I. Background

Traffic intersections are well-known targets for monitoring because of their high number of reported accidents and collisions.[1] There are three parts to concern in traffic intersection analysis system: Pedestrian, Vehicle and Conflicts between them. Vision-based pedestrian detection methods, behavior analysis have previously been studied[2]. Recently, vision-based detection, tracking, and behavior analysis for vehicles have been reviewed[3]. In this project, we will consider pedestrian as well as vehicle at intersections and also predict the possibility of pedestrian and vehicle conflicts.

Fig. 1 shows the process of pedestrian and vehicle recognizing system, and several methods can be used to achieve each goal.
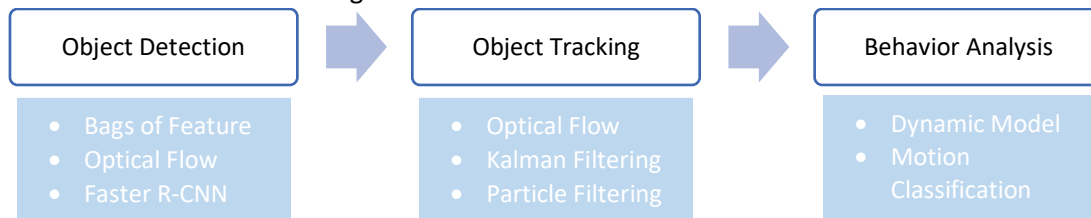
Object Detection	Object Tracking	Behavior Analysis
• Bags of Feature • Optical Flow • Faster R-CNN	• Optical Flow • Kalman Filtering • Particle Filtering	• Dynamic Model • Motion Classification

Fig. 1

(1) Object Detection

Several methods can be used to do pedestrian and vehicle detection, such as Support Vector Machines (SVM) in [4] for pedestrian and in [5] for vehicle. In [6], a convolutional neural network (CNN) classifier is used to automatically learn appropriate features and obtain improved detection performance. And in [7], the faster region-based convolutional neural networks(Faster R-CNN) is an elegant and effective solution where proposal computation is nearly cost-free given the detection network's computation.

(2) Object Tracking

Optical Flow (OF) can be used to do object detection and tracking, it can track vehicles by directly measuring the new position and the displacement of interest points[8]. In [9], vehicles' positions and velocities are estimated using Kalman filtering. Particle filtering has also been widely used for monocular tracking in the image plane [5].

(3) Behavior Analysis

In [10], a method of modeling each object's environment with its own occupancy grid and measuring the motion is proposed, this method is based on Baysian Networks for predicting the likelihood of motion. Classification of behavior is performed by using a variety of techniques [3]. In [11] and [12], we can also use optical flow to detect overtaking behavior.

II. Related Work Review

A. Pedestrian Oriented

In [13], a vision-based tracking system which delivers count and behavior analysis of pedestrians is proposed.

It improves contextual fusion system [14], which uses Gaussian mixture model (GMM), by adding local binary pattern(LBP) to detect pedestrian. GMM [15] is an adaptive background subtraction method and is used to detect moving pedestrians in a scene by creating the adaptive background. LBP is the appearance-based method which can recognize a target directly from an image [13]. For tracking system, it benefits from the cooperation of OP and

bipartite graph matching of detections. The bipartite-graph matching can be used to initialize the track with the detected pedestrians given by previous detection system. Then the initialized track uses OP to handle partial occlusion. This paper also provides scene configuration to detecting and tracking systems by defining mix areas and typical paths.

The results of detecting pedestrian in this is paper are shown in Fig. 2. Detection results are evaluated using true positive rate (TPR). Fig. 3 illustrates the detection results evaluation of state-of-art methods shown in [16] , which uses miss rate of detecting pedestrian.

| INT	Method	Green		Red		Total	
		TPR	J	TPR	J	TPR	J
	GMM	0.14	0.13	0.25	0.21	0.22	0.19
INT 1	LBP	0.34	0.20	0.22	0.13	0.25	0.15
	Haar	0.41	0.10	0.34	0.08	0.36	0.09
	GMM+Haar [7]	0.30	0.29	0.25	0.23	0.27	0.24
	GMM+LBP	0.32	0.31	0.20	0.19	0.23	0.22
	GMM+Haar+ LBP	0.46	0.43	0.35	0.31	0.38	0.34
	GMM	0.08	0.06	0.18	0.15	0.13	0.11
INT 2	LBP	0.64	0.18	0.61	0.20	0.64	0.19
	Haar	0.70	0.06	0.48	0.05	0.59	0.06
	GMM+Haar [7]	0.46	0.25	0.53	0.27	0.50	0.26
	GMM+LBP	0.52	0.41	0.58	0.41	0.55	0.41
	GMM+Haar+ LBP	0.56	0.30	0.64	0.33	0.60	0.31

Fig. 2
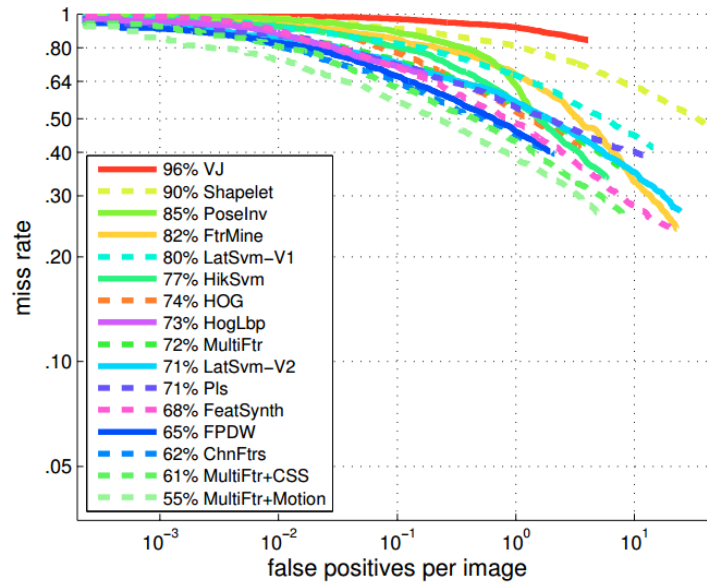
Fig. 3

For tracking evaluation, there are five criteria to evaluate the performance of the tracking system quantitatively:
1) Number of mostly tracked (MT) trajectories: more than 80% of the trajectory is tracked. The value should be high. 2) Number of mostly lost (ML) trajectories: more than 80% of the trajectory is lost. The value should be low. 3) Number of fragments (FG) of trajectories: the

generated trajectory is between 80% and 20% of the ground truth. 4) Number of false trajectories (FT): trajectories corresponding to no real object. The value should be low. 5) The frequency of identity switches (IS): identify exchanges between a pair of result trajectories. The value should be low.

The comparison of optical flow (OF) with proposed tracking methods is shown in Fig. 4.

Intersection	Tracking Method	GT	MT	ML	FG	FT	IS
INT 2	OF [4], [5]	14	4	7	19	30	3
	OF (GMM+LBP)	14	8	4	5	17	2
	OF (GMM+Haar+LBP)	14	11	2	4	20	2
INT 3	Optical flow [4], [5]	80	10	64	18	21	6
	OF (GMM+LBP)	80	20	53	73	3	5
	OF (GMM+Haar+LBP)	80	34	38	56	15	4

Fig. 4

B. Vehicle Oriented

The work in [17] presents a vision-based vehicle tracking system designed for improved intersection behavior analysis, this system can evaluate the count, speed and waiting time of vehicle's turning movement.

Vehicle detection is performed by the GMM, which can also address lighting changes, repetitive motions from clutter and long-term scene changes. For tracking, it uses bipartite graph to find the nearest objects between two successive frames, and applies dynamic model for track matching and updates. The improvement of this work is tracking the vehicle using 5 defined zones at the intersection, including 4 cardinal directions along with center of the intersection and using longest common subsequence (LCSS) to do path recognition. Because instead of a one-to-one mapping between all points in two trajectories, some points with no good match can be ignored, which gives the track path good performance [18]. The behavior analysis is performed by reconstructing the path, therefore the system can find out the movements such as stopping, turning etc. Speed measurements are calculated using trajectories and the frequency of speed belongs to each category, which is {Stop, Slow, Normal, Speeding}. And the waiting time can also be estimated by counting frames in different speed profiles during tracking.

Fig. 5 The five zones in the intersection

The counting accuracy of this system, which uses zone and LCSS, comparing to manual counting is 84%, and the previous method which uses zone only to resolve the tracking issue has the accuracy of 64% [17].

C. Conflict Oriented

In [19], it presents an automated video analysis system that: 1) detects and tracks road users in a traffic scene2) identifies important events that may lead to collisions; 3) calculates several severity conflict indicators. This study is unique in its attempt to extract conflict indicators from video sequences in a fully automated way.

This work uses feature-based detection [20] and tracking system [21], because it can handle partial occlusion. This system applies a threshold on the maximum speed of each road user to discriminate between pedestrians and motorized road users in mixed traffic. This method classifies bicyclists as motorized road users as well which is inaccurate, so it will be improved in the future by using object classifiers based on background subtraction and image appearance. The system detects all events constituted by the pairs of pedestrians and vehicles that are in the traffic scene simultaneously. And the indicators can be detected by the algorithm in [19] appendix 1, which is based on position, velocity of both pedestrian and vehicle.

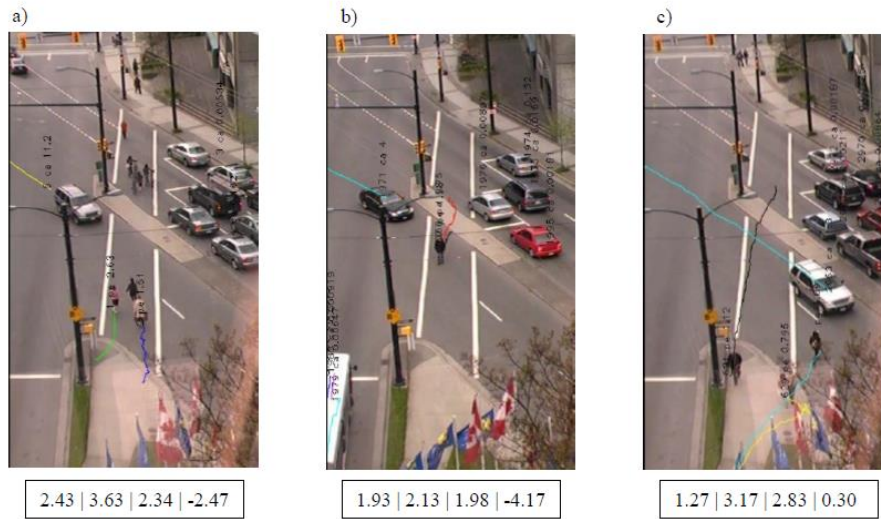| 2.43 | 3.63 | 2.34 | -2.47 | | 1.93 | 2.13 | 1.98 | -4.17 | | 1.27 | 3.17 | 2.83 | 0.30 |

Fig. 6 Sample of automatically detected important events with the road users' trajectories. The numbers under each image are respectively the min TTC (seconds), PET (seconds), maximum DST (m/s2), and min GT (s). In the images, the road user speed is indicated in m/s.

Four conflict indicators are calculated in this study, Time-to-Collision (TTC), Post-Encroachment Time (PET), Gap time (GT), Deceleration-to-Safety Time (DST). A combination of the four conflict indicators enables the system to automatically capture 89.5% of the conflicts and 71.7% of important events while detecting 54.5% of undisturbed passage events as important events. This system can be improved by enhancing the accuracy of detecting and tracking road users in more crowded traffic scenes.

III. Reference

[1]   "Intersection collision warning system.pdf." .

[2]   T. Gandhi and M. M. Trivedi, "Pedestrian Protection Systems: Issues, Survey, and Challenges," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 3, pp. 413–430, Sep. 2007.

[3]   S. Sivaraman and M. M. Trivedi, "Looking at Vehicles on the Road: A Survey of Vision-Based Vehicle Detection, Tracking, and Behavior Analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1773–1795, Dec. 2013.

[4]   F. Suard, A. Rakotomamonjy, A. Bensrhair, and A. Broggi, "Pedestrian Detection using Infrared images and Histograms of Oriented Gradients," in *2006 IEEE Intelligent Vehicles Symposium*, 2006, pp. 206–212.

[5]   H. T. Niknejad, A. Takeuchi, S. Mita, and D. McAllester, "On-Road Multivehicle Tracking Using Deformable Object Model and Particle Filter With Improved Likelihood Estimation," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 748–758, Jun. 2012.

[6]   M. Szarvas, A. Yoshizawa, M. Yamamoto, and J. Ogata, "Pedestrian detection with convolutional neural networks," in *IEEE Proceedings. Intelligent Vehicles Symposium, 2005.*, 2005, pp. 224–229.

[7]   S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PP, no. 99, pp. 1–1, 2016.

[8]   Y. Zhu, D. Comaniciu, V. Ramesh, M. Pellkofer, and T. Koehler, "An integrated framework of vision-based vehicle detection with knowledge fusion," in *IEEE Proceedings. Intelligent Vehicles Symposium, 2005.*, 2005, pp. 199–204.

[9]   S. Bota and S. Nedevschi, "Tracking multiple objects in urban traffic environments using dense stereo and optical flow," in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2011, pp. 791–796.

[10]   D. Kasper *et al.*, "Object-Oriented Bayesian Networks for Detection of Lane Change Maneuvers," *IEEE Intell. Transp. Syst. Mag.*, vol. 4, no. 3, pp. 19–31, Fall 2012.

[11]   F. Garcia, P. Cerri, A. Broggi, A. de la Escalera, and J. M. Armingol, "Data fusion for overtaking vehicle detection based on radar and optical flow," in *2012 IEEE Intelligent Vehicles Symposium*, 2012, pp. 494–499.

[12]   A. Geiger and B. Kitt, "Object flow: A descriptor for classifying traffic motion," in *2010 IEEE Intelligent Vehicles Symposium*, 2010, pp. 287–293.

[13]   M. S. Shirazi and B. Morris, "Vision-based pedestrian monitoring at intersections including behavior crossing count," in *2016 IEEE Intelligent Vehicles Symposium (IV)*, 2016, pp. 1022–1027.

[14]   M. S. Shirazi and B. Morris, "Contextual Combination of Appearance and Motion for Intersection Videos with Vehicles and Pedestrians," in *Advances in Visual Computing*, 2014, pp. 708–717.

[15]   Z. Zivkovic and F. van der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," *Pattern Recognit. Lett.*, vol. 27, no. 7, pp. 773–780, May 2006.

[16]   P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian Detection: An Evaluation of the State of the Art," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 743–761, Apr. 2012.

[17]   M. S. Shirazi and B. T. Morris, "Vision-Based Turning Movement Monitoring:Count, Speed Waiting Time Estimation," *IEEE Intell. Transp. Syst. Mag.*, vol. 8, no. 1, pp. 23–34, Spring 2016.

[18]   B. Morris and M. Trivedi, "Learning trajectory patterns by clustering: Experimental studies and comparative evaluation," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 312–319.

[19]   K. Ismail, T. Sayed, N. Saunier, and C. Lim, "Automated Analysis of Pedestrian-Vehicle Conflicts Using Video Data," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2140, pp. 44–54, Dec. 2009.

[20]   N. Saunier and T. Sayed, "A feature-based tracking algorithm for vehicles in intersections," in *The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)*, 2006, pp. 59–59.

[21]   X. Ma, N. Fan, Z. He, and W. Yang, "Real-Time Affine Tracking Using Re-located Lucas-Kanade Algorithm," in *2015 Third International Conference on Robot, Vision and Signal Processing (RVSP)*, 2015, pp. 35–38.