(d) inferred surfaces and junctions, and the associated disparity assignments

(e) inferred surfaces, surface discontinuities and region boundaries

(f) a rectified, texture mapped view of the recovered scene description

Figure 5.7 (continued) The book scene
5.2.3 Unique disparity assignment

Using the computed inlier saliency values, we then impose the uniqueness constraint on the correspondence data by removing non-salient correspondences in a non-iterative manner. Among the correspondences, the least salient one is identified. This disparity assignment, together with the neighborhood information associated with it, is removed, unless it is the only one along the lines of sight in either image. We repeat this process until all the correspondence data are examined for uniqueness. Note that since local feature matching may fail to extract some of the correct correspondences, this uniqueness enforcement cannot remove all the wrong matches, as illustrated in figure 5.7(c) for the book scene.

5.2.4 Salient surface extraction

Once the more plausible correspondences are selected among the initial feature matches, we infer the underlying scene surfaces by spreading the computed saliency and continuity information throughout the entire disparity space: for each voxel in the disparity array, the density and distribution of the correspondence data in the neighborhood is collected and analyzed by applying tensor voting. Each location is then assigned a tensor which encodes its saliency and its orientation information as in the first phase. Locations with highest saliency in the local neighborhood are identified as being on a surface or on a surface discontinuity. The actual surfaces and discontinuity curves are extracted using a modified Marching process [51].

A dense saliency tensor field is obtained after all input data have cast their votes, from which salient surfaces, junction curves and junction points can be extracted by a non-maximal suppression process [77] modified from the marching process [51]. Figure 5.8 depicts a slice of the inferred surface saliency for the book example, in which
brighter indicate higher saliency. Note that although we use a specific surface model in
our estimation, the estimation errors due to model misfit are incorporated as orientation
uncertainties at all locations and are absorbed in the non-maximal suppression process.

After this phase, the surfaces are segmented, but their boundaries are not accurately
located. In fact, surfaces always overextend due to the absence of inhibitory process
during voting. As an example, figure 5.7(d) (left) shows the inferred salient surface and
junctions obtained from the disparity assignments depicted in figure 5.7(c). Also shown
(right) are the correct correspondence so obtained.

5.2.5 Region trimming
In this last phase, we rectify the surface overexpansion problem by imposing the
uniqueness constraint in the entire disparity space. To illustrate the symptom on the
book scene, we compute the product of the input images and their corresponding dis-
parity images. The results are shown in figure 5.9. Since overexpansion only happens
at occluding boundaries, and the associated occluded boundaries where scene surfaces
stop abruptly, we can identify problem areas by finding surface patches which project
onto the same pixel in one of the two images. Moreover, as each inferred overlapping
surface region is partitioned by the corresponding occluding boundaries into two areas
with correct matches all lying on one side, occluding boundaries and spurious regions
can be inferred by analyzing the distribution of correspondence data along the inferred

Figure 5.8 A cut of the inferred surface saliency
surfaces. However, as shown by Sara and Bajcsy in [71], intensity-based stereo matching often shifts the location of the occluding boundaries according to the contrast of the corresponding edge in the images. It is therefore necessary to incorporate monocular information from the images when inferring occluding boundaries.

We hence proceed to trim down the overexpanded surface as follows (with illustration on the book scene in Figure 5.10):

1. locate overlapping surface regions in the inferred surfaces.
2. Edge segments obtained in the preprocessing phase are backprojected onto each of the overlapping surface regions.
3. Using tensor voting, the distributions of the correspondence data along the surfaces are analyzed and are used to assign to each instance of edge segment a vector which encodes both its saliency as a occluding boundary and its direction of occlusion.