Julian Jandeleit

Camera Pose Estimation for Ushichka Dataset Status Report 3

IMAGE REFERENCE FRAME

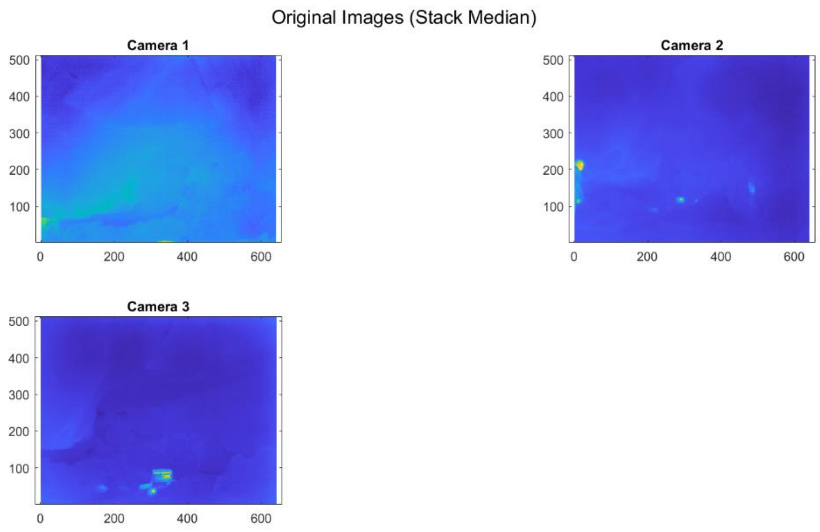

Figure 1

Figure 1 shows the base images as experimented with in this report. They are created as the median of the image stack of each camera respectively. This removed most evidence of both people who are present in the scene. Note that, in contrast to the previous reports, the origin of the image is in the bottom left, as is expected in the theoretical model. This mismatch was the cause of many problems in the past.

This problem was detected when comparing manually selected points and calibration points when applying the fundamental matrix (also computed from calibration).

RECTIFICATION

In this step, I was able to use the calibration from report 2 to compute a working rectification such that the epipolar lines of corresponding points are aligned on the same y coordinate. The rectification was computed by using a variation of the "compact algorithm for rectification of stereo pairs" [1] by Fusiello et. al. It computes two homographies that rotate the cameras to align their x-axes with the baseline connecting them. The rectified images then get translated by a vector t, such that the rectified images both lie in positive coordinates. This makes sure they can be warped backwards into an image represented by a matrix. Figure 2 shows the rectified images for pairs Camera 1 / Camera 3 and Camera 2 / Camera 3. Matching structures now lie at the same height but on different x-coordinates.

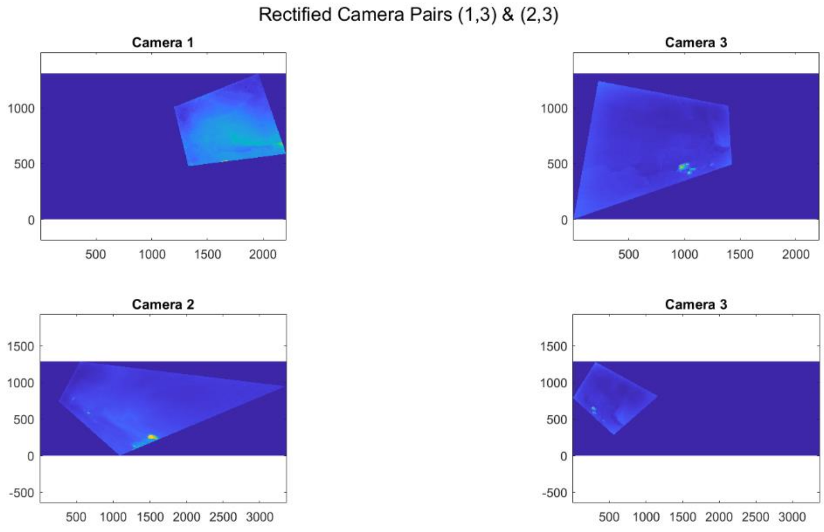

Figure 2

DENSE CORRESPONDENCE

The rectified images can now be used to compute a disparity map which contains the difference in x-direction of corresponding points. The rectification is necessary as it reduces the problem to a search along a one-dimensional scanline, which dramatically reduces computation cost. I tried several algorithms:

- Naïve patch comparison using sum of squared differences, implemented by myself. This was way to computationally expensive, as it probably didn't make use of MATLAB's internal parallelism. The results were also not very meaningful. Using the census-transform instead of sum of squared differences might improve quality.
- Semi global matching from MATLAB Toolbox only produces star and line shaped results that are completely unrelated to the image. The problem might be that it only allows a maximum disparity of 128 pixels. When looking at the rectified images in Figure 2, you can see that even when only looking at the warped image regions, the disparity is larger than 128.
- Block matching from MATLAB Toolbox provides useful results, at least in stereo pair (1,3). was reached. It was ±1024 for pair (1,3) and ±2048 for pair (2,3). Those ranges subjectively produced the most sensible result.

Figure 3 shows the block-matching result in rectified and original reference frame. The disparities are computed on the rectified image and the warp back into the original frame of reference. Pixels that are dark blue are NaN values, where no correspondence was found (lowest possible threshold was used). Bright yellow are values where the maximum disparity range is reached. The border-like structure visible in the disparity images might be the result of the transition between NaN values and the rectified image.

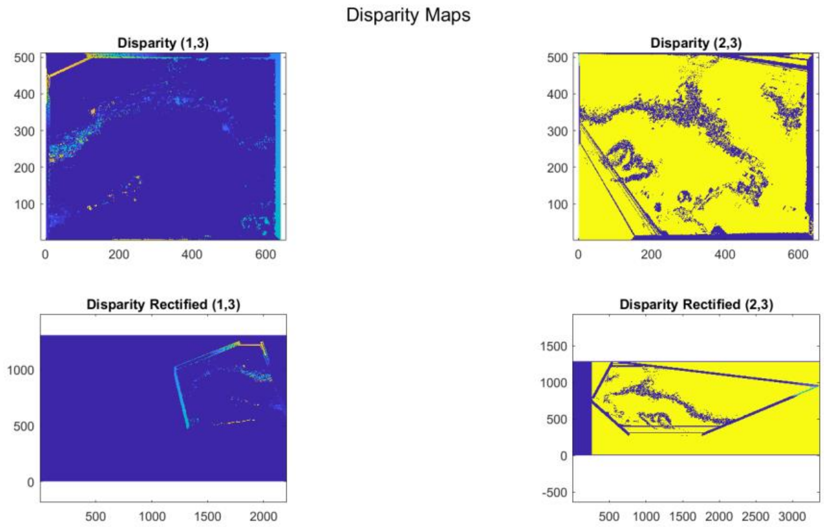

Figure 3

The resulting structures in the disparity image can be identified in the original images. However more correspondences and removing the border around the image are desirable, especially as the actual structure recovered from the second stereo pair is currently represented by NaN values instead of disparities.

3D POINT CLOUD

The disparities from the first disparity map could be projected into world space using the calibrated projection matrix for camera 1. The result was a ceiling-of-a-cave like structure, but the "border lines" dominate the result, so a 2D view for illustration purposes is not practical.

OUTLOOK

To me, the primary question now is how to improve the quality of the disparity map and to make it work reasonable on all camera pairs. I believe there is still room for significant improvement, although I'm not sure how big advances can still be made using conventional methods.

REFERENCES

[1] A. Fusiello, E. Trucco und A. Verri, „A compact algorithm for rectification of stereo pairs," *Machine Vision and Applications*, 2000.