AI3604— Computer Vision

Fall 2024

Homework #1

Due: 11:59 PM, Sunday, November 3, 2024

You can find relevant knowledge in book《Computer Vision: Algorithms and Applications》！

Only basic Python, Numpy, and OpenCV functions are allowed, unless otherwise specified. This rule applies in particular with the use of OpenCV. Basic image IO functions such as `cv2.imread`, `cv2.imwrite` etc. are allowed. If you are unsure of an allowable function, please ask on OC or ask TA before assuming! You will get no credit for using a “magic” function to answer any questions where you should be answering them. When in doubt, ASK!

Please use python 3 for the programming problems. You will be provided with template Python files containing functions for you to complete. You can implement your own helper functions if necessary. However, do **NOT** change the input and output signature of the original functions in the instructions. For all functions that take or return images, make sure to handle the actual arrays. Do not pass filenames around.

In terms of grading, we would not like to judge students by how good the programming results are. Our purpose is to help the students to learn more from the homework, so do not be afraid of your programming results. **As long as you're serious about it, we will give you a fair grade.**

Written Assignment

1) Consider a pinhole camera with perspective projection.

    a. (10 points) Consider a circular disk that lies in a plane parallel to the image plane. The disk does not necessarily lie on the optical axis. What is the shape of the image of the disk? Show your work and derive your equations.

    b. (10 points) In class, we discussed how to find the vanishing point of lines. Suppose I now consider a family of lines that all lie in a single plane. Where on the image would the vanishing points of ALL lines on the plane Ax+By+Cz+D = 0 lie? The coordinate frame is located at the pinhole with the z-axis pointing towards the image and the effective focal length is ***f***.

First, start with a simple case where A=C=D=0 and B=1 which defines a plane extending horizontally from the optical center. Write down three different line directions in this plane and work out where they vanish. Include your work in your submission.

Second, do the same with the plane B=C=D=0 and A = 1. Again include your work in your submission.

c. (5 points) Now try to define a general relationship (this is challenging!) that relates the parameters of the plane to the parameters of the vanishing points of all lines that fall in that plane.

Programming Assignment

!!! Clarification on Coordinate Definitions:

For problem 1, in order to follow the definition of orientation used in the slides, we assume the origin is at the bottom left of the image, the x axis points to the right, while the y axis points upwards. Therefore, column indices directly correspond to x coordinates, but conversions on the row indices are needed to get the correct y coordinates.

For problem 2 and 3, simply use the column indices as x coordinates and the row indices as y coordinates. This corresponds to the convention where the origin is at the top left of the image, the x axis points to the right, while the y axis points downwards. This convention is more widely used in the vision community today.

!!! Include design choices and output images in your write-up (put in the same .pdf file with your written assignment):

(1) For each required function, list the main ideas behind your code and the choices you made. Examples include (but are not limited to):

- What is the threshold value you used to get a good binary image?
- What are the kernel values in your Sobel edge detector?

(2) For each required function, if the output is an image, include it in your write-up; if it is a number or a short array, report it in your write-up; if it is a long array (> 10 elements), report the first 10 elements in your write-up.

(3) Please keep your report clear but brief. It should not be more than three pages. Do not waste your time and energy on meaningless things.

1) Our goal is to develop a vision system that characterizes two-dimensional objects in images. Given an image such as **two_objects.png**, we would like our vision system to determine how many objects are in the image and to compute their positions and orientations. This information is valuable as it can be used by a robot to sort and assemble manufactured parts.

   The task is divided into three parts, a)-c), each corresponding to a Python function. You need to complete the code snippets marked with "`#TODO`" in **p1_object_attributes.py**.

   In this problem, you are provided with a driver program in the `main` function. The code can be run with the following command, which specifies the image to process and the thresh_val described in a):

```
python3 p1_object_attributes.py two_objects 128
```

   a. (10 points) Write a Python function that converts a gray-level image to a binary one using a threshold value. The binary image should be 255 if intensity >= thresh_val else 0.

```
def binarize(gray_image, thresh_val):
  # TODO
  return binary_image
```

   b. (15 points) Implement the sequential labeling algorithm that segments a binary image into connected regions:

```
def label(binary_image):
  # TODO
  return labeled_image
```

   Note that you may have to make two passes of the image to resolve possible equivalences in the labels. In the "labeled" output image each object region should be painted with a different gray-level: the *gray-level assigned to an object is its label*. The labeled images can be displayed to check the results produced by your program. Note that your program must be able to produce correct results given any binary image. You can test it on the images given to you.

Note: You can either implement the two-pass version of the algorithm in the notes, or a (somewhat simpler but slower) recursive algorithm as described here: https://courses.cs.washington.edu/courses/cse373/00au/chcon.pdf

c. (15 points) Write a Python function that takes a labeled image and computes a list of object attributes (which in a real application could be used to locate and identify an object).

```
def get_attribute(labeled_image):
  # TODO
  return attribute_list
```

Each element of the attribute list should be a dictionary with the following keys: position, orientation, and roundedness. The position should be a dictionary with keys: 'x' and 'y'. The origin is defined as the bottom left pixel of the image. Please use radian for orientation, and all numbers are floats.

The program will write the images to files and print their properties. Test your function on the images **many_objects_1.png** and **many_objects_2.png.**

2) Your task here is to implement Hough transform to find circles in an image. We provide **data/coins.png** as an example image. You need to complete the code snippets marked with "`#TODO`" in **p2_hough_circles.py**. In this problem, you need to implement your own driver program to load images and save results.

a. (10 points) First you need to find the locations of edge points in the image. Complete function `detect_edges`. The input to the function is a 2 dimensional uint8 array representing a grayscale image. The output should be a 2 dimensional float array where the intensity at each point is proportional to the edge magnitude. The output array should have the same size as the input array (pad input image with zero when necessary).

Use Sobel masks to implement the edge detector. You may **NOT** use existing edge detectors in OpenCV. However, you can compare your output to the output of those detectors for verification. (In fact, this is a good idea to test your code!)

```
def detect_edges(image):
  # TODO
```

```
  return edge_image
```

b. (15 points) Next, you need to implement the Hough Transform for circle detection. Complete the function `hough_circles`. The first input argument of the function is the edge magnitude map you got from your edge detector. The second argument is a threshold on the edge magnitude. Thresholding your edge magnitude map with this value gives a boolean map indicating whether each pixel is an edge point or not. Experiment with different threshold values so only strong edges are kept.

We represent a circle with its center coordinates (*x*, *y*) and its radius *r*. Therefore, your accumulator array should be 3 dimensional. We assume that the *x* and *y* ranges are the same as the input image and the resolutions are both 1 pixel. The radius range and resolution are specified by the third input argument of the function, which is an array of possible radius values, e.g. [20, 21, … , 39, 40].

The function should return the thresholded edge image and the accumulator array. Experiment with different edge threshold. In your driver program, save the output edge image to **output/coins_edges.png** when you are satisfied with the result.

```
def hough_circles(edge_image, edge_thresh, radius_values):
  # TODO
  return thresh_edge_image, accum_array
```

c. (10 points) To find circles in the image, scan through the accumulator array looking for parameter values that have high votes. Complete function `find_circles`. Here again, a threshold must be used to distinguish strong circles.

The function should return a list of 3-tuples where each element (*r*, *y, x*) represents the radius and the center coordinates of a circle found by your program.

Besides, draw these circles on a copy of the original image using OpenCV function `cv2.circle` (with `color=(0, 255, 0)` and `thickness=2`). Return the resulting image as well. Experiment with different vote threshold. In your driver program, save the output image to **output/coins_circles.png** when you are satisfied with the result.

```
def find_circles(image, accum_array, radius_values,
                 hough_thresh):
  # TODO
```

```
return circles, circle_image
```