2004 IEEE Intelligent Vehicles Symposium
University of Parma
Parma, Italy • June 14-17, 2004

Standard Platform for Sensor Fusion on Advanced Driver Assistance System using Bayesian Network

Naoki Kawasaki, Uwe Kiencke

Abstract— **In this paper, a new architecture for sensor fusion for advanced driver assistant system (ADAS) is proposed.**

This architecture is based on Bayesian Network and plays the role of platform for integrating various sensors such as Lidar, Radar and Vision sensors into sensor fusion systems. This architecture has the following 3 major advantages:

(1) It makes structure and signal flow of the complicated fusion systems easy to understand

(2) It increases the reusability of the sensor algorithm modules

(3) It achieves easy integration of various sensors with different specifications

These advantages are confirmed by vehicle test.

I. Introduction

UNTIL now, Advanced Driver Assistance System (ADAS) such as ACC (Adaptive Cruise Control) and FCAAS (Frontal Collision Avoidance Assistance System) have been developed. These systems are using frontal monitoring sensors, such as millimeter wave (MMW) radar, laser radar (LIDAR) or vision sensor (video camera), to recognize preceding cars and other objects.

It is very difficult to fulfill very high accuracy requirements from those systems by only single sensor. Therefore the sensor fusion method is a hot topic. It takes advantage of the fact, that different sensors have different advantages by weighting each sensor depending on its performance. For example, the fusion of MMW radar and vision sensor is an effective combination [1]. Because the former has a good accuracy in longitudinal distance measurement, but poor in lateral while the properties of the latter are the exact opposite. Thus, this combination provides good accuracy in both positional measurements.

Many sensor fusion methods have been suggested in the literature using averaging, majority rule, Kalman filtering or Bayesian estimation [2][3], and higher accuracy is achieved already. But those conventional algorithms tend to be quite ad hoc. They depend on the specific combination of sensor devices, algorithms or application systems they were designed for. In the future, there will be many new sensors and upgraded recognition algorithms, which will cause a huge number of combinations. Therefore, the modular and general fusion architecture is an important topic.

In this paper, sensor fusion architecture using a BN is presented. This general probabilistic method can treat all kinds of relations with statistical data in the estimation.

In chapter 2, general usage of BN estimation is explained. Chapter 3 introduces the BN into the application of MMW radar and vision sensor fusion. Then chapter 4 explains the new architectural structure we propose in this paper.

Manuscript received November 28, 2003.

Naoki Kawasaki is with DENSO CORPORATION, Kariya, Aichi, 448-8661 Japan (e-mail: kawasaki@rd.denso.co.jp).

Uwe Kiencke is head of the Institute of Industrial Information Technology, Universität Karlsruhe, D-76187 Karlsruhe, Germany (e-mail: kiencke@iiit.uni-karlsruhe.de).

II. Bayesian Network Estimation

A. How to build a model

A "Bayesian Network (BN)", which is also called "belief net" or "causal network", is a visualized modeling technique for statistical dependencies between variables and also works as a probabilistic estimation machine [4][5].

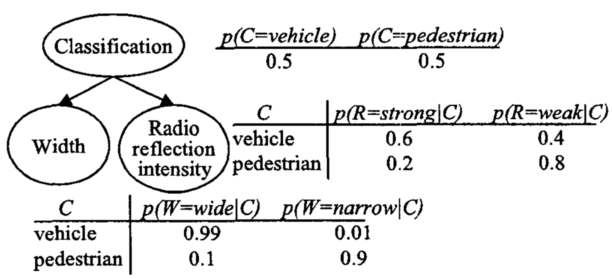


Fig.1. Sample BN model for classification estimation

Fig. 1 shows an example of a BN model. In common BN model consists of nodes and arrows (DAG: Directed Acyclic Graph). A node is a probabilistic variable, which consist of exhaustive and mutually exclusive states. An arrow is the causal relationship between nodes, from cause to effect.

In this example, all variables are implemented with 2 states and it is assumed that target object is either a vehicle or a pedestrian.

Every node has a database inside, which shows the strength of dependency between the parent nodes and itself. In Fig.1, those are expressed as tables *p(C)*, *p(R|C)*, and *p(W|C)*, and called conditional probability tables (CPT). Those databases are built from historical knowledge about how often the state occurred given fixed parent node's condition. In this paper, lower case "*p(X)*" means database of BN, such as CPT.

B. How to calculate the estimation

Given a model structure and conditional probability tables (CPT), the probabilistic estimation is now available. Probabilities are assigned to all states of the classification nodes. Namely, if *P(C=pedestrian|given data)* > *P(C=vehicle|given data)*, the target is very likely a pedestrian (*C=ped*). For example, if the sensor detects that the radio reflection intensity is weak in Fig.1, *P(C=ped|R=wea)* must be calculated.

To calculate it, the fundamental rule for probability calculus can be used, which is also called Bayes' rule.

$$P(a,b) = P(a \mid b)P(b) = P(b \mid a)P(a)$$
$$P(b \mid a) = \frac{P(a,b)}{P(a)} \quad \left(= \frac{P(a \mid b)P(b)}{P(a)} \right) \tag{1}$$

It is expressed it as the following general inference form

$$P(Searched|Known) = \frac{\sum_{Free} P(Searched, Known, Free)}{P(Known)} \tag{2}$$

where *P(Known)* is scalar normalization factor. Now we get

$$P(C{=}ped \mid R{=}wea) = \frac{\sum_{W} P(C{=}ped, R{=}wea, W)}{P(R{=}wea)}. \tag{3}$$

Here we have to calculate the joint probability function *P(C,R,W)* with 3 variables *C*, *R*, and *W*. BN offers a very useful theorem called "chain rule" for a more compact representation of *P(C,R,W)*.

$$P(all\ nodes) = \prod_{all\ nodes} p(node \mid parent\ nodes\ of\ that\ node) \tag{4}$$

Using this rule, the joint probability function is described as

$$P(C,R,W) = p(C) \cdot p(R \mid C) \cdot p(W \mid C) \tag{5}$$

where the right terms are the given conditional probability tables (CPT). Now (3) derives following equation.

$$P(C{=}ped \mid R{=}wea) = \frac{\sum_{W} p(C{=}ped) \cdot p(R{=}wea \mid C{=}ped) \cdot p(W \mid C{=}ped)}{P(R{=}wea)}. \tag{6}$$

From this equation the probability *P(C=ped| R=wea)* can be calculated as 0.667, meanwhile *P(C=veh| R=wea)* is 0.333. Comparing them, the result *P(C=ped|R=wea)* > *P(C=veh| R=wea)* can be given,

By introducing additional acquired data, the probability estimation may be improved. If the width is additionally observed as narrow, the answer can now be calculated as

$$P(C{=}ped \mid R{=}wea, W{=}nar) = \frac{p(C{=}ped) \cdot p(R{=}wea \mid C{=}ped) \cdot p(W{=}nar \mid C{=}ped)}{P(R{=}wea, W{=}nar)}. \tag{7}$$

Now *P(C=ped|R=wea, W=nar)* is 0.994, which is more accurate than before.

If a sensor's performance changes, only the corresponding node's CPT has to be changed and other CPTs or model structure can be left untouched. In Fig.1, the utilization of a different radar sensor might cause some change in the reflection intensity resulting in a different CPT of that node. However the structure of the model and other nodes' tables do not have to be modified.

To treat continuous variables such as length or position, probability density functions (PDF) can be assigned to these probabilities. Instead of CPT, conditional probability density function (CPDF) is used. In this paper, PDF and probability function are expressed as upper case "*P(X)*", while the databases of BN such as CPT and CPDF are expressed as lower case "*p(X)*".

C. BN sensor fusion

Fig.2(a) shows a typical example of the sensor fusion BN model structure. There are two sensors A and B which both detect the continuous variable "lateral position" *Xa* and *Xb* respectively, to estimate true lateral position. CPDF is stored in each node. The example of CPDF *p(X)* and *p(Xa|X)* are shown in Fig.2(b) and (c).

p(Xa|X) shows the statistically acquired probability distribution database: "if the true lateral position *X* is given, the sensor A outputs the data *Xa* in such distributions."

Sensor CPDF could be expressed by a combination of Gaussian and uniform distribution. If the sensor detects objects with a reliability of 50%, the half of the sensor outputs is wrong and has no relationship between the aimed target's true position *X* and the acquired data *Xa*. This can be expressed by a uniform probability distribution.

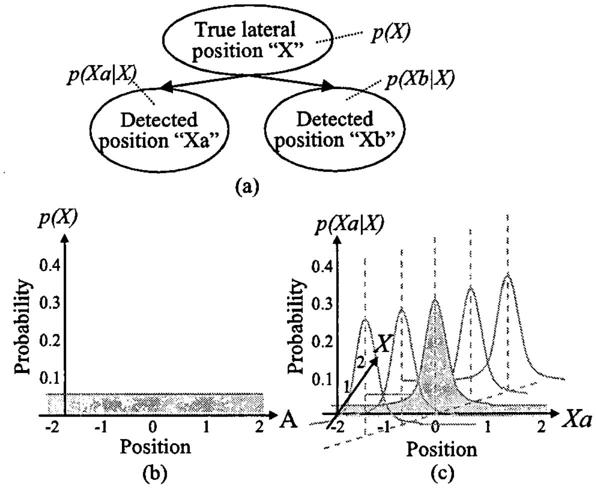


Fig.2. BN causality model for lateral position estimation

D. *Likelihood Merger*

Basic sensor fusion is done by a BN such as in Fig.2. From this function of the BN model and calculation, an important idea of the fusion can be evolved, called "likelihood merger."

The equation of the fusion calculation to acquire the PDF $P(X|Xa=a,Xb=b)$ is shown below.

$$
\begin{aligned}
P(X \mid Xa{=}a, Xb{=}b) &= \frac{P(X, Xa{=}a, Xb{=}b)}{P(Xa{=}a, Xb{=}b)} \\
&= \frac{p(X)\cdot p(Xa{=}a \mid X)\cdot p(Xb{=}b \mid X)}{\int_X P(X, Xa{=}a, Xb{=}b)\, dX} \\
&= \frac{1}{Z}\, p(Xa{=}a \mid X)\cdot p(Xb{=}b \mid X) \qquad (8)
\end{aligned}
$$

where $p(X)$ is assumed to be constant, and the denominator is also constant. So they are expressed as a normalization factor Z.

It can be said from (8) that the fusion result $P(X|Xa=a,Xb=b)$ is calculated by merging $p(Xa=a|X)$ and $p(Xb=b|X)$, which are the database CPDFs $p(Xa|X)$ and $p(Xb|X)$ where Xa and Xb are given as a and b. In other words, by fixing variable Xa, the CPDF $p(Xa=a|X)$ is transferred into "likelihood" of X. It means that "how likely X is when the sensor A detected a." In this paper, this kind of likelihood is also called "unitable CPDF" depending on X.

Fig.3 shows the idea of this "likelihood merger." There are two unitable CPDFs $p(Xa=a|X)$ and $p(Xb=b|X)$ in Fig.3(a), which are calculated from CPDFs and detected data a and b. They are simply multiplied to get the resulting PDF $P(X|Xa=a,Xb=b)$ like Fig.3(b). The highest peak of that is the most likely result of X.

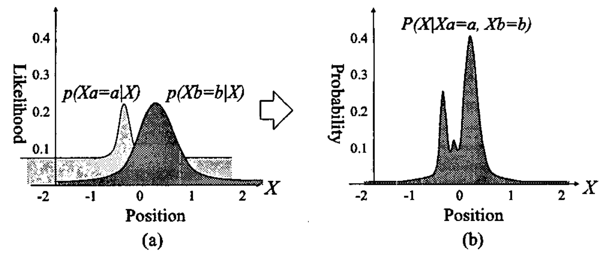


Fig.3. Example of likelihood merging

If the third sensor is added in this system, the third child node will be added to the model in Fig.2(a), while the existing nodes are untouched. Then the third unitable CPDF will also appear in Fig.3(a), and the peak of the answer PDF Fig.3(b) will be steeper. This means that the estimation will be more precise.

This idea is the essence of the Bayes theorem, and it will be used in chapter IV.

III. Bayesian Network Sensor Fusion Model

This BN estimation technique is applied in a real sensor fusion practice. The experimental subject is a data fusion system of MMW radar plus vision sensor for tracking vehicles on a road.

A. *System configuration and inputs*

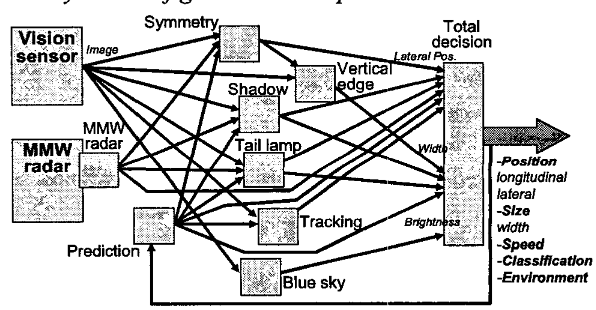


Fig.4. Data flow diagram of experimental sensor system

Fig.4 shows the data flow diagram of the experimental sensor system using two devices, MMW radar and vision sensor. There are many object-detection algorithms, which are called recognition IPs (Intellectual Properties). They output the target object properties such as lateral center position or width. All outputs of the recognition IPs are collected and fused at "Total decision IP," where BN is implemented. Some image recognition IPs are using MMW IP output information, but it is only for preconditioning.

Here are brief explanations about the recognition IPs, the inputs of the BN:

1) Symmetry IP: searches the symmetry of the object image, and detects the lateral position of the object.
2) Vertical edge IP: sums up the luminance gradient of the image in vertical direction, and detects the width of the

object.

3) Shadow IP: searches the dark spot in the image, and detects the center position/width of the object.

4) Tail lamp IP: searches the corresponding pair of white spot in the image, and detects the center position/width of the object.

5) Tracking IP: looks for the previously detected image of the object in the current image, and detects the center position of the object.

6) Blue sky IP: detects the environmental brightness by scanning the upper part of the image.

7) MMW radar IP: analyzes radar wave reflection and detects the lateral and longitudinal position, and longitudinal relative velocity of the object.

8) Prediction IP: predicts the object position, width, etc. using previous estimation result.

Each IP also outputs the detection intensity value, which means "how far the recognition was successful." For example of the Symmetry IP, it outputs the value "how symmetrical the image was."

B. Bayesian Network fusion model

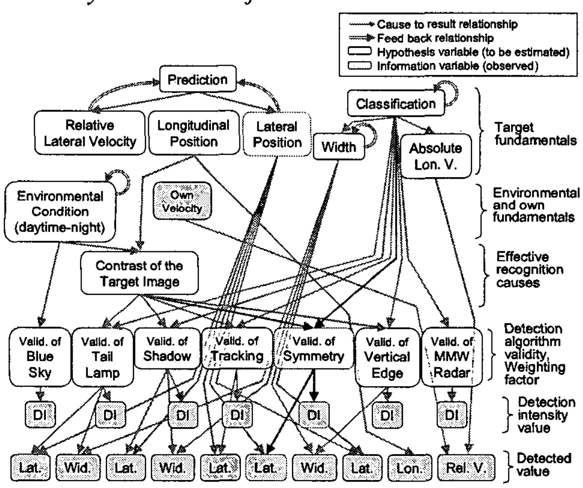


Fig.5. Bayesian Network model for experimental sensor fusion

Fig.5 shows the BN model for this experimental sensor fusion as an algorithm within the "Total decision IP" in Fig.4. This model consists of several layers, which help categorizing of the nodes. Upper layers are the true positions or properties of the object that are the targets of the estimation. Lower layer are the detected values that are caused by upper layer nodes.

The core part of this model is e.g. "Lateral Position" node and its children "Lat." nodes. This structure means just like Fig.2: several IPs detect the lateral position of the same object.

On the other hand, each recognition IP has the Boolean node "detection algorithm validity (weighting factor)," which is the most characteristic point of this model. This factor means "how much the detection algorithm is adequate for the current situation and target object," and "how much this IP is reliable now." If the probability of the state "yes" of this factor is higher, it is more reliable, and the acquired data will have a bigger impact on the final result. If that probability is zero, there is no relationship between the true and the acquired data.

This weighting factor is determined by other related nodes. For example in Fig.5, the Symmetry IP's weighting factor depends on the "Contrast of the Target Image," "Classification" and "Detection Intensity." It means: this IP may be not trustworthy if it is night and the image is not clear, or if the target object is something which is not necessarily symmetrical such as a road structure, or if the symmetry of the image is weak. This model part represents and defines the reliability causality of each IP.

The weighting factor node is also connected with "detected lateral position" node, which is at the lowest layer in Fig.5. This means: it affects the unitable CPDF (likelihood), the impact of detected value on the true value. If it is night condition and the weighting factor is low, the unitable CPDF $p(X=x|A)$ in Fig.3(a) must be adapted according to the condition as shown in Fig.6. The lower the weighting factor is, the more the uniform distribution becomes dominant.

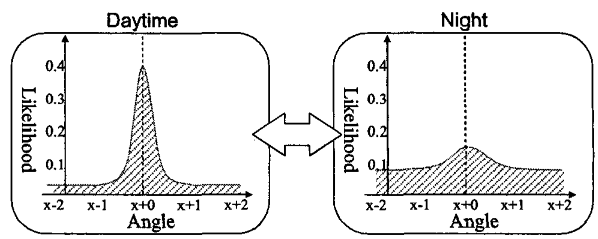


Fig.6. Example of likelihood adaptation

BN allows a clear description of the fusion algorithm strategies. It based on the detection mechanism using causal model, therefore it is very easy to understand, analyze weak points, add/remove sensors, and improve the fusion algorithm. This advantage is very important as the models tend to become huge.

IV. ARCHITECTURAL STRUCTURE

A. Architectural problem

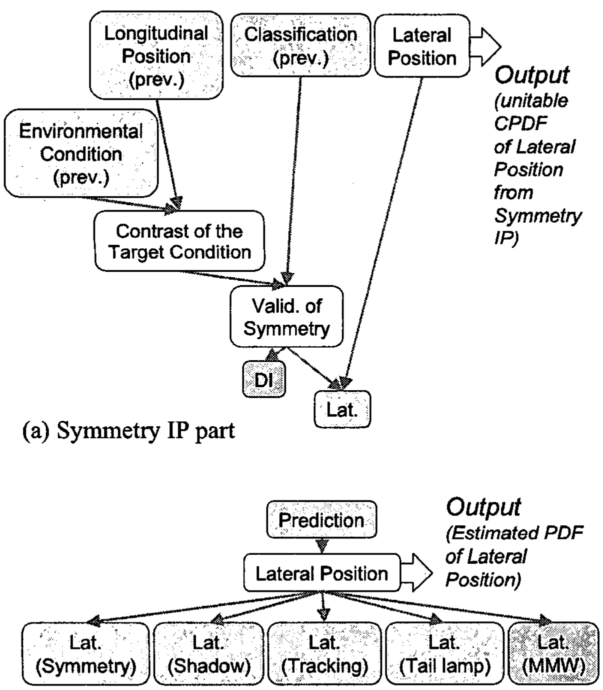
The above BN sensor fusion algorithm achieves good estimation accuracy. However, it still has an architectural problem. If only a single recognition IP algorithm is changed, we also have to change the Total decision IP. This is because the information concerning the recognition IP's performance is not stored in the respective IP, but in the Total decision IP as CPDF.

B. Modular Bayesian Network Architecture

In order to reduce the development effort, a modular BN architecture is proposed.

To eliminate the generation of unitable CPDFs from the Total decision IP, the BN model is separated into recognition IP parts and a Total decision IP part. Fig.7(a) shows the part of the separated models for Symmetry IP as an example of recognition IP parts. This model shows how to output the unitable CPDF as explained in chapter III-B. This means that each IP considers its own accuracy and reliability, using previously estimated and currently acquired information, and sends the unitable CPDF to the Total decision IP. Those functions are now inside of the "Symmetry" block in Fig.4, not in the "Total decision" block.

Fig.7(b) is the partial model extracted from Total decision IP part BN model. This part is now simplified because the unitable CPDFs are now calculated in the recognition IP parts, which means that the function is distributed. Only the "likelihood merger" calculation, as explained in chapter II-D, and the peak search are done in Total Decision IP. The Total decision IP block is now independent from sensor devices. All inputs are likelihoods of the estimation target variables, and Total decision IP does not care what kinds of devices are used.



(a) Symmetry IP part

(b) Total decision IP part (partial)

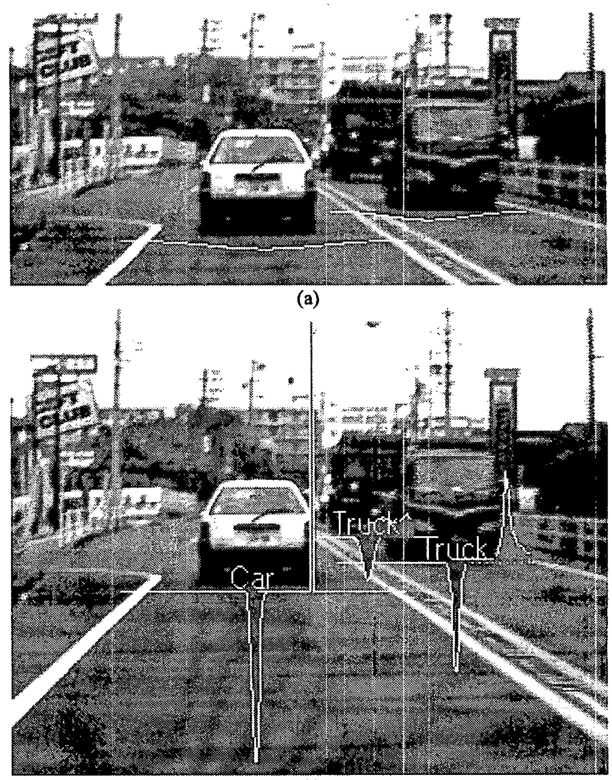
Fig.7. Examples of the separated BN models

This architecture makes the fusion system quite modular. When an IP algorithm changes, the IP developer must only change its CPDF database. This means that the only corresponding recognition IP block in Fig.4 is changed, and the Total decision IP block does not have to be modified.

This modularity offers reusability. If a recognition IP is removed, only corresponding node in Fig.7(b) has to be removed. This means, one of unitable CPDFs is deleted from equation (8). It is very easy operation.

Addition of new sensor or recognition IP is also easily treated in this architecture, if it outputs the data as united CPDF format. Such "platform function" will be an important ability for sensor fusion architecture, because it allows easy integration of various sensors from various sensor suppliers.

V. Experiment and Result



(a)

(b)

Fig.8. Estimation results

The proposed architecture and algorithm was validated on real data.

Fig.8(a) and (b) show the estimated lateral position PDF expressed as a downward graph under the car image. Fig.8(a) is the result using the output of MMW radar IP alone, whereas Fig.8(b) is the fusion result using every IP. In Fig.8(b), the estimated width PDF is also shown as an upward peak, which indicates the right edge position of the object.

The lateral position PDF peek of (b) is steeper than that of (a). It means that the fusion result is more reliable than MMW radar alone. Generally speaking, it was checked that fusion result achieves better accuracy than MMW alone, and almost same accuracy with the conventional fusion algorithm.

It was also checked that such addition or removal of the IP results is easily treated by simply adding or removing the corresponding unitable CPDF from multiplication at Total decision IP part. Therefore it is also said that this fusion algorithm is quite modular.

VI. CONCLUSION

This paper offered new architecture of the sensor fusion by use of modular Bayesian Network.

Bayesian Network describes the fusion system in a causality model, which makes the fusion algorithm easy to understand. This merit is important if the fusion algorithm is big and complicated. It also helps to modify the algorithm when the system is changed. The addition of the new sensors and recognition algorithms will be easily treated by adding new nodes in the existing BN model, while the existing nodes are untouched.

This architecture defines "unitable CPDF" as a standardized output format for all sensors and recognition algorithms, and synthesizes them in a mathematically well-defined, modular way. Therefore, the reusability of each recognition algorithm is greatly increased. Additionally, sensors from different suppliers can be easily combined on this standard platform.

Validation was done using real data. The accuracy of the merged data was as good as conventional fusion algorithm, while the modularity is drastically increased.

Higher accuracy will be achieved mainly by adding new helpful recognition algorithms, or by improving existing recognition algorithms. But a more detailed model and a more precise database of BN will also make the estimation more powerful and accurate.

A disadvantage of BN is their need for calculation performance. In the above application, this problem was reduced by separating the model into the modular Bayesian Networks we proposed. Further research can be done on this topic.

REFERENCES

[1] A. Gern, U. Franke, and P. Levi, DaimlerChrysler Research, Stuttgart, "Robust Vehicle Tracking Fusing Radar and Vision", Intelligent Vehicle Symposium 2000, IEEE

[2] C. Coué, Th. Fraichard, P. Bessière, and E. Mayer, "Multi-Sensor Data Fusion Using Bayesian Programming: an Automotive Application", Intelligent Vehicle Symposium 2002, IEEE

[3] B. Steux, C. Laurgeau, L. Salesse, D. Wautier, "Fade: a vehicle detection and tracking system featuring monocular color vision and radar data fusion", Intelligent Vehicle Symposium 2002, IEEE

[4] R. O. Duda, Pattern Classification, 2nd ed. John Wiley & Sons, 2001, ISBN 0-471-05669-3

[5] F. V. Jensen, Bayesian Networks and Decision Graphs. New York: Springer-Verlag, 2001, ISBN 0-387-95259-4