



HAL
open science

An efficient alternative approach for home furniture detection and localization by an autonomous mobile robot

Oscar Alonso-Ramirez, Yaser Aguas-Garcia, Antonio Marin-Hernandez,
Homero V. Rios-Figueroa, Michel Devy

► **To cite this version:**

Oscar Alonso-Ramirez, Yaser Aguas-Garcia, Antonio Marin-Hernandez, Homero V. Rios-Figueroa, Michel Devy. An efficient alternative approach for home furniture detection and localization by an autonomous mobile robot. 2015 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC 2015) , Nov 2015, Ixtapa, Mexico. 10.1109/ROPEC.2015.7395157 . hal-01355111

HAL Id: hal-01355111

<https://hal.science/hal-01355111>

Submitted on 22 Aug 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Efficient Alternative Approach for Home Furniture Detection and Localization by an Autonomous Mobile Robot

Oscar Alonso-Ramirez, Yaser Aguas-Garcia,
Antonio Marin-Hernandez, Homero V. Rios-Figueroa
Research Center on Artificial Intelligence
Universidad Veracruzana
Sebastián Camacho No. 5, 91000, Xalapa, Mexico
Email: oscalra_820 at hotmail.com,
yaser0.7 at gmail.com, [anmarin, hrios] at uv.mx

Michel Devy
LAAS-CNRS
Université de Toulouse
7 avenue du Colonel Roche, F-31077
Toulouse Cedex, France.
Email: devy at laas.fr

Abstract—In order for service robots to help humans in daily home tasks, they need to have a better understanding of their environment. Therefore, detection and localization of home furniture becomes a very important capability for them. This paper presents an approach to detect typical home furniture by an RGB-D sensor mounted on a mobile robot. The approach is based on the analysis of discriminant features extracted from very easy to compute measures distributions. Over an offline learning phase, each piece of furniture is modeled according to their distributions of: height, color (H and S components) and normals. Then a sequence of distributions analysis are applied to the scene for selecting, according to learned models, the pieces of furniture with a high probability of being present. The point cloud is segmented according to model analysis and then segmented points are projected to the floor plane for clustering and noise removal. According to previous analysis and footprints of segmented points clouds, regions are then classified as a possible piece of furniture. Finally, the orientation and localization of the detected furnitures are obtained, using the footprint and their neighbor regions. The proposed approach has been proved to be very efficient in order to be incorporated on mobile robotic platforms.

I. INTRODUCTION

Over the last few years, service robots research community have shown an increasing interest to develop methods and algorithms to bring closer humans and robots in daily tasks. For example, services to help elderly or disabled people at their homes are some of the more important goals. Yet, robots still struggle to perform efficiently their tasks on real, dynamic human environments. Moreover real scenarios scene understanding is still a challenging task.

Nowadays robots can use SLAM methods for mapping environments, or specific regions, and safely navigate in them. In case of indoor environments, robots can use fixed maps to add flexibility and to reduce computational cost, since it can be considered that there are many characteristics that doesn't change, i.e. walls, wardrobes or cupboards. However, problems may come when objects in the environment change position, are added or disappear from sight.

In this work, it is considered a robot that is localized on a map containing only fixed characteristics of the environment.



Fig. 1. The mobile platform UVerto and the apartment environment

This map can be provided by: a) a CAD model made by an architect or designer, or b) the robot has previously made a map of the empty environment, (i.e. just when moving to a new apartment).

Being able to understand the scene but primarily, detecting specific pieces of furniture and locating them in the environment helps the robot to improve decision making tasks, navigation and localization. On one hand, it adds details to the previous map, and it also allows the robot to extract cognitive information about a given region of the house; so the robot can discriminate between rooms (i.e. the bedroom has a bed, the dinning room has a table and chairs).

We use a mobile robot platform called UVerto equipped among others sensors, with a Kinect like RGB-D perception sensor mounted on a pan/tilt unit (Fig. 1).

The robot is localized by an AMCL implementation under a fixed map containing only no-removable features of the environment. A system for transformation between different reference frames is available on the robot. It is used a package from ROS called *TF*, which provide transformation between

different reference frames (one or more for each sensor or actuator). Since the robot is localized, sensor data can be transformed from the sensor frame to the base-link frame of the robot, and then to the world reference frame.

This paper is organized as follow; in next section (II), we present some of the most significant related works. In section III, we described the offline analysis and feature extraction process done for each piece of furniture, followed by section IV, where the proposed methodology to find each furniture in a scene is presented. In section V the results and evaluations of the proposed methodology are shown and finally in section VI, our conclusions and future work are described.

II. RELATED WORK

Many approaches have been exploited over the last few years for object detection. Most of these techniques have been focused on small objects, suitable for robot manipulation, however, not very efficient for furniture detection.

According to Drost et al. [1], the diverse approaches for recognition of free-form objects in 3D, can be divided in global and local approaches. On one hand, global techniques are good for detecting shapes such as planes, cylinders and spheres; which can be a good approach for detecting furniture. However, they often need a heavy segmentation process of the scene and typically are neither very precise nor fast. On the other hand, the local approaches proved to be very efficient, but they depend on local surface information so they are sensitive to occlusions and the quality and resolution of the data acquired. Taking into account the advantages and disadvantages of both approaches, in [1] it is proposed a hybrid technique which creates a global model description using an oriented point pair feature and matches that model locally using a voting scheme similar to the Generalized Hough Transform.

Salas-Moreno et al. [2] follow the proposed approach in [1] to detect objects (chairs and tables) and create an object oriented 3D SLAM with them. They report that their 3D object recognition approach lead to some advantages in terms of efficient and semantic scene description for real time SLAM.

However, if this approach is used to detect other pieces of home furniture, its accuracy decreases, mainly because their characteristics could be very similar and the oriented points couples would be similar (i.e. the dining table, the small table and the bed, have very similar oriented point couples).

In [3], Günther et al. present an approach to create a semantic map based on the recognition and localization of furniture objects. They use a model-based method for the recognition of the objects, where they first detect and classify planar regions and later they match the object with CAD model. The classification of the surfaces is done by a Web Ontology Language (OWL-DL) which contain the features and the rules that encode the structural model of the objects.

Wünstel et al. [4] presented an object recognition system based on the object's prototypical geometry. They model the objects as a directed acyclic graph, where each node correspond to different segments of the objects at different heights. Each segment is characterized by height, minimum

and maximum size and the relationships between them correspond to weighted edges. They were able to successfully identify chairs and tables within an office environment.

Liu et al. [5] emphasizes the importance of an advance environment model as a requirement for an intelligent robot to perform its daily tasks. So, they propose the implementation of a semantic map using a parametric model to approximate the geometry of the perceived environment, with 3D object localization. Kasper et al. [6] deduce semantical information from the detection and localization of objects, based on their spatial relationships between them.

In [7] a method is proposed for creating a household environment map with geometric models for environment manipulation. The object detection is performed by comparing clusters from the point cloud, with the database of objects models, using a version of ICP. The creation of the environment was successful, but all process were working asynchronously which makes it non-viable for real time applications.

In [8], Rusu et al. propose an algorithm for building environment object maps for a kitchen environment. The objects modeled include cupboards, tables and drawers, all of them relevant for manipulation tasks. This work was later expanded in [9] to construct a hybrid map that contains the objects and uses semantic information to add triangular meshes used for collision avoidance.

In [10] was presented an algorithm to identify and detect home furniture based on color, 3D position and 3D normals. Analyzing a small set of random points with a set of artificial neural networks, the method was able to avoid the processing of the complete frame, making the process run in real-time. This paper presents an alternative methodology based on fast an easy to compute measures distributions, which allows an efficient furniture detection and localization.

III. FEATURES EXTRACTION AND MODEL GENERATION

To distinguish one kind of furniture from another, it is necessary to compare relevant characteristics which allows to discriminate between them. However, looking to avoid time consuming and/or computationally inefficient processes, we propose to model each piece of furniture using easy to extract measurements, computed at each 3D point (similar to a local approach). Then, we model measures distribution for each piece of furniture, by extracting significant shapes over these distributions (similar to a global approach).

At this stage, we are interested only in significant peaks on a given distribution, which will be modeled by Gaussian curves (equation 1) individually for each furniture.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (1)$$

There are three properties that we measure and analyze: the height, the color components HS (from HSV color space) and the normals. Although each of these features may not be totally discriminant by themselves, together they increase the possibility for a good classification.

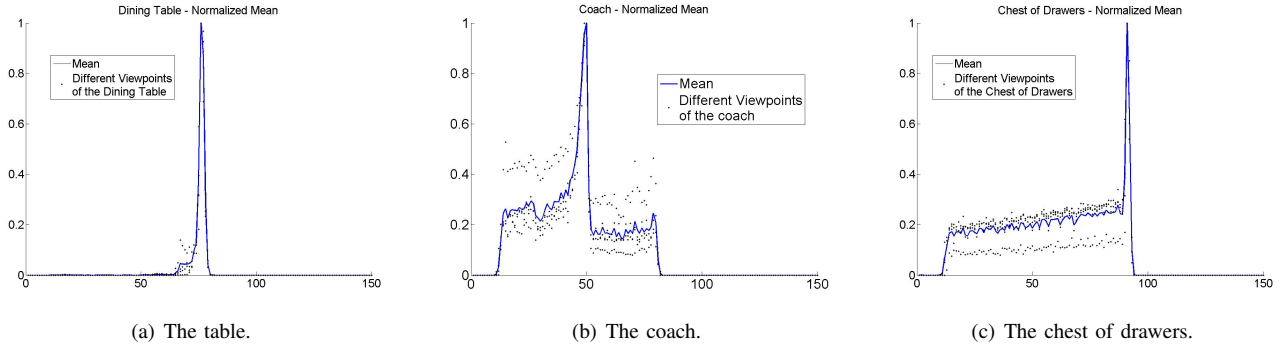


Fig. 2. Height distribution (histograms) of some furniture pieces

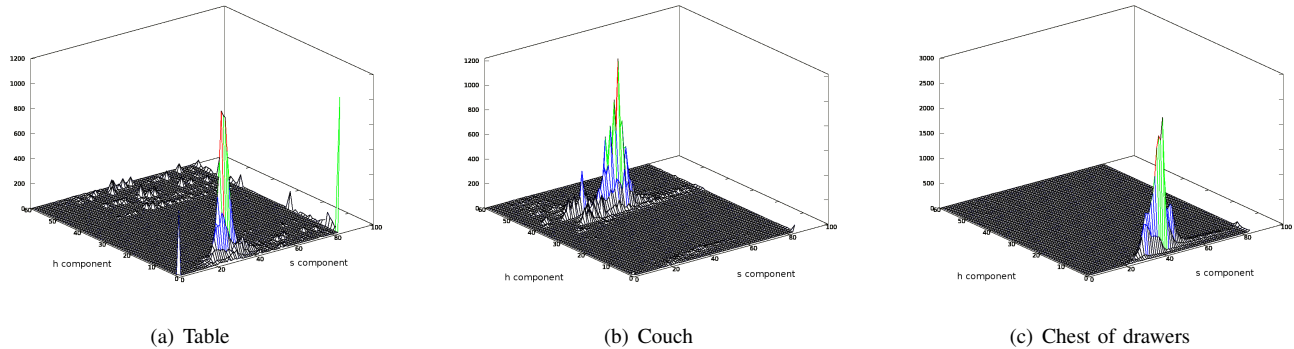


Fig. 3. HS color distribution of some furniture pieces

In order to compute accurate measurements, we have acquired different point clouds at different angles (rotation along z axis, perpendicular to the floor) for each isolated object. We maximize the size of objects inside the field of view of the RGB-D camera, to have an upper limit, in the number of points of the object in the image, and to model them, under these conditions.

A. Height Distribution

The shapes of most of the furniture can be approximated by simple geometric forms (e.g. boxes, cylinders, etc.). However, the horizontal planes are a common characteristic for most of these objects; depending on the kind of furniture, these planes serve to place smaller items or to sit or lie on them.

These horizontal planes are validated as a discriminating feature by the proposed distributions of heights measurements. In figure 2, we show the height distributions for three pieces of furniture found in our environment. The height distribution is obtained as an histogram of the heights of all points in the object, according to the floor reference frame.

We use the TF package from ROS to transform the point cloud from the camera-view frame to the robot's base frame, eliminating any configuration of the pan/tilt unit (PTU), so that all the points belonging to an horizontal plane will have the same height.

On figure 2, curves described only by points correspond to the measures at different points of view (angles) of the

object, and continuous lines are the average of these curve distributions.

The peaks on height distributions correspond to points over the main horizontal plane. While other shapes correspond to other characteristics (e.g. the vertical planes for the chest of drawers), at this stage we only model the highest peak (as a feature) on this distribution.

B. Color Distribution

Color is a popular characteristic for detecting or tracking an object, however it can be a very noisy characteristic. The color of an object, as it is perceived by a camera, can easily change with even the smallest changes in the illumination conditions.

For the color distribution, a two-dimensional histogram has been built (figure 3). In order to have a good color representation we have choose to use the HSV color space, without considering the value (V) component. We approximate the higher areas of the histogram with a 2D Gaussian distribution and we use the projected ellipses to define the characteristic regions for the color of each piece of furniture.

C. Normals Distribution

Normals can be very useful to detect planes, but they can also be computationally very expensive. In this work we calculate the normals to obtain the orientation of a point (and the points next to it) with respect to the z axis. With those angles we can determinate if a point belongs to an horizontal or a vertical plane (Fig. 4).

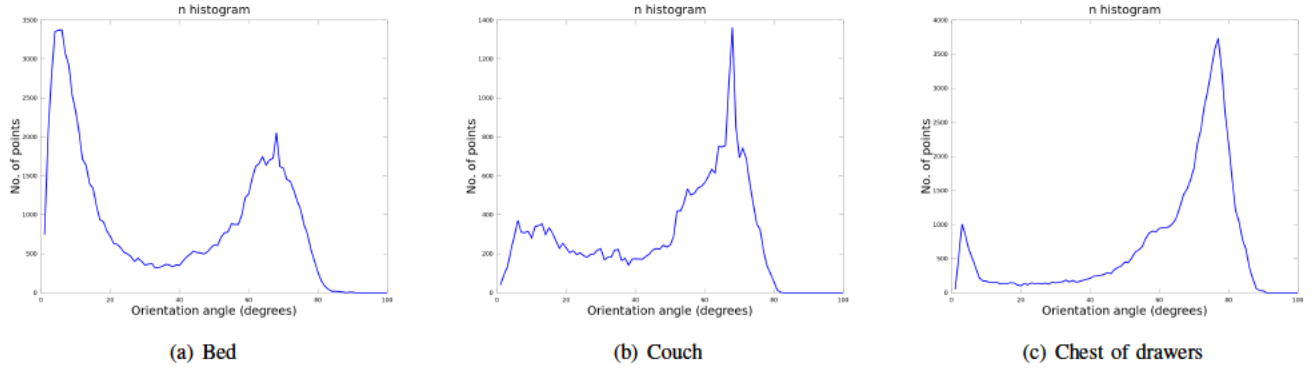


Fig. 4. Normals distribution of some furniture pieces

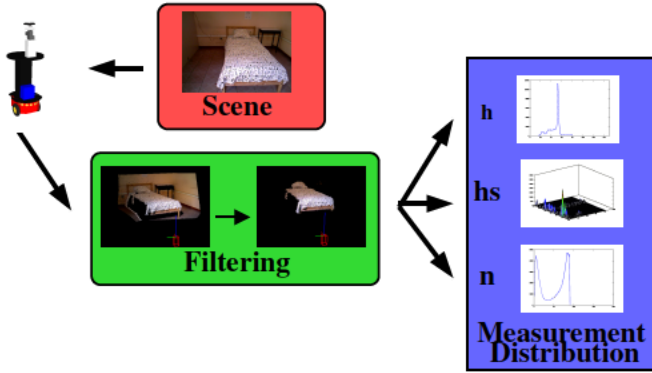


Fig. 5. Elements for model construction

D. Furniture modeling from measurement distribution

The scheme in figure 5 shows the process and elements required for the model construction of each piece of furniture. At this stage each model is created by hand based on the observed significant features on the measurement distributions of each furniture. For example, in the case of the bed, the model doesn't include color because it can drastically change from one quilt to another. In case of the couch, the model includes color and the orientation of the points (normals) above and below the main horizontal plane. Due to lack of space, the details of this models have not been included.

IV. FURNITURE RECOGNITION AND LOCALIZATION

As previously described, it is assumed that the robot has a map of the empty environment and it can localize itself on it. So, it is possible to remove, from a scene point cloud, the fixed features on the environment, i.e. floor, walls and/or fixed furnitures. A scene acquisition is composed from points belonging to furnitures, and all other objects present at this moment on the scene. While the robot navigate on the environment it can incrementally incorporate furniture models to their map, providing a better localization.

The proposed scheme for the furniture detection is shown in figure 6. The first step is to calculate the measure distributions for the scene, in a similar way as the model construction.

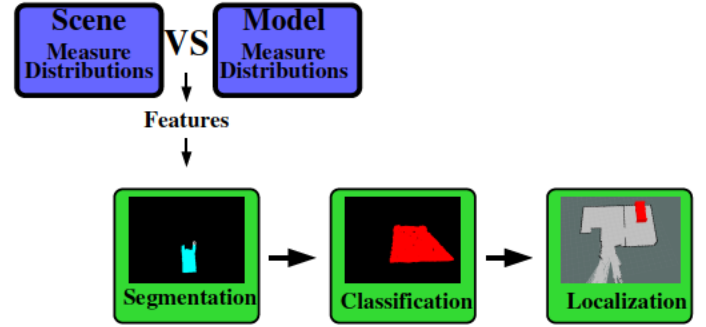


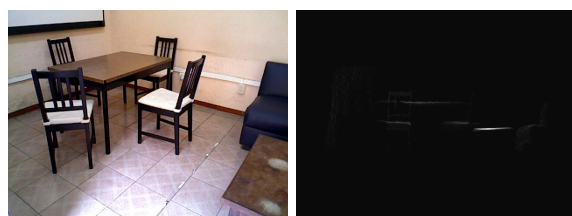
Fig. 6. Proposed scheme

Beginning with the height distribution, we select as features the salient peaks. If we were facing a single piece of furniture, we could compare directly the curve of the histogram obtained from the scene with the curves obtained in the learning phase for each object. However in the environment, very often we are going to find more than one piece of furniture together and even some other objects. In these cases, the height distribution will correspond to the accumulation of different measurements from each object present on the scene.

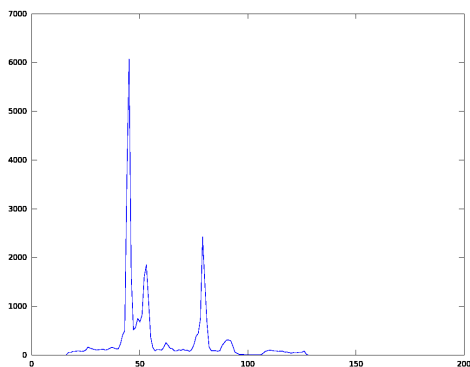
The problem here is to segment correctly the points corresponding to different objects according to the height distribution. To solve this, we search for the higher peaks on the histogram (those who are above two times the mean) and we tag them depending on their closeness to the peaks on the models of each furniture.

The next step is to project the points belonging to each peak onto the floor plane. These points correspond to the main horizontal planes of each objects. For example, in the the figure 7(a) is shown the image of a scene; in figure 7(b) are shown the points projected onto an imaginary vertical plane to visualize the correspondence of the horizontal planes obtained from the height distribution shown in figure 7(c).

A clustering algorithm is applied to the floor projected points, to separate different objects. Each cluster is now a possible piece of furniture. The color and the normals information of the clusters is then analyzed to verify if they lie within the measure distribution of the furniture models indicated by the tag assigned by the height analysis. A search for new points



(a) Scene (b) height projection



(c) Scene height histogram

Fig. 7. Horizontal planes found on a scene

that could be part of the same object is performed.

Depending on the model created for each piece of furniture, we could change the search area. For example, if the horizontal plane we are analyzing is tagged as possible couch, we will search for points above the plane to find the back of the couch. If we are looking for a table, we will search for the table legs, below the plane. The color and orientation of these new points must be within the measurement distributions modeled for each piece of furniture; i.e. the points above a couch plane must have the same color and an orientation almost perpendicular to the horizontal plane.

A final validation is performed by measuring the area of the projected points from each cluster. If this measure is less than a percentage of the area of the corresponding model, it is considered that the information is not enough to validate the model. Using the information from the adjusted model and its projection to the floor plane, the position on the map is obtained.

V. RESULTS AND EVALUATION

The experimental environment consists of two pieces of apartment: a bedroom and a dining/living room (sometimes referred to as a T2 apartment). The bedroom contains a bed and a chest of drawers, and the dining/living room contains a coffee table, a dining table and chairs. For the experiments, we had the robot navigating through both rooms.

The figure 8 shows some results of one of the different sequences done by the robot. Each row corresponds to a different time of the sequence, from t_0 for the upper row, t_i for the rows in between to t_f for the last row. The first column shows the RGB images obtained. The second column shows

the main horizontal planes detected in each scene. And in the last column are placed the detected pieces of furniture on their estimated position on the map.

As it can be seen in the images of the sequence shown in the figure 8, at t_0 the robot detects a bed, places it on the map and continues its navigation. At time t_i , the robot continues its navigation, detecting and adding models to the map. The process continues incrementally until time t_f , where the last pieces of furniture are added and the map is completed. According to the obtained results it has been observed that the variations on the positions for a piece of furniture have a variation of less than 5cm, except for one piece who has a variation of 10cm. And the variations on the orientation for all the cases were inferior to 5%.

VI. CONCLUSIONS AND FUTURE WORK

The proposed methodology has proved to be fast and efficient, so it can be considered as a good alternative to the other approaches. The errors obtained are within an acceptable range, considering a map constructed under a 5cm occupancy grid. This approach is now running on our mobile robot platform. So the robot is able to detect furniture pieces and place them correctly on the map.

Processing time is very low, under .5s in the worst case, to locate and place all the furniture given a complex scene, so the robot can do diverse tasks while processing point clouds. As future work is proposed to increase the number and the types of furniture, as the size of the environment. Also it can be considered to add other measurements that could be discriminant characteristics.

REFERENCES

- [1] B. Drost, M. Ulrich, N. Navab, and S. Ilic, "Model globally, match locally: Efficient and robust 3d object recognition," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, June 2010, pp. 998–1005.
- [2] R. Salas-Moreno, R. Newcombe, H. Strasdat, P. Kelly, and A. Davison, "Slam++: Simultaneous localisation and mapping at the level of objects," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, June 2013, pp. 1352–1359.
- [3] M. Günther, T. Wiemann, S. Albrecht, and J. Hertzberg, "Model-based furniture recognition for building semantic object maps," *Artificial Intelligence*, no. 0, pp. –, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S000437021400157X>
- [4] M. Wunzel and R. Moratz, "Automatic object recognition within an office environment," in *Computer and Robot Vision, 2004. Proceedings. First Canadian Conference on*, May 2004, pp. 104–109.
- [5] Z. Liu, W. Wang, D. Chen, and G. von Wichert, "A coherent semantic mapping system based on parametric environment abstraction and 3d object localization," in *Mobile Robots (ECMR), 2013 European Conference on*, Sept 2013, pp. 234–239.
- [6] A. Kasper, R. Jakel, and R. Dillmann, "Using spatial relations of objects in real world scenes for scene structuring and scene understanding," in *Advanced Robotics (ICAR), 2011 15th International Conference on*, June 2011, pp. 421–426.
- [7] Y. Kakiuchi, R. Ueda, K. Okada, and M. Inaba, "Creating household environment map for environment manipulation using color range sensors on environment and robot," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, May 2011, pp. 305–310.
- [8] R. B. Rusu, N. Blodow, Z. Marton, A. Soos, and M. Beetz, "Towards 3d object maps for autonomous household robots," in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*. IEEE, 2007, pp. 3191–3198.

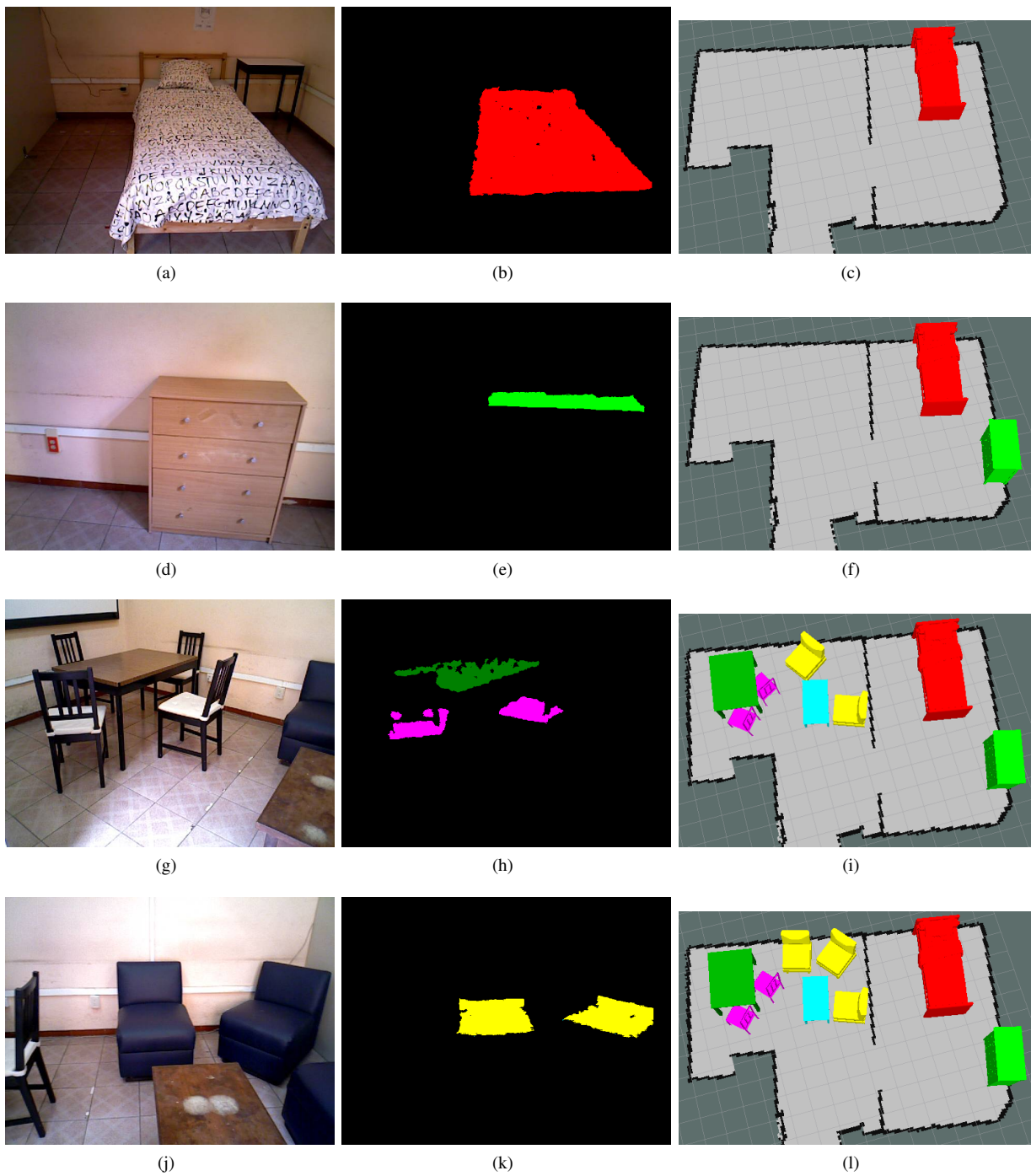


Fig. 8. Results from a sequence acquired by the robot

- [9] R. B. Rusu, Z. C. Marton, N. Blodow, A. Holzbach, and M. Beetz, "Model-based and learned semantic object labeling in 3d point cloud maps of kitchen environments," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, 2009, pp. 3601–3608.
- [10] O. Alonso-Ramirez, A. Marin-Hernandez, M. Devy, and F. Montes-Gonzalez, "Indoor home furniture detection with rgb-d data for service robots," in *Electronics, Communications and Computers (CONIELECOMP), 2014 International Conference on*, Feb 2014, pp. 172–177.