16-720 Computer Vision
Spring 2020

Homework 6

Alex Wu, chichiaw@andrew.cmu.edu

1 Part I - Calibrated Phtometric Stereo

1.1 a

Intensity at each pixel :

$$I = \frac{\rho}{\pi} I \vec{n} \cdot \vec{l}$$

Projected area:

$$dA \cdot cos\theta$$

θ is the angle between $\vec{n}$ and $\vec{l}$

Since the surface is assumed to be a Lambertian surface, light emits uniformly in all direction. Therefore the intensities are functions of lightsources and surface normals.

1.2 b

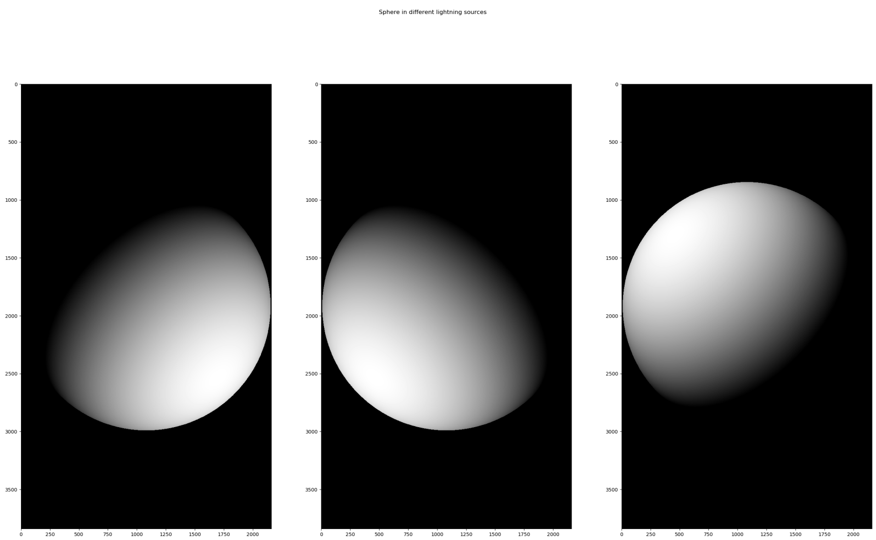


Figure 1: Sphere Rendering Results

1.4 d

$$I = L^T B$$
I dim: 7 x P, L^T dim :7 x 3, I dim :3 x P

Theory:

Rank(AB) $\leq$ min(Rank(A), Rank(B))

Proof:

A is a m x n matrix, B is a n x r matrix
column space C(A) = $\{Ax | x \in R^n\}$
column space C(AB) = $\{ABy | y \in R^r\}$

if C(AB) > C(A)
then there exists a $y \in R^m$, that is in C(AB) but not in C(A)
which implies $\exists z \in R^r$ s.t. ABz = y
$y \in R^{nx1} \therefore y \in C(A)$, which is contradict to the assumption
$\therefore$ Rank(AB) $\leq$ min(Rank(A), Rank(B))

The singular values are [1.26289066, 1.61659626, 2.414729, 9.22148403, 13.16260675, 79.36348099]. There are actually six singular values in result. However, one could argue that there are some uncertainties due to the existence of noise. The first three singular values can be regarded as zero, which then will agree with the rank-3 requirements.

1.5 e

$$
\begin{aligned}
&I = L^T B \\
&LT = LL^T B \\
&(LL^T)^{-1} LI = B \text{ pseudo normal} \\
&A = (LL^T)^{-1} L \quad y = B \\
&\text{Albedos} = \text{magnitude of B, Normals} = \frac{B}{|B|}
\end{aligned}
$$

1.6 f

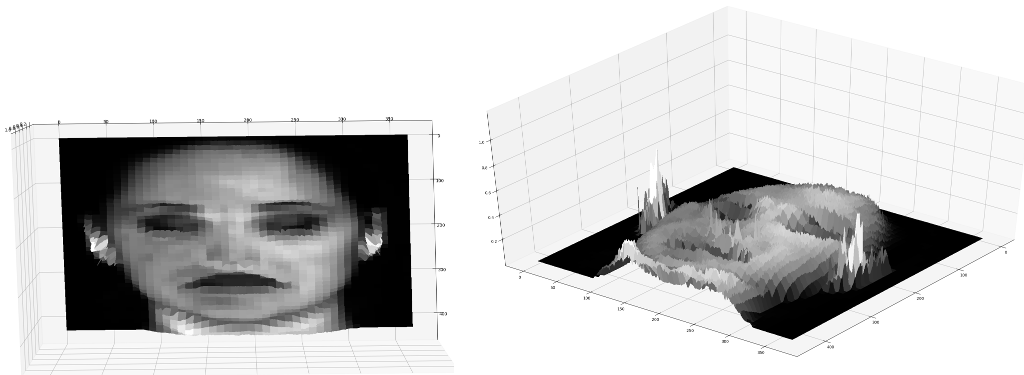


Figure 2: Albedos Visualization

The albedo values are reasonable in most areas within the image. The only unusual features are the enormous high values at both earlobes. After further inspection into the original images, it appears that the woman in the images were wearing ear rings. In some images, there are shinny spots on the metallic material which gives rise to the unusual features in albedo visualization results. We made the assumption that the surface were all lambertian, and this failed at estimating the albedos at the ear rings metallic surfaces.

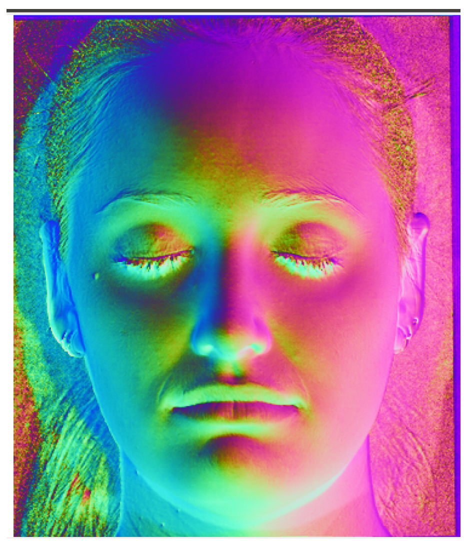


Figure 3: Normal Visualization

The normals match suprisingly well to the expectation of the curvature of the face. It even shows some features that are less obvious in the original images, such as the dark hair of the women.

1.7 g

normal of a surface is denoted as $\vec{n} = \{n_1, n_2, n_3\}$
a surface can be represented in the form $n_1 x + n_2 y + n_3 z = d$

$$\therefore z = f(x, y) = \frac{d - n_1 x - n_2 y}{n_3}$$

$$\therefore \frac{\partial f}{\partial x} = -\frac{n_1}{n_3}, \quad \frac{\partial f}{\partial y} = -\frac{n_2}{n_3}$$

1.8 h

Q:Are the g's calculated from two procedures the same?
A: The two procedures gave out the same result of g.

Q:How can we modify the gradients you calculated above to make g_x and g_y non-integrable?
A: If we simply make each row in g_x or each column in g_y different from each other, then it will make g_x and g_y non-integrable.

Q:Why may the gradients estimated in the way of (h) be non-integrable?
A: Because we assumes the gradients across each row (for g_x) and across each column (for g_y) are the same. It is not always the case in arbitrary matrices.

1.9 i

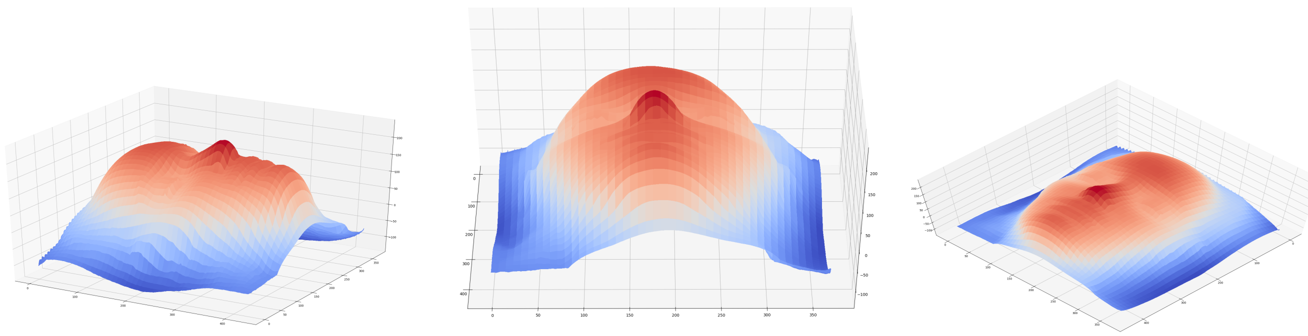


Figure 4: Shape Estimation

2 Part II - Uncalibrated photometric stereo

2.1 a

$I = L^T B$

Compute the SVD of I: $I = UWV^T$

Here we are using best rank-3 approximation:

we create a U_3 by taking the first 3 columns of U

we create a V_3 by taking the first 3 columns of V

we create a W_3 by taking the upper left 3x3 blocks of W

$\therefore$ One possible decomposition is $L^T = U_3 W_3^{1/2}, \quad B = W_3^{1/2} V_3^T$

2.2 b

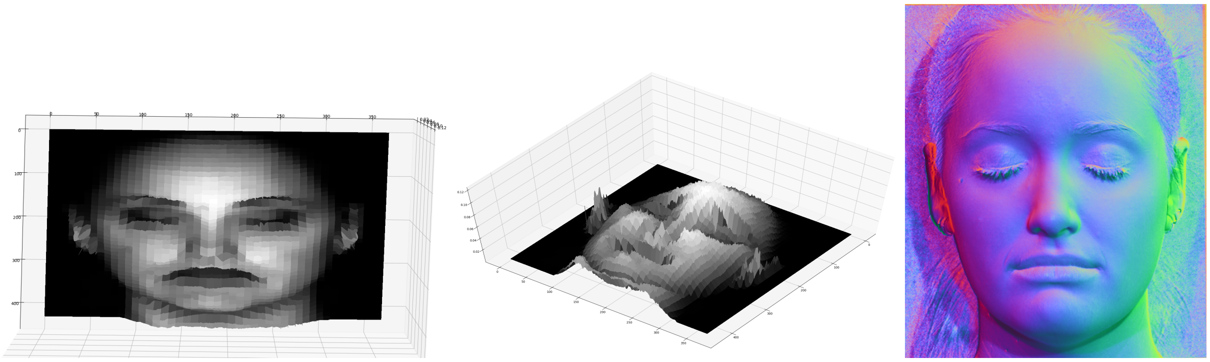


Figure 5: Albedos and Normals results for uncalibrated phtometric stereo

2.3 c

This estimated light source results are based on the factorization method in problem 2a. The estimated light sources are different from the ground truth light sources.

```
ground truth lighting
 [[-0.1418 -0.1804 -0.9267]
 [ 0.1215 -0.2026 -0.9717]
 [-0.069  -0.0345 -0.838 ]
 [ 0.067  -0.0402 -0.9772]
 [-0.1627  0.122  -0.979 ]
 [ 0.      0.1194 -0.9648]
 [ 0.1478  0.1209 -0.9713]]
estimated lightning
 [[ 0.51365956 -0.25905223  0.8179522 ]
 [ 0.21945604  0.5011738   0.83705667]
 [ 0.17201655 -0.19993045  0.96459221]
 [-0.00455701  0.16486523  0.98630558]
 [-0.07244299 -0.54213005  0.83716607]
 [-0.25795646 -0.13171486  0.95713618]
 [-1.8830081   0.79271078  3.3525448 ]]
```

Figure 6: Light Source Comparison

```
ground truth lighting
 [[-0.1418 -0.1804 -0.9267]
 [ 0.1215 -0.2026 -0.9717]
 [-0.069  -0.0345 -0.838 ]
 [ 0.067  -0.0402 -0.9772]
 [-0.1627  0.122  -0.979 ]
 [ 0.      0.1194 -0.9648]
 [ 0.1478  0.1209 -0.9713]]
estimated lightning
 [[ 0.82399461 -0.34782899  0.44726712]
 [ 0.39701972  0.75889616  0.51618985]
 [ 0.42257506 -0.41109482  0.80773224]
 [-0.01253822  0.37967735  0.925034  ]
 [-0.13392744 -0.83889215  0.52756365]
 [-0.59950759 -0.2562197   0.75824938]
 [-0.62008579  0.21849616  0.37632601]]
```

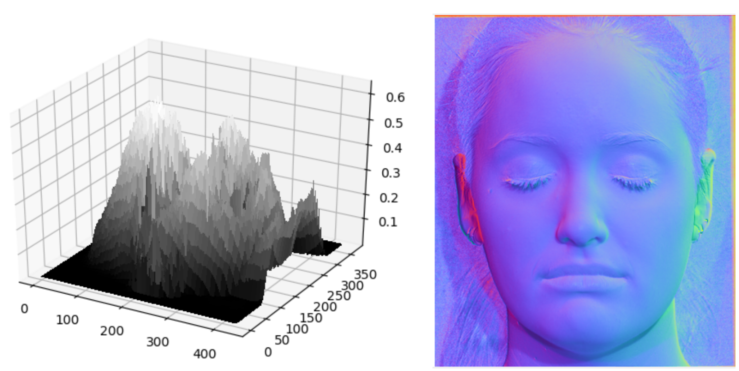


Figure 7: Results from different matrix factorization methods. $L^T = U_3, \quad B = W_3 V_3^T$

2.4 d

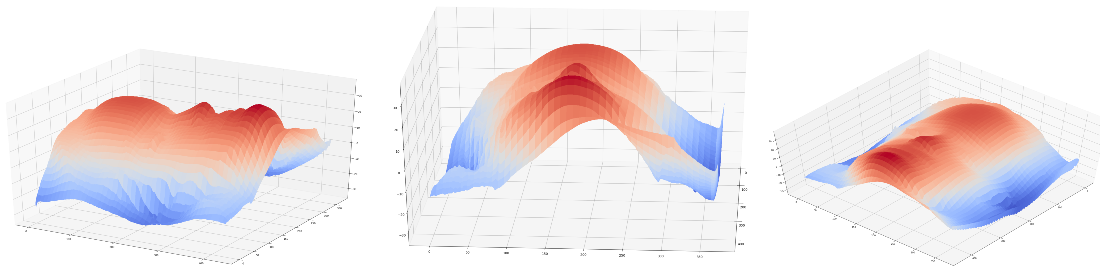


Figure 8: Reconstruction of the face based on the uncalibrated photometric stereo

The results look like a human face, but are distorted in some areas. Compared to the results in problem 1, the face reconstruction results here are less symmetric and are not able to show features of the ears.

2.5 e

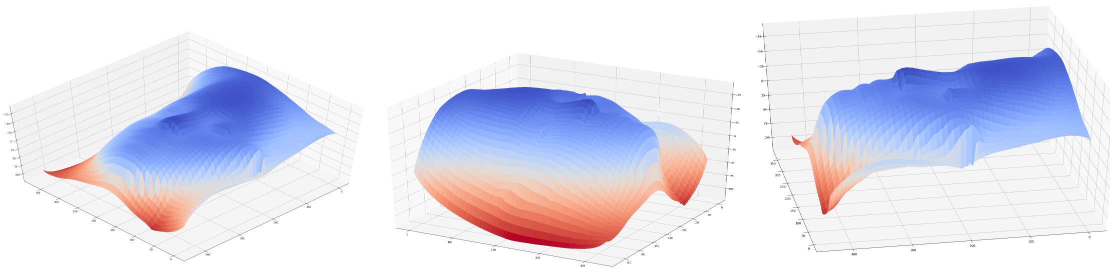
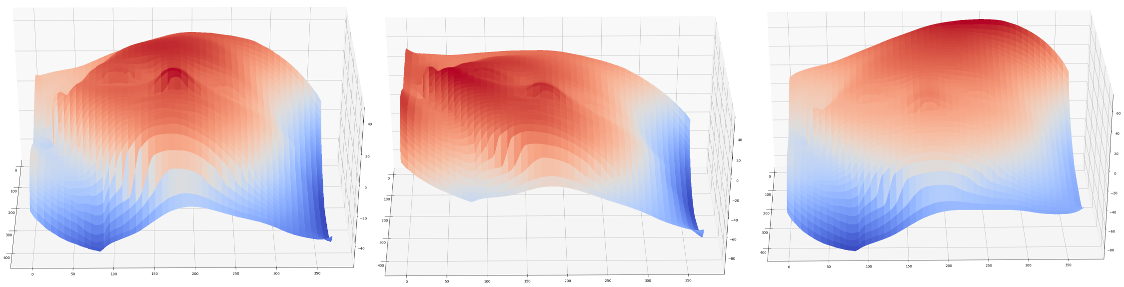


Figure 9: Reconstruction of the face based on the uncalibrated photometric stereo

The results after enforcing integrability are more symmetric than the results in the previous question, and they look more similar to the the results in calibrated photometric stereo. However, the results are still less detailed in some areas.

2.6 f



(a) $\lambda = 1, \nu = 0.01, \mu = 0.01$ (b) $\lambda = 1, \nu = 0.01, \mu = 1$ (c) $\lambda = 1, \nu = 1, \mu = 0.01$

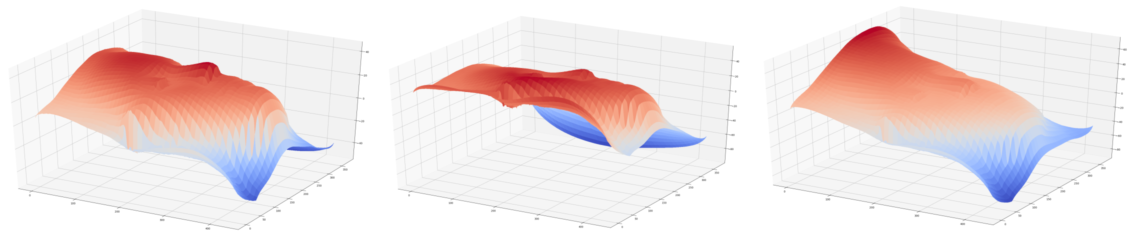


Figure 10: Reconstruction of the faces with different parameters

Q: Looking at these, what is your guess for why the bas-relief ambiguity is so named?
A: The actual shapes of the surfaces have less depth then one would think when they are viewed from an orthogonal perspective.

Q: How the three parameters affect the surface?
A: By looking at the ambiguity equation: $f(x, y) = \lambda f(x, y) + \mu x + \nu y$. Increasing λ will elongate the reconstructed shape, vise versa. If one increases the ν and μ in the G matrix, the reconstructed surface will be flattened. Noted, increase ν will decrease the difference along the lateral axis(ear-to-ear). Increase μ will decrease the difference along the vertical axis(forehead-to-chin)

2.7 g

Q:How would you go about designing a transformation that makes the estimated surface as flat as possible?
A: As mentioned in the previous 2f, smaller λ, larger μ and larger ν will flatten the surface.

2.8 h

Q:Will acquiring more pictures from more lighting direections help resolve the ambiguity?
A: No. The ambiguity remains even if we acquire more pictures. The ambiguity can only be resolved if we have additional information about the albedo or the strength of the light sources.

3 PartIII - Homework Feedback

3.1 a

1) The first two problems in question 1 are not stated very clearly. People asked many questions to clarify many details on piazza, and I personally took quite a few time to get my head around. Maybe a more clear figure of the setups would ease the pain.
2) It would be nice if there is a note at the beginning of each problem stating whether it is a code or write-up problem.