

Dynamic Vision for Driver Assistance Systems

“Cars with Eyes”

Tobi Vaudrey

Imagine a futuristic world where cars drive themselves. In New Zealand alone, there have been about driving 5000 fatalities over the past 10 years. Think of the reduction in accidents as cars are always looking at the road and making decisions based on its perception of the environment. The car will never get drunk, never fall asleep at the wheel, be always attentive and never suffer from road rage! This seems like a far-distant possibility that is only in the minds of science fiction writers, such as those of *Star Trek*. But this reality is perhaps not so far away.

My research is on vision based driver assistance systems, in other words, I try and create eyes for cars. We are still a while off the future depicted in the introduction, but there are smart cars that do take control of vehicles already, even if the driver is not aware. One of the first safety systems was the seat-belt. This simple device was able to reduce road-tolls dramatically. Since this revolutionary development, further systems have been introduced to prevent drivers from harming themselves. Many of the engineered passive safety systems, such as crumple zones, aim to reduce damage to the driver in a crash. Later pseudo-active systems, ASB (anti-lock breaking system) and ESP

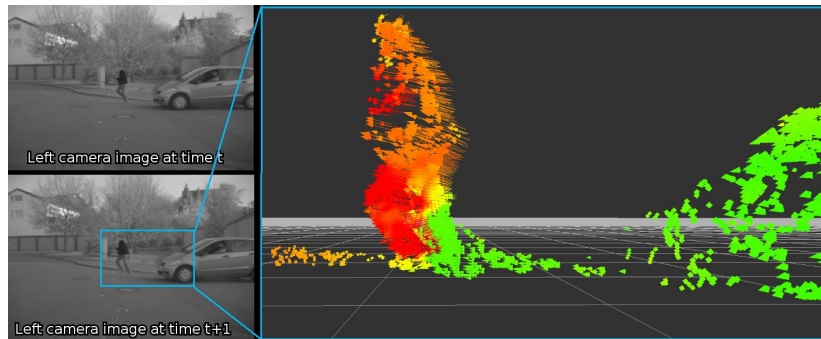


Figure 1: A person running from behind a vehicle, can you tell what is happening from these images alone? (Left: images taken from camera at two sequential instances. Right: Motion and distance analysis using “scene flow”, red = moving fast, green = not moving.)

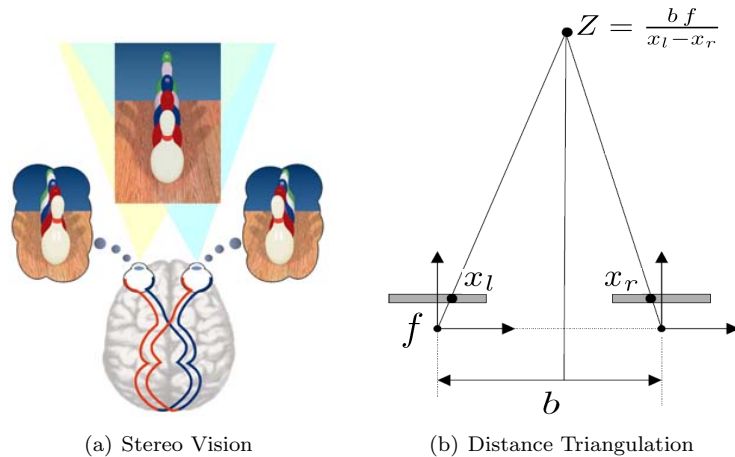


Figure 2: How to see using stereo vision. (a) shows the what the different images look like between left and right “eyes”. (b) shows how triangulation works to find distance.

(electronic stability program), aim to prevent the crash happening by stopping a car sliding off the road and maintaining control in critical situations. The latest reports from engineers at Daimler and other big car companies have said We have gone as far as we can with rubber-to-road technology. Therefore, any new systems need to be truly active; predict and prevent crashes.

But do we prevent these crashes? We need to make the vehicles aware by giving them eyes! My specific research topic is trying to achieve dynamic vision for cars, and make the information reliable for situation analysis; I try and analyse the current situation in front of a moving vehicle, and try to predict the situation in the future. The algorithms I am developing aim to analyse the 3D motion and position of objects in a scene (*scene flow*), as this is the most basic information needed for decision making. Furthermore, this information needs to be very reliable if we are to put our lives in the hands of a computer!

Take a look at Figure 1. Can you tell what is happening from 1 image alone? Can you tell the approximate distance to the person? In this case, the answer to both these questions, for a human, is yes. You can tell that the person is running (from their stance). You can also probably tell that they are about 20 metres away (from their relative size). These perceptions are model based and use our huge brain to access our memory to analyse a possible solution. Unfortunately, computers are nowhere near as smart as humans, so analysing a situation using model is very complex, and currently near impossible.

I rely on using a more simple approach, but still inspired from biology, called *stereo-vision*. The only information provided to the computer is from two cameras, mounted looking out the windscreen of a vehicle. Two images, from

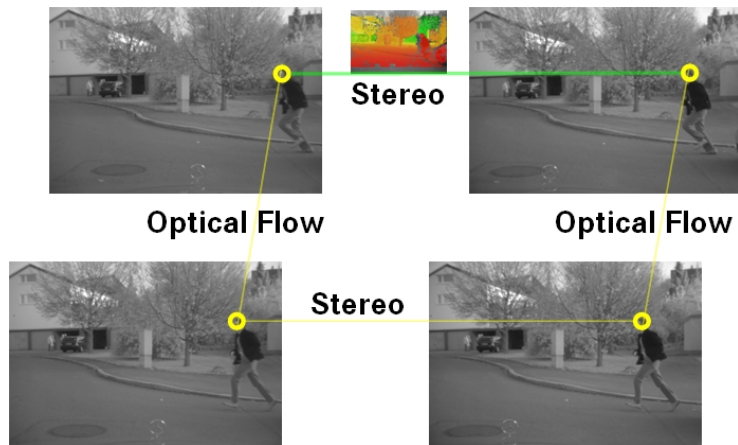


Figure 3: Scene flow schema. Using stereo images from two sequential instances, you are able to calculate the scene-flow, see Figure 1

slightly different angles, are taken at exactly the same time (see Figure 2(a)). These images are then processed, measuring the differences between them. Finally, the differences between the left and right image can be used to triangulate the distance to each point, as long as the epipolar geometry (relative position between the cameras) is known (see Figure 2(b)). This is not a new concept, but the community is still looking for ways to improve results for stereo-vision. This provides an answer to question 2 above (how far?), however from this we can not tell anything about the dynamics of the environment (question 2 above).

To get motion information from images, you need an image sequence. By analysing sequential images, you can obtain motion information using *optical flow*. This give you the 2D motion information of a scene. It is obvious from here that 2D motion will not be good enough in an environment such as driving. So why not fuse the two information? Use a sequence of stereo-images to provide 3D information, and then use optical flow to generate 3D scene flow. This is where my research lies, I have already proved the viability of such an approach. Now I aim at making this more and more robust to common problems in driver assistance.

I hope that my research will one day be in every vehicle, enabling a safer environment on the road, reducing accidents, and for those who want it, taking the hassle out of the commute to and from work or getting you home from the pub safely.