16-720 Computer Vision: Photometric Stereo

Yanjia Duan

November 2020

1 Calibrated photometric stereo

Q1 (a)

The equation for surface radiance is

$$I = \frac{\rho_d}{\pi} K cos\theta = \frac{\rho_d}{\pi} K \overrightarrow{n} \cdot \overrightarrow{l}$$

where I is the brightness viewed from camera, ρ_d is the absorption rate (i.e. albedo), K is the source intensity, $\overrightarrow{n}$ is the surface normal vector and $\overrightarrow{l}$ is the incident direction.

The dot product comes from the cosine value of the angle between the source light and the surface normal. The $cos\theta$ is related to the surface area that receives the light. The larger the area is, the smaller the $cos\theta$ is, and since the energy of the light keeps the same, the brightness on the surface area will be lower. Therefore, we use $cos\theta$ to represent the projected area.

The viewing direction doesn not matter because the object has body reflection. The light is reflected to everywhere, so wherever you view the object, you receive the same amount of light. Therefore the viewing direction does not matter.

Q1 (b)

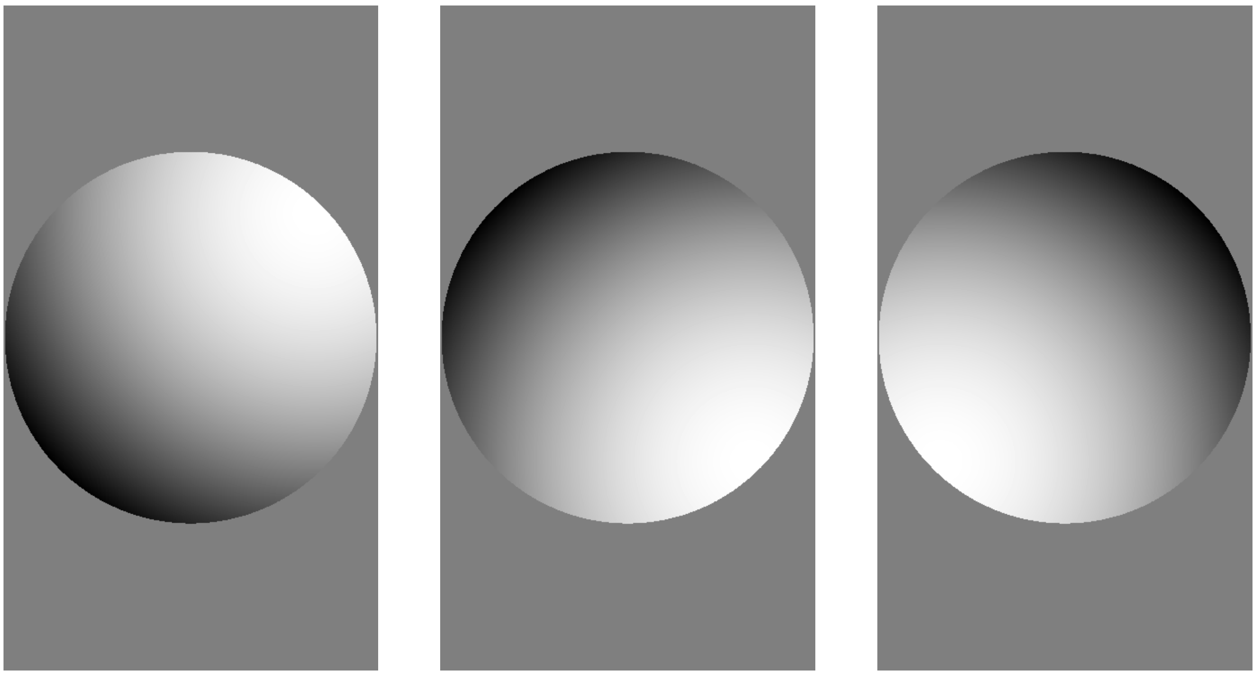

Figure 1: n-dot-l lighting for a sphere under incoming light directions $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$, $(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$, $(-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$.

Q1 (c)

```python
def loadData(path = "../data/"):
    I = None
    L = None
    s = None

    for i in range(1, 8):
        img = imread(path + 'input_' + str(i) + '.tif', dtype=np.uint16)
        img_xyz = rgb2xyz(img)
        img_luminance = img_xyz[:, :, 1]
        if I is None:
            s = img_luminance.shape
            I = img_luminance.flatten().reshape(1, -1)
        else:
            I = np.vstack((I, img_luminance.flatten().reshape(1, -1)))
    L = np.load('../data/sources.npy').T
    return I, L, s
```

Figure 2: Screenshot of function `loadData`.

Q1 (d)

Since $\mathbf{I} = \mathbf{L}^T\mathbf{B}$, $\mathbf{L}$ is a 3×7 matrix and $\mathbf{B}$ is a $3\times P$ matrix, $\mathbf{L}$ and $\mathbf{B}$ have at most 3 ranks. Since $\mathbf{L}$ is a set of seven 3D incoming light directions and $\mathbf{B}$ is a set of 3D pseudonormals, the three ranks correspond to x, y, z coordinates in 3D space, which are independent of one another. Therefore, $\mathbf{L}$ and $\mathbf{B}$ have exactly 3 ranks and their multiplication $\mathbf{I}$ should be 3-rank.

Singular values of $\mathbf{I}$:

```
[79.36348099 13.16260675 9.22148403 2.414729 1.61659626 1.26289066 0.89368302]
```

This results agree with the rank-3 requirement that the first three singular values are much larger than the rests. It shows that singular vectors corresponding to the first three singular values can express most information of $\mathbf{I}$. The rest singular values are very small and their singular vectors barely contain information of $\mathbf{I}$.

Q1 (e)

The general linear equation is $\mathbf{Ax} = \mathbf{y}$. We want to solve $\mathbf{I} = \mathbf{L}^T\mathbf{B}$ to get $\mathbf{B}$. Therefore, $\mathbf{B} = \mathbf{x}$, $\mathbf{A} = \mathbf{L}^T$, $\mathbf{y} = \mathbf{I}$.

Q1 (f)

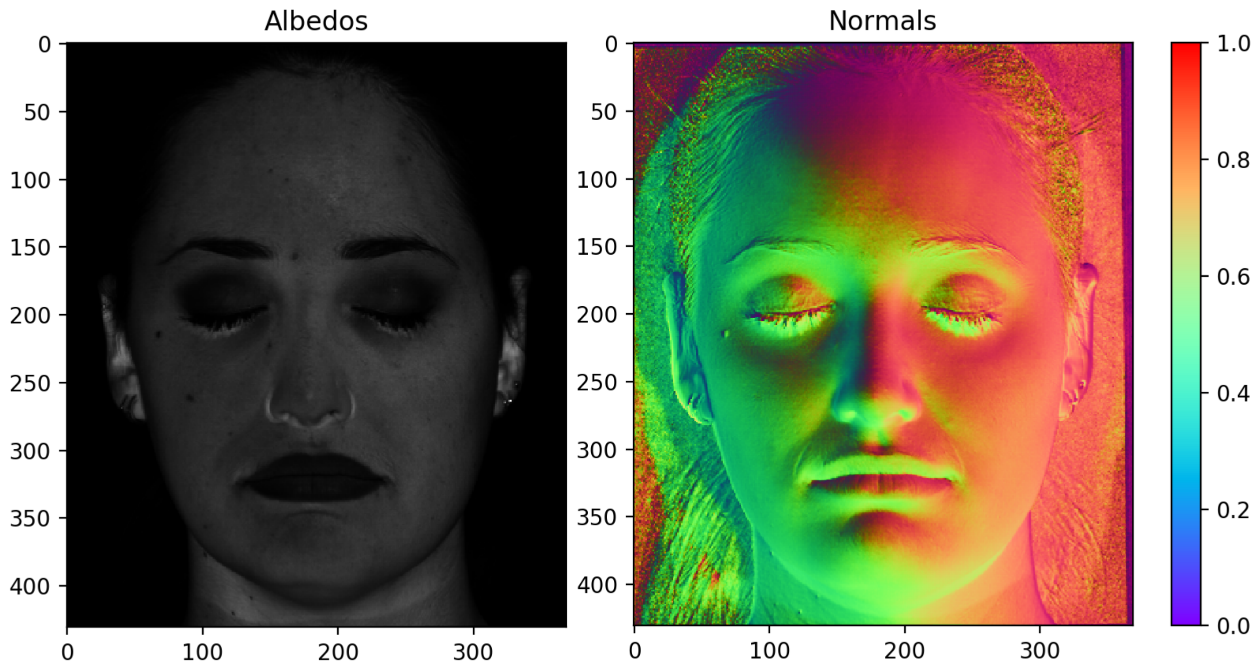

Figure 3: Albedos and normals from calibrated photometric stereo.

Albedo

Theoretically, for a uniform fully reflective Lambertian object, surfaces whose normals in similar directions to the lighting directions will get higher $cos\theta$ (namely, lower albedo values) and vice versa. Under the assumption that "images are taken at lighting directions such that there is not a lot of occlusion or normals being directed opposite to the lighting", most normals should get a high $cos\theta$ and therefore low albedos. However, in the left figure in Figure 3, albedos underneath the nose and around the ears have large values and appear to be brighter. The reason might be that the surface normals underneath the nose and around ears are almost perpendicular to all lighting conditions, which make the $cos\theta$ close to zero. Although the image intensities $\mathbf{I}$ of these shadow areas are small, $cos\theta$ is almost zero therefore albedos could still be very large. So the shadow areas appear to be brighter in the albedo image.

Normals

The normal image in Figure 3 looks reasonable that the bottom-left half of the face has one color and the upper-right half has another color. The reason is that the left half of the face has its normals point left and the upper lip has its normals point down, so they all have green colors. The right half of the face has its normals point right and the lower lip has its normals point up, so they all have red colors.

Q1 (g)

The normal $\mathbf{n}$ at (x, y) is perpendicular to the tangent plane where the *partial derivatives* of f at (x, y) live. So we construct two vectors V_x and V_y that are on the tangent plane.

For f_x, suppose that x increments by Δx, which leads to z incrementing by Δz. V_x is a vector of the increment:

$$\begin{aligned} V_x &= (x + \Delta x, y, z + \Delta z) - (x, y, z) \\ &= (\Delta x, 0, \Delta z) \\ 0 &= \mathbf{n} \cdot V_x \\ &= (n_1, n_2, n_3) \cdot (\Delta x, 0, \Delta z) \\ &= n_1 \Delta x + n_3 \Delta z \end{aligned}$$

$$n_1 + n_3 \frac{\Delta z}{\Delta x} = 0 \Rightarrow n_1 + n_3 f_x = 0 \Rightarrow f_x = -\frac{n_1}{n_3}$$

Similarly for f_y, suppose that y increments by Δy, which leads to z incrementing by Δz. V_y is a vector of the increment:

$$\begin{aligned} V_y &= (x, y + \Delta y, z + \Delta z) - (x, y, z) \\ &= (0, \Delta y, \Delta z) \\ 0 &= \mathbf{n} \cdot V_y \\ &= (n_1, n_2, n_3) \cdot (0, \Delta y, \Delta z) \\ &= n_2 \Delta y + n_3 \Delta z \end{aligned}$$

$$n_2 + n_3 \frac{\Delta z}{\Delta y} = 0 \Rightarrow n_2 + n_3 f_y = 0 \Rightarrow f_y = -\frac{n_1}{n_3}$$

Q1 (h)

For the 2D discrete function $g = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$, its x and y gradients g_x and g_y are

$$g_x = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \qquad g_y = \begin{bmatrix} 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \end{bmatrix}$$

For reconstruction, given $g(0,0) = 1$

1. Using g_x to construct the first row of g will get $g = \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$, then use g_y to construct the rest will get $g = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$.

2. Using g_y to construct the first column of g will get $g = \begin{bmatrix} 1 \\ 5 \\ 9 \\ 13 \end{bmatrix}$, then use g_x to construct the rest will get $g = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$.

The reconstruction results are the same.

If g is discrete, it will always be integrable. If g is continuous and its second parital derivatives are not continuous, namely $g_{xy} \neq g_{yx}$, it will be non-integrable at the point where $g_{xy} \neq g_{yx}$. For example,

$$g = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

where $g_{xy}(0,0) = -1$ and $g_{yx}(0,0) = 1$. If we first use g_x and then use g_y to construct, we will get $\int_y (\int_x g_{yx} dx) dy$. If we first use g_y and then use g_x to construct, we will get $\int_x (\int_y g_{xy} dy) dx$. Since $g_{xy} \neq g_{yx}$, the results will be different. Therefore, g at (0,0) is non-integrable.

Q1 (i)

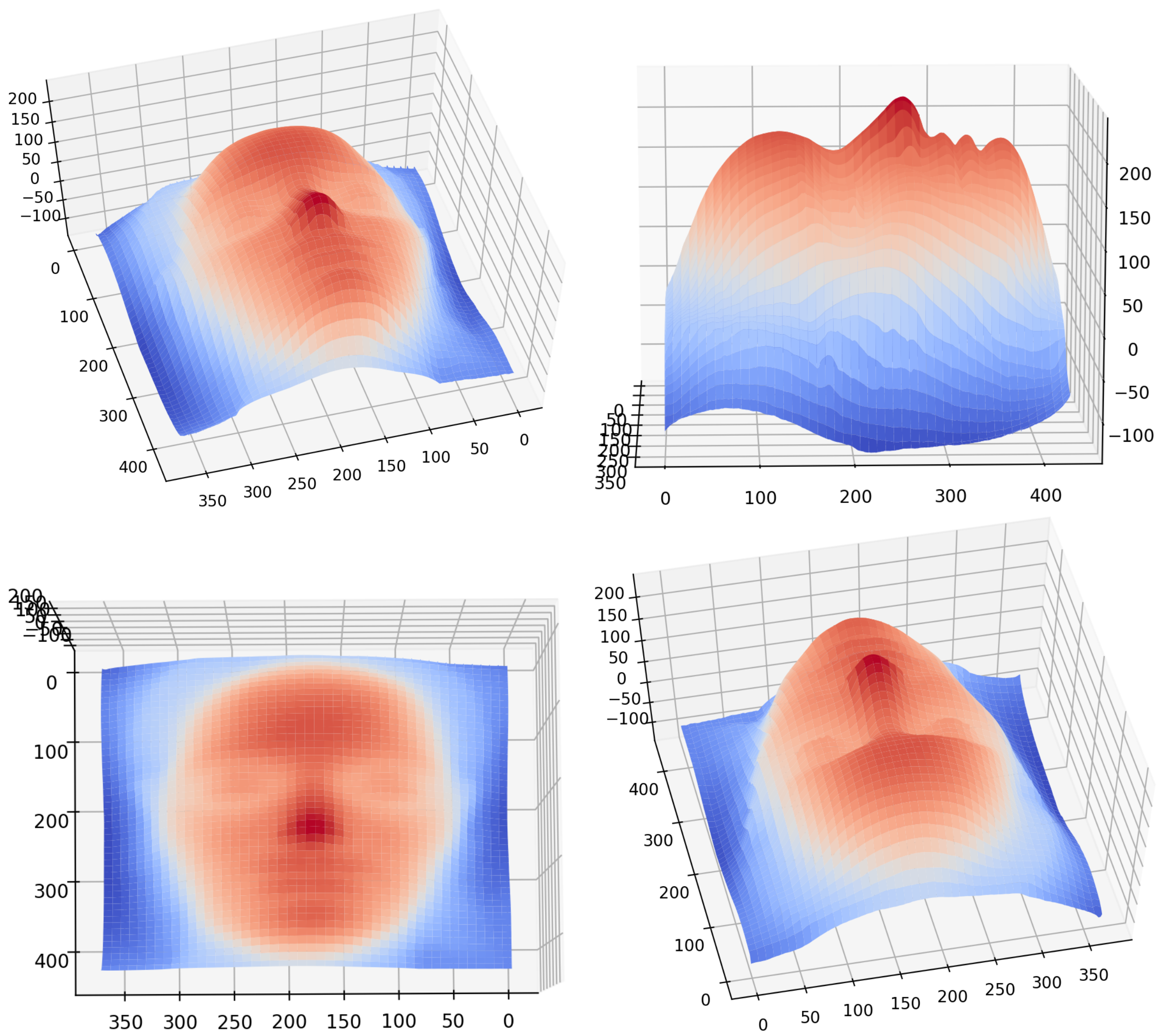

Figure 4: 3D reconstruction result from calibrated photometric stereo.

Q2 (a)

Given $\mathbf{I} = \mathbf{L}^T\mathbf{B}$ and the singular value decomposition of $\mathbf{I}$ is $\mathbf{I}$=$\mathbf{U}\Sigma\mathbf{V}^T$. By estimating $\hat{\mathbf{L}}$ and $\hat{\mathbf{B}}$, since the rank of $\mathbf{I}$ is 3, we keep the top 3 singular values in Σ and set the rests to zeros and get $\hat{\Sigma}$. Therefore,

$$\hat{\mathbf{L}}^T = \mathbf{U}\hat{\Sigma}^{\frac{1}{2}} \Rightarrow \hat{\mathbf{L}} = (\hat{\Sigma}^{\frac{1}{2}})^T\mathbf{U}^T$$

$$\hat{\mathbf{B}} = \hat{\Sigma}^{\frac{1}{2}}\mathbf{V}^T$$

Q2 (b)

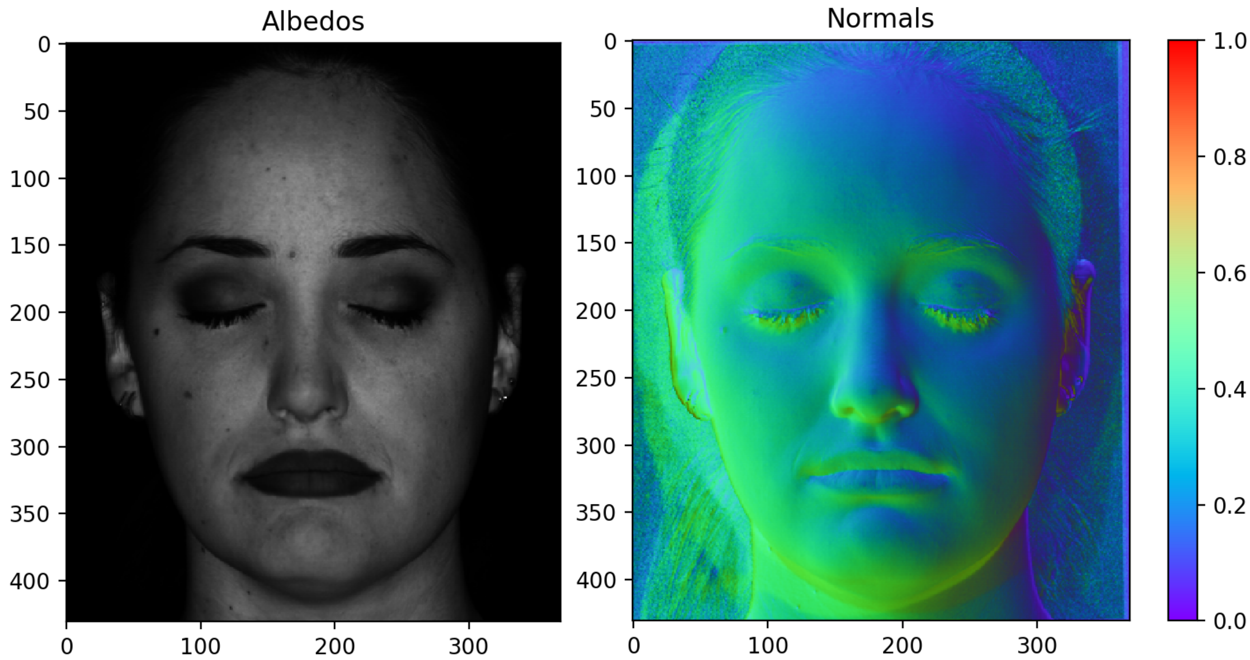

Figure 5: Albedos and normals from uncalibrated photometric stereo.

Q2 (c)

Estimated $\hat{\mathbf{L}}$:

```
[[-2.99267472 -3.86998525 -2.40803005 -3.74500806 -3.59135539 -3.38666635
  -3.3525448 ]
 [ 0.94780484 -2.31708946  0.49911094 -0.62599426  2.32568155  0.46605103
  -0.79271078]
 [ 1.87934697  1.01461663  0.42942606 -0.01730299 -0.3107729  -0.91273581
  -1.8830081 ]]
```

Ground truth $\mathbf{L}_0$:

```
[[-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]]
```

$\hat{\mathbf{L}}$ and $\mathbf{L}_0$ are totally different.
One possible change is to insert $\mathbf{A}^{-1}\mathbf{A}$ to the equation:

$$\hat{\mathbf{I}} = \hat{\mathbf{L}}^T \hat{\mathbf{B}}$$

$$\begin{aligned} \iff \hat{\mathbf{I}} &= \hat{\mathbf{L}}^T \mathbf{A}^{-1}\mathbf{A}\hat{\mathbf{B}} \\ &= (\mathbf{A}^{-T}\hat{\mathbf{L}})^T \mathbf{A}\hat{\mathbf{B}} \end{aligned}$$

which makes $\hat{\mathbf{L}}$ becomes $\mathbf{A}^{-T}\hat{\mathbf{L}}$ and $\hat{\mathbf{B}}$ becomes $\mathbf{A}\hat{\mathbf{B}}$.

Q2 (d)

The result doesn't look like a face as shown below:

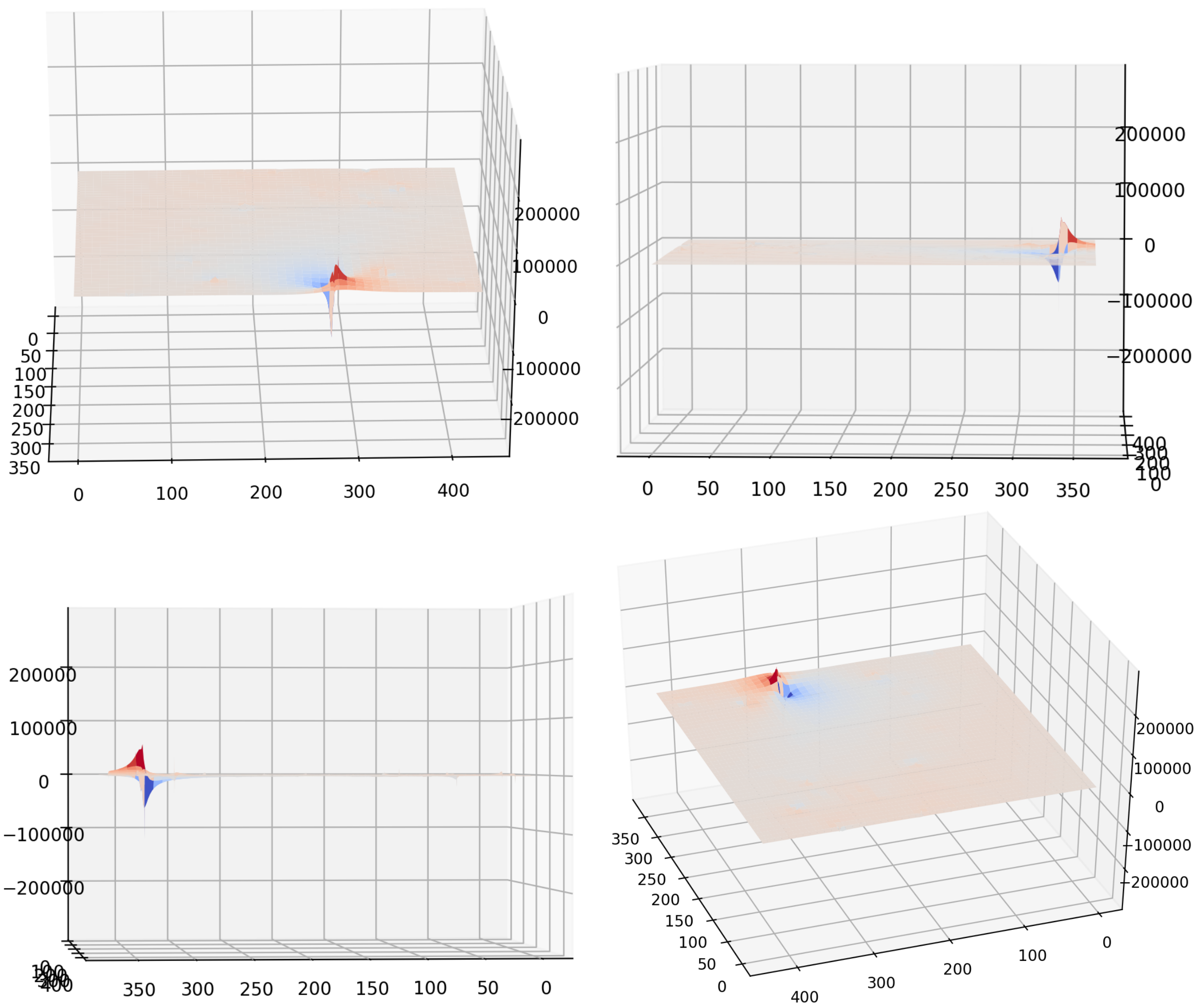

Figure 6: 3D reconstruction result (attempt 1) from uncalibrated photometric stereo.

Q2 (e)

The result looks like the one generated by calibrated photometric stereo.

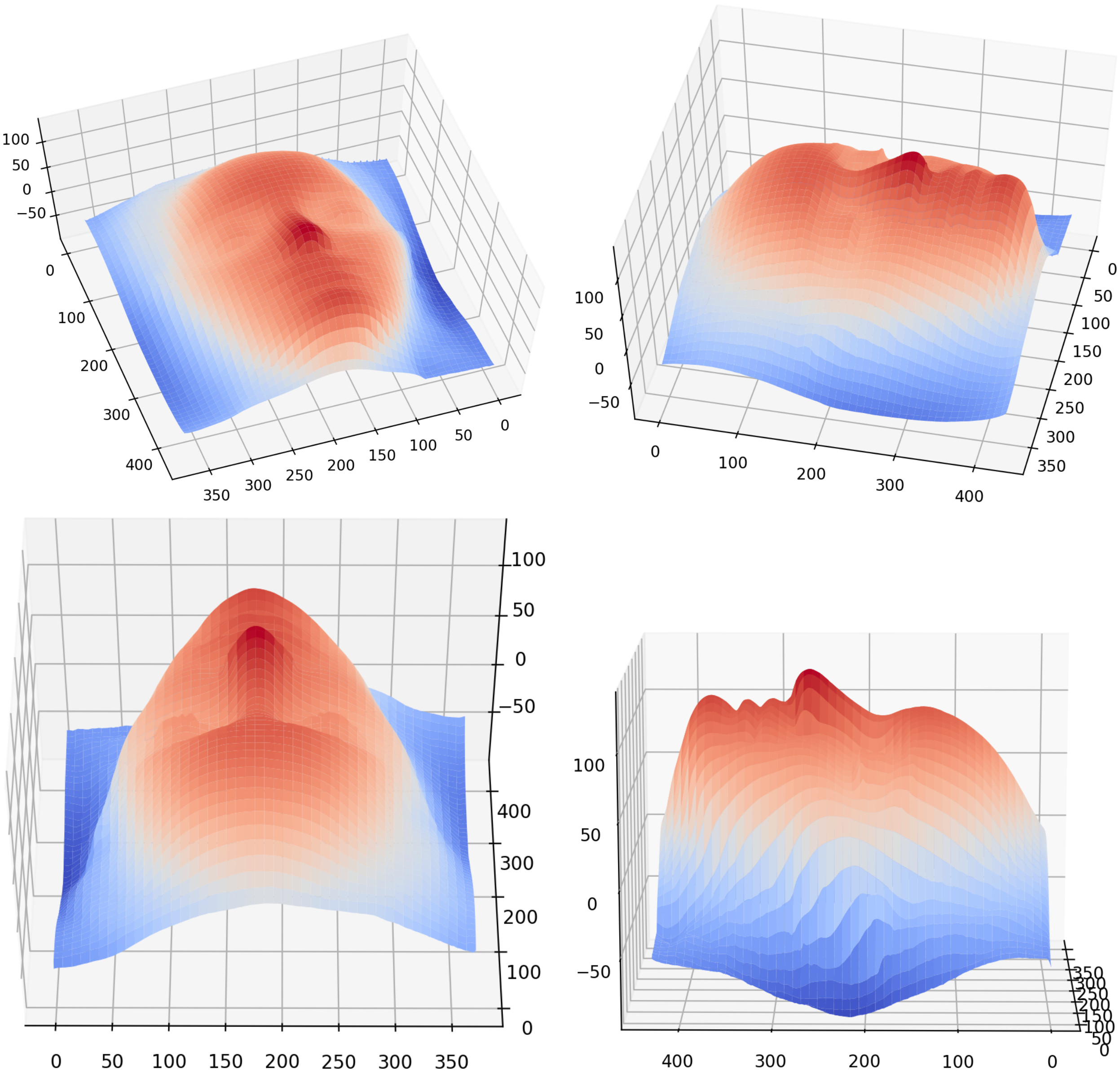

Figure 7: 3D reconstruction result (attempt 2) from uncalibrated photometric stereo.

Q2 (f)

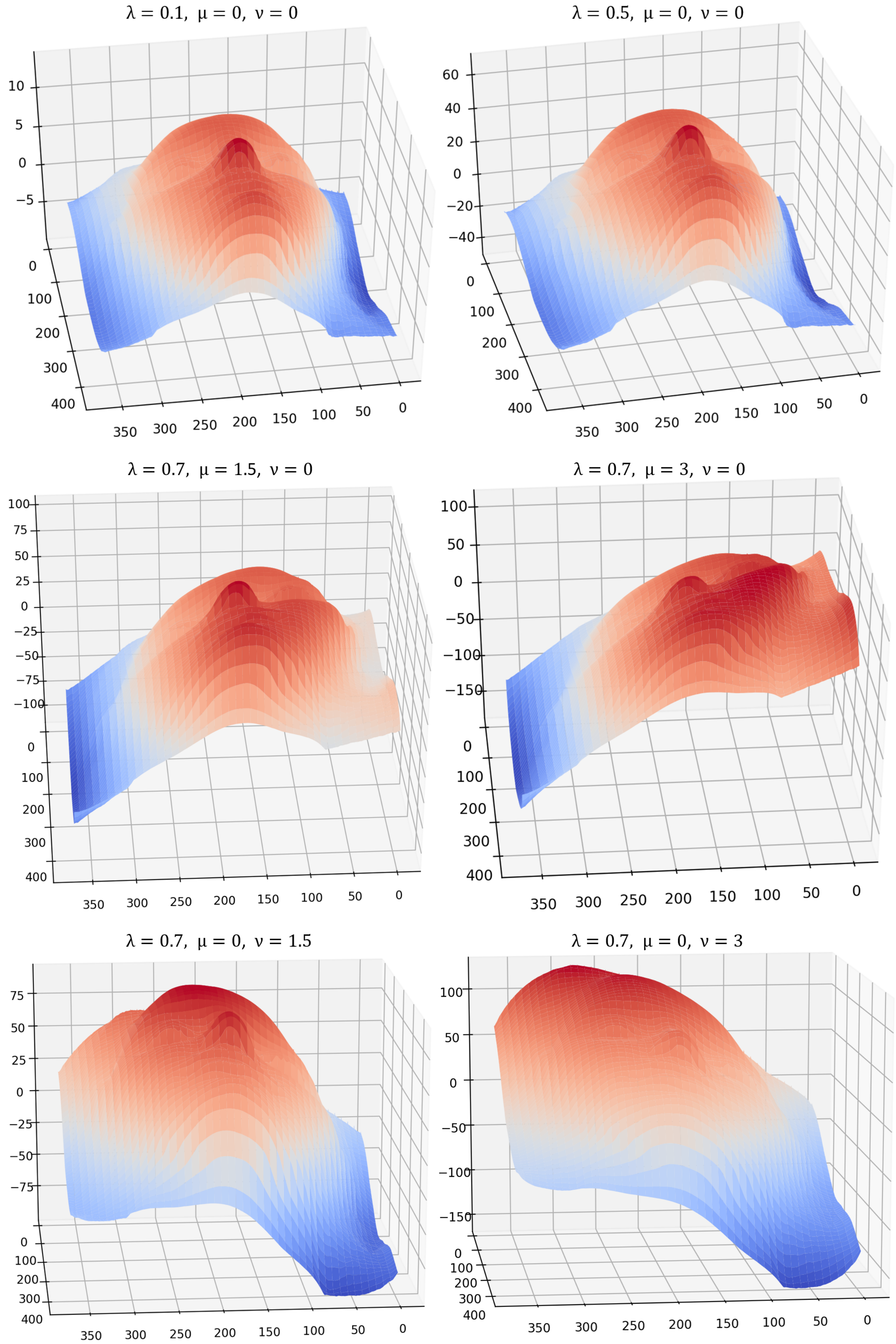

Figure 8: Bas-Relief ambiguity.

The ambiguity is named bas-relief ambiguity because a 3D object with different degrees of flatteness could produce a set of same-looking images under arbitrary lighting sources. A low-relief sculpture looks like a full-relief one when we view it at the front. We cannot resolve the flatteness so we call it bas-relief ambiguity.

λ affects the flatteness of the surface along the **z** axis, μ tilts the surface along the **x** axis, ν tilts the surface along the **y** axis. Smaller λ shrinks the surface along the **z** axis more. The larger the μ is, the surface along the **x** axis will be scaled more and looks more tilted. The larger the ν is, the surface along the **y** axis will be scaled more and looks more tilted.

Q2 (g)

Since the transformation from surface f to new surface $\bar{f}$ is

$$\bar{f}(x, y) = \lambda f(x, y) + \mu x + \nu y$$

the smaller the λ is, the flatter the surface will be along the **z** axis. We could use $\lambda = 0.0001$ or other close-to-zero values to get a flattest possible surface.

Q2 (h)

With the fixed carema view, acquiring more pictures from more lighting directions will not help resolve the ambiguity. At best, one ambiguity, λ, cannot be resolved that people cannot determine how flat the object is. Because for arbitrary light sources, there always exist a bas-relief transformation that a transformed object and transformed light sources can generate the same-looking images.