3D INDUSTRIAL RECONSTRUCTION BY FITTING CSG MODELS TO A COMBINATION OF IMAGES AND POINT CLOUDS

Tahir Rabbani, Frank van den Heuvel *

Section of Photogrammetry and Remote Sensing, Faculty of Aerospace Engineering, Delft University of Technology, Kluiverweg 1, 2629 TS Delft, The Netherlands
Email: t.rabbani@lr.tudelft.nl, F.A.vandenHeuvel@lr.tudelft.nl

Commission V, WG V/I

KEY WORDS: Industrial Reconstruction, Point Cloud, Laser Scanning, Photogrammetry, Recognition, Registration, Modelling, Automation

ABSTRACT:

We present a method for 3D reconstruction of industrial sites using a combination of images and point clouds with a motivation of achieving higher levels of automation, precision, and reliability. Recent advances in 3D scanning technologies have made possible rapid and cost-effective acquisition of dense point clouds for 3D reconstruction. As the point clouds provide explicit 3D information, they have a much higher potential for the automation of reconstruction. However, due to the measurement principle employed by laser scanners and their limited point density, the information on sharp edges is not very reliable. It is precisely where images have superiority over point clouds. In addition images are required for visual interpretation, texture mapping, and modelling parts not visible in the point clouds. Moreover, image acquisition is more flexible, and the cost and time required for it is much lower than that of laser scanning, making their combined use essential for a cost-effective solution. These reasons led us to develop a modelling strategy that uses both images and point clouds in combination with a library of CAD primitives found in industrial scenarios represented as CSG (Constructive Solid Geometry) objects. The modelling pipeline in our algorithm starts from point clouds as the main data source for automation. First of all we segment the point cloud using surface smoothness and detect simple objects like planes and cylinders using Hough Transform. This is followed by fitting of CSG objects to a combination of segments. These fitted CAD models are used as registration targets for adding more scans to the project. Additionally, by fitting the projected edges to image gradients we register images to point clouds. Once we have a registered data set, manual measurements are added to images to model missing parts and to increase the reliability of modelling for portions where laser data is known to be noisy. The final phase is similar to bundle adjustment in traditional Photogrammetry as there we estimate pose and shape parameters of all CSG objects using all image measurements and points clouds simultaneously. We name this final phase Integrated Adjustment as it integrates all available information to determine the unknown parameters. The results of applying this method to data from an industrial site are presented showing the complementary nature of point cloud and image data. An analysis of improvement in quality of 3D reconstruction shows the benefits of the adopted approach.

1. INTRODUCTION

As built CAD models of industrial sites are required for many purposes like maintenance, documentation, and training. Moreover, current research is focusing on applying Virtual and Augmented Reality for providing various services for training and operation in industrial environments. The implementation of these technologies requires accurate 3D models of industrial environments for various sub-processes like tracking and alignment of virtual and real objects. One such project on which the Section of Photogrammetry and Remote Sensing at TU Delft has been working since 2001 focuses on using Augmented Reality for providing various training services to industrial users (STAR, 2003). In contrast to Virtual Reality where everything has to be modelled explicitly, Augmented Reality is more flexible as it uses a mixture of real-time video and virtual objects and humans. As a result more realistic scenarios and services can be easily implemented without requiring explicit 3D models for all the objects. At the same time the requirement for aligning the video frames to 3D objects becomes more critical. This necessitates more accurate 3D geometric models for the objects present in the surrounding environment, which are used as targets during tracking and alignment.

Traditional techniques for modelling industrial environments use point and line Photogrammetry as it is much faster and convenient compared to manual surveys. An improved approach suitable for Photogrammetric modelling of industrial environments was presented by (Vosselman et al., 2003) which uses fitting of image edge measurements to back-projected contours of the CSG object in the image. This eliminates the measurement to CAD model conversion stage, which is required for point and line Photogrammetry based approaches. Additionally, inclusion of various internal and external geometric and parametric constraints greatly reduces the degrees of freedom. Thus the number of the required manual measurements is also reduced. Still, this process requires a lot of manual work, which is the major cost in any modelling project.

The prospects of implementing any automatic strategy for industrial environments using only images are very dim; there are three main reasons for that. Firstly, there is no explicit 3D information in images; at least two images having good

* Corresponding author.
intersection are needed along with the information about corresponding points to calculate 3D coordinates. Secondly, the clutter which is a universal feature of most industrial sites combined with uncontrolled lighting makes the automatic object detection much more difficult. Thirdly, the information in images is contained mainly on the edges, as there the contrast is usually much better, while in the absence of some distinguishable marks the information about the 3D geometry of the rest of the surface is at best minimal. This becomes a major limitation as in industrial environments the curved objects are universally present, and for these objects only edge-localized information is not enough for automatic detection and fitting. All these limitation are successfully resolved by laser scanning techniques, where we get direct 3D information in the form of a point cloud making the job of object detection and fitting much easier. Furthermore, recent advances in 3D scanning technologies have made possible high-speed acquisition of dense and accurate point clouds at moderate costs (Laser Scanner Survey, 2003).

The strengths of Laser Scanning do not mean that images lose all their utility. Actually the fact that images provide accurate information on edges becomes a source of strength if both image and point cloud data are simultaneously used. Most of the currently available laser scanners are using techniques based on either triangulation or time of flight. In point clouds acquired using either of them the data on edges is noisier compared to that on the surface of the scanned object. This has to do with the angle between the surface normal and the laser beam, which changes very rapidly near the edges of the object, making precise point cloud acquisition very difficult. Additionally, in the case of time of flight multiple reflected pulses lead to averaging of range measurements. This is especially true for the step edges. These limitations of scanning technologies make images a complementary source of information, especially on the edges of the objects. Acquisition of images as a supporting data source is not a problem, as cameras are still much faster and cheaper than currently available laser scanners. Additionally, images are much better for visual interpretation and are required for producing texture-mapped models for realistic visualization.

Based on the above observation of the complementary nature of images and point clouds the modelling strategy that we have developed uses both data sources simultaneously and thus exploits all available information to achieve a more accurate estimation as well as higher levels of automation. The rest of the paper is organised as follows. In Section 2, we give a summary of the modelling pipeline. Section 3 provides details of fitting CSG objects to point clouds and to image edges. We present fitting results on an industrial scenario in Section 4, along with a discussion about the improvement in estimation accuracy using two experiments of fitting on single objects. Finally, we conclude in Section 5 and propose some directions for future work.

2. MODELLING PIPELINE

Our modelling pipeline is shown in Fig. 1. As it was noted in the introduction we are using both images and point clouds as data sources. We start from an initial approximate scan-to-scan registration using Iterative Closest Point method (Besl and McKay, 1992). The registration obtained from this pre-processing stage is used until objects have been recognized and fitted. Then this initial registration is refined in the final Integrated Adjustment using object-based registration (Dijkman and Heuvel, 2003). For image registration or exterior orientation, image edge to back-projected CSG model contour fitting is used, during which only image exterior orientation parameters are adjusted and object parameters of the modelled objects are kept fixed.

For next stages of segmentation and object recognition only point cloud data is used, as in contrast to images it provides explicit 3D information, and thus has better chances of achieving automation. This is especially true for the reconstruction of industrial sites as due to their man-made origin presence of well-defined CAD primitives can be expected. For example as reported by Nourse et al. (1980) 85% of objects found in industrial scenes can be approximated by planes, spheres, cones and cylinders. This percentage rises to 95% if toroidal surfaces are included in the set of available primitives (Requicha and Voelcker, 1982; Petitjean, 2002).

Using point clouds we take a two-step approach, consisting of segmentation followed by Hough transform based object detection. In the first step we use a simple region growing based segmentation using what we call Smoothness Constraint. It is based on the assumption that most of the surfaces in industrial environments can be expected to be smooth with their surface normals changing rapidly only on the object edges. First of all we estimate the surface normal for each point in the point cloud using plane fitting to the points within a small neighbourhood. This is followed by the stage of region growing in which we keep on adding points to one region until the angle between normals exceeds a specified threshold. Actually, segmentation and object recognition are two related problems, because if we know the type and location of objects, segmentation is reduced to selecting points having a low distance from the object surface; and similarly if we have a perfect segmentation, the object recognition is just a matter of surface fitting and finding the surface which gives minimum error of fit. Most of the segmentation approaches to date haven’t been able to achieve a high success rate (Hoover et al. 1996; Min et al., 2000). The segmentation approach we use leads usually to under-segmented results, with multiple objects being assigned to one segment. The following object recognition stage detects the planes and cylinders present in the segments using a Hough Transform. As presence of multiple objects and outliers is not a problem for the Hough transform we are able to recover from the errors of the preceding stage of segmentation. The object
detection is currently limited to cylinders and planes, but in many industrial environments they can account for more than 70% of the objects.

Next is the Constraint Detection stage where objects which have been multiply detected are combined together, the cylinders which might be connected are found, and the presence of curves between pipes in proximity is hypothesized, and then checked against the point cloud. The process of combining various segments and assigning them to CSG objects from the Object Catalogue is currently manual. But in future, we plan to make it automatic.

The next stage of surface fitting assumes that the combination of preceding segmentation and object recognition stages have resulted in correctly labelled points, and we know which points belong to which CSG model. Similarly in images the point measurements (either manual or automatic) are correctly assigned to their corresponding CSG objects. The details of fitting the selected CSG objects to the point cloud and the image measurements as well as their combination are discussed in the following sections.

3. MODEL FITTING

3.1 Fitting of CSG model to point clouds

The problem we are addressing can be formulated as follows. We have a set of points, which are sampled from some object that can be approximated by the given CSG model. We want to estimate those values of the parameters for the CSG object, which minimize the sum of the squares of the orthogonal distance of the points from the surface of the model i.e.,

$$\min \sum_{i=1}^{N} \Omega[p_i, \Gamma(\tau_1,\tau_2,\ldots,\tau_M)]$$

(1)

$$\Omega$$ defines the shortest distance of a given point $$p_i$$ to the surface of the CSG model $$\Gamma$$ which has $$M$$ shape and pose parameters given by $$\tau_1,\tau_2,\ldots,\tau_M$$. The point cloud consists of $$N$$ points, $$p_1,p_2,\ldots,p_N$$ (Fig. 2(a)).

To solve this non-linear least-squares problem we need a method to find the value of the distance function $$\Omega$$ in (1) and its partial derivatives with respect to the CSG parameters i.e.

$$\frac{\partial \Omega}{\partial \tau_1}, \frac{\partial \Omega}{\partial \tau_2},\ldots,\frac{\partial \Omega}{\partial \tau_M}$$

(2)

The calculation of $$\Omega$$ is a difficult problem, as due to the bounded surfaces used by the CSG specification it is not possible to have closed form analytical expressions. For a comparison of different numerical methods for its computation we refer the reader to Rabbani and van den Heuvel (2004). Here we use ACIS (2004), which is a commercial geometric modelling engine to compute $$\Omega$$. Similarly, the partial derivatives are estimated numerically using finite differences. As noted by Dennis and Schnabel (1996), for sufficiently small step-size, the results obtained from the finite difference approximation of the partial derivatives for the least-squares solution are indistinguishable from the analytical ones.

For minimizing the function (1) with respect to parameters of the CSG model we use Levenberg-Marquardt method (Björk, 1996; Press et al., 1996). Starting from an initial estimate of CSG parameters $$\Gamma_0$$, at each iteration we get an adjustment given by:

The calculation of distances for fitting (a) $$\Omega$$ for a Point Cloud, the model is shown in yellow, the red arrows from green points to model surface indicate the distance (b) $$\Psi$$ for an Image, the measurements are in green, the back-projected model is in yellow, and red arrows indicate their distance in image space.

$$\Delta \Gamma = (J^TJ + \lambda I)^{-1}(J^TD)$$

(3)

$$\Gamma_l = \Gamma_0 - \Delta \Gamma$$

(4)

where $$J$$ is the Jacobian matrix and $$D$$ is the distance vector

$$J_l = \frac{\partial \Omega}{\partial \tau_l}$$

(5)

$$D = \Omega(p, \Gamma_0)$$

(6)

$$\Omega$$ is the distance of the $$l$$th point from the CSG surface, and $$\tau_l$$ is the $$l$$th parameter of the CSG tree. In (3) above $$\lambda$$ is the Levenberg-Marquardt parameter. When $$\lambda = 0$$ Newton step is taken while for $$\lambda \to \infty$$ results in steepest descent step.

We are using quaternions (Shoemake 1985) for the specification of rotation as they provide a singularity free representation. This means we have four rotation parameters with one constraint i.e.:

$$q_1^2 + q_2^2 + q_3^2 + q_4^2 = 1$$

(7)

The constraint in (7) cannot be enforced during the adjustment, as Levenberg Marquardt is an unconstrained optimization method. This means that we have an over-parameterisation and the resulting matrix of normal equations can be singular. To avoid the resulting numerical problems we use Singular Value Decomposition (Golub, 1996) for inverting the matrix in (3). This way if there is a rank deficiency we take the column corresponding to minimum singular value out of the matrix system and thus get a minimum norm solution.

3.2 Fitting of CSG Model to images

The use of CAD models for fitting to images was pioneered by Lowe (1991). He estimated the pose and the shape parameters by minimizing the distance of the visible edges from the hidden-line projection of the estimated model. Vosselman et al. (2003) extended and modified this approach for fitting CSG objects to image gradients and point measurements for industrial reconstruction. They also used internal and external geometric constraints to reduce the number of degrees of freedom and thus the required image measurements. We follow their fitting approach for images, with one exception that we don’t know a priori the correspondence between image measurements and back projected edges of the CSG model. Due to this missing information we follow an iterative procedure, where before each iteration for fitting, the measurements are assigned to the closest edge.
defines the shortest distance of a given measurement in an image to the closest edge of the back projected CSG. It is also a non-linear least squares problem and is similar to that of minimizing \( \Omega \) discussed in the above section. We need partial derivatives with respect to CSG parameters i.e., \( \frac{\partial \Psi}{\partial \tau_1}, \frac{\partial \Psi}{\partial \tau_2}, \ldots, \frac{\partial \Psi}{\partial \tau_M} \).

Although analytic expressions for the estimation of the partial derivatives for some of the CSG objects are given by Ermes et al (1999), here we estimate them numerically using finite differences. The final estimation uses Levenberg-Marquardt in combination with Singular Value Decomposition. The details are similar to the ones discussed in Section 3.1.

### 4. FITTING EXPERIMENTS

As it was said in the introduction, images and point clouds provide complementary sources of information, and by their combination we can expect better estimation accuracy. Edges of the object where laser scanner usually provide noisy data are captured best in the images. Additionally, while fitting bounded objects point clouds do not contain enough information about determining the bounds, whereas by providing the full edge outline images fix the bounds. For example in the case of a cylinder usually the closing lids on both sides are not scanned either because they are not visible due to the connections with other surrounding objects, or because it is not convenient to place the scanner in a position where the lids are visible. As a result we expect the length of the cylinder to be poorly determined by such a point cloud. In contrast the measurements in the image provide points on the edges and thus help improve the precision of the length estimate.

To demonstrate the complementary nature of the information coming from images and point clouds we will do some fitting experiments on two test objects. Each object will be fitted three times, first using only point cloud, then using only image measurements and finally a combination of both. The point clouds we will use were captured using a Cyrax scanner. We assume standard deviation of 5mm for each point. The images were captured using a Nikon CoolPix camera having a resolution of 5 mega pixels and using a fixed focal length of 7.34 mm. The standard deviation for image measurements is taken to be 1 pixel.

#### 4.1 Cylinder fitting

The arrangement we used for the first experiment is shown in Fig. 3. A cylinder is scanned from the front, and images are taken from three different positions. We see back-projected hidden lines in yellow, points measured on edges in red, while the sub-sampled point cloud is shown in white. A cylinder is taken to be 1 pixel. As it was said in the introduction, images and point clouds do not contain enough information about determining the bounds, whereas by providing the full edge outline images fix the bounds. For example in the case of a cylinder usually the closing lids on both sides are not scanned either because they are not visible due to the connections with other surrounding objects, or because it is not convenient to place the scanner in a position where the lids are visible. As a result we expect the length of the cylinder to be poorly determined by such a point cloud. In contrast the measurements in the image provide points on the edges and thus help improve the precision of the length estimate.

Thus the fitting problem reduces to the estimation of those values of the CSG parameters, which minimize the sum of the squares of the orthogonal distance of the image measurements from the back-projected contour of the CSG model. The back projection must have a mechanism for hidden-line removal, so that the effects of self and external occlusions are taken into account. We follow the second approach, and use ACIS (2004), which is a commercial geometric modelling engine to compute the hidden line projection of the model in the image. For an example see Fig. 2(b).

![Fig. 3: Cylinder fitting experiment (a) Point Cloud (b-d) Images with back-projected model in yellow, point measurements in red, and sub-sampled point cloud in white](image)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Image (1)</th>
<th>Image (2)</th>
<th>Image (3)</th>
<th>Point Cloud</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>52.269</td>
<td>10.326</td>
<td>9.923</td>
<td>0.838</td>
<td>0.762</td>
</tr>
<tr>
<td>Y</td>
<td>163.45</td>
<td>14.168</td>
<td>13.629</td>
<td>1.335</td>
<td>0.862</td>
</tr>
<tr>
<td>Z</td>
<td>12.740</td>
<td>3.489</td>
<td>3.467</td>
<td>119.90</td>
<td>1.654</td>
</tr>
<tr>
<td>t0</td>
<td>0.056</td>
<td>1.0E-2</td>
<td>1.0E02</td>
<td>2.28E-3</td>
<td>2.0E-3</td>
</tr>
<tr>
<td>t1</td>
<td>0.065</td>
<td>2.1E-2</td>
<td>2.1E2</td>
<td>3.96E-3</td>
<td>2.9E-3</td>
</tr>
<tr>
<td>t2</td>
<td>5.283</td>
<td>2.538</td>
<td>2.534</td>
<td>0.359</td>
<td>0.237</td>
</tr>
<tr>
<td>Length</td>
<td>10.591</td>
<td>3.121</td>
<td>3.120</td>
<td>(\infty)</td>
<td>1.857</td>
</tr>
<tr>
<td>Radius</td>
<td>169.57</td>
<td>16.166</td>
<td>15.522</td>
<td>0.634</td>
<td>0.565</td>
</tr>
</tbody>
</table>

Table 1: Standard deviation for Cylinder fitting experiment

Each point measurement in the image gives us a ray in 3D. Given a set of images with measured points we want to estimate those values for CSG parameters that result in the minimum distance between all these rays and the estimated CSG model. Alternatively, the ray to body distance can be calculated in image space. There we have to compute the distance in pixels between an image measurement and the closest back-projected contour of the CSG model. The back projection must have a mechanism for hidden-line removal, so that the effects of self and external occlusions are taken into account. We follow the second approach, and use ACIS (2004), which is a commercial geometric modelling engine to compute the hidden line projection of the model in the image. For an example see Fig. 2(b).

Thus the fitting problem reduces to the estimation of those values of the CSG parameters, which minimize the sum of the squares of the orthogonal distance of the image measurements from the back-projected contour of the CSG model. The back projection must have a mechanism for hidden-line removal, so that the effects of self and external occlusions are taken into account. We follow the second approach, and use ACIS (2004), which is a commercial geometric modelling engine to compute the hidden line projection of the model in the image. For an example see Fig. 2(b).

The problem of minimizing \( \Psi \) is also a non-linear least squares problem and is similar to that of minimizing \( \Omega \) discussed in the above section. We need partial derivatives with respect to CSG parameters i.e., \( \frac{\partial \Psi}{\partial \tau_1}, \frac{\partial \Psi}{\partial \tau_2}, \ldots, \frac{\partial \Psi}{\partial \tau_M} \).

Although analytic expressions for the estimation of the partial derivatives for some of the CSG objects are given by Ermes et al (1999), here we estimate them numerically using finite differences. The final estimation uses Levenberg-Marquardt in combination with Singular Value Decomposition. The details are similar to the ones discussed in Section 3.1.

![Fig. 3: Cylinder fitting experiment (a) Point Cloud (b-d) Images with back-projected model in yellow, point measurements in red, and sub-sampled point cloud in white](image)
As the z-axis is aligned with the length of the cylinder there is a very high correlation between both of them. As a result the estimation of z-position is also very weak compared to the estimation of x and y position. But if we combine the point cloud with measurements from the images (Table 1, column “Both”) the situation improves dramatically as the edges in the images provide enough information about the length and the resulting standard deviations are much lower, indicating much better estimation precision.

Cylinder axis can be specified using two parameters, but as we are using 3 without enforcing the constraint there is an over-parameterisation. Although the standard deviation of axis parameters look quite good, but due to over parameterisation their correlation is very high. For example the correlation between t1 and t2 is 0.52, which indicates that the low values of standard deviations are due to some numerical effects.

As expected as we use more images the standard deviation of parameter estimation goes down. It also shows that even a single image in combination with a good scan can lead to significant improvement in the estimation of those parameters which are not well-determined from the point cloud.

### 4.2 Box Fitting

The second example is that of a box, with only two of its faces fully scanned. Additionally, three images are taken from different positions (Fig.4). The box has 10 parameters, 3 for the position, 4 for the rotation, and 3 for the sizes. Again, similar to the example of cylinder discussed above, we have an over-parameterisation for rotation, as we use 4 instead of required 3 parameters, and cannot enforce the constraint. Again we find a very high correlation between different q parameters that lowers the confidence in the otherwise low standard deviation values. For example the correlation between q0 and q1 is 0.511.

In the absence of points on all faces of the box, it is not possible to reliably determine the size parameters of the box. That’s what we see in the standard deviation resulting from fitting using only point clouds (Table 2), where the standard deviation for y and z sizes is ∞ meaning that they could not be estimated. The value of standard deviation for x size is low only because of the coordinate system chosen for the box, which has its origin in the left corner. This fixes the position of right side and thus the x size is also determined. Due to high correlation between z-position and z-size, its estimation is also bad.

Once again, we see from the last column of the Table 2 that the inclusion of image measurements leads to a much better estimation of size and position parameters.

Both of these examples prove our thesis, that although point clouds contain direct 3D information, which is very useful for automatic object recognition, the final adjustment must use a combination of both data sources to account for missing or noisy information in point clouds.

### 4.3 Modelling of an industrial site

We applied the presented methodology for making 3D model of an industrial site shown in Fig. 5. Seven scans were made using a Cyrax laser scanner. Each scan consisted of one million points with a standard deviation of 5mm. Additionally about 60 images were taken from different positions. Following the modelling pipeline discussed in Section 2 we started with approximate registration using ICP. The approximately registered scan was segmented using Smoothness constraint based region growing. Cylinders and planes were automatically detected using the Hough transform, and then used to refine the scan-to-scan registration. For images the orientation was approximated using vanishing points. This was followed by scan-to-image registration using a few image measurements and keeping all object parameters fixed, while estimating only exterior orientation parameters of the images.

The process of combining automatically detected cylinders and planes to full CSG objects as well as the process of adding measurements to images was done manually. Once we have image measurements as well as segmented points clouds, we proceed with the Integrated adjustment using both data sources simultaneously. This integrated adjustment minimizes the sum of square of the distances of point cloud from the model surface and sum of square of the image measurement distance from the back projected edges of the model, while estimating the pose and shape parameters of the CSG object as well as the registration parameters of the individual scans and exterior orientation of the images. This process is an extension of the idea of bundle adjustment in traditional Photogrammetry but is...
much more general, as it uses both point cloud and image measurements in big adjustment.

This integrated adjustment was applied to the test scenario shown in Fig. 5. Only cylinders, boxes and tori were used from the catalogue of CSG objects. The results of the fitting are shown as a 3D Model in Fig. 5(d).

5. CONCLUSIONS

We have presented a modelling technique for fitting CAD models described as CSG primitives to measurements in images and point clouds. While the point clouds are excellent for automatic object recognition, the comparison of improvement in the standard deviation of the estimated parameters clearly shows that images have a complementary role as they provide more information on the edges and help fix the bounds of models where point clouds fail to do so. In future we plan to extend the fitting procedure from point measurements to edge and curve measurements in images. Similarly different strategies for automatic object recognition in industrial environments using a combination of imagery and point clouds will also be investigated.

REFERENCES:


