

RS: Epipolar Plane Image Analysis

Application to SkySat videos

Quentin CHAN-WAI-NAM

February 28, 2018

1 Choice of the article and objective

We first investigated two different articles that describe two different methods for estimating depths maps from videos.

- [2] describes an algorithm that estimates the optical flow given two images using a global optimization scheme minimizing a data attachment term and a regularization using the total variation of the flow. The idea is that the brightness I of single points along their trajectories should be constant in time, leading to the optical flow constraint equation

$$\nabla I \cdot \mathbf{u} + \frac{\partial I}{\partial t} = 0$$

where \mathbf{u} is the optical flow (the velocity vector field). The article then introduces a regularization on \mathbf{u} (its total variation) and reformulates the problem in order to adapt to discrete sequences of images so that the attachment term consists in minimizing some L^1 term. In order to solve the subsequent global optimization problem, the authors propose a numerical scheme based on alternate optimization scheme. Finally, the authors investigate the influence of the several parameters of the algorithm – noticeably the weight λ of the data attachment term – on the precision of the estimation of the optical flow and the sensitivity to noise.

In our case, the optical flow computed with this algorithm could be interpreted as some disparity measurement. An interesting point is that the computation of the optical flow can be done on any pair of images, even if not rectified.

- [1] describes a method for computing precise and exhaustive depth maps using “light fields”, that is a dense set of images captures along a linear path. By concatenating one line of the rectified images together, one obtains an “epipolar-plane image” (EPI), in which a single scene point appears as a linear trace whose slope is related to its distance to the camera. Thus, by estimating these slopes, one can reconstruct the depth of each point of the scene.

The authors use very high definition images, so that the article includes several implementation details in order to ensure computational feasibility, both in terms of space (sparse representation of light fields) and computational power. For instance, they prefer local optimization near object boundaries and propagation to nearby areas in a fine-to-coarse approach to global optimization on

the whole image. In the end, the method seems relatively fast, precise and robust to inconsistencies and outliers like noise or temporary occlusions.

We chose to work on the latter article. The objective of the project is then to implement the method presented by Kim et al. in [1] and investigate how it performs with videos taken from SkySat. We will first investigate the case in which the images from the video are pre-rectified. Then we will investigate more complicated situations, for instance increasing the density of the images and considering non-pre-rectified sequences.

References

- [1] Changil Kim, Henning Zimmer, Yael Pritch, Alexander Sorkine-Hornung, and Markus Gross. Scene reconstruction from high spatio-angular resolution light fields. *ACM Trans. Graph.*, 32(4):73:1–73:12, July 2013.
- [2] Javier Sánchez Pérez, Enric Meinhardt-Llopis, and Gabriele Facciolo. TV-L1 Optical Flow Estimation. *Image Processing On Line*, 3:137–150, 2013.