

# 混合高斯模型和帧间差分相融合的自适应背景模型

刘鑫<sup>1)</sup> 刘辉<sup>1)</sup> 强振平<sup>2)</sup> 耿续涛<sup>1)</sup>

<sup>1)</sup>(昆明理工大学信息工程与自动化学院,昆明 650051) <sup>2)</sup>(西南林学院计算机与信息科学系,昆明 650224)

**摘要** 提出了运动目标检测中背景动态建模的一种方法。该方法是在 Stauffer 等人提出的自适应混合高斯背景模型基础上,为每个像素构建混合高斯背景模型,通过融入帧间差分把每帧中的图像区分为背景区域、背景显露区域和运动物体区域。相对于背景区域,背景显露区中的像素点将以大的更新率更新背景模型,使得长时间停滞物体由背景变成运动前景时,被遮挡的背景显露区被快速恢复。与 Stauffer 等人提出的方法不同的是,物体运动区不再构建新的高斯分布加入到混合高斯分布模型中,减弱了慢速运动物体对背景的影响。实验结果表明,在有诸多不确定性因素的序列视频中构建的背景有较好的自适应性,能迅速响应实际场景的变化。

**关键词** 背景建模 混合高斯模型 运动目标检测 帧间差分

中图法分类号:TP391.4 文献标识码:A 文章编号:1006-8961(2008)04-0729-06

## Adaptive Background Modeling Based on Mixture Gaussian Model and Frame Subtraction

LIU Xin<sup>1)</sup>, LIU Hui<sup>1)</sup>, QIANG Zhen-ping<sup>2)</sup>, GENG Xu-tao<sup>1)</sup>

<sup>1)</sup>(Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650051)

<sup>2)</sup>(Computer and Information Department, Southwest Forestry University, Kunming 650224)

**Abstract** In this paper a dynamic background modeling approach for moving objects detection is proposed. This model is based on mixture Gaussian model suggested by Stauffer et al. It constructs a mixture Gaussians Model for each pixel. In sequence frames subtracting the model classify the pixels in each frame into background area, uncovered background area and moving objection area. In order to quick restore the background covered by stagnated objects when they move again, the model set the update rate in uncovered background area larger than which in background area. Compare to the Stauffer's model, our model moving objection area no longer creates new Gaussian distribution, so it can avoid classifying slow-moving objects to the background. The experimental result indicate that our model has preferable adaptive performance to the scene with many uncertain factors, and correspondence quickly.

**Keywords** background model, mixture Gaussian model, moving object detection, frames subtraction

## 1 引言

视觉监控系统中的现实场景往往是错综复杂、变化无常的,对监控场景进行背景建模是目标检测、跟踪、分类和行为理解等后续处理的基础,如何准确地对监控场景进行建模具有积极的研究意义。

随着全球对智能视觉监控的需求日益增加,许多关于智能视觉监控的大规模研究项目已经在欧美国家展开,已有多篇论著详细介绍了针对不同应用条件的背景模型。Ridder 等人对每个像素构建一个模型,每个像素有一个 Kalman 滤波,使得系统能适应场景中的灯光改变<sup>[1]</sup>。这种方法有一个像素自动阈值,但是背景依然自适应变化得很缓慢,不能很好地处理多层背景的情况。Wren

等人<sup>[2]</sup>的 Pfinder 实时系统对跟踪对象用了一个多类的统计模型,但是背景模型每个像素只有一个高斯模型,在初始化一个没有运动物体的背景后,这个系统能得到一个很好的结果,但是在户外场景,随着时间的迁移、光线的变化,这种算法并没有得到成功的跟踪结果。

针对静止摄像机的情况,本文的多维混合高斯和帧间差分相融合的自适应背景模型是首先选取一帧作为背景帧,在 Stauffer 等人提出的背景模型<sup>[3,4]</sup>基础上为每个像素构建多维混合高斯分布模型,运用帧间差分法对相邻两帧图像进行差分处理,通过自适应阈值把每帧中像素点区分为背景区域和变化区域;再将变化区域中的点与其多维混合高斯分布进行匹配,进而区分出背景显露区域和运动物体区域;相对于背景区域,背景显露区中的像素点将以大的更新率更新多维混合高斯分布模型;而运动物体区域中的点不再构建新的高斯分布加入到多维混合高斯分布模型中;在学习新的场景信息后更新背景。该方法有效弥补了 Stauffer 等人提出的自适应混合高斯背景模型对于长时间停滞物体由背景变成前景运动物体时,背景模型更新速度较慢的问题;同时算法也减弱了慢速运动物体对背景模型的影响。实验结果表明,即使光线变化、车辆人流混乱的场合中,背景生成也是稳定的,具有较好的自适应性,能做到实时的户外运动目标检测与跟踪。

## 2 背景模型

### 2.1 自适应多维混合高斯背景模型

Stauffer 等人提出的背景模型<sup>[4]</sup>对每个像素点建立  $K$  个多维高斯分布来混合模拟该点的背景值,对于像素点  $(x_0, y_0)$ ,  $t$  时刻的观察值  $\mathbf{x}_t = (r_t, g_t, b_t)^T$ ,属于背景的概率

$$p(\mathbf{x}_t) = \sum_{i=1}^K \omega_{k,t} \eta(\mathbf{x}_t, \boldsymbol{\mu}_{t,i}, \boldsymbol{\Sigma}_{t,i}),$$

$$\eta(\mathbf{x}_t, \boldsymbol{\mu}_{t,i}, \boldsymbol{\Sigma}_{t,i}) = \frac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}_{t,i}|^{1/2}} \times e^{-\frac{1}{2}(\mathbf{x}_t - \boldsymbol{\mu}_{t,i})^T \boldsymbol{\Sigma}_{t,i}^{-1} (\mathbf{x}_t - \boldsymbol{\mu}_{t,i})} \quad (1)$$

为第  $i$  个高斯分布的概率密度数,  $\omega_{i,t}$  ( $\sum_{i=1}^K \omega_{i,t} = 1$ ) 为分布对应的权值,  $K$  个高斯分布总是按照优先级  $\rho_{i,t} = \frac{\omega_{i,t}}{\sigma_i}$  从高到低的次序排列<sup>[4]</sup>。  $K$  的取值为

5~7,  $K$  可以根据计算机的存储容量和运算能力进行调整。假设 rgb 3 通道是相互独立的(以牺牲一些精确度为代价,避免大量的运算消耗在矩阵倒置中),并且有相同的方差。于是均值  $\boldsymbol{\mu}_{i,t} = (\boldsymbol{\mu}_r, \boldsymbol{\mu}_g, \boldsymbol{\mu}_b)_{i,t}^T$ ; 协方差矩阵  $\boldsymbol{\Sigma}_{i,t} = \sigma_{i,t}^2 \mathbf{I}$  ( $\mathbf{I}$  为单位矩阵)。

背景建模过程如下:

#### (1) 高斯分布模型匹配

新观察值  $\mathbf{x}_t$  与它的按优先级  $\rho_{i,t}$  高低排列的  $k$  ( $1 \leq k \leq K$ ) 个高斯分布进行匹配,匹配的条件为  $|\mathbf{x}_t - \boldsymbol{\mu}_{i,t}| < 2.5\sigma_i$ 。若都不匹配,当  $k < K$  时,增加一个新的高斯分布,当  $k = K$  时用新的高斯分布代替最小优先级  $\rho_{k,t}$  的高斯分布。新的高斯分布以  $\mathbf{x}_t$  作为均值,并初始一个大的方差<sup>[4]</sup>,这里取 50。

#### (2) 混合高斯模型的更新

高斯分布权值更新:  $\omega_{i,t} = (1 - \alpha)\omega_{i,t-1} + \alpha(M_{i,t})$ , 匹配的分  $M_{i,t} = 1$ , 其余分布  $M_{i,t} = 0$ 。对匹配的高斯分布的均值和方差进行更新,不匹配的分保持不变的分布保持不变。

$$\boldsymbol{\mu}_{j,t} = (1 - \beta)\boldsymbol{\mu}_{j,t-1} + \beta\mathbf{x}_t$$

$$\sigma_{j,t}^2 = (1 - \beta)\sigma_{j,t-1}^2 + \beta(\mathbf{x}_t - \boldsymbol{\mu}_{j,t})^T (\mathbf{x}_t - \boldsymbol{\mu}_{j,t})$$

$\alpha\beta$  为学习率

$$\beta = \alpha\eta(\mathbf{x}_t | \boldsymbol{\mu}_{t,k}, \sigma_{t,k}) \quad (2)$$

#### (3) 生成背景模型

重新对混合高斯分布按优先级  $\rho_{i,t}$  由大到小排列<sup>[4]</sup>,取前  $B$  个高斯分布联机的生成背景:

$$\mathbf{x}_{bg} = \sum_{i=1}^B \omega_{k,t} \boldsymbol{\mu}_{t,i}$$

$$B = \underset{b}{\operatorname{argmin}} \left( \sum_{k=1}^b \omega_{k,t} > T \right) \quad (3)$$

取  $T = 0.85$ , 如果  $T$  选择一个小的值,则背景通常用一个高斯分布来描述;如果  $T$  选择一个比较大的值,背景由多个分布来混合描述。

由  $K$  均值算法,  $\omega_{i,t}$  等价于每个高斯分布作为背景的先验概率,每个分布的概率密度函数为  $\eta(\boldsymbol{\mu}_{t,i}, \boldsymbol{\mu}_{t,i}, \boldsymbol{\Sigma}_{t,i})$  (均值  $\boldsymbol{\mu}_{t,i}$  的概率)<sup>[5,6]</sup>。由贝叶斯判定准则得:

$$R = \frac{p(BG | \mathbf{x}_{bg})}{p(FG | \mathbf{x}_{bg})} = \frac{p(\mathbf{x}_{bg} | BG)}{p(\mathbf{x}_{bg} | FG)}$$

$$= \frac{\sum_{i=1}^B \omega_{i,t} \eta(\boldsymbol{\mu}_{t,i}, \boldsymbol{\mu}_{t,i}, \boldsymbol{\Sigma}_{t,i})}{\sum_{j=B}^K \omega_{j,t} \eta(\boldsymbol{\mu}_{t,j}, \boldsymbol{\mu}_{t,j}, \boldsymbol{\Sigma}_{t,j})}$$

$$= \frac{\sum_{i=1}^B \frac{\omega_{i,t}}{(\sigma_i)^3}}{\sum_{j=1}^K \frac{\omega_{j,t}}{(\sigma_j)^3}} \quad (4)$$

分布是按照  $\rho_{i,t} = \frac{\omega_{i,t}}{\sigma_i}$  由大到小排列的,所以

$R > 1$ , 即后验概率  $p(BG | \mathbf{x}_{bg}) > p(FG | \mathbf{x}_{bg})$ , 背景生成是有效的。

### 3 融入帧间差分的背景更新

背景模型的更新是背景模型设计中很重要的环节。需要注意的是背景的变化可能是由以下原因引起的:①光照变化等因素引起的背景本身颜色的变化;②物体由运动转为静止,被“融入”到背景中;③物体由静止转为运动时,“逃离”背景,背景模型对运动目标要有较强的抗干扰能力,因为在背景模型的更新过程中,对背景模型上的每点而言都是受到了一个颜色序列的“训练”;④对静止的背景的“训练”是我们所希望的;⑤运动目标的“训练”是不希

望看到的。

因为 Stauffer 等人提出多维混合高斯分布背景模型<sup>[4]</sup>本身有多个高斯分布,判定前景、背景并不单单依赖于某个高斯分布,更依赖于各个分布的权值和优先级。对①②,它可以通过引入新的高斯分布解决;对④,由于只有相匹配的高斯分布的高斯参数才得到更新,所以受运动物体的干扰也不是那么严重。但它的不足依然存在,因为它并没有对③中静止目标的情况作特别处理,被物体覆盖的背景显露出来,如果背景模型不能迅速跟上实际背景的变化,检测结果中就会出现大范围的噪声或不合理的长时间静止前景物体,如静止的汽车启动后,长时间留下“影子”。也没有考虑到⑤中的不同“训练”之间的区别,特别是当体积较大物体运动较慢时,这种长时间的“训练”可能会引起错误的检测结果(如图1(b)生成的背景中,大型慢速车辆留下的“划痕”)。为了解决③和⑤,我们通过融入帧间差分对多维混合高斯分布背景模型进行了相关改进。

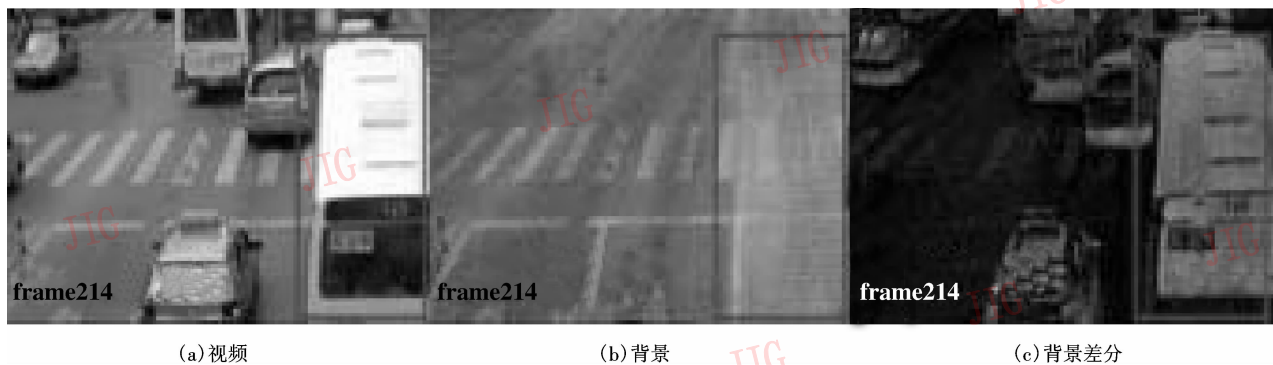


图1 大型慢速车辆留下的“划痕”

Fig. 1 Slow and lager moving object left “scratch”

#### 3.1 融入帧间差分的背景学习率

在前文描述的多维混合高斯分布背景模型中,学习率  $\beta$  的大小由学习率  $\alpha$  决定,所以背景的更新速度主要由  $\alpha$  决定。如果  $\alpha$  取得较小,适应环境变化的能力就低,需要足够长的时间才能适应环境的变化;如果  $\alpha$  取得较大,适应环境变化能力强,但容易引入噪声,前景运动物体对背景模型的影响也较大,因此针对不同的区域采取不同的更新策略<sup>[4,7]</sup>。

#### 3.2 变化区域的检测

帧间差分能够检测出相邻两帧间发生了变化的区域( $a_c$ )。这个区域实际上包括运动物体在前一帧所覆盖的背景区域即背景显露区( $a_{bg}$ ),运动物

体现在所覆盖的区域,即当前帧中就是运动物体本身( $a_m$ )。设  $f_{t-1}$  和  $f_t$  为视频序列中连续的两帧图像,将这两帧图像进行差分处理,对于视频图像中点  $x_t(i, j)$ , 判定规则如下:

$$\begin{cases} \text{if } |\vec{f}_t(i, j) - \vec{f}_{t-1}(i, j)| < T & x_t(i, j) \in a_{bg} \\ \text{if } |\vec{f}_t(i, j) - \vec{f}_{t-1}(i, j)| \geq T & x_t(i, j) \in a_c \\ T = T_c + T_1 & T_1 = \frac{1}{N} \sum |\vec{f}_t(i, j) - \vec{f}_{t-1}(i, j)| \end{cases} \quad (5)$$

式中,  $T_c$  为颜色变化阈值,取  $T_c = 30$ ,  $T_c$  的选取直接影响运动目标的提取与分割(运动目标轮廓的描

绘)。 $T_c$  取值小, 会引入大量的噪声点,  $T_c$  取值大, 小运动物体不能被检出。自适应添加项  $T_1$  表示整体光线的变化, 如果图像光照变化较小, 添加项的值趋近于零<sup>[8]</sup>。如果图像光照变化明显, 则添加项有明显增大, 这样能使阈值  $T$  适应环境光照的突变情况。 $a_{bg}$  为差分后得到的背景区域,  $a_c$  为差分后得到的变化区域(如图 2(c),  $a_c$  为 522 帧与 521 帧差分的结果)。

### 3.3 背景显露区域和运动区域的区分

将分割出来的运动变化区域  $a_c$  中的像素点  $x_i(i, j)$  与各自的前  $M$  个按优先级  $\rho_{i,t}$  由大到小进行排序的高斯分布去匹配。匹配方法为

$$\text{if } |\overline{f_i(i, j)} - \mu_{t-1, n}| < 2.5\sigma_i \quad 0 < n \leq M$$

$$x_i(i, j) \in a_{ubg}$$

其中,  $\mu_{t-1, n}$  为该像素点  $(t-1)$  时刻的第  $n$  个高斯分布的均值。

如果跟前  $M$  个分布都不匹配, 则  $x_i(i, j) \in a_m$ 。 $M$  的取值方式为

$$\text{if } B = 1, M = k \quad \text{else } M = B$$

其中,  $B$  为该像素点  $(t-1)$  时刻生成背景所取的高斯分布个数;  $k$  为该像素点现有的分布数。(如图 2(d), 匹配的像素点用白色表示, 得到背景显露区  $a_{ubg}$ ; 图 2(e) 区分出来的运动区域  $a_m$  和背景显露区  $a_{ubg}$ )。

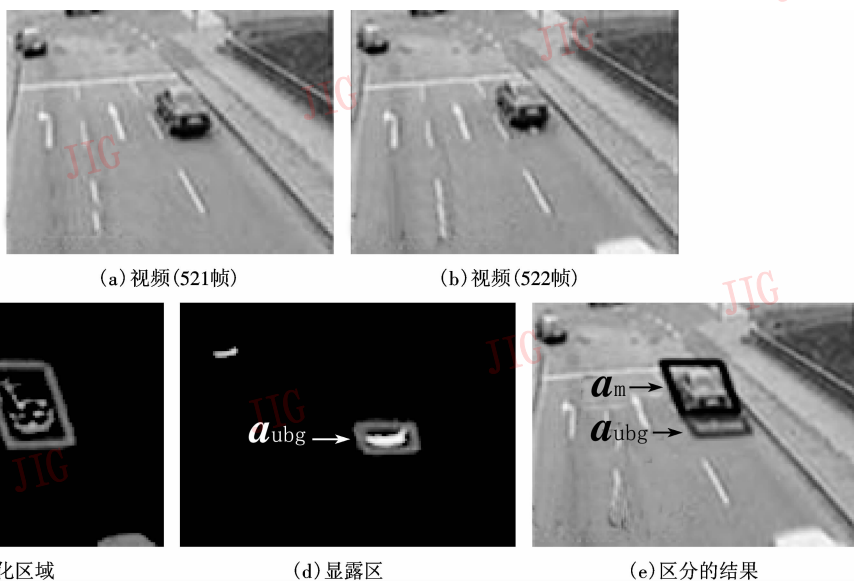


图 2 区分出不同区域

Fig. 2 Distinguish different areas

### 3.4 不同区域 $a$ 的取值

赋予背景区域( $a_{bg}$ )中像素点对应的多维混合高斯分布模型小的更新率  $\alpha_{bg}$ , 保持背景区域的稳定性, 避免引入噪声。赋予背景显露区域( $a_{ubg}$ )中像素点对应的多维混合高斯分布模型大的更新率  $\alpha_{ubg}$ , 这样匹配的高斯分布(即显露的背景对应的分布)将获得较大权值  $\omega$ , 其对应的方差  $\sigma^2$  将变小, 其优先级  $\rho$  迅速变大, 而其余的高斯分布随着权值  $\omega$  的减小, 优先级  $\rho$  将减小。综上匹配的分影响背景的程度将加大, 显露的背景迅速地恢复, 有效地解决了③。

运动区域( $a_m$ )中的像素点因为没有跟任何一个多维高斯分布匹配, 不再构建新的高斯分布加入到多维混合高斯分布模型中。这样将有利于减少运动目标对背景的影响, 有效地解决了⑤。

## 4 实验结果

依照本文算法分别对  $688 \times 565$ 、 $768 \times 576$  的视频图像帧进行了背景建模。取背景学习率  $\alpha_{bg} = 0.002$ ,  $\alpha_{ubg} = 0.1$ , 在普通 PC 机(Pentium4 3GHz CPU)上, 本文的背景建模算法达到了 15 fps 的处理速度。

图 3 是对背景模型的更新策略改进前后的实验对比结果。改进前背景中车 a 留下了“影子”a' 缓慢的“消失”。背景差分时 a 离开后的影子 a' 始终都被检测出来。改进后的背景中车 a 留下的影子 a' 迅速“消失”。图 4 是在光照变化, 人、车流多的交通视频序列上建立的背景模型, 实验表明在复杂多变的真实场景中建立的背景模型有较好的自适应性和

鲁棒性,能迅速响应实际场景的变化。

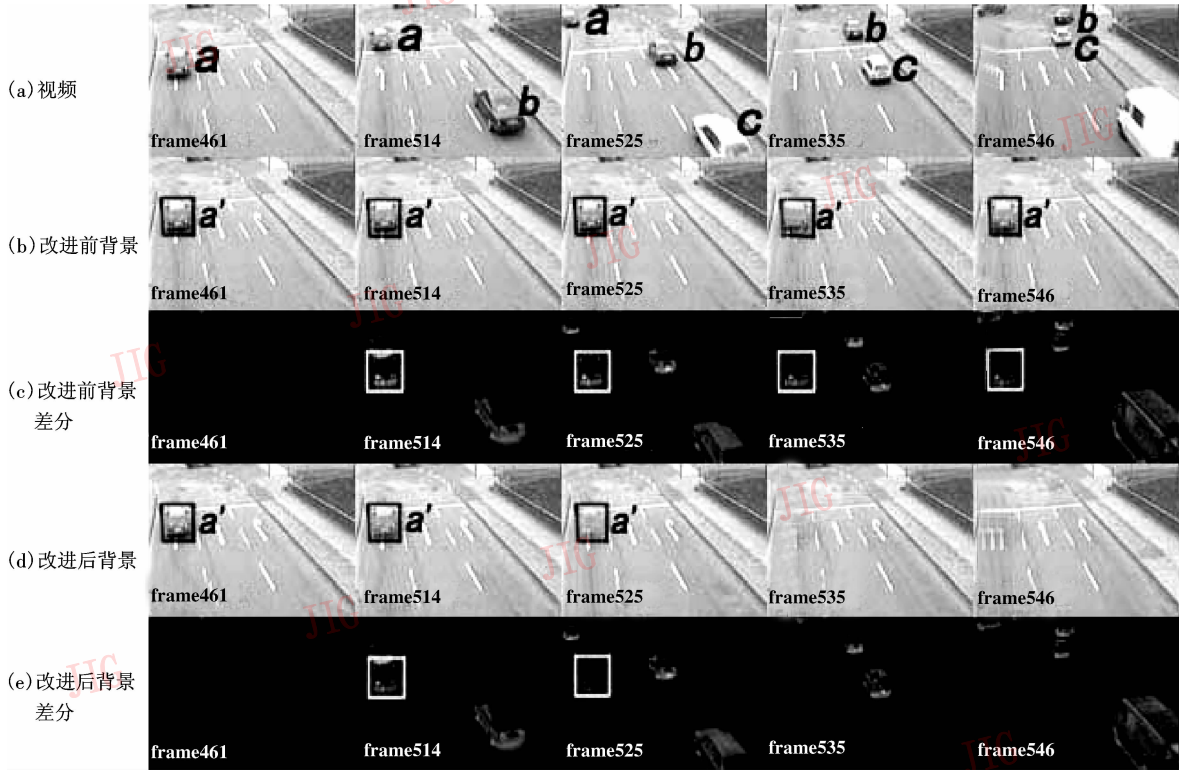


图 3 算法改进前后效果比较

Fig. 3 The comparison of Stauffer's algorithm and improved algorithm

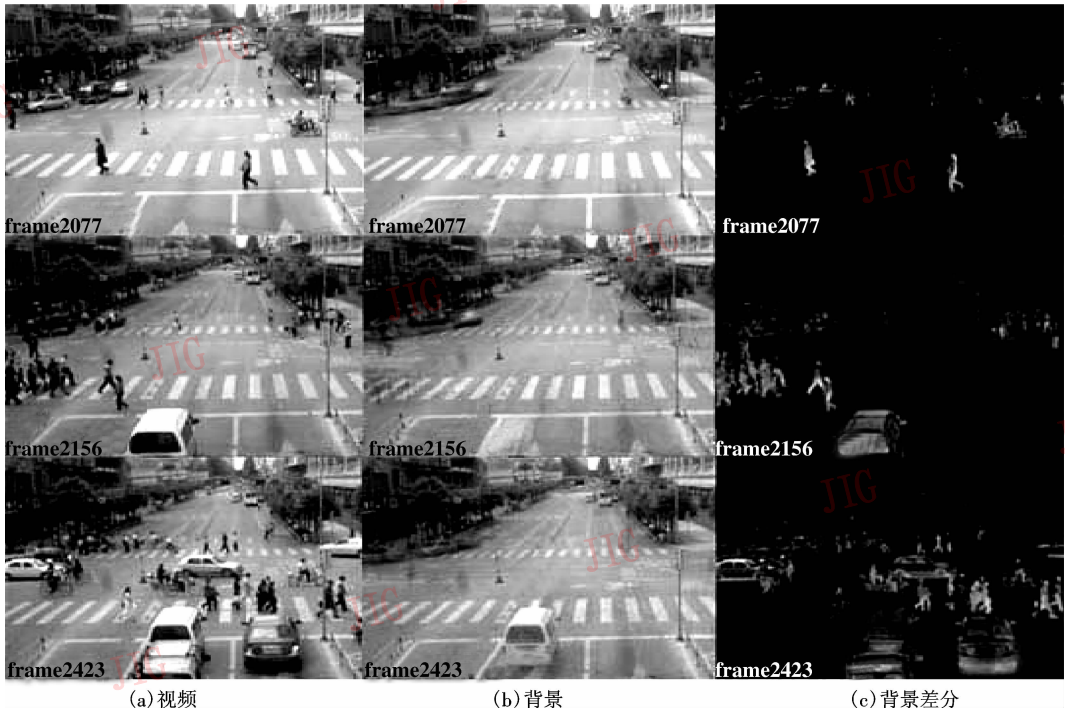


图 4 复杂的场景

Fig. 4 Complex scenes

## 5 结 论

提出了一种静止摄像机条件下对视频序列进行背景建模的方法。方法中采用多维混合高斯和帧间差分相融合的自适应背景模型,对背景区域、背景显露区域和运动目标区域三者采取了不同的背景更新策略,方法有效地弥补了 Stauffer 等人提出的自适应混合高斯背景模型对于长时间停滞物体由背景变成前景运动物体时,背景模型更新速度较慢的问题;同时算法也减弱了慢速运动物体对背景模型的影响。实验结果表明,即使光线变化、车辆人流混乱的场合中,该方法可以迅速响应实际背景的变化,生成的背景能准确的反映场景信息。

### 参考文献 (References)

- Ridder C, Munkelt O, Kirchner H. Adaptive background estimation and foreground detection using Kalman-filtering [A]. In: Proceedings of the Int'l Conference on Recent Advances Sinmechatronics [C], Istanbul, Turkey, 1995:193 ~ 199.
- Wren C, Azarbayejani A, Darrell T, *et al.* Pfunder: real-time tracking of the human body [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, **19**(7):780 ~ 785.
- Friedman N, Russell S. Image segmentation in video sequences: A probabilistic approach [A]. In: Proceedings of the 13th Annual Conference on Uncertainty in Artificial Intelligence [C], Providence, Rhode Island, 1997:175 ~ 181.
- Stauffer C, Grimson W. Adaptive background mixture models for real-time tracking [A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C], Fort Collins, Colorado, USA, 1999:246 ~ 252.
- Zivkovic Z. Improved adaptive Gaussian mixture model for background subtraction [A]. In: Proceedings of the 17th International Conference on Pattern Recognition [C], Cambridge, United Kingdom, 2004, **2**:28 ~ 31.
- Hayman E, Eklundh J. Statistical background subtraction for a mobile observer [A]. In: Proceedings of the 9th International Conference on Computer Vision [C], Nice, France, 2003: 67 ~ 74.
- Liu Ya, Ai Hai-zhou, Xu Guang-you. Moving object detection and tracking based on background subtraction [J]. Information and Control, 2002, **31**(4): 315 ~ 319. [刘亚, 艾海舟, 徐光祐. 一种基于背景模型的运动目标检测与跟踪算法 [J]. 信息与控制, 2002, **31**(4): 315 ~ 319.]
- Zhu Ming-han, Luo Da-yong, Cao Qian-xia. Moving objects detection algorithm based on two consecutive frames subtraction and background subtraction [J]. Computer Measurement & Control, 2005, **13**(3): 215 ~ 217. [朱明早, 罗大庸, 曹倩霞. 帧间差分与背景差分相融合的运动目标检测算法. 计算机测量与控制 [J], 2005, **13**(3): 215 ~ 217.]