# Home Furniture Detection by Geometric Characterization by Autonomous Service Robots

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

# Home Furniture Detection by Geometric Characterization by Autonomous Service Robots

Oscar Alonso-Ramirez[1], Antonio Marin-Hernandez[1], Homero V. Rios-Figueroa[1] and Michel Devy[2]
[1]*Artificial Intelligence Research Center, Universidad Veracruzana, Sebastián Camacho No. 5, Xalapa, Ver., 91000, Mexico.*
[2]*CNRS; LAAS; Université de Toulouse; 7 avenue du Colonel Roche, F-31077 Toulouse Cedex, France.*
{*oscalra_820 at hotmail.com, [anmarin, hrios] at uv.mx, devy at laas.fr*}

Keywords: Service Robots, Furniture Detection

Abstract: Service robots are nowadays more and more common on diverse environments. In order to provide useful services, robots must not only identify different objects but also understand their use and be able to extract characteristics that make useful an object. In this work, a framework is presented for recognize home furniture by analyzing geometrical features over point clouds. A fast and efficient method for horizontal and vertical planes detection is presented, based on the histograms of 3D points acquired from a Kinect like sensor onboard the robot. Horizontal planes are recovered according to height distribution on 2D histograms, while vertical planes with a similar approach over a projection on the floor (3D histograms). Characteristics of points belonging to a given plane are extracted in order to match with planes from furniture pieces in a database. Proposed approach has been proved and validated in home like environments with a mobile robotic platform.

## 1 INTRODUCTION

When someone thinks of a fully functional service robot, it is very common to imagine a robot performing different tasks the same way humans do. Although many advances has been achieved from diverse groups around the world, e.g. in the fields of mobile robot localization, path planning and human-robot interaction; cognitive representation of environments is still a challenge. For a robot to understand that a chair has a surface to sit down, surface that it is closely linked with the property of being a chair, but that in some cases could be used to place objects, it is an abstraction very difficult to do for a robot.

Many works for detecting objects do not consider characteristics of high level that allow to understand the use or the properties of an object. Dealing with such a problem, in this paper it is proposed to analyze furniture pieces by extracting some of the principal characteristics of them, we focus particularly on pieces of furniture that can be moved or relocated by humans or the same robot while doing his tasks, e.g. cleaning beside or under a couch or bed. Furniture fixed to the environment as wardrobes or cabinets that do not move, are out of the scope of this work.

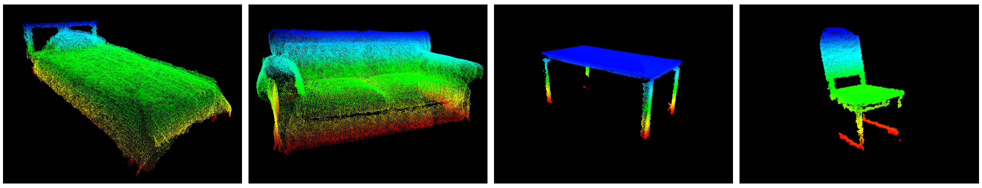
Along this work it is proposed to model and identify objects by its geometrical properties, i.e. horizontal or vertical planes, supports (legs), etc. By modeling individually each part of the furniture the robot could infer in the future, that a flat surface like the main plane of a table has the same characteristics as the horizontal plane of the chest of drawers and then suppose that is possible to pose objects on it.

This paper is organized as follow, on next section (2) some of the most important related works are described. Section 3 refers to planes characteristics extractions for furniture model representation and on section 4 those characteristics are used to identify objects in a data base. On section 5 are shown some experiments and finally on section 6 are given the conclusions and future work.

## 2 RELATED WORK

In the recent years, a lot of works for detection and recognition of objects have been developed, commonly working with 3D information, which gives to the robots more information but also increase the processing required. In order to keep this processing from becoming untreatable, new techniques have been created.

There are different ways to represent the 3D data of the objects, choosing the better representation will depend on the application. In (Nießner et al., 2013), a



(a) (b) (c) (d)

Figure 1: Examples of point cloud models: (a) bed, (b) couch, (c) table and (d) a chair.

simple spatial hashing scheme is presented for large and fine scale reconstruction. This helps for more efficient object reconstruction and real time updates. In (Wu et al., 2015) is proposed to represent the 3D shape of objects as a probability distribution of binary variables on a voxel grid using a Convolutional Deep belief Network. They are able to recognize and reconstruct objects based on their own dataset.

In (Wahl et al., 2003) is introduced a four-dimensional feature invariant to translation and rotation that captures the intrinsic geometrical relations between pairs of oriented surface points. In (Drost et al., 2010) is proposed an off-line global model description based on oriented point pair features. For the recognition phase they match the object and obtain its pose based on a voting scheme. This method, reported in (Salas-Moreno et al., 2013) for an object oriented SLAM technique, has shown good results when the objects occupy most of the field of view, but fails when the objects are distant or partially occluded.

Most of these techniques are able to capture small details on the objects and in general they work better for small objects, when the pieces of furniture of a home are analyzed, it is necessary to focus on bigger features like planes.

As stated in (Swadzba and Wachsmuth, 2014), it is reasonable to represent a 3D scene as a collection of planes since a typical indoor environment mostly consist of planar surfaces. In (Trevor et al., 2012) they also highlighting the importance of the planar surfaces as landmarks.

In (Günther et al., 2013) the furniture is represented as a set of planar structures that "have a certain size, orientation, height above ground and spatial relation to each other". A faster alternative to plane segmentation was presented in (Holz et al., 2012), using integral images and taking advantage of the structured point cloud from RGB-D cameras. In (Alonso-Ramirez et al., 2015) was presented an approach to a fast horizontal planes detection also based on structured point clouds.

In (Günther et al., 2017) is presented a classification technique based on semantic models from furniture objects based on their planar surface.

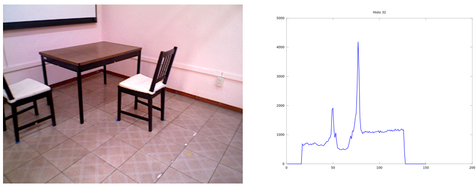
Taking advantage of statistical properties of the 3D world and the contextual relationships between the objects in the world, (Lin et al., 2013) present a technique for object detection and scene understanding. Based on (Carreira and Sminchisescu, 2012) to find object regions and later classify them. By using a conditional random field (CRF) model they integrate appearance, geometry, object relations with the environment and with other objects and they avoid some of the problems with feature-based approaches like pose variation, object occlusion or illumination changes.

These works highlight how plane detection combined with semantic information can provide a good representation of the environments for the robots. The proposed approach characterize the planes of the furniture and extract some characteristic information to classify them, information that could be used in the future for semantics.

# 3 Obtaining and characterizing the PCD models for the furniture

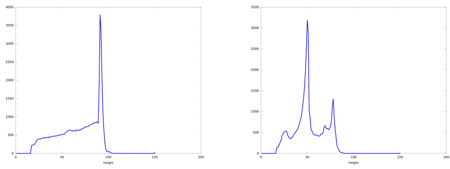
To be able to identify the furniture, the robot requires to previously model each one of the pieces of furniture in the environment. At this stage of the research, a specific model for each piece of furniture has been created and not a general model, however this could be incorporated later.

To obtain the models, there are two ways; to get the model from the internet or to construct the model ourselves. Our approach requires for the model to be a point cloud. It is possible to extract one from the online models (depending on its format) but is very likely some extracted points correspond from surfaces not visible for the robot, for example the lower side of a tables plane. In order to have a model similar to the point clouds the robot will see, we decided to construct the models ourselves from several point clouds from different views of the furniture.



(a) (b)

Figure 2: In (a) an indoor scene and in (b) its height histogram



(a) (b)

Figure 3: Examples of height histograms for (a) a chest of drawers and (b) a couch.

## 3.1 Creating a PCD model

We took several images around each piece of furniture. Since the robot was localized in the environment, all those images were in the same reference frame. Small errors in alignment are corrected with an ICP algorithm. Then, the different views were merged and down-sampled to obtain the point cloud model of the furniture. The figure 1 shows some examples of the PCD models obtained, the point cloud color represents the $z$ value of each point.

In order to characterize the obtained models, it is necessary to extract the planes from the PCD. Most of the methods for plane extraction use computationally expensive variations of RANSAC-like algorithms. However, a quick analysis of the distribution of the points can be enough.

Using an RGB-D camera mounted on a PTU in the robot, point clouds can be extracted and they can be transformed from the camera reference frame to the robot's base frame or to the world reference frame. Under these circumstances, detecting horizontal and vertical planes on the point clouds can be simplified by analyzing the concentration of the points, avoiding a complex mathematical approach.

## 3.2 Characterizing the horizontal planes

To extract the horizontal planes, the approach presented in (Alonso-Ramirez et al., 2015) was followed.

The first step is to construct a height histogram. Since the point cloud is transformed to the robot's base coordinate frame, a horizontal plane will concentrate a considerable amount of points at the same hight. In order to extract the planes from the histogram we need to search for the higher peaks.

The figure 2 shows an example of a height histogram, the figure 2(a) shows the RGB data of the point cloud that is being analyzed and the figure 2(b) shows the height histogram. It can be observed how there are two outstanding peaks on the histogram, which correspond to the height of the table and the height of the seats of the chairs.

To find the outstanding peaks, first the curve of the histogram is reduced with the DouglasPeucker algorithm (a curve simplification algorithm) to reduce the noise in the signal by eliminating small variations without increasing the bin size. Then a search is performed to find the larger peaks in the curve. All the points within those peaks are retrieved and projected to a floor plane. Then, a clustering algorithm is performed on the projection in order to separate different planes that can exist at the same high for example the seats of different chairs.

The Figure 3, shows the characteristic curves for some models of pieces of furniture. However this information is not enough for the robot to classify them since the histograms can change significantly during the execution due to noise and partial views. In order to characterize each plane detected the following characteristics are extracted.

**center** The 3D point center of the plane points

**orientation** Defining an horizontal or vertical plane

**height** The average height of the points.

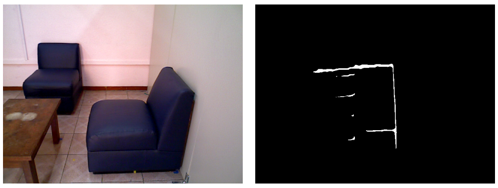
**height deviation** The standard deviation of the point's height.

**area** Area of the plane projected

**PCA eigenvectors and eigenvalues** Eigenvectors and eigenvalues resulting of a principal component analysis (PCA) to the plane points

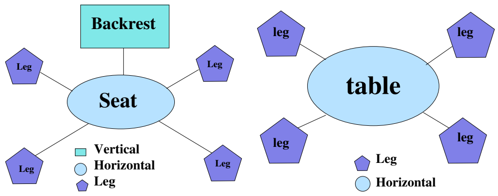
This information allow to perform a better analysis of the planes, which leads to a better classification. The horizontal planes of the furniture have different height, area, and some adjust better to a mathematical plane than others, for example the points from a table have a smaller height deviation than the points on the horizontal plane of a bed.

However, more information needs to be extracted from the point clouds apart from the horizontal planes, to complete the furniture models. In a different way as the presented on (Alonso-Ramirez et al., 2015), along this work the color or normal histograms are not used.



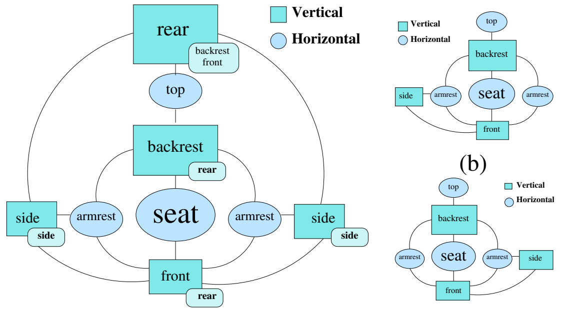
(a) (b)

Figure 4: In (a) indoor scene and (b) its vertical projection




(a) (b)

Figure 5: Graph for (a) a chair and (b) a table




(a) (b) (c)

Figure 6: In (a) a complete couch graph and in (b) the left and in (c) the right visual sub-graphs

## 3.3 Characterizing the vertical planes

For detecting the vertical planes, the same approach for the horizontal planes can be extended. A two dimensions histogram will be constructed, projecting all 3D points to a floor plane and then counting the number of points that lie at the same coordinates on the projection. All the points on a vertical plane projected to the floor plane will produce a line on the projection. Therefore retrieving all the points in those lines will extract the vertical planes.

In order to make the lines in the projection more remarkable, a smoothing filter has been applied and then morphological operators of dilation and erosion. Figure 4 shows an example of a vertical projection from a particular scene where it can be seen the walls project the bigger lines, smaller lines for the side, the front and the back of the couches and small dots for the table legs.

Since the lines in the projection are not always straight and their width can be variable, the Hough algorithm did not provide good results. To extract the lines a clustering algorithm is performed, the clusters obtained are approximated to polygons for reducing its number of points. A convex hull is calculated to determinate, based on its size, if the cluster is one single line or it has two o more lines together, for example two walls joining in a corner or a piece of furniture placed against the wall. If that is the case, the lines are separated on the joints. The points belonging to the planes are retrieved the same way as they were retrieved in the horizontal planes.

The vertical planes or regions detected will be considered as secondary characteristics to the horizontal regions. They will bring new information and together with the horizontal planes will complete the model of the furniture. The same characteristics calculated for the horizontal planes are calculated for the verticals. In this case, based on the PCA eigenvectors and eigenvalues it can be make a subdivision to the vertical planes and identify the legs of some furniture like tables and chairs.

## 3.4 Furniture Characterization

Once the models were created and their planes were separated and characterized, a graph is constructed. Each plane will be a node and the arcs will represent adjacency between two planes. The principal node will be the horizontal plane of the furniture or the biggest horizontal plane in cases where there are more than one. In Figure 5 are shown some examples of the graphs of the furniture's model. The node's shape represent it's type, the ellipses nodes correspond to horizontal planes, the rectangles to vertical planes and the pentagons are the furniture supports (legs).

These graphs contain all the planes of each piece of furniture. However, when we analyze an image from a scene the robot is unable to see all the planes since they can be on opposite sides. In the figure 6(a) is shown the graph for the couch. In the lower right corner of each node there is a list of the opposite planes, for example, if the robot sees the rear plane, depending on the angle maybe he can see the seat, but he can not see the backrest or the front plane.

Based on these viewing restrictions sub-graphs are created, each one showing the visible planes from the piece of furniture from different point of views. In the figures 6(b) and 6(c) there are sub-graphs from the couch viewed from the front left side and from the front right side respectively.

# 4 Scene analysis

Once a furniture's models database has been constructed and the robot is wandering in the environment, it is required to identify the furniture present in the scene. Once the robot takes a point cloud to analyze, the first step is to eliminate the floor by filtering all the points corresponding to a height smaller to $10cm$. And then, as described in the previous section, the horizontal and vertical planes are extracted.

The characteristics of the horizontal planes are compared with the main horizontal planes of the models. The probability of a plane to belong to each type of furniture is based on the equation (1).

$$p_{plane}(x|h,\sigma,a) = \frac{1}{2}ph + \frac{1}{4}p\sigma + \frac{1}{4}pa \tag{1}$$

where $h$, $\sigma$ and $a$ represent the heigh, the height deviation and the area of the plane. The probability is the weighted sum of the similarity of the attributes to the model's as described in the equations (2), (3) and (4). To classify the plane, all the categories with the probability higher than certain threshold are chosen.

$$ph = 1 - |h_m - h_x|/h_{range} \tag{2}$$

$$p\sigma = 1 - |\sigma_m - \sigma_x|/\sigma_m \tag{3}$$

$$pa = 1 - |a_m - a_x|/a_m \tag{4}$$

Once the horizontal plane has been classified as belonging to a furniture, a search is performed and the planes which are close to it are combined on a graph. This graph from the scene will be compared with the sub-graphs of the probable models.

First, it is necessary to match the nodes from the scene graph to the model's graph. The node for the main horizontal plane is automatically matched, a search of configurations is performed to match the rest of the nodes based on its characteristics and its position with respect to the main plane.

Once all the nodes from the scene graph are matched to the nodes of the model graph, their similarity is calculated in a similar way as the probability for the horizontal planes. Then, a weighted sum of the similarities of all the nodes, is performed. The weight for each pair of nodes will be the percentage of node from the total area of the model sub-graph.

$$p_{object} = \sum_{i}^{N} w_i p_{plane}(x) \tag{5}$$

# 5 Results in Furniture classification

The figure 7 shows some of the results obtained. The images on the left side are the RGB of the point cloud analyzed, the images in the center show the planes obtained and the images on the right show the resulting graphs and their classification.

The figure 7(a) shows a scene where two pieces of furniture can be found. It can be observed that two horizontal planes were detected. For the plane labelled as "H00" (which correspond to a coffee table) there were two plane probabilities higher than the threshold of 0.6, the higher probability for the coffee table and a smaller one for the couch. The results can be observed on the table 1. In this case, no adjacent planes were found, so the constructed graph will consist only on the main node. Once it is compared to the graphs for the coffee table and the couch, it is confirmed a higher probability for the former and then it is labelled.

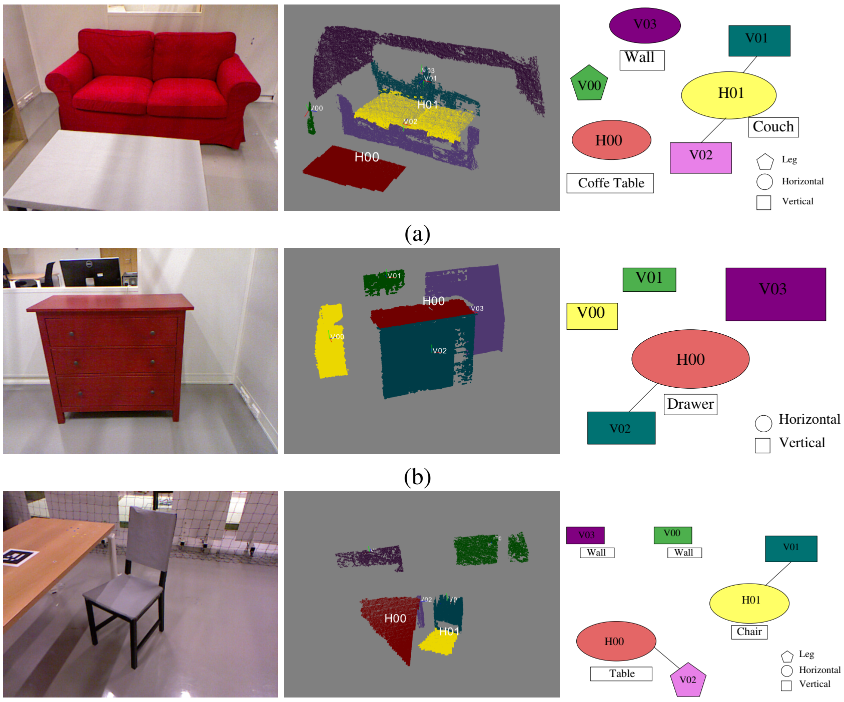
Table 1: Example for graph classification

| Main Node | Node Prob | Adjacent Nodes | Node Prob | Graph Probability |
|---|---|---|---|---|
| H00 | CoffeTable 0.8238 | none | – | 0.5517 |
| H00 | Couch 0.6903 | none | – | 0.1344 |
| H01 | Couch 0.9375 | V01 | backrest 0.7703 | 0.6241 |
| | | V02 | front 0.7960 | |
| H01 | Bed 0.7651 | V01 | no match – | 0.5895 |
| | | V02 | front 0.6496 | |

For the second plane "H01", corresponding to a couch, there were also two plane probabilities higher than the threshold, and there were two adjacent planes found. The results can be observed in table 1 were the higher object probability correspond to the couch. In figure 7(b) and 7(c) other results for different types of furniture can be found.

# 6 Conclusion

Along this work, an approach for home furniture detection, has been presented. The proposed approach is based on the characterization of geometric entities of the diverse pieces of furniture. As described, the main characteristic for these furniture pieces is its horizontal planes but also vertical planes and legs are extracted. A neighbourhood graph is then constructed in order to identify the correct piece of furniture present in the scene. The approach has been validated and proven their effectiveness on a home like environment with a robot moving in it. In future work, the characterization are going to be used to let the robot to infer about similar properties of geometric entities, e.g. the horizontal planes.




(a)

(b)

(c)

Figure 7: Results for the furniture detection

# REFERENCES


Alonso-Ramirez, O., Aguas-Garcia, Y., Marin-Hernandez, A., Rios-Figueroa, H. V., and Devy, M. (2015). An efficient alternative approach for home furniture detection and localization by an autonomous mobile robot. In *2015 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, pages 1–6.

Carreira, J. and Sminchisescu, C. (2012). Cpmc: Automatic object segmentation using constrained parametric min-cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(7):1312–1328.

Drost, B., Ulrich, M., Navab, N., and Ilic, S. (2010). Model globally, match locally: Efficient and robust 3d object recognition. In *IEEE CVPR'10*, pages 998–1005.

Günther, M., Wiemann, T., Albrecht, S., and Hertzberg, J. (2013). Building semantic object maps from sparse and noisy 3d data. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2228–2233.

Günther, M., Wiemann, T., Albrecht, S., and Hertzberg, J. (2017). Model-based furniture recognition for building semantic object maps. *Artificial Intelligence*, 247:336 – 351. Special Issue on {AI} and Robotics.

Holz, D., Holzer, S., Rusu, R. B., and Behnke, S. (2012). *Real-Time Plane Segmentation Using RGB-D Cameras*, pages 306–317. Springer Berlin Heidelberg, Berlin, Heidelberg.

Lin, D., Fidler, S., and Urtasun, R. (2013). Holistic scene understanding for 3d object detection with rgbd cameras. In *IEEE ICCV'13*, pages 1417–1424.

Nießner, M., Zollhöfer, M., Izadi, S., and Stamminger, M. (2013). Real-time 3d reconstruction at scale using voxel hashing. *ACM Trans. Graph.*, 32(6):169:1–169:11.

Salas-Moreno, R. F., Newcombe, R. A., Strasdat, H., Kelly, P. H. J., and Davison, A. J. (2013). Slam++: Simultaneous localisation and mapping at the level of objects. In *IEEE CVPR'13*, pages 1352–1359.

Swadzba, A. and Wachsmuth, S. (2014). A detailed analysis of a new 3d spatial feature vector for indoor scene classification. *Robotics and Autonomous Systems*, 62(5):646 – 662. Special Issue Semantic Perception, Mapping and Exploration.

Trevor, A. J. B., Rogers, J. G., and Christensen, H. I. (2012). Planar surface slam with 3d and 2d sensors. In *IEEE ICRA'12*, pages 3041–3048.

Wahl, E., Hillenbrand, U., and Hirzinger, G. (2003). Surflet-pair-relation histograms: a statistical 3d-shape representation for rapid classification. In *Fourth International Conference on 3-D Digital Imaging and Modeling, 2003. Proceedings.*, pages 474–481.

Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., and Xiao, J. (2015). 3d shapenets: A deep representation for volumetric shapes. In *IEEE CVPR'15*, pages 1912–1920.