Computer Vision, 5th assignment

Mathis Lamarre

Silhouette extraction

In order to extract the silhouette, a simple thresholding technique is used. The pixels whose intensity value is greater than the threshold are set to 1, the others to 0.

A B

Fig. 1: Silhouette extracted from the images by thresholding. A shows the pixels with value 1 or 0. B shows the original intensity of the chosen pixels.

Finding an optimal threshold is hard as some parts of the images have a high intensity even though they are of no interest, like the floor or the edges of the table. It is not a real problem as they will not be in the bounding box anyway. Some parts of the statue are also problematic as they have a low intensity, such as the bottom of the arm. The results of Fig.1 are obtained with $threshold = 111$.

Volume of interest

After trying several values, the following bounding box was chosen:
$min_x = 0.25, min_y = -0.11, min_z = -1.8$, $max_x = 2.07, max_y = 1.1, max_z = 2.5$

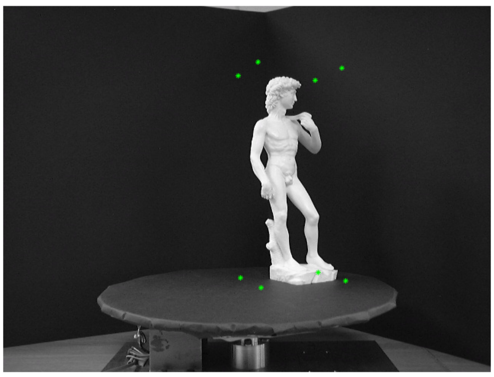

Fig. 2: Image of the statue with the corners of the bounding box displayed.

As seen on Fig.2, the box contains the entire statue and is a tight fit.

Visual hull

In order to compute the visual hull, we first define a 3D-grid filled with zeros which as the dimensions we chose for the volume. Then, for each camera, we iterate over all the voxels. We compute the projection of the volume point corresponding to the voxel in world coordinates with the T matrix provided. Then, we project this point on the 2D plan of the camera with the P matrix specific to each camera. Finally, after normalizing it, we check if this 2D point is in the silhouette previously extracted and if such is the case, we add 1 to the voxel.

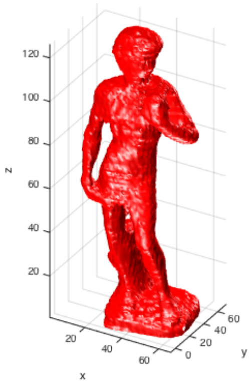

Fig. 3: 3D iso-surface extracted from all the cameras with a 64x64x128 grid.

The method is quite satisfactory and shows good results (see Fig.3). This iso-surface was extracted with a threshold of 17, which is inferior to the number of cameras.

Improvements

The Shape from Silhouettes method is very straight-forward and gives good results with certain conditions, like in our case. Nevertheless, it has some drawbacks.

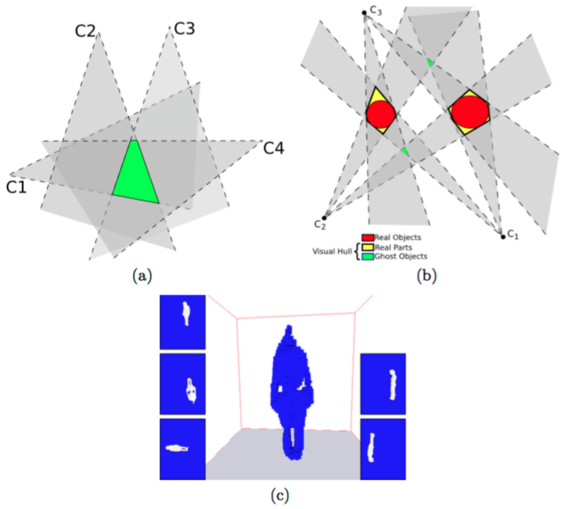

Fig. 4: Example of drawbacks, figure from [1]. (a) shows the intersection of viewing cones. (b) shows ghost objects and real objects. (c) shows the dependance between silhouette extraction and reconstruction.

One of the main limitations is the number of cameras and their placement. This method can only extract objects that are in the intersection of the camera viewing cones (see Fig.4-a). We didn't encounter this problem because we had lots of pictures with the statue rotating around so we always had it in sight and from lots of angles. However, when using only two camera angles as in Fig.5, the results are much worse.

Fig. 5: 3D iso-surface from only 2 cameras.

Another problem is the occurence of ghost objects that do not correspond to any actual matter. As seen on Fig.4-b, when there are several objects, the visual cones of some cameras can intersect at at spot where there is no real object. In our case, it can be created by the fact that the arms are away from the body and results in a ghost object below the hand (see Fig.6).

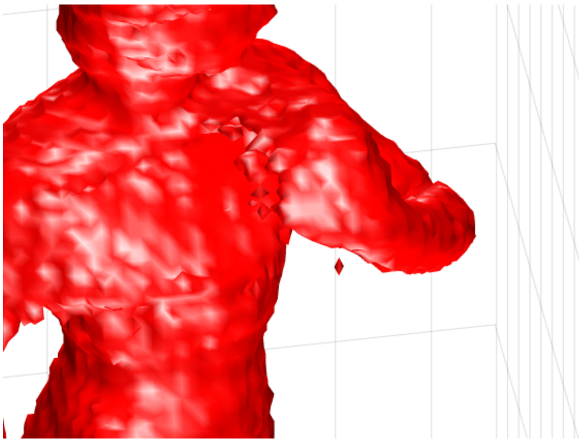

Fig. 6: Ghost object seen below the hand due to the fact the statue is not compact.

The accuracy of the Silhouette from Shape is also limited by the accuracy of the silhouette extraction (see Fig.4-c). In our case, we could not find of perfect threshold but it was not a big issue as the immediate surrounding of the statue was really dark, and the other clear parts (like the floor or the edges of the table) were far away from the bounding box. However, in a case where it is hard to distinguish the object of interest and its surrounding, the silhouette extraction is critical. A way to remedy this would be to use a more robust method than a simple thresholding, such as a feature extraction (with SIFT for example) to define the corners.

Some shapes, such as a cavity on the thorax of the statue, can not be perceived with the Shape from Silhouette technique as they will always be in the overall silhouette of the object, no matter the angle. A way to extract them would be to use more information of the image, like the shading, shadows, texture or light sources.

References

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