

Klaus Dietmayer

---

## 20.1 Introduction

In the case of highly-automated and fully-automated driving it is necessary for the vehicle itself to recognize the limitations of its machine perception, as well as the functional limitations of processing modules based on this perception to react adequately. While simulator studies of highly-automated driving have shown that realistic transfer times to the driver of between 5 and 10 s can be assumed [1, 2] before the driver can reliably take over the driving task again, with fully-automated driving a human would not provide any backup whatsoever. In the case of functional limitations, the vehicle would have to be able to achieve an intrinsically safe state completely by itself. However, potential transfer times of 5 s and more require extensive autonomy of the vehicle, if only for a limited time, in order to be able to bridge this time period reliably under all circumstances.

To be able to achieve this degree of autonomy, the vehicle must perceive its surroundings, interpret them appropriately and be able to derive and execute reliable actions continuously. Technically, this task is carried out by individual processing modules that build on each other. A simplified representation of the relationships is shown in Fig. 20.1.

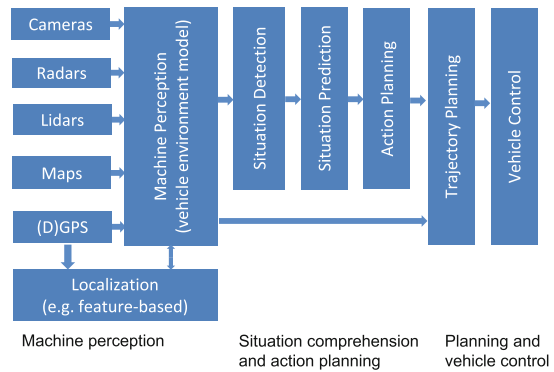
The machine perception of the vehicle's surroundings is enabled by various sensors, such as cameras or radar sensors, incorporated into the vehicle. Further information about the static driving environment is usually added from very precise digital maps. However, this can only be used when the vehicle knows its exact position. Therefore, the vehicle also requires a self-localization functional module for the map matching. The result of the machine perception is a dynamic vehicle environment model in which the vehicle itself and all other road users are represented by individual dynamic motion models. This

---

K. Dietmayer (✉)

Universität Ulm, Institute of Measurement, Control and Microtechnology,  
89081 Ulm, Germany  
e-mail: klaus.dietmayer@uni-ulm.de

**Fig. 20.1** General structure of the information processing for automated vehicle driving. Image rights: Author has copyright



should also contain all the relevant infrastructure elements such as traffic signs and traffic lights, as well as structuring elements such as traffic islands and curbstones, road markings for dividing traffic lanes, closed areas or pedestrian crossings.

Based on this vehicle environment model, during the situation recognition, all the individual components are set in relation to each other in order to generate a machine interpretation of the scene from the dependencies of the individual elements. In the situation prediction module based on this, various possible developments of the scene over time, also known as episodes, are calculated in advance and evaluated as to the probability of their occurrence. Therefore, in this document an episode refers to a possible specific development over time for a detected traffic scene, whereby the time horizon lies within the range of a few seconds. On the basis of this situational information, the module based on this determines the higher-level action planning. For example, it could stipulate driving around an obstacle or overtaking a slower vehicle. For the execution of the plans, possible trajectories of the vehicle are calculated with a typical time horizon of 3–5 s and are evaluated in terms of safety and comfort. The optimal trajectory based on the criteria that can be stipulated is executed by the vehicle control. The processing procedure described is repeated continuously, usually in line with the data capture of the sensors, so that the vehicle is able to react to the actions and reactions of other road users.

The description of this technical process chain clearly shows that a failure of the machine perception would immediately lead to uncertainties in the situation evaluation of such a magnitude that reliable safe action planning and action execution would no longer be possible. The degradation of the machine interpretation of the scene and the action planning and action execution based on this depends on the situation; however, reliable prediction would typically not be able to exceed 2–3 s. Therefore, it is evident that a minimum perception capability is required even for highly-automated driving due to the significantly greater transfer times to the driver. A complete failure of the machine perception must be avoided in all circumstances, though of course this also applies to the modules based on it and the vehicle control with its sensors and actuators, which are not within the focus of this document, however.

Therefore, the question is whether limitations in the operation of the machine perception can be detected or even predicted, and if they are, over what period of time. In this context, the following sections will discuss the state of the technology of known methodical approaches, and on this basis derive possible research questions.

---

## 20.2 Machine Perception

### 20.2.1 Scope and Characteristic

As described in the previous section, the task of machine perception is to reliably detect all the other road users relevant to the operation of the automated driving, and to assign them correctly to the traffic infrastructure. This is particularly necessary because, for example, a pedestrian at the side of the road presents a different potential risk than one who is using a separate pedestrian walkway running parallel to the road.

For the machine perception sensors based on camera and radar and/or lidar technology are used. More detailed information on the operation and design of these sensors can be found in [3], for example. Cameras provide a 2D representation of a 3D scene in the form of high-resolution gray-scale or colored images, from which image processing methods can be used to extract individual objects when there is sufficient contrast or differentiation in the texture. However, the object distance can only be determined with mono cameras based on assumptions that often lead to errors, such as a flat surface. Although stereo cameras also enable the object distance to be determined by means of the disparity image, the accuracy decreases quadratically as the distance increases. With the currently prevailing base distances for the stereo arrangements and the resolution of the cameras, measuring ranges of up to around 50 m are possible without the error margin increasing to such an extent that functions could no longer make any use of the data.

On the other hand, radar and also lidar sensors provide distance measuring data that is comparatively very accurate and also practically distance-independent in terms of the measuring error margin. However, due to their low angle resolution, they are less accurate in capturing the contours, i.e. the external dimensions of objects. This applies in particular to radar sensors. Additionally, radar and lidar sensors do not provide any texture information. Due to these different measuring properties, the different sensor types are generally used in combination to create the machine perception. This is referred to as sensor data fusion.

The combined sensor data enables moving and static objects, but also road surface markings, for example, to be categorically detected and physically measured. The possible measuring dimensions depend on the specific sensor set-up. Typical physical measured data that can be captured includes the dimensions of an object used for a box model with length, width and height, as well as its position absolutely in the world or relative to the vehicle. In the case of moving objects, the object speeds and object accelerations are added to this data. More difficult to determine, and generally very unreliably, from

external sensor measurements is the yaw rate or yaw angle of other road users. Without vehicle-to-vehicle communication, these variables can only be determined reliably for one's own vehicle.

However, for the subsequent situation evaluation and situation prediction, not only the physical measurement of the objects is required, but also information about what class of object is involved. For example, a pedestrian and a motorcyclist differ in terms of their possible degrees of freedom of movement and also their possible movement dynamic. Also, depending on the context and constellation, road surface markings can have different meanings. Therefore, it is necessary also to determine the semantic meaning of the objects detected from the sensor data, or from other information sources such as a digital map. In the context of the machine perception, this operation is known as a classification step, but it is a component of the machine perception.

While humans are able to assign a semantic meaning to the visual perceptions very quickly and nearly without errors, this is still a comparatively difficult task for the machine perception with the current state of the technology. The known classification algorithms are always based on more or less complex models of expected object classes, which are either learned automatically from examples or are specified manually. These models then display, as discriminately as possible, characteristics that can be captured with the available sensors, so that a distinction can be made between the object classes that occur. However, it also becomes clear that object classes that are not trained in advance cannot be identified semantically with the methods known at present. Due to their significantly greater capabilities, learning classification algorithms have become widely accepted.

A machine perception with semantic information is only technically possible in the context of driver assistance systems and automated driving because the driving area is well structured and limited to a few object classes. Additionally, only a rough class differentiation is relevant for situation recognition and situation prediction. With the current state of the technology, it is sufficient to be able to distinguish between the pedestrian, cyclist, passenger car and truck or bus classes with respect to moving objects. Additionally, there are stationary obstacles, but these are usually assigned to a residue class along with the non-classifiable objects.

For the correct assignment of the classified objects to the traffic infrastructure, it is also necessary to be able to identify reliably, with the correct semantic meaning, road surface markings, blocked areas, stop lines, traffic light systems and traffic signs. As this complex classification task is not yet possible with the required degree of reliability, highly accurate and comprehensively attributed digital maps are used as a support, based on the state of the technology. Knowing its own position, the automated vehicle can use these maps to identify the stationary objects and markings expected in the sensors' field of vision, together with their semantic meaning. The sensors then only have to verify that the objects are present.

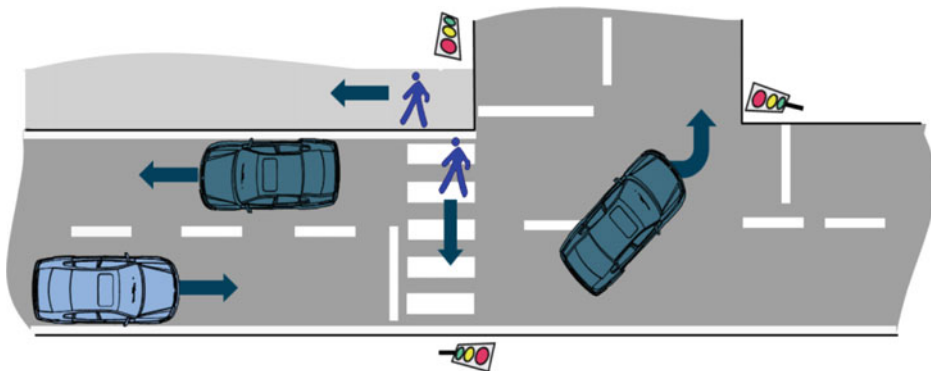
A disadvantage of this approach is that a highly accurate localization of the vehicle is required, for which a standard GPS localization is not sufficient, and the map must always be up to date. For this reason, the goal is to develop technical solutions in the future that will no longer require highly accurate, up-to-date maps.

### 20.2.2 Characteristics of Environment Models

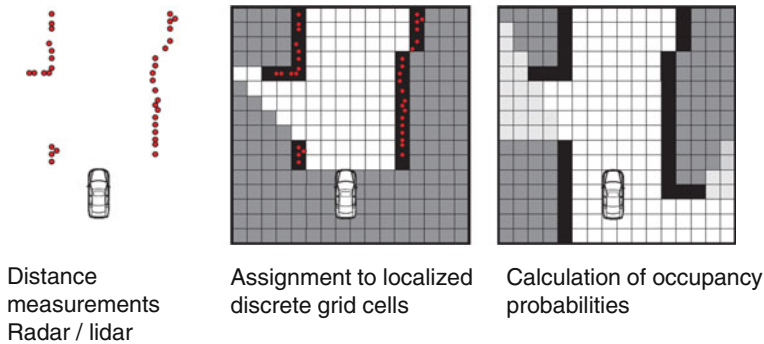
The machine perception is used to create a dynamic environment model. Two main representations are known: object-based and grid-based forms. Both forms of representation can also be combined.

An object-based vehicle environment model is a dynamic data structure in which all the relevant object and infrastructure elements in the vicinity of the vehicle are represented correctly in space and time. As explained above, the capturing and tracking over time of the objects and infrastructure elements is performed continuously by suitable, usually fused on-board sensors such as cameras, radars, lidars, and with the additional use of highly accurate digital maps. Figure 20.2 shows an example of components that incorporate an environment representation.

Which objects and structure elements are relevant for automated driving mainly depends on the driving task to be performed, the complexity of which increases starkly, starting from simple motorway scenarios via country roads to inner-city traffic. In the object-based representation, all the other road users relevant to the representation, the relevant infrastructure elements and one's own vehicle itself are described by means of a separate dynamic object model, usually a time-discrete state space model. The states of this model, such as position, speed or 2D/3D object dimensions, are continuously updated in line with the sensor measurements. Furthermore, there is continuous capturing of the road surface markings and traffic signs, as well as the status of the traffic light systems.



**Fig. 20.2** Schematic diagram of the object-based vehicle environment representation. All the relevant objects are detected, classified and correctly assigned to the infrastructure. Image rights: Author has copyright



**Fig. 20.3** Schematic diagram of the structure of a grid-based representation of the vehicle's surroundings. In the simplest case, this only contains static obstacles. Image rights: Author has copyright

A grid-based representation uses raster maps to divide the stationary environment into localized cells of equal size. The vehicle moves across this localized 2D or 3D grid and the on-board, the sensors then only supply information as to whether specific cells are free and thus can be driven on freely, or whether there is an obstacle in the respective cell. Additionally, the state in which no information on cells is available can also be modeled. This type of depiction is mainly suitable for representing static scenarios and static obstacles. It does not require any model hypotheses about the object classes to be expected and can therefore be categorized as very resistant to model errors. Figure 20.3 shows the basic procedure. Further information on grid-based representations can be found in [4–8].

## 20.3 Methods for Dealing with Uncertainties of Machine Perception

### 20.3.1 Uncertainty Domains

As described in Sect. 20.2, the machine perception is made up of different task scopes. These are, on the one hand, detecting static and dynamic objects and physically measuring them as precisely as possible, and on the other, assigning the correct semantic meanings to the detected objects. In the context of these tasks, the following three uncertainty domains exist for the machine perception:

1. State uncertainty

State uncertainty describes the uncertainty in the physical measured variables, such as size, position and speed, and is a direct consequence of measuring errors in the sensors and sensor signal processing that cannot always be avoided.

## 2. Existence uncertainty

Existence uncertainty describes the uncertainty as to whether an object detected by the sensors and transferred to the representation of the surroundings actually exists at all. Errors of this kind can occur due to deficiencies in the signal processing algorithms or incorrect measurements by the sensors themselves.

## 3. Class uncertainty

This refers to uncertainty with regard to the correct semantic assignment, which can be caused by deficiencies in the classification procedure or insufficiently accurate measured data.

In order to facilitate automated driving, it is necessary to reliably detect any uncertainties or errors in the various domains and, if possible, even to be able to predict them. In the current state of the technology, uncertainties are handled, almost without exception, using methods based on Bayes' theorem [9–11] or on the generalization of same, the Dempster-Shafer theory [12]. The advantage of these methods is that they allow the uncertainty domains to be handled using a totally probabilistic and therefore mainly heuristic-free approach.

In the narrow sense, the uncertainty domains named above only apply to the on-board sensors for now. However, errors in the information from a digital map or in the data obtained via Car2x communication can also be categorized. Car2x communication in particular can harbor additional sources of error due to possible variable latency times in the transfer of data and the possibility of imprecisely known uncertainty evaluations of the sending sources. However, the effects can still be assigned to the three uncertainty domains named, and therefore we will not go into further detail here.

### 20.3.2 State Uncertainty

The state uncertainty of a detected object is described, in accordance with Bayes' theorem, by means of a probability density function which can be used to determine the most probable total and individual state and also, with a certain probability, possible variations from this. In the case of a multi-dimensional, normally distributed probability density function, the state uncertainty is completely represented by a covariance matrix.

In estimating static variables, such as the vehicle dimensions, their state uncertainty can be reduced progressively by means of repeated measurements. The estimated value based on the available measurements converges with the true values, as long as there is no systematic sensor error, e.g. in the form of an offset. For the estimation of dynamic, time-changeable states such as the object position or the object speed, due to the movement of the object between the measuring times, there is no convergence with a true value. Therefore, for the evaluation of the quality of the state estimation, it is stipulated that the mean error is zero and the uncertainty as low as possible.

The basic procedure for handling state uncertainties is the general Bayes filter [9]. With this, the estimated state of an object and the related uncertainty are represented by a multi-dimensional probability density function  $p$  (PDF):

$$p_{k+1}(x_{k+1}|Z_{1:k+1}).$$

In general, it depends on all the measurements  $Z_{1:k+1} = \{z_1, \dots, z_{k+1}\}$  available at time  $k + 1$ . This is expressed by means of the selected notation of a conditional probability, i.e. the probability for the state of system  $x$  is conditional upon measurements  $Z$ .

The motion model of an object captured by the sensors for the period between two consecutive measurements is described by a motion equation of the form

$$x_{k+1|k} = f(x_k) + v_k$$

whereby  $v_k$  represents an additive disturbance variable representing possible model errors. The motion equation expresses in which state, such as location, speed and direction of motion, the object will probably find itself at the next point in time. Alternatively, this motion equation can also be expressed by means of a Markov transition probability density:

$$f_{k+1|k}(x_{k+1}|x_k).$$

The Markov transition probability density is ultimately only another mathematical notation for the same model assumptions. To keep the equations practically calculable, it is common to presuppose a Markov property of the first order. This property expresses simplistically that the future state of a system only depends on the last known state and the current measurement, not on the entire history of measurements and states. Therefore, the Markov property of the first order is a presupposed system property. In our specific case, the predicted state  $x_{k+1}$  of the object before the new measurement is available only depends on the last determined state  $x_k$ , as this implicitly comprises the entire measurement history  $Z_{1:k} = \{z_1, \dots, z_k\}$ .

The prediction of the current object state  $x_k$  to the next measuring time  $k + 1$  is basically carried out based on the Chapman-Kolmogorov equation

$$p_{k+1|k}(x_{k+1}|x_k) = \int f_{k+1|k}(x_{k+1}|x_k)p_k(x_k)dx_k.$$

This is denoted as a prediction step of the Bayes filter.

The measuring process of the sensors can generally be described as a measurement equation in the form

$$z_{k+1|k} = h_{k+1}(x_{k+1}) + w_{k+1}$$



The measurement function  $h(\cdot)$  describes how measurements and state variables are related. For example, if a state variable can be measured directly, then  $h(\cdot)$  is a 1:1 mapping. Here, the stochastic disturbance variable  $w_{k+1}$  represents a possible measuring error. An alternative mathematical representation of the measurement equation is the likelihood function

$$g(z_{k+1}|x_{k+1}).$$

If the current measurements  $z_{k+1}$  are available, the probability density function of the object state is updated. The current estimate of the state is calculated using the Bayes formula

$$p_{k+1}(x_{k+1}|z_{k+1}) = \frac{g(z_{k+1}|x_{k+1})p_{k+1|k}(x_{k+1}|x_k)}{\int g(z_{k+1}|x_{k+1})p_{k+1|k}(x_{k+1}|x_k)dx}.$$

This second step to incorporate the current measurement is known as the innovation step.

The recursive estimation procedure briefly described by the prediction step and the innovation step is known as the general Bayes filter, and all the methods and implementations of the stochastic state estimation commonly used today are based on this. Along with the process and measurement equations, the procedure only requires an a priori PDF for object state  $p_0(x_0)$  at time  $k = 0$ . However, it is not efficient to implement the filter in this general form.

With the assumption of normally distributed measurement signals and linear models, the Kalman filter [13] enables a simple analytical implementation of the general Bayes filter. As a Gaussian distribution is completely described by its first two statistical moments, i.e. the mean value and the related covariance matrix, the temporal filtering of these two moments represents a mathematically exact solution. The Kalman filter can be applied to systems with non-linear process or measurement equations by using the Extended Kalman Filter (EKF) [11] or the Unscented Kalman Filter (UKF) [14]. While the EKF linearizes the system equations by using a Taylor series approximation, the objective of the UKF is a stochastic approximation by using what are known as sigma points [14].

Independently of the specific implementation, all the procedures based on the general Bayes filter have in common that they continuously supply a probabilistic measure for the uncertainty of the physical variables determined from the sensor data. This enables the reliable detection of sensor failures, but also of degeneration in the capabilities of individual sensors. For example, if the measured data of individual sensors deviates significantly, i.e. outside the variation range to be expected statistically, there is a corresponding reduction in capability.

However, it must be remembered that a reduction in the capability of a sensor can only be detected after it has occurred. Apart from trend indication in the case of slow degeneration, it is not possible to make any prediction of the future perception capability in relation to the state uncertainty.

### 20.3.3 Existence Uncertainty

For the performance of automated driving, existence uncertainty is at least as relevant as state uncertainty. It expresses the probability that the object in the representation of the vehicle's surroundings actually corresponds to a real object. For example, emergency braking of an automated vehicle should only be triggered in the case of a very high existence probability for a detected obstacle.

While the estimation of state uncertainties using Bayes estimation methods is well founded in theory, the existence probability in today's systems is still mostly determined on the basis of a heuristic quality measure. An object is taken as confirmed if the quality measure exceeds a sensor- and application-dependent threshold. For example, the quality measures are based on the number of measurements that have confirmed the object, or simply the interval between the initialization of the object and the current point in time. Often the state uncertainty of the object (Sect. 20.3.2) is also used for the validation.

An approach with a better theoretical foundation is the estimation of a probability-based existence probability. This firstly requires a definition of the specific object existence. While in some applications, all real objects are taken to be existent, the object existence can also be limited to the objects that are relevant in the current application. Additionally, a limitation to the objects that can also be detected with the current sensor setup is also possible. In contrast to a threshold procedure, this determination of the existence probability enables a probability-based interpretation option. For example, an existence probability of 90 % means that there is a 90 % probability that the measurement history and the motion pattern of the object were created by a real object. Consequently, the action planning of the automated vehicle can use these probabilities when evaluating alternative actions.

A known algorithm for calculating an existence probability is the Joint Integrated Probabilistic Data Association (JIPDA) procedure, also based on the Bayes filter, which was first introduced in 2004 by Musicki and Evans [15]. This procedure additionally uses the detection and false alarm probabilities of the sensors, which are presumed to be known.

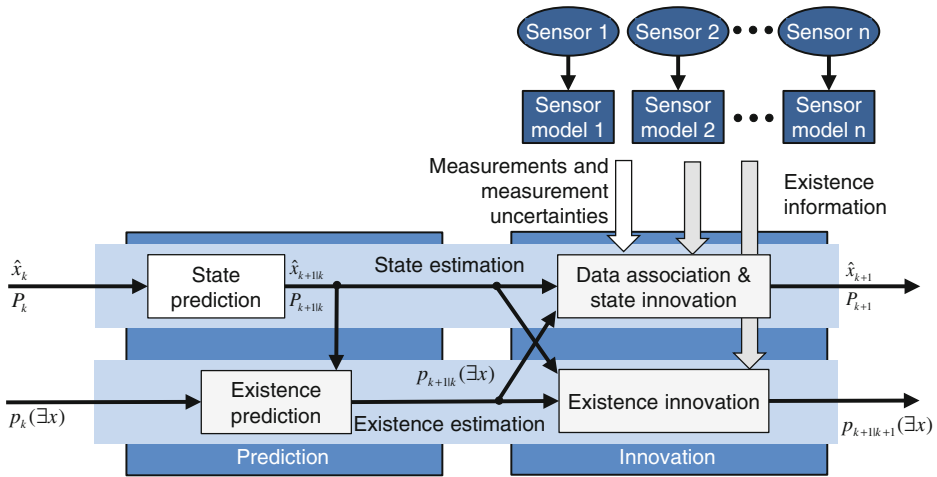
The calculation of the current object existence probability is performed similarly to state estimation in the Kalman filter in a prediction step and an innovation step. The existence prediction is performed using a Markov model of the first order. The predicted existence of an object is given by the Markov chain

$$p_{k+1|k}(\exists x) = p_S p_k(\exists x) + p_B p_k(\bar{\exists} x)$$

whereby the probability  $p_s$  represents the persistence probability of the object and  $p_B$  the probability for the occurrence of an object in the sensor capture area. Consequently, the probability for the disappearance of an object is given by  $1 - p_s$ . In the innovation step, the a posteriori existence probability  $p_{k+1}(\exists x)$  is calculated. It essentially depends on the number of current measurements that confirm the existence of the object.

As the persistence probability of an object depends on the current object state and the posteriori existence probability in turn depends on the data associations, the JIPDA filter can be interpreted as the coupling of two Markov chains shown in Fig. 20.4. The upper Markov chain represents the state prediction and innovation known from the Kalman filter, while the lower Markov chain represents the prediction and innovation of the existence probability. For details on the JIPDA procedure and its specific formulation with regard to the applications in automotive applications, see [16] for example. Current multi-object tracking procedures only developed in the last few years also enable integrated object-specific existence estimation. For further information on this, see [9, 17, 18, 19–21].

With regard to the functional behavior of the existence estimation, the same limitations apply as with the state estimation. A probabilistic measure for the specific existence of the object is continuously supplied. Therefore, sensor failures during operation can also be detected reliably in this uncertainty domain. However, a prediction of future capability is not possible here either.



**Fig. 20.4** Structure of a JIPDA filter as coupling of two Markov chains. Image rights: Author has copyright

### 20.3.4 Class Uncertainty

Classification procedures for determining the object class, i.e. the determination of semantic information, are structured very sensor-specifically. Due to the significantly higher information content, image-based procedures are more common in the classification sector. A differentiation is made for learning procedures in which the classifier trains offline using positive and negative examples and then can recognize the trained object classes to a greater or lesser extent during online operation. The characteristics used in the training are either specified or are implicitly generated themselves in the learning process. Methodically, two basic approaches have established themselves in the learning procedures. On the one hand there are cascaded procedures based on Viola and Jones [22] or methods based on different neuronal networks [23, 24].

A more classical but also common approach is to specify from the sensor data as many deterministic characteristics as possible for the different classes, such as length, width or speed, and to determine for these the class-specific statistical variation areas. The mean value of the individual characteristics, including the variation range, is approximated by means of a normal distribution, for example. Following this, based on the current measured values and the known characteristic distributions, the most probable class from a Bayesian perspective is determined. If different sensors are to be used in combination which can each only capture individual characteristics of the total set, the Dempster-Shafer theory [12] can be used for the class determination, because it allows “non-knowledge” to be considered as well. However, these procedures are generally less powerful than learning procedures, and will presumably continue to diminish in importance as a result.

A disadvantage of all the classification procedures named is that no theoretically substantiated probabilities can be determined for the current quality of the classification. At present no comprehensive theoretical basis exists for this. The output of the classifiers is currently only an individual reliability measure that can be standardized to the value range of 0–1. It does not represent a probability in the narrower sense, and therefore different algorithms are not comparable in this regard. Image-based trained classifiers differ so greatly from characteristic-based procedures comprising lidar and radar sensors that standardized treatment of them will not be easy to achieve.

### 20.3.5 Summary

The explanation in the previous sections make clear that the machine perception is made up of three fundamental uncertainty domains, namely the state uncertainty, the existence uncertainty and the class uncertainty. All domains have a direct influence on the capability of the machine perception. If the uncertainties are too great, whereby it must be defined which uncertainties are tolerable for specific functions, it is no longer possible to drive an automated vehicle reliably.

**Table 20.1** Uncertainty domains of machine perception and their methodical handling

	State uncertainty	Existence uncertainty	Class uncertainty
Characteristic	Uncertainty in the state variables such as object position, object speed, etc.	Uncertainty whether an object captured by a sensor really exists	Uncertainty of class membership (e.g. passenger car versus truck)
Cause	Stochastic measuring error of the sensor technology used	Detection uncertainties of individual sensors, e.g. camera, lidar or radar	Classification uncertainties of the algorithms/limitations of individual sensors
Modeling	Probabilistic; expected value with variances/covariances	Probabilistic by means of detection probabilities	No persistent method; at present mainly heuristic
Methods	Closed theory via general Bayes filter (e.g. Kalman filter variant)	Closed theory coupled with estimate of state uncertainty (e.g. JIPDA filter)	Feature-based: Bayes, Dempster-Shafer Learning-based: Neuronal networks, cascaded procedures (Viola Jones, etc.)
Prediction of the future	Generally no; limited possibility using trend indications	Generally no	Generally no

What is problematic is that a future higher uncertainty, and thus a greater error probability, cannot be predicted in time. While the currently known methods for estimating state and existence uncertainties do enable a current estimation of the capability of the machine perception, in principle it is not possible to predict degeneration in the capability of individual sensors or even a failure of components. Only a trend indication is possible. Table 20.1 summarizes the results once more.

## 20.4 Implications for the Machine Perception Capability Prediction

As was explained and substantiated in the previous sections, the future development of the machine perception capability of an automated vehicle cannot be predicted with sufficient confidence. By no means can the perception capability for the period of 5–10 s required to transfer the driving duties to a human be reliably predicted under all circumstances, as is stipulated as a backup option in highly-automated driving. Moreover, a fully-automated vehicle would have to be able to achieve an intrinsically safe state autonomously, for which even a longer period of time would be required in some cases than for the case of a driver taking over. Although there are certainly a number of options for predicting the future limitation of the perception capability based on external conditions such as

imminent camera glare due to the low position of the sun and sensor limitation due to the onset of rain, snow or fog banks, these are special scenarios that also require extremely reliable context information. Therefore, in principle, the prediction of the perception capability is not a general option for ensuring the necessary reliability for automated driving.

However, as described above, there already exist theoretically substantiated methods and procedures for continually monitoring the current machine perception capability and being able to detect system failures and the degradation of individual components reliably and quickly. Therefore, machine perception systems must be designed in such a way that sensor redundancy is provided which ensures that sufficient perception capability remains either until the transfer to the driver or, in the case of a fully-automated vehicle, until an intrinsically safe state is achieved if individual components break down. Thus, a complete failure of the machine perception must not occur.

Redundancies such as these are basically provided by multi-sensor systems that are used in parallel and combine information from various sensors and sensor principles. For example, if radar and lidar sensors are incorporated, they both supply distance measurement data, but of different quality and in a different sensor capture area. The weather dependencies of the sensor principles are also different. However, due to the similarity of the measured data, they can provide mutual support or also mutual compensation if a component breaks down, with a slight loss of measuring quality for the overall system. Additionally, only through this usage of independent sensor principles is it possible to achieve the highest safety level in accordance with ASIL D, which is required for the operational safety in automated driving.

A redundancy can also be planned and provided easily for cameras. For example, if a camera in a stereo camera system breaks down, the second camera of the stereo system remains available for the classification tasks and the detection of road markings. Only a distance estimate is then no longer available from stereo data and would have to be compensated by means of lidar or radar sensors, for example. Of course, a prerequisite for this redundancy is that the processing hardware and the underlying software of the individual cameras would have to be set up independently, i.e. redundantly. Alternatively, an additional mono camera could be incorporated, including its own processing hardware and software. Therefore, redundancy concepts such as these can enable a minimum perception capability to be always maintained in the automated vehicle, even if individual components break down.

Automated vehicle control is based on the current machine perception and on the prediction of the current traffic situation. With the state of the technology, the latter is essentially performed by means of a simple prediction of the current motion behavior of the objects into the future. Due to the large number of possible and non-predictable events, especially the reactive actions of other road users, the uncertainties increase so starkly after around 2–3 s that reliable trajectory planning is no longer possible on this basis. Therefore, the situation prediction cannot reliably bridge the period for transferring the vehicle control back to the driver in highly-automated driving, or for achieving an

intrinsically safe state in fully-automated driving, if the machine perception no longer continuously updates the vehicle environment model.

However, due to his/her driving experience, a human is also only capable of anticipating the overall situation for around 2–3 s into the future with a degree of reliability [25]. But, because a human perceives and interprets his/her environment quasi-continuously, this brief prediction horizon is completely sufficient for reacting adequately and de-escalatory in practically all situations, and for avoiding accidents as a general rule. This should also be possible for automated vehicles, whereby here of course additional latencies and uncertainties in the perception must be considered. The prerequisite, as mentioned above, is a guaranteed minimum capability for the machine perception.

However, for the overall operation, it is essential that the automated vehicle does not put itself into a technically insolvable situation in the first place. The permissible criticality of the situation must always correspond to the current machine perception capability. What must be considered in particular here are suddenly occurring failures and the resulting spontaneous reduction in the machine perception capability. Within the relatively reliable prediction period for the situation development of 2–3 s, the automated vehicle must be able to adjust its driving behavior to the altered machine perception capability. A simple example would be driving on its own lane. If the sensor range is reduced by technical failures or weather factors, the vehicle must be able to adjust its speed to the current situation within the validity of the prediction, and this represents a reliably solvable technical problem.

While this simple situation is easy to describe and analyze, at present it is generally not known how critical situations arise, and what distinguishes these in terms of the capability of an automated vehicle. In any case, in establishing the reliability of automated vehicles, driving a predefined number of kilometers does not ensure that the resulting dataset contains all the possible critical situation developments (episodes). Consequently, it is not possible to ensure the operating reliability in this way, regardless of the fact the mileage required to statistically prove the very low error rates would be neither practically nor economically feasible.

Therefore, a possible research task in the future would be to find a suitable mathematical representation of random episodes that would then provide the range of all possible episodes. Based on this description, what are known as Monte Carlo simulations, for example, can be used to structure the entire episode range into critical and uncritical subareas, in order to draw a conclusion about required specific tests. A possible methodical approach for this would be rejection sampling, whereby every sample represents a complete episode. Starting from basic episodes, which are distinguished by different road types (1, 2, or 3 lanes per direction, oncoming traffic), for example, or by the number of vehicles in the near vicinity, similar situations are generated through the statistical variation of the episode parameters. When there are a sufficient number of samples, it is to be expected that the episode range has been completely covered. In the process, every episode used is tested for its physical feasibility, and irrelevant episodes are discarded. The remaining episodes are then tested as to whether, for example, critical time

gaps or spaces arise between objects. The suitable criteria for this must also be defined. The identification and prioritizing of critical situations are carried out by means of subsequent clustering in the episode range.

The goal of such a procedure would be to use this hierarchical approach to determine a quantity of potentially critical episodes that is as complete as possible but still manageable. This would then be analyzed using simulated data as to its controllability by the highly-automated system at different levels of machine perception capability. For example, individual ranges, capture angles and detection rates could be modeled for a sensor setup in the vehicle in order to then systematically analyze the consequences for the behavior of the vehicle for critical episodes. This analysis can be carried out initially for a fully-functioning system, and then under the assumption of a failure of individual components.

Another research question that is open is the possibility of a more reliable situation prediction that would use context information and hypotheses about the future behavior of the road users to enable longer prediction periods. Such a procedure would be justified to the extent that our entire traffic system is based on the cooperation of the road users. Of course, a disadvantage of this would be that uncooperative behavior, or simply the errors of other road users, could not be expected nor included in the action planning of an automated vehicle. In this respect, such approaches do not enable a significant extension of the reliably predicted time period, but they can still support planning algorithms. Additionally, it should be noted that in many situations manual drivers do not have a chance to react appropriately when other drivers do not behave correctly or make unforeseen driving errors. Therefore, excessive demands should not be made of automated vehicles in this regard. However, naturally this is also a question of society's consensus with respect to the permissible potential risk involved in a new technology.

---

## 20.5 Summary

The existing methods for the state and existence estimations are based on a closed, well-founded theory and enable in-line, reliable evaluation of the current quality of the machine perception capability. This makes it possible to detect complete failures of individual sensors as well as a gradual degeneration in the sensor technology and/or perception.

However, the procedures do not enable a prediction of the future perception capability, and only a linear extrapolation of detected trends is conceivable. The reliability and quality of the evaluation of the machine perception capability depends on the available sensor models, and error models in particular, which are sensor- and manufacturer-specific. The perception systems alone do not have a sufficient prediction capability that could reliably cover a time horizon of between 5 and 10 s, as is currently envisaged for returning the highly-automated system to the driver. However, this is probably not even necessary for the reliable behavior of an automated vehicle. What is decisive for the controllability of



situations during automated driving is a sufficient number of physically implementable and reliable trajectories for the automated vehicle. These are essentially defined by the spatial proximity of limiting objects to one's vehicle and the available drivable free space. Therefore spatial proximity must be incorporated into key figures for evaluating criticality, while also taking into account uncertainties in the perception and the number of physically possible, reliable trajectories. The currently available machine perception capability must also be considered here. Such sufficiently coordinated and theoretically founded criticality measures do not exist at present.

A situation prediction into the future over a period of 2–3 s will not provide a conclusive result in purely model-based, probabilistic extrapolation, as every development of the situation becomes possible after this period. A possible approach, and a future research question, would be the context-related, hypothesis-based temporal extrapolation based on known and evaluated situations stored in a knowledge base. With an existing knowledge base, the current situation can then be evaluated continuously with regard to the assumed outcome. No reliable methods exist for this at present, and in some cases not even any ideas of how it could be implemented. However, this seems to be a path which can be taken.

The advances in situation prediction are extremely difficult to foresee. However, significantly more capable methods can presumably be expected only in a time horizon of 10 years or more.

**Open Access** This chapter is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, duplication, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, a link is provided to the Creative Commons license and any changes made are indicated.

The images or other third party material in this chapter are included in the work's Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in the work's Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt or reproduce the material.

---

## References

1. Damböck, D.; Bengler, K.; Farid, M.; Tönert, L.: Übernahmezeiten beim hochautomatisierten Fahren. Tagungsband der VDI-Tagung Fahrerassistenz in München, Jahrgang 15, pp. 16ff, 2012.
2. Gold, C.; Damböck, D.; Lorenz, L.; Bengler, K.: "Take over!" How long does it take to get the driver back into the loop? In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 57, No. 1, pp. 1938-1942, 2013.
3. Winner, H.; Hakuli, S.; Wolf, G. (Eds.): Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort, 2. Auflage (2012), Teubner Verlag.
4. Nuss, D.; Reuter, S.; Konrad, M.; Munz, M.; Dietmayer, K.: Using grid maps to reduce the number of false positive measurements in advanced driver assistance systems. IEEE Conference on Intelligent Transportation Systems (ITSC), 2012, pp.1509-1514.

5. Bouzouraa, M.; Hofmann, U.: Fusion of occupancy grid mapping and model based object tracking for driver assistance systems using laser and radar sensors. *IEEE Intelligent Vehicles Symposium (IV)*, 2010, pp. 294-300.
6. Coué, C.; Pradalier, C.; Laugier, C.; Fraichard, T.; Bessiere, P.: Bayesian Occupancy Filtering for Multitarget Tracking: an Automotive Application. *International Journal of Robotics Research*, 2006, Vol. 25.1, pp.19-30.
7. Konrad, M.; Szczot, M.; Schüle, F.; Dietmayer, K.: Generic grid mapping for road course estimation. *IEEE Intelligent Vehicles Symposium (IV)*, 2011, pp. 851-856.
8. Schmid, M.R.; Mählich, M.; Dickmann, J.; Wünsche, H.-J.: Dynamic level of detail 3d occupancy grids for automotive use. *IEEE Intelligent Vehicles Symposium (IV)*, 2010, pp. 269-274.
9. Mahler, R.: *Statistical Multisource-Multitarget Information Fusion*. Artech House, Boston, 2007.
10. Bar-Shalom, Y.; Tse, E.: Tracking in a cluttered environment with probabilistic data association. *Automatica*, 1975, Vol. 11, pp. 451-460.
11. Bar-Shalom, Y.; Fortmann, T.: *Tracking and Data Association*. Academic Press, Boston, 1988.
12. Shafer, G.: *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
13. Kalman, R.: A new approach to linear filtering and prediction problems. *Transactions of the ASME – Journal of Basic Engineering*, 1960, Vol. 82, pp. 35-45.
14. Julier, S.; Uhlmann, J.; Durrant-Whyte, H.: A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 2000, Vol. 45.3, pp. 477-482.
15. Musicki, D.; Evans, R.: Joint Integrated Probabilistic Data Association: JIPDA, *IEEE Transactions on Aerospace and Electronic Systems*. 2004, Vol. 40.3, pp. 1093-1099.
16. Munz, M.: *Generisches Sensorfusionsframework zur gleichzeitigen Zustands- und Existenzschätzung für die Fahrzeugumfelderfassung*. PhD Thesis, Ulm University, 2011.
17. Mahler, R.: Multitarget Bayes filtering via first-order multitarget moments. *IEEE Transactions on Aerospace and Electronic Systems*, 2003, Vol. 39.4, pp. 1152-1178.
18. Nuss, D.; Stuebler, M.; Dietmayer, K.: Consistent Environmental Modeling by Use of Occupancy Grid Maps, Digital Road Maps, and Multi-Object Tracking. *IEEE Intelligent Vehicles Symposium (IV)*, 2014.
19. Reuter, S.; Wilking, B.; Wiest, J.; Munz, M.; Dietmayer, K.: Real-time multi-object tracking using random finite sets. *IEEE Transactions on Aerospace and Electronic Systems*, 2013, Vol. 49.4, pp. 2666-2678.
20. Vo, B.-T.; Vo, B.-N.: Labeled Random Finite Sets and Multi-Object Conjugate Priors. *IEEE Transactions on Signal Processing*, 2013, Vol. 61, 3460-3475.
21. Reuter, S.; Vo, B. T.; Vo, B. N.; Dietmayer, K.: The labeled multi-Bernoulli filter. *Signal Processing*, *IEEE Transactions on*, 62(12), pp. 3246-3260, 2014.
22. Viola, P.; Jones, M.: Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition*, 2001. *CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on* (Vol. 1, pp. I-511). IEEE.
23. Egmont-Petersen, M.; de Ridder, D.; Handels, H.: Image processing with neural networks—a review. *Pattern recognition*, 35(10), pp. 2279-2301.
24. Zhang, G. P. : Neural networks for classification: a survey. *Systems, Man, and Cybernetics, Part C: Applications and Reviews*, *IEEE Transactions on*, 30(4), pp. 451-462.
25. Godthelp, H.; Milgram, P.; Blaauw, G. J.: The development of a time-related measure to describe driving strategy. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 26(3), 1984, pp.257-268.