Place detection through analysis of descriptor sequences of 3D point clouds

Mihankhah, Ehsan

2018

Mihankhah, E. (2018). Place detection through analysis of descriptor sequences of 3D point clouds. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/73719>

<https://doi.org/10.32657/10356/73719>

PLACE DETECTION THROUGH ANALYSIS OF DESCRIPTOR SEQUENCES OF 3D POINT CLOUDS

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

EHSAN MIHANKHAH

School of Electrical and Electronic Engineering

A thesis submitted to the Nanyang Technological University in partial fulfilment of the
requirement for the degree of Doctor of Philosophy

2018

To my parents and my wife, Saeedeh, with endless love

Acknowledgments

I would like to sincerely thank my supervisor Professor Wang Danwei, for his aspiring guidance and valuable advice during all the stages of the research work in commendable patience. Besides technical lessons, I definitely learned a lot from his supportive and friendly attitude.

I also would like to present my sincere thanks to my family for their endless love and support throughout my life; my parents and my wife, Saeedeh, in special.

This research was partially supported by the ST Engineering – NTU Corporate Lab through the NRF corporate lab@university scheme. I would like to thank the due

Abstract

Place awareness is a critical component of safe decision-making in autonomous mobile robotics. The significance of place awareness is more appreciable in large-scale outdoor environment in which, safety measures, mission planning, and responsive intelligent behaviour depend on individual circumstances of each specific place. It is necessary for high-level mission planning and autonomous decision-making module to be aware of current place in order to invoke place-specific task objectives in a large-scale mission for long-term autonomy. In this regard, detection of transition between places (e.g. indoor to outdoor transition), or detection of specific scenarios (e.g. travelled from room to road via the main corridor), seriously assists the autonomous mission planning in complex environment.

The importance of place awareness in self-localization and mapping, as a fundamental requirement of autonomous mobile robots, inspired the development of a group of place detection techniques from this perspective. Recognition of a revisited place is the crucial requirement for cancelation of accumulated mapping drifts. Besides, identification of similar places contributes to selection of proper frame for merging the maps generated by multiple mobile robots. Moreover, map segmentation to meaningful places provides a more semantically informative model of the environment. Place recognition is also attended by developers of image retrieval algorithms. In this second approach, given an image of a place and a pool of unordered images from different places, image retrieval techniques identify best matching images from the pool to the image of interest, or alternatively, some more advanced methodologies classify unordered images into groups of related places. This approach for place recognition branches from the general image recognition and image classification research, which is dedicated to recognition of landmarks and invariant image features under changing environmental situation and viewpoint changes. There is also a third approach to place recognition that relies on external equipment, such as GPS, or relies on processing the available place-specific signals, such as WiFi access point signal, which has demonstrated acceptable performance under the assumption of perfect coverage. In this

research, place detection using on-board equipment is studied and the external equipment are used only as supplementary facilities. Thus, the third approach stays out of the attention of this thesis.

Through the first approach, researchers face major challenges to maintain metrically accurate map of the environment in large-scale outdoor environment. On the other hand, second approach confronts environmental changes, such as illumination, which highly affect feature selection and feature matching. Moreover, in the first approach, place recognition, which is supposed to assist the mapping module, depends itself on the mapping and localization module. Therefore, once mapping and localization loses accuracy for any reason, place recognition risks total failure. Second approach, which is free from metrical accuracy debate, confronts memory storage demand in long run, since these techniques need to store the feature models of different scenes, in form of a multi-dimensional vector structure, in a codebook for succeeding matching purpose. Moreover, the second approach corresponds a place to one snapshot, which should necessarily capture enough landmark features. It is important to note that, in general, a place might or might not contain a specific landmark, and therefore, it might or might not be representable through a single snapshot. Therefore, it is necessary to look for an alternative solution, which does not rely on metrical accuracy of environment model, and is not sensitive to feature-related concerns.

A new methodology for place detection is proposed in this thesis in order to cover the aforementioned concerns through a different approach to place detection problem. The proposed methodology identifies places as meaningful subsets of three dimensional environments, in which, every observation made at different locations of the place is consistently described by similar place descriptors. To compare one place to another, unlike the common approach of comparing two respective descriptor structures, two matrices consisted of descriptor similarities between different observations made from each place are compared. This novel comparison approach enhances the detection confidence. The descriptor used for place characterization is the list of eigenvalues of Laplace-Beltrami operator applied to the polygonised surface meshes. These meshes are generated from the three dimensional point cloud data captured by the sensory system. The mentioned descriptor is non-feature-based, and is applied to the entire captured input. Although the mentioned descriptor has been applied for object

identification before, mainly because limited perturbation in studied object results to limited perturbation in the object descriptor, to the best of our knowledge, through this thesis, this is the first time it is applied for characterization of places. Since the motivation for this research is to enhance place awareness in a long-term robot autonomy in large scale mixed outdoor and indoor environment exploration missions, it is assumed that the robot motion is continuous. This assumption makes possible to study the sequence of places and the pattern of descriptor changes in order to detect a revisited place with high confidence. Moreover, study of sequences enables recognition of revisited scenarios of motion as well. “Pattern of descriptor changes” is another new concept proposed through this research. Moreover, sensitivity of place detection to field of view limitation is addressed in this thesis, and a sensory system is suggested to achieve the desired outcomes.

The advantages of the proposed methodology include independence from metrical accuracy, elimination of feature selection and feature association challenges, storage efficiency, and applicability for large-scale outdoor scenarios in addition to insensitivity to environment changes. In GPS denied environment, these properties make the proposed methodology preferable for meaningful environment segmentation, identification of revisited places, and recognition of revisited scenarios, which form the fundamental requirements for identification of loop-closure frames in self-localization and mapping, identification of proper frames for merging multiple maps generated by multiple robots, autonomous start-up location estimation, and autonomous place-specific task assignment. Moreover, whenever GPS data is available, automatic labelling of visited places becomes possible. If the observations are time-stamped, it becomes possible for the detection system to deliver the accurate timing for recognition of revisited places. This facilitates the analysis of other sensory inputs captured at the identified time for any further possible mission-specific analysis.

CONTENTS

Acknowledgments	i
Abstract	iii
List of Figures	xi
List of Tables	xiii
List of Acronyms	xv
1 Introduction	1
1.1 Background	1
1.2 Related work	2
1.2.1 Defining the term “Place”	2
1.2.2 Place detection and field of view limitation	4
1.2.3 Place detection equipment	4
1.3 Motivation	12
1.4 Proposed place detection approach	14
1.4.1 Detection procedure	14
1.4.2 Input acquisition	14
1.4.3 Place descriptor for characterization	15
1.4.4 Environment segmentation	15
1.4.5 Similarity between places	16
1.4.6 Identification of similar scenarios	16
1.4.7 Identification of revisited places	16
1.4.8 Place detection and place recognition	16
1.5 Contributions	17
1.5.1 A novel approach to place detection	17
1.5.2 Place characterization with Laplace eigenvalues	17
1.5.3 Meaningful environment segmentation	18
1.5.4 Recognition of revisited places through recognition of revisited scenarios	18
1.6 Outline of the thesis	19
2 Place Descriptor	21
2.1 Classification of descriptors	21
2.1.1 Features	21
2.1.2 Appearance	23
2.1.3 Domains	23
2.2 Proposed descriptor	24
2.3 Uniformly dense digitization across the entire usable field of view	25
2.3.1 Existing place detection sensors	25
2.3.2 Sensory system setup	27

2.4	Descriptor derivation procedure	28
2.4.1	Input data preparation	29
2.4.2	Mesh generation	30
2.4.3	Eigenvalues of Laplace-Beltrami operator	31
2.5	Properties of DNA descriptor	33
2.5.1	Resolution	34
2.5.2	Noise and occlusion	35
2.5.3	DNA of sub-parts	39
2.5.4	Rotation and translation	42
2.5.5	Connectivity, change of boundary	42
2.6	Chapter conclusion	46
3	Place Characterization	47
3.1	Uniqueness of DNA	48
3.1.1	Comparison between surface DNAs	49
3.1.2	Comparison between a surface DNA and a place DNA	52
3.1.3	Comparison between place DNAs	53
3.2	Percentage of matching eigenvalues	56
3.3	Similarity measure	57
3.3.1	Common comparison techniques	58
3.3.2	Proposed comparison technique	58
3.4	DNA Length	61
3.5	General parameters	61
3.6	Consistency of DNA	61
3.7	Place characterization through DNA analysis	64
3.8	Deformation of DNA along the continuous motion	66
3.9	Second characterization case study	67
3.10	Third characterization case study	68
3.11	Chapter conclusion	70
4	Environment Segmentation	71
4.1	Distinctive subsets vs gradual changes	71
4.2	Marking a place in table of similarities	73
4.3	Segmentation case study	75
4.4	Segmentation procedure	79
4.4.1	Sensitivity to similarity threshold	79
4.4.2	Scan labels	79
4.5	Chapter summary	80
5	Detection of Similar and Identical Places through Sequence Analysis	81
5.1	Cross similarity	82
5.2	Identification of similar places; a case study	83
5.3	Sequences	85
5.3.1	Proper sequences	87
5.4	Sequence analysis	89
5.4.1	Sequence pattern matching	89
5.5	Discussion on sequence matching	90
5.5.1	Selection of templates and images	90
5.5.2	Detection confidence	90
5.5.3	Other template matching techniques	91

5.5.4	Reverse sequences	91
5.6	Best matching frames	91
5.6.1	Diagonal elements of cross similarity place matrices	92
5.7	Second sequence analysis case study	94
5.8	Procedure for recognition of revisited places and revisited scenarios	102
5.9	Chapter conclusion	102
6	Comparison with Existing Methodologies	105
6.1	Definition of a place	106
6.2	Sensing	106
6.3	Description	107
6.4	Comparison, storage, and recall	107
6.5	Identification of similar places and identical places through sequence analysis	108
7	Conclusion and Future Work	111
7.1	Conclusion	114
7.1.1	Advantages of the presented methodology	114
7.1.2	Applications	115
7.1.3	Trade-offs to be considered	116
7.2	Future directions	118
	Author's Publications	121
	Author's Previous Publications and Contributed Papers	123
	References	125

List of Figures

Fig. 1-1 – Left: Common place detection procedure; Right: Proposed place detection procedure.....	15
Fig. 2-1 – Sensory system.....	28
Fig. 2-2 – Sample 3D point clouds generated by the proposed sensory system	28
Fig. 2-3 – Descriptor derivation procedure.....	29
Fig. 2-4 – Left: Raw 3D Point Cloud; Right: Resultant polygonised mesh.....	31
Fig. 3-1 – An example of iso-spectral surfaces introduced by [150]	48
Fig. 3-2 – Illustration of the DNAs of similar shapes and different shapes	51
Fig. 3-3 – An ideal place containing five different surfaces; five different observations are made from the studied place.....	55
Fig. 3-4 – Sensor range and distance travelled; intersection of two spheres is the common volume visible from both view-points.	57
Fig. 3-5 – Pavement; this place is also studied as place 3 in 3.7	62
Fig. 3-6 – DNA plots of four observation depicted in Fig. 3-6. Each DNA is consisted of smallest 1500 non-zero eigenvalues.....	62
Fig. 3-7 – Four different observations from the same place. Observation points are few meters distant from each other.	63
Fig. 3-8 – Studied place 1; roofed drop off zone	64
Fig. 3-9 – Studied place 2; between residential buildings	64
Fig. 3-10 – DNAs of observations made from studied places 1, 2, and 3.....	66
Fig. 4-1 – When similarity threshold is a larger number, transition frames are excluded from place matrices.....	74
Fig. 4-2 – When similarity threshold is a smaller number, transition frames are included in place matrices.....	74
Fig. 4-3 – When similarity value is not too large or small, some of the transition frames are included in place matrices.	74
Fig. 4-4 – The path travelled by the robot in segmentation experiment	75
Fig. 4-5 – Starting point and a snapshot from the environment where segmentation experiment was conducted	75
Fig. 4-6 – Eight place matrices matched with the corresponding places on the map	77
Fig. 4-7 – Procedure for environment segmentation to meaningful places	80

Fig. 5-1 – Place matrices 1, 2, 3, and 4, along with three of the six possible cross similarity matrices $C_{(1)(3)}$, $C_{(1)(4)}$, and $C_{(3)(4)}$	83
Fig. 5-2 – Experiment for identification of similar and identical places.....	84
Fig. 5-3 – Place matrices 1, 2, 3, and 4, along with four sequences of place matrices $S_{(1)(2)}$, $S_{(2)(3)}$, $S_{(3)(4)}$, and $S_{(1)(2)(3)}$	87
Fig. 5-4 – Second sequence analysis case study route	94
Fig. 5-5 – Procedure for recognition of revisited places and revisited scenarios.....	103
Fig. 7-1 - Place detection procedure	112
Fig. 7-2 – Place detection procedure (continued)	113

List of Tables

Table 2-1 – Differences between shape DNA and place DNA.....	25
Table 2-2 – Exact and numeric approximation of the eigenvalues of flat rectangle.....	35
Table 2-3 – Initial eigenvalues of rectangle compared to the eigenvalues of pyramid.....	36
Table 2-4 – Initial eigenvalues of original and noisy rectangles	37
Table 2-5 – Initial non-zero eigenvalues of original rectangle and noisy versions.....	37
Table 2-6 – Initial eigenvalues of original, noisy, and occluded rectangles	38
Table 2-7 – Initial eigenvalues of pyramid and duplicated pyramid	40
Table 2-8 – Initial non-zero eigenvalues of ideal elements; eigenvalues less than 20 are listed.	41
Table 2-9 – Initial non-zero eigenvalues of an ideal place; eigenvalues less than 20 are listed (colour code and labels provided for easier comparison with Table 2-8)	41
Table 2-10 – Eigenvalues of rotated objects.....	42
Table 2-11 – Eigenvalues of translated objects	43
Table 2-12 – Initial eigenvalues of single pyramid, connected double pyramid, and disconnected double pyramid.....	44
Table 2-13 – Initial eigenvalues of single pyramid, single-side-connected double pyramid, and all-sides-connected double pyramid.....	45
Table 3-1 – DNA of six different surfaces.....	50
Table 3-2 – Surface DNA comparison result; inflated DNAs are compared.....	51
Table 3-3 – DNA of an ideal place compared to the DNA of Rock as a sub-part.	54
Table 3-4 – DNA comparison of five different observations from an ideal place	55
Table 3-5 – Comparison between ideal place DNA and rectangle DNA.....	60
Table 3-6 – Comparison between ideal place DNA and noisy rectangle DNA.....	60
Table 3-7 – General experiment parameters	61
Table 3-8 – Similarity among DNAs of different place observations shown in Fig. 3-6.	64
Table 3-9 – Multiple observations from three different places.....	65
Table 3-10 – Similarity of observations made from three different places.....	67
Table 3-11 – Similarity of observations made from LT 26 and LT 28.....	68
Table 3-12 – Panoramic photos of LT 26 and LT 28.....	68
Table 3-13 – Similarity of observations made from LT 24 and LT 28.....	69
Table 3-14 – Panoramic photos of LT 24 and LT 28.....	69

Table 4-1 – Pattern of table of similarities in presence of overlapping sub-parts.....	72
Table 4-2 – Table of similarities of the segmentation experiment; 53 scans were captured along the path	76
Table 4-3 – Table of similarities of segmentation experiment after merging highly overlapped place matrices, 8 place matrices are identifiable in the table.	77
Table 4-4 – Place matrices and the corresponding places on the map.....	78
Table 5-1 – Table of similarities for the experiment that includes revisited places.....	85
Table 5-2 – Table of similarities with sequence matrices and a cross sequence matrix.	88
Table 5-3 – Cross sequence matrix $CS_{(4)(9)-(5)(10)}$ is used to identify best matching snapshot...	92
Table 5-4 – Matching snapshots	93
Table 5-5 – Parameters of second sequence analysis case study	95
Table 5-6 – Environment segmentation result in second sequence analysis case study	96
Table 5-7 – Cross similarity matrices marked in second sequence analysis case study	97
Table 5-8 – Sequence matrices and cross sequence matrices marked for second sequence analysis case study	98
Table 5-9 – Pattern comparison of sequences in second sequence analysis case study.....	99
Table 5-10 – Revisited scenario matrix zoomed (second sequence analysis case study)	99
Table 5-11 – Best matching frames in second sequence analysis case study	100
Table 5-12 – Photos from places studied in second sequence analysis case study	101
Table 6-1 – Comparison between the proposed method and conventional methods	108
Table 6-2 – Comparison between the proposed method and conventional methods (Continued I).....	109
Table 6-3 – Comparison between the proposed method and conventional methods (Continued II).....	110

List of Acronyms

2D: Two-Dimensional

3D: Three-Dimensional

ARPACK: ARnoldi PACKAge

BRIEF: Binary Robust Independent Elementary Features

CAD: Computer Aided Design

doG: difference-of-Gaussian

GLOH: Gradient Location-Orientation Histogram

GNSS: Global Navigation Satellite System

GPS: Global Position System

IMU: Inertial Measurement Unit

LiDAR: Light Detection and Ranging

LoG: Laplacian-of-Gaussian

PCA: Principle Component Analysis

PCL: Point Cloud Library

RADAR: RAdio Detection And Ranging

ROS: Robot Operating System

SIFT: Scale Invariant Feature Transform

SLAM: Simultaneous Localization And Mapping

SURE: Surface Entropy

SURF: Speeded-Up Robust Features Descriptor

Chapter 1

Introduction

1.1 Background

Localization in mobile robotics has been an active research area for many years. Different methodologies for self-localization have been introduced in the literature. Each methodology is specific to certain mission objective. The most well-known methodology for indoor localization of mobile robots is SLAM¹ [1]. SLAM requires no prior information about the environment. Localization engine takes platform's initial pose as the reference and then estimates the relative position using one of the numerous techniques of SLAM [2-4]. SLAM logic is based on exploring and registering the current observation at the latest estimated location (creating the map), which in a loop helps making the estimation for the next updated location of the robot. The estimated location is occluded with some level of uncertainty, which can propagate through time and result in drifted map and incorrect estimated location in long run, especially when robot never observes an evidence of a revisited location to close the loop and cancel the drifts. The efficiency of SLAM methodology has been tested, and different techniques have been compared, countless number of times in the literature [5]. The problem is still an open research topic for large-scale outdoor missions and long-term autonomy.

On the other hand, for outdoor localization, GPS², which provides absolute position, has been proven to be a reliable localization solution for road navigation missions [6]. However, GPS loses accuracy when the receiver module is not fully covered by the satellites, and this makes GPS-based localization limited to open outdoor environments, which are not affected by large buildings or trees or other sources of occlusion.

¹ - Simultaneous Localization And Mapping

² - Global Position System

In our past mobile robotic projects, whenever the mission was to explore a large outdoor on and off-road area, assuming the area had been accessible prior to the main mission, given the task to find certain target objects inside and outside potential buildings, it was sufficient to hard code GPS points in order to navigate to the potential buildings. It was also possible to explore inside of the buildings using SLAM. However, it was still necessary to identify different places, and it was also necessary to identify transition between places especially in GPS denied regions. This objective could not be fully handled through SLAM in outdoor environment due to aforementioned limitations. This has been the motivation for starting the place detection research as an individual module to support mission planning in large-scale environment consisted of both indoor and outdoor places in a long-term multi robot exploration mission, especially in GPS denied condition.

1.2 Related work

1.2.1 Defining the term “Place”

The objective of place detection necessitates different definition for places. Place detection methodologies, inspired by way-point-navigation and exploration research, describe the environment as nodes and links. Robot objective is to travel from one node to another via a link. In this approach nodes are taken as zero dimensional places [7] and links are routes between places. Discordantly, in [8], where environment semantics are studied, place is defined as “the abstraction of a region” and a region is a “two-dimensional subset of the environment”. Place detection methodologies inspired object detection, resulted to a new topic in localization called SLAM++ [9]. According to this technique, a place can be identified through the identification of objects. From a different perspective, according to [10], gateways cluster the environment to different places. A gateway between places occurs when appearance configuration changes significantly. With similar assumptions, work [11] used image sequencing partitioning to form a topological graph node of places. Although different objectives result to different definitions for places, only few major place detection approaches are identifiable in the literature. The place definition used in this thesis is proposed in section 1.3.

1.2.1.1 Definition of places through boundaries

Some place detection methods such as [12] are based on metrically accurate mapping techniques, define places as a space confined by mathematically defined boundaries. In this case, the simplified boundaries are in form of bounding boxes [13] or bounding circles, spheres, or ellipsoids [14]. Since places are not necessarily square-shaped or spherical-shaped, these segmentation techniques cannot include places with complex shapes, or at least they cannot accurately define complex-shaped places. A space confined by the intersecting planes or curved surfaces could be a more accurate definition. However, calculation of several such surfaces becomes tedious and computationally expensive when the studied place is in shape of a complex volume. In this case, the most accurate and efficient calculation of place boundaries could be achieved by polygonization of the surrounding map points of a specific volume (place) [15]. However, this definition requires substantial amount of memory to store the boundary definition for each place. Therefore, although this technique provides very accurate definition for a place, it is not practically possible to be used for large-scale outdoor applications.

1.2.1.2 Definition of places through landmarks

Other place detection methodologies such as [16, 17] are based on feature extraction and matching. They identify places by recognizing features of major landmarks in a snapshot of a place [18]. In this case, major landmarks are detected followed by feature matching, and places are identified based on detected landmarks. Place is mapped directly to major landmarks and indirectly mapped to specific set of features. Therefore, in this approach, a place is defined by a set of landmarks or by a set of features. There is no dispute that landmark detection can be handled through feature processing. However, although some specific places contain unique landmarks, in general, a place may or may not contain a major landmark.

Moreover, a place might have any area and shape. Therefore, a place might or might not be entirely visible through a single snapshot. Therefore, definition of a place through landmark processing also has the shortcoming to include large places, or places which cannot be well-represented by a single frame due to their shapes.

1.2.2 Place detection and field of view limitation

An important issue, which highly affects place detection through image processing, is the field of view limitation. This is more related to sensing rather than processing. For example, to identify a place through landmark processing, it is required for the specific landmark, which identifies the place, to be sufficiently captured by the sensing device [19, 20]. Otherwise, even though the observation is made in the same location, since the landmark is not visible and hence not detectable, place cannot be identified. A place does not have any specific direction [8]. Therefore, place definitions influenced by presumed directional observations have shortcoming to fulfil the logical meaning of a place.

1.2.3 Place detection equipment

Place detection overlaps with geo-position sensing [21], SLAM [1], and scene detection [22]. Each topic studies places with different assumptions and through different approaches. In geo-position sensing research geographical positions are readily available as absolute pose of the receiver in the studied environment and are matched to prior knowledge about the places at certain locations. On the other hand, SLAM researchers see places as additional semantic information that assist the loop closure and map merging. Scene detection researchers attach places to landmarks and appearance-based features [23-26] of the studied image. First approach incorporates external equipment such as GPS, beacons, and other similar facilities. The other two approaches use on-board equipment such as cameras, LiDARs¹, IMUs², and other sensors as the primary means for place detection. Therefore, place detection is divided to two main categories from this perspective.

1.2.3.1 Place detection through external equipment

In this approach, for small scale and indoor applications, beacons are used to identify places[27]. Beacons can identify places with an acceptable accuracy. In indoor environment where it is possible to install beacons, place detection problem can be

¹ - Light Detection and Ranging

² - Inertial Measurement Unit

handled through this technique. Scalability, coverage, and overlap are the main concerns of this technique.

For outdoor and large scale environment such as roads, GPS (or GNSS¹) can be used instead [28]. Since GPS depends on satellites for position estimation, the receiver equipment mounted on the moving vehicle is the only facility required to accomplish positioning requirements. Therefore, the technique is highly scalable and the accuracy of advanced models are perfect in ideal satellite coverage situation [29]. However, whenever the receiver equipment is not covered by satellites, e.g. the area close to the buildings and trees, positioning through this technique becomes inaccurate [30, 31]. The method is not, generally, usable in indoor environment since no satellite coverage is expected in covered environment. Another concern for incorporation of methods relying on external equipment is the high vulnerability to external attacks [32].

GPS and beacons are not the only external equipment used for identification of places. Other available signals in the environment can also be used for identification purpose. Fingerprint-based location identification [33] is a research topic which studies sound signatures, motion patterns of people and environment elements, availability and signal strength of WiFi access points, and other place-specific factors.

1.2.3.2 Place detection through on-board equipment

On-board sensory system is referred to the sensors that do not depend on external facilities and equipment. Independence from external infrastructure eliminates the coverage concern for the mobile platform. Since, in this research, place detection is studied in mixed indoor and outdoor environments, independence from the external equipment and infrastructure becomes an indispensable necessity. Therefore, in this research, only on-board equipment are used for detection purpose.

Despite the mentioned advantage, independence from external equipment introduces a major concern in the analysis of sensor readings. Perception made through analysis of on-board sensor readings is made in the robot coordinate frame. In such analysis, any attempt to integrate the input data, results to accumulation of uncertainties, which stand out in form of drift in the long run. Therefore, in the analysis of sensor readings in robot

¹ - Global Navigation Satellite System

coordinate frame, either a drift compensation mechanism should be considered, or the analysis should not involve integration over sensor readings. In the proposed methodology, which targets large-scale environment for long-term autonomy, accumulation of consequent sensor reading is avoided to address this major concern.

Although it is possible to use many on-board equipment such as IMU, wheel encoder, compass, camera, LiDAR, RADAR¹, and etc., for relative position estimation, identification of places through on-board equipment is commonly achieved through visual place detection techniques [18]. This means, cameras and LiDARs are used as the primary sensors and other sensors serve as supplementary equipment to overcome the possible limitations. In visual place detection visual fingerprints are studied [18].

Visual place detection incorporates different image processing techniques for identification. Images can be in form of 2D camera frames (colour intensity information) [34], 3D-range point cloud data (depth information) [35-37], or RGBD frames (both colour intensity and depth information) [38, 39]. Visual place detection is a well-attended research. Regardless of problem formulation, assumptions, and processing techniques, all methodologies follow five general steps of input acquisition, data preparation, description, comparison, followed by the final step, which is to apply comparison results in order to achieve the detection objectives.

Although the proposed place detection methodology is solely based on analysis of on-board equipment, if mapping the local perception to global frame would be required, it is possible to use external sensors and equipment, GNSS in especial, as supplementary equipment.

1.2.3.2.1 Input acquisition

2D cameras are the simplest and the least expensive solution for visual input acquisition[40]. Although focal point adjustment, illumination effect, distortion, and small field-of-view are major concerns when basic models are used, the low price of the device is still a strong motivation, which contributes to facilitating the commercialization of the research. Advanced cameras have fewer limitations. Auto-focus function is added to most modern cameras through different engineering

¹ - RAdio Detection And Ranging

techniques. Distortion-free lenses can be added to cameras too. Panoramic view cameras and spherical view cameras provide extended field-of-view. However, including all advanced functions in a camera increases the price of the sensor.

3D cameras add depth information to pixels [16]. When depth information is available, more advanced and more precise processing techniques can be applied to the input information. Depth information is usually added either through stereo vision techniques [41] or through time of flight measurement techniques [42].

LiDAR is another widely used sensor for visual place recognition that generates binary point cloud images. Depth information generated through basic LiDARs is generally more precise and consistent, even in long ranges, compared to the depth information estimated by cameras. LiDARs are less affected by environmental conditions such as illumination. LiDARs are costlier compared to cameras. Most commercial LiDARs are made for collision avoidance applications. The commercial LiDAR models, which can produce a proper point cloud for identification purpose, are limited and the few existing models are costly.

While methodologies that use cameras as the primary sensor are generally oriented towards colour analysis, LiDAR-based methodologies [36] are more towards geometric analysis.

Challenges such as distortion compensation, uniform density of input data, field-of-view limitations, and sensitivity to environmental conditions, can be handled at sensor level. However, these issues are sometimes addressed at processing level as well.

1.2.3.2.2 Data Preparation

Useful information should be extracted from the raw input data. This step includes outlier removal, resolution adjustments, colour normalization, geometry normalization, segmentation, feature extraction, sensor fusion, and any other processes required to generate the set of information which is believed to contain the intrinsic information about the input image.

In case of camera frames, vibration compensation, illumination compensation, distortion compensation, down sampling, selection of image patches, and feature extraction are the steps commonly followed. The ultimate goal is to identify 2D or 3D

key-features. It is also possible to map 2D features in 3D [43], or to map 3D features in 2D [44].

Feature selection is a challenging topic with heated debate. A good feature should be repeatedly identifiable and should be invariant against scaling, rotation, shifting, deformations, and noise [45]. The feature should not be only specific to a limited category of images. In short, a good feature should be consistent, invariant, and general. Several techniques are available for feature detection. Simplest visual features are edges [46] and corners [47]. These features are used as the primary interest points in complex feature detectors. The Hessian feature detector finds pixels that exhibit high derivative values in two orthogonal directions [48]. The detector computes the Hessian for each pixel. Feature point is the point at which the determinant of the Hessian becomes maximum. Harris detector [47] is another detector which finds geometric features. The Harris detector searches for image regions where the second moment matrix has large eigenvalues. The resulting points correspond to corner-like structures. LoG¹ [49], and a close approximation of it, DoG² [50], are also well-known feature detectors widely used in detection algorithms. However, there are other techniques that combine these two with other measures. For example, Harris-Laplacian detector [51] combines the Harris corner detection with LoG. The resulting feature points are reported to be fewer compared to LoG and DoG but the selected points are more resilient against rotation, scaling, illumination changes, and input noise [52]. If the Harris-Laplace is applied to Hessian, the detector is called Hessian-Laplace [53], which produces slightly higher number of selected points but with similar quality compared to original Harris-Laplacian.

In case of binary point cloud data, noise removal, down sampling, segmentation, normal vector assignment, and feature extraction are the general data preparation steps. In one common approach, feature extraction for 3D point clouds is followed by creating polygonised meshes [54]. In [54], features are extracted through second difference operator applied to mesh edges. Similarly, the work [55] extracts feature lines inside meshes by applying third order derivatives to the surface mesh. Curvature extremums

¹ - Laplacian-of-Gaussian

² - difference-of-Gaussian

are other possible features which can be extracted from polygonised meshes [56]. Meshless feature extraction is another direction in feature selection. The work [57] uses region growing method. It starts by segmenting the point cloud into clusters. Then regions of sharp features are identified. This is the initial step for reconstruction of feature lines. Sharp features were selected through other techniques too [58, 59]. Principle component Analysis (PCA) is one of the efficient techniques to extract salient features from point clouds. Extracted features are vectors along the principle axes in the space which encapsulate statistic summary about the point cloud being studied [60]. PCA determines the axis along which the studied data points have the highest variance (the eigenvector associated with the largest eigenvalue) and lowest variance (the eigenvector associated with the smallest eigenvalue). This method can be applied to multi-dimensional data. The 3D version is used in computer vision for segmentation [61], as well as for feature detection for the purpose of object classification[62] and place recognition[63]. It should be noted that the eigenvalues studied by PCA are different from the ones used in proposed descriptor of this thesis. Other 3D point cloud feature extraction methods are covered in [64].

Depending on the methodology, some of the above mentioned preparation steps might be skipped, or some other steps might be added.

1.2.3.2.3 Description

Image descriptor is the model or data structure that describes an image. Image descriptor is a set of information, usually in form of a numeric structure, which encapsulates intrinsic information about the studied image. When descriptor operates on image features, image descriptor is sometimes referred to as feature descriptor. However, for the case of holistic description [65] or non-feature-based description [66] of images, it is logical to use the general term “image descriptor”. In this thesis, “place descriptor” or “descriptor” is used as the general term to refer to all types of data structures that encapsulate desired description of studied places. A good descriptor should be robust against noise, should be unique, and should be computationally efficient.

Literature review reveals that majority of detection techniques, including place detection techniques [67], use the SIFT¹ descriptor [22] or a modified version of it. In SIFT algorithm, local extrema of difference-of-Gaussian operator is used for selection of key-points. Feature vectors are extracted for key-points. SIFT algorithm calculates the gradient magnitudes and their orientations in a small region around each key-point. Then, a set of orientation histograms is created for sub-regions. In this case, the descriptor is a vector that contains values of all orientation histograms for each key-point. GLOH² is an extension of the SIFT descriptor that is designed to enhance robustness and distinctiveness of description [68]. SURF³ descriptor [69] is an alternative to SIFT. It is reported to be computationally more efficient than SIFT [52]. SURF descriptor uses “Haar wavelets” [70] instead of Gaussian derivatives. The SURF algorithm similarly operates on feature regions but instead of generating intensity histograms, it generates a structure that includes intensity statistics along 2D coordinate directions. There are different group of descriptors that work on content distribution instead. For example, instead of intensity values, geometric histograms [71] or shape context [72] were incorporated into similar procedure to form an image descriptor. “Shape context” works similar to SIFT but operates on Canny edges [46]. BRIEF⁴ is another descriptor which conducts small number of intensity difference tests to convert image patches to binary strings [73]. In each image patch, any pixel with higher intensity compared to the adjacent pixel is marked 1, and otherwise marked 0. Binary vectors are compared through calculation of hamming distance [74]. Texture can also be the studied feature of the studied image. This parameter can be better studied in frequency domain. Gabor filter [75] and Wavelet [76] are the old signal processing tools in frequency domain that also have been used for image description [77], and more specifically for place description [78]. “Gist” is a holistic image descriptor that uses Gabor filters to extract information from images [77]. Other than these types, spectral nature of images was used for description too. For example, in [79] central moments are used to characterize shape and intensity distribution of colour images.

¹ - Scale Invariant Feature Transform

² - Gradient location-orientation histogram

³ - Speeded-Up Robust Features Descriptor

⁴ - Binary Robust Independent Elementary Features

1.2.3.2.4 Comparison

Regardless of the approach for descriptor generation, in order to find similarity between descriptors, descriptor structures should be compared. A comparison method can be as simple as finding the Euclidean distance between descriptor vectors, or it can be much more complicated. Using the bag-of-words model [80] is the most referenced comparison methodology in image identification applications as well as in visual place recognition [81]. Bag-of-words model associates descriptor vectors of image patches to visual words. Afterwards, word processing techniques are used to compare images by comparing the visual words in their descriptor structures. Bag-of-words model eliminates the geometric and spatial relevance among image features. Apart from this, probabilistic calculations can also be used for descriptor comparison. Work [82] combines expectation minimization with “mixture of Gaussian conditional densities” for descriptor comparison. It is also possible to use the combination of bag-of-words model and probabilistic model [83]. Comparison method is very much related to the type of descriptor used. For example, work [84] studies the sequence of features, and therefore the author modified a biologically inspired DNA comparison methodology [85] for descriptor comparison.

1.2.3.2.5 Detection objectives

The ultimate purpose of place recognition research is to detect if the current observation of environment was re-observed or not. This becomes possible through comparison of descriptors. Comparison results should be incorporated in the specific application for which the detection procedure was designed. Many place detection techniques are highly inspired by SLAM. Although place detection is studied as an independent module, it can fulfil some important requirements of SLAM. Place detection provides loop closure clues for SLAM [86-88]. Moreover, it can be used to generate the topological map of the environment [23]. In a biologically inspired mapping methodology, place cell models of rat hippocampus is used for visual navigation without maintaining any metrically accurate map [89]. Place recognition is also used for map merging purposes [90] in multi robot exploration missions. It is also used in pure image retrieval applications [91, 92] to classify database records consisted of place images (e.g. [93]). SLAM++ may also be used directly for place detection through identification of map semantics. It can also be used to enhance other place detection

techniques by eliminating false detections. For example, kitchen should be eliminated from the list of possible results when a car is detected in the scene. In order to use SLAM++ for elimination of false positive detection in outdoor environment, detection of objects in urban environment has been attended by several research works such as [94]. In [94], 3D detection of objects from point clouds data is addressed. The entire point cloud is segmented into clusters. Subsequently, clusters are searched for features. Features are described through the number of points, estimated volume, average height, standard deviation in height, and the standard deviations in the two principle horizontal directions. Feature vectors represent specific objects in the environment.

1.3 Motivation

Place awareness is a critical component of safe decision-making in autonomous mobile robotics. Importance of place awareness is more appreciable in large-scale outdoor environment when safety measures, mission planning, and responsive intelligent behaviour depend on individual circumstances of each specific place. Besides, recall of revisited places assists the generations of loop closure and map merging clues for mapping components of mobile robots in large-scale outdoor exploration missions [88].

Place detection is commonly taken equivalent either to “snapshot pairing through viewpoint tolerant feature matching” [95] or to a “metrically described volume” [3]. While the former definition is more inspired by the research in image processing direction, which has indeed proven promising performance for object detection, the latter definition is heavily towards SLAM direction, which recently has resulted to outstanding accuracies in modelling indoor and small-scale environments. In the former approach, places are treated as “snapshots with distinct features”, and in the latter approach, a place is defined as “district confined within a metrically described volume”.

The former approach encounters challenges such as feature selection, feature pairing, storage limitations, viewpoint effects, scaling, and handling the changes of outdoor environmental conditions. However, in the latter approach, maintaining a metrically accurate model of the environment becomes very challenging, and processing such model becomes computationally expensive in large-scale outdoor environment.

Therefore, it is necessary to address place detection problem through a different approach. The new approach provides a new definition for “place” in the first instance.

A different sensing methodology is required to overcome viewpoint-related challenges in place detection. Non-feature-based-description and independence from metrical accuracy should stand out as other key properties in the new approach to overcome the aforementioned shortcomings. Nevertheless, since place detection is applied to large-scale outdoor environment, matching places should be concluded from a more confident approach than frame-to-frame comparison, being yet more computationally efficient than comparing new observations with the entire map of the environment. Therefore, an entirely new procedure is required to address place detection as an individual task independent from object/scene detection and metrically accurate self-localization modules in autonomous decision-making system.

Limitations of existing approaches for place detection are summarized as follows.

- Accurate definition of places through their mathematically defined boundaries becomes tedious, computationally expensive, and storage demanding for large-scale missions. Bounding primitives are not suited to include places with complex shapes.
- Defining places as snapshots containing specific landmarks has limitations to include places with no specific landmarks.
- Place detection analysis, which restrict places to single frames, are troubled to include large places, and complex-shaped places that cannot be entirely observed through a single snapshot.
- Place is not a directional measure and therefore, place detection techniques which presume specific observation direction, invalidate the logical meaning of a place.
- Feature selection and feature matching are open challenges. Place detection through feature processing is heavily affected by feature quality. Selection and association of features are the bottlenecks of feature based place detection.
- Maintaining the metrically accurate map of the environment is very challenging for large scale outdoor environment in long term autonomy.

This thesis proposes a new place detection methodology that overcomes the abovementioned shortcomings. Suggested definition for places is inspired by the work [96]. According to [96], a place is “somewhere distinctive relative to other nearby locations, according to some associated sensory information known as a place signature

or a place description”. Our proposed definition of a place is an extension of this definition, which addresses the consideration to all the above listed concerns. In this thesis, place is referred to as:

“A subset of a three dimensional environment in which, any partial observation across the entire field of view, holds consistent descriptor that is distinctive compared to its surrounding”.

This comprehensive description of a “place” illustrates the scope of the work to be presented in the thesis. The main objective of the thesis is to detect revisited places. Identified places are characterized according to the specified description. To achieve this, a proper sensing technique and a novel procedure are presented as supporting components for the suggested holistic and non-feature-based place detection methodology.

1.4 Proposed place detection approach

1.4.1 Detection procedure

The common place detection procedure and the suggested place detection procedure are depicted in the diagrams of Fig. 1-1. As illustrated, the suggested procedure does not require image segmentation and feature selection steps. Image segmentation, feature selection, and feature association are the main challenges in existing place detection techniques. These challenges are the sources of sensitivity to view point changes and sensitivity to environmental changes. Since the proposed method is non-feature-based, there is no need to store and update the structure conversion database (e.g. codebook in case of using bag-of-words model). This database is reported to have included 100 million code word-to-feature records in 20km traversal [91].

1.4.2 Input acquisition

An affordable sensory system, which can produce uniformly dense and long-range 3D point cloud across the entire usable field of view, is designed and made to generate the desired input data.

Colour images are inconsistent against environmental changes (e.g. illumination, and seasonal changes). Range images are independent from such variances. Existing 3D

point cloud sensors are either too expensive, or they are not made to produce uniformly dense 3D point clouds, and also they are mostly limited to capture a slice of the surrounding volume.

1.4.3 Place descriptor for characterization

Shape DNA [97] is the a sequence of eigenvalues of Laplace-Beltrami operator over a surface mesh. Since shape DNA does not assume existence of any specific type of features, incorporation of same idea for characterization of places overcomes the shortcomings of appearance-based feature matching techniques to include all outdoor places even the places with no specific landmarks. This is the primary reason that this research incorporates place DNA as the descriptor for place detection. Incorporation of DNA descriptor for place detection applications is reported by this research for the first time.

Fig. 1-1 – Left: Common place detection procedure; Right: Proposed place detection procedure

1.4.4 Environment segmentation

The environment being explored is a continuous 3D volume. However, it is possible to divide the environment to meaningful subsets called places. Places are distinctive compared to the surrounding. Instead of corresponding a place to a single descriptor structure, the proposed methodology corresponds places to a matrix of similarity

values. This matrix contains the similarity analysis of all the descriptors of different observations made from each place.

1.4.5 Similarity between places

To compare the similarity between two places, in the proposed methodology, every observation made from both places are compared against each other. This means, unlike the common practice of comparing two descriptor structures (which represent only two snapshots) for detection, two matrices, each containing several descriptor comparisons, are analysed.

1.4.6 Identification of similar scenarios

Comparing all observations through matrix comparison enables the detection of similar sequence of traversal along similar places by identifying the similar templates. Through template matching, a repeated experience of traversal through several places can be detected.

1.4.7 Identification of revisited places

Since a place is identified with a group of observations made out of it, a revisited place is identified by cross comparison of two groups of observations that are highly similar to each other. Sequence analysis enables comparison of patterns, which contributes to the enhancement of detection confidence. If necessary, it is also possible to pinpoint one scan from initial visit, which highly matches another scan in a subsequent visit. Thus, revisited place can either be identified through all the observations made from the place, or through one pair of highly similar scans; each scan made during an individual visit to the same place.

1.4.8 Place detection and place recognition

In this thesis, place detection is referred to the procedure of meaningful segmentation of the travelled route into the meaningful subsets called “places”. The focus of first four chapters of the thesis is on place detection.

Chapter 5 focuses on recognition of revisited places and revisited scenarios. Therefore, chapter 5 is about place recognition. However, whenever the term “place detection” is

used in chapter 5, it is used shortly in replacement for the term “place detection followed by recognition of revisited places and scenarios”.

The ultimate goal of the thesis is recognition of revisited places and scenarios, which is performed followed by analysis over the detected places along the travelled route. Therefore, whenever the general “place detection” term is used to refer to the overall research work, it is used in replacement for the term “place detection followed by recognition of revisited places and scenarios” procedure.

1.5 Contributions

In this research, a place detection methodology is proposed, which is novel in the approach and procedure. Moreover, some of the techniques used in each step of the procedure are new as well.

1.5.1 A novel approach to place detection

A procedure for identification of places through their place-similarity-matrices and through analysis of the “pattern of place sequences” is novel. Place –similarity-matrices are comprised of descriptor comparison among all the observations made from each place.

Instead of frame-to-frame comparison which is the common approach of scene detection techniques, the presented methodology studies the pattern of similarity values by comparing a set of partial observations from a place to a set of partial observations from another place. Therefore, through this approach, identified place is beyond one snapshot. The identified place, in the proposed approach, is a subset of the studied environment which is distinctive compared to the surrounding. In addition to comparison of several observations, in order to find a revisited place, pattern of place descriptor changes in consequent places are taken into consideration.

1.5.2 Place characterization with Laplace eigenvalues

To the best of our knowledge, this is the first time eigenvalues of Laplace operator are used to characterize places in a three dimensional analysis. Through incorporation of the mentioned descriptor, similarity of consequent observations can be studied confidently without extracting the features and without investigating the input details.

The fundamental analysis presented in chapter 3 is the indispensable material required to illustrate the superior capacity of Laplace eigenvalues for place characterization. This analysis never appeared in the literature before.

Common place detection approaches incorporate different feature extraction techniques combined with different data association methods. Inconsistency of features in outdoor environment, and data association shortcomings in large-scale exploration missions, advocate the incorporation of non-feature-based and non-appearance based technique to solve this specific problem. Eigenvalues of Laplace operator are believed to encapsulate intrinsic information about the curvature of the observed surfaces and therefore, they can be used as the preferred descriptor for place detection application.

1.5.3 Meaningful environment segmentation

Based on the definition of places matrices (explained in chapter 4), a new algorithm is proposed for environment segmentation into meaningful subsets.

In this work, places are identified as “meaningful segment of the environment” in accordance with the definition provided in section 1.3. This makes the proposed work different from conventional techniques which define places as snapshots with distinctive features. A place in this work is identified by a “set of partial observations with a consistent descriptor”. Real experiments support the fact that this segmentation technique yields to division of the studied environment to meaningful subsets (places). This treatment is another novel aspect of the research. Resultant segmentation provides a confident signal for detection of transition from one place to another.

1.5.4 Recognition of revisited places through recognition of revisited scenarios

Recognition of revisited scenario is another novel aspect of the work. This functionality is primarily used to narrow the search scope for recognition of revisited places.

A scenario is a sequence of places travelled during an exploration mission. In the proposed methodology, this concept is characterized by a template of similarity values. If a subsequent experience matches a sequence template, a scenario is identified as revisited scenario. Therefore, recalling a scenario through this methodology provides strong clues for loop closure detection and suggests proper map merging frames.

Identification of revisited places, which is performed after sequence matching, is strongly confident since the proposed procedure takes into account the comparison between many similarity values as well as the pattern of similarity values.

1.6 Outline of the thesis

Chapter 2 of the thesis explains the procedure for generation of place descriptor from input data. Hardware of a sensory system required for generation of input data is illustrated in chapter 2 as well. Key properties of the proposed place descriptor, which make it an ideal choice for the presented place detection methodology, are presented subsequently. In chapter 3, through examples, environment characterization using the proposed descriptor is studied. Chapter 4 incorporates the presented fundamentals for environment segmentation. Chapter 5 explains sequence analysis which facilitates recognition of revisited places, recalling revisited scenarios, and identification of best matching frames of two different visits to the same place. Proposed methodology is compared to conventional place recognition methodologies in chapter 6. Thesis is concluded in chapter 7 and potential future directions for this research are explained.

Chapter 2

Place Descriptor

In computer vision, a descriptor is a numeric structure from the set of key-features of the studied image. In ideal situation, the descriptor can fully describe the entire image. To remain analogous with the definition provided in section 1.3, noting that a place cannot necessarily be represented by one image, a proposed definition for a place descriptor is as follows:

“A numeric model encapsulating intrinsic information for unique identification of a place”.

For the case of feature-based descriptors, the numeric model represents key-features that describe the shape, colour, texture, and other appearance-related or geometric-related properties of the place. However, for place description, other methods that do not rely on the appearance of the studied phenomena should be preferred. This chapter addresses the respective concerns that contribute to making such preference. Moreover, a preferred descriptor is proposed and some key properties of the proposed descriptor are explained. Whenever the mathematical proof is available to support a property, proper references are provided. However, some of the properties can be inferred from experimental results. For such cases, the inferred property is explained through an example.

2.1 Classification of descriptors

2.1.1 Features

Numerous computer vision techniques are available for object detection [18]. These techniques extract the list of key-features of the object as object-descriptor. Most feature-based general image descriptors are extensions of such 2D object descriptors. As discussed in section 1.2.1, viewpoint invariance is a critical requirement for

successful place detection. Work [22] introduces a scale invariant descriptor which is not sensitive to rotation and translation. The descriptor introduced in [22] is the output of the analysis on histograms of gradients of Harris-corners. This well-known object descriptor is abbreviated SIFT. SIFT descriptors have been used for place detection as well (e.g. in [73]). SURF is another feature-based descriptor, which is based on forming the Hessian matrix at selected points that are not necessarily Harris corners. SURF is used in [17] as descriptor for visual place detection. Work [98] introduces a place descriptor, which is based on 3D sparse visual features. The features are visual key-points, mainly the corners, introduced in [99]. In [98], Description procedure is initiated by detection of landmarks. Followed by that, the visual features are extracted, and descriptor is formed. Descriptor is similar to 3D Gestalt descriptor [100-102]. Work [103], introduces surface entropy features (SURE) for place recognition purpose. SURE features are introduced in [104]. The feature is based on the statistics of the distribution of local surface normal vectors. If RGB data is available, then texture information is also appended to the feature. Distance from the object is another component stored in the feature structure. Followed by that, work [103] incorporates bag-of-words model to store and compare features.

Invariant feature selection is an ongoing challenge in computer vision. Feature-based-description confronts these challenges as well. Feature selection is not required for another group of descriptors. Examples of such non-feature-based descriptors are Wavelet descriptors [76-78], descriptors based on central moments [79], heat kernel signature [105, 106] and shape DNA [97].

View-point invariance is the primary property for a good place descriptor. Few of the aforementioned feature-based descriptors are believed to properly handle viewpoint inaccuracies as far as the inaccuracy is small compared to the dimension of the area being studied. The required viewpoint-invariance tolerance for place detection application is in scale of several meters, which is beyond the acceptable tolerance of the above mentioned techniques. Although 3D feature selection [107] enhances the viewpoint-invariance tolerance, sensitivity of features to environmental condition still remains an issue. Therefore, it can be concluded that non-feature-based descriptors are preferred for place description especially when large-scale outdoor environment is to be studied.

2.1.2 Appearance

All the feature-based descriptors in section 2.1.1 are extracted from the appearance (e.g. colour, or pattern) of the studied phenomena. Therefore, they are naturally sensitive to environmental condition. Such sensitivity introduces inconsistency in feature selection and feature matching in large-scale outdoor environment. Other than the appearance-based descriptors, there are other descriptors which study lines[108], planes [109], centroids[110], corners and intersections [111], and combination of these features. Another example of such features are the statistical summary of data segments extracted through PCA technique[63].

Such Non-appearance-based descriptors are less sensitive to environmental conditions and therefore they are preferred for place detection.

2.1.3 Domains

One common approach for image description is to apply feature selection on image patches instead of the entire image [63, 73]. Feature-based description techniques mentioned in section 2.1.1 follow this approach. Therefore, such descriptors are called local feature descriptors as well. However, image description can be done through holistic approaches as well. The author of paper [112] believes that human brain identifies places based on the global features of the scene. The author argues that human brain can recognize a place just by catching a brief glimpse, which may take even less than a second. Brain cannot recognize the details of the scene in this very short time. Therefore, place recognition mechanism in human brain is based on analysis of global features rather than local features. Holistic descriptors process the entire content of the image. For example, the colour histogram of the entire image forms the descriptor in [23], and the descriptor is used for classification of panoramic 2D colour images. Gist is another holistic image descriptor [112]. Gist descriptor is generated through operations on statistical summary of the spatial layout properties of the scene. Gist is used for place detection in [78] as well. Paper [113] presents a surface descriptor which is based on a local height field representation of a surface along with a transformation which adds rotational invariance to descriptor properties. Calculated height map, for each vertex, relies on Euclidean distance measurements. However, according to [114], this makes the method sensitive to scaling.

While local feature descriptors are more sensitive to illumination and environmental changes, holistic approaches are more sensitive to viewpoint changes [18].

2.2 Proposed descriptor

When input data is in form of point cloud, it is possible to use holistic and non-appearance-based descriptors without going through feature extraction procedure. In [115], the author introduced the Laplace spectra as finger print for image processing. The technique is the implementations of the mathematics presented in [116]. The descriptor is used for CAD model classification. In [116], the author presented the mathematics for derivation of discrete Laplace-Beltrami approximation along with eigenvalue and eigenfunction approximation methods for pattern recognition applications. The author also suggested studying heat kernel patterns for classification of low dimensional manifolds in high dimensional spaces. Results of [115] were later used in [117-119] for shape studies of grey matter of brain structures. The research demonstrated superior performance of the method in medical imaging application, especially for identification of populations of similar shapes. The research has been ongoing, from 2005 to 2015, for brain imaging and the recent work is published in [120]. In [97] the same author used the term “Shape DNA” to refer to the list of eigenvalues of Laplace-Beltrami operator. Through [121], we introduced a modified version of this idea to characterize places. The proposed descriptor demonstrated high level of viewpoint- invariance for place characterization. In this thesis, the proposed descriptor is referred to as the “Place-DNA”.

Definition - Shape DNA or Surface DNA

- List of eigenvalues of Laplace-Beltrami operator applied to the surface mesh model of the shape is called the shape DNA. In this thesis, shape DNA and surface DNA are taken equivalent.

A shape denotes a closed object with no boundaries or an open object with boundaries. Surface is a connected open shape with boundaries. In the proposed detection methodology, it is only possible to see the visible parts of the shapes that are in direct line of sight with the sensor. Therefore, only surfaces stay in the scope of the thesis. Whenever the term “shape” is used in the thesis, it is taken equivalent to a surface.

Definition - Place DNA

- List of eigenvalues of Laplace-Beltrami operator applied to the surface mesh model of all the observable surfaces in a place.

It can be inferred from the definitions of surface DNA and place DNA that a surface is a connected sub-part of a place. Table 2-1 summarizes the differences between shape DNA and place DNA.

Table 2-1 – Differences between shape DNA and place DNA

Criteria	CAD model / Shape DNA	Places / Place DNA
Structure	Eigenvalues of the Laplace operator over the studied shape/object	union of the eigenvalues of Laplace operator over the sets of surfaces forming the place
Domain	digitization of perfect CAD models that are studied at any desirable resolution	point clouds captured from real world subjected to limited resolution, which imposes different resolutions at different depth
Concerns	deformation and scaling	noise, occlusion, resolution, partial visibility

2.3 Uniformly dense digitization across the entire usable field of view

As explained in section 1.2.2, sensing technique affects place detection. In this section, limitations of commercially available sensors for place detection are explained, and a sensory system is suggested that overcomes the mentioned limitations.

2.3.1 Existing place detection sensors

Regardless of the place detection technique applied, cameras and LiDARs are the primary sensors for visual place detection.

2.3.1.1 Camera

Cameras are probably the most common sensors used for environment perception. Even basic models of cameras, which are inexpensive and widely available in the market, can produce rich input data. However, basic models have limited field of view and they produce images with distorted edges. These shortcomings are compensated in more professional models through distortion-less lenses. Panoramic view cameras [122], spherical view cameras [123] and hyperbolic mirrors for omnidirectional vision [124] are made to address field of view limitations. However, all these solutions require distortion recovery in software level. However, after the recovery process, resultant image does not include depth information. Models that include depth information have limited field of view [125, 126]. Depth accuracy of common commercial models is acceptable only for short distances. Long-range models are very costly. To the best of our knowledge, an affordable commercial camera, which can produce long-range depth information from the entire field of view, does not exist up to now. Besides, colour is the most valuable data produced by cameras. For outdoor applications, illumination and environmental factors highly affect the colour information captured by the camera. To conclude, and with respect to the place definition provided in section 1.3, commercially available camera models, without being fused by other sensors, may not produce the input data which can entirely fulfil the requirements of outdoor perception in every environmental situation.

2.3.1.2 LiDAR

Accurate long-distance depth measurement is achievable even by inexpensive models of LiDARs. LiDARs are also not sensitive to environmental changes and illumination. High-resolution 2D-LiDARs are commercially available [127]. When depth is the only information available for place detection, 2D information is not sufficient. Commercially available 3D-LiDARs, which are primarily made for collision avoidance, combine planar measurements across few rotated planes to generate 3D depth data [116]. Such data frame has sub-degree angular resolution inside each plane, but planes are several degrees rotated from each other. The result is a point cloud with non-uniform resolution in different directions. There are few models which can produce uniformly dense 3D point clouds [117], but those sensors either do not cover the entire field of view, or they are very costly [120].

2.3.2 Sensory system setup

Limitations of commercially available, yet affordable sensors are summarized as follows.

- Photos captured by camera are affected by illumination and environmental changes.
- Cameras have limited field of view; the special models that cover the entire field of view are 2D and they require distortion correction.
- Depth information from cameras is accurate for short-distances. 3D cameras do not cover entire field of view.
- Some 3D LiDARs produce non-uniform range information, and other models have limited field of view.

In order to generate a 3D dense, accurate, and uniform scan across the entire field of view, which is also not affected by environmental changes and is free from distortion debate, a LiDAR-based sensory system is proposed in this thesis. The hardware is a servo-mechanism that rotates a 2D-LiDAR. The idea of rotating a 2D-LiDAR for 3D point cloud generation was presented in several other research works before with various mounting directions and different rotation axes according to the requirements of the research [128, 129]. For place detection, the suggested setting is the 2D-LiDAR laid on horizontal plane, to be rotated around vertical axes.

The mechanism shown in Fig. 2-1, is a Hokuyo UTM-30LX (2D-LiDAR) [115] rotated by a servo motor HerkuleX DRS-601 [130]. The resolution of servo motor is 0.165 degree per step. Sweep range of LiDAR is 270 degrees with 0.25 degree steps.

When servo motor rotates, 2D-LiDAR scans are registered in the corresponding servo motor angle. The entire 3D environment is scanned when the servo motor rotates 180 degrees. Sample point clouds, generated by the proposed sensory system, are shown in Fig. 2-2.

Since the sweep angle of 2D-LiDAR is 270 degrees, a conic volume under the sensor is excluded from the resultant point cloud. This volume includes the robot body and small surface around the robot, which does not contain valuable information about the studied place. Therefore, using this sensory system, accurate, long-range, and uniformly dense 3D point cloud can be captured across the entire usable field of view.

Fig. 2-1 – Sensory system

Fig. 2-2 – Sample 3D point clouds generated by the proposed sensory system

2.4 Descriptor derivation procedure

The diagram in Fig. 2-3 summarizes the procedure for place DNA derivation.

Fig. 2-3 – Descriptor derivation procedure

2.4.1 Input data preparation

The raw data captured by the sensory system is a 3D point cloud across the entire usable field of view. The point cloud is uniformly dense in all directions. However, point cloud density is not uniform in depth dimension. Which means far objects are scanned with lower resolution compared to the close objects. Although increased density enhances calculation accuracy, it oppositely affects computational costs. Therefore, neither the lower resolution for far measurements, nor the higher resolution for close measurements are desirable. In order to achieve the desired density, which is also more uniform in depth dimension, points are uniformly down-sampled to voxels. Please note that better down sampling approaches exist which take into account the content of the point cloud [131]. Adaptive refinement approaches allow content aware down sampling. While flat surfaces can be highly down sampled, non-flat parts of the point cloud can be sampled with higher resolution. However, content aware down sampling

imposes further computational costs to the entire process. Study of the trade-off between content-aware down sampling benefits and the imposed computation cost is beyond the scope of this work. It will be considered to be further explored in the future work of this research.

After down sampling, resultant voxels are checked for outliers. Small cluster of voxels are taken as minor elements and are removed from the cloud.

2.4.2 Mesh generation

Mesh generation is required for eigenvalue derivation. The closed-form solution for eigenvalues and eigenfunctions of Laplace-Beltrami operator can be derived for rectangles, circles, spheres and other simple primitive shapes [132, 133]. But for complex surfaces, numeric methods should be used to approximate eigenvalues [134], and mesh generation is one direction to solve this problem. Several methods were suggested in the literature of the eigenvalue approximation. Major methods are:

- a) Approximation of continuous surface out of discrete data [130]
- b) Direct calculation on discrete data [135]
- c) Surface mesh analysis [118]

Method *a* is not recommended since approximating a continuous surface from discrete data is computationally expensive. After finding the analytic description for the continuous surface, approximated function might be so complicated that analytic solution to eigenvalue problem becomes tediously complex. In this situation, a numeric method should be used instead, and the initial approximation of continuous surface becomes only an extra overhead for the computation.

In the work [135] the mathematics of method *b* is presented. However, this mathematics is very complicated and we could not find any other work implementing this mathematics.

Mesh generation through polygonization of 2D and 3D point clouds have been studied for many years and several good implementations are available in existing programming libraries (such as in PCL [125]). In this approach, 3D point cloud is converted to a polygonised mesh. Derivation of Laplace-Beltrami operator over this mesh is possible through easier mathematics [136, 137].

In this thesis, polygonised surface mesh is generated from the scanned 3D-point cloud through the modified version of the polygonization technique explained in [138]. subsequently, triangulated mesh is converted to quadrangles. By connecting the center of each triangle to the middle point of each side, three quadrangles are made. There has been two reasons for this conversion. Firstly, this technique uniformly increases the mesh resolution and enhances eigenvalue approximation accuracy. Secondly, the eigenvalue solver used in this thesis takes quadrangled mesh inputs. After polygonazation, the 3D point cloud is converted to several surface mesh segments. Small segments are taken as outliers and are eliminated from calculations. A sample of resultant polygonised mesh is depicted in Fig. 2-4.

Fig. 2-4 – Left: Raw 3D Point Cloud; Right: Resultant polygonised mesh

2.4.3 Eigenvalues of Laplace-Beltrami operator

Laplace Beltrami operator over a manifold m is defined as the divergence of the gradient

$$\Delta m = \text{div}(\nabla(m)) = \frac{1}{\sqrt{g}} \partial_i (\sqrt{g} g^{ij} \partial_j), \quad (2.1)$$

where

$$G = \{g_{ij}\} \quad (2.2)$$

is the first fundamental form of m and

$$g \triangleq \det(g_{ij}), \quad (2.3)$$

$$g^{ij} \triangleq (g^{-1})_{ij}. \quad (2.4)$$

For the coordinate function $f(x, y, z)$, Laplace-Beltrami operator reduces to the simple Laplacian

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}. \quad (2.5)$$

This operator has a discrete set of eigenvalues. The Laplace eigenvalue problem is defined as

$$-\Delta u = \lambda u. \quad (2.6)$$

Where in (2.6), u is the eigenfunction associated with the eigenvalue λ . It can be shown that the eigenvalues are real numbers [134]. List of eigenvalues for each scan is a set of numbers increasing from zero to infinity.

To handle Laplace-Beltrami operator on discrete surface meshes several techniques are suggested in the literature such as:

- Graph Laplacian [139]
- Cotangent Weights [140]
- Finite Element Method [141]

A very good summary of these techniques can be found in [142]. Finite element method is the chosen technique for this research. This is mainly because of the availability of mature mathematics and programming libraries for solving differential equations through this technique, and because of the highly accurate results this method can produce compared to the other techniques. Using finite element method [143], the eigenvalues of (2.6) are equal to the eigenvalues of the generalized eigenvalue problem

$$-Au = \lambda Bu. \quad (2.7)$$

In (2.7), A and B are sparse positive definite (or semi-definite) symmetric matrices which are calculated using the finite element method with linear form functions. Unlike [143] which works on triangulated meshes, quadrangled meshes are processed in this thesis. Deal II library [144] is used to solve Laplace eigenvalue problem with finite element method. This library converts the differential equation eigenvalue problem (2.6) to a linear algebraic version in (2.7). Arnoldi Method [145], is used to solve the

linear algebraic eigenvalue problem in (2.7). Arnoldi method can calculate desired number of smallest eigenvalues from (2.7) as well. This technique is completely described in [146] and the code is perfectly implemented in Fortran (ARPACK [147]), MATLAB (eig function) and C++ (ARPACK++) libraries. C++ wrapper of deal II over ARPACK is used in this thesis to derive desired number of eigenvalues.

In [97] the author used the term Shape DNA to describe the sequence of eigenvalues of Laplace-Beltrami operator. DNA descriptor has properties, which makes it the proper choice of descriptor for place characterization. These properties are studied hereafter.

2.5 Properties of DNA descriptor

The author of [97] compared the shape DNA of several CAD models and managed to classify them based on shape DNA similarity. Place DNA, which was introduced in section 2.2, describes places. As mentioned in section 2.2, surfaces (or shapes) are connected sub-parts of a place.

The point cloud generated by the sensory system, introduced in section 2.3.2, is only consisted of points scanned from the surfaces of 3D objects in direct line of sight with the LiDAR. The parts behind the objects are invisible to the sensor as well as the parts beyond the sensor measurement range. This means, in majority of the cases, the entire closed 3D shape cannot be observed through a single scan. Therefore, the scope of this work becomes different from [97] which assumes uniform access to the entire object. With the proposed scanning technique, the boundaries of the scanned surfaces (2D surfaces of 3D shapes) are not necessarily the boundaries of the real 3D object. This is unlike the work in [97], which works on the entire mesh of a single studied volume with known boundaries (mostly closed shapes without any boundaries), and without any concerns for view-point changes, digitization resolution, or occlusion. In this thesis, only partial observations from several volumes in each scan can be accessed. Profile of the surfaces, number of surfaces, their boundaries, and the capturing resolution for each surface, can change with minor or major view-point changes.

For closed objects studied in [97], the differential equation can be solved assuming Neumann or Dirichlet boundary conditions and they produce the same results (at least for the eigenvalues). Generally, for 2D objects in 3D space with known boundaries, eigenvalue analysis is valid for both Neumann and Dirichlet conditions. However, the

results can be different. In this research, boundaries of the surfaces cannot be assumed to match the real boundaries of the objects. In this situation, Dirichlet condition is not a proper choice since it assumes constant values for the function at boundaries. Neumann condition is a better choice because it can, at least, cover the limitation of sensor range and, in many cases [134], it is equivalent to natural boundary conditions, and therefore it can be entirely eliminated from calculations. The mathematical concept of Laplace eigenvalue analysis is provided in many reference books such as [134]. Although Neumann condition is not the perfect choice for all cases, but we did not find a better boundary condition to fully cover the logic of floating boundaries of scanned surfaces. Selection of better boundary condition assumptions is left as an open problem. However, the conducted experiments in real application were supportive of Neumann boundary condition to be the reasonable choice in this situation.

Shape DNA of few primitive shapes are studied hereafter to explain the properties of DNA descriptor.

2.5.1 Resolution

First few eigenvalues of the rectangle are listed in Table 2-2. It can be seen that, by increasing the resolution, approximated eigenvalues calculated by numeric method converge to the exact values calculated from analytic method. This means the resolution has direct effect on calculation precision. However, increasing the resolution increases the computational costs. Therefore, if the objective is to study the overall similarities of point clouds, it is not necessary to study them in very high resolution. The exact eigenvalues of the rectangle, given in the last column of Table 2-2, are calculated in [148] through

$$\lambda_{i+j} = \frac{i^2 + j^2}{4} \pi^2, \text{ for } i, j \geq 0. \quad (2.8)$$

In (2.8), i and j are eigenvalue indices respective to x and y coordinates.

Selection of the appropriate resolution depends on the application. In the case of this thesis, for place detection application, the captured 3D point cloud is down sampled to 30cm voxel grid, evenly in all directions. This resolution is good enough to capture the overall shape of the visible surfaces in outdoor environment.

For most of the analysis presented in this section, the flat rectangle and the pyramid with a fixed resolution are studied. The first few eigenvalues of them are listed in Table 2-2.

Table 2-2 – Exact and numeric approximation of the eigenvalues of flat rectangle

	Low Resolution	High Resolution	Exact Value From Analytic Solution
Eigenvalue 0	3.35E-15	1.74E-14	0
Eigenvalue 1	3	2.49927	2.4674011
Eigenvalue 2	3	2.49927	2.4674011
Eigenvalue 3	6	4.99854	4.934802201
Eigenvalue 4	12	10.3866	9.869604401
Eigenvalue 5	12	10.3866	9.869604401
Eigenvalue 6	15	12.8859	12.3370055
Eigenvalue 7	15	12.8859	12.3370055
Eigenvalue 8	24	20.7733	19.7392088
Eigenvalue 9	24	20.7733	19.7392088

2.5.2 Noise and occlusion

In Table 2-3, although the eigenvalues of the perturbed rectangles are not the same as those of original shape, but the list is a perturbed version of the list of eigenvalues of the original shape. This shows that the perturbation in shape results to perturbation in shape DNA. Repeated eigenvalues in the original list reflect symmetry in the shape (this is mathematically proven in [134]). With one missing element, the shape loses the symmetry and so the repeated eigenvalues disappear. However, since the noisy (perturbed) samples are nearly symmetric, eigenvalues with very small differences can be spotted.

Table 2-3 – Initial eigenvalues of rectangle compared to the eigenvalues of pyramid

	Flat Rectangle	Pyramid
Eigenvalue 1	2.49927	2.27907
Eigenvalue 2	2.49927	2.27907
Eigenvalue 3	4.99854	4.92196
Eigenvalue 4	10.3866	6.26958
Eigenvalue 5	10.3866	9.13991
Eigenvalue 6	12.8859	11.6944
Eigenvalue 7	12.8859	11.6944
Eigenvalue 8	20.7733	16.1591

The difference between the two noisy rectangles in Table 2-3 is not a rotation. If the noise level is increased, it is expected that perturbation from original DNA is increased too. In Table 2-4 the DNA of original shape is compared to the DNA of noisy versions with increased levels of noise. Comparing the DNAs suggests that the higher level of noise in the shape results to higher level of perturbation in the DNA.

Property 1 - Perturbation

- Perturbation in shape results to perturbation in DNA. This has been reported in [130] as well.

(inferred) Property 2 - Consistency

- Shapes with similar structures have similar DNAs.

Noisier version of rectangle is compared with less noisy versions and the original rectangle in Table 2-4.

Table 2-4 – Initial eigenvalues of original and noisy rectangles

	Original	One Element Missing	Another Element Missing
Eigenvalue 0	1.74E-14	4.38E-14	2.11E-14
Eigenvalue 1	2.49927	2.36249	2.31861
Eigenvalue 2	2.49927	2.48068	2.38032
Eigenvalue 3	4.99854	4.80005	4.84292
Eigenvalue 4	10.3866	9.33621	9.39302
Eigenvalue 5	10.3866	9.42884	10.0364
Eigenvalue 6	12.8859	11.8677	12.1772
Eigenvalue 7	12.8859	12.1147	12.4568
Eigenvalue 8	20.7733	19.397	19.1295

Table 2-5 – Initial non-zero eigenvalues of original rectangle and noisy versions

	Flat Rectangle	Noisy (V1)	Noisy (V2)	Noisy (V3)
Eigenvalue 0	2.46938	2.36249	2.31861	2.20035
Eigenvalue 1	2.46938	2.48068	2.38032	2.27113
Eigenvalue 2	4.93877	4.80005	4.84292	4.46532
Eigenvalue 3	9.90135	9.33621	9.39302	7.90938
Eigenvalue 4	9.90135	9.42884	10.0364	8.47732
Eigenvalue 5	12.3707	11.8677	12.1772	10.6631
Eigenvalue 6	12.3707	12.1147	12.4568	12.1568
Eigenvalue 7	19.8027	19.397	19.1295	18.2968

(inferred)Property 3 - Noise affects the extent of similarity

- Higher perturbation (in form of noise) results to lower DNA similarity.

In presence of noise, consistency of DNA structure is maintained until the structure of shape is maintained. DNA pattern changes when the shape structure is changed. In real application, this situation can occur due to occlusion.

Table 2-5 compares the original, noisy, and occluded rectangles. The pattern of eigenvalues listed in the first two columns of Table 2-5 are similar. It can still be seen that the DNA of the highly noisy rectangle is still similar to the DNA of original rectangle. However, for occluded rectangle (L-shape surface), the structure of the shape has changed and so the pattern of the DNA is no more similar to the pattern of the original shape DNA. Still, some of the eigenvalues of occluded rectangle are exact copies of original rectangle. So, still some level of similarity can be inferred.

Table 2-6 – Initial eigenvalues of original, noisy, and occluded rectangles

	Original	Large hole – noisy rectangle	L Shape –occluded rectangle
Eigenvalue 0	1.74E-14	2.17E-14	-3.44E-15
Eigenvalue 1	2.49927	2.05271	1.51652
Eigenvalue 2	2.49927	2.09934	3.5847
Eigenvalue 3	4.99854	4.74576	10.3866
Eigenvalue 4	10.3866	9.29632	10.3866
Eigenvalue 5	10.3866	10.6043	11.9143
Eigenvalue 6	12.8859	12.2441	13.2568
Eigenvalue 7	12.8859	13.0977	20.7733
Eigenvalue 8	20.7733	19.977	24.0405

(Inferred) Property 4 - Occlusion changes DNA structure

- Occlusion that changes the shape structure, substantially changes the shape DNA. Minor occlusion that does not change the structure results to perturbation in shape DNA.

In summary, noise and occlusion in a shape can change the shape DNA. This change is in form of *perturbation of eigenvalues* or in form of *altered DNA structure*. But in any case, so far that the noisy or occluded shape has noticeable similarity with the original shape, similarities can be tracked in DNA of original and altered versions of the shape.

2.5.3 DNA of sub-parts

Place is consisted of several surfaces generated out of different clusters of points in a scan. Each surface has its own DNA. Place DNA is the union of DNAs of its sub-parts.

When shape *A* is repeated *N* times in place *B*, eigenvalues of *A* are *N* times repeated in the DNA of *B*. An example is given in Table 2-6, where the DNA of single pyramid and double pyramid shapes are listed.

Property 5 - Repeated sequence of eigenvalues

- Repeated sequence of eigenvalues in place DNA denotes the repetition of a sub-part.

In Table 2-7 and Table 2-8, DNA of an ideal place and DNA of all sub-parts are listed. Colour codes and titles assigned to eigenvalues in these two tables are provided for better illustration of the comparison between the values. All the eigenvalues listed in Table 2-7 are repeated in Table 2-8.

(Inferred)Property 6 - Union of eigenvalues

- If DNA of shape *A* is listed in “*list a*”, and DNA of *B* is listed in “*list b*”, then DNA of the place *C* that contains *A* and *B* will be the “*list c*” which is the union of *list a* and *list b*.

Table 2-7 – Initial eigenvalues of pyramid and duplicated pyramid

	Single	Double
Eigenvalue 0	-2.82E-15	-5.12E-14
Eigenvalue 1	2.27907	9.29E-14
Eigenvalue 2	2.27907	2.27907
Eigenvalue 3	4.92196	2.27907
Eigenvalue 4	6.26958	2.27907
Eigenvalue 5	9.13991	2.27907
Eigenvalue 6	11.6944	4.92196
Eigenvalue 7	11.6944	4.92196
Eigenvalue 8	16.1591	6.26958
Eigenvalue 9		6.26958
Eigenvalue 10		9.13991
Eigenvalue 11		9.13991
Eigenvalue 12		11.6944
Eigenvalue 13		11.6944
Eigenvalue 14		11.6944
Eigenvalue 15		11.6944
Eigenvalue 16		16.1591
Eigenvalue 17		16.1591

Table 2-8 – Initial non-zero eigenvalues of ideal elements; eigenvalues less than 20 are listed.

Wave		Pyramid		Flat Rectangle		Hills		Rock	
Label	Eigenvalue	Label	Eigenvalue	Label	Eigenvalue	Label	Eigenvalue	Label	Eigenvalue
W1	1.86	P1	2.28	F1	2.50	H1	1.32	R1	2.06
W2	1.86	P2	2.28	F2	2.50	H2	1.32	R2	2.21
W3	3.69	P3	4.92	F3	5.00	H3	2.45	R3	4.66
W4	9.00	P4	6.27	F4	10.39	H4	8.51	R4	7.53
W5	9.00	P5	9.14	F5	10.39	H5	8.51	R5	8.29
W6	11.13	P6	11.69	F6	12.89	H6	10.49	R6	10.99
W7	11.13	P7	11.69	F7	12.89	H7	10.49	R7	11.17
W8	18.40	P8	16.16			H8	11.03	R8	16.68
W9	18.40	P9	19.11			H9	11.03	R9	18.49
W10	19.15	P10	19.11			H10	11.14	R10	19.27
						H11	12.16		
						H12	19.45		

Table 2-9 – Initial non-zero eigenvalues of an ideal place; eigenvalues less than 20 are listed (colour code and labels provided for easier comparison with Table 2-8)

Ideal Place											
Label	H1	H2	W1	W2	R1	R2	P1	P2	H3	F1	F2
Eigenvalue	1.32	1.32	1.86	1.86	2.06	2.21	2.28	2.28	2.45	2.50	2.50
Label	W3	R3	P3	F3	P4	R4	R5	H4	H5	W4	W5
Eigenvalue	3.69	4.66	4.92	5.00	6.27	7.53	8.29	8.51	8.51	9.00	9.00
Label	P5	F4	F5	H6	H7	R6	H8	H9	W6	W7	H10
Eigenvalue	9.14	10.39	10.39	10.49	10.49	10.99	11.03	11.03	11.13	11.13	11.14
Label	R7	P6	P7	H11	F6	F7	P8	R8	W8	W9	R9
Eigenvalue	11.17	11.69	11.69	12.16	12.89	12.89	16.16	16.68	18.40	18.40	18.49
Label	P9	P10	W10	R10	H12						
Eigenvalue	19.11	19.11	19.15	19.27	19.45						

2.5.4 Rotation and translation

DNAs of rotated version of the same shape are listed in Table 2-9. It can be seen that rotation does not change the DNA.

In Table 2-10, two double pyramid sets are compared. One set is consisted of distant pyramids and the other consists of a zero distance double pyramid set. Please note that the zero distance set is not connected. Table 2-10 makes it clear that the DNA does not change when position of the objects are changed.

Property 7 - Position and orientation of a shape do not affect its DNA

- Eigenvalue analysis is *space independent*. This means, position and orientation of a shape in space does not affect the DNA [130].

Table 2-10 – Eigenvalues of rotated objects

	L Shape	Rotated L Shape
Eigenvalue 0	-3.44E-15	-5.62E-14
Eigenvalue 1	1.51652	1.51652
Eigenvalue 2	3.5847	3.5847
Eigenvalue 3	10.3866	10.3866
Eigenvalue 4	10.3866	10.3866
Eigenvalue 5	11.9143	11.9143
Eigenvalue 6	13.2568	13.2568
Eigenvalue 7	20.7733	20.7733
Eigenvalue 8	24.0405	24.0405

2.5.5 Connectivity, change of boundary

The place DNA being only the union of the DNA of its sub-parts was illustrated by the examples presented in section 2.5.3. However, the connectivity among sub-parts does affect the place DNA. Examples provided so far studied the shapes with disconnected sub-parts.

Table 2-11 – Eigenvalues of translated objects

	Distant	Zero Distance
Eigenvalue 0	-5.12E-14	-4.67E-14
Eigenvalue 1	9.29E-14	-8.11E-15
Eigenvalue 2	2.27907	2.27907
Eigenvalue 3	2.27907	2.27907
Eigenvalue 4	2.27907	2.27907
Eigenvalue 5	2.27907	2.27907
Eigenvalue 6	4.92196	4.92196
Eigenvalue 7	4.92196	4.92196
Eigenvalue 8	6.26958	6.26958
Eigenvalue 9	6.26958	6.26958
Eigenvalue 10	9.13991	9.13991
Eigenvalue 11	9.13991	9.13991
Eigenvalue 12	11.6944	11.6944
Eigenvalue 13	11.6944	11.6944
Eigenvalue 14	11.6944	11.6944

In Table 2-11, DNA of single pyramid object is compared to the DNA of a *disconnected double pyramid set in zero distance* and the DNA of a *connected double pyramid set*. Although the disconnected double pyramid looks very much like connected double pyramid set, the difference is that the disconnected pyramids have 81 individual nodes each (disconnected double pyramid set has total 162 nodes), but connected pyramids have 9 common nodes at the connection line (connected pyramid set has total 153 nodes). Analysis of the Table 2-11 clears that the shape DNA of disconnected double set has exactly two copies of the single pyramid eigenvalues, and nothing more. However, for the case of connected double pyramid set, although the shape DNA reflects inheritance from a single pyramid since it contains all the eigenvalues of the single pyramid, the connected object has actually a new identity as an individual shape, and therefore, it contains additional eigenvalues to describe its unique identity.

Different attachments can result to different shape-DNAs. For example, in Table 2-12, attachment between two pyramids with only one common side is compared to the attachment of two pyramids with four common sides. Although the shape-DNAs are different (please see Table 2-12), both shape-DNAs still contain all the eigenvalues of the single pyramid.

Table 2-12 – Initial eigenvalues of single pyramid, connected double pyramid, and disconnected double pyramid

	Original	Zero Distance	Attached
Eigenvalue 0	-2.82E-15	-4.67E-14	-4.94E-14
Eigenvalue 1	2.27907	-8.11E-15	0.480441
Eigenvalue 2	2.27907	2.27907	2.27907
Eigenvalue 3	4.92196	2.27907	2.27907
Eigenvalue 4	6.26958	2.27907	2.83822
Eigenvalue 5	9.13991	2.27907	4.35021
Eigenvalue 6	11.6944	4.92196	4.92196
Eigenvalue 7	11.6944	4.92196	6.26958
Eigenvalue 8	16.1591	6.26958	7.54159
Eigenvalue 9		6.26958	7.81867
Eigenvalue 10		9.13991	9.13991
Eigenvalue 11		9.13991	11.6944
Eigenvalue 12		11.6944	11.6944
Eigenvalue 13		11.6944	12.8249
Eigenvalue 14		11.6944	13.3005
Eigenvalue 15		11.6944	16.1591
Eigenvalue 16		16.1591	17.21
Eigenvalue 17		16.1591	19.1054

Table 2-13 – Initial eigenvalues of single pyramid, single-side-connected double pyramid, and all-sides-connected double pyramid

	Original	1 Attachment	4 Attachments
Eigenvalue 0	-2.82E-15	-4.94E-14	1.48E-13
Eigenvalue 1	2.27907	0.480441	2.27907
Eigenvalue 2	2.27907	2.27907	2.27907
Eigenvalue 3	4.92196	2.27907	2.72614
Eigenvalue 4	6.26958	2.83822	4.92196
Eigenvalue 5	9.13991	4.35021	6.26958
Eigenvalue 6	11.6944	4.92196	9.13991
Eigenvalue 7	11.6944	6.26958	9.5103
Eigenvalue 8	16.1591	7.54159	9.5103
Eigenvalue 9		7.81867	11.6944
Eigenvalue 10		9.13991	11.6944
Eigenvalue 11		11.6944	15.0154
Eigenvalue 12		11.6944	16.1591
Eigenvalue 13		12.8249	17.3711
Eigenvalue 14		13.3005	19.1054
Eigenvalue 15		16.1591	19.1054
Eigenvalue 16		17.21	21.3964
Eigenvalue 17		19.1054	24.0395

2.6 Chapter conclusion

This chapter proposed eigenvalues of Laplace-Beltrami operator, which is a holistic and non-feature-based descriptor, as the preferred place descriptor. The proposed descriptor is preferred because of the properties that makes it superior for detection of minor and major changes in studied place. The descriptor is invariant against rotation and translation. Most importantly, perturbation (minor changes) in the studied place, results to perturbation in the DNA. This means, if the DNA that describes a new observation from the environment is not majorly different from the previous DNA, it can be concluded that the new observation is still made from the same place. Likewise, major change of DNA is expected to denote transition to a new place. Since the proposed descriptor does not depend on features, it is believed to be more robust against environmental changes in outdoor environment. Elimination of feature detection and feature matching eliminates the necessity of memory allocation for feature descriptor database storage. This enhances the scalability of the entire place detection methodology. Chapter 3 studies the proposed descriptor in place detection application, initially by characterization of different indoor and outdoor places.

Chapter 3

Place Characterization

In section 2.5, properties of DNA descriptor were presented. In this chapter, DNA descriptor with the presented properties is used to characterize places as meaningful subsets of the environment.

In this thesis, environment is referred to the surrounding of the route travelled by the robot. An environment is comprised of several places with different shapes and dimensions. Each place is comprised of connected sub-parts called surfaces. DNA comparison can be made at surface level or at place level. This chapter explains characterization of places through their DNAs. Different observation from the same place are made at different viewpoints. Visible surfaces are different from each viewpoint, and therefore the DNA of the different observations from a place are not identical, but they are yet similar. DNA similarity analysis provides the means for identification of individual places through their DNA.

To the best of our knowledge, incorporation of DNA descriptor for place characterization never appeared in the literature up to now. Although most of the material covered in chapter 2, and the material presented in this chapter up to section 3.1.2, can be understood or can be inferred from the previous research work which first was introduced by [97] and later got enhanced by others, e.g. [149], the analysis succeeding section 3.1.2 never appeared in the literature. The analysis presented in the existing literature is very similar to the one explained in section 3.1.1 where ideal computer generated shapes are compared with each other. However, the analysis that illustrates the comparison of “partial observations from a group of surfaces” and subsequently, the analysis of point clouds generated through digitization of real-world places, form the indispensable fundamentals that promote DNA description to the preferred technique for place characterization. Concerns such as occlusion and

attachment of sub-parts are specific to place DNA, and are out-of-debate in shape-DNA analysis.

3.1 Uniqueness of DNA

The work [150] introduces iso-spectral plane domains and surfaces. One simple example of such surfaces is shown in Fig. 3-1. Later the work [151] mathematically proved that only certain convex planar regions with analytic boundary hold unique set of eigenvalues. Without the mentioned assumption, it is sometimes possible to synthesize two iso-spectral surfaces. However, without accurate synthesis, for the surfaces generated through polygonization of complex point clouds, the possibility of visiting iso-spectral surfaces is negligible. Therefore, according to the extremely small possibility of visiting iso-spectral surfaces in scope of the work presented in this thesis, uniqueness of place DNA is a valid assumption for place detection application.

According to [152], shapes with different perimeter, area, or genus have indeed different spectrum. However, it is possible to synthesize two iso-spectral shapes (which of course have the same area, perimeter, and genus in the first place) through transplantation techniques [153]. In this research, surfaces are generated from point clouds captured from the real world. Therefore, encountering non-similar surfaces with exact same area, same perimeter and same genus, which also hold the same spectrum, is of extremely negligible probability. Place is consisted of several surface. Although it is possible to think of different sets of surfaces with similar total spectrum [150], the occurrence of such situation is of even extremely smaller probability in real application. Formulation of this probability is beyond the scope of this research and we could not find any other work in the literature that quantifies this small probability.

Fig. 3-1 – An example of iso-spectral surfaces introduced by [150]

3.1.1 Comparison between surface DNAs

Uniqueness of DNA implies that DNA of non-identical surfaces are different. However, it is expected that similarity between surfaces could be inferred from comparison of their DNAs. This statement is further discussed through an example.

Table 3-1 lists the DNA of six different surfaces. Only first 30 non-zero eigenvalues are listed for each DNA.

Definition - Inflated DNA

- Inflated DNA is generated by assignment of tolerance to eigenvalues of the original DNA.

The attempt to find exact eigenvalues of one surface inside the DNA of another surface results to no possible matches since DNA is unique for each surface. However, according to property 1, perturbed version of a shape has perturbed version of its DNA. Perturbed version of DNA can be simulated by assignment of some tolerance to its eigenvalues. If for each eigenvalue λ a tolerance of $\varepsilon\%$ of λ is assumed, inflated version of DNA is generated. In this situation, the comparison between inflated surface DNAs (perturbed surface DNAs) reflects similarity between surfaces. For example, for $\varepsilon=3$, DNA comparison result for the six studied surfaces is according to Table 3-2. In this example, 35 smallest non-zero eigenvalues are compared. The number provided in each cell of Table 3-2 is the percentage of matching eigenvalues when inflated DNAs are compared.

It can be learned from Table 3-2 that identical surfaces show 100% DNA match (through original or inflated DNA comparison). Moreover, similar surfaces have high percentage of matching eigenvalues. For example, flat rectangle and noisy rectangle show 80% matching eigenvalues when their inflated DNAs are compared.

Inference 1 - Similar and identical surfaces

- Two surface DNAs can be compared by counting their matching eigenvalues. Unless two surfaces are exactly identical, exact matching DNAs are not expected. However, similarity between surfaces can be studied through comparison of the inflated version of their DNAs.

Table 3-1 – DNA of six different surfaces

	Wave	Pyramid	Flat Rectangle	Noisy Rectangle	Hills	Rock
1	1.86	2.28	2.50	2.30	1.32	2.06
2	1.86	2.28	2.50	2.36	1.32	2.21
3	3.69	4.92	5.00	4.61	2.45	4.66
4	9.00	6.27	10.39	8.64	8.51	7.53
5	9.00	9.14	10.39	9.23	8.51	8.29
6	11.13	11.69	12.89	11.37	10.49	10.99
7	11.13	11.69	12.89	12.88	10.49	11.17
8	18.40	16.16	20.77	19.60	11.03	16.68
9	18.40	19.11	24.87	22.66	11.03	18.49
10	19.15	19.11	24.87	25.45	11.14	19.27
11	20.39	24.04	27.37	25.91	12.16	23.26
12	21.29	25.63	27.37	28.58	19.45	25.10
13	30.33	29.19	35.26	32.54	20.23	29.48
14	31.25	29.72	35.26	35.91	22.89	30.17
15	31.25	29.72	48.00	45.52	22.89	34.22
16	34.85	37.20	48.00	46.40	29.89	37.17
17	34.85	41.38	49.74	47.44	29.89	41.17
18	37.47	45.38	50.50	51.20	31.04	43.01
19	39.71	45.38	50.50	52.89	35.67	44.31
20	51.54	48.12	58.39	54.48	37.55	48.12
21	51.54	52.82	58.39	63.85	45.09	50.13
22	54.70	56.22	72.87	71.18	45.09	53.64
23	57.34	56.22	72.87	72.33	47.29	56.66
24	57.34	62.68	82.07	80.84	47.29	58.20
25	59.30	62.68	82.07	82.23	49.33	61.19
26	59.30	68.13	84.57	82.71	49.33	67.06
27	61.73	72.43	84.57	87.03	50.07	70.76
28	74.08	74.77	92.46	90.16	51.07	75.92
29	79.98	81.78	92.46	92.03	58.65	78.99
30	79.98	81.78	96.00	98.78	60.41	80.30

Fig. 3-2 – Illustration of the DNAs of similar shapes and different shapes

Table 3-2 – Surface DNA comparison result; inflated DNAs are compared

	100	31.43	34.29	31.43	37.14	45.71
31.43	100	28.57	42.86	25.71	54.29
34.29	28.57	100	80	31.43	31.43
31.43	42.86	80	100	28.57	37.14
37.14	25.71	31.43	28.57	100	40
45.71	54.29	31.43	37.14	40	100

To illustrate the similarity between the DNAs of similar and different shapes, three of the shapes of Table 3-1 are further studied. Fig. 3-2 draws the initial of the DNA of two very similar shapes (rectangle, and noisy rectangle) along with the initial of the DNA of a different shape (hills). Two important results can be inferred from Fig. 3-2. First is the uniqueness of DNA even for very similar shapes. It can be seen in Fig. 3-2 that even though rectangle is only different from noisy rectangle in a single tile, the DNAs are not exactly overlapping. The second result is that the similar, but not exactly identical, shapes have DNAs with similar growth rate. Moreover, with a small tolerance, many matching eigenvalues can be tracked between the two DNAs. In another word, the DNA of the similar shape is consisted of perturbed version of the set of eigenvalues of the original shape's DNA. On the other hand, a different shape holds a DNA with totally different growth rate, and therefore fewer matching eigenvalues can be tracked when DNAs of two shapes are compared. Comparison of two sets of numbers can be done through many different techniques. The proposed technique for DNA descriptor comparison is presented in section 3.3.2 in details.

3.1.2 Comparison between a surface DNA and a place DNA

As explained through property 6, DNA of a place is the union of multiple surface DNAs. A surface being a connected sub-part of a place, it is expected that all the eigenvalues of the surface DNA could be found in the DNA of the place.

In Table 3-3, DNA of an ideal place is compared to the DNA of Rock as a sub-part. In the given example, the eigenvalues of Rock at positions 1 to 9 are exactly repeated in the DNA of the place in positions 5, 6, 13, 17, 18, 28, 34, 41, and 44. Therefore, given only the DNA of a place and the DNA of one of its sub-parts, without knowing anything about the appearance or structure of the place and the surface, it can be checked whether the studied place contains the specified surface. Exact values can be compared in ideal situation, but in real situation, inflated DNAs should be compared instead. The result of such comparison becomes more accurate when "point clouds to be compared" have higher number of points, and the surface meshes generated out of them are more complex. An example of comparison between the inflated DNAs of a place and a surface is presented in section 3.3.2.

Inference 2 - Inclusiveness

- If all the eigenvalues of a surface DNA are found in a place DNA, the place contains the surface.

3.1.3 Comparison between place DNAs

The example provided in section 3.1.2 can be used as a directive for comparison of two place DNAs. Comparison between two place DNAs can be taken as checking whether most of the surface DNAs contained in one place DNA can be found in another place DNA or not. An illustrative example is provided as follows.

Fig. 3-3 depicts an ideal place that contains five different surfaces. Five different observations were made from this place such that, due to some occlusion, in each observation, only four of the surfaces are visible. Table 3-4 summarizes the result of comparison among the place DNAs generated from different observations, when DNAs are consisted of the first 150 non-zero eigenvalues.

Please note that the DNAs are compared through comparison of their eigenvalues and not through comparison among their sub-parts. This means, the analysis does not include the initial phase of breaking the place to sub-parts. Observations are compared as a whole without taking into account which sub-parts are contained. It can be learnt from Table 3-4 that although different observations from a place are not identical, inflated DNAs of different observations still demonstrate many matching eigenvalues.

Inference 3 - Different observations from a place

- If most of the eigenvalues of a place DNA match the eigenvalues of another place DNA, two DNAs refer to observations from similar places (or maybe to two observation from the same place).

Table 3-3 – DNA of an ideal place compared to the DNA of Rock as a sub-part.

Ideal Place DNA											
Label	Eigenv1	Eigenv2	Eigenv3	Eigenv4	Eigenv5	Eigenv6	Eigenv7	Eigenv8	Eigenv9	Eigenv10	Eigenv11
Eigenvalue	1.32	1.32	1.86	1.86	2.06	2.21	2.28	2.28	2.45	2.50	2.50
Label	Eigenv12	Eigenv13	Eigenv14	Eigenv15	Eigenv16	Eigenv17	Eigenv18	Eigenv19	Eigenv20	Eigenv21	Eigenv22
Eigenvalue	3.69	4.66	4.92	5.00	6.27	7.53	8.29	8.51	8.51	9.00	9.00
Label	Eigenv23	Eigenv24	Eigenv25	Eigenv26	Eigenv27	Eigenv28	Eigenv29	Eigenv30	Eigenv31	Eigenv32	Eigenv33
Eigenvalue	9.14	10.39	10.39	10.49	10.49	10.99	11.03	11.03	11.13	11.13	11.14
Label	Eigenv34	Eigenv35	Eigenv36	Eigenv37	Eigenv38	Eigenv39	Eigenv40	Eigenv41	Eigenv42	Eigenv43	Eigenv44
Eigenvalue	11.17	11.69	11.69	12.16	12.89	12.89	16.16	16.68	18.40	18.40	18.49

Rock DNA											
Label						Eigenvalue					
R1						2.06					
R2						2.21					
R3						4.66					
R4						7.53					
R5						8.29					
R6						10.99					
R7						11.17					
R8						16.68					
R9						18.49					

In real application, ideal surfaces are not available. Therefore, presented comparison should be applied to inflated DNAs of places instead. Number of eigenvalues to be studied for fair similarity analysis, percentage of matching eigenvalues, and the tolerance for matching eigenvalues cannot be determined analytically. However, there are techniques which can be used as directives for selection of initial values for these parameters, and fine tuning should be done through trial and error. Selection of initial values for these parameters is discussed subsequently.

Fig. 3-3 – An ideal place containing five different surfaces; five different observations are made from the studied place.

Table 3-4 – DNA comparison of five different observations from an ideal place

Observations from the studied place					
	100	78	75.33	75.33	78
	78	100	80	76	81.33
	75.33	80	100	75.33	82
	75.33	76	75.33	100	75.33
	78	81.33	82	75.33	100

3.2 Percentage of matching eigenvalues

According to the definition of a place, provided in section 1.3, different partial observations from a place are supposed to have similar descriptors (DNAs). However, percentage of matching eigenvalues to be used as the threshold for marking two observations as similar, can be anticipated according to the experiment parameters.

It can be assumed that the difference between two observations from a place is caused by viewpoint displacements. The uniform distribution of data across the entire usable field of view, and the fact that the place descriptor is consistent in pure rotations, concludes that the pure rotation of the sensor in the environment does not affect the comparison. Hence, translation is the only factor that causes the difference between two observations. If d is the distance travelled between two consequent observations, and r is the sensor measurement range, in the ideal situation, maximum visible area from both consequent view-points is the area of the intersection of two spheres. Area of the intersection surface is less than the area of one entire sphere (see Fig. 3-4). The proportion of the intersection area to the area of the sphere is the anticipated similarity between two consequent scans in ideal situation, which can be calculated from

$$\frac{\text{Area of the intersection}}{\text{Area of a full sphere}} = \frac{4\pi r(r-d)}{4\pi r^2} = \frac{r-d}{r}, \quad (3.1)$$

where r is sensor measurement range, and d is the distance between two observation points. It is also possible to compare the volume of intersection and the entire sphere. In this case, the similarity is given through

$$\frac{\text{Volume of the intersection}}{\text{Volume of a full sphere}} = \frac{\frac{\pi}{12}(4r+d)(2r-d)^2}{\frac{4}{3}\pi r^3} = \frac{(4r+d)(2r-d)^2}{16r^3}. \quad (3.2)$$

Fig. 3-4 – Sensor range and distance travelled; intersection of two spheres is the common volume visible from both view-points.

For all the scenarios studied in this thesis, sensor range is 30 meters and the distance between two consequent samples are in scale of few meters (e.g. 4 to 7 meters). Therefore, with the values assumed for r and d , the anticipated similarity between two consequent observations made from the same place can be roughly estimated. This gives a clue for selection of proper similarity threshold for comparison between two DNAs. Certainly, since the ideal situation cannot be assumed in real application, the selected threshold value should be closer to the lower band of the anticipated similarity value range.

In order to compare two DNAs, a similarity measure is required. If the similarity measure is above the anticipated similarity threshold, compared DNAs are considered similar.

3.3 Similarity measure

Descriptor comparison is the procedure that influences the entire similarity analysis. Therefore, it is very important to think of a proper comparison method. Comparison method should be selected based on the nature of the descriptor. While one comparison method might have proven best performance for one type of descriptor, it might not be necessarily the best choice for another type.

3.3.1 Common comparison techniques

Among the comparison methods used in the literature, bag-of-words model [154] is widely used to compare descriptors. In feature-based description, vector of feature descriptors represents one code-word, and a codebook is the collection which converts code-words as a text, to numeric vectors as descriptors, and vice versa. This way, each image is mapped to a collection of words with no specific order. When two images are compared, the words are compared, and the number of matching words is the similarity measure for comparison. Bag-of-words is a special case of N-gram model [155] which takes into account the sequence of N words instead of comparing unordered words. This way, the order of features contributes to comparison as well.

Since features are out of the debate for the proposed description technique, there is no need to maintain the conversion database (e.g. codebook in case of using bag-of-words model). This database is reported to have included 100 million code word-to-feature records in 20km traversal [91]. A typical code-word is consisted of 128 floating point numbers [45]. Therefore, this model confronts scalability concerns in outdoor applications in the long run.

Comparison method is very much related to the descriptor type. For example, work [84] studies the sequence of features and introduces a modified biologically inspired DNA comparison methodology [85] for descriptor comparison. Since place DNA is a sorted list of eigenvalues, with the presented definition, sequence of eigenvalues does not play a meaningful role for place DNA comparison.

3.3.2 Proposed comparison technique

With respect to the explanation provided in section 3.3.1, none of the conventional comparison methods is more preferable to counting matching eigenvalues from inflated DNA comparison. The proposed comparison technique is the enhanced version of the one explained in section 3.1.1 and is described subsequently through an example.

Table 3-5, visualizes the matching eigenvalues of ideal place and rectangle. It is clear from the comparison chart that for every eigenvalue in rectangle DNA, an exact match can be found in ideal place DNA. It can be shown that this is also true for larger eigenvalues of rectangle that are omitted from this example only for illustration simplicity. According to the discussion provided in section 3.1.2, when a place contains

one rectangle, DNA of rectangle is included in the DNA of the place. In this analysis, counting the number of matching eigenvalues concludes to the similarity measure.

```

• INPUT DNA1, DNA2 as array of floating point numbers
• MATCHED_CASES = 0
• TOLERANCE = 0.01
• FOR i = 1 to LENGTH_OF (DNA1)
  ○ FOR j = 1 to LENGTH_OF (DNA2)
    ▪ m = j @ ARGMIN (ABS (DNA1 (i) -DNA2 (j)))
  ○ LOOP j
  ○ FOR k = 1 to LENGTH_OF (DNA1)
    ▪ n = k @ ARGMIN (ABS (DNA1 (k) -DNA2 (m)))
  ○ LOOP k
  ○ IF (n=i) THEN
    ▪ IF (ABS (DNA1 (i) -DNA2 (m)) <
      TOLERANCE * ((DNA1 (i) +DNA2 (m)) /2) THEN
      • MATCHED_CASES ++;
    ▪ ENDIF
    ▪ REMOVE (DNA1 @ i)
    ▪ REMOVE (DNA2 @ m)
    ▪ i = 0
  ○ ELSE
    ▪ i = n
    ▪ CONTINUE
  ○ ENDIF
• LOOP i
• OUTPUT MATCHED_CASES

```

Pseudocode 3-1 – Pseudocode for calculation of the number of matched eigenvalues

Table 3-6, visualizes the eigenvalues of ideal place and DNA of a noisy rectangle. It can be seen that no exact matches can be found. According to DNA property 2 and property 3, since ideal place contains a rectangle, it is expected that for every eigenvalue of noisy rectangle DNA, a close (but not equal) eigenvalue can be found in ideal place. In this case, counting the exact matches does not help. Instead, matching eigenvalues of inflated DNAs are counted. However, inflation intensity is not fixed for all the eigenvalues. Pseudocode 3-1 illustrates the comparison procedure through assignment of inflation value to eigenvalues. In this procedure, first step is to identify the closest pairs of eigenvalues between the compared DNAs. For all the eigenvalues of one DNA, the closest eigenvalue from the other DNA is found. Then the median μ of each pair is calculated. If the eigenvalues of the pair are within the distance of “ $\varepsilon\%$ of μ ” from μ , the pair denotes a matched case. In this thesis, ε equals 1 for all real situation case studies. “Matched cases” variable, counted according to Pseudocode 3-1, is taken as the

similarity measure. This value indicates the similarity between any arbitrary DNA1 and DNA2. It is possible to normalize the resultant number by total number of eigenvalues in DNAs. This way, the similarity value can be represented in percentage and so the anticipated similarity introduced in section 3.2 can be used as the threshold for detection of two similar DNAs.

Table 3-5 – Comparison between ideal place DNA and rectangle DNA

Table 3-6 – Comparison between ideal place DNA and noisy rectangle DNA

3.4 DNA Length

The more complex the surface is, the longer the list of eigenvalues should be for sufficient description of the surface. However, experiments show (also reported in [143]) that a cut-off from the list of eigenvalues can approximate the DNA. Since finding all the eigenvalues of a large point cloud can be computationally expensive, approximated DNA with limited number of eigenvalues is preferred. However, very short DNA cannot fully represent the content of the studied volume. Proper length of DNA for optimized performance depends on the application, digitization resolution, and the complexity of the studied volume. This length should be found through trial and error. According to the sensory system and down sampling value applied in the experiments presented in this thesis, length of 1500 to 2000 non-zero eigenvalues is found to produce acceptable approximation. Experiments with lengthier DNA did not change the result significantly, and the shorter DNA could not differentiate the places.

3.5 General parameters

For all the experiments presented in this thesis, the set of parameters are tuned according to the guidelines explained so far, and listed in Table 3-7. Other parameters are specified for each experiment based on experiment setup.

Table 3-7 – General experiment parameters

Parameter	Value
Down sampling / voxelization grid size	30cm
Minimum point cluster size /outlier removal	100 elements
Eigenvalue comparison tolerance	1% of the average size of closest eigenvalue pairs
DNA Length	1500 to 2000 of non-zero eigenvalues

3.6 Consistency of DNA

In this section, consistency of DNA for different observations made from the same place at different observation points is studied. Fig. 3-7 depicts four different observations from a place shown in Fig. 3-5. The robot carrying the sensory system,

introduced in section 2.3.2, travelled along the pavement next to a building. Robot scanned the environment every few meters. DNAs of all scans are plotted in Fig. 3-6.

Fig. 3-5 – Pavement; this place is also studied as place 3 in 3.7

Fig. 3-6 – DNA plots of four observation depicted in Fig. 3-5. Each DNA is consisted of smallest 1500 non-zero eigenvalues.

Fig. 3-7 – Four different observations from the same place. Observation points are few meters distant from each other.

DNAs are compared according to the similarity measure explained in Pseudocode 3-1 and similarity percentages are reported in Table 3-8.

Table 3-8 – Similarity among DNAs of different place observations shown in Fig. 3-6.

	Scan 1	Scan 2	Scan 3	Scan 4
Scan 1	100	83.8	87.53	82.8
Scan 2	83.8	100	84.53	86.6
Scan 3	87.53	84.53	100	83.67
Scan 4	82.8	86.6	83.67	100

The values in Table 3-8 are in the anticipated similarity band introduced in section 3.2.

This result confirms that the DNAs of several observations made from the same place are consistent.

3.7 Place characterization through DNA analysis

This section explains how DNA is used to characterize individual places. Table 3-9 depicts observations from three different places shown in Fig. 3-8, Fig. 3-9, and Fig. 3-5. The latter is the same place studied in section 3.6.

Fig. 3-8 – Studied place 1; roofed drop off zone

Fig. 3-9 – Studied place 2; between residential buildings

Table 3-9 – Multiple observations from three different places

Place1_1	Place2_1	Place3_1

Place1_2	Place2_2	Place3_2

Place1_3	Place2_3	Place3_3

Place1_4	Place2_4	Place3_4

DNAs of all the observation shown in Table 3-9 are plotted in Fig. 3-9.

Fig. 3-10 – DNAs of observations made from studied places 1, 2, and 3

It can be seen from the plots in Fig. 3-10 that DNAs of different observations from each of the three studied places demonstrate similar and consistent growth rate pattern. The growth rate pattern is notably different for DNAs of different places. Similar growth rate pattern results to higher number of matching eigenvalues. Therefore, it is expected that the DNAs calculated for observation from the same place show higher similarity percentage when compared against each other. Likewise, it is expected that the DNAs of observations from different places show lower number of matching eigenvalues. Similarity between all the observations made from all three places is calculated and reported in Table 3-10.

The values in Table 3-10 support the fact that DNAs of the observations made from the same place are more similar to each other than to the DNAs of observations made from other places. This means, DNA of a place is a proper measure to characterize the place.

3.8 Deformation of DNA along the continuous motion

Although the result is promising in the sense that it shows DNA analysis can be used to classify observations from different places, the test still does not reflect the performance of DNA analysis in a continuous travel throughout several places. In another word, since change of environment is gradual in a continuous travel, one may

question the possibility of incorporation of DNA analysis to identify individual places in a continuous travel. For example, the last observation made from the second place studied in section 3.7, is slightly different from the other three observations since a new building enters the sensor measurement range. Although the DNA still holds majority of the place information, it should also reflect the entrance of the newly visited major element. Should this observation be taken as part of the second studied place, or should it be taken as the transition frame to a new place depends on the specified similarity threshold. Detailed analysis of transition detection is provided in chapter 4 of this thesis.

Table 3-10 – Similarity of observations made from three different places.

	Place1_1	Place1_2	Place1_3	Place1_4	Place2_1	Place2_2	Place2_3	Place2_4	Place3_1	Place3_2	Place3_3	Place3_4
Place1_1	100.00	85.87	85.67	85.13	52.67	52.20	53.20	56.40	34.33	35.73	33.87	35.67
Place1_2	85.87	100.00	86.00	85.27	52.13	52.93	53.80	56.93	34.33	35.93	34.20	36.07
Place1_3	85.67	86.00	100.00	85.60	52.47	52.13	53.20	56.53	34.27	35.73	33.53	35.53
Place1_4	85.13	85.27	85.60	100.00	51.80	51.80	52.73	56.27	33.13	35.00	33.40	35.33
Place2_1	52.67	52.13	52.47	51.80	100.00	83.67	82.00	78.40	60.80	62.93	61.13	64.13
Place2_2	52.20	52.93	52.13	51.80	83.67	100.00	82.07	77.47	61.20	62.87	61.20	64.07
Place2_3	53.20	53.80	53.20	52.73	82.00	82.07	100.00	78.73	59.40	62.33	59.80	62.47
Place2_4	56.40	56.93	56.53	56.27	78.40	77.47	78.73	100.00	55.07	57.60	55.73	58.20
Place3_1	34.33	34.33	34.27	33.13	60.80	61.20	59.40	55.07	100.00	83.80	87.53	82.80
Place3_2	35.73	35.93	35.73	35.00	62.93	62.87	62.33	57.60	83.80	100.00	84.53	86.60
Place3_3	33.87	34.20	33.53	33.40	61.13	61.20	59.80	55.73	87.53	84.53	100.00	83.67
Place3_4	35.67	36.07	35.53	35.33	64.13	64.07	62.47	58.20	82.80	86.60	83.67	100.00

3.9 Second characterization case study

In section 3.7, studied places are too different. Therefore, the difference between places are clear. In another experiment, two sets of scans were captured from two different lecture theatres containing similar objects in different settings. The first lecture theatre

(LT 28) has rectangular floor with seats placed along straight lines in rows. There are 17 seats in each row. The floor of the other lecture theatre (LT 26) is a sector of a circle with the board mounted near the center and seats placed along the sectors in the way that the closer row to the board has 10 seats and the furthest row has more than 40 seats. Similarity comparison result is reported in Table 3-11. Although in this study, places contain similar objects, the geometry of the rooms are quite different. Therefore, the difference is still clear. This can be verified from the numbers reported in Table 3-11 too. Panoramic view photos of the studied places are shown in Table 3-12.

Table 3-11 – Similarity of observations made from LT 26 and LT 28

	LT 26 Scan 1	LT 26 Scan 2	LT 26 Scan 3	LT 26 Scan 4	LT 26 Scan 5	LT 26 Scan 6	LT 28 Scan 1	LT 28 Scan 2	LT 28 Scan 3	LT 28 Scan 4	LT 28 Scan 5
LT 26 Scan 1	100.00	92.80	88.00	88.67	92.13	90.67	75.00	68.60	78.13	81.20	77.93
LT 26 Scan 2	92.80	100.00	90.80	90.53	89.13	93.07	72.33	65.13	74.93	78.00	74.80
LT 26 Scan 3	88.00	90.80	100.00	93.67	85.33	92.80	68.67	61.87	71.40	74.47	71.40
LT 26 Scan 4	88.67	90.53	93.67	100.00	85.27	93.40	68.07	61.67	71.60	74.27	71.07
LT 26 Scan 5	92.13	89.13	85.33	85.27	100.00	87.20	78.67	71.47	81.40	84.67	81.60
LT 26 Scan 6	90.67	93.07	92.80	93.40	87.20	100.00	70.20	63.60	73.07	76.07	73.07
LT 28 Scan 1	75.00	72.33	68.67	68.07	78.67	70.20	100.00	88.67	91.40	89.07	92.20
LT 28 Scan 2	68.60	65.13	61.87	61.67	71.47	63.60	88.67	100.00	85.40	83.47	85.93
LT 28 Scan 3	78.13	74.93	71.40	71.60	81.40	73.07	91.40	85.40	100.00	92.67	94.67
LT 28 Scan 4	81.20	78.00	74.47	74.27	84.67	76.07	89.07	83.47	92.67	100.00	92.67
LT 28 Scan 5	77.93	74.80	71.40	71.07	81.60	73.07	92.20	85.93	94.67	92.67	100.00

Table 3-12 – Panoramic photos of LT 26 and LT 28

Place	Panoramic photo
LT 26
LT 28

3.10 Third characterization case study

In this study, two lecture theatres with similar geometry were studied. Both studied lecture theatres have rectangular-shaped floor with similar dimensions. One lecture

theatre (LT 24) has 18 seats in each row and the other (LT 28) has 17 seats in each row and an extra walkway between seats. The difference is quite invisible from sensor perspective when scans are captured from the walkway between seats. This can be verified from the similarity values reported in Table 3-13 as well. In this study, all the scans made from two different lecture theatres show high level of similarity to each other in the way that the analysis does not help classifying the places. Panoramic view photos of the studied places are shown in Table 3-14.

Table 3-13 – Similarity of observations made from LT 24 and LT 28

	LT 24 Scan 1	LT 24 Scan 2	LT 24 Scan 3	LT 24 Scan 4	LT 28 Scan 1	LT 28 Scan 2	LT 28 Scan 3	LT 28 Scan 4	LT 28 Scan 5
LT 24 Scan 1	100.00	95.07	83.53	91.33	86.00	79.40	89.40	92.47	89.53
LT 24 Scan 2	95.07	100.00	82.67	90.87	85.40	78.33	88.73	92.00	88.60
LT 24 Scan 3	83.53	82.67	100.00	87.80	92.47	91.80	89.60	87.27	89.80
LT 24 Scan 4	91.33	90.87	87.80	100.00	90.20	84.00	92.93	94.87	93.20
LT 28 Scan 1	86.00	85.40	92.47	90.20	100.00	88.67	91.40	89.07	92.20
LT 28 Scan 2	79.40	78.33	91.80	84.00	88.67	100.00	85.40	83.47	85.93
LT 28 Scan 3	89.40	88.73	89.60	92.93	91.40	85.40	100.00	92.67	94.67
LT 28 Scan 4	92.47	92.00	87.27	94.87	89.07	83.47	92.67	100.00	92.67
LT 28 Scan 5	89.53	88.60	89.80	93.20	92.20	85.93	94.67	92.67	100.00

Table 3-14 – Panoramic photos of LT 24 and LT 28

Place	Panoramic photo
LT 24
LT 28

Although place classification can be made through DNA comparison when places are different, classification of too similar places, which are not exactly the same, is not always possible only by comparing DNA structures. In this situation, higher resolution

is required to capture more details from the scene. Please note that the motivation for this research is to study the places visited during a continuous travel. Analysis of “DNA structure changes along the sequence of places” helps to identify too similar places that are not identical. This analysis is explained in chapter 5 of the thesis and a similar situation is studied in section 5.7 when two similar corridors (office areas) are identified as different places through sequence analysis in a continuous travel through several places.

3.11 Chapter conclusion

Material covered in this chapter prepared the fundamentals of incorporating DNA analysis for place detection. Below is the list of main results of the chapter, which facilitate the further place detection analysis presented in the thesis.

- a) For comparison of two studied DNAs, closest pairs of eigenvalues within an acceptable tolerance are counted. Percentage of similar eigenvalues is used as the similarity measure.
- b) Place DNA is consistent among all the observations made from the same place.
- c) Place DNA is a compressed description of the studied place that encapsulates intrinsic information and uniquely characterizes individual places.

However, DNA of two different places may show relatively high level of similarity. Although this can be taken as an opportunity for identification of “similar, but not identical” places, it might sound as a warning for false place recognition as well. Although transition frames are good examples of this situation, other scenarios, such as the one explained in section 3.10, are also possible. Chapter 5 introduces sequence analysis to minimize false recognition in similar situations.

Chapter 4

Environment Segmentation

In chapter 3, characterization of places using their DNA descriptors was explained. Classification of DNAs of several places was studied in section 3.7. However, presented experiment only included separate places with zero overlapping areas. In this chapter, environment segmentation to multiple meaningful subsets (places) is discussed. Experiments presented in this chapter include the cases of overlapping areas among the places. Considerations regarding transition frames are discussed in details. Please note that in this thesis, visible surrounding region of the experiment trail is referred to as the studied environment. Segmentation of environment to meaningful subsets is the primary step for recognition of revisited places. Recognition of revisited places is covered in chapter 5 of the thesis.

4.1 Distinctive subsets vs gradual changes

While moving along a place, all the major surrounding objects change gradually. No new major object is expected to be seen before the robot enters a new place. Therefore, according to DNA Property 2, changes in DNA elements are up to the extent that the consistency of DNA structure is maintained for all different observations from the place. However, when a new major object enters the scene, DNA structure starts to transform to a new consistent structure. This change stands for detection of transition to a new place. It is expected for DNA changes to be gradual for the observations made about the border where the change of place occurs. This gradual change brings into the picture a concern for differentiating neighbouring places. Since every two neighbouring observations are slightly different, separation of adjacent places seems to be a challenge. In short, the question is:

“How is it possible to distinguish distinctive places while they contain partial overlaps that cause gradual DNA transformation?”

Here, the delicate point relies in different types of DNA changes. There are two major types of DNA changes:

- DNA changes which are caused by observation point changes within a place.
- DNA changes which are caused by change of major sub-parts of the scans.

The first group of changes do not transform the DNA structure. This can be seen in the experiments presented in chapter 3. However, second group of changes produce transition frames between two adjacent places. DNAs transform from one stable structure to another stable structure through one or few transition frames. At transition stage between two adjacent places, it is expected to see the similarity pattern shown in Table 4-1.

Table 4-1 – Pattern of table of similarities in presence of overlapping sub-parts

	Place1_1	Place1_2	Place1_3	Place1_4	T-Frame 1	T-Frame 2	Place2_1	Place2_2	Place2_3	Place2_4
Place1_1	Identical	Very similar	Very similar	Very similar	Similar	Similar	Different	Different	Different	Different
Place1_2	Very similar	Identical	Very similar	Very similar	Similar	Similar	Different	Different	Different	Different
Place1_3	Very similar	Very similar	Identical	Very similar	Similar	Similar	Different	Different	Different	Different
Place1_4	Very similar	Very similar	Very similar	Identical	Similar	Similar	Different	Different	Different	Different
T-Frame 1	Similar	Very similar	Very similar	Very similar	Identical	Similar	Similar	Similar	Similar	Similar
T-Frame 2	Similar	Similar	Similar	Similar	Similar	Identical	Similar	Similar	Similar	Similar
Place2_1	Different	Different	Different	Different	Similar	Similar	Identical	Very similar	Very similar	Very similar
Place2_2	Different	Different	Different	Different	Similar	Similar	Very similar	Identical	Very similar	Very similar
Place2_3	Different	Different	Different	Different	Similar	Similar	Very similar	Very similar	Identical	Very similar
Place2_4	Different	Different	Different	Different	Similar	Similar	Very similar	Very similar	Very similar	Identical

Identical
Very similar
Similar
Less similar
Different

Number of transition frames depends of the amount of overlap between two adjacent places. Since in the experiments of chapter 3, studied places had zero overlaps, sharp changes occur between different groups of observations in the table of similarities and no transition frames are seen. The situation is not expected to be the same when a mobile robot travels along a path including several distinct places. When the explored environment is large enough, several places are expected to be visited during the mission. Many of these places have overlapping areas. In another word, in a mobile robotic mission, some observations are expected to belong to bordering areas between adjacent places. The DNA of such data frames includes information about the

intersecting places. This DNA shows high level of similarity to all intersecting places, but this value is not as high as the similarity values among the non-bordering DNAs. By the way, even in a continuous travel, it is possible that the overlapping sub-parts between two adjacent places would be very brief. One such example is the transition from one area to another area through a separating door. In this situation, no transition frames are expected between two adjacent places.

4.2 Marking a place in table of similarities

According to the definition, a place is a subset of environment in which all the descriptors generated from observations made from different locations of the subset follow a consistent structure. In a continuous travel along a path, after registering the similarity values in the table of similarities through sporadic sampling, a place can be identified as a square sub-matrix along the main diagonal of the table of similarities. The value of all the elements of such sub-matrix are greater than a specified threshold.

In experiments presented in this thesis, sporadic sampling stands for scanning the environment every 4~7 meters.

Definition - Place matrix

- Place matrix is a square matrix and represents multiple observations made from the same place at different viewpoints. Dimension of the place matrix equals the number of observations made from the represented place. Place matrix appears as a sub-matrix along the main diagonal of the similarity matrix. Values of all the elements of a place matrix are greater than a specified similarity threshold value.

Depending on the selected similarity threshold value, some or all of the transition frames can be included or excluded from the place matrix. Fig. 4-1, Fig. 4-2, and Fig. 4-3 show how three different choices of similarity threshold affect place matrices. Increasing the threshold value results to marking smaller places. On the other hand, decreasing the threshold, helps visualizing the overlapping sub-parts of adjacent places. However, threshold value is only used to mark the boundaries of the place matrices. It does not affect further analysis on the table. Therefore, sensitivity of place detection analysis to similarity threshold value is very minimal.

4.3 Segmentation case study

In this section the procedure for environment segmentation to meaningful subsets (places) is studied through an experiment.

A robot moves along a path shown in Fig. 4-4. Travel length is greater than 200 meters. Robot scanned the environment every 4~6 meters. Total of 53 observations were made during this experiment. Fig. 4-5 depicts a snapshot from the start point of the experiment.

Fig. 4-4 – The path travelled by the robot in segmentation experiment

Fig. 4-5 – Starting point and a snapshot from the environment where segmentation experiment was conducted

DNA of each observation is consisted of 2000 non-zero eigenvalues. Similarity values are calculated using 1% tolerance. Similarity threshold is taken 82%, which is a typical value in the range of anticipated similarity (80%~87%) calculated according to section

[illegible]

76

real environment map (compare Table 4-3 with Fig. 4-6). Table 4-4 explains how place matrices correspond to meaningful places on the real map.

Table 4-3 – Table of similarities of segmentation experiment after merging highly overlapped place matrices, 8 place matrices are identifiable in the table.

	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8	Column 9	Column 10	Column 11	Column 12	Column 13	Column 14	Column 15	Column 16	Column 17	Column 18	Column 19	Column 20	Column 21	Column 22	Column 23	Column 24	Column 25	Column 26	Column 27	Column 28	Column 29	Column 30	Column 31	Column 32	Column 33	Column 34	Column 35	Column 36	Column 37	Column 38	Column 39	Column 40	Column 41	Column 42	Column 43	Column 44	Column 45	Column 46	Column 47	Column 48	Column 49	Column 50	Column 51	Column 52	Column 53	Column 54	Column 55	Column 56	Column 57	Column 58	Column 59	Column 60	Column 61	Column 62	Column 63	Column 64	Column 65	Column 66	Column 67	Column 68	Column 69	Column 70	Column 71	Column 72	Column 73	Column 74	Column 75	Column 76	Column 77	Column 78	Column 79	Column 80	Column 81	Column 82	Column 83	Column 84	Column 85	Column 86	Column 87	Column 88	Column 89	Column 90	Column 91	Column 92	Column 93	Column 94	Column 95	Column 96	Column 97	Column 98	Column 99	Column 100																					
Row 1	2000	1855	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251	263	275	287	299	311	323	335	347	359	371	383	395	407	419	431	443	455	467	479	491	503	515	527	539	551	563	575	587	599	611	623	635	647	659	671	683	695	707	719	731	743	755	767	779	791	803	815	827	839	851	863	875	887	899	911	923	935	947	959	971	983	995	1000															
Row 2	1695	2000	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251	263	275	287	299	311	323	335	347	359	371	383	395	407	419	431	443	455	467	479	491	503	515	527	539	551	563	575	587	599	611	623	635	647	659	671	683	695	707	719	731	743	755	767	779	791	803	815	827	839	851	863	875	887	899	911	923	935	947	959	971	983	995	1000															
Row 3	1517	1771	2000	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251	263	275	287	299	311	323	335	347	359	371	383	395	407	419	431	443	455	467	479	491	503	515	527	539	551	563	575	587	599	611	623	635	647	659	671	683	695	707	719	731	743	755	767	779	791	803	815	827	839	851	863	875	887	899	911	923	935	947	959	971	983	995	1000														
Row 4	1431	1683	1804	2000	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251	263	275	287	299	311	323	335	347	359	371	383	395	407	419	431	443	455	467	479	491	503	515	527	539	551	563	575	587	599	611	623	635	647	659	671	683	695	707	719	731	743	755	767	779	791	803	815	827	839	851	863	875	887	899	911	923	935	947	959	971	983	995	1000													
Row 5	1243	1495	1616	1740	1860	1979	2000	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251	263	275	287	299	311	323	335	347	359	371	383	395	407	419	431	443	455	467	479	491	503	515	527	539	551	563	575	587	599	611	623	635	647	659	671	683	695	707	719	731	743	755	767	779	791	803	815	827	839	851	863	875	887	899	911	923	935	947	959	971	983	995	1000										
Row 6	1067	1319	1440	1564	1683	1804	1923	2000	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251	263	275	287	299	311	323	335	347	359	371	383	395	407	419	431	443	455	467	479	491	503	515	527	539	551	563	575	587	599	611	623	635	647	659	671	683	695	707	719	731	743	755	767	779	791	803	815	827	839	851	863	875	887	899	911	923	935	947	959	971	983	995	1000									
Row 7	891	1143	1264	1385	1507	1628	1749	1869	1989	2000	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251	263	275	287	299	311	323	335	347	359	371	383	395	407	419	431	443	455	467	479	491	503	515	527	539	551	563	575	587	599	611	623	635	647	659	671	683	695	707	719	731	743	755	767	779	791	803	815	827	839	851	863	875	887	899	911	923	935	947	959	971	983	995	1000							
Row 8	715	967	1088	1209	1329	1449	1569	1689	1809	1929	2000	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251	263	275	287	299	311	323	335	347	359	371	383	395	407	419	431	443	455	467	479	491	503	515	527	539	551	563	575	587	599	611	623	635	647	659	671	683	695	707	719	731	743	755	767	779	791	803	815	827	839	851	863	875	887	899	911	923	935	947	959	971	983	995	1000						
Row 9	539	791	912	1033	1154	1274	1395	1515	1635	1755	1875	1995	2000	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251	263	275	287	299	311	323	335	347	359	371	383	395	407	419	431	443	455	467	479	491	503	515	527	539	551	563	575	587	599	611	623	635	647	659	671	683	695	707	719	731	743	755	767	779	791	803	815	827	839	851	863	875	887	899	911	923	935	947	959	971	983	995	1000				
Row 10	363	615	736	857	978	1098	1219	1339	1459	1579	1699	1819	1939	2000	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251	263	275	287	299	311	323	335	347	359	371	383	395	407	419	431	443	455	467	479	491	503	515	527	539	551	563	575	587	599	611	623	635	647	659	671	683	695	707	719	731	743	755	767	779	791	803	815	827	839	851	863	875	887	899	911	923	935	947	959	971	983	995	1000			
Row 11	187	439	560	681	802	923	1043	1164	1284	1404	1524	1644	1764	1884	1999	2000	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251	263	275	287	299	311	323	335	347	359	371	383	395	407	419	431	443	455	467	479	491	503	515	527	539	551	563	575	587	599	611	623	635	647	659	671	683	695	707	719	731	743	755	767	779	791	803	815	827	839	851	863	875	887	899	911	923	935	947	959	971	983	995	1000	
Row 12	9	360	481	602	723	843	964	1084	1204	1324	1444	1564	1684	1804	1924	1999	2000	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251	263	275	287	299	311	323	335	347	359	371	383	395	407	419	431	443	455	467	479	491	503	515	527	539	551	563	575	587	599	611	623	635	647	659	671	683	695	707	719	731	743	755	767	779	791	803	815	827	839	851	863	875	887	899	911	923	935	947	959	971	983	995	1000
Row 13	23	444	565	686	807	928	1048	1169	1289	1409	1529	1649	1769	1889	1999	2000	1771	1683	1595	1507	1419	1331	1243	1155	1067	979	891	803	715	627	539	451	363	275	187	99	11	23	35	47	59	71	83	95	107	119	131	143</

Fig. 4-6 – Eight place matrices matched with the corresponding places on the map

Table 4-4 – Place matrices and the corresponding places on the map

Place matrix number	A photo from the place	Description on the real map
1		Fire fighter zone between building and open-air parking
2		Under the building
3		Entrance of open-air parking
4		On the road between the elevated ground and entrance of roofed parking
5		On the road between the building and elevated ground

6		On the road between the building and bus-stop
7		On the road between the building and excavated ground
8		On the road between the sheltered parking and excavated ground

4.4 Segmentation procedure

Having the table of similarities as the input, environment segmentation to meaningful places is the procedure shown in Fig. 4-7.

4.4.1 Sensitivity to similarity threshold

As explained before, change of similarity threshold can slightly affect this result in the sense that some scans made at the border of places (transition frames) get excluded from place matrices, or oppositely, overlap between adjacent place matrices is increased. However, these changes are minimal and they only affect the bordering elements of place matrices.

4.4.2 Scan labels

The labels used for scans in all the tables of similarities studied so far contained no specific information about the time or geo-location at which the scan was captured. By inclusion of such information in labels, more information can be extracted from the table. For example, if scans are time-stamped, it is possible to check for the duration of travel inside a specific place. If scans are geo-tagged (e.g through GPS readings), knowing the scan range of the sensor (30 meter in the tests conducted in this thesis),

3D geometry of the places can be roughly estimated. Moreover, geo-tags can help auto labelling the place matrices by automatically corresponding the scans to map semantics.

Fig. 4-7 – Procedure for environment segmentation to meaningful places

4.5 Chapter summary

Environment segmentation incorporates the characterization technique, introduced in previous chapters, to direct the research towards real application. Through the segmentation case study, the first tangible outcome of the earlier definitions and procedures was explained. Place matrices in the table of similarities correspond to meaningful places on the map. Places identified on the map agree with the definition provided in section 1.3. Environment segmentation is resulted from the similarity analysis among the DNAs of different observations from the same place. Similarity among different places is studied in chapter 5.

Chapter 5

Detection of Similar and Identical Places through Sequence Analysis

Chapter 4 explained the procedure of dividing the studied environment to meaningful places by identifying place matrices. Place matrices hold the similarity analysis results from the comparison between DNAs of multiple observations from a place. Therefore, place matrices describe self-similarity among observations made from the same place at different viewpoints.

In this chapter, similarity between different places is studied. Although two places might have similar DNAs, they might be identical places or they might be highly similar places. This concern was earlier noticed in section 3.10, where two similar lecture-theatres were compared. Identification of revisited places is a critical functionality for autonomous mobile robots. Therefore, it is necessary to make distinction between similar and identical places. The proposed solution is to study a place as a part of a sequence. Place DNAs transform gradually from one stable structure to another stable structure when robot travels from one place to another. Pattern of DNA changes carries valuable information. Analysis of this pattern is used as the mechanism for making distinction between similar and identical places.

Analysis of sequences can be found in related research work such as [86, 156] in which camera frame sequences were used to enhance SLAM. However, in the mentioned works and other similar works, speed of vehicle is a concern which demonstrates the sensitivity of such analysis to scan-rate. However, the sequence analysis presented in this chapter includes the analysis of sub-matrices of the table of similarities, which can be compared even when scan-rate changes throughout the mission.

5.1 Cross similarity

Table of similarities encapsulates similarity values resulted from comparison of all observation made from the studied environment. It is expected for the similarity values on the main diagonal of this table to be exactly 100% since diagonal elements represent self- similarity calculation. However, any off-diagonal value is the result of comparison between two different scans. It is a common practise to form table of similarities based on some similarity measure and look for large off-diagonal elements in the table of similarities to pin point highly similar scans [88]. However, scan to scan comparison is prone to false positive detections (wrong detection of two scans as similar). In this thesis, proposed similarity analysis is not based on processing individual scans. Instead, place matrices are compared. Because place matrices encapsulate similarity comparison among several observations from each place, comparing place matrices is expected to be less prone to false detection. While place matrices, introduced in section 4.2, describe self-comparison (comparing DNAs of different observation from the same place), there are off-diagonal matrices which describe the cross similarity between two different places.

Definition - Cross similarity matrix

- A cross similarity matrix encapsulates DNA comparison result between each of the observations made from one place to every observation made from another place. Cross similarity matrix (not necessarily square-matrix) is an off-diagonal sub-matrix of the table of similarities. Cross similarity matrix between places A and B , is consisted of the elements of the table of similarities in the same rows of place matrix A , and in the same columns of place matrix B .

Cross similarity matrix describes the similarity between two different places. Fig. 5-1 illustrates cross similarity matrix location in the table of similarities. All place matrices and three of the six possible cross similarity matrices are shown in Fig. 5-1.

Since the table of similarities is symmetric, cross similarity matrices of upper triangle could be found in lower triangle (transposed) as well.

In Fig. 5-1, $C_{(1)(4)}$ is the only cross similarity matrix that all its elements are greater than similarity threshold. This means, all the observations made from place 1 are highly similar to all the observations made from place 4 (true for all 12 possible combination

of observation pairs). Likewise, none of the observations made from place 1 are highly similar to any of the observations made from place 3 (true for all 9 possible combination of observation pairs included in $C_{(1)(3)}$). Therefore, it could be inferred from Fig. 5-1, that place 1 and place 4 are similar. Whether these places are identical or not can be determined through sequence analysis, which is discussed later in this chapter.

Fig. 5-1 – Place matrices 1, 2, 3, and 4, along with three of the six possible cross similarity matrices $C_{(1)(3)}$, $C_{(1)(4)}$, and $C_{(3)(4)}$.

5.2 Identification of similar places; a case study

In this section, similar places which may represent revisited places are identified in an experiment. Fig. 5-2 shows a scenario in which a mobile robot enters a building from the main entrance, exits the building from the opposite side, turns around the building, enters the building from another entrance, travels along part of the building which was

travelled before, exits the building again, and proceeds to the road. Total distance travelled is more than 350 meters. 85 scans were captured during this experiment.

Fig. 5-2 – Experiment for identification of similar and identical places

In this experiment, it is desired to study the table of similarities to pinpoint the place matrices that are revisited place candidates. Table 5-1 shows the table of similarities for this experiment. All place matrices and the cross similarity matrices which contain elements larger than the similarity threshold are marked. Table 5-1 shows the sub-matrices after merging the ones with more than 50% overlap. 1700 is taken as the similarity threshold (85%) and rest of the experiment parameters are identical to the experiment conducted in section 4.3. Although not all cross similarity matrices correspond to revisited places, they are the clues for identification of revisited places. For example, since $C_{(2)(5)}$ contains above the threshold values, places 2 and 5 are similar and might represent a revisited place. This means these two place matrices are similar but the corresponding places might or might not be identical. Please note that if individual scans were used for identification of revisited places, many candidates could be found. Through processing the cross similarity matrices, search scope is limited to few possibilities.

Table 5-1 – Table of similarities for the experiment that includes revisited places.

Figure 1 displays a large grid of numbers, representing a 2D lattice. The grid is divided into several regions by colored lines and labels. The regions are labeled with $C_2(5)$, $C_2(10)$, $C_3(6)$, $C_4(9)$, $C_5(11)$, $C_5(10)$, $C_6(9)$, $C_7(12)$, and 10. The grid is also labeled with numbers 1 through 11 in the top left corner. The grid is divided into several regions by colored lines and labels. The regions are labeled with $C_2(5)$, $C_2(10)$, $C_3(6)$, $C_4(9)$, $C_5(11)$, $C_5(10)$, $C_6(9)$, $C_7(12)$, and 10. The grid is also labeled with numbers 1 through 11 in the top left corner.

5.3 Sequences

Up to this point, self-similarity and cross-similarity were discussed. Both taking into account only the elements of the table with higher-than-threshold values. However, the table of similarities holds additional information in the elements containing lower similarity values. Low-value elements of the table of similarities can contribute to detection of place matrices that represent revisited places. These elements are studied as elements of a sequence matrix.

Definition - Sequence matrix

- A sequence matrix represents the scenario of traveling through multiple places. Sequence matrix is a square-matrix along the main diagonal of the table of similarities. It contains multiple consequent place matrices as well as the transition frames between them. In table of similarities, a sequence matrix starts from the first element of one place matrix and terminates to the last element of the last place matrix of the sequence.

Sequence of place matrices includes two or more place matrices. Fig. 5-3 illustrates place matrices and some of the possible sequences of place matrices. Elements of the sequence matrix can hold large or small values. Large values are mainly the ones inside place matrices. Most of the other elements of the sequence matrix have lower-than-threshold values.

All the analysis presented so far operated on the elements of the table of similarities as if the values are either higher-than-threshold or lower-than-threshold. In this sense, each matrix element was treated as a binary number. It is very important to note that the values inside the table of similarities are numbers ranging zero to DNA length (or ranging 0 to 100 when values are normalized). Therefore, the table of similarities in this sense is similar to a grayscale bitmap in which the intensity value of each pixel equals the similarity value stored in the corresponding matrix element.

It is also possible to compare sequence of places through analysis of cross similarity among them.

Definition - Cross sequence matrix

- Cross sequence matrix represents similarity between two scenarios. In table of similarities, Cross sequence matrix is an off-diagonal sub-matrix of similarity matrix and contains a sequence of cross similarity place matrices. It starts from the first element of one cross similarity place matrix and terminates to the last element of the last cross similarity place matrix in a sequence.

To better illustrate the role of sequence matrices in place recognition analysis, experiment of section 5.1 is further studied in the subsequent section. Table 5-2 shows two possible sequence matrices and one cross sequence matrix.

Fig. 5-3 – Place matrices 1, 2, 3, and 4, along with four sequences of place matrices $S_{(1)(2)}$, $S_{(2)(3)}$, $S_{(3)(4)}$, and $S_{(1)(2)(3)}$.

5.3.1 Proper sequences

It is possible to identify many sequences. However, there are only few cases which contribute to identification of revisited places. These proper sequences are described here. The procedure starts from a cross sequence matrix. Assume an arbitrary cross sequence matrix $CS_{(r1)(c1)-...-(rn)(cn)}$ is a sequence of n cross similarity matrices where r indices refer to row contributing place matrices and c indices refer to column contributing place matrices.

Table 5-2 – Table of similarities with sequence matrices and a cross sequence matrix.

Proper selection of cross sequence matrix is the one with similar r and c sequential patterns. For example r indices 3, 4, 5 are three consequent indices and c indices 9, 10, 11 are also three consequent indices. Therefore, $CS_{(3)(9)-(4)(10)-(5)(11)}$ is a proper cross sequence matrix which should be further analysed. However, r indices 3, 5, 6 and c indices 9, 13, 12 do not follow any specific orders in row indices nor in column indices. Therefore, $CS_{(3)(9)-(5)(13)-(6)(12)}$ is not a proper cross sequence matrix.

In Table 5-2, the shown candidate cross similarity matrices take their row numbers from place matrices 2, 3, 4, 5, 6, and 7, and their column numbers from place matrices 6, 9,

10, and 12. Among all the possible combinations, only for the cross sequence $CS_{(4)(9)-(5)(10)}$ the situation is so that the row-contributing place matrices are consequent place matrices (4 and 5), and the column contributing place matrices are consequent place matrices (9 and 10) too. These sequences are proper choices for pattern checking. No other such combinations could be found if other cross sequence matrices are considered. For example, another possible selection, $CS_{(3)(6)-(4)(9)}$, refers to consequent row contributing place matrices (3 and 4), but the column contributing place matrices (6 and 9) are not consequent place matrices. Therefore, $CS_{(3)(6)-(4)(9)}$ is not a proper selection for further analysis.

5.4 Sequence analysis

Please note that in Table 5-2 observations are time stamped and the stamp is shown as the title of each scan. Sequences are compared based on their pattern.

5.4.1 Sequence pattern matching

In order to explain the process of matching the patterns of sequences, experiment of section 5.1 is further studied using the sequences shown in Table 5-2. In this experiment, pattern of $CS_{(4)(9)-(5)(10)}$ should be matched with $S_{(4)(9)}$ and $S_{(5)(10)}$ sequence matrices. One possible matching technique is to calculate matching score through normalized two-dimensional cross-correlation. One implementation of this function is provided in OpenCV Library [157]. The function calculates the cross correlation according to the equation

$$R(x, y) = \frac{\sum_{(x,y)} (T(x,y) \cdot I(x+x, y+y))}{\sqrt{\sum_{(x,y)} (T(x,y))^2 \cdot \sum_{(x,y)} (I(x+x, y+y))^2}}. \quad (5.1)$$

In (5.1), R is a 2D matrix, which holds correlation score when template T slides over image I by x steps along rows and by y steps along columns. Then,

$$\max_{(x,y)}(R) \quad (5.2)$$

outputs the largest element of the matrix R , which indicates the best matching score of template T in image I . It is possible to stretch template T to match the size of image I before calculating the cross correlation if required.

For the conducted experiment, $CS_{(4)(9)-(5)(10)}$ is taken as a grayscale template and is matched with $S_{(4)(9)}$ and $S_{(5)(10)}$ as grayscale bitmaps images. $CS_{(4)(9)-(5)(10)}$ matches to $S_{(4)(9)}$ by 97.26% and to $S_{(5)(10)}$ by 96.93%. Moreover, $S_{(5)(10)}$ is matched to $S_{(4)(9)}$ by 96.69%. Therefore, place matrices 4 and 9 represent a revisited place (identical places), and places 5 and 10 represent another revisited place. Sequences $S_{(4)(9)}$ and $S_{(5)(10)}$ represent revisited scenarios.

$CS_{(2)(5)-(3)(6)}$ is another proper sequence. However, the pattern analysis shows matching score of 77.86% when $CS_{(2)(5)-(3)(6)}$ is compared to $CS_{(2)(3)}$ and it shows only 66.12% matching score when compared to $CS_{(5)(6)}$. Even $CS_{(2)(3)}$ and $CS_{(5)(6)}$ show only 67.16% matching score. Therefore, $CS_{(2)(5)-(3)(6)}$ is not a revisited scenario. Thus, no places included in this sequence are not revisited places.

5.5 Discussion on sequence matching

5.5.1 Selection of templates and images

In real application, many place matrices can be identified in the table of similarities. Therefore, many sequence matrices can be identified as well. However, section 5.3.1 explains the procedure to identify the proper sequence matrices. Search procedure should start from cross sequence matrices and by checking the pattern of the indices of row contributing and column contributing place matrix indices. Through this procedure, pattern checking should be applied to very few templates that meet the specified criteria.

5.5.2 Detection confidence

Table elements greater than similarity threshold can be called “similar observations”, and elements less than the threshold can be called “different observations”. Recognition of revisited places through sequence pattern analysis, involves checking the similarity between place matrices, checking the difference among place matrices, and the pattern of transition frames. Therefore, in the case study of section 5.4, beside the similarity between places, differences between places 4 and 5, and differences between places 9

and 10, contribute to place detection. Therefore, when sequence patterns are studied, similarities and differences among places and transitions are all considered in detection process. For the case of conducted experiment, 30 observations from the environment are involved. 224 similarity values (some higher than threshold and some lower than threshold) and their sequential patterns are processed for recognition of revisited places. In terms of detection confidence, this is an advantage of this method over the methods which detect only through comparison of individual snapshots. Pattern analysis of the cross sequence matrix means to study the pattern of DNA transformations along the mission.

5.5.3 Other template matching techniques

Many other template matching methods exist. Other techniques were not tested in this research. It will be considered as a minor future work.

5.5.4 Reverse sequences

In the conducted experiment $CS_{(6)(9)-(5)(10)}$ is another possible cross sequence matrix which refers to a reverse sequence of row contributing place matrices (6 and 5) and a forward sequence of column contributing place matrices (9 and 10). This is also a proper selection since it refers to the possibility of robot traveling first in one direction, and next time in the opposite direction. In this situation, it is necessary to do template matching with flipped version of the template. In the case of conducted experiment, matching score after flipping the template is less than 65%, which is not high and therefore does not conclude to recognition of revisited places.

5.6 Best matching frames

Although presented detection methodology does not use snapshot matching to detect revisited places, it is possible to find matching snapshots through analysis of identified revisited places. In the conducted experiment, $CS_{(4)(9)-(5)(10)}$ is the cross sequence matrix which holds the information about revisited sequences and places. This matrix is shown in Table 5-3. In this matrix, $C_{(4)(9)}$ and $C_{(5)(10)}$ are the cross similarity places matrices which refer to revisited places. The largest elements along the main diagonal of $C_{(4)(9)}$ and $C_{(5)(10)}$ refer to the frames (snapshots) with highest level of similarity. These elements are shown in Table 5-3. As mentioned earlier, elements of table of similarities

in this experiment are time-stamped. Once the largest element along the main diagonal is found, row label and column label are the times when the matching observations were made.

5.6.1 Diagonal elements of cross similarity place matrices

Cross similarity matrices are not necessarily square matrices. Therefore, diagonal elements are defined through a bilinear formula. For a matrix with m rows and n columns, diagonal element at row i , is in the column j so that

$$j = \text{ROUND}\left(\frac{n}{m}i\right). \quad (5.3)$$

In (5.3), *ROUND* function outputs the nearest integer. In Table 5-3, diagonal elements of cross similarity place matrices are shaded (yellow) and the largest element is shaded darker (red).

Table 5-3 – Cross sequence matrix $CS_{(4)(9)-(5)(10)}$ is used to identify best matching snapshot

	20_55_37	20_56_20	20_57_05	20_57_49	20_58_37	20_59_25	21_00_11	21_00_57	21_01_44	21_02_34	21_03_21	21_04_04	21_04_48	21_05_34
20_16_22	1858	1808	1730	1770	1758	1060	1256	1538	1403	1367	1387	1360	1287	1372
20_17_14	1817	1849	1778	1832	1696	1016	1208	1472	1353	1310	1335	1313	1233	1316
20_18_06	1726	1793	1854	1812	1607	942	1122	1393	1250	1217	1247	1217	1149	1226
20_19_01	1802	1852	1800	1853	1672	998	1187	1468	1329	1283	1310	1285	1208	1290
20_19_57	1825	$C_{(4)(9)}$			1736	1792	1091	1293	1577	1437	1388	1424	1390	1321
20_20_50	1637	1573	1589	1550	1765	1264	1496	1788	1638	1601	1625	1594	1515	1609
20_21_46	1127	1064	1006	1050	1233	1836	1772	1451	1606	1661	1634	1660	1745	1646
20_22_36	1316	1266	1201	1245	1440	1603	1840	1676	1826	1877	1871	1860	1858	1880
20_23_37	1521	1472	1401	1440	1655	1392	1615	1896	1757	1724	1760	1718	1640	1730
20_24_29	1378	1309	1245	1293	1492	1540	1788	1745	1880	1886	1899	1887	1806	1878
20_25_18	1389	1328	1268	1309	1519	1526	1757	1757	1898	1862	1892	1851	1776	1862
20_26_11	1436	1361	1305	1351	1555	1478	1711	1795	1855	1822	1849	1815	1734	1817
20_27_03	1375	1315	1247	1300	1493	1541	1777	1732	1872	1875	1898	1862	1784	1873
20_27_54	1385	1337	1269	1311	1505	1525	1757	1753	1891	1855	1889	1856	1772	1870
20_29_01	1314	1260	1197	1237	1434	1621	1857	1673	1801	1859	1820	1854	1878	1843
20_30_01	1377	1327	1256	1300	1504	1533	1767	1738	1883	1864	1892	1866	1776	1866

The matching snapshots from Table 5-3 are depicted in Table 5-4. According to Table 5-3, first matching pairs of snapshots show about 93% of DNA similarity (1858 out of 2000 pairs of eigenvalues), and the second matching pair show about 95% DNA

similarity (1898 out of 2000 pairs of eigenvalues). 3D visualization of point clouds captured at the specified times, shown in Table 5-4, confirms the similarity of the corresponding point clouds.

Table 5-4 – Matching snapshots

3D Point cloud - First visit	3D Point cloud – Second visit	Place Photo
<p>Place 4 Time: 20:16:22</p>	<p>Place 9 Time: 20:55:37</p>

<p>Place 5 Time: 20:27:03</p>	<p>Place 10 Time: 21:03:21</p>

5.7 Second sequence analysis case study

The travelled path of another combination of indoor/outdoor experiment is shown in Fig. 5-4. Some places, which are referred to in subsequent analysis, are labelled and rest of the places are not labelled to reduce confusion. Please note that the path shown is only for illustration since the actual experiment is conducted one level below the top floor. The length of the path is more than 650 meters. 103 scans were captured and the largest distance between scans is about seven meters. Experiment parameters are summarized in Table 5-5.

Fig. 5-4 – Second sequence analysis case study route

According to section 3.2, anticipated similarity in ideal situation is a value between 76% and 85%. For this experiment, the similarity threshold value of 78% is selected for place analysis. The reason for selection of lower threshold value (compared to previous experiments) is that the distance between the scans in this experiment is larger than previous experiments.

Table 5-6 shows the place matrices that were identified through applying the procedure explained in section 4.2. Place matrices are labelled in accordance with Fig. 5-4.

Cross similarity place matrices are marked in Table 5-7 according to the procedure explained in section 5.1.

Table 5-5 – Parameters of second sequence analysis case study

Parameter	Value
Travel lenght	More than 650m
Number of scans	103
Maximum distance between scans	About 7m
Anticipated similrity	76% to 85%
Similarity threshold	78%

In Table 5-8, only the sequence matrices and cross similarity sequence matrices which qualify the properties explained in section 5.3.1 are shown, and the matrix labels are simplified for easier follow up. By applying 2D cross correlation function, pattern of matrices *A*, *B*, and *C* returns a matching index of more than 90%. However, checking the patterns of *D*, *E*, and *F* returns values less than 56%. The results are summarized in Table 5-9. Therefore, Place matrices included in sequences *A* and *B* highly qualify the criteria to be revisited places, and sequences *A* and *B* are revisited scenarios.

On the other hand, Although *D* and *E* are highly similar (they refer to two long staff offices corridors), but they are unlikely to be revisited places since they do not belong to similar sequences.

Since the values reported in the table of similarities are time-stamped, zooming into matrix *C* (see Table 5-10) helps identifying the best matching frames in two different visits to the same place according to the procedure explained in section 5.6. Two pairs of matching frames are shown in Table 5-11. It can be seen that the shown point clouds are slightly rotated and shifted but they are highly similar.

Photos from labelled places of this experiment are listed in Table 5-12.

35	45	55	65	75	85	95	05	15	25	35	45	55	65	75	85	95
36	46	56	66	76	86	96	06	16	26	36	46	56	66	76	86	96
37	47	57	67	77	87	97	07	17	27	37	47	57	67	77	87	97
38	48	58	68	78	88	98	08	18	28	38	48	58	68	78	88	98
39	49	59	69	79	89	99	09	19	29	39	49	59	69	79	89	99
40	50	60	70	80	90	00	10	20	30	40	50	60	70	80	90	00
41	51	61	71	81	91	01	11	21	31	41	51	61	71	81	91	01
42	52	62	72	82	92	02	12	22	32	42	52	62	72	82	92	02
43	53	63	73	83	93	03	13	23	33	43	53	63	73	83	93	03
44	54	64	74	84	94	04	14	24	34	44	54	64	74	84	94	04
45	55	65	75	85	95	05	15	25	35	45	55	65	75	85	95	05
46	56	66	76	86	96	06	16	26	36	46	56	66	76	86	96	06
47	57	67	77	87	97	07	17	27	37	47	57	67	77	87	97	07
48	58	68	78	88	98	08	18	28	38	48	58	68	78	88	98	08
49	59	69	79	89	99	09	19	29	39	49	59	69	79	89	99	09
50	60	70	80	90	00	10	20	30	40	50	60	70	80	90	00	10
51	61	71	81	91	01	11	21	31	41	51	61	71	81	91	01	11
52	62	72	82	92	02	12	22	32	42	52	62	72	82	92	02	12
53	63	73	83	93	03	13	23	33	43	53	63	73	83	93	03	13
54	64	74	84	94	04	14	24	34	44	54	64	74	84	94	04	14
55	65	75	85	95	05	15	25	35	45	55	65	75	85	95	05	15
56	66	76	86	96	06	16	26	36	46	56	66	76	86	96	06	16
57	67	77	87	97	07	17	27	37	47	57	67	77	87	97	07	17
58	68	78	88	98	08	18	28	38	48	58	68	78	88	98	08	18
59	69	79	89	99	09	19	29	39	49	59	69	79	89	99	09	19
60	70	80	90	00	10	20	30	40	50	60	70	80	90	00	10	20
61	71	81	91	01	11	21	31	41	51	61	71	81	91	01	11	21
62	72	82	92	02	12	22	32	42	52	62	72	82	92	02	12	22
63	73	83	93	03	13	23	33	43	53	63	73	83	93	03	13	23
64	74	84	94	04	14	24	34	44	54	64	74	84	94	04	14	24
65	75	85	95	05	15	25	35	45	55	65	75	85	95	05	15	25
66	76	86	96	06	16	26	36	46	56	66	76	86	96	06	16	26
67	77	87	97	07	17	27	37	47	57	67	77	87	97	07	17	27</

The figure is a detailed floor plan of the 12th floor of the Taipei City Hall. It shows various rooms and their occupancy counts. The plan is color-coded: red for 'Office area in S2a', blue for 'Office area in S2a-b2', green for 'Study area', and yellow for 'After S2 entrance'. The plan includes labels for 'Before S1 entrance', 'Staircase entrance', and 'Elevator S1 entrance'. The occupancy counts are listed in a grid format, with columns representing different areas and rows representing specific rooms or sections. The plan is divided into several sections, including 'Office area in S2a', 'Office area in S2a-b2', 'Study area', and 'After S2 entrance'.

Table 5-7 – Cross similarity matrices marked in second sequence analysis case study

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	
201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	
301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	
401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	
501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	
601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	
701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	
801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	
901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	

Table 5-8 – Sequence matrices and cross sequence matrices marked for second sequence analysis case study

[illegible]

Table 5-9 – Pattern comparison of sequences in second sequence analysis case study

Matrices Compared through Cross Correlation Function	Matching Index
A, C	90.37
B, C	90.48
A, B	92.83
D, F	55.56
E, F	52.30
D, E	37.55

Table 5-10 – Revisited scenario matrix zoomed (second sequence analysis case study)

	2017_Jul_21_19:26:46	2017_Jul_21_19:27:33	2017_Jul_21_19:28:20	2017_Jul_21_19:29:06	2017_Jul_21_19:29:54	2017_Jul_21_19:30:56	2017_Jul_21_19:39:43	2017_Jul_21_19:41:12	2017_Jul_21_19:43:12	2017_Jul_21_19:43:57	2017_Jul_21_19:44:43	2017_Jul_21_19:45:30	2017_Jul_21_19:46:18	2017_Jul_21_19:47:06	2017_Jul_21_19:47:54	2017_Jul_21_19:48:44
2017_Jul_21_17:08:32	79	84	82	79	77	82	83	80	78	66	83	84	66	42	77	73
2017_Jul_21_17:10:08	93	89	67	64	63	67	67	65	63	52	67	68	53	31	62	59
2017_Jul_21_17:10:56	86	91	75	71	70	75	75	73	70	59	75	77	60	36	70	66
2017_Jul_21_17:11:44	82	87	80	76	75	79	80	77	75	63	80	81	63	39	74	71
2017_Jul_21_17:12:53	79	85	83	79	78	82	83	81	78	66	83	84	66	42	77	74
2017_Jul_21_17:13:38	59	64	86	90	92	87	86	89	92	88	86	85	89	60	93	95
2017_Jul_21_17:14:27	58	62	85	89	90	85	85	88	90	90	85	83	91	63	91	94
2017_Jul_21_17:15:15	57	61	84	88	90	84	84	87	89	90	84	82	91	63	90	94
2017_Jul_21_17:16:02	59	63	85	90	92	86	85	89	91	88	86	84	89	61	92	95
2017_Jul_21_17:17:38	64	68	92	95	94	92	92	94	95	82	92	91	83	56	94	91
2017_Jul_21_17:18:59	59	64	87	91	92	87	87	90	92	87	87	85	88	60	92	95
2017_Jul_21_17:20:03	56	60	83	87	89	83	83	86	88	92	83	81	92	64	89	93
2017_Jul_21_17:20:50	55	59	81	85	87	82	81	84	87	93	81	80	94	66	87	91
2017_Jul_21_17:21:38	58	61	85	88	91	85	85	88	90	90	85	83	91	63	90	94
2017_Jul_21_17:22:34	73	78	89	86	84	89	89	87	84	72	89	91	73	47	84	80
2017_Jul_21_17:23:20	56	60	83	87	89	84	83	86	89	92	83	81	92	64	89	93
2017_Jul_21_17:24:08	29	32	48	52	54	49	49	51	53	66	49	47	65	95	54	58
2017_Jul_21_17:24:56	57	61	84	88	89	84	83	86	89	91	84	82	91	63	90	93
2017_Jul_21_17:26:24	60	64	87	91	93	88	87	90	92	87	87	86	88	60	93	95
2017_Jul_21_17:27:36	49	53	74	78	80	74	74	77	79	93	74	73	93	74	80	84

Table 5-11 – Best matching frames in second sequence analysis case study

3D Point cloud First visit	3D Point cloud Second visit
<p>Time: 17:17:38</p>	<p>Time: 19:39:43</p>
<p>Time: 17:20:50</p>	<p>Time: 19:43:57</p>

Table 5-12 – Photos from places studied in second sequence analysis case study

Place Label	A photo from the place
After S2 entrance
Study area
Staircase entrance
Before S1 entrance
Office area in S2.2-b2
Office area in S2a

5.8 Procedure for recognition of revisited places and revisited scenarios

Fig. 5-5 explains the procedure of processing the table of similarities for recognition of revisited places, revisited scenarios, and the most similar observations (snapshot frames) between two different visits of robot to the same place.

5.9 Chapter conclusion

This chapter presented the analysis of patterns in table of similarities. Conducted experiments demonstrated the main idea of the research to see places beyond single snapshots. The results confirm the superior capacity of place DNA to explain environment gradual changes in an outdoor mission. The results also highlight the advantages of matrix comparison and pattern analysis over conventional frame to frame comparison approaches. The proposed methodology concludes to highly confident recognition of revisited places, revisited scenarios, and similar snapshot frames.

One advantage of the presented technique is the low sensitivity to scan rate. For example, in the experiment explained in 5.4, the first visits to place 4 takes four observations made in every 4 meters. The second visit to the same place (place 9) lasts three observations made in every 5 meters. Even though the different visits to the same place are different to this extent, the methodology, confidently detects the revisited place and even the best matching snapshots. Same situation can be seen for place 5 with eight available observation, which is matched to place 10 with seven available observations.

Recognition of revisited scenarios is another outcome of DNA pattern analysis. This functionality enables description of “traveling from place A to place C via place B”. Scenario detection is a high-level description of motion, which is applicable to autonomous decision-making.

Place detection analysis through the presented technique can go beyond what was described in this thesis. Chapter 6 presents some future directions of this research and introduces some application for the presented methodology.

Fig. 5-5 – Procedure for recognition of revisited places and revisited scenarios

Chapter 6

Comparison with Existing Methodologies

The conventional methodology of place recognition has been attended for the past few decades and several research works have been carried out to enhance the results based on the conventional methodology. As studied in Chapter 1 and Chapter 2, most of the existing place recognition methodologies are extensions of feature-matching object recognition techniques. Precious and extensive effort has been made to optimize every single step of these techniques. The focus of place recognition community has been over the bottlenecks. Creative and well-engineered solutions were suggested in the literature to handle these bottlenecks. Therefore, after many years of work, and through the effort made by many researchers, efficient algorithms are available for place recognition following the conventional approach under certain assumptions. The main contribution of this thesis has been to address these bottlenecks, and to propose a different approach, which avoids confrontation with these bottlenecks. In this thesis, the presented case studies are typical implementations of the idea, which help better illustration of the steps, to achieve the place recognition goal, through the proposed methodology. Numeric comparison of the results does not help to illustrate the advantages and disadvantages of conventional methodologies and the proposed methodology for two main reasons. First, the philosophy of place recognition, which is initiated at the definition level, are different in these two approaches. Therefore, the comparison between the outputs is significantly dependent on the application, and so comparing the numbers is not fair. Secondly, the conventional method has been perfectly optimized through decades, and further optimizations or amendments made to any step, directly reflect the enhancements through comparison of the results of amended versions. However, in this thesis, development of the methodology, which involves studying the bottlenecks and providing solutions to prevent facing them, has been the concern. Presented case study implementations require further engineering

work to reach real-time performance. Therefore, the comparison of presented non-optimized implementations of the proposed methodology with results of the mature implementations of conventional methods does not lead to fair conclusions. This section presents a qualitative comparison between the conventional methodology and the proposed methodology to demonstrate a fair picture from the advantages and trade-offs of each methodology. This comparison is summarized in Table 6-1, Table 6-2, and Table 6-3.

6.1 Definition of a place

Existing definitions for place (discussed in section 1.2.1), restrict the studied places to observations that include outstanding landmarks, or contain enough number of certain type of features. Recognition of landmarks, or detection of a group of certain type of features is taken equivalent to recognition of the studied places. A single frame can identify a place.

The proposed definition for the place (in section 1.3) includes all places that include or exclude outstanding landmarks or specific features. Observations, showing consistent descriptors belong to one place. Several observations from each place are necessary to identify a place.

6.2 Sensing

The conventional approach for place recognition is to “look at the place from an external view point” by “pointing the sensor towards the place” through a “certain field of view”. This approach is the root of “sensitivity of the procedures to view-point changes”. The common sensors such as inexpensive monocular cameras and commercially available LiDARs can be used in this approach.

The proposed methodology recommends the view-point to be one internal point of the studied place. Therefore, the proposed sensing technique “looks from a view-point inside of the studied place” to the entire surrounding, through “the entire field of view”. This treatment grounds the basis for “view-point independence” property, which is completed by choosing a proper description technique. Commercially available sensors to produce dense, accurate and long range 3D point cloud to cover the entire field of

view in real-time do not exist up to now. However, the trend of LiDAR products is towards generation of such outputs [158].

6.3 Description

“Invariant feature selection” and “feature association” are two major concerns in most of the conventional methodologies. It is assumed that an image is perfectly represented through its features. Therefore, it is assumed that fair comparison of observations is achieved by comparison of their feature-points only. Because descriptor is formed out of feature vectors, it is also assumed that enough number of certain type of features are available in every observation.

The proposed description methodology is “non-feature-based”. Therefore, feature selection and feature association steps are eliminated. Every bit of the surrounding environment of the view-point is involved in description of the place. It is hereby restated that any non-feature-based descriptor other than the one proposed in this thesis can replace the proposed descriptor, if it has the properties explained in section 2.5 and 3.1 of the thesis. Although fast descriptor generation through this approach can be achieved by parallel computations due to the nature of mathematics involved, no real-time implementation is available yet.

6.4 Comparison, storage, and recall

The conventional method for storage of descriptors is the formation of a Bag-of-words dictionary. Bag-of-words model has proven remarkable performance, and so it has been the most accepted model for storage and recall of feature descriptors. The advantage of the bag-of-words model is elimination of order among the features, which contributes to reduction of sensitivity of feature-based methods to point-of-view changes. However, the size of this database is reported to grow large [91], to the extent that storage and retrieval of entries is only possible through indexing techniques and large-data access models [159]. In addition, comparison of places in conventional methods is the result of comparison between the query frame and previously visited frames. Frame-to-frame comparison is the source of false detections.

Proposed comparison method does not require the formation and maintenance of any database. Comparison between the descriptors is reduced to counting the number of

similar eigenvalues in the place DNA structures through the inflated DNA comparison method explained in section 3.3.2. Definition of place matrices (as explained in section 4.2), enables the methodology to make comparison between “all the observations made from the query place” and all the observations made from each of the previously visited places. This group-to-group comparison ensures that different spots (observations at one view-point) of one place are compared to different spots of another place, and this approach helps to reduce false detections.

6.5 Identification of similar places and identical places through sequence analysis

Study of places in sequences is one way to check whether two similar places are identical places or not.

In conventional methods, sequence comparison can be done over sequences of frames [156]. A long sequence of frames should be compared and correspondences should be checked to match sequences of observations. Matching discrepancies in the long chain make it challenging to identify similar sequences.

In the proposed methodology, introduction of place matrices, cross similarity place matrices, place matrix sequences, and cross similarity place matrix sequences (as explained in section 5.1 of the thesis), enables the methodology to compare the pattern of sequences (as explained in section 5.4.1 of the thesis) even in presence of partial discrepancies. However, generation of helper matrices adds computational overhead.

Table 6-1 – Comparison between the proposed method and conventional methods

Methodology Criteria	Conventional methodologies		Proposed methodology	
	Advantages	Disadvantages	Advantages	Disadvantages
Definition of a place	<ul style="list-style-type: none"> One snapshot from a place describes the place 	<ul style="list-style-type: none"> Not all places contain outstanding landmarks or enough number of specific type of features 	<ul style="list-style-type: none"> Includes all places including or excluding landmarks 	<ul style="list-style-type: none"> Necessitates several observations from a place

through sequence analysis

Table 6-2 – Comparison between the proposed method and conventional methods (Continued I)

Methodology Criteria	Conventional methodologies		Proposed methodology	
	Advantages	Disadvantages	Advantages	Disadvantages
Sensing	<ul style="list-style-type: none"> • Inexpensive cameras and commercially available LiDARs can be used. • Pointing the sensor towards the studied place makes it possible to recognize the place ahead. 	<ul style="list-style-type: none"> • Limited field of view which contributes to view-point dependence of the entire process 	<ul style="list-style-type: none"> • Dense and precise long-range 3D point cloud across the entire field of view helps dealing with view-point invariance concern. • Looking around the viewpoint makes it possible to recognize the place where the sensor is located. 	<ul style="list-style-type: none"> • No commercial sensor can produce such real time output
Description	<ul style="list-style-type: none"> • Several options for different environments • Computationally optimized feature handling algorithms are available 	<ul style="list-style-type: none"> • Assumes existence of enough number of certain type of features in the environment. • Invariant feature selection and feature association are open research issues. 	<ul style="list-style-type: none"> • Non-feature based description can be applied to wider range of places. • Every piece of input data is used for description. 	<ul style="list-style-type: none"> • Procedure for descriptor generation is not real time at the moment. However, since it only involves linear algebraic mathematics, fast parallel implementations can be prepared.
Descriptor comparison, storage, and retrieval	<ul style="list-style-type: none"> • Bag-of-words model eliminates the order of features, which contributes to reduction of sensitivity to view point • One frame of observation is enough to recognize a place 	<ul style="list-style-type: none"> • Bag-of-words model can grow huge for large scale and long-term autonomy. • Frame to frame comparison is prone to false detection. 	<ul style="list-style-type: none"> • Group comparison (through place matrices) produces stronger evidence for similarity between places. • No requirement to store vocabulary/ dictionary. • Fast comparison of descriptors through counting the matching inflated DNA components. 	<ul style="list-style-type: none"> • Several observations are required to form a place matrix

Table 6-3 – Comparison between the proposed method and conventional methods (Continued II)

Methodology Criteria	Conventional methodologies		Proposed methodology	
	Advantages	Disadvantages	Advantages	Disadvantages
Sequence analysis to make distinction between similar and identical places	<ul style="list-style-type: none"> • Can be applied to sequences of frames; Therefore, limited computation is required. 	<ul style="list-style-type: none"> • Long chain of frames should be compared for correspondence checking. In presence of few discrepancies in the studied chain, sequence detection becomes challenging. 	<ul style="list-style-type: none"> • Is applied to groups of observations (place matrices). Introduction of cross similarity place matrices, sequence matrices, and cross similarity place matrices facilitate comparison of the pattern of group sequences which is robust against few discrepancies in comparison procedure. 	<ul style="list-style-type: none"> • Pre-processing computational overhead to identify helping matrices
Others	<ul style="list-style-type: none"> • Many mature implementations exist. • Large community is working to enhance different steps. • Optimum for short-term and mid-term autonomy. 		<ul style="list-style-type: none"> • Only linear algebraic mathematics is involved. Therefore, fast parallel implementations are possible. • Suitable for large scale combined indoor/outdoor applications for long-term autonomy in structured and unstructured places. 	<ul style="list-style-type: none"> • Real-time implementation does not exist yet.

Chapter 7

Conclusion and Future Work

A new methodology for place detection was proposed in this thesis. Research work started by proposing a new definition for a place, which includes all types of places even the ones with no specific landmarks. An affordable sensory system to capture A new methodology for place detection was proposed in this thesis. Research work started by proposing a new definition for a place, which includes all types of places even the ones with no specific landmarks. An affordable sensory system to capture uniformly sense point clouds along the entire usable field of view was suggested. Reasons for preference of Non-feature based and holistic descriptor were discussed. Eigenvalues of Laplace-Beltrami operator was used as the place descriptor, and it was called place DNA. This selection is due to the properties that make place DNA ideal for this application. The most important of which is that the perturbation in studied point cloud results to perturbation in DNA. Table of similarities was arranged to include the similarity between all observations made during the mission. Place matrices were used to represent distinctive places. Since comparison based on place matrices includes the study of many comparisons among different observations from places, results of the analysis is more confident than the result of conventional methods which compare individual frames for detection. Recognition of revisited scenarios was made possible through matching the templates of sequences. Place detection, which involves places in a sequence, is of even higher confidence since the pattern of DNA changes is included in the analysis beside place matrix similarity comparison. Low sensitivity of the technique to scan rate was explained through experiments. This low sensitivity facilitates detection of best matching observation frames in different visits to the same place, even if the scan rate changes in different visits.

Place detection procedure is summarized in the diagram shown in Fig. 7-1 and continued in Fig. 7-2.

Fig. 7-1 - Place detection procedure

Fig. 7-2 – Place detection procedure (continued)

7.1 Conclusion

The proposed place detection methodology addresses a range of shortcomings of conventional place detection techniques from the perquisites such as definition, sensing, and place description, to implementation concerns such as comparison, and recognition of revisited places and revisited scenarios. The advantages, which make the proposed methodology preferred for place detection in large-scale outdoor environment, are explained subsequently and a list of applications are explained followed by that.

7.1.1 Advantages of the presented methodology

- Presented place definition covers all types of places including the large and complex shaped places irrespective to existence of any specific landmark.
- Point cloud capturing system outputs a high density uniform long-range point-cloud across the entire usable field of view.
- No feature selection and data association is required. Therefore, sensitivity to environmental factors are minimized.
- Detection method depends solely on on-board equipment. No external devices are involved and therefore, device coverage issues are out of debate.
- Detection method is independent from metrically accurate mapping module and therefore it is free from accumulated error and drifting debate.
- DNA is a compact descriptor, which is highly consistent against observation-point inaccuracies.
- Table of similarities is a compact summary of similarity among all the observations made throughout the mission. Processing the table of similarities in front end is much faster than comparison of point cloud data. However, the corresponding point cloud data can be retrieved if required.
- Place detection based on place matrix comparison is more confident than conventional frame to frame comparison approach. Because comparison of place matrices involves comparison of every observation made from one place to all the observations made from another place.
- Matching the template of sequences includes comparison of place matrices, similarity values, and differences among all the participating places and

transition frames. Therefore, template matching further increases the detection confidence.

- Sequence matching enables description of scenarios beside the recognition of revisited places.
- Methodology has low sensitivity to scanning rate. Places scanned with different rates are comparable. Methodology has low sensitivity to other parameters such as similarity tolerances and thresholds.
- Other complimentary data such as geo-location and time-stamp can be easily added to the point cloud data to extend the application scope.

7.1.2 Applications

Developed methodology is applicable to many mobile robotics applications. Few cases are explained as follows.

7.1.2.1 Map semantics

Upon successful mapping of outdoor environment, map segmentation to meaningful subsets adds semantic information to the map.

7.1.2.2 Loop closure clue for SLAM

Since the explained methodology is independent from SLAM module, recognition of revisited places and the best matching frames provides loop closure clues for SLAM if the observed point clouds are properly time-stamped.

7.1.2.3 Map merging clues

If multiple robots separately map an area without knowing the initial position and orientation of each other, initial guess for the merging frame of multiple maps can be generated by the presented place detection technique if the table elements are properly time-stamped.

7.1.2.4 Transition detection

Assume an extensive exploration mission in which the robot is supposed to enter and explore several building along the road. Road navigation can be handled through GPS way point navigation. Navigation inside the buildings can be handle through 2D SLAM. The problem is the transition from outdoor to indoor. GPS reading is inaccurate near the buildings, and GPS is not expected to work inside the buildings at all. 2D SLAM is

not reliable for outdoor environment. Therefore, it is required to describe the transition from outdoor to indoor through a technique that is independent from GPS and SLAM. If the DNA table for transitions exists as prior knowledge, once the transition scenario as a template matches the experienced sequence, transitions can be detected and handled.

7.1.2.5 Position initialization in GPS denied environment

If a robot starts traveling along a road from which the DNA table is available as prior knowledge, robot can guess its initial position through the explained place detection technique especially if it passes through few places.

7.1.2.6 Drift detection on localization failure

UAV localization is highly dependent on GPS readings. If GPS is lost, the flying vehicle should go through a safety handling procedure. Unlike the case of ground vehicles, which can stop when motion command is not received, flying vehicle can drift easily in presence of wind. In order to check if the flying vehicle is still in the same place, or it has drifted to another place, the proposed place characterization methodology can be used. Until the DNA structure is unchanged, it is confirmed that the UAV is still in the same place. Once the DNA structure changes, it is identified that UAV has drifted to another place.

7.1.2.7 Geo-location estimation in GPS-denied environment

If DNA table of a scenario is available as prior knowledge, and if the table data is properly geo-tagged, geo-location of robot can be estimated through the presented place detection technique.

7.1.3 Trade-offs to be considered

Like any other detection technique, incorporation of the proposed methodology involves trade-offs.

7.1.3.1 Physical meaning of eigenvalues

DNA descriptor is consisted of eigenvalues of Laplace operator over a mesh manifold. According to [152], eigenvalues of the Laplace operator are the squares of the frequencies of vibration of the studied surface which correspond to one standing wave on the surface domain. The general solution of Laplace eigenvalue problem is the

superposition of all the possible solutions. Therefore, it is not easy to visualize the meaning of each eigenvalue in the DNA structure. This is in spite of the fact that some physical measures, such as area, perimeter, curvature, and connectivity, can be studied through calculations over some of the eigenvalues as discussed in [160]. Anyway, eigenvalues do not represent (at least directly) a visualizable or physically sensible measure. From this perspective, in comparison, feature-based description methods are superior in the sense that they process the physically meaningful and visually presentable measures.

7.1.3.2 Modern processing tools and libraries

Feature-based description has been well-attended by many different computer vision research topics. Therefore, many modern processing libraries are already developed to process the features with multi-CPU and GPU-enabled implementations. Since the proposed methodology is new, parallel processing implementations are yet to be introduced. DNA derivation faces inversion of large matrices, which require long processing time through single thread implementations. Since the objective of this research is limited to development of the methodology, implementation optimization, which also involves multi-core and GPU-enabled implementation, is left as open issue for future extensions. Therefore, implementation of existing feature-based techniques currently result to faster outputs compared to the proposed methodology.

7.1.3.3 Quantitative comparison with other methodologies

Since this technique is a totally new approach to place detection, quantitative comparison of the proposed methodology with existing techniques is not possible.

7.1.3.4 Initial detection

Since the proposed methodology involves processing the place matrices and sequences, at start-up, it is necessary for the methodology to receive number of inputs during the motion, before any detection can be made. In comparison with frame-based detection methodologies, which detect based on one single input frame, the frame-based detection is more efficient during the start-up phase.

7.1.3.5 Sensing

Although the philosophy supporting the proposed definition of place necessitates the type of sensing introduced in section 2.3.2, up to this moment, no commercial sensing

system can produce the required input data in real time. However, LiDAR technology is moving towards production of multi plane LiDAR sensors for recognition purposes. Up to this moment, 128 planes, it the best to be covered in real time [158].

7.2 Future directions

This methodology can be extended and optimized in many of the steps of the procedure shown in Fig. 7-1 and Fig. 7-2. The most important items are listed as follows.

- **Sensory system** – the current sensory system is suggested because the affordable commercial solution does not exist. Current sensory system imposes few seconds of stop to capture each point cloud. Therefore, sensor design is one of the issues to be considered.
- **Point cloud data** – how accurate and dense the point cloud should be, how long the sensor range should be, how frequent the scan rate should be, are the questions that require pushing the methodology to the limits. It is necessary to conduct many tests and analysis to find such tolerances.
- **Input data pre-processing** – content aware down sampling should be studied to find the increased performance in expense of increased computation cost.
- **Eigenvalue problem solution** – Mesh generation is a computationally costly process. Meshless methods should be studied. Lower computation in expense of lower accuracy should be analysed.
- **Sub-part analysis**– Currently, the detection does not study the DNA of each sub-part as an individual descriptor. It is possible to define a descriptor as a vector of DNA arrays. It is expected that matching sub-part DNAs would result to lower computation since smaller matrices are involved for eigenvalue computation. Moreover, it is expected that matching sub-parts could facilitate provision of more semantic information about the studied point cloud.
- **Computation techniques** – Single CPU serial computation should change to multi core GPU computation if online computation is required.
- **Other signals** – other input data, such as colour, can be added to calculations.
- **Complimentary sensors** – other than time-stamp and geo-location tags, other sensory data can be added to the table of similarities entries to extend the application.

Author's Publications

Conference Papers:

1. E. Mihankhah and D. Wang, "Environment characterization using Laplace eigenvalues," in 2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV), 2016, pp. 1-6.
2. E. Mihankhah and D. Wang, "Avoiding to Face The Challenges of Visual Place Recognition," Accepted At Intelligent Systems Conference 2018, London, UK, 2018.

Journal Papers:

3. E. Mihankhah and D. Wang, "Non-Feature-Based Place Detection through Segmentation of The Travelled Route to Subsets ," 2018, In review process, *The Journal of Field Robotics*.

Patents:

1. E. Mihankhah and D. Wang, "Identification of Zones through Sequence Analysis," NTU Ref: PAT/395/16/17/SG PRV, Singapore provisional patent application number 10201707160W, Dec 2016, Singapore

Author's Previous Publications and Contributed Papers

1. E. Mihankhah, A. Kalantari, E. Aboosaeedan, H. D. Taghirad, S. Ali, and A. Moosavian, "Autonomous staircase detection and stair climbing for a tracked mobile robot using fuzzy controller," in 2008 IEEE International Conference on Robotics and Biomimetics, 2009, pp. 1980-1985.
2. E. Mihankhah, H. D. Taghirad, A. Kalantari, E. Aboosaeedan, and H. Semsarilar, "Line matching localization and map building with least square," in 2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2009, pp. 1734-1739.
3. Y. Wang, D. Wang, and E. Mihankhah, "Navigation of multiple mobile robots in unknown environments using a new decentralized navigation function," in 2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV), 2016, pp. 1-6.
4. J. Pellenz, A. Jacoff, T. Kimura, E. Mihankhah, R. Sheh, and J. Suthakorn, "RoboCup rescue robot league," in Robot Soccer World Cup, 2014, pp. 673-685.
5. R. Sheh, T. Kimura, E. Mihankhah, J. Pellenz, S. Schwertfeger, and J. Suthakorn, "The robocuprescue robot league: Guiding robots towards fieldable capabilities," in Advanced Robotics and its Social Impacts (ARSO), 2011 IEEE Workshop on, 2011, pp. 31-34.
6. S. A. A. Moosavian, A. Kalantari, H. Semsarilar, E. Aboosaeedan, and E. Mihankhah, "ResQuake: a tele-operative rescue robot," Journal of mechanical design, vol. 131, p. 081005, 2009.
7. A. Kalantari, E. Mihankhah, and S. A. A. Moosavian, "Safe autonomous stair climbing for a tracked mobile robot using a kinematics based controller," in Advanced Intelligent Mechatronics, 2009. AIM 2009. IEEE/ASME International Conference on, 2009, pp. 1891-1896.
8. A. A. Aghamohammadi, H. D. Taghirad, A. H. Tamjidi, and E. Mihankhah, "Feature-Based Laser Scan Matching For Accurate and High Speed Mobile Robot Localization," in EMCR, 2007.

References

1. Thrun, S., W. Burgard, and D. Fox, *Probabilistic robotics*. 2005: MIT Press.
2. Aulinas, J., et al., *The SLAM problem: a survey*, in *Proceedings of the 2008 conference on Artificial Intelligence Research and Development: Proceedings of the 11th International Conference of the Catalan Association for Artificial Intelligence*. 2008, IOS Press. p. 363-371.
3. Durrant-Whyte, H. and T. Bailey, *Simultaneous localization and mapping: part I*. IEEE Robotics & Automation Magazine, 2006. **13**(2): p. 99-110.
4. Bailey, T. and H. Durrant-Whyte, *Simultaneous localization and mapping (SLAM): part II*. IEEE Robotics & Automation Magazine, 2006. **13**(3): p. 108-117.
5. Cadena, C., et al., *Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age*. IEEE Transactions on Robotics, 2016. **32**(6): p. 1309-1332.
6. Hamid, M.H.A., et al. *Navigation of mobile robot using Global Positioning System (GPS) and obstacle avoidance system with commanded loop daisy chaining application method*. in *2009 5th International Colloquium on Signal Processing & Its Applications*. 2009.
7. Dudek, G., et al., *Robotic exploration as graph construction*. IEEE Transactions on Robotics and Automation, 1991. **7**(6): p. 859-865.
8. Kuipers, B., *The Spatial Semantic Hierarchy*. Artificial Intelligence, 2000. **119**(1): p. 191-233.
9. Salas-Moreno, R.F., et al. *SLAM++: Simultaneous Localisation and Mapping at the Level of Objects*. in *2013 IEEE Conference on Computer Vision and Pattern Recognition*. 2013.
10. Kortenkamp, D., L.D. Baker, and T. Weymouth. *Using Gateways To Build A Route Map*. in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. 1992.
11. Korrapati, H., et al. *Image Sequence Partitioning for outdoor mapping*. in *2012 IEEE International Conference on Robotics and Automation*. 2012.
12. Lee, J.H., et al., *Place recognition using straight lines for vision-based SLAM*. 2013. 3799-3806.
13. Volkhardt, M., et al., *Detection of Lounging People with a Mobile Robot Companion*, in *Intelligent Robotics and Applications: 4th International Conference, ICIRA 2011, Aachen, Germany, December 6-8, 2011, Proceedings*,

- Part II*, S. Jeschke, H. Liu, and D. Schilberg, Editors. 2011, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 328-337.
14. Janet, J.A., et al. *Using a hyper-ellipsoid clustering Kohonen for autonomous mobile robot map building, place recognition and motion planning*. in *Neural Networks, 1997., International Conference on*. 1997.
 15. Bloomenthal, J., *Skeletal Design of Natural Forms*, in *Department of Computer Science*. 1995, The University of Calgary: Calgary, Alberta, Canada.
 16. Cadena, C., et al., *Robust Place Recognition With Stereo Sequences*. IEEE Transactions on Robotics, 2012. **28**(4): p. 871-885.
 17. Murillo, A.C., J.J. Guerrero, and C. Sagues. *SURF features for efficient robot localization with omnidirectional images*. in *Proceedings 2007 IEEE International Conference on Robotics and Automation*. 2007.
 18. Lowry, S., et al., *Visual Place Recognition: A Survey*. IEEE Transactions on Robotics, 2016. **32**(1): p. 1-19.
 19. Ranganathan, A., *Detecting And Labeling Places Using Runtime Change-Point Detection*. 2011, Google Patents.
 20. Ranganathan, A., *PLISS: labeling places using online changepoint detection*. Auton. Robots, 2012. **32**(4): p. 351-368.
 21. Song, T., et al., *Enhancing GPS With Lane-Level Navigation to Facilitate Highway Driving*. IEEE Transactions on Vehicular Technology, 2017. **66**(6): p. 4579-4591.
 22. Lowe, D.G. *Object recognition from local scale-invariant features*. in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*. 1999.
 23. Ulrich, I. and I. Nourbakhsh. *Appearance-based place recognition for topological localization*. in *Robotics and Automation, 2000. Proceedings. ICRA '00. IEEE International Conference on*. 2000.
 24. Cummins, M. and P. Newman. *Accelerated appearance-only SLAM*. in *2008 IEEE International Conference on Robotics and Automation*. 2008.
 25. Cummins, M. and P. Newman, *FAB-MAP: Probabilistic localization and mapping in the space of appearance*. The International Journal of Robotics Research, 2008. **27**(6): p. 647-665.
 26. Cummins, M. and P. Newman, *Appearance-only SLAM at large scale with FAB-MAP 2.0*. The International Journal of Robotics Research, 2011. **30**(9): p. 1100-1123.
 27. Radoi, I., et al. *Indoor Positioning inside an Office Building Using BLE*. in *2017 21st International Conference on Control Systems and Computer Science (CSCS)*. 2017.

28. Barnard, M.E., *The Global Positioning System*. IEE Review, 1992. **38**(3): p. 99-102.
29. Gu, Y. and S. Kamijo. *GNSS positioning in deep urban city with 3D map and double reflection*. in *2017 European Navigation Conference (ENC)*. 2017.
30. Lutsenko, V.I., et al. *Methods and algorithms of correction of propagation factor influence on errors of measurement coordinates receivers GNSS*. in *2012 22nd International Crimean Conference "Microwave & Telecommunication Technology"*. 2012.
31. Axell, E., E.G. Larsson, and D. Persson. *GNSS spoofing detection using multiple mobile COTS receivers*. in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2015.
32. Fan, Y., et al., *A Cross-Layer Defense Mechanism Against GPS Spoofing Attacks on PMUs in Smart Grids*. IEEE Transactions on Smart Grid, 2015. **6**(6): p. 2659-2668.
33. Vo, Q.D. and P. De, *A Survey of Fingerprint-Based Outdoor Localization*. IEEE Communications Surveys & Tutorials, 2016. **18**(1): p. 491-506.
34. Tron, R. and R. Vidal, *Distributed 3-D Localization of Camera Sensor Networks From 2-D Image Measurements*. IEEE Transactions on Automatic Control, 2014. **59**(12): p. 3325-3340.
35. Zhongyang, Z., L. Yan, and J. Wang. *LiDAR point cloud registration based on improved ICP method and SIFT feature*. in *2015 IEEE International Conference on Progress in Informatics and Computing (PIC)*. 2015.
36. Paul, R. and P. Newman. *FAB-MAP 3D: Topological mapping with spatial and visual appearance*. in *2010 IEEE International Conference on Robotics and Automation*. 2010.
37. Cole, D.M. and P.M. Newman. *Using laser range data for 3D SLAM in outdoor environments*. in *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006*. 2006.
38. Zhu, H., J.B. Weibel, and S. Lu. *Discriminative Multi-modal Feature Fusion for RGBD Indoor Scene Recognition*. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.
39. Labbé, M. and F. Michaud. *Online global loop closure detection for large-scale multi-session graph-based SLAM*. in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2014.
40. Civera, J., et al. *Towards semantic SLAM using a monocular camera*. in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2011.
41. Konolige, K., et al., *View-based Maps*. Int. J. Rob. Res., 2010. **29**(8): p. 941-957.

42. Newcombe, R.A., et al. *KinectFusion: Real-time dense surface mapping and tracking*. in *2011 10th IEEE International Symposium on Mixed and Augmented Reality*. 2011.
43. Sattler, T., B. Leibe, and L. Kobbelt, *Fast image-based localization using direct 2D-to-3D matching*, in *Proceedings of the 2011 International Conference on Computer Vision*. 2011, IEEE Computer Society. p. 667-674.
44. Sattler, T., B. Leibe, and L. Kobbelt, *Improving image-based localization by active correspondence search*, in *Proceedings of the 12th European conference on Computer Vision - Volume Part I*. 2012, Springer-Verlag: Florence, Italy. p. 752-765.
45. Hassaballah, M. and A.I. Awad, *Detection and Description of Image Features: An Introduction*, in *Image Feature Detectors and Descriptors : Foundations and Applications*, A.I. Awad and M. Hassaballah, Editors. 2016, Springer International Publishing: Cham. p. 1-8.
46. Canny, J., *A Computational Approach to Edge Detection*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986. **PAMI-8**(6): p. 679-698.
47. Harris, C. and M. Stephens. *A combined corner and edge detector*. in *Alvey vision conference*. 1988. Manchester, UK.
48. Beaudet, P.R. *Rotationally invariant image operators*. in *Proceedings of the 4th International Joint Conference on Pattern Recognition*. 1978.
49. Lindeberg, T., *Feature Detection with Automatic Scale Selection*. International Journal of Computer Vision, 1998. **30**(2): p. 79-116.
50. Lowe, D.G., *Distinctive Image Features from Scale-Invariant Keypoints*. International Journal of Computer Vision, 2004. **60**(2): p. 91-110.
51. Mikolajczyk, K. and C. Schmid, *Scale & Affine Invariant Interest Point Detectors*. International Journal of Computer Vision, 2004. **60**(1): p. 63-86.
52. Grauman, K. and B. Leibe, *Visual Object Recognition*. Synthesis Lectures on Artificial Intelligence and Machine Learning, 2011. **5**(2): p. 1-181.
53. Mikolajczyk, K., et al., *A Comparison of Affine Region Detectors*. International Journal of Computer Vision, 2005. **65**(1): p. 43-72.
54. Hubeli, A. and M. Gross. *Multiresolution feature extraction for unstructured meshes*. in *Visualization, 2001. VIS '01. Proceedings*. 2001.
55. Hildebrandt, K., K. Polthier, and M. Wardetzky, *Smooth feature lines on surface meshes*, in *Proceedings of the third Eurographics symposium on Geometry processing*. 2005, Eurographics Association: Vienna, Austria. p. 85.
56. Watanabe, K. and A.G. Belyaev, *Detection of Salient Curvature Features on Polygonal Surfaces*. Computer Graphics Forum, 2001. **20**(3): p. 385-392.

57. Demarsin, K., et al., *Detection of closed sharp edges in point clouds using normal estimation and graph theory*. Comput. Aided Des., 2007. **39**(4): p. 276-283.
58. Weber, C., S. Hahmann, and H. Hagen, *Sharp Feature Detection in Point Clouds*, in *Proceedings of the 2010 Shape Modeling International Conference*. 2010, IEEE Computer Society. p. 175-186.
59. Fleishman, S., D. Cohen-Or, and C.T. Silva, *Robust moving least-squares fitting with sharp features*. ACM Trans. Graph., 2005. **24**(3): p. 544-552.
60. Abdi, H. and L.J. Williams, *Principal component analysis*. Wiley Interdisciplinary Reviews: Computational Statistics, 2010. **2**(4): p. 433-459.
61. Han, Y., X.-C. Feng, and G. Baciú, *Variational and PCA based natural image segmentation*. Pattern Recognition, 2013. **46**(7): p. 1971-1984.
62. Malagón-Borja, L. and O. Fuentes, *Object detection using image reconstruction with PCA*. Image and Vision Computing, 2009. **27**(1): p. 2-9.
63. Wang, J., Z. Dodds, and W.L. Miranker, *Principal Component Analysis for Place Recognition*. Journal of Neural, Parallel, and Scientific Computations, 1996.
64. Weber, C., S. Hahmann, and H. Hagen. *Methods for Feature Detection in Point Clouds*. in *Visualization of Large and Unstructured Data Sets - Applications in Geospatial Planning, Modeling and Engineering (IRTG 1131 Workshop)*. 2010. Bodega Bay, CA, United States: Schloss Dagstuhl-Leiniz-Zentrum fuer Informatik.
65. Murillo, A.C. and J. Kosecka. *Experiments in place recognition using gist panoramas*. in *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*. 2009.
66. Singhal, A., N. Srivastava, and R. Mishra. *Hiding Signature In Colored Image*. in *2006 International Symposium on Communications and Information Technologies*. 2006.
67. Stumm, E., C. Mei, and S. Lacroix. *Probabilistic place recognition with covisibility maps*. in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2013.
68. Mikolajczyk, K. and C. Schmid. *A performance evaluation of local descriptors*. in *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings*. 2003.
69. Bay, H., T. Tuytelaars, and L. Van Gool, *SURF: Speeded Up Robust Features*, in *Computer Vision – ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I*, A. Leonardis, H. Bischof, and A. Pinz, Editors. 2006, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 404-417.
70. Goswami, J.C. and A.K. Chan, *Fundamentals of Wavelets: Theory, Algorithms, and Applications*. 2011: Wiley Publishing. 359.

71. Ashbrook, A.P., et al., *Robust recognition of scaled shapes using pairwise geometric histograms*, in *Proceedings of the 6th British conference on Machine vision (Vol. 2)*. 1995, BMVA Press: Birmingham, United Kingdom. p. 503-512.
72. Belongie, S., J. Malik, and J. Puzicha, *Shape matching and object recognition using shape contexts*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002. **24**(4): p. 509-522.
73. Calonder, M., et al., *BRIEF: Computing a Local Binary Descriptor Very Fast*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012. **34**(7): p. 1281-1298.
74. Pappalardo, F., et al. *HAMFAST: Fast Hamming Distance Computation*. in *2009 WRI World Congress on Computer Science and Information Engineering*. 2009.
75. Gabor, D., *Theory of communication*. Electrical Engineers - Part I: General, Journal of the Institution of, 1947. **94**(73): p. 58.
76. Vetterli, M. and J. Kovačević, *Wavelets and subband coding*. 1995: Prentice-Hall, Inc. 488.
77. Oliva, A. and A. Torralba, *Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope*. International Journal of Computer Vision, 2001. **42**(3): p. 145-175.
78. Yang, L. and Z. Hong. *Visual loop closure detection with a compact image descriptor*. in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2012.
79. Van Gool, L., T. Moons, and D. Ungureanu, *Affine / photometric invariants for planar intensity patterns*, in *Computer Vision — ECCV '96: 4th European Conference on Computer Vision Cambridge, UK, April 15–18, 1996 Proceedings, Volume I*, B. Buxton and R. Cipolla, Editors. 1996, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 642-651.
80. Baeza-Yates, R.A. and B. Ribeiro-Neto, *Modern Information Retrieval*. 1999: Addison-Wesley Longman Publishing Co., Inc. 513.
81. Ho, K.L. and P. Newman, *Detecting Loop Closure with Scene Sequences*. International Journal of Computer Vision, 2007. **74**(3): p. 261-286.
82. Vlassis, N. and B. Krose. *Mixture conditional density estimation with the EM algorithm*. in *1999 Ninth International Conference on Artificial Neural Networks ICANN 99. (Conf. Publ. No. 470)*. 1999.
83. Angeli, A., et al. *Real-time visual loop-closure detection*. in *2008 IEEE International Conference on Robotics and Automation*. 2008.
84. Ho, K. and P. Newman. *Multiple map intersection detection using visual appearance*. in *International conference on computational intelligence, robotics and autonomous systems*. 2005.

85. Smith, T.F. and M.S. Waterman, *Identification of common molecular subsequences*. Journal of Molecular Biology, 1981. **147**(1): p. 195-197.
86. Newman, P. and H. Kin. *SLAM-Loop Closing with Visually Salient Features*. in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. 2005.
87. Gutmann, J.S. and K. Konolige. *Incremental mapping of large cyclic environments*. in *Computational Intelligence in Robotics and Automation, 1999. CIRA '99. Proceedings. 1999 IEEE International Symposium on*. 1999.
88. Ho, K.L. and P. Newman, *Loop closure detection in SLAM by combining visual and spatial appearance*. Robotics and Autonomous Systems, 2006. **54**(9): p. 740-749.
89. Giovannangeli, C., P. Gaussier, and G. Desilles. *Robust Mapless Outdoor Vision-Based Navigation*. in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2006.
90. Li, Z., S.R.u.N. Jafri, and R. Chellali. *Visual place recognition for multi-robots maps merging*. in *2012 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. 2012.
91. Schindler, G., M. Brown, and R. Szeliski. *City-Scale Location Recognition*. in *2007 IEEE Conference on Computer Vision and Pattern Recognition*. 2007.
92. Mohan, M., et al. *Environment selection and hierarchical place recognition*. in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. 2015. IEEE.
93. Blanco, J.-L., F.-A. Moreno, and J. Gonzalez, *A collection of outdoor robotic datasets with centimeter-accuracy ground truth*. Autonomous Robots, 2009. **27**(4): p. 327.
94. Golovinskiy, A., V.G. Kim, and T. Funkhouser. *Shape-based recognition of 3D point clouds in urban environments*. in *2009 IEEE 12th International Conference on Computer Vision*. 2009.
95. Satkin, S. and M. Hebert. *3DNN: Viewpoint Invariant 3D Geometry Matching for Scene Understanding*. in *2013 IEEE International Conference on Computer Vision*. 2013.
96. Kuipers, B. and Y.-T. Byun, *A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations*. Robotics and Autonomous Systems, 1991. **8**(1): p. 47-63.
97. Reuter, M., F.-E. Wolter, and N. Peinecke, *Laplace–Beltrami spectra as ‘Shape-DNA’ of surfaces and solids*. Computer-Aided Design, 2006. **38**(4): p. 342-366.
98. Cieslewski, T., et al. *Point cloud descriptors for place recognition using sparse visual information*. in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. 2016.

99. Leutenegger, S., et al., *Keyframe-based visual-inertial odometry using nonlinear optimization*. The International Journal of Robotics Research, 2015. **34**(3): p. 314-334.
100. Bosse, M. and R. Zlot. *Place recognition using keypoint voting in large 3D lidar datasets*. in *2013 IEEE International Conference on Robotics and Automation*. 2013.
101. Bosse, M. and R. Zlot, *Keypoint design and evaluation for place recognition in 2D lidar maps*. Robotics and Autonomous Systems, 2009. **57**(12): p. 1211-1224.
102. Bosse, M. and R. Zlot, *Place Recognition Using Regional Point Descriptors for 3D Mapping*, in *Field and Service Robotics: Results of the 7th International Conference*, A. Howard, K. Iagnemma, and A. Kelly, Editors. 2010, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 195-204.
103. Fiolka, T., et al. *Distinctive 3D surface entropy features for place recognition*. in *2013 European Conference on Mobile Robots*. 2013.
104. Fiolka, T., et al., *SURE: Surface Entropy for Distinctive 3D Features*, in *Spatial Cognition VIII: International Conference, Spatial Cognition 2012, Kloster Seeon, Germany, August 31 – September 3, 2012. Proceedings*, C. Stachniss, K. Schill, and D. Uttal, Editors. 2012, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 74-93.
105. Litman, R. and A.M. Bronstein, *Learning Spectral Descriptors for Deformable Shape Correspondence*. IEEE Trans. Pattern Anal. Mach. Intell., 2014. **36**(1): p. 171-180.
106. Sun, J., M. Ovsjanikov, and L. Guibas, *A concise and provably informative multi-scale signature based on heat diffusion*, in *Proceedings of the Symposium on Geometry Processing*. 2009, Eurographics Association: Berlin, Germany. p. 1383-1392.
107. Se, S., D. Lowe, and J. Little, *Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks*. The international Journal of robotics Research, 2002. **21**(8): p. 735-758.
108. Zhang, Z. and O.D. Faugeras, *Determining motion from 3D line segment matches: a comparative study*. Image and Vision Computing, 1991. **9**(1): p. 10-19.
109. Grimson, W. and T. Lozano-Perez. *Model-based recognition and localization from tactile data*. in *Robotics and Automation. Proceedings. 1984 IEEE International Conference on*. 1984.
110. Choudhary, S., et al. *SLAM with object discovery, modeling and mapping*. in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. 2014. IEEE.
111. Ok, K., D.-N. Ta, and F. Dellaert, *Vistas and Wall-Floor Intersection Features: Enabling Autonomous Flight in Man-made Environments*. 2012.

112. Oliva, A. and A. Torralba, *Building the gist of a scene: The role of global image features in recognition*. Progress in brain research, 2006. **155**: p. 23-36.
113. Maximo, A., et al., *A robust and rotationally invariant local surface descriptor with applications to non-local mesh processing*. Graphical Models, 2011. **73**(5): p. 231-242.
114. Qiu, A., L. Younes, and M.I. Miller, *Intrinsic and extrinsic analysis in computational anatomy*. NeuroImage, 2008. **39**(4): p. 1803-1814.
115. Peinecke, N., F.-E. Wolter, and M. Reuter, *Laplace spectra as fingerprints for image recognition*. Computer-Aided Design, 2007. **39**(6): p. 460-476.
116. Belkin, M. and P. Niyogi, *Laplacian eigenmaps for dimensionality reduction and data representation*. Neural computation, 2003. **15**(6): p. 1373-1396.
117. Reuter, M., et al., *Laplace-Beltrami Eigenvalues and Topological Features of Eigenfunctions for Statistical Shape Analysis*. Comput Aided Des, 2009. **41**(10): p. 739-755.
118. Niethammer, M., et al., *Global Medical Shape Analysis Using the Laplace-Beltrami Spectrum*. Med Image Comput Comput Assist Interv, 2007. **10**(Pt 1): p. 850-7.
119. Reuter, M., et al. *Global Medical Shape Analysis Using the Volumetric Laplace Spectrum*. in *Cyberworlds, 2007. CW '07. International Conference on*. 2007.
120. Wachinger, C., et al., *BrainPrint: a discriminative characterization of brain morphology*. Neuroimage, 2015. **109**: p. 232-48.
121. Mihankhah, E. and D. Wang. *Environment characterization using Laplace eigenvalues*. in *2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. 2016.
122. Szeliski, R., *Video mosaics for virtual environments*. IEEE Computer Graphics and Applications, 1996. **16**(2): p. 22-30.
123. Nister, D. *A minimal solution to the generalised 3-point pose problem*. in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*. 2004.
124. Nayar, S.K. *Catadioptric omnidirectional camera*. in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 1997.
125. *Point Cloud Library (PCL)*. 2016; Available from: <http://pointclouds.org/>.
126. Weingarten, J. and R. Siegwart. *3D SLAM using planar segments*. in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*. 2006.
127. Pathak, K., et al., *Fast Registration Based on Noisy Planes With Unknown Correspondences for 3-D Mapping*. IEEE Transactions on Robotics, 2010. **26**(3): p. 424-441.

128. Schadler, M., J. Stückler, and S. Behnke. *Multi-resolution surfel mapping and real-time pose tracking using a continuously rotating 2D laser scanner*. in *2013 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. 2013.
129. Xiao, J., et al., *Three-dimensional point cloud plane segmentation in both structured and unstructured environments*. *Robotics and Autonomous Systems*, 2013. **61**(12): p. 1641-1652.
130. Cipriano, G., J. G. N. Phillips, and M. Gleicher, *Multi-Scale Surface Descriptors*. *IEEE Transactions on Visualization and Computer Graphics*, 2009. **15**(6): p. 1201-1208.
131. Staten, M.L. and N.L. Jones, *Local refinement of three-dimensional finite element meshes*. *Engineering with Computers*, 1997. **13**(3): p. 165-174.
132. Rosenberg, S. and S. Rosenberg, *The Laplacian on a Riemannian Manifold*. 1997: Cambridge University Press.
133. Zelditch, S., *Local and global analysis of eigenfunctions on Riemannian manifolds*, in *Handbook of geometric analysis*. 2008, International Press: Somerville, MA.
134. Babuška, I. and J. Osborn, *Eigenvalue problems*, in *Handbook of Numerical Analysis*. 1991, Elsevier. p. 641-787.
135. Belkin, M., J. Sun, and Y. Wang, *Constructing Laplace Operator from Point Clouds in $\mathbb{R}^d$* , in *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*. 2009. p. 1031-1040.
136. Rosenberg, S., *The Laplacian on a Riemannian Manifold: An Introduction to Analysis on Manifolds*. London Mathematical Society Student Texts. 1997, Cambridge: Cambridge University Press.
137. Maintrot, M., *Finite Element Method on Riemann Surfaces and Applications to the Laplacian Spectrum*. 2012, EPFL.
138. Marton, Z.C., R.B. Rusu, and M. Beetz. *On fast surface reconstruction methods for large and noisy point clouds*. in *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*. 2009.
139. Mohar, B., *Laplace eigenvalues of graphs—a survey*. *Discrete Mathematics*, 1992. **109**(1): p. 171-183.
140. Meyer, M., et al., *Discrete Differential-Geometry Operators for Triangulated 2-Manifolds*, in *Visualization and Mathematics III*, H.-C. Hege and K. Polthier, Editors. 2003, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 35-57.
141. Boffi, D., *Finite element approximation of eigenvalue problems*. *Acta Numerica*, 2010. **19**: p. 1-120.
142. Guoliang, X. *Convergent discrete Laplace-Beltrami operators over triangular surfaces*. in *Geometric Modeling and Processing, 2004. Proceedings*. 2004.

143. Reuter, M., *Laplace Spectra for Shape Recognition*. 2006: Books on Demand GmbH.
144. Bangerth, W., et al., *The deal.II library, Version 8.3*. 2016, 2016. **4**(100): p. 11.
145. Arnoldi, W.E., *The Principle of Minimized Iterations in The Solution of The Matrix Eigenvalue Problem*. Quarterly of Applied Mathematics, 1951. **9**(1): p. 17-29.
146. Demmel, J., et al., *Templates for the solution of algebraic eigenvalue problems: a practical guide*, ed. B. Zhaojun. 2000: Society for Industrial and Applied Mathematics. 410.
147. Lehoucq, R.B., D.C. Sorensen, and C. Yang, *ARPACK Users' Guide*. ARPACK Users' Guide. 1998.
148. Kuttler, J.R. and V.G. Sigillito, *Eigenvalues of the Laplacian in Two Dimensions*. SIAM Review, 1984. **26**(2): p. 163-193.
149. Gao, Z., Z. Yu, and X. Pang, *A compact shape descriptor for triangular surface meshes*. Computer-Aided Design, 2014. **53**: p. 62-69.
150. Gordon, C., D. Webb, and S. Wolpert, *Isospectral plane domains and surfaces via Riemannian orbifolds*. Inventiones mathematicae, 1992. **110**(1): p. 1-22.
151. Zelditch, S., *Spectral determination of analytic bi-axisymmetric plane domains*. Geometric & Functional Analysis GAFA, 2000. **10**(3): p. 628-677.
152. Chapman, S.J., *Drums That Sound the Same*. The American Mathematical Monthly, 1995. **102**(2): p. 124-138.
153. Bérard, P., *Transplantation Et Isospectralité II*. Journal of the London Mathematical Society, 1993. **s2-48**(3): p. 565-576.
154. Sivic, J. and A. Zisserman, *Efficient Visual Search of Videos Cast as Text Retrieval*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009. **31**(4): p. 591-606.
155. Dunning, T., *Statistical identification of language*. 1994, Computing Research Laboratory, New Mexico State University.
156. Milford, M.J. and G.F. Wyeth. *SeqSLAM: Visual route-based navigation for sunny summer days and stormy winter nights*. in *2012 IEEE International Conference on Robotics and Automation*. 2012.
157. Bradski, G. *The OpenCV Library*. 2000; Available from: <http://opencv.org/>.
158. Velodyne VLS 128. 2017; Available from: <http://velodynelidar.com/blog/128-lasers-car-go-round-round-david-hall-velodynes-new-sensor/>.
159. Galvez-López, D. and J.D. Tardos, *Bags of Binary Words for Fast Place Recognition in Image Sequences*. IEEE Transactions on Robotics, 2012. **28**(5): p. 1188-1197.

160. Kac, M., *Can One Hear the Shape of a Drum?* The American Mathematical Monthly, 1966. **73**(4): p. 1-23.