COL783: Digital Image Analysis

Assignment 2: Face Expression Transfer and Face Swap

October 15, 2023

Shashank Krishna Vempati (2022AIY7509)

Introduction

The report consists of 2 parts:

- Face Expression Transfer
- Face Swap

Part 1: Face Expression Transfer

In this part of the assignment, we undertake a process to transfer facial expressions from one individual to another. This process is analogous to techniques in image-to-image translation, such as PixelCNN and VAEs, which are capable of generating new images or data that closely resemble the provided examples. In our case, we accomplish facial expression transfer through a method known as image warping. For performing the tasks in Part 1, we took 5 sets of (I1, I2, I3) images with monochromatic background as input images.

(a) I1

(b) I2

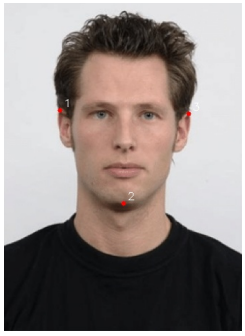
(c) I3

Figure 1: Input Images

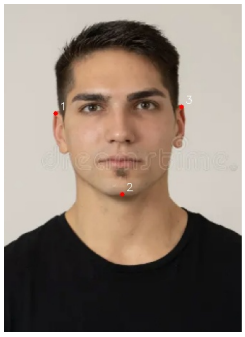
1.1: Warp Matrix based expression transfer for one triangle

In this task, we conducted expression transfer for one triangle between images using warp matrix. We carefully selected three anchor points in each image to ensure semantic correspondence across

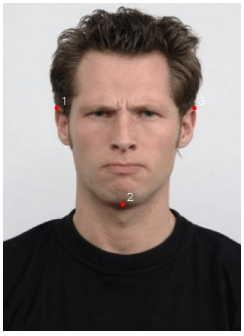
all three. These points formed anchor point triangles, enabling triangle-to-triangle mapping. We computed the warp matrix (H1) between the corresponding anchor point triangles of I1 and I2, which was applied to warp I1 into I1', and the same warp matrix was employed for I3, yielding I3'. A secondary warp matrix (H2) was then computed between the triangles of I1' and I3', allowing us to warp the corresponding triangle in I2, producing I4. Since we are currently working with just one triangle, we may not observe significant expression transfer. However, we intend to use this technique in subsequent experiments to achieve more noticeable expression transfers.



(a) Points on I1

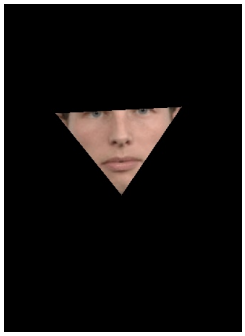


(b) Points on I2

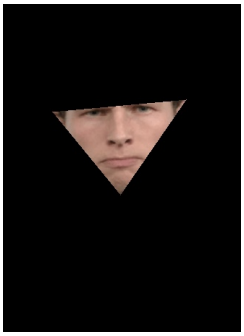


(c) Points on I3

Figure 2: Selected Points



(a) I1' triangle



(b) I3' triangle

Figure 3: Cropped Images after applying H1

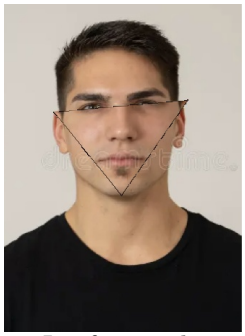
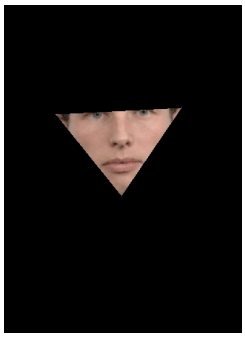


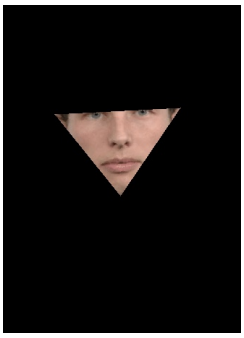
Figure 4: Final Image I4 after applying H2 for one triangle

1.2 Barycentric coordinate-based expression transfer for one triangle

Within our experimentation, we maintain the same image sets (I1, I2, and I3) and ensure the consistency of our previously chosen anchor points. Employing the Barycentric coordinate system, we perform pixel mapping inside the triangles of I1 and I3 to align them with the corresponding vertices of the triangle selected in I2, resulting in the generation of I1' and I3'. Subsequently, we once again employ the Barycentric coordinate system to map the pixels within the triangle of I1', aligning its corners with those chosen in I3' to produce the resulting image, I4.



(a) I1' triangle



(b) I3' triangle

Figure 5: Cropped Images after applying Barycentric method for one triangle

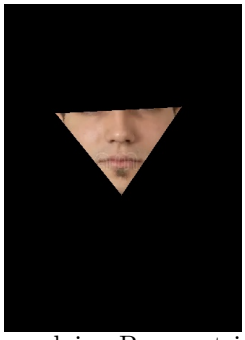


Figure 6: Final Image I4 triangle after applying Barycentric method between I1' and I3' triangles

1.3 Expression transfer over the entire image

For transferring expression over entire image, I recursively follow the methodology established in Task 1 and Task 2 for each of these corresponding triangles to achieve expression transfer. The result of this process is I4, where the entire image has undergone expression transfer.

1.3.1 Expression transfer over the entire image using Task 1 (Warp Matrix)

Technique 1: As explained in write-up (Triangle to triangle warping)

1. *Points Selection:* Selected corresponding points on images I1, I2, and I3 to facilitate expression transfer.

2. *Triangulation:* Utilized the SciPy module for barycentric triangulation on all three images (I1, I2, and I3) to create triangles. This step defined regions for subsequent warping.

3. *Warping from I1 to I2:* Computed warp matrices (H1) for sets of corresponding triangles between I1 and I2. Applied these matrices to image I1, yielding the transformed image I1'.

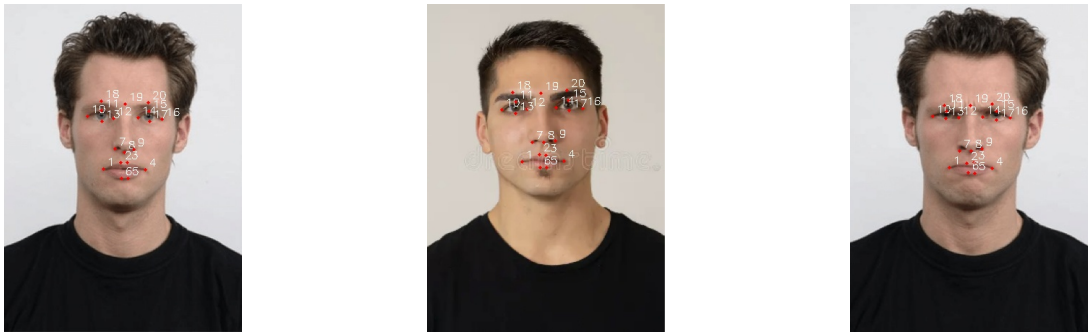
4. *Warping from I3 to I2:* Assumed the same set of points selected on I1 and applied the corresponding warp matrices to match triangles in I2. This process warped I3 to align with I2, resulting in the image I3'.

5. *New Point Correspondences:* Derived fresh coordinates on I3' by utilizing the pre-selected points from I3 and applying the appropriate warp matrices based on the allocation of points within specific triangles.

6. *Warping from I1' to I3':* Calculated warp matrices of corresponding triangles between images I1' and I3', defining the transformation between corresponding Delaunay triangles.

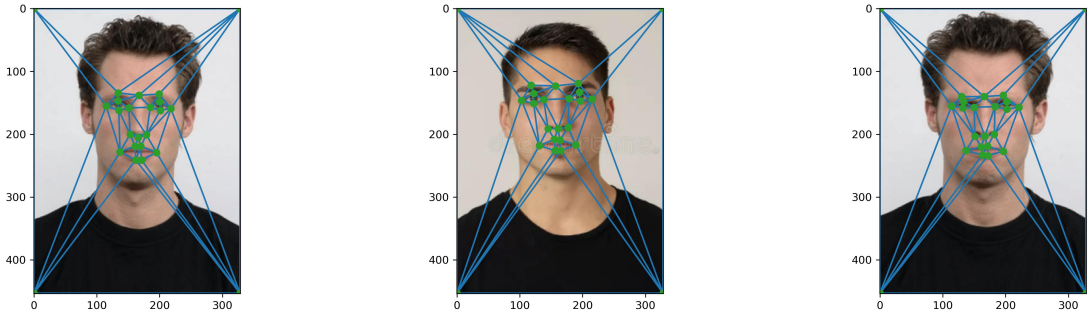
7. *Expression Transfer to I2:* Applied the warp matrices (H2) to image I2 to obtain the transformed image, I4.

This transformation ensured that all Delaunay triangles on I2 were warped according to H2, effectively transferring the expression from I3 to I2.



(a) I1 with 20 points (b) I2 with 20 points (c) I3 with 20 points

Figure 7: Input images with marked points (20 corresponding points)



(a) I1 with Delaunay (b) I2 with Delaunay (c) I3 with Delaunay

Figure 8: Delaunay Triangulation on I1, I2 and I3 w.r.t corresponding points

(a) I1' after applying H1

(b) I3' after applying H1

Figure 9: I1' and I3' after applying H1 on both images

Figure 10: I4 after applying H2 (20 correspondences)

Drawbacks of Technique 1:

As we are applying H1 computed between I1 and I2 on I3 to get I3', we assume that corresponding points chosen on I1 and I3 are at similar positions so that relative information is intact after the warping procedure. So this technique fails if the facial features are not at aprroximately at same positions in both images I1 and I3.

Additionally, the forward warping process can potentially introduce gaps or overalps in the resulting image because it doesn't guarantee that all pixels will be mapped. In contrast, utilizing backwarping to map and then warping the image can effectively eliminate these gaps or voids in the image. In the above technique, I used backward warping for consistency in the resultant images.

Technique 2: Global Alignment followed by Triangulation and Warping

1. *Correspondence Selection:* Initial step involved selecting corresponding points on images I1, I2 and I3 to establish the basis for expression transfer.

2. *Global Alignment of I1 and I3:* Performed global alignment between I1 and I2 using warp matrix. This process resulted in the creation of image I1'. Similarly, performed global alignment between I3 and I2 which resulted in the image I3'.

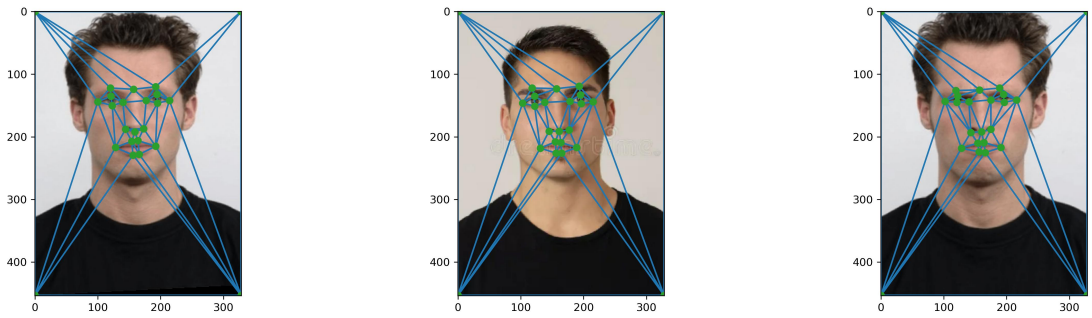
3. *Delaunay Triangulation on I2:* Applied Delaunay triangulation to the selected points on image I2. This step created a network of triangles.

4. *Mapping Triangles from I1' to I3':* For each triangle in I1', mapped the corresponding triangle in I3'. This mapping was achieved using warp matrix between corresponding triangles of both images.

5. *Expression Transfer to I4:* The result of the above step i.e mapping of triangles from I1' to I3' is used on I2 to get I4 and hence transferring the relative information between I1' and I3' to I2.

In this revised procedure, the initial alignment focuses on preserving the local information of I3 by aligning images I1 and I3 relative to I2. Subsequently, the established point correspondences effectively capture the relative information between I1 and I3. This captured information is then leveraged to transfer the expression to I2, resulting in the generation of I4.

Same set of facial landmarks are used from Technique 1.



(a) I1 with Delaunay (b) I2 with Delaunay (c) I3 with Delaunay

Figure 11: Delaunay Triangulation on I1, I2 and I3 w.r.t corresponding points

(a) I1' after global alignment (b) I3' after global alignment

Figure 12: I1' and I3' after performing global alignment

Figure 13: I4 after warping corresponding triangles of I1' and I3'

Note: The expression transfer in Technique 1 did not yield the expected results in few scenarios. Consequently, I made slight adjustments to the method which is similar to Barycentric approach, and found that it is working much better in cases where I1 and I3 have faces at different positions.

Note 2: we solely implement global modifications to align I1 and I3 with I2, yielding I1' and I3', and subsequently apply global adjustments between I1' and I3', then extend these adjustments to I2 to generate I4, the result will not encompass expression transfer. Global adjustments primarily entail warping the image while considering the overall perspective, which doesn't facilitate expression transfer.

1.3.2 Expression transfer over the entire image using Task 2 (Barycentric System)

1. *Initial Image Alignment: Global Warping* The first step involved performing a global alignment of both I1 and I3 with respect to I2. This resizing aimed to harmonize the proportions and features of I1 and I3 with those of I2, for proper transfer of facial expression. Global warping encompasses not just resizing but also involves rotation and translation. Hnece, we obtain I1' and I3'.

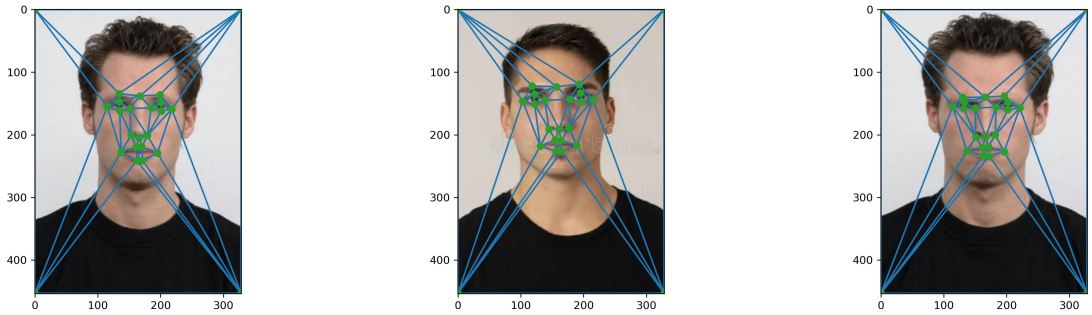
2. *Selection of Correspondences:* Correspondence points were established in the three images: P1: Set of points on I1 P2: Set of correspondences on I2. P3: Set of correspondences on I3.

3. *Resized Correspondence Points:* These points are properly transformed according to warped images in the initial step. P1' was the new location of P1 in the warped I1. P3' was the new location of P3 in the warped I3.

4. *Displacement of P2:* To achieve alignment, I displaced points P2 in the reference image I2, resulting in a new set of points, P4. The displacement was determined based on the difference vector between P3' and P1'. This process was repeated for all control points, ensuring that the relative information between I1' and I3' is captured.

5. *Modified Points for I4:* With the adjusted points P4 for all correspondence points in I2, I proceeded to generate a new image, I4. In the process of generating the image, we took into account the pixels contained within each triangle of the Delaunay triangulation. Barycentric coordinates played a crucial role in calculating the pixel values within these triangles (reverse mapping of pixels is employed), ultimately giving rise to the transformed image, I4.

Here also, I assumed the same set of 20 points chosen in Technique 1.



(a) I1 with Delaunay (b) I2 with Delaunay (c) I3 with Delaunay

Figure 14: Delaunay Triangulation on I1, I2 and I3 w.r.t corresponding points

(a) I1' after global alignment (b) I3' after global alignment

Figure 15: I1' and I3' after performing global alignment

Figure 16: I4 after applying Barycentric system

Conclusion

When the facial position within the image remains relatively stable, the expression transfer appears consistent across various scenarios, whether utilizing warp matrices or the barycentric system. In both approaches, I implemented reverse mapping techniques on the pixel level to ensure a seamless transition, preventing gaps or overlaps in the resulting image. Ablations will give much deeper idea on how each technique is different.

Part 2: Face Swap

The main aim of this section is to develop a robust and visually appealing face swapping technique, building upon the principles of face warping. This face swapping process will focus exclusively on the facial region of the source and target images, excluding areas like hair and neck.

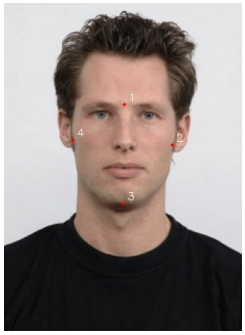
(a) I1

(b) I2

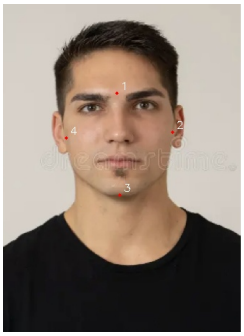
Figure 17: Input Images

2.1 Global Pose Correction

To initiate the face swap process with global pose correction, I began by selecting four points symmetrically positioned on the forehead, right ear, chin, and left ear of both images, I1 and I2, establishing correspondences. Subsequently, I performed a complete image warp of I1 based on I2 and reciprocally warped I2 based on I1 with gloval affine matrix. This served as the initial step in the face swapping procedure.



(a) I1 (4 points marked)



(b) I2 (4 points marked)

Figure 18: Point Selection for Global Pose Correction

(a) I1 after Pose Correction

(b) I2 after Pose Correction

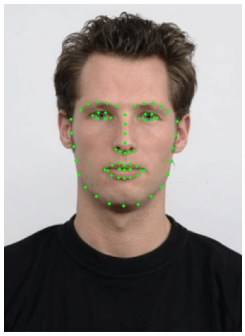
Figure 19: Pose Corrected Images

2.2 Obtain Facial Landmarks

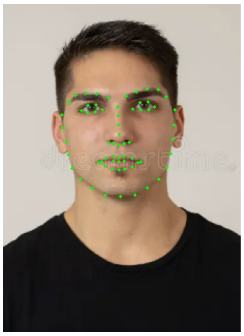
The next in this process is to obtain corresponding facial landmarks from both images I1 and I219. These landmarks include key points on the face, such as the eyes, nose, and mouth. Additionally, boundary points are selected to define the region of the face to be swapped, excluding areas like the neck and hair.

The implementation has two options:

- Automatic facial landmark detection using the dlib module.
- Manual selection of facial features if automatic detection is not preferred.



(a) I1 facial landmarks (Automatic selection)



(b) I2 facial landmarks (Automatic selection)

Figure 20: Selected landmarks on the images (Using dlib module)

2.3 Delaunay Triangulation followed by Warping

This step involves Delaunay triangulation applied to the boundary points of both images (I1 and I2) followed by image warping of corresponding triangles. The procedure is divided into two sub-steps:

Delaunay Triangulation on I1: Triangles are formed based on the boundary points and facial landmarks of I1. These triangles are then mapped to corresponding triangles in I2 using point correspondences. The facial features are transferred between the corresponding triangles using barycentric coordinates, resulting in an intermediate face of I1 that is aligned with the pose of I2.

Delaunay Triangulation on I2: Similar to the above step, triangles are formed on the boundary points and facial landmarks of I2 and mapped to corresponding triangles in I1. The facial features are transferred from I2 to the intermediate face of I1 as shown below.

(a) I1 after Delaunay (b) I2 after Delaunay

Figure 21: Delaunay Triangulation on the selected landmarks

(a) I1 without face (b) I2 wihtout face

Figure 22: Images without facial landmarks

(a) I1 warped according to I2 (b) I2 warped according to I1

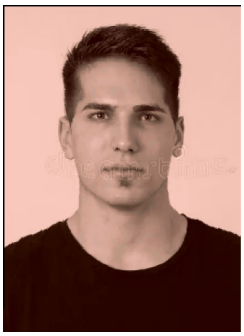
Figure 23: Warped faces (Ready to be swapped)

2.4 Color adjustment using swatch

In the step involving color adjustment using the swatch, we adjust the colors of the replaced face to closely resemble those of the original image. The swatch selection process is manual, and we randomly choose 50 colors from the swatch. Subsequently, we execute color transfer from

the swatch to the substituted face by considering intensity values (L channel of the LAB color format for both the swatch and the face). We then overlay the facial segment onto the image, generating a new image with the face swapped. This procedure is applied to both I1 and I2, resulting in the swapped images shown below.

(a) I1 colored using swatch from I2

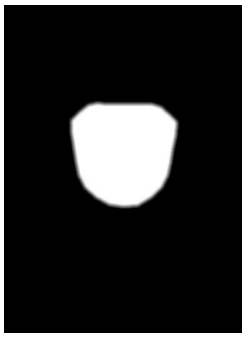


(b) I2 colored using swatch from I1

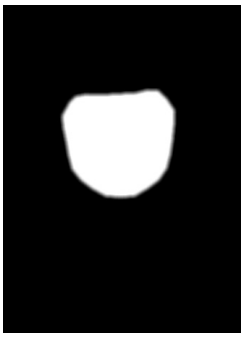
Figure 24: Coloring of Images using Swatch

2.5 Smoothing for the boundary

In the face swapping process, we might notice strange-looking areas near the face, where it doesn't match up well with the rest. To fix this, I added a smooth mask to the warped face and used a technique called alpha blending. This makes the transition from the face to the surrounding areas smoother and more natural.

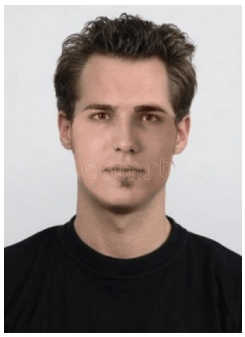


(a) I1 mask



(b) I2 mask

Figure 25: Mask with smooth boundaries for alpha blending



(a) I1 after swapping

(b) I2 after swapping

Figure 26: Images after swapping and color matching (Final Outputs)

2.6 Additional Step : Histogram Matching

Sometimes, the lighting in the two images can be quite different. In such situations, we can choose to use histogram matching on the warped images for the specific swapping part of the image. This helps ensure that the brightness levels match up, and we can then enhance the overall appearance by transferring colors using color swatches, resulting in improved final outputs. Added as option in the code.

Link to Drive for all files

Google Drive Link containing all Files with Images

Ablation Study

Part 1: Face Expression Transfer

I performed ablations by varying the number of points chosen for Delaunay triangulation, reporting results using five different sets of points, ranging from 1 point (in addition to the four default corner points) to 30+ points.

1 point: When a single point is considered, it doesn't transfer the expression itself but rather introduces sharp deformations to the image in some cases.

5 points: The effect of five points in the warping process is substantial, resulting in a sharp alteration of the image. While this may yield unusual or unexpected output, it also provides some visual cues or indications of the direction in which the image is being transformed.

10 points: Using ten points in the warping process leads to a smoother transition, reducing some of the sharp and unconventional changes in the image. However, it's important to note that even with ten points, it may not be adequate to achieve a complete transformation of the image with all the desired modifications.

20, 30+ points: Increasing the number of points, such as 20 or more, contributes to a more effective expression transfer, resulting in reduced uneven changes. When a precise and appropriate set of corresponding points is chosen, the outcome of the expression transfer is notably improved, leading to more faithful and refined image transformations.

Set 2:

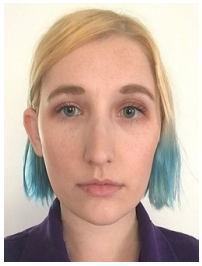
(a) I1

(b) I2

(c) I3

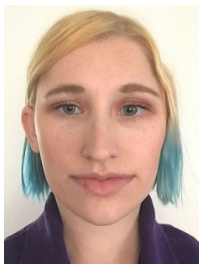
Figure 27: Input images

(a) Task 3 using warp matrix

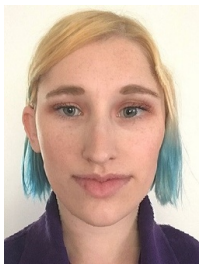


(b) Task 3 using barycentric coordinates

Figure 28: 1 point

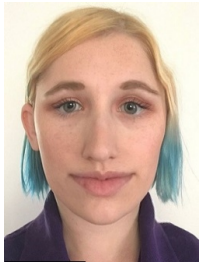


(a) Task 3 using warp matrix



(b) Task 3 using barycentric coordinates

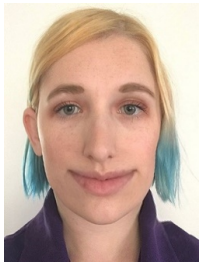
Figure 29: 5 points



(a) Task 3 using warp matrix

(b) Task 3 using barycentric coordinates

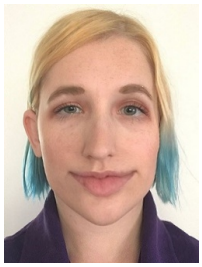
Figure 30: 10 points



(a) Task 3 using warp matrix

(b) Task 3 using barycentric coordinates

Figure 31: 20 points



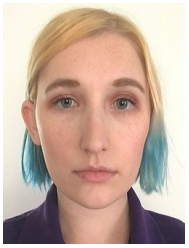
(a) Task 3 using warp matrix

(b) Task 3 using barycentric coordinates

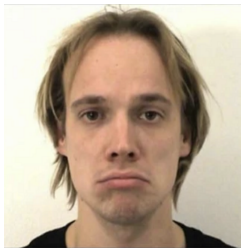
Figure 32: 30 points

Set 3:

(a) I1

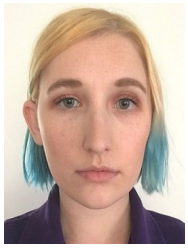


(b) I2

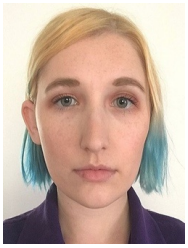


(c) I3

Figure 33: Input images

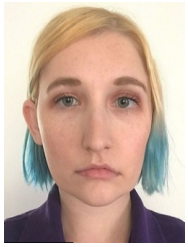


(a) Task 3 using warp matrix

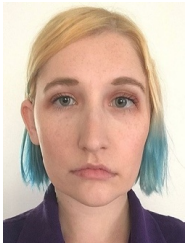


(b) Task 3 using barycentric coordinates

Figure 34: 1 point

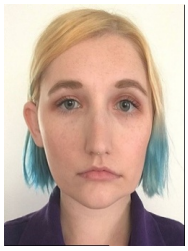


(a) Task 3 using warp matrix

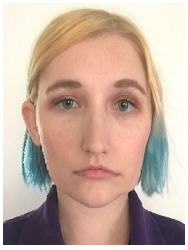


(b) Task 3 using barycentric coordinates

Figure 35: 5 points

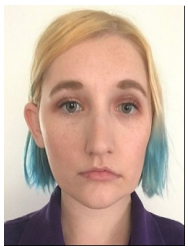


(a) Task 3 using warp matrix

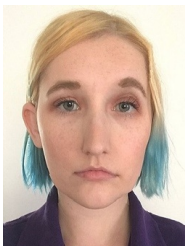


(b) Task 3 using barycentric coordinates

Figure 36: 10 points

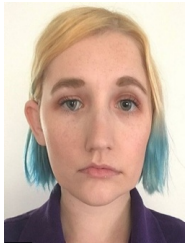


(a) Task 3 using warp matrix

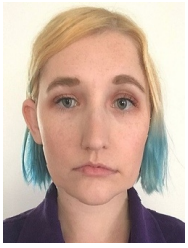


(b) Task 3 using barycentric coordinates

Figure 37: 20 points



(a) Task 3 using warp matrix



(b) Task 3 using barycentric coordinates

Figure 38: 30 points

Set 4: Face position changed slightly and hence Barycentric coordinates based results are better

(a) I1

(b) I2

(c) I3

Figure 39: Input images

(a) Task 3 using warp matrix

(b) Task 3 using barycentric coordinates

Figure 40: 20 points

Set 5: In this case barycentric method worked better as we did global pose correction according to I2 and then took the corresponding points whereas its the not the case for warp matrix based method

(a) I1

(b) I2

(c) I3

Figure 41: Input images

(a) Task 3 using warp matrix

(b) Task 3 using barycentric coordinates

Figure 42: 20 points

Figure 43: Task 3 using warp matrix (Technique 2)

Note: Technique 2 on set5 worked better than Technique 1 because of the fact that, I included global pose correction before processing the image further. Hence the relative information stayed intact. So there are no visible unnecessary distortions in the image.

Part 2: Face Swap

Set 2:

(a) I1

(b) I2

Figure 44: Input Images

(a) I1 swapped

(b) I2 swapped

Figure 45: Images after face swapping (Automatic point selection using dlib)

Set 3:

(a) I1

(b) I2

Figure 46: Input Images

(a) I1 swapped

(b) I2 swapped

Figure 47: Images after face swapping (Automatic point selection using dlib)

References

1. Face Morphing Article
2. Lecture related to Transformations
3. Reference Material
4. Delaunay Triangulation