Chapter 6

Experimental Results

6.1 Introduction

In this chapter, we aim at experimentally evaluating the performance of the approach described in the preceding chapter, regarding computational efficiency, accuracy and robustness on a variety of target object configurations. The objects we considered for performing our experiments, are box-like objects very frequently encountered in super markets and distribution centers. We categorized the particular objects according to descenting degree of rigidity as follows: (**i**) sacks (bags), (**ii**) pillows, (**iii**) box-like objects wrapped in transparent foil, (**iv**) rigid card board boxes. The dimensions of the graspable surface of the objects range between 108mm × 111mm (small bags), and 173mm × 354mm (pillows). Non rigid objects like sacks are more difficult to be segmented than rigid objects (e.g. cardboard boxes), and the reason for this is that the former objects may slightly locally deform: Such deformations are not accounted for by our modeling entities. In addition, the edge maps originating from non rigid objects are considerably more noisy than the maps obtained from rigid objects.

For every object category, we manually created 10 configurations. Besides, we created 5 configurations comprising objects belonging to different categories. For each configuration a range image was acquired, and the objects were recovered by our segmentation approach in a Pentium 4 2.8 GHz PC, using the parameter values of table 5.2. A part of the obtained results are illustrated in figures 6.1 until 6.14. In each of these figures, the first row shows the intensity and range image of the configuration. The second row presents the results of the edge detection process: The left image of the second row, shows the detected edge points superimposed on the range image, and the right image the edge map. The third row presents the results of the seed placement process: The left image of the second row shows the three dimensional superquadric models corresponding to the seeds, and the right image the boundaries of these models embedded on the edge map. Finally, the last row presents the results of our recovery approach: The left image of the last row shows the recovered models in three dimensions superimposed on the range image, and the right image their boundaries embedded in the two dimensional edge map.

The presentation of these results is organized in sections. Section 6.2 presents results related to configurations of bags, section 6.3 shows results related to configurations of box-like pillows wrapped in a foil. Similarly, sections 6.4, 6.5 and 6.6 shows results extracted from configurations containing box-like objects wrapped in a transparent foil, card board boxes and configurations combining all objects respectively. Finally, in 6.7, a summary of the results presented in the previous sections is given. We have to note at this point, that in all cases expect the section dealing with card board boxes, the system parameters μ_1, μ_2 controlling the degree of integration of region and boundary information within the boundary finding process where set to $\mu_1 = \mu_2 = 0.5$. Especially for the rigid boxes we set $\mu_1 = 1, \mu_2 = 0$, because of the high quality of boundary information.

For all configurations related to each of the sections, we extracted measurements concerning the computational efficiency, robustness and accuracy, which are presented in table form in the beginning of each section. For assessing our system's robustness, we measured the number of *true-positive*, *false-positive* and *false-negative* responses of our system. A system's response is considered to be true-positive, when a recovered object actually exists in the image in the particular position, false-positive, when a recovered object does not exist in the image, and false negative when a graspable object exists in the image but is not detected by our system. We did not considered true-negative responses in our assessment, because in every case at least one graspable object exists in the scene. Apparently, the bigger the number of true-positive responses, and the lower the number of both false-positive and false-negative responses is, the more robust the system will be.

As discussed in the preceding chapter, and more specifically in section 5.3.2.2, the decisions regarding the existence of a graspable object in the image, are taken by the post processing module of our approach, that is the decisions taken by this module influence the robustness of our system to the greatest extend. In the context of this module, a recovered model is assumed to correspond to a graspable object, if the fitting error residual of the superquadric model to the range points belonging to the object of interest is small, and if the fitting error residual of the model's boundary to the edge map is small as well. Given that our objects can be satisfactory well described by superquadric models, the accuracy of boundary information plays the primary role in taking correct decisions. Hence, incomplete or missing boundary information substantially reduces the robustness of our system. The most typical case in which boundary information is incomplete, is when the objects are placed very tightly one after another in distinct layers. In this case, the resolution of the sensor is not high enough to capture the changes in depth, which will allow for boundary detection by our edge detection module. After extensive experimentation, we found that boundary information is complete in such cases, when the distance between the objects is not less than 1 cm. This was taken into consideration when generating object configurations for assessing the robustness of our system: When we created configurations with similar objects in the same layer (see figures 6.1, 6.6, 6.7, 6.11) we avoided placing objects too close to each other. Measurements on our system's robustness are shown in tables 6.2, 6.5, 6.8, 6.11. The first row of each table, shows the total number of graspable objects in all images of the corresponding section. The second, third and fourth row the number of true positive, false positive, and false negative system responses.

For assessing the accuracy of our system, we performed *manual* segmentation of the input range images, that is, for all true positive system responses, we isolated all regions in the input images corresponding to the exposed surfaces of the recovered objects. We then devised two measures for measuring accuracy: The first, employs depth information to measure similarity, and will be hereinafter denoted to as d_{3D}. This measure expresses the average euclidean distance of the recovered model to the points of the region corresponding to the particular model. The second, which will be hereinafter referred to as d_{2D}, expresses the similarity between the boundary of the recovered model, and the boundary of the manually segmented region corresponding to this model. In order to compute this measure, we firstly consider the boundary of the image coordinates of the manually segmented region of a graspable object. Then, we generate the euclidean distance transformed image of the boundary. Subsequently, we embed the boundary of the related recovered model to the transformed image. d_{2D} equals the average euclidean distance of the pixels on the boundary of the model to the pixels of the boundary of the manually segmented region. Apparently the lower the value of both d_{2D} and d_{3D} is, the more accurate the recovery is.

We additionally assessed the accuracy of our system by comparing the parameters of the recovered model to a *reference* model. Given a manually segmented region of points to which a recovered model corresponds, the reference model is acquired by *manually* fitting a box-like superquadric to the region. In particular, we compared the parameters related to the size (a_1,a_2) and the pose of the models ($p_x, p_y, \phi, \theta, \psi$). Note that as discussed in section 5.3.2.2, it is not possible to robustly determine the height of the model a_3, using one view of the configuration. This is the reason why this parameter is not considered in the comparison.

The assessment of the system's accuracy is presented in the tables 6.3, 6.6, 6.9, 6.12 of this chapter. More specifically, the first subtable (a) of each table, shows the averaged values for the distance measures $\bar{d}_{3D}$, and $\bar{d}_{2D}$ per graspable object. The second subtable (b), shows the average differences in the parameter values between the reference and the recovered models.

Finally, tables 6.1, 6.4, 6.7, 6.10, illustrate computational efficiency measurements. The first row of each table, shows the average time required for executing the hypothesis generation stage of our approach. The second row shows the average time required for the execution of the hypothesis refinement stage. The third row shows the average time required for the recovery of a single object. Note, that since in our system the recovery of a model parameters is performed separately for each seed, the object recovery, the average time for executing our system in a multi processor computer, where the recover of each seed is assigned to a separate processor is the some of the time needed to perform seed generation and the average time required for recovering a single seed.

6.2 Sacks (bags)

Modules	**Duration (sec)**
Hypothesis generation	3
Hypothesis refinement	125
Hypothesis refinement per model	20

Table 6.1: Computational efficiency for bag (sack) configurations

Measures	**Values**
Graspable objects	48
True positives	43
False positives	4
False negatives	5

Table 6.2: Robustness for bag (sack) configurations

Distance measures	d_{3D} (mm)	d_{2D} (pixels)
Values	9.1	1.3

(a) Average distances between recovered models and manually segmented regions

Parameters	a_1(mm)	a_2(mm)	p_x(mm)	p_y(mm)	p_z(mm)	ϕ(deg)	θ(deg)	ψ(deg)
Values	6.81	9.67	10.43	4.91	9.96	1.23	5.17	1.46

(b) Average difference in pose parameters between recovered and reference models

Table 6.3: Accuracy for bag (sack) configurations

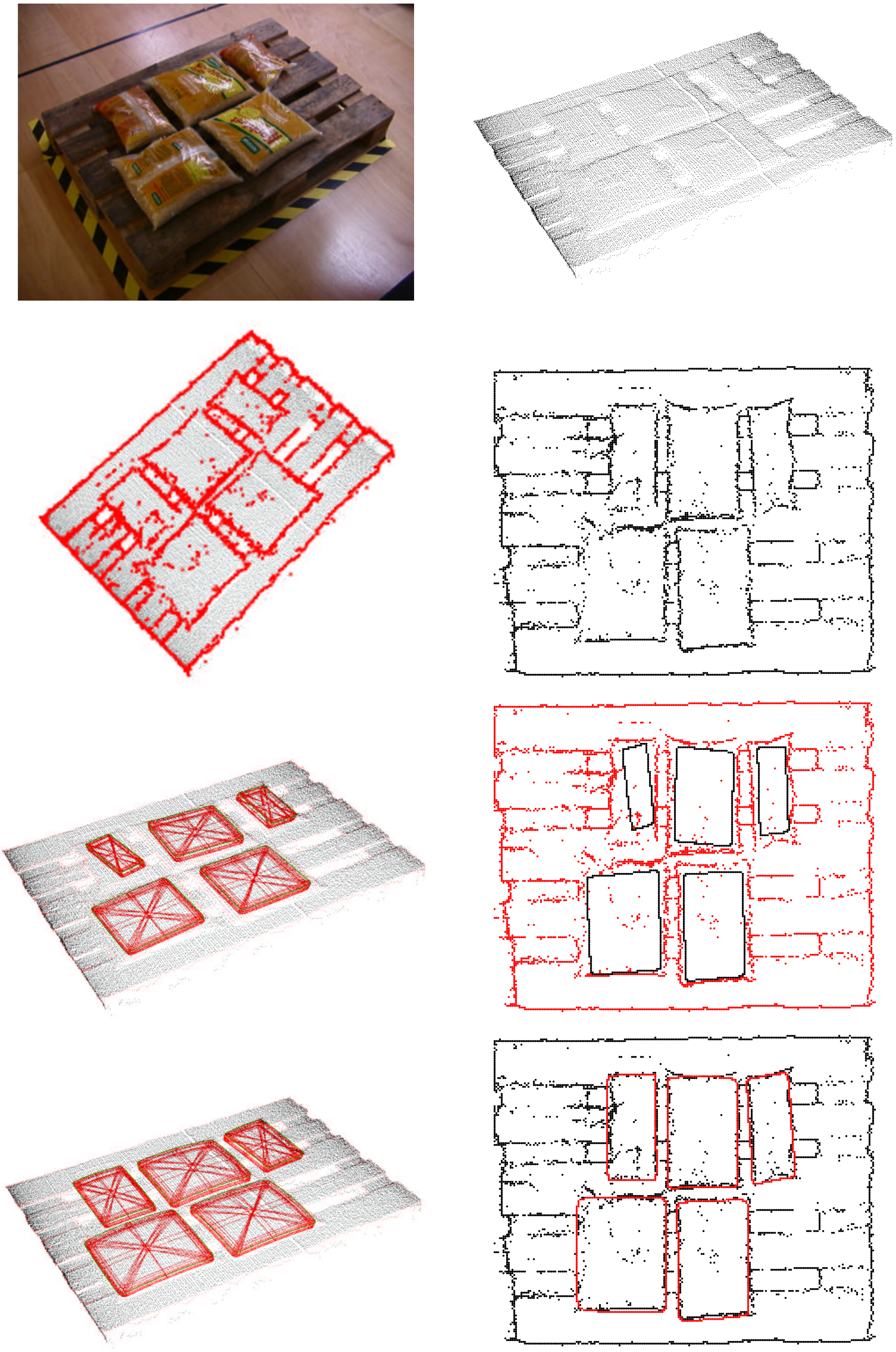

Figure 6.1: Sacks 1

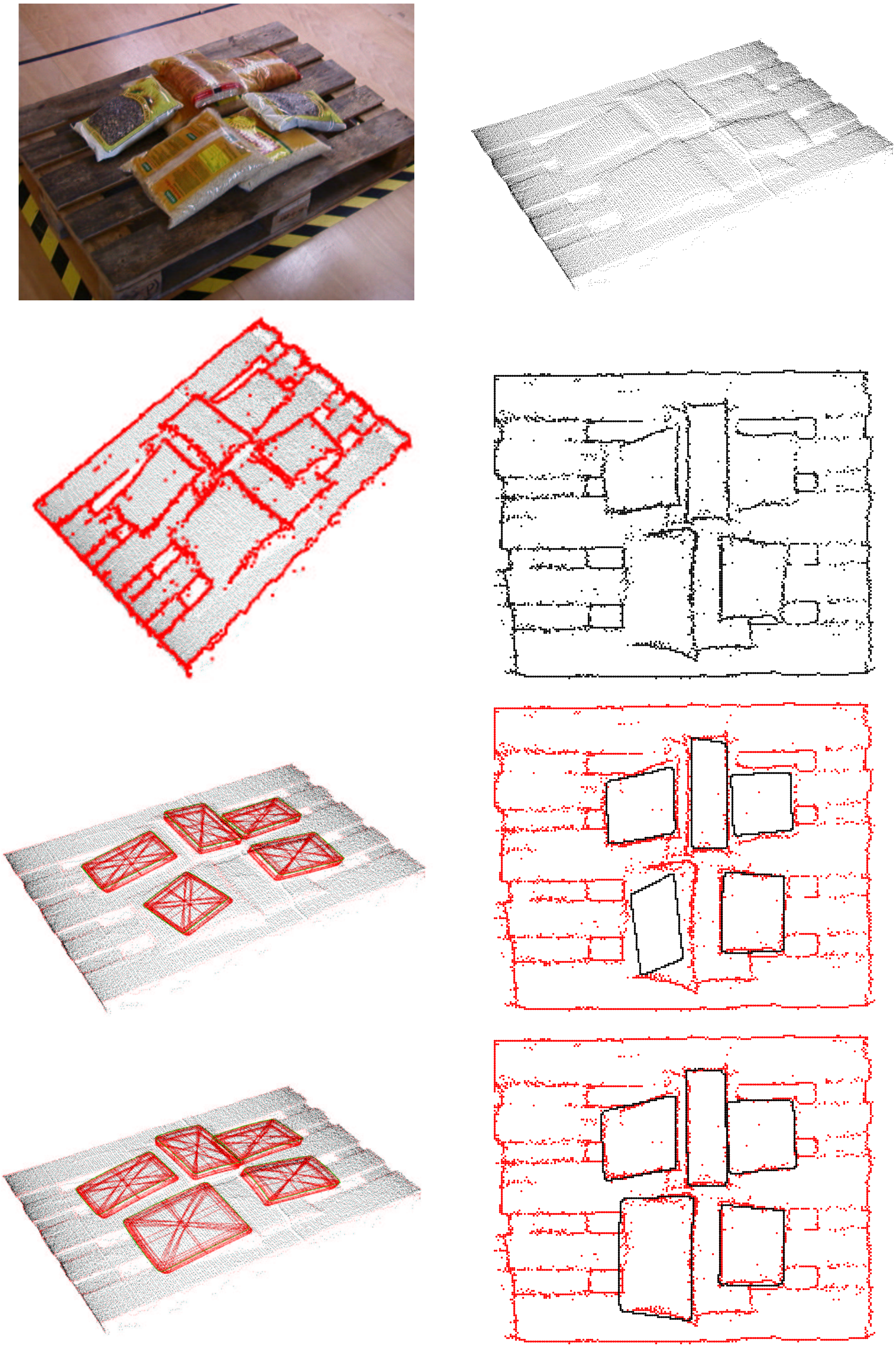

Figure 6.2: Sacks 2

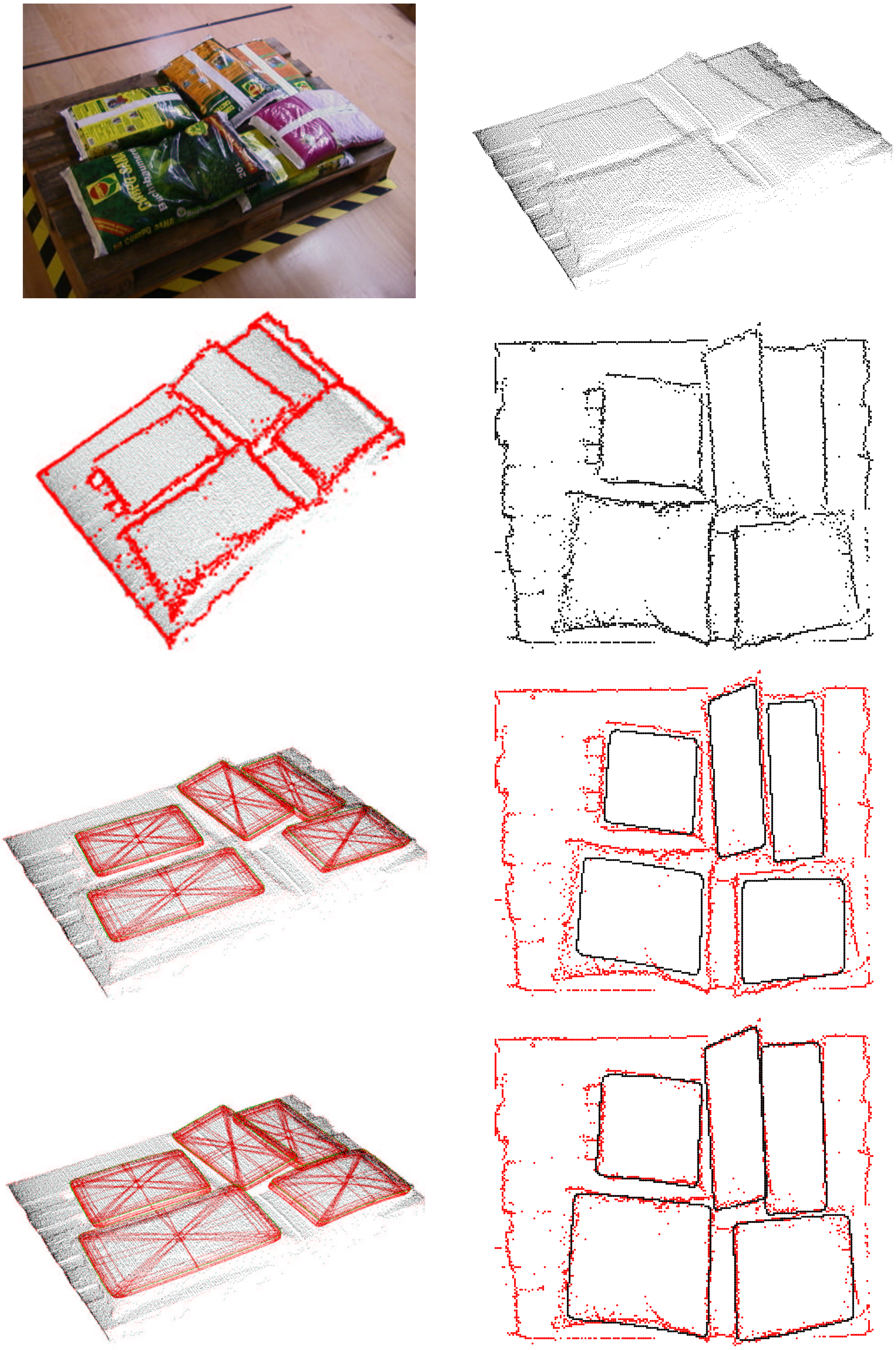

Figure 6.3: Sacks 3

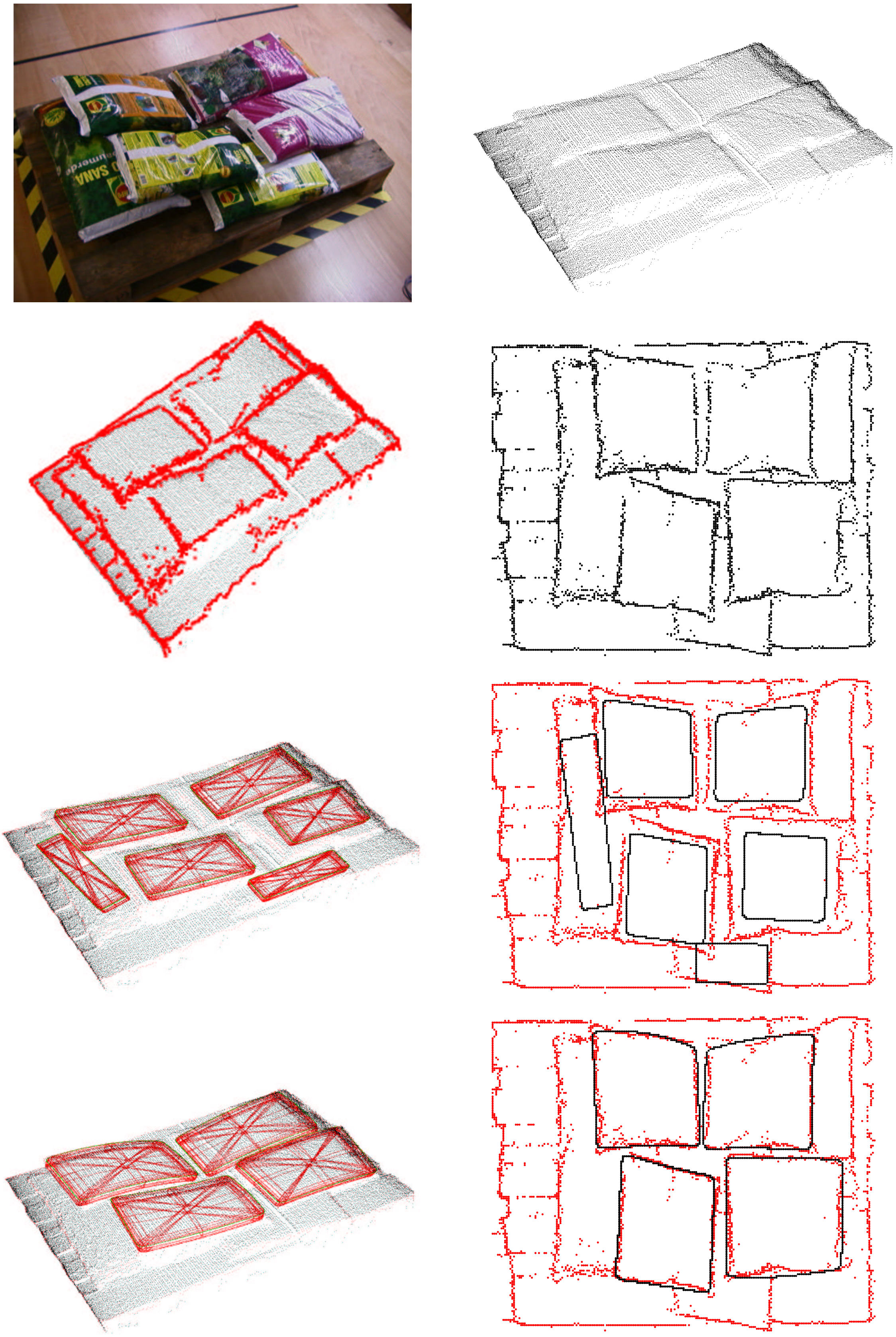

Figure 6.4: Sacks 4

6.3 Box-like pillows

Modules	**Duration (sec)**
Hypothesis generation	5
Hypothesis refinement	74
Hypothesis refinement per model	22

Table 6.4: Computational efficiency for pillow configurations

Measures	**Values**
Graspable objects	28
True positives	25
False positives	0
False negatives	3

Table 6.5: Robustness for pillow configurations

Distance measures	d_{3D} (mm)	d_{2D} (pixels)
Values	9.9	1.35

(a) Average distances between recovered models and manually segmented regions

Parameters	a_1(mm)	a_2(mm)	p_x(mm)	p_y(mm)	p_z(mm)	ϕ(deg)	θ(deg)	ψ(deg)
Values	6.37	2.45	9.70	7.89	8.14	2.15	2.85	0.64

(b) Average difference in pose parameters between recovered and reference models

Table 6.6: Accuracy for pillow configurations

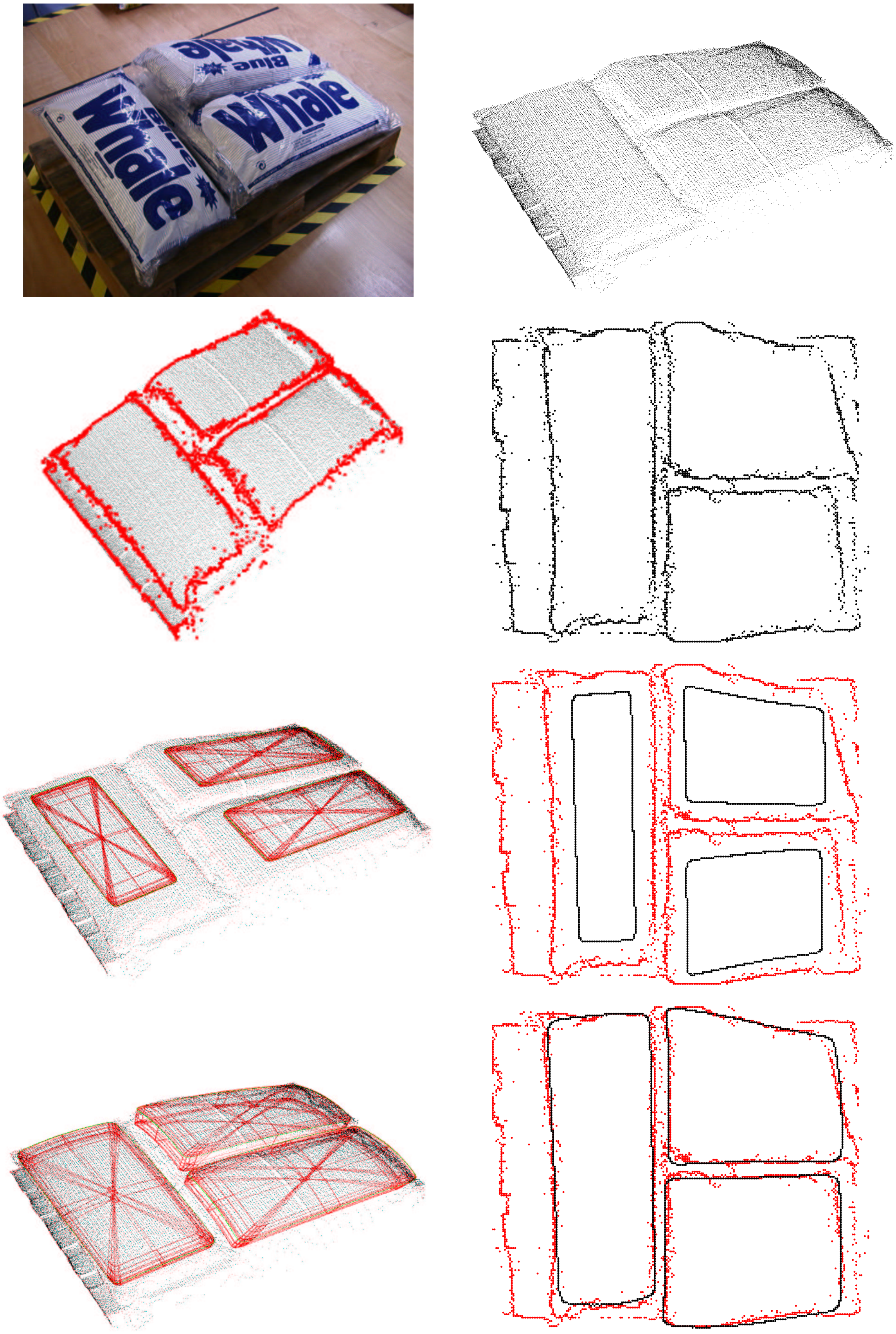

Figure 6.5: Pillows 1

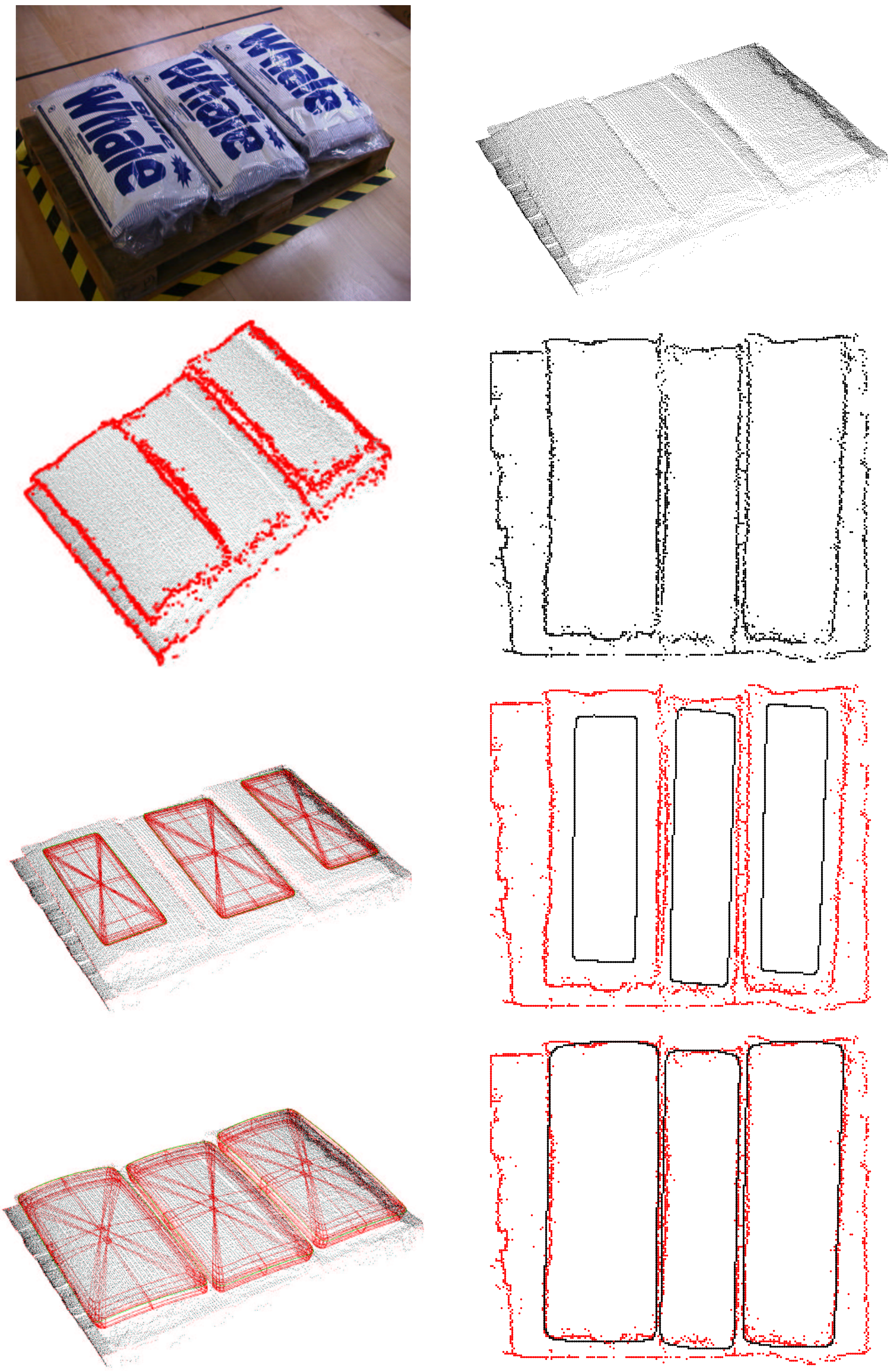

Figure 6.6: Pillows 2

6.4 Box-like objects

Modules	Duration (sec)
Hypothesis generation	2.5
Hypothesis refinement	118
Hypothesis refinement per model	15

Table 6.7: Computational efficiency for box-like object configurations

Measures	Values
Graspable objects	51
True positives	47
False positives	0
False negatives	4

Table 6.8: Robustness for box-like object configurations

Distance measures	d_{3D} (mm)	d_{2D} (pixels)
Values	8.40	0.90

(a)Average distances between recovered models and manually segmented regions

Parameters	a_1(mm)	a_2(mm)	p_x(mm)	p_y(mm)	p_z(mm)	ϕ(deg)	θ(deg)	ψ(deg)
Values	4.20	10.7	6.71	2.48	10.63	2.32	0.45	0.91

(b) Average difference in pose parameters between recovered and reference models

Table 6.9: Accuracy for box-like object configurations

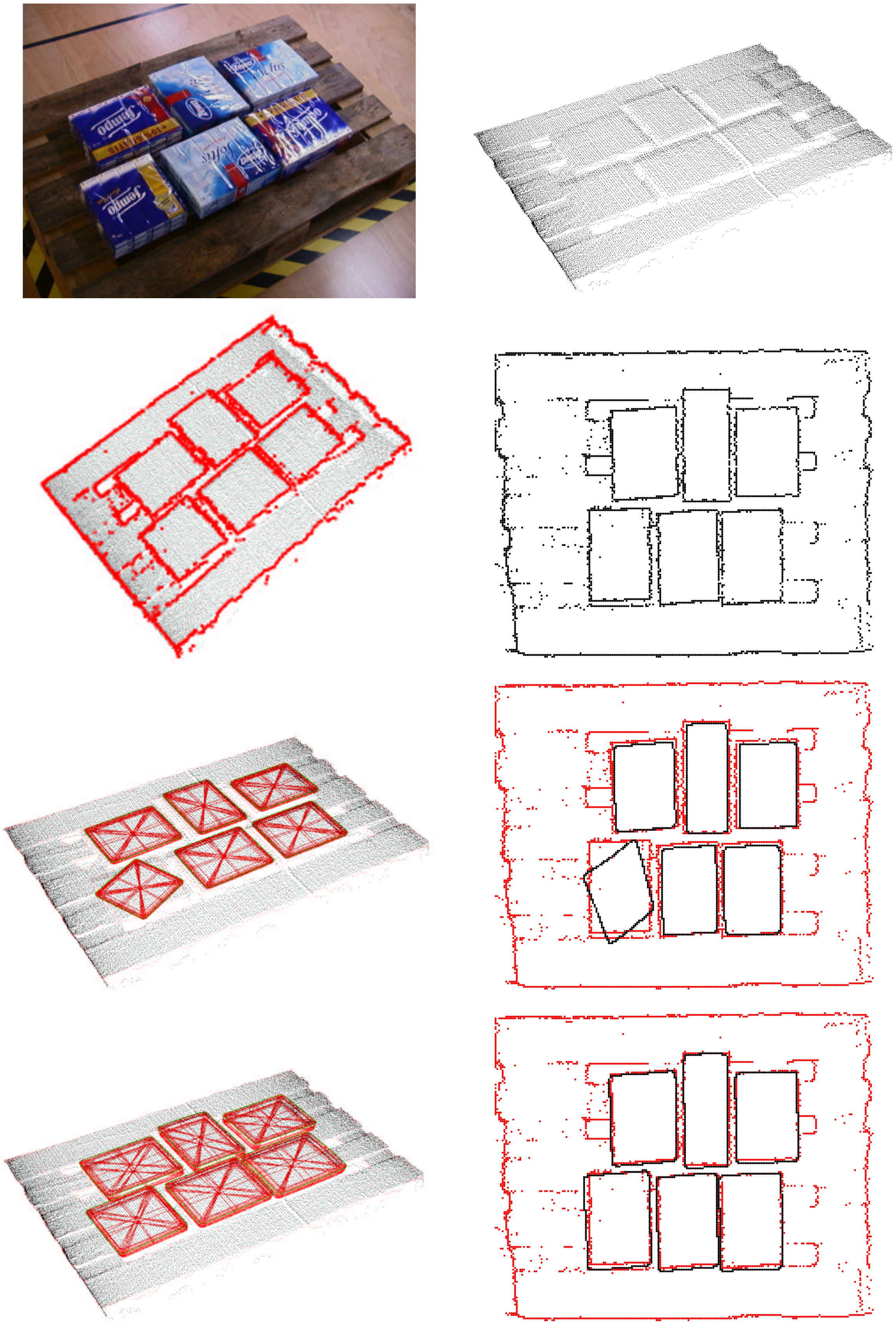

Figure 6.7: Box-like 1

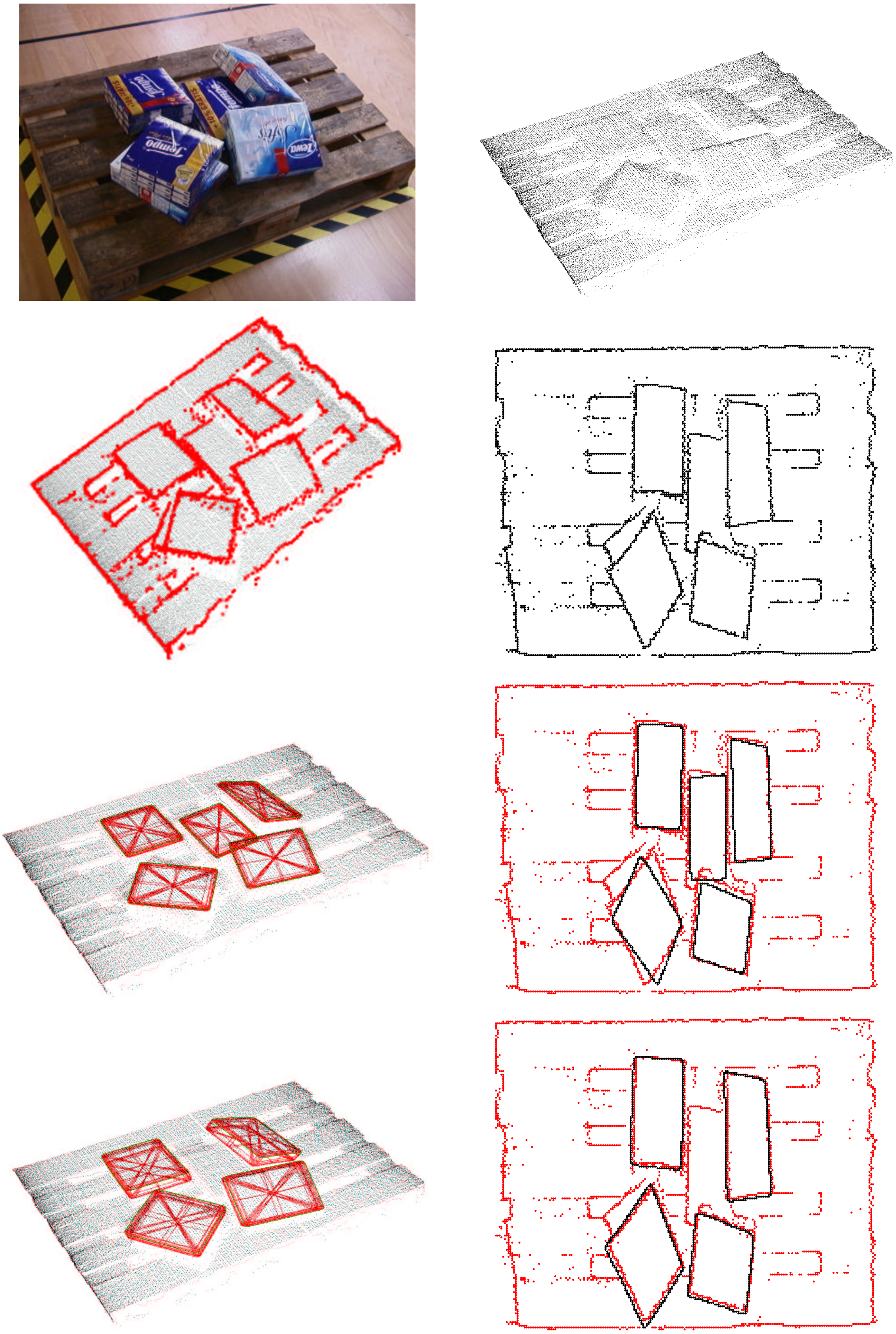

Figure 6.8: Box-like 2

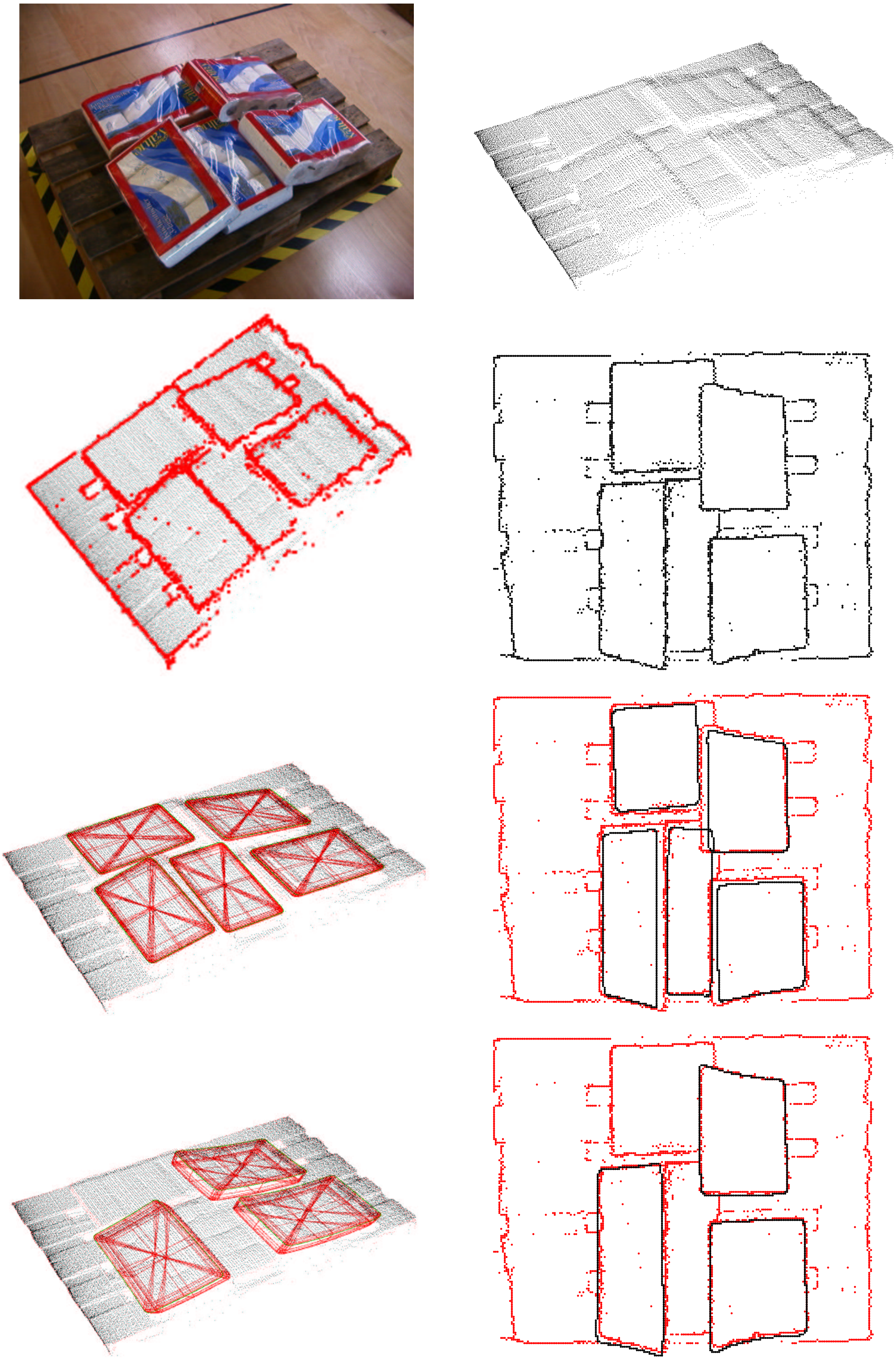

Figure 6.9: Box-like 3

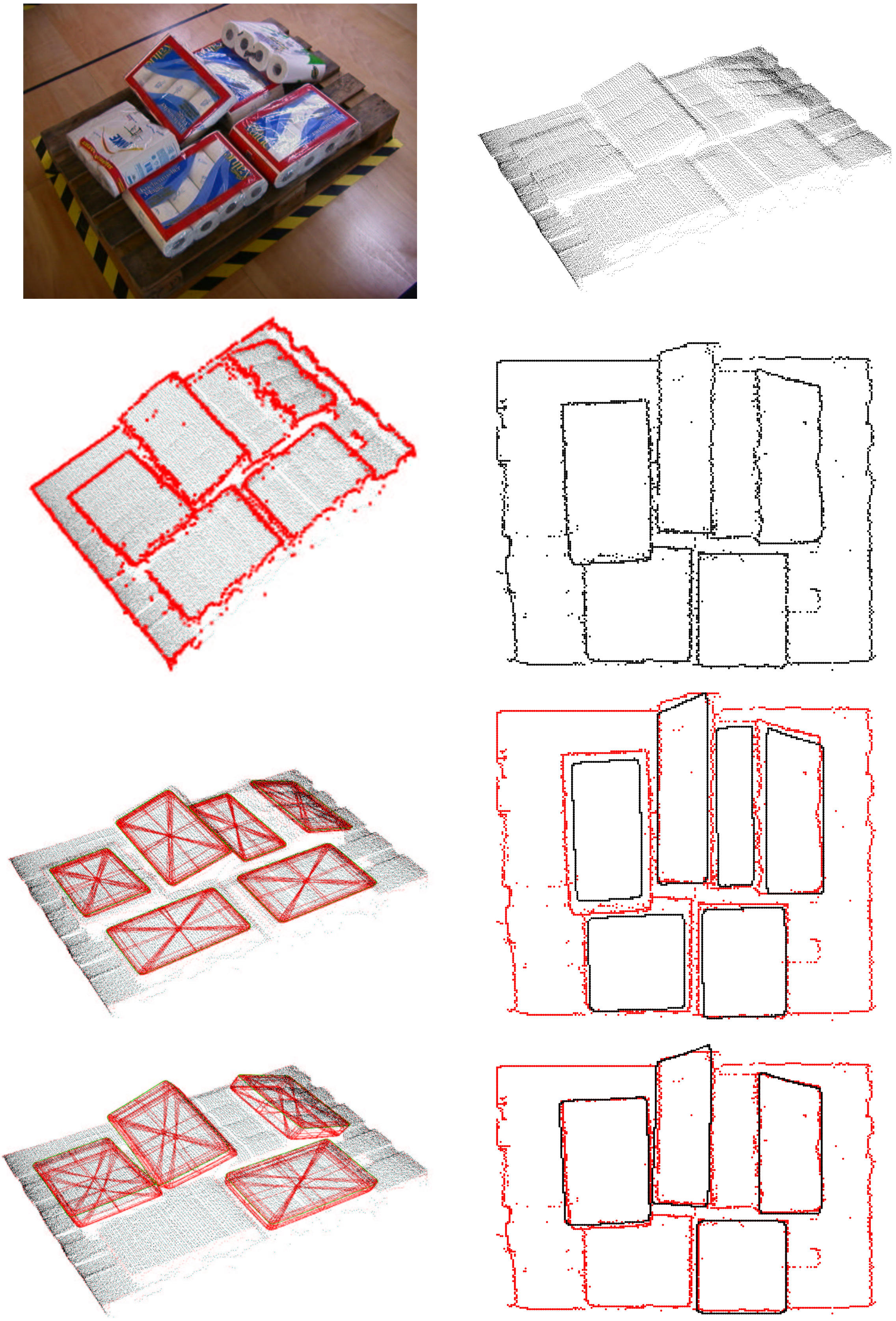

Figure 6.10: Box-like 4

6.5 Card board boxes

Modules	**Duration (sec)**
Hypothesis generation	2
Hypothesis refinement	50
Hypothesis refinement per model	5

Table 6.10: Computational efficiency for rigid box configurations

Measures	**Values**
Graspable objects	47
True positives	44
False positives	2
False negatives	3

Table 6.11: Robustness for rigid box configurations

Distance measures	d_{3D} (mm)	d_{2D} (pixels)
Values	7.3	0.5

(a) Average distances between recovered models and manually segmented regions

Parameters	a_1(mm)	a_2(mm)	p_x(mm)	p_y(mm)	p_z(mm)	ϕ(deg)	θ(deg)	ψ(deg)
Values	8.0	3.09	5.12	1.92	8.14	1.15	0.57	1.11

(b) Average difference in pose parameters between recovered and reference models

Table 6.12: Accuracy for rigid box configurations

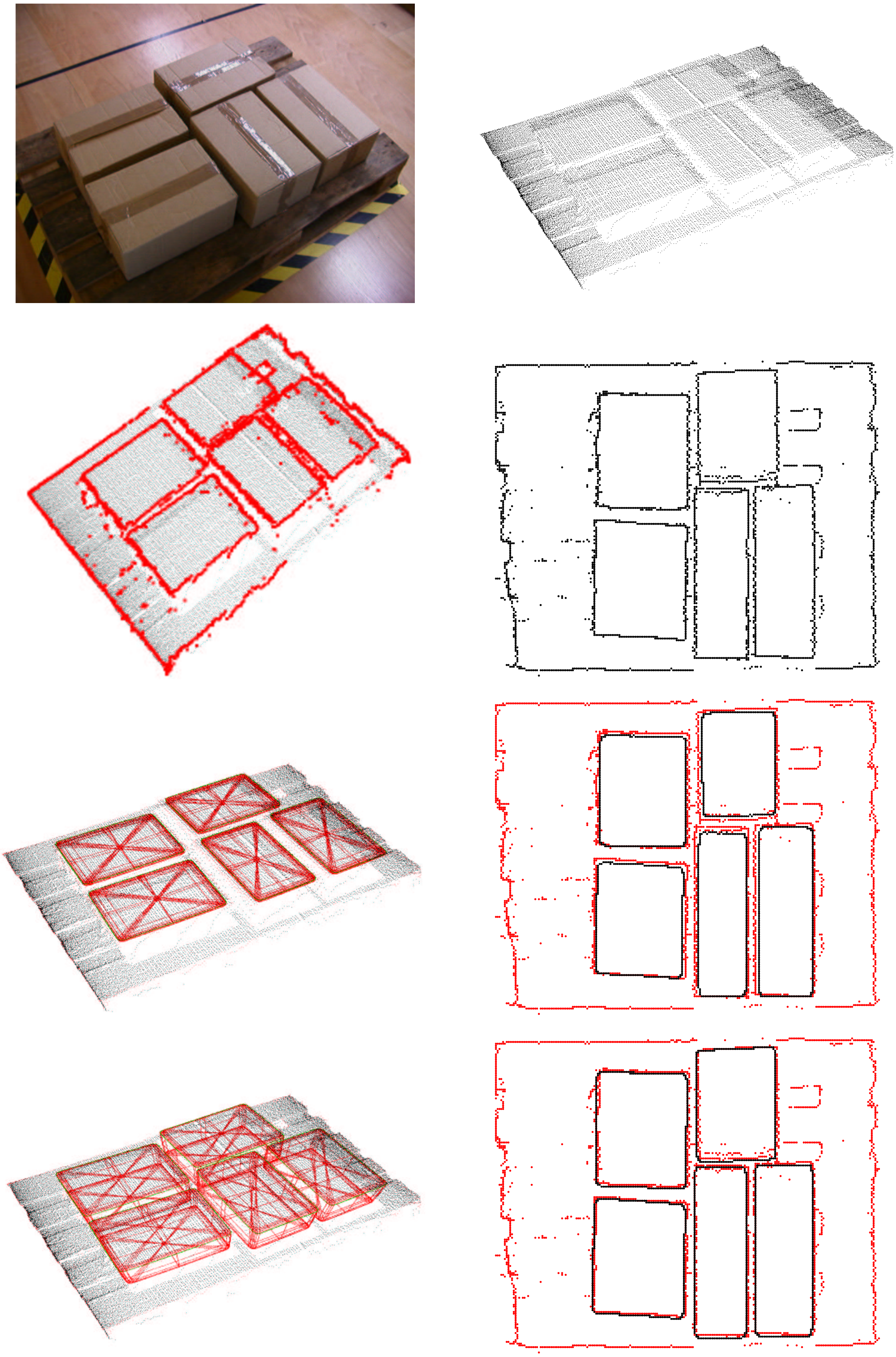

Figure 6.11: Boxes 1

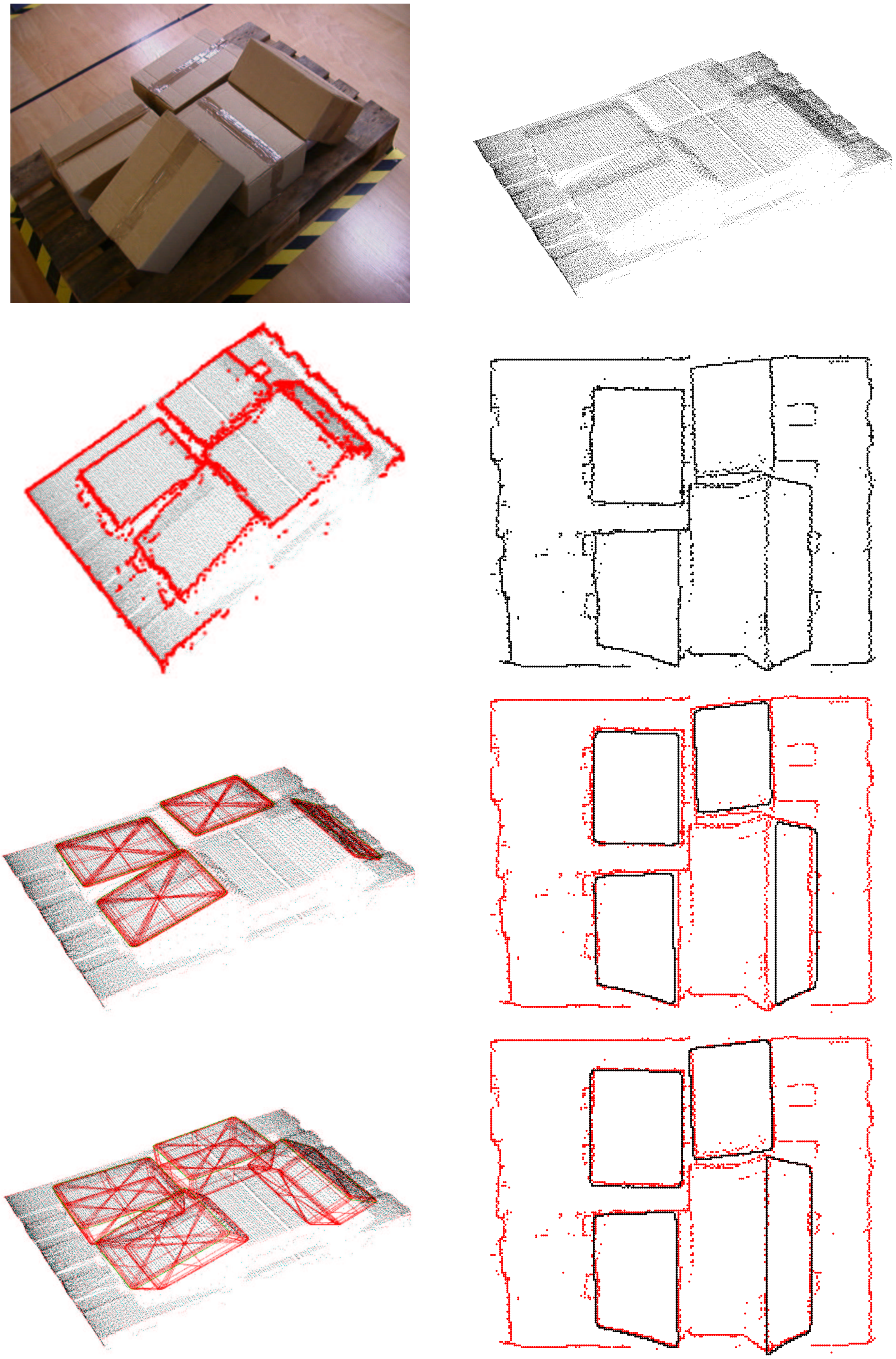

Figure 6.12: Boxes 2

6.6 Mixed

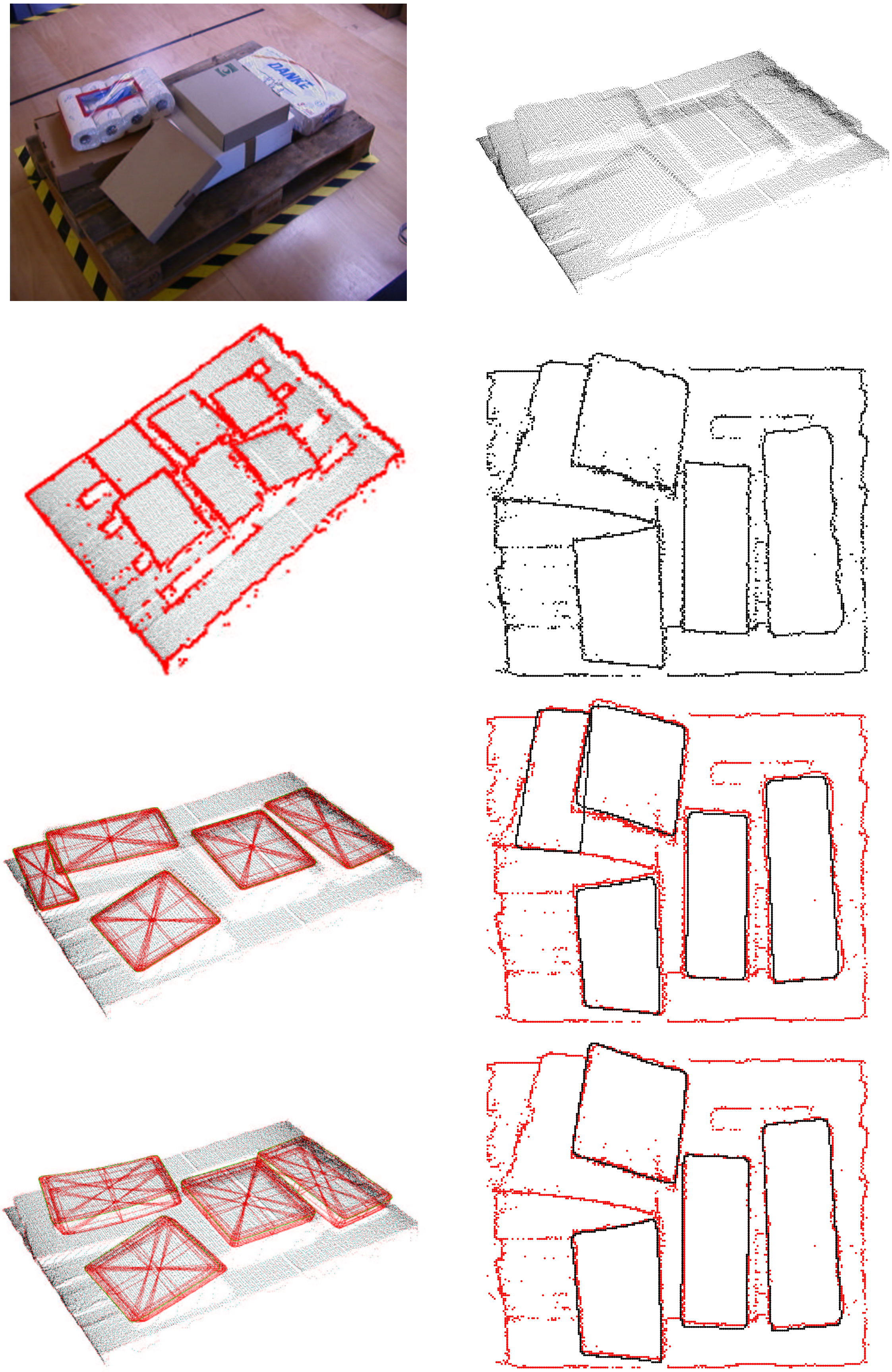

Figure 6.13: Mixed 1

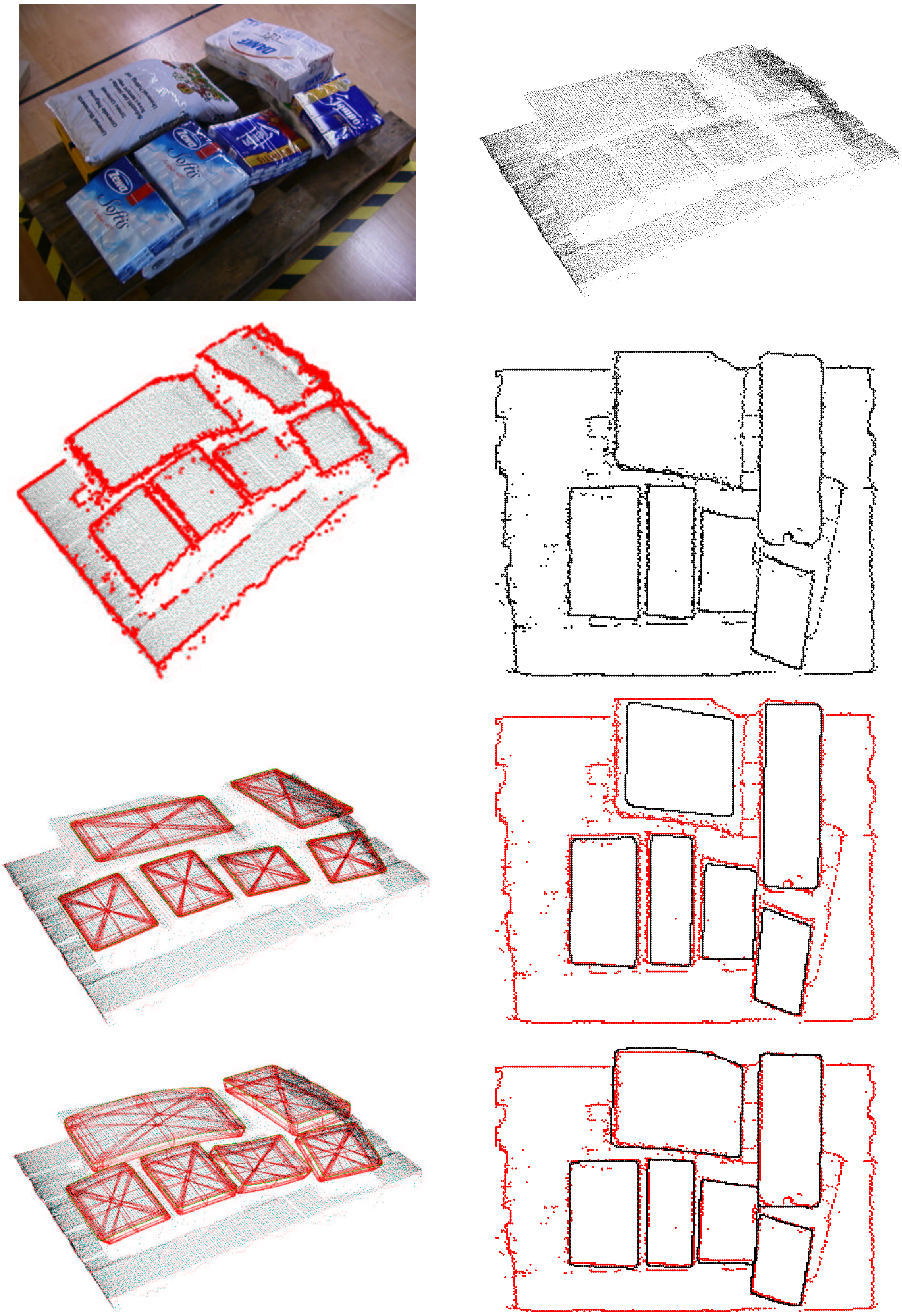

Figure 6.14: Mixed 2

6.7 Conclusions

From the results presented in the preceding paragraphs, we draw the conclusion that the performance of our approach depends on the rigidity of the objects examined. Actually, when the issue of computational efficiency is considered, our strategy is slower when dealing with non-rigid objects like sacks, pillows, or box-like objects, (for which the average processing time per image was 125, 74, and 118 seconds respectively and the average recovery time per graspable object 20, 22, and 15 seconds respectively) than when dealing with rigid boxes (with average processing time per image 50 seconds, and average recovery time per graspable object 5 seconds respectively). Concerning robustness, our system gives more false positive responses when dealing with sacks (4), than dealing with all other objects. Finally, concerning accuracy, both indeces d_{3D} and d_{2D} have higher values when non-rigid objects are dealt with: For sacks and pillows their values where 9.10 mm, 1.3 pixels and 9.90 mm, 1.35 pixels respectively, while for rigid boxes $7,30$mm, $0,50$pixels. As mentioned in chapter 5, as well as in the first paragraph of this chapter, the reason why this happens is that our globally deformable modeling entities cannot satisfactory describe the local deformations occurring to the surface of the non-rigid objects. Furthermore, local deformations of the non rigid objects is the cause of the generation of noisy edge maps. However, the reduction in the performance of our system when non rigid objects are addressed is slight, so that satisfactory handling of all kinds of box-like objects is enabled.

The quantitative measurements shown in the preceding paragraphs are summarized in the tables 6.13, 6.14, and 6.15. More specifically, table 6.13 shows averaged computational efficiency measurements for *all* the object configurations examined, 40 in total. The average time for image analysis is 91.5 seconds and the average time for the recovery of a graspable objects is 15.5 seconds. Table 6.14, demonstrates our system's robustness: In the 40 examined images, 174 graspable objects exist in total, 159 of which have been successfully recovered. In other words, the percentage of the true positive responses was more than 91% . Note, that the system gave false negative responses in only 6 cases. Finally, table 6.15 illustrates averaged segmentation accuracy measurements. More specifically, the index d_{3D} was 8.68mm per graspable object on the average, while d_{2D} was a little bit more than one pixel per graspable object. The average difference between the dimensions of the exposed surfaces of the reference and the recovered model per graspable object was 6.35mm for the x axis, and 6.48mm for the y axis of the model coordinate system respectively. These results clearly demonstrate that our strategy is a robust, computationally efficient, and accurate solution to the object recovery problem. The output of our approach on additional object configurations is exposed in appendix A.

Modules	**Duration (sec)**
Hypothesis generation	3.13
Hypothesis refinement	91.75
Hypothesis refinement per model	15.5

Table 6.13: Average computational efficiency

Measures	**Values**
Graspable objects	174
True positives	159
False positives	6
False negatives	15

Table 6.14: Robustness

Distance measures	d_{3D} (mm)	d_{2D} (pixels)
Values	8.68	1.01

(a) Average distances between recovered models and manually segmented regions

Parameters	a_1(mm)	a_2(mm)	p_x(mm)	p_y(mm)	p_z(mm)	ϕ(deg)	θ(deg)	ψ(deg)
Values	6.35	6.48	7.99	4.3	9.21	1.74	2.26	1.03

(b) Average difference in pose parameters between recovered and reference models

Table 6.15: Average accuracy