2010 IEEE International Conference on Robotics and Automation
Anchorage Convention District
May 3-8, 2010, Anchorage, Alaska, USA

Heteroscedastic Gaussian Processes for Data Fusion in Large Scale Terrain Modeling

Shrihari Vasudevan, Fabio Ramos, Eric Nettleton and Hugh Durrant-Whyte
Australian Centre for Field Robotics (ACFR)
The University of Sydney, NSW 2006, Australia
shrihari.vasudevan@ieee.org,{f.ramos,e.nettleton,hugh}@acfr.usyd.edu.au

Abstract— **This paper presents a novel approach to data fusion for stochastic processes that model spatial data. It addresses the problem of data fusion in the context of large scale terrain modeling for a mobile robot. Building a model of large scale and complex terrain that can adequately handle uncertainty and incompleteness in a statistically sound manner is a very challenging problem. To obtain a comprehensive model of such terrain, typically, multiple sensory modalities as well as multiple data sets are required. This work uses Gaussian processes to model large scale terrain. The model naturally provides a multi-resolution representation of space, incorporates and handles uncertainties appropriately and copes with incompleteness of sensory information. Gaussian process regression techniques are applied to estimate and interpolate (to fill gaps in unknown areas) elevation information across the field. In this work, the GP modeling approach is extended to fuse multiple, multi-modal data sets to obtain a best estimate of the elevation given the individual data sets. The individual data sets are treated as different noisy samples of the same underlying terrain. Experiments performed on sparse GPS based survey data and dense laser scanner data taken at mine-sites are reported.**

I. Introduction

Large scale terrain mapping is an essential problem in a wide range of applications, from space exploration to mining and more. For autonomous robots to function in such high-value applications, an efficient, flexible and high-fidelity representation of space is critical. The key challenges in realizing this are that of dealing with the problems of uncertainty, incompleteness and handling highly unstructured terrain. Uncertainty and incompleteness are virtually ubiquitous in robotics as sensor capabilities are limited. The problem is magnified in a field robotics scenario due to sheer scale of the application (for instance, a mining or space exploration scenario).

State-of-the-art surface mapping methods employ representations based on triangular tesselations. This process, however, does not have a statistically sound way of incorporating and managing uncertainty. The assumption of statistically independent data is a further limitation of many works that have used these approaches. While there are several interpolation techniques known, the independence assumption can lead to simplistic (simple averaging like) techniques that result in inaccurate modeling of the terrain. Further, the limited perceptual capabilities of sensors renders most sensory data incomplete. In [1], a Gaussian process based terrain modeling approach is proposed that provides a multi-resolution representation of space, incorporates and manages uncertainty in a statistically sound way and handles spatially correlated data in an appropriate manner. This work builds on the GP modeling approach. It proposes a data fusion method for data that are modeled using Gaussian processes. It demonstrates the benefits of the data fusion approach in the context of overcoming sensor limitations.

Typically, sensory data is incomplete due to the presence of entities that occlude the sensors view. This is compounded by the fact that every sensor has a limited perceptual capability i.e. a limited range and limited applicability (eg. one sensor may not be usable in a certain area that another sensor can be accessed with). Thus, most large scale modeling experiments would ideally require multiple sensory snapshots and multiple sensors to obtain a more complete model. These sensors may also have different characteristics (accuracies). The problem is in fusing these multiple and multi-modal sensory data sets - this is the theme of the paper. Terrain data can be obtained using numerous sensors including 3D laser scanners and GPS. 3D laser scanners provide dense and accurate data whereas a GPS based survey typically comprises of a relatively sparse set of well chosen points of interest. The experiments reported in this work use data sets obtained from both these sensors to develop an integrated picture of the terrain.

The contribution of this work is a novel approach to fusing multiple, multi-modal data sets towards obtaining a comprehensive model of the terrain under consideration. The fusion technique itself is generic and applicable as a general Gaussian process fusion methodology. The fusion approach is based on the underlying principles of the Gaussian process itself and hence is well founded. Experiments conducted using real data sets obtained from GPS and laser scanner based surveys in real application scenarios (mining) are reported in support of the proposed approach.

II. Related Work

State-of-the-art representations used in applications such as mining, space exploration and other field robotics scenarios as well as in geospatial engineering are typically limited to elevation maps ([2] and [3]), triangulated irregular networks (TIN's) ([4] and [5]), contour models and their variants or combinations ([6] and [7]). Each of these methods have their own strengths and preferred application domains. The former two are more popular in robotics. All of these representations, in their native form, do not handle spatially correlated data effectively and do not have a statistically correct way of incorporating and managing uncertainty.

Gaussian processes [8] (GP's) are powerful non-parametric learning techniques that can handle these issues. They produce a scalable multi-resolution model of the large scale terrain under consideration. They yield a continuous domain representation of the terrain data and hence can be sampled at any desired resolution. They incorporate and handle uncertainty in a statistically sound way and represent spatially correlated data in an appropriate manner. They model and use the spatial correlation of the given data points to estimate the elevation values for other unknown points of interest. In an estimation sense, GP's provide the best linear unbiased estimate [9] based on the underlying stochastic model of the spatial correlation between the data points. They basically perform an interpolation methodology called *Kriging* [10] which is a standard interpolation technique used in the mining industry. GP's thus handle both uncertainty and incompleteness effectively.

Recently, Gaussian processes have been applied in the context of terrain modeling - see [11] and [1]. The former work is based on using a non-stationary equivalent of a stationary squared exponential covariance function [12] and incorporates kernel adaptation techniques to adequately handle both smooth surfaces as well as inherent (and characteristic) surface discontinuities. It introduces the idea of a "hyper-GP", using a stationary kernel, to predict the most probable length scale parameters to suit the local structure. It also proposes to model space as an ensemble of GP's to reduce computational complexity. The latter work [1], proposes the use of non-stationary kernels (neural network) to model large scale discontinuous spatial data. It shows that using a suitable non-stationary kernel can directly result in modeling local structure and smoothness. It also proposes a local approximation methodology to address scalability issues relating to the application of this approach to large scale data sets. This approximation technique is based on an efficient hierarchical representation (KD-tree) of the data. It also compares performances of GP's based on stationary (squared exponential) and non-stationary (neural network) kernels as well as several other standard interpolation methods applicable to elevation maps and TIN's, in the context of large scale terrain modeling.

This paper builds on the work presented in [1]. It extends the GP terrain modeling approach to handle multiple multi-modal data sets. It treats the data-fusion problem as a problem of combining different noisy data samples of the common entity being modeled. In the Machine Learning community, this idea is referred to as heteroscedastic GP's ([13], [14], [15] and recently, [16]). Both [13] and [16] are particularly relevant to this work. They model the noise variance using a separate GP in addition to the GP governing the noise free output. While Goldberg et al [13] follow use Markov Chain Monte-Carlo techniques (MCMC) to estimate the posterior noise variance, Kersting et al [16] propose a maximum-likelihood approach (faster) using an EM like iterative optimization procedure to compute the noise variances. The fusion approach presented in this paper is based on the concept of heteroscedastic GP's. It treats individual terrain data sets as homoscedastic in nature but different data sets considered together form a heteroscedastic system.

Two other related works that attempt the problem of data fusion in the context of Gaussian processes include [17] and [18]. While the former is based on similar assumptions to this work, it bears a "hierarchical learning" flavor to it in that it essentially demonstrates how a GP can be used to model an expensive process by (a) modeling a GP on an approximate or cheap process and (b) using the many input-output data from the approximate process and the few samples available of the expensive one together in order to learn a GP for the expensive process. The work [18] attempts to generalize arbitrary transformations on GP priors through linear transformations. It hints at how this framework could be used to introduce heteroscedasticity and how information from different sources could be fused. However, specifics on how the fusion can actually be performed are beyond the scope of the work.

The contribution of this work is a novel method of fusing multiple multi-modal large scale terrain data sets into an integrated model. The approach is similar to the idea of heteroscedastic GP's as presented in [13] and [16]. However, this work builds up on the ideas proposed in [1], it does not use a separate GP to model noise and does not rely on computationally expensive MCMC based approaches. The approach is tailored towards handling large data sets ($\sim$1 million data points per data set) and thus relies heavily on the local approximation methods. Experiments conducted on multiple multi-modal data sets taken from real application scenarios (mine sites) are reported and findings discussed. Note that this work develops only the fusion methodology. The registration of individual data sets to a common reference frame is assumed given for this work.

III. Approach

A. Problem Definition

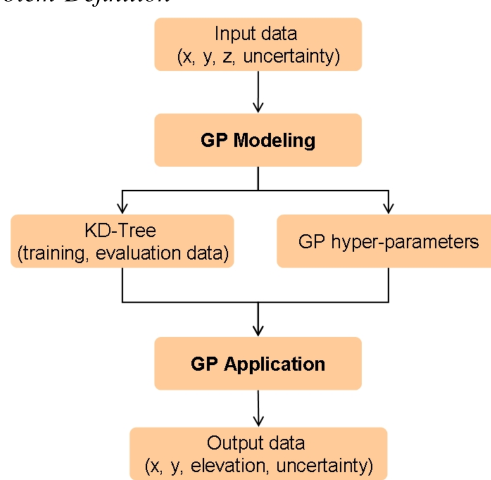

Fig. 1. The overall process of learning and using Gaussian processes to model terrain. Data from any sensor along with noise estimates are provided as the input to the process. The modeling step then learns the appropriate Gaussian process model (hyper-parameters of the kernel) and the result of the modeling step are the GP hyperparameters together with the training/evaluation data. These are then used to perform GP regression to estimate the elevation data across any grid of points, of the users interest. The application step produces a 2.5D elevation map of the terrain as well as an associated uncertainty for each point.

Gaussian process fusion in the context of combining multiple data sets for large-scale terrain modeling can be understood as follows. Given multiple sensors (GPS, laser, etc.) that provide terrain data sets, each spanning a certain area and comprising of a set of points (x, y, z) in 3D space, the objective of this work is to develop a single multi-resolution probabilistic representation of the terrain. This will enable better handling of typical individual sensor limitations such as incompleteness due to occlusions and limited perceptual capability (range and applicability). The prior work [1] detailed the basic step of modeling a single data set from a single sensor using Gaussian processes and applying it to estimate terrain elevation in an area/resolution of interest. This is summarized in Figure 1.

B. Fusing multiple data sets

The fusion methodology is based on two underlying ideas

1) Data from the same entity can be modeled using a single set of GP hyperparameters with just the noise parameter varying between data sets. Thus, the data sets are considered as different noisy samples of a common terrain that has to be modeled.
2) The fusion problem is treated as a standard GP regression/estimation problem with data having different noise parameters. The formulation is similar to the heteroscedastic GP formulation described in [13] and [16].

Given multiple data sets (possibly multi-modal) of the terrain being modeled, the objective is to estimate the elevation at a point given all prior data sets as well as the respective GP's (hyperparameters) that are used to model them. This can be specified as

$$\mathbb{E}[\, f_*(\mathbf{x}_*)\,]\ ,\ var(f_*(\mathbf{x}_*))\ |\ X_i\ ,\ GP_i\ ,\ x_*\ , \qquad (1)$$

where X_i are the given data sets, GP_i are their respective GP model hyperparameters and i varies from 1 to the number of data sets available.

The joint distribution of any finite number of random variables of a GP is Gaussian. Thus, the joint distribution of the training outputs $\mathbf{z}$ and test outputs f_* given this prior can be specified by

$$\begin{bmatrix} \mathbf{z} \\ f_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(X,X)+\Sigma & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*) \end{bmatrix}\right), \qquad (2)$$

where

$$\mathbf{z} = [\ \mathbf{z}_1\ ,\ \mathbf{z}_2\ ,\ \mathbf{z}_3\ ,\ \dots\ ,\ \mathbf{z}_n\ ]'$$

is the output elevation values of the selected training data from the individual data sets,

$$X = [\ X_1\ ,\ X_2\ ,\ X_3\ ,\ \dots\ ,\ X_n\ ]$$

are the input location values of the selected training data from the respective individual data sets and

$$\Sigma = diag[\ \sigma^2_{1\{1\ \dots\ N_1\}}\ ,\ \sigma^2_{2\{1\ \dots\ N_2\}}\ ,\ \sigma^2_{n\{1\ \dots\ N_n\}}\ ]$$

is a diagonal noise matrix where each noise term is repeated as many times as the number of training data taken from the corresponding data set, denoted by N_1 , N_2 , ... , N_n. The experiments performed in this paper use the neural network kernel, although, any kernel ([8]) may be used so long as the same kernel is used for modeling each of the individual data sets. For N training points and N_* test points, $K(X, X_*)$ denotes the $N \times N_*$ matrix of covariances evaluated at all pairs of training and test points. $K(X, X)$, $K(X_*, X_*)$ and $K(X_*, X)$ can be similarly defined.

$$\bar{f}_* = K(X_*, X)\,[K(X,X) + \Sigma]^{-1}\,\mathbf{z} \qquad (3)$$

$$\text{cov}(f_*) = K(X_*, X_*) - K(X_*, X)[K(X,X)+\Sigma]^{-1}\,K(X, X_*) \qquad (4)$$

The mean and variance of the elevation estimate can thus be obtained by applying Equations 3 and 4, incorporating multiple data sets in the component terms as shown before. This estimate is the conditional estimate at a desired point given the multiple and possibly multi-modal data sets and their respective GP models. The data sets may thus be fused to generate integrated and comprehensive terrain models.

C. Mathematical Properties

1) Batch estimator: Equations 3 and 4 provide the batch fusion estimator, ie. they provide the conditional mean and variance in elevation given all the data sets taken together.

2) Decrease in uncertainty: It can be shown that the formalism adopted in this paper guarantees that with the addition of data sets (any number, from any sensor), the uncertainty in the fused elevation estimate cannot increase. If the new or incoming data set has relevant information for the prediction at a query point in the first data set, the posterior uncertainty will decrease; if there is no relevant information (assume, for instance, no points are selected from successive data sets for a particular query point), the uncertainty will remain same.

The detailed derivation is not included here due to paucity of space, however, it is based on the following idea - without loss of generality, the difference between the posterior uncertainty using a single data set to that obtained using two data sets can be shown to be a positive semi-definite matrix. This change in uncertainty will be based on the "information-gain" provided by points of the successive data sets to the prediction at a query point using the first data set. For two data sets, this is specified by

$$\alpha_{21} = \begin{bmatrix} K_{11}^{-1}K_{12}\tilde{K}^{-1}K_{21}K_{11}^{-1} & -K_{11}^{-1}K_{12}\tilde{K}^{-1} \\ -\tilde{K}^{-1}K_{21}K_{11}^{-1} & \tilde{K}^{-1} \end{bmatrix} \qquad (5)$$

where $\tilde{K} = K_{22} - K_{21}K_{11}^{-1}K_{12}$, $K_{11} = K(X_1, X_1) + \sigma_1^2 I$ represents the covariance matrix of the training data selected (for a query point) from the first data set. $K_{12} = K(X_1, X_2)$, $K_{21} = K(X_2, X_1)$ and $K_{22} = K(X_2, X_2) + \sigma_2^2 I$ can be similarly defined. This term is always positive semi-definite guaranteeing that the uncertainty will either remain the same or decrease but never increase. This condition is used in the experiments (Section IV) to verify if the fusion actually occurs.

3) Recursive estimator:

$$E(f_*|X_1, X_2) = E(f_*|X_1) + [K(X_*, X_1), K(X_*, X_2)]\, \alpha_{21}\, [\mathbf{z_1}, \mathbf{z_2}]' \quad (6)$$

$$\mathrm{cov}(f_*|X_1, X_2) = \mathrm{cov}(f_*|X_1) - [K(X_*, X_1), K(X_*, X_2)]\, \alpha_{21} \begin{bmatrix} K(X_1, X_*) \\ K(X_2, X_*) \end{bmatrix} \quad (7)$$

The information gain term (Equation 5) can be used to derive the conditional mean and variance in a recursive form, respectively specified by Equations 6 and 7. This form enables a recursive fusion process wherein the previous best estimate (and its uncertainty) together with the information gain from the new data can be used to derive the new fused elevation estimate and its uncertainty. Equation 7 also shows that the uncertainty will not increase with the addition of new data.

D. Learning the hyperparameters

Learning of hyperparameters is based on the maximum marginal likelihood framework demonstrated in [1]. Only one set of hyperparameters are used with the noise parameter alone varying across data sets. These hyperparameters may be obtained by selecting training data from each of the data sets and doing a "joint learning" exercise. This method is computationally expensive and is limited by the computational resources available.

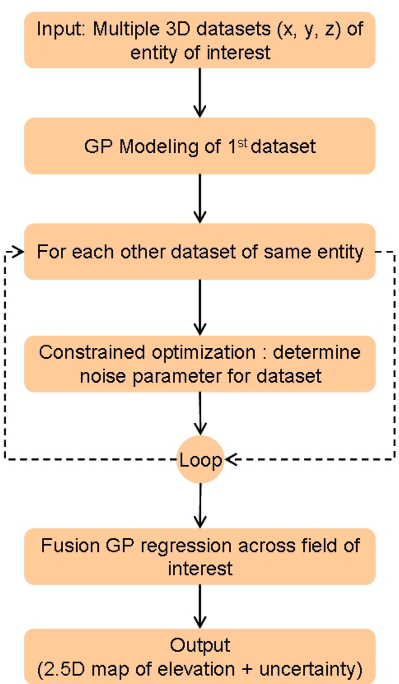

Fig. 2. The fusion process: For a first data set - a full GP learning procedure is done as in [1]. Thereafter for every successive data set of the same entity, the hyperparameters of the first GP are fixed and only the noise parameter is adjusted to model the individual data sets. The end result is a single set of GP hyperparameters along with a set of noise parameters corresponding to each of the data sets. The estimation (GP application) process estimates the expected elevation at a query point given the different noisy data obtained from the respective data sets, in its vicinity.

An alternative approach to GP-hyperparameter learning in the context of this paper is shown in Figure 2. Here, the first data set in consideration is modeled using standard GP modeling procedure described in [1]. Thereafter, for each other data set, a constrained optimization method is adopted where the hyperparameters obtained before are re-used, the noise hyperparameter alone being modified in order to model the successive data sets. The ordering of the data sets is irrelevant. This approach is akin to modeling successive data sets in terms of the existing data set and then fusing them using Equations 3 and 4. The advantages of this approach include the bounding of the computational complexity and the ability to run the optimization operations in parallel. The former is particularly important in the context of this work.

E. Point selection strategy

The prior work [1] was concerned with the GP modeling of a single terrain data set. It used a KD-tree based "moving window"/nearest neighbor approximation in the GP application/inference stage. The work described in this paper attempts to fuse information from the first data set with multiple other data sets. This involves the selection of a *relevant* subset of points from the second and successive data sets for the GP regression process. While the default option of using the same KD-tree approximation on every other data set exists, this work attempted to perform a more systematic analysis of point selection methods from successive data sets. The following methods were tested -

1) Initial point selection
   a) The default KD-tree based local approximation that selects the n nearest neighbors to the query point.
   b) The distance between the query point and all n points from the first data set used for prediction is computed. For successive data sets, *all points* within the maximum value of the distance are used.
   c) Similar approach to 1b but the median value of the distance was used. The median (and not mean) was chosen as it was thought to be less sensitive to exceptional cases.
2) Distance based filtering - This is applied only to approach 1a from the initial point selection. The objective was to filter out points from successive data sets which were not in the same neighborhood as the points from the first data set. This would occur when there are inaccuracies in registration or when sensors with very different characteristics (resolution, noise) are used. The following two options were tested:
   a) The distance between the query point and all n points from the first data set used for prediction is computed. For successive data sets, only the subset of the n nearest neighbors that lie within the maximum value of the distance are used.
   b) Similar approach to 2a but the median value of the distance was used.
3) Entropy based filtering - An entropy based selection method would choose only points that would significantly reduce the uncertainty in the resulting elevation

estimate (maximum decrease in entropy or maximum information gain [19]). The following two options were tested:

a) Single-shot entropy computation - This computes the entropy reduction for each point considered individually, taken with the points from the first data set. Subsequently, points which reduce entropy most (empirically set threshold) are chosen. The number of points does not exceed the number of points chosen from the first data set - n.
b) Incremental entropy computation - This computes entropy reduction incrementally. In each cycle, points from the first data set as well as points selected from previous cycles are taken together with all other (not yet selected) points individually. The point which reduces entropy most is then selected for the current cycle. In this way, the final set of points *together* reduce entropy the most.

Entropy reduction was computed as the log of the ratio of the posterior uncertainty of a query point after including a certain training/evaluation point, to its uncertainty without including the training/evaluation point in consideration. It is a measure of the information content in the training/evaluation point in consideration, in relation to performing GP regression at a particular query point. The entropy change of the q^{th} point of the p^{th} data set is specified by Equation 8

$$I_{pq} \;=\; \log\left(\frac{k(x_*|X_1, X_{p,q})}{k(x_*|X_1)}\right), \tag{8}$$

where $k(x_*|X)$ is the posterior uncertainty (of the estimate obtained from GP regression) of query point x_*, obtained from Equation 4 as

$$k(x_*|X) = k(x_*, x_*) - k(x_*, X)[K(X, X) + \Sigma]^{-1}k(X, x_*)$$

Thirteen relevant combinations of these methods (1a only, 1a + 2a, 1a + 2b, 1a + 3a, 1a + 3b, 1a + 2a + 3a, 1a + 2a + 3b, 1a + 2b + 3a, 1a + 2b + 3b, 1b + 3a, 1b + 3b, 1c + 3a, 1c + 3b) were tested in the context of fusing (a) a sparse data set with a dense one, (b) a dense data set with a sparse one and (c) two dense data sets. The data sets were real mining data sets (also used in Section IV) taken with GPS (sparse) and RIEGL laser scanner (dense) sensors of a mine pit. Paucity of space necessitates only mentioning the final conclusions.

No significant gain in prediction accuracy was achieved by using approaches 1b or 1c, but computational complexity significantly increased. Distance based filtering reduced prediction error as some of the nearest neighbors from successive data sets were not situated within the local region spanned by the points from the first data set. Using very distant points lead to smoothing and an increase in error. Both 2a and 2b performed similarly with the latter producing only marginally better results due to a more conservative rejection of neighboring data points. The experiments in this work use 2a to enable a balance between point selection and the resulting prediction error. Both entropy based selection methods also resulted in improved performance. Approach 3b was too expensive even with respect to 3a for it to be applied to large scale experiments. While a combination of 1a, 2a (or 2b) and 3a proved effective in terms of prediction accuracy, the gain obtained was not justified by the increase in computational complexity (and time) as a result of using 3a. This was particularly apparent when sampling the GP model at a desired resolution; typically, this would involve performing GP regression for as many as one million points to get the final elevation map. Thus, the final choice for point selection for this work was a combination of a nearest neighbor point selection (1a) together with a distance based filtering step (2a).

F. Heteroscedastic Prediction

Equations 3 and 4 provide the mean elevation estimate and uncertainty (covariance) of GP sampled at points X_* assuming no noise in the query points. If the predictions need to be made at query points that are as uncertain as the data at hand, Equation 4 will be modified to

$$\begin{aligned}\mathrm{cov}(f_*) = \;& K(X_*, X_*) + R(X_*) - \\ & K(X_*, X)[K(X, X) + \Sigma]^{-1}\, K(X, X_*)\end{aligned} \tag{9}$$

Here $R(X_*)$ represents the noise or uncertainty of the query points themselves. Typically this is not known. For heteroscedastic GP regression, estimation of the noise hyperparameters of the data points as well as the query points is a key issue. The works [13] and [16] deal with the problem by maintaining two GP's - one to estimate the quantity of interest given the *expected* values of the noise parameters (in addition to the data sets and GP hyperparameters) and the other GP to estimate the noise hyperparameters given the data points and query points. The former is a straightforward application of Equations 3 and 9. The latter GP is the key issue as it provides the noise values to the former GP. Both [13] and [16] make an intuitive approximation - the noise values obtained from the second GP are approximated by their expected values. This work adopts the same idea but implements it differently. As the query points can be assumed to be as noisy as the training data and the fact that this work adopts a local approximation methodology towards GP regression [1], the query points are assigned a noise value that is the expected value of the noise terms of data taken from the individual data sets. The individual noise terms are learnt as before (through either joint learning or constrained optimization). Thus,

$$R(x_*) \;=\; \frac{\Sigma_{i=1}^{n} N_i \sigma_i^2}{\Sigma_{i=1}^{n} N_i}$$

where $x_* \,\epsilon\, X_*$ is a query point, N_i are the number of training data points chosen from the i^{th} data set (through the point selection step, Section III-E) and σ_i^2 is the noise variance of the (homoscedastic) GP modeling the individual data sets. An inverse distance based weighted average of the noise values could also be used in this context.

IV. Experiments

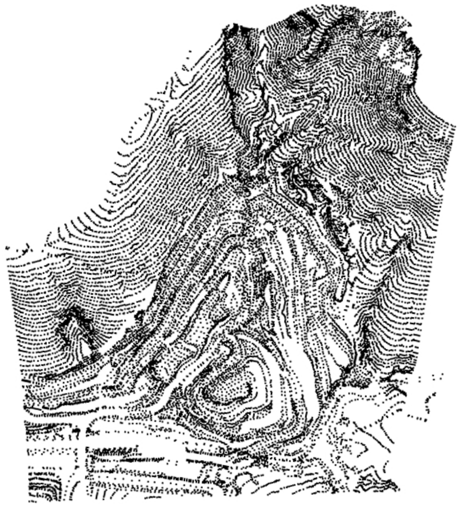

Fig. 3. GPS data: GPS based survey taken at Mt. Tom Price. It is a sparse data set consisting of 34,530 points spread over 1437.2 m x 1879.5 m x 380.5 m

Fig. 4. Laser data 1: RIEGL Laser scanner based dense scan taken at Mt. Tom Price. It consists of over 850,000 points spread over 2146.6 m x 2302.1 m x 464.3 m

Fig. 5. Laser data 2: RIEGL Laser scanner based dense scan taken at Mt. Tom Price. It consists of about 400,000 points spread over 1416.6 m x 2(

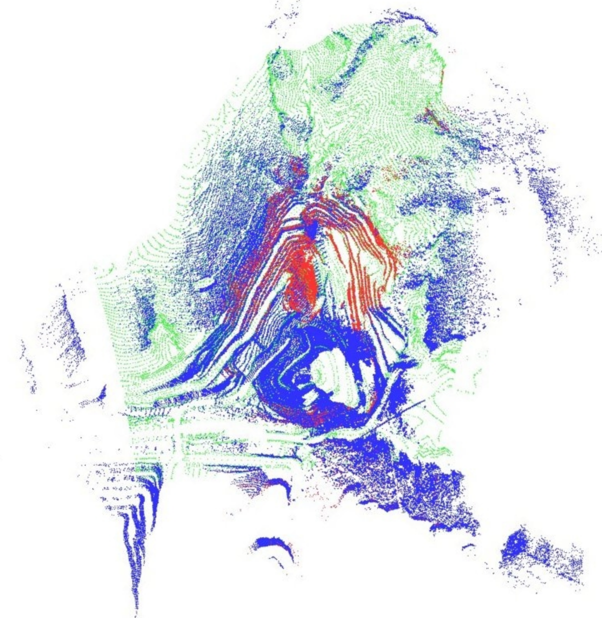

Fig. 6. The three (GPS survey and two laser scans) data sets overlaid on one another for a clearer picture of the site in consideration. The points in blue represent Laser scan 1 (Figure 4), the points in red represent the second laser scan (Figure 5) and finally, the points in green represent the GPS survey data (Figure 3).

Three kinds of experiments were performed: (1) using simulated data (fusing different noisy samples from a sine wave), (2) fusing multiple uni-sensor (RIEGL laser scanner) data sets of a mine pit and finally, (3) fusing multiple multi-sensor (RIEGL laser scanner and GPS survey) of a large mine pit. Due to paucity of space, only the last experiment is reported here.

Three data sets of the same area and of different characteristics were acquired from Mt. Tom Price mine in Western Australia. The first was a dense wide area (2146.6 m x 2302.1 m x 464.3 m) RIEGL laser scan comprising of over 850,000 points. The second was sparse GPS Survey having only about 34,530 points spread over 1437.2 m x 1879.5 m x 380.5 m. The third data set was a dense (about 400,000 points) RIEGL laser scan spread over a relatively smaller area as compared to the first scan (1416.6 m x 2003.4 m x 497.8 m). Figure 3, 4 and 5 respectively depict the data sets. Figure 6 depicts the three data sets overlaid on each other to clarify the overall picture of the terrain in consideration.

The objective was to demonstrate the benefits of GP data fusion using these data sets. The sparse GPS data is first modeled alone, then fused with the first laser data set and then the pair are fused with the third laser data set. The same test is repeated starting with the dense wide area scan and then fusing the two other data sets. In both cases the root mean squared error (RMSE) and average change of uncertainty of a randomly selected set of test points different from the training/evaluation data is evaluated after each step of the fusion process. The results of the fusion process are summarized in Tables I and II. The former describes results of the GPS-Laser fusion process whereas the latter describes the Laser-GPS fusion process. Figures 7 and 8 depict the surface map and uncertainty estimates obtained after fusing the GPS data with the two laser scanner data sets.

As shown in Table I and II, the uncertainty decreases with each successive fusion step. Thus, the required condition for fusion occurs. In Table I, the uncertainty reduction is more significant when the sparse GPS data is fused with the first dense laser scan; when the second dense laser scan is also fused, the gain in information is less than before - this is intuitive and expected. Further, it is observed that the RMSE also reduces with each fusion step. This clearly justifies the benefits of data fusion in such a context. In Table II, the RMSE of laser data 1 improves with the fusion of the GPS data because of the latter's more uniform spread. But the gain in uncertainty is more significant with the inclusion of Laser data 2 in the fusion process. This is because of the much higher density information provided by the laser scanner data set, albeit in a smaller area as compared to the first laser scanner data. Even in Table II, both the uncertainty and RMSE decrease with each successive fusion step. This clearly demonstrates the occurrence and benefits of data fusion in the context.

Note that as mentioned in Section III-D, the order in which the data sets are fused is irrelevant. The idea behind the experiments performed was to form an "intial state" based on any one data set and progrssively improve knowledge of that particular state using each other data set. The RMSE's obtained in each case depend on the points selected from the first data set - this obviously varies with different "first" data sets and hence the different values for the RMSE's in Tables I and II. However, to demonstrate the data fusion, the same set of points are evaluated after each fusion step to show a progressive improvement in the knowledge of the elevation of these points. Note also that this work only considers noise in the observations or output data and focuses on the data fusion problem. Noise in the input data (x, y) may be incorporated as demonstrated in works such as [20].

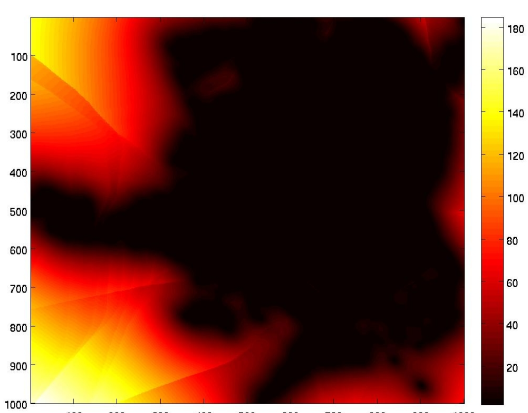

Fig. 8. Uncertainty (in meters) of the predicted elevation map obtained from the GP fusion of the GPS data and the two laser scanner data sets. Fringe areas that are not well supported by the individual data sets observe high prediction uncertainty.

V. Conclusion

Gaussian processes were used to model complex, large scale terrain. This work proposed a novel approach to fusing multiple, multi-modal data sets to obtain a comprehensive model of the terrain under consideration. The fusion technique itself is generic and applicable as a general Gaussian process fusion methodology. It is based on the underlying principles of Gaussian processes and hence is well founded. The approach proposed and demonstrated in this work treated the fusion problem as a standard Gaussian process regression problem using different noisy samples of the terrain to be modeled. Experiments conducted using real data sets obtained from GPS and Laser scanner based surveys in real application scenarios (mining) demonstrate the viability of the proposed technique. The resulting model obtained naturally provides a multi-resolution representation of large scale terrain, effectively handles both uncertainty and incompleteness in a statistically sound way, provides a powerful basis to handle correlated spatial data in an appropriate manner and finally, provides a statistically correct means of overcoming the limited perceptual capabilities of sensors used in large scale field robotics scenarios.

Acknowledgments

This work has been supported by the Rio Tinto Centre for Mine Automation and the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government. The authors acknowledge the support of Annette Pal, James Batchelor, Craig Denham, Joel Cockman and Paul Craine of Rio Tinto.

References

[1] S. Vasudevan, F. Ramos, E. Nettleton, and H. Durrant-Whyte, "Gaussian Process Modeling of Large Scale Terrain," *Journal of Field Robotics*, vol. 26(10), 2009, http://www-personal.acfr.usyd.edu.au/shrihari/svjfr09.pdf.

[2] S. Lacroix, A. Mallet, D. Bonnafous, G. Bauzil, S. Fleury, M. Herrb, and R. Chatila, "Autonomous rover navigation on unknown terrains: Functions and Integration," *International Journal of Robotics Research (IJRR)*, vol. 21(10-11), pp. 917–942, 2002.

[3] R. Triebel, P. Pfaff, and W. Burgard, "Multi-Level Surface Maps for Outdoor Terrain Mapping and Loop Closing," in *International Conference on Intelligent Robots and Systems (IROS)*, Beijing, China, October 2006.

[4] J. Leal, S. Scheding, and G. Dissanayake, "3D Mapping: A Stochastic Approach," in *Australian Conference on Robotics and Automation*, November 2001.

[5] I. Rekleitis, J. Bedwani, D. Gingras, and E. Dupuis, "Experimental Results for Over-the-Horizon Planetary exploration using a LIDAR sensor," in *Eleventh International Symposium on Experimental Robotics*, July 2008.

[6] H. Durrant-Whyte, "A Critical Review of the State-of-the-Art in Autonomous Land Vehicle Systems and Technology," Sandia National Laboratories, USA, Tech. Rep. SAND2001-3685, November 2001.

[7] I. D. Moore, R. B. Grayson, and A. R. Ladson, "Digital terrain modelling: A review of hydrological, geomorphological, and biological applications," *Hydrological Processes*, vol. 5-1, pp. 3–30, 1991.

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

[9] P. K. Kitanidis, *Introdcution to Geostatistics: Applications in Hydrogeology*. Cambridge University Press, 1997.

[10] G. Matheron, "Principles of Geostatistics," *Economic Geology*, vol. 58, pp. 1246–1266, 1963.

TABLE I

GP Fusion: Mt. Tom Price data sets (GPS - laser scanner fusion)

Fusion sequence (Data Sets)	Root Mean Squared Error (RMSE) (6000 test points) (mean $\pm$ std. dev. in meters)	Average change in Uncertainty (with respect to previous case) (std. dev. in meters)
GPS data only	3.30 $\pm$ 4.26	
GPS data & Laser data set 1	2.19 $\pm$ 2.63	-0.6106 (no cases of an increase in uncertainty)
GPS data, Laser data set 1 & Laser data set 2	2.13 $\pm$ 2.59	-0.1889 (no cases of an increase in uncertainty)

TABLE II

GP Fusion: Mt. Tom Price data sets (Laser scanner - GPS fusion)

Fusion sequence (Data Sets)	Root Mean Squared Error (RMSE) (4000 test points) (mean $\pm$ std. dev. in meters)	Average change in Uncertainty (with respect to previous case) (std. dev. in meters)
Laser data set 1 only	1.78 $\pm$ 2.29	
Laser data set 1 & GPS data	1.27 $\pm$ 2.22	-0.0264 (no cases of an increase in uncertainty)
Laser data set 1, GPS data & Laser data set 2	1.02 $\pm$ 1.58	-0.0830 (no cases of an increase in uncertainty)

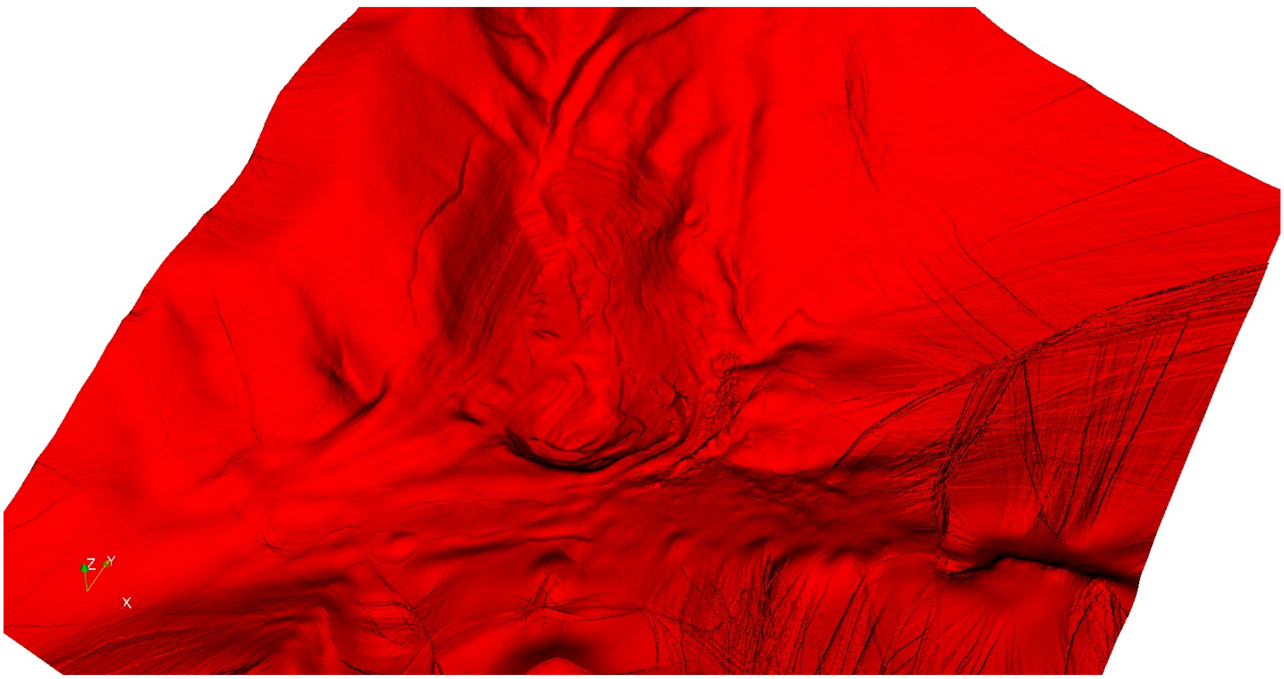

Fig. 7. Output of GP Fusion algorithm applied to the Tom Price data sets (GPS data and the two laser scanner data sets). The test data comprises of 1 million points. The surface map of the output elevation map is depicted in the image.

[11] C. Plagemann, S. Mischke, S. Prentice, K. Kersting, N. Roy, and W. Burgard, "A Bayesian regression approach to terrain mapping and an application to legged robot locomotion," *Journal of Field Robotics*, vol. 26(10), 2009.

[12] C. J. Paciorek and M. J. Schervish, "Nonstationary Covariance Functions for Gaussian Process Regression," in *Advances in Neural Information Processing Systems (NIPS) 16*, S. Thrun, L. Saul, and B. Schölkopf, Eds. Cambridge, MA: MIT Press, 2004.

[13] P. W. Goldberg, C. K. I. Williams, and C. M. Bishop, "Regression with Input-dependent Noise: A Gaussian Process Treatment," in *Advances in Neural Information Processing Systems (NIPS) 10*, M. I. Jordan, M. J. Kearns, S. A. Solla, and L. Erlbaum, Eds. Cambridge, MA: MIT Press, 1998.

[14] M. Yuan and G. Wabha, "Doubly Penalized Likelihood Estimator in Heteroscedastic Regression," Department of Statistics, University of Wisconsin, Madison, WI, USA, Tech. Rep., 2004.

[15] Q. Le, A. Smola, and S. Canu, "Heteroscedastic Gaussian Process Regression," in *International Conference on Machine Learning (ICML)*, 2005.

[16] K. Kersting, C. Plagemann, P. Pfaff, and W. Burgard, "Most Likely Hetereoscedastic Gaussian Process Regression," in *International Conference on Machine Learning (ICML)*, 2007.

[17] M. El-Beltagy and W. Wright, "Gaussian processes for model fusion," in *International Conference on Artificial Neural Networks (ICANN)*, 2001.

[18] R. Murray-Smith and B. Pearlmutter, *Deterministic and Statistical Methods in Machine Learning, LNAI 3635*. Springer-Verlag, 2005, ch. Transformations of Gaussian Process priors, pp. 110–123.

[19] N. Lawrence, M. Seeger, and R. Herbrich, "Fast Sparse Gaussian Process Methods: The Informative Vector Machine," in *Advances in Neural Information Processing Systems (NIPS) 15*, S. T. S. Becker and K. Obermayer, Eds. Cambridge, MA: MIT Press, 2003, pp. 609–616.

[20] S. O'Callaghan, F. Ramos, and H. Durrant-Whyte, "Gaussian Process Occupancy Maps Incorporating Sensor and Location Uncertainty," in *International Conference on Robotics and Automation (ICRA)*, 2010.