摘　要

排队是人类生活中十分常见的现象，当公共服务不能同时满足多个人的需求时，人们就要进行排队等待。人数估计作为排队现象中最基本、最关键的问题，具有重要的实际应用价值。例如，根据本文提出的方法实现的出租车站排队人数估计系统，可以实时查询车站等待人数，从而合理有效的调配出租车。

本文采用计算机视觉领域的多种技术，从已有的监控视频或实时视频流中估计排队区域内的人数。解决这个问题的难点在于实际环境下复杂的拍摄视角、目标遮挡以及复杂多变的人体姿态对人数估计的影响。国内外针对人数估计的研究大多基于纹理特征、局部特征或动态特征来估计人数，应用于广场、街道等比较大的场景，视角比较广，人在区域内所占范围较小。而本文关注的排队场景（如车站、银行、超市等）视角较窄，遮挡较为严重，目前的相关研究成果并不多见。

本文针对排队场景中存在的难点，利用计算机视觉领域的检测、跟踪、背景减除、深度估计等技术，设计并实现了若干个排队人数估计方法。本文的主要贡献如下：

(1) 提出一种基于人脸检测的人数估计方法。对于从侧前方拍摄的队列，采用基于人脸检测的方法来估计人数。本文使用的人脸检测器是基于 Haar 特征，利用瀑布型分类器来进行检测，并利用基于 LAB 特征的分类器作为预处理加速检测过程。采用队列分段的方法有效去除遮挡，提高对多变人体姿态的鲁棒性；尝试掩模处理来提高估计方法的速度与精度。

(2) 提出一种基于人体检测的人数估计方法。对于从正前面或背面拍摄的队列，使用基于人体检测的方法来估计人数。本文利用的是 HOG 特征和级联分类器来训练人体检测器，利用检测结果的颜色信息和时空连续性去除不稳定误检。

(3) 提出一种基于排队区域的人数估计方法。当排队区域比较固定时，可以使用基于排队区域的方法估计人数。把排队区域划分成检测区域、排队区域和非排队区域。利用背景减除和跟踪手段对目标进行跟踪，当其完全进入或离开排队区域后，更新排队人数。

(4) 提出一种基于深度信息的人数估计方法。针对顶部部署摄像头，排队队列比较混乱的情景，本文提出了基于深度信息的人数估计方法。通过 Kinect 结构光得到深度信息后，利用高度差做阈值处理，对前景中人体轮廓的连通区域进行统计，从而进行人数估计。

关键词： 排队人数估计；目标检测；视频监控；人脸检测；人体检测；深度估计；Kinect

People Counting in Real-World Queue Lines

Zihan Zhao (Computer Application Technology)
Directed By Xilin Chen, Professor

Queuing is very common in our life, when the public resources cannot meet the needs of all the people at the same time, people will queue up. As a basic and crucial issue in real-world queue line scenario, the problem of people counting is of important application values. For example, with the people counting system developed in this thesis, we can know the real-time number of people in an taxi station, and then dispatch the taxies more efficiently.

In this paper, a variety of computer vision techniques are used to solve the problem of people counting, where the inputs are either existing or live video streams. The difficulties of the problem include: various angles of shooting, heavy occlusion, and complex poses of people. Recent research on people counting are mostly based on texture features, local features or dynamic features. Most systems are applied in relatively large scenes like squares, streets and so on. The queuing scenario of small region and heavy occlusion (such as railway stations, banks, supermarkets, etc.), which is the focus on this thesis, has not been studied well.

To address the difficulties existing in people counting for real-world queuing scenes, we integrate computer vision techniques of detection, tracking, background subtraction and depth estimation to design and implement a people counting system. The main contributions of this paper are as follows:

(1) This paper presents a people counting method based on face detection. For the videos with front view peoples in the queue, we use face detection method to estimate the number of people. The face detector we used is based on Haar features and using cascade classifiers for rapid detection. The method uses queue segmentation method to effectively alleviate the influence of occlusions.

(2) This paper presents a people counting method based on body detection. For the videos with side view or back view people in the queue, we try body detector to estimate the number of people. In this paper, we use HOG (Histogram of Oriented Gradient) features and cascade classifiers to train the human body detector. The method also uses people identity and space-time continuity to remove the unstable false positives.

(3) This paper presents a region-based people counting method when the queuing area is relatively fixed. The method divides the area into the observing area, queuing area and non-queuing area. We try to use background subtraction and tracking to track the target, when

it completes entering or leaving the area, we update the estimated number of people.

(4) This paper presents a people counting method based on depth information for the videos with top view people in the queue. In this paper, we use Kinect to get the depth information of the scenes. After preprocessing and thresholding, the foreground connected regions are estimated as the number of queuing people.

Key Words: people counting in real-world queue lines; target detection; video surveillance; face detection; human detection; depth estimation; Kinect