Style Transfer for Headshot Portraits

Team Name - **CONNECTED COMPONENTS**
Project ID - 5

Team Members-
1) Chirag Parikh – 2022900005 PGSSP
2) Kethu Sesha Sarath Reddy – 2020102028 ECE
3) Aravapalli Dheeraj Murugan Sai – 2020102020 ECE
4) Neeraj Sai Ramana Veerla – 2021121008 LCD

Mentor TA: Manasvi Vaidyula

Repo URL: https://github.com/Digital-Image-Processing-IIITH/dip-m22-project-connected-components

Main goal of the project:

- Style transfer of head portrait (robustly transfer the local statistics) , by using a reference image while matching properties such as the local contrast and the overall lighting direction while being tolerant to the unavoidable differences between the faces of two different people. This can allow one to easily reproduce the look of renowned artists.
- Additionally, because artists sometimes produce entire headshot collections in a common style, we show how to automatically find a good example to use as a reference for a given portrait, enabling style transfer without the user having to search for a suitable example for each input.

Problem Definition:

What is the problem? (Motivation):

- When we look at the photos of celebrities taken by professionals under perfect conditions , we wish our head portraits to look like them in style, even professional photographers find it different to take many photos of a similar style.
- Given an input unprocessed headshot and a model headshot by an artist, we describe an automatic algorithm to transfer the visual style of the model onto the input.

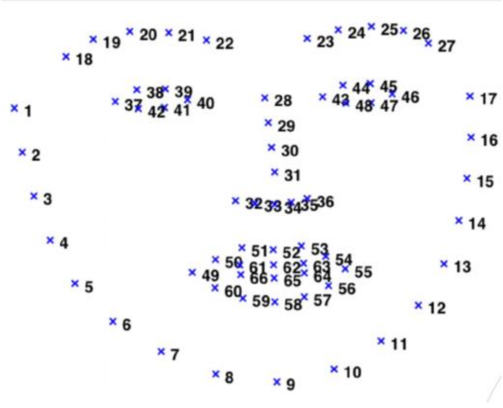
Procedure

- Dense correspondence of Example to Input (Warp)
- Multiscale transfer of Local contrast
- Additional Post processing
- Automatic Selection of the Example in a collection

Dense correspondence of Example to Input (Face Warp):

1.) Estimating facial landmarks

- ➢ We find 66 facial landmarks for a head image using an algorithm from the paper below.

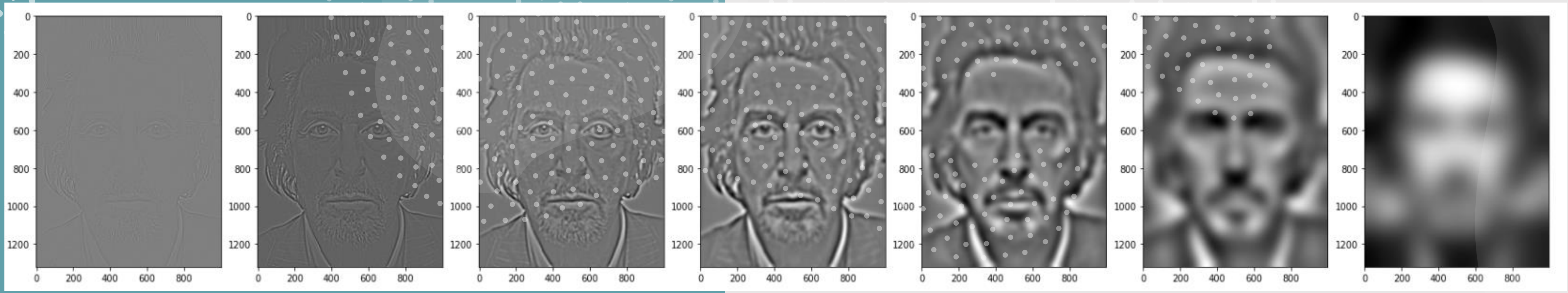


2.) Point-to-Point correspondence, Delaunay Triangulation and Affine transform-based alignment of eyes and mouth:

•We roughly align the eyes and mouth of the example with those of the input using an affine transform using an algorithm from the paper below.

3.) Region-to-Region correspondence:

•We morph the example to the input using the segments on the face template using an algorithm from the paper below.



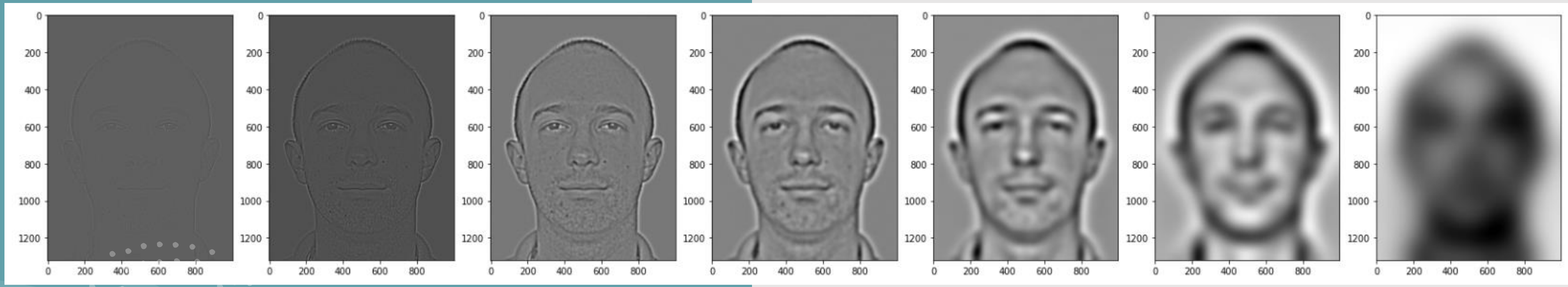
Multiscale Transfer of Local Contrast

Decompose the input and example images into multiscale Laplacian stacks.

$$L_\ell[I] = I - I \otimes G(2) \quad \text{if } \ell = 0$$

$$L_\ell[I] = I \otimes G(2^\ell) - I \otimes G(2^{\ell+1}) \quad \text{if } \ell > 0$$

1.)Multiscale decomposition (Laplacian):



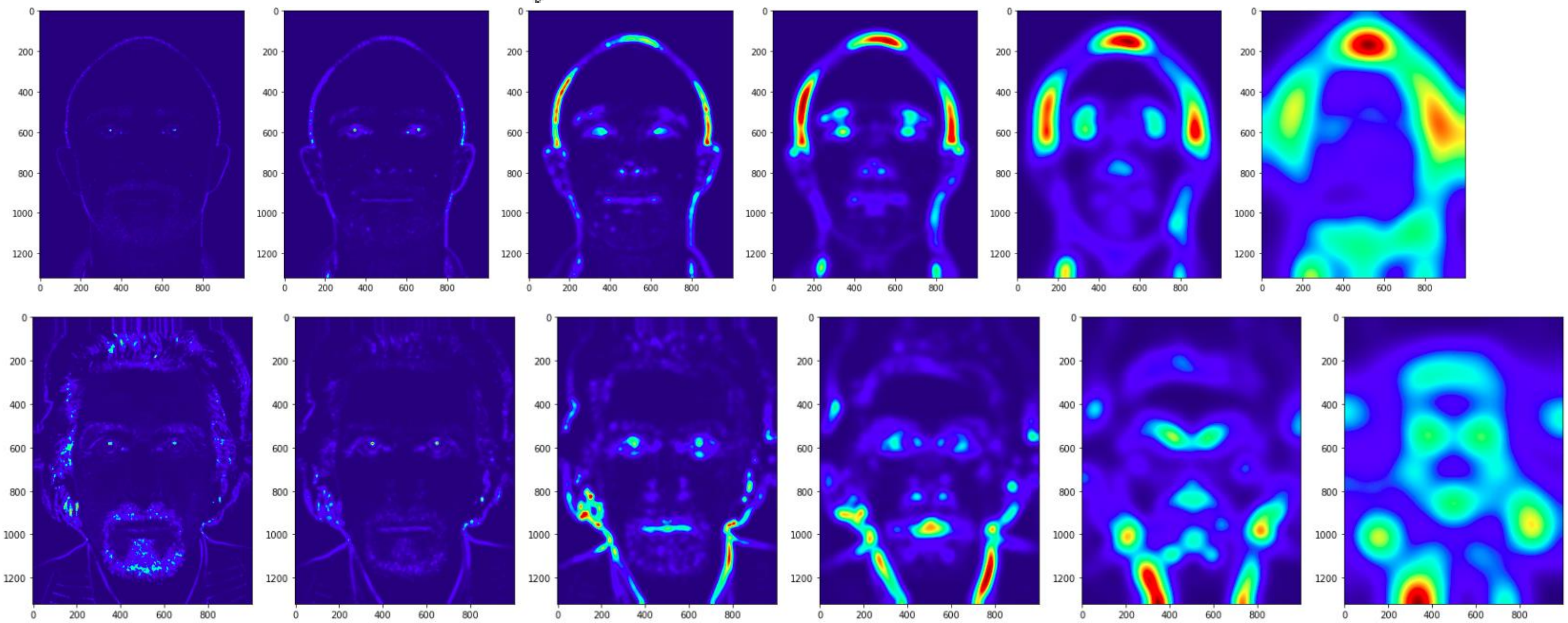
2.) Local Energy (Power):

We estimate the local energy S in each subband by the local average of the square of subband coefficients.

$$S_\ell\ [I] = L\ 2\ \ell\ [I] \otimes G(2\ \ell+1\ )$$

Now, we use the dense correspondence(wrap) on local energy.

$$S^{\hat{}}[E] = W(S_\ell[E])$$



3) Gain Map (without robust implementation):

We modify the input subbands so that they get the same energy distribution as the example subbands.

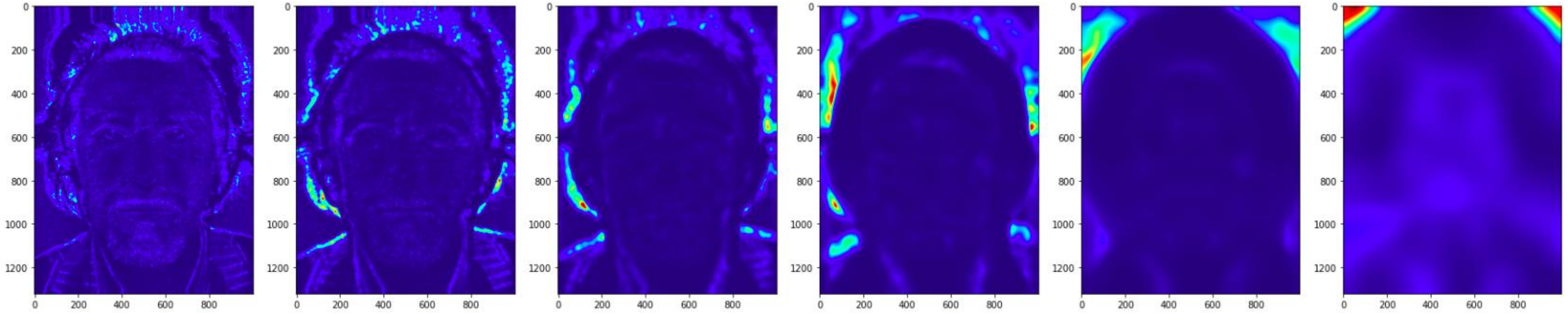
$$L_\ell[O] = L_\ell[I] \times \text{Gain}$$

$$\text{with Gain} = \sqrt{S_\ell[E]/S_\ell[I] + \epsilon}$$

It can introduce artifacts (moles or glasses in reference image) , to avoid it we robust implementation.

$$\text{RobustGain} = \max(\min(\text{Gain}, \theta_h), \theta_l) \otimes G(\beta 2\ell)$$

After this step we get style transfer in grayscale.



4) Gain Map (with robust implementation):

We modify the input subbands so that they get the same energy distribution as the example subbands.

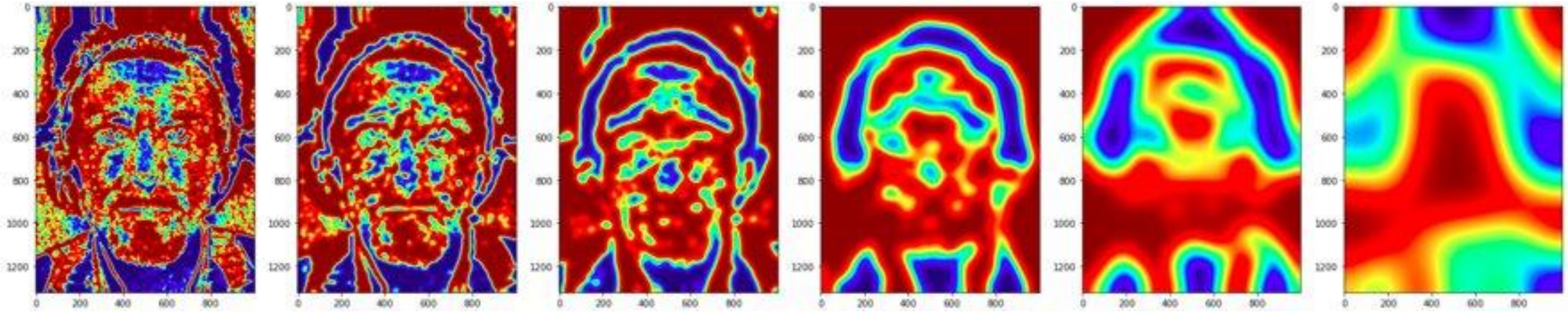
$$L_\ell[O] = L_\ell[I] \times \text{Gain}$$

with $\text{Gain} = \sqrt{S_\ell[E]/S_\ell[I] + \epsilon}$

It can introduce artifacts (moles or glasses in reference image) , to avoid it we robust implementation.

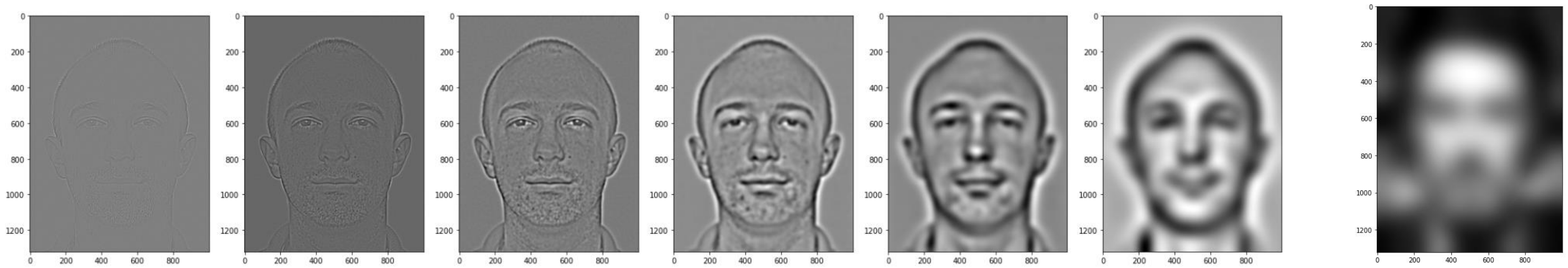
$$\text{RobustGain} = \max(\min(\text{Gain}, \theta_h), \theta_l) \otimes G(\beta 2^\ell)$$

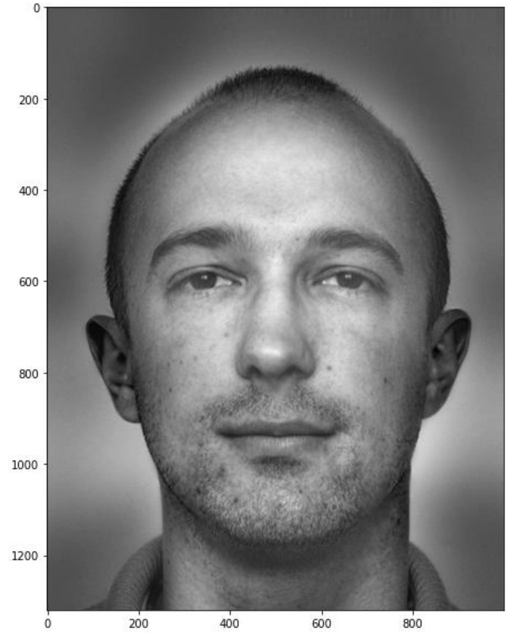
After this step we get style transfer in grayscale.



5)Aggregating the stacks

- We multiple the gain maps with the individual stacks and summing all the Laplacian stacks along with the residue to get the final style transfer image.





4.) Dealing with colors:

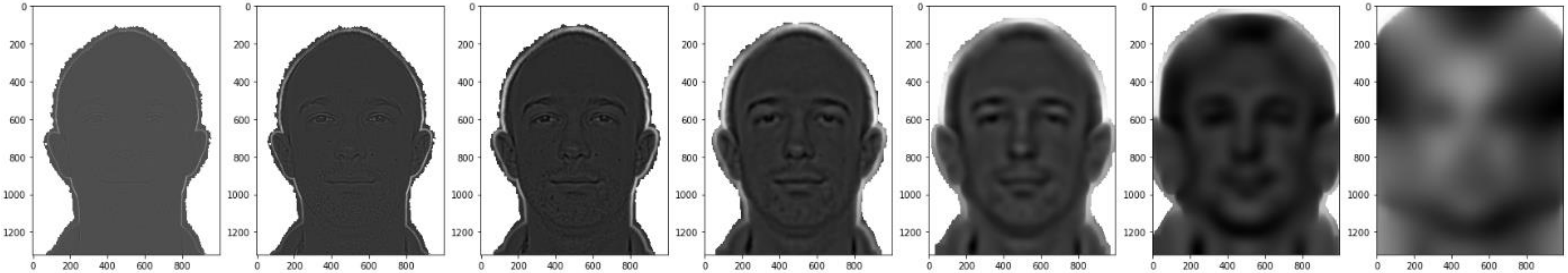
- We work in the CIE-Lab color space as it is more confirming to human perception, and process each channel independently using the algorithm that we just described.

5.) Using a face mask (GrabCut + Matting Laplacian) to preserve shape:

We truncate the Gaussian convolutions so that they only consider values within the mask.

- we run GrabCut [Rother et al. 2004] initialized with a face detection result to find a binary mask.
- we refine using the Matting Laplacian

$$Image \otimes G \;=\; (Image \times Mask) \otimes G \div \; Mask \otimes G$$

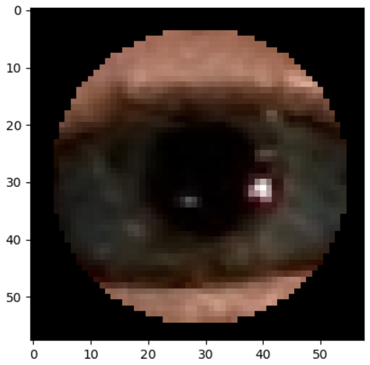


Mask obatained

Additional Post processing – Eye Highlights

1.) Locate iris using circular Hough Transform:

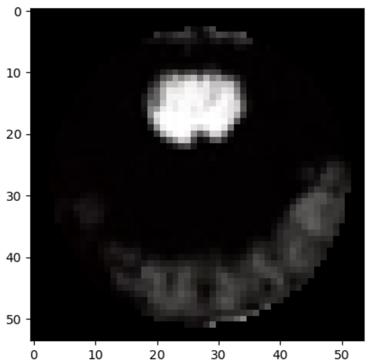
We used Hough Transform for detecting circle from a rectangular window of eye extracted from the landmark points and from the many circles detected we take the circle which is closer to the eye center.

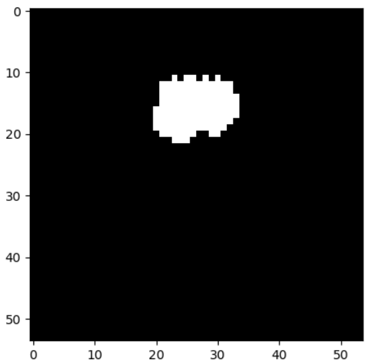


2.) K – Means with thresholding and also connected components:

Then, we create an approximate segmentation into iris, highlight, and pupil by running a k-means algorithm on the pixel colors with k = 3.

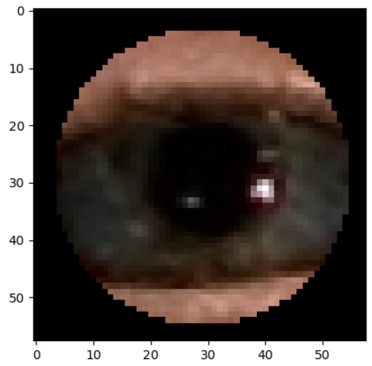
Additional observation-This can also be done by connected components.

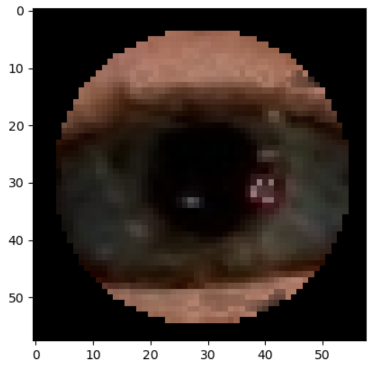




3.) Replacement by inpainting:

- Then, we erase the detected pixels in the inout image and fill in the hole using inpainting.

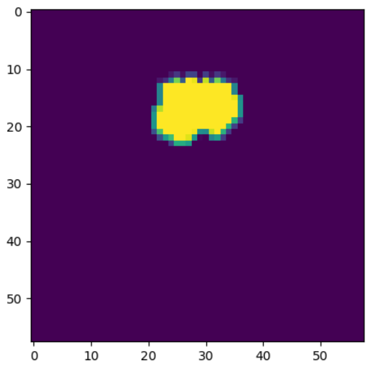


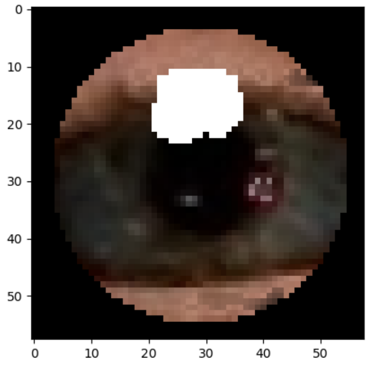


4.) Alpha Matting:

We then resize the reference eye image to the input eye image and perform alpha matting from the paper below.This step is failing because the eye width open for inpuy image is less than reference image.

LEVIN, A., LISCHINSKI, D., AND WEISS, Y. 2008. A closed-form solution to natural image matting. IEEE Trans. Pattern Analysis and Machine Intelligence 30, 2, 228–242

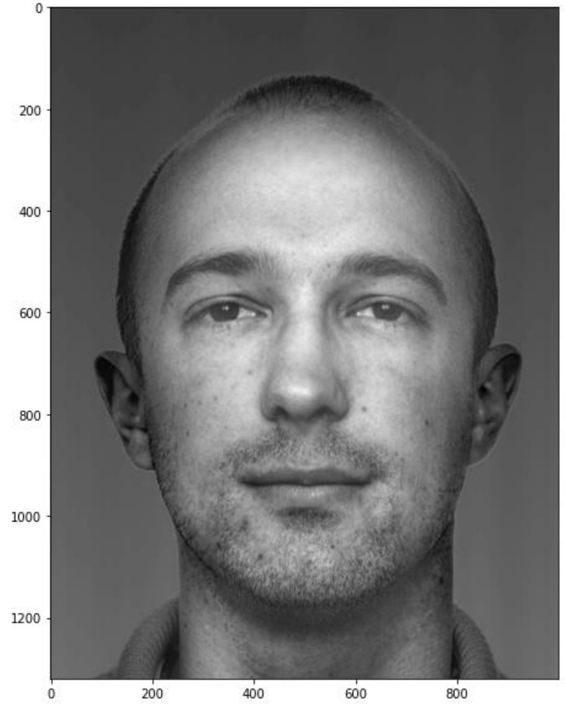




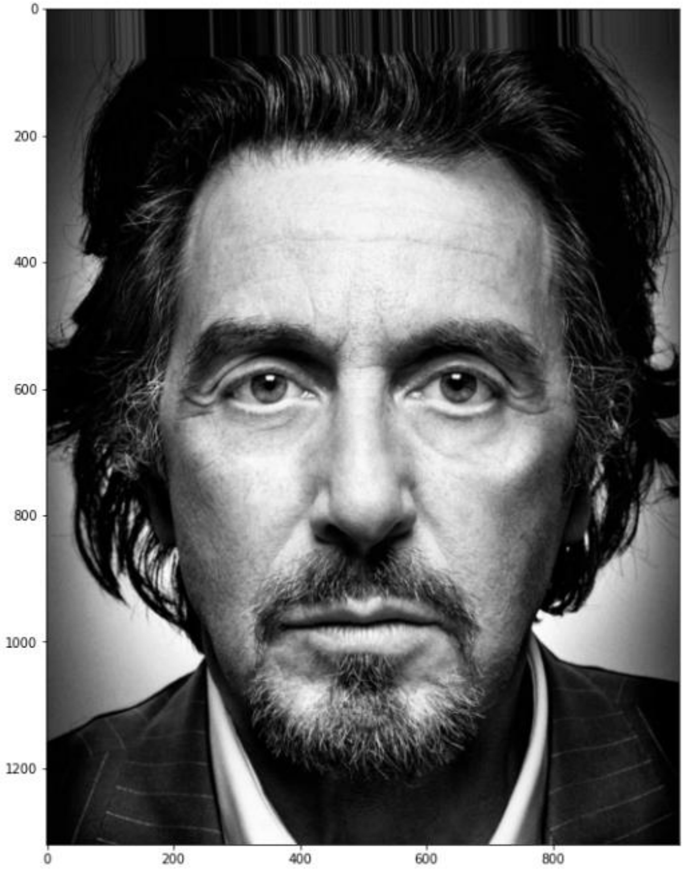
Additional Post processing – Background

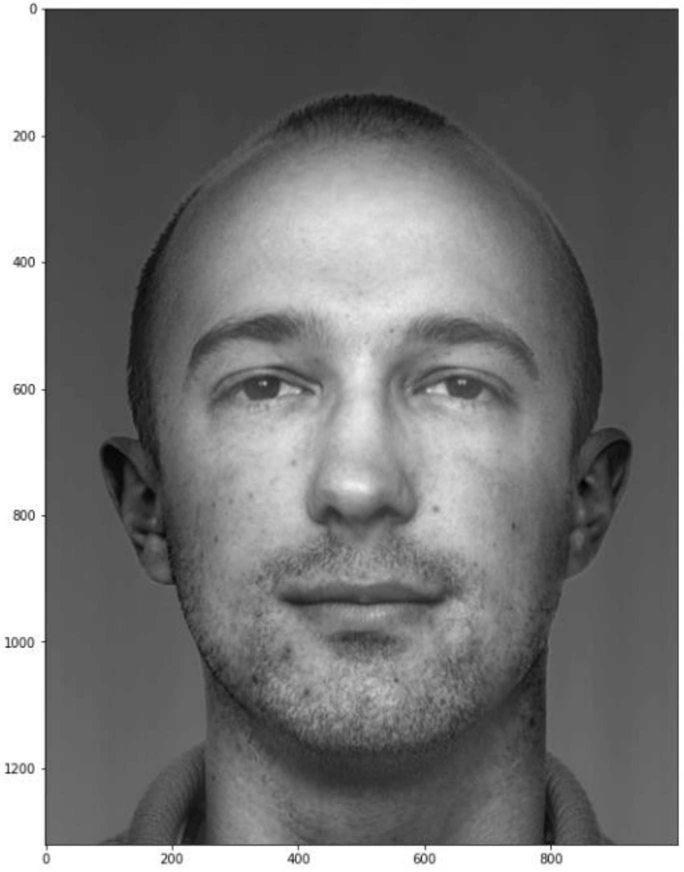
1.) Inverted face mask based background replacement (inpainting if required to fill any holes):

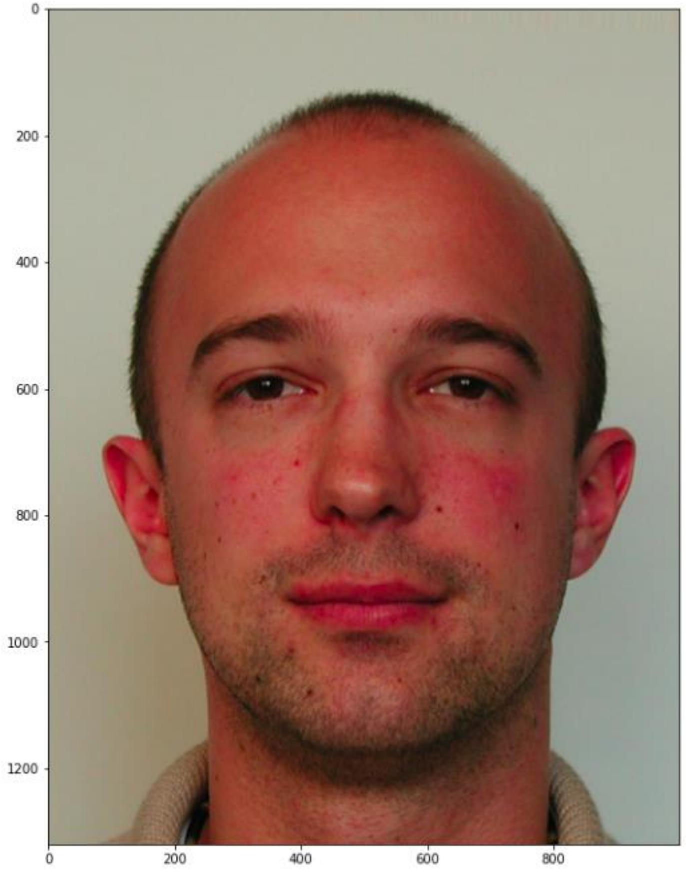
We use the previously computed masks to extract the example background and replace the input background with it. If needed, we extrapolate the missing data using inpainting is also done.



Final Output



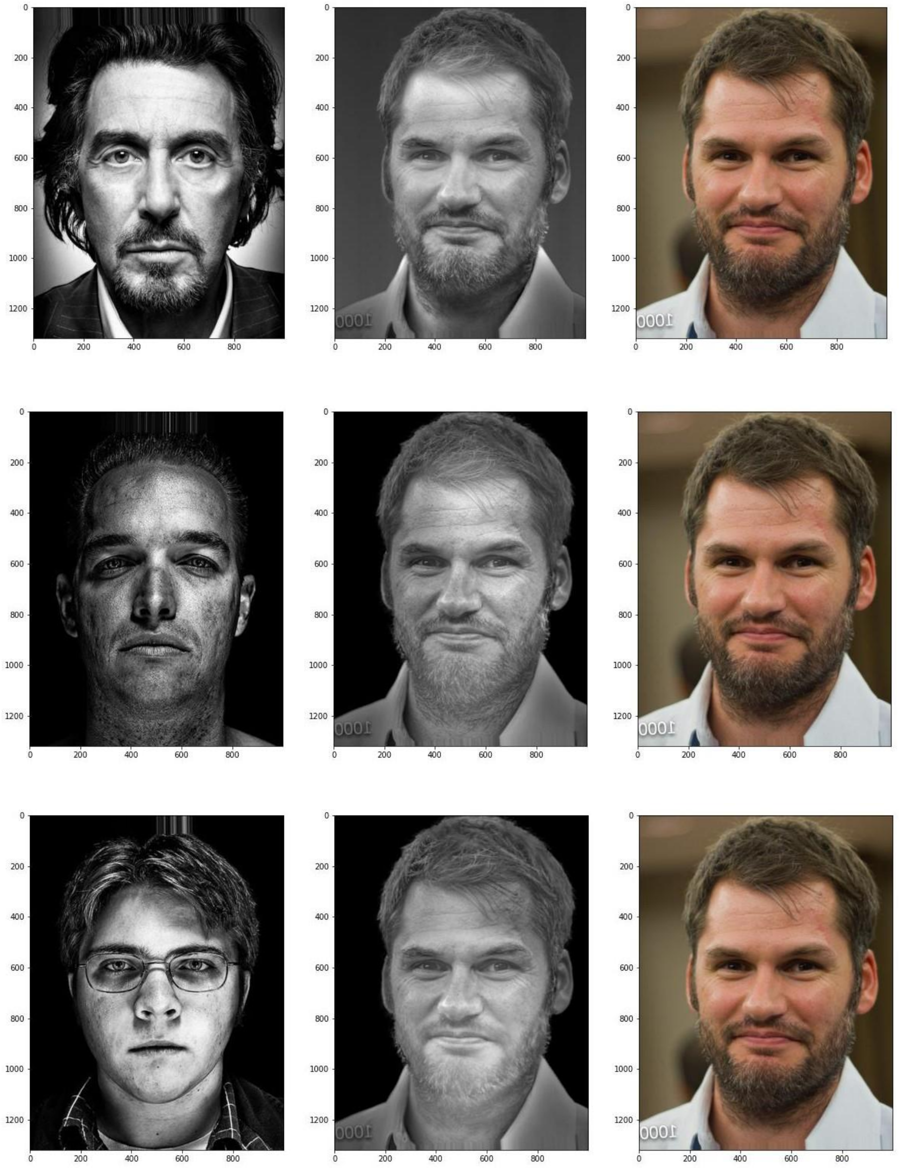




Automatic Selection of the Example in a Collection:

- It is an automatic algorithm to select a suitable example among a collection of consistently stylized headshots.
- We concatenate S_ℓ over all scales to get the feature vector representing a face image, and use the normalized cross correlation between the two feature vectors as the similarity function (paper claim it is better than L2).

0.769 0.726 0.647



Results

Running time

Running time claimed by the paper- 12 seconds

Our approch- 8 seconds

Work Division:

We have 4 major parts in our project:

1.) Dense Correspondence – Done by Chirag

2.) Multiscale Transfer of Local contrast – Done by Neeraj, Chirag and Dheeraj

3.) Eye Highlights – Done by Sarath,Neeraj

4.) Automatic selection – Done by Dheeraj