Estimating Rigid Motions from Stereo

David Demirdjian

Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: This work addresses the problem of 3D rigid motion estimation from stereo. The appealing feature of this approach is that it directly uses disparity images and do not require explicit 3-D (Euclidean) reconstructions.

Motivation: We assume that the stereo rig has parallel cameras and show, in that case, the geometric and topological properties of the disparity images. Then we introduce a rigid transformation (called *d-motion*) that maps two disparity images of a rigidly moving object. We show how it is related to the Euclidean rigid motion and a motion estimation algorithm is derived.

Previous Work: Many approaches for ego-motion and motion estimation with a stereo rig have a common structure. First, image points are matched in the image pair. A disparity image is obtained, and a scene reconstruction is performed in the Euclidean space. Then the rigid motion is estimated based on 3-D point correspondences.

Unfortunately using the 3-D Euclidean space to estimate rigid motions fails to give optimal solutions because reconstructions performed in 3-D Euclidean space have non homogeneous and non isotropic noise (as wrongly approximated in many standard SVD- or quaternion-based approaches). One should instead use methods that deal with non homogeneous and non isotropic noise [4, 3] or methods that look for an optimum solution for both structure and motion, like *bundle-adjustment* [1].

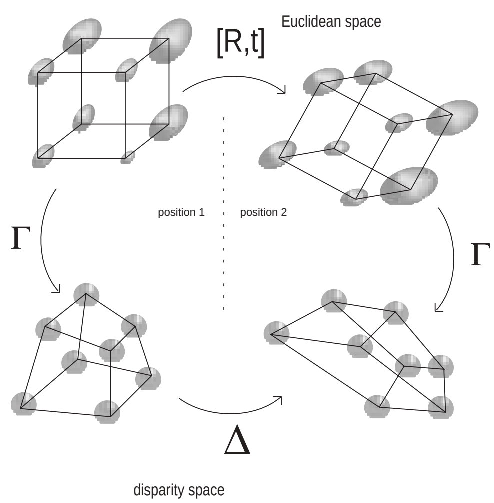

Figure 1: Euclidean reconstruction and motion of a cube *vs.* reconstruction and motion in the disparity space.

Approach: The disparity images have a well-behaved noise (theoretically isotropic for parallel stereo images) and can be used instead of the 3-D Euclidean space for motion estimation in this paper. The disparity images are related to the 3-D Euclidean reconstruction by a projective transformation. We show that two disparity images of a rigid scene are related by a transformation (that we call *d-motion*) that is fully described in [2].

There are many theoretical advantages of estimating the motion from the disparity space and d-motion. Minimizing the 'transfert error' gives an accurate estimation of the motion because, for parallel camera stereo rigs, the noise of points in the disparity space is isotropic (and nearly homogeneous when the reconstruction is restricted to well textured points). Therefore such a minimization gives a (statistically) quasi-optimal estimation.

Future Work: A topic of ongoing and future work is the use of multi-modal noise distribution to model disparity uncertainties. Introducing that model in a stereo matching algorithm would give multiple-hypothesis disparity images, where each image pixel could have one or multiple disparities. A robust algorithm should be able to estimate the d-motion from two multiple-hypothesis disparity images corresponding to a rigid moving object.

References:

[1] D. Brown. The bundle adjustment - progress and prospect. In *XIII Congress of the ISPRS,* 1976.

[2] David Demirdjian and Trevor Darrell. Motion estimation from disparity images. Technical Report AI Memo 2001-009, MIT Artificial Intelligence Laboratory, May 2001.

[3] B. Matei and P. Meer. Optimal rigid motion estimation and performance evaluation with bootstrap. In *Proceedings Computer Vision and Pattern Recognition Conference*, 1999.

[4] N. Ohta and K. Kanatani. Optimal estimation of three-dimensional rotation and reliability evaluation. In *Proceedings of the 5th European Conference on Computer Vision*, 1998.