

# 使用单目摄像头的实时行人计数算法

靳 晶, 万卫兵, 方 涛

(上海交通大学图像处理与模式识别研究所, 上海 200240)

**摘 要:** 为解决智能视频监控系统通道入口处的行人计数问题, 设计在高斯背景建模下的行人检测计数算法。在该算法中, 摄像头垂直放置在入口处的门顶上。为了降低计算复杂度, 提高计算精度, 设置一个感兴趣的检测区域, 在该区域中通过对行人的检测及跟踪进行计数, 并采用队列模型处理以提高检测效果。实验结果表明, 该算法可以有效地对入口行人进行检测计数。

**关键词:** 单目摄像头; 行人计数; 高斯模型

## Real-time Pedestrians Counting Algorithm Using Single Camera

JIN Jing, WAN Wei-bing, FANG Tao

(Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai 200240)

**【Abstract】** Based on background segmentation by mixed model of Gaussian distributions, a real-time people counter using blobs is built to get the pedestrian flow passing through a gate or a door. The proposed algorithm mounts the camera vertically and sets a region of interest to detect and track pedestrians. A queue module is built to help the system achieve good performance. Experimental results show that effective detection and counting accuracy can be achieved.

**【Key words】** single camera; pedestrians counting; Gaussian model

### 1 概述

通道入口处的客流统计是视频监控系统中应用前景很好的一项技术。通过对客流的分析, 可以方便商场、展馆等场所制定决策。基于图像处理技术的实时行人计数技术, 通过对计算机处理彩色摄像头传来的数据进行处理, 能高效、准确、可靠地进行计数。

目前已经有很多论文在这方面进行了研究。有一些方法采用单摄像头进行分析, 在文献[1]中, 通过垂直放置摄像头避免可能出现的人与人之间的重叠问题。还有一些研究是基于多摄像头或使用立体视觉进行分析的, 文献[2]使用 3D 模型帮助处理视频图像, 但计算太复杂, 满足不了实时的要求。

对于单摄像头的通道行人计数, 主要从下面 2 个方向进行研究:

(1) 根据行人的某些特征, 如头部、步态或衣服颜色进行研究。在文献[3]中, 通过检测行人的头部进行计数。在文献[4]中, 使用了颜色直方图进行分析。通常, 这些方法是有效的, 但当行人佩戴帽子或穿了与背景相同颜色的衣物时, 就会产生误差。

(2) 使用背景分割, 找出运动前景, 然后根据分析检测到的运动前景(行人)进行计数。在文献[5]中是对平行拍摄的灰度视频进行处理, 采用 blob 进行分析, blob 为标示提取出的运动物体的矩形框。

本文采用背景分割方法, 摄像头垂直安装在通道顶上, 如图 1 所示。但不同于通常对整个图片帧进行分析的做法, 本文设置了一个感兴趣的检测区域, 在此区域中采用分割效果好但较耗时的高斯混合模型进行背景分割, 保证了实时、准确地检测行人。同时, 对检测出的行人通过一个队列模型帮助跟踪, 克服了摄像头采集图片帧过程中产生的噪点, 也避免了算法初始化阶段可能产生的误判断。实验表明, 在不

太拥挤的情况下, 计数较准确。

### 2 算法原理

在本文中, 摄像头垂直安装于大门和通道的顶上, 如图 1 所示。

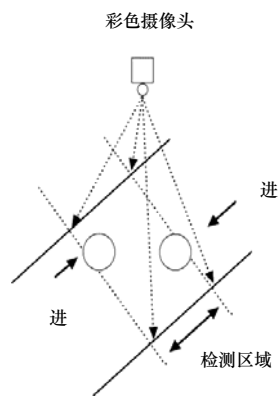


图 1 系统设置

#### 2.1 检测模型

本文采用高斯混合模型<sup>[6]</sup>进行运动物体分割, 其中, 某一个像素点为  $x_n$  的概率表示为

$$p(x_n) = \sum_{j=0}^K \omega_j \eta(x_n; \theta_j) \quad (1)$$

其中,  $\omega_i$  为第  $i$  个分量  $c_i$  的归一化权重, 满足:

$$\sum_{i=1}^K \omega_i = 1 \quad (2)$$

$\eta(x; \theta_k)$  是第  $k$  个高斯元素的分布:

**作者简介:** 靳 晶(1982—), 男, 硕士, 主研方向: 机器视觉; 万卫兵, 讲师; 方 涛, 教授、博士生导师

**收稿日期:** 2009-05-27 **E-mail:** jin\_jing@sjtu.edu.cn



$$\eta(x; \theta_k) = \eta(x; \mu_k; \sigma_k) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{(x-\mu_k)^2}{2\sigma_k^2}} \quad (3)$$

其中,  $\mu_k, \sigma_k$  分别为第  $k$  个高斯的期望和方差。

本模型采用  $k=5$  个高斯分析每一个像素点, 每一个高斯所处理的对象是图像的 R, G, B 信息。高斯混合模型包含 2 个处理过程。

(1) 检测运动物体。对于彩色图像, 每个像素点都有 R, G, B 3 个分量。检测是通过将当前帧的每个像素点与  $k$  个高斯模型进行比较, 如式(4)所示, 如果像素点  $x_j$  不能被任何一个已有的分布描述:

$$x_j \notin [\mu_i - 2.5\sigma_i, \mu_i + 2.5\sigma_i], i=1, 2, \dots, k \quad (4)$$

则说明  $x_j$  是前景, 反之, 说明是背景。

(2) 高斯混合参数的更新。对每一帧都必须更新混合高斯模型中的参数, 使其适应环境变化, 有更好的检测效果。背景是那些证据数目较多、较为稳定的分布, 相应地体现为  $\omega$  较大、 $\sigma$  较小, 所以, 使用

$$p_i = \frac{\omega_i}{\sigma_i} \quad (5)$$

作为高斯分量,  $c_i$  是背景分布可能性的度量。

对于  $k$  个高斯分布, 按照  $\omega/\sigma$  从大到小进行排序, 并按照式(6)选择前  $B$  个分布作为背景。

$$B = \arg \min_b (\sum_{i=1}^b \omega_i > T) \quad (6)$$

其中,  $T$  是一个根据经验选定的阈值。本文选择  $T=0.7$ 。

由于检测出的像素点可能为背景, 也可能为前景, 因此有 2 种更新策略。当  $k$  个高斯模型中没有匹配检测的像素点时, 则最小可能的高斯将被当前的像素取代, 当前像素点的 R, G, B 的值将作为相应的均值  $\mu$ , 方差则采用初始的较高的方差  $\delta$ , 权值  $\omega$  取一个较低的权值。当存在相应的高斯模型匹配检测的像素点时, 第  $k$  个分布的权值将作如下调整:

$$\omega_{k,t} = (1-\alpha)\omega_{k,t-1} + \alpha(M_{k,t}) \quad (7)$$

其中, 当  $t$  时  $k$  匹配当前的样本,  $M_{k,t}$  为 1; 否则,  $M_{k,t}$  为 0。

不匹配的均值  $\mu$  和均值  $\delta$  将保持原来的值。匹配的  $k$  高斯的参数  $\mu$  将采用下式进行更新:

$$\mu_{k,t} = (1-\rho)\mu_{k,t-1} + \rho x_{k,t} \quad (8)$$

方差由下式更新:

$$\sigma_{k,t}^2 = (1-\rho)\sigma_{k,t-1}^2 + \rho(x_{k,t} - \mu_{k,t})^2 \quad (9)$$

在本文中,  $\rho=0.05$ 。实验表明混合高斯模型能有效地检测运动物体, 同时相当灵敏, 一些轻微的阴影也容易被检测出来, 当在户外进行检测时, 必须考虑阴影的情况, 并且高斯混合模型处理量会随着图片像素点数量的增多而增大。

## 2.2 行人跟踪

当获得运动行人的区域后, 如图 2 中的圆形区域所示, 本文使用矩形框描述行人, 这样有助于系统对行人进行处理, 否则, 如使用图 3 中小矩形内的曲线进行处理, 则处理过程变得复杂。本文根据行人在上下帧是否重叠进行跟踪。代表运动行人的矩形框在前后帧上是否重叠取决于摄像头和算法的处理速度。由于本文检测、跟踪的对象是行人, 行人速度相对较慢, 经过实验发现, 使用图 2 所示的通过上、下帧的重叠性进行跟踪完全满足要求。而如摩托车等相对较快的运动物体, 上一帧的摩托车区域可能和下一帧的不重叠, 导致跟踪错误。

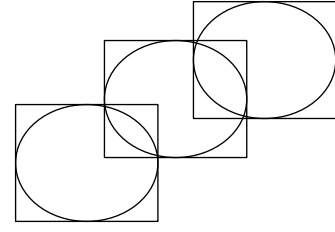


图 2 矩形框跟踪示意图



图 3 矩形框内的曲线

## 3 基于本文算法的行人计数

### 3.1 系统结构

本文系统的流程如图 4 所示。

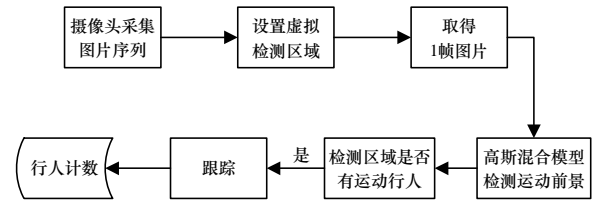


图 4 系统流程

### 3.2 虚拟检测区域的建立

系统的目的是检测并统计入口处的客流, 同时由于高斯混合模型的复杂度, 因此整个图像所有的像素点运用混合高斯模型将耗费很多运算时间并可能达不到实时的要求。另外, 对于入口处周围路过的一些行人, 并不需要进行计数, 如果对他们也进行处理, 势必影响处理速度和结果。所以, 系统在场景中设置了一个虚拟的检测区域, 如图 3 中在进门入口处深色矩形框所示, 只有对进入此区域的行人, 系统才进行处理并计数。

### 3.3 队列模型

在使用高斯混合模型进行背景分割时, 发现对视频序列或者实时处理的摄像头数据, 检测结果中会存在一些噪点。小的、孤立的噪点可以通过形态学、阈值处理滤掉, 但有一些噪点会聚在一起, 如果不对其进行处理, 则系统会误认为是行人, 这将对结果有很大的影响。

通过观察, 发现这些噪点是一些随机的噪声, 并不集中在某一固定位置, 据此, 在系统中设置了一个队列模型, 队列长度为  $N$ , 如图 5 所示, 队列中存放了连续帧中检测到的行人信息, 这些信息可以是位置、大小、颜色等, 便于跟踪, 甚至更深入的处理。检测过程中出现的噪声连续在同一位置出现的概率很低, 不会形成一个连续队列, 通过判断, 本系统认为这是一个噪声, 不对其进行跟踪。通过队列模型, 在  $N$  帧中建立了联系, 使系统判断更加准确。在本文中,  $N=8$ 。





图 5 行人队列

当程序开始时,背景检测模块有一个初始化的过程,通常的做法是设置一个初始化时间 $t$ ,这个时间会和背景检测模块的参数 $\rho$ 有关,但如果 $t$ 设置过长,有些通过的行人会漏检, $t$ 设置过短,会造成一些错误检测。引入队列模型后的实验表明,即使不设置初始化时间,系统也会自动滤掉初始化阶段造成的错误检测,达到更好的效果。

### 3.4 多人情况

当 2 个人或更多人同时进入检测区域时,单目摄像头不能区分进入的是几个人。从经验可知,2 个人的面积肯定比 1 个人面积大,即使是一个人的面积,由于体形、衣服的关系,可能会有大小之差。为了能处理这种情况,采用了文献[4]中的经验公式:

$$PN = 1, N \leq P_{Ni} < 1.4N \quad (10)$$

$$PN = 2, 1.4N \leq P_{Ni} < 2.6N \quad (11)$$

$$PN = 3, N \leq P_{Ni} < 3.6N \quad (12)$$

$$PN = 4, 3.6N \leq P_{Ni} < 4.8N \quad (13)$$

$$PN = 5, N \leq P_{Ni} < 5.8N \quad (14)$$

其中, $PN$ 表示人数; $N$ 表示一个人的像素点个数; $P_{Ni}$ 表示 $PN$ 个人所要的像素点数。当人数超过 5 个时,上述式子将不再准确。

### 3.5 实验结果

为了测试算法结果,搭建了测试平台。其中,使用网络摄像头垂直放置于入口处的门顶上,见图 1。以采集的视频作为样本,视频图像大小为  $320 \times 240$ ,速率为每秒 25 帧,在 1.6 GHz CPU 和 1.5 GB 内存的 P4 PC 上进行实验。

图 6~图 8 为实验结果,其中,大的长方形矩形区域为虚拟的检测区域,可根据实际情况进行设置;小的矩形框代表了行人,左下角为行人编号,据此进行计数。图 6 为室内的测试结果;图 7 为室外的测试结果;图 8 为对偶然出现的车辆的测试结果,当经过的摩托车或自行车速度不是很快时,依然能正确计数。

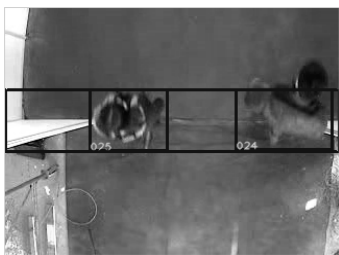


图 6 室内检测效果

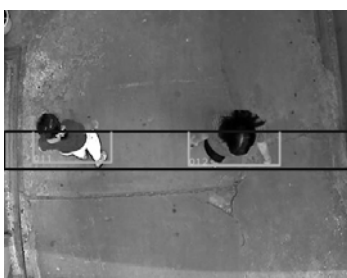


图 7 室外检测效果



图 8 偶然出现车辆情况的处理

在实验场景中 4 对本文算法做了 5 测试,结果如表 1 所示。

表 1 检测结果

环境	描述	实际人数	系统计数	正确率/(%)
1	室内 行人双向通过大门, 间隔 1 m 左右	50	48	96.0
2	室内 行人双向通过大门, 间隔 0.5 m 左右	56	51	91.0
3	室外 行人双向通过通道, 间隔 1 m 左右	48	44	91.0
4	室外 行人双向通过通道, 间隔 0.5 m 左右	44	35	79.5

从实验结果可以看出,室内的实验结果好于室外的结果。在室外环境下,当光线突然变化时,本系统会短暂性失灵,同时也会产生错误的计数,但本系统马上会更新,重新开始正常运行;当行人通过通道时,会使其周围的光线发生变化,可能产生一些细微的阴影,当阴影叠加到其他检测到的行人身上,也对系统有一些影响,会造成误检。同时垂直放置的摄像头并不能完全抑制行人重叠的产生,当发生行人从重叠到分开或相反的过程时,本系统并未进行处理,这也会产生计数上的误差,下一阶段,将对这类问题进行研究。

### 4 结束语

本文实现了基于单目视觉的客流检测系统,该系统将摄像头放置于通道或者大门入口处,通过设置一个虚拟的检测区域,使用高斯混合模型提取运动物体,借助队列模型,对其进行跟踪分析,然后对跟踪结果进行计数。实验表明,该系统可以实时有效地对监控场景进行计数。

### 参考文献

- [1] Rossi M, Bozzoli A. Tracking and Counting Moving People[C]//Proc. of IEEE International Conference on Image Processing. Austin, Texas, USA: IEEE Computer Society, 1994: 212-216.
- [2] Zhao Tao, Nevatia R. Tracking Multiple Humans in Crowded Environment[C]//Proc. of CVPR'04. Washington D. C., USA: IEEE Press, 2004.
- [3] Zhang Xiaowei, Sexton G. Automatic Human Head Location for Pedestrian Counting[C]//Proc. of the 6th IEEE International Conference on Image Processing and Its Applications. [S. l.]: IEEE Press, 1997: 535-540.
- [4] Haritaoglu I, Harwood D, Davis L S. W4: Real-time Surveillance of People and Their Activities[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 22(8): 809-830.
- [5] Chen Thou-ho, Chen Tsong-yi, Chen Zhi-xian. An Intelligent People-flow Counting Method for Passing Through a Gate[C]//Proc. of 2006 IEEE Conference on Robotics, Automation and Mechatronics. Bangkok, Thailand: IEEE Press, 2006: 1-6.
- [6] Stauffer C, Grimson W E L. Learning Patterns of Activity Using Real-time Tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 747-757.

编辑 张正兴