INEXPENSIVE AND ROBUST 3D MODEL ACQUISITION SYSTEM FOR THREE-DIMENSIONAL MODELING OF SMALL ARTIFACTS

Ulaş Yılmaz, Oğuz Özün, Burçak Otlu, Adem Mulayim, Volkan Atalay

{ulas, oguz, burcak, adem, volkan}@ceng.metu.edu.tr

Department of Computer Engineering,
Middle East Technical University,
Ankara, Turkey, TR-06531

KEY WORDS: acquisition, calibration, digitization, modeling, registration, triangulation, high resolution, three-dimensional, distortion, image, Internet, object, sequence, texture, web-based, photo-realism, virtual reality.

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. Proposed system enables robustly modeling of small artifacts with high resolution geometry and surface appearance. Such artifacts may have handles, holes, concavities, cracks, etc. The proposed system enables robustly modeling of these properties also. The models obtained by this method are stored in Virtual Reality Modeling Language format which enables the transmission and publishing of the models easier over Internet.

1 INTRODUCTION

Realistic looking three dimensional (3D) models are important components of virtual reality environments. They are used in many multimedia applications such as 3D virtual simulators, 3D educational software, 3D object databases and encyclopedias, 3D shopping over Internet and 3D teleconferencing. The virtual environments enable the user interact with the models. The user can rotate, scale, even deform the models; observe the models under different lighting conditions; change the appearance (color, material, etc.) of the models; observe the interaction of a model with the other models in the environment.

The most common way of creating 3D models for virtual reality environments is manual design. This approach is very suitable for the creation of the models of non-existing objects. However, it is cost expensive and time consuming. Furthermore, the accuracy of the designed model for a real object may not be satisfying. In fact, the number of vertices of a simple model varies between 10000 and 100000, which is relatively high for manual design. In a second approach, geometry of real objects is acquired using a system that captures directly 3D data. Such systems are constructed using expensive equipments such as laser range scanners, structured light; touch based 3D scanners, or 3D digitizers. In most of these active scanning systems, the texture of the model is not captured while the geometry of the object is acquired precisely as a set of points in the 3D space. This set can then be converted to polygonal model representations for rendering (Soucy et al., 1996). In a third approach, the model of a real object is reconstructed from its two dimensional (2D) images. This technique is known as *image-based modeling*. Even using an off-the-shelf camera, considerably realistic looking models with both geometry and texture is reconstructed (Niem and Wingbermühle, 1997, Matsumoto et al., 1997, Schmitt and Yemez, 1999, Pollefeys et al., 2000, Mülayim and Atalay, 2001, Yılmaz et al., 2003, Mülayim et al., 2003).

Recent advances in computer vision in addition to photogrammetry make it possible to acquire high resolution 3D models of scenes and objects. One of the most popular applications has emerged as digitizing historical and cultural heritance such as *Digital Michelangelo* and *ACOHIR* (Levoy et al., 2000, ESPRIT, 1998). However, reconstruction of a complex rigid object from its 2D images is still a challenging computer vision problem under general imaging conditions. Without *a priori* information about the imaging environment (camera geometry, lighting conditions, object and background surface properties, etc.), it becomes very difficult to infer the 3D structure of the captured object. For practical purposes, the problem can be simplified by using controlled imaging environments. In such an environment, camera makes a controlled motion around the object, and background surface and lighting are selected to reduce the specularities on the acquired images (Niem and Wingbermühle, 1997, Matsumoto et al., 1997, Schmitt and Yemez, 1999, Mülayim et al., 1999, Lensch et al., 2000, Ramanathan et al., 2000, Yılmaz et al., 2003, Mülayim et al., 2003).

Figure 1: Image acquisition system consists of a turn table, a camera and a computer.

In this study, an inexpensive and robust 3D model acquisition system is described. The system is more suitable for acquiring the 3D models of small artifacts such as such as cups, trinkets, jugs and statues. Image acquisition system consists of a turn table, a digital camera and a computer as shown in Figure 1. The system works as follows. The object to be modeled is placed on a turn table. By rotating the table, images of the object are acquired. This image sequence is then calibrated. Object silhouettes are segmented from the background. Each silhouette is back projected to an open cone volume in the 3D space. By intersecting these open volumes, a coarse 3D model volume of the object is obtained. This coarse volume is further carved in order to get rid of excess volume on the concave parts. The surface appearance for model is also recovered from the acquired images. The model is considered as a surface composed of particles. Color of each particle is recovered from the images with an algorithm that computes the photo consistent color for each particle. The resultant appearance is stored in a texture map, while the shape is stored in a triangular mesh. Overall system diagram is illustrated in Figure 2.

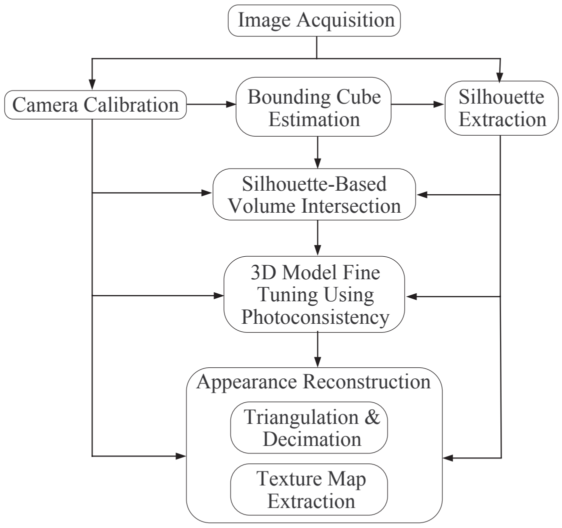


Figure 2: Overall system diagram.

The main advantage of proposed approach is its low computational complexity. Most of the methods in the literature require a large amount of time and labor for reconstruction. However, the complexities of our algorithms are reasonable. This enables building real time graphical user interfaces on top of these algorithms, by which real time interactive modifications can be done on the reconstructed models. This approach is very suitable for generating 3D models of small artifacts with high resolution geometry and surface appearance. Such artifacts may have handles, holes, concavities, cracks, etc. The proposed approach enables robustly modeling of these properties also. Furthermore, we are currently developing algorithms for 3D modeling from auto-calibrated images. The models obtained by this method are stored in Virtual Reality Modeling Language (VRML) format which enables the transmission and publishing of the models easier.

The organization of the paper is as follows: we first describe our camera calibration and geometry reconstruction processes 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, and the paper concludes with Section 5.

2 CAMERA CALIBRATION AND GEOMETRY RECONSTRUCTION

In order to compute the parameters of the camera, we use a multi-image calibration approach (Mülayim and Atalay, 2001). 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 (Mülayim et al., 1999). 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 (Mülayim et al., 2000).

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 (Mülayim and Atalay, 2001). Algorithm 1 which is mostly inspired from Matsumoto et. al. (Matsumoto et al., 1999) 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 3 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.

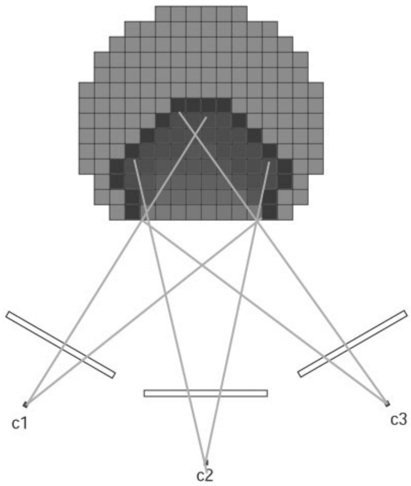
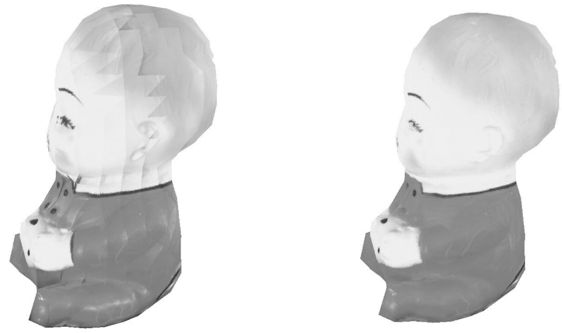


Figure 3: Darkest colored voxels get the highest voting.

3 APPEARANCE RECONSTRUCTION

There exist several studies for appearance reconstruction of 3D models from real images (Niem and Broszio, 1995, Genç and Atalay, 1999, Lensch et al., 2000, Neugebauer and Klein, 1999). In most of these studies, the model is represented as a triangular wireframe, 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 4, since adjacent triangles can be associated with different source images. Applying low pass filter on the boundaries cannot come up with a global solution.

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 (Schmitt and Yemez, 1999, Szelisky and Tonnesen, 1992). 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 as shown in Figure 5. Each particle's color is extracted from the images independently. 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) (b)

Figure 4: (a) Discontinuities on the triangle boundaries; (b) reconstruction using particles.

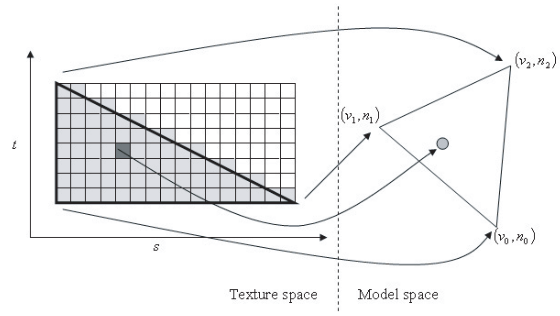


Figure 5: Association of particles with texture elements.

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. 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 4-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.

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 EXPERIMENTAL RESULTS

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 6, the algorithms are successful in removing the highlights which possibly occur at the image acquisition step. Reconstruction results given in Figure 7 show that objects which have holes and arbitrary shapes can also be modeled by this method. Concavities can succesfully be detected and carved in our method as shown in Figure 8 and Figure 9.

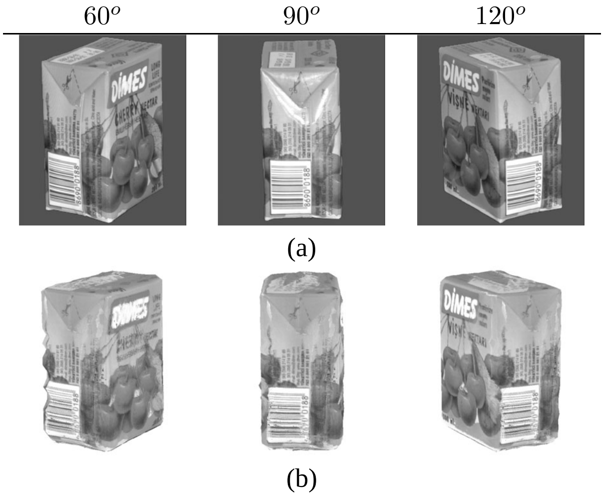


Figure 6: (a) Sample images, (b) sample reconstructions for the "box" object.

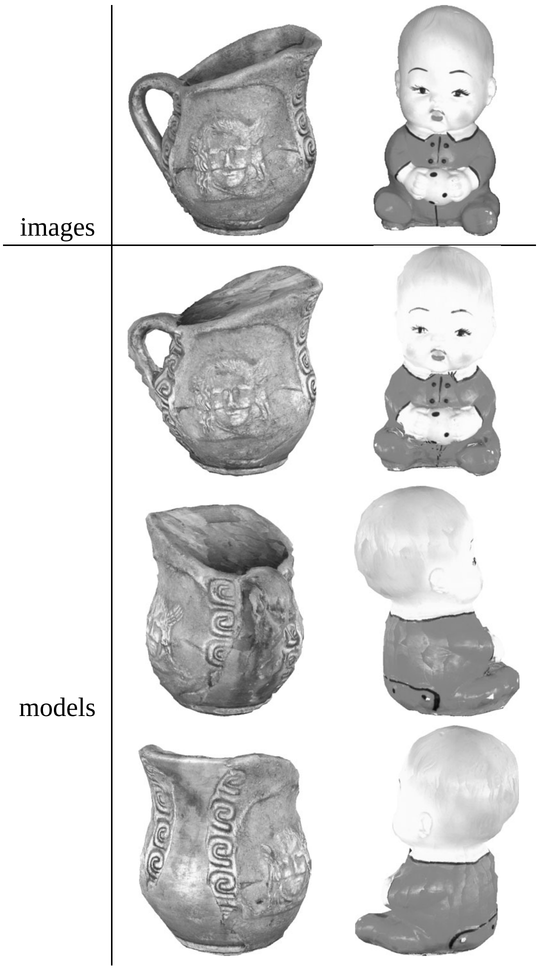


Figure 7: Sample reconstructions for a "cup"object and a small artifact, "toy" object.

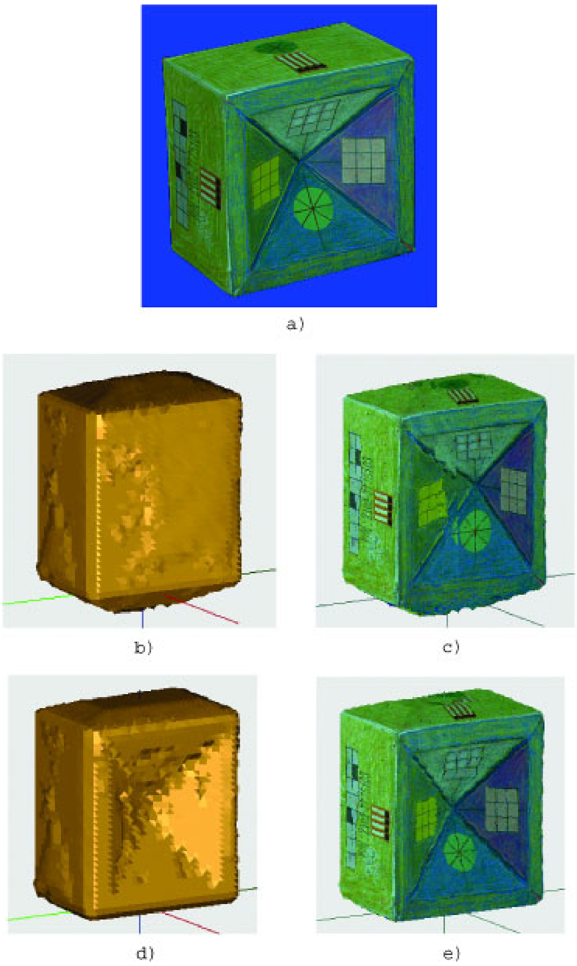


Figure 8: Sample reconstructions for a "concave box" object.

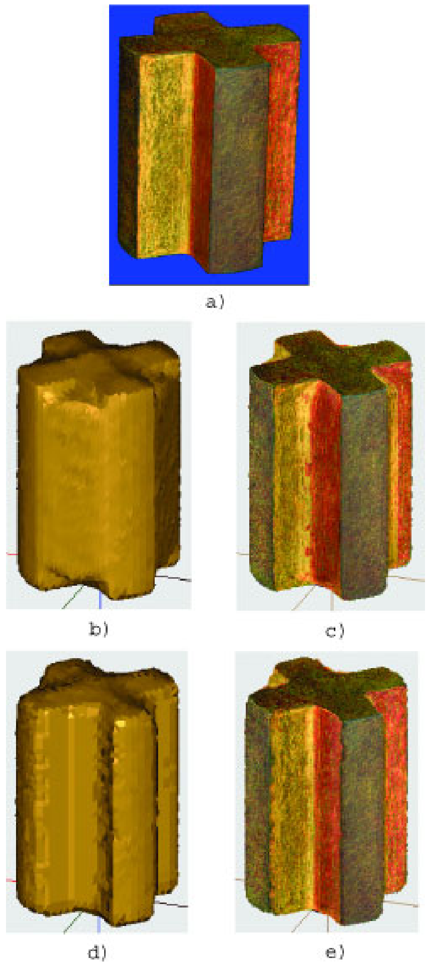


Figure 9: Sample reconstructions for a "concave block" object.

5 CONCLUSIONS

An inexpensive and robust 3D model acquisition system which is suitable for acquiring the 3D models of small artifacts such as such as cups, trinkets, jugs and statues, is described. Image acquisition system consists of a turn table, a digital camera and a computer. Objects are placed on the turn table, and their images are acquired by rotating the table. Having calibrated the image sequence, a coarse model is reconstructed by silhouette-based approach in which silhouettes of the objects in the images are back projected to an open cone volume in the 3D space and these open volumes are intersected to find the space of the object in space. This coarse volume is further carved in order to get rid of excess volume on the concave parts. The surface appearance for model is also recovered from the acquired images. To guarantee the continuity on the surface appearance, the concept of particles is successfully adopted in order to construct the proper texture map.

The algorithms in the framework of our system are fast and robust in removing the illumination artifacts due to image acquisition; reconstructing the objects with holes and arbitrary shapes; detecting and removing concavities. Thus, proposed system is very suitable for generating 3D models of small artifacts with high resolution geometry and surface appearance. Such artifacts may have handles, holes, concavities, cracks, etc. The models obtained by this method are stored in VRML format which enables the transmission and publishing of the models easier.

REFERENCES

ESPRIT, 1998. Acohir: Accurate colour high resolution recording and reproduction. http://www.mmrg.ecs.soton.ac.uk/projects/acohir.html.

Genç, S. and Atalay, V., 1999. Texture extraction from photographs and rendering with dynamic texture mapping. In: 10th International Conference on Image Analysis and Processing.

Lensch, H. P. A., Heidrich, W. and Seidel, H.-P., 2000. Automated texture registration and stitching for real world models. In: Pacific Graphics, pp. 317–337.

Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J. and Fulk, D., 2000. The digital michelangelo project: 3d scanning of large statues. In: SIGGRAPH.

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

Matsumoto, Y., Terasaki, H., Sugimoto, K. and Arakawa, T., 1997. A portable three-dimensional digitizer. In: International Conference on Recent Advances in 3D Digital Imaging and Modeling, pp. 197–204.

Mülayim, A. Y. and Atalay, V., 2001. Multibaseline stereo coreection for silhouette-based 3d model reconstruction from multiple images. In: SPIE, Three-Dimensional Image Capture and Applications IV, pp. 24–25.

Mülayim, A. Y., Yemez, Y., Schmitt, F. and Atalay, V., 1999. Rotation axis extraction of a turn table viewed by a fixed camera. In: Vision Modeling and Visualisation, pp. 35–42.

Mülayim, A. Y., Yılmaz, U. and Atalay, V., 2003. Silhouette-based 3d model reconstruction from multiple images. IEEE Transactions on Systems, Man and Cybernetics Part B: Cybernetics Special Issue on 3-D Image Analysis and Modeling. To appear.

Mülayim, A. Y., Özün, O., Atalay, V. and Schmitt, F., 2000. On the silhouette based 3d reconstruction and initial bounding cube estimation. In: Vision Modeling and Visualisation, pp. 11–18.

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

Niem, W. and Broszio, H., 1995. Mapping texture from multiple camera views onto 3d object models for computer animation. In: International Workshop on Stereoscopic and Three Dimensional Imaging, pp. 99–105.

Niem, W. and Wingbermühle, J., 1997. Automatic reconstruction of 3d objects using a mobile monoscopic camera. In: International Conference on Recent Advances in 3D Digital Imaging and Modeling, pp. 173–180.

Pollefeys, M., Koch, R., Vergauwen, M. and Gool, L. V., 2000. Automated reconstruction of 3d scenes from sequences of images. ISPRS Journal Of Photogrammetry And Remote Sensing 55(4), pp. 251–267.

Ramanathan, P., Steinbach, E. and Girod, B., 2000. Silhouette-based multiple-view camera calibration. In: Vision, Modeling and Visualization, pp. 3–10.

Schmitt, F. and Yemez, Y., 1999. 3d color object reconstruction from 2d image sequences. In: International Conference on Image Processing, pp. 65–69.

Soucy, M., Godin, G. and Rioux, M., 1996. A texture mapping approach for the compression of the colored 3d triangulations. The Visual Computer 12(10), pp. 503–514.

Szelisky, R. and Tonnesen, D., 1992. Surface modeling with oriented particles. SIGGRAPH Computer Graphics 26(2), pp. 185–194.

Yılmaz, U., Mülayim, A. Y. and Atalay, V., 2003. An image-based inexpensive 3d scanner. International Journal of Image and Graphics 3(2), pp. 235–263.