Estudio bibliográfico de percepción 3D para tareas robóticas de manipulación de objetos.
A survey on 3D Perception for robot object manipulation tasks

Javier I. Peralta Sáenz.* Federico Ruiz Ugalde.**

ARCOS-Lab (Autonomous Robots and Cognitive Systems Laboratory)
INII (Instituto de Investigaciones en Ingeniería)
Universidad de Costa Rica
**(Tel: 2511-5146 ; e-mail: javier.peraltasaenz@ucr.ac.cr).*
*** Coordinador ARCOS-Lab (Tel: 2511-5141; e-mail: federico.ruizugalde@ ucr.ac.cr).*

Resumen: Un sistema confiable de manipulación de objetos requiere de percepción sofisticada. Aspectos importantes de este sistema es ser capaz de detectar la forma, tamaño y pose de los objetos a manipular. Estas capacidades se han mostrado antes en la literatura pero sin demostrar altos grados de confiabilidad en ambientes no controlados. Las soluciones actuales tienen muchas limitaciones para lograr un manejo adecuado de los objetos. El objetivo principal de este estudio es revisar las técnicas de percepción más relevantes al problema, sus características y limitaciones. Esto será de utilidad para dirigir el trabajo futuro a los aspectos que un robot (como el desarrollado en el ARCOS-Lab) necesita para manipular objetos en ambientes humanos no controlados cotidianos, como las cocinas.

Palabras clave: robótica, manipulación de objetos, percepción, segmentación semántica, estimación de pose, estimación de parámetros

Abstract: A reliable object manipulation system requires a sophisticated perception. Important aspects that are required for this type of manipulation are been able to detect the shape, size and pose of the objects to manipulate. These capabilities have been shown before but their reliability is usually not high enough to achieve proper object manipulation in uncontrolled environments. Many are the limitations of current solutions and many are the requirements for this type of manipulation. The main goal of this survey is to study the current most relevant methods, their features and limitations. This will be useful to concentrate future work on the missing aspects that our robots will need to achieve a reliable object manipulation in unconstrained every day human environments such as kitchens.

Keywords: robotics, object manipulation, perception, semantic segmentation, pose estimation, parameter estimation

1. INTRODUCTION

We would like to create robots capable of assisting humans on a variety of tasks on regular human environments and in the presence of humans and other actors. Such robots would require far more reliable and sophisticated capabilities than the current widely deployed industrial robots - unlike production lines, human everyday environments are unconstrained, objects are never located on the same places nor orientations, unfamiliar objects appear all the time and vary greatly in shape and uses and, on top of it all, humans and other actors move the objects around in an unpredictable fashion when compared with the orderly motion on factories.

In order for these robots to be useful they need to interact with the world around them, which calls for advance and reliable manipulation capabilities, hopefully matching humans in skill, which is a high bar - humans are capable of reacting quickly and using a wide array of tools to perform varied tasks. For our machines to have any hope of matching us, it needs to be able to see, interpret and understand the environment and the objects that compose it - to be able to do a simple task such as grasping an object, the robots needs not only to be able to locate that object in relation to itself, but also to determinate the object's geometry. Furthermore, depending in what the robot intends to do with the object, it may need to grasp it differently; it's not the same, passing a bottle of water to a human, than to open that bottle. Therefore, the robots need to understand what type of the object it is trying to manipulate, what are its intended uses and how is it going to be used on the next few actions, another issue is that the world is dynamic, it changes at a fast pace even if the robot takes no action, so the robot needs also to be able to track to a certain degree these changes when detected and also expect that observations made some time ago may no longer be valid. A robot with the ability to manipulate our everyday objects would also require a reliable perception system capable of extracting and processing all the information needed, potentially in real time.

Given that the world is 3D, it's useful for the perception systems to be 3D as well; the robot need the full 3D geometry, position and orientation of an object in order to plan and take any manipulating actions on it. 3D perception is challenging and different than its 2D counterpart. Point clouds, the main data structure used for pure 3D, are a collection of x,y,z coordinates representing points in space - unlike digital images, they are unorganized, two points that are contiguous on the data structure are not on the real 3D space, making it difficult to process efficiently. Another problem is that they are not uniformly sampled; there are regions in space which have more corresponding points than other regions, even if they both have about the same amount of objects in them. There also may be extra data depending on the sensor used, in addition to x,y,z, color information or intensity of the laser reflected may be included on the point cloud. Other problems that may arise which are more traditional are the noise of the sensors, occlusion (every object hides half of it self), transparent items that are not easy to sample with most 3D sensors, and clutter. It may also be necessary to fuse data from different types of sensors in order to complement or complete all the sufficient object information for manipulation. This sensor fusion may not be trivial.

We believe that the capabilities of detecting the position of objects, their geometry and the state, especially dimensions are critical to the success of manipulation of any kinds and this should be as reliable as possible. That is why on this paper we would like to focus on three subjects within the field of 3D perception: semantic segmentation, pose estimation and object parameter estimation. Semantic Segmentation is the problem of answering the question "what object is this?". Pose estimation algorithms helps to understand where is the object located and what's it's orientation. Finally object parameter estimation is the task of determining the state of the object. We believe these subjects to be key for the kind of complex manipulation we would like our robot to perform.

On the rest of this paper we will explain the general problem of how to perceive objects for manipulation task and later on more detail, some of the most particular techniques that are important parts of the object perception problem. As will be shown in this paper, there are many problems not yet completely solved and there are many methods not discussed in this paper because of a space constraint. Therefore, we will concentrate our study to some of the most relevant examples, to the above described problem, of the different techniques used and not a comprehensive review of each of the subjects.

2. HOW TO PERCEIVE OBJECTS FOR MANIPULATION

On this section, we will describe some known approaches to the problem and some of their possible limitations.

A particular approach that could allow skilled manipulation we are interested in is presented by(Ruiz Ugalde 2015). For the perception system, a mix of markers and artoolkit was used. This system is very reliable most of the time, however it requires each object to be tag with the marker, imposing a constrain on the environment. Today and in the near foreseeable future, regular everyday objects are not tagged, making this approach limited on real environments at best.

Other methods are described by (Rusu 2009). In this work, the author argues that it cannot be expected that any particular algorithm performs perfectly on all conditions, therefore he proposes to set a type of general algorithmic steps for general tasks and have a pool of algorithms that can be interchanged for each of those steps. The rest of the work consists on a algorithm for each task based both on machine learning and object model fitting, and some scenarios to test those algorithms. In general, the author claims good results, but he acknowledges that the algorithms need to be tested on a larger set of environments. Another pending issue is a method to select the right algorithm for a given task.

Another approach to perception is presented by (Klank 2012). Just like on the previous work, the author argues that no algorithm works perfectly for every task, however one of the key contributions of his work is a method to select an algorithm for a given task. The rest of the work describes a way to estimate the six degree of freedom pose of objects. This process uses images and the 3D models of the objects, the work explains how to match this to the environment and how to acquire new models via internet databases. This systems impose a constraint: all objects most be on their "upright" regular position, otherwise the search tree used to generate all 2D projections of the 3D objects grows too large. Another disadvantage is that the models should be of exactly the objects on the robot's environment. It is possible that the model can vary from the real object, but the performance and robustness of the algorithm may diminish. Also, it isn't clear how much the algorithm fails when the model is not exactly the same as the object.

For a radically different approach to manipulation, NVIDIA is trying use deep learning to transfer simulation scenarios to the real world on (Chebotar et al. 2018). The system would probably require a simpler perception system, as the deep learning neural network encodes the knowledge of 3D space and object physics, probably a simple digital camera is enough. On the downside of this method, is that it requires a large amount of simulations for each task, which means a large amount of computational power and time.

3. SEMANTIC SEGMENTATION

The goal in segmentation is to divide the 3D data into regions that belong to different entities, for example to divide the foreground and the background or the different objects on a scene from each other. In Semantic Segmentation, we need to separate objects from each other and also identify what is the object - for example a task in semantic segmentation is finding which points on a point cloud are part of a mug, or a table. We could also use more vague classification, for example finding the points that belong to a sphere, a cylinder or some form of geometric primitive.

Semantic segmentation has uses beyond object manipulation, therefore much research on the area is not directly applicable to the object manipulation problem. In particular, semantic segmentation of digital images is not useful to our problem. There are also many methods to do the segmentation and this is usually associated with pattern recognition/machine learning. We will review to families of methods, the ones based older and more traditional algorithms of machine learning and pattern recognition and those based on deep learning.

3.1 Traditional Machine Learning and Pattern Recognition

The techniques of machine learning and pattern recognition are older than those of deep learning, so there is a wider variety of better studied methods available. The typical approach has some pre-processing (noise removal, data registration, segmenting the data into smaller regions, etc) and then some features are extracted from the data and normally a previously trained classifiers deals with the semantic segmentation. The features used exploit geometric, positional, statistical or chromatic property of the objects to allow a proper classification.

Among these papers, the work of the Intelligent Autonomous Systems Laboratory from Technische Universitlit Munchen is specially interesting since their perception systems are designed for robots working on human environments. Some contributions include a new feature: the global fast point feature histogram (Rusu et al. 2009), work on creating algorithms for segmentation (Goron et al. 2012) and creating semantic maps of kitchens (Rusu et al. 2008; Rusu et al. 2008; Radu Bogdan Rusu et al. 2007; Blodow et al. 2011), , and experimental validations of those techniques. A lot of the work on these papers is covered by (Rusu 2009) and (Klank 2012) which was already discussed on a previous section. There is also work on mixing 2D and 3D and different sensors to improve performance and accuracy on (Marton et al. 2011).

While not necessarily designed specifically for robotic manipulation, there are also other works that create classifiers and features that could be useful for our application (Nissler, Marton, and Suppa 2013; Weinmann et al. 2015; Vo et al. 2015; Lai et al. 2011; Kriegel et al. 2013; Hausman et al. 2013; Gupta, Arbeláez, and Malik 2013; Herbst, Ren, and Fox 2011; Kim and Sukhatme 2014; Huang and You 2013; Nuricumbo et al. 2016).

In general, most of these methods are supervised, meaning they require sufficient training data which may not be easily accessible.

3.2 Deep Learning

Deep learning is a recent area of machine learning which bases itself on convolutional neural networks (CNN). This area is incredibly popular at this point in time, thanks to its great results on hard 2D computer vision applications.

Applying deep learning to point clouds is recent, with one of the first CNN designed in 2017 on (Qi et al. 2016). There has been various papers proposing improvements to the overall PointNet architecture, and a completely different approach presented by (Jiang, Wu, and Lu 2018). There are also attempts to reduce the 3D problem into a 2D one, but retaining the 3D annotations as a result, for example in (Boulch et al. 2018), various multi-view 2D snapshots are generated from the point cloud. Another approach is (McCormac et al. 2016), which uses tracking from odometry to transform the 2D sequence into 3D space. There is a specific survey for deep learning on semantic segmentation in (Garcia-Garcia et al. 2017).

While deep learning could potentially surpass more traditional methods, they require a large dataset for training to perform well, even larger than those required by traditional machine learning. Another problem is that CNNs for 3D are not invariant to rotations and scaling, but the typical workaround is to train it on various transformations of the same object, making the required dataset even larger. There also have been limitation regarding the size of point clouds that can be processed.

4. POSE ESTIMATION

Pose estimation is the problem of determining the object's position and orientation on 3D space. In order to do that we need to find 6 variables, 3 translational, and 3 rotational, given that objects have six degrees of freedom on 3D space.

There are multiple approaches, one presented by (Klank 2012), which was presented on a previous section. There are various variations to the general idea of searching the full pose on digital images as presented on [Riedel, Marton, and Kriegel (2016); Muja et al. (2011); Zhe Cao, Sheikh, and Banerjee (2016); Aldoma et al. (2011); Kouskouridas et al. (2012); Azad et al. (2011); Cornelius, Kragic, and Eklundh (2005); Ekvall, Kragic, and Hoffmann (2005); Kouskouridas et al. (2015);], they all share the disadvantage that extra information must be known about the objects and they all need a way to create all 2D projections of the 3D objects in order to work.

Deep learning is also applied in this area, a survey of deep learning for pose estimation is presented on (Li et al. 2016). That survey concludes that, even though there are promising directions for further work, none of the various deep learning pose estimation based methods is robust and fast enough to deal with all situations due to occlusions, scale of the scenes, pose changes, and multi-instance objects.

6. PARAMETER ESTIMATION

Parameter estimation is the problem of knowing the value of some property of an object. For grasping, the dimensions and centroids are the most important properties that can be deduced from visual sensor data. Given the manipulation system of our group, we are interested on treating objects based geometric primitives - spheres, cylinders, boxes, etc - or a combination of them.

On this line of work RANSAC is a classical algorithm to use, along with that (Rusu 2009) uses a machine learning SVM approach to deal with objects, which later a mesh is generated, however the tests ran on synthetically created objects. Another issue is that the mesh grid created can only work for pick and place kind of problems, since only an approximation of the surface is known and no other parameters.

(Vu et al. 2018) developed a variation of RANSAC called GCSAC to match cylinder and spheres. However, the experiments show large error rates up to 81.01% and the method requires complete scan of the object, which is not generally available to the robot because of occlusion.

6. CONCLUSIONS

There is interesting work on many perception problems needed for robot object manipulation, however a lot of work is still needed to bring the level of reliability needed. Much of the work on recent years is based on deep learning, which, for 3D data, is still very much new, with problems due to the large amounts of data required and processing speed needed for a robot to run this algorithms in real time.

Some of the contributions of the above work include (among others):

- Deep Learning based methods to deal with pose estimation and semantic segmentation.
- Traditional Machine learning methods to deal with all 3 problems.
- Various features that perform well on 3D.
- Work on aquireing models form the internet and the environment
- The use of data structures like semantic maps to categorize the environment

Even though the methods shown on the paper can perform well (given the right constrains) there is work that still needs to be done:

- Either the right datasets needs to be collected in order to train both traditional supervised methods and deep learning based ones or a way get better results from small dataset needs to be done. None supervised methods take a wide hit on accuracy and performance compared to the supervised ones.
- Deep Learning based methods don't scale well on real, unconstrained environments.
- Most of the methods require testing on more varied scenarios.

7. ACKNOWLEDGEMEMTS

We would like to thank El Instituto de Investigación en Ingenería (INII) and the School of Electrical Engineering (EIE) of la Universidad de Costa Rica of there support on equipment and space to work, Rectoría de la Universidad de Costa Rica for their support on getting our robot, to Vicerrectoría de investigación of Universidad de Costa Rica for their support and funding of our research work and to Postgrade School of Electrical Engineering (PPIE) of their support during the master's program. Without their contributions none of the work, of our group including this paper, would be possible.

REFERENCES

Aldoma, A., M. Vincze, N. Blodow, D. Gossow, S. Gedikli, R. B. Rusu, and G. Bradski. 2011. "CAD-Model Recognition and 6DOF Pose Estimation Using 3D Cues." In 2011 Ieee International Conference on Computer Vision Workshops (Iccv Workshops), 585–92. https://doi.org/10.1109/ICCVW.2011.6130296.

Azad, P., D. Münch, T. Asfour, and R. Dillmann. 2011. "6-Dof Model-Based Tracking of Arbitrarily Shaped 3D Objects." In 2011 Ieee International Conference on Robotics and Automation, 5204–9. https://doi.org/10.1109/ICRA.2011.5979950.

Blodow, N., L. C. Goron, Z. Marton, D. Pangercic, T. Rühr, M. Tenorth, and M. Beetz. 2011. "Autonomous Semantic Mapping for Robots Performing Everyday Manipulation Tasks in Kitchen Environments." In 2011 Ieee/Rsj International Conference on Intelligent Robots and Systems, 4263–70. https://doi.org/10.1109/IROS.2011.6094665.

Boulch, Alexandre, Joris Guerry, Bertrand Le Saux, and Nicolas Audebert. 2018. "SnapNet: 3D Point Cloud Semantic Labeling with 2D Deep Segmentation Networks." Computers & Graphics 71: 189–98. https://doi.org/https://doi.org/10.1016/j.cag.2017.11.010.

Chebotar, Yevgen, Ankur Handa, Viktor Makoviychuk, Miles Macklin, Jan Issac, Nathan D. Ratliff, and Dieter Fox. 2018. "Closing the Sim-to-Real Loop: Adapting Simulation Randomization with Real World Experience." CoRR abs/1810.05687. http://arxiv.org/abs/1810.05687.

Cornelius, H., D. Kragic, and J. -. Eklundh. 2005. "Object and Pose Recognition Using Contour and Shape Information." In ICAR '05. Proceedings., 12th International Conference on Advanced Robotics, 2005., 613–20. https://doi.org/10.1109/ICAR.2005.1507472.

Ekvall, Staffan, Danica Kragic, and Frank Hoffmann. 2005. "Object Recognition and Pose Estimation Using Color Cooccurrence Histograms and Geometric Modeling." Image Vision Comput. 23 (11): 943–55. https://doi.org/10.1016/j.imavis.2005.05.006.

Garcia-Garcia, Alberto, Sergio Orts-Escolano, Sergiu Oprea, Victor Villena-Martinez, and José García Rodríguez. 2017. "A Review on Deep Learning Techniques Applied to Semantic Segmentation." CoRR abs/1704.06857. http://arxiv.org/abs/1704.06857.

Goron, L. C., Z. Marton, G. Lazea, and M. Beetz. 2012. "Robustly Segmenting Cylindrical and Box-Like Objects in Cluttered Scenes Using Depth Cameras." In ROBOTIK 2012; 7th German Conference on Robotics, 1–6.

Gupta, S., P. Arbeláez, and J. Malik. 2013. "Perceptual Organization and Recognition of Indoor Scenes from Rgb-d Images." In 2013 Ieee Conference on Computer Vision and Pattern Recognition, 564–71. https://doi.org/10.1109/CVPR.2013.79.

Hausman, K., F. Balint-Benczedi, D. Pangercic, Z. Marton, R. Ueda, K. Okada, and M. Beetz. 2013. "Tracking-Based Interactive Segmentation of Textureless Objects." In 2013 Ieee International Conference on Robotics and Automation, 1122–9. https://doi.org/10.1109/ICRA.2013.6630713.

Herbst, E., X. Ren, and D. Fox. 2011. "RGB-d Object Discovery via Multi-Scene Analysis." In 2011 Ieee/Rsj International Conference on Intelligent Robots and Systems, 4850–6. https://doi.org/10.1109/IROS.2011.6095116.

Huang, J., and S. You. 2013. "Detecting Objects in Scene Point Cloud: A Combinational Approach." In 2013 International Conference on 3D Vision - 3DV 2013, 175–82. https://doi.org/10.1109/3DV.2013.31.

Jiang, Mingyang, Yiran Wu, and Cewu Lu. 2018. "PointSIFT: A Sift-Like Network Module for 3D Point Cloud Semantic Segmentation." CoRR abs/1807.00652. http://arxiv.org/abs/1807.00652.

Kim, D. I., and G. S. Sukhatme. 2014. "Semantic Labeling of 3D Point Clouds with Object Affordance for Robot Manipulation." In 2014 Ieee International Conference on Robotics and Automation (Icra), 5578–84. https://doi.org/10.1109/ICRA.2014.6907679.

Klank, Ulrich Friedrich. 2012. "Everyday Perception for Mobile Manipulation in Human Environments." Dissertation, München: Technische Universität München.

Kouskouridas, R., A. Amanatiadis, S. A. Chatzichristofis, and A. Gasteratos. 2015. "What, Where and How? Introducing Pose Manifolds for Industrial Object Manipulation." Expert Systems with Applications 42 (21): 8123–33. https://doi.org/https://doi.org/10.1016/j.eswa.2015.06.039.

Kouskouridas, R., T. Retzepi, E. Charalampoglou, and A. Gasteratos. 2012. "Ontology-Based 3D Pose Estimation for Autonomous Object Manipulation." In 2012 Ieee International Conference on Imaging Systems and Techniques Proceedings, 476–81. https://doi.org/10.1109/IST.2012.6295586.

Kriegel, S., M. Brucker, Z. Marton, T. Bodenmüller, and M. Suppa. 2013. "Combining Object Modeling and Recognition for Active Scene Exploration." In 2013 Ieee/Rsj International Conference on Intelligent Robots and Systems, 2384–91. https://doi.org/10.1109/IROS.2013.6696691.

Lai, K., L. Bo, X. Ren, and D. Fox. 2011. "Sparse Distance Learning for Object Recognition Combining Rgb and Depth Information." In 2011 Ieee International Conference on Robotics and Automation, 4007–13. https://doi.org/10.1109/ICRA.2011.5980377.

Li, W., Y. Luo, P. Wang, Z. Qin, H. Zhou, and H. Qiao. 2016. "Recent Advances on Application of Deep Learning for Recovering Object Pose." In 2016 Ieee International Conference on Robotics and Biomimetics (Robio), 1273–80. https://doi.org/10.1109/ROBIO.2016.7866501.

Marton, Zoltan-Csaba, Dejan Pangercic, Nico Blodow, and Michael Beetz. 2011. "Combined 2D-3D Categorization and Classification for Multimodal Perception Systems." Int. J. Rob. Res. 30 (11): 1378–1402. https://doi.org/10.1177/0278364911415897.

McCormac, John, Ankur Handa, Andrew J. Davison, and Stefan Leutenegger. 2016. "SemanticFusion: Dense 3D Semantic Mapping with Convolutional Neural Networks." CoRR abs/1609.05130. http://arxiv.org/abs/1609.05130.

Muja, M., R. B. Rusu, G. Bradski, and D. G. Lowe. 2011. "REIN - a Fast, Robust, Scalable Recognition Infrastructure." In 2011 Ieee International Conference on Robotics and Automation, 2939–46. https://doi.org/10.1109/ICRA.2011.5980153.

Nissler, C., Z. Marton, and M. Suppa. 2013. "Sample Consensus Fitting of Bivariate Polynomials for Initializing Em-Based Modeling of Smooth 3D Surfaces." In 2013 Ieee/Rsj International Conference on Intelligent Robots and Systems, 4228–34. https://doi.org/10.1109/IROS.2013.6696962.

Nuricumbo, Jorge René, Haider Ali, Zoltán-Csaba Márton, and Marcin Grzegorzek. 2016. “Improving Object Classification Robustness in Rgb-d Using Adaptive Svms.” Multimedia Tools Appl. 75 (12): 6829–47. https://doi.org/10.1007/s11042-015-2612-7.

Qi, Charles Ruizhongtai, Hao Su, Kaichun Mo, and Leonidas J. Guibas. 2016. “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation.” CoRR abs/1612.00593. http://arxiv.org/abs/1612.00593.

Radu Bogdan Rusu, N. Blodow, Z. Marton, A. Soos, and M. Beetz. 2007. “Towards 3D Object Maps for Autonomous Household Robots.” In 2007 Ieee/Rsj International Conference on Intelligent Robots and Systems, 3191–8. https://doi.org/10.1109/IROS.2007.4399309.

Riedel, S., Z. Marton, and S. Kriegel. 2016. “Multi-View Orientation Estimation Using Bingham Mixture Models.” In 2016 Ieee International Conference on Automation, Quality and Testing, Robotics (Aqtr), 1–6. https://doi.org/10.1109/AQTR.2016.7501381.

Ruiz Ugalde, Federico. 2015. “Compact Models of Objects for Skilled Manipulation.” Dissertation, München: Technische Universität München.

Rusu, Radu Bogdan. 2009. “Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments.” Dissertation, München: Technische Universität München.

Rusu, Radu Bogdan, Zoltan Csaba Marton, Nico Blodow, Mihai Dolha, and Michael Beetz. 2008. “Towards 3D Point Cloud Based Object Maps for Household Environments.” Robotics and Autonomous Systems 56 (11): 927–41. https://doi.org/https://doi.org/10.1016/j.robot.2008.08.005.

Rusu, R. B., A. Holzbach, M. Beetz, and G. Bradski. 2009. “Detecting and Segmenting Objects for Mobile Manipulation.” In 2009 Ieee 12th International Conference on Computer Vision Workshops, Iccv Workshops, 47–54. https://doi.org/10.1109/ICCVW.2009.5457718.

Rusu, R. B., Z. C. Marton, N. Blodow, M. E. Dolha, and M. Beetz. 2008. “Functional Object Mapping of Kitchen Environments.” In 2008 Ieee/Rsj International Conference on Intelligent Robots and Systems, 3525–32. https://doi.org/10.1109/IROS.2008.4650972.

Vo, Anh-Vu, Linh Truong-Hong, Debra F. Laefer, and Michela Bertolotto. 2015. “Octree-Based Region Growing for Point Cloud Segmentation.” ISPRS Journal of Photogrammetry and Remote Sensing 104: 88–100. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2015.01.011.

Vu, Hai, Van-Hung Le, Thi Thanh Thuy Nguyen, Thi-Lan Le, and Thanh-Hai Tran. 2018. “Fitting Cylindrical Objects in 3-d Point Cloud Using the Context and Geometrical Constraints.” In.

Weinmann, Martin, Boris Jutzi, Stefan Hinz, and Clément Mallet. 2015. “Semantic Point Cloud Interpretation Based on Optimal Neighborhoods, Relevant Features and Efficient Classifiers.” ISPRS Journal of Photogrammetry and Remote Sensing 105: 286–304. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2015.01.016.

Zhe Cao, Y. Sheikh, and N. K. Banerjee. 2016. “Real-Time Scalable 6DOF Pose Estimation for Textureless Objects.” In 2016 Ieee International Conference on Robotics and Automation (Icra), 2441–8. https://doi.org/10.1109/ICRA.2016.7487396.A., M.