CS 484, Fall 2017
Homework Assignment 1: Binary Image Analysis

Due: October 31, 2017

The goal of this assignment is to find objects of interest in images using binary image analysis techniques.

Question 1 (20 pts)

Write your own implementations of the morphological `dilation` and `erosion` operations. Your programs should input a binary image (as a matrix) and a structuring element (also as a matrix), and produce a binary image (another matrix) as the result of the operation.

You should generate the structuring element as a binary image with an arbitrary shape. Given the structuring element, your code should implement the `dilation` and `erosion` operations using the definitions given in the course slides. Note that the structuring element should be created (as a matrix) outside and given as an input to the dilation/erosion codes so that your code can work with any kind of structuring element.

You are free to use any programming language for the implementation. The representation of the image data and the structuring element data (using data structures such as arrays) will depend on your choice of the language. You MUST write your own implementations of these two morphological operations. Code from other sources is NOT allowed for this part of the assignment (as an exception, you can use the `strel` function in Matlab or the `getStructuringElement` function in OpenCV to generate the arrays containing the structuring elements). Contact the instructor or the TA if you need help with image I/O using different programming languages and environments.

Submit: Well-documented source code in ASCII format for `dilation` and `erosion` operations. Also cite the definition you used for the implementation in the code documentation.

Question 2 (40 pts)

License plates from different countries are shown in Figure 1. The objects of interest are the letters and digits that are part of the license plate number. The license plate images contain other parts such as logos, signs, country-specific marks, etc., but we are interested in only the letters and digits that belong to the plate number.

The first step is to find a threshold that will produce a binary image that has most of the objects separated into distinct regions. You can empirically select a threshold by trial-and-error. Note that, images have red, green, and blue color bands so that you can select one of them

Figure 1: License plate images.

or can convert images to grayscale in order to apply thresholding. (A grayscale image can be obtained as the average of the three RGB bands.)

The second step is to use the morphological operators (`dilation` and `erosion`) you have written, to help isolate the objects of interest while removing everything else. Once you have a decent binary image, use connected components labeling to produce a labeled image that has a distinct label (integer) for each letter and digit. Note that, you MUST have only one function that takes a license plate image and returns the labeled result so that the same procedure (i.e., the same threshold and the same sequence of morphological operations with the same structuring elements) is applied to all license plate images.

You MUST use the `dilation` and `erosion` code from the first question. You CAN use other sources for thresholding, arithmetic and logical operators, connected components labeling, and image I/O. Note that you are NOT allowed to use any operator other than the ones listed above.

Submit: The results of thresholding, morphological operations, and connected components labeling (total of three images) for each license plate, as well as a description of the particular sequence of operations you used to obtain these results (as a text or pdf file). You are also expected to provide a discussion of the results (e.g., which steps were easy and which were more difficult, what was possible and what was not).

Question 3 (40 pts)

The goal of video surveillance is to find new objects in a scene by comparing a new image to a known static background image. The scene is monitored using a camera and the frames of the recorded video sequence are compared to the background image.

Example frames from two video sequences are shown in Figures 2 and 3. Both of these sequences were taken using a static camera (i.e., the camera was not moving) monitoring an intersection on a university campus. Therefore, the differences between a new frame and the background frame are caused by external factors. The objects of interest are the moving cars and people. The frames were taken from the IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS) (`http://www.cvg.reading.ac.uk/slides/pets.html`) in 2000 and 2001.

The first problem is to find the changed areas in a frame. This can be done using the background subtraction technique. Use the first frame in each sequence as the background frame, subtract this background frame from each of the remaining frames, and threshold the difference image to find the changing pixels. Since the original images have RGB bands, you can simply convert them to grayscale and do the subtraction using the grayscale images.

Remember that the changes can occur because of cars or people that appear or disappear but also because of moving leaves, shadows, lighting changes, etc. These noisy changes must be removed from the output. Therefore, the second problem is to use the morphological operators (`dilation` and `erosion`) you have written, to help remove these noisy changes, and to obtain the objects of interest, i.e., the cars and people, as individual connected components. You are expected to obtain the objects of interest as complete as possible (i.e., no under-detection or no over-detection).

You MUST use the `dilation` and `erosion` code from the first question. You CAN use other sources for RGB to grayscale conversion, thresholding, arithmetic and logical operators, connected components labeling, and image I/O. Note that you are NOT allowed to use any operator other than the ones listed above.

Submit: For each sequence, for three frames (the other one being used as the background), the individual results of background subtraction, thresholding, morphological operations, connected components labeling (total of four images for each frame), as well as a description of

(a) Frame 1

(b) Frame 2

(c) Frame 3

(d) Frame 4

Figure 2: Frames from the first surveillance video sequence.

(a) Frame 1

(b) Frame 2

(c) Frame 3

(d) Frame 4

Figure 3: Frames from the second surveillance video sequence.

the particular sequence of operations you used to obtain these results (as a pdf file). You are also expected to provide a discussion of the results (e.g., which steps were easy and which were more difficult, what was possible and what was not).

Notes:

1. This assignment is due by midnight on Tuesday, October 31, 2017. You should upload your solutions as a **single archive file** that contains your **code, resulting image files and descriptions of how you obtained them** using the online submission form on the course web page before the deadline. Please see the course syllabus for a discussion of the late homework policy as well as academic integrity. If you have any questions about what is allowed and what is not allowed in a solution, please check the course syllabus on the course web page.

2. Make sure that the image formats you use to submit your results do not use lossy compression (e.g., do not use jpeg; you can use png, bmp, pgm or ppm).