# Computer Vision Assignment 9: Shape from X

*Charlotte Moraldo*

## 1. Silhouette Extraction

To extract the silhouette of the statue from the image, thresholding is used: pixels with intensity above the threshold are set to 1, and the ones below are set to 0.
Finding the best threshold isn't always easy. Indeed, some parts of interest are consisted of pixels with lower intensity than the rest of the silhouette, for example the areas with more shadow like the part under the left arm of the statue. On the opposite, some parts that do not interest us are of higher intensity, such as the floor or the edges of the table. After several tries, I ended up choosing a threshold equal to **threshold=110**, giving the silhouettes shown in figure 1 below:

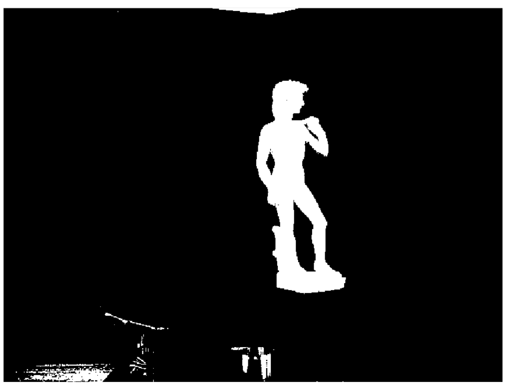


**Figure 1:** Shows pixels with value 0 or 1

**Figure 2:** Shows pixels above threshold shown with their original intensity

## 2. Volume of Interest

The following bounding box was chosen:

$$[min_x \ min_y \ min_z; \ max_x \ max_y \ max_z] = [0.2 \quad -0.1 \quad -1.8 \ ; \quad 2.3 \ \ 1.4 \ \ 2.5]$$

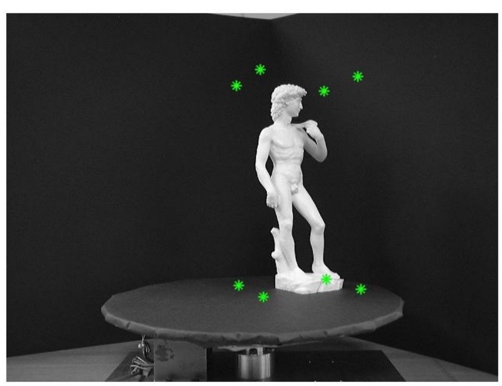


**Figure 3:** Image of the statue with volume box corners displayed in green

As seen on figure 3, the statue is tightly contained in the box delimited by the green crosses. As suggested in the slides, I decided to define a volume of sizes 64x64x128.

## 3. Visual Hull

To compute the visual hull, we begin by defining a 3D-grid filled with zeros and with the same dimensions as the volume defined in point 2. We then go over all the voxels, for each camera. The projection of each volume point (voxel in world coordinates) is computed using the given matrix $T$, and this point is projected on the bidimensional plan of the camera using the camera matrix $P$. After normalizing this 2D point, we check if it is in the extracted silhouette, and if yes, we add 1 to the corresponding voxel in the 3D-grid.
This method works pretty well, as can be seen in the screenshots below showing different views of the 3D iso-surface extracted:

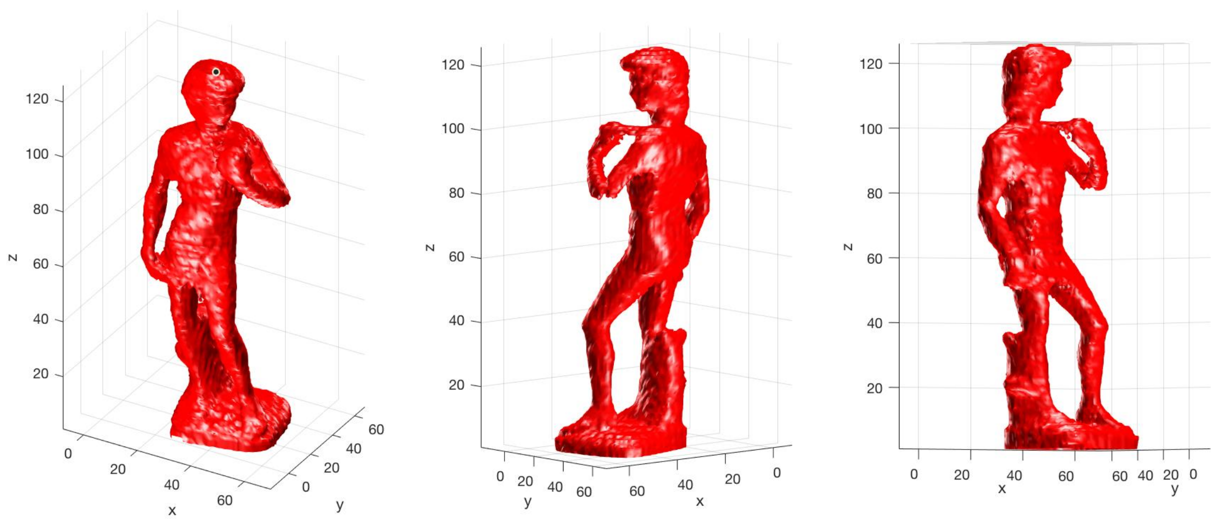
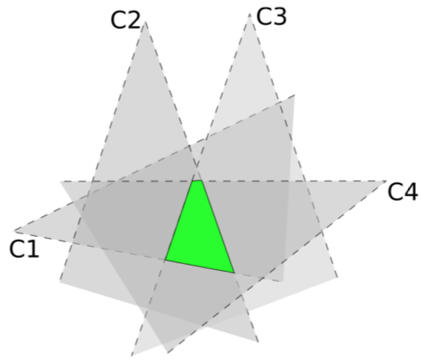


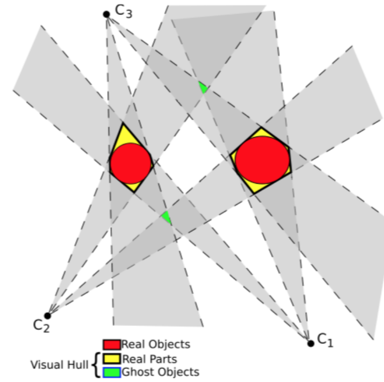

**Figure 4:** Different views of the 3D iso-surface extracted from all the cameras

## 4. Improvements

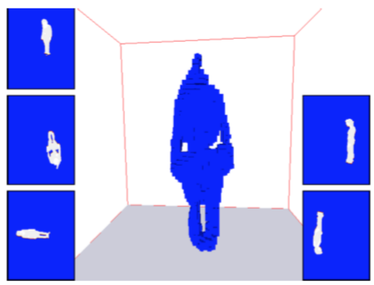
The shape from silhouettes method is a very straight-forward and robust approach to capture dynamic 3D scene in real-time. However, the volumes reproduced may have many drawbacks in different situations than the one studied in this assignment. Some examples are illustrated in the following figure, extracted from a paper about 'Silhouettes Fusion for 3D Shapes Modeling with Ghost Object Removal' [1].




(a)




(b)



(c)

**Figure 5:** Main drawbacks of the Shapes from Silhouettes algorithm [1]

As explained in [1], one of the main drawbacks is the number of cameras and the way they are placed. The shapes from silhouettes algorithm only extract structures that are in the intersection of the camera's fields of view (green area in Fig5-a). In this assignment, this issue was not encountered as the statue was rotating around and a lot of pictures were taken, all from an angle such that the statue was well seen. We always had the statue in sight and from a lot of different angles.
To overcome this, [1] suggests using a subset of cameras in order to determine whether a 3D point is in the object of interest or not, unlike with the shape from silhouette method that uses the complete set of cameras.

The accuracy of the shape from silhouette method is also limited by the accuracy of the silhouette extraction (Fig 5-c) [1]. In this assignment, the silhouette extraction through thresholding wasn't that complicated because the object of interest (the statue) has in general a much higher intensity than the background and the other elements of the picture. Elements such as the table edges or the floor which could have compromised the silhouette extraction due to their lighter color were discarded by choosing a tight volume box around the statue. However, in an image where it is harder to distinguish the object of interest from its background and surroundings, silhouette extraction becomes much harder.
An idea to overcome this issue would be to use a silhouette extraction method more robust than just a simple thresholding, such as feature extraction (using SIFT for instance) in order to define the corners. [1] suggests adopting a probabilistic approach in order to process data issued from the different images.

Another important drawback of this method is the apparition of ghost objects [1]. Ghost objects are 3D areas that do not actually correspond to any part of the extracted silhouette. As explained in Fig5-b, when there are several objects, the field of view of some cameras can intersect in an area where there is actually no real object. This phenomenon can also be noticed in our 3D iso-surface computed in part 3 (Fig.6), but for a different reason: it could be due to the fact that the arms are quite far away from the body, resulting in these ghost objects below the hand.

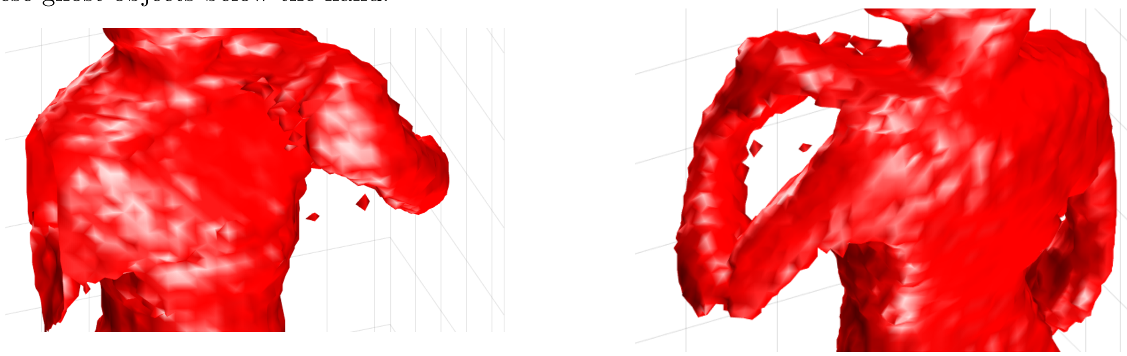


**Figure 6:** Ghost objects in the 3D iso-surface

To overcome this problem, [1] proposes a way to identify ghost objects, the concept being that if a pixel contained in a silhouette is derived from exactly one 3D connex part, then that connex part belongs to a real object.

Some parts of the structure, such as the thorax or the back of the statue, are not processed by shapes from silhouettes. Indeed, this method only looks at the overall silhouette of the object of interest, and not the details. A way to extract them would be to use more detailed information such as depth, structure, shadings, etc. For instance, shapes from silhouettes could be combined with stereo matching, which infers the depth information of a structure based on points correspondences obtained from image pairs (as seen in assignement 6). This is proposed in [2], in order to refine the shape and extract more precise geometric models.

## 5. References

[1] Michoud B., Guillou E., Briceo H.M., Bouakaz S. (2008). Silhouettes Fusion for 3D Shapes Modeling with Ghost Object Removal. LIRIS - CNRS UMR 5205, Universite Claude Bernard Lyon 1

[2] Lin H.-Y., Wu J.-R. (2008). 3D Reconstruction by Combining Shape from Silhouette with Stereo. 1-4. 10.1109/ICPR.2008.4761016.