Homework #1 Single-site Operations and Morphology Operations

CAP 5415, Computer Vision, fall, 2016
Department of Computer Science, Florida State University

Points: 95
Due: At the beginning of class on Monday, September 19, 2016
Submission: Hardcopy along with source codes/programs attached

Problem 1 (20 points) Write a set of functions to 1) compute the histogram of an image and 2) compute and apply the point operation to a given image so that its histogram will be as flat as possible. The last function accepts one additional parameter, which is the number of different values in the resulting image. Then use your functions to do the following on the given image (available at http://www.cs.fsu.edu/~liux/courses/cap5415-2016/assignments/pictures/fsu_Westcott.jpg).

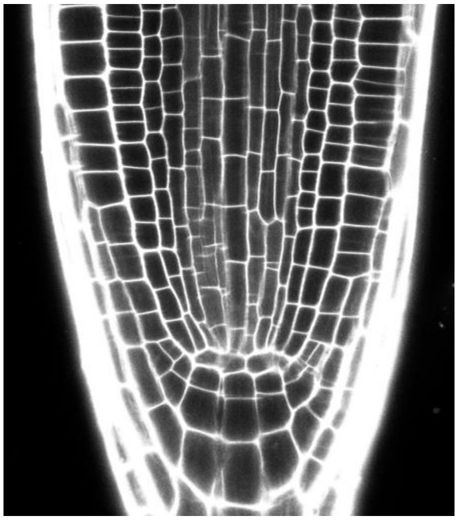
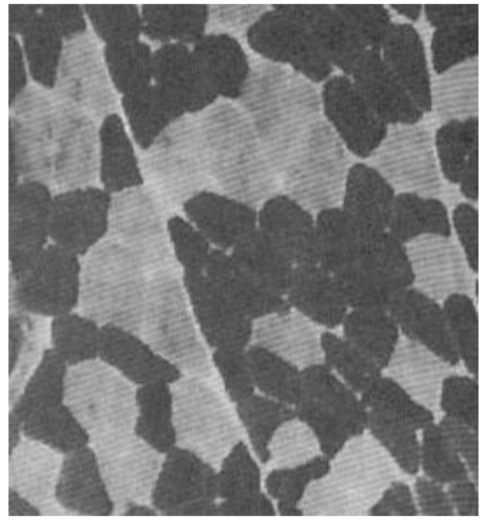
1) Compute and plot its histograms of each channel. Are the histograms consistent with your perception of the color distributions in the image? Give a brief justification.

2) Transform the above image so that its histogram is as flat as possible in each channel. Here the resulting image should have 32 different values in each channel. You need to show the histograms of the original image and the histograms of the resulting image.

Problem 2 (**20 points**) Answer the following questions regarding Otsu's method for thresholding as we discussed in class.

1) Give an incremental implementation of the Otsu's method, where the variables are updated for the next threshold based on the previous results.

2) Use your function to segment the following images (available at http://www.cs.fsu.edu/~liux/courses/cap5415-2016/class-only/root-image.pgm and http://www.cs.fsu.edu/~liux/courses/cap5415-2016/class-only/muscle_fiber.pgm), i.e., to segment the images into foreground and background. You need to give the optimal threshold explicitly and the corresponding segmentation result.

Problem 3 (25 points) Content-based image retrieval (CBIR) is potentially critical for finding similar images in large image datasets with medical and other applications. One of the early CBIR methods is to use color histograms to represent images and compare them by comparing their color histograms using histogram intersection. To quantify performance of a CBIR system, two measures are typically used: precision and recall. For a query image, suppose *n* images are returned to the user, the precision is the percentage of *n* images that are relevant and the recall is the ratio of the number of relevant images among the *n* images to the total number of relevant images in the whole dataset. Note for each *n*, we have a precision measure and a recall measure and by varying *n* (from 1 to all the images in the dataset), we can generate a precision-recall plot for each query image.

As a naïve CBIR system, suppose we have three classes (labeled as c1, c2, and c3) and each class has 10 images (named "0.ppm" to "9.ppm" under a subdirectory named by the corresponding class label) and we assume that for a query image all the images in the same class are relevant. Write necessary functions to compute and plot the precision-recall plot for image "1.ppm" in the three classes (i.e., the precision-recall plot for "c1/1.ppm", "c2/1.ppm", and "c3/1.ppm" respectively).

The images (under subdirectories "c1", "c2", and "c3") for this assignment are available on the class web site at http://www.cs.fsu.edu/~liux/courses/cap5415-2016/class-only with the same file names and you can not post them on the web.

Problem 4 (15 points) A common way to visualize the effect of a geometric transformation is by deforming a regular grid using the transformation (for example, see Figure 3.51 in the book). Using a regular grid of $-1 \leq x, y \leq 1$ with a step size of 0.1, show the effect of the following spatial transformation:

1) Rotation of -30°.

2) Scaling of $x' = 0.75x$ and $y' = 0.60y$.

3) Affine transformation of $\begin{array}{l} x' = 1.25x - 0.25y + 0.15 \\ y' = 0.25x + 0.80y - 0.25 \end{array}$.

4) Perspective transformation of $\begin{aligned} x' &= \dfrac{x}{2y/3+4/3} \\ y' &= \dfrac{1.5y}{2y/3+4/3} \end{aligned}$.

5) **Extra credit option (5 points)**. Thin-plate spline: Use the thin-plate spline to generate the deformed grid in Figure 3.51 (d) using the correspondences given in Figure 3.51 (b). (Matlab implementation of computing thin-plate spline given correspondences is available at http://www.cs.fsu.edu/~liux/courses/cap5415-2016/class-only/RPM-TPS/ctps_gen.m and the function of computing the thin-plate spline transformation is available at http://www.cs.fsu.edu/~liux/courses/cap5415-2016/class-only/RPM-TPS/crbf_warp_pts.m).

Problem 5 (**15 points**) A camera takes an image I of a penny, a dime, and a quarter lying on a white background and the coins do not overlap. Suppose that thresholding creates a binary image B successfully with 1 for the coin regions and 0 for the background. You are given the known diameters the coins d_p, d_d, and d_q in pixels (note that $d_d < d_p < d_q$). Using morphology operations (dilation, erosion, opening, and closing) and logical and set operations (AND, OR, NOT, and set difference), show how to produce three binary output images P, D, and Q. P should contain just the penny, D should contain just the dime, and Q should contain just the quarter.

Extra credit option (**5 points**): Take an image of three real coins (a penny, a dime, and a quarter) and apply your operations on the image. You need to show the original image, and the generated binary image by thresholding, the three resulting images, and any necessary intermediate images to show your steps. You can use an image available at http://www.cs.fsu.edu/~liux/courses/cap5415-2016/class-only/quarter-penny-dime.JPG .