Robust and Fast Modelling of 3D Natural Objects from Multiple Views

Wolfgang Niem
Universität Hannover
Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung
Appelstr. 9A, D–30167 Hannover, Germany
email: niem@tnt.uni–hannover.de

ABSTRACT

An algorithm for the robust and fast automatic construction of a 3D model of any real object using images from multiple views is presented. The images are taken from a real object rotating in front of a stationary calibrated CCD TV camera. The object silhouettes extracted from the input images, the related turntable positions and camera orientation are used to construct the volume model of the real object by applying the method of occluding contours. A keypoint in performing this method is a proper volume representation, characterized by low complexity and suitability for a fast computation of volume models. In the presented approach, each volume model is described by pillar–like volume elements (pillars) ensuring a computational complexity proportional to the size of the real object surface and enabling a fast and simple construction of the volume model. The fast performance is due to the simple projection feasibility of those pillars and the easy–to–perform intersection test for the object silhouette with the projected pillars. Results with real image sequences have confirmed the robustness of the developed algorithm even for the modelling of real objects with highly detailed and complex surfaces and the use of imperfect object silhouettes.

1 INTRODUCTION

This paper presents a new algorithm for the robust and fast automatic construction of a 3D model of real objects using silhouettes from multiple views. Fields for applications are computer aided design (CAD) where automatically generated model objects can be used as a first draft input for the designer; robotics, where the scene has to be modelled before it can be manipulated; and virtual reality where the generation of the virtual scene can be facilitated by providing a large library of model objects. For all these applications it is important that the acquisition of the model objects can be done with simple equipment and in reasonable time.

For the automatic construction of 3D models, principally two kinds of methods can be distinguished. Active methods use structured light or a laser scanner. Passive methods use image sequences of a camera. The passive methods require less equipment and are therefore easier to handle. Besides, they have no principal restriction for using them in outdoor scenes as in case of active methods which are restricted with respect to the maximum object size. On the other hand, the passive methods require more sophisticated algorithms. The approach for modelling real objects from multiple views presented in this paper makes a contribution to the passive methods.

Basic work for the construction of 3D models from multiple views was done by Martin and Aggarwal [8]. They introduced the method of ”occluding contours” which uses the real object silhouettes extracted from the real images and orthographic projection for the construction of a volume model. Busch [1] used a polygon approximation of the object silhouette and constructed a volume model from arbitrary views using a voxel representation. Jackins and Tanimoto [5], and Meagher [9] introduced an efficient volume representation, the octree representation, which describes the object as a tree of recursively subdivided cubes. Chien [3] constructed an octree from three orthographic projections. Potmesil [11] and Szeliski [13] used arbitrary views and perspective projection to construct an octree. Both of these approaches project the octree cubes into the image plane to perform the intersection test between the projected cube and the object silhouette.

Two principal approaches for the intersection test of the model volume with the object silhouettes can be distinguished : (a) volume elements can be projected into the image plane and tested for intersection with the object silhouette or (b) object silhouettes can be backprojected onto the volume and tested for intersection. The former approach requires intersection testing in two dimensions whereas the latter one requires intersection testing in three dimensions. Therefore, the more viable first approach is chosen in the presented work.

One critical point in former work is the intersection test of the projected octree volume cubes with the object silhouette; a cube projected into the image plane can form a hexagon which can lead to a time consuming intersection test. The approximation of this hexagon by a bounding box as done by several authors can lead to reconstruction errors. Moreover, some authors [1],[12],[10] approximate the object silhouette with a polygon, which can lead to reconstruction errors as well, especially in case of complex

silhouettes with holes. The speed of computation also depends on the way the central projection is performed; for that reason the choice of a suitable camera model is important.

In this approach, the volume model is described by volume pillars, ensuring a complexity proportional to the size of the real object surface and enabling a fast and simple intersection test of the projected volume elements with the object silhouette. The camera model used is the one introduced by Yakimovski and Cunningham [15] which is suited for a fast implementation of the central projection.

In Section 2 the camera model used for the modelling is defined. Section 3 describes the measurement environment and the system calibration. In Section 4 the principles are described how to generate volumes from silhouettes. Section 5 introduces the proposed algorithm for the integration of the object silhouettes using a new kind of volume elements. Section 6 presents and discusses results obtained with the developed new algorithm.

2 THE CAMERA MODEL

For the processing of the input images, the real camera must be represented by a mathematical model which describes the physical behaviour of the real camera sufficiently exact. A camera model, which is suited for 3D–modelling applications, is the one introduced by Yakimovski and Cunningham[15]. It assumes the camera to be geometrically linear, an assumption which is reasonable, considering the linear sensor array and the high quality of the used lens. Thus, it is possible to use the laws of central projection for the object acquisition. Additionally, a linear distortion and a shift of the optical center with respect to the image center can be taken into account.

For the derivation of the main projection equations, a world coordinate system (WCS) related to the real objects as shown in Fig. 1 a, and an image coordinate system (ICS) related to the image plane as shown in Fig. 1 b are defined. The following vectors are used:

$\vec{c} = (c_x, c_y, c_z)^T$: vector to the focal center in WCS
$\vec{a} = (a_x, a_y, a_z)^T$: unit vector in the direction of the optical axis in WCS
$\vec{h'} = (h_x', h_y', h_z')^T$: unit vector parallel to the horizontal axis of the image plane in WCS
$\vec{v'} = (v_x', v_y', v_z')^T$: unit vector parallel to the vertical axis of the image plane in WCS
$\vec{P} = (P_x, P_y, P_z)^T$: vector to a real–world point in WCS
$\vec{T} = (I,J)^T$: vector to an image point in ICS
$\vec{T}_h = (X_h, Y_h)^T$: vector to the optical center in ICS

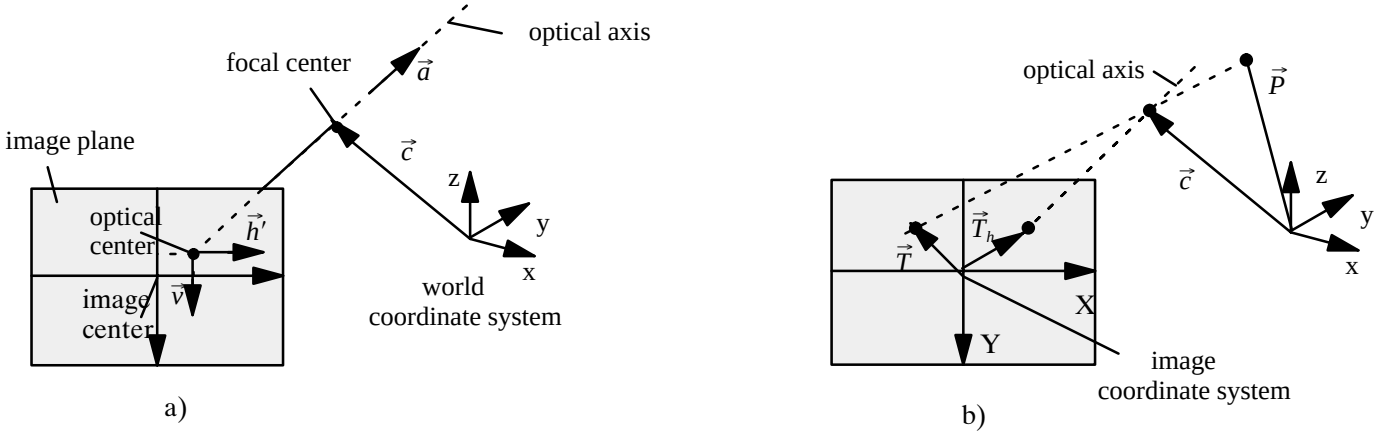

Fig. 1 : (a) Camera model
(b) Projection of a real–world point $\vec{P}$ into an image point $\vec{T}$

The unit vectors $\vec{a}$, $\vec{h'}$,$\vec{v'}$ are perpendicular to each other and form a right–handed system. Its origin is located in the optical center of the image which must not inevitably be the same as the origin of the ICS. This is taken into into account by the introduction of the vector $\vec{T}_h = (X_h, Y_h)^T$ describing the difference vector between image center and optical center. Using this system and the focal

length f of the camera, the projection of a real–world point $\vec{P}$ into an image point $\vec{T} = (I,J)^T$ as shown in Fig. 1 can be performed by

$$I = \frac{(\vec{P} - \vec{c}) \cdot f \cdot \vec{h'}}{(\vec{P} - \vec{c}) \cdot \vec{a}} + X_h \tag{1}$$

$$J = \frac{(\vec{P} - \vec{c}) \cdot f \cdot \vec{v'}}{(\vec{P} - \vec{c}) \cdot \vec{a}} + Y_h \quad , \tag{2}$$

The focal length f is scaled (using the physical size of a CCD element) in order to calculate with an image pixel size of 1mm x 1mm. Using the properties of the scalar product, the equations (1) and (2) can be transformed to

$$I = \frac{(\vec{P} - \vec{c}) \cdot (f \cdot \vec{h'} + \vec{a} \cdot X_h)}{(\vec{P} - \vec{c}) \cdot \vec{a}} \tag{3}$$

$$J = \frac{(\vec{P} - \vec{c}) \cdot (f \cdot \vec{v'} + \vec{a} \cdot Y_h)}{(\vec{P} - \vec{c}) \cdot \vec{a}} \tag{4}$$

If we define

$$\vec{h} = f \cdot \vec{h'} + \vec{a} \cdot X_h \tag{5}$$

$$\vec{v} = f \cdot \vec{v'} + \vec{a} \cdot Y_h \quad , \tag{6}$$

the equations (3) and (4) can be simplified to

$$I = \frac{(\vec{P} - \vec{c}) \cdot \vec{h}}{(\vec{P} - \vec{c}) \cdot \vec{a}} \tag{7}$$

$$J = \frac{(\vec{P} - \vec{c}) \cdot \vec{v}}{(\vec{P} - \vec{c}) \cdot \vec{a}} \tag{8}$$

The main advantage of the camera description with the $\vec{c}, \vec{a}, \vec{h}, \vec{v}$ system in contrary to other alternatives like the system used by Tsai [14] is the easy–to–perform projection of a real world point into the image plane of a CCD camera. All geometric manipulations described in the next sections are performed using this $\vec{c}, \vec{a}, \vec{h}, \vec{v}$ system and will be referred to as cahv–camera model.

3 MEASUREMENT ENVIRONMENT AND SYSTEM CALIBRATION

The environment used for this work consists of a stationary CCD TV camera in front of a turntable, which can be rotated in controlled steps. The background is of one colour in order to facilitate the segmentation of the real object against the background. Before the acquisition of a real object on the turntable can be started, the geometric relation between the CCD camera and the turntable has to be computed, i.e. the system has to be calibrated. The calibration of this image acquisition system supplies the extrinsic parameters of the camera, which are the position and orientation of the camera with respect to the real world, and the intrinsic parameters, which are the focal length and the shift of the optical center. For that purpose, a precisely known test pattern consisting of black circles on a white background is placed on the turntable as shown in Fig. 2. The size and position of the test pattern are measured. The detection of the corresponding ellipses in the image taken by the CCD camera and the subsequent estimation of all parameters describing the acquisition process is performed using a calibration method like the one proposed by Tsai [14].

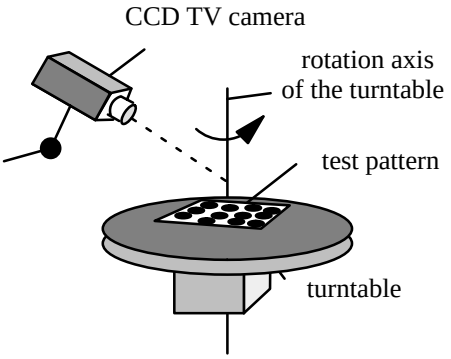

Fig. 2 : Arrangement for system calibration

Calculating those parameters for two known positions of the turntable allows the computation of the rotation axis as well. The world coordinate system is defined that one coordinate axis is parallel to the rotation axis and the origin is the intersection of the rotation axis with the plane of the turntable. Thus the cahv–parameters for any rotation angle of the turntable can be computed.

4 VOLUMES FROM SILHOUETTES

An approach for the construction of volume models using multiple views was first described by Martin and Aggarwal [8]. They introduced the method of ”occluding contours” which uses the object silhouettes and the related camera parameters to construct a volume model of the real object. The principle of this method can be divided into three steps.

In a first step, the silhouette of the real object must be extracted from the input images as shown in Fig.3. In controlled environments the segmentation of the real object against the background can be facilitated by using a background of one colour. In this case, a color segmentation technique like the one used in TV studios (”blue screen technique”) can be used to get the object silhouette.

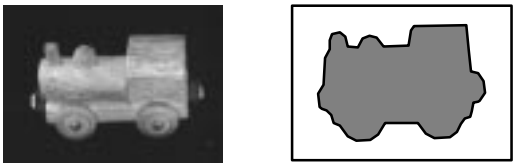

Fig. 3 : Segmentation of the input image

In a second step, a bounding pyramid is constructed using the focal point of the camera and the silhouette as shown in Fig. 4 a. The convex hull of this pyramid is formed by the rays of sight from the camera focal point through different contour points of the object silhouette. For each view point such a pyramid is constructed, and each pyramid can be seen as a first approximation of the volume model.

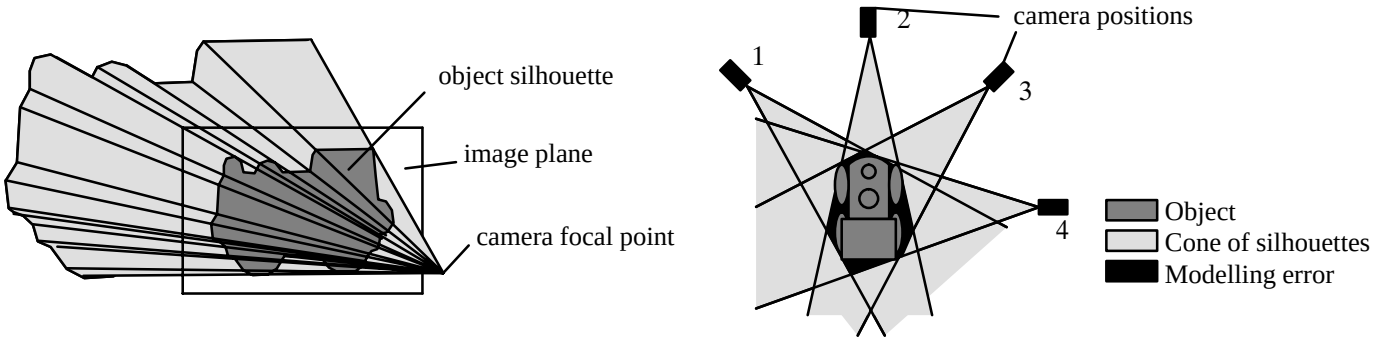

Fig. 4 : (a) Construction of a bounding pyramid, (b) Two dimensional description of the intersection of the bounding pyramids

In a last step, all the pyramids from the different view points are intersected in the 3D–world and form the final approximation of the volume model. This is performed with the knowledge of the camera positions which give the information of the geometrical relation of the pyramids. In Fig. 4b a two dimensional description of the intersection of the bounding pyramids is shown exemplary.

The limitations of such a silhouette–based approach are shown in Fig. 4b: the principal inability of the algorithm to detect object concavities (like the hollow in a cup) and the dependency of the approximation accuracy from the number and the choice of the viewpoints. The resulting volume model obtained by this method is only exact in case of infinite number of views and convex objects. In order to minimize the modelling error using only a finite number of views, the views must be chosen carefully.

5 INTEGRATION OF SILHOUETTE INFORMATION USING PILLAR–LIKE VOLUME ELEMENTS

A keypoint in performing the method of ”occluding contours” is a proper volume representation, characterized by low complexity and suitability for a fast computation of volume models. One often used representation are octrees [2],[3],[5],[6],[9],[10],[11],[13], which describe the object as a tree of recursively subdivided cubes down to the finest resolution. The volume model is constructed projecting the octree cubes consecutively into the image planes of every viewpoint and testing them for intersection with the object silhouettes.

One crucial point of this method is the intersection test with the object silhouette; a cube projected into the image plane can form a hexagon which can lead to a time consuming intersection test. The approximation of this hexagon by a bounding box as done by several authors [11],[13] can lead to reconstruction errors.

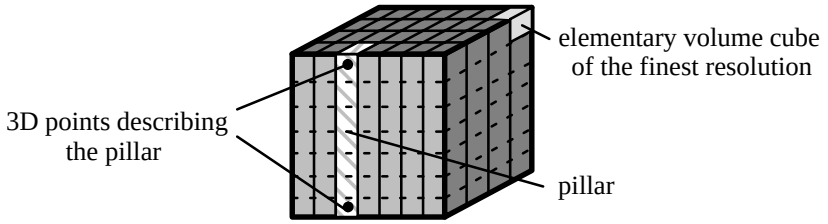

Fig. 5 Volume representation using pillar–like volume elements (pillars). Each pillar is described by the 3D positions of the center of the elementary volume cubes on the top and the bottom of the pillar.

In this paper, a new volume representation is introduced. The volume is decomposed into pillar–like volumes (pillars) which are built of elementary volume cubes of the finest resolution. Each of those pillars is completely described by the position of the center points of the cubes on the top and the bottom of the pillar as shown in Fig. 5. The complexity of this representation is proportional to the object surface area (measured in units of the finest resolution).

The presented algorithm starts with the definition of an initial volume that surely contains the object. Actually, a volume cube composed of a set of pillars as shown in Fig. 5 is used for that purpose. Each of the pillars is projected consecutively into the image planes from the different views containing the object silhouettes and tested for intersection with the object silhouettes. The algorithm works as follows:

Projection of the 3D points describing the pillar into the image plane with the silhouette (Fig. 6 a)
Using the definitions of section 2, the equations (7) and (8) and the definition of

$\vec{P}_1 = (P_{1x}, P_{1y}, P_{1z})^T$, $\vec{P}_2 = (P_{2x}, P_{2y}, P_{2z})^T$: vectors to the two 3D points describing the pillar in WCS
$\vec{T}_1 = (I_1, J_1)^T$, $\vec{T}_2 = (I_2, J_2)^T$: vectors to the projections of $\vec{P}_1$, $\vec{P}_2$ in the image plane in ICS,

the projection of the 3D points describing the pillar into the image plane is obtained as

$$I_i = \frac{(\vec{P}_i - \vec{c}) \cdot \vec{h}}{(\vec{P}_i - \vec{c}) \cdot \vec{a}} \quad , \text{i} = 1, 2 \tag{9}$$

$$J_i = \frac{(\vec{P}_i - \vec{c}) \cdot \vec{v}}{(\vec{P}_i - \vec{c}) \cdot \vec{a}} \quad , \text{i} = 1, 2 \tag{10}$$

Connecting the resulting image points with a 2D line (Fig. 6 b)
A scan–conversion algorithm for lines [4] computes the coordinates of the pixels that lie on or near an ideal, infinitely thin straight line through $\vec{T}_1$, $\vec{T}_2$ imposed on a 2D raster grid. This line represents the projection of a volume pillar in the image plane.

Subdividing and reducing the 2D line into new line segments by eliminating the pixels outside the silhouette (Fig. 6 c)
The pixels describing the 2D line are tested for intersection with the object silhouette; pixels outside the silhouette are eliminated, the pixels left can form new and reduced line segments as shown in Fig 6 c. Each of those new line segments can be again described by two points in the image plane.

Subdividing and reducing the pillar corresponding to the pixels describing the new 2D line segments (Fig. 6 d)
The obtained 2D line segments correspond to the parts of the volume pillar which are inside the volume model of the real object. For that reason, the volume pillar is subdivided in new and reduced pillar segments. For that purpose, the pillar is represented by a 3D line $\vec{L}_p$which is described by the points $\vec{P}_1$ on the top and $\vec{P}_2$ on the bottom of the pillar as

$$\vec{L}_p = \vec{P}_1 + \lambda_1(\vec{P}_2 - \vec{P}_1) \qquad , 0 < \lambda_1 < 1 \qquad (11)$$

Using the $\vec{c},\vec{a},\vec{h},\vec{v}$ system, the line of sight through any image point $\vec{T}_x = (I_x J_x)^T$ is given as

$$\vec{L}_l = \vec{c} + \lambda_2(\vec{h} - I_x \cdot \vec{a}) \times (\vec{v} - J_x \cdot \vec{a}) \qquad , 0 < \lambda_2 < +\infty \quad . \qquad (12)$$

The lines of sight through the 2D points describing a line segment are intersected with the 3D line representing a pillar. The intersection points describe the reduced pillar segment. Because of numerical errors, this intersection may not exist. For that reason, the point on the line $\vec{L}_p$ is computed for which the distance to the line of sight is minimum. Introducing the vector $\vec{e}$ perpendicular to $\vec{L}_p$ and $\vec{L}_l$ a vector equation with the three unknowns $\lambda_1,\lambda_2,\lambda_3$ can be derived which gives a unique solution for the point on $\vec{L}_p$which has a minimal distance to $\vec{L}_l$.

$$\vec{L}_p = \vec{L}_l + \lambda_3\vec{e} \qquad , -\infty < \lambda_3 < +\infty \qquad (13)$$

Applying those steps consecutively to each pillar, the approximation of the volume model represented by a set of pillars will improve for each processed silhouette. The approximation accuracy of a convex real object is limited only by the resolution of the silhouettes, the accuracy of the camera parameters and the number and choice of viewpoints.

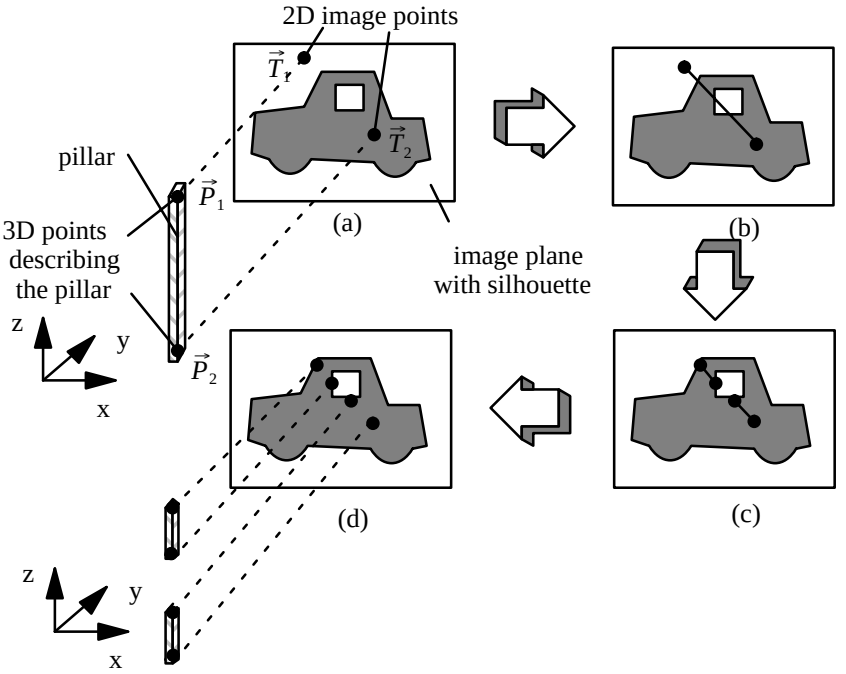

Fig. 6 Volume construction using pillars
(a) Projection of the 3D points $\vec{P}_1$, $\vec{P}_2$ describing the pillar into the image plane with the silhouette
(b) Connecting the resulting 2D image points with a 2D line
(c) Subdividing and reducing the 2D line into new line segments by eliminating the pixels outside the silhouette
(d) Subdividing and reducing the pillar corresponding to the pixels describing the new 2D line segments

6 RESULTS

The described algorithm was tested on a number of real image sequences. For that purpose, the real object was placed on a turntable and the turntable was rotated in 10^0 steps. For each of the 36 views, a 720 x 576 image was taken from the real object with a stationary calibrated CCD camera. The rotation angle accuracy of the turntable was about 0.1^0. Input sequences were taken not just from simply shaped objects but also from natural objects with complex surfaces. Fig. 7 shows exemplary original images from a dinosaur, a flower and a teapot. The object silhouettes obtained as results from the segmentation step using a color segmentation technique are shown in Fig. 8. A typical segmentation error is shown in Fig. 9c.

a) dinosaur b) flower c) teapot

Fig. 7 : Original images from real objects

a) dinosaur b) flower c) teapot

Fig. 8 : Object silhouettes

The subsequent processing of the volume was performed using a resolution in space of 160^3 unit cubes for the bounding box of each real object. A rendered side and a top view of the resulting volume models of a dinosaur, a teapot and a flower are shown in Fig. 9, 10, and 11. The results emphasize the suitability of the developed algorithm even for real objects with a complex and detailed surface like the dinosaur and the flower and its robustness against the use of imperfect silhouettes as shown in Fig. 9c.

The computational complexity of the algorithm has been compared with algorithms proposed by Szeliski [13] and Potmesil [11], which both use an octree construction algorithm. As the algorithms were not available on our workstations, the complexity analysis is based on published results [13]. Silhouettes of a synthetic generated sphere were computed from 32 known view points and the volume model was computed for each approach in a resolution of 64^3 unit cubes. The number of intersection tests of the projected volume primitives with the object silhouettes of each view were counted and summed up for each approach. The volume primitives are cubes of different size in case of octrees, pillars build of elementary cubes of the finest resolution in the presented approach. The results of this comparison are shown in Table 1. The proposed algorithm uses significant less tests than the former proposed.

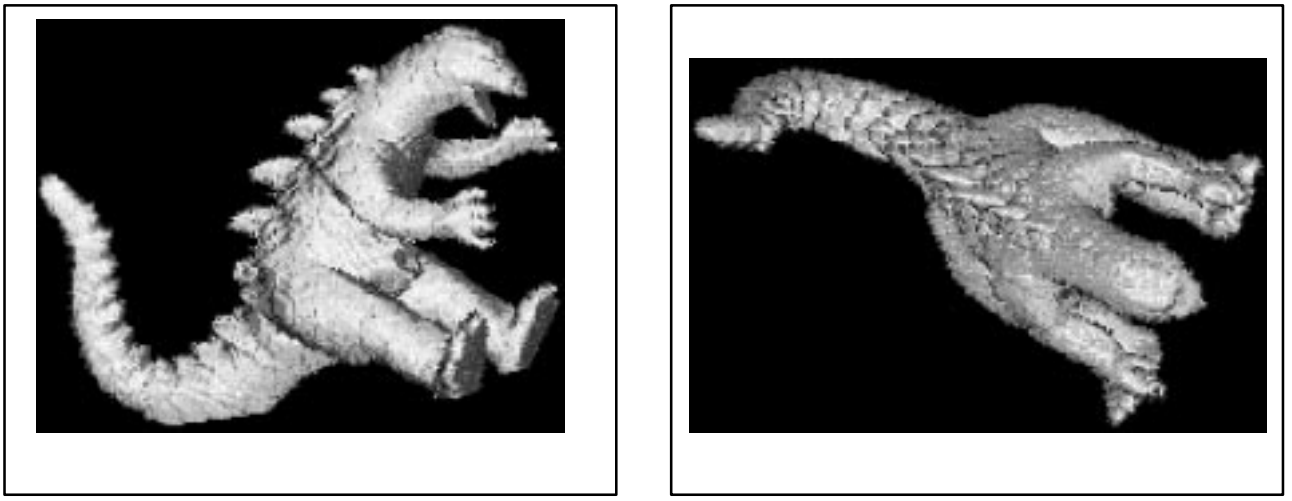

Fig. 9 Side and top view of the rendered volume model of a dinosaur

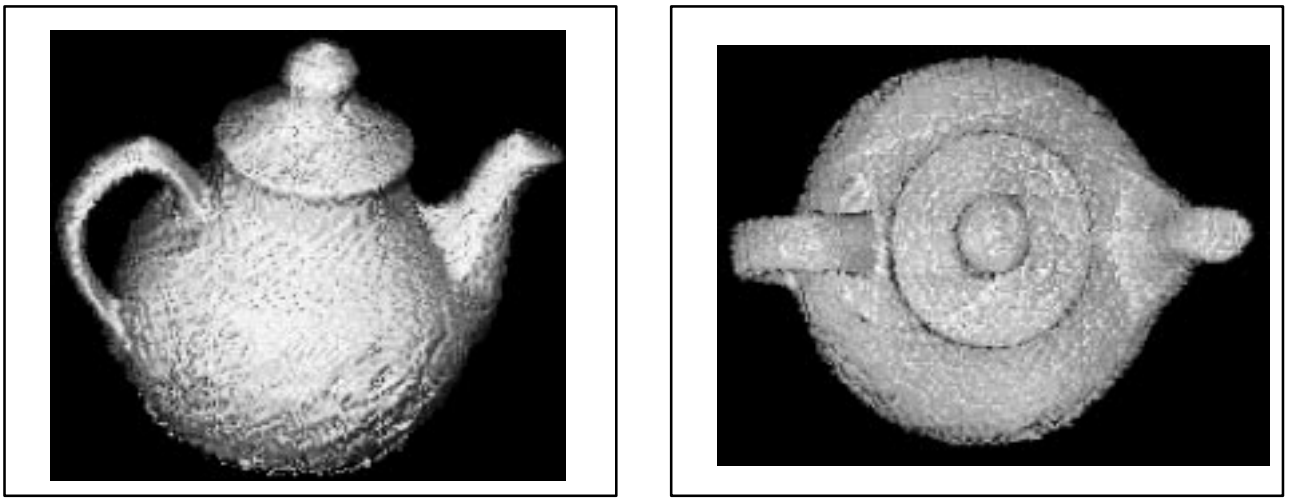

Fig. 10 Side and top view of the rendered volume model of a tea pot

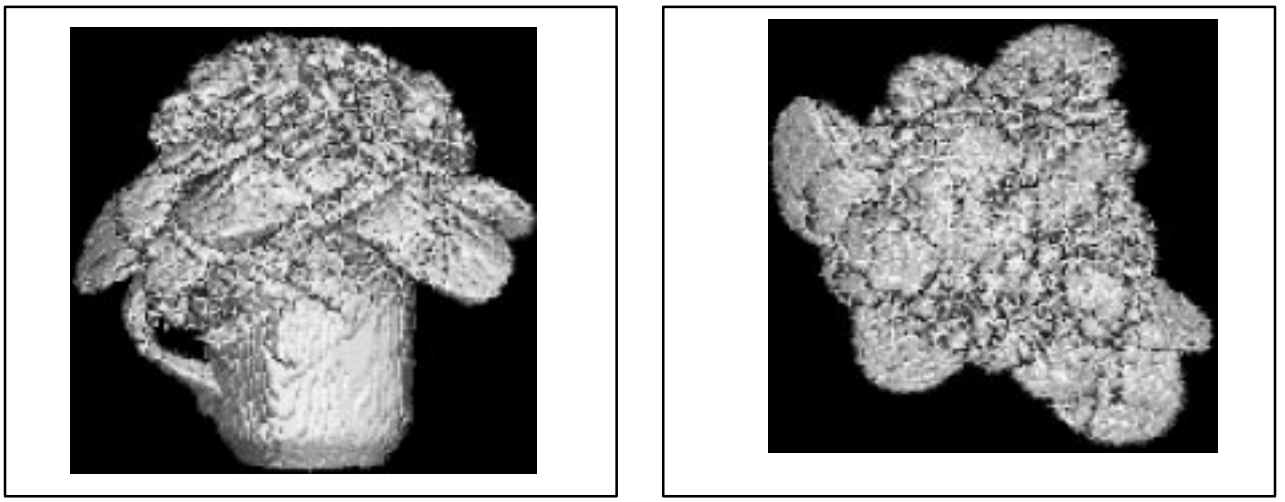

Fig. 11 Side and top view of the rendered volume model of a flower

The operations required for a single intersection test for a projected cube and a projected pillar with the object silhouette are compared in Table 2; the intersection test of a projected pillar with the object silhouette seems to be the simpler one. Comparing the operations in table 2 and assuming that the last two operations of each test will be of balanced complexity, the performance of the proposed test is increased roughly by a factor of four. For more detailed comparisons, benchmark tests have to be carried out in future work.

Benchmark tests on a SUN SPARC10/41 workstation for different real objects are shown in Table 3. Taking into account that one revolution of the turntable for the acquisition of 36 images takes about three minutes, the computation of the volume models can be performed in real–time.

ALGORITHM	NUMBER OF TESTED CUBES	NUMBER OF TESTED PILLARS
a	531054	—
b	429508	—
c	—	108192

Table 1 : Comparison of the number of intersection tests of projected volume primitives with the object silhouettes for the construction of a sphere in a resolution of 64^3 unit cubes.
a) Algorithm proposed by Potmesil (intersecting projected cubes with the object silhouette)
b) Algorithm proposed by Szeliski (intersecting projected cubes with the object silhouette)
c) Presented algorithm (intersecting projected pillars with the object silhouette)

OPERATIONS FOR THE INTERSECTION TEST OF A PROJECTED CUBE WITH THE OBJECT SILHOUETTE	OPERATIONS FOR THE INTERSECTION TEST OF A PROJECTED PILLAR WITH THE OBJECT SILHOUETTE
– 8 perspective projections of 3D points into the image plane – approximation of a hexagon in the image plane with a bounding box – intersection test of a 2D box with the object silhouette	– 2 perspective projections of 3D points into the image plane – scanning of a 2D line in the image plane for silhouette edges – 2 backprojections of edge points

Table 2: Required operations for a single intersection test for a projected cube and a projected pillar with the object silhouette

MODEL	DINOSAUR	TEAPOT	FLOWER
CPU TIME [S]	93	112	141

Table 3: Computation time for different volume models in a resolution of 160^3 unit volume cubes using 36 images with a resolution of 720x576 pixel on a SUN SPARC10/41 station

7 CONCLUSIONS

An algorithm for the robust and fast automatic construction of a 3D model object of the real object using images from multiple views is presented. It is based on the method of ”occluding contours” which uses real object silhouettes extracted from input images and the related camera parameters to construct a volume model of the real object. A keypoint in performing this method is a proper volume representation, distinguished by low complexity and suitability for a fast computation of volume models.

For that purpose, a new volume representation is proposed which describes the volume with pillar–like volume elements (pillars). Those pillars can easily be projected into the image plane and be tested for intersection with the real object silhouettes; the parts of the projected pillars outside the silhouette are eliminated, such that they are reduced or subdivided. The corresponding pillars of the volume model are reduced or subdivided according to their projections. The volume model is constructed by consecutively projecting each of the pillars into the image planes with the silhouettes of each view. The final set of pillars represents the volume model of the real object. Top and bottom of each pillar pillar are positioned on the object surface. The used camera model is the one introduced by Yakimovski and Cunningham, suited for a fast computation of the central projection.

The comparison of complexity with former proposed algorithms has confirmed, that the proposed algorithm requires significant less intersection tests of projected volume primitives with the object silhouettes to construct a volume model, and that the use of pillars

simplifies those intersection tests. Coarse tests have confirmed an increase of the performance by a factor of four. For more detailed comparisons, benchmark tests have to be carried out in future work.

Results with real image sequences have confirmed the robustness of the proposed algorithm even for the modelling of real objects with complex and detailed surfaces using imperfect silhouettes. The resulting models were constructed in reasonable time with simple equipment with an accuracy sufficient for computer graphic applications. The construction of a toy dinosaur in a resolution of 160^3 unit volume cubes from 36 input images which have a resolution of 720 x 576 takes 93 seconds on a SUN SPARC10 workstation.

However, the approach still has limitations like previous silhouette–based modelling approaches, such as the principal inability to detect concavities in the objects (like the hollow in a cup). For that reason, future work will concentrate on the integration of more sophisticated shape–from–motion algorithms.

8 ACKNOWLEDGEMENTS

This work was supported by the RACE II project R2052 ”MONA LISA”. Partners in this project are Thomson–CSF/LER, BBC, Daimler Benz, DVS Digitale Videosignale GmbH, Queen Mary and Westfield College, Siemens AG, University of Balearic Islands, and VAP Video Art Production GmbH.

9 REFERENCES

[1] H. Busch, ”Ein Verfahren zur Oberflächenmodellierung dreidimensionaler Objekte aus Kamerabildfolgen”, *Ph.D. thesis, Universität Hannover,* 1991.

[2] H. H. Chen and T. S. Huang, ”A survey of construction and manipulation of octrees”, *Comp. Vision Graphics Image Process.*, Vol.43, 1988, pp. 409 – 431.

[3] C. H. Chien and J. K. Aggarwal, ”Identification of 3D objects from multiple silhouettes using quadtrees/octrees ”, *Comp. Vision Graphics Image Processing* 36, 1986, pp. 256–273.

[4] J.D. Foley et al., ”Computer Graphics: Principles and Practice”, Addison Wesley, 1992.

[5] C. L. Jackins and S. L. Tanimoto, ”Oct–trees and their Use in Representing Three–dimensional Objects”, *Comp. Graphics Image Processing.* 14, 1980, pp. 249–270.

[6] Lavakusha, A. K. Pujari, and P. G. Reddy, ”Linear octrees by volume intersection”, *Comp. Vision Graphics Image Process.,* Vol 45, 1989, pp.371–379.

[7] H. C. Longuet–Higgins, ‘‘A computer algorithm for reconstructing a scene from projections”, *Nature,* Vol. 293, 1981, pp. 133–135.

[8] W. N. Martin and J. K. Aggarwal , ‘‘Volumetric Descriptions of Objects from Multiple Views” ,*Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI–5, 1983, pp. 150–159.

[9] D. Meagher, ”Geometric modelling using octree encoding”, *Computer Graphics Image Process.*, Vol 19, 1982, pp. 129–147.

[10] H. Noborio, S. Fukada, and S. Arimoto, ”Construction of the octree approximating three–dimensional objects by multiple views”, *Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI–10(6), 1988, pp. 769–782.

[11] M. Potmesil, ”Generating octree models of 3D objects from their silhouettes in a sequence of images”, *Comp. Vision Graphics Image Proc.* 40, 1987, pp. 1–29.

[12] S. K. Srivastava, and N. Ahuja, ”Octree generation from object silhouettes in perspective views”, *Comp. Vision Graphics Image Proc.* 49, 1990, pp. 68–84.

[13] R. Szeliski, ”Rapid Octree Construction from Image Sequences”, *CVGIP: Image Understanding*, Vol. 58, No 1, July, pp. 23–32, 1993.

[14] R. Y. Tsai, ”A Versatile Camera Calibration Technique for High–Accuracy 3D Machine Vision Metrology Using Off–the–Shelf TV Cameras and Lenses”, *Journal of Robotics and Automation*, Vol. RA–3. No.4, August 1987, pp. 323–344.

[15] Y.Yakimovsky, R. Cunningham, ”A System for Extracting Three–Dimensional Measurements from a Stereo Pair of TV Cameras”, *Computer Graphics and Image Processing*, Vol. 7, pp. 195 –210, 1978.