Reconstruction of Three Dimensional Models from Real Images *

Ulaş Yılmaz, Adem Mülayim, Volkan Atalay
Department of Computer Engineering,
Middle East Technical University,
Ankara, Turkey, TR-06531

Abstract

An image based model reconstruction system is described. Real images of a rigid object acquired under a simple but controlled environment are used to recover the three dimensional geometry and the surface appearance. Based on a multi-image calibration method, an algorithm to extract the rotation axis of a turn-table has been developed. Furthermore, this can be extended to estimate robustly the initial bounding volume of the object to be modeled. The coarse volume obtained, is then carved using a stereo correction method which removes the disadvantages of silhouette based reconstruction by photoconsistency. The concept of surface particles is adapted in order to extract a texture map for the model. Some metrics are defined to measure the quality of the reconstructed models.

1 Introduction

An image-based model reconstruction method is described. Using an off-the-shelf camera, considerably realistic looking models of a 3D rigid object can be reconstructed from its 2D images [11, 1, 7, 6, 12]. The goal of this study is to investigate the image-based reconstruction of 3D graphical models of real objects in a controlled imaging environment and present the work done in our group for such a reconstruction. Although many parts of the whole system have been well-known in the literature and in practice, there are contributions of this study which can be stated as follows: developing a vision based calibration algorithm and its comparison, extraction of the rotary axis, estimation of the initial bounding cube of the acquired object, removing the disadvantages of the reconstruction approach based on silhouettes and volume intersection using photoconsistency, use of particles for a smooth appearance recovery. The organization of the paper is as follows: we first describe our geometry reconstruction process in the following section. Section 3 gives the detailed description of our appearance recovery algorithms. Results obtained in the framework of our study are given in Section 4. The paper concludes with Section 5.

*This study is partially supported by TÜBİTAK (Scientific and Technical Research Council of Turkey) under the grant EEEAG-199E024, by METU Graduate School of Natural and Applied Sciences and by METU Research Grant. Adem Y. Mülayim's study is supported by TÜBİTAK Integrated PhD Program.

2 Geometry reconstruction

In order to compute the parameters of the camera, we use a multi-image calibration approach [6]. Our acquisition setup is made up of a rotary table with a fixed camera as shown in Figure 1. The rotation axis and distance from the camera center to this rotation axis remain the same during the turns of the table. Based on this idea, we have developed a vision based geometrical calibration algorithm for the rotary table [8]. Furthermore, we can compute very easily the distance between the rotation axis of the table with respect to the camera center which in fact facilitates the calculation of the bounding cube [7].

Figure 1. System setup.

Once the bounding volume is obtained, carving this volume by making use of the silhouettes, a coarse model of the object is computed. This volume has some extra voxels which in fact should not exist. In this context, we have implemented a stereo correction algorithm which removes these extra voxels using photoconsistency [6]. Algorithm 1 which is mostly inspired from Matsumoto et. al. [5] outlines the process.

Algorithm 1 Computing the photoconsistent voxels.

reset all photoconsistency values of the voxels in V_{object} to max photoconsistency value
for all image i in the image sequence **do**
  for all visible voxels in image i **do**
    produce a ray from camera optic center
    find max photoconsistent voxel on the ray
    for all voxels between the max photoconsistent voxel and camera optic center **do**
      reduce voxel photoconsistency votes
    end for
  end for
end for
for all voxel v in voxel space V_{object} **do**
  if the photoconsisency of v is less than a threshold **then**
    remove v from V_{object}
  end if
end for

In the algorithm, each voxel in the object voxel space V_{object}, starts with a high photoconsistency vote value; that is each voxel on the model generated by the silhouette based reconstruction is assumed to be on the real object surface. Each view i is then processed in the following manner. For each view i, rays from the camera center c_i through the voxels seen from that view i are traversed voxel by voxel. Each voxel on the ray is projected onto the images $i-1, i, i+1$ and the voxel's photoconsistency value is calculated using texture similarity measures among the projection regions on the images $i-1, i, i+1$. Then, the voxel with maximum photoconsistency value is found and all the voxels existing between this voxel and the camera center c_i loose votes in an increasing order as they become closer to c_i. This process is repeated for all the rays which can be generated from the view i. When the process is performed for all the views, excess voxels caused by the silhouette based reconstruction lose most of their initial photoconsistency votes. Then by thresholding, this excess volume is carved. Figure 2 explains better the idea. In this figure, darkest colored voxels get the highest voting; i.e. they have the maximum texture similarity according to the algorithm. The color of voxel shows its vote.

3 Appearance reconstruction

There exist several studies for appearance reconstruction of 3D models from real images [10, 2, 4, 9]. In most of these studies, the model is represented as a triangular wireframe,

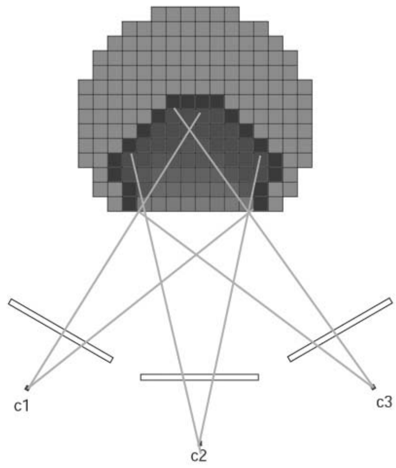
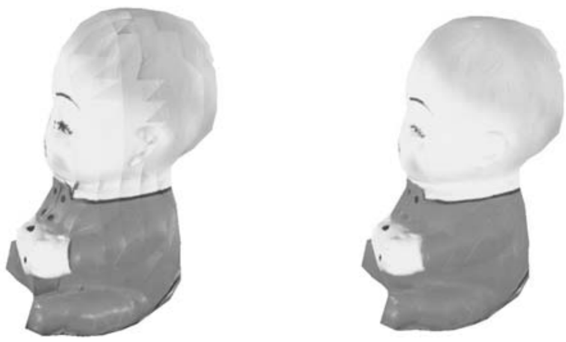


Figure 2. Darkest colored voxels get the highest voting.

and each triangle is associated with one of the images for texture extraction. The method causes discontinuities on the triangle boundaries as shown in Figure 3-a, since adjacent triangles can be associated with different source images. Applying low pass filter on the boundaries cannot come up with a global solution.



(a) (b)

Figure 3. (a) Discontinuities on the triangle boundaries, (b) reconstruction using particles.

In this study, 2D texture mapping is used but to reduce the drawbacks due to the lack of third dimension information, the concept of surface particles is adapted [12, 13]. An abstraction is done on the actual representation of the model: the model is considered to be a surface composed of particles with three attributes: position, normal and color. While reconstructing the appearance of the model, instead of associating triangles to images, particles are associated with images for texture extraction. This is what makes the proposed method superior to the others: since a triangle is not necessarily textured from a single image, there are not discontinuities on the triangle boundaries due to the fact of being textured from different images. Each particle on the surface is associated with a pixel on the texture map, and the color information of the texture map is recovered by this means.

A particle is not necessarily visible in all of the images in the sequence, it can be occluded or a back face in some of them. So, the extraction takes place at two steps: visibility check and color retrieval. Visibility check is performed using the particle normal and the particle position by simple hidden surface removal and occlusion detection algorithms. The particle is projected on the source images in which it is visible, and a set of candidate color values, $C = \{c_0, ..c_{M-1}\}$ are collected. In this study, the candidate color values are fused in order to produce the most photoconsistent appearance. Before assigning a color value to a particle, it is decided whether the information extracted from the source images is photoconsistent or not. The photoconsistency is defined in Definition 1. The value of Θ is empirical and set to 60. It is expected that the values of a photoconsistent set concentrates around the view-independent color of the particle. This method is very suitable for removing illumination artifacts as shown in Figure 3-b. However, if the geometry of the object is not constructed precisely, the photoconsistency criteria will fail for most of the particles, which will cause irregularities on the surface appearance.

Definition 1 *Let the extracted color values for a given particle be $C = \{c_0, c_1, ..c_{M-1}\}$ for an image sequence $S = \{I_0, I_1, ..I_{N-1}\}$. The color of the particle for this sequence is photoconsistent if*

1. *There exists at least two images in S in which the particle is not occluded.*
2. *The particle is not on the background in any of the images in S.*
3. *$C_\sigma < \Theta$ where C_σ is the standard deviation of intensity values of the colors in C.*

The overall appearance reconstruction process is described in Algorithm 2: by projecting the particle on the source images, a set of candidate colors is extracted. If the reconstructed set is photoconsistent then the color of the particle is selected as the median of this set. If a particle is occluded in all of the images or a photoconsistent color cannot be extracted from the sequence, there occurs regions whose appearance cannot be recovered on the model. The colors of the particles in these regions are interpolated using the colors of the adjacent particles.

Algorithm 2 Recovering the color of a particle.

```
reset the candidate color set of the particle to empty set
for all images in the sequence do
  if particle is visible in the image then
    project particle on the image
    insert the extracted color in C
  end if
end for
if C is photoconsistent then
  set the color of the particle to the median of C
else
  set the color of the particle to the color of the nearest
  particle whose color is consistent.
end if
```

4 Error analysis and experimental results

In addition to visual comparison, the quality of the reconstructed models is measured with a set of methods. These methods are based on the comparison of original images of the object with rendered corresponding images using the reconstructed model. The images that are used for reconstruction (training images) should not be used for quality measurement (test images). If the training images are also used as the test images, the measurements will be biased. Instead, some of the acquired images are not used in reconstruction but kept as test images to obtain more reliable quantitative analysis. Generally, visual similarity cannot be described by computational similarity. Although there is no universal and accepted method to tell how similar two images are, there still exists in the literature the following methods: Hausdorff distance (HD), root-mean-square error ($RMSE$), direct difference error (DDE), and normalized cross correlation ratio (CO) [3].

The experiments are performed on a personal computer with 512 MB of RAM, Intel PIII 800Hz CPU and 32MB frame buffer. The images are captured with a 2/3” Color Progressive scan CCD camera at a resolution of 1294x1030. While rendering the final model, texture mapping is performed using the routines provided by OpenGL. As shown in Figure 4, the algorithms are successful in removing the highlights. Some further results for two objects are shown in Figure 5. Also the error analysis for the “box” object is given in Table 1.

Table 1. Error analysis for the “box” object.

	HD_{chess}	HD_{city}	$NCCR$	$RMSE$	DDE
$Error$	14.50 %	24.46 %	1.17 %	12.26%	7.38 %

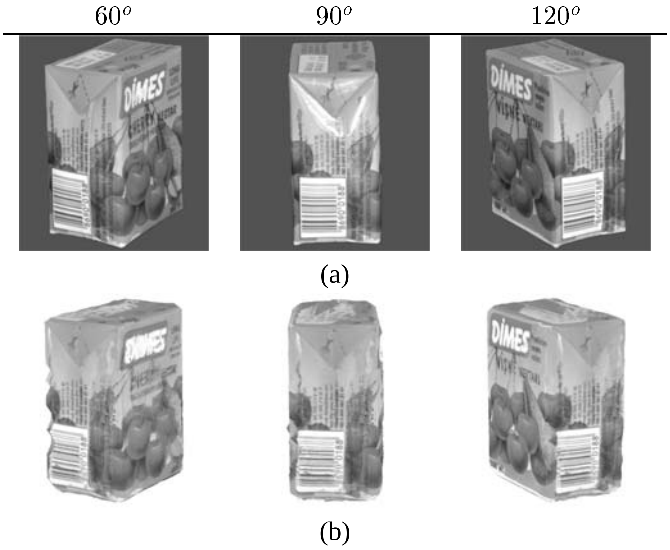


Figure 4. (a) Sample images, (b) sample reconstructions for the "box" object.

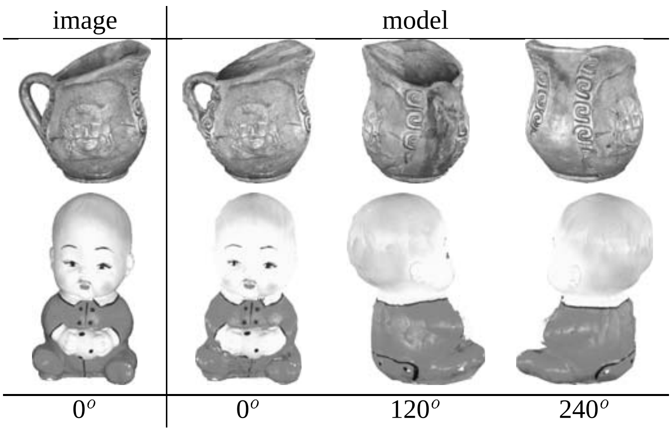


Figure 5. Sample reconstructions.

5 Conclusion

We describe an image based model reconstruction system using real images of rigid objects. Several contributions have been made in this study. A vision based camera calibration algorithm is developed and it is used in the extraction of the rotation axis. The bounding cube of the object is estimated using the extracted rotation axis. To eliminate the drawbacks of the silhouette based reconstruction, a multibaseline stereo correction algorithm is developed by using photoconsistency. The concept of surface particles is adapted in order to extract a texture map for the model. Furthermore, some common metrics are used to measure the quality of the reconstructed models.

References

[1] P. E. Debevec, C. J. Taylor, and J. Malik. Modeling and rendering architecture from photographs: A hybrid geometry and image-based approach. In *Proceedings of SIGGRAPH'96, Computer Graphics Proceedings, Annual Conference Series*, pages 11–20, August 1996.

[2] S. Genc and V. Atalay. Texture extraction from photographs and rendering with dynamic texture mapping. In *Proceedings of the 10th International Conference on Image Analysis and Processing*, September 1999.

[3] V. D. Gesu and V. Starovoitov. Distance-based functions for image comparison. *Pattern Recognition Letters*, 20(2):207–214, 1999.

[4] H. P. A. Lensch, W. Heidrich, and H.-P. Seidel. Automated texture registration and stitching for real world models. In *Proceedings of Pacific Graphics*, pages 317–337, October 2000.

[5] Y. Matsumoto, K. Fujimura, and T. Kitamura. Shape-from-silhouette/stereo and its application to 3-d digitizer. In *Discrete Geometry for Computer Imagery Conference, Lecture Notes in Computer Science 1568*, pages 177–188, 1999.

[6] A. Y. Mulayim and V. Atalay. Multibaseline stereo coreection for silhouette-based 3d model reconstruction from multiple images. In *Proceedings SPIE, Three-Dimensional Image Capture and Applications IV*, pages 24–25, January 2001.

[7] A. Y. Mulayim, O. Ozun, V. Atalay, and F. Schmitt. On the silhouette based 3d reconstruction and initial bounding cube estimation. In *Proceedings of Vision Modeling and Visualisation*, pages 11–18, November 2000.

[8] A. Y. Mulayim, Y. Yemez, F. Schmitt, and V. Atalay. Rotation axis extraction of a turn table viewed by a fixed camera. In *Proceedings of Vision Modeling and Visualisation*, pages 35–42, November 1999.

[9] J. P. Neugebauer and K. Klein. Texturing of 3d models of real world objects from multiple unregistered photographic views. *Computer Graphics Forum*, 18(3):245–256, September 1999.

[10] W. Niem and H. Broszio. Mapping texture from multiple camera views onto 3d object models for computer animation. In *Proceedings of International Workshop on Stereoscopic and Three Dimensional Imaging*, pages 99–105, September 1995.

[11] W. Niem and J. Wingbermhle. Automatic reconstruction of 3d objects using a mobile monoscopic camera. In *Proceedings of International Conference on Recent Advances in 3D Imaging and Modeling*, pages 173–180, May 1997.

[12] F. Schmitt and Y. Yemez. 3d color object reconstruction from 2d image sequences. In *Proceedings of International Conference on Image Processing*, pages 65–69, October 1999.

[13] R. Szelisky and D. Tonnesen. Surface modeling with oriented particles. *Computer Graphics (SIGGRAPH'92)*, 26(2):185–194, July 1992.