

## 复杂场景中的全景密度估计方法

贺 昆<sup>1</sup>, 刘 舟<sup>2\*</sup>, 戚鹿宁<sup>1</sup>, 杨 恒<sup>2</sup>, 朱 桐<sup>1</sup>, 刘燕伟<sup>1</sup>, 周计美<sup>2</sup>

(1. 公安部 警卫局, 北京 100031; 2. 北京正安融翰技术有限公司, 北京 100088)

(\* 通信作者电子邮箱 liuzhou441@163.com)

**摘 要:** 为了克服传统密度估计方法受限于算法配置工作量大、高等级密度样本数量有限等因素无法大规模应用的缺点, 提出一种基于监控视频的全景密度估计方法。首先, 通过自动构建场景的权重图消除成像过程中射影畸变造成的影响, 该过程针对不同的场景自动鲁棒地学习出对应的权值图, 从而有效降低算法配置工作量; 其次, 利用仿真模拟方法通过低密度等级样本构建大量高密度等级样本; 最后, 提取训练样本的面积、周长等特征用于训练支持向量回归机(SVR)来预测每个场景的密度等级。在测试过程中, 还通过二维图像与全景地理信息系统(GIS)地图的映射, 实时展示全景密度分布情况。在北京北站广场地区的深度应用结果表明, 所提全景密度估计方法可以准确、快速、有效地估计复杂场景中人群密度动态变化。

**关键词:** 全景监控; 人群密度估计; 灰度共生矩阵; 支持向量回归; 智能视觉监控

**中图分类号:** TP391.4 **文献标志码:** A

### Panoramic density estimation method in complex scene

HE Kun<sup>1</sup>, LIU Zhou<sup>2\*</sup>, QI Luning<sup>1</sup>, YANG Heng<sup>2</sup>, ZHU Tong<sup>1</sup>, LIU Yanwei<sup>1</sup>, ZHOU Jimei<sup>2</sup>

(1. VIP Protection Department, Ministry of Public Security of the People's Republic of China, Beijing 100031, China;

2. Beijing Zheng An Rong Han Technology Company Limited, Beijing 100088, China)

**Abstract:** The traditional density estimation method cannot be applied in large-scale because of large configuration workload and limited number of high-level density samples. To overcome that problem, a panoramic density estimation method was proposed using video surveillance. Firstly, a method for automatically obtaining the perspective map was introduced, which effectively processed the influence of projective distortion and reduced the workload of deployment. Secondly, high-level density samples were generated from low-level samples using the simulation method. Finally, a Support Vector Regression (SVR) was trained for predicting the density of each scenario after extracting features such as area and perimeter. Moreover, in the detection process, the panoramic crowd density distribution was demonstrated through the mapping relation from 2D image to Geographic Information System (GIS) map. The application of the method in the square of Beijing North Railway Station shows that it can deal with complex scenes effectively and robustly.

**Key words:** panoramic video surveillance; crowd density estimation; grey level dependence matrix; Support Vector Regression (SVR); intelligent video surveillance

## 0 引言

在人口流动频繁且社会活动不断增加的今天, 由于人群密度过大而造成人员伤亡的事件屡见不鲜。因此, 在日常生活中, 对于人们经常出入的地铁、车站、超市等地点的人群密度进行估计是十分必要的。人群密度估计可以应用于社会、经济的很多方面, 它不仅可以解决人群拥挤带来的安全问题, 还可以应用于交通数据统计和交通调度中, 提高城市公交系统的利用率和效率。因此, 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通过监控场景的闭路电视人工进行

监控, 费时费力且缺乏客观性。随着计算机和图像处理技术的发展和广泛应用, 智能化的人群密度监控系统成为人们研究的重点。

在人流密度估计的研究中, 按研究对象的不同可以分为三类: 基于像素特征的方法、基于个体特征的方法以及基于人群团块(group)的方法。

基于像素特征的方法就是基于像素级特征进行密度估计, 例如文献[1-2]通过背景差分的方法来提取前景人群占据的空间大小, 以及运用边缘提取的方法检测人群对象的边缘长度, 通过这两类像素级特征对人群密度进行估计。这种方法较为简单, 计算量小, 当人群密度较低时效果较好, 但是

**收稿日期:** 2013-12-20; **修回日期:** 2014-02-21。 **基金项目:** 公安部重点研究计划项目(201202ZDYJ014); 北京市科技新星计划项目(Z131101000413083); 北京市科技计划项目(Z131100005313004)。

**作者简介:** 贺昆(1969-), 男, 北京人, 工程师, 主要研究方向: 安防监控、公安保卫; 刘舟(1981-), 男, 湖南株洲人, 博士研究生, 主要研究方向: 模式识别、计算机视觉、图像处理; 戚鹿宁(1979-), 男, 北京人, 助理工程师, 主要研究方向: 安防监控、公安保卫; 杨恒(1983-), 男, 河南商丘人, 硕士研究生, 主要研究方向: 模式识别、计算机视觉、图像处理; 朱桐(1983-), 男, 北京人, 助理工程师, 硕士研究生, 主要研究方向: 安防监控、公安保卫; 刘燕伟(1983-), 男, 山东青岛人, 助理工程师, 硕士研究生, 主要研究方向: 安防监控、公安保卫; 周计美(1984-), 女, 河北邢台人, 硕士研究生, 主要研究方向: 模式识别、计算机视觉、图像处理。

当人群密度较高时,由于人群遮挡严重会导致结果误差很大。之后为了处理遮挡问题,文献[3-5]基于特征点检测、特征点跟踪以及轨迹聚类的方法进行人数统计,这类方法能够有效处理拥挤场景,但是计算复杂度非常高,不适合监控等对实时性能要求很高的场景。

基于个体特征的方法中常用的是基于行人检测的方法获得场景中每个人的位置信息,通过累计场景中的人数获得密度估计信息。这类方法对人流不大的场景有非常好的估计效果,但是当拥挤非常严重时,往往不能准确检测出场景中的每个人,误差会很大。之后也有基于人头和人脸<sup>[6-7]</sup>等个体特征信息进行密度估计的方法,这类基于人体部分特征的方法可以有效缓解部分遮挡的问题,但该方法通常对观测角度和姿态等比较敏感。

基于人群团块(group)的方法是一种将运动目标检测的结果或者图像中的感兴趣区域(Region Of Interest, ROI)作为一个整体去估计的方法。该方法通常包含特征提取与分类模块,例如文献[8]通过提取灰度共生矩阵等纹理特征来描述图像中的ROI,并通过高斯回归方法估计人流密度。该类方法实时性好,准确率高,本文也是基于该类方法进行研究。

基于人群团块的方法在理想情况下能够获得较高的精度,但是在实际应用中确实较难得到应用,究其原因主要有以下几点:

- 1) 算法针对不同的场景配置过程复杂,难于大规模推广。
- 2) 对于绝大多数场景,人流密度很高的情况很少出现,因此可用于描述高等级人流的样本很少,很难精确估计高等级的情况。
- 3) 绝大多数密度估计方法都是基于分镜头进行分析,无法展现全场景的密度分布情况。

本文针对上述传统密度估计方法的不足,进行了以下改进:

1) 通过行人检测方法统计出场景中各个位置行人的尺寸信息,并自动构建场景的权重图用于处理射影畸变造成的影响。该过程可以大幅减少算法的配置过程,有利于大规模应用。同时还提出了一种可以鲁棒估计权重图的方法,用于处理背景中可能存在的行人误检测问题。

2) 算法通过仿真的方式构建高等级人流时的样本,该方法将有效补充样本不足无法精确描述对应等级的缺点。

3) 通过二维图像与地理信息系统(Geographic Information System, GIS)地图之间的映射关系,构建全景人流密度估计,可以实时动态掌握全景人流密度的分布和变化趋势。

本文算法已经在北京北站站前广场地区得到了成功应用,具有很高的精度和实用性。

## 1 人群密度估计算法

本文算法框架如图1所示,主要包含两个部分:离线学习过程和实时密度估计过程。离线学习过程中首先通过自学习的方式自动获取场景的畸变矫正权值图,用于矫正特征提取过程因为透视变化造成的特征变化(例如:图像近大远小所造成的面积变化)。自动获得权值图后,特征矫正过程与文献[9]相似。在获取训练模型的过程中,首先利用仿真方法通过低密度样本构建高密度样本数据。在获得大量学习样

本后,将提取特征用于训练支持向量回归(Support Vector Regression, SVR)模型。在实时密度估计中,首先对场景进行运动目标提取并针对运动目标区域提取特征,之后基于训练过程获得SVR模型估计人流密度,最后通过二维图像和全景GIS的映射关系将分镜头数据映射至全景GIS地图中进行全景展示。

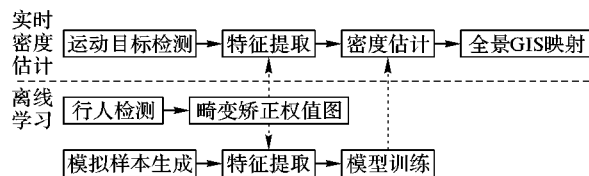


图1 算法流程

### 1.1 模拟样本生成

在人流密度估计的实际应用中,通常很难获得高等级密度的样本用于模型的学习,在这里将利用低等级时获得的样本数据仿真场景中高等级的样本数据。其具体步骤如下:

- 1) 将图像分为  $N \times M$  个子区域。
- 2) 运动目标检测,并进行连通域分析。将检测到的运动目标及其掩码按其所在子区域分别进行保存,保存的样本示例如图2所示。



图2 样本示例

- 3) 去除保存样本中人数大于1的样本数据。
- 4) 人群所在区域选择。首先确定模拟样本中含有人数  $n$ , 之后随机从  $N \times M$  个子区域中选择  $n$  个区域用于假定人群可能存在的位置。注意为了产生更好的训练样本,  $n$  个子区域互不相同。

5) 训练样本生成。在每个选择的子区域随机选择一个保存的样本数据与黑色背景模块进行融合,不同子区域的样本若存在重叠时,重叠关系随机。图3展示了模拟的样本以及相应掩码的示例。

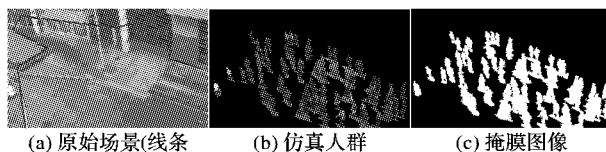


图3 仿真样本示例

### 1.2 畸变矫正权值图

由于透视投影变化的影响会使三维空间中相同大小的人在图像的不同位置会有很大的差异,距离摄像头越近人通常会越大。反映在人流密度估计方面,透视变化的影响通常会造相同数量的人群在图像的不同位置具有不同的特征,例如:面积、纹理和边缘等。与文献[9]类似,假设权值图在场景中符合二次曲线分布。

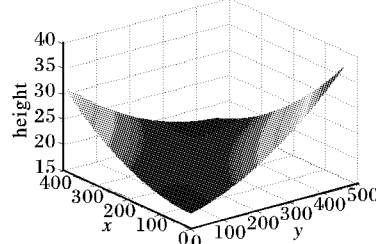
传统的权值图获取方法,例如文献[9],需要针对不同场景手工标定同一个人图像在不同位置的大小信息,花费大量的人力物力,严重影响了该功能的大范围推广。本文通过行人检测的方法<sup>[10]</sup>自动获取行人在场景中不同位置大小的统计信息,然后通过二次曲线拟合的方法自动获取各个权值分布的曲线方程。最后曲线中的各个点将会依据曲线的最大值进行归一化处理。在特征矫正过程中,与文献[9]一致,基于

面积的特征将会除以对应位置处权值,而对于边缘特性的特征,例如边缘数量、方向直方图,统计时将除以对应位置处权值的均方根。

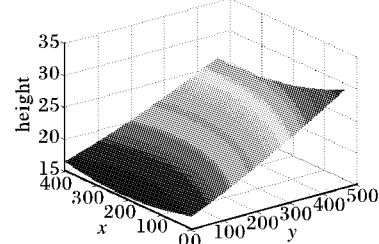
在实际场景的应用中,由于背景的纹理信息通常不会发生改变,因此如果一旦在背景的某个位置由于纹理相似出现了行人检测的误报,该误报将会在相近的位置频繁出现,导致了大量的噪声样本(这样的噪声数量远远高于其他位置样本数量),如图4(a)所示。这些样本会使二次曲线拟合失败,如图4(b)所示(图4(b)和(c)中的坐标原点对应于图像的左上



(a) 场景图片中行人检测结果



(b) 改进前的拟合结果



(c) 改进后的拟合结果

图4 行人检测中的误检测以及二次曲线拟合结果

### 1.3 运动目标检测

混合高斯模型<sup>[11]</sup>是目前背景建模较为成功的方法之一,能够较好地适应背景的动态变化(如晃动的树叶、变化的光照等),其原理是通过自适应统计每个像素在时间域上的分布信息,构造每个像素的时间域上的多高斯背景模型,即对于固定位置相机采集的图像,用 $K$ 个高斯成分组成的混合高斯模型表示像素在时间域上的概率分布模型,即图像中像素 $j$ 在时刻 $t$ 取值为 $x_j$ 的概率为:

$$P(x_j) = \sum_{i=1}^K \omega_{j,t}^i \cdot \eta(x_j; \mu_{j,t}^i, \Sigma_{j,t}^i) \quad (1)$$

其中: $\omega_{j,t}^i$ 表示在时刻 $t$ 时像素 $j$ 的混合高斯模型中第 $i$ 个高斯模型的权重; $\mu_{j,t}^i$ 和 $\Sigma_{j,t}^i = (\sigma_{j,t}^i)^2 I$ 分别表示第 $i$ 个高斯模型的均值和协方差,这里 $\sigma$ 表示标准差, $I$ 表示单位矩阵, $\eta$ 表示高斯概率密度函数。

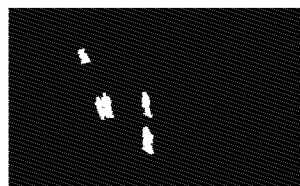
像素 $j$ 的混合高斯模型描述了其特征值 $x_j$ 在时间域上的概率分布,在判断识别前背景时,认为属于成功的背景模型的特征值为背景;反之则可认为是前景。所谓成功的背景模型由以下方式产生,根据每个高斯模型的权重与标准差之比排序后,取前 $B_j$ 个高斯模型作为成功的背景模型, $B_j$ 根据式(2)计算:

$$B_j = \arg \min_b \left( \sum_{i=1}^b \omega_{j,t+1}^i > T \right) \quad (2)$$

阈值 $T$ 为认为当前像素点的特征值在时域上属于背景的概率值。图5展示了一些运动目标检测结果。



(a) 原始图片



(b) 前景检测结果

图5 前景检测示例

### 1.4 特征提取

在构建畸形矫正图以及获得运动目标检测区域后,对整个运动目标区域提取特征用于密度估计。这里所用的特征包

括基于像素的特征以及基于纹理的特征。基于像素信息的特征,如面积,较为简单,计算量小,当人群密度较低的时候效果较好,但是当人群密度较高时,严重的遮挡会导致结果误差很大。此时,基于纹理的特征能够有效克服上述缺点。

本文所用的特征包括:面积、周长、边缘方向直方图、面积周长比、边缘像素点数量、灰度共生矩阵、Minkowski 维度<sup>[12]</sup>。这些特征在一定程度上可以反映视频中人群的数量以及遮挡的关系。

为了克服以上问题,将场景图片分成不同的子区间,对各个子区间中行人检测的样本数据进行聚类,并以聚类中心作为最终的样本数据利用最小二乘法进行曲线拟合。这样的处理方式可以有效减少因为背景纹理的因素造成的频繁误检测对最终权值曲线拟合的影响。图4(c)显示了改进后的拟合结果。从图4可看出,拟合的鲁棒性有了大幅提升。

括基于像素的特征以及基于纹理的特征。基于像素信息的特征,如面积,较为简单,计算量小,当人群密度较低的时候效果较好,但是当人群密度较高时,严重的遮挡会导致结果误差很大。此时,基于纹理的特征能够有效克服上述缺点。

本文所用的特征包括:面积、周长、边缘方向直方图、面积周长比、边缘像素点数量、灰度共生矩阵、Minkowski 维度<sup>[12]</sup>。这些特征在一定程度上可以反映视频中人群的数量以及遮挡的关系。

### 1.5 密度估计

在获取特征之后,通过支持向量回归机(SVR)来获取当前帧的密度估计值。支持向量回归机建立在统计学习理论之上,是结构风险最小化一致性原则的一种具体实现结构。在模型预测中,引用支持向量回归机提高了传统模型的抗干扰能力,扩大了其应用范围,取得了较为理想的预测结果。用支持向量回归机建立图像特征和图像人数等级的数学关系。根据训练样本的特征训练SVR模型,然后SVR模型根据输入图像特征估计图像人数等级。SVR模型预测图像等级的表达式为:

$$f(x) = \sum_{i=1}^n (a_i - a_i^*) K(x_i, x) + b \quad (3)$$

其中: $a_i, a_i^*$ 为拉格朗日乘子, $K(x_i, x_j)$ 为满足 Mercer 条件的核函数, $x_i$ 是支持向量, $x$ 是图像特征, $f(x)$ 是预测的图像等级数。 $f(x)$ 输出为连续值,最后通过各等级阈值化得到预测等级输出结果。

### 1.6 全景 GIS 映射

传统的密度分析方法无法给客户展示监控场景的全景密度分布。在这里通过插值的方法构建二维图像中像素点与全景GIS图中的坐标关系,并将二维图像上对应运动区域的密度结果映射至GIS图上进行全景展示。

为了节省计算量,首先将图像分成 $N \times M$ 个子块,之后通过手工标记的方式获得二维图像四个顶点对应于全景GIS图的位置并通过插值的方式获得子块顶点所对应的GIS图坐标。在密度实时估计过程中,图像中对应于前景区域的子块顶点将会被给予所估计的密度值,而非前景区域将会给予0



值。对应子块顶点的密度值将会通过之前获取的映射关系投影至二维 GIS 图上,并通过不同等级所对应不同的色度条进行实时展示。在二维 GIS 图中,其余的点将通过插值的方式获得其等级值,对于重叠区域将求取等级平均值。这里需要注意的是由于是进行全景展示,场景中不同位置处的分镜头应该具备一致的分级标准。图6展示了全景密度估计效果以及相关的分镜头密度估计情况,在图中用不同颜色代表不同的密度等级。

上述算法除了可用于人流密度外还可以很方便地推广用于估计车流密度。GIS 图左下角就是车流密度估计的结果。

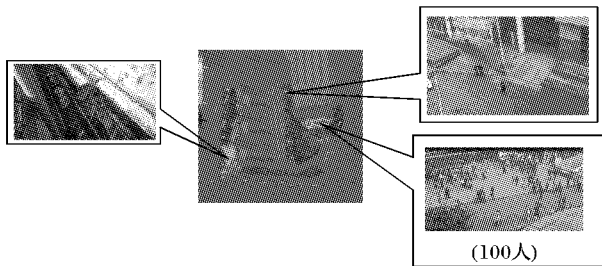


图6 全景密度估计示例

## 2 实验结果及分析

在实际测试中,以北京北站的数据作为测试数据测试本文算法的有效性以及精确度。根据客户需求定义场景的人流密度等级如表1所示。

表1 密度等级定义

密度等级	人流密度
一级(低密度)	$\leq 5$
二级(中低密度)	5 ~ 15
三级(中高密度)	15 ~ 30
四级(高密度)	30 ~ 60
五级(报警密度)	$> 60$

图7展示了密度估计结果的示例,估计的等级数值在相应的包含运动目标的子区域顶点以及左上角进行了展示(原始视频高宽比为16:9)。

在实际应用中,通常高等级密度的样本较难获取,因此在测试过程中对第四等级和第五等级生成了仿真样本。本文方法与文献[9]方法进行的比较(文献[9]中并未采用仿真样本进行密度估计)如表2所示。从表2中可看出本文方法在复杂环境下对密度估计具有较高的估计精度。同时,在采用仿真样本后,在低密度时估计性能基本不变的情况下,高密度时的密度估计性能有了大幅度提高,已满足实用要求。从表2还可看出,本文方法在未采用仿真样本时的精度达到甚至略

表2 测试性能

密度等级	样本数			本文方法的正确率/%		文献[9]方法的正确率/%
	训练	仿真	测试	采用仿真样本	仅采用真实样本	仅采用真实样本
一级	800	0	800	92.25	92.50	91.90
二级	800	0	800	93.50	93.75	93.85
三级	800	0	800	91.25	90.50	90.40
四级	100	700	200	90.10	71.50	70.00
五级	80	720	150	90.70	65.30	67.30
平均正确率				92.80	89.27	89.10

微超过了文献[9]方法,从而也说明了本文自动获取畸形矫正权值图的有效性。



图7 密度估计示例

## 3 结语

本文实现了一种智能化人群密度估计算法,提出了一种高等级下的样本生成策略、一种全自动鲁棒获取畸形矫正权值图方法以及一种密度全景展示方案。本文研究内容已经在北京北站广场区域得到应用。对超过2000幅图像的实验结果证明,该方法可达到90%以上的正确识别率,同时配置方式简单,能够满足智能化人群密度监控的要求并在实际场景中得到了应用。

在之后的工作中将进一步优化回归算法提高估计精度,从而进一步实现对人数的精确估计而不仅仅是对人数的分级估计。

(下转第1723页)

时间复杂度高的问题,本文在 Criminisi 算法的基础上,提出了基于随机映射的快速图像修复算法,该算法搜索与待修复区域在结构和纹理相似的样本区域,去除冗余的样本搜索空间,优化了基于置信因子和边缘信息的优先级计算方法,改进了最优相似块的计算方法,增强了图像结构传播的正确性。实验结果表明,该方法不仅提高了修复速度,而且提高了图像修复效果。但是,由于该贪婪算法并非每一步都能选择到最优的样本匹配块,一旦出现错误就会造成误差的严重累积,影响图像修复的效果。因此,如何挖掘图像中其他有益信息,智能加强结构传播的正确性还待进一步研究。

#### 参考文献:

- [1] BERTALMIO M, SAPIRO G, CASELLES V, *et al.* Image inpainting [C]// Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM Press, 2000: 417–424.
- [2] EFROS A A, LEUNG T K. Texture synthesis by non-parametric sampling [C]// ICCV 1999: Proceedings of the 7th IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 1999, 2: 1033–1038.
- [3] CRIMINISI A, PEREZ P, TOYAMA K. Region filling and object removal by exemplar-based image inpainting [J]. IEEE Transactions on Image Processing, 2004, 13(9): 1200–1212.
- [4] CHENG W-H, HSIEH C-W, LIN S-K, *et al.* Robust algorithm for exemplar-based image inpainting [EB/OL]. [2012-10-10]. [http://cmlab.csie.org/new\\_cml\\_website/media/publications/Cheng-2005-RAE.pdf](http://cmlab.csie.org/new_cml_website/media/publications/Cheng-2005-RAE.pdf).
- [5] LIU J, LU D. Image inpainting via weighted optimization [J]. Journal of Image and Graphics, 2011, 16(4): 528–532. (刘建明, 詹东明. 采用加权优化的图像修复 [J]. 中国图象图形学报, 2011, 16(4): 528–532.)
- [6] MICHAEL A. Synthesizing natural textures [C]// Proceedings of the 2001 Symposium on Interactive 3D Graphics. New York: ACM Press, 2001: 217–226.
- [7] LU P, CHEN M, MA L, *et al.* Fast structural image completion [J]. Journal of Image and Graphics, 2010, 15(6): 931–935. (路平, 陈敏刚, 马利庄, 等. 快速结构化图像修复 [J]. 中国图象图形学报, 2010, 15(6): 931–935.)
- [8] TONG X, ZHANG J, LIU L, *et al.* Synthesis of bidirectional texture functions on arbitrary surfaces [J]. ACM Transactions on Graphics, 2002, 21(3): 665–672.
- [9] BARNES C, SHECHTMAN E, FINKELSTEIN A, *et al.* Patch-Match: a randomized correspondence algorithm for structural image editing [J]. ACM Transactions on Graphics, 2009, 28(3): Article No. 24.
- [10] HE K, SUN J. Computing nearest-neighbor fields via propagation-assisted KD-trees [C]// CVPR 2012: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2012: 111–118.
- [11] BARNES C, SHECHTMAN E, GOLDMAN D B, *et al.* The generalized patchmatch correspondence algorithm [C]// Computer Vision — ECCV 2010, LNCS 6313. Berlin: Springer-Verlag, 2010: 29–43.
- [12] KOPPEL, KIENZLE W, DRUCKER S, *et al.* Quality prediction for image completion [J]. ACM Transactions on Graphics, 2012, 31(6): Article No. 131.
- [13] KOMODAKIS N, TZIRTAS G. Image completion using efficient belief propagation via priority scheduling and dynamic pruning [J]. IEEE Transactions on Image Processing, 2007, 16(11): 2649–2661.
- [1] DAVIES A C, YIN J H, VELASTIN S A, *et al.* Crowd monitoring using image processing [J]. IEEE Electronics and Communication Engineering Journal, 1995, 7(1): 37–47.
- [2] CHO S Y, CHOW T W S, LEUNG C T. A neural-based crowd estimation by hybrid global learning algorithm [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 1999, 29(4): 535–541.
- [3] BROSTOW G J, CIPOLLA R. Unsupervised Bayesian detection of independent motion in crowds [C]// CVPR 2006: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2006: 594–601.
- [4] RABAUD V, BELONGIE S J. Counting crowded moving objects [C]// Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2006: 705–711.
- [5] KOPF M. A contribution to traffic state estimation by means of image processing methods [J]. Information Sciences and Technologies Bulletin of the ACM Slovakia, 2013, 9(5): 41–48.
- [6] WU B, NEVATIA R. Detection of multiple, partially occluded humans in a single image by Bayesian combination of edgelet part detectors [C]// ICCV 2005: Proceedings of the Tenth IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2005: 90–97.
- [7] LIN S-F, CHEN J-Y, CHAO H-X. Estimation of number of people in crowded scenes using perspective transformation [J]. IEEE Transactions on System, Man, and Cybernetics, Part A: Systems and Humans, 2001, 31(6): 645–654.
- [8] ZHANG E, CHEN F. A fast and robust people counting method in video surveillance [C]// Proceedings of the 2007 International Conference on Computational Intelligence and Security. Piscataway: IEEE Press, 2007: 339–343.
- [9] CHAN A B, LIANG Z-S J, VASCONCELOS N. Privacy preserving crowd monitoring: counting people without people models or tracking [C]// CVPR 2008: Proceedings of the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2008: 1–7.
- [10] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]// CVPR 2005: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2005: 886–893.
- [11] 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.
- [12] MARANA A N, da FONTOURA COSTA L, LOTUFO R A, *et al.* Estimating crowd density with Minkowski fractal dimension [C]// Proceedings of the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Piscataway: IEEE Press, 1999, 6: 3521–3524.

(上接第 1718 页)

#### 参考文献: