Concept of Operations of Next-Generation Traffic Control Utilizing Infrastructure-Based Cooperative Perception

Preprint

Stanley Young,[1] Erik Bensen,[1] Lei Zhu,[2] Christopher Day,[3] J. Sam Lott,[4] Rimple Sandhu,[1] Charles Tripp,[1] and Peter Graf[1]

1 National Renewable Energy Laboratory
2 University of North Carolina at Charlotte
3 Iowa State University
4 Automated Mobility Systems, LLC

Presented at the ASCE International Conference on Transportation & Development
Seattle, Washington
May 31–June 3, 2022

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5400- 81978
June 2022

Concept of Operations of Next-Generation Traffic Control Utilizing Infrastructure-Based Cooperative Perception

Preprint

Stanley Young,[1] Erik Bensen,[1] Lei Zhu,[2] Christopher Day,[3] J. Sam Lott,[4] Rimple Sandhu,[1] Charles Tripp,[1] and Peter Graf[1]

1 National Renewable Energy Laboratory
2 University of North Carolina at Charlotte
3 Iowa State University
4 Automated Mobility Systems, LLC

Suggested Citation
Young, Stanley, Erik Bensen, Lei Zhu, Christopher Day, J. Sam Lott, Rimple Sandhu, Charles Tripp, and Peter Graf. 2022. *Concept of Operations of Next-Generation Traffic Control Utilizing Infrastructure-Based Cooperative Perception: Preprint*. Golden, CO: National Renewable Energy Laboratory. NREL/CP-5400-81978. https://www.nrel.gov/docs/fy22osti/81978.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5400- 81978
June 2022

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the Laboratory Directed Research and Development (LDRD) Program at NREL. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

Concept of Operations of Next-Generation Traffic Control Utilizing Infrastructure-Based Cooperative Perception

Stanley E. Young, Ph.D., P.E.,[1] Erik A. Bensen,[2] Lei Zhu, Ph.D.,[3] Christopher Day, Ph.D.,[4] J. Sam Lott,[5] Rimple Sandhu, Ph.D.,[6] Charles Tripp, Ph.D.,[7] Peter Graf, Ph.D.[8]

[1] Center for Integrated Mobility Sciences, National Renewable Energy Laboratory, Golden, CO 80401; email: Stanley.Young@nrel.gov
[2] Computational Science Center, National Renewable Energy Laboratory, Golden, CO 80401; email: Erik.Bensen@nrel.gov
[3] Department of Systems Engineering and Engineering Management, University of North Carolina at Charlotte, Charlotte, NC 28223; email: Lei.Zhu@uncc.edu
[4] Department of Civil, Construction, and Environmental Engineering, Iowa State University, Ames, IA 50011; email: cmday@iastate.edu
[5] Automated Mobility Systems, LLC, Houston, TX; email: jsamlott.amsllc@gmail.com
[6] National Wind Technology Center, National Renewable Energy Laboratory, Golden, CO 80401; email: Rimple.Sandhu@nrel.gov
[7] Computational Science Center, National Renewable Energy Laboratory, Golden, CO 80401; email: Charles.Tripp@nrel.gov
[8] Computational Science Center, National Renewable Energy Laboratory, Golden, CO 80401; email: Peter.Graf@nrel.gov

ABSTRACT

This paper provides a system architecture for an infrastructure-based cooperative perception fusion engine for next-generation traffic control. This engine will provide a complete state-space digital representation with measurable accuracy to support a wide-range of applications. The architecture includes inputs, functional flow, data standardization recommendations, outputs, and supported applications. The cooperative perception engine addresses critical needs with respect to accelerating the benefits of automation through intelligent roadway infrastructure, which complements and accelerates connected and automated vehicle (CAV) technology. The cooperative perception acquires and fuses information from sensors (radar, LiDAR, and cameras) and CAVs to perceive roadway traffic states of moving objects, creates a complete 3D digital representation of that state-space, and communicates it to downstream application such as intelligent signal control, safety and energy applications, and cooperate driving applications. The intelligent roadway infrastructure approach, as opposed to a vehicle-centric approach, is more scalable because it can be deployed to the roughly 300,000 signalized intersections more readily than the over 300 million vehicles in the United States, and accrues early-stage benefits equitable to all roadway users addressing safety, equity, fuel efficiency, and greenhouse gas reduction.

INTRODUCTION

Vehicle automation promises significant societal benefits, but the pace of adoption is hindered by a lack of robust vehicle-to-infrastructure connectivity. SAE International recognizes that a single vehicle cannot achieve full perception of its surroundings in complex roadway environments and stresses the importance of cooperative driving automation (CDA) to “enable mobility applications that are not achievable by individual automated driving system-operated vehicles operating independently of each other." Further, SAE explains that "sharing information can be used to increase safety, efficiency, and reliability of the transportation system, and that may serve to accelerate the deployment of driving automation in on-road motor vehicles" (SAE 2020). This acceleration of vehicle automation is critical to decarbonization, as automated vehicles are widely thought to offer the greatest potential for electrifying the transportation system. Automated vehicles, and more generally automated mobility, also play a pivotal role in providing a more equitable transportation system, offering a high-quality alternative to the vehicle ownership model for lower-income workers and people who cannot drive, and by filling in the gaps in the nation's transit system.

Infrastructure-based cooperative perception significantly advances the state of CDA by developing and deploying a computational engine, and then employing that engine to enhance the performance of connected and automated vehicles (CAVs) and enabling the next generation of traffic control that equitably provides benefits to all modes and road users. Having a robust infrastructure participant is a critical and necessary component for successful CDA applications, and a critical gap in current research and development. The competitive commercial automated vehicle market has necessitated closely guarded corporate intellectual property, leaving little fundamental research of spatial perception in the public realm. Recognizing the need for robust infrastructure involvement, the U.S. Department of Transportation initiated Cooperative Automation Research Mobility Applications (CARMA) Streets as part of its CARMA research and development program; however, successfully implementing key CDA applications (such as eco-approach/departure, signalized optimization, and vulnerable road user protection) requires robust, low-latency sensor fusion with high-fidelity, quantified accuracy computations and communications from an infrastructure-based cooperative perception computational engine. The National Renewable Energy Laboratory's cooperative perception project fills this gap.

This initiative combines sensor fusion and high-performance computing to create an open-source, referential cooperative perception engine that can be utilized through the CARMA Streets framework. Even though the CARMA Streets framework anticipated an active collaborative infrastructure element (such as a traffic controller communicating signal phase and timing), a full infrastructure-based cooperative perception computational engine embeds the highest-quality common digital representation shared with all roadway participants, including pedestrians, micromobility, and transit users, extending the benefits of advanced technology to the most vulnerable road users. The ability to equitably address the needs of underrepresented modes and road users is a second critical gap addressed by infrastructure-based cooperative perception. In collaboration with CARMA and other industry and jurisdictional partners, the National Renewable Energy Laboratory has developed a concept of operations for the cooperative perception initiative as outlined in this publication—an architecture that can support multiple real-world applications.

This paper lays out the architecture and elements of next-generation intelligent roadway infrastructure, the necessary physical elements, and the logical functional block diagrams, and it also begins to address critical data and communications issues. Our objective is not only to document this system-level architecture for ongoing research and development efforts, but also to share it broadly within the CDA and traffic engineering circles as a proposed systems model with standardized open interface protocols for next-generation traffic control that is extensible to many other applications. The paper is organized as follows:

- Literature review as it relates to infrastructure intelligence and the need and methods for creating a full state-space perception model of all agents in a roadway domain

- Overarching system functional architecture of inputs, analysis, and outputs to enable applications
- Data interface specifications to provide standards-based application development
- Other considerations with respect to resilient, fail-safe, cybersecure, and equitable intelligent roadway infrastructure architecture.

LITERATURE REVIEW

The literature review covers two perspectives: (1) the motivation and evidence for the need for an infrastructure-based cooperative perception engine, both from the perspective of the CDA initiative and to enable a host of other infrastructure-centric applications; and (2) supporting methods and tools for enabling the cooperative perception engine.

The Role of Infrastructure-Based Cooperative Perception

From a safety perspective, intersections (and specifically signalized intersections) are disproportionally represented spatially in traffic statistics. Based on 2009 traffic safety data, of all crashes, 40% occur at intersections, and 21% are at signalized intersections. For fatal and injury crashes, intersections are even more disproportionally represented, with 46% occurring at intersections, and 24% specifically at signalized intersections (FHWA 2014).

Transportation safety has been extensively studied with respect to societal economic impacts by transportation agencies and policy officials. The embodied energy impact of transportation safety is also studied by developing an energy equivalence of safety framework (Zhong 2020; Zhu 2019) to provide a holistic view of the long-term energy and fuel consequences of motor vehicle crashes, incorporating both induced congestion and impacts from lost human productivity resulting from injury and fatal accidents, and the energy content resulting from all consequences and activities from a crash. The method utilizes a ratio of gross domestic product to national energy consumed in a framework that bridges the gap between safety and energy, leveraging extensive studies of the economic impact of motor vehicle crashes. The energy costs per fatal, injury, and property-damage-only crashes in 2017 were found to be approximately 200,000, 4,400, and 440 gasoline gallons equivalent, respectively, which are significantly greater than impacts from induced congestion alone. (Zhong 2020) This indicates the energy improvement potential of safety-oriented, infrastructure-based cooperative perception technologies.

A 2020 workshop on infrastructure spatial sensing at intersections hosted by the Transportation Research Board invited speakers to present their work covering various aspects of infrastructure-based sensing (e.g., safety, efficiency, automation). Infrastructure-based spatial sensing and perception technology encompasses the deployment of advanced spatial sensors that can detect and track all objects in a field of view—for example, mounting light detection and ranging (LiDAR) sensors, radar, and video imaging at an intersection and fusing the data to produce a 3D dynamic operational awareness. The workshop touched on the wide range of possible benefits from infrastructure-based spatial sensing (Zhu 2020). Topics included intersection collision avoidance resulting from dilemma zone monitoring, traffic and pedestrian applications from LiDAR deployment at intersections, estimating the energy benefits of crash avoidance, deconflicting roadway intersections for automated vehicle "interlocking," advancements in video detection for traffic applications at intersections, and leveraging sensor-based detection for signal operations in Colorado Springs, Colorado.

Although vehicle automation promises significant potential for societal benefits, vehicle perception is limited largely due to small fields of view, occlusion, and low measurement density. However, Arnold et al. (2020) proposes that multimodal sensing in place of sensing isolated to a single point of view can cost-effectively address these issues. One such manner of creating this multimodal perception is through placing sensors on existing infrastructure. In 2019, Seebacher et al. demonstrated the preliminary benefits of using infrastructure perception to augment autonomous vehicle perception in Austrian motorways. They found the best results for high velocity and low traffic volume on the highways, indicating the need for further development of sensor fusion techniques for high traffic density areas like intersections.

Methods and Tools

With increases in the sophistication of sensors and the trend for sensor vendors to preprocess the raw sensor data and return object detection information, the cooperative perception engine will focus on two primary steps: associating objects detected by the sensors with new or previously tracked objects and fusing the sensor outputs for a maximal accuracy state estimation.

The object association step is an example of a data association problem. Data association can be tackled in many ways. Algorithms like k-means cluster data points into groups of similar points that are associated with a tracked object, whereas other algorithms like the nearest neighbor method compute distances in state-space and use this distance to determine which tracked object the data point is closest to, or if the data point is far enough away to be a new object. More sophisticated data association algorithms like the nearest neighbor standard filter use a Bayesian formalism to compare the proximity of the probability density functions (PDF) of the predicted object location from the state-space motion model and the PDF of the observed location from sensors. More sophisticated methods like the probabilistic data association filter (Bar-Shalom 1975) and joint probabilistic data association filter enable detailed handling of the underlying uncertainties in the state-space framework. All these techniques are already being successfully applied to the traffic context (Yuan 2015; Yuan 2017).

Once sensor observations are attributed to existing or new objects, state estimation via Bayesian filtering processes allows for probabilistic merging of the predicted state PDF and the noisy sensor observations. These processes require a state-space model structure, which consists of two stochastic models: the motion model and the measurement model. The measurement model relates the measurement from a sensor and to the object state-space, whereas the motion model is used to model the physical movement of an object in the intersection from one time step to a future time step. Importantly, both models are stochastic and depend on random variables corresponding to the sensor measurement error and dynamics model error, respectively. Depending on the nature of the two models, various Bayesian filters are applicable:

- Kalman filter: Requires linear motion and observation models with additive Gaussian errors (Peng et al. 2009)
- Extended Kalman filter: Can handle minor nonlinearities in either or both models by linearizing the models about the current estimate; assumes additive Gaussian errors (Wang 2007)
- Sampling-based filters (unscented Kalman filter, particle filter): Uses sparse sampling or Monte Carlo sampling methods to represent the state-space PDF; allows these filters to handle strong nonlinearities with non-Gaussian errors and sparse data, but they are much more computationally expensive (Li 2003).

Two simulation environments, the Car Learning to Act (CARLA) vehicle simulator (Dosovitskiy 2017) and the Simulation of Urban MObility (SUMO) (Lopez 2018), are leveraged to test out different motion models and their corresponding filtering algorithms to determine the highest-accuracy digital representations of an intersection.

Architectures to provide fusion of sensor data from infrastructure and vehicle sensors have been described (Gabb 2019) using similar methods, but are primarily inspired by augmenting vehicle-based sensors for automated vehicle (AV) and connected vehicle (CV) applications somewhat exclusively, as opposed to the creation of an independent infrastructure agent as an authoritative publisher of the state-space within its geo-spatial domain, and targeted toward serving infrastructure-based applications, and augmenting AV and CV as secondary.

FUNCTIONAL SYSTEM ARCHITECTURE

The cooperative perception engine will leverage available data from both infrastructure-based sensors (such as LiDAR, radar, and video processing) and cooperatively shared information such as from CAVs and connected vehicle data sources. Although advanced signal control, eco-approach and departure, and enhanced safety are natural applications of the highly accurate state-space model, the vision is for the cooperative perception engine to support a wide variety and growing number of applications in the roadway infrastructure and smart city space. Imagined applications include safety-affirmative signaling, a needed function for safe passage of automated vehicle fleets for public mobility. Also, the complete digital state-space representation would provide needed input for both curb management and the volume and state of charge of electric vehicles to use as inputs in grid and building energy integration (not shown in Figure 1).

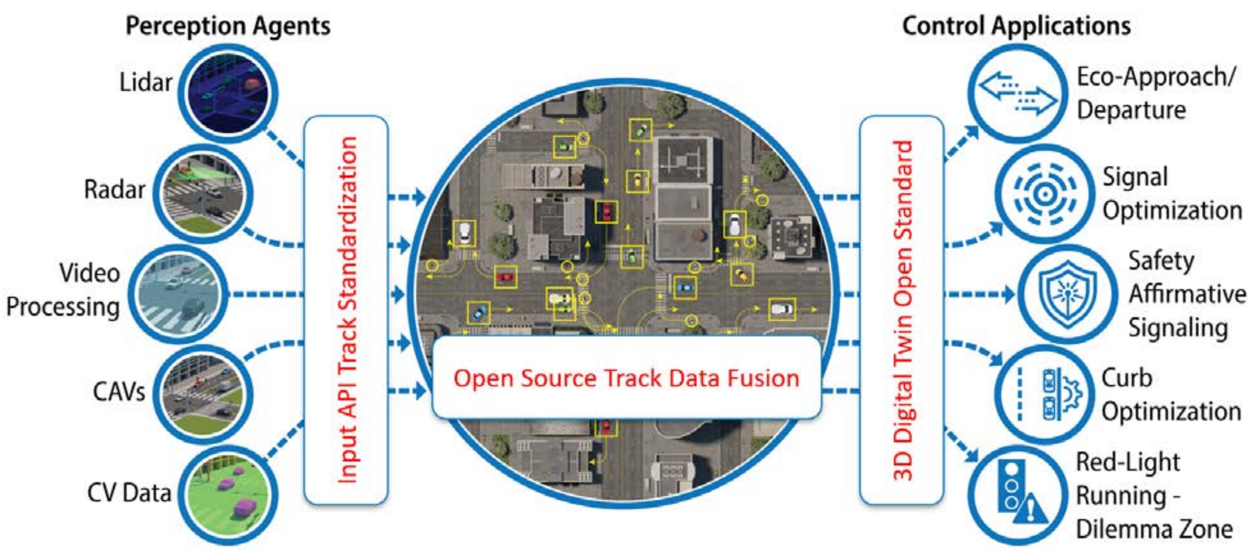

Figure 1. Conceptual function of the cooperative perception engine.

The central image in Figure 1 represents the fusion engine taking in data from available sources and creating the highest-accuracy digital representation of the physical space and all moving objects. The functional building blocks of the fusion engine are further illustrated in Figure 2. Data from each of the information sources or sensors arrive at the fusion engine for preprocessing, in which each data source is assigned a sensor model based on type. For example, radar has good depth (radial) accuracy, but typically less accurate angular precision. LiDAR provides points clouds with good spatial precision, but poor object identification capability.

Video image processing provides stronger object identification (e.g., vehicle, pedestrian, bicycle, heavy truck) but less precise spatial attributes. Similarly, a connected vehicle and its data have unique characteristics that are captured in a sensor model that allows the fusion engine to properly weight the data from each source. Preprocessing also typically transforms all sensor data into a common coordinate system and handles any time synchronization that may need to be accounted for.

With data in a common format and their fundamental accuracy characteristics understood, they are fed forward into the first stage of data fusion—data association (i.e., associating data from different sensors that represent common moving objects). This is informed by the previous state-space solution, in that the algorithm will anticipate data from sensors for any moving objects already in the domain of interest (sometimes referred to as the field of view for single-sensor applications). New objects at the periphery of the domain may be detected for the first time (and corresponding sensor data grouped), and existing objects moving out of the domain may not have any new associated sensor data assigned.

With sensor data grouped by object, this information is fed forward into the dynamic model processing for each object. In this step, the new data for each object are combined with the previous state-space solution from a previous time step, to optimally estimate both the location (latitude and longitude) and velocity of each object. In addition to this state-space vector solution, each object is assigned confidence (or correspondingly the estimated error) in each dimension of the state-space solution. This also includes identification of the object (e.g., vehicle, pedestrian, public transit bus). Note the information concerning the confidence of each parameter in the state-space solution is as critical for advanced applications as is the estimate of the state-space parameter itself. Without the corresponding confidence measures (inversely the estimation of error in the parameter), the quality of the state-space solution would be unknown.

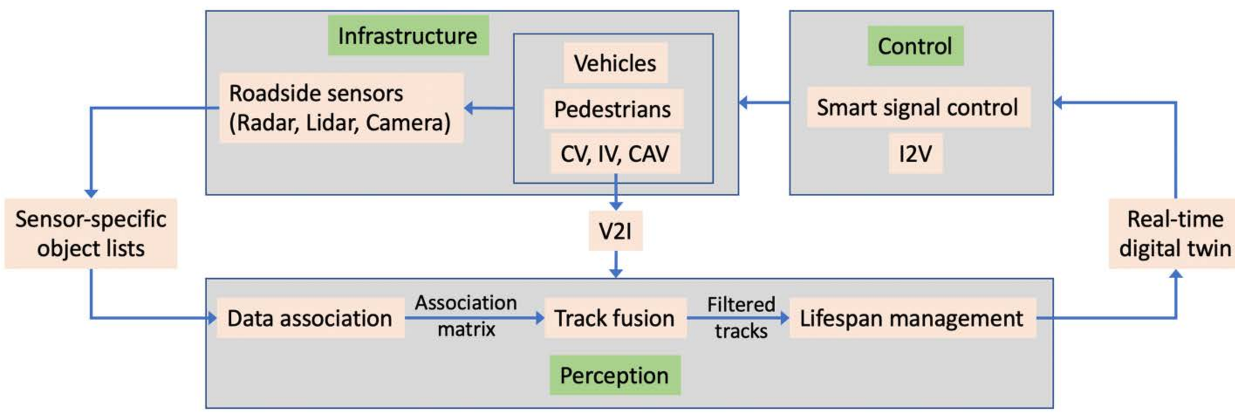

Figure 2. Functional flow of the data fusion engine of the infrastructure-based cooperative perception engine.

The digital state-space model is transmitted as output from the cooperative perception engine to downstream applications or other CDA agents. Achieving a verified accurate state-space model is key to enabling new applications. Without verified accuracy, applications would be limited to "convenience" or "opportunistic" functions—i.e., functions that would not have adverse impacts on the system (particularly safety) if the input information were incorrect. Current traffic control practice is of this nature. If a presence detector (such as a loop detector) fails, traffic control may be inefficient, but it does not result in a potential safety hazard. An example of a safety-critical application requiring verifiable state-space accuracy is early termination of pedestrian crossing phase. Currently, pedestrian crossing clearance times—and thus minimum duration of the pedestrian crossing signal for traffic lights—are based on walking speeds of the slowest pedestrians, perhaps elderly or handicapped. If a signal controller senses that the pedestrian crossed much faster, the control decision to end the pedestrian signal early is safety-critical. If the information related to pedestrian clearance was in error, and the controller shortened the pedestrian crossing signal, it would create a hazard to the pedestrian, and potential liability to the traffic control authority should the hazard result in a crash. The following applications describe the criticality and specifications of the digital state-space model for a variety of applications, the advanced control that could be enabled, and the associated benefits.

Signal Optimization

Currently, other than within some central business districts, most traffic signals operate with the use of some form of detection that is used to actuate the traffic signal by registering the presence of demand for specific movements and determining when to terminate the green for movements currently being served. The information used to drive actuation consists of binary ("on/off") occupancy states of detection zones. Typical zones are 40 to 60 ft in length, residing near the stop line, or shorter length detectors placed upstream. This offers a very limited picture of the overall traffic situation, as illustrated in Figure 3. In this example, three different vehicles pass through a section of roadway, but a long detection zone only reports a single period of vehicle presence in the zone, so the signal system has no way of knowing how many vehicles were actually present or what they were doing. Signal control methods have had to operate within these constraints for nearly 100 years. Although innovations have made the most of the data, entirely new modes of operation could potentially be enabled by the introduction of positions and speeds of vehicles and other roadway users.

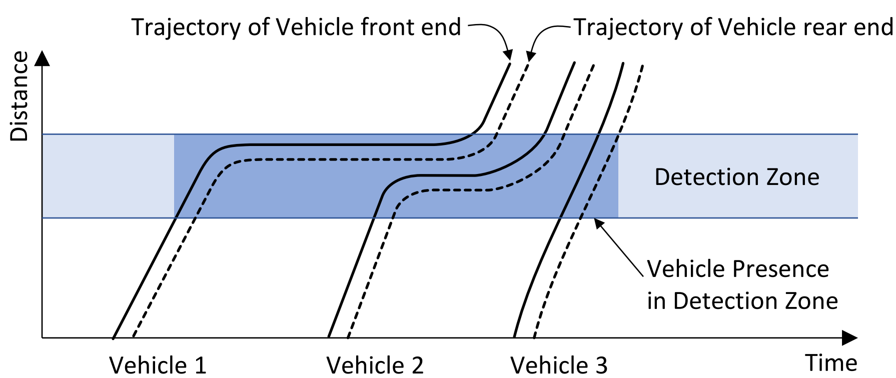

Figure 3. Information from existing detectors compared with vehicle trajectories.

For example, the actuation process for terminating green tries to determine when the flow rate drops below a certain level as a trigger for ending green. This indicates that standing queues have most likely cleared. However, the occupancy state of a 40- to 60-ft zone does not permit the measurement of flow rates, so the gap between vehicles is measured instead; whenever one vehicle leaves the detector, a "passage time" must expire. If no other vehicle arrives before the passage time elapses, then the green will be terminated. If the actual positions between vehicles could be recorded, the green could be terminated immediately after the last vehicle before the gap passes the point where it is unlikely to stop if the signal turns yellow, saving several seconds per cycle. Small inefficiencies like this can have enormous cumulative improvement potentials in a device that operates continuously.

Eco-Approach and Departure

In addition to adjusting signal operation to better respond to traffic, the same stream of information can be used to adjust traffic to better respond to the signal control. Essentially, if a vehicle application knows when a downstream red signal is likely to turn green, the application can advise the driver to proceed at a certain speed that would allow it to arrive at the signal when it is green, rather than stopping on red, which would save energy and reduce emissions by avoiding unnecessary deceleration and acceleration. This is one use of signal phase and timing messages broadcast by roadside units in a connected vehicle ecosystem. However, these data do not convey the presence of queues at the intersection, and unless there are connected vehicles in the queue, the signal is unlikely to know the length of a queue (existing detection cannot generally identify the queue length in real time). Improved intersection perception would bridge this information gap by allowing the direct measurement of queue lengths.

Safety-Affirmative Traffic Signaling

Early demonstrations of automated mobility solutions revealed excessive hazard exposure at intersections. Some early AV technology pilots expressly deployed proprietary sensing equipment at key intersections, while most either avoided complex intersections, particularly those with left hand turns that present a significant safety hazard, or required manual control through the intersection. These concerns led to safety-affirmative signaling control concepts required of next generation advanced traffic signaling along with the organizational framework of an appropriate authority having jurisdiction to assess and mitigate hazards. Such an approach will efficiently allow for multiple fleets of AVs to provide public mobility services, operating within an "automated mobility district" (Lott 2021a). The next generation of advanced traffic control will take on the role of safety management arbiter of AV moments through a complex roadway intersection, as has been done in communication-based train control providing guideway junction "interlocking" for the last 50 years in automated guideway transit systems (Lott 2021b), with each vehicle given permission to proceed through the intersection only when a safe travel path is ensured.

Without safety-affirmative signaling, AVs approaching complex roadway intersections may reduce speed, creating a safety hazard for other human-operated vehicles desiring to travel at higher speeds. Speed volatility resulting from human-operated vehicles as they maneuver around the slow-moving AVs will increase the risk of crashes and degrade safety for other more vulnerable roadway users. As such, a fundamental objective of safety affirmative signalizing will be to increase the operating speeds of AVs to match the general traffic flow conditions. Furthermore, the cooperative perception combined with CDA *intent sharing* messages will enhance the accuracy of the cooperative perception model's forecast of future traffic state. *Prescriptive* CDA messages may then be used to expressly guide an AV's trajectory through complex roadway intersections, with each individual AV receiving instructions on adjustments to its operating speed (higher or lower) as well as lane changes to more safely de-conflict its path from other objects.

DATA INTERFACE STANDARDIZATION FOR INFRASTRUCTURE APPLICATIONS

Operationally, the cooperative perception stack for infrastructure differs structurally from data fusion architectures for vehicles in that the solution to the state space, the "digital twin" is shared outwardly to enable applications beyond its domain. As such, any uncertainties, and confidence in the state-space solution must be clearly communicated such that downstream applications rightfully apply the information, knowing the error bounds and confidence in the digital twin.

Input data from sensors also require standardized data interfaces to create a nonproprietary processing chain from input to output. This is illustrated in Figure 1 above, emphasizing the need for standards-based interface design.

Standard input application programming interfaces (API) between any physical sensor and the cooperative perception engine will allow not only for competitive multivendor procurement, but also enable a stable referential open-source track data fusion engine, rather than having to develop custom input/output for every vendor and model of sensor. This architecture assumes that each sensor will perform initial processing of raw data to object lists, and then convey object list and associated attributes to the cooperative perception engine. In addition to the identification of objects, probabilistic measures that track certainty and precision are also highly desirable to seed the cooperative perception track fusion engine with appropriate parameter weighting, though such information is not absolutely mandatory as such information can be derived over time with calibration. Certainty and precision information from existing vendor sensor APIs is typically absent.

At the output of the cooperative perception is a 3D representation of the intersection domain containing the location and trajectory of every object in the field of view. The fundamental output can be envisioned as a fused object list conveyed in a 3D coordinate reference system. It is essential that this representation convey probability measures of accuracy and certainty, as well as health and calibration information of all input sensors. However, few downstream applications will require this full digital twin output in its most granular, and data-intensive state. For example, traffic controllers will require knowledge of vehicle queues, positions, and trajectories, at time sequences much less frequent than the native digital twin output. In lieu of communicating object lists expressed in a referential 3D coordinate system, the state of the roadway and its vehicles need to be associated with approaches and lanes. A traffic controller will also need assurance of the accuracy of the overall digital twin, but not at the level of individual vehicle error characteristics. Similarly, other applications may require only a subset of the full digital twin, and at frequencies and spatial granularities significantly less than the native digital twin output. As such, multiple standard digital interfaces that provide required data to various application genres are needed both to prompt independent application development (so developers of each application need not recreate that entire cooperative perception software stack and associated APIs), and to manage communication channels and bandwidth appropriately.

The intent of the standardization is not to limit innovation or competition, but rather the opposite. The diagram in Figure 1 calls out an open-source referential track fusion algorithm. However, a commercial algorithm that provides track fusion from standardized input to standardized output—perhaps with added fidelity, performance, or functionality—may be used. Similarly, sensor performance in terms of greater accuracy, range, or granularity are accommodated, with the improved specifications communicated as part of the API standard, and subsequent improvements in the resulting digital twin.

DISCUSSION AND CONCLUSION

The system architecture for next-generation traffic control utilizing infrastructure-based cooperative perception presented in this paper presents a feasible evolution to incorporate trajectory-aware sensing into intelligent roadway infrastructure, and in so doing, the infrastructure becomes an active and vital agent in CDA. This is critical from a number of vantage points. The sensor limitations from any one platform, or any one physical location (be it a vehicle or mounting position on roadside furniture), is incapable of providing all the spatial data necessary for safety management and roadway navigation in a robust way. By fusing together individual sensor, and incorporating CAV data as available, from multiple vantage points, the resulting digital state-space model not only becomes more accurate but also provides the capability to discern and measure the accuracy of the composite picture, enabling safety-critical applications and providing protection against single points of failure, cybersecurity risks, and bad actors in the system. Furthermore, this approach allows for nearer term benefits in an equitable fashion at the 300,000+ signalized roadway intersections in contrast to the nearly 300 million U.S. vehicle fleet. Signalized intersections are disproportionally represented with respect to crashes and excessive energy use in the United States. Also, the benefits that accrue from infrastructure-based perception, and subsequent control applications are more extensible to all roadway users than a vehicle centric CAV-only deployment.

In summary, this intelligent roadway infrastructure, with emphasis on an authoritative infrastructure-based cooperative perception computational engine (as opposed to vehicle-based), complements other efforts by providing a digital state-space representation of all objects in the roadway domain to a measurable degree of accuracy. This intelligent roadway infrastructure framework is also an active cooperative driving automation (CDA) agent, and is targeted as a referential open-source platform within the CARMA Streets model hosted by the Federal Highway Administration. The intelligent roadway infrastructure cooperative perception fusion engine utilizes both advanced infrastructure-based sensing and available CAV information to provide a complete state-space digital representation to support a wide range of applications, including advanced traffic signal control and other cooperative applications that address safety, equity, fuel efficiency, and greenhouse gas emissions reduction. This framework allows multiple application developers to innovate without the need to redevelop a cooperative perception software stack, and to do so based on measurable confidence in the digital twin data representation of the intersection. The next generation of traffic signal control will not only need to coordinate right-of-way (red, yellow, green signaling), but also be the safety management arbiter of moments in a multimodal environment that includes vulnerable road users (pedestrians and micromobility) along with AVs and traditional vehicles. The intelligent roadway infrastructure cooperative perception approach more equitably distributes advanced technology benefits to all road users, and not just those with capacity to own an operative advanced technology vehicle.

ACKNOWLEDGMENTS

The authors would like to thank our strong collaborators and partners, including the City of Colorado Springs, Iowa State University, University of North Carolina, Econolite, Quanergy, and others. This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. This work was supported by the Laboratory Directed

Research and Development (LDRD) Program at NREL. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

REFERENCES

Arnold, Eduardo, et al. "Cooperative perception for 3D object detection in driving scenarios using infrastructure sensors." *IEEE Transactions on Intelligent Transportation Systems* (2020).

Bar-Shalom, Yaakov, and Edison Tse. "Tracking in a cluttered environment with probabilistic data association." *Automatica* 11.5 (1975): 451-460.

Dosovitskiy, Alexey, et al. "CARLA: An open urban driving simulator." *Conference on robot learning*. PMLR, (2017).

Federal Highway Administration (FHWA). *Signalized Intersections Informational Guide, Second Edition.* Publication No. FHWA-SA-13-027, July (2013) https://safety.fhwa.dot.gov/intersection/signal/fhwasa13027.pdf.

Gabb, M., H. Digel, T. Müller and R. Henn. "Infrastructure-supported Perception and Track-level Fusion using Edge Computing," *2019 IEEE Intelligent Vehicles Symposium (IV)*, (2019), pp. 1739-1745, doi: 10.1109/IVS.2019.8813886.

Lott, J.S., and S. E. Young. "A Safety and Management Framework to Enable Automated Mobility Districts in Urban Areas." Proceedings of the ASCE 2021 International Conference on Transportation and Development (2021a).

Lott, J. Sam, Stanley Young, and Lei Zhu. *Safe Operations at Roadway Junctions-Design Principles from Automated Guideway Transit*. SAE Technical Paper No. 2021-01-1004, (2021b).

Lopez, P.A., et al. "Microscopic Traffic Simulation using SUMO," *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, (2018), pp. 2575-2582, doi: 10.1109/ITSC.2018.8569938.

Li, Rong, and V.P. Jilkov. *Survey of maneuvering target tracking*. Part I. Dynamic Models. IEEE Transactions on Aerospace and Electronic Systems, 39(4):1333–1364, (2003).

Peng, Depin, Xingquan Zuo, Jianping Wu, Chunlu Wang, and Tianle Zhang, "A Kalman Filter based information fusion method for traffic speed estimation." *2009 2nd International Conference on Power Electronics and Intelligent Transportation System (PEITS)*, (2009), pp. 374-379, doi: 10.1109/PEITS.2009.5406993.

SAE International. *Taxonomy and Definitions for Terms Related to Cooperative Driving Automation for On-Road Motor Vehicles*. SAE International, Warrendale, PA, (2020). https://www.sae.org/standards/content/j3216_202005/.

Seebacher, S., et al. “Infrastructure data fusion for validation and future enhancements of autonomous vehicles' perception on Austrian motorways,” *2019 IEEE International Conference on Connected Vehicles and Expo (ICCVE)*, (2019), pp. 1-7, doi: 10.1109/ICCVE45908.2019.8965142.

Wang, Yibing, Markos Papageorgiou, Albert Messmer, (2007) “Real-Time Freeway Traffic State Estimation Based on Extended Kalman Filter: A Case Study.” *Transportation Science* 41(2):167-181. https://pubsonline.informs.org/doi/10.1287/trsc.1070.0194.

Yuan, T., et al., “Object Matching for Inter-Vehicle Communication Systems—An IMM-Based Track Association Approach With Sequential Multiple Hypothesis Test,” in *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 12, pp. 3501-3512, Dec. (2017), doi: 10.1109/TITS.2017.2723894.

Yuan, T. et al., “Track-to-track association for object matching in an inter-vehicle communication system,” Proc. SPIE 9596, Signal and Data Processing of Small Targets 2015, 959609, September (2015).

Zhong, Zijia, Lei Zhu, and Stanley Young. “Approximation Framework of Embodied Energy of Safety: Insights and Analysis.” *Energies*, (2020). 13 (16): 4230.

Zhu, Lei, Stanley E. Young, and Christopher M. Day. *Exploring First-Order Approximation of Energy Equivalence of Safety at Intersections*. National Renewable Energy Laboratory, Golden, CO, (2019). https://www.nrel.gov/docs/fy19osti/73405.pdf.

Zhu, Lei. *2020 Transportation Research Board Workshop: Infrastructure Spatial Sensing at Intersections: The Potential to Enhance Safety and Enable Automation*. National Renewable Energy Laboratory, Golden, CO, (2020). https://www.nrel.gov/docs/fy20osti/76090.pdf.