

一种改进的 ViBe 背景提取算法

张夏清, 茅耀斌

(南京理工大学 自动化系, 江苏 南京 210094)

摘要:背景生成算法通过提取视频场景背景的方法,检测出运动的前景对象,在智能视频监控应用中具有重要的意义。针对 ViBe 背景提取算法,文中提出一种改进方法。在样本采集方面,提出了利用高斯分布在像素点的邻域中进行采样的方法;在样本更新方面,提出自适应概率的随机子采样方法;在背景生成方面,构建样本集的像素直方图,提出了通过颜色直方图和 Mean Shift 迭代进行背景生成的方法,并结合时空域的像素信息进行前景检测。具体描述了改进的 ViBe 背景生成算法,并通过实验比较了改进算法与 ViBe 背景提取算法的效果。结果表明,改进的 ViBe 算法在背景生成和前景提取方面均达到了更好的效果。

关键词:ViBe 算法;背景生成;目标检测;颜色直方图;均值漂移

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2016)07-0036-04

doi:10.3969/j.issn.1673-629X.2016.07.008

An Improved ViBe Background Generation Method

ZHANG Xia-qing, MAO Yao-bin

(Department of Automation, Nanjing University of Science and Technology,
Nanjing 210094, China)

Abstract: Background generation and foreground detection play a significant role in the field of intelligent video surveillance. By extracting the background from video scenes, the algorithm can detect the movement of the foreground objects. An improved background generation method based on ViBe background extraction is proposed. In terms of sample collection, samples are taken in the neighborhood of every pixel according to a Gaussian distribution. Regarding sample updating, an adaptive probability of random sub-sampling is presented. When it comes to background generation, by constructing a color histogram of the sample pixels, a background generation method is put forward based on color histograms and Mean Shift iteration. And the foreground is detected with the pixel information in both time and space domain. It specifically describes an improved ViBe background generation method in this paper, which is compared with the original algorithm by means of experiments. The results show that the proposed algorithm is more effective in both background generation and foreground extraction.

Key words: ViBe algorithm; background generation; object detection; color histogram; Mean Shift

0 引言

背景生成和前景对象检测是智能视频处理过程中的一个关键步骤,也是视觉监控研究的难点和热点之一。在生成背景的基础上,将运动的前景对象从变化的场景中提取出来,这对于目标跟踪、分类、行为理解以及视频浓缩、检索等智能监控领域中的应用都十分重要。

背景生成与前景提取从 20 世纪 90 年代开始被广泛研究,至今已经有很多成熟的方法,同时由于视频场

景和目标运动的复杂性,各方法也存在一定的不足。Stauffer 等首先提出了混合高斯背景建模的方法^[1],利用多个高斯分布混合的方式进行建模。Elgammal 等提出了一种基于核密度估计的非参数化背景建模方法^[2],通过计算像素差分的中值估计核概率密度函数的方差。Liu 等提出了一种影响因素描述的模型^[3],并用均值漂移算法^[4]计算最可靠背景状态。Kim 等提出了 codebook 编码本^[5]的背景建模方法,将属于背景的样本通过量化加入码本,以代表较长时间序列上的

收稿日期:2015-11-17

修回日期:2016-03-09

网络出版时间:2016-06-22

基金项目:国家自然科学基金资助项目(61472188)

作者简介:张夏清(1991-),女,硕士研究生,研究方向为视频图像处理、模式识别;茅耀斌,副教授,博士,通信作者,研究方向为视频图像处理、模式识别。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20160622.0845.070.html>

背景模型。

Barnich 在文献[6]中提出了 Visual Background extractor(视觉背景提取,后文称 ViBe)算法,这是一种典型的基于区域的背景提取方法,在前景检测方面效果较好,但算法对初始帧具有很大的依赖性,容易引入鬼影区域,造成目标错检,且没有提出确切的背景生成方法。文献[7]在(R,G,I)标准颜色空间进行目标检测;文献[8-10]利用帧间差法改进目标检测效果;文献[11-12]分别采用自适应的背景更新和前景判断阈值进行鬼影抑制。

针对 ViBe 算法中存在的问题,文中提出一种改进的背景提取方法:

- (1)在背景样本集构建方面,结合时间和空间的采样方法,提出基于高斯分布的邻域样本采集方式;
- (2)在样本更新方面,采用随机子采样的方法,提出自适应子采样概率的样本更新策略;
- (3)提出利用 Mean Shift 迭代得到背景收敛点的方法,并提出基于颜色直方图的背景初值确定以及更新判断准则,从而进行背景生成。

1 ViBe 算法

1.1 基本原理

ViBe 算法的重要思想是基于像素邻域信息的背景样本随机采样。设图像点 x 处的像素值为 $v(x)$,则该点处大小为 N 的背景样本集记作:

$$M(x) = \{v_1, v_2, \dots, v_N\} \quad (1)$$

$M(x)$ 则可以通过随机采样其时空域像素值而得到。

定义 $S_R(v(x))$ 为以 x 为中心、 R 为半径的球形区域,对新的像素值 x ,若其 $S_R(v(x))$ 区域包含了 $\#_{\min}$ 个以上的背景样本值,则将其归为背景,并更新其背景样本,即背景判断准则表示为:

$$\# \{S_R(v(x)) \cap M(x)\} \geq \#_{\min} \quad (2)$$

其中, R 、 $\#_{\min}$ 均为参数。

在背景样本集的构造方面,ViBe 算法在图像点的邻域像素中进行随机采样,由于利用了相邻像素的时空分布特性,仅仅通过一帧图像即可完成样本集初始化。在样本更新方面,ViBe 采用随机子采样的更新方法,背景样本以 $1/\varphi$ 的随机子采样概率更新图像点的样本集及其邻域点的背景样本集。

1.2 算法的改进思路

背景样本集的采集和更新,是背景生成和前景检测算法中的关键问题,关系着算法的实际效果。文中基于 ViBe 算法在这两方面都进行了相应的改进。在背景采集方面,提出基于高斯分布的方法,在像素点的时空域中进行采样,可以控制时间域和空间域像素在

背景样本集中占的比例,更能适应不同的实际场景;在背景更新方面,由于 ViBe 的初始化容易产生目标错检与鬼影现象,因此文中提出采用自适应概率的随机子采样方法,利用时变的随机子采样概率,可以改善因初始帧像素经多次采样且长期不更新而引起的错检问题。

通过研究发现,对于静止摄像头拍摄的视频,其背景可以定义为在一段较长时间内不会变化的场景。对单个像素位置来说,在大的样本集下,背景可以看成是在该点处出现概率密度最大的像素值。因此,文中采用颜色直方图和 Mean Shift 结合的方法进行背景生成,具体改进如下:

- (1)基于高斯分布进行邻域样本采集,利用图像点的时空信息进行背景样本集的构建;建立样本集上像素的颜色直方图,寻找最大直方图位,并计算用于 Mean Shift 迭代的起点像素值。
- (2)利用自适应的随机子采样概率进行背景样本的更新,并更新样本直方图。
- (3)根据直方图判断是否执行 Mean Shift 迭代过程,将收敛得到的最大概率密度点作为背景值;并根据 2.5 节的检测策略进行前景判断。

2 改进的 ViBe 算法

2.1 背景样本的采样

在样本采集方面,经典的建模算法是将图像中点的历史像素值作为背景样本值,逐帧采样组成样本集。进一步的研究表明,背景不仅和历史像素值有关,也和图像点邻域位置的像素有关^[6],即背景样本集应包含点的时间以及空间的像素值信息,利用时空像素值采样的方法构造背景样本集。

文中在构造背景样本集时利用了像素的时间和空间信息,不同于文献[6]在邻域中以随机等概率采样的方式,文中采用高斯分布的随机概率来采样以图像点为中心的 $N_c \times N_c$ 邻域。以 $P(x_0, y_0)$ 为中心的二维高斯分布 $N(0, \sigma^2)$ 如下:

$$G(x, y) = \frac{1}{s\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma^2}} \quad (3)$$

其中, $(x, y) \in N_c(x_p)$, 是点 P 的 $N_c \times N_c$ 邻域; s 是归一化系数,使 $\sum_{(x,y) \in N_c(x_p)} G(x, y) = 1$ 。

定义背景样本集的时空信息比例为 R_{st} , 表示时间信息样本在背景样本集中所占的比例,即

$$R_{st} = G(x_0, y_0) = \frac{1}{s\sqrt{2\pi\sigma^2}} \quad (4)$$

R_{st} 由高斯分布的方差以及邻域的大小决定。ViBe 算法^[6]中的样本是通过在 3×3 的像素邻域中等

比例采样而得到的,样本的重复采集对复杂场景容易产生较为严重的误检现象。文献[13]提出四菱形邻域的采样方法,将样本限制在大小为 5 的像素区域内,而这并不能适应变化的实际场景。文中提出的高斯邻域采样方法,在利用区域的时空像素信息的基础上,通过改变高斯分布的方差可以调节 R_s ,即控制采集样本时邻域像素所占的比例,从而达到适应不同应用场景的目的。

在构造背景样本集时,每个图像点的样本以 $G(x, y)$ 的概率从邻域 $N_c(x_p)$ 中选取,同样地,在样本更新的过程中,也分别利用该概率随机地更新当前点及其邻域点的样本集。

2.2 背景样本的更新策略

在样本更新方面,一般有两种典型方法:一种是基于先入先出的思想,另一种是基于随机子采样的更新方法,如文献[6]提出的历史样本剩余时间的指数衰减方法。虽然基于随机子采样的样本更新方法在前景提取方面取得了很好的效果,但由于初始帧的多次子采样,使得初始帧的信息占主导,又由于子采样概率的限制,需要更多的样本来缓解样本集的侧重问题,在这过程中就容易出现鬼影现象。

针对上述问题,文中提出自适应的随机子采样概率 φ_t 。根据视频帧数,在小样本情况下尽可能利用现有样本来构造和更新样本集,而在样本充足的情况下则以随机概率执行子采样。设 t 为当前视频帧数,则

$$\varphi_t = \begin{cases} \alpha/t, & t < N_\varphi \\ \varphi_N, & t \geq N_\varphi \end{cases} \quad (5)$$

其中, α 和 N_φ 为系数,与背景样本集的大小有关。

2.3 基于 Mean Shift 的背景生成

ViBe 算法^[6]可以较好地检测到视频中的前景目标,但是没有给出具体的背景生成方法。文中利用 Mean Shift 方法进行背景生成,提出利用直方图方法进行迭代起点的定位,以及迭代区域的变化检测。通过均值漂移方法进行背景生成达到了较好的效果,因此首先介绍基于 Mean Shift 的背景生成方法。

根据概率密度函数,将核密度估计函数写为:

$$f(x) = \frac{C_{k,d}}{Nh^d} \sum_{i=1}^N k\left(\left\|\frac{x - x_i}{h}\right\|^2\right) \quad (6)$$

其中, $C_{k,d}$ 是归一化系数。

令 $g(x) = -k'(x)$, 且令 $\nabla f(x) = 0$, 可得 Parzen 窗中概率密度的极大值点:

$$x = \sum_{i=1}^N x_i g\left(\left\|\frac{x - x_i}{h}\right\|^2\right) / \sum_{i=1}^N g\left(\left\|\frac{x - x_i}{h}\right\|^2\right) \quad (7)$$

Mean Shift 通过令 $\nabla f(x) = 0$ 进行迭代,不断得到新的中心,收敛到局部概率密度最大的点,从而得到背

景值。

2.4 基于颜色直方图的初值计算

Mean Shift 是一种基于核密度函数迭代求极值的方法,其运行效率与计算精度会依赖于所选取的迭代起点,因此,迭代初值的正确选取十分重要。文中利用颜色直方图的方法计算直方图的最大值,并将最大直方图的中值位作为 Mean Shift 的迭代起点。通过计算颜色直方图的方法进行背景值的粗定位,每个直方图的位数为 m 。

$$p^{(u)} = \frac{\sum_{i=1}^N \delta[h(x_i) - u]}{N}, u = 1, 2, \dots, m \quad (8)$$

$$U_{\max_u} = \operatorname{argmax}_u p_c^{(u)} \quad (9)$$

其中, $p = \{p^{(u)}\}_{u=1,2,\dots,m}$ 表示样本像素值的直方图,通过计算最大直方位,分别得到每个通道对应的像素位 U_{\max_u} ,而迭代起点 $x_0 = (r, g, b)$ 设为最大直方位所对应的像素中值。以像素范围为 256 的 RGB 图像为例,若直方图宽度为 d ,即 $m = (256/d)^3$,则

$$x_{0c} = U_{\max_u} \cdot d + d/2 \quad (10)$$

颜色直方图是根据背景样本集中的点建立的,随样本集的更新而更新,而直方图的最大像素值 x_0 表明了概率密度最大值所处的局部区域。因此,文中算法的基本思想是,通过颜色直方图定位到背景的迭代起点 x_0 ,再利用 Mean Shift 进行迭代直至收敛,得到的最大概率密度点即为背景值。

2.5 前景检测策略

文中算法通过 Mean Shift 方法进行背景生成,由于背景是样本集中概率密度最大的点,因此,越靠近背景值的样本,为背景的概率就越大。因此,根据样本值与当前背景的距离来判断是否为前景。设 Mean Shift 半径为 $R = h/2$,背景判断准则:

$$\text{background}(x) = \begin{cases} 1, & \|x - x_{bg}\|_2 \leq R_T \\ 0, & \|x - x_{bg}\|_2 > R_T \end{cases} \quad (11)$$

其中, $R_T = \gamma R$ 为判断阈值, γ 为比例系数,当新样本与背景值距离大于 R_T 时判断为前景,如图 1 所示。

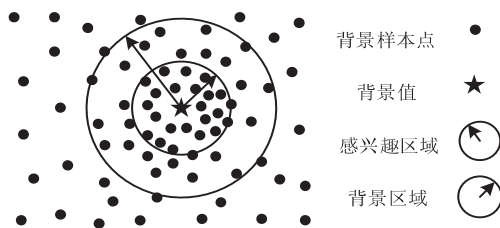


图 1 背景判断示意图

3 实验

3.1 实验环境与定性结果

文中实验在 Intel® Core(TM) i5 CPU、2.53 GHz、

内存为 4 GB 的笔记本上进行,操作系统为 Windows 7,开发环境是 Visual Studio 2010,利用 C++进行算法实现。将文中的改进算法与 ViBe 算法^[6]进行实验比较。为评估文中算法的有效性,需要选择具有人工标注的真实前景目标数据的视频序列进行实验分析。文中选取了三个实验室的 4 段广泛使用的视频序列:Room、Highway3(<http://arma.sourceforge.net/shadows/>),Office(<http://www.changedetection.net/>),Synthetic111(<http://bmc.iut-auvergne.com/?q=node/7>)。定性的实验结果如图 2 所示。

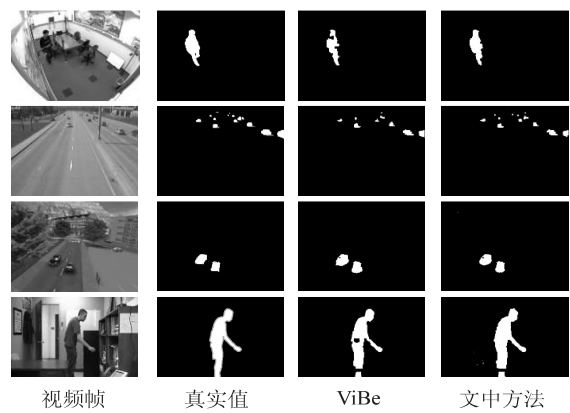


图 2 文中改进算法与 ViBe 算法对比

注:从上到下分别为 Room,Highway3,Synthetic111 和 Office

另一方面,文中提出了基于高斯分布的自适应随机子采样概率,能改善 ViBe 算法中的鬼影问题。以视频序列 highwayII(<http://arma.sourceforge.net/shadows/>)为例,第 25 帧图像的前景检测结果如图 3 所示。

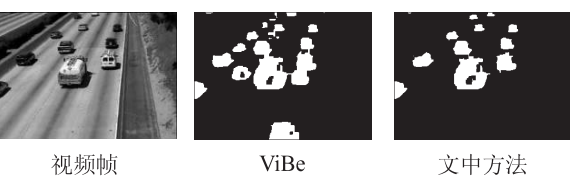


图 3 算法关于鬼影问题的比较

由图中可以看到,ViBe 算法由于初始帧的影响,在检测结果中出现了前景的“鬼影”现象,而该问题在文中算法得到了很好的抑制。

3.2 定量评估

为定量评估文中算法的检测效果,采用文献[13]所述的精度度量标准进行比较。一共 4 种精度值,分别为:Recall, Precision, F_1 , Similarity。其中,Recall 表示真实前景像素的检出率;Precision 表示真实前景像素的检准率; F_1 也称为品质因数,用于衡量检出率和检准率的总体水平;Similarity 是一种基于像素的相似性度量方法。

文中的定量评估实验分别在上述的 4 个视频序列上进行,分别计算两种背景生成算法在评测数据上的 4 个不同角度的精度值,结果如表 1~4 所示。

表 1 Room 的定量评估结果

算法	Recall	Precision	F_1	Similarity
ViBe ^[6]	0.80	0.93	0.85	0.75
文中算法	0.86	0.87	0.86	0.75

表 2 Highway3 的定量评估结果

算法	Recall	Precision	F_1	Similarity
ViBe ^[6]	0.65	0.98	0.76	0.64
文中算法	0.72	0.98	0.82	0.71

表 3 Synthetic111 的定量评估结果

算法	Recall	Precision	F_1	Similarity
ViBe ^[6]	0.89	0.84	0.86	0.76
文中算法	0.90	0.91	0.90	0.82

表 4 Office 的定量评估结果

算法	Recall	Precision	F_1	Similarity
ViBe ^[6]	0.36	0.71	0.37	0.30
文中算法	0.88	0.75	0.81	0.68

通过一系列的研究和实验,算法可以衡量为,在不牺牲目标检准率的前提下,尽量达到好的目标检出率^[14]。也就是说,好的算法应具有好的目标检测率,包括 Recall 和 Precision;与此同时, F_1 和 Similarity 用于衡量总体的检测水平,表示检出率和检准率的一致性,同样也需要尽量达到好的结果。从表中可见,通过定量的实验评估,在同等条件下,文中算法均能达到较好的目标检测效果和较高的目标检测率,这也验证了文中算法的可行性和有效性。

4 结束语

针对视频的背景生成和前景检测问题,文中提出了一种改进的 ViBe 算法。提出基于高斯分布的像素时空域采样方式,并将其用于背景样本集的构造,结合自适应概率的随机子采样方法,使得算法在初始帧和小样本的情况下,能较好地完成样本集的构造和更新。基于背景样本集,算法通过计算颜色直方图确定 Mean Shift 的迭代起点,并结合直方图判断是否进行背景迭代,从而收敛得到背景值。最后,将文中算法和 ViBe 算法的实验效果进行比较,通过定性和定量评估可知,文中算法能取得更好的背景生成和目标检测效果。

参考文献:

[1] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking[C]//Proceedings of the IEEE computer society conference on computer vision and pattern recognition. Los Alamitos, CA, USA: IEEE Computer Society, 1999:246-252.

数进行搜索,得到能量效率的最优值。仿真结果表明,该算法切实可行,能有效提高系统的能量效率,降低了算法的复杂度。

参考文献:

[1] Malta V. Telecommunication support for the protection of the environment[C]//Proc of world telecommunication development conference. [s. l.]:[s. n.],1998.

[2] Fettweis G P,Zimmermann E. ICT energy consumption trends and challenges[C]//Proc of 11th international symposium on wireless personal multimedia communication. [s. l.]:[s. n.],2008:1-4.

[3] Ge Xiaohu,Hu Jinzhong,Wang Chengxiang,et al. Energy efficiency analysis of MISO-OFDM communication systems considering power and capacity constraints[J]. Mobile Networks and Applications,2012,17(1):29-35.

[4] 喻的雄,蔡跃明,吴 丹,等. OFDMA 上行链路中基于博弈论的子载波和功率分配算法[J]. 电子与信息学报,2010,32(4):775-780.

[5] 张冬梅,徐友云,蔡跃明. OFDMA 系统中线性注水功率分配算法[J]. 电子与信息学报,2007,29(6):1286-1289.

[6] Bohge M,Gross J,Meyer M,et al. Dynamic resource allocation in OFDM systems: an overview of cross-layer optimization principles and techniques[J]. IEEE Network Magazine,2007,21(1):53-59.

[7] Jia Yizhen, Wang Youzheng, Lu Jianhua. Joint allocation of frequency and space resources for a cooperative OFDMA cellular system[C]//Proc of international conference on wireless communications, networking and mobile computing. [s. l.]:[s. n.],2009.

[8] Miao Guowang,Himayat N,Li G Y. Energy-efficient link adaptation in frequency-selective channels[J]. IEEE Transactions on Communications,2010,58(2):545-554.

[9] Christian I,Fettweis G P. Energy-efficient link adaptation with transmitter CSI[C]//Proc of wireless communications and networking conference. [s. l.]:IEEE,2011:1381-1386.

[10] Christian I,Chong Zhijiat,Jorswieck E,et al. Framework for link-level energy efficiency optimization with informed transmitter[J]. IEEE Transactions on Wireless Communications,2012,11(8):2946-2957.

[11] Schaible S. Fractional programming[J]. European Journal of Operations Research,1983,27(1):39-54.

[12] Schaible S,Ibaraki T. Fractional programming[J]. European Journal of Operational Research,1983,12(4):325-338.

[13] Dinkelbach W. On nonlinear fractional programming[J]. Management Science,1967,13(7):492-498.

[14] Ibaraki T. Parametric approaches to fractional programs[J]. Mathematical Programming,1983,26(3):345-362.

[15] Schaible S. Fractional programming II: on Dinkelbach's algorithm[J]. Management Science,1976,22(8):868-873.

[16] Boyd S, Vandenberghe L. Convex optimization[M]. Cambridge:Cambridge University Press,2004.

[17] Xiao Xiao,Tao Xiaoming,Jia Yizhen,et al. An energy-efficient hybrid structure with resource allocation in OFDMA networks[C]//Proc of wireless communications and networking conference. [s. l.]:IEEE,2011:1466-1470.

[18] Seong K,Mohseni M,Cioffi J M. Optimal resource allocation for OFDMA downlink systems[C]//Proc of IEEE international symposium on information theory. [s. l.]:IEEE,2006:1394-1398.

(上接第 39 页)

[2] Elgammal A M,Harwood D,Davis L S. Non-parametric model for background subtraction[C]//Proceedings of the 6th European conference on computer vision. Berlin,BE,GER:Springer-Verlag,2000:751-767.

[3] Liu Y,Yao H,Gao W,et al. Nonparametric background generation[J]. Journal of Visual Communication & Image Representation,2007,18(3):253-263.

[4] Comaniciu D,Meer P. Mean Shift: a robust approach toward feature space analysis[C]//Proc of IEEE transactions on pattern analysis and machine intelligence. Piscataway,NJ,USA:IEEE Computer Society,2002:603-619.

[5] Kim K,Chalidabhongse T H,Harwood D,et al. Real-time foreground-background segmentation using codebook model[J]. Real-time Imaging,2005,11(3):172-185.

[6] Barnich O,Vanogenbroeck M. ViBe: a universal background subtraction algorithm for video sequences[J]. IEEE Transactions on Image Processing,2011,20(6):1709-1724.

[7] 严红亮,王福龙,刘志煌. 基于 ViBe 算法的改进背景减去

法[J]. 计算机系统应用,2014,23(6):130-134.

[8] 胡小冉,孙 涵. 一种新的基于 ViBe 的运动目标检测方法[J]. 计算机科学,2014,41(2):149-152.

[9] 赵光明,韩 光,李晓飞,等. 基于融合帧间差的改进 Vibe 方法[J]. 计算机技术与发展,2015,25(3):76-80.

[10] 吴剑舞,翁玲瑜,童 怀. 一种基于改进 ViBe 的运动目标检测方法[J]. 计算机与现代化,2015(7):50-54.

[11] 王 辉,宋建新. 一种基于阈值的自适应 Vibe 目标检测算法[J]. 计算机科学,2015,42(6A):154-157.

[12] 周 晓,赵 锋,朱艳林. 基于 ViBe 的室外动态背景闪烁像素噪声消除方法[J]. 计算机应用,2015,35(6):1739-1743.

[13] 吕嘉卿,刘立程,郝禄国,等. 基于视觉背景提取的自适应运动目标提取算法[J]. 计算机应用,2015,35(7):2029-2032.

[14] Lucia M,Alfredo P. A self-organizing approach to background subtraction for visual surveillance applications [J]. IEEE Transactions on Image Processing,2008,17(7):1168-1177.