A Space-Carving Approach to Surface Estimation

Santosh K. Divvala, J. Andrew Bagnell, Martial Hebert
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213
{santosh,dbagnell,hebert}@ri.cmu.edu

Abstract

In this project, the problem of detecting the surface of an unknown arbitrarily-shaped scene from a set of points is considered. This problem is usually formulated as the estimation of a function whose zero set represents the surface. Most existing approaches solve this problem by employing *implicit* surface representations to recover the surface from the given data points. The proposed approach differs from existing methods as it not only uses data points but also utilizes the additional geometric *ray*-based information often ignored by earlier formulations. By posing this problem as an infinite-constraint minimization task, a maximum-margin based formulation is developed and solved using a kernel-based online subgradient method. Experimental results validate the proposed approach.

1 Introduction

In recent years, 3D geometry has gained increasing popularity as the new form of digital media content. Due to advances in sensor technology, it is now feasible to acquire highly detailed 3D scans of complex scenes to obtain millions of data points at high sampling rates over large spatial extents. This ability to acquire high-resolution depth information brings with it the possibility of using 3D geometric data to construct detailed shape models to address challenging problems in computer graphics [3] and robotics [9].

Inferring a surface from millions of points sampled with noise is a non-trivial task however, for which a variety of methods have been proposed. Simple approaches such as least squares have little practical value due to the instability of the obtained solutions. The class of implicit or level set surface representations [9, 11] is a rather large one, although other methods have also been suggested. However a major problem with these methods is that - as the complexity of the surface increases they require increasing number of parameters, far exceeding the number of available data points (which makes the fitting problem ill-posed).

In this work, a different perspective to the surface fitting problem is studied by framing it as a classification task. The problem of surface fitting to a set of points given an additional set of points labeled as interior or exterior to the surface is considered. In this setting, the surface estimation task lends itself to a max-margin based representation, allowing non-linear/non-parametric and high dimensional surfaces to be estimated using kernel functions.

In addition, this work, for the first time, proposes the use of the complete geometric *ray*-based information about the data points. Most often, when data points are sampled from a sensor, the information about the ray connecting the sampled point to its source location is implicitly available. By incorporating this information as multiple set of constraints into the above max-margin framework, an improved solution to the surface fitting problem can be achieved. Such a formulation is attractive due to the stability of the obtained solutions and the range of functional forms that could be included. It must be mentioned that the idea of utilizing ray information about the data points has been earlier used in the computer vision community in the context of space carving [3, 8], where calibrated views of a scene were used to carve a 3D space of voxels to reconstruct the actual scene.

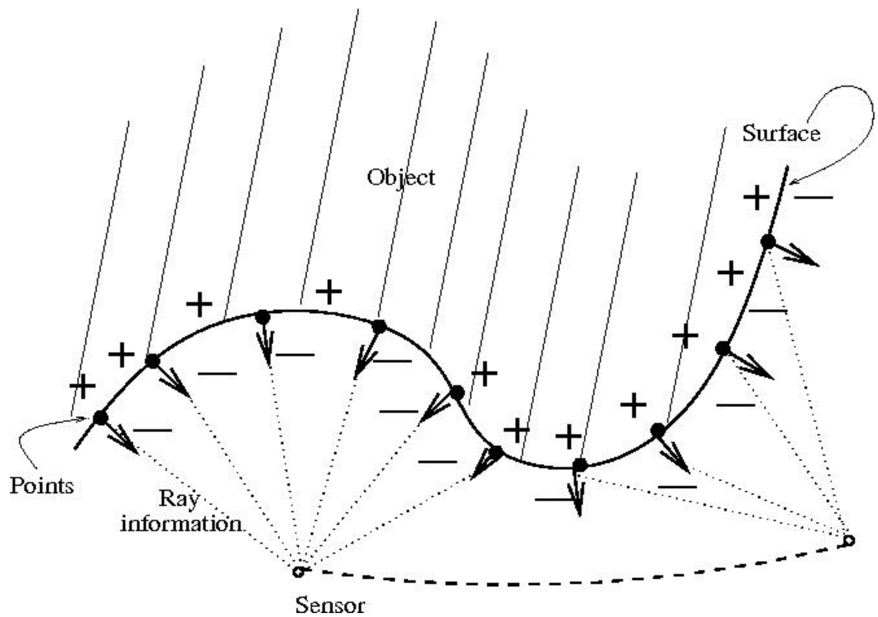

Figure 1: Space carving approach to surface estimation: By defining interior and exterior points with respect to the given data points, the surface estimation problem is formulated as a two-class data classification task.

The rest of the report is organized as follows. Section. 2 describes the proposed maximum-margin approach to surface fitting and then explains a subgradient-based method for solving it. Section. 3 presents experimental results using the proposed method. Section. 4 concludes the report.

2 Proposed Approach

In this section, we first briefly review the standard linear Support Vector Machine (SVM) classification formulation and then proceed to develop the proposed max-margin based approach to surface estimation.

Support Vector Classification The formulation of hard-margin SVM classifier is developed by considering the problem of two-class linear data classification [10]. That is, given a data-set $\{(x_i, y_i)\}_{1 \leq i \leq n}$ of training samples $x_i \in \mathcal{R}^d$ belonging to classes labeled by $y_i \in \{-1, 1\}$, the task is to find a hyper-plane that separates the two classes. The SVM approach to this problem finds the separating hyper-plane with maximum distance to x_i, while preserving the separation of the two classes. If the hyper-plane is denoted by the set $\{x \in \mathcal{R}^d | < w, x > +b = 0\}$, then finding the optimal parameters $w \in \mathcal{R}^d$, $b \in \mathcal{R}$ is equivalent to minimizing the loss function $< w, w >$ subject to $y_i(< w, x > +b) \geq 1, i = 1, ..., n$. The attractive property of this formulation is the derivation of the dual of the above primal problem results in an equivalent optimization problem but one in which the data samples appear only in the form of inner products with one another. As a result of this, the *kernel trick* [10] can be applied to find the non-linear separating hyper-surface in a high-dimensional space.

2.1 Maximum Margin based Formulation

We have a set of points of the form $\{(p_1, p_2, ..., p_n)\} \subset \mathcal{R}^d$ where each p_i is a data point lying on the hyper-surface, sampled by the sensor located at $c \in \mathcal{R}^d$. Given this set of data points, we define a set $\{(x_i, y_i)\}_{1 \leq i \leq N} \in X \times Y$ of labeled points $x_i \in \mathcal{R}^d$ with associated labels $y_i \in \{-1, 1\}$, where x_i is interior to the hyper-surface of interest if $y_i = 1$ and exterior if $y_i = -1$. The points x_i are defined along along the ray joining p_i's to c as shown in Fig. 1.

The goal is to learn the function that maps the data samples to their labels $f : X \rightarrow Y$. The zero set of this function $f(x) = 0$ corresponds to the surface that needs to be estimated. The output f of the learning algorithm is usually referred as an *hypothesis* and the set of all possible hypothesis is denoted by $\mathcal{H}$. In the following, we will assume $\mathcal{H}$ is a *reproducing kernel Hilbert space* (RKHS) [7]. This means that there exists a kernel $k : X \times X \rightarrow \mathcal{R}$ and a dot product $< ., . >_{\mathcal{H}}$ such that k has the reproducing property

$$< f, k(x, .) >_{\mathcal{H}} = f(x) \quad \forall x \in X, \tag{1}$$

and all $f \in \mathcal{H}$ are linear combinations of kernel functions. The inner product $< ., . >_{\mathcal{H}}$ induces a norm on $f \in \mathcal{H}$ as

$$||f||_{\mathcal{H}} = < f, f >_{\mathcal{H}}^{\frac{1}{2}} . \tag{2}$$

Given the above notations, we begin with the following convex optimization problem:

$$min \;\; ||f||^2_{\mathcal{H}} \;\; subject \; to \tag{3}$$

$$\begin{aligned} f(x_i) &\geq 1 \quad \text{if } y_i = 1 \\ max\; f(x_i) &\leq -1 \quad \text{if } y_i = -1. \end{aligned} \tag{4}$$

The latter constraint expresses the fact that there exists an infinite set of constraints representing the negative data points which correspond to the points lying on the ray connecting the data point p_i to the sensor c. As it would be sufficient for the closest data point to be correctly classified, this corresponds to taking the least negative of the points to be less than -1.

Adding slack variables ξ_i to (3), we have the following optimization problem:

$$min \;\; \frac{\lambda}{2}||f||^2_{\mathcal{H}} + \frac{1}{N}\sum_i \xi_i \;\; subject \; to \tag{5}$$

$$\begin{aligned} f(x_i) &\geq 1 - \xi_i^+ \quad \text{if } y_i = 1 \\ max \;\; f(x_i) &\leq -1 + \xi_i^- \quad \text{if } y_i = -1. \end{aligned} \tag{6}$$

where $\lambda \geq 0$ is a hyper-parameter that trades off constraint violations for margin maximization. Finally we note that the constraints in this convex program are tight (equality holds at the optimum) so we can place them directly into the objective. Doing so we arrive at the following unconstrained optimization problem:

$$min \;\; R(x, y, f) \;\; \equiv \;\; min \;\; \frac{\lambda}{2}||f||^2_{\mathcal{H}} + \frac{1}{N}\sum_i r(x_i, y_i, f(x_i)), \tag{7}$$

where

$$r(x_i, y_i, f(x_i)) = \begin{cases} max(0, max(f(x_i) + 1)) & \text{if } y_i = \text{-1} \\ max(0, 1 - f(x_i)) & \text{if } y_i = 1 \end{cases} \tag{8}$$

Equation (7) characterizes our entire problem of surface estimation. Recall that in doing so, we have converted an optimization problem with a convex objective function and a combination of linear and non-linear constraints (non-linear due to the *max* operator) into an equivalent implicit convex optimization problem. We could attempt to solve this problem either using a constraint generation or by a variable generation approach (*i.e.,* by taking its dual). Here we employ a kernel-based sub-gradient method to solve it, partly motivated by the fact that there have been efficient and fast algorithms developed for its computation [5].

2.2 On-line Kernel-based Sub-gradient Method

As a direct generalization of gradient descent to non-differentiable convex objective functions, the sub-gradient method iteratively computes and steps in the negative direction of a gradient-like vector known as a sub-gradient. Recall that sub-gradient is defined using a tangent-like lower bound property of the function at the point in question. The gradient is the unique sub-gradient at any point of differentiability, while there exists a continuum of sub-gradient vectors at points of non-differentiability. In the latter case, any one is chosen arbitrarily (the choice does not make any difference in theory or practice). We will in fact employ an on-line variant of the sub-gradient method as using a batch method has problems of dealing with increasing complexity with increasing number of observations and high training time.

The basic idea is to update the current solution by following the negative of the sub-gradient so as to move towards the optimum solution:

$$f_{t+1} \leftarrow f_t - \eta_t \frac{\partial R(f, x_t, y_t)}{\partial f}, \tag{9}$$

where $\eta_t > 0$ is the *learning rate*. In order to evaluate the gradient in above equation, we note that the evaluation function $f \mapsto f(x_i)$ has the property as mentioned in (1), and therefore

$$\frac{\partial r(x_t, y_t, f(x_t))}{\partial f} = r'(x_t, y_t, f(x_t))k(x_t, .), \tag{10}$$

where from (7), we have

$$r'(x_t, y_t, f(x_t)) = \begin{cases} -y_t & \text{if } y_t f_t(x_t) \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

The above equation expresses the fact that the gradient has to be considered only if a point violates the constraint. Since $\frac{\partial ||f||^2_{\mathcal{H}}}{\partial f} = 2f$, the update (9) reduces to

$$f_{t+1} \leftarrow (1 - \eta_t \lambda) f_t - \eta_t r'(x_t, y_t, f(x_t))k(x_t, .), \tag{12}$$

with the zero initial hypothesis being $f_1 = 0$. For the purpose of practical computations, one can rewrite f_t as a kernel expansion [6]

$$f_t(x) = \sum_{i=1}^{t-1} \alpha_i k(x_i, x) \quad x \in X, \tag{13}$$

where the coefficients α are given by

$$\alpha_i = \begin{cases} -\eta_t(-y_t) & \text{if } i = t \\ (1 - \eta_t \lambda)\alpha_i & \text{if } 1 \leq i < t \end{cases} \tag{14}$$

In summary, following from the above derivation, we have the overall algorithm as described in Algorithm. 1. Theoretical analysis, as described in [5], shows that this simple intuitively appealing sub-gradient algorithm performs well for several learning problems.

Given: Dataset $(x_i, y_i)_{i \in N}$, regularization parameter λ, learning rate η_t, max iterations T
Initialize: $t = 1$
repeat
    Pick any point i randomly
    $f_t(.) = \sum_{j=1}^{t-1} \alpha_j (1 - \eta_t \lambda)^j k(x_j, .)$
    $x_i^* = max(1 - y_i f_t(x_i))$
    $\alpha_t = -\eta_t r'(x_i^*, y_i^*, f_t(x_i^*))$
    $\alpha_j = (1 - \eta_t \lambda)\alpha_j$ for $j \in [1, t-1]$
    $t = t + 1$
until ($t < T \,||\, (\alpha_{t-k} : \alpha_t) < \epsilon$)
return **f** $= (f_1, f_2, ...)$

Algorithm 1: On-line Kernel-based Sub-gradient Method for Surface Estimation

3 Experimental Results and Analysis

In this section, we present experimental results on simulated as well as real 3D laser and camera data to evaluate the performance of the proposed approach. In our experiments, we employed the exponential Radial Basis Function (RBF) kernel

$$k(x, x') = e^{(-\frac{||x-x'||^2}{2\sigma^2})}, \tag{15}$$

which is the most commonly used kernel. The advantage of this kernel is that it can project the data into an infinite dimensional space and thus can deal with arbitrarily complex structures. In all the cases, the data-sets were normalized and a fixed bandwidth of $\sigma = 0.25$ was used. The value of λ was set to $\frac{1}{200N}$ where N is the size of the dataset and η_t was set to $\frac{1}{2\sqrt{t}}$. The on-line sub-gradient algorithm was terminated when the value of the risk functional $R(.)$ was less than a small threshold value. The points lying on the interior and the exterior were generated as $(1+\gamma)p_j - \gamma c$, $(1-\gamma)p_j + \gamma c$ for a small γ value respectively. The computation of the sub-gradient requires solving (4), which finds the most violated point amongst the continuum of points lying along the ray for each data sample. Note that this essentially boils down to performing a line search on the ray $\beta x_i + (1-\beta)c$ with $\beta \in [0, 1]$, where x_i denotes the data point and c is the sensor location. The line search was performed using MATLAB's `fmincon` command.

Simulated Data To validate the proposed formulation and solution, we first evaluated the algorithm on a simple simulated dataset consisting of points lying on a circle as shown in Fig. 2. In the figure, the interior (positive) examples are displayed using '+' symbol while the exterior (negative) samples are displayed using '*'. Notice that using the above proposed max-margin formulation and on-line sub-gradient solution, we are able to estimate the separating hyperplane (zero set).

Although a good surface estimate is obtained, we observe that the inclusion of the ray-based information is not quite useful in this case. To demonstrate the utility of introducing the ray-based constraints, the experiment was repeated on a more complex data setting as shown in Fig. 3. Observe that the solution obtained without using ray-based constraints is erroneous due to the extraneous peaks (To confirm this erroneous behavior, the same experiment was performed using [2] as shown in Fig. 3(b)). However by introducing the ray-constraints, we get rid of the spurious bumps of the estimated surface.

Laser Data In this experiment, the laser data collected by a mobile robot exploring the CMU campus was used (See Fig. 4). The points on the interior (positive examples) and

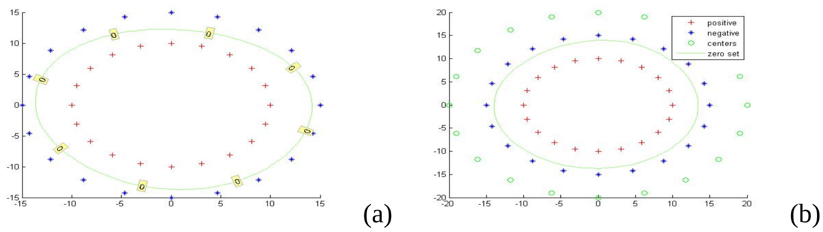

Figure 2: Validation of the proposed max-margin formulation and sub-gradient solution on a simple dataset (a) Zero-set estimated without using ray-information (b) With using ray-information.

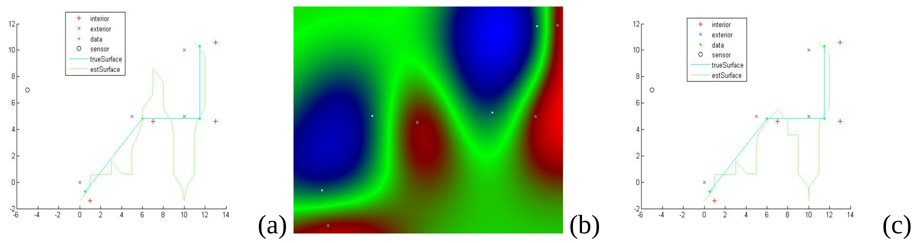

Figure 3: Introducing ray-based constraints helps in achieving an improved surface fit (a) Zero-set estimated without using ray-information (b) Similar result obtained using [2] (c) With using ray-information. Notice that when the ray-constraints are introduced the extraneous bump of the surface is eliminated.

exterior (negative examples) were generated as described earlier. The resultant surface fit obtained after running the proposed approach is shown in Fig. 4(c,d). It can be observed that the ray-based constraints help in achieving a better surface fit.

Camera (Stereo) Data The final experiment was performed on a set of $3D$ points reconstructed using a calibrated stereo image pair [1] as shown in Fig. 5. This dataset is quite huge and dense. For computational purposes, a smaller subset of points randomly sampled from the original dataset were used as the data points for this experiment (To deal with larger datasets, an efficient implementation involving kd-trees [4] will be required. The running time for a dataset of 100 points is approximately one minute). Notice that a decent surface fit is obtained by using a sparse subset of points (around less than 0.01% of the original data points was used!). Additionally, the use of ray-based constraints helps in improving the fit surface.

4 Conclusion

The problem of detecting the surface of an unknown arbitrarily-shaped scene from a point cloud is an important task for several applications. In this project, a new formulation based on a max-margin framework was proposed. Also a solution based on on-line kernel-based sub-gradient method was described. Experimental results validate the proposed formulation and demonstrate the superiority of the results upon inclusion of the ray-based constraints.

Acknowledgments Thanks to Ranjith Unnikrishnan for providing access to the laser data used in the experiments.

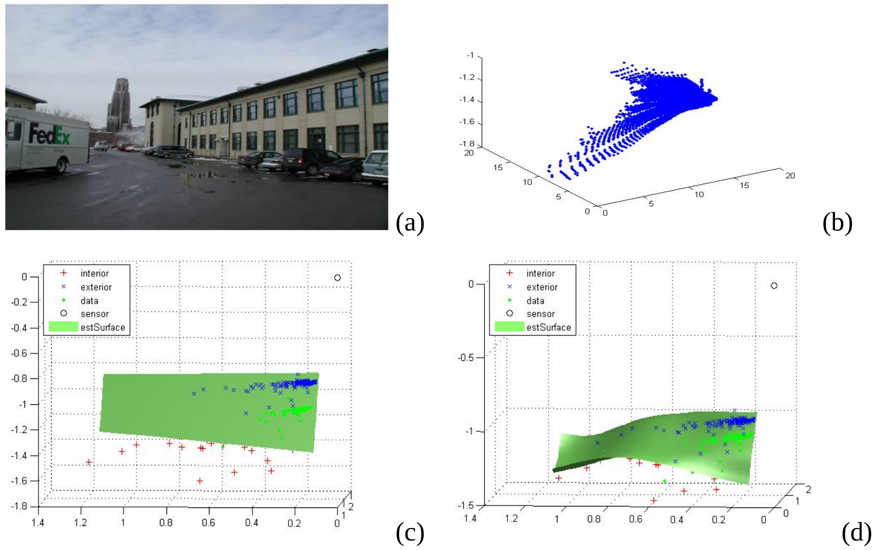

Figure 4: Experiment using 3D laser data collected by a mobile robot exploring CMU campus (a-b) Scene and the Laser Data (c) Surface fit without using ray information (d) Improved result upon incorporation of ray-based constraints

References

[1] http://vision.middlebury.edu/stereo/data/.

[2] http://www.site.uottawa.ca/ gcaron/svmapplet/svmapplet.html.

[3] K. Kutulakos and S. Seitz. A theory of shape by space carving. *International Conference on Computer Vision*, 1999.

[4] A. Moore. *Efficient memory-based learning for robot control*. PhD thesis, University of Cambridge, 1990.

[5] N. D. Ratliff, J. A. Bagnell, and M. A. Zinkevich. (Online) Subgradient methods for structured prediction. *Artificial Intelligence and Statistics*, 2007.

[6] B. Scholkopf, R. Herbrich, and A. J. Smola. A generalized representer theorem. *Annual Conference on Computational Learning Theory*, 2001.

[7] B. Scholkopf and A. J. Smola. *Learning with kernels*. MIT Press, Cambridge, 2001.

[8] S. Seitz and C. Dyer. Photorealistic scene reconstruction by voxel coloring. *Computer Vision and Pattern Recognition*, 1997.

[9] R. Unnikrishnan. Statistical approaches to multi-scale point cloud processing. Technical report, Carnegie Mellon University, USA, 2008.

[10] V. N. Vapnik. *The Nature of Statistical Learning Theory*. Springer Verlag, 1995.

[11] C. Walder, B. Scholkopf, and . Chapelle. Implicit surface modelling with a globally regularised basis of compact support. Technical report, Max Planck Institute for Biological Cybernetics, Germany, 2006.

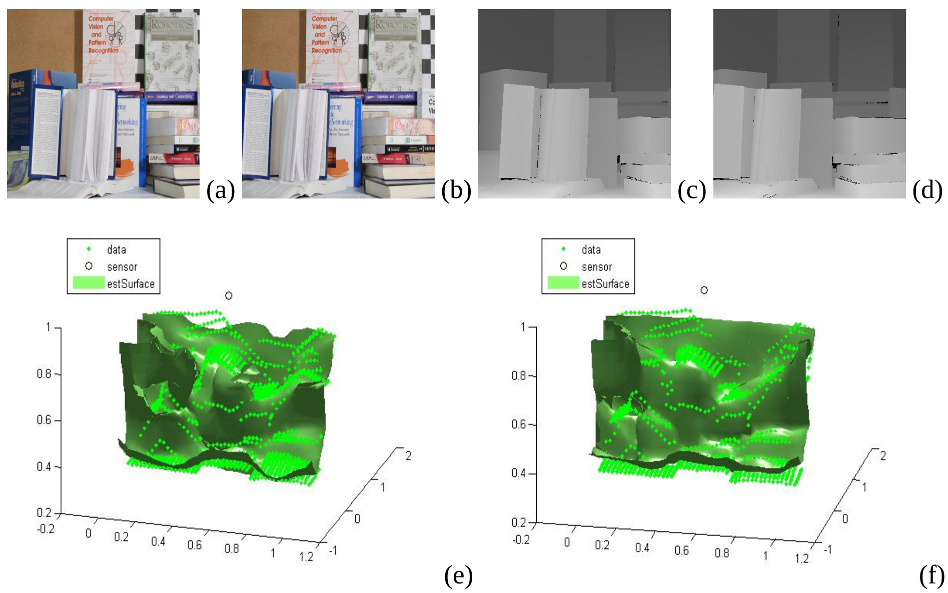

Figure 5: Experiment using camera (stereo) data (a-b) Images (c-d) Dense depth maps (e) Surface fit without using ray information (f) Improved result upon incorporation of ray-based constraints