Yue Pan
yuepan@student.ethz.ch

Assignment 8

Computer Vision
December 6, 2019

Task 1: Silhouette extraction

The parameters finally used in this exercise are listed in Table 1.

Table 1: Parameter settings for assignment 8

Parameter	Value
Silhouette Threshold	115
Bounding Box	[0.1 -0.1 -1.8 ; 2.2 1.3 2.4]
Volume Resolution	$128 \times 128 \times 256$
Volume Threshold	17

To distinguish the sculpture from the background, I adopted the binary thresholding with different threshold. Finally, the silhouette's threshold is set as 115 so that the silhouette of the sculpture is completely extracted without the background (floor and the table). The results are shown in Fig.1 and Fig.2.

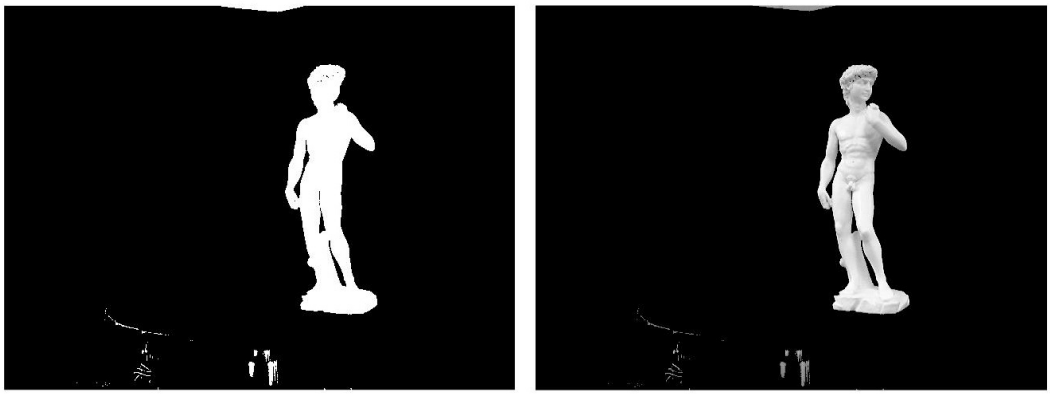


Figure 1: Silhouettes 1: Left: Binary image with the silhouettes' threshold; Right: Show pixels above the threshold with its original gray value.

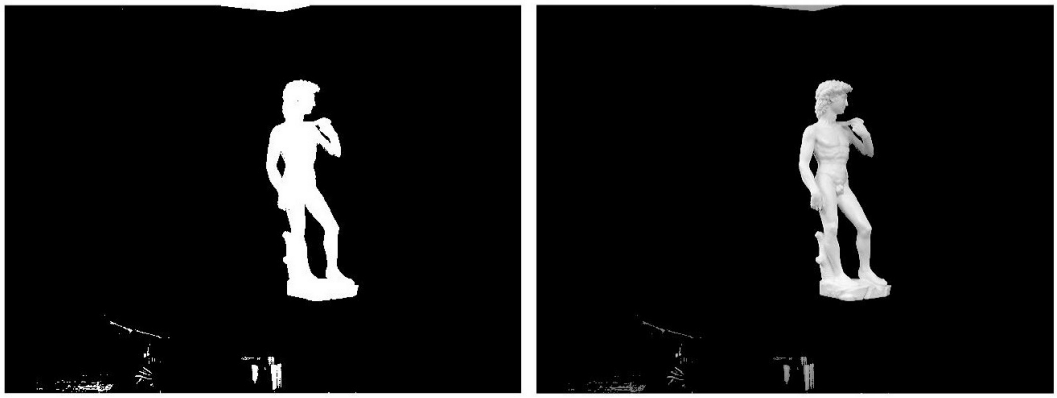


Figure 2: Silhouettes 2: Left: Binary image with the silhouettes' threshold; Right: Show pixels above the threshold with its original gray value.

Task 2: Volume of interest

The volume of interest here should be neither too small to omit some part of the sculpture or too large to make the searching area too large. So I manually clicked the approximate bounding box corners on one of the images and then project back the 2D pixel to the 3D world coordinate system using the camera projection matrix. I keep refining it and finally get the bounding box corners as following: [0.1 -0.1 -1.8 ; 2.2 1.3 2.4]. For this bounding box setting, the result shown in Fig.3 can be drawn. The bounding box is rendered in yellow and the corners are rendered in green.

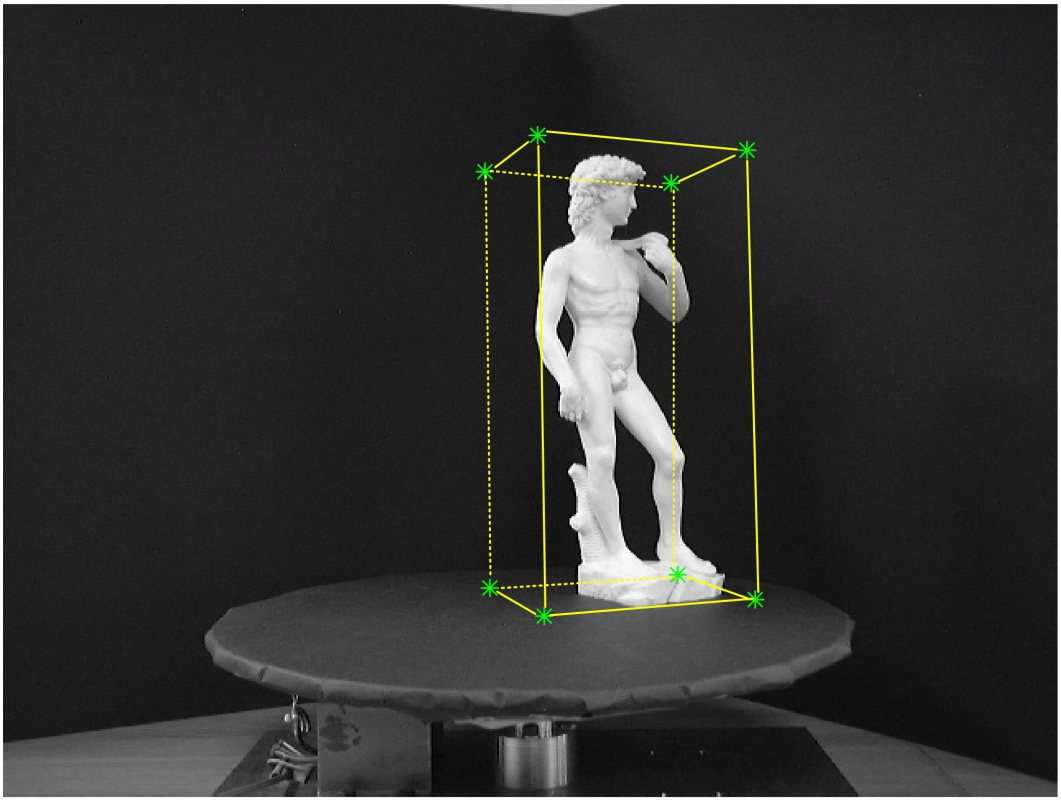
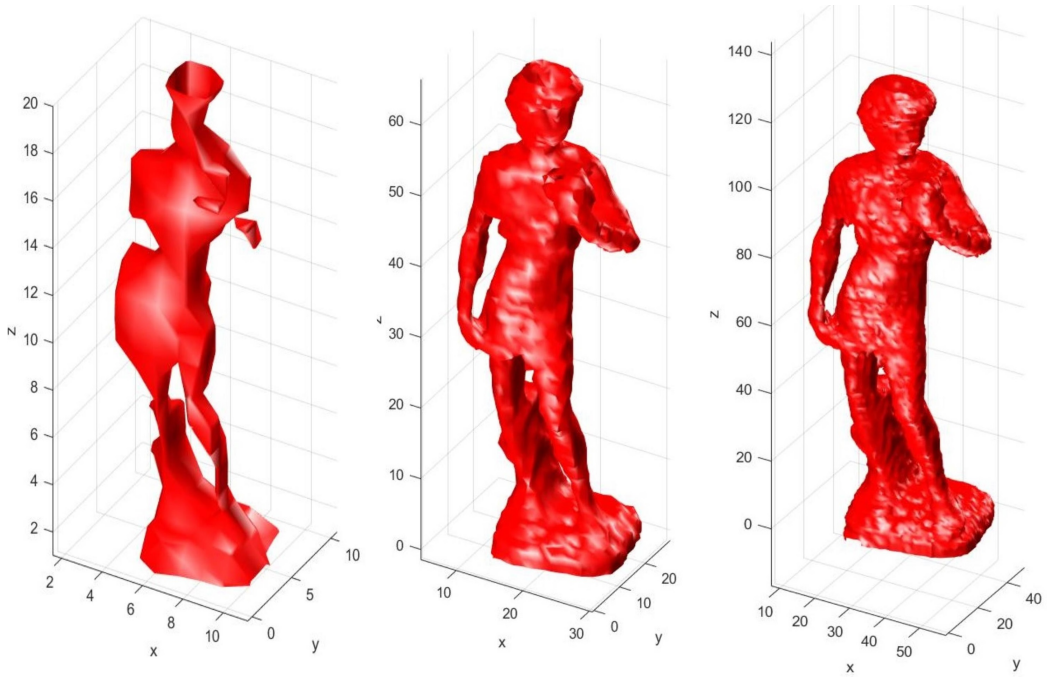


Figure 3: Image of the statue with the bounding box corners projected to the image.

Task 3: Visual hull

In this task, the volume of interest in task 2 is divided into a lot of small voxels which has the same size. The scores of these voxels are initialized as zero. Then for each camera, we go over all of these voxels. If the projected 2D coordinate of the voxel is inside the extracted silhoette, we add 1 to the score of the voxel. Then after processing all the cameras, the voxels whose score is not smaller than the volume threshold (set as 17) would be used to generate the 3D surface of the sculpture. I try different resolution of the voxel and finally set the volume resolution as 64x64x128. 3D surface results for different volume resolution are shown in Fig.4, in which we can tell the higher the resolution, the more meticulous 3D surface would be generated. The results for volume resolution (64x64x128) from different view points are shown in Fig.5.



(a) $10 \times 10 \times 20$ (b) $32 \times 32 \times 64$ (c) $64 \times 64 \times 128$

Figure 4: Generated 3D surface with different volume resolution

Task 4: Improvements

Summary:

3D reconstruction methods based on shape from silhouettes deal with the reconstruction of 3D solid models by volume intersection from a set of binary silhouette images to construct the approximated visual hull of the target object.

Questions:

What is the main drawback of the shape from silhouettes approach ?

According to (Mulayim et al. 2003), one of the main drawback of the silhouette-based volume intersection approach is the additional excess volume coming from insufficient number of viewing positions and from the concavities on the object to be modeled. As for the surface concavities problem, it is even impossible to be solved with infinite number of surrounding images. For the cases dealing with multiple silhouettes using limited number of silhouette images, the problem of ghost objects (generated 3D shape that are not corresponding to any silhouette) may occur.

Secondly, the silhouettes of the interested object are not easy to be extracted accurately from the background and surroundings for some challenging scenarios. A single threshold of gray value may not work under such circumstance. Some image segmentation methods may be helpful to deal with this problem.

Thirdly, the interested volume (bounding box of the interested object) is also hard to determine. To avoid omit some part of the object, we'd better select a larger interested of volume. Under such condition, if the requirement of voxel resolution is high and photos are taken from a lot of perspective, the algorithm would be very time-consuming.

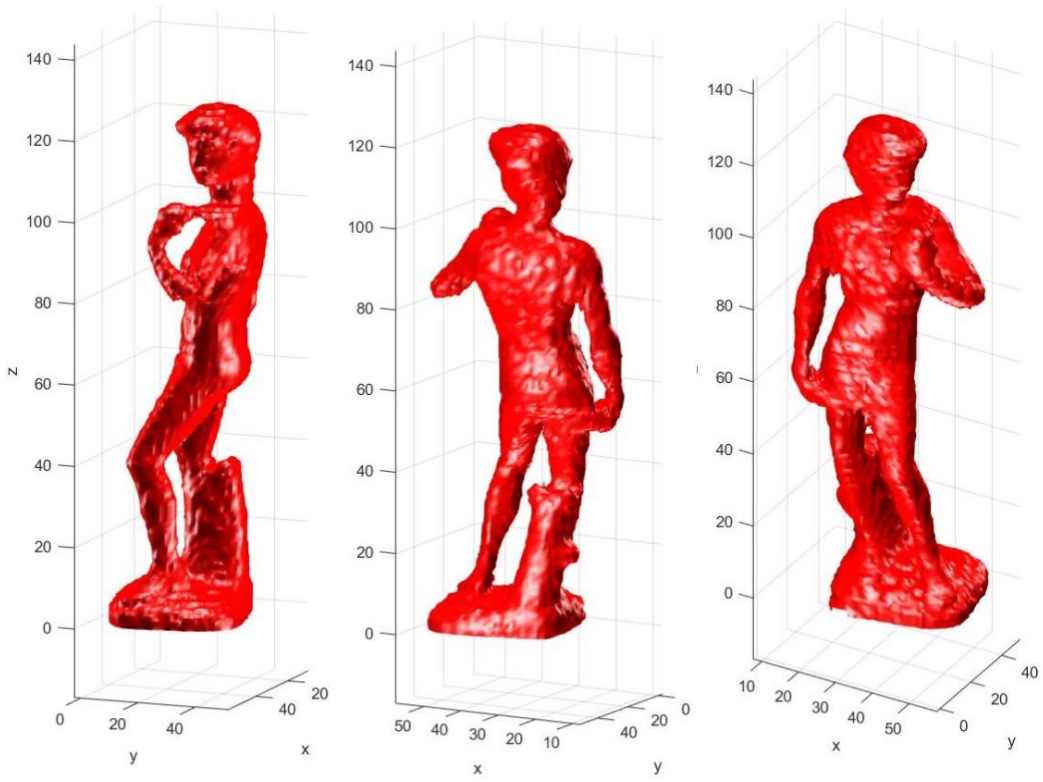


Figure 5: Different perspectives of the generated 3D surface

Is there additional information in the images that could be used?

Firstly, we can try to adopt the RGB information of the images to extract the silhouettes more accurately because RGB values provide more information than gray value when doing the image segmentation.

Besides, the shape reconstruction problem can also be solved using structure from motion and multi-view stereo since the camera matrices are given. Then maybe we can use the model generated by multi-view stereo to refine the shape generated by shape from silhouettes method, which is introduced in (Lin and Wu, 2008).

References:

[1] Mülayim A. et al., 2003, Silhouette-based 3D Model Reconstruction from Multiple Images, IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS, PART B

[2] Lin H. and Wu J., 2008, 3D Reconstruction by Combining Shape from Silhouette with Stereo, ICPR 2008