

Sensor fusion as an enabling technology for safety-critical driver assistance systems

Richard Altendorfer, Sebastian Wirkert, Sascha Heinrichs-Bartscher

Driver Assistance Systems

TRW Automotive

Germany

ABSTRACT

Driver assistance systems are incorporating more and more advanced safety functions. As these functions have to react quickly and reliably in emergency situations with a false alarm rate close to zero a high integrity of the environmental perception is required. This elevated level of signal integrity can be achieved by data fusion, where the information of several, in general heterogeneous sensors is combined to obtain a better model of the environment in terms of accuracy, object integrity, object identity, etc. As an example, we demonstrate the power of sensor fusion by an automatic emergency brake (AEB) system whose environmental perception is based upon a video camera and a radar sensor. In particular we discuss the improvement of kinematic attributes such as object lateral distance as well as the object's confidence or probability of existence.

INTRODUCTION

Automotive driver assistance systems (DAS) such as adaptive cruise control (ACC) or lane departure warning (LDW) need to perceive the environment using exteroceptive sensors. As DAS become more sophisticated and move from comfort applications such as ACC to safety-critical applications such as automatic emergency braking (AEB), the requirements regarding perception as well as failure modes are becoming more stringent. An automatic emergency braking system can slow down a vehicle with a deceleration of up to $-1g = -9.8m/s^2$ in case of a perceived collision with another vehicle. This intrusive action is meant to minimize the impact of such a collision. An erroneous deployment of an automatic emergency braking maneuver, e. g. the automatic braking of a car with maximum deceleration going at top speed on an empty highway is considered unacceptable to the driver. In addition to the discomfort and annoyance of the driver, such erroneous deployment can be dangerous and even lethal to the occupants of the braking as well as the following vehicle in case of a rear-end collision. Hence qualifying conditions must be imposed such that erroneous deployments or false alarms are kept at an absolute minimum - a typical requirement is 1 false alarm per 1 million driven kilometers.

Those qualifying conditions can be applied at every component of the driver assistance system: sensors, perception, controller, and actuation, see fig. 1. In this paper we will focus on the estimation module of the perception component. There, those qualifying conditions can deal with the estimated relative dynamics of a driving situation, i. e. how the perceived objects move with respect to the own (ego) vehicle and whether any of those poses a threat to the ego vehicle. Common criteria are an estimated time-to-collision (TTC), see e. g. [1] or a collision probability see e. g. [2]. On the other hand, estimated objects need not be what they seem to be, e. g. an object could be classified by an identity estimation module as a truck whereas in reality this object corresponds to a passenger car. Even worse, an estimated object might correspond to a real object irrelevant to the vehicle threat assessment, e. g. a soda can, or it can even not correspond to any real object at all - a non-existing object based on spurious radar reflections for example.

Both problems, namely the correct estimation of relative dynamics and an estimation of the confidence or quality of the estimated object can be ameliorated if not completely solved by sensor fusion. This paper is organized as follows: in the first section the types and benefits of sensor fusion will be discussed in general as well as for the special case of DAS. In the next section two improvements due to

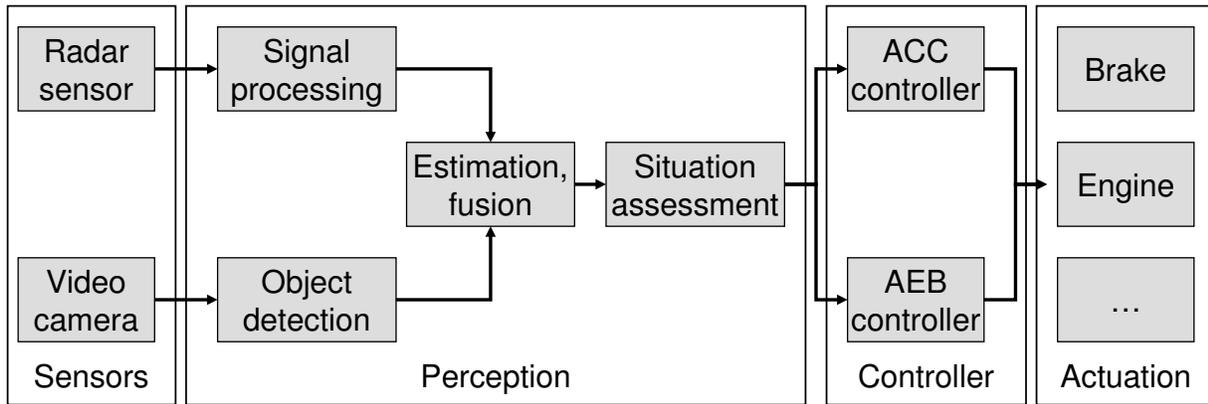


Figure 1: Simplified block diagram for a radar and video based ACC and AEB system. In the situation assessment module e. g. the ACC track is selected from the set of all tracks as well as the track for which a collision is most likely.

sensor fusion for an AEB system based upon a radar and a video sensor will be discussed in detail: the improvement in estimation of the relative dynamics (kinematic state estimation) and the improvement of a confidence measure for each track.

SENSOR FUSION

WHAT IS SENSOR FUSION? Sensor fusion has its origin in military applications where information from several sources is combined to obtain a better or more complete picture of a battle field or combat situation. Some of the pioneering work including an abstract classification of sensor fusion levels - the JDL model of sensor fusion [3] - was initiated by the US Department of Defense.

A commonly cited definition of information fusion was formulated in an editorial of the journal Information Fusion [4]: “Information fusion encompasses theory, techniques and tools conceived and employed for exploiting the synergy in the information acquired from multiple sources (sensor, databases, information gathered by human, etc.) such that the resulting decision or action is in some sense better (qualitatively or quantitatively, in terms of accuracy, robustness, etc.) than would be possible if any of these sources were used individually without such synergy exploitation.”

Sensor fusion as a subfield of information fusion is the combining of sensory data or data derived from sensory data such that the resulting information is in some sense better than would be possible when these sources were used individually. In particular, sensor fusion deals with the environmental estimation using multiple sensors, see e. g. [5].

BENEFITS AND TYPES OF SENSOR FUSION Sensor fusion offers the following general benefits:

- | | |
|----------------------|--|
| Robustness | redundancy in multiple sensors makes a system more robust in case of partial failure |
| Extended coverage | if the ranges of multiple sensors do not overlap or only partially overlap, the joint coverage is extended |
| Increased confidence | a measurement of one sensor is confirmed by measurements of other sensors covering the same domain |
| Increased accuracy | combining measurements from multiple sensors covering the same domain increases the accuracy of the measured quantities such as distance, speed, etc |

Three general types of sensor fusion can be distinguished:

| | |
|---------------|---|
| Complementary | Complementary sensor fusion combines sensory data of sensors whose sensor ranges or fields of view (FOV) do not overlap. An example is the fusion of a forward-looking radar with a backward-looking radar. Generally, fusing complementary data is easy, since the data from independent sensors can be appended to each other. |
| Competitive | Competitive sensor fusion combines sensory data of sensors whose sensor ranges are identical or have a large overlap and whose sensor data are of the same type e. g. angle. Competitive sensor fusion can be used for fault tolerance, e. g. an object is confirmed if it is seen by at least two out of three sensors, or for increased confidence in an object, e. g. if an object has been seen by all available sensors. |
| Cooperative | Cooperative sensor fusion combines sensory data of sensors whose sensor ranges have a large overlap in order to obtain new, additional information. Their sensor data need not necessarily be of the same type. This additional information can still be of the same type, e. g. angle, but with an increased accuracy that would not have been achievable with only one sensor; or it can be of a new type, e. g. in inertial navigation the pitch and roll angles can be estimated once speed measurements are available. |

More concretely, we have listed salient benefits of sensor fusion for automatic emergency brake (AEB). The sensor configuration under consideration consists of a forward-looking radar and in addition a video sensor. On the right-hand side the corresponding fusion type is indicated.

| | |
|--|--|
| <ul style="list-style-type: none"> • Fewer false alarms <li style="margin-left: 20px;">– Lower confidence for objects detected by <i>only one</i> sensor <li style="margin-left: 20px;">– More accurate estimation of <i>non-colliding</i> vehicle trajectories <ul style="list-style-type: none"> - Object geometry using video width - Better lateral distance | <p>fusion type</p> <p>competitive, cooperative</p> <p>complementary</p> <p>cooperative</p> |
| <ul style="list-style-type: none"> • More true alarms <li style="margin-left: 20px;">– Higher confidence for objects detected by <i>both</i> sensors <li style="margin-left: 20px;">– More accurate estimation of <i>colliding</i> vehicle trajectories <ul style="list-style-type: none"> - Object geometry using video width - Better lateral distance | <p>competitive, cooperative</p> <p>complementary</p> <p>cooperative</p> |

In the next section we will focus on improvements in lateral distance and in an object's confidence by sensor fusion.

SENSOR FUSION FOR SAFETY-CRITICAL DAS

IMPROVING KINEMATIC STATE ESTIMATION BY SENSOR FUSION

State estimation The kinematic estimation module which is situated between the signal processing/object detection and the situation assessment module as in fig. 1 computes an estimated state $\hat{\xi}$ on the basis of the current and all previous (noisy) sensor measurements z . This estimate should be close to the real, physical state ξ . Knowledge of the state is necessary as it is used by the DAS controller component. For a radar-based driver assistance system such as ACC, the state vector typically contains the relative Cartesian position, velocity, and acceleration, see fig. 2. If a scanning radar is used or video measurements are present, the vehicle's width w is commonly

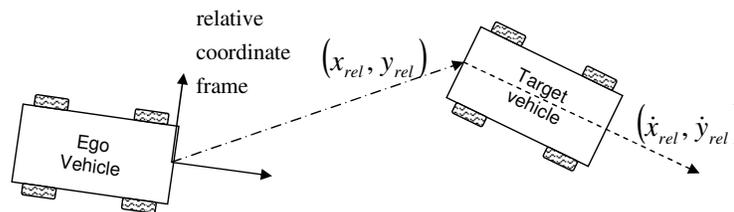


Figure 2: Definition of relative (ego vehicle centered) coordinate system.

included:

$$\xi = (x_{rel}, y_{rel}, \dot{x}_{rel}, \dot{y}_{rel}, \ddot{x}_{rel}, \ddot{y}_{rel}, w)^\top \quad (1)$$

Accelerations are usually included in the state vector because they are required by an ACC controller.

In order to estimate this state, sensory information is necessary. A radar sensor can measure the relative position of a radar target in polar coordinates and in addition the radial velocity exploiting the relativistic Doppler effect, see e. g. [6]:

$$z_{radar} = \begin{pmatrix} r_{rel} \\ \phi_{rel} \\ \dot{r}_{rel} \end{pmatrix} \quad (2)$$

A video sensor with object detection capability, on the other hand, typically determines the presence of an object, e. g. vehicle rear views, by applying for example classifiers to rectangular subregions of a video frame, see e. g. [7]. As can be seen in fig. 3, a detected object is represented by a tight rectangle around a vehicle rear. The four edges of the rectangle in the 2D pixel image correspond to four

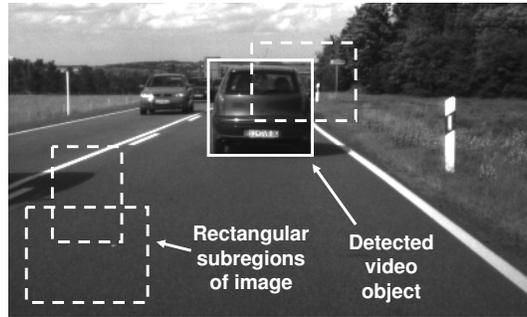


Figure 3: Video object detection using rectangular subregions.

angles under which the vehicle is seen from the camera's point of view. Hence we arrive at

$$z_{video} = \begin{pmatrix} \alpha_{left} \\ \alpha_{right} \\ \alpha_{top} \\ \alpha_{bottom} \end{pmatrix} \quad (3)$$

While the left and right angle or their average can be used directly for the localization of the object in the horizontal plane, the top and bottom angle must be transformed using simple geometric relations, see fig. 4, in order to obtain the object's height and distance. As an object's height is rarely estimated we will not consider the top angle from this point on.

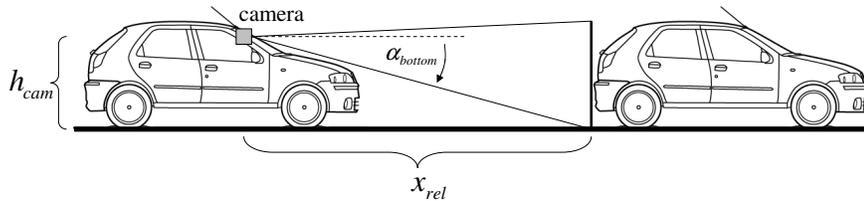


Figure 4: Simplified distance estimation from video object detection.

The figure reveals that an accurate distance computation relies on the following assumptions

- the ego vehicle and the target vehicle are on a planar surface
- the camera height over the surface is known
- the ego vehicle's pitch and roll angles are known
- the lower edge of the rectangle is on the surface vertically below the rearmost edge of the vehicle

Unfortunately, these assumptions do not hold in many driving situations. On vertically or horizontally curved roads the first assumption is violated. As the vehicle's suspension system allows for vertical motion as well as pitching and rolling the camera height and the pitch and roll angles are time-varying and cannot be measured directly in general. In addition, sometimes video classifiers place the lower edge of the rectangle at the vehicle's bumpers or where the car's shadow hits the surface. Although developers of video object detection systems have taken great pains in order to compensate for the above errors, the estimated distances from a video object detection system can still be off by 20% and more.

The relation between measurement and state is given by the output function $h(\xi, u) = z$ with u being an optional input or control vector. Given the state vector (1), the radar measurement vector (2), and the video measurement vector (3) (omitting α_{top} as discussed above) we obtain

$$h_{radar}(\xi) = \begin{pmatrix} \sqrt{x_{rel}^2 + y_{rel}^2} \\ \arctan\left(\frac{y_{rel}}{x_{rel}}\right) \\ \frac{x_{rel}\dot{x}_{rel} + y_{rel}\dot{y}_{rel}}{\sqrt{x_{rel}^2 + y_{rel}^2}} \end{pmatrix} = z_{radar} = \begin{pmatrix} r_{rel} \\ \phi_{rel} \\ \dot{r}_{rel} \end{pmatrix} \quad (4)$$

$$h_{video}(\xi) = \begin{pmatrix} x_{rel} \\ y_{rel} \\ \frac{w}{x_{rel}} \end{pmatrix} = z_{video} = \begin{pmatrix} "h_{cam} / \tan \alpha_{bottom}" \\ \frac{1}{2} (\tan \alpha_{left} + \tan \alpha_{right}) \\ (\tan \alpha_{left} - \tan \alpha_{right}) \end{pmatrix} \quad (5)$$

Here the video angles have been transformed to arrive at algebraically simpler expressions. Finally, the first entry of z_{video} has been put in quotation marks as this simple expression is never used due to the number of assumptions discussed above that go into it. Devising better estimates of this expression is part of the intellectual property of suppliers of video-based object detection systems.

As the state vector evolves over time, a model of the state dynamics in terms of a differential equation $\dot{\xi} = f(\xi, u)$ or in terms of already integrated discrete dynamics $\xi_{k+1} = F(\xi_k, u_k, \Delta t_k)$ with $\xi_k = \xi(t_k)$, $u_k = u(t_k)$ and $\Delta t_k = t_{k+1} - t_k$ is also required. The simplest dynamics for the state vector (1) is a decoupled¹ constant acceleration (CA) model without control input u_k

$$\xi_{k+1} = F(\xi_k, \Delta t_k) = \begin{pmatrix} x_{rel} + \Delta t_k \dot{x}_{rel} + \frac{(\Delta t_k)^2}{2} \ddot{x}_{rel} \\ y_{rel} + \Delta t_k \dot{y}_{rel} + \frac{(\Delta t_k)^2}{2} \ddot{y}_{rel} \\ \dot{x}_{rel} + \Delta t_k \ddot{x}_{rel} \\ \dot{y}_{rel} + \Delta t_k \ddot{y}_{rel} \\ \ddot{x}_{rel} \\ \ddot{y}_{rel} \\ w \end{pmatrix} \quad (6)$$

although more accurate models have been developed that take into account the ego vehicle's dynamics, see e. g. [8].

Together, the state dynamics and the output equation constitute a dynamical system

$$\xi_{k+1} = F(\xi_k, u_k, \Delta t_k) \quad (7)$$

$$z_{k+1} = h(\xi_{k+1}, u_{k+1}) \quad (8)$$

The task is now to derive a "best" estimate $\hat{\xi}_{k+1}$ to the real physical state ξ_{k+1} given the dynamical model F and the output function h . Unfortunately, noise, i. e. disturbing influences, which are too numerous and too complex to be modeled in detail, is inevitable in the measurement process. Also, the assumed dynamical model is never accurate in all driving scenarios as for example the actions of the target vehicle's driver such as accelerating or steering are not known. Those two sources of uncertainty are commonly modeled by separate stochastic noise processes - the measurement noise and the process noise.

Clearly, due to the presence of noise and modeling uncertainties, an estimator can only try to minimize but not eliminate the estimation error $\|\hat{\xi}_{k+1} - \xi_{k+1}\|$. While the design of estimators - in particular for specialized settings and highly non-linear systems - is still an active area of research, the Kalman filter² for linear dynamical systems and the extended Kalman filter (EKF) for non-linear systems (see e. g. [9]) has emerged as the industry standard for tracking and estimation during the past 40 years. As can be seen from the two non-linear output functions (4) and (5) an extended Kalman filter is required for those sensors. In the following, when referring to a Kalman filter, a Kalman filter or an extended Kalman filter is implied as dictated by the (non-)linearity of the dynamical system.

Radar-video fusion for driver assistance systems The Kalman filter is not only a means for estimation, it is also commonly used for sensor fusion. In so-called "sensor-to-track fusion," raw, unsynchronized measurements of several sensors are combined in a central estimate or track as in fig. (5) where sensor-to-track fusion is exemplified by the combination of a radar and a video sensor.

¹Decoupled means that the model dynamics in x -dimension is independent of the dynamics in y dimension.

²Filter is used here in its general meaning of a system with inputs and outputs whose output depends on the current input as well as previous inputs, not in its narrower

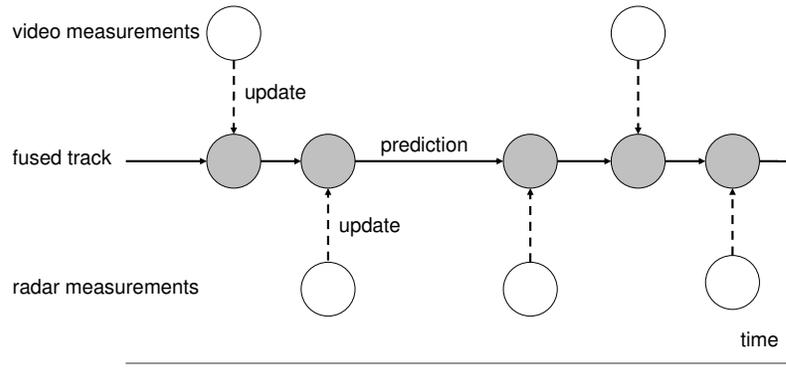


Figure 5: Block diagram of sensor-to-track fusion of a radar and a video sensor. The prediction and update steps can be realized using a Kalman filter.

The improved estimation accuracy that can be achieved by sensor fusion critically depends upon the nature of the sensor measurements in terms of their measurement vector z and their corresponding measurement accuracy in terms of standard deviations or variances. The question of whether the estimated quantities of a Kalman filter can converge to the true values can be formalized using the system-theoretic concept of *observability*. One of the necessary conditions for a Kalman filter to converge is that its dynamical system be observable, see e. g. [10]. Roughly speaking, a system is observable if its state vector at a certain time can be deduced from the current and a series of previous measurements.

It is obvious that without a video sensor or scanning radar the vehicle’s width is unobservable. Although it can be shown that even without a radar sensor using video only the dynamical system is observable, the Kalman filter’s estimate would be highly inaccurate due to the large uncertainties in the video distance measurement as discussed above. For this reason driver assistance systems using only video sensors have not yet been mass-produced for functions that require accurate distance information such as ACC.³ It is for this same reason that we will not treat the radar and video sensor on an equal footing in sensor fusion. In particular video measurements will only be used to update already existing tracks that have been initialized by radar but will not be allowed to initialize new tracks within the common field of view. Also, the video distance measurement will be assigned a low weight in the track update and therefore will not strongly influence and possibly corrupt the track’s estimated distance. In the next subsection we will show experimental results on the improvements in estimation accuracy using video-radar sensor fusion.

Experimental results For an AEB system we have implemented a fusion module using a 24GHz radar and a video camera with object detection capabilities. We have evaluated the improvement in the estimation accuracy of the lateral distance y by fusing video and radar information in comparison to tracking using radar sensor data only.

In the top plot of fig. 6 the lateral distances as delivered by radar tracking only and by a fused estimate from radar and video sensor data are shown. In addition, the ground truth lateral distance from a reference system is plotted. In the bottom plot of fig. 6 the magnitude of the estimation error with respect to the ground truth lateral distance is shown. It is obvious that the estimation accuracy of the lateral distance y has been improved dramatically by sensor fusion. The ground truth data were obtained from two coupled DGPS-aided inertial navigation systems mounted in the ego and the target vehicle. With this reference system a position accuracy of up to 2cm can be realized.

IMPROVING AN OBJECT’S CONFIDENCE BY SENSOR FUSION In this section we will focus on quantities that estimate the confidence in an estimated object or quality of an estimated object. They should allow the distinction whether an estimated object corresponds to a real object or whether it corresponds to random noise or clutter. Sometimes those quantities can also be interpreted as a *probability of existence*. We show how we can derive such a probability of existence using a prominent probabilistic framework, namely Bayes estimation. This framework can also be used to fuse measurements from several sensors and thus deliver a better estimate. For an alternative approach to computing a probability of existence using a generalization of Bayes estimation we refer to [12]. Furthermore,

meaning of a (low-pass) frequency filter. Indeed, as noise is modeled best by a stochastic process, the application of pure frequency filters has been proven inefficient in estimation.

³Note that in principle it is possible to track an object and estimate its distance with bearing or azimuthal angle information only (bearings-only-tracking, BOT, see e. g. [11]). This, however, requires assumptions about the dynamics between the tracker and the tracked object that are generally not met in an automotive setting. Bearings-only-tracking is used e. g. by submarines and fighter aircraft where for stealth purposes only passive sensing is employed.

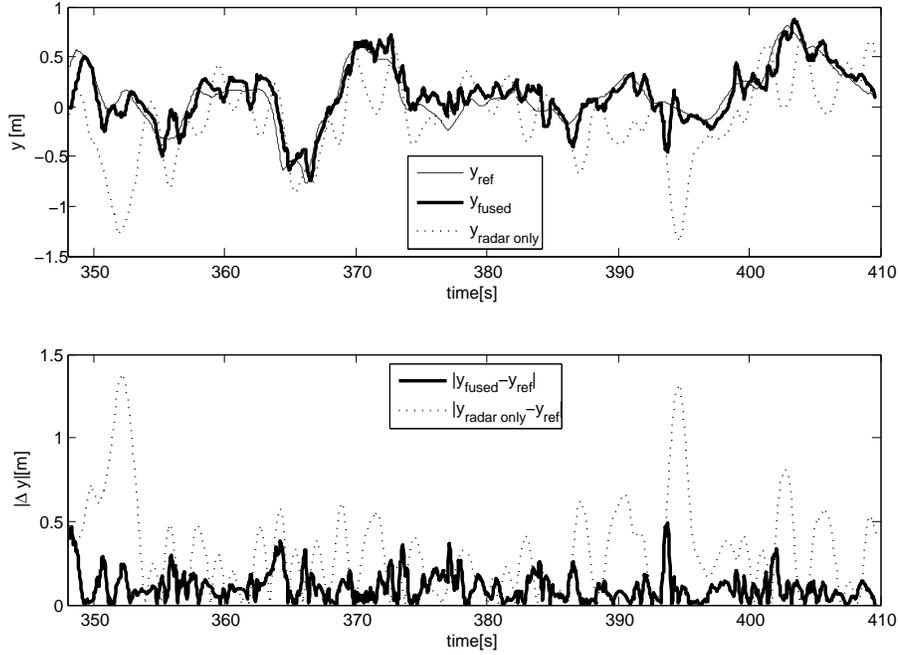


Figure 6: Top: Lateral distance from radar tracking only and from radar-video fusion as well as ground truth lateral distance. Bottom: Estimation errors with respect to ground truth data using radar-video fusion and using radar tracking only.

we extend our concept of probability of existence to a *localized* probability of existence, meaning we reason how probable it is for an object to exist in a certain region.

Bayes estimation In estimation we want to determine an in some sense best or optimal estimate of a state ξ at time t_k ; this state can be a kinematic vector containing position, velocity, etc. of an object, or it can be a single hypothesis such as whether an estimated object actually exists in reality, i. e. the hypothesis of existence. The estimate is based upon all observations or measurements up to the time t_k of the object characterized by ξ . Because of the uncertainty caused by inevitable noise sources, the state can only be characterized by a probability distribution over the state space. Bayes estimation (for an overview see e. g. [13]) computes the probability density function (pdf) of the state ξ at time t_k given all measurements $Z^k = \{z_k, z_{k-1}, \dots, z_0\}$ that have been associated with this object over time: $p(\xi_k|Z^k)$. Bayes estimation can be performed iteratively, i. e. given the estimate $p(\xi_k|Z^{k-1})$ based upon all previous measurements Z^{k-1} and given the current measurement characterized by its measurement likelihood $p(z_k|\xi_k)$, the new estimate $p(\xi_k|Z^k)$ is

$$p(\xi_k|Z^k) = \frac{p(z_k|\xi_k)p(\xi_k|Z^{k-1})}{p(z_k|Z^{k-1})}. \quad (9)$$

Here the conditional probability $p(z_k|Z^{k-1})$ is a normalization factor.

In our case, the state ξ_k we are interested in is a binary hypothesis, namely whether the object exists or does not exist at time-step k : $\xi_k \in \{\exists_k, \nexists_k\}$. By using this binary state in eq. (9) and transforming to the so-called logarithmic odds (LO) the Bayes iteration simplifies considerably:

$$\begin{aligned} LO(\exists_k|Z^k) &= \ln \frac{p(\exists_k|Z^k)}{p(\nexists_k|Z^k)} \\ &= \ln \frac{p(\exists_k|Z^{k-1})}{p(\nexists_k|Z^{k-1})} + \ln \frac{p(z_k|\exists_k)}{p(z_k|\nexists_k)} \\ &= LO(\exists_k|Z^{k-1}) + LLR(\exists_k|Z^k). \end{aligned} \quad (10)$$

This iteration formula is the basis for the track score formalism discussed in [14, 15]. It was originally developed as a means to evaluate

different hypotheses of associating measurements to tracks. However, it is commonly used for track management, i. e. for track initiation, confirmation, and deletion.

Here, we will use this iteration for the estimation of a probability of existence. An estimator usually consists of a (temporal) prediction of the state and an update or correction of the state due to new sensory information (predictor-corrector structure). The update step is given by Bayes estimation whereas the prediction step requires the specification of a dynamical model.⁴ As illustrated in fig. 7 the estimation of the kinematic state ξ_k as discussed in section “Improving kinematic state estimation by sensor fusion” and the estimation of the probability of existence proceeds in parallel and with analogous steps. The combination of both estimates to form a localized probability of existence will be explained below.

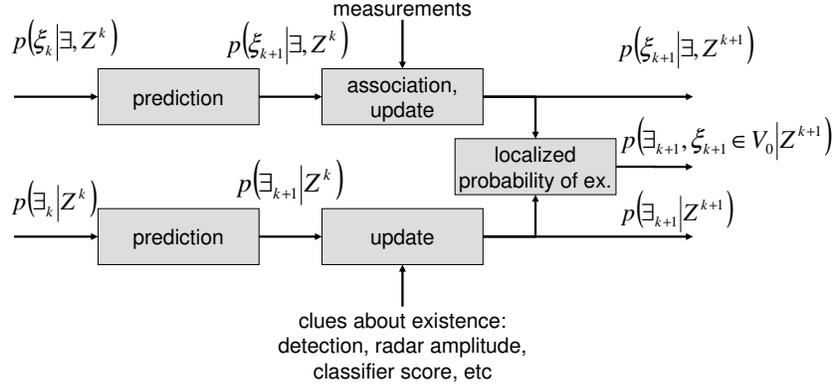


Figure 7: Separate estimation of object state ξ and existence \exists using the factorization from eq. (14).

First we discuss the update step which has to incorporate clues about an object’s existence that are specific to radar and video sensors.

UPDATE For the incorporation of new sensory information using the Bayes update step eq. (10) the LLR, i. e. the measurement likelihood ratio, must be provided for radar and video measurements.

Measured radar data contain also the signal amplitude a_k in addition to the kinematic quantities given in eq. (2). As the radar amplitude can provide strong clues about an object’s existence, the LLR reads

$$LLR_{radar}(\exists_k | z_k) = \ln \frac{p(a_k | \exists_k)}{p(a_k | \nexists_k)}. \quad (11)$$

Assuming a Swerling I target fluctuation model and Gaussian clutter, explicit expression for those likelihoods can be derived, see e. g. [6, 16]. If no radar measurement has been received or associated (\emptyset) we need to specify $\ln \frac{p(\emptyset | \exists_k)}{p(\emptyset | \nexists_k)}$, see e. g. [16, 12] for an expression for the common cell-averaging constant false alarm rate (CA-CFAR) detector.

Likewise, such measurement likelihoods must be generated for video measurements if a probability of existence is to be computed using video. In the automotive domain where computational power is limited and a fast detection rate is essential it is common to use cascaded classifiers for video detection. By cascading i. e. serially concatenating classifiers, a first, fast layer can sort out most false detections. Subsequent, more complicated layers can result in a high true detection rate [17]. The output of these classifiers typically is a classifier score s_k . Hence the measurement likelihood ratio reads

$$LLR_{video}(\exists_k | z_k) = \ln \frac{p(s_k | \exists_k)}{p(s_k | \nexists_k)}. \quad (12)$$

One possibility of obtaining those likelihoods is to compare the classifier score with ground truth (i. e. manually labeled) data and determine empirical probabilities. Those probabilities can then be retrieved from a look-up table, for example. Another possibility is the direct generation of a probability of existence from the classifier without resorting to stored, off-line data. This requires the extension of a cascaded classifier to a probabilistic boosting tree (PBT) [18] which has been discussed in [19] in the context of an automotive application.

PREDICTION The dynamical model has to be chosen carefully with respect to the task at hand. In the track score formalism [14, 15], no dynamical model was assumed, i. e. that a true target does not spontaneously change to clutter from one time step to the next. In [20], an object’s birth and destruction was explicitly modeled. In [21] the dynamics for the probability of existence was modeled by a Markov

⁴This is in complete analogy to a Kalman filter. Indeed, the Kalman filter can be derived from Bayes estimation.

chain and depended upon the sensor ranges and object occlusions such that for example the transition probability from existence to non-existence was higher outside the sensor field of view. The rationale for this is that the opposite of existence included non-existence, non-perceivability, and non-relevance.

On the other hand, we want to purely model an object’s existence without any occlusion or relevance reasoning. In automotive scenarios the existence of a traffic participant does not actually change over time unless it e. g. crashes and disintegrates into pieces; this holds even if it moves outside a sensor’s FOV. However, we still want to decrease the probability of existence from one time step to the next independent of any measurements - this is a conservative, precautionary measure with the purpose of counteracting unmodeled effects and modeling errors that lead to a too high probability of existence. For a consistent framework where the transition probabilities add up to one we therefore model the prediction stage as a complete Markov chain as in [22] - albeit with very small birth and destruction probabilities: $p(\exists_{k+1}|\nexists_k) \ll 1, p(\nexists_{k+1}|\exists_k) \ll 1$.

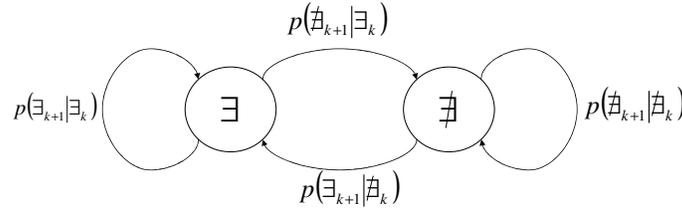


Figure 8: Markov model for probability of existence dynamics.

Obviously, by using such a Markov chain the probability of existence is limited to a value smaller than one.

LOCALIZED PROBABILITY OF EXISTENCE Up to this point, our confidence measure indicates how much confidence we have that the estimated track actually corresponds to an existing object. We now extend the confidence measures from the probability of a binary hypothesis $p(\exists|Z^k)$ to a joint probability density function $p(\exists, \xi|Z^k)$ where ξ is a state vector containing position, speed, etc, of the object because it is hard to interpret existence without position.

The fact that an estimated object exists with a probability of 99% without reference to a position has little practical meaning as its existence at a distant galaxy is generally irrelevant. For the decision to execute an automatic emergency braking maneuver, however, the information that an object exists with 99% probability in a certain region in front of the car is highly relevant. This is achieved by “marginalization” over a certain volume V_0 of the kinematic state as discussed in [12]:

$$p(\exists, \xi \in V_0|Z^k) = \int_{V_0} p(\exists, \xi|Z^k) d\xi \quad (13)$$

Using Bayes’ theorem the joint pdf $p(\exists, \xi|Z^k)$ can be factorized into

$$p(\exists, \xi|Z^k) = p(\exists|Z^k)p(\xi|\exists, Z^k). \quad (14)$$

This means that the kinematic state estimate is conditional on the existence of the tracked object and is undefined if the tracked object does not exist [23]. Then the “marginalization” simplifies to

$$\begin{aligned} p(\exists_k, \xi_k \in V_0|Z^k) &= \int_{V_0} p(\exists_k, \xi_k|Z^k) d\xi_k \\ &= p(\exists_k|Z^k) \int_{V_0} p(\xi_k|\exists_k, Z^k) d\xi_k \end{aligned} \quad (15)$$

A high uncertainty in the kinematic state estimation will result in a large state covariance which in turn will reduce the integral in (15).

Experimental results We have evaluated the above algorithm in an automotive collision mitigation braking (CMB) system based on a 24GHz radar sensor with a cycle time of 40ms and a video camera with object detection capabilities. The system can autonomously brake in case of a perceived impending collision with a deceleration of up to $-0.5g = -4.9m/s^2$. In order to meet the stringent limits on the false alarm rate, the localized probability of existence $p(\exists, \xi \in V_0|Z^k)$ must exceed a certain threshold as a qualifying condition for collision mitigation braking.

The following plot depicts the logarithmic odds $LO_k = \ln \frac{p(\exists|Z^k)}{1-p(\exists|Z^k)}$ during a track’s life using radar and video updates. For better illustration of the Bayes estimator’s effect the localized probability of existence is not shown here as variations of the position covariance would be superimposed on the variations due to the measurement likelihoods.

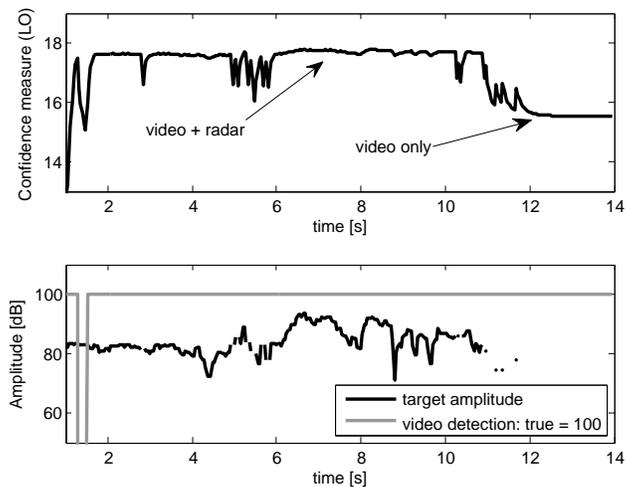


Figure 9: Fused LO of a sample track over its life time with updates from radar and video measurements.

It can be seen that an object that is seen by both radar and video sensor has a higher LO and hence a higher probability of existence than an object that is seen by only one of the sensors. This is a substantial improvement over heuristic rules such as “object must be detected by both sensors in at least 8 of the last 10 sensor cycles” used for a collision mitigation braking. In this formalism the confidence or probability of existence gain can be quantified using measurement likelihoods from individual sensors and the update is derived formally as a Bayes estimator.

SUMMARY AND CONCLUSIONS

In this paper we have outlined some of the benefits of sensor fusion for driver assistance systems. In particular we have argued that sensor fusion is an enabling technology for safety-critical DAS such as automatic emergency brake that entails a massive intrusion into the driver’s control of the vehicle and can endanger the occupants of the ego as well as the target vehicle in case of a malfunction or false alarm. We have shown the improvement in estimation accuracy of the lateral distance which will benefit the selection or identification of a relevant collision object. We have also shown how an object’s probability of existence can be calculated using Bayesian estimation and how the localization of this measure leads to a relevant criterion for automatic emergency braking. Using sensor fusion an object’s probability of existence will reach higher values and hence a higher object integrity. In general, the improvement in kinematic state estimation due to sensor fusion and the increase in an object’s confidence or probability of existence due to independent detections of the same object by different, ideally heterogeneous sensors, will lead to more true alarms and fewer false alarms.

REFERENCES

- [1] R. van der Horst and J. Hogema. Time-to-collision and collision avoidance systems. In *Proceedings of the sixth ICTCT workshop, Salzburg, Austria*, pages 1–12, 1993.
- [2] M. M. Muntzinger, S. Zuther, and K. Dietmayer. Probability estimation for an automotive pre-crash application with short filter settling times. In *Proceedings of IEEE Intelligent Vehicles Symposium*, pages 411–416, 2009.
- [3] A.N. Steinberg, C.L. Bowman, and F.E. White. Revisions to the JDL data fusion model. In *Proceedings of SPIE, Sensor Fusion: Architectures, Algorithms, and Applications III*, volume 3719, pages 430–441, 1999.
- [4] B.V. Dasarathy. Information fusion - what, where, why, when, and how? *Information Fusion*, 2(2):75–76, 2001.
- [5] D.L. Hall and S.A.H. McMullen. *Mathematical techniques in multisensor data fusion*. Artech House, 2004. second edition.
- [6] N. Levanon. *Radar Principles*. Wiley, 1988.
- [7] Z. Sun, G. Bebis, and R. Miller. On-road vehicle detection: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(5):694–711, 2006.
- [8] R. Altendorfer. Observable dynamics and coordinate systems for automotive target tracking. In *Proceedings of IEEE Intelligent Vehicles Symposium*, pages 741–746, 2009.

- [9] M.S. Grewal and A.P. Andrews. *Kalman Filtering*. Wiley, 2001.
- [10] Y.K. Song and J.W. Grizzle. The extended Kalman filter as a local asymptotic observer for nonlinear discrete-time systems. *Journal of Mathematical Systems, Estimation and Control*, 5(1):59–78, 1995.
- [11] V. Aidala and S. Hammel. Utilization of modified polar coordinates for bearings-only tracking. *IEEE Transactions on automatic control*, 28(3):283–294, 1983.
- [12] R. Altendorfer and S. Matzka. A confidence measure for vehicle tracking based on a generalization of Bayes estimation. In *Proceedings of IEEE Intelligent Vehicles Symposium*, pages 766–772, 2010.
- [13] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. MIT Press, 2005.
- [14] S. Blackman. *Multiple-Target Tracking with Radar Applications*. Artech House, 1986.
- [15] S. Blackman and R. Popoli. *Design and analysis of modern tracking systems*. Artech House, Boston, 1999.
- [16] G. V. Weinberg. Estimation of false alarm probabilities in cell averaging constant false alarm rate detectors via Monte Carlo methods. Technical Report DSTO-TR-1624, Defence Science and Technology Organisation (Australia), 2004.
- [17] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 511–518, 2001.
- [18] Z. Tu. Probabilistic boosting-tree: learning discriminative models for classification, recognition, and clustering. In *Proceedings of IEEE International Conference on Computer Vision*, pages 1589–1596, 2005.
- [19] R. Schweiger, H. Hamer, and O. Löhlein. Determining posterior probabilities on the basis of cascaded classifiers as used in pedestrian detection systems. In *Proceedings of IEEE Intelligent Vehicles Symposium*, pages 1284–1289, 2007.
- [20] R.W. Sittler. An optimal data association problem in surveillance theory. *IEEE Transactions on Military Electronics*, 8(2):125–139, 1964.
- [21] M. Mählisch, M. Szczot, O. Löhlein, M. Munz, and K. Dietmayer. Simultaneous processing of multitarget state measurements and object individual sensory existence evidence with the joint integrated probabilistic data association filter. In *Proceedings of WIT 2008: 5th International Workshop on Intelligent Transportation*, 2008.
- [22] D. Musicki, R. Evans, and S. Stankovic. Integrated probabilistic data association. *IEEE Transactions on automatic control*, 39(6):1237–1241, 1994.
- [23] D. Musicki. Track score and target existence. In *Proceedings of the 9th International Conference on Information Fusion*, pages 1–7, 2006.

CONTACT INFORMATION

Richard Altendorfer, Sebastian Wirkert, Sascha Heinrichs-Bartscher
 Driver Assistance Systems
 TRW Automotive
 Carl-Spaeter-Str. 8
 56070 Koblenz
 Germany