Automatic Calibration of 2-D and 3-D Point Correspondences

Daniel Aisen, Marcia Lee, Wenqi Shao

Background

Automating the calibration of 2-D to 3-D feature correspondences is at the forefront of image processing research. The challenge behind this problem is, given an image of a scene and scan data from a 3-D laser scanner from a slightly different perspective, to find a perfect mapping between the two. At the present, point and line features from a 2-D image must be manually specified to correspond to 3-D sensor scan data. Each mapping requires that a person select a small number of corresponding points ($n > 2$) from the two sources. This current standard requires more human input and time than is feasible on a large scale, and it is also prone to error.

Automated feature correspondence can be used, for example, in 3-D reconstruction of an image. Our project could also be applied to pose estimation, a method of deducing the vantage point from which an image was taken. It can also be used for object recognition by using relative depth information to identify element contours. The process of segmenting images by objects at different distances, such as separating the foreground from the background of an image, would also benefit from this work. This particular project was completed in the context of augmenting a calibration toolbox that currently uses manually specified 2-D to 3-D correspondences and would greatly benefit from speedier computer automation.

Methods

2-D and 3-D source data was collected in the form of an image and a depth map from several complicated scenes containing an unobstructed rectilinear calibration target. The camera and laser scanner were mounted in a rigid setup such that point correspondences are fixed across all pairs of scene data. A preliminary mapping between the two sources was calculated based on this distance.

Points of interest are selected from the 3-D data via a proprietary algorithm developed by Paul Baumstarck. This algorithm identifies likely rectilinear regions by clustering the 3-D data into planar segments, fitting hulls to these segments, converting them into best-fit rectilinear hulls, and finally preprocessing the results to identify good planes and points. The algorithm identifies a subset of rectilinear hull corners that are most likely to be hard corners, and these points are used for developing correspondences.

Potential rectilinear object corners are selected from the 2-D image Harris corner map, which provides a confidence level for every pixel in an image representing the likelihood of that pixel being a hard corner. Otsu's method, a thresholding algorithm, is then applied to this confidence map to identify an optimal minimum cutoff for selecting high confidence corners. Blob analysis was used to remove regions over-crowded with points, as these regions present too many degrees of freedom of matching.

Potential correspondences between 2-D and 3-D points are identified using the RANSAC method. The algorithm begins with the preliminary mapping. Each trial chooses a subset of three to five 3-D points of interest at random and matches each one to a random one of its most likely corresponding Harris corners based on the preliminary mapping. Each trial produces some calibration between the two data sources based on these correspondences. The quality of this calibration is then determined based on some heuristic function. Because the preliminary mapping is within a small margin of the ideal mapping, if the rotational and translational values are far from the preliminary values, the calibration is filtered out. A large number of trials are performed, and correspondences that produce good heuristic values are noted. Once several valid correspondences are identified across all sets of data, they can be combined to generate an optimal calibration.

In order to test potential correspondences, a heuristic function must be developed to evaluate the quality of a given calibration. A proper heuristic function should give a good value for an accurate calibration and a poor value for an arbitrary invalid projection. The integrity of the heuristic is tested via a method called bootstrapping, in which the value outputted after applying the heuristic function to our preliminary projection is compared to that of random projections. Finally, the heuristic undergoes testing on a proper manually chosen correspondence between 3, 4 and 5 points to determine whether the heuristic value improves for a proper mapping.

For each calibration, the heuristic function finds the best one-to-one mapping between the projected 3-D points of interest and the filtered Harris corners and then sums the distances between each pair. For any 3-D points that do not map onto the image, the maximum possible distance between a point on the image and its nearest Harris corner is used. The smaller the value produced by the heuristic function, the higher the quality of the calibration.

Results

Applying Harris corner selection to the image, and then thresholding the resultant confidence matrix, produced a reasonable set of 6408 corners to work with. However, there were several problematic regions that contained multiple nearby corners. After applying blob analysis these regions were successfully identified and removed, leaving only 1317 corners to work with. It is worth noting that many of these corners appear in clumps of pixels around a single element corner.

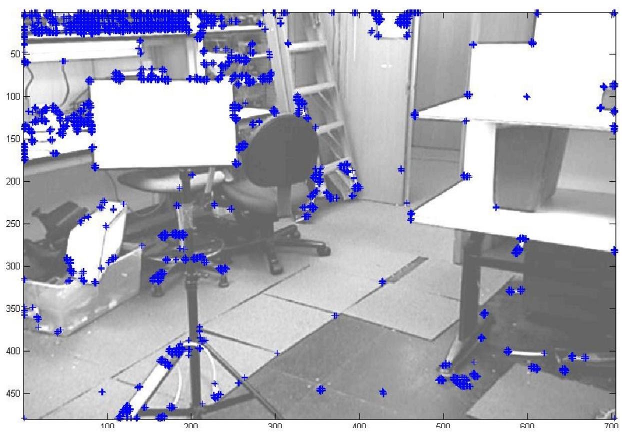

Figure 2: Harris corners after thresholding.

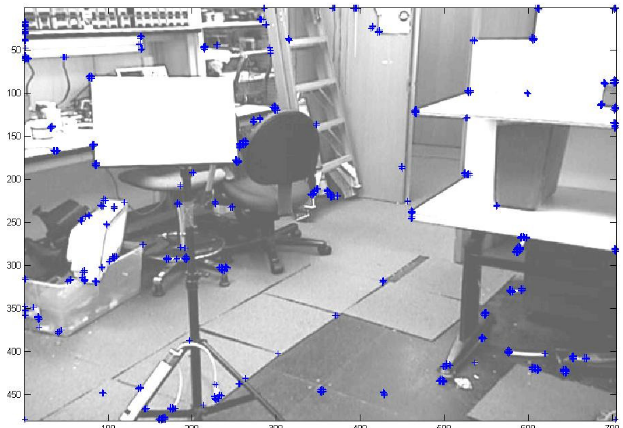

Figure 1: Harris corners after thresholding and blob analysis.

Once a reasonable set of 2-D and 3-D points of interest had been established, the heuristic function was implemented as described in the methods. Bootstrapping the heuristic by evaluating random projections produced promising results. While the preliminary calibration had a heuristic value of 435.3, a test of 100 random calibrations produces an average heuristic value of 1722.4 with a standard deviation of 254.7. The best random calibration had a value of 525.5, significantly worse than that of the preliminary calibration.

Figure 3: Original calibration. Heuristic = 435.3.

Figure 4: Random calibration. Heuristic = 933.0.

After successful bootstrapping, the heuristic was further tested by manually identifying proper correspondences between 3-D hulls and 2-D objects. These tests produced mixed results. The manual calibration based on three proper correspondences produced a calibration of low quality with a heuristic value of 747.7. A manual calibration based on 4 correspondences was moderately successful and had a value of 455.9. A manual calibration based on 5 correspondences was also reasonable-looking and had a heuristic value of 466.1.

Figure 5: Calibration based on 3 points – TL, BL, BR of easel. Heuristic = 747.7.

Figure 6: Calibration based on 4 points – TL, BL, BR of easel, BL of desk. Heuristic = 455.9.

Based on the performance of the heuristic, the RANSAC algorithm was performed for 20 iterations of 1000 trials, trying to find four proper correspondences between our 22 interesting 3-D points and our 1317 Harris corners. For each randomly chosen interesting hull corner, we examined its ten nearest Harris corners and chose one at random. If there was exactly one proper correspondence for each of the 3-D points of interest, and if this point was one of its ten closest Harris corners, then the likelihood of a random correspondence succeeding would be $1/10^4$. However, as not every interesting 3-D point has a corresponding 2-D corner (i.e. the top right corner of the easel), and as the corresponding corner might not be one of the nearest ten, in actuality the likelihood of success is significantly lower. In our test, the algorithm did not find a superior calibration, although more extensive testing may have led to positive results.

Future Research

One possible extension to our 2-D corner selection may be to consolidate pixel clusters around single element corners. This would provide many fewer points and greatly increase the chance of success via the RANSAC method. Additionally, as even manually corresponding a subset of points between the two data sources does not necessarily improve the mapping, it may be reasonable to develop a more sophisticated heuristic function that can evaluate the validity of just the corresponded points, rather than the projection as a whole. Furthermore, alternatives to the RANSAC method, such as prioritizing certain correspondences based on distance, might improve run-time and allow for more extensive testing.

Acknowledgements

This project was completed under the close guidance of the brilliant Paul Baumstarck.

References

Dornaika, F. and Garcia, C., Robust camera calibration using 2D to 3D feature correspondences, Proceedings of the International Symposium SPIE --Optical Science Engineering and Instrumentation, Videometrics V, Volume 3174, pages 123--133, San Diego, Ca., July 1997.

Harris, C.G. and Stephens, M.J.. "A combined corner and edge detector", Proceedings Fourth Alvey Vision Conference, Manchester, pp 147-151, 1988.

Kovesi, P. Harris corner detector. http://www.csse.uwa.edu.au/~pk/Research/MatlabFns/. Last accessed 12 Dec 08.

Noble, A., "Descriptions of Image Surfaces", PhD thesis, Department of Engineering Science, Oxford University 1989, p45.

Quan, L. and Lan, Z., Linear N-Point Camera Pose Determination, PAMI(21), No. 8, August 1999, pp. 774-780.