

# 利用鱼眼相机对人群进行运动估计

胡学敏<sup>1</sup> 郑 宏<sup>2</sup> 郭 琳<sup>1</sup> 熊饶饶<sup>1</sup>

1 湖北大学计算机与信息工程学院,湖北 武汉,430062  
2 武汉大学电子信息学院,湖北 武汉,430072

**摘 要:**人群运动估计是人群行为分析的重要步骤。特定场景的人群运动分析和监控,是维护公共安全和社会稳定的一个必要措施,也是视频监控领域的一个研究难点。利用鱼眼相机视场大、无视觉盲区的优点,提出了一种基于特征点光流的人群运动估计方法。首先,采用一种基于面积反馈机制的混合高斯背景差分方法,对原始视频图像进行预处理,并利用圆拟合的方法获取兴趣区域;其次,为了在保证准确描述人群目标的同时提高算法的实时性,提出一种基于边缘密度非均匀采样的人群特征点提取方法来描述运动的人群目标,并利用 Lucas & Kanade 光流法计算光流场;最后,为了解决远近人群的尺寸大小不一致的问题和鱼眼相机的畸变问题,采用鱼眼相机的透视加权模型,计算人群运动加权统计直方图,获取人群在鱼眼图像中的全局运动方向和速度。实验结果表明,针对密集的人群,该方法能有效、实时地估计人群的运动方向和速度,为人群行为分析提供有力的研究基础。

**关键词:**鱼眼相机;特征点;光流;透视加权模型;人群运动估计;方向直方图;速度直方图  
**中图法分类号:**P231;TP751 **文献标志码:**A

人群密集的场所,如果不能在人流高峰期及时地进行分流疏散,极易酿成灾祸。2015 年元旦上海外滩的踩踏事件就是一个反面例子。因此,针对特定场所分析人群的运动行为,获取运动参数,是公共安全领域的一个重要课题。

目前典型的运动估计方法主要基于简单匹配技术(如块匹配法)、频域技术(如 Gabor 滤波法)和微分技术(如光流法)三种技术。块匹配法可以实现人群的运动估计,但容易陷入局部最优<sup>[1,2]</sup>。Gabor 滤波能获取较准确的运动矢量,但是较为耗时<sup>[3]</sup>。光流法在理想的情况下能够检测独立运动的对象,不需要预先知道场景的任何信息,可以精确地计算出运动物体的速度,但对于光线的变化较敏感<sup>[4-6]</sup>。另外,由于光流法的计算量大,如果对前景图像每个点都进行光流的计算的话,难以保证算法实时性;而取特征点计算稀疏光流则难以精确提取轮廓复杂的运动目标的形状。考虑人群运动估计的实时性和准确性,以及人群形状和大小的多变性,块匹配和 Gabor 滤波难以应用于人群运动估计。光流法要满足人群运动估计的实时性,则需解决光线变化以及寻找到合适的特

征点来兼顾准确性与实时性的问题。

现有人群运动估计方法大多基于传统枪式相机。传统相机具有视场小、盲点多等缺点,运动人群目标容易脱离监控范围。鱼眼相机是一种典型的全景相机,具有视场大,无盲点和无需图像拼接等优点,现主要用于对少数行人或者交通的监控<sup>[7,8]</sup>,目前少有研究者利用鱼眼相机对人群运动进行整体分析。

本文利用鱼眼相机,提出一种基于特征点光流的人群运动估计方法。首先,对输入图像序列进行预处理;其次,为了在保证准确描述人群目标的同时提高算法的实时性,提出一种基于边缘密度采样的人群特征点提取方法,并利用 Lucas & Kanade 法计算光流场;最后,采用鱼眼相机的透视加权模型,计算人群运动矢量,获取全局运动方向和速度。实验结果表明,本文提出的方法能有效、实时地估计人群的运动方向和运动速度。

## 1 基于鱼眼相机的人群运动估计

本文提出的基于鱼眼相机的人群运动估计计算

收稿日期:2016-12-06  
项目资助:湖北省大学生创新训练项目(201510512041)。  
第一作者:胡学敏,博士,主要从事图像处理和计算机视觉研究。huxuemin2003@163.com  
通讯作者:郭琳,博士。10837330@qq.com

法流程图如图 1 所示。首先,对输入的图片序列进行预处理,包括兴趣区域(region of interest, ROI)的划分和人群目标的分割;其次,为了保证准确性和降低计算复杂度,本文提出了一种基于边缘密度采样的特征点提取算法,对人群目标进行特征点提取;然后针对特征点,利用 Lucas & Kanade 光流法计算光流场;最后,为了有效、准确地计算全局的人群运动参数,采用基于鱼眼相机的透视加权模型,并利用此模型计算全局的人群运动方向和速度。因此,本文提出的人群运动估计方法包含预处理、特征点光流计算和运动矢量计算三个部分。

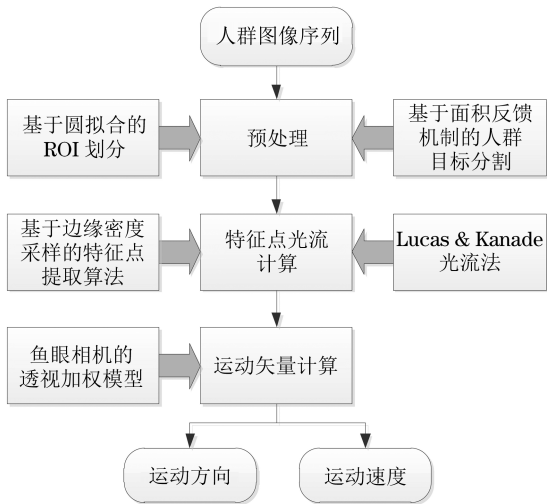


图 1 基于鱼眼相机的人群运动估计流程图  
Fig.1 Flowchart of Crowd Motion Estimation Approach Based on the Fisheye Camera

1.1 预处理

本文使用一种基于半径约束的最小二乘圆拟合方法对鱼眼图像进行 ROI 划分<sup>[9]</sup>。图 2(a)中的圆为圆拟合的结果。拟合得到的圆心( $x_c, y_c$ )和半径  $R$  即为有效像素区域的中心坐标点和半径。ROI 的划分方法为,到图像中心点的距离小于  $k \times R$  的区域划分在 ROI 内,其他区域划分到 ROI 以外,其中  $k$  为经验值。由于透视加权模型基于“所有人群目标分布在同一个水平面上”这个前提条件<sup>[10]</sup>,因此需要将楼梯等不在同一个水平

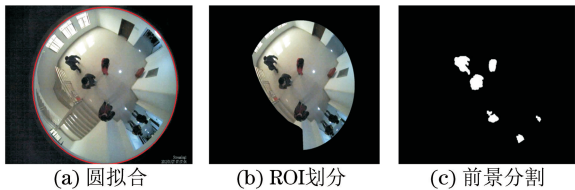


图 2 ROI 划分和前景分割结果示意图  
Fig.2 Results of Preprocessing and Foreground Segmentation

面上的区域人为地划分到 ROI 以外。最终 ROI 划分结果如图 2(b)所示。

常用运动分割的方法有帧间差分法、背景差分法和光流法三种。为解决光线变化问题,本文采用基于面积反馈机制的混合高斯背景差分法进行目标分割<sup>[11]</sup>。提取后的人群目标如图 2(c)所示。

1.2 特征点光流计算

一般情况下,前景区域计算得到的光流矢量存在噪声。如图 3(b)所示,白色箭头代表光流场的方向。故需要找到一种特征点的提取方法,减少光流矢量的计算量。

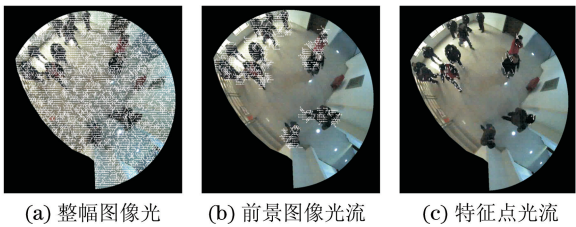


图 3 对不同区域计算的光流矢量对比图  
Fig.3 Comparison Images for Optical Flow in Different Regions

典型特征点提取方法有 Harris 角点、尺度不变特征变换(scale-invariant feature transform, SIFT)特征点和边缘点等。Harris 角点和 SIFT 特征点能有效描述目标的特征,但计算量较大<sup>[10,12]</sup>。本文采用 Canny 算子来检测人群目标的边缘,结果如图 4(c)所示。图 4 中的边缘像素点比较多,直接对这些像素点进行光流矢量的计算,计算量仍然较大;而通过增大边缘提取时的阈值来减少边缘像素点的个数,则可能会导致重要边缘信息的丢失。如图 5 所示,通过采样

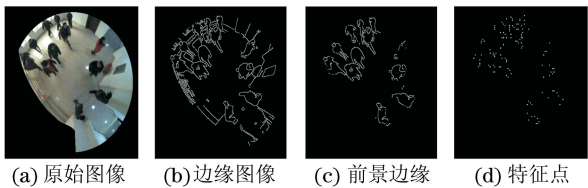


图 4 特征点提取结果图  
Fig.4 Images for Feature Point Extraction

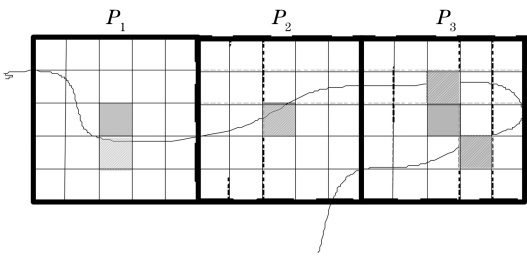


图 5 基于边缘密度的非均匀采样示意图  
Fig.5 Non-uniform Sampling Based on Edge Density

来减少边缘点的个数。假设窗口滑动以窗口的中心点为基准点,粗实线为窗口处于三个不同的位置,分别以  $P_1$ 、 $P_2$ 、 $P_3$  来表示,细实线表示窗口中的网格,窗口中灰色像素点为窗口中心点,中间横跨 3 个窗口的曲线表示边缘像素点。如果使用  $5 \times 5$  的模板进行均匀采样,则图 5 所示的 3 个窗口只有 1 个像素点被采样到( $P_2$  窗口的中心点)。显然,多数边缘信息丢失。为了兼顾准确性和实时性,本文提出一种基于边缘密度采样的特征点提取方法。其基本思想是边缘像素点密集的地方采样多,边缘密集的地方采样少。当窗口滑动到某一个位置的时候,对此窗口中边缘像素点采样的方法如下。

首先,统计当前窗口中边缘像素点的个数。假设当前窗口中包含  $m$  个边缘像素点,则通过  $m$  和窗口尺寸  $n$  来判断此窗口中需要采集的边缘像素点的个数  $m'$ ,  $m'$  可以通过式(1)计算得到,其中“ $\lfloor \cdot \rfloor$ ”表示向下取整:

$$m' = \lfloor m/n \rfloor \quad (1)$$

其次,判断  $m'$  的值是否为 0。如果  $m' = 0$ ,则再判断窗口中心点是否为边缘像素点。如果是,则采集此边缘像素点作为当前窗口的采样点,  $m'$  加 1;如果不是,则当前窗口不采集任何边缘像素点,窗口滑动到下一个位置。如果  $m'$  不为 0,则计算此窗口中所有边缘像素点到窗口中心点之间的距离,记为  $d_1, d_2, \dots, d_m$ 。将  $d_1 \sim d_m$  按照从小到大的顺序进行排序,然后采集前  $m'$  个距离对应的边缘像素点作为当前窗口中的采样点。

使用基于像素密集度的非均匀采样方法,可以在边缘像素密集时采样较多,在边缘像素稀疏时采样少,既能够保持边缘信息不丢失,又能够减少边缘点的个数,有效去除像素间冗余。如图 5 所示,使用本文的方法进行采样时,当窗口在  $P_1$  和  $P_2$  位置时各能采集到 1 个样本点;当窗口在  $P_3$  位置时,能采集到 2 个样本点,采集的样本点用阴影线表示。当一个窗口中的样本点采集完毕后,窗口以  $n$  的步长滑动到下一个位置继续采集,直到搜索完整幅图像为止,可得到图像的特征点。

针对特征点,本文采用 Lucas & Kanade 光流法计算光流场<sup>[4]</sup>,就可以得到特征点的光流矢量。图 3(c)即为基于特征点的光流矢量示意图。通过使用特征点来计算光流,不仅提高了实时性,而且还有效去除了大部分的光流噪声。

### 1.3 运动矢量计算

在得到人群目标的特征点光流之后,通过每

一个特征点的光流计算人群的整体运动矢量,即运动方向和速度。假设  $P_k = \{(x_{1k}, y_{1k}), (x_{2k}, y_{2k}), \dots, (x_{nk}, y_{nk})\}$  为使用本文的特征点提取算法对第  $k$  帧图像提取的一组特征点集合,  $(x_{ik}, y_{ik})$  为第  $i$  个特征点在第  $k$  帧的坐标,  $n$  为特征点的个数。通过 Lucas & Kanade 光流法的计算,可以得到在第  $k+1$  帧图像上与之相对应的特征点集合  $P_{k+1} = \{(x_{1k+1}, y_{1k+1}), (x_{2k+1}, y_{2k+1}), \dots, (x_{nk+1}, y_{nk+1})\}$ ,  $(x_{ik+1}, y_{ik+1})$  为第  $i$  个特征点在第  $k+1$  帧的坐标。对于第  $i$  个特征点,其运动的位移  $S_i$  和方向  $\theta_i$  可以通过式(2)、(3)计算得到<sup>[4]</sup>:

$$S_i = \sqrt{(x_i^k - x_i^{k+1})^2 + (y_i^k - y_i^{k+1})^2} \quad (2)$$

$$\theta_i = \arctan \frac{y_i^{k+1} - y_i^k}{x_i^{k+1} - x_i^k} \quad (3)$$

式中,位移  $S_i$  的单位为像素;方向  $\theta_i$  的单位为弧度。在得到第  $i$  个特征点的运动位移以后,结合视频的帧率  $f$  (单位为帧/s),通过式(4)就可以计算出第  $i$  个特征点的运动速度  $V_i$ :

$$V_i = S_i f \quad (4)$$

式中,运动速度  $V_i$  的单位为像素/s。

在获得每个特征点的运动方向和速度之后,需要获取人群的整体运动方向和速度。鱼眼图像中,同等数量的人群,离图像中心远的区域人群目标较小,其特征点也较少;离图像中心近的区域人群目标较大,其特征点也较多。如果统计时把远近不一的特征点按相同的权值计算,势必会导致较大的误差<sup>[13]</sup>。因此,在计算人群整体运动方向和速度的时候,需要对鱼眼图像的不同区域进行加权。文献[14]针对鱼眼相机,建立了一种透视加权模型,并进行人群密度和人数的估计,有效解决了静态人群监控的问题。本文在此模型的基础上,提出人群的运动估计方法,旨在解决动态人群监控的问题。依据文献[14],令图像中心点  $P_G$  的权值为  $w_G = 1$ ,则图像中离  $P_G$  点的距离为  $r$  的像素点的权值  $w_r$  为:

$$w_r = \left| \sec \left( \frac{r\epsilon}{f} \right) \right| \quad (5)$$

式中,  $\epsilon$  为成像靶面尺寸;  $f$  为焦距;  $r$  为当前点到图像中心点到像素距离;  $\sec$  为正割函数。依据式(5)可以计算出鱼眼图像中每一个像素点的权值。

为了获取人群运动方向和速度,本文对运动矢量进行加权统计。将鱼眼图像中人群的运动方向分为 4 个方向,依次为  $[0, 90^\circ)$ 、 $[90^\circ, 180^\circ)$ 、 $[180^\circ, 270^\circ)$ 、 $[270^\circ, 360^\circ)$ ,对应区间编号为 1、2、



3、4。考虑到鱼眼图像的透视结构,每一帧图像的统计运动方向直方图  $H_\theta$  和每个方向的瞬时速度  $V_\theta$  的数学表达式可以通过式(6)~(7)来计算:

$$H_\theta(t) = \sum_{i=1}^n (\omega_r(x_i, y_i) \delta(t - \lfloor \frac{2\theta_i}{\pi} \rfloor)) \quad (6)$$
$$V_\theta(t) = \frac{1}{H_n(t)} \sum_{i=1}^n (V_i \omega_r(x_i, y_i) \delta(t - \lfloor \frac{2\theta_i}{\pi} \rfloor)) \quad (7)$$

式中,  $\omega_r(x_i, y_i)$  为鱼眼图像在点  $(x_i, y_i)$  的透视

加权值,计算方法同式(5);  $t$  的取值为 0、1、2、3,依次代表 4 个方向区间;  $\delta$  表示冲激函数;  $H_n(t)$  也为运动方向直方图,但是不带加权,可以由式(8)计算得到:

$$H_n(t) = \sum_{i=1}^n \delta(t - \lfloor \frac{2\theta_i}{\pi} \rfloor) \quad (8)$$

统计结果的雷达图如图 6 所示。图 6 中第 1 行为视频 2 某帧结果,第 2 行为视频 1 平均结果,第 3 行为视频 2 平均结果。

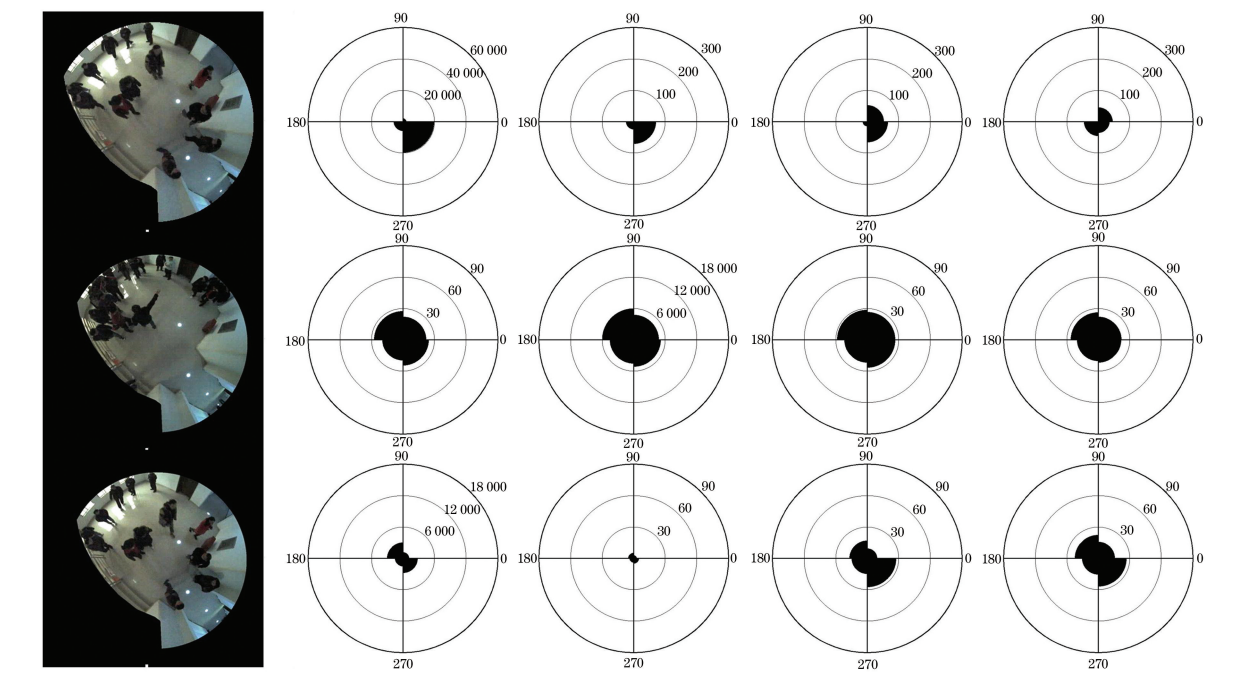


图 6 人群运动估计结果  
Fig.6 Images of Crowd Motion Estimation

2 实验结果及分析

本文实验视频拍摄所用鱼眼相机型号为 GeoVision GV-FE110。相机离地面的高度为 3.03 m,主光轴与地面的角度为 90°。视频帧率为 15 fps,尺寸为 1 280×1 024。本文实验的视频共两段,测试视频 1 帧数为 877,其中人数比较密集,人群围绕大厅做绕环运动。测试视频 2 的帧数为 749,其中人群较为稀疏,且人群在大厅中的左上角和右下角做往返运动。除了求实时的人群运动方向和速度外,还分别对两段测试视频的人群运动方向和速度求均值,如式(9)、式(10)所示,单位分别为像素和像素/s:

$$\overline{H}_\theta(t) = \frac{1}{N_f} \sum_{i=1}^{N_f} H_\theta(i, t) \quad (9)$$

$$\overline{V}_\theta(t) = \frac{1}{N_f} \sum_{i=1}^{N_f} V_\theta(i, t) \quad (10)$$

式中,  $\overline{H}_\theta(t)$  和  $\overline{V}_\theta(t)$  分别为人群的运动方向均值和速度均值;  $t$  的取值为 0、1、2、3,依次代表 4 个方向区间;  $H_\theta(i, t)$  为第  $i$  帧图像、第  $t$  个方向区间的运动方向直方图的幅值;  $V_\theta(i, t)$  为第  $i$  帧图像、第  $t$  个方向区间的瞬时运动速度;  $N_f$  为视频的总帧数。

将运动历史图像(motion history image, MHI)方法<sup>[15]</sup>与透视加权模型结合进行人群运动估计,并与本文的方法进行对比,结果如图 6 和表 1 所示。对比的结论如下。

1) 采用透视加权模型,利用鱼眼图像透视信息,准确地对远近各区域的人群进行实时的运动方向和速度的估计,达到了实时分析人群整体运动的目的。图 6 中,两种方法得到的人群运动都是以向右下角运动为主,而本文的方法效果更明显;运动速度上,本文方法受到干扰明显要少。

2) 本文方法能有效地估计一段时间内人群整

体平均运动方向和速度。图 6 和表 1 显示,视频 1 中人群的平均运动方向和速度在 4 个方向上接近。视频 1 中的人群在大厅中做绕圈运动;视频 2 中人群的平均运动方向主要以左上和右下方为主,平均速度在这两个方向较快,与实际情况相符。

3) 本文提出的方法不仅能准确描述人群的特征,保证结果的准确性,还能有效地降低运算量,提高系统的实时性。表 1 显示本文方法平均每帧运行时间约 20 ms,比基于 MHI 的方法要快。

表 1 人群运动速度估计和运行时间结果

Tab.1 Results of Crowd Speed Estimation and the Running Time

视频	方向区间	OL	MHI
1	1	22.35	20.46
	2	31.49	29.96
	3	23.55	22.78
	4	29.00	27.86
	平均每帧时间/s	0.021	0.094
2	1	20.31	19.98
	2	40.41	38.64
	3	18.89	17.44
	4	42.13	40.45
	平均每帧时间/s	0.019	0.088

### 3 结 语

本文提出了一种基于特征点光流的人群运动估计方法。首先对原始图像进行 ROI 划分和人群目标分割;其次,为了保证准确性和实时性,提出一种基于边缘密度采样的人群特征点提取方法,并计算特征点的光流场;最后,采用鱼眼相机的透视加权模型,计算人群运动方向和速度的加权统计直方图,获取全局运动方向和速度。实验结果表明,本文方法能有效、实时地估计人群的运动参数。

本文获取的运动速度并不是人群真正的运动速度,而只是人群在图像上的运动速度;由于鱼眼相机的短焦特征,本文方法只有在室内进行人群运动分析才能取得较好的效果。未来工作将集中在如何获取实际的人群运动速度以及人群行为分析方面。

参 考 文 献

[1] Davies A C, Yin J H, Velastin S A. Crowd Monitoring Using Image Processing [J]. *Electronics & Communication Engineering Journal*, 1995, 7 (1): 37-47

[2] Xu J B, Po L M, Cheung C K. Adaptive Motion

Tracking Block Matching Algorithm for Video Coding [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1999, 9 (7): 1 025-1 029

[3] Lee T S. Image Representation Using 2D Gabor Wavelets [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1996, 18 (10): 959-971

[4] Lucas B D, Kanade T. An Iterative Image Registration Technique with an Application to Stereo Vision [C]. *International Joint Conference on Artificial Intelligence*, San Francisco, CA, USA, 1981

[5] Ihaddadene N, Djeraba C. Real-time Crowd Motion Analysis [C]. *International Conference on Pattern Recognition*, Tampa, Florida, USA, 2008

[6] Cao T, Wu X, Guo J, et al. Abnormal Crowd Motion Analysis [C]. *IEEE International Conference on Robotics and Biomimetics*, Piscataway, NJ, USA, 2009

[7] Posada L F, Narayanan K K, Hoffmann F, et al. Floor Segmentation of Omni-Directional Images for Mobile Robot Visual Navigation [C]. *IEEE/RSJ Conference on Intelligent Robots and Systems*, Taipei, China, 2010

[8] Houben S, Komar M, Hohm A, et al. On-vehicle Video-based Parking Lot Recognition with Fisheye Optics [C]. *International IEEE Conference on Intelligent Transportation Systems*, New York, USA, 2013

[9] Liu K, Zhou F Q, Zhang G J. Radius Constraint Least Square Circle Fitting Method and Error Analysis [J]. *Journal of Optoelectronics & Laser*, 2006, 17(5): 604-607

[10] Kovacs A, Sziranyi T. Improved Harris Feature Point Set for Orientation-sensitive Urban-Area Detection in Aerial Images [J]. *Geoscience and Remote Sensing Letters*, 2013, 10 (4): 796-800

[11] Hu Xuemin, Zheng Hong, Guo Lin. Intelligent Monitoring for Dense Crowd Using a Fisheye Camera [J]. *Geomatics and Information Science of Wuhan University*, 2013, 39 (10): 1 236-1 240 (胡学敏, 郑宏, 郭琳.利用鱼眼相机对密集人群进行智能监控 [J].*武汉大学学报·信息科学版*, 2013, 39 (10): 1 236-1 240)

[12] Zhang Zhengpeng, Jiang Wanshou, Zhang Jing. A Gross Error Detection Method of Vehicle-borne Cubic Panoramic Image Sequence [J]. *Geomatics and Information Science of Wuhan University*, 2014, 39 (10): 1 208-1 212 (张正鹏,江万寿,张靖.车载立方体全景影像匹配点的粗差检测方法 [J].*武汉大学学报·信息科学版*, 2014, 39 (10): 1 208-

1 212)

[13] Chang Qinglong, Xia Hongshan. Pedestrians Gathering Detection Based on Normalized Foreground and Two-dimension Joint Entropy [J]. *Geomatics and Information Science of Wuhan University*, 2013, 38 (9): 1 126-1 130 (常庆龙, 夏洪山. 利用归一化前景和二维联合熵的人群聚集检测方法 [J]. 武汉大学学报·信息科学版, 2013, 38 (9): 1 126-1 130)

[14] Hu Xuemin, Zheng Hong, Zhang Qing. Crowd Monitoring for Subway Platform Based on Weighted Area from Perspective [J]. *Geomatics and Information Science of Wuhan University*, 2012, 37 (3): 343-347 (胡学敏, 郑宏, 张清. 利用加权面积透视变化对地铁站台进行人群监控 [J]. 武汉大学学报·信息科学版, 2012, 37 (3): 343-347)

[15] Gupta R, Jain A, Rana S. A Novel Method to Represent Repetitive and Overwriting Activities in Motion History Images [C]. *International Conference on Communications and Signal Processing*, Melmaruvathur, India, 2013

Crowd Motion Estimation Using a Fisheye Camera

HU Xuemin<sup>1</sup> ZHENG Hong<sup>2</sup> GUO Lin<sup>1</sup> XIONG Raorao<sup>1</sup>

1 School of Computer Science and Information Engineering, Hubei University, Wuhan 430062, China

2 School of Electronic Information, Wuhan University, Wuhan 430072, China

**Abstract:** Crowd motion estimation is an important part of crowd action analysis. Crowd motion Analysis in special places is a necessary action for maintaining the safety and social stability in public place and there is a research difficulty in the field of intelligent video monitoring. Existing approaches for crowd motion estimation based on traditional cameras have the limitation of small field-of-view and more blind spots. This paper proposes a crowd motion estimation approach based on the feature point optical flow employing the advantages of large field-of-view and no blind spot of fisheye cameras. Firstly, the original images are preprocessed using the method of background difference based on Gaussian Mixture Model with area feedback, and the region of interest (ROI) is obtained by circle fitting. Secondly, a feature point extraction method based on non-uniform sampling of edge density is presented to describe the moving crowd for improving the real-time performance as the same time as ensuring the accuracy of describing the crowd. And then the optical flow field is calculated using the method by Lucas & Kanade. Finally, a perspective weight model of the fisheye camera is developed to weighting the compute the motion vector and the motion direction and speed of the crowd in fisheye camera images in order to solve the issues of the size differences of the crowd in long and short distances and the distortion of fisheye images in this paper. The experimental results show that the proposed approach is effective and feasible for estimating the motion speed and orientation of the crowd in dense crowd. In addition, the proposed method provides an important research basis for crowd behavior analysis.

**Key words:** fisheye camera; optical flow; feature point; perspective weight model; crowd motion estimation; orientation histogram; speed histogram

**First author:** HU Xuemin, PhD, specializes in image processing and computer vision. E-mail: huxuemin2003@163.com

**Corresponding author:** GUO Lin, PhD. E-mail: 10837330@qq.com

**Foundation support:** Undergraduate Innovation Training Foundation of Hubei Province, No. 201510512041.