



(12) 发明专利申请

(10) 申请公布号 CN 104885452 A

(43) 申请公布日 2015. 09. 02

(21) 申请号 201380051257. 3

代理人 吕晓章

(22) 申请日 2013. 06. 25

(51) Int. Cl.

(66) 本国优先权数据

H04N 17/00(2006. 01)

PCT/CN2012/080632 2012. 08. 27 CN

(85) PCT国际申请进入国家阶段日

2015. 03. 31

(86) PCT国际申请的申请数据

PCT/CN2013/077846 2013. 06. 25

(87) PCT国际申请的公布数据

W02014/032463 EN 2014. 03. 06

(71) 申请人 汤姆逊许可公司

地址 法国伊西莱穆利诺

(72) 发明人 张帆 廖宁 谢凯 陈志波

(74) 专利代理机构 北京市柳沈律师事务所

11105

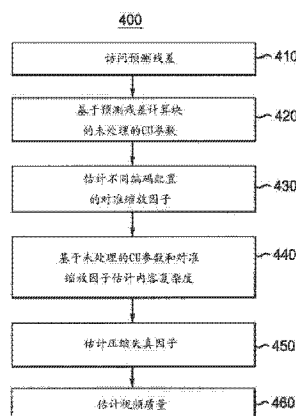
权利要求书2页 说明书8页 附图5页

(54) 发明名称

用于估计视频质量评估的内容复杂度的方法和装置

(57) 摘要

为了估计视频的内容复杂度,计算预测残差的能量。当视频较不复杂和更好预测时,预测残差通常较小。预测残差的尺度还取决于编码配置,例如即使当内容非常类似并且因此具有类似的感知到的内容复杂度时,I画面也通常具有比P和B画面更大的预测残差。为了更准确地反映内容复杂度,对于不同编码配置估计对准缩放因子。基于预测残差的能量和对准缩放因子,可以估计总内容不可预测性参数,以计算视频的压缩失真因子。结合切片和冻结失真因子的压缩失真因子可以用于估计视频的视频质量度量。



1. 一种用于评估被包括在比特流中的视频的视频质量的方法,包括下列步骤:
响应于视频中的画面的块的预测残差确定(420)内容不可预测性参数;
响应于块的编码配置缩放(440)内容不可预测性参数;以及
响应于缩放的内容不可预测性参数确定(460)表示被包括在比特流中的视频的视频质量的质量度量。
2. 如权利要求1所述的方法,其中内容不可预测性参数指示对块进行预测的难度,并且其中缩放的内容不可预测性参数的值和与该块具有类似内容的另一块的第二内容不可预测性参数的值基本上相同。
3. 如权利要求1所述的方法,其中编码配置包括画面类型和GOP(画面群组)结构中的至少一个。
4. 如权利要求1所述的方法,其中确定内容不可预测性参数的步骤包括下列步骤:
确定预测残差的能量;以及
响应于编码配置缩放预测残差的能量。
5. 如权利要求4所述的方法,其中基于预测残差遵循伽马分布来确定缩放步骤中使用的缩放因子。
6. 如权利要求5所述的方法,其中基于编码配置的伽马分布的形状参数与另一编码配置的另一形状参数相同,对于多个编码配置确定相应的缩放因子。
7. 如权利要求1所述的方法,其中对于视频中的多个块中的每一个确定相应的内容不可预测性参数,并且其中响应于多个块的内容不可预测性参数确定视频的总内容不可预测性参数。
8. 如权利要求1所述的方法,还包括:
执行监视比特流的质量、响应于质量度量调整比特流、基于质量度量创建新的比特流、调整用于传送比特流的分发网络的参数、基于质量度量确定是否保持比特流以及选择解码器处的错误隐藏模式中的至少一个。
9. 如权利要求1所述的方法,其中确定质量度量的步骤进一步响应于分辨率和帧速率中的至少一个。
10. 一种用于评估被包括在比特流中的视频的视频质量的装置(500),包括:
解码器(520),其被配置为访问视频的画面的块的预测残差并且确定块的编码配置;
内容不可预测性参数计算器(534),其被配置为
响应于块的预测残差确定内容不可预测性参数;以及
响应于块的编码配置缩放内容不可预测性参数;以及
质量预测器(550),其被配置为响应于缩放的内容不可预测性参数确定表示被包括在比特流中的视频的视频质量的质量度量。
11. 如权利要求10所述的装置,其中内容不可预测性参数指示对块进行预测的难度,并且其中缩放的内容不可预测性参数的值和与该块具有类似内容的另一块的第二内容不可预测性参数的值基本上相同。
12. 如权利要求10所述的装置,其中编码配置包括画面类型和GOP(画面群组)结构中的至少一个。
13. 如权利要求10所述的装置,其中内容不可预测性参数计算器确定预测残差的能量

以及响应于编码配置缩放预测残差的能量。

14. 如权利要求 13 所述的装置,其中内容不可预测性参数计算器基于预测残差遵循伽马分布来确定缩放因子。

15. 如权利要求 14 所述的装置,其中基于编码配置的伽马分布的形状参数与另一编码配置的另一形状参数相同,对于多个编码配置确定相应的缩放因子。

16. 如权利要求 10 所述的装置,其中对于视频中的多个块中的每一个确定相应的内容不可预测性参数,并且其中响应于多个块的内容不可预测性参数确定视频的总内容不可预测性参数。

17. 如权利要求 10 所述的装置,还包括:

视频质量监视器 (640、650、660),其被配置为执行监视比特流的质量、响应于质量度量调整比特流、基于质量度量创建新的比特流、调整用于传送比特流的分发网络的参数、基于质量度量确定是否保持比特流以及选择解码器处的错误隐藏模式中的至少一个。

18. 如权利要求 10 所述的装置,其中内容不可预测性参数计算器进一步响应于分辨率和帧速率中的至少一个确定质量度量。

19. 一种计算机可读存储介质,在其上存储用于根据权利要求 1-9 评估被包括在比特流中的视频的视频质量的指令。

用于估计视频质量评估的内容复杂度的方法和装置

[0001] 相关申请的交叉引用

[0002] 本申请要求 2012 年 8 月 27 日提交的 WO 国际申请号 PCT/CN2012/080632 的权益。

技术领域

[0003] 本发明涉及一种视频质量测量,并且更具体地,涉及一种用于确定比特流级别的视频质量度量的方法和装置。所确定的视频质量度量可以随后用于例如调整编码参数,或提供接收器侧的所要求的视频质量。

背景技术

[0004] 通常,视频内容纹理越多或越复杂,人眼可以容忍的视频内容中的伪像越多。也就是说,当通过人眼观看视频内容时,视觉伪像可以被视频内容本身屏蔽。人眼的该属性被认为是屏蔽属性或屏蔽效应。因此,可以在视觉质量评估中考虑内容复杂度。

[0005] 在我们先前的工作中,我们估计内容不可预测性(CU)参数以指示内容复杂度。在 F. Zhang、N. Liao、K. Xie 和 Z. Chen 的、名称为“Method and apparatus for video quality measurement”的共同拥有的 PCT 申请中(PCT/CN11/002096,代理人案号 PA110061,在下文中“Zhang1”),其教导通过引用具体合并于此,我们公开了一种用于使用量化参数(QP)预测视频质量的方法,该参数通过取决于内容不可预测性的校正函数而调整。

[0006] 在 F. Zhang、N. Liao、K. Xie 和 Z. Chen 的、名称为“Video quality measurement”的另一共同拥有的 PCT 申请中(PCT/CN2011/082870,代理人案号 PA110050,在下文中“Zhang2”),其教导通过引用具体合并于此,我们公开了一种用于使用从比特流推导的参数(例如,量化参数、内容不可预测性参数、丢失块的比例、传播块的比例、错误隐藏距离、运动向量、冻结持续时间和帧速率)来估计压缩失真因子、切片失真因子和冻结失真因子的方法。

发明内容

[0007] 本原理提供一种用于评估被包括在比特流中的视频的视频质量的方法,包括如下所述的下列步骤:响应于视频中的画面的块的预测残差确定内容不可预测性参数;响应于块的编码配置缩放内容不可预测性参数;以及响应于缩放的内容不可预测性参数,确定表示被包括在比特流中的视频的视频质量的质量度量。本原理还提供一种用于执行这些步骤的装置。

[0008] 本原理还提供一种用于评估被包括在比特流中的视频的视频质量的方法,包括如下所述的下列步骤:响应于视频中的画面的块的预测残差和编码配置确定内容不可预测性参数,其中内容不可预测性参数指示对块进行预测的难度;响应于编码配置缩放内容不可预测性参数,其中缩放的内容不可预测性参数的值与具有类似内容的另一块的第二内容不可预测性参数的值基本上相同;以及响应于确定的内容不可预测性参数确定表示被包括在比特流中的视频的视频质量的质量度量。本原理还提供一种用于执行这些步骤的装置。

[0009] 本原理还提供一种计算机可读存储介质,在其上存储用于根据上述方法评估被包括在比特流中的视频的视频质量的指令。

附图说明

[0010] 图 1A 是分别描绘被编码为 I、P 和 B 画面的三个原始画面的图示示例,图 1B 是描绘与图 1A 对应的未处理 (raw) 的 CU 参数的图示示例,以及图 1C 是描绘根据本原理的实施例生成的对准的 CU 参数的图示示例。

[0011] 图 2A 是分别描绘被编码为 I、P 和 B 画面的另外三个原始画面的另一图示示例,图 2B 是描绘与图 2A 对应的未处理的 CU 参数的图示示例,以及图 2C 是描绘根据本原理的实施例生成的对准的 CU 参数的图示示例。

[0012] 图 3 是分别描绘 I、P 和 B 画面的未处理的 CU 参数分布的图示示例。

[0013] 图 4 是描绘根据本原理的实施例的基于内容复杂度估计视频质量的示例的流程图。

[0014] 图 5 是描绘可以与本原理的一个或多个实现方式一起使用的视频质量测量装置的示例的框图。

[0015] 图 6 是描绘可以与本原理的一个或多个实现方式一起使用的视频处理系统的示例的框图。

具体实施方式

[0016] 已知方差和熵可以用于测量自然图像的复杂度。对于一般的信号序列,方差量化对比度或能量,而熵评估随机性或不确定性。在 A. B. Watson、R. Borthwick 和 M. Taylor 的“Image quality and entropy masking”SPIE Proceedings, vol. 3016, 1997 中描述的工作提出除了对比度和随机性之外的、被称为学习能力的另一与复杂度相关的特性可以用于测量内容复杂度。

[0017] 学习能力可以被认为与可预测性类似。也就是说,信号可以通过邻近信号表示的可能性越高,其可预测性和学习能力越高。在视频编码中,帧内预测和帧间预测通常用于消除视频中的冗余。具体地,可以从先前编码的块预测当前块,并且可以对当前块和预测块之间的差(即,预测残差),而非当前块本身进行编码。通常,当帧内预测或帧间预测工作良好时,也就是说,当可以从先前编码的块良好地预测当前块时,预测残差变小和/或变少。因此,视频编码中的预测残差的能量可以是视频信号多难预测(即,内容不可预测性)的良好指示器。在本实施例中,我们使用内容不可预测性来表示内容复杂度。

[0018] 当基于预测残差定义内容可预测性参数时,内容可预测性参数可以被计算为测量对视频信号进行预测的难度的参数。因为在压缩域中可获得预测残差,因此可以在没有完全解码或获取解码的图像像素的情况下测量内容可预测性参数,并且因此提高使用 CU 参数,例如 QoE(经验质量)评估和视频指数的应用的计算效率。

[0019] 大多数现有视频压缩标准,例如 H. 264 和 MPEG-2,使用宏块(MB)作为基本编码单元。因此,下列实施例使用宏块作为基本处理单元。然而,本原理可以被适配为使用不同大小的块,例如 8×8 块、 16×8 块、 32×32 块或 64×64 块。

[0020] 在一个实施例中,对于每一个宏块,可以使用预测残差的方差或能量计算 CU 参

数,该方差或能量理论上等于平方去量化 DCT 系数的和减去去量化 DC 系数的平方和。当没有对宏块执行预测时, CU 参数可以被计算为像素的能量。也就是说,逆 DCT (IDCT) 对于计算 CU 参数不是必要的。给定从比特流解析的 QP 和量化 DCT 系数,可以以足够的精度估计去量化 DCT 系数以便计算 CU 参数。

[0021] 在示例中,在 Zhang1 中,当 H. 264/AVC 用于编码时,第 r 个宏块的 CU 参数可以被计算为:

$$[0022] \quad CU_r = \frac{(0.625 \times 2^{QP_r/6})^2}{K_{MB}} \left[\sum_{l \in r} \sum_{k \in l} DCT_{k,l}^2 - \left(\sum_{l \in r} \frac{DCT_{0,l}}{|l|_r} \right)^2 \right], \quad (1)$$

[0023] 其中常数 K_{MB} 是 H. 264/AVC 的宏块中像素的总数,即 256; $0.625 \times 2^{QP_r/6}$ 是关于 H. 264/AVC 的 QP 值计算量化步骤的近似公式; l 为第 r 个宏块中的分区 (partition) 做索引; $|l|_r$ 表示其中分区的总数,例如对于 4×4 块分区为 16 分区以及对于 8×8 块分区为 4 分区;非负 k 为第 r 个宏块的第 l 个分区中的量化 DCT 系数做索引;以及 $DCT_{0,l}$ 指代第 l 个分区中的量化的 DC 系数。

[0024] 上面的 CU 计算可以应用于任何大小的块 (或例如 HEVC 的其他标准中的编码单元 / 预测单元) 和块内的任何分区。在本应用中,例如等式 (1) 中计算的、基于预测残差的能量计算的内容不可预测性参数被表示为未处理的 CU 参数。

[0025] 预测技术根据编码画面类型而变化。例如, I 画面仅支持帧内预测并且可以具有比 P 和 B 画面更大的预测残差。由于帧间预测, P 画面通常具有比 I 画面更小的预测残差。因为双向帧间预测, B 画面通常具有比 I 和 P 画面更小的预测残差。也就是说,不同类型的画面包含不同数值尺度的预测残差,即使对应的画面可以具有类似内容 (例如,画面属于相同场景,或画面之间的相似度度量小于阈值)。因此,被计算为画面上的预测残差能量的未处理的 CU 参数不仅取决于视频内容,而且取决于画面类型,并且因此可能不能准确地反映画面的复杂度。

[0026] 图 1A 分别示出被编码为 I、P 和 B 画面的三个示例原始画面。图 1B 示出对应的未处理的 CU 参数。为了可视化, CU 参数被对数地映射为像素亮度 (即,图 1B 中的像素的亮度对应于图 1A 中的宏块的 CU 参数的值)。像素越亮, CU 参数越大。虽然三个原始画面在图像内容上非常类似,并且因此在感知的内容复杂度上非常类似,但是 I 画面具有最大未处理的 CU 参数并且 B 画面具有最小未处理的 CU 参数。在图 2A 和 2B 中示出的另一示例中也进行类似观察。

[0027] 图 3 示出不同画面类型的宏块的未处理的 CU 参数值的示例分布。概率分布函数通过具有 15 的区间 (bin) 宽度 (即,未处理的 CU 参数值的间隔) 的标准化直方图来近似。在 576 行视频的连续 250 帧中,对来自 I、P 和 B 画面的全部宏块分别计算未处理的 CU 参数。总地,未处理的 CU 参数对于 I 画面最大,并且对于 B 画面最小。从 I、P 到 B 画面,概率分布曲线的尾部变得更细。图 3 的图例列出 Γ 分布参数,其单独地适用于每种类型的画面。在该示例中,我们观察到形状参数 (I 画面为 0.033, P 画面为 0.041 并且 B 画面为 0.049) 相当类似,而缩放参数 (I 画面为 4916, P 画面为 1086 并且 B 画面为 324) 非常不同。

[0028] 根据图 1-3 中示出的示例,我们观察到即使当感知的内容复杂度非常类似时,未处理的 CU 参数可以随着画面类型显著变化。为了更准确地测量内容复杂度,本原理针对一种考虑到上文所述的感知到的变化的用于改进估计的内容不可预测性参数的方法和装置。

具体地,对未处理的 CU 参数进行缩放以与在另一编码配置中获得的 CU 参数对准,使得对准的 CU 参数更接近地反映感知到的内容复杂度。当另一块具有与当前块类似的内容时,在 CU 参数对准之后,在一个编码配置中的当前块的对准的 CU 参数的值与在另一编码配置中该另一块的对准的 CU 参数的值基本上相同。

[0029] 所提出的 CU 参数对准方法基于下文中讨论的两个一般假设。

[0030] 假设 1 宏块的未处理的 CU 参数值遵照独立伽马分布 (Γ 分布)。 Γ 分布的、具有形状参数 k 和缩放参数 θ 的、被表示为 $Y \sim \Gamma(k, \theta)$ 的随机变量 Y 具有概率分布函数:

$$[0031] \quad \text{对于 } y \geq 0 \text{ 和 } k, \theta > 0, f(y; k, \theta) = \frac{1}{\theta^k \Gamma(k)} y^{k-1} e^{-\frac{y}{\theta}} \quad (2)$$

[0032] 假设 1 基于宏块的未处理的 CU 参数被定义为预测残差或像素的方差的观察结果,以及下列引理:

[0033] 引理:如果 $x_1, x_2, \dots, x_n$ 为独立同分布 i.i.d. $\mathcal{N}(\mu, \sigma^2)$ 随机变量,则它们的方差,即 $\sum_{i=1}^n (x_i - \frac{1}{n} \sum_{i=1}^n x_i)^2$ 遵照伽马分布的特殊情况 $\Gamma(\frac{n-1}{2}, 2\sigma^2)$ 。

[0034] 严格地说,相邻像素和残差都不是独立高斯分布的,因此我们作出它们的方差遵照具有宽松的形状参数 k 的广义 Γ 分布 $\Gamma(k, 2\sigma^2)$, 其中 k 不再准确地等于 $(n-1)/2$ 的小心的近似。通常,相邻像素的相关性和稀疏性使得 k 比 $(n-1)/2$ 小得多。然而,尽管编码配置不同,但是对于相同视觉内容,这种相关性应相对一致。因此,我们进行下列假设。

[0035] 假设 2 在不同配置下, (相同或类似内容的) 未处理的 CU 参数分布具有常量形状参数 k 和可能不同的缩放参数 θ 。

[0036] 基于上述假设, CU 对准的问题变为估计不同配置的缩放参数 θ 。假设存在两种配置 X 和 Y , 估计问题可以通过 CU 训练集合 $\{x\}_X$ 和 $\{y\}_Y$ 上的分布参数的最大似然估计来解决。

$$[0037] \quad \arg \max_{k, \{\theta\}} \prod_{i=1}^{n_X} f(x_i; k, \theta_X) \prod_{i=1}^{n_Y} f(y_i; k, \theta_Y) \quad (3)$$

[0038] 其中 x 和 y 分别表示在不同配置 X 和 Y 下的 MB 的未处理的 CU 参数值, 并且 n_X 和 n_Y 分别是集合 $\{x\}_X$ 和 $\{y\}_Y$ 的总数。

[0039] 随后,我们可以对于每一种配置对准 MB 方式 (MB-wise) 的 CU, 并且获得对准的 MB 方式的 CU (被表示为 $\widehat{CU}$), 例如:

[0040]

$$\widehat{CU}_i = \begin{cases} CU_i \theta_Y / \theta_X & \text{如果使用 } X \text{ 配置 } MB_i \\ CU_i & \text{如果使用 } Y \text{ 配置 } MB_i \end{cases} \quad (4)$$

[0041] 在下文中,我们提供一种用于解决如等式 (3) 中描述的估计问题的示例实施例。参考等式 (2) 和 (3), 对数似然为

[0042]

$$\begin{aligned} \ell(k, \{\theta\}) = & (k-1) \left(\sum_i \ln x_i + \sum_i \ln y_i \right) - (n_x + n_y) \ln \Gamma(k) \\ & - \left(\sum_i x_i / \theta_x + \sum_i y_i / \theta_y \right) - k(n_x \ln \theta_x + n_y \ln \theta_y). \end{aligned}$$

[0043] 关于 θ_x 将对数似然的导数设置为零, 得到:

[0044]

$$\frac{\partial \ell}{\partial \theta_x} = \sum_i x_i / \theta_x^2 - k n_x / \theta_x = 0,$$

[0045] 并且导致 θ_x 的最大似然估计:

$$[0046] \quad \hat{\theta}_x = \sum_i x_i / (k n_x).$$

[0047] 类似地, θ_y 的最大似然估计可以被推导为:

$$[0048] \quad \hat{\theta}_y = \sum_i y_i / (k n_y).$$

[0049] 因此, 在没有估计 k 的情况下, 我们具有:

$$[0050] \quad \hat{\theta}_y / \hat{\theta}_x = \frac{n_x \sum_i y_i}{n_y \sum_i x_i} \quad (5)$$

[0051] 也就是说, 等式 (5) 提供配置 X 和 Y 的缩放参数 θ 之间的比例, 其中比例也被称为对准缩放因子, 以对未处理的 CU 参数进行缩放, 以便在配置 X 和 Y 之间将它们对准。

[0052] 在一个示例中, 当 ITU-T P. NBAMS (视频媒体流传输质量的参数非侵入式比特流评估) 数据库用于训练时, 我们推导对准缩放因子并且估计视频序列的内容复杂度作为视频序列的全部画面中的全部可用 MB 的对准的 CU 参数的平均值。在该示例中, 内容复杂度被细化为 CU 的加权组合:

[0053]

$$z_c = \frac{1}{T} \left(\sum_{t \in \{I\}} \sum_{r \in t} \frac{CU_r}{20.6 \times |r|_t} + \sum_{t \in \{P\}} \sum_{r \in t} \frac{CU_r}{3.52 \times |r|_t} + \sum_{t \in \{B\}} \sum_{r \in t} \frac{CU_r}{|r|_t} \right) + c, \quad (6)$$

[0054] 其中 $t \in \{I\}$, $t \in \{P\}$ 和 $t \in \{B\}$ 分别表示 I 画面、P 画面和 B 画面, CU_r 为第 r 个 MB 的未处理的内容不可预测性参数值, T 为画面的总数, c 为常数, 并且 $|r|_t$ 表示第 r 个 MB 中的分区的总数。常数 20.6 为 I 画面和 B 画面之间的对准缩放因子, 并且常数 3.52 为 P 画面和 B 画面之间的对准缩放因子。

[0055] 对于图 1A 和 2A 的示例, 对准的 CU 参数分别在图 1C 和 2C 中示出。与图 1B 和 2B 中示出的未处理的 CU 参数相比较, 对准的 CU 参数跨越不同画面类型上具有更类似的缩放。因此, 对准的 CU 参数可以更好地反映内容复杂度。

[0056] 除了画面类型之外, 其他编码配置, 例如使用的压缩标准、最大连续 B 画面、GOP (画面群组) 结构、即时刷新率和帧内预测的预测方向的总数可以影响预测效率并且从而会影响 B、P 或 I 画面的平均未处理的 CU 参数值, 特别是当不频繁地使用 I 帧时 (即, 使用长 GOP)。此外, 未处理的 CU 参数值还取决于视频分辨率和帧速率。例如, 更高的分辨率或帧速率通常导致更低的未处理的 CU 参数值。因此, 需要跨越不同配置确定对准。在一个实施例中, 取决于编码配置, 可以使用不同训练集合并且因此可以相应地估计对准缩放因子。

例如,关于分辨率的幂函数和关于帧速率的幂函数可以用于对准,其可以保持对准的 CU 参数与配置独立并且仅与内容有关。

[0057] 例如等式 (6) 中的 z_c 的视频序列的估计的内容不可预测性参数可以用于替换 Zhang1 (PCT/CN11/002096) 的等式 (1) 中的 Q_c 用于测量视频质量,或者 Zhang2 (PCT/CN2011/082870) 的等式 (3) 中的 CU_T 用于计算压缩失真因子。也就是说,压缩失真因子可以被计算为:

$$[0058] \quad d_c = (\log z_c)^{b_1} \times (51 - QP_T)^{b_2}, \quad (7)$$

[0059] 其中 QP_T 为视频剪辑的平均 QP, 并且 b_1 和 b_2 为常数。

[0060] 结合压缩失真因子和其他失真因子 (例如,切片失真因子和冻结失真因子),对于视频序列可以获得总视频质量度量。因为预测残差在比特流中可用,因此根据本原理的视频质量测量可以在比特流级别上实现。

[0061] 使用 z_c 和其他参数,质量度量可以被计算为:

[0062]

$$q = \frac{MOS_{ub} - MOS_{lb}}{1 + \alpha (a_c x_c^{b_{c0}} z_c^{b_{c1}} R^{b_{c2}} F^{b_{c3}} + a_f x_f^{b_{f0}} z_f^{b_{f1}} + a_s x_s^{b_{s0}} z_s^{b_{s1}})^{\beta}} + MOS_{lb}, \quad (8)$$

[0063] 其中,输出变量 q 是预测的质量分数,常数 MOS_{ub} 和 MOS_{lb} 分别是 MOS (平均意见分数) 的上限和下限,即 5 和 1; α , β , $\{a\}$ 和 $\{b\}$ 是模型参数 ($a_c = 1$ 不变); 下标 c 、 f 和 s 分别指示压缩、冻结和切片损失; 变量 $\{x\}$ 和 $\{z\}$ 是模型因子并且还通常被称为特征,该特征从视频数据中提取; R 是宽度 * 高度 / 256; 并且 F 是帧速率。具体地, $\{x\}$ 和 $\{z\}$ 分别是与每一种类型的损失相关联的关键因子和公共变量,例如 x_c 是压缩损失的关键因子,并且 z_s 是切片损失的公共变量。

[0064] 图 4 示出根据本原理的实施例的基于内容复杂度测量视频质量的示例方法 400。方法 400 在步骤 410 开始,其中从例如比特流访问预测残差。在步骤 420,计算未处理的 CU 参数例如作为等式 (1) 中示出的预测残差的能量。在步骤 430,例如使用等式 (5) 估计不同编码配置的对准缩放因子。基于未处理的 CU 参数和对准缩放因子,在步骤 440 处估计对准的 CU 参数以估计内容复杂度。在步骤 450,可以基于对准的 CU 参数和其他编码参数 (例如,量化参数) 计算压缩失真因子。结合压缩失真因子和切片和 / 或冻结失真因子,可以在步骤 460 例如使用等式 (8) 估计总视频质量度量。

[0065] 除了视频质量测量之外,内容不可预测性参数还可以用在纹理识别、对象识别或场景识别中。

[0066] 图 5 描绘可以用于生成视频质量度量的示例性视频质量测量装置 500 的框图。装置 500 的输入包括包含比特流的传输流。输入可以是包含比特流的其他格式。在系统级处的接收器确定在接收到的比特流中的分组丢失。

[0067] 解复用器 510 解析输入流以获得基本流或比特流。解复用器 510 还将关于分组丢失的信息传递给解码器 520。解码器 520 解析必要信息,包括每一个块或宏块的 QP、画面类型、变换系数和运动向量,以便生成用于估计视频质量的参数。解码器还使用关于分组丢失的信息来确定视频中的哪些宏块丢失。解码器 520 被表示为部分解码器以强调并未执行完全解码,即视频并未被重构。

[0068] 使用从解码器 520 解析的 MB 级 QP, QP 解析器 533 获得画面和整个视频剪辑的平均 QP。使用从解码器 520 获得的变换系数,变换系数解析器 532 解析该系数。内容不可预测性参数计算器 534 基于变换系数、量化参数和编码配置计算单个画面和整个视频剪辑的内容不可预测性参数。使用关于哪些宏块丢失的信息,丢失 MB 标签器 531 标记哪一个 MB 丢失。进一步使用运动信息,传播 MB 标签器 535 标记哪些 MB 直接或间接使用丢失块用于预测(即,哪些块被错误传播影响)。使用块的运动向量,MV 解析器 536 计算 MB、画面和整个视频剪辑的运动参数。其他模块(未示出)可以用于确定错误隐藏距离、冻结持续时间和帧速率。

[0069] 压缩失真预测器 540 估计压缩失真因子,切片失真预测器 542 估计切片失真因子,并且冻结失真预测器 544 估计冻结失真因子。基于估计的失真因子,质量预测器 550 估计总视频质量度量。

[0070] 当允许额外的计算时,解码器 570 解码画面。解码器 570 被表示为完全解码器并且其将重构画面并且必要时执行错误隐藏。马赛克检测器 580 对重构的视频执行马赛克检测。使用马赛克检测结果,丢失 MB 标签器 531 和传播 MB 标签器 535 更新相关参数,例如丢失块标志和传播块标志。纹理屏蔽估计器 585 计算纹理屏蔽加权。纹理屏蔽加权可以用于对失真进行加权。

[0071] 视频质量测量装置 500 可以用在例如 ITU-T P. NBAMS 标准中,其在两种应用场景中,即 IPTV 和移动视频流传输,其分别也被称为 HR(高分辨率)场景和 LR(低分辨率)场景,对视频质量评估模型工作。两种场景之间的差别的范围从视频内容的时空分辨率和编码配置到传输协议和观看条件。

[0072] P. NBAMS VQM(视频质量模型)的输入是具有全部传输分组报头(UDP/IP/RTP 或 UDP/IP/RTP/TS)的编码的视频比特流。输出是客观 MOS 分数。P. NBAMS 工作的主要目标应用是监视机顶盒(STB)或网关中的视频质量。P. NBAMS 模式 1 模型仅使用比特流信息,并且模式 2 模型可以解码部分或全部视频序列,并且像素信息除了用于解析比特流信息之外,还用于视觉质量预测,以便改进预测准确度。

[0073] 参考图 6,示出可以应用上述特征和原理的视频传输系统或装置 600。处理器 605 处理视频并且编码器 610 编码视频。从编码器生成的比特流通过分发网络 620 传送到解码器 630。例如装置 500 的视频质量监视器或视频质量测量装置可以用在不同阶段。

[0074] 在一个实施例中,视频质量监视器 640 可以由内容创建器使用。例如,在决定诸如模式决定或比特速率分配的编码参数时,可以由编码器使用估计的视频质量。在另一示例中,在编码视频之后,内容创建器使用视频质量监视器来监视编码的视频的质量。如果质量度量不满足预定义的质量水平,则内容创建器可以选择重新编码视频以改进视频质量。内容创建器还可以基于质量对编码的视频进行评级并且相应地对内容进行收费。

[0075] 在另一实施例中,视频质量监视器 650 可以由内容分发器使用。视频质量监视器可以被放置在分发网络中。视频质量监视器计算质量度量并且将它们报告给内容分发器。基于来自视频质量监视器的反馈,内容分发器可以通过调整带宽分配和访问控制来改进它的服务。

[0076] 内容分发器还可以向内容创建器发送反馈以调整编码。注意,改进编码器处的编码质量可能不一定改进解码器侧的质量,因为高质量编码的视频通常要求更多带宽并且留

下更少带宽用于传输保护。因此,为了实现解码器处的最佳质量,应考虑编码比特速率和用于信道保护的带宽之间的平衡。

[0077] 在另一实施例中,视频质量监视器 660 可以由用户设备使用。例如,当用户设备搜索因特网中的视频时,搜索结果可以返回许多视频或到与请求的视频内容对应的视频的许多链接。搜索结果中的视频可以具有不同质量水平。视频质量监视器可以计算这些视频的质量度量并且决定选择存储哪个视频。在另一示例中,解码器关于不同错误隐藏模式估计隐藏的视频的质量。基于估计,可以由解码器选择提供更好的隐藏质量的错误隐藏。

[0078] 这里描述的实现方式可以被实现为例如方法或过程、装置、软件程序、数据流或信号。即使仅在单一形式的实现方式的上下文中讨论(例如,仅作为方法讨论),讨论的特征的实现方式也可以被实现为其他形式(例如,装置或程序)。装置可以被实现为例如适当的硬件、软件和固件。方法可以被实现为例如装置,该装置例如是总地指代处理设备的处理器,其包括例如计算机、微处理器、集成电路或可编程逻辑设备。处理器还包括通信设备、例如计算机、蜂窝电话、便携式/个人数字助理(“PDA”),以及有助于终端用户之间的信息通信的其他设备。

[0079] 提及本原理的“一个实施例”或“一实施例”或“一个实现方式”或“一实现方式”以及它们的其他变型指的是结合实施例描述的特定特征、结构、特性等被包括在本原理的至少一个实施例中。因此,在整个说明书的各个地方出现的短语“在一个实施例中”或“在一实施例中”或“在一个实现方式中”或“在一实现方式中”以及任何其他变型的表述不一定全都指代相同实施例。

[0080] 此外,本申请或它的权利要求可能提及“确定”各种信息片段。确定信息可以包括例如估计信息、计算信息、预测信息或从存储器取得信息中的一个或多个。

[0081] 此外,本申请或它的权利要求可能提及“访问”各种信息片段。访问信息可以包括例如接收信息、取得信息(例如,从存储器)、存储信息、处理信息、传送信息、移动信息、复制信息、擦除信息、计算信息、确定信息、预测信息或估计信息中的一个或多个。

[0082] 此外,本申请或它的权利要求可能提及“接收”各种信息片段。与“访问”类似,接收意图作为广义术语。接收信息可以包括例如访问信息或取得信息(例如,从存储器)中的一个或多个。此外,“接收”通常以某种方式在例如存储信息、处理信息、传送信息、移动信息、复制信息、擦除信息、计算信息、确定信息、预测信息或估计信息的操作期间涉及。

[0083] 本领域技术人员将清楚,实现方式可以产生被格式化为承载可以例如存储或传送的信息的各种信号。信息可以包括例如用于执行方法的指令或由所述实现方式中的一个产生的数据。例如,信号可以被格式化为承载所述实施例的比特流。这种信号可以被格式为例如电磁波(例如,使用频谱的射频部分)或基带信号。格式化可以包括例如编码数据流和利用编码的数据流调制载波。信号承载的信息可以是例如模拟或数字信息。已知信号可以在各种不同的有线或无线链路上传送。信号可以被存储在处理器可读介质上。

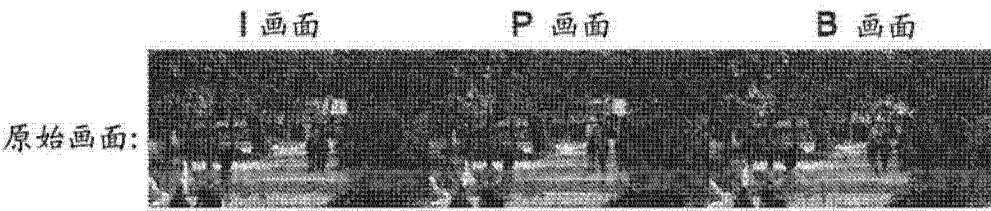


图 1A



图 1B



图 1C

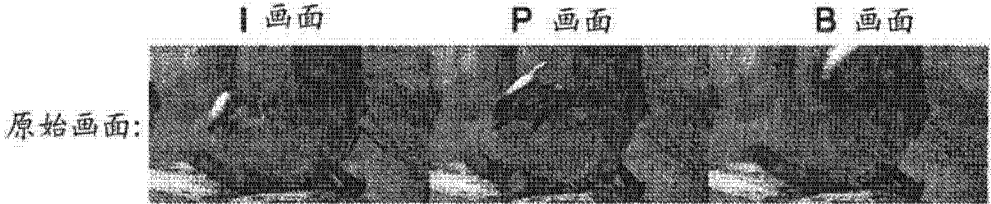


图 2A



图 2B

对准的
CU映射:

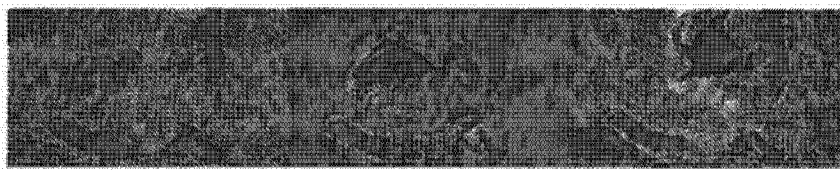


图 2C

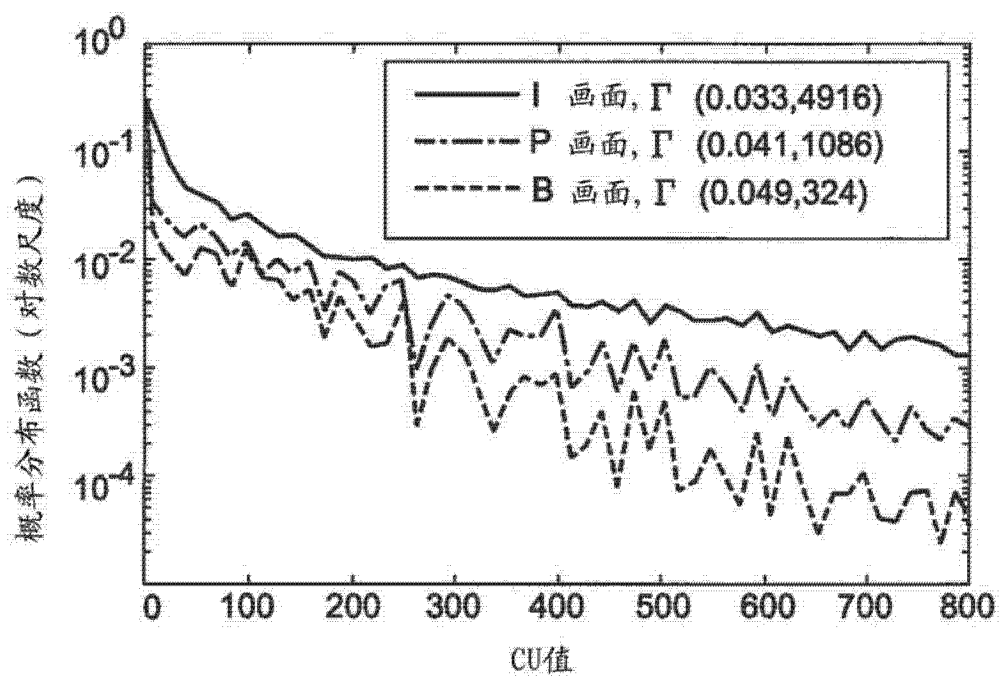


图 3

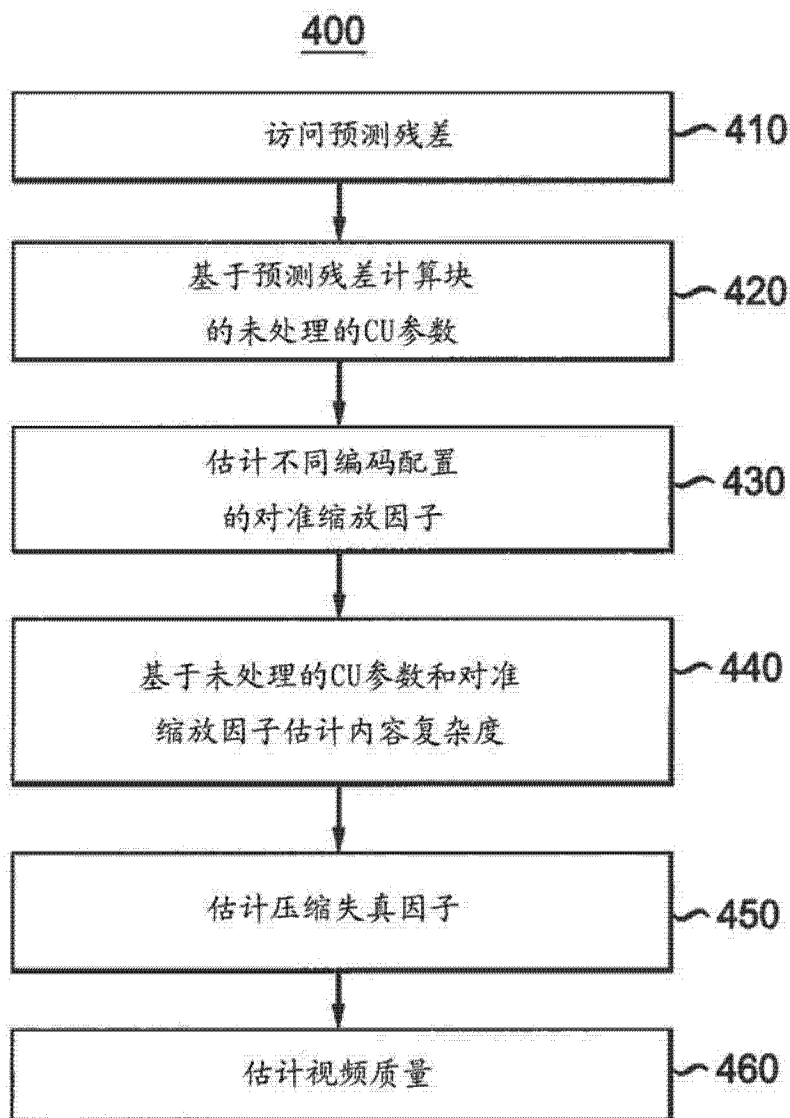


图 4

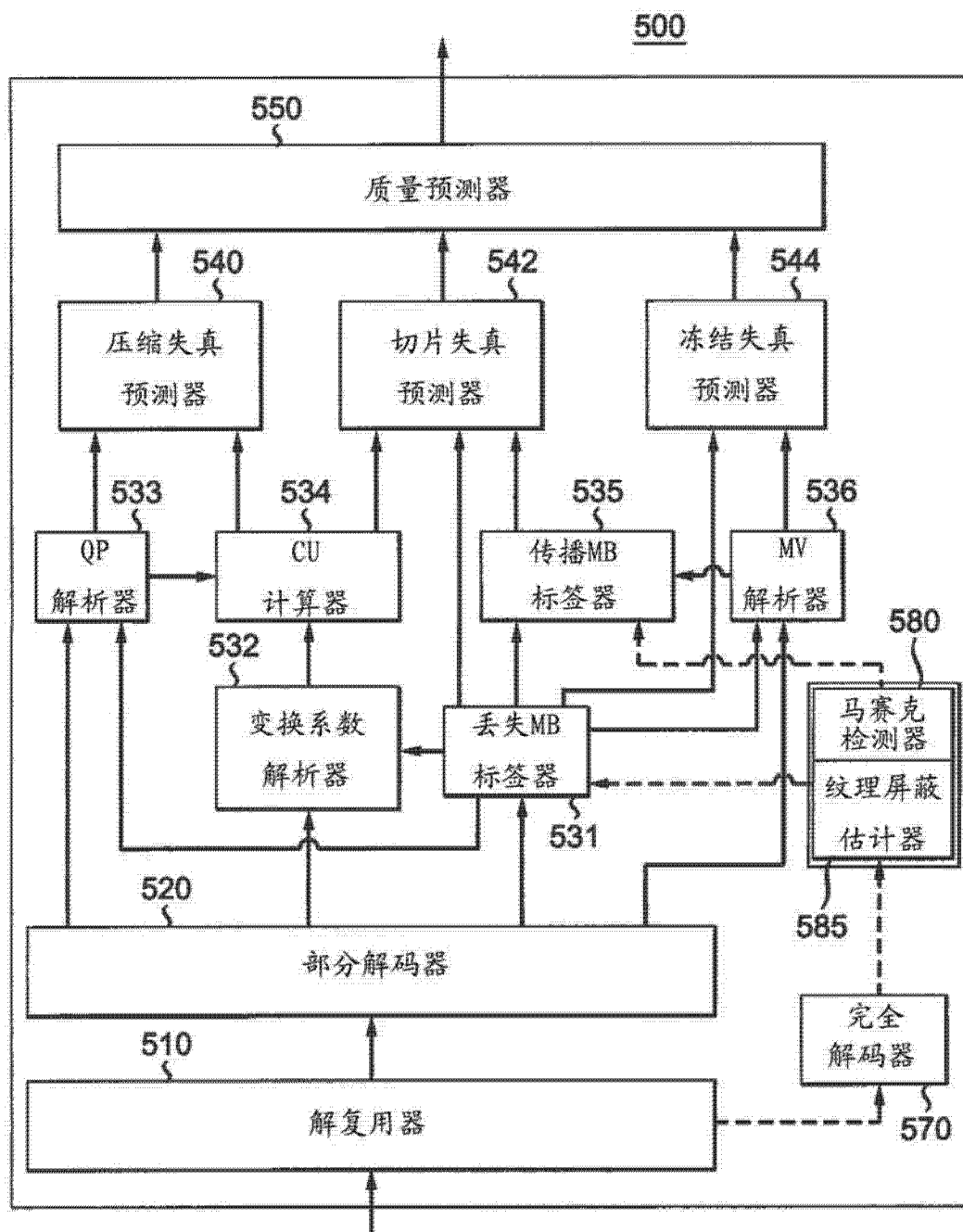


图 5

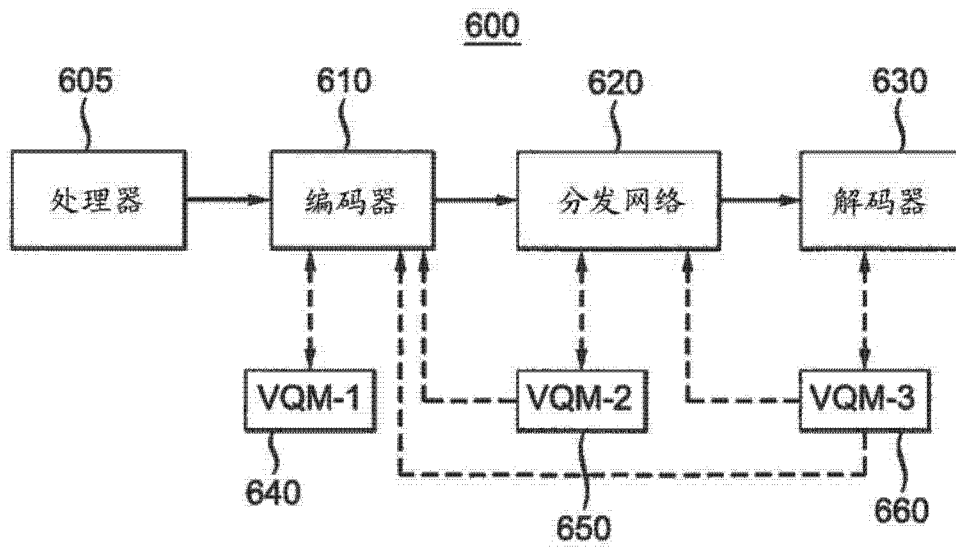


图 6