Assignment 1

CS 543 - Computer Vision

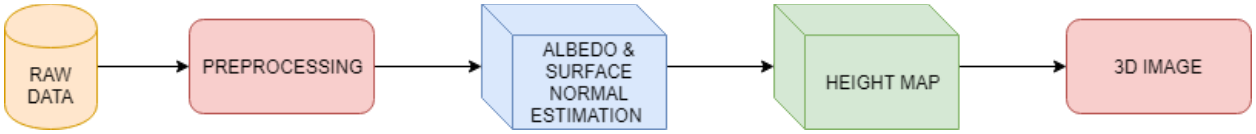
Shape from Shading

Naman Shukla (UIN : 656868559, netID : namans2)

---

1. Introduction:

The aim is of this assignment is to reconstruct the shape of the person's face provided in the Yale face database. This consists of 64 images each of four subjects. The overall flow of the assignment is as follow:



1.1 Assumptions for the above process:

- The subject is lambertian model i.e. the subject follows the lambert's law.
- A complete diffuse reflection by the surface.
- A local shading model (each point on a surface receives light only from sources visible at that point)
- A set of known sources of lights with their respective directions.
- A set of pictures of an object, obtained in exactly the same camera/object configuration but using different sources
- Orthographic projection.

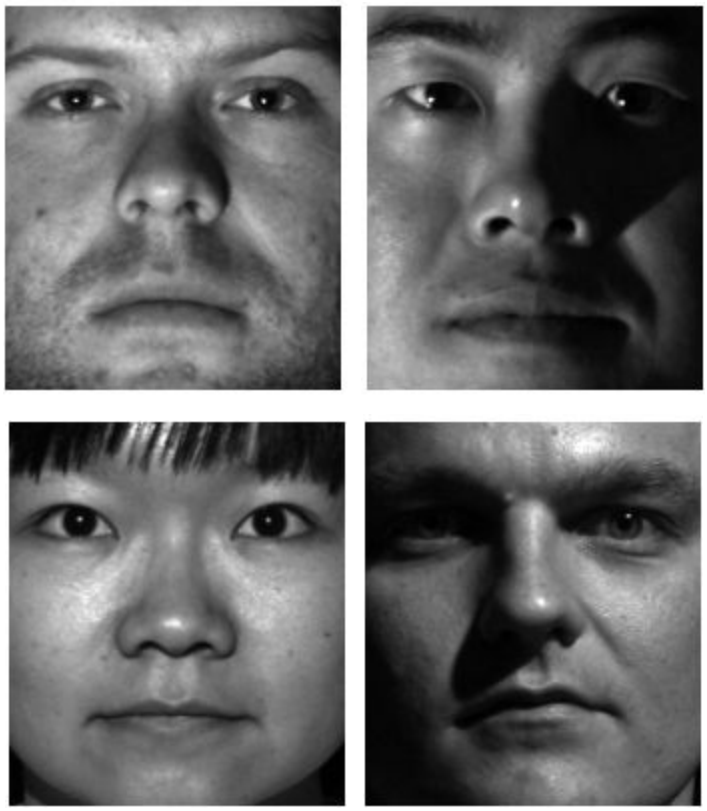
2. Raw Data:

The Yale face database provides images in the PGM files.

3. Processing:

We need to preprocess the data by subtracting ambient image from each images followed by normalization. This process converted the images from 0-255 to 0-1 scale (thresholding to the negative values are also done).

Shown below are the first images out of 64 of each subject after the processing.

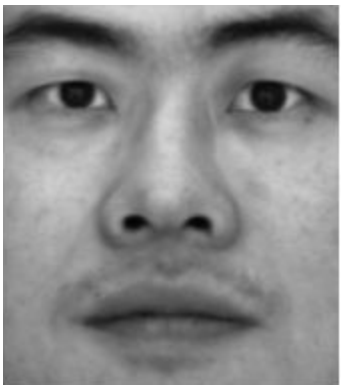
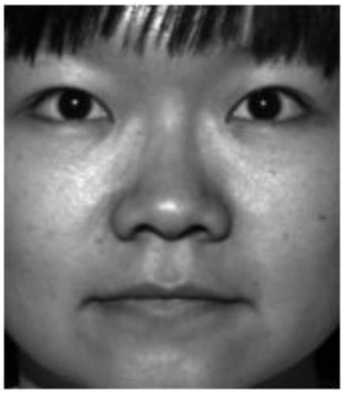


4. Albedo and Surface Normal Estimation:

4.1 Algorithm:

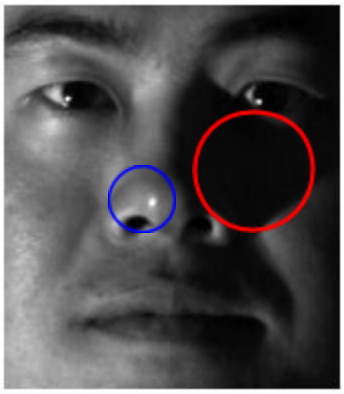
1. Convert the processed images [N x h x w x 3] ➡ [N x (npics)]
2. Find the least square solution of lambertian model. (OLS)
3. Take square root of sums of the OLS solution ➡ reshape [h x w] and set it to albedo image
4. Divide OLS solution with albedo image ➡ reshape [h x w x 3]and set it to Surface Normal.

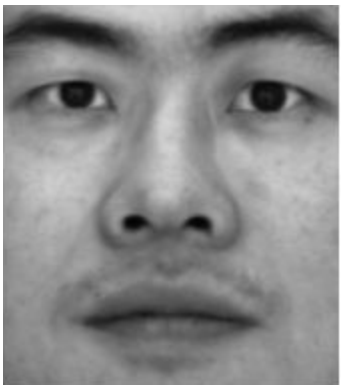
4.2 Results :

Subject	Original Processed Image	Albedo Image
Subject 1 yaleB01		
Subject 2 yaleB02		
Subject 3 yaleB05		

Subject 4 yaleB07		

4.3 Unit Analysis:





The **blue circle** in the processed image is the shiny area (specular reflection) near the nose that is completely blended in the albedo image. The **red circle** is the shadow region in the processed image. The shadow is also removed. The albedo image is symmetric on the left and right side of the face. This indicates that Lambert's law is holding.

5. Height Map

We now have the albedo image and the surface normals for each pixel in the image. The surface normal is of the form [N_1 N_2 N_3]. Given below are four different algorithms for finding heights of each pixel.

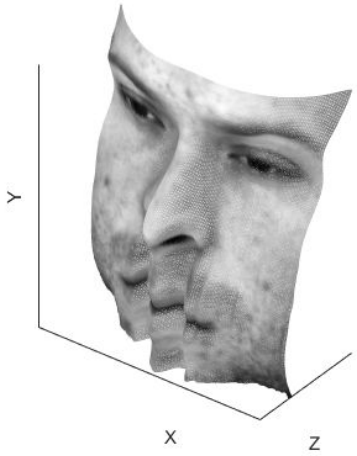
5.1 Algorithms:

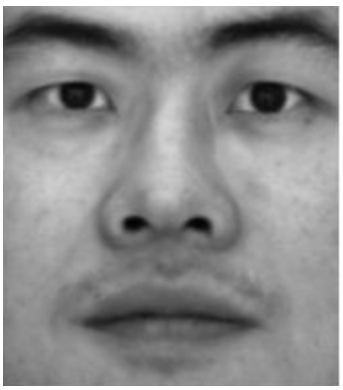
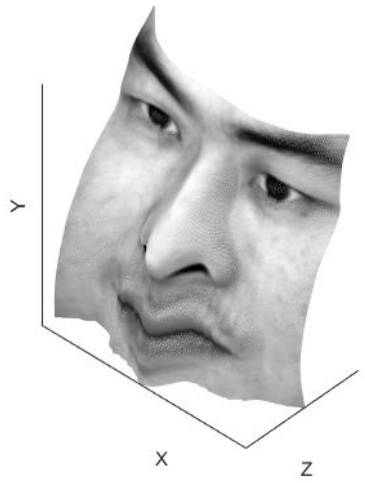
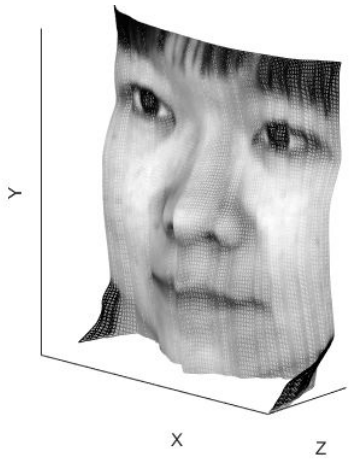
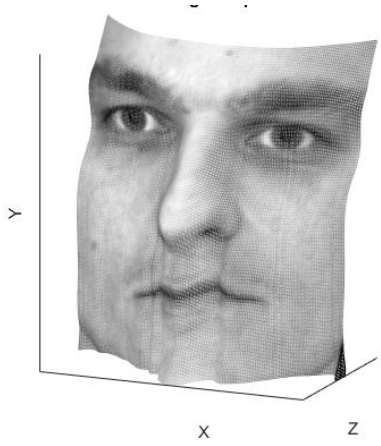
5.1.1 Column method:

1. Create partial gradient matrix
    a. Matrix p : [N_1 / N_3]
    b. Matrix q : [N_2 / N_3]
2. Initialize top left corner of height map to zero.
3. For each pixel in the left column of height map
   height value = previous height + corresponding q value
   end
4. For each row
   For each element of the row except for leftmost
   height value = previous height + corresponding p value
   end
   end

NOTE: The above method is implemented without any loops

Results:

Subject / Comments	Albedo	3D Reconstructed Face
Subject 1 **Comments :** Able to reconstruct the upper face portion successfully. **Artifacts observed:** Zig-zag surface near lips Pointed nose		

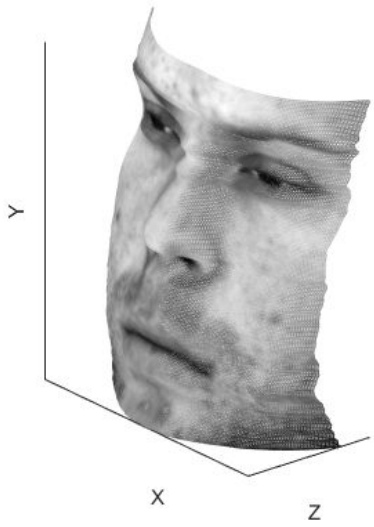
Subject 2 **Comments:** Able to reconstruct most of the left portion of the face. **Artifacts observed:** Unusual right side of the face Pointy lips at middle region		
Subject 3 **Comments:** Able to construct most regions of the face. **Artifacts observed:** Visible vertical lines Surface near lip area is not smooth		
Subject 4 **Comments:** Eyes and cheeks are properly reconstructed. **Artifacts observed:** Uneven surface near lips Shape of nose		

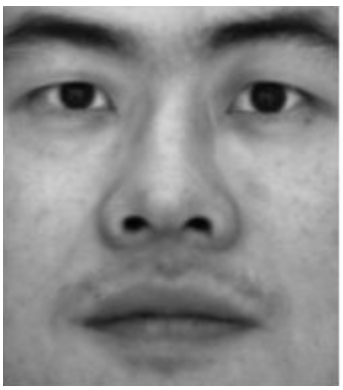
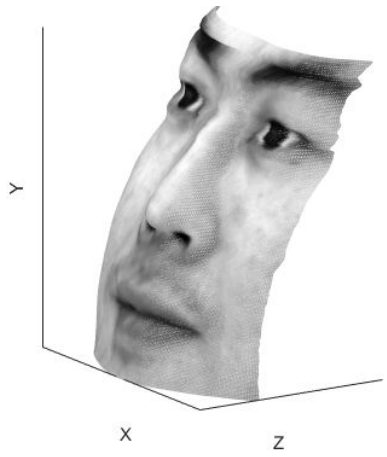
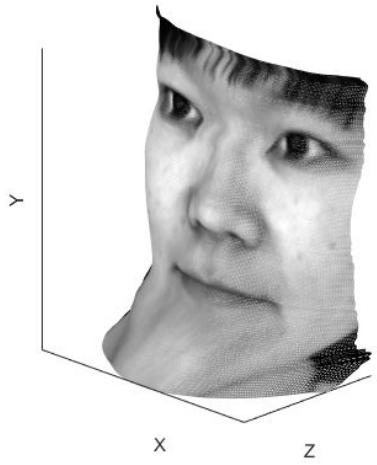
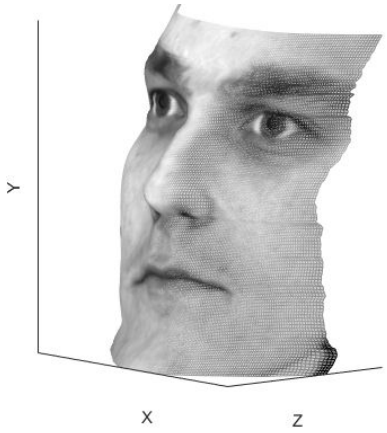
5.1.2 Row method:

1. Create partial gradient matrix
    a. Matrix p : [N_1 / N_3]
    b. Matrix q : [N_2 / N_3]
2. Initialize top left corner of height map to zero.
3. For each pixel in the left row of height map
   height value = previous height + corresponding p value
   end
4. For each column
   For each element of the column except for leftmost
   height value = previous height + corresponding q value
   end
   end

NOTE: The above method is implemented without any loops

Results:

Subject / Comments	Albedo	3D Reconstructed Face
Subject 1 **Comments**: The image is better than the column algorithm but looks like stretched along x-axis. **Artifacts observed**: Nose is not smooth. Lips are flatten.		

Subject 2 **Comments:** The image is better than column method. **Artifacts observed:** Forehead tilted backwards. Eyes stretched along y-axis. Slight discontinuity near eyebrows		
Subject 3 **Comments:** Not much improvement from the column method. **Artifacts observed:** Unusually stretched chin area Flat cheeks		
Subject 4 **Comments:** Great improvements from column method. Overall better structure. **Artifacts observed:** Stretched chin area. Horizontal lines visible		

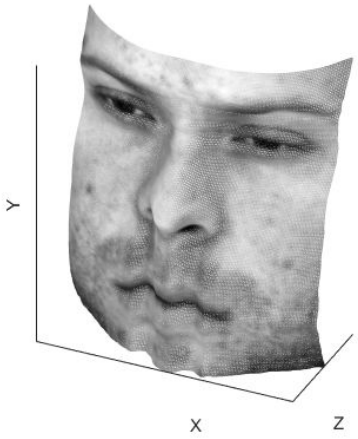
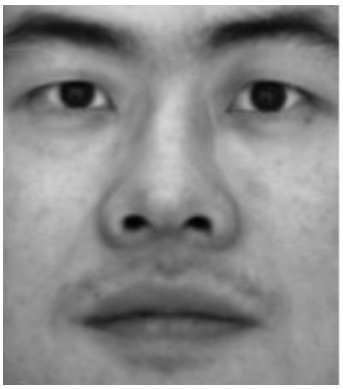
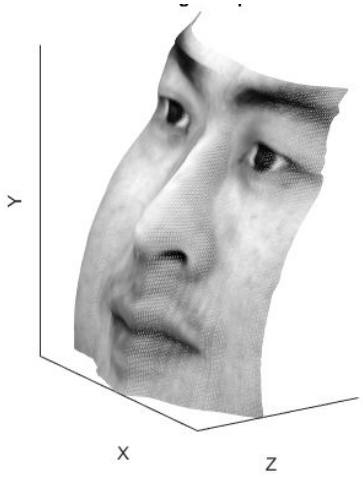
5.1.1 Average method:

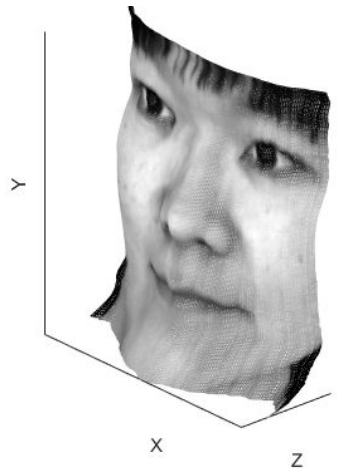
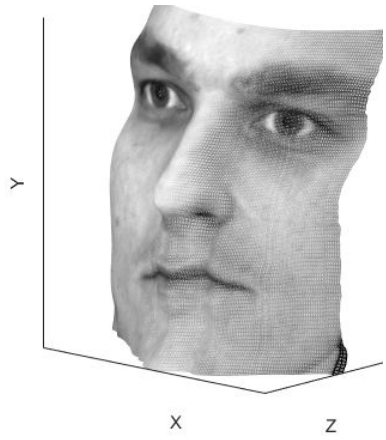
1. Create partial gradient matrix
    a. Matrix p : $[N_1 / N_3]$

b. Matrix q : [N_2 / N_3]

2. Initialize top left corner of height map to zero.
3. Repeat method 1 ➡ set height 1
4. Repeat method 2 ➡ set height 2
5. Take average of the step 3 and step 4

Results:

Subject / Comments	Albedo	3D Reconstructed Face
Subject 1 **Comments:** Better results than previous two methods. **Artifacts observed:** Presence of uneven lip surface		
Subject 2 **Comments:** Much better structure overall. **Artifacts observed:** Forehead still slightly tilted backwards		

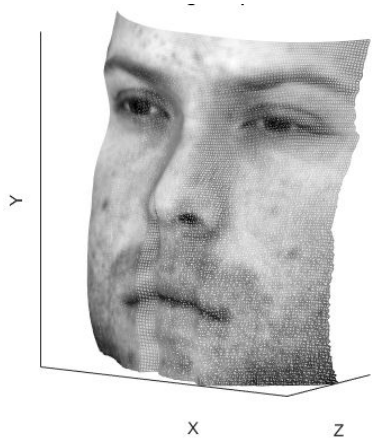
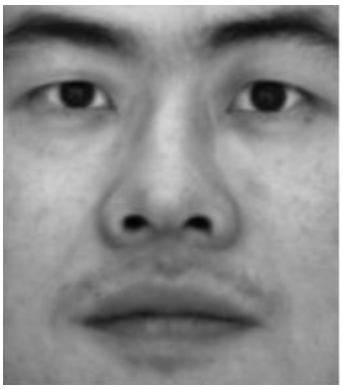
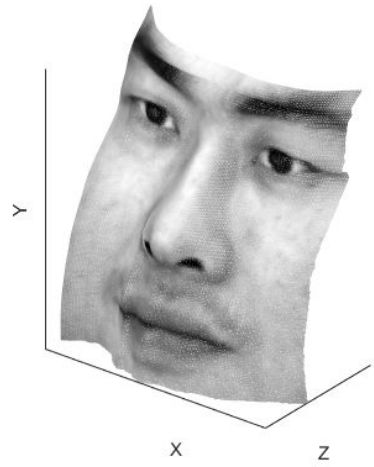
Subject 3 **Comments:** Better result than row method. **Artifacts observed:** Elongated chin		
Subject 4 **Comments:** Much better from above two methods. **Artifacts observed:** Pointed lips		

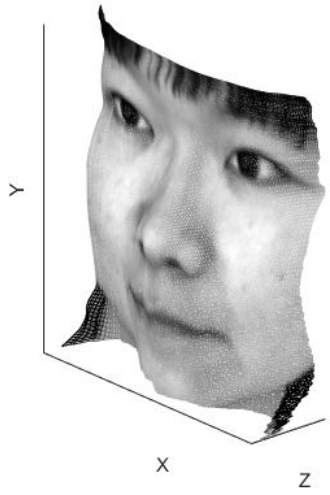
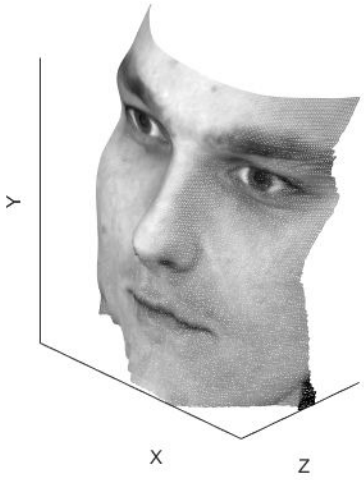
5.1.1 Random path method:

1. Create partial gradient matrix
    a. Matrix p : [N_1 / N_3]
    b. Matrix q : [N_2 / N_3]
2. Initialize top left corner of height map to zero.
3. Set number of paths
4. For each row
    For each column
        For number of paths
            Generate random binary string of length equal to the sum of the coordinates of target pixels.
            Starting from top left and move to target with random string as direction guide
                If '1 : down' ➡ height value = previous height + corresponding q value
                else : 'left' ➡ height value = previous height + corresponding p value
        End
    End

End

Results:

Subject / Comments	Albedo	3D Reconstructed Face
Subject 1 **Comments:** Improved overall. **Artifacts observed:** Pocket around the middle part of the lips		
Subject 2 **Comments:** Normalized the overall unevenness. **Artifacts observed:** Right cheek is slightly elevated.		

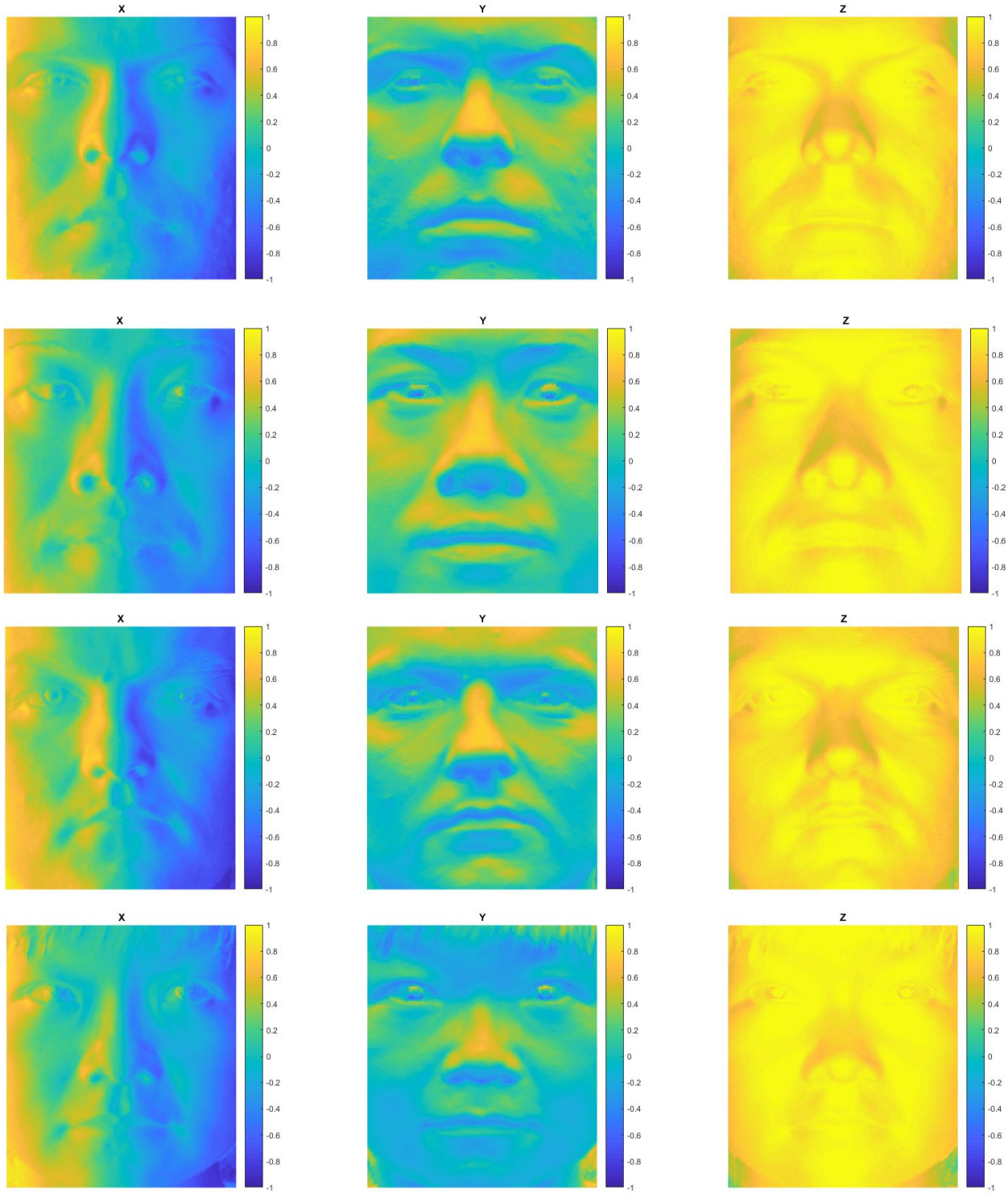
Subject 3 **Comments:** Most effective reconstruction of this subject. **Artifacts observed:** Minute defects around the neck region		
Subject 4 **Comments:** Best facial reconstruction overall. **Artifacts observed:** Rough surface (due to the random paths generated)		

6. Running Time Analysis:

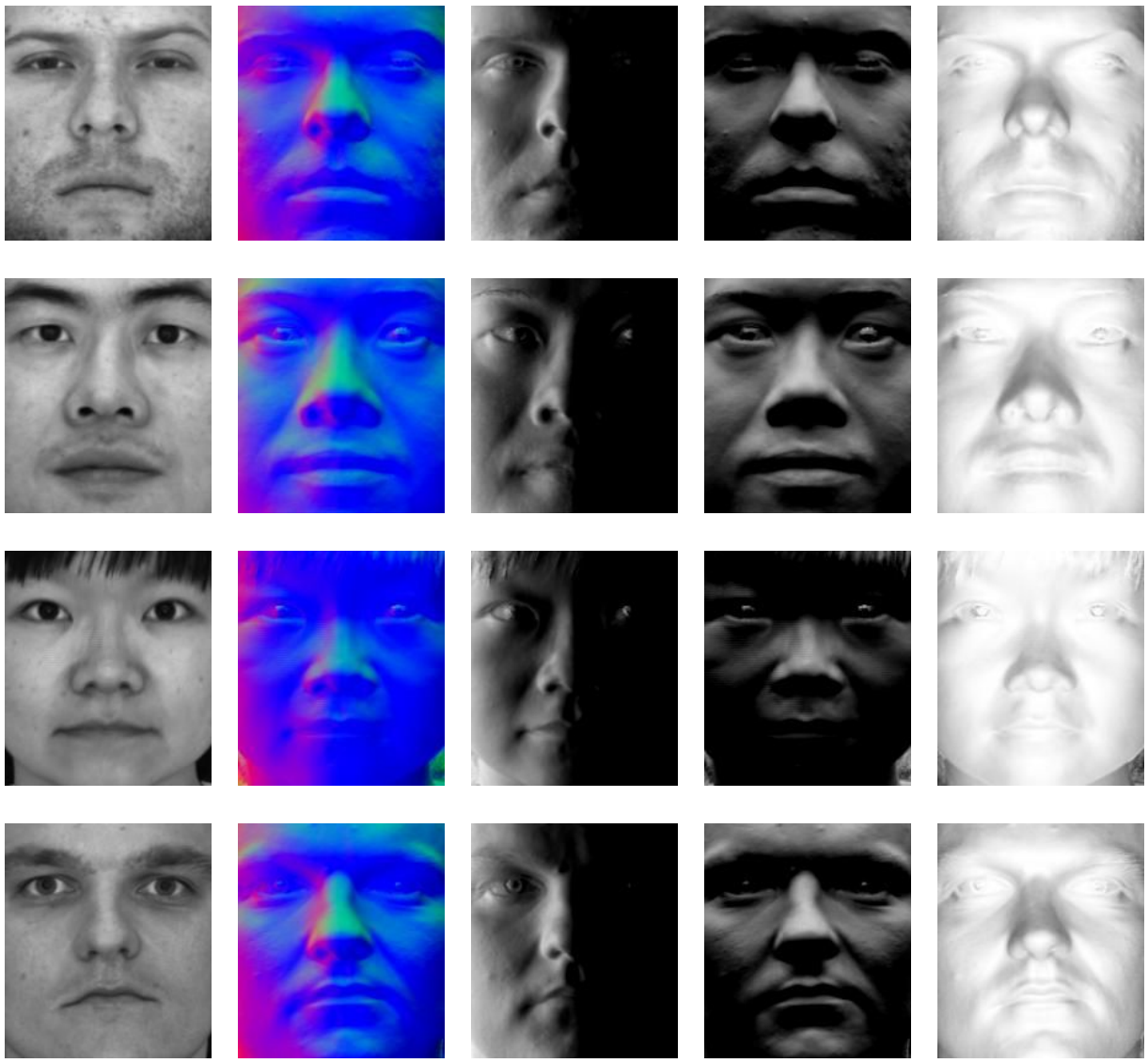
The optimal value of random paths is chosen from the tradeoff between time complexity and accuracy. Hence, I have used 50 random path which gives best results.

	Column	Row	Average	Random path
Subject 1	0.000778	0.000680	0.001830	7.983256
Subject 2	0.000746	0.000684	0.001439	8.294034
Subject 3	0.000706	0.000936	0.001507	7.994182
Subject 4	0.000715	0.000699	0.001993	8.105913

8. Heat Maps:



9. Normals



10. Discussion:

The best profile is obtained from the random path and for the subject 4 (yaleB07). We have able to obtain the 3D model of face which has very high resemblance to the actual face of the subject. Below are few key reasons and the limitations for not obtaining the shape of the face exactly same as ground truth.

10.1 Violation of the assumption:

- <u>Presence of specular reflection:</u>
  Most of the images have shiney area present in it, due to specular reflection. This is a violation of assumption (2).

- Shadows in the Images:
  Some of the images have the shadows present in it. We have not included the role of the shadow in out lambertian model.

- Texture of the skin is not perfectly rough:
  The texture of the skin is not perfect rough that is the reason for the violation of lambertian model as of does not produce diffuse reflection but specular reflection (diffuse reflection violation → not a prefect lambertian model)

10.2 Scope of error:

- Camera position:
  The position of camera and its movement could be a reason for cumulative error in the implementation. This might result in non orthographic images within the set.

- Subject Movement:
  We can observe that the subject is not perfectly still throughout the experiment. The eyes movement (contracted due to light source glare) is very apparent. This is one major source of error.

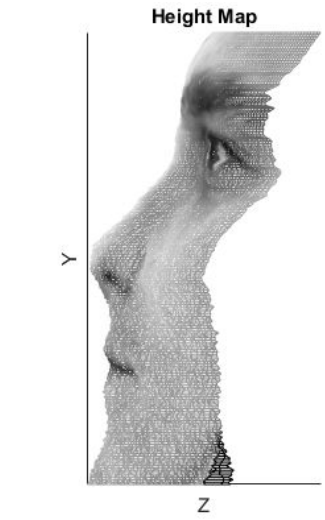
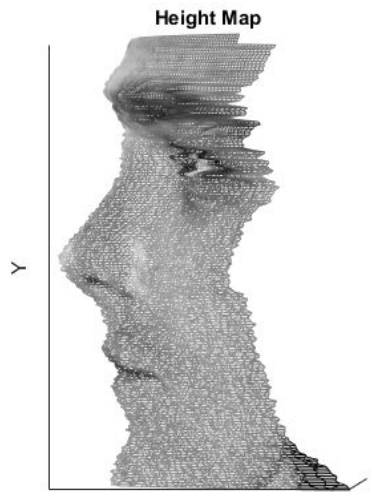
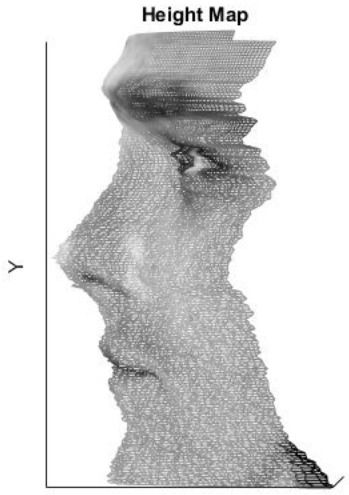
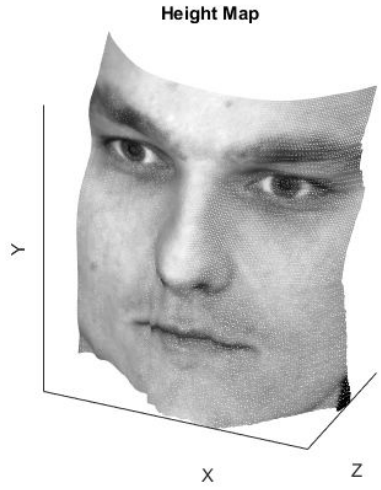
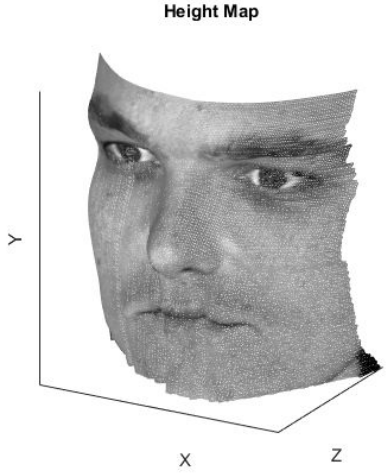
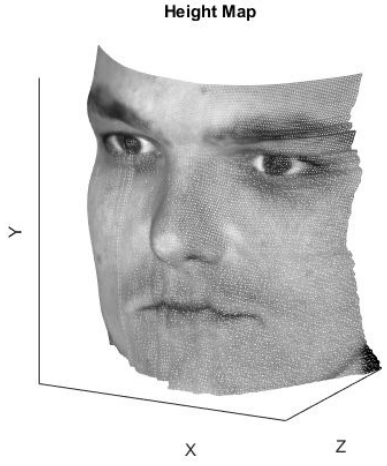
- Instrument Artifacts:
  Few images have bold lines present due to camera defect (artifact).

- Pseudo Random Number:
  The random number generated for the random path is pseudo random. (minor scope of non randomness)

BONUS : Implementing threshold based sampling and additional filter feature

1. Shadow Removal: As there are many images which have dominant shadow effect which contribute to imperfect final result, this threshold based sampling is technique to remove those images. Threshold based sampling works on predefined threshold set for each image. Each image in the dataset has given a score based on the average grayscale value. Only those pictures are allowed to participate in the model formulation which can pass this threshold value. An optimal value of 0.30 is set as threshold in our case.

2. Curvature Enhancement: An attempt to enhance the curvature of some features like eyes and nose of the subject. I have used gaussian and sharp filter (median filter will not work because it will return ‘NAN’ for most of the height values as the possibility of gradient being zero is very high) and below are the results for the over all implementation.

No sampling and No filter	Sampled + sharp filter	Sampled + gaussian filter
		
		

Observations:

- It is very apparent from the side profile that threshold based sampling and filtering makes the constructed model more realistic and hence closer to ground truth.
- Sharp filtered image have more fuzziness as compared to gaussian filtered image.
- The eye in the sharp filtered image is more pointy but gaussian curvature is more realistic (although little extended)
- Minor artifacts are still present along the edges of the 3D image.
- Specularity is getting enhanced through sampling and it is reflected in the final 3D image as well (look near the nose).