CIS581: Computer Vision and Computational Photography
Project 2: Face Morphing and Blending
Due: Oct. 17, 2017 at 3:00 pm

Instructions

- This is an **individual** project. 'Individual' means each student must hand in their **own** answers, and each student must write their **own** code in the homework. It is admissible for students to collaborate in solving problems. To help you actually learn the material, what you write down must be your own work, not copied from any other individual. You **must** also list the names of students (maximum two) you collaborated with.
- You **must** submit your code online on Canvas. We recommend that you can include a README.txt file to help us execute your code correctly. Please place your **code**, **resulting images and videos** into the top level of a single folder (no subfolders please!) named `<Pennkey>_Project2.zip`
- Your submission folder should include the following:
  - your .m or .py scripts for the required functions.
  - .m or .py scripts for generating the face morphing video.
  - any additional .m files with helper functions you code.
  - the images you used.
  - .avi files generated for each of the morph methods in face morphing.
- This handout provides instructions for two versions of the code: MATLAB and Python. You are free to select **either one of them** for this project.
- Feel free to create your own functions as and when needed to modularize the code. For MATLAB, ensure that each function is in a separate file and that all files are in the same directory. For python, add all functions in a helper.py file and import the file in all the required scripts.
- **Start early!** If you get stuck, please post your questions on Piazza or come to office hours!

Overview

This project focuses on image morphing techniques. You will produce a "morph" animation of your face into another person's face. **This is mandatory**. You can also morph your face into anything else that you wish. Use your creativity here!

You will need to generate 60 frames of animation. You can convert these image frames into a .avi movie.

A morph is a simultaneous warp of the image shape and a cross-dissolve of the image colors. The cross-dissolve is the easier part; controlling and doing the warp is the harder part. The warp is controlled by defining a correspondence between the two pictures. The correspondence should map eyes to eyes, mouth to mouth, chin to chin, ears to ears, etc., to get the smoothest transformations possible.

The triangulation used in Task 2 can be computed in any way you like, or can even be defined by hand. A Delaunay triangulation is a good choice. Recall that you need to generate only one triangulation and use it on both the sets of points. We recommend computing the triangulation at the midway shape (i.e. mean of the two point sets) to lessen the potential triangle deformations.

Notice that, apart from the functions described below, it is also required of you to implement the full script to generate face morphing videos using the triangulation model and the Thin Plate Spline model. Do **not** create your videos using the MATLAB Command Window or the IPython Console or terminal.

1 Defining Correspondences

Goal First, you will need to define pairs of corresponding points on the two images by hand (In general: The more the points, the better the morph).

For MATLAB, `cpselect` is recommended. For Python, `ginput` is recommended. There may be other functions that help you implement this functionality.

[im1_pts, im2_pts] = click_correspondences(im1, im2)

- (INPUT) im1: H1 $\times$ W1 $\times$ 3 matrix representing the first image.
- (INPUT) im2: H2 $\times$ W2 $\times$ 3 matrix representing the second image.
- (OUTPUT) im1_pts: N $\times$ 2 matrix representing correspondences coordinates in first image.
- (OUTPUT) im2_pts: N $\times$ 2 matrix representing correspondences coordinates in second image.

2 Image Morph Via Triangulation

Goal You need to write a function that produces a warp between your two images using point correspondences.

[morphed_im] = morph_tri(im1, im2, im1_pts, im2_pts, warp_frac, dissolve_frac)

- (INPUT) im1: H1 $\times$ W1 $\times$ 3 matrix representing the first image.
- (INPUT) im2: H2 $\times$ W2 $\times$ 3 matrix representing the second image.
- (INPUT) im1_pts: N $\times$ 2 matrix representing correspondences in the first image.
- (INPUT) im2_pts: N $\times$ 2 matrix representing correspondences in the second image.
- (INPUT) warp_frac: 1 $\times$ M vector representing each frame's shape warping parameter.
- (INPUT) dissolve_frac: 1 $\times$ M vector representing each frame's cross-dissolve parameter.
- (OUTPUT) morphed_im: Data structure containing output images.
  - In MATLAB, it should be an M elements cell array, where each element is a morphed image frame.
  - In Python, it should be a 4 dimensional NumPy array, with dimensions being Num $\times$ H $\times$ W $\times$ 3, where Num is the number of images.

In particular, images im1 and im2 are first warped into an intermediate shape configuration controlled by warp_frac, and then cross-dissolved according to dissolve_frac. For interpolation, both parameters lie in the range [0,1]. They are the only parameters that will vary from frame to frame in the animation. For your starting frame, they will both equal 0, and for your ending frame, they will both equal 1.

Given a new intermediate shape, the main task is to map the image intensity in the original image to this shape. As explained in class, this computation is best done in `reverse`:

- For each pixel in the target intermediate shape(called shape B from this point on), determine which triangle it falls inside. You could use inbuilt functions for this, or implement your own barycentric coordinate check. **Recommended functions for MATLAB and Python will be posted on Piazza shortly.**

- Compute the barycentric coordinate for each pixel in the corresponding triangle. Recall, the computation involves solving the following equation:

$$\begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{1}$$

  where `a, b, c` are the three corners of triangle, `(x, y)` the pixel position, and α, β, γ are its barycentric coordinate. Note you should only compute the matrix $\begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ 1 & 1 & 1 \end{pmatrix}$ and its inverse only once per triangle.

- Compute the corresponding pixel position in the source image: using the barycentric equation(eq. 1), but with three corners of the same triangle in the source image $\begin{pmatrix} a^s_x & b^s_x & c^s_x \\ a^s_y & b^s_y & c^s_y \\ 1 & 1 & 1 \end{pmatrix}$, and plug in same barycentric coordinate (α, β, γ) to compute the pixel position (x^s, y^s, z^s). You need to convert the homogeneous coordinate (x^s, y^s, z^s) to pixel coordinates by taking $x^s = x^s/z^s, y^s = y^s/z^s$.

- Copy back the pixel value at x^s, y^s the original (source) image back to the target (intermediate) image. You can round the pixel location (x^s, y^s) or use bilinear interpolation.

3 Gradient Domain Blending

For this part of the project, you will be blending images in the gradient domain as described in the paper Poisson Image Editing by Patrick Perez, Michel Gangnet and Andrew Blake. It is a gradient-domain processing technique with numerous applications such as blending, non-photorealistic rendering, contrast enhancement, texture flattening and tone-mapping.for automatically and seamlessly blending two images together. The paper discusses this technique in Section 3 named Seamless Cloning, which you are strongly advised to thoroughly read and understand before starting this project.

The goal is to seamlessly blend an object from a source image into a target image. The simplest method would be to just copy and paste the pixels from one image directly into the other. However, this will create apparent seams, even if the backgrounds are alike. We need to get rid of these seams without visually tampering the source image.

Human vision is found to be more sensitive to gradients than absolute image intensities. We formulate this problem as finding values for the output pixels that maximally preserve the gradient of the source region without altering any of the background pixels.

To begin with, we define the image that we are changing as the **target image**, the image region that we cut and want to clone as the **source region**, and the pixels in the target image that will be seamlessly cloned with the source image as the **replacement pixels**.

- Image Interpolation using a Guidance Vector Field

$$\min_f \iint_\Omega |\nabla f - \mathbf{v}|^2 \text{ with } f|_{\partial\Omega} = f^*|_{\partial\Omega} \tag{2}$$

  where $\nabla = [\frac{\partial}{\partial x}, \frac{\partial}{\partial y}]$ is the gradient operator, f is the function of the blending image, f^* is the function of the target image, $\mathbf{v}$ is the vector field or the gradient field of the source image, Ω is the region of blending and $\partial\Omega$ is the boundary of the blending region.
  We solve this interpolation problem (Poisson equations) for each color channel independently.

- Discrete Poisson Solver
  The variational problem in equation (6) is discretized to obtain a quadratic optimization problem.

$$\min_{f|_\Omega} \sum_{\langle p,q \cap \Omega \neq \emptyset \rangle} (f_p - f_q - v_{pq})^2, \text{ with } f_p = f^*_p \text{ for all } p \in \partial\Omega \tag{3}$$

where N_p is the set of 4-connected neighbors for pixel p, $\langle p,q \rangle$ denote a pixel pair such that $q \in N_p$, f_p is the value of f at p and $v_{pq} = g_p - g_q$ for all $\langle p,q \rangle$.

The solution satisfies the following simultaneous linear equations:

$$\text{for all, } p \in \Omega, |N_p| f_p - \sum_{q \in N_p \bigcap \Omega} f_q = \sum_{q \in N_p \bigcap \partial\Omega} f_q^* + \sum_{q \in N_p} v_{pq} \tag{4}$$

$$|N_p| f_p - \sum_{q \in N_p} f_q = \sum_{q \in N_p} v_{pq} \text{ for pixels interior to } \Omega, \text{ i.e. } N_p \subset \Omega \tag{5}$$

We need to solve for f_p from the given set of simultaneous linear equations. If we form all the f_p as a vector x, then the given set of equations can be converted into a linear system of $Ax = b$. Please do note that not all of f_q is unknown. It is possible that $q \in N_p$ and also $q \in \partial\Omega$, in which case, $f_q = f_q^*$ and it becomes a known parameter.

3.1 Align the Source Image and Create its Mask

First, you need to align the source image and target image. Please use any image editor to adjust the size and position of the source image, ensuring that the region of the target image you want to replace is well-aligned with the source image. Then, save the resized source image and the coordinate of its top left corner as an offset. From now on, source image will refer to the resized source image.

Complete the following function to create an image mask - a logical matrix representing the pixels you want to replace in the source image. A value of 1 means that we will be using the pixel whereas a value of 0 means that the pixel will not be used.
We recommend using MATLAB's function `imfreehand` and `createMask`.

```
[mask]=maskImage(img)
```

- (INPUT) `img`: $h \times w \times 3$ matrix representing the source image.
- (OUTPUT) `mask`: $h \times w$ matrix representing the logical mask

3.2 Index the Pixels

The intensity of the replacement pixels in the target pixel can be found using the linear system $Ax = b$. But, not all the pixels need to be computed. Only the pixels masked as 1 in the logical mask will be used to blend. In order to reduce the number of calculations, you need to index the replacement pixels such that each element in x represents one replacement pixel. As shown in Figure 3, the yellow locations are the replacement pixels (indexed from left to right).

Complete the following function to obtain the indexes of the replacement pixels:

```
[indexes] = getIndexes(mask, targetH, targetW, offsetX, offsetY)
```

- (INPUT) `mask`: The logical matrix $h \times w$ representing the replacement region.
- (INPUT) `targetH`: The height of the target image, h'
- (INPUT) `targetW`: The width of the target image, w'
- (INPUT) `offsetX`: The x-axis offset of the source image with respect to the target image.
- (INPUT) `offsetY`: The y-axis offset of the source image with respect to the target image.
- (OUTPUT) `indexes`: $h' \times w'$ matrix representing the indices of each replacement pixel. The value 0 means that is not a replacement pixel.

(a) Source Image

(b) Target Image

(c) Blended Image

3.3 Compute the Coefficient Matrix

As described in the section 3.2, the intensities of the replacement pixels are obtained by solving $Ax = b$. In this section, you need to generate the the Coefficient Matrix A. Please note that the Coefficient Matrix is of size $N \times N$, where N is the number of replacement pixels. In order to reduce the memory of this matrix, you will have to use a sparse matrix.
Complete the following function to compute the Coefficient Matrix:

```
[coeffA] = getCoefficientMatrix(indexes)
```

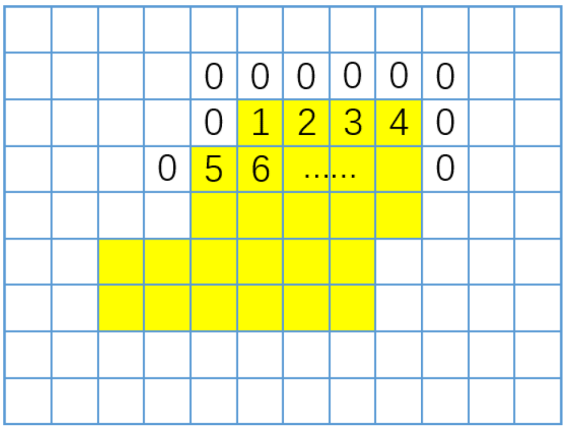
- (INPUT) `indexes`: $h' \times w'$ matrix representing the indices of each replacement pixel.
- (OUTPUT) `coeffA`: an $N \times N$ sparse matrix representing the Coefficient Matrix, where N is the number of replacement pixels.

3.4 Compute the Solution Vector

Complete the following function to generate the solution vector b in the linear system $Ax = b$.

```
[solVectorb] = getSolutionVect(indexes, source, target, offsetX, offsetY)
```

- (INPUT) `indexes`: $h' \times w'$ matrix representing the indices of each replacement pixel.
- (INPUT) `source`: $h \times w$ matrix representing one color channel of the source image.
- (INPUT) `target`: $h' \times w'$ matrix representing one color channel of target image.
- (INPUT) `offsetX`: The x-axis offset of the source image with respect to the target image.



3.2 Indexing the Pixels

- (INPUT) `offsetY`: The y-axis offset of the source image with respect to the target image.
- (OUTPUT) `solVectorb`: $1 \times N$ vector representing the solution vector.

3.5 Seamlessly Clone the Image

Once you have obtained A and b as stated above, solve for vector x. You will need to replace the pixels in question with the updated intensity i.e. clone the image and obtain the resulting image. Complete the following function to obtained the composite image:

```
[resultImg] = reconstructImg(indexes, red, green, blue, targetImg)
```

- (INPUT) `indexes`: $h' \times w'$ matrix representing the indices of each replacement pixel.
- (INPUT) `red`: $1 \times N$ vector representing the intensity of the red channel replacement pixel.
- (INPUT) `green`: $1 \times N$ vector representing the intensity of the green channel replacement pixel.
- (INPUT) `blue`: $1 \times N$ vector representing the intensity of the blue channel replacement pixel.
- (INPUT) `targetImg`: $h' \times w' \times 3$ matrix representing the target image.
- (OUTPUT) `resultImg`: $h' \times w' \times 3$ matrix representing the resulting cloned image.

3.6 Wrapper Function

After you complete all the above functions, you will need to write a wrapper function and a demo script. In this function named `seamlessCloningPoisson.m`, call `getIndexes.m`, `getCoefficientMatrix.m`, `getSolutionVect.m` and `reconstructImg.m` and solve the linear system. We recommend the MATLAB function `mldivide`.

```
[resultImg] = seamlessCloningPoisson(sourceImg, targetImg, mask, offsetX, offsetY)
```

- (INPUT) `sourceImg`: $h \times w \times 3$ matrix representing the source image.
- (INPUT) `targetImg`: $h' \times w' \times 3$ matrix representing the target image.
- (INPUT) `mask`: The logical matrix $h \times w$ representing the replacement region.
- (INPUT) `offsetX`: The x-axis offset of the source image with respect to the target image.
- (INPUT) `offsetY`: The y-axis offset of the source image with respect to the target image.
- (OUTPUT) `resultImg`: $h' \times w' \times 3$ matrix representing the resulting cloned image.

Finally, write a script to generate your blended image using `seamlessCloningPoisson.m` and `maskImage.m`

Please use your creativity while creating cloned images.

4 Extra Credit Challenge:

The following challenges are extra credit challenges. Implementation of the following challenges are optional.

NOTE: The following challenges are the ones that we have in mind so far. We may release more before the deadline. The latest potential release date is a week before the deadline. Stay tuned on Piazza!

• 4.1 Image Morphing Via Thin Plate Spline

Goal: Implement the same function as in Task 2 except using the Thin Plate Spline model.

For this part, you need to compute thin-plate-spline that maps from the feature points in the intermediate shape (B) to the corresponding points in original image (A). Recall you need two of them - one for the x coordinate and one for the y coordinate. A thin plate spline has the following form:.

$$f(x,y) = a_1 + a_x \cdot x + a_y \cdot y + \sum_{i=1}^{p} w_i U\left(||(x_i,y_i) - (x,y),||\right) \tag{6}$$

where $U(r) = -r^2 log(r^2)$.

We know there is some thin-plate spline (TPS) transform that can map the corresponding feature points in image (B) back to image (A), using the same TPS transform we will transform all of the pixels of image (B) into image (A), and copy back their pixel value.

You need to implement three functions:

- Thin-plate parameter estimation:

```
[a1,ax,ay,w] = est_tps(interim_pts, source_pts)
```

  - (INPUT) interim_pts: N $\times$ 2 matrix, each row representing corresponding point position (x,y) in second image.
  - (INPUT) source_pts: N $\times$ 1 vector representing corresponding point position x or y in first image.
  - (OUTPUT) a1: double, TPS parameter.
  - (OUTPUT) ax: double, TPS parameter.
  - (OUTPUT) ay: double, TPS parameter.
  - (OUTPUT) w: N $\times$ 1 vector, TPS parameters.

Recall the solution of the TPS model requires solving the following equation:

$$\begin{pmatrix} K & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} w1 \\ w2 \\ \cdots \\ wp \\ ax \\ ay \\ a1 \end{pmatrix} = \begin{pmatrix} v1 \\ v2 \\ \cdots \\ vp \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{7}$$

where

$$K_{ij} = U(||(x_i,y_i) - (x_j,y_j)||), \tag{8}$$

$v_i = f(x_i, y_i)$, and i^{th} row of P is $(x_i, y_i, 1)$. K is a matrix of size p by p, and P is a matrix of size p by 3. In order to have a stable solution you need to compute the solution using:

$$\begin{pmatrix} w1 \\ w2 \\ \dots \\ wp \\ ax \\ ay \\ a1 \end{pmatrix} = inv(\begin{pmatrix} K & P \\ P^T & 0 \end{pmatrix} + \lambda \star I(p+3, p+3)) \begin{pmatrix} v1 \\ v2 \\ \dots \\ vp \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{9}$$

where $I(p+3, p+3)$ is an identity matrix of size $p+3$ and $\lambda \geq 0$ (usually close to zero).
NOTE: You need to compute two TPS, one by plugging in the x-coordinates as v_i, and one by plugging in the y-coordinates as v_i.

- `morphed_im = obtain_morphed_tps(im_source, a1_x, ax_x, ay_x, w_x, a1_y, ax_y, ay_y, w_y, interim_pts, sz)`
  - (INPUT) im_source: $H_s \times W_s \times 3$ matrix representing the source image.
  - (INPUT) a1_x, ax_x, ay_x, w_x: the parameters solved when doing est_tps in the x direction.
  - (INPUT) a1_y, ax_y, ay_y, w_y: the parameters solved when doing est_tps in the y direction.
  - (INPUT) interim_pts: $N \times 2$ matrix, each row representing corresponding point position (x, y) in target image.
  - (INPUT) sz: 1×2 vector representing the target image size (H_t, W_t).
  - (OUTPUT) morphed_im: $H_t \times W_t \times 3$ matrix representing the morphed image.
- `morphed_im = morph_tps(im1, im2, im1_pts, im2_pts, warp_frac, dissolve_frac)`
  - (INPUT) im1: target image
  - (INPUT) im2: source image
  - (INPUT) im1_pts: correspondences coordinates in the target image
  - (INPUT) im2_pts: correspondences coordinates in the source image
  - (INPUT) warp_frac: a vector contains warping parameters
  - (INPUT) dissolve_frac: a vector contains cross dissolve parameters
  - (OUTPUT) morphed_im: a set of morphed images obtained from different warp and dissolve parameters. The size should be [number of images, image height, image Width, color channel number]

In this step, you need transform all the pixels in image (B) by the TPS model, and read back the pixel value in image (A) directly. The position of the pixels in image A is generated by using equation 2.

• 4.2 Face Blending:

Using the sequence of images during image morphing, seamlessly clone a source face onto a target face and create a video. The recommended format is .avi.

5 Test and Submission

- Use different kinds of images. Face to face morphing is mandatory but get creative! Morph objects to objects, even face to objects. Creative morphs will receive the honor of public recognition on Piazza.
- We have provided a test script for MATLAB and Python for this project. Extract the contents of the test script to the same directory as your functions and run `Test_script.m/py` in MATLAB/Python. When grading, we'll be calling your functions in the same manner, so make sure they work as you'd expect on the sample in the test script.
- Collect all your source code files and test images into a folder named as `<Pennkey>_Project2`. Zip this folder and submit it to Canvas. Any break in this rule will lead to a failure in the test script. Only submit codes pertaining to your language of implementation. For example: If you choose to do the project in Python, do not submit the MATLAB folder containing the MATLAB starter codes.