

[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] 发明专利说明书

专利号 ZL 200410077709.8

[45] 授权公告日 2008年6月4日

[11] 授权公告号 CN 100393133C

[22] 申请日 2004.12.30

[21] 申请号 200410077709.8

[73] 专利权人 中兴通讯股份有限公司

地址 518057 广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦A座6层

[72] 发明人 鲁照华 叶云 刘继兴 吴畏

[56] 参考文献

JP2000-115749A 2000.4.21

US5768533A 1998.6.16

US2004/0258155A1 2004.12.23

CN1133535A 1996.10.16

审查员 裴素英

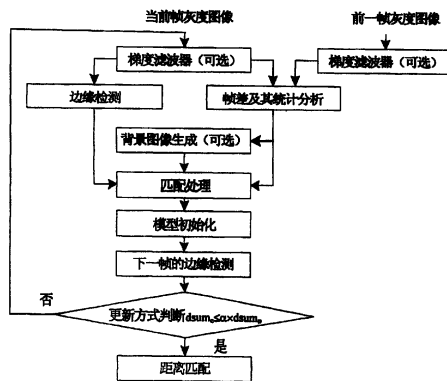
权利要求书2页 说明书8页 附图1页

## [54] 发明名称

一种视频序列中运动物体的分割方法

## [57] 摘要

本发明属于一种视频序列中运动物体的分割方法。为了克服物体运动快慢因素的影响，本方法包括如下步骤：(a) 求取当前帧与前一帧图像的帧差图像；以该帧差图像中各个像素点为中心开一个  $N * N$  大小的窗口，利用见右 (1) 式，三式求得一二值运动图像，该二值运动图像经过填充后得到一填充二值图像；(b) 求取当前帧的边缘图像；将该边缘图像的所有边缘点对应于填充二值图像内  $f_i = 255$  的像素所构成的集合  $E_i$ ，同经过形态滤波的上述填充二值图像的边缘像素集合  $E_p$  根据见右 (2) 式，形成初始化模型  $O_{ei}$ ；(c) 根据已得到的若干帧差图像和下一帧图像的分析结果，对下一帧图像进行模型更新。



$$m_i = (\sum_{x \in E_i} dy_x) / M, \sigma_i^4 = (\sum_{x \in E_i} (dy_x - m_i)^4) / M, f_i = \begin{cases} 255 & \sigma_i^4 > T \\ 0 & \text{其他} \end{cases} \quad (1)$$

$$O_{ei} = \{e \in E_i \mid \min_{x \in E_p} \|e - x\| \leq Th_{em}\} \quad (2)$$

1. 一种视频序列中运动物体的分割方法，其特征在于其包括如下步骤：

(a) 求取当前帧与前一帧图像的帧差图像；以该帧差图像中各个像素点为中心开一个  $N*N$  大小的窗口，利用下面 (1)、(2)、(3) 式求得一二值运动图像，该二值运动图像经过填充后得到一填充二值图像；

$$m_i = \left( \sum_{k \in s_i} dy_k \right) / M \quad (1)$$

$$\sigma_i^4 = \left( \sum_{k \in s_i} (dy_k - m_i)^4 \right) / M \quad (2)$$

$$f_i = \begin{cases} 255 & \sigma_i^4 > T \\ 0 & \text{其他} \end{cases} \quad (3)$$

其中， $dy_k$  是帧差图像中像素的灰度值， $s_i$  为以像素  $i$  为中心的窗口， $M$  为窗口所包含像素的数目  $N*N$ ， $T$  为一阈值， $f_i=255$  时表示像素  $i$  属于运动物体， $f_i=0$  时表示像素  $i$  属于背景， $m_i$  为包含  $N*N$  个像素点的帧差图像块中所有像素点的均值， $\sigma_i^4$  为包含  $N*N$  个像素点的帧差图像块中所有像素的四阶矩；

(b) 求取当前帧的边缘图像；将该边缘图像的所有边缘点中对应于填充二值图像内  $f_i=255$  的像素所构成的集合  $E_i$ ，同经过形态滤波的上述填充二值图像的边缘像素集合  $E_p$  根据式 (4) 形成初始化模型  $O_{ci}$ ；

$$O_{ci} = \{ e \in E_i \mid \min_{x \in E_p} \|e-x\| \leq Th_{init} \} \quad (4)$$

其中， $e$  为边缘图像的所有边缘点中对应于填充二值图像内  $f_i=255$  的像素所构成的集合  $E_i$  中的元素， $x$  为形态滤波得到的填充二值图像的边缘像素集合  $E_p$  中的元素， $Th_{init}$  为一阈值；

(c) 根据已得到的若干帧差图像和下一帧图像的分析结果，对下一帧图

像进行模型更新。

2. 如权利要求 1 所述的视频序列中运动物体的分割方法,其特征在于在步骤(a)中的当前帧与前一帧图像如果有阴影则经过梯度滤波处理,如果没有阴影则不经过梯度滤波处理,然后进行帧差图像的求取。
3. 如权利要求 1 所述的视频序列中运动物体的分割方法,其特征在于在步骤(b)中的当前帧与前一帧图像如果有阴影则经过梯度滤波处理,如果没有阴影则不经过梯度滤波处理,再进行边缘图像的求取。
4. 如权利要求 1 所述的视频序列中运动物体的分割方法,其特征在于在步骤(a)中该填充二值图像是通过将当前帧的背景图像与上述二值运动图像经过逻辑与的运算并填充后得到。
5. 如权利要求 1 所述的视频序列中运动物体的分割方法,其特征在于在步骤(b)中的当前帧的边缘图像是通过 Sobel 算法、Watershed 算法、域值法、聚类法、贝叶斯法、或利用 Canny 算子对当前帧图像进行边缘检测的方法获得。
6. 如权利要求 1 所述的视频序列中运动物体的分割方法,其特征在于式(4)中的  $Th_{min}$  的范围为 1~3 个像素。
7. 如权利要求 1 所述的视频序列中运动物体的分割方法,其特征在于在步骤(c)中根据下式(5),对下一帧图像进行模型更新:若式(5)成立,则采用距离匹配算法实现模型更新,若式(5)不成立,则通过实施步骤(a)、(b)实现模型更新;

$$dsum_c \leq \alpha \times dsum_p \quad (5)$$

其中,  $dsum_p$  为当前帧差图像之前若干帧差图像的所有灰度像素绝对值和的帧平均值,  $dsum_c$  为当前帧差图像的所有灰度像素绝对值和,  $\alpha$  为一乘数因子。

8. 如权利要求 7 所述的视频序列中运动物体的分割方法,其特征在于该距离匹配算法为 hausdorff 距离算法或欧氏距离匹配算法。

## 一种视频序列中运动物体的分割方法

### 【技术领域】

本发明属于多媒体通信领域中的数字视频压缩技术，尤其涉及 MPEG-4 标准中视频序列中运动物体的分割方法。

### 【背景技术】

传统的视频压缩标准 MPEG-1、MPEG-2、H.261、H.263，均采用带有运动补偿的预测编码与 DCT 变换编码相结合的混合矩形帧压缩编码方法，没有对视频场景进行分割或分析，属于低级压缩编码技术。随着多媒体应用的快速增长和基于内容的交互需求，这些基于帧压缩编码技术的视频标准无法满足现有要求。为此，MPEG 专家组制订了新的基于内容的视频压缩国际标准—MPEG-4。通过引入视频对象（VO）、视频对象平面（VOP）的概念，将输入视频序列的每一帧分割成相应形状的图形区域，使每一帧视频对象都代表语义上有意义的对象或感兴趣的视频内容。视频分割，实质上就是从视频序列中抽取运动物体的形状信息，是实现基于内容压缩编码、多媒体内容描述及智能信号处理等技术的关键。

根据分割准则的不同，传统的视频分割方法主要有两大类：空域一致性和变化检测方法。空域一致性方法的基本处理步骤是先利用形态滤波器简化要处理的矩形视频帧，再借助 watershed 等方法将整个矩形帧分成若干个小区域，然后通过区域合并形成最后的对象区域。变化检测方法通过相邻两帧的帧差信息来检测运动物体的位置和形状，利用空域、时域信息优化检测的结果。因为运动物体的运动是不同于背景的，空域一致性方法在不知道运动信息的情况下，单纯利用空域信息分割背景，会浪费大量的计算资源，所以变化检测方法更为

高效，但是传统的变化检测方法会受到物体运动快慢、阴影效应、背景误判等因素的影响，常用的去除这些不利因素的方法所需的计算量很大，且效果不佳。

### 【发明内容】

为此，本发明要解决的技术问题是提供一种快速高效，且不受物体运动快慢因素的影响的视频序列中运动物体的分割方法。

为解决上述技术问题，本发明提供一种视频序列中运动物体的分割方法，其包括如下步骤：

- (a) 求取当前帧与前一帧图像的帧差图像；以该帧差图像中各个像素点为中心开一个  $N*N$  大小的窗口，利用下面 (1)、(2)、(3) 式求得一二值运动图像，该二值运动图像经过填充后得到一填充二值图像；

$$m_i = \left( \sum_{k \in s_i} dy_k \right) / M \quad (1)$$

$$\sigma_i^4 = \left( \sum_{k \in s_i} (dy_k - m_i)^2 \right) / M \quad (2)$$

$$f_i = \begin{cases} 255 & \sigma_i^4 > T \\ 0 & \text{其他} \end{cases} \quad (3)$$

其中， $dy_k$  是帧差图像中像素的灰度值， $s_i$  为以像素  $i$  为中心的窗口， $M$  为窗口所包含像素的数目  $N*N$ ， $T$  为一阈值， $f_i=255$  时表示像素  $i$  属于运动物体， $f_i=0$  时表示像素  $i$  属于背景， $m_i$  为包含  $N*N$  个像素点的帧差图像块中所有像素点的均值， $\sigma_i^4$  为包含  $N*N$  个像素点的帧差图像块中所有像素的四阶矩；

- (b) 求取当前帧的边缘图像；将该边缘图像的所有边缘点中对应于填充二值图像内  $f_i=255$  的像素所构成的集合  $E_i$ ，同经过形态滤波的上述填充二值图像的边缘像素集合  $E_p$  根据式 (4) 形成初始化模型  $O_{ci}$ ；

$$O_{ci} = \{e \in E_i \mid \min_{x \in E_p} \|e-x\| \leq Th_{mit}\} \quad (4)$$

其中， $e$  为边缘图像的所有边缘点中对应于填充二值图像内  $f_i=255$  的像素所构成的集合  $E_i$  中的元素， $x$  为形态滤波得到的填充二值图像的边缘像素集合  $E_p$  中的元素， $Th_{mit}$  为一阈值；

(c) 根据已得到的若干帧差图像和下一帧图像的分析结果，对下一帧图像进行模型更新。

其中，步骤(a)中的当前帧与前一帧图像如果有阴影则经过梯度滤波处理，如果没有阴影则不经过梯度滤波处理，然后进行帧差图像的求取。步骤(b)中的当前帧与前一帧图像如果有阴影则经过梯度滤波处理，如果没有阴影则不经过梯度滤波处理，再进行边缘图像的求取。在步骤(a)中该填充二值图像可通过将当前帧的背景图像与上述二值运动图像经过逻辑与的运算并填充后得到。在步骤(b)中的当前帧的边缘图像是通过 Sobel 算法、Watershed 算法、域值法、聚类法、贝叶斯法、或利用 Canny 算子对当前帧图像进行边缘检测的方法获得。

此外，在步骤(c)中可根据下式(5)，对下一帧图像进行模型更新：若式(5)成立，则采用距离匹配算法实现模型更新，若式(5)不成立，则通过实施步骤(a)、(b)实现模型更新；

$$dsum_c \leq \alpha \times dsum_p \quad (5)$$

其中， $dsum_p$  为当前帧差图像之前若干帧差图像的所有灰度像素绝对值之和的帧平均值， $dsum_c$  为当前帧差图像的所有灰度像素绝对值之和。上述距离匹配算法可为 hausdorff 距离算法或欧氏距离匹配算法， $\alpha$  为一乘数因子。

本方法在步骤(a)中，通过选用四阶矩求取二值运动图像，可以有效地消除物体运动快慢对分割结果的影响。另外，对于阴影区域的边缘存在大的梯度值时，利用梯度滤波可以使阴影区域的灰度值变的非常小，从而显著减少阴影的影响，同时也可以消除由于相机增益变化或照明条件变化对视频序列产生的影响。此外，通过背景图像与帧差及统计分析过程中得到的二值运动图像经过逻辑

辑与运算并填充得到的幅填充二值图像参与下面地匹配运算，也可以有效地提高分割精度。

### 【附图说明】

下面结合附图及实施例对本发明进行详细说明：

图 1 是本发明的方法流程图。

### 【具体实施方式】

本发明的基本构思是基于模式识别和目标追踪，通过对当前帧差图像的统计分析，即判断每个像素的运动状态，得到一幅二值运动图像；同时当前帧经边缘检测后得出当前帧的边缘图像，然后将上述所得图像根据一定准则进行匹配运算，所得结果通过形态滤波处理后得到运动物体初始化模型，然后根据先前已得到的若干帧差图像和下一帧差图像的统计分析决定对下一帧图像是采用距离匹配方法还是重复模型初始化的过程实现模型更新。

下面结合附图和具体实施例对本发明的构思做进一步说明。

首先请参照图 1，本发明方法的流程主要包括帧差及其统计分析、边缘检测、匹配处理、模型初始化和模型更新等模块。本发明方法的具体实现可分为以下两个步骤：模型初始化和模型更新。

第一步模型初始化主要由帧差及其统计分析、边缘检测、匹配处理模块组成。

#### 1) 帧差及其统计分析

在该步骤中，首先求取当前帧与前一帧图像的帧差图像。其中，该当前帧与前一帧图像如果有阴影则经过梯度滤波处理，如果没有阴影则不经过梯度滤波处理，然后进行帧差图像的求取。本发明利用统计规律抽取帧差图像空间差异大的区域，以检测出运动物体。通过对相邻两帧图像取灰度差的绝对值，得

到帧差图像。然后以该图像中各个像素点为中心开一个  $N*N$  大小的窗口， $N$  的大小一般为 3 个像素，也可根据图像大小和处理器的处理能力适当变化。再利用下式 (1)、(2)、(3) 确定该像素是静止的还是运动的，得到一幅二值运动图像。

$$m_i = \left( \sum_{k \in s_i} dy_k \right) / M \quad (1)$$

$$\sigma_i^4 = \left( \sum_{k \in s_i} (dy_k - m_i)^2 \right) / M \quad (2)$$

$$f_i = \begin{cases} 255 & \sigma_i^4 > T \\ 0 & \text{其他} \end{cases} \quad (3)$$

其中， $dy_k$  是帧差图像中像素的灰度值。 $s_i$  代表以像素  $i$  为中心的窗口。 $M$  为窗口所包含像素的数目  $N*N$ 。 $T$  为一阈值，一般情况下取值为 23，也可根据具体情况改变  $T$  值大小。 $f_i=255$  表示像素  $i$  属于运动物体， $f_i=0$  时表示像素  $i$  属于背景。

本实施例中选用四阶矩可以有效地消除物体运动快慢对分割结果的影响。该二值运动图像经过填充后，便得到一幅填充二值图像。

## 2) 边缘检测

本步骤主要是求取当前帧的边缘图像。同样，在本步骤中，如果当前帧与前一帧图像有阴影则需经过梯度滤波处理，如果没有阴影则不经过梯度滤波处理，然后进行帧差图像的求取。当前帧的边缘图像可以采用边缘检测算法，如 Sobel 算法、Watershed 算法、域值法、聚类法、贝叶斯法等分割方法得到，也可利用结合高斯去噪和梯度锐化方法优点的 Canny 算子对当前帧图像进行边缘检测，提取当前帧所有边缘像素点来得到。

## 3) 匹配处理

本步骤是将当前帧通过边缘检测算法得到的所有边缘点中对应于填充二值



图像内灰度值为 255 的像素都被保留下来，以构成集合  $E_i$ 。该集合  $E_i$  同经过形态滤波的填充二值图像的边缘像素集合  $E_p$  利用下式 (4) 进行最小距离判断，以形成初始化模型  $O_{ci}$ 。

$$O_{ci} = \{e \in E_i \mid \min_{x \in E_p} \|e-x\| \leq Th_{init}\} \quad (4)$$

式 (4) 中， $Th_{init}$  一般取 1 到 3 个像素。

在模型初始化过程中，也可根据视频序列情况、分割精度等因素有选择地加入消除阴影效应的梯度滤波器以及提高分割精度的背景图像生成模块。

在许多实际应用中，运动物体的阴影可能出现在某个场景的背景区域中。当物体运动时，其阴影的形状也会发生变化，这就使从背景区域中分割运动物体变的相当困难。实验发现，在通常情况下，阴影区域的灰度值总有一种梯度的变化趋势，因此，当阴影区域的边缘存在大的梯度值时，利用梯度滤波可以使阴影区域的灰度值变的非常小，这将显著减少阴影的影响，同时也可以消除由于相机增益变化或照明条件变化对视频序列产生的影响。因此，可有选择地对当前帧进行梯度滤波处理，其中梯度滤波的实现可采用传统的灰度域值统计、形态滤波等数字图像处理方法实现。

采用的梯度滤波器可由形态膨胀、腐蚀运算实现，具体公式如下：

$$G = (I \oplus B) - (I \ominus B) \quad (5)$$

其中， $G$  为梯度图像， $I$  为输入视频帧， $B$  为形态运算中的结构元素。另外梯度滤波器也可以利用灰度域值统计等数字图像处理方法实现。

另外，在许多实际应用中，运动物体总是在一定背景区域中运动的，通过分析当前帧差图像之前的若干帧差图像，可判断任意位置处像素的亮度变化情况。若某位置处像素在这些图像中都被判断为静止的，则该像素点属于背景。同时通过反复的帧差图像分析可得到更新的背景图像，也可以利用帧与帧之间的运动信息来提取视频序列的背景。该背景图像与帧差及统计分析过程中得到的二值运动图像经过逻辑与运算并填充后，得到一幅填充二值图像，属于运动

物体的像素对应的值为 255，属于背景的对应的值为 0。以该填充二值图像参与下面地匹配运算，可提高分割精度。因此，在本实施例中可以有选择地加入背景图像生成模块。

当背景变化不大时，采用背景图像生成技术可以改善视频分割的图像质量。如果背景变化很大，则可采用 MPEG-4 中的 sprite 技术来获得整个序列的大背景图像，然后再利用 MPEG-4 中的技术将某帧相应的背景找到，然后进行分割运算。

在视频会议系统中，如果与会者的背景全部都是固定颜色的，例如蓝色，则只需进行帧差及其统计分析这个步骤，就可以简单快速完成视频序列中运动物体的分割，而不需要进行其它步骤。

第二步是模型更新。在视频序列中，物体会不断运动，因此在每一帧中必须对目标模型进行更新。实质上就是从当前帧所有边缘像素点构成的二值图像中找到新模型。根据对帧差图像的统计分析决定对该帧图像是采用距离匹配方法还是采用重复第一步模型初始化的过程实现模型更新。

### 1) 更新方式判断

模型更新方法可以通过下面的公式 (6) 进行判断。若公式 (6) 成立，说明物体在进行低速运动，则采用距离匹配算法实现模型更新。反之若公式 (6) 不成立，说明物体运动比较快，则可采用第一步中的方法实现模型更新。

$$dsum_c \leq \alpha \times dsum_p \quad (6)$$

其中， $dsum_p$  为当前帧差图像之前若干帧差图像的所有灰度像素绝对值之和的帧平均值， $dsum_c$  为当前帧差图像的所有灰度像素绝对值之和。

### 2) 距离匹配算法

距离匹配算法可以采用常用的距离匹配算法，例如 hausdorff 距离算法，或简单的欧氏距离匹配算法。

距离匹配算法的实现过程如下：设  $O_q$  为前一帧  $q$  的模型， $E_{q+1}$  为当前帧  $q+1$

排除属于更新背景图像的边缘像素点的剩余边缘像素点集合， $O_{q+1}$ 为当前帧  $q+1$  的模型。该方法的目的是从  $E_{q+1}$  中找到  $O_{q+1}$ 。

$$h(O_q, E_{q+1}) = \min_{e \in E_{q+1}} \|o - e\|, \quad o \in O_q \quad (7)$$

对属于前一帧  $q$  的模型  $O_q$  每一模型点  $o$ ，利用式 (7) 在当前帧  $q+1$  的边缘像素点集合  $E_{q+1}$  中找到与  $o$  距离最近的点  $e$ 。 $O_q$  的更新模型  $O_{q+1}$  即为  $E_{q+1}$  与  $O_q$  的最佳匹配，即符合式(7)的边缘像素点集合  $E_{q+1}$  的点集合。

需要说明的是，上述说明仅是对本发明较佳实施例的详细描述，叙述仅为说明本发明的可实现性及其突出效果，具体特征并不能用来作为对本发明的技术方案的限制，本发明的保护范围应以本发明所附权利要求书为准。

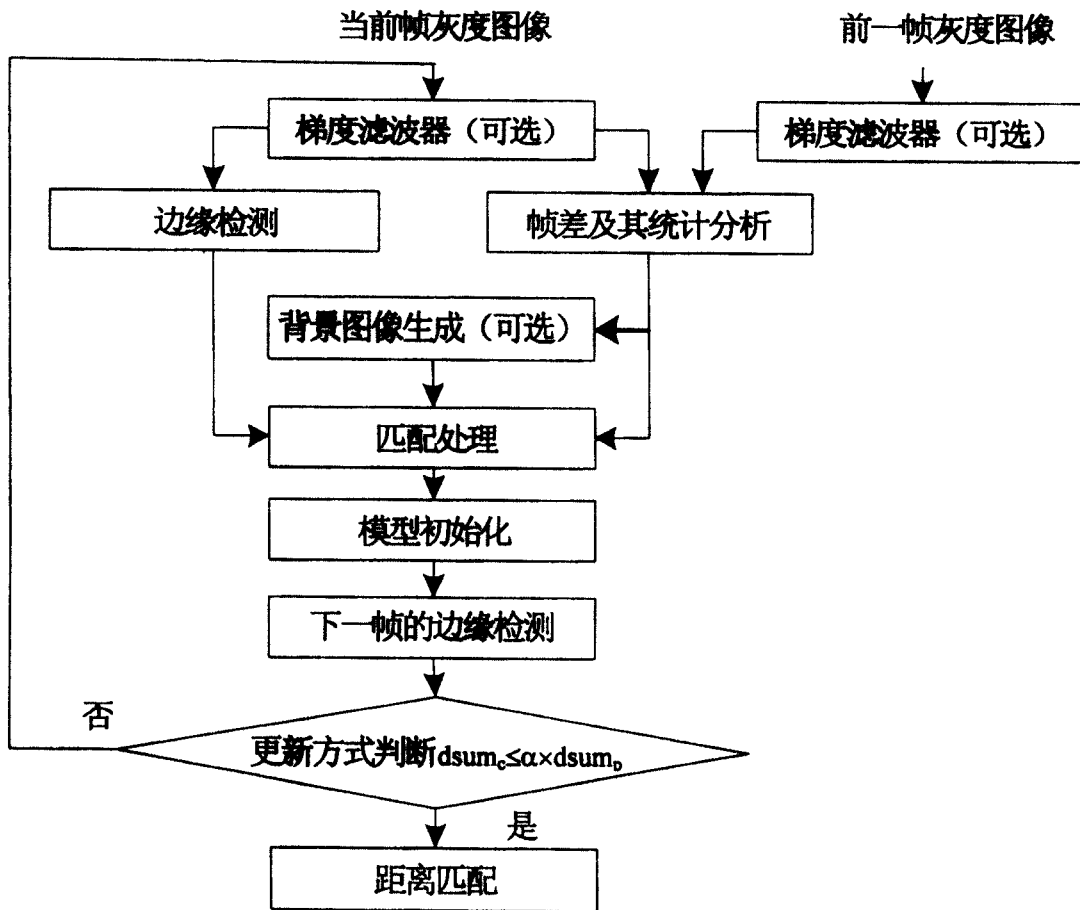


图 1