

文章编号:1007-2780(2011)06-0831-05

基于分块背景建模的运动目标检测技术

刘 翔¹, 周 桢²

(1. 武汉科技大学 工程训练中心, 湖北 武汉 430065, E-mail: liuxiang@wust.edu.cn;
2. 武汉市广播影视局播送中心, 湖北 武汉 430015)

摘 要: 研究了视频监控系统中的目标检测技术, 在分析比较几种目标检测算法的基础上提出了一种使用分块处理的方式建立初始背景模型的方法。该方法对视频序列图像中的噪声有较好的抑制作用, 并且能够有效地克服因平均而造成背景模糊和阴影残留的缺点。为适应背景变化, 在背景更新方法中结合三帧差分寻找像素值变化较大的像素点进行自适应更新。仿真结果表明, 在背景减除中使用该方法能够快速、完整、准确地检测出运动目标。

关 键 词: 运动目标检测; 背景建模; 背景更新; 背景减除

中图分类号: TP391.41 **文献标识码:** A **DOI:** 10.3788/YJYXS20112606.0831

Moving Objectives Detecting Technology Based on Block Processing of Background Model Establishment

LIU Xiang¹, ZHOU Zhen²

(1. Engineering Training Centre of Wuhan University of Science and Technology, Wuhan 430065, China, E-mail: liuxiang@wust.edu.cn;
2. Broadcast-Transmission Centre of Wuhan Broadcasting-Television Bureau, Wuhan 430015, China)

Abstract: The moving objectives detecting technology in video surveillance system was studied. On the base of analysis and comparison of several target detection algorithm, a way of using block processing method to establish the initial background model was proposed. The method had a good inhibitory effect on the noise in video sequences, and it can overcome the shortcomings of fuzzy and residual shadow caused by mean method; In order to meet the change of background, the three-frames differences method was used to find those pixels that had a significant difference, then these pixels were updated adaptively. Simulation experiment showed that, this method can be used in background subtraction algorithm. The moving targets can be detected rapidly, completely and accurately.

Key words: moving objectives detecting; background modeling; background updating; background subtraction

1 引 言

运动目标检测 (Moving Objectives Detecting, MOD) 是计算机视觉应用和无人值守监控

及自动化实时报警系统中的关键技术之一。MOD 需要将运动目标的信息从图像序列中检测出来, 从而为后续的跟踪建立模型, 完成跟踪和识别的最初处理。现有的运动目标检测算法

主要有帧间差分^[1]、光流法^[2]和背景相减法^[3]。帧间差分法利用图像序列中连续两帧或几帧来检测运动目标,实现方法简单且无需建立背景,对场景中光线的变化也不敏感,易于实时监控处理;但是其检测结果往往容易产生“空洞”,无法提取出完整的运动目标。光流法给图像中的每个像素赋予一个速度矢量,在某个时刻,图像上的点与三维物体上的点一一对应。如果图像中无运动目标,则光流矢量是连续变化的;如果有运动物体存在,则运动物体的速度矢量必然和邻域背景速度矢量不同,从而可以检测出运动物体。光流法可以获得完整的运动信息,但是它采用迭代方式求解,计算时间很长,无法满足系统实时性的要求。背景相减法的基本思想是将图像看成是前景与背景的组合,然后从图像中将背景减除就可以得到完整的运动目标。该方法简单、实时性好、能够检测出完整的运动信息,是目前应用最广泛的一种运动目标检测方法,也是本文选用它作为运动检测方法的原因。然而由于场景环境的复杂性,获得一幅没有运动目标的背景图像并不简单。随着时间的推移,背景图像的更新也是需要考虑的问题。

为了获得一幅“干净”的背景图像,本文提出了一种改进的基于图像分块处理的背景生成方法。该方法以分块中边缘点数与均值背景中相应位置上边缘点数的差作为判断该块属于背景还是前景的依据。在背景更新过程中,使用三帧差分检测发生变化的像素,在分割差分图像确定运动变化区域时引入了一个修正因子对差分后的自适应阈值进行修正,以减少扰动对背景更新的干扰。实验结果证明,本文提出的两种方法在背景生成与更新中取得了良好的效果。

2 背景建模

背景是一幅不含运动目标的静态图像。如果在某个场景的某个时段内没有运动物体进入,并且在这一时段的场景内没有扰动,场景的光照强度也没有发生变化,那么该时段内任何一幅场景的图像都可以作为背景。但在实际情况中,由于环境的复杂性,几乎不可能找到一幅满足上述描述的背景。在运动检测中使用的背景都是通过某种方法建立起来的背景模型。常用的背景建模的

方法有均值法^[4-5]、中值法^[6-7]和高斯模型法等^[8-9]。

(1) 均值法背景建模

运动目标的像素值与背景之间存在较大的差异,并且运动目标在某一位置停留的时间一般不会太长,所以可以将一段时间内某位置像素的均值作为该位置背景的背景像素值,这就是均值法的基本思想。用数学式表示均值法为:

$$B(x, y) = \frac{\sum_{n=0}^{N-1} I_n(x, y)}{N} \quad (1)$$

在式(1)中, $B(x, y)$ 表示背景, $I_n(x, y)$ 表示视频序列中的第 n 帧图像, N 表示帧数。均值法计算简单,但当场景内的运动目标较多时, N 的取值必须足够大才能得到理想的背景。 N 值越大,获得的背景就越模糊,而且建立背景的时间也会增加,不利于实时处理。

(2) 中值法背景建模

在一段时间内,场景中某个位置上背景像素会在大多数时间被观察到,所以可将这段时间内的视频序列各帧对应位置的像素点取中值作为背景像素。中值法用数学式可以表示为:

$$B(x, y) = \text{median}[I_0(x, y), I_1(x, y), I_2(x, y) \cdots I_{N-1}(x, y)] \quad (2)$$

在式(2)中, $B(x, y)$ 表示背景,median 表示取中值。对于中值法来说, N 的值不宜太小,否则,若背景像素在多帧内都被运动目标遮挡,背景建模将会失败。

(3) 高斯模型背景建模

高斯背景建模分为单高斯模型和混合高斯模型两种情况。用一个高斯分布 $N(\mu, \sigma^2)$ 去拟合视频序列中相应位置的同一个像素称为单高斯模型。其中 μ 是序列帧相应位置像素值的均值, σ^2 是方差。如果 $\mu - 3\sigma < I_n(x, y) < \mu + 3\sigma$, 则认为背景像素,否则即为运动目标。如果用多个高斯分布拟合视频序列中相应位置的同一个像素则称为混合高斯模型。这种方法给每个高斯分布赋予一定的初始权值和优先级,再根据观测的像素与建立的高斯模型的匹配情况对这些权值和优先级进行更新。最后,将各高斯分布按新的优先级由高到低进行排列,取前面的 B 个高斯分布表征背景模型。若观测像素与这 B 个分布中的任意一个匹配则为背景,否则为运动目标。与单高斯

模型相比,混合高斯模型更能反映出现实场景的变化情况,但是混合高斯模型对每个像素都用多个高斯分布进行描述,计算量大,不利于实时监控的需要。

(4) 本文使用的背景建模方法

针对上述一些建模方法的优缺点,本文提出了一种新的分块处理的背景建模方法,其具体步骤是:

Step1:将视频序列的前 N 帧用均值法得到一幅均值背景图像 B_m ;

Step2:对这 N 帧序列中的每一幅图像以及 Step1 中得到的 B_m 用文献[10]的形态学边缘检测的方法提取边缘并分割成二值图像,将结果分别记为: I_n^b , ($n=0,1,\dots,N-1$) 和 B_m^b , 其中下标 n 表示帧号;

Step3:采用分块处理的方法将 I_n^b 与 B_m^b 分成 8×8 的小块,若图像序列的宽度记为 W ,高度记为 H ,则 I_n^b 与 B_m^b 都可以分为 $\text{int}(W/8) \times \text{int}(H/8)$ 块, $\text{int}(\ast)$ 表示取整运算。用 $I_n^b\{x\}$, $B_m^b\{x\}$, $n=0,1,\dots,N-1, x=1,2,\dots,\text{int}(W/8) \times \text{int}(H/8)$ 分别表示某帧中一小块和均值背景中一小块。例如:第3帧中的第8块即为 $I_2^b\{8\}$;

Step4:统计每帧各小块中边缘点的个数,并将其记为: $\text{NUM_}I_n^b\{x\}$, 它表示第 $n-1$ 帧中第 x 小块内边缘点的数量。同样统计 B_m^b 中各块边缘点数,记为: $\text{NUM_}B_m^b\{x\}$;

Step5:计算 $I_n^b\{x\}$ 内边缘点数与 B_m^b 中对应块 $B_m^b\{x\}$ 的边缘点数之差,如果差值的最小值小于某个阈值 Q_{th} ,则记录下 n 和 x 的值,并将均值背景中对应位置的小块 $B_m^b\{x\}$ 用 $I_n^b\{x\}$ 替换。 $I_n^b\{x\}$ 表示第 $n-1$ 帧中的第 x 块。如果差值的最小值都大于 Q_{th} ,则 $B_m^b\{x\}$ 保持不变。用数学式表示即为:

$$B\{x\} = \begin{cases} B_m^b\{x\} & \text{if } |\min(\text{NUM_}I_n^b\{x\} - \text{NUM_}B_m^b\{x\})| > Q_{th} \\ I_n^b\{x\} & \text{if } |\min^*(\text{NUM_}I_n^b\{x\} - \text{NUM_}B_m^b\{x\})| \leq Q_{th} \end{cases} \quad (3)$$

式中“ $\ast$ ”表示若满足 $\min(\text{NUM_}I_n^b\{x\} - \text{NUM_}B_m^b\{x\})$ 的值不止一个,且都小于阈值 Q_{th} 时,还要对满足条件的各帧与 $B_m^b\{x\}$ 进行求取相关系数的运算,取相关系数最大的那一帧替代 $B_m^b\{x\}$ 。

本文在检测边缘中使用了形态学的方法。

与经典的边缘检测算法相比,形态学方法对噪声的抑制更强,运算速度更快。如果在视频序列中有运动物体存在,则在各帧的相应位置上就会检测出较多的边缘点,相反边缘点的数量就少。边缘点较少的区域被认为是与背景较接近的区域,而对于有较多边缘点的区域则认为是前景区域,对前景区域使用平均背景进行替换。实验表明,采用本文方法进行背景建模时,背景中绝大多数小块取自各帧中对应位置,只有少数小块取自背景均值。因此,用这种方法建立的背景比均值法建立的背景更加清晰;而与混合高斯建模的方法相比,这种算法更加简单,计算量更小。

3 背景更新

由于光照、场景变化等各种因素的影响,背景不可能在建立之后就一成不变。要使背景模型在一段时间之内能够适应环境的变化必须对初始模型不断地进行更新。背景更新的实质是用当前帧中发生了变化的信息去修正使用过去帧建立的模型。一种常用的自适应背景更新方法采用式(4)进行:

$$B_n(x, y) = \alpha \cdot B_{n-1}(x, y) + (1 - \alpha) \cdot I_n(x, y) \quad (4)$$

在式(4)中, $B_n(x, y)$ 是当前时刻经过更新后得到的背景; $B_{n-1}(x, y)$ 是前一时刻建立的背景; $I_n(x, y)$ 是用于修正前一时刻背景的当前帧; α 称为学习速率,其值可以通过实验确定,一般取 0.995。

式(4)是用当前帧中每个位置上的像素值与原来背景中相应位置的像素值加权平均值替代原背景对应位置上的值。但是在实际的环境中,前后两帧图像的差异通常不是很大,对于那些差异很小且不妨碍正确检出运动目标的像素或是没有发生变化的像素无需更新,只需更新那些发生较大变化的像素值。

帧间差分对运动变化十分敏感。因此本文使用了三帧差分法来检测发生明显变化的像素,然后只对这些像素用式(4)进行更新。具体步骤如下:

Step1:计算当前帧与前后两帧之间的差分,得到两幅差分图像;

$$\text{Diff}^-(x, y) = |I_n(x, y) - I_{n-1}(x, y)| \quad (5)$$

$$\text{Diff}^+(x, y) = |I_n(x, y) - I_{n+1}(x, y)| \quad (6)$$

Step2: 对 $\text{Diff}^-(x, y)$ 、 $\text{Diff}^+(x, y)$ 进行二值分割, 得到二值图像 $\text{Diff}_B^-(x, y)$ 和 $\text{Diff}_B^+(x, y)$;

在阈值取值的问题上可以使用文献[11]中自适应最佳阈值的方法以提高分割的鲁棒性和准确性。但是这种方法求得的最佳阈值在光照突变时的效果往往不好, 因为光线突变会使前后两帧中像素的均值变化, 原来归于背景的像素可能会在突变后归于前景。这就使文献[11]中用于迭代计算最佳阈值的条件发生了变化。为了抵消这种变化引起的错分, 本文将连续两帧像素变化的平均值作为一个修正因子对最佳阈值进行修正, 修正因子定义为式(7)、(8):

$$\Delta^- = \frac{1}{M \times N} \times \sum_{x=0}^{n-1} \sum_{y=0}^{M-1} \left| \text{Diff}^-(x, y) - E[\text{Diff}^-(x, y)] \right| \quad (7)$$

$$\Delta^+ = \frac{1}{M \times N} \times \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \left| \text{Diff}^+(x, y) - E[\text{Diff}^+(x, y)] \right| \quad (8)$$

其中, M 、 N 分别表示图像的宽和高, $E[\text{Diff}^-(x, y)]$ 与 $E[\text{Diff}^+(x, y)]$ 分别表示前向和后向差分图像的均值。

$$E[\text{Diff}^-(x, y)] = \frac{1}{N \times M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} |\text{Diff}^-(x, y)| \quad (9)$$

$$E[\text{Diff}^+(x, y)] = \frac{1}{N \times M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} |\text{Diff}^+(x, y)| \quad (10)$$

经过修正后的阈值为:

$$T^- = Q_{th}^- + \Delta^- \quad (11)$$

$$T^+ = Q_{th}^+ + \Delta^+ \quad (12)$$

Step3: 确定当前帧中发生运动或像素变化剧烈的区域 $R(x, y)$, $R(x, y) = \text{Diff}_B^-(x, y) \cap \text{Diff}_B^+(x, y)$;

Step4: 进行背景模型更新, 对于 $R(x, y)$ 区域内的像素直接使用前一时刻的背景, 对于其他区域则按式(4)进行更新。

$$B_n(x, y) = \begin{cases} B_{n-1}(x, y), & \text{if } R(x, y) = \text{TRUE}; \\ \alpha \cdot B_{n-1}(x, y) + (1 - \alpha) \cdot I_n(x, y), & \\ \text{if } R(x, y) = \text{FALSE}. \end{cases} \quad (13)$$

4 运动目标检测

获得了背景图像后就可以使用背景减除的方法进行运动目标的检测。设当前帧图像为 $I_n(x, y)$, 当前的背景为 $B_n(x, y)$, 经过背景减除, 并二值化后的结果为:

$$\text{Obj}(x, y) = \begin{cases} \text{TRUE}, & \\ \text{if } |I_n(x, y) - B_n(x, y)| > T; \\ \text{FALSE}, & \\ \text{if } |I_n(x, y) - B_n(x, y)| \leq T. \end{cases} \quad (14)$$

其中, T 是判别阈值, 可以通过实验或文献[11]的方法得到。经过二值化后的图像可能在目标区域内出现“空洞”或在目标区域之外存在噪声, 因此有必要对式(14)的结果做进一步的后处理, 后处理主要包括连通域检测、空洞填充、膨胀和腐蚀运算。经过这些后处理之后, 就能得到令人满意的运动检测效果。

5 实验结果及说明

本文实验视频序列为一段室外交通视频, 视频分辨率大小为 320×240 , 仿真平台为 Matlab7.0, 计算机配置为 Intel Pentium Dual 1.6 GHz 处理器, 1 Gbyte 内存。根据本文的算法, 取前 100 帧进行背景建模。实验结果如图 1 所示。

图 1(a) 是视频序列的初始帧, 初始场景是一幅前景与背景混合的图像; (b) 是 100 帧建立的均值背景, 可以看到 (b) 图比较模糊, 同时在 (b) 右上角椭圆表示的区域中存在明显的物体运动轮廓留下的痕迹; (c) 是本文提出的方法建立的背景模型, 由于 (c) 中绝大多数小块都是来自各帧中, 所以 (c) 与 (b) 相比更加清晰。同时右上角处的痕迹也消失了; (d) ~ (f) 是 125 ~ 127 连续的 3 帧, 在这 3 帧中出现了运动目标; (g) 是用本文方法检测出的第 126 帧中发生变化的像素区域; (h) 是背景在第 126 帧更新的结果; (i) 是减除背景及一系列后处理得到的结果; (j) 是用 (i) 作为模板在第 126 帧中找到的运动目标。

从图 1 可见, 本文提出的基于分块重构背景的背景建模方法建立了较好的背景模型, 能够准确、有效地检测出运动目标。

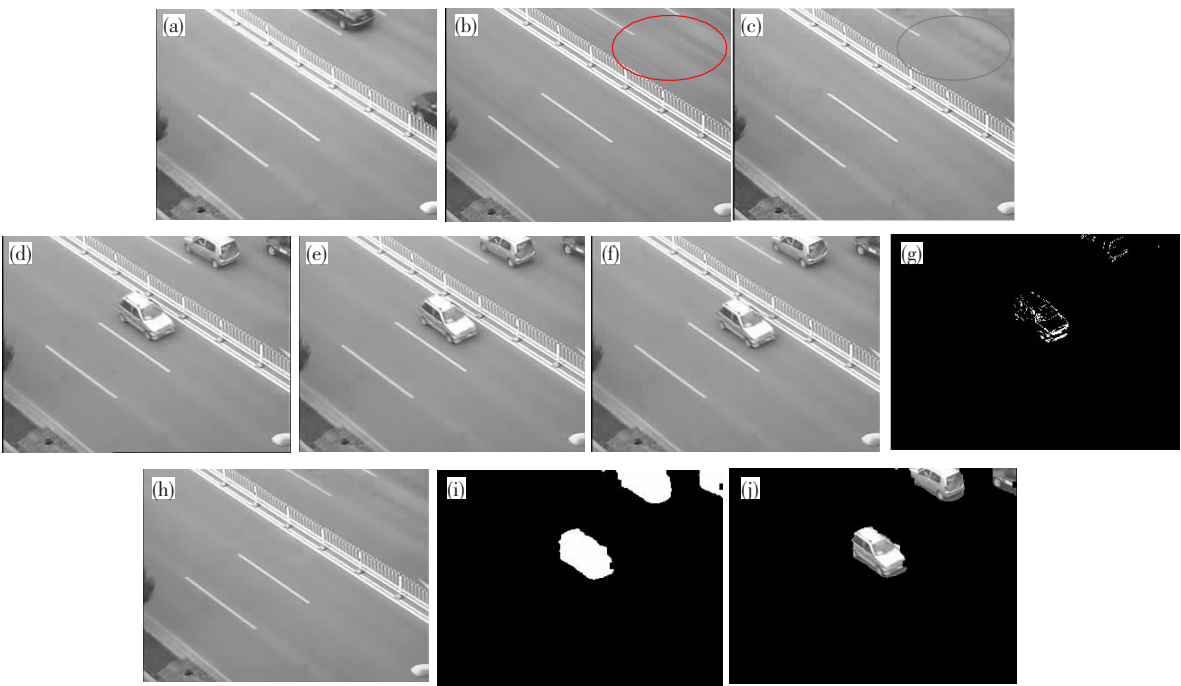


图 1 仿真实验结果

Fig.1 Simulation experiment results

6 结 论

研究了在运动目标检测的背景减除法中初始背景的建立以及更新的方法。通过比较分析目前常用的几种背景建模方法的优缺点,提出了一种基于分块重构背景的背景建模方法。新算法实时

性好,易于实现,克服了以往一些算法的缺点,能够建立较好的背景模型。在对背景模型更新时,综合了三帧差分检测对运动敏感的特点,并在对差分图像的分割中引入最佳阈值的修正因子抵消因光线突变造成的影响。实验结果证明,本文的方法能够准确、有效地检测出运动目标。

参 考 文 献:

[1] Lipton A, Fujiyoshi H, Patil R S. Moving target classification and tracking from real-time video [J]. *IEEE Workshop Application of Computer Vision*, 1998, 17(9):8-14.

[2] Barron J. Fleet D, Beauchemin S. Performance of optical flow techniques [J]. *International J. Computer Vision*, 1994, 12(1):42-77.

[3] Jain R. Difference and accumulative difference pictures in dynamic scene analysis[J]. *Image and Vision Computing*, 1984,2(2):99-108.

[4] 吴亮,马宏兴. 交通信息的背景提取技术分析 [J]. *西安科技大学学报*, 2010, 30(1):117-122.

[5] 刘扬,张云峰,董月芳. 复杂背景下抗遮挡的运动目标跟踪算法[J]. *液晶与显示*, 2010,25(6):890-895.

[6] 苏礼坤,陈怀新. 中值滤波的视频背景更新 [J]. *光电工程*, 2010, 37(1):131-135.

[7] 李轶博,李小兵,周娴. 基于 FPGA 的快速中值滤波器设计与实现[J]. *液晶与显示*, 2010,25(2):292-296.

[8] 康晓晶,吴谨. 基于高斯背景建模的目标检测技术 [J]. *液晶与显示*, 2010, 25(3):454-459.

[9] 刘士荣,王凯,邱雪娜. 基于自适应混合高斯模型全方位视觉目标检测 [J]. *浙江大学学报(工学版)*, 2010,44(7):1387-1393.

[10] 范立南,韩晓微,王忠石,等. 基于多结构元的噪声污染灰度图像边缘检测研究 [J]. *武汉大学学报(工学版)*, 2003, 36(3):86-90.

[11] Gonzalez R C, Woods R E, Eddins S L. 数字图像处理(MATLAB 版)[M]. 阮秋琦,译. 北京:电子工业出版社, 305-306.