IMAGE BASED MODELING ON A CONSTRAINED AXIS

Laura Wieme
Rochester Institute of Technology
Rochester, NY
ljw5439@rit.edu

Abstract

Given a set of multiple calibrated photographs, silhouettes can be extracted to form a coarse approximation of 3D geometry. This image based model can further be refined by using feature based photoconsistency constraints and iterative bundle adjustment to correlate and refine the details found in the data set. Image based modeling has practical applications in visual effects, medical imaging, and prototyping.

1. Background

Image based modeling refers to the technique of reconstructing solid models of complex three-dimensional objects from a series of multiple calibrated photographs. The ultimate goal is to obtain a water tight model that preserves all the cavities and fine detail of the original object. For sake of this paper, it is assumed that the targeted data sets are constrained along one axis of rotation. This generally consists of images captured from a turn table since these data sets are best for obtaining a single high quality model. Since the cost of high quality digital imaging equipment has declined, image based modeling has become a very active and accessible research area.

2. Obtaining a Basic Mesh

Before being able to advance to more current refinement methods, a generalized volume of the photographed object must be obtained. This initial step has been pruned down to three common approaches, all of which assume that silhouettes have previously been extracted from the data set. The first method, space carving, builds a volume by creating a grid of voxels around projected surface of the object. The term carving refers to the process of iteratively removing voxels with a best fit approach until a desirable level of detail is reached. This method is appropriate for creating a 3D visualization of a model and simplifies the texturing process, but it is not suitable for applications in which a polygonal mesh is desired. Also, structures created from voxels lack structural information, such as intersection curves, that are required for refinement algorithms. The second approach utilizes the common global optimization technique, graph cuts, to find an optimal surface. This method proves to be slightly better for polygonal mesh creation, but since extrema are removed, the created model tends to lack detail. The oldest, yet still frequently used approach is to create a visual hull. A visual hull refers to the process of projecting cones through each silhouette image. The resulting intersections produce the outer approximation of the object's surface.

Since this method closely models similar techniques familiar to graphics programmers, such as raytracing, this approach is easily implemented and creates a reasonable watertight polygonal mesh. It is because of these advantages that the visual hull technique was chosen for this study.

3. Visual Hull Approach

The presented visual hull procedure is based off of the work of Lazebnik, Furukawa, and Ponce. As previously stated, the visual hull is achieved by projecting solid geometric cones through the back of each silhouette image and finding the intersections. Lazebnik improves upon the basic algorithm by describing the resulting hull as a generalized polyhedron rather than the traditional generalized volume. Faces are defined as the surface patches of individual cones, edges are intersection

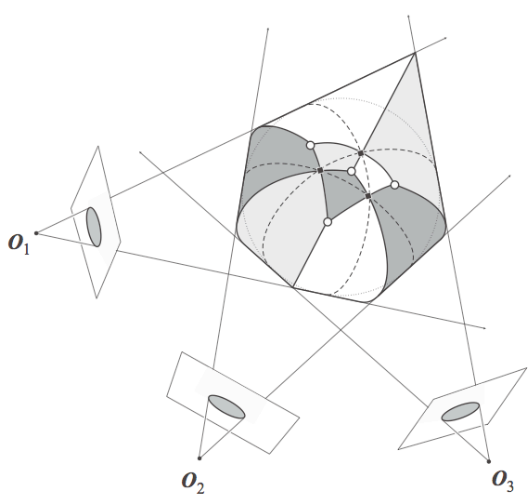


Diagram from Lazebnik

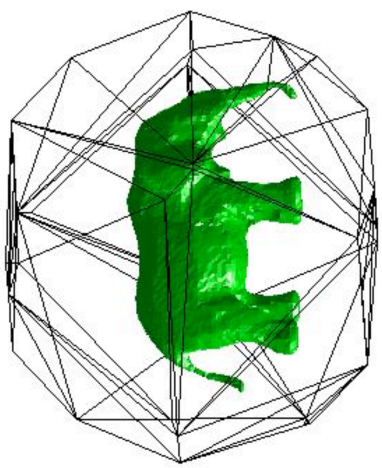
curves of cones, and vertices are located where three or more faces come in contact. Finally, Lazebnik defines the boundary of the surface as the union of "cone strips." Despite a boundary being define, one elementary graphical problem remains to create a proper 3D representation. In projective geometry, the generated surfaces lack a definition of front vs back required to accurately represent the object. Lazebnik solves the projection problem by adapting her previously proposed method of *oriented projective differential geometry*. The algorithm begins by sequentially producing each visual cone. With the calculation of each new cone, points that fall outside of this new cone are clipped by the definition of the visual hull. As a second iteration, intersection curves are clipped using the previous cones for comparison. This clipping procedure results in a boundary representation, but not yet a cohesive polygonal model. Lazebnik then uses epipolar tangency to trace an epipolar line to find crucial points and their sign values. The sign values determine back and front facing surfaces previously defined as the projection problem. A branch structure of the crucial points is created to be used as coordinates. Pairs of coordinates are then triangulated to create continuous polygonal strips, achieving the basic visual hull.

4. Visual Hull Results

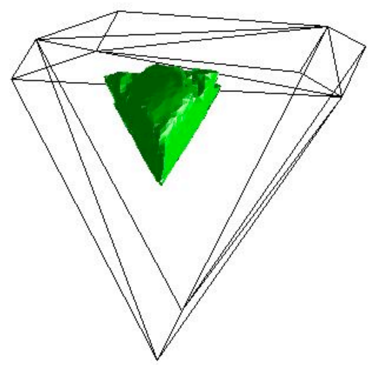
The first results are from a data set specifically designed for visual hull calculation. Preprocessed silhouettes and camera calibration data were provided.

Reconstruction of a elephant with 10 views

Partial reconstruction of the dinosaur with 6 views



Visualization of the view boundaries



Visualization of the missing views

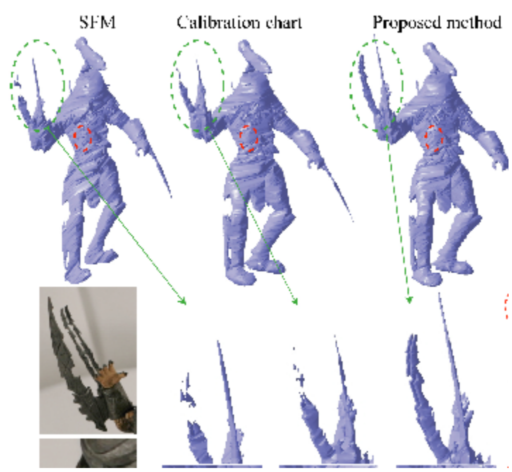
The following results are from a trial of the implementation of a Middlebury multiview stereo data set of a ceramic dinosaur. The method was able to build the mesh from the first six input images, but would fail if any subsequent views were added.

5. Refinement Methods

More recent image based modeling techniques aim to reach higher accuracy and detail by applying a refinement method to the basic visual hull. Of particular interest is the recent patch based refinement approach developed by Furukawa. Furukawa's approach builds off of the previously mentioned visual hull algorithm by reconstructing oriented points densely covering the objects surface. Using the visual hull for surface approximation, image features are matched across the multiple views based on visibility. Sparse bundle adjustment software is then used to improve camera calibration. Bundle adjustment is an optimization problem in which the 3D mesh coordinates are refined to minimize the reprojection error between the the image points and our predicted locations. Often the sum of squared differences algorithm is used for these minimization calculations. Bundling refers to the process of grouping rays from each feature that then converge to each camera's center. Since this paper focuses on constrained data sets, the bundle adjustment procedure can be simplified because our rotation matrix is constrained to one axis. This also produces more accurate results because there is a lower chance of error or variation in specifying our camera parameters. Initial feature detection is determined by visibility information and surface correspondence from the visual hull using epipolars. Only a sparse set of these points are stored since bundle adjustment does not require a dense data set to produce accurate results. The next step is to refine the detected features by texture comparison. Since the previous visual hull method defines an orientation normal for each point, a texture feature patch can be formed for comparison with other patches. This process is iterative and with each pass the centers of the patches are shifted and refined into the optimal placement of their 3D position. Bundle adjustment is then used as the final refinement.

6. Furukawa's Results

Furukawa's image based modeling approach has produced accurate and detailed results. By refining a series of surface patches based on texture, his method is able to produce extreme photometric consistency. His details are further refined using bundle adjustment to ensure the proper feature placement according to the camera calibration. Furukawa's method particularly stands out when reconstructing fine extremities that other methods may only create a blocky representations or erode away. Below are some of his results.



Fine detail Preservation

Face and hair reconstructed from texture

7. Conclusion

Image based modeling is a much more complex area of research that it initially seems. There are multiple areas for improvement and the field is active and growing. Due to the newness of modern approaches and the general size of the image based modeling problem, I have not yet achieved the level of detail I initial desired. Now armed with a better understanding of the reconstruction process, further work is required to reach the accuracy of Furukawa's results.

References

[1] Furukawa, Y. and Ponce, J. 2009. Accurate Camera Calibration from Multi-View Stereo and Bundle Adjustment. *Int. J. Comput. Vision* 84, 3 (Sep. 2009), 257-268. DOI= http://dx.doi.org/10.1007/s11263-009-0232-2

[2] Furukawa, Y. and Ponce, J. 2009. Carved Visual Hulls for Image-Based Modeling. *Int. J. Comput. Vision* 81, 1 (Jan. 2009), 53-67. DOI= http://dx.doi.org/10.1007/s11263-008-0134-8

[3] Lazebnik, S., Furukawa, Y., and Ponce, J. 2007. Projective Visual Hulls. *Int. J. Comput. Vision* 74, 2 (Aug. 2007), 137-165. DOI= http://dx.doi.org/10.1007/s11263-006-0008-x

[4]Middlebury Multiview Stereo. *vision.**middlebury**.edu/mview/* February 24, 2009.

[5]Wong, K. H., Ming, M., and Chang, Y. 2004. 3D Model Reconstruction by Constrained Bundle Adjustment. In *Proceedings of the Pattern Recognition, 17th international Conference on (Icpr'04) Volume 3 - Volume 03* (August 23 - 26, 2004). ICPR. IEEE Computer Society, Washington, DC, 902-905. DOI= http://dx.doi.org/10.1109/ICPR.2004.8
