

人群计数问题研究综述^{*}

——从基于图像的方法到基于无线网络的方法

潘轶功¹⁺

¹(中国科学技术大学 计算机科学与技术学院, 安徽省 合肥市 231000)

+通讯作者: 潘轶功(SA16011050), E-mail: pikaqiu@mail.ustc.edu.cn

摘 要: 人群计数, 指的是计算或估计一个地区内的人数, 有人员密集场所的人流控制、旅游景区人流引导等诸多应用。传统的方案是基于视频(或图像)的, 后来随着手机的普及, 基于无线设备的有源感知人群计数方法逐渐兴起, 克服了视频方法具有视觉盲区等固有的缺点。现在, 随着生活中的无线网络的大规模普及, 基于无线网络的无源感知方法开始出现, 这种方法具有不需要人群携带设备, 成本低廉的优点。**据我了解, 这是第一篇按照历史发展轨迹, 从图像方法、有缘感知、无缘感知三个方面全面介绍人群计数问题的综述。**

关键词: 人群计数 无线网络 图像识别 有源感知 无源感知

引言

人群计数, 指的是计算或估计一个地区内的人数, 在许多应用中起着重要的作用, 比如在旅游景区内的人流导引问题和人流密集场所的控制问题。人群计数问题的解决方案应该以更广的扩展性和更低限度地干扰人群(比如不需要设备的方案)为目标。

基于视频(或图像)的人群计数方案[1][2][3][4][16]已经被广泛应用在许多公共场所, 基于图像的解决方案的优点是不需要设备。但是, 这个方法具有一些固有的缺点。首先, 摄像机只能固定在一个地方, 导致出现许多监控盲区。二是视频、相机质量受环境光影响很大, 在雾霾天和夜晚, 图像质量会很差。第三, 人流密集区域人与人在图像中重叠严重, 将导致计数结果严重偏差。此外, 使用相机拍摄会涉及隐私保护问题。

基于设备的解决方案[5][6][7][8][9][10]通常利用无线电设备定位物体, 如 RFID 标签, 手机, 传感器节点等。这些基于设备的方法要求人们携带某些可被监视的设备, 这显然限制了使用范围。为每一个在公共区域的大众, 分配设备是不切实际的, 成本也很高, 可行性很差。

近几年出现了一些无装置的基于无线网络的方法[11][12][13][14]。大多数方法采用 RSS 指纹进行定位, 这是很容易获得的现成的无线设备。但是, 这种方法有个缺点是需要实地考察, 耗费, 费时费力, 且容易受环境影响。为了避免实地考察, 可通过 RSS 信号强弱传播距离以定位对象。不幸的是, 信号强度随距离的衰减模型在复杂的室内环境下表现不佳。最近的主要趋势之一是探索细粒度的基于物理层信息的在多路径环境中定位和人群计数问题。另外, [15]中提出一种基于来自基于 OFDM 系统的信道状态信息(CSI)的测量的计数方法。

本文有以下贡献点:

- 1、据我了解, 这是按照历史发展轨迹, 从图像方法、有缘感知、无缘感知三个方面归纳整理、介绍人群计数问题的第一篇综述。
- 2、本文对图像方法、有缘感知、无缘感知的优缺点进行了对比。

^{*}作者简介: 潘轶功(1993.12.19—), 性别男, 浙江衢州人, 硕士研究生在读, 主要研究领域为数据挖掘, 社会计算, 程序语言理论, 量子计算。本论文使用《软件学报》的模板。

基于图像的方法

基于视频（或图像）的人群计数方案已经被广泛应用在许多公共场所。

在早期的研究[1]中，人群中人的数目用前景像素的函数来表示，而前景图像通过去除使用参考图像的背景而获得。但是如果背景改变它可能失败。

[2]提出了一种更成熟的系统，通过使用单个图像，能估计在一个复杂的拥挤场景中的人数。该系统的主要步骤有识别头状轮廓和人群规模的估计。首先，提取头状的特征区域轮廓用的是 Haar 小波变换(Haar wavelet transform)(HWT)技术，然后使用支持向量机(SVM)将这些特征区域分类为头部的轮廓。其次，使用计算机视觉的透视变换技术以更准确地估计人群大小。这篇文章中还构造了一个虚拟的模型世界用来测试这个系统。

[3]提出了一种基于一组扩展的轨迹特征的估计移动相同物体的方法。一般的识别密集对象运动的方法为保证充分的通用性，密集对象的运动分析，需要识别和分割每个移动的实体，这些任务的难度很大。而当对象被限制为相同种类且为刚性或半刚性时，可以通过聚类跟踪的特征点区分密集物体。这篇文章提出的方法基于高度并行化的版本 KLT 跟踪器，以便将视频处理成一组特征轨迹。虽然这样的轨迹集可用于运动分析的，但它们的长度不等性和运动的不完整性给后期的处理带来了困难。为了解决这个问题这篇文章提出了一个简单的调节空间和时间轨迹的手段。最终给出了估计密集人群中的移动物体的数量关于时间的函数。类似地，[16]提出了一种基于启发式和基于形状估计移动群体人口的方法。这两种方法假设对象是移动的。

[4]对前面的工作做了改进，提出了一种新的方法来解决人群计数的问题。这种方法移动和静止的人群都适用，该方法组合了 MID (Mosaic Image Difference)的前景分割算法 HOG (Histograms of Oriented Gradients) 头-肩检测算法来准确地估计观察区域的人数。该论文进行了许多实验，结果都表明该方法的有效性。

基于设备的方法

基于设备的方法通常使用无线电设备定位物体，如 RFID 标签，手机，传感器节点等。

[5]提出了一个名叫 Cricket 的位置计算系统，这个系统是基于耦合的射频信号和超声信号来计算位置的。该系统具有保护用户隐私，无需集中管理，适应网络异质性和低成本的特点。具体地说，该系统不依赖任何集中式管理，可用于任何类型的网络连接，并且制作每个 Cricket 设备的成本低于 10 美元。

[6]提出了名叫 LANDMARC 的用于定位室内物体的系统，该系统使用无线电频率识别技术 (RFID) 定位室内的物体。主要优点是它在整体上提高了定位的精度。实验结果表明，RFID 技术用于室内物体位置的定位是可行的和低成本，虽然 RFID 技术不是为室内定位而设计的。

智能手机数量的迅猛发展提供了识别逻辑位置的新机会。手机的各种传感器、相机、麦克风可以监测环境声音，加速度计等传感器可以用于推断用户的运动，通常由地方的性质决定。通过结合这些光学，声学 and 运动属性，可以实现逻辑位置的识别。相邻商店中的用户可以在逻辑上分离，即使他们的位置非常接近。[7]提出了名叫 SurroundSense 的系统，完成了上述工作。51 家不同的商店的实验结果表明，SurroundSense 可以实现的正确率为 87%。

[8]提出了一种使用手机扫描环境中的蓝牙设备估计人群密度的方法。该论文扩展了更高级的功能，利用了之间几个靠近的设备，使用的特征不直接取决于环境中的绝对设备数量。利用慕尼黑 Oktoberfest 三天节日产生的数据集完成的实验显示该方法的精度超过 80%，比简单计数的方法提高 30%。

在[9]中，论文作者提出了一个基于音频音调的进行计数解决方案，而麦克风和扬声器在大多数手机上都有，因此实际应用价值广泛。实验表明该方法能够实现高达 90%的精度，而且比使用 WiFi 的方法减少消耗 81%的能量。

[10]中建立了集成众多模块的端到端的室内定位系统，第一次实现了米级室内定位系统，具有价格低廉，手机位置独立，用户自适应，易于部署的特点。我们对用户进行的广泛的实验结果表明，精度能达到 1.5m。

基于无线网络的方法

无源定位的概念（DfP）在[11]中首先被提出来。DfP 系统是能够检测，跟踪和识别不携带任何电子设备的也不主动报告定位的物体。系统通过监控和处理来定位物体。DfP 系统的应用包括检测、跟踪和保护公共财产，如管道，铁路等。这篇文章中描述了实现 DfP 系统的架构和面临的挑战。我们通过描述显示系统的可行性，提出了两种入侵检测技术用于跟踪单个物体。实验结果表明该系统可以达到非常高的查准率和查全率。这篇文章认为需要不同的研究方向的人来应对实现 DfP 系统的挑战。

在[12]中，作者提出一种被动估计人群数目的系统，这种系统能够轻松地安装在现有的基础架构上。准确地说，作者是使用 RSSI（接收信号强度指示符）和 LQI（链路质量指标）这两个指标来估计的，而这两个指标可以通过无线网络轻松获得。

[13]中，使用了无线传感器网络（WSN）弥补相机的不足。作者的方法是使用迭代过程使其在每个时隙中包含两个相位。在检测步骤中，将人群密度分成不同的水平根据使用 K-means 算法的 WSNs 获得的 RSSI 数据。在校准步骤中，消除噪声和其他基于空间 - 时间相关性的偏差估计。此外，该论文中提出的算法通过了真实世界实验使用的 16 传感器节点和大规模模拟。结果表明：该算法具有准确，高效和一致的性能。

基于射频的无设备无源定位，这种技术通过观察人们如何扰乱室内空间中的无线电波的从而导出其位置。但众所周知的多路径效应使这个问题非常具有挑战性，因为在一个复杂的环境，拥有足够的知识以准确地模拟对象对周围无线电的影响是不切实际的。在[14]中，开发了使用基于指纹的概率分类的方法，和基于判别分析的方法。

最后[15]中提出了FCC方法，一个无源的基于信道状态信息（CSI）的人群计数的方法。该文从理论上讨论了人的位置变化与无线信道状态的变化关系。在设计方法的过程中，作者的一个主要挑战是寻找一种稳定的单调函数来表征人群数量与无线信道状态的关系。为此，文中提出了一个指标，CSI矩阵的非零元素的百分比。作者使用现成的FCC，使用IEEE 802.11n设备评估其性能在现实世界场景中做了广泛实验。结果表明FCC优于现有技术方法，具有更好的精度，可扩展性和可靠性。

结论

图像方法、有源感知方法、无源感知方法各有优缺点，实际使用时应看具体情况。下面通过表1对比这三种方法的方式结束这篇论文。

方法	诞生年代	优点	缺点
图像方法	上个世纪	被测人群不需要设备	存在视觉盲区 受环境影响大 会涉及隐私保护问题
基于设备（有源感知）	2000年	准确、精度高	需要设备可行性差 成本高
基于无线网络（无源感知）	2007年	成本低 不需要设备 通过使用新的方法，精度在逐渐提高	某些方法精度不高

References:

- [1] Velastin S A, Yin J H, Davies A C, et al. Automated measurement of crowd density and motion using image processing[C]//SEVENTH INTERNATIONAL CONFERENCE ON ROAD TRAFFIC MONITORING AND CONTROL, 26-28 APRIL 1994. PROCEEDINGS.(IEE CONFERENCE PUBLICATION 391). 1994.
- [2] Lin S F, Chen J Y, Chao H X. Estimation of number of people in crowded scenes using perspective transformation[J]. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 2001, 31(6): 645-654.
- [3] Rabaud V, Belongie S. Counting crowded moving objects[C]//2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06). IEEE, 2006, 1: 705-711.
- [4] Li M, Zhang Z, Huang K, et al. Estimating the number of people in crowded scenes by mid based foreground segmentation and head-shoulder detection[C]//Pattern Recognition, 2008. ICPR 2008. 19th International Conference on. IEEE, 2008: 1-4.
- [5] Priyantha N, Chakraborty A, Balakrishnan H. The Cricket location-support system[C]// International Conference on Mobile Computing and NETWORKING (MobiCom'01). ACM, 2001:32-43.
- [6] Ni L M, Liu Y, Lau Y C, et al. LANDMARC: indoor location sensing using active RFID[J]. Wireless networks, 2004, 10(6): 701-710.
- [7] Azizyan M, Constandache I, Roy Choudhury R. SurroundSense: mobile phone localization via ambience fingerprinting[C]//Proceedings of the 15th annual international conference on Mobile computing and networking (MobiCom'09). ACM, 2009: 261-272.
- [8] Weppner J, Lukowicz P. Collaborative crowd density estimation with mobile phones[J]. Proc. of ACM PhoneSense, 2011.
- [9] Kannan P G, Venkatagiri S P, Chan M C, et al. Low cost crowd counting using audio tones[C]//Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems. ACM, 2012: 155-168.
- [10] Li F, Zhao C, Ding G, et al. A reliable and accurate indoor localization method using phone inertial sensors[C]//Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp'12). ACM, 2012: 421-430.
- [11] Youssef M, Mah M, Agrawala A. Challenges: device-free passive localization for wireless environments[C]//Proceedings of the 13th annual ACM international conference on Mobile computing and networking (MobiCom'07). ACM, 2007: 222-229.
- [12] Nakatsuka M, Iwatani H, Katto J. A study on passive crowd density estimation using wireless sensors[C]//The 4th Intl. Conf. on Mobile Computing and Ubiquitous Networking (ICMU 2008). 2008.
- [13] Yuan Y, Qiu C, Xi W, et al. Crowd density estimation using wireless sensor networks[C]//Mobile Ad-hoc and Sensor Networks (MSN), 2011 Seventh International Conference on. IEEE, 2011: 138-145.
- [14] Xu C, Firner B, Zhang Y, et al. Improving RF-based device-free passive localization in cluttered indoor environments through probabilistic classification methods[C]//Proceedings of the 11th international conference on Information Processing in Sensor Networks. ACM, 2012: 209-220.
- [15] Xi W, Zhao J, Li X Y, et al. Electronic frog eye: Counting crowd using wifi[C]//IEEE INFOCOM 2014-IEEE Conference on Computer Communications. IEEE, 2014: 361-369.
- [16] Kilambi P, Ribnick E, Joshi A J, et al. Estimating pedestrian counts in groups[J]. Computer Vision and Image Understanding, 2008, 110(1): 43-59.