的解码部分所使用的计算不同。如上所述，即使在出现高次谐波低位 DCT 失配抖动时，这种失配也会导致漂移(还增加噪声)。MPEG-2 中的普通实践是使 I 帧的位置相互接近(约每半秒)，以再校正这种漂移以及限制差错的扩展。然而，I 帧是相当无效率的，通常，耗费的位数约为 P 帧位数的 3 倍那么多，以及 B 帧的 5 倍那么多。在“调准”到运动画面序列期间，I 帧还形成再起点和参考点。然而，可以有用地增加它们在存

储媒体上的出现频率到数秒(例如, 平均在约 1 到 5 秒的范围内)以提高效率, 如果没有 IDCT 失配的话。

通过使解码器的 IDCT 功能的数值算法与编码器的解码部分使用 IDCT 功能的数值算法确切地匹配可以排除 IDCT 失配。由于从一个 P 帧到下一个 P 帧通过对它的差信号(它包括 IDCT 精度限制的差)编码而出现的固有反馈机构, 在每个 P 帧处自动校正了这些匹配 IDCT 功能中的任何精度的限制。数值算法的“确切匹配”意味着把输入变换成输出的算法的这些部分应该应用相同的分辨率来进行乘法和加法功能、相同的数值表示、相同的精度等。然而, 就计算机程序或集成电路实施而论, 数值算法不需要相同。因此, 例如, 可以使用不同的计算机语言和二进制产生模式(例如, 解译对编译)。

因此, 通过 IDCT 功能中的足够精度可以得到高质量编码。然而, IDCT 功能不需要极高的精度。例如, 在 MPEG-4 视频参考软件中, 使用双宽度浮点(64 位) IDCT 实施。完全不需要这样, 由于 16 位 IDCT 实施足以提供对于高达 12 一位动态范围编码所需要的改进。可以使用诸如 32 位浮点实施之类的大于 16 位的编码器和解码器 IDCT 实施(精确匹配), 以扩展动态范围达 16 位(这导致在小于百分之零点的对数步长中大于十亿到 1 的动态范围, 超过人类视觉极限)。因此通过精确匹配编码器和解码器 IDCT 实施, 本发明大大地减少了 IDCT 实施所需要的计算量, 同时排除了 IDCT 失配的问题。此外, 反直觉地, 使用精确匹配编码器和解码器 IDCT 实施, 即使具有增加的中间精度, 实际上也增加总效率(即, 更高的平均压缩比), 由于可以使耗费位的 I 帧在时间上隔开更远(例如, 平均在约 1 到 5 秒的范围内)。的确, 通过实际上无限制的时间可以使 I 帧隔开, 仅受到的限制是能够跳转到程序中间或校正从有损分布信道产生的误差所要求的。

图 4 是根据本发明的 MPEG 类解码方法的方框图。把高精度编码输入位流 402 施加于具有与输入位流的精度匹配的“中间”处理精度的去量化功能 404。把结果较佳地施加于与对应的编码器的 IDCT 功能 310 精确匹配的 IDCT 功能 406。IDCT 功能 406 输出与所有以前内部处理步骤的中间精度(例如, 16 位)相同的有符号的值 408。这个输出包括 I 帧 410, 或与来自以前帧 412 或后继帧 414(两者都具有相同的精度)的数据组合而产生的新帧 416。

此外, 应该排除所有形式的抖动, 因此减小了噪声和伪象。尤其, 应该排除来自舍入的抖动(除了全精度之外, 诸如 16 位, 即, 舍入第 17 位), 以及来

自 IDCT 失配的高次谐波的低位的抖动。还有，在较佳实施例中，在解码期间的任何最终颜色空间转换步骤期间(例如，YUV 到 RGB 或其它转换，诸如 YUV 4:2:0 到 YUV 4:2:2，用于观看、利用或存储经转换图象)，使用附加的中间象素精度，仅在最终步骤处进行舍入。

应该注意，可以使用在图 3 和图 4 中示出的高精度技术对标准精度输入(例如，MPEG-2 使用的 8-位精度输入)进行编码和接着进行解码。尽管结果的质量没有象对较高精度输入进行编码和解码的质量那么高，但是对于现有的 MPEG 编码和解码仍是一种改进。此外，通过在中间处理和存储期间增加精度可以改进编码和解码两种处理。不需要使这种精度相等以得到改进，但是当编码和解码的解码部分在精度和数值算法方面精确匹配时，使改进最优化。

下面是对改进压缩图象质量的本发明的较佳实施例的一个概括：

- 1) 在中间处理期间保存比输入精度多的精度位(例如在编码期间从 RGB 到 YUV 转换步骤保存更多精度位，以及从 IDCT 步骤的更多精度位)。
- 2) 存储所增加的中间精度结果。
- 3) 任选地，在编码器和解码器中利用精确匹配的 IDCT 实施。
- 4) 任选地，排除所有形式的抖动。
- 5) 在解码期间的这种颜色空间转换步骤期间，利用附加的象素精度，仅在最终步骤处舍入。

动态范围和对比度扩展

本发明人已经在前面讨论了逐渐增加基于象素的图象表示的比色范围和动态范围，例如，见 Gray Demos 在 1990 年 10 月的 SMPTE 会议上提出的，在 SMPTE 杂志(1990 年 10 月，第 10 卷，第 10 期)发表的“*The Use of Logarithmic and Density Units for Pixels(象素的对数单位和密度单位的使用)*”。还有，见 Gray Demos 在 1993 年 10 月的 SMPTE 会议上提出的，并在进展和预发稿版上发表的“*An Example Representation for Image Color And Dynamic Range Which Is Scalable, Interoperable, and Extensible(象素压缩和可定标的、可中间操作的、以及可扩展的动态范围的示例表示)*”。

使用对数表示象素值有许多优点。例如，YUV 编码方法计算 U 作为 $R-Y$ ，计算 V 作为 $B-Y$ 。在对数表示中，U 变成 R/Y 而 V 变成 B/Y ，在 1993 年 SMPTE 文章“*An Example Representation for Image Color And Dynamic Range Which*

Is Scalable, Interoperable, and Extensible(象素压缩和可定标的、可中间操作的、以及可扩展的动态范围的示例表示)”中制造的术语中，两者都是“相同发光”，如上所述。即，在照射改变下，U和V两个通道都不包含信号，如果它们具有恒定的色调的话。在使用U和V通道对颜色编码时，这提供了高效率。此外，在总亮度的很宽的动态范围上得到这种色调编码效率。对数表示还提供容易的系统测量和定标方法，以及在极宽的亮度范围上在感知上是一致的。

表1表示在9到14位/象素范围中各种位数的范围和色调分辨率能力。从该表可以看到，使用对数表示，可以使用小于16位精度而达到在亮度范围、颜色范围以及色调分辨率(按1/8百分数步长)中的跨度为约10,000,000—到—1的范围的人的视觉范围。

总对比度范围	所需要的位数			
	1%步长	.5%步长	.25%步长	.125%步长
165: 1	9	10	11	12
30000: 1	10	11	12	13
700,000,000: 1	11	12	13	14

使用整个、半个、四分之一和八分之一百分数的对数步长所需要的位数

表 1

当前高质量电子摄像机和放映机能够在约10位/象素每颜色的动态范围和颜色范围处操作。例如，Polaroid//Philips LDK 9000摄像机提供具有宽颜色范围的低噪声图象。这个摄像机具有11微米的CCD象素尺寸以及约25,000个电子的完整的较优良电子计数。较大的传感器尺寸是极可能的，具有增加完整的较优良电子计数的潜力，以对成千上万或成百万的电子进行计数。在图象投射方面，具有10位伽马输入的微镜放映机能够得到具有适当色调差别的1000:1的动态范围，从而到达电影胶片的质量。虽然最佳胶片可以得到具有高色调保真度和宽颜色范围的较宽动态范围(约3000:1)，但是现在有可能预见数字成象和呈现在竞争并最终超过这种性能。

如上所述，通过保留在中间处理期间的经扩展的精度可以显著地提高运动图象压缩的质量。还可以利用相同的机构，较大地扩展可以有效压缩的图象信

息的动态范围，例如，如果保留中间精度的 14 位，则这个位范围可以按 $1/8$ % 对数步长来表示 7000,000,000:1 的动态范围。

还有用的是进行不包括在参考文章中的附加观察，亮度的区分是局部的。因此不可能区分紧邻极亮目标的暗阴影中的小的亮度变化。因此，只需要保留相对于图象所在的不同区域中的局部亮度的色调和动态范围区分。然而，图象的不同部分，可能具有暗阴影，可能离实质上看到细节区分的亮区域足够远，需要在局部亮度显示的色调范围中的相应的细节。

可以把结合了中间处理期间经扩展精度的使用的这些概念施加于运动图象压缩。尤其，一旦把保留在压缩系统中的中间精度扩展到较多位数，诸如 13、14、15 或 16 位，则这个经扩展的精度还能够显示宽动态范围图象。此外，为了得到有效的压缩，可以利用在亮区域中的色调区分的感知极限相对阴影中经扩展的区分的关系。

在 MPEG 类压缩系统中，通过“量化参数”或“QP”来确定色调区分。把 QP 分成 DCT 频率系数以减少对运动图象流编码所需要的位数。在解码期间，在计算 IDCT 之前使 QP 乘以 DCT 系数。因此，虽然在频率(DCT)空间中应用 QP，它仍表示色调精度参数。

根据上述区域区分特征，在高亮度区域中可以利用高 QP 而没有可见的清晰度或色调的损失，由于色调区分是相对于全亮度的。然而，在暗图象阴影区域中，必须利用低 QP 来提供精确的色调精度。

在 MPEG-1、MPEG-2 以及 MPEG-4 中利用 1-32 范围中的线性 QP 因子是普通的。而这适合于诸如 MPEG-2 提供的 8 位动态范围，对于更多位数(诸如 10 位或 12 位)，或对于较宽的动态范围，这个范围是不够的。在 MPEG-2 和 MPEG-4 中，有可能从一个宏模块到下一个宏模块改变 QP。这一般是一种机构，通过该机构调节位速率使之保持恒定位速率。较高的 QP 产生较少的编码位，而较低的 QP 产生较多的编码位。因此，在 1-32 的范围中改变 QP 是所需要的一切，以便在诸如 MPEG-1 或 MPEG-2 之类的 8 位能力的有限精度系统中保持给定的恒定位速率。然而，对于 10 位或 12 位精度，如在 MPEG-4 中，情景变化量较宽地改变(高情景强度)，并且需要低的恒定位速率，32 个可能值的 QP 范围可能就不够。对于 10 位的 32 个值的 QP 范围等效于 8 位的 8 个值的 QP 范围，只有四分之一的范围可用于诸如 MPEG-2 之类的 8 位编码系统。对于诸如 MPEG-4 之类的 12 位编码系统，32 个值的范围等效于 8 位编码的 2 个值的 QP

范围，只有QP范围的开始的十六之一可用于8位编码系统。

因此，在一般情况下扩展QP的范围是有用的。然而，注意，使用高精度中间编码和/或解码连同QP的范围和QP值之间的直接对应(即，值x与显示数x相同；因此，值14等于显示数14)可以导致高质量压缩和解压缩。

如果要求改变在帧中的QP，则希望在QP中保持较少的步长数目(诸如32个或64个值，或某些相似的小数目)，由于在MPEG-4中把对每宏模块的QP改变进行编码所需要的位限制为2个单元。如果每帧只改变一次QP，或每幅或相似大的结构改变一次，则QP的值的数目可能较大。

如果需要，通过取得动态范围的感知对数特性的优点，可以把QP的有效范围扩展到在较少数目显示代码中的确定值(即，在压缩和解压缩期间实际应用的QP值)的较宽范围。相应地，可以使用非线性查找表把较少数目的显示QP代码(诸如1到32，或1到128的范围)映射到确定QP值(诸如1到128、1到256、1到1024或1到4096)的更宽范围。在如此的表中，低QP代码输入几乎是一一对地映射以确定QP值。例如，QP代码1到4可能映射来确定1到4的QP输出值。然而，在对数模型中，映射将逐渐变得更陡峭，致使QP代码8可能映射来确定QP值16，QP代码16可能映射来确定QP值64，而QP代码32可能映射来确定QP值256。注意，MPEG-2对于从1-31的线性QP模式提供双步长QP模式，即映射每个从1-31的代码到它的值的两倍(即，到2-62)，以及把代码1-31映射到确定值1-112的非线性QP模式。在MPEG-2中，这些大的确定QP值导致使用8位象素值的极粗糙的编码。例如，62和112的QP值分别对应于8位值中只有两或一位的编码。因此，使用这些值的任何编码图象将具有极差的质量。

另一方面，如果每帧或每幅或其它大的结构规定QP一次，则不限制可用于QP的位数，并且可以在任何适当的位数(包括16位、32位、32位浮点，甚至更多的位数)中的极宽范围上完整地表示QP值。然而，宽动态范围图象的概念化框架是如此的，图象某些部分的亮度高，需要高的确定QP值，而其它部分亮度低，需要低的确定QP值。因此，使用在区域基础上有效地规定QP值的一种方法是有益的。如果限制QP代码的范围(诸如1到32，如同在MPEG-4中)，则允许QP值每宏模块变化 ± 2 单位的编码系统的现有机构(诸如MPEG-4)是足够的。然而，如果需要大的QP值范围，则规定区域性QP值的其它简单方法也是适当的和有益的。

因此，使用区域性变化的 QP 值以允许极宽动态范围显示是十分普通的，所述极宽动态范围显示是可以高度压缩的，并在视觉上与原始运动图象不能区分。

必须注意除以确定 QP 值之后的可用于编码显示的位数(还把 DCT 输出被 QP 除称为“量化”)。在量化之后，必须把余留的位编码成位流。除了内部帧和内部宏模块情况之外的经编码的位表示以前或后继帧，以及当前帧和当前宏模块中的最佳运动矢量预测匹配之间的差异。这个量化的 DCT—编码的不同的编码表示将确定可以得到的压缩比。

在 MPEG-2 和 MPEG-4 中，最大编码值范围是 ± 2047 (受到 VLC 表表示的限制)。这对应于 8 位的未量化的精度。因此，对于 10 位图象的未量化的(即 $QP = 1$)编码，有可能超过这个最大编码范围达一个 4 倍的因子。这可能发生，如果最佳匹配预测器块包括 DC 中的满刻度黑到白过渡，或等效的满刻度 AC 系数改变(诸如从平坦—灰度区域预测的清晰的黑白边缘)。最优化的预测器难得提供如此差的匹配，因此在这 10 位例子情况中，如果有过的话，也是偶尔要求满刻度编码的。然而，在 12 位运动图象或在甚至更宽的 14 或 16 位图象中，经常会超过范围极限，诸如 ± 2047 。在容易扩展 ± 2047 极限的同时，在概念上有益的是试图限制进行编码的位的平均数。通过使用 QP 值能直接使位的平均数以及最大编码范围两者减少。例如，为 4 的一个 QP 扩展了用 ± 2047 编码表示的可得到的动态范围，以包括 10 位编码的所有可能情况，以及 12 位运动画面编码的所有可能情况。因此可以看到，扩展 QP 值到更高数目的值，诸如 16、32 或 64，可以进一步扩展可以通过诸如 ± 2047 范围之类的受限制编码位的数目表示的运动图象动态范围。这种更高的 QP 值不但扩展范围，而且还减少进行编码的位的平均数，因此产生了高度的压缩。

这种概念的关键是在宽动态范围图象中的高 QP 值对应于不需要精细的色调区分的具有高亮度的图象区域。通过保持帧的扩展精度，以及通过在亮区域中利用高 QP 值和暗区域中利用低 QP 值而对宽动态范围图象进行编码的这种方法可以得到高压缩比，具有在感知上与原始编码不能区分的完美的编码。

因此，可以扩展 MPEG 类压缩技术的效用，用于压缩运动的宽动态范围图象。当摄像机、放映机以及其它图象系统部件扩展它们的范围能力时，可以施加这种压缩方法而得到高效的压缩。因此，在图象系统中的将来许多代的技术改进上，所产生的压缩系统是可以扩展的。

概括地说，本发明的一个方面包括通过扩展 QP 值的范围而把高精度压缩施加于宽动态范围图象。可以通过直接增加 QP 值的范围，或间接地通过非线性变换(诸如一种功能或查找表)而完成扩展，所述非线性变换把 QP 代码的小范围映射到用于除(压缩)或乘(解压缩)的 QP 值的较宽范围。本发明的另一个方面是根据区域性信息，或通过检查压缩期间可得到的信息(例如，DCT 系数，或候选 QP 值的给定数目产生的位数，从所述候选 QP 值选择一个适当的值)，或两种确定方法的组合来确定这种宽范围 QP 值。

还可以把经扩展的中间处理精度和经扩展的 QP 值范围应用于运动图象中的扩展的对比度范围区域和降低的对比度范围区域两者的编码。例如，在阴霾的日子使前景目标具有高对比度是很普通的，但是具有随距离降低的对比度。远离的目标经常具有极低的对比度。诸如在建筑物窗口后面以及汽车挡风玻璃后面的情景之类的其它普通情况，也由于玻璃和玻璃的反射而具有降低的对比度。反射也呈现降低的对比度。

可以把经扩展的精度和 QP 值范围的原理施加于图象的对比度降低区域以扩展精度，用该精度对低对比度部分进行压缩编码。正如对于暗区域使用低 QP 值和对于亮区域使用高 QP 值的动态范围扩展的情况一样，低对比度区域(通常较远)可以利用低 QP 值，而高对比度区域(通常是前景)利用高 QP 值。如此，如果摄像机传感器具有足够的位分辨率，则当解压缩时，完全可以区分在空中的、或地面上的、或阴霾日子的远目标。接着可以人工地增强它们的对比度，从而显示具有正常动态范围的清晰图象。

虽然当前摄像机和胶片限制到约 10 位的伽马或对数动态范围，将来的摄像机十分可能具有更高的精度。除了扩展动态范围之外，这种经扩展的摄像机图象亮度区分对于观看低对比度区域中的细节是有用的。正如摄像机的情况一样，当放映机扩展它们的动态范围和它们的最大亮度时，有可能区分这个经扩展的范围中的低对比度细节。在宽动态范围放映机上比在有限动态范围的观看监视器上更容易看到诸如落下的雨滴之类亮度中的精细变化。容易区分通过宽动态范围放映机显示的目标，因为它对于观看者具有较宽范围的亮度变化，而在计算机 CRT(阴极射线管)显示器上，亮度变化范围较小。因此，当摄像机和显示器扩展它们的动态范围和色调区分(即，增加附加的分辨率位使之超过当前 10 位的能力)时，希望不但扩展动态范围而且还扩展对比度区分。

重要的是，与支持经扩展的动态范围的技术相同技术还支持低对比度区域

的高区分编码。尤其，以与通常进行特别编码(以提高 DC 项编码效率)的 DC 系数的方式不同的方式，把 QP 值应用于 DCT 输出的 AC 系数。因而 AC 系数的标度自然地大约复盖了主要 DC 值。例如，在模糊灰暗的低对比度区域具有在 DC 模糊平均值周围的低幅度 AC 系数。因此，施加低 QP 值将自然地保存在低对比度区域中的精细的色调变化。正如经扩展的动态范围的情况一样，高 QP 值允许高对比度前景区域的正常编码。

为了调节 QP 值使之适合于暗区域和模糊区域，还适合于正常对比度的亮度区域，应该相对于每个画面区域的对比度和亮度而区域性地确定 QP 值。如果对于每个宏模块设置 QP 值，致使在一个区域中的每个宏模块产生近似相同的位数，则也可以自动地确定 QP 值。对于具有包括暗区域和低对比度区域的宽动态范围图象，把恒定的位数给予每个宏模块将自动地使亮度和对比度的整个范围上的表示最优化。然而，还希望把比提供给低细节区域的位数更多的位提供给高细节区域，以及向运动区域提供比静态区域更多的位数。

通过检查每个宏模块中的 DCT 系数的相对幅度可以自动地确定每个宏模块的 QP 值。可以向包括表示细节的 DCT 系数和表示运动的 DCT 系数的宏模块提供比那些包括表示低细节或低变化(运动)的系数的相对权重的宏模块更多的位。然而，还须要考虑摄像机传感器的噪声，因为噪声的性能就象变化(运动)和细节(高频系数)两者一样。当与真的宽动态范围和合适的低噪声电平的高区分传感器一起使用时，相对于 DCT 系数自己加权的 DCT 系数可以形成合适的指示符(indicator)，用于自动地设置 QP 值。尤其，较大的 DCT 系数产生较大的 QP 值。相应地，可以完全跟据经验确定 DCT 系数和所要求的对应 QP 值之间的映射或相关。

诸如最大区域亮度和对比度之类简单的区域性算法是可以用来确定(或与其它机构一起帮助确定)合适的 QP 值的另一种方法。通过区域性细节幅度(画面的高频静态)测量算法还可以提供附加的信息。每种方法具有它自己的优点。然而，在内部宏模块的情况中，DCT 系数自己足以确定 QP 值。即，DCT 系数是与预测的宏模块的运动组合的细节的量度，所以使用独立的细节量度(诸如平行的内部 DCT 变换)可以帮助从当前帧图象宏模块的细节(在运动补偿之后的静态，诸如具有慢运动摄像机的草的叶片(grass of blades))析出运动细节变化(诸如雨滴或水平方向上运动的水波浪)。使用 DCT 系数本身来表示 QP 的简单性使之成为特别吸引人的实际实施。

图 5 是概括在图象压缩期间扩展动态范围和/或对比度的一种较佳方法的流程图。

步骤 500: 以宽动态范围或宽对比度范围的源画面开始。

步骤 502: 如果需要, 扩展量化参数(QP)代码组的有效值范围。例如, 这可以通过下列技术中之一来完成:

1) 把 QP 代码的组从额定范围(一般是 32 级)扩展到较大的实际范围(128、1024 或 4096 级, 或无论什么适合于图象范围的)。因此, 额定值直接表示可能值的经扩展的范围。

2) 使用非线性查找表或映射函数, 以非线性的方式使额定 QP 代码相关成值的较大有效范围。在低值处, 映射一般是线性的, 但是当值向一般范围最大值增加时, 增加有效的 QP 乘和除步长增加大小。例如, 使用非线性查找表或映射功能可以扩展 32 或 64 个代码, 以产生具有诸如 128、1024、4096 或无论什么适合于图象范围之类的较大最大值的较大有效范围。

步骤 504: 确定在进行压缩的一个图象的每个宏模块的应该编码的 QP 值, 最好使用下列方法中之一:

1) 使用算法确定合适的 QP 值, 以从运动图象流的分析确定每个帧中的局部图象区域的对比度, 在每个帧中的局部动态范围, 在每个帧中的局部细节幅度以及一个或多个顺序帧之间的局部运动(如上所述)。

2) 在量化之前, 根据在压缩过程中产生的信息, 根据所产生的位数(对于 QP 的许多候选值), 和 DCT 系数的幅度和频率确定 QP 值。

3) 应用来自 1) 和 2) 的信息的组合, 利用区域性信息以及从压缩过程产生的信息确定每个宏模块的 QP 值。

步骤 506: 如上所述, 对于所有中间处理使用经扩展的精度, 以使用从步骤 504 确定的 QP 值对图象进行压缩。可以按需要存储或发送经压缩的图象和相关联的标称 QP 代码, 所述相关联的标称 QP 代码对应于在压缩期间使用的所确定的 QP 值。

步骤 508: 使用如上所述的高精度解压缩, 对存储的或发送的图象进行解压缩, 使之成为供各种应用的宽动态范围、宽对比度范围、高分辨率图象。如果需要, 把这种解压缩的相关联的标称 QP 代码映射回对应的确定的 QP 值。这种应用包括电影和运动的家庭和戏院显示, 把存储的图象存档, 运动图象显示的商务应用, 政府应用(例如, 监督、军事命令和控制)以及等等。可以在宽动

态范围显示器装置上观看经解压缩的图象, 和/或为了提供最优化分析, 用作从(或需要)高质量宽动态范围图象而得益的使用算法的图象分析的源(这种算法不属于本揭示的范围)。

实施

可以在硬件或软件、或它们的组合(例如, 可编程逻辑阵列)中实施本发明。除非另行规定, 不使作为本发明的一部分包括在内的算法与任何它计算机或其它设备固有地相关。尤其, 可以使用具有根据这里的学说编写的程序的各种通用用途机器, 或可以更方便地构成更专用(例如, 集成电路)的设备, 以执行所需要的方法步骤。因此, 可以在一个或多个可编程计算机系统(每个计算机系统包括至少一个处理器、至少一个数据存储系统(包括易失性和非易失性存储器 and/或存储单元)、至少一个输入装置或端口以及至少一个输出装置或端口)上执行的一个或多个计算机程序中实施本发明。把程序代码施加于输入数据以执行这里描述的功能以及产生输出信息。按已知的方式把输出信息施加于一个或多个输出装置。

可以按任何要求的计算机语言(包括机器、汇编或高级过程的、逻辑的或面向对象的编程语言)来实施每个如此的程序, 以与计算机系统通信。在任何情况中, 语言可以是编译语言或解译语言。

最好把每个如此的计算机程序存储在、或下载到可以通过通用途或专用的可编程计算机读出的存储媒体或器件(例如, 固态存储器或媒体、或磁性媒体或光学媒体)中, 当计算机系统读出存储媒体或器件时, 用于配置和操作计算机以执行这里描述的过程。还可以考虑实施本发明的系统作为配置有计算机程序的计算机可读存储媒体, 如此配置的存储媒体导致计算机系统在规定和预定的方式中操作以执行这里描述的功能。

已经描述了本发明的许多实施例。但是, 可以理解, 可以进行各种修改而不偏离本发明的精神和范围。例如, 上述某些步骤可以有独立的排序, 因此可以以与所描述的次序不同的次序来执行。相应地, 其它实施例在下列权利要求书的范围内。

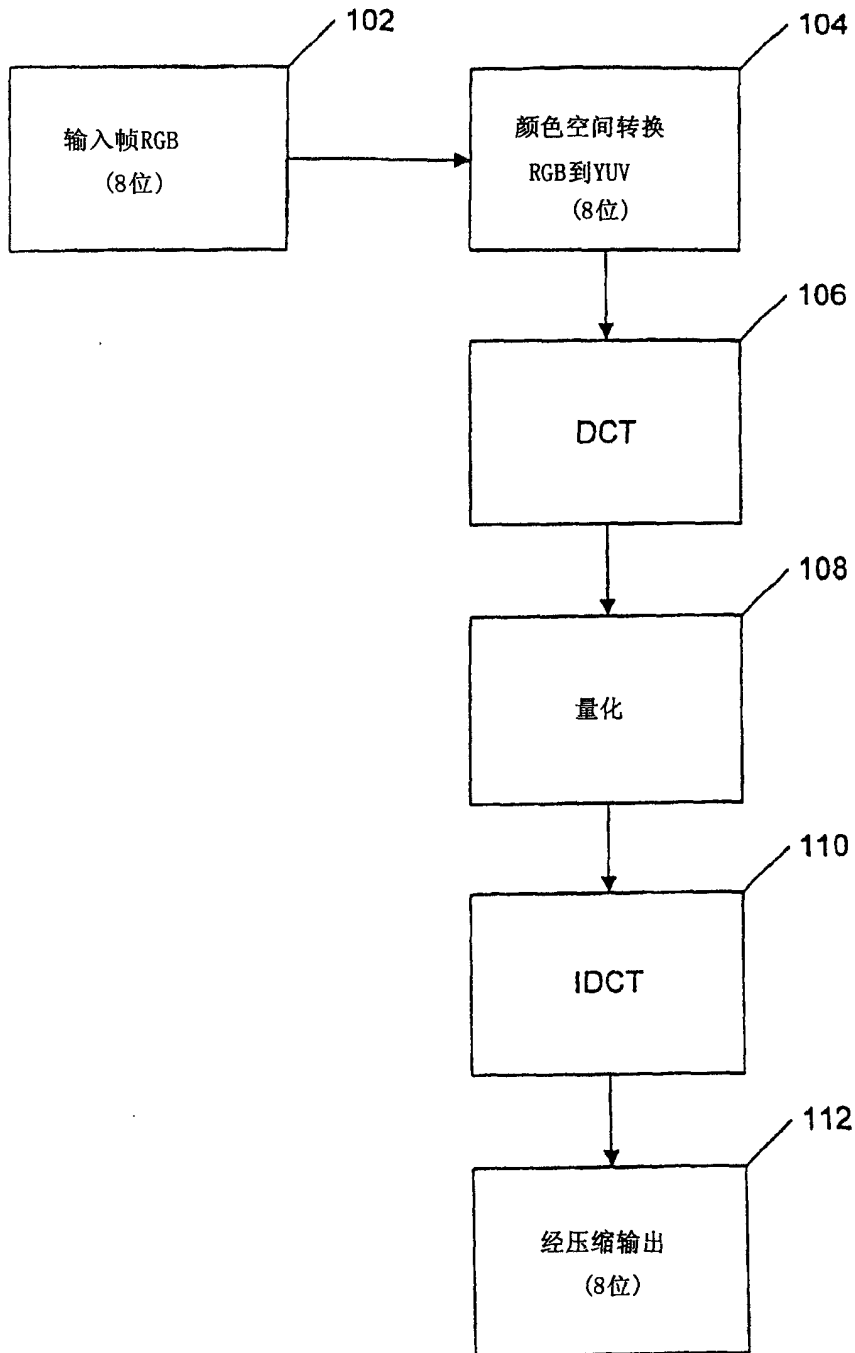


图 1

现有技术MPEG-2编码器

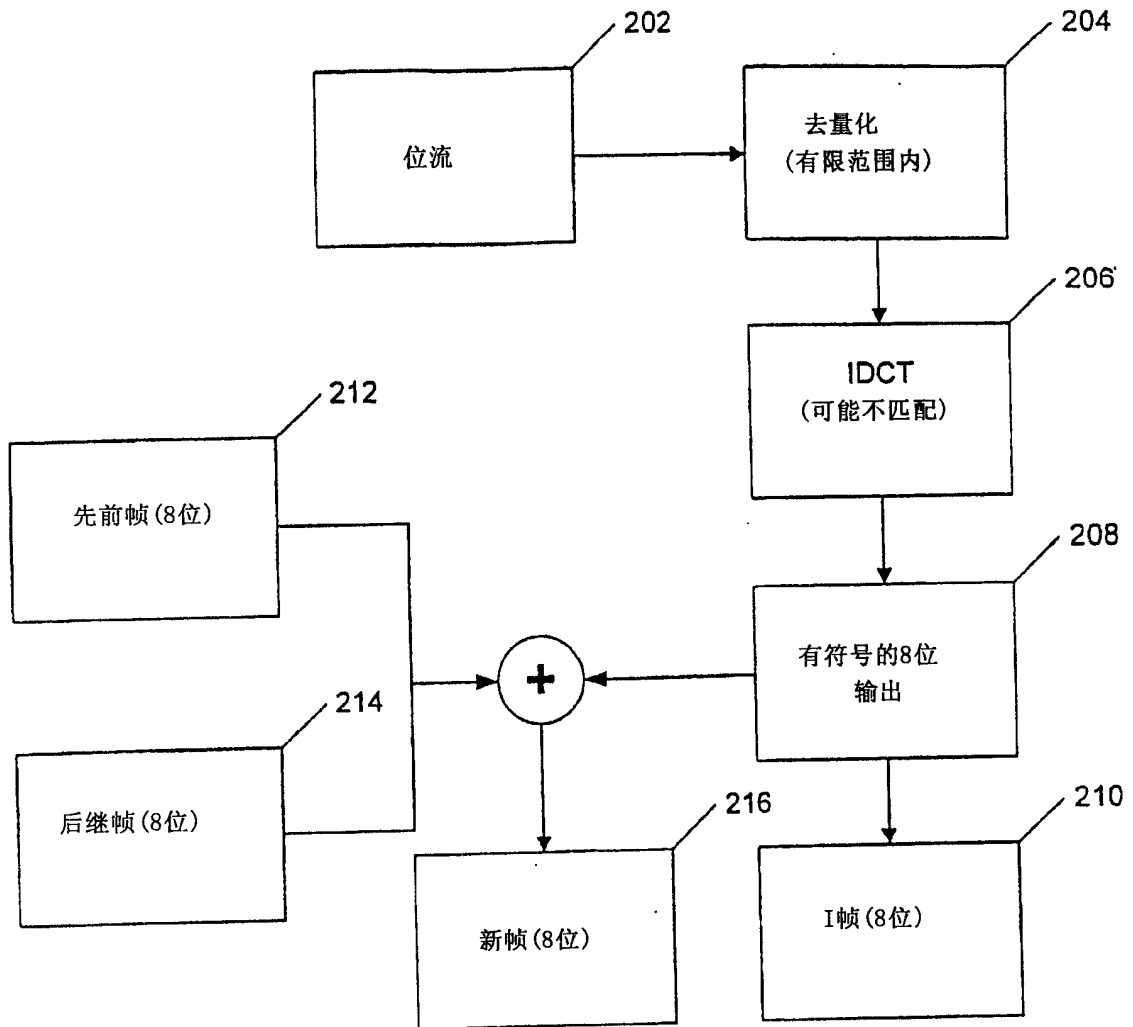


图 2
现有技术MPEG-2编码器

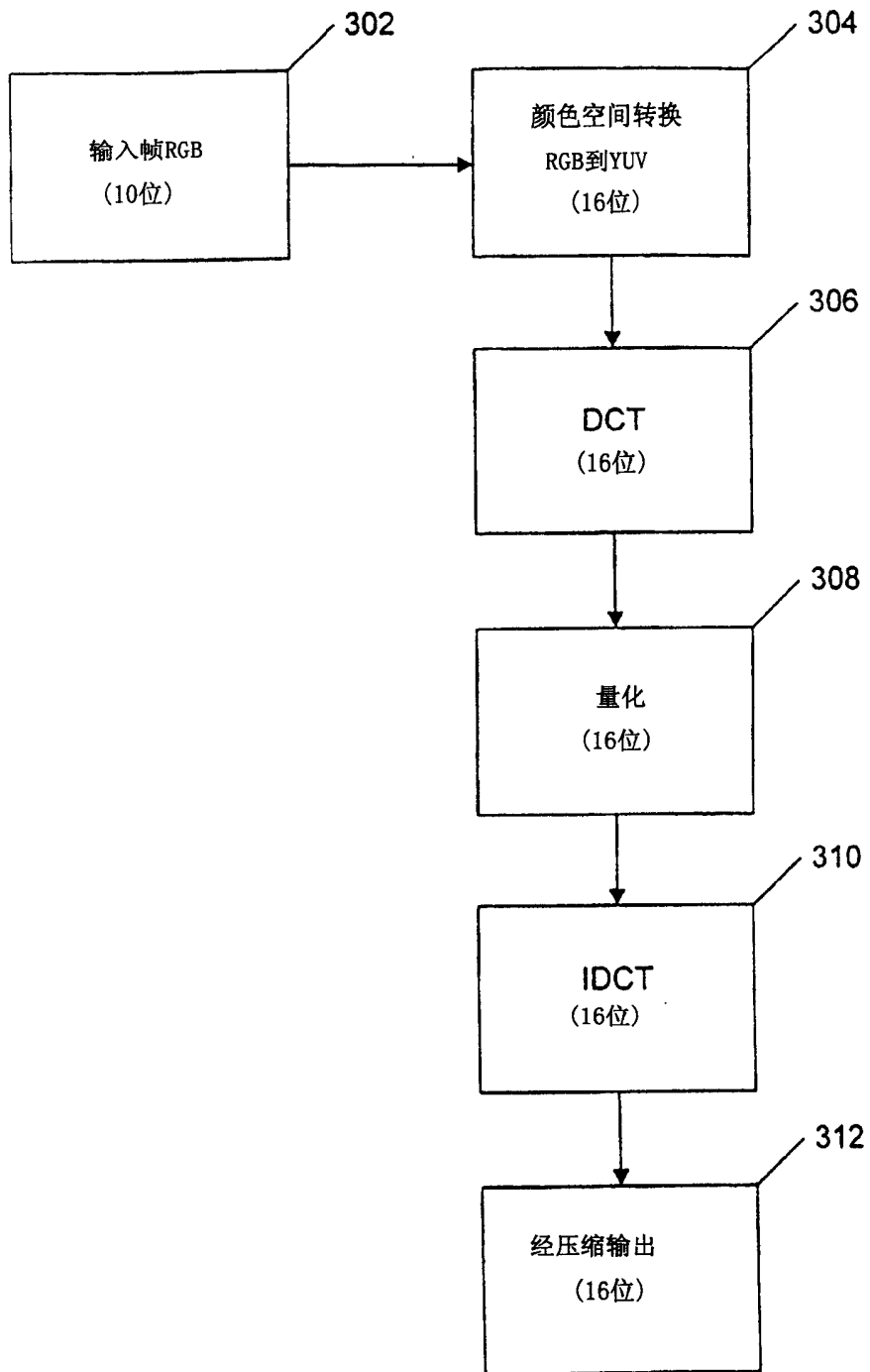


图 3

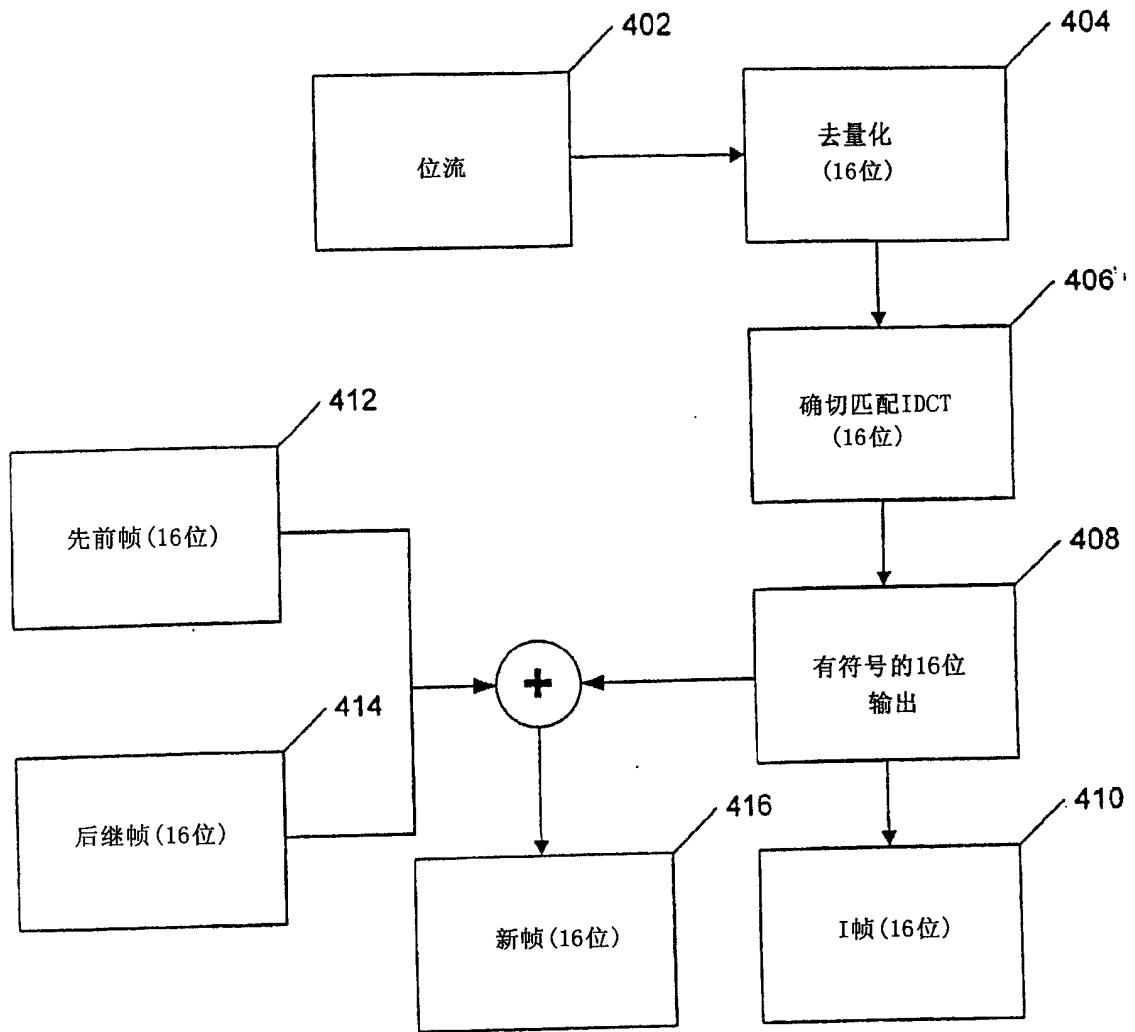


图 4

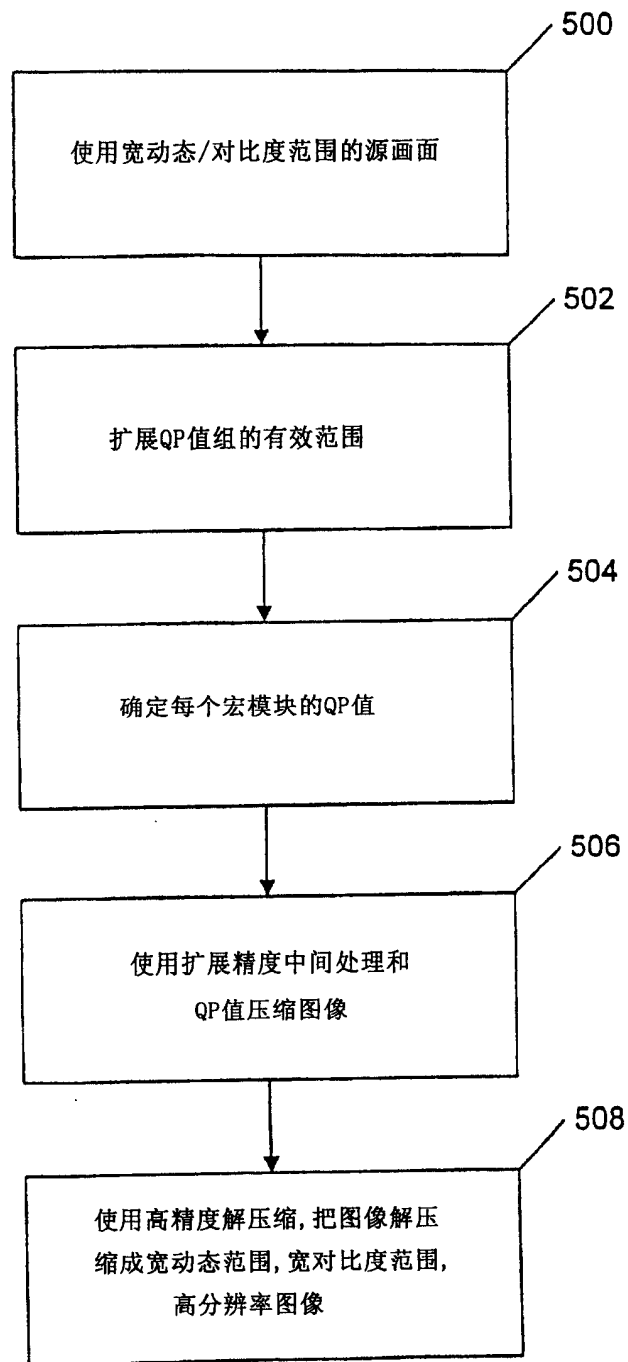


图 5