



XI CONGRESSO NACIONAL DE ENGENHARIA MECÂNICA
DE 07 A 11 DE AGOSTO DE 2022, TERESINA-PI, BRASIL

CONEM2022-0324

ADVANTAGES OF SENSORS AGGREGATION IN AUTONOMOUS CARS

Ehrenbrink, Albano Leo, albano@ehrenbrink.com
Dutra, Max Suell Dr.-Ing., max@mecanica.coppe.ufrj.br

Federal University of Rio de Janeiro - PEM / COPPE / UFRJ - Program of Mechanical Engineering
Centro de Tecnologia, Bloco G/ 204 - PO Box: 68.503 - 21945-970 - Cidade Universitária, RJ, Brasil

Abstract. *This article studies the advantages of using several sensors in autonomous vehicles. Nonetheless, the variation of the uncertainties of different sensors in different scenarios, mainly in the transition between scenarios, causes a decrease in the reliability of the decision-making model. This work shows the advantages and disadvantages of each type of sensor typically used in autonomous car projects, the evolution of assistant drive systems, and an understanding of how the liability increases overtime for these systems. This work also makes a proposition how aggregate sensors inputs for a better interpretation of the scenario with the consequent increase in the robustness of the decision-making model. Finally, this article also highlights the importance of sensor aggregation mainly in everyday situations that present significant challenges to computer vision as well as systems that use: LiDAR (laser imaging, detection, and ranging), RGBD (Red, Green, Blue, and Distance), or stereo vision to determine the distance of objects to estimate the vehicle location known as SLAM (Simultaneous localization and mapping). The solution for these challenges includes the benefit of aggregation of sensors that reads different wavelengths and are susceptible to various obstacles limitations. Despite this, the interpretation of the car's behavior or location can contribute to choosing what sensor is more appropriate to increase the reliability of the decision-making model.*

Keywords: *sensor aggregation, autonomous vehicles, computer vision, LiDAR, SLAM*

1. INTRODUCTION

The autonomous car concept dates to 1956 when presented at the MOTOROMA fair. Over these 66 years, several sensors have been incorporated into the automobile industry to offer comfort and safety to bring us closer to this concept. Still, they are always limited to certain specific functions.

In 2004, the USA agency DARPA (Defense Advanced Research Projects Agency) launched competitions such as the Grand Challenges and the Urban Challenges. These competitions brought significant advances with sophisticated sensors such as LiDAR (laser imaging, detection, and ranging) and solutions for locating the vehicle to make decisions in autonomous systems.

In 2008 China, with its China Development Master Plan, launched the 中国智能车未来挑战赛 (China Smart Car Future Challenge) with advances in computer semantic understanding through computer vision.

This approach to semantic understanding tries to fill the gaps that the lack of information necessary to interpret the environment is present. However, sometimes purely reading the values of the different sensors limits the understanding of the scenario for decision making.

The limitation to achieving higher levels of automation is still present. The leading causes for this limitation are the actual computational limit that some solutions require and the type of sensor used as the primary source of information. Conventional RGB cameras and RGBD distance sensors, or even LiDAR sensors, cannot adequately interpret the scenario without sensor aggregation.

A daily example is in the transition of scenarios such as when entering or exiting tunnels in daylight or when intersecting other cars at night in opposite directions with high beams. Unfortunately, the image's Dynamic Range significantly limits the information that can extract from these images with visual computing.

Another example is LiDAR sensors in rainy or foggy weather or any other obstacle that does not affect driving safety, such as falling leaves or a plastic bag. Due to the number of points acquired by LiDAR, it is impossible to make an adequate semantic interpretation of the object or scenario for the proper decision making.

Even with the improvement of the sensors, the aggregation of different sensors readings brings benefits to the understanding of the scenario and adequate decision-making.

Ultimately the sensor aggregation brings benefits because they are based on different readings and mainly on different wavelengths that compensate for the temporary prejudicated readings of some sensors.

Due to advances in sensors and systems and with several companies working on the massification and lowering the prices of solutions, the NHTSA (National Highway Traffic Safety Administration) ranked the groups of vehicle automation in 2013 to coordinate this initiative into four levels. This classification was updated by the SAE (Society of Automotive Engineers) in 2014 to 6 levels with the standard J3016, ORAD (2021).

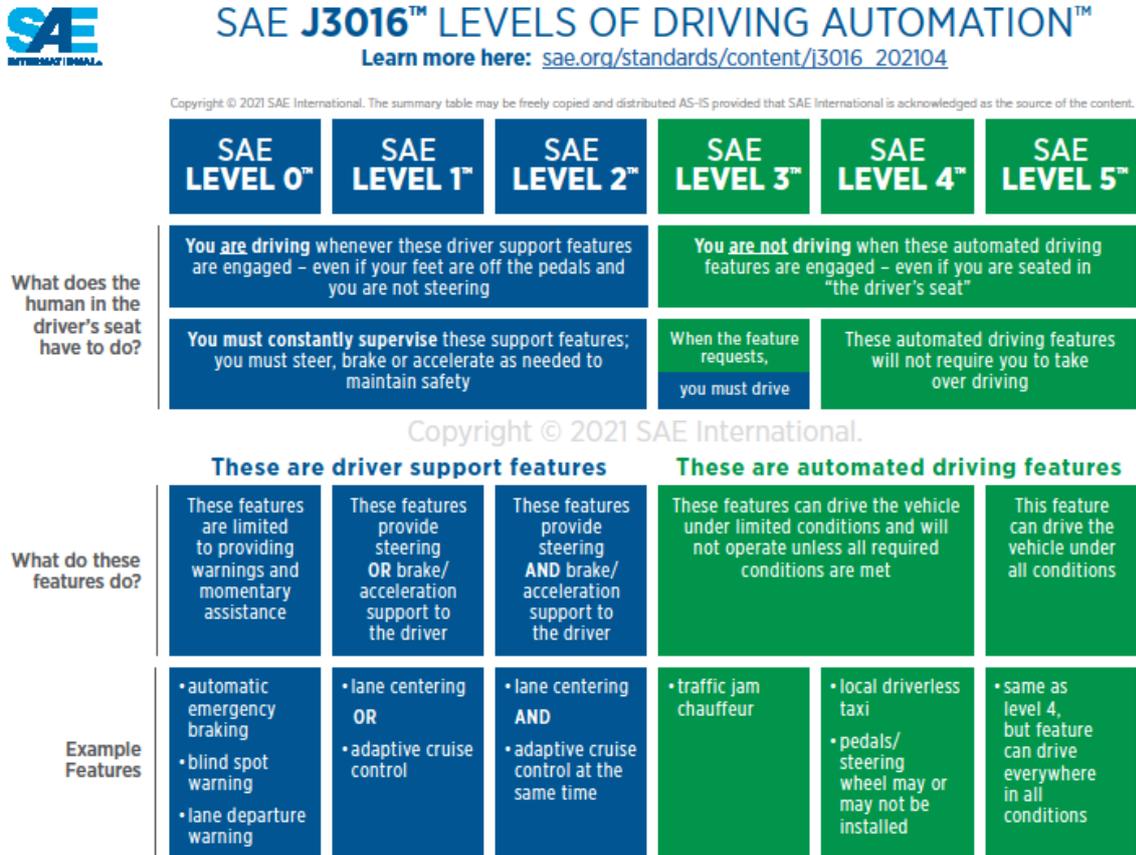


Figure 1. SAE J3016 levels of driving automation

This classification expresses the level of automation by the user's perception and not in the sensors needed to implement it or the engineering model robustness required. Moreover, the use of individual sensors hardly achieves a higher classification due to the lack of information that fills gaps and increases the severity of consequences of high levels of automation.

This article, in Section 2, will list the different sensors, their strengths, and their weaknesses. Then, section 3 will propose a way of semantically aggregating sensors. Finally, in section 4, conclusions and suggestions for future work are presented.

2. SENSORS

2.1. Classical deterministic systems

In automation, the commands rely on the reading of sensors. Depending on the severity of the consequences, as in cars, the sensors must be consistent enough not to trigger a sequence of unwanted events that lead to dangerous situations.

For this reason, engineers develop driver alert systems without acting on the vehicle in the initial phase. For example, park distance control plays a sound or displays a scale to notify the driver of the proximity of other obstacles between 0,2 meters and 1,60 meters (8 inches and 63 inches), Bosch GMBH (2007).

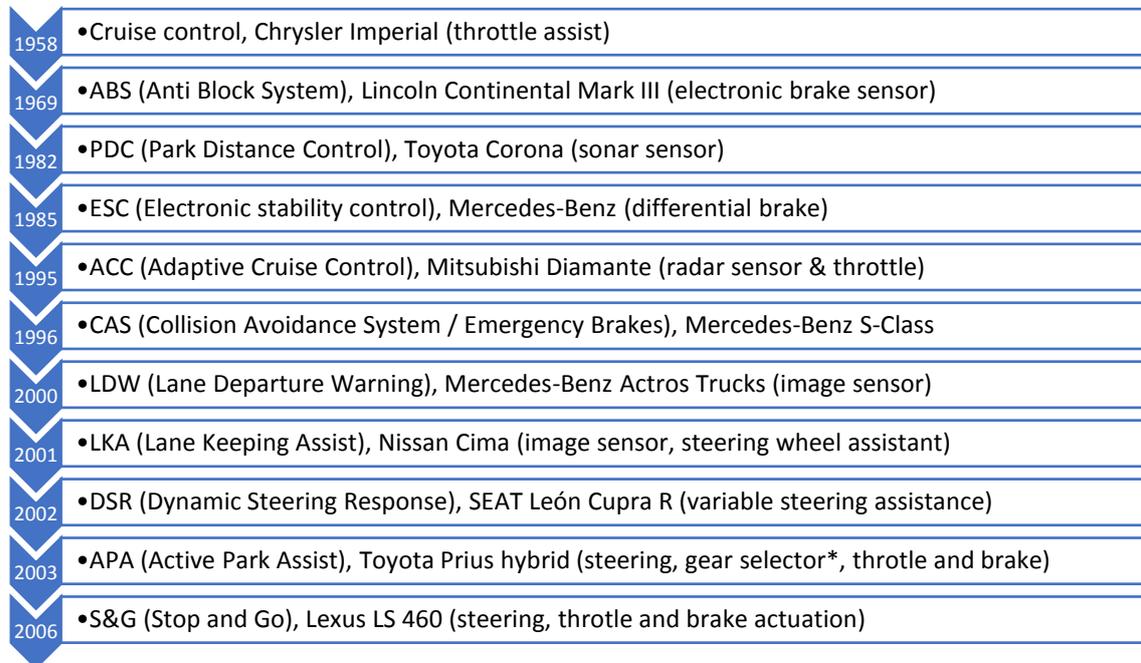
Engineers implement the actuation assistant after the increase in detection reliability. This assistant assumes part of the effort and delegates the authority over the command to the driver.

To increase command authority, scenarios where sensor reading errors occur, are considered. For example, engineers implement systems where misreading leads to an actuation that drives the vehicle to safety before systems where the driver needs to override assistance to bring a safe condition. An example of this is collision avoidance systems, where a wrong reading of an obstacle slows the vehicle while the failure to read an obstacle remains the driver's responsibility.

2.2. The history of sensors

For decades many sensors have been developed to assist the driver. These sensors have two goals: to decrease the driver's workload and increase safety.

The development of sensors starts in the laboratory, evolves into patents, tests into prototypes, and finally, becomes products. The image below chronologically summarizes the implementation of some sensors in vehicles available on the market when overcoming the technical challenges of reliability and production cost.



*Toyota Prius hybrid uses electric motors to reverse speed (does not use a gearbox).

Figure 2. Some vehicle's automation systems evolution 1958 - 2006

In the figure above, just the last two systems actuate the vehicle autonomously in a limited scenario where the speed is low enough to allow steering or the command is small enough that the driver has time to take over the control if necessary.

These two scenarios are not required a profound understanding of the environment. Instead, just identifying the parking space or monitoring the distance of another vehicle in front is necessary.

2.3. Statistical semantic systems

Comparing the vehicles systems evolution with the SAE automation level stands out different approaches.

The systems described in the last section have deterministic characteristics. Each set of reads leads to a single action. The SAE accepts a new approach where statistical models and uncertainty scenarios are conceivable.

This approach allows us to interpret the world semantically and brings a new level of services.

This semantic interpretation approach is essential to the autonomous system. Otherwise, it would require a costly standardization of the scenario like roads, traffic lights, road signs, and avoiding unknown objects with fences. In other words, it will be better to build a railroad.

The statistical approach bothers some parts of society because the uncertainty is now classified as acceptable if the results are safer than vehicles without the system. In contrast, the classical deterministic system tries to avoid any ambiguity.

As mentioned in the introduction of this article, the initial challenge was to drive a vehicle autonomously from point A to point B bypassing some obstacles.

So, it is needed to identify the drivable zone and route the optimal path to drive between the initial position and the desired destination. With GPS (Global Position System) to establish the initial and desired position and LiDAR (Laser Imaging, Detection, And Ranging) to map obstacles, it is possible to specify the initial path, Gao (2018).

When obstacles block GPS signal, it is possible to use IMU (Inertial Measurement Unit) and SLAM (Simultaneous Localization and Mapping) technics. The SLAM uses the LiDAR system to compare distances with a previously recorded track to determine the position.

Driving with IMU, Lidar and SLAM are like walking in the dark at home. You remember where is the desired destination. Your arms can confirm your position, but some toys may be undetected, or a location imprecision may lead you to stub your pinky toe on the furniture. It is better than just using GPS and LiDAR, but there are challenges too.

The use of cameras brings a lot of information, but simple cameras do not read distances. For example, even stereo cameras cannot read distances from distant objects, and RGBD (Red, Green, Blue, and Distance) cameras have lower details that miss faraway obstacles like LiDAR.

The benefit of information is that you can try to interpret the images with a CNN (Convolutional Neural Network), Fukushima and Miyake (1982) and Qi, Charles R., et al. (2018). If you can identify the object in your frame and know its size, you can estimate the distance of the obstacle.

Note that a CNN results in a statistical certainty of the object. The distance is also estimated. Other information derived from this estimative, like the obstacle's relative speed and consequent predictable path, are inferred, and all these uncertainties are Bayesian analysis.

This analysis can use LSTM (Long Short-Term Memory), Schmidhuber, and Hochreiter (1997) to increase the CNN assertiveness. This CNN analysis with LSTM is significant when another object temporarily blocks part of the previously identified object.

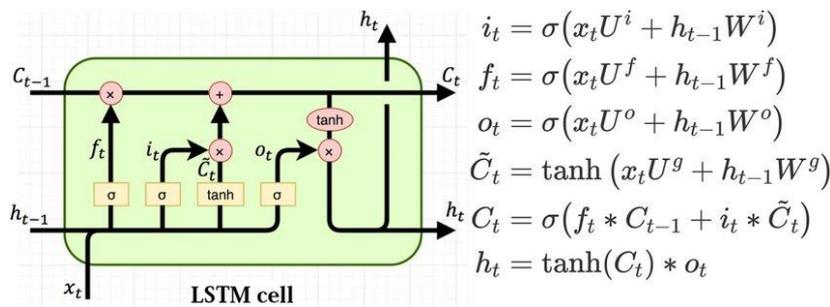


Figure 3. LSTM - Long Short-Term Memory diagram

2.4. Strength and Weakness of sensors

Each kind of sensor reads some value. However, depending on the physics behind the source and the environment, some sensors are better than others, Cheng (2011).

Table 1 - Sensors

	Resolution	Distance	Weakness		Notes
			Environment	Source	
Sonar	Medium	Short		Soft materials	Suitable for traffic jams & parking
Radar	Low	Long		Non-metallic	Suitable for foggy or smoke
Camera	Very good	Medium	White Balance, Foggy	Low light	Good to understand the context
Thermal Cam	Good	Medium	Rainy, Foggy	Reflective	Suitable for low light environment
LiDAR	Medium	Long	Rainy, Foggy	Reflective	Suitable for measuring distance
RGBD Cam	Good	Medium	Rainy, Foggy		Suitable for measuring distance
GPS/WAAS	Good		Indoor places, lane under the bridge		Suitable for position measure
IMU	Medium		Loss of precision over time		Good for accelerations measures
Odometer	Excellent		Loss of precision on slippery surfaces		Good for travel distance measures
Transponder	Excellent	Long	It depends on external sensors emitters		Suitable to avoid traffic jams

Although each situation indicates the most appropriate sensor, a less suitable sensor can assume the function when there is difficulty obtaining the value by the most indicated sensor—for example, the vehicle's position inside a tunnel.

Although the primary purpose of GPS is to obtain the position, it will not be able to do so inside a tunnel. Instead, you can use the odometer or a transponder to get this information, such as the Bluetooth beacons installed in the tunnel and used by some navigation applications.

3. AGGREGATING SENSORS

The goal of using sensors is to identify the drivable zone. In some scenarios, the sensors may not generate readings or readings with very high variability. One of the methodologies is their unification through normalized averages or sums known as sensor fusion. However, this unification technique suffers from not identifying and purging the least reliable sensor, reducing the reliability of the decision model.

To understand the classification of the drivable zone, see Figure 4 - Drivable zone in front of UFRJ Chancellor's Building.

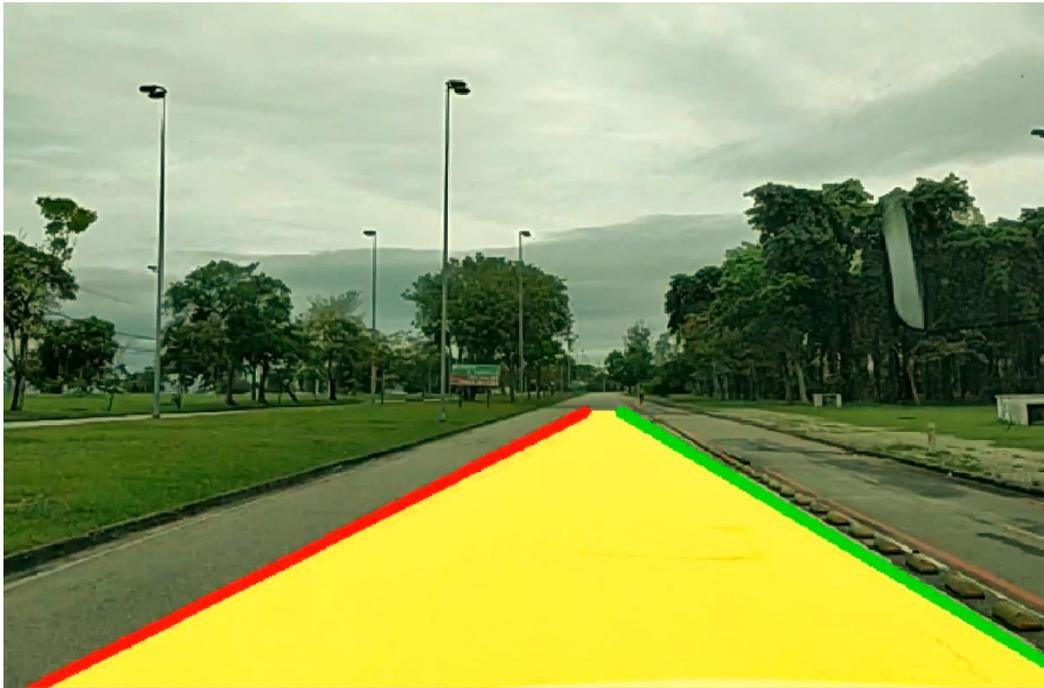


Figure 4 - Drivable zone in front of UFRJ Chancellor's Building

For example, If the system group the radar and camera readings when traveling on a highway on a sunny day. The road signs over the roads tend to be interpreted as obstacles by the radar, resulting in "phantom brakes" Herpel (2008) and Wu (2002).

On the other hand, training a neural network to discard the radar reading in these situations will lead to this signal being discarded or having very little weight on the decision-making model. Dumping the radar reading may seem correct, but an accident will occur if the vehicle encounters a truck crossing the road or a truck jackknifing.

This article proposes cross-interpretation of sensors and the constant identification of the scenario and sensors' limitations due environment. This cross-interpretation can adapt the weights in the decision-making model to avoid unwanted situations.

In the example described above, using the cross-interpretation, the camera should identify the traffic sign on the highway and purge a spurious signal from the radar for having identified the cause of the false positive.

In a scenario where there are limitations to the camera, such as fog or dynamic range limitation, other sensors, like the radar, gain relevance and must be considered.

In a third scenario, on the same high-speed lane but in a traffic jam, sonar, a sensor not usually used in this environment, gains relevance because it has better accuracy at short distances than radar. Again, as noted, interpreting the vehicle's situation and the limitations present to the sensors increases the assertiveness of the system.

3.1. IMPROVING THE RELIABILITY OF THE DECISION-MAKING MODEL

As described, many parameters can interfere with obtaining the desired analysis, so this article suggests the following order of procedures to increase the reliability of the decision-making model.

- 1- Obtain the location with GPS, interpret the environment by the camera and thermal vision, and measure the distances with LiDAR, radar, and sonar of the nearby environment.
- 2- Analyze the data collected to classify the limitation of each sensor in the environment and guesses the place where the vehicle is traveling (highway, tunnel, urban circuit with obstacles coming from different directions).
- 3- Establish the sensors' priority order, considering the sensors' limitations and the type of location. Then, plan the route (trajectory optimization, energy, reinforcement learning) and perform the control on the vehicle.
- 4- Non-priority sensors, classified in step 2, continuously monitor the environment, informing warnings to the model. The primary sensors must classify these warnings as non-relevant when distinguishing the cause or transmitting the alert to the responsible driver.
- 5- In case of warnings transmitted to the driver increase or the classification of step 2 changes, the system establishes a new priority order of the sensors. Then, the system performs a further analysis, or a warning is transmitted to the driver, informing the unsafe condition.

4. CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

Sensors with higher resolution, such as cameras, generate enough points to, with an instant image, generate enough inputs to feed a convolutional neural network that produces a probabilistic interpretation of the understanding of the image. The lower the number of points acquired, the lower the power to interpret what is being measured.

However, even with the high number of information present in a camera, the analysis becomes segmented into several segregated temporal spaces. Even using LSTM short memory techniques, we still have an analysis of raster images.

In future work, the interpretation of the environment with the construction of a three-dimensional vector model will be able to aggregate the information from sensors with lower resolution in a much more effective way. Thus, allowing the allocation of values from other sensors in the spatial domain of the interpretation of the environment.

Building a vector model allows for filling the gaps the sensors present with lower resolution. It also has the potential to reduce spurious readings, increasing the reliability of the decision model.

5. ACKNOWLEDGMENTS

I am thankful to UFRJ for the opportunity to improve my knowledge. The Robotics Laboratory at COPPE/UFRJ provides ideal conditions for research and development, and Professor Dr.-Ing. Max Suell Dutra is a true hero who serves as an example, overcomes difficulties, guides his students, and provides vast knowledge and experiences vital to success.

I also want to thank my family, friends, and colleagues who have helped me significantly, giving advice, being companions, and providing support and motivation in this journey that requires effort and dedication.

As Isaac Newton said in 1675, if we see further, it is by standing on the shoulders of Giants.

6. REFERENCES

- Bosch GMBH, Robert, 2022 "Automotive Handbook 11th edition" *Wiley*. pp. 1813-1815.
- Cheng, Hong, 2011. "Autonomous intelligent vehicles: theory, algorithms, and implementation." *Springer Science & Business Media*
- Fukushima, Kunihiko, and Sei Miyake, 1982. "Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition." *Competition and cooperation in neural nets*. Springer, Berlin, Heidelberg. pp. 267-285.
- Gao, H., Cheng, B., Wang, J., Li, K., Zhao, J., and Li, D., 2018. "Object Classification Using CNN-Based Fusion of Vision and LIDAR in Autonomous Vehicle Environment," in *IEEE Transactions on Industrial Informatics*, vol. 14, no. 9, pp. 4224-4231, Sept. 2018, DOI: 10.1109/TII.2018.2822828.
- Herpel, Thomas, Lauer, Christoph, 2008. "Multi-sensor data fusion in automotive applications.", *3rd International Conference on Sensing Technology*. IEEE, pp. 206-211.
- ORAD (On-Road Automated Driving committee), 2021. "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles.", https://www.sae.org/standards/content/j3016_202104/
- Schmidhuber, Jürgen, & Hochreiter, Sepp, 1997. "Long short-term memory." *Neural Computation*, 9(8), pp.1735-1780.
- Qi, Charles R., et al., 2018, "Frustum pointnets for 3d object detection from rgb-d data." *Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Wu, Huadong, Siegel, Mel, Stiefelwagen, Rainer and Yang, Jie, 2002, May. "Sensor fusion using Dempster-Shafer theory." *Proceedings of the 19th IEEE Instrumentation and Measurement Technology Conference (IEEE Cat. No. 00CH37276)* (Vol. 1, pp. 7-12). IEEE.