

RaSpectLoc: Raman SPECTroscopy-dependent robot LOCALisation

Christopher Thirgood^{1a}, Oscar Mendez^{1a}, Erin Chao Ling^{1b}, Jon Storey², Simon Hadfield^{1a}

Abstract—This paper presents a new information source for supporting robot localisation: material composition. The proposed method complements the existing visual, structural, and semantic cues utilized in the literature. However, it has a distinct advantage in its ability to differentiate structurally [23], visually [25] or categorically [1] similar objects such as different doors, by using Raman spectrometers. Such devices can identify the material of objects it probes through the bonds between the material’s molecules. Unlike similar sensors, such as mass spectroscopy, it does so without damaging the material or environment. In addition to introducing the first material-based localisation algorithm, this paper supports the future growth of the field by presenting a gazebo plugin for Raman spectrometers, material sensing demonstrations, as well as the first-ever localisation data-set with benchmarks for material-based localisation. This benchmarking shows that the proposed technique results in a significant improvement over current state-of-the-art localisation techniques, achieving 16% more accurate localisation than the leading baseline.

The code and dataset will be released at: <https://github.com/ThirgoodC/RaSpectLoc>

I. INTRODUCTION

Mobile robots have historically relied on depth sensors for localisation tasks, more recently, visual (RGB) sensors have become ubiquitous. However, these traditional RGB approaches face challenges in urban environments where large planar regions of uniform colour often dominate and distinct visual landmarks are lacking. This is especially problematic in self-similar spaces such as hotels with identical rooms and rotational symmetric floorplans.

A novel solution is to sense the material composition of the environment, which can help distinguish visually similar objects, such as different doors. Additionally, small deviations in material composition, such as impurities in concrete or uneven levels of corrosion, can serve as local landmarks. These observations can be made and the material composition identified using sensors such as mass spectrometers and Raman spectrometers.

In this paper, we propose a new localisation approach that leverages the capabilities of Raman spectroscopy. Raman probes are commonly used in physics as active sensors to analyze at a range, without damaging the subject matter. They work by analyzing molecular interactions and bonds through light scattering. This results in a Raman spectrum

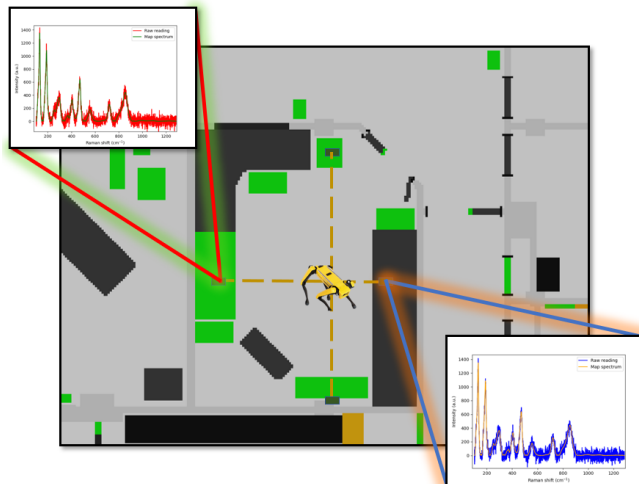


Fig. 1: Visualisation of the RaSpectLoc system in action. Spectral readings are compared between the environment and Map along a number of bearing vectors.

featuring peaks corresponding to specific molecular bond vibrations.

We present an approach that uses spectral responses produced by a Raman spectrometer as its central sensing unit. It can work with or without range information and demonstrates superior performance to vision and depth-based methods. Furthermore, it effectively handles inaccuracies introduced in raw responses from the sensor.

Fig. 1 demonstrates a visualisation of RaSpectLoc, where spectra are compared against each other within a material map.

In summary, the contributions of this paper are:

- 1) A novel approach in mobile robotics that utilizes material-based spectroscopic data for localisation.
- 2) A plug-in for the popular simulator, Gazebo, to simulate Raman spectrometers.
- 3) A new dataset containing maps and recordings of material composition from a test environment.
- 4) A benchmarked performance evaluation of our approach which provides a far greater average trajectory error compared to current state-of-the-art localisation approaches that utilize RGBD sensors.

II. LITERATURE REVIEW

Mobile robot applications have traditionally used scan-matching approaches to localise within a known map. These approaches have historically been dominated by particle filter-based methods, such as AMCL [23] and GMapping [26] which have been popular solutions due to the integration with the ROS navigation stack [4]. The technique

^{1a}CVSSP, Computer Science and Electronic Engineering, University of Surrey, Guildford, Surrey, United Kingdom {c.thirgood, s.hadfield, o.mendez}@surrey.ac.uk

^{1b}Surrey Institute for People-Centered Artificial Intelligence, University of Surrey, Guildford Surrey, United Kingdom chao.ling@surrey.ac.uk

²Industrial 3D Robotics, Tonbridge, Kent, UK jstorey@i3drobotics.com

is range-based (RMCL) and is highly adaptable, adjusting the number of hypotheses generated by the particle filter as confidence in the result improves. However, AMCL may fail if the map and the sourced environment differ. Many ‘Simultaneous Localisation and Mapping’ (SLAM) systems intend to solve this by generating a map at run-time and identifying ‘loop closures’ to solve for drift [2], [3] for static environments. For typical dynamic scenes such as industrial work-spaces [5], [19] these approaches have attempted to track the pose using an apriori map and secondary short-term map at two different timescales. The problem with such a method is that it is too computationally complex for mobile robotics to reliably localize for long periods with accumulated errors. Our approach aims to resolve this problem via the use of efficient spectral comparison functions which can provide continuous similarity scores. These are robust to misalignments.

Many RMCL approaches seek to enhance MCL localisation for mobile robots by using 3D point clouds generated from 3D LiDAR systems. Maken et al. [20] utilized Iterative Closest Point (ICP) with LiDAR sensors for pose estimation and localisation with MCL. Although improvements were evident in the results, the accuracy of the covariance estimation of the ICP output poses is dramatically affected by positioning accuracy. In recent deployments, machine learning models have become common in localisation systems that utilize point clouds to improve accuracy and reduce covariance estimation. For example, SegMap proposed by Dube et al. [7] uses a Convolutional Neural Network (CNN) to identify and segment geometric primitives in point clouds. However, the use of local and global descriptors leads to slow performance due to the expensive descriptor similarity networks. As a solution, SegSemMap [6] proposed by Cramariuc et al. enriches point clouds with CNN segmentation overlays from RGB cameras and adds them to the map. These approaches still require a compact and computationally efficient representation of the objects for segmentation, providing a high representation strength. Although SegMap and SemSegMap do not use an MCL approach, many segmentation-based MCL approaches have been used in combination with geometry and range, like Hendriks et al. [16]. Another example is 3D geometry in BIM models which can be used to estimate the camera pose. The models are then segmented to gain local information about the area using range readings from a 2D range-based sensor. Mendez et al. proposed SeDAR [1] which aims to simplify these models with a more robust approach to RMCL by removing the depth readings altogether. With a semantic understanding of the environment through a simple CNN, the weights of the particles are adjusted based on the likelihood of particle observations in the segmented map. In contrast to the discrete semantic label comparisons used in SeDAR [1], our proposed material comparison offers a much finer level of matching granularity. Furthermore, computing likelihoods directly from the Raman spectra in the environment without visual classification networks significantly reduces computational costs.

There have been attempts to extend the applicability of AMCL to non-visual sensing modalities. One such example is seen with haptic sensors, which can provide touch-based sensing of the environment [32]. Buchanan et al. [8] have proposed to use data segments from the surroundings with a neural network to update MCL from a map of different terrains. However, this approach requires extensive data on the environment’s terrain and prolonged pre-training, which can be challenging. Another unconventional sensing approach was proposed by Serrano et al. [9], who used the Wi-Fi signal strength of mobile devices for MCL. Although this approach is dependent on the Wi-Fi signal data strength, it can be unreliable in complex indoor environments where signals are absorbed through many walls before being sensed by the receiver. Visible Light Positioning (VLP) sensors have also been used, with W. Guan et al. [10] being able to identify objects in a room using LED patterns from lights and other devices. Despite these unconventional sensing methods, they still face challenges with even minor changes in the environment, leading to significant inaccuracies in the MCL algorithm. In contrast, Raman Spectrometers have the advantage of treating every part of the environment as an identifying landmark due to its material fingerprint.

Raman spectrometers are remote scanning instruments often used to identify materials or substances. Noise in the sensor comes in three forms, shot, dark and read noise which couples to the Poisson distribution of the data, background noise and electronics of the device, respectively [21]. Such noise can be overwhelmed by gaussian noise from the device in the process of scanning with modern probes. To identify or recognise similar Raman Spectra, most approaches use a large database with neural networks producing long training and inference times [14], [30]. In contrast, our approach simplifies and extends this by using the likelihood between stored spectra in the map and the latest spectra read from the environment. De-noising methods such as those proposed by Lussier et al. and Horgan et al. [11], [14] show the benefits of using deep-learning models to remove noise, but can also be done via other methods suggested by Zhao et al. [17]. Analytical likelihood similarity algorithms have been proposed to identify the type of material from the peaks of the spectra as proposed by Foster et al. [12]. We propose new Raman spectral similarity functions in RaSpectLoc by providing new algorithms while still delivering a system with low computational complexity. The most effective similarity functions for Raman spectra necessitate peak classification through algorithmic convolution [15]. This is infeasible for MCL approaches as they tend to take too long to process each hypothesis.

The main focus of Raman spectrometry in the literature is remote sensing of chemicals using mobile robots [13], whereas we go further and apply this research for a robot localisation approach. This paper proposes a novel approach, utilizing innovative localisation techniques and algorithmic similarity functions for mobile robots equipped with under-utilized Raman Spectral sensors.

III. PROBLEM DEFINITION

Monte-Carlo Localisation by Dellaert et al. [23] provides a framework which performs matching to generate pose hypotheses in a map. The process is as follows:

- 1) Particles are sampled, either uniformly around a space or in a Gaussian around a pose hypothesis
- 2) Particles are propagated through a motion model, typically the robot's odometry with added Gaussian noise.
- 3) Each particle is given a weight based on the accuracy of its observations against the map, typically matching a range and bearing scan line to an existing map
- 4) A re-sampling step is performed, proportional to the particle's likelihood, before the process repeats.

The current pose $x_t \in SE(2)$, of the robot, can be estimated as a set of possible samples: $\mathbb{S}_t = \{s_t^i; i = 1 \dots N\}$, given the wheel odometry measurements: $\mathbb{U}_t = \{u_j; j = 1 \dots t\}$ and depth sensor measurements: $\mathbb{Z}_t = \{z_j; j = 1 \dots t\}$ and a 2D map \mathbb{V} . If all previous odometry and depth measurements are equally weighted, the posterior probability $P(s_t^i | \mathbb{U}_t, \mathbb{Z}_t, \mathbb{V})$ can be decomposed into an online sequential process:

$$P(s_t^i | \mathbb{U}_t, \mathbb{Z}_t, \mathbb{V}) = P(s_t^i | u_t, s_{t-1}^i) P(\mathbb{Z}_{t-1} | \mathbb{U}_{t-1}, s_{t-1}^i, \mathbb{V}) \quad (1)$$

The motion model in the localisation process is determined by the odometry measurements received from the robot, represented by u_t . This information is used to "shift" the particles, assigning a likelihood based on the probability of the final position given the measured odometry. The particles are propagated based on u_t , with Gaussian noise added to model sensor noises, as follows:

$$P(s_t^i | u_t, s_{t-1}^i) \sim \mathcal{N}(u_t + s_{t-1}^i, \Upsilon_t) \quad (2)$$

where Υ_t is the covariance of the odometry and \mathbb{N} is a normal distribution across the dimensions of $SE(2)$.

The sensor model $P(\mathbb{Z}_{t-1} | \mathbb{U}_{t-1}, s_{t-1}^i, \mathbb{V})$ measures how well the range scan fits the pose hypothesis. The probability of each range-scan (r_t) is estimated under the assumption all scans are independent. Two sensor models which are commonly used with AMCL are the "Beam" and "Likelihood-field" models. The Beam model is a raycasting operation where a ray is cast starting from hypothesis s_t^i along the bearing θ_t^k and terminates when an occupied cell is reached. The likelihood is estimated as:

$$P_R(\mathbb{Z}_t^k | s_t^i) = \exp\left(-\frac{(r_t^k - r_t^{k*})^2}{2\sigma_o^2}\right) \quad (3)$$

where σ is the variance of the sensor measurement noise and $r_t = r_t^k \forall k = 0 \dots N$. The likelihood field model uses a field similar to a chamfer map to quickly estimate the distance to the nearest geometry in the floorplan, eliminating the need for costly raycasting operations. The chamfer map is defined as:

$$C_{i,j} = \min_{k,l} \|[i - k, j - l]\|, \mathbb{V}_{k,l} \neq 0 \quad (4)$$

where $C_{i,j}$ is the cost value of the chamfer map at position (i, j) , while \min is a minimum value function and points

(k, l) which iterates over the neighbors of (i, j) . In the likelihood-field model, assuming a Gaussian error distribution, the weight of each particle, s' , can be estimated as:

$$P_r(\mathbb{Z}_t^k | s_t^i, \mathbb{V}) = \exp\left(-\frac{\delta_o^2}{2\sigma_o^2}\right) \quad (5)$$

where δ_o is the value obtained from the distance map and σ_o is dictated by the noise characteristics of the sensor. During runtime, the endpoint of each bearing-range tuple is computed for each pose hypothesis, and its probability is related to the distance stored in the chamfer map. The likelihood field model is faster and produces more accurate results compared to the beam model as it is more robust to orientation errors.

RaSpectLoc is similar to AMCL but replaces many of the operations listed above for material-based sensing. Our method includes adapting a floorplan for Raman spectra, a novel spectral-based raycasting and a particle weight calculation method for each hypothesis.

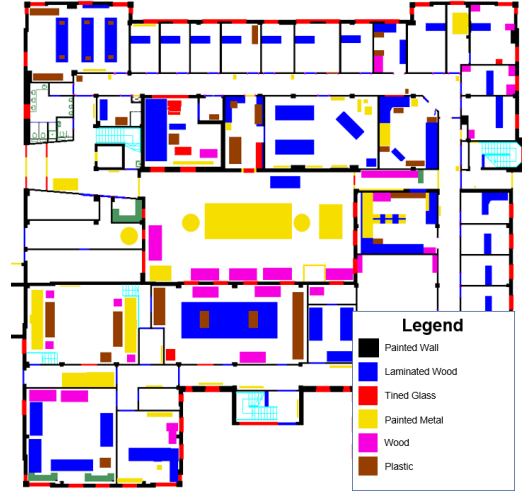


Fig. 2: Material Map used in experimentation

IV. METHODOLOGY

Fig. 3 illustrates an overview of the RaSpectLoc system. The system requires input odometry of the robot and a materials map with the Raman spectra embedded. After initialisation, section IV-B details how Raman spectra are accumulated around the robot. Section IV-C describes how the weight is calculated based on the spectral input from the mounted probe of the mobile robot. The particles propagate according to the motion model of Eq. 2. The system then repeats this process measuring the materials in the local environment around the robot with the Raman Probe and comparing them against the map.

A. Material Floorplans

The RaSpectLoc system offers a more innovative solution, utilizing spectral data obtained from a Raman probe. The system compares this data against previously recorded Raman spectra that are incorporated into the map. In this work, each cell on the floorplan is associated with a unique Raman spectrum measurement. This map is used in the ray

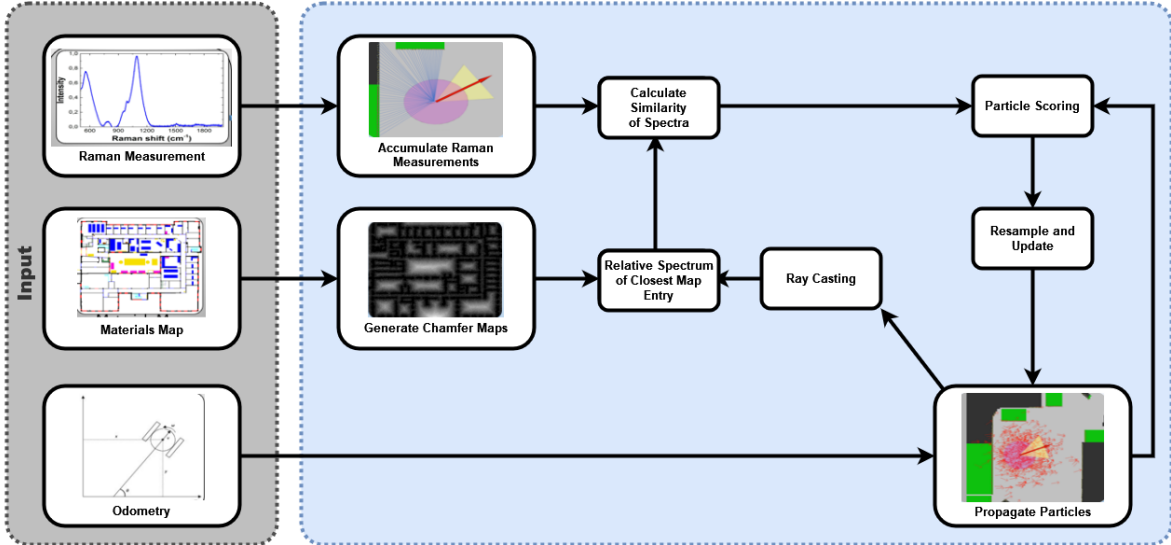


Fig. 3: RaSpectLoc system diagram

casting operation to determine the likelihood between the data and ends of the rays. The materials represented in our floorplan for this work include various spectra for painted walls, laminate wood, painted metals, wood and plastic. These materials are broadly grouped based on similarity for visualization purposes, depicted in Fig.3.

The map is converted to an occupancy grid when initially uploaded to a map server. If \mathbb{M} is a set of possible 2D positions, the map can then be defined as $\mathbb{V} = \{v_m; m \in \mathbb{M}\}$. Each embedded spectrum is formally described by $\mathbb{I}_i = \{i_0, \dots, i_n\}$ where i reflects the intensity of a particular wavelength.

B. Raman Sensor



Fig. 4: RP1000 Raman Probe [31]

At the time t , the sensor on the robot generates a message consisting of a tuple of ranges, bearings, and Raman spectra, \mathbb{Z}_t . Modern Raman spectra allow a non-invasive technique with background noise scattering removal for high-accuracy readings with little post-processing for smoothed spectral results. RaSpectLoc can also operate without an explicit range sensor (i.e. only using bearing and spectra tuples) as explained in section IV-D. In this case, raycasting is

used as with the “beam” sensor model, and the range component of the likelihood is omitted. The measurements are configured as $\mathbb{Z}_t = [\langle r_t^k, \Theta_t^k, \mathbb{I}_t^k \rangle; k = 1 \dots k]$, which represents the range from the depth camera, the bearing along the scan line, and the Raman spectrum respectively. The scan line is assumed to be parallel to the ground plane and aligned with the horizontal axis of the mobile robot’s sensor mounting point. Fig. 1 shows a simulation performed using our gazebo plug-in, highlighting the scan line and mounting point. Raman measurements are taken by rotating the joint around the robot to complete a single scan message with k measurements.

Our approach in RaSpectLoc aims to tackle the challenges faced in localisation for modular robots, which are currently dominated by geometry primitive recognition methods that require higher-end hardware. We leverage the similarity between spectra that are read from the environment after normalisation and baseline correction is performed on them. The information we gather is converted into chamfer maps, enabling the RaSpectLoc system to calculate the similarity for each particle. Despite the advanced technology, current Raman probes still have a level of noise that must be accounted for when computing the similarity between observation and map. The shot, readout, thermal background, and baseline noise are Poisson distributed and assumed to be proportional to the square root of the number of photons detected. Most Raman systems require proximity to the target in question. The IS-Instruments RP1000 probe, coupled to a HES2000, provides the capability of making Raman measurements at ranges over 1m. Such a probe is ideal for robot deployment and an ideal candidate for this application.

The noise in low response areas of spectra substantially affects the results using many comparison functions. When there is a significant number of (photonic) events, Poisson noise is indistinguishable from Gaussian noise as described by Larkin [27] and also by Lewis et al. [28]. Therefore,

we use a Squared exponential mapping function to convert spectrum distance to a likelihood, P_s , mapped between $0 \rightarrow 1$:

$$P_s(\mathbb{Z}_t | s_t, \mathbb{V}) = \exp\left(\frac{-f(\mathbb{I}_t^k, \mathbb{I}^m)^2}{K}\right) \quad (6)$$

where $f(\mathbb{I}_t^k, \mathbb{I}^m)$ is the distance function of the spectra \mathbb{I}_t^k and \mathbb{I}^m which is the spectra at endpoint of a raycasted beam m . While K is a constant used to scale the different distance metrics into the same range. Our approach supports previous state-of-the-art mathematical similarity functions for spectra [15], in addition to originally proposed methods.

C. Spectral Similarity Functions

In RaSpectLoc, various similarity functions can be used to compare Raman spectra and calculate the particle weights. We will describe these functions in detail and their advantages and disadvantages. The similarity functions described in this section are the Spectral Linear Kernel, Modified Euclidean distance [15], Wasserstein distance [18], Kullback-Leibler distance and Spectral Angle Mapping (SAM).

1) *Spectral Linear Kernel*: The spectral linear kernel (SLK), proposed by Conroy et al. [18], was designed for comparing Raman spectra. It can be expressed by:

$$f_{slk}(\mathbb{I}_t^k, \mathbb{I}^m) = \sum_{i_n^k \in \mathbb{I}_t^k} \left(i_n^k \cdot i_n^m + \sum_{j=n-W}^{j=n+W} (i_n^k - i_j^k)(i_n^m - i_j^m) \right) \quad (7)$$

where \mathbb{I}^k and \mathbb{I}^m are the input spectra. The kernel considers the original intensity values at each wave number and includes the difference between the intensity of its neighbouring points on the spectrum. The windowed difference product helps compare the relative shapes of the two spectra. However, relative to other functions mentioned in this paper it is slow and results can be poor when calculating the similarity over long windows of non-similar regions in the two spectra.

2) *Modified Euclidian Spectral Similarity metric (Mod. L2)*: The approach introduced by Khan et al. [15] is an alternative to the Euclidian similarity calculated between each intensity. This approach gives equal importance to the peaks of each spectrum but also rewards spectra that have shorter distances between the peaks. The squared distance between intensities of the spectra can be calculated via $D_{\mathbb{I}}^i = (\mathbb{I}_t^k - \mathbb{I}_t^m)^2$. This can then be used under the conditions in equation (7) to calculate the Mod. L2 distance:

$$f_{me}(\mathbb{I}_t^k, \mathbb{I}^m)^2 = \begin{cases} \sum_{i_n^k \in \mathbb{I}_t^k} \frac{1}{w} * D_{\mathbb{I}}^i, & \text{if } x_i \neq 0 \text{ AND } z_i > x_i \\ \sum_{i_n^k \in \mathbb{I}_t^k} w * D_{\mathbb{I}}^i, & \text{if } x_i = 0 \text{ AND } z_i > x_i \\ \sum_{i_n^k \in \mathbb{I}_t^k} D_{\mathbb{I}}^i, & \text{otherwise} \end{cases} \quad (8)$$

The function above calculates a distance between the spectra using a weight, w , found by:

$$w = \frac{\max(\mathbb{I}_t^k)}{1 - \max(\mathbb{I}_t^k)} \quad (9)$$

The weight w is utilized as a penalty/reward scheme for relative changes in individual intensities at each wavelength. The drawback of this function is that the weight does not work when the $\max Q$ value is less than 0.5. To resolve this issue, the weight is inverted to correct the error from the original equation.

3) *Wasserstein Distance*: The Wasserstein distance, also known as the Earth Mover's Distance, is a metric that quantifies the effort required to transform one distribution into another. It was suggested in previous literature to be an effective method of evaluating the similarity of spectra by Gao et al. [28]. Despite its accuracy, the Wasserstein distance is complex and slow compared to the other functions mentioned in this section. This is explored in Section V.

$$f_W(\mathbb{I}_t^k, \mathbb{I}^m) = \min\{E(t) | t : \mathbb{I}_t^k \rightarrow \mathbb{I}^m\} \quad (10)$$

Where t is a transportation plan that maps the elements of \mathbb{I}_t^k to the elements of \mathbb{I}^m , and $E(t)$ is the cost of the transportation plan t , which is typically represented as the sum of the distances between the corresponding vector elements in \mathbb{I}_t^k and \mathbb{I}^m .

4) *Kullback-Leibler Divergence*: The KL-Divergence is a highly efficient yet effective distance to compare two equally binned distributions. Formally it is a generalisation on the L2 Norm and is defined as the

$$f_{kl}(\mathbb{I}_t^k || \mathbb{I}^m) = \sum_{i_n^k \in \mathbb{I}_t^k} i_n^k \log \frac{i_n^k}{i_n^m} \quad (11)$$

The KL-divergence quantifies the information lost when approximating \mathbb{I}_t^k with \mathbb{I}^m . It is a non-symmetric metric, meaning that $f_{KL}(\mathbb{I}_t^k || \mathbb{I}^m) \neq f_{KL}(\mathbb{I}_t^m || \mathbb{I}^k)$. The KL divergence is sensitive to tiny differences between distributions, which is desirable in the context of Raman spectra, where little variations in intensity values can indicate the presence of different materials.

5) *Spectral Angle Mapping*: The Spectral Angle Mapping (SAM) function is a similarity metric commonly used to calculate the angle between two spectra in a high-dimensional space. Each dimension represents a different Raman shift. The angle between the two spectra shows their similarity, with a smaller angle indicating higher similarity. Formally the function can be expressed by:

$$f_{sam}(\mathbb{I}_t^k || \mathbb{I}^m) = \cos^{-1} \left(\frac{\mathbb{I}_t^k \cdot \mathbb{I}^m}{\|\mathbb{I}_t^k\| \|\mathbb{I}^m\|} \right) \quad (12)$$

SAM is founded on the idea that similar spectra should have similar shapes and relative intensities across different Raman shifts benefiting from the similarity angle returned around the shape of each peak. Furthermore, the function is also scale-invariant, meaning it is not affected by differences in the absolute intensity of the spectra. However, SAM is quite

sensitive to noise and variations of the baseline of the Raman spectra since they affect the shape of the peaks. SAM also fails to consider the intensity of the spectra and assumes spectra are linearly mixed. This assumption may not be the case for some materials in the environment.

D. Sensor Models

We adapt both the “beam” and “likelihood-field” sensor models for use with Raman Spectra.

The Beam model depends on ranges as part of the weight for each particle. This is because the beam model depends on the depth of the scan from the environment and the end of the beam projected into the map, \mathbb{V} . While the probability of an observation given the map and pose for the beam model can be expressed by a weighted sum of the two probabilities:

$$P_r(\mathbb{Z}_t^k | s_t^i, \mathbb{V}) = \epsilon_R P_R(\mathbb{Z}_t^k | s_t^i, \mathbb{V}) + \epsilon_m P_m(\mathbb{Z}_t^k | s_t^i, \mathbb{V}) \quad (13)$$

For the Beam-model of RaSpectLoc, the likelihood of a particle is defined by two user-defined weights, ϵ_R and ϵ_m . When ϵ_m is zero, the likelihood is equivalent to the standard RMCL. The optimal weight configuration is evaluated in Section V. Unlike range scanners, the standard deviation σ' of equation 2 cannot be linked to the sensor’s physical properties. Instead, it is estimated from the prior of each material spectra on the map, avoiding the need for tuning. However, the Beam Model has drawbacks to its approach as it lacks smoothness since it depends on the resolution of the map. This is worse for mobile robots as it becomes more memory dependent to increase the number of beams and resolution of the map.

The likelihood-field model provides a smoother approach providing gradients between each cell. Furthermore, it can use ranges, bearing and material likelihood or only bearing and material likelihoods. The likelihood field model in our approach generates a chamfer map to the nearest occupied cell for each material based on equation 2:

$$C_{i,j}(\mathbb{I}) = \min_{k,l} (C_{k,l}(\mathbb{I}) + f(\mathbb{I}_{i,j}, \mathbb{I}_{k,l})) \quad (14)$$

Where f is the chosen spectral likelihood function mentioned in Section III-B. For every $\mathbb{Z}_t = [(\Theta_t^k, \mathbb{I}_t^k); k = 1 \dots k]$, the raycasting is performed. Instead of comparing the range with the map, the spectral likelihood can be used to calculate the cost:

$$P_r(\mathbb{Z}_t^k | s_t^i, \mathbb{V}) = P_m(\mathbb{Z}_t^k | s_t^i, \mathbb{V}) \quad (15)$$

This method is a combination of the beam and likelihood model. Hence, in Eq. 13, when ϵ_R is zero, the approach solely relies on the Raman spectra information in the floor-plan and material similarity likelihood.

The nearest occupied cell is identified through similar AMCL range-based raycasting. The distances obtained are utilized to improve the smoothness of Eq. 13, which suggests that the likelihood of the spectra similarity is proportional to the distribution of material similarity in terms of angle. As a result, this approach is scale-invariant, provided that the aspect ratio of the map is maintained.

V. EXPERIMENTS AND RESULTS

Our research focuses on MCL-based localisation using Raman Spectra from a single probe. This section presents the overall performance improvements of our system over other state-of-the-art localisation approaches. This dataset will be released to the community to support future research. RaSpectLoc can choose between the beam or the likelihood field model which are described in Section IV-D. The ablation and parameter exploration section, V-B, illustrates the difference in the performance of RaSpectLoc due to changes in likelihood functions and configurations of ϵ . Whereas the Quantitative Results section V-C will compare RaSpectLoc with various state-of-the-art baselines. These include, AMCL [23], SeDAR [1], visual pose recognition (PoseNet [25]), range-based SLAM (GMapping [26]) and Monocular SLAM (ORB_SLAM 3 [24]).

Spectral Similarity Function Analysis (m)						
Function	RMSE	Mean	Median	SD	Min	Max
SLK	0.18	0.17	0.16	0.07	0.03	0.38
Mod.L2	0.14	0.13	0.12	0.06	0.03	0.41
Wasserstein	0.17	0.14	0.13	0.09	0.02	0.44
KL.D	0.15	0.13	0.11	0.08	0.02	0.34
SAM	0.18	0.16	0.14	0.09	0.03	0.42

TABLE I: Spectral Similarity Function Experiments

A. Experimental Simulation Setup

The following sections evaluate the performance of RaSpectLoc using a dataset consisting of 5 mobile robot trajectories around a University. The dataset includes Raman Spectra, RGBD images and odometry information. Additionally, a material-embedded floorplan with Raman spectra at each coordinate is provided.

To quantitatively evaluate the performance of RaSpectLoc and state-of-the-art methods, the Absolute Trajectory Error (ATE) and the Relative Pose Error (RPE) metrics presented by Strum et al. [22] are used. The ATE calculates the rigid transformation estimation between the two paths. The RPE evaluates the relative rotation and relative translation for all points in the path, ignoring the effects of drift over time in the trajectory. The RMSE, mean, median, standard deviation, minimum distance and maximum distance values of the residual position error are provided for the resulting ATE and RPE of the experiments. For evaluation in Table I, AMCL, SeDAR [1] and RaSpectLoc is given a coarse initialisation with standard deviations of 2.0m in (x, y) and 2.0 radians in θ . The system was run with a maximum of 1000 particles placed around the covariance ellipse. The error was recorded for each new set of accumulated Raman spectra into the published tuple \mathbb{Z}_t .

B. Ablation and parameter exploration

This section focuses on the performance of the RaSpectLoc system using various spectral similarity functions and optimal weights, ϵ , to accurately compare against modern localisation approaches. We evaluate the performance using the ATE and RPE of the produced trajectories with the likelihood-field model. Additionally, we consider the accuracy of determining the correct spectrum after adding

Function	Translational RMSE(m)	Rotational RMSE(deg)	Translational Mean(m)	Rotational Mean(deg)	Translational Median(m)	Rotational Median(deg)	Translational SD(m)	Rotational SD(deg)	Translational Max(m)	Rotational Max(deg)
Wasserstein	0.06	2.02	0.05	1.39	0.04	0.91	0.03	1.46	0.22	9.67
SLK	0.07	2.32	0.06	1.58	0.05	0.97	0.04	1.70	0.20	11.9
KL.D	0.05	2.03	0.04	1.34	0.04	0.81	0.02	1.52	0.16	10.86
Mod.L2	0.06	2.23	0.05	1.5	0.04	0.95	0.03	1.64	0.17	10.17
SAM	0.07	2.34	0.06	1.57	0.05	0.96	0.04	1.73	0.31	9.94

TABLE II: Spectral Similarity Function Analysis (RPE)

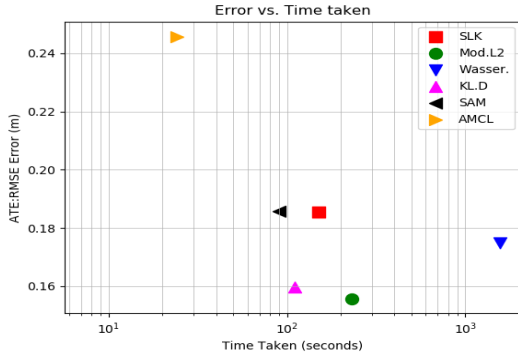


Fig. 5: ATE:RMSE error against time taken to complete the trajectory

Weights		RMSE	Mean	Median	SD	Min	Max
ϵ_R	ϵ_m						
0.2	0.8	0.19	0.17	0.15	0.09	0.01	0.43
0.4	0.6	0.19	0.17	0.16	0.09	0.04	0.38
0.5	0.5	0.18	0.15	0.13	0.10	0.03	0.50
0.6	0.4	0.22	0.20	0.20	0.10	0.08	0.47
0.8	0.2	0.21	0.19	0.19	0.11	0.08	0.45

TABLE III: Weight experiments ablation tests

Raman-probe shot noise. This approach should provide a more realistic evaluation of the similarity functions compared to post-processed spectra that do not contain these intensity artefacts or noise. For Eq. 12 the value of β is set to 1. The results in Tables I and II indicate that the Mod. L2 function produces the best accuracy and performance for RaSpectLoc. The results of the Mod. L2 function is because it provides a descriptive way of examining the similarity between peaks with a reward and punishment scheme. In testing, the Wasserstein distance takes significantly longer to process than other functions, often leading to performance issues during periods of high message frequency. This is highlighted in Fig. 5 by the logarithmic axis for the time taken to complete a 5min 55sec rosbag recording. One obvious limitation of RaSpectLoc from this graph is the increased run-time compared to AMCL, although this comes with a significant reduction in error.

As shown in Table III, the performance of RaSpectLoc varies with the weight distribution. The best-performing weights are 0.5 for the material and ranges.

Compared to Table I the results still show that a ‘spectra-only’ approach outweighs the combined range and spectral data approach seen in Table III. This result is likely due to geometric cues in a hallway environment. These are relatively weak compared to the fine-grained spectral data that provide

a greater understanding of the current pose for the system. Findings in these experiments are highlighted in Section V-C, where we will compare RaSpectLoc’s performance against other state-of-the-art systems.

C. Quantitative Results

Tables IV and V showcase the results of our tests against various state-of-the-art localisation approaches in terms of Absolute Trajectory Error (ATE) and Relative Pose Error (RPE) metrics. Experiments are compared with ϵ_R and ϵ_m set to 0.5 and 0.5, respectively, as per the optimal configuration found in Section V-B for our combined localisation model with the Mod. L2 likelihood function (Section IV-C2). These tests are performed using the same RGB images and depth scan messages along a single trajectory. This separation allows us to compare our results against three non-MCL approaches: ORBSLAM, GMapping and PoseNet.

Approach	RMSE	Mean	Median	SD	Min	Max
AMCL	0.24	0.21	0.2	0.11	0.04	0.95
SeDAR	0.19	0.16	0.14	0.10	0.02	0.55
ORBSLAM3 (MONO)	7.17	6.54	6.81	2.93	0.55	11.57
GMapping	0.71	0.63	0.57	0.31	0.32	1.43
Posenet	4.64	2.58	1.42	3.85	0.02	25.66
RaSpectLoc (combined)	0.18	0.15	0.13	0.10	0.03	0.50
RaSpectLoc (Materials only)	0.14	0.13	0.12	0.06	0.03	0.41

TABLE IV: ATE: Baseline localisation comparisons

Approach	RMSE	Mean	Median	SD	Max
AMCL	0.07/2.68	0.05/1.68	0.04/0.68	0.05/2.09	0.37/11.89
SeDAR	0.06/2.21	0.04/1.5	0.04/0.9	0.03/1.72	0.2/12.47
ORBSLAM3 (MONO)	0.42/14.7	0.41/5.78	0.41/0.86	0.10/13.58	0.77/61.60
GMapping	0.40/26.65	0.18/13.05	0.31/8.5	0.35/23.23	1.49/92.65
Posenet	4.72/47.70	2.23/28.40	0.89/9.8	4.16/38.33	28.96/179.6
RaSpectLoc (combined)	0.06/2.31	0.06/1.52	0.05/0.94	0.03/1.74	0.21/9.77
RaSpectLoc (Materials)	0.05/2.03	0.04/1.34	0.04/0.81	0.02/1.52	0.16/10.86

TABLE V: RPE: Baseline localisation comparisons

Our experiments revealed that RaSpectLoc consistently outperforms the other approaches in accuracy, robustness, ATE and RPE metrics. The results showed that our method was able to accurately estimate the 3-DoF pose of the robot even in challenging environments with noise and dynamic obstacles. Furthermore, our experiments highlighted the ability of RaSpectLoc to scale to larger environments and handle noise-riddled Raman Spectra, making it a highly competitive solution for real-world applications.

VI. CONCLUSION

In conclusion, this paper has presented a novel approach for mobile robot localisation that harnesses the power of Raman spectra. With its ability to recognise subtle material-composition landmarks, RaSpectLoc can outperform traditional RGBD systems. This recognition has significant implications for localisation in challenging environments, evident

from our practical experimental section (V-C). The future of this field is even more promising, as we envision a future where robots can exploit spectral data to navigate and map hazardous environments, such as nuclear decommissioning sites. Such a SLAM system would provide a more detailed understanding of the environment while reducing the risk posed to human workers. Raman spectra could also assist in detecting anomalies in these high-risk locations with the potential to revolutionize the field and greatly benefit society.

VII. ACKNOWLEDGEMENTS

This work was partially supported by Industrial 3D Robotics (I3D), IS-Instruments and partially funded by the EPSRC under grant agreement EP/S035761/1.

REFERENCES

- [1] Mendez, O., Hadfield, S., Pugeault, N. and Bowden, R. (2020) SeDAR: reading floorplans like a human—using deep learning to enable human-inspired localisation. *International Journal of Computer Vision*, 128(5), pp. 1286–1310. (doi: 10.1007/s11263-019-01239-4)
- [2] D. Talwar and S. Jung, "Particle Filter-based Localization of a Mobile Robot by Using a Single Lidar Sensor under SLAM in ROS Environment," 2019 19th International Conference on Control, Automation and Systems (ICCAS), 2019, pp. 1112–1115, doi: 10.23919/ICCAS47443.2019.8971555.
- [3] Thale, Sumegh and Prabhu, Mihir and Thakur, Pranjali and Kadam, Pratik. (2020). ROS-based SLAM implementation for Autonomous navigation using Turtlebot. *ITM Web of Conferences*. 32. 01011. 10.1051/itmconf/20203201011.
- [4] Stanford Artificial Intelligence Laboratory et al., 2018. *Robotic Operating System*, Available at: <https://www.ros.org>.
- [5] S. Z. Ahmed, V. B. Saputra, S. Verma, K. Zhang and A. H. Adiwahono, "Sparse-3D Lidar Outdoor Map-Based Autonomous Vehicle Localization," 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2019, pp. 1614–1619, doi: 10.1109/IROS40897.2019.8967596.
- [6] A. Cramariuc et al., "SemSegMap – 3D Segment-based Semantic Localization," 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2021, pp. 1183–1190, doi: 10.1109/IROS51168.2021.9636156.
- [7] Dube, Renaud & Cramariuc, Andrei & Dugas, Daniel & Nieto, Juan & Siegwart, Roland & Cadena, Cesar. (2018). *SegMap: 3D Segment Mapping using Data-Driven Descriptors*. 10.15607/RSS.2018.XIV.003.
- [8] Buchanan, R., Bednarek, J., Camurri, M. et al. Navigating by touch: haptic Monte Carlo localization via geometric sensing and terrain classification. *Auton Robot* 45, 843–857 (2021). <https://doi.org/10.1007/s10514-021-10013-w>
- [9] Serrano, Oscar & Cañas Plaza, José & Matellán, Vicente & Rodero, Luis. (2014). Robot localization using WiFi signal without intensity map.
- [10] W. Guan, S. Chen, S. Wen, Z. Tan, H. Song and W. Hou, "High-Accuracy Robot Indoor Localization Scheme Based on Robot Operating System Using Visible Light Positioning" in *IEEE Photonics Journal*, vol. 12, no. 2, pp. 1–16, April 2020, doi: 10.1109/JPHOT.2020.2981485.
- [11] Conor C. Horgan, Magnus Jensen, Anika Nagelkerke, Jean-PhillipeSt-Pierre, Tom Vercauteren, Molly M. Stevens, and Mads S. Bergholt, "High-throughput molecular imaging via deep learning enabled Raman spectroscopy", *Analytical Chemistry* 2021, 93, 48, 15850–15860.
- [12] Gerasimov, V. A., Gurovich, A. M., Kostrin, D. K., Selivanov, L. M., Simon, V. A., Stuchenkov, A. B., Paltcev, A. V., and Uhov, A. A. (2016). Raman spectroscopy for identification of woodspecies. *Journal of Physics: Conference Series*, 741, 012131. <https://doi.org/10.1088/1742-6596/741/1/012131>.
- [13] Foster, M, Wharton, M, Brooks, W, Goundry, M, Warren, C, Storey, J. Remote sensing of chemical agents within nuclear facilities using Raman spectroscopy. *J Raman Spectrosc.* 2020; 51: 2543– 2551. <https://doi.org/10.1002/jrs.6016>
- [14] Félix Lussier, Vincent Thibault, Benjamin Charron, Gregory Q. Wallace, Jean-Francois Masson, Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering, *TrAC Trends in Analytical Chemistry*, Volume 124, 2020, 115796, ISSN 0165-9936, <https://doi.org/10.1016/j.trac.2019.115796>.
- [15] Shehroz S. Khan, Michael G. Madden, New similarity metrics for Raman spectroscopy, *Chemometrics and Intelligent Laboratory Systems*, <https://doi.org/10.1016/j.chemolab.2012.03.007>.
- [16] R. W. M. Hendriks, P. Pauwels, E. Torta, H. P. J. Bruyninckx and M. J. G. van de Molengraft, "Connecting Semantic Building Information Models and Robotics: An application to 2D LiDAR-based localization," 2021 IEEE International Conference on Robotics and Automation (ICRA), 2021, pp. 11654–11660, doi: 10.1109/ICRA48506.2021.9561129.
- [17] Zhao, X.Y. & Liu, G.Y. & Sui, Y.T. & Xu, M. & Tong, L.. (2020). Denoising method for Raman spectra with low signal-to-noise ratio based on feature extraction. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. 250. 119374. 10.1016/j.saa.2020.119374.
- [18] J. Conroy, A.G. Ryder, M.N. Leger, K. Hennessy, M.G. Madden, "Qualitative and quantitative analysis of chlorinated solvents using Raman spectroscopy and machine learning", *Proc. of SPIE, the International Society for Optical Engineering*, 2005, pp. 131–142.
- [19] R. Valencia, J. Saarinen, H. Andreasson, J. Vallvé, J. Andrade-Cetto and A. J. Lilienthal, "Localization in highly dynamic environments using dual-timescale NDT-MCL", *Proc. ICRA*, pp. 3956–3962, May 2014.
- [20] Maken FA, Ramos F, Ott L. Bayesian iterative closest point for mobile robot localization. *The International Journal of Robotics Research*. 2022;41(9-10):851-874. doi:10.1177/02783649221101417.
- [21] Jahn, I.J.; Grjasnow, A.; John, H.; Weber, K.; Popp, J.; Hauswald, W. Noise Sources and Requirements for Confocal Raman Spectrometers in Biosensor Applications. *Sensors* 2021, 21, 5067. <https://doi.org/10.3390/s21155067>
- [22] Sturm, J., Engelhard, N., Endres, F., Burgard, W., Cremers, D. (2012). A benchmark for the evaluation of RGB-D SLAM systems. In *International conference on intelligent robots and systems (IROS)* (pp. 573–580). IEEE/RSJ. <https://doi.org/10.1109/IROS.2012.6385773>.
- [23] Dellaert, F., Fox, D., Burgard, W., & Thrun, S. (1999). Monte Carlo localization for mobile robots. In *International conference on robotics and automation (ICRA)* (pp. 1322–1328). IEEE, Detroit.
- [24] Campos, C., Elvira, R., Rodríguez, J. J., Montiel, J. M., & Tardós, J. D. (2020). ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM. *arXiv*. <https://doi.org/10.1109/TRO.2021.3075644>
- [25] Kendall, A., Grimes, M. & Cipolla, R. (2015). PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization.. *ICCV* (p./pp. 2938–2946), : IEEE Computer Society. ISBN: 978-1-4673-8391-2
- [26] Giorgio Grisetti, Cyrill Stachniss, and Wolfram Burgard: Improved Techniques for Grid Mapping with Rao-Blackwellized Particle Filters, *IEEE Transactions on Robotics*, Volume 23, pages 34–46, 2007
- [27] P. J. Larkin, "Infrared and Raman Spectroscopy: Principles and Spectral Interpretation," Oxford: Elsevier, 2011.
- [28] I. R. Lewis and H. G. M. Edwards (eds.), "Handbook of Raman Spectroscopy," Bristol: Taylor & Francis, 2001.
- [29] J. Gao, Y. Hu, and Y. Zhao, "A novel approach for raman spectra comparison: Wasserstein distance with local constraints," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 239, 2020, 118448.
- [30] Liu, J., Osadchy, M., Ashton, L., Foster, M., Solomon, C.J. and Gibson, S.J. (2017). Deep convolutional neural networks for Raman spectrum recognition: a unified solution. *The Analyst*, 142(21), pp.4067–4074. doi:<https://doi.org/10.1039/c7an01371j>.
- [31] Marsh-Collings, S. (n.d.). Raman Probes. [online] IS-Instruments. Available at: <https://is-instruments.com/raman-probe/> [Accessed 4 Jul. 2023].
- [32] M. Łysakowski, M. R. Nowicki, R. Buchanan, M. Camurri, M. Fallon and K. Walas, "Unsupervised Learning of Terrain Representations for Haptic Monte Carlo Localization," 2022 International Conference on Robotics and Automation (ICRA), Philadelphia, PA, USA, 2022, pp. 4642–4648, doi: 10.1109/ICRA46639.2022.9812296.