COL 783: Assignment 4
Robust Forms Processing

Due date: Friday, 1 November 2019

OVERVIEW

In this assignment, you will work on processing photos and scans of printed form in order to extract the form fields and the written information entered in them. These images have been shared with us by Prof. Aaditeshwar Seth and Prof. Chetan Arora, and are images of immunization records which need to be digitized. Since many of the images have been taken by lay participants in uncontrolled environments, they have varying orientation, illumination, contrast, etc. which makes them harder to process automatically. What we want to do is to extract the form fields from each image, and detect the written information in each field.

GOALS

This assignment is a bit different from the previous ones: it is much more application-driven and open-ended. You are permitted to use built-in routines for all of the topics discussed in class, i.e. edge detection, the Hough transform, thresholding and other segmentation techniques, morphological operations, etc. The challenge is to find a good sequence of operations to apply for each task, and choose the best parameters for each, to get as accurate results as possible.

Ideally, we would like a method that is fully automatic, that is to say, the same code should work well on all the input images without requiring per-image parameter tuning. If this is not possible (as may well be the case), show results on all the images using the best fixed set of parameters you can find, and for any failure cases, show any better results you can get by manual tuning. To simplify the parameter choice, you may want to resize the images to a consistent size in your code before further processing.

Download the dataset of images from here. The images are of three different types: (i) scanned, (ii) photos of printed-out forms, and (iii) photos of forms in booklets. Since these three sets have rather different characteristics, you may use different parameters for each set, but try to use the same parameters for every image within a given set.

TASKS

Alignment

Automatically detect the orientation of the form, and rotate the image to straighten it. You can use a built-in function to perform the rotation, but be careful to choose an interpolation scheme that does not introduce aliasing.

(For photographed forms, there can also exist some perspective distortion which cannot be removed simply by rotation. We will ignore this issue in this assignment; in the provided dataset I have tried to include only images with minimal perspective distortion.)

Form field segmentation

Design an algorithm to segment out the form fields, i.e. the white boxes in which information is supposed to be entered. Note that these are generally rectangle-shaped light regions surrounded by darker regions or edges. Find a segmentation operation or sequence of operations that isolates all (or most) of the fields as separate connected regions. You may need additional steps to remove extra regions that do not have the right shape.

Character detection

Some of the fields have characters written in them, while others do not. Find a way to isolate these characters and label them as distinct connected components. (Of course, if two characters have been written in a connected way, you are not expected to separate them.)

Extra credit

The form fields also include pre-printed slashes for the dates to be entered, and these slashes can show up in the character segmentation as well. Can you design a method to automatically remove them, even if they are overlapping with the written characters? You could use the fact that they always appear in the same location in each form field.

SUBMISSION AND EVALUATION

As usual, submit your code and a report describing your work. Since this is a more open-ended assignment, it is especially important to describe your algorithm in detail and discuss the reasons for your relevant design and implementation choices. Feel free to include other approaches you tried and why they did not work as well as your final solution. Include images demonstrating the results of each task of the assignment.

Students working in groups of two are expected to show results on all of the images in the provided dataset. Students working alone may show results on only two images of each type, but must include at least one image with significant rotation.

You will be graded on the accuracy of the results, the amount of manual fine-tuning required, and the comprehensiveness of the report.