Camera Viewpoint Control for the Automatic Reconstruction of 3D Objects

Wolfgang Niem, Marcus Steinmetz

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

ABSTRACT

An algorithm for the image dependent control of the camera viewpoint is presented, which is applied to the automatic reconstruction of 3D objects. For computer animation applications, "shape from silhouettes" using equally distributed viewpoints is an often used reconstruction technique. With respect to the local reconstruction errors, the use of equally distributed camera views is unfavourable for arbitrary shaped objects. For that reason, a camera viewpoint control is introduced, which purposefully rotates a turntable with the 3D object depending on the trace of the silhouette contour points over the rotation angle. This trace provides information about the location of object planes and gives a measure for the expected local 3D reconstruction errors. It turns out, that the new algorithm reduces the remaining 3D reconstruction errors up to 70 % compared to algorithms without viewpoint control using the same number of viewpoints.

1 INTRODUCTION

Recently the need for realistic 3D models of real objects is growing in a variety of computer animation applications. Virtual studio production is one of those applications and has been followed by the European project MONA LISA [2] (RACE 2052). MONA LISA was completed at the end of 1995, involved 9 partners and resulted in the ELSET virtual studio (http://www.dcs.qmw.ac.uk/research/parallel/monalisa/). The goal of MONA LISA was the integration of image analysis and synthesis techniques for the mixing of 3D virtual scenes and real actors. Those virtual background scenes are composed of different virtual 3D models, whose manual creation is time consuming and therefore cost expensive. For that reason, one task of MONA LISA was the development of algorithms performing the reconstruction of 3D objects automatically with simple equipment and in reasonable time.

One of the 3D reconstruction algorithms developed in MONA LISA aims at the reconstruction of separated objects and is based on the evaluation of silhouettes from multiple camera views [1][4][5] (shape from silhouettes, method of occluding contours). Beside the inability of this algorithm to detect object concavities, reconstruction errors occur due to the limited number of silhouettes used for the reconstruction process. Especially for the reconstruction of objects containing planes as well as for objects with a distinctive direction of expansion and varying curvature, the use of equally distributed camera views is unfavourable. The resulting reconstruction errors heavily depend on a proper choice of the evaluated viewpoints.

For that reason, the reconstruction algorithm has been improved by using a controlled vision approach [3][6] (also referenced as "active vision"). Therefore a camera viewpoint control is introduced, which purposefully rotates a turntable with the 3D object depending on the trace of the silhouette contour points over the rotation angle.

In the second section the basics of shape from silhouettes are described. The properties of the trace of silhouette contour points $C(\phi)$ are discussed in the third section. In section four an assessment of the local 3D reconstruction error with respect to $C(\phi)$ is performed. The fifth section explains the applied viewpoint control, while in section six some results are presented.

2 SHAPE FROM SILHOUETTES

One approach for the reconstruction of 3D objects using silhouettes from multiple camera views is called "shape from silhouettes "or "method of occluding contours" [1][4]. The principle of this approach can be divided into three steps.

In a first step, the silhouette of the real object is extracted from the camera images. In controlled environments, this segmentation of the real object against the background can be facilitated by using a background of uniform colour.

In a second step, a bounding pyramid is constructed using the focal point of the camera and the silhouette. The convex hull of this pyramid is formed by the lines of sight from the camera focal point through different contour points of the object silhouette.

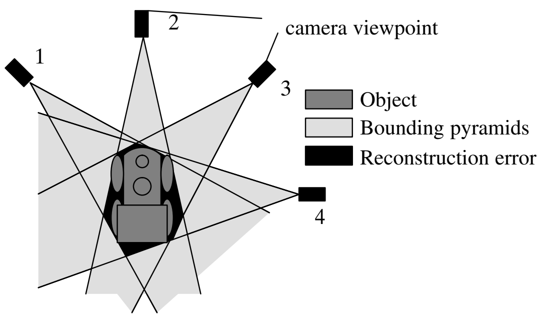

Fig. 1 Two dimensional description of the intersection of the bounding pyramids

Finally, the pyramids are intersected and form an approximation of the bounding volume (Fig. 1). The resulting reconstruction error heavily depends on the number and choice of viewpoints. Therefore it seems to make sense to introduce a viewpoint control.

3 TRACING CONTOUR POINTS

The change of the silhouette contours over the rotation angle and its relation to the viewed local shape of the 3D objects has been investigated only by few studies [6]. This paper will give a connection between the trace of silhouette contour points and the expected local reconstruction error.

For simplicity, orthogonal projection of the real world into the image plane (which can almost be achieved by using a lens with a long focal length) and a camera orientation perpendicular to the rotation axis is assumed. Furthermore, only one cross section of a 3D object parallel to the rotational plane is considered for clarity; the projection of this cross section results in a horizontal line in the image plane.

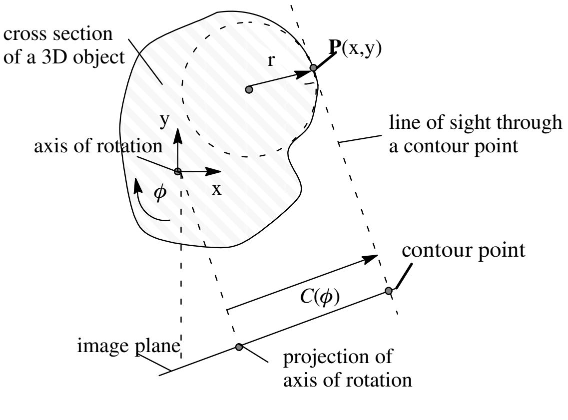

Fig. 2 Trace of silhouette contour points as the angle–varying distance $C(\phi)$ from the projection of the rotational axis to a silhouette contour point; the radius of curvature at the corresponding 3D point $\boldsymbol{P}(x, y)$ is r.

While the object rotates, a silhouette contour point results from the projection of an object surface point $\boldsymbol{P}(x, y)$ (Fig. 2). The trace of silhouette contour points is now defined as the angle–varying distance $C(\phi)$ from the projection of the rotational axis to a silhouette contour point. It provides information about the 3D object shape.

For the developed viewpoint control the properties of $C(\phi)$ are investigated. Especially the characteristics of $C(\phi)$ for 3D objects with a smooth convex surface and with plane surface regions are explored in the following.

$C(\phi)$ for 3D objects with a smooth convex surface

For smooth convex surfaces the trace of silhouette contour points $C(\phi)$ is differentiable for all ϕ. As derived in [6], for this case the radius r of the local surface curvature can be calculated from $C(\phi)$ by

$$r \quad = \quad C(\phi) + \frac{\partial^2 C(\phi)}{\partial \phi^2} \qquad (1).$$

$C(\phi)$ for 3D objects with plane surface regions

A plane surface region of a 3D object results in a non–smoothness in $C(\phi)$, where $C(\phi)$ is not differentiable (Fig.3). This results in a discontinuity in $C^{'}(\phi) = \frac{\partial C(\phi)}{\partial \phi}$ where Eq. (1) can not be applied. This discontinuity can also be observed for concave or occluded convex surface regions.

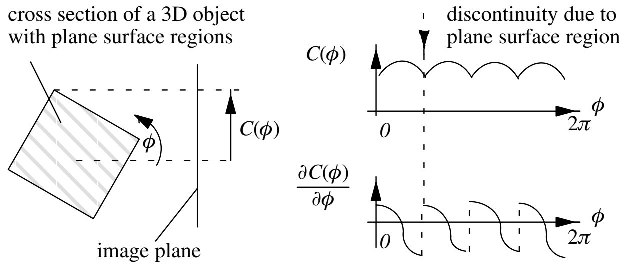

Fig. 3 Example of $C(\phi)$ for 3D objects with plane surface regions

4 LOCAL 3D RECONSTRUCTION ERROR

Due to the limited number of evaluated silhouettes the reconstruction of a 3D object results in a local 3D reconstruction error. For plane object surface regions, this error can be reduced to zero by a proper choice of viewpoints (Fig.4).

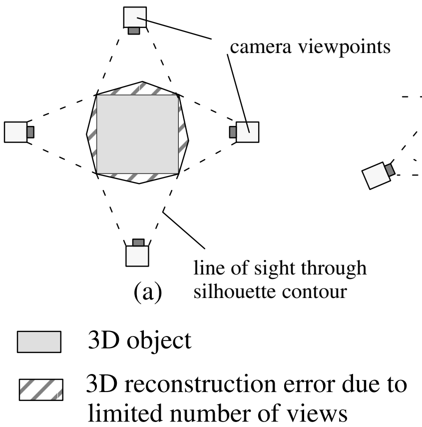

Fig. 4 Cross section of a 3D object with plane surface regions and two different sets of camera viewpoints (a) unfavourable choice and (b) favourable choice of camera viewpoints.

For object regions with a smooth convex surface, this error depends on the angle $\Delta\phi$ between two consecutive viewpoints and the radius r of curvature of the observed 3D surface part.

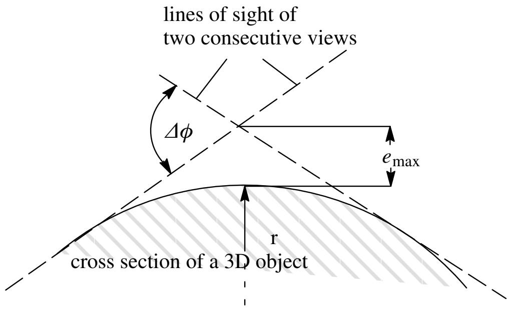

Fig. 5 Illustration of the local reconstruction error $e_{\max}$

For the assessment of this error, in Fig. 5 the radius r of curvature is assumed to be locally constant. Then it follows for the maximum local reconstruction error $e_{\max}$ as defined in Fig. 5:

$$e_{\max} = r\left[\frac{1}{cos(\frac{\Delta\phi}{2})} - 1\right] \approx r\frac{(\Delta\phi)^2}{8} \quad ,\Delta\phi < 30^o \qquad (2)$$

It turns out, that the error grows linearly with the radius r of curvature and quadratically with the angle $\Delta\phi$; for surface regions with a low curvature therefore the local reconstruction error is larger than for surface regions with high curvature.

5 VIEWPOINT CONTROL

As a consequence of the discussion in the previous sections, the algorithm to be developed aims at a viewpoint control with respect to an upper limit of the maximum local reconstruction error. Hence, the implemented analysis of $C(\phi)$ considers

- the detection of plane object regions in order to find camera viewpoints, which reduce the local reconstruction error almost completely and
- the evaluation of the local reconstruction error in order to ensure equally distributed reconstruction errors over the rotation angle ϕ.

The developed viewpoint control consists of the following steps:

1. Acquisition of a first set of camera views from equally distributed viewpoints

From this set, an initial trace of silhouette contours $C(\phi)$ is computed.

2. Detection of plane object regions

As pointed out in section 4, a plane surface region of the 3D object results in a non–smoothness of $C(\phi)$ and hence in a discontinuity of its first derivative $C'(\phi)$. As shown in Fig.3, a plane surface region coincides with a large jump from negative to positive values of $C'(\phi)$.

This jump is detected from discrete values of $C(\phi)$ by using difference equation calculus. Iteratively, new viewpoints are introduced close to the detected jump in order to make the detection more precise.

3. Evaluation of maximum local reconstruction error

By combining Eq. (1) and Eq. (2) the maximum local reconstruction error $e_{\max}$ can be assessed directly from $C(\phi)$:

$$e_{\max} = \frac{(\Delta\phi)^2}{8}\left(C(\phi) + \frac{\partial^2 C(\phi)}{\partial\phi^2}\right) \qquad (3).$$

Expanding Eq. (3) in a difference equation it follows for the assessed maximum local reconstruction error $\tilde{e}_{\max}$:

$$\tilde{e}_{\max} = \frac{1}{8}\big(C(\phi - \Delta\phi) + C(\phi + \Delta\phi) + (\Delta\phi^2 - 2)C(\phi)\big)$$
(4)

which can be computed from three consecutive values of $C(\phi)$ belonging to discrete angle steps.

Intermediate camera views are acquired for viewpoints, where the calculated maximum local reconstruction error $\tilde{e}_{\max}$ exceeds a predefined threshold.

6 RESULTS

The described algorithm has been tested for real objects placed on a computer controlled turntable in front of a calibrated CCD camera. A camera lens with a long focal lens has been used in order to approximate orthogonal projection. The results with and without viewpoint control have been compared for the objects ”box” and ”rugby ball”, using 24 viewpoints each (Fig. 6).

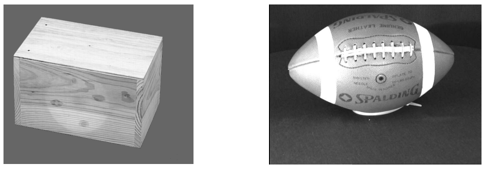

Fig. 6 Original images of the objects ”box” and ”rugby ball”

For comparison, a volume model has been calculated for the selected viewpoints for both algorithms and compared with the true volume. In Fig. 7 a cross section parallel to the rotational plane of the volume model ”box” is shown for both algorithms. Clearly, the viewpoint control results in a better reconstruction of the plane object regions.

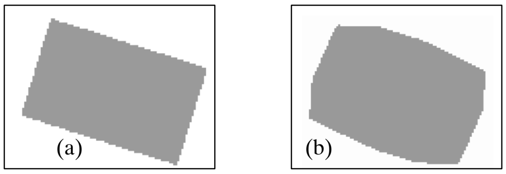

Fig. 7 Cross section parallel to the rotational plane of the volume model ”box” reconstructed (a) with viewpoint control (b) without viewpoint control

For numerical comparison, the reconstruction error for both algorithms have been evaluated and normalized to the reconstruction errors of the algorithm without viewpoint control (Fig. 8). It shows, that the reconstruction error is reduced by 70 % for the object ”box” and 30% for the object ”rugby ball” by using viewpoint control.

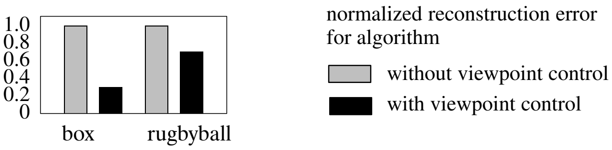

Fig. 8 Comparison of the normalized reconstruction error for algorithms with and without viewpoint control.

7 CONCLUSIONS

This paper presents an algorithm for the camera viewpoint control applied to the automatic reconstruction of 3D objects. It is based on a reconstruction approach called ”shape from silhouettes”. In contrast to formerly presented algorithms, it uses a control strategy for the choice of favourable camera viewpoints with respect to the maximum local reconstruction error.

For this purpose, the relation between the trace of silhouette contour points $C(\phi)$ over the rotation angle and the surface points of a 3D object has been investigated. It turns out, that for smooth convex surfaces the maximum local reconstruction error can be directly computed from $C(\phi)$. For surfaces with plane regions the location of the planes can be detected from $C(\phi)$, as they result in discontinuities in its first derivative with respect to the rotation angle.

The implemented viewpoint control takes a first set of camera views from equally distributed viewpoints. From this set, plane surface regions are detected from values of $C(\phi)$ at discrete angle steps, whereby locally new viewpoints are introduced in order to make the detection more precise. Afterwards, the maximum local reconstruction error is calculated for all viewpoints and intermediate camera views are acquired as long as the maximum local reconstruction exceeds a predefined threshold.

The algorithm has been tested with real 3D objects and has been compared with algorithms without viewpoint control by using the same number of viewpoints. It could be shown, that the new algorithm reduces the remaining 3D reconstruction errors up to 70 % depending on the kind of object.

REFERENCES

[1] H. Busch, ”Automatic Modelling of Rigid 3D Objects Using an Analysis by Synthesis System”, *SPIE Proceedings ”Visual Communication and Image Processing IV”*, 1989

[2] L. Blonde, ”The MONA LISA Project” , *Proceedings on the European Workshop on Combined Real and Synthetic Image Processing for Broadcast and Video Production,* VAP Media Centre, Hamburg, 1994

[3] K.N. Kutulakos and C.R. Dyer, ”Recovering Shape by Purposive Viewpoint Adjustment”, *International Journal of Computer Vision”*, 12:2/3. pp. 113–136, 1994.

[4] W. Niem, ”Robust and Fast Modelling of 3D Natural Objects from Multiple Views”, *SPIE Proceedings ”Image and Video Processing II”*, Vol. 2182, 1994, pp. 388–397.

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

[6] J.Y. Zheng, ”Acquiring 3–D Models from Sequences of Contours”, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 16, No. 2, February 1994.