Sensor and Data Fusion for Intelligent Transportation Systems

Lawrence A. Klein

SPIE PRESS
Bellingham, Washington USA

Library of Congress Cataloging-in-Publication Data

Names: Klein, Lawrence A., author.
Title: Sensor and data fusion for intelligent transportation systems / Lawrence A. Klein.
Description: Bellingham, Washington, USA : SPIE Press, [2019] | Includes bibliographical references.
Identifiers: LCCN 2019001144 | ISBN 9781510627642 (softcover) | ISBN 1510627642 (softcover) | ISBN 9781510627659 (PDF) | ISBN 1510627650 (PDF) | ISBN 9781510627666 (ePub) | ISBN 1510627669 (ePub) | ISBN 9781510627673 (Kindle) | ISBN 1510627677 (Kindle)
Subjects: LCSH: Intelligent transportation systems. | Multisensor data fusion. | Motor vehicles–Automatic control. | Traffic congestion. | Algorithms.
Classification: LCC TE228.3 .K537 2019 | DDC 388.3/12–dc23
LC record available at https://lccn.loc.gov/2019001144

Published by

SPIE
P.O. Box 10
Bellingham, Washington 98227-0010 USA
Phone: +1 360.676.3290
Fax: +1 360.647.1445
Email: books@spie.org
Web: http://spie.org

Printed in the United States of America.
First Printing.
For updates to this book, visit http://spie.org and type "PM305" in the search field.

Table of Contents

Preface

Sensor and Data Fusion for Intelligent Transportation Systems was prepared to give undergraduate and graduate students, researchers, and traffic management professionals a multidisciplinary description of sensor and data fusion and the benefits it brings to the transportation community, especially for intelligent transportation systems (ITS). Although sensor and data fusion processes are introduced through the lens of the military-oriented U.S. Joint Directors of Laboratories (JDL) data fusion model and the Data Fusion Information Group (DFIG) enhancements, analogous language for its application to traffic management is afforded at each data fusion processing level. This model was selected because of its ability to expose all of the generally acknowledged facets of data fusion.

The book focuses on data fusion functions and data processing algorithms. Sensor fusion architectures, although discussed, are not treated in as much detail. The discussions explore techniques that enhance the interpretation of information gathered from a diverse mixture of sensors and other data sources (for instance, floating cars, connected and cooperative vehicles, self-driving vehicles, cellular telephone messages, mobile device locations accessible through Bluetooth® communications, location data from global navigation satellite system devices, and automatic license plate and toll-tag readers) that help characterize the traffic environment. Its often-complex nature is exacerbated by a mix of different types of vehicles, changes in traffic flow characteristics, appearance of unexpected objects such as pedestrians darting across a roadway, inclement weather, vehicles changing lanes, and roadside structures or poor visibility conditions that interfere with the normal observation of traffic patterns and the gathering of needed data. Bayesian inference, Dempster–Shafer evidential reasoning, and Kalman filtering data fusion algorithms that combine data from infrastructure-based sensors and other sources are described in detail with illustrative examples of their application to ITS. Other algorithms applicable to fusion of traffic data are also discussed, including parametric templates, artificial neural networks, cluster algorithms, voting methods, knowledge-based expert systems, and fuzzy logic.

The manuscript is based, in part, on short courses and semester-length courses taught by the author at several universities and conferences for many years. The impetus for the book came from a desire to prepare a treatment of selected subjects pertaining to sensor and data fusion for novice and more experienced practitioners from universities; transportation institutes; and local, regional, state, and multi-state or multi-national agencies and consulting companies. The intent is to provide an understanding of the functions of multilevel data fusion processes suitable for ITS, knowledge of the kinds of information and parameters required to implement particular data fusion algorithms, and recognition of pertinent issues that may affect the operation of a data fusion system. Personnel engaged in research may be particularly interested in the sections that deal with ongoing needs in data fusion research.

Readers desiring additional information concerning sensor and data fusion are referred to *Sensor and Data Fusion: A Tool for Information Assessment and Decision Making*, Second Edition, SPIE Press, Bellingham, Washington (2012) [doi: 10.1117/3.928035], which is written by the author. Further material regarding freeway traffic management centers, ITS, sensors and other data acquisition devices, data requirements, sensor testing, automated vehicles and vehicle systems, connected vehicle programs in the U.S. and elsewhere, systems engineering, and National Intelligent Transportation System Architectures is found in *ITS Sensors and Architectures for Traffic Management and Connected Vehicles*, Taylor and Francis, Boca Raton (2018), which is also written by the author.

I wish to thank Maijian Qian for assisting in verifying several of the equations that appear in Chapter 5; Dara Burrows, Senior Editor SPIE Press, and Tim Lamkins, SPIE Press Manager, who facilitated the editing and publication of the book and afforded me their usual courtesy; and finally, the reviewers for providing valuable feedback that greatly enhanced the material that appears here.

Lawrence A. Klein
June 2019

Chapter 1
Introduction

Transportation systems require reliable and accurate data for monitoring and managing operations that maximize the safety and efficiency of the roadway network. The proliferation of connected (or cooperative) and automated vehicles, which are capable of providing data regarding vehicle speeds and volumes, congestion and collision avoidance maneuvers, and inclement weather and other potentially dangerous road conditions, create both a major opportunity and a major challenge for public agencies and private companies that support intelligent transportation systems (ITS). The opportunity arises from the diversity of data available to traffic and transportation management personnel, while the challenge arises from the need to gather, process, analyze, and store the vast quantities of data that will flow to the traffic management agencies, especially as big data become available. These emerging and fertile data sources, along with the spread of Bluetooth® and IP-based (cellular and Wi-Fi) communications technologies, will increase the travelers' proclivity for accurate road traffic information.

Predictions by Gettman, Hales, Voss, and Tumati[1] for the connected city of 2021 forecast growth of 400% in connected travelers, 300% in connected vehicles, and 25% in connected infrastructure. As the connected city continues to grow, the amount of data stored within the system will need to grow from terabytes to petabytes in 10 years—explaining the importance of scalability (i.e., a system's ability to increase and/or decrease storage resources as needed). If all of this emerging data (such as from connected travelers, connected vehicles, and new types of connected infrastructure) related to traffic operations are stored, the cumulative storage of a typical traffic management agency could be in the many thousands of terabytes by 2026.[2] A challenge is to develop data fusion methods that can cope with these vast quantities of data.

Sensors that monitor traffic flow at a given point are often ineffective in supplying the data required by modern transportation management systems. For example, origin–destination (OD) pair data needed for planning purposes and vehicle density studies are not readily available from point sensors.

Global positioning system (GPS) and other global navigation satellite systems' location devices, cell phone tracking through media access control (MAC) address readers, probe vehicles, automatic license-plate readers (ALPRs), toll-tag readers, and trucking industry transponders are increasingly supplementing the information provided by conventional traffic flow sensors.

When multisource data are complementary in nature and an output from one data source does not imply nor is implied by an output from another (i.e., the data sources are conditionally independent), data fusion can provide operational benefits to traffic management personnel. These benefits include utilizing the data to obtain a more accurate description of the road and traffic conditions by decreasing the uncertainty present in the individual sources of data, extending spatial and temporal coverage areas, accessing data not usually available from conventional roadside sensors such as OD pairs and connected vehicle data (for instance, weather and road conditions from in-vehicle sensors that automatically turn on windscreen wipers or initiate traction control measures), and providing the ability to function in inclement weather. The fusion of data and information from multiple sources is consequently a well-adapted answer to the operational needs of traffic management centers and traffic information providers, allowing them to achieve their goals more effectively.[3] Data fusion opportunities also exist for processing archived data such as traffic volumes by time-of-day, day-of-week, month, or season, and recurring special events. This offline information, together with real-time sensor data, is often useful in predicting traffic trends and forecasting the need for new roadways or other travel modalities.

Sensor and Data Fusion for Intelligent Transportation Systems introduces readers to the roles of several data fusion processes as defined by the Joint Directors of Laboratories (JDL) data fusion model and the Data Fusion Information Group (DFIG) enhancements, data fusion algorithms, and noteworthy applications of data fusion to ITS. Additionally, the monograph offers detailed descriptions of three of the widely applied data fusion techniques and their relevance to ITS (namely, Bayesian inference, Dempster–Shafer evidential reasoning, and Kalman filtering), and indicates directions for future research in the area of data fusion. The focus is on data fusion algorithms rather than on sensor and data fusion architectures, although the book does summarize factors that influence the selection of a fusion architecture and several architecture frameworks.

1.1 Applications to ITS

Effective design of an ITS requires a systems approach that incorporates sensors and other data-gathering devices and a variety of communications technologies into the concept exploration, architecture selection, hardware and data processing design, testing, operation, and performance evaluation of the fabricated system. Systems engineering helps assure that a proposed architecture will satisfy the goals and objectives of the project. Typical projects that benefit from this method include freeway designs that employ ramp metering, information dissemination to travelers, managed lanes, and active traffic management; arterials that utilize traffic adaptive signal control, signal priority for transit, freight, and emergency vehicles, and parking guidance; integrated corridors that coordinate and seek to optimize traffic flow on arterials and limited-access highways; road-weather systems; and the construction and staffing of transportation management centers, especially those that collocate personnel from different agencies.

Several data fusion algorithms are already prevalent in ITS applications. These include Bayesian inference, Dempster–Shafer evidential theory and some of its modifications, artificial neural networks, fuzzy logic, knowledge-based expert systems, and vehicle and pedestrian tracking based on the Kalman filter or extended Kalman filter (EKF), Monte Carlo techniques, and particle filters. The Bayesian and Dempster–Shafer approaches belong to the class of feature-based parametric algorithms. They directly map parametric data (e.g., features) into a declaration of identity. Physical models are not used. Artificial neural networks belong to the class of feature-based information theoretic techniques that transform or map parametric data into an identity declaration. No attempt is made to directly model the stochastic aspects of the observables. Fuzzy logic and knowledge-based expert systems are examples of cognitive-based approaches that attempt to emulate and automate the decision-making processes used by human analysts. The Kalman filter and its nonlinear-motion counterparts are examples of physical models since the kinematics of the objects being tracked are modeled. Physical models replicate object discriminators—in this case, position, velocity, and sometimes acceleration—that are easily observable or calculable.

ITS subsystems include advanced transportation management systems (ATMS), automatic incident detection (AID) (a subset of ATMS), advanced traveler information systems (ATIS), advanced driver assistance systems (ADAS), and commercial vehicle operations (CVO), all of which gather data and information from different sources. Data fusion techniques can therefore be developed to combine the data and obtain a better understanding of what the data represent, e.g., the types of objects or the situation (for instance, a roadway incident) giving rise to the data.

Incident detection algorithms that automatically detect incidents, accidents, and other road events requiring emergency responses have existed for more than three decades. Most of the algorithms rely on loop detector data. However, these algorithms exhibit mixed success. Interest in incident detection algorithms has renewed partly because of the availability of new sensors and data sources such as probe and connected vehicles and cellular telephone reporting. However, data fusion involving combinations of hard (i.e., sensor) and soft (i.e., voice) data is not without risks and is a subject of current research. Notwithstanding the preceding comment, AID is typical of the class of problems that can be enhanced by data fusion techniques.

ATIS employ a variety of automatic data collection techniques to assist in understanding traffic conditions and derive relevant indicators that support traveler guidance. Traveler information is often presented as travel time or, in the worst case, as a road closure or extended delay due to an incident. In this context, travel time is used as a measure of impedance (or cost) for route choice strategies. However, conventional traffic sensors that measure the prevailing traffic conditions on an urban road may be ineffective at this task because of a sparsity of sensors or their inoperability. The proliferation of alternative data and information sources (e.g., surveillance cameras, GPS, MAC address readers, ALPRs, connected and automated vehicles, and cell phone reports) provide data that supplement traditional sensor measurements. These complementary devices also have the potential to improve the accuracy of travel-time estimates. As a result, travel-time estimation becomes a candidate data fusion application.

Improving traveler safety is a primary function of ITS. The increased availability of ADAS and collision avoidance systems (CAS) is indicative of the growth in active safety devices that complement traditional passive ones such as seat belts and air bags. ADAS help alert a driver to traffic and other hazards surrounding the vehicle in pre-crash situations. Their input data frequently come from a variety of sensors including radar, lidar, ultrasonic, and video imaging sensors. Fusion is exploited to combine these data and alert the driver to potentially dangerous situations. CAS have the ability to notify not only drivers of potential incident situations, but also traffic management personnel who can take actions such as reducing speed limits on the affected highways or lanes, closing lanes, and notifying state highway police to implement a traffic break or initiate other accident-preventive procedures.

An important item in the field of transportation planning and control is OD estimation from counts undertaken on specific links of the transportation network. An estimation of a most likely OD matrix is then derived from the counts. Traffic flow forecasting is of increasing importance to traffic surveillance, facility management, and departments that plan for new roadways. Many traffic flow prediction schemes of the past were based on classic autoregressive models, especially time series techniques. Connected vehicle, GPS, MAC address, and ALPR data are now available to enhance the acquisition of OD pair information.

Modern transportation systems require accurate information concerning the position and orientation of vehicles to forecast congestion and incident precursors such as wrong-way drivers. This application has been abetted by the ubiquitous nature of GPS for location and navigation services. However, when satellite signals are blocked by tall buildings and trees, or are corrupted by electromagnetic interference or refraction as they propagate through the atmosphere, loss or degradation of the GPS signal occurs. In such situations, the estimation of position is degraded at best or, at worst, is impossible to obtain as the device is unable to acquire a signal. Inertial navigation systems (INS) that determine the location of a vehicle rely on dead reckoning and can be used to complement GPS data when the latter are degraded or unavailable. However, INS are subject to integration drift caused by the accumulation of small errors in the measurement of acceleration and angular velocity that manifest as larger errors in the position estimate. Data fusion offers a complementary approach that combines the benefits of the GPS and INS techniques, namely, calibration of the INS by the GPS signals when they are available, and position and angle updates provided by the INS when the GPS signal is blocked.

Another area closely related to ITS is autonomous or self-driving vehicles. The need for data fusion in these vehicles was noted by Raj Rajkumar, Professor of Electrical and Computer Engineering at Carnegie Mellon University, at a forum sponsored by the National Academy of Engineering in 2018.[4] Prof. Rajkumar remarked that some companies are not relying on GPS to provide location data to the autonomous vehicles they manufacture because there are many GPS-denied environments, such as under bridges, inside tunnels, in urban canyons, and in densely forested areas. Instead, these companies are working toward the identification of landmarks to locate the vehicle. However, certain places have very few landmarks. Rajkumar observed that "There are pros and cons to every approach you can take. My own philosophy is that you need to fuse together as many things as you can [to] get the best of all possible worlds."[4] Furthermore, the components of an autonomous vehicle can fail, so redundancy must be built into the vehicle. This presents another opportunity to incorporate data fusion into such a vehicle.

1.2 Data, Information, and Knowledge

Data, information, and knowledge are terms found throughout this book. Their relation to each other is illustrated in Fig. 1.1 in the form of a triangle whose foundation is the data that evolves into information and finally into knowledge through further processing, interpretation, and comprehension.[5] Data are the individual observations, measurements, and primitive messages from the lowest level of abstraction. Data are obtained from human communication, text

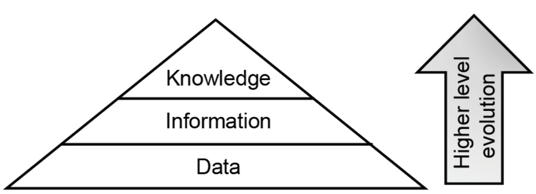

Figure 1.1 Evolution of data to information and knowledge.

messages, electronic queries, or devices that sense phenomena. Evidence consists of relevant data or specific elements of the overall data set.

Information is represented by organized sets of data. Organization may occur through sorting, classifying, and indexing and linking data to place data elements in relational context for subsequent searching and analysis.

Finally, knowledge, or foreknowledge (i.e., predictions or forecasts), evolves from information that is analyzed, understood, and explained. Once understood, knowledge provides a degree of comprehension of the static and dynamic relationships among data objects, the ability to model structures, and an understanding of past and future behavior of those objects.

1.3 Summary of Book Contents

This book contains five additional chapters. Chapter 2 includes several common definitions of sensor and data fusion, and presents an argument for its application to traffic management. The relevance of multisensor data fusion is due to the following factors: its value in combining and interpreting information gathered from a complex environment characterized by the presence of different types of vehicles; the often rapid changes in traffic flow characteristics; unexpected objects such as debris, pedestrians, or animals darting across a roadway; vehicles changing lanes; and roadside structures or weather effects that interfere with the normal observation of traffic patterns and the gathering of needed data. Chapter 2 describes the processes typically associated with data fusion and their benefits with the aid of the JDL data fusion model and the DFIG augmentations. Although this model was originally developed with military-oriented applications in mind, its concepts and approaches are shown to be directly applicable to traffic management.

The traffic management data fusion strategies discussed in the chapter utilize data from traffic flow sensors installed along the roadway; devices that access toll-tag and electronic screening and truck transponders, ALPRs, and Bluetooth-device MAC address readers used to track and re-identify vehicles; floating cars that provide a traffic management center emissions information in addition to normal traffic flow parameters linked to the vehicle's location via GPS; crowdsourcing applications and personal device monitoring that measure travel behavior of pedestrians and bicyclists; and, in the not too distant future, connected and cooperative vehicle data associated with vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and vehicle-to-pedestrian (V2P) applications. A summary of factors that influence the selection of a sensor and data fusion architecture and several architecture classification schemes are provided, although this is not the primary focus of the chapter. The emphasis is on data fusion algorithms and processes, and their applications to ITS and traffic management. Consequently, the discussions of the taxonomies and several of the algorithms used for object detection, classification, identification, and state estimation and tracking utilize transportation-related examples to illustrate their relevance. The higher-level data fusion processes of situation and impact assessment, performance refinement, and user refinement identify behavior patterns (in this case, traffic behavior), associate entities and events, predict future behaviors and their time frames, assess the situation generating the collected data, refine the fusion process, and address issues concerning human interpretation of the results of the data fusion process. The chapter concludes by describing a variety of sensor and data fusion applications to ITS and its subsystems.

Chapter 3 discusses the underlying principles of Bayesian inference, the assumptions typically associated with its application, and several examples that illustrate its relevance to traffic management. Common prerequisites for Bayesian analyses are the ability to: (1) compute or model the likelihood functions for each sensor or information source and object in the scenario of interest, (2) use a set of mutually exclusive hypotheses and calculate the *a priori* probabilities that the hypotheses are true, (3) discard past evidence or data once the posterior probability is calculated for the current period of interest, and (4) obtain data from conditionally independent sensors. Bayesian inference is widely employed for detection, classification, and identification of objects and events related to traffic management. It provides travel-time estimation, automatic incident detection, and decision support. The examples described in the chapter are identification of vehicle type with emission spectra data collected by a sensor, incident detection utilizing an influence diagram and a joint sensor report, truck classification using a multiple classifier system in which two of the processes are Bayesian, and maneuver-based trajectory prediction and criticality assessment by means of causal Bayesian networks.

Chapter 4 examines Dempster–Shafer evidential theory, a pseudo-probability–based data fusion classification algorithm. This method finds application when the sensors (or, more generally, the information sources) contributing data cannot associate a 100% probability of certainty to their output decisions. The sensors must function as independent sources of information concerning the presence of the objects or events of interest. Knowledge from multiple sensors about events (called propositions) is combined using Dempster's rule to find the intersection or conjunction of the propositions and their associated probability, which is expressed as probability mass, sometimes called a basic probability assignment. The chapter summarizes several modifications to the original Dempster–Shafer theory that have been proposed to accommodate situations where the information sources are in conflict. Perhaps the most difficult part of applying Dempster–Shafer theory in its original or modified forms is obtaining probability mass functions. Five methods for developing these functions are discussed. An example is given of the application of Dempster–Shafer reasoning to create a decision support system that enhances a traffic manager's understanding of the conditions that give rise to the collected data.

The Kalman filter described in Chapter 5 provides an optimal state estimate for linear systems as long as the dynamics of the tracked object and measurement noise are accurately modeled. The filter estimates an object's state (i.e., its position, velocity, and acceleration) at some future time, e.g., the predicted time of the next observation, and then updates that estimate using noisy measurements of the state. Process noise present in the kinematics of the object must also be modeled and of sufficient magnitude to keep the Kalman gain large enough to ensure that the tracker does not ignore the current measurement data and simply dead reckon the track based on past history. Moreover, the Kalman filter offers an estimate of the tracking error statistics through the state error covariance matrix. Also discussed are the α-β filter, which is applicable when the tracked object moves with constant velocity, an interacting multiple model to describe potential vehicle motion on a roadway, and the extended Kalman filter, which is applicable to nonlinear systems.

Chapter 6 reviews the state of the practice of data fusion in traffic management applications and the need for continued research in several areas that include reliability and quality of input data to the fusion system, security of data fusion systems, fusion of sensor or hard data with human-generated or soft data, fusion system performance assessment, the adaptive nature of measures of performance (MoPs), dependency of MoPs on the data fusion scenario, obtaining accurate ground truth data, use of commercial database management and operating systems tools, and designing for worst-case scenarios to properly specify processing and communications requirements. The chapter concludes with a summary of the information needed to apply Bayesian inference, Dempster–Shafer evidential theory, artificial neural networks, voting logic, fuzzy logic, and Kalman filtering data fusion algorithms to object detection, classification, identification, and state estimation.

References

1. D. Gettman, K. Hales, A. Voss, A. Toppen, and B. Tumati, "Integrating Emerging Data Sources into Operational Practice: State of the Practice Review," FHWA-JPO-16-424, Kimley Horn and Associates, Phoenix, Dec. (2016).
2. D. Gettman, A. Toppen, K. Hales, A. Voss, S. Engel, and D. El Azhari, "Integrating Emerging Data Sources into Operational Practice—Opportunities for Integration of Emerging Data for Traffic Management and TMCs," FHWAJPO-18-625, Kimley Horn and Associates, Phoenix, Nov (2017).
3. N.-E. El Faouzi and L. A. Klein, "Data Fusion in Intelligent Traffic and Transportation Engineering: Recent Advances and Challenges," in

Multisensor Data Fusion: From Algorithm and Architecture Design to Applications, H. Fourati, Ed., Chapter 32, CRC Press, Boca Raton, Florida (2015).
4. National Academy of Engineering, *Autonomy on Land and Sea and in the Air and Space: Proceedings of a Forum*, The National Academies Press, Washington, D.C., pp. 2–3 (2018) [doi: 10.17226/25168. Also available at http://nap.edu/25168. Accessed Aug. 7, 2018].
5. J. Roy, "Combining elements of information fusion and knowledge-based systems to support situation analysis," *Proc. SPIE* **6242**, 624202 (2006) [doi: 10.1117/12.663896].