[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
H04N 7/26 (2006.01)
H04N 7/50 (2006.01)

[12] 发明专利申请公布说明书

[21] 申请号 200580019971.X

[43] 公开日 2007 年 5 月 23 日

[11] 公开号 CN 1969558A

[22] 申请日 2005.6.6
[21] 申请号 200580019971.X
[30] 优先权
[32] 2004.6.18 [33] US [31] 60/581,280
[86] 国际申请 PCT/US2005/019772 2005.6.6
[87] 国际公布 WO2006/007285 英 2006.1.19
[85] 进入国家阶段日期 2006.12.18
[71] 申请人 汤姆逊许可公司
地址 法国布洛涅
[72] 发明人 亚历山德罗斯·米歇尔·托拉彼斯
吉尔·麦克唐纳·瓦布斯 尹 鹏
[74] 专利代理机构 北京康信知识产权代理有限责任公司
代理人 章社杲

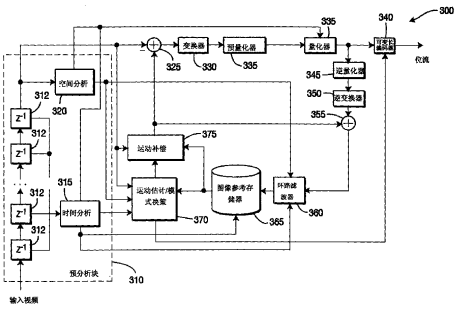
权利要求书 4 页 说明书 19 页 附图 3 页

[54] 发明名称

用于视频编码优化的方法和设备

[57] 摘要

一种用于对与多个图像相对应的视频信号数据进行编码的编码器及相应的方法。该编码器包括重叠窗口分析单元(310)，该重叠窗口分析单元用于利用多个重叠分析窗口，对与视频信号数据相对应的多个图像的至少一些执行视频信号数据的视频分析，并基于视频分析的结果来修改视频信号数据的编码参数。



1. 一种用于对与多个图像相对应的视频信号数据进行编码的编码器，该编码器包括重叠窗口分析单元（310），用于利用多个重叠分析窗口，对与所述视频信号数据相对应的所述多个图像的至少一些执行所述视频信号数据的视频分析，并基于所述视频分析的结果来修改所述视频信号数据的编码参数。

2. 根据权利要求 1 所述的编码器，其中所述重叠窗口分析单元（310）利用空间－时间分析来执行所述视频信号数据的视频分析。

3. 根据权利要求 2 所述的编码器，其中所述重叠窗口分析单元（310）利用图像编码类型信息、边缘信息、均值信息和方差信息的至少之一，来执行所述空间－时间分析、拉格朗日参数和量化参数的修改、以及死区处理的至少之一。

4. 根据权利要求 3 所述的编码器，其中所述重叠窗口分析单元（310）利用绝对差和方差来修改所述量化参数。

5. 根据权利要求 1 所述的编码器，其中所述重叠窗口分析单元（310）利用初遍编码来执行所述视频信号数据的视频分析。

6. 根据权利要求 1 所述的编码器，其中所述重叠窗口分析单元（310）利用空间－时间分析和初遍编码两者，来执行所述视频信号数据的视频分析。

7. 根据权利要求 6 所述的编码器，其中所述重叠窗口分析单元（310）利用图像编码类型信息、边缘信息、均值信息和方差信息的至少之一，来执行所述空间－时间分析、拉格朗日参数和量化参数的修改、以及死区处

理的至少之一。

8.根据权利要求7所述的编码器，其中所述重叠窗口分析单元（310）利用绝对差和方差来修改所述量化参数。

9.根据权利要求1所述的编码器，其中所述视频信号数据包括多个帧，所述多个帧的每一帧都代表相应的图像，并且所述重叠分析单元执行所述视频分析、以便只考虑先前已编码的图像。

10.根据权利要求1所述的编码器，其中所述编码参数包括以下至少之一：片类型，图像和图像组（GOP）编码结构和顺序，加权参数，量化值和死区处理，拉格朗日参数，参考数，参考顺序和处理，帧/场图像和宏块参数，去块参数，帧间块大小，帧内空间预测，及直接模式。

11.根据权利要求1所述的编码器，其中所述重叠窗口分析单元（310）多次迭代地执行所述视频分析，并基于先前所产生的分析统计特性来修改所述编码参数和分析统计特性之一。

12.根据权利要求1所述的编码器，其中所述重叠窗口的每一个都具有P个图像的窗口大小、以及与所述窗口大小相关联的重叠大小，并且所述重叠窗口分析单元基于先前所产生的分析统计特性来修改所述窗口大小和所述重叠大小。

13.一种用于对与多个图像相对应的视频信号数据进行编码的方法，包括以下步骤：

利用多个重叠分析窗口，对与所述视频信号数据相对应的所述多个图像的至少一些执行（425、430）所述视频信号数据的视频分析；以及

基于所述视频分析的结果来修改(435)所述视频信号数据的编码参数。

14. 根据权利要求 13 所述的方法，其中所述执行步骤利用空间－时间分析来执行所述视频信号数据的视频分析。

15. 根据权利要求 14 所述的方法，其中所述执行步骤和修改步骤分别利用图像编码类型信息、边缘信息、均值信息和方差信息的至少之一，来执行所述空间－时间分析、拉格朗日参数和量化参数的修改、以及死区处理的至少之一。

16. 根据权利要求 15 所述的方法，其中利用绝对差和方差来修改所述量化参数。

17. 根据权利要求 13 所述的方法，其中所述执行步骤利用初遍编码来执行所述视频信号数据的视频分析。

18. 根据权利要求 13 所述的方法，其中所述执行步骤利用空间－时间分析和初遍编码两者，来执行所述视频信号数据的视频分析。

19. 根据权利要求 18 所述的方法，其中所述执行步骤和修改步骤分别利用图像编码类型信息、边缘信息、均值信息和方差信息的至少之一，来执行所述空间－时间分析、拉格朗日参数和量化参数的修改、以及死区处理的至少之一。

20. 根据权利要求 19 所述的方法，其中利用绝对差和方差来修改所述量化参数。

21. 根据权利要求 13 所述的方法，其中所述视频信号数据包括多个帧，所述多个帧的每一帧都代表相应的图像，并且所述执行步骤执行所述视频分析、以便只考虑先前已编码的图像。

22.根据权利要求 13 所述的方法，其中所述编码参数包括以下至少之一：片类型，图像和图像组（GOP）编码结构和顺序，加权参数，量化值和死区处理，拉格朗日参数，参考数，参考顺序和处理，帧/场图像和宏块参数，去块参数，帧间块大小，帧内空间预测，及直接模式。

23.根据权利要求 13 所述的方法，其中所述执行步骤多次迭代地执行所述视频分析，并且所述修改步骤基于先前所产生的分析统计特性来修改所述编码参数和分析统计特性之一。

24.根据权利要求 13 所述的方法，其中所述重叠窗口的每一个都具有窗口大小、以及与所述窗口大小相关联的重叠大小，并且所述执行步骤包括基于先前所产生的分析统计特性来修改所述窗口大小和所述重叠大小的步骤。

用于视频编码优化的方法和设备

交叉引用相关申请

本申请要求2004年6月18日提交的美国临时申请序号No.60/581,280的权益，其在此整个被引入作为参考。

技术领域

本发明一般涉及视频编码器和解码器，并且尤其涉及一种用于视频编码优化的方法和设备。

技术背景

多遍视频编码方法已用于许多视频编码架构中，如MPEG-2（运动图像专家组2）和JVT（联合视频工作组）/H.264/MPEG AVC（高级视频编码），以便获得更好的编码效率。这些方法背后的思想是，通过执行分析并收集能够用于将来迭代中的统计特性以试图提高编码性能，来设法利用几次迭代来对整个序列进行编码。

两遍编码方案已经用于几种编码系统中，包括MICROSOFT®（微软）WINDOWS MEDIA®和REALVIDEO®编码器。根据这种编码方案，编码器首先利用一些初始预定的设置来对整个序列执行初遍编码，并收集关于该序列内每个图像的编码效率的统计特性。在完成该过程之后，通过同时考虑先前所产生的统计特性来再一次重新处理并编码整个序列。这可以大大提高编

码效率，甚至允许我们满足某些预定的编码约束或要求，例如对于编码后的流满足给定的比特率约束。这是因为，编码器现在更加了解整个视频序列或图像的特征，从而能够更加适当地选择参数，如将要用于编码的量化器、死区处理（deadzoning）等。能够在该第一遍编码期间收集的、并且能够用于该用途的一些统计特性是每个图像的位数、空间活性（即平均标准化宏块方差和均值）、时间活性（即运动矢量/运动矢量方差）、失真（例如均方差（MSE））等。虽然利用这些方法能够大大提高编码性能，但是这些方法也倾向于具有很高的复杂性，只能脱机使用（先对整个序列编码、然后执行第二遍编码），而不适于实时编码器，并且不一定考虑到可以从第一步编码推断的所有可能统计特性。

发明内容

本发明致力于解决现有技术的这些及其它缺点和不利条件，本发明涉及一种用于视频编码优化的方法和设备。

根据本发明的一方面，提供了一种用于对与多个图像相对应的视频信号数据进行编码的编码器。该编码器包括重叠窗口分析单元，该重叠窗口分析单元用于利用多个重叠分析窗口，对与视频信号数据相对应的多个图像的至少一些图像执行视频信号数据的视频分析，并基于视频分析的结果来修改该视频信号数据的编码参数。

根据本发明的另一方面，提供了一种用于对与多个图像相对应的视频信号数据进行编码的方法。该方法包括以下步骤：利用多个重叠分析窗口，对与视频信号数据相对应的多个图像的至少一些图像执行视频信号数据的视频分析，并基于视频分析的结果来修改视该频信号数据的编码参数。

根据以下连同附图一起阅读的示范性实施例的详细描述，本发明的这

些及其它方面、特征和优点将变得很清楚。

附图说明

根据以下示例图可以更好地理解本发明，其中：

图 1 示出了根据本发明原理的、基于窗口的示范性两遍编码架构的框图；

图 2 示出了根据本发明原理的、在变换和量化期间死区处理的影响的曲线图；

图 3 示出了根据本发明原理的编码器的框图；以及

图 4 示出了根据本发明原理的示范性编码过程的流程图。

具体实施方式

本发明涉及一种用于视频编码优化的方法和设备。有利的是，本发明允许视频编码器在给定的特定比特率下，以大大提高的主观和客观质量来压缩视频序列。这是通过对与仍待编码的 N 个随后图像相比的当前图像执行简单分析，通过视频序列的非因果处理而得以实现的。然后，编码器利用该分析结果，对将要用于当前图像的编码的编码参数（包括但不限于图像/片类型、量化器、阈值处理参数、拉格朗日算子λ等）作出更好的决定。和对整个序列执行两遍或多遍编码以获得更好编码性能的几种现有技术系统不同，本发明较简单，从而对复杂性具有较小的影响。本发明的原理也可以和其它多遍编码策略一起使用，以获得更高的效率。以类似的方式，可以创建因果系统（使用 M 个先前已编码的图像）。

按照本发明的原理，首先只对整个序列的重叠图像窗口子集进行分析。基于所产生的统计特性，来适当地调节每个图像的编码参数。这些编码参数可以包括但不限于：图像/片类型决策（I、P、B）、帧/场决策、B 图像距离、图像或宏块（MB）量化值（QP）、阈值处理系数、拉格朗日参数、色度偏移、加权预测、参考图像选择、多块大小决策、熵参数初始化、帧内模式决策、去块滤波器参数等。可以利用可能需要不同复杂性代价的分析方法来执行图像/宏块分析，包括完全第一遍编码、具有空间分析的简单第一遍运动估计，乃至简单的时间与空间分析度量，包括但不限于方差、图像差等。此外，重叠图像窗口（及重叠图像）可以根据需要变大或变小（变多或变少），由此提供不同的延迟/性能折衷。

本说明书举例说明本发明的原理。因而，应该理解，本领域技术人员将能够设计出虽然在此未明确描述或示出、但是能具体实施本发明原理且被包括在本发明精神和范围内的各种配置。

在此列举的所有例子和条件语言都试图用于教学，以帮助读者理解本发明人所提出的、有助于促进本技术领域发展的本发明的原理和概念，并且应该被认为不限于这些具体列举的例子和条件。

而且，在此列举本发明的原理、方面和实施例及本发明的特定例子的所有语句，都试图包括本发明的结构和功能等效物。另外，这种等效物试图包括当前已知的等效物以及将来开发的等效物，即与结构无关的、执行相同功能的任何被开发元件。

因而，例如本领域技术人员应该理解，在此给出的框图代表具体实施本发明原理的说明性电路的概念图。类似应该理解，任何流程表、流程图、状态转移图、伪码等，都代表实质上可以在计算机可读介质中表示、从而可通过计算机或处理器来执行的各种过程，而不管这种计算机或处理器是否明确示出。

通过利用专用硬件以及能够与适当软件关联地执行软件的硬件，可以提供图中所示的各种元件的功能。当通过处理器来提供这些功能时，可以通过单个专用处理器、通过单个共享处理器、或通过其中一些可以被共享的多个单独处理器，来提供这些功能。而且，术语“处理器”或“控制器”的显式使用不应被认为专门指能够执行软件的硬件，而是可以隐含地包括但不限于数字信号处理器（DSP）硬件、用于存储软件的只读存储器（ROM）、随机存储器（RAM）和非易失性存储器。

也可以包括其它常规和/或定制的硬件。类似地，图中所示的任何开关都只是概念性的。如从上下文更具体地理解的，可以通过程序逻辑的操作、通过专用逻辑、通过程序控制和专用逻辑的交互、乃至手动操作的可以由实施者选择的特定技术来实现它们的功能。

在本发明权利要求中，被表示为用于执行指定功能的装置的任何元件都试图包括执行该功能的任何方式，例如包括：a)执行该功能的电路元件的组合；或 b)任何形式的软件，因而包括和用于执行该软件的适当电路相结合以执行该功能的固件、微码等。如这些权利要求所定义的本发明在于，以权利要求所要求的方式来组合并集合由所列举的各种装置所提供的功能性。因而，申请人认为能够提供这些功能性的任何装置都等效于在此所示出的那些装置。

根据本发明的原理，公开了一种新的多遍编码架构，与在每一遍编码期间都考虑整个视频序列或独立窗口的以前方法不同，该新的多遍编码架构对允许在相邻窗口之间重新使用先前确定的特征的重叠窗口，执行每遍编码。虽然由于能够利用少得多的步骤来实现最优编码，而得到更低的成本/复杂性、以及更小的存储器需求/低等待时间，但是该架构仍然能够获得多遍编码的好处，如大大提高的视频质量。考虑到由于相邻窗口之间的相似性，编码器有可能甚至在第一遍编码期间就决定了最佳参数、从而不

需要更多的迭代进行最终编码，所以该特征对于实时编码应用尤为重要。

转到图 1，一般用附图标记 100 来表示基于窗口的两遍编码架构。处理/分析窗口的大小为W_p个图像，而相邻两组之间允许的重叠大小为W_o个图像。第一窗口的处理将提供某些初始统计特性，这些初始统计特性能够用于为该窗口内的所有帧确定一组初步编码特征。更具体地说，如果使用两遍方案，则可以基于所产生的参数，立即对也不属于将来窗口的所有帧进行编码。然而，该信息能够立即用于该将来窗口的处理/分析。例如，在该窗口的处理期间这些参数可用作初始种子，并且考虑到大多数序列中所存在的高度时间相关，这些参数能够改进分析。更重要的是，可以基于新产生的统计特性，对用于由于W_o的选择而也属于先前窗口的该窗口中的初始帧的编码参数，进行进一步改进/限制。例如在处理整个序列或M个相邻窗口之后，如果使用较多迭代/遍，这基本上允许更快地收敛到最优解。显然，取决于编码器的能力或需求，时间窗口可以尽可能的大或小，同时也可利用不同窗口大小（更大或更小的W_o和W_p）来执行这种方案的迭代。

在我们的多遍方案的预分析步骤期间，可以使用多种不同的准则。这种准则可以取决于编码器架构的复杂性约束，并且可以考虑从简单的空间—时间方法（包括但不限于边缘检测、纹理分析度量以及绝对图像差）到更复杂的策略（包括但不限于离散余弦变换（DCT）分析、第一遍帧内编码、运动估计/补偿、乃至完全编码）。也可以通过增加或减少分析和/或重叠窗口，来调节等待时间。

作为这种系统的例子，在该分析期间可以计算以下准则：

对于窗口W_p内的每个图像k，进行以下计算：

(i) 对于位置(i,j)处的每个宏块，如下计算均值$MBmean(k,i,j)$：

$$MBmean(k,i,j)=\frac{1}{B_W\times B_H}\sum_{y=0,x=0}^{y=B_H-1,x=B_W-1}c[k,i\times B_W+x,j\times B_H+y]$$

(ii) 如下计算均方值 *MBsqmean*(*k*, *i*, *j*)：

$$MBsqmean(k,i,j)=\frac{1}{B_W\times B_H}\sum_{y=0,x=0}^{y=B_H-1,x=B_W-1}\left(c[k,i\times B_W+x,j\times B_H+y]\right)^2$$

(iii) 如下计算方差值 *MBvariance*(*k*, *i*, *j*)：

$$MBvariance(k,i,j)=MBsqmean(k,i,j)-(MBmean(k,i,j))^2$$

(iv) 对于整个图像，如下计算平均宏块均值AMM_k：

$$AMM_k=\frac{1}{PMB_W\times PMB_H}\sum_{j=0,i=0}^{j=PMB_H-1,i=PMB_W-1}MBmean(k,i,j)$$

(v) 如下计算平均宏块方差AMV_k：

$$AMV_k=\frac{1}{PMB_W\times PMB_H}\sum_{j=0,i=0}^{j=PMB_H-1,i=PMB_W-1}MBvariance(k,i,j)$$

(vi) 如下计算图像方差PV_k：

$$PV_k=\frac{1}{PMB_W\times PMB_H}\sum_{j=0,i=0}^{j=PMB_H-1,i=PMB_W-1}MBsqmean(k,i,j)-AMM_k^2$$

其中*c*[*x*, *y*]对应于位置(*x*, *y*)的像素值，PMB_W和PMB_H分别是用宏块表示的图像宽度和高度，以及B_W和B_H是当前图像中每个宏块的宽度和高度（通常为$B_W=B_H=16$）。

此外，也可以如下计算以下相对于图像 *m*（例如 *m*=*k*+1）的时间特征：

(I) 如下计算平均绝对图像差$MAPD_{k,m}$：

$$MAPD_{k,m}=\frac{1}{PMB_W\times PMB_H\times B_W\times B_H}\sum_{y=0,x=0}^{y=PMB_H\times B_H-1,x=PMB_W\times B_W-1}|c[k,x,y]-c[m,x,y]|$$

(II) 如下计算平均绝对加权图像差 $MAWPD_{k,m}$：

$$MAWPD_{k,m}=\frac{1}{PMB_W\times PMB_H\times B_W\times B_H}\sum_{y=0,x=0}^{y=PMB_H\times B_H-1,x=PMB_W\times B_W-1}\left|c[k,x,y]-\frac{AMM_k}{AMM_m}c[m,x,y]\right|$$

(III) 如下计算平均绝对偏移图像差 $MAWPD_{k,m}$：

$$MAWPD_{k,m}=\frac{1}{PMB_W\times PMB_H\times B_W\times B_H}\sum_{y=0,x=0}^{y=PMB_H\times B_H-1,x=PMB_W\times B_W-1}\left|c[k,x,y]-c[m,x,y]+AMM_k-AMM_m\right|$$

(IV) 如下计算均方图像差 $MSPE_{k,m}$：

$$MSPE_{k,m}=\frac{1}{PMB_W\times PMB_H\times B_W\times B_H}\sum_{y=0,x=0}^{y=PMB_H\times B_H-1,x=PMB_W\times B_W-1}\left(c[k,x,y]-c[m,x,y]\right)^2$$

(V) 如下计算图像方差之绝对差 $APVD_{k,m}$：

$$APVD_{k,m}=\left|PV_k-PV_m\right|$$

可以计算的其它空间－时间特征是：直方图的绝对差，绝对差的直方图，k 与 M 之间的 χ^2 度量，使用任何（乃至多个）边缘算子（包括但不限于 canny、sobel 或 prewitt 边缘算子）的 k 的边缘，乃至用于检测序列交错特性的基于场的度量。可能有用并且可以从以上推断的另外两种统计信息是，当前图像到最接近的过去（$last_idistance_k$）及最接近的将来（$next_idistance_k$）已编码帧内图像的距离，如例如用图像数、编码顺序或图像顺序计数（poc）所度量的。通过考虑场景变化/镜头检测器（shot detector）和/或默认的图像组（GOP）结构，可以增强这些统计特性。可以利用原始或重构图像来计算时间特征（例如，如果本发明应用于多遍编码实施中），同时这些度量的计算也可考虑运动估计/补偿。

基于以上度量，编码器可以决定修改与编码过程相关的某些图像、宏块乃至子块参数。这些包括诸如以下的参数：量化值（QP），死区/阈值处理系数，用于宏块编码以及帧与场间图像级决策的拉格朗日值，去块滤波器参数，编码和参考图像排序，场景/镜头（包括但不限于淡入淡出/溶解/

擦除/闪烁等）检测，GOP 结构等。

在本发明的一个说明性实施例中，当对$cur_slice_type_k$片类型的图像k进行编码时，如下考虑以上参数以执行图像QP修改。在该实施例中，将$distance_{k,k+1}$看作用图像数表示的相邻两个图像之间的距离：

```
if (next_idistance_k > 3 && cur_slice_type_k == I_Slice)
{
   if (PV_k<1 && MAPD_{k,k+1}<1 && last_idistance_k > 5*distance_{k,k+1})
     QP_k = QP_k-4
   else if (MAPD_{k,k+1}<3 && (k==0 // last_idistance_k > 5*distance_{k,k+1}))
     QP_k = QP_k-3
   else if (MAPD_{k,k+1}<10)
     QP_k = QP_k-2
   else if (MAPD_{k,k+1}<15)
     QP_k = QP_k-1
}
else if (AMV_k>10 && AMV_k<60)
{
   if (PV_k<500 && next_idistance_k > 3*distance_{k,k+1})
   {
     if (MAPD_{k,k+1}<10 && AMV_k<35 && last_idistance_k > 2*distance_{k,k+1})
          QP_k = QP_k-2
     else
          QP_k = QP_k-1
}
else if (PV_k<1500 && next_idistance_k > 0)
{
     if (MAPD_{k,k+1}<25)
               QP_k = QP_k-1
   }
}
     else if (MAPD_{k,k+1}==0 && next_idistance_k > 3*distance_{k,k+1} && last_idistance_k
>4*distance_{k,k+1})
   QP_k = QP_k-2
else (((MAPD_{k,k+1}<2 && next_idistance_k > 3*distance_{k,k+1} && last_idistance_k
>2*distance_{k,k+1})
      || last_idistance_k >30) && next_idistance_k > 5)
{
   if (MAPD_{k,k+1}<1)
          QP_k = QP_k-3
   else if (MAPD_{k,k+1}<4)
          QP_k = QP_k-2
```

else if ($MAPD_{k,k+1}$<10)

$QP_k = QP_k$-1

}

在以上实施例中，没有考虑先前图像或邻近的过去图像是否由于以上规则而已经更新了其QP。这可能导致不必要地更新QP值，在率失真（RD）性能方面这可能是不希望的。为此，与图像类型无关地将参数$last_idistance_k$更新为等于最后经过QP调节的图像的值。

类似地，可以使用宏块/块方差、均值和边缘统计特性来确定局部编码参数。例如，对于位置(i, j)处的宏块的拉格朗日算子λ的选择，可以考虑以下规则：

if ($cur_slice_type_k$!= B_Slice)

{

if ($contains_edges(k, i, j)$)

$$\lambda = 0.5 \times 2^{\frac{QP-12}{3}}$$

else if ($cur_slice_type_k$ == I_Slice)

{

if ($MBvariance(k, i, j)$<15 || $MBvariance(k, i, j)$>60)

$$\lambda = 0.58 \times 2^{\frac{QP-12}{3}}$$

else if ($MBvariance(k, i, j)$>=15 && $MBvariance(k, i, j)$<=40)

$$\lambda = 0.65 \times 2^{\frac{QP-12}{3}}$$

else

$$\lambda = 0.60 \times 2^{\frac{QP-12}{3}}$$

}

else // *cur_slice_type*$_k$ == P_Slice

{

if (MBvariance(k, i, j)<15 || *MBvariance(k, i, j)*>60)

$$\lambda = 0.60 \times 2^{\frac{QP-12}{3}}$$

else if (MBvariance(k, i, j)>=15 && *MBvariance(k, i, j)*<=40)

$$\lambda = 0.70 \times 2^{\frac{QP-12}{3}}$$

else

$$\lambda = 0.65 \times 2^{\frac{QP-12}{3}}$$

}

}

else

{

bscale=max(2.00, min(4.00, (*QP* / 6.0)));

if (*contains_edges(k, i, j)*)

$$\lambda = 0.65 \times bscale \times 2^{\frac{QP-12}{3}}$$

else

{

if (MBvariance(k, i, j)<15 || *MBvariance(k, i, j)*>60)

$$\lambda = 0.68 \times bscale \times 2^{\frac{QP-12}{3}}$$

else if (MBvariance(k, i, j)>=15 && *MBvariance(k, i, j)*<=40)

$$\lambda = 0.72 \times bscale \times 2^{\frac{QP-12}{3}}$$

else

$$\lambda = 0.70 \times 2^{\frac{QP-12}{3}}$$

}

if (nal_reference_idc == 1)

$$\lambda = 0.80 \times \lambda$$

}

对于用于残差编码的量化值或阈值处理系数的选择，可以作出类似的决策。更具体地说，如下执行 H.264 中系数 *W* 的量化：

$$Z = \mathrm{int}\left(\left\{|W| + f \times (1 << q_bits)\right\} >> qbits\right) \cdot \mathrm{sgn}(W)$$

其中 *Z* 是最终量化值，而 *q_bits* 则基于当前宏块的量化器 QP。*fx(1<<q_bits)* 项用作量化过程的舍入项，该项“最佳地”应该等于 1/2x(1<<q_bits)。现在转到图 2，一般用附图标记 200 来表示变换和量化期间死区处理的影响。在图 2 中，零周围的区间被称为死区。用以下两个参数来表征死区量化器：零箱（bin）宽度(2s-2f)和箱外宽度(s)，如图 2 所示。常常把通过 f 执行的死区优化，用作获得良好率失真性能的有效方法。然而，众所周知，该过程期间死区的引入（即 *f* 项的减小）可能通常使附加的比特率减小，同时对质量产生小的影响。对于缺少较高分辨率素材的细节（及胶片颗粒信息）的较低分辨率内容，尤其如此。虽然可以使用 *f*=1/2，但这也可能造成比特率的较大幅度增加，并可能损害 RD 评价方面的性能。

考虑到不同的频率比其它更为重要，一种备选方法将是考虑该观测结果、以提高性能。不是对所有变换系数都使用固定的 *f* 值，而是实质上用矩阵方法来考虑不同值，其中基于频率位置来选择每个死区参数。因此，现在可以如下计算 Z：

$$Z = \mathrm{int}\left(\left\{|W| + f(i,j)\times(1 << q_bits)\right\} >> qbits\right)\cdot \mathrm{sgn}(W)$$

其中 i 和 j 对应于块变换系数内的当前列或行。阵列 f 现在可以取决于片或宏块的类型，并且也可取决于当前块的纹理特征（方差或边缘信息）。例如，如果块包含边缘、或具有低方差特征，则重要的是不引入由于死区处理所产生的更多人为产物，因为这些人为产物将更加可见。另一方面，具有高空间活性的块可以掩蔽更多的人为产物，并且可以在不对质量产生大的影响的情况下增加死区处理。也可以取决于当前块是否为将来图像中的块提供任何有用信息（即，当前块内的任何像素是否用于预测其它像素），来改变死区处理。

例如，如果使用 4×4 变换，则可以使用以下死区处理矩阵：

```
if (cur_slice_type_k == I_Slice)
{
   if (MBvariance(k,i,j)<15 || MBvariance(k,i,j)>60)
```

$$f = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/3 \\ 1/2 & 1/2 & 1/2 & 1/3 \\ 1/2 & 1/2 & 1/3 & 1/4 \\ 1/3 & 1/3 & 1/4 & 1/5 \end{bmatrix}$$

```
   else if (MBvariance(k,i,j)>=15&&MBvariance(k,i,j)<=40||contains_edges(k,i,j))
```

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

```
   else
```

$$f = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 & 1/3 \\ 1/2 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \end{bmatrix}$$

```
}
```

```
else if (cur_slice_typek == P_Slice)
{
   if (MBvariance(k,i,j)<15 || MBvariance(k,i,j)>60)
```

$$f=\begin{bmatrix} 1/3 & 2/7 & 4/15 & 2/9 \\ 2/7 & 4/15 & 2/9 & 1/6 \\ 4/15 & 2/9 & 1/6 & 1/7 \\ 2/9 & 1/6 & 1/7 & 2/15 \end{bmatrix}$$

```
   else if (MBvariance(k,i,j)>=15&&MBvariance(k,i,j)<=40||contains_edges(k,i,j))
```

$$f=\begin{bmatrix} 1/2 & 1/3 & 2/7 & 2/9 \\ 1/3 & 4/15 & 2/9 & 1/6 \\ 2/7 & 2/8 & 1/6 & 1/7 \\ 2/9 & 1/6 & 1/7 & 2/15 \end{bmatrix}$$

```
   else
```

$$f=\begin{bmatrix} 2/5 & 1/3 & 4/15 & 2/9 \\ 1/3 & 4/15 & 2/9 & 1/6 \\ 4/15 & 2/9 & 1/6 & 1/7 \\ 2/9 & 1/6 & 1/7 & 2/15 \end{bmatrix}$$

```
}
else // B_slices
{
```

$$f=\begin{bmatrix} 1/4 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/7 \\ 1/6 & 1/6 & 1/7 & 1/7 \\ 1/6 & 1/7 & 1/7 & 1/7 \end{bmatrix}$$

```
}
```

在某些条件下，可能编码器不可能利用将来帧来执行时间分析。在这种情况下，可以通过只考虑先前已编码的图像、并假定将来图像具有类似的时间特征，来执行时间分析。例如，如果当前图像具有高度相似性（例如$MAPD_{k,k-1}$小），则假定待编码的下一个图像的相似性（$MAPD_{k,k+1}$）也将小。

因而，在用*(k, k-1)*代替所有索引*(k, k+1)*的同时，可以基于已经可得的信息来修改编码参数。

现在转到图 3，一般用附图标记 300 来表示视频编码器。视频编码器 300 的输入以信号通信方式连接到预分析块 310 的输入。预分析块 310 包括多个帧延迟器 312，这多个帧延迟器 312 以信号通信方式互连，以致这多个帧延迟器 312 的每一个都顺序地串联，并且全都通过并行信号路径而并联。该并行信号路径也以信号通信方式连接到时间分析器 315 的输入。所串联连接的、距编码器 300 的输入最远的最后一个帧延迟器 312 的输出，以信号通信方式连接到空间分析器 320 的输入、第一求和点 325 的反相输入、运动补偿器 375 的第一输入、以及运动估计器/模式决策块 370 的第一输入。第一求和点 325 的输出以信号通信方式连接到变换器 330 的输入。变换器 330 的输出以信号通信方式连接到量化器 335 的第一输入。量化器 335 的输出以信号通信方式连接到可变长编码器 340 的第一输入以及逆量化器 345 的输入。可变长编码器 340 的输出是视频编码器 300 的外部可用输出。逆量化器 345 的输出以信号通信方式连接到逆变换器 350 的输入。逆变换器 350 的输出以信号通信方式连接到第二求和点 355 的非反相第一输入。第二求和点 355 的输出以信号通信方式连接到环路滤波器 360 的第一输入。环路滤波器 360 的输出以信号通信方式连接到图像参考存储器 365 的第一输入。图像参考存储器 365 的输出以信号通信方式连接到运动估计器/模式决策块 370 的第二输入以及运动补偿器 375 的第二输入。运动估计器/模式决策块 370 的第一输出以信号通信方式连接到可变长编码器 340 的第二输入。运动估计器/模式决策块 370 的第二输出以信号通信方式连接到运动补偿器 375 的第三输入。运动补偿器 375 的输出以信号通信方式连接到第一求和点 325 的非反相输入以及第二求和点 355 的非反相第二输入。空间分析器 320 的第一输出以信号通信方式连接到量化器 335 的第二输入。空间分析器 320 的第二输出以信号通信方式连接到环路滤波器 360 的第二

输入、运动估计器/模式决策块370的第三输入、以及第一求和点325的非反相输入。时间分析器 315 的第一输出以信号通信方式连接到量化器 335 的第二输入。时间分析器 315 的第二输出以信号通信方式连接到运动估计器/模式决策块370的第四输入。时间分析器315的第三输出以信号通信方式连接到环路滤波器360的第三输入以及图像参考存储器365的第二输入。

在时间分析步骤期间考虑一组图像，这样决定几个参数，包括：片类型决策、GOP结构、加权参数（通过运动估计器/模式决策块370）、量化值和死区处理（通过量化器335）、参考顺序和处理（图像参考存储器365）、图像编码排序、帧/场图像级自适应决策、乃至去块参数（环路滤波器360）。类似地，对每个编码后的帧执行空间分析，这样可以类似地影响量化和死区处理（量化器335），拉格朗日参数及片类型决策（运动估计器/模式决策块 370），帧间/帧内模式决策，帧/场图像级和宏块级自适应决策以及去块处理（环路滤波器360）。

现在转到图 4，一般用附图标记 400 来表示用于对视频信号数据进行编码的示例过程。该过程能够通过在每次迭代中收集并更新所需统计特性，来多次分析或编码相同的位流。这些统计特性用于随后的每一遍编码中，以便在给定的视频特征或用户要求下、通过修改编码器参数来提高编码性能。具体而言，要利用大小为（N, M）的窗口、对k帧（即把未存储的图像排除在外）编码L遍（在此也称为“重复”和“迭代”），其中N是窗口内的总帧数，且M是相邻窗口之间的重叠帧数。用变量frm来索引待编码的帧，而用变量w_{index}来索引窗口内的当前位置。

该过程包括开始块405，它将控制传递给功能块410。功能块410将序列大小设为k，将重复次数设为L，将变量i设为 0，并将控制传递给功能块415。功能块 415 将窗口大小设为N，将重叠大小设为M，将变量frm设为零（0），并将控制传递给功能块420。功能块420将变量w_{index}设为零（0），并

将控制传递给功能块 425。因而，应该理解，对于每一遍编码都对窗口参数进行初始化。这允许使用不同的窗口大小，乃至基于先前的分析步骤来修改它们（例如，如果检测到场景变化，则可以据此调节*N*和*M*，以便只包括完整的场景）。

功能块 425 通过考虑窗口内的所有*N*帧，对待处理的每个窗口执行时间分析，产生时间统计特性（$tstat_{i,frm...frm+N-1}$），并可选地利用当前统计特性、来修改或改进来自前面几遍编码或前面编码步骤的统计特性。然后，功能块 425 将控制传递给功能块 430。功能块 430 对具有索引*frm*（当前窗口内的w_{index}）的帧执行空间分析，直到不再满足条件$w_{index}<N-M$为止，并将控制传递给功能块 435。功能块 435 基于时间和空间分析结果来对这些帧进行编码，产生/收集在需要进行多遍编码时可使用的编码器统计特性，并将控制传递给功能块 440。

功能块 440 递增变量*frm*和w_{index}的值，并将控制传递给决策块 445。决策块 445 判定变量*frm*是否小于*k*。

如果变量*frm*小于*k*，则控制被传递给决策块 450，决策块 450 判定w_{index}是否小于(*N-M*)。否则，如果变量*frm*不小于*k*，则控制被传递给决策块 455，决策块 455 判定*i*是否小于*L*。

如果w_{index}小于(*N-M*)，则控制被传递给功能块 430。否则，如果w_{index}不小于(*N-M*)，则控制被传回给功能块 420。

如果 *i* 不小于 *L*，则控制被传回给功能块 415。否则，如果 *i* 小于 *L*，则控制被传递给结束块 460。

现在将描述按照本发明的各说明性实施例的本发明的许多相关优点/特征中的一些优点/特征。例如，一个优点/特征是，提供一种基于待编码内容的受约束但重叠的窗口来执行视频分析、且利用该信息来修改编码参

数的编码设备和方法。另一个优点/特征是，在视频分析中使用空间－时间分析。又一个优点/特征是，对于视频分析考虑初遍编码。而且，另一个优点/特征是，在视频分析中联合考虑空间－时间分析和初遍编码。而且，另一个优点/特征是，利用图像编码类型、边缘、均值和方差信息的至少之一，来执行空间分析、拉格朗日参数修改、量化和死区处理。又一个优点/特征是，利用绝对差和方差来修改量化参数。另外，另一个优点/特征是，所执行的视频分析只考虑先前已编码的图像。此外，另一个优点/特征是，所执行的视频分析用于决定几种编码参数的至少之一，包括但不限于：片类型决策，GOP 和图像编码结构及顺序，加权参数，量化值和死区处理，拉格朗日参数，参考数，参数顺序和处理，帧/场图像和宏块决策，去块参数，帧间块大小决策，帧内空间预测，以及直接模式。而且，另一个优点/特征是，可以通过考虑先前所产生的统计特性以修改编码参数或分析统计特性，利用多次迭代来执行视频分析。而且，另一个优点/特征是，基于先前所产生的分析统计特性来修改窗口大小和重叠窗口区。

相关领域的普通技术人员根据在此给出的教导可以容易地弄清本发明的这些及其它特征和优点。应该理解，可以用各种形式的硬件、软件、固件、专用处理器或它们的组合，来实施本发明的教导。

更为优选地，把本发明的教导实施为硬件和软件的组合。而且，优选地把软件实施为确实可在程序存储单元上具体实施的应用程序。该应用程序可以被上载到包括任何合适架构的机器上，并被该机器所执行。优选地，在具有硬件如一个或多个中央处理单元（CPU）、随机存取存储器（RAM）和输入/输出（I/O）接口的计算机平台上，实现该机器。该计算机平台也可包括操作系统和微指令码。在此描述的各种过程和功能可以是，可由 CPU 执行的微指令码一部分、或应用程序一部分、或它们的任何组合。另外，其它各种外围单元如附加数据存储单元和打印单元，可以连接到该计算机平台。

进一步应该理解，因为附图中所描述的一些组成系统部件和方法优选地是用软件来实现的，所以系统部件或过程功能块之间的实际连接可以随本发明的编程方式而有所不同。给定在此的教导，相关领域普通技术人员将能够预期本发明的这些及类似实施或配置。

虽然在此参考附图描述了说明性实施例，但是应该理解，本发明不限于这些具体实施例，并且相关领域普通技术人员可以在不脱离本发明的范围或精神的情况下进行各种变动和修改。所有这种变动和修改都包括在所附的权利要求所阐述的本发明的范围内。

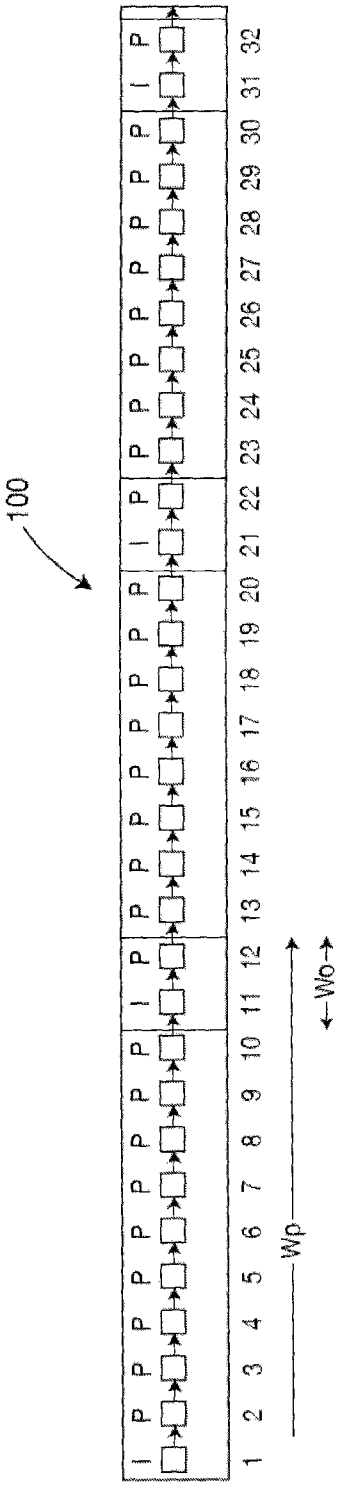

100
I P P P P P P P P P P I P P P P P P P P P P I P
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
Wp
Wo

图 1

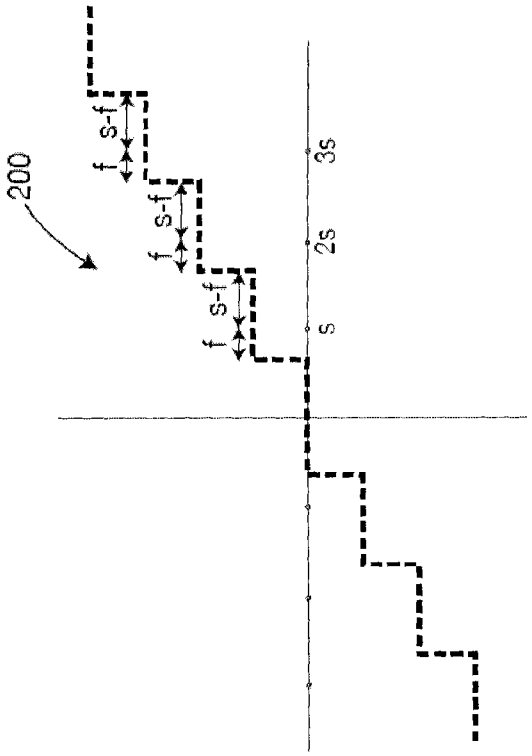

200
f s-f
f s-f
f s-f
s
2s
3s

图 2

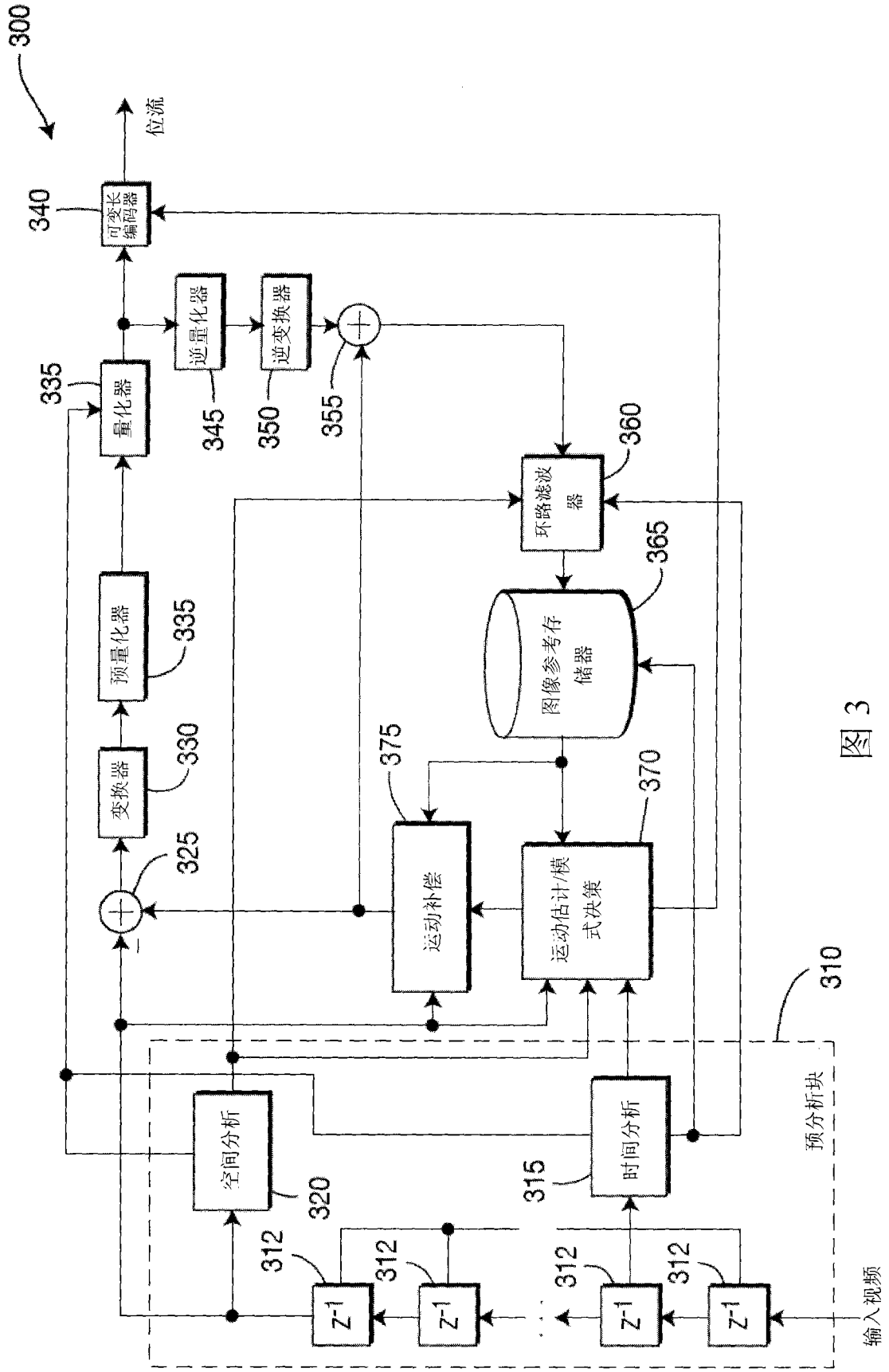

300
位流
340
可变长编码器
335
量化器
345
逆量化器
350
逆变换器
355
360
环路滤波器
365
图像参考存储器
335
预量化器
330
变换器
325
375
运动补偿
370
运动估计/模式决策
310
预分析块
320
空间分析
315
时间分析
312
Z-1
输入视频

图 3

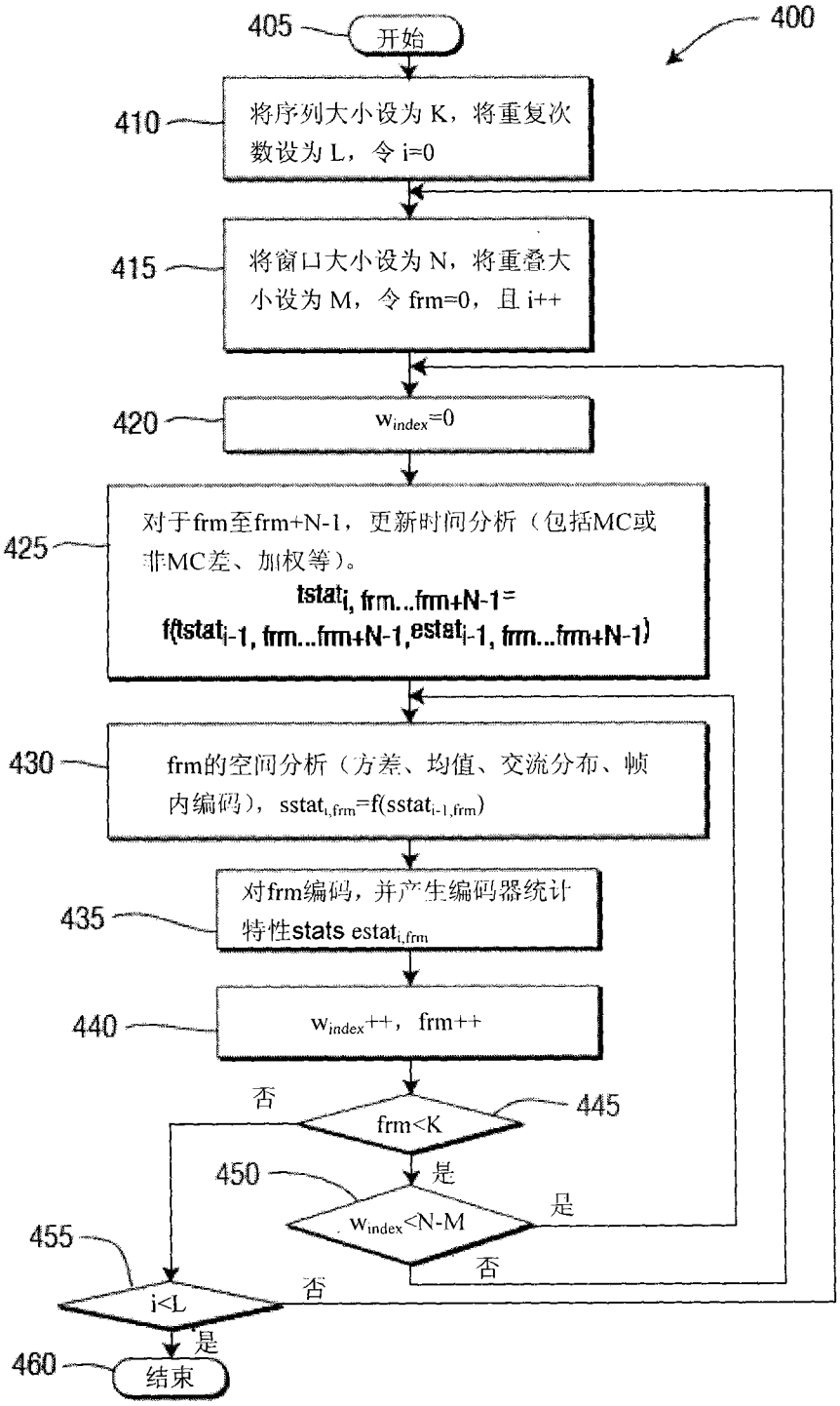

400
405
开始
410
将序列大小设为 K，将重复次数设为 L，令 i=0
415
将窗口大小设为 N，将重叠大小设为 M，令 frm=0，且 i++
420
w_{index}=0
425
对于frm至frm+N-1，更新时间分析（包括MC或非MC差、加权等）。
$tstat_{i, frm...frm+N-1}$=
$f(tstat_{i-1, frm...frm+N-1}, estat_{i-1, frm...frm+N-1})$
430
frm的空间分析（方差、均值、交流分布、帧内编码），$sstat_{i,frm}=f(sstat_{i-1,frm})$
435
对frm编码，并产生编码器统计特性stats $estat_{i,frm}$
440
w_{index}++，frm++
445
frm<K
否
是
450
w_{index}<N-M
是
否
455
i<L
否
是
460
结束

图 4