

一种基于区域划分的人群密度估计快速方法^{*}

张 燕 黎 宁 刘福美

(南京航空航天大学信息科学与技术学院 南京 210016)

摘 要 公共场景监控下的人群密度估计是公共安全管理中的一个重要内容。针对大场景监控情况,研究了一种基于区域划分的智能化人群密度估计快速方法。在帧差提取出的彩色视频背景的基础上,按视觉比例将大场景分成子区域,然后对某个子区域采用像素统计的方法实现人群密度估计,其余子区域采用自适应比例的方法实现人群密度快速估计。实验结果表明该方法在实际应用中简单、有效、快速,能为公共场所的预警系统提供有力的帮助。

关键词 人群密度估计;彩色背景提取;区域划分;自适应比例

中图分类号 TP391

A Fast Crowd Density Estimation Based on Region Division

Zhang Yan Li Ning Liu Fumei

(College of Information Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016)

Abstract Crowd density estimation in public scene surveillance is an important issue of public security. A intellectualized method of fast crowd density estimation based on region division was studied to resolve the people number estimation of full big scene. The full big scene was divided into some needed sub-areas according to the visual scale after the color background had been extracted by the frame difference. The method of pixel statistics was used to estimate the density of a certain sub-area, and then a method of adaptive scale was used to realize the fast estimation of the other sub-areas. Experimental result shows that this method is simple, effective and fast in practical application. And it can also provide powerful help to the public warning system.

Key Words crowd density estimation, color background extraction, region division, adaptive scale

Class Number TP391

1 引言

近年来随着城市人口密度的急剧增大,许多公共基础设施,如机场、地铁站、车站等经常会迎来短期的人流高峰^[1]。人群的高度拥挤容易引起各种突发事件,如果不能对人群迅速有效地疏散分流,不仅会影响居民日常生活,更可能造成重大的人员伤亡和财产损失。因此对公共基础设施进行人群密度估计是十分必要的,人群密度估计有着广泛的应用前景和研究价值。传统的人群密度估计是通过闭路电视对不同场景进行人工监控,由监控室的

工作人员对监视场景进行人工判断。传统方法不仅费时费力,而且监控人员容易因为疲惫等原因而忽视监控场景的突发情况,从而造成不可挽回的后果^[2]。本文利用闭路电视监控系统^[3],使用智能化方法对监控场景的人群密度进行快速估计。

Davies 和 Chow 提出了基于像素统计的图像处理方法,通过背景减的方法计算前景人群所占的像素数,对人群密度进行估计^[4]。该方法简单有效,其中背景提取的实际效果和人群人数的计算实时性是实现人群密度估计最重要的部分。在一些大监控场景中,整个大场景的训练工作量往往很

^{*} 收稿日期:2010 年 10 月 14 日,修回日期:2010 年 11 月 24 日

基金项目:民航基金项目(编号:E1002-071B)资助。

作者简介:张燕,女,硕士研究生,研究方向:图像处理、模式识别、计算机视觉等。黎宁,女,副教授,研究方向:图像处理、目标识别与跟踪和计算机视觉等。刘福美,女,硕士研究生,研究方向:图像处理、数字视频处理等。

大,因此对整个大场景直接使用基于像素统计的方法进行人群密度估计往往不具有现实意义。针对这些实际问题,本文根据场景大小将场景分成 N 块子区域分别处理,某个子区域直接采用基于像素统计的方法进行人群密度估计,其余子区域利用自适应比例方法实现人群密度的快速估计。

2 彩色视频背景帧差提取

传统提取视频背景的原理^[5]是:对于视频流中的前后两帧灰度图像,如果两帧图像的对应像素点灰度变化不大(小于某个阈值 T),则认为此像素点是背景区的像素点;如果两帧图像的对应像素点灰度变化较大(大于某个阈值 T),则认为此像素点是目标区的像素点。彩色视频图像不能简单使用传统提取灰度背景的方法。本文提出了彩色视频的背景帧差提取方法:将前后两帧彩色图像对应像素点的 RGB 三分量分别单独运用灰度图像的处理方法,然后再将得到的属于背景像素点的 RGB 三分量组合成彩色背景。

彩色视频的背景帧差提取算法:取一个时间段视频的 n 帧图像序列,为了完整地提取彩色视频背景,通常取 $n \geq 100$ 。在 R 分量上:将前后相邻两帧同一像素点的 R 分量值相减,若绝对差值小于阈值 T ,则认为是该点是背景区域像素点,将该点的 R 分量

值加入到背景像素 R 分量累计结果 CBr 中,否则认为是目标区域像素点。重复进行 $n-1$ 次帧差,统计该像素点判为背景区域像素点的次数 Nr ,则彩色视频背景中该像素点的 R 分量值 Br 等于 R 分量累计结果 CBr 除以次数 Nr 。 G 和 B 分量采用同样的处理方法,最后将 RGB 三分量组合成彩色背景。

以 R 分量为例解释算法:

$$CBr(i,j)=$$

$$\begin{cases} CBr(i,j) & |f1r(i,j)-f2r(i,j)| \geq T \\ CBr(i,j)+f2r(i,j) & |f1r(i,j)-f2r(i,j)| < T \end{cases} \quad (1)$$

$$Nr(i,j)=$$

$$\begin{cases} Nr(i,j) & |f1r(i,j)-f2r(i,j)| \geq T \\ Nr(i,j)+1 & |f1r(i,j)-f2r(i,j)| < T \end{cases} \quad (2)$$

$$Br(i,j)=CBr(i,j)/Nr(i,j) \quad (3)$$

式(1)和式(2)中, $f1r(i,j)$ 和 $f2r(i,j)$ 分别代表前一帧和后一帧像素点 (i,j) 处的 R 分量值。从第一帧到第 n 帧共进行 $n-1$ 次帧差, $CBr(i,j)$ 代表的是背景像素点 (i,j) 处 R 分量累计结果, $Nr(i,j)$ 代表的是 $n-1$ 次帧差中像素点 (i,j) 被判为背景像素点的次数,均初始化为0。根据式(3),得出彩色视频背景中像素点 (i,j) 的 R 分量值 $Br(i,j)$ 。阈值 T 的选择有一定的随意性,主要靠经验或者大量试验总结,本论文中取10。



图 1 帧差提取的彩色视频背景



图 2 背景减前的人群图像

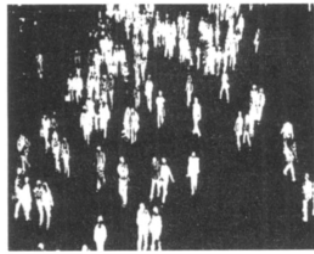


图 3 背景减二值化的人群图像

图 1 是帧差提取的彩色视频背景,图 2 是第 980 帧背景减之前的人群图像,图 3 是第 980 帧图像背景减并二值化后的人群图像。

3 区域划分的自适应比例快速估计人数

选择遮挡现象不严重的大场景作为本文研究对象,并且针对由于摄像机光轴与大地平面所成角度使得人群在图像上呈现近大远小的现象,提出了基于区域划分的自适应比例快速估计人数的方法。

3.1 场景子区域纵向划分

在一幅较大的场景中,如果直接采用基于 Da-

vies 等人提出的人群人数和像素数成正比的关系来估计整个大场景的人群人数,训练样本时需人工数出整个大场景中的实际人数。由于整个场景较大,训练样本时的工作量往往很大。针对这一问题,首先将大监控场景按照视觉比例(具体比例见下文 3.2.1)由近至远划分成实际监控面积基本相等的子区域,然后只需对其中某个子区域训练样本,其余子区域可采用下文提出的基于区域划分的自适应比例快速估计人数的方法。大场景的区域划分效果如图 4 所示。

3.2 自适应比例快速估计人数

同一目标处于同一场景的不同位置时,在图像

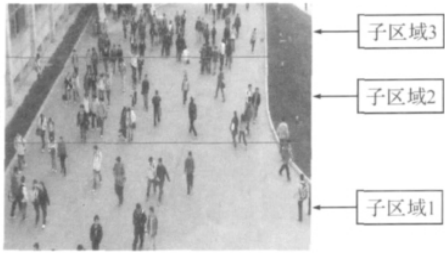


图 4 大场景的区域划分效果图

上成像大小也会不同。当目标距离摄像头较近时，成像比较高比较宽，即所占像素数比较多，反之较少。为了提高估计人数的准确率，本文提出一种自适应比例的方法，即依据摄像机的成像规律对同一基准目标进行高度和宽度比例缩放，以适应图像中不同区域的人群目标。

3.2.1 摄像机二维成像高度比例

根据计算机立体视觉测量技术中经常采用的针孔模型^[6]，得到图 5 所示的二维成像高度比例模型。图 5 中， (Y_c, Z_c) 为摄像机坐标系，原点 O_c 定义在摄像机的光心， Z_c 轴为摄像机的光轴； (Y_w, Z_w) 为世界坐标系， Z_w 轴垂直于水平面，将原点 O_w 定义在 Z_w 轴与 Y_w 轴的交点处； H 为摄像机与地面的垂直拍摄高度， θ 为摄像机与地面的拍摄角度，拍摄到的图像垂直于 Z_c 轴。

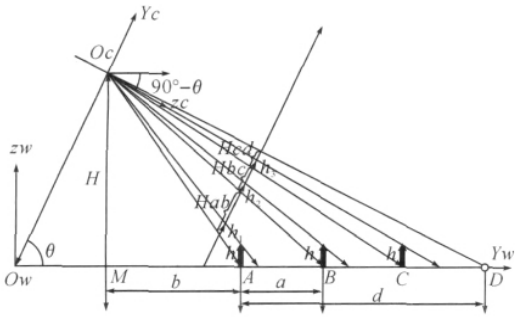


图 5 摄像机二维成像高度比例模型

通过在实际场景中的测量，按照离摄像机的距离由近及远均匀标定出三个区域，如大场景地面区域长度范围是图 5 中 AD 线段，三个均匀长度范围区域是 AB, BC, CD 线段，然后计算出 AB, BC, CD 线段在成像图像中对应的各个区域范围纵向长度 H_{ab}, H_{bc} 和 H_{cd} ，并以此作为 3.1 节中场景子区域纵向划分的比例关系。假设地面上一个高为 h 的基准目标分别站在等距离间隔的 A, B, C 三点位置，这三个位置处的基准目标对应于成像照片中的成像长度分别为 h_1, h_2 和 h_3 。

根据图 5 中给出的摄像机坐标系与世界坐标系之间的几何关系，可以推导出：

$$\frac{h_2}{h_1} = \frac{(bH+cH-ch)(b+c)(a+b)}{(ah+bH+cH-ch)(a+b+c)b},$$
$$\frac{h_3}{h_1} = \frac{(bH+cH-ch)(b+c)(2a+b)}{(2aH+bH+cH-ch)(2a+b+c)b} \quad (4)$$

其中 $a = \frac{d}{3}, c = H \cot \theta$ 。

3.2.2 摄像机二维成像宽度比例

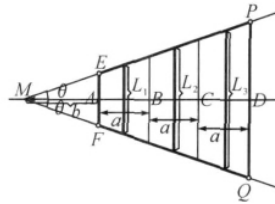


图 6 摄像机二维成像宽度比例模型

当目标距离摄像头较近时，目标在图像上成像比较宽^[7]，即所占像素数比较多，反之较少。图 6 中， EF 线段是大场景中距离摄像头最近的区域边缘的

实际宽度， PQ 是距离摄像头最远的区域边缘的实际宽度。通过在实际场景中的测量，测出三个均匀区域 AB, BC 和 CD 中间部分的实际宽度 L_1, L_2, L_3 ，并以此计算基准目标在三个均匀区域的摄像机二维成像宽度比例关系。

根据图 6 中给出的几何关系，可以推导出：

$$\frac{L_2}{L_1} = \frac{b + \frac{3a}{2}}{b + \frac{a}{2}}, \frac{L_3}{L_1} = \frac{b + \frac{5a}{2}}{b + \frac{a}{2}} \quad (5)$$

其中， a, b 定义与摄像机二维成像高度比例模型一致。

3.2.3 统计像素估计人群人数

对背景减得到的二值化人群图像计算出由近到远的三大区域（即图像纵向上由下到上三个区域）的前景人群像素总数。采用基于人群人数和像素数成正比的关系来估计人群密度的方法，首先只要人工数出训练样本中每幅图像的任一区域（本论文采用区域 1）的人数，然后运用最小二乘法^[8]拟合出一条对应于该区域的实际人群人数和像素数的关系直线。其余区域不需要经过训练样本拟合出相应的关系直线，直接利用基准目标在不同区域所成的成像大小比例关系进行自适应比例快速估计人数，即分别将子区域 2 和子区域 3 计算出的二值化人群像素总数乘以相应的成像大小比例关系就可以得到二值化人群像素数。

成像大小比例关系由二维成像高度和宽度比例导出：

$$Ratio_{21} = \frac{h_1 \times L_2}{h_2 \times L_1}, Ratio_{31} = \frac{h_1 \times L_3}{h_3 \times L_1} \quad (6)$$

其中， $Ratio_{21}, Ratio_{31}$ 分别代表同一基准目标在

子区域 1 中通过二值化得到的像素数与在子区域 2 和子区域 3 中得到的像素数之比。

4 实验结果及分析

对南航江宁校区拍摄的学生下课高峰时的视频进行抽帧,从抽帧图像序列中选取 1800 帧作为实验图像,如图 2 所示。为了体现算法的真实有效性,训练样本使用前 799 帧,测试样本使用后 1001 帧。

按图 4 将图像划分为 3 个子区域,用帧差提取的彩色视频背景对每帧测试样本图像进行背景减并二值化,根据拟合直线计算每帧图像子区域 1 的人群人数,子区域 2 和子区域 3 采用基于区域划分的自适应比例方法快速估计人数。图 7 表示的是子区域 1 的拟合直线,图 8 表示从第 800 帧到第 1800 帧大场景图像的实际总人数和估计总人数之间的曲线关系。从图 8 可以看出,本文算法实现的大场景人群密度估计与实际人群密度非常接近,高密度时由于人群重叠现象过于严重从而导致误差较大,低密度时估计的大场景人群总人数与实际总人数基本一致。

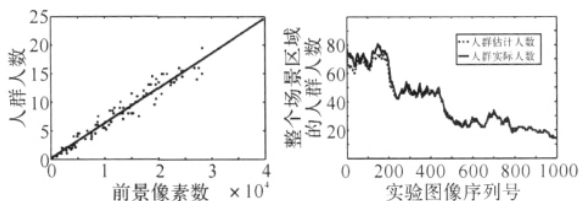


图 7 子区域 1 的
拟合直线

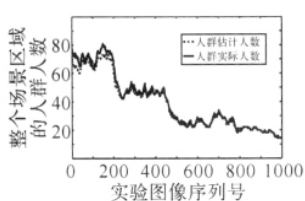


图 8 大场景下测试帧的
人数变化(共 1001 帧)

考虑到人群运动的速度有限,所以一秒钟处理两帧图像已经足够,实验表明,每帧图像的平均处理时间为 0.125s,因此可以满足实时性。

5 结语

在帧差提取的彩色视频背景基础上,首先根据视觉比例将大监控场景合理划分为若干子区域,然

后根据摄像机二维成像模型得出划分区域的自适应比例,从而快速实现对大场景下的智能化人群密度估计。实验结果表明该方法在实际应用中简单、有效、迅速,能够实时发现公共基础设施中的异常情况,为公共场所的安全预警提供有效的帮助。在遮挡现象十分严重的情况时采用基于像素统计的方法效果不是十分理想,可以采用纹理分析^[9]等其他方法来对人群密度估计做进一步完善的研究。

参考文献

- [1] 梁英宏,王知衍,曹晓叶,等. 视频图像理解在客流统计中的应用[J]. 计算机工程与设计,2008,29(5):1203~1206
- [2] 衣淑凤,黄祥林,沈兰荪. 智能化人群监控技术研究[J]. 测控技术,2003,22(5):22~24
- [3] Davies A C, Yin J H, Velastin S A, et al. Crowd monitoring using image processing[J]. IEE Electronics and Communication Engineering Journal,1995,7(1):37~47
- [4] 胡波,李晓华,沈兰荪. 场景监控中的人群密度估计[J]. 电路与系统学报,2007,12(6):19~22
- [5] 李娜,方卫宁. 基于视频流的地铁人群目标识别[J]. 北京交通大学学报,2006,30(1):96~99
- [6] SHA Ling, LU Chao-hui. A Camera Calibration Method for Stereoscopic Vision Measurement[J]. Machinery,2003,41(5):10~11
- [7] Stefan Huwer, Heinrich Niemann. Adaptive Change Detection for Real-Time Surveillance Applications [C]//Proceedings of Third IEEE International Workshop 2000. Dublin: IEEE Computer Society Press, 2000:37~46
- [8] A. Chan, Z. S. Liang, N. Vasconcelos. Privacy preserving crowd monitoring: Counting people without people models or tracking[J]. CVPR,2008:1~7
- [9] Wu Xinyu, Liang Guoyuan, Lee Ka Keung, et al. Crowd Density Estimation Using Texture Analysis and Learning[C]//IEEE International Conference on Dec, 2006:214~219

(上接第 24 页)

- [8] 周相兵. 基于 Ontology 的语义 Web 服务聚合自动机研究及应用[C]//Proceedings of the 27th Chinese Control Conference, Kunming, Yunnan, China, IEEE Press, 2008:719~723
- [9] Alexander Maedche. Ontology learning for the semantic web[M]. Kluwer Academic Publishers: Boston/Dordrecht/London,2002:11~21
- [10] LI YUHUA, BANDAR Z, McLEAD D. An approach for measuring semantic similarity between words using

multiple information sources[J]. IEEE Transactions on Knowledge and Data Engineering,2003,15(4):871~882

- [11] 陈友玲,张永阳,孙亚南,等. 基于 Petri 网的敏捷化生产调度建模方法[J]. 计算机集成制造系统,2010,16(3):598~602
- [12] Yu-Liang Chi, Hsun-Ming Lee. A formal modeling platform for composing web services[J]. Expert Systems with Applications,2008,34(2):1500~1507