EE 569: Homework #4

Issued: 11/3/2013 Due: 11:59PM, 12/01/2013

General Instructions:

Please refer to Homework Guidelines and MATLAB Function Guidelines for more information about how to complete and submit the homework. Also, refer to the USC policy on academic integrity and penalties for cheating and plagiarism - these rules will be strictly enforced. **If you make any assumptions about a problem, please clearly state them in your report.**

Problem 1: Face Warping (35 %)

In this problem, you will use spatial warping techniques to morph an image into another image.

1(a) Back to Baby (Basic: 12%)

In this part, you are given two images, "drew.raw" and "baby_drew.raw", which are two faces of the same person. The size of each image is 300 x 350 as shown in Figure 1.

(a) "drew.raw" (b) "baby_drew.raw"

Figure 1: Component image for Problem 1(a)

Before conducting facial morphing, the coordinates of key features points on each face need to be properly aligned. **Please select proper control points for the alignment of facial features.** The control points will divide the image into different triangular regions, and spatial warping will be conducted in each sub-region.
If the feature (e.g. eye, nose, lips) you select is located at (x_m, y_m) in the 'drew.raw' image and (x_t, y_t) in

'baby_drew.raw', then after spatial warping they should be both align at $(\frac{x_m+x_t}{2}, \frac{y_m+y_t}{2})$. Figure 2 shows the warping process for the feature "eye-center-point". ("ECP" - the middle point between two eyes)

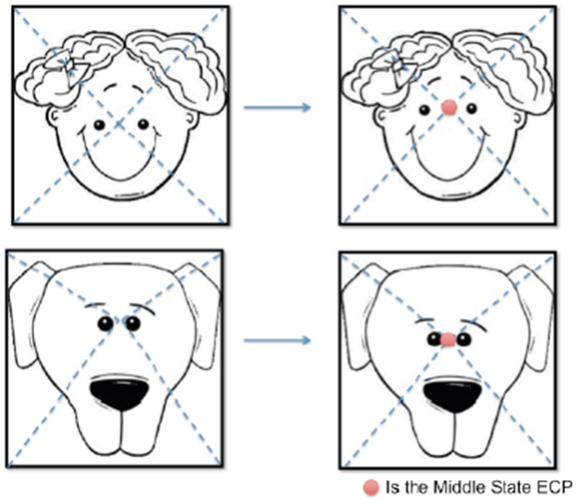


Figure 2: Illustration of using eye-center-point as control point for spatial warping.

Describe the details of your selected control points in your report. This should include what types of feature you have selected and the total number of control points. Print out your result for both images after facial warping.

1(b) Face Morphing (Basic: 5%)

After performing face warping, the key feature points of both images are well aligned. To complete facial morphing, the next step is to do cross-dissolving between the two images, which is described as following:

$$I_{out}(i,j) = (1-\alpha) \cdot I_{drew}(i,j) + \alpha \cdot I_{baby\_drew}(i,j), \quad 0 \le \alpha \le 1 \tag{1}$$

Increasing the value of α from 0 to 1, you will see the gradual transformation from "drew.raw" to "baby_drew.raw". Please print out the intermediate image for $\alpha = 0.2$, 0.5 and 0.8, and discuss your morphing result. Is the transformation smooth and reasonable? Why? How could the morphing result be further improved?

1(c) Advanced Morphing: From Bruce Banner to Hulk (Advanced: 18%)

Your implemented morphing algorithm should be robust enough to easily handle other morphing tasks. Please use the control points (same feature, but different coordinates) to conduct morphing from "bruce_banner.raw" and "hulk.raw", which are shown in Figure 3. Please print out the intermediate image for $\alpha = 0.2$, 0.5 and 0.8. Discuss about the weak points of your algorithm based on the result.

As you might have already discovered, the approach described in 1(a)-(b) still has much room for further improvement. A better facial morphing algorithm would allow more feature points to be aligned with each other. We encourage you to improve the facial morphing result of the same images using your own approach. Your method could be based on an existing approach, but it needs to demonstrate improved result. Please describe your approach in detail, show the results, and discuss how/why it outperforms the previous one. You will be graded by your effort and morphing result.

(a) "bruce_banner.raw"

(b) "hulk.raw"

Figure 3: Component image for Problem 1(a)

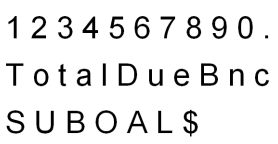
Problem 2: Optical Character Recognition (35 %)

Your friend Tommy Trojan requests you to implement an optical character recognition system, which he plans to integrate into his smartphone app. He is interested in recognizing the total amount paid in a "restaurant bill" and calculating the 'gratuity' that needs to be added to the total amount. He asks you to develop an OCR system by answering the following questions.

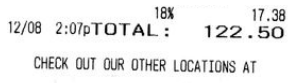
2(a) OCR Segmentation and Training (Basic: 15%)

Tommy gives you his training image which has some symbols - (some English alphabets, all the digits and some special symbols). The training data is shown in "training-bill.raw". Using this training data, please build an OCR program using the following steps.

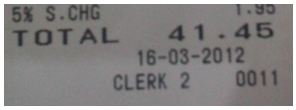
1. First, segment the different characters by your own program.
2. For each segmented image, learn its properties. Here are some properties taught in class or described in the textbook. You may use these properties as the features: [1] Area, [2] Perimeter, [3] Euler Number, [4] Circularity, [5] Spatial Moment, [6] Symmetry, and [7] Aspect Ratio.
3. Learn a decision tree for the segmented symbols using the features extracted in the above step.



(a) training-bill.raw



(b) testOCR-bill1.raw



(c) testOCR-bill2.raw

Figure 4: OCR training and testing

Please describe the features built for your OCR program, and illustrate your decision tree of classification in the report.

2(b) OCR Testing (Basic: 10%)

After your OCR system is built, please apply it to two simple test images ("testOCR-bill1.raw" and "testOCR-bill2.raw"). Since each document image contains multiple symbols of different fonts and sizes,

you may have to perform a primitive segmentation to separate each symbol in the test images. Note: segmentation of the image should be done automatically using your own program.
Please find the "Total amount" shown in the two test images and show your results in the report. Please list any assumptions you make for this problem. How good is your system?

2(c) OCR Testing on 'Restaurant bill' (Advanced: 10%)

Tommy wants to test your OCR system built in part (a) to read the "restaurant bills" captured on your smartphone, so that your program can directly add ' x% gratuity' to the total amount. For this task, he gives you a real "restaurant bill" as a shown in figure 5. He asks you to find the answers for the following questions using your OCR system.

1. Find the 'total amount' present on the restaurant bill.
2. Find the 'amounts' for all the items present in the bill.

[Hints will be provided during the discussions]

Figure 5: "Real Restaurant Bill" test image

Problem 3: Image Segmentation (30 %)

In Homework No. 3, you implemented texture segmentation. In this assignment, you will perform image segmentation for general images as shown in Figure 6.

(a) seg1.raw (b) seg2.raw (c) seg3.raw

Figure 6: Images for segmentation.

3(a) Image segmentation using K-means (Basic: 15%)

In this part, you will implement K-means clustering on color features of image pixels to accomplish the image segmentation task.

1. Let $[R(i)\ G(i)\ B(i)]^T$ be a feature vector for pixel i consisting of its red, green and blue intensity values. Please use this feature vector to segment the images using k-means clustering algorithm. Please try several cluster numbers for each image in Figure 6.

2. Propose a new feature for each pixel and use it along with $[R(i)\ G(i)\ B(i)]^T$ for image segmentation using k-means. (Note: You cannot use pixel location $i = (x, y)$ as a feature).

3(b) Image segmentation using mean shift filtering (Advanced: 15%)

Mean shift is one of the state-of-the-art image segmentation techniques. It was first proposed by Comaniciu and Meer [1] in 2002. In this part, you will conduct image segmentation with mean shift filterig on the three test images. Please implement (write your own code) mean-shift filtering (see Section 4.1.1 of [1]) with a Gaussian kernel function. Select a couple values for the kernel parameters and discuss their effect on the segmentation results. Compare the obtained results with those in Part 3(a) and explain the reason for superior performance of mean shift filtering.
Hint: The algorithm would be presented in the discussion section.

3(c) Image segmentation using mean shift clustering (Bonus: 10%)

Implement an image segmentation method using mean shift clustering as described in Section 4.2.1 of [1] and apply it to the three test images.

Appendix:

Problem 1: Face Warping

drew.raw	$300 \times 350 \times 3$, 8-bit depth, inteleaved RGB image
baby_drew.raw	$300 \times 350 \times 3$, 8-bit depth, inteleaved RGB image
bruce_banner.raw	$400 \times 400 \times 3$, 8-bit depth, inteleaved RGB image
hulk.raw	$400 \times 400 \times 3$, 8-bit depth, inteleaved RGB image

Problem 2: Optical Character Recognition

training-bill.raw	$600 \times 300 \times 3$, 8-bit depth, interleaved RGB image
testOCR-bill1.raw	$272 \times 82 \times 3$, 8-bit depth, interleaved RGB image
testOCR-bill2.raw	$901 \times 316 \times 3$, 8-bit depth, interleaved RGB image
restaurant-bill.raw	$397 \times 564 \times 3$, 8-bit depth, interleaved RGB image

Problem 3: Image Segmentation

seg1.raw	$481 \times 321 \times 3$, 8-bit depth, interleaved RGB image
seg2.raw	$481 \times 321 \times 3$, 8-bit depth, interleaved RGB image
seg3.raw	$481 \times 321 \times 3$, 8-bit depth, interleaved RGB image

Sample Codes Provided

readraw.m	MATLAB source code provided to read in grayscale raw image files
writeraw.m	MATLAB source code provided to output grayscale raw image files
readraw.cpp	C++ code provided to read in and output grayscale raw image files

Reference Images

All images used in this homework were downloaded either from Google image search (www.google.com/imghp) or USC-SIPI image database (sipi.usc.edu/database).

References

[1] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 5, pp. 603–619, 2002.