Adaptive Data Compression for Robot Perception

Mike Smith and **Ingmar Posner** and **Paul Newman**
Oxford University Mobile Robotics Group
Oxford, UK
{mike, hip, pnewman}@robots.ox.ac.uk

Abstract

This paper concerns the creation of an efficient, continuous, non-parametric representation of surfaces implicit in 3D laser data as typically recorded by mobile robots. Our approach explicitly leverages the probabilistic nature of Gaussian Process regression to provide for a principled, adaptive subsampling which automatically prunes redundant data. The algorithm places no restriction on the complexity of the underlying surfaces and enables predictions at arbitrary locations and densities. We present results using real and synthetic data and show that our approach attains decimation factors in excess of two orders of magnitude without significant degradation in fidelity of the workspace reconstructions.

1 Introduction

Robot perception in complex and dynamic environments is reliant on the timely processing of low-level sensor data to provide a faithful and detailed impression of the current workspace. In recent years, 3D point clouds obtained from purpose-built laser range finders have become a particularly popular environment representation in robotics: they provide a rich source of information regarding 3D geometry. Unfortunately, the data represent only a sparse, oftentimes non-uniform sampling of the world with a manyfold redundancy in information provided per datum: for example, flat surfaces will often be sampled as densely as more complex objects. While the high degree of redundancy results in an unnecessarily inflated computational cost, the non-uniform, sparse nature of the data provides a challenge for standard perception tasks such as object detection and classification or surface reconstruction. Much is to be gained, therefore, by a similarly faithful, yet more compact, continuously queryable representation of 3D environments. The development of such a representation is the subject of this paper.

We restrict our attention to point clouds generated by a simple 3D data acquisition system commonly found in robotics: a single 2D laser scanner is pushed through the workspace and a point cloud is formed by aggregation. Based on this system, Section 3 describes the parameterisation adopted in this work. Our representation is based on a non-parametric method which simultaneously generates a continuous representation of the workspace from discrete laser samples and decimates the data set retaining only locally salient samples. At the heart of our work lies Gaussian Process (GP) regression based, in this case, with support from a finite, time varying region. This sliding-window approach to GP regression is outlined in Section 4. Inspired by the basic mechanisms of GP sparsification [Quiñonero-Candela and Rasmussen, 2005; Csató and Opper, 2002; Seeger *et al.*, 2003], the data in the support region are chosen using the GP predictive distribution itself. We refer to this process as *active sampling* and describe it in Section 4.1. Active sampling explicitly retains data in regions of high complexity and therefore allows us to attain decimation factors in excess of two orders of magnitude without significant degradation in fidelity. Fig. 1 shows a typical example of the output of our system.

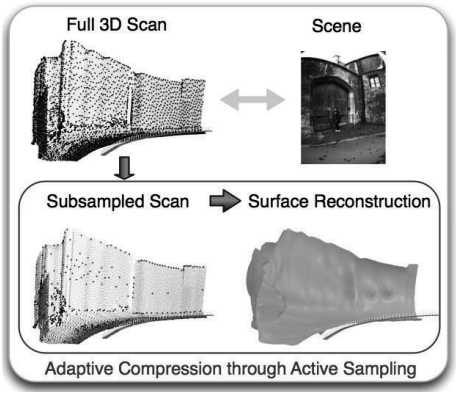

Figure 1: A typical result of our approach: a dense 3D point cloud representing an urban scene is subsampled such that only information rich data are retained (black points in the bottom left image are those that are selected). These form the basis for a continuous surface reconstruction.

2 Related Works

The requirement of robots to operate in large, unstructured environments provides ample incentive for research into suitable models of the workspaces encountered. A large body of work addresses the problem of 3D surface reconstruction using meshing techniques (see, for example, [Früh and

Zakhor, 2003; Hoppe *et al.*, 1993]) where every datum of a point cloud forms the vertex of a polygon. Heuristics are used to achieve the minimum number of vertices required for a suitable representation. Problems are encountered when the data arise from a non-homogeneous sampling of the workspace and/or coverage is incomplete – both are frequently the case in our problem domain. Alternative approaches address the problem by fitting geometric primitives such as planes to subsets of the raw data [Hähnel *et al.*, 2003; Triebel *et al.*, 2005]. This requires a strong prior on which primitives will represent the workspace well. GP models have also become a popular choice in recent years to model the terrain traversed by a robot. This is due to their ability to handle incomplete data in a principled, probabilistic fashion, Examples of such approaches include [Lang *et al.*, 2007; Plagemann *et al.*, 2008; Vasudevan *et al.*, 2009]. However, these works do not address the data compression problem.

While we share the common goal of accurate 3D surface reconstruction with the literature above, this work bears closest relation to that of Gaussian Beam Processes [Plagemann *et al.*, 2007]. The authors model laser range data on a per-scan basis using a GP model to regress range on bearing. However, while our model also regresses on range, we achieve an implicit model of the entire workspace through a sliding window approach and active data selection. This provides significant advantages above and beyond a mere extension of Gaussian Beam Processes to the 3D case.

Our active sampling strategy has been inspired by the body of work addressing GP sparsification, where active subsampling strategies have been used to select information-rich data through use of information theoretic criteria [Krause *et al.*, 2008; Deisenroth *et al.*, 2009; Seeger *et al.*, 2003]. Our sparsification approach is particularly similar to [Seeger *et al.*, 2003]. However, we exploit the time-sequential nature of laser data to form an exact and inexpensive predictive distribution for use in our decision criterion.

3 The Push-Broom 3D Laser System and its Configuration Space

The 3D laser system considered in this work consists of a simple 'push-broom' configuration: a 2D laser sensor is moved through space along an arbitrary path (see Fig. 2-left). A 3D point cloud is constructed by agglomeration of the individual scans. Our aim is to form an implicit representation of workspace surfaces by processing the data gathered. Specifically, we pose queries of points in (x, y, z) on the workspace's surface as range queries along arbitrary rays emanating from the sensor at a point along the sensor's trajectory. Let a beam from the laser be parameterised as a point $\mathbf{q} \in \mathbb{S} \times \mathbb{R}$ where $\mathbf{q} = [\theta, t]^T$. θ denotes the 1D angular position of the laser beam (*not the whole sensor unit*) and t denotes the timestamp of the laser scan. We describe $\mathbf{q}$ as a point in *sensor configuration space* $\mathcal{Q}$ as illustrated in Fig. 2 and refer to it as a particular configuration of the laser sensor. Sensor configuration space provides a natural domain for range regression since it is closely related to the state of the laser and not that of the rest of the robotic system. The mapping from sensor configuration space to range is necessarily

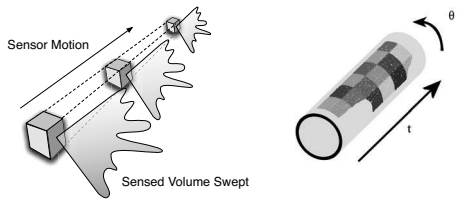

Figure 2: An illustration of the sensor system considered in this work (left) and the corresponding sensor configuration space $\mathcal{Q}$ (right). Every point $\mathbf{q}$ on the manifold in the configuration space is parameterised by time t and angular position of the laser beam θ.

a well formed function – our laser sensor associates a single range r with any point in $\mathcal{Q}$ such that

$$G : \mathbb{S} \times \mathbb{R} \to \mathbb{R}, \mathbf{q} \mapsto r \tag{1}$$

Consider now a mapping from $\mathcal{Q}$ to the Euclidean workspace $\mathcal{W}$

$$E'(\mathbf{q}) = E(V(\mathbf{q}), G(\mathbf{q})) \tag{2}$$

where:

$$V : \mathbb{S} \times \mathbb{R} \to \mathbb{S}^3 \times \mathbb{R}^3, \mathbf{q} \mapsto \mathbf{p} \tag{3}$$

$$E : (\mathbb{S}^3 \times \mathbb{R}^3) \times \mathbb{R} \to \mathbb{R}^3, (\mathbf{p}, r) \mapsto \mathbf{x} \tag{4}$$

For every point $\mathbf{q}$ in $\mathcal{Q}$, $V(\mathbf{q})$ provides the six d.o.f. pose of the sensor's laser beam $\mathbf{p} \in \mathbb{S}^3 \times \mathbb{R}^3$ at the time a measurement is taken – represented as roll, pitch, yaw and position. $E(\mathbf{p}, r)$ maps a six d.o.f. laser beam pose $\mathbf{p}$ and a scalar range measurement r to a single point $\mathbf{x}$ in 3D Euclidean space. This parameterisation naturally eschews the non-functional relation between elevation z and (x, y) location that is commonly found in terrain mapping formulations. By keeping each operation distinct, we also decouple robot trajectory estimation $V(\mathbf{q})$, from that of the regression of the laser data $G(\mathbf{q})$[1]. This permits the independent relaxation of the sensor trajectory (which, for example, could be in response to loop closure events).

To form an estimate of $G(\mathbf{q})$ at any arbitrary position in $\mathcal{Q}$ we turn to Gaussian Process Regression. Given a set of measurements $\mathcal{D} = \{(\mathbf{q}_i, r_i)\}_{i=1}^N$ at a query point $\mathbf{q}^*$ we can obtain a predictive distribution $p(r^*|\mathbf{q}^*, \mathcal{D})$. In the following section we present a brief summary of how this predictive distribution is obtained from a sliding window in sensor configuration space.

4 Sliding-Window Gaussian Process Regression

Gaussian Processes (GPs) provide for non-parametric probabilistic regression. A GP consists of a collection of jointly Gaussian distributed random variables and describes a distribution over latent functions underlying observations. It is fully specified by mean $\mu(\mathbf{q})$ and covariance $k(\mathbf{q}, \mathbf{q}')$ functions. In our application we are concerned with estimating the mapping $G(\mathbf{q})$ corresponding to these latent functions. Given known ranges $\mathbf{r}$ from different configurations $Q = \{\mathbf{q}_i\}_{i=1}^N$

[1]Note that the regression will always be dependent on the ground-truth trajectory of the vehicle, but that no knowledge of this trajectory is required in our algorithm.

and a query point $\mathbf{q}^*$ with corresponding unknown target range r^* we can write:

$$\begin{bmatrix} \mathbf{r} \\ r^* \end{bmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu(Q) \\ \mu(\mathbf{q}^*) \end{pmatrix}, \begin{bmatrix} K(Q,Q) + \sigma_m^2 I & \mathbf{k}(Q,\mathbf{q}^*) \\ \mathbf{k}(Q,\mathbf{q}^*)^T & k(\mathbf{q}^*,\mathbf{q}^*) + \sigma_m^2 \end{bmatrix} \right) \tag{5}$$

In the stationary case, each element of K is derived from a suitably chosen distance metric $d = \|\mathbf{q} - \mathbf{q}'\|$ between two corresponding points $\mathbf{q}$ and $\mathbf{q}'$ in $\mathcal{Q}$. We explicitly account for noise in the training observations $\mathbf{r}$ through an additive white noise process of strength σ_m along the diagonal entries of K[2]. The derivation of the mean $\mathbb{E}[r^*]$ and covariance $\mathbb{V}[r^*]$ of the predictive distribution $p(r^*|\mathbf{q}^*, \mathcal{D})$ for a deterministic $\mu(\mathbf{q}) = 0$ (as is commonly used [Vasudevan *et al.*, 2009]) are standard and can be found, for example, in [Rasmussen and Williams, 2006]

$$\mathbb{E}[r^*] = \mathbf{k}(Q,\mathbf{q}^*)^T (K(Q,Q) + \sigma_m^2 I)^{-1} \mathbf{r} \tag{6}$$

$$\mathbb{V}[r^*] = k(\mathbf{q}^*,\mathbf{q}^*) + \sigma_m^2 - \mathbf{k}(Q,\mathbf{q}^*)^T (K(Q,Q) + \sigma_m^2 I)^{-1} \mathbf{k}(Q,\mathbf{q}^*) \tag{7}$$

Throughout this work we use a member of the Matérn class of covariance functions as advocated in [Stein, 1999]. We note, however, that many others, including non-stationary covariance functions [Lang *et al.*, 2007], could be adapted and substituted in its place. The Matérn class is dependent on a shape parameter ν which regulates the smoothness of the interpolation. It equates to the more standard exponential covariance function as a special case when $\nu = \frac{1}{2}$, and the squared exponential as $\nu \to \infty$. As suggested in [Rasmussen and Williams, 2006], we explored several common choices of $\nu = \left\{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \infty\right\}$ over a number of workspaces varying in complexity. We found that $\nu = \frac{3}{2}$ consistently produced accurate surface reconstructions for a variety of support set sizes and length scales. Although the smoother covariance functions $\nu = \left\{\frac{5}{2}, \infty\right\}$ performed well for simple workspaces, they seemed over constrained in complex scenarios, and vice versa for the rough covariance function $\nu = \frac{1}{2}$. Thus

$$k(\mathbf{q},\mathbf{q}')_{matern\,\nu=\frac{3}{2}} = \sigma_p^2 (1 + \frac{d\sqrt{3}}{l}) \exp(-\frac{d\sqrt{3}}{l}) \tag{8}$$

where σ_p^2 is the process noise, l is the length scale and $d = \|\mathbf{q} - \mathbf{q}'\|_{\mathcal{Q}}$ denotes the geodesic distance between $\mathbf{q}$ and $\mathbf{q}'$ in $\mathcal{Q}$, which for this sensor configuration is the ℓ_2 norm.

In the absence of data, predictions using $G(q)$ tend to the mean function $\mu(\mathbf{q})$. Thus far, this mean function has been specified a priori without accounting for the data. This leads to suboptimal behaviour particularly near the boundaries of the active window. Therefore, we modify our approach by introducing a stochastic mean function such that the GP models the residuals between a local model of the mean derived from data and the support set. This view of GP regression has been used successfully in robotics applications such as [Nguyen-Tuong and Peters, 2010] and, in our case, provides substantial gains in terms of decimation factor. Due to space limitations the details of this approach have been omitted and the reader is referred to [Smith *et al.*, 2010] and [Smith *et al.*, 2011] for a detailed description.

The quantities of data we consider render the application of a single monolithic GP infeasible – time complexity of a naïve implementation of GP regression is cubic in N, the size of the dataset $\mathcal{D}$. Instead, for each prediction we enforce a fixed support window size n, formed from the closest (in terms of the ℓ_1 norm from t to t_*) measurements that have been actively accepted by our algorithm to guarantee constant time operation. The time-sequential nature of laser data ensures that this support window slides across $\mathcal{Q}$ as the robot progresses along its trajectory.

We are interested in minimising the computational complexity of our algorithm for practical applications. Therefore, we now devise an active sampling strategy to intelligently determine a salient subset of $\mathcal{D}$, $\mathcal{D}'$ to include in our active window. The advantages of this are two-fold: we can achieve a significant compression of $\mathcal{D}$, and we can re-use the inversion in Equations 6 and 7 across multiple predictions, thus decreasing computational cost. Where we must incorporate new measurements into the active window we update, rather than re-calculate the inversion. This lowers the overall worst case prediction cost to $\mathcal{O}(n^2)$.

4.1 Active Sampling Using KL Divergence

Measurements are actively selected for inclusion in the support region of the GP regressor based on the information they provide with respect to the current model. Specifically, this decision is taken based on whether the current model's range prediction $r^* \sim \mathcal{N}(\mu_g, \sigma_g^2)$ differs significantly from each measurement $r_m \sim \mathcal{N}(\mu_m, \sigma_m^2)$ at $\mathbf{q}^*$. We use the KL divergence between the two distributions to ascertain if the average additional information required to specify r_m as a result of using r^* (instead of the true r_m) is greater than a threshold κ. This threshold, naturally modulated by environment complexity, affects the rate of compression that is achieved. A sensitivity analysis of our method with a changing κ is presented in Section 6. The KL divergence between two 1D Gaussian distributions has a closed form solution:

$$D_{KL}(\mathcal{N}_m || \mathcal{N}_g) = \frac{1}{2}\left[log_e \frac{\sigma_g^2}{\sigma_m^2} + \frac{\mu_g^2 + \mu_m^2 - 2\mu_g\mu_m + \sigma_m^2}{\sigma_g^2} - 1 \right] \tag{9}$$

On adoption of any new measurement into the active support set an *introspection step* analyses the closest (in terms of ℓ_1 norm in $\mathcal{Q}$) r_m that had previously been marked as redundant to determine if this assessment has changed in light of the new information. On adoption of this measurement further introspection steps are applied for the next closest redundant r_m until there are no further adoptions. This approach is demonstrated in Figure 3. Importantly, it is the introspection steps that allow both sides of discontinuities to be retained given the final sampling of the opposite side of the discontinuity. Figure 4 demonstrates the results of this reverse sweep, and the typical $\mathcal{D}'$ that is stored and used for subsequent predictions instead of $\mathcal{D}$. In the worst case, where there are truly complex sections of the workspace, our algorithm performs as well as a naïve implementation by using

[2] In this paper we have used $\sigma_m^2 = 0.01m^2$.

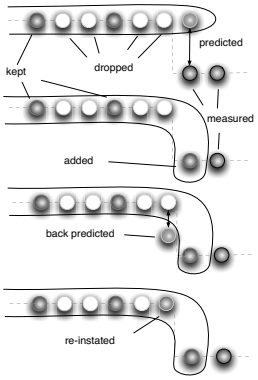

Figure 3: An illustration of the active sampling process in 1D. Each node represents a measurement: black indicates data accepted by the forward pass; white nodes have been rejected; blue nodes are yet to be observed; green nodes are predictions made by our algorithm and red nodes are data accepted during introspection steps. A predicted observation has a KL divergence from the actual measurement greater than κ (top). Hence, it is adopted into the active support set (2nd row). The algorithm then makes a back prediction which is now in error with the measurement (3rd row). The measurement, which was previously rejected, is therefore accepted into the active region (bottom).

all $\mathcal{D}$, while maintaining the ability to automatically subsample simple scenes. Note that this subsampling provide an information rich subset of $\mathcal{D}$ which is desirable for common applications such as data registration, where careful selection of measurements can increase accuracy and robustness [Rusinkiewicz and Levoy, 2001].

5 Experimental Setup

We now analyse system performance using both synthetic data and a real dataset collected from an outdoor urban environment. In particular, we investigate how a changing κ affects the compression rate ($\frac{|\mathcal{D}'|}{|\mathcal{D}|}\,\%$) of the original point cloud.

The real data used are part of the publically available New College data set [Smith *et al.*, 2009]. Data were gathered from a two-wheeled Segway RMP~200 platform with two vertically aligned SICK LMS291-S14 laser sensors mounted on the side of the vehicle. Pose estimates were obtained using visual odometry from a forward facing stereoscopic camera [Newman *et al.*, 2009]. Throughout, we use the Matérn class of covariance functions, a length scale l of $8\,units$, a process variance σ_p^2 of $0.05\,m^2$ and an active window size of 200 measurements. These parameters were determined empirically to produce consistently accurate surface representations, for a given compression rate, as discussed in Section 4. The reconstruction error (in metres) is defined as the ℓ_1 norm in range between a hold out set of the real measurements and corresponding predictions at the same point in $\mathcal{Q}$. Predictions are made as the centre of the active window passes the hold out set in $\mathcal{Q}$ and are conditioned on the current active set.

6 Results

Figure 5 depicts typical subsampling as κ is varied. For aggressive decimation factors of 1,000 (corresponding to $\kappa = 3.18\,nats$) to we achieve a modest mean error of $0.3\,m$. With $\kappa = 0.3\,nats$ errors are comparable with measurement precision of $0.015\,m$ and we need retain only one sixth of the original data.

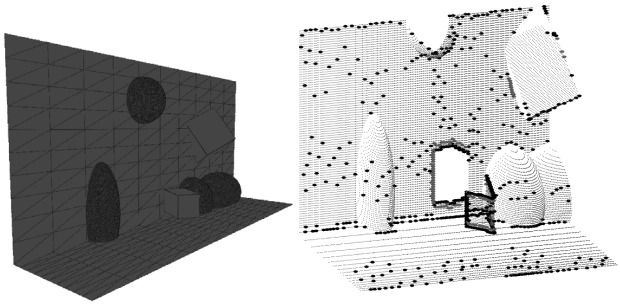

Figure 4: The result of applying our active sampling approach to data from a synthetic environment. (Left) The CAD model of the environment. (Right) The synthetic data. Blue indicates all laser measurements, black denotes measurements selected on the forward pass of the algorithm, and red are measurements that have been chosen through introspection steps. Here $\kappa = 0.8\,nats$. Note that the discontinuity caused by the floating cube in the centre of the scene is densely sampled on both sides.

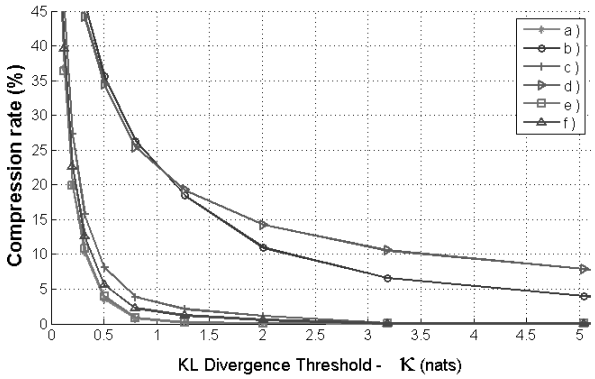

Figure 6: Compression rate, measured as the percentage of $\mathcal{D}$ selected by our algorithm ($\frac{|\mathcal{D}'|}{|\mathcal{D}|}\,\%$), versus κ for exemplars (a) to (d) in Fig. 7. (e) and (f) refer to results on the synthetic environment shown in Fig. 4. For (e) only the wall and ground plane are used. (f) uses the entire model. A lower compression rate results from higher thresholds and scenes that are less complex.

In Figure 6 results are collated across a range of κ for the four exemplars, a) through d) (see Figure 7) and the two synthetic cases, e) and f) (see Figure 4). Intuitively, as the threshold is increased the amount of data retained decreases. This is accompanied by an increase in reconstruction error as depicted by the box plots in Figure 7. As one can discern from the images and CAD models, the relative positions of the curves correspond to the scene complexity: the more complex the scene, the greater the compression rate. Scenes b) and d) are the most complex, with noisy foliage, measurements of ceilings behind window panes and discontinuities as great as $5\,m$, compared to that of the $1\,m$ discontinuity the person in c) presents. In all cases, these complex regions have been sampled most heavily – the outline of the person can be recognised in the Euclidean plot of c) in Figure 7.

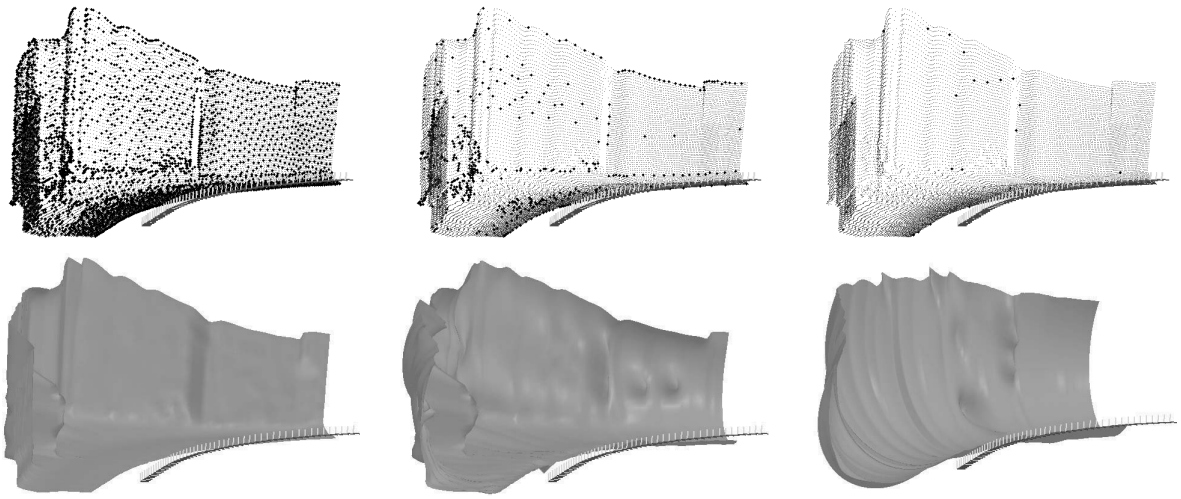

Figure 5: The subsampling process for various values of KL divergence threshold, κ, for test case c) plotted in 3D Euclidean space. The original measurements are in blue. Measurements selected by our active sampling algorithm are in black. The bottom row shows surface representations using only the measurements selected by our algorithm.

7 Conclusions

This paper presents an overview of an adaptive 3D point-cloud compression algorithm first proposed in [Smith *et al.*, 2010]. The approach leverages a Gaussian Process framework to provide a continuous representation of the implicit surfaces underlying 3D laser point clouds commonly encountered in robot perception. Adaptive data compression is achieved via an information theoretic selection criterion applied in a sliding window. The resulting algorithm decimates point clouds of simple workspaces by factors in excess of two orders of magnitude without significant degradation in fidelity. The computational complexity of the algorithm proposed is squared in the size if of the active window, which is constant, rather than cubic in the size of the data set. For a more detailed description of this work as well as an investigation into the use of non-stationary covariance functions that are used for the GP regression at the heart of the system, the reader is referred to [Smith *et al.*, 2010] and [Smith *et al.*, 2011].

8 Acknowledgements

This work was funded through an Engineering and Physical Sciences Research Council CASE award with BAE Systems, and by the European Commission under grant agreement number FP7-231888-EUROPA. We would also like to thank Winston Churchill for provision of the CAD models used throughout this paper.

References

[Csató and Opper, 2002] L. Csató and M. Opper. Sparse online Gaussian processes. *Neural Computation*, 14(3):641–668, 2002.

[Deisenroth *et al.*, 2009] M.P. Deisenroth, C.E. Rasmussen, and J. Peters. Gaussian process dynamic programming. *Neurocomputing*, 72(7-9):1508–1524, 2009.

[Früh and Zakhor, 2003] C. Früh and A. Zakhor. Constructing 3d city models by merging aerial and ground views. *IEEE Comput. Graph. Appl.*, 23(6):52–61, 2003.

[Hähnel *et al.*, 2003] D. Hähnel, W. Burgard, and S. Thrun. Learning compact 3D models of indoor and outdoor environment with a mobile robot. *Robotics and Autonomous Systems*, 44(1):15–27, 2003.

[Hoppe *et al.*, 1993] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle. Mesh optimization. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 19–26. ACM New York, NY, USA, 1993.

[Krause *et al.*, 2008] A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *The Journal of Machine Learning Research*, 9:235–284, 2008.

[Lang *et al.*, 2007] T. Lang, C. Plagemann, and W. Burgard. Adaptive non-stationary kernel regression for terrain modeling. In *Robotics: Science and Systems (RSS)*, 2007.

[Newman *et al.*, 2009] P.M. Newman, G. Sibley, M. Smith, M. Cummins, A.R. Harrison, C. Mei, I. Posner, R. Shade, D. Schroeter, D.M. Cole, and I. Reid. Navigating, recognising and describing urban spaces with vision and laser. *IJRR*, 28(11-12):1406–1433, November 2009.

[Nguyen-Tuong and Peters, 2010] D. Nguyen-Tuong and J. Peters. Using Model Knowledge for Learning Inverse Dynamics. In *IEEE International Conference on Robotics and Automation*, 2010.

[Plagemann *et al.*, 2007] C. Plagemann, K. Kersting, P. Pfaff, and W. Burgard. Gaussian beam processes: A

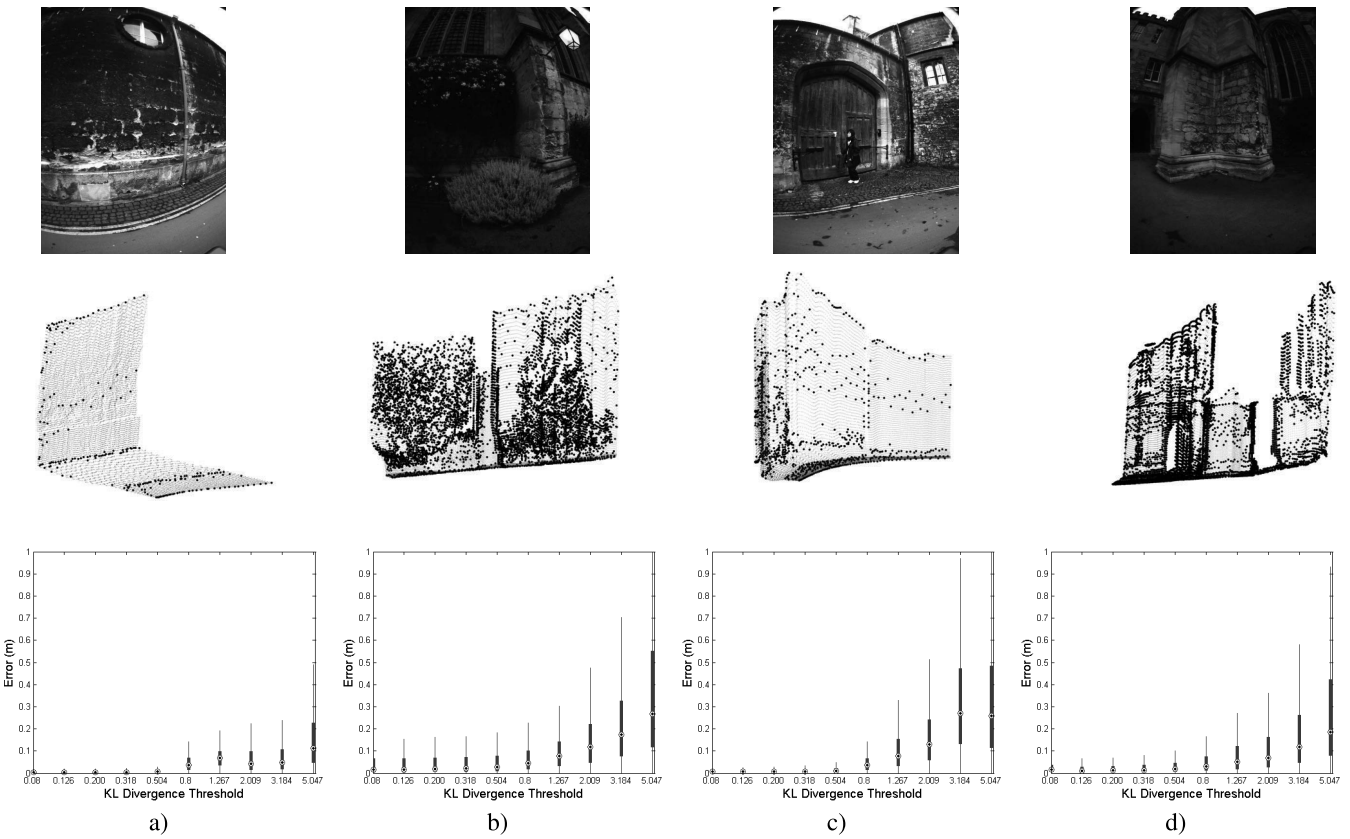

Figure 7: Real-world exemplars (column-wise). Images (top) were captured at the same time as the laser data (2nd row). Blue points represent the raw data. Black points indicate actively sampled data ($\kappa = 0.8\,nats$). (Note that, for ease of interpretation, images and scans are not exactly aligned and sometimes vary in aspect angle and field of view relative to each other.) Our approach retains the majority of data in salient regions: at the intersections of planes, around the outline of a person (to the left in the scan in part (c)), on bushes and other complex geometries. The box plots (bottom row) denote median and interquartile range for each of the logarithmically spaced thresholds in Figure 6. Whiskers extend a maximum of one and a half times the interquartile range.

nonparametric bayesian measurement model for range finders. In *Robotics: Science and Systems (RSS)*, 2007.

[Plagemann *et al.*, 2008] C. Plagemann, S. Mischke, S. Prentice, K. Kersting, N. Roy, and W. Burgard. Learning predictive terrain models for legged robot locomotion. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2008.

[Quiñonero-Candela and Rasmussen, 2005] J. Quiñonero-Candela and C.E. Rasmussen. A unifying view of sparse approximate gaussian process regression. *J. Mach. Learn. Res.*, 6:1939–1959, 2005.

[Rasmussen and Williams, 2006] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006.

[Rusinkiewicz and Levoy, 2001] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Proc. 3DIM*, pages 145–152, 2001.

[Seeger *et al.*, 2003] M. Seeger, C.K.I. Williams, and N.D. Lawrence. Fast forward selection to speed up sparse Gaussian process regression. In *Workshop on AI and Statistics*, volume 9, 2003.

[Smith *et al.*, 2009] M. Smith, I. Baldwin, W. Churchill, R. Paul, and P.M. Newman. The new college vision and laser data set. *International Journal for Robotics Research (IJRR)*, 28(5):595–599, May 2009.

[Smith *et al.*, 2010] M. Smith, I. Posner, and P. Newman. Efficient Non-Parametric Surface Representations Using Active Sampling for Push Broom Laser Data. In *Proceedings of Robotics: Science and Systems*, Zaragoza, Spain, 2010.

[Smith *et al.*, 2011] M. Smith, I. Posner, and P. Newman. Adaptive compression for 3d laser data. *To appear in the International Journal for Robotics Research (IJRR) - Special Issue on RSS VI*, 2011.

[Stein, 1999] M. L. Stein. *Interpolating Spatial Data*. Springer-Verlag, 1999.

[Triebel *et al.*, 2005] R. Triebel, W. Burgard, and F. Dellaert. Using hierarchical em to extract planes from 3d range scans. In *Int. Conf. Robotics and Automation (ICRA)*, 2005.

[Vasudevan *et al.*, 2009] S. Vasudevan, F. Ramos, E. Nettleton, H. Durrant-Whyte, and A. Blair. Gaussian process modeling of large scale terrain. In *Int. Conf. Robotics and Automation (ICRA)*, 2009.