摘要

多目标追踪是计算机视觉领域重要的研究课题之一。前人的工作大多针对静止的监控摄像机，然而，近年来，智能交通、自动驾驶等技术的兴起使得人们迫切地需要解决复杂交通场景中第一视角下的多目标追踪问题。

由于摄像机处于复杂的运动中，同时较低的视角导致严重的遮挡，以及复杂的环境变化等因素，使得移动摄像机下的多目标追踪问题比传统的静止摄像机中的追踪问题要困难得多。同时，第一人称视角下的多目标追踪又有很多它所独有的特点和规律：物体近大远小的透视规律，摄像机视角与地面平行且高度一般不发生改变，被追踪的物体往往都是处于地面上等。

本文沿用了受到广泛认可的，通过数据关联对物体检测结果进行归类的方法来完成多目标追踪。从数学上阐述了在考虑摄像机运动信息的情况下如何将多目标追踪问题转化为最小费用流问题。

针对复杂交通场景中第一视角下的多目标追踪问题和传统的静止监控摄像机下多目标追踪问题的不同，本文在目前对静止摄像头下效果比较好的多目标追踪方法的基础上做出了一些改进。这些改进包括：在度量物体位置和尺度的变化规律时考虑摄像机的运动，通过计算逆透视变换后物体的实际大小估计物体检测结果的可靠性，学习新的透视视角下物体轨迹起点和终点先验模型的。此外，本文还评估了外观相似性度量对结果的影响。

通过在KITTI数据集上进行对比实验，证明了这些改进的有效性和重要性。

关键词：多目标追踪，最小费用流，摄像机运动

FIRST-PERSON MULTIPLE OBJECT TRACKING IN COMPLEX TRAFFIC SCENES

Zhao Zhang (Computer Applied Technology)
Directed by Prof. Yizhou Wang

ABSTRACT

Multiple object tracking is one of the most important problems in computer vision. Many existing tracking approaches reply on the assumption that camera is static, which is surveillance. However, tracking multiple objects from the first-person viewpoint, e.g., cameras are installed on the vehicles which may move very fast, is also very important, especially for developing autonomous vehicles and intelligent traffic system.

This problem brings many new challenges compared to traditional multi-object tracking problem. Because the camera can move very fast and change direction abruptly, the targets may move in a different way. The objects may get severely occluded by other objects due to the low viewpoint. In complex traffic scenes, the illumination changes can be large, and the background is dynamic and clutter. All these new challenges make previous tracking approaches relying on static surveillance camera fail on this problem.

However, tracking multiple objects from the first-person viewpoint has its own pattern and specialty. The further target is from the camera, the smaller it presents in the video. The camera faces the front and its height keeps in the same level. The targets which are needed to be tracked usually stand on the ground.

First, I adopt the tracking-by-detection approach and demonstrate how to transfer the multiple object tracking problem into min-cost flow problem considering the motion of camera.

Next, I re-define the cost from three aspect as improvements, considering the differences between moving camera under first-person view and surveillance camera. By considering the camera motion, the similarity measure can handle the change of scale and position of objects when the camera moves in any pattern. The likelihood which indicates whether the detection candidate is true is evaluated by the object scale prior in real world after inversing the perspective transformation. The prior probability of the initialization and termination of a

trajectory is learned under first-person view, correlating to the position and scale of detection candidates in images. We also consider different methods to calculate appearance relationship to improve the tracking performance.

Experimental results on KITTI dataset demonstrate the advantages of our improvements.

KEYWORDS: Multiple object tracking, Min-cost flow, Camera motion