Homework 6

Thomas Zhang (AndrewID: tczhang)

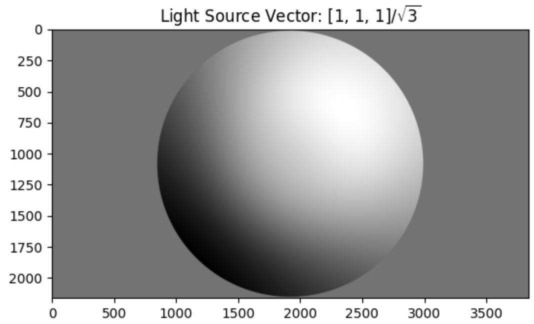
December 2023

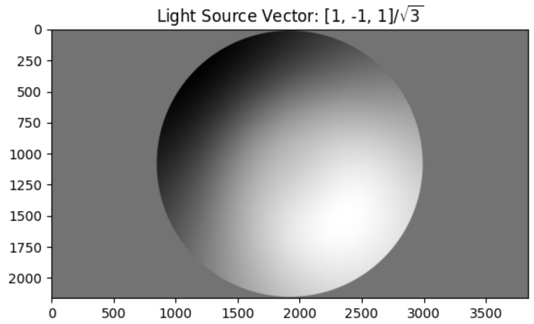
Question 1

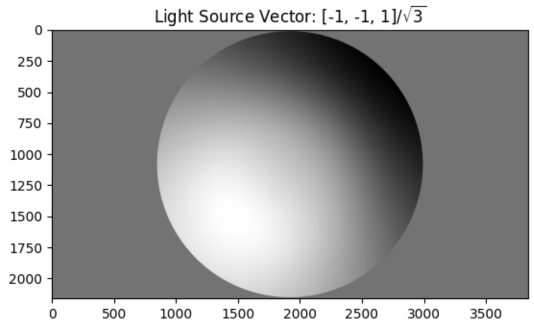
a)

$\overrightarrow{l}$ is the normalized vector representing the direction of the light source. $\overrightarrow{n}$ is the normal vector to the object surface. $\overrightarrow{n} \cdot \overrightarrow{l}$ equivalent to $cos(\theta)$ where θ is the angle between the normal vector $\overrightarrow{n}$ and the light vector $\overrightarrow{l}$. $\overrightarrow{n} \cdot \overrightarrow{l}$ is used when using the Lambertian case of photometric stereo. Under Lambertian conditions $L = \frac{\rho}{\pi} I cos(\theta)$ where L is the surface radiance and I is the light source intensity. $\frac{\rho}{\pi}$ is the albedo constant resulting from the Lambertian BRDF assumption. The viewing direction doesn't matter because the Lambertian assumption is that points appear at the same brightness from all directions (camera position doesn't matter). In other words, the Lambertian assumption ignores specular reflection (which depends on viewing angle), and only considers diffuse reflection (which does not depend on viewing angle).

b)







c)

```python
from skimage import color

all_images = []
for i in range(7):
    im_path = path + f'input_{i+1}.tif'
    rgb_image = cv2.imread(im_path, -1)
    xyz_image = color.rgb2xyz(rgb_image)
    all_images.append(xyz_image[:, :, 1].flatten())

    # print('rgb_image:', rgb_image.shape, rgb_image.dtype)
    # print('xyz_image:', xyz_image.shape, xyz_image.dtype)
    # print(all_images[-1].shape)

I = np.stack(all_images, axis = 0)
L = np.load(path + 'sources.npy').T
s = rgb_image.shape[:-1]
# print('I:', I.shape)
# print('L:', L.shape, L.dtype)
# print('s:', s)

return I, L, s
```

d)

The expected rank of $\mathbf{I}$ is 3 because $\mathbf{L}$ is a 3×7 matrix that ideally has rank 3 and $\mathbf{B}$ is a $3 \times P$ matrix that also has rank 3. Thus $\mathbf{I} = \mathbf{L}^T\mathbf{B}$ should be a $7 \times P$ matrix with rank 3.

The singular values of $\mathbf{I}$ we are given are: 72.40617702, 12.00738171, 8.42621836, 2.23003141, 1.51029184, 1.17968677, 0.84463311. These singular values do not agree with the premise that $\mathbf{I}$ should have rank 3. If $\mathbf{I}$ was rank 3 then there would only be 3 non-zero singular values. The reason why $\mathbf{I}$ is not rank 3 is because the images captured are not in an ideal state. For the rank of $\mathbf{I}$ to be 3, the conditions in each image must be identical except for the moving light source and the surface can't have too much specular reflection. Realistically, this is difficult to achieve, thus the rank of $\mathbf{I}$ is larger than 3.

e)

$$\text{Given: } \mathbf{I} = \mathbf{L}^T\mathbf{B}$$

$$\text{Let } \mathbf{I} = \begin{bmatrix} I_{1,1} & \dots & I_{1,p} \\ \vdots & \ddots & \vdots \\ I_{n,1} & \dots & I_{n,p} \end{bmatrix} \text{ and } \mathbf{L} = \begin{bmatrix} L_{1,1} & \dots & L_{1,n} \\ L_{2,1} & \dots & L_{2,n} \\ L_{3,1} & \dots & L_{3,n} \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} B_{1,1} & \dots & B_{1,p} \\ B_{2,1} & \dots & B_{2,p} \\ B_{3,1} & \dots & B_{3,p} \end{bmatrix}$$

Then $\mathbf{I} = \mathbf{L}^T\mathbf{B}$ can be rewritten as:

$$\begin{bmatrix} I_{1,1} & \dots & I_{1,p} \\ \vdots & \ddots & \vdots \\ I_{n,1} & \dots & I_{n,p} \end{bmatrix} = \begin{bmatrix} L_{1,1} & L_{2,1} & L_{3,1} \\ \vdots & \vdots & \vdots \\ L_{1,n} & L_{2,n} & L_{3,n} \end{bmatrix} \begin{bmatrix} B_{1,1} & \dots & B_{1,p} \\ B_{2,1} & \dots & B_{2,p} \\ B_{3,1} & \dots & B_{3,p} \end{bmatrix}$$

Which can further be rewritten as:

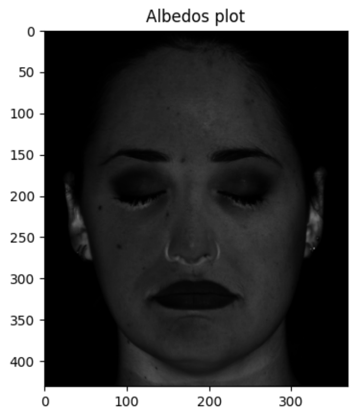
$$\begin{bmatrix} I_{1,1} \\ I_{1,2} \\ \vdots \\ I_{1,p} \\ I_{2,1} \\ I_{2,2} \\ \vdots \\ I_{2,p} \\ I_{n,1} \\ I_{n,2} \\ \vdots \\ I_{n,p} \end{bmatrix} = \begin{bmatrix} L_{1,1} & L_{2,1} & L_{3,1} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & L_{1,1} & L_{2,1} & L_{3,1} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & L_{1,1} & L_{2,1} & L_{3,1} \\ L_{1,2} & L_{2,2} & L_{3,2} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & L_{1,2} & L_{2,2} & L_{3,2} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & L_{1,2} & L_{2,2} & L_{3,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ L_{1,n} & L_{2,n} & L_{3,n} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & L_{1,n} & L_{2,n} & L_{3,n} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & L_{1,n} & L_{2,n} & L_{3,n} \end{bmatrix} \begin{bmatrix} B_{1,1} \\ B_{2,1} \\ B_{3,1} \\ B_{1,2} \\ B_{2,2} \\ B_{3,2} \\ \vdots \\ B_{1,p} \\ B_{2,p} \\ B_{3,p} \end{bmatrix}$$

So if we want to write $\mathbf{I} = \mathbf{L}^T\mathbf{B}$ in the form of $\mathbf{Ax} = \mathbf{y}$, we let $\mathbf{A}$ be a $n * p \times 3p$ matrix, x be a $3p$ long vector and y be a $n * p$ long vector. The forms of $\mathbf{A}$, x and y are written below.

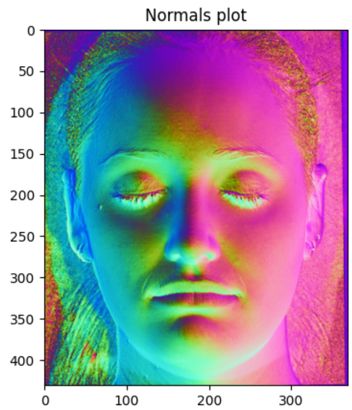
$$\mathbf{A} = \begin{bmatrix} L_{1,1} & L_{2,1} & L_{3,1} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & L_{1,1} & L_{2,1} & L_{3,1} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & L_{1,1} & L_{2,1} & L_{3,1} \\ L_{1,2} & L_{2,2} & L_{3,2} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & L_{1,2} & L_{2,2} & L_{3,2} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & L_{1,2} & L_{2,2} & L_{3,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ L_{1,n} & L_{2,n} & L_{3,n} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & L_{1,n} & L_{2,n} & L_{3,n} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & L_{1,n} & L_{2,n} & L_{3,n} \end{bmatrix}$$

$$x = \begin{bmatrix} B_{1,1} \\ B_{2,1} \\ B_{3,1} \\ B_{1,2} \\ B_{2,2} \\ B_{3,2} \\ \vdots \\ B_{1,p} \\ B_{2,p} \\ B_{3,p} \end{bmatrix} \qquad y = \begin{bmatrix} I_{1,1} \\ I_{1,2} \\ \vdots \\ I_{1,p} \\ I_{2,1} \\ I_{2,2} \\ \vdots \\ I_{2,p} \\ I_{n,1} \\ I_{n,2} \\ \vdots \\ I_{n,p} \end{bmatrix}$$

f)



For the most part, the albedo constant is the same accross the entire face in the image. However there are bright spots around the nose and ears of the face in the picture. This is not match the assumption that the albedo is constant throughout the picture. This is caused by the reflections of light in the picture at areas where there are corners.



The gradients for the most part match the expected curvature of a face.

g)

Given the 3D depth map $z = f(x, y)$. We can also represent the 3D depth map as a function of x, y and z in the form $F(x, y, z) = z - f(x, y) = 0$. The normal to a point on a 3D surface is the gradient of $F(x, y, z) = 0$.

$$\begin{aligned}\nabla F(x, y, z) &= (\frac{dF}{dx}, \frac{dF}{dy}, \frac{dF}{dz}) \\ &= (\frac{d}{dx}z - f(x, y), \frac{d}{dy}z - f(x, y), \frac{d}{dz}z - f(x, y)) \\ &= (-\frac{df}{dx}, -\frac{df}{dy}, 1)\end{aligned}$$

Given that (n_1, n_2, n_3) is the normal to the surface at point (x, y). Then the gradient at (x, y) should be some scalar multiple of (n_1, n_2, n_3). If we scale the normal vector by $\frac{1}{n_3}$, then the normal vector becomes $(\frac{n_1}{n_3}, \frac{n_2}{n_3}, 1)$. This matches the form of the gradient. Hence:

$$-\frac{df}{dx} = \frac{n_1}{n_3} \qquad -\frac{df}{dy} = \frac{n_2}{n_3}$$

h)

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

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

Reconstruct the first row from g_x.

$$g = \begin{bmatrix} 1 & 2 & 3 & 4 \\ g_{2,1} & g_{2,2} & g_{2,3} & g_{2,4} \\ g_{3,1} & g_{3,2} & g_{3,3} & g_{3,4} \\ g_{4,1} & g_{4,2} & g_{4,3} & g_{4,4} \end{bmatrix}$$

Reconstruct the rest of g using g_y.

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

Reconstruct the first column from g_y.

$$g = \begin{bmatrix} 1 & g_{1,2} & g_{1,3} & g_{1,4} \\ 5 & g_{2,2} & g_{2,3} & g_{2,4} \\ 9 & g_{3,2} & g_{3,3} & g_{3,4} \\ 13 & g_{3,2} & g_{3,3} & g_{3,4} \end{bmatrix}$$

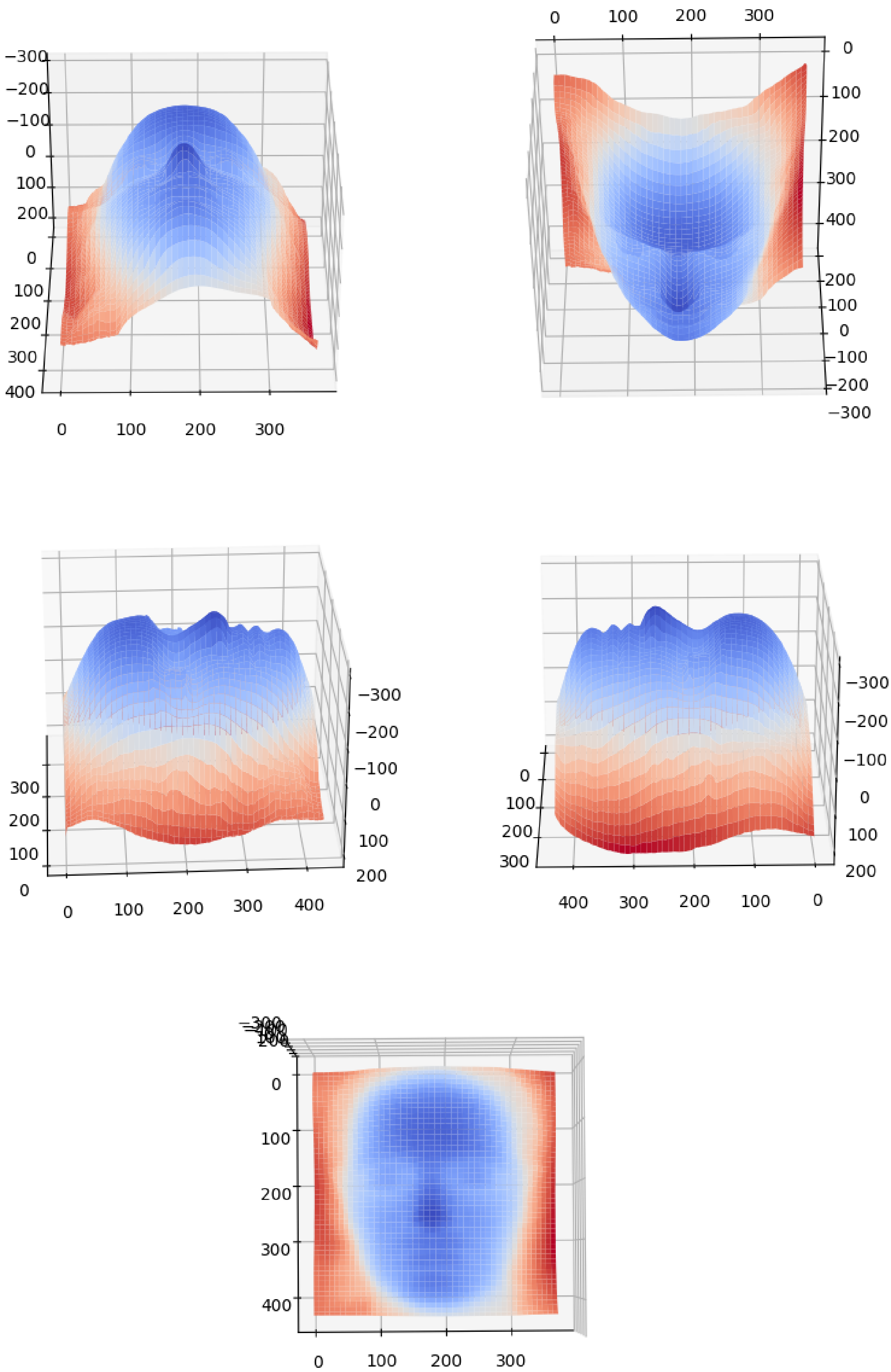
Reconstruct the rest of g using g_x.

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

To make g_x and g_y non-integratable, we would need to add some function $f(x, y)$ to g_x and g_y where $f(x, y)$ is some function that is not fully dependent on g.

The gradients calculated in (g) may be non-integratable if the normals calculated are inaccurate. Given that the normals were calculated using the Lambertian assumption, this may happen because there is a non-insignificant amount specular reflection causing the brightness to be higher or lower than it would be if the surface were an ideal Lambertian surface.

i)



−300
−200
−100
0
100
200
0
100
200
300
400
0
100
200
300
0
100
200
300
0
100
200
300
400
200
100
0
−100
−200
−300
−300
−200
−100
0
100
200
300
200
100
0
0
100
200
300
400
−300
−200
−100
0
100
200
0
100
200
300
400
300
200
100
0
0
100
200
300
400
0
100
200
300

Question 2

a)

Let $\mathbf{I} = \mathbf{L}^T\mathbf{B}$ be approximated by $\hat{\mathbf{I}} = \hat{\mathbf{L}}^T\hat{\mathbf{B}}$. Where rank($\hat{\mathbf{I}}$)=3.
Using SVD we can approximate the rank-3 approximation of $\mathbf{I}$.

$$\mathbf{I} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

where $\mathbf{U}$ is a $n \times n$ matrix, Σ is a $n \times p$ matrix, and $\mathbf{V}$ is a $p \times p$ matrix. n is the number of images = number of light source locations. p is the number of pixels in an image.
Let $\mathbf{U}_3$ be the first 3 columns of $\mathbf{U}$, $\mathbf{\Sigma}_3$ be the upper left 3×3 diagonal matrix of the 3 largest singular values, and $\mathbf{V}_3$ be the first 3 columns of $\mathbf{V}$. Then:

$$\hat{\mathbf{I}} = \mathbf{U}_3\mathbf{\Sigma}_3\mathbf{V}_3^T$$

In the expression $\hat{\mathbf{I}} = \hat{\mathbf{L}}^T\hat{\mathbf{B}}$, $\hat{\mathbf{I}}$ is a $n \times p$ matrix. $\hat{\mathbf{L}}$ is a $3 \times n$ matrix. And $\hat{\mathbf{B}}$ is a $3 \times p$ matrix.
There are multiple ways to get $\hat{\mathbf{I}} = \mathbf{U}_3\mathbf{\Sigma}_3\mathbf{V}_3^T$ in the form of $\hat{\mathbf{I}} = \hat{\mathbf{L}}^T\hat{\mathbf{B}}$.
One way of doing so is letting

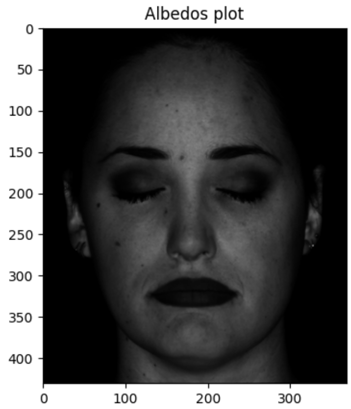
$$\hat{\mathbf{L}}^T = \mathbf{U}_3\mathbf{\Sigma}_3 \qquad \hat{\mathbf{B}} = \mathbf{V}_3^T$$

$\mathbf{U}_3\mathbf{\Sigma}_3$ is a $n \times 3$ matrix times a 3×3 matrix resulting in a $n \times 3$ matrix and $\mathbf{V}_3^T$ is a $p \times 3$ matrix. Thus:

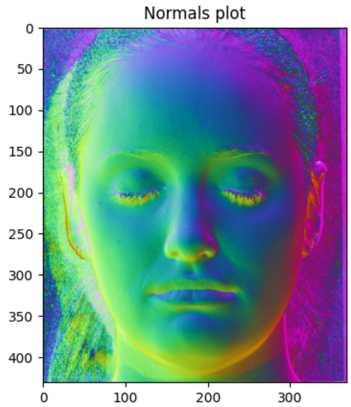
$$\hat{\mathbf{L}}^T = \mathbf{U}_3\mathbf{\Sigma}_3 \qquad \hat{\mathbf{B}} = \mathbf{V}_3^T$$

can be used to represent the equation $\hat{\mathbf{I}} = \hat{\mathbf{L}}^T\hat{\mathbf{B}}$.

b)



There appears to be less bright spots compared to the albedo image generated in part 1.



The magnitude of the normals appear to be different from the normal image generated in part 1. However the face is still recognizable and the vectors still appear to be pointing in directions that make sense.

c)

$$L_0 = \begin{bmatrix} -0.1418 & 0.1215 & -0.069 & 0.067 & -0.1627 & 0. & 0.1478 \\ -0.1804 & -0.2026 & -0.0345 & -0.0402 & 0.122 & 0.1194 & 0.1209 \\ -0.9267 & -0.9717 & -0.838 & -0.9772 & -0.979 & -0.9648 & -0.9713 \end{bmatrix}$$

$$L = \begin{bmatrix} -1838.6756 & -2487.4061 & -1426.11 & -2381.8964 & -2261.4064 & -2117.235 & -2090.765 \\ 580.5752 & -1519.3042 & 306.662 & -389.8274 & 1547.8414 & 329.1821 & -475.6262 \\ 1220.9793 & 640.31 & 287.8695 & -15.0264 & -152.5687 & -584.3694 & -1257.9945 \end{bmatrix}$$

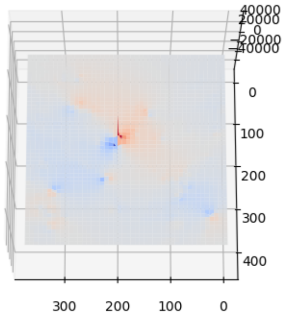
The $\mathbf{L}$ and $\mathbf{L}_0$ matrices are not the same.
Another way of factorizing I would be to let

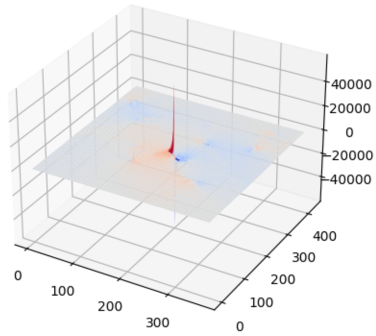
$$\hat{\mathbf{I}} = \mathbf{U}_3\mathbf{\Sigma}_3^{1/2}\mathbf{\Sigma}_3^{1/2}\mathbf{V}_3^T$$
$$\hat{\mathbf{L}}^T = \mathbf{U}_3\mathbf{\Sigma}_3^{1/2} \qquad \hat{\mathbf{B}} = \mathbf{\Sigma}_3^{1/2}\mathbf{V}_3^T$$

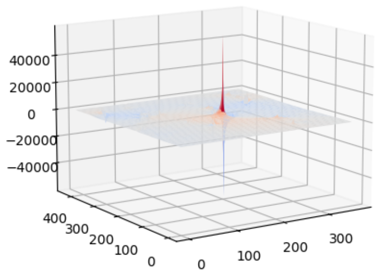
Where $\mathbf{\Sigma}_3^{1/2}$ is a 3×3 diagonal matrix where the values on the diagonal are the square root of the diagonal in $\mathbf{\Sigma}_3$. These still satisfy the dimension constraints of $\hat{\mathbf{L}}^T$ and $\hat{\mathbf{B}}$ in $\hat{\mathbf{I}} = \hat{\mathbf{L}}^T\hat{\mathbf{B}}$.

d)

The image does not look like a reconstructed face.

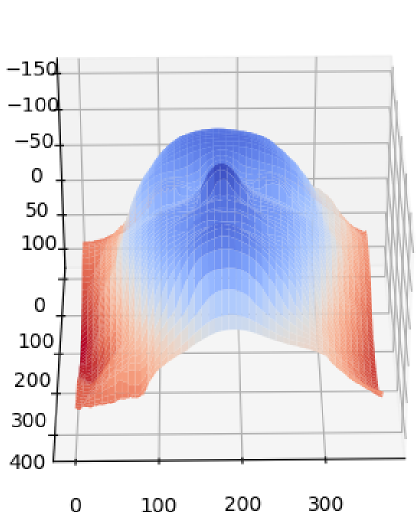


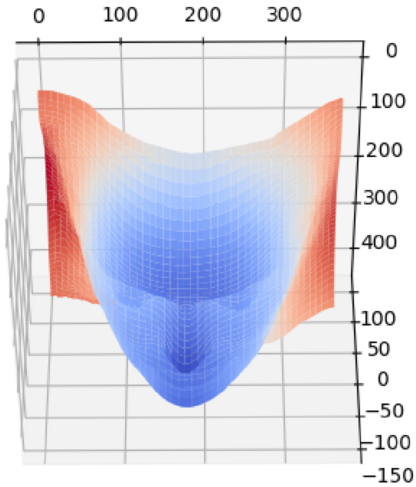


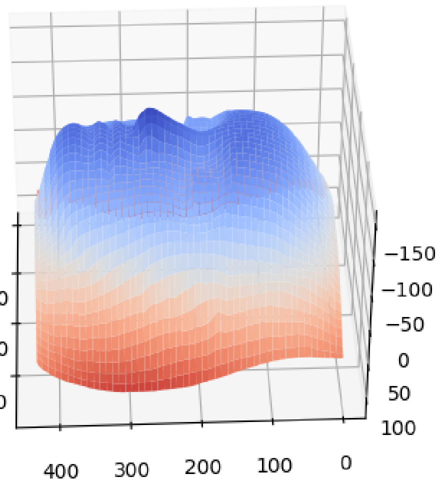


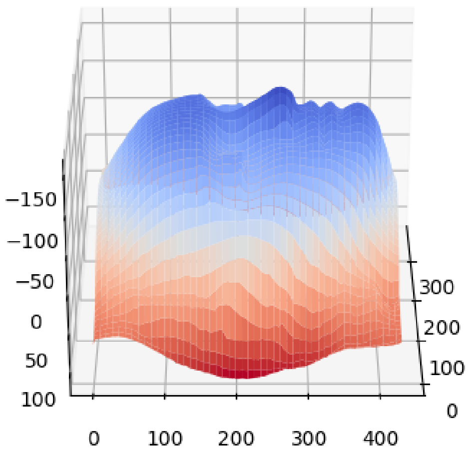
e)

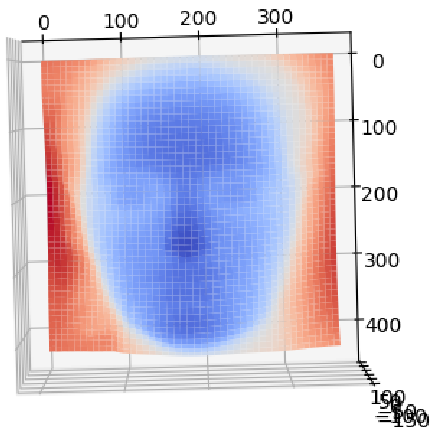
The surface looks very similar to the output created via calibrated photometric stereo. The shape is very similar, the biggest difference is that the scale of the surface in the calibrated photometric stereo is roughly 2 times that of the uncalibrated photometric stereo.





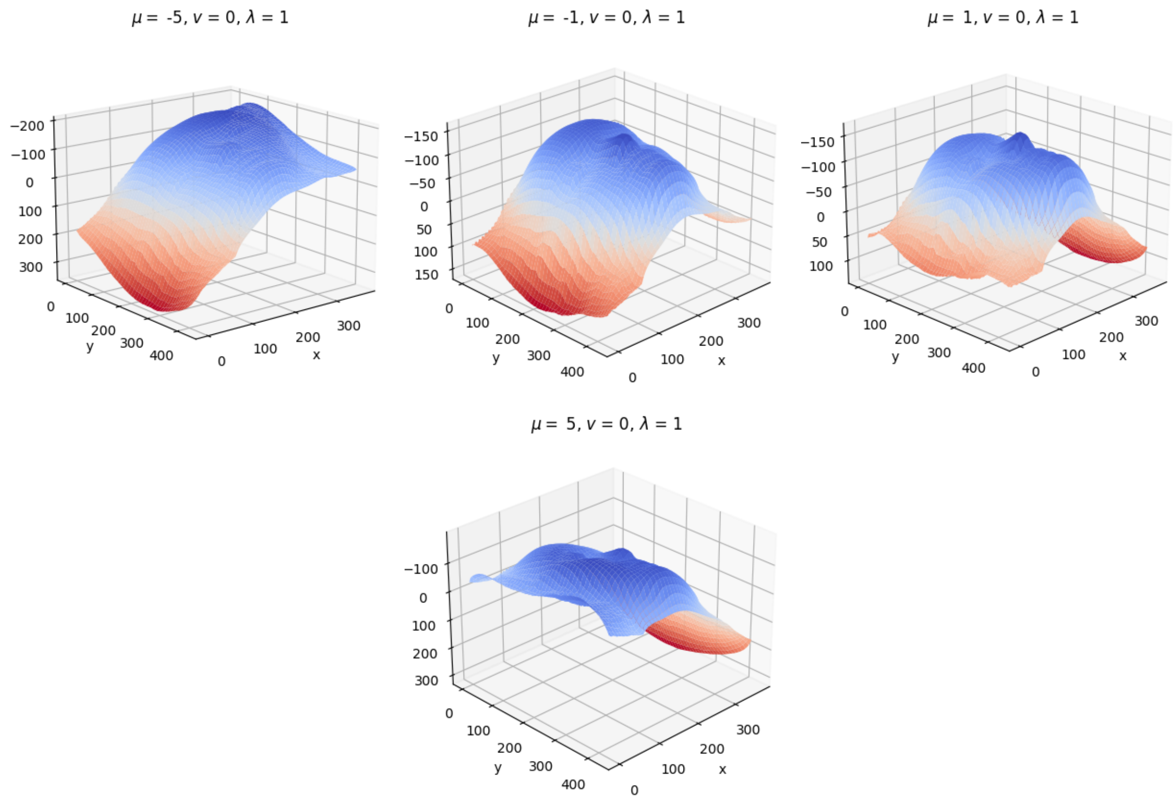




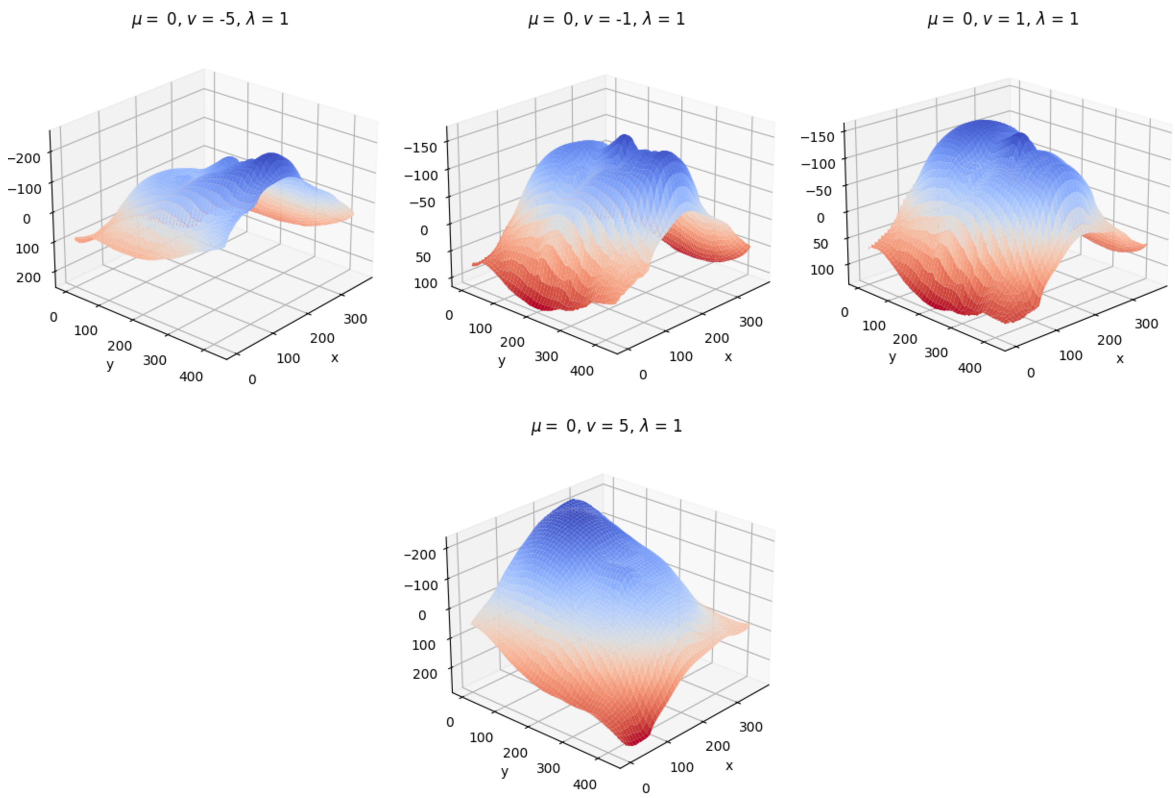


f)

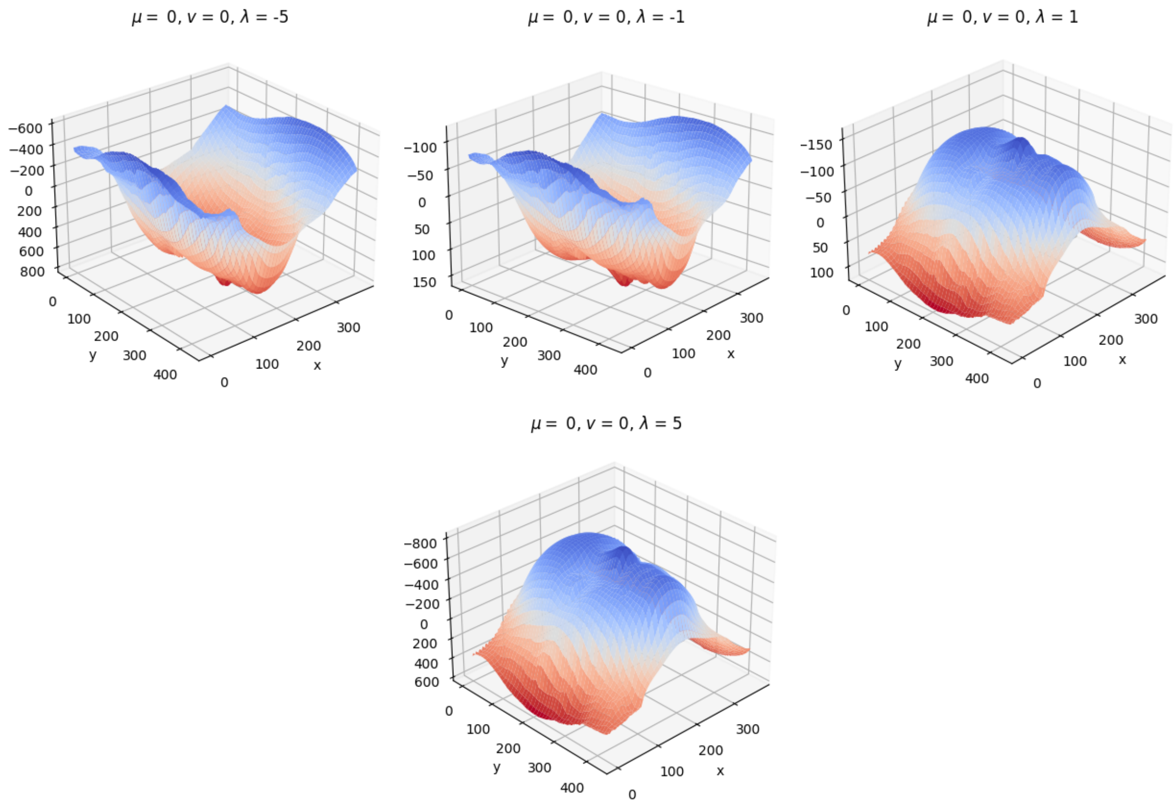
Modifying u in the G matrix causes the surface to "tilt" along the x direction. Negative values of u cause the surface to tilt towards the positive x, while positive values of u cause the surface to tilt towards negative x.



Modifying v in the G matrix causes the surface to "tilt" along the y direction. Negative values of v cause the surface to tilt towards the positive y, while positive values of v cause the surface to tilt towards negative y.



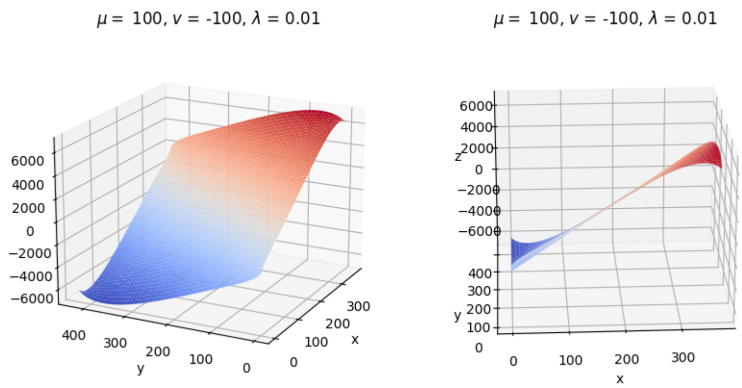
Modifying λ in the G matrix causes the difference in bumps and dips in the surface to be amplified or reduced. The larger the magnitude of λ the larger the amplification. Negative λ values cause the surface to invert.



The observations above support that multiplying $\mathbf{B}$ by $\mathbf{G}^{-T}$ is the same thing as scaling the surface by λ then adding the plane $z = ux + vy$ to the surface. The bas relief ambiguity is likely named so because the surface begins to look similar to bas reliefs (a sculptural relief where the projection from the surface is slight and no part of the model is undercut) when u and v increase in magnitude.

g)

To make the surface as flat as possible we can lower the magnitude of λ to almost 0. Then make u a really large positive number and v a really large negative number. Alternatively, make the magnitude of λ small, u a large negative number and v a large positive number.



h)

No. The ambiguity arises when estimating the direction of the light sources and the surface normal from the rank 3 approximation of $\mathbf{I}$. The ambiguity arises because there are multiple possible light directions + surface normal combinations that can produce a particular pixel intensity on a Lambertian surface. The transformation of the light can be represented by a 3×3 matrix. With surface integrability constraints the 9 degrees of ambiguity of the 3×3 matrix, drops to 3 degrees of ambiguity. Mathematically, the ambiguity is arises when factoring $\hat{\mathbf{I}}$ into $\hat{\mathbf{L}}^T$ and $\hat{\mathbf{B}}$. Regardless of how many images we take, this ambiguity will still exist because there will always be multiple ways to represent $\hat{\mathbf{L}}^T$ and $\hat{\mathbf{B}}$ as expressions of U_3, Σ_3 and V_3 where $\mathbf{I}_3 = U_3 \Sigma_3 V_3^T$ is the SVD of the rank 3 approximation of $\mathbf{I}$.