



(12) 发明专利

(10) 授权公告号 CN 101426135 B

(45) 授权公告日 2011. 07. 20

(21) 申请号 200810173512. 2

(22) 申请日 2008. 10. 29

(30) 优先权数据

2007-280940 2007. 10. 29 JP

2008-040983 2008. 02. 22 JP

2008-266831 2008. 10. 15 JP

(73) 专利权人 佳能株式会社

地址 日本东京都大田区下丸子 3-30-2

(72) 发明人 中山文贵

(74) 专利代理机构 北京怡丰知识产权代理有限公司

公司 11293

代理人 任之光 郭召道

(51) Int. Cl.

H04N 7/26 (2006. 01)

(56) 对比文件

CN 1531347 A, 2004. 09. 22, 全文.

JP 2005323312 A, 2005. 11. 17, 全文.

JP 2006262075 A, 2006. 09. 28, 全文.

CN 1926878 A, 2007. 03. 07, 全文.

JP 11196417 A, 1999. 07. 21, 全文.

审查员 雷云珊

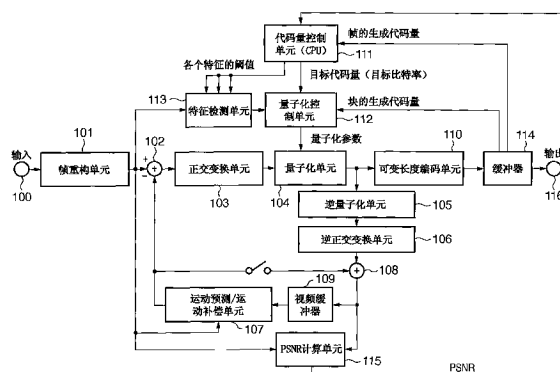
权利要求书 2 页 说明书 12 页 附图 11 页

(54) 发明名称

编码装置及其控制方法

(57) 摘要

本发明提供了一种编码装置及其控制方法, 所述编码装置包括: 检测单元, 被配置为确定多个块中的各个块的图像特征, 并且检测各个块中的视觉劣化是否明显; 确定单元, 被配置为基于所述检测单元的检测结果确定各个块的量子化参数; 变换单元, 被配置为执行对所述块的正交变换并且用所述确定的量子化参数执行量子化; 编码单元, 被配置为对所述变换单元的变换结果执行可变长度编码; 逆变换单元, 被配置为对所述变换单元的变换结果进行逆变换, 以生成局部解码图像; 以及计算单元, 被配置为计算参照值来改变用于确定所述图像特征的标准, 其中, 所述检测单元通过根据所述参照值来改变所述标准, 以检测具有明显视觉劣化的块。



1. 一种编码装置,其将输入图像分割为多个块,并且对所述图像进行编码,所述编码装置包括:

检测单元,被配置为确定所述多个块中的各个块的图像特征,并且检测视觉劣化明显的块,其中,所要检测的块是平坦块、边缘块和肌色块中的至少一个;

确定单元,被配置为确定所述检测单元所检测的各个块的量子化参数;

变换单元,被配置为执行对所述多个块的正交变换并且用所确定的量子化参数执行量子化;

编码单元,被配置为对所述变换单元的变换结果执行可变长度编码;

逆变换单元,被配置为对所述变换单元的变换结果进行逆变换,以生成局部解码图像;
以及

计算单元,被配置为基于所述输入图像和所述局部解码图像计算所述多个块的各个的峰值信噪比,并且根据所述峰值信噪比来改变用于确定所述图像特征的标准,

其中,所述检测单元根据所述计算单元改变的所述标准,来检测视觉劣化明显的块。

2. 根据权利要求1所述的编码装置,其中所述峰值信噪比代表所述图像的劣化度。

3. 根据权利要求1或2中任意一项所述的编码装置,其中如果所述具有明显视觉劣化的块包含于所述输入图像的平坦区域内,

则所述检测单元计算所述块中的像素的分散度,基于所述分散度的大小确定所述块是否包含于所述平坦区域内,并检测所述具有明显视觉劣化的块作为所述平坦块。

4. 根据权利要求3所述的编码装置,其中所述检测单元通过使用第一阈值作为所述标准并将所述第一阈值与所述分散度对比,来确定所述块是否包含于所述平坦区域内。

5. 根据权利要求1或2中任意一项所述的编码装置,其中如果所述具有明显视觉劣化的块至少是所述输入图像的边缘的一部分,

则所述检测单元进一步将所述块分割为多个子块、计算所述子块之间的像素的分散度的差、基于所述差的大小确定所述块是否包含于所述边缘中并检测所述具有明显视觉劣化的块作为所述边缘块。

6. 根据权利要求5所述的编码装置,其中,

所述检测单元通过使用第二阈值作为所述标准并将所述第二阈值与所述差对比,来确定所述块是否至少是所述边缘的一部分,并且

如果所计算的峰值信噪比高于目标峰值信噪比,则增加所述第二阈值的值。

7. 根据权利要求1或2中任意一项所述的编码装置,其中,如果所述具有明显视觉劣化的块包含于所述输入图像中的肌色区域内,

则所述检测单元计算所述块中的肌色像素的数量,基于所述数量的大小确定所述块是否至少是所述肌色区域的一部分,并检测所述具有明显视觉劣化的块作为所述肌色块。

8. 根据权利要求7所述的编码装置,其中,

所述检测单元通过使用第三阈值作为所述标准并将所述第三阈值与所述数量对比,来确定所述块是否包含于所述肌色区域,并且

如果所计算的峰值信噪比高于目标峰值信噪比,则降低所述第三阈值的值。

9. 一种控制编码装置的方法,所述编码装置将输入图像分割为多个块并且对所述图像进行编码,所述方法包括:

确定所述多个块中的各个块的图像特征,并检测视觉劣化明显的块,其中,所要检测的块是平坦块、边缘块和肌色块中的至少一个;

确定在所述检测中所检测的各个块的量子化参数;

执行对所述多个块的正交变换,并且用所确定的量子化参数执行量子化;

对所述正交变换的变换结果执行可变长度编码;

对所述正交变换的变换结果进行逆变换,以生成局部解码图像;以及

基于所述输入图像和所述局部解码图像计算所述多个块的各个的峰值信噪比,并且根据所述峰值信噪比来改变用于确定所述图像特征的标准,

其中,在所述检测中,根据在所述计算中改变的所述标准,来检测视觉劣化明显的块。

编码装置及其控制方法

技术领域

[0001] 本发明涉及编码装置及其控制方法。

背景技术

[0002] 随着近年来多媒体的发展,提出了各种运动图像压缩编码方法。典型的例子是 MPEG-1、2 和 4,以及 H. 264。在压缩编码处理中,将包含在运动图像中的原图像(图像)分割为称作块的预定区域,并且对各个分割的块实施运动补偿/预测以及 DCT 变换处理。对于运动补偿/预测,通过对编码图像数据进行局部解码来获取基准图像。因此,解码处理甚至在编码中也是必需的。

[0003] 当图像按照 MPEG 压缩和编码时,代码量通常会根据作为图像本身特征的空间频率特征、场景以及量子化标度值而大幅变化。代码量控制是通过实现具有这样编码特征的编码装置来允许以高图像质量获取解码图像的重要技术。

[0004] 作为代码量控制算法的一种,通常使用 TM5(测试模型 5)。TM5 代码量控制算法包含下述的三个步骤。在下列三个步骤中控制代码量以在每个 GOP(画面组)中确保预定的比特率。

[0005] (步骤 1)

[0006] 确定下面将要编码的画面(picture)的目标代码量。通过下式计算作为当前 GOP 中可用代码量的 R_{gop} :

[0007] $R_{gop} = (n_i + n_p + n_b) * (bits_rate / picture_rate) \quad \dots (1)$

[0008] 其中, n_i , n_p 和 n_b 为当前 GOP 中剩余 I、P 和 B 画面的数量, $bits_rate$ 为目标比特率, $picture_rate$ 为画面率。

[0009] 基于以下编码结果获取 I、P 和 B 画面的复杂度 X_i , X_p 和 X_b :

[0010] $X_i = R_i * Q_i$

[0011] $X_p = R_p * Q_p$

[0012] $X_b = R_b * Q_b \quad \dots (2)$

[0013] 其中, R_i , R_p 和 R_b 为通过对 I、P 和 B 画面进行编码而获取的代码量, Q_i , Q_p 和 Q_b 为在 I、P 和 B 画面中所有宏模块中的 Q 标度的平均值。基于等式 (1) 和 (2), I、P 和 B 画面的目标代码量 T_i , T_p 和 T_b 可以由下式获取 :

[0014] $T_i = \max \{ (R_{gop} / (1 + ((N_p * X_p) / (X_i * K_p)) + ((N_b * X_b) / (X_i * K_b)))) , (bit_rate / (8 * picture_rate)) \}$

[0015] $T_p = \max \{ (R_{gop} / (N_p + (N_b * K_p * X_b) / (K_b * X_p))) , (bit_rate / (8 * picture_rate)) \}$

[0016] $T_b = \max \{ (R_{gop} / (N_b + (N_p * K_b * X_p) / (K_p * X_b))) , (bit_rate / (8 * picture_rate)) \}$

... (3)

[0017] 其中, N_p 和 N_b 为在当前 GOP 中 P 和 B 画面的剩余数量, 常数 $K_p = 1.0$, $K_b = 1.4$ 。

[0018] (步骤 2)

[0019] 三个虚拟缓冲器分别用于 I、P 和 B 画面, 以管理通过等式 (3) 获取的目标代码量

和生成代码量的差。反馈各个虚拟缓冲器的数据积累量,并基于该数据积累量为下面将要编码的宏模块设定 Q 标度参照值,从而使实际生成代码量接近于目标代码量。例如,如果当前画面类型为 P 画面,则可以通过基于下式的运算处理来获取目标代码量和生成代码量之间的差:

$$[0020] \quad d_{p,j} = d_{p,0} + B_{p,j-1} - ((T_p * (j-1)) / MB_cnt) \quad \dots (4)$$

[0021] 其中,下标 j 为画面中的宏模块数, $d_{p,0}$ 为虚拟缓冲器的初始充满度 (initial fullness), $B_{p,j}$ 为到第 j 个宏模块时的总代码量, MB_cnt 为画面中的宏模块的数量。通过下式用 $d_{p,j}$ (以下称作“ d_j ”) 来获取第 j 个宏模块中的 Q 标度参照值:

$$[0022] \quad Q_j = (d_j * 31) / r \quad \dots (5)$$

$$[0023] \quad \text{这里 } r = 2 * \text{bits_rate/picture)rate} \quad \dots (6)$$

[0024] (步骤 3)

[0025] 执行基于编码目标宏模块的空间活动性来最终确定量子化标度以获取满意的视觉特征 (即,高解码图像质量) 的处理。

$$[0026] \quad ACT_j = 1 + \min(vblk1, vblk2, \dots, vblk8) \quad \dots (7)$$

[0027] 其中, vblk1 至 vblk4 为具有帧结构的宏模块中的 8×8 子块中的空间活动性, vblk5 至 vblk8 为具有场结构的宏模块中的 8×8 子块中的空间活动性,空间活动性可以通过下式计算:

$$[0028] \quad vblk = \sum (P_i - Pbar)^2 \quad \dots (8)$$

$$[0029] \quad Pbar = (1/64) * \sum P_i \quad \dots (9)$$

[0030] 其中, P_i 为第 i 个宏模块中的像素值,等式 (8) 和 (9) 中的 $\sum$ 表示对 $i = 1$ 到 64 的计算。通过等式 (7) 获取的 ACT_j 通过下式标准化:

$$[0031] \quad N_ACT_j = (2 * ACT_j + AVG_ACT) / (ACT_j + AVG_ACT) \quad \dots (10)$$

[0032] 其中, AVG_ACT 为之前被编码的画面中的 ACT_j 的参照值,并且通过下式最终计算量子化标度 (Q 标度值) MQUANT_j:

$$[0033] \quad [0033] \quad MQUANT_j = Q_j * N_ACT_j \quad \dots (11)$$

[0034] [0034] 根据上述 TM5 算法,步骤 1 中的处理分配给 I 画面大的代码量。大的代码量分配给画面中视觉上明显劣化的平坦块 (具有低空间活动性),这样可以在抑制图像质量劣化的同时,进行预定比特率之内的代码量控制和量子化控制。

[0035] 这里还提供一种类似 TM5 根据图像特征控制量子化的方法,从而改进视觉特征 (日本特开 11-196417 号专利公报)。

[0036] 上述的 TM5 方法从各个宏模块中提取特征,然后基于该特征在改变量子化参数的同时执行自适应量子化,从而执行量子化控制来达到预定的目标代码量。

[0037] 在日本特开 11-196417 号专利公报中,如果具有高复杂度并且需要提高量子化参数的块的数量较少,则即使生成的代码量增加,具有高复杂度的块的劣化也将变得显著。因此,抑制了自适应量子化。这同样适用于具有低复杂度并且需要降低量子化参数的块的数量较多的情况。

[0038] 发明内容

[0039] 但是,在上述算法中,图像质量会因各帧之间存在 / 不存在自适应量子化而发生变化。

[0040] 即使在包含多个需要降低量子化参数的块中,一些诸如平坦块的块由于块的特征所以代码量也不增加。因此,对于本应执行自适应量子化的图像却不能执行自适应量子化。

[0041] 本发明提供一种通过根据编码图像的劣化度和各个块的特征实行自适应量子化来抑制图像质量劣化的编码技术。

[0042] 根据本发明实施例的一个方面,本发明涉及一种编码装置,其将输入图像分割为多个块,并且对所述图像进行编码,所述编码装置包括:检测单元,被配置为确定所述多个块中的各个块的图像特征,并且检测视觉劣化明显的块,其中,所要检测的块是平坦块、边缘块和肌色块中的至少一个;确定单元,被配置为确定所述检测单元所检测的各个块的量子化参数;变换单元,被配置为执行对所述多个块的正交变换并且用所确定的量子化参数执行量子化;编码单元,被配置为对所述变换单元的变换结果执行可变长度编码;逆变换单元,被配置为对所述变换单元的变换结果进行逆变换,以生成局部解码图像;以及计算单元,被配置为基于所述输入图像和所述局部解码图像计算所述多个块的各个的峰值信噪比,并且根据所述峰值信噪比来改变用于确定所述图像特征的标准,其中,所述检测单元根据所述计算单元改变的所述标准,来检测视觉劣化明显的块。

[0043] 根据本发明实施例的另一方面,本发明涉及一种编码装置控制方法,所述编码装置将输入图像分割为多个块并且对所述图像进行编码,所述方法包括:确定所述多个块中的各个块的图像特征,并检测视觉劣化明显的块;确定在所述检测中所检测的各个块的量子化参数;执行对所述多个块的正交变换,并且用所确定的量子化参数执行量子化;对所述正交变换的变换结果执行可变长度编码;对所述正交变换的变换结果进行逆变换,以生成局部解码图像;以及基于所述输入图像和所述局部解码图像计算所述多个块的各个的峰值信噪比,并且根据所述峰值信噪比来改变用于确定所述图像特征的标准,其中,在所述检测中,根据在所述计算中改变的所述标准,来检测视觉劣化明显的块。

[0044] 根据本发明实施例的又一个方面,本发明涉及一种编码装置,该编码装置包括:变换/编码单元,被配置为对分割为多个块的输入图像数据的各个块进行变换和编码,并且输出变换系数数据;特征检测单元,被配置为从所述多个块中提取具有明显视觉劣化的特征的块,并输出帧中所提取的块的数量;量子化控制单元,被配置为根据来自所述特征检测单元的所提取的块的所述数量,确定各个块的量子化参数;以及量子化单元,被配置为根据由所述量子化控制单元确定的所述量子化参数,针对各个块对从所述变换/编码单元输出的所述变换系数数据进行量子化。

[0045] 本发明的其它特征将在以下参照附图的示例性实施例的描述中得到明确。

[0046] 附图说明

[0047] 图 1 是示出根据本发明第一实施例的编码装置的构成的例子的框图;

[0048] 图 2 是说明根据本发明第一实施例的特征检测单元的处理的框图;

[0049] 图 3A 至 3C 是示出根据本发明的第一实施例的将检测单元的强度转换为等级的函数的图;

[0050] 图 4 是用于说明根据本发明第一实施例的 PSNR 的图;

[0051] 图 5A 至 5C 是说明根据本发明第一实施例的 PSNR 和阈值之间的关系图;

[0052] 图 6 是示出根据本发明第二实施例的编码装置的构成的例子的框图;

[0053] 图 7 示出了用于说明根据本发明第二实施例的闪烁发生(flickeroccurrence)的

机理的图；

[0054] 图 8A 至 8C 是用于说明根据本发明第二实施例的闪烁发生度和阈值之间的关系的图；

[0055] 图 9 是示出根据本发明第三实施例的代码量控制单元 111 和特征检测单元 113 的示意性构成的框图；

[0056] 图 10A 和 10B 是用于说明具有不同特征比例的图像例子的图；以及

[0057] 图 11A 至 11C 是示出在帧中所检测出的宏模块的数量和反应参数之间的关系的图。

具体实施方式

[0058] 现在将参照附图描述本发明的实施例。

[0059] 参照图 1 至 3C 描述第一实施例。图 1 是示出根据本实施例的编码装置的总体构成的例子的框图。图 2 和图 3A 至 3C 是用于说明根据本实施例的特征检测的图。图 4 是用于说明 PSNR 的图。图 5A 至 5C 是示出 PSNR 和阈值之间的关系的图。

[0060] 图 1 示出了根据本发明实施例的编码装置。该装置可以作为诸如数码摄像机的视频和音频信号记录装置来实现。该装置支持诸如 MPEG(运动图像专家组)或者 H. 264/AVC(高级视频编码)的编码方法。

[0061] 参照图 1, 通过将运动图像分割为预定的块所生成的输入信号 100, 被输入到编码装置。例如在 MPEG 中, 使用 16×16 或者 8×8 块。通过编码方法确定大小。以下将块称作“宏模块”。

[0062] 编码装置包含帧重构单元 101, 加/减单元 102, 正交变换(DCT)单元 103, 量子化单元 104, 逆量子化单元 105, 逆正交变换(逆 DCT)单元 106, 以及运动预测/运动补偿单元 107。编码装置还包含加/减单元 108, 作为帧存储器的视频缓冲器 109, 可变长度编码单元 110, 代码量控制单元 111, 量子化控制单元 112, 特征检测单元 113, 以及暂时保持来自编码装置的流的缓冲器 114。编码装置还包含基于输入图像和局部解码图像计算 PSNR(峰值信噪比)的 PSNR 计算单元 115, 以及来自编码装置的输出信号 116。PSNR 是代表在编码中图像的劣化度的指标, 其通过称作 dB(分贝)的单位来表示。用输入图像和局部解码图像之间的差的平方和来计算 PSNR。如果作为计算结果的 PSNR 值较大, 则编码劣化较小。如果 PSNR 值较小, 则编码劣化较大。在本实施例中, 对各个宏模块计算 PSNR, 并且将包含在图像中的所有宏模块的 PSNR 值的和用作计算的 PSNR。

[0063] 下面将描述根据本实施例的编码装置的操作。在帧重构单元 101 中根据编码画面类型对输入信号 100 进行重构。将经过重构之后的图像分割为多个块。当编码画面为帧内编码(内编码)方法时, 正交变换单元 103 对各个宏模块中的信号执行正交变换。量子化单元 104 对作为正交变换结果获取的 DCT 系数进行量子化。

[0064] 当编码画面为帧间编码画面(间编码)方法时, 逆量子化单元 105 对已编码画面执行逆量子化, 并且逆正交变换单元 106 对已编码画面执行逆正交变换。经过这些逆变换, 生成局部编码图像。运动预测/运动补偿单元 107 执行关于待编码的画面的运动预测和运动补偿。加/减单元 102 计算与局部解码图像的差值。正交变换单元 103 进行差值的正交变换。量子化单元 104 对 DCT 系数进行量子化。

[0065] 无论编码为帧内编码还是帧间编码,均通过可变长度编码单元 110 对通过量子化单元 104 获取的量子化信号进行编码。已编码的信号作为输出信号 116 输出。

[0066] 下面将描述代码量控制单元 111 和量子化控制单元 112。代码量控制单元 111 基于 GOP 中的画面的位数量将已分好的位数量 (assigned bit amount) 分配给 GOP 中的各画面,包括分配目标画面。按照在 GOP 中编码画面的顺序重复分配,以设定各画面的目标代码量。代码量控制单元 111 还向特征检测单元 113 提供各个特征的阈值。

[0067] 下面,量子化控制单元 112 确定量子化参数。为了使各画面的目标代码量与实际生成代码量相匹配,基于虚拟缓冲器的容量通过反馈控制获取从代码量控制单元 111 输出的各个宏模块的生成代码量,来作为量子化标度参照值。根据该量子化标度参照值,基于通过特征检测单元 113 计算出的活动性,用等式 (11) 确定将要被量子化单元 104 使用的量子化参数。上述操作对应于背景技术中所描述的步骤 1 至 3。

[0068] 下面将参照图 2 和 3A 至 3C 描述特征检测单元 113 的处理。特征检测单元 113 基于代码量控制单元 111 提供的阈值确定图像特征,并且检测各个宏模块中的视觉劣化是否明显。在本实施例中,将平坦块、边缘块和肌色块作为其中视觉劣化明显的块的例子。下面将说明所有的平坦块、边缘块和肌色块。但是,可以对它们其中一个或者它们的组合来实施本发明。

[0069] 如图 2 所示,特征检测单元 113 包含平坦检测单元 201,边缘检测单元 203,以及肌色检测单元 205。这些检测单元从代码量控制单元 111 分别接收平坦阈值 202 (th1F, th2F),边缘阈值 204 (th1E, th2E),以及肌色阈值 206 (th1S, th2S)。这些阈值用作确定宏模块是属于平坦块、边缘块还是肌色块的标准。在根据各个检测单元计算的等级进行正规化之前,最小值计算单元 207 计算活动性。正规化单元 208 在正规化之前,通过将预定的反应参数应用于活动性来计算正规化的活动性。各个参数是可变的。

[0070] 图 3A 至 3C 是用于说明将检测单元的强度转换为等级的函数的图。图 3A 示出用于平坦检测单元的函数 301。图 3B 示出用于边缘检测单元的函数 302。图 3C 示出用于肌色检测单元的函数 303。

[0071] 根据输入各个检测单元的阈值将输入特征检测单元 113 的各个块的图像信号转换为等级。下面将描述各个检测单元的处理。

[0072] 平坦检测单元 201 计算块信号的分散值,并且将该分散值与阈值进行对比。如果分散值等于或小于 th1F,则等级为 1。如果分散值等于或大于 th2F,则等级为无穷。分散值代表块中的像素值的分散度。随着等级的降低,块中的像素值的分散值变小。即,块趋近于平坦图像。如果分散值落入阈值 th1F 和 th2F 之间,则基于连接 (th1F, 1) 和 (th2F, pre_avg) 的函数来计算等级。注意“pre_avg”是紧邻前面的编码画面的活动性的平均值。最初的阈值 th1F 和 th2F 包含在通过由 CPU 形成的代码量控制单元 111 为各个帧设定的平坦阈值 202 (第一阈值) 中。

[0073] 边缘检测单元 203 将块信号进一步分割为 8×8 子块,并且计算各个子块的分散值。边缘检测单元计算分散值的最大值和最小值之间的差,并且将该差与阈值进行对比。如果该差等于或大于 th2E,则等级为 1。如果差等于或小于 th1E,则等级为无穷。随着等级的降低,子块之间的分散值差增大。即,块趋近于边缘图像。如果差落入阈值 th1E 和 th2E 之间,则基于连接 (th1E, pre_avg) 和 (th2E, 1) 的函数计算等级。注意“pre_avg”是紧邻的

前面的编码画面的活动性的平均值。阈值 $th1E$ 和 $th2E$ 包含在通过由 CPU 形成的代码量控制单元 111 为各个帧设定的边缘阈值 204 (第二阈值) 中。

[0074] 肌色检测单元 205 基于块信号的亮度分量和色差分量对肌色像素的数量 进行计数。肌色检测单元将该数量与阈值进行对比。如果该数量等于或大于 $th2S$, 则等级为 1。如果该数量等于或小于 $th1S$, 则等级为无穷。随着等级的降低, 肌色像素的数量增大。即, 块趋近于肌色图像。如果该数量落入阈值 $th1S$ 和 $th2S$ 之间, 则, 基于连接 ($th1S$, pre_avg) 和 ($th2S$, 1) 的函数计算等级。注意“ pre_avg ”是紧邻的前面的编码画面的活动性的平均值。阈值 $th1S$ 和 $th2S$ 包含在通过由 CPU 形成的代码量控制单元 111 为各个帧设定的肌色阈值 206 (第三阈值) 中。

[0075] 最后, 最小值计算单元 207 将通过检测单元 201、203 和 205 计算出的等级中最小的一个确定为正规化之前的活动性。正规化单元 208 将该活动性正规化。如果来自检测单元的所有的等级均为无穷, 则将块的分散值确定为正规化之前的活动性。

[0076] 在本实施例中, 可以用下列方法基于通过 PSNR 计算单元 115 计算的 PSNR 来改变作为标准的各个阈值。该处理将参照图 4 和 5A 至 5C 来描述。

[0077] 图 4 是横坐标代表时间纵坐标代表 PSNR 的图。虚线表示作为预定目标的 PSNR (目标 PSNR)。实线表示通过 PSNR 计算单元 115 计算出的 PSNR (计算的 PSNR)。如果该计算的 PSNR 高于目标 PSNR, 则图像质量确定为高。如果该计算的 PSNR 低于目标 PSNR, 则图像质量确定为低。

[0078] 根据代码量的方面, 下述 PSNR 的解释也是可以的。当计算的 PSNR 高于目标 PSNR 时, 将足够的代码量分配给将要编码的图像, 并且同时保证在代码量方面的稳定状态。从而可以通过特征检测来检测更多的边缘和肌色块并分配更大的代码量。

[0079] 另一方面, 当计算的 PSNR 低于目标 PSNR 时, 不将足够的代码量分配给将要编码的图像, 并且同时生成在代码量方面的不稳定状态。因此, 如果通过特征检测检测出了太多的边缘和肌色块, 则在将不足的代码量分配给包含高频的块的同时, 将大的代码量分配给边缘和肌色块。即使当执行相对粗糙的量子化时, 包含高频的块中的视觉劣化也应该不明显。但是, 如果所分配的代码量太小, 则在包含高频的块中的图像质量将严重劣化。此外, 可能发生缓冲器故障。

[0080] 在本实施例中, 根据目标 PSNR 和计算的 PSNR 之间的差来控制将要提供给各个检测单元的阈值。图 5A 至 5C 是横坐标代表从目标 PSNR 中减去计算的 PSNR 所获取的差 (PSNR 差), 并且纵坐标代表各个检测单元阈值的图。

[0081] 图 5A 示出表示关于 PSNR 差的平坦阈值 202 中的变化的图 501。横坐标 代表 PSNR 差, 纵坐标代表与图 3A 中所示的函数 301 中的 $th1F$ 和 $th2F$ 相对应的平坦阈值。即, 阈值 $th1F$ 或 $th2F$ 根据 PSNR 差与图 501 相对地变化。

[0082] 图 5B 示出表示边缘阈值 204 的变化的图 502。横坐标代表 PSNR 差, 纵坐标代表与图 3B 中所示的函数 302 中的 $th1E$ 和 $th2E$ 相对应的边缘阈值。即, 阈值 $th1E$ 或 $th2E$ 根据 PSNR 差与图 502 相对地变化。

[0083] 图 5C 示出表示肌色阈值 206 的变化的图 503。横坐标代表 PSNR 差, 纵坐标代表与图 3C 中所示的函数 303 中的 $th1S$ 和 $th2S$ 相对应的肌色阈值。即, 阈值 $th1S$ 或 $th2S$ 根据 PSNR 差与图 503 相对地变化。

[0084] 如图 4 中的区域 401, 如果通过 PSNR 计算单元 115 计算出的计算的 PSNR 高于目标 PSNR, 则图像质量稳定并且可以进一步改进。因此, 改变阈值以检测更多的边缘块和肌色块。如图 5A 至 5C 所示, 考虑到目标 PSNR 和计算的 PSNR 之间的分离度 (劣化度), 优选将阈值连续地变化。对于平坦块, 因为即使在量子化参数改变时代码量也不是大量地增加, 所以阈值不变。

[0085] 如图 4 中的区域 402, 如果计算的 PSNR 低于目标 PSNR, 并且劣化度较高, 则图像质量通常较差。即使当将大的代码量分配给边缘块时, 也不能使图像质量改进。因此, 对于边缘块, 提高阈值 (Th1E 和 Th2E), 从而只检测强边缘, 即视觉上特别明显的边缘。肌色块通常包含诸如人的脸部的重要区域。因此, 即使当计算的 PSNR 大大低于目标 PSNR 并且劣化度高于预定的程度时, 很多肌色块也要通过降低阈值 (Th1S 和 Th2S) 来检测。在这种情况下, 尽管在除了脸部的任何部位都发生了劣化, 但是通过提高肌色块的图像质量仍可以改进总体视觉印象。

[0086] 如图 5A 至 5C 所示, 考虑到目标 PSNR 和计算的 PSNR 之间的分离度, 优选地将阈值连续地变化。各个检测单元所使用的方法不限于以上所描述的。除了分散值以外, 还可以通过例如频率转换等获取用于检测的元素。

[0087] 如上所述, 根据本实施例的编码装置通过根据编码图像的劣化度来改变将要提取的块的数量, 使在生成的代码量稳定的状态下, 改进视觉劣化明显的块的图像质量成为可能。即使当编码图像的劣化度较高时, 也将较大的代码量分配给尤其应该抑制其劣化的块, 从而给人以图像质量较好的感觉。

[0088] (第二实施例)

[0089] 在第一实施例中, 特征检测单元 113 通过使用代表图像劣化度的 PSNR, 来确定在各个宏模块中的视觉劣化是否明显。在第二实施例中, 依据在再现的图像中是否可能发生闪烁, 来控制将要提供给各个检测单元的阈值。

[0090] 图 6 是示出根据第二实施例的编码装置的总体构成的例子的框图。其基本构成除了在图 1 的构成中增加了检测闪烁的机构之外, 与第一实施例的构成相同。特征检测单元 613 的构成与图 1 中的特征检测单元 113 相同, 因此将参照图 2 描述。

[0091] 参照图 6, 编码装置包含帧重构单元 601, 加 / 减单元 602, 正交变换 (DCT) 单元 603, 量子化单元 604, 逆量子化单元 605, 逆正交变换 (逆 DCT) 单元 606, 以及运动预测 / 运动补偿单元 607。编码装置还包含加 / 减单元 608, 用作帧存储器的视频缓冲器 609, 可变长度编码单元 610, 代码量控制单元 611, 量子化控制单元 612, 特征检测单元 613, 以及暂时保持来自编码装置的流的缓冲器 614。编码装置还包含基于输入图像和局部解码图像计算 PSNR (峰值信噪比) 的 PSNR 计算单元 615, 以及来自编码装置的输出信号 616。上述构成与图 1 中相同。编码装置还包含帧特征检测单元 620, 帧运动检测单元 621, 以及闪烁检测单元 622。特征检测单元 613 提取各个宏模块的特征。而帧特征检测单元 620 提取各个帧的特征, 因此其表现是不同的。以下将描述这些单元的功能。

[0092] 帧特征检测单元 620 基于来自输入终端 600 的图像数据, 计算将要编码的图像的复杂度 (作为帧活动性)。在第二实施例中, 将图像数据的 AC 分量的量, 优选高频分量的量, 作为复杂度。更具体地讲, 将一个帧的图像数据分割为预定大小的块, 并且计算各个块的分散度。所有块的计算的分散度的总和作为高频分量的量。通过诸如 DCT (离散余弦变

换)或者阿达马变换(Hadamard transformation)的频率转换获取的高频成分可以代替分散度来使用。

[0093] 帧运动检测单元 621 使来自输入终端 600 的邻近帧的图像数据相关联,并且计算将要编码的整个图像的运动量。更具体地讲,将一个帧的图像分割为预定大小的块,并且在移动一个相邻帧图像的坐标的同时,在各个块中计算相关度最高的坐标移动量。将在帧的各自块中所计算的移动向量的总和定义为帧间运动量。帧间运动量代表全局运动(全局向量),并且可以用上述方法之外的任何方法计算。

[0094] 与第一实施例一样,装置具有基于输入图像和局部解码图像来计算编码失真量的 PSNR 计算单元 615。PSNR 计算单元 615 基于来自输入终端 600 的图像数据和局部解码图像数据(从加法器 608 输出的图像数据)来计算各个宏模块的 PSNR。随后,PSNR 计算单元 615 输出一个帧中的宏模块的 PSNR 值的总和(作为最终的 PSNR)。这里所计算的 PSNR 是编码图像(即,至少紧邻将要编码的图像的前面的图像)的 PSNR。

[0095] 闪烁检测单元 622 接收来自帧特征检测单元 620 的帧活动性、来自帧运动检测单元 621 的帧间运动量以及来自 PSNR 计算单元 615 的 PSNR。闪烁检测单元 622 根据这三个参数检测将要编码的图像是否可能发生闪烁。

[0096] 以下将参照图 7 描述闪烁发生的机理。图 7 示出用于说明根据本实施例的闪烁发生的机理的图。图 7 示出在使用帧内编码和帧间编码之后的亮度信号。图 7 中的图 701 示出 I 画面的再现信号等级中的时间率(time-rate)变化。横坐标代表时间(或者帧),纵坐标代表再现视频信号等级。亮度平均值的虚线表示视频信号的平均值。这也适用于图 702 和 703。峰值亮度的虚线表示视频信号的亮度的峰值。这也适用于图 702 和 703。附加在平坦视频信号上的噪音分量的峰值亮度可以通过帧内编码在某种程度上重构。这是因为在上述代码量控制算法中,分配给 I 画面的代码量大于其它剩余画面类型的代码量。

[0097] 图 7 中的图 702 示出在诸如 P 或 B 画面等通过帧间编码来编码的画面的再现信号等级中的时间率变化。横坐标代表时间(或者帧),纵坐标代表再现视频信号等级。在具有高复杂度的图像中帧之间的关联度变低。因此,在正常编码中,在 P 和 B 画面中,帧之间的差信息量增加。这使得编码中的视频信号劣化,从而不能对峰值亮度进行重构。在图 702 中,与图 701 中所示的 I 画面相比峰值亮度发生变化。因此,如图 703 所示,在运动图像再现的时刻发生亮度闪烁。横坐标代表时间(或者帧),纵坐标代表再现视频信号等级。

[0098] 但是,不是所有的大峰值亮度差都被感知为闪烁。人的视觉特征更容易检测带有微小运动的图像中的劣化,而不是带有大的运动的图像中的劣化。这就是为什么在静止图像的平坦块中包含干扰噪音的原因。因此,在带有大的运动的图像中不能检测出峰值亮度差,并且闪烁极少被感知。如果峰值亮度差不是在带有大的运动的图像中生成,而是在带有微小运动的图像中生成,则闪烁是明显的。

[0099] 以上描述了闪烁生成的机理。

[0100] 如上所述,闪烁生成的条件是编码之后不同画面类型之间的亮度峰值发生了变化,并且图像具有微小运动。亮度峰值在编码之后在满足以下两个条件的图像中生成:1)图像的复杂度高,以及 2)编码图像劣化。第一个条件意味着通过帧特征检测单元 620 计算的高频分量的量较大。第二个条件意味着通过 PSNR 计算单元 615 计算的 PSNR 较低。

[0101] 如果一个条件不满足,则不生成亮度峰值差。例如,包含很多高频的图像具有高复

杂度。即,图像满足第一条条件。但是,如果比特率较高,则编码图像不劣化,因此没有亮度峰值差生成。另一方面,如果编码图像劣化,则满足第二条条件。但是,如果比特率较低,则即使图像的复杂度较低,编码图像也会发生劣化。在这种情况下,没有亮度峰值差生成。图像带有微小运动的条件是整个图像的运动量小。这意味着通过帧运动检测单元 621 计算的帧间运动量小。

[0102] 闪烁检测单元 622 可以基于高频分量的量、帧间运动量以及 PSNR 计算闪烁发生度。如果高频分量的量大于其参照值、PSNR 小于其参照值并且帧间运动量小于其参照值,则在将要编码的图像中发生闪烁的概率较高。因此,闪烁发生度较高。相反,如果不满足这些条件,则闪烁发生度较低。闪烁发生度可以通过任意方法来计算。但是,闪烁发生度是离散值。这个值越大,则闪烁发生度越高。

[0103] 图 8A 至 8C 是用于说明根据本实施例的闪烁发生度和阈值之间的关系图。图 8A 至 8C 中的每个图示出横坐标代表闪烁发生度,纵坐标代表检测单元的阈值。图 8A 示出表示平坦阈值 202 关于闪烁发生度变化的图 801。横坐标代表闪烁发生度,纵坐标代表与图 3A 中所示的函数 301 中 $th1F$ 和 $th2F$ 相对应的平坦阈值。即,阈值 $th1F$ 和 $th2F$ 根据闪烁发生度与图 801 相对地变化。

[0104] 图 8B 示出表示边缘阈值 204 关于闪烁发生度变化的图 802。横坐标代表闪烁发生度,纵坐标代表与图 3B 中所示的函数 302 中的 $th1E$ 和 $th2E$ 相对应的边缘阈值。即,阈值 $th1E$ 和 $th2E$ 根据闪烁发生度与图 802 相对地变化。

[0105] 图 8C 示出表示肌色阈值 206 关于闪烁发生度变化的图 803。横坐标代表闪烁发生度,纵坐标代表与图 3C 中所示的函数 303 中的 $th1S$ 和 $th2S$ 相对应的肌色块阈值。即,阈值 $th1S$ 或 $th2S$ 根据闪烁发生度与图 803 相对地变化。

[0106] 如果基于闪烁检测单元 622 的结果闪烁发生度较高,并且相当大的代码量分配给了劣化明显的区域,则用于包含高频分量的区域的代码量变小,因而发生强烈闪烁。为防止这种情况,对于边缘块,将阈值($th1E$ 和 $th2E$) 提高到只检测强边缘,即,视觉特别明显的边缘。对于肌色块,将阈值($th1S$ 和 $th2S$) 提高到只检测强肌色,即,视觉特别明显的肌色。

[0107] 如果基于闪烁检测单元 622 的结果闪烁发生度较低,则图像只包含少数高频成分,并且 PSNR 足够高。从而进一步将代码量分配给劣化明显的区域。因此,对于边缘块,通过降低阈值($th1E$ 和 $th2E$) 检测到许多边缘。对于肌色块,通过降低阈值($th1S$ 和 $th2S$) 检测到许多肌色。

[0108] 对于平坦块,因为即使当量子化参数改变时,代码量也不会大量地增加,所以平坦阈值($Th1F$ 和 $Th2F$) 不依赖于闪烁发生度而变化。

[0109] 如上所述,根据本实施例的编码装置使其可以通过根据编码图像的闪烁发生度来改变将要提取的块的数量,来改进在生成的代码量稳定的状态下视觉劣化明显的块的图像质量。即使当编码图像的闪烁发生度较高时,也将较大的代码量分配给尤其应该抑制其劣化的块,从而给人以图像质量较高的印象。

[0110] (第三实施例)

[0111] 根据本实施例的编码装置的构成与图 1 相同。但是,在本实施例中,编码装置不需要包含 PSNR 计算单元 115。图 9 是示出特征检测单元 113 的示意性构成的框图。图 10A 和 10B 以及图 11A 至 11C 是说明其操作的图。

[0112] 图 9 是示出根据本实施例的代码量控制单元 111 和特征检测单元 113 的示意性构成的框图。平坦检测单元 1201 来自帧重构单元 101 的各个宏模块图像数据检测平坦块的度（等级）。边缘检测单元 1202 来自帧重构单元 101 的图像数据检测边缘的度（等级）。肌色检测单元 1203 来自帧重构单元 101 的图像数据检测肌色的度（等级）。检测单元 1201、1202 和 1203 的检测特征或者等级转换的函数与图 3A 至图 3C 相同。

[0113] 最小值计算 / 特征确定单元 1204 计算通过检测单元 1201、1202 和 1203 所计算出的等级的最小值，并且将该最小值等级（正规化之前的活动性）输出至正规化单元 1205。如果通过检测单元 1201、1202 和 1203 所计算出的所有的等级均为无穷，则最小值计算 / 特征确定单元 1204 将宏模块的分散度作为正规化之前的活动性输出。最小值计算 / 特征确定单元 1204 还将宏模块中的等级的最小值信息或者代表所有等级均为无穷的信息，作为宏模块的特征信息发送至反应参数选择单元 1210。

[0114] 正规化单元 1205 通过将预定的反应参数应用于来自最小值计算 / 特征确定单元 1204 的正规化之前的活动性，来计算正规化活动性。反应参数是确定量子化参数的值的参数。大的反应参数使量子化参数更小。

[0115] 反应参数选择单元 1210 提供将要被正规化单元 1205 在正规化处理中使用的反应参数的信息。使用与宏模块的特征相对应的反应参数。反应参数选择单元 1210 根据从最小值计算 / 特征确定单元 1204 接收的特征信息，选择将要使用的反应参数。在本实施例中，平坦反应参数 1206、边缘反应参数 1207 和肌色反应参数 1208 分别对应于平坦块（检测单元 1201）、边缘块（检测单元 1202）以及肌色块（检测单元 1203）来准备，每个参数用由特征信息代表的最小值来表示。当宏模块不属于平坦块、边缘块和肌色块中的任意一个时，一般反应参数 1211 也作为可用的反应参数来准备。反应参数选择单元 1210 根据由获取的特征信息所代表的最小值，选择反应参数 1206、1207 和 1208 中的一个。如果特征信息代表所有的等级均为无穷，则反应参数选择单元 1210 选择一般反应参数 1211。

[0116] 积分单元 1209 对通过检测单元 1201、1202 和 1203 对于各个类型所检测到的宏模块的数量进行计数，并且将所提取的块的数量提供给量子化控制单元 112。各个反应参数和积分单元 1209 之间的关系将参照图 10A 至 11C 描述。

[0117] 图 10A 和图 10B 是用于说明具有不同特征率的图像例子的图。图 10A 示出在特征检测时帧中包含 20% 平坦块，20% 边缘块，以及 20% 肌色块的图像。图 10B 示出包含 30% 平坦块，20% 边缘块，以及 40% 肌色块的图像。不属于上述任何一个类型的宏模块归为“其它”。平坦并且具有肌色的块，例如人的脸颊，属于平坦块。边缘并且具有肌色的块，例如脸部轮廓，属于边缘块。

[0118] 图 11A 至 11C 是示出在帧中检测的宏模块的数量与反应参数之间的关系的图。图 11A 示出平坦反应参数。横坐标代表检测的平坦块的数量，纵坐标代表反应参数。图 11B 示出边缘反应参数。图 11C 示出肌色反应参数。

[0119] 在平坦块，尽管代码量不会变化很大，但图像质量会随着量子化参数的变化而改变。即，当相同的量子化参数应用于图 10A 和 10B 所示的图像例子时，整个的平坦块的代码量不会发生大的变化。因此，在本实施例中，如图 11A 所示，平坦块的反应参数不会随着检测的宏模块的数量发生大的变化。

[0120] 另一方面，在边缘块和肌色块中，图像质量随着量子化参数的变化而改变，并且代

码量也随着量子化参数的变化而发生大的改变。通过降低边缘块的量子化参数可以降低蚊状噪声。但是,生成的代码量也将随之增加。同样,通过降低肌色块的量子化参数可以降低脸部的劣化。但是,生成的代码量也将随之增加。如果检测太多的边缘块和肌色块,大的代码量将分配给边缘块和肌色块,而分配给包含高频的块的代码量将减少。即使当执行相对粗糙的量子化时,包含高频的块中的视觉劣化也应该是不明显的。但是,如果分配的代码量太小,则在包含高频的块中,图像质量会严重地劣化。此外,由于整个图像的代码量也增加,因此比率变得高于预定比率。

[0121] 在本实施例中,根据检测的宏模块的数量来改变边缘块和肌色块的反应参数。如图 11B 和 11C 所示,如果检测为边缘块或肌色块的宏模块的数量较小,则反应参数增加,从而降低量子化参数。如果检测为边缘块或肌色块的宏模块的数量较大,则反应参数减小,从而防止量子化参数变得太低。图 11B 和 11C 示出具有相同趋势的曲线。但是,该曲线可以任意地设定。

[0122] 图 10A 中所示的图像例子中检测到的边缘块的数量与图 10B 中所示的相同。因此,用相对大的反应参数来确定量子化参数。但是,图 10B 中所示的图像例子中检测到的肌色块的数量大于图 10A 中所示的。因此,在图 10A 中所示的图像例子中,用相对大的反应参数来确定量子化参数。在图 10B 中所示的图像例子中,用相对小的反应参数来确定量子化参数。

[0123] 各个检测单元的检测方法不限于以上所描述的。除了分散度,检测要素可以通过,例如,频率转换来获取。在本实施例中,将要检测的块的类型为平坦块、边缘块以及肌色块。但是,其它任意类型的块都可以用作构成要素。如果该类型(例如平坦块)造成即使在量子化参数变化时,代码量也没有大的改变,则采用与用于平坦块的相同的控制。如果该类型造成在量子化参数变化时,代码量有大的改变,则采用与用于边缘块或肌色块的相同的控制。

[0124] 注意本发明能够应用到包含单一装置的设备或包含多个设备的系统。

[0125] 进而,本发明能够通过以下方式来实现:向系统或装置直接地或间接地提供实现了上述实施例的功能的软件程序,并且通过所述系统或设备的计算机来读取并执行所提供的程序代码。在此情况下,只要该系统或设备具有本程序的功能,则实施方式不必依靠程序。

[0126] 因此,由于本发明的功能能够由计算机实现,所以安装在计算机上的程序代码也实现了本发明。换句话说,本发明的权利要求也涵盖以实现本发明的功能为目的的计算机程序。

[0127] 在此情况下,只要所述系统或装置具有所述程序的所述功能,则可以以任何形式实现该程序,例如,目标代码、解释器执行的程序或提供给操作系统的脚本数据。

[0128] 可以用于提供程序的存储介质的例子是软盘、硬盘、光盘、磁光盘、CD-ROM、CD-R、CD-RW、磁带、非易失型存储卡、ROM 和 DVD (DVD-ROM、DVD-R 或 DVD-RW)。

[0129] 对于提供所述程序的方法,可以通过使用客户端计算机的浏览器来连接到因特网上的网站,并可以将本发明的计算机程序或该程序的可自动安装的压缩文件下载到诸如硬盘的记录介质上。此外,本发明的所述程序还可以通过以下方式提供:将构成所述程序的程序代码分割成多个文件,并从不同的网站下载这些文件。换句话说,本发明的权利要求还涵盖:给多个用户下载实现本发明的功能的程序文件的 WWW(万维网)服务器。

[0130] 还可以将本发明的程序加密并存储到诸如 CD-ROM 的记录介质上,然后将该记录介质发给多个用户,允许满足某种要求的用户通过互联网从网站下载解密密钥信息,并允许这些用户通过使用该密钥信息来将加密的程序解密,从而将该程序安装在用户计算机上。

[0131] 除了通过计算机执行所读取的程序来实现根据所述实施例的上述功能外,在该计算机上运行的操作系统等也可以执行部分或者全部的实际处理,使得能够通过该处理实现上述实施例的功能。

[0132] 此外,在将从存储介质读取的程序写入插入计算机的功能扩展板或者设置在连接到计算机的功能扩展单元中的存储器之后,安装在该功能扩展板或者功能扩展单元中的 CPU 等执行部分或者全部的实际处理,使得通过该处理实现上述实施例的功能。

[0133] 虽然参照示例性实施例对本发明进行了描述,但是应当理解的是,本发明并不限于所公开的示例性实施例。应当对权利要求的范围给予最宽泛的解释,以包括所有变体、等同结构和功能在内。

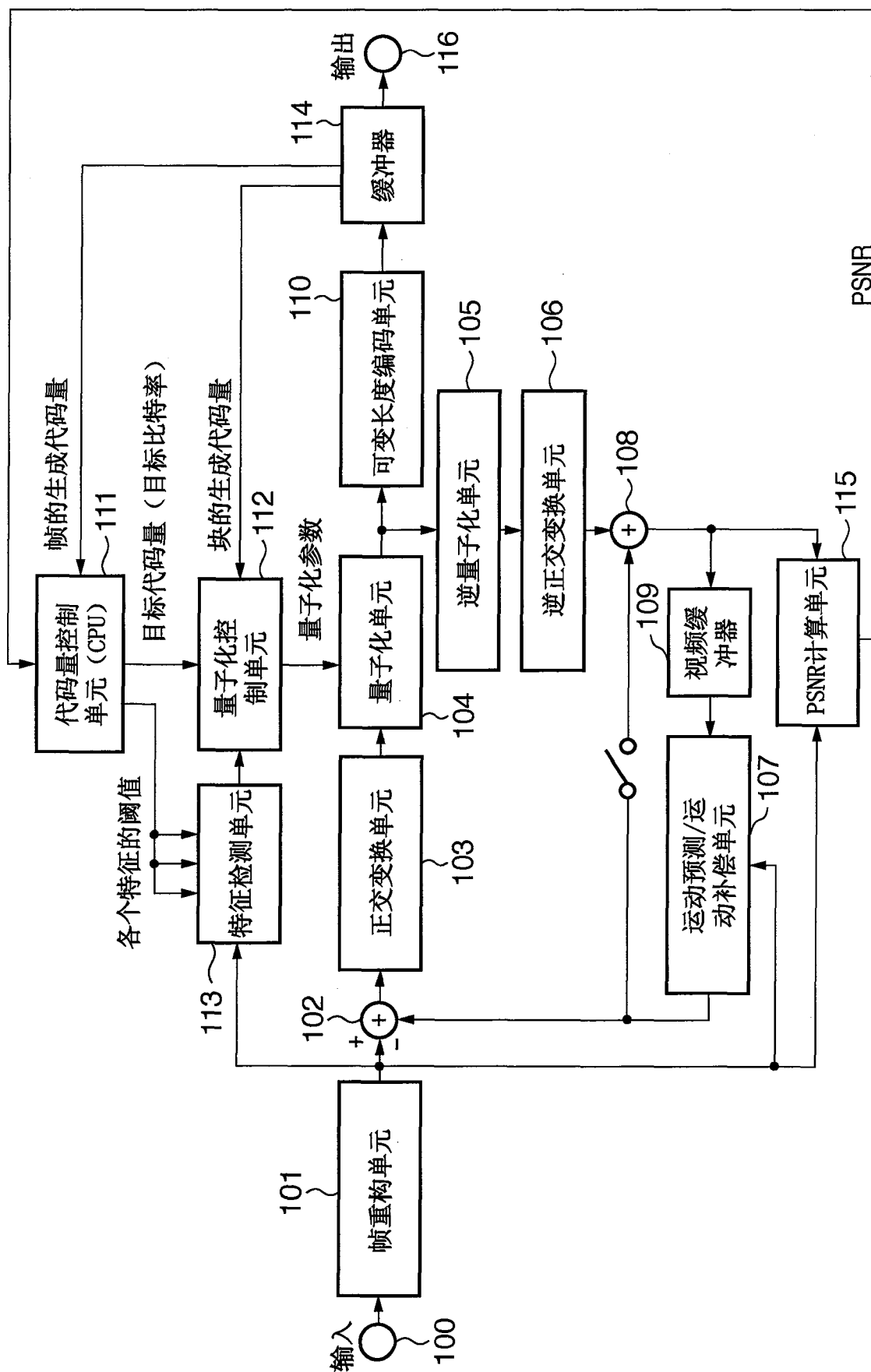


图 1

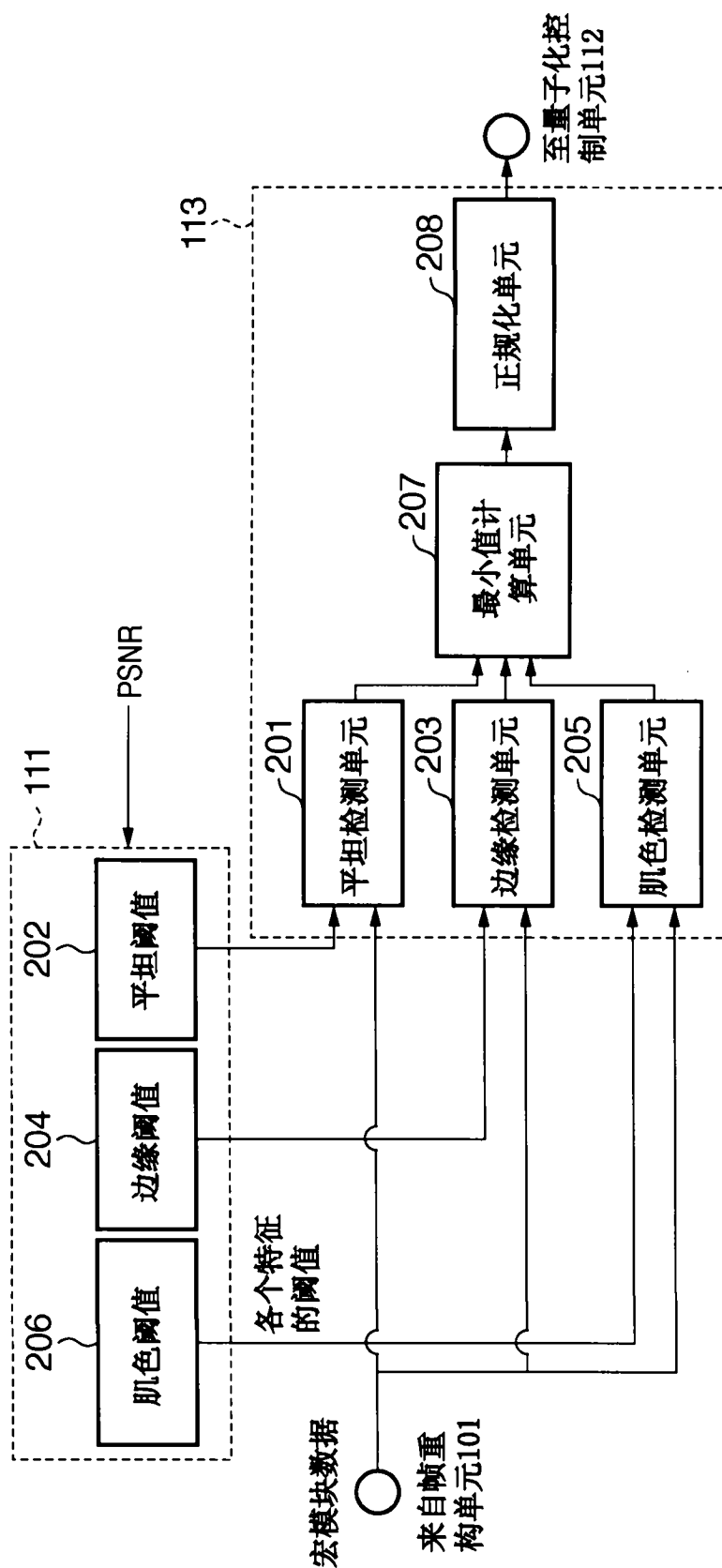


图2

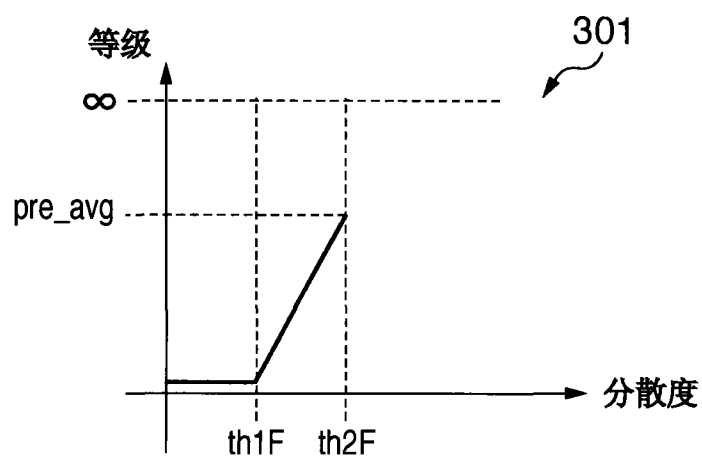


图 3A

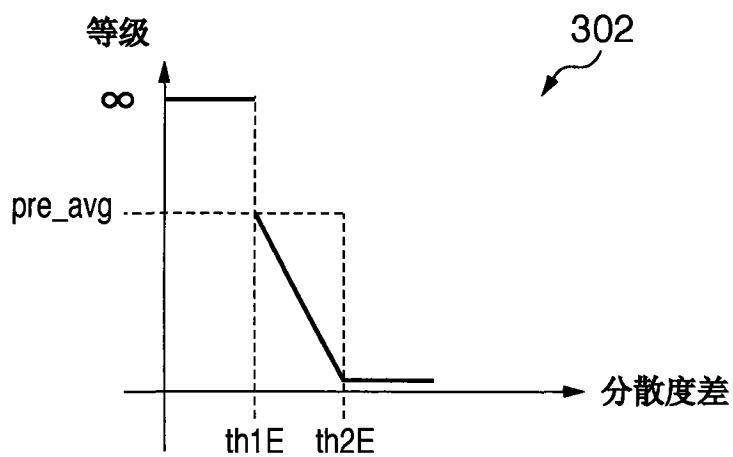


图 3B

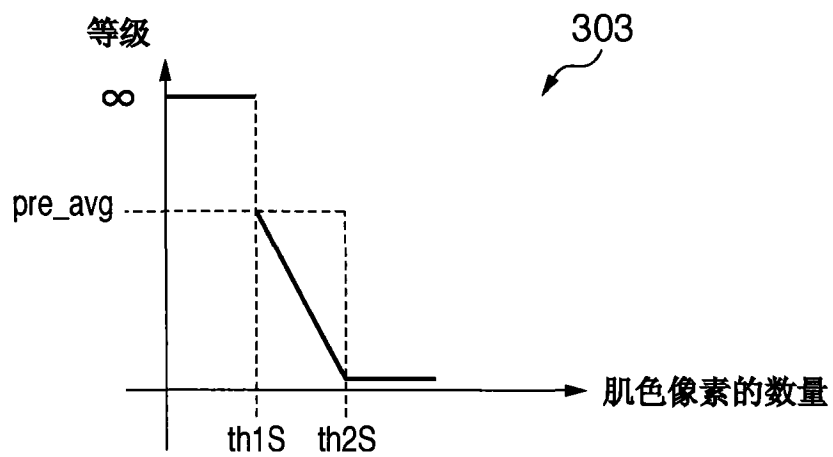


图 3C

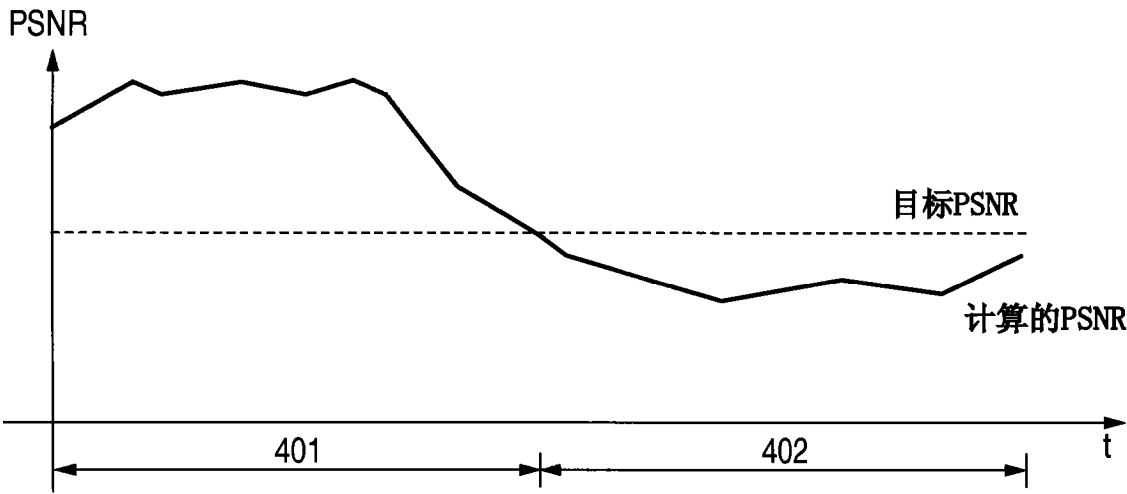


图 4

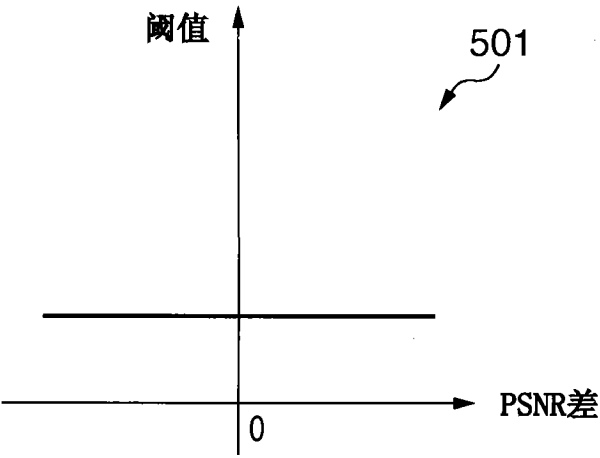


图 5A

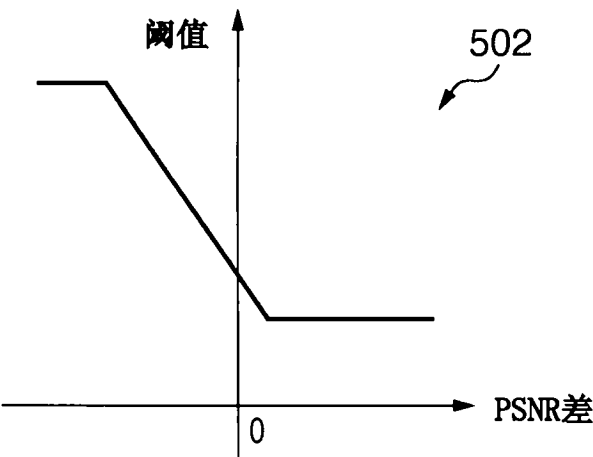


图 5B

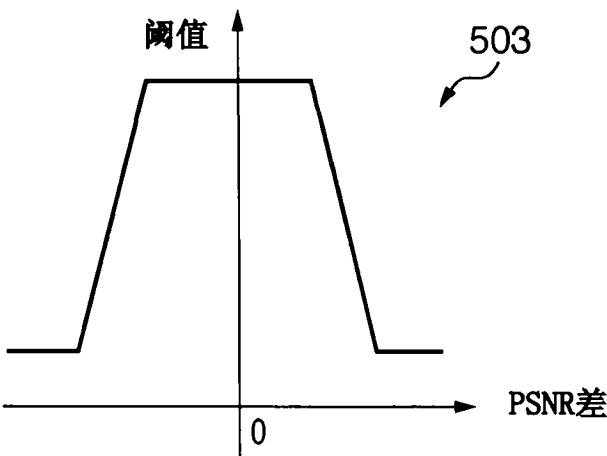


图 5C

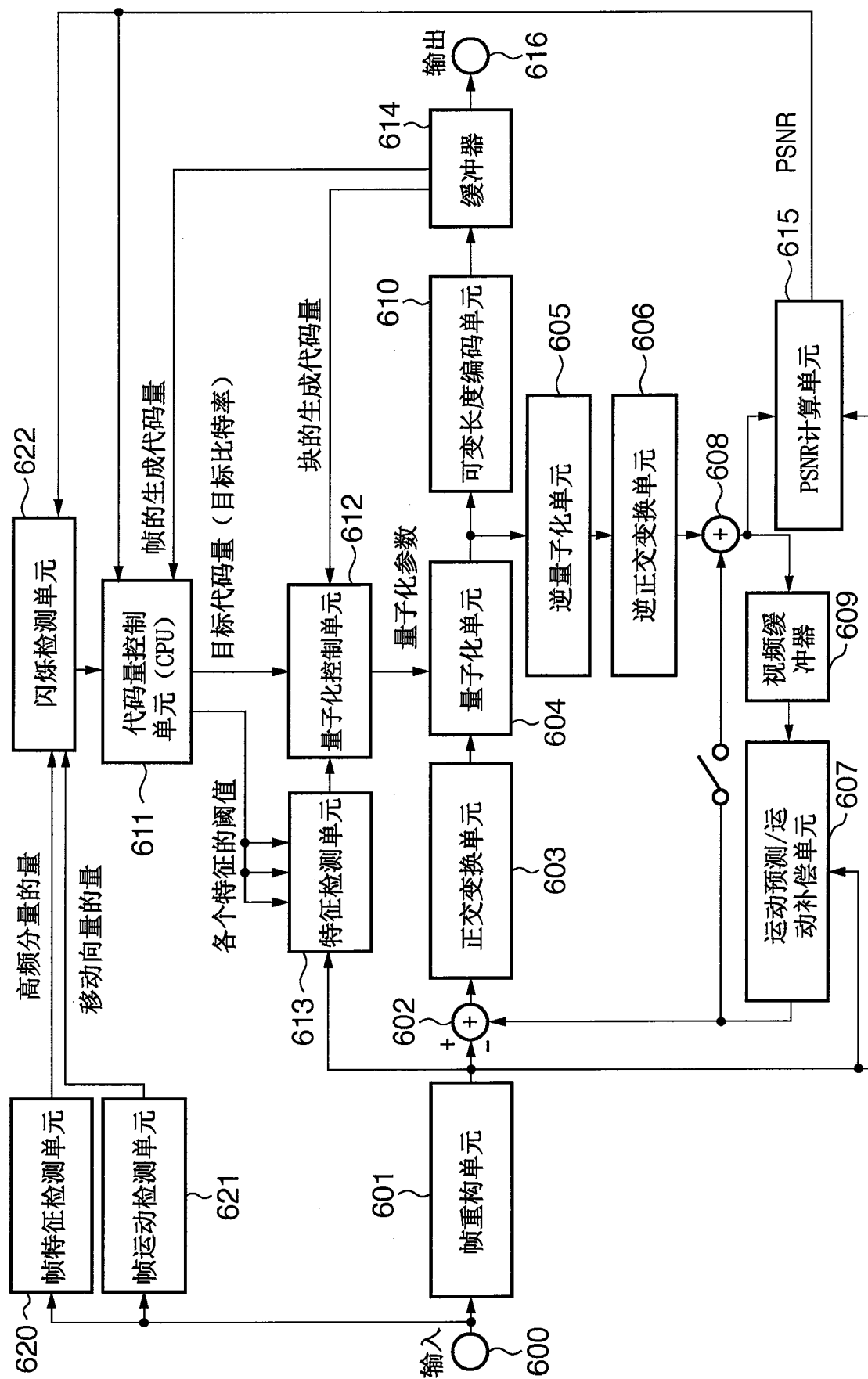


图 6

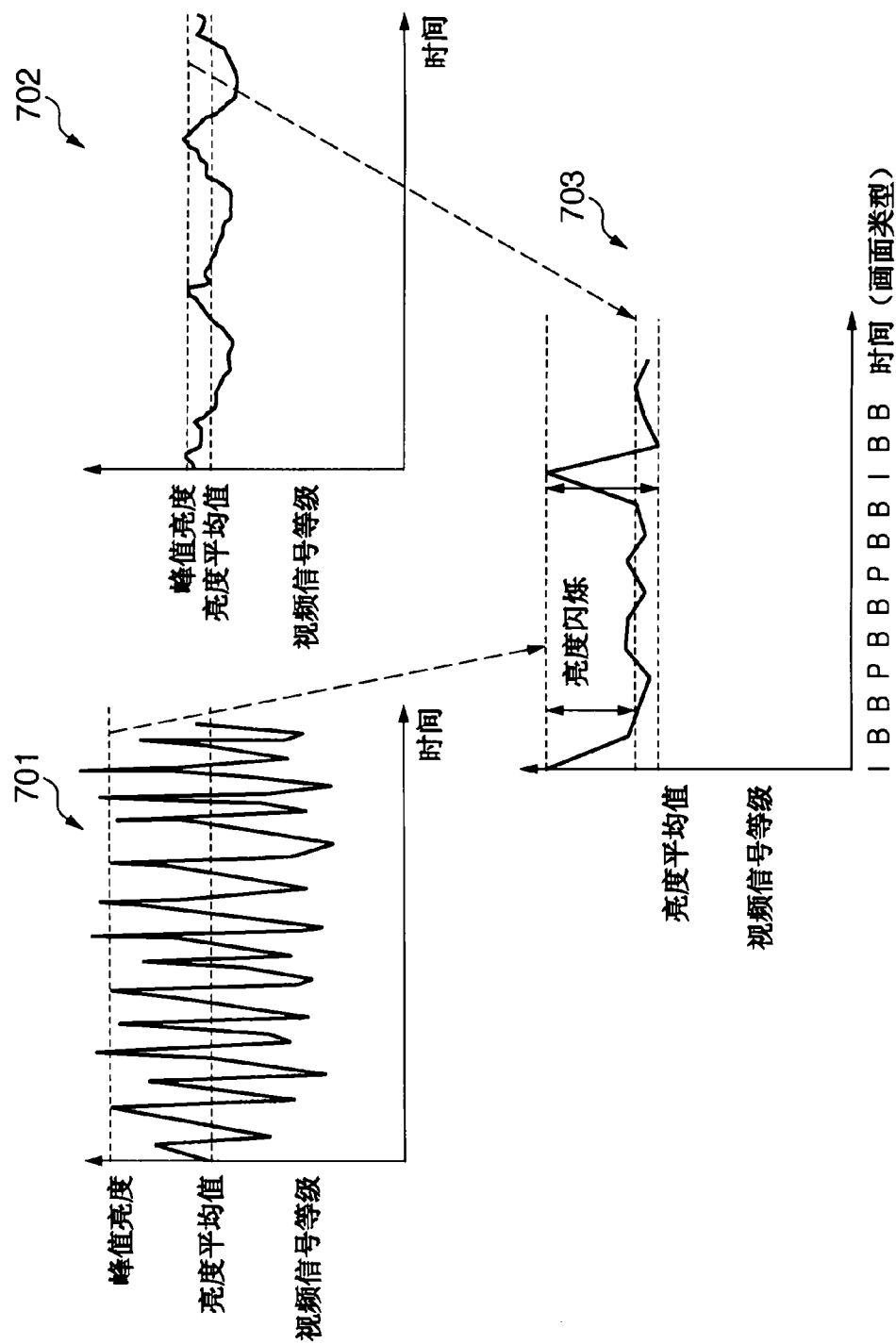


图7

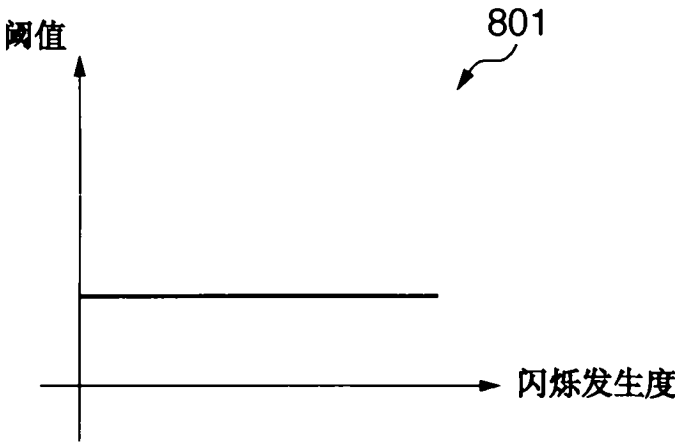


图 8A

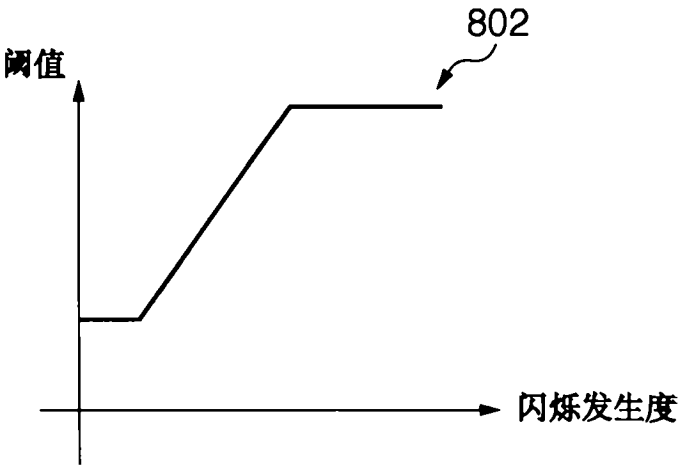


图 8B

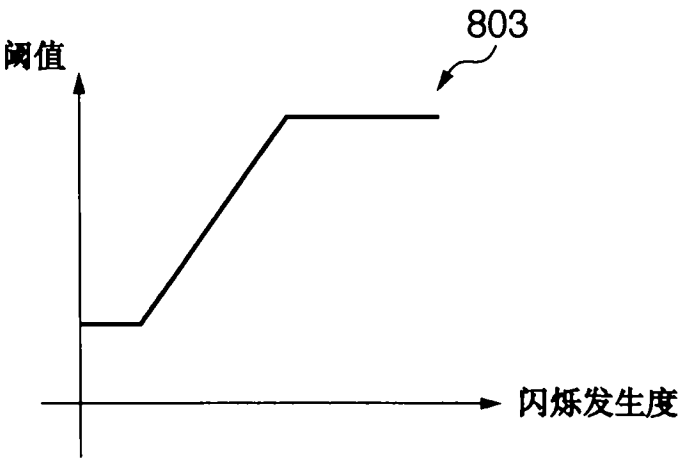


图 8C

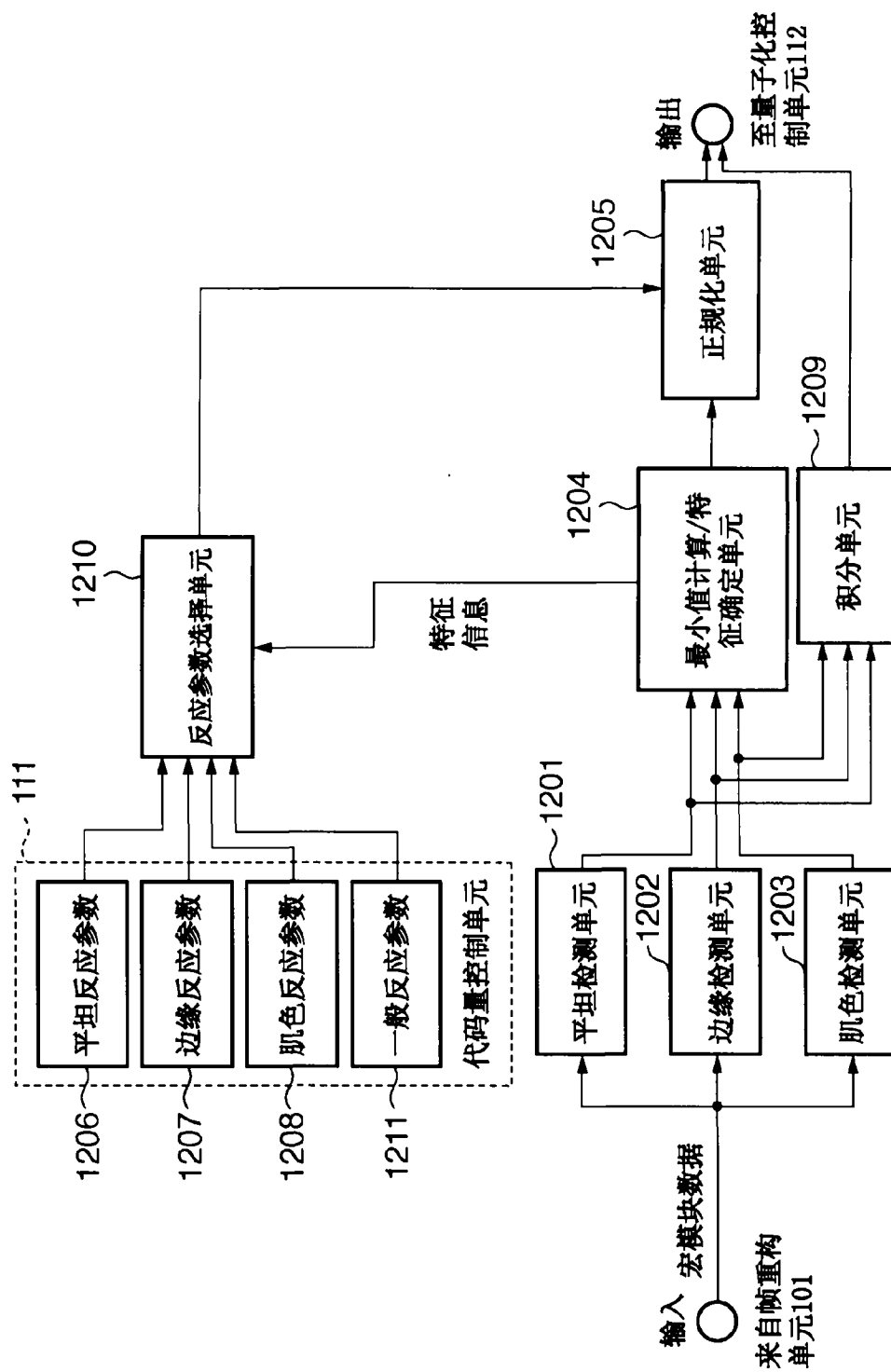


图9

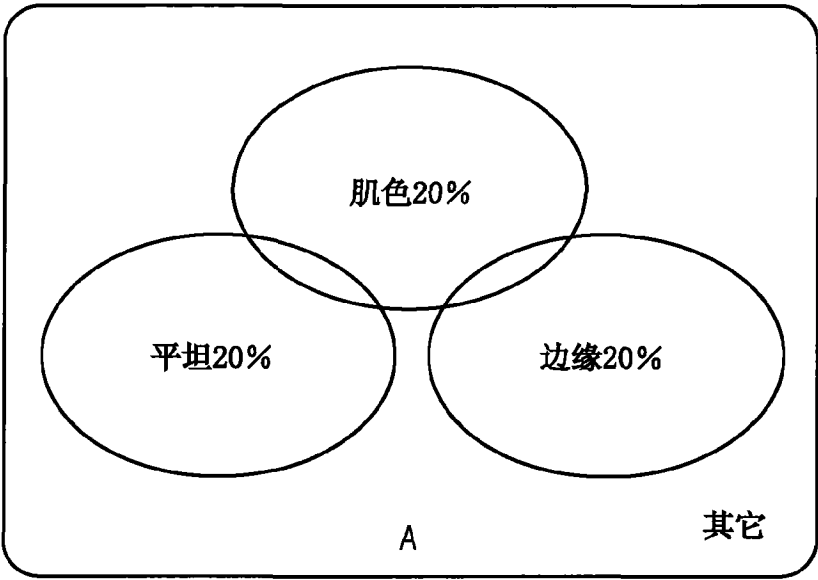


图 10A

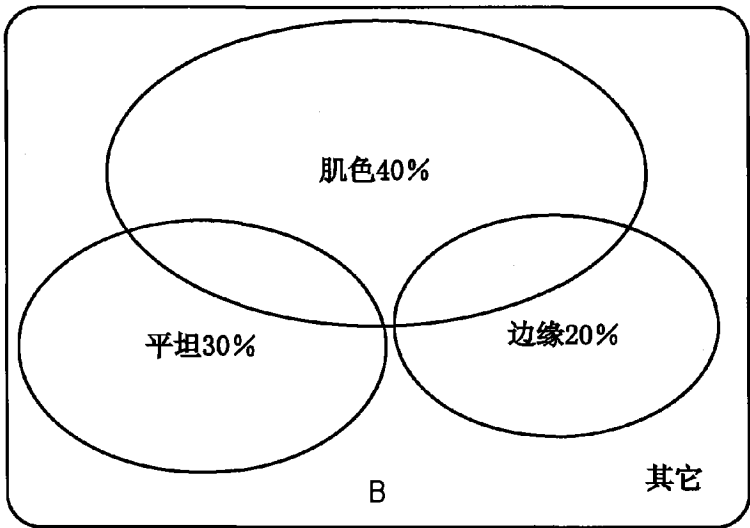


图 10B

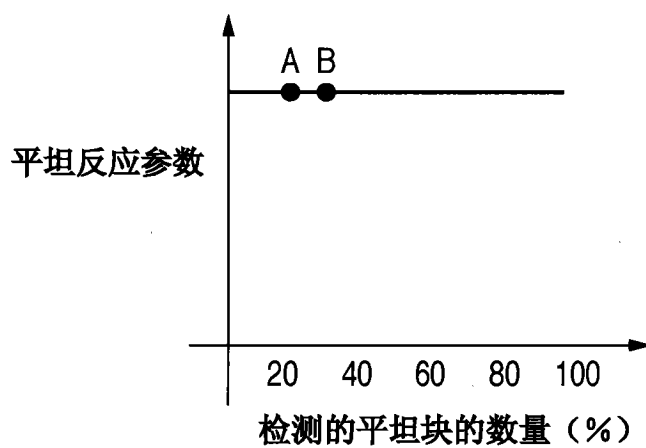


图 11A

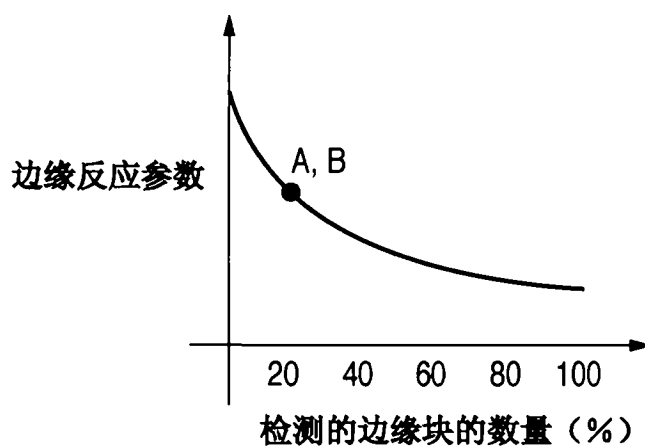


图 11B

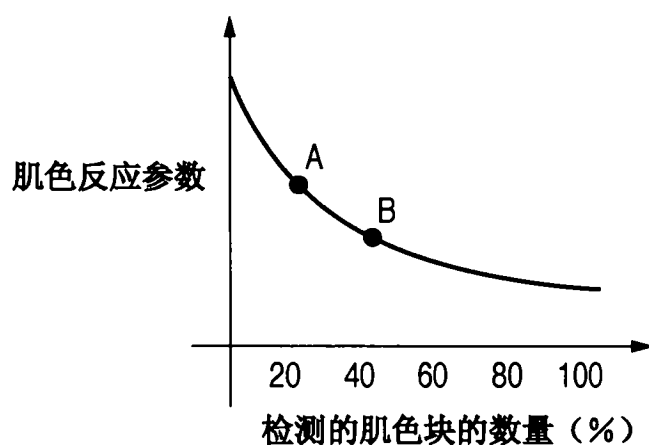


图 11C