Stereo-Matching Using Semi-Global Matching and Census Transform Cost

# Theoretical Background

The semi-global matching technique is a ubiquitous method for stereo analysis. It has been proven multiple times that it can produce in real-time acceptable results. The algorithm was first proposed by Heiko Hirschmuller. The original proposed algorithm suggests using the Mutual Information as matching cost, but others may be used as well. In my implementation I have used Census Transform to create the cost matrix. Hirschmuller suggested the calculation of costs locally first, creating an initial cost matrix which specifies for each pixel the cost of matching at a given disparity:

and then to aggregate them in 8/16 paths. The matching cost for a pixel p at disparity d over a path in direction r in the image is defined recursively:

The final matching cost for each pixel is the sum of matching costs along the chosen directions:

The above formula adds a penalty P1 to disparities which only differ by one along the path and a disparity P2 for disparities which differ by more than 1. P1 should be smaller than P2. This way, inconsistencies along the path are penalized.

The disparity image is computed by selecting for each pixel the disparity with the lowest matching cost.

Better results can be achieved by checking for left-right consistencies. This is done by first matching the images in normal way and then by matching the matched image to the base image. The two maps are compared, and the pixels are set to invalid if they differ more than by 1.

# Implementation Considerations

The algorithm is implemented in modern C++ using the OpenCV image processing framework. The user selects the left and right images, which are to be matched.

For the census cost a class is created which takes the maximum disparity named CensusTransformer. This class first performs a census transform using a window of 5 x 5 and encoding the result into an int32. The cost matrix is then calculated using the Hamming Distance over the transformed image for each disparity in the range [-max\_disparity, +max\_disparity]. The cost matrix’s rows are the image pixels locations and the rows are the disparity range.

The semi-global algorithm is implemented in the SgmMatcher class. This class can be configured for future-proofing to be used with the Mutual Information matching cost. The depth maps’ computation consists of the following steps:

* Initial cost computation using Census transform
* Cost aggregation
* Disparity computation for both images
* Median filter application
* Outliers invalidation

The cost aggregation is carried out in 8 directions in the image space: NE, N, NW, W, SW, S, SE, E. The cost aggregation in a direction begins with the initialization of the border pixels cost for each disparity with the values of the initial cost matrix. On each direction the current pixel’s cost is calculated using the above formula. The final cost for a pixel is the sum of costs in all directions. The starting pixels for a given direction are the border pixels located in the opposed direction in image space, e.g. for the N direction the cost aggregation begins from the bottom row.

The disparity computation is just a straight iteration over the disparities for each pixel in the image in the cost matrix and the one with the lowest cost is selected as the disparity.

The median filter applied over the image eliminates some of the random noise. The outliers are invalidated by checking the value for the disparity in the base and matching images. In the matching image the corresponding pixel on the epipolar line, calculated from the disparity is used.

The disparity image is then displayed using the library. The disparity image is also normalized to have a better contrast for the visualization.

# Results

The result for the cones image in Figure 1 is shown in Figure 2.

Figure Cones

Figure Cones Disparity

This method gave some mediocre results. This is probably because of the small window used in the census transform and the small number of directions for the cost aggregation. Better results would probably be obtained when using the Mutual Information score as well.

# References

1. H. Hirschmuller, Stereo Processing by Semi-Global Matching and Mutual Information
2. H. Hirschmuller, Semi-Global Matching – Motivation, Development and Applications