[19] 中华人民共和国国家知识产权局

[51] Int. Cl[7]
H04N 7/24
G06T 7/20 H04N 7/01

[12] 发明专利说明书

[21] ZL 专利号 99103294.2

[45] 授权公告日 2004 年 10 月 27 日

[11] 授权公告号 CN 1173571C

[22] 申请日 1999.3.31 [21] 申请号 99103294.2

[30] 优先权

[32] 1998.4.14 [33] EP [31] 98400918.3

[71] 专利权人 汤姆森多媒体公司

地址 法国布洛里

[72] 发明人 弗朗西斯·勒克莱尔

审查员 李意平

[74] 专利代理机构 中科专利商标代理有限责任公司

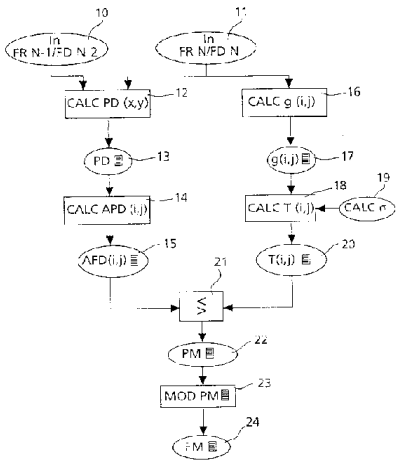
代理人 朱进桂

权利要求书 3 页 说明书 9 页 附图 4 页

[54] 发明名称 检测一序列视频图象中的静止区域的方法

[57] 摘要

本发明公开了一种用来检测一序列视频图象中的静止区域的方法，它包括以下步骤：计算连续帧或具有相同奇偶性的连续场的相应象素(30)之间的象素差异；针对视频帧的象素的象素块(B(i，j))累计象素差异的幅度；计算象素块的阈值 T(i，j)；把阈值 T(i，j)用于相应在象素块的累计象素差异，提供静止区域的一个初步映象图(PM)；通过消除孤立的静止象素块和小组的相邻静止象素块，修正所述初步映象图(PM)。



ISSN1008-4274

1、一种用来检测一序列视频图象中的静止区域的方法，其特征在于，它包括以下步骤:

a)计算连续帧或具有相同奇偶性的连续场的相应象素（30）之间的象素差异;

b)针对视频帧的象素的象素块（B（i，j））累计象素差异的幅度;

c)计算象素块的阈值T(i，j)，其中所述的每个象素块的阈值T(i，j)被匹配于

i)当前场或帧的一个估计的噪声水平（σ），和

ii)一个象素块梯度的量值（g（i，j）），并且所述的阈值水平T(i，j)是通过如下公式计算的:

$T(i, j) = N(\alpha\sigma + \beta g(i, j))$

其中 N 是一个象素块中的象素数，数值α和β是基于如何确定估计的噪声水平σ和象素块g(i，j)所选择;

d)把阈值 T（i，j）用于相应的象素块的累计象素差异，提供静止区域的一个初步映象图（PM);

e)通过消除孤立的静止象素块和小组的相邻静止象素块，修正所述初步映象图（PM)。

2、如权利要求 1 所述的方法，其特征在于，象素块梯度 g(i，j) 的计算方式是从水平 $g_x(i, j)$ 与垂直 $g_y(i, j)$ 梯度组份中选取大的那个，即:

$g(i, j)=MAX[g_x(i, j), g_y(i, j)]$

其中,i，j是一帧或场的一行和一列中的象素块数,水平梯度组份$g_x(i, j)$和垂直梯度组份 $g_y(i, j)$ 被分别计算为水平相邻的象素的亮度差异的绝对值的平均和垂直相邻的象素的亮度差异的绝对值的平均，计算是根据象素块B(i，j)的灰度值L(k，l)对N个象素P(k，l)进行的，即:

$$\begin{cases} g_x(x,y)=\dfrac{1}{N}\sum\limits_{L(k,l)\in B(i,j)}\left|L(k,l)-L(k-1,l)\right| \\ g_y(x,y)=\dfrac{1}{N}\sum\limits_{L(k,l)\in B(i,j)}\left|L(k,l)-L(k,l-1)\right| \end{cases}$$

。

3、如权利要求 2 所述的方法，其特征在于，如果所述帧的噪声水平σ代表一个由灰度水平表示的标准偏差，α被取值为 1.4，β被取值为 1/8。

4、如权利要求 1 所述的方法，其特征在于，把象素块 B(i, j)的累计象素差异 APD(i, j)与相应的阈值 T(i, j)进行比较，而且，为每个象素块 B(i, j)指定一个代表该象素块是静止或非静止的的数值，最好按以下公式进行:

$$\begin{cases} PM(i,j)=1 & \text{如果} APD(i,j)\le T(i,j) \\ PM(i,j)=0 & \text{如果} APD(i,j)>T(i,j) \end{cases}$$

其中 PM(i, j)是静止区域的一个初步映象图的一个输入，数值 1 代表静止象素块，0 代表非静止象素块。

5、如权利要求 1 所述的方法，其特征在于，所述静止区域的初步映象图被如下处理:

a)如果当前象素块 B(i, j)已经在静止区域的初步映象图中被划分为非静止的，即 PM(i, j) = 0，那么，相应的输入在静止区域的最终映象图中也被划分为非静止的，即 FM(i, j) = 0;

b)如果当前象素块 B(i, j)已经被划分为静止的，则进行如下步骤:

i)造成一个以当前象素块（42）为中心的、具有 M_X 个象素块的水平邻近区域（41）;

ii)数出在这个邻近区域内被划分为静止态的象素块的数目 S_X;

iii)分析比值 S_X / M_X 是否小于一个预先设定的阈值 T_X，如果是这样，那么在静止区域的最终映象图中把当前象素块划分为非静止的，即 FM(i, j) = 0;

iv)造成一个以当前象素块为中心的、具有 M_Y 个象素块的垂直邻近区域（40）;

v)数出在这个邻近区域内被划分为静止态的象素块的数目 S_Y;

vi)分析比值 S_Y / M_Y 是否小于一个预先设定的阈值 T_Y，如果是这样，那么在静止区域的最终映象图中把当前象素块划分为非静止的，即 FM(i, j) = 0。

6、如权利要求 5 所述的方法，其特征在于，其中所述的数值 M_X、M_Y、T_X、T_Y 与象素块的尺寸、帧或场的尺寸相匹配。

7、如权利要求 1 所述的方法，其特征在于，它们是应用于一种用于视频图象的运动补偿编码的设备的方法。

8、如权利要求 1 所述的方法，其特征在于，它们是应用于一个用于视频图象的运动补偿编码的设备，尤其用于 50Hz 至 100Hz 上变频、50Hz 至 60Hz 标准转换、或 60Hz 至 50Hz 标准转换的方法。

9、如权利要求 1 所述的方法，其特征在于，它们是应用于一种用于隔行扫描到逐行连续扫描转换的视频信号处理设备的方法。

检测一序列视频图象中的静止区域的方法

技术领域

本发明涉及一种用来检测一序列数字视频图象中的静止区域的方法。这样一种方法通常可以被用于视频信号处理设备中，来检测一序列视频图象的中的运动情况。更具体地说，本发明的一种应用是瞬时视频格式转换领域，这种瞬时视频格式转换可以是运动补偿型的，也可以不是。这包括用于视频信号上变频转换的算法，如在装有 50 Hz 至 100Hz 上变频装置的电视接收机中。这一应用领域包括低端上变频算法和高端上变频算法，在低端上变频算法中，简单的场（半帧）/帧重复被用来生成中间场/帧，在高端上变频算法中，采用运动补偿技术对中间场/帧进行计算。另一个应用领域是视频标准转换，主要是 50 Hz(赫兹)至 60Hz 和 60Hz 至 50Hz 的转换。其它重要的应用领域是 MPEG 视频编码和隔行扫描到逐行连续扫描的转换。

背景技术

有了一序列视频图象中的静止区域的映象图，才可以实现对非运动区域的象素的具体处理。其效果例如可以是这样的，在采用运动补偿技术用于插入图象（转换技术）或用于复构图象（压缩技术）时保持一幅图象中的静止区域的完整分辨率。关于静止区域的信息通常可用于检测视频图象序列中的运动的所有算法，因为零运动的区域就是视频广播场景中按统计结果出现最频繁的区域。

在一个上变频转换或视频标准转换中，输出序列的分辨率可以在非运动区域进行优化，其方式是将运动补偿处理关闭并简单地将源象素复制到输出中。

在 MPEG 视频解码领域，传输被划分为静止区域的象素块中的最小量

的信息，可以节省比特速率。的确，如果运动补偿作用于两帧或具有同样奇偶性的两个场，对于静止象素块来说，只要传输象素块为静止的信息就足够了。这意味着象素块的运动矢量为 0，而且所有的预测误差（这种说法是针对 MPEG 视频标准）应该为 0（它们可能并不正好是 0，但如果确定了象素块是静止的，那么很明显，非零预测误差代表着不应被传输的噪声组份）。

另一个好处来自于一种自然现象，即，人的眼睛对静止区域的缺陷比对运动区域中的缺陷更敏感，这样，有关图象中静止区域的位置的信息显得非常重要，因为只要简单地复制静止区域的源象素，因运动补偿处理所造成的缺陷就可以避免了。

发明内容

本发明的目的在于提供一种用于检测一序列视频图象中的静止区域的可靠的算法。这一目的是通过本发明的一个方面所提供的方法来实现的。

本发明所提供的方法原则上包括 5 个步骤:

第一，计算连续帧或具有相同的奇偶性的连续场相应的象素之间的象素差异;

第二，针对视频图象的象素块，累计象素差异的幅度;

第三，针对所述象素块，计算累计象素差异的阈值;

第四，把所述阈值与累计象素差异进行比较，得到静止象素块的一个初步映象图;

第五，通过消除孤立的静止象素块或小组的相邻静止象素块，对静止区域的初步映象图进行清理。

本发明在隔行扫描源条件下可以取得具有良好可靠性的静止象素块的检测，因为它在具有相同奇偶性的两个场上操作。这样，也避免了由于隔行扫描效果而带来的不需要的扰动。通过累计相应象素的差异值的幅度，有噪声存在条件下静止区域检测的强度得到提高，因为噪声组份的随机变化被消除了。

其中，通过把象素块的阈值匹配于帧或场的一个估计的噪声水平和一个象素块梯度的量，本方法的可靠性进一步提高。考虑象素块梯度的原因

是摄象机即使发生非常小的移动，也会在存在很高空间频率的区域造成巨大的象素差异。所以用来确定阈值的公式包括一个与象素块梯度成正比的“惩罚”项。

此外，给出了考虑到这一因素后计算阈值的一个简单的公式。如果在一个象素块中存在许多较高空间频率组份与/或明显的噪声，那么，将要与象素块的累计象素差异进行比较的相应阈值将会高。有了这些方法，划分一个象素块为静止的或非静止的决定就比正常化了，就是说，可以不受噪声和图象内容的影响。

此外，提供了检测象素块梯度的一个最佳实施例。

此外，给出了用于计算阈值的公式中的比例常数α、β的最佳实施例。这些数值是针对估计噪声水平σ被作为一个标准偏差来计算、并用灰度水平进行表达的情况而优化的。

此外，给出了把阈值用于累计帧差异的一个简单的实施例。

在本发明的另一方面，提供了用来对静止区域的初步映象图进行清理的优选方案。其中，孤立的静止象素块或小组的邻近象素块被从静止/非静止象素块的初步映象图中消除。

此外，提出了可以采用本发明技术方案的重要的设备。

附图说明

对本发明的实施例的详细描述是结合下述附图进行的，其中：

图 1 是用来检测一个视频图象序列中的静止区域的算法的流程图；

图 2a）是用来解释相应象素之间的象素差异的计算方法的连续视频帧的示意图；

图 2b）是用来解释具有相同奇偶性的场中的相应象素之间的象素差异的计算方法的连续视频场的示意图；

图 3 是帧被再分成等大小的象素块的帧的片段；

图 4 是静止区域的一个初步映象图的片段；

图 5 是用于静止区域的一个初步映象图的清理过程的水平与垂直邻近区域的示例；以及

图 6 是与相应于图 4 的一个静止区域最终映象图的片段。

具体实施方式

本发明所提供的用于检测两幅连续视频图象中的静止区域的新方法需要一个前提，即视频图象必须是以数字形式存在的，就是说，模拟源的视频图象必须被进行离散取值和数字化处理。如此生成的数字象素值被存储于帧或场存储器中，这是本领域所公知的内容。如果源视频信号比如是一个标准 NTSC 信号，那么离散取值的频率是 13.5MHz，帧重复速率是每秒 30 帧。一个水平线周期持续 63.5 毫秒。所以，在每个水平线中生成 858 个象素。由于 NTSC 制式中每帧含有 525 个水平扫描线，一帧就有 858 × 525 = 450450 个象素。通常用 8 个比特来表示一个象素，对于帧存储器来说，一个帧存储器必须提供 858 × 262.5 = 225225 字节的存储容量。

本发明的方法适用于逐行扫描和隔行扫描视频图象。其差别将在下面加以详述。

本发明所提供的用来检测连续视频图象中的静止区域的新方法含有几个必须一一顺次进行的步骤。所有的步骤都可以由一个处理单元来执行，该处理单元比如可以含有一个执行一个相应的计算机程序的微处理器。下面结合图 1 对这样的一个程序加以解释。毋需解释的是，该方法还可以用一个专用集成电路来实施，如一个具有用来执行不同的步骤和任务的专用硬件的 ASIC 电路。

在步骤 10 中，对于逐行连续扫描视频系统，第一帧 N－1 被读入一个第一帧存储器；对于隔行扫描视频系统来说，一个奇/偶场被读入一个第一场存储器。在步骤 11 中，对于连续扫描视频系统来说，下一帧 N 被读入一个第二帧存储器；对于隔行扫描视频系统来说，与步骤 10 中被读入的场具有同样的奇偶性的下一个奇/偶场被读入一个第二场存储器。当这些步骤完成的时候，两个平行的分支程序开始被执行。

采用第一种对图象进行模型化的方法，图象可以被模型化为一个代表图象信息的第一信号组份与一个噪声组份的叠合。在非移动图象部分，相应象素间的差异正是场间或帧间噪声差异的样本。在移动图象部分，由一个特定的象素位置所表示的场景目标中的坐标点在一幅图象与下一幅图象中是不同的，所以一个额外的信号差异组份将被加入到噪声差异中。因此，

从均值角度来说，移动部分的差异幅度大于静止部分的差异幅度。这就是为什么本发明的方法中象素差异被与阈值进行比较以区分静止与移动象素的原因。

回到图 1 中的流程图，第一分支含有步骤 12 至 15，在这些步骤中，对两个帧/场的差异进行分析。第二个分支含有步骤 16 至 21，在这些步骤中，对阈值进行计算以决定图象的哪个部分是静止的或非静止的。

先对含有步骤 12 至 15 的第一分支进行说明。这里所提出的方法是基于对连续帧/场的象素内的差异进行分析。所说的差异被定义为占据同一位置的象素的灰度的差异，包括两种情况:

逐行连续输入源条件下的两个连续帧，

隔行输入源条件下的具有相同奇偶性的两个场，即两个场周期分开的两个场。

在步骤 12 中，对存储于帧/场存储器中的两个帧或两个场进行象素差异计算。图 2 的上半部分给出的是两个连续帧 N－1 和 N。其中，帧 N－1 所关心的是前一帧，而帧 N 所关心的是当前帧。这些帧都含有一定数目的线，如图 2 所示。每线都含有一定数目的象素，前面已有所描述。在帧 N－1 和帧 N 中都分别给出了一个象素 30。这两个象素的差异可以由下列公式来计算:

$$PD(x,\ y)=L^{N}(x,\ y)-L^{N-1}(x,\ y)$$

其中，$L^{N}(x,\ y)$代表当前帧中(x, y)位置象素 30 的亮度值，$L^{N-1}(x,\ y)$代表前一帧中(x, y)位置象素 30 的亮度值。

在图 2 的下半部分，给出了隔行扫描视频图象条件下的 3 个连续的场 N－2、N－1、N。场 N－2 与场 N 具有相同的奇偶性，而场 N－1 具有相反的奇偶性。这也可以通过图 2 中的实线与虚线清楚地看出来。仍然用参考序号 30 来表示将要计算差异值的象素，这一次的计算用下面的公式来进行:

$$PD(x,\ y)=L^{N}(x,\ y)-L^{N-2}(x,\ y)$$

在步骤 13 中，象素差异计算的结果被存储于相应的存储器位置。存储有最旧的帧 N－1/场 N－2 的帧/场存储器可以用于这一用途。在存储器中等于是存储了一幅象素差异的映象图。

步骤 14 中，象素差异的幅度被累加到预先确定的象素块上，可以是交迭的，也可以是不交迭的。图 3 给出了非交迭象素块中帧或场的再分示例。在这个例子中，每个象素块含有 16×8 个象素。自然，其它的例子也可用，如 8×8 象素块等。采用非交迭式象素块简化了存储器组织。

对于每个象素块B(i, j),计算组成象素块的每个象素的象素差异PD(i, j)的绝对值的总和，作为累计象素差异 APD(i, j)。这提高了在有噪声存在的情况下静止象素块检测的鲁棒性，因为图象中的噪声组份的随机变化可以被累计操作的过程消除。

在步骤15 中，累计象素差异值 APD(i, j)被收集于相应的存储器位置。注意，数字 i 和 j 决定了帧或场内象素块的位置。

下面对图 1 中含有步骤 16 至 21 的第二分支加以说明。在此分支中，计算前述每个象素块的阈值 T(i, j)。注意，在这个分支中，只考虑当前帧或场的象素数据。在步骤 16 中，计算一个“象素块梯度”g(i, j)。所谓的象素块梯度是指存在于当前象素块 B(i, j)中的空间梯度。在本发明的最佳实施例中，g(i, j) 的计算方式是从水平 $g_x(i, j)$ 与垂直 $g_y(i, j)$ 梯度组份中选取大的那个，用如下公式表示:

$$g(i, j)=MAX[g_x(i, j), g_y(i, j)]$$

其中，水平梯度组份 $g_x(i, j)$被估计为水平相邻的象素的亮度差异的绝对值的平均，垂直梯度组份 $g_y(i, j)$ 被估计为垂直相邻的象素的亮度差异的绝对值的平均，计算的对象包括组成当前象素块 B(i, j)的 N 个象素 P(k, l)，其灰度用 L(k, l)表示。计算公式如下:

$$\begin{cases} g_x(x,y) = \dfrac{1}{N} \sum_{L(k,l)\in B(i,j)} |L(k,l) - L(k-1,l)| \\ g_y(x,y) = \dfrac{1}{N} \sum_{L(k,l)\in B(i,j)} |L(k,l) - L(k,l-1)| \end{cases}$$

在步骤 17 中，如此算得的象素块梯度被存储于相应的存储位置。只所以考虑象素块梯度是因为即使摄象机发生非常小的移动，到了存在很高的空间频率的区域，也会导致巨大的象素差别。

在步骤 18 中计算阈值。阈值 T(i, j)被设计成与象素块梯度相匹配，相匹配的方式是它包含一个与象素块梯度成正比的“惩罚项”。这些阈值还被设计成与当前帧或场的噪声水平的估计值相匹配。噪声水平是被作为步骤 18 的一个外部输入。噪声水平的估计是在视频处理领域一个众所周知的问题。例如，号码为 EP－A－0 562407 的专利申请就公开了一种用于噪声测量的方法与装置。所以，在此不必详细解释噪声估计的过程。但上述专利申请可以作为参考。另外一个有用的信息是，本专利申请的发明人在本专利申请的同一申请日提交了另一份有关噪声估计的一种新方法的发明申请，其内部参考号是 PF980012。

设一个象素块中的象素数是 N，σ是可能的噪声水平的估计值， 每个阈值都可以由σ和 g(i, j)的加权平均来获得:

$T(i, j) = N(\alpha\sigma + \beta g(i, j))$

数值α和β取决于噪声水平σ和象素块 g(i, j)被表达的精确度。出于此目的，一个由不同数值的α和β所组成的表格可以被存储于存储器中。如果按照本发明的最佳实施例算得 g(i, j)，如上所述，并且如果σ代表了被表达成了一个灰度水平的标准偏差，那么α和β的最佳选择分别是 1.4 与 1/8。

在步骤 21 中，这些阈值被收集到相应的存储器位置中。

步骤 22 中，针对每个象素块 B(i, j)把累计象素差异 APD(i, j)与阈值 T(i, j)进行比较，以提供静止区域的一个初步映象图 PM。该映象图的每个象素 PM(i, j)都与一个象素块 B(i, j)相关联并代表一个为这个象素块所采取的二重（静止的或非静止的）决定。按照惯常的作法，数值 1 代表静止象素块，0 代表非静止象素块，PM(i, j)可以由如下公式给出:

$$\begin{cases} PM(i,j)=1 & \text{如果} APD(i,j) \le T(i,j) \\ PM(i,j)=0 & \text{如果} APD(i,j) > T(i,j) \end{cases}$$

比较是在步骤 22 内进行的，其结果在步骤 23 被输入存储器。当步骤 23 执行完成后，在存储器中已经有了一个初步的静止区域映象图。图 4 所示是这样一个初步映象图的示例。

值得注意的是，在多个非静止象素块所组成的区域内会有一些静止的象素块。通常这些象素块的尺寸远远小于图象中的真实物体的相应尺寸。这种情况下，孤立的象素块或小组孤立的象素块一定是误检测的结果，不能代表静止物体。所以，在初步映象图中的这些静止象素块相关联的象素必须有表示非静止象素块的象素来替代。

出于这种目的，在下一个步骤 24 中执行一个对映象图进行腐蚀的步骤，即把孤立的静止象素块或小组的相邻静止象素块去掉，就象进行对映象图进行腐蚀一样，从而提供一个最后的静止区域映象图 FM。这一步骤主要起作用于象素块的尺寸远远小于图象中真实物体的尺寸的情况。

在本发明的最佳实施例中，腐蚀过程的实现过程是处理初步映象图中的每个输入（等同于当前场或帧中的每个象素块），具体过程如下：如果当前象素块 B(i, j)已经被划分为非静止的，即 PM(i, j) = 0，那么，也设定 FM(i, j) = 0（即，什么也不做），而如果 PM(i, j) = 1，则：

1. 考虑一个以当前象素块为中心的水平邻近区域，并把 M_X 个象素块作为水平邻近它的象素块；

2. 数出在这个邻近区域内被划分为静止态的象素块的数目 S_X。

3. 如果比值 S_X / M_X 小于一个预先设定的阈值 T_X，那么设定 FM(i, j) = 0，即说明这一象素块为非静止的。

4. 考虑一个以当前象素块为中心的垂直邻近区域，并把 M_Y 个象素块作为垂直邻近它的象素块；

5. 数出在这个邻近区域内被划分为静止态的象素块的数目 S_Y。

6. 如果比值 S_Y / M_Y 小于一个预先设定的阈值 T_Y，那么设定 FM(i, j) = 0。

数值 M_X 、M_Y 、T_X 、T_Y 应该与象素块和图象的尺寸相匹配。图 5 给

出了用于在 NTSC、PAL、SECAM 和 16×8 象素的标准电视图象中对映象图进行腐蚀的过程所采用的水平与竖直邻近区域的一个示例。参考标号 40 表示含有 3 个相邻象素块的竖直邻近区域。参考标号 41 表示含有 5 个相邻象素块的竖直邻近区域。当前象素块由参考标号 42 表示。在这个例子中，T_X 值被确定为 3/5，T_Y 值被确定为 2/3。

对于位于场或帧的边界的象素块来说，可以改变对初步映象图进行腐蚀的步骤。例如，对于顶部和底部边界的象素块来说，步骤 4 至 6 可以被省略。对于左右边界的象素块来说，步骤 1 至 3 可以被省略。

在映象腐蚀步骤 24 之后，在步骤 25 中获得静止象素块的最终映象图。图 6 所示即是在对图 4 中的初步映象图进行清理之后而获得的最终映象图。

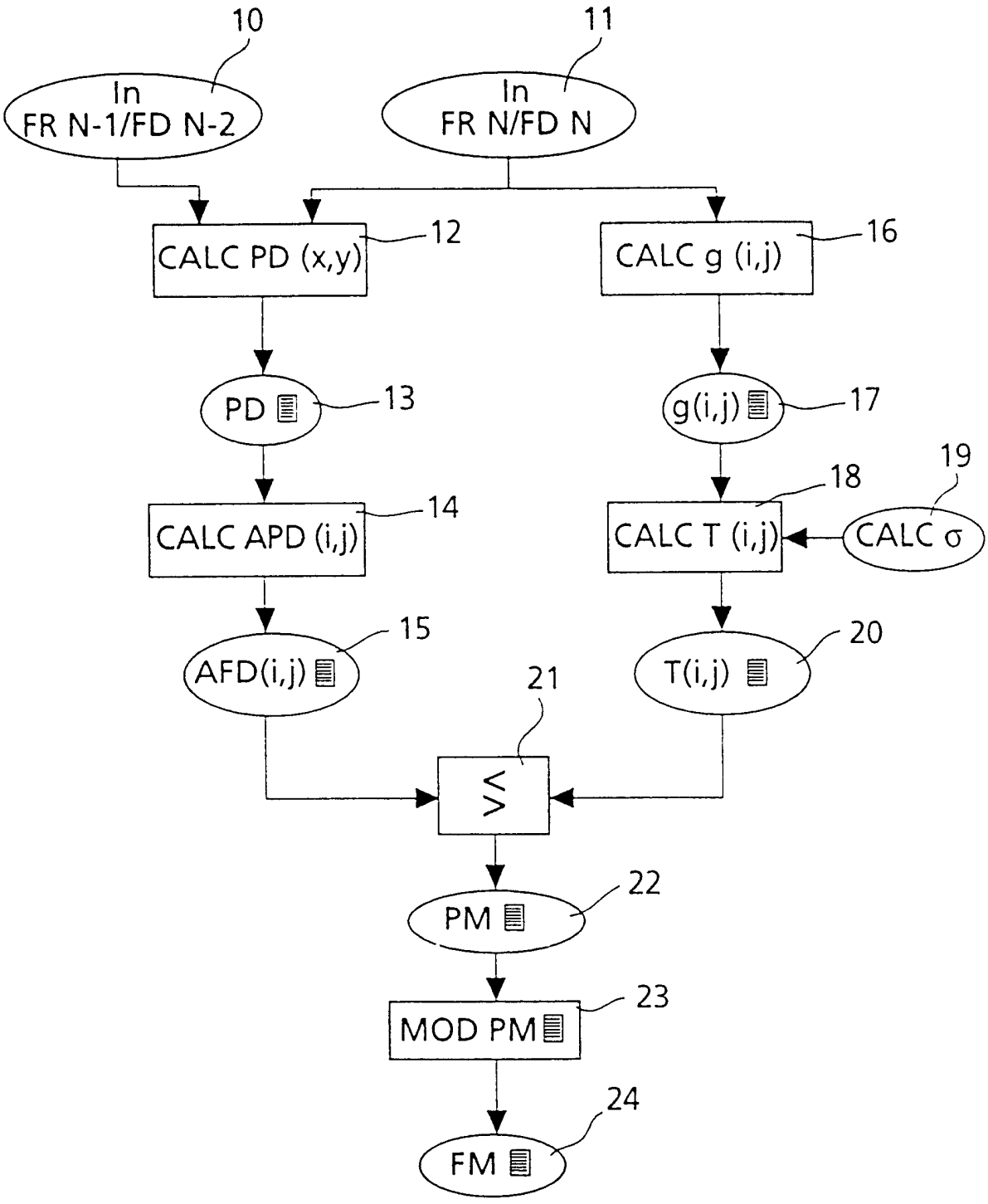

10
In
FR N-1/FD N-2
11
In
FR N/FD N
CALC PD (x,y)
12
CALC g (i,j)
16
PD
13
g(i,j)
17
19
18
CALC APD (i,j)
14
CALC T (i,j)
CALC σ
AFD(i,j)
15
T(i,j)
20
21
≤
>
PM
22
MOD PM
23
FM
24

图 1

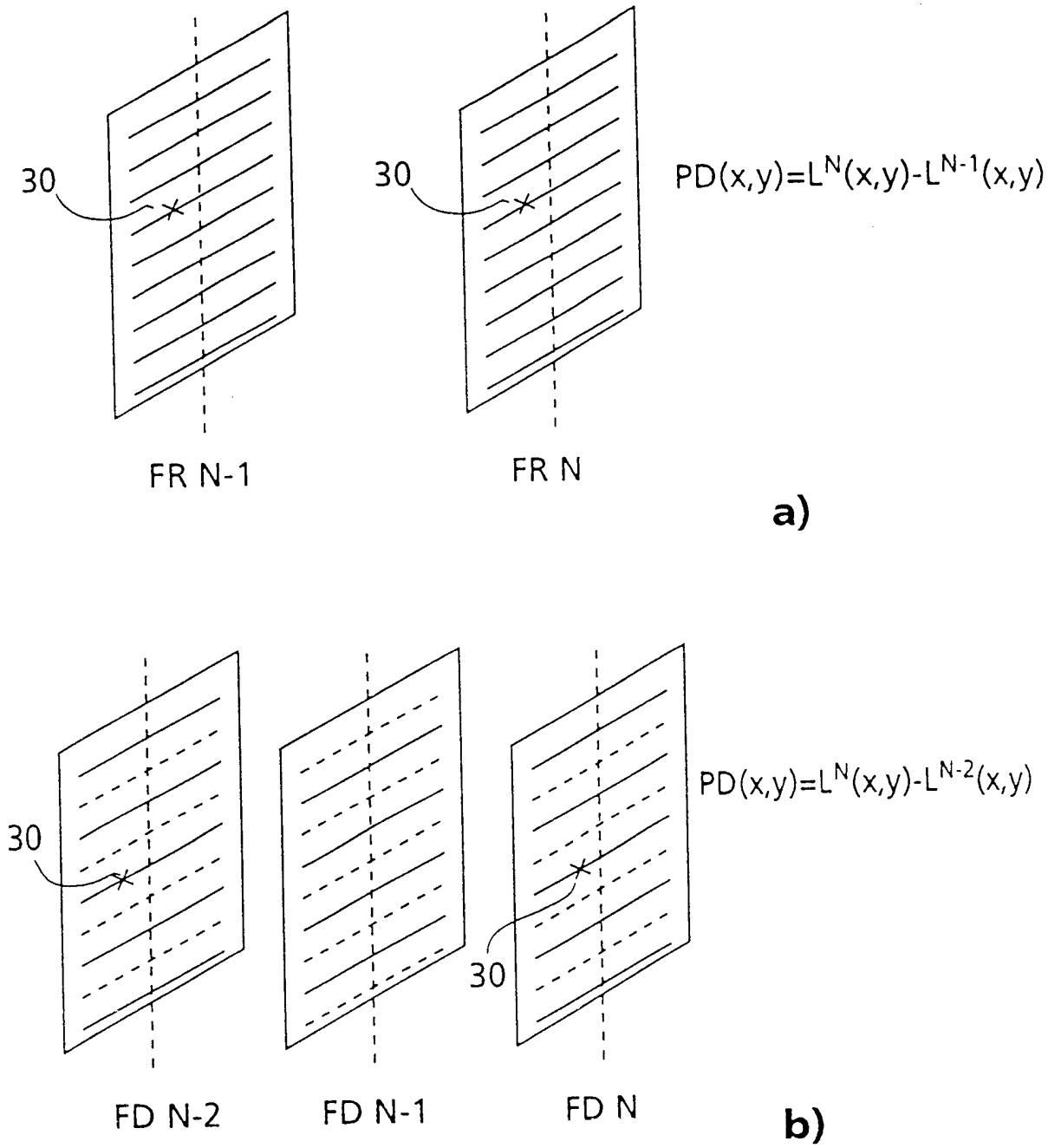

30
30
PD(x,y)=L^N(x,y)-L^N-1(x,y)
FR N-1
FR N
a)
30
PD(x,y)=L^N(x,y)-L^N-2(x,y)
30
FD N-2
FD N-1
FD N
b)

图 2

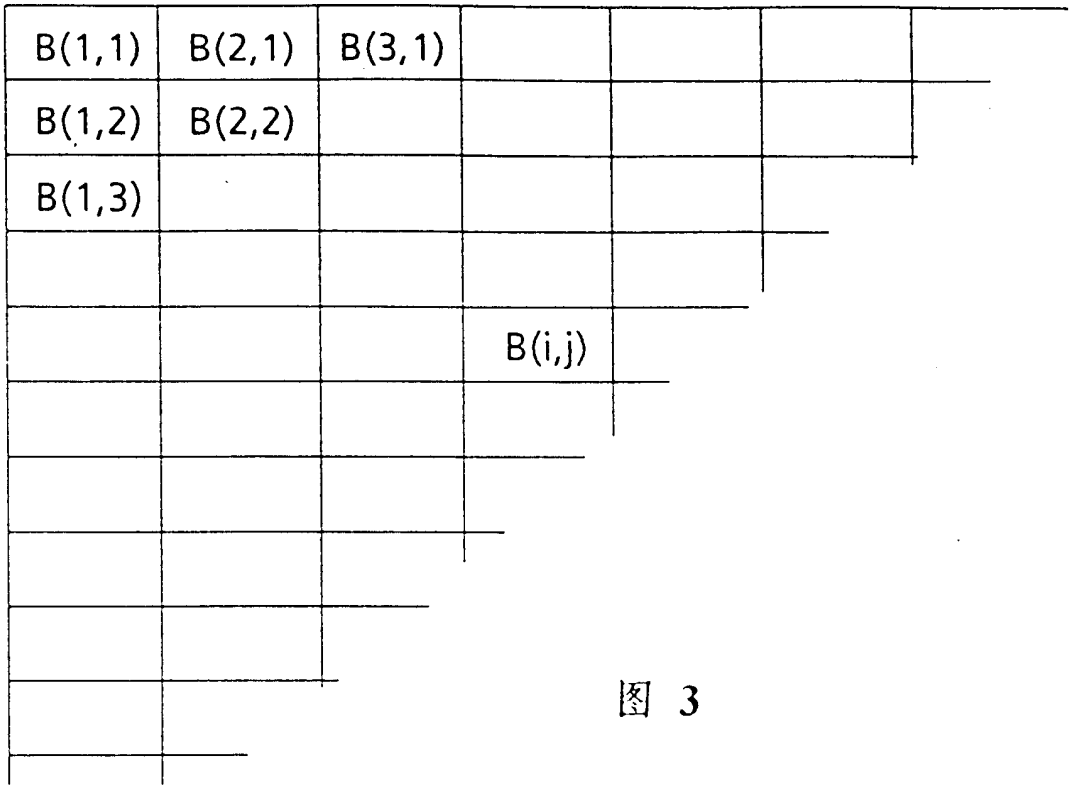


图 3

0	0	1	0	0	0	1	1	1	1
0	0	0	1	0	0	1	1	1	1
0	0	0	0	0	0	0	1	1	1
0	0	0	0	0	1	0	1	1	1
0	0	1	0	0	0	0	0	1	1
0	0	0	0	0	0	0	0	1	1

PM

图 4

0	0	0	0	0	0	1	1	1	1
0	0	0	0	0	0	1	1	1	1
0	0	0	0	0	0	0	1	1	1
0	0	0	0	0	0	0	1	1	1
0	0	0	0	0	0	0	0	1	1
0	0	0	0	0	0	0	0	1	1

FM

图 6

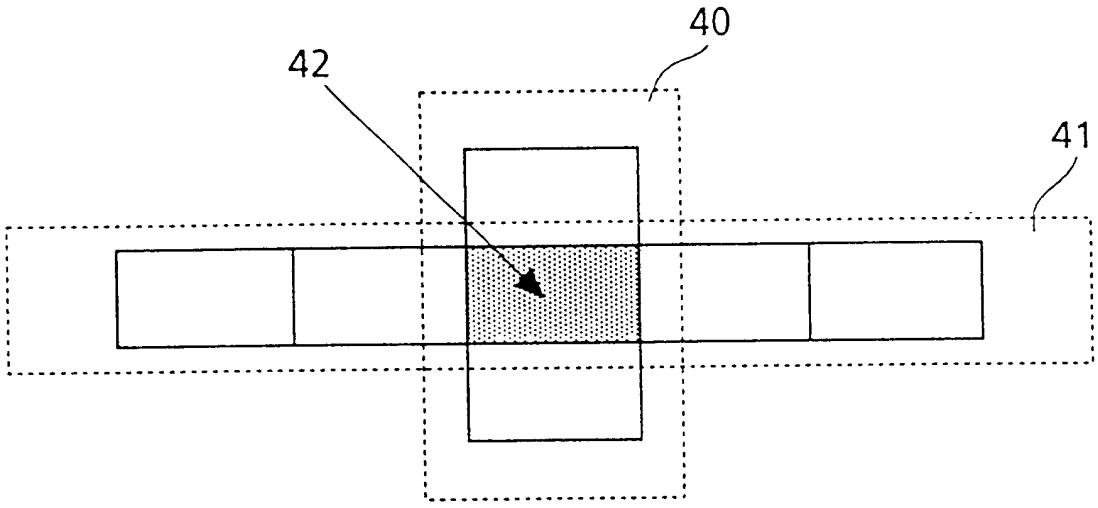


图 5