Master's thesis description
2018-2019

Artificial Intelligence: Sensor-fusion with On-board Cameras for Autonomous Navigation

Dieter Balemans

Supervisors: Simon Venneste, Jens de Hoog, Sigfried Mercelis, Peter Hellinckx

02/01/2019

Universiteit Antwerpen

Abstract

Autonomous vehicles use a large number of sensors. Commonly used sensors in autonomous applications are LiDAR and camera sensors. In order to handle all the data provided by these sensors in a optimal way multisensor data fusion can be used. In the field of data fusion there is a standard model proposed by the Joint Directors of Laboratories. This model consists of four processing levels. These levels are: object refinement, situation refinement, threat refinement and process refinement. Most state of the art research is situated in the object refinement level. In this master thesis a system will be proposed that will situated on both the object refinement and threat refinement level. The system will fuse LiDAR and camera data. The purpose of this system is to achieve higher confidence in sensor measurements and a quicker response to possible threats.

Contents

1 Master's thesis description

1.1 Problem Description

Modern autonomous vehicles use a large number of sensors. Besides LiDAR and radar, many also use on-board cameras to sense the environment. Every sensor has its own use case. However, sensors often provide information about the same features. For example, with data from a LiDAR sensor we can determine a distance between an object and the front of the vehicle. Alternatively, also radar data can be used for this. Knowing this, we can combine multiple sensors to compare the same features and improve the confidence in these features. Every sensor has its own advantages and disadvantages. The purpose of sensor fusion is to combine sensors in such a way the advantages of each sensor are combined and the disadvantages are suppressed. This concept is not new as living organism have the capability of using multiple senses to learn about the environment. The brain fuses all this information in order to perform a decision task. This concept is called multisensor data fusion. [1]

As Hall [2] describes in is paper the Joint Directors of Laboratories have standardized terminology related to data fusion. They proposed a functionally oriented model which consists of different levels. These levels are: object refinement (level one), situation refinement (level two), threat refinement (level three), and process refinement (level four).

If you take a look at most state of the art applications in the field of LiDAR and camera fusion, you can conclude most applications are situated on the first level. For example a United States patent exists [3] that protects a method for fusing radar/camera object data and LiDAR scan points. In this method object files are created in which objects are represented with their properties like position, orientation, and velocity.

In research from the University of Toronto and Uber Advanced Technologies Group [4] a method was developed to detect 3D objects using LiDAR and camera data to perform very accurate localization. In other research from Zhang et al. [5] a vehicle detection system was developed that consisted of two phases: a hypothesis generation and verification phase. This is a commonly used method for fusing LiDAR and camera sensors.

1.2 Research question / Thesis statement

This master thesis handles multisensor data fusion using LiDAR and camera sensors. The main research question of this thesis is: How can we apply sensor fusion algorithms for object refinement and threat refinement in autonomous applications?

1.3 Approach / Methods

The plan is to create a system that takes the LiDAR and camera data as input and outputs refined objects and a threat level. The refined object will consist of the objects location, velocity, heading, and classification. The threat level will be signal that will allow the vehicle to react quickly in emergency situations. The objects output can be classified as a level one output, the threat level as a level three output. It is worth noting the process described here is a provisional

process. The details of the process can change in the course of the master thesis. The process itself will consist of three phases: a preprocessing phase, a fusion phase, and a post-processing phase. An overview of the process is visible in figure 1. In the preprocessing phase both the LiDAR data and camera data will be used to detect objects. In the case of LiDAR data these objects needs to be extracted from a point cloud. These objects will form clusters in the this cloud. This has already been achieved in research from Douillard et al. [6] and research from Johansson and Wellenstam [7]. In case of the camera data NVIDIA tools [8] and existing deep learning algorithms will be used. In the fusion step the detected objects from both inputs need to matched with each other. Therefore, a transformation step will be needed to transform both data sets to a consistent coordinate system. After this step the objects from both data sets are fused.
Lastly in the postprocessing step, the fused objects can be used for properties extraction. The most important properties of the objects that can be extracted are position, velocity, heading and the classification. For extracting the velocity and heading multiple frames will be needed. This means the system needs to be able to track objects over different frames. A frame is the current input at a specific moment in time. Thus the system will track objects in time. An important criteria for this to work is that both data streams need to be synchronized. The level one output consists of the objects identity (classification) along with all its properties. The detected objects will be updated as long as they are detected. Based on the objects properties it will be possible to predict next states of objects, and infer a threat level.
The evaluation of this process will consist of a hybrid environment, both simulation and a real world data will be used.

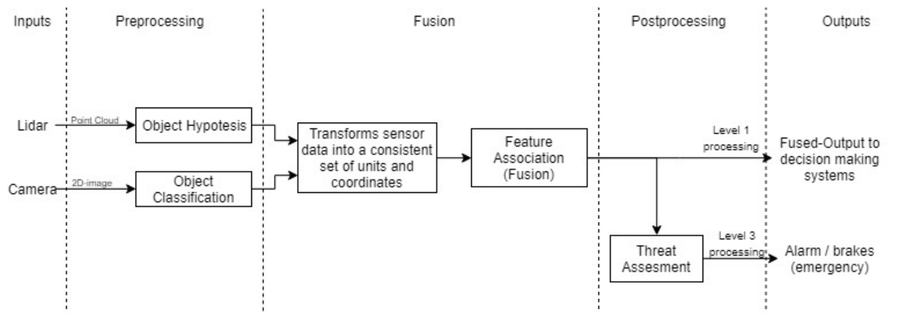

Figure 1: Schematic of fusion system with outputs on different levels of the JDL model

1.4 Preliminary results and discussion

At the moment there are no results to report.

2 Work plan including time table

Begin	Deadline	Summary
01/10/2018	06/01/2019	Research of the state of the art and writing a survey paper. Also this document is written in this period.
01/02/2019	24/02/2019	Further preliminary research. In this period all details of the system need to be determined.
24/02/2019	17/03/2019	Implementation of the preprocessing step. This includes the clustering of the LiDAR points and object classification using the camera data.
17/03/2019	31/03/2019	Implementation of the fusion step. This is the step where both data sets need to be transformed and the detected objects need to be associated with each other.
31/03/2019	30/04/2019	Implementation of the postprocessing step. In this step the object refinement and threat detection take place.
01/05/2019	01/06/2019	Finish implementation and start evaluation. At the end, process the results in the final paper and prepare the presentation.

3 Implications of research

The main purpose is to develop a fusion algorithm between LiDAR and camera sensors. The state of the art applications in the field of LiDAR and camera fusion are currently mostly only working on the first level of the JDL model (object refinement) [2]. In this research we extend this to the higher levels. Especially the threat refinement level will be researched. The purpose will be assessing a threat level using multiple input sensors. Furthermore will all the developed methods be tested in a hybrid environment between simulation and the real world.

4 Contacts

4.1 Internal supervisor(s)

Simon Venneste, +3232659009, simon.vanneste@uantwerpen.be
Jens de Hoog, +3232658942, jens.dehoog@uantwerpen.be
Sigfried Mercelis, +3232651698, siegfried.mercelis@uantwerpen.be
Peter Hellinckx, +3232651686, peter.hellinckx@uantwerpen.be

4.2 Externals supervisor(s)

None

4.3 Students

Dieter Balemans, +32479534808, dieter.balemans@student.uantwerpen.be

5 List of references

References

[1] D. P. Mandic, D. Obradovic, A. Kuh, T. Adali, U. Trutschell, M. Golz, P. De Wilde, J. Barria, A. Constantinides, and J. Chambers, "Data Fusion for Modern Engineering Applications: An Overview," in *Artificial Neural Networks: Formal Models and Their Applications – ICANN 2005* (W. Duch, J. Kacprzyk, E. Oja, and S. Zadrożny, eds.), (Berlin, Heidelberg), pp. 715–721, Springer Berlin Heidelberg, 2005.

[2] D. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 6–23, 1997.

[3] Z. Shuqing, "System and method for fusing radar/camera object data and LiDAR scan points," feb 2016.

[4] M. Liang, B. Yang, S. Wang, and R. Urtasun, "Deep Continuous Fusion for Multi-sensor 3D Object Detection," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11220 LNCS, pp. 663–678, 2018.

[5] F. Zhang, D. Clarke, and A. Knoll, "Vehicle detection based on LiDAR and camera fusion," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1620–1625, IEEE, oct 2014.

[6] B. Douillard, J. Underwood, N. Kuntz, V. Vlaskine, A. Quadros, P. Morton, and A. Frenkel, "On the segmentation of 3D LIDAR point clouds," in *2011 IEEE International Conference on Robotics and Automation*, pp. 2798–2805, IEEE, may 2011.

[7] T. N. Johansson and O. Wellenstam, "LiDAR Clustering and Shape Extraction for Automotive Applications," 2017.

[8] NVIDIA, "Nvidia drive developer." https://developer.nvidia.com/driveworks.

[9] F. E. White, "JDL, Data Fusion Lexicon," *Technical Panel For C3*, vol. 15, no. 0704, p. 15, 1991.