

分块帧差和背景差相融合的运动目标检测*

高美凤, 刘 娣

(江南大学 轻工过程先进控制教育部重点实验室, 江苏 无锡 214122)

摘 要: 提出了一种分块帧差和背景差相融合的运动目标检测方法。该方法利用图像分块建立初始背景模型, 将视频图像划分为多个子块, 对帧间差分图像的各子块进行自适应阈值检测, 完成运动目标的粗分割, 采用双阈值背景差分法和邻域背景差分法对粗分割出来的运动区域进行细分割。背景采用自适应更新方法, 能够克服光照变化和背景干扰。实验结果表明, 该方法运算速度快、鲁棒性好, 能够准确检测出运动目标。

关键词: 运动检测; 背景重建; 分块帧差; 自适应阈值; 背景差分; 邻域背景差分

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1001-3695(2013)01-0299-04
doi:10.3969/j.issn.1001-3695.2013.01.077

Moving object detection based on consecutive blocks frame difference and background subtraction

GAO Mei-feng, LIU Di

(Key Laboratory of Advanced Process Control for Light Industry of Ministry of Education, Jiangnan University, Wuxi Jiangsu 214122, China)

Abstract: This paper proposed a novel approach for moving detection, which employed blocks frame difference and background subtraction. It used block processing method to establish the initial background model and it divided the sequence image of video into several blocks which was detected by self-adaptive threshold inter-frame difference, and separated roughly the motion region. Then it carried on fine-grained segmentation through double-threshold background subtracting for the motion region, and used the adaptive background updating to overcome the light changing and background interference. Experimental results indicate that the method has the characteristics of fast operation and great robustness, and it can detect the moving object effectively.

Key words: moving object detection; background reconstruction; blocks frame difference; self-adaptive threshold; background subtraction; neighborhood background subtraction

0 引言

运动目标检测是视频处理的关键技术,它涵盖了图像处理、模式识别与信号处理、控制理论、生物科学等多学科的理论知识,并在智能监控系统、人机交互系统等领域得到了广泛应用^[1]。对于固定摄像机捕捉的视频图像,运动目标检测的主要方法有光流法^[2]、帧间差分法^[3]、背景差分法^[4]等。其中,光流法的计算量较大,对硬件要求较高,一般很难实现对目标的实时检测;帧间差分法尽管能有效地去除静止的背景,但往往提取的目标比较粗糙,比实际的运动目标轮廓要大,并且目标中会出现空洞现象;而背景差分方法中对背景建模和背景更新要求较高,需要快速地建立背景图像并对背景进行及时更新,以保证具有对光照、干扰等环境变化的适应能力。

背景差分法一直是运动目标检测技术中研究较为广泛的方法,其关键是背景的构建和更新。常用的背景建模方法可以分为基于像素的和基于块的背景建模方法。基于像素的背景建模方法主要是根据每个像素在时间域上的分布信息构造背景模型,如 Stauffer 等人^[5]提出对单个像素的灰度值采用混合

高斯建模(MoG),该方法是按照高斯分布对每个像素建立模型,并通过基于回归滤波的在线 EM 近似方法对模型参数进行更新,它能鲁棒地克服光照变化、树枝摇动等造成的影响。基于块的背景建模方法是将每帧图像分割为多个重叠或非重叠的小块,然后使用每个块的特征构造背景模型。与单个像素相比,块能够获得更多的空间分布信息,使得基于块的背景建模方法对场景的局部变化不敏感。但是,基于分块的目标检测方法获得的目标比较粗糙,不适于对运动目标分割要求较高的场合,如果能够充分地融合背景在时间及空间上的分布信息,将基于像素的和基于分块的检测方法结合起来,则可以获得更为鲁棒的检测结果。Chen 等人^[6]提出了一种分级的背景建模方法,在利用基于块的背景模型对运动目标检测的基础上,用基于像素的背景模型对运动目标作进一步检测,以获得准确的目标外形。文献[7]研究分析已有的高斯分量数目自适应算法,同时利用对背景已经提取出的区域信息,在不包含目标的区域内采用基于像素块的检测方法,提高了检测速度。

本文提出了一种基于分块帧差和背景差相融合的运动目标检测方法。该方法利用图像分块原理建立初始背景模型,运

收稿日期: 2012-05-18; 修回日期: 2012-07-04 基金项目: 国家自然科学基金资助项目(61104092)

作者简介: 高美凤(1963-),女(蒙古族),江苏镇江人,副教授,博士,主要研究方向为信号处理、无线传感网技术(mfgao2008@gmail.com); 刘娣(1986-),女,硕士研究生,主要研究方向为图像处理、目标检测与跟踪。

动目标检测时,a)对相邻帧各子块进行自适应阈值差分,完成运动区域的粗分割,克服了帧差法提取目标空洞的现象;b)采用双阈值背景差分法和邻域背景差分法对粗分割出来的运动区域进行目标细分,精确检测出运动目标,从而将基于像素的和基于分块的目标检测方法结合起来;c)在背景更新时,采用一种自适应的背景更新策略,能够有效地适应不同光照条件下的检测场景,有效地克服了运动检测中的噪声干扰。实验部分比较了本文方法与传统的混合高斯模型方法以及文献[8]的方法在多个典型场景下的运动目标检测性能。

1 初始背景建模

基于像素的背景建模以像素为单位,每个像素被看成是相互独立的随机变量,被单独地决策为前景或是背景。这种方法虽然能够获得较细致的运动目标形状,但是没有充分利用像素在空间上的分布信息。事实上,对于没有目标的区域,图像的结构具有稳定性,对单个像素进行分析会产生冗余,而对相邻区域内的多个像素进行整体处理自然会提高背景建模的效率。所以可以利用一种图像分块处理的背景建模算法,其具体步骤如下:

- a)读入视频序列前 N 帧图像 $I_n (n=1,2,\dots,N)$,用均值法建立背景图像 I_0 。
- b)令迭代参数 $n=0$ 。
- c)对输入图像 I_n 进行 Sobel 边缘提取并分割成二值图像记为 I_n^* ,将 I_n 和 I_n^* 均分成 8×8 的子块 $I_n(i)$ 和 $I_n^*(i)$ (i 为子块的序列号),遍历图像 I_n^* ,求每一子块中的边缘点的个数 $\text{num}\{I_n^*(i)\}$ 。
- d)当 $n=0$ 时, $n=n+1$,转入 c);否则转入 e)。
- e)计算对应子块边缘点数差值 $d_n(i) = |\text{num}\{I_n^*(i)\} - \text{num}\{I_0^*(i)\}|$,当 $n=1$ 时,令 $d_{\min}(i) = d_1(i)$;当 $n>1$ 时, $d_{\min}(i) = \min\{d_n(i), d_{\min}(i)\}$ 。遍历子块; $n=n+1$,转入 c)至 $n=N$ 结束。
- f)如果 $d_{\min}(i) < th$,记 $d_{\min}(i)$ 取值所对应的迭代参数为 k ,则令 $I_0(i) = I_k(i)$;否则, $I_0(i)$ 不变。
- g)将 I_n 整合为图像,结束算法。

从上述步骤可以看出,背景建模方法结合了分块法和均值法,算法简单、计算量较小。利用均值法得到的背景虽然模糊,但是用 Sobel 检测的边缘点个数较少,可以认为不含运动目标。当运动目标存在时,检测的边缘点个数相对增加,因此可以将检测出来的边缘点个数进行比较,如果两者的最小差值小于阈值 th ,那么可以认为该子块与背景最为接近,用该子块替代均值背景中对应的子块;反之,均值背景中的子块保持不变。实验证明,在利用一段图像序列进行背景建模时,均值背景中的大部分子块都被图像序列中对应的子块替换了,所以这种背景建模方法能够克服均值背景法模糊的缺点。图 1 为均值法、文献[8]方法、高斯法和分块法分别建立的背景(均取图像序列前 50 帧)。

从图 1 中可以看出,均值法和高斯法建立的背景有拖影现象;文献[8]方法和本文方法建立的背景较好,但本文的算法更简单,运算速度较快。

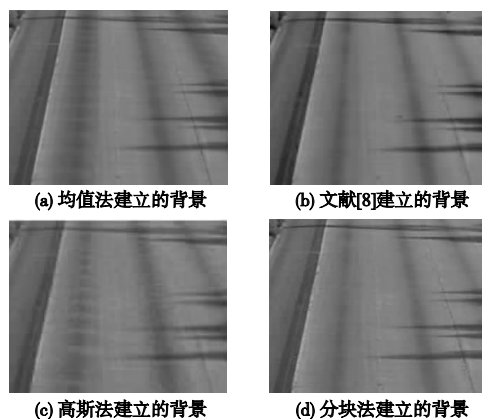


图1 构建背景的比较

2 目标检测

在摄像机位置固定不动的情况下,检测运动目标的背景也是基本固定的,只会产生局部细微的变化,其他大部分的背景保持不变。一般情况下,背景中会有许多像素点具有相同的像素值,如背景中的墙壁、马路、天空,运动目标的介入只是引起背景像素的局部变化,背景结构是相对稳定的。本文在序列图像分块的基础上,利用帧差法能有效分离背景的优势,对各子块图像进行帧差,完成对运动图像进行粗分割,分离出运动区域和非运动区域,然后对粗分割出来的运动区域采用背景差分法进行细分,检测出精确的运动目标。

2.1 分块帧差粗分割

设 $I_k (k=1,2,\dots)$ 为图像序列, $I_k(i,j)$ 表示第 k 帧图像在 (i,j) 点的像素值。帧间差分图像表示为

$$D_k(i,j) = |I_k(i,j) - I_{k-1}(i,j)| \quad (1)$$

在普通帧间差分法中,若 $D_k \leq T$ 即图像的变化不大,则认为无运动目标;若 $D_k > T$ 即图像的变化足够大,则认为有运动目标。其中, T 为检测阈值。式(1)反映了所有像素的共同特征,本文需要检测的是运动目标引起的像素变化,这是一种局部特征,因此需要一个对局部变化较敏感的判别式。

把相邻两帧图像 I_k 和 I_{k-1} 的差分图像 D_k 划分为 $m \times n$ 个子块,划分后的子块表示为 $D_k(p,q) (p=1,2,\dots,\text{int}(M/m), q=1,2,\dots,\text{int}(N/n))$,然后对每一差分分子块求和,并根据检测阈值 T_1 来判断子块属于前景或是背景。

$$D'_k(p,q) = \sum_{i=(p-1)m+1}^{pm} \sum_{j=(q-1)n+1}^{qn} D_k(i,j) \quad (2)$$

其中: $D_k(i,j)$ 表示差分图像 (i,j) 点的像素值; $D'_k(p,q)$ 表示差分图像 (p,q) 子块的像素值之和。如果 $D'_k(p,q) \leq T_1$,那么子块内的灰度变化没有超出噪声水平,则判定该子块为非运动区域,将子块内的像素值全部置 0;如果 $D'_k(p,q) > T_1$,那么子块内的灰度变化超出噪声水平,即有运动目标,则判定该子块为运动区域,并将子块内的像素值全部置 1。

$$D_k(i,j) = \begin{cases} 0 & D'_k(p,q) \leq T_1 \\ 1 & D'_k(p,q) > T_1 \end{cases} \quad (3)$$

在这一步中检测阈值 T_1 的确定是关键,在本文算法中把噪声和干扰认为是均匀分布的,可以定义为

$$T_1 = \frac{mn}{MN} \sum_{p=1}^{\text{int}(M/m)} \sum_{q=1}^{\text{int}(N/n)} D'_k(p,q) \quad (4)$$

当图像的信噪比高或背景波动较大时, $D_k'(p, q)$ 的值较大, T_1 的值也较大;反之 T_1 的值也就较小。 T_1 的这种自适应性能确保阈值能够将噪声、干扰和运动目标正确区分。

2.2 背景差分分割

经过分块帧间差分图像后,虽然可以将运动目标区域分割出来,但同时也将运动区域扩大了。为了更精确地检测出运动目标,本文采用双阈值背景差分法,只对分块帧差分出来的运动区域作处理。设置高阈值 T_H 和低阈值 T_L ,分别对差分图像进行二值化,得到两幅图像 D_k^H 和 D_k^L ,再对两图像作与操作,算法表示为

$$D_k^H(i, j) = \begin{cases} 1 & \text{if } |I_k(i, j) - B_k(i, j)| > T_H \text{ \& } D_k(i, j) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

$$D_k^L(i, j) = \begin{cases} 1 & \text{if } |I_k(i, j) - B_k(i, j)| > T_L \text{ \& } D_k(i, j) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

$$D_k^{BS}(i, j) = D_k^H(i, j) \cap D_k^L(i, j) \quad (7)$$

在室外环境下,背景常常表现为邻域运动,为了避免将背景本身的运动检测为运动目标,可以考虑像素邻域的背景差分,表示为

$$D_k^{NBS}(i, j) = \begin{cases} \min |I_k(i, j) - B_k(i + l_1, j + l_2)| & \text{if } D_k^{BS}(i, j) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

其中: l_1 和 l_2 为邻域半径,通常取 1 或 2。将 D_k^{NBS} 二值化可以得到最终的目标区域,二值化阈值为 T_2 。

经过背景差分的细分分割后,所获得运动目标的二值化模板中仍然有可能包含噪声点和空洞现象,因此还必须进行形态学操作等后续处理,才能得到真正精确的运动目标模板。

图 2 为运动目标检测粗分割和细分分割效果图。其中, (c) 为采用自适应阈值分块差分图像结果, (d) 为其细分分割结果; (e) (g) 为采用固定阈值分块差分图像结果, (f) (h) 分别为对应的细分分割结果。从粗分割结果可以看出,采用固定阈值分割图像的效果受阈值影响较大,阈值要手动设置,凭经验取值;而自适应阈值分割图像不但效果较好,还可以适应不同环境下图像的检测。从细分分割结果可以看出,本文采用的双阈值背景差分法和邻域背景差分法检测出来的运动目标更精确,噪声更小。



图2 运动目标检测结果

3 背景模型更新

由于光照、场景变化等各种因素的影响,背景模型建立后不可能一成不变。要使背景能够对外界环境的变化具有自适应性,从而提高运动检测的鲁棒性,必须对背景进行更新。背景更新即是用当前帧中发生变化的信息去修正过去帧建立的

背景模型。可以利用当前帧 I_k 和背景图像 B_k ,采用自适应背景更新模型^[9]得到更新背景图像 B_{k+1} ,更新模型为

$$B_{k+1}(i, j) = \begin{cases} B_k(i, j) & D_k^{NBS}(i, j) = 0 \\ \alpha \times B_k(i, j) + (1 - \alpha) \times I_k(i, j) & D_k^{NBS}(i, j) = 1 \end{cases} \quad (9)$$

其中: α 为更新权值,直接影响背景模型更新的快慢, α 越大表示背景更新的速度越慢;反之,则越快。为了对光照变化进行背景的自适应更新,取相邻两帧图像的灰度值的均值分别为 E_k 和 E_{k-1} ,则 α 计算方法为

$$\alpha = \beta \left(1 - \frac{|E_k - E_{k-1}|}{\max(E_k, E_{k-1})} \right) \quad (10)$$

其中: β 是常量,可以设置在 $[0.05, 0.4]$ 内。

4 实验结果与分析

为验证算法的有效性,本文对复杂场景下的视频进行了实验,实验环境为 CPU E3200/2.4 GHz、内存为 1 GB 的 PC 机,使用 MATLAB 7.0 编程实现。本文分别用视频序列连续帧和五个典型视频序列对算法进行了测试,其中五个视频序列包括两个室内场景和三个室外场景。视频序列存在着运动模糊、背景运动和噪声较大的情况。

分别使用文献[8]方法、GMM 方法和本文方法对这些场景中运动目标进行检测。图 3 给出了连续帧的检测结果,图 4 给出了每个序列的典型图像检测结果,检测结果未进行任何形态学处理。参数设置为: $th = 5$; $m \times n = 8 \times 8$; $T_H = 15$; $T_L = 10$; $T_2 = 20$; $l_1 = l_2 = 2$; $\beta = 0.05$ 。

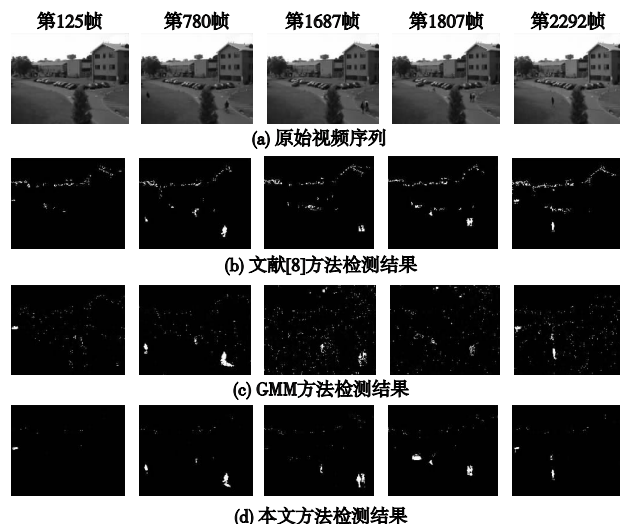


图3 视频序列连续帧检测结果的比较

图 3 是停车场的一段复杂场景,视频序列的背景中存在摇动的树木、汽车的移进移出和着光照的缓慢变化的情况,而且视频噪声干扰较大。从检测结果可以看出,本文方法比文献[8]方法和 GMM 方法检测的结果噪声小,运动目标更完整。文献[8]方法由于采用固定阈值差分图像,很容易受到噪声的干扰,目标粗分割时出现缺块现象,导致目标检测不完整。GMM 方法检测噪声较大,对物体的移进移出不能及时反应。本文方法采用自适应阈值分块差分图像和双阈值背景差分图像,能够降低场景的噪声干扰,对场景中物体的移近移出作出及时反应。

由图 4 可以看出,对于不同场景,本文方法也能够准确、完

整地分割出运动目标,本文方法与文献[8]方法、GMM方法相比具有较好的鲁棒性。

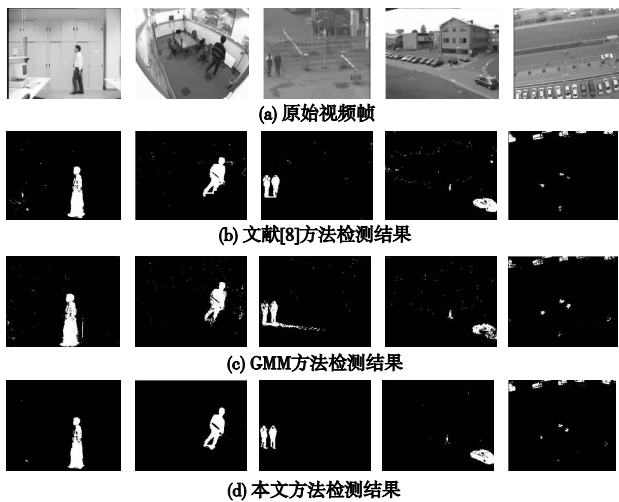


图4 多个视频序列检测结果的比较

为了定量分析算法性能,从图4所示的序列2中手工分割出109幅包含准确目标区域的掩模图像作为比较基准,并采用文献[10]中定义的查全率(recall)、查准率(precision)和 F 量测三个指标对文献[8]、GMM和本文算法的检测结果进行定量比较。设 t_p 表示检测目标正确像素个数, f_p 表示将背景错误检测为目标的像素个数, f_n 表示将运动目标错误检测为背景的像素个数。查准率、查全率和 F 量测分别定义为

$$\text{precision} = \frac{t_p}{t_p + f_p}, \text{ recall} = \frac{t_p}{t_p + f_n}, F = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

其中,查准率反映的是虚警率,值越大,虚警率越低;查全率反映的是漏警率,值越大,漏警率越低; F 量测反映的是整个算法的综合性能。

利用三种算法的检测结果,计算序列每帧的三个性能指标,并绘制成曲线图进行比较,如图5所示。

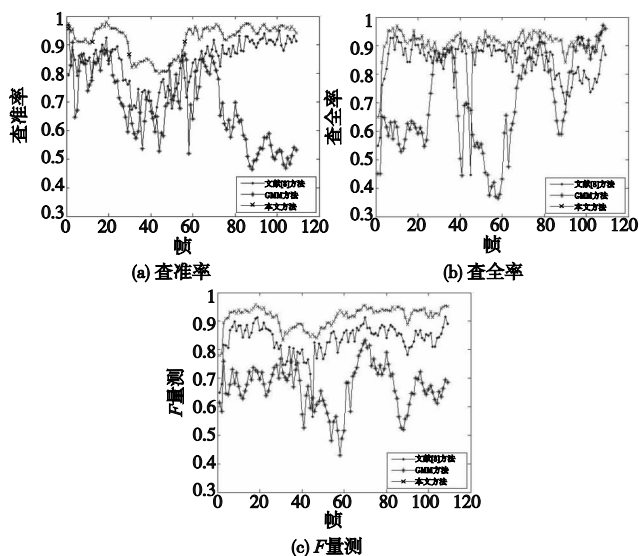


图5 三种方法检测性能曲线

从图5可以看出,本文方法的查准率高于文献[8]和GMM方法;查全率略高于文献[8],明显高于GMM方法;本文算法的综合性能明显高于文献[8]和GMM方法。

最后给出了三种方法在图4所示的五个图像序列上平均处理时间的对比,结果如表1所示。

表1 三种方法处理时间对比/帧/s

图像序列	文献[8]方法	GMM方法	本文方法
camera1 (320 × 240)	0.051	1.596	0.063
camera2 (320 × 240)	0.056	1.537	0.068
camera3 (352 × 288)	0.065	2.027	0.078
camera4 (384 × 288)	0.073	2.180	0.086
camera5 (320 × 240)	0.052	1.520	0.062

从表1中可以看出,本文方法的平均处理时间比文献[8]略长,比GMM方法要快。这是因为本文利用图像分块处理,对运动目标进行粗分割,过滤了大部分的背景区域,这种区域级的处理方法比GMM方法的像素级处理方法要快。虽然采用的自适应阈值和双阈值分割图像方法比文献[8]时间上略有增加,但是换来了更好的检测效果。影响本文方法速度的主要因素是分块的数目、样本图像数目和分块闭运算。

5 结束语

运动目标检测在视频监控等很多领域中具有广泛的应用价值,稳定可靠的运动目标检测算法是计算机视觉高层理解的重要前提。本文提出的分块帧差和背景差相融合的运动目标检测方法,在有效地重建背景的基础上,通过分块处理减少了目标检测的噪声,在差分图像时采用了自适应阈值和双阈值分割图像,不仅能够准确完整地提取运动目标,而且对光线、背景干扰不敏感,具有良好的鲁棒性。该方法提高了目标检测的处理速度,能够满足实时的检测目的。

参考文献:

- [1] 甘新胜,赵书斌. 基于背景差的运动目标检测方法比较分析[J]. 指挥控制与仿真,2008,30(3):45-50.
- [2] SENST T, EVANGELIO R H, SIKORA T. Detecting people carrying objects based on an optical flow motion model[C]//Proc of IEEE Workshop on Applications of Computer Vision. Washington DC:IEEE Computer Society,2011:301-306.
- [3] WENG Mu-yun, HUANG Guo-ce, DA Xin-yu. A new interframe difference algorithm for moving target detection[C]//Proc of the 3rd International Congress on Image and Signal Processing. 2010:285-289.
- [4] MOHAMED S S, TAHIR N M, ADNAN R. Background modeling and background subtraction performance for object detection[C]//Proc of the 6th International Colloquium on Signal Processing and Its Applications. 2010:236-241.
- [5] STAUFFER C, GRIMSON W E L. Learning patterns of activity using real-time tracking[J]. IEEE Trans on Pattern Analysis & Machine Intelligence,2000,22(8):747-757.
- [6] CHEN Yu-ting, CHEN Chu-song, HUANG Chun-rong, et al. Efficient hierarchical method for background subtraction[J]. Pattern Recognition,2007,40(10):2706-2715.
- [7] 陈明生,梁光明,孙即祥,等. 利用时空背景模型的快速运动目标检测方法[J]. 中国图象图形学报,2011,16(6):1002-1007.
- [8] 胡建华,杨帆,徐健健. 基于分块差分与背景减除的运动检测[J]. 电子测量技术,2007,30(10):49-53.
- [9] LEE S, LEE J, HAYES M H, et al. Adaptive background generation for automatic detection of initial object region in multiple color-filter aperture camera-based surveillance system[J]. IEEE Trans on Consumer Electronics,2012,58(1):104-110.
- [10] MADDALENA L, PETROSINO A. A self-organizing approach to background subtraction for visual surveillance applications[J]. IEEE Trans on Image Processing,2008,17(7):1168-1177.