

Yihao Qian

Team A: Aware

Teammates: Amit Agarwal Harry Golash

Menghan Zhang Zihao (Theo) Zhang

ILR08

March.03, 2016

Individual Progress

During the last week, I was in charge of improving the stereo vision algorithm and researching on object tracking system.

Stereo Vision:

I changed our stereo vision system from SGBM (Semi Global Block Matching) to Efficient Large Scale Stereo Matching (ELAS). Compared to the SGBM, ELAS is a dense method used to compute the depth. ELAS can also provide smoother depth estimation. You may find the detailed description of how the algorithm works in the following paragraphs.

Note: Calibration and un-distortion parts are same as last ILR

Calibration:

The purpose of calibration is to compute the extrinsic and intrinsic parameters for both of the cameras. The calibration process also computes the relative position and rotation of two cameras. The most important thing is to compute the radial distortion and tangential distortion of the cameras.

Un-distortion:

Using the distortion parameters that we get from the calibration process, we could un-distort the images captured from two cameras. The reason why we need to un-distort the images is that the stereo vision system assumes the cameras are pin-hole structure. However, for modern cameras, usually the cameras are not pin-hole structure. That's the reason why we need to un-distort the image, so that the un-distorted image looks the same as the image captured by a pin-hole structure camera. Otherwise, it may lead to severe error.

Matching:

The program first selects support matching point from the image. The support matching point means the matching points that has strong confidence and less ambiguity. The standard for selecting support point is as follow:

1. The confidence of matching should be bigger than 0.95
2. The ratio for the second highest confidence point / highest confidence point should be less than 0.5
3. The depth difference between different support points should not be so large

Using the above standard, the algorithm could select the support matching pairs from the images. Those matching pairs' depth information could be used as a prior information for computing the depth for other points.

Then the algorithm uses Delaunay triangulation, interpolation and feature matching to compute the depth for each point.

You may find the ELAS's depth image in figure 1. You may find SGBM's depth image in figure 2.

Object Tracking:

The requirement for object tracking are as follow:

1. Tracking multiple objects
2. Work in real time
3. Track pedestrians and vehicles

I searched online for different kinds of object tracking algorithms. Most of the open source object tracking systems are based on Matlab. The only one C++ algorithm I found is called GOTURN. This is an algorithm that could work in real time. However, this algorithm could only track one object at one frame. We are planning to modify the algorithm in the next few weeks. You may find how the algorithm works in the following paragraphs. You may find the structure of GOTURN in figure 3.

1. The convolutional networks of GOTURN are trained offline, using millions of video on the Internet. Those convolutional networks are used to learn the deep feature for the object that the algorithm is going to track. This enables the network to track on generic objects
2. After training the network, the network sends the bounding box of previous frame and potential bounding box of current frame to the network. The network could learn to identify the object in the current frame and adjust the corresponding bounding box in the current frame.

This algorithm ranks high in the object tracking bench mark. Those are the reason why we choose GOTURN for our object tracking system.

Challenges:

Object tracking:

At this stage, our object tracking system is behind the schedule. We need to build the object tracking system as soon as possible. After that, we could start to build Compiling package:

I faced some challenge on compiling the package. The error says the path for caffe directory is not found. After seeking help from Alex, I found out I need to set the path manually and link the corresponding library.

Teamwork:

Zihao worked on GPS subsystem. He first found the Arduino code for GPS online and transplanted the code to Ubuntu system. Harry and Amit worked on getting data from radar. They could achieve the raw data from radar. The next thing they are going to do is to decode the information. Menghan built visualization tool. She used OpenCV to show the position for pedestrian and vehicle.

Plans:

In the next week. We will build object tracking system. The first stage is to run the demo on our computer. The next thing is to combine object detection system, stereo vision system with GUI.

Figures:

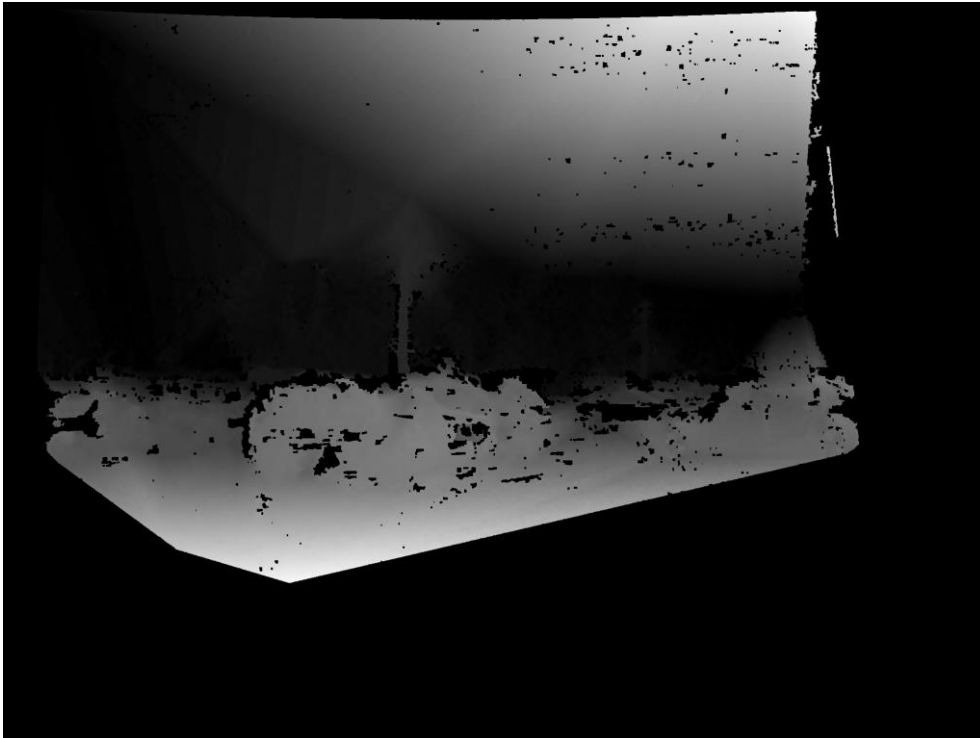


Figure 1. Depth Map for Parking Lot (ELAS)

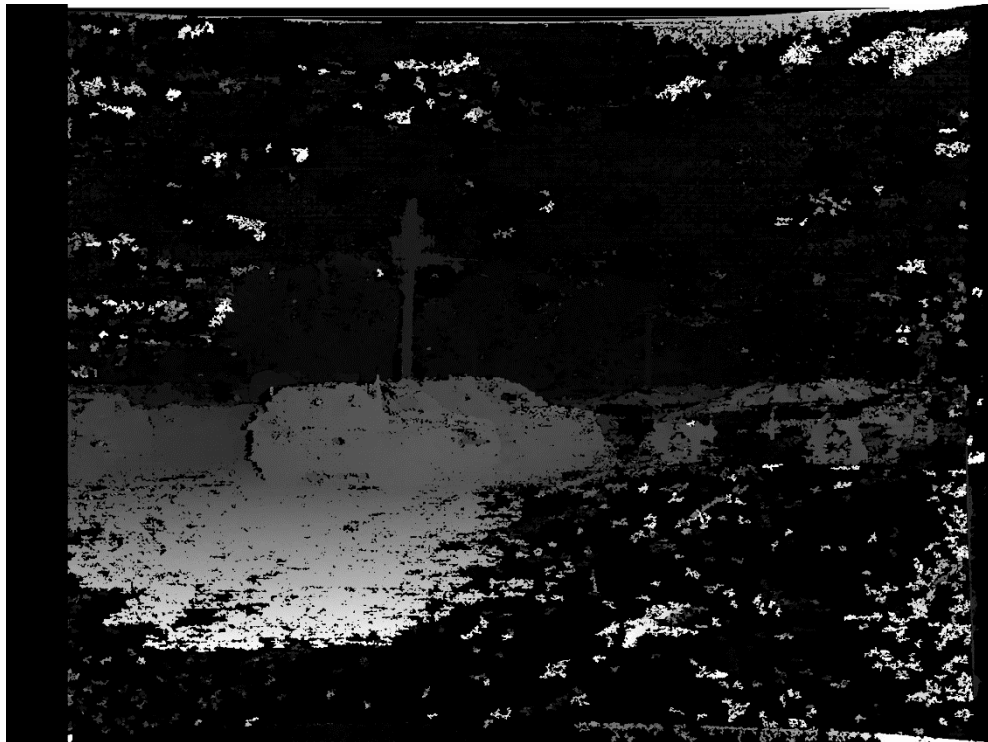


Figure 2. Depth Map for Parking Lot (SGBM)

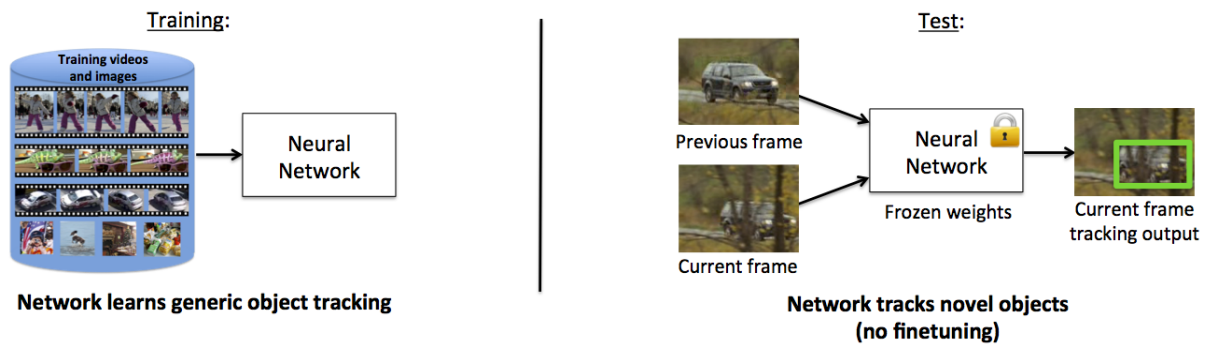


Figure 3. Architecture for GOTURN

Reference: Generic Object Tracking Using Regression Networks, David Held, Sebastian Thrun, Silvio Savarese