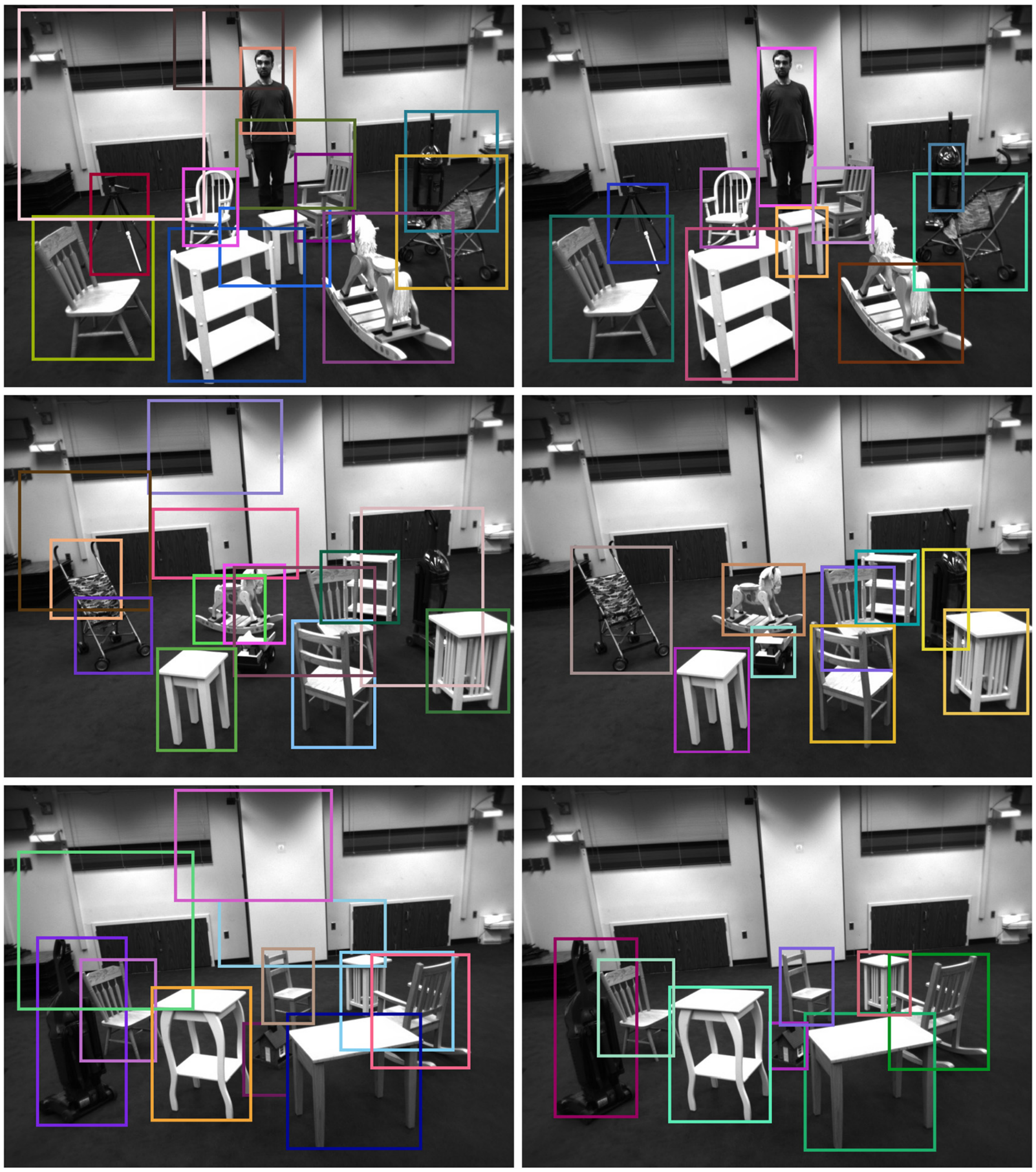

Fig. 3 Representative results comparing *K*-means (left column) to symmetry-based object localization (right column).

Sec. 2.4. When the objects are too far from the camera, then the stereo system degenerates to single view geometry. For example, an object 0.3-m deep, placed 4 m from our stereo camera setup, has only 2 pixels of disparity difference between front and back. Therefore, when constructing the corpus, objects were placed between 1.5 and 4.5 m from the stereo camera.

2-D ground truth was specified as a set of bounding rectangles in the left-camera-image of each pair of images. Each bounding rectangle was drawn by hand around the regions containing the individual exemplars. The rectangles were axis-aligned such that they have horizontal and vertical edges.

4.1 *K-Means*

Since mirror symmetry is calculated in 3-D, we applied *K*-means clustering to an unstructured 3-D point-cloud. A disparity map was calculated as per Sec. 3. The symmetry-based algorithm considered only binocular correspondences that coincided with the canny edge maps generated from each image. This sparse point cloud was used to reduce the search space for symmetrical correspondences; however, it also reduces noise from texture-based artifacts typical in stereo reconstructions. In order to do a fair comparison, we restricted the "*K*-means" point cloud to the same set of 3-D points that were used to generate two-view mirror symmetry quadruples. The method described by Caliński and Harabasz[40] was used to automatically determine the value of *K*: the number of objects in the image. Once clusters were determined, the 3-D points were projected back to the image plane of the left camera, and bounding rectangles with horizontal and vertical sides were calculated.

4.2 *Comparison Function*

Bounding rectangles were compared as follows. Intersection-over-union[41] [Eq. (13)] was used to calculate the best matches