Spatio-temporal modeling of vehicle surroundings

Thomas Schön (Uppsala University), Jacob Roll (Autoliv) and Dave Zachariah (Uppsala University)

February 3, 2017

1 Background

Deep learning [LeCun et al., 2015] is a quickly evolving research area. Based on numerous examples (training data), a deep neural network can, in theory, learn virtually any function, with applications such as describing the content of diverse images, playing games (chess, Go, Poker), understanding speech, or finding user preferences in social networks. An application that has received much attention in recent years is vehicle active safety and autonomous driving. With the help of vision and other sensors, the autonomous driving system needs to get a representation of its surroundings, suitable to determine how to control the vehicle. This is the topic of this PhD project.

Autoliv is the world’s largest automotive safety supplier with sales to all the leading car manufacturers in the world. We develop, manufacture and market protective systems such as airbags, seatbelts, steering wheels, pedestrian protection systems, child seats, passive safety electronics and active safety systems. Our leading market position in automotive safety includes a global market share of approximately 39% in passive safety and more than 20-25% in active safety.

At Autoliv Vision Tech Center in Linköping, we focus on vision-based systems for active safety and autonomous driving, based on cutting edge image processing technology. We have 400 employees and consultants working on tomorrow’s technology.

The statistical Machine Learning group at Uppsala University work on both basic and applied research in order to automate the extraction of knowledge and understanding from data. We mainly work with probabilistic models for dynamical systems and their surroundings, focusing on developing new models and how these models can be learnt from data.

The statistical Machine Learning group at Uppsala University and Autoliv have a longstanding collaboration and this PhD project is a natural progression that is fully in line with existing collaborations. Some earlier joint results in the same direction as this project are illustrated in the left plot in Figure 1.

Figure 1: Left: Illustration of a spatial model (here a so-called conditional random field) of the road surface in front of the vehicle that is learnt on-line based on information from a stereo camera and other on-board sensors. The vehicle surroundings have also been classified into various categories as indicated using different colours. Right: Example of Autoliv’s existing semantic segmentation using deep learning, illustrated on freely available data from Cordts (2016). Each pixel in the image is assigned a specific object class (car, road, person, traffic sign, pole, vegetation, etc.)
**2 Project description**

The goal of the PhD project is to develop new representations (models) capable of explicitly summarizing and explaining the vehicle surrounding based on data from cameras and other internal sensors on-board the vehicle. The problem of automatically learning these representations lies at the heart of this problem. One of the key lessons from modern Machine Learning is that flexible models often give the best performance. We will in this work make use of and further develop two state-of-the-art models and techniques, namely spatio-temporal models and deep learning.

Spatio-temporal models (Cressie and Wikle, 2011; Wikle, 2015) offer a natural basis since they are capable of representing not only changes over time, but also the spatial structure of a dynamical system. The on-board sensors (including for example cameras, inertial sensors and wheel speed sensors) provide information that will be used in learning spatio-temporal models, describing not only the motion of the host vehicle, but also the immediate surroundings of the vehicle.

Deep learning has revolutionized computer vision and machine learning and it is extensively used in autonomous driving applications. Still there are important developments needed, both at the basic research level and from a more applied perspective. We will in this project let the application drive the basic research. More specifically, new spatio-temporal models will be developed and in particular their interface with the video data from the cameras. This data will be handled using deep neural networks, capable of turning the camera images into representations suitable for important tasks such as object tracking (e.g. other cars, pedestrians, bicyclists, animals, etc.) and free space detection (recall Figure 1). Recent semantic segmentation results by Pinheiro et al. (2015, 2016) offer fertile ground for interesting developments. The combination of deep learning and spatio-temporal models should provide an improved scene understanding that can be used to detect approaching obstacles, warn the driver if a collision is imminent, or automatically take control of the vehicle if necessary.

The outcomes of this project will have solid impact on the autonomous vehicles currently being developed. There is also a wide range of other problem areas that have a lot to gain from the automated learning of new and more capable spatio-temporal models.

**3 The team**

The supervisors for this project are:

- Thomas Schön, Dept. of Information Technology, Uppsala University, thomas.schon@it.uu.se
- Jacob Roll, Autoliv Electronics AB, jacob.roll@autoliv.com
- Dave Zachariah, Dept. of Information Technology, Uppsala University, dave.zachariah@it.uu.se

Geographically the work will be carried out both within the concept development group at Autoliv in Linköping and within the statistical Machine Learning group at Uppsala University.

**References**


