Leveraging Experience for Large-Scale LIDAR Localisation in Changing Cities

Will Maddern, Geoffrey Pascoe and Paul Newman

***Abstract*— Recent successful approaches to autonomous vehicle localisation and navigation typically involve 3D LIDAR scanners and a static, curated 3D map, both of which are expensive to acquire and maintain. In this paper we propose an experience-based approach to matching a local 3D swathe built using a push-broom 2D LIDAR to a number of prior 3D maps, each of which has been collected during normal driving in different conditions. Local swathes are converted to a combined 2D height and reflectance representation, and we exploit the GPU rendering pipeline to densely sample the localisation cost function to provide robustness and a wide basin of convergence. Prior maps are incrementally built into an experience-based framework from multiple traversals of the same environment, capturing changes in environment structure and appearance over time. The LIDAR localisation solutions from each prior map are fused with vehicle odometry in a probabilistic framework to provide a single pose solution suitable for automated driving. Using this framework we demonstrate real-time centimetre-level localisation using LIDAR data collected in a dynamic city environment over a period of a year.**

I. Introduction

Autonomous consumer vehicles have progressed from a distant goal a decade ago at the first DARPA Grand Challenge [1] to a common sight in some parts of the world, with Google and other research groups demonstrating fully autonomous prototypes that have driven hundreds of thousands of miles [2], [3], [4]. The most successful of these prototypes combine 3D LIDAR scanners with high-accuracy GPS+INS systems to localise with centimetre-precision in curated, globally consistent 3D prior maps [5].

A major obstacle preventing widespread deployment of autonomous road vehicles is the prohibitive cost of the sensor systems. At over £160,000 at the time of writing, a typical autonomous vehicle sensor suite consisting of a Velodyne 3D LIDAR scanner [6] and an Applanix POS-LV GPS+INS system [7] remains far beyond the price range of sensors for consumer vehicles. This high cost has lead researchers to investigate cheaper sensor alternatives such as cameras; however, despite recent advances in exploiting consumer cameras for autonomous vehicle localisation [8], [9], [10], [11], many challenges remain before a truly reliable vision-based localisation system is available for widespread use.

A secondary cost not often considered in the domain of automated driving is the costs involved in building and maintaining maps suitable for autonomous vehicle operation. While Google Street View [12] serves as proof that it is possible to produce 3D street-level maps on a truly global scale with dedicated survey vehicles, these surveys only capture the structure and appearance of the environment once every few years. Surveying environments of this scale at a rate required to capture scene variation relevant for autonomous driving (such as changes in road surface, buildings, parked vehicles, local vegetation and seasonal variation) would be prohibitively expensive for a map provider, and any economic benefits of selling automated driving systems may be offset by the constant need to re-survey the environment that these autonomous vehicles operate in.

Instead, we propose that an "experience-based" approach

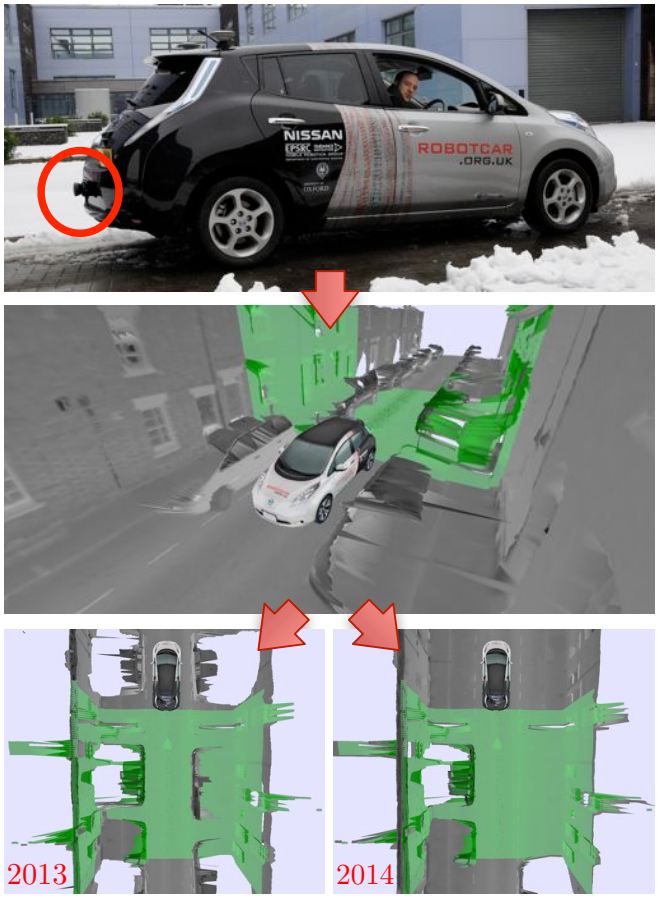


Fig. 1. Experience-based 2D push-broom LIDAR localisation in changing city environments. A single 2D LIDAR is mounted to the rear bumper of the car (top). Using odometry, a local 3D swathe (green) can be produced (middle), which can then be simultaneously matched to multiple prior experiences (bottom) collected over the lifetime of the vehicle. New experiences are continuously collected in areas of high uncertainty, such as in the presence of dynamic objects or changing structure, maintaining localisation performance in complex city environments over time-scales of multiple years.

Authors are from the Mobile Robotics Group, Dept. Engineering Science, University of Oxford, UK. {wm,gmp,pnewman}@robots.ox.ac.uk

[13], as shown in Fig. 1, is a natural method of dealing with environmental change over time in the domain of autonomous vehicles. Rather than defining a single static global map that all vehicles must use (and that must be updated frequently), each vehicle independently builds and maintains a set of "experiences", capturing representations specific to its operating environment. As the vehicle traverses a route a number of times, it will intrinsically capture richer representations of more dynamic locations as required - since map creation is driven by the task-oriented metric of localisation performance, it will record only the information necessary to maintain consistent accuracy across all locations. Once sufficient experiences are accumulated, the autonomous vehicle may be sufficiently confident to, for example, offer to take over control on certain sections of a frequently-traversed route.

In this paper we present a probabilistic experience-based approach to localisation with 2D push-broom LIDAR sensors. We adapt the approach of [14] and build locally metric 3D "swathes" using odometry information, eliminating our requirement for a 3D LIDAR. We use low-cost GPS observations as a "weak localiser" to find the most relevant experiences for the current swathe, then localise the swathe within the experience with a robust sample-based method. If the current swathe is not matched to any experience with sufficient accuracy, a new experience is created to capture more detail of the environment in difficult localisation conditions. We present a GPU implementation that can provide real-time localisation at 5Hz and pose estimates at 40Hz, and demonstrate large-scale localisation using over 50km of data collected in a dynamic city environment over a period of a year.

II. Related Work

LIDAR-based localisation for autonomous vehicle applications has been addressed by several authors in the past decade. Early approaches using a traditional horizontal planar 2D LIDAR configuration [15], [16] produced impressive localisation performance at medium scales, but suffered from out-of-plane effects such as ground-strike. Later approaches [17] actuated the 2D laser to produce local 3D pointclouds for localisation, to avoid the limitations of a fixed planar LIDAR configuration.

The availability of 3D LIDAR sensors for the DARPA Urban Challenge [18] enabled larger-scale robust localisation approaches. In [19], a Velodyne LIDAR was used to build a 2D orthographic reflectivity map of road surfaces, and centimeter-accurate localisation was achieved using a particle filter. This approach was extended in [20] to incorporate reflectivity variance across multiple scans in a probabilistic framework. A similar approach presented in [21] used a 2D height-map representation for navigating in a multi-level parking garage.

To reduce the sensing cost of a LIDAR-based localisation system for autonomous vehicles, [14] presented an approach based on a 2D LIDAR in a push-broom configuration. By integrating vehicle odometry over a short window, multiple 2D laser scans were combined to form a 3D representation of path the vehicle has recently traversed. This local 3D map was compared to a global 3D map using a grid-based histogram approach, yielding centimetre-level accuracy and increased robustness in comparison to a GPS+INS system. This approach was extended in [22] to use 2D LIDARs for both vehicle odometry and localisation relative to a prior map.

The methods mentioned so far all localise relative to a single static global map, either incrementally constructed using a SLAM framework or offline using an optimisation approach. However, environments for autonomous vehicles contain highly dynamic objects (vehicles, pedestrians) as well as features that change gradually over time (parked vehicles, construction, seasonal changes), and a static global map will fail to capture this variation. Rather than attempting to combine multiple distinct representations of the same environment into a single static map, the experience-based framework of [13], [23] simply stores these representations, dubbed "experiences", and attempts to localise relative to multiple experiences simultaneously. This approached produced impressive localisation performance over a 3-month period at different times of day and in different weather conditions, however still suffers from the limitations of stereo-camera-based approaches, namely strong dependence on scene illumination and a narrow convergence basin [24].

III. Map and Swathe Construction

In this section we present our approach for local 3D pointcloud construction from a 2D LIDAR sensor for the purposes of localisation against a prior 3D map. We define a 2D LIDAR scan $\mathbf{s}(t)$ at time t as follows:

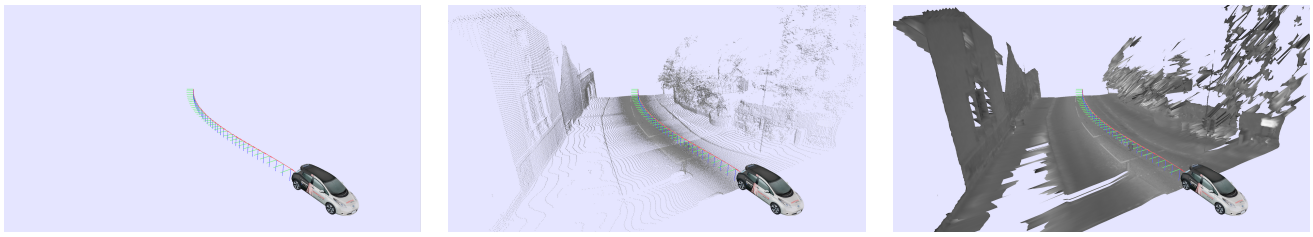
$$\mathbf{s}(t) = \{d_1 \dots d_m, r_1 \dots r_m\} \tag{1}$$

where d_i is the laser distance measurement (in metres) for beam i, and r_i is the corresponding reflectance measurement in the infrared spectrum. $\mathbf{S} = \{\mathbf{s}(t_0) \dots \mathbf{s}(t)\}$ describes a collection of such scans over the time period $[t_0, t]$. To compose the 2D LIDAR scans $\mathbf{S}$ into a local 3D representation, we must estimate the trajectory of the vehicle over a short time period as follows:

$$\dot{\hat{\mathbf{x}}}(t) = \hat{v}(t) \begin{bmatrix} \cos\left(\int_{t_0}^{t} \hat{w}_z(t)\, dt\right) \\ \sin\left(\int_{t_0}^{t} \hat{w}_z(t)\, dt\right) \\ \sin\left(\int_{t_0}^{t} \hat{w}_x(t)\, dt\right) \end{bmatrix} \tag{2}$$

where $\hat{v}(t), \hat{w}(t)$ are the translational and rotational velocities of the vehicle estimated using inertial sensors, wheel odometry or visual odometry. By integrating $\dot{\hat{\mathbf{x}}}(t)$ over the period $[t_0, t]$, we produce the estimated continuous $\mathbb{SE}(3)$ pose $\hat{\mathbf{x}}(t)$. We then define the "swathe" $\mathbf{Q}$ as the local 3D pointcloud produced by projecting 2D scans $\mathbf{S}$ along the continuous-time trajectory $\hat{\mathbf{x}}(t)$ as follows:

$$\mathbf{Q}_{t_0}^{t} = g(\hat{v}, \hat{w}, \mathbf{S})\,\big|_{t_0}^{t} \tag{3}$$



(a) (b) (c)

Fig. 2. Local 3D swathe construction from 2D laser scans and vehicle trajectory. (a) By integrating vehicle velocities over a short period, a local $\mathbb{SE}(3)$ estimate of the vehicle trajectory can be made over a short window. (b) Recent push-broom LIDAR scans are then projected at locations along the trajectory, forming a locally consistent 3D point cloud of the environment. (c) For matching on the GPU, successive push-broom LIDAR scans are stitched together forming a local 3D mesh representation.

where the operator $g(.)$ defines the pointcloud projection function along the trajectory. Fig. 2 illustrates the process of building the 3D pointcloud from 2D scans along the vehicle trajectory $\hat{\mathbf{x}}(t)$.

For localisation, we require a prior 3D pointcloud against which we can compare the swathe $\mathbf{Q}$. In contrast to other methods [20], [11] which construct a globally consistent 3D map of the environment, we only require locally consistent maps to compare the swathe against. Hence, we can apply the same process as above to construct swathes using historical data. We denote each of these historical swathes as an individual map $\mathbf{m}^i = \langle \mathbf{Q}^i, \mathbf{z}^i \rangle$, where $\mathbf{z}^i$ is the GPS observations recorded along the trajectory used to construct $\mathbf{Q}^i$. The collection of all maps $\mathbf{M}$ is as follows:

$$\mathbf{M} = \{\mathbf{m}^1 \dots \mathbf{m}^n\} \tag{4}$$

By storing multiple maps of the same environment gathered in successive traversals, we can simultaneously localise against multiple "experiences" of the environment under different conditions, increasing robustness at the cost of additional storage space. Although each map is not globally consistent, swathes built from short time periods $[t_0, t]$ are sufficient for relative localisation against prior trajectories, which enables teach-and-repeat [24] and relative-frame planning approaches.

IV. Localisation

Given a swathe $\mathbf{Q}_k$ built at the discrete update interval k and a set of prior maps $\mathbf{M}$, we can frame the process of estimating the pose of the vehicle $\hat{\mathbf{x}}_k$ as a maximum-a-posteriori estimation problem as follows:

$$\hat{\mathbf{x}}_k = \arg\max_{\mathbf{x}_k} p\left(\mathbf{x}_k \left| \mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{z}_k, \mathbf{M}, \mathbf{Q}_k \right.\right) \tag{5}$$

where $\mathbf{x}_{k-1}$ is the pose of the vehicle at the previous interval $k-1$, $\mathbf{u}_k$ is the motion experienced between update intervals and $\mathbf{z}_k$ is the GPS observation (if available) at interval k. Applying Bayes' rule yields the following relationship after removing unnecessary conditionals:

$$p\left(\mathbf{x}_k \left| \mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{z}_k, \mathbf{M}, \mathbf{Q}_k \right.\right) \propto p\left(\mathbf{x}_k \left| \mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{z}_k \right.\right) \prod_i p\left(\mathbf{Q}_k \left| \mathbf{x}_k^i, \mathbf{m}^i \right.\right). \tag{6}$$

The first term $p\left(\mathbf{x}_k \left| \mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{z}_k \right.\right)$ is the *location prior*, which models the likelihood of a pose using incremental odometry and GPS observations. The second term $p\left(\mathbf{Q}_k \left| \mathbf{x}_k^i, \mathbf{m}^i \right.\right)$ is the *observation likelihood*, which models the likelihood of producing the observed swathe $\mathbf{Q}_k$ given a hypothesised pose $\mathbf{x}_k^i$ within the prior map $\mathbf{m}^i$. The product of the observation likelihood term for all i permits the simultaneous comparison to multiple prior maps $\mathbf{m}^i$. The two terms will be examined in detail below.

A. Location Prior

The location prior term from Equation 6 can be further expanded as follows:

$$p\left(\mathbf{x}_k \left| \mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{z}_k \right.\right) \propto p\left(\mathbf{x}_k \left| \mathbf{x}_{k-1}, \mathbf{u}_k \right.\right) \prod_i p\left(\mathbf{z}_k \left| \mathbf{x}_{k-1}^i \right.\right) \tag{7}$$

where $p\left(\mathbf{x}_k \left| \mathbf{x}_{k-1}, \mathbf{u}_k \right.\right)$ is the motion model of the vehicle and $p\left(\mathbf{z}_k \left| \mathbf{x}_{k-1}^i \right.\right)$ is the GPS observation likelihood for each previous pose $\mathbf{x}_{k-1}^i$ in the local frame of map $\mathbf{m}^i$ (which will be discussed further in the following section). The odometry update $\mathbf{u}_k$ can be computed by integrating the continuous-time vehicle velocity from Equation 2 as follows:

$$\mathbf{u}_k = \int_{t_{k-1}}^{t_k} \dot{\hat{\mathbf{x}}}(t)\, dt \tag{8}$$

where t_{k-1} and t_k are the timestamps at update intervals $k-1$ and k respectively.

B. Weak GPS Localisation

The GPS observation likelihood term in Equation 7 is not intended to provide centimetre-accurate location priors; rather, it is used as a *weak localiser*, providing information on which maps $\mathbf{m}^i$ will be relevant for the current swathe $\mathbf{Q}_k$. By making use of occasional, inaccurate GPS observations when they are available, we sidestep the need for loop closure algorithms to deal with global initialisation and the "kidnapped robot" problem, but equally we do not rely on GPS for accurate location estimates for path planning and vehicle behaviour.

As our maps are only locally metric, both localisation accuracy and map quality decrease as the distance between the current pose estimate in the map $\mathbf{x}_k^i$ and the location

of the GPS observation $\mathbf{z}_k$ increases. To represent this, we approximate the GPS observation as a Gaussian with covariance matrix $\mathbf{\Sigma_z}$, then further inflate the covariance based on the measurement as follows:

$$\bar{\mathbf{\Sigma}}^i_{\mathbf{z}_k} = \mathbf{\Sigma_z} + \left[\mathbf{z}_k - h\left(\mathbf{x}^i_{k-1}, \mathbf{m}^i\right)\right]\left[\mathbf{z}_k - h\left(\mathbf{x}^i_{k-1}, \mathbf{m}^i\right)\right]^{\mathrm{T}} \tag{9}$$

where $\bar{\mathbf{\Sigma}}^i_{\mathbf{z}_k}$ is the new covariance of the GPS location estimate for map $\mathbf{m}^i$ and the function $h\left(\mathbf{x}^i_{k-1}, \mathbf{m}^i\right)$ produces the expected GPS observation at location $\mathbf{x}^i_{k-1}$ in map $\mathbf{m}^i$. Therefore, only maps $\mathbf{m}^i$ containing prior GPS observations $\mathbf{z}^i$ close to the current GPS observation $\mathbf{z}_k$ will yield acceptably low location prior uncertainties, making them relevant for swathe-based localisation as discussed below.

C. Swathe Localisation

While the location prior and weak GPS localisation alone may yield an acceptable localisation solution for some applications, true centimetre-level accuracy can be obtained by making use of 3D map and swathe information. However, existing methods for comparing 3D pointclouds are often computationally expensive [25], highly susceptible to initialisation noise and data association errors [26], [27], or a combination of the two. Instead, we follow the approach of [19], [21], [14] and project our 3D pointcloud onto a 2D plane, yielding a 2.5D height and reflectance representation.

The x-y plane on which to perform the 2D projection can be extracted in a number of ways: by fitting a ground plane to laser data, transforming the pointcloud relative to the local gravity vector using an inertial measurement unit or simply assuming the vehicle is locally horizontal. More important is the consistency of the method, such that the same ground plane is used each time the vehicle revisits a location.

We can factor the observation likelihood term from Equation 6 by sampling from the swathe $\mathbf{Q}_k$ as follows:

$$p\left(\mathbf{Q}_k \left| \mathbf{x}^i_k, \mathbf{m}^i\right.\right) = \prod_j p\left(\mathbf{q}_j \left| \mathbf{x}^i_k, \mathbf{m}^i\right.\right) \tag{10}$$

where each sample $\mathbf{q}_j = \langle h_j, r_j \rangle$ and h_j, r_j are the height and reflectance sampled from location j. The sample likelihood term $p\left(\mathbf{q}_j \left| \mathbf{x}^i_k, \mathbf{m}^i\right.\right)$ can then be directly computed in log-likelihood form as follows:

$$\begin{aligned} -\log\left[p\left(\mathbf{q}_j \left| \mathbf{x}^i_k, \mathbf{m}^i\right.\right)\right] \propto \\ \left[\frac{h_j - H_j\left(\mathbf{x}^i_k, \mathbf{m}^i\right)}{\sigma_{\mathbf{h}}}\right]^2 + \left[\frac{r_j - R_j\left(\mathbf{x}^i_k, \mathbf{m}^i\right)}{\sigma_{\mathbf{r}}}\right]^2 \end{aligned} \tag{11}$$

where the functions $H_j(.)$ and $R_j(.)$ produce the expected height and reflectance values at sample location j in the map, and $\sigma_{\mathbf{h}}, \sigma_{\mathbf{r}}$ are the standard deviations of height and reflectance respectively, based on intrinsic sensor noise from the LIDAR scanner. Note that in contrast to methods that use either reflectance only [19], [20] or height only [21], [14], we combine both sources of information into a single cost function to improve robustness.

D. Localisation Covariance

While the localisation problem in Equation 5 could be solved using an optimisation-based approach to yield a maximum likelihood estimate for the vehicle pose $\hat{\mathbf{x}}_k$, for autonomous driving applications it is often desirable to also know the uncertainty of the localisation estimate in order to influence vehicle behaviour (e.g. reducing speed in areas of high uncertainty, or reverting control to the driver of the vehicle). Inspired by [28], we further sample the observation likelihood term $p\left(\mathbf{Q}_k \left| \mathbf{x}^i_k, \mathbf{m}^i\right.\right)$ at a series of poses $\mathbf{x}^i_{(j)}$ to produce a mean offset pose $\bar{\mathbf{x}}^i_k$ and uncertainty $\bar{\mathbf{\Sigma}}^i_k$ as follows:

$$\begin{aligned} K &= \sum_j \mathbf{x}^i_{(j)} \mathbf{x}^{i\mathrm{T}}_{(j)} p\left(\mathbf{Q}_k \left| \mathbf{x}^i_{(j)}, \mathbf{m}^i\right.\right) \\ u &= \sum_j \mathbf{x}^i_{(j)} p\left(\mathbf{Q}_k \left| \mathbf{x}^i_{(j)}, \mathbf{m}^i\right.\right) \\ s &= \sum_j p\left(\mathbf{Q}_k \left| \mathbf{x}^i_{(j)}, \mathbf{m}^i\right.\right) \\ \bar{\mathbf{x}}^i_k &= \frac{1}{s}u, \; \bar{\mathbf{\Sigma}}^i_k = \frac{1}{s}K - \frac{1}{s^2}uu^{\mathrm{T}} \end{aligned} \tag{12}$$

The resulting covariance matrix $\bar{\Sigma}^i_k$ captures both the intrinsic noise of the sensor and swathe representation as well as the uncertainty in data association between swathe and map. This serves to provide a realistic estimate of the localisation uncertainty of the vehicle and a wide basin of convergence, but at the cost of increased computational complexity during sampling. The computational cost is addressed in Section V.

E. Map Update

Although the use of multiple maps allows us to robustly localise against multiple representations of the same environment under different conditions, there will come occasions when no map is sufficiently similar to the current swathe to provide an acceptable localisation estimate. Indeed, this will occur whenever the vehicle traverses a previously unmapped location. Rather than attempting to generalise from existing maps, we simply add the current swathe and GPS observations to the set of maps $\mathbf{M}$ according to the following condition:

$$\det\left(\hat{\mathbf{\Sigma}}_k\right) > \rho_{\max} \; : \; \mathbf{m}^{n+1} \mapsto \langle \mathbf{Q}_k, \mathbf{z}_k \rangle \tag{13}$$

where $\hat{\mathbf{\Sigma}}_k$ is the current localisation uncertainty and $\rho_{\max}$ is a threshold for the maximum acceptable uncertainty. By dynamically adding live sensor data to the map only when localisation performance is insufficient, we are performing "experience-based" navigation [13]; this ensures we capture sufficiently rich representations of dynamic environments without map storage requirements that are linear with operation time.

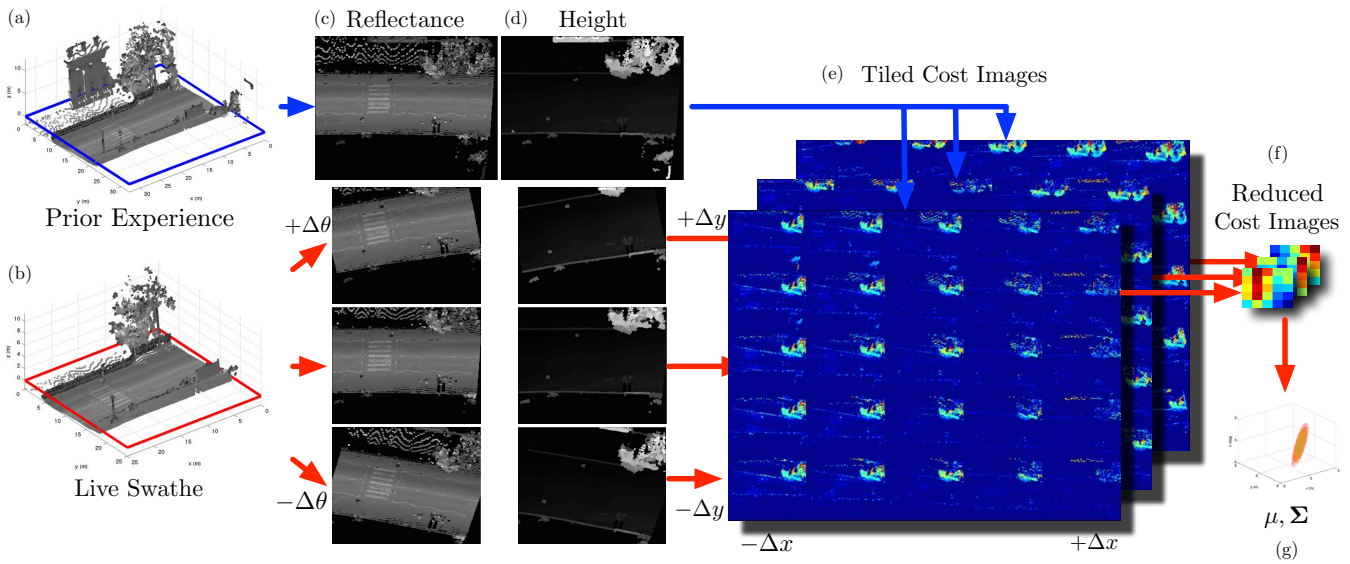


Fig. 3. GPU-accelerated swathe-based localisation process. The prior 3D map (a) and local 3D swathe (b) are both converted to a mesh representation and copied to GPU memory where they are both rasterised to 2D reflectance (c) and height (d) images using a top-down orthographic projection. The prior map is rasterised at slightly higher resolution, and the swathe is rasterised at a number of orientation offsets $\theta_1 \dots \theta_n$. The height at each pixel of the reflectance images is extracted from the z-buffer at no additional rendering cost. A tiled cost image (e) is formed by computing the squared difference between reflectances and heights of the map and swathe at a series of additional translational offsets $(x_1, y_1) \dots (x_n, y_n)$. The tiled cost images are reduced to low-resolution cost images (f) using a *mipmap reduction*, where each pixel in the cost image equals the average cost of the entire tile at a particular x, y, θ offset. Finally, the cost images are stacked to form a cost volume, and a mean and covariance is fit to the volume (g), forming the swathe localisation estimate.

V. GPU Implementation

While the sampling methods presented in Equations 10 and 12 produce high-quality likelihood and uncertainty estimates, they require a significant computational cost to evaluate. However, since the samples are independent, it is possible to exploit modern GPU processors to simultaneously evaluate a large number of observation likelihoods in parallel. In particular we exploit the *mipmap reduction* pipeline [29] to efficiently reduce large tiled cost images with as many as 2M pixels to low-resolution reduced cost images (typically 16 x 16 pixels) in under a millisecond, where each pixel represents the cost of a specific (x, y, θ) offset. The following steps outline the process of efficiently computing $\bar{\mathbf{x}}_k^i$ and $\bar{\Sigma}_k^i$:

1) Compute 3σ bounds for x, y, θ from $\bar{\Sigma}_{k-1}^i$.
2) Convert the sparse map and swathe pointcloud to a dense mesh representation.
3) Project the swathe mesh at a series of orientations $\theta_1 \dots \theta_n$ covering the 3σ bound.
4) Rasterise the map and swathe to form 2D orthographic height/reflectance images.
5) Compute the likelihood $p\left(\mathbf{q}_j \left| \mathbf{x}_k^i, \mathbf{m}^i \right.\right)$ for each pixel in the height/reflectance images for a series of index offsets $(x_1, y_1) \dots (x_n, y_n)$ covering the 3σ bound, forming a tiled cost image.
6) Reduce the tiled cost image to a stacked reduced cost image using an efficient *mipmap reduction*.
7) Fit a mean and covariance to the tiled cost image as per Equation 12 to yield $\bar{\mathbf{x}}_k^i$ and $\bar{\Sigma}_k^i$.

This process is illustrated in Fig. 3. For the above process, implemented in OpenGL and running on a 2012 MacBook Pro with an Nvidia GT650M GPU, swathe localisation can be performed at 5Hz for each map in parallel, providing a corrected odometry signal at 40Hz.

VI. Experimental Setup

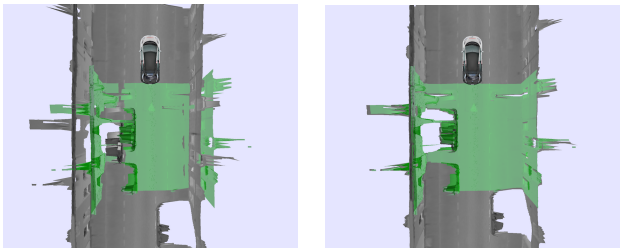
In this section we present our experimental approach to demonstrating long-term localisation with an experience-based 2D push-broom LIDAR approach.

A. Experimental Data

Our experimental platform is the Oxford University Robotcar, an autonomous Nissan LEAF, depicted in Fig. 1. The LEAF is equipped with a SICK LMS-151 laser scanner in push-broom configuration mounted unobtrusively on the rear bumper. Vehicle odometry is provided by shaft encoders on each wheel, providing velocity estimates at 40Hz. Low-cost GPS observations were simulated by querying the NovAtel SPAN-CPT GPS+INS system for the raw, unfiltered GPS-only position estimates at 1Hz.

The experimental dataset consists of six traversals of a an approximately 8km route through central Oxford collected between July 2013 and August 2014, illustrated in the map section of Fig. 7. The traversals were made on public roads at different times of day (from 5am to 7pm) and in different traffic conditions. GPS signals were significantly degraded in areas of tree cover and narrow streets. There is some variation in the route over time, as different routes became available or were closed due to construction work.

Ground truth position estimates for experiments of this duration is challenging even for a tightly-coupled GPS+INS system, as noted in [20]. Instead we make use of a stereo camera mounted on the vehicle, and perfom an offline multi-session pose graph optimisation combining visual odometry



(a) GPS+INS (b) Ground truth

Fig. 4. Local 3D swathe (green) projected on to a map using (a) the location from the GPS+INS system and (b) the metric ground truth derived from a globally optimised pose graph based on hand-corrected stereo visual odometry. Over a period of days or weeks, the GPS+INS position estimate can drift by several metres, rendering it unsuitable for both online autonomous driving and an accurate ground truth for offline localisation benchmarking.

TABLE I

LOCALISATION ALGORITHM PARAMETERS

Symbol	Parameter	Value
$t - t_0$	Swathe Period	10s
$\mathbf{\Sigma_z}$	GPS Covariance	5m in x, y
$\sigma_\mathbf{h}$	Height Std Dev	0.5m
$\sigma_\mathbf{r}$	Reflectance Std Dev	100
n	Number of x, y, θ offsets	16
$\rho_{\max}$	Max Uncertainty Threshold	0.1

and loop closures provided by FAB-MAP [30] to form a globally consistent ground-truth metric map. While this produces superior results to the GPS+INS system as illustrated in Fig. 4, it is an expensive offline process requiring manual verification of many loop closures and is thus only suitable for generating ground-truth poses for benchmarking an online localisation algorithm.

B. Localisation Algorithm Details

For localisation, we use the push-broom SICK LMS-151 LIDAR in combination with wheel odometry (to estimate vehicle velocities $\hat{v}(t), \hat{w}(t)$) and GPS observations derived from the SPAN-CPT. Table I lists the parameters used for the localisation algorithms:

VII. RESULTS

In this section we examine the localisation performance and number of experiences recorded while driving in a changing city environment over a period of a year.

A. Localisation Accuracy

Of chief importance for autonomous driving is the accuracy of the localisation estimate. Fig. 5 shows the localisation error distribution relative to metric ground truth for the final traverse, demonstrating low longitudinal (0.38m), lateral (0.07m) and heading (0.43^o) RMS errors. Additionally, the localisation performance of the GPS+INS system (with inertial filter) was evaluated relative to metric ground truth, illustrating that even high-cost inertial navigation systems are subject to gradual drift over long periods of time, rendering them unsuitable for long-term autonomy over multiple years.

For autonomous driving applications it also important that the localisation uncertainty estimate does not underestimate

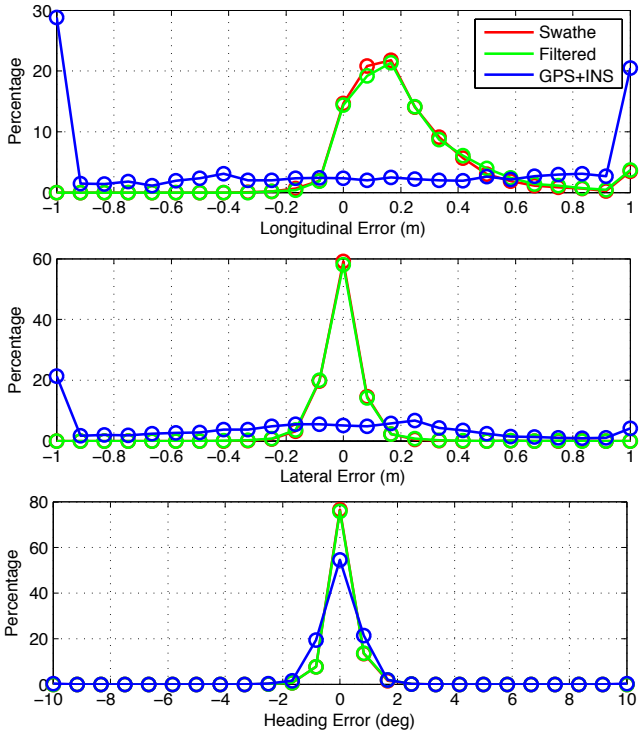


Fig. 5. Longitudinal, lateral and heading error distributions relative to metric ground truth for the localisation system on the final traversal. The GPS+INS system (blue) maintains heading accuracy but is subject to significant translational drift, rendering it useless for autonomous driving. The swathe-only solution (red) is most accurate but only provides updates at 5Hz. The 40Hz filtered solution (green) maintains accuracy while providing high frequency updates. Due to the structure of the road the longitudinal errors (0.38m RMS) are higher than the lateral (0.07m RMS) and heading (0.43^o RMS) errors, but all are sufficiently low for autonomous driving.

the true error, as this may lead to over-confident vehicle behaviour in uncertain environments. To evaluate the covariance estimates of the swathe-based localisation algorithm, we use the *normalised estimation error squared* (NEES) [31], which characterises the consistency of a state estimator. The NEES score ϵ_k is computed as follows:

$$\epsilon_k = (\mathbf{x}_k - \hat{\mathbf{x}}_k)^\mathsf{T} \hat{\mathbf{\Sigma}}_k^{-1} (\mathbf{x}_k - \hat{\mathbf{x}}_k) \tag{14}$$

where $\hat{\mathbf{x}}_k$ and $\hat{\mathbf{\Sigma}}_k$ are the estimated location and uncertainty at update k and $\mathbf{x}_k$ is the true location from ground truth. Over all k the set of NEES scores ϵ will follow a chi-squared distribution, and $\hat{\mathbf{\Sigma}}$ can be deemed a *conservative* estimate of the uncertainty if the following condition is satisfied:

$$E[\epsilon] < \dim(\mathbf{x}). \tag{15}$$

For the $\mathbb{SE}(2)$ localisation problem, the expected NEES score $E[\epsilon]$ must fall below the state vector dimension of 3 for it to yield conservative estimates of the uncertainty. Fig. 6 shows the distribution of NEES scores for the swathe localisation experiment. The expected NEES score of 1.13 indicates that the localisation uncertainty estimates provided by the filtering framework slightly overestimate the true error, and therefore are conservative.

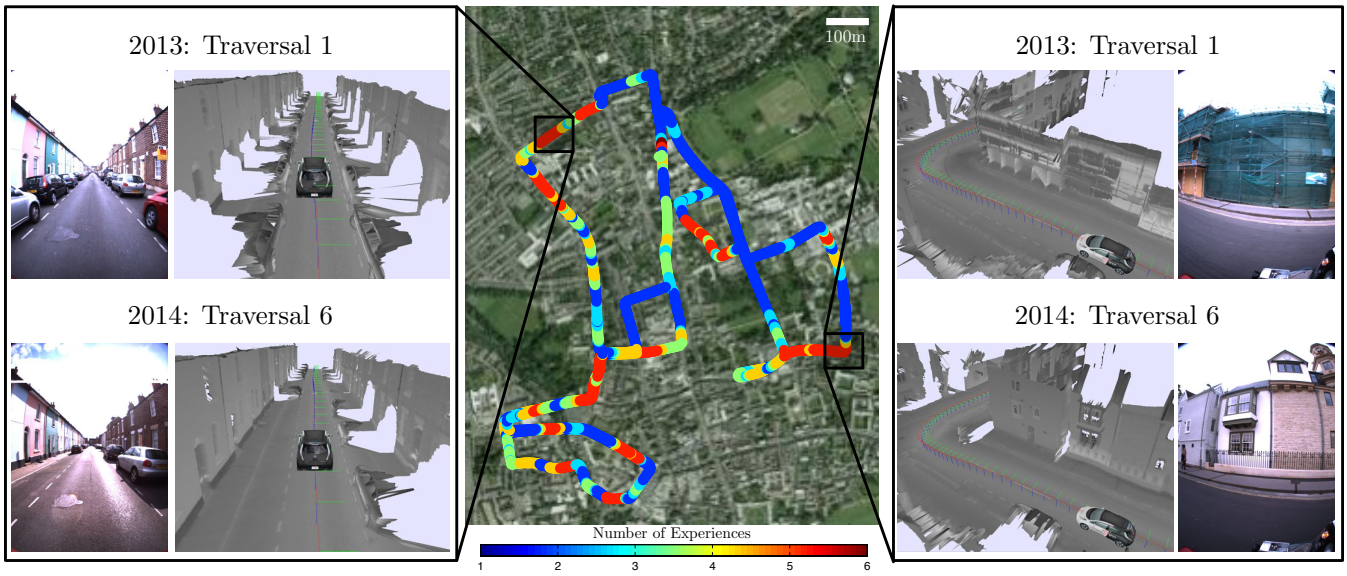


Fig. 7. Test environment with number of experiences recorded for each location after six traversals, along with example experiences in more dynamic parts of the environment. Although a single experience is sufficient for many locations, many parts of the environment require between two and five experiences to represent change over time, including short-term variation such as parked cars (left), pedestrians and vehicle traffic, as well as long-term changes in the environment such as the construction of buildings (right). Images logged by onboard cameras are shown for illustration purposes only. Note that the only location with six experiences is in fact the starting point of the trajectory; the filter reports a high localisation uncertainty when initialising using GPS (and thus records a new experience) but quickly converges once the first swathe localisation is made.

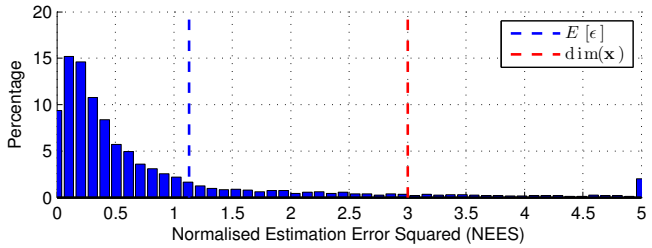


Fig. 6. NEES distribution for the final traversal relative to metric ground truth. The expected value of 1.13 is well below the state dimension of 3, therefore the estimated localisation uncertainty is consistently conservative, and unlikely to confidently assert that the vehicle is in an incorrect location.

B. Experience-based Localisation

For each traversal of the environment, new experiences were recorded at locations where the current set of experiences did not sufficiently match the local 3D swathe. Fig. 7 shows where experiences were recorded along the trajectories, illustrating selected dynamic locations where more experiences were required.

Table II lists statistics for each successive traversal of the test environment, including the cumulative distance travelled, the longitudinal, lateral and heading RMS errors relative to metric ground truth, the average NEES estimate, and the percentage of the route retained as new experiences. As expected, the entire first traversal is recorded as new experiences for all locations in the environment. Each subsequent traversal records fewer experiences, indicating that the set of accumulated experiences yield sufficiently rich representations of the environment to provide consistently good localisation performance over the entire traversal. Additionally, since the solutions from multiple maps are combined in the filtering framework, the RMS localisation errors decrease from 0.5m, 0.3m and 1.5^o after the first traversal to 0.4m, 0.07m and 0.4^o after the sixth traversal. Finally, the expected NEES score $E[\epsilon]$ remains consistently below the state dimension of 3, indicating that the localisation uncertainty estimate remains conservative and does not become over-confident over time; a crucial requirement for long-term operation in dynamic environments.

TABLE II
EXPERIENCE-BASED LOCALISATION PERFORMANCE

Traversal	Cumulative Distance	$\mathbf{x}_{\text{RMS}}$	$\mathbf{y}_{\text{RMS}}$	θ_{RMS}	$E[\epsilon]$	New Experiences
1	8.02km	–	–	–	–	100%
2	16.98km	0.49m	0.32m	1.51^o	1.96	38.14%
3	25.59km	0.44m	0.25m	0.47^o	1.62	22.71%
4	34.57km	0.55m	0.27m	0.67^o	1.75	14.08%
5	43.47km	0.45m	0.16m	0.61^o	1.28	7.34%
6	51.64km	0.38m	0.07m	0.43^o	1.13	1.09%

VIII. CONCLUSIONS

Dealing with structural change in an environment is critical requirement for long-term operation of autonomous vehicles in cities. The range of variation, from pedestrians, vehicles, parked cars, seasonal change and even construction of new buildings is something no static map approach could adequately represent. In this paper we demonstrated real-time, continuously-improving, centimetre-accurate localisation with low-cost 2D LIDAR sensors, with large-scale experiments spanning a year of operation in a dynamic city environment. We presented a probabilistic method of combining localisations from multiple prior maps to improve accuracy, and an efficient implementation that runs in real-time on commodity GPU hardware. We believe that learning from experience is the key to enabling true life-long autonomy for mobile robots in complex, changing cities.

REFERENCES

[1] C. Urmson, J. Anhalt, M. Clark, T. Galatali, J. P. Gonzalez, J. Gowdy, A. Gutierrez, S. Harbaugh, M. Johnson-Roberson, H. Kato, *et al.*, "High speed navigation of unrehearsed terrain: Red team technology for grand challenge 2004," *Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-04-37*, 2004.

[2] S. Thrun, "What we're driving at," 2010. [Online]. Available: http://googleblog.blogspot.co.uk/2010/10/what-were-driving-at.html

[3] M. Bertozzi, L. Bombini, A. Broggi, M. Buzzoni, E. Cardarelli, S. Cattani, P. Cerri, A. Coati, S. Debattisti, A. Falzoni, *et al.*, "Viac: An out of ordinary experiment," in *Intelligent Vehicles Symposium (IV), 2011 IEEE*. IEEE, 2011, pp. 175–180.

[4] J. Ziegler, P. Bender, M. Schreiber, H. Lategahn, T. Strauss, C. Stiller, T. Dang, U. Franke, N. Appenrodt, C. Keller, *et al.*, "Making bertha drive? an autonomous journey on a historic route," *Intelligent Transportation Systems Magazine, IEEE*, vol. 6, no. 2, pp. 8–20, 2014.

[5] J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Z. Kolter, D. Langer, O. Pink, V. Pratt, *et al.*, "Towards fully autonomous driving: Systems and algorithms," in *Intelligent Vehicles Symposium (IV), 2011 IEEE*. IEEE, 2011, pp. 163–168.

[6] Velodyne Lidar. Velodyne HDL-64E Product Description. [Online]. Available: http://velodynelidar.com/lidar/hdlproducts/hdl64e.aspx

[7] Applanix Corp. Applanix POS-LV Product Description. [Online]. Available: http://www.applanix.com/products/land/pos-lv.html

[8] J. Ziegler, H. Lategahn, M. Schreiber, C. G. Keller, C. Knoppel, J. Hipp, M. Haueis, and C. Stiller, "Video based localization for bertha," in *Intelligent Vehicles Symposium Proceedings, 2014 IEEE*. IEEE, 2014, pp. 1231–1238.

[9] A. Broggi, P. Cerri, S. Debattisti, M. C. Laghi, P. Medici, M. Panciroli, and A. Prioletti, "Proud-public road urban driverless test: Architecture and results," in *Intelligent Vehicles Symposium Proceedings, 2014 IEEE*. IEEE, 2014, pp. 648–654.

[10] W. Maddern, A. Stewart, C. McManus, B. Upcroft, W. Churchill, and P. Newman, "Illumination invariant imaging: Applications in robust vision-based localisation, mapping and classification for autonomous vehicles," in *Proceedings of the Visual Place Recognition in Changing Environments Workshop, IEEE International Conference on Robotics and Automation (ICRA)*, Hong Kong, China, May 2014.

[11] R. W. Wolcott and R. M. Eustice, "Visual localization within lidar maps for automated urban driving," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2014, pp. 176–183.

[12] D. Anguelov, C. Dulong, D. Filip, C. Frueh, S. Lafon, R. Lyon, A. Ogale, L. Vincent, and J. Weaver, "Google street view: Capturing the world at street level," *Computer*, vol. 43, no. 6, pp. 32–38, 2010.

[13] W. Churchill and P. Newman, "Practice makes perfect? managing and leveraging visual experiences for lifelong navigation," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Minnesota, USA, May 2012.

[14] I. Baldwin and P. Newman, "Road vehicle localization with 2d push-broom lidar and 3d priors," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 2611–2617.

[15] M. Bosse, P. Newman, J. Leonard, M. Soika, W. Feiten, and S. Teller, "An atlas framework for scalable mapping," in *Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on*, vol. 2. IEEE, 2003, pp. 1899–1906.

[16] M. Bosse and R. Zlot, "Map matching and data association for large-scale two-dimensional laser scan-based slam," *The International Journal of Robotics Research*, vol. 27, no. 6, pp. 667–691, 2008.

[17] ——, "Continuous 3d scan-matching with a spinning 2d laser," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 4312–4319.

[18] M. Buehler, K. Iagnemma, and S. Singh, *The DARPA urban challenge: autonomous vehicles in city traffic*. springer, 2009, vol. 56.

[19] J. Levinson, M. Montemerlo, and S. Thrun, "Map-based precision vehicle localization in urban environments." in *Robotics: Science and Systems*. Citeseer, 2007.

[20] J. Levinson and S. Thrun, "Robust vehicle localization in urban environments using probabilistic maps," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010, pp. 4372–4378.

[21] R. Kummerle, D. Hahnel, D. Dolgov, S. Thrun, and W. Burgard, "Autonomous driving in a multi-level parking structure," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 3395–3400.

[22] I. Baldwin and P. Newman, "Laser-only road-vehicle localization with dual 2d push-broom lidars and 3d priors," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, 2012, pp. 2490–2497.

[23] W. Churchill and P. Newman, "Experience-based Navigation for Long-term Localisation," *The International Journal of Robotics Research (IJRR)*, 2013.

[24] P. Furgale and T. D. Barfoot, "Visual teach and repeat for long-range rover autonomy," *Journal of Field Robotics (JFR)*, vol. 27, no. 5, pp. 534–560, 2010.

[25] T. D. Stoyanov, M. Magnusson, H. Andreasson, and A. Lilienthal, "Fast and accurate scan registration through minimization of the distance between compact 3d ndt representations," *The International Journal of Robotics Research*, p. 0278364912460895, 2012.

[26] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Robotics-DL tentative*. International Society for Optics and Photonics, 1992, pp. 586–606.

[27] A. Censi, "An icp variant using a point-to-line metric," in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*. IEEE, 2008, pp. 19–25.

[28] E. B. Olson, "Real-time correlative scan matching," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 4387–4393.

[29] L. Williams, "Pyramidal parametrics," in *ACM Siggraph Computer Graphics*, vol. 17, no. 3. ACM, 1983, pp. 1–11.

[30] M. Cummins and P. Newman, "FAB-MAP: Probabilistic Localization and Mapping in the Space of Appearance," *The International Journal of Robotics Research*, vol. 27, no. 6, pp. 647–665, 2008.

[31] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: theory, algorithms and software*. John Wiley & Sons, 2004.