



(12) 发明专利

(10) 授权公告号 CN 101835048 B

(45) 授权公告日 2012. 09. 26

(21) 申请号 201010174145. 5

审查员 张伯约

(22) 申请日 2010. 04. 30

(30) 优先权数据

12/750, 401 2010. 03. 30 US

(73) 专利权人 香港应用科技研究院有限公司

地址 中国香港新界沙田香港科学园科技大道西二号生物资讯中心三楼

(72) 发明人 颜庆义 马林 湛伟权 刘雨

(74) 专利代理机构 深圳新创友知识产权代理有限公司 44223

代理人 江耀纯

(51) Int. Cl.

H04N 7/26 (2006. 01)

H04N 7/50 (2006. 01)

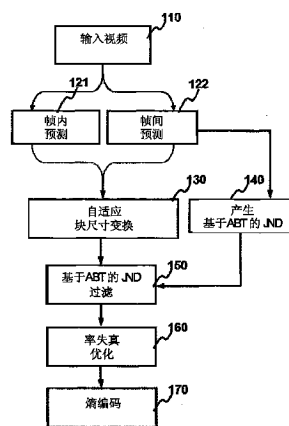
权利要求书 2 页 说明书 11 页 附图 5 页

(54) 发明名称

通过基于自适应块尺寸变换 ABT 的最小可觉差 JND 模型进行视频编码的方法和装置

(57) 摘要

本发明涉及通过基于自适应块尺寸变换 (ABT) 的最小可觉差 (JND) 模型对视频编码的方法和装置。为了建立最小可觉差模型, 使用空间内容信息 (SCI) 来表示宏块及其子块之间的空间外观相似度, 并且使用运动特性距离 (MCD) 来表示宏块及其子块之间的运动特性相似度。对帧内编码帧, 基于宏块 SCI 的平衡策略被用来产生基于 ABT 的 JND 模型。对帧间编码帧, 基于宏块 SCI 和 MCD 的平衡策略被用来产生基于 ABT 的 JND 模型。使用基于 ABT 的 JND 模型, 帧内每个块的残差系数被滤波以获得一组缩减的残差系数用于传输, 但不会降低视觉质量。



1. 一种视频编码装置,包括:

一个提供视频信号的单元;

一个预测模式确定单元,其依照帧预测模式确定视频信号里的帧是帧间编码帧还是帧内编码帧;

一个基于自适应块尺寸变换ABT的最小可觉差JND模型产生单元,其产生一个基于ABT的JND模型,该模型包括由基于视觉感受模型的心理物理实验确定的一个或多个JND阈值;

和一个基于ABT的JND滤波器,其通过去除一个或多个低于JND阈值的残差系数并从一个或多个大于JND阈值的残差系数中减去JND阈值来滤波一个或多个残差系数,从而输出一个含有被滤波的残差系数的信号。

2. 根据权利要求1所述的视频编码装置,其中:

基于空间内容信息,基于ABT的JND模型产生单元为帧内编码帧计算基于ABT的JND模型。

3. 根据权利要求2所述的视频编码装置,其中:

如果块及其子块之间的空间内容信息一致,就选择一个较大块尺寸的基于ABT的JND模型。

4. 根据权利要求2所述的视频编码装置,其中:

空间内容信息是由块类型确定,而块类型是由帧的每个块里的边缘像素数量决定的。

5. 根据权利要求1所述的视频编码装置,其中:

基于空间内容信息和时间特性,基于ABT的JND模型产生单元为帧间编码帧计算基于ABT的JND模型,时间特性是由块及其子块之间的运动特性距离确定,运动特性距离由以下等式确定:

$$MCD = \sqrt{\sum_{i=0}^{N-1} [(MV_{s_i_x} - MV_{M_x})^2 + (MV_{s_i_y} - MV_{M_y})^2]} / N, \text{ 其中 } MCD \text{ 是运动特性距离, } MV_{M_x} \text{ 和 } MV_{M_y} \text{ 分别是宏块的水平方向和垂直方向的运动矢量, } MV_{s_i_x} \text{ 和 } MV_{s_i_y} \text{ 分别是子块水平方向和垂直方向的运动矢量, } i \text{ 表示子块的编号, } N \text{ 代表一个宏块里面包含的子块的数量。}$$

6. 根据权利要求5所述的视频编码装置,其中:

如果空间内容信息一致并且时间特性一致,就选择一个较大块尺寸的基于ABT的JND模型。

7. 根据权利要求5所述的视频编码装置,其中:

空间内容信息是由块类型确定,而块类型是由帧的每个块里的边缘像素数量决定。

8. 一种对视频信号进行编码的方法,包括:

使用权利要求1所述的视频编码装置,处理视频信号。

9. 一种对视频信号进行编码的方法,包括:

提供一个将被编码的视频信号;

依照帧预测模式确定视频信号里的帧是帧间编码帧还是帧内编码帧;

产生一个基于自适应块尺寸变换ABT的最小可觉差JND模型,其包括由基于视觉感受模型的心理物理实验确定的一个或多个JND阈值;和

通过去除一个或多个低于 JND 阈值的残差系数并从一个或多个大于 JND 阈值的残差系数减去 JND 阈值,滤波一个或多个残差系数,从而输出一个具有被滤波的残差系数的信号。

10. 根据权利要求 9 所述的视频信号编码方法,还包括:

基于空间内容信息,计算帧内编码帧的基于 ABT 的 JND 模型。

11. 根据权利要求 10 所述的视频信号编码方法,其中:

如果块及其子块之间的空间内容信息一致,就选择一个较大块尺寸的基于 ABT 的 JND 模型。

12. 根据权利要求 10 所述的视频信号编码方法,其中:

空间内容信息是由块类型确定的,而块类型是由帧的每个块里的边缘像素数量决定。

13. 根据权利要求 9 所述的视频信号编码方法,还包括:

基于空间内容信息和时间特性,计算帧间编码帧的基于 ABT 的 JND 模型,时间特性是由块及其子块之间的运动特性距离确定,运动特性距离由以下等式确定:

$$MCD = \sqrt{\sum_{i=0}^{N-1} [(MV_{s_i_x} - MV_{M_x})^2 + (MV_{s_i_y} - MV_{M_y})^2]} / N,$$

其中 MCD 是运动特性距离, MV_{M_x} 和 MV_{M_y} 分别是宏块的水平和垂直方向的运动矢量, $MV_{s_i_x}$ 和 $MV_{s_i_y}$ 分别是子块水平和垂直方向的运动矢量, i 表示子块的编号, N 代表一个宏块里面包含的子块的数量。

14. 根据权利要求 13 所述的视频信号编码方法,其中:

如果空间内容信息一致并且时间特性一致,就选择一个较大块尺寸的基于 ABT 的 JND 模型。

15. 根据权利要求 13 所述的视频信号编码方法,其中:

空间内容信息是由块类型确定的,而块类型是由帧的每个块里的边缘像素数量决定。

通过基于自适应块尺寸变换 ABT 的最小可觉差 JND 模型进行视频编码的方法和装置

【技术领域】

【0001】 本发明通常涉及数字视频处理,特别涉及包括编码和解码的数字视频编码。

【背景技术】

【0002】 由于摄像装置的进步,无论何时何地都能够产生越来越多的视频(图像序列,也被称为运动图像)。此外,为了追求更高的分辨率,视频尺寸不断增加,从而涉及到大量数据需要处理。这给诸如显示、传输和存储视频的过程带来了巨大挑战。所以,需要一种视频编码技术,使得视频在不牺牲其质量的情况下能够被压缩。

【0003】 至于质量,期望发生较少的失真,传统上失真通常是由量化度量如均方差(MSE)和峰值信噪比(PSNR)来衡量的。但是,这些量化度量并不能反映观看者的感受以及失真在观看者眼中的严重程度。因此,通过考虑人类如何视觉感受事物的特性,在不带来额外视觉失真的前提下,需要进一步提高压缩比。换言之,人类如何视觉感受图像的主观试验非常重要。所以,需要保持用户的视觉感受,同时更有效地进行视频处理。

【0004】 此外,按照视频标准的发展,越来越多的特征被结合到当前和未来的视频编准里。其中一个特征是允许视频编码过程里有各种块尺寸,例如,自适应块尺寸变换(ABT)。通过考虑图像内容,这有助于改善图像质量。所以,需要在视频编码时能够进行ABT,例如,与视频标准如H.264/AVC兼容。

【发明概述】

【0006】 基于空间内容信息和运动特性距离,本发明产生基于ABT的最小可觉差(JND)模型,其通过残差滤波方法可以应用于视频编码系统。JND模型也被称为JND方案。由于JND模型能够以矩阵格式表示,JND模型也被称为JND图。

【0007】 本发明其中一个优点是通过使用基于ABT的JND模型进行视频编码能够节省高达50%的比特率,而不会降低主观视觉质量。本发明使用人类如何视觉感受事物的特性以降低比特率,而不会带来可觉察的视觉失真。一个模拟人类视觉感受事物过程的系统被称为视觉感受模型(VPM)。一个视觉感受模型的例子是最小可觉差(JND)模型。建立基于ABT的JND模型,并提供JND阈值给帧内的每个块,如果非零残差系数的数值小于JND阈值,非零残差系数被滤波。如果非零残差系数的数值大于相应的JND阈值,将从非零残差系数中减去相应的JND阈值。

【0008】 就本发明JND模型的特性而言,不同于图像域内的现有JND模型,或频率域内产生的基于固定块尺寸DCT的JND模型,本发明采用基于不同块尺寸的基于ABT的JND模型。另外,基于ABT的JND模型在不同块尺寸变换时考虑人类如何视觉感受事物的特性。基于ABT的JND模型是根据空间内容信息(SCI)和运动特性距离(MCD)而确定的。

【0009】 就本发明JND模型的应用而言,不同于现有JND模型用于图像/视频水印和质量评估,本发明使用JND模型以增强视频编码效率并降低比特率。

【0010】 此外,依照一个预设的图像质量度量,本发明使用基于ABT的JND模型进行视频编

码,而不需要调整量化矩阵。

[0011] 本发明的一个方面是将固定尺寸的 JND 模型扩展到不同块尺寸。不同块尺寸的 JND 模型被用来描述人类对不同尺寸变换如何视觉感受事物的特性。

[0012] 本发明的另一个方面是确定空间内容信息 (SCI),其描述宏块及其子块之间的空间外观相似性。

[0013] 本发明的另一个方面是确定运动特性距离 (MCD),其描述宏块及其子块之间的运动特性相似性。

[0014] 本发明的另一个方面是基于宏块的 SCI 使用图像或帧内的不同块尺寸变换之间的平衡策略。

[0015] 本发明的另一个方面是基于宏块的 SCI 和 MCD 使用帧间的不同块尺寸变换之间的平衡策略。

[0016] 本发明的另一个方面是对不同尺寸变换进行残差滤波以降低比特率并保持相同的视觉质量。

[0017] 本发明的其它方面也将披露,如以下所述的实施例。

【附图说明】

[0018] 以下将参照附图,详细描述本发明的目的、方面和实施例,其中:

[0019] 图 1 显示本发明一个视频编码的典型实施例的流程图;

[0020] 图 2 显示一个产生基于 ABT 的 JND 模型的典型实施例的流程图;

[0021] 图 3 显示一个基于 ABT 的 JND 滤波的典型实施例的流程图;

[0022] 图 4 显示在一个实施例里分别在一个 16x16 块和一个 4x4 块内分配测试 DCT 子带。

[0023] 图 5 描述宏块及其相应子块的运动矢量以解释 MCD 一致性概念用于帧间的时间平衡策略。

[0024] 图 6 描述一个视频编码装置的典型实施例的模块示意图。

【发明详述】

[0026] 图 1 显示本发明提供的一个视频编码典型实施例的流程图。在一个实施例里,视频编码实在一个输入视频 110 上进行,输入视频 110 被编码。为了编码输入视频 110,计算输入视频 110 里每帧内不同块尺寸的每个块的 JND 模型。可能的块尺寸的例子包括但不限于 4x4、8x8 和 16x16。

[0027] 对每个宏块,基于帧内预测 121 或帧间预测 122,计算预测值。例如,当帧是 P 帧或 B 帧时,使用帧间预测 122。帧内预测 121 是基于不同块尺寸如 4x4、8x8 和 16x16,并也被称为不同块尺寸的帧内预测。帧间预测 122 是基于不同块尺寸如 4x4、8x8 和 16x16,并也被称为不同块尺寸的帧间预测。

[0028] 差值块 (difference block) 是由预测和块之间的差异而产生。差值块也被称为残差 (residual) 块。根据残差块的块尺寸,残差块被分割成不同构造的非重叠块用于进行编码。同时计算以下信息:

[0029] (1) 块类型,如帧间或帧内;

[0030] (2) 块分类,如平面、边缘或纹理;和

[0031] (3) 不同尺寸的块的运动矢量。

[0032] 以上信息被用来计算最小可觉差 (Just-Noticeable Difference, JND) 模型 140。依照其中一个实施例,根据 Jia, Y. ;Lin, W. ;Kassim, A. A. 所著的 Estimating Just-Noticeable Distortion for Video (IEEE Transactions on Circuits and Systems for Video Technology, Vol. 16, No. 7, pp. 820-829, Jul. 2006) 一文中的 Jia 的模型,在此其披露通过引用结合到本发明,假设 JND 模型 140 最初是专为 8x8 离散余弦变换 (DCT) 块开发的。8x8DCT 块的 JND 模型 140 可以被扩展到以下所述的其它块尺寸。

[0033] 以 X 表示将被编码的块, X_p 是 X 的预测。如以下等式 (1) 所示,残差块 X_r 被变换、量化和编码。

$$[0034] \quad X_r = X - X_p \quad (1)$$

[0035] 残差块 X_r 通过自适应块尺寸变换 (ABT) 130 进行变换,也被称为不同块尺寸变换,如通过不同块尺寸离散余弦变换 (DCT) 以计算被变换的残差块 Y 。被变换的残差块 Y 在以下等式 (2) 定义:

$$[0036] \quad Y = T(X_r) \quad (2)$$

[0037] 其中 T 表示在残差块 X_r 上的 ABT。被量化的变换的残差块 Y_q 在以下等式 (3) 定义:

$$[0038] \quad Y_q = Q(Y) \quad (3)$$

[0039] 其中 Q 表示量化运算。

[0040] 对块 X , 计算基于 ABT 的 JND 模型 140。基于 ABT 的 JND 模型 140, 由 $J_x(u, v)$ 表示, 是一个包括 JND 阈值的矩阵或图, 超过该 JND 阈值, 观察者就能注意到每个残差系数的任何失真。JND 阈值的计算将在以下讨论。

[0041] 以 $Y(u, v)$ 表示残差块 X_r 的第 (u, v) 位置的残差系数。如果 $Y(u, v)$ 的幅值小于 $J_x(u, v)$, 则可以丢弃变换的残差系数 $Y(u, v)$ 。如果 $Y(u, v)$ 的幅值大于 $J_x(u, v)$, 变换的残差系数 $Y(u, v)$ 减去 $J_x(u, v)$ 。变换的残差系数 $Y(u, v)$ 的处理也被称为一个滤波步骤 150 (基于 ABT 的 JND 滤波)。在一个实施例里, 基于 ABT 的 JND 滤波步骤 150 在量化之前实施, 并且滤波计算如以下等式 (4) 定义:

$$[0042] \quad Y_j(u, v) = \begin{cases} \text{sign}(Y(u, v)) \times (|Y(u, v)| - J_x(u, v)) & \text{if } |Y(u, v)| > J_x(u, v), \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

[0043] 其中 Y_j 是 JND 滤波变换块。Sign 函数确保在滤波之后 $Y(u, v)$ 仍然是负或正。由于残差系数被滤波, 通过以下等式 (5) 计算重建块 X_j' :

$$[0044] \quad X_j' = X_p + T^1(Q^{-1}(Y_{j,q})) \quad (5)$$

[0045] 其中 $Y_{j,q}$ 是被量化的 Y_j 。

[0046] 在重建块 X_j' 和块 X 之间的失真 d 由以下等式 (6) 定义:

$$[0047] \quad d = \|X_j' - X\|_1 \quad (6)$$

[0048] 其中失真 d 等于重建块和初始块之间的绝对差总和 (SAD)。其它差值度量, 诸如但不限于平方差总和 (SSD) 或绝对变换差总和 (SATD), 也可以被用于率失真优化 (RDO) 160。

[0049] 以 E 表示初始块和重建块之间的 DCT 变换差, 并且从以下等式 (7) 计算 E :

$$[0050] \quad E = T(X_j' - X) \quad (7)$$

[0051] E 由以上等式 (4) 的 J_x 滤波, 由以下等式 (8) 计算滤波:

$$[0052] \quad E_j(u, v) = \begin{cases} \text{sign}(E(u, v)) \times (|E(u, v)| - J_x(u, v)) & \text{if } |E(u, v)| > J_x(u, v), \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

[0053] 对“滤波”，进行以下操作：如果 $E(u, v)$ 的幅值小于 $J_x(u, v)$ ，可以丢弃 DCT 变换差 $E(u, v)$ 。如果 $E(u, v)$ 的幅值大于 $J_x(u, v)$ ，则 DCT 变换差 $E(u, v)$ 减去 $J_x(u, v)$ 。然后，如下等式 (9) 所示，从 E_j 的逆变换，计算失真 d_j ，用于率失真优化 160：

$$[0054] \quad d_j = ||T^{-1}(E_j)||_1 \quad (9)$$

[0055] 如果能够进行率失真优化 160，选择的块的预测模式是最小化 RD（率失真）成本。例如，RD 成本是由以下等式 (10) 定义：

$$[0056] \quad R = d + \lambda L \quad (10)$$

[0057] 其中 λ 是 Lagrangian 乘子，而 L 是利用预测模式对块进行编码的实际比特长度。例如，如果帧间预测的 RD 成本低于帧内预测，那么就选择帧间预测作为该块的预测模式。

[0058] 基于 ABT 的 JND 滤波步骤 150 可以看作是一个自适应量化步骤，其中每个块的量化装置是由其空间和时间特性确定的。因为基于 ABT 的 JND 滤波的块将经历正常的编码程序，比特流完全与 H.264 标准兼容，并且能够通过任何解码器准确地进行解码。

[0059] 选择的变换模式与被滤波的残差系数一起，通过熵编码 170 进行编码。在进行不同块尺寸变换和基于 ABT 的 JND 滤波之后，计算宏块的总失真及其所需比特。根据宏块的总失真及其所需比特，依照率失真优化而确定最佳变换模式（最佳变换尺寸）。基于 ABT 的 JND 模型 140 与变换模式有关，所以不需要编码额外信息。

[0060] 图 2 显示一个典型实施例的产生基于 ABT 的 JND 模型的流程图。为了产生基于 ABT 的 JND 模型，需要确定块类型 210。如果是一个帧间块，一个根据空间和时间特性的平衡策略是合适的。如果是一个帧内块，仅空间平衡策略就合适了。在一个实施例里，使用 8×8 基于 DCT 的 JND 模型，并可以扩展到不同块尺寸的基于 DCT 的 JND 模型，如 16×16 基于 DCT 的 JND 模型和 4×4 基于 DCT 的 JND 模型。图像的自适应的基于块尺寸变换 (ABT) 的最小可觉差 (JND) 模型的详情在 Ma, Lin ; Ngan, King N. 所著的 Adaptive Block-Size Transform Based Just-Noticeable Difference Profile for Images (Proceedings of the 10th Pacific Rim Conference on Multimedia: Advances in Multimedia Information Processing, Session 13: Advanced Image Processing Techniques ; Page 1208-1218 ; 2009) 一文里有描述，在此其披露通过引用结合到本发明。

[0061] 在图像域或变换域内可以确定一个自动的 JND 模型，如 DCT 和离散小波变换 (DWT) 或这两种方案的组合。为了准确地描述人类如何视觉感受事物，在变换域内产生的一个 JND 模型，其也被称为基于子带 (subband-based) 的 JND，可以集合所有主要的影响因子。主要影响因子包括但不限于空间对比敏感度函数 (CSF)、亮度适应效应和对比度掩盖 (contrast masking)。

[0062] ABT 允许不同的变换尺寸。依照普通策略采用不同的变换尺寸，即较大变换会提供较好的能量压缩和较好的细节保留，而较小变换可以避免压缩期间的振铃现象。如果使用不同的变换尺寸，本发明允许基于 ABT 的 JND 模型使用人类如何视觉感受事物的不同表现。在扩展期间，进行心理物理实验来参数化空间 CSF 以考虑人类如何视觉感受事物的特性。图像内容也被自适应地调整以确定采用哪一个变换尺寸来产生基于 ABT 的 JND 模型。

[0063] 对一个帧内块，使用空间平衡策略 220 来确定空间内容信息 (SCI) 在块内是否一

致。根据空间平衡策略 220, 如果一个 16×16 块的空间内容信息一致, DCT 的块尺寸被选择为 16×16 。 16×16 DCT 被用来产生基于 ABT 的 JND 模型。根据基于 16×16 块尺寸的心理物理实验 221, 产生的基于 ABT 的 JND 模型是基于 16×16 的块尺寸。如果 16×16 块的空间内容信息不一致, 那么 16×16 块被分割成 8×8 块。随后, 对每个 8×8 块, 根据空间平衡策略 222, 如果 8×8 块的空间内容信息一致, 就选择 DCT 的块尺寸为 8×8 。 8×8 DCT 被用来产生基于 ABT 的 JND 模型。产生的基于 ABT 的 JND 模型是基于 8×8 的块尺寸。如果 8×8 块的空间内容信息不一致, 就选择 DCT 的块尺寸为 4×4 , 4×4 DCT 被用来产生基于 ABT 的 JND 模型。根据基于 4×4 块尺寸的心理物理实验 223, 产生的基于 ABT 的 JND 模型是基于 4×4 的块尺寸。

[0064] 对一个帧间块, 使用空间和时间平衡策略 230 来确定 16×16 块的空间内容信息 (SCI) 是否一致, 以及 16×16 块的运动特性距离 (MCD) 是否一致。依照空间和时间平衡策略 230, 如果 SCI 和 MCD 一致, 就选择 DCT 的块尺寸为 16×16 。 16×16 DCT 被用来产生基于 ABT 的 JND 模型。根据基于 16×16 块尺寸的心理物理实验 231, 产生的基于 ABT 的 JND 模型是基于 16×16 的块尺寸。如果 SCI 和 MCD 不一致, 16×16 块被分割成 8×8 区块。随后, 对每个 8×8 块, 使用空间和时间平衡策略 232 来确定 8×8 块的 SCI 是否一致以及 8×8 块的 MCD 是否一致。依照空间和时间平衡策略 232, 如果 SCI 和 MCD 一致, 就选择 DCT 的块尺寸为 8×8 。 8×8 DCT 被用来产生基于 ABT 的 JND 模型。如果 SCI 和 MCD 不一致, 就选择 DCT 的块尺寸为 4×4 。 4×4 DCT 被用来产生基于 ABT 的 JND 模型。根据基于 4×4 块尺寸的心理物理实验 233, 产生的基于 ABT 的 JND 模型是基于 4×4 的块尺寸。

[0065] 空间平衡策略

[0066] 在心理物理学中, 最小可觉差 (JND) 是某一特定感官刺激所能察觉的最小差别。JND 也被看作是差别阈限或差分阈值。通常, JND 模型能够以以下等式 (11) 表述:

$$T_{JND} = T_{JND_s} \times T_{JND_t} \quad (11)$$

[0068] 其中 T_{JND_s} 是空间 JND, 而 T_{JND_t} 是时间 JND 调制因子。在帧间帧, 对于处理不适合运动估计的块, 也需要帧内预测。因为这些块依赖于其空间邻近块多于时间帧。因此, 这些块被看作仅包含空间信息, 并通过遵循空间平衡策略的空间 JND 模型进行模拟。

[0069] 在 DCT 域里的空间 JND 模型是由以下等式 (12) 确定:

$$T(m, n, i, j) = T_{\text{basic}}(i, j) \cdot \alpha_{\text{lum}}(m, n) \cdot \alpha_{\text{cm}}(m, n, i, j) \quad (12)$$

[0071] 其中 (m, n) 表示图像里 DCT 块的位置, (i, j) 表示 DCT 子带 (i, j) , T_{basic} 是由空间对比敏感度函数 (CSF) 产生的基本可视度阈值; α_{lum} 是亮度适应性; α_{cm} 是对比度掩盖; 而 $T(m, n, i, j)$ 是获得的空间 JND 阈值。

[0072] 基本可视度阈值

[0073] 基本可视度阈值 T_{basic} 被定为空间频率的对比敏感度的倒数, 其是在空间频率上人类视觉感受的敏感度, 并通过以下等式 (13) 进行模拟:

$$H(\omega) = (a + b\omega) \cdot \exp(-c\omega) \quad (13)$$

[0075] 其中 ω 是指定的空间频率。由于 JND 阈值是由等式 (13) 模拟的敏感度的倒数, 基本 JND 阈值由以下等式 (14) 模拟:

[0076]

$$T_{basic}(m, n, i, j) = \frac{s}{\phi_i \phi_j} \cdot \frac{\exp(c\omega_{i,j})/(a + b\omega_{i,j})}{\gamma + (1 - \gamma) \cdot \cos^2 \varphi_{ij}} \quad (14)$$

[0077] 其中 $s = 0.25$ 表示求和影响因子, 并且 ϕ_i, ϕ_j 是如以下等式 (15) 所定义的 DCT 正态因子:

$$\phi_m = \begin{cases} \sqrt{1/N}, & m = 0 \\ \sqrt{2/N}, & m > 0 \end{cases} \quad (15)$$

[0079] N 表示 DCT 块的尺寸。 φ_{ij} 表示相应 DCT 系数的方向角度, 如以下等式 (16) 所示:

$$\varphi_{ij} = \arcsin(2\omega_{i,0}\omega_{0,j}/\omega_{i,j}^2) \quad (16)$$

[0081] 其中 $\omega_{i,j}$ 是相应的 DCT 子带 (i, j) 的空间频率, 由以下等式 (17) 定义:

$$\omega_{i,j} = (1/(2N))\sqrt{(i/\theta_x)^2 + (j/\theta_y)^2} \quad (17)$$

[0083] 其中 θ_x 和 θ_y 分别是水平和垂直视觉角度, 由以下等式 (18) 提供:

$$\theta_x = \theta_y = 2 \cdot \arctan(1/(2 \cdot R_d \cdot P_h)) \quad (18)$$

[0085] 其中 R_d 表示观看距离与图像高度的比率, 而 P_h 是图像高度 (像素)。参数 (a, b, c) 在以下段落进行定义和计算。

[0086] 亮度适应性

[0087] 亮度适应性 α_{lum} , 被用来表示对比不同亮度有多容易地察觉到差别。例如, 在黑暗或明亮区域内有一个较高的可视度阈值, 以表示要察觉到任何差别会更困难些, 而在中等亮度区域内有一个较低的可视度阈值, 以表示更容易察觉到任和差别。亮度适应性是基于 DCT 块的平均强度 I_{ave} , 由以下等式 (19) 给定:

$$\alpha_{lum} = \begin{cases} (60 - I_{ave})/150 + 1 & I_{ave} \leq 60 \\ 1 & 60 < I_{ave} < 170 \\ (I_{ave} - 170)/425 + 1 & I_{ave} \geq 170 \end{cases} \quad (19)$$

[0089] 其中 I_{ave} 是整个 DCT 块的平均强度值。

[0090] 对比度掩盖

[0091] 对比度掩盖 α_{cm} 被用来表示在出现另一个视觉分量时一个视觉分量可视度的降低。例如, 在一个纹理区域内将较难见到噪声, 而在一个光滑区域或边缘区域内将更容易见到噪声。因此, α_{cm} 依块分类的不同而不同, 由以下等式 (20) 给定:

$$\alpha_{cm} = \begin{cases} \Psi, & \text{for } (i^2 + j^2) \leq 16 \quad \text{平面和边缘块} \\ \Psi \cdot \min(4, \max(1, (\frac{C(m, n, i, j)}{T_{basic}(m, n, i, j) \times \alpha_{lum}(m, n, i, j)})^{0.36}))), & \text{其他} \end{cases} \quad (20)$$

[0093]

$$\text{其中 } \Psi = \begin{cases} 1, & \text{平面和边缘块} \\ 2.25, & \text{for } (i^2 + j^2) \leq 16 \quad \text{纹理块} \\ 1.25, & \text{for } (i^2 + j^2) > 16 \quad \text{纹理块} \end{cases}$$

[0094] 在一个实施例里, 块分类是在图像域内实施的。首先, 在图像域内使用 Canny 边缘探测器标记边缘像素。其次, 基于块内的边缘像素数目, 块可以被分类成三种类型, 即分别

是 PLANE (平面)、EDGE (边缘)、和 TEXTURE (纹理)。例如, 对一个 16×16 宏块, 依照以下等式 (21) 定义块分类:

$$[0095] \quad Block_Categorization = \begin{cases} PLANE & \sum_{Edge} < 16 \\ EDGE & 16 \leq \sum_{Edge} \leq 52 \\ TEXTURE & \sum_{Edge} > 52 \end{cases} \quad (21)$$

[0096] 其中 \sum_{Edge} 是在一个给定宏块里边缘像素的数目。

[0097] 在一个实施例里, 时间 JND 模型由以下等式 (20) 确定:

[0098]

$$T_{JND_T} = \begin{cases} 1, & f_s < 5cpd \ \& \ f_t < 10Hz \\ 1.07^{(f_t-10)}, & f_s < 5cpd \ \& \ f_t \geq 10Hz_t \\ 1.07^{f_t}, & f_t > 5cpd_s \end{cases} \quad (22)$$

[0099] 其中 $f_t = f_{sx}v_x + f_{sy}v_y$, f_{sx} , f_{sy} 分别是水平空间频率和垂直空间频率, 而 v_x , v_y 分别是沿着水平方向和沿着垂直方向在视网膜平面上物体运动的速度。

[0100] 不同块尺寸变换之间的 JND 方案的平衡策略

[0101] 在扩展 8×8 到 16×16 以及 4×4 之后, 将获得三个基于不同块尺寸变换的 JND 方案。较小块尺寸的 DCT 将产生较好的能量压缩和细节信息保留, 而较大块尺寸的 DCT 能够更有效地表示局部图像块的内容和特性。

[0102] 帧内的空间平衡策略

[0103] 表格 1 (在 16×16 、 8×8 和 4×4 JND 方案中间的平衡策略)

[0104]

16x16 宏块类型	8x8 子块类型	JND 方案
PLANE	所有 8x8 子块是 PLANE	16x16 宏块 PLANE
	否则	表格 2
EDGE	所有 8x8 子块是 EDGE	16x16 宏块 EDGE
	否则	表格 2
TEXTURE	所有 8x8 子块是 TEXTURE	16x16 宏块 TEXTURE
	否则	表格 2

[0105] 对帧内, 图像首先被分割成 16×16 宏块。块分类被采用作为平衡不同 JND 方案的基础。如表格 1 所示, 第一列表示 16×16 宏块类型, 而第二列表示 8×8 子块分类。第三列是产生 JND 方案的决定。例如, 如果 16×16 宏块被标记为 PLANE, 而在宏块内的所有 8×8 子块也被标记为 PLANE, 则采用 16×16 基于 DCT 的 JND 模型, 因为其具有能量压缩和细节信息保留的优点。否则, JND 方案确定过程将继续, 如以下表格 2:

[0106] 表格 2 (在 16×16 、 8×8 和 4×4 JND 方案中间的平衡策略)

[0107]

8x8 子块类型	4x4 子块类型	JND 方案
PLANE	所有 4x4 子块是 PLANE	8x8 子块 PLANE
	否则	4x4 子块
EDGE	所有 4x4 子块是 EDGE	8x8 子块 EDGE
	否则	4x4 子块
TEXTURE	所有 4x4 子块是 TEXTURE	8x8 子块 TEXTURE
	否则	4x4 子块

[0108] 例如,如果 8×8 子块被标记为 PLANE,并且在 8×8 子块内的所有 4×4 子块也被标记为 PLANE,根据 8×8 子块的不同信息和特性,如块类型是否是 PLANE 或 EDGE 或 TEXTURE,则采用 8×8 基于 DCT 的 JND 模型用于 16×16 宏块内的每个 8×8 子块。否则,通过考虑 4×4 子块的不同内容信息和特性,如区块类型是否是 PLANE 或 EDGE 或 TEXTURE,采用 4×4 基于 DCT 的 JND 模型用于每个 4×4 子块。

[0109] 帧间的时间平衡策略

[0110] 块及其子块之间的时间特性被用于帧间的平衡策略。在一个实施例里,运动矢量被用来表示物体的运动特性。运动特性距离 (MCD) 表示一个宏块及其子块之间的运动特性相似性。MCD 由以下等式 (23) 确定:

$$[0111] \quad MCD = \sqrt{\sum_{i=0}^{N-1} [(MV_{s_i_x} - MV_{M_x})^2 + (MV_{s_i_y} - MV_{M_y})^2]} / N \quad (23)$$

[0112] 其中 MV_{M_i} 是宏块的运动矢量, MV_{s_i} 是其子块的运动矢量。

[0113] 如果 MCD 小于阈值,空间内容信息一致,那么将使用一个较大的基于 DCT 的 JND。否则,将采用较小的基于 DCT 的 JND。

[0114] 图 5 描述宏块及其相应子块的运动矢量以解释 MCD 一致性概念用于帧间的时间平衡策略。图 5 显示一个当前帧 520 及其参考帧 510。在当前帧 520 里的宏块 521 有一个运动矢量,指向参考帧 510 里的相应宏块 511。当前帧 520 里的宏块 522 有一个运动矢量,指向参考帧 510 里的相应宏块 512。图中显示,与宏块 522 的子块的那些运动矢量相比,宏块 521 的子块的运动矢量与宏块 521 的运动矢量更加一致。宏块 522 的 MCD 大于宏块 521 的 MCD,甚至大于阈值,都是可能的。

[0115] 心理物理实验

[0116] 图 4 显示一个实施例里分别在一个 16×16 块和一个 4×4 块内示例分配测试 DCT 子带。在一个 16×16 块 410 的实施例里,测试 DCT 子带是由阴影单元表示,如单元 411,其分配如附图所示。在一个 4×4 块 420 的实施例里,测试 DCT 子带由阴影单元表示,如单元 421,其分配如附图所示。如图 4 所示,测试 DCT 子带被分配,使得至少一个测试 DCT 子带安排在每行和每列上,测试 DCT 子带覆盖低频、中频和高频,并且至少一个 DCT 子带被安排在每个对角线上。噪声被分别加到 16×16 和 4×4 DCT 的某些空间频率子带上。对每个测试 DCT 子带,基于预设的衡量,对每个子带选择多个幅度的噪声。对是否能够感受到噪声,观察

者提供意见和看法。如果能察觉到噪声,将降低噪声幅度。相反,如果感受不到噪声,将提高噪声幅度。在实验之后,能够获得所选择 DCT 子带的最小可觉差阈值。

[0117] JND 模型有方向性,其表示为对水平和垂直频率分量的最大敏感度和对对角线分量的最小敏感度。

[0118] 在参数化期间,至少一个 DCT 子带被安排在每行和每列上,至少一个 DCT 子带被安排在每个对角线上,并且 DCT 子带最好覆盖低频、中频和高频。

[0119] 采用最小均方误差来拟合获得的 JND 阈值,如以下等式 (24) :

$$[0120] \quad (a, b, c) = \arg \min_{\omega_{ij}} \sum [T_{\omega_{ij}} - T_{basic}(m, n, i, j)]^2 \quad (24)$$

[0121] 其中 $T_{\omega_{ij}}$ 是从心理物理实验获得的 JND 阈值。

[0122] 图 3 显示一个基于 ABT 的 JND 滤波的典型实施例的流程图。在一个实施例里,依照不同变换尺寸的 DCT, 4×4 残差系数 301、 8×8 残差系数 302、和 16×16 残差系数 303 是从 ABT 300 获得。

[0123] 通过产生 JND 模型 310,而产生选择的 JND 模型 311,其可以是 4×4 JND 模型、 8×8 JND 模型、或 16×16 JND 模型。此外,每个变换尺寸有两种类型的 JND 模型,即帧内宏块的空间 JND 模型和帧间宏块的时间 JND 模型。

[0124] 为了使 JND 模型与 ABT 变换兼容,选择的 JND 模型通过 JND 扩展 312 被扩展到基于不同块尺寸的 JND 模型。在 JND 扩展 312 里,一些信息诸如不同块的块类型、块分类和运动矢量被用来产生每个块的基于不同块尺寸的 JND 模型。因此,获得了 4×4 JND 模型 321、 8×8 JND 模型 322、和 16×16 JND 模型 323 用于滤波变换的残差系数。

$$[0125] \quad \hat{E}_{typ} = \begin{cases} 0, & \text{if } E_{typ} \leq JND_{typ} \\ \text{sign}(E_{typ}) \cdot (|E_{typ}| - JND_{typ}), & \text{otherwise} \end{cases} \quad (25)$$

[0126] 等式 (25) 表示滤波过程。“typ”表示用于 DCT 变换和 JND 模型的块尺寸。 E_{typ} 是根据不同块尺寸的变换的残差系数, JND_{typ} 是产生的 JND,其可以是 4×4 JND 模型 321、 8×8 JND 模型 322、或 16×16 JND 模型 323。 \hat{E}_{typ} 是被滤波的系数。

[0127] 在滤波之后,获得 4×4 被滤波的残差系数 331、 8×8 被滤波的残差系数 332、或 16×16 被滤波的残差系数 333。残差系数与预测模式信息一起将由率失真优化 (RDO) 340 进行处理以选择最佳预测模式。因此,当率失真被激活时,选择用于块的预测模式即是最小化率失真成本的那个预测模式。在一个实施例里,率失真成本被定义为 $R = d + \lambda L$ 。

[0128] 其中 λ 是 Lagrangian 乘子,而 L 是对那个预测模式的块进行编码的实际比特长度。

[0129] 图 6 描述本发明提供的一个视频编码装置的典型实施例的模块示意图。以下所述的各个部件可以被实施为一个或多个处理器。预测模式确定单元 601 对输入视频数据进行运作,根据帧预测模式确定视频信号里的帧是帧间编码帧还是帧内编码帧。ABT 单元 603 在输入视频数据上执行自适应块变换以产生各种块尺寸的 DCT 变换信号。另一方面,基于 ABT 的 JND 模型产生单元 602 基于以下上述信息产生一个基于 ABT 的 JND 模型:

[0130] (1) 块类型,如帧间或帧内;

[0131] (2) 块分类,如平面、边缘或纹理;和

[0132] (3) 不同尺寸块的运动矢量。

[0133] 基于 ABT 的 JND 模型包括一个或多个 JND 阈值,这些阈值是由基于视觉感受模型的心理物理实验确定的。根据上述空间平衡策略,基于 ABT 的 JND 模型产生单元 602 基于空间内容信息计算帧内编码帧的基于 ABT 的 JND 模型。根据上述空间和时间平衡策略,基于 ABT 的 JND 模型产生单元 602 也基于空间内容信息和时间特性计算帧间编码帧的基于 ABT 的 JND 模型。

[0134] 通过丢弃一个或多个低于基于 ABT 的 JND 模型的 JND 阈值的残差系数,基于 ABT 的 JND 滤波单元 604 滤波一个或多个残差系数,从而输出一个被滤波的残差系数的信号。基于被滤波的残差系数,如果率失真优化单元激活时,通过降低率失真成本,率失真优化单元 605 确定率失真成本以选择最佳预测模式。

[0135] 接着,在对被滤波的残差系数进行编码之后,熵编码单元 606 计算被编码的数据。

[0136] 本发明实施例能够以软件、硬件、应用逻辑的形式实施,或以软件、硬件和应用逻辑的组合而实施。软件、应用逻辑和 / 或硬件可以被安置在集成电路芯片、模块或存储器里。如果期望的话,部分软件、硬件和 / 或用用逻辑可以被安置在集成电路芯片里,部分软件、硬件和 / 或用用逻辑可以被安置在模块里,并且部分软件、硬件和 / 或用用逻辑可以被安置在存储器里。在一个典型实施例里,应用逻辑、软件或指令集合保留在任何一种传统的非临时性计算机可读媒质里。

[0137] 在本说明书里描述的过程和逻辑流程可以通过一个或多个可编程处理器运行一个或多个计算机程序,运行输入数据并产生输出,以执行以上功能。过程和逻辑流程也可以通过专用逻辑电路如 FPGA(现场可编程栅极阵列)或 ASIC(专用集成电路)来运行。

[0138] 在本说明书里描述的装置或设备可以通过可编程处理器、计算机、芯片系统或其组合来实施,运行输入数据并产生输出。装置或设备可以包括专用逻辑电路如 FPGA(现场可编程栅极阵列)或 ASIC(专用集成电路)。装置或设备也可以包括硬件、为计算机程序产生运行环境的代码,如构成处理器固件、协议堆栈、数据库管理系统、运行系统、跨平台运行环境的代码,如虚拟机或一个或多个虚拟机的组合。

[0139] 适合执行计算机程序的处理器包括诸如通用和专用微处理器,以及任何类型数字计算机的任何一个或多个处理器。通常,处理器从只读存储器和 / 或随机存取存储器接收指令和数据。计算机元件通常包括一个用来执行指令的处理器、以及一个或多个用来存储指令和数据的存储器装置。

[0140] 在本说明书里描述的计算机可读媒质可以是任何与指令执行系统、装置或设备如计算机一起连接以保留、存储、通信、传播或传输所用指令的任何媒质或装置。计算机可读媒质可以包括计算机可读存储媒质,其可以是任何与指令执行系统、装置或设备如计算机一起连接以保留或存储所用指令的媒质或装置。计算机可读媒质可以包括所有形式的非易失性存储器、媒质和存储器装置,包括如半导体存储器装置,如 EPROM、EEPROM、和闪存装置,磁碟如内置硬碟或可移除碟、磁光碟、以及 CD-ROM 和 DVD-ROM 碟。

[0141] 计算机程序(也被称为程序、软件、软件应用程序、脚本或代码)可以任何编码语言撰写,包括可编译或可解释语言、可陈述或程序语言,并且能够以任何形式运行,包括以一个单机程序或适合用于计算环境里的一个模块、组件、子程序、对象或其它单元。计算机程序可以在一个计算机或多个计算机上运行,计算机可以安置在单个地点或分布在多个地点,并通过通信网络互连。

[0142] 在本说明书里描述的实施例和 / 或特征可以实施在一个计算系统里,其包括一个后端部件如数据服务器,或其包括一个中间件部件如应用程序服务器,或其包括一个前端部件如具有用户界面或 Web 浏览器的客户计算机,由此用户能够与本说明书里所述的实施例进行互动,或该计算机系统可以是一个或多个这种后端、中间或前端部件的任何组合。系统部件可以通过任何形式或通过数字数据通信的方法如通信网络而进行互连。通信网络的例子包括局域网 (LAN) 和广域网 (WAN),如互联网。

[0143] 计算系统可以包括客户机和服务器。客户机和服务器通常互相远离,并通常通过一个通信网络进行互动。客户机和服务器的关系凭借各个计算机上运行的计算机程序发生联系,并互相有一个客户机 - 服务器的关系。

[0144] 整个说明书包括许多具体实施细节。这些具体实施细节并不意味着是对本发明范围的限制,而是作为本发明特别实施例的具体特征描述。

[0145] 在各个实施例上下文里描述的某些特征也可以被合并,并作为单个实施例。相反地,在单个实施例的上下文里描述的几个特征也可以单独或以任何合适的亚合并方式实施在多个实施例里。此外,尽管特征可以被描述以某些方式合并,甚至初始作如此声明,来自所述组合或声明组合的一个或多个特征在某些情况下可以被排除出组合,并且声明组合可能涉及一个亚组合或一个亚组合的变化。尽管本发明各个方面在独立权利要求里有阐述,但本发明的其它方面还包括来自实施例和 / 或从属权利要求的特征以及独立权利要求特征的其他组合,且并不是仅仅在权利要求里明确阐述的组合。

[0146] 在本说明书里描述的某些功能可以以不同次序和 / 或互相同时执行。而且,如果期望的话,一个或多个上述功能可以是可选的或可以合并。

[0147] 以上描述提供了本发明的典型实施例,但不应该被看作是限制性的。此外,还有可能对其作出改变和修改,而不会脱离如在所附权利要求里所定义的本发明范围。

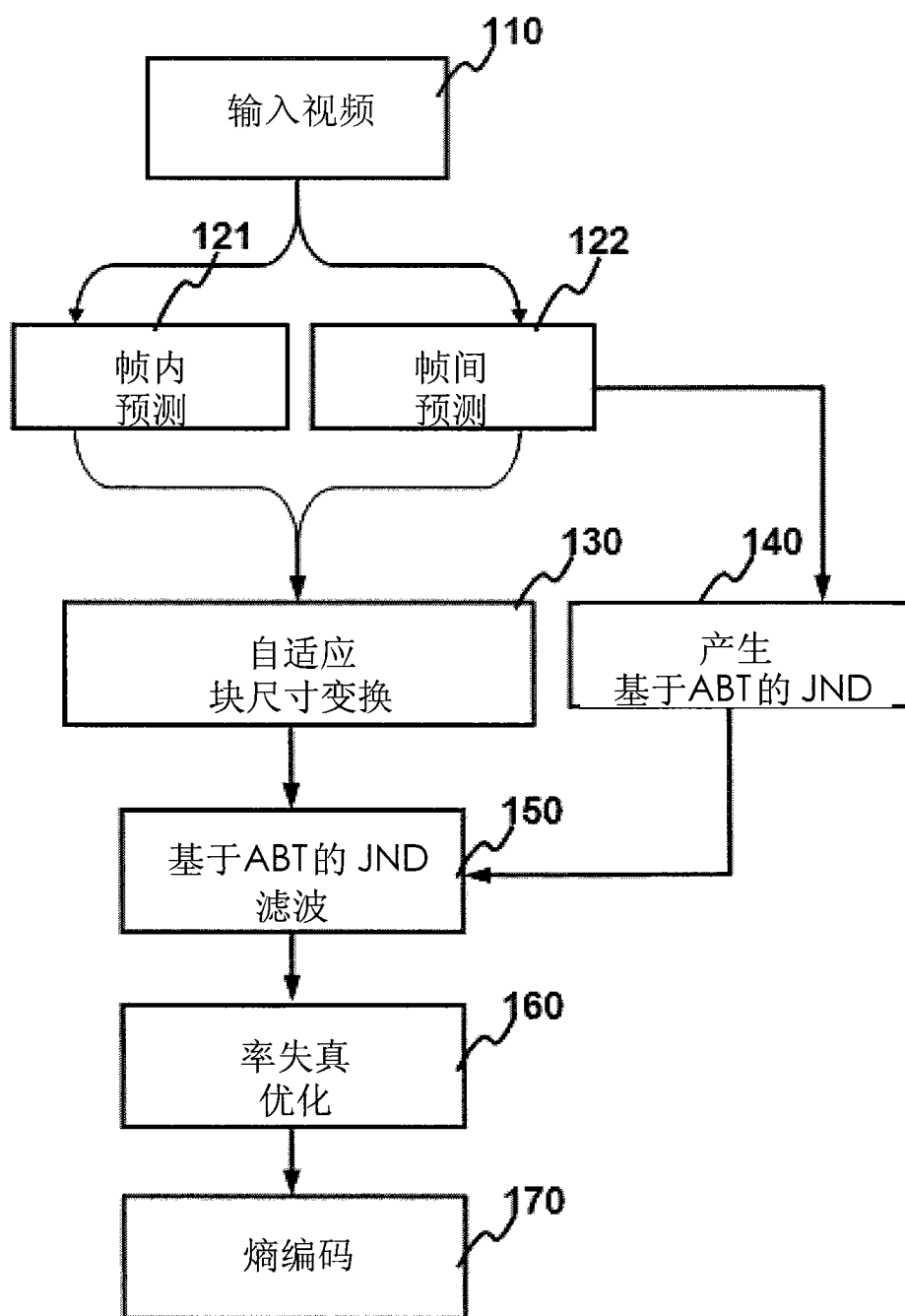


图 1

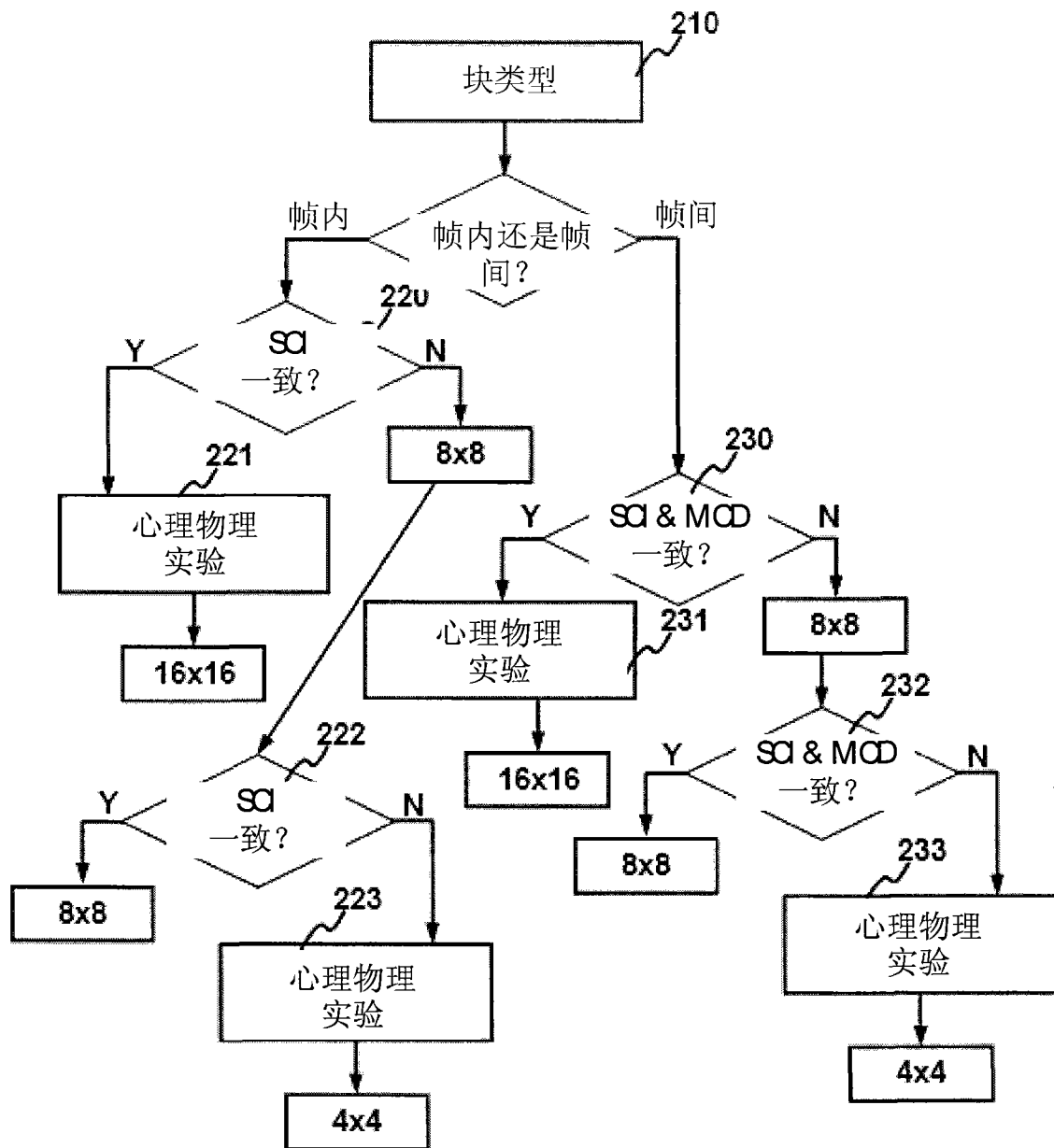


图 2

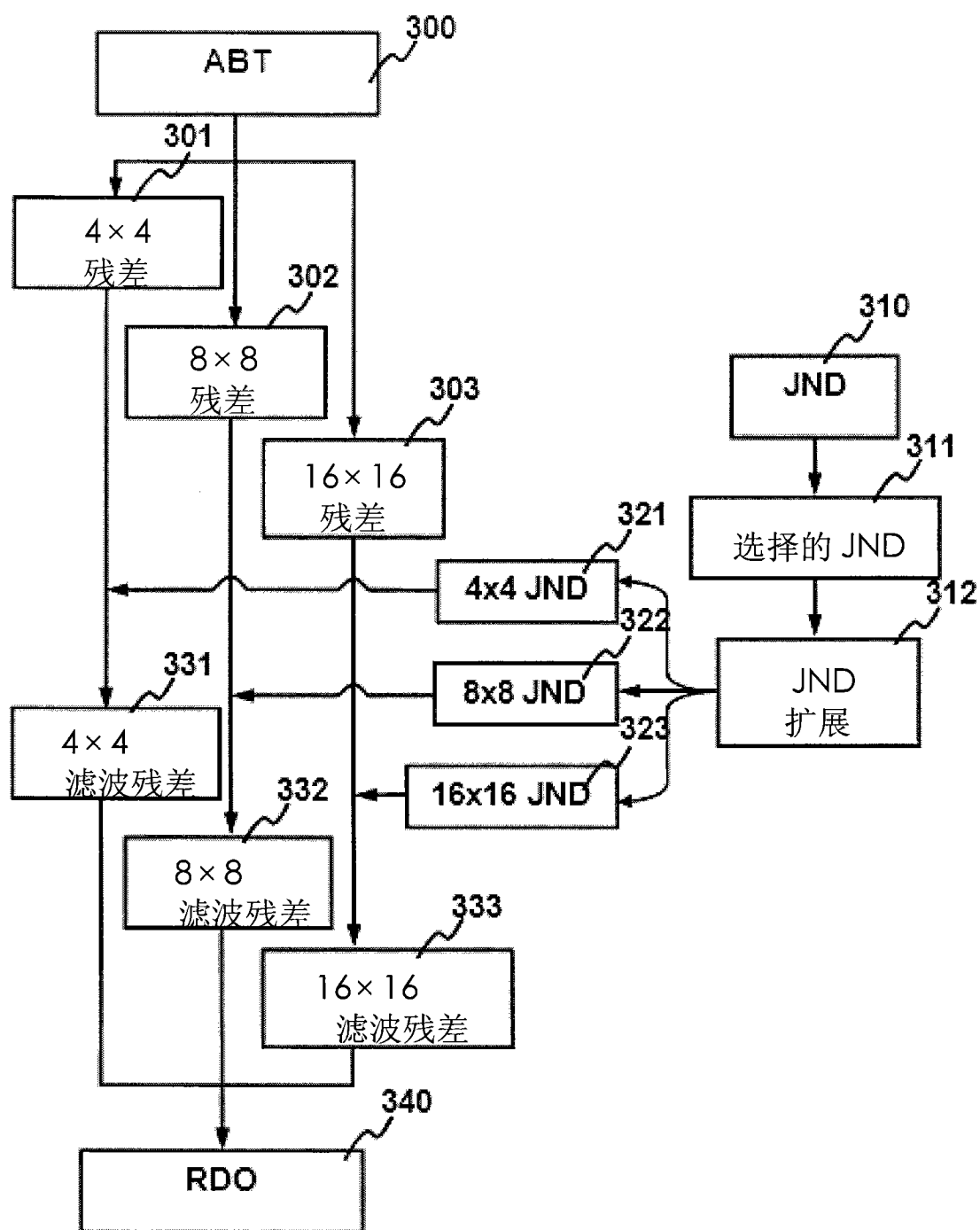


图 3

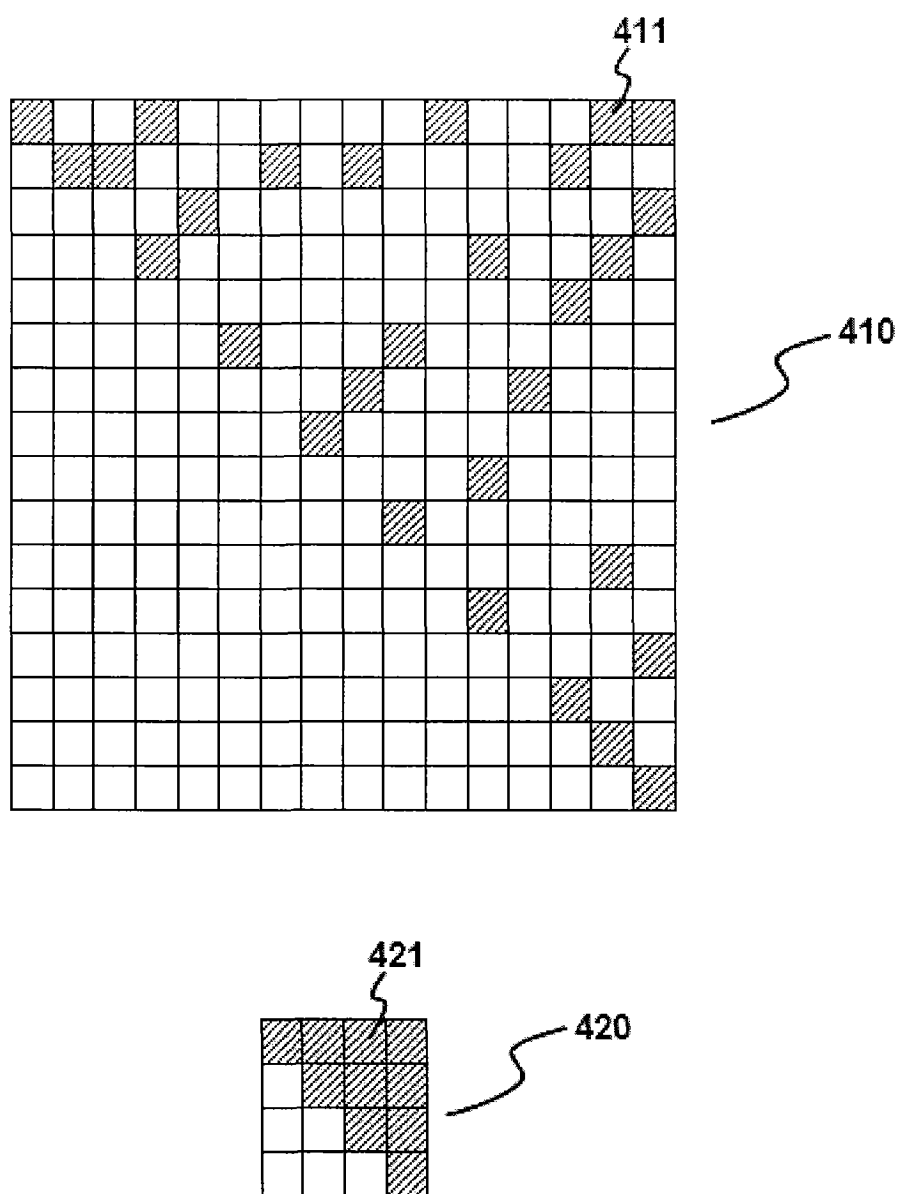


图 4

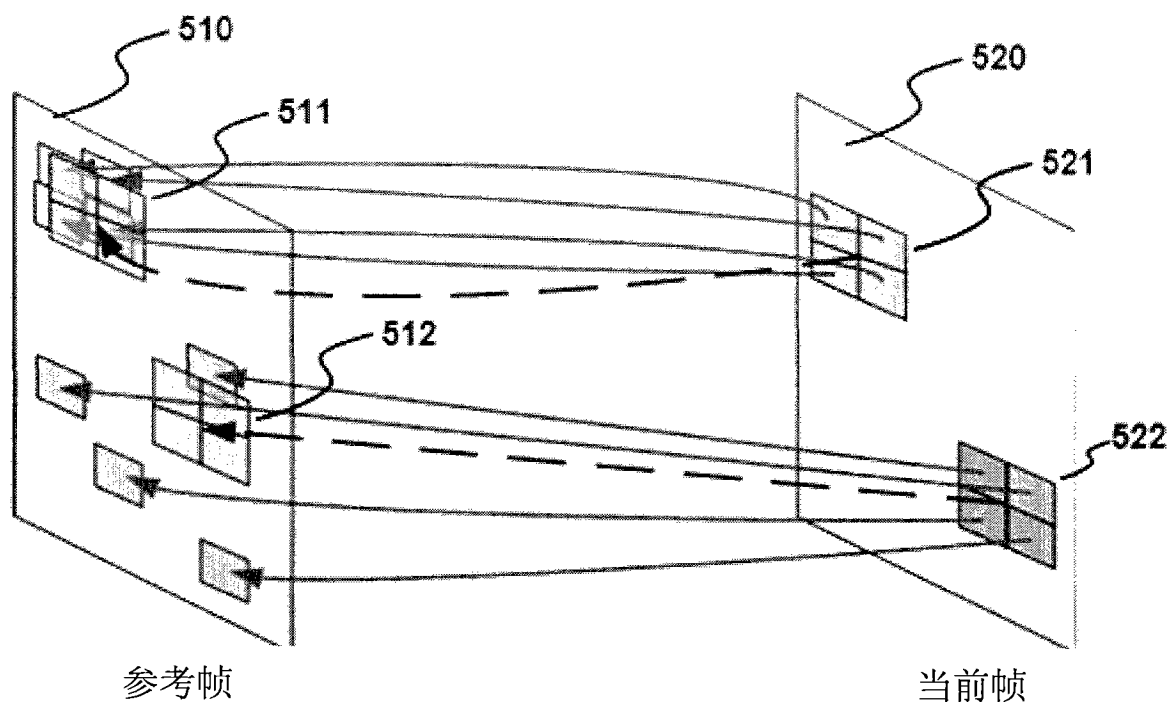


图 5

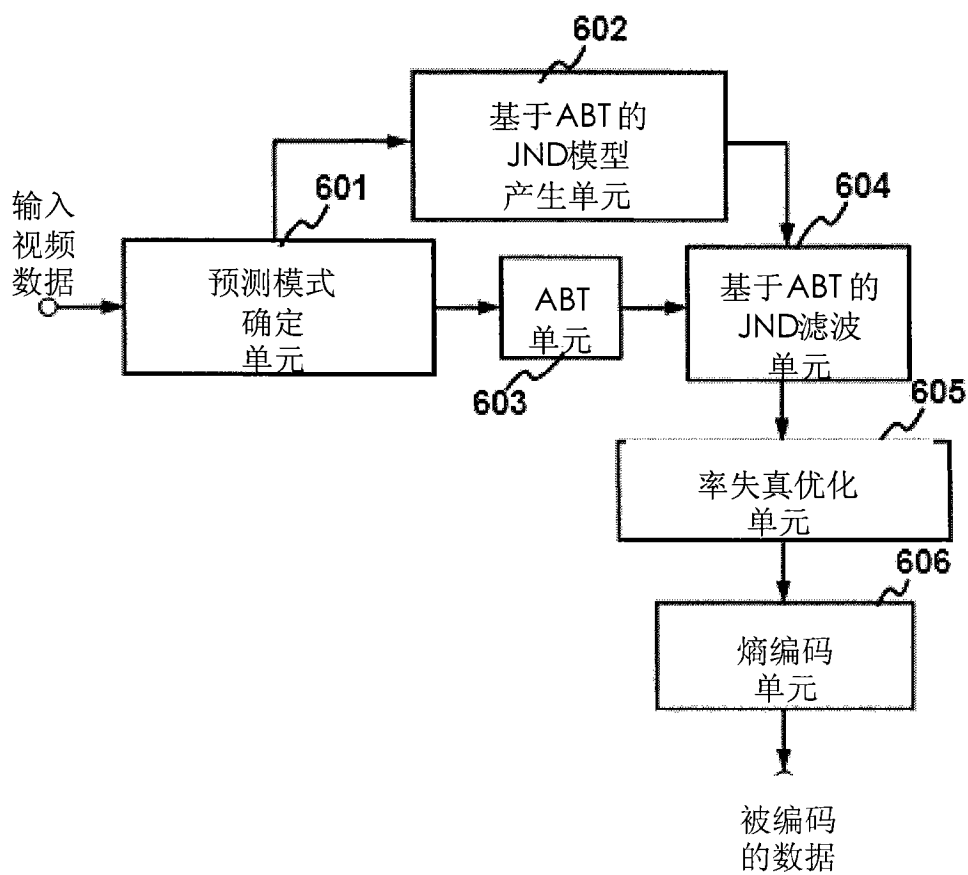


图 6