Project Proposal

Stereo Reconstruction Using Image Rectification and PatchMatch

1 Abstract

Stereo matching and 3D reconstruction is the foundation of 3D computer vision. It can be differentiated into an active or passive process, that is used to acquire depth information from images. While in active stereo matching information is gained by projecting light into the scene, passive stereo matching is more similar to human vision: It requires two cameras with internal and external calibration permitted and both intrinsic and extrinsic parameters known to operate. With this binocular system, underlying geometric properties are used to extract depth information and reproduce a 3D representation of any object point represented as a pixel pair in both cameras' image planes.

The main challenge in stereo matching is solving the pixel correspondence problem, that is, finding points of the same object in both images. Since we know the exact pose of both cameras relative to each other, we can exploit the underlying geometry as follows: The search for corresponding pixels can be restricted to the so-called epipolar line that is formed by the triangulation method, when connecting one object point to both camera centers. By principle, the corresponding pixels in both image planes must lie on this line. The process of finding these pairs can be improved by rectifying both images first, such that the resulting epipolar line is horizontal, thus allowing for more efficient row-based iteration through the images [2].

Many different algorithms were proposed for the subsequent correspondence search based on sparse and dense feature detection and matching. For usual dense methods, the block-matching approach considers a small n-by-n neighbourhood of each pixel, which is used for efficient descriptor build-up. Each pixel is matched to obtain a dense representation of corresponding pixel pairs. At this step, a trade-off is faced between accuracy and reliability of the resulting image: By choosing different patch sizes, one can include more of the surrounding region, thus making the descriptor more robust to accidental matches. On the other hand, you lose the precision of using smaller regions by having more matches at a higher resolution. An optimal value needs to be found by evaluating various sizes.

The particular method used in this work is the PatchMatch algorithm [1]. The assumption here is that the performance of the correspondence search can be optimized by finding a few good pixel matches in both images, thus making the processing of the whole scene no longer necessary. By using random sampling of locations in both frames, reliable patch matches can be found and propagated quickly to their surroundings. This is only possible due to natural coherence in the imagery, from which a less dense repre-

sentation of correspondences is obtained. The resulting pixel matches are then used to reconstruct depth information using epipolar geometry, as stated above.

The resulting pipeline is reproduced in this project. An overview can be seen in figure 1.

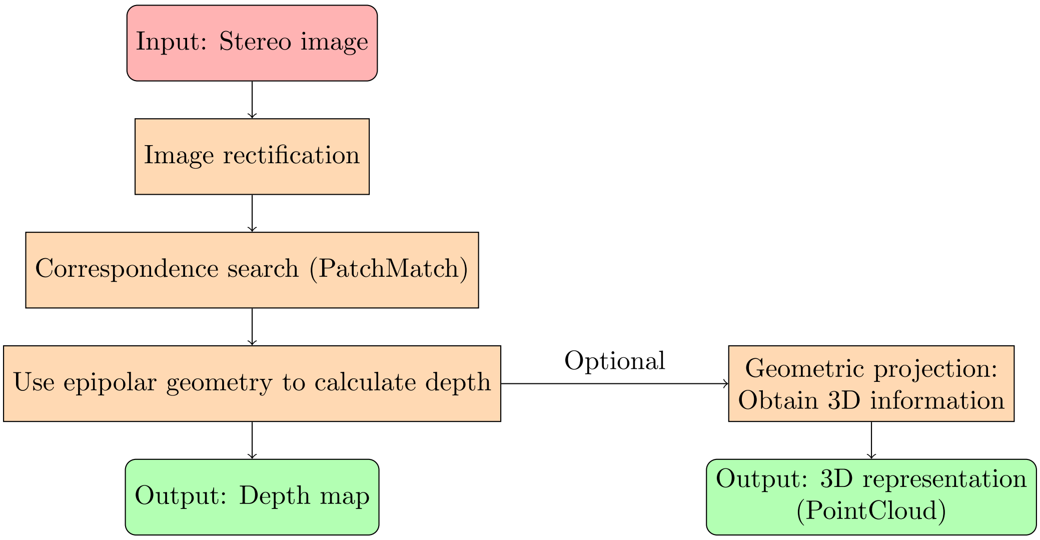

Figure 1: Pipeline overview.

2 Milestones

- CW 24: Setup of stereo camera rig. Installing drivers, understanding the API and receiving frames from the camera.
- CW 25: Implementation of image rectification.
- CW 26: Implementation of PatchMatch algorithm.
- CW 27: Implementation of depth calculation and representation.
- CW 28: Optimization of algorithm parameters (e.g., patch size)
- CW 29: Gathering results and preparation of poster/final talk.

3 Requirements

- Stereo RGB-camera with output of both images
  - alternatively: system of two fixed and calibrated cameras

4 Team

| Lukas Knak | Evgenija Pavlova | Gregor Wiese | Raphael Wild |

References

[1] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan B Goldman. PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 28(3), August 2009.

[2] C. Loop and Zhengyou Zhang. Computing rectifying homographies for stereo vision. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, volume 1, pages 125–131 Vol. 1, June 1999.