BBM 418 - Computer Vision Laboratory

HACETTEPE UNIVERSITY
Department of Computer Engineering

Programming Assignment-1
Spring 2024-2025
Dr. Nazli Ikizler Cinbis

Due Date: 23:59 on Thursday, March 20th, 2025

Perspective Correction using Edge and Line Fitting

Document dewarping and recognition involve digitally correcting skewed or perspective-distorted scanned documents. When documents are captured with mobile cameras or scanners, they may not be perfectly aligned, leading to geometric distortions such as warping, folding, or perspective issues. You can check some examples on Figure 1.

In your assignment, you will try to solve distortion problems with Hough Transform and RANSAC. For the given image of a document (e.g., a book page or a receipt), you will detect its edges and fit quadrilateral contours using Hough Transform and RANSAC.

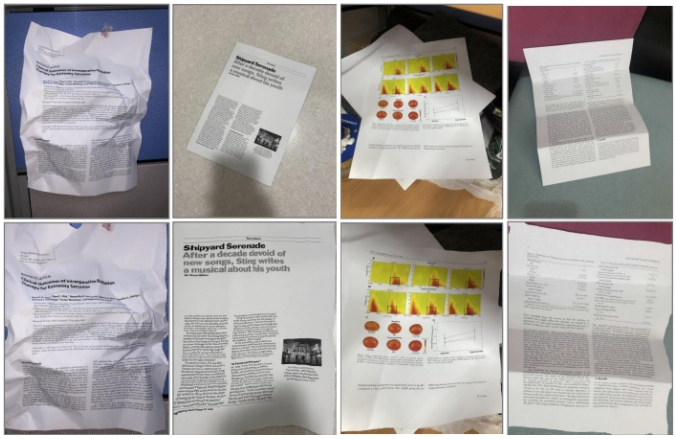

Figure 1: Some examples for distorted and corrected images with perspective correction using edge and line fitting.

Hough Transform and RANSAC

Hough Transform: The Hough Transform is a voting-based method used to detect potential line objects in an image (Figure 2). It achieves this by:

- Identifying potential line objects from given points.
- Counting the potential line objects detected.
- Assigning points to the corresponding line objects.

The transform accumulates votes for each potential line passing through the edge points and identifies lines with the most votes.

Figure 2: An example for Hough Transform on a sample image from the dataset.

RANSAC (Random Sample Consensus): RANSAC is an iterative algorithm used to robustly fit a model to a set of data points containing outliers. When combined with edge detection and Hough Transform, RANSAC helps refine line detection by repeatedly selecting random subsets of points, fitting a line model, and evaluating how well it matches the remaining data. This approach is particularly effective in scenarios with noise or outliers, ensuring that only the inlier points contribute to the final line model.

Dataset

Dataset [1] consists of images of some documents(See Figure 3). For your assignment, you will use a smaller version (Download link). The dataset has the following folders;

- digital (ground-truth images)
  - curved, fold, incomplete, perspective, random, rotate
- distorted
  - curved, fold, incomplete, perspective, random, rotate

Figure 3: Sample images for ground truth and distorted document images.

The Implementation Details

1. Firstly, you are expected to implement Hough Transform for finding line objects for images.
2. Then, you need to use RANSAC to improve matches for line objects obtained from Hough Transform.
3. It is important that you will find quadrilaterals on your assignment.
4. Also, you need to use geometric transformations to improve the document's frontal look.
5. You will also write your own code for Hough Transform, RANSAC, and geometric transformation methods. You can use NumPy or other libraries for basic tasks.
6. Dataset includes 6 different classes of distorted images (curved, fold, incomplete, perspective, random, rotate). You will use these classes to compare your method's quality.
7. [**Important**] This assignment is not a classification task, you will use classes for comparison of the method you use. You can see that all images include the paper and some images paper is folded, crumpled, etc.
8. For comparing ground truth and obtained image use Structural Similarity Index (SSIM) measurement. For SSIM you can use the scikit-learn library.

The Report

- You will report results, comparisons, plots, etc. on the PDF file. (Latex format will be shared on Piazza.)
- You need to explain your findings in your own words.
- Explain the methods, why you use the Hough Transform and RANSAC methods, and how to improve the result of what you do.
- Explain which geometric transformations you use on your assignment.

- Compare your results with ground-truth images. Do your results look good (plot some images)? You can add some good results and bad results for each class.
- Compare SSIM results, and what can be done to improve the result.

What to Hand In

Your submission format will be:

- b<studentNumber>.ipynb (iPython notebook)
- b<studentNumber>.py (Convert your ipynb file to py file)
- report.pdf

Archieve this folder as **b<studentNumber>.zip** and submit to `https://submit.cs.hacettepe.edu.tr`.

Grading

The assignment will be graded out of 100:

- Code (50 points): Hough Transform (15), RANSAC (15), geometric transformations (10), comments (5), well-written code (5)
- Report (50 points)

Academic Integrity

All work on assignments must be done individually unless stated otherwise. You are encouraged to discuss with your classmates about the given assignments, but these discussions should be carried out abstractly. That is, discussions related to a particular solution to a specific problem (either in actual code or in the pseudocode) will not be tolerated. In short, turning in someone else's work, in whole or in part, as your own will be considered a violation of academic integrity. Please note that the former condition also holds for the material found on the web as everything on the web has been written by someone else.

References

[1] `https://sg-vilab.github.io/event/warpdoc/`