

[12] 发明专利申请公开说明书

[21] 申请号 99810653.4

[43] 公开日 2001 年 10 月 10 日

[11] 公开号 CN 1317125A

[22] 申请日 1999.8.2 [21] 申请号 99810653.4

[30] 优先权

[32] 1998.8.7 [33] DE [31] 19835845.8

[86] 国际申请 PCT/DE99/02406 1999.8.2

[87] 国际公布 WO00/08601 德 2000.2.17

[85] 进入国家阶段日期 2001.3.7

[71] 申请人 英芬能技术公司

地址 德国慕尼黑

[72] 发明人 哲真·番德尔 格罗·巴斯

诺伯特·欧特尔

[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

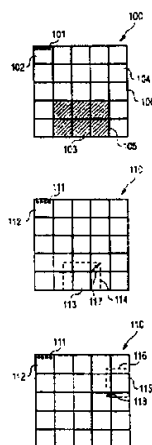
代理人 杨国旭

权利要求书 2 页 说明书 8 页 附图页数 3 页

[54] 发明名称 对具有象素的数字图像进行运动估算的方法和装置

[57] 摘要

对具有象素的数字图像进行运动估算的方法和装置,其中,象素被分组到图象块。象素被分组到至少第一图象区域和第二图象区域。第一运动估算在第一检索区域被实现,以便确定一个第一运动矢量。此外,第二运动估算在第二检索区域被实现,以便确定一个第二运动矢量。第一检索区域和第二检索区域大小不同。



ISSN 1008-4274

权 利 要 求 书

1. 一种对具有象素的数字图像进行运动估算的方法，
其中象素被分组到图象块；

5 其中象素被分组，以便至少形成第一图象区域和第二图象区域；

在第一检索区域，至少为在第一图象区域中的第一图象块实行第一运动估算，以便确定第一个运动矢量，利用它，通过比较在前面图象中的第一图象块或后面图象中的第一图象块，第一图象块的移动被描述；

10 其中，在第二检索区域，至少为在第二图象区域的第二图象块实行第二运动估算，以便确定第二个运动矢量，利用它，通过比较在前面的原来图象中的第二图象块或通过比较在后面的后继图象中的第二图象块，第二图象块的移动被确定；

其中，第一检索区域和第二检索区域大小不同。

15 2. 按照权利要求 1 所述的方法，

其特征在于：第一检索区域和/或第二探测区大小是预定图象质量的函数，通过它第一图象块和/或第二图象块被编码。

3. 按照权利要求 2 所述的方法，

20 其特征在于：第一检索区域和/或第二探测区的大小按照量化参数的函数而变化，利用它，第一图象块和/或第二图象块被量化。

4. 按照权利要求 1-3 所述的方法，被用于数字图象的编码。

5. 按照权利要求 4 所述的方法，

其中运动矢量变长度编码被实现，

25 其中，若干个不同的表被用于变长编码，所述表中存储了变长编码的代码。

6. 按照权利要求 5 所述的方法，其特征在于：该表匹配到运动矢量的最大长度。

7. 一种对具有象素的数字图像进行运动估算的装置，
具有一个处理器，它被设定以便实行下述功能：

象素被分组到图象块，

象素被分组，以便至少形成第一图象区域和第二图象区域，

在第一检索区域，至少为第一图象区域中的第一图象块实行第一运动估算，以便确定第一运动矢量，利用它，通过比较在前面图象中的第一图象块或后面图象中的第一图象块，第一图象块的移动被描述；

进一步，在第二检索区域，至少为第二图象区域中的第二图象块实行第二运动估算，以便确定第二运动矢量，利用它，通过比较在前面的原来图象中的第二图象块或通过比较在后面的后继图象中的第二图象块，第二图象块的移动被确定；

10 第一检索区域和第二检索区域大小不同。

8、按照权利要求 7 所述的装置，

其特征在于：该处理器被如此设置，以使第一检索区域和/或第二检索区域的大小是预先确定的图象质量的函数，利用它，第一图象块和/或第二图象块被编码。

15 9、按照权利要求 8 所述的装置，

其特征在于：该处理器被如此设置，以使第一检索区域和/或第二检索区域大小按照量化参数的函数变化，通过它第一图象块和/或第二图象块被量化。

10、按照权利要求 7-9 所述的装置，它们被用于图象编码设备。

20 11、按照权利要求 7-9 所述的装置，它们被用于图象编码设备，

其特征在于：该处理器被如此设置，以使

运动矢量变长编码被实现，

若干储存的不同表用于变长编码，所述表中存储有变长编码的代码。

25 12、按照权利要求 11 所述的装置，

其特征在于：该处理器被如此设置，以使该表匹配到运动矢量的最大长度。

说 明 书

对具有象素的数字图像
进行运动估算的方法和装置

5

本发明涉及对具有象素的数字图像进行的运动估算。

这样的方法不同于参考文献[1]。

在参考文献[1]的运动估算方法中，将要进行运动估算的数字块中的象素被分组成为若干图象块。

10 对于在图象中的每一个图象块，在大小可以预先确定的检索区域内进行尝试，以确定图象块中一个与检索区域同样大小的编码信息最相似的区域，该编码信息包括在进行运动估算的图象块中。

在下文中，名词“编码信息”意味着与一个象素相关的亮度信息(亮度值)或颜色信息(色度值)。

15 为此，在先前图象中，基于图象块在先前图象中的位置，对于大小可以预先确定的区域(检索区域)中的每一个位置，具有相应大小的包括与图象块中的象素数量相同的象素的区域被形成；并且，将要进行运动估算的图象块的编码信息和在先前图象中的不同区域的编码信息之间的绝对偏差和或偏差的平方和被形成。最佳匹配的区域，这也就是说和值最小的区域，被视作匹配图象块，并且在先前图象的“最佳”
20 区域和该图象块之间图象块位置的移动被确定。这个移动被称为运动矢量。

本发明针对上述问题，提供了一种可以减少为编码数字图象的运动估算的运动矢量所需总位数的方法和装置。

25

具有独立权利要求的特征的方法和装置，解决了这个问题。

在对具有象素的数字图像进行运动估算的方法中，象素被分组到图象块中。象素至少被分组到第一图象区域和第二图象区域。在第一检索区域，至少为在第一图象区域的第一图象块实行第一运动估算，以便确定第一个运动矢量，利用它，通过比较在先前图象中的第一图

象块或后面的图象中的第一图象块，第一图象块的移动被描述。进一步，在第二检索区域，至少为在第二图象区域中的第二图象块实行第二运动估算，以便确定第二个运动矢量，利用它，通过比较在前面的原来图象中的第二图象块或通过比较在后面的后继图象中的第二图象块，第二图象块的移动被确定。第一检索区域和第二检索区域大小不同。

对具有象素的数字图像进行运动估算的装置，具有一个处理器，它被如此设定以便执行下述步骤：

象素被分组到图象块，

象素被分组以便至少形成一个第一图象区域和第二图象区域，

在第一检索区域，至少为在第一图象区域中的第一图象块实行第一运动估算，以便确定第一个运动矢量，利用它，通过比较在先前图象中的第一图象块或后面的图象中的第一图象块，第一图象块的移动被描述；

进一步，在第二检索区域，至少为在第二图象区域中的第二图象块实行第二运动估算，以便确定第二个运动矢量，利用它，通过比较在前面的原来图象中的第二图象块或通过比较在后面的后继图象中的第二图象块，第二图象块的移动被确定；

第一检索区域和第二检索区域大小不同。

本发明可以减少为压缩视频数据传输而要求的数据量，因为运动矢量的大小可自适应地符合质量的要求，所以即使在非常小的检索区域，也没有明显的图象质量的主观印象的下降，例如，在只要求低质量的区域中。检索区域中运动矢量的最大尺寸也因此相对小，这导致运动矢量编码的位数被减小。

从本发明可以明显的看出：不同大小的检索区域被用于该图象区域中的图象块的运动估算，本发明可以根据质量灵活的减少运动矢量编码所要求的数据量。

从从属权利要求中可得出本发明的进一步改进。

一个改进规定第一检索区域和/或第二检索区域的大小是预先确定

的图象质量的函数，利用它，第一图象块和/或第二图象块被编码。

以这种方式，一种限制检索区域的方法被确定，它根据要求的图象质量减少需求的数据量。

5 在一个实施例中，一个十分简单的确定检索区域大小的标准是一个量化参数，通过它第一图象块和/或第二图象块被量化。

进一步的改进规定一定数量的表格，存储不同长度的编码的代码，它们用于运动矢量的不同的长度编码，这样导致进一步减少视频数据的传输所要求的数据量。

10 本发明的一个实施例将在下面详细介绍，并在下面的图中图示，其中：

图 1a 到 1c 显示了图象和先前图象的略图，其中，图示了本发明的基本原理；

图 2 显示了具有两个计算机、一个摄像机和一个显示器的装置，通过它视频数据被编码、传送、解码和显示；

15 图 3 显示了基于块的数字图象的编码设备的略图。

图 2 显示了包括两台计算机 202，208 和一个摄像机 201 的装置，显示了图象的编码、视频数据的传输和图象的解码。

20 摄像机 201 通过连线 19 连接到第一计算机 202。摄像机 201 传送它拍摄的图象 204 到第一计算机 202。第一计算机 202 拥有第一处理器 203，它通过总线 218 连接到帧存储器 205。在第一计算机 202 的第一处理器 203 执行一种图象编码方法。以这种方式，编码的视频数据 206 从第一计算机 202 通过通讯链路 207，最好是电缆或是无线路径，被传送到第二计算机 208。第二计算机 208 拥有第二处理器 209，它通过总线 210 连接到帧存储器 211。通过第二处理器 209，图象解码方法
25 被实行。

第一计算机 202 和第二计算机 208 具有各自的显示器 212 或 213，通过它们视频数据 204 被显示。输入单元，最好是键盘 214 或 215 和计算机鼠标 216 或 217，被分别提供给第一计算机 202 和第二计算机 208。

通过连线 219 从摄像机 201 传送到第一计算机 202 的视频数据 204 是时域数据, 而通过通讯链路 207, 从第一计算机 202 传送到第二计算机 208 的数据 206 是频域视频数据。

编码的视频数据被显示在显示器 213。

5 图 3 显示了根据 H.263 标准(参见[5])执行基于块的图象编码方法的装置的略图。

10 将编码的并具有被顺序数字化的图象的视频数据流提供到图象编码单元 301。数字化的图象被再细分为宏块 302, 每一个宏块包括 16x16 象素。宏块 302 包括 4 个图象块 303, 304, 305 和 306, 每一个图象块包括 8x8 象素, 它们的照度值(亮度值)被指定。此外, 每一个宏块 302 包括具有被分配给该象素(色信息, 色饱和度)的色度值的 2 个色度块 307 和 308。

图象中的块包括亮度值(=亮度), 第一色度值和第二色度值。这时, 亮度值, 第一色度值和第二色度值被称作颜色值。

15 图象块被提供给变换编码单元 309。在差分图象编码时, 从先前图象的图象块编码的值被减去在那时编码的图象块的值, 只有差分构成信息 310 被提供给变换编码单元(离散余弦变换, DCT)309。为此, 宏块 302 通过链路 334 被传信到运动估算单元 329。在变换编码单元 309, 将被编码的图象块或差分图象块的频谱系数 311 被形成, 并被提
20 供给量化单元 312。

量化的频谱系数 313 在反馈路径上被提供给扫描单元 314 和反量化单元 315。使用扫描方法, 例如"z 字形"扫描方法, 在为此目的而提供的熵编码单元 316 中, 对扫描的频谱系数 332 执行熵编码。熵编码的频谱系数作为编码的视频数据 317 通过一个信道, 最好是电缆或无线
25 通路, 被传送到解码器。

量化的频谱系数 313 的反向量化在反向量化单元 315 中被实行。以这种方法得到的频谱系数 318 被提供给反向变换编码单元 319(反向离散余弦变换, IDCT)。重建的编码值(和差分编码值)320 以差分构成模式被提供给加法器 321。一旦运动补偿被进行, 加法器 321 也接收

从先前图象获得的图象块的编码值。加法器 321 用于形成储存在帧存储器 323 中的重建的图象块 322。

重建的图象块 322 的色度值 324 从帧存储器 323 被提供到运动补偿单元 325。对于亮度值 326，在提供的内插单元 327 中实现内插。该内插最好用于使包括在各个图象块中的亮度值扩大四倍。亮度值 328 不仅被提供到运动补偿单元 325，也被提供到运动估算单元 329。运动估算单元 329，通过链路 334 分别接收将被编码的宏块(16x16 象素)的图象块。在运动估算单元 329，根据内插的亮度值("基于半象素的运动估算")进行运动估算。

运动估算的结果是表示从先前图象选出的宏块到将被编码的宏块 302 的位置移动的运动矢量 330。

与由运动估算单元 329 确定的宏块相关的亮度信息和色度信息通过运动矢量 330 被改变，被减去宏块 302 的编码值(参见数据路径 231)

运动估算得到了具有 2 个运动矢量分量的运动矢量 330，即，分别沿着第一方向 x 和第二方向 y 的第一运动矢量分量 BV_x ，和第二运动矢量分量 BV_y 。

$$BV = \begin{pmatrix} BV_x \\ BV_y \end{pmatrix}$$

运动矢量 330 被分配到图象块。

图 3 中的图象编码单元为所有图象块和图象宏块都提供运动矢量 330。

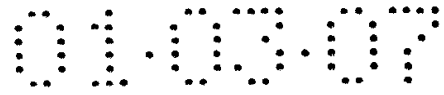
图 1a 显示了将用图 3 所示设备编码的数字图象 100。

数字图象 100 具有被分配有编码信息的象素。

象素 101 被分组到图象块 102。图象块 102 被分组到第一图象区域 105 和第二图象区域 106。

在下文中，假设在第一图象区域 105 的质量要求比在第二图象区域 106 的质量要求更严格。

为第一图象区域 105 中的第一图象块 103 实行运动估算。为此，



在先前图象和/或在后继图象 110 中，定义了第一检索区域 114。

基于形状、大小同于第一图象块的起始区 113，在每种情况下，以下误差 E 被确定，它变化是一个象素或几分之一或多个象素(例如用半象素(半象素运动估算))，利用它，在各种情况下，起始区 113 被移

$$E = \sum_{i=1}^n \sum_{j=1}^m (x_{i,j} - y_{i,j})^2,$$

5 动：

其中

-i, j 是连续的指数，

-n 是在第一图象块中沿第一方向的象素数，

-m 是在第一图象块中沿第二方向的象素数，

10 -x_{i, j} 是在第一图象块中在位置 i, j 的象素的编码信息，

-y_{i, j} 是通过相应的运动矢量移动的，在先前图象相应位置的象素编码信息。

15 在先前图象，对于每一个移动 110，误差 E 被计算，并且从该移动(=运动矢量)的图象块中，误差 E 具有最小值的、一个被选为与第一图象块 103 最相似的一个。

在本实施例中，在每种情况下，检索区域在水平垂直方向跨越 4 象素间隔，大约是相应于在先前图象 110 第一图象区域的第一图象块位置的起始位置 113。因此，这时将被编码的第一运动矢量 117 的最大尺寸是 $4\sqrt{2}$ 象素间隔(参见图 1b)。

20 图 1c 显示了在第二图象区域 106 中的第二图象块 104 的第二运动估算。该运动估算的基本过程也被描述为如上述的第二运动估算的过程。

对于第二运动估算，第二检索区域 116 更小，因为在第二图象区域 106 的图象质量要求不如第一图象区域 105 的严格。

25 因此，第二检索区域 116 的大小只是从起始位置 115 向每一个方向的 2 个象素 116。将为第二图象块 104 编码的第二运动矢量 118 的最大尺寸是 $2\sqrt{2}$ 。

从这个例子可见，编码第二运动矢量 118 需要的计算量比编码第一运动矢量 117 显著减小。

根据这个例子，在该实施例中的图象块的探测区大小取决于量化参数，它显示了用于先前图象 100 编码的量化步骤。

5 探测区大小 S 用下式获得：

$$S=15-QP/2$$

其中

-S 是检索区域大小，

-QP 是量化参数。

10 量化参数 QP 是包括在 H.263 的一般标题数据中的因数，并用作为量化初始值。

为图象块而确定的检索区域大小 S 随量化参数 QP 的变小而增大，这相应于高的图象质量。

15 包括具有不同长度范围的不同长度的运动矢量的编码的若干个表，用于运动矢量的可变长编码。

量化参数 QP 用于选择一个变长度编码表，其中变长度编码域值与检索区域大小 S 匹配，因此达到运动矢量的最大长度。

上述示例性实施例的一些变形被描述如下。

运动估算的类型以及相应的类似形成测度方法，与本发明无关。

20 例如，以下公式也可形成误差 E：

$$E = \sum_{i=1}^n \sum_{j=1}^m |X_{i,j} - Y_{i,j}|,$$

25 它进一步显示了，为了进一步减少要求的数据量，在许多情况下，只传送运动矢量已经足够了，而不传送在运动补偿差分图象的形成时产生的误差信号。

很明显，在本发明中，不同尺寸的检索区域被用于该图象区域中的图象块的运动估算，其结果是，对于该运动矢量的编码，可按照质量以需要的数据率来减少数据量。

下面的出版物被本文引用:

[1]ITU-T Draft Recommendation H.26, Video Coding for Low Bitrate Communication, May, 1996.

说明书附图

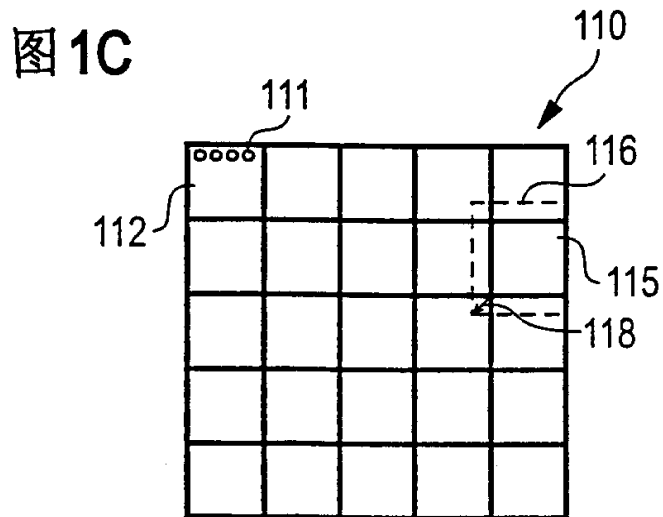
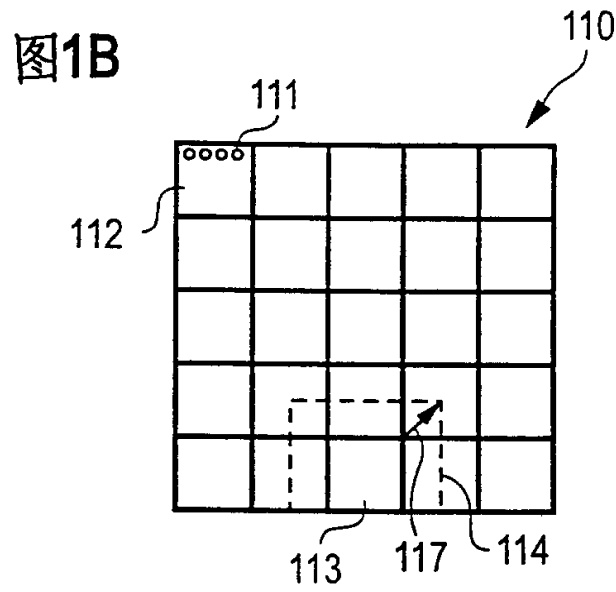
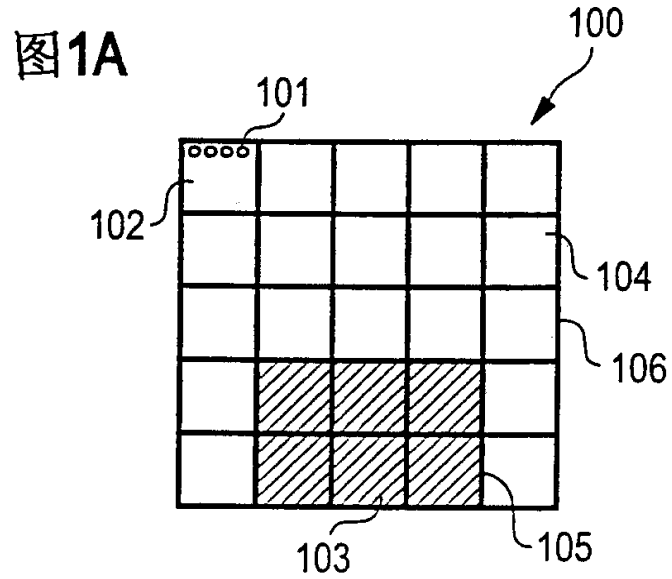


图 2

