Figure 5.8 shows matching output for 4 other stereo pairs. The matcher was run twice on the `test` and `blocks` images in order to recover the appropriate epipolar geometry and disparity range (results of second pass shown). For the `roof` images, these were obtained by hand-matching four points, and for the `lab` scene, using the INRIA corner matching system; ground truth calibration data were not available. In all cases a linear epipolar constraint is used. Results are good, considering that no surface shape constraint has yet been imposed; however there are a number of false correspondences — typically unconnected short segments without a true match which become associated with one another more or less randomly.

Reconstruction is not attempted at this stage.

5.4.8 Complexity analysis

Feature extraction. Line segment extraction is based upon edge detection, which is quite slow on general-purpose hardware due to the convolution stage. This takes a constant time for a given image area, and is independent of the number of features. For simple 'blocks world' scenes, edge detection is the slowest part of the entire stereo algorithm (taking about 10 seconds per image).

Monocular relations. Although bucketing is used to reduce the complexity in simple scenes, in clutter the search for related pairs reverts to $\mathcal{O}(n^2)$. Let the number of relations per line segment be r (usually much less than n).

Enumeration of candidate matches. As above, this has complexity $\mathcal{O}(n^2)$. Let the number of candidate matches be nm.

Enumeration of FRIENDS. This involves enumerating the pairs of candidate matches for each pair of segments related in one image. Complexity is therefore $\mathcal{O}(nrm^2)$ which in clutter tends to $\mathcal{O}(n^3)$.

Constraint propagation. The uniqueness constraint makes this efficient,[7] since each iteration need only traverse the list of candidate matches once ($\mathcal{O}(nm)$) to find the subset of at most n with more support than any uniqueness-RIVAL. Each of these must be compared with $\mathcal{O}(n)$ other promoted matches to test for RIVALRY, and the number of iterations is itself bounded above by n. In the

---

[7] Even where multiple collinear matches abound, their number is implicitly limited by the criteria for a candidate match, which each must meet.

absence of FRIENDS, overall complexity is between $\mathcal{O}(n^2)$ and $\mathcal{O}(n^3)$. With f FRIENDS per match, time is required to adjust the support of the FRIENDS of matches which are destroyed, with complexity $\mathcal{O}(nmf)$.

It is noted that the complexity depends not only on the number of line segments in the images but also on the number of *junctions* and other relations between segments (which also affects the number of FRIENDS). Where this is small, overall complexity is $\mathcal{O}(n^3)$ but in pathological images (such as when all line segments radiate from a point), r tends to n, f tends to nm, and complexity approaches $\mathcal{O}(n^4)$.

5.5 Coplanarity grouping

We now consider the integration of plane grouping into the stereo system, using the paradigm of [34, 118] and others. Coplanar sets of segments are identified by consensus with an affine transformation between views. In uncalibrated stereo it is important that coplanarities be detected before full 3-D reconstruction is attempted, so as to reduce the disparity errors caused by epipolar misalignment. Rather than occurring after correspondence, plane hypothesis and grouping are incorporated into the cooperative matching stage and introduce a shape constraint (see table 5.2, p98).

5.5.1 Plane hypothesis formation

A plane hypothesis is an affine transformation between views, defined by a set of three (or four, if there is parallelism) matching segments called a *seed*. It would be computationally expensive to enumerate all the triplets of candidate matches, so seeds are generated only from promoted matches, and heuristics employed to choose the seeds most likely to lie on planar facets.

Three forms of seed were considered:

- **Triangle form** (figure 5.9a). This consists of three matching segments, whose intersections in two images lie close to corresponding epipolar lines (i.e. they are pairwise coplanar, allowing for epipolar misalignment), and whose direction vectors in ($u, v,$ *disparity*) space form a triple product close to zero (i.e. their directions in 3-D are approximately coplanar). To reduce the search space, we require that two of the intersections be junctions in both images. A triangle gives a minimal definition of an affine transformation between views.

- **Parallel form** (figure 5.9b). Because many facets are rectangular, a second form of seed was allowed, in which two segments are parallel, and may even be horizontal, and are joined by a third segment so that the intersections in two images lie close to corresponding epipolar lines. The lines are likely to be coplanar but do not define a unique affine transform. Therefore an endpoint or junction with a *fourth* segment (which need not itself be coplanar) is required to fully define the transformation.

- **Parallelogram form** (figure 5.9c). Parallelograms in the images are often produced by rectangular facets in the world, but due to edge fragmentation these do not always yield seeds of the above form. Therefore a variation of the parallel seed was introduced, formed from two pairs of parallel segments which meet at two opposite junctions and define more than half of the perimeter of a parallelogram. The junctions and one of the other intersections are used to define the affine transformation.

Plane hypotheses are systematically enumerated out of the promoted matches as these emerge from the correspondence process; each is then tested for consensus with previously promoted matches as well as with candidate matches.

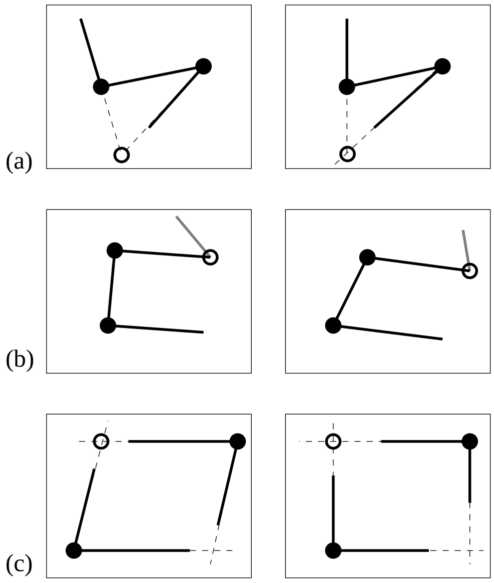


Figure 5.9: Forms for plane seeds, consisting of 3 or 4 matching segments: (a) triangle form; (b) parallel form; (c) parallelogram. In each case, two intersections (blobs) and an additional endpoint or intersection (circle) provide a minimal basis for the affine transformation.

5.5.2 Hypothesis testing and coplanar segment support

To test for coplanarity, promoted or candidate matches are tested for consensus with the affine transformation defined by the seed (figure 5.10). The test consists of 'transferring' the line defined by the segment from one view into the image coordinate frame of the other, and comparing it with the matching segment. The test is performed symmetrically in both directions between the views. No assumption is made about endpoint correspondences,[8] but the segments are required to 'overlap' on the plane by at least 33% of each segment's length.

Although Canny's algorithm can detect step edges to sub-pixel accuracy [14], it can produce correlated errors which do not manifest themselves in the residual errors after line fitting. Rather than propagating the residual errors, a simple threshold of 2.0 pel normal offset at the endpoints was chosen (cf. [52]). Candidate matches consistent with one or more hypotheses received extra support in the matching process (table 5.2).

Complexity

The number of plane hypotheses is proportional to the number of triples of connected line segments, which is approximately nr^2. Let there be h plane hypotheses. The testing of promoted and candidate matches for consensus with each plane hypothesis increases the computational complexity of the matching/grouping algorithm by $\mathcal{O}(nmh)$. In most scenes, h/n is small and complexity varies $\mathcal{O}(n^3)$, so the speed of the algorithm is not significantly reduced.

Results

The above forms of plane seed were successful at identifying most of the planes in the test images without generating many false hypotheses, although planes without three connected edges were missed. Many of the facets produced several hypotheses, but these did not always return the same consensus set; and occasionally a segment was wrongly grouped as belonging to two or more conflicting hypotheses. Such problems are inherent in any threshold-based test, in which noisy inliers cannot be perfectly separated from nearby outliers. Figure 5.11 shows successful plane seeds

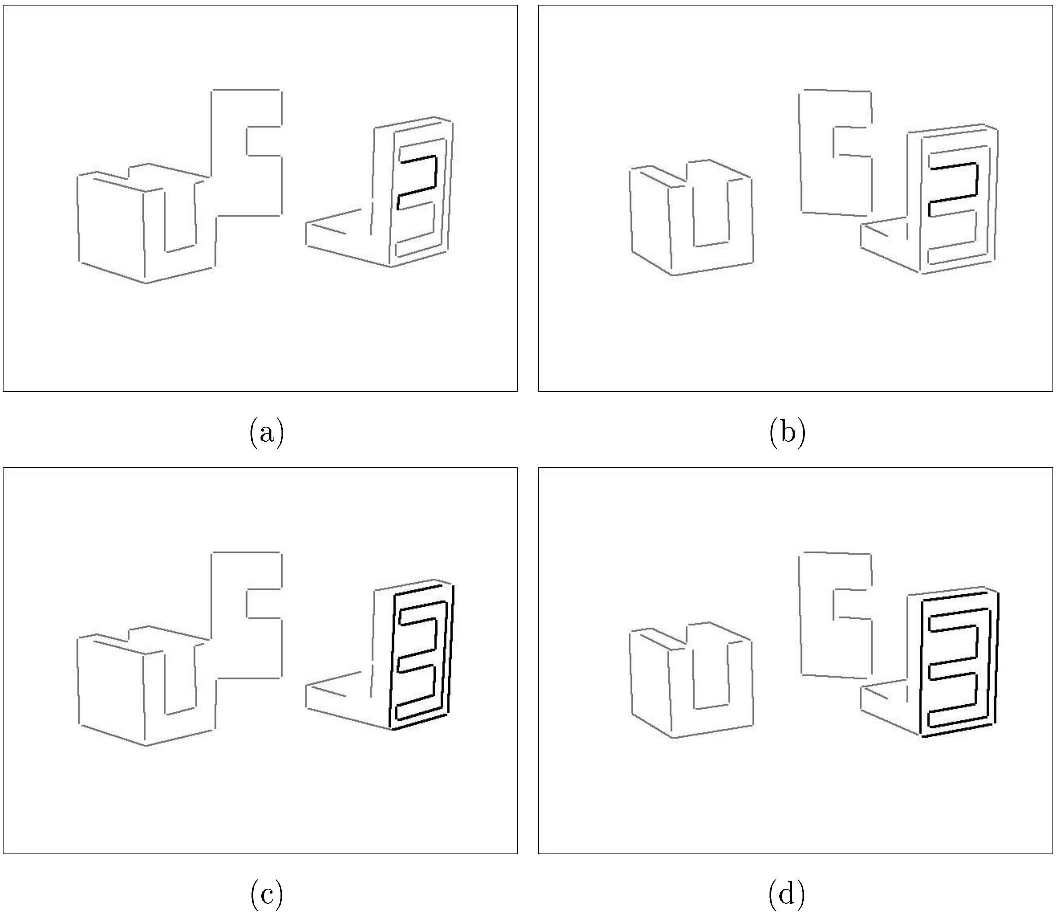
---

[8] Segments at a small angle ($< 10^\circ$) to the epipolar lines must also have at least *one* endpoint consistent with the transformation. This inelegant extra constraint is necessitated by the degenerate behaviour of horizontal lines.

and their consensus sets on the `cube` pair.

Despite the varying accuracy of coplanarity grouping with different seeds, it was found that candidate matches belonging to a coplanar group were almost invariably correct matches — we conclude that grouping by common affine transformation is valuable as a *matching constraint* as well as a cue to reconstruction. By giving these matches extra support in the disambiguation process, a significant improvement was seen in the accuracy of correspondence.

Figure 5.12 shows the matched segments on the `test`, `roof`, `lab` and `blocks` scenes which were consistent with one or more plane hypotheses — matches are sparser than figure 5.8, but more reliable.



(a) (b)

(c) (d)

Figure 5.10: Plane grouping by consensus: (a,b) a 'seed' used to generate a plane hypothesis; (c,d) segments consistent with the affine transformation.

```
• for each pair of line-segments intersecting corresponding epipolar bands
    • evaluate the pair as a candidate match
    • if criteria met, calculate intrinsic support
• for each candidate match
    • enumerate FRIENDS (using monocular figural relations)
    • each match gives extra +ve support to its FRIENDS
• while candidate matches remain
    • enumerate winning matches with more support than any RIVAL
    • sort winners by support difference over nearest RIVAL
    • for a proportion of the least ambiguous winners
        – promote winning match to a confirmed match
        – for each of its RIVALS
            ∗ withdraw support from its FRIENDS
            ∗ destroy the rival match
        – give extra support to FRIENDS
        – enumerate plane seeds that can be formed with promoted matches
    • for each new plane seed
        – for each other candidate or promoted match
            ∗ test for consensus with affine transformation
            ∗ give +ve support to coplanar candidate matches
        – if there are $\geq 4$ supporting matches, record plane hypothesis
```

Table 5.2: The combined matching/grouping algorithm, which forms line segment matches, propagates FRIEND and RIVAL constraints, and generates plane hypotheses which help to guide the matching process.

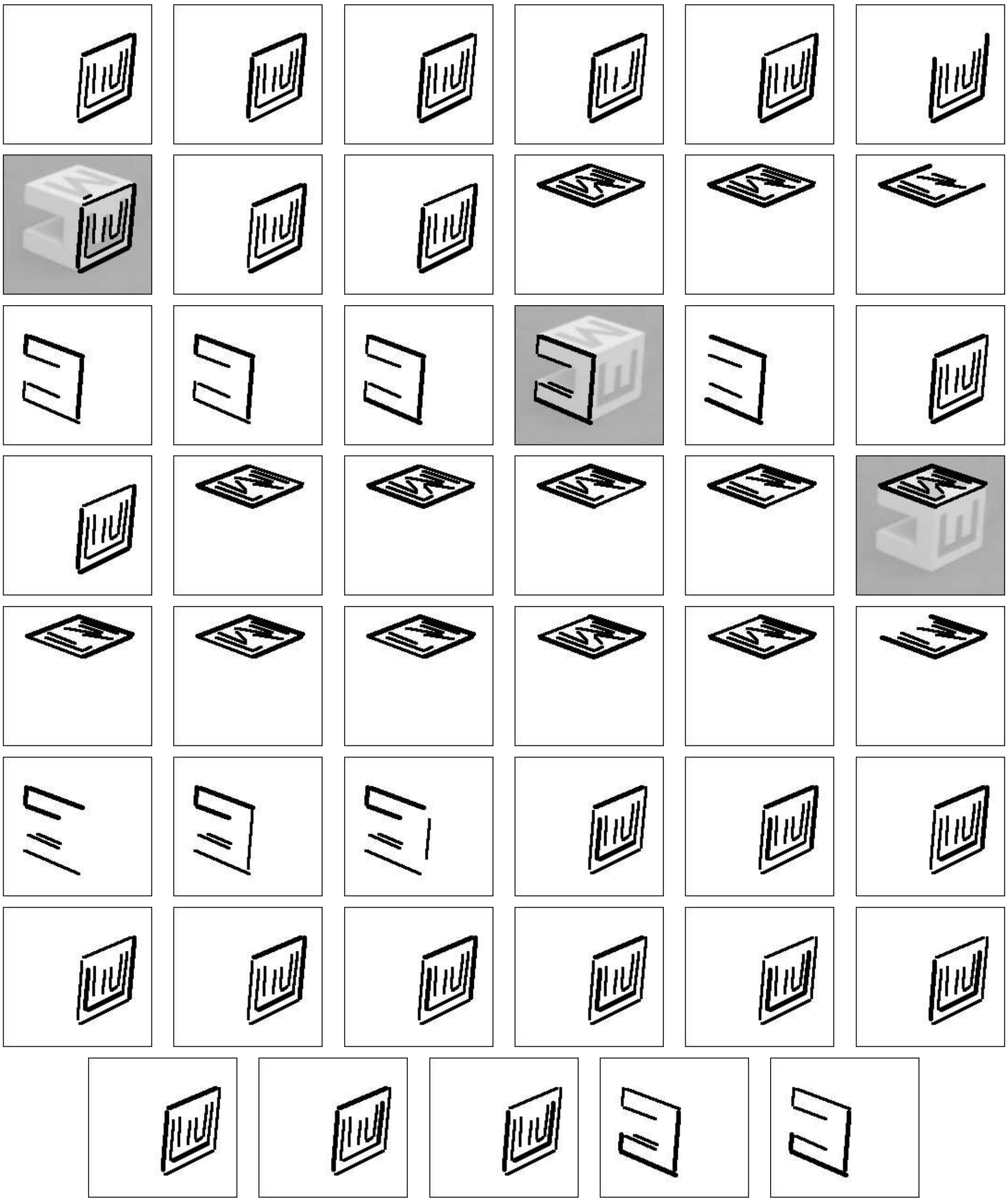


Figure 5.11: Plane seeds (bold) and final matches consistent with the affine transformation for the `cube` pair. The seed on each plane with the largest consensus set is shown against the original image (left view). Apparent duplicates are where different junctions have been used to disambiguate a parallel seed.

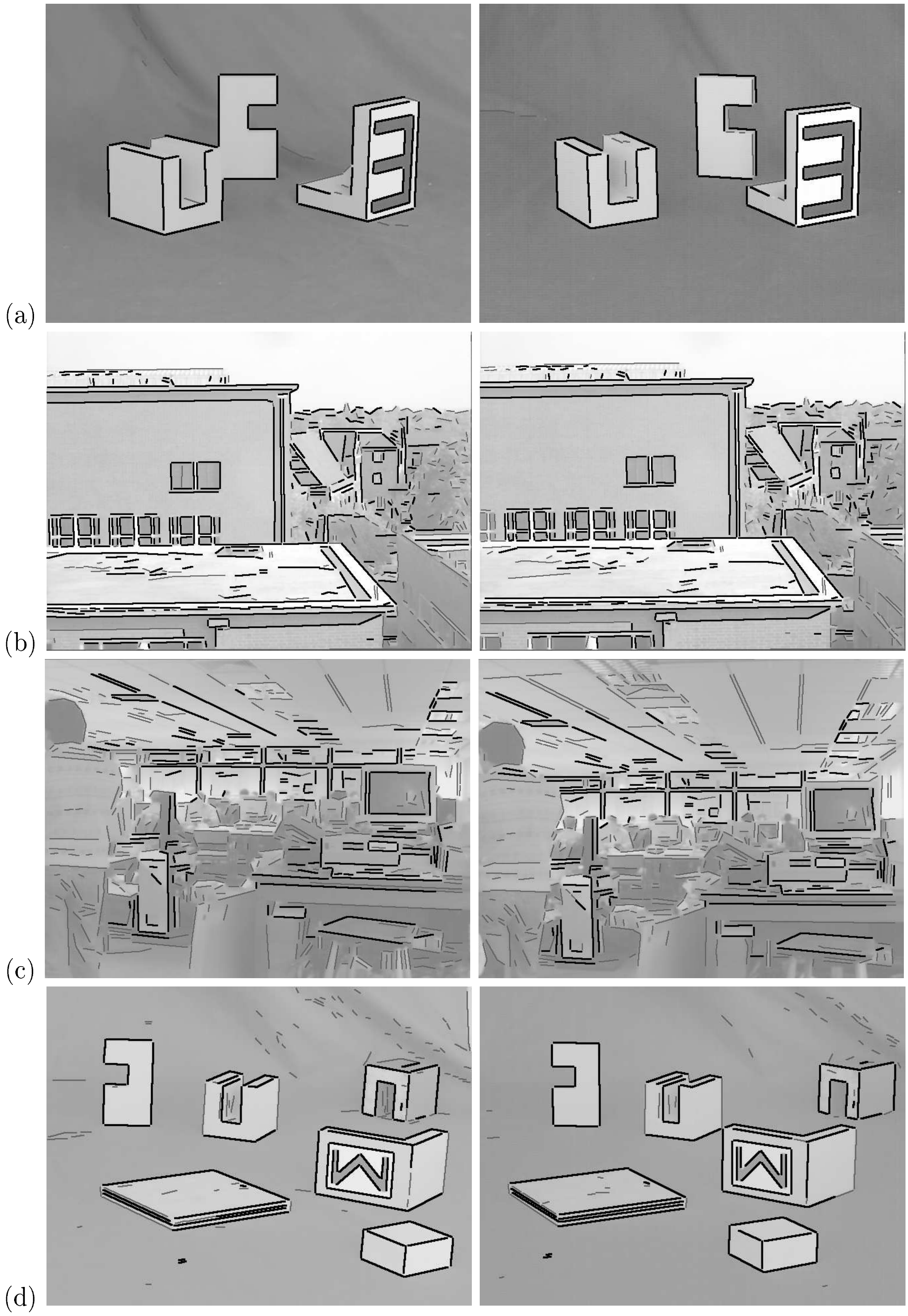


Figure 5.12: Matching results for the `test`, `roof`, `lab` and `blocks` scenes, showing only matches consistent with one or more plane hypotheses (in black).

5.5.3 Hypothesis selection

We have now obtained reliable correspondences for line segments, and a set of plane hypotheses consisting of an affine transformation between views. This may be quite large, and contain multiple hypotheses for each plane. We must therefore select an appropriate set of hypotheses to form a global segmentation of the scene into planar groupings.

If we can assume that the scene contains a discrete set of planes, rather than gently curving surfaces, we would expect the consensus sets of plane hypotheses to exhibit *convexity* in the neighbourhood of each plane — that is, the most representative model for each surface will be the one with the largest number of supporting segments, and hypotheses further from the actual plane model will have smaller consensus sets.

The following rule is therefore applied to select locally optimal plane models, and reduce the number of hypotheses: **If a plane hypothesis has at least half of its consensus set of edges in common with another hypothesis, which has a larger or equal consensus set, the first hypothesis is discarded.** To prevent unnecessary deletions, the hypotheses are first sorted into descending order of support.[9] The rule is then applied in turn to each pair until all duplicates have been deleted.

Results

The `cube` images are correctly segmented into the three planes marked in figure 5.11. Note that one short edge segment is incorrectly assigned to both the top and right faces of the cube. Figures 5.13 to 5.16 show the plane models recovered from the `test`, `lab`, `roof` and `blocks` scenes.

The `test` and `blocks` images yield a small set of distinct planes, which include all of the prominent planes in the scenes. Several of the edge segments are (correctly) grouped as belonging to two adjacent facets; there are also a few ‘accidental’ groupings of unconnected segments which are approximately coplanar with another facet. Grouping accuracy is noticeably coarser for edge segments close to the horizontal (the approximate direction of the epipolar lines).

---

[9] Where hypotheses have the same size of consensus set, they are ordered according to the total *length* of supporting edge segments.

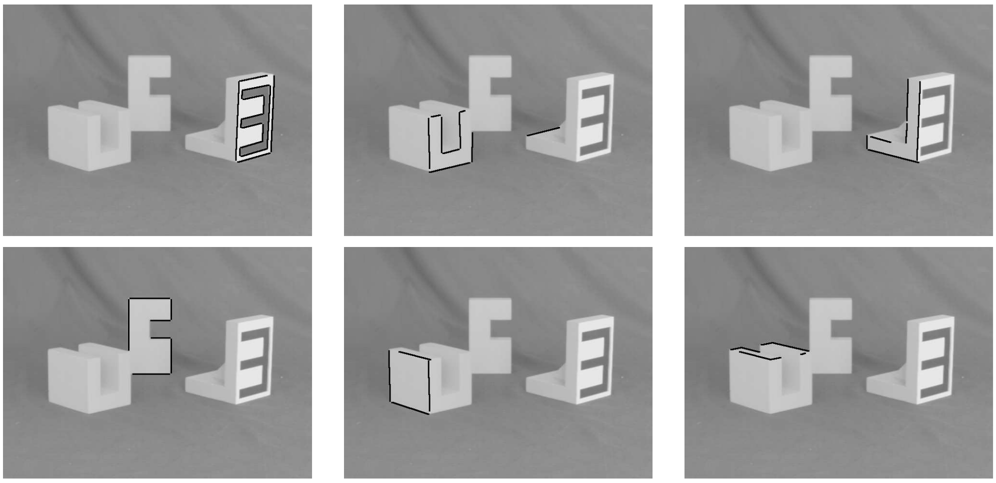


Figure 5.13: Planes recovered in the `test` scene

Figure 5.14: Planes recovered in the `lab` scene