Proceedings of the 16th International IEEE Annual Conference on
Intelligent Transportation Systems (ITSC 2013), The Hague, The
Netherlands, October 6-9, 2013

TuD5.4

Observing On-Road Vehicle Behavior:
Issues, Approaches, and Perspectives

Sayanan Sivaraman[1], Brendan Morris[2], and Mohan Trivedi[1]

Abstract—In this work, we review recent works comprising an emerging field of intelligent transportation: behavior analysis of vehicles. The ITS community has approached this topic both from vehicle-based and infrastructure-based sensing. In both cases, motion is the key indicator required for behavioral characterization, with accurate long-term prediction being the ultimate goal. However, the popular methods for behavior characterization differ between the sensing methodologies. Vehicle-based sensing tends to focus on spatio-temporal measurements coupled with various features for accurate estimation of object state. In contrast, infrastructure-sensing tends to avoid attempting high resolution estimation of vehicle state and prefers to utilize patterns learned in aggregate for constrained estimation. This review focuses on vision-based sensing and provides highlights of state-of-the art methods used in surveillance, and on-road vision modalities. We provide discussion and comment on future directions in the field.

I. INTRODUCTION

Modern governments invest heavily in the installation and maintenance road networks. Integral to economic health, roadways allow transportation of goods to market, commuting, and leisure travel. Worldwide, road transport costs are currently growing at twice the rate of GDP [1]. A significant cost associated with roadways derives from congestion and traffic accidents. According to the World Bank, between 1 and 3 percent of global GNP is spent on costs associated with traffic accidents [1]. Automotive collisions annually kill tens of thousands of peoples are killed on the roads each year, and most fatal crashes feature more than one vehicle [2]. As research in sensing and environmental perception progresses, there is great potential to save lives and cost, by developing intelligent characterization of on-road behavior, and tools for long-term prediction of critical events.

Given the large volume of data available, researchers have looked at characterizing the behavior of vehicles on the road from two distinct vantage points. Intelligent vehicles researchers primarily use vehicles equipped with sensors, like cameras, lidar, and radars, to sense and perceive the on-road environment, tracking other vehicles on the road. Many infrastructure and surveillance researchers mainly use static cameras to track vehicles. These two bodies of research examine similar problem domains with different sets of data, the primary distinction moving vs. static sensing. Interestingly,

modeling and algorithmic approaches also differ between the two bodies of research.

In this work, we review the characterization of on-road vehicle behavior, from both the vehicle-base, and infrastructure-based frames of reference. We highlight the differing assumptions, data fidelities, modeling, and algorithmic approaches that researchers have taken in this quickly-evolving field. With advanced sensors integrated both in infrastructure and on vehicles, researchers have access to new insight and opportunities to improve transportation quality in myriad ways. We detail the state-of-the art in each of these domains, and provide perspectives on future directions in this body of research.

II. VEHICLE-BASED BEHAVIOR ANALYSIS

Analysis of the behavior of tracked vehicles has emerged as an active and challenging research area. While considerable research effort has been dedicated to on-road detection and tracking of vehicles in images and video, the highest level of semantic interpretation lies in characterizing the behavior of vehicles on the road.

Research studies in this area take a variety of approaches to characterize on-road behavior [3]. Certain studies try to categorize observed vehicle behavior as normal or abnormal [4]. Other studies try to identify specific maneuvers, such as overtaking [5], turning [6], or lane changes [7]. Most recently, studies in the literature have tried to make long term classification and prediction of vehicle motion. Trajectory modeling approaches try to predict vehicle motion up to 2 seconds ahead, based on models of typical vehicle trajectories [8], [9].

Broadly speaking, we categorize studies that address the characterization of on-road vehicle behavior based on four main criteria. Firstly, we consider the role of context in the analysis of on-road behavior, loosely defined to encompass considerations such as urban driving vs. highway driving, or intersection vs. non-intersection driving. Secondly, we consider the identification of pre-specified maneuvers, such as turning, lane change, or overtaking maneuvers of tracked vehicles on the road. Thirdly, we consider the use of trajectories, long-term sequences of positions and velocities, in characterizing on-road behavior. Finally, we consider classification and modeling found in the literature.

A. Context

Context plays a vital role in many studies. Using context can help narrow the scope of the sensing problem, while increasing robustness. The motion model used in [14] models

TABLE I
REPRESENTATIVE WORKS IN ON-ROAD BEHAVIOR ANALYSIS

Research Study	Context-Specific or General?	Maneuver Detection	Trajectory-based?	Classification/ Inference	Description
Diaz-Alonso et al., 2008 [10]	General	Overtaking vehicles	No	Template matching score	Detection and tracking of overtaking vehicles
Cherng et al., 2009 [4]	Context-specific	No	No	Neural network	Dynamic visual model of typical on-road behavior; saliency used to detect unusual and critical situations.
Barth and Franke, 2010 [6]	Context-specific	Turning behavior	No	Interacting multiple model likelihood	Velocity and yaw-rate estimation used to infer turning behavior of oncoming vehicles.
Geiger and Kitt, 2010 [11]	Context-specific	No	No	Support vector machine	Histograms of scene flow used to classify intersection vs. non-intersection driving environment.
Hermes et al., 2010 [12]	Context-specific	No	Yes	Augmented particle filter	Vehicle motion is matched to 44 prototypes using QRLCS distance.
Sivaraman et al., 2011 [13]	Context-specific	No	Yes	hidden Markov model	Unsupervised clustering of observed on-road trajectories.
Kasper et al., 2011 [7]	General	Lane change	No	Dynamic Bayesian network	Dynamic Bayesian network is used to predict lane changes of other vehicles.
Garcia et al., 2012 [5]	General	Overtaking vehicles	No	Optical flow direction, intensity	Optical flow is used to detect overtaking vehicles.

the distribution of vehicles in the image plane, using it as a prior probability on vehicle detections. The vehicle detection in [14] can be viewed as a detection-by-tracking approach, enabled by spatio-temporal modeling of the driving context. In [11], histograms of scene flow vectors are used to classify the driving environment as intersection or non-intersection driving, modeling the driving context using spatio-temporal information. In [4], a dynamic visual model is developed of the driving environment, with saliency alerting the system to unusual and critical on-road situations.

B. Specific Maneuver Identification

A body of work has been dedicated to the detection of specific maneuvers of vehicles on the road. In [10], overtaking behavior is detected, by detecting vehicles in the blind spot of the ego-vehicle. Overtaking behavior is specifically detected in [15], this time with the camera pointing forward, and vehicles detected as they overtake in front of the ego-vehicle. In [5], overtaking behavior is also detected in front of the ego-vehicle, using a fusion of vision and radar. Overtaking behavior is also detected in [16], also for vehicles in front of the ego-vehicle. In these studies, the overtaking maneuver is detected by virtue of detecting the vehicle, as the search space includes only vehicles that are in the process of overtaking.

By contrast, specific maneuvers are identified in other works via inference on tracking information. In [17], the turning behavior of tracked vehicles is identified by measuring the yaw rate using extended Kalman filtering. Using the yaw rate in the vehicle motion model, the system is able to detect turning behavior. In [6], turning behavior is further addressed, using interacting multiple models to characterize the motion of the oncoming vehicle. The model with the highest likelihood, based on observations, characterizes the

turning behavior of the oncoming vehicle, with a transition probability handling change of states. Turning behavior is addressed in [18] by solving the vehicle's pose, with respect to the ego vehicle using clustering of 3D points.

C. Trajectory Analysis

The use of vehicle trajectories, to characterize and learn on-road vehicle behaviors, has emerged in the past few years. A trajectory is typically defined as a data sequence, consisting of several concatenated state vectors from tracking, meaning an indexed sequence of positions and velocities over a given time window. Using a time window of 1 second, for example, can mean trajectories consisting of 25-30 samples, depending on the frame rate of the camera.

D. Classification and Modeling

On-road vehicle behavior is modeled in [19] as a Markov process, and inferred using a dynamic Bayesian network, based on tracking observations. However, the experimental evaluation is performed using simulation data. In [7], the lane change behavior of tracked vehicles is modeled using a dynamic Bayesian networks, and the experiments are performed on real-world vision data.

In [20], variational Gaussian mixture modeling is used to classify and predict the long-term trajectories of vehicles, using simulated data. In [13], highway trajectories are recorded using stereo-vision, and clustering is performed to model the typical trajectories encountered in highway driving, with classification performed using hidden Markov modeling. In [8], trajectories are classified using a rotation-invariant version of the longest common subsequence as the similarity metric between trajectories. Vehicle trajectories are used to characterize behavior at roundabouts in [12], using the

QRLCS metric to match observed trajectories to a database of pre-recorded trajectories. Similar work is carried out in [21] for vehicles at intersections.

In studies that explicitly classify vehicle behavior, we see a preponderance of generative modeling. In [20], Gaussian mixture modeling is used, which provides distribution over the prediction, complete with a point-estimate [the conditional mean], and a covariance to convey uncertainty. In [6], the likelihood of the interacting multiple model tracking is used to classify the tracked vehicle's turning behavior, complete with a transition probability. In [7], Bayesian networks are used for classifying the vehicle's behavior, and predicting the vehicle's lane change. In [13], hidden Markov modeling is used to model each of the prototypical trajectories, learned using clustering. In [19], the vehicle behavior is also modeled as a Markov process, with observations coming from the vehicle's instantaneous state vector. Table I highlights some representative works in vision-based on-road behavior analysis.

III. INFRASTRUCTURE-BASED BEHAVIOR ANALYSIS

Early infrastructure systems were installed to collect traditional traffic parameters – density, flow, speed – used by transportation engineers to manage capacity. Recently there has been a shift towards the use of video sensors [22], [23] due to rich information content in video, vastly improved computer vision processing, and the desire to have video available for monitoring at transportation management centers.

Recently, there has been a major push in the vision community to move away from instantaneous counts and measurements and better utilize the vast amount knowledge encoded in traffic video feeds. In this vein, research has focused on data mining, machine learning, and pattern recognition to leverage the big data in transportation more effectively. A key constraint in behavior understanding algorithm design from infrastructure video arises from far-field surveillance. Typically, video resolution is low and vehicles are small in size making it extremely difficult to extract complex descriptors such as pose with real accuracy. Instead, only simple features such as position and velocity, can be used to characterize behavior.

A. Unsupervised Contextualization

One shortcoming of kinematic and dynamic motion models for predicting future state evolution in time is their performance typically degrades quickly with increasing prediction time-horizons. This is particularly apparent when motion is complex (e.g. u-turn) or when some internal state or intention motivates motion (decision to turn at an intersection). In these cases, the dynamic models are not able to be precisely modeled. By observing motion, over time the typical motion patterns can be learned and used as a priori knowledge for prediction. The advantage of these techniques is that they are generally applicable and do not require manual re-training for new scenes or scenarios.

Motion pattern learning research can be divided into two categories 1) trajectory-based and 2) motion-based. Trajectory-based algorithms augment the typical vehicle detection and tracking analysis pipeline with an additional trajectory learning module while motion-based algorithms do not rely on accurate tracking (just motion primitives) to learn patterns.

B. Trajectory Patterns

The key questions that arise in trajectory learning are how to compare trajectories of varying length and how to ensure that all semantically meaningful patterns are extracted in a completely unsupervised fashion. Various trajectory specific similarity measures (or distances) have been proposed to make trajectory clustering compatible with traditional clustering algorithms. Comparisons have shown that the clustering algorithm is not as important as selection of the type of similarity measure [24]. However, it is still unclear how to best extract meaningful patterns, however, recently iterative clustering techniques are studied to enable time dependent patterns that are able to adapt to new data and changing conditions [25]–[27].

The growing hidden Markov model is utilized in [25] in order to have an online adaptive model that changes with conditions rather than based on a fixed observation time. Voronoi tessellation is used to build up a topological map of spatial states which are then connected through the growing HMM. Rather than explicitly model trajectory patterns, this work looks at the probability of goals (the eventual endpoint of a trajectory) and gives a distribution of look ahead states. Interestingly, the patterns can be learned incrementally and in parallel with prediction.

In [28], a three level hierarchical learning framework was developed to robustly cluster and model trajectory patterns. The first level accounted for the goals (entries and exits) of the scene, the second level performed spectral clustering of spatial trajectories, while the final level modeled the resulting clusters with HMMs that included spatial and dynamic information. This system provided a cluster adaption scheme based upon maximum likelihood linear regression as well as a batch update but did not have a true incremental learning framework to handle time-varying patterns.

Non-parametric Bayesian modeling in the form of a dual hierarchical Dirichlet Process (Dual-HDP) has been utilized for unsupervised trajectory analysis and semantic region modeling [26]. In this approach, trajectories were considered as documents and the trajectory observations words and both the number of activity categories (patterns) and semantic regions (shared sub-patterns) were automatically learned from the data. This was further extended to the Dynamic Dual-HDP to allow for dynamic update of models and online detection of abnormal activities. In addition, this worked showed results from both vision-based trajectories as well as radar tracks demonstrating the general applicability of trajectory learning.

In [29], object tracking is performed but each observation is treated independently. The image space is vector quantized based on the density of track points resulting in a Voronoi tessellation with each cell representing an alphabet entry. Trajectories are then remapped into strings using the P-spectrum kernel to count common substrings for graph cuts spectral clustering to provide an behavior hierarchy.

Very recent work by Hu et. al [27] also used the Dirichlet process mixture model (DPMM) to cluster, model and retrieve trajectories in an incremental fashion. Further, the time-varying information contained in a trajectory is modeled using the time-sensitive Dirichlet process mixture model (tDPMM) over sub-trajectories.

For more historical perspective, readers are directed to the 2008 survey on trajectory learning [30].

C. Motion Grouping

The previous section highlighted a number of methods that were able to learn scene context from observation of trajectory data. However, the quality of those methods are highly dependent on robust tracking of vehicles which is inherently difficult in general due to noise, changing lighting conditions, shadows, and occlusion. Another class of unsupervised learning methods have utilized low-level motion cues which can be easily obtained, such as optical flow, and focused instead on developing more complex modeling to group motion into behavior. These motion based learning techniques borrow from topic-model literature used for web-based document clustering literature [35]. Video is segmented into clips (a document) and each clip can be described with a small vocabulary of words (motion vectors) and the collection of words expresses the topics inherent in the video clip (the prevalent behaviors) [27], [31]–[34]. These topic models are unsupervised and learn through co-occurrence of visual features without manual labeled training examples. In addition topic models are hierarchical Bayesian models that can model simple actions and more complicated ones jointly at various hierarchical levels which make them well suited for general application in various infrastructure settings.

In [31], dense optical flow vectors are computed as motion descriptors to avoid tracking. Video is divided into 1 sec clips and all optical flow vectors are clustered using k-means into a mixture of Gaussian distributions. Mixture components were treated as nodes in a graph connected though time based on a reachability criterion that relied on a constant motion model between proximal clips. The resulting Gaussian chains defined the distribution of a motion pattern were used for inference based on Kullback-Leibler divergence.

A two level Latent Dirichlet Allocation (LDA) topic model was used to learn scene behaviors. The first level learned single-agent motion while the second LDA level used the first level output to learn interactions over multi-agents. This hierarchy enabled anomaly detection at both levels for every video frame rather than for clips [32].

In 2012, a framework to trade-off computational complexity of the Bayesian topic model processes with scalability for large datasets was introduced [33]. This work used a fast and simple rank-1 robust principle component analysis (RPCA-R1) for foreground pixel detection. The counts of foreground pixels in 16×16 blocks was used as input words to a DPMM for learning. The DPM provided scalability to large datasets through incremental learning and inference.

In [34], a word-document hierarchical topic model is used to model video clips of optical flow words. Sparse topical coding (STC) is used to efficiently represent the clips with only a sparse set of motion patterns.

The interested reader is directed to the recent review of topic models for action recognition by Wang [35] for more complete treatment.

IV. DISCUSSION AND FUTURE DIRECTIONS

A. On-Road Behavior Analysis

On-road behavior analysis is a relatively immature, but rapidly developing area of research. Advances in this area seek to address issues such as identification of maneuvers, characterization of vehicle behavior, and long-term motion prediction. Only in the past few years have vehicle detection and tracking methodologies become sufficiently mature to enable the exploration of these deeper questions.

In characterizing vehicle behavior, a major challenge will be in identifying erratic, abnormal, or aggressive driving by other vehicles. While identifying specific maneuvers can be formulated as a well-defined problem, abstractly characterizing another vehicle's behavior remains an open question. Given the tracking of vehicles with respect to their own lanes, weak cues such as a vehicle's veering within its lane, or crossing over lane boundaries could be used. More likely, research studies will try to characterize normal driving behavior for a given context in a data-driven manner, and identify abnormal trajectories by measuring the model fit of an observed vehicle trajectory [13].

Long term motion prediction requires an estimate of the vehicle's motion 1-2 seconds, or 25-50 frames, ahead of time, outside the capabilities of conventional filtering techniques. Long-term motion classification and prediction will involve further research into learning and modeling of vehicle trajectories. An enhanced understanding of vehicle trajectories will allow on-board systems to infer the intent of other vehicle's drivers, based on sequential tracking measurements from vision-based systems.

There should be movement towards understanding the motion and behavior of other vehicles as independent traffic agents. To this end, we foresee learned trajectory models working in concert with established tracking approaches like interacting multiple models. A full vehicle motion understanding engine would include multiple trackers with distinct motion models to estimate vehicle state in the short-term, interacting multiple models to identify vehicle maneuvering in the medium term, and trajectory learning to predict vehicle motion in the long term. Associated issues, such as data windowing, and online model updates, will also need to be addressed.

TABLE II
REPRESENTATIVE WORKS IN INFRASTRUCTURE BEHAVIOR ANALYSIS

Research Study	Input Data	Classification/ Inference	Description
Vasquez et al., 2008 [25]	Trajectory	growing HMM	Online adaptive modeling of the different ways to move from scene goals.
Saleemi et al., 2010 [31]	Optical Flow	GMM	Gaussian flow distributions are learned for video clips and chained together across clips to define a pattern distribution.
Morris and Trivedi, 2011 [28]	Trajectory	HMM	Three-level learning hierarchy accounting for goals, spatial paths, and dynamic behaviors.
Wang et al., 2011 [26]	Trajectory	Dynamic Dual-HDP	Dynamic Dual-HDP non-parametric Bayesian model to automatically model activity categories and semantic regions without specifying the number of topics and with online update of model.
Song et al., 2011 [32]	Optical Flow	LDA	Two level LDA topic model learned first for single-agent motion which is input to second level LDA for multi-agent interactions.
Rana et al., 2011 [33]	Foreground Pixels	DPMM	Fast rank-1 robust PCA used for foreground detection with counts of pixels in blocks used as input for DPMM learning which enables incremental learning and inference.
Noceti and Odone , 2012 [29]	Trajectory	Common Strings	Image space quantized with Voronoi tessellation to build an alphabet for representing a trajectory as a string.
Fu et al., 2012 [34]	Optical Flow	LDA	Sparse topical coding used for efficient learning and representation of the topic model.
Hu et al., 2013 [27]	Trajectory	tDPMM	Dirichlet process mixture model is used for unsupervised clustering and modified to handle temporal structure and ordering inherent in a trajectory sequence.

B. Infrastructure-Based Behavior Analysis

Perhaps the most natural extension to traditional traffic analysis one can envision with the use of video technology is advanced conflict analysis. Rather than waiting for rare collision events (and annual reports that may lack critical information), with traffic cameras, critical locations can be monitored in real-time. Safety can then be analyzed in a proactive manner based on the interactions between vehicles. A hierarchy can be constructed utilized to classify interactions based on the interactions with the rare collisions on the top, the safe interactions on the bottom (the large majority), and conflicts in the middle. Conflicts, the situation when two vehicles could collide without intervention, can be used as a surrogate safety measure [36].

Surrogate safety analysis is only possible with accurate prediction of future events. Prediction is difficult because of the complex dynamics and control with vehicular traffic; a driver will change the kinematic state based on an internal plan or intentions. Numerical techniques, such as Markov Chain Monte Carlo or multiple hypothesis tracking, can be used to enumerate many possible future outcomes as well as attempting to guess the driver intention, however, these are computational expensive [37]. Instead of complex motion models, patterns learned from motion or trajectory clustering can be used to generate likely future trajectories. These clusters would represent the possible intentions of a driver.

It is noteworthy to mention that the increased emphasis in trajectory pattern analysis for driving behavior understanding has been fueled by the emergence of two popular datasets: 1) NGSIM [38] and 2) MIT intersection [39]. These datasets have been used by many of the cited papers for behavior prediction and anomaly detection. This sharing of benchmark data is incredibly important for improving research results because it provides fair comparison and lowers the barrier of entry into the area.

C. Combining Infrastructure and Vehicle-Based Analysis

The promise of new vehicle to vehicle (V2V) and vehicle to infrastructure (V2I) communication technology will usher in a new age of cooperative algorithms. Processing and analysis will be shared between vehicles and the infrastructure to provide enhanced environmental perception and situational awareness. For example, in simulation, intersection cameras with wide-area coverage have be used to augment the blind-spots in an in-vehicle camera system to improve warning systems [40].

Work is already underway using floating car data to learn about behaviors. Public willingness to share GPS data from phones and navigation systems, will provide massive data to be mined for insight into flow and congestion, while providing rich behavioral information. This data could be used for personalized reports of traffic patterns or directed advertisements. But, it can also be used in aggregate to learn trends – most popular routes between destinations, crime prevention [41], [42], etc.

V. CONCLUDING REMARKS

This manuscript has reviewed recent progress in understanding traffic behavior both from vehicle-based and infrastructure-based sensing. We have reviewed the state-of-the art in each domain, contrasting data resolution, modeling, and algorithmic approaches between the two problem domains. We have also provided perspective on future directions in vehicle-based behavior characterization, infrastructure-based behavior characterization, and open questions in the integration of data from each domain. The authors believe that this will be a fertile research ground in coming years, and that there are ample opportunities to integrate infrastructure and vehicle-based analysis.

REFERENCES

[1] "World bank development report, roads and high-ways: Road safety," 2013. [Online]. Available: http://www.worldbank.org/transport/roads/safety.htm

[2] Department of Transportation - NHTSA, "Traffic safety facts," 2011. [Online]. Available: http://www-nrd.nhtsa.dot.gov

[3] S. Sivaraman and M. Trivedi, "Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis," *Intelligent Transportation Systems, IEEE Transactions on*, 2013.

[4] S. Cherng, C.-Y. Fang, C.-P. Chen, and S.-W. Chen, "Critical motion detection of nearby moving vehicles in a vision-based driver-assistance system," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 10, no. 1, pp. 70 –82, march 2009.

[5] 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 *Intelligent Vehicles Symposium (IV), 2012 IEEE*, june 2012.

[6] A. Barth and U. Franke, "Tracking oncoming and turning vehicles at intersections," in *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, sept. 2010, pp. 861 –868.

[7] D. Kasper, G. Weidl, T. Dang, G. Breuel, A. Tamke, and W. Rosenstiel, "Object-oriented bayesian networks for detection of lane change maneuvers," in *Intelligent Vehicles Symposium (IV), 2011 IEEE*, june 2011, pp. 673 –678.

[8] C. Hermes, C. Wohler, K. Schenk, and F. Kummert, "Long-term vehicle motion prediction," in *Intelligent Vehicles Symposium, 2009 IEEE*, june 2009, pp. 652 –657.

[9] B. Morris and M. Trivedi, "Unsupervised learning of motion patterns of rear surrounding vehicles," in *Vehicular Electronics and Safety (ICVES), 2009 IEEE International Conference on*, 2009.

[10] J. Diaz Alonso, E. Ros Vidal, A. Rotter, and M. Muhlenberg, "Lane-change decision aid system based on motion-driven vehicle tracking," *Vehicular Technology, IEEE Transactions on*, vol. 57, no. 5, pp. 2736 –2746, sept. 2008.

[11] A. Geiger and B. Kitt, "Objectflow: A descriptor for classifying traffic motion," in *IEEE Intelligent Vehicles Symposium*, June 2010.

[12] C. Hermes, J. Einhaus, M. Hahn, C. Wo andhler, and F. Kummert, "Vehicle tracking and motion prediction in complex urban scenarios," in *Intelligent Vehicles Symposium (IV), 2010 IEEE*, june 2010.

[13] S. Sivaraman, B. T. Morris, and M. M. Trivedi, "Learning multi-lane trajectories using vehicle-based vision," *International Conference on Computer Vision, Workshop on Computer Vision for Vehicle Technology*, 2011.

[14] A. Jazayeri, H. Cai, J. Y. Zheng, and M. Tuceryan, "Vehicle detection and tracking in car video based on motion model," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 12, no. 2, pp. 583 –595, june 2011.

[15] Y. Zhu, D. Comaniciu, M. Pellkofer, and T. Koehler, "Reliable detection of overtaking vehicles using robust information fusion," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 7, no. 4, pp. 401 –414, dec. 2006.

[16] J. Wang, G. Bebis, and R. Miller, "Overtaking vehicle detection using dynamic and quasi-static background modeling," in *Computer Vision and Pattern Recognition - Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on*, june 2005, p. 64.

[17] A. Barth and U. Franke, "Estimating the driving state of oncoming vehicles from a moving platform using stereo vision," *IEEE Trans. Intell Transp. Syst.*, vol. 10, no. 4, Dec. 2009.

[18] B. Barrois, S. Hristova, C. Wohler, F. Kummert, and C. Hermes, "3d pose estimation of vehicles using a stereo camera," in *Intelligent Vehicles Symposium, 2009 IEEE*, june 2009, pp. 267 –272.

[19] T. Gindele, S. Brechtel, and R. Dillmann, "A probabilistic model for estimating driver behaviors and vehicle trajectories in traffic environments," in *Intelligent Transportation Systems (ITSC)*, sept. 2010, pp. 1625 –1631.

[20] J. Wiest, M. Hoffken, U. Kresel, and K. Dietmayer, "Probabilistic trajectory prediction with gaussian mixture models," in *Intelligent Vehicles Symposium (IV), 2012 IEEE*, june 2012, pp. 141 –146.

[21] E. Kafer, C. Hermes, C. Woandhler, F. Kummert, and H. Ritter, "Recognition and prediction of situations in urban traffic scenarios," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, aug. 2010.

[22] N. Buch, S. Velastin, and J. Orwell, "A review of computer vision techniques for the analysis of urban traffic," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 12, no. 3, pp. 920–939, 2011.

[23] B. Tian, Q. Yao, Y. Gu, K. Wang, and Y. Li, "Video processing techniques for traffic flow monitoring: A survey," in *Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on*, 2011, pp. 1103–1108.

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

[25] D. Vasquez, T. Fraichard, and C. Laugier, "Incremental learning of statistical motion patterns with growing hidden markov models," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 10, no. 3, pp. 403–416, 2009.

[26] X. Wang, K. Ma, G.-W. Ng, and W. Grimson, "Trajectory analysis and semantic region modeling using nonparametric hierarchical bayesian models," *International Journal of Computer Vision*, vol. 95, no. 3, pp. 287–312, 2011. [Online]. Available: http://dx.doi.org/10.1007/s11263-011-0459-6

[27] W. Hu, X. Li, G. Tian, S. Maybank, and Z. Zhang, "An incremental DPMM-based method for trajectory clustering, modeling, and retrieval," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 5, pp. 1051–1065, 2013.

[28] B. Morris and M. Trivedi, "Trajectory learning for activity understanding: Unsupervised, multilevel, and long-term adaptive approach," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 11, pp. 2287–2301, 2011.

[29] N. Noceti and F. Odone, "Learning common behaviors from large sets of unlabeled temporal series," *Image and Vision Computing*, vol. 30, no. 11, pp. 875 – 895, 2012.

[30] B. Morris and M. Trivedi, "A survey of vision-based trajectory learning and analysis for surveillance," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 18, no. 8, pp. 1114–1127, 2008.

[31] I. Saleemi, L. Hartung, and M. Shah, "Scene understanding by statistical modeling of motion patterns," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 2010, pp. 2069–2076.

[32] L. Song, F. Jiang, Z. Shi, and A. Katsaggelos, "Understanding dynamic scenes by hierarchical motion pattern mining," in *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, 2011, pp. 1–6.

[33] S. Rana, D. Phung, S. Pham, and S. Venkatesh, "Large-scale statistical modeling of motion patterns: a bayesian nonparametric approach," in *Indian Conference on Computer Vision, Graphics and Image Processing*, 2012.

[34] W. Fu, J. Wang, Z. Li, H. Lu, and S. Ma, "Learning semantic motion patterns for dynamic scenes by improved sparse topical coding," in *Multimedia and Expo (ICME), 2012 IEEE International Conference on*, 2012, pp. 296–301.

[35] X. Wang, "Action recognition using topic models," in *Visual Analysis of Humans*, T. B. Moeslund, A. Hilton, V. Krger, and L. Sigal, Eds. Springer London, 2011, pp. 311–332. [Online]. Available: http://dx.doi.org/10.1007/978-0-85729-997-0_16

[36] N. Saunier, T. Sayed, and K. Ishmail, "Large-scale automated analysis of vehicle interactions and collisions," *Transportation Research Record: Journal of the Transportation Research Board*, 2010.

[37] M. G. Mohamed and N. Saunier, "Motion prediction methods for surrogate safety analysis," *J. Transport. Research Board*, 2012.

[38] (2012) NGSIM community. US DOT's Federal Highway Administration. [Online]. Available: http://ngsim-community.org/

[39] X. Wang, X. Ma, and W. Grimson, "Unsupervised activity perception in crowded and complicated scenes using hierarchical bayesian models," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 3, pp. 539–555, Mar. 2009.

[40] K. Minoura and T. Watanabe, "Driving support by estimating vehicle behavior," in *Pattern Recognition (ICPR), 2012 21st International Conference on*, 2012, pp. 1144–1147.

[41] L. Liao, D. Fox, and H. Kautz, "Learning and inferring transportation routines," *Artificial Intelligence*, 2007.

[42] L. Teng, Q. Shaojie, T. Changjie, L. Liangxu, L. Taiyong, and W. Jiang, "E3tp: A novel trajectory prediction algorithm in moving objects databases," *Intelligence and Security Informatics*, pp. 76–88, 2009.