

# 一种复杂环境下运动目标检测的背景构建方法

王 枚<sup>1,3</sup> 苏光大<sup>2</sup> 王国宏<sup>3</sup>

(1 烟台职业学院图像处理研究所,烟台市滨海中路 2018 号,264670)  
(2 清华大学电子工程系,北京市清华园,100084)  
(3 海军航空工程学院电子信息工程系,烟台市二马路 188 号,264001)

**摘 要:**针对复杂环境下背景构建困难等问题,提出了一种彩色图像依照目标运动状态实时进行更新的背景构建方法。该方法通过定义彩色图像距离,根据当前图像与背景图像之间的距离变化定义运动状态矩阵,从而利用状态矩阵结合当前帧图像的彩色特征更新背景。该方法无需假设背景模型,能够适应各类复杂的突变,通过获取包含多类目标的实测彩色图像实验验证了算法的有效性和准确性。  
**关键词** 运动目标检测;背景差;运动状态矩阵;更新  
**中图法分类号:**P273.3

基于视频序列的运动目标检测一直以来都是机器视觉、智能监控系统和视频跟踪等领域的研究重点。现有的运动目标检测方法有时间差分法、光流法、背景差法等。其中,背景差法是固定的单摄像头监控中应用最广的运动目标检测方法。常见的背景差的背景构建方法主要有以下几种:① 基于复杂背景建模方法<sup>[1-4]</sup>;② 基于知识的方法<sup>[5,6]</sup>,如经典的均值、中值、运动平均等方法;③ 基于运动检测的方法,如运动光流法、显著性运动检测方法<sup>[7]</sup>等。方法①是对较长时间内的背景图像数据进行分析,但到目前为止,还没有一种模型可以描述各类运动的视频背景,并且随着视频图像帧数量的增大,算法复杂性和计算耗时都无法达到实时性要求;方法②可以较好地表示视频序列的整体特征,但缺少目标运动的时间特性;方法③是基于运动区域内的目标特征进行研究,不需要构造背景图像,但采用光流法的计算复杂性大,也无法达到工程的实用性要求。从整个背景构建方法的特点可以看出,由于算法的复杂度不断增加,而实际视频监控系统对实时性又不断提高,因此,一个快速的背景构建算法往往是决定监控系统性能的关键。本文从彩色视频帧图像的

时间序列考虑,跟踪目标的运动状态,由此定义运动状态矩阵,并实时更新彩色背景,能够很好地描述存在突变等扰动的复杂背景变化。

## 1 本文方法

视频序列中的每帧图像包含背景图像和运动图像两部分,设  $I_i(x,y)$  ( $i=1,2,\cdots,m$ ) 为图像序列中的第  $i$  帧图像,  $Bg_i(x,y)$  为该图像的背景图像,  $O_i(x,y)$  为该图像的运动图像。一般来说,当摄像机固定不动,背景图像基本不变化时,当前图像与背景图像和运动目标之间具有如下关系:

$$\begin{cases} I_i(x,y) = Bg_i(x,y) + O_i(x,y) \\ Bg(x,y) = \sum_{i=1}^m Bg_i(x,y) \end{cases} \quad (1)$$

其中,变量  $i$  ( $i=1,2,\cdots,m$ ) 代表图像的当前帧数,即当前图像在视频序列中的顺序; $m$  是序列图像的总帧数,代表视频序列长度。根据该思路设计,则视频图像目标检测的背景差方法可表示为:

$$O_i(x,y) = I_i(x,y) - Bg(x,y) \quad (2)$$

图 1 为背景差方法运动目标检测过程。其中,图 1(a)为视频序列任一帧图像,图 1(b)为本

文方法利用 40 帧序列图像获得的背景图像;图 1(c)是图 1(a)与 1(b)的差分图像;图 1(d)是经数学形态学去除小面积后运动目标的检测结果。由该过程可以看出,背景差方法的核心就是背景图像构建。

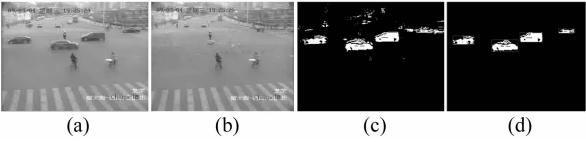


图 1 背景差方法运动目标检测过程  
Fig. 1 Process of Motion Object Detection of Background Subtraction

颜色是目标特征的重要描绘子,为表示视频序列中两帧彩色图像之间的差异,首先选取代表彩色图像每个像素的特征  $f(i) = \bigcup_{i \in [R,G,B]} f(i)$ ,即定义两幅彩色图像  $f_1$  和  $f_2$  之间的距离  $D_{f_1,f_2}$  为:

$$D_{f_1,f_2} = \|f_1 - f_2\| / (\|f_1\| + \|f_2\|) = \frac{[(r_1 - r_2)^2 + (g_1 - g_2)^2 + (b_1 - b_2)^2]^{1/2}}{(r_1^2 + g_1^2 + b_1^2)^{1/2} + (r_2^2 + g_2^2 + b_2^2)^{1/2}} \quad (3)$$

其中,  $r_1, g_1, b_1, r_2, g_2, b_2$  分别是彩色图像  $f_1, f_2$  的 RGB 分量值。

为了表示目标的运动状态,用当前图像与背景图像之间的变化次数来描述运动状态,定义目标运动状态矩阵  $S_i(x, y)$  为:

$$S_i(x, y) = \begin{cases} S_{i-1}(x, y) + 1, D_{I_i, Bg_{i-1}} < T \\ S_{i-1}(x, y) - 1, D_{I_i, Bg_{i-1}} \geq T \end{cases} \quad (4)$$

其中,  $D_{I_i, Bg_{i-1}}$  为当前图像  $I_i$  与由前  $i-1$  帧构建的当前背景图像  $Bg_{i-1}$  之间的距离;  $T$  为图像距离阈值;  $i=1, 2, \dots, m$ 。

从本质上讲,基于帧时序运动状态矩阵的背景构建方法就是跟踪目标运动状态,实时更新背景的方法。该方法具体描述如下:

1) 初始化背景及运动状态矩阵。视频序列图像中由于有目标运动,造成背景被遮挡部分也随之变化。采用计算当前图像与背景图像之间的距离判断是否有运动状态存在。首先对背景图像  $Bg(x, y)$  及目标运动状态矩阵  $S(x, y)$  及帧时序控制变量  $i$  进行初始化:背景初始图像  $Bg_0(x, y) = 0$ ;初始运动状态矩阵  $S_0(x, y) = 0$ ;帧时序控制变量  $i = 1$ 。

2) 根据视频序列帧时序变化,跟踪背景和目的一系列变化,并实时更新运动状态矩阵。方法为:按照帧时序逐一读取序列图像,在像素级层次,计算帧序列中当前图像与背景图像之间的距离,若有变化,判断有目标运动,更新对应的状态

矩阵,并记录变化次数,若连续  $N$  帧图像均有变化,将最后一帧图像中该像素点的特征值更新到背景图像中;若无变化,背景图像不更新。该方法可以实时地检测到背景的突变,包括原来运动的目标停止变化为背景,或原来作为背景的静止目标突然运动成为了前景目标等情况。

算法过程如下:开始读取第  $i$  帧图像  $T_i$ ,按照式(3)计算图像  $T_i$  与前一背景图像之间的距离  $D_{T_i, Bg_{i-1}}$ ,进行目标运动状态检测,根据  $D_{T_i, Bg_{i-1}}$  的结果确定运动状态矩阵  $S_i(x, y)$ ,并依照  $S_i(x, y)$  确定当前帧背景,构造状态矩阵  $Bc_i(x, y)$  和背景图像更新状态矩阵  $Bu_i(x, y)$ :

$$Bc_i(x, y) = \begin{cases} 1, S_i(x, y) < 0 \\ 0, S_i(x, y) \geq 0 \end{cases}$$
$$Bu_i(x, y) = \begin{cases} 1, S_i(x, y) > N, D_{T_i, Bg_{i-1}} < T \\ 0, \text{其他} \end{cases} \quad (5)$$

其中,变量  $N$  代表目标状态变化的次数,本文  $N$  取 7;  $T$  为图像距离阈值,本文取 0.06;  $i=1, 2, \dots, m$ 。接着,根据  $Bc_i(x, y)$  和  $Bu_i(x, y)$  的状态,为表示背景突变,依次获取第  $i$  帧背景图像  $Bg_i$ :

$$Bg_i(x, y) = T_i(x, y) \cdot Bc_i(x, y) + Bg_{i-1}(x, y) \cdot [1 - Bc_i(x, y)], i = 1, 2, \dots, m \quad (6)$$

最后,更新运动状态矩阵为:

$$S_i(x, y) = S_{i-1} \cdot [1 - 2Bc_i(x, y)], i = 1, 2, \dots, m \quad (7)$$

3) 根据以上背景图像更新状态矩阵  $Bu_i(x, y)$  和当前帧背景图像  $Bg_i(x, y)$ ,将当前帧背景图像(仍表示为  $Bg_i(x, y)$ )更新为:

$$Bg_i(x, y) = T_i(x, y) \cdot Bu_i(x, y) + Bg_i(x, y) \cdot [1 - Bu_i(x, y)], i = 1, 2, \dots, m \quad (8)$$

4) 前帧序列图像处理完毕后,背景图像已按照运动状态矩阵的变化进行了实时更新,接着读取下一帧图像,调整帧时序控制变量  $i=i+1$ ,若  $i \leq m$ ,转向第 2)步;否则,整个序列图像处理结束,保留图像  $Bg_m(x, y)$  为最终构建的背景图像  $Bg(x, y)$ 。

5) 根据背景图像,采用背景差方法,提取任意帧图像中的运动目标,并进行二值化处理:

$$\text{Object}_r(x, y) = \begin{cases} 1, D_{T_r, Bg} \geq T \\ 0, D_{T_r, Bg} < T \end{cases} \quad (9)$$

其中,  $r$  是视频图像序列中的任一帧序号。

2 实验结果与分析

选取山东烟台南大街一段交通监控视频中连

续 40 帧大小为 288 像素×768 像素的 RGB 彩色图像进行性能测试比较。

1) 首先选取 40 帧图像,分别采用均值背景方法、PCA 提取主成分方法、文献[8]背景更新方法与本文方法分别构建背景图像,其中,PCA 方法所用耗时最高,远远高于其他三种方法。所构建的背景结果见图 2。

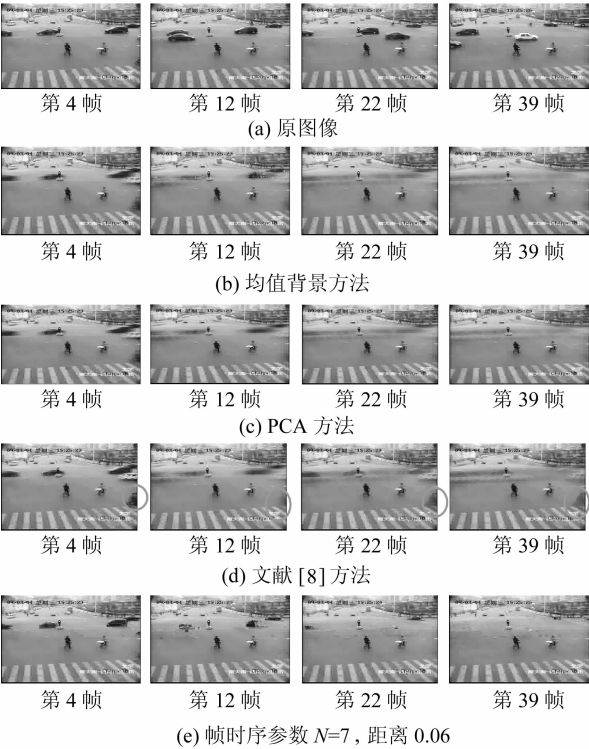


图 2 不同方法构建的背景图像比较

Fig. 2 Background Images Comparison of Difference Methods

从图 2 可以看出,均值背景法构建的图像能够很好地保留图像序列的总体特性,但对于图像中运动速度慢的目标对背景图像有较大影响,在背景中能够看出路面上目标运动的痕迹;PCA 构造背景图像能够体现出序列图像的主要背景,与均值背景类似,仍能看出目标运动的痕迹;文献[8]方法采用  $m_{t+1} = \alpha x_t + (1 - \alpha)m_t$  和  $\sigma_{t+1}^2 = \alpha(x_t - m_t)^2 + (1 - \alpha)\sigma_t^2$  进行均值和方差更新,其中,

$$\alpha = \begin{cases} 0, & x \notin \text{background} \\ 0.05, & x \in \text{background} \end{cases}$$

但是若背景中目标被移动或者长时间固定在某一位置,结果背景会在新、旧两个地方检测到目标(如图 2 所示);而帧时序背景图像方法很好地保留了背景特征,因为它用帧序列中当前像素点的值进行背景更新,对目标运动速度快慢均能很好地进行描述,并可以实时地反映当前帧图像变化的情况。

2) 对背景突变更新性能进行比较。如图 2 中右上角有一运动速度相对较慢的汽车,从 12 帧、22 帧中可以看出,其运动距离相对于其他目标而言是非常近的,在 10 帧距离中,该区域中部分像素应被看成是停止运动的背景,从 PCA 和均值方法构造的背景中很难看出这一变化特征,而在本文方法中,背景图像中可以保存一段时间该区域特征的原貌,直至 39 帧确认其为真实目标后,将其从背景图像中更新掉,即本文方法对复杂背景的变化适应性能较强。

3) 对不同构建方法获得的背景结果进行图像质量评价,测试在视频序列中构建背景图像的变化情况。目前,对于彩色图像的客观评价标准主要依赖主观视觉感受,对于相应的灰度图像性能,采用基于统计特性的客观评价指标(即均值、标准偏差以及基于信息量的客观评价指标)进行性能比较。图 3 中,均值代表图像的平均亮度度量,由图 3(a)可见,PCA 方法均值最高,而均值背景高于文献[8];帧时序方法处于居中偏上,图像质量较 PCA、均值以及文献[8]方法要好。由图 3(b)可见,帧时序方法标准偏差要高于均值和 PCA 方法,与文献[8]方法接近,30 帧及以后构建的背景图像质量优于文献[8]。由图 3(c)可见,四种方法在背景构建开始部分相差不大,但本文方法从第 20 帧开始,熵值要远远大于其他三种方法。由图 3(d)可以看出,PCA、均值以及文献[8]方法很接近,但基于帧时序的背景构建方法对突变情况的适应性能要高于其他三种背景构建方法。

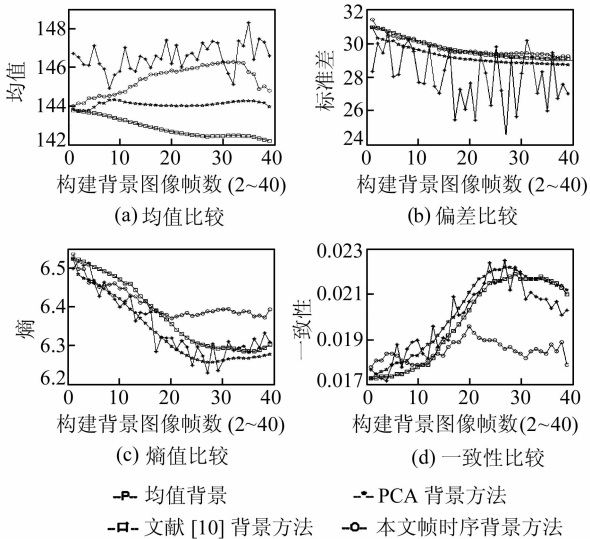


图 3 不同方法构建背景图像性能比较

Fig. 3 Background Images Performance Comparison of Different Method

3 结 语

根据视频图像目标的运动状态,采用帧时序运动状态矩阵实时描述当前图像及背景的变化情况,通过定义彩色图像距离,判断序列帧图像与背景之间的变化,实时更新背景图像,该背景构建方法无需假设背景模型,不需要借助大量的序列帧图像,主要从目标运动状态出发,考虑复杂背景的实时变化,对各类突变适应能力强,实时性好。其缺点是忽略了序列图像的总体特性。该方法适用于各类场景复杂视频动态目标检测的背景图像构建。

参 考 文 献

[1] Stauffer C, Grimson E. Learning Patterns of Activity Using Real-time tracking[J]. IEEE Transactions on Pattern Recognition and Machine Intelligence, 2000, 22(8):747-757

[2] Stauffer C, Grimson W E L. Adaptive Background Mixture Models for Real-time Tracking [J]. Computer Vision and Pattern Recognition, 1999 (2): 246-252

[3] Elgammal A, Harwood D, Davis L. Non-parametric Model for Background Subtraction [C]. The IEEE International Conference on Computer Vision

Frame-Rate Workshop, Kerkyra, Greece,1999

[4] Liu Z, Wang X Y. Segmenting Foreground from Similarly Colored Background[J]. Optical Engineering. 2008,47(7): 1-11

[5] Cucchiara R, Grana C, Piccardi M. Detecting Moving Objects, Ghosts and Shadows in Video Streams [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003,25(10):1 337-1 342

[6] Cucchiara R, Grana C, Piccardi M, et al. Detecting Objects, Shadows and Ghosts in Video Streams by Exploiting Color and Motion Information[C]. IEEE International Conference on Image Analysis and Processing, Palermo, Italy, 2001

[7] Tian Y L, Lu M, Hampapur A. Robust and Efficient Foreground Analysis for Real-time Video Surveillance[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 2005

[8] Hansung K, Ryuuki S, Itaru K. Robust Foreground Extraction Technique Using Background Subtraction with Multiple Thresholds [J]. Optical Engineering, 2007,46(9):1-12

第一作者简介:王枚,副教授,博士,主要研究方向为图像处理及模式识别。  
E-mail:wangmei336@163.com

Background Building Method of Moving Objects  
Detection in Complex Scenes

WANG Mei<sup>1,3</sup> SU Guangda<sup>2</sup> WANG Guohong<sup>3</sup>

(1 Laboratory of Image Processing, Yantai Vocational College, 2018 Middle Binhai Road, Yantai 264670,China)

(2 Department of Electronic Engineering, Tsinghua University, Tsinghuayuan, Beijing 100084,China)

(3 Department of Electronic Information, Naval Aeronautical Engineering Institute, 188 Erma Road, Yantai 264001,China)

**Abstract:** We present a background image building method to deal with the problem of color images motion object detection in real-time. The distance between color images is defined to compare the frame image and the background image according to the frame sequences. The motion state matrix is extracted to update background image in pixels level. The method need not background models and fit variety changes. Actual color images were used to test the method. The experimental results show that our method is effective and accurate.

**Key words:** motion object detection;background subtraction;motion state matrix;update

About the first author: WANG Mei, associate professor, Ph.D, majors in image processing and pattern recognition.  
E-mail: wangmei336@163.com