Automatic Surface Scanning of 3D Artifacts

Shiaofen Fang, Basil George and Mathew Palakal

Department of Computer and Information Science
Indiana University Purdue University Indianapolis, Indianapolis, IN 46202

Abstract—**This paper describes an automatic 3D surface scanning technique using a 3D scanner. It allows the acquisition of a complete surface model of a 3D artifact without any manual registration and human interference. A two-pass approach is applied using a rotary table. In the first pass, a sequence of 2D images of the artifact are collected using a small rotation step. An image similarity measure is then taken to compare adjacent images to assess the differences between consecutive images to establish the optimal scanning angles. In the second pass, 3D scans are taken using the scanning angles derived in the first pass. An Iterative Closest Point (ICP) algorithm is employed to provide precise data alignment of neighboring scans. A new space encoding algorithm is developed to filter the over sampling areas for smoother polygonization. This technique is being implemented in a digital library project for the reconstructions of 3D computer models for artifacts from a collection in several Museums.**

Index Terms—**Surface Scanning, Image Analysis, Digital Library.**

I. INTRODUCTION

Laser scanning technology is becoming increasingly popular recently as the scanning devices become more reliable, cheaper, faster and more portable. One of the most common applications of this technology is the acquisition of 3D models (as polygon meshes) from real world 3D objects. One prominent example is the Digital Michelangelo project [1], which includes the complete scans of a collection of statues by Michelangelo. An important application of 3D laser scanning is digital library or virtual museum, where valuable 3D artifacts may be digitized and displayed in 3D for web access and virtual exhibition. Other applications include reverse engineering, rapid prototyping, archaeology, forensic imaging, etc.

Because laser scanner is essentially a rangefinder, it can only scan one projection plan at a time. Multiple scans are therefore needed to reconstruct a complete surface of the 3D object. The scans will need to be registered, aligned and stitched together to form a single polygon mesh surface. There are two issues involved in this process: view selection and stitching. In most scanning system, the decision about the number of scans taken and where (angles) these scans are taken is often made arbitrarily. For automatic scanning using a rotary table, for instance, this would mean a uniform rotation with a pre-determined rotation step. Since the geometric complexity of an object is not uniform in all directions, this may lead to either over-scanning (more scanning time and unnecessary stitching errors) or under-scanning (missing features). A more efficient approach would require a view selection process that automatically and quickly selects the best views for efficient and adaptive scanning.

A second issue that needs to be addressed is the stitching of the neighboring scans. Most existing stitching techniques require manual assistance to provide initial alignment. Markers attached to the object are often used to facilitate the initial matching of neighboring scans. Applications such as digital library, however, usually involve large numbers of artifacts that require a speedy scanning process. As a result, manual or even semi-automatic stitching is often not acceptable. Markers are often not feasible either, since most of these artifacts are highly valuable and cannot be touched with markers. Hence, automatic 3D stitching technique is highly desirable. Unfortunately, a general automatic 3D stitching is very difficult and error prone. This is because global optimization is difficult to achieve without a good initial alignment.

In this paper we propose an automatic 3D scanning technique using a rotary table for 3D artifacts. Our technique includes a image analysis based view selection method and an automatic stitching algorithm. This method is suitable for the types of artifacts that can be scanned through a 360 degree rotation. Our goal is to build a system that can automatically construct a polygon mesh surface of an artifact without human interference or markers on the object. This system is currently being used in a digital library project for reconstructing 3D computer models by scanning the artifacts from collections in several Museum.

In the following, we will first, in Section 2, discuss previous literatures related to 3D scanning and stitching, and then describe the systems setup and hardware configurations in Section 3. The view selection method will be discussed in Section 4, and the stitching algorithm will be presented in Section 5. In Section 6, we will provide several scanning results, and we will conclude the paper with a few final remarks and future work in Section 7.

II. RELATED WORK

A previous work in automatic scanning is presented in [2], A robot arm was used to position and control the artifact. A uniform scanning sequence is initially applied with a few added scans afterwards based the stitching results. In [3], a preliminary

uniform scanning is first applied to build a coarse model, which is then used for the planning of finer scanning. Similar work, mostly in the robotics field for scanning path planning, can be found in [4-6].

Prior work related to 3D stitching can be classified into pairwise alignment methods and global alignment methods. In pairwise methods [7,8], two adjacent scans are aligned at a time. An initial alignment needs to be estimated first, and then refined using the Iterative Closest Point (ICP) algorithm. The initial estimate often requires manual assistance or markers on the object. Polygon stitching or point cloud filtering is not discussed. In global alignment [9,10], multiple scans are registered at the same time to achieve a global optimization. The Iterative Closest Point algorithm again is often used to carry out the optimization process. These methods also require manual assistance for initial alignment, and are generally less robust.

Curless and Levoy (1996) [11] proposed a volumetric method for integration of the range images which consists of a cumulative weighted signed distance function which works with one range image at a time to first scan-convert it to a distance function and then combine this with data already acquired. Two other volumetric methods were proposed in systems presented in [12,13].

Soucy and Laurendeau [14] describe a method to classify overlapping data regions using Venn diagrams, followed by re-parameterization and merging of regions. Rutishauser, et al [15] establishes consensus surface positions using errors along the sensor's lines of sight and then perform a re-tessellation to include redundant data points. These methods are considered structured data algorithms and several parametric approaches have been proposed, most of them operating on range images that have been polygonized. Our approach may be classified as an unorganized points algorithms. A key benefit of the unorganized points algorithms is the fact that they do not make any prior supposition about the connectivity of the points.

These methods also require manual assistance for initial alignment, and are generally less robust. Also in these methods, scans are generally first stitched and then polygonized. This might cause the loss of significant features on the object due to removal of the points during noise point removal.

The technique presented in this paper is a pairwise technique. But the initial alignment is achieved automatically using a rotary table. A novel point-filtering method is used to ensure uniform sampling for better polygon mesh surfaces.

III.   SYSTEM SETUP

Our system comprises of a turntable onto which an object is placed, and a 3D scanner that can take one range scan at a time. The turntable can be automatically rotated by a fixed angle at each step. A scan is initiated automatically each time the tunable is rotated until a 360 degree circle is complete. The stitching algorithm is first applied to the first two scans to generate a master scan. The subsequent scans will all be stitched with the master scan one at a time to eventually form a complete surface.

3.1 The Rotary Stage

The turntable used in the project is the RT-12 Rotary Positioning Table along with an MD2-A stepper motor which is controlled using an MD-2 stepper motor driver (See Figure 1(a)).

The rotary table has a 1 degree rotational resolution and can rotate at a speed of 45 degrees per second. The turntable is controlled by writing values to the printer port to which it is connected. Values written to this port are used to turn the stepper motor attached to the table on or off or make it rotate in either direction by the specified number of steps. Rotating the stepper motor 135 steps makes the turntable rotate by 20 degrees in the clockwise direction. Attached to the turn table is a marker that hangs exactly above the center of the turntable. This marker can be moved vertically and can be heightened or lowered so that it is visible in scans. The marker is a disc of plastic, one centimeter in diameter that always faces the scanner.
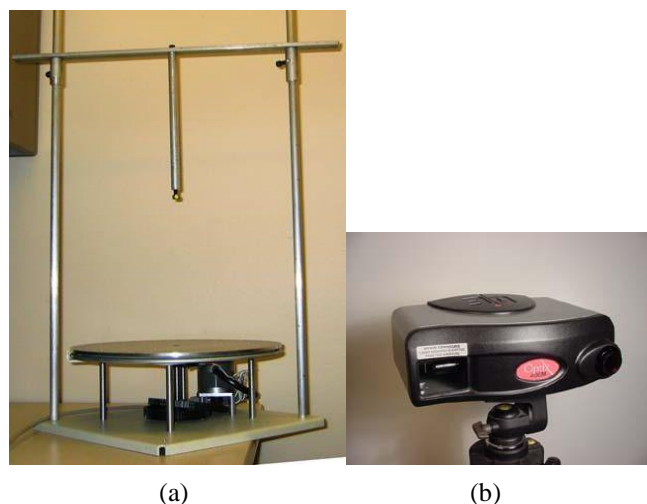

(a)   (b)

Fig. 1. (a) The rotary table with the marker positioning rig; (b) The Optix 400M scanner.

3.2 The Scanner

The 3D scanner used to perform the scanning is an Optix 400M (Fig. 1(b)).The scanner is connected to an external digital camera in order to provide high quality color texture information.

The scanner is controlled using a COM object installed by scanner software that allows most languages to communicate with the scanner. The Com object has methods within it which may be called to set various parameters of the scan, such as vertical resolution, which sets the number of vertical scan lines per scan; gain control, which specifies the sensitivity of the 3D scanner's camera; and Z Range cutoff planes which specify the scanner's depth of vision.

The scanner has a resolution of 0.007" at a 12" distance and 0.015" at a 26" distance, and a field of view of 30 degree.

IV.   VIEW SELECTION

A two-pass scanning method is employed. In the first pass, a set of scanning views (rotation angles) are selected by collecting and analyzing a sequence of 2D images of the artifact based on image similarity. This view selection approach differs from the traditional view planning methods in that view planning process is done with 2D images, which are taken using an external digital camera attached (and pre-calibrated) to the scanner.

   The purpose of this view selection process is to determine the optimal set of rotation angles at which 3D scans are taken that would provide the complete coverage of the surface features of the object. Occlusion analysis can be applied to these images to determine the critical views. But occlusion analysis using 2D images without 3D geometric information is unreliable and time-consuming. For a fast estimate of the scanning angles, we believe a simpler image similarity comparison would be sufficient, particularly if a pre-defined upper bound of the angle step is used. The 2D image acquisition process in this first pass is straightforward. The rotary table will rotate with a pre-defined small rotation step (5 degree in our experience), and a high resolution digital image is taken at every rotation step. With these images, we implemented two image similarity measures: edge count comparison and texture analysis of edge images.

   In edge count comparison, Edge detection using Sobel operator is first applied to each image. Texture information has been widely used in image retrieval by evaluating image similarities using texture features. We implemented a texture feature analysis for image similarity measure similar to the method given in [16]. The texture analysis is applied to images after edge detection, as it provides better edge pattern analysis. Fig. 2 shows some examples of images before and after the edge detection. Edges are then segmented to provide edge segments. The edge segments in an image are then labeled. The edge segment difference is then computed for two adjacent images as the sum of the numbers of edge segments present in one image that is not present in the other and vice versa.

   The similarity measurements generated from the above technique is used as the basis for view selection. This is done by first choosing an image difference threshold, δ , which is defined as the maximum acceptable difference measure between the images of two consecutive scans. The scanning angles are then selected such that the cumulative image difference between two consecutive scans is less than δ .

   Fig. 3(a), (b) and (c) show the results of view selection with the same δ values using the edge segment count comparison for objects of varying complexity. Fig. 3(a) shows the view selection result in the camel sculpture which has a uniformly complex shape, Fig. 3(b) shows the view selection result in a cup with a handle which has a non-uniform shape and Fig. 3(c) shows the view selection result in a water bottle which has a uniformly simple shape.

   As is expected, for the camel sculpture, the number of suggested scans and suggested scanning angles are the greatest, due to the complexity of its shape. The plot for the cup shows an increased number of suggested scans at the location of the

handle of the cup while the plot for the water bottle shows a uniform scan angle of forty degrees which is set as the maximum allowable difference in suggested scan angles.
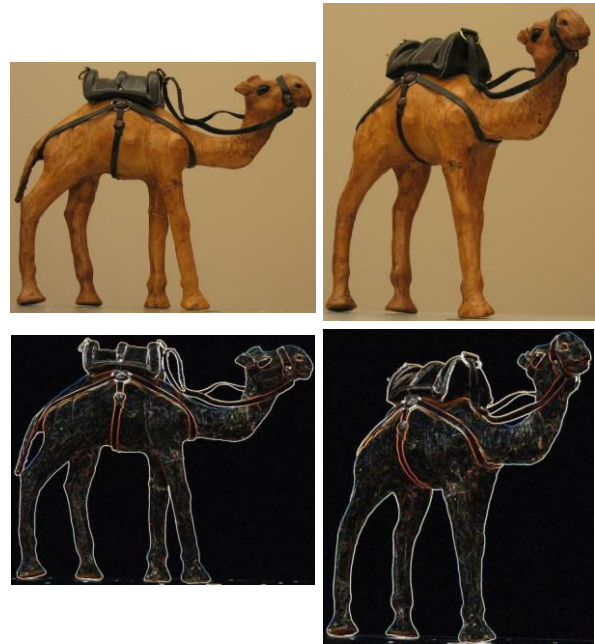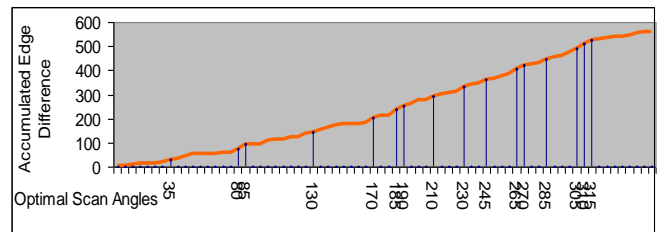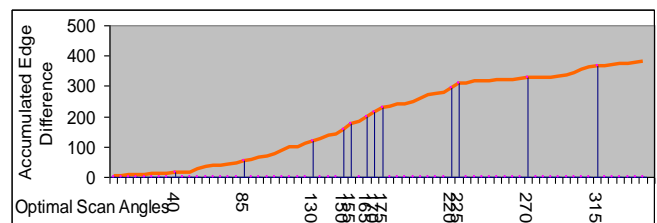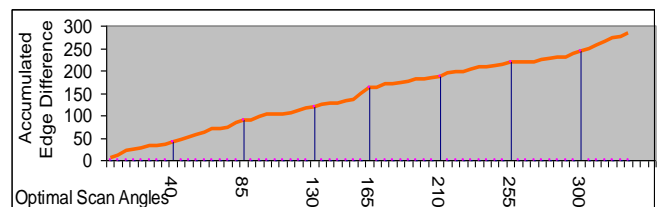

Fig. 2. Examples of images of a camel before and after edge detection

(a)

(b)

(c)

Fig. 3. View selection: Accumulated Edge segment count comparison for (a) uniformly complex (b) non-uniform and (c) uniformly simple shapes.

V.   STITCHING

Based on the views selected in the first pass, a scan will be taken each time the tunable is rotated. Each scan produces a set of dense points, or a point cloud, that can be polygonized to generate a polygon mesh surface. But since polygon stitching is difficult, we decide to merge all the point clouds together first and then polygonize the final point cloud only once. There are three steps in this process: (1) initial alignment; (2) refinement, and (3) point filtering.

5.1 Initial Alignment

The point clouds are processed sequentially in the order they were received. A marker is defined on the central axis of the rotary table as the highest point of the system. When a point cloud is processed, the highest point is identified, which represents the rotational axis.

The coordinates of the marker point provide a reference point of the rotation axis. Since the rotation angle is known from the rotary table, a rigid transformation can be easily derived by concatenating a translation matrix and a rotation matrix. This transformation will then be applied to the second point cloud to generate an initial alignment of the two scans. Fig. 4 shows an example of two scans before and after the initial alignment.
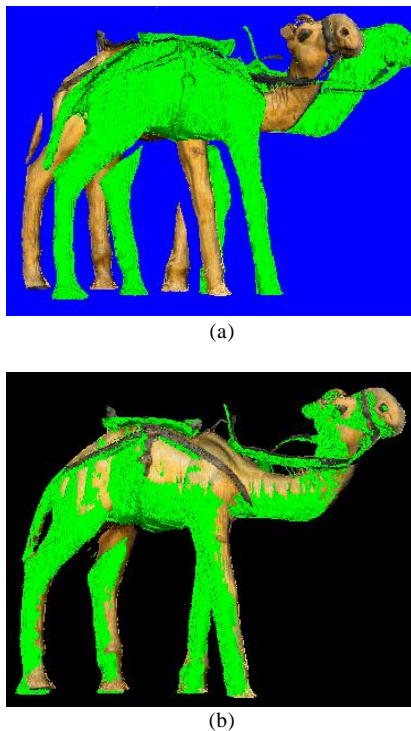

(a)

(b)

Fig. 4 The point clouds of two scans (a) before alignment; (b) after alignment

5.2 Refinement

There are errors in the initial alignment. The errors come from two sources. First, there is error in detecting the reference point. Second, the rotary table has a rotational resolution of 1 degree. Therefore the rotation itself carries some error. The residual distances between the point clouds results in an inaccurate rendering of the surface. In order to rectify this problem, the point clouds have to be further aligned using a finer alignment process.

The refinement is done using the popular Iterative Closest Point (ICP) algorithm [17,7,8]. The iterative closest point (ICP) is an algorithm utilized to match two clouds of points. It iteratively calculates the transformation (translation, and rotation) between two raw point clouds. The algorithm works in three phases: it first establishes a correlation between pairs of features in the two structures that are to be aligned based on their closeness to each other; it then estimates the rigid transformation that best maps the second cloud of the pair onto the first such that the mean squared error (MSE) given by the sum of the squared distances between the corresponding points is minimized; and it then applies that transformation to all features in the second structure. These three steps are then reapplied until convergence is concluded when the MSE falls below a given threshold or no further improvement can be achieved. The cost function is defined as the sum of the least square distances from the vertices of one face dataset to the other face surface, i.e.

$$f(M,T) = \sum_i \|x_i - Mp_i - T\|^2 \qquad (1)$$

where M is a 3 by 3 rotation matrix and T is a translation matrix. Although simple, the algorithm works quite effectively when given a good preliminary estimate as by the initial alignment.

5.3 Point Filtering

Two adjacent scans often have significant overlaps after alignment. Directly constructing polygons from these point clouds will lead to in improper triangulations along the boundaries of the overlapping regions. Since there are twice as many points in the overlapping regions as the rest of the areas, the density of the polygon mesh is also very uneven, leading to difficulties in later processing and rendering of the surface. To correct this problem, a filtering procedure is designed to remove the extra points to ensure that the sampling density in the overlapping regions is the same as the rest of the surface.

Without any prior assumption about the orders of the points in each scan, this problem would require excessive search to identify neighborhood information so that sampling density can be calculated. To expedite this range searching process, we employ a spatial encoding technique to quickly remove points that are too close together.

For each combined point cloud, the bounding box of the point cloud is subdivided into a regular 3D grid as a volumetric space. The size of each voxel represents the desired density of the point cloud. The set of points will be hashed into the 3D grid one at a time, using the X, Y, and Z coordinates of the point as the hashing values to place the point into the appropriate voxel. Each voxel can store at most one point. If a point is hashed into a voxel that already has a point in it, one of the two points will be removed. Usually, the point that is closer to the center of the

voxel should be retained. This way, we only go through the point set once to generate a uniformly sampled point cloud. Fig. 5 shows two point clouds before and after the filtering in a 2D space. After all point clouds are aligned and filtered, a polygon mesh will be constructed using a Delaunay triangulation to form the final surface model.
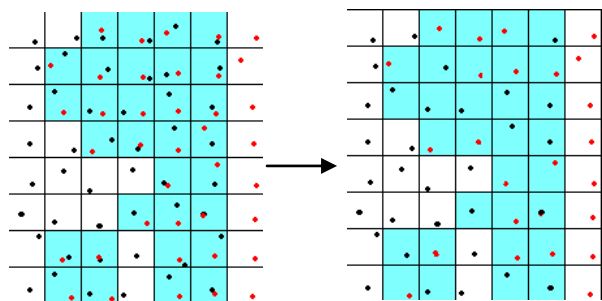

Fig. 5. Point Cloud Filtering

VI. IMPLEMENTATION

The scan automation application is written in C++ using Microsoft Foundation Classes for the user interface which utilizes a slider to obtain user input. Figure 6 shows the overall structure of the automatic scanning algorithm.
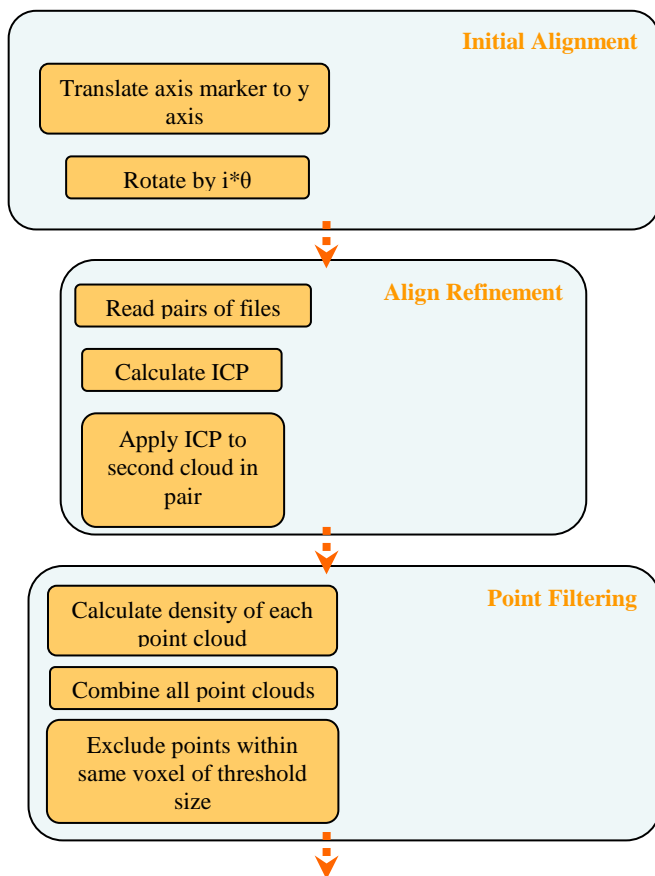

Figure 6: Flow chart of the scan processing algorithm

The technique developed in this research will be applied to a digital library system CLIOH. As a subproject in CLIOH,

artifacts from various museums such as the Museum of Macedonia are scanned and processed to construct 3D computer models for preservation, virtual museum, and online exhibition applications.

In Fig. 7, the example in Fig. 4 is shown with the original object, the stitched point cloud, and the final rendering result. Fig. 8 and 9 show two additional scanning examples.
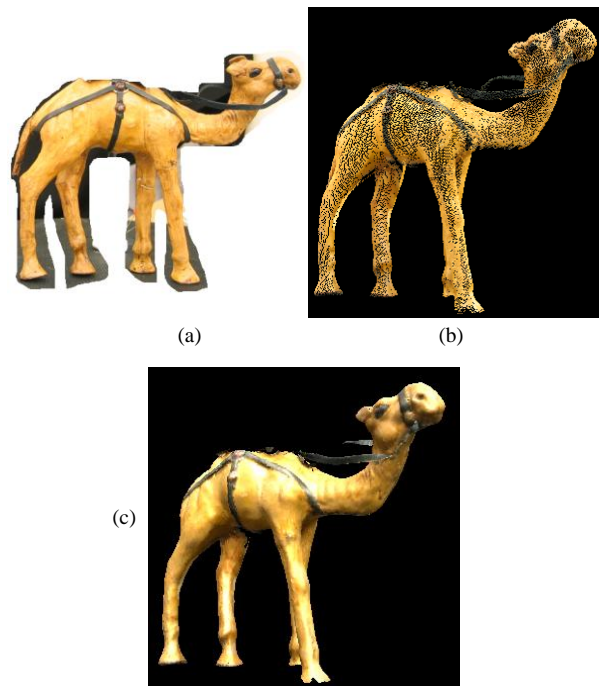

(a)   (b)

(c)

Fig. 7. (a) The original object; (b) The stitched point cloud; (c) The final rendering of the reconstructed model.

(a)   (b)

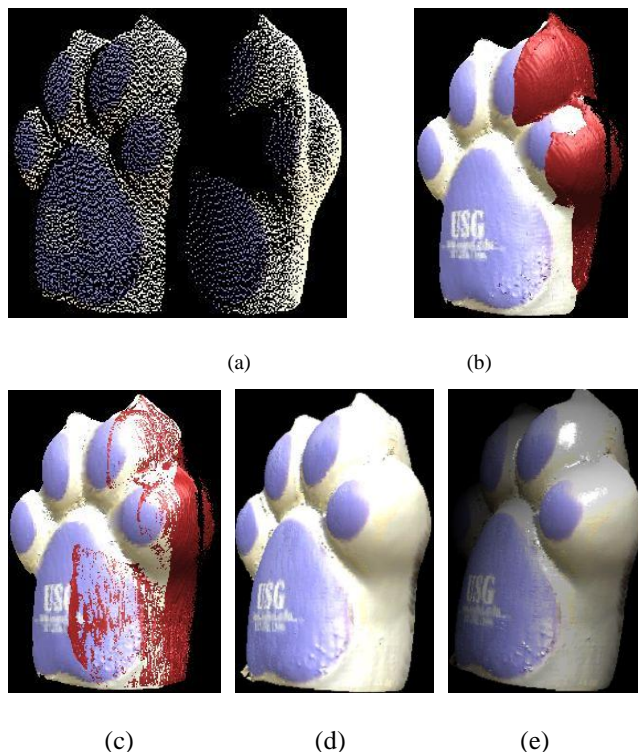(c)   (d)   (e)

Fig. 8. (a) Two separate scans; (b) After the initial alignment; (c) After refinement; (d) Final rendering; (e) The original object.
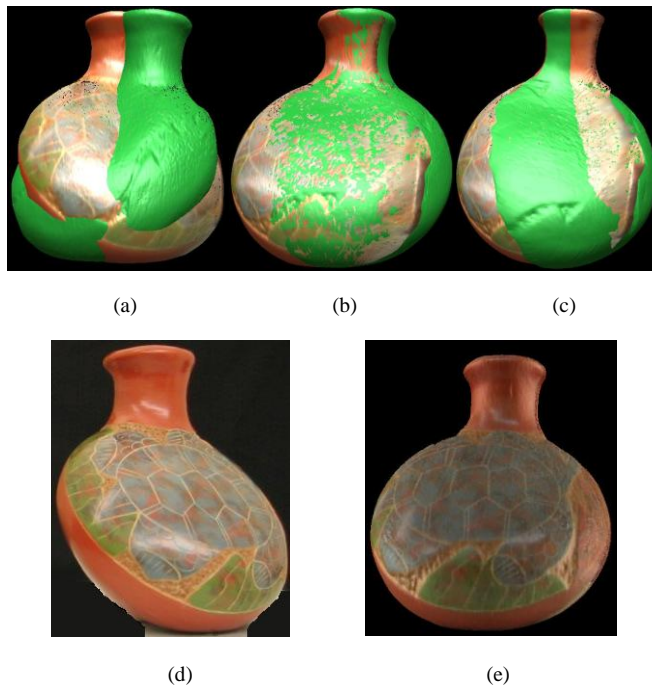
Fig. 9. (a) Two separate scans; (b) After the initial alignment; (c) After refinement; (d) Final rendering; (e) The original object.

VII.   CONCLUSION

This paper explores some of the issues involved in automatically digitizing real world objects, primarily, the issues of view selection and the stitching of dense point clouds. The 2D image based view selection is fast and easy to implement, and provides a quantitative guidance for adaptive scanning. The stitching process is highly accurate and automatic.

   In the future we would like to explore more sophisticated image analysis tools for more accurate occlusion analysis in view selection. We will also try to extend this technique to larger scale scanning problems such as multiple objects and real-world scenes.

REFERENCES

[1] M. Levoy, et al, "The Digital Michelangelo Project: 3D Scanning of Large Status," Proceedings of SIGGRAPH'00, pp. 131–144, New Orleans, July, 2000.

[2] M. Callieri, A. Fasano, G. Impoco, P. Cignoni, R. Scopigno, G. Parrini, G. Biagini, "RoboScan: An Automatic System for Accurate and Unattended 3D Scanning," 3dpvt, pp. 805-812, Second International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT'04), 2004.

[3] W. Scott, G. Roth, and J-F Rivest, " Performance-Oriented View Planning for Automatic Model Acquisition. Proc. 31st International Symposium on Robotics, 314-319, 2000.

[4] C. Connolly, "The Determination of Next Best Views", In CRA'85, 432-435.

[5] J. Maver and R. Bajcsy, "Occlusions as a Guide for Planning the Next View, PAMI, 15(5), 417-433, 1993.

[6] A. Papadopoulos-Orfanos and F. Schmitt, "Automatic 3D Digitization Using a Laser Rangefinder with a Small Field of View", in 3DIM97, 1997.

[7] P. Besl and N. McKay. A method for registration of 3d shapes. Trans. Patt. Anal. Mach. Intell. 14(2), 239–256, 1992

[8] Y. Chen and G. Medioni, "Object Modeling by Registration of Multiple Range Images, Image and Vision Computing, 10(3), 145-155, 1992.

[9] R. Bergevin, M. Soucy, H. Gagnon, D. Laurendeau, "Towards a General Multi-View Registration techniqueakso", IEEE Tran. Pattern recognition and Machine Intelligence, 18(5), 540-547, 1996.

[10] A. Fusiello, U. Catellani, L. Ronchetti, V. Murino, "Model Acquisition by Regisration of Multiple Acoustic Range Views", Proc. EECV, 805-822, 2002.

[11] Curless, B., and Levoy, M., 1996. A volumetric method for building complex models from range images. Proc. Of SIGGRAPH'96, August 4-9, New Orleans, pp. 303-312

[12] Pulli, K., Duchamp, T., Hoppe, H., McDonald, J., Shapiro, L.,and Stuetzle, W., 1997. Robust meshes from multiple rangemaps. *Int. Conf. on 3D Digital Imaging and Modeling*, Ottawa,May 12-15, pp. 205-211

[13] Hilton, A., and Illingworth, J., 1997. Multi-resolution geometricfusion. *Int. Conf. on 3D Digital Imaging and Modeling*, Ottawa,May 12-15, pp. 181-188.

[14] M. Soucy and D. Laurendeau, "Multi-Resolution Surface Modeling from Multiple Range Views," *Proc. IEEE Computer Visiion and Pattern Recognition Conf. '92,* pp. 348-353, 1992.

[15] M. Rutishauser, M. Stricker, and M. Trobina. Merging range images of arbitrarilyshaped objects. In *Proceedings 1994 IEEE Computer Society Conference onComputer Vision and Pattern Recognition*, pages 573–580, June 1994.

[16] Selim Aksoy and Robert M. Haralick, "Using Texture in Image Similarity and Retrieval", in Texture Analysis in Machine Vision, M. Pietikainen, Ed., vol. 20, pp. 129-149, 2000.

[17] M. Greenspan and G. Godin. A nearest neighbor method for efficient ICP. In Proceedings of the 3rd International Conference on 3-D Digital Imaging and Modeling (3DIM01), 2001.

Shiaofen Fang is a Professor of Computer Science with the Department of Computer and Information Science at Indiana University Purdue University Indianapolis (IUPUI). His research interests include computer graphics, scientific and biomedical visualization, geometric modeling, and medical imaging.

Basil George graduated with a MS degree in Computer Science from Indiana University Purdue University Indianapolis. He is currently working as a … at …. His research interests are in scientific visualization and computer graphics and animation.

Mathew Palakal is a Professor of Computer Science and Informatics at Indiana University Purdue University Indianapolis (IUPUI). He also serves as the Associate Dean of School of Informatics at IUPUI. His research interests include information management, intelligent systems, bioinformatics and neural networks.