



(12) 发明专利

(10) 授权公告号 CN 101653007 B

(45) 授权公告日 2011.06.08

(21) 申请号 200880011102.6

(22) 申请日 2008.03.26

(30) 优先权数据

11/693,958 2007.03.30 US

(85) PCT申请进入国家阶段日

2009.09.30

(86) PCT申请的申请数据

PCT/US2008/003942 2008.03.26

(87) PCT申请的公布数据

W02008/121281 EN 2008.10.09

(73) 专利权人 伊斯曼柯达公司

地址 美国纽约州

(72) 发明人 P·W·琼斯

(74) 专利代理机构 中国专利代理(香港)有限公

司 72001

代理人 王岳 王忠忠

(51) Int. Cl.

H04N 7/30(2006.01)

(56) 对比文件

US 2001/0014120 A1, 2001.08.16, 全文.

CN 1526204 A, 2004.09.01, 全文.

US 6614942 B1, 2003.09.02, 全文.

CN 1663258 A, 2005.08.31, 全文.

Damon M Chandler etc. Contrast based quantization and rate control for wavelet coded images. 《Image Processing. 2002. Proceedings. 2002 International Conference》. 2002, 第3卷 III-233 - III-236.

Majid Rabbani etc. an overview of the JPEG 2000 still image compression standard. 《Signal Processing: Image Communication》. 2002, 第17卷(第1期), 3-48.

审查员 李鹏

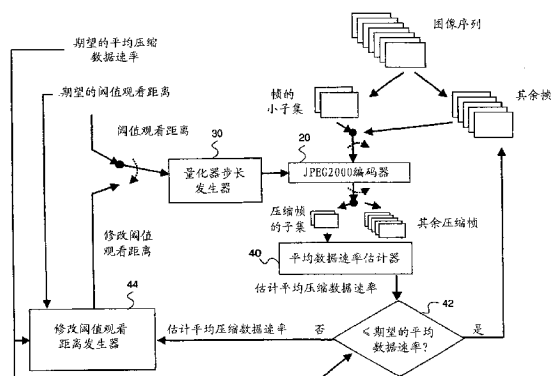
权利要求书 2 页 说明书 16 页 附图 19 页

(54) 发明名称

控制压缩数据量的方法

(57) 摘要

公开了一种用于压缩一组数字图像以便在整个组内实现平均压缩数据速率的方法。从该组图像中选择图像的子集。指定阈值观看距离和期望的平均压缩数据速率,并根据阈值观看距离来确定压缩参数。用压缩参数来压缩图像的子集以产生压缩数据,并对于来自图像的子集的压缩数据计算平均压缩数据速率。将图像的子集的平均压缩数据速率与期望的平均数据速率相比较。如果图像的子集的平均压缩数据速率小于或等于期望的平均压缩数据速率,则使用被用于图像子集的压缩参数来压缩不在图像子集中的来自图像序列的其余图像。否则,将阈值观看距离修改为更大的距离并重复压缩过程。



1. 一种用于压缩一组连续数字图像以实现用于该组图像的平均压缩数据速率的方法，包括步骤：

- (a) 指定阈值观看距离和期望的平均压缩数据速率；
- (b) 从连续图像组中选择图像的小子集；
- (c) 应用阈值观看距离作为用于计算压缩参数的输入；其中，所述压缩参数是量化器步长；
- (d) 用所述压缩参数来压缩图像的所述小子集以产生压缩数据；
- (e) 计算图像的小子集的平均压缩数据速率；
- (f) 将图像的小子集的平均压缩数据速率与期望的平均压缩数据速率相比较；
- (g) 如果图像的小子集的平均压缩数据速率小于或等于期望的平均压缩数据速率，则在不重新压缩图像的小子集的情况下使用被用于图像的小子集的压缩参数来压缩来自图像组的不在图像的小子集中的其余图像；
- (h) 否则，修改所述阈值观看距离并重复上述步骤 (c) 至 (h)。

2. 如权利要求 1 所述的方法，其中，在通过图像的小子集的第一遍压缩之后的修改阈值观看距离的步骤 (h) 还包括使用平均压缩数据速率和所述阈值观看距离来确定速率失真函数的估计斜率并使用该估计斜率来计算将实现期望的平均压缩数据速率的修改阈值观看距离。

3. 如权利要求 2 所述的方法，其中，在通过图像的小子集的第二遍压缩之后的修改阈值观看距离的步骤 (h) 还包括使用来自第一遍压缩的平均压缩数据速率和阈值观看距离及来自第二遍压缩的平均压缩数据速率和阈值观看距离来确定速率失真函数的计算斜率并使用该计算斜率来计算将实现期望的平均压缩数据速率的修改阈值观看距离。

4. 如权利要求 3 所述的方法，其中，在通过图像的的子集的至少第三遍压缩之后的修改阈值观看距离的步骤 (h) 还包括使用来自前一遍压缩的平均压缩数据速率和阈值观看距离及来自当前遍压缩的平均压缩数据速率和阈值观看距离来确定速率失真函数的计算斜率并使用该计算斜率来计算将实现期望的平均压缩数据速率的修改阈值观看距离。

5. 如权利要求 1 所述的方法，其中，所述阈值观看距离被输入到对比度灵敏度函数计算器中以用于计算压缩参数。

6. 一种用于产生一组连续图像的恒定感知图像质量的 JPEG2000 可变位率 (VBR) 压缩方法，所述方法同时还使压缩数据量最小化以实现前述图像质量，其中，所述方法包括步骤：

- (a) 指定阈值观看距离和期望的平均压缩数据速率；
- (b) 从连续图像组中选择图像的小子集；
- (c) 应用阈值观看距离作为用于计算压缩参数的量化器步长发生器的输入；
- (d) 用所述压缩参数来压缩所述图像的小子集以产生压缩数据；
- (e) 计算图像的小子集的平均压缩数据速率；
- (f) 将图像的小子集的平均压缩数据速率与期望的平均压缩数据速率相比较；
- (g) 如果图像的小子集的平均压缩数据速率小于或等于期望的平均压缩数据速率，则在不重新压缩图像的小子集的情况下使用被用于图像的小子集的压缩参数来压缩来自所述图像组的不在图像的小子集中的其余图像；

(h) 否则,修改所述阈值观看距离并重复上述步骤 (c) 至 (h)。

7. 如权利要求 6 所述的方法,其中,在通过图像的小子集的第一遍压缩之后的修改阈值观看距离的步骤 (h) 还包括使用所述平均压缩数据速率和所述阈值观看距离来确定速率失真函数的估计斜率并使用该估计斜率来计算将实现期望的平均压缩数据速率的修改阈值观看距离。

8. 如权利要求 7 所述的方法,其中,在通过图像的小子集的第二遍压缩之后的修改阈值观看距离的步骤 (h) 还包括使用来自第一遍压缩的平均压缩数据速率和阈值观看距离及来自第二遍压缩的平均压缩数据速率和阈值观看距离来确定速率失真函数的计算斜率并使用该计算斜率来计算将实现期望的平均压缩数据速率的修改阈值观看距离。

9. 如权利要求 8 所述的方法,其中,在通过图像的子集的至少第三遍压缩之后的修改阈值观看距离的步骤 (h) 还包括使用来自前一遍压缩的平均压缩数据速率和阈值观看距离及来自当前遍压缩的平均压缩数据速率和阈值观看距离来确定速率失真函数的计算斜率并使用该计算斜率来计算将实现期望的平均压缩数据速率的修改阈值观看距离。

10. 如权利要求 6 所述的方法,其中,所述阈值观看距离被输入到对比度灵敏度函数计算器中以用于计算压缩参数。

控制压缩数据量的方法

技术领域

[0001] 本发明总体上涉及数字图像处理领域,尤其涉及用于在 JPEG2000 图像压缩中高效地控制压缩数据速率 (compressed data rate) 的方法。

背景技术

[0002] JPEG2000 是提供一组丰富的特征以满足不同应用的需要的尖端图像压缩技术。这些特征由于形成 JPEG2000 压缩系统的各种组件的性质而可以实现。下文提供了对这些组件的简要说明,且更完整的说明可以在 Signal Processing:Image Communication, Vol. 17, pp. 3-48, 2002 中的由 M. Rabbani 和 R. Joshi 所著的题为 “An overview of the JPEG 2000 stillimage compression standard” 的论文中找到。

[0003] 图 1 举例说明 JPEG2000 压缩系统中的基本组件。原始图像数据被发送到预处理器 10,在那里,可以对原始图像进行若干修改。一种可能的修改是称为不可逆颜色变换 (ICT) 的颜色变换。此变换被设计为将红-绿-蓝 (RGB) 图像数据转换为亮度 (Y) 和两个色度 (Cb、Cr) 分量以改善压缩性能。然后,使用离散小波变换 12 来变换经预处理的数据,离散小波变换 12 将空间图像数据转换成对应于多个空间频率子带的小波系数。用具有恒域 (deadzone) 的均匀量化器 14 将该小波系数量化。量化是以由已量化的小波系数重构的图像中的误差为代价降低用来表示小波系数的数据量的多对一映射。对于每个频率子带 b 的量化程度由与子带相关的量化器步长 (图 1 中表示为 $Q(b)$) 来确定。在 JPEG2000 系统中常常指定用于量化的单个基础步长,然后通过使用小波变换 (即从每个频率子带的频带增益因数) 的性质来导出对应于各种频率子带的一组量化器步长。可替换地,可以对频率子带指定一组明确的量化器步长以便更精确地控制量化。

[0004] 然后,将已量化的小波系数划分成小代码块,并使用自适应二进制算术编码器 (adaptive binary arithmetic encoder) 16 将来自每个代码块的量化系数编码为位面 (bitplane)。此步骤称为第一层 (Tier-1) 编码。最后,通过使用位流组织器 (bitstream organizer) 18 以各种方式布置已编码代码块数据来产生压缩代码流。此步骤称为第二层 (Tier-2) 编码。在第二层编码中,可以丢弃对应于已量化的小波系数的一个或多个位面的已编码数据,这将以解压缩图像中的误差增加为代价降低压缩数据量。从给定小波系数中丢弃位面等同于该系数的较粗 (coarser) 量化。如果将最初应用于系数的量化器步长表示为 Δ , 则丢弃 k 个最低有效位面之后的有效量化器步长为 $2^k \cdot \Delta$ 。在美国专利 6,650,782 中由 Joshi 等人描述了系数位面的丢弃与所得量化程度之间的这种关系。

[0005] JPEG2000 中所实现的灵活性以高计算复杂性为代价,且期望在仍提供期望水平的图像质量的同时降低这种复杂性。这种高计算复杂性在运动图像应用中尤其是个问题,在运动图像应用中,在序列中存在对于编码系统来说必须以高吞吐量来编码以便节省时间和计算资源的大量帧。JPEG2000 编码器中的计算负荷的重要部分在使用算术编码器来将已量化的小波系数编码以形成压缩数据的第一层编码期间发生。在第一层编码之后,常常在第二层编码期间丢弃大量的压缩数据以实现用于压缩代码流的预指定大小,或实现在解压缩

期间重构的图像中的预指定失真水平。改变压缩数据量的此过程称为速率控制。

[0006] JPEG2000 的一种常见速率控制方法是后压缩速率失真优化 (post-compression rate-distortion optimization, PCRD-opt)。PCRD-opt 的某些变体可在大多数 JPEG2000 软件和硬件编码器中找到,且完整的说明可以在 Kluwer Academic Publishers, Boston, MA, 2002 的、D. S. Taubman 和 M. W. Marcellin 的题为 JPEG2000 Image Compression Fundamentals, Standards and Practice 的一书的第 339 ~ 348 页中找到。应用 PCRD-opt 的典型目的是与图像内容无关地实现压缩代码流的固定大小。产生固定或恒定压缩大小 (compressed size) 的压缩系统称为恒定位率 (CBR)。在采用 PCRD-opt 的典型 JPEG2000 实施方式中,小波系数被细 (finely) 量化,这在第一层编码期间产生大量压缩数据。PCRD-opt 算法通过分析丢弃数据对压缩代码流大小的影响和在解压缩图像中导致的所产生失真来确定应丢弃哪些数据以满足预指定大小。目的是使在实现压缩数据的期望最终大小的条件下的总失真最小化。使用通过 PCRD-opt 算法细量化的小波系数的优点是精确地控制压缩数据量以便代码流大小尽可能接近于预指定限度而不超出限度。然而,这种方法也是耗时的,因为在第一层编码中由细量化系数来形成压缩数据需要相当多的计算,而大部分压缩数据在随后的速率控制过程期间被丢弃。这是过度编码的问题。还请注意,对过压缩数据的评估还增加第二层编码期间的处理时间。

[0007] 虽然通常使用 PCRD-opt 算法来使经受最大压缩代码流大小方面的约束的总失真最小化,但是还可以使用 PCRD-opt 算法的变体来使经受对总失真的约束的压缩代码流大小最小化。此类系统将产生近似恒定的失真且压缩代码流大小将随着图像内容而波动。试图通过允许压缩大小随图像内容而变来产生恒定质量的压缩系统称为可变位率 (variable bitrate, VBR)。相比于 CBR 系统,通常优选 VBR 系统,这是因为对于内容改变的图像系统 (ensemble) 的相当图像质量,VBR 系统通常比 CBR 产生更小的平均压缩大小。然而,从计算的观点出发,如果在第一层编码中的算术编码器之前将变换系数细量化,则 VBR 系统仍遭受过度编码的问题。

[0008] Becker 等人 (美国申请 2005/0100229) 讲授了用于使运动图像序列中的过度编码量最小化的一种方法。然而,这种方法需要精确地估计某些参数,并需要附加逻辑作为 PCRD 优化的替代或扩展。大多数 JPEG2000 编码器将需要修改以实施 Becker 等人所讲授的方法,且该方法还不适合于压缩单独的图像帧。更重要的是,当图像内容快速改变时,诸如在运动序列中的场景边界处发生的那样,则这种方法可能失效。

[0009] 无论系统是 CBR 还是 VBR,当前系统中的另一个限制是通常用来量化图像质量的失真度量是均方误差 (MSE)。MSE 是对原始图像的像素值与解压缩图像的像素值之间的数值差的测量。与 MSE 相关的度量是使用分贝 (dB) 单位指定的峰值信噪比 (PSNR)。这些度量在数学上是方便的,但是众所周知的是它们并不总是与感知的质量相关。结果,即使获得特定的 MSE 或 PSNR,也不能保证给定水平的感知图像质量。对于许多应用,目标可能是视觉无损质量,即当在指定观看条件下由人类观看者来观看解压缩图像和原始图像这两者时,解压缩图像看起来与原始图像相同。用于压缩图像的 MSE 或 PSNR 的知识未提供关于在其下将获得视觉无损质量的观看条件的任何信息。

[0010] 例如,给定具有 40dB 的 PSNR 的压缩图像,不可能预先知道什么样的观看条件将为图像产生视觉无损质量。可能在指定的观看条件下感知质量不是视觉无损的,或者可能获

得视觉无损质量,但使用了比实现该质量所需的更多的压缩数据。第一种情况导致所感知的图像质量比所期望的低,而第二种情况导致压缩代码流比视觉无损质量所需的大。因此,作为质量度量,严格依赖于 MSE 或 PSNR 是不够的。这种情况在希望产生已知水平的感知图像质量、同时还将图像压缩至期望质量所需的最低数据量的 VBR 系统中尤其如此。

[0011] 在某些 JPEG2000 应用中出现的另一问题是需要满足每个图像的最大压缩代码流大小。一种此类应用是数字电影 (DCinema) 压缩。作为美国主要演播室财团的数字电影倡导联盟 (Digital Cinema Initiative, DCI) 最近已开发了一套规范以保证 DCinema 设备制造商之间的一定水平的互操作性。一种规范提倡对每个图像的最大压缩代码流大小的约束,有对可以用来表示图像中的每种颜色分量 (color component) 的最大压缩数据量的其他限制。对于每秒 24 帧的图像序列中的每个图像帧,对最大压缩代码流大小的约束是每帧 1,302,083 个字节。可以将对最大代码流大小的这种约束等效地指定为 250 兆位 / 秒 (Mbps) 的瞬时最大压缩数据速率。图像帧中的每种颜色分量的压缩数据量可以是最大值 1,041,066 字节,对于每秒 24 帧的图像序列来说其等效于 200Mbps 的瞬时压缩数据速率。这些规范主要由当前 JPEG2000 解压缩系统的能力促成。虽然 CBR 系统可以始终通过诸如 PCRD-opt 的速率控制算法来满足这些数据约束,但其由于与 VBR 系统相比通常较大的 CBR 的平均代码流大小而不是优选解决方案。因此,期望具有可能地最大程度上提供恒定的感知质量、同时仍满足对每个图像的最大压缩数据量的任何约束的 VBR 系统。

[0012] 除对图像序列中的帧分量 (frame component) 或每个图像帧的这些瞬时速率约束之外,还可能期望约束整个序列的压缩代码流的总大小。这样,可以确保给定持续时间的电影将安装到具有某容量的存储介质上,例如两小时的电影被压缩到 160 千兆字节 (Gbyte) 的硬盘驱动器上。如果要通过通信网络来发送压缩电影内容而不是将其存储在物理介质上以便从内容提供商输送到电影院,则压缩代码流的总大小同样重要。压缩代码流的较小总大小意味着可以将数据更快且更具成本效益地发送到剧院。无论使用物理存储介质还是通信网络,优选压缩解决方案都是 VBR 编码,因为对于相同的总代码流大小,其将比 CBR 编码产生更高的总体图像质量。

[0013] 对压缩代码流的总大小的约束等效于对整个序列上的平均数据速率的约束,为方便起见,以下讨论将使用平均数据速率来代替总文件大小。例如,如果一个人具有长度为 2 小时的电影且希望将压缩电影安装到 160 千兆字节 (Gbyte) 的硬盘驱动器上,则平均压缩数据速率必须小于 182Mbps。由于可变位率编码,所以某些帧将使用更高的瞬时压缩数据速率 (在服从 DCI 的代码流的情况下达到最大 250Mbps),而其它帧将使用基本小于 182Mbps 的瞬时数据速率。尽管瞬时压缩数据速率改变,但目的是对于每个帧都具有恒定的图像质量。

[0014] 为了在保证恒定图像质量的同时用 VBR 系统来实现期望的平均数据速率,需要执行某种类型的多遍 (multipass) 编码。这是因为预先并不知道图像内容的复杂度,因此预先也不知道对于给定压缩参数组的压缩数据速率。如果需要用不同的压缩参数来反复压缩整个内容直至实现期望的平均数据速率,则工作流程效率将大大降低,因此需要一种更高效的方法。可以利用 PCRD-opt 算法来剪裁 (trim) 压缩帧,但这也是耗时的,因为必须首先收集序列中的所有帧的速率失真统计值,然后确定丢弃哪些数据以便在同样保持 MSE 或 PSNR 所量化的恒定图像质量的同时满足平均速率要求。

[0015] 改善效率的一种方法是使用压缩参数的初始设置来仅仅处理图像序列中的数据子集,并随后使用所得的平均速率作为所有帧的平均速率的估计。如果估计的平均速率大于期望速率,则修改压缩参数并再次压缩图像数据的子集。一旦用该图像数据子集满足了对期望平均数据速率的约束,则可以用所确定的压缩参数来压缩所有图像帧。Honsinger 等人的 US 6,356,668(被转让给 Kodak)描述了用 JPEG 压缩进行高效速率控制的方法,其使用单个图像内稀疏采样的图像区域来估计整个图像的速率。另外,Honsinger 等人的专利描述了阈值观看距离(threshold viewing distance)作为质量度量在确定 JPEG 压缩参数中的使用,及其在构造速率失真(R-D)曲线中的使用,如果需要附加的迭代来满足目标数据速率,则所述速率失真曲线可以用于调节压缩参数。

[0016] 然而,Honsinger 等人提出的这种现有技术方法在应用于数字电影应用中用于图像序列压缩的 JPEG2000VBR 压缩时具有限制。首先,利用 JPEG2000 编码,图像的稀疏空间采样由于小波变化的性质而不具有高效率(与在 JPEG 压缩中使用的 DCT 相比)。其次,Honsinger 的方法可能需要很多次的压缩迭代以构造 R-D 曲线和实现期望的平均数据速率。每次压缩迭代需要附加的计算,这可能导致整体压缩过程效率低。

[0017] 因此,需要一种用于图像序列的 JPEG2000VBR 压缩系统,其:(1)使过度编码的量最小化以便降低计算复杂性;(2)产生压缩代码流,其以计算上高效的方式满足对单独帧和帧分量及整个压缩代码流两者的最大压缩数据量的约束;以及(3)产生恒定的感知图像质量,同时使压缩数据量最小化以实现该水平的图像质量。

发明内容

[0018] 本发明旨在克服上文阐述的问题中的一个或多个。简而言之,根据本发明的一个实施例,公开了一种用于压缩一组数字图像以便在整个组上实现平均压缩数据速率的方法,包括步骤:

[0019] a) 指定阈值观看距离和期望的平均压缩数据速率;

[0020] b) 从图像组中选择图像子集;

[0021] c) 根据所述阈值观看距离来确定压缩参数;

[0022] d) 用所述压缩参数来压缩所述图像子集以产生压缩数据;

[0023] e) 计算来自所述图像子集的压缩数据的平均压缩数据速率;

[0024] f) 将所述图像子集的平均压缩数据速率与期望的平均压缩数据速率相比较;

[0025] g) 如果所述图像子集的平均压缩数据速率小于或等于期望的平均压缩数据速率,则使用被用于所述图像子集的压缩参数来压缩来自所述图像组但不在所述图像子集中的其余图像;

[0026] h) 否则,修改所述阈值观看距离并重复步骤 c) ~ h)。

[0027] 本发明的有利效果

[0028] 本发明具有以下优点:

[0029] 本发明的方法的优点在于其使 JPEG2000 压缩中的过度编码最小化以便实现改善的计算效率。

[0030] 本发明的方法的另一优点在于其可以提供通过阈值观看距离来量化的已知水平的感知图像质量。

[0031] 本发明的方法的又一个优点在于可以用对于指定的观看条件提供视觉无损图像质量所需的最小压缩数据量来获得恒定的感知质量。

[0032] 本发明的方法的再一个优点在于其可以高效地实现图像帧序列的期望平均压缩数据速率。

[0033] 在阅读结合附图而进行的以下详细说明时,本发明的这些及其它特征和优点将变得对于本领域的技术人员来说显而易见,在附图中示出并描述了本发明的说明性实施例。

附图说明

[0034] 虽然本说明书总结出特别指出并明确主张本发明的主题的权利要求书,但请相信,通过结合附图而进行的以下说明,将更好地理解本发明,在附图中:

[0035] 图 1 是举例说明现有技术 JPEG20000 压缩系统的基本组件的框图;

[0036] 图 2 是举例说明误差从 JPEG2000 编码器到显示图像的传递的框图;

[0037] 图 3 是用于人类视觉系统的对比度灵敏度函数 (contrast sensitivityfunction) 的示例;

[0038] 图 4 是以画面高度为单位的观看距离的例示;

[0039] 图 5 是观看距离对对比度灵敏度函数的影响的例示;

[0040] 图 6 是将对比度灵敏度函数值映射到各个小波子带的示例;

[0041] 图 7 是由作为到显示器的输入的一个代码值的增量 (delta) 根据输入电平 (input level) 产生的对比度 (contrast) 增量的示例;

[0042] 图 8 是观看者响应和编码器/解码器系统及显示器的表征的例示;

[0043] 图 9 是举例说明使用观看条件参数和编码器/解码器系统及显示器的表征来生成量化器步长的本发明的框图;

[0044] 图 10 是本发明的量化器步长发生器的框图;

[0045] 图 11 是举例说明使用观看条件参数来生成失真加权因数 (distortionweighting factor) 的本发明的框图;

[0046] 图 12 是举例说明将图像序列压缩至期望的平均数据速率的本发明的框图;

[0047] 图 13 是不同类型的图像序列内容的速率失真函数的示例的例示;

[0048] 图 14 是对于 1.5 屏幕高度的阈值观看距离来说平均压缩数据速率与速率失真函数的斜率 (slope) 之间的关系示例;

[0049] 图 15 是在 1.5 屏幕高度的阈值观看距离的情况下使用第一遍压缩的平均压缩数据速率来估计速率失真函数的斜率的例示;

[0050] 图 16 是使用估计的速率失真函数来修改用于第二遍压缩的阈值观看距离的例示;

[0051] 图 17 是使用所计算的速率失真函数来修改用于第三遍压缩的阈值观看距离的例示;

[0052] 图 18 是用最终阈值观看距离在第三遍压缩之后实现期望的平均压缩数据速率的例示;以及

[0053] 图 19 是使用所计算的速率失真函数来修改用于第四遍压缩的阈值观看距离的例示。

具体实施方式

[0054] 本说明特别地针对构成依照本发明的装置的一部分或更直接地与其协作的元件。应理解,未具体示出或描述的元件可以采取本领域的技术人员所熟知的各种形式。

[0055] 如先前所讨论的, JPEG2000 压缩系统的计算复杂性可能由于在第一层编码中发生的算术编码而相当高。在典型的 JPEG2000 压缩系统中,将量化器步长选择为产生细量化系数,细量化系数生成将被编码的大量数据。然而,在第二层编码中形成压缩代码流期间,许多经算术编码的数据随后被丢弃以实现期望的压缩代码流大小。这是过度编码的问题,其中将所有量化系数数据编码,但并不是将其全部包括在最终代码流中。

[0056] 本发明通过在第一层编码中只生成有限的量化系数数据量作为到算术编码器中的输入来解决过度编码的问题。这样,存在较少的必须由算术编码器来处理的数据,并且随后,还存在较少的必须在第二层编码期间在组织最终代码流时评估的数据。然而,在确定生成什么量化系数数据时必须非常小心,因为如果量化过于粗糙,完全有可能在解压缩图像中引入显著的伪影。为了防止伪影的引入,本发明使用人类视觉系统的模型连同预期的观看条件一起来确定用于小波频率子带的量化器步长。这些量化器步长在限制所产生的过度编码数据量、同时仍保持期望水平的感知图像质量方面是最佳的。观看条件包括诸如观看距离(即从观看者到显示器的距离)、显示设备的大小和明亮度的参数。

[0057] 将人类视觉系统模型应用于图像压缩在现有技术中众所周知。Daly 等人在 U. S. 4, 780, 761 中讲授了一种方法,其根据人类视觉系统对空间频率的灵敏度的二维模型来量化变换系数。在这种方法中,通过离散余弦变换(DCT)来产生变换系数。在包括 Joshi 等人的 U. S. 6, 650, 782 和 U. S. 6, 668, 090 二者在内的现有技术中也已描述了在 JPEG2000 的情况下将人类视觉系统模型用于量化小波系数。然而,这些现有技术方法中没有一种解决了 JPEG2000 中的过度编码问题或提供满足对最大压缩数据量的约束的有效手段。在描述本发明如何解决这些问题之前,描述人类视觉系统模型和如何可以将其应用于图像压缩是有用的。

[0058] 参照图 2,举例说明了误差从 JPEG2000 压缩到显示图像的传递。使用具有用于各种小波频率子带的量化器步长的 JPEG2000 编码器 20 来压缩原始图像数据。量化在已量化小波系数中引入相对于未量化小波系数的误差。将已量化小波系数编码并最后传送到 JPEG2000 解码器 22, JPEG2000 解码器 22 根据已解码的量化系数来重构图像。已量化系数中的误差导致解压缩图像代码值相对于原始图像代码值的误差。当解压缩图像被发送到显示设备 24 时,人类观看者可以观看到显示图像。该显示图像由于解压缩图像代码值的误差而包含对比度误差。对比度在人类心理物理学研究领域中具有特殊意义,且稍后不久将定义该术语。给定如刚刚描述的那样的组合的 JPEG2000 编码器/解码器和显示系统,目的是确定将在显示器处造成人类观看者不可检测的对比度误差的量化器步长。

[0059] 对比度是信号的亮度变化相对于其平均亮度的度量。根据正在观看的信号的具体特性,可以以各种方式来定义对比度。一种很常见的定义是:

$$[0060] \quad \text{Contrast} = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}} = \frac{\Delta L}{2 \cdot L_{\text{mean}}}, \quad (1)$$

[0061] 其中, $L_{\max}$ 是信号的最大亮度, $L_{\min}$ 是信号的最小亮度, ΔL 是最小与最大亮度之间

的差,且 L_{mean} 是平均亮度。对比度的这种定义称为 Michelson 对比度,且其特别适合于亮度以正弦方式变化的信号。虽然由小波压缩技术产生的伪影不是简单的正弦曲线,但我们已发现 Michelson 对比度在实施本发明方面起到相当好的作用。

[0062] 对于具有空间频率 f 的正弦激励 (stimulus), 可以通过心理物理学实验来确定观看者检测到该激励的存在所需的最小对比度。此最小可检测对比度称为该频率的对比度阈值,我们将其表示为 $C_t(f)$ 。对比度灵敏度被简单地定义为 $1/(\text{对比度阈值})$, 其允许将对比度灵敏度函数 (CSF) 定义为:

$$[0063] \quad \text{CSF} = 1/C_t(f) \quad \forall f > 0. \quad (2)$$

[0064] 图 3 举例说明典型的对比度灵敏度函数,其中以视差 (visualsubtense) 的周/度 (cycles/degree) 为单位给出空间频率。如本领域的技术人员所熟知的那样,对比度灵敏度函数的形状随着诸如平均亮度水平、显示器的噪声水平、取向等的许多观看条件参数而变。

[0065] 在应用 CSF 时尤其重要的一个观看条件参数是观看距离。一旦指定了预期的观看距离和显示器特性,则可以将 CSF 的周/度单位映射成显示器上的物理单位,诸如 Jones 等人的 US 6,349,151 中概述的周/毫米 (cycles/mm) 等。然而,周/毫米的单位是特定显示器的函数,且已发现在本发明的实施例中便于将 CSF 映射到周/画面高度的单位。这样,将显示像素的大小的测量单位与观看距离相组合,且只需要指定从显示器到观看者的就画面高度数目而言的观看距离以建立 CSF 空间频率轴的适当缩放比例 (scaling)。在图 4 中针对两个画面高度的观看距离举例说明了就画面高度而言的观看距离的这种概念。在图 4 中将显示图像的总对角 (subtended angle) 表示为 θ 。可以通过以下简单计算将周/度的单位转换成周/画面高度:

$$[0066] \quad \text{周/画面高度} = (\text{周/度}) (\theta \text{ 度/画面高度}) \quad (3)$$

[0067] 例如,在两个画面高度处观看的图像将具有总对角 $\theta = \tan^{-1}(1/2) = 26.6^\circ$ 。可以对任何期望的观看距离进行类似计算,例如在一个画面高度的观看距离处的显示器总对角为 $\tan^{-1}(1/1) = 45^\circ$ 。

[0068] 在图 5 中针对 0.5、1.0、和 2.0 画面高度 (PH) 的观看距离举例说明了 CSF 到周/画面高度的映射和画面高度形式的观看距离的所得效果。在观看者远离显示器移动时,CSF 被映射到较低的空间频率,且存在对较高空间频率的较低灵敏度。对于电影应用,常常使用术语屏幕高度来代替画面高度,其中假设图像被投影为其高度近似充满电影屏幕,所以术语画面高度和屏幕高度基本上是等效的,且对于本发明,在本文中将其视为等效。可替换地,有时以与显示像素大小有关的尺度指定观看距离。例如,如果图像具有 1000 像素的高度尺度,则还可以将两个画面高度的观看距离描述为 2000 像素,其中像素与显示在显示器上的像素具有相同的尺度。用来描述观看距离的画面高度、屏幕高度和显示器像素的单位可容易地互换,但是根据观看应用的领域相比于其它单位而选择一个单位可能是有利的。

[0069] 给定观看距离和诸如显示器亮度的其它观看条件参数,可以根据诸如 Daly 等人 US 4,780,761 中描述的 CSF 模型的 CSF 模型来计算适当的 CSF 曲线。随后,将 CSF 值 (基本上是连续函数) 映射到离散小波子带。然而,每个小波子带包含一定范围的空间频率,因此必须选择单个频率来代表整个子带。图 6 举例说明使用每个子带的中心频率的映射;本示例中的 CSF 是用于 2.0 屏幕高度的观看距离。其它映射也是可能的,例如使用比中心频率低的频率,以便在定义对每个子带中出现的误差的灵敏度时更稳妥。

[0070] 重要的是注意在将 CSF 应用于本发明时,使用具有大于指定观看距离的观看距离的所有 CSF 的包络 (envelope)。在图 5 中针对 0.5PH 的观看距离举例说明了低频 CSF 包络的示例。此包络的使用具有消除 CSF 的低频衰减 (roll-off) 的效果,以便已针对预期的观看距离 D 而被压缩的图像对于大于 D 的所有观看距离仍将是视觉无损的。所述预期的观看距离 D 也称为阈值观看距离,因为任何更近的观看距离可以显示出经解压缩的显示图像中存在的对比度误差。使用阈值观看距离及其它相关观看条件导出的量化器步长称为阈值量化器步长。阈值观看距离是指定图像质量的精确方式,因为其提供有意义的物理量,该物理量指示在什么样的观看条件下压缩图像与原始未压缩图像相比将是视觉无损的。这与严格依赖于 MSE 或 PSNR 度量的技术形成对照,该技术几乎不提供关于图像对特定显示条件的适合性的信息。

[0071] 一旦知道了对应于小波子带的 CSF 值,则可以计算相应的量化器步长组,其将产生针对指定观看条件的视觉无损质量。为了执行此计算,存在必须表征的两种其它关系。

[0072] 首先是就当对显示器进行输入时由代码值变化产生的显示图像的对比度变化而言显示器的表征 (characterization)。具体地说,需要知道一个代码值 (CV) 的输入增量的显示对比度增量。如所证明的那样,此关系是一个 CV 增量发生时的信号电平 (signal level) 的函数,且在图 7 中针对以 2.6 的伽玛值 (gamma) 显示的 12 位数据举例说明了此关系。可以看到,在低 CV 处 (阴影区) 产生最大对比度变化,但观看者更可能适合于较高的亮度水平。因此,可以选择中尺度 (mid-scale) 输入电平来建立对比度 /CV 关系。

[0073] 其次是就对于给定量化器步长在解压缩图像中产生的 CV 变化而言 JPEG2000 编码器 / 解码器系统的表征。具体地说,需要知道将在解压缩图像中产生一个 CV 增量的用于每个小波子带 b 的量化器步长 $Q(b)$ 。可以由直接数值计算或通过简单的实验来确定这些值。将产生一个 CV 增量的特定量化器步长将取决于 JPEG2000 实施方式以及小波滤波器 (例如在服从 DCI 的系统的情况下为 9-7 滤波器)。

[0074] 在确定这两种表征之后,可以将其与 CSF 值组合以确定实际阈值量化器步长以便进行视觉无损压缩。参照图 8,可以看到,由在解压缩图像中引起一个 CV 增量 ($Q(b)/CV$) 的用于每个小波子带 b 的量化器步长来表征 JPEG2000 编码器 / 解码器系统;由一个 CV 输入增量 (对比度 /CV) 的对比度增量来表征显示器,且由每个子带 b 的阈值对比度 ($C_t(b)$) 来表征观看者,可以通过对等式 2 求逆 (invert) 而由 CSF 来确定该阈值对比度 ($C_t(b)$)。给定这些关系,用于给定小波子带 b 的阈值量化器步长 $Q_t(b)$ 为:

$$[0075] \quad Q_t(b) = \frac{[Q(b)/CV] \cdot [C_t(b)]}{[Contrast/CV]} \quad (4)$$

[0076] 如图 9 所示,除图 1 所描绘的 JPEG2000 编码器元件之外,本发明还包括量化器步长发生器 30。量化器步长发生器 30 生成将被应用于小波系数量化器 14 中的频率子带的量化器步长。输入量化器步长发生器中的是观看条件参数 (其包括以画面高度、屏幕高度、及其它类似物为单位的观看距离)、一个 CV 输入增量的对比度增量的显示器表征、以及在压缩图像中产生一个 CV 增量的量化器步长的 JPEG2000 编码器 / 解码器系统表征。

[0077] 参照图 10,如图 9 所示的量化器步长发生器 30 包括:CSF 计算器 30a,其将观看条件参数用作输入;小波子带空间频率映射器 (mapper) 30b,其使用 CSF 值来确定每个小波子带的对比度阈值 $C_t(b)$;以及量化器步长计算器 30c,其使用显示器表征 (对比度 /CV)、

JPEG2000 编码器 / 解码器表征 ($Q(b)/CV$)、以及对比度阈值 ($C_t(b)$) 利用在等式 4 中给出的公式来计算用于小波子带的量化器步长。

[0078] 本发明具有以独立于显示器的特定大小的画面高度、屏幕高度或类似物为单位来指定观看距离的优点。在诸如使用数字电影系统的电影放映的应用中,可以用本发明将电影压缩至阈值观看距离所指定的质量水平,且客户确信无论显示电影的屏幕大小如何,图像质量将是相同的。另一个优点是可以在不进行任何修改的情况下用大多数 JPEG2000 编码器来实施本发明,因为大多数编码器被设计为接受用户指定的量化器步长作为输入值。

[0079] 在本发明的一个实施例中,指定预期的观看条件并根据 CSF 模型和如前所述的其它系统表征来确定相应的阈值量化器步长。这些步长用来压缩图像,以便在预期的显示条件下观看经解压缩的显示图像和原始未压缩图像这两种图像时,经解压缩的显示图像将具有与原始未压缩图像相同的感知质量,即对于预期的显示条件来说压缩是视觉无损的。在这种方法中,过度编码被明显最小化,因为阈值量化器步长将产生期望的质量水平所需的最小信息量,因此第一层编码将仅对此最小信息量进行编码。这种压缩系统内置地是 VBR 系统,因为最终代码流中的压缩数据量将取决于图像内容。

[0080] 然而,即使用这种 VBR 系统,仍然可能存在对每个图像的最大压缩数据量有约束的情况,如同对于图像序列的 JPEG2000 压缩的服从 DCI 的代码流的情况一样。为了满足此类约束,本发明在第二层编码期间包括诸如 PCRD-opt 的附加速率控制算法以确保满足所有大小约束。可以将压缩数据量与预指定的最大量相比较,且如果超过该最大量,则第二层编码期间的速率控制可以降低压缩数据量。对于本发明所述的视觉无损 VBR 系统,这些最大大小约束通常大于实际压缩大小,因为对于大部分图像内容来说基于视觉的量化将产生小的压缩大小。因此,第二层编码期间的对附加压缩数据的剪裁通常不会发生,且即使在最坏的情况下该剪裁也是最小的,因此过度编码是最小的。

[0081] 在由于代码流大小约束而发生压缩数据剪裁的情况下,不再保证所得的图像对于预期的观看条件来说是视觉无损的。然而,作为量化图像质量上的可能损失的一种方式,可以使用 US 6,650,782 中所描述的方法对于此类经修改的图像计算新的阈值观看距离。然而,对于使用 DCI 指定大小约束的 DCinema 应用,极少针对典型观看条件而裁剪压缩数据,此外,即使当发生裁剪时,裁剪的量也很小。因此,实际上,这种 VBR 方法提供几乎恒定的感知质量,对于大多数压缩的图像,该感知质量对于指定阈值观看距离(和所有更大的观看距离)来说是视觉无损的。

[0082] 虽然本发明的上述实施例通过降低计算复杂性而克服了现有技术的限制,但这些实施例仍可以在第二层编码期间利用 PCRD-opt 或类似的速率控制算法来保证满足对最大压缩代码流大小的约束。这是满足大小约束的有利方式,因为大多数 JPEG2000 编码器包括某种形式的 PCRD-opt,因此,可以在不进行进一步修改的情况下用此类编码器来实施本发明。

[0083] 如前所述,PCRD-opt 使用 MSE 作为失真度量,而 MSE 可能并不总是与感知质量很好地相关。然而,PCRD-opt 和 MSE 在本发明的情况下在剪裁编码数据中的使用与其在产生细量化小波系数的典型 JPEG2000 编码器中相比很少成为问题。在这些现有技术方法中,每次编码判定都是基于丢弃压缩数据时的 MSE(及相应的速率变化),且不能保证此类编码判定将产生对于最终压缩代码流的大小来说最佳的感知图像质量。相反,本发明在量化阶段

期间进行大部分编码判定,结果,被传递至第一层编码并随后到达第二层编码的数据已被限制为具有非常强的感知相关性的有限数据量。

[0084] 在用本发明的 VBR 编码的情况下,第二层编码中的剪裁量对于大多数图像来说是无关紧要的,因此所得的压缩图像将提供几乎恒定的感知质量。为了保证在第二层编码期间进行的任何编码判定就图像质量而言是最佳的,还可以在第二层编码期间采用视觉模型。这可以通过在计算 PCRD-opt 算法或类似的方法的失真时使用基于视觉的失真加权因数来实现。可以将失真加权因数分配给每个小波频率子带,然后使用该加权因数来确定由给定子带中的特定系数贡献的 MSE 的降低。图 11 中举例说明了此过程,其中观看条件是失真加权因数发生器 32 的输入参数,失真加权因数发生器 32 生成将在第二层编码 18 中被 PCRD-opt 或类似的算法使用的基于视觉的加权因数。用来确定加权因数的人类视觉系统模型通常基于前文所述及图 5 和 6 所示的 CSF 模型。然而,应注意的是用于每个频率子带的量化器步长和加权因数不是相同的值,步长和加权因数也不一定跨越所有子带而共享相同的关系。这是因为 MSE 的量化和视觉加权过程基本上是不同的,因此对于两个过程来说各子带的关系是不同的。

[0085] 将基于视觉的加权因数用于 MSE 失真计算在现有技术中众所周知,但将阈值观看距离用于确定本发明中实施的基于视觉的小波系数量化和基于视觉加权 MSE 的后续编码判定两者提供相比于现有技术的优点。如前文所讨论的,通过使用对于预期的观看条件的视觉模式来限制过度编码量,所得的压缩代码流对于那些观看条件来说就感知质量和大小而言几乎是最佳的。如果需要剪裁压缩数据以满足对最大压缩数据量的限制,则 MSE 的视觉加权可以帮助保证丢弃的数据将对由较小压缩代码流所产生的感知图像质量具有最小的影响。

[0086] 如图 11 所示,可以使用被用于确定量化器步长的相同观看条件作为到失真加权因数发生器 32 的输入。使用由相同观看条件导出的基于视觉的加权因数的效果是对应于较高空间频率的失真通常被加权为较低值,这意味着在 PCRD-opt 的速率失真优化期间将优先丢弃用于较高空间频率的压缩数据。这产生可能具有比期望锐度略低的锐度的图像,但其使引入令人讨厌的伪影的可能性最低。与不自然的编码伪影相比,观看者更能容忍图像锐度的轻微损失。

[0087] 然而,如图 11 中进一步举例说明的那样,使用一个或多个可选观看条件参数作为到失真加权因数发生器 32 的输入以产生用于失真计算的加权因数也可能是有利的。这种方法的示例是使用比用来生成量化器步长的阈值观看距离大的观看距离来确定加权因数。用这种方法,在满足压缩大小约束时,对应于较高空间频率的压缩数据比用于较低空间频率的数据更有可能被丢弃。这对于已将系数量化至阈值水平的 VBR 编码来说特别重要。任何附加的压缩误差都将在具有刚刚描述的视觉加权的较高频率处发生,这等效于根据比原始指定的阈值观看距离更大的阈值观看距离来进行量化。同样,这种数据缩减(reduction)丢弃感知相关性最低的压缩数据。

[0088] 将非常大的加权因数应用于较低空间频率子带以完全保证不从这些子带中丢弃数据可能是更有利的。当已将低频子带量化至视觉无损水平时,优选的是不将任何附加失真引入这些频率子带中,因为所得的图像质量可能受到严重影响。然而,当计算基于视觉的量化器步长和加权因数时无论使用相同还是不同的观看条件,视觉加权因数的效果都将保

证在剪裁压缩数据以满足最大代码流大小约束时丢弃感知相关性最低的数据。

[0089] 在前述讨论中,尚未提及通常组成图像的颜色分量。在本发明的另一实施例中,对于每种颜色分量确定单独的量化器步长组,其中针对总体压缩/解压缩/显示链(chain)调节每个组。例如,被压缩为RGB色彩通道的图像的三组最佳量化器步长不同于已用ICT进行预处理以便在压缩之前将RGB数据转换成YCbCr数据的图像的最佳量化器步长组。这是因为在将压缩数据解压缩、必要时用逆(inverse)ICT进行处理、并转换成显示器上的彩色信号时,对于每种色彩通道被引入已量化小波系数中的误差以各种方式被放大或减小。对于表示对应于不同原色(colorprimary)组(诸如XYZ而不是RGB)的图像颜色分量的数据和具有不同原色组(primary set)(例如RGB P3或RGB 709原色)的显示器,在不同色彩通道中发生误差可感知性方面的类似变化。同样地,期望基于不同的色彩通道的特性来调整用于任何MSE计算的视觉加权因数,因为某些色彩通道在感知上更重要,例如绿色色彩通道,而其它的具有较低感知相关性,例如蓝色色彩通道。给定特定图像色彩表示和显示器色彩特性的知识,可通过直接数值计算或简单的观看者实验来确定不同色彩通道的最佳量化器步长和加权因数。

[0090] 除对单个帧或帧分量的压缩数据量的约束之外,还可能存在对用来表示组成图像序列的所有图像帧的压缩数据总量的约束。如前所述,对图像序列的压缩数据总量的约束允许确信给定持续时间的电影将安装到具有一定容量的存储介质上,例如两小时的电影被压缩到160千兆字节的硬盘驱动器上。对总代码流大小的约束在将通过通信网络来发送压缩数据的系统中也是有利的,诸如压缩电影内容到电影院的卫星传递。

[0091] 对压缩代码流的总大小的约束等效于对整个序列上的平均压缩数据速率的约束,为方便起见,以下讨论将使用平均数据速率来代替总文件大小。例如,如果一个人具有长度为2小时的电影且希望将压缩电影安装到160千兆字节的硬盘驱动器上,则平均压缩数据速率必须小于182Mbps。由于可变位率编码,所以某些帧将使用更高的瞬时压缩数据速率(在服从DCI的代码流的情况下达到最大250Mbps),而其它帧将使用基本小于182Mbps的瞬时数据速率。尽管瞬时压缩数据速率改变,但目的是对于受到平均压缩数据速率约束的每个帧都实现恒定的感知图像质量,且以高效的方式这样做以防止工作流程效率方面的任何损失。

[0092] 在本发明的实施例中同样通过使用阈值观看距离的图像质量度量来达到这一目的。参照图12,用户提供期望的平均压缩数据速率,以及期望的阈值观看距离作为初始质量规格。除前述阈值观看距离之外,用户还可以提供其它观看条件参数。所述期望阈值观看距离被提供为前文已描述并在图10中示出的量化器步长发生器30的输入,以计算初始压缩参数组,即量化器步长。用户还提供期望的平均压缩数据速率和期望的阈值观看距离作为稍后不久将描述的修改阈值观看距离发生器44的输入。另外,期望的平均压缩数据速率也被提供作为同样在稍后不久将描述的比较逻辑42的输入。

[0093] 再次参照图12,组成原始图像序列的图像帧组被分成两个子集。被表示为子集#1的一个子集只包含少量的帧,通常是帧总数的1~2%,而被表示为子集#2的另一子集包括不在小子集中的全部其余帧。可以通过对图像序列的随机时间取样来选择小子集帧(子集#1),但是已发现固定时间采样间隔在实践中起到很好的作用且比随机采样间隔更容易实现。还可以使用其它采样策略,诸如根据特定场景内容来选择帧,如果此类信息可用的

话。

[0094] 利用 JPEG20000 编码器 20, 使用初始压缩参数组来整体地压缩子集 #1 中的每个帧以产生压缩帧的子集。使用平均数据速率估计器 (estimator) 40 来计算子集 #1 的压缩帧的平均压缩数据速率, 其中, 所述平均压缩数据速率仅仅是与子集 #1 中的各个帧相关的压缩数据速率的平均值。使用子集 #1 的平均压缩数据速率作为整个序列的平均压缩数据速率的估计。这种稀疏时间采样方法适用于运动画面数据, 因为帧间相关性通常很高, 且贯穿图像序列的稀疏时间采样将提供对整个序列的平均数据速率的非常好的估计。然而, 应注意的是形成图像帧的子集以估计平均压缩数据速率的这种方法还可以应用于任何图像组, 不仅仅是运动图像序列, 只要对于小子集中的图像和该组中的其余图像存在图像内容的足够相似性即可。

[0095] 使用比较逻辑 42 来将估计的平均压缩数据速率与期望的平均压缩数据速率相比较。如果估计的平均数据速率小于或等于期望的平均数据速率, 则可以利用 JPEG2000 编码器 20 使用与用于子集 #1 的压缩参数相同的压缩参数来压缩其余帧的子集 (子集 #2), 以产生一组已压缩的其余帧。压缩帧 (来自子集 #1) 的小子集和已压缩的其余帧 (来自子集 #2) 的子集一起组成整个压缩帧组。请注意, 如果在该第一遍压缩时满足平均速率约束, 则不需要重新压缩子集 #1, 因此如果对于子集 #1 而言实现期望的平均压缩数据只需要一遍压缩, 则不存在总压缩时间上所引起的开销。

[0096] 如果估计的平均数据速率大于期望的平均数据速率, 则使用修改阈值观看距离发生器 44 来修改阈值观看距离。具体地说, 增加阈值观看距离, 这将通过允许在由量化器步长发生器 30 用较大阈值观看距离来计算新的量化器步长时引入更多误差来缩小压缩文件大小 (并因此而降低平均数据速率)。当在期望的较近观看距离处观看内容时, 较大阈值观看距离的使用将在较高空间频率内引入可检测误差, 但主观测试显示 在 DCinema 环境中的典型观看条件的情况下由于观看距离的变化而引起的质量损失相当小。

[0097] 由速率失真 (R-D) 函数来控制使用修改阈值观看距离发生器 44 对阈值观看距离进行修改以满足给定的平均数据速率。在视觉无损 VBR 系统中, 通过以屏幕高度为单位 (或如前所述的相关测量单位) 的阈值观看距离来量化失真, 并由诸如 Mbps 的适宜单位的平均压缩数据速率来量化所述速率。R-D 函数取决于图像内容, 且在图 13 中针对三种不同类型的 2K 分辨率电影内容示出了 R-D 函数的示例。试验内容 1 是因为其不包含噪声而相当易于压缩的源自计算机图形的内容, 而试验内容 2 和 3 源自胶片, 且更加难以压缩, 因为它们包含胶片颗粒噪声。还可以考虑其它类型的内容, 诸如数字式照相机来源, 其将产生另一种 R-D 函数。除内容的起源之外, 实际场景内容也将影响 R-D 函数。例如, 序列中的帧可以包含大的微细节区域, 这些帧将比具有大的均匀区域的帧更加难以压缩。

[0098] 在图 13 中的 R-D 函数示例中, 无论图像内容的来源如何, 平均压缩数据速率与阈值观看距离之间的关系是高度线性的。通常, R-D 函数不是直线, 但 DCinema 应用中所关心的阈值观看距离和平均数据速率两者都只跨越有限的范围, 且线性模型通常足以模拟该关系。这意味着如果具有给定内容的 R-D 曲线上的任何两个点, 则修改阈值观看距离发生器 44 可以线性地内插或外插以找到将实现期望的平均压缩数据速率的阈值观看距离。

[0099] 然而, 通过帧的小子集 (子集 #1) 的第一遍压缩只提供 R-D 函数上的一个点。在完成通过帧的小子集的第二遍压缩之前, R-D 曲线的线性函数不是已知的。因此在执行第二遍

压缩时出现的问题是当使用量化器步长发生器 30 来计算用于第二遍的新的量化器步长时阈值观看距离应由修改阈值观看距离发生器 44 调整多少？理想的是，子集 #1 的第二遍压缩将实现满足期望的平均速率约束的估计的平均压缩数据速率，以便使得不需要其他的压缩遍数。为了确定将实现期望的平均压缩数据速率的修改阈值观看距离，必须知道 R-D 函数的斜率以表征线性关系，但此斜率在很大程度上取决于前文所述和图 13 所示的内容的类型。可以基于内容和捕捉设备（例如 CGI 动画、高速与低速起源胶片 (origination film) 的关系等）的表征来进行直观推断。然而，更简单且更高效的方法是使用来自帧的小子集（子集 #1）的第一遍压缩的平均压缩数据速率来估计 R-D 曲线的斜率。

[0100] 例如，考虑始终使用对应于 1.5 屏幕高度的阈值观看距离的压缩参数来对子集 #1 执行第一遍压缩的系统。对于不同类型的内容，可以预先确定 1.5 屏幕高度处的平均数据速率与 R-D 曲线的斜率之间的关系。图 14 中举例说明了此预定关系的示例，其将图 13 所示的数据用于三组试验内容。请注意，为方便起见，图 14 中的纵坐标被绘制为 $-(R-D \text{ 函数的斜率})$ 。实际上，将对许多不同组的试验内容计算 R-D 函数和相关斜率以得出比图 14 所示更稳健的估计器，图 14 所示的估计器只基于三个序列。另外，如果不希望对于初始阈值观看距离使质量规格局限于诸如 1.5 屏幕高度的单个值，则可以对于许多初始观看距离预先确定平均压缩数据速率与 R-D 斜率之间的关系。

[0101] 在示例性观看距离为 1.5 屏幕高度的情况下给定来自帧的小子集（子集 #1）的第一遍压缩的平均数据速率，类似于图 14 所示函数的函数可以被修改阈值观看距离发生器 44 用来估计正在被压缩的内容的线性 R-D 函数的斜率。例如，图 15 举例说明使用 1.5 屏幕高度的观看距离的对帧的小子集的第一遍压缩，所述观看距离产生 130Mbps 的平均压缩数据速率。使用如图 14 所示的预定关系，帧的此小子集的 R-D 函数的估计斜率是 -0.0164 。

[0102] 现在，假设期望的平均压缩数据速率是 110Mbps 而不是在使用 1.5 屏幕高度的阈值观看距离的示例中实际实现的 130Mbps。必须将阈值观看距离从 1.5 屏幕高度的初始值修改为较大的阈值观看距离，并随后对于通过图像帧的小子集（子集 #1）的第二遍压缩计算相应的量化器步长。使用 R-D 函数的估计斜率，由修改阈值观看距离发生器 44 使用下式来计算修改阈值观看距离：

[0103] 修改阈值观看距离 = 当前阈值观看距离 + (R-D 曲线的斜率) $\times$ (期望的平均数据速率 - 当前平均数据速率) (5)

[0104] 其中，当前阈值观看距离和当前平均数据速率是与子集 #1 的第一遍压缩相关的值。图 16 举例说明如何确定修改阈值观看距离以便对于图像帧的小子集估计第二遍压缩以产生期望的平均压缩数据速率的示例。使用本示例中的 R-D 函数的估计斜率 ($= -0.0164$)，实现 110Mbps 的期望数据速率的修改阈值观看距离是 $[1.5 + (-0.0164)(110 - 130)] = 1.83$ 屏幕高度。使用此修改阈值观看距离对第二遍压缩计算量化器步长。用修改阈值观看距离所得的来自子集 #1 的第二遍压缩的平均数据速率应非常接近于期望的平均数据速率。如果该平均速率小于或等于期望的平均速率约束，则可以用该修改阈值观看距离来压缩其余帧（子集 #2），并用最小的开销完成整个序列的整个压缩过程。

[0105] 如果在子集 #1 的第二遍压缩后仍然超过期望的平均数据速率，则再次修改阈值观看距离。然而，不是与如图 16 中一样使用估计斜率，而是使用来自头两遍的实际结果来在修改阈值观看距离发生器 44 中计算 R-D 曲线的斜率，即：

[0106]

$$R-D \text{ 曲线的斜率} = \frac{(\text{来自当前遍的阈值观看距离} - \text{来自前一遍的阈值观看距离})}{(\text{来自当前遍的平均数据速率} - \text{来自前一遍的平均数据速率})} \quad (6)$$

[0107] 其中,所述当前遍(current pass)和所述前一遍(previous pass)分别是帧的小子集(子集#1)的第二遍和第一遍。继续前一示例,假设第二遍的实际平均压缩数据速率是115Mbps,而不是期望的110Mbps。使用子集#1的第一遍和第二遍的实际平均压缩速率和相应阈值观看距离,所计算的R-D斜率是 $[(1.83-1.50)/(115-130)] = -0.0220$ 。使用此计算出的斜率,可以再次由修改阈值观看距离发生器44使用等式5来修改阈值观看距离,其中,当前阈值观看距离和当前平均数据速率现在是与第二遍压缩相关的值。本示例中的修改阈值观看距离被确定为 $[1.83+(-0.0220)/(110-115)] = 1.94$ 屏幕高度。由量化器步长发生器30使用修改阈值观看距离来计算新的压缩参数,并对帧的小子集(子集#1)执行第三遍压缩。图17以图形方式举例说明了如何使用所计算的R-D斜率来确定修改阈值观看距离以便对于图像帧的小子集估计第三遍压缩以产生110Mbps的期望的平均压缩数据速率。

[0108] 在大多数情况下,如图18所示,第三遍压缩足以保证满足平均数据速率约束。如果第三遍未满足速率约束,则使用由修改阈值观看距离发生器44使用等式5和6用每次附加的遍数所计算出的更新R-D斜率来对子集#1执行附加的遍数,其中,平均压缩数据速率和阈值观看距离被更新以反应当前遍压缩和前一遍压缩。继续前一示例,假设在1.94屏幕高度处第三遍之后的实际平均压缩数据速率是112Mbps而不是期望的110Mbps。现在所计算出的R-D斜率被更新为 $[(1.94-1.83)/(112-115)] = -0.0367$ 。本示例中用于第四遍的修改阈值观看距离被确定为 $[1.94+(-0.0367)(110-112)] = 2.01$ 屏幕高度。在图19中以图形方式举例说明了用于调整观看距离以使用第四遍压缩实现平均数据速率的过程。可以使用阈值观看距离的相同调整方式来执行附加的压缩遍数,直至实现期望的平均压缩数据速率为止。对于子集#1,一旦实现了期望的平均压缩数据速率,则使用相同的阈值观看距离来压缩其余帧(子集#2)。

[0109] 虽然不太可能,但也可以在子集#1的压缩遍数之一期间满足平均速率约束,但该平均压缩数据速率基本小于期望的平均速率。这可能导致图像质量比期望平均速率下的可能图像质量低,因此目的是实现非常接近于期望速率而不超过它的平均速率。如果发生这种情况,子集#1的另一遍压缩可能是有利的,且可以使用等式5和6来计算修改阈值观看距离,该修改阈值观看距离将实现更接近于期望的平均速率的平均压缩数据速率。

[0110] 如果子集#1仅由帧的1~2%组成(贯穿整个图像序列而被选择,以得到图像内容的良好代表),则根据所需遍数,多遍速率控制的总开销通常低于5~10%。如果必须实现给定的平均压缩文件大小,则这可能是整个工作流程上可接受的损失。

[0111] 使用这种基于视觉的方法来约束总代码流大小的另一个优点是可以将最终阈值观看距离的报告返回给用户。这为用户提供在约束平均数据速率时估计压缩序列的图像质量的手段。例如,如图18的示例所示,期望的阈值观看距离可能是1.5屏幕高度,但实际上期望的平均数据速率是使用1.91屏幕高度的观看距离实现的。通过报告1.91屏幕高度的最终阈值观看距离,用户现在知道平均速率约束对图像质量的影响。如果已经执行适当的观看者研究以建立各种实际观看距离处的阈值观看距离变化与感知质量变化之间的

关系,则也可以将阈值观看距离的这种改变转换成质量的最小可觉察差 (just noticeable difference) (Δ JND) 的增量。

[0112] 因此,提供的是一种用于以比现有编码器更高效的方式控制 JPEG2000 图像的压缩大小的方法,且该方法还保证压缩数据的大小满足对最大压缩数据量的任何约束。在至少一个实施例中,该方法用以非常高效的方式满足对单独帧或帧分量的压缩数据量和对整个图像序列的压缩数据量的约束的能力提供组成图像序列的图像的几乎恒定的感知图像质量。

[0113] 此外,可以以硬件和软件两种方式用大多数现有 JPEG2000 编码器来实施本发明。仅由的要求是编码器:(1) 接受明确的量化器步长作为到量化器的输入,以及(2) 提供用于在超过预定的最大限度的情况下降低压缩数据量的手段,其中此类手段包括但不限于诸如 PCRD-opt 的常见速率控制技术。

[0114] 应理解,本发明的计算机程序产品可以利用众所周知的图像处理算法和过程。因此,本说明书将特别针对构成本发明的方法的一部分或更直接地与本发明的方法协作的那些算法和过程。因此,应理解,本发明的计算机程序产品实施例可以包含对实施方式有用的在本文中未具体示出或描述的算法和过程。此类方法和过程是传统的,且在本领域的普通技术人员所知的范围内。

[0115] 此类算法和系统的其它方面、以及本发明的计算机程序产品所涉及或协作的用于产生或处理图像的硬件和/或软件在本文中未具体示出或描述,且可以选自本领域已知的此类算法、系统、硬件、组件、和元件。

[0116] 用于执行本发明的方法的计算机程序可以存储在计算机可读存储介质中。此介质可以包括例如:诸如磁盘(诸如硬盘驱动器或软盘)或磁带等磁性存储介质;诸如光盘、光带、或机器可读条形码等光学存储介质;诸如随机存取存储器(RAM)或只读存储器(ROM)等固态电子存储设备;或被用来存储计算机程序的任何其它物理设备或介质。用于执行本发明的方法的计算机程序还可以存储在经由因特网或其它存储介质连接到图像处理器的计算机可读存储介质上。本领域的技术人员将很容易认识到还可以在称为专用集成电路(ASIC)或诸如现场可编程门阵列(FPGA)等可编程数字逻辑芯片的硬件或固件中构造此类计算机程序产品的等价物。

[0117] 部件列表

[0118] 10 预处理器

[0119] 12 离散小波变换

[0120] 14 具有恒域的均匀量化器

[0121] 16 自适应二进制算术编码器

[0122] 18 位流组织器

[0123] 20 JPEG2000 编码器

[0124] 22 JPEG2000 解码器

[0125] 24 显示设备

[0126] 30 量化器步长发生器

[0127] 30a CSF 计算器

[0128] 30b 小波子带空间频率映射器

- [0129] 30c 量化器步长计算器
- [0130] 32 失真加权因数发生器
- [0131] 40 平均数据速率估计器
- [0132] 42 平均压缩数据速率比较逻辑
- [0133] 44 修改阈值观看距离发生器

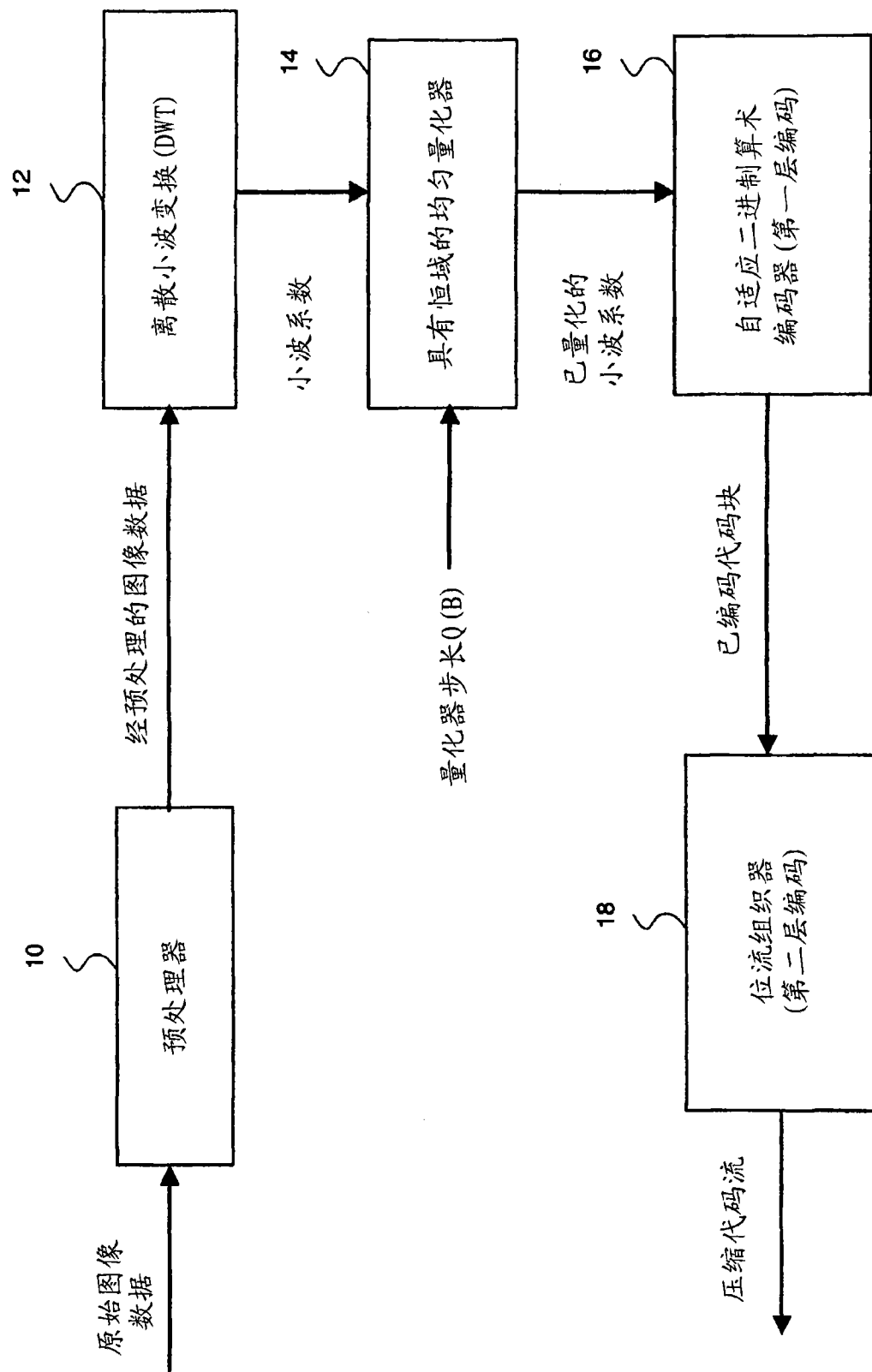


图 1

(现有技术)

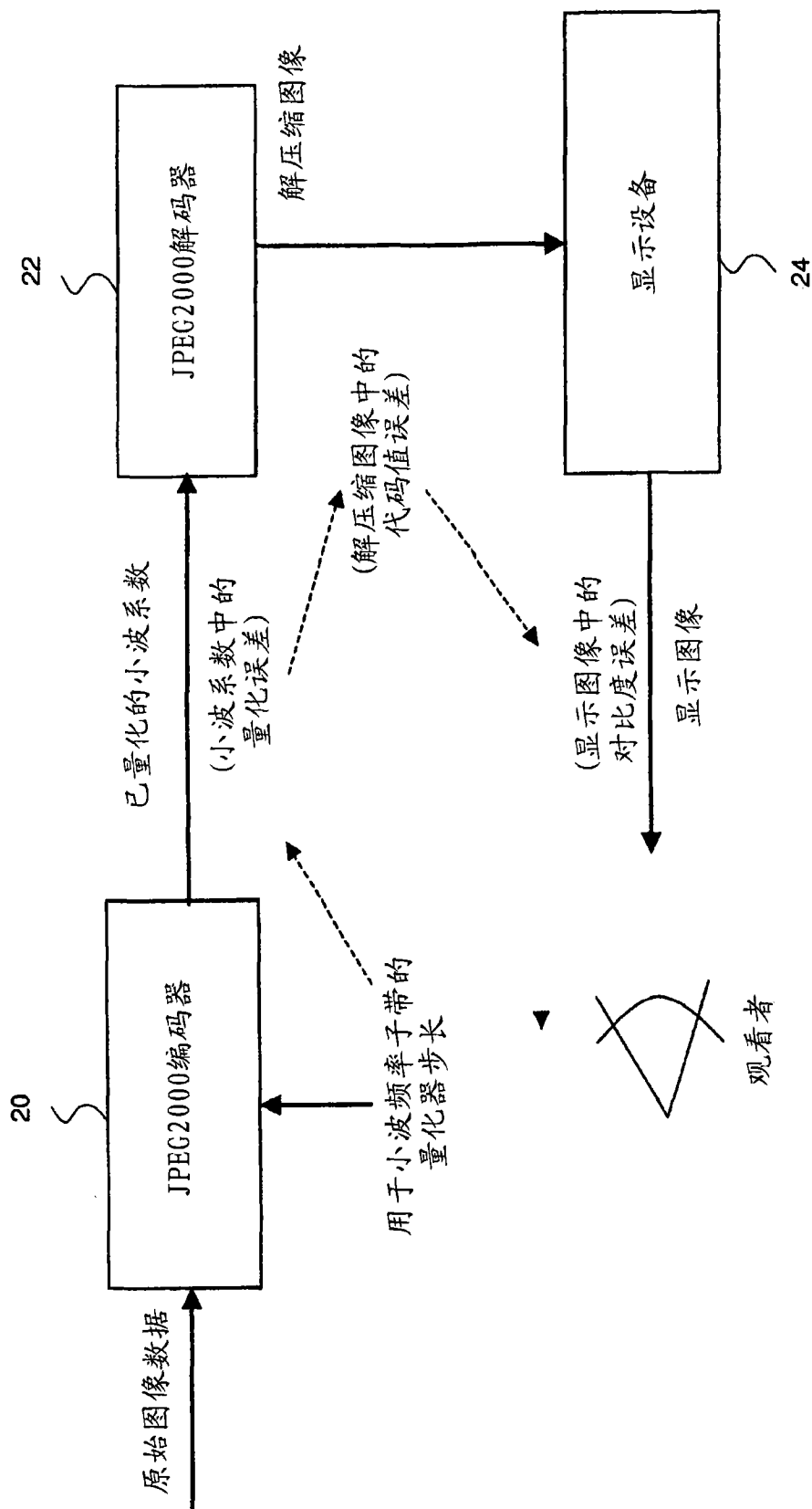


图 2

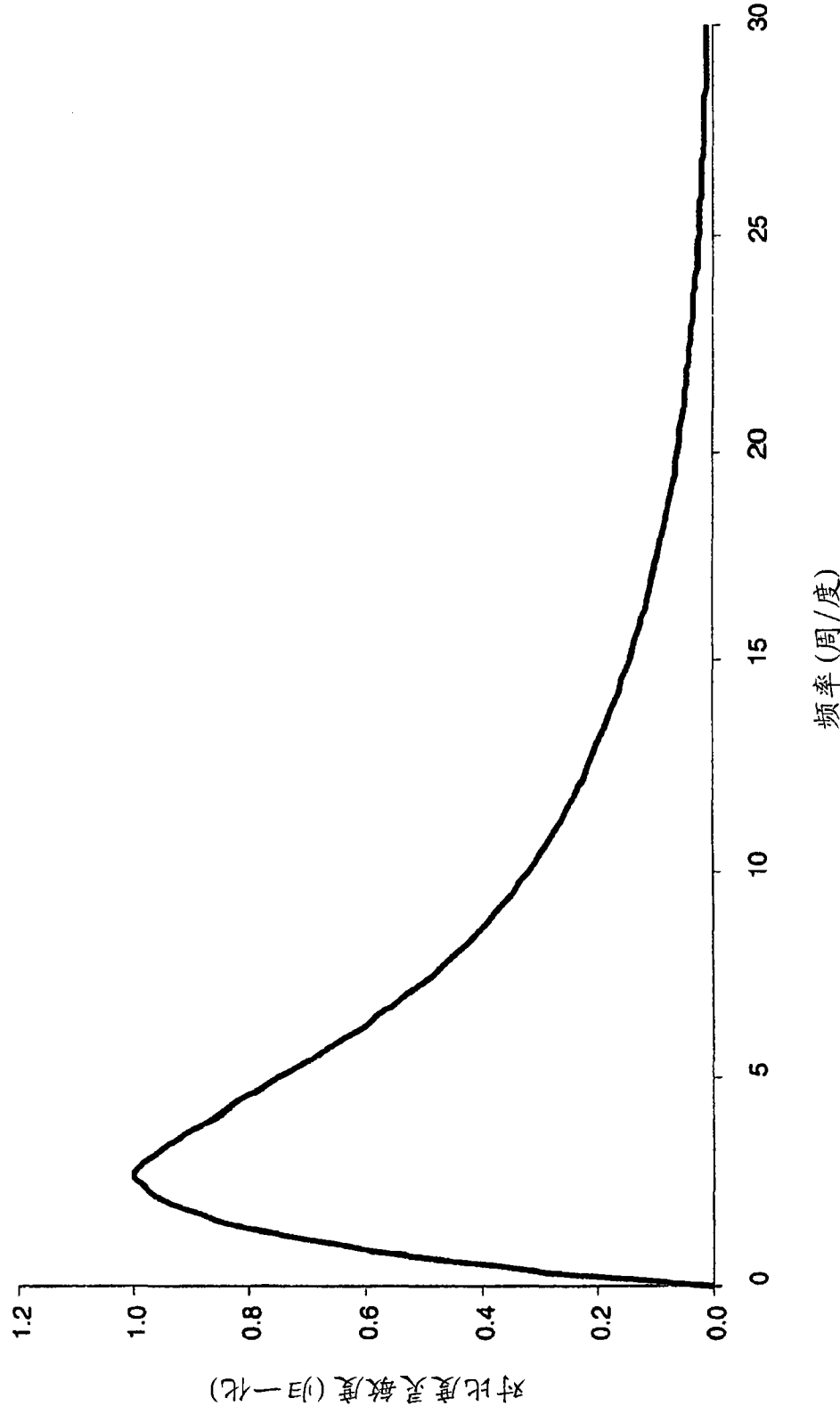


图 3

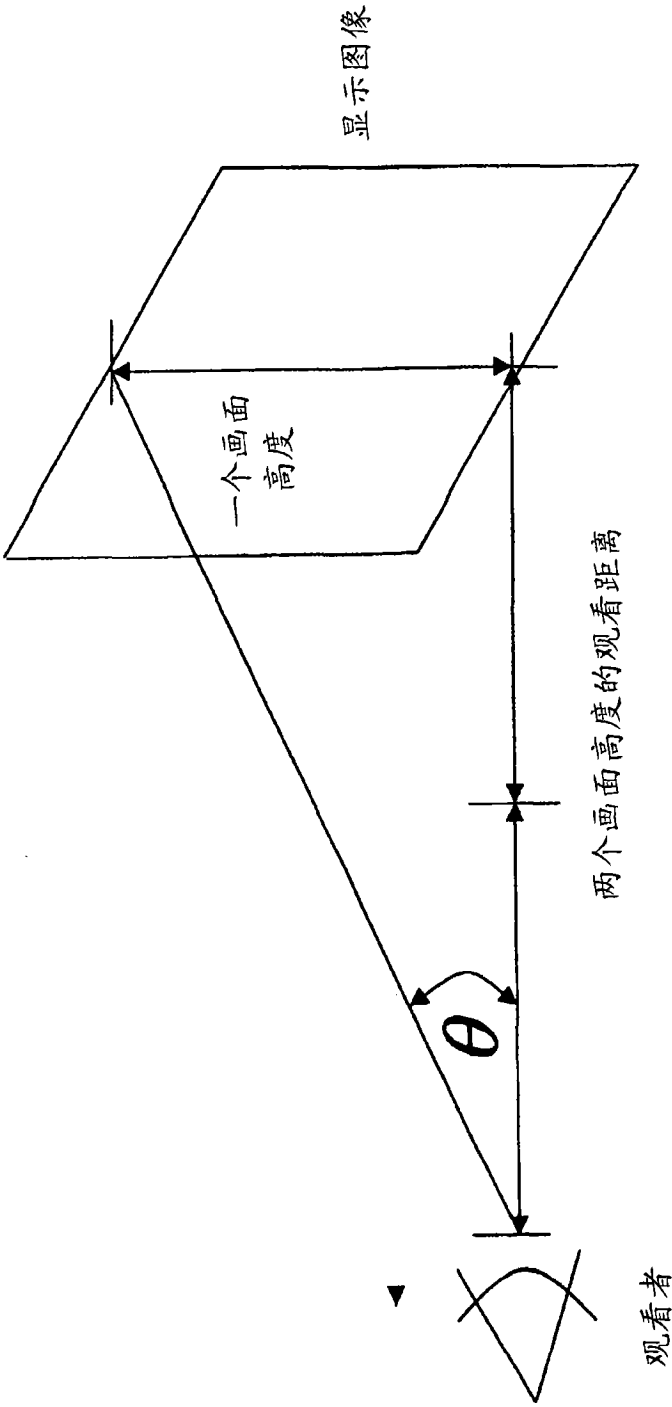


图 4

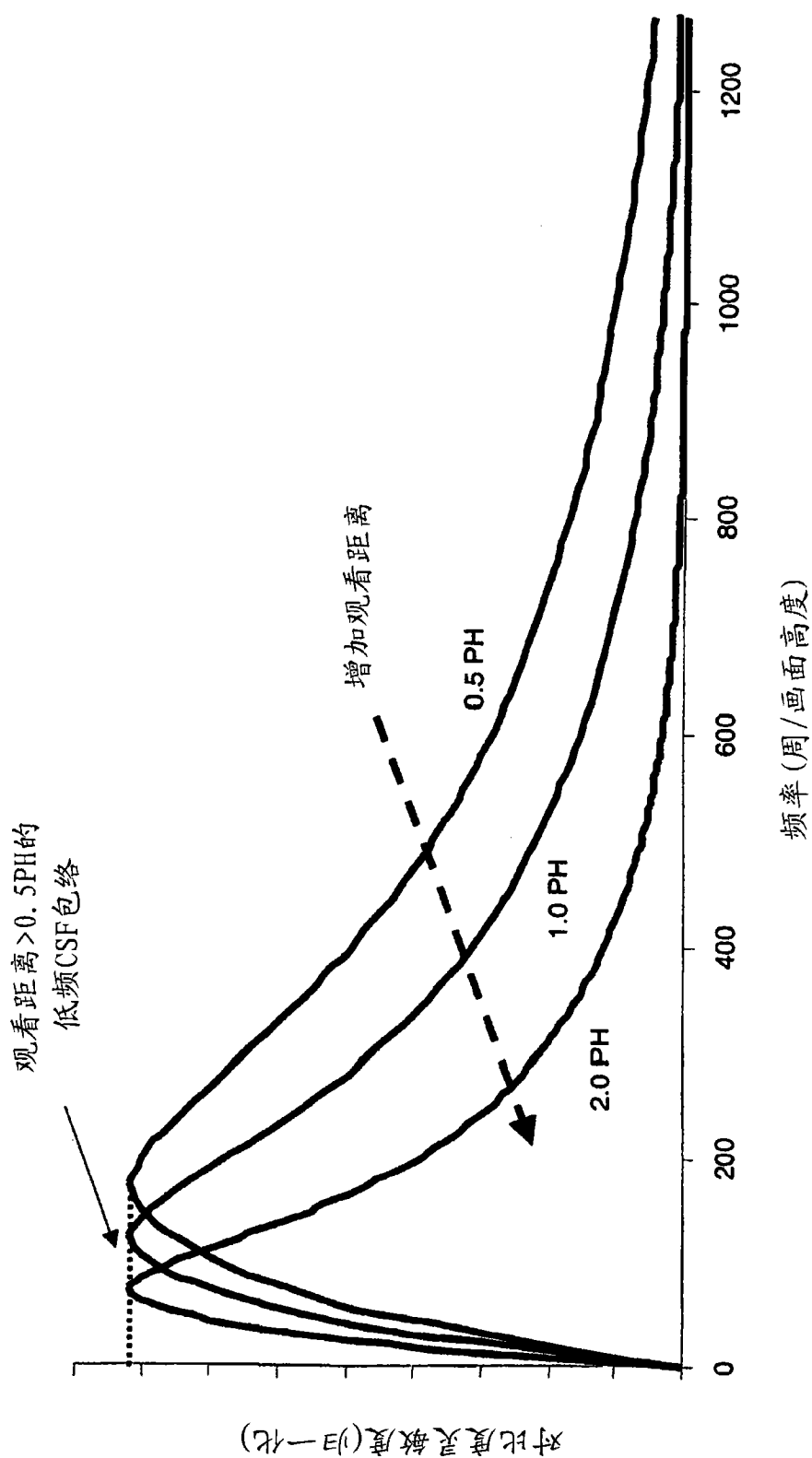


图 5

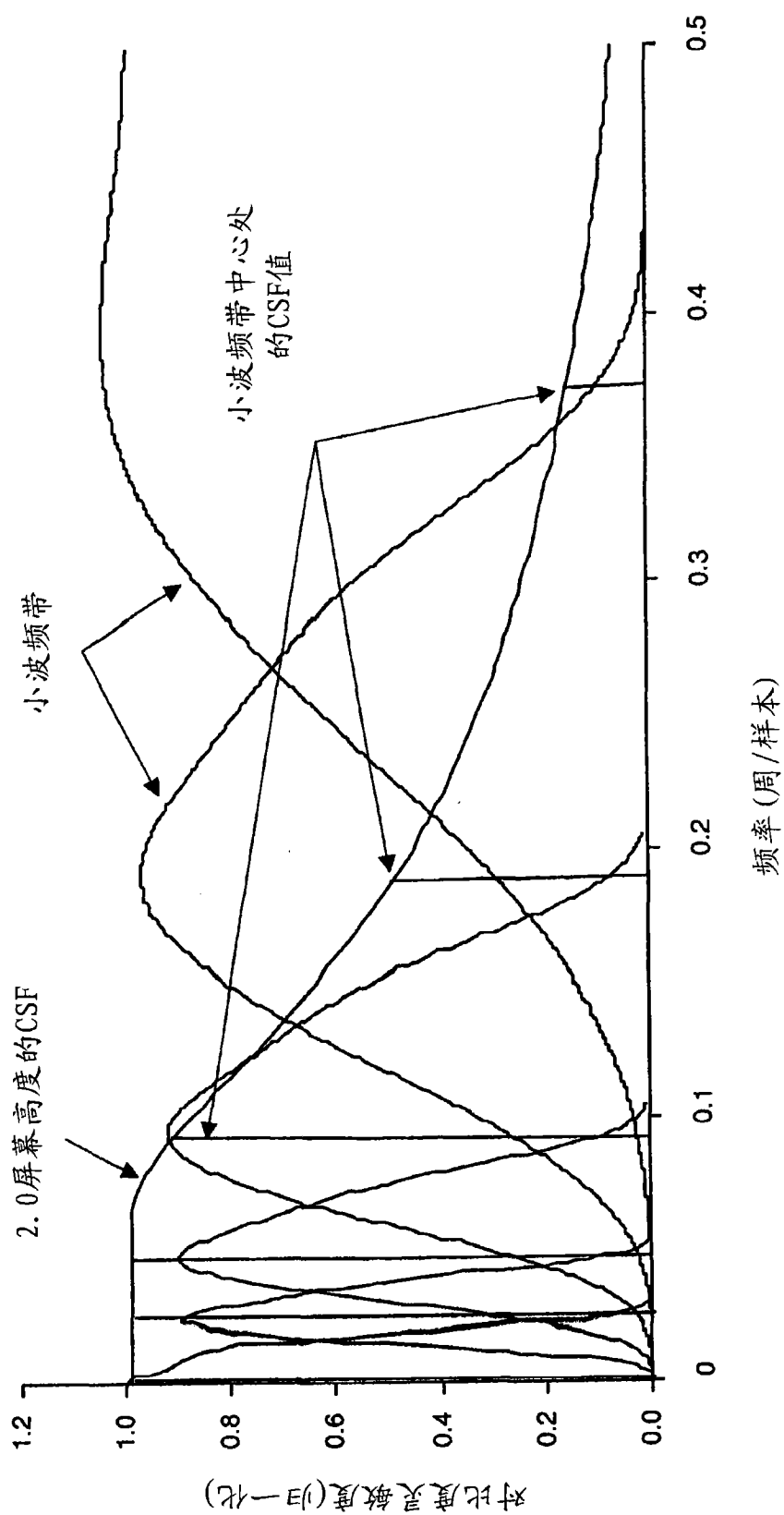


图 6

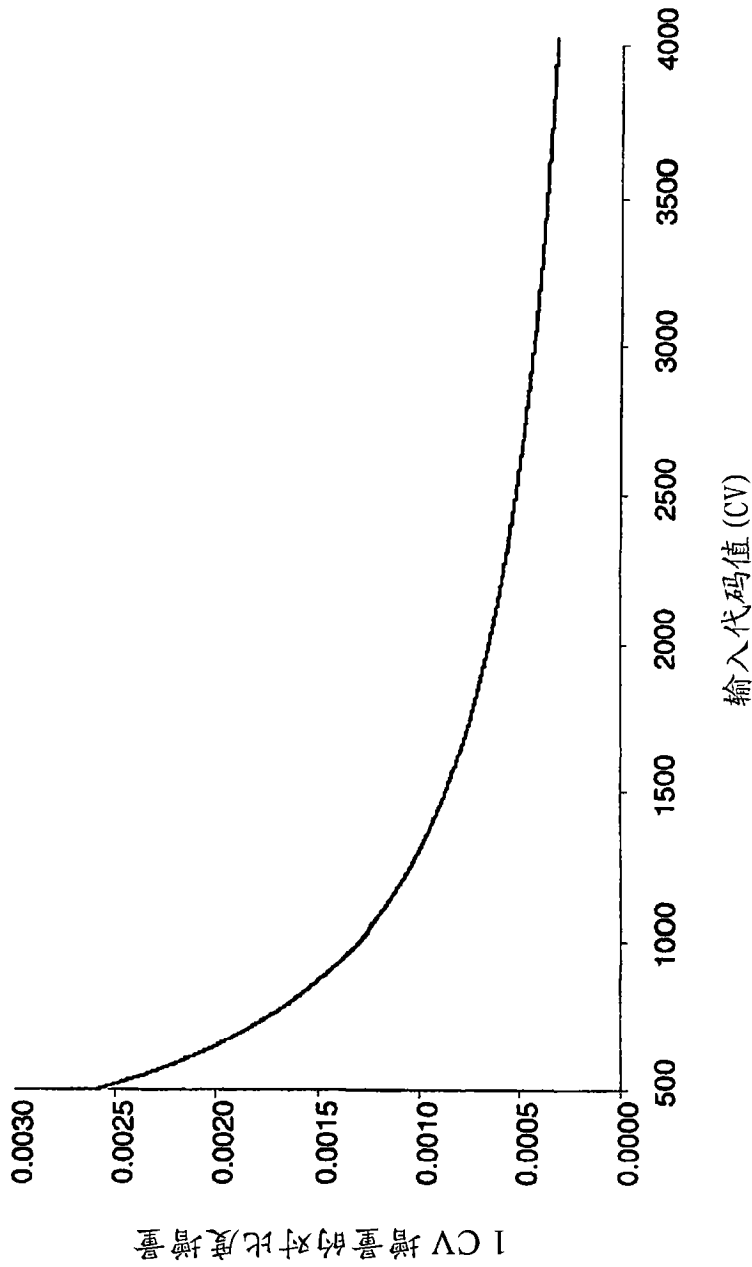


图 7

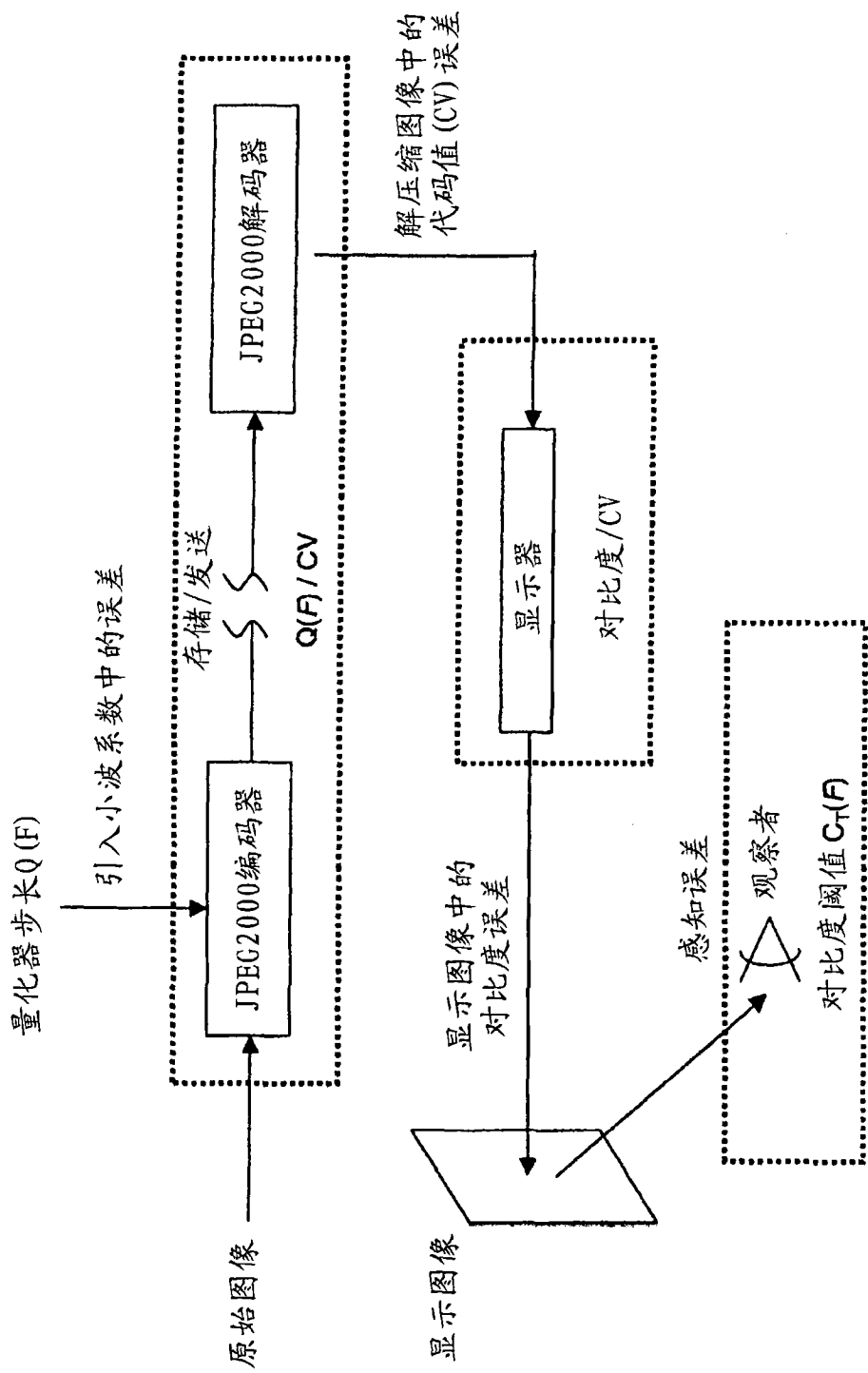


图 8

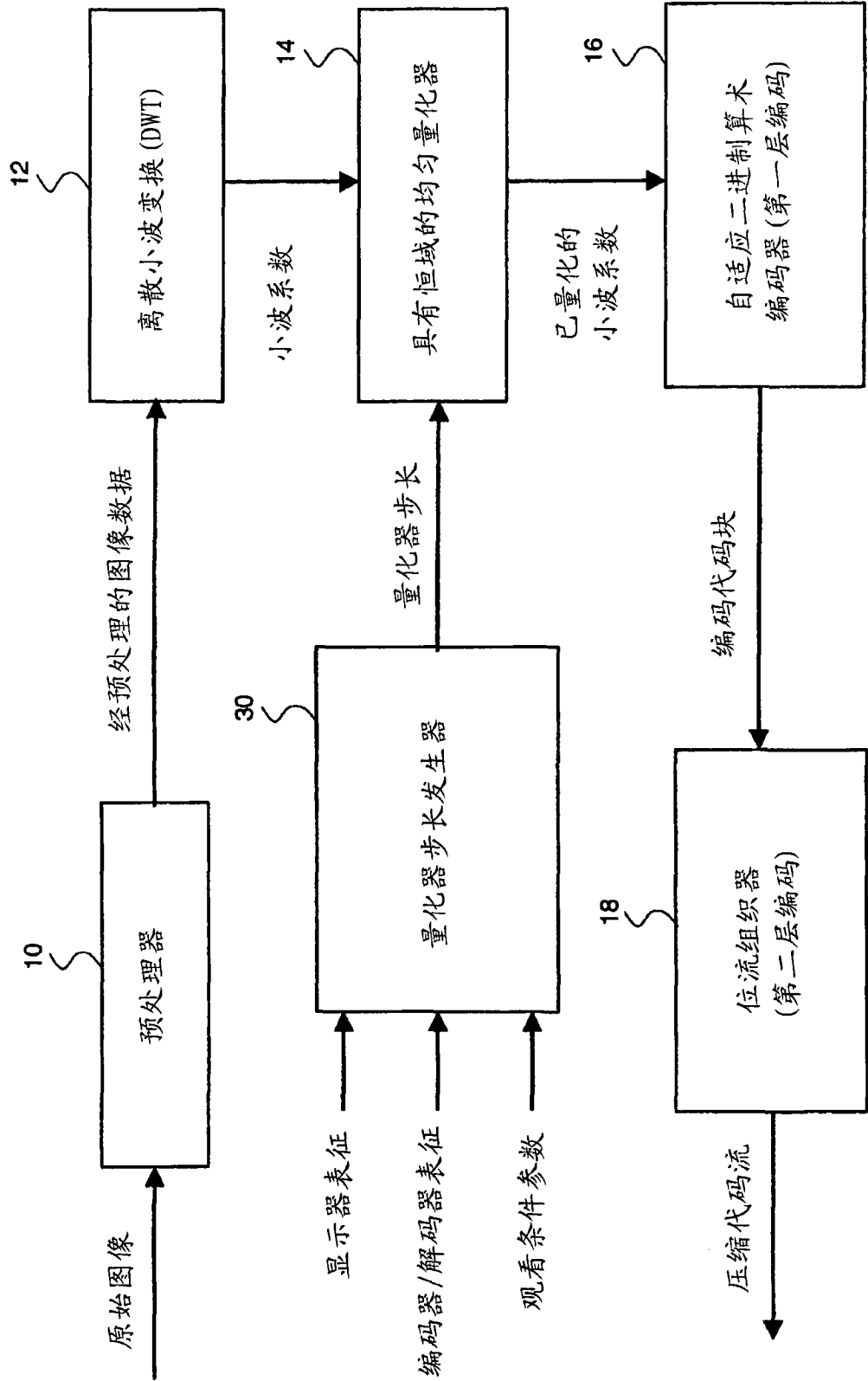


图 9

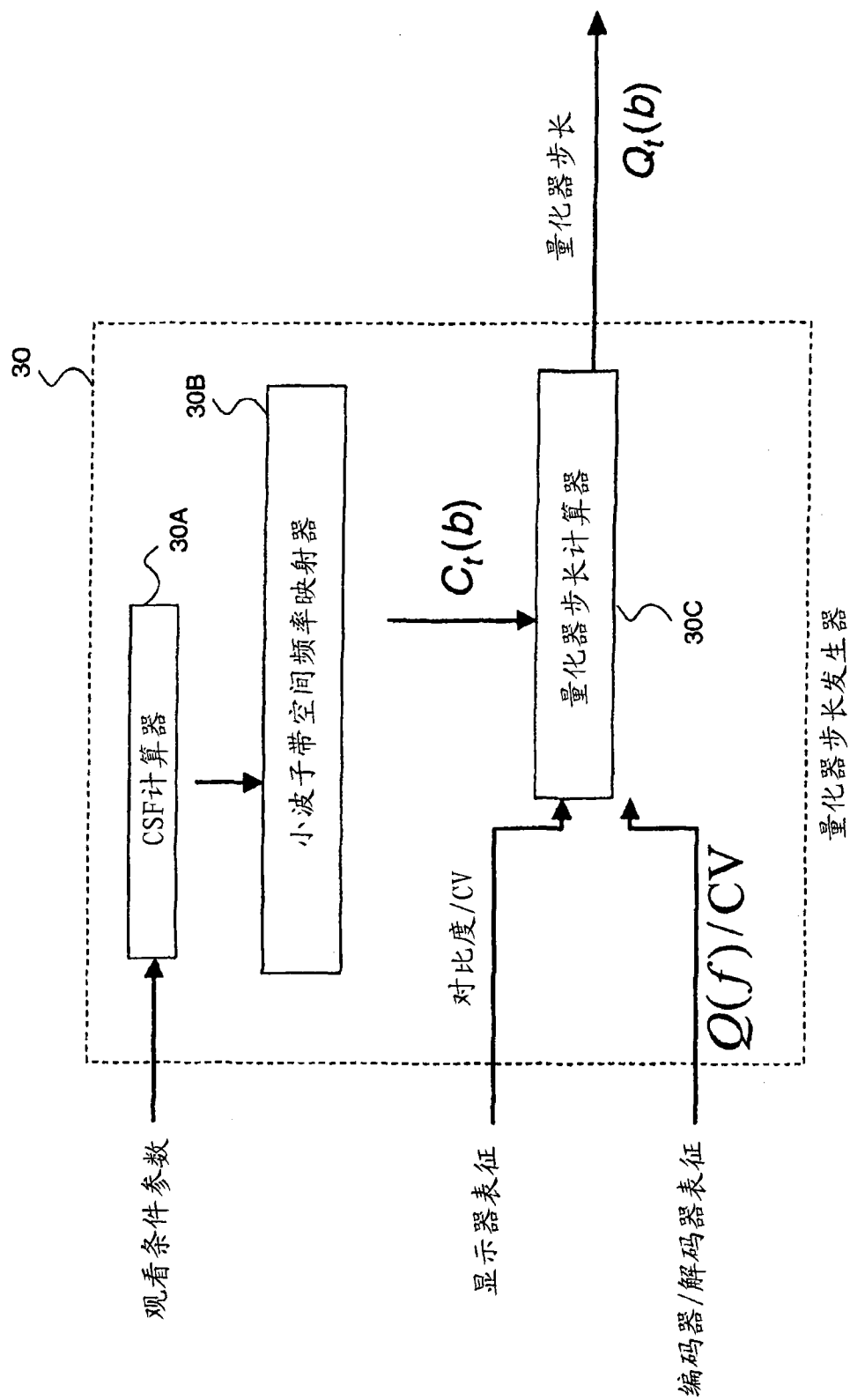


图 10

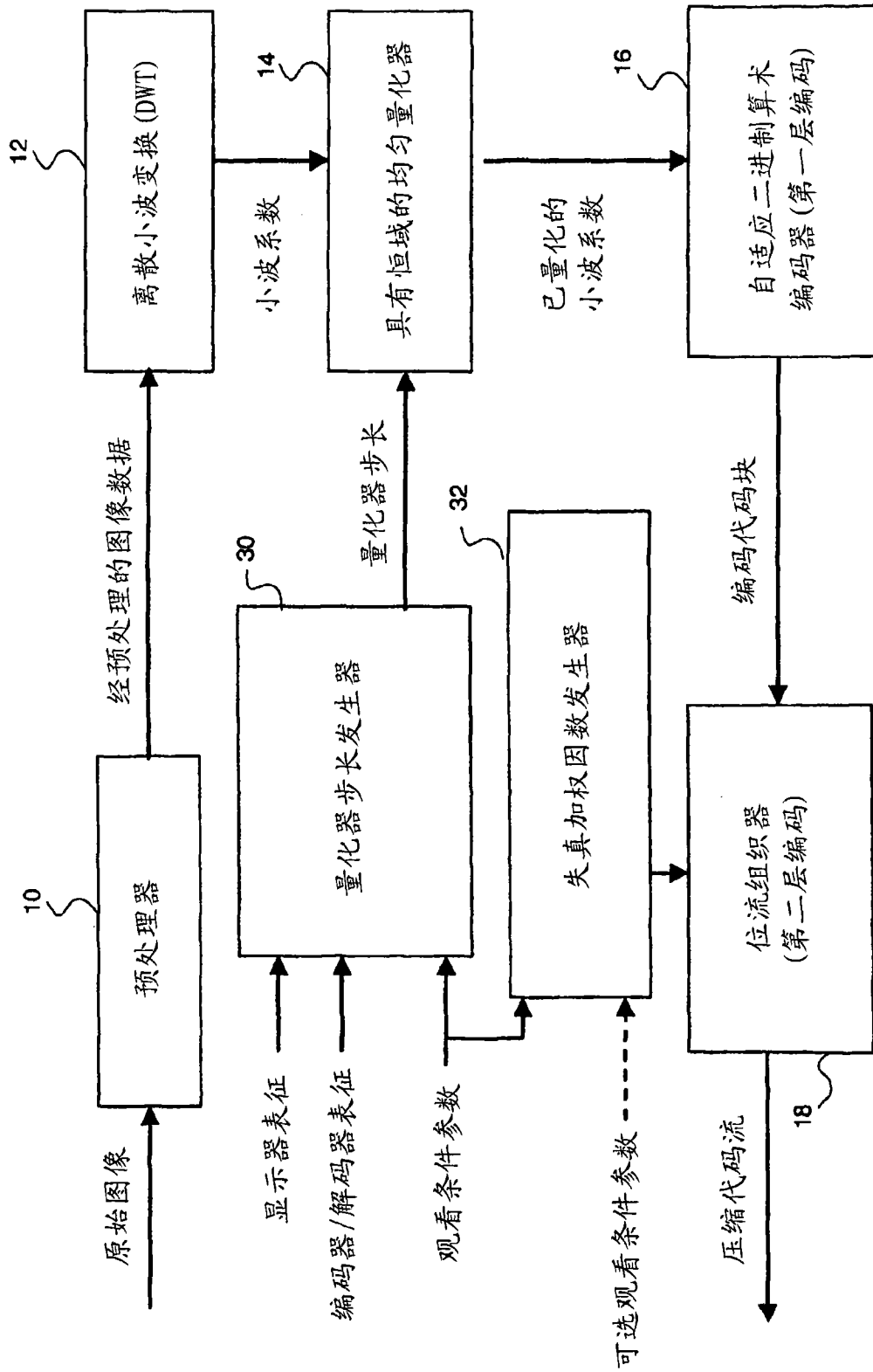


图 11

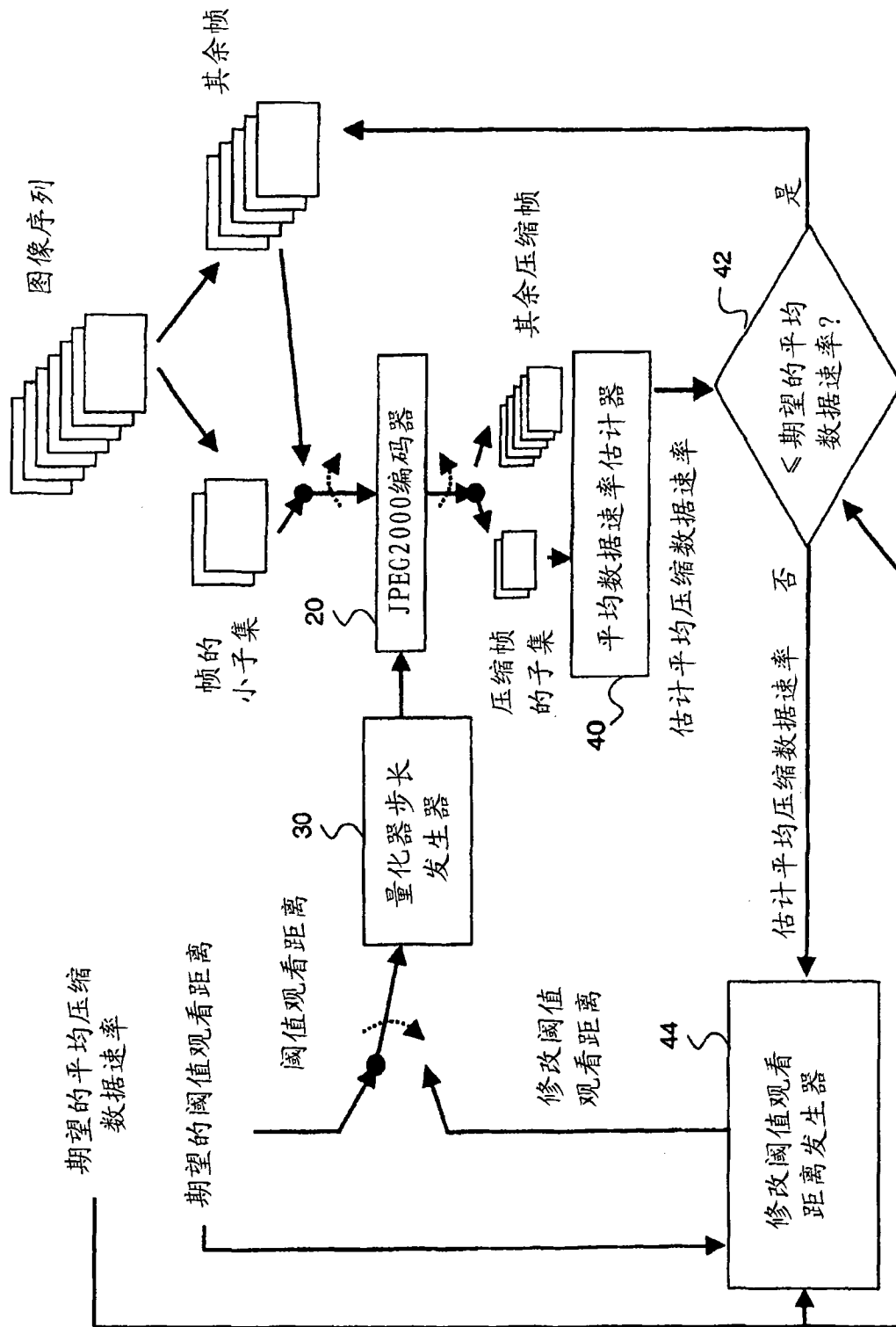


图 12

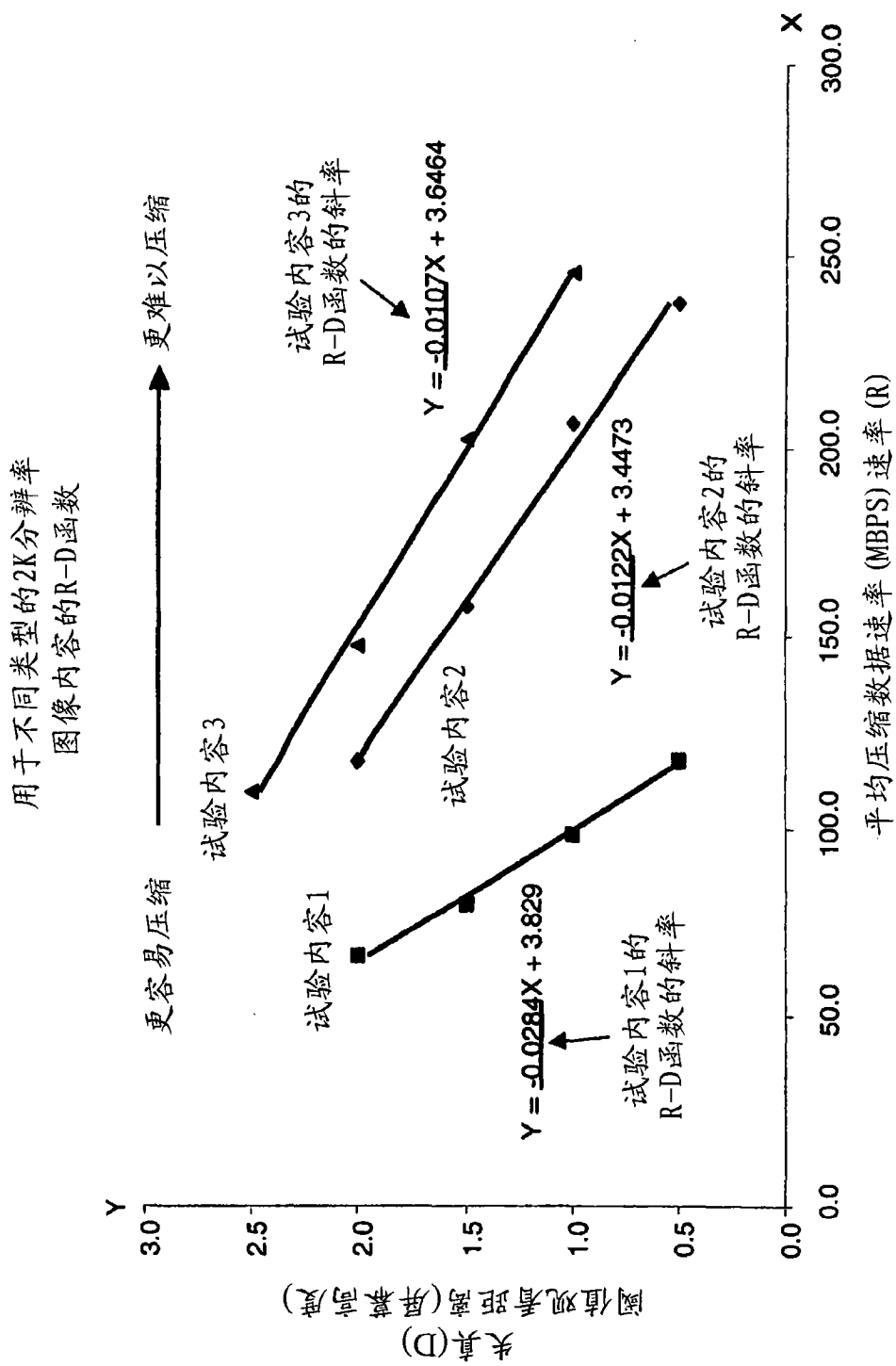


图 13

1.5屏幕高度的阈值观看距离的R-D函数的估计斜率
对比平均压缩数据速率

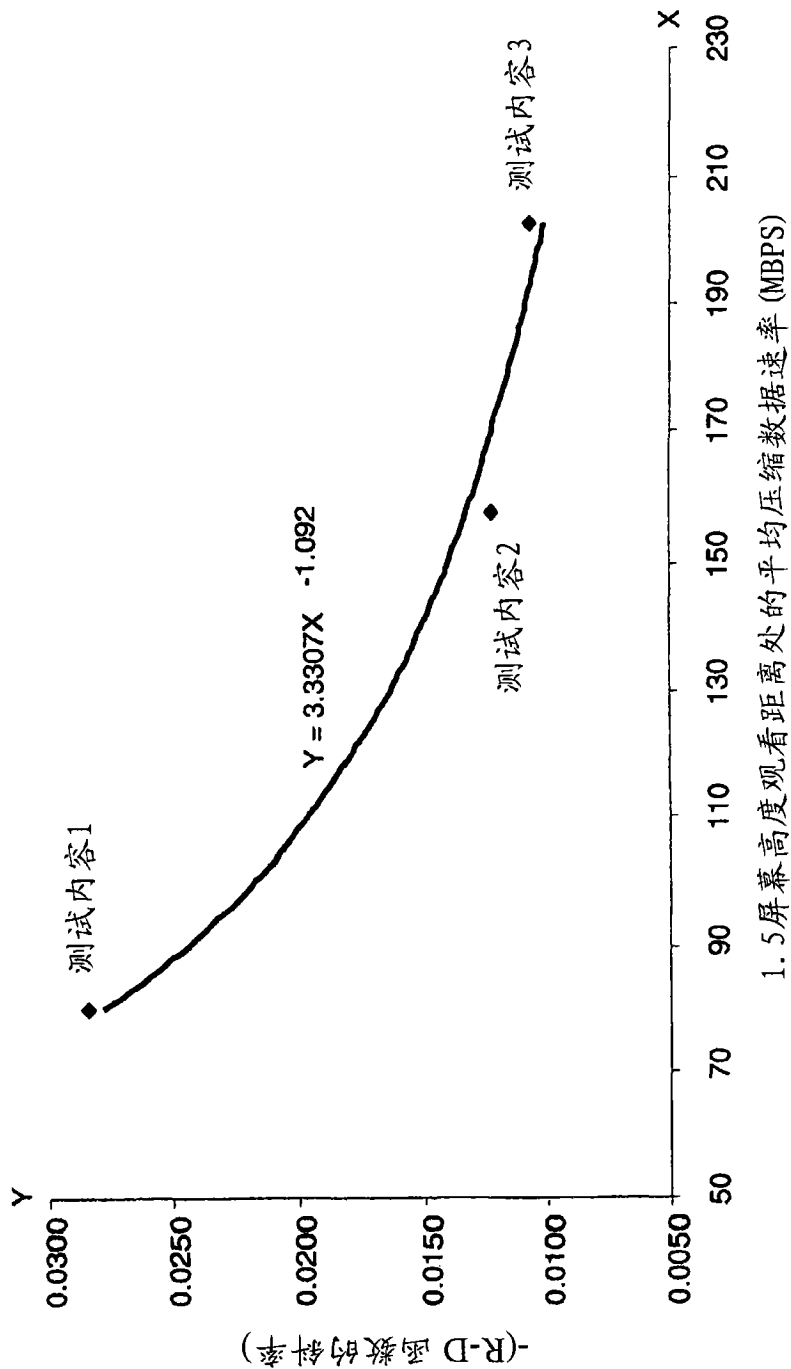


图 14

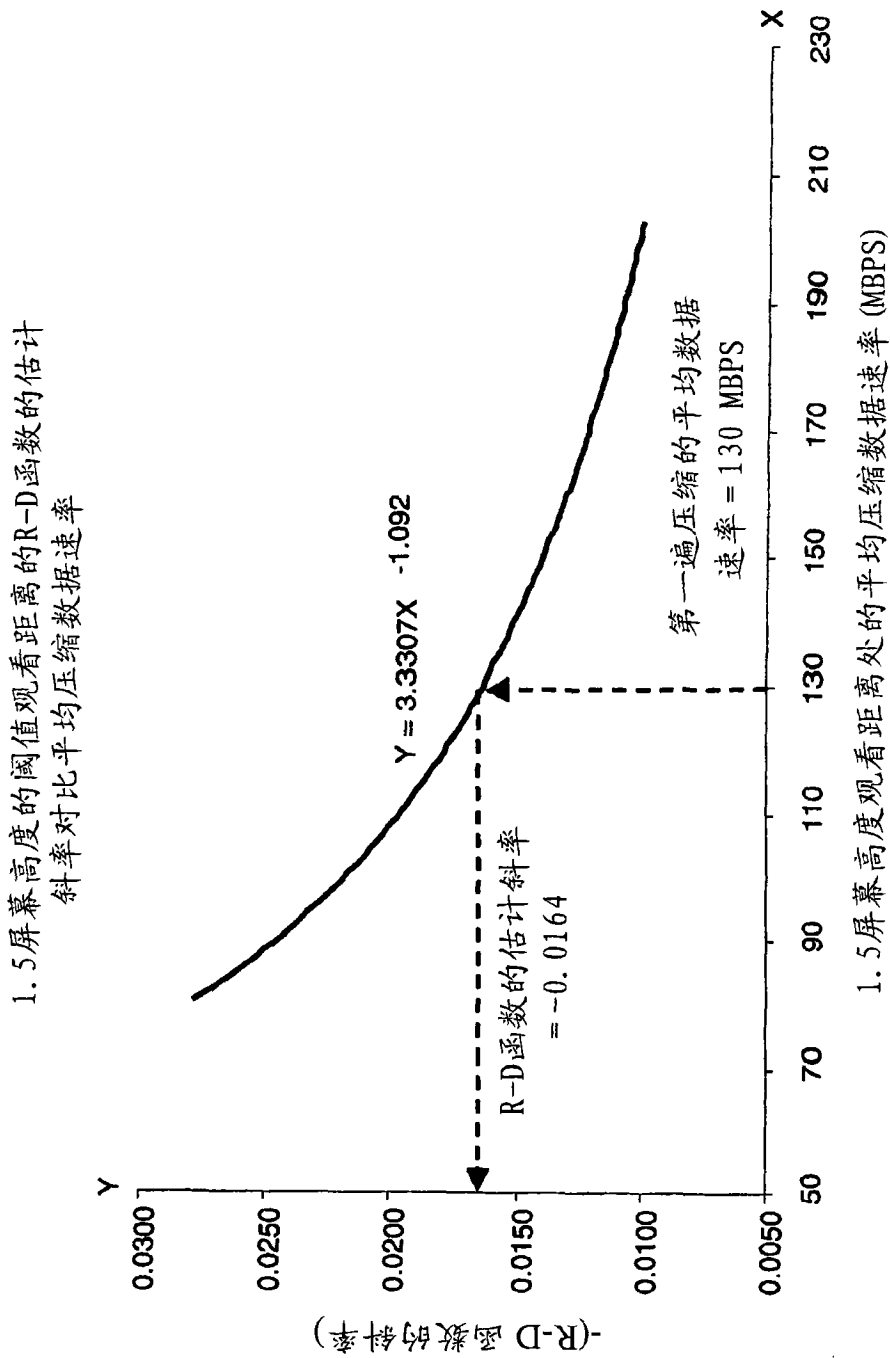


图 15

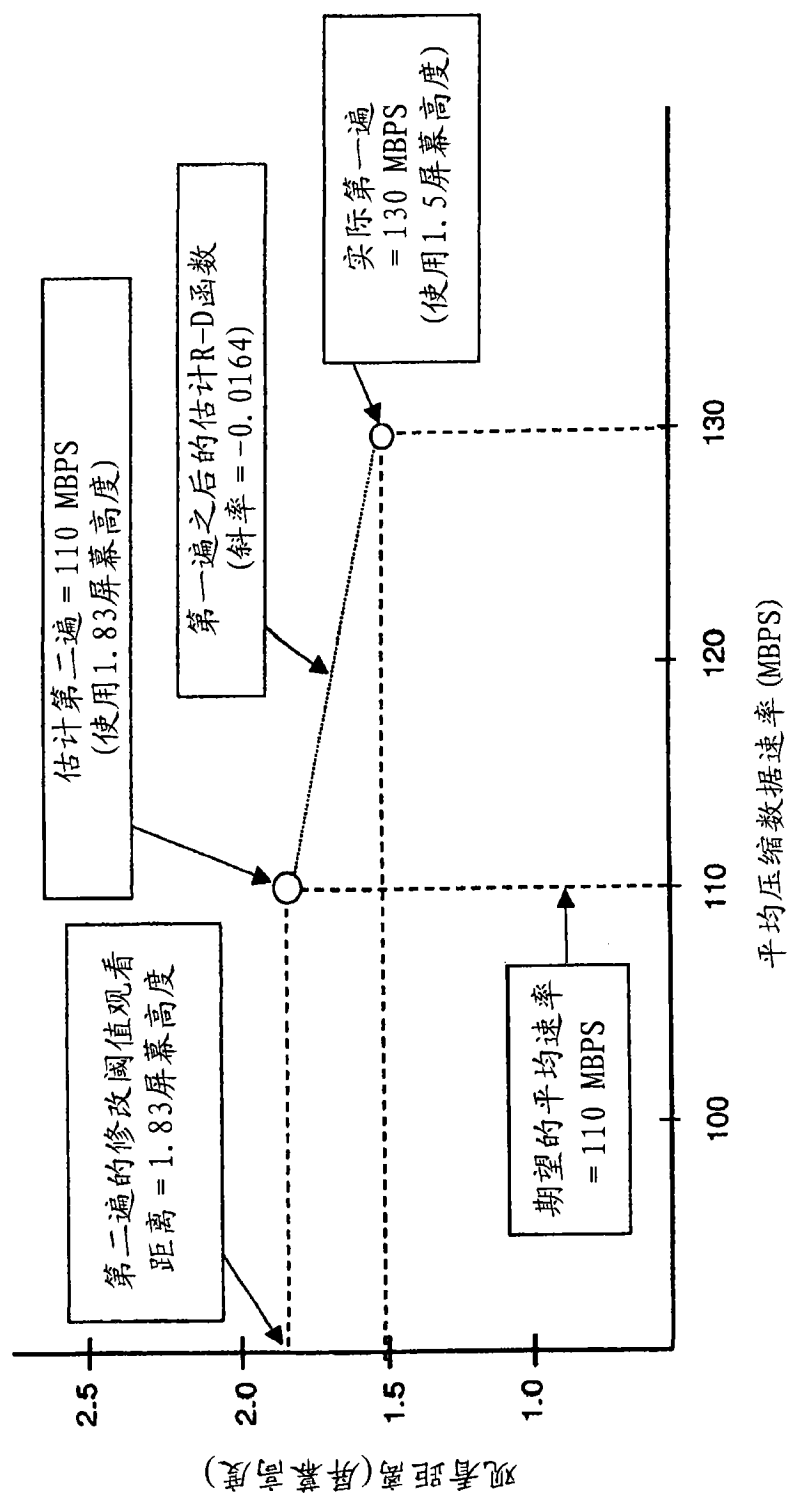


图 16

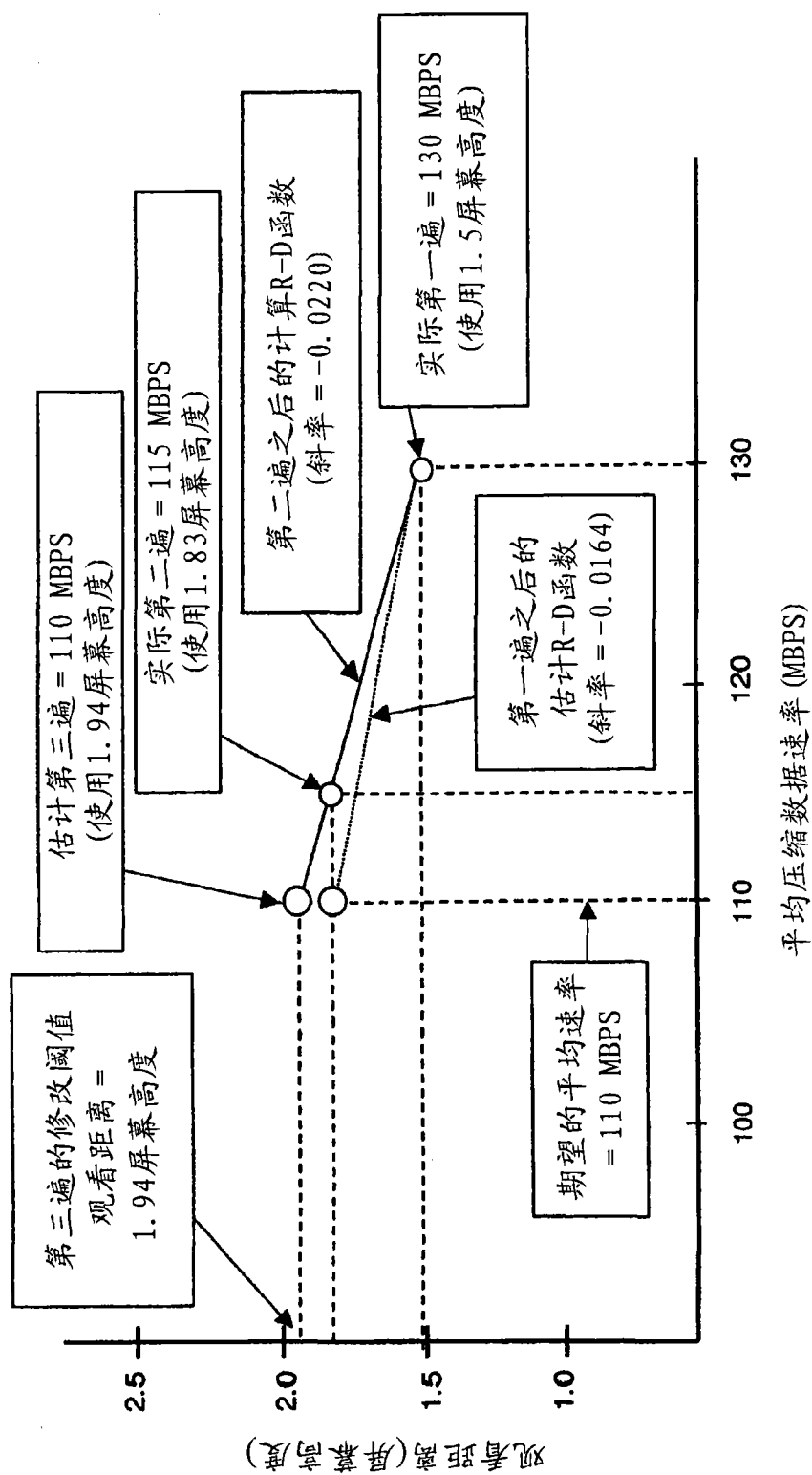


图 17

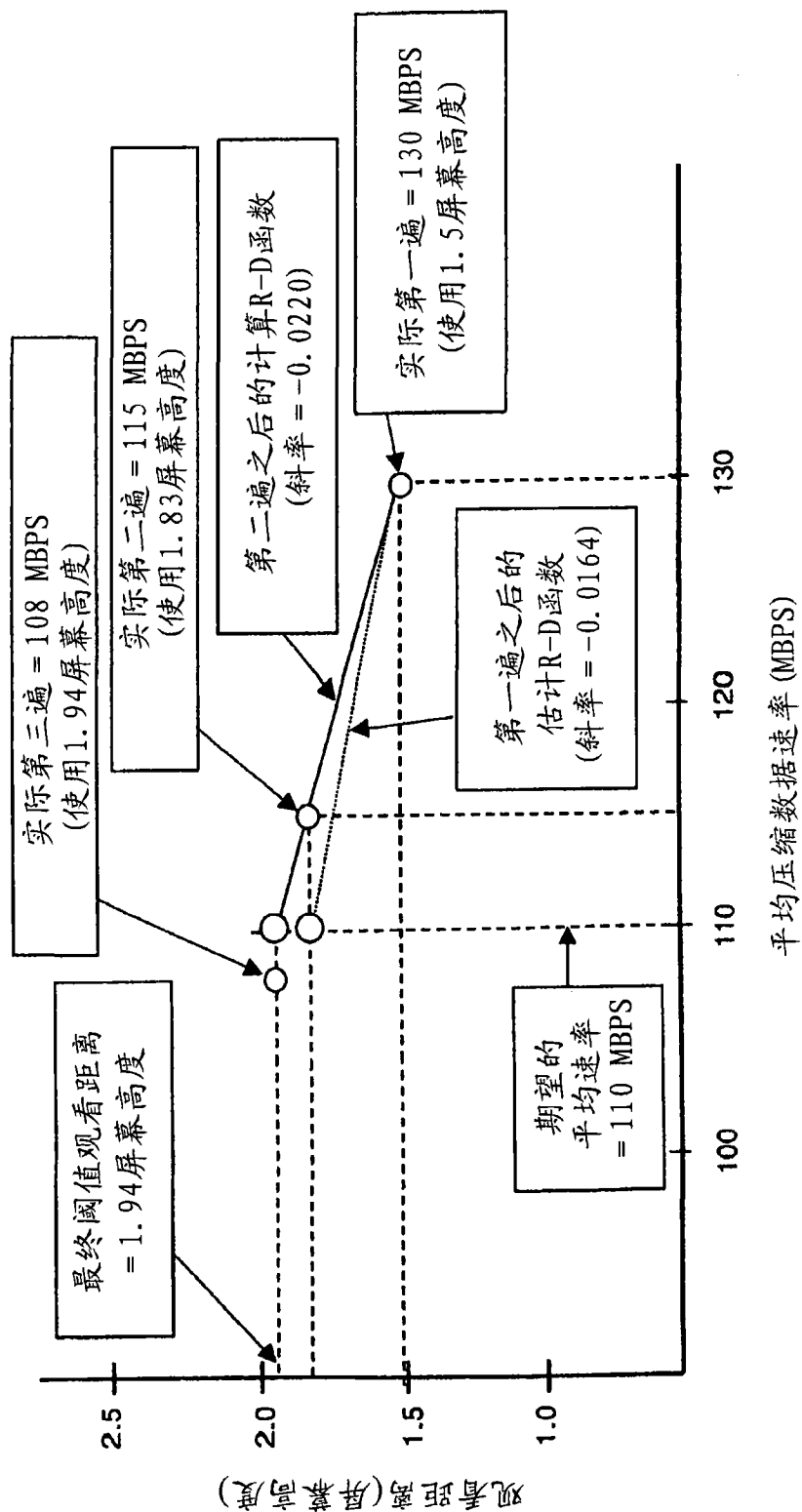


图 18

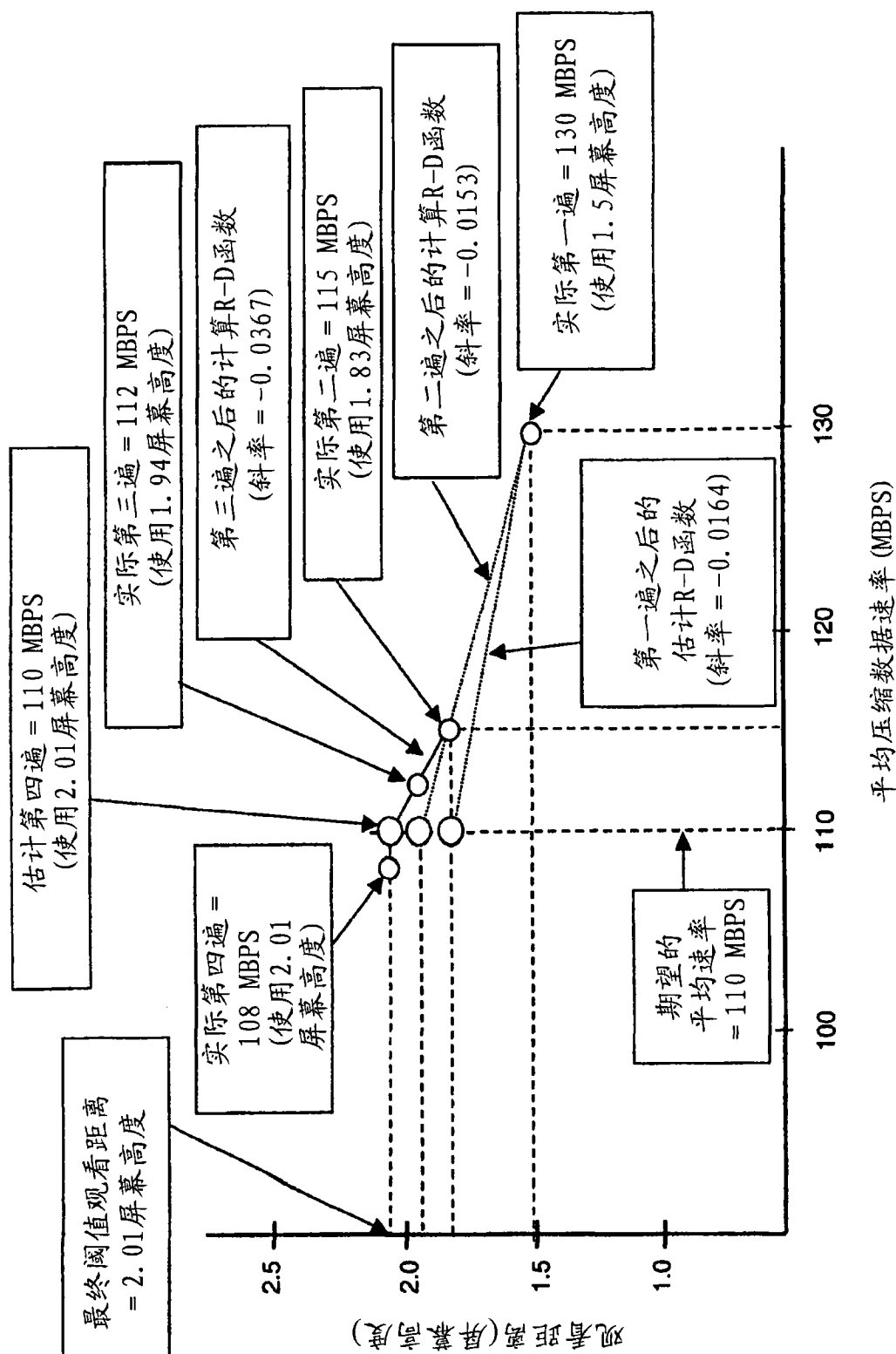


图 19