

# 长期视频监控系统的多分布模型背景差方法\*

岑 峰 戚飞虎 陈茂林

(上海交通大学计算机科学与工程系, 上海, 200030)

**摘要** 提出在长期视频监控系统中采用背景差进行运动目标提取时算法所要满足的基本要求, 并提出了一种能够满足这些要求的背景差方法. 该方法用色度、亮度空间的多个分布模型来建立背景模型, 描述彩色视频图像的背景像素点及其统计特性, 在对背景模型更新时将均值、方差的更新速率和多个模型的更替速率分开. 对像素值属于多个分布模型的情况, 用最小相似距离确定要更新的模型. 该方法利用提取的前景像素点信息反馈以检测光强的突变, 利用亮度信息消除运动目标的阴影. 实验证明该方法很好的满足了长期视频监控系统的要求.

**关键词** 背景差, 背景模型, 视频监控, 目标跟踪.

## MULTI-DISTRIBUTION MODEL FOR BACKGROUND SUBTRACTION IN LONG-TERM VIDEO SURVEILLANCE SYSTEM \*

CEN Feng QI Fei-Hu CHEN Mao-Lin

(Department of Computer Science & Engineering, Shanghai Jiaotong University, Shanghai 200030, China)

**Abstract** The basic performance requirement of exploiting background subtraction approach in long-term color video surveillance system was described. To meet the requirement, an algorithm was also proposed. In this approach, by using the multi-distribution models in lightness and chromaticity spaces, the background model was built. The multi-distribution models were then updated using independent mean and covariance updating rate and model replacing rate. When a pixel could be represented by more than one distribution model, the model which has the minimum similarity distance was updated. The approach also involves the detection of the sudden changes in illumination by the feedback of the information of foreground pixels and the suppression of shadow influenced by lightness information. Experiments show that the method satisfies the demands of long-term video surveillance.

**Key words** background subtraction, background model, video surveillance, object tracking.

### 前言

视频监控系统主要使用固定摄像机对一固定场景进行监控, 因此背景差方法在智能化视频监控系统中是一种重要的运动目标分割方法. 背景差方法的基本思想是通过输入图像与背景模型进行比较的方法分割运动目标, 背景差的难点是对背景模型的维持与更新<sup>[1]</sup>, 实际上, 场景中的背景很复杂, 存在各种各样的干扰, 且背景是随时间不断变化的, 背景的模式及其更新应能处理这些干扰和变化.

背景差通常分为低层次处理和高层次处理两个步骤: 低层次处理检测出与背景模型不同的像素, 高

层次处理对低层次处理检测出的像素进一步处理, 消除噪声, 提取出运动目标.

近几年随着计算机性能的提高, 使智能化视频监控系统的实用成为可能, 人们对背景差方法作了大量的研究, 有基于统计模型的方法, 如高斯模型<sup>[2,3]</sup>、混合高斯模型<sup>[6]</sup>、非参数化模型<sup>[5]</sup>等, 基于预测的方法如卡尔曼滤波器方法<sup>[4]</sup>、Wiener 滤波器方法等<sup>[1]</sup>. 但由于现实场景的复杂性, 没有一种方法是完善的.

本文以长期视频监控为应用背景, 要保证视频监控能够长期运行需要背景差算法具有很强的鲁棒性, 需要满足以下要求: (1) 能适应背景随时间的缓

\* 国家自然科学基金(编号 60072029)资助项目  
稿件收到日期 2001-05-29, 修改稿收到日期 2001-07-27

\* The project supported by the National Natural Science Foundation of China (No. 60072029)  
Received 2001-05-29, revised 2001-07-27

慢变化,如在一天中不同时间的光照变化;(2)能适应背景物体的变化,如场景中移入新的物体,背景中的物体移出场景等变化;(3)背景模型能描述背景中的一些较大扰动,如树叶晃动、显示器屏幕闪烁等;(4)能检测出光照的突然变化,并能在尽量短的时间内适应这种变化;(5)为了能够更好的跟踪识别物体,在背景差算法中应能尽量消除运动目标的阴影的影响。

本文的背景模型采用彩色空间的多个分布模型描述彩色视频图像的背景像素点及其统计特性.对背景模型更新时将均值、方差的更新速率和多个模型更替的速率分开.对像素值属于多个分布模型的情况,用最小相似距离确定要更新的模型.该模型可以很好地描述存在扰动的复杂背景,适应背景的实时变化.本方法还利用提取的前景像素点信息反馈以检测光强的突变,利用亮度信息尽可能地消除运动目标阴影的影响,很好地满足了长期视频监控系统的要求。

## 1 背景模型及更新

### 1.1 背景模型

人的视觉系统通过物体的亮度和色彩来区别不同的物体,在监控系统中由摄像机输入的图像是由  $R, G, B$  3 个分量组成的,混合了色度的亮度信息.区分色度和亮度信息能更好的描述物体.我们首先将输入图像由  $R, G, B$  空间转换到亮度、色度空间  $(s, r, g)$ ,其中亮度坐标  $s$  定义为<sup>[5]</sup>

$$s = (R + G + B) / 3$$

色度信息由  $R, G, B$  的相对确定为

$$r = R / (R + G + B),$$

$$g = G / (R + G + B),$$

$$b = B / (R + G + B).$$

$r, g, b$  称为色度坐标,并有  $r + g + b = 1$ .色度空间是二维的,只有两个色度坐标独立,我们选  $r, g$  作为色度空间坐标。

固定摄像机输入视频图像序列为  $\{1, 2, \dots, t, \dots\}$ ,第  $t$  帧视频图像中像素  $X = (x, y)$  的输入信号为  $I(X, t) = (I_s(X, t), I_r(X, t), I_g(X, t))$ .如果没有运动目标存在,则视频图像相对静止,每一像素点随时间变化都服从一定的统计模型.我们对每一个像素点分别建立统计模型,称为该点的分布模型,用该点的均值

$$\mu(X, t) = (\mu_s(X, t), \mu_r(X, t), \mu_g(X, t)) \quad (1)$$

和均方差

$$D(X, t) = (D_s(X, t), D_r(X, t), D_g(X, t)) \quad (2)$$

描述这个分布模型(为简单起见假设每一像素  $s, r, g$  分量互不相关).树叶晃动、水面晃动、显示器闪烁等背景物体对应的像素点取值变化并不存在单一的分布模式<sup>[6]</sup>,对此我们用多个分布模型  $\{P_1(X, t), \dots, P_i(X, t), \dots, P_N(X, t)\}$  进行描述.其中  $P_i(X, t) = \{\mu_i(X, t), D_i(X, t)\}$ ,下标  $i$  表示第  $i$  个模型,  $N$  表示最大模型个数,通常取值 3~5.

用多个分布模型描述像素点时,像素点取值属于不同模型的概率不同.设属于第  $i$  模型的概率为  $h_i(X, t)$ .如果运动目标存在,运动目标对应的像素点会出现一个新的分布模型,但这个模型是暂时、偶然出现的,其  $h_i(X, t)$  很小.因此我们对各模型按  $h_i(X, t)$  大小排序,取前面的  $L$  各作为背景分布模型,其中  $L$  定义为

$$L = \min_i \left( \sum_{j=1}^i h_j(X, t) > T \right). \quad (3)$$

式(3)中  $T$  为一全局阈值,决定了背景分布模型个数的多少.  $T$  越大则背景分布模型越多,反之则越少,  $L+1 \sim N$  个分布模型称为瞬态分布模型。

### 1.2 背景更新

随着时间推移背景会发生变化,如光照变化、背景物体移入移出等,背景模型必须加以更新以反映这些变化.背景模型的更新包括对每个分布模型的更新和各分布模型之间的更替两部分。

输入视频图像序列并不需要每一帧都用来对背景模型进行更新.因此我们对视频序列进行  $W$  抽样,即每  $W$  帧中抽取一帧对背景模型进行更新,  $W$  为抽样速率.设用第  $t$  帧对背景模型进行更新,下文说明如何对像素点  $X$  的背景模型进行更新。

#### 1.2.1 分布模型更新

首先,将输入信号  $I(X, t)$  与  $N$  个分布模型进行比较,然后对匹配模型进行更新.如果

$$|I_j(X, t) - \mu_j(X, t)| < \tau D_j(X, t), \quad j = s, r, g \quad (4)$$

则  $I(X, t)$  与模型  $P_i$  匹配.其中  $\tau$  为正常数,是一个全局阈值,下标  $i$  表示第  $i$  个分布模型,下标  $j$  表示在  $(s, r, g)$  空间中的分量.如果  $I(X, t)$  同时与多个  $P_i$  匹配,我们选择概率大、方差小、与  $I(X, t)$  相差小的分布模型进行更新,即满足相似距离  $d_i(X, t)$  为最小的分布模型进行更新.  $d_i(X, t)$  定义为

$$d_i(X, t) = \sum_{j=s, r, g} \frac{|I_j(X, t) - \mu_j(X, t) D_j(X, t)|}{h_j(X, t)}. \quad (5)$$

对匹配模型  $P_i$  按下式进行更新:

$$\begin{cases} \mu_{ij}(X, t+1) = (1-\alpha)\mu_{ij}(X, t) + \alpha I(X, t), \\ D_{ij}(X, t+1) = \min((1-\beta)D_{ij}^2(X, t) + \\ \beta(I(X, t) - \mu_{ij}(X, t))^2)^{1/2}, D_{\max}). \end{cases} \quad (6)$$

其中  $\alpha \in (0, 1)$  为均值更新因子, 决定均值的更新速率;  $\beta \in (0, 1)$  为方差更新因子, 决定方差的更新速率;  $D_{\max}$  是所有模型中最大方差的估计值, 是方差的全局上限. 更新因子取值越大则更新速度越快. 在背景分布模型中均值  $\mu$  描述了背景物体反射光的实际值, 方差  $D$  描述了背景物体反射光投影到摄像机平面, 并经过摄像机这一线性系统变换引入的噪声程度. 背景物体是时刻变化的, 时间上越近的视频图像越能反映这种变化, 因此,  $\alpha$  应取较大值. 由最近的视频图像决定  $\mu$  的取值, 以及时反映背景物体的变化. 摄像系统是相对稳定的, 引入的噪声程度也相对稳定, 因此,  $\beta$  取较小值. 由过去较长一段时间的视频图像决定  $D$  的取值.

### 1.2.2 分布模型更替

在长期视频监控中, 背景物体经常会发生变化, 如物体的移入和移出, 背景物体自身发生变化等. 因此多分布模型必须对各分布模型的出现概率进行更新, 使长时间没有出现的背景分布模型逐渐转变为瞬态分布模型最终被新出现的分布模型所代替, 从而反映背景物体的变化. 各分布模型的概率可按如下式更新:

$$h_i(X, t+1) = (1-\gamma)h_i(X, t) + \gamma H_i(t), \quad (7)$$

其中  $P_i$  与  $I(X, t)$  相匹配时  $H_i(t) = 1$ ,  $P_i$  与  $I(X, t)$  不匹配时  $H_i(t) = 0$ ,  $\gamma$  为更新因子, 决定了分布模型的更替速度, 然后对各分布模型按  $h_i(X, t)$  大小排序. 如何将背景物体发生的这种变化与运动目标区分开是一个相当困难的问题. 将背景物体发生的变化与运动目标完全区分开来是高层次视觉处理的任任务, 但在背景差这一低层次运动目标检测算法中, 背景模型要能够及时反映背景变化, 并尽量避免把运动目标融入背景模型.  $\gamma$  取值越大分布模型的更替速度越快, 慢速运动的物体会被融入背景模型, 会增加负的虚假检测; 反之, 分布模型的更替速度越慢, 背景发生的变化将不能及时得到反映, 会增加正的虚假检测. 应针对具体应用场合确定  $\gamma$  的取值.

如果  $I(X, t)$  不与任何  $P_i$  相匹配, 表明出现了一个新的分布模型. 新的分布模型以  $I(X, t)$  为均值,  $D_{\max}$  为方差, 替换  $h_i(X, t)$  最小的分布模型. 然

后对  $h_i(X, t)$  进行归一化处理, 并重新计算出背景分布模型个数  $L$ .

在实际监控系统中常常会遇到输入信号光强突然变化的情况, 如开关灯时, 或具有光圈自动调整功能的摄像机在光圈自动调整时. 在这种情况下光强的变化是全局性的, 如果下文 2.1 节检测到的前景像素目标点占整个视频图像的百分比超过一定比率  $\lambda_1$ , 则认为光强发生了突变. 此时, 为了让背景模型尽快适应这种变化, 可以提高抽样率  $W$ , 同时增大  $\gamma$  加快分布模型的更替速率, 直到背景模型适应这种突变. 如果前景像素目标点整个视频图像的百分比小于一定比率  $\lambda_2$  ( $\lambda_2 < \lambda_1$ ), 则认为背景模型适应了这种突变. 实际应用证明中取  $\lambda_1 = 0.5$  和  $\lambda_2 = 0.3$  即可以取得较好的效果.

## 2 运动目标提取

### 2.1 像素变化检测

对运动目标提取的低层次处理是在像素级上检测出与背景不同的像素, 本文所提出的多分布背景模型就是检测出与所有背景分布模型都不匹配的像素点. 对于像素点  $X$ , 用  $I(X, t)$  分别与该像素点的  $L$  个背景分布模型相比较. 如果  $I(X, t)$  与  $P_i$  满足式(4), 那么  $I(X, t)$  与  $P_i$  相匹配. 如果  $I(X, t)$  与所有的  $L$  个  $P_i$  都不匹配, 则认为点  $X$  为前景目标像素点.

### 2.2 阴影检测

在智能化监控系统中运动目标的阴影会在很大程度上影响后续跟踪和识别的正确率, 因此我们希望在目标检测阶段尽量消除阴影, 以提高后续跟踪、识别的正确率.

阴影在图像中可以看作半透明区域<sup>[7]</sup>. 漫反射物体被阴影遮挡时其表面物理反射特性并没发生任何改变, 只是光照度发生改变. 因此图像中阴影遮挡部分的色度 ( $r, g$ ) 相对于背景没有发生变化, 而亮度  $s$  相对背景增大或减小. 设  $t$  时刻图像中前景目标像素点集合为  $\Sigma$ , 阴影部分像素点集合为  $\Omega$ , 则:

$$\Omega = \{X | X \in \Sigma, m \leq \frac{I_s(X, t)}{\mu_s(X, t)} \leq n, \text{ 且存在与 } X \text{ 在 } (r, g) \text{ 空间满足式(4)的背景分布模型}\}.$$

其中  $I(X, t)$  为像素点  $X$  在  $t$  帧取值,  $m$  为阴影导致的亮度最大衰减比率,  $n$  为阴影造成的物体亮度增大的最大比率,  $\mu_s$  为背景物体没有阴影遮挡时的期望值. 由于每一像素点的值存在多个背景分布模型, 因此用  $(r, g)$  空间与  $X$  匹配的各背景分布

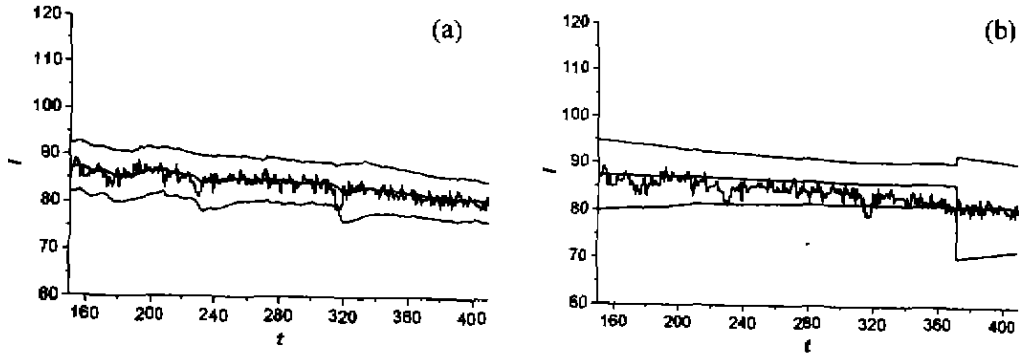


图1 更新算法比较  
 (a)本文算法 (b)混合高斯模型  
 Fig.1 Comparison of algorithms  
 (a)the present method (b)mixed-Gaussian model

模型均值的加权平均值作为  $\mu_s$ , 即

$$\mu_s(X, t) = \left( \sum_{i \in Z} \frac{h_i(X, t)}{H(X, t)} \mu_{is}(X, t) \right) / M. \quad (8)$$

其中  $h_i(X, t)/H(X, t)$  为权重,  $H(X, t) = \sum_{i \in Z} h_i(X, t)$ ,  $Z$  为在  $(r, g)$  空间与  $X$  匹配的背景分布模型的集合,  $M$  为集合  $Z$  中背景分布模型的个数.

2.3 目标提取

通过与各背景模型像比较,在像素级低层次处理上我们得到了前景目标的像素点集合,必须对这些集合作进一步的高层次处理,去除噪声才能完成目标提取.首先我们通过 2.2 节的方法除去阴影部

分的像素点,然后采用传统的形态滤波方法对提取出的像素点进行滤波,抑制噪声,最后删除含很少像素点、不可能是运动目标的孤立区域,进一步消除噪声,完成运动目标提取.

3 实验结果及讨论

文献[6]中的混合高斯模型同本文的模型一样,都是对同一像素点建立多个分布模型.但混合高斯模型是基于高斯分布的,其更新算法中更新因子基于高斯分布,分布模型对背景的变化适应慢,同时也不具备对光强突变和阴影的检测.

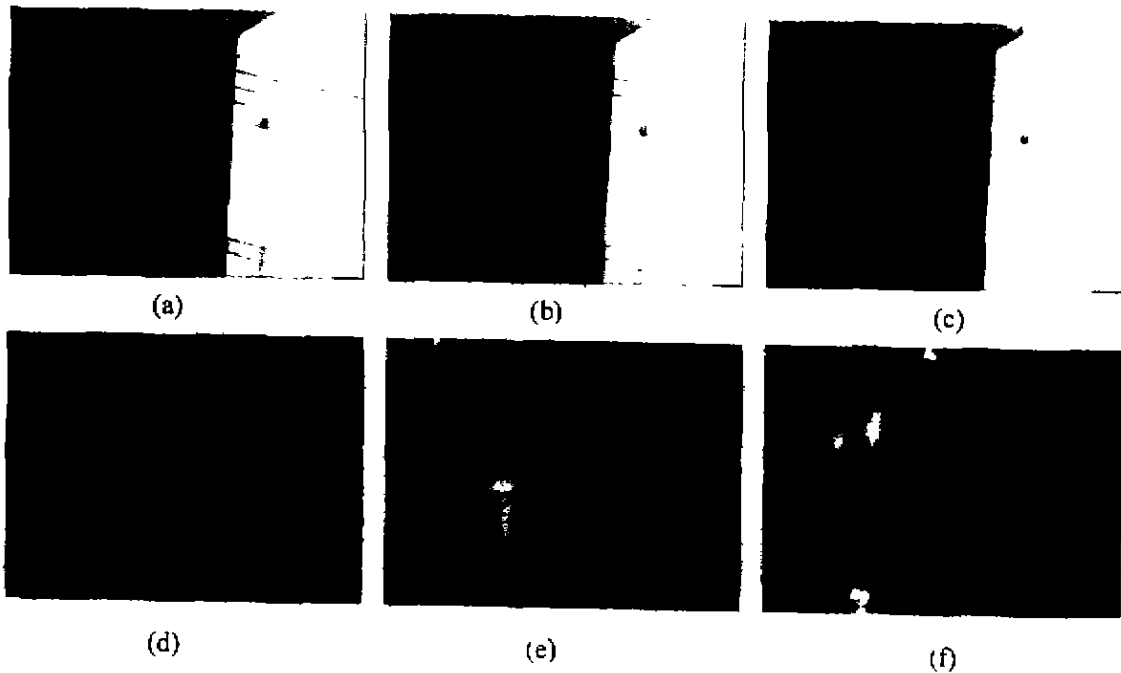


图2 背景物体发生变化时的目标分割  
 Fig.2 Object segmentation when the background changes



图3 复杂背景中的运动目标分割  
Fig.3 Moving object segmentation in clutter scene

图1为两种更新算法对亮度变化的适应曲线,图1(a)和(b)最上方和最下方两条曲线分别为模型的上下阈值,中间的两条曲线为模型的均值和输入的亮度信号.两种算法采用相同的模型更替速率,即对于运动目标的最小运动速度采用相同的下限.从图中可以看出本文算法对于较快的光强渐变,能够及时适应,而混合高斯模型方法由于更新因子是基于高斯分布的,所以模型更新没能及时适应这种变化,而是用新的模型来适应.在新旧模型交替过程中会出现很多噪声.

图2为采用本文算法对雨天中背景物体发生变化时的目标检测.图2(a)、(b)、(c)为视频图像,(d)、(e)、(f)分别为对应的运动目标检测结果.(b)中出现了一辆驶入的汽车停在背景中,从图中可以看出本文算法模型很好的适应了背景发生的变化,并有效的消除了目标的阴影干扰.

图3为本文模型对有较大扰动的复杂背景的目标检测结果,有效的探测出了在树林中行走的人.

对开关灯等光强突变情况也进行了实验,实验表明在检测到光强变化几秒钟之内,本文的模型就可以很好的适应这种光强突变,正确检测运动目标.

我们在PⅢ733上实现了以上算法,每秒钟可以处理46帧 $384 \times 288$ 的彩色视频图像,满足了实用性的要求.

## REFERENCES

- [1] Toyama K, Krumm J, Brumitt B, et al. Wallfollower: principles and practice of background maintenance. *Inter. Conf. on Computer Vision*, 1999. Corfu, Greece, 224—261
- [2] Huwer S, Niemann H. Adaptive change detection for real-time surveillance application. *3rd IEEE Inter. Workshop on Visual Surveillance*, Dublin, Ireland, July 2000, 37—45
- [3] Wern C R, Azarbayejani A, Darrell T, et al. Pfinder: real-time tracking of human body. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 1997, 19: 780—785
- [4] Ridder C, Munkelt O, Kirchner H. Adaptive background estimation and foreground detection using Kalman-Filtering. In: *Proc. International Conf. on recent advances in mechatronics, ICRAM'95*, Istanbul, Turkey, Aug. 1995: 193—199
- [5] Elgammal A, Harwood D, Davis L. Non-parametric model for background subtraction. *ICCV Frame Rate Workshop*, 1999: 246—252
- [6] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking. *IEEE Conf. on CVPR*, 1999:
- [7] Rosin P L, Ellis T. Image difference threshold strategies and shadow detection. *Proc. of the Sixth British Machine Vision Conference*, 1995, 347—356