3D VOLUMETRIC RECONSTRUCTION AND CHARACTERIZATION OF OBJECTS FROM UNCALIBRATED IMAGES

Teresa C. S. Azevedo*, João Manuel R. S. Tavares** and Mário A. P. Vaz**

* INEGI – Instituto de Engenharia Mecânica e Gestão Industrial,
LOME – Laboratório de Óptica e Mecânica Experimental
FEUP – Faculdade de Engenharia da Universidade do Porto, PORTUGAL
e-mail: teresa.azevedo@fe.up.pt

** INEGI, LOME
FEUP, DEMEGI – Departamento de Engenharia Mecânica e Gestão Industrial
e-mail: {tavares, gmavaz}@fe.up.pt

ABSTRACT
Three-dimensional (3D) object reconstruction using only bi-dimensional (2D) images has been a major research topic in Computer Vision. However, it is still a complex problem to solve, when automation, speed and precision are required. In the work presented in this paper, we developed a computational platform with the main purpose of building 3D geometric models from uncalibrated images of objects. Simplicity and automation were our major guidelines to choose volumetric reconstruction methods, such as *Generalized Voxel Coloring*. This method uses photo-consistency measures to build an accurate 3D geometric model, without imposing any kind of restrictions on the relative motion between the camera used and the object to be reconstructed. Our final goal is to use our computational platform in building and characterize human external anatomical shapes using a single off-the-shelf camera.

KEY WORDS
Computer Vision, 3D Reconstruction, Volumetric Methods, Calibration

1. Introduction

Computer Vision is continuously developing new and more robust computational methodologies to automatically extract useful information from images. For example, the 3D reconstruction of objects from images is one research fields of Computer Vision in which the scientific community has been made considerable efforts, especially in the last decades. Applications of 3D reconstruction methodologies range several distinct domains, like industry, medicine, virtual reality, security, among others.
Usually, contactless methods used to recover the 3D geometry of an object are divided in two classes:

- active: they require the projection of some kind of energy or relative motion between the camera used and the object to be reconstructed in order to obtain information about its shape;
- passive: they do not consider the projection of any type of energy neither relative motion.

Both methods have advantages and disadvantages, so the decision on which to use depends on the application case, on the properties of the object to reconstruct and/or on the precision and accuracy required.

1.1 Human body 3D reconstruction

3D geometric reconstructions of the human body was initially developed for ergonomic purposes in the automotive and aeronautics industry. The used models were consisted by articulated skeletons, and thus the body was just represented by simple geometric primitives, such as cylinders or parallelepipeds, ([1]).
Currently, 3D models of the human body are much more realistic and are of great interest in a variety of application fields, such as the cinematographic industry, virtual reality, clothing industry and biomedical applications. Usually, high quality 3D models of a static human body are obtained using commercial *scanners* ([2]), generally expensive but easy to handle.
Other usual 3D human reconstruction techniques use single- or multi-cameras photogrammetry (e.g. [3]) or structure-from-motion methodologies (e.g. [4]).

1.2 Volumetric methods

Recently, volumetric methods were developed, which are a low-cost and versatile alternative to the above mention 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 well suitable for objects with smooth surfaces or capable to suffer from occlusion problems ([5]).
The first volumetric methods proposed were called *shape-from-silhouettes* or *shape-from-contours* (e.g. [6], [7]), and were based on the *visual-hull* concept, ([8]). They combine object's silhouette images with camera's calibration information, to build a set of visual rays in the scene space for all silhouette points. These visual rays define generalized cones within which the object to reconstruct is guaranteed to be. The intersection of these cones is commonly referred as *visual hull*, Figure 1. The major drawback of these silhouette-based methods is that they cannot deal with concavities on the object.
Recently, volumetric methods use *color consistency* to verify if a certain voxel belongs to the object's surface, Figure 2. Thus, they are not only able to build objects with complex geometry, as they also generate colour models without the need of an extra step for coloration.
The first method to use a color consistency measure was the *Voxel Coloring*, ([9]). Others well known methods are the *Space Carving*, ([10]), *Generalized Voxel Coloring* ([11]), *Roxels* ([12]), etc.
The accuracy of the reconstructions built using volumetric methods, depends on the number of images used, positions of each viewpoint associated with them, accuracy of the camera's calibration procedure used and on the complexity of the object's shape.

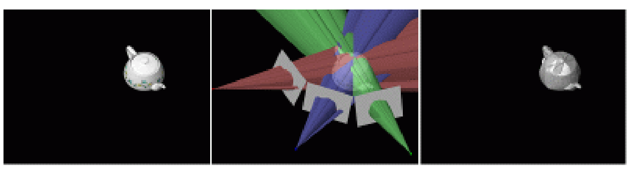

Figure 1. Left to right: from the original object to its *visual hull* ([13]).

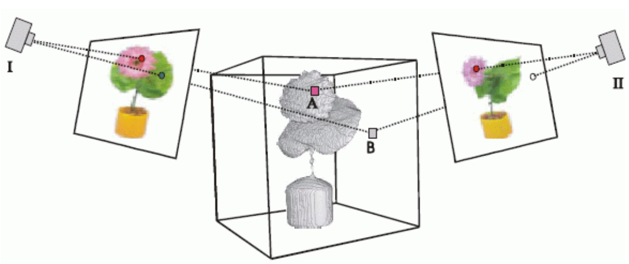

Figure 2. Color consistency ([14]): both projections of voxel A have the same color, so it is consistent; projections of voxel B have different colors, so it is inconsistent and, consequently, should not belong to the object's surface.

2. Used methodology

Our computational platform, named *3D Builder*, integrates several useful Computer Vision techniques for obtain the 3D reconstruction of objects using the methodology indicated in Figure 3. The next subsections will give some insights for each step of considered in our methodology.

2.1 Image acquisition

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

- a first one, acquired moving freely in 3D space a planar chessboard calibration pattern;
- for the second sequence, the object to reconstruct is placed on a simple turntable device, with the same chessboard pattern beneath it, Figure 4; keeping the camera untouched, the second sequence of images is acquired, spinning the turntable device until a full rotation is performed.

In our methodology, is not made any restriction on the number of images acquired, neither the rotation angle, between two consecutive images of the second image sequence, need to be known or even constant.

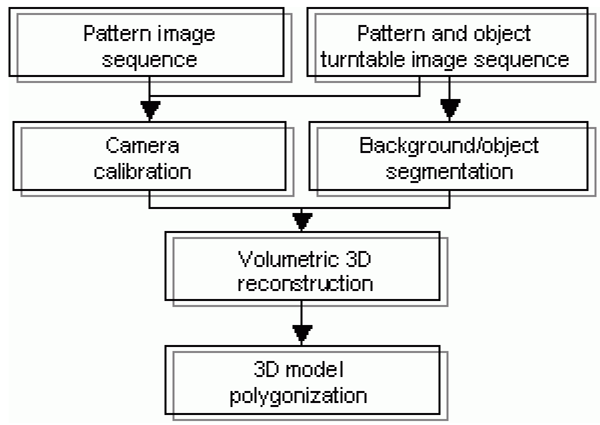

Figure 3. Methodology followed to obtain the 3D reconstruction of objects from uncalibrated images.

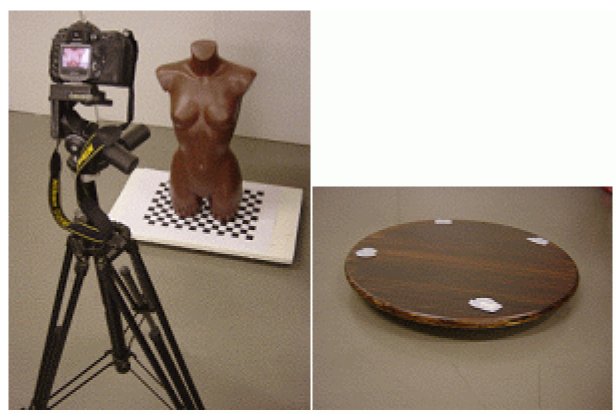

Figure 4. Left: camera and object layout. Right: simple turntable device used to reconstruct objects from images.

2.2 Camera calibration

Camera's calibration means to find the transformation that maps the 3D world in the associated 2D image space. In *3D Builder*, the calibration procedure integrated is based on the *Zhang*'s algorithm ([15]). Thus, intrinsic and extrinsic camera's parameters are determined using the two acquired image sequences: intrinsic parameters (focal length and principal point) and distortion parameters (radial and tangential) are obtained from the first image sequence; using the second image sequence, the extrinsic parameters (rotation and translation) associated with each image to use in the reconstruction process are determined.

2.3 Background/object segmentation

Even when the scene background has low color variation, the photo-consistency criteria may not be sufficient for correct 3D reconstructions, ([16]). Also, since the used calibration pattern will rotate 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 be reconstructed. Thus, to prevent this to happen, on the second image sequence we need to perform an object's segmentation. For that, we use some basic image processing tools, like binarization by threshold value and morphologic operator. The obtained silhouette information will allow that can be removed all voxels that do not reproject inside all silhouette images.

2.4 Volumetric 3D reconstruction

Using the second image sequence and associated silhouette images, combined with the parameters of the camera used determined by the calibration procedure, the object's model can be build using the *Generalized Voxel Coloring – GVC* method ([17]). 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 polygonization and smoothing

Finally, the volumetric model obtained is polygonized and smoothed using the *Marching Cubes* algorithm ([18]). Basically, this algorithm extracts a polygonal surface from the volumetrical data. Thus, it proceeds through the voxelized model, and, for each voxel, 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 on two real objects: a "hand" and a "human" torso.

Has described on subsection 2.1, using a simple off-the-shelf CCD camera, two image sequences were acquired for each object, Figures 5 and 6.

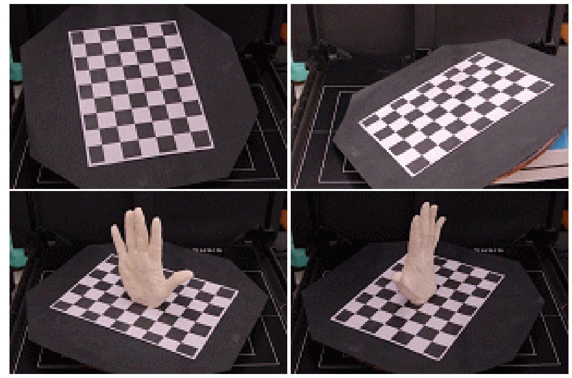

Figure 5. Two images of the sequence used for the intrinsic parameters calibration (above) and two images of the sequence used for the extrinsic calibration (below) for the hand case.

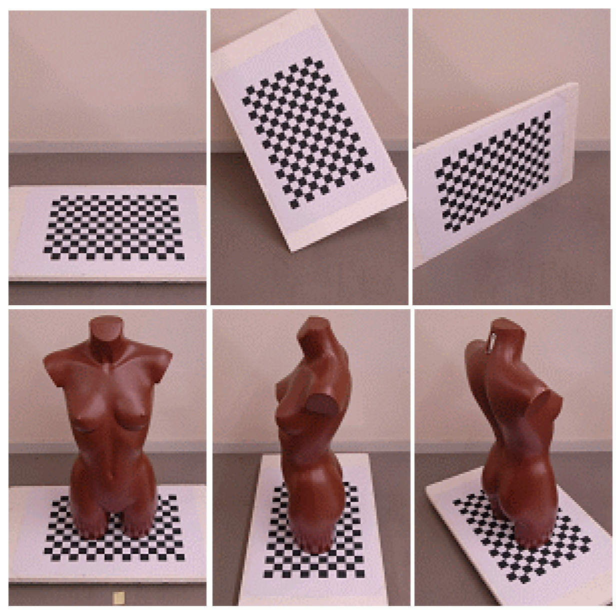

Figure 6. Three images of the sequence used for the intrinsic parameters calibration (above) and three images of the sequence used for the extrinsic calibration parameters (below) for the torso case.

3.1 Calibration results

For both objects considered, the results of the extrinsic calibration procedure are displayed in Figures 7 and 8. Thus, the two 3D graphics used represent the camera's positions of all images used, considering the world coordinate system fixed on the chessboard pattern and the camera rotating around the object.
Another way to verify the accuracy of the calibration results obtained is to reproject the 3D points from the chessboard pattern in all images of the second sequence considered. The standard deviations of the reprojection errors (in pixels) for the hand and torso models cases are represented in Figures 9 and 10, respectively. In these figures, the influence of the number of pattern points used

on the accuracy of the calibration results can be easily verified.

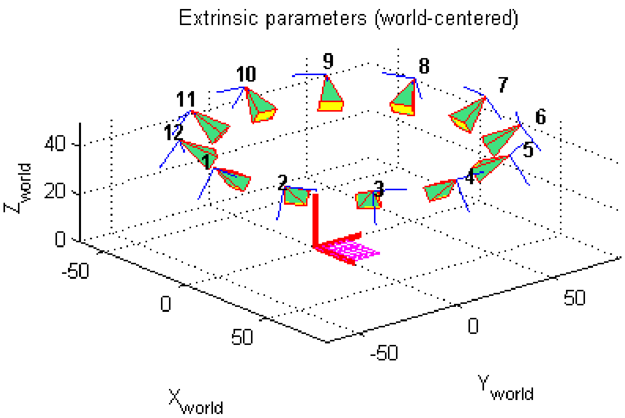

Figure 7. Extrinsic parameters representation for the image sequence of the hand case.

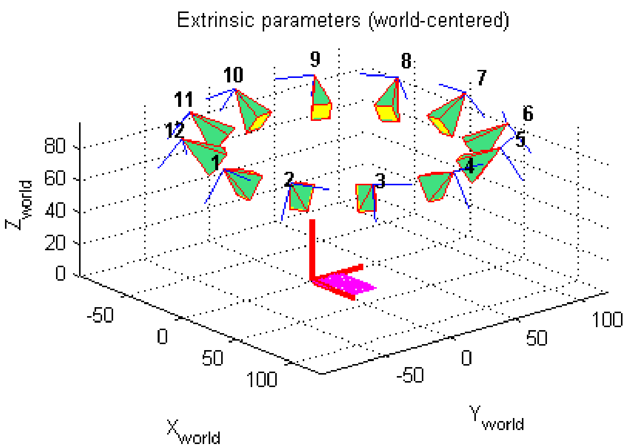

Figure 7. Extrinsic parameters representation for the image sequence of the torso case.

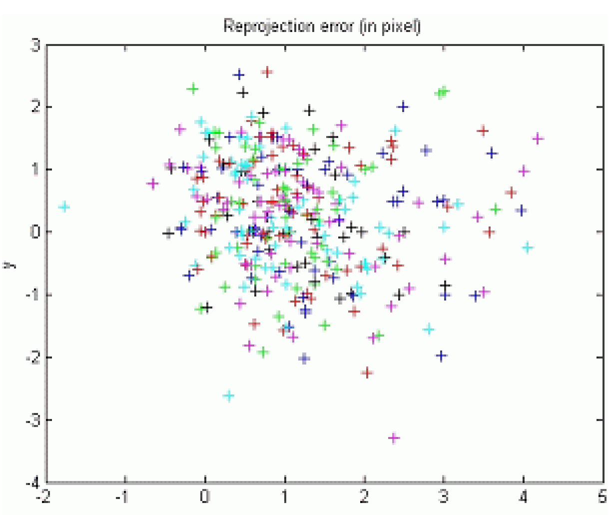

Figure 9. Reprojection errors for the image sequence of the hand case.

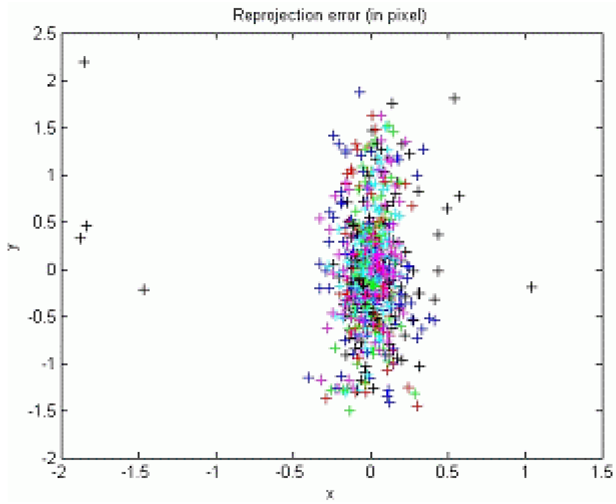

Figure 10. Reprojection errors for the image sequence of the torso case.

3.2 Segmentation results

As previously referred, using some usual algorithms of image processing, like *Sobel*'s operator for edge detection and other morphology functions, like flood-fill to fill separated regions, it was possible to obtain reasonably good silhouettes for both objects, Figure 11 and 12.

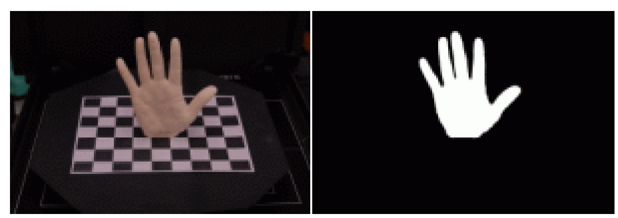

Figure 11. An example of the segmentation on one image of the hand sequence case: left, original image; right, segmentation result.

Figure 12. An example of the segmentation on one image of the torso sequence case: left, original image; right, segmentation result.

3.3 3D Reconstruction results

With the second image sequences and respective silhouette images, combined with the camera's calibration parameters, both objects' models were successfully built, polygonized and smoothed, as it can be seen in Figures 13 and 14. In the reconstructions done, an initial cube with 200 voxels per side was used.

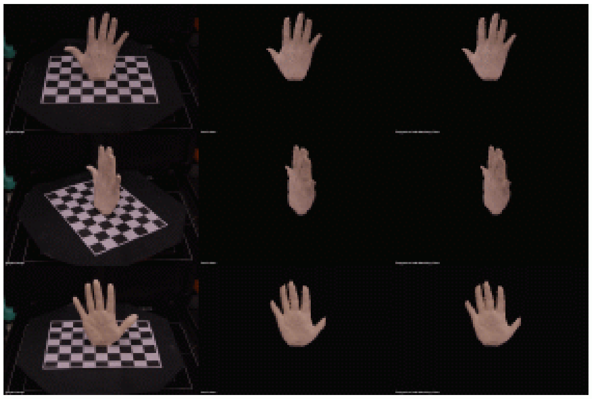

Figure 13. Three viewpoints (by row) of the 3D reconstruction obtained for the hand case. Left: original images. Middle: voxelized 3D model. Right: polygonized and smoothed 3D model.

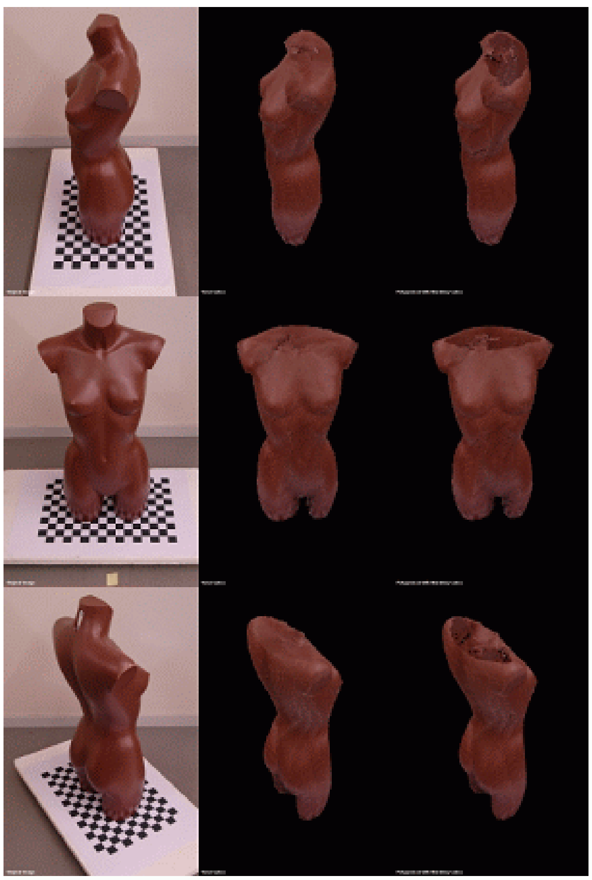

Figure 14. Three viewpoints (by row) of the 3D reconstruction obtained for the torso case. Left: original images. Middle: voxelized 3D model. Right: polygonized and smoothed 3D model.

Table 1 shows the final number of voxels and triangles of the 3D models obtained.

As it can be noticed in Figure 14, the torso object was not completely reconstructed. Maybe one possible reason for this is the fact that the used *GVC* implementation considers that the camera is align with the object's center, when the respective extrinsic parameters are relatively to the calibration pattern's center instead. This situation can cause the fail the reconstruction of objects with considerable height.

Another problem verified is that the torso reflects the calibration pattern on its surface. As consequence, the inferior volume of the reconstructed 3D model is not very accurate, both in terms of its shape and in color.

From the voxelized 3D model obtained, some geometrical measures can be determined, as height, length, width, centroid coordinates and volume. Tables 2 and 3 resume these values for both reconstructed objects. The obtained geometrical measures are compared with the real ones in Table 4.

Object	Hand	Torso
# voxels	11982	39596
# triangles	40685	122776

Table 1. Final number of voxels and triangles of both reconstructed 3D models.

Height (cm)	18.7513		
Length (cm)	11.4927		
Width (cm)	4.8390		
Centroid (cm)	x	y	z
	15.255	15.0654	19.0864
Volume (cm^3)	271.4245		

Table 2. Geometrical measures determined from the reconstructed 3D model for the hand case.

Height (cm)	66.8534		
Length (cm)	31.0868		
Width (cm)	20.7246		
Centroid (cm)	x	y	z
	33.5540	28.2606	35.1998
Volume (cm^3)	1914.3		

Table 3. Geometrical measures determined from the reconstructed 3D model for the torso case.

Object	Hand	Torso
Height (cm)	17.7	81.5
Length (cm)	15.5	40.0
Width (cm)	4.6	21.5

Table 4. Real geometrical properties of the two objects considered.

3. Conclusions and Future Work

The main final goal of this work was to develop a computational platform to reconstruct human external anatomical shapes, which are complex in shape, very smooth and low texturized. Starting with two uncalibrated image sequences, the used camera is calibrated, images containing the object to be reconstructed are segmented, in order to obtain the 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.

Although the obtained 3D reconstructions were relatively good, their quality is highly dependent on the accuracy of the camera's calibration results and in the object's segmentation on the used images. These dependencies limit the use of our computer platform *3D Builder* in controlled environments, because of the low texturized background restriction. Moreover, the complexity of the object's shape and the reflectance of its surface are aspects that must be considerer for more accurate 3D reconstructions. We can conclude that in controlled environments the methodology used is capable to obtain adequate 3D static reconstructions of human's external anatomical shapes from images. Its major contribution may be the fact that it is fully automatic and suitable for many real applications.

The future work will concern the implementation of auto-calibration procedures and the 3D reconstruction of non-rigid objects, like is considered in ([19]). As it was also verified that the *GVC* method is very computational demand and consequently slow, which can be unsuitable for many interactive and/or real-time applications, another possible future task could be the implementation of a coarse-to-fine approach (e.g. [20], [21]).

Acknowledgements

This work was partially done in the scope of the 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] M. Dooley, Anthropometric modeling programs - a survey, *IEEE Computer Graphics and Applications*, 2 (9), 17-25, 1982.

[2] F. Remondino, 3D Reconstruction of Static Human Body with a Digital Camera, *Videometrics VII, SPIE Electronic Imaging, Santa Clara (California), USA*, 5013, 38-45, 2003.

[3] N. D'Apuzzo, Modeling human faces with multi-image photogrammetry, *SPIE Proceedings of Three-Dimensional Image Capture and Applications V, San Jose, USA*, 4661, 191-197, 2002.

[4] T. Kanade, P. Rander and P. J. Narayanan, Virtualized reality: Constructing virtual worlds from real scenes, *IEEE Multimedia*, 4 (1), 34-47, 1997.

[5] G. Slabaugh, W. B. Culbertson, T. Malzbender and R. Shafer, A survey of methods for volumetric scene reconstruction from photographs, *International Workshop on Volume Graphics, New York, USA*, 21-22, 2001.

[6] L. Massone, P. Morasso and R. Zaccaria, Shape from occluding contours, *SPIE Conference on Intelligent Robots and Computer Vision, Cambridge, MA, USA*, 521, 114-120, 1985.

[7] T. Fromherz and M. Bichsel, Shape from multiple cues: Integrating local brightness information, *4th International Conference for Young Computer Scientist, Beijing, China*, 855-862, 1995.

[8] A. Laurentini, The visual hull concept for silhouette-based image understanding, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16 (2), 150-162, 1994.

[9] S. N. Seitz and C. R. Dyer, Photorealistic Scene Reconstruction by Voxel Coloring, *IEEE Conference on Computer Vision and Pattern Recognition Conference, San Juan, Puerto Rico*, 1067-1073, 1997.

[10] K. N. Kutulatos and S. M. Steiz, A Theory of Shape by Space Carving, *Technical Report TR692, Computer Science Department, University of Rochester, New York, USA*, 1998.

[11] G. G. Slabaugh, W. B. Culbertson and T. Malzbender, Generalized Voxel Coloring, *Workshop on Vision Algorithms, Corfu, Greece*, 100-115, 1999.

[12] J. S. D. Bonet and P. Viola, Roxels: Responsibility weighted 3D volume reconstruction, *IEEE 7th International Conference on Computer Vision, Kerkyra, Corfu, Greece*, 1, 415-425, 1999.

[13] L. Yang, 3D Surface Reconstruction from 2D Images, *Center for Visual Computing, Instructional Computing, Stony Brook University, EUA*, 2003.

[14] Y. Kuzu and O. Sinram, 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.

[15] Z. Zhang, A Flexible New Technique for Camera Calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22 (11), 1330-1334, 2000.

[16] K. Sande, A Practical Setup for Voxel Coloring using off-the-shelf Components, *Bachelor Project, Universiteit van Amsterdam, Netherlands*, 2004.

[17] M. Loper, Archimedes: Shape Reconstruction from Pictures - A Generalized Voxel Coloring Implementation, 2002.

[18] W. E. Lorensen and H. E. Cline, Marching cubes: A high resolution 3D surface construction algorithm, *International Conference on Computer Graphics and Interactive Techniques, ACM Press, New York, USA*, 21 (4), 163-169, 1987.

[19] S. Vedula, S. Baker, S. Seitz and T. Kanade, Shape and motion carving in 6D, *IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, SC, USA*, 2, 592–598, 2000.

[20] R. Szeliski, Rapid octree construction from image sequences, *Computer Vision, Graphics and Image Processing: Image Understanding*, 58 (1), 23-32, 1993.

[21] A. C. Prock and C. R. Dyer, Towards real-time voxel coloring, *DARPA - Image Understanding Workshop, Monterey, CA, USA*, 315-321, 1998.