Vision and Image Processing:
Assignment 3 Photometric Stereo

Mostafa Elshamy, Tom Vig, Andreas Hammer

December 2024

1 Modelling

1.1 Lambert's law

For a generic pixel p, light source s and normal, Lambert's law states that:

$$I(p) = \rho(\mathbf{x})\,\mathbf{s}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x})$$

Where $\mathbf{x}$ is the corresponding surface point, and $\rho(\mathbf{x})$, $\mathbf{s}(\mathbf{x})$, $\mathbf{n}(\mathbf{x})$ are the albedo, light source direction and surface normal at point $\mathbf{x}$.

1.2 Dealing with self shadows

Lambert's law can be modified to deal with self shadows by ensuring that the product of the light source direction and the surface normal is a non-negative value:

$$I(p) = \rho(\mathbf{x})\, \max\left(0, \mathbf{s}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x})\right)$$

1.3 Difference between self and cast shadows

- **Self-shadows:** Occur on the same surface where light cannot reach due to the surface angle.
- **Cast shadows:** Created when one object blocks light from reaching another surface. This is a non-local phenomenon and is therefore not addressed by Lambert's law.

1.4 Modelling limits of Lambert's law

Other than not addressing cast shadows, Lambert's law assumes a matte surface where light is evenly scattered in all directions. This does not account for specular reflection (e.g. mirrors) or complex materials with mixed reflection behaviors (e.g. milk).

1.5 Photometric Stereo

Woodham's approach to Photometric Stereo utilizes multiple images $I_1, I_2, \ldots, I_k$ taken with a fixed camera and varying light directions. Through Lambert's law, we get the following system of equations:

$$I_i(p) = \rho(\mathbf{x})\,\mathbf{s}_i \cdot \mathbf{n}(\mathbf{x}), \quad i = 1, 2, \ldots, k$$

By using at least three light sources, this system of equations can be solved to estimate the surface normal and albedo.

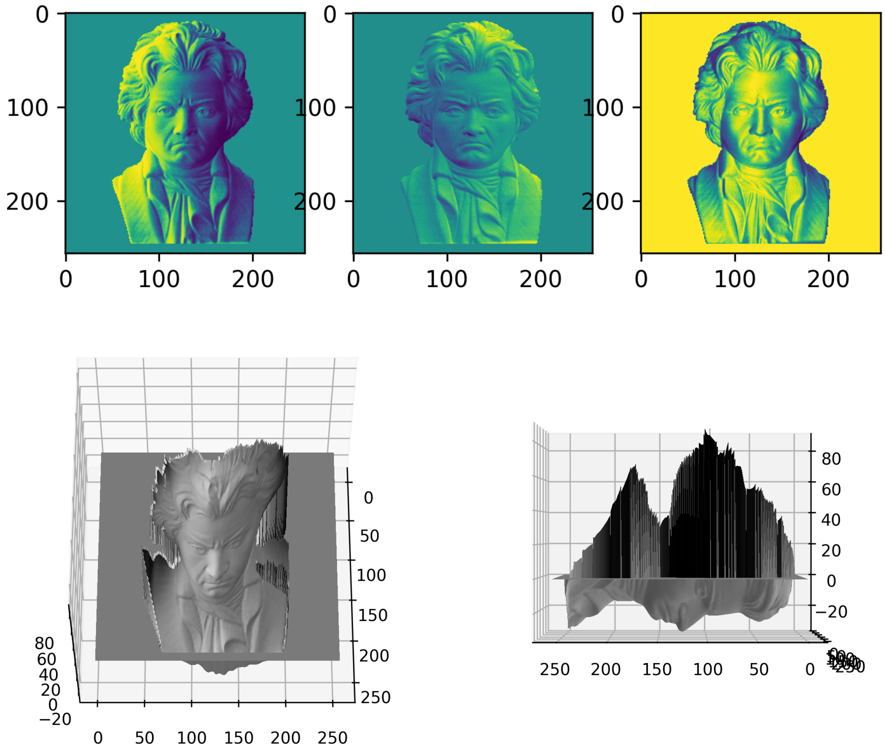

Figure 1: Woodham on Beethoven dataset

1.6 Using RANSAC

RANSAC is a method to fit a model with a minimal amount of outliers. In image modeling, we can use it to estimate a pixel's normal and albedo in the following way:

1. Randomly sample a subset of intensity values from the set of images provided by the Photometric Stereo.
2. Estimate normals and albedo using this subset.
3. Compare the Estimated normals and albedo to the observed intensities in all the images.
4. Repeat the process multiple times and select the model with the highest number of inliers (intensities that agree with the model up to a threshold).

2 Beethoven and Matte Vase Datasets

Both the Beethoven and Matte Vase datasets underwent the same Woodham processing, with figures 1 and 2 displaying the three normals, as well as their 3-dimensional surface construction. The albedo-constructed field shows a relatively accurate 3D reconstruction of Beethoven's face. It show's details such as the furrowing of his brows and fine lines around his mouth. Likewise for the matte vase, the 3D depth construction shows a smooth surface, as is expected.

Focusing now on the normalized images (top three in both figures), we can see that the n1,n2 images show variations in the object in the X and Y directions, respectively. Again, the clean datasets result in clear images, that show the smoothness of the matte vase, and the details in Beethoven's face. The third normal image highlights the parts of the object that are closest to the camera, emphasizing the middle bulge in the vase, and Beethoven's nose and forehead.

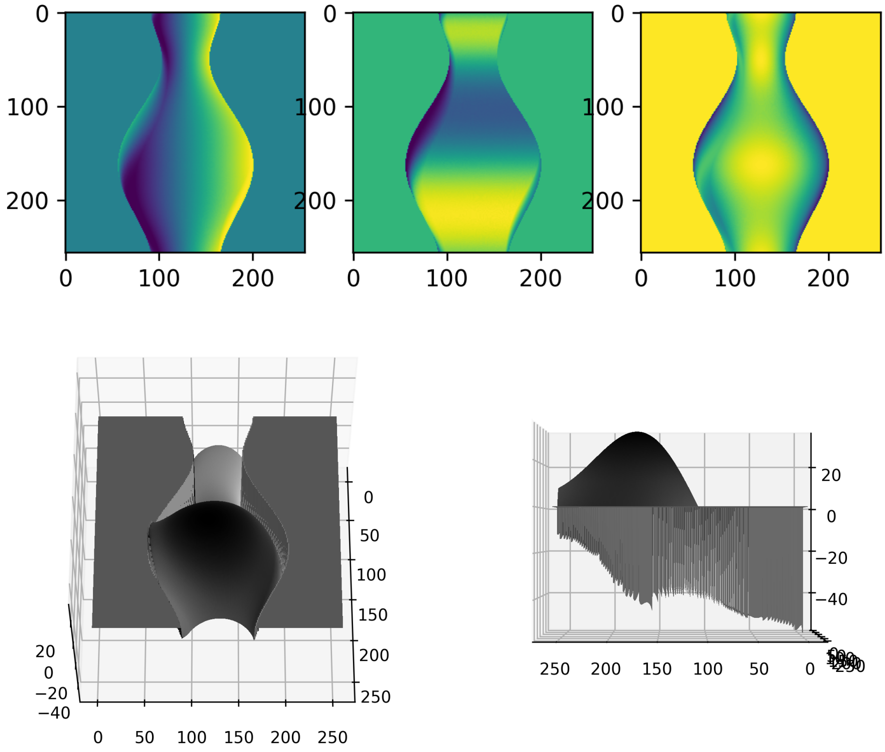

Figure 2: Woodham on mat vase dataset

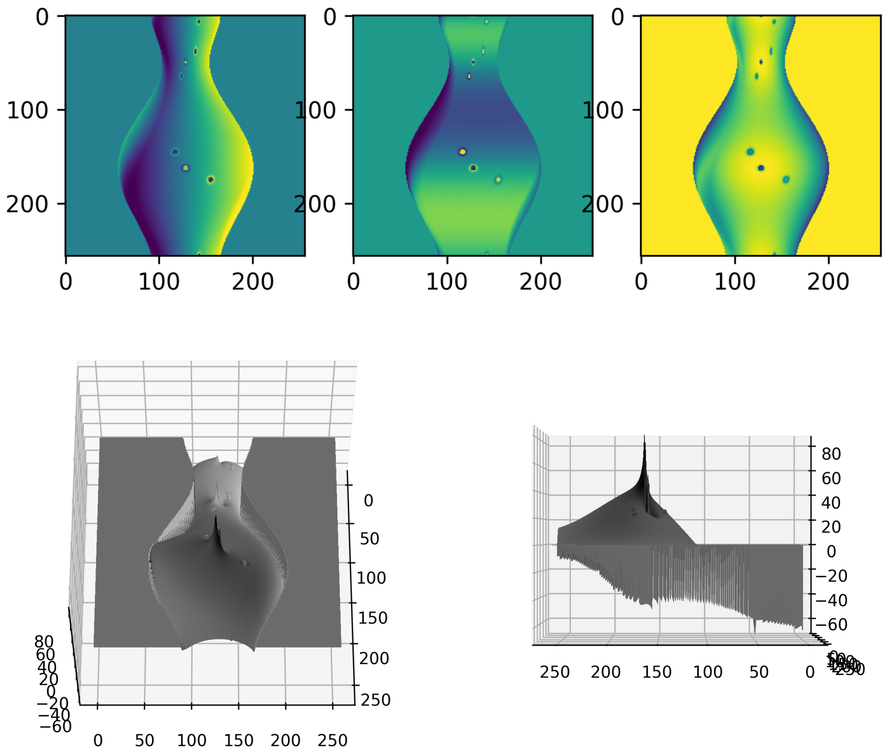

Figure 3: Woodham on shiny vase dataset

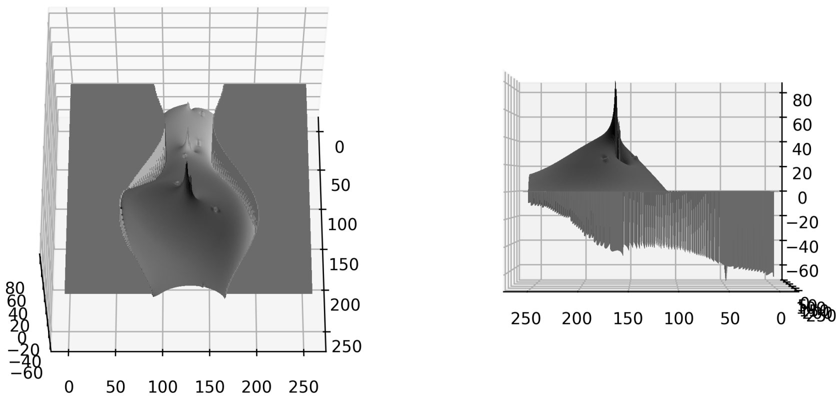

Figure 4: RANSAC on shiny vase dataset

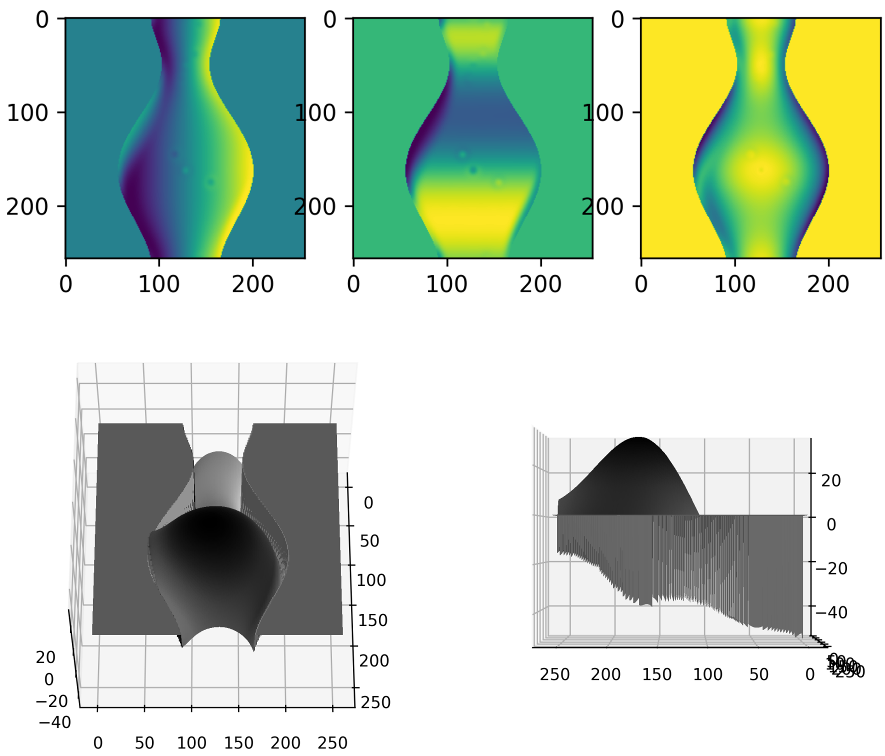

Figure 5: RANSAC and smoothing on shiny vase dataset

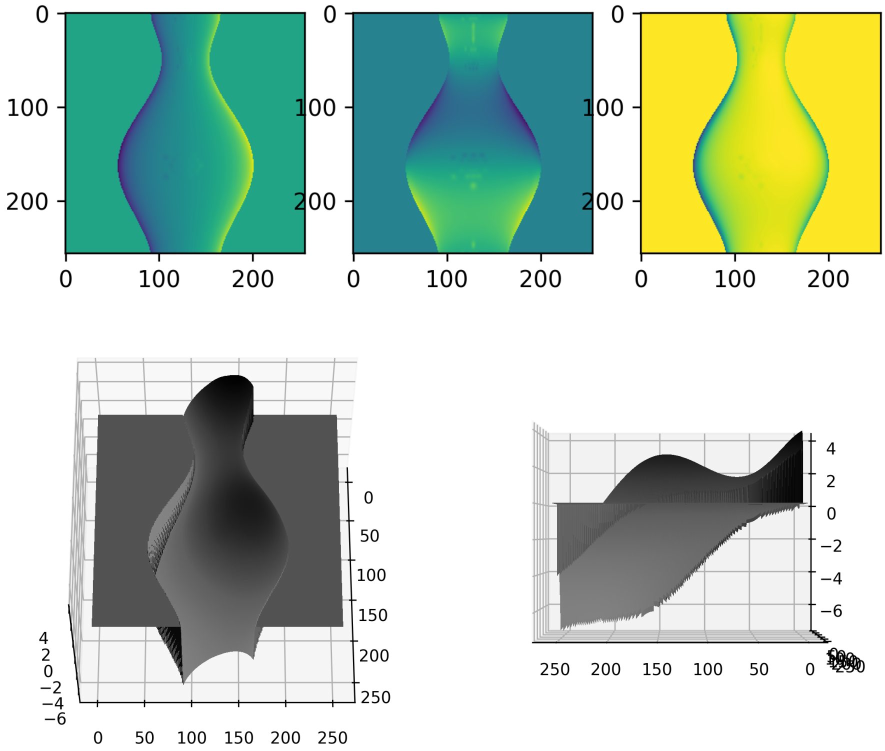

Figure 6: Woodham on shiny_vase2 dataset

3 Shiny Vase Dataset

On figure 3 we can see that the Woodham processing was applied similarly to the previous section. However, due to the shiny surface, the results turn out quite different. We can see on the 3D surface plotted at the bottom 2 images of figure 3, that the shiny spots on the vase image have translated to large peaks in the graph. This is also reflected in the three normals. Particularly on the upper right image, the gleam of the shiny spots of the vase causes an inaccuracy in both the 3D image and and the three perspective images. This appears to be quite similar after applying the RANSAC filtering algorithm in figure 4 with threshold of 2.0 (moderate-low for this algorithm). However, a major increase in accuracy when applying both RANSAC and smoothing. In figure 5, we can see the impact of using both a RANSAC filtering with threshold 2.0, as well as 100 iterations of the smoothing algorithm. This results in a much more accurate result, both in terms of the Albedo normalized images, but also in the 3D construction. The reflection points are still slightly visible on the n1 and n2 Albedo images, but not on the n3 output, which shows protrusions towards the camera. This indicates that the RANSAC algorithm has successfully estimated the lack of depth of the reflection points.

4 Shiny Vase 2 Dataset

The results of this task are similar to that on the Shiny Vase dataset. Figure 6 shows a well smoothed result using only Woodham approximation. We felt it unnecessary to present the RANSAC nor RANSAC and smoothed images, as they did not improve much upon the initial Woodham estimation.

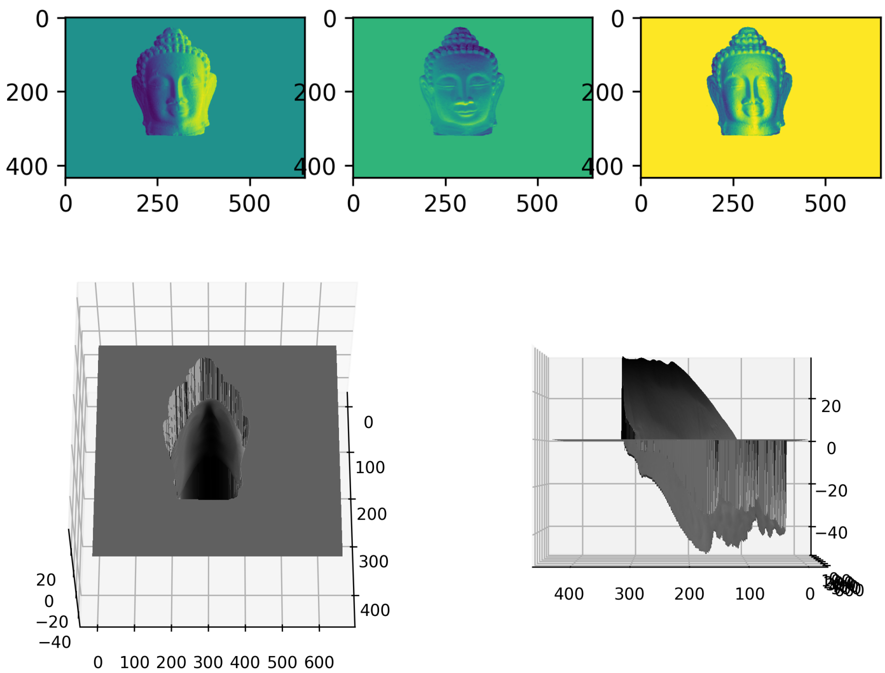

Figure 7: Woodham on Buddha dataset

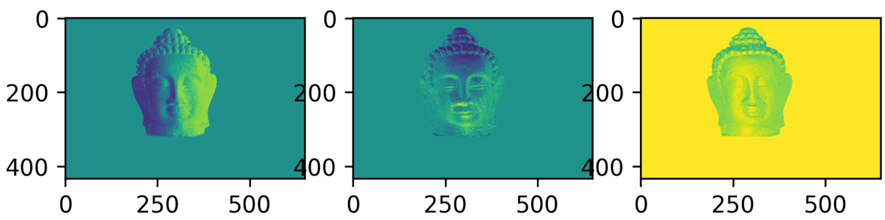

Figure 8: RANSAC on Buddha dataset

5 Buddha Dataset

The results of applying the process described in the assignment can be shown in figure 7. Here we see the Woodham estimation was somewhat accurate, though much of the detail in the Buddha statue's face has been lost. There are many shadows caused by lighting inconsistencies, causing a blurry image. After applying RANSAC with a threshold of 25.0, we see much more fine detail. Increasing the threshold further did not provide significantly improved results.

6 Face Dataset

For the face dataset, we have many more images to work with. Observe figure 9. The Albedo perspectives are the result of using a RANSAC threshold of 10.0, allowing for some surface detail to show, whilst remaining the most obvious glare and noisy artifacts. After applying smoothing for 100 iterations, figure 10 becomes overly smoothed, loosing much detail of the RANSAC results. The depth map created by the RANSAC algorithm alone is sufficient for this data set, though the nature of this being a real dataset results in visible graininess in the Albedo perspective images that is not present in the clean data sets.

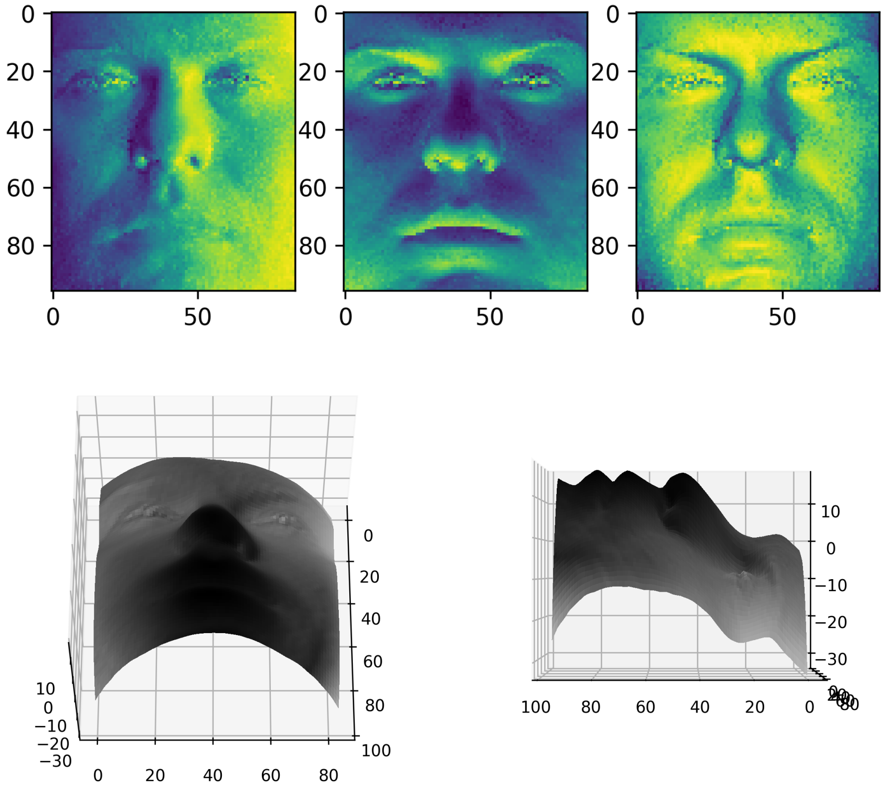

Figure 9: RANSAC on Face dataset

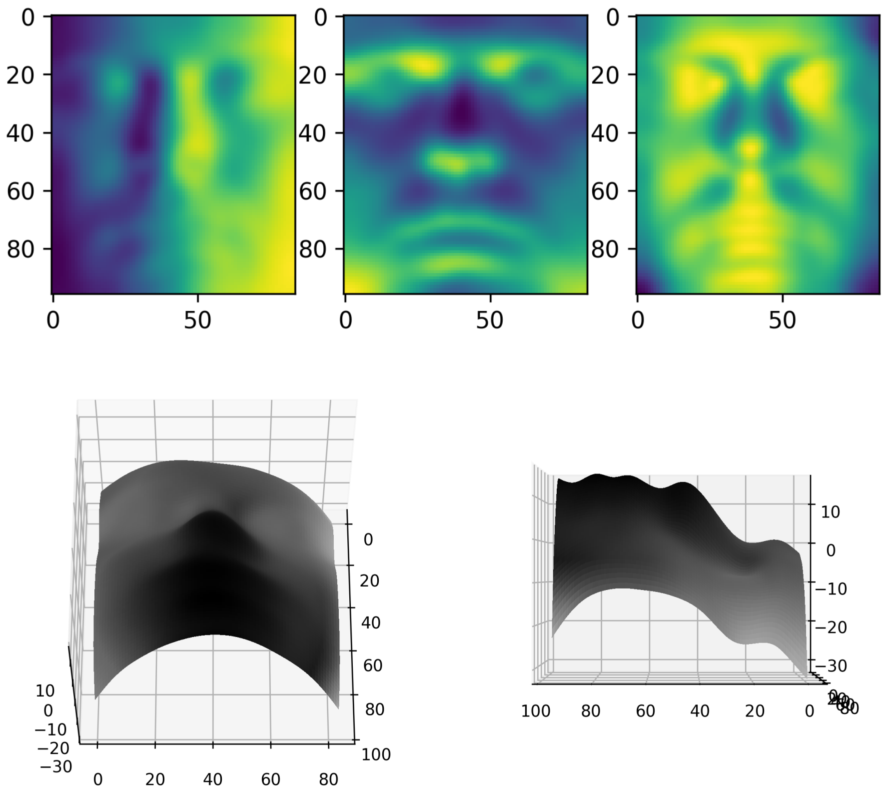

Figure 10: RANSAC and smoothing on Face dataset