For: koch
Printed on: Tue, Jul 11, 1995 17:19:32

Document: hawaii
Last saved on: Tue, Jul 11, 1995 15:35:30

SHAPE ADAPTATION FOR MODELLING OF 3D OBJECTS IN NATURAL SCENES

C.−E. Liedtke, H. Busch, R. Koch
Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung
Universität Hannover, D−3000 Hannover 1, Germany

Abstract

Rigid 3D objects are modelled automatically from an image sequence taken by a camera which is rotated around the object. The image sequence was recorded using a calibrated camera which allows to measure the camera positions and to estimate the true object size. The 3D object shape is obtained in two steps. The object silhouettes are employed to find the enclosing volume of the object. The volume is converted into a flexible surface representation and the 3D shape is refined based on the texture information of the object surface. Finally texture mapping is applied to generate a highly realistic 3D model of the object.

1. Introduction

A procedure to automatically model arbitrary shaped objects from a series of images is presented. The object is observed from different known view points. An automatic calibration process serves to determine the camera positions in advance. The objects are assumed to be opaque and fully visible from all view points and the scene illumination is diffuse. The objects need to be roughly convex, but minor concavities are allowed. Besides these conditions no a priory knowledge is necessary.

2. Shape from Multiple Silhouettes

For the generation of the initial object shape the 2D−silhouettes of all input images are employed to generate a 3D volume segmentation. A bounding volume is defined and the object is assumed to be inside this volume. The volume is partitioned as a 3D voxel space, where each voxel represents a volume element of the 3D−space. The silhouettes from the different camera positions are projected through the bounding volume. The common volume between all silhouette projections and the bounding volume is the enclosing volume of the 3D object. Fig. 1 demonstrates the volume segmentation of a toy cow. The quality of this initial shape determination depends on the number of views and the complexity of the object shape.

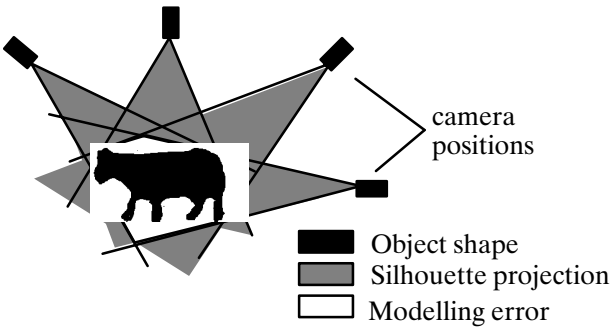

Fig. 1. Volume segmentation from multiple silhouettes

3. Flexible surface models

To further enhance the model shape the voxel representation is converted into a flexible surface representation using a closed mesh of triangles. A triangulation method which creates a near ideal flexible mesh has been developed and published by Busch [2]. The triangular mesh is built across the surface of the voxel representation through an adaptive, locally controlled growing process. Fig. 2 shows a wireframe representation of the modelled toy cow. The cow was modelled using 12 different views placed equally around the object.

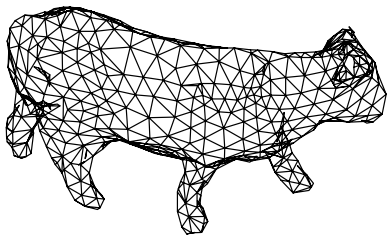

Fig. 2. Wireframe representation of cow modelled from 12 different views

4. Adaptation of Model Shape

In the previous chapters methods have been described to derive a first estimate of the model shape from a sequence of images based on object silhouettes. In a further adaptation step the modelling of the shape is refined based on the photometric information from the input image sequence using an analysis by synthesis method. From each image the texture which provides the best quality is projected onto the model object using texture mapping. A synthetic image sequence is computed from the model and compared to the real sequence. Differences between real and synthesized images are due to shape errors. Minimizing the image differences through shape corrections adapts the model shape towards the real object shape. The shape adaptation process is described in detail by Kappei [3] and Liedtke et. al.[4] and will be explained shortly in the following section.

4.1 Determination of the Shape Parameter

Each vertex of the triangular mesh can be shifted in space to adapt the surface shape of the modelled object. It is therefore necessary to estimate the shift vector for each vertex from the image sequence. Fig. 3 explains the geometry used for the shape adaptation. Epipolar geometry between two camera positions $\mathbf{C}$ and $\mathbf{C}_1$ is applied. A displacement s$\mathbf{V}$ of Vertex $\mathbf{P}$ along the line of sight $\mathbf{V}$ yields a displacement D in the image plane of camera $\mathbf{C}_1$.

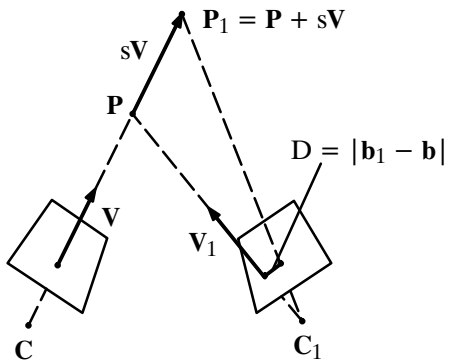

Fig. 3. Shape adaptation through displacement analysis

When the displacement D is only few pixels in the image plane and the illumination does not change, a brightness constancy constraint can be assumed between the views $\mathbf{C}$ and $\mathbf{C}_1$ and a linear relationship between the spatial displacement $\mathbf{b}_1 - \mathbf{b}$ and the corresponding image intensities $I(\mathbf{b}_1) - I(\mathbf{b})$ can be formulated:

$$I(\mathbf{b}_1) - I(\mathbf{b}) = \mathbf{G}(\mathbf{b}_1 - \mathbf{b})$$

with $\mathbf{b} = (b_x, b_y, f)^T$ image plane projection of $\mathbf{P}$

and $\mathbf{G} = (G_x, G_y, 0)^T$ spatial image plane gradient

Substituting the perspective projections of $\mathbf{P}_1$ and $\mathbf{P}$ for $\mathbf{b}_1$ and $\mathbf{b}$ and setting $\mathbf{P}_1 = \mathbf{P} + s\mathbf{V}$ as in Fig. 2, the shape parameter s can directly be calculated:

$$s = \frac{P_z^2[I(\mathbf{b}_1) - I(\mathbf{b})]}{fP_z\mathbf{GV} - fV_z\mathbf{GP} - P_zV_z[I(\mathbf{b}_1) - I(\mathbf{b})]}$$

4.2 Shape Improvement by Sequence Analysis

When more than two different views $\mathbf{C}_i$ of a point $\mathbf{P}$ on the surface of the object are available, then a position estimate of $\mathbf{P}_i$ can be made from each view. Under ideal conditions, the estimates would be the same and identical to the true position, i.e. $\mathbf{P}_1 = \mathbf{P}_2 = ... = \mathbf{P}_i$. Since this is not the case under real conditions, a certainty measure a_i for each position estimate $\mathbf{P}_i$ is introduced. The most reliable estimate $\mathbf{P}_N$ for the position $\mathbf{P}_i$ from N observations can be made by a linear superimposition of the individual estimates $\mathbf{P}_i$ weighed by their certainty factors a_i:

$$\mathbf{P}_N = \left(\sum_{i=1}^{N} a_i\mathbf{P}_i\right) / \left(\sum_{i=1}^{N} a_i\right)$$

The results of the automatic modelling of a 3D model from N=12 images is demonstrated in Fig. 4. A computer generated image of the 3D model of the textured cow is shown.With this approach it is possible to model complex shapes and generate highly realistic image sequences from the model with the computer.

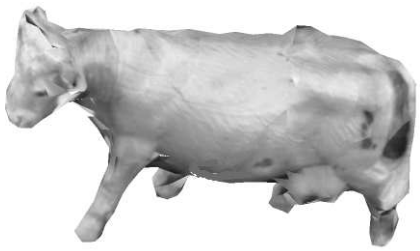

Fig. 4. Computer generated view of the modelled toy cow

5. Acknowledgements

The work has been supported by the Forschungsinstitut der DEUTSCHEN BUNDESPOST TELECOM.

6. References

[1] R. Szeliski, "Real−Time Octree Generation from Rotating Objects" and "Shape from Rotation", Cambridge Research Laboratory, Technical Report Series, Cambridge, USA, 1990.

[2] H. Busch, "Automatic Modelling of Rigid 3D Objects Using an Analysis by Synthesis System", SPIE Sym-

posium on Visual Communications and Image Processing, Philadelphia, USA, 1989.

[3] F. Kappei,"Modellierung und Rekonstruktion bewegter dreidimensionaler Objekte in einer Fernsehbildfolge", Ph.D.−Thesis, Universität Hannover, 1989.

[4] C.−E. Liedtke, H. Busch, R. Koch,"Automatic Modelling of 3D Moving Objects from a TV Image Sequence", SPIE/SPSE Symposium on Electronic Imaging, Santa Clara, USA, 1990.