Towards lifelong mapping in pointclouds

University of Maastricht
panos@almende.org

Abstract

The thesis discuses the application of dependent Dirichlet processes and how such tools can help improve the quality of clustering algorihms in the recently introduced data representation objects, pointclouds. A novel method to point cloud clustering is introduced its mechanics are analysed. Finally Its potential weaknesses and what steps can be taken to remedy them are analysed.

Contents

List of Figures

List of Tables

I. Introduction

Simultaneous localization and mapping is one of the fundamental problems of autonomous systems[1]. In order for a robot to be truly autonomous, it must have the ability to enter an area and infer its structure. To that direction, a lot of effort has been put in algorithms that are able to map static environments. With solutions like EKF-SLAM[2], FastSlam[39] and GraphSLAM[3] robots are now able to efficiently map static environments. The logical extention to methods that can map static environments is methods that remove this restriction. The idea of lifelong robot learning is not new and has been introduced as a general concept to the literature by Sebastian Thrun [5]. Konolige et al.[6] specifically focus on lifelong learning in mapping and the utility such methods would have. In the PhD thesis of Walcott [7] long term mapping is decomposed to 4 basic subproblems:

- Continuously incorporate new information.
- Address the problem of tractability for growing DPG
- Representation of the environment should include the history of the map as changes occur with the passage of time.
- Detect changes and update the map online

The first two problems can be though of as compression problems as the map increases over time whereas the latter ones can be though of as dynamic environment problems. Methods of tackling those problems vary according to what sensors a robot uses to perform the mapping. In this project the focus will be directed in methods that use RGBD devices to perform SLAM like Microsoft's Kinect.

Since its introduction in 2010 Microsoft Kinect[25] has revolutionized RGBD devices with its low price range and high quality sensors. It came as no surprise that research in point clouds, the representation system of Kinect sensor readings, has increased since. Many libraries that enable the user to perform tasks from feature extraction to plane segmentation[12] in pointclouds are currently available. In the field of robotics, many teams are using the Kinect sensors to perform simultaneous localization and mapping[13]. The goal of this thesis is to introduce a novel approach to tackle the compression problem of long term mapping methods that use the Kinect device by using Bayesian non parametric methods.

Dirichlet processes and Dirichlet process mixture models [26] are the cornerstone of Bayesian non parametric statistics. The strength of those models lies in the fact that they allow the model's mixture components to grow as much as needed so as to best fit the data. The dynamic number of components in combination with the highly resilient priors leads to very flexible models that can be used in a very large area of applications from topic modeling[42] to speaker diarization[44].

The main motivation behind this project is to use such methods as a means of compressing the information provided by the environment. In that direction, finding a way to robustly clustering a point cloud into semantically sound "chunks" of structure seems a reasonable starting point. This leads to the direction of object based SLAM, which is a domain where objects are used as reference points to perform the mapping.

In this paper, a novel EKF SLAM algorithm that takes point clouds as visual input will be implementeed. Its ability to compress the data a point cloud introduces will be presented and analysed; Furthermore, directions on how the method could be extended to tackle the first two subproblems of Walcott's thesis will be given in the discussion.

The rest of the paper is structured as follows. Section II will present relevant literature review, Section III will introduce the theories behind the model, Section 4 IV will define the model, Section 5 V will show experimental results of the method. Finally, Section 6 VI will end up with a discussion on the methods strengths and weaknesses.

II. Literature review

Related research will be focused on 4 general sub fields of related literature.

- Object based SLAM or semantic slam
- Point Cloud Object segmentation
- Non-parametric clustering methods
- The correspondence problem in SLAM

More specifically, object based SLAM is crucial due to the fact that point clouds are used as input and from such raw input objects representations must be extracted. Methods that use such approaches to perform SLAM are then needed. The second part of the research is focused on point cloud representations. This part of the research is mostly focused on what features or meta features need to be taken into account so that the reduced representation is optimal. The third part of the research is focused on non-parametric Bayesian methods and the clustering tools they provide. Such tools are important as they can be used to provide novel approaches to object segmentation within a point cloud. Finally, research is focused on the correspondence problem in SLAM. As one of the fundamental problems that need to be solved in order to have robust SLAM algorithms, it is imperative the correspondence problem be solved efficiently. In that extent and due to the unique representation of our objects, a novel approach in the correspondence between objects in a SLAM problem is given.

I. Object based SLAM

Object based SLAM or semantic slam methods proposed to the literature focus in domain specific solutions of a particular problem. Salas-Moreno et al [14] define a method of performing object based slam for specific objects. The objects are identified by camera that is on top of the robot. By having a model of pretrained objects SLAM can be performed on environments the robot knows what objects to expect. The disadvantage of that method is that object models have be to be well defined and there is a small number of such objects. Castle et al. use object recognition to perform object based SLAM with the use of a hand-held cameras. Selvatici et al [15] use a similar approach while exploiting structural information such as object height and position within the room. That way a couch that is a large object situated in floor level is easier to be recognized. Choudhary et al. [17] use point clouds and an object database to match objects currently seen with known objects within their database. They use omnimaper [27] as their mapping method and as a representation a combination of the downsampled voxel grids with additional normal and curvature information.

Finally, all their operations are done in the non-planar components of the point cloud. Jensfelt et al [15] present an object based approach to SLAM where the robot can manipulate the objects of the map. They use camera pictures as input and receptive Field Histogram as the method to abstract the camera input and extract features for their object matching algorithm. Their approach is proposed as a solution to a service robot scenario. MonoSLAM [19] introduces a method of performing slam using a monocular camera.

What all the aforementioned methods have in common is that they approach the problem of object based slam as a classification task. Objects need to be semantically understood before they are processed. The approach introduced in this paper considers the environment to be a collection of chunks. So having specific enough environment descriptors should lead to the robot being able to operate in a label free environment. This would remove the need of having to classify objects but would also increase the time it takes to extract features from the environment as features are the base of the unsupervised object discovery. Seongyong Koo et al. [20] introduce a method of unsupervised object individuation from RGB-D image sequences. They cluster their initial cloud into candidate objects using Eucledian clustering and proceed to extract features like the Euclidian distance(L2) and the Kullback-Leibler distance between point cloud objects. They use IMFT to solve their tracking problem.

II. Point Cloud Object segmentation

In the problem at hand, augmenting the information each cloud containts is crucial for a successful method of object representation. Research towards object segmentation in point clouds is focusing on calculating meta information regarding the points and applying some heuristic function to see if the points could belong in the same segment of the cloud. Trevor et al. [29] take positional information, Euclidean distances and the normal of points to as input to their heuristic and output segments that are part of the same object. PCL library [12] introduces methods like Euclidean clustering and conditional Euclidean clustering that use a number of heuristics that take normal as well as curvature information to extract segments in the point cloud that represent objects. Furthermore, a there is a lot of resarch on segmentation of point clouds in scenes, the emphasis is usually on extracting geometric primitives [30], [31] using cues like normals and curvature. Rabbani et al [23] introduce a new method of object segmentation using KNN as their base algorithm. They also present a very informative literature review along with the strengths and weaknesses of existing methods. Finally Triebel et al. [32] introduce a general clustering framework that does not rely on plane segmentation. Instead of segmenting the plane by using classical approaches like RANSAC or MLASAC they introduce a framework where they make no assumptions regarding plane data.

III. Non Parametric Bayesian methods

Bayesian non-parametric methods are the cornerstone of Bayesian statistics. In this project the focus was directed towards the clustering methods that are being introduced by those tools. Radford M. Neal [37] with his paper regarding MCMC methods for Dirichlet process mixture models made the definitive step towards Dirichlet process mixture models(DPMM's) reaching mainstream sucess. Since then, a variety of approaches for inference on such models has been introduced with Statistical inference and MCMC methods, and Variational inference being two prominent ones. Variational inference for DPMM's was introduced by Jordan et al. [38] and it introduces deterministic tools to perform inference and approximate a posterior distribution and marginals. Both methods have strengths and weaknesses and many tools have been established by using the two approaches as their base. Blei et al. [42] introduced LDA as a method to perform topic modelling. Teh et al [40] introduce a hierarchy on the inference process by introducing the Hierarchical Dirichlet process. Particle filter approaches have also been established. Doucet et al. [41] introduce Sequential Monte Carlo as a fast way to approximate inference. Inference on Dirichlet process mixtures is a very active research field and covering it is beyond the scope of this report. In this project SMC samplers where used due to their robustness as well as their inherent extensiveness.

IV. Correspondence

In its general definition, the correspondence problem refers to the problem of ascertaining which parts of one image correspond to which parts of another image, where differences are due to movement of the camera, the elapse of time, and/or movement of objects in the photos. Under the semantic SLAM context, it refers to the problem of identifying objects as objects that have been encountered before during the mapping process. Towards that direction Cree et al. [34] create a histogram of line segments of each landmark and compute their root mean square error. They then proceed to calculate their RGB signature to calculate the distance between different landmarks. Low et al. [35] match Scale Invariant Feature Transform (SIFT) features, an approach which transforms image data into scale-invariant coordinates relative to local features. Lamon et al [36] store a database of fingerprints which indicate the location in the robot's environment. The features are ordered and stored at a database at as they appear in the robot's immediate surroundings. A new fingerprint is computed for each new view and matched against existing ones.

The approach presented in this paper takes as input features as the aforementioned methods do and is similar to [20] as parts of the point cloud are being clustered and fit into a distribution. The features that are used for the cloud representation are an extension of the features presented in [33] with the addition of extra angular information present in the points of the cloud. The distance among clusters can be then represented as one among their distributions and there has been extensive search on statistical distribution distance. Distances like Hellinger, KL divergence, Euclidian, Mahalanobis can all be taken into account when performing the object matching. The robustness of the method can also be increased by using tracking methods like IMFT. The novelty lies in its completely probabilistic mechanism as the clustering is done by using SMC to the augmented feature space. Using distributions as a means of representing objects within the point cloud is a form of compression as objects are represented by a distribution which is smaller in size and easier to maintain and expand. Finally, instead of using heuristics, a novel approach to distribution correspondence is introduced through the means of a random forest algorithm.

III. Theory background

Relevant theory will be divided into 2 major sections

- Dependent Dirichlet processes
- SLAM

I. Dependent dirichlet process mixture models

Dependent Dirichlet processes(DDP) remove the restriction of exchangeable data. Data are now being given dependencies which could be temporal, positional etc. The DDPs are a natural extension of the DP's in domains where data cannot be considered exchangeable. They where introduced by MacEachern [43] and have been widely used since. The main motivation behind using such methods is the imediate extension they provide to dynamic environments. Since data on every frame are dependent on data of the previous frame, using such methods is intuitively straightforward.

IV. Model definition

The models introduces a Pipeline in the layer of landmark detection of the EKF-SLAM method. All the operations from feature extraction to clustering and object matching will be performed at that layer. More specifically the pseudocode for the method is the following:

Algorithm 1 Landmark Layer

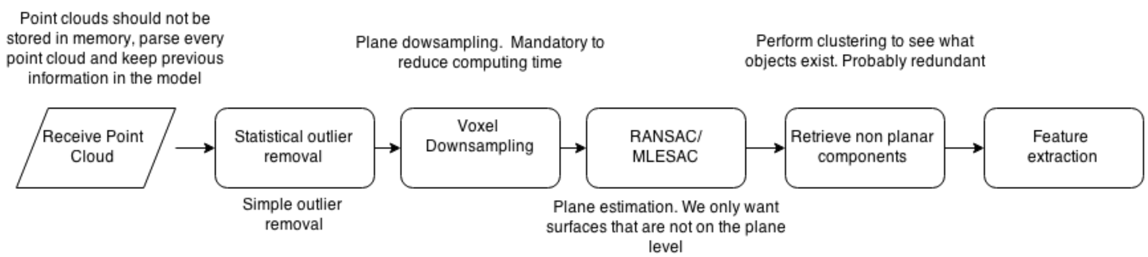
```
procedure getLandmarkId(pointCloud, timepoint, existingLandmarks)   ▷ Post transformation
    initialize(landMarkIds)                                         ▷ Initialize empty list
    pointCloudReduced ← extractMetaFeatures(pointCloud)             ▷ Cloud preprocessing
    features ← extractMetaFeatures(pointCloudReduced)
    landmarks ← cluster(features)                                   ▷ Cluster using features
    for landmarks as landmark do
        (probability, landId) ← getBestLandmarkCorrespondence(landmark, existingLandmarks)
        if probability > threshold then
            addLandmarks(landMarkIds, landId)                       ▷ Return known landmark
        else
            newLandID ← addLandmarkDB(landmarkDB, landmark)         ▷ Add get new id
            addLandmarks(newLandID)                                 ▷ Add landmark
        end if
    end for
    return landMarkIds                                              ▷ Return added landmark
end procedure
```

This top level description has a lot of implied steps so a line by line description will be provided:

Method input: The method takes as input a point cloud. The post transformation comment has to do with the fact that the cloud expected is the one after all the frame transformations are done in the tf layer.

Lines 3-4: Feature extraction is done through the pcl [12] library. An initial point cloud reduction is mandatory to increase the speed of the process. A voxel grid is used to reduce the dataset size. A leaf size of 0.04cm produces a good tradeoff between precision and speed. The object representation approach is similar to [17]. Instead of using the CSHOT descriptor, pcl's fpfh [22] is being used. Fast point feature histogram(fpfh) represents an angular signature between a point and its neighbors. In that way we end up with a 23 dimensional angular signature of the angular information between a point and its neighbors. Since the information are given in the form of a histogram, classical statistic solutions regarding the distances between two points can be taken into account. In our solution EMD,Hellinger,and KL divergence are being computed in the pipeline. Color information is being encoded with an approach similar to [33]. The colour spectrum is discretized and what is extracted is the count of different colour signatures between a point and its k nearest neighbors. Finally positional information is also given as input to the algorithm. The pipeline is presented in figure fig. 1. What the algorithm outputs is then a vector of $\mathbf{x} = (x_s, x_c, x_a)$ where s represents a 3x1 space information, c colour information 27x1 and a angular information whose dimensionality depends on the distances computed(in our case 3x1).

Figure 1: Point cloud modification pipeline.

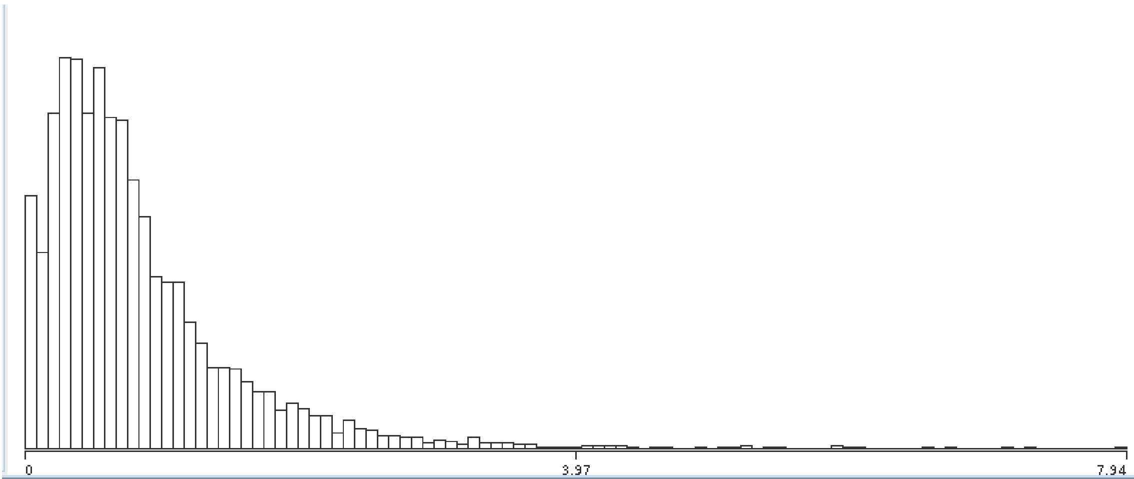


Lines 5: The clustering takes place in this line. The input of the method is the feature vector calculated in the previous steps. An SMC sampler is used as was presented in the theory section. The data are modeled as following a distribution $G_0(\theta_k) = Normal(x_n^s|\mu, \Sigma) * Mult(x_n^c|\delta) * exp(x_n^a|\lambda)$ Where Normal is a 3 dimensional Gaussian distribution with mean μ and covariance Σ, mult is a multinomial with probability vector δ_t and exp is an exponential with rate λ. The exponential distribution was chosen after the data analysis showed the exponential trend on the data follow as can be seen in figure fig. 2

Lines 6-12: The correspondence of the previous landmarks to current ones happens in theses lines.The getBestLandmarkCorrespondence function is a random forest implementation. An offline model has been trained to recognize correlation between distances and data correspondence. This approach is similar to [20] and many the efficiency of the method depends on the training to be general enough. The random forest outputs the probability of a landmark having been encountered before; if it has, it is being added to the landmarks list to be send for update in the EKF, otherwise a new landmark is added and its ID is then updated to the list.

Lines 15: The algoritm returns the list of landmarks given the pointcloud and all the previous landmarks the robot has encountered.

Figure 2: Exponential trend in KL distances.



V. Results

VI. Conclusion and future work

A novel method for SLAM by using compressed representations of objects in a point cloud is introduced. Its strengths and weaknesses are presented. Future work could include

- Improved tracking by using IMRF
- Extend to dynamic environments
- Extend the representation used

Table 1: Example table

Name		
First name	Last Name	Grade
John	Doe	7.5
Richard	Miles	2

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

$$e = mc^2 \tag{1}$$

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

VII. Discussion

I. Subsection One

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

II. Subsection Two

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

References

[1] Thrun, S. (2002). Probabilistic robotics. Communications of the ACM, 45(3), 52-57.

[2] Bailey, T., Nieto, J., Guivant, J., Stevens, M., & Nebot, E. (2006, October). Consistency of the EKF-SLAM algorithm. In Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on (pp. 3562-3568). IEEE.

[3] Thrun, S., & Montemerlo, M. (2006). The graph SLAM algorithm with applications to large-scale mapping of urban structures. The International Journal of Robotics Research, 25(5-6), 403-429.

[4] Figueredo, A.J. and Wolf, P. S.A. (2009). Assortative pairing and life history strategy - a cross-cultural study. *Human Nature*, 20:317–330.

[5] Thrun, S., & Mitchell, T. M. (1995). Lifelong robot learning. The Biology and Technology of Intelligent Autonomous Agents, 165-196.

[6] Konolige, K., & Bowman, J. (2009, October). Towards lifelong visual maps. In Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on (pp. 1156-1163). IEEE.

[7] Walcott, A. (2011). Long-term robot mapping in dynamic environments (Doctoral dissertation, Massachusetts Institute of Technology).

[8] Hjort, N. L., Holmes, C., MÃijller, P., & Walker, S. G. (Eds.). (2010). Bayesian nonparametrics (Vol. 28). Cambridge University Press.

[9] MacEachern, S. N. (2000) Dependent dirichlet processes. Unpublished manuscript, Department of Statistics, The Ohio State University.

[10] Barber, D. (2012) Bayesian reasoning and machine learning.

[11] Neiswanger, W., Wood, F., & Xing, E.The dependent dirichlet process mixture of objects for detection-free tracking and object modeling. In Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics (pp. 660-668) (2014, August)

[12] Rusu, R. B., & Cousins, S. (2011, May). 3d is here: Point cloud library (pcl). In Robotics and Automation (ICRA), 2011 IEEE International Conference on (pp. 1-4). IEEE.

[13] LabbÃľ, M., & Michaud, F. (2011, September). Memory management for real-time appearance-based loop closure detection. In Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on (pp. 1271-1276). IEEE.

[14] Salas-Moreno, R. F., Newcombe, R. A., Strasdat, H., Kelly, P. H., & Davison, A. J. (2013, June). Slam++: Simultaneous localisation and mapping at the level of objects. In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on (pp. 1352-1359). IEEE.

[15] Selvatici, A. H., & Costa, A. H. (2008). Object-based visual slam: How object identity informs geometry.

[16] Castle, R. O., Gawley, D. J., Klein, G., & Murray, D. W. (2007, April). Towards simultaneous recognition, localization and mapping for hand-held and wearable cameras. In Robotics and Automation, 2007 IEEE International Conference on (pp. 4102-4107). IEEE.

[17] Choudhary, S., Trevor, A. J., Christensen, H. I., & Dellaert, F. (2014, September). SLAM with object discovery, modeling and mapping. In Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on (pp. 1018-1025). IEEE.

[18] Jensfelt, P., Ekvall, S., Kragic, D., & Aarno, D. (2006, September). Augmenting slam with object detection in a service robot framework. In Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on (pp. 741-746). IEEE.

[19] Davison, A. J., Reid, I. D., Molton, N. D., & Stasse, O. (2007). MonoSLAM: Real-time single camera SLAM. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 29(6), 1052-1067.

[20] Koo, S., Lee, D., & Kwon, D. S. (2014, September). Unsupervised object individuation from RGB-D image sequences. In Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on (pp. 4450-4457). IEEE.

[21] Cichocki, A., & Amari, S. I.Families of alpha-beta-and gamma-divergences: Flexible and robust measures of similarities. Entropy, 12(6), 1532-1568.

[22] Fast point feature histogram.Rusu, R. B., Blodow, N., & Beetz, M. (2009, May). Fast point feature histograms (FPFH) for 3D registration. In Robotics and Automation, 2009. ICRA'09. IEEE International Conference on (pp. 3212-3217). IEEE.

[23] Rabbani, T., van den Heuvel, F., & Vosselmann, G. (2006). Segmentation of point clouds using smoothness constraint. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(5), 248-253.

[24] Caron, F., Davy, M., & Doucet, A. (2012) Generalized Polya urn for time-varying Dirichlet process mixtures. arXiv preprint arXiv:1206.5254.

[25] Zhang, Z. (2012) Microsoft kinect sensor and its effect. MultiMedia, IEEE, 19(2), 4-10.

[26] Wainwright, M. J., & Jordan, M. I. (2008). Graphical models, exponential families, and variational inference. Foundations and Trends in Machine Learning, 1(1-2), 1-305.

[27] A.Trevor, J.Rogers, and H.Christensen. Omnimapper: A modular multimodal mapping framework. In IEEE International Conference on Robotics and Automation (ICRA), 2014

[28] Koo, S., Lee, D., & Kwon, D. S. (2013, November). Multiple object tracking using an rgb-d camera by hierarchical spatiotemporal data association. In Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on (pp. 1113-1118). IEEE.

[29] Trevor, A. J., Gedikli, S., Rusu, R. B., & Christensen, H. I. (2013). Efficient organized point cloud segmentation with connected components. Semantic Perception Mapping and Exploration (SPME).

[30] Unnikrishnan, R., & Hebert, M. (2003, October). Robust extraction of multiple structures from non-uniformly sampled data. In Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on (Vol. 2, pp. 1322-1329). IEEE.

[31] Rabbani, T., van den Heuvel, F., & Vosselmann, G. (2006). Segmentation of point clouds using smoothness constraint. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(5), 248-253.

[32] Triebel, R., Shin, J., & Siegwart, R. (2010, June). Segmentation and unsupervised part-based discovery of repetitive objects. In Robotics: Science and Systems (Vol. 2).

[33] Neiswanger, W., Wood, F., & Xing, E. (2014, August). The dependent dirichlet process mixture of objects for detection-free tracking and object modeling. In Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics (pp. 660-668).

[34] Cree, M. J., Jefferies, M. E., & Baker, J. T. Using 3D Visual Landmarks to Solve the Correspondence Problem in Simultaneous Localisation and Mapping.

[35] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2), 91-110.

[36] Lamon, P., Tapus, A., Glauser, E., Tomatis, N., & Siegwart, R. (2003, October). Environmental modeling with fingerprint sequences for topological global localization. In Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on (Vol. 4, pp. 3781-3786). IEEE.

[37] Neal, R. M. (2000). Markov chain sampling methods for Dirichlet process mixture models. Journal of computational and graphical statistics, 9(2), 249-265.

[38] Blei, D. M., & Jordan, M. I. (2006). Variational inference for Dirichlet process mixtures. Bayesian analysis, 1(1), 121-143.

[39] Montemerlo, M., Thrun, S., Koller, D., & Wegbreit, B. (2002). FastSLAM: A factored solution to the simultaneous localization and mapping problem. AAAI/IAAI, 593-598.

[40] Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2006). Hierarchical dirichlet processes. Journal of the american statistical association, 101(476).

[41] Doucet, A., De Freitas, N., & Gordon, N. (2001). An introduction to sequential Monte Carlo methods (pp. 3-14). Springer New York.

[42] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. the Journal of machine Learning research, 3, 993-1022.

[43] MacEachern, S. N. (2000). Dependent dirichlet processes. Unpublished manuscript, Department of Statistics, The Ohio State University.

[44] Fox, E. B., Sudderth, E. B., Jordan, M. I., & Willsky, A. S. (2011). A sticky HDP-HMM with application to speaker diarization. The Annals of Applied Statistics, 5(2A), 1020-1056.