

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

H04N 7/26

H04N 7/30



[12] 发明专利申请公开说明书

[21] 申请号 02815583.1

[43] 公开日 2005 年 3 月 9 日

[11] 公开号 CN 1593064A

[22] 申请日 2002.6.13 [21] 申请号 02815583.1

[30] 优先权

[32] 2001.6.15 [33] US [31] 09/882,753

[86] 国际申请 PCT/US2002/018962 2002.6.13

[87] 国际公布 WO2002/104039 英 2002.12.27

[85] 进入国家阶段日期 2004.2.9

[71] 申请人 高通股份有限公司

地址 美国加利福尼亚州

[72] 发明人 S·戈文德斯瓦米 K·塞加拉简

J·A·莱文

[74] 专利代理机构 上海专利商标事务所

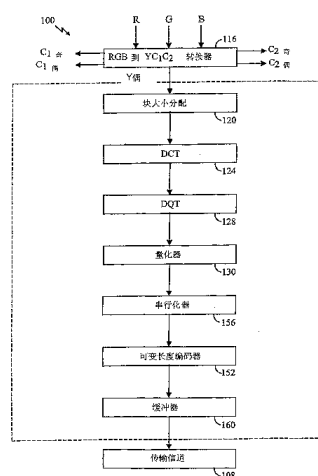
代理人 李家麟

权利要求书 5 页 说明书 12 页 附图 10 页

[54] 发明名称 可配置的图形优化器

[57] 摘要

描述了在数字影院系统中对基于频率的图像数据进行串行化的装置和方法。编辑了至少一组数据，它们可能被表示为 16×16 的数据块。或者，编辑了数据帧。该组数据被分成四组，每组都可能被表示为 8×8 块。四个 8×8 数据块的每一个都用 Z 字形扫描、垂直扫描以及/或者水平扫描来进行串行化。



1. 数字影院系统中，一种对基于频率的图像数据进行串行化的方法，所述方法包括：

编辑至少一组数据，它们可以被表示为一个 16×16 块；

把该组数据分成可被表示为 4 个 8×8 块的组；

对 4 个 8×8 数据块的每一个进行串行化。

2. 如权利要求 1 所述的方法，其特征在于，所述串行化包括对 4 个 8×8 数据块的每一个进行 Z 字形扫描。

3. 如权利要求 1 所述的方法，其特征在于，所述串行化包括对 4 个 8×8 数据块的每一个进行垂直扫描。

4. 如权利要求 1 所述的方法，其特征在于，所述串行化包括对 4 个 8×8 数据块的每一个进行水平扫描。

5. 如权利要求 1 所述的方法，其特征在于，所述编辑至少一个组包括：编辑一个数据帧，它可以被表示为多个 16×16 块。

6. 如权利要求 1 所述的方法，其特征在于，所述基于频率的图像数据被分成 Y、Cb 和 Cr 彩色分量。

7. 如权利要求 6 所述的方法，其特征在于，所述 Y、Cb 和 Cr 彩色分量进一步被分成偶的和奇的彩色分量。

8. 数字影院系统中，一种压缩数字图像的方法，所述图像包括像素数据，所述像素数据被分成彩色分量，所述方法包括以下动作：

读取像素数据的一组彩色分量；

产生一块大小分配，以把像素的该组彩色分量分成像素数据的子块；

把像素数据的子块变换成相应的频域表示；以及

把频域表示按比例缩放成数据流，其中缩放动作是基于与图像质量相关的质量度量；

从数据流编辑至少一组数据，它们可能被表示为一个 16×16 块；

把该 16×16 的数据组分成可被表示为 4 个 8×8 块的组；以及

对 4 个 8×8 数据块的每一个进行串行化。

9. 如权利要求 8 所述的方法，其特征在于，所述按比例缩放动作还包括：向所述像素数据的子块提供频率加权的掩模的动作，使得频率加权掩模强调对人类视觉系统更为敏感的图像部分，并且较不强调对人类视觉系统较不敏感的图像部分。

10. 如权利要求 8 所述的方法，其特征在于，所述按比例缩放动作还包括：根据图像质量对像素数据的子块进行量化的动作。

11. 如权利要求 8 所述的方法，其特征在于，所述质量度量是信噪比。

12. 如权利要求 8 所述的方法，其特征在于，所述变换动作执行离散余弦变换。

13. 如权利要求 8 所述的方法，其特征在于，所述变换动作执行离散余弦变换，其后是差分四叉树变换。

14. 如权利要求 8 所述的方法，其特征在于，所述彩色分量是 Y、Cb 和 Cr 彩色分量。

15. 如权利要求 14 所述的方法，其特征在于，所述 Y、Cb 和 Cr 彩色分量被分成偶的和奇的彩色分量。

16. 数字影院系统中，一种对基于频率的图像数据进行串行化的装置，所述装置包括：

编辑至少一组数据的装置，该组数据可以表示为一个 16×16 的块；

把该组数据分成可被表示为 4 个 8×8 块的组的装置；

对 4 个 8×8 数据块的每一个进行串行化的装置。

17. 如权利要求 16 所述的装置，其特征在于，所述用于串行化的装置包括对 4 个 8×8 数据块的每一个进行 Z 字形扫描的装置。

18. 如权利要求 16 所述的装置，其特征在于，所述用于串行化的装置包括对 4 个 8×8 数据块的每一个进行垂直扫描的装置。

19. 如权利要求 16 所述的装置，其特征在于，所述用于串行化的装置包括对 4 个 8×8 数据块的每一个进行水平扫描的装置。

20. 如权利要求 16 所述的装置，其特征在于，所述编辑至少一个组的装置包括：编辑被表示为多个 16×16 块的数据帧的装置。

21. 如权利要求 16 所述的装置，其特征在于，所述基于频率的图像数据被分成 Y、Cb 和 Cr 彩色分量。

22. 如权利要求 21 所述的装置，其特征在于，所述 Y、Cb 和 Cr 彩色分量进一步被分成偶的和奇的彩色分量。

23. 数字影院系统中，一种压缩数字图像的装置，所述图像包括像素数据，所述像素数据被分成彩色分量，所述装置包括：

用于读取一组像素数据的装置；

用于产生一块大小分配来把该组像素分成像素数据的子块的装置；

用于把像素数据的子块变换成相应的频域表示的装置；以及

用于把频域表示按比例缩放成数据流的装置，其中缩放动作是基于与图像质量相关的质量度量；

用于从数据流编辑至少一组数据的装置，该组数据可能被表示为一个 16×16 块；

用于把该 16×16 的数据组分成可被表示为 4 个 8×8 块的组的装置；以及

用于对 4 个 8×8 数据块的每一个进行串行化的装置。

24. 如权利要求 23 所述的装置, 其特征在于, 所述变换动作执行离散余弦变换。

25. 如权利要求 23 所述的装置, 其特征在于, 所述变换动作执行离散余弦变换, 其后是差分四叉树变换。

26. 如权利要求 23 所述的装置, 其特征在于, 所述彩色分量是 Y、Cb 和 Cr 彩色分量。

27. 如权利要求 26 所述的装置, 其特征在于, 所述 Y、Cb 和 Cr 彩色分量被分成偶的和奇的彩色分量。

28. 数字影院系统中, 一种用于对基于频率的图像数据进行串行化的装置, 所述装置包括:

编辑器, 用于编辑可被表示为 16×16 块的至少一组数据;

分割器, 用于把该组数据分成可被表示为 4 个 8×8 块的组;

串行化器, 用于对 4 个 8×8 数据块的每一个进行串行化。

29. 如权利要求 28 所述的装置, 其特征在于, 所述串行化器进一步包括一 Z 字形扫描器, 用于对 4 个 8×8 数据块的每一个进行 Z 字形扫描。

30. 如权利要求 28 所述的装置, 其特征在于, 所述串行化器进一步包括一垂直扫描器, 用于对 4 个 8×8 数据块的每一个进行垂直扫描。

31. 如权利要求 28 所述的装置, 其特征在于, 所述串行化器进一步包括一水平扫描器, 用于对 4 个 8×8 数据块的每一个进行水平扫描。

32. 如权利要求 28 所述的装置, 其特征在于, 所述编辑器用于对表示为多个 16×16 块的数据帧进行编辑。

33. 如权利要求 28 所述的装置, 其特征在于, 所述基于频率的图像数据被分成 Y、Cb 和 Cr 彩色分量。

34. 如权利要求 33 所述的装置, 其特征在于, 所述 Y、Cb 和 Cr 彩色分量进一步被分成偶的和奇的彩色分量。

35. 数字影院系统中, 一种用于压缩数字图像的装置, 所述图像包括像素数据, 所述装置包括:

阅读器, 用于读取一组像素数据;

发生器, 用于产生一块大小分配, 以把像素组分成像素数据的子块;

变换器, 用于把像素数据的子块变换成相应的频域表示;

比例缩放器, 用于把频域表示按比例缩放成数据流, 其中缩放动作是基于与图像质量相关的质量度量;

编辑器, 用于从数据流中编辑至少一组数据, 该组数据可被表示为一个 16×16 块;

分割器, 用于把该 16×16 的数据组分成可表示为 4 个 8×8 块的组; 以及
串行化器, 用于对 4 个 8×8 数据块的每一个进行串行化。

可配置的图形优化器

发明背景

I. 发明领域

本发明涉及图像处理和压缩，尤其涉及压缩图像的可配置的图形优化程序。

II. 相关技术描述

数字图像处理在数字信号处理领域内具有显著地位。人类视觉的重要性已经促进了数字图像处理的领域和学科内的极大兴趣和进展。在视频信号的发送和接收领域，譬如那些用于投影胶片或电影的领域，对图像压缩作出了各种改进。许多当前和提议的视频系统利用了数字编码技术。该领域的方面包括图像编码、图像恢复、以及图像特征选择。图像编码代表试图以有效的方式发送数字通信信道的图片，利用尽可能少的比特来使所需的带宽最小，而同时把失真维持在某些限制内。图像恢复代表努力恢复对象的真实图像。在通信信道上被发送的已编码图像已经被各种因素而失真。降级的来源可能在最初从对象创建图像时已引起。特征选择是指选择图片的某些属性。较广泛环境中的识别、分类以及判决可能要求这种属性。

视频的數字编码，譬如数字影院内的编码，是受益于改进的图像压缩技术的领域。数字图像压缩可能一般被分为两类：无损和有损的方法。无损图像无须任何信息损失而被恢复。有损的方法包括不能恢复某些信息，这取决于压缩率、压缩算法的质量、以及算法的实现。一般而言，有损压缩方法被认为是获得效能成本划算的数字影院方法所期望的压缩率。为了达到数字影院质量级别，压缩方法应该提供在视觉上无损的性能级别。这样，尽管由于压缩过程而有数学的信息损失，然而该损失引起的图像失真对于正常观察条件下的观察者来说应该是难以察觉的。

已经为其它应用研发了现有的数字图像压缩技术，也就是为了电视系统。这种技术已作出了设计折衷以适用于想要的应用，但不满足影院演出所需的质量要求。

数字影院压缩技术应该提供常看电影的人先前体验过的视觉质量。理想的是，数字影院的视觉质量应该试图超越高质量版本的拷贝胶卷。与此同时，压缩技术应该具有实用的高编码效率。如这里所定义的，编码效率是指使被压缩的图像质量满

足一定质量级别所需的比特率。此外，系统和编码技术应该具有内部灵活性来容纳不同格式，并且应该是效能成本划算的；即，小型且有效的解码器或编码器处理。

许多可用的压缩技术提供了显著的压缩级别，但是导致视频信号质量的降级。一般而言，用于传递被压缩的信息的技术要求以恒定的比特率传递被压缩的信息。

能提供显著压缩级别同时保持期望的视频信号质量的一种压缩技术使用了已编码的离散余弦变换(DCT)系数数据的自适应大小的块和子块。该技术下文中被称为自适应块大小离散余弦变换(ABSDCT)方法。该技术在美专利号 5021891 中公开，该专利题为“自适应块大小图像压缩方法和系统(Adaptive Block Size Image Compression Method and System)”，它被转让给本发明的受让人并且通过引用被结合于此。DCT 技术也在美专利号 5107345 中公开，该专利题为“自适应块大小图像压缩方法和系统(Adaptive Block Size Image Compression Method and System)”，它被转让给本发明的受让人并且通过引用被结合于此。而且，美专利号 5452104 中讨论了 ABSDCT 技术与差分二叉树变换技术的组合，该专利题为“自适应块大小图像压缩方法和系统(Adaptive Block Size Image Compression Method and System)”，它也被转让给本发明的受让人并且通过引用被结合于此。这些专利中公开的系统使用了称为“帧内”编码的编码，其中图像数据的每帧都被编码，而不考虑其它帧的内容。通过使用 ABSDCT 技术，可达到的数据速率可能从约每秒 1.5 千兆比特降低到约每秒 50 兆比特，而没有信号质量的可辨别的降级。

ABSDCT 技术可能用于压缩黑色和白色或彩色图像，或者表示该图像的信号。彩色输入信号可能是 YIQ 格式，Y 是亮度或辉度采样，I 和 Q 是每个 4×4 像素块的色度(即彩色)采样。也可能使用其它已知格式，譬如 YUV、 $YCbCr$ 或 RGB 格式。由于眼睛对彩色的低空间敏感性，大多数研究表明，水平和垂直方向上以 4 倍对彩色分量的子采样是合理的。因而，视频信号可以用 4 个亮度分量和 2 个色度分量来表示。

通过使用 ABSDCT，视频信号一般被分段成像素块用于处理。对于每个块而言，亮度和色度分量被传递至块交织器。例如， 16×16 (像素)块可以被提供给块交织器，块交织器排序或组织每个 16×16 块内的采样，以产生数据的块和复合子块，用于离散余弦变换(DCT)分析。DCT 运算符是把时间采样的信号转换成相同信号的频率表示的一种方法。通过转换成频率表示，已经示出 DCT 技术允许非常高的压缩程度，因为可以设计量化器利用图像的频率分布特性。在优选实施例中，一个 16×16 的 DCT 应用于第一次排序，四个 8×8 的 DCT 应用于第二次排序，16 个 4×4 的 DCT

应用于第三次排序，而 64 个 2×2 的 DCT 应用于第四次排序。

DCT 操作降低了视频源内固有的空间冗余。在执行了 DCT 之后，大多数视频信号能量会集中在少数的 DCT 系数中。可能使用附加变换，差分二叉树变换(DQT)，来降低 DCT 系数间的冗余。

对于 16×16 块以及每个子块而言，分析 DCT 系数值和 DQT 值(如果使用了 DQT)来确定对块或子块进行编码所需的比特数。然后，选择要求最少比特数的块或子块的组合来编码以表示图像分段。例如，可能选择 2 个 8×8 子块、6 个 4×4 子块和 8 个 2×2 子块来表示图像分段。

然后，适当地顺序排列所选的块或子块的组合，组成一个 16×16 块。在准备传输时，DCT/DQT 系数值接着可能经受频率加权、量化和编码(譬如可变长度编码)。尽管上述 ABSDCT 技术执行得非常好，然而它是计算密集的。因此，该技术的小型硬件实现可能很难。

可变长度编码以游程长度和大小的形式来完成。其它压缩方法，譬如联合图像专家组(JPEG)或运动图像专家组(MPEG-2，在整个被处理的块大小上使用了标准的 Z 字形扫描方法。然而，通过使用 ABSDCT，根据数据块内的方差产生不同的块大小。因而，在整个被处理的块大小上，标准的 Z 字形扫描方法不总是最佳的。而且，每个块大小上的标准 Z 字形扫描方法在硬件上可能难以实现。此外，Z 字形图形对于给定的块或帧来说不总是最佳图案。因此，需要一种方法和装置来确定最佳图案。

发明概述

本发明的实施例提供了用于最佳图案确定器的装置和方法。一实施例中，最佳图案可以逐帧地配置。另一实施例中，使用了预定块大小的缺省图案，不考虑由自适应块大小离散余弦变换(ABSDCT)技术所确定的实际块大小。

本发明是图像压缩的基于质量的系统和方法，它使用了离散余弦变换系数数据的自适应大小的块和子块以及基于质量的量化标度因子。一像素数据块被输入编码器。编码器包括一块大小分配(BSA)元件，它对输入的像素块分段以用于处理。块大小分配是基于输入块和进一步细分的块的方差。通常，具有较大方差的区域被细分成较小的块，具有较小方差的区域不被细分，只要块和子块的均值落在不同的预定范围内。这样，首先，块的方差阈值根据其均值从其标称值被修改，然后把块的方差与阈值相比较，如果方差大于阈值，块就被细分。

向变换元件提供了块大小分配, 变换元件把像素数据变换成频域数据。仅对通过块大小分配而选择的块和子块执行变换。然后, 变换数据通过量化和串行化而经受按标度变化。变换数据的量化是基于图像质量度量的, 譬如调节对比度的标度因子、系数计数、码率失真、块大小分配的密度和/或过去的标度因子。串行化是基于创建相同值的最长可能的游程长度。一实施例中, 使用固定块大小的 Z 字形扫描来对数据进行串行化, 以产生数据流, 而不考虑块大小分配。另一实施例中, 块大小为 8×8 。在准备传输时, 可以用可变长度编码器对数据流进行编码。已编码的数据通过传输信道被发送到解码器, 在解码器处为了准备显示而重建像素数据。

另一实施例中, 描述了数字影院系统中对基于频率的图像数据进行串行化的方法。至少一组数据被编辑, 它们可能被表示为一个 16×16 的数据块。或者, 一个数据帧被编辑。该组数据被分成四组, 每组都被表示为 8×8 的块。4 个 8×8 数据块的每一个都用 Z 字形扫描、垂直扫描和/或水平扫描来进行串行化。

因而, 实施例的一方面是在 8×8 块内使用固定的扫描图形来处理数据块, 不考虑实际的块大小分配。

实施例的另一方面是逐帧地确定并实现最佳扫描技术。

实施例的另一方面是向用户提供可配置的扫描图形。

附图简述

通过下面提出的结合附图的详细描述, 本发明的特征、性质和优点将变得更加明显, 附图中相同的元件具有相同的标识, 其中:

图 1 是基于质量的图像处理系统的编码器部分的框图, 该系统结合了基于方差的块大小分配系统以及本发明的方法;

图 2 是基于质量的图像处理系统的解码器部分的框图, 该系统结合了基于方差的块大小分配系统以及本发明的方法;

图 3 是说明基于方差的块大小分配中所涉及的处理步骤的流程图;

图 4a 说明了示例性的块大小分配;

图 4b 说明了贯穿 16×16 块大小的 Z 字形扫描图形;

图 4c 说明了每个可变块大小内的 Z 字形扫描图形;

图 5a 说明了独立于实际块大小的 8×8 块的 Z 字形扫描图形;

图 5b 说明了独立于实际块大小的 8×8 块内实现的不同扫描图形;

图 6a 说明了串行化过程的实施例; 以及

图 6b 说明了串行化过程的另一实施例。

优选实施例的详细描述

为了便于数字信号的数字传输并且享受相应的好处，一般需要采用某些形式的信号压缩。为了实现产生图像时的高压缩，维持图像的高质量也很重要。而且，对于微型硬件的实现计算效率是期望的，这在许多应用中都是重要的。

在详细说明本发明一实施例之前，应该理解本发明不限于下面描述中提出的或者附图中所述的组件的结构和排列细节。本发明能有其它实施例，并且以各种方式实现。而且，可以理解，这里所用的措辞和术语都是为了描述的目的，而不应被视为是限制性的。

一实施例的一方面所采用的图像压缩是基于离散余弦变换(DCT)技术，譬如共同待批的美国专利申请“Contrast Sensitive Variance Based Adaptive Block Size DCT Image Compression”中公开的技术，该申请序列号为 09/436085，于 1999 年 11 月 8 日提交，被转让给本发明的受让人并且通过引用被结合于此。一般而言，要在数字域内被处理的图像会由像素数据组成，像素数据被分成不重叠块的阵列，大小为 $N \times N$ 。对每个块执行二维 DCT。二维 DCT 由下列关系式所定义：

$$X(k,l) = \frac{\alpha(k)\beta(l)}{\sqrt{N \cdot M}} \sum_{m=0}^{N-1} \sum_{n=0}^{M-1} x(m,n) \cos\left[\frac{(2m+1)\pi k}{2N}\right] \cos\left[\frac{(2n+1)\pi l}{2M}\right], 0 \leq k, l \leq N-1$$

其中 $\alpha(k), \beta(k) = \begin{cases} 1, & \text{如果 } k = 0 \\ \sqrt{2}, & \text{如果 } k \neq 0 \end{cases}$ ，以及

$x(m,n)$ 是 $N \times M$ 块内的像素位置 (m,n) ，以及

$X(k,l)$ 是相应的 DCT 系数。

由于像素值是非负的，因此 DCT 分量 $X(0,0)$ 总是正的，并且通常具有最多能量。实际上，对于一般图像而言，大多数变换能量都集中在分量 $X(0,0)$ 附近。该能量紧缩属性使 DCT 技术成为具有吸引力的压缩方法。

图像压缩技术使用了对比度自适应编码来实现进一步的比特率降低。已经观察到，大多数自然图像的组成是：平的相对缓慢变化的区域、以及像对象边界和高对比度纹理这样的繁忙区域。对比度自适应编码方案通过向繁忙区域分配更多比特而向较不繁忙的区域分配较少比特而利用了该因素。

对比度自适应方法使用了帧内编码(空间处理)而不是帧间编码(空间—时间处理)。除了复杂的处理电路以外，帧间编码本身要求多个帧缓冲器。在许多应用

中,实际实现需要降低的复杂度。在会使空间一时间编码方案崩溃且性能表现极差的情况下,帧内编码也是有用的。例如,由于机械快门所产生的积分时间相对较短,因此每秒 24 帧的影片会落在这个类别。短积分时间允许较高的时间混叠程度。帧帧相关的假设对于快速运动而破坏,因为它变得急动。当包含了 50Hz 和 60Hz 两个电源线频率时,帧内编码也容易标准化。电视目前以 50Hz 或 60Hz 发送信号。帧内方案作为数字方法的使用可以适用于 50Hz 和 60Hz 操作两者,或者通过折衷帧速率相对空间分辨率,甚至适用于每秒 24 帧的影片。

为了图像处理目的,在被分成不重叠块的阵列的像素数据上执行 DCT 操作。注意到,尽管这里讨论的块大小尺寸为 $N \times N$,然而可以预想,可以使用各种块大小。例如,可以使用 $N \times M$ 的块大小,其中 N 和 M 都是整数, M 或大于或小于 N 。另一重要方面是块可被分成至少一层子块,譬如 $N/i \times N/i$ 、 $N/i \times N/j$ 、 $N/i \times M/j$ 等等,其中 i 和 j 是整数。而且,这里讨论的示例性块大小是 16×16 的像素块,具有 DCT 系数的相应块和子块。还可以预想到,可以使用各种其它整数值,譬如既可以是偶数又可以是奇数,如, 9×9 。

图 1 和图 2 说明了图像处理系统 100,该系统结合了可配置串行化器的概念。图像处理系统 100 包括一编码器 104,它压缩接收到的视频信号。被压缩的信号使用传输信道或物理媒介 108 来发送,并且被解码器 112 所接收。解码器 112 把接收到的已编码数据解码成图像采样,后者可以被显示。

通常,图像被分成多个像素块用于处理。彩色信号可能用 RGB 到 YC_1C_2 转换器 116 而从 RGB 空间被转换到 YC_1C_2 空间,其中 Y 是亮度分量, C_1 和 C_2 是色度(即彩色)分量。由于人眼对彩色的低空间敏感性,因此许多系统在水平和垂直方向上以 4 倍对 C_1 和 C_2 分量进行子采样。然而,子采样不是必要的。称为 4:4:4 的全分辨率图像在某些应用中可能非常有用或必要,譬如在覆盖“数字影院”的应用中。两个可能的 YC_1C_2 表示是, YIQ 表示和 YUV 表示,两者都是本领域公知的。也可能采用 YUV 表示的变体,称为 $YCbCr$ 。这可能进一步被分裂成奇和偶分量。因而,一实施例中,使用了表示 Y -偶、 Y -奇、 Y 、 Cb -偶、 Cb -奇、 Cr -偶、 Cr -奇。

在一优选实施例中,奇和偶 Y 、 Cb 、 Cr 分量的每一个都不作子采样地被处理。这样, 16×16 像素块 6 个分量的每一个输入被提供给编码器 104。为了说明目的,说明了 Y -偶分量的编码器 104。对于 Y -奇分量、以及奇的和偶的 Cb 和 Cr 分量使用了类似的编码器。编码器 104 包括一块大小分配元件 120,它在准备视频压缩时执行块大小分配。块大小分配元件 120 根据块内图像可观察的特性来确定 16×16

块的块分解。块大小分配根据 16×16 块内的活动性而以四叉树方式把每个 16×16 块细分成较小的块，譬如 8×8 、 4×4 和 2×20 。块大小分配元件 120 产生一四叉树数据，称为 PQR 数据，它的长度在 1 到 21 比特之间。因此，如果块大小分配元件确定要分解 16×16 块，则设定 PQR 数据的 R 比特，R 比特后是 Q 数据的 4 个附加比特，它们对应于 4 个分解后的 8×8 块。如果块大小分配确定要细分任一 8×8 块，则添加每个被细分的 8×8 块的 P 数据的 4 个附加比特。

现在参照图 3，提供了一流程图，示出块大小分配元件 120 的操作细节。使用块的方差作为要细分块的判决的度量。从步骤 202 开始，读取一个 16×16 的像素块。步骤 204 中，计算该 16×16 块的方差 v_{16} 。方差计算如下：

$$\text{var} = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_{i,j}^2 - \left(\frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_{i,j} \right)^2$$

其中 $N=16$ ， $x_{i,j}$ 是 $N \times N$ 块内的第 i 行、第 j 列。步骤 206 中，如果块的均值在两个预定值之间，那么首先修改方差阈值 T_{16} 来提供新的阈值 T'_{16} ，然后把块方差与新阈值 T'_{16} 相比较。

如果方差 v_{16} 不大于阈值 T_{16} ，则在步骤 208 中， 16×16 块的起始地址被写入临时存储器，然后把 PQR 数据的 R 比特设为 0 以指明不细分 16×16 块。然后，算法读取下一个 16×16 像素块。如果方差 v_{16} 大于阈值 T_{16} ，则在步骤 210 中，把 PQR 数据的 R 比特设为 1 以指明 16×16 块要被细分成 4 个 8×8 块。

如步骤 212 所示，4 个 8×8 块 ($i=1:4$ 被顺序地考虑用于进一步的细分。对于每个 8×8 块而言，在步骤 214 中计算方差 v_{8_i} 。步骤 216 中，如果块的均值在两个预定值之间，那么首先修改方差阈值 T_8 来提供新的阈值 T'_8 ，然后把块方差与该新阈值相比较。

如果方差 v_{8_i} 不大于阈值 T_8 ，则在步骤 218 中， 8×8 块的起始地址被写入临时存储器，并且把相应的 Q 比特 Q_i 设为 0。然后处理下一个 8×8 块。如果方差 v_{8_i} 大于阈值 T_8 ，则在步骤 220 中，把相应的 Q 比特 Q_i 设为 1 以指明 8×8 块要被细分成 4 个 4×4 块。

如步骤 222 所示，4 个 4×4 块 ($j_i=1:4$ 被顺序地考虑用于进一步的细分。对于每个 4×4 块而言，在步骤 224 中计算方差 $v_{4_{ij}}$ 。步骤 226 中，如果块的均值在两个预定值之间，那么首先修改方差阈值 T_4 来提供新的阈值 T'_4 ，然后把块方差与该新阈值相比较。

如果方差 $v_{4_{ij}}$ 不大于阈值 T_4 ，则在步骤 228 中，写入 4×4 块的起始地址，并且把相应的 P 比特 P_{ij} 设为 0。然后处理下一个 4×4 块。如果方差 $v_{4_{ij}}$ 大于阈值 T_4 ，则在步骤 230 中，把相应的 P 比特 P_{ij} 设为 1 以指明 4×4 块要被细分成 4 个 2×2 块。此外，4 个 2×2 块的地址被写入临时存储器。

阈值 T_{16} 、 T_8 和 T_4 可能是预定的常数。这被称为硬判决。或者，可能实现自适应判决即软判决。软判决根据 $2N \times 2N$ 块的平均像素值改变方差的阈值，其中 N 可以是 8、4 或 2。因而，平均像素值的函数可以被用作阈值。

为了说明目的，考虑以下示例。令 Y 分量的预定方差阈值对于 16×16 、 8×8 以及 4×4 块分别为 50、1100 和 880。换言之， $T_{16}=50$ ， $T_8=1100$ ，且 $T_4=880$ 。令平均值的范围为 80 到 100。假定 16×16 块所计算的方差为 60。由于 60 大于 T_{16} ，且均值 90 在 80 和 100 之间，因此 16×16 块被细分成 4 个 8×8 子块。假定为 8×8 子块计算的方差为 1180、935、980 和 1210。由于两个 8×8 块具有超出 T_8 的方差，因此这两个块再被细分以产生总的 8 个 4×4 子块。最终，假定 8 个 4×4 块的方差为 620、630、670、610、590、525、930 和 690，相应的均值为 90、120、110、115。由于第一个 4×4 的块的均值落在范围 (80, 100) 内，因此其阈值会被降低到 $T'_4=200$ ，这低于 880。所以，该 4×4 块会与第七个 4×4 块一样被细分。所产生的块大小分配在图 4a 中示出。图 4b 中示出相应的二叉树分解。此外，图 4c 中示出由这个块大小分配所产生的 PQR 数据。

注意到使用了类似的步骤来为彩色分量 Y-奇、 C_b -偶、 C_b -奇、 C_r -偶和 C_r -奇分配块大小。彩色分量可以水平、垂直或在两个方向上被抽取。

此外，注意到尽管已经把块大小分配描述成自顶向下的方法，其中首先评估最大的块(本例中是 16×16 ，然而也可以取代地使用自底向上的方法。自底向上方法会先评估最小的块(本例中为 2×2)。

回过头参照图 1，PQR 数据以及所选块的地址被提供给 DCT 元件 124。DCT 元件 124 使用 PQR 数据对所选的块执行适当大小的离散余弦变换。仅仅所选的块需要经历 DCT 处理。

图像处理系统 100 还包括 DQT 元件 128，用于降低 DCT 的 DC 系数中间的冗余度。在每个 DCT 块的左上角遇到 DC 系数。通常，DC 系数相比 AC 系数是大的。大小差异使得难以设计有效的可变长度编码器。因而，降低 DC 系数间的冗余是有利的。

DQT 元件 128 对 DC 系数执行二维 DCT，每次取 2×2 。从 4×4 块内的 2×2 块

开始, 对 4 个 DC 系数执行二维 DCT。该 2×2 DCT 被称为 4 个 DC 系数的差分二叉树变换, 即 DQT。接着, 使用 8×8 块内 DQT 的 DC 系数以及 3 个相邻的 DC 系数来计算下一级 DQT。最终, 使用 16×16 块内 4 个 8×8 块的 DC 系数来计算 DQT。这样, 在 16×16 块内, 有一个真实的 DC 系数, 其余是对应于 DCT 和 DQT 的 AC 系数。

变换系数(DCT 和 DQT 两者)被提供给量化器用于量化。在优选实施例中, 用频率加权掩模(FWM)和量化标度因子来量化 DCT 系数。FWM 是与输入 DCT 系数的块维数相同的频率权重的表格。频率权重对不同的 DCT 系数应用不同的权重。权重被设计成强调具有对人类视觉系统更为敏感的频率内容的输入采样, 而不强调具有对视觉系统较不敏感的频率内容的采样。权重也可以根据像观察距离等因素来设计。

权重是根据经验数据选择的。ISO/IEC JTC1 CD 10918 中公开了为 8×8 DCT 系数设计权重掩模的方法, 该文题为“Digital compression and encoding of continuous-tone till images – part 1: Requirements and guidelines”, 由国际标准组织于 1994 年公布, 通过引用被结合于此。通常, 设计两个 FWM, 一个用于亮度分量, 一个用于色度分量。通过抽选而获得块大小为 2×2 、 4×4 的 FWM 表。通过对 8×8 块的 FWM 表的内插而获得 16×16 的 FWM 表。标度因子控制了被量化的系数的质量和比特率。

这样, 每个 DCT 系数都按照以下关系被量化:

$$DCT_q(i,j) = \left\lfloor \frac{8 * DCT(i,j)}{fwm(i,j) * q} \pm \frac{1}{2} \right\rfloor$$

其中, $DCT(i, j)$ 是输入 DCT 系数, $fwm(i, j)$ 是频率加权掩模, q 是标度因子, 而 $DCT_q(i, j)$ 是经量化的系数。注意到根据 DCT 系数的符号, 括号内的第一项上舍入或下舍入。DQT 系数也用适当的加权掩模来进行量化。然而, 可以使用多个表格或掩模, 并将其应用于 Y、Cb 和 Cr 分量的每一个。

然后, 用量化器 130, 或标度因子元件, 对像素数据块和频率加权掩模进行定标。在一优选实施例中, 有对应于平均比特率的 32 个标度因子。不像 MPEG2 这样的其它压缩方法, 平均比特率是根据所处理图像的质量来控制的, 而不是根据目标比特率或缓冲状态来控制。

经量化的系数被提供给扫描串行化器 152。串行化器 152 扫描经量化系数的块以产生经量化系数的经串行化的流。可以采用 Z 字形扫描、列扫描、或行扫描。也可以选择许多不同的 Z 字形扫描图案, 以及除了 Z 字形之外的图案。一优选技术为 Z 字形扫描采用了 8×8 块大小, 然而也可以使用其它大小。

参照图 4 和 5, 这里描述了不同的扫描技术。图 4b 说明了对整个 16×16 块 400 的 Z 字形扫描。在基于频率的块中, 譬如 DCT, 对值进行编码并表示, 使得 DC 值处在左上角, AC 值在接近右下角时减小。因此, 不考虑 16×16 块内的块大小分配, 整个 16×16 块上 Z 字形扫描的扫描技术导致编码的低效。换言之, 这样的 Z 字形扫描导致相同值的较短的游程长度。

图 4c 说明了较佳的扫描技术, 利用了给定块内系数排列的顺序。每个块 404、406、408、410、412、414、416、418、420、422、424、426 和 428 都采用了各自的 Z 字形扫描。一实施例中, 每个块都采用不同的扫描图案, 譬如垂直或水平的、或反向 Z 字形。尽管该实施例在保留最大游程长度方面非常好, 然而为每个块计算各自的 Z 字形扫描在计算上更加密集, 并且在硬件上难以实现。

因而, 已经确定, 像图 5a 和 5b 中所述的扫描实现那样的实现可能在使游程长度最大以及硬件实现容易性之间的最佳取得平衡。图 5a 说明了一个 16×16 块 500, 它通过块大小分配被细分成块 504、506、508、510、512、514、516、518、520、522、524、526 和 528。一实施例中, 不考虑 BSA 中止, 采用了对 16×16 块的每个 8×8 象限的 Z 字形扫描。这样, 块 504、506、508 和 510 被一 Z 字形扫描串行化, 块 512 被一 Z 字形扫描串行化、块 514 被一 Z 字形扫描串行化、而块 516、518、520、522、524、526 和 528 被一 Z 字形扫描串行化。

图 5b 说明了一个 16×16 块 550, 它通过块大小分配被细分成块 554、556、558、560、562、564、566、568、570、572、574、576 和 578。在该实施例中, 对 16×16 块的每个 8×8 象限采取不同的扫描类型。通过评估 8×8 块内的值并且确定哪种扫描方法是最有效的, 从而确定所采用的扫描类型。例如, 在图 5b 中, 为块 554、556、558、560 采取水平扫描, 块 562 被一 Z 字形扫描串行化, 块 564 被一垂直扫描串行化, 而块 566、568、570、572、574、576 和 578 被一 Z 字形扫描串行化。另一实施例中, 最佳扫描方法是逐帧确定的, 这与以逐块的基础相对。与逐块方法相比, 逐帧地确定最佳扫描方法计算上较不密集。

图 6a 说明了发生串行化的过程 600。读取一组数据 604。由于被读取的数据是基于可变块大小的, 因此被读入的数据不具有统一大小或长度。数据被编辑 608, 或者被构造为可表示为 16×16 块的形式。然后, 数据被分成 4 个 8×8 块大小 612。然后对每个 8×8 块执行 Z 字形扫描 616。然后, 数据被路由至缓冲器 620)。

图 6b 说明了串行化的另一实施例 650。读取一数据帧 654。评估数据帧已确定最佳的串行化技术 658。根据评估, 采取 Z 字形扫描 662、垂直扫描 664、或水

平扫描 668。在根据扫描方法之一进行串行化之后，数据被路由至缓冲器 672。

回过头参照图 1，经串行化的、经量化的系数的流被提供给可变长度编码器 156。可变长度编码器 156 可能使用零值的游程长度编码，其后是哈夫曼编码。该技术在上述美国专利号 5021891、5107345 和 5452104 中详细讨论，这些专利都通过引用被结合与总结于此。游程长度编码器取得经量化的系数并从非零值系数中分出零值。零值被称为游程长度值，并且经受哈夫曼编码。非零值被分开地进行哈夫曼编码。

经量化系数的修正哈夫曼编码也是可行的，它用于优选实施例中。这里，在 Z 字形扫描之后，游程长度编码器会确定每个 8×8 块内的游程长度/大小对。这些游程长度/大小对接着经过哈夫曼编码。

哈夫曼编码是从图像的测得的或理论的统计量中设计的。已经观察到，大多数自然图像的组成是由：平坦的相对缓慢变化的区域、以及像对象边界和高对比度纹理这样的繁忙区域。通过向繁忙区域分配更多比特而向平坦区域分配较少比特，具有像 DCT 这样的频域变换的哈夫曼编码器采用了这些特征。通常，哈夫曼编码器利用查找表来对游程长度和非零值进行编码。一般使用多个表格，本发明优选的是 3 个表格，然后可以根据需要采用 1 个或 2 个表格。

编码器 104 所产生的压缩的图像信号可以用缓冲器 160 临时存储，然后用传输信道 108 发送到解码器 112。包含块大小分配信息的 PQR 数据也被提供给解码器 112。解码器 112 包括一缓冲器 164 和一可变长度解码器 168，后者对游程长度值和非零值进行解码。

可变长度编码器 168 的输出被提供给反向串行化器 172，后者按照所采用的扫描方案对系数排序。例如，如果使用了 Z 字形扫描、垂直扫描和水平扫描的混合，则反向串行化器 172 会根据所采用的扫描类型适当地重排系数。反向串行化器 172 接收 PQR 数据来帮助把系数适当地排序成复合系数块。

用选择器 174 把复合块提供给反向量化器 176，用于撤销由量化器标度因子和频率加权掩模的使用所产生的处理。

然后，如果应用了差分四叉树变换，则把系数块提供给 IDQT 元件 186，其后是 IDCT 元件 186。否则，系数块被直接提供给 IDCT 元件 190。IDQT 元件 186 和 IDCT 元件 190 对系数进行反变换以产生一个像素数据块。然后，像素数据可能必须被内插、转换成 RGB 形式、然后为了将来的显示而被存储。

因而，给出了用于图像压缩的系统和方法，它根据像素方差执行块大小分配。

基于方差的块大小分配提供了几个好处。由于在确定块大小之后执行离散余弦变换，因此实现了高效的计算。仅需对所选的块执行计算上密集的变换。此外，块选择过程是有效的，因为计算像素的方差值数学上很简单。基于方差的块大小分配的另一好处是它是基于感觉的。像素方差是块内活动性的度量，并且指示了边缘、纹理等等的存在。它能比像素平均值这样的度量更好地捕获块的细节。这样，本发明的基于方差的方案向具有较多边缘的区域分配较小的块，向较平坦的区域分配较大的块。因此，在重建的图像中可以实现出众的质量。

作为例子，各种说明性的逻辑框、流程图、以及结合这里所公开的实施例描述的步骤可以用以下来实现：具有应用专用集成电路(ASIC)的硬件或软件、可编程逻辑器件、微处理器、离散门或晶体管逻辑、离散硬件组件、譬如寄存器和 FIFO、执行一组固件指令的处理器、以及常规的可编程软件和处理器，或者它们的组合。处理器最好是微处理器，但或者，处理器可以是任一常规处理器、控制器、微控制器或状态机。软件可以驻留在 RAM 存储器、闪存、ROM 存储器、寄存器、硬盘、可移动磁盘、CD-ROM、DVD-ROM 或本领域已知的任何其它形式的存储媒介内。

提供了优选实施例的前面描述，使得本领域的任何技术人员都能制造或使用本发明。对于本领域的技术人员而言，对于本发明的各种修改是显而易见的，这里所发现的基本原理也可以应用于其它实施例，而不使用创造性的能力。因此，本发明不是为了限制这里所示的实施例，但是符合与这里所公开的原理和新颖性特征一致的最宽泛的范围。

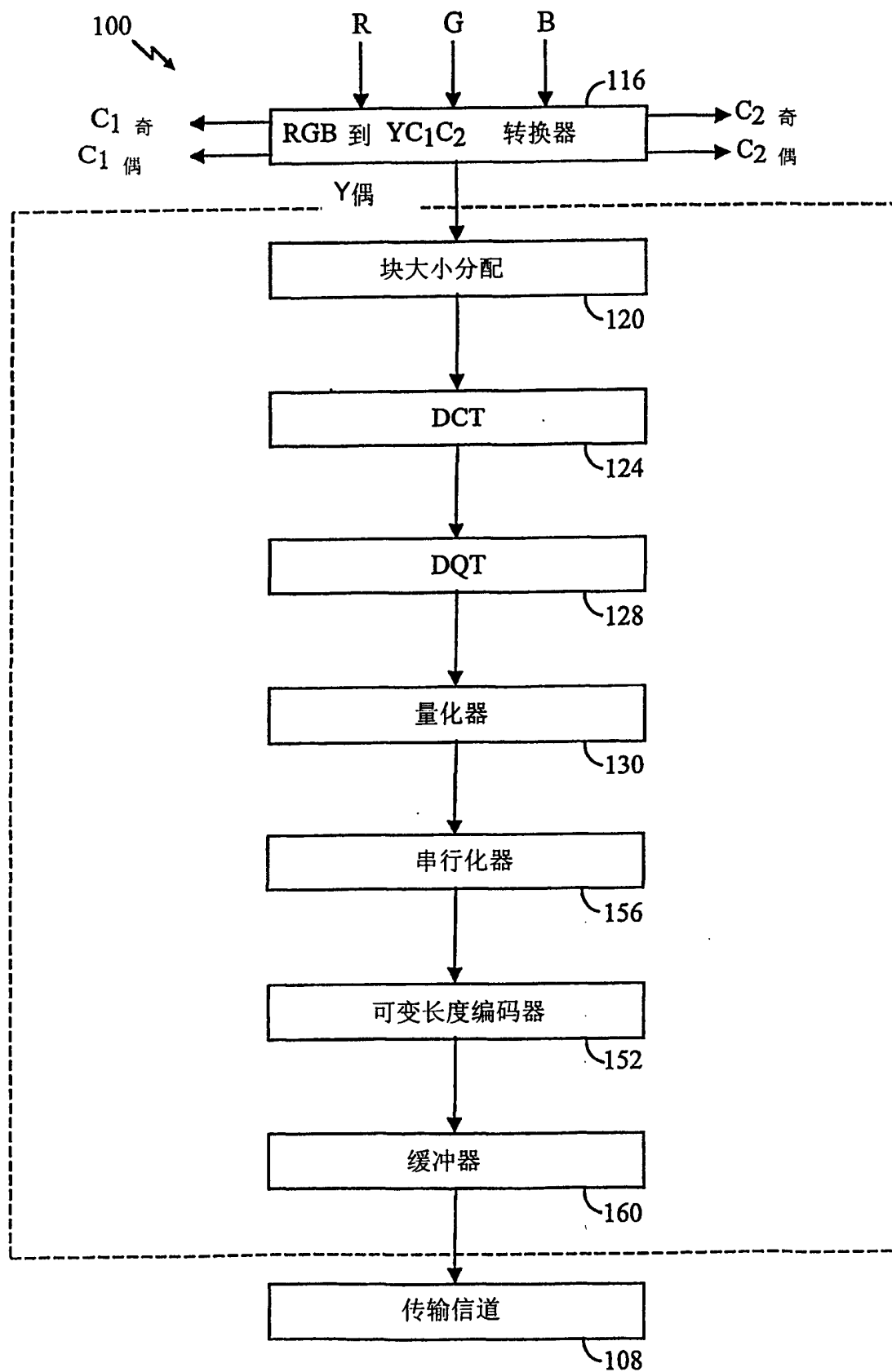


图 1

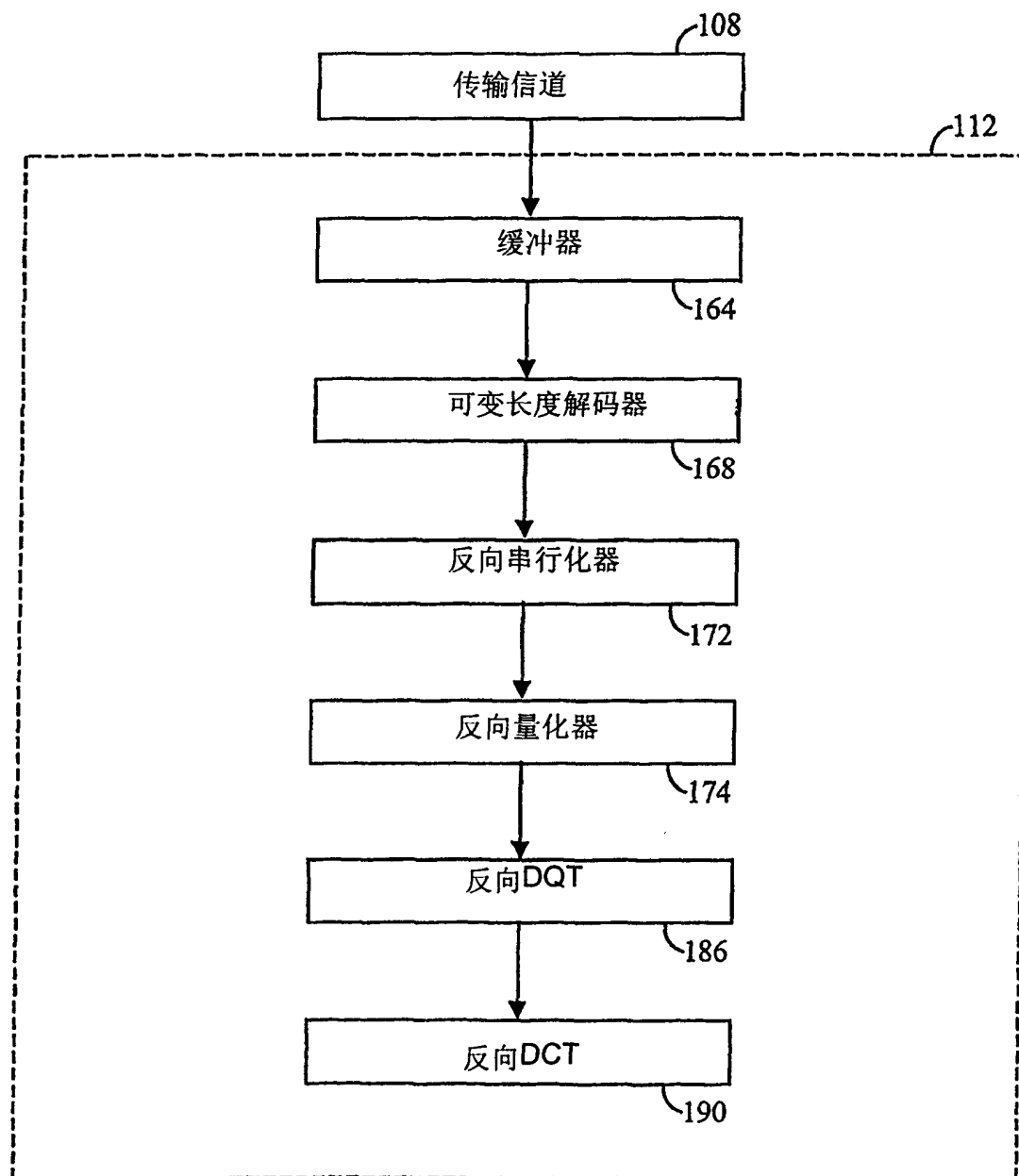


图 2

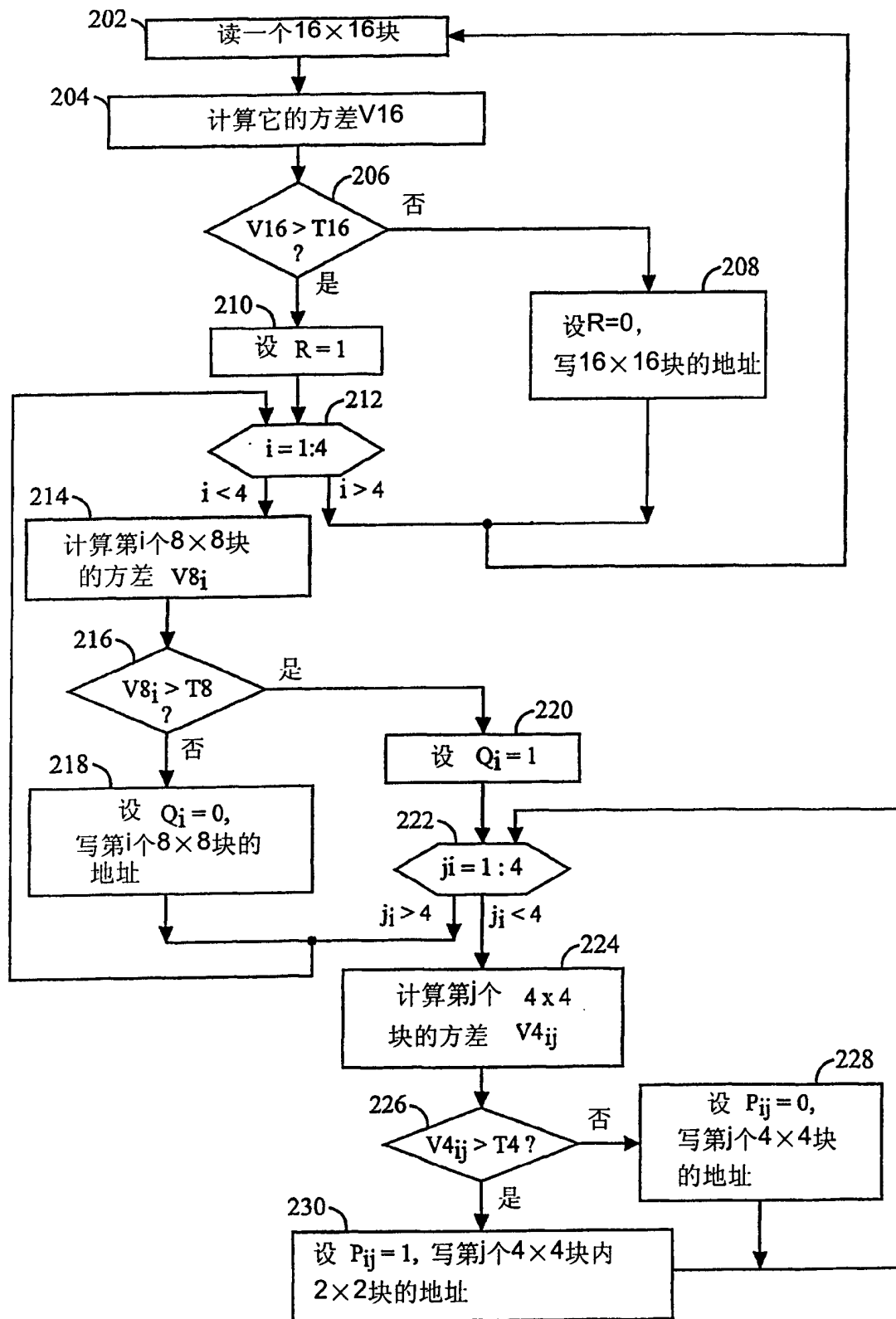


图 3

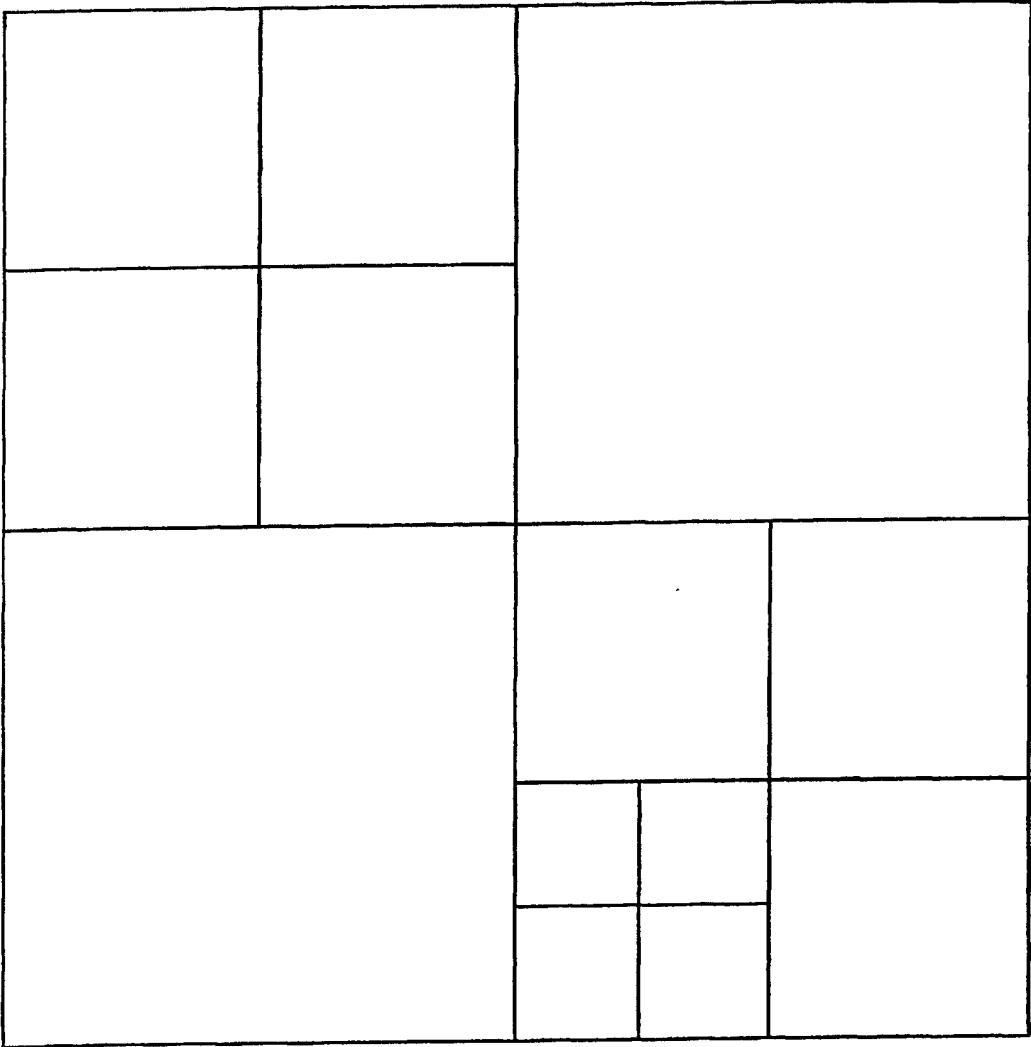
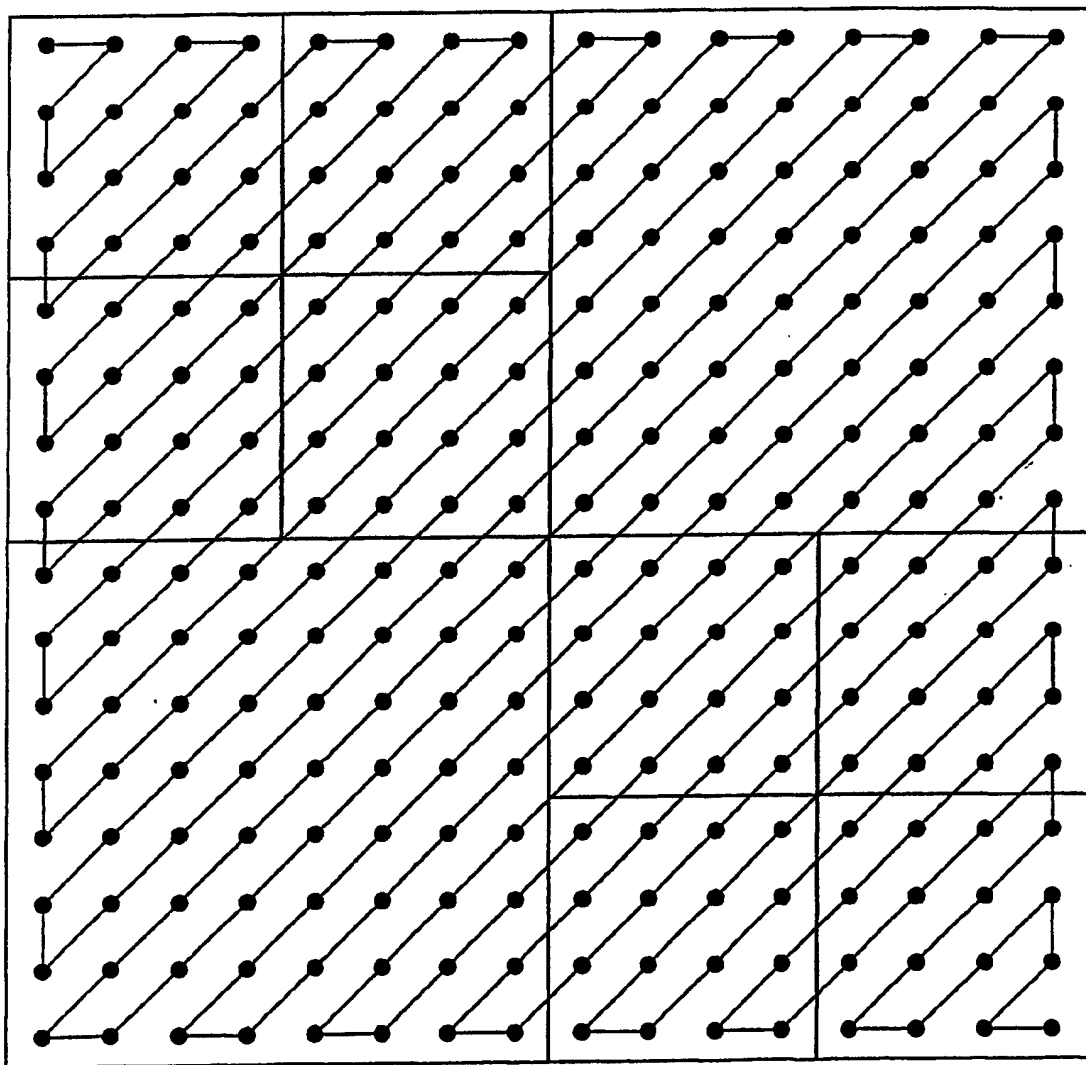
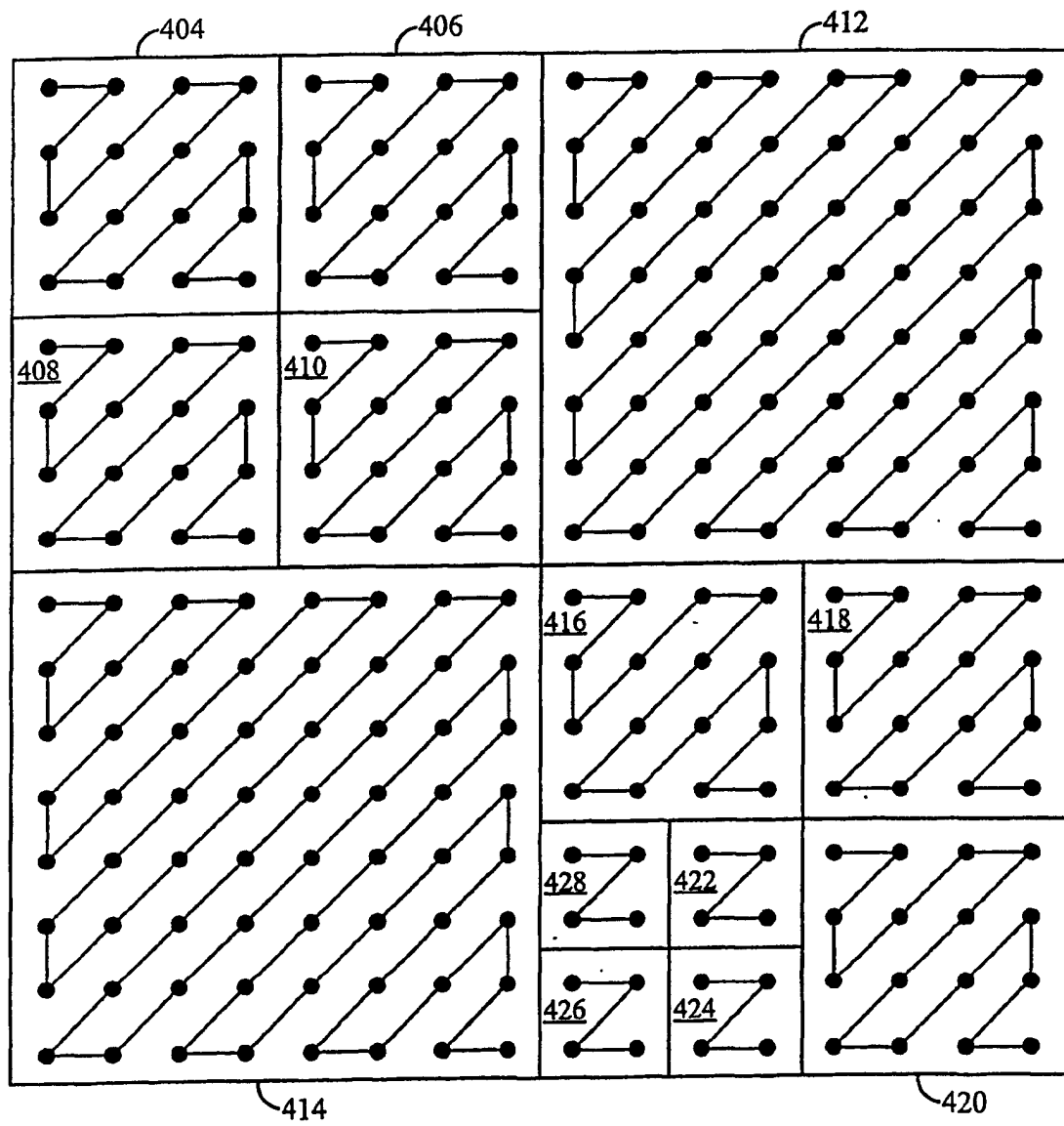


图 4a



402 ↗

图 4b



403 ↗

图 4c

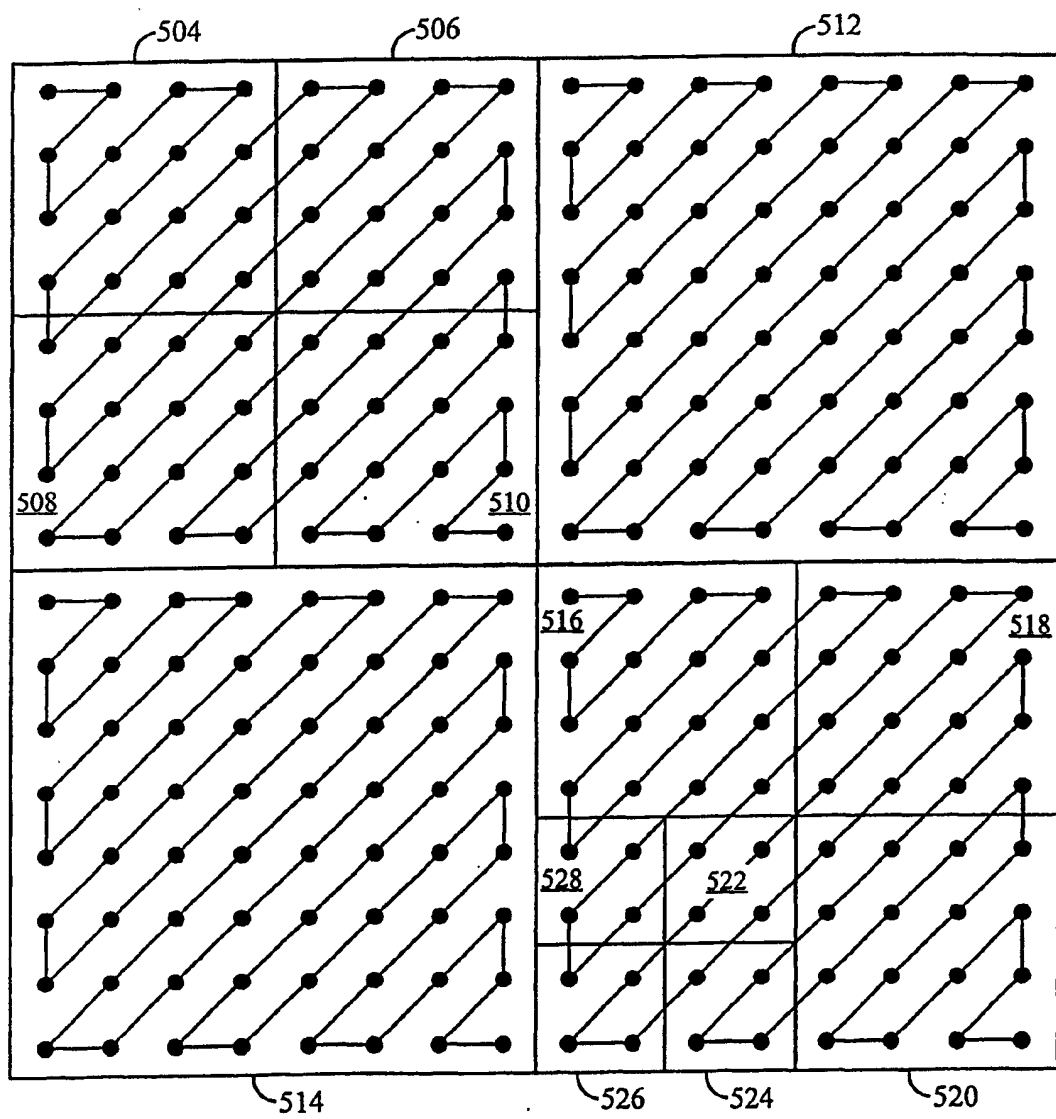


图 5a

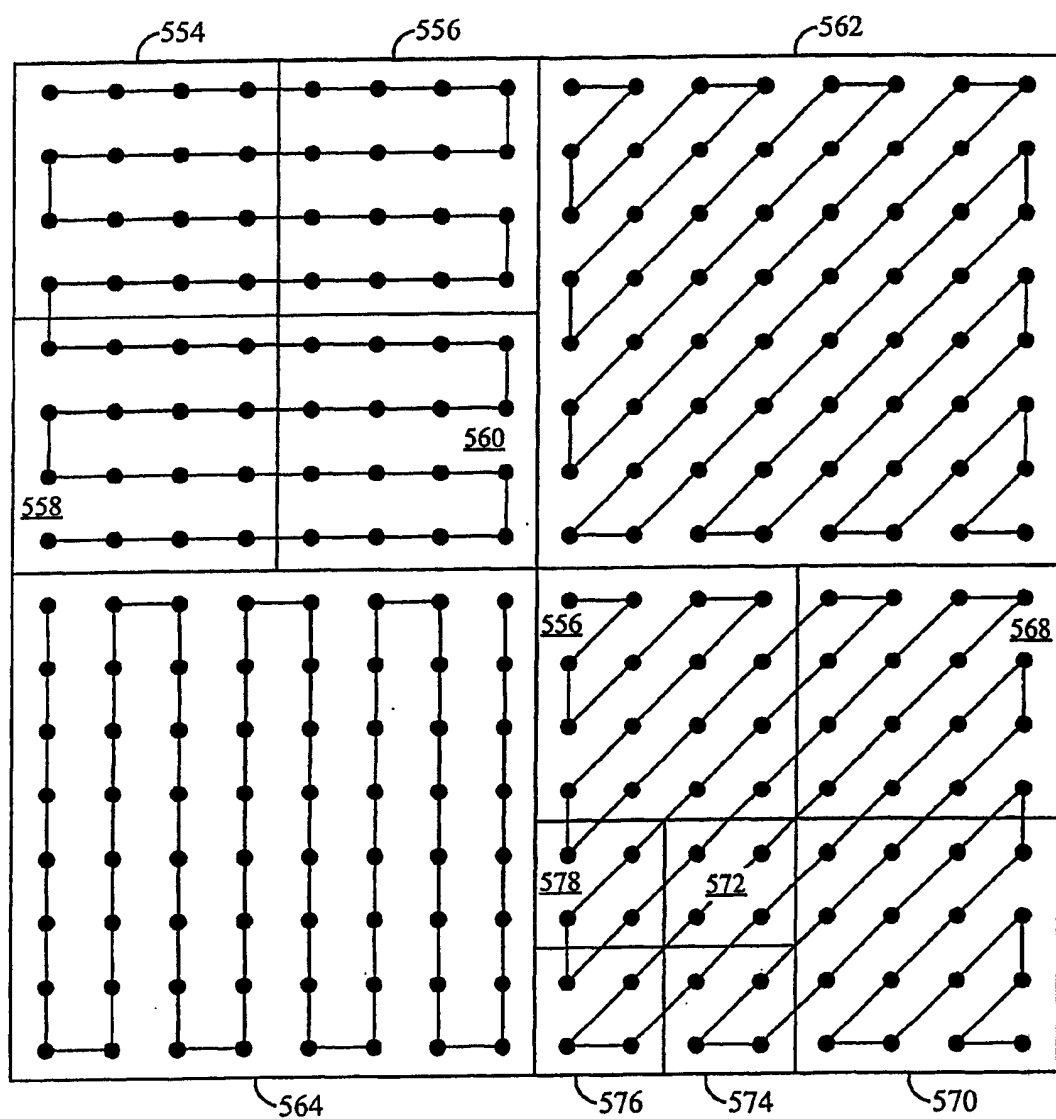
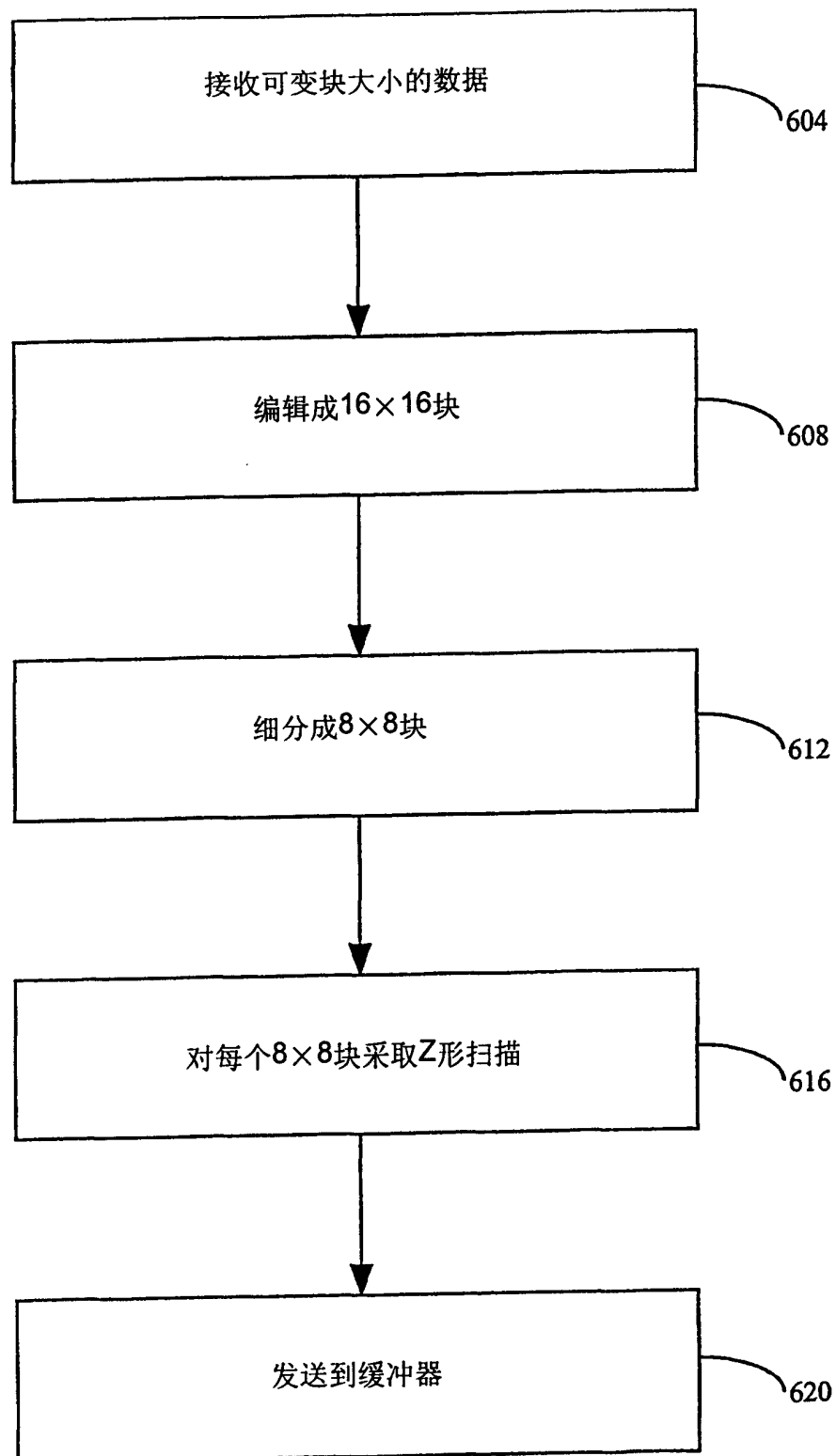
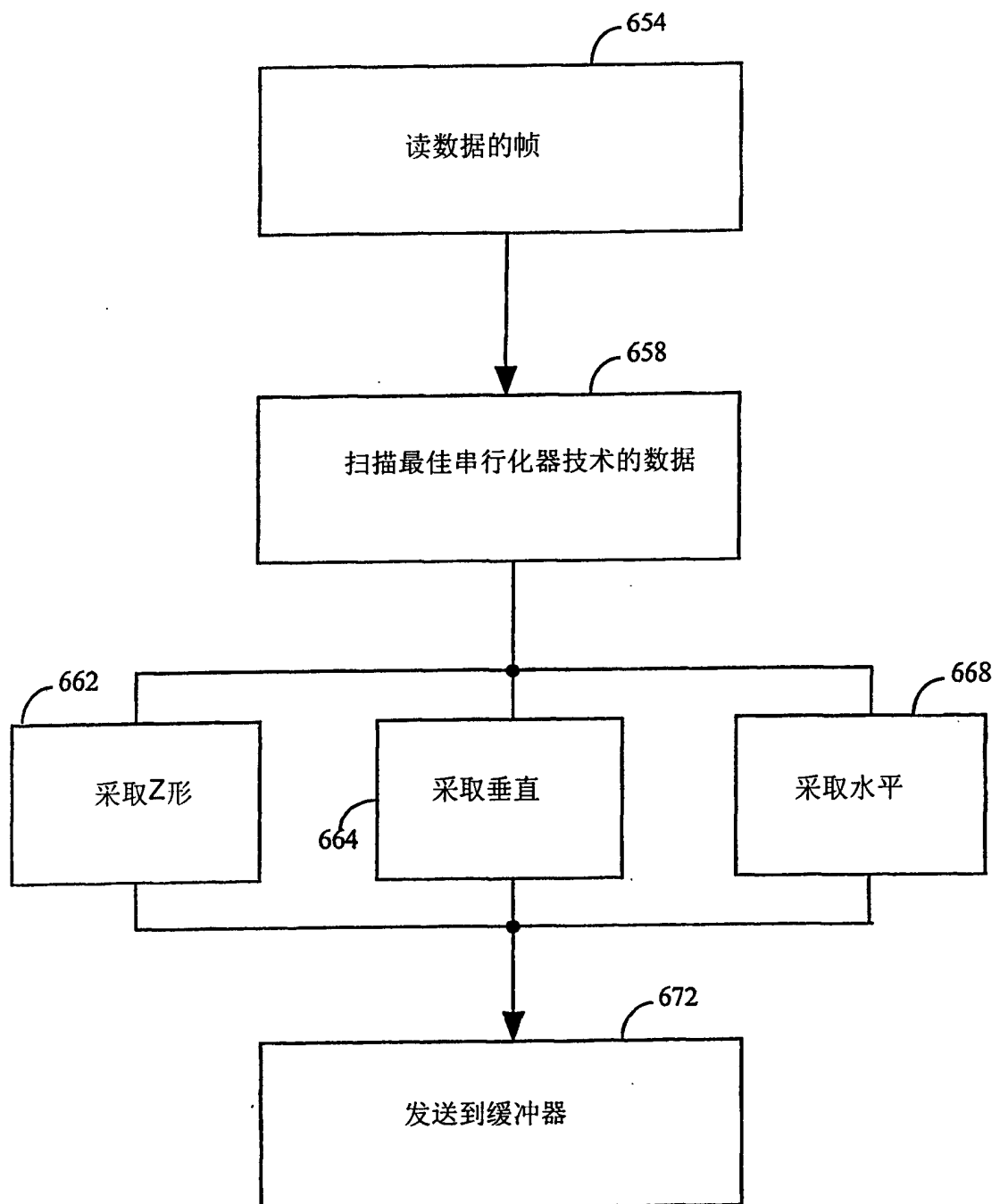


图 5b



600

图 6a



600 ↗

图 6b