16-720 A Computer Vision

Assignment 6

Sourojit Saha

sourojis

Collaborated with Sohan Kulkarni(sohank) and Arjun Sengupta(arjunsen)

1. Calibrated photometric stereo

1.a Understanding $\boldsymbol{n}$-dot-$\boldsymbol{l}$ lighting

$\vec{n}$ is the normal vector to the plane of area dA. $\vec{l}$ is the direction of incident light rays on the area dA. $\vec{v}$ is the vector of reflected light along the direction towards the camera. The amount of light incident on the surface is proportional to the cosine of the angle between the $\vec{n}$ and $\vec{l}$. This can also be found by taking the dot product of $\vec{n}$ and $\vec{l}$. The viewing angle does not matter because we are considering Lambertian surface which has a constant BRDF.

1.b Rendering $\boldsymbol{n}$-dot-$\boldsymbol{l}$ lighting

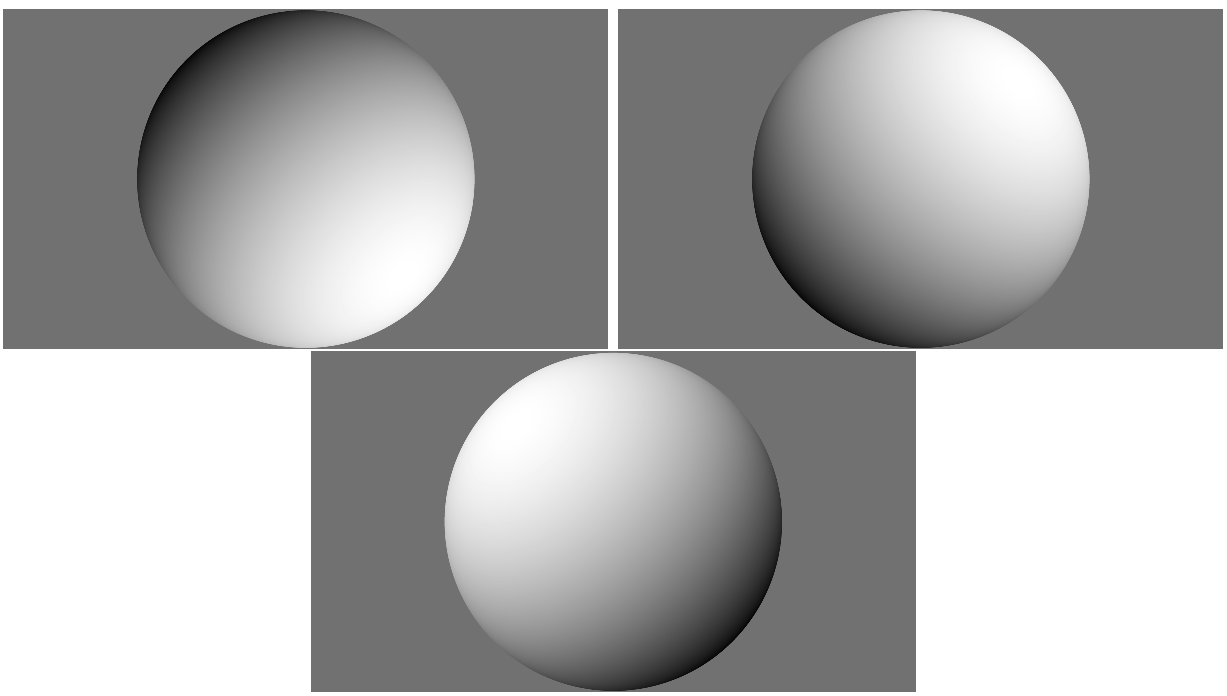

Figure 1: Rendering n-*dot*-l *lighting*

1.c Loading data

```
def loadData(path = "../data/"):
    I = None
    L = None
    s = None
    img_path = path + 'input_1.tif'
    temp = Image.open(img_path)
    temp_array = np.array(temp)
    I = np.zeros((0, temp_array.shape[0]*temp_array.shape[1]))
    print ("Shape of I ", I.shape)
    for i in range(1, 8):
        img_path = path + 'input_' + str(i) + '.tif'
        dummy_img_array = np.asarray(Image.open(img_path), dtype = np.
    uint32)
        dummy_img_colorsp = rgb2xyz(dummy_img_array)
        dummy_vector = dummy_img_colorsp[:, :, 1]
        dummy_vector = dummy_vector.flatten()
        I = np.vstack((I, dummy_vector))
        if i==7:
            r, c, _ = dummy_img_array.shape
            s = [r, c]

    L = np.load(path + 'sources.npy')
    L = L.T

    return I, L, s
```

1.d Initials

B is a set of vector (pseudonormal) in 3-dimensional space. Hence, the maximum possible rank of B is 3. Similarly, L gives the direction of light source in 3-dimensional space and hence can have a maximum rank of 3. Since $I (= L^T B)$ is a linearly related to B via L^T, I should also have a rank of 3. The singular values of I are:

$Singular \quad values \quad of \quad I : [2.86634305e-04, \quad 3.41436112e-05, \quad 2.39722784e-05,$ $7.36215196e-06, \quad 5.56202966e-06, \quad 4.39464744e-06, \quad 3.59846261e-06]$

As we see, the I matrix is has a rank 7. This is because of noise. I can be forced to be rank 3, by setting all the singular values, except the top 3 singular values, to 0 as: $I = U \times \Sigma_{new} \times V^T$.

1.e Estimating pseudonormals

Given:

$$L^T B = I \Rightarrow L L^T B = L I \tag{1}$$

Comparing with:

$$Ax = y \tag{2}$$

We get:

$$\begin{aligned} A &= L L^T \\ x &= B \\ y &= L I \end{aligned} \tag{3}$$

Solving, we get:

$$B = (L L^T)^{-1} (L I) \tag{4}$$

1.f Albedos and normals

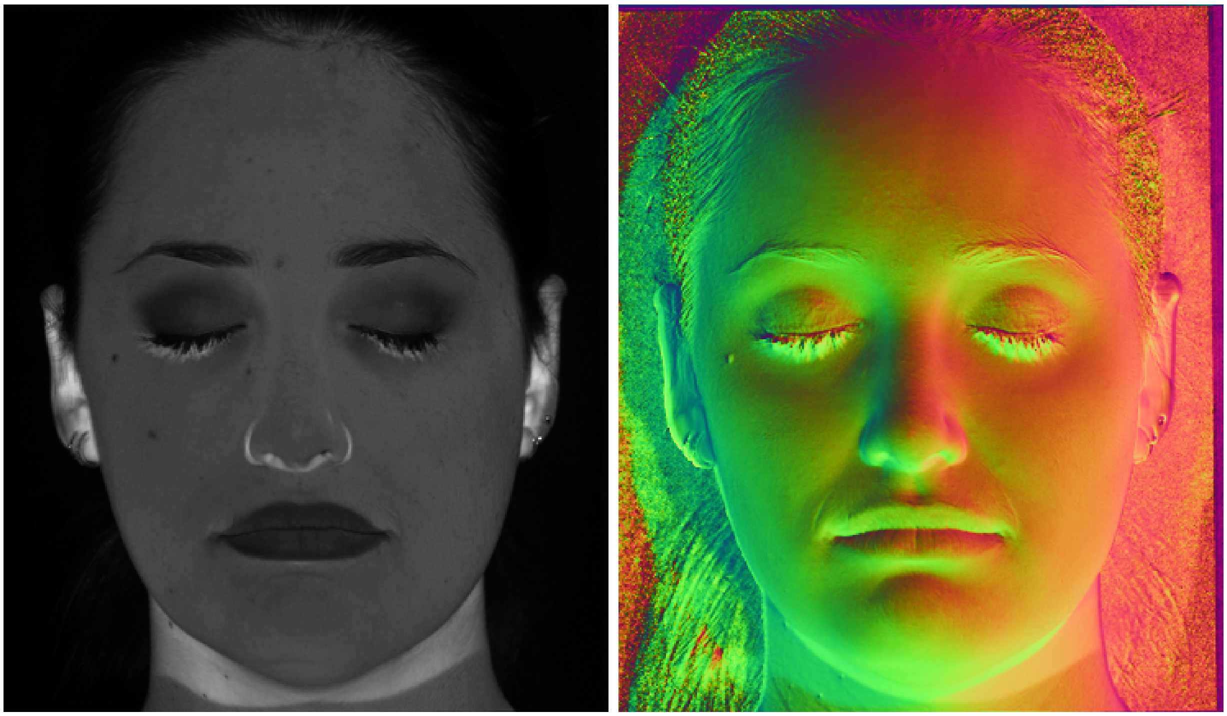

Figure 2: Visualization of albedos(left) and normals(right).

In the albedos image, we see unusual brightness at the ears, below the eye lashes, around the nose and in the neck region. As albedos contains the information regarding the reflectance of the surface, such brightness is expected due to the skin there being smooth, and/or high oil on the skin.

The normals image match the expectation of the curvature of the image.

1.g Normals and depth

We know:

$$\frac{\partial f(x,y)}{\partial (x)} = z_{x+1,y} - z_{x,y} \tag{5}$$

From the diagram, we an see that:

$$\begin{aligned} V_1 = (1, 0, z_{x+1,y} - z_{x,y}) \Rightarrow N \cdot V_1 = 0 \\ n_1 + n_3(z_{x+1,y} - z_{x,y}) = 0 \Rightarrow \tfrac{\partial f(x,y)}{\partial (x)} = \tfrac{-n_1}{n_3} \end{aligned} \tag{6}$$

Similarly:

$$\frac{\partial f(x,y)}{\partial (y)} = z_{x,y+1} - z_{x,y} \tag{7}$$

From the diagram, we an see that:

$$\begin{aligned} V_2 = (1, 0, z_{x,y+1} - z_{x,y}) \Rightarrow N \cdot V_2 = 0 \\ n_2 + n_3(z_{x,y+1} - z_{x,y}) = 0 \Rightarrow \tfrac{\partial f(x,y)}{\partial (y)} = \tfrac{-n_2}{n_3} \end{aligned} \tag{8}$$

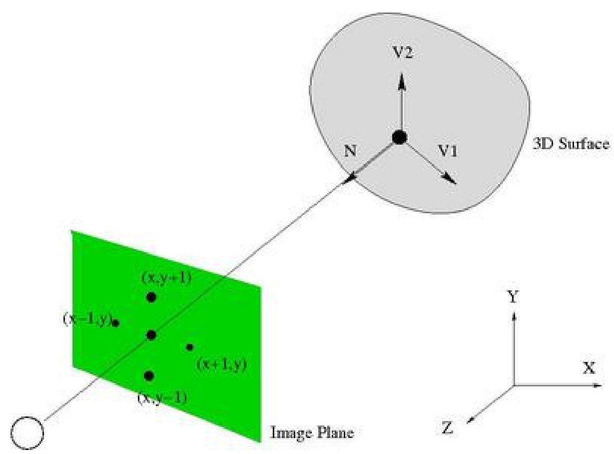

Figure 3: Ref. diagram.

1.h Understanding integrability of gradients

Given:

$$g = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \tag{9}$$

Hence:

$$g_x = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{10}$$

$$g_y = \begin{bmatrix} 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \end{bmatrix} \tag{11}$$

Given $g(0,0) = 1$, we can re-construct g using g_x and g_y as:

- Using g_x, we get the first row of $g = \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$. There after, using g_y, we can compute the whole g matrix as:

$$g = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \tag{12}$$

- Using g_y, we get the first column of $g = \begin{bmatrix} 1 & 5 & 9 & 13 \end{bmatrix}^T$.There after, using g_x, we can compute the whole g matrix as:

$$g = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \tag{13}$$

We see both the methods give the same results.

We can modify the g_x as $g_x = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. Given $g(0,0) = 1$, we can re-construct g using g_x and g_y as:

- Using g_x, we get the first row of $g = \begin{bmatrix} 1 & 3 & 4 & 5 \end{bmatrix}$. There after, using g_y, we can compute the whole g matrix as:

$$g = \begin{bmatrix} 1 & 3 & 4 & 5 \\ 5 & 7 & 8 & 9 \\ 9 & 11 & 12 & 13 \\ 13 & 15 & 16 & 17 \end{bmatrix} \tag{14}$$

- Using g_y, we get the first column of $g = \begin{bmatrix} 1 & 5 & 9 & 13 \end{bmatrix}^T$. There after, using g_x, we can compute the whole g matrix as:

$$g = \begin{bmatrix} 1 & 3 & 4 & 5 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \tag{15}$$

We see that this gives us different g matrix, and are hence not integratable.

1.i Shape estimation

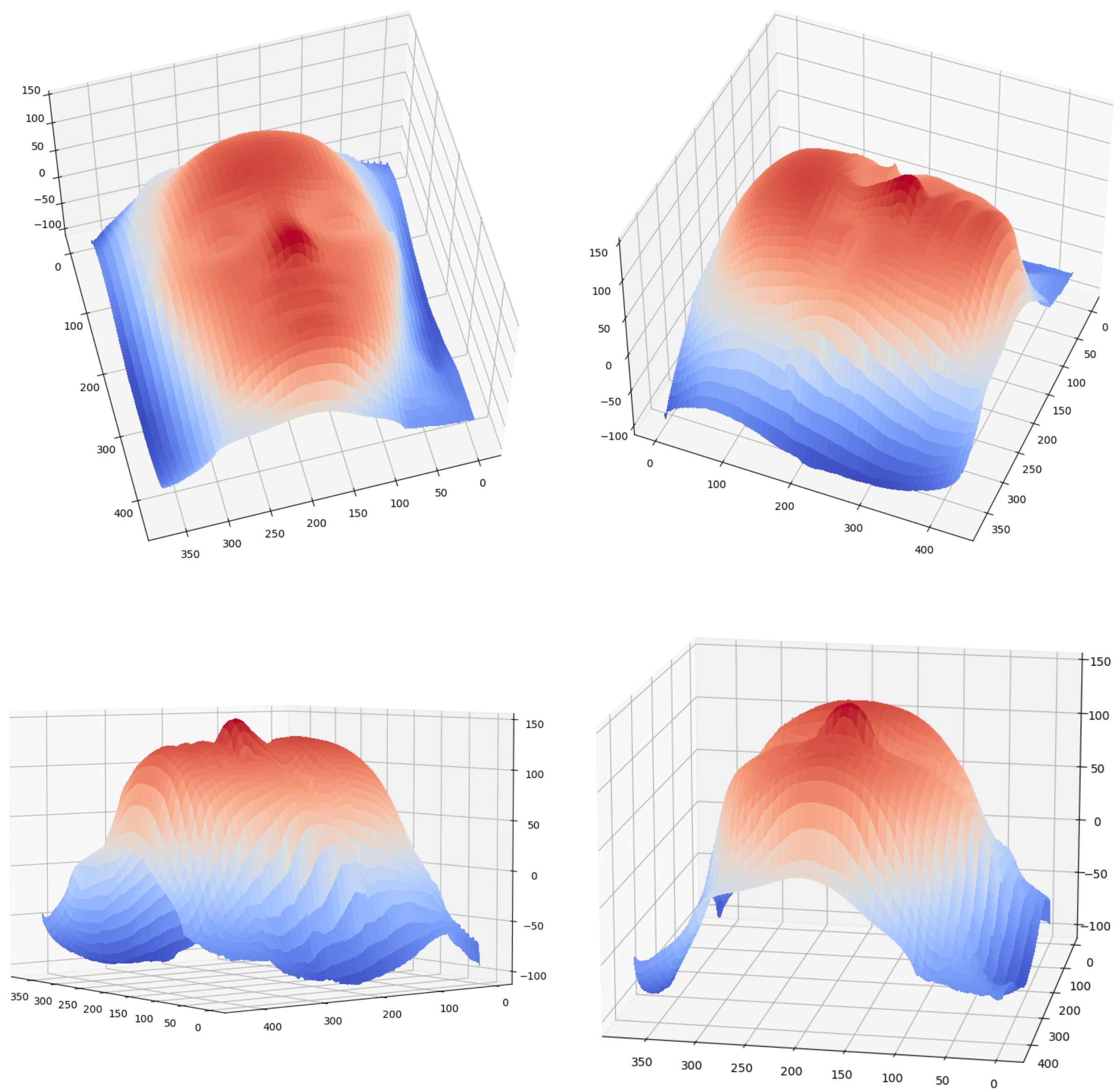

Figure 4: Shape Visualization.

2. Uncalibrated photometric stereo

2.a Uncalibrated normal estimation

The matrix I can be decomposed using SVD. To enforce rank 3 constraint, all the singular values of I, except the top 3 singular values, should be set to 0. There after we can compute $\hat{I}$ as $\hat{I} = U \times \Sigma_{new} \times V^T$. This gives $\hat{I}$ that has rank 3. To get $\hat{L}$ and $\hat{B}$, we have to decompose $\hat{I}$ again using SVD, which will give us $\hat{U}$, $\hat{\Sigma}$ and $\hat{V}^T$. Now, we can create $\hat{L}$ and $\hat{B}$ in the following way:

- $\hat{L} = \hat{U} \times \hat{\Sigma}$ and $B = \hat{V}^T$
- $\hat{L} = \hat{U}$ and $B = \hat{\Sigma} \times \hat{V}^T$

There can be other ways to combine $\hat{U}$, $\hat{\Sigma}$ and $\hat{V}^T$. Two possible ways are shown above.

2.b Calculation and visualization

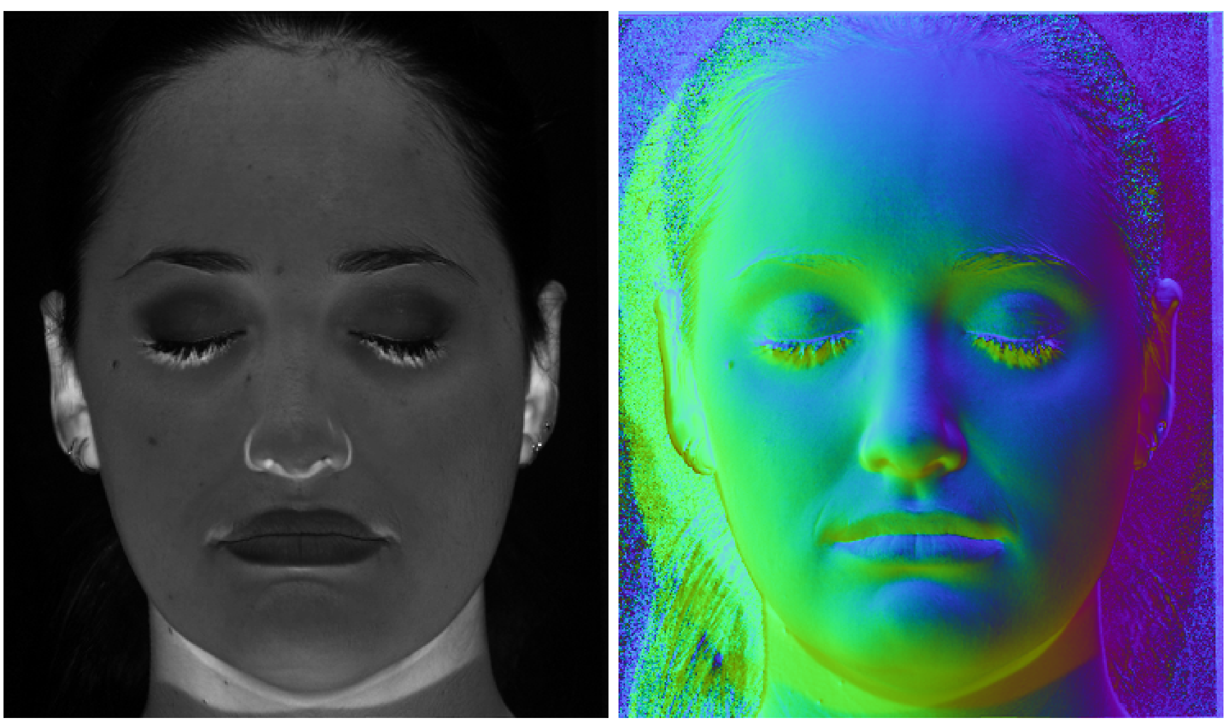

Figure 5: Visualization of albedos and normals.

2.c Comparing to ground truth lighting

Ground truth L is:

```
L:
 [[-0.1418  0.1215 -0.069   0.067  -0.1627  0.      0.1478]
 [-0.1804 -0.2026 -0.0345 -0.0402  0.122   0.1194  0.1209]
 [-0.9267 -0.9717 -0.838  -0.9772 -0.979  -0.9648 -0.9713]]
```

Figure 6: Ground truth L.

Estimated L_{new} is:

```
L_new:
 [[-4.10576404e-07 -4.61021422e-07 -3.75677696e-07 -4.64604653e-07
  -4.54995684e-07 -4.44670353e-07 -4.39982146e-07]
 [ 4.89027863e-08 -8.71305675e-08  2.47399274e-08 -2.71761002e-08
   7.80132455e-08  1.21313840e-08 -3.97005315e-08]
 [ 6.09891951e-08  3.85347255e-08  1.02825952e-08 -1.33677778e-09
  -2.00906083e-08 -3.21223776e-08 -5.14178086e-08]]
```

Figure 7: Estimated L.

The values of L and L_{new} are different. As, in (a), $\hat{L}$ and $\hat{B}$ can be changed by combining $\hat{\Sigma}$ to either $\hat{U}$ or $\hat{V}^T$. Also, $\hat{\Sigma}$ can be decomposed by taking square root and combining them to $\hat{U}$ and $\hat{V}^T$.

2.d Reconstructing the shape, attempt 1

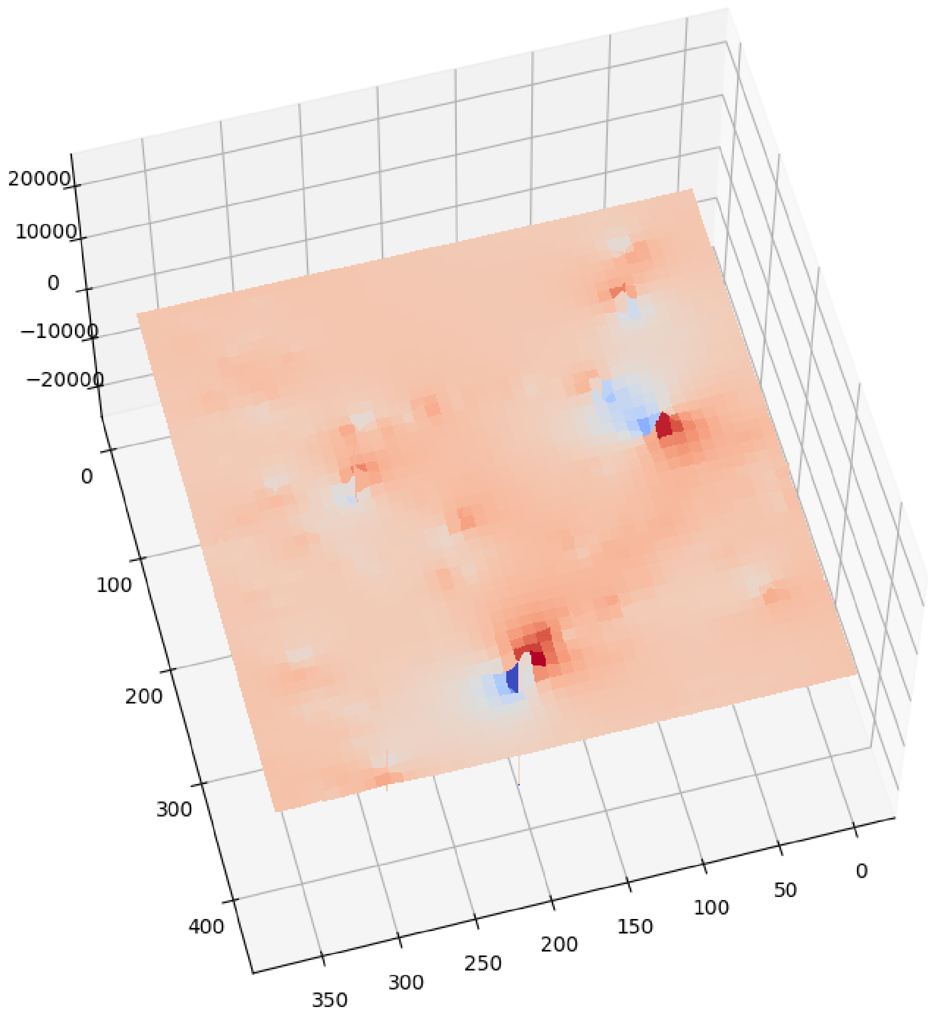

Figure 8: Attempt 1

2.e Reconstructing the shape, attempt 2

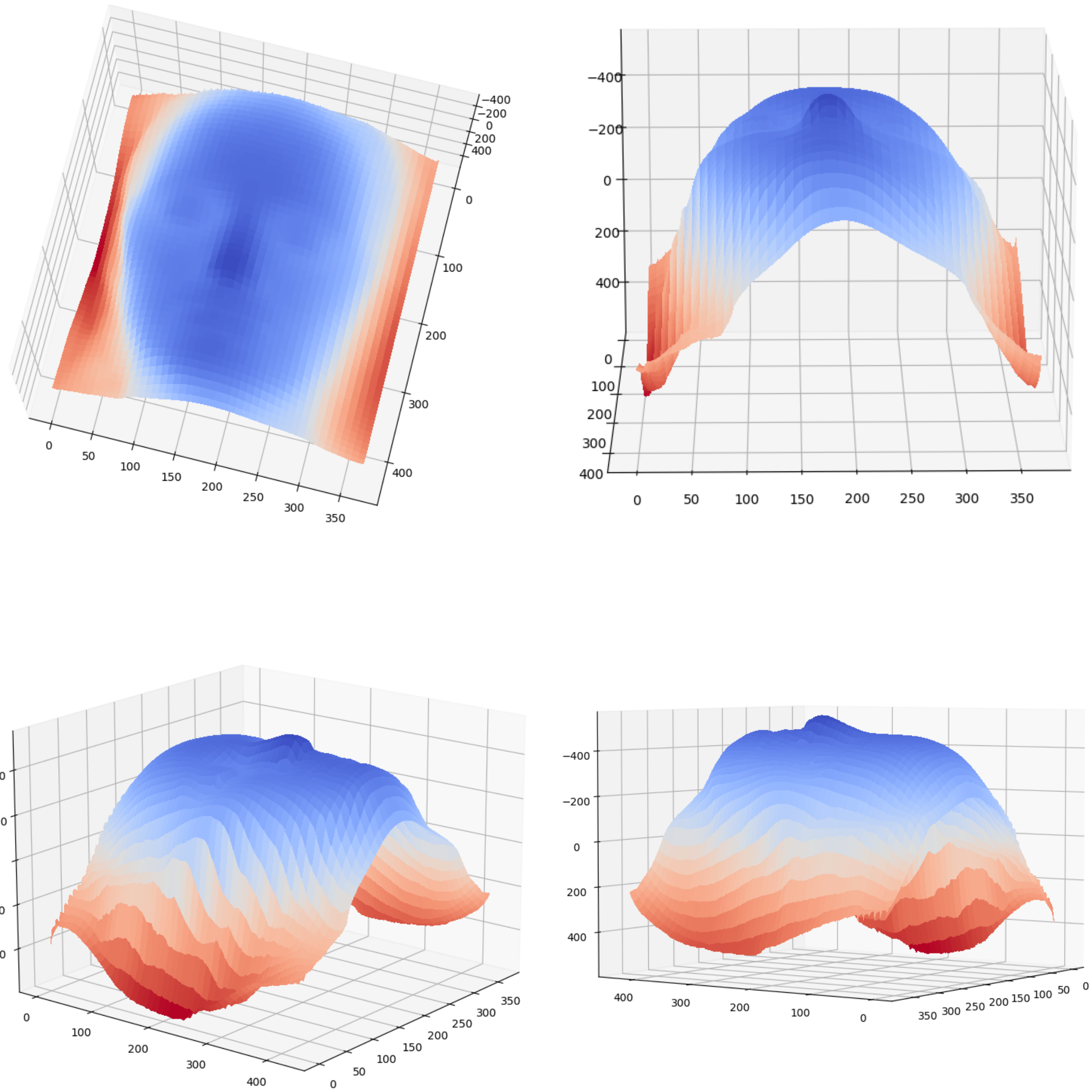

Figure 9: Attempt 2

This reconstruction looks like a face of a person.

2.f Why low relief ?

Varying μ causes the face to flatten and tilt to one side. Varying the ν causes the face to flatten/distort and tilt in another direction. Varying λ cases the scale to change(note the axes), and distorts the face. The bas-relief is so named, because, the 3-d structures have a low elevation from the background. When viewed from the front, the structure looks like a normal 3-dimensional sculpture, but when viewed from the side, it makes no sense.

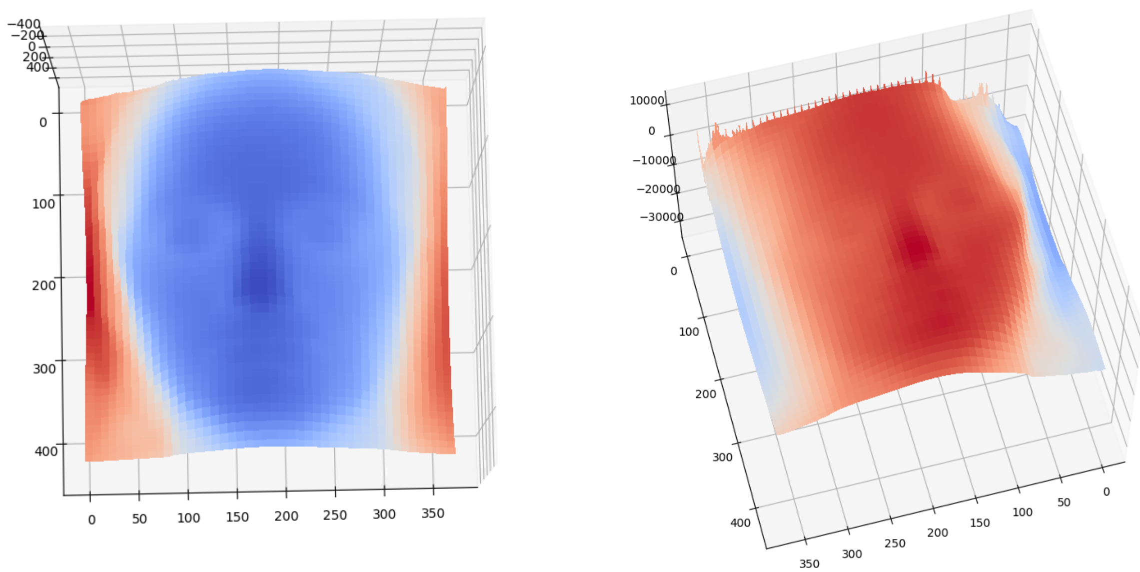

Figure 10: Reconstruction with different μ , ν and λ: $[0, 0, 1]$ *and* $[15, 0, 1]$

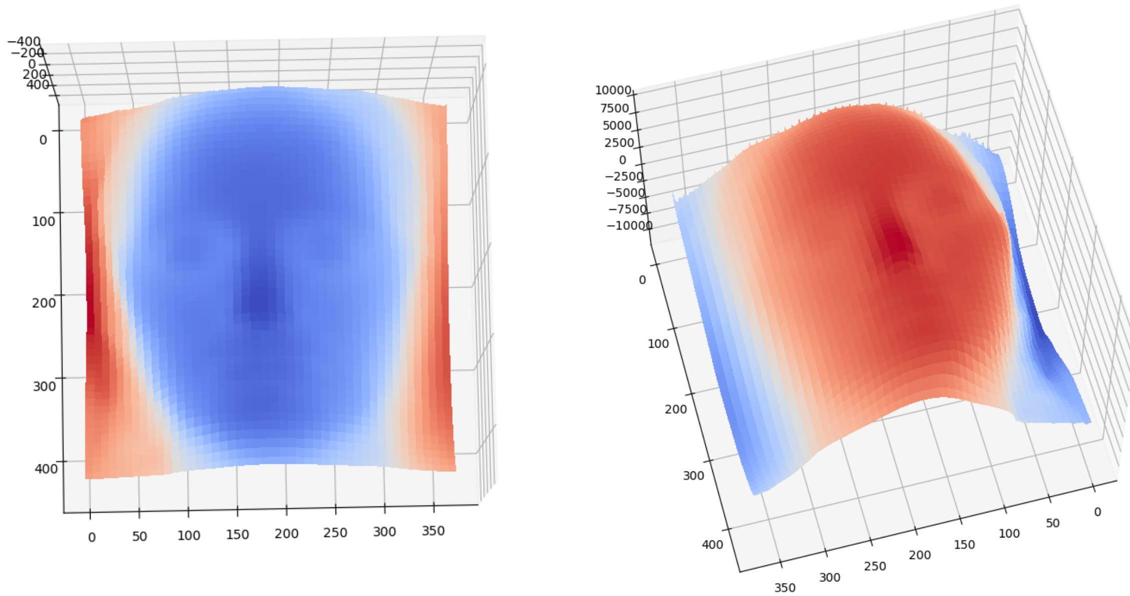

Figure 11: Reconstruction with different μ , ν *and* λ*:* $[0, 0, 1]$ *and* $[0, 15, 1]$

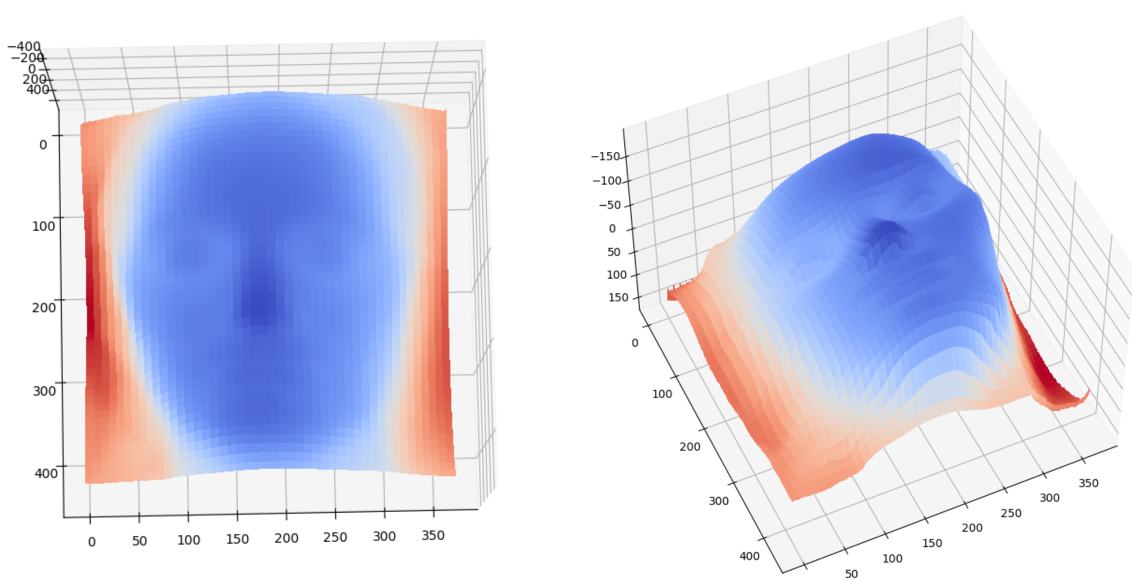

Figure 12: Reconstruction with different μ , ν *and* λ*:* $[0, 0, 1]$ *and* $[0, 0, 15]$

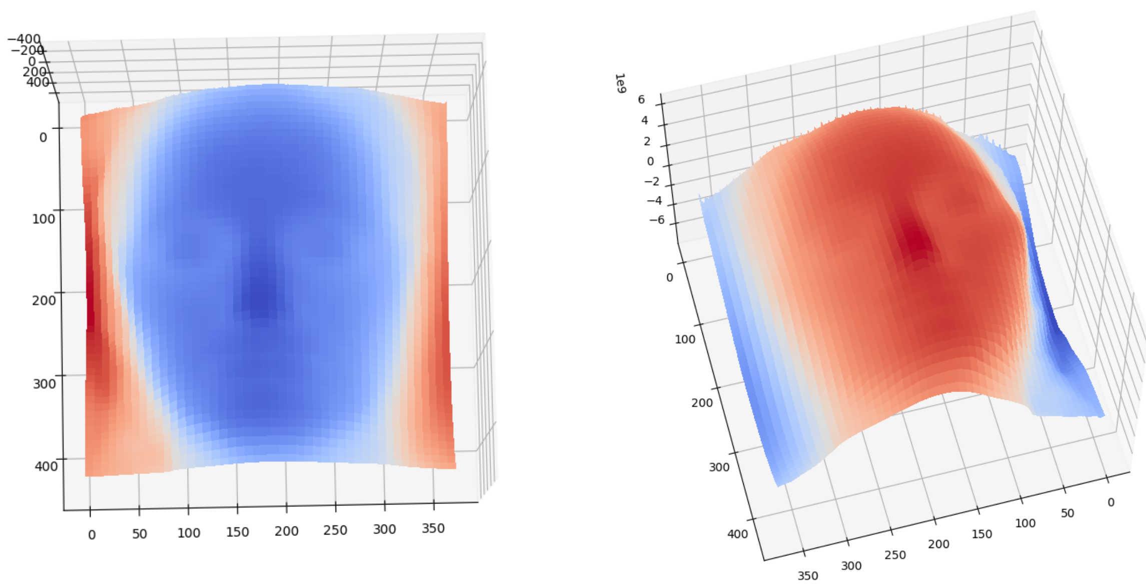

Figure 13: Reconstruction with different μ *,* ν *and* λ*:* $[0, 0, 1]$ *and* $[0, 0, 1e-07]$

2.g Flattest surface possible

To make the surface flat, choose a high value of μ, ν and λ. To make the reconstruction flat, I have set the value of $\mu = 15$ and $\nu = 15$, and $\lambda = 15$. The results are given below:

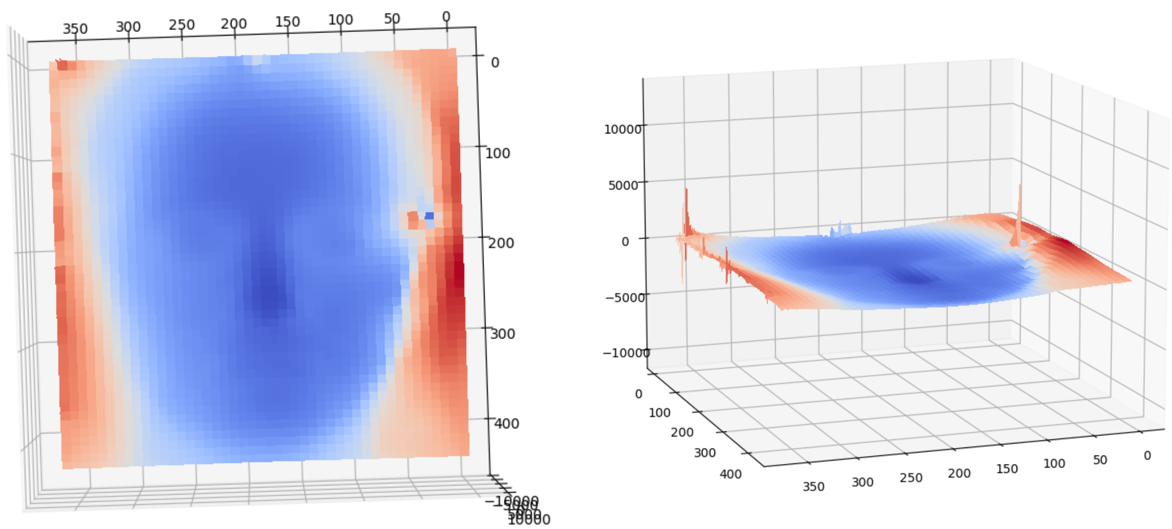

Figure 14: Flat reconstruction of face.

2.h More measurements

More measurements will not help in solving the ambiguity. Unless we know the direction of light source, we cannot predict the actual shape with certainty.