

Know Your Limits: Embedding Localiser Performance Models in Teach and Repeat Maps

Winston Churchill, Chi Hay Tong, Corina Gurău, Ingmar Posner and Paul Newman

Abstract—This paper is about building maps which not only contain the traditional information useful for localising — such as point features — but also embeds a spatial model of *expected* localiser performance. This often overlooked second-order information provides vital context when it comes to map use and planning. Our motivation here is to improve the performance of the popular Teach and Repeat paradigm [1] which has been shown to enable truly large-scale field operation. When using the taught route for localisation, it is often assumed the robot is following exactly, or is sufficiently close to, the original path, enabling successful localisation. However, what happens if it is not possible, or not desirable to exactly follow the mapped path? How far off the beaten track can the robot travel before it gets lost? We present an approach for assessing this localisation area around a taught route, which we refer to as the *localisation envelope*. Using a combination of physical sampling and a Gaussian Process model, we are able to accurately predict the localisation performance at unseen points.

I. INTRODUCTION

We need robots to be useful over large spatial and temporal scales. For many application domains, the Teach and Repeat paradigm [1] offers a way forward and this paper considers how the framework can be extended even further. By manually “teaching” the robot a desired route (e.g. by driving it through the environment), it can autonomously repeat the route itself many times. The core attraction of Teach and Repeat is the opportunity it affords to duck the issue of building a single-coordinate frame map (which is still hard at arbitrary scales). It is the route network that matters, not a Cartesian projection of it. Put differently, some mobile robotics applications allow us to care not about the geometry of the internal world model as long as the vehicle can *reliably and repeatably* use it to traverse through and operate in its workspace. This is the point we wish to emphasise in this work. We want to adorn the Teach and Repeat map with a locally varying spatial model which explicitly captures the ability of the map to support localisation despite deviations (planned or otherwise) from the taught route. In this way we can inform planning and control decisions to reduce the chances of localisation failure.

In the original Teach and Repeat approach, it was expected that the robot travelled sufficiently close to the taught path to enable localisation. Any localisation failures were assumed to come from external factors, e.g. the appearance of the scene changing relative to the mapped appearance. The question we would like to ask the map is: “how far can we travel away from the map and still remain localised?” This information

Authors are members of the Mobile Robotics Group, University of Oxford, Oxford, England; {winston, chi, corina, hip, pnewman}@robots.ox.ac.uk

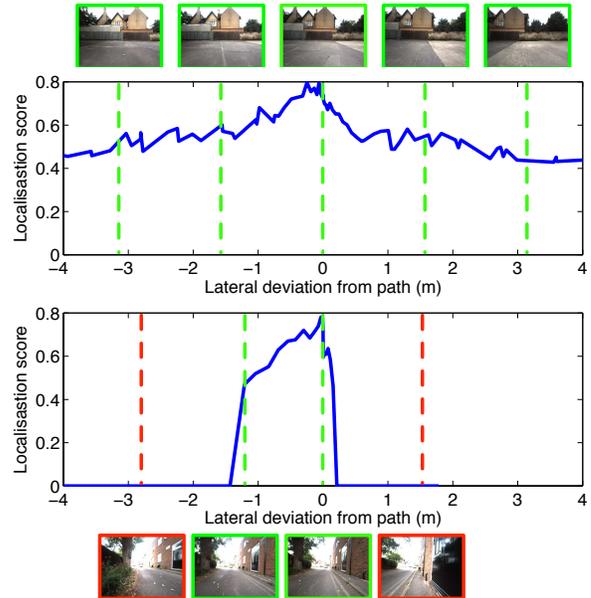


Fig. 1. Not all places are created equal in the eyes of a localiser. Localisation performance is strongly affected by the local environment. The top graphic demonstrates a space where large lateral deviations from the map can be tolerated by our localisation system, while the lower graphic gives an example where only a small, asymmetric envelope is usable. Green borders indicate successful localisations, while red are failures. In this paper we look to capture this envelope of localisation via Gaussian Process regression, informed by physical sampling.

is invaluable, as it informs the robot how far it can venture. For example, can the robot successfully navigate around obstacles in the taught route? Which parts of the route require precise navigation (as it is easy to get lost in these areas), and which can be driven approximately as the localisation is robust to large deviations from the path? Being able to answer these questions would also make for a more informed planner. Venturing too far is undesirable as once the robot is lost, it must back-track to somewhere known and try again (if that is even possible).

The ability to localise against a visual map is a function of many parameters. One of the most influential is the position of the robot relative to the most relevant map point, hence why it has typically been assumed the robot drives along the previously mapped route. Other factors include the shape of the trajectory, the structure of the scene, the type of localisation system used and appearance change. Two examples of localisation performance as a function of lateral offset from the map are shown in Figure 1. Note how the localisation score is different and asymmetric about the

centre line in each place.

In this work, we will look to model the envelope of a single visual teach pass. We will refer to the area around a taught path that can be successfully localised in as the *localisation envelope*. While this additional information would be useful for any navigation system, this work focuses on stereo vision with sparse features. Our approach is formed from two parts. The first models the envelope using a Gaussian Process (GP). This takes as input a position relative to the path, and returns a localisation score, as well as its associated uncertainty. This uncertainty is vital for consumers of the map, such as planners. The second part involves physically sampling the area around the path, to provide training data for the GP.

II. RELATED WORKS

This work looks to compute the localisation envelope surrounding a visual map. The closest related work is the stereo vision Teach and Repeat system from Furgale and Barfoot [1]. They calculate the localisation envelope (they refer to it as the *convergence properties*) by manually placing the camera at continually divergent positions until failure at several places along the route. They then average the results to get a single, constant envelope at all locations. We are looking to extend this in two ways. First we aim to create a place-dependent model of the envelope that allows us to express the variability of the envelope along the path. Second, we propose that the model be continually updated, either through explicit (potentially autonomous) sampling of the route, or subsequent repeat runs. Exploiting the mobility of the robot to physically sample the localisation envelope is similar to the work by Stenning *et al.* [2], who use their robot to physically embody a rapidly exploring random tree (RRT) [3].

Also extending the Teach and Repeat paradigm is the recent work of Krüsi *et al.* [4] who implement a laser-based system. They show that the laser-based approach retains higher accuracy when navigating around obstacles. However, their study only focuses on accuracy, and does not explicitly model the limits of the system. This information is useful for planning.

In addition to localisation quality, our localisation envelope encodes a measure of uncertainty. The presence of uncertainty has been a topic of great interest in the planning community. Melchior and Simmons account for uncertainty in the planning domain by using Particles in the RRT algorithm [5]. Lunders *et al.* use Chance Constrained RRTs to fold in uncertainty, using probabilistically informed feasibility checks to assess path validity. Prentice and Roy present belief road maps [6], a variant on probabilistic road maps [7], that can efficiently produce plans in belief space and results in improved performance and accuracy. While there exist several algorithms to include the localisation uncertainty into planning, it is often not done in practice, as localisation schemes typically do not offer this information to the planners.

One of the key factors in our sparse-feature-based system is the ability to return correct matches given our choice of

feature detector and descriptor. Feature detector/descriptor pairs that support matching under larger image distortions may potentially allow for a larger localisation envelope. A significant amount of research has been spent on developing and assessing the performance of various algorithms [8][9]. Our work may facilitate a more intuitive comparison.

III. PRELIMINARIES

We begin by briefly describing our Teach and Repeat system, for which we are trying to estimate the localisation envelope. It is similar to the approach described by Furgale and Barfoot [1]. At its heart is a sparse-feature, stereo visual odometry (VO) system. VO computes the ego-motion between two stereo frames and has been developed by many researchers over several decades [10][11][12][13]. Stereo VO is popular in robotics due to its ability to provide metric pose estimates, and while the cumulative positional error can grow without bound, over a local window they are sufficiently accurate for robotic tasks [14]. It is also typically more accurate than wheel odometry [13][15].

There are many variants on the core VO idea. Our implementation uses FAST [16] corners combined with BRIEF descriptors [17]. While the BRIEF descriptor does not offer the same level of robustness to rotation as gradient-based methods such as SIFT [18] and SURF [18], it is significantly faster at extracting and matching. This means that our VO system can run at frame rate on a modest CPU-only machine. RANSAC [19], followed by non-linear least squares optimisation is used to compute the ego-motion.

In addition to VO, the teach phase also saves keyframes, which contain 3D landmarks with associated BRIEF descriptors, at regular intervals while the robot is manually driven along a desired route. The exact same VO pipeline, including descriptor matching and egomotion estimation, is then used for localisation when repeating the route. The only difference is that the latest image from the camera — the live frame — is matched against a stored keyframe, not the previous live frame. This outputs a 6 Degree of Freedom (DoF) pose, \mathbf{p} , relative to the teach pass. For a more detailed description of Teach and Repeat, the reader is referred to [1][20].

IV. APPROACH

A. Modelling

Given a visual map, \mathcal{M} , we wish to estimate its localisation envelope. We describe a position relative to the path as a feature vector \mathbf{x} , which is defined as a function of both the map and pose: $\mathbf{x} = f(\mathcal{M}, \mathbf{p})$. Given this feature vector, we wish to predict y , a score of localisation quality at \mathbf{p} . In this work, y is the normalised number of matched keyframe landmarks and is a measure of localisation quality. A score of one indicates that all keyframe landmarks will be matched and will result in high localisation quality, while a score of zero suggests that no matches will be found and that localisation will fail. Additionally, we want a measure of uncertainty on y , as we may want to incorporate the certainty of localisation quality in planning.

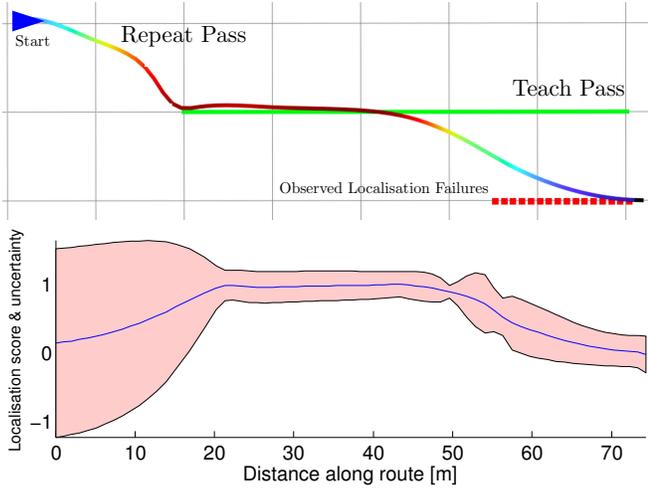


Fig. 2. The top figure shows a taught route (straight green line) and a set of observed localisation failures (red dots). A planned repeat pass (multicoloured line) is also shown, with blue indicating a low localisation score, and red a high localisation score. The bottom graph shows the distance along the route versus the localisation score, with accompanying uncertainty. At the beginning of the repeat pass the robot is situated away from the path. As a result the localisation score is low, but the uncertainty is high. As the trajectory approaches the path, the localisation score improves with increasing confidence. Towards the end of the route the trajectory passes through an area where localisation failures have previously been observed. In this case, the localisation score drops, but the confidence remains high due to the previous observations. There is a temporary increase in uncertainty as the path moves between observations.

To achieve this we use Gaussian Process regression. GPs offer supervised non-parametric learning [21] and are attractive for two reasons. Firstly, they do not require an explicit model of how the localisation envelope will behave. Secondly, they naturally handle uncertainties. We begin with a brief overview of the GP framework.

A GP is described by a mean function, $\mu(\mathbf{x})$, and a covariance function, $k(\mathbf{x}, \mathbf{x}')$. A set of input data $\mathcal{T} = \{\{\mathbf{x}_i, y_i\}_{i=1}^n\}$ is used to train the GP. The set of all input vectors is denoted as $X = \{\mathbf{x}_i\}_{i=1}^n$, and similarly, $\mathbf{y} = \{y_i\}_{i=1}^n$. Then, given a set of query points X_* , we wish to determine \mathbf{y}_* . The joint distribution over all variables is described by

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}_* \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu(X) \\ \mu(X_*) \end{bmatrix}, \begin{bmatrix} K(X, X) & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix} \right), \quad (1)$$

where $K(X, X_*)$ is the $n \times n_*$ matrix of all covariance pairs for X, X_* . The other $K(\cdot)$ matrices are similarly defined. We account for observation noise by adding $\sigma_n^2 I$ in $K(X, X)$ and $K(X_*, X_*)$, where σ_n is the observation noise, and I is the identity matrix.

As is often done with GPs, we assume the mean function, $\mu(\mathbf{x})$, is zero. This is appropriate in our case, as in the absence of any observations it is best to be conservative and assume we cannot localise at an arbitrary, unseen point in space.

We estimate \mathbf{y}_* by conditioning our prediction on the training data. This results in the well known formula for

Gaussian Process regression:

$$p(\mathbf{y}_* | \mathbf{y}, X, X_*) \sim \mathcal{N}(\bar{\mathbf{y}}_*, \mathbb{V}[\mathbf{y}_*]), \quad (2)$$

where the best estimate of \mathbf{y}_* is the mean of the distribution

$$\bar{\mathbf{y}}_* = K_* K^{-1} \mathbf{y}, \quad (3)$$

and the variance associated with this value is

$$\mathbb{V}[\mathbf{y}_*] = K_{**} - K_* K^{-1} K_*^T, \quad (4)$$

where we have used the abbreviation $K = K(X, X)$, $K_* = K(X_*, X)$ and $K_{**} = K(X_*, X_*)$.

The final key piece of the GP framework is the choice of covariance function, $k(\cdot)$. As we assumed a zero mean function, the behaviour of our GP is entirely defined by $k(\cdot)$. There exists a range of options, which allows for many different behaviours. It allows us to encode how we expect the input data, \mathbf{x} and \mathbf{x}' to relate to each other. Typically, $k(\cdot)$ also requires a set of free parameters, represented by the vector θ . In the GP framework these are referred to as *hyperparameters*. The best choice of θ is achieved by maximising the log marginal likelihood

$$\log p(\mathbf{y} | \mathbf{x}, \theta) = -\frac{1}{2} \mathbf{y}^T K^{-1} \mathbf{y} - \frac{1}{2} \log |K| - \frac{n}{2} \log 2\pi. \quad (5)$$

In this work we are interested in modelling the performance of our localisation system as we deviate from the taught path. While our VO system is capable of full 6 DoF pose estimation, we will only assess the localisation performance of the system in the first two degrees, which we refer to as x_1 and x_2 . This is because ground based planners typically only operate in 3 DoF Cartesian space, $[x_1, x_2, \theta]$. In this work we make the simplifying assumption that the robot is always parallel with the path. While this is a limitation of this study, our experience is that the localisation envelope for rotation is relatively small. This is because pure rotation of the camera rapidly changes the viewpoint of the imaged scene, which leads to localisation failure.

We parametrise our teach pass with a cubic spline. Given this representation, for a query point \mathbf{p}_* , we can compute the closest point on the spline. This provides us with the signed cross track error to the teach trajectory, s , a point on the spline, defined by the spline parameter d , and the local curvature, c , at that point. This allows us to capture two important factors affecting localisation, the lateral offset to the path (with s) and the paths current curvature (with c). We can also uniquely identify a point on the path via d . These three parameters form the input vector to our GP framework:

$$\mathbf{x} = [s \ d \ c]^T. \quad (6)$$

For $k(\cdot)$, we chose the commonly used Matérn covariance function [21], with $\nu = \frac{3}{2}$. This takes the form

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \left(1 + \sqrt{3} M \tau \right) \exp \left(-\sqrt{3} M \tau \right), \quad (7)$$

where $\tau = \|\mathbf{x} - \mathbf{x}'\|_2$ is the l_2 -norm between the two input vectors. Our hyperparameter vector is defined as

$$\theta = [l_s \ l_d \ l_c \ \sigma_f]^T, \quad (8)$$

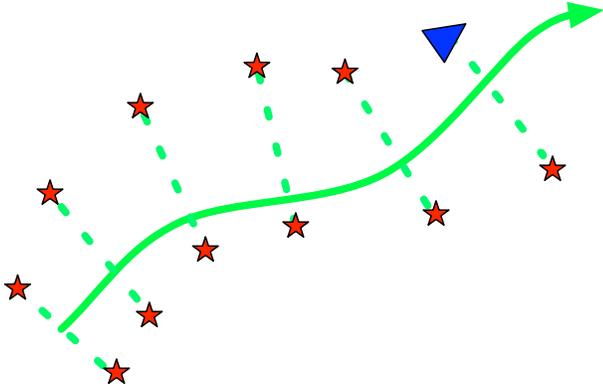


Fig. 3. Physical sampling of the localisation envelope. This figure shows the original teach trajectory (solid green line) and the robot (blue triangle) taking samples of the localisation envelope along the route. At regular intervals the robot deviates (green dashed lines) until it experiences a localisation failure (red stars). This operation can be performed autonomously as the robot can use the repeat machinery to traverse the map, and VO is sufficiently accurate to recover from temporary localisation failures. In this manner a more informed envelope can be discovered by the robot without requiring additional manual teaching.

where l_s , l_d and l_c refer to the length scales for the signed cross track error, the spline distance and local curvature respectively. An additional term encoding the overall relation between points is provided by σ_f . This is used to construct the diagonal matrix M , which encodes the three length scales by

$$M = \begin{bmatrix} l_s^{-2} & 0 & 0 \\ 0 & l_d^{-2} & 0 \\ 0 & 0 & l_c^{-2} \end{bmatrix}. \quad (9)$$

The Matérn covariance function is stationary, meaning that it is only a function of the relative distance between the two inputs, but is unaffected by the global position of the inputs in the feature space. The values of θ are selected by maximising equation 5.

Using the GP machinery, we can use the keyframes in the visual map and any additional sample points to estimate the localisation envelope. An example of this is shown in Figure 2. A teach pass is shown in the top figure (green line) with additional observations of localisation failure (red dots). A sample repeat trajectory is also depicted, where the varying colours indicate the localisation score: blue corresponds to a low value and red to a high score. The lower graphic shows the localisation score against the distance along the route, along with the associated uncertainty. Note that when far from the observations of the localisation envelope the uncertainty is high, but near observations the uncertainty reduces.

The GP also allows us to continually refine our model of the envelope. As more observations are taken, either from physical sampling (discussed in Section IV-B) or straight forward repeat runs of the environment, they can be incorporated to produce more informed estimates about the map.



Fig. 4. Overhead of the teach trajectory (orange line). The urban environment offers a variety of features that affect the resulting localisation envelope, including wide open, narrow and asymmetric spaces.

B. Physical Sampling

While the GP framework is useful for modelling the localisation envelope, its predictions are only as informed as the provided training data. If a teach pass is the only data available, estimates further away from the original trajectory are more unreliable. This is reflected by the GP’s uncertainty. We are also limited by the fact that the actual envelope is likely to be complex. If we are able to gather more observations of the envelope, we can improve our GP model.

We propose physical sampling of the localisation envelope on subsequent passes of the trajectory by intentionally deviating from the route until failure. After the initial teach pass, the robot is allowed to repeat the route using the normal repeat localisation framework. However, at intervals along the route, it takes detours until it encounters a localisation failure, at which point it returns to the path and continues onwards. If done autonomously, the VO should be sufficiently accurate to allow the robot to return to the localisation envelope after a failure. The physically-sampled deviations from the path implicitly capture features of the world, such as occlusions and the degree of viewpoint change. This sampling process can be thought of as a secondary part of the teach phase. Although in this work we opt for regular sample of the teach pass, more complex search strategies could be used. Figure 3 demonstrates our proposed sampling strategy and the output.

While this explicit expenditure of effort to perform sampling can be used to bootstrap the GP, it is also possible to use straight forward repeats of the teach pass as input to the GP.

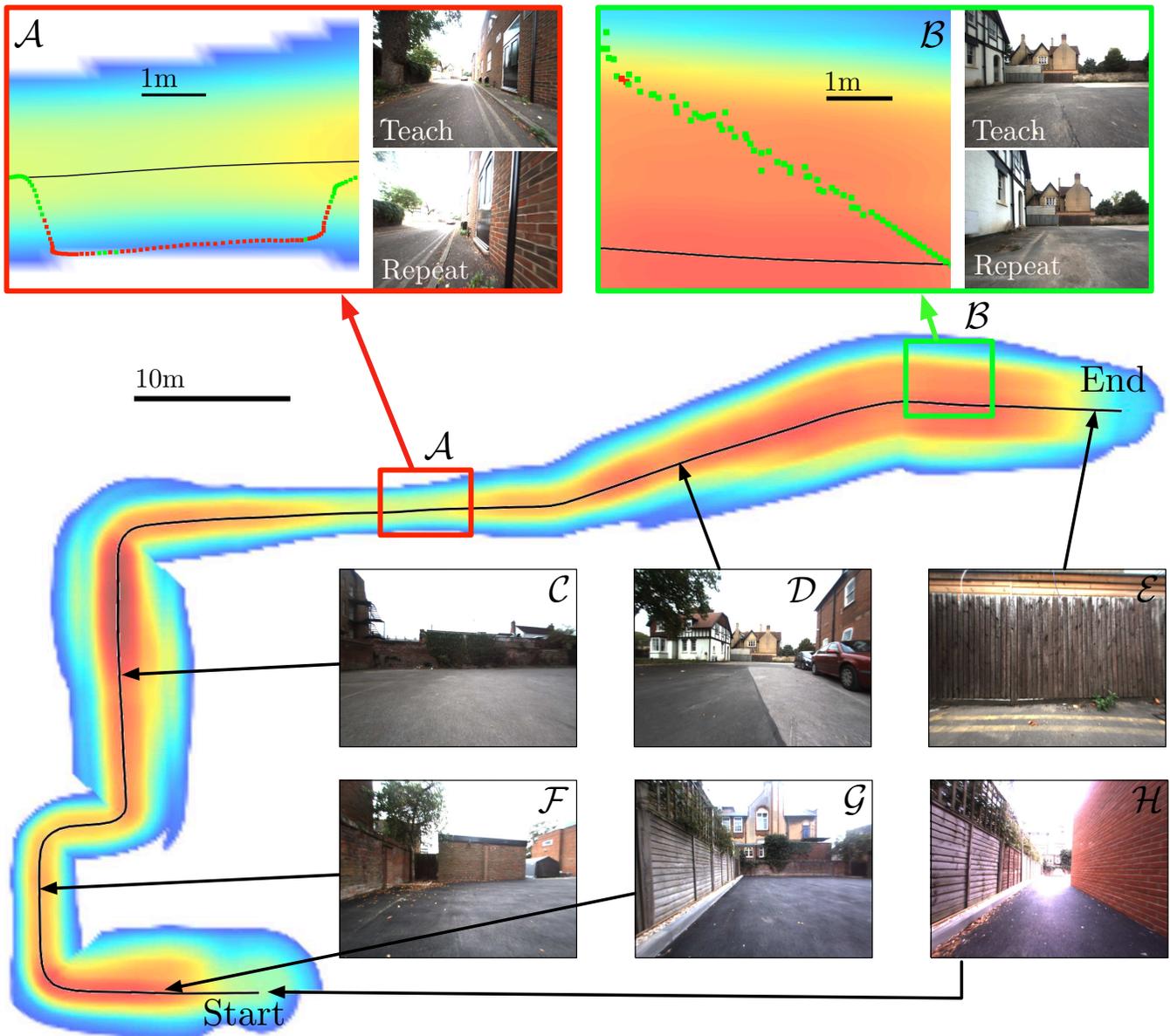


Fig. 5. The teach pass (black line), with the surrounding localisation envelope. Blue indicates low localisation scores, while red marks high scoring areas. Note how the envelope is not constant, or symmetric along the route. Inserts for areas A and B give detailed examples of the envelope, along with one of the six repeat trajectories. The success or failure of the repeat trajectory’s localisations are indicated by green or red dots. In area A , the localiser fails with less than 1m of deviation from the taught route, due to the significant scene change caused by the presence of the wall. Conversely in area B , the localiser is successful several meters away, as the scene appearance does not change significantly despite large deviations. Areas C - H give other examples for how the localisation envelope varies with the teach route. See main text for details.

V. RESULTS

To evaluate our envelope modelling, we collected several datasets from an urban environment in central Oxford. These centre around a 110m teach pass, as shown in Figure 4. The route travels through a variety of structurally different spaces, which affect the shape of the envelope. After teaching the route, a sampling repeat pass was performed in a manner similar to Figure 3. At regular intervals (approximately every 1 to 2 meters on the straights, and at the start, mid-point and exit of every corner) lateral deviations were performed from the teach pass until the localisation failed. This survey dataset served as the training data, \mathcal{T} , for the GP model. Six

subsequent repeat runs were performed, which deviated to varying degrees from the teach pass. The most conservative approximately followed the teach pass, as would be expected in the original Teach and Repeat paradigm. The most divergent intentionally deviated a long way from the path, with the specific aim to get lost. These six repeats served as our test data. All data were collected with a Bumblebee2 stereo camera.

The output of the trained GP can be seen in Figure 5. The teach pass is shown as the black line, and the surrounding colour map is computed by sampling the GP at regular

TABLE I

PREDICTION OF GP FOR LOCALISATION SUCCESS OR FAILURE. THE ACTUAL LOCALISATION RESULTS FOR EACH REPEAT ARE USED AS GROUND TRUTH.

	Frames	Precision	Recall	F ₁
Repeat 1	1111	95.2%	100.0%	97.6%
Repeat 2	1061	99.0%	100.0%	99.5%
Repeat 3	1043	85.6%	96.0%	90.5%
Repeat 4	1040	99.0%	100.0%	99.5%
Repeat 5	1010	93.1%	99.0%	96.0%
Repeat 6	1064	99.3%	100.0%	99.7%
Overall	6329	95.5%	99.3%	97.4%

intervals¹. Red indicates a high localisation score, while blue indicates a low score. The lower right grid of images show examples of the scene at different points along the teach pass, with their location indicated by arrows. In the upper part of the figure, two sections of the route are enhanced, \mathcal{A} and \mathcal{B} . In each, one of the repeat passes is shown, where the poses are indicated by either green dots for successful localisations, or red dots for failures. The position of the red dots are approximated by integrating the VO, which is assumed to be accurate over short distances. This is proven in area \mathcal{A} , where the integrated VO position and localisation from the rejoined path agree. To the right of each insert, two images are shown, with one from the teach run and one from the repeat.

The key point of this figure is to show how the localisation envelope varies with the teach route. In area \mathcal{A} , the localisation envelope is relatively small with less than 1m in cross track error resulting in failure. Compare this to area \mathcal{B} , where the system can still localise several meters away from the teach pass. Looking at the images to the right of each insert, it is apparent that this is caused by how the local scene appearance changes as a function of position. In area \mathcal{A} , modest translations result in large changes in appearance due to the proximity to the wall. Conversely, in area \mathcal{B} , large deviations result in modest appearance changes.

Following this point, it is also interesting to consider the sample images along the taught route. Areas \mathcal{C} and \mathcal{D} are relatively open, so the envelope around those parts of the route is wide. Area \mathcal{H} at the start of the route has a very narrow corridor which results in an initially uncertain GP. By place \mathcal{G} , the space has opened out, resulting in a wider basin. Also note here the asymmetric nature of the envelope, caused by the presence of the fence on one side, and open space on the other. Curiously in space \mathcal{F} , the envelope is also relatively narrow. This is likely caused by the presence of the wall to the left, which results in significant viewpoint change when deviating in either direction. Finally, towards the end of the route in area \mathcal{E} , we see the envelope narrow as we approach a wall. This makes sense, as the closer we get to the wall, the smaller the deviation required to cause the imaged scene to change, and therefore localisation to fail.

¹The rough edges stem from the sampling frequency. In this image we sampled every 0.2m.

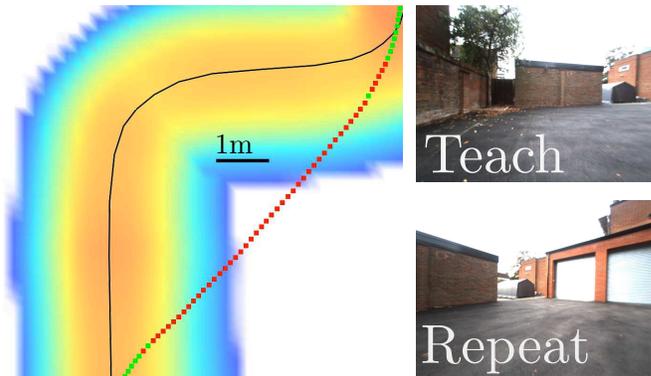


Fig. 6. A subsection near area \mathcal{F} (see Figure 5) is shown on the left, along with repeat 3. During this run, the corner was intentionally cut, resulting in a camera direction significantly different to that of the taught route. Example images from the teach and the repeat are shown on the right. This change in camera angle causes localisation failures, even though the $[x_1, x_2]$ position is inside the envelope.

To assess the performance of the GP, we treat the six repeats as test cases (remember that the GP is trained with the single sampling run). Each of the six test repeats are compared against the taught map, and the localisation results recorded. The poses along the repeats are used as input to the GP to generate localisation score predictions, y_* . However, assessing the localisation quality, i.e. how displaced the estimate is from ground truth, is difficult without an accurate external ground truth source. Since we could not rely on GPS, as its solution quality suffers in canyon-like environments, under overhanging trees and it can drift over time [22], we thresholded the scores and treated the predictions as binary classifications: did the GP predict a successful localisation or failure. The choice of binary classification is due to the fact that our localisation system does not degrade gracefully. The limitation of the binary signal may be overcome by utilising the uncertainty provided by the GP for this value. We compare this to what happened when the repeat ran in the Teach and Repeat system, which we take as ground truth. We tabulate the results for each run individually, and overall, in Table I. Frames refers to the number of stereo frames in each dataset, and the Precision, Recall and F₁ scores are computed by thresholding y at a value of 0.4. This value was selected by plotting a precision-recall curve and selecting a suitable operating point.

The choice of the localisation score, y , is an important one, as it must give an indication of the localiser's expected performance. Here we have chosen the normalised number of matched keyframe landmarks, but the localiser offers several other outputs (e.g. RMS error). While our current choice of y is a good predictor of being able to localise (Table I), it is not perfect. In areas \mathcal{A} and \mathcal{B} (Figure 5) there are repeat poses which succeed despite the GP predicting they would fail, and visa versa. This may be why the repeat runs in Table I have a higher recall compared to precision. One area to investigate further is the choice of y .

Repeat 3 was one of the more divergent repetitions of the taught route, where we intentionally cut two of the three

major corners. Part of the trajectory can be seen on the left side of Figure 6. When cutting the corner, the heading error relative to the teach pass becomes significant. This difference in angle results in a different section of the scene being imaged, as shown by the images on the right side of the figure. This results in localisation failures, even though the $[x_1, x_2]$ position is initially inside the predicted localisation envelope. This intentional divergence from the model is why repeat 3 performs poorly compared to the other repeats in Table I.

VI. CONCLUSIONS

In this paper we have presented a way to capture the localisation envelope of a visual teach pass in the framework of Teach and Repeat. The modelling is performed using a Gaussian Process, which takes as input samples of the localiser's performance around the taught route. These can come from explicit sampling, or simply subsequent repeat passes. This place-dependent spatial model of localiser performance provides crucial additional information for when the map is subsequently used for planning and autonomy. We have demonstrated our approach in an urban environment, and validated the model's predictions against a number of repeat runs. Future extensions are to include rotation in the model, investigate the influence of different feature descriptors on the envelope, test in different environments and incorporate scene structure into the feature vector.

While we have applied the idea of localisation envelope to our stereo vision based system, it would be equally applicable to other approaches, such as monocular vision, topological localisers and laser-based systems. By continually improving our understanding of the localisation envelope, we can know the limits of our system.

VII. ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of this work by the EPSRC Leadership Fellowship, grant EP/J012017/1, the European Community's Seventh Framework Programme under grant agreement no FP7-610603 (EUROPA2) and the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/J012017/1.

REFERENCES

- [1] P. Furgale and T. D. Barfoot, "Visual teach and repeat for long-range rover autonomy," *Journal of Field Robotics*, vol. 27, no. 5, pp. 534–560, 2010.
- [2] B. E. Stenning, C. McManus, and T. Barfoot, "Planning using a network of reusable paths: A physical embodiment of a rapidly exploring random tree," *Journal of Field Robotics*, vol. 30, no. 6, pp. 916–950, 2013. [Online]. Available: <http://dx.doi.org/10.1002/rob.21474>
- [3] S. M. LaValle and J. J. Kuffner, "Randomized kinodynamic planning," *The International Journal of Robotics Research*, vol. 20, no. 5, pp. 378–400, 2001.
- [4] P. Krsti, B. Bcheler, F. Pomerleau, U. Schwesinger, R. Siegwart, and P. Furgale, "Lighting-invariant adaptive route following using iterative closest point matching," *Journal of Field Robotics*, pp. n/a–n/a, 2014. [Online]. Available: <http://dx.doi.org/10.1002/rob.21524>
- [5] N. Melchior and R. Simmons, "Particle rrt for path planning with uncertainty," in *IEEE International Conference on Robotics and Automation*, April 2007, pp. 1617–1624.
- [6] N. Roy and S. Prentice, "The belief roadmap: Efficient planning in belief space by factoring the covariance," *The International Journal of Robotics Research*, 2009.
- [7] L. Kavraki, P. Svestka, J.-C. Latombe, and M. Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration spaces," *Robotics and Automation, IEEE Transactions on*, vol. 12, no. 4, pp. 566–580, Aug 1996.
- [8] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1/2, pp. 43–72, 2005.
- [9] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, Oct 2005.
- [10] H. Moravec, "Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover," in *Tech Report CMU-RI-TR-80-03, Robotics Institute, Carnegie Mellon University & Doctoral Dissertation, Stanford University*, September 1980, no. CMU-RI-TR-80-03.
- [11] D. Nistér, O. Naroditsky, and J. Bergen, "Visual Odometry for Ground Vehicle Applications," *Journal of Field Robotics*, vol. 23, 2006.
- [12] M. Maimone, Y. Cheng, and L. Matthies, "Two years of Visual Odometry on the Mars Exploration Rovers: Field Reports," *Journal of Field Robotics*, vol. 24, no. 3, pp. 169 – 186, March 2007. [Online]. Available: <http://dx.doi.org/10.1002/rob.v24:3>
- [13] D. Scaramuzza and F. Fraundorfer, "Visual Odometry: Part I - The First 30 Years and Fundamentals," *IEEE Robotics Automation Magazine*, vol. 18, no. 4, pp. 80 –92, December 2011.
- [14] G. Sibley, C. Mei, P. Newman, and I. Reid, "A System for Large-Scale Mapping in Constant-Time using Stereo," *International Journal of Robotics Research*, 2010.
- [15] T. Barfoot, "Online Visual Motion Estimation using FastSLAM with SIFT Features," *International Conference on Intelligent Robots and Systems*, 2005.
- [16] E. Rosten, G. Reitmayr, and T. Drummond, "Real-time Video Annotations for Augmented Reality," in *Advances in Visual Computing, LNCS 3840*, December 2005, pp. 294–302.
- [17] M. Calonder, V. Lepetit, M. Ozuysal, T. Trzinski, C. Strecha, and P. Fua, "BRIEF: Computing a Local Binary Descriptor Very Fast," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011.
- [18] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," in *International Journal of Computer Vision*, 2004.
- [19] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Communications of the ACM*, vol. 24, p. 381395, 1981.
- [20] W. Churchill and P. Newman, "Experience-based Navigation for Long-term Localisation," *The International Journal of Robotics Research (IJRR)*, 2013.
- [21] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, 2006.
- [22] I. Baldwin and P. Newman, "Road vehicle localization with 2d push-broom lidar and 3d priors," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Minnesota, USA, May 2012.