

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

H04N 7/50

H04N 7/30



[12] 发明专利说明书

[21] ZL 专利号 98801430.0

[45] 授权公告日 2004 年 1 月 28 日

[11] 授权公告号 CN 1136734C

[22] 申请日 1998.7.6 [21] 申请号 98801430.0

[30] 优先权

[32] 1997. 7. 29 [33] EP [31] 97401823.6

[32] 1998. 5. 26 [33] EP [31] 98401247.6

[86] 国际申请 PCT/IB98/01031 1998.7.6

[87] 国际公布 WO99/07159 英 1999.2.11

[85] 进入国家阶段日期 1999.5.27

[71] 专利权人 皇家飞利浦电子有限公司

地址 荷兰艾恩德霍芬

[72] 发明人 E·菲尔特 D·凯瑟

审查员 梁军丽

[74] 专利代理机构 中国专利代理(香港)有限公司

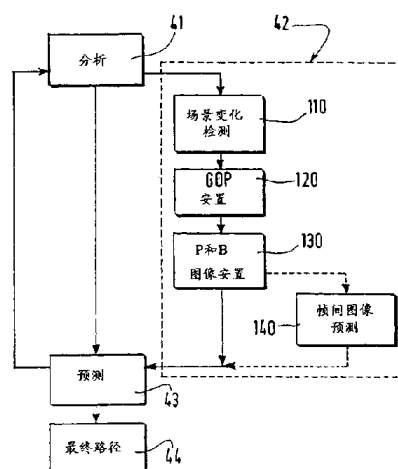
代理人 王勇 王岳

权利要求书 2 页 说明书 14 页 附图 6 页

[54] 发明名称 可变比特率视频编码方法及相应的视频编码器

[57] 摘要

按照本发明的一种可变比特率视频编码方法，包括一个反复的过程，该过程包括：第一的分析步骤，第二的预测步骤，以及最后的控制步骤，调整关于上述目标比特率的步长；按照本发明所述方法，在一次反复的分析和预测步骤之间，提供一个图象重新排列的步骤。应用：具有有限容量的存储介质的 MPEG-2 编码器。



1. 一种可变比特率视频编码方法，包括一个迭代过程，该过程包括：第一的分析步骤，利用一个不变的量化步长对与一个图象序列相对应的比特流编码；第二的预测步骤，根据一个预先确定的目标比特率预测将用来对所述比特流编码的量化步长；最后
5 的控制步骤，调整关于上述目标比特率的步长；所述方法其特征在于：在一次反复的分析和预测步骤之间，它包含一个图象重新排列的步骤；所述图象重新排列步骤顺序地包含第一的场景变化检测子步骤，第二的分配子步骤，第三的最佳放置子步骤。
- 10 2. 根据权利要求 1 的方法，其特征在于：所述场景变化检测子步骤包含一个在序列的两个相继图象中执行的相关性操作，以及用于显示一可能的场景变化出现的决策操作。
3. 根据权利要求 2 的方法，其特征在于：所说的相关性操作是基于图象复杂性估算。
- 15 4. 根据权利要求 2 的方法，其特征在于：所说的相关性操作是基于图象块与先前参考图象的参考块的比较。
5. 根据权利要求 1 至 4 中任何一个的方法，其特征在于：所说最佳放置子步骤包括：第一长期优化操作，用于在几个图象组之上寻找这些图象组之间的最佳距离；第二短期优化操作，以在几个图象组
20 中寻找预测的且内插的图象的最佳位置。
6. 根据权利要求 5 的方法，其特征在于：所述分配子步骤是基于 Viterbi 算法的执行，该算法允许为图象的持续组选择一个最佳的尺寸，同时将这样一组图象的开始与一个场景变化相匹配。
7. 根据权利要求 5 的方法，其特征在于：在两个连续迭代之间
25 的图象类型的顺序发生变化的情况下提供附加的帧内图象预测步骤。
8. 一种可变比特率视频编码器，包括第一编码分支、预测分支、及一个控制电路，所述第一编码分支至少包括一个离散余弦变换电路，量化电路，可变长度编码电路和输出缓冲器，所述预测分支包括
30 反向量化电路，反向离散余弦变换电路，加法器，缓冲器，运动估算电路，运动补偿电路和减法器，所述控制电路包括一个控制环，所述控制环包括：

-
- 第一计算电路，其中所述控制环的输出乘以给定因数；
第二计算电路，其中产生对量化步长的综合估算；
将所述第一和第二计算电路的输出相加的加法器；
变换电路，和所述加法器串联，从而为所有在前的图象给出积累
- 5 比特率；
- 比较器，用于将所述积累比特率和所述积累预测比特率比较；
积分器，积分器的输出被利用于控制量化步长值；
所述的编码器还包括在编码分支和预测分支之间的图象重新排列装置，该图象重新排列装置包括场景变化检测电路、分配电路和最佳
- 10 放置电路。

可变比特率视频编码方法 及相应的视频编码器

5 本发明涉及一个可变比特率视频编码方法，其中包括一个反复的过程：该过程包括第一的分析步骤，利用一个不变的量化步长对相应于一个图象序列的比特流编码；第二的预测步骤，根据一个预先确定的目标比特率预测将用来为上述比特流编码的量化步长；最后的控制步骤，用于调整与上述目标比特率相关的步长。本发明还涉及可以执行上述方法的一个相应的视频编码器。

正如《MPEG 视频编码：基本指南介绍》（S.R. Ely, BBC 研究与发展报告，BBC-RD-1996/3, 1-10 页）所述，MPEG 行为始于 1988 年，目标是为视频及音频信号的数字化压缩确定标准。第一个目标是为数字存储媒体定义一个视频运算法则，例如 CD-ROM（高密盘只读存储器），但是所得标准也应用于交互式 CD 系统（CD-I）。此标准允许以 1 至 15 兆比特/秒的比特率传送并存储图象数据，且此标准基于这样一个数据压缩，这个数据压缩是通过使用用于减少时间冗余的基于块的运动补偿及一个用于减少空间冗余的离散余弦转化（DCT）而获得的。

20 在常规 CD 标准中，例如 CD-I 和 CD-ROM，传送比特率是固定的并且因此仅可以对图象以一个不变的比特率编码。新标准例如数字化多功能碟片（DVD）允许以可变比特率（VBR）传送数据：为保持一个不变的品质，与对包含较少信息场景的编码相比，可以用一个较高的比特率对复杂的场景编码。

25 本发明的一个目的是提出一个 VBR 视频编码方法，采用这种方法可利用编码器输出比特流的最小比特率来获得编码序列的不变品质。

对此，本发明涉及一种可变比特率视频编码方法，包括一个迭代过程，该过程包括：第一的分析步骤，利用一个不变的量化步长对与一个图象序列相对应的比特流编码；第二的预测步骤，根据一个预先确定的目标比特率预测将用来对所述比特流编码的量化步长；最后的控制步骤，调整关于上述目标比特率的步长；所述方法其特征在于：

在一次反复的分析和预测步骤之间,它包含一个图象重新排列的步骤;所述图象重新排列步骤顺序地包含第一的场景变化检测子步骤,第二的分配子步骤,第三的最佳放置子步骤。

- 在最佳实施例中,所述场景变化检测子步骤包含一个关连操作,
5 此操作在序列的连续图象之间执行和决策操作,以显示一个可能的场景变化的出现。

本发明的另一个目的是提出一个可执行所述编码方法的 VBR 视频编码器。

- 对此,本发明涉及一种可变比特率视频编码器,包括第一编码分
10 支、预测分支、及一个控制电路,所述第一编码分支至少包括一个离散余弦变换电路,量化电路,可变长度编码电路和输出缓冲器,所述预测分支包括反向量化电路,反向离散余弦变换电路,加法器,缓冲器,运动估算电路,运动补偿电路和减法器,所述控制电路包括一个控制环,所述控制环包括:

- 15 第一计算电路,其中所述控制环的输出乘以给定因数;
第二计算电路,其中产生对量化步长的综合估算;
将所述第一和第二计算电路的输出相加的加法器;
变换电路,和所述加法器串联,从而为所有在前的图象给出积累
比特率;

- 20 比较器,用于将所述积累比特率和所述积累预测比特率比较;
积分器,积分器的输出被利用于控制量化步长值;

所述的编码器还包括在编码分支和预测分支之间的图象重新排列装置,该图象重新排列装置包括场景变化检测电路、分配电路和最佳放置电路。

- 25 结合以下的描述及附图,现将本发明的优点详细解释如下:

图 1 显示了一组图象中的一些图象

图 2 和图 3 图示了此典型图象组在图象的放置顺序和它的传送顺序之间的区别;

- 图 4 和图 5 分别显示了根据本发明编码方法的主要步骤及可以
30 执行所述方法的视频编码器的相应结构;

图 6 显示了在两个连续的 B 图象之间场景变化的序列中一些图象及相连的运动矢量。

图 7 和图 8 展示了两个检测系数 DCR 和 DCL 的值,对其计算是为了定量估计运动矢量统计数, B 图象的宏块用此统计数确定关于 IBBP 和 PBBP 图象组的两个 B 图象的场景删节的三个可能位置;

图 9 展示了 P 图象的一个相似检测系数 DCP。

5 图 10 图示了仅使用 B 图象的场景变化检测方法的性能。

图 11 给出了用来解决图象组(或 GOPs)的最佳安置问题的 Viterbi 算法的决策值。

图 12 图示了 GOP 大小的加权函数(加权值被用来评价 GOPs 的大小)的例子。

10 图 13 展示了一个控制回路,它用来执行所述编码方法的最终控制步骤。

在描述按照本发明的编码方法之前,我们先回忆一下 MPEG-2 标准的一些基本原则。为了支持在广泛领域内的应用,这一标准具有一定弹性。这种弹性的取得是通过对合适于所有应用要求的轮廓及层次加以定义。一个轮廓是 MPEG-2 标准的一个子集,仅用来支持一个给定的应用类别所需的特征,而一个层次定义对比特流参数施加的一套限制。

应用于包含三个分量(Y, U, V)的彩色图像的 MPEG-2 压缩的基本步骤涉及一些被分为小部分和宏块的图像,其中包括亮度和色度块。这些步骤是:运动估算及补偿(基于 16 象素×16 行宏块),离散余弦变换(基于 8 象素×8 行块),步长编码。

定义了三种图象。对帧内图象(I 图象)进行了编码且不涉及其他图象。使用从一个过去 I 或 P 图象运动补偿预测对预测图象(或 P 图象)编码,并且双向预测图象(或 B 图象)利用过去和未来的 I 或 P 图象作运动补偿对。运动信息以运动矢量的形式给出,此运动矢量是通过执行块与块的匹配搜寻得来的(其中大量实验偏差在编码器中被检验,基于在正被编码的块与预测之间最小误差的测量选出最佳者。)

30 图 1 展示了在运动矢量的基础上 P 和 B 图象是怎样被定义的,这些不同的图象通常是以重复序列的形式出现的,如上所述,这些不同图象被称为图象组或 GOP,且包含一个 I 图象及下一个 I 图象出现前的所有后序图象。一个典型的 GOP 在图 2 中以陈列方式展示(黑箭头

与前向预测对应，白箭头与后向预测对应，序列及预测周期性重复），在图 3 中以传送顺序（P4，P7，I10，I13 指明重新排列的帧），所述顺序是不同的，以便在解码侧实现根据将来图象作后向预测。

- 5 一个规则的 GOP 结构可用两个参数描述，N 和 M。被定义为 GOP 大小的参数 N（见图 2）是所述 GOP 图象的数量，即两个 I 图象之间的图象数加一。参数 M 是 P 图象的间距，或（与之相同）相邻 B 图象数加一。在图 1 至图 3 所示的例子中，M=3，N=9。很明显其他组合可能是：

图象显示次序	N	M
IPPPPPIPP	6	1
IBPBPBPBI	8	2
IBBPBBPBBPBBIBBP	12	3

N 和 M 互相独立地选择。

- 10 决定了采取宏块补偿的种类后，通过从原来的宏块中减去被估算的宏块获得此宏块的每一个象素的预测误差。而后为一个 8×8 象素的块，对预测误差进行一个 DCT 操作（因此每一个宏块有六个 DCT 变换：亮度成份有四个，色度成份有两个），由此获得的频率成份被量化。量化步长决定了被解码图像的比特率和失真：如果量化是粗糙的，为图象编码只须很少的比特，但最终的质量很低，而如果量化大小精细，为图象编码须很多的比特，但最终的质量很高。由于人眼对高频不象对低频那样敏感，对高频成份使用较粗糙量化具有优越性（实际上，为了获得取决于频率的量化，一个加权矩阵将应用于一个基本宏块量化参数：在所述加权量化后许多特别是处于高频的系数等于零。）
- 15 20

然后每一个块被锯齿地扫描，对所得的表编码。通过决定一组（A，NZ）进行步长编码，其中 A 代表连续零的数量（0 至 63），NZ 代表以下非零系数的振幅。

- 25 按照此组（A，NZ）的出现频率，此组被赋予一个可变长度编码，（一个普通的组合（A，NZ）将被赋予一个短的可变长度编码，而不时常发生的一组被赋予一个长的可变长度编码。）

以上所述功能在以下文献中被归纳《普通视频应用的混合扩展 MPEG 视频编码算法》，C. T. Chen 等，信号处理：图像通信 5(1993)，21

至 37 页, 2.4 部分, 其中包括一个被普通化了的 MPEG-2 编码器的方案。现提出的 VBR 编码方法的目标是使用从几个连续的分析路径中执行的先前编码步骤中获得的信息, 执行图象类型的适应分配, 它可以使最终比特流的大小最小化, 以使一个固定的容量恰好符合存储媒介 (例如一个 DVD)。图 4 给出了所述方法的图表, 图 5 给出了带有可执行所述方法的图象重组的视频编码器的结构。

编码方法分为四步骤, 41, 42, 43 和 44。第一步骤 41 是一个分析步骤, 其中用一个不变的量化步长 Q 为一个图象序列编码 (所以有一个不变的品质)。这一步骤的末尾产生一个规则的 MPEG-2 适应的比特流, 但在所述步骤结束前不知道, 这样被处理 (即该序列中被编码的比特总数除以那一序列中图象总数) 的全部序列的平均比特率不能满足对一个比特流特定大小所须的限制。

第二步 42 是图象重排步骤, 在分析步骤 41 后执行。这一重新排列的步骤可分为三个子步骤 110 至 130。图象分配最优化的任务可分为两部分。第一步是改进 I 图象的放置, 它等同于 GOP 分配的最优化, 而第二步是 B 与 P 图象的最有效放置。

第一与第二子步骤 110 和 120 构成所述优化任务的第一部分。很清楚不使用一个序列连续图象之间的时间相关性的 I 图象耗费的比特率最大。另外, 他们必须允许随机访问序列, 且随机访问对许多应用都很重要。而且, 由于通常所须的是快速随机访问, 必须注意不可超过 I 图象间的给定最大距离 (例如最多 12 个图象)。当一个场景变化发生时该剪接处前后的图象 (左右图象) 是不相关连的。新场景中的第一 P 图象的运动补偿未被很好的执行, 所以它的比特率大约是一个 I 图象。这样一个 I 图象可以取代所述 P 图象在不需额外带宽的情况下被放置了。然后 I 图象优化放置的策略必须在可能的时候分配 I 图象在新场景的开始。

第一子步骤 110 是一个场景变化检测子步骤, 可允许这种分配。为了检测场景变化, 将检测序列中后续图象的相关性 (最好在运动补偿之后): 如果两个相邻图象几近无联系, 很可能新图象从第二个开始。在基础的 MPEG-2 编码过程中, 几个参数给出了一些关于连续图象间相关性的信息:

- P 或 B 图象的复杂性: 然而复杂性与关连性间的连接并不总被验证

- （低比特率有时是由于与参照图象的高度相关性或由于低帧内复杂性，例如全黑图象）；
- 通过比较图象的宏块 MB 和其参考宏块（总是由运动补偿单元提供，无论块编码的类型是帧内的还是帧间的）获得了相关性的更好估算值
- 5 （虽然要花费更多的计算时间）：利用平方误差失真 $d(MB)$ 的计算进行比较，例如可以由以下等式（1）给出：

$$d(MB) = \frac{1}{255} \sum_{i=0}^{i=255} (P(i) - M(i))^2 \tag{1}$$

- 其中 $P(i)$ 是被分析的宏块 MB 的像素， $M(i)$ 是参考宏块的像素；
- 10 运动估算是宏块导向的，对一个图象的预测使用几个运动补偿选择：

图象类型	运动补偿选择
I	帧内（既无运动补偿）
P	帧内
P	向前的
P	不补偿
B	帧内
B	向前的
B	向后的
B	内插的

- 运动补偿统计数字可以传递关于图象相关性的信息：如果宏块
- 15 是内部编码的，则与参考图象的相关性是低的，反之亦然。
- 所执行的实施例用所述运动补偿来检测场景的变化，以仅仅使用 B 图象的情况为例，图 6 展示带有场景变化的序列 PBBP 的图象及运动矢量，此场景变化在两个 B 图象（断箭表明较少的相关参考图象的宏块被用于预测相关图象，相关性是较低的）之间举例说明。当场景剪
- 20 接处在两个 B 图象间出现时，第一个几乎仅用先前 P 图象作为参考图象，因为它基本上与以后的 P 图象不相关。同样的，第二个 B 图象几乎与先前 P 图象无关，且仅以后面的 P 图象作为参照。

在一个三个图象的组合中（例如举例说明的（PBB 或 IBB）），一个场景变化可被放在相邻 B 图象之前、之间、或之后。下表给出场景剪接处的三个可能位置的运动补偿。第一 B 图象的大多数宏块及第二图象的大多数宏块使用此运动补偿：

位置	第一图象	第二图象
之前	向后	向后
之间	向前	向后
之后	向前	向前

5

（从 B 图象的角度看运动补偿的方向）

为了定量评价上述的运动矢量统计数，将计算两个检测系数 DCL 和 DCR（DC 为检测系数，L 和 R 为“左”和“右”，MC 为“运动补偿”）：

$$DCL = \frac{(\text{帧内}) + (\text{向后}MC)}{(\text{向前}) + (\text{内插}MC)} \quad (2)$$

10

$$DCR = \frac{(\text{帧内}) + (\text{向前}MC)}{(\text{向后}) + (\text{内插}MC)} \quad (3)$$

对于 P 图象，检测系数可以类似地定义为：

$$DCP = \frac{(\text{帧内}MC)}{(\text{向前}MC) + (\text{非}MC)} \quad (4)$$

15

图 7、图 8 表明了连续 B 图象的左右检测系数 DCR 及 DCL 的值，场景剪接处清楚地与尖峰对应。同样地，P 图象的检测系数 DCP 如图 9 所示（很明显，P 图象的运动矢量提供的信息不如 B 图象提供的可靠。）

20 例如在两个 B 图象之后出现的场景变化的情况下，即在第二 B 图象与后面参考图象之间（在所设第二 B 图象右侧），分析后的 B 图象的少数宏块向后被补偿或以内插值替换，因为所述 B 图象和后续参考图象之间的相关性低并且他们中的大多数是被帧内的或向前的运动补偿：所以检测系数 DCR 的值很高，而检测系数 DCL 的值并没有被增

加（相反地，在第一 B 图象左侧的场景变化的情况中，在先前参考图象与所述 B 图象之间，DCL 有一个很高的值，DCR 维持小值，而如果在 M 图象的块中没有场景变化出现，DCL 及 DCR 均有一个小的值）。

实际上，为了获得一个场景变化的单一的、对称的指示，将对两个检测

5 测值之间的差分 DDV 进行计算，得到：

$$DDV = DCL - DCR \quad (5)$$

也就是说：

$$DDV = \frac{(\text{帧内}) + (\text{向后MC})}{(\text{向前}) + (\text{内插MC})} - \frac{(\text{帧内}) + (\text{向前MC})}{(\text{后向}) + (\text{内插MC})} \quad (6)$$

$$DDV = \frac{(NBMB \text{ 每幅图象}) * (\text{后向MC} - \text{向前MC})}{(\text{向前} + \text{内插}) * (\text{向后} + \text{内插})} \quad (7)$$

10 为每个三图象组 IBB 或 PBB 的每个 B 图象计算被称为运动补偿比率的差分 DDV。正如所假设的那样，每组有不多于一个的场景变化，测量这种场景变化的可能性的决策值 DVL 的决定是通过为两个相邻 B 图象加上 DDV 的绝对值确定的：

$$DVL = \frac{|DDV(1)| + |DDV(2)|}{2} \quad (8)$$

15 数 1 和 2 表明比率是与两个后序图象的第一或第二图象有关。关于双向图象的场景变化的准确位置可以通过看两个比率的标记来决定。

- 如果 DDV (1) 和 DDB (2) 大于 0，场景变化在第一 B 图象之前已出现；

20 - 如果 DDV (1) 大于 0，DDV (2) 小于 0，场景变化出现在两个 B 图象之间；

- 如果 DDV (1) 和 DDV (2) 小于 0，场景变化出现在两个 B 图象之后。

25 图 10 展示了仅使用 B 图的场景变化检测方法的执行。每一个 IBB 或 PBB 组的一个决策值被计算了出来，可以看出：

- 决策值的尖峰与真实场景剪接处处于同一位置；

- 宏块 n°50 左右的噪音由序列中的光效应造成，它干扰运动估算算法，进而干扰按照运动补偿的场景剪接处预测；

- 被检测的视频序列的最后一部分基本上是停滞的图象（图象几乎是

同样的)：所以未定义使用那一个运动补偿，因为对全部补偿类型，参考宏块是相同的，虽然没有场景变化出现，决策值因此有一个不可忽略的值（为减少误差的场景变化预测的危险，考虑除 B 图象之外的 P 图象运动矢量的统计数是很有用的：如果 P 图象的检测系数很低，
5 三个先前图象中没有图象变化出现）。

第二子步骤 120 是 GOP 放置子步骤。一个 GOP 的优化安置由两个相互抵触的目标决定：

(a) 第一个是为 GOP 选择一个最佳的大小：如果一个 GOP 太小，就会浪费比特，因为对 I 图象的安置比必须的比特耗费多，而如果一个 GOP
10 太大，随机访问被削弱；

(b) 第二个是将 GOP 的开始图象与一个场景变化相匹配。

因此 GOP 安置的问题是以一种最优方式安排 GOPs，当遇到限制 (a)，(b) 时（既考虑到一个所述 GOP 最大及最小尺寸，在新场景的开始处，开始一个新的 GOP）。为解决这个最优化问题，使用了一个 Viterbi 算法：对于每一个路径将纠正 GOP 最佳尺寸的偏移，而将
15 鼓励在一个 GOP 的开始处包括的一个可能的场景变化。所有决策值的积累数决定着为每一个图象所选择的路径。

此算法检测了序列的 GOP 的最优开始位置。每一图象有一个所附的场景变化决策值，它描述了在各自位置的场景剪接处的可能性；如果决策值大，就意味着在那一位置场景变化的可能性很高，所以安置
20 一个新的 GOP 有利可图。然而，由于 GOP 的尺寸不得太大或太小，在 GOP 的始点之间的转移（即 GOPs 的大小）也被加权。

在图 11 中，垂直线代表图象，参数 S 描述各自图象的场景变化的可能性（假设一个三图象组中只出现一个场景变化，所以每一个 PBB 或 IBB 块中只存在一个决策值 $D_i(N)$ ），W 参数是评价 GOP 尺寸（同
25 最佳尺寸接近的尺寸较合适，例如 12）的加权值。以图象结束的路径的决策值被计算如下：

$$D_i(N) = C_{i-N} + W(N) \quad (9)$$

N 是所考虑 GOP 的大小， C_{i-N} 是所有权数 S 的数量，W 为从图象“1”
30 到图象“i-N”的 GOPs 的最优放置。决策值最高的 GOP 尺寸被选了出来。加权函数 $W=f(N)$ 是二次方程式，所以 W 随 GOP 尺寸与最佳 GOP 尺寸平方差成比例减少，如图 12 示例了 GOP 尺寸的加权函数（所提

出的加权函数使两个图象变化之间全部的 GOP 有大约相同的尺寸：因此，如果两个图象变化有 16 个图象的距离，两个尺寸是 8 的 GOP 被放置而不是一个尺寸是 10 另一个尺寸是 6 的两个 GOP）。

到目前为止，人们并未考虑到图象的传送顺序与显示顺序是不同的。如果认为 GOP 的开始处被分配在场景变化后的第一个图象，那么 GOP 中的 M 图象的第一组也在场景剪接处开始。然而，I 图象是要被放置的块中的最后一个图象。若 $M=3$ ，则 GOP 中的头两个图象被编码为 B 图象，且只有第三个图象是 I 图象。所以，为了保证场景变化后的第一个图象是 I 图象而不是 B 图象，GOP 的开始可以由一个或两个位置转换到左边。

第三个子步骤 130 是 P 与 B 图象放置子步骤。考虑到取决于时间的参数 M 的优化，以适应的方式搜寻 B 和 P 图象的最佳位置确实允许比特流的最小化，这是在为序列编码时所需的。增加 M 的值就增加了 P 图象的比特率，但是使用了更省比特率的 B 图象而不是 P 图象。所以以后序图象间的相关性对于优化最重要的参数，他实际上被划分为两个子任务：

(a) 长期优化，为了在几个 GOP 上找到最佳的 M

(b) 短期优化，当考虑到图象间相关性的局部变化时，将在 GOP 内部找到 B 和 P 图象的最佳位置

对于长期优化，必须注意到，如果连续图象之间的相关性系数趋向于一，无论是选择 B 还是 P 图象都无所谓，因为在任何情况下都几乎没有系数比特剩余，而如果所述关连很低，运动补偿就不会工作。在这些极端情况下（分别是一个停滞图象和不关连图象），哪一个 M 是合适的，并不明显。在其他情况下，通常可以说对于低相关性序列，小 M 较好，对于一个高度相关的序列，大 M 更好一些。如果对大量场景进行实验，就会取得 M 长期优化的最好结果。

对于短期优化，M 可以在每个 GOP 内部任意变化，这使得使用图象间相关性的短期变化成为可能，其目的是使比特率最小化。下表给出了一个 M 短期优化的例子，显示了场景变化之前 M 的选择：

场景 1				场景 2	
位置	1	2	3	4	
M = 1	P	P	P	I	

M = 2	P	P	B (同 P 一样)	I
M = 3	P	B (同 P 一样)	B	I

很明显，场景变化之前的 B 图象只可以被前向预测。在新场景之前 M=1 还是 2 并没有很大区别，因为场景变化之前的 B 图像的运作象 P 图象一样；很明显 M=3 的一个选择是不太好的，因为在位置 3 的 B 5 图象在位置 1 使用一个参考图象（先前的 P 图象），从而远离两个位置。由于图象间的相关性随它们间距离的增加而减少，M=3 在位置 3 处的 B 图象比特率比 M=1 或 2 在同一位置的图象比特率高。

第三步 43 是预测步骤，用以预测量化步长 Q，依据明确的目标比特率对比特流编码就必须使用该步骤。预测步骤一完成，如果需要就 10 可以重复分析步骤 41（图 4 中的箭头），以此获得对 Q 的更准确估算（然而，一个好的预测通常是在几次例如 2 次运转之后获得的）。

由于第二步末尾中存在的量化步长 Q 仅是一个估算值，如果使用所说预测值为每一个图象编码，总比特预算就不会被准确匹配。一个最终步骤 44 确保严格遵守对总平均比特率的限制。为确保最终输出 15 比特流确实是理想大小，执行一个量化步长控制程序。此程序基于一个控制回路，此控制回路依赖于预测的与真实的比特率的比较。在最终步骤中为每一个图象编码后，控制程序对消耗的比特总数与允许数量进行比较。如果消耗了比预算多的比特，量化步长就增加了，随后图象的比特率就减少了。如果消耗了比预算少的比特，Q 减少，比特 20 率增加，总目标比特率最终严格匹配。

所说的 VBR 编码方法可以在一个编码器中执行，此编码器的结构如图 5 所示，其中每个块与一个在控制器 55 监视下执行的一个特殊函数对应。所示编码器包括一个输入缓冲器 51，一个减法器 549，一个 DCT 电路 521，一个量化电路 522，一个可变长度编码电路 523， 25 及一个输出缓冲器 524。电路 521 到 524 包括与预测分支 53 相关的编码分支 52 的主要元素，此预测分支 53 包括一个反向量化电路 531，一个反向的 DCT 电路 532 及一个预测子系统。预测子系统本身包括一个加法器 541，一个缓冲器 542，一个运动估算电路 543（此估算基于存在于缓冲器 51 输出的输入信号的分析），一个运动补偿电路 544 30 （其输出信号被向后传递到加法器 541 的第二输入端），减法器 549

(为向编码分支传送差分,接收缓冲器 51 的输出信号及运动补偿电路 544 的输出信号)。

所示编码器的输出被传送到包括执行最终步骤 44 的控制回路的控制器 55。用于 VBR 编码器最终路径的所述控制回路的主要元素在图 13 中显示。正如已解释的那样,必须在这个最终编码路径中调整量化大小,以此保证算子提供的总目标比特率严格匹配。所述回路首先包括一个第一计算电路 131,其中回路的输出(即积累预测误差)乘以一个因数 KP。因数本身等于一个常数 QC(由算子选择)乘以一个加权因数 Q_{int}/APG ,其中 Q_{int} 是一个 Q 及 APG(N 图象 GOP 比特的总数)的综合估算。

然后加法器 133 将电路 131 的输出 Q-prop 与在第二计算电路 132 输出处的信号 Q_{int} 相加,此电路 132 用于产生一个 Q 的综合估算。基于存储在电路 134 中的等式 $R=f_i(Q)$ (在加法器 63 的输出的量化因数 Q 与比特率 R 之间),一个反相电路 134 为所有先前图象给出积累的比特率。将以此所得的积累比特率在比较器 135 中与在比较器第二个输出端的积累预测比特率比较,且经过在电路 136 中的综合后被使用,以相应地修改 Q。

以上所示的 VBR 编码策略是对先前 VBR 编码器的一个改进,因为它获得了解码序列视在参数的一个更好量化。传统的 VBR 编码器在编码一个图象时调整量化参数,因此预测比特率与每一个图象相匹配。从而它们允许量化参数 Q 在一个图象内部变化,且无法获得该图象的不变空间品质。无论对图象的比特率的预测正确与否,这种品质的变化都会发生。对于所提出的 VBR 编码,一个图象的 Q 是不变的,且视频序列中的任意图象的空间品质都是不变的。如果对图象比特率和量化步骤宽度的估算是正确的, Q 在适应性量化及主观变形之前,对序列的所有宏块保持不变。由于量化步骤宽度与图象比特率仅仅是被估算了, Q 的一个变化及由此而来的序列品质的变化在图象之间出现,但是在一些分析路径后,图象平均 Q 的变化通常在 1%之下。

除获得不变的帧内图象品质之外,新 VBR 策略还涉及一些其它重要方面:

- 以重复的方式通过增加分析路径的数量来改进量化因素的预测是可能的:在分析过后,如果与所须目标比特率的偏离仍旧很高,可使

用先前编码路径中的结果计算量化因素的一个更好的估算;

- 由于新 VBR 编码策略预测 Q , 可以利用以另一个图象顺序执行的分析路径, 而不用预测的路径: 这在旧策略中是不可能的, 它是新编码概念的一个主要优点;

- 5 - 如果在最后的路径中, Q 的变化及随之而来的质量变化大的不可接受, 最终的步骤可作为对 Q 及相应路径比特率的预测的一个分析路径: 使用这一特点, 就可以发觉一个可执行所有需要编码的路径的一个编码器, 直到输出比特流的性能在由算子定义的一定限制之内;
- 由于控制回路有一个综合的特点, 短期比特率预测误差相互抵偿:
- 10 所以系统的图象型的预测误差并不严重影响所建议 VBR 编码器的性能。

很明显, 本发明并不局限于前面所述, 在不脱离发明的范围内有一些变化和改进。例如, 一个可选择的第四子步骤(图 4 中的标号 140, 并用点状线显示)可以被包括到重新排列步骤 42。为了在最后一步

15 44 的最终路径中以一个给定的比特率 $R(t)$ 为序列编码, 必须预测一个目标量化步骤宽度 Q 和目标图象比特率 $R(I)$ 。执行一个分析路径的唯一要求是有一个 Q 的预测值。由于在第一分析步骤 41 中没有使用针对 Q 的控制系统, 无须预测图象目标。对于比特率 $R(I)$ 和步骤宽度 Q 的估算, 需要先前编码路径的量化因素及图象比特率。然而,

20 如果图象类型的顺序在两个路径间变化, 在两个路径中的不同图象类型对一个序列的相同图象编码。

如果认为执行分析路径是使用 $N=12$ 和 $M=3$, 而为预测路径编码是使用 $N=8$ 和 $N=2$, 相应的图象放置如下表所示:

N	M	图象显示次序
12	3	BBIBBPBBPBBPBBIB
8	2	BIBPBPBPBIBPBBBP

- 25 其中第二个图象在第一路径中被编码为 B 图象, 在第二路径中被编码为 I 图象。由于使用比特率预测与第一分析路径中的图象类型相同的图象的目标比特率, 如果图象在第一路径中被编码为 B 图象, 就预测第二路径的 B 图象比特率。在一个被修改了的图象顺序的情况下, 预测图象比特率就没有用处了。

- 由于图象重组后目标图象比特率的预测是不可能的，图象重新排列后并不能直接执行最终编码路径。所以在最终编码路径之前必须执行一个分析路径：因此依据本发明，在那种情况下对于 VBR 编码器至少需要三个编码路径。为了保证最终路径的预测图象比特率不会偏差太大，可以提供一个帧内图象预测附加子步骤，它将估算当分析路径以新的而不是旧的图象顺序执行时图象拥有的比特率。选择性的附加子步骤 140 使用图象比特率的时间相关性。
- 5

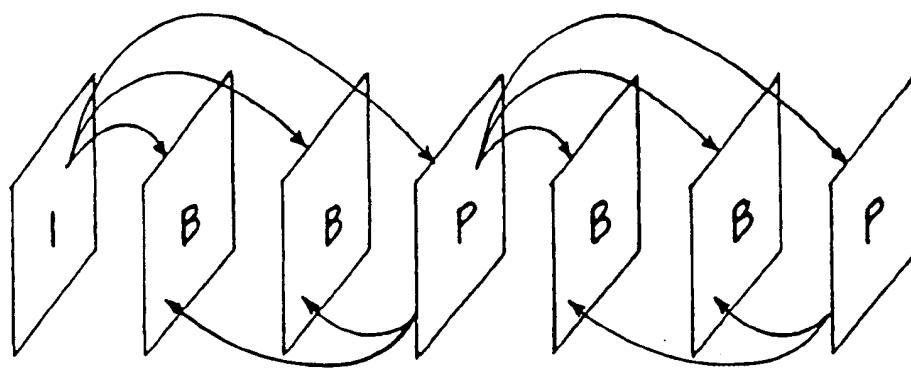


图 1

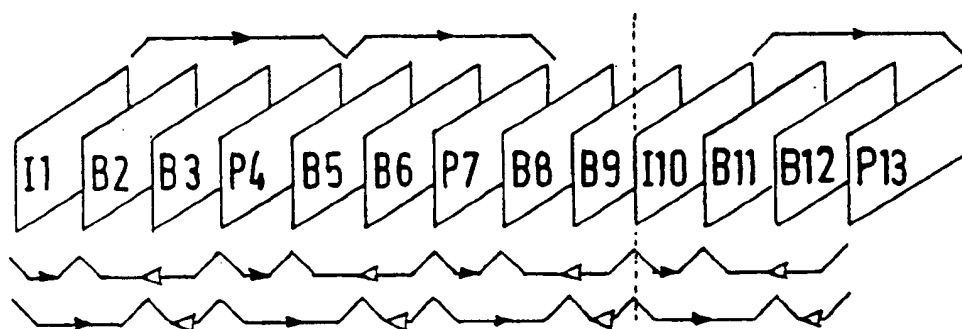


图 2

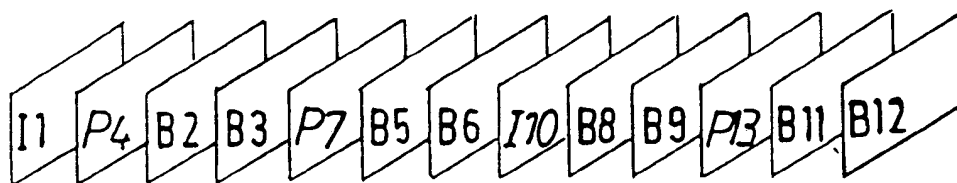


图 3

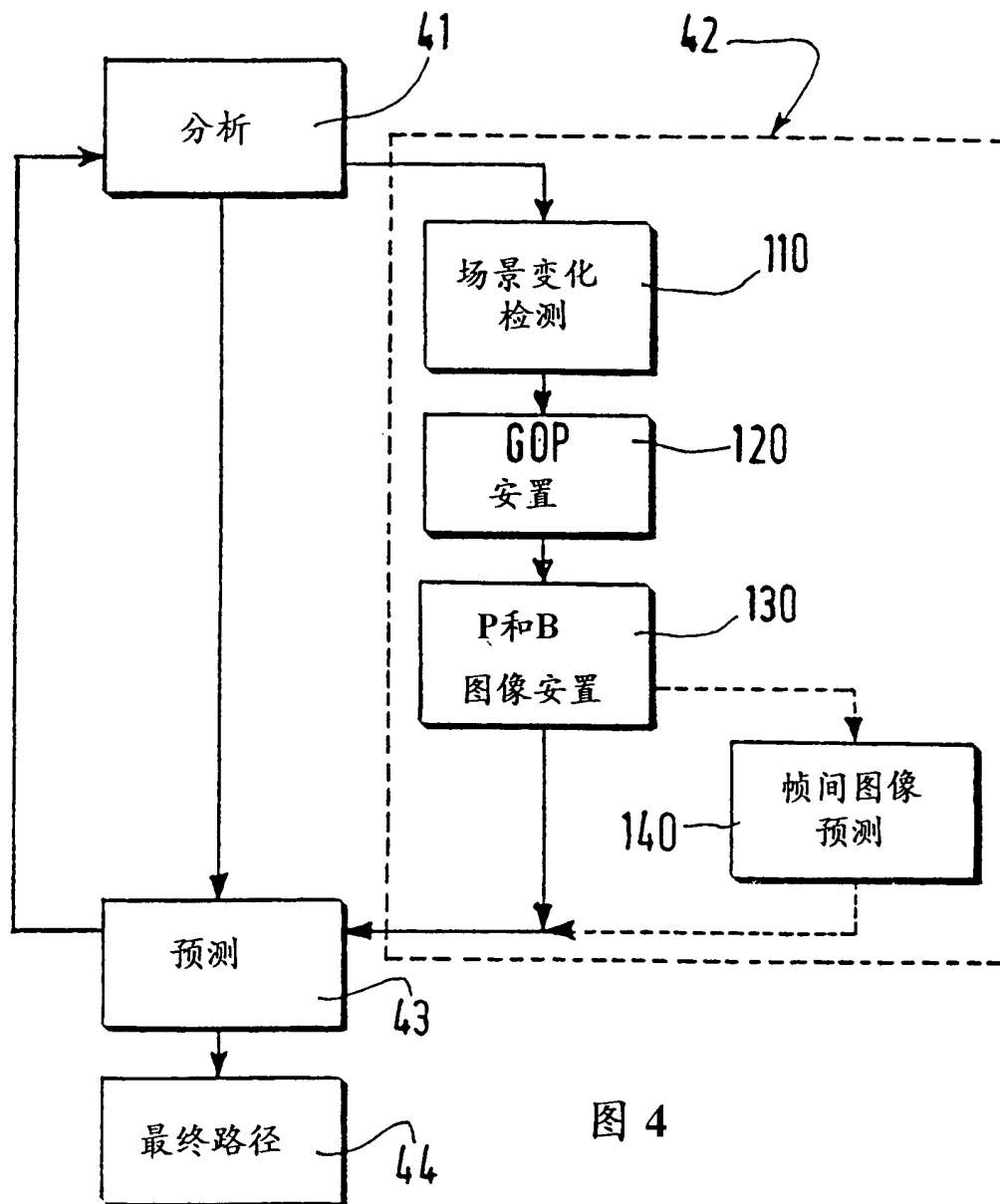


图 4

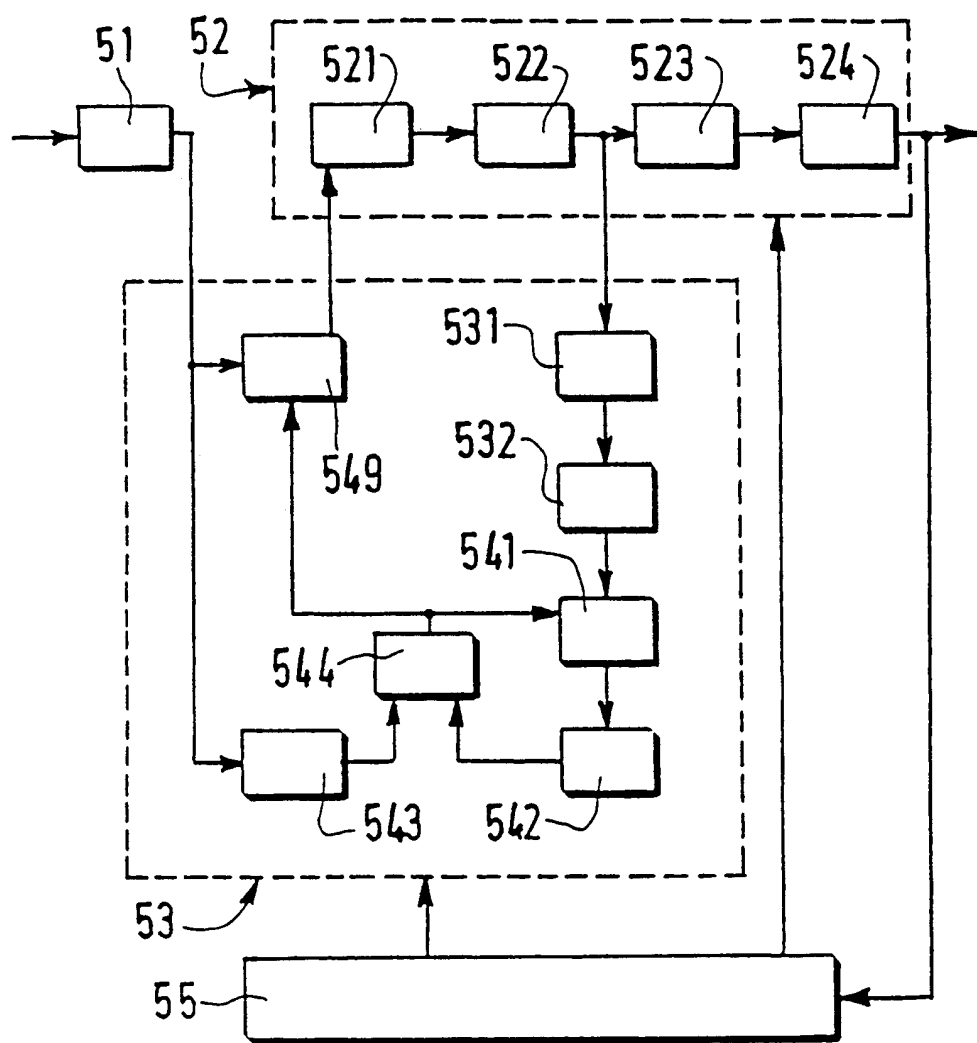


图 5

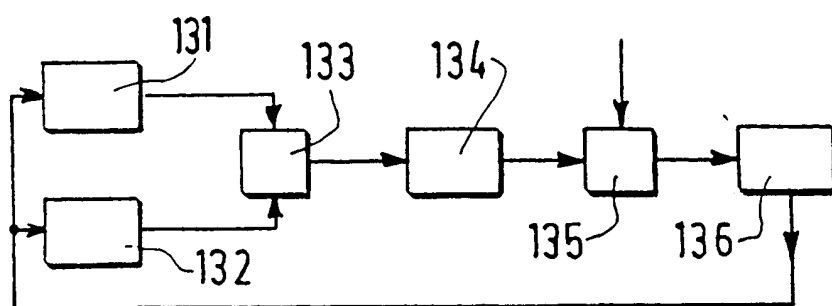


图 13

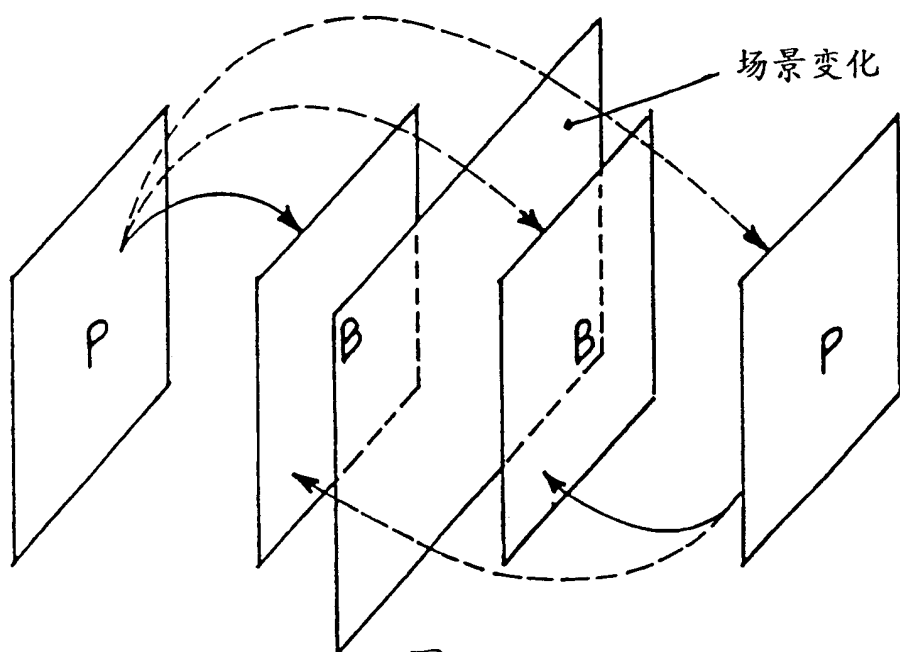


图 6

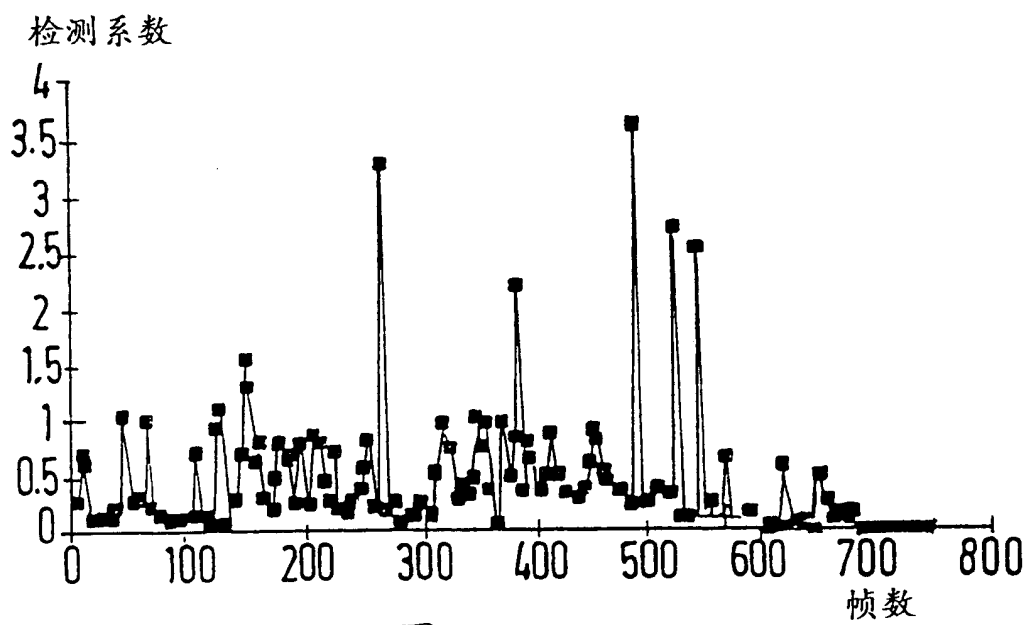


图 9

检测系数左

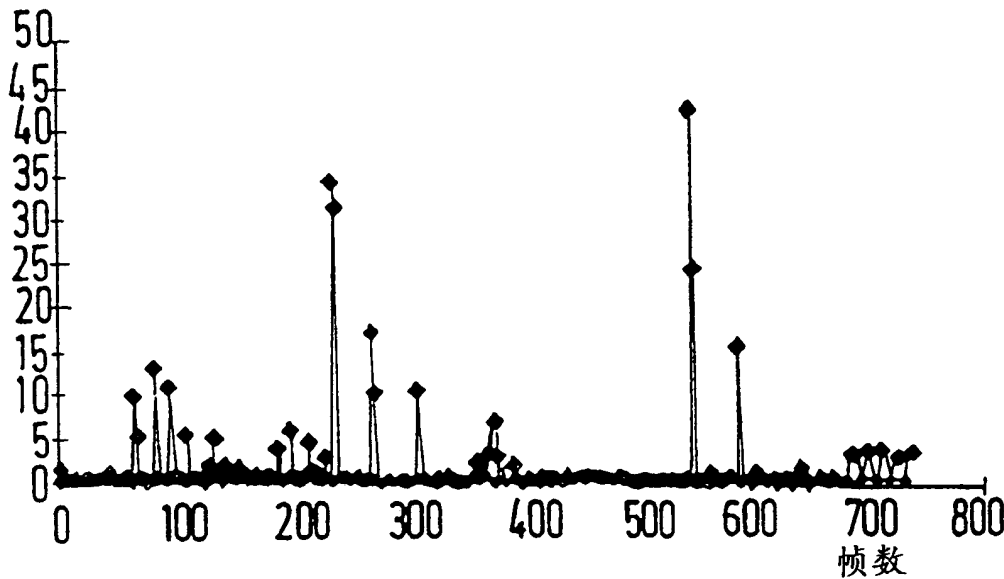


图 7

检测系数右

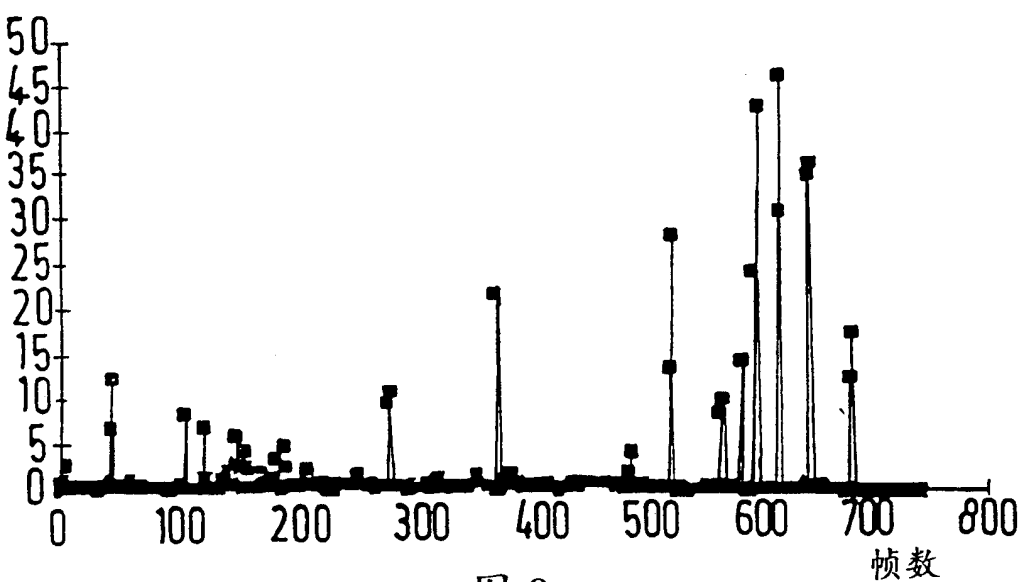


图 8

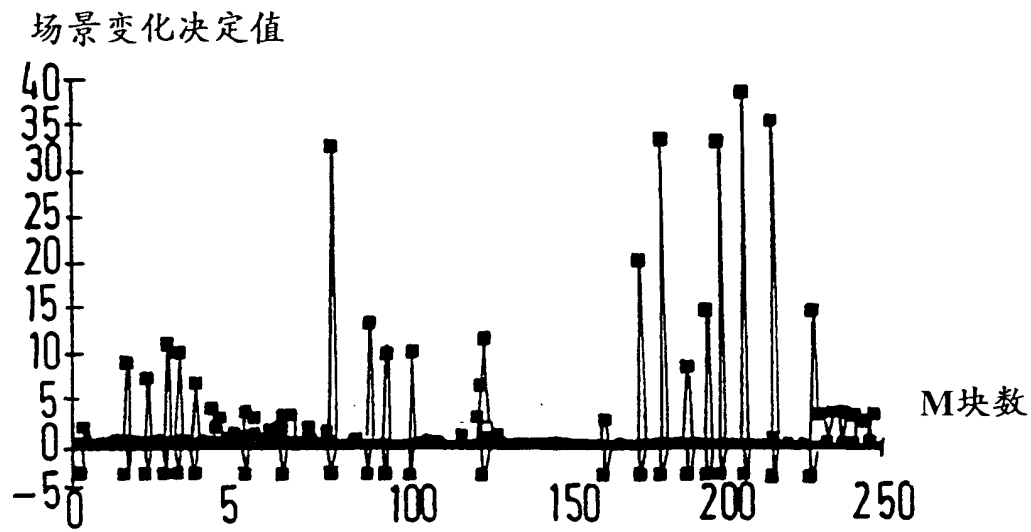


图 10

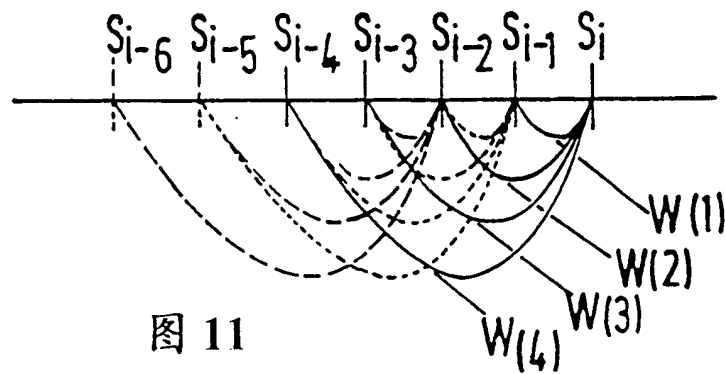


图 11

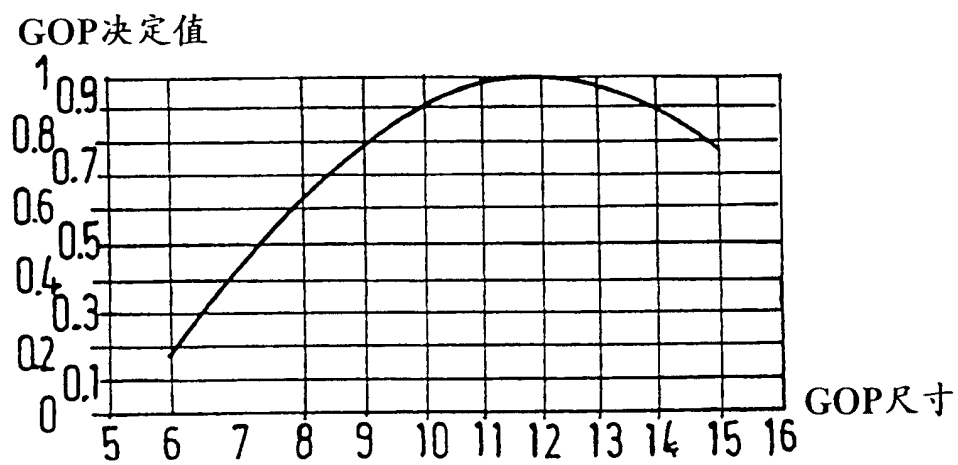


图 12