Motivation

Text documents, such as reports and journals, in different languages usually require people fluent in those languages to read and translate them manually. This could take a long time to do. Why not have it automated by a program?

Solution: Build an image processing system to process text document images taken with a phone/camera and translate them text.

The structure of different written languages are different and Thai is one language that involves more subtle complex written patterns than languages like English. It becomes a challenging problem to tackle.

Processing Pipeline and Methods

Preprocess

- Load camera image
- Load font templates
- Locally Adaptive Thresholding
- Noise Removal (filtering)

Processing

- Image segmentation/region labeling (identify characters)
- Hough Transform (fix orientation)
- SIFT Features
- SVD + PCA
- Eigenvector Comparison

Translate

- XOR + Resize (template matching)
- Weighted distance scoring
- Homography + RANSAC

Results

- Correctly classify:
- Misclassify:

Experimental Results

<table>
<thead>
<tr>
<th>Detection Method</th>
<th>XOR + resize + weighted distance score</th>
<th>SVD + PCA</th>
<th>SIFT + RANSAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clean Test Document (same font)</td>
<td>91%</td>
<td>73%</td>
<td>64%</td>
</tr>
<tr>
<td>Clean Test Document (different font)</td>
<td>80%</td>
<td>36%</td>
<td>36%</td>
</tr>
<tr>
<td>Test Camera Document (same font)</td>
<td>95%</td>
<td>69%</td>
<td>59%</td>
</tr>
<tr>
<td>Test Camera Document (different font)</td>
<td>83%</td>
<td>36%</td>
<td>18%</td>
</tr>
</tbody>
</table>

NOTE: Table shows the percentage of correct character detection

- The testing documents we used contains actual sentences from Thai newspaper and articles. This is so that it will represent the frequency of each characters being used
- For XOR method, same characters are either consistently correctly classified or consistently misclassify. This is not the case for SVD and SIFT

Related Work/ Reference