# Robust Moving Object Tracking and Trajectory Prediction for Visual Navigation in Dynamic Environments

Yi-Sheng Hsieh, Yu-Chi Su, and Liang-Gee Chen, *Fellow, IEE*E
DSP/IC Design Lab., Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan

***Abstract*-- We propose a robust and efficient moving object detection system with trajectory prediction mechanism for mobile visual applications in dynamic environments. The navigation system not only supports intelligent moving object detection and collision avoidance, but also provides automatic trajectory prediction mechanism to create a safe environment for the user. To take into account recognition challenges in practice, the system addresses problems of camera shake and image blurs caused by mobile camera and fast moving objects, respectively. Experimental results show our system supports 97.1% detection rate for fast moving objects and achieves high accuracy of trajectory prediction with error estimation less than 16.5 cm.**

## I. INTRODUCTION

Considering safety issues, obstacle detection and collision avoidance are critical issues for mobile vision applications, like robotic vision, intelligent vehicles, and wearable visual navigation. For a mobile agent, automatic trajectory analysis and prediction for coming moving objects, such as pedestrians or cars, are basic and important abilities for safety. However, tracking trajectories of moving objects for mobile agents is challenging due to the joint effect of global camera motion and various object motions. Several works have been presented to detect moving obstacles for mobile applications [1]-[5]. ESS et al. developed a mobile vision method for multi-person tracking [4]. Yuan et al. proposed a statistical approach for moving object motion analysis [5].

In this paper, we proposed a robust and efficient visual navigation system with intelligent moving object tracking and trajectory prediction mechanism. The system aims to provide the user semantic and useful information for collision avoidance in dynamic environments. In addition, we addresses serious problems of recognition instability caused by unwanted camera shake on a rough road and image blurs of fast moving objects. In addition, the characteristic of the proposed algorithm is suitable for hardware implementation to achieve real-time performance.

## II. PROPOSED METHOD

The overall state diagram of the system is shown in Fig 1. The input video sequence is processed through motion analysis and classification to separate different motions among various moving objects in a dynamic environment. The following is a weighting 2D multi-cue object tracking mechanism considering both extracted moving object motions and object appearance detection results. Finally, the depth of the target object is estimated to calculate its 3D position relative to the user. Further, the future trajectory of the object is predicted to provide useful information to the user.

### *A. Motion Analysis and Classification*

To distinguish various moving objects for a mobile agent in a dynamic environment, a statistical motion analysis approach is employed as the first step of the framework. For each frame, global camera motion needs to be estimated before extracting motions for individual moving objects. This is because the motion of a moving object captured by the camera mounted on a mobile agent is composed of global camera motion and individual object motions.

A feature-based technique is used to extract global camera motion from a frame. We take the advantage of the optical flow which matches corresponding feature points between two consecutive frames. In this paper, Lucas–Kanade method is adopted. It solves the basic optical flow equations for all the pixels in their neighborhoods by the least squares criterion and is less sensitive to noise. The major orientation of optical flow in the generated motion field can be considered as global camera motion since feature points in static background usually occupy most area of the whole picture. Next, object motions are separated from extracted optical flow by removing the effect of global camera motion or unwanted camera shake.

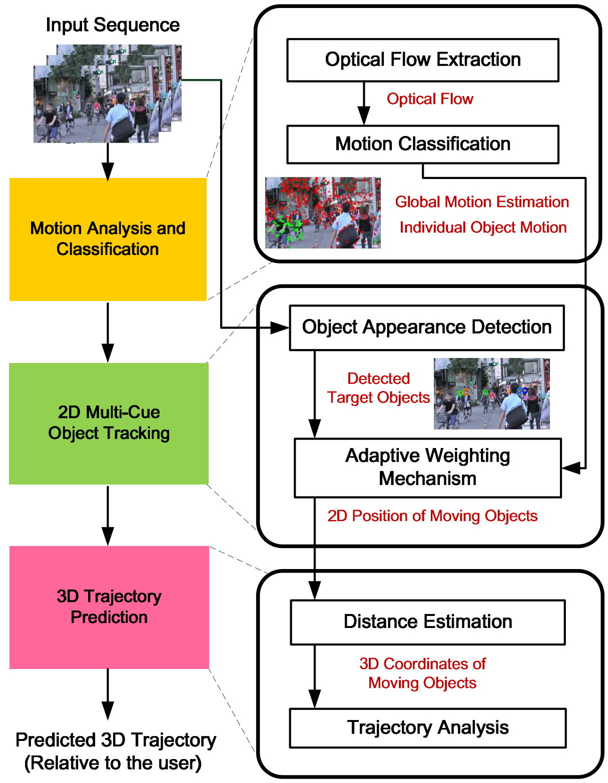


Fig. 1 System block diagram

To distinguish different moving objects, motion classification starts by statistically classifying object motions into several class levels according to corresponding positions, orientations and magnitudes. By maximizing inter-class differences among various individual motions, thresholds for motion clustering are automatically set. In this way, individual object motion for each object can be inferred.

### *B. 2D Multi-Cue Object Tracking*

A weighting 2D multi-cue tracking technique combined with object appearance detection and motion tracking is used to provide robust moving object detection. We design an adaptive weighting mechanism to dynamically adjust impacts of objects motions and appearances. The idea of weighting mechanism comes from the fact that object appearance detection reveals its weakness when detecting image blurs caused by fast motions of moving objects. In this situation, object motions extracted from the previous step are able to assist in tracking positions for target objects more accurately. On the other hand, even if motions of fast moving objects can be successfully extracted by the motion analysis step, it usually fails to distinguish motions of slow moving objects from plenty of global motions from background pixels, especially when suffering from severe camera shake. In this case, appearance detection has better performance to clearly recognize an object.

As to object appearance detection, various classifiers, such as pedestrian, bike, or vehicle detectors, can be adopted in the framework. In our work, we focus on detecting pedestrians and bike riders who move at either slow or high speed. Face detection is performed to capture the existence of a target object. This is because potential dangerous moving objects are pedestrians or bike riders who move toward the user. Thus, face detection can be an effective approach for object appearance detection in our scenarios.

The classified motions of moving objects and outputs of object appearance detection are both fed into the weighting mechanism to track exact positions of target objects. Weighting parameters are locally adjusted according to motion speed for different moving objects.

### *C. 3D Trajectory Prediction*

The depth information of a moving object is determined based on the prior knowledge of relationship between the face size of the object appearing in the image and the actual distance between the object and the camera. We assume that face sizes of various persons differ only in a negligible scale. With the depth information and 2D positions from object tracking results, 3D coordinates of moving objects can be calculated.

When all the procedures illustrated above are done, the next frame is processed by the system in a similar way. Every 10 frames, a set of data regarding 3D positions of moving objects during this period are collected. The data can be used to analyze the trajectory for each object and further predict its moving direction in the future. Trajectory prediction is also useful when occlusion happens, which usually misses tracking information. Most importantly, trajectory prediction in this work is employed to provide semantic information for users to let them understand moving directions of potential threats.

## III. EXPERIMENTAL RESULTS AND ANALYSIS

Two experiments are conducted to evaluate accuracy of the proposed system. All the sequences tested in experiments are captured from a 24fps CMOS front-mounted camera with 1920×1080 resolution.

Table 1 shows the improvement on detection rate involving the weighting 2D multi-cue object tracking mechanism. With the proposed adaptive weighting mechanism that considers both object motion and appearance, the recognition performance is increased from 90.3% to 97.1%. Fig. 2 depicts the enhancement. From Fig.2 (a), we can see that a bike rider is not detected by object appearance detection due to the motion blur resulted from its moving motion. Fortunately, motion analysis, as shown in Fig. 2 (b), detects the existence of this fast moving object. Combining these two detection methods, the system achieves higher accuracy.

(a) (b)

Fig. 2 Results of different methods. (a) Detection with object appearance approach. (b) Detection with motion analysis and classification.

TABLE I
ACCURACY COMPARISON

| Detection method | Accuracy (%) |
|---|---|
| Object appearance detection alone | 90.3 |
| 2D multi-cue object tracking | 97.1 |

The second experiment evaluates the trajectory accuracy of the proposed system. In this scenario a user wearing a head-mounted camera moves forward and at the same time a pedestrian walks toward him. To measure the correctness of the proposed trajectory mechanism, both the user and the pedestrian walk along their predefined paths with trajectories marked in advance. According to the predicted trajectory for the pedestrian, at each timestamp the average distance between the estimated location and the actual position where the pedestrian passes is less than 16.5 cm. The estimation error can be negligible when moving in outdoor environments. This experimental result demonstrates that our system provides effective information of trajectory prediction to the user.

TABLE II
TIME COMPLEXITY ANALYSIS

| Module | Proportion (%) |
|---|---|
| Motion analysis and classification | 26.6 |
| 2D multi-cue object tracking | 69.4 |
| 3D trajectory prediction | 0.005 |
| Others | 3.97 |

Table II summaries the workload ratio of each module by software implementation for the proposed system. This result shows motion analysis and 2D multi-cue object tracking dominate 96% of the whole process. The whole process of the system takes about 3 second to deal with a frame. Fortunately, most of the algorithms adopted in these two modules, such as Lucas–Kanade optical flow and Adaboost-based face recognition approach, have characteristics of parallelization. Thus, our system is suitable for hardware implementation. To achieve low power and real-time performance for portable devices, hardware acceleration for critical functions is effective.

## IV. CONCLUSION

We propose a robust and efficient moving object detection system with trajectory prediction mechanism for mobile visual applications. To overcome degrading recognition accuracy due to camera shake and image blurs of fast moving objects, we apply statistical motion analysis and weighting 2D multi-cue tracking approaches for robust detection. Our system supports 97.1% detection rate for fast moving objects. High accuracy of trajectory prediction with error estimation less than 16.5cm is also achieved. In addition, the hardware-efficient characteristic of the proposed algorithms can be massively parallelized as hardware circuit for low power consumption and real-time performance.

## REFERENCES


[1] Robert L. Williams II and Jianhua Wu, "Dynamic Obstacle Avoidance for an Omni-Directional Mobile Robot", *Journal of Robotics*, Volume, 2010

[2] Chaoxia Shi, Yanging Wang, Jingyu Yang, "A local obstacle avoidance method for mobile robots in partially known environment", *Journal Robotics and Autonomous Systems*, Volume 58 Issue 5, May, 2010

[3] Andreas ESS, Konrad Schindler, Bastian Leibe, Luc Van Gool "Object Detection and Tracking for Autonomous Navigation in Dynamic Environments", *I. J. Robotic Res.,* pp. 1707-1725, 2010

[4] Andreas ESS, Bastian Leibe, Konrad Schindler, and Luc Van Gool, "A Mobile Vision System for Robust Multi-Person Tracking", *2011 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 1-8, 2008.

[5] Chunrong Yuan, Isabell Schwab, Fabian Recktenwald and Hanspeter A. Mallot, "Detection of Moving Objects by Statistical Motion Analysis", *The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, October 18-22, 2010, Taipei, Taiwan