Figure-Ground Organization using 3D Symmetry

AARON MICHAUX, VIJAI JAYADEVAN, EDWARD DELP, and ZYGMUNT PIZLO
Purdue University, West Lafayette, Indiana

Abstract

We present a novel approach to object localization using mirror symmetry as a general purpose and biologically motivated prior. 3D symmetry leads to good segmentation because (i) almost all objects exhibit symmetry, and (ii) configurations of objects are not likely to be symmetric unless they share some additional relationship. Furthermore, psychophysical evidence suggests that the human vision system makes use symmetry in constructing 3D percepts, indicating that symmetry may be important in object localization. No general purpose approach is known for solving 3D symmetry correspondence in 2D camera images, because few invariants exist. Therefore, to test symmetry as a clustering mechanism, we disambiguated the symmetry correspondence problem with the epipolar geometry of the binocular correspondence problem in order to simplify both. Mirror symmetry is a spatially global property that is not likely to be lost in the spatially local noise of binocular depth maps. Furthermore, each of these two correspondence problems provides non-overlapping constraints that makes it easier to solve both at once rather that each individually. We tested our approach on a corpus of 60 images collected indoors with a stereo camera system. K-means clustering was used as a baseline for comparison. The informative nature of the symmetry prior makes it possible to cluster data without a priori knowledge of which objects may appear in the scene, and without knowing how many objects there are in the scene.

Introduction

A typical forward model of vision takes a 3D world and projects it onto a 2D surface (retina, or camera sensor). Marr (1982) defined vision as the process of transforming that 2D array of information into a 3D description of the world. This 3D description is, of course, inherently underdetermined, making vision an ill-posed inverse problem. Tsotos (2011) points out that the human vision system copes with this by solving a computationally tractable approximation of the general vision problem – through the use of priors. Given the superior speed and flexibility of the human vision system, we believe in the importance of studying human-vision priors by applying them to the domain of computer vision. Li et al. (2013) make the case that symmetries, and in particular 3D symmetry, play a central role in 3D shape perception in humans. We therefore developed an algorithm that performs FGO by localizing objects in a scene using 3D symmetry as a shape prior.

There is no general solution to solving the symmetry correspondence problem in 2D images; however, it is possible to disambiguate it with binocular information. That is, pairs of symmetrically corresponding points in each image must also obey the binocular epipolar geometry if they represent the same pair of veridical (3D symmetric) points. We call these four points – two in each image – a *mirror symmetric two-view quadruple*, or simply *quadruple* for short. This allows for an operational definition of a 3D object as a maximal set of quadruples that obey a single plane of symmetry. We then extend this definition to FGO with the operational definition that an FGO consists of the set of non-overlapping 3D object hypotheses with the maximum collective number of quadruples. This definition of FGO is a form of clustering, so we called it *symmetry based clustering* (SBC).

Experimental Results

SBC was tested on a new corpus of 60 pairs of images of collections of 23 different exemplars, such toys, people, and childrens' furniture. Images were acquired with a Point Grey Bumblebee®2 1394a stereo camera system with a 66° horizontal field of view. The exemplars were placed upright at random locations on plain carpet, providing minimal background clutter.

We used K-means as a baseline for comparison, because it can naturally take advantage of 3D information available in binocular information, and also because of its long history, and its reliance on distance metrics commonly used in clustering problems. We used the well established Caliński and Harabasz (1974) method to automatically determine K, the number of objects in the scene. The experiment, therefore, compared symmetry as an informative prior, versus euclidean 3D distance as an uninformative prior, in FGO.

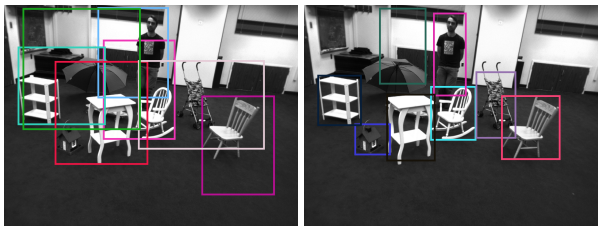

Figure 1: Left: K-means clustering results, versus, Right: 3D symmetry based object localization.

SBC performance was superior than K-means, in its ability to correctly determine the numbers of objects ($p < 0.01$), and correctly localize individual objects, as tested by intersection-over-union against ground-truth annotations greater than 50%.

This basic investigation of SBC indicates that more sophisticated research will prove useful in solving FGO, which is a fundamental and non-trival early stage of processing in the human vision system.

References

Caliński, T., and Harabasz, J. 1974. A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods* 3(1):1–27.

Li, Y.; Sawada, T.; Shi, Y.; Steinman, R.; and Pizlo, Z. 2013. Symmetry is the sine qua non of shape. In *Shape perception in human and computer vision*. Springer. 21–40.

Marr, D. 1982. Vision: A computational investigation into the human representation and processing of visual information, henry holt and co. *Inc., New York, NY* 2.

Tsotsos, J. K. 2011. *A computational perspective on visual attention*. MIT Press.