

Research Article

An Iterative Surface Evolution Algorithm for Multiview Stereo

Yongjian Xi and Ye Duan

Department of Computer Science, University of Missouri, Columbia, MO 65211, USA

Correspondence should be addressed to Ye Duan, duanye@missouri.edu

Received 2 August 2009; Revised 16 December 2009; Accepted 3 March 2010

Academic Editor: Kenneth K. Y. Wong

Copyright © 2010 Y. Xi and Y. Duan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a new iterative surface evolution algorithm for multiview stereo. Starting from an embedding space such as the visual hull, we will first conduct robust 3D depth estimation (represented as 3D points) based on image correlation. A fast implicit distance function-based region growing method is then employed to extract an initial shape estimation based on these 3D points. Next, an explicit surface evolution will be conducted to recover the finer geometry details of the recovered shape. The recovered shape will be further improved by several iterations between depth estimation and shape reconstruction, similar to the Expectation Maximization (EM) approach. The experiments on the benchmark datasets show that our algorithm can obtain high-quality reconstruction results that are comparable with the state-of-art methods, with considerable less computational time and complexity.

1. Introduction

Despite significant advancement in interactive shape modeling, creating complex high-quality realistic looking 3D models from scratch is still a very challenging task. Recent advancement in 3D shape acquisition systems such as laser range scanners and encoded light projecting system has made directly 3D data acquisition feasible [1]. These active 3D acquisition systems however remain expensive. Meanwhile, the price of digital cameras and digital video cameras keeps decreasing while the quality is improving every day, partially due to the intense competition in the huge consumer market. Furthermore, huge amounts of images and videos are added in internet sites such as Google, and so forth. Every day, a lot of which could be used for multiview image-based 3D shape reconstruction [2].

To date, there have been a lot of researches conducted in the area of multiview image-based modeling. The recent survey by Seitz et al. [3] gives an excellent review of the state of arts in this area. As summarized by [4], most of the existing algorithms follow a two-stage approach: (1) conduct depth estimation based on local groups of input images; (2) fuse the estimated depth values into a global watertight 3D surface estimation. The depth estimation step

is often based on image correlation [5]. The main differences between existing algorithms are in the second stage, the data fusion step, which can be divided into two categories. The first type of data fusion reconstructs the 3D surface by conducting volumetric data segmentation using global energy minimization approaches such as graph cut [6–11], level-set [12–16], or deformable models [5, 17–19]. Recently, people have proposed other types of data fusion algorithms that are based on local surface growing and filtering [2, 20, 21]. Without global optimization, these types of data fusion algorithms can be computationally more efficient [22, 23].

Our algorithm also follows this two-stage process. We proposed an iterative refinement scheme that iterates between the depth estimation step and the data fusion step. This is similar in spirit of the Expectation Maximization (EM) algorithm. Moreover, we propose a novel outlier removal algorithm based on anisotropic kernel density estimation. Our data fusion algorithm integrates the fast implicit region growing with the high-quality explicit surface evolution; thus it is both fast and accurate.

The rest of the paper is organized as follows. In Section 1.1 we discuss the main differences between our approach and related existing works. Section 2 describes the

details of our algorithm. The benchmark data evaluation is shown in Section 3. The paper concludes in Section 4.

1.1. Comparison with Related Works. Our work is most related to the works of Hernández and Schmitt [5] and Quan et al. [16, 24]. Hernández et al. proposed a deformable model-based reconstruction algorithm [5] that achieves one of the highest-quality reconstruction [3]. The depth estimation of [5] is conducted by rectangular window-based normalized cross-correlation (NCC). The estimated depth values are then discretized into an octree-based volumetric grid. Finally a gradient vector flow-based deformable model is applied to the volumetric grid to reconstruct the 3D surface.

Our depth estimation follows the similar pipeline of [5], with several modifications to further improve its efficiency. We will describe these modifications in Section 2.2. Furthermore, unlike [5], we represent the depth estimations as 3D points whose accuracy is not restricted by the resolution of the volumetric grid. Quan et al. [16, 24] also represent the estimated depth values as 3D points. However, unlike our method, they do not have an explicit outlier removal. Instead they rely on level-set-based surface evolution with high-order smoothness terms such as Gaussian/mean curvature to overcome noises, which may create surfaces that maybe too smooth to represent finer geometry details of the original object. Most recently, Campbell et al. [4] proposed an outlier removal algorithm based on the Markov Random Field (MRF) model which can achieve very impressive reconstruction results. On the other hand, our outlier removal algorithm is based on kernel density estimation and is conducted on 3D unorganized points instead of the 2D image space of [4].

To summarize, the main contributions of this paper are. (1) a novel iterative refinement scheme between the depth estimation and the data fusion, (2) a novel anisotropic kernel density estimation based outlier removal algorithm, (3) a novel data fusion algorithm that integrates the fast implicit distance function-based region growing method with the high-quality explicit surface evolution.

2. Algorithm

The entire algorithm (Figure 1) consists of the following five main steps:

- (1) visual hull construction,
- (2) 3D point generation,
- (3) outlier removal,
- (4) implicit surface evolution,
- (5) explicit surface evolution.

Starting from an initial shape estimation such as the visual hull (Step 1), we will use this shape estimation to generate more accurate 3D points based on image correlation-based depth estimation (Step 2), which can then be used to create a better shape estimation (Step 3 to Step 5). In practice, two to three iterations between Step 2 and Step 5 will be sufficient

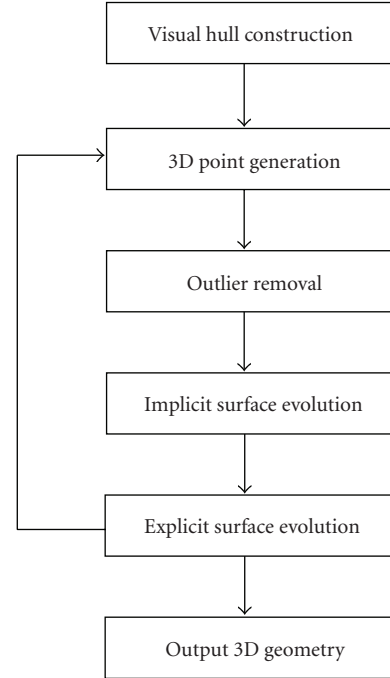


FIGURE 1: Flow-chart of the algorithm.

to create a very good shape estimation. Figure 2 is a 2D illustration of the reconstruction process. Figures 3, 4, 5, and 6 show the corresponding intermediate steps of one iteration of the 3D reconstruction process for the four benchmark datasets of [25], dino sparse ring, dino ring, temple sparse ring, and temple ring, respectively.

2.1. Visual Hull Construction. The first step of our algorithm is to obtain an initial shape estimation by constructing a visual hull. Visual hull is an outer approximation of the observed solid constructed as the intersection of the visual cones associated with all the input cameras [26]. A discrete volumetric representation of the visual hull can be obtained by intersecting the cones generated by back projecting the object silhouettes from different camera views. An explicit shape representation can be obtained by iso-surface extraction algorithms such as Marching Cubes [27].

2.2. 3D Points Generation. Once we had an initial explicit shape estimation, we will proceed to 3D depth estimation. First, we need to estimate the visibility of the initial shape with respect to all the cameras. We use OpenGL to render the explicit surface into the image planes of each individual cameras and extract the depth values from the Z-buffer. Given a point on the surface, its visibility with respect to a given camera can then be decided by comparing its projected depth value into the image plane of the given camera with the corresponding depth value stored in the Z-buffer.

Our depth estimation is based on the Lambertian assumption; that is, if a point belongs to the object surface, its corresponding 2D patches in the image planes of its visible cameras should be strongly correlated. Hence starting

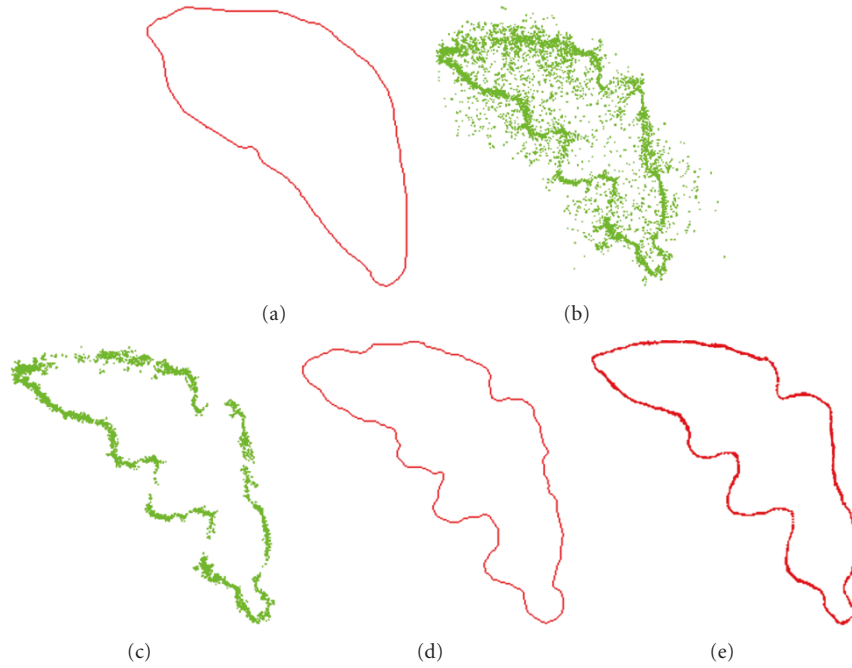


FIGURE 2: A 2D illustration of the whole reconstruction pipeline: (a) visual hull; (b) points generated by depth estimation; (c) after outlier removal; (d) shape estimated by implicit region growing; (e) refined shape estimation by explicit surface evolution.

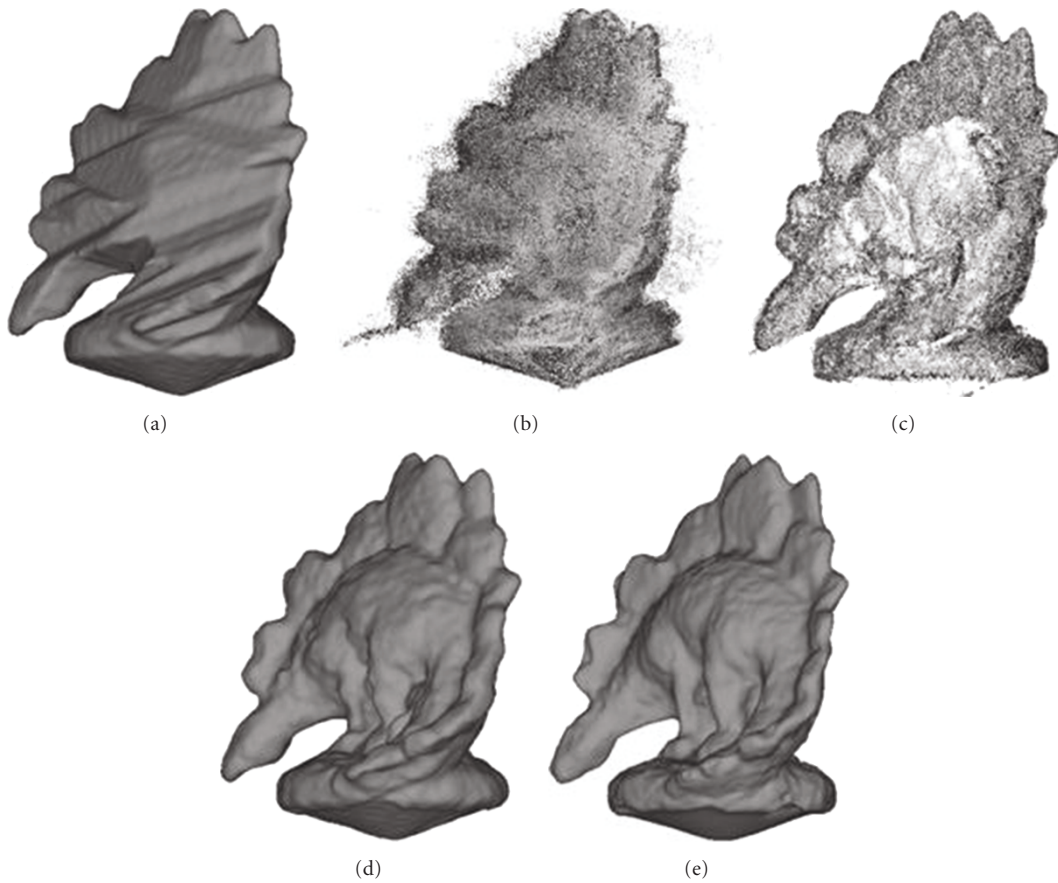


FIGURE 3: Intermediate steps of one iteration of the 3D reconstruction process for the dino sparse ring dataset of [25]. From top left clockwise, visual hull, 3D points generated by depth estimation, after outlier removal, shape estimated by implicit region growing, and refined shape estimation by explicit surface evolution.

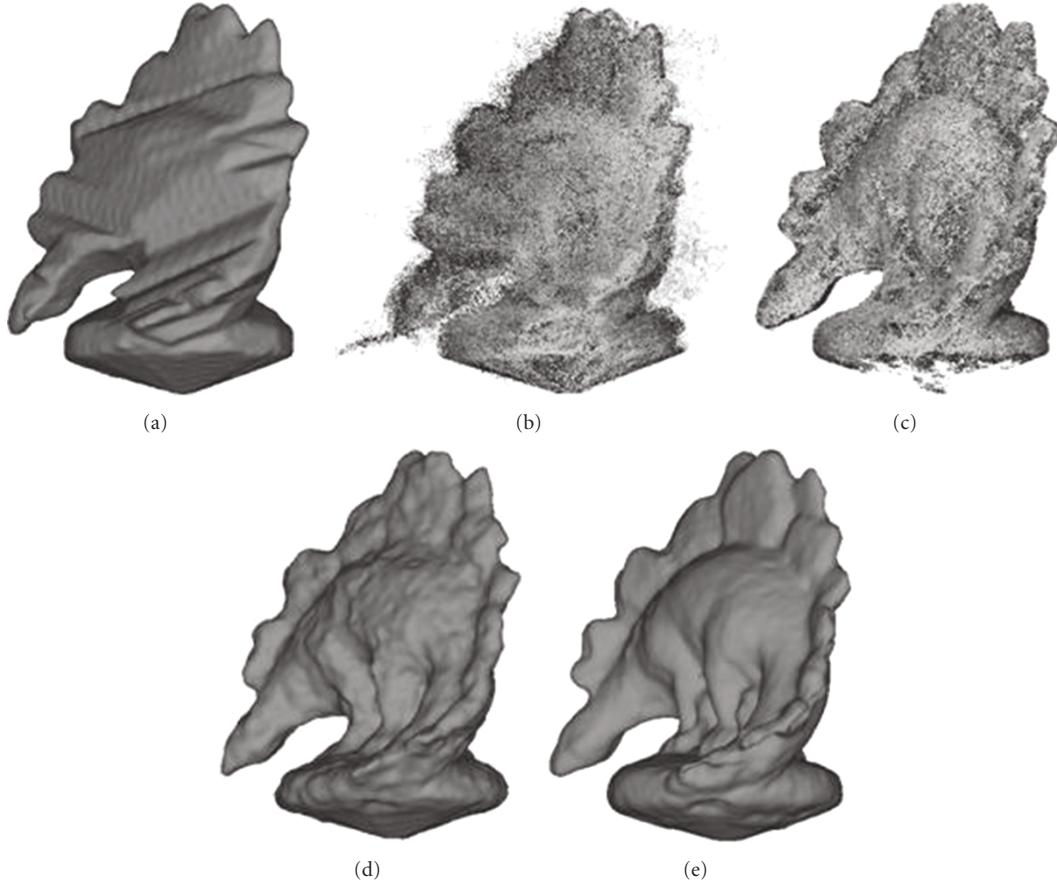


FIGURE 4: Intermediate steps of one iteration of the 3D reconstruction process for the dino ring dataset of [25]. From top left clockwise, visual hull, 3D points generated by depth estimation, after outlier removal, shape estimated by implicit region growing, and refined shape estimation by explicit surface evolution.

from a point on the object surface, we can conduct a line search along a defined search direction to locate the best position whose correlation between the corresponding 2D image patches of different visible cameras is the maxima within a certain search range. This idea is first proposed by [5]. Our paper follows the same principle with several modifications. In the following, we will briefly describe our depth estimation method as well as the main differences between our method and the method of [5].

Given a point on the initial surface, we will select a set of (up to) five “best-view” visible cameras based on the point’s estimated surface normal. Each camera in the selected set will serve as the main camera for once. The search direction is defined as the optical ray passing through the optical center of the main camera and the given point. We will uniformly sample the optical ray within a certain range of the given point, and for each sampled position, we will project it into the image planes of the main camera and another camera in the set, respectively. Rectangular image patches centered at the projected locations of the two image planes will be extracted, and the correlation between the two image patches will be computed by similarity measures such as the normalized cross-correlation (NCC) [5].

For a set of five “best-view” cameras, a total of 20 correlation curves will be generated. For each of the correlation curves, the best position (i.e., the point with the highest correlation value) will be selected as the depth estimation. The depth estimations will be represented as 3D points, which will be processed further to construct a new shape estimation of the object.

The main differences between our implementation and the method of [5] are the following. First, we start the line search from every point on the explicit object surface. The line search in [5] is initiated from every image and the correlation is computed with all the other images, which could be computationally more expensive than ours. Secondly, in [5], for each set of correlation curves computed using the same search direction and the same main camera, only one representative depth estimation is used. While in our method, we avoid this potentially premature averaging by using the depth estimations from all the correlation curves, and postpone the outlier pruning into the subsequent outlier removal step. Thirdly, in [5], the depth estimations are stored in an octree-based volumetric grid, while we store them as discrete points whose accuracy is not restricted by the grid size.

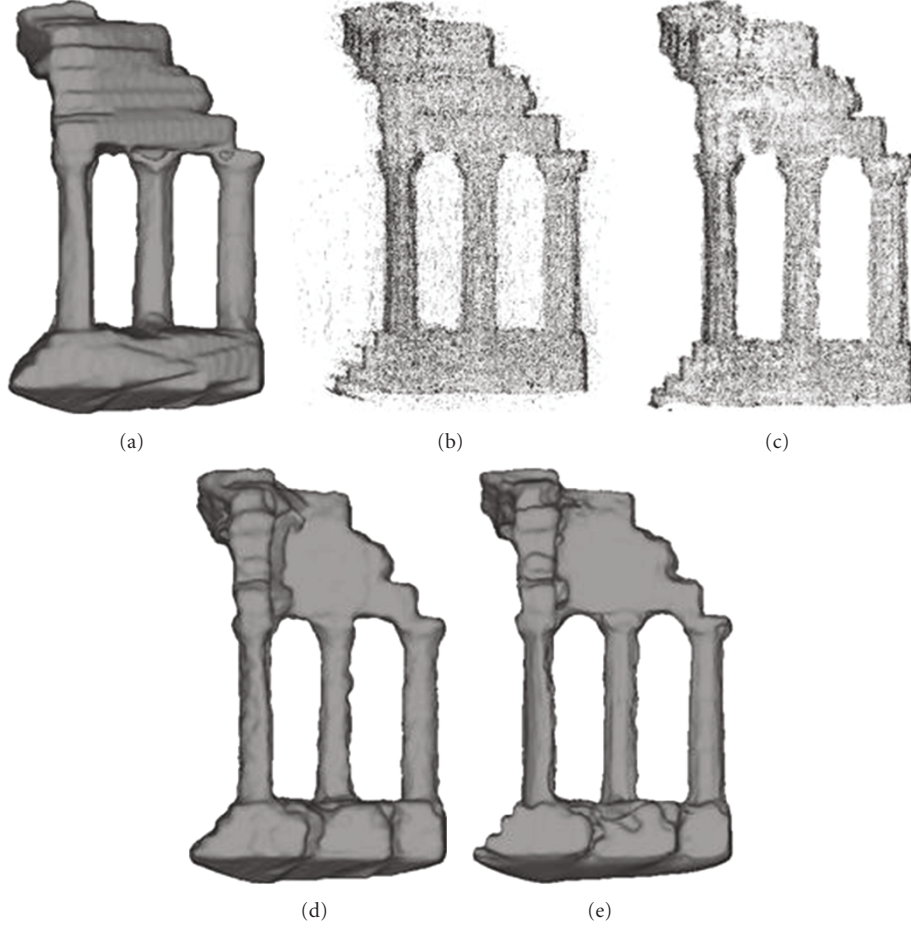


FIGURE 5: Intermediate steps of one iteration of the 3D reconstruction process for the temple sparse ring dataset of [25]. From top left clockwise, visual hull, 3D points generated by depth estimation, after outlier removal, shape estimated by implicit region growing, refined shape estimation by explicit surface evolution.

2.3. Outlier Removal. Points generated by the above depth estimation step may contain outliers (points that do not belong to the object surface) that have to be removed. Since the real object surface is unknown, it is hard to specify a general criterion to detect outliers. In this paper, we propose to employ Parzen-window-based nonparametric density estimation method for outlier removal.

Given n data points x_i , $i = 1, \dots, n$ in the d -dimensional Euclidean space R^d , the multivariate kernel density estimate obtained with kernel $K(x)$ and window radius h (without loss of generality, let us assume $h = 1$ from now on), computed in the point x , is defined as

$$f(x) = \frac{C_{k,d}}{n} \sum_{i=1}^n k(\|x - x_i\|^2), \quad (1)$$

where $\|x\|$ is the L_2 norm (i.e., Euclidean distance metric) of the d -dimensional vector x . There are three types of commonly used spherical kernel functions $K(x)$: the Epanechnikov kernel, the uniform kernel, and the Gaussian kernel [28].

For 3D point cloud obtained by depth estimation, the outliers tend to spread in the space randomly, while “real”

(we use a quotation here to emphasize the fact that the real surface is unknown) surface points will spread along a thin shell which encloses the real surface object. In other words, the distribution of the outliers is relatively isotropic, while the distribution of the real surface points is rather anisotropic. Hence in this paper, we propose to employ an anisotropic ellipsoidal kernel-based density estimation method for outlier removal. More specifically, for anisotropic kernel, the L_2 norm $\|x - x_i\|$ in (1), which measures the Euclidean distance metric between two points x and x_i , will be replaced by the Mahalanobis distance metric $\|x - x_i\|_M$:

$$\|x - x_i\|_M = \left((x - x_i)^t H^{-1} (x - x_i) \right)^{1/2}, \quad (2)$$

here H is the covariance matrix defined as

$$H = DD^T, \quad (3)$$

$$D = (x_1 - x, x_2 - x, \dots, x_n - x).$$

Geometrically, $(x - x_i)^t H^{-1} (x - x_i) = 1$ is a three-dimensional ellipsoid centered at x , with its shape and orientation

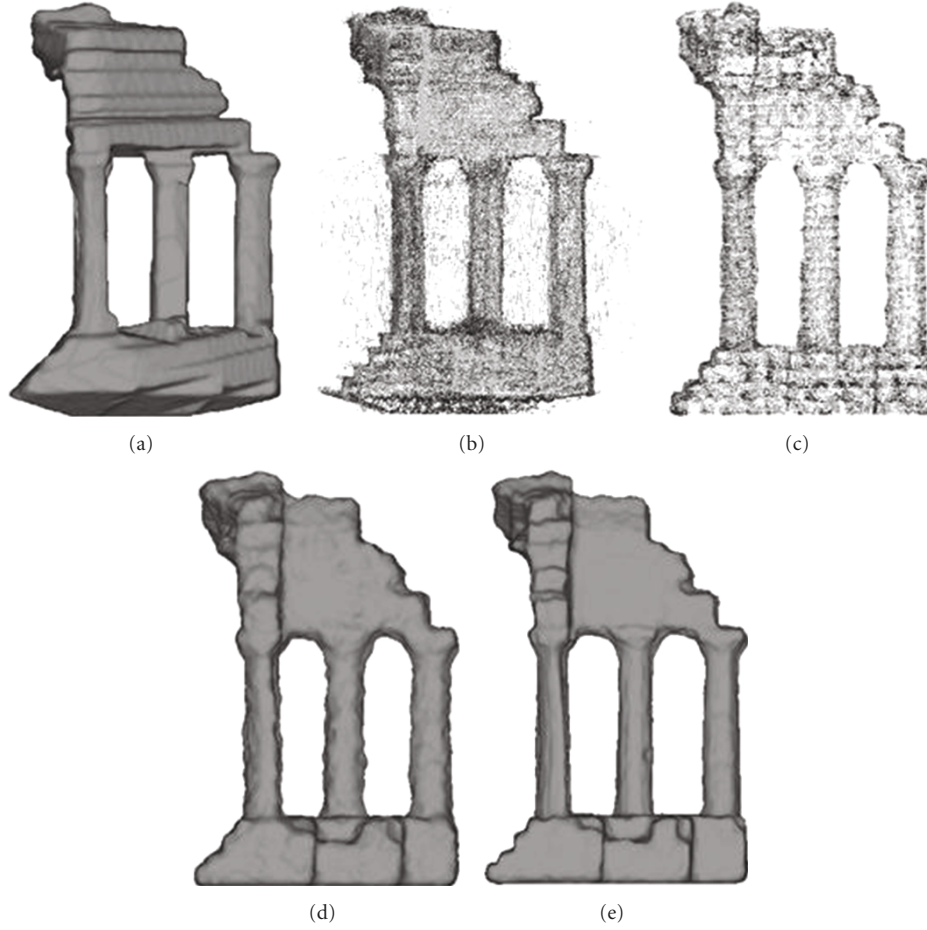


FIGURE 6: Intermediate steps of one iteration of the 3D reconstruction process for the temple ring dataset of [25]. From top left clockwise, visual hull, 3D points generated by depth estimation, after outlier removal, shape estimated by implicit region growing, refined shape estimation by explicit surface evolution.

defined by H . Using Single Value Decomposition (SVD), the covariance matrix H can be further decomposed as

$$H = UAU^T, \quad (4)$$

with

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \quad (5)$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the three eigenvalues of the matrix H , and U is an orthonormal matrix whose columns are the eigenvectors of matrix H .

To compute the anisotropic kernel-based density, we will apply an ellipsoidal kernel E of equal size and shape on all the data points. The orientation of the ellipsoidal kernel E will be determined locally. More specifically, given a point x , we will calculate its covariance matrix H by points located in its local spherical neighborhood of a fixed radius. (Without loss of generality, we will assume the radius is 1, which can be done by normalizing the data by the radius). The U matrix of (4) calculated by the covariance analysis is kept unchanged

to maintain the orientation of the ellipsoid. The size and shape of the ellipsoid will be modified to be the same as the ellipsoidal kernel E by modifying the diagonal matrix A as

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & r \end{bmatrix}, \quad (6)$$

where r is half of the length of the minimum axis of the ellipsoidal kernel E .

After the density value is estimated, we will remove all the points whose estimated density value is smaller than a user-defined threshold. The remaining points will be passed into the subsequent implicit surface evolution step and as long as the outlier removal step does not create very big holes, the implicit surface evolution will be able to create a watertight 3D surface of the object. Figure 7 shows the 3D outlier removal results under different user-defined thresholds. Figure 7(a) is the original point clouds obtained by the aforementioned depth estimation step. The next four images Figures 7(b)–7(e) are the outlier removal results under different user-defined thresholds: 40, 60, 80, and 160,

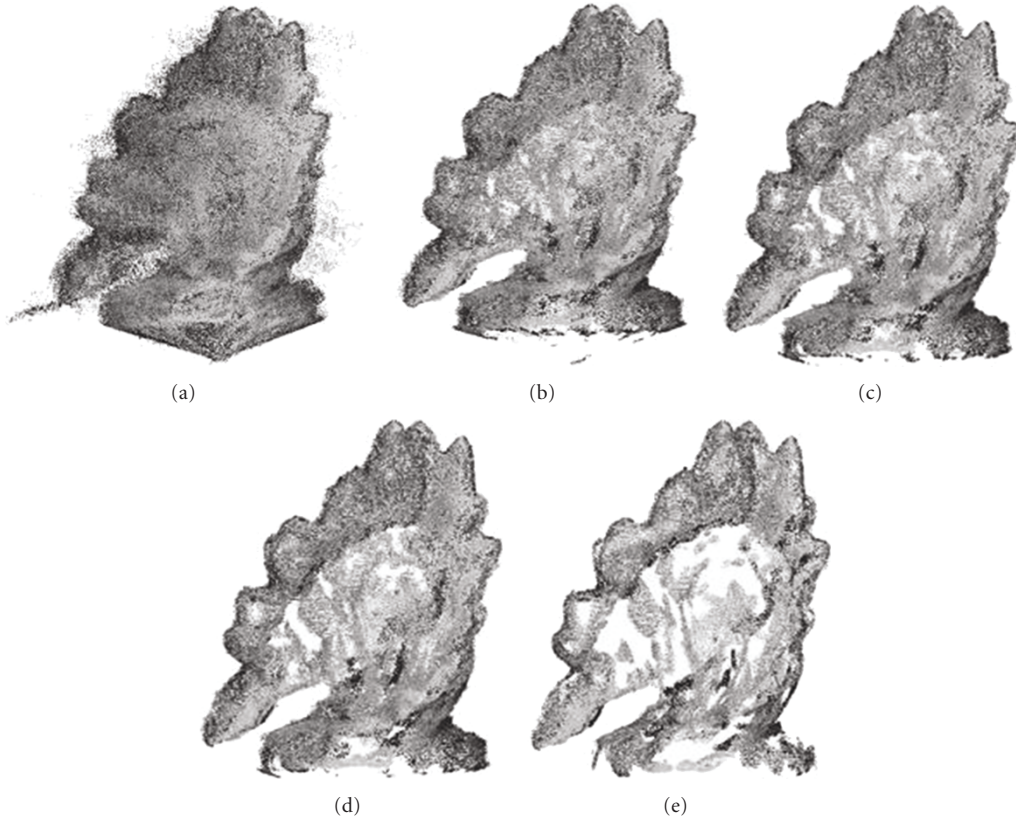


FIGURE 7: Outlier removal results under different user-defined thresholds. (a) original points obtained by depth estimation from the dino sparse ring data. From (b) to (e) are the outlier removal results under different user-defined thresholds: 40, 60, 80, and 160, respectively.

respectively. Among these four outlier removal results, the first three data (Figures 7(b)–7(d)) are all acceptable to the subsequent implicit surface evolution step (Section 2.4) to construct a watertight 3D surface. However the implicit surface evolution step might fail to create a single watertight surface of the object for the fourth data in Figure 7(e) as the threshold is set too high thus creating very big holes in the data.

2.4. Implicit Surface Evolution. After outlier removal, the remaining 3D points will be used to reconstruct the 3D surface of the object. The shape estimation is conducted into two steps. First, a fast implicit distance function-based region growing method—tagging algorithm [29]—is employed to create a coarse shape estimation from the 3D points. Next, an explicit surface evolution step is applied to recover the finer geometry details of the object. We will briefly review the tagging algorithm in the following, for more details please refer to the original paper in [29]. The explicit surface evolution method will be discussed in the next section.

The basic idea of tagging algorithm is to identify as many correct exterior grid points as possible and hence provide a good initial implicit surface, which is represented as an interface that separates the exterior grid points from the interior grid points. There are two main steps in the original

tagging algorithm. First, we will compute a volumetric unsigned distance field based on the 3D points. This is done by the aforementioned fast sweeping method [30]. Once we had the volumetric unsigned distance field, the tagging algorithm will iteratively grow the set of exterior grid points and stop at the boundary of the object. The algorithm can start from any initial exterior region that is a subset of the true exterior region, for example, an outmost corner grid point of the bounding volume, and iteratively tag all the grid points as exterior or interior points based on the comparison of the closeness to the object boundary between the current grid points and its neighboring interior grid points.

2.5. Explicit Surface Evolution. The shape estimation obtained by the implicit tagging algorithm will be converted to explicit mesh by the marching cubes algorithm [27], which will then serve as the initial shape for the subsequent explicit surface evolution step to further improve the geometry accuracy of the shape reconstruction. The surface evolution is guided by energy-optimization-based partial differential equations (PDEs). Classical PDEs such as minimal surface flow [31] usually includes a second-order curvature term to improve the robustness against noise. However it may also prevent the surface evolution to recover finer geometry details. In this paper, we choose the

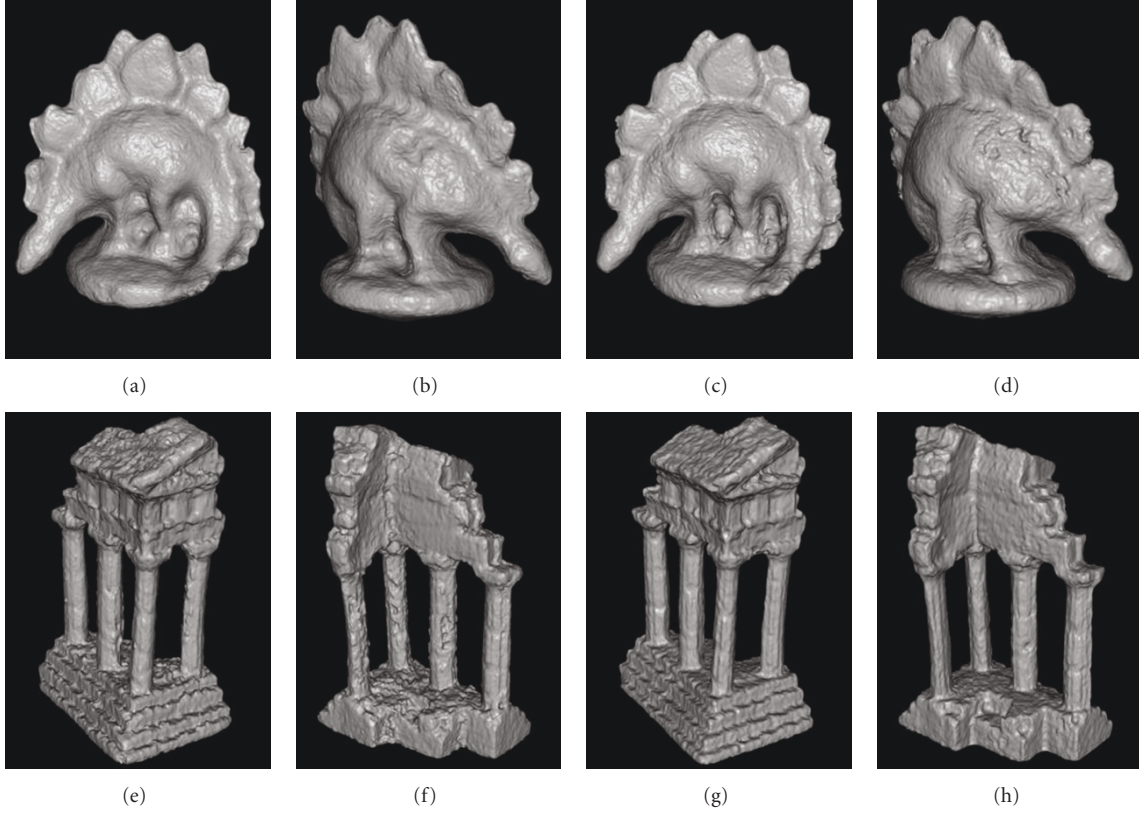


FIGURE 8: 3D rendering of our reconstruction results running on the four Benchmark datasets of [25]. From top to bottom, dino sparse ring, dino ring, temple sparse ring, and temple ring.

TABLE 1: Running time and reconstruction accuracy.

Dataset	Running time (mins : secs)	No. of input images	Accuracy
Temple ring	33 : 17	47	98.9%
Temple sparse ring	29 : 06	16	96.8%
Dino ring	36 : 45	48	97.7%
Dino sparse ring	32 : 01	16	97.6%

simple convection equation to guide the explicit surface evolution:

$$\begin{aligned} \frac{\partial S}{\partial t} &= g(S)\vec{N}, \\ g(S) &= f'(S), \end{aligned} \quad (7)$$

where $S = S(t)$ is the 3D evolving surface, t is the time parameter, $g(S)$ is speed function and is defined as the derivative of $f(S)$, which is the point-based density estimation calculated by (1). \vec{N} is the surface normal vector. The final reconstructed 3D shape is then given by the steady-state solution of the equation $S_t = 0$. Since the speed function g is dynamically calculated at each time step based on the local points distribution, the accuracy of our evolution method will not be limited by the grid resolution as other volumetric image based surface evolution methods such as in [5].

3. Benchmark Data Evaluation

We had applied our algorithm to the four benchmark datasets: temple ring, temple sparse ring, dino ring, and dino sparse ring from [25]. Table 1 shows the running time and the reconstruction accuracy obtained from the evaluation site [25]. The running time is based on a Pentium D Desktop PC with CPU 2.66 GHz, 2 GB RAM. Figure 8 shows the 3D rendering of our final reconstruction results copied from the evaluation website. Our result is listed under the name “SurfEvolution”.

4. Conclusion and Future Work

In this paper, we propose an iterative surface evolution algorithm for 3D shape reconstruction from multiview images. The proposed novel iterative refinements between image correlation-based 3D depth estimation and surface evolution-based shape estimation can significantly reduce

the computational time and improve the accuracy of the final reconstructed surface. The benchmark evaluation results are comparable with the state-of-art methods.

Currently, our method utilizes the visual hull for initial estimation. This requires image segmentation that may be difficult for some images. We would like to relax this requirement in the future. This might be possible since our algorithm uses the iterative refinement which should be able to start from any coarse shape such as a bounding box or a convex hull.

Acknowledgments

The authors are very grateful for Seitz et al. [3] for providing them the datasets used in the paper and Daniel Scharstein for helping them evaluating the result on the benchmark datasets. Research was supported in part by the Leonard Wood Institute in cooperation with the U.S. Army Research Laboratory and was accomplished under Cooperative Agreement # LWI-281074, and by the NSF Grant no. CMMI-0856206. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Leonard Wood Institute, the Army Research Laboratory, the Army Research Office, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

References

- [1] Y. Wang, X. Huang, C. S. Lee, et al., "High resolution acquisition, learning and transfer of dynamic 3-D facial expressions," *Computer Graphics Forum*, vol. 23, no. 3, pp. 677–686, 2004.
- [2] M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. M. Seitz, "Multi-view stereo for community photo collections," in *Proceedings of the 11th IEEE International Conference on Computer Vision (ICCV '07)*, Rio de Janeiro, Brazil, October 2007.
- [3] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, vol. 1, pp. 519–526, July 2006.
- [4] N. Campbell, G. Vogiatzis, C. Hernández, and R. Cipolla, "Using multiple hypotheses to improve depth-maps for multi-view stereo," in *Proceedings of the European Conference on Computer Vision (ECCV '08)*, pp. 766–779, 2008.
- [5] C. Hernández and F. Schmitt, "Silhouette and stereo fusion for 3D object modeling," *Computer Vision and Image Understanding*, vol. 96, no. 3, pp. 367–392, 2004.
- [6] G. Vogiatzis, C. Hernández, P. H. S. Torr, and R. Cipolla, "Multi-view stereo via volumetric graph-cuts and occlusion robust photo-consistency," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 12, pp. 2241–2246, 2007.
- [7] M. Goesele, B. Curless, and S. Seitz, "Multi-view stereo revisited," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, pp. 2402–2409, July 2006.
- [8] A. Hornung and L. Kobbelt, "Hierarchical volumetric multi-view stereo reconstruction of manifold surfaces based on dual graph embedding," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, pp. 503–510, July 2006.
- [9] G. Vogiatzis, P. Torr, and R. Cipolla, "Multi-view stereo via volumetric graph-cuts," in *Proceedings of Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, pp. 391–398, San Diego, Calif, USA, July 2005.
- [10] S. N. Sinha and M. Pollefeys, "Multi-view reconstruction using photo-consistency and exact silhouette constraints: a maximum-flow formulation," in *Proceedings of 10th IEEE International Conference on Computer Vision (ICCV '05)*, pp. 349–356, October 2005.
- [11] V. Kolmogorov and R. Zabih, "Generalized multi-camera scene reconstruction using graph cuts," in *Proceedings of the European Conference on Computer Vision (ECCV '02)*, vol. 3, pp. 82–96, 2002.
- [12] H. Jin, S. Soatto, and A. J. Yezzi, "Multi-view stereo reconstruction of dense shape and complex appearance," *International Journal of Computer Vision*, vol. 63, no. 3, pp. 175–189, 2005.
- [13] O. Faugeras and R. Keriven, "Variational principles, surface evolution, PDE's, level set methods, and the stereo problem," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 336–344, 1998.
- [14] S. Soatto, A. Yezzi, and H. Jin, "Tales of shape and radiance in multi-view stereo," in *Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV '03)*, pp. 974–981, Nice, France, October 2003.
- [15] H. Jin, S. Soatto, and A. Yezzi, "Multi-view stereo beyond Lambert," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '03)*, vol. 1, pp. 171–178, Madison, Wis, USA, July 2003.
- [16] M. Lhuillier and L. Quan, "A quasi-dense approach to surface reconstruction from uncalibrated images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 418–433, 2005.
- [17] Y. E. Duan, L. Yang, H. Qin, and D. Samaras, "Shape reconstruction from 3D and 2D data using pde-based deformable surfaces," in *Proceedings of the European Conference on Computer Vision (ECCV '04)*, vol. 3, pp. 238–251, May 2004.
- [18] C. Hernandez and F. Schmitt, "Multi-stereo 3D object reconstruction," in *Proceedings of 3D Data Processing Visualization and Transmission*, pp. 159–166, Padova, Italy, June 2002.
- [19] Y. Furukawa and J. Ponce, "Carved visual hulls for image-based modeling," in *Proceedings of the European Conference on Computer Vision (ECCV '06)*, vol. 3951, pp. 564–577, Graz, Austria, May 2006.
- [20] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multi-view stereopsis," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, July 2007.
- [21] M. Habbecke and L. Kobbelt, "A surface-growing approach to multi-view stereo reconstruction," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, June 2007.
- [22] P. Merrell, A. Akbarzadeh, L. Wang, et al., "Real-time visibility-based fusion of depth maps," in *Proceedings of the 11th IEEE International Conference on Computer Vision (ICCV '07)*, Rio de Janeiro, Brazil, October 2007.
- [23] D. Bradley, T. Boubekeur, and W. Heidrich, "Accurate multi-view reconstruction using robust binocular stereo and surface meshing," in *Proceedings of the 26th IEEE Conference*

- on *Computer Vision and Pattern Recognition (CVPR '08)*, Anchorage, Alaska, USA, July 2008.
- [24] L. Quan, J. Wang, P. Tan, and L. Yuan, "Image-based modeling by joint segmentation," *International Journal of Computer Vision*, vol. 75, no. 1, pp. 135–150, 2007.
 - [25] "The multi-view stereo evaluation," <http://vision.middlebury.edu/mview>.
 - [26] A. Laurentini, "The visual hull concept for silhouette-based image understanding," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 2, pp. 150–162, 1994.
 - [27] W. E. Lorensen and H. E. Cline, "Marching cubes: a high resolution 3D surface construction algorithm," *Computer Graphics*, vol. 21, no. 4, pp. 163–169, 1987.
 - [28] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
 - [29] H. K. Zhao, S. Osher, and R. Fedkiw, "Fast surface reconstruction using the level set method," in *Proceedings of IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pp. 194–201, Vancouver, Canada, July 2001.
 - [30] H. K. Zhao, S. Osher, B. Merriman, and M. Kang, "Implicit and nonparametric shape reconstruction from unorganized data using a variational level set method," *Computer Vision and Image Understanding*, vol. 80, no. 3, pp. 295–314, 2000.
 - [31] V. Caselles, R. Kimmel, G. Sapiro, and C. Sbert, "Three dimensional object modeling via minimal surfaces," in *Proceedings of the European Conference on Computer Vision (ECCV '96)*, vol. 1, pp. 97–106, Cambridge, UK, April 1996.