Automated 3d recording of archaeological pottery[1]

Martin Kampel and Robert Sablatnig

Pattern Recognition and Image Processing Group
Institute for Automation, Vienna University of Technology,
Favoritenstr. 9/183/2, A-1040 Vienna, Austria
Fax: +43 (1) 58801 183 92; email: {kampel, sab}@prip.tuwien.ac.at

ABSTRACT

At excavations a large number of sherds of archaeological pottery is found. Since the documentation and administration of these fragments represent a temporal and personnel effort, we construct a computer aided documentation system for archaeological fragments to form the basis for a subsequent semi-automatic classification and reconstruction.

In archaeology the determination of the exact volume of arbitrary vessels is of importance since this provides information about the manufacturer and the usage of the vessel. The technique used for the 3d-acquisition is shape from silhouette. The major problem of the 3d surface reconstruction using a turntable is the varying resolution in direction to the camera due to the varying rotation of object points in respect to the rotational axis of the turntable.

Therefore, the optimal set of views for capturing the viewable object surface is a set of angles, which guarantees a uniform object resolution. In this paper we present a technique, which estimates the next angle dynamically, depending on the entropy of the silhouette actually acquired. The relation proposed guarantees a uniform object resolution on one side and a minimal number of acquisition steps on the other side. The method has been tested on synthetic and real data with reasonably good results. The paper concludes with a presentation of results and an outlook on future work.

KEYWORDS: 3d acquisition, automated recording and archivation

INTRODUCTION

Ceramics are one of the most widespread archaeological and are a short-lived material. This property helps researchers to document changes of style and ornaments. Therefore, ceramics are used to distinguish between chronological and ethnic groups. At excavations a large number of ceramic fragments, called sherds are found. These fragments are photographed, measured, drawn (called documentation) and classified. Up to now documentation and classification have been done manually which means a lot of routine work for archaeologists and a very inconsistent representation of the real object. First, there may be errors in the measuring process. (Diameter or height may be inaccurate), second, the drawing of the fragment should be in a consistent style, which is not possible since a drawing of an object without interpreting it is very hard to do.

Since the documentation and administration of these fragments represent a temporal and personnel effort, we construct a documentation system for archaeological fragments to form the basis for a subsequent semi-automatic classification and reconstruction [8, KS99]. Because the conventional method [12, 14] for documentation is unsatisfactory, we developed an automated 3d-object acquisition with respect to archaeological requirements.

The optimal set of views for capturing the complete object surface depends on the unknown shape of the surface and the degree of freedom of the equipment used for the acquisition. Our acquisition system consists of one turntable, two CCD-cameras and two lasers. The turntable used allows movements of the objects of one degree of freedom in controlled steps. Therefore, the determination of the next sensor position will be reduced to the estimation of the next rotation angle. The technique used for the acquisition is shape from silhouette [16, 13]. It is a robust approach for the automatic reconstruction of the 3d shape, which extracts the silhouette from the input images in a step and reconstructs a volume cone for each silhouette in a second step. Finally, the intersection of the volume cones results in the bounding volume [20]. Shape from Silhouette offers a unique 3d-voxel based data representation even for objects with a complex surface like the handle of a vessel.

The major problem of the 3d surface reconstruction using a turntable is the varying resolution in direction to the camera due to the varying rotation of object points in respect to the rotational axis of the turntable [10]. Therefore, the optimal set of views for capturing the viewable object surface is a set of angles, which guarantees a uniform object resolution.

In order to guarantee a uniform object resolution, we have to calculate the next angle dynamically, depending on the entropy of the silhouette actually acquired. High entropy

[1]This work was supported in part by the Austrian Science Foundation (FWF) under grant P13385-INF, the European Union under grant IST-1999-20273 and the Federal Ministry of Education, Science and Culture.

means that the selection of a wide angle leads to the loss of information. Therefore, we need a relation between the entropy and the angle, which guarantees a minimal loss of information for the next acquisition step.

In this paper we present a technique, which estimates such a relation in order to calculate the next viewing angle dynamically. The relation proposed guarantees a uniform object resolution on one side and a minimal number of acquisition steps on the other side. The paper is organized as follows: In Section 2 we describe the acquisition system used and in Section 3 we explain how we estimate a uniform object resolution with minimal acquisition steps. Experimental results are described in Section 4 and we conclude with a summary and future work in section 5.

ACQUISITION SYSTEM

The acquisition system consists of the following devices:

- 1 turntable with a diameter of 50 cm, which can be rotated about the z-axis, used to move the object of interest through the acquisition area. The position desired can be specified with an accuracy of 0.05 degrees
- 2 red lasers to illuminate the scene for the next view planning process. One mounted on the top (distance to rotation plane is 45 cm), one beside the turntable (distance to the rotation center is 48 cm). Both lasers are extended with cylindrical lenses to spread the laser beam into one illuminating plane. The laser light plane intersects with the object surface, forming one laser stripe.
- 1 CCD-camera (b/w) used for next view planning (shape from structured light - **sfsl**), with a 16 mm focal length, a resolution of 768x572 pixels, and a distance of 40 cm to the rotation center. The angle between the camera normal vector and the rotation plane is approximately 45 degrees. A frame grabber card is used to connect the camera to a PC.

An important issue is the illumination of the object observed, which should be clearly distinguishable from the background, independent from the object's shape or the type of its surface. For that reason back-lighting [1] is used. A large (approximately. 50x40 cm) rectangular lamp is put behind the turntable (as seen from the camera) Figure 1.

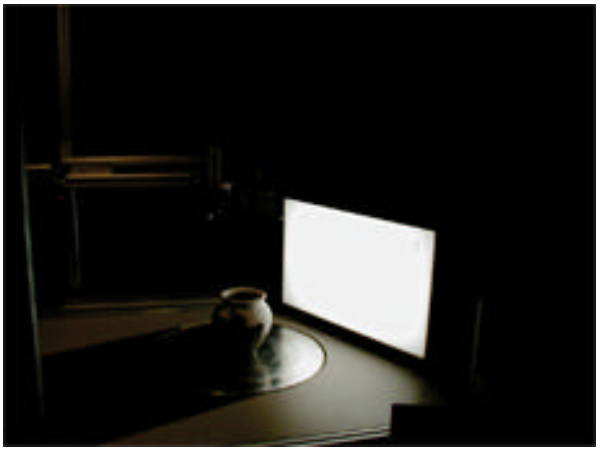


Figure 1: back-lighting for sfs

- 1 CCD-camera (b/w) for the 3d-acquisition process (shape from silhouette - **sfs**), with a 16 mm focal length, a resolution of 768x572 pixels, and a distance of 70 cm to the rotation axis. The angle between the camera normal vector and the laser plane is approximately 90 degrees. A frame grabber card is used to connect the camera to a PC.
- 1 Intel Pentium PC under Linux operating system.

Figure 2 depicts the acquisition system. Prior to any acquisition, the system is calibrated in order to determine the inner and outer orientation of the cameras and the rotational axis of the turntable. The calibration method used was exclusively developed for the Shape from Silhouette algorithm presented and it is described in detail in [2, 18].

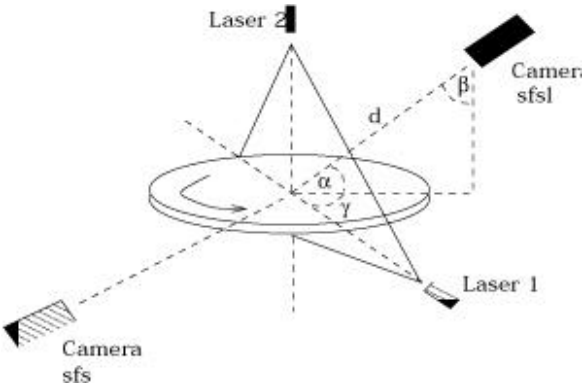


Figure 2: Acquisition system

SENSOR PLANNING

The estimation of the next best view (NBV) [6] can be divided into three categories:

- Minimization of occlusions: Occlusions are interpreted as filled polygons. Then, a set of

sensor positions and angles relative to the objects surface are computed for each pixel of these polygons. The result of this step is a set of intervals from which the polygon pixel is visible. Decomposing these intervals gives the next sensor position. Whereas [6, 5] analyses range images, [11] uses volume models, to solve the NBV-Problem.

- Analyzing the geometric properties of the surface: In [7] the surface structure is given by triangulated surface points. The surface is completed by stepwise refinement. Regions, which show highly structured parts will be scanned with higher density than regions which show low structured parts.

- Heuristic search and objective functions: The set of next sensor positions is reduced by applying a heuristic search. The best position is estimated by maximizing an objective function [15, 19].

The result of the first category approaches is information about the position and orientation of occlusions in the scene. This information is two or three-dimensional. To reduce occlusions, a system, which allows a movement of more than one degree of freedom, is needed. The turntable allows a movement about the z-axis and the next best sensor position will be estimated by the analysis of the surface structure. The notion of the next "best" sensor position can be defined in two ways:

- The system should achieve a minimal number of acquisition positions and steps to reconstruct the object of interest.

- Computing those acquisition positions and directions, which give the best reconstruction results.

In this work, we develop a system, which achieves a minimal number of acquisition steps by accomplishing given accuracy requirements.

Adaptive Image Acquisition

One of the major problems of the 3d surface reconstruction using a turntable is the varying resolution in the direction to the camera, due to the varying distance of object points to the rotational axis of the turntable [3, 4]. The varying resolution leads to a loss of surface features. Figure 3 shows two examples to the loss of information. Sampling the object with a constant angle of 20 degrees (Figure 3a) we loose one corner of the square. On the other side, sampling the object with a lower constant angle, the loss of information is less (Figure 3b) than sampling with a higher constant angle. Therefore, the accuracy of reconstructions can be improved by decreasing the sampling angle, whereas the effort of the acquisition process will be increased.

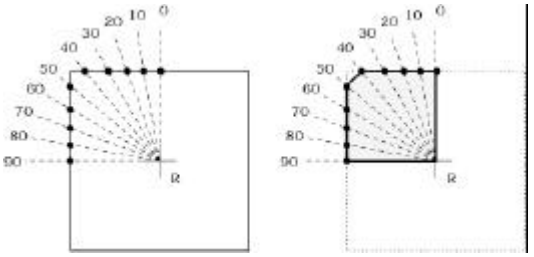


Figure 3a: Sampling with equiangular steps (10 degrees)

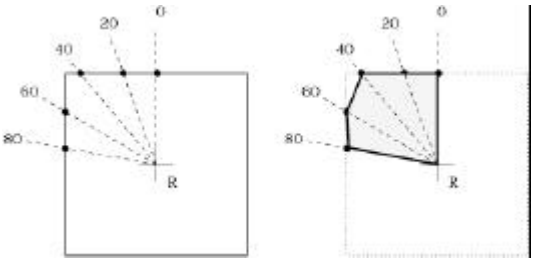


Figure 3b: Sampling with equiangular steps (20 degrees)

Complexity

The maximum number of acquisition steps depends on the camera resolution. Therefore the minimum angle is given by

$$f_{\min} = \arctan\frac{r}{A} \qquad \textbf{Equation 1}$$

where r is the distance of one surface point to its center of revolution and A is the resolution of the camera. The maximum number of acquisition steps is then given by

$$N_{\max} = \frac{360}{f_{\min}} \qquad \textbf{Equation 2}$$

Let n be the desired number of maximal acquisition steps. The complexity of the NBV-problem is then given by

$$C_{NBV} = \sum_{i=1}^{n} \binom{N_{\max}}{i} \qquad \textbf{Equation 3}$$

Equation 3 is viewed as the more general set theory problem of finding a minimal number of subsets that completely covers a set [17], which is in the class of NP complete problems and therefore only solvable with polynomial effort.
The limitation to a directed movement and the definition of a maximum angle $f_{\max}$ reduces the complexity, because the number of possible next positions will be decreased.

Computing the Next Sensor Position

In order to estimate the sensor position, we have to

calculate the distance of the captured surface points relative to the axis of rotation. The calculation of the next rotational angle is given by the following expressions:

- Defining and calculating a distance function: Let L be a set of back transformed surface points P given by one acquisition step. For each of these points, we calculate the Euclidean distance d_{norm} to its axis of rotation. $R_{axis} = R + v \cdot S$. d_{norm} is given by

$$d_{norm} = \frac{\left| S \; x \; (P - R) \right|}{|S|} \quad \textbf{Equation 4}$$

- Defining and calculating the gradient of one surface point: Let $P_{i,\max}$ be the surface point with maximum Euclidian distance d_{norm} to R_{axis} in the ith acquisition step. The gradient g_i is calculated by the following algorithm (see also Figure 4):
  1. Estimation of $d_{i,\max}$: This is explicitly given by $P_{i,\max}$.
  2. Estimation of the surface point P_{i-1} (acquired in the previous step $i-1$) with the same z-component as $P_{i,\max}$ and calculation of $d_{i-1}(P_{i-1})$.
  3. Computation of the approximated gradient g_a between d_{i-1} and $d_{i,\max}$.
  4. Computation of the gradients angle a_a of g_a.
  5. Estimation of the surface point P_i with the same z-component as the point $P_{i-1,\max}$ and calculation of $d_i(P_i)$.
  6. Computation of the approximated gradient g_b between $d_{i-1,\max}$ and d_i.
  7. Computation of the gradients angle a_b of g_b.
  8. Estimation of the region with the highest entropy. This region is denoted by $\max(a_a, a_b)$.

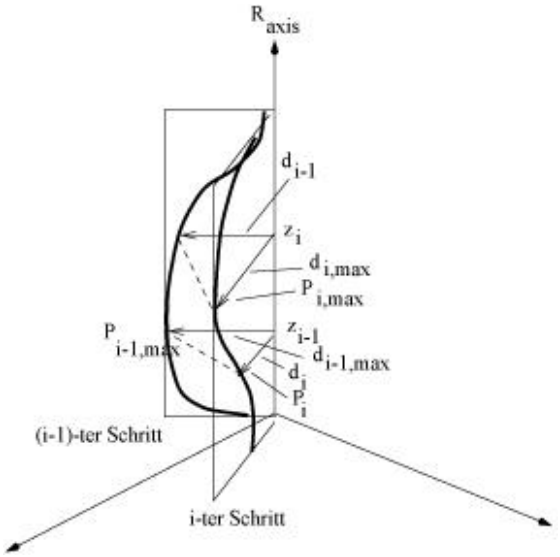


Figure 4: Calculating the next viewing angle

- Calculation of the next rotation angle: The gradient value could be positive, negative or zero, depending on increasing, decreasing or unchanging entropy. Table 1 shows the calculation of the relative change $f_{i,rel}$ and the absolute angle $f_{i,abs}$ depending on the sign of the gradient g_i, the gradient angle a_i and the relative change $f_{i-1,rel}$ of the (I-1)-t^{h} step, where the threshold t_a encodes the geometric under which the system sampling density will be increased.

g_i	a_i	$f_{i-1,rel}$	$f_{i,rel}$	$f_{i,abs}$
> 0	$> t_a$	$> f_{\min} \cdot 2$	$\frac{f(P_i) - f(P_{i-1})}{2}$	$f_{i-1} - f_{i.rel}$
		$< f_{\min} \cdot 2$	$f_{\min}$	$f_{i-1} - f_{i.rel}$
		$= f_{\min} \cdot 2$	$f_{\min}$	$f_{i-1} + f_{i.rel}$
> 0	$\leq t_a$	-	$f_{i-1,rel}$	$f_{i-1} + f_{i.rel}$
< 0	$> t_a$	$> f_{\min} \cdot 2$	$\frac{f(P_i) - f(P_{i-1})}{2}$	$f_{i-1} - f_{i.rel}$
		$< f_{\min} \cdot 2$	$f_{\min}$	$f_{i-1} - f_{i.rel}$
		$= f_{\min} \cdot 2$	$f_{\min}$	$f_{i-1} + f_{i.rel}$

< 0	$\leq t_a$	$< f_{max}/2$	$f_{i-1,rel} \cdot 2$	$f_{i-1} + f_{i.rel}$
		$\geq f_{max}/2$	f_{max}	$f_{i-1} + f_{i.rel}$
$= 0$	$= 0$	$< f_{max}/2$	$f_{i-1,rel} \cdot 2$	$f_{i-1} + f_{i.rel}$
		$\geq f_{max}/2$	f_{max}	$f_{i-1} + f_{i.rel}$

Table 1: Next rotation angle

3d acquisition

We combine two different techniques for 3d acquisition: First of all, a shape from structured light system is used to estimate the next best view and to determine highly structured parts of the object. The second system is a shape from silhouette system, which is used to estimate the surface and build a volume model of the object of interest.
We define an iterative process, which combines the two approaches, by the following steps (see also Figure 5, which depicts this process):

1. sfsl-Image acquisition: The scene is captured by the sfsl-CCD-camera. The result is a grayscale image, which shows the intersection between the laser plane and the object, which forms a line.
2. sfsl-Feature extraction: The line shown in the camera image is extracted. The result is a set of 2d points.
3. sfsl-Registration: The set of 2d points extracted in the previous step is transformed from the world coordinate system in its object coordinates.
4. sfsl-Integration: Each registered point is integrated into the existing model computed and integrated at the previous iterations of the acquisition process.
5. Next View Planning: The next viewing angle is computed based on the algorithm shown in the previous section and the turntable moves to the calculated absolute angle.

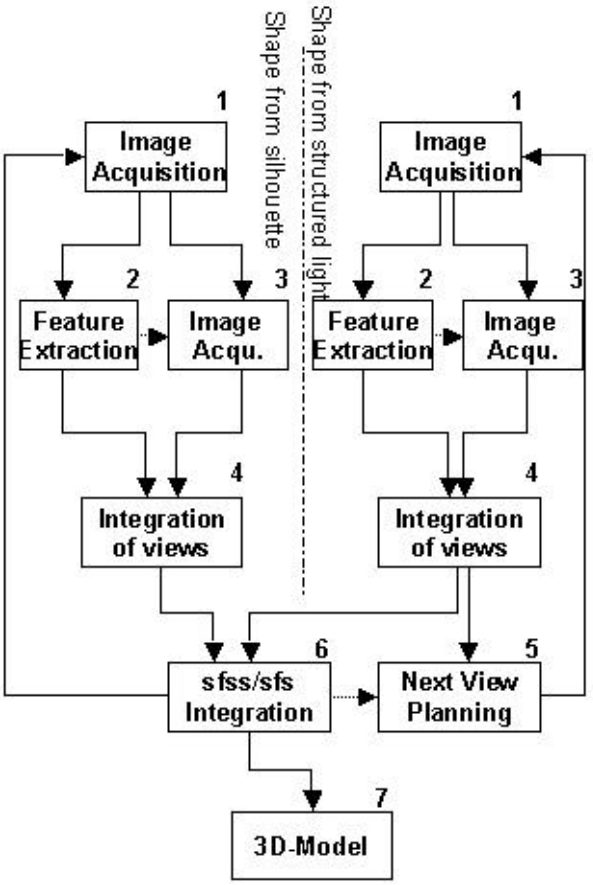


Figure 5: Iterative procedure for a combined sfs/sfsl-approach

The shape from silhouette process, which runs simultaneously with the sfsl-process is defined by the following steps:

1. sfs-Image acquisition: The scene is captured by the sfs-CCD-camera. The result is a grayscale image, which shows the objects silhouette of the current acquisition step.
2. sfs-Feature extraction: The silhouette shown in the camera image is extracted. The result is a set of 2d points.
3. sfs-Registration: An octree node is projected into the image plane by projecting all of its eight vertices. Octree coordinates of a vertex can be determined by the path from the root node of the octree to the node the vertex belongs to. We need to transform the octree coordinates of a vertex into image coordinates. This transformation is performed in three steps, each of which, if we use homogeneous coordinates [9] to represent points in all three coordinate systems, can be described by a 4×4 *or* 4×3 transformation matrix.
4. sfs-Integration: The result of the projection of an octree node into the image plane are image coordinates of all of the vertices of the node's

corresponding cube. In the general case, the projection of a node looks like a hexagon.

5. sfsl/sfs-Integration: In this step, the surface points and the volume model will be integrated. Therefore the intersection of the two models is generated.
6. 3d-model visualization.

The process repeats until the turntable revolves one complete rotation.

RESULTS

Figure 6 shows the reconstruction of an archaeological amphora using the sfsl-approach. The symmetry axis of the pottery and the rotational axis of the turntable are roughly justified. The minimum angle was defined as f_{min} = 4deg and the maximum angle as f_{max} = 12deg. Analyzing the reconstructed data shows a displacement of 1.8 mm in x-direction and 2.1 mm in y-direction. Therefore, the object was sampled with varying relative angles in 36 steps.

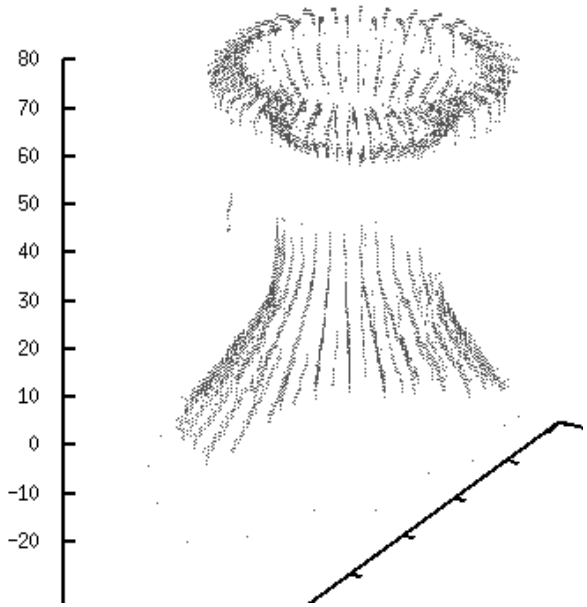


Figure 6a: Recorded pottery reconstructed

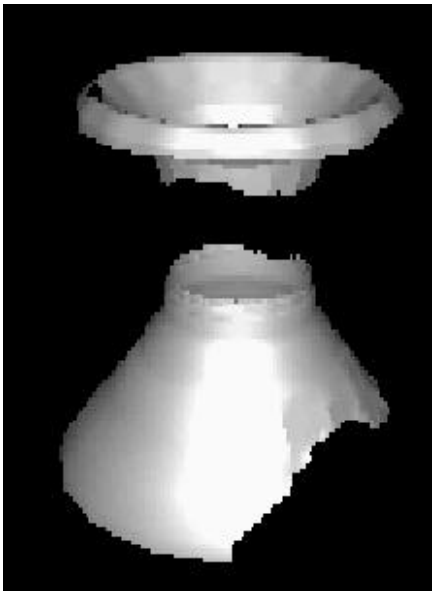
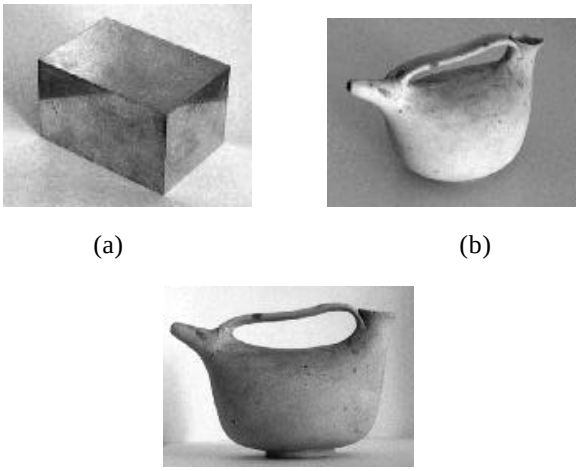


Figure 6b: Recorded pottery rendered

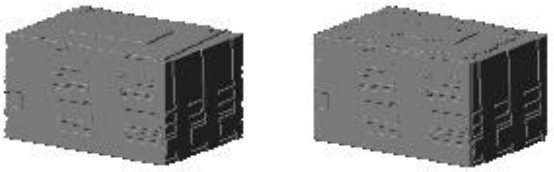
We built three octree models for three different objects (see Figure 7) by using 4, 12 and 36 input views. In a first step we used a uniform angle between two consequent views, with octree resolution of 2563 voxels. Figure 8 shows the 3D models constructed and Table 2 gives the octree statistics and the size comparison between the real objects and their constructed models. In a second step we used our NBV-strategy in order to get a minimal number of steps to reconstruct the object of interests. First results have shown that the number of views can be reduced by 32%, which reduced the computation time by 50%.



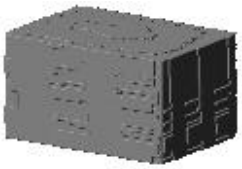
(a) (b)

(c)

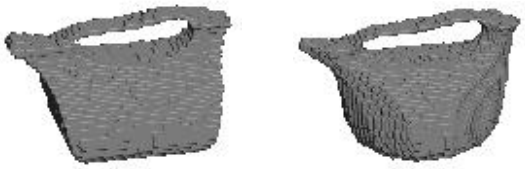
Figure 7: Real objects: a metal cuboid (a) and two pots (b) and (c)



(a): 4 views (b): 12 views

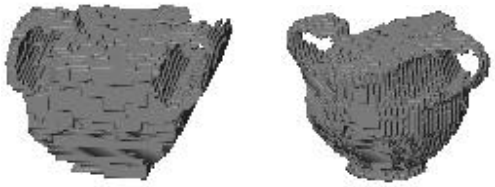


(c): 36 views



(a): 4 views (b): 12 views

(c): 36 views



(a): 4 views (b): 12 views

(c): 36 views

Figure 8: Constructed models of real objects with different number of views (a: 4, b:12 and c: 36 views)

object	Dimensions [mm]	Voxel size [mm]	# nodes	CPU time [sec]
Cuboid real	100.2 x 60.1 x 70.3	-	-	-
Cuboid 4 views	109.0 x 62.1 x 77.3	0.59	84065	9.76
Cuboid 4 views	109.0 x 62.1 x 77.3	0.59	161625	15.15
Cuboid 4 views	109.0 x 62.1 x 77.3	0.59	310649	36.10
Pot1 real	141.2 x 93.7 x 84.8	-	-	-
Pot 1 4 views	148.4 x 93.8 x 89.1	0.78	275177	12,30
Pot 1 12 views	145.3 x 93.8 x 89.1	0.78	486585	23.47
Pot 1 36 views	145.3 x 93.8 x 85.9	0.78	857753	67.28
Pot 2 real	114.2 x 87.4 x 114.6	-	-	-
Pot 2 4 views	118. x 89.1 x 115.6	0.78	257897	12.57
Pot 2 12 views	117.2 x 89.1 x 114.1	0.78	424201	23.32
Pot 2 36 views	117.2 x 89.1 x 114.1	0.78	726689	65.10

Table 2: Octree statistics for real data with varying number of views

The results with both synthetic and real input data show that there is a certain minimal octree resolution required to obtain an accurate model of an object, especially for highly detailed objects, like the two pots used for tests with real images. To increase the octree resolution would not improve the results of our tests, because the projection of a single voxels would be less than half the pixels size. Concerning the number of input views used for obtaining a model of an object, it turned out that beginning from 12 views, the constructed model does not change a lot - in our tests the octrees built from 12 views were almost the same as the ones built from 36 views, except that they took much less time to construct.

The results with synthetic data, where we had a perfect transformation matrix, showed that the error in the dimensions of the model lies within or is slightly higher than the error introduced through the minimal voxel size. The error with real data depends additionally on the accuracy of the calibration algorithm.

CONCLUSION

We have presented a next-view-planning technique combined with shape-from-silhouette in order to reduce the computational effort in 3d surface reconstruction. The work was performed in the framework of the documentation of a ceramic fragments. First results have shown that the number of views can be reduced by 32%, which reduced the computation time by 50% without decreasing the quality of reconstruction. The surface structure is preserved since high structured parts of the surface are sampled with higher density than unstructured parts.

It is part of continuing research efforts to improve the results from multiple, various objects since the technique has some drawbacks. The first one refers more to the calibration algorithm, which makes many simplifying

assumptions about the acquisition system, the one about the optical axis of the camera lying exactly in the turntable. However, it showed to be a very good approximation, which greatly simplifies the calibration algorithm.
For archaeological applications, the object surface has to be smoothed in order to be applicable to ceramic documentation, for classification, however, the accuracy of the method presented is sufficient since the projection of the decoration can be calculated and the volume estimation is much more precise than the estimated volume performed by archaeologists.

REFERENCES

1. R. M. Haralick and L.G. Shapiro. Glossary of computer vision terms. *Pattern Recognition*, 24(1), 1991, 69-93

2. M. Kampel and S. Tosovic, Turntable calibration for automatic 3D-reconstruction, in *Applications of 3D-Imaging and Graph-based Modelling, Proceedings of the 24th Workshop of the Austrian Association for Pattern Recognition (ÖAGM)*, 2000, 25-31

3. C. Liska, Das adaptive Lichtschnittverfahren zur Oberflächenrekonstruktion mittels Laserlicht, *Tech. Rep.* PRIP-TR-055, PRIP, TU Wien, 1999.

4. C. Liska, Sablatnig R., "Estimating the Next Sensor Position based on Surface Characteristics", in: "Proc. of 15th International Conference on Pattern Recognition, Barcelona", Vol.1, 2000, 538-541

5. J. Maver, Necessary Views for a Coarse Representation of a Scene, in *Proceedings of the 13th ICPR* - Track A, 1996, 936-940

6. J. Maver and R. Bajcsy, Occlusions as a Guide for Planning the Next View, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15, May 1993, 417-432

7. M. Milroy, C. Bradley, and G. Vickers, Automated laser scanning based on orthogonal cross sections, *Machine Vision and Applications* 9, 1996, 106-119

8. C. Menard and R. Sablatnig, Computer based Acquisition of Archaeological Finds: The First Step towards Automatic Classification, in *Interfacing the Past, Computer Applications and Quantitative Methods in Archaeology*, H. Kamermans and K. Fennema, eds. no. 28, March 1996, Analecta Praehistorica Leidensia, (Leiden), 413-424

9. V. S. Nalwa, *A Guided Tour Of Computer Vision*, Addison-Wesley, 1993.

10. W. Niem, Error analysis for silhouette-based 3d shape estimation from multiple views, in *Proc. of Intl. Workshop on Synthetic-Natural Hybrid Coding and Three-Dimensional* Imaging, N. Sarris and M. Strintzis, eds., 1997, pp. 143-146

11. D. Papadopoulos-Orfanos and F. Schmitt, Automatic 3-D Digitization Using a Laser Rangefinder with a Small Field of View, in *Proceedings of the International Conference on Recent Advances in 3-D Digital Imaging and Modelling*, pp. 60-67, (Ottawa, Canada), May 1997.

12. C. Orton, P. Tyers, A. Vince. *Pottery in Archaeology*, 1993.

13. M. Potsemil. Generating octree models of 3D objects from their silhouettes in a sequence of images. In *Computer Vision, Graphics, and Image Proecessing*, 1990, 40:68-84

14. C.M, Sinopoli. Approaches to Archeological Ceramics. New York, 1991

15. V. Sequeira, J. G. M. Goncalves, and M. I. Ribeiro, Active View Selection for Efficient 3D Scene Reconstruction, in *Proceedings of the 13th International Conference on Pattern Recognition* - Track A, 1996, 815-819

16. R. Szeliski. Rapid octee construction from image sequences. *CVGIP: Image Understanding*, 58(1): July 1993, 23-32

17. G. H. Tarbox, Planning for Complete Sensor Coverage in Inspection, *Computer Vision and Image Understanding* 61, January 1995, 84-111

18. S. Tosovic, Lineare Hough-Transformation und Drehtellerkalibrierung, *Tech. Rep. PRIP-TR-59*, Institute of Computer Aided Automation, Pattern Recognition and Image Processing Group, Vienna University of Technology, Austria, 1999.

19. H. Zha, K. Morooka, T. Hasegawa, and T. Nagata, Active Modeling of 3-D Objects: Planning on the Next Best Pose (NBP) for Acquiring Range Images, in *Proceedings of the International Conference on Recent Advances in 3-D Digital Imaging and Modeling*, May 1997, 68-75

20. S. Vedula, S. Baker, S. Seitz, and T. Kanade. Shape and motion carving in 6D. In CVPR00, volume 2, 2000, 592-598.

ABOUT THE AUTHORS

Martin Kampel was born in Steyr, Austria, in 1968. He studied data technologies (1990-1993) and computer science (1992-1999) at the Vienna University of Technology. Since 1996 he is working at the Pattern Recognition and Image Processing Group, Institute of Computer Aided Automation in Vienna, engaged in research, teaching and administration. His research

interests are 3D-Vision, Computer graphics and Virtual Archeology. Currently he is working on his PhD thesis on 3D mosaicing.
kampel@prip.tuwien.ac.at

Robert Sablatnig was born in Klagenfurt, Carinthia, Austria, in 1965. He received the M.Sc. degree (Diplom Ingenieur) in Computer Science (Computer Graphics, Pattern Recognition & Image Processing) in 1992 and the Ph.D. degree in Computer Science in 1997 from the Vienna University of Technology.
He is an assistant professor (Univ.Ass.) of computer vision at the Pattern Recognition and Image Processing Group, Institute of Computer Aided Automation, engaged in research, project leading, and teaching. His research interests are Applications in Industrial Inspection, Automatic Visual Inspection, Machine Vision and 3D Computer Vision.
sab@prip.tuwien.ac.at