Unsupervised Learning for Subterranean Junction Recognition Based on 2D Point Cloud

Sina Sharif Mansouri, Farhad Pourkamali-Anaraki, Miguel Castao Arranz, Ali-akbar Agha-mohammadi, Joel Burdick, and George Nikolakopoulos

Abstract— This article proposes a novel unsupervised learning framework for detecting the number of tunnel junctions in subterranean environments based on acquired 2D point clouds. The implementation of the framework provides valuable information for high level mission planners to navigate an aerial platform in unknown areas or robot homing missions. The framework utilizes spectral clustering, which is capable of uncovering hidden structures from connected data points lying on non-linear manifolds. The spectral clustering algorithm computes a spectral embedding of the original 2D point cloud by utilizing the eigen decomposition of a matrix that is derived from the pairwise similarities of these points. We validate the developed framework using multiple data-sets, collected from multiple realistic simulations, as well as from real flights in underground environments, demonstrating the performance and merits of the proposed methodology.

1. INTRODUCTION

Recent technological advances in the field of robotics lead to the deployment of Micro Aerial Vehicles (MAVs) in challenging environments such as underground mine navigation [1], search and rescue mission [2], large infrastructure inspections [3], to name a few. MAVs have the potential to navigate rapidly in complex, unpredictable, and diverse subterranean environments for collecting various data types or saving lives.

Despite the broad spectrum of technological advances in the field of robotics, autonomous navigation in subterranean environments is still an ongoing quest. Subterranean environments face unique challenges, including irregular geological structures, complex geometry, lack of Global Positioning System (GPS) signal, the general absence of illumination, constrained passages, and unpredictable topology (see Figure 1 for illustration). Complex geometry and the presence of multiple junctions hinder the navigation mission. Junction detection is a critical task to enable safer and more optimal overall mission execution. For this task, a 2D Lidar equipped on the MAVs is used for sensing the environment. The high-end 3D lidars can provide 3D point cloud, and it can be used for junction detection. However, we are trying to remove the dependency on high end and highly expensive sensors, replacing them with lower cost solutions. Visual sensors cannot provide sufficient information regarding the complex geometry of subterranean environments due to lack of illumination. Furthermore, burdening a MAV with optical sensors in all directions and a bright light source will reduce the flight time, and require a large computational resource to post-process the visual feedback.

The 2D point cloud extracted from a 2D lidar is used in this paper to obtain important information about the number of junctions. The introduced method relies on a family of unsupervised learning techniques that compute a spectral embedding of the original data points by forming similarity graphs. To be more specific, spectral clustering algorithms typically consist of three steps [4]. The first step involves forming a similarity graph that encodes information about all pairwise similarities among data points in the given 2D point cloud. A popular approach for defining such similarities is to employ the Radial Basis Function (RBF) [5]. The next step is to compute the eigenvalue decomposition of the normalized Laplacian matrix that will provide a non-linear mapping of the original point cloud [6]. A key feature of our proposed framework is that the number of zero eigenvalues of the Laplacian reveals the number of connected components of the constructed similarity graph. The third step is to use the K-means clustering algorithm [7] to partition the 2D point cloud. Hence, our proposed method partitions the transformed 2D point cloud into connected components or clusters for identifying the number of junctions.

A. Background & Motivation

Navigation based on visual sensors or 2D/3D lidar has received significant attention in recent years in different application scenarios, e.g., [8], [9]. One of the essential capabilities for successful autonomous navigation in unknown and uncertain environments is environmental awareness, such as identifying environment geometries, obstacles, junctions, dead-ends, etc.

Several works have considered the use of visual sensors, lidar, and satellite images for detecting road intersections in uncertain environments [10]. For example, the previous work [11] developed an automatic road junction detection technique using airborne lidar data as well as the direction and width of road branches. Synthetic aperture radar systems are used for detecting L- and T- shape road junctions.
in [12] based on Markov optimization. The authors of [13] developed a machine learning technique based on the 3D point cloud generated by a laser rangefinder to detect road junctions. The introduced method was evaluated on two different data-sets from Germany and Brazil roads. The studied frames were classified into two groups of “junctions” or “roads”. In [14], junction detection is used to improve map localization in areas with weak or erroneous GPS signals such as forests or urban canyon. The proposed method detected crossroads using the lidar data by identifying free spaces between obstacles. However, most of these methods cannot be used in mines, as the map of such environments is not readily available, and existing methods are only suitable for certain shapes of junctions. Therefore, the prior work on using the lidar data for identifying subterranean junctions has demonstrated only limited success.

Towards using visual feedback for junction detection, in [15], a Convolutional Neural Network (CNN) binary classifier was developed for outdoor road junction detection. The method is experimentally evaluated on a commercially available MAV Bebop 2 from Parrot. In [16], a road intersection detection module has been proposed as binary classification, using the Long-Term Recurrent Convolutional Network (LRCN) architecture to identify relative changes in outdoor features and eventually detecting intersections. In [17], the authors proposed an architecture that combines CNN, Bidirectional LSTM [16] and Siamese [18] style distance function learning for junction recognition in videos. In this work, the authors evaluated their approach on different data-sets spanning various experimental scenarios. Towards junction detection in underground mines, in [19], the authors proposed transfer learning with AlexNet [20]. The method is limited to identifying only three different junctions and relied on the looking forward camera heavily. Moreover, these methods mostly consider binary classifiers, while in real-life scenarios more complex types of junctions exist and the junction recognition should recognize the different types of junctions. Additionally, most of these works have been evaluated and tuned in out-door environments and with proper illumination, i.e., using rich data about the surrounding of the platforms.

In this work, we propose to identify the number of junctions in uncertain subterranean environments using an unsupervised learning framework. Cluster analysis is one of the most fundamental problems in machine learning for finding groups of similar data points without any supervision [21]. Many existing clustering methods such as K-means clustering learn hidden structures from data points that are connected within convex boundaries. However, in this work, our goal is to develop a framework for identifying junctions without posing restrictive assumptions. Thus, we propose to use a family of unsupervised learning methods known as spectral clustering [22]. The superior performance of spectral clustering stems from the ability to exploit non-linear pairwise similarities between data points. As a result, spectral clustering has found applications in many domains, including image segmentation and genomic data analysis [23], [24].

### B. Contributions

Based on the state-of-the-art mentioned above, we list the main contributions of this work in the following.

The first contribution of this work is to develop a new framework for recognizing subterranean junctions using spectral clustering, which is a principled unsupervised learning technique. The proposed method allows us to uncover intrinsic structures from the 2D point cloud extracted from a lidar. Unlike previous techniques, the proposed method applies to various settings without requiring any prior knowledge on the number of junctions and their topology. Furthermore, our method does not require any pre-training and work well when working with small sample sizes. Hence, our introduced framework is suitable for a wide variety of settings with complex geometry.

In addition to our algorithmic contributions, the second contribution originates from a comprehensive evaluation of the proposed method in simulation environments with complex geometry and data-sets collected from the autonomous flight of the MAV in real underground mines in Sweden. As we will see, the obtained results demonstrate the effectiveness of our proposed framework on situational awareness of real-world subterranean environments.

The rest of the article is structured as follows. Section II provides a brief overview of detecting junctions in subterranean environments. Section III presents our proposed methodology for identifying the number of intersections or junctions. Then, in Section IV, we evaluate the proposed method on various simulated and real-world data-sets. Finally, Section V concludes this article by summarizing our findings and offering some future research directions to improve our framework further.
II. Problem Statement

This article considers lightweight aerial scouts, which are disposable low-cost robots with the task of fast exploration of unknown subterranean environments for collecting data and provide preliminary feedback for higher-level mission planners. These robots should be able to effectively navigate in such challenging environments based on the collected local information. It should be highlighted that most of the works investigating MAV navigation in subterranean environments consider high-end platforms [25]. In fact, these methods rely on a global map or are evaluated in tunnels without multiple branches [26]. However, uncertainties and drift over time in localization or map will result in wrong directions, and the underground tunnel has multiple branches and existing multiple branches [26]. Thus, our method provides navigation capabilities to detect the number of junctions in man-made tunnels based on the point cloud information. It should be highlighted that most of the works investigating MAV navigation in subterranean environments based on the collected local information. Such challenging environments based on the collected local information. In this case, the number of junctions in man-made tunnels based on the local point cloud information which is obtained from 2D lidars. Thus, our method provides navigation capabilities with resource-constrained solutions. The point cloud is in a body fixed frame of the MAV, and the MAV is centered of the point cloud.

![Image 1](image1.png)

Fig. 2. Top-view concept image of a subterranean environment with multiple branches, while body fixed frame of the MAV is depicted by $\mathbb{B}$.

III. Junction Detection

In this section, we discuss our proposed methodology for identifying junctions in underground environments by utilizing the acquired 2D point clouds, extracted from a 2D lidar. Let $X \in \mathbb{R}^{n \times 2}$ be a data matrix comprising of $n$ data points. Thus, each row of $X$ corresponds to one data point in $\mathbb{R}^2$ from the given 2D point cloud. Without any prior information, our goal is to identify the number of junctions, which is equal to the number of walls. In this paper we target the junction recognition with the use of the point cloud generated by only one revolution of the lidar. A wall will be registered as a set of data points with geographical proximity. The goal of this paper is therefore to utilize clustering techniques in order to identify the number of walls and their position. While one may attempt to use simple unsupervised learning techniques, such as $K$-means clustering, a challenge that we face in our problem of interest and it will be further analyzed in Section II, is to handle complex and non-linear structures. Hence, we propose to use spectral clustering techniques that will allow us to capture such complex structures.

In the following, we explain the three main steps of spectral clustering in detail. Given the data matrix $X$, we first construct a similarity graph. We denote the undirected similarity graph by $G = (V,E,W)$, where we have a set of $n$ nodes $V$ that are interconnected by a set of edges $E$. The matrix $W \in \mathbb{R}^{n \times n}$ is the adjacency matrix of this graph such that $W(i,j) \geq 0$ and $W$ is symmetric. A common approach to construct the similarity graph is to use the RBF as a measure of similarity. For some user-defined $\sigma > 0$, we compute the pairwise similarity between $x_i$ and $x_j$ (recall that $x_i$ refers to the $i$-th row of $X$):

$$W(i,j) = \exp(-\sigma \|x_i - x_j\|_2^2), \forall i,j \in \{1,\ldots,n\}. \quad (1)$$

The next task is to compute a spectral embedding of the original data points $x_1,\ldots,x_n$. Using the adjacency matrix $W$, we form the normalized Laplacian matrix $L \in \mathbb{R}^{n \times n}$ in the following form:

$$L = I - D^{-1/2}WD^{-1/2}, \quad (2)$$

where $I$ is the identity matrix and $D \in \mathbb{R}^{n \times n}$ is a diagonal degree matrix, i.e., $D(i,i) = \sum_{j=1}^{n} W(i,j)$. Hence, the matrix $D^{-1/2}$ can be efficiently formed by computing the inverse square root of the entries of $D$ on the main diagonal.

The eigenvalue decomposition of $L$ provides valuable insights about the structure of the similarity graph $G$. First, it is worth pointing out that every eigenvalue of the Laplacian matrix is non-negative, i.e., the matrix $L$ is Positive Semi-Definite (PSD). Thus, we can represent all $n$ eigenvalues of $L$ in non-decreasing order as follows:

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n. \quad (3)$$

Furthermore, the number of connected components in the graph $G$ is equal to the multiplicity of the 0 eigenvalue [27]. Hence, spectral clustering does not require any prior knowledge of the number of components.

Let $k$ be the estimated number of connected components using the above eigenvalue decomposition of the Laplacian. The subsequent task is to partition the original data points $x_1,\ldots,x_n$ into $k$ clusters. To this end, we compute the eigenvectors associated with the $k$ smallest eigenvalues of $L$. Next, we form a new matrix $U \in \mathbb{R}^{n \times k}$ by concatenating these $k$ eigenvectors column-wise. Hence, we can view the $i$-th row of $U$, i.e., $u_i$, as the spectral embedding of the original data point $x_i$. We can thus look for the connected data points in the transformed $k$-dimensional data-set $u_1,\ldots,u_n$ (instead of the original feature space $\mathbb{R}^2$). As mentioned before, this step is essential for many complex data-sets as the original data points may not be linearly separable [28].
The last step is to perform the $K$-means clustering algorithm on the non-linearly transformed data points $u_1, \ldots, u_n$. The corresponding optimization problem has the following form:

$$\min_{C} f(C, U) = \sum_{i=1}^{n} \min_{c \in C} \|u_i - c\|_2^2,$$

where $C$ represents a set of $k$ cluster centroids or representatives. Solving this optimization problem is known to be NP-hard and thus computationally intractable. However, there are approximation algorithms, such as the $K$-means++ algorithm, to solve this problem efficiently. As a result, we will find the partitioning of the embedded data points $u_1, \ldots, u_n$ into $k$ clusters. Note that this assignment can be easily translated to the original data points $x_1, \ldots, x_n$. The proposed algorithm is summarized in Algorithm 1.

Algorithm 1: Junction Detection via 2D Point Cloud

| Data: A set of $n$ data points $x_1, \ldots, x_n$ in $\mathbb{R}^2$ |
| Parameter: $\sigma > 0$ |
| Result: Number of junctions and clustering $X$ |
| 1 Form the adjacency matrix $W \in \mathbb{R}^{n \times n}$ using a non-linear similarity function, such that $W_{ij} = \exp(-\sigma \|x_i - x_j\|_2^2)$; |
| 2 Form the normalized Laplacian matrix $L = I - D^{-1/2}WD^{-1/2}$; |
| 3 Compute the eigenvalue decomposition of $L$ and sort the eigenvalues in non-decreasing order; |
| 4 Find the number of zero eigenvalues $k$; |
| 5 Perform $K$-means++ clustering on the top $k$ eigenvectors $u_1, \ldots, u_n$ in $\mathbb{R}^k$; |

IV. RESULTS

A. Simulation Evaluations

It should be highlighted that access to subterranean environments is typically limited. Thus, initially for evaluating the proposed method with complex geometry, random range measurements from 2D lidars are generated in the MATLAB [29]. The 2D lidar is considered with maximum range of 15 m and with 360 measurements for each revolution. The obtained point cloud from 2D lidar measurements are fed to the proposed method, the number of junctions are indicated, and the walls are clustered and color coded. All simulations have been performed on a single core on a computer with an Intel Core i7-6600U CPU, 2.6GHz and 8GB RAM and $\sigma = 1.5$, while the average computation time for all scenarios is 0.2s. Figure 3 depicts the obtained results, while the number of junctions are written above the figures, and the MAV is considered at the center of the point cloud. The proposed method successfully detects all junctions and walls when considering various types of complex geometry for the simulated 2D point clouds.

B. Experimental Data-sets

The data-sets have been collected from autonomous navigation of the MAV in two different underground mines, while the platform is equipped with a 2D lidar [1], [26]. The first location is an underground tunnel with a few branches, and the second one is an underground mine with multiple branches. Both environments are located in Sweden. The width and height of the first and second areas are approximately $3.5 \times 3$ and $6 \times 4$ m respectively.

Figure 4 depicts the irregular geometry of the first environment, while three specific areas are chosen to evaluate the proposed method, the first and second areas are tunnels with unstructured walls, and the third area has three junctions. The point cloud extracted from 2D lidar is fed to our proposed method and Figure 5 shows the obtained results. The method detects the number of junctions correctly, even though the point cloud extracted from experimental data-set is sparse compare to simulation results. Additionally, it should be highlighted that the tunnel without any branch considers two junctions as the platform has two options to navigate. In the future work, the obtained junctions should be matched to avoid counting same one multiple times and the proposed method should be evaluated in the closed loop with navigation and mission planner modules.

In the second case, the MAV autonomously navigates in an underground mine, however the width of the area is larger than the first case and the walls have wet surfaces. The areas with three and four junctions are chosen for evaluating the proposed method as depicted in Figure 6. Conditions such as the large width and wet wall surfaces result to in a sparse point cloud from 2D lidar range measurements. As it can be seen in Figure 7, our proposed method detects the number of junctions and walls correctly. Future work should investigate the development of robust methodologies, as dust can be misclassified as junctions. However, in our case, the areas we studied have negligible amount of dust.

V. CONCLUSION AND FUTURE WORK

This article proposed a novel method for the problem of junction detection in subterranean environments. The proposed method uses a non-linear pairwise similarity function to form a graph and find the number of connected components. The proposed method is computationally efficient and detects the number of junctions and walls with average computation time of 0.2s, while the detected walls can be used for wall following methods and number of junctions for local or high level planners. The method is successfully evaluated on multiple random generated 2D point clouds and extracted point clouds from autonomous navigation of the MAV in underground mines. Future works include matching the detected junctions for loop closure and to avoid counting same junctions multiple time, and evaluation of the proposed method in closed loop of navigation and planning components. The proposed unsupervised learning approach in this article will be also extended in the future work to handle 3D point clouds.
Fig. 3. Simulation results for identifying the number of junctions in different scenarios using spectral clustering. Each wall is indicated with a different color and the number of junctions is written above each figure (best viewed in color).

Fig. 4. The images of the MAV navigation in three different environments of an underground tunnel in Sweden.

Fig. 5. The detected walls and junctions from a 2D point cloud utilizing the proposed framework, with experimental data-set from autonomous MAV navigation in an underground tunnel, while each wall is indicated with a different color.

Fig. 6. The images of the MAV navigation in three different environments of an underground mine in Sweden.
Fig. 7. The detected walls and junctions from 2D point cloud using the proposed method, with experimental data-set from autonomous MAV navigation in an underground mine, while each wall is indicated with a different color.

REFERENCES

[1] S. S. Mansouri, P. Karvelis, C. Kanellakis, D. Komniak, and G. Nikokoulopoulos, “Vision-based mav navigation in underground mine using convolutional neural network,” in IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society, vol. 1. IEEE, 2019, pp. 750–755.
[2] T. Tomic, K. Schmid, P. Lutz, A. Domel, M. Kassecker, E. Mair, I. L. Grixa, F. Rues, M. Suppa, and D. Burschka, “Toward a fully autonomous uav: Research platform for indoor and outdoor urban search and rescue,” IEEE robotics & automation magazine, vol. 19, no. 3, pp. 46–56, 2012.
[3] S. S. Mansouri, C. Kanellakis, E. Fresk, D. Komniak, and G. Nikokoulopoulos, “Cooperative coverage path planning for visual inspection,” Control Engineering Practice, vol. 74, pp. 118–131, 2018.
[4] U. Von Luxburg, “A tutorial on spectral clustering,” Statistics and computing, vol. 17, no. 4, pp. 395–416, 2007.
[5] F. Pourkamali-Anaraki and S. Becker, “Improved fixed-rank Nyström approximation via QR decomposition: Practical and theoretical aspects,” Neurocomputing, vol. 363, pp. 261–272, 2019.
[6] N. Tremblay and A. Loukas, “Approximating spectral clustering via sampling: a review,” in Sampling Techniques for Supervised or Unsupervised Tasks, 2020, pp. 129–183.
[7] F. Pourkamali-Anaraki and S. Becker, “Preconditioned data sparsification for big data with applications to PCA and K-means,” IEEE Transactions on Information Theory, vol. 63, no. 5, pp. 2954–2974, 2017.
[8] C. Kanellakis and G. Nikolakopoulos, “Survey on computer vision for uavs: Current developments and trends,” Journal of Intelligent & Robotic Systems, pp. 1–28, 2017.
[9] T. D. Barfoot, C. McManus, S. Anderson, H. Dong, E. Beerepoot, C. H. Tong, P. Furgale, J. D. Gammell, and J. Enright, “Into darkness: Visual navigation based on a lidar-intensity-image pipeline,” in Robotics Research. Springer, 2016, pp. 487–504.
[10] D. Chaudhuri, N. Kushwaha, and A. Samal, “Semi-automated road in an underground mine, while each wall is indicated with a different color.

MATLAB, version 9.6. (R2019a). Natick, Massachusetts: The MathWorks Inc., 2019.