Volumetric Object Reconstruction using Generalized Voxel Coloring
Teresa C. S. Azevedo, João Manuel R. S. Tavares, Mário A. P. Vaz
International Workshop on Combinatorial Image Analysis 2008, Buffalo State College, Buffalo, NY, USA, April 7-9, 2008
Image Analysis – From Theory and Applications, ISBN: 978-981-08-0228-8, Research Publishing Services, pp. 135-144

Volumetric Object Reconstruction using Generalized Voxel Coloring

Teresa Azevedo[1], João Manuel R. S. Tavares[2] and Mário Vaz[2]

[1] INEGI – Instituto de Engenharia Mecânica e Gestão Industrial,
LOME – Laboratório de Óptica e Mecânica Experimental
FEUP – Faculdade de Engenharia da Universidade do Porto, Portugal
[2] FEUP/LOME/DEMEGI – Departamento de Engenharia Mecânica e Gestão Industrial
{teresa.azevedo, tavares, gmavaz}@fe.up.pt

Abstract. This paper presents a volumetric approach for three-dimensional (3D) object reconstruction. Building 3D models of objects from 2D images is still a very difficult task. However, the built 3D models can have different applications in many areas, such as in industrial inspection, virtual reality, and medicine, among others. In our work, an uncalibrated turntable image sequence of the object to be reconstructed is acquired. Then, 3D reconstruction is performed using *Generalized Voxel Coloring* (GVC), a volumetric method based on photo-consistency criteria. Two experimental objects, with different topologies, were reconstructed. The robustness of the GVC method was verified by varying some parameters, like the number of images used and the influence of the objects' silhouettes considered.

Keywords: 3D reconstruction, camera calibration, silhouettes, volumetric methods.

1 Introduction

Building three-dimensional (3D) models of objects has been a major research topic in Computer Vision, especially in the last decades. Mainly, the Computer Vision research community is concerned with the development of theories and methods for the automatic extraction of useful information from images. Thus, an effort has been made to build those 3D models directly from real-world scenes with high realism. The input images used in the reconstruction process can be acquired in many ways, such as from video recorders, single or multiple cameras or scanner devices. Since most Computer Vision algorithms require considerable computational resources, there is always a trade-off between hardware, software, price and speed.

1.1 Methods for 3D reconstruction

The methods comprised for 3D reconstruction of objects are usually classified into contact or non-contact methods, Fig. 1.

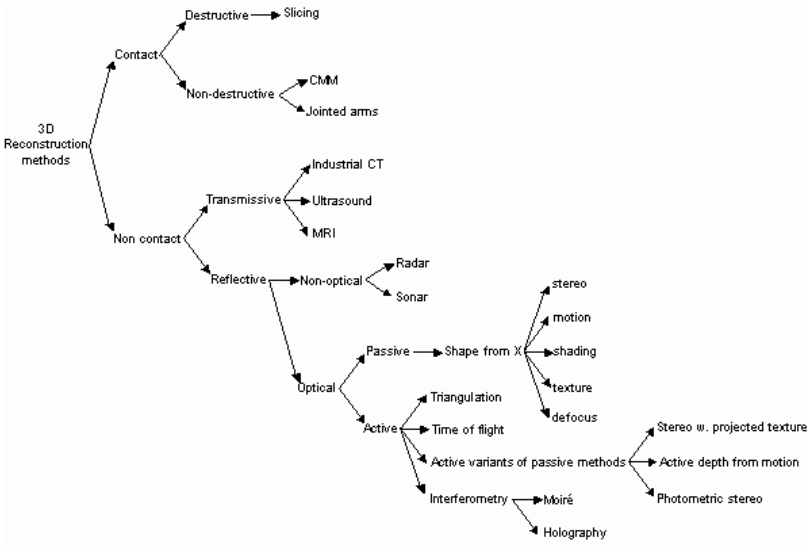

Fig. 1. Usual division of 3D reconstruction methods (adapted from [27]).

Contact-based methods can achieve high accuracy rates and are suitable for a wide range of applications. However, these methods involve mechanical movement from one measurement point to the next, across the entire object's surface, and thus they collect a set of sparse data points. This process can leave some critical surface areas unverified. In addition, the procedure of object scanning might modify or even damage the object, due to contact pression, [9].

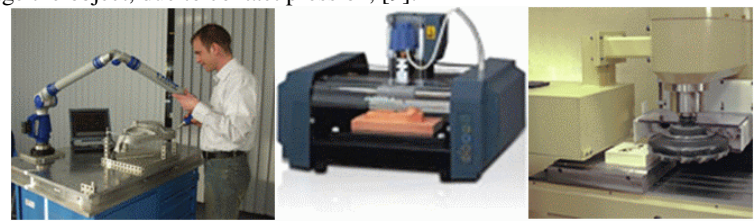

Fig. 2. Some industrial contact-based devices used for 3D objects reconstruction (obtained, from left to right, from [6], [18] and [4]).

Nowadays, the generation of 3D models is mainly accomplished using non-contact optical methods, Fig. 3. Most common non-contact methods use image data, range sensors or a combination of both.

Range sensors collect distance measurements from a known reference coordinate system to the surface points on the object to be reconstructed. They have been very well-liked when highly detailed models are required. However, they are spatially limited, expensive and some systems of this class do not provide color information about the reconstructed object. Range-based technologies that employ laser scanning and structured light are the most widely used. Main contribution factors for the success of these range-based methods are their precision, simplicity of use and the

wide variety of software packages available to process and edit the acquired data, in order to build the 3D models and to determine characteristic measures from the latter.

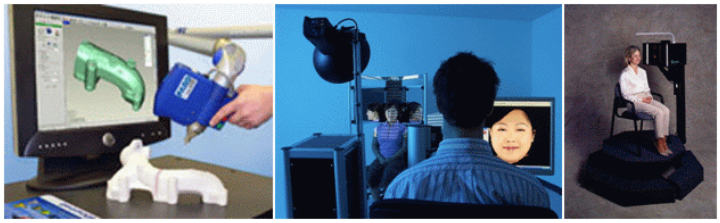

Fig. 3. Some industrial non-contact devices used for the 3D objects reconstruction (obtained, from left to right, from [7], [1] and [5]).

Image-based methods are widely used, in particular in industrial applications, architectural objects or for precise terrain and city modeling. The required images can be acquired from video sequences or photo-cameras, which are acquisition systems currently very accessible and of low cost. However, these methods require that relevant features are available in the input images, which sometimes is not possible due to occlusions or lack of texture.

1.2 Volumetric methods

Recently, volumetric methods were developed, which are image-based 3D reconstruction approaches. Since volumetric (or *voxel*-based) methods work on the 3D objects' space and do not require a matching process between the images used in the reconstruction process, they are most suitable for objects with smooth surfaces and also for objects which are capable to suffer from occlusion problems, [22].

The first volumetric methods proposed were called *shape-from-silhouettes* or *shape-from-contours* (e.g. [16], [8]), and were based on the *visual-hull* concept, [13]. They combine silhouette images with calibration information of the used cameras, to build a set of visual rays in the scene space for all silhouette points. These rays define visual cones where the object is guaranteed to be. The intersection of these visual cones is commonly referred to as *visual hull*, Fig. 4. The major drawback of silhouette-based methods is that they cannot deal with concavities on the object's surface.

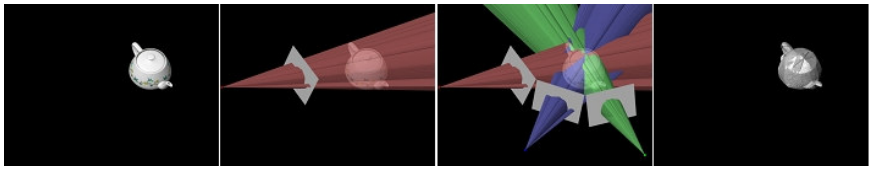

Fig. 4. Left to right: from the original object to its visual hull ([28]).

Recently, volumetric methods use *color consistency* to verify if a certain voxel belongs to the object's surface, Fig. 5. Thus, they are not only able to build objects with complex geometry, but also generate color models without the need of an extra coloration step. *Voxel Coloring* ([20]) was the first method to use a color consistency

measure. Other well known methods are: *Space Carving* ([11]), *Generalized Voxel Coloring* ([23]) and *Roxels* ([3]). The accuracy of the 3D models, built using volumetric methods, depends on the number of images used, the positions of each associated viewpoint, the accuracy of the camera's calibration procedure and the object's shape complexity.

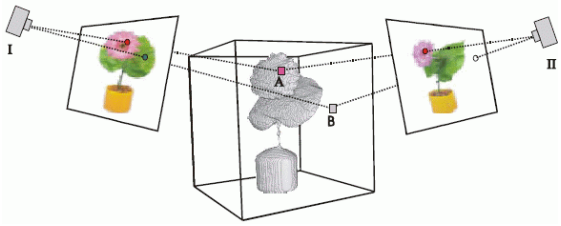

Fig. 5. Color-consistency check ([12]): both projections of voxel A have the same color, so it is consistent; projections of voxel B have different colors, thus, it should not belong to the object's surface model.

2 Used Methodology

The next subsections will give some insights of each step considered in our methodology, Fig. 6.

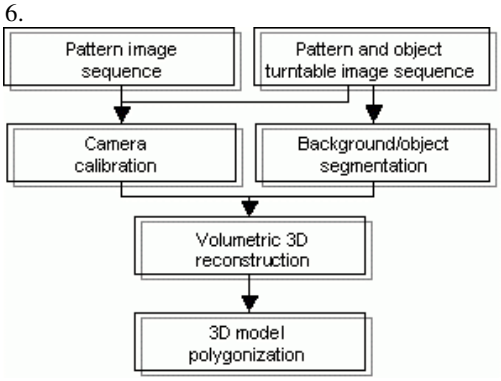

Fig. 6. Methodology used in this work to do 3D reconstruction of objects from images.

2.1 Image Acquisition

In our methodology, it is necessary to acquire two image sequences:

- the first is acquired by moving a planar chessboard calibration pattern freely in front of a static camera;
- the second image sequence is acquired with the same chessboard pattern and the object to be reconstructed on top of it, both placed on a simple turntable device; then, keeping the camera untouched, the turntable device is spinned until a full rotation is performed.

2.2 Camera Calibration

Camera calibration is the process of finding the transformation that maps the 3D world in the associated 2D image space. The calibration procedure implemented is based on *Zhang*'s camera calibration algorithm, [29]. Thus, intrinsic parameters (focal length and principal point) and distortion parameters (radial and tangential) are obtained from the first sequence. Then, using the second sequence, the extrinsic parameters (rotation and translation) are determined, for each image to be used in the reconstruction process.

2.3 Image Segmentation

Even when the scene background has low color variation, the photo-consistency criteria may not be sufficient for accurate 3D reconstructions, [19]. Moreover, since the used calibration pattern rotates along with the object to be reconstructed, it will not be considered has background and, consequently, will be reconstructed as if it was part of the object. To prevent this from happening, we need to perform segmentation between object and background on the second image sequence. This is done by first removing the red and green channels from the original RGB images and, finally, using image binarization through threshold value. The silhouette information obtained will allow all voxels, which do not reproject inside all silhouette images, to be removed from the reconstruction volume.

2.4 Volumetric Reconstruction

Using the second image sequence, the associated silhouette images and camera's parameters, determined by the calibration procedure, the object's model can be built using the GVC method, [14]. Briefly, this reconstruction process is initialized with a 3D bounding box of voxels containing in it the object to be reconstructed. For each voxel of that box, a pixel collection is built, containing all pixels that the referred voxel projects onto. The 3D shape of the desired object is then constructed by removing (*carving*) voxels that are not photo-consistent with its pixel collection color values. Moreover, a voxel is carved whenever it projects outside one of the silhouette images from which it is visible.

2.5 Model Smoothing

Finally, the volumetric model obtained is polygonized and smoothed using the *Marching Cubes* algorithm, [15]. Basically, this algorithm extracts a polygonal surface from the input volumetric data. Thus, for each voxel of the built 3D model, it determines the polygon(s) needed to represent the part of the isosurface that passes through the referred voxel. The individual polygons are then fused into the desired surface.

3 Experimental Results

Our methodology was tested using two real objects, with different topologies: a parallelepiped and a human hand model. As described in previous section 2.1, using a single off-the-shelf CCD camera, two image sequences were acquired for each object, Fig. 7 and Fig. 8.

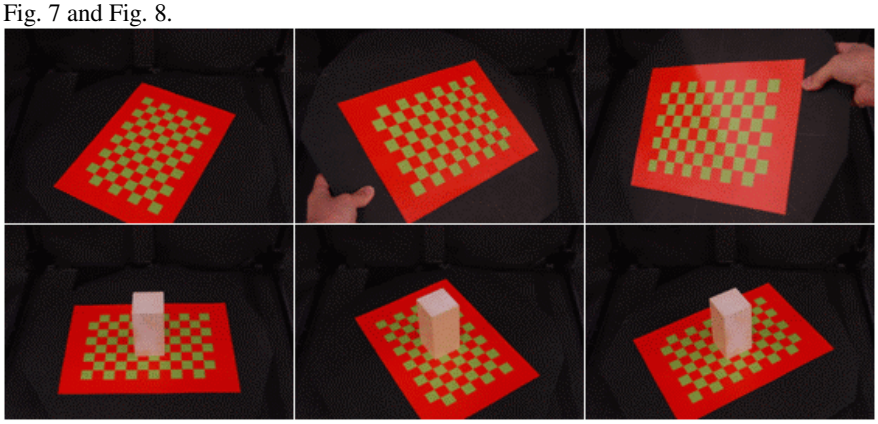

Fig. 7. Above: three images of the sequence used to obtain the intrinsic camera parameters. Below: three images of the sequence used to obtain the extrinsic camera parameters and to reconstruct the parallelepiped object.

Fig. 8: Three images of the sequence used to obtain the extrinsic camera parameters and to reconstruct the human hand model object.

For both objects being considered, the accuracy of the camera calibration procedure was checked in two ways: first, by building a 3D graphical representation of the camera's extrinsic parameters, Fig. 9; secondly, by measuring the reprojection error of the calibration pattern points into all images of the second image sequence, Table 1.

The effectiveness of our segmentation method enabled us to obtain good silhouette images for both objects, see Fig. 10.

With the second image sequences and respective silhouette images, combined with the camera's calibration parameters, the models for both objects were successfully built, polygonized and smoothed, as shown in Fig. 11 and Fig. 12.

To test the robustness of the GVC method, we changed some of its parameters, like the number of images used and the quality of the silhouettes considered when building the 3D model. In Fig. 13 can be seen a viewpoint of the reconstructed model for both objects considered in this work, obtained without using silhouette images. In

this figure, can be noticed the simultaneous reconstruction of the object and the used calibration pattern. Also, it becomes clear the erroneous reconstruction of the parallelepiped, were its surface is not perfectly Lambertian, that is, were the pattern is reflected into the surface of the parallelepiped.

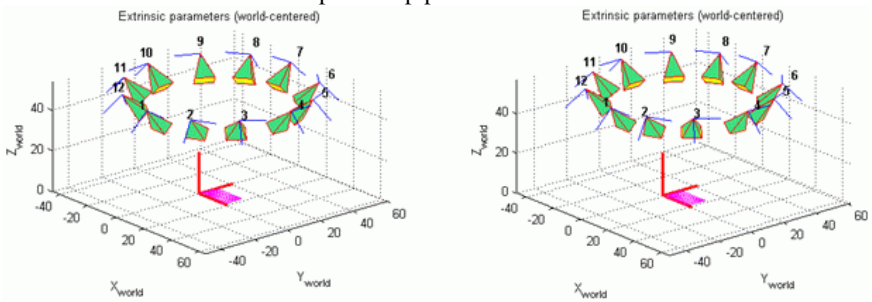

Fig. 9. Graphical representation of the camera's extrinsic parameters, for the parallelepiped (left) and the human model (right) objects.

Table 1. Reprojection error of the pattern points into all images of the second image sequence.

Object	Reprojection error (in pixels)			
	average		standard deviation	
	x	*y*	*x*	*y*
Parallelepiped	-1.24e-04	-2.67e-05	0.545	0.594
Hand model	-7.31e-05	-2.61e-05	0.673	0.840

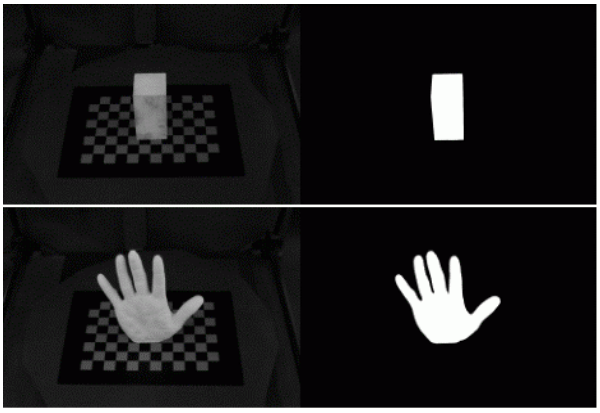

Fig. 10. A segmented image of the parallelepiped (above) and the hand model (below): on the left, the blue channel from the original image; on the right, silhouette obtained from binarization using a threshold value.

The second test was to reduce the number of images used in the reconstruction process. Initially, 12 images were acquired for the second sequence of both objects. Keeping only one third of the original images (the 3rd, 6th, 9th and 12th image), both objects were reconstructed. With the parallelepiped object, no major differences were observed, when comparing to the previously reconstructed model, left on Fig. 14.

However, with the hand model object, the obtained 3D model was a lot more imprecise, right on Fig. 14.

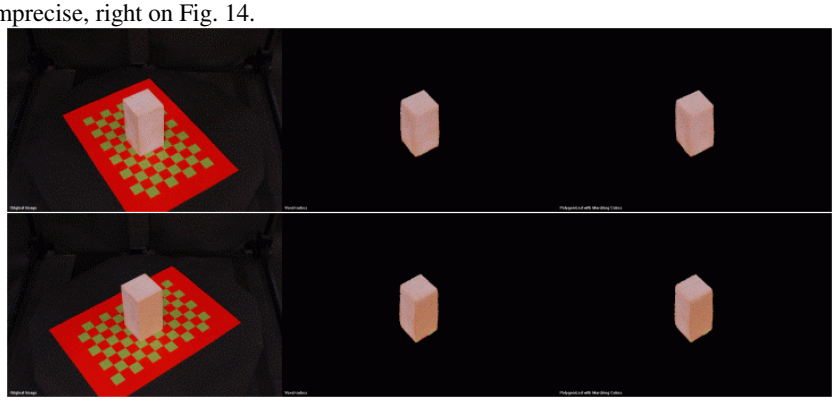

Fig. 11. Two different viewpoints of the 3D model obtained for the parallelepiped object: left, original image; middle, voxelized 3D model; right, polygonized and smoothed 3D model.

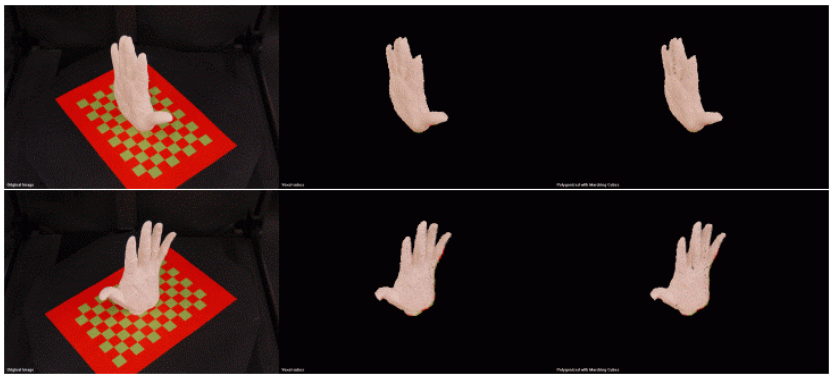

Fig. 12. Two different viewpoints of the 3D model obtained for the hand model object: left, original image; middle, voxelized 3D model; right, polygonized and smoothed 3D model.

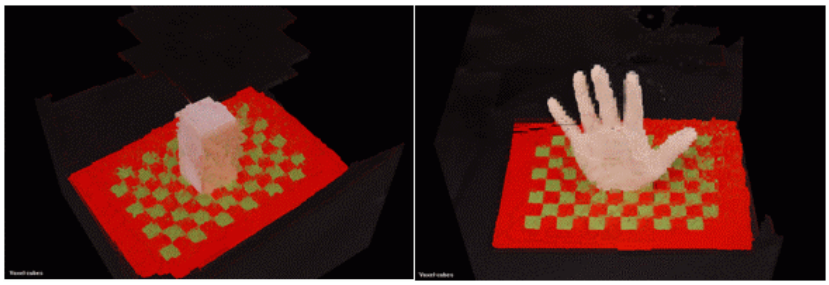

Fig. 13. One viewpoint of the voxelized 3D models obtained without using silhouettes: left, parallelepiped case; right, hand model case.

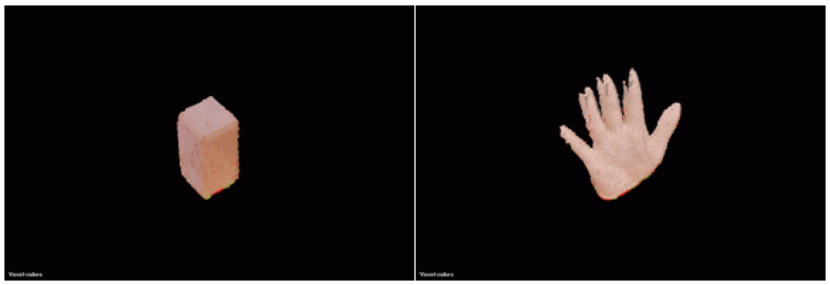

Fig. 14. One viewpoint of the voxelized 3D models obtained with only four input images: left, parallelepiped case; right, hand model case.

Our last experimental test was to modify the silhouette images, using the dilation and erosion morphological operators. With these altered silhouettes, we could analyze the influence of the image segmentation accuracy on the reconstruction results. On the dilation case, the reconstructed 3D model includes some of the calibration pattern, right on Fig. 15. Several characteristic measures are compared between the real parallelepiped object, the model built using the original silhouettes and both models built using the eroded and dilated silhouettes, Fig. 16. All built models have bigger height than the real parallelepiped object, Fig. 17, because the photo-consistent visual hull tends to be larger than the real objects, [21], Fig. 18.

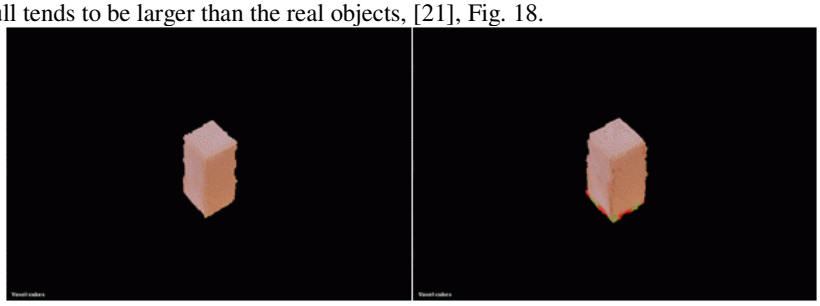

Fig. 15. The same viewpoint of two parallelepiped 3D models obtained with the eroded silhouettes (left) and the dilated silhouettes (right).

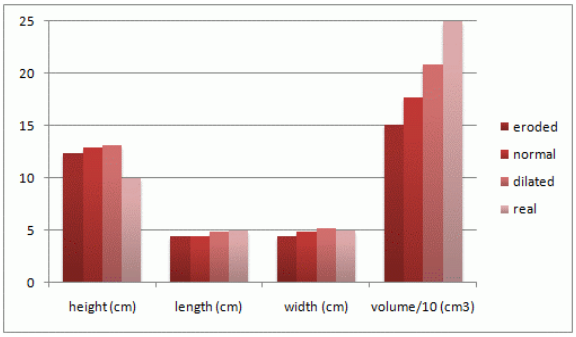

Fig. 16. Characteristic measures obtained from the 3D models (built using the original, eroded and dilated silhouettes) and the real parallelepiped object.

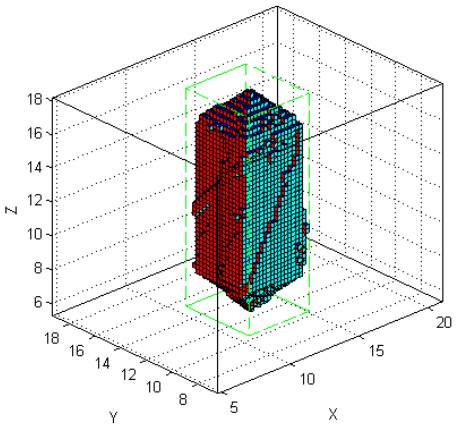

Fig. 17. Graphical representation of the voxelized 3D model of the parallelepiped object, built with the original silhouette images.

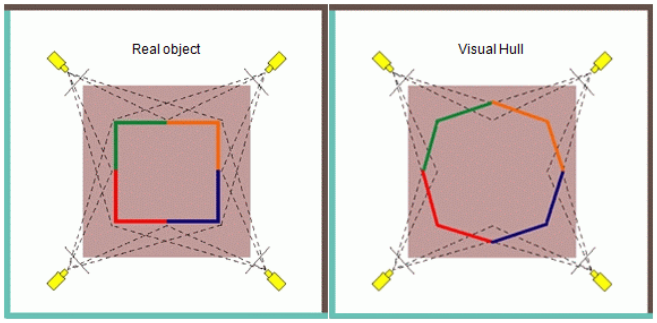

Fig. 18. Difference between the real object (left) and the visual hull defined by four viewpoints (image adapted from **[21]**).

4 Conclusions and Future Work

The final goal of this work was to test an image-based volumetric reconstruction method, denominated *Generalized Voxel Coloring*. Starting with two uncalibrated image sequences, the used CCD camera is calibrated, the obtained images are segmented, in order to obtain the object's silhouette information, and a 3D volumetric model is built. From this model, a polygonized and smoothed surface can be extracted as well as some geometrical measures.

The models built using the GVC method were quite similar and closer to the real objects; this applies to shape and color. However, the quality of the obtained 3D reconstructions is highly dependent on the accuracy of the camera calibration procedure and on the quality of the objects segmentation in the used images. These dependencies limit the use of the GVC methodology in controlled environments,

because of the low texturized background restriction. Moreover, the complexity of the objects shape and the reflectance of their surface are aspects that must be considered for more accurate 3D reconstructions. In resume, we can conclude that in controlled environments the GVC methodology is capable to obtain adequate 3D static reconstructions of objects from images. In addition, its major contribution may be the fact that it is fully automatic and suitable for many real applications.

Our future work will pass through the implementation of auto-calibration methods (e.g. [25], [10], [2]) and through the 3D reconstruction of non-rigid objects (e.g. [26]). It was also observed that the GVC method is very computationally demanding and slow, which can be unsuitable in many applications. Thus, other future tasks may concern the implementation of a coarse-to-fine approach (e.g. using *octrees*), like the one proposed in [24] or [17], or the development of parallel implementations.

5 Acknowledgements

This work was partially done in the scope of project "Segmentation, Tracking and Motion Analysis of Deformable (2D/3D) Objects using Physical Principles", with reference POSC/EEA-SRI/55386/2004, financially supported by *FCT – Fundação para a Ciência e a Tecnologia* from Portugal.

References

1. 3D-Shape, FaceSCAN3D, 3D-Shape GmbH, www.3d-shape.com/produkte/face_e.php (retrieved 12th June 2007).
2. Agapito, L. d., Hayman, E. and Reid, I.: Self-calibration of a rotating camera with varying intrinsic parameters, British Machine Vision Conference, Southampton, UK, pp. 105-114 (1998).
3. Bonet, J. S. D. and Viola, P.: Roxels: Responsibility weighted 3D volume reconstruction, IEEE 7th International Conference on Computer Vision, Kerkyra, Corfu, Greece, 1, pp. 415-425 (1999).
4. CGI, CGI-CSS, Cross-Sectional Scanning, CGI® - Capture Geometry Internally, www.cgiinspection.com/css_process.cfm (retrieved 11th June 2007).
5. Cyberware, Head & Face Color 3D Scanner, Cyberware, Inc., www.cyberware.com/products/scanners/ps.html (retrieved 12th June 2007).
6. Faro, FaroArm, FARO Swiss Holding GmbH, http://www.faro.com/content.aspx?ct=us&content=pro&item=2 (retrieved 11th June 2007).
7. Faro, LaserScanArm, FARO Swiss Holding GmbH, www.faro.com/content.aspx?ct=us&content=pro&item=1 (retrieved 12th June 2007).
8. Fromherz, T. and Bichsel, M.: Shape from multiple cues: Integrating local brightness information, 4th International Conference for Young Computer Scientist, Beijing, China, pp. 855-862 (1995).
9. Gershon, R. and Benady, M.: Noncontact 3-D measurement technology enters a new era, (2001), http://www.qualitydigest.com/sept01/html/3d.html (retrieved 11th June 2007).
10. Hartley, R. I.: Self-Calibration from Multiple Views with a Rotating Camera, European Conference on Computer Vision, Stockholm, Sweden, 1, pp. 471-478 (1994).

11. Kutulatos, K. N. and Steiz, S. M.: A Theory of Shape by Space Carving, Technical Report TR692, Computer Science Department, University of Rochester, New York, USA, (1998).
12. Kuzu, Y. and Sinram, O.: Photorealistic Object Reconstruction Using Voxel Coloring and Adjusted Image Orientations, American Society for Photogrammetry and Remote Sensing/American Congress on Surveying and Mapping /International Federation of Surveyors Conference, Washington DC., USA, (2002).
13. Laurentini, A.: The visual hull concept for silhouette-based image understanding, IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(2), pp. 150-162 (1994).
14. Loper, M.: Archimedes: Shape Reconstruction from Pictures - A Generalized Voxel Coloring Implementation, (2002), http://matt.loper.org/Archimedes/ (retrieved August 2006).
15. Lorensen, W. E. and Cline, H. E.: Marching cubes: A high resolution 3D surface construction algorithm, International Conference on Computer Graphics and Interactive Techniques, ACM Press, New York, USA, 21(4), pp. 163-169 (1987).
16. Massone, L., Morasso, P. and Zaccaria, R.: Shape from occluding contours, SPIE Conference on Intelligent Robots and Computer Vision, Cambridge, MA, USA, 521, pp. 114-120 (1985).
17. Prock, A. C. and Dyer, C. R.: Towards real-time voxel coloring, DARPA - Image Understanding Workshop, Monterey, CA, USA, pp. 315-321 (1998).
18. Roland, MDX-15/20 Desktop 3D Scanning and Milling, Roland® Advanced Solutions Division, www.rolanddga.com/asd/resources/pdf/brochure_MDX20.pdf (retrieved 11th June 2007).
19. Sande, K.: A Practical Setup for Voxel Coloring using off-the-shelf Components, Bachelor Project, Universiteit van Amsterdam, Netherlands, (2004).
20. Seitz, S. and Dyer, C. R.: Photorealistic Scene Reconstruction by Voxel Coloring, Computer Vision and Pattern Recognition Conference, San Juan, Puerto Rico, pp. 1067-1073 (1997).
21. Slabaugh, G., Culbertson, W. B., Malzbender, T., et al.: Improved voxel coloring via volumetric optimization, Technical Report 3, Center for Signal and Image Processing, Georgia Institute of Technology, USA, (2000).
22. Slabaugh, G., Culbertson, W. B., Malzbender, T., et al.: A survey of methods for volumetric scene reconstruction from photographs, International Workshop on Volume Graphics, New York, USA, pp. 21-22 (2001).
23. Slabaugh, G. G., Culbertson, W. B. and Malzbender, T.: Generalized Voxel Coloring, Workshop on Vision Algorithms, Corfu, Greece, pp. 100-115 (1999).
24. Szeliski, R.: Rapid octree construction from image sequences, Computer Vision, Graphics and Image Processing: Image Understanding, 58(1), pp. 23-32 (1993).
25. Triggs, B.: Autocalibration and the Absolute Quadric, IEEE Conference on Computer Vision and Pattern Recognition, Puerto Rico, USA, pp. 609-614 (1997).
26. Vedula, S., Baker, S., Seitz, S., et al.: Shape and motion carving in 6D, IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, SC, USA, 2, pp. 592-598 (2000).
27. Velho, L., Carvalho, P. C., Soares, E., et al.: Fotografia 3D, IMPA - 25º Colóquio Brasileiro de Matemática, Rio de Janeiro, Brazil, (2005).
28. Yang, L.: 3D Surface Reconstruction from 2D Images, Center for Visual Computing, Instructional Computing, Stony Brook University, EUA, (2003).
29. Zhang, Z.: A Flexible New Technique for Camera Calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), pp. 1330-1334 (2000).