

[19]中华人民共和国专利局

[51]Int.Cl⁶



[12] 发明专利说明书

H04N 7/12

H04N 11/02

H04L 29/00

[21] ZL 专利号 92111430.3

[45]授权公告日 1997 年 10 月 29 日

[11] 授权公告号 CN 1036303C

[22]申请日 92.10.9 [24]颁证日 97.8.16

[21]申请号 92111430.3

[30]优先权

[32]91.10.31[33]US[31]07/786,233

[73]专利权人 国际商业机器公司

地址 美国纽约

[72]发明人 塞瑟尔·奥古斯托·贡扎里斯

多夫·N·莱姆

艾里克·N·维西托

[74]专利代理机构 中国国际贸易促进委员会专利商标

事务所

代理人 范本国

[56]参考文献

EP0280313

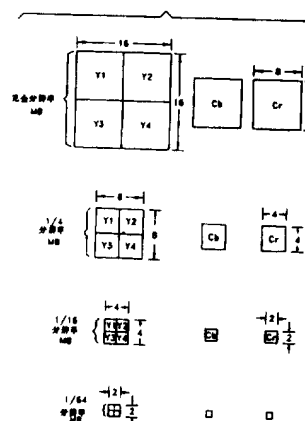
审查员 陈 源

权利要求书 2 页 说明书 24 页 附图页数 11 页

[54]发明名称 具有多分辨率特征的运动视频信号压缩系统

[57]摘要

本发明为用于处理一视频图象数据串以产生对应于分辨率或比特串标度复用的数据的视频表示的系统和方法。这种表示可在所有分辨率和比特串标度下保存 MPEG-1ISO 标准的基本宏观块 (MB) 的身份。MB 与一系列可影响包括在 MPEG-1 压缩数据串中的开销数据量的特性有关, 因此通过在多个分辨率和比特串标度下保存 MB 身份, 这些标度就可分享这种开销, 从而只需要在数据串中将其包括一次。



权 利 要 求 书

1.一种从包含多级空间分辨率标度数据的压缩视频信息中以目标空间分辨率对运动图象进行译码的方法，其特征在于包括如下步骤：

(a)从一个基础空间分辨率标度开始对宏观块属性和转换系数数据进行解压缩，

(b)对对应于分级中下一个更高的空间分辨率标度的宏观块属性的细微数据和附加的转换系数数据进行解压缩，

(c)对于分级中更高的空间分辨率标度，重复步骤(b)，直至达到目标空间分辨率标度，以及

(d)将解压缩后的转换系数数据加以组合，形成一种运动图象的转换表示。

2.如权利要求1所述的方法，其特征在于，上述的宏观块属性可以适当地加以修改，以便符合上述目标空间分辨率标量的空间分辨率标置。

3.如权利要求1所述的方法，其特征在于还包括以下步骤：当发现上述宏观块中的信息有差错时忽略至少部分压缩信息。

4.一种生成运动图象的压缩视频数据表示的方法，所述表示中包括多级图象空间分辨率标度，其特征在于包括如下步骤：

提供对应于基础空间分辨率标度的表示运动图象的压缩视频数据，所述的数据中包括对于全部分级共同的宏观块属性以及相应的基础分辨率转换数据，

提供对应于至少一个更高空间分辨率标度的附加数据，所述的附加数据中包括相应的更高分辨的转换数据，

其中，对应于分级中的各个空间分辨率标度的转换数据被用来以可以预测的方式生成分级中下一更高空间分辨率标度的更高分辨率转换数据。

5.如权利要求4的方法，其特征在于：上述的对应于更高空间分辨率标度的附加数据中包括基础分辨率标度下的宏观块属性的细微数据。

说 明 书

具有多分辨率特征的运动视频

信号压缩系统

本发明涉及数据压缩领域,更具体地说,是涉及在多种标度下压缩和解压数字运动视频信号的系统和技術。該技術扩展了与国际标准组织的运动图象专家组 (*International Standards Organization's Moving Picture Experts Group*) (*MPEG*) 提出的 *MPEG* 标准类似的方法。

数字传送网络、数字存贮介质、超大规模集成装置、以及视频和音频信号的数字压缩方面的技术进展,都趋向于使数字视频信号的存贮和传送在广大的应用范围内变得更加经济。由于数字视频信号的存贮和传送对众多应用都是至关重要的,且视频信号的未压缩表示要求存贮量大,故在此先进技术中数字视频压缩技术非常关键。因此,在过去十年中出现了几个用于数字视频信号压缩的国际标准,且更多的在正在出现中。这些标准适用于在各种应用中传送及存贮压缩数字视频信号的方法,包括:可视电话和电话会议;同轴和光纤网络及地面广播和直接广播卫星上的高质量数字电视传送;及 *CD-ROM*、数字录音带、及光盘装置上的交互作用的多媒介产品。

这些标准中的几个包括基于压缩技术的公共核的方法,如 CCITT (国际电报及电话咨询委员会, *Consultative Committee on International Telegraphy and Telephony*) 建议 H. 120、CCITT 建议 H. 261、及 ISO/IEC MPEG 标准。此 MPEG 法是由运动图象专家组 (MPEG) 建立的,该组织是国际标准组织 (ISO) 和国际电技术协会 (IEC) 的联合技术委员会的一部分。该 MPEG 委员会正为视频和音频信号的多重复用压缩表示建立一草案标准。该标准规定了压缩比特串的句法及在一定空间分辨率下解译数字视频信号的方法。该草案标准将被称作 MPEG—1 标准或算法,以区别于该委员会时下正在讨论的更新的算法。此 MPEG—1 草案标准在 1991 年 5 月的文件 ISO/IEC JTC1/SC2 WG11 MPEG 91/090 中作了描述。

由于本发明可被用于扩展 MPEG—1 解码器的功能,以从同一压缩比特串产生多种视频分辨率,故将提到 MPEG—1 视频压缩算法中的一些有关的部分。但应注意的是,本发明还可用于其他的、具有 MPEG 法的某些特征的视频编码算法。

MPEG—1 视频信号压缩算法

首先应理解的是,任何数据体,如书中的一页,一幅图象、一段讲话、或一系列视频信号,它们的压缩均可被视为一系列步骤,包括:1) 将该数据体分解成一组符号;2) 用在某些情况下有最小长度的二进数串来表示这些符号;3) 以确定的顺序连接这些数串。步骤 2 和 3 是

无失真的,即在逆变换时原始数据是可复原的,而且 步骤 2 也叫熵编码。(参见 T. BERGER 的 *Rate Distortion Theory*, NJ: Prentice—Hall, 1977; R. McELIECE 的 *The Theory of Information and Coding*, Reading, MA: Addison—Wesley, 1971; D. A. HUFFMAN 的 *A Method for the Construction of Minimum Redundancy Codes*, “*Proc. IRE*”, pp 1098—1101, September 1952; G. G. LANGDON 的 “*An Introduction to Arithmetic Coding*”, *IBM J. Res. Develop*, Vol(28), pp. 135—149, March 1984)。步骤 1 通常即可是无失真的,也可是失真的、多数视频压缩算法是失真的。成功的失真压缩算法消掉多余和无关的信息、在错误看上去不明显时能容许较多的错误、并且刻意表现人类观察者非常注意的序列部分。在 MPEG—1 法中采用的步骤 1 的技术可称为预测/内插运动补偿混合 PCT/DPCM 编码 (*Predictive/interpolative motion—compensated hybrid DCT/DPCM coding*)。Huffman 编码,也叫可变长度编码(见上面所列的 Huffman 1952 年的文章)被用在步骤 2 中。尽管 MPEG—1 标准实际上是一解码器和压缩比特串句法的规范,但为说明方便起见,以下的 MPEG—1 规范描述主要是从编码器的角度进行的。

MPEG—1 视频标准规定了数据存贮介质的视频编码表示,如 1991 年 MPEG 委员会草案 ISO—IEC JTC1/SC2/WG11 MPEG CD—11172 中所规定的那样。该算法是为在无交错分量视频信号中

使用而设计的,尽管它可通过把两个依次的交错场合成一个图象而扩展应用到交错视频图象上。每个图象有三个分量:亮度(y)、红色差(Cr)和兰色差(Cb)。 Cr 和 Cb 分量各自在垂直及水平方向的取样均只有 y 分量的一半。另外,此算法是以单一的视频分辨率水平进行的。

MPEG—1 序列的分层结构

一个 MPEG—1 数据串由一视频数串和一音频数串组成,这些数串与系统信息及其他可能出现的比特串一起,被装入一可视为有分层结构的系统数据串中。在 MPEG—1 数据串的视频层中,压缩数据被进一步分层。最高的一层是视频序列层,它包含控制信息及整个序列的参量。对其他层的组成的描述会有助于理解本发明。MPEG—1 视频分层结构的这些层在图 1 至 4 中显示。在这些图中:

图 1 显示了图象组(GOP);

图 2 显示了图象的宏观块划分;

图 3 显示了图象(例)的片划分;

图 4 显示了宏观块的块划分。

这些层关系到压缩算法操作及压缩比特率组合。如前所述,最高层是包含控制信息及整个序列的参量的视频序列层。在下一层中,序列被分成依次图象的组,每个被称为图象组(GOP)。在图 1 中概括地显示了这种层。解码可始于任何 GOP 的开始处,基本上与前面的 GOP 无关。GOP 中的图象数目是没有限制的,所有 GOP 中的图

象数目也不必是相同的。

第三或图象层是一幅图象。图 2 中一般地显示了该层。解码可始于任一 *GOP* 的开始处；各图象的亮度分量基本被分成 16×16 个区，而色差分量则被分成与 16×16 个亮度区在空间上同处 (*Co-sited*) 的 8×8 个区。同处的亮度区和色差区一同构成了第五层，叫做宏观块（以下称为 *MB*）。

在图象和 *MB* 层之间是第四或“片”层。各片由任意或选定数目的依次 *MB* 组成。一个图象内或不同图象中的片不必是大小一致的。它们的大小可以仅有几个宏观块，也是延伸过几列 *MB* (如图 3 所示)。

一个 *MB* 是一基本层，它可联系有各种特性，这在下面可以看出。如图 4 所示，*MB* 的基本结构包括四个亮度块和两个度块。在 *MPEG-1* 中，所有这些块的大小均为 8×8 。本发明的一个目的，是在多个图象分辨率下保存 *MB* 的结构和特性（不一定是其大小）。

在一个 *GOP* 内，会出现三种类型的图象。各类图象的区别在于所用的压缩方法。“内模式”图象或 *I* 图象是与其他图象独立地压缩的。尽管 *I* 图象的间距没有上限，但希望它们的频繁地在整个序列中，以利于随机存取及其他模式的操作。各 *GOP* 必须以一个 *I* 图象开头、且其他的 *I* 图象可出现在此 *GOP* 中。其他类型的图象，即预计运动补偿图象 (*P* 图象) 和双向运动补偿图象 (*B* 图象) 将在下面对运动补偿的讨论中描述。在图 5 中给出了一般的说明。

运动补偿

大多数视频信号序列都表现出依次的图象之间的高度相关性。在编码图象之前消除这种冗余的一种有利方法是“运动补偿”。运动补偿需要有装置来模拟并测定一个场面中的运动。在 *MPEG-1* 中, 各图象被分成宏观块, 且把各 *MB* 与一或多个预计图象中处于相同大致空间位置的 16×16 个区相比较。预计图象中与某些场面中的 *MB* 最匹配的区被作为预计。*MB* 的空间位置与预计者的空间位置间的差别被称为“运动矢量”。这样, 用于一 *MB* 的运动估计和补偿就是运动矢量和运动补偿差宏观块。这些一般可得比原有 *MB* 本身更多地压缩。用正向的 (即在顺序中在时间上是向前的) 单一预计图象来预计运动补偿的图象, 叫做 *P* 图象。

在 *MPEG-1* 中, *P* 图象和其预计图像间的时间间隔可大于一个图象间隔。对 *P* 图象之间或 *I* 图象与 *P* 图象间的图象, 除了时间上向前的预计外, 也可采用时间上反向的预计, 这种图象叫双向运动补偿图象 (*B* 图象)。对 *B* 图象, 除了正向和反向的预计, 还允许帧内插运动补偿, 其中的预计为来自以前预计图象的块与来自未来预计图象的块的平均。在此情况, 需要两个运动矢量。

采用双向运动补偿导致了一个两级运动补偿结构, 如图 5 所示。各箭头表示用与点相接触的图象对触到箭头头部的图象所作的补偿。各 *P* 图象均用前一 *P* 图象 (或 *I* 图象, 如实际可能出现的) 进行运动补偿。各 *B* 图象均用紧邻它的前或和后一幅 *P* 或 *I* 图象进行

运动补偿。这些预计图象有时叫做“锚”图象。对锚图象 间距及 I 图象们的间距, MPEG-1 中都没作限定。实际上, 在一个整个序列中, 这些参数不必是常数。若以 N 表示 I 图象间的距离而以 M 表示 P 图象间的距离, 则图 5 所示的序列有 $(N, M) = (9, 3)$

因此应当理解, MPEG-1 序列由一系列 I 图象组成, 这些 I 图象间可夹有一或多个或根本没有 P 图象。这些 I 和 P 图象间可根本没有或夹有一或多个 B 图象, 在后一种情况下它们是锚图象。

MB 的变换和量化(quantization)

一种非常有用的图象压缩技术是变换编码。(参见 N. S. JAYANT 和 P. NOLL 的 "digital Coding of Waveforms, principles and Applications to Speech and video", Englewood Cliffs, N. J; Prentice-Hall, 1984 和 A. G. TESCHER, "Transform Image Transmission Techniques, pp 113-155, New York, NY: Academic Press 1979.) 在 MPEG-1 和其他几种压缩标准中, 分离余弦变换(DCT)是所适用的变换 (见 K、R、RAO 和 P、YIP 的 "Discrete Cosine Transformation, Algorithms, Advantages, Applications", San Diego, CA: Academic Press, 1990, 和 N. AHMED、T. NATRAJAN 和 K、R、RAO 的 "Discrete cosine Transformation, " IEEE Transactions on Computers, pp. 90-93, January 1974)。比如, 一幅 I 图象的压缩, 是通过在 MB 内取亮度及色度象素 (pixel) 的块的 DCT、量化该 DCT 的系数、并对其结

果作 *Huffman* 编码而实现的。类似的原则也适用于 *P* 和 *B* 图象的压缩，只是在这些情况下，*DCT* 可被用于一个 *MB* 中的象素块之间的差和它们的相应运动补偿预计。*DCT* 把 $n \times n$ 个象素转换成 $n \times n$ 组变换系数。*DCT* 在压缩变换中是非常有用的，因为它倾向于把象素块数据的能力集中到几个 *DCT* 系数之中，而且这些 *DCT* 系数几乎是相互独立的。象几个国际压缩标准，*MPEG-1* 法采用了一个 8×8 的 *DCT* 块，这时应于 *MB* 内的块的大小。本发明的一个目的，是利用更大或更小的 *DCT*，从而变量 *MB* 的大小，以支持有多重分辨率的图象。

下一个步骤是 *DCT* 系数的量化，这在 *MPEG-1* 法中是主要的失真源。用 C_{mn} 表示 *DCT* 系数的二维阵列中的元素，其中 m 和 n 可从 0 至 7，则除了舍位和近似修正之外，量化是通过用 $W_{mn}XQP$ 来除各个 *DCT* 系数而实现的，其中 W_{mn} 是加权因子，而 QP 是量化参数。加权因子 W_{mn} 使较粗略的量化能被用于看上去较次要的系数。这些权可以有两组，一组用于 *I* 图象，另一组用于 *P* 和 *B* 图象。常规的权可在视频序列层中传送。量化参量 QP 是在 *MPEG-1* 中权衡质量和比特率的主要手段。应注意的是，在一幅图象中不同 *MB* 的 QP 可以不同。还应注意到，在本发明中，可选择为其他大小的 *DCT* 提供单独的加权矩阵，或提供在数字上关联的、大小不同的加权矩阵，以利于解码。

在量化之后，各 *MB* 的 *DCT* 系数信息得到组织并用 *Huff-*

man 码进行编码。此步骤的细节对理解本发明并不是重要的,因此这里就不作说明了,但其进一步的情况可参见前面引用的 *HUFFMAN*1952 年的文章。

与运动补偿有关的宏观块特性

可以想见,有三种运动补偿可被用于 *MB*:正向、反向和帧内插运动补偿。编码器须选择这些模式中的一种。对某些 *MB*,没有一种运动补偿能给出精确的预计。在这种情况下,可将此 *MB* 选出,以作内模式编码,就象对 *I* 图象那样。这样,根据运动补偿模式,*MB* 可以是下列型式的:

- * 正向
- * 反向
- * 帧间内插
- * 帧内内插

在 *P* 图象中,根据运动矢量的值,*MB* 可以是带零运动矢量或非零运动矢量类型的。这些类型连同所需运动矢量数据与各 *MB* 一起被编码,作为开销 (*overhead*) 数据。例外是忽略的 *MB*,如下面所要描述的。以下说明与变换及量化有关的宏观块特性。

如前所述,*QP* 参量可随不同的 *MB* 而改变,当这种改变出现时,用附加的 *MB* 类型表示应该使用新的 *QP*。该新 *QP* 值本身与 *MB* 一同传送。

当时 *MB* 中的块作 *DCT* 及量化后,有些块会只含有零。这些块

不用再作编码,并用所谓编码块图案码来标示。这种码代表额外的,开销项目。

最后,每当 *MB* 不含新信息时,都可将其略去。为传达这个信息,把一 *MB* 地址与各未省略 *MB* 一同传送。

在此应注意,*MB* 们携带着一系列的特性,这些特性可通过使各编码 *MB* 包含开销数据而得到描述。本发明的一个目的,是在多种量度下保存 *MB* 们的身份,从而使经常项目只用包含一次,当然在某些参数(如运动矢量的精确度)的精细化方面除外。

因此,从前面对 *MPEG-1* 视频信号法的描述,应当理解到 *MPEG-1* 的目的,是规定于视频序列的压缩比特串的句法,并规定在单一的空间分辨率水平下解码该序列所用的方法。本发明所要解决的问题,是扩展 *MPEG-1* 的句法及解码方法规范,从而可以多种比例(*scale*)对数字视频序列进行解码和编码。为这些目的,有两种定标(*scaling*)类型:

1. 分辨率定标:指产生一种比特率的能力,这种比特串可通过选择该比特串的不同部分而以多种空间分辨率进行解码。此特征在要把多个视频窗口显示地完全分辨率屏幕上的某些应用中是需要的。另外,它的有用之处还在于,它可使解码器能在各种复杂程度下使用,从而有可能用极简单的解码器来解译较低的空间分辨率。

2. 比特串定标:指产生一种比特串的能力,该比特串中的某些编码比特可被省去,而有用的图象仍能产生。分辨率可定标算法也是比

特串可定标法;但在此比特串定标是在更窄的意义上理解的。比特串可定标性总是用来指扩展到完全分辨率。这在有些压缩比特串信息因噪声而受损坏时,对适当降低扩展视频信号的质量是非常有用的。

本发明的目的,是通过提供一种支持由 8×8 DCT 分量的等级编码定标的比特串和分辨率的装置和方法,来扩展 MPEG-1 解码方法,以支持这两种形式的定标。这种方法是灵活的:一个或多个这些分辨率标度可叠加起来,以支持许多的应用。分级层的具体数目取决于编码器的设计人或应用要求。另外,本发明还解决扩展 MPEG-1 的句法及方法以在不同的分级层上保存宏观块特性的问题。

关于本发明涉及的等级编码技术,有众多的技术文章。这些文献中许多涉及次带(subband)编码和锥形(pyramid)编码。与此主题有关的最近的两本书是:

* “Subband Image Coding”, J. W. WOODS, editor, Kluwer Academic Publishers, 1991.

* “Digital Image Compression Techniques”, M. RABBANI and P. JONES, SPIE Optical Engineering, Bellingham, WA, USA, 1991.

下列三个文件据信是与本发明最相关的先有技术:

* “Alternative to the hierarchical scheme”, Ch. GUILLEMOT, T. N'GUYEN, and A. LEGER, ISO/JTC1/SC2/WG8,

JPEG N—260, February 1989。

* “Setup of CCIR 601 multi — purpose coding scheme , ”
PTT RESEARCH, the Netherlands, ISO/IEC JYC1/ SC2/ WG11
MPEG 91/051, May 1991。

* “Compatible Coding of CCIR 601 Image : Predict the prediction error , ” PTT RESEARCH, the Netherlands, ISO/JTC1/ SC2/ WG11 MPEG . 91/114, August 1991。

第一个文件描述了压缩多分辨率静止图象的分级方案,其中用低分辨率图象的 *DCT* 系数来预计较高分辨率的 *DCT* 系数。该方案与本发明的不同之处是 *DCT* 的大小总是相同的。另外,也没有考虑采用压缩被分成有共同特性的象素的块的影象的方案。第二个和第三个(尤其是第三个)文件描述的方案在应用于两层分辨率标度时与本发明相类似。这些文件中的该方案的目的,是利用具有两级分辨率标度,以产生与 *CCIR* 601 格式和 *MPEG—1 SIR* 格式相兼容的图象。但没有试图将该技术扩展到在分辨率及比特串方面可定标的编码器。另外,在两个标度水平只未保持宏观块的特性,特别是在较高分辨率标度下的运动补偿矢量被定在 16×8 块的水平,而不是 36×16 的 *CCIR* 601 宏观块水平。这意味着各分辨率标度并不都具有这一特性。另外,在如何处理其他特性上也未给出细节。

与前述先有技术系统和方法不同,本发明的目的在于提供一种用于压缩图象信号序列的灵活句法和编/解码方案,这些使得能在多

种空间分辨率和比特串标度下对视频信号进行解码。本发明的另一目的,是提供用于以扩展 MPEG-1 标准的已有句法及方法的方式来支持多种标度的系统和算法。这种扩展是在只加上最少额外经常信息的情况下进行的。

本发明的再一个目的,是提供与本发明的系统和算法相符合的编码器和解码器实施方案。

本发明包括用于以产生图象数据的一种表示的方式处理一系列视频图象数据的系统和方法,此表示根据分辨率及比特串标度对数据进行多路。该表示可在所有分辨率及比特串标度下保存 MPEG-1 的基本宏观块(MB)结构的身份。图 6 显示 MB 身份是如何通过在四个分辨率水平上定标而得以保存的。保存此身份的重要性在于 MB 是与一系列特性相联系的,这些特性影响着包含在 MMPEG-1 压缩数据串中的开销数据量。通过在多种分辨率及比特串标度下保存 MB 身份,这些标度就能分享这些开销数据,从而只需将其在数据串中包含一次。

保存 MB 身份大大简化了除最高分辨率以外的所有分辨率标度下的运动测定矢量数据的导出。对应于任何分辨标度的运动矢量数据基本上都可通过适当降低最高分辨率运动矢量数据的标度而从其导出。比如,1/4 分辨率的 x 和 y 运动矢量分量是相应的完全分辨率分量的 1/2。另一方面,完全分辨率运动矢量可通过适当提高较低分辨率运动矢量的标度而导出。在任一情况下,可把附加的修正加在

较高的分辨率标度,以改善运动矢量数据的精度。

本发明的第二个方面是 *MB* 的编码方法也得到了保存。如前所述,在 *MPEG-1* 中,一个 *MB* 被分成六个 8×8 的亮度和色度信息块,每块均用 8×8 分之余弦度换(*DCT*)编码。在本技术中,各标定的 *MB* 也被分成六个亮度和色度信息块,各块用适当大小的 *DCT* 编码。这种对图 6 中 $R\ 1/4$ 分辨的 *MB*,将用 4×4 大小的 *DCT*。另外还应注意,这种 4×4 大小的 *DCT*,作为一个例子,还可以多种方式从相应的 8×8 *DCT* 系数导出,这样就不必直接执行 4×4 的 *DCT*。

最后,本发明的第三个方面涉及产生以最终分辨率标度重建图象所需的 *DCT* 块数据的方法。在本发明的技术中,任何分辨率或比特串标度下的 *DCT* 系数数据,均由采用标度等级较低的 *DCT* 数据作为预测值的标准微分编码技术,进行了编码。

图 1—4 显示了 *MPEG-1* 数据串的视频压缩层中的压缩数据层,即图 1 描述了帧或图象组(*GOP*),图 2 描述了图象的宏观(*MB*)分割,图 3 描述了作为例子的一个帧的分割,图 4 描述了一宏观块的块分割。

图 5 显示了在 *MPEG-1* 标准中采用的一个 *GOP* 中的图象中的两级运动补偿。

图 6 显示如何根据本发明来把宏观块定标在各种标度。请注意也可把宏观块定标到高于该图中所示的完全分辨率的分辨率上。

图 7 显示了对应于图 6 中所示标度的 *DCT* 系数的分级预测。

图 8 是可解译图 6 和 7 的三或四个分辨率标度的解码器的框图。注意如输出只需一个分辨率标度时可略去若干个框。

图 9 是具有比特串可定标性的解码器的框图。

图 10 是带预测前的非量化的分级解码器的框图。

图 11 是可与本发明连用的灵活可定标视频压缩编码器方案的框图。

图 12a 是可用于图 11 的灵活可定标视频压缩编码器方案的变换单元一种型式。

图 12b 是可用于图 11 是灵活可定标视频压缩编码器方案的变换单元的另一种型式。

图 13a 是可用于图 11 的灵活可定标视频压缩编码方案的分级预测单元的一种型式。

图 13b 是可用于图 11 的灵活可定标视频压缩编码器方案的分级预测单元的另一种型式。

在给出本发明具体实施例的描述之前,应解释的是,在涉及数字视频信号的大量应用,对视频信号压缩算法提出了很大相互矛盾的要求,这些矛盾表现在标准的兼容、编码器和解码器方案的复杂性、功能、图象质量等方面。不能指望用一种编码法来满足所有这些相互矛盾的要求,但可要求一种灵活性可用来满足这些相互矛盾的要求中的许多个,同时仍保存不同设计方案之间的广泛兼容性。据信,例

如,虽然不同的应用会要求不同的编码器方案,实现能解译所有符合该灵活设计的所有比特串的单一解码器是有可能的。此外,相信这种通用解码器应不会太复杂。本发明一个目的就是提供这样的通用解码器。

若干种应用将得益于上述的分辨率及比特串可定标性特性。本发明的方法和装置支持用 8×8 DCT 分量分级编码分辨率和比特串定标。虽然也可采用其他大小的 DCT,最佳实施例仍以 8×8 DCT 开始,因为它为几种标准压缩算法所选用。

对分辨率定标,本发明可被用于提供多达四级的分辨率。最低的可能分辨率是通过编码 8×8 DCT 块的左上分量的等值量来达到的,该分辨率是原分辨率的 $1/64$ 。原分辨率的 $1/16$ 的分辨率,可通过编码 DCT 块的左上 2×2 个系数的等值量而获得。编码左上 4×4 个系数的等值量则得到原分辨率 $1/4$ 的分辨率。最后,编码所有 8×8 个系数将给出完全的影象分辨率。

给定一固定的分辨率,则本发明通过以同样的分辨率对多个分级层进行编码,来支持比特串定标。但是,随着量化因子越来越精细,这些层将导致具有相同空间分辨率和改进的质量的影象。在此情况下,分级结构中的第一个和最低的层将得到粗略的量化编码,而较高的层将得到越来越精确的量化编码。

该方案是灵活的:可将一或多个这些分级层按分辨率或精度增加的顺序叠置起来,从而可用在某一级分辨率下重建的系数去预测

下一级分辨率下的相应系数。除了最低的分级层,任何层的变换系数均相对其预测进行差分编码。例如,可选用一编码器来产生只包含完全以及 1/16 分辨率标度的数据的比特串。在这种情况下,1/16 分辨率层的系数将被用来预测完全分辨率下的系数。或者,对比特串定标,某一层中的粗略量化的 8×8 DCT 系数组可被用来预测下一层中的相应 8×8 系数组,其中后者的量化更为精确。

本发明的方案的基本特性,是 MPEG-1 的宏观块(MB)的身份在所有分辨率和精度层上均得到保存。图 6 显示了对四层分辨率编码时的这一特性。保存该身份的重要性在于,如在 MPEG-1 算法时提到的,MB 是关系到一系列特性的,这些特性将决定包含在压缩数据串中的开销数据的量。保存 MB 身份使得能在所有分级层中重复使用这种开销数据。例如,对应于任何分辨率标度的运动矢量数据,可通过适当定标,从最高的分辨率运动矢量数据导出。

标定层的多路化

虽然不是本发明的内容,但也应注意的是,在传送和存贮之前,各标度的数据均要用编码装置作为多路处理。实施这种多路处理可有多种选择。例如,可按分辨率和精度从低到高的顺序,把各级的完整图象连接起来。应理解的是,该数据包含代表与所处理的影象的象元或象素有关的信息。这些信号是用适当电子装置处理的通常的电信号,这些装置可以是摄象机、计算机、以及本领域人员完全熟悉和理解的辅助设备。

分辨率和比特串的标度的量化

在最佳实施例中,各分级层的量化采用 *MPEG-1* 中规定的相同的矩阵。这样,各分辨率标度的 *DCT* 数据均从完全分辨率 8×8 *DCT* 矩阵导出。若 *MPEG-1* 权重量化矩阵用 Q_8 表示,则较低分辨率 *DCT* 的权重矩阵的导出如下:

量子化子	<i>DCT</i>	因子
Q_1	1×1	$1/8Q_8$
Q_2	2×2	$1/4Q_8$
Q_4	4×4	$1/2Q_8$

应注意,本发明也包括在各标度采用类属量化矩阵的情况,但此时解码装置将变得更复杂。

分辨率和比特串标度的分级预测

在最佳实施例中,用分级层中的 *DCT* 系数来预测分级结构中下一层的相应系数。在图 7 中显示了一个例子,其中显示了 4 个分辨率层的分级结构。预测算法在此为适当定标的系数的简单适当定标的系数一对一映射。但是,应注意的是,这里又一次可在复杂性加大的情况下采用其他的预测算法。

对分级层速率控制的规定

在最佳实施例中 *MPEG-1* 的量化参数 *QP* 被用在分级结构的最低层。分级结构其他层的 *QP* 参数均参照此较低层 *QP* 来规定。例如,一高层的 *QP* 参数可被定为较低层 *QP* 的两倍。

可定标解码器实施方案

本发明包括具有可变数目的分级层的方案。然而,为说明方便起见,将描述两个三级解码器,它们体现了分辨率和比特串定标的特性。这种解码器装置,如图 8 所示,支持 2×2 (低)、 4×4 (中)、和 8×8 (高)分辨率标度。只有一个分辨率标度的解码器,可通过消去图 8 中与实现该分辨率标度无关的方框来实现。在对三个分辨率标度的压缩数据作熵解码和信号分离之后,对每一 8×8 块数据都有相应的 2×2 和 4×4 块数据,所有这些都是在建立最终的 8×8 DCT 系数矩阵所必需的。

在最佳实施例中,进行了如下步骤来得到最终的完全分辨率 8×8 DCT 系数。在用 2×2 层原 qp_2 量化参数进行非量化之后,低分辨率的 2×2 块被用作对相应 4×4 块的四个最低级系数的预测。将这些预测值与非量化的 4×4 系数相加,其中 4×4 系数的非量化是由 qp_4 量化系数进行的。前述的求和结果类似地被用作相应 8×8 块的 16 个最低级系数的预测。这些预测值被加到非量化的 8×8 系数上,其中 8×8 系数的非量化是由 qp_8 量化参数进行的。

注意 DCT 系数只是在重建最终的系数矩阵时由量化参数来去量化的。由权重量化矩阵进行的去量化只是在一旦达到最终分辨率时才需要。这种特征之所以可能,是因为在最佳实施例中,各标度下的权重矩阵是比例相关的,如上面所述。

现在可利用 MPEG-1 技术,包括运动补偿预测,用最后的 $8 \times$

8 系数矩阵来重建完全分辨率图象。参见图 8, 可看出在完全分辨率水平下, 16×16 MCP 单元表示在 MB 上运行的类属 MPEG-1 运动补偿预测单元; IDCT 框是执行标准 MPEG-1 8×8 逆变换原单元; 而 $Q8^{-1}$ 框表示用相应 MPEG-1 标重矩阵进行反量化的单元。

为在其他分辨率标度下进行视频图象解码, 要进行类似的过程, 只是把预测值与非量化系数数据相加的步骤终止于需要解码的分辨率上。用权重矩阵(用 $Q4^{-1}$ 和 $Q2^{-1}$ 表示)作非量化的操作, 采用了最佳实施例的定标矩阵。逆变换操作是用适当大小的变换进行的。这样, 要在 $1/4$ 分辨率下解码, 就要用 4×4 的逆 DCT。适于在所支持的三个“定标的”分辨率下解码的一维 DCT 矩阵是:

$$DCT(1 \times 1) = 1$$

$$DCT(2 \times 2) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$DCT(4 \times 4) = \frac{1}{2} \begin{bmatrix} 1.00 & 1.31 & 1.00 & 0.54 \\ 1.00 & 0.54 & -1.00 & -1.31 \\ 1.00 & -0.54 & -1.00 & 1.31 \\ 1.00 & -1.31 & 1.00 & -0.54 \end{bmatrix}$$

$DCT(1 \times 1)$ 和 $DCT(2 \times 2)$ 都是不难的, 并易于实施, 即使是在软件中也是如此。

应注意, 其他分辨率下的 MCP 单元具有相同的运动矢量数

据, MV , 以在不同分辨率标度下产生用于标定的 MB 的运动补偿预测。应特别注意的是, 当采用运动补偿技术时, 应适当标定完全分辨率矢量, 以与解码器分辨率相匹配。这已在前面描述过。

在图 9 中显示了实施比特串定标的解码器。该解码器的运行与图 8 中的很相似, 不同之处是只采用了 8×8 操作, 以产生质量更高的输出图象。因此, 可用单一硬件来在实体上构成不同的单元, 而不同的分级层可共用这些单元。

最后, 在图 10 中显示了另一种解码器实施例, 它不是本发明的最佳实施例, 但仍属于本发明。该解码器的运行方式与图 8 的相同。然而, 它允许在各分级层采用任何的量化权重矩阵。它还构成从较低分级层预测 DCT 系数的类属装置。在此方面, P_2 被用来从 2×2 系数预测 4×4 系数, 而 P_4 被用来从 4×4 系数预测 8×8 。可采用的预测算法有多种, 包括最佳实施例中的一对一映射。

编码器实施方案

与上述本发明解码器兼容的编码器的实施方案可有许多种, 但作为例子, 这里将只描述两种这类编码器, 每种都是只为分辨率定标而设计的。在图 11 中显示了三分辨率层编码器的一般结构。该编码器分为三部分。第一部分为一变换单元, 它接收数字视频信号输入并输出用于三个分辨率层的 DCT 数据: $d(8 \times 8)$ 、 $d(4 \times 4)$ 、和 $d(2 \times 2)$ 。第二部分是一分级预测单元, 它接收变换单元的 DCT 输出, 并输出所有分辨率层下的量化差分 DCT 数据: $q(8 \times 8)$ 、 $q(4 \times 4)$ 。该

输出在第三部分中被多路化并得到熵编码,以产生最终的压缩视频信号。该预测单元还产生重建的 DCT 数据: $b(8 \times 8)$ 、 $b(4 \times 4)$ 、和 $b(2 \times 2)$,该数据被反馈到变换单元,以完成通常的混合变换编码回路。

图 12a 显示了变换单元的一个简单实施方案。在此实施方案中, 8×8 层包含也是 $MPEG-1$ 编码器一部分并且更一般地是运动补偿混合变换编码器的一部分的元素。上支路包括:产生运动补偿预测差的加法器(Σ);一向前的 $8 \times 8 DCT$ 变换,以产生所述预测差的 8×8 变换系数;用于由权重矩阵($Q8$)进行量化的单元。输出是一组部分量化的 DCT 系数 $d(8 \times 8)$ 。返回或反馈支路接收一组部分重组 DCT 系数 $b(8 \times 8)$,并随后用下列单元对其进行处理:一用于权重矩阵($Q8^{-1}$)的去量化器和重建预测差的逆 DCT 变换器($IDCT8 \times 8$);一加法器,用于把重建的预测差与运动补偿预测值相加,从而重建最初的图象数据;用于产生下一图象的预测的运动补偿预测(MCP)单元。在 $MPEG-1$ 中,该 MCP 单元在 $16 \times 16 MB$ 上运动(如图所示),但也可采用其他尺寸的块。在此变换单元实施方案中, $d(2 \times 2)$ 和 $d(4 \times 4)$ 系数是从相应的 $d(8 \times 8)$ 系数中简单地摘取出来。应注意的是,也可通过把其他处理或加权算法作用到 8×8 系数上来导出 $d(2 \times 2)$ 和 $d(4 \times 4)$ 系数。

注意由于在较低分辨率标度下没有反馈环路,在这些分辨率标度此编码器将产生量化和运动补偿误差的累积。但这种误差在开始

新的图象组时将被自动复位到零。虽然这种误差累积限制了较低分辨率层的质量,但编码器的简化使这一途径仍具吸引力。特别地,如果所需的仅是比特串定标,则这一途径就可满足全部需要。

图 12b 中显示了变换单元的另一实施方案。此方案与图 12a 中所示的类似,但 $d(2 \times 2)$ 和 $d(4 \times 4)$ 是由类似但完全独立并在各自分辨率标度运行的环路产生的。在此情况下 H_4 和 H_2 单元被用来在每种情况下把输入视频信号的分辨率滤波并降低 $1/4$ 。以此方式,各层均接收具有近似分辨率的输入。所有操作,如 DCT、量化、和 MCP,均按该层的分辨率来定标。

在复杂性增大的情况下,此方案通常比图 12a 的单元能产生质量更好的图象。在此情况下,编码误差的积累将不超过一个图象周期。但应注意, 16×16 运动估计的结果可以是所有分辨率环路所共有的,因为在本发明中,运动矢量所有标度下的 MB 所共有的一个特性。此实施方案更适合于那些低分辨率图象的质量较为重要的应用。

图 13a 显示了分级预测单元的一个实施方案。首先,在加法器中产生 4×4 和 8×8 层的分级预测差。随后,所有层均由各自的量化参数量化,其结果作为 $q(2 \times 2)$ 、 $q(4 \times 4)$ 、及 $q(8 \times 8)$ 输出。这些结果还由相应的量化参数进行反量化并随后在其他两个加法器中被加起来,以产生部分重建的 $d(2 \times 2)$ 、 $b(4 \times 4)$ 、及 $b(8 \times 8)$ 数据,这些数据被馈回到变换单元。图 13b 显示了可用作分级预测单元另一种实

施方案的同样元件的重新设置。

可以理解,本领域原技术人员可设计出其他符合本发明解码方法和方案的编码方案。

说明书附图

图1

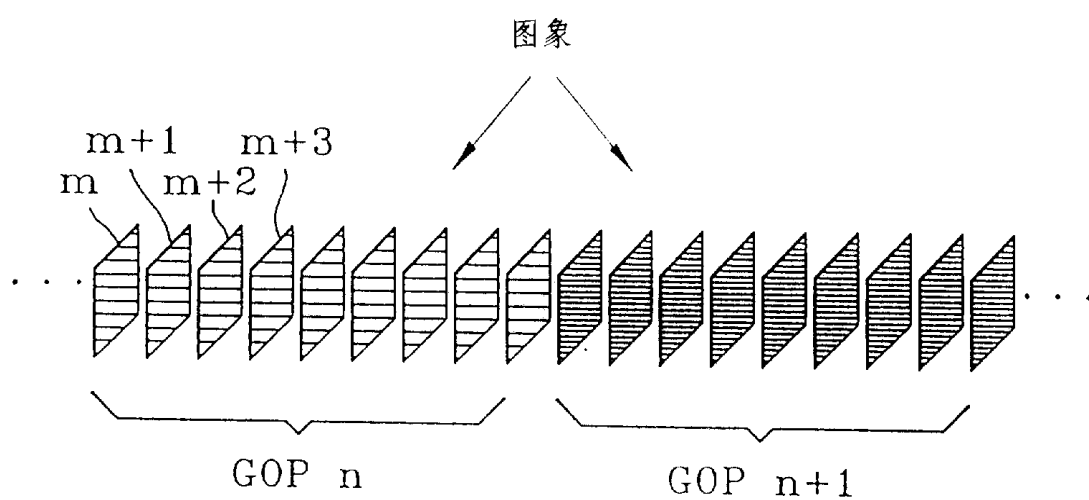


图2

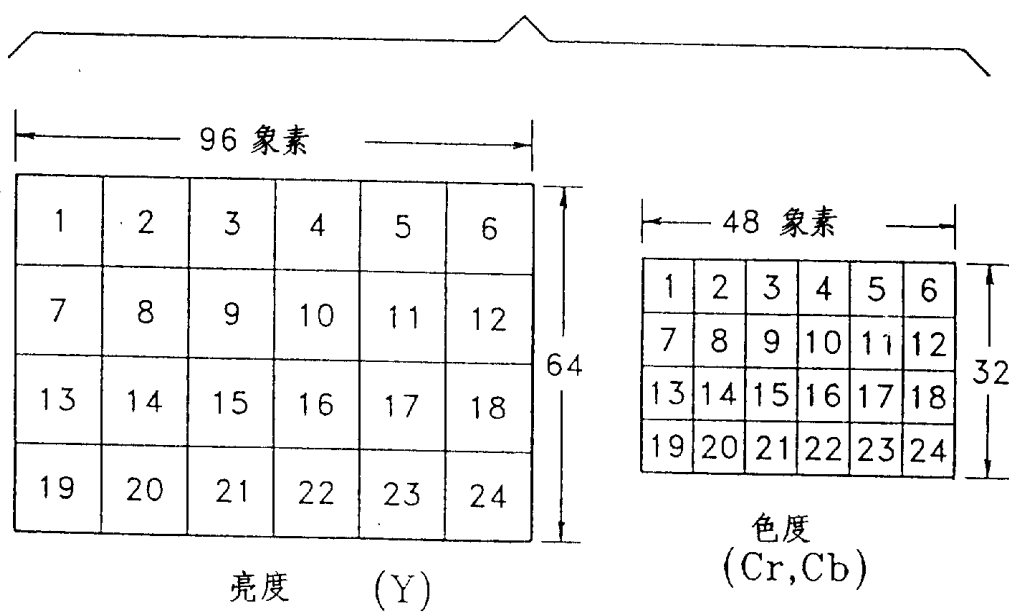


图3

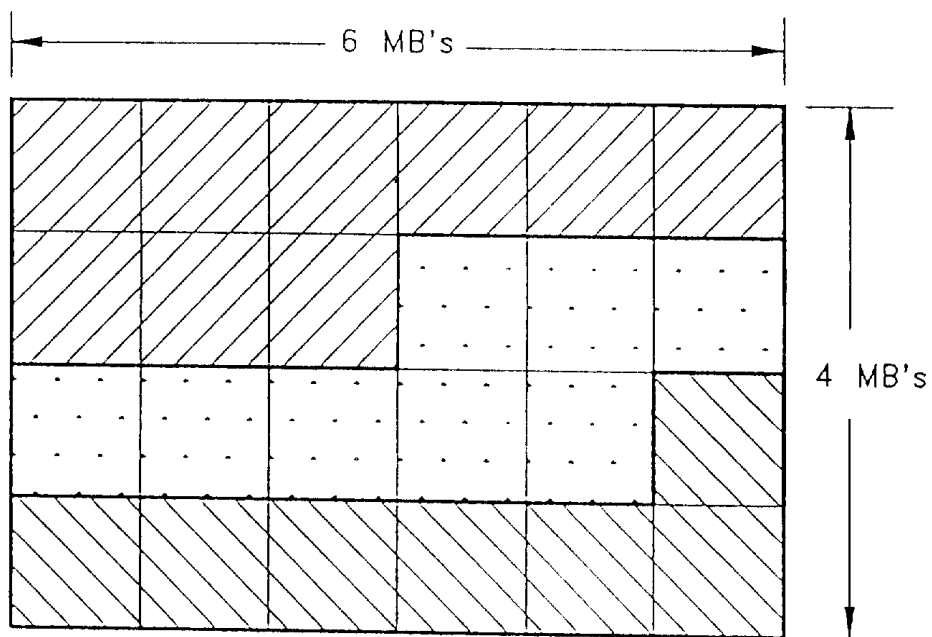
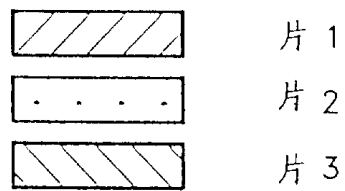


图4

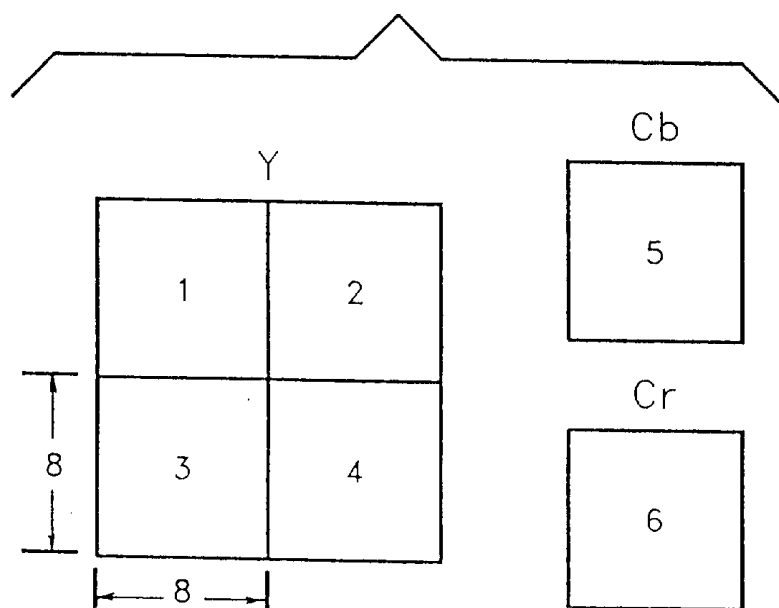
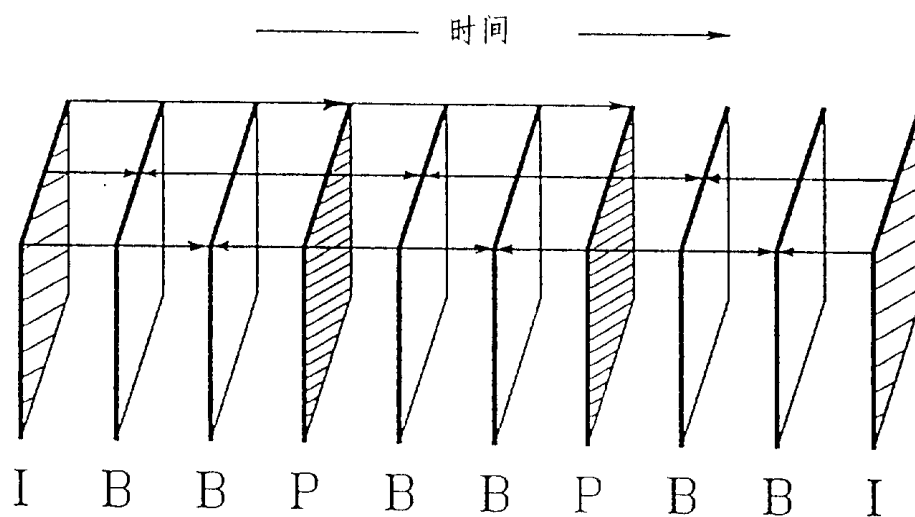


图5



$$N=9 \quad M=3$$

N = I 帧的间距

M = P 帧的间隔

图 6

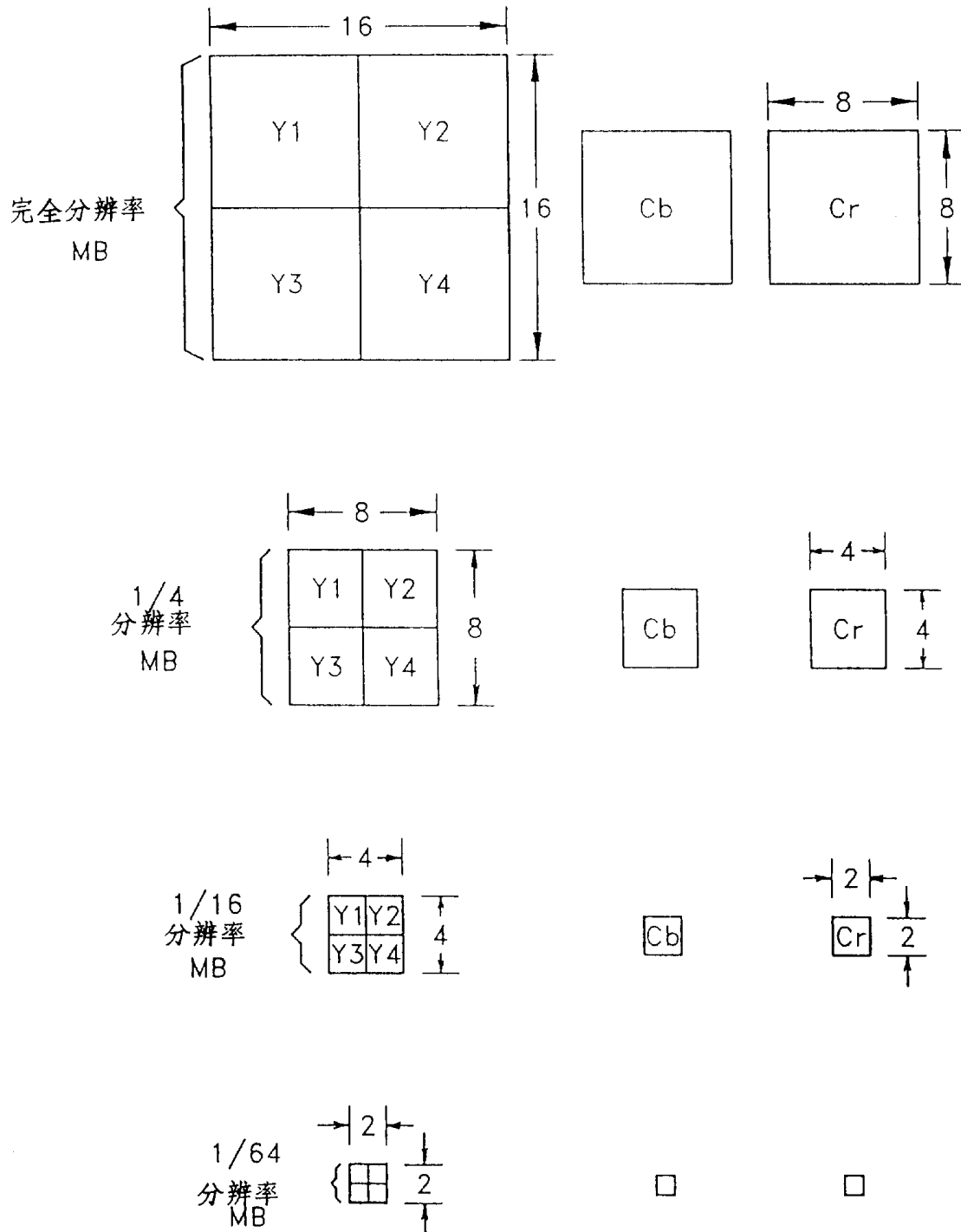


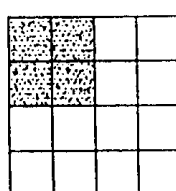
图7

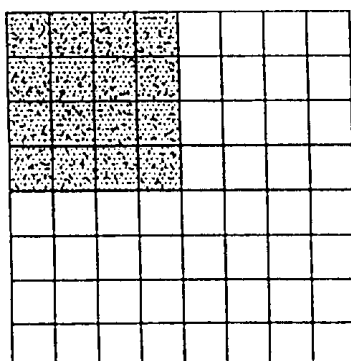
C_1  } 1/64 分辨率

 预测 DCT 系数

 非预测 DCT 系数

C_2  } 1/16 分辨率

C_4  } 1/4 分辨率

C_8  } 完全分辨率

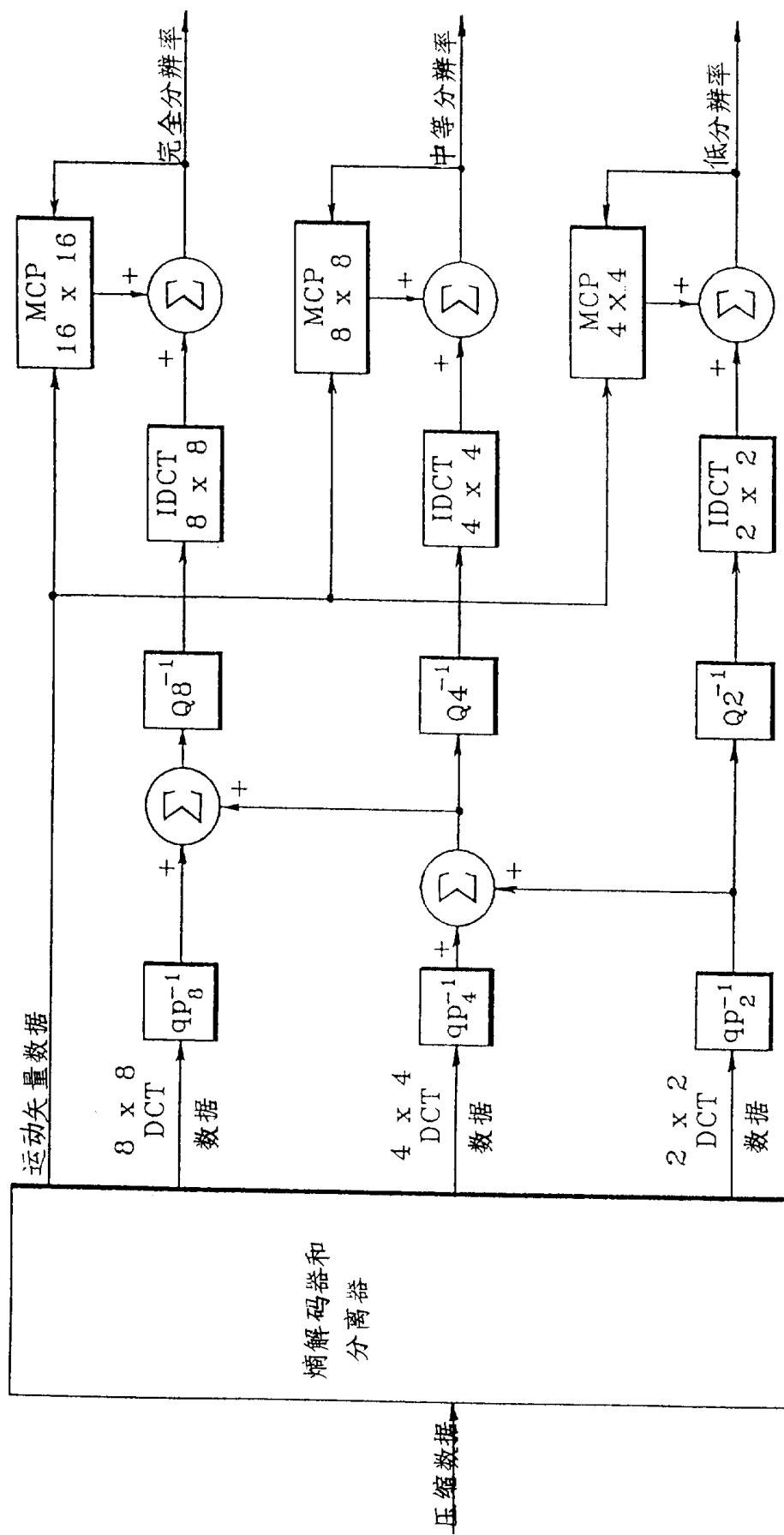


图8

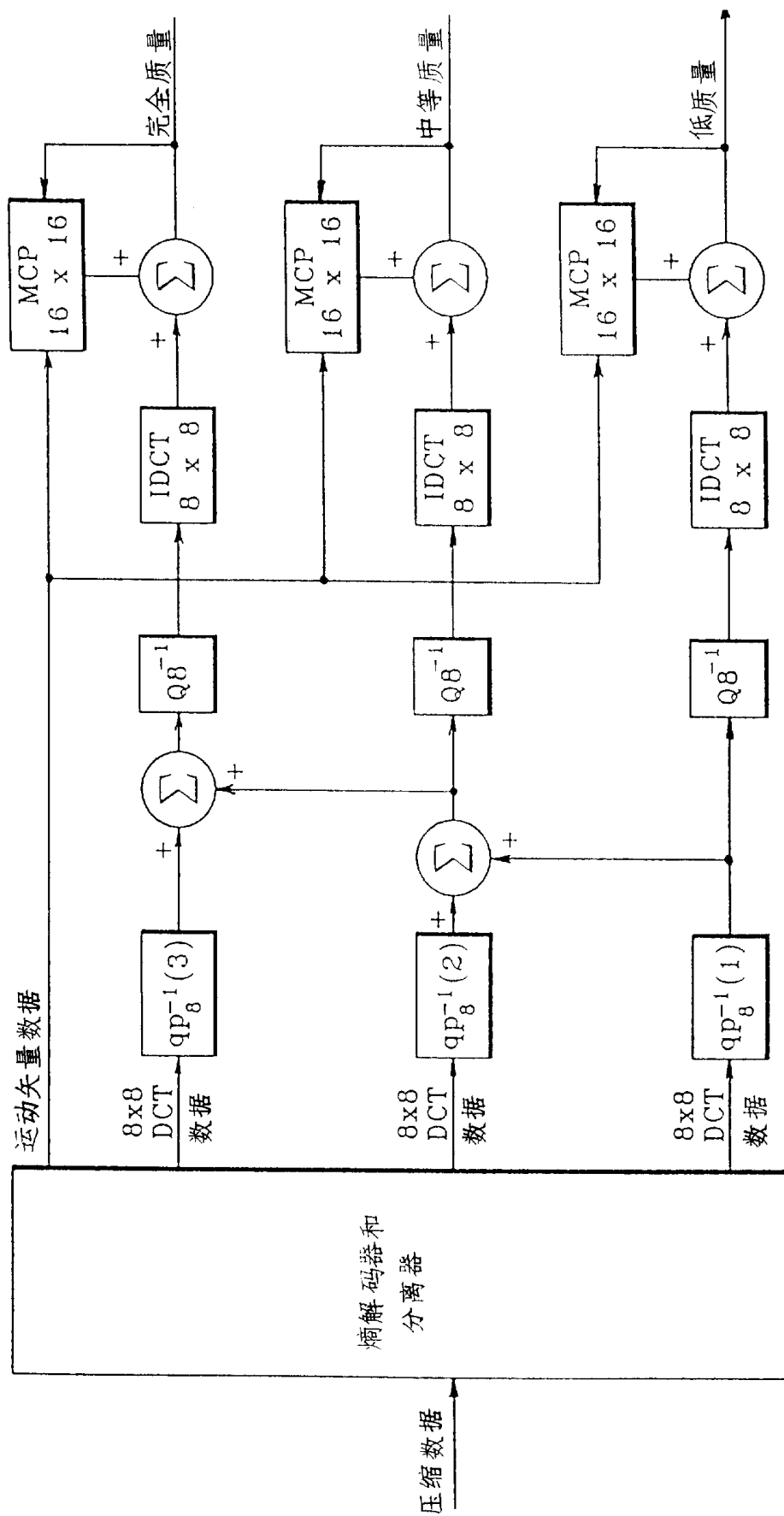


图9

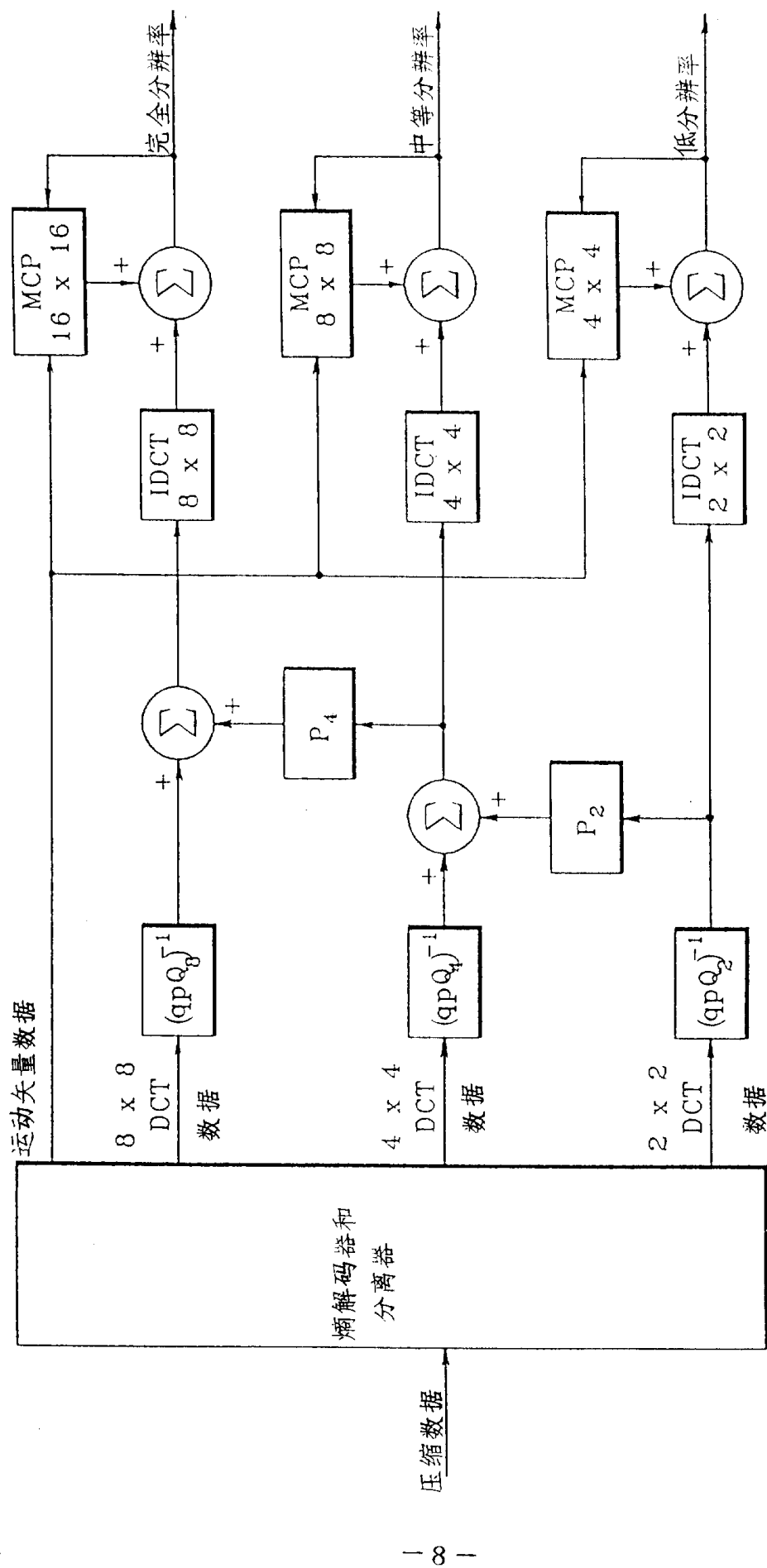


图10

图11

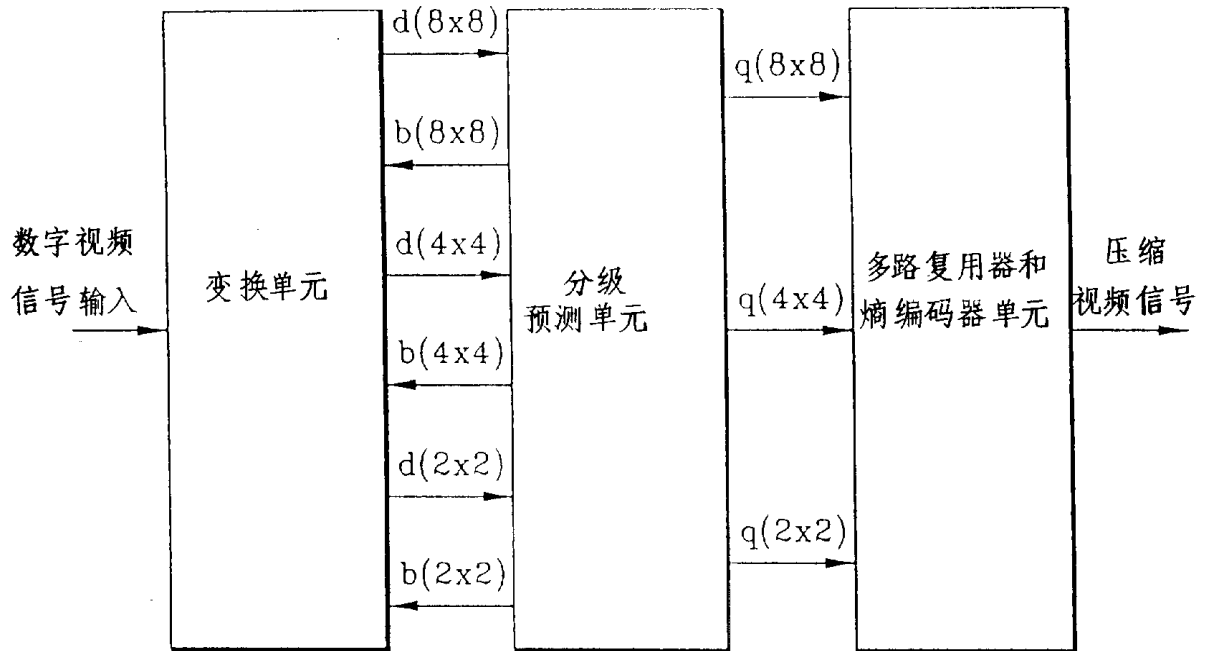
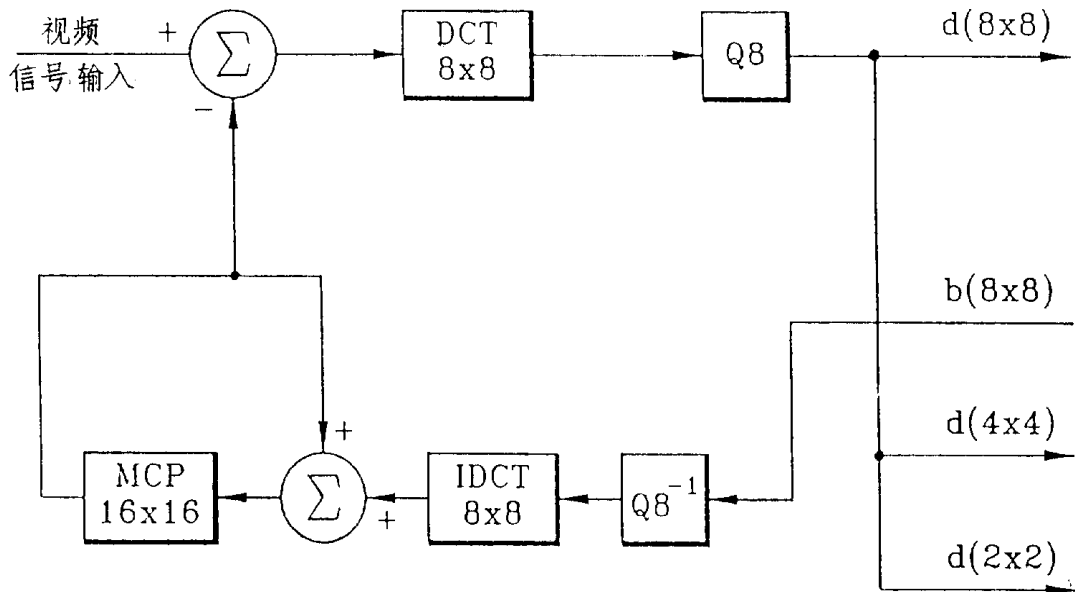


FIG. 12a



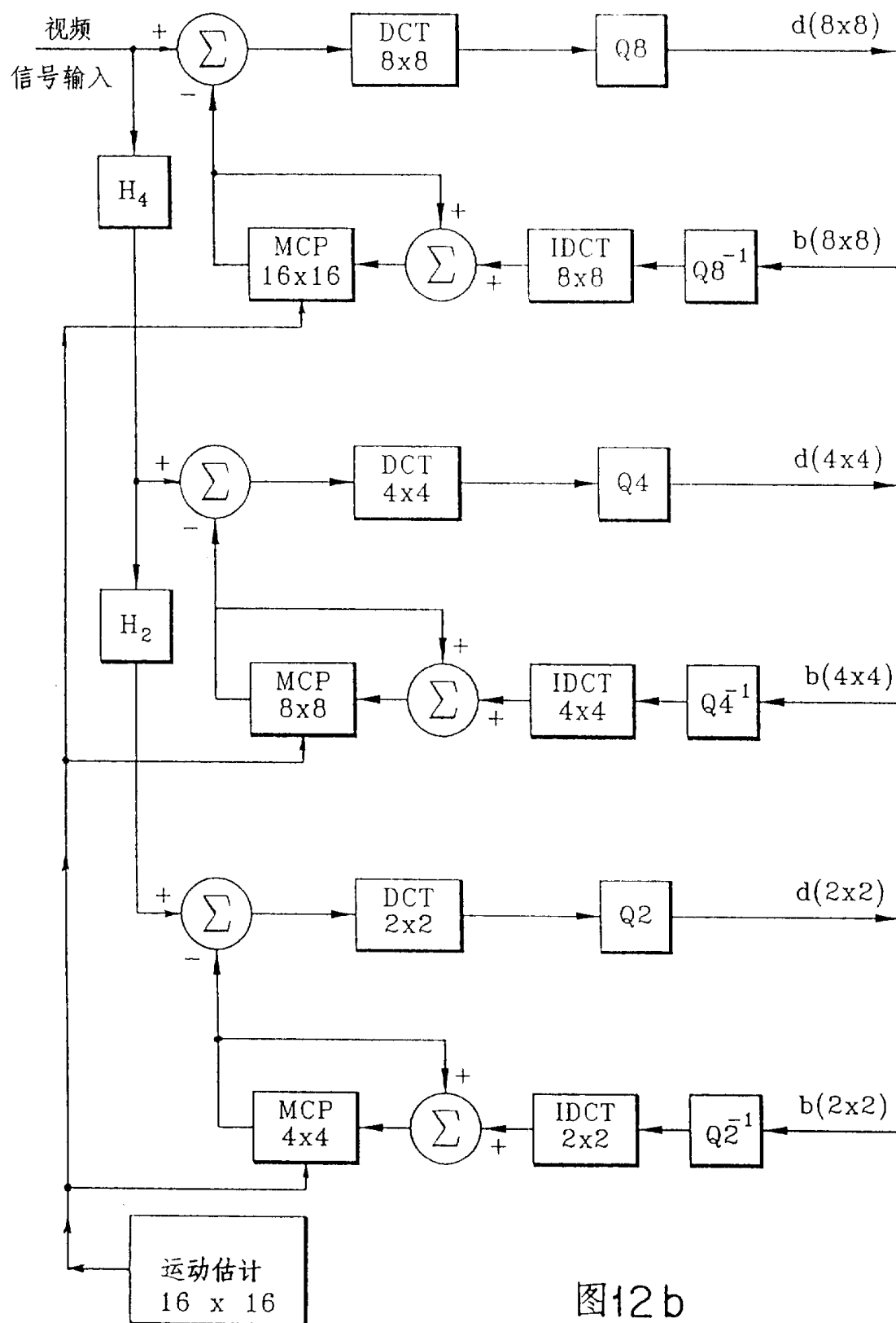


图12 b

图13a

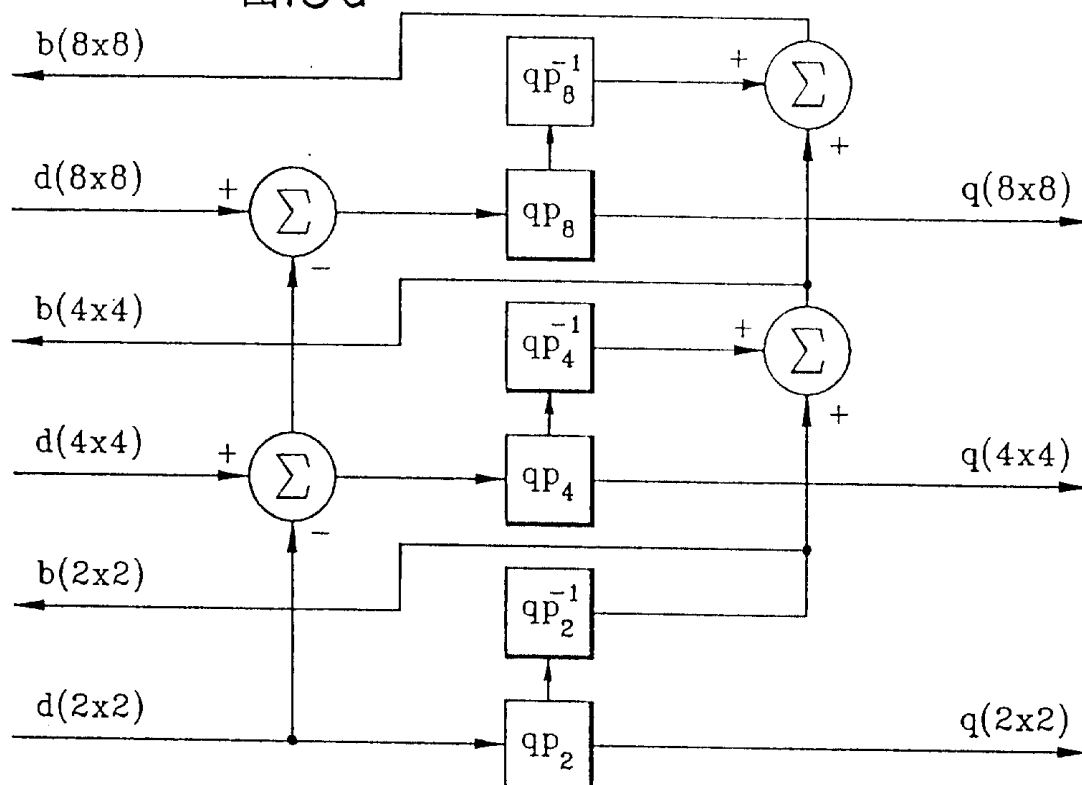


图13b

