Integrated Stereovision for an Autonomous Ground Vehicle

2005 SURF Report

Erik Johannesson, Lund Institute of Technology
Faculty Mentor: Professor Richard Murray
Co-Mentor: Kristo Kriechbaum

Abstract

This report describes development of the stereovision sensing system on an autonomous vehicle competing in the DARPA Grand Challenge, a desert race for robotic vehicles. A basic stereovision system was in place when the project started, but it had not been tested on the new vehicle. It turned out a lot of problems needed to be solved before stereovision was ready to be integrated into the full system. New algorithms were designed and implemented to improve camera focus and stereovision processing, and to remove noise in the data. A lot of testing was performed to ensure data quality before integration into the sensor suite of the vehicle. Results show that, even though the vehicle is capable of obstacle avoidance and navigation using only stereovision sensing, the quality of stereovision data needs to be improved

Introduction

To stimulate research and innovation in the field of autonomous ground vehicles, the DARPA (Defense Advanced Research Projects Agency) announced in March 2003 that it would arrange a competition known as the DARPA Grand Challenge (DGC). The competition would be an offroad race for autonomous vehicles held in the desert. A first race was held a year later, which none of the participants were able to complete. A second race was held on the 8th of October 2005.[1] A team consisting of mainly Caltech undergraduates, named Team Caltech, competed in both races. The work described in this report was performed during the summer of 2005 and the remaining time after the summer before the race.

Team Caltech's vehicle for the 2005 DGC, called Alice, is a 2005 Ford E-350 van that has been modified by Team Caltech sponsor Sportsmobile, Inc. for offroad usage. Modifications include custom suspension and transmission, plus electrical power and air handling systems.[2] In order for Alice to drive autonomously, she has been equipped with actuators, a computer cluster and a range of sensors, including four cameras used for stereovision processing.

Problem Description and Approach

Alice uses internal maps of the surrounding terrain to plan her driving. The objective of stereovision is to help create these maps using images from cameras in stereo pairs. Given knowledge on how the three-dimensional world is transformed into the images, the images are used to reconstruct geometrical information about the world. Finding the parameters that describe the transformation is known as *camera calibration*. Alice's stereovision system uses a software called Small Vision System (SVS) from Videre Design Inc. that performs major parts of the calibration and the stereo processing.

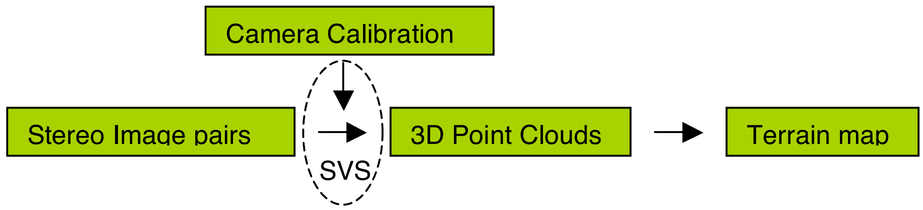

Figure 1: Simple flowchart of stereovision processing in Alice

In the beginning of the summer of 2005, the stereovision system that was used in 2004 had not yet been tested on Alice. In order for stereovision to be used in the 2005 race, there was need for new hardware, better calibration, good parameter settings, system integration, and lots of testing.

The approach carried out was to try to break down all the problems into separate issues and to always work on the issue that seemed to be the biggest obstacle towards a well-functioning integrated system. Due to the pressing deadline (the race), simple solutions were often chosen rather than more elegant, but time-consuming, ones. In general, solutions were first tested in simulation or in simple terrain (i.e. a parking lot) before attempting the desert environment.

System Design and Analysis

Camera Essentials

The old camera lenses were not sturdy and thus inappropriate for rough driving. It was therefore decided to purchase new lenses. The focal length of the new lenses was selected to be 2.8 mm and 8 mm for the short and long range camera pairs respectively. In combination with baselines (the distance between a camera pair) of 0.6 m and 1.5 m, this approximately gives range resolutions of 0.2 m at 12 m and 37 m range for short and long range respectively.

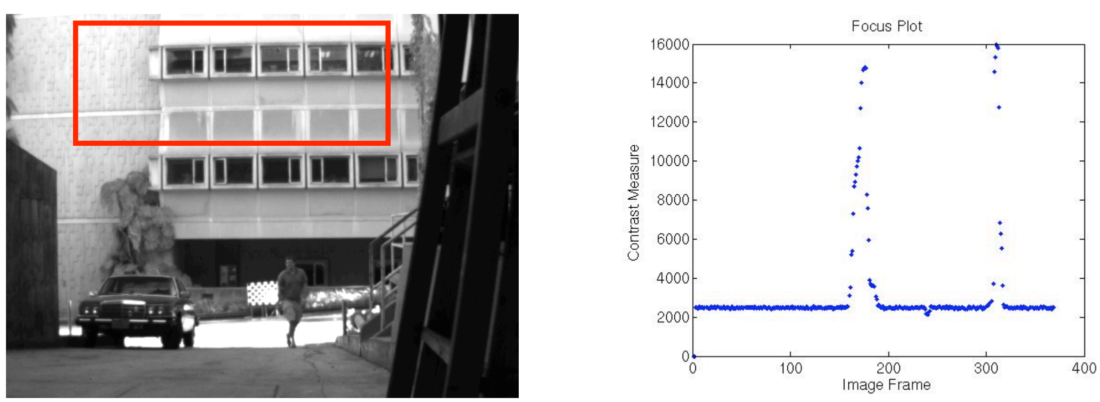

Figure 2: The plot shows a measure of the contrast in the red box as the focus was changed.

To achieve optimal focus at the desired ranges visual inspection of the images was not considered to be sufficient. A semi-automatic focus algorithm was implemented in MATLAB. The algorithm continuously computes a simple contrast measure (the sum of the absolute values of the 100 largest intensity differences) on a selected part of the image, and plots the measure in real-time as the focus knob on the lens is manually turned. With feedback from the plots, it was easy to find optimal focus settings.

There were severe problems with camera calibration. Although SVS reported decent calibration errors, it seemed impossible to get good data. Eventually, a software bug that caused the processing to ignore the calibration information was discovered. Once found, it was easy to fix, but it severely slowed down the progress.

Figure 3: New and old calibration targets

When calibrating the cameras, a checkerboard calibration target is used to get images of a known geometry. As this checkerboard, glued on a foam board, was brought out to field tests in the desert, it eventually became slightly bent. Since the calibration software assumes a plane this rendered the target unusable and a new target had to be made*.

Processing improvement

It was discovered that there was a significant amount of noise in the stereovision data. To remove the noise, a filter was implemented on one of the intermediate results in the process: the disparity image. The disparity image corresponds to the left image in the stereo pair, but the intensity (the *disparity*) is inversely proportional to the estimated range of the pixel. The noise was clearly visible as scattered pixels of different ranges, not corresponding to real objects in the images. Since one can assume that objects in the world are largely continuous, while noise is more random, it was decided to filter out small regions of connected pixels, in the sense that neighbouring pixels are seen as connected if the disparity difference is small. The filter was named the "blob filter" and was implemented using a flood-filling algorithm that uniquely colors connected regions (blobs) and then removes all pixels with a color that has less than a certain number of occurences (hence removing small regions). Both the threshold for region size and the tolerance level for connectedness was implemented as adjustable parameters. The idea for the algorithm was based on a blob coloring algorithm by Ballard & Brown[3].

* Material for the new target was acquired by Tami Reyda.

Figure 4: The top image, a typical desert road, is taken by the left long range camera. The middle image is the disparity image before the blob filter is applied – notice the spurious objects in the upper part of the image. The bottom image is the disparity image after the blob filter is applied (minimum region size = 350, connectedness tolerance = 8).

With the blob filter in place it was possible to get a higher signal-to-noise ratio than before. Then, to find the best balance between amount of data and risk of noise (resulting in spurious obstacles in the map) the stereo processing parameters had to be tuned. These parameters control the way SVS computes the point clouds, e.g. by changing thresholds that are used to remove noise. The tuning was done by collection of log data (large sets of image pairs) from the desert that was replayed several times in simulation with different parameter settings. The parameters were chosen by inspection of accuracy in both disparity images and in the resulting maps.

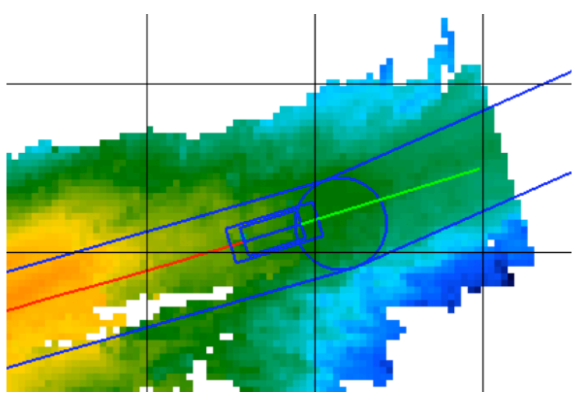

Figure 5: Example of elevation map from short range stereo (blue is higher)

Processing speed is important for the quality of the maps. If the frame rate is too low, then there may be gaps in the map and Alice's reaction time becomes longer. One way to increase the frame rate is to only process a part of the images. This technique, called *subwindowing*, does not necessarily lead to loss of information, since e.g. the sky often takes up a large portion of each image. The code to do this was originally written last year[4], but it had to be rewritten due to an upgrade of SVS. To increase the processing speed, the code for the coordinate transformation of the point cloud (from the camera frame to UTM coordinates) was also rewritten[*].

System Integration

Good quality of maps was required before attempting autonomous operation of Alice using stereovision. Before integration it was therefore verified that accurate maps could be acquired in simulation on log data and with the stereo cameras put on a table in front of a set of obstacles. A lot of time was also spent tuning parameters and looking at the data produced by stereovision while driving Alice in the desert, with stereovision data not being fed into the planning software.

Once the ability to create accurate maps was demonstrated, it was attempted to use stereovision data along with other sensor data. Although there were not too many errors such as spurious obstacles, a problem seemed to be that stereovision often pulled Alice to a certain side of the road. The reason was that no data was received in areas of low texture (because of the way stereo processing works) and it was common for the desert roads to have one side with more texture than the other. Since the planning software prefer to drive trough areas with more sensory data (all other things alike), this unnecessary bias was introduced. However, using only stereovision for terrain sensing, Alice was able to quite successfully navigate through obstacles in complex situations – on the first attempt!

Though, after a while of testing, it became apparent that stereovision in some situations could produce spurious holes, up to 3 meters deep, in the map. These holes were usually located in the middle of the road. By looking at log data it was concluded that these holes was the result of a mismatch of parallel wheel tracks on the ground (in the stereo processing, a correlation filter is used to find similar-looking points in pairs of images). The same phenomenon was later seen for fenceposts and series of the same logotype on the walls in the National Qualification Event (a part of the competition which took place just before the race).

Unfortunately there was not enough time to implement a nice solution to this problem, since it likely would have required a replacement of SVS as the stereo processing software. The mismatching problem in combination with the tendency for Alice to drift to one side of the road instead led to a change in how stereovision was used. Instead of trying to use it to make dense terrain maps it was used for detection of positive obstacles. This proved to be a good complement to the LADAR (Laser Range Finder) sensors that were also used. The LADARs are better at providing dense terrain maps, but stereovision has the advantage that it always sees the full height of obstacles. By using stereovision in this way, the advantage was kept even though the problems with erroneous or incomplete data no longer had an impact.

The change to only use stereovision for positive obstacle detection was simply implemented by a filter that removed all the points in the point cloud that were outside some given volume. A lower bound was set so that no points below 0.3 m and 0.4 m above the ground was allowed for short and long range respectively. This filter also allowed for removal of points above a certain height, which prevents interpretation of tunnels as high obstacles. An additional advantage of the filter

---

[*] Part of this work was done by Jeremy Leibs

was the possibility to put constraints on the range of allowed points. This was useful because the accuracy tended to be low far away.

Conclusions and Future Work

Stereovision is a hard problem to solve in practice. There are many practical issues to deal with to get good data when driving with high speed in desert terrain. The accuracy and speed of Alice's stereovision sensors need to improve significantly. Nevertheless, it has been shown that stereovision can be sufficient for navigation around several obstacles.

A possible direction for improving the system is to find an alternative to SVS. A nice project would be to implement a new stereo processing engine. The processing could then be customized for the application and possibly remove the need for some of the ad hoc fixes such as only using stereovision for positive obstacle detection. If the processing was also done using hardware acceleration the speed could be dramatically improved. Then, the cameras (now 640x480 resolution) could be replaced by ones with higher resolution, rendering an increase in accuracy.

Another issue that needs to be fixed is to construct a more robust calibration target, so that it does not break when brought to the desert. This could be achieved by some kind of reinforcement on the back or sides.

Acknowledgements

I would like to thank all the members of Team Caltech for working so hard and always being available to help with any questions. A special thanks to Jeremy Gillula who was previously responsible for stereovision and provided a lot of invaluable help. Thanks to Maria Koeper for handling purchasing of new equipment with great efficiency and care.

I am of course very grateful to Professor Richard Murray for making it possible for me to come to Caltech and work in this project. Also thanks to Tony and Sandie Fender for giving me great food while doing so.

References

[1] DARPA Grand Challenge site (2005). Retrieved 10/26/2005 from http://www.grandchallenge.org/

[2] Murray, R. et al (2005) *DARPA Technical Paper: Team Caltech*. Retrieved 10/25/2005 from http://www.darpa.mil/grandchallenge/TechPapers.html

[3] Ballard, D. & Brown, C. (1982) *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall. Retrieved 8/19/2005 from http://homepages.inf.ed.ac.uk/rbf/BOOKS/BANDB/bandb.htm, page 151.

[4] Gillula, J. (2004) Development of Enhanced Algorithms for Vision based Obstacle Avoidance for an Unmanned Ground Vehicle. SURF Final Report.