Two Correspondence Problems Easier Than One

AARON MICHAUX and ZYGMUNT PIZLO
Purdue University, West Lafayette, Indiana

Abstract

Computer vision research rarely makes use of symmetry in stereo reconstruction despite its established importance in perceptual psychology. Such stereo reconstructions produce visually satisfying figures with precisely located points and lines, even when input images have low or moderate resolution. However, because few invariants exist, there are no known general approaches to solving symmetry correspondence on real images. The problem is significantly easier when combined with the binocular correspondence problem, because each correspondence problem provides strong non-overlapping constraints on the solution space. We demonstrate a system that leverages these constraints to produce accurate stereo models from pairs of binocular images using standard computer vision algorithms.

System Architecture

Corpus images (as described below) are processed through a three-stage pipeline. In stage 1, a disparity map is calculated using *Sum of Absolute Differences*. As described in Li et al. (2012), this disparity map is used to estimate a region of interest for a given piece of furniture by fitting a 3-orthotope (i.e., box) to a subset of 3D points. Two binary edge maps are then calculated using the canny edge detector. The disparity map is examined further to estimate 3D locations for every point on the edge maps of the two images. This sparse point cloud is used as the basis for processing stages 2 and 3.

Stage 2 processing makes a precise estimate of all three parameters of the veridical plane of symmetry for the imaged piece of furniture. Here we take advantage of the non-overlapping constraints provided by the symmetry and binocular correspondence problems. The two edge maps are examined to find sets of four points – two from each image – that simultaneously obey the constraints of both correspondence problems. Specifically, two points in each image must be col-linear with the vanishing point, and further, points from alternate images must form pairs that obey the epipolar constraint. The collection of these sets of points is referred to as the inlier set. (Inliers are thresholded by their reprojection error.) Nelder-Mead is then used to find the two parameters of the vanishing point that produce a local maxima in the number of inliers. The axes of the fitted 3-orthotope are used to estimate the initial vanishing point; however, it is also possible to use random vanishing points to segment the image into spatially local sets of symmetric points.

The constraints in stage 2 produce recognizable but sparse point-wise 3D figures. Stage 3 leverages this information – specifically, the full four parameters of the symmetry plane – to find a more complete figure under the assumption that it is made from geometric primitives. RANSAC is used in a greedy search to fit symmetric 3D line segments to binocularly corresponding points, simultaneously across both images. Self-symmetric straight lines require special treatment, and are recovered by identifying the end-points of the inlier set after it has been projected onto each image plane.

The Toy Furniture Corpus

The *toy furniture corpus* is a set of 90 pairs of images of 11 different pieces of toy furniture, acquired with a Point Grey Bumblebee®2 1394a stereo camera system. The furniture was centered on plain carpet in each image, providing minimal background clutter.

Experimental Results

Stage 2 symmetry planes are accurately calculated, even for complicated objects. Furthermore, almost all line segments are recovered without the characteristic noise present in binocular reconstructions. There are some false-positive lines, but they are always short, and complicate the figure. When line segments are missed, it is invariably because of occlusion – a problem we do not yet address. An example result is given in figure 1.

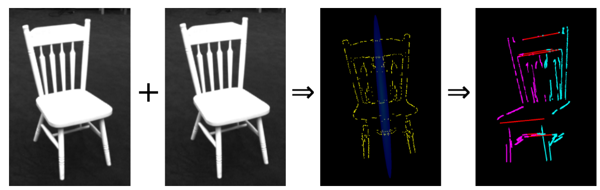

Figure 1: Two binocular input images and the last two stages of the pipeline. Center-right shows the stage 2 symmetry plane estimate. The final image shows the fitted lines from stage 3. Line color indicates symmetry, i.e., the "left" and "right" side of the figure.

Conclusions and Further Research

Although initial results from this procedure are promising, further processing is required to produce complete objects. An intriguing possibility involves applying a richer set of geometric constraints, such as finding symmetric planar curves within both correspondence problems. Line segments and curves might be refined further by stitching together end-points – where the reprojective error warrants it. Thus end-points and corners are revealed, a notoriously difficult problem to solve. Locating end-points and object planes over-constrains the estimation of geometric primitives parameters. This implies that collections of primitives can be optimized over a lower dimensional subspace. A future research goal is to algorithmically find the set of primitives that explains the visual evidence using the fewest degrees of freedom – a mathematically rigorous conception of *shape*.

References

Li, Y.; Sawada, T.; Latecki, L. J.; Steinman, R. M.; and Pizlo, Z. 2012. A tutorial explaining a machine vision model that emulates human performance when it recovers natural 3d scenes from 2d images. *Journal of Mathematical Psychology* 56(4):217–231.