Automatic Shape Reconstruction of Rigid 3-D Objects from Multiple Calibrated Images

Gerald Eckert
Institut fuer Theoretische Nachrichtentechnik und Informationsverarbeitung, University of Hannover,
Appelstr. 9a, 30167 Hannover, GERMANY
Tel: +49 511 762-5320; fax: +49 511 762-5333
e-mail: eckert@tnt.uni-hannover.de

ABSTRACT

We present a method for reconstructing the shape of a rigid 3-D object based on multiple calibrated images from an arbitrary camera setup. The algorithm is robust and able to reconstruct concavities. It is based on the idea to reconstruct a photo-consistent shape by carving and coloring voxels from a volume description of the object. Coloring and carving steps are carried out with respect to a color verification from the input images for each visible voxel. As compared to other approaches, the threshold used for the color verification is adapted automatically. Information included in the input images is processed and applied simultaneously without any explicit integration or assignment to the volume. The approach is evaluated using camera images of a real scene.

1 Introduction

For the visualisation of rigid 3-D objects in multimedia or internet applications, techniques for an automated creation of the object model are required. The easiest way for a user to get information about an object is to take images of the object from different viewpoints with a camera. The camera can be calibrated using existing methods [6]. The problem is to reconstruct the object from these calibrated camera images.

Disparity based depth estimation (stereo) by block-matching [1] is a suitable solution for a binocular image pair, because disparity is examined only between two images per time. Processing more than two images requires to merge the depth information of different depth maps. This step is not trivial and the depth maps often contain regions where no correspondence was found.

Our goal is the reconstruction from multiple calibrated camera images. The shape-from-silhouette approach [2] calculates the intersection of backprojected object silhouettes. This algorithm achieves robust results but an explicit segmentation step is required. The result is the observed visible hull of the object. Concavities in the object can not be reconstructed.

A class of algorithms that are able to reconstruct concavities is based on a volume description of the object by voxels. The voxels are colored or carved according to the color of the input images at the locations of the projected voxel into the images [3, 4, 5]. We present an algorithm from this class that combines the ability to reconstruct concavities of the voxel coloring algorithms with a robustness close to the shape-from-silhouette algorithm.

2 Reconstruction of a 3-D object

The reconstruction algorithm aims at the reconstruction of the maximal photo-consistent shape [3], i.e. the largest shape reproducing the input images when colored accurately.

The proposed algorithm consists of five steps:

- Voxel-volume initialization,
- Determination of the voxel visibility,
- Color verification,
- Carving of inconsistent voxels,
- Adaptation of the applied threshold.

2.1 Voxel-volume Initialization

For initialization of the voxel-volume we use a shape-from-silhouette algorithm [2] in order to obtain a suitable volume initialization. This initial shape is consistent with all evaluated object silhouettes.

2.2 Determination of the voxel visibility

Due to occlusions a voxel is not visible in each image. Each voxel on the object surface is projected into each input image. The visibility of the voxel from an image is verified by calculating possible intersections of the projection ray with other voxels. A voxel that is visible from at least one image is considered to be a visible voxel.

2.3 Color Verification

The colors at the locations of a projected voxel in the visible input images are added to a set of colors $\mathbf{C} = \{C_1, \ \ldots, \ C_n\}$. A voxel that is visible from at least two images is going to be color verified. Therefore each color

C_i of the set of colors $\mathbf{C}$ is compared to each other color C_j.

$$\Delta C_{ij} = \|C_i - C_j\| < d \qquad (1)$$

If the difference ΔC_{ij} is less than an assigned threshold d then the color pair is considered as photo-consistent. The voxel is defined to be non-photo-consistent if there exists one pair of colors in the set $\mathbf{C}$ that is not photo-consistent.

We used the RGB-color space with an euclidian color norm.

$$\|C\| = \sqrt{R(C)^2 + G(C)^2 + B(C)^2} \qquad (2)$$

2.4 Carving of inconsistent voxels and Coloring of consistent voxels

After the color verification for each visible voxel, the non-photo-consistent voxels are removed from the object. Remaining voxels are colored with the mean color of the set $\mathbf{C}$. Due to the voxel removal, the visible surface at the location of a carved voxel moves one voxel inside the object.

The steps visibility determination, color verification, voxel carving and voxel coloring are iteratively repeated. The reconstruction terminates if no non-photo-consistent voxel can be carved from the volume, that means the maximal photo-consistent shape is reconstructed. Due to the visibility determination for each voxel, an explicit plane sweep is not necessary and the algorithm works for arbitrary camera setups. The initialization with the shape-from-silhouette algorithm results in an initial object that is already consistent with the object silhoutte. It avoids unnecessary iteration steps for carving voxels outside the visible hull when initalizing with the bounding box of an object.

2.5 Threshold adaptation

The algorithm reconstructs a photo-consistent object referred to the chosen threshold d and the color norm $\|C\|$. This color norm is based on a color space which has been used.
The reconstruction thus depends on

- the threshold d and,
- the used color space and color norm.

The idea is to realize an automatic adaptation of the threshold d so that the algorithm looses its dependency on an user-defined value.

We have to consider that the task of the algorithm is to reconstruct an object that reproduces the input images. For a solution we take this condition as constraint for the reconstruction process by a comparison between the input images and the reproduced images rendered from the reconstructed and colored object. The outcome of this comparison is an error criterion Q for the current reconstruction. We choose a normalized euclidean distance between the K pixels of the input images I_{in} and the rendered images I_{rec} from the reconstruction and we applied the same color norm as for the color verification.

$$Q = \frac{\sum_{k=1}^{K} \|I_{in,k} - I_{rec,k}\|^2}{\sum_{k=1}^{K} \|I_{in,k}\|^2} \qquad (3)$$

We regard the criterion Q as function of the used threshold d. When decreasing d from high to low values, we expect $Q(d)$ to decrease monotonically until a minimum is reached. If we continue decreasing d we expect the criterion $Q(d)$ to increase due to damaging the object by removing too many voxels. The local mimimum of $Q(d = d_{opt})$ first reached when decreasing d corresponds to a maximal photo-consistent shape. Therefore, it is quite likely, that the reconstructed object has correctly reconstructed concavities and is still consistent with the input silhouettes. Due to this, the robustness of the reconstruction is comparable to that of the shape-from-silhouette algorithm.

$$d_{opt} = \max\left\{d \mid Q(d) < Q(d+h)\right\} \qquad (4)$$

We define the reconstruction result as the object reconstructed with the threshold d_{opt}.

3 Reconstruction Results

As an example for a real scene, we used an egyptian statue which has strong concavities in the regions of the cavernous eyes. We took a sequence of 36 input images of the object standing on a turntable. Three of them are shown in figure 1 (left column).

In figure 3 the initial object that results from the shape-from-silhouette algorithm is shown in voxel representation (uncolored and colored). The concavities of the original object are missing. Rendered images of the same object are shown in figure 1 (middle column).
Here photo-inconsistency is observable in the region of the eyes (blurred colors) due to the missing concavities.

The criterion $Q(d)$ during the reconstruction process can be seen in figure 2. The threshold d was decreased in steps of one unit and the first local minimum regarded as reconstruction result is reached with $d_{opt} = 133$. In this case the first local minimum is the global minimum. The right part of the curve (for $d > d_{opt}$) corresponds with the reconstruction process and is quite smooth while the part that represents the damaging of the object due to a small threshold (for $d < d_{opt}$) shows a steep increase.

The reconstructed object is shown in figure 4. The concavities have been reconstructed while the silhouette was not violated. The same object is shown in figure 1 (right column), where the achieved photo-consistency can be seen. The input images are now projected on the

Figure 1: *left column: Input images, middle column: Rendered images of the inital object, right column: Rendered images of the reconstructed object*

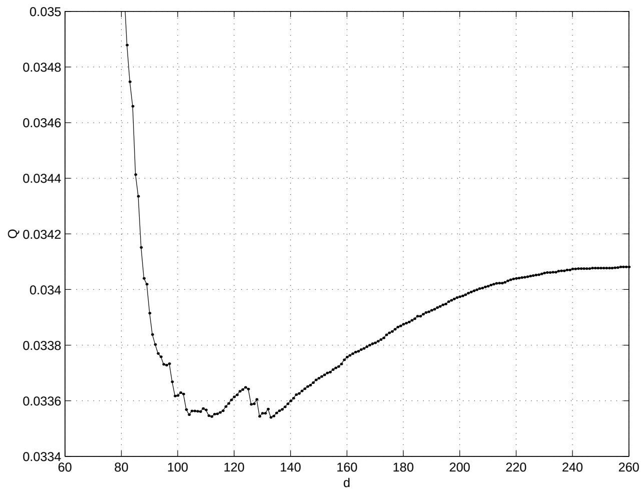

Figure 2: *Result for $Q(d)$ during the reconstruction process*

correctly reconstructed object shape producing sharp renderings.

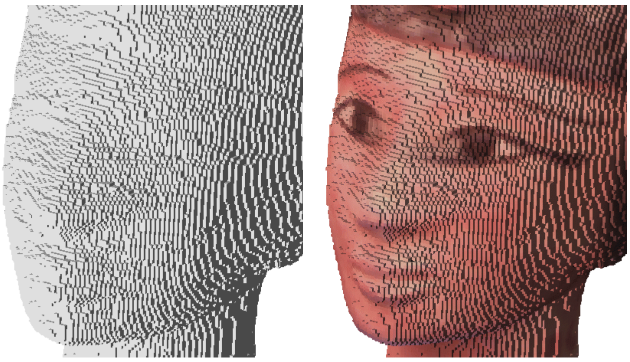

Figure 3: *Initial object*

4 Conclusion

In this paper we presented an algorithm that achieves an automated volume reconstruction of a rigid 3-D object based on mutiple calibrated images from arbitrary camera setups. The voxels of an initialization with the shape-from-silhouette algorithm are carved or colored depending on a color verification of the input images at the locations of the projected voxels. An used threshold adapts automatically by minimizing an error criterion. The reconstructed object is photo-consistent with the input images and can be used for visualization purposes.

References

[1] Lutz Falkenhagen, "Hierarchical Block-Based Disparity Estimation Considering Neighbourhood Constraints", International workshop on SNHC and 3D Imaging, September 5-9, 1997, Rhodes, Greece.

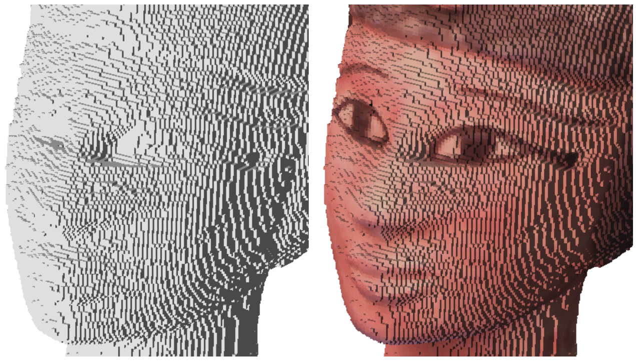

Figure 4: *Reconstructed object*

[2] Wolfgang Niem, "Automatische Rekonstruktion starrer dreidimensionaler Objekte aus Kamerabildern", PhD Thesis, University of Hannover, 1999.

[3] Kiriakos N. Kutulakos, Steven M. Seitz, "A Theory of Shape by Space Carving", University of Rochester, CS Technical Report No. 692, 1998.

[4] Steven M. Seitz, Charles R. Dyer, "Photorealistic Scene Reconstruction by Voxel Coloring", Int. J. Computer Vision, to appear 1999.

[5] Peter Eisert, Eckehard Steinbach und Bernd Girod, "Multi-hypothesis, Volumetric Reconstruction of 3-D Objects from Multiple Calibrated Camera Views", ICASSP 99, pp. 3509-3512, Phoenix, 1999.

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