Bulletin of Networking, Computing, Systems, and Software – www.bncss.org, ISSN 2186–5140
Volume 1, Number 1, pages 9–12, January 2012

Robot Mapping Using *k*-means Clustering Of Laser Range Sensor Data

Ankit A. Ravankar, Yohei Hoshino, Takanori Emaru, Yukinori Kobayashi
Division of Human Mechanical Systems and Design, Graduate School of Engineering
Hokkaido University
Sapporo, Japan

Abstract— **In this paper we discuss the map building technique for mobile robot using *k*-means clustering using laser ranger sensor. Clustering is an efficient technique to group together data set to obtain accurate maps for autonomous mobile robots. We discuss the *k*-means clustering algorithm in detail with experimental results and how this method can be used to obtain straight line maps for indoor environments. We discuss our results with different sizes of clusters and for data set with noise. Results confirm that k-means clustering can be used to obtain straight line maps for mobile robots.**

Keywords-Clustering, k-means algorithm, Robot Mapping, Mobile Robots.

I. INTRODUCTION

Simultaneous localization and mapping (SLAM) is a technique that is used in the field of autonomous mobile robots for building a map of an unknown environment or to update a map within a known environment, while at the same time keeping the track of its current location. Numbers of solutions to SLAM problem are proposed over the last decade [1]-[5], the most popular being the Kalman Filter based approach which is a linear estimate based technique [6]. However there are always errors associated with every model, as for each observation the error accumulates over a time giving inaccurate map or position estimate. Sensors are prone to errors and generate noise which adds to the complexity of the problem. So prior processing of the sensor data is essential to generate accurate maps.

Clustering is one of the techniques that is based on grouping together data that are similar. Its application are wide in the field of data mining and data statistical data analysis like machine learning, patter recognition, image analysis and bioinformatics.

Various clustering techniques have been implemented for SLAM like Iterative closest point (ICP) method for making clusters of data points obtained from the sensor and for data reduction [7].

For indoor mapping, we discuss how *k*-means clustering can be effectively used with advantages and limitations.

II. K-MEANS CLUSTERING

A. *k-means clustering*

k-means clustering is a method of cluster analysis which aims to partition a set of data points into *k* clusters in which each observation belongs to the cluster with the nearest mean [8],[9],[11]. Here *k* is a positive integer number. The grouping is done by minimizing the sum of squares of distances between data and the corresponding cluster centroid. Thus the purpose of *k*-means clustering is to classify the data. The objective function that *k*-means optimizes is

$$\text{kmeans(X,C)} \leftarrow \sum_{i=1}^{n} \min \left[x_i - c_j \right]^2 : \ j \in \{1.....k\} \tag{1}$$

B. *Algorithm*

This part describes the algorithm for *k*-means clustering. The input to the algorithm is a set of points generated by the laser distance sensor. The algorithm is very straightforward and works in following steps:

Step 1: Begin with a decision on the value of *k*= number of clusters.

Step 2: Put any initial partition that classifies the data into *k* clusters.

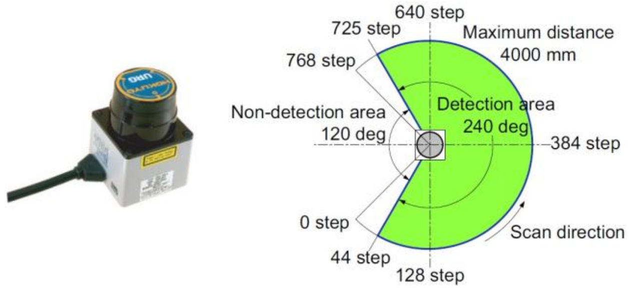
Step 3: Take each sample in sequence and compute its distance from the centroid of each of the clusters. If a sample is not currently in the cluster with the closest centroid, switch the sample to that cluster and update the centroid of the cluster gaining the new sample and the cluster losing the sample.

Step 4: Repeat Step 3 until convergence is achieved, that is until the data set generates no new centroid.

After generating the centroids for different clusters, we can join these clusters to obtain straight line maps. The number of clusters will greatly affect the resultant map.

III. SYSTEM SPECIFICATION

The scanning laser range sensor used during experiment is manufactured by Hokuyo Co. Ltd. This sensor was chosen for its accuracy, processing and reliability [10]. The model used is Hokuyo URG-04LX shown in Figure 1. The scan angle of the sensor is 240 degrees with angular resolution of 0.36 degrees. Range detection is 60mm-400mm with ± 2.5 %. The speed of scanning is 100 millisec/scan. Required power voltage is 5 VDC±5%. It uses a semiconductor laser diode of λ=785 nm and laser power less than 0.8 mW. The weight of the sensor is 160 g. The specifications of the sensor are summarized in table I.



(a) URG–04LX (b) Mesurement range

Figure 1. Laser range sensor (URG-04LX , Hokuyo Automatic Co.Ltd.)

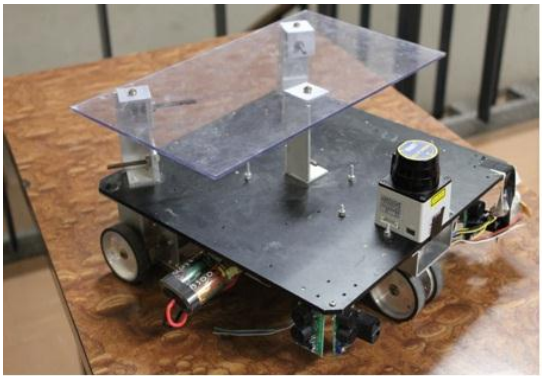


Figure 2. Robot(Plat F1, Japan System Design Co. Ltd)

TABLE1
LASER RANGE SENSOR SPECIFICATIONS

Laser range sensor	URG-04LX
Distance range	60 mm-4000 mm
Accuracy	20-1000 mm:±25mm
	1000-4000 mm: ±2.5%
Distance resolution	3.8 mm
Scan angle	240 degree
Angular resolution	1.875 degree

The algorithm was computed in Java language (version 1.6) on a 32 bit, Windows 7 operating system with 2 GB RAM.

Figure 2 shows the robot used during experiment (Plat F1, Japan System Design Co. Ltd) with URG-04LX laser sensor.

IV. EXPERIMENTAL RESULTS

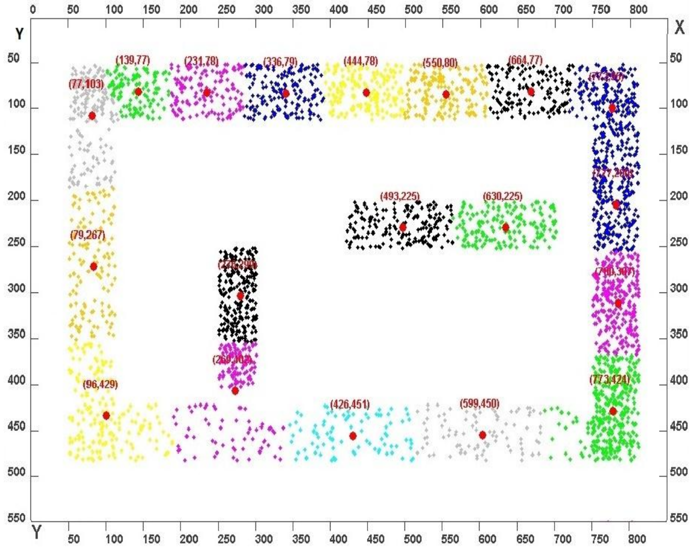


Figure 3. Result of k-means clustering with 20 clusters

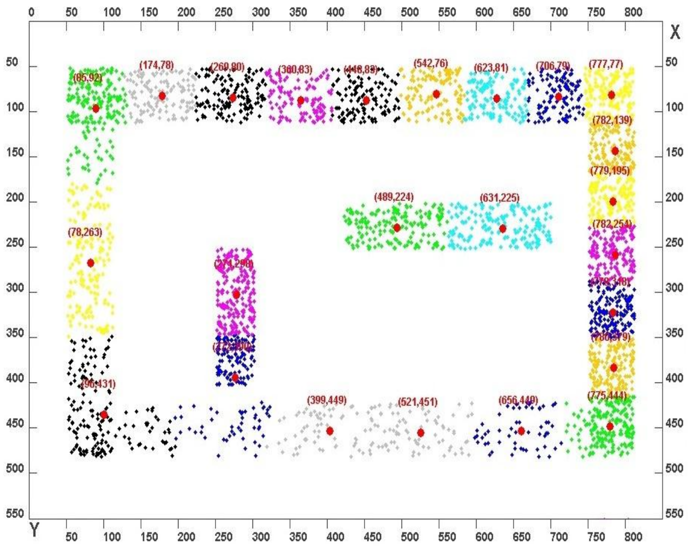


Figure 4. Result of k-means clustering with 25 clusters

The experimental environment built for the map generation consists of straight line walls. The robot with sensor is moved inside the area to generate data points. After applying *k*-means algorithm to the set of data points obtained from the laser range sensor, the set of points are divided into clusters with centroids. For obtaining the desired number of clusters, initial *k* centroids are defined and the clustering is performed. From the results in Figure 3 with 20 clusters, it shows how the entire sets of data points are divided into small clusters with a mean for every set, shown in the figure by red dot. For accurate mapping more number of clusters are desired so that straight lines can be drawn through the mean points. Figure 4 and Figure 5 shows results obtained for the number of clusters as 25 and 30 respectively. The initial number of clusters to be input plays an important role to obtain good results. Smaller the size of *k* will generate an undesirable result. The position or coordinates of the mean points are later extracted to join and get a line map of the environment.

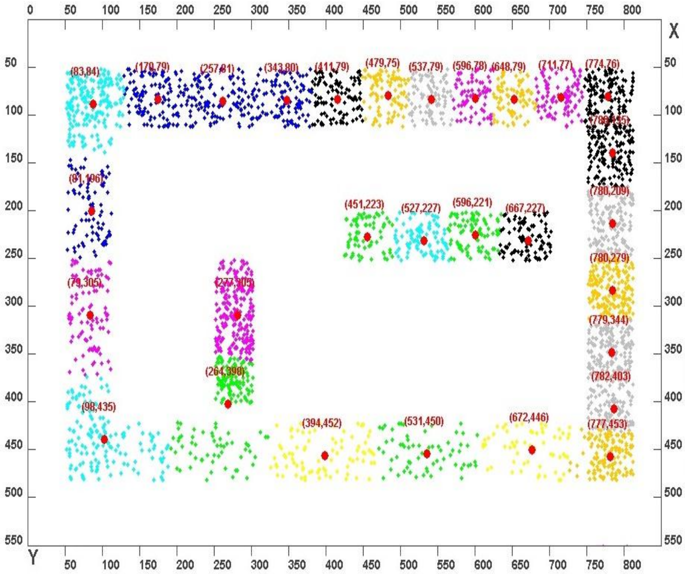


Figure 5. Result of k-means clustering with 30 clusters

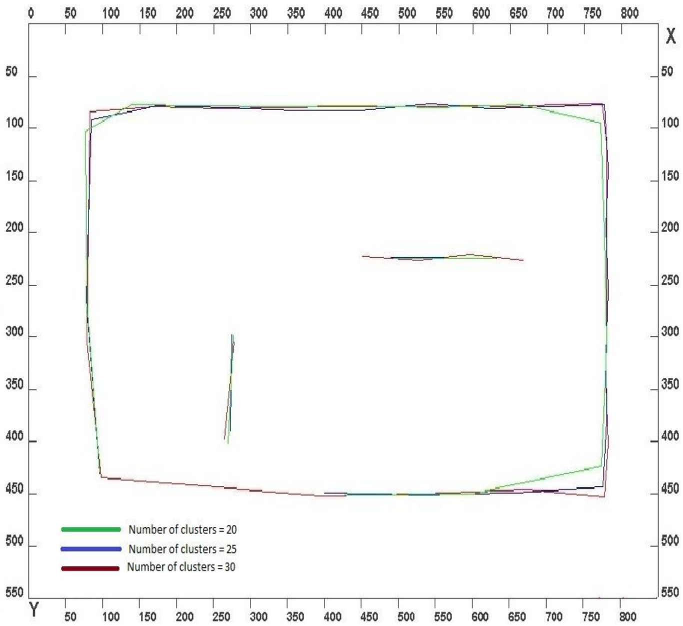


Figure 6. Final line map obtained for varying cluster sizes, different colours respresent different cluster sizes

Figure 6 shows superimposed result after joining the mean points for cluster size of 20, 25 and 30. It is evident that higher number of clusters produces good line maps compared to lower. It is to be noted that the mean points may not lie on the same line and hence will not always generate straight lines. This happens mainly because of noise present in the observation points. More advanced algorithms can be used like Hough Transform and Singular Value Decomposition to produce more accurate straight line maps [13], [14]. Also noise can be reduced prior to running the clustering algorithm.

\

V. LIMITATIONS

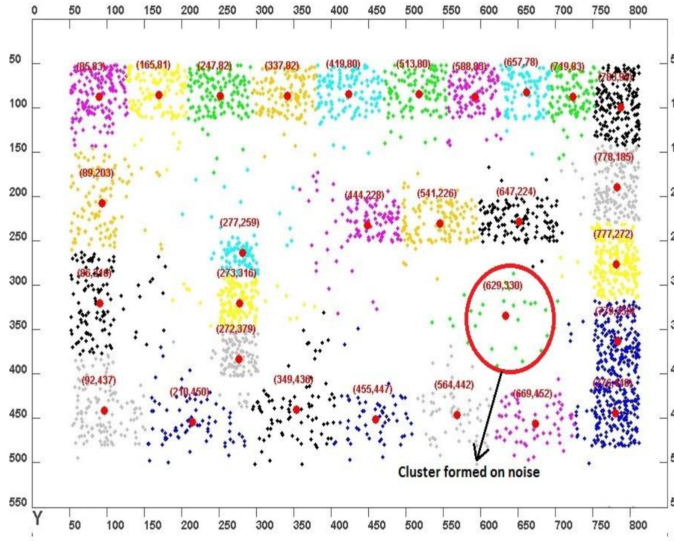


Figure 7. Result of k-means clustering on noise data(Red circled cluster represnts noise cluster with mean)

As discussed earlier, noises present in the sensor data can disorient the cluster formation and give errors. Figure 7 shows results obtained after applying *k*-means algorithm on sensor data with noise. The red circled cluster shows a noise cluster with its mean which is unwanted. Thus it is very important to reduce noise before applying the clustering technique. More efficient algorithms and filtering techniques can be one of the solutions to this problem [11], [12].

Also *k*-means algorithm requires user to input number of clusters to be formed. Thus predicting how many clusters can generate a good result is one of the limitations. Thus previous information about sensor and environment can result in effective clustering.

VI. CONCLUSION AND FUTURE WORKS

In this paper, we have discussed how clustering can be an efficient method to generate maps for mobile robots in an indoor environment. Using *k*-means algorithm, we showed that this method is very simple, powerful and efficient technique to generate straight line maps. Results obtained from the experimental results proved how the number and size of the clusters can be varied to generate different size clusters with mean points. We also showed how straight lines maps can thus be obtained from joining these mean points and discussed results for different number of clusters. As for the limitation, noise affects the clustering and hence noise reduction and data reduction is helpful in getting an accurate clustering.

The immediate future work would be to apply more efficient clustering algorithms and combining them to produce maps with higher accuracy. Also improving the computational efficiency of the algorithm and applying it in more complex environment would be the work in the future.

REFRENCES

[1] J.-G. Kang, S.-Y. An, S. Kim, and S.-Y. Oh, "A new approach to simultaneous localization and map building with learning: Neoslam (neuro-evolutionary optimizing)," in Computational Intelligence in Robotics and Automation (CIRA), 2009 IEEE International Symposium on, dec. 2009, pp. 278–284.

[2] J. Guivant and E. Nebot, "Optimization of the simultaneous localization and map-building algorithm for real-time implementation," Robotics and Automation, IEEE Transactions on, vol. 17, no. 3, pp.242–257, Jun 2001.

[3] S. Thrun, W. Burgard, D. Fox, H. Hexmoor, and M. Mataric, "A probabilistic approach to concurrent mapping and localization for mobile robots," in Machine Learning, 1998, pp. 29–53.

[4] R. Smith, M. Self, and P. Cheeseman, "A stochastic map for uncertain spatial relationships," in Proceedings of the 4th international symposium on Robotics Research. Cambridge, MA, USA: MIT Press, 1988, pp. 467–474. [Online]. Available:http://dl.acm.org/citation.cfm?id=57425.57472.

[5] M. Montemerlo and S. Thrun, "Simultaneous localization and mapping with unknown data association using fastslam," 2003, pp. 1985–1991.

[6] A. Rao, J. Mullane, W. Wijesoma, and N. Patrikalakis, "Slam with adaptive noise tuning for the marine environment," in OCEANS, 2011 IEEE - Spain, june 2011, pp. 1–6.

[7] W. Tingting, Z. Chongwei, B. Wei, and W. Mulan, "Mobile robot simultaneous localization and mapping in unstructured environments," in Informatics in Control, Automation and Robotics (CAR), 2010 2nd International Asia Conference on, vol. 1, march 2010, pp. 116 –120.

[8] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability, L. M. L. Cam and J. Neyman, Eds., vol. 1. University of California Press, 1967, pp. 281–297.

[9] S. P. Lloyd, "Least squares quantization in pcm," IEEE Transactions on Information Theory, vol. 28, pp. 129–137, 1982.

[10] S. Fu, H. ying Liu, L. fang Gao, and Y. xian Gai, "Slam for mobile robots using laser range finder and monocular vision," in Mechatronics and Machine Vision in Practice, 2007. M2VIP 2007. 14th International Conference on, dec. 2007, pp. 91–96.

[11] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," 2000.

[12] G. Hamerly and C. Elkan, "Alternatives to the k-means algorithm that find better clusterings," in Proceedings of the eleventh international conference on Information and knowledge management, ser. CIKM '02. New York, NY, USA:ACM,2002,pp.600–607.[Online].Available: http://doi.acm.org/10.1145/584792.584890

[13] G. A. Borges and M.-J. Aldon, "Line extraction in 2d range images for mobile robotics," J. Intell. Robotics Syst., vol. 40, pp. 267–297, July 2004. [Online]. Available: http://dx.doi.org/10.1023/B:JINT.0000038945.55712.65.

[14] C. Fern´andez, V. Moreno, B. Curto, and J. A. Vicente, "Clustering and line detection in laser range measurements," Robot. Auton. Syst., vol. 58, pp. 720–726, May 2010. [Online]. Available: http://dx.doi.org/10.1016/j.robot.2009.10.008