ECE549 / CS543 Computer Vision: Assignment 1

Instructions

1. Assignment is due at **11:59:59 PM on Thursday Feb 20 2020**.
2. Course policies: http://saurabhg.web.illinois.edu/teaching/ece549/sp2020/policies.html.
3. Submission instructions:
   (a) A single `.pdf` report that contains your work for Q1, Q2 and Q3. For Q1 and Q2 you can present your responses in the space provided on pages 2, 3 and 4 (directly typing on PDF, or via LaTeX, or by hand-writing on a print out and scanning it). For Q3 your response should be electronic (no handwritten responses allowed). You should respond to the questions 3(c), 3(d)(i), 3(d)(ii), 3(e), and 3(f) individually and include images as necessary. Your response to Q3 in the PDF report should be self-contained. It should include all the output you want us to look at. You will not receive credit for any results you have obtained, but failed to include directly in the PDF report file. PDF file will need to be submitted to https://www.gradescope.com (Entry Code: **MVZDNW**), and you will need to tag your PDF with where your response to each of the question is.
   (b) You also need to submit code for Q3 in the form of a single `.ipynb` file (with output cleared). Code will need to be submitted to compass2g.
   (c) The LaTeX source is available: http://saurabhg.web.illinois.edu/teaching/ece549/sp2020/mp/mp1-latex.tgz.
   (d) We reserve the right to take off points for not following submission instructions.

Change log

v0	02/06/2020	Creation.

1. **Vanishing Points and Vanishing Lines [10 pts]**[1]**.** Consider a plane defined by $\mathbf{N}^T\mathbf{X} = d$, that is undergoing perspective projection with focal length f. Show that the vanishing points of lines on this plane lie on the vanishing line of this plane.

all points lie on the plane satisfy:

$$N_x X + N_y Y + N_z Z = d$$

$$N^T X = \begin{bmatrix} N_x \\ N_y \\ N_z \end{bmatrix}^T \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = d$$

any line on the plane can be represented as:

$$N_x(A_x - B_x) + N_y(A_y - B_y) + N_z(A_z - B_z) = 0$$

its projection is:

$$N_x \frac{f(A_x - B_x)}{A_z - B_z} + N_y \frac{f(A_y - B_y)}{A_z - B_z} + f N_z = 0$$

the vanishing point of this line is:

$$N_x \frac{A_x - B_x}{A_z - B_z} + N_y \frac{A_y - B_y}{A_z - B_z} + N_z = 0$$

the vanishing line of plane $N^T X = d$ is:

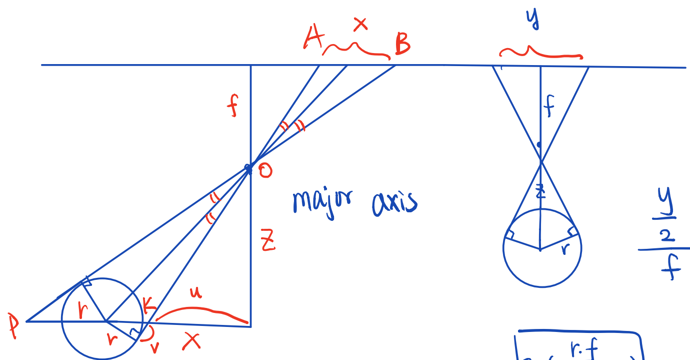
$$N_x x + N_y y + N_z = 0$$

which lies on the vanishing line of this plane

[1] Adapted from Jitendra Malik.

2. **Sphere Under Perspective Projection [30 pts].**[2]

   (a) **[20 pts]** Under typical conditions, the silhouette of a sphere of radius r with center $(X, 0, Z)$ under planar perspective projection is an ellipse. Show that the eccentricity of this ellipse is $\frac{X}{\sqrt{X^2+Z^2-r^2}}$. Recall that, under perspective projection a point (X, Y, Z) in 3D space maps to $\left(f\frac{X}{Z}, f\frac{Y}{Z}\right)$ in the image, where f is the distance of the image plane from the pinhole.
   Hint: There are different ways you can solve this. One line of attack would be to compute the lengths of major and minor axes of the projected ellipse, and compute eccentricity via $e = \sqrt{1 - \left(\frac{\text{length of minor axis}}{\text{length of major axis}}\right)^2}$, but there could be other simpler alternatives as well.



$$\frac{\frac{y}{2}}{f} = \frac{r}{z}$$

$$y = \boxed{2 \cdot \left(\frac{r \cdot f}{\sqrt{z^2-r^2}}\right)}$$

minor axis

$$\frac{v}{u} = \frac{r}{z} \qquad v = \frac{ur}{z}$$

$$r^2 + v^2 = (x-u)^2$$

$$r^2 + \left(\frac{ur}{z}\right)^2 = (x-u)^2$$

$$u = \frac{xz^2 \pm zr\sqrt{z^2+x^2-r^2}}{z^2-r^2}$$

$$PK = \frac{2zr\sqrt{z^2+x^2-r^2}}{z^2-r^2}$$

$$x = \frac{f}{z} PK = \frac{f}{z}(PK)$$

$$x = \boxed{\frac{2rf\sqrt{x^2+z^2-r^2}}{z^2-r^2}} \qquad \frac{x}{PK} = \frac{f}{z}$$

$$e = \sqrt{1 - \frac{(2rf)^2}{z^2-r^2} \quad \frac{(z^2-r^2)^2}{(2rf)^2(x^2+z^2-r^2)}}$$

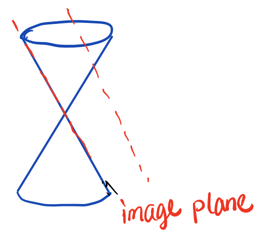
$$= \sqrt{1 - \frac{z^2-r^2}{x^2+z^2-r^2}} = \boxed{\frac{x}{\sqrt{x^2+z^2-r^2}}}$$

[2] Adapted from Jitendra Malik.

(b) **[10 pts]** Are there circumstances under which the projection could be a parabola or hyperbola? If yes, write down the conditions on X, Z, r and f, for parabola and hyperbola respectively; if no, explain why.

yes.

parabola:

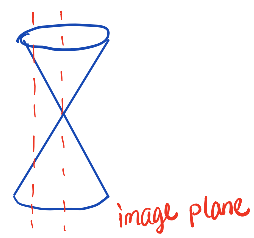


$e = 1$

$X^2 + Z^2 - r^2 = X^2$

$\boxed{Z = r}$

hyperbola:

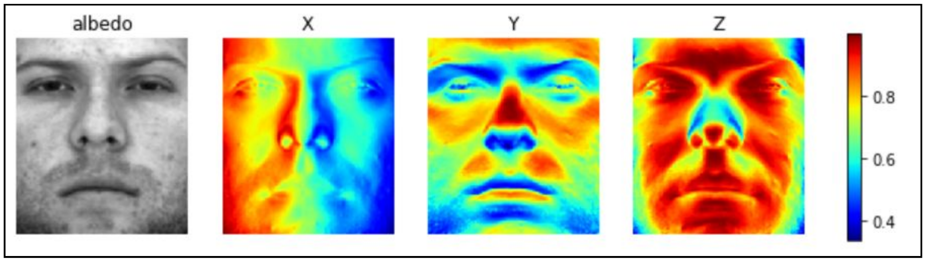


$e > 1$

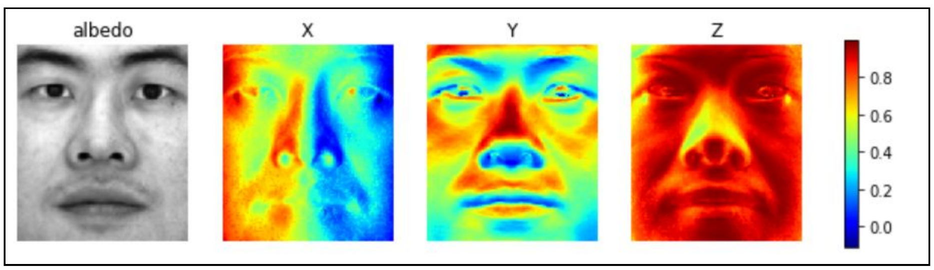
$X^2 + Z^2 - r^2 < X^2$

$\boxed{r > Z}$

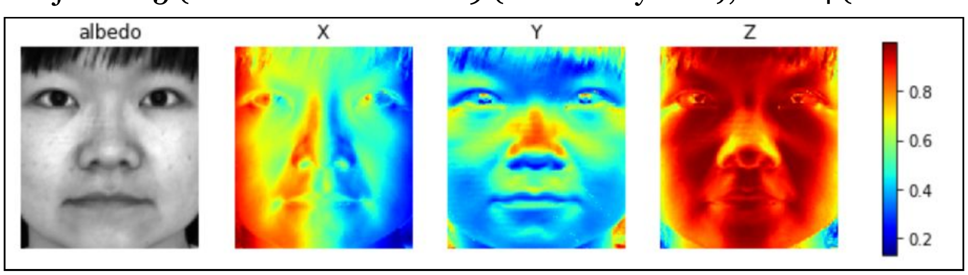
3.c)
Subject B01
(mean residual = 1.82 (divided by h*w), 0.0284 (divided by h*w*N))



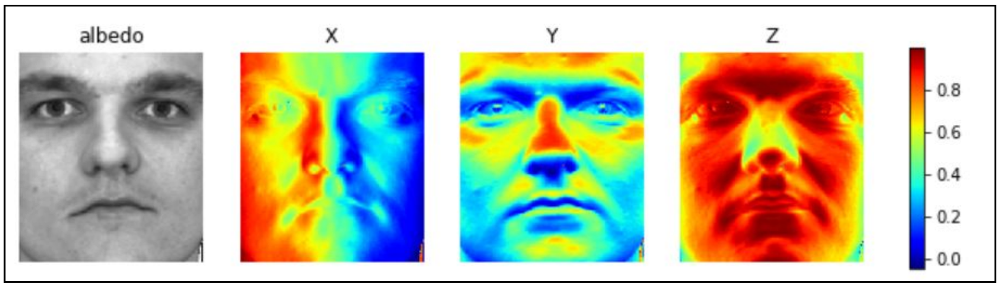
Subject B02 (mean residual = 1.93 (divided by h*w), 0.0301 (divided by h*w*N))



Subject B05 (mean residual = 1.69 (divided by h*w), 0.0264 (divided by h*w*N))



Subject B07 (mean residual = 1.46 (divided by h*w), 0.0228 (divided by h*w*N))

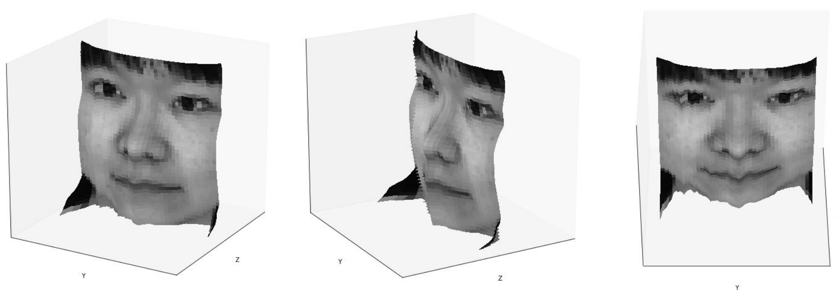


3.d)
Subject B05
i.

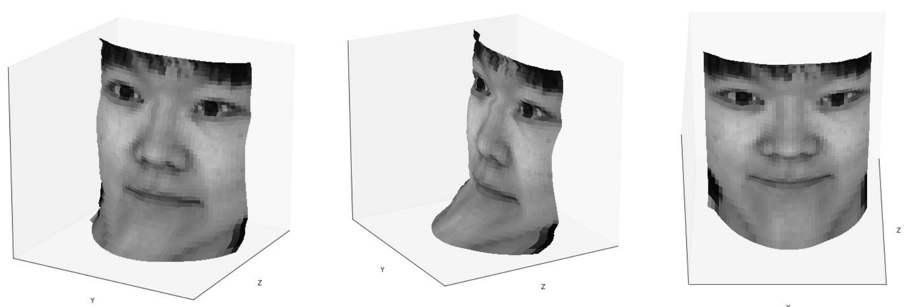
The difference between the different integration methods can be found through both their implementations and the resulting images. The first row is integrated in the 'row' method, while the first column is integrated in the 'column' method. In the resulting images of the 'row' method, the subject's face seems to be stretched sideways, and horizontal features of the face look more prominent. In the resulting images of the 'column' method, the chin, especially, looks more stretched out, and the vertical features of the face become more prominent this time. The 'average' method, since it is implemented by averaging the results from both of the previous methods, yields results that are simply the average of those derived by using the 'row' and the 'column' methods. Finally, in the 'random' method, the integration path of each pixel is randomly decided, and the results derived by using this method have higher visual qualities. The images are able to more closely represent the ground truth.

Different integration methods were tried on images of subject B05. Visually inspected, results derived by using the 'random' method are the best compared to those derived by other methods. However, the ‘random method’ took much more time than the other methods, so there is a tradeoff between runtime and quality of the resulting images. This method outperforms other methods because of two main reasons. Firstly, it resembles the ‘average’ method by taking the average of values derived from 10 random integration paths (10 iterations), which enhances the quality of the resulting images. Secondly, the contributions from the rows’ and the columns’ integral values are treated equally when the image is generated: this also explains why it yields the best result.

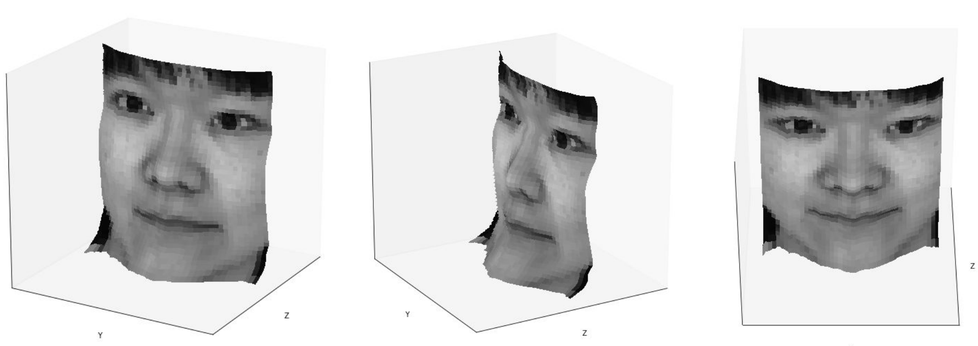
method: ‘row’ runtime: 0.00158 sec



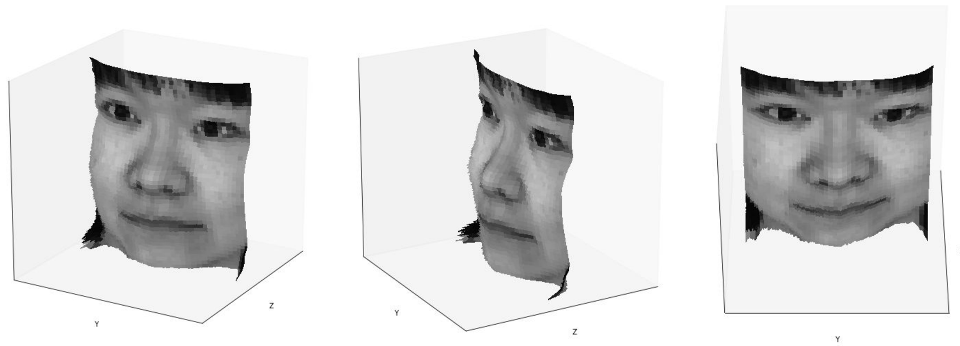
method: ‘column’ runtime: 0.00172 sec



method: ‘average’ runtime: 0.00202 sec

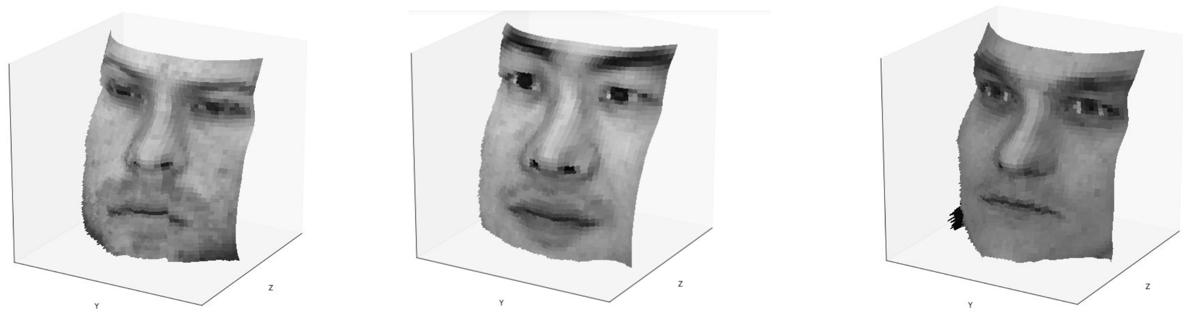


method: ‘random’ runtime: 234.3038404 sec



ii.

Subject B01 Subject B02 Subject B07

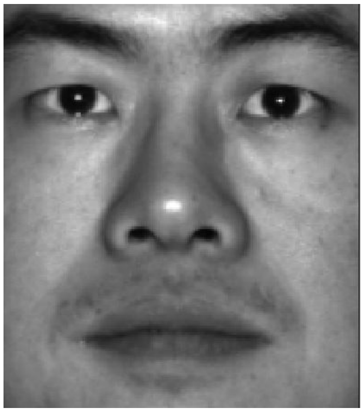


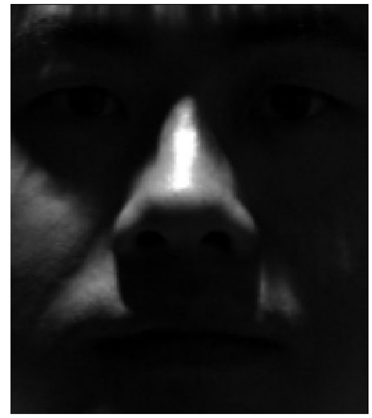
3.(e)

The assumptions of photometric stereo are:

- A Lambertian object
- A local shading model
- A set of known light source directions
- A set of pictures of an object, obtained in exactly the same camera/object configuration but using different sources
- Orthographic projection

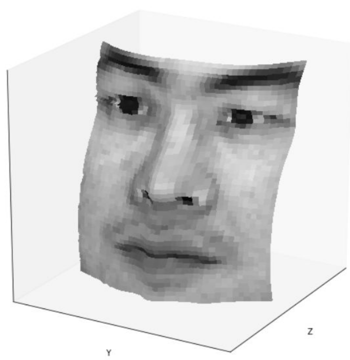
A Lambertian object must have Lambertian reflectance where the fraction of incident light is the same from all directions. Human faces exhibit most of Lambertian reflectance properties, but images might still show specular effects, especially when skin is oily or overly smooth. In the dataset provided, there are images with obvious specular reflectance areas as well as shadows which might contribute to errors in the results. We will take images of the subject YaleB02 as an example. In the images below, there are a few areas of specular reflectance around the nose and checks. Also, since pictures are taken for a human subject, it is almost impossible to keep all images the same: pixels are slightly shifted across different images of the same subject.



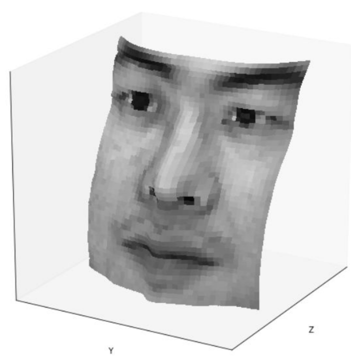


After some of the images that strongly violate the assumptions of Lambertian reflectance were removed, the results derived from 3D reconstruction are shown below. The 3D reconstruction with subset shows slightly better results with less artifacts. However, the face appears to be wider since the subset does not have as many images as before.

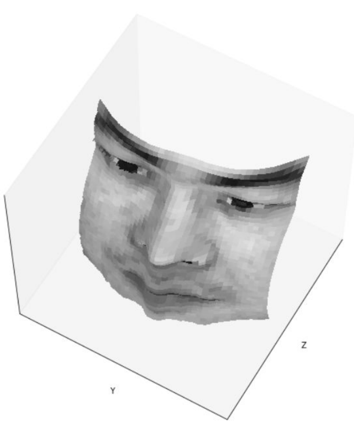
albedo image



with subset



with all images



with subset

with all images