



[12] 发明专利说明书

专利号 ZL 02143205.8

[45] 授权公告日 2005 年 12 月 7 日

[11] 授权公告号 CN 1231068C

[22] 申请日 2002.9.18 [21] 申请号 02143205.8

[30] 优先权

[32] 2001.9.18 [33] US [31] 09/955,577

[71] 专利权人 微软公司

地址 美国华盛顿州

[72] 发明人 H·S·马尔渥

审查员 张 军

[74] 专利代理机构 上海专利商标事务所有限公司

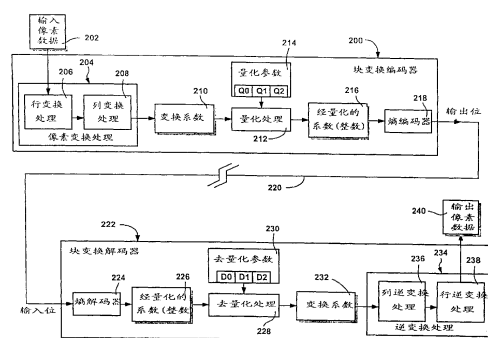
代理人 钱慰民

权利要求书 4 页 说明书 13 页 附图 3 页

[54] 发明名称 图像和视频编码用的改进的块变换和量化

[57] 摘要

图像或视频编码和解码用的一种改进方法和块变换，其中：变换和逆变换矩阵是这样确定的，即编码和解码时显著地减少计算复杂性。如，解码期间在将去量化变换系数转为输出像素信息的 2 维逆变换中，每个系数变换只需 4 次加法和 1 次移位操作，都在 16 位运算中进行。因为通过使用根据各系数位置从 3 张表中选用的 1 张表，编码期间的量化和解码期间的去量化(16 位)具有对其他变换乘法因子已作补偿的参数值(除了在变换和逆变换处理中可由移位操作完成的这些 2 的幂的乘法操作之外(如，2 或 1/2))，所以变换可提供正确的结果。与其他已知的变换相比，其计算复杂性大大减少，并对压缩和质量没有不利影响。



1. 一种对视频或图像像素数据进行解码的方法，其特征在于，包括：
接收一组经量化的变换系数；
把经量化变换系数经去量化转换成去量化的变换系数；和
把逆变换处理应用于去量化的变换系数，从中重构信息，所述逆变换处理应用于具有值的逆变换计算，使得在所述值中的一个值和所述逆变换计算中去量化的变换系数中一个系数之间的每个乘法操作可由一个移位操作来完成。
2. 如权利要求 1 所述的方法，其特征在于，所述对经量化系数去量化包括将每个经量化变换系数与从至少一组参数中所选择的一个参数相乘。
3. 如权利要求 2 所述的方法，其特征在于，至少有 2 组参数可访问，其中，经量化的变换系数是逻辑地排列于块中，且对于每个经量化变换系数，对所述组中的一组的选择是根据该经量化变换系数在块中的位置。
4. 如权利要求 3 所述的方法，其特征在于，有 3 组参数可访问。
5. 如权利要求 2 所述的方法，其特征在于，经量化变换系数和参数值能以 16 位运算进行所述去量化。
6. 如权利要求 1 所述的方法，其特征在于，把逆变换处理应用于去量化变换系数是在 16 位运算中完成。
7. 如权利要求 1 所述的方法，其特征在于，所述逆变换处理对应于 2 维变换，该逆变换计算值是基于下述矩阵：

$$\begin{bmatrix} 1 & 1 & 1 & 1/2 \\ 1 & 1/2 & -1 & -1 \\ 1 & -1/2 & -1 & 1 \\ 1 & -1 & 1 & -1/2 \end{bmatrix}$$

8. 如权利要求 1 所述的方法，其特征在于，进一步包括，在接收经量化变换系数组之前：

把变换处理应用于输入信息，从中构成变换系数，所述变换处理应用于具有值的变换计算，使得在所述变换计算值中的一个值和所述变换计算中变换系数中一个系数之间的每个乘法操作可由一个移位操作来完成；和

把变换系数量化为经量化变换系数。

9. 如权利要求 8 所述的方法，其特征在于，所述变换处理对应于 2 维变换，其中所述变换计算值是基于下述矩阵：

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

10. 如权利要求 1 所述的方法，其特征在于，由存储在计算机可读媒体上的计算机可执行指令来执行所述方法。

11. 一种提供视频或图像数据的系统，其特征在于，包括：

(1) 变换处理，把变换应用于信息，以从中构成变换系数，变换处理应用具有值的变换计算，使得在所述变换计算值中的一个值和所述变换计算中变换系数中一个系数之间的每个乘法操作可由一个移位操作来完成；和

(2) 量化处理，把变换系数量化为一组经量化整数变换系数；

一种机构，把来自块变换编码器的量化整数变换系数提供给块变换解码器；和

块变换解码器，包括：

①去量化处理，把经量化的整数变换系数经去量化转换成去量化的变换系数；以及

②逆变换处理，从所述经去量化的变换系数中产生经重构的信息，逆变换处理应用具有值的逆变换计算，使得在所述逆变换计算值中的一个值和所述逆变换计算中去量化的变换系数中一个系数之间的每个乘法操作可由一个移位操作来完成。

12. 如权利要求 11 所述的系统，其特征在于，所述变换处理对应于 2 维的行变换和列变换处理。

13. 如权利要求 12 所述的系统，其特征在于，所述行变换和列变换处理以及变换计算值是依据下述矩阵：

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

14. 如权利要求 11 所述的系统，其特征在于，所述量化处理是通过从一参数组中选择一参数来换算各变换系数。

15. 如权利要求 11 所述的系统，其特征在于，所述量化处理通过对每个变换系数，从 3 组参数中所选的 1 组参数中选择一个参数，来换算变换系数，所选择的组是根据所述变换系数相对于其他变换系数的相对位置而选择。

16. 如权利要求 11 所述的系统，其特征在于，提供量化整数变换系数的机构包括块变换编码器中的熵编码器，把块变换编码器以逻辑方式接至块变换解码器的至少 1 个通信媒体或器件，以及块变换解码器中的熵解码器。

17. 如权利要求 11 所述的系统，其特征在于，所述去量化处理通过从一参数组中选择一参数来换算各量化整数变换系数。

18. 如权利要求 11 所述的系统，其特征在于，所述去量化处理通过对每个量化整数变换系数从 3 组参数中所选的 1 组参数中选择一个参数，来换算量化整数变换系数，所选择的组是根据所述量化整数变换系数相对于其他量化整数变换系数的相对位置而选择的。

19. 如权利要求 11 所述的系统，其特征在于，所述逆变换处理对应于 2 维的列逆变换和行逆变换处理。

20. 如权利要求 19 所述的系统，其特征在于，所述列逆变换和行逆变换处理以及逆变换计算值是根据下述矩阵：

$$\begin{bmatrix} 1 & 1 & 1 & 1/2 \\ 1 & 1/2 & -1 & -1 \\ 1 & -1/2 & -1 & 1 \\ 1 & -1 & 1 & -1/2 \end{bmatrix}$$

21. 一种用于视频或图像像素数据进行解码的方法，其特征在于，包括：
接收一组量化变换系数：

把量化变换系数经去量化转换为去量化的变换系数；和

把逆变换处理应用于去量化的变换系数，所述逆变换处理根据矩阵示出值对所述去量化变换系数应用逆变换计算：

$$\begin{bmatrix} 1 & 1 & 1 & 1/2 \\ 1 & 1/2 & -1 & -1 \\ 1 & -1/2 & -1 & 1 \\ 1 & -1 & 1 & -1/2 \end{bmatrix}$$

其中，所述逆变换计算包括多个乘法操作，并且，所述多个乘法操作中的每个操作都可以通过一次移位操作执行。

22. 如权利要求 21 所述的方法，其特征在于，所述去量化包括：选择各量化变换系数，和对各选择的量化变换系数：

根据所述量化变换系数相对于其他量化变换系数的逻辑相对位置选择 1 组参数；

在选中的组中查找一参数；以及

通过将量化变换系数乘以此参数，将量化变换系数换算为去量化变换系数。

图像和视频编码用的改进的块变换和量化

发明领域

本发明一般涉及图像和视频信号的编码和解码，特别涉及到一种用于编码和解码图像视频信号的改进的块变换和逆变换，以及量化和去量化方法。

发明背景

离散余弦变换（DCT）一般用于块变换的图像和视频（图像序列）编码，例如 JPEG 和 MPEG。其他原因是，就图像视觉质量而言，DCT 可将随机信号（例如，对应于图像数据的信号）分离为较重要的低频部分和次要的高频部分，这样便于压缩。

举例来说，H. 26L 是一种使用 4×4 像素块构成帧的图像和/或视频的编码技术，与采用 8×8 像素块的 MPEG 和 JPEG 不同，H. 26L 通过使用除帧间估计和运动补偿（即，帧际编码）外，还利用同帧中存在的其他块的预测信息（即，帧内编码），从而使采用 4×4 像素块取得了相对较高质量的压缩。一般说来，为了完成帧内预测，H. 26L 编码器要告诉相应的 H. 26L 解码器，在构成帧中哪些以前的像素块被用作预测基础，并识别 6 种可能的预测器（公式）中的 1 个来确定新块像素（从以前帧内的块的像素进行识别）。这会引致预测误差，但此误差也供给解码器，以校正新块。预测误差信息利用块变换（DCT）加以编码，将编码结果送至解码器，用于再次计算误差信息，包括通过逆变换来校正预测的块。

图像和视频的编码或解码需要大量的处理能力。众所周知，编码器具有的处理能力要大大地超过多个将数据转换为图像的解码器的处理能力，因为解码器通常是在用户设备中提供，如，图像和视频的解码可发生在电视机顶端盒内，个人数字助理内，袖珍式个人计算机和更先进的蜂窝式电话内。

因此，考虑图像和视频的编码和解码方法时，保持解码的简单性是很重要的，纵然这意味着编码必须在计算上做得更复杂些。任何简化解码的措施都是需要的，在这样做时，编码和解码的处理不会对现有技术基本方法的压缩水平、最终图像质量和解码器需要的其他资源量起到不利影响。

发明内容

简言之，本发明提供一种改进的方法，系统和块变换，使得编码器和解码器两者都可大大地简化对图像和视频的计算复杂性。而同时，压缩、图像/视频质量和其他资源所受到的影响可以忽略。

特别是，提供一种对离散余弦变换（DCT）采用新的整数近似的正交块变换和相应的逆变换，以及其他可显著地减少编码器和解码器计算复杂性的计算改变。事实上，在一个实施例中，在编码器和解码器，每个系数的变换操作次数已减少至在16位运算中的4次加法和1次移位（而H.26L规定的变换要在32位运算中有4次加法和3次乘法）。因为编码期间的量化（换算和舍入至整数）和解码期间的去量化（通过使用根据各系数位置所选择的3张表之一）具有对其他变换乘法系数进行补偿的参数值（除了在变换和逆变换过程中由移位操作所完成的乘法外），所以本发明能正确地变换像素信息（如，误差校正数据）。

而且，解码器在解码期间对各经量化的系数进行1次16位乘法，将经量化的系数去量化，成为变换系数，替代（如H.26L所需的）32位的乘法运算。在许多设备中，就一般图像和视频来说，本发明的数据处理的好处与H.26L相比是显而易见的，特别在解码器上，如果产生质量和/或压缩的任何损失，也是可忽略的。

因此，本发明提供了一种对视频或图像像素数据进行解码的方法，包括：接收一组经量化的变换系数；把经量化变换系数经去量化转换成去量化的变换系数；和把逆变换处理应用于去量化的变换系数，从中重构信息，所述逆变换处理应用于具有值的逆变换计算，使得在所述值中的一个值和所述逆变换计算中去量化的变换系数中一个系数之间的每个乘法操作可由一个移位操作来完成。

另一方面，本发明提供了提供视频或图像数据的系统，包括：

（1）变换处理，把变换应用于信息，以从中构成变换系数，变换处理应用具有值的变换计算，使得在所述变换计算值中的一个值和所述变换计算中变换系数中一个系数之间的每个乘法操作可由一个移位操作来完成；和

（2）量化处理，把变换系数量化为一组经量化整数变换系数；

一种机构，把来自块变换编码器的量化整数变换系数提供给块变换解码器；和块变换解码器，包括：①去量化处理，把经量化的整数变换系数经去量化转换成去量化的变换系数；以及②逆变换处理，从所述经去量化的变换系数中产生经重构的信息，逆变换处理应用具有值的逆变换计算，使得在所述逆变换计算值中的一个值和所述逆变换计算中去量化的变换系数中一个系数之

间的每个乘法操作可由一个移位操作来完成。

再一方面，本发明提供了一种用于视频或图像像素数据进行解码的方法，包括：接收一组量化变换系数；把量化变换系数经去量化转换为去量化的变换系数；和把逆变换处理应用于去量化的变换系数，所述逆变换处理根据矩阵示出值对所述去量化变换系数应用逆变换计算：

$$\begin{bmatrix} 1 & 1 & 1 & 1/2 \\ 1 & 1/2 & -1 & -1 \\ 1 & -1/2 & -1 & 1 \\ 1 & -1 & 1 & -1/2 \end{bmatrix}$$

其中，所述逆变换计算包括多个乘法操作，并且，所述多个乘法操作中的每个操作都可以通过一次移位操作执行。

从下面结合附图的详细描述中，其他好处和优点就是明显的，其中：

附图说明

图 1 是示出包含本发明的示范计算机系统的方块图；

图 2 是示出根据本发明的一个方面，用于将视频或图像数据像素编码成经编码的视频，和将经编码的视频解码为像素信息的部件的方块图，包括变换，量化，去量化和逆变换。

图 3 示出根据本发明的一个方面所表示的改进变换矩阵的蝶形结构。

图 4 示出根据本发明的一个方面所表示的改进逆变换矩阵的蝶形结构。

具体实施方式

图 1 示出可实施本发明的一个适合操作环境 120 的例子，特别是对图像和/或视频数据的解码。操作环境 120 只是适合操作环境的一个例子，并不是要对本发明应用范围或功能性提出任何限止。适用于本发明的其他一些众所周知的计算系统、环境和/或配置包括（但不局限于）：个人计算机、服务器计算机、手持式和膝上型设备、多处理器系统、基于微处理的系统、可编码的用户电子设备、网络 PC 机、小型计算机、主计算机，包括上述任何系统或设备的分布计算环境等等。例如，图像和/或视频图像数据的编码很可能经常在 1 台具有比现今手持式个人计算机更强处理能力的计算机上进行，但这并不是说不能在示范设备上上进行编码，或在更强处理能力的设备上上进行解码。

本发明在一般计算机可执行的指令的范围中描述，如程序模块，可由 1 个或多个计算机或其他设备来执行。一般说来，程序模块包括执行特定任务或实施特定的抽象数据类的例行程序、程序、对象、组成部分、数据结构等等。一般程序模块功能性可按各种实例的要求进行组合或分布。

计算设备 120 一般至少包括一些形式的计算机可读媒体。计算机可读媒体可以是由计算设备 120 能存取的任何可获得的媒体。举例来说，但不限于，计算机可读媒体可包括计算机存储媒体和通信媒体。计算机存储媒体包括以任何方法或技术来存储信息所实现的易失性和非易失性、可移动和不可移动的媒体，如计算机可读指令、数据结构、程序模块或其他数据。计算机存储媒体包括，但不限于，RAM、ROM、EEPROM、快闪存储器或其他存储器技术、CD-ROM、数字式多用途盘（DVD）或其他光存储器、磁盘、磁带、磁盘存储器或其他磁存储器件，或可用于存储所需信息，并由计算设备 120 可存取的任何其他媒体。通信媒体一般包括计算机可读指令、数据结构、程序模块或经调制的数据信号中的其他数据如载波或其他传送机制，并包括任何信息发送媒体。术语“经调制的数据信号”就是指具有其特性组中 1 个特性或多个特性或可按关于对信号中信息编码的一种方式改变的信号。举例来说，并不局限于，通信媒体包括有线媒体如：有线网络或直接有线连接，和无线媒体如：声频、RF、红外或其他无线媒体。上述的任何组合也必须包括在计算机可读媒体的范围内。

图 1 示出一种这样的手持式计算设备 120 的功能部件，包括处理器 122、存储器 124、显示器 126 和键盘 128（它可以是实际的或虚拟的键盘）。存储器 124 一般包括易失性存储器（如 RAM）和非易失性存储器（如 ROM、PCMCIA 卡等等）两者。操作系统 130 驻留在存储器 124 内，并在处理器 122 上执行，如微软公司 Window® CE 操作系统或其他操作系统。

1 个或多个应用程序 132 加载入存储器 124，并在操作系统上 130 运行。应用程序的例子包括电子邮件程序、调度程序、PIM（个人信息管理）程序、字处理程序、电子表格程序、因特网浏览器程序等等。手持式个人计算机 120 还包括加载入存储器 124 的通知管理程序 134，此程序在处理器 122 上执行。通知管理程序 134 处理诸如来自应用程序的通知要求。

手持式个人计算机 120 具有由 1 个或多个电池实现的电源 136。电源 136 进一步包括外接电源，对机内电池过载（override）或充电，如 AC 适配器或供电的连接底座。

图 1 示出的示范手持式个人计算机 120 具有 3 种类型的外部通知机构: 1 个或多个发光二极管 (LED) 140 和 1 个音频发生器 144。这些器件可直接耦合至电源 136, 这样, 当这些器件激活时, 即使手持式个人计算机处理器 122 和其他组成部分为了节省电池电源而关机, 这些器件仍在通知机构所指定的时期内保持开启。LED 140 最好无限止地保持开启, 直至用户采取行动。注意现在这类音频发生器 144 耗用当今手持式个人计算机的电池电源太多, 所以当系统休眠时或在激活后有限时期内, 将它置于关闭。

编码和解码

就背景所述, 离散余弦变换 (DCT) 利用线性变换公式 $X=Hx$ 将长度 N 矢量 x 映射为新的变换系数矢量 X , 其中 H 的第 k 行和第 n 列中的元素由下式规定:

$$H_{kn} = H(k, n) = c_k \sqrt{\frac{2}{N}} \cos\left((n+1/2)\frac{k\pi}{N}\right)$$

其中 $k=0, 1, \dots, N-1$, 而 $n=0, 1, \dots, N-1$, 同时当 $k>1$ 时, $c_0=\sqrt{2}$ 而 $c_k=1$ 。DCT 矩阵是正交的, 因此, 其逆向等于其转置, 即 $x=H^T X=H^T X$ 。

DCT 的一个缺点是 $H(k, n)$ 项是无理数, 所以整数输入数据 $x(n)$ 将映射为无理变换系数 $X(k)$ 。因此, 使用数字计算机时, 当直接变换和逆变换以级联方式计算时, 输出数据就不完全相等于输入数据。换言之, 如果以 $X=Hx$ 来计算, 且 $u=\text{round}(H^T X)$, 然而对所有 n 来说, $u(n)=x(n)$ 并不为真。然而通过引入合适的换算因子 α, γ , 例如 $x=\gamma Hx$ 而 $u=\text{round}(\alpha H^T X)$, 则当 α 和 γ 选择得足够大时, 对几乎所有 n , $u(n)=Gx(n)$, 其中 G 是整数, 但这也不能保证是正确的结果。

在运动补偿的视频编码器中, 例如, 过去解码的帧数据用作为用于生成当前帧的预测信息的参考信息。因此, 作为编码的一部分, 编码器生成其自己的经解码的帧, 因而编码器需要计算逆变换。如果应用公式 $u=\text{round}(\alpha H^T x)$, 则不同处理器中的不同浮点格式和舍入方法会导致不同的结果。结果是, 在编码器上的解码数据对解码器上的解码数据之间将产生偏差 (它们具有不同处理器), 因为编码器是基于对帧块进行预测/运动信息, 而此帧块越来越少, 如解码器所产生的帧块那样, 因此每 1 新帧的图像越来越差。

解决数据偏差的一个办法是用只具有整数的矩阵来近似矩阵 H , 就可消除舍入误差。如果 H 行是正交的并具有相同的范数 (平方和), 然后接下来其遵循对所有整数 x , 可以以整数运算正确计算 u 。换言之, 当直接变换由 $X=Hx$ 来计算, 逆变换

由 $u=H^T X$ 来计算时, 得出结果为 $u=Gx$, 其中 G 是整数, 等于 H 中任何行的平方范数。

对 DCT 生成整数近似值的一种方法是使用通用公式:

$$Q(k, n) = \text{round}(\alpha H(k, n))$$

其中 α 为换算参数。

在 H. 26L 视频压缩标准中, 图像是由 4×4 像素块组成, 故在 H. 26L 的 DCT 公式中 $N=4$ 。DCT 矩阵的结果等于:

$$H = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ c & s & -s & -c \\ 1 & -1 & -1 & 1 \\ s & -c & c & -s \end{bmatrix}$$

其中 $c = \sqrt{2} \cos(\pi/8)$ 和 $s = \sqrt{2} \sin(\pi/8)$ 。

目前版本 H. 26L 中变换矩阵是设定 $\alpha = 26$ 取得的, 此值计算为:

$$Q_0 = \text{round}(26H) = \begin{bmatrix} 13 & 13 & 13 & 13 \\ 17 & 7 & -7 & -17 \\ 13 & -13 & -13 & 13 \\ 7 & -17 & 17 & -7 \end{bmatrix}$$

使用换算参数值 26, Q_0 行和列相互是正交的 (即, 任何两列的内积为零), 所有行和列具有 1 个范数 676。事实上, 对于 $\alpha < 50$ 时, 只有 $\alpha = 2$ 或 $\alpha = 26$ 才能提供具有相等范数的行的正交矩阵。然而 $\alpha = 2$ 的解决方法所得到的压缩效果不好, 并且更大的 α 值并不令人满意, 因为计算直接变换的结果 $X = Q_0 x$ 时, 所需的计算复杂性会增加 (如, 字长)。因而在 H. 26L 中迄今还是选用 $a = 26$ 。

逆变换由 $x' = Q_0 X$ 来规定, 故可用整数运算来计算。按上述规定, $X' = 676x$, 即, 重构的数据 x' 等于原始数据 x 由整数增益 676 可放大 (它是 Q_0 中任何行的范数)。

按照本发明一个方面, 提供一种采用整数近似值的块变换矩阵, 这些近似值是正交的, 但具有不同的值, 能在编码和解码时显著地减少计算复杂性。更具体地说, 使用 $\alpha = 2.5$, 生成下列矩阵:

$$Q_0^{\Delta} = \text{round}(2.5H) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

注意: Q_0 行是相互正交的, 当它们范数不同时 (行 0 和行 2 的范数等于 4; 而

行 1 和行 3 的范数等于 10)，如下所述，以一种不明显增加计算复杂性的方法进行量化和去量化处理。

回到附图 2，显示出块变换编码器 200 和块变换解码器 222，能通过一些类型媒体和/或中间器件 220 加以连接。众所周知，编码器 200 不需要直接将其输出用于解码器 222，这样，图 2 中标出的线 220 可虚拟地代表任何媒体或器件，如网络电缆、无线传输媒体、电话线、存储器件、路由器和/或它们虚拟的任何组合。

一般说来，编码器 200 通过可分的 2 维变换 204 将每个 $N \times N$ 输入像素数据 202 块进行变换来工作。具体地说，首先由行变换处理 206 来变换块中行像素数据，然后行变换的结果（图 2 中未特别示出）通过列变换处理 208 变换成结果变换系数 210。换言之，2 维 DCT 对像素块各行基本上执行 1 维 DCT，接着是，对由 1 维 DCT 在行上产生的各列像素块进行 1 维 DCT。行和列的变换也可以以相反的次序进行，而获得的结果相同。

在 H. 26L 中，规定的变换矩阵形成下列公式，以计算变换系数：

$$\begin{aligned} A &= 13a + 13b + 13c + 13d \\ B &= 17a + 7b - 7c - 17d \\ C &= 13a - 13b - 13c + 13d \\ D &= 7a - 17b + 17c - 7d \end{aligned}$$

其中，[abcd]在 4X4 块中首先代表行像素值，在行变换后，代表这些行变换后数据值的列像素值。但在实际中，这些公式可以简化，DCT 矩阵具有递归结构，可减少所需的乘法和加法的运算次数。然而，上述矩阵需要至少 4 次加法和 3 次乘法来计算各像素的变换系数。而且，因为每个像素可能是一个带符号的 9 位数值，当行和列因子（增益等于范数 676）相乘时，计算要求 32 位运算，而单独一次 32 位运算要花费与多次 16 位运算相同的时间。当这些考虑因素一般在编码期间并不明显时，倘若应用 H. 26L 规定的这些矩阵值，在解码期间还是会产生额外操作和 32 位运算，因此，开销就很大。

与 H. 26L 规定的矩阵相反，使用如上所述的同样的通用公式表示法，本发明的块变换矩阵为：

$$Q_b^{\Delta} = \text{round}(2.5H) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

它提供下列公式来计算变换系数：

$$\begin{aligned} A &= a + b + c + d \\ B &= 2a + b - c - 2d \\ C &= a - b - c + d \\ D &= a - 2b + 2c - d \end{aligned}$$

此公式/矩阵可最优化工作的部分理由是通过简单地改变换算中使用的值，使乘法因子基本上在量化阶段得以处理，从而替代原来在变换阶段中进行各个乘法。结果，利用此公式的唯一的乘法是乘以 2 的幂，在处理器中此操作是通过简单的左移操作来完成的，并不是真正的乘法。以 u, v, y 和 z 作为辅助参数，其中“ $\ll 1$ ”就是指左移 1 位（等效于乘以 2，但计算更快），上述公式可简化成下述公式：

$$\begin{aligned} u &= a + d; \\ v &= b + c; \\ y &= b - c; \\ z &= a - d; \\ A &= u + v; \\ C &= u - v; \\ B &= y + (z \ll 1); \\ D &= z - (y \ll 1); \end{aligned}$$

通过图 3 示出的蝶形结构（其中 2 点之间的线指示和数，而线上的数字（几个 -1、2 和 -2 指出乘法因子），查找这些变换系数实际上仅需要对每个给定的像素进行 4 次加法和 1 次移位，它与 H.26L 规定的因子相比，可提供显著的性能改进。因为系数值更低，所有变换操作都可在 16 位运算中完成（带有 9 位像素数据）。特别是，使用上述规定的变换矩阵 Q_b ，在 2 维 2-D 变换后，最大信号放大倍数为 36。由此，2 维变换后输出系数比输入要延伸 6 位。因此，对 9 位的输入来说，输出系数具有 15 位的动态范围，也就是 2 维直接变换可用 16 位运算来计算。而且，如下所述，由于操作次数大大减少，而每次操作只需 16 位运算所带来的性能提升在采用逆变换进行解码期间，可再次重现性能提高，这就大大提高了解码性能。

一旦算出变换系数 210，通过对此值换算和舍入它们至它们最近的整数对块系数 210 进行量化。这由图 2 中的量化过程 212 示出，它从量化参数 214 的量化值中进行选择，将变换系数 210 换算为量化系数 216。如下所述，因为在改进的矩阵中不是有 1 个范数，而是 3 个范数，3 种量化表 Q_0, Q_1 和 Q_2 包括这些值，这取决于要量化的变换系数在块中的位置。

换算和舍入的量化公式如下，但它必须用 32 位精度加以计算：

$$L = [K \times A(QP, r) + fX] \gg 20$$

其中 L 是量化系数， K 是变换系数， $A(QP, r)$ 是由量化参数 QP 和 r 所索引的换算因子，其中 r 识别应用哪张表（ Q_0, Q_1 或 Q_2 ）， fX 在 $[0-0.5] \times 2^{20}$ 范围内（ fX 具有

和 K 相同的符号)，而 $\gg 20$ 是指右移 20 位（除以 1,048,576），以使数按比例下降。量化过程因此引入误差。注意，虽然应用 32 位运算进行量化，但这仅是编码所需，当解码时就没有那么重要（其中去量化只需要 16 位精度，如下所述）。

索引 r 可从 3 张量化表 Q_0, Q_1, Q_2 选出 1 张用于根据块中系数 k 的位置来调整 3 个不同的范数：

如果系数来自下列位置 $\{(0, 0), (0, 1), (1, 0), (1, 1)\}$ 中之一，则 $r=0$ （使用 Q_0 表）；

如果系数来自下列位置 $\{(0, 2), (0, 3), (1, 2), (1, 3), (2, 0), (2, 1), (3, 0), (3, 1)\}$ 中之一，则 $r=1$ ，使用 Q_1 表；

如果系数来自下列位置 $\{(2, 2), (2, 3), (3, 2), (3, 3)\}$ 中之一，则 $r=2$ ，使用 Q_2 表。

在右移 20 位后，各个经量化结果 L 拟合到 16 位整数中。

在示范实施例中，根据下列表格，换算因子 $A(QP, r)$ 取决于量化参数 QP 和系数位置组 r ：

```
A(QP=0..31, r=0) = {104858, 93418, 83226, 74146,
66056, 58849, 52429, 46709, 41613, 37073, 33028,
29425, 26214, 23354, 20806, 18536, 16514, 14712,
13107, 11677, 10403, 9268, 8257, 7356, 6554, 5839,
5202, 4634, 4129, 3678, 3277, 2919};
```

```
A(QP=0..31, r=1) = {66318, 59082, 52636, 46894, 41778,
37220, 33159, 29541, 26318, 23447, 20889, 18610,
16579, 14771, 13159, 11723, 10444, 9305, 8290, 7385,
6580, 5862, 5222, 4652, 4145, 3693, 3290, 2931, 2611,
2326, 2072, 1846};
```

```
A(QP=0..31, r=2) = {41943, 37367, 33290, 29658, 26422,
23540, 20972, 18684, 16645, 14829, 13211, 11770,
10486, 9342, 8323, 7415, 6606, 5885, 5243, 4671, 4161,
3707, 3303, 2942, 2621, 2335, 2081, 1854, 1651, 1471,
1311, 1168}.
```

上述特定值是设计用来满足 H.26L 中设计技术规范，量化步骤规模必须在表索引中每 6 次递增时，加倍。对于其他视频或静止图像编码应用，要考虑所需的保真度增量，设计另外的量化表。注意，为了将量化表存储于编码器中，需要有额外空间来存储 3 张表，而不是 1 张表，但是该额外空间量仅为 64 字节，故微不足道。

再看图 2，经量化过程后，量化系数 216 馈至熵编码器 218，进一步减少块编码时所需的位数。熵编码器（和解码器）是大家熟知的，故这儿不再累述，因为本发明操作前，基本上与熵编码无关，而操作后基本上也与熵解码无关。

考虑在块变换解码器 222 中解码时，有时不论如何传递，熵编码的输出位馈至熵解码器 224 作为输入位。一般说来，这种熵解码器是已知的，因此在其他可能的操作中，它足以说明熵解码器 224 对给定的块再生量化系数 226。注意，一般编码和解码是不能有损的，即馈至熵编码器 218 的量化系数 216 与由熵解码器 224 所产生的量化系数 226 是一致的。

一般说来，块变换解码器 222 反映块变换编码器 200 的操作，虽然如下所述，且为了和本发明一致，提供一种改进的逆变换矩阵，能应用 16 位运算于整个解码过程中，这样，大大简化了解码器的计算复杂性。

一旦量化系数 226 从熵解码器 224 中取出，每个量化系数 L 由去量化处理 228 转换为重构（去量化）值 K' ，去量化处理执行下述公式：

$$K' = L \times B(QP, r)$$

其中换算因子 B 取决于编码时使用的 QP 索引，而 r 通过从 3 张表 $D0$, $D1$ 或 $D2$ 中选用的 1 张表的 1 个参数来决定去量化参数 230（以调整 3 个范数）。注意，如上面有关量化所述， r 能从块中正在去量化的系数的位置中推导出来，虽然从编码器中送出 r 值是可行的（这能减少解码器上计算的复杂性），但这要求对每个系数至少要多送出 2 位，以便标识 r 。

这样，换算因子 $B(QP, r)$ 取决于编码时所用的 QP 量化参数索引和系数位置组 r 。在示范实施中，根据下表规定相关性：

$B(QP=0..31, r=0) = \{80, 90, 101, 113, 127, 143, 160, 180, 202, 226, 254, 285, 320, 359, 403, 453, 508, 570, 640, 718, 806, 905, 1016, 1140, 1280, 1437, 1613, 1810, 2032, 2281, 2560, 2874\};$

$B(QP=0..31, r=1) = \{101, 114, 127, 143, 161, 180, 202, 227, 255, 286, 321, 361, 405, 454, 510, 572, 643, 721, 810, 909, 1020, 1145, 1285, 1443, 1619, 1817, 2040, 2290, 2570, 2885, 3239, 3635\};$

$B(QP=0..31, r=2) = \{128, 144, 161, 181, 203, 228, 256, 287, 323, 362, 406, 456, 512, 575, 645, 724, 813, 912, 1024, 1149, 1290, 1448, 1625, 1825, 2048, 2299, 2580, 2896, 3252, 3650, 4095, 4596\}.$

当使用编码表时，上述的特定值是设计满足 H.26L 的设计技术规范的，即量化步骤的规模必须是索引表中每增量为 6 就加倍。对其他视频或静止图像编码应用，考虑到所需保真度的增量，可设计其他的量化表。

可以理解的是，虽然这些换算值中的每个值都足够低，以保证解码器只需要 16 位运算，但去量化公式又使系数换算回成为较大值。必须注意的是，因为 1 个表中的各项与其他 2 个表中相应的 QP 索引的项在数学上是相关的（基于范数），

就可能换个办法，根据 r 值作出适当的数学调整来只使用 1 张表的 QP 值，而不是用 3 张表。然而，在小表中查找更为有效，并且以每个项 2 字节空间来存储 3 张表的 32 项代替（H.26L 变换规定所需的）每项以 4 位空间来存储 1 张表的 32 项所需的字节数是 64 个附加字节，这在现代计算设备中是可忽略的，特别考虑到由去量化生成的变换系数拟合到 16 位字中，而不是在 32 位字中，因而可减少所需的存储器总量。

图 2 中又示出，一旦经量化系数 226 通过上述公式和表进行去量化后，就产生重构变换系数 232 的 4×4 矩阵。根据这些变换系数 232，通过将它们馈至分别由列和行逆变换处理 236, 238 组成的 2 维逆变换处理 234 中去，就能生成重构像素。

为了与本发明相符，不采用平常所用的逆变换 $Q_i = Q_b^T$ ，而是允许 16 位去量化和逆变换计算，本发明规定下列逆变换矩阵供使用：

$$Q_i = \begin{bmatrix} 1 & 1 & 1 & 1/2 \\ 1 & 1/2 & -1 & -1 \\ 1 & -1/2 & -1 & 1 \\ 1 & -1 & 1 & -1/2 \end{bmatrix}$$

注意， Q_i 诸列是相互正交的，但它们范数是不同的。然而这可通过在选择去量化参数中所用的“ r ”值来处理，如上所述。而且，注意由 Q_b 生成 Q_i 时， Q_b 是转置的，通过列 1 和列 3 乘以 $1/2$ 。再之，在 D0、D1 和 D2 表中已经存在的 QP 值已为此而事先加以调整，故不需要任何附加计算复杂性以进行补偿。而且，使用此矩阵在计算逆变换时必须进行的唯一“乘法”是乘以 $1/2$ ，这实际上是通过高效率的右移操作来完成的。注意，这种不精确的除法引起的少量噪声基本上对速率失真性能没有影响。

当 Q_i 如上规定，其最大增益等于 4。因此，2 维逆变换过程的动态范围只扩展 4 位，故可在 16 位运算中计算。

图 2 示出的列逆变换处理 236 和行逆变换处理 238（在列逆变换处理的结果上操作）生成输出像素数据 240。至此，通过使用上述逆变换矩阵和调用 $[A \ B \ C \ D]$ 为特定的逆变换的行或列，相应的逆变换值 $[a' \ b' \ c' \ d']$ 由下式算出：

$$\begin{aligned}
 u &= A + C; \\
 v &= A - C; \\
 y &= (B \gg 1) - D; \\
 z &= (D \gg 1) + B; \\
 a' &= u + z; \\
 b' &= v + y; \\
 c' &= v - y; \\
 d' &= u - z;
 \end{aligned}$$

其中， u, v, y 和 z 是辅助变量，而 $\gg 1$ 是指右移 1 位（等效于乘以 $1/2$ ）。相似于编码，通过图 4 所示的蝶形结构，这些公式实际上对每个系数已减少至 4 次加法操作和 1 次移位操作，都在 16 位运算中进行。最后，逆变换处理 234(图 2)以 2^7 因子来换算重构的像素值（通过 1 次 7 位右移操作来完成）。

在 7 位移位后，重构的像素值具有 9 位范围，所以逆变换可在 16 位运算中计算。上述设置的去量化表 D0、D1 和 D2 是这样的，即去量化公式在以 16 位运算进行计算时不会溢出。下述表格总结了使用本发明变换和逆变换在计算上与 H. 26L 所规定的操作相比存在的一些优点：

变换	每像素操作次数，直接变换	量化	去量化	每像素操作次数，逆变换
H. 26L 所规定	4 次加法，3 次乘法，在 32 位运算中进行	1 次乘法，在 32 位运算中进行	1 次乘法，在 32 位运算中进行	4 次加法，3 次乘法，在 32 位运算中进行
本发明	4 次加法，1 次移位，在 16 位运算中进行	1 次乘法，在 32 位运算中进行	1 次乘法在 16 位运算中进行	4 次加法，1 次移位，在 16 位运算中进行

现在很好理解，使用本发明所带来的这种节省是很显著的，特别在 16 位的处理器上。例如，实际上，一种典型的 16 位 PDA 处理器通过应用本发明，可以在解码方面具有速度提高 1 倍左右的改善。

而且，已经对变换/逆变换对质量和压缩的影响作了测试。编码增益的结果（一般定义为信噪比提高）说明：使用本发明的变换方式与 H26L 变换相比，以测试数据的结果来看，损失仅为 0.01dB，这是可以忽略的，事实上有可能要比实际的视频信号信息的损失小得多，如像素预测误差。而且，相对于 H. 26L 规定的变换，在 32 个量化参数的每个参数上作了总的性能测试，结果只是 $\pm 0.5\%$ 范围的差异，而

平均值又十分接近零，因此表明本发明的变换和 H. 26L 规定的变换一样好。

从上面详细描述来看，现提供一种改进的变换和逆变换矩阵分别用于图像或视频的编码和解码，比起其他现有的变换来说，可大大地减少计算复杂性，并对压缩和质量没有不利影响。不需要变换的乘法，但仍能取得正确的结果，因为除了在变换和逆变换处理中移位操作外，量化和去量化参数值都对这些乘法因子进行补偿。由于使用此参数值，编码时变换操作和解码时去量化和变换操作都能在 16 位运算中完成（像素数据由 9 位或更少位来表示）。

本发明可以有各种修改和其他结构，在图中示出说明实例例，并在上面作了详述。然而，应该理解的是，并不是要将发明局限于所揭示的具体形式内，相反，本发明是要包括包含在本发明的精神和范围内的所有修改、替代结构和同等物。

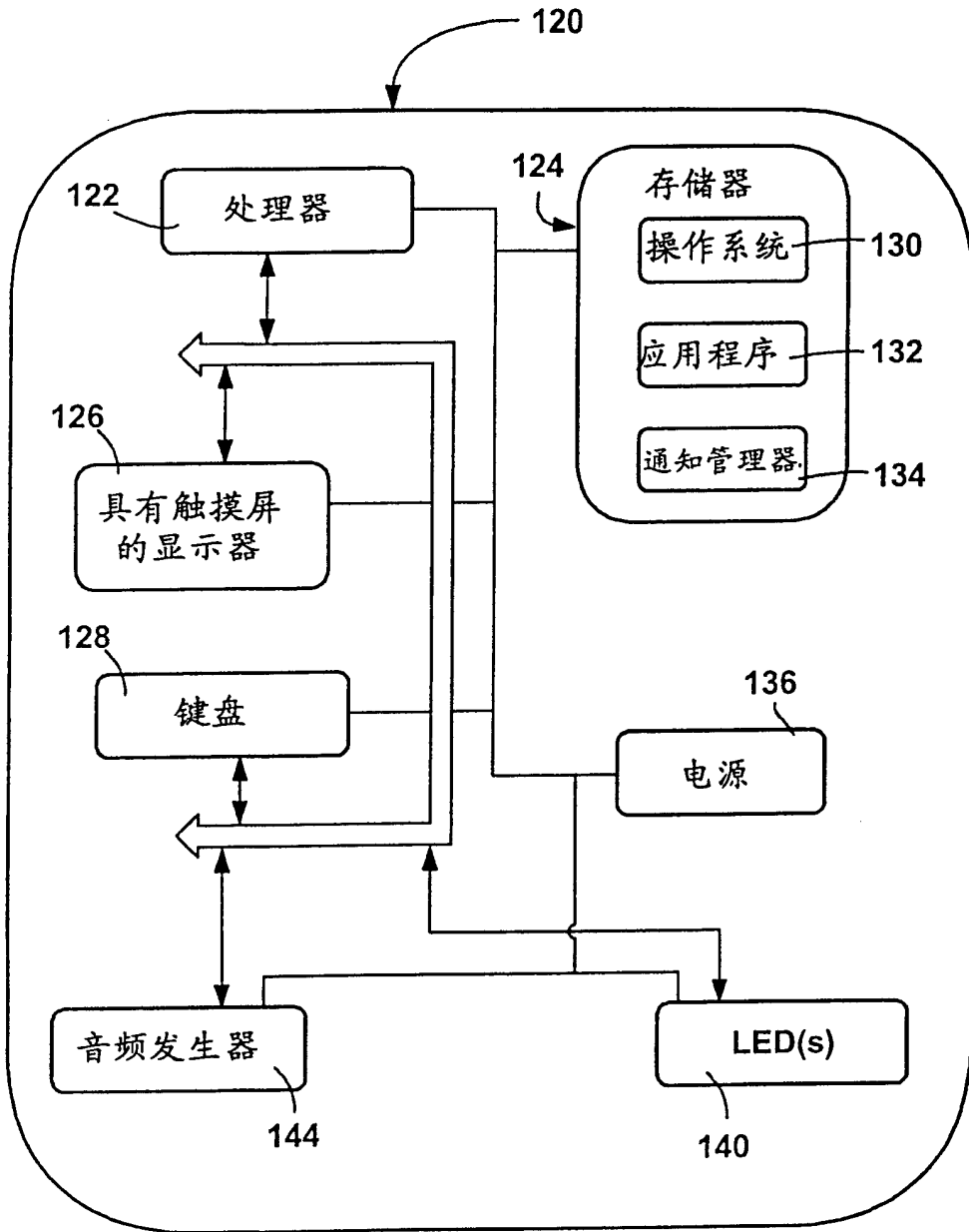


图 1

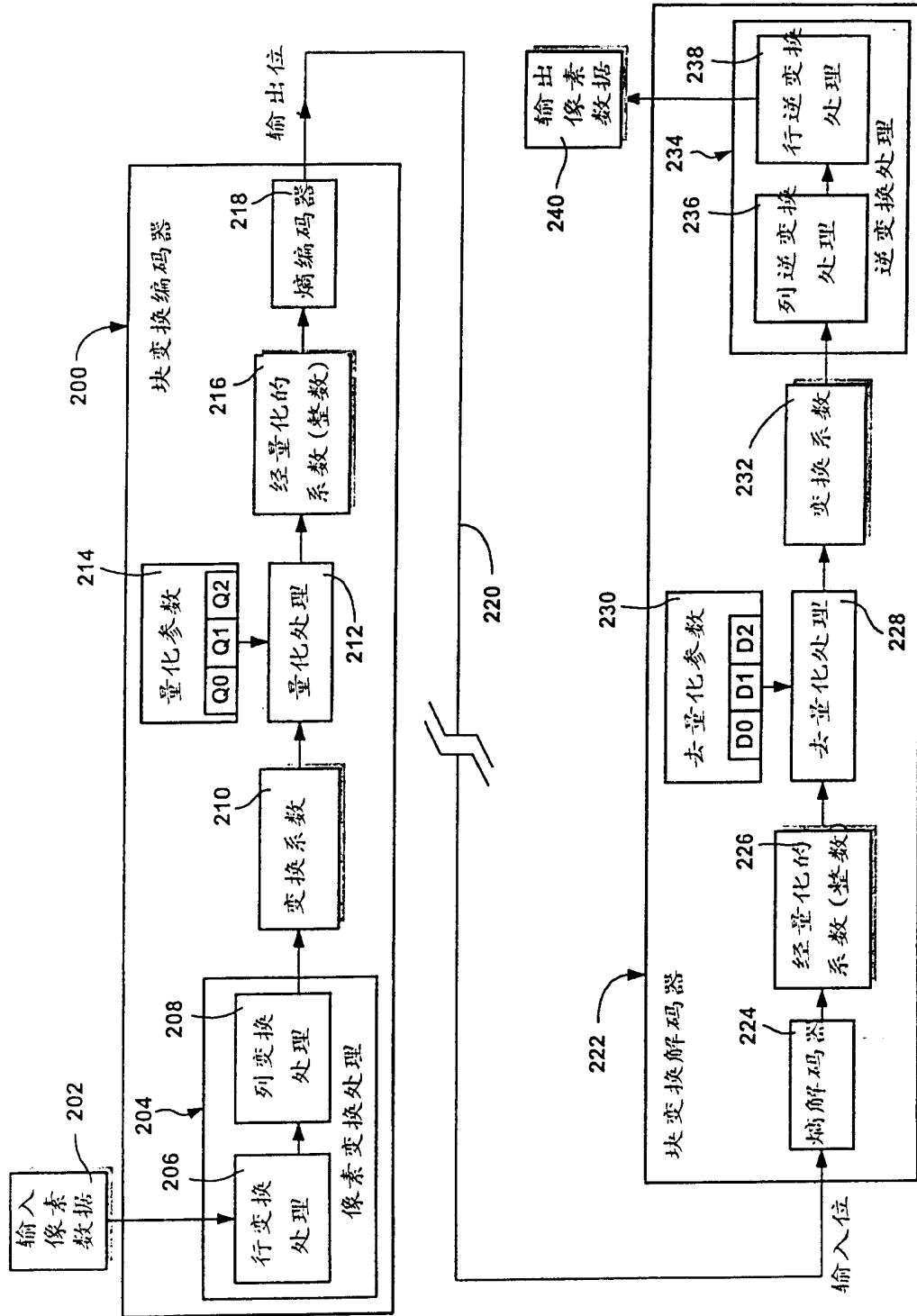


图 2

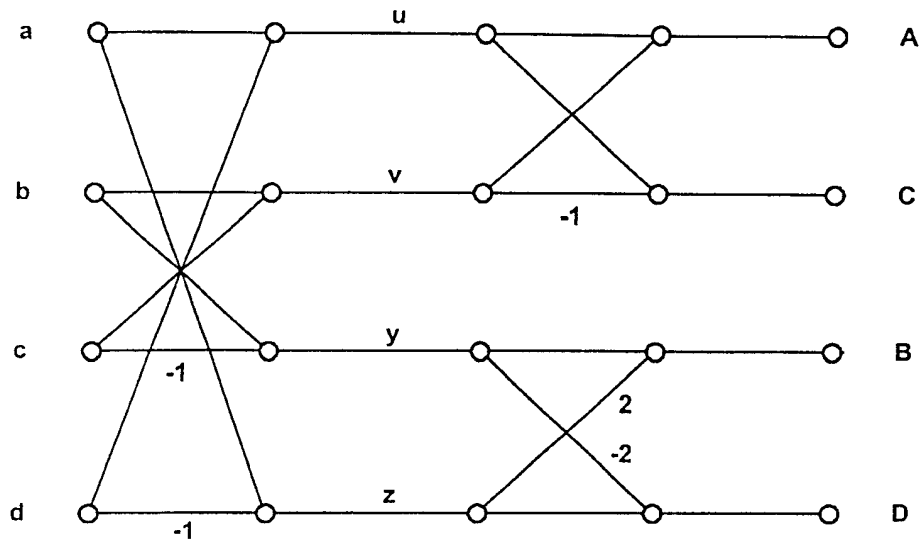


图 3

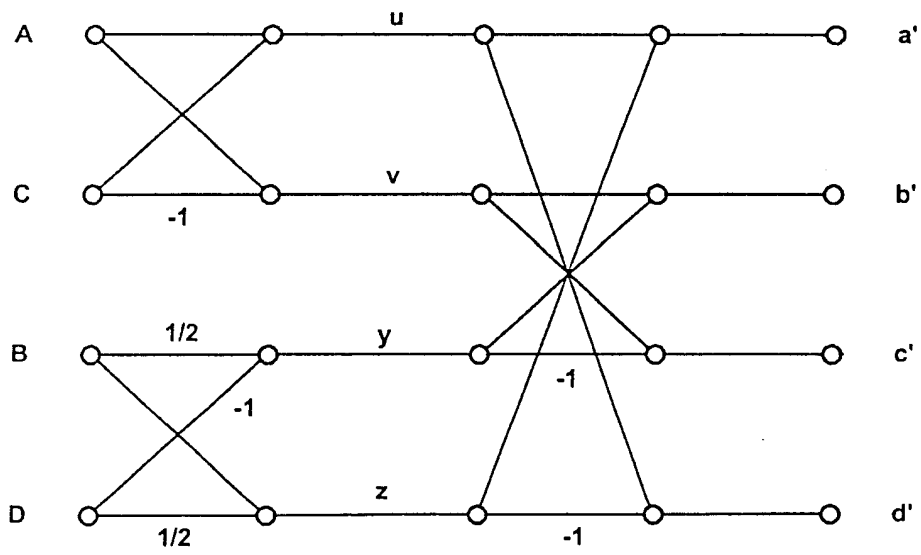


图 4