Handwriting Recognition Using PCA

Authors: Ismail Makhlouf

                Ahmed Ghobashy

Objective

The goal of this implementation of Optical Character Recognition is the detection of handwritten letters on paper and/or whiteboard. The simplest application for this is the quick extraction of Professor notes from the whiteboard or student notes from a notebook using pattern analysis techniques, rather than the traditional, tedious method of copying by hand or photocopy machine. To that end, a number of various attempts have been made. Other applications include recognition of addresses on postal mail and/or recognizing handwritten signatures.

Data Collection

Several datasets have been manually constructed by the authors across three mediums; lined paper, white paper, and whiteboard using a whiteboard marker. Images were captured primarily using an iPhone 5 8 megapixel camera. Furthermore, digitally constructed characters were also experimented with to test the robustness of the classifier. The data sets consisted of Latin (English) uppercase characters as well as Arabic numerals in separate data sets. Each data set (or even subset) was subject to separate lighting, orientation, and zoom conditions, thus providing various capture scenarios. The Latin character set(s) consisted of around 60 samples, and the numerical character set(s) consisted of around 20 samples.

Feature Extraction and Segmentation

As the input to the OCR system is received in the form of raw images, items of interest (in this case, the characters themselves) must be detected, segmented, preprocessed and extracted to ensure accuracy and proper testing. It should be noted, however, that the authors had control over the collection of the data and thus the quality of the input (in part). As such, even with changes in lighting, angle, etc., the authors attempted to collect characters in as clean a form as possible, with the curves making up the characters as thick as possible (hence the marker). After the images were collected, the characters were segmented by manual cropping. This process was undergone for both the training and testing sets.
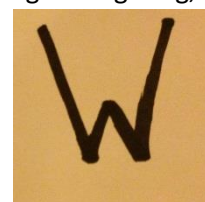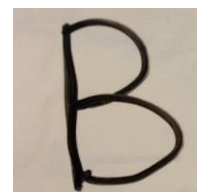
Preprocessing

A number of preprocessing steps were required to improve the quality of the samples. Firstly, the samples were converted to grayscale and reduced to a standard size of 64x64. A nonlinear (median) filter was then applied to remove impulse noise from the sample, which would be common on a whiteboard, for example, which would have residual markings on its surface.

Methodology

As mentioned in the title, the main recognition technique used in this implementation is Principal Component Analysis (mentioned later). However, due to previous experience with PCA, it loses accuracy when the input data does not come from a controlled environment.

Turk and Pentland, for example, captured images of faces in a controlled manner, at specific angles and facial orientations, with minimal background interference. The nature of this implementation, and its resulting application, does not, unfortunately, lend itself to this type of control. As such, a number of techniques must be employed to compliment the use of PCA.

Radon Transform

 "The Radon transform is the projection of the image intensity along a radial line oriented at a specific angle. The radial coordinates are the values along the x' - axis, which is oriented at ϑ degrees counter clockwise from the x -axis. The origin of both axes is the center pixel of the image[1]." The Radon transform is a mathematical operation very similar to the Linear Hough transform, in that it projects lines onto another space, forming an accumulator with peaks that correspond to certain line parameters.

The Radon transform is a very useful tool for line detection, and is used widely in the fields of radar imaging, geophysical imaging, nondestructive testing, and medical imaging. This transform is useful for OCR because it is very good at extracting lines and/or curves from very noisy images.
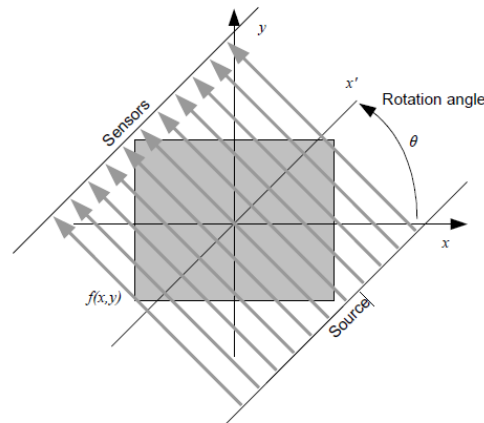
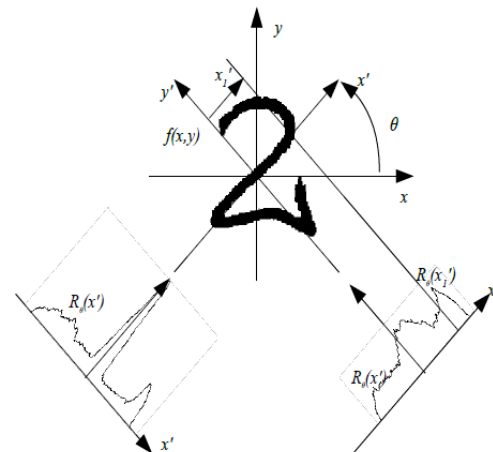Fig 3. Single projection at a specified rotation angle

Fig 5. Geometry of the Radon Transform

Principal Component Analysis

PCA is used to transform data onto a new space such that the greatest variance by any projection of the data lies on the first coordinate of the new space (First Principal Component), the second greatest variance on the second coordinate, etc. This technique can be used to reduce dimensionality by retaining only the characteristics of the data that contribute the most to the variance.

Typically, when performing supervised classification such as this, PCA is performed on the dataset of images of the training/test items. However, since the Radon transform filters out noise and detects lines well, it may yield a more unique, and less corrupted, representation of each data item. Thus, we perform the transform on each training/testing item and

---

[1] Miciak, Mirosław . *Character Recognition Using Radon Transformation and Principal Component Analysis in Postal Applications*. Proceedings of the International Multiconference on Computer Science and Information Technology, pp. 495 – 500

perform PCA on the transform to reduce its dimensionality and project the data points corresponding to each transform into a lower dimensional space. The distance from each testing sample to every class (in this case, every training image) is then measured, and the test sample closest to a certain class is classified as such (given that we've established that the test item is, indeed, a character.

Challenges

Unfortunately, the techniques mentioned above were not applied sufficiently to determine how effective they actually are in classifying characters. This is due to a number of challenges faced:

- Lighting, orientation, scale and zoom
- Differences in line thickness within and between the letters; thin lines were especially problematic
- Segmentation was manual and thus unstandardized
- Training taking too long
- Sample size too small
- Choice of number of eigenvalues
- Difficulty of visualizing operations on images when in Radon transform
- Similarity between Radon transforms (naked eye)
- Similarity between different characters
- Data set availability

Radon Transforms – Very Similar
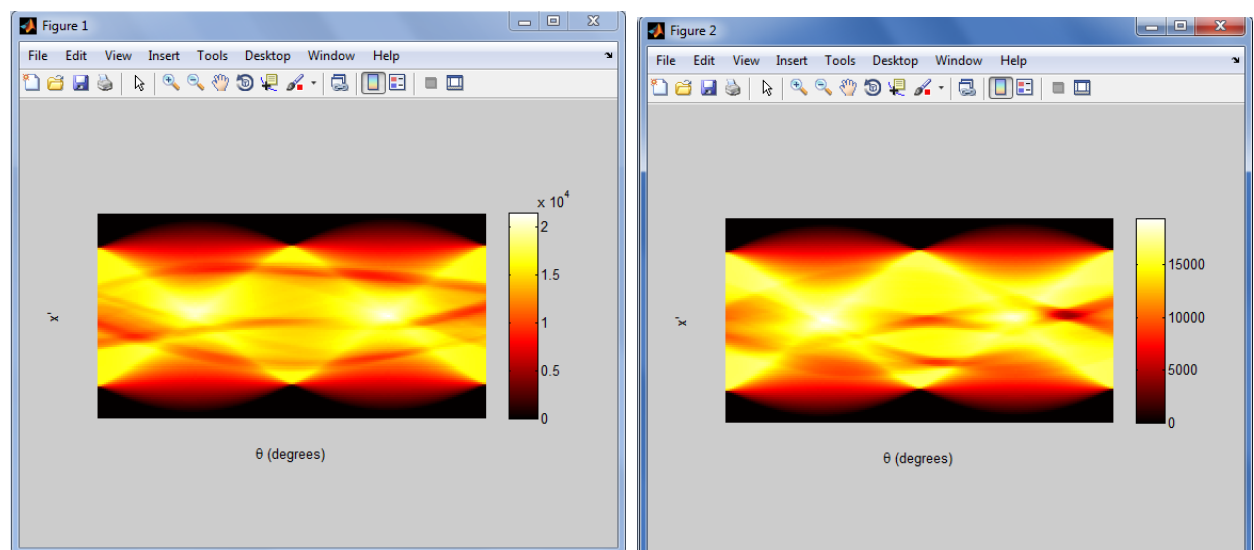
Results

As the OCR system has not yet been completed, performance statistics have not quite been developed yet. However, rough rate of recognition for numerical characters has been found at around 22.2%. Characters are more complicated, as there are 26 classes, many of which have nested shapes of other letters, so their accuracy is yet to be determined.

Tools for Improvement

The key to this recognition system is increased preprocessing. To normalize the input letters, for example, we can use the characters center of gravity to find the angle of rotation and scale coefficient and acting accordingly. In addition, we can vary the subdimension used for PCA to heuristically determine its ideal value. Furthermore, we can combine PCA with more robust classifiers to yield better classification rates. We can also use the General Hough Transform instead of radon transform, and thus tailor the accumulator to detect the exact pattern of desired shapes. In addition, we can use edge detection (was attempted, but did not succeed) more aggressively to further extract desired shapes with clarity, and apply further filtration to the edge map itself.