

一种基于改进的卡尔曼滤波的人群密度检测算法

杨清永¹, 谢晓霞²

¹天津中德应用技术大学, 软件与通信学院, 天津

²天津白泽技术有限公司, 天津

收稿日期: 2022年3月13日; 录用日期: 2022年4月14日; 发布日期: 2022年4月21日

摘要

提出了一种改进的基于卡尔曼滤波技术的人群密度检测算法, 该算法通过减背景法识别人群, 使用改进的基于卡尔曼滤波技术进行背景更新, 通过阈值统计的方法增强识别出的前景人群的效果, 最后在此基础上基于像素统计信息进行人群密度计算。实验结果表明, 该算法实现简单, 可以有效的解决卡尔曼滤波技术更新背景时的空洞现象, 比较准确地计算出人群密度。

关键词

视频智能分析, 人群识别, 人群密度, 卡尔曼滤波

A Crowd Density Estimation Based on Improved Kalman Filtering

Qingyong Yang¹, Xiaoxia Xie²

¹School of Software and Communication, Tianjin Sino-German University of Applied Sciences, Tianjin

²Tianjin Baize Technology Co., Ltd., Tianjin

Received: Mar. 13th, 2022; accepted: Apr. 14th, 2022; published: Apr. 21st, 2022

Abstract

The crowd density estimation is the basis for the abnormal behavior detection of crowded scene or group events in intelligent video surveillance system. A crowd density estimation algorithm is given which is based on improved Kalman filtering. The given algorithm will detect crowd by using background subtraction, and update the background by the improved Kalman filtering, and then it

will enhance the foreground crowd by statistic method. Last, it will calculate the crowd density. The experimental result shows that it is simple and can solve the holes in the background caused by Kalman filtering effectively, and the crowd density calculated is accurate.

Keywords

Intelligent Video Analysis, Crowd Detection, Crowd Density, Kalman Filtering

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 概述

近年来, 公共场合的安全问题受到越来越多的关注, 若某一区域内的人群高度拥挤则容易造成踩踏等群体事件, 给社会带来巨大损失。通过视频智能分析技术, 计算监控场景的人群密度, 及时提供预警信息, 是预测群体事件的基础, 因此, 人群密度计算、人数统计等智能监控技术受到越来越多的关注和研究[1] [2]。

人群密度检测的基础是进行目标人群识别。利用运动目标在视频图像序列中边缘、颜色、纹理等方面的差异, 检测出运动目标人群[3] [4]; 利用边缘检测算法, 获取人的轮廓、位置等运动状态[5] [6], 是实现人群密度检测、人数统计、群体事件检测等的基础, 但易受到干扰、光线变化、遮挡等因素的影响, 常见算法有减背景法、帧差法和光流法。卡尔曼滤波器[7] [8]利用前一时刻信号与噪声的状态空间的估计值来预测当前时刻的估计值, 能够在较大的噪声中还原原来的信号, 是减背景法中一种常见的背景更新算法[9]。本文采用减背景法来识别视频序列里的目标人群, 利用卡尔曼滤波实现背景更新, 使用均值法解决使用卡尔曼滤波进行背景更新时容易产生空洞现象的情况, 在正确更新背景的基础上, 通过统计阈值法增强识别出的前景人群的效果, 最后在此基础上进行基于像素统计信息的人群密度计算。

2. 卡尔曼滤波

卡尔曼滤波是一种时变线性系统递归滤波器, 是一个最优化自回归数据处理算法。卡尔曼滤波器可以解决很多领域的问题, 可应用在通信、控制、军事等领域, 其基本思想是: 以最小均方误差为最佳估计准则, 利用当前时刻的系统状态和控制量的估计值, 能够从一系列存在测量噪声的数据中预测下一时刻系统的状态, 并且当前时刻和下一时刻系统状态都是最优值[10]。卡尔曼算法可以程序化实现, 十分适合处理实时数据, 因此近年来被应用于计算机视觉处理领域, 例如目标识别, 目标跟踪等, 是一种高效最优算法。

3. 改进的基于卡尔曼滤波的人群密度检测算法

3.1. 基于改进的卡尔曼滤波的背景更新算法

减背景法是一种简单有效的目标识别方法。卡尔曼滤波是一种常用的减背景法中背景更新算法, 该算法对视频中的背景图像使用时域低通滤波进行更新, 可将图像的快速变化过程中的缓变部分被分离出来, 整个背景更新过程相当于一个视频序列的多帧降噪过程[11]。使用卡尔曼滤波进行背景更新采用如下公式(1)和(2)。

$$B(i, j, t+1) = B(i, j, t) + [\alpha_1 * (1 - M(i, j, t)) + \alpha_2 * M(i, j, t)] * D(x, y, t) \quad (1)$$

$$D(x, y, t) = I(i, j, t) - B(i, j, t) \quad (2)$$

其中: $B(i, j, t)$ 和 $B(i, j, t+1)$ 分别是第 t 帧和 $t+1$ 帧的背景图像, $D(x, y, t)$ 表示 t 帧图像 $I(i, j, t)$ 和背景图像 $B(i, j, t)$ 的差值, $M(i, j, t)$ 是当前运动目标图像被二值化以后的结果, α_1, α_2 是一组常量, 取值在 0 和 1 之间。

该方法将运动前景也带入了背景更新函数中, 因此在前景对象出现频率不高时运行效果较好, 但一旦背景被目标遮挡一段时间后, 该算法将会失效, 导致检测的目标出现拖尾现象, 严重时甚至会扩大到整个区域, 并且算法中 $M(i, j, t)$ 是由 $D(x, y, t)$ 经过一系列计算而来, 因此这种自反馈机制将导致滤波效果不稳定, 最终影响背景更新的准确性。为解决该问题, 我们对该算法进行改进, 在自反馈的过程中, 不再直接通过前景直接迭代, 而是通过计算多帧均值来替代原算法中的前景数据。令 $G(x, y, t)$ 为 t 帧内像素 (x, y) 的灰度值; $C(x, y, t)$ 为近 15 帧图像的均值。定义中间变量 $N(x, y, t)$, 我们按照如下公式(3)和(4)进行背景更新。

$$N(x, y, t) = \begin{cases} \left(1 - \alpha + \alpha \frac{C(x, y, t)}{B(x, y, t)}\right) G(x, y, t) & G(x, y, t) \neq 0 \\ C(x, y, t) \alpha & G(x, y, t) = 0 \end{cases} \quad (3)$$

$$B(x, y, t+1) = \begin{cases} N(x, y, t) & M(x, y, t) = 1 \\ G(x, y, t) & M(x, y, t) = 0 \end{cases} \quad (4)$$

其中: 若 $N(x, y, t) > 255$ 则令 $N(x, y, t) = 255$ 。其中, α 为比例因子, 本实验取值为 0.1, 改进算法中, 背景更新充分考虑了视频时域特征变化因素, 利用 15 帧图像均值信息来更新背景图像, 因此可以很好的解决背景遮挡和光线突变对背景更新的影响。

3.2. 基于统计信息的人群识别

在进行完背景更新后, 接下来进行目标人群识别, 传统的方法是直接用前景图减去背景图像就得到要检测的目标, 但这种方法经测试最终会将部分前景作为背景而识别不出来, 因此, 我们在计算前景时设定阈值 T , 然后求取当前图像和图像均值的方差, 当方差小于阈值时, 将该像素判定为背景, 而当方差大于阈值时, 将该像素判定为前景, 具体方法如式(5), 其中 n 取 15, T 取值为 5。

$$M(x, y, t) = \begin{cases} G(x, y, t) - B(x, y, t) & \frac{1}{n} \sum_n (G(x, y, t) - C(x, y, t))^2 \leq T \\ G(x, y, t) & \frac{1}{n} \sum_n (G(x, y, t) - C(x, y, t))^2 > T \end{cases} \quad (5)$$

3.3. 基于像素特征的人群密度计算

识别出人群之后, 最后进行密度计算, 密度计算可以采用基于像素特征的方法和基于纹理特征的方法来实现[12]。基于像素特征的密度计算方法主要通过减背景法来提取前景人群占据的空间大小, 该方法简单, 计算量小, 但有遮挡时会影响人群检测效果; 基于纹理特征的人群密度技术方法计算量大, 可以解决遮挡问题, 但不适合背景复杂的场景。本文中为提高运算速度, 采用基于像素特征的处理方法来判断人群密度, 在计算出前景目标人群后, 首先对人群边缘进行增强, 然后进行平滑和腐蚀填充, 最后根据像素特征对人群密度进行估计。令 $Density(t)$ 为 t 帧的人群密度, 其计算公式如(6)所示。

$$\text{Density}(t) = \sum_{M(x,y,t)} \frac{1}{W \times H} \quad (6)$$

其中, W 和 H 分别表示图像的宽度和高度。

4. 实验结果

4.1. 测试环境

实验采用的是 Intel(R)Core(TM)i3CPU@2.7 GHz, 集成显卡, 操作系统为 Windows。测试用视频为某路口白天的真实监控录像数据, 视频图像分辨率为 352×288 , 帧率为 25 帧/秒。

4.2. 效果对比测试

1) 卡尔曼算法改进前后的背景更新效果对比

首先进行算法改进前后的背景效果对比测试, 随机选取视频中的第 1、20、88 和 180 帧视频进行测试, 原始图像如图 1(a)所示, 使用改进前的卡尔曼滤波算法进行的背景更新结果如图 1(b)所示。

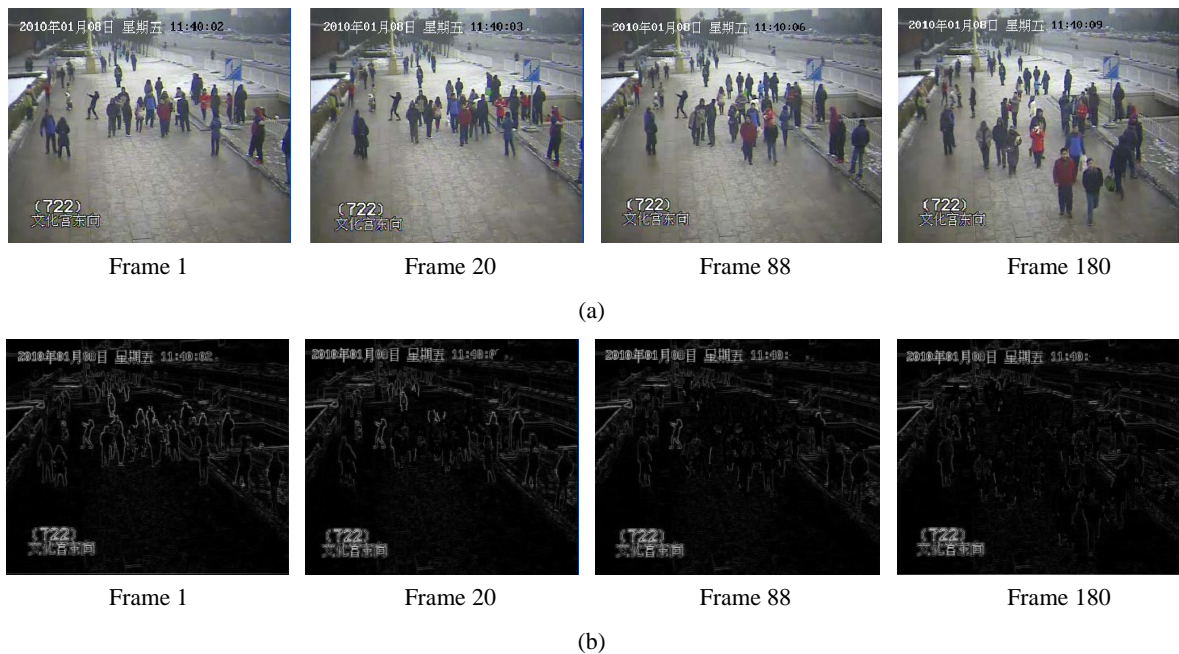


Figure 1. The effect of background update by using Kalman algorithm. (a) Raw image; (b) Background image

图 1. 使用卡尔曼算法进行背景更新。(a) 原始视频图像; (b) 背景图像

从图 1(b)的第 20、88 和 180 帧可以看出, 由于卡尔曼算法更新时会利用前景信息, 因此, 部分前景图像会被认为是背景图像, 因此导致一些背景图像中出现空洞现象, 并随时间面积逐渐扩大。接下来测试使用改进后的卡尔曼算法进行运算, 背景更新的结果如图 2(a)所示, 可以看出使用改进后算法得到的背景更新结果已经解决了原算法背景更新空洞问题, 每一帧背景图像中, 背景轮廓都比较完整。

使用改进的卡尔曼滤波算法继续进行前景人群检测, 得到的前景人群如图 2(b)所示, 从该图可以看出, 基本识别出了场景里的人群而滤掉了场景里的固定背景, 识别效果较为准确。

2) 本文算法和常见的目标识别算法时间对比

应用本文算法和常见的帧差法、高斯混合和光流法进行识别对比实验, 测试处理一帧图像的平均时间见表 1。由表 1 可见, 本文算法较高斯混合和光流法两种算法在计算时间上要快很多, 比差分法稍慢。



Figure 2. The effect of improved Kalman filter algorithm. (a) The effect of background update; (b) The effect of crowd recognition

图 2. 改进的卡尔曼滤波算法运行效果。(a) 背景更新效果; (b) 人群识别效果

Table 1. The comparison of operation time

表 1. 运算时间对比

算法	本文	高斯混合	光流法	差分法
时间	<3.5 毫秒	13 毫秒	大于 100 毫秒	<1 毫秒

3) 人群密度计算测试

接下来进行人群密度计算测试, 在视频文件中分别选取人数较少、一般、较多的帧, 每种情况分别选 4 帧, 具体帧图像如下图 3 所示。

分别使用本文算法和差分法计算人群密度, 图 3 中每帧图像的人群密度检测结果如表 2 所示。对应每帧图像可看出使用本文在各种场景下计算出的人群密度和实际情况基本相符, 都比较准确, 而使用差分法计算出的人群密度结果不稳定, 并且和目测的密度差别较大。

Table 2. The result of crowd density calculation

表 2. 人群密度计算结果

帧号	人数较少		帧号	人数中等		帧号	人数较多	
	密度			密度			密度	
	本文	差分		本文	差分		本文	差分
2781	0.220595	0.343257	431	0.339301	0.248086	7307	0.478210	0.344648
2969	0.187816	0.143801	565	0.313447	0.300219	7314	0.480025	0.329023
3051	0.168087	0.200925	2215	0.335316	0.148220	7336	0.460089	0.364554
3285	0.265862	0.344736	2390	0.368065	0.265783	10360	0.411606	0.289201

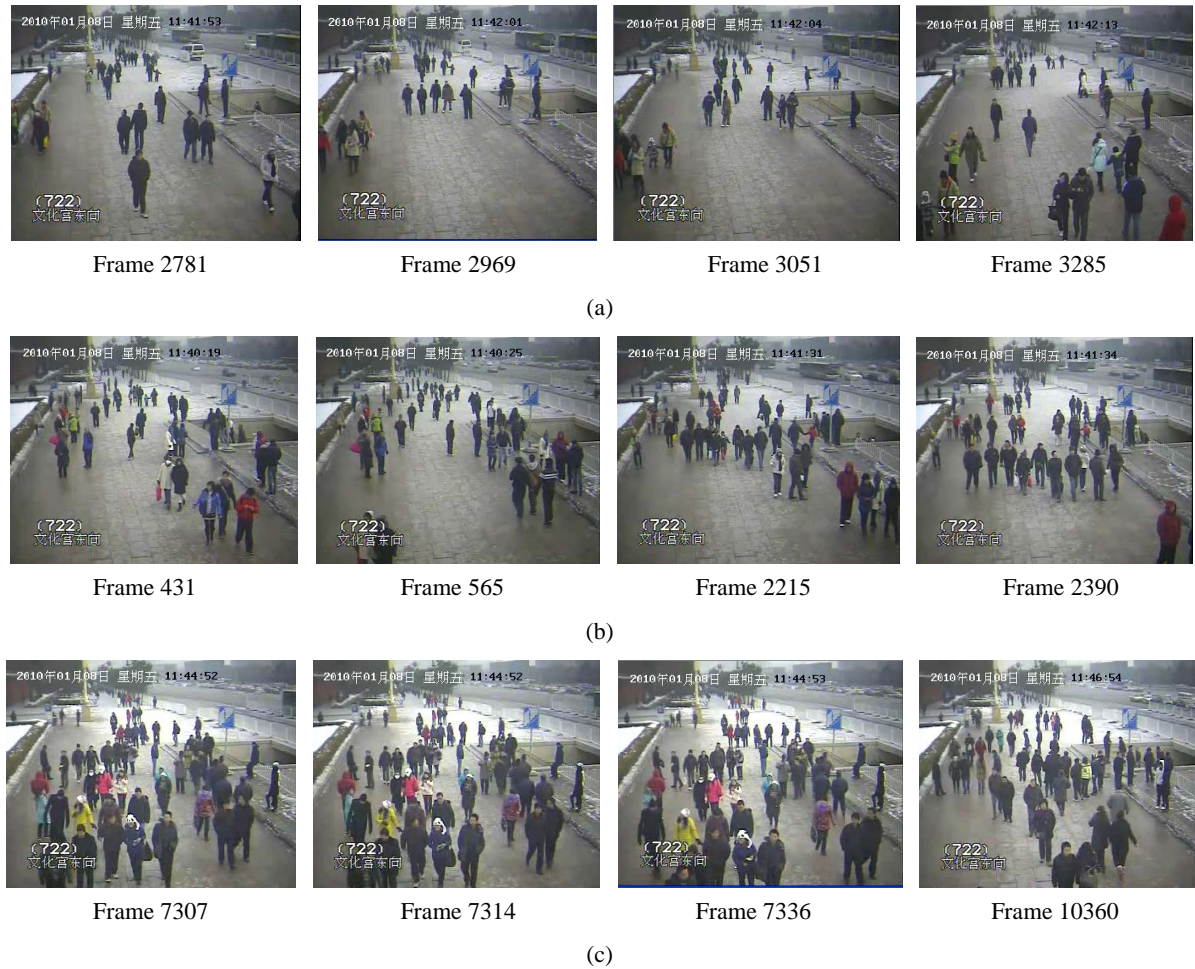


Figure 3. Crowd density detection images. (a) Fewer people; (b) Medium number; (c) More people

图 3. 人群密度检测图像。(a) 人数较少; (b) 人数一般; (c) 人数较多

5. 结论

视频监控系统中, 人群密度检测技术是群体事件和人群异常行为判断的前提, 准确的识别是后续各种智能分析工作的关键。但由于目标人群的识别易受到干扰、光线变化、阴影、遮挡等因素的影响而导致检测不准确或失败, 所以好的运动人群目标检测算法具有相当重要的意义。本文对视频序列中人群运动目标检测进行了研究和探讨, 使用改进的卡尔曼算法进行背景更新, 解决了背景空洞问题, 使用均值方法提高了人群的检测效果, 最后根据像素特征对人群密度进行估计计算。通过实验验证, 该算法可以准确识别出场景内的运动人群, 计算出的密度也较为准确, 但前景识别效果图显示仍有噪点存在, 因此今后应继续优化目标人群识别效果, 提高人群密度计算的准确性。

参考文献

- [1] 赵宏伟, 徐亮, 王治, 等. 基于尺度融合的密集人群计数[J]. 计算机系统应用, 2021, 30(10): 1-11.
- [2] 朱弈霖. 基于视频的公共场所人群异常行为检测系统的设计与实现[D]: [硕士学位论文]. 南京: 东南大学, 2020.
- [3] 潘尚考, 刘光杰, 刘伟伟, 等. 基于目标检测的轨道交通客流密度估计系统及其方法[P]. 中国专利, CN112633210A, 2021.
- [4] 李章维, 胡安顺, 王晓飞. 基于视觉的目标检测方法综述[J]. 计算机工程与应用, 2020, 56(8): 1-9.

-
- [5] 李平, 喻洪流. 基于两种分类标准的目标检测算法综述[J]. 计算机应用研究, 2021, 38(9): 2582-2589.
- [6] 徐光祐, 曹媛媛. 动作识别与行为理解综述[J]. 中国图象图形学报, 2009, 14(2): 190-196.
- [7] Ali, A. and Terada, K. (2010) A General Framework for Multi-Human Tracking Using Kalman Filter and Fast Mean Shift Algorithms. *Journal of Universal Computer Science*, **16**, 921-937.
- [8] 章学静, 陈禾, 杨静. 结合卡尔曼滤波和 Mean Shift 的抗遮挡跟踪算法[J]. 北京理工大学学报, 2013, 33(10): 1056-1061.
- [9] 程旭. 复杂场景下的目标跟踪算法研究[D]: [博士学位论文]. 南京: 东南大学, 2015.
- [10] 梁超, 王亮, 刘红云. 基于扩展卡尔曼滤波的实时视觉 SLAM 算法[J]. 计算机工程, 2013, 39(8): 231-234+238.
- [11] Kim, Z.W. (2008) Real Time Object Tracking based on Dynamic Feature Grouping with Background Subtraction. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2008*, June 2008, 1-8.
<https://doi.org/10.1109/CVPR.2008.4587551>
- [12] 周波, 李俊峰. 结合目标检测的人体行为识别[J]. 自动化学报, 2020, 46(9): 1961-1970.