To appear in IEEE CVPR 2013
Portland, Oregon

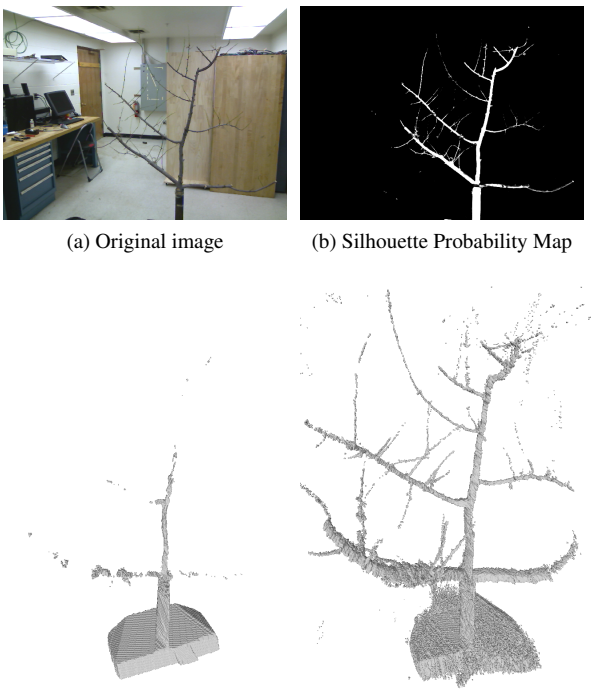
Shape from Silhouette Probability Maps: reconstruction of thin objects in the presence of silhouette extraction and calibration error

Amy Tabb
USDA-ARS-AFRS and Purdue University
Kearneysville, West Virginia and West Lafayette, Indiana
amy.tabb@ars.usda.gov

Abstract

This paper considers the problem of reconstructing the shape of thin, texture-less objects such as leafless trees when there is noise or deterministic error in the silhouette extraction step or there are small errors in camera calibration. Traditional intersection-based techniques such as the visual hull are not robust to error because they penalize false negative and false positive error unequally. We provide a voxel-based formalism that penalizes false negative and positive error equally, by casting the reconstruction problem as a pseudo-Boolean minimization problem, where voxels are the variables of a pseudo-Boolean function and are labeled occupied or empty. Since the pseudo-Boolean minimization problem is NP-Hard for nonsubmodular functions, we developed an algorithm for an approximate solution using local minimum search. Our algorithm treats input binary probability maps (in other words, silhouettes) or continuously-valued probability maps identically, and places no constraints on camera placement.

The algorithm was tested on three different leafless trees and one metal object where the number of voxels is 54.4 million (voxel sides measure 3.6 mm). Results show that our approach reconstructs the complicated branching structure of thin, texture-less objects in the presence of error where intersection-based approaches currently fail.[1]



(a) Original image (b) Silhouette Probability Map

(c) Visual Hull reconstruction (d) Our reconstruction

Figure 1. **[Best viewed color.]** The goal of this paper is to reconstruct the shape of leafless trees in the presence of silhouette extraction and camera calibration error. **Left**: VH reconstruction. **Right**: reconstruction using our algorithm. Results throughout this document are displayed with Meshlab (http://meshlab.sourceforge.net, developed with support from the 3D-CoForm project).

1. Introduction

The reconstruction of thin, texture-less objects such as leafless trees is a necessary step for agricultural applications such as robotic pruning. In these applications, the exact shape of the object is not known a priori, but it is known that the objects do not contain concavities.[2] For these reasons, we pursue a Shape from Silhouette (SfS) method for reconstructing thin, texture-less objects from images.

The visual hull (VH) is a type of SfS reconstruction generated by intersecting the backprojected silhouette view-

[1] Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer.

[2] A concavity in the 3D sense is an egg-shaped depression in the surface of the object.

ing cones of N cameras [16]. The VH reconstruction can be implemented in fast algorithms for either voxel-based or polyhedral-based representations ([5],[7],[9],[17],[19],[20]), and allows the use of silhouettes when other features are not available or reliable.

In real-world applications, silhouette or camera calibration error is often present, and the VH approach is ill-equipped to deal with this error, particularly false negatives (false negatives are pixels that are marked background but represent the target object). Reconstructions of thin objects are particularly sensitive to the effects of silhouette errors, whether from silhouette extraction errors, noise, or camera calibration error, because the two-dimensional projections of thin objects may be only a few pixels wide. As a result, small errors in silhouette extraction can have large effects on the accuracy of a VH reconstruction. In addition, voxel and pixel resolution settings, even when silhouettes are accurate, can produce VH reconstructions that fail to reconstruct many portions of thin objects.

When a reconstruction resembles the original object, we call it a *representative reconstruction*. Given these preliminaries, a statement of our problem is: assuming silhouette and camera calibration error is present, generate representative reconstructions of thin, texture-less objects from silhouette images. We consider a solution to this problem to be the final step of a three-dimensional reconstruction system, as the various characterisitics of the objects we consider means that photoconsistency approaches will recover little useful information about the shape of the object.

The set of input silhouettes may take one of two different forms. In the first, silhouette images are binary-valued. We call this problem the Shape from Inconsistent Silhouette (SfIS) problem.[3] The second is Shape from Silhouette Probability Maps (SfSPM); instead of binary silhouettes, silhouette probability maps (SPM) are given as input, where the probalities that the pixels represent the object are continuously-valued. SPMs are used to avoid making early committments to one label or the other. In this paper, SfIS is considered a special case of the SfSPM problem, where the probabilites are restricted to binary values.

Our approach to SfSPM is to penalize false positive and false negative SPM error equally. To that end, we give a pseudo-Boolean error function ($f : \mathbb{B}^n \mapsto \mathbb{R}$, where $\mathbb{B} = \{0, 1\}$ and $\mathbb{R}$ denotes the set of real numbers) that characterizes the match between the SPMs and the reconstruction as the pixel-by-pixel differences between the SPMs and the image of the reconstruction. This error function penalizes false positive and false negative error equally, unlike the VH approach.

The error function is non-submodular, and to minimize a non-submodular pseudo-Boolean function is NP-Hard (unless P=NP).[4] Consequently, we focus on local minimum search methods to find representative reconstructions from SPMs, and describe a local minimum search algorithm that uses heuristics developed for SfSPM. Experimental results on three different trees and one metal object are shown in §4. These results show that our solution is a great improvement over the existing intersection-based approaches, as it reconstructs the complicated branching structure of trees, even under silhouette extraction and camera calibration error.

Our contributions to the state-of-the-art on SfIS, SfSPM, and the reconstruction of thin, textureless objects are as follows:

1. The formulation of the SfSPM (and by extension, SfIS) problem as a pseudo-Boolean optimization problem where false negative and false positive error is equally weighted.

2. Introduce local minimum search algorithms of pseudo-Boolean optimization to the SfSPM problem and show how heuristics developed for SfSPM allow for lower values of the error function to be found.

3. A reconstruction method that produces representative reconstructions of thin, texture-less objects in the presence of silhouette extraction and camera calibration error.

1.1. Related work

Recent works on SfIS and SfSPM have sought to compensate for the problems of VH approaches by delaying decisions about a voxel's label until more information about the voxel can be gained. All of these works use a voxel-based representation for the reconstruction. They can be divided into three main categories: sensor fusion, probabilistic, and minimization of silhouette inconsistency approaches.

In the sensor fusion approach, an observation is represented using a sensor model, and then the model information is fused to determine voxel occupancy probabilities. Franco and Boyer [8] used a forward sensor model for each pixel in order to jointly infer voxel occupancy probabilities from all pixels. Díaz-Más et al [6] fused sensor uncertainty models using Dempster-Shafer theory. Guan et al [11] detected static occluding objects by using a Bayesian sensor formulation for voxels, and later, fusing the information to determine the location of occluders.

The next category concerns probabilistic methods. Cheung et al [4] proposed a projection test called SPOT, or Sparse Pixel Occupancy Test, which attempts to increase speed and reduce the effects of segmentation noise. Landabaso et al [15] extended the ideas of [4] concerning projection tests with the Sampled Projection Test with a different

[3] What we call SfIS is the same as Landabaso et al's SfS-IS (Shape from Silhouette with Inconsistent Silhouettes) [15].

[4] See §2.1 for a more in-depth discussion.

error function. Also, in [15], the authors propose a method called the *unbiased hull* that compensates for silhouette error by classifying voxels that backproject to a minimum number of silhouettes $T^\star$ as occupied (while taking into account occlusion by the visual hull). $T^\star$ is found by minimizing an error function that describes the probability of voxel misclassification.

Our work is most closely aligned with that of Haro and Pardás [13]. Their approach is to minimize an approximation of the silhouette inconsistency error, where the function used depends on whether the SfIS or SfSPM problem is being considered. They consider voxels as continuously-valued variables in the range of 0 and 1, and seek a minimum by gradient descent. Binary voxel labels are determined by a threshold. In comparison, our approach uses a closed-form, exact silhouette inconsistency error function which is identical for SfIS and SfSPM and considers voxel labels as Boolean during the minimization process, and is not dependent on the setting of parameters or thresholds.

2. Formulation of the pseudo-Boolean error function

We represent the difference between SMPs and images of any reconstruction as a closed-form pseudo-Boolean function.

The pseudo-Boolean error function is formulated over a set of voxels $\mathbf{x} = \{x_0, x_1, ..., x_{n-1}\}$. A voxel is *empty* if its label is 1, and *occupied* if its label is 0. Note that this labeling is the reverse of most other literature on this subject, and is only done to make the function simpler. For instance, if we had labeled empty voxels 0 and occupied voxels 1, all of the variables (x_i) would have been negated ($\bar{x}_i$). By simply switching the labels, the formulation of the error function is more straightforward as it consists exclusively of non-negated variables.

Let an individual pixel in a SPM be p_i. r_i is the value of the reconstruction image pixel at the same location as p_i. Since the labels of occupied and empty are reversed for voxels, we also reverse the usual labeling for pixels. Consequently, the value of p_i represents the probability that p_i is viewing background. For example, if $p_i = 0$, then p_i is viewing the object with one hundrend precent probability $P(p_i = object) = 1$, and if $r_i = 0$, then r_i back-projects to the reconstruction where voxels are labeled occupied.[5]

Given voxel labeling $\mathbf{x}$, the value of reconstruction pixels can be found. Let S_{p_i} be the set of voxels that are intersected by a viewing ray from pixel p_i. Then the label of reconstruction image pixel r_i is

$$r_i = \prod_{x_a \in S_{p_i}} x_a \tag{1}$$

In other words, $r_i = 1$, representing background, only if all voxels viewed by p_i are empty.

The Silhouette Inconsistency Error (SIE) function represents the differences between reconstruction images and silhouette probability maps. For SPMs and reconstruction image pair of pixels p_i and r_i,

$$SIE(p_i, r_i) = |p_i - r_i| \tag{2}$$

Before presenting the general formula for the SfSPM problem, we will first represent $|p_i - r_i|$ as a pseudo-Boolean function by considering the two cases of the SfIS problem: when p_i is 0 (a silhouette pixel) or when p_i is 1 (a non-silhouette pixel), and substituting for r_i as in Eq 1:

$$SIE(p_i, \mathbf{x}) = \begin{cases} \prod_{x_a \in S_{p_i}} x_a & p_i = 0, \\ 1 - \prod_{x_a \in S_{p_i}} x_a & p_i = 1 \end{cases} \tag{3}$$

We now represent $|p_i - r_i|$ as a pseudo-Boolean function for the general case that p_i is continuously-valued ($p_i \in [0, 1]$):

$$SIE(p_i, \mathbf{x}) = (1 - p_i) \prod_{x_a \in S_{p_i}} x_a + p_i(1 - \prod_{x_a \in S_{p_i}} x_a) \tag{4}$$

$$SIE(p_i, \mathbf{x}) = p_i - (1 - 2p_i) \prod_{x_a \in S_{p_i}} x_a \tag{5}$$

For the special case that $p_i \in \{0, 1\}$, Eq. 5 is equivalent to Eq. 3.

SIE for a set of input images $\mathbb{I}$ is the sum of the SIE error of the individual pixels as in Eq 6.

$$SIE(\mathbb{I}, \mathbf{x}) = \sum_{p_i \in \mathbb{I}} SIE(p_i, \mathbf{x}) \tag{6}$$

As mentioned in §1, the SIE portion of the cost function treats false positive and false negative errors equally. We will now show how traditional intersection-based approaches treat these types of error, with reference to our error function SIE, as applied to the SfIS problem.

First we split $SIE(\mathbb{I}, \mathbf{x})$ into two parts: the false positive error (FP) and the false negative error (FN). If a pixel p_i is part of the silhouette (0) and the reconstruction image pixel r_i is 1, then p_i is a false positive. The opposite case is a false negative. Then the false positive and false negative error is

[5] Pixel probabilities are binary for the SfS-IS problem, and continuously-valued for the SfSPM problem.

$$FP(\mathbb{I}, \mathbf{x}) = \sum_{p_i \in \mathbb{I}, p_i = 0} SIE(p_i, \mathbf{x}) \tag{7}$$

$$FN(\mathbb{I}, \mathbf{x}) = \sum_{p_i \in \mathbb{I}, p_i = 1} SIE(p_i, \mathbf{x}) \tag{8}$$

$$SIE(\mathbb{I}, \mathbf{x}) = FP(\mathbb{I}, \mathbf{x}) + FN(\mathbb{I}, \mathbf{x}) \tag{9}$$

In the VH approach, the false negative error is zero: all non-silhouette pixels project to empty voxels. We can conclude that the VH approach minimizes false negative error, setting it to zero, while ignoring false positive error. We can represent the VH approach in terms of a pseudo-Boolean function as follows, where M is a very large constant, such as the number of pixels in all images.

$$SIEvh(\mathbb{I}, \mathbf{x}) = FP(\mathbb{I}, \mathbf{x}) + M \cdot FN(\mathbb{I}, \mathbf{x}) \tag{10}$$

The global minimum of $SIEvh$ is the VH reconstruction, where false positive and false negative errors are unequally weighted. As a result, false negative pixels have a disproportionally large impact on the VH reconstruction as compared to false positive pixels.

2.1. Complexity of minimizing SIE

Provided that there is at least one pixel in the SPMs where $P(p_i = object) > P(p_i = background)$, SIE is nonsubmodular. Finding the global minimum of nonsubmodular pseudo-Boolean functions is a NP-Hard problem ([3]). We mention that the global minimum of submodular pseudo-Boolean functions can be found in polynomial time by using graph cut methods, such as the algorithm by Kolmogorov and Zabin [14] or QPBO [12]; QPBO may also find some labels for variables of nonsubmodular functions when the function is quadratic. Even after reduction to a quadratic pseudo-Boolean function, graph cut methods such as QPBO [12] are unable to label any voxel for SIE in Eq. 6, even when there is a small number of voxels, such as 50. For these reasons, in the next section we describe an approximation solution to $\min_{\mathbf{x}} SIE(\mathbf{x})$ that finds a local minimum given an initial voxel labeling.

3. Local minimim search

The search from a local minimum is a method borrowed from the optimization community. A more in-depth discussion can be found in [3], which is the source of our discussion on the topic.

First we begin with a definition of a local minimum for a pseudo-Boolean function f. For a labeling of voxels $\mathbf{x}$, there is a neighborhood $\mathcal{N}$ of other labelings $\mathbf{y}$, where $\mathbf{y}$ is equal to $\mathbf{x}$ except that the label of one voxel differs between the two labelings. A particular labeling $\mathbf{x}$ is called a local minimum if there are no other labelings $\mathbf{y}$ in the neighborhood of $\mathbf{x}$ that have a lower value of f than $\mathbf{x}$ does. In other words, $\mathbf{x}$ is a local minimum if and only if $f(\mathbf{x}) \leq f(\mathbf{y}) \quad \forall \mathbf{y} \in \mathcal{N}(\mathbf{x})$.

This property of local minima can be stated in terms of partial first derivatives as follows. Let $\frac{\partial f}{\partial x_i}$ be the partial first derivative of f with respect to x_i. Then, $\mathbf{x}$ is a local minimum if and only if for each voxel x_i of $\mathbf{x}$ the following is satisfied:

$$x_i = \begin{cases} 1 & \text{if } \frac{\partial f}{\partial x_i}(\mathbf{x}) \leq 0 \\ 0 & \text{if } \frac{\partial f}{\partial x_i}(\mathbf{x}) \geq 0 \end{cases} \tag{11}$$

In order to find a local minimum given an initial labeling $\mathbf{x}^{(0)}$, the labels of individual voxels are changed until Eq 11 is satisfied for all voxels. This process to find a local minimum $\mathbf{x}_{min}$ for f, when $f = SIE$, is presented in pseudocode by algorithm LOCAL-MIN-SEARCH, Alg 1.

Algorithm 1 LOCAL-MIN-SEARCH($f(\cdot)$), $\mathbf{x}^{(0)}$, c)

Require: $c = f(\mathbf{x}^{(0)})$
$n =$ number of voxels
$k = 0$
while $\mathbf{x}^{(k)}$ is not a local minimum **do**
  for all $i = 0$ to n **do**
    $condition_0 = \frac{\partial f}{\partial x_i}(\mathbf{x}^{(k)}) < 0$ and $x_i^{(k)} = 0$
    $condition_1 = \frac{\partial f}{\partial x_i}(\mathbf{x}^{(k)}) > 0$ and $x_i^{(k)} = 1$
    if $condition_0$ or $condition_1$ **then**
      $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$
      $x_i^{(k+1)} = \neg x_i^{(k)}$
      $c = c - |\frac{\partial f}{\partial x_i}(\mathbf{x}^{(k)})|$
      $k = k + 1$
    end if
  end for
end while
return local minimum $\mathbf{x}^{(k)}$, value of f at $\mathbf{x}^{(k)}$, c

Local minimum search is similar to Iterated Conditioning Modes, or ICM, which has been used to minimize energy functions in vision ([2], [10], and comparison to other techniques for energy minimization [21]).

3.1. Local minimum searches for SfSPM

For any value c, where $SIE(\mathbf{x}) = c$, there are many different labelings $\mathbf{y}$ such that $SIE(\mathbf{y}) = c$. This observation is similar to that of the traditional VH theory as presented by Laurentini [16]. In the VH theory, there are many labelings of voxels that are silhouette consistent. However, the VH is chosen to be the labeling with the greatest volume. In our alteration of the local minima search, we also specify that any local minima $\mathbf{x}_{min}$ where $SIE(\mathbf{x}_{min}) = c$ have the greatest volume labeling out of all labelings $\mathbf{y}$ where

$SIE(\mathbf{y}) = c$. We refer to this as the *greatest volume property*.

We require that the local minimum labeling $\mathbf{x}_{min}$ has the greatest volume property because of the following situation, as shown in Fig 2a. In that figure, Camera 0 has no silhouette error, as silhouette pixels from Camera 0 project to at least one of the three occupied voxels. Camera 1 has some false negative error from voxel v. If voxel v is removed as in Fig 2b, the false negative error for Camera 1 will decrease by e_1, but the false positive error for Camera 0 will increase by e_0, $e_0 > e_1$. Because of this type of deadlock, voxel v will never be removed during algorithm LOCAL-MIN-SEARCH.

Since the local minimum search is done iteratively, thin protrusions and isolated voxels in the local minimum, like those in Fig. 2a are common, and the local minimum search would frequently stall on these types of labelings. However, by altering the local minimum search to require that local minima have the greatest volume property, it is possible to avoid getting trapped in minima with high values of SIE.

We will illustrate this process in Fig. 2c. Here, the voxel labeling has the same value of SIE, c, as in Fig. 2a, though Fig. 2c represents the maximal labeling for $SIE(\mathbf{x}) = c$. Then when we test whether or not to change voxel v's label to empty, we can see in Fig. 2d that the value of SIE decreases by e_1, since the false negative error is removed for Camera 1 and Camera 0 has no error.

To alter LOCAL-MIN-SEARCH for SfSPM, we simply change the conditions on line 6 so that voxels with $\frac{\partial f}{\partial x_i}(\mathbf{x}^{(k)}) \geq 0$ and $x_i^{(k)} = 1$ will have their labels changed. We call this altered algorithm LOCAL-MIN-SEARCH-SFSPM.

4. Experiments

4.1. Implementation details

We have implemented LOCAL-MIN-SEARCH-SFSPM in C++, on a machine with 8 Intel Xeon processors at 2.4 GHz and 36 GB RAM. The success of many optimization search methods depends heavily on the starting point of the search, and we found that LOCAL-MIN-SEARCH-SFSPM is sensitive to the initial labeling $\mathbf{x}^{(0)}$. In our implementation, we set $\mathbf{x}^{(0)}$ to the visual hull labeling. Algorithm LOCAL-MIN-SEARCH-SFSPM will also return different results depending on the order in which voxels are tested (Alg. 1, line 4 has the voxel indices tested in increasing order). We found that randomizing the voxel indices for testing on every iteration (at the **while** loop on line 3 of Alg. 1) produced better results with lower values of SIE than a fixed ordering.

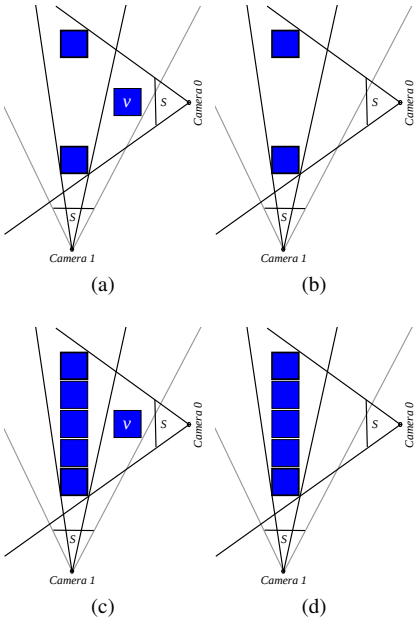


Figure 2. **[Best viewed in color]** Example in 2D illustrating the need for an altered local minimum search for SfSPM. There are two 1D cameras, Camera 0 and Camera 1. The regions of both cameras that represent silhouette regions are denoted with an S, and the viewing rays on the boundary of S are black lines. Camera 0's entire image consists of silhouette pixels, while Camera 1 has two non-silhouette regions, one either side of a central silhouette region. Gray lines represent the boundary of Camera 1.

4.2. Experiments on trees and thin metal object

Our tests were conducted on three different leafless trees with different branching characteristics. The trees were removed from a research orchard plot, defoliated for use in the tests, and mounted in a stand. We also experimented with a less complicated object: a metal pole with a copper coil suspended from the pole. All of these experiments used the same camera configuration, which consisted of 10 low-cost webcameras and 20 industrial cameras, both with image size 640x480 pixels, mounted on one wall and ceiling and pointed toward the area where the trees and metal object would be placed. The webcams had a larger field of view than the industrial cameras, and inspecting the camera calibration matrices for both types of cameras showed that the focal length in terms of pixel dimensions for the webcameras was half that of the industrial cameras. Because the size of the tree was large (roughly 2 meters high by 2 meters wide), many cameras only viewed a portion of the tree.

We acquired silhouettes and silhouette probability maps through background subtraction (reviews and evaluations are given in [1] and [18]). We acquired 50 images of the scene without the objects and modeled each pixel as one Gaussian distribution. Then, we compared images acquired

with the object in the scene and the background model, then assigned a probability that each pixel represented the object. For the silhouette probability maps, we discretized the continuous range $[0, 1]$ into 256 values for ease in displaying the silhouette probability maps as unsigned 8-bit integer images. The silhouette probability maps were thresholded at 0.5 to produce silhouettes for the experiments. In this manner, we were able to compare the results of our algorithm for the SfIS problem by using the silhouettes and the SfSPM problem by using the silhouette probability maps, when the silhouettes and silhouette probability maps are derived from the same background model.

Since we used a background subtraction method to extract silhouettes, silhouette error resulted from the following: the target object matched the background color, shadow effects, thin object regions relative to pixel size, and image sensor noise. Camera calibration was performed using a custom-made calibration rig with seven different calibration patterns so that all cameras could be calibrated to the same world coordinate frame. However, the calibration error was dependent on the camera's location with respect to the calibration rig and the images used for calibration, and some cameras did have higher calibration error than others (average error for external parameters ranged from 0.2-10 pixels, and the webcameras had greater calibration error than the industrial cameras, as was expected). While this error was small, since the objects to be tested were quite thin, the camera calibration error did have a large effect on the quality of a visual hull reconstruction.

The number of voxels in SIE was 54.4 million (voxel sides measured 3.6 mm) and the number of terms was 8.6 million. The degree of SIE was 975 for all of the tests. The LOCAL-MIN-SEARCH-SFSPM algorithm was implemented in parallel and run times ranged from 16-40 minutes, depending on the dataset. The number of iterations for the **while** loop at line 3 of LOCAL-MIN-SEARCH was 14-16, depending on the dataset.

4.3. Evaluation and Discussion

Results for the four datasets are shown in Figs. 3 - 6. The reconstructions using our approach for both the SfIS and SfSPM problems offer a great improvement over the traditional visual hull approach. Our approach reconstructs very complicated branching structure in the presence of small camera calibration error and silhouette extraction error. A negative aspect of our approach is the small noisy regions in the reconstruction, which could be removed with some post-processing techniques.[6]

While the values of SIE may differ by $100,000$ or more for the SfIS versus SfSPM problems, we found that the reconstructions are largely the same. We hypothesize that the small differences between the SfIS and SfSPM reconstructions are an artifact of the random order for voxel testing (as mentioned §4.1) rather than of using binary versus continuously-valued silhouette probability maps.

We compared our approach to SPOT [4] and Unbiased Hull from Landabaso et al [15], where the choice of projection test was SPOT. While SPOT works well for noisy silhouettes of thicker objects such as humans, SPOT does not reconstruct branching structure. The unbiased hull approach was able to reconstruct more of the large-scale features of the tree than SPOT, but it introduced artifacts near the boundaries of image viewing regions and misses small-scale features of branches. In the absence of occlusion by the visual hull, which is typical in our datasets, the unbiased hull approach classifies voxels that backproject to a minimum number of silhouettes $T^\star$ as occupied; $T^\star$ is selected by exhaustively evaluating the cost function for each possible value of $T^\star$. Larger-scale details exert a greater influence over the value of the cost function than fine details in the unbiased hull approach, so larger branches are reconstructed and smaller branches are not. Our reconstructions are more representative of the original object than the intersection-based comparison methods because our method does not assume that one threshold, such as $T^\star$, will be applicable to all voxels.

5. Closing remarks

We presented an algorithm for reconstructing thin, textureless objects from silhouette probability maps or silhouettes by formulating the difference between input and reconstruction images as a pseudo-Boolean minimization problem. We were able to reconstruct the objects with greater fidelity than the tradiional visual hull technique. Our approach is flexible in that it is not dependent on the number of cameras or the characteristics of those cameras. As a result, we were able to successfully use a mixture of high-cost and low-cost cameras with varying degrees of accuracy for reconstruction.

References

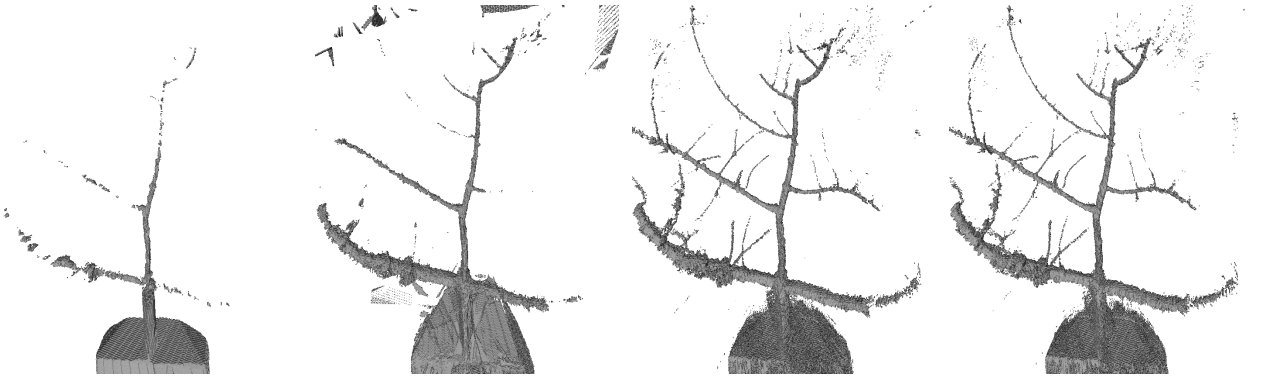
[1] Y. Benezeth, P. Jodoin, B. Emile, H. Laurent, and C. Rosenberger. Review and evaluation of commonly-implemented background subtraction algorithms. In *ICPR 2008: 19th International Conference on Pattern Recognition*, pages 1 –4, dec. 2008.

[2] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48(3):259–302, 1986.

[3] E. Boros and P. L. Hammer. Pseudo-boolean optimization. *Discrete Applied Mathematics*, 123(1-3):155 – 225, 2002.

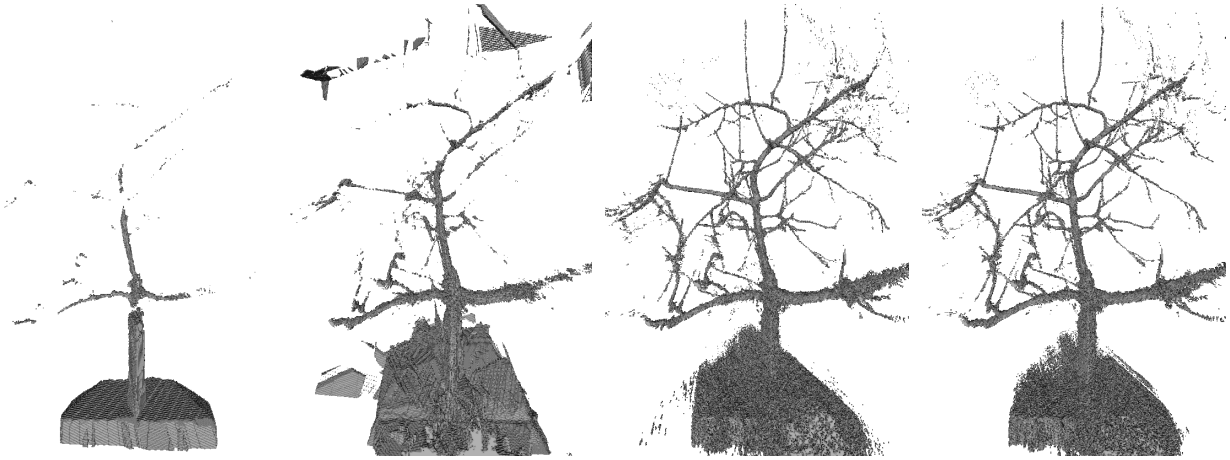
[4] G. K. Cheung, T. Kanade, J.-Y. Bouguet, and M. Holler. A real time system for robust 3d voxel reconstruction of human motions. *CVPR '00: Proceedings of the IEEE Interna-*

[6] Our results are shown without postprocessing other than to smooth the voxels into a surface.



(a) SPOT reconstruction (b) Unbiased hull reconstruction (c) SfIS reconstruction using our method (d) SfSPM reconstruction using our method

Figure 3. Dataset Tree A.



(a) SPOT reconstruction (b) Unbiased hull reconstruction (c) SfIS using our method (d) SfSPM using our method

Figure 4. Dataset Tree B.

tional Conference on Computer Vision and Pattern Recognition, 2:2714, 2000.

[5] C. Chien and J. Aggarwal. Volume/surface octrees for the representation of three-dimensional objects. *Computer Vision Graphics and Image Processing*, 36:100–113, 1986.

[6] L. Díaz-Más, R. Muoz-Salinas, F. Madrid-Cuevas, and R. Medina-Carnicer. Shape from silhouette using dempster-shafer theory. *Pattern Recognition*, 43(6):2119 – 2131, 2010.

[7] J.-S. Franco. *Three-dimensional modeling from silhouettes*. PhD thesis, Institut National Polytechnique de Grenoble, 2005.

[8] J.-S. Franco and E. Boyer. Fusion of Multi-View Silhouette Cues Using a Space Occupancy Grid. In *ICCV '05: Proceedings of the Tenth IEEE International Conference on Computer Vision*, pages 1747–1753, Beijing, China, 2005. IEEE Computer Society.

[9] J.-S. Franco and E. Boyer. Efficient polyhedral modeling from silhouettes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(3):414–427, March 2009.

[10] D. M. Greig, B. T. Porteous, and A. H. Seheult. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society. Series B (Methodological)*, 51(2):pp. 271–279, 1989.

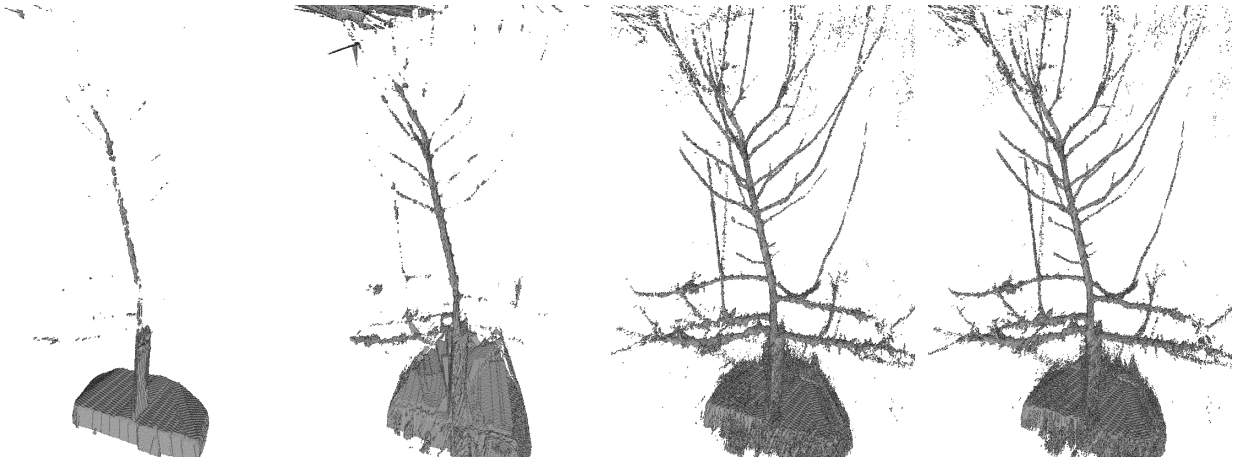
[11] L. Guan, J.-S. Franco, and M. Pollefeys. 3D Occlusion Inference from Silhouette Cues. In *CVPR '07: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pages 1–8, United States, June 2007.

[12] P. Hammer, P. Hansen, and B. Simeone. Roof duality, complementation and persistency in quadratic 01 optimization. *Mathematical Programming*, 28:121–155, 1984. 10.1007/BF02612354.

[13] G. Haro and M. Pardàs. Shape from incomplete silhouettes based on the reprojection error. *Image and Vision Computing*, 28(9):1354 – 1368, 2010.

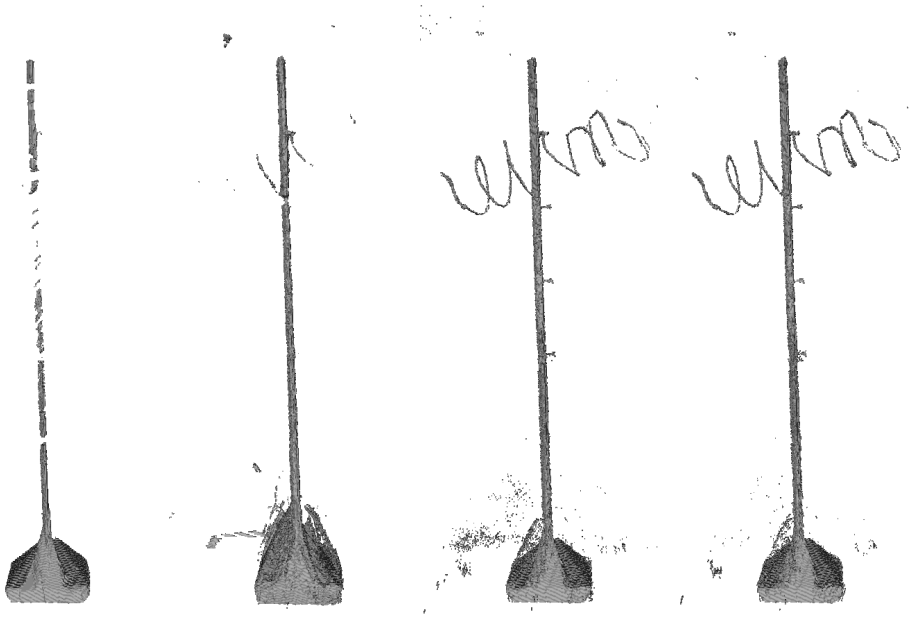
[14] V. Kolmogorov and R. Zabin. What energy functions can be minimized via graph cuts? *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(2):147–159, Feb. 2004.

[15] J.-L. Landabaso, M. Pardàs, and J. R. Casas. Shape from inconsistent silhouette. *Computer Vision and Image Under-*



(a) SPOT reconstruction (b) Unbiased hull reconstruction (c) SfIS using our method (d) SfSPM using our method

Figure 5. Dataset Tree C.



(a) SPOT reconstruction (b) Unbiased hull reconstruction (c) SfIS using our method (d) SfSPM using our method

Figure 6. Dataset with the metal pole and copper coil.

standing, 112(2):210 – 224, 2008.

[16] A. Laurentini. The visual hull concept for silhouette-based image understanding. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(2):150–162, 1994.

[17] W. Martin and J. Aggarwal. Volumetric description of objects from multiple views. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 5(3):150–158, 1983.

[18] M. Piccardi. Background subtraction techniques: a review. In *2004 IEEE International Conference on Systems, Man and Cybernetics*, volume 4, pages 3099 – 3104 vol.4, oct. 2004.

[19] M. Potmesil. Generating octree models of 3d objects from their silhouettes in a sequence of images. *Comput. Vision Graph. Image Process.*, 40(1):1–29, 1987.

[20] R. Szeliski. Rapid octree construction from image sequences. *CVGIP: Image Underst.*, 58(1):23–32, 1993.

[21] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for markov random fields with smoothness-based priors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(6):1068–1080, 2008.