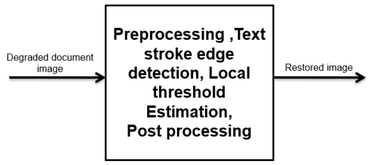
Degraded Document Image Restoration

Harshini Keerthi Vasan, Keerthi S, SRM Easwari Engineering College

Project Goal

- ❖ Restore the degraded documents using novel document image processing techniques.
- ❖ Reduce the binarization error by minimum parameter tuning using contrast image construction.
- ❖ Minimize the computational time for text stroke edge pixel detection using sobel edge detection algorithm



Motivation

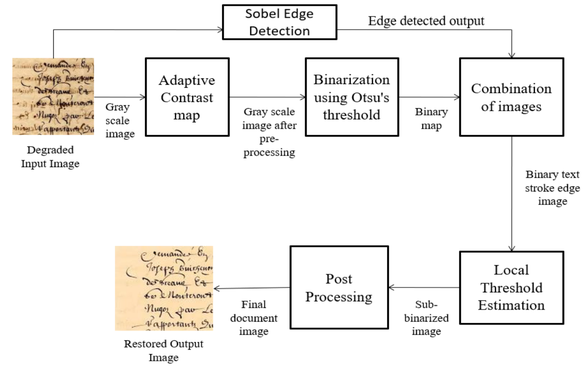
- ❖ The historical documents represent the main support of knowledge transfer.
- ❖ Documental degradation results in diminished robustness with high computational complexity.
- ❖ The image binarization system will improve the degree of success in character segmentation and recognition such as in Optical Character Recognition(OCR).

Issues in Existing System

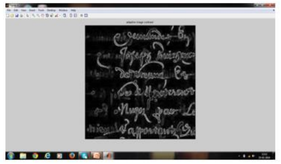
- ❖ The existing system cannot work on degraded document images with complex background.
- ❖ Computational complexity is expensive: $O(N^3)$ for an N*N image.
- ❖ Foreground text is not retained in an area which has low contrast.

Key Idea

- ❖ A binarization system is developed to produce a refined document image with superior quality using Sobel Edge Detection algorithm.
- ❖ Adaptive contrast map is used as an advanced technique to alter the pixel range according to the contrast and gradient of the input image.

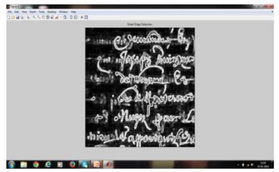


- ❖ Uses Sobel Edge Detection for Text Stroke Edge Pixel Detection.
- ❖ The binarized image obtained by using Otsu's global thresholding method is combined with the edge detected image obtained by using sobel edge detector.

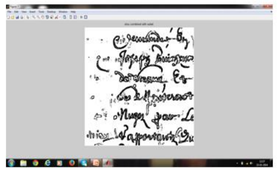


1. Contrast Enhanced Image

2. OTSU Threshold Image



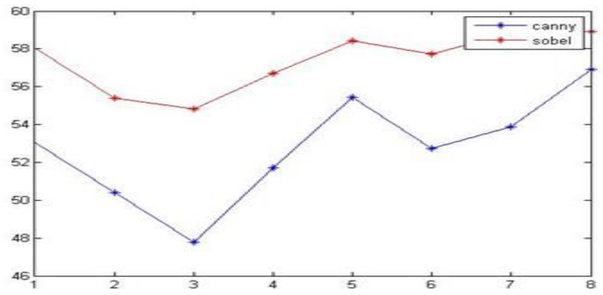
3. Sobel Edge Detected Image

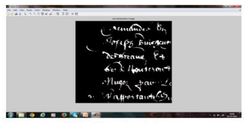


4. Combined Image

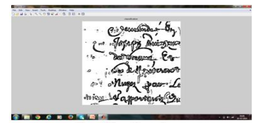
Results

- ❖ Performance Analysis for this system is carried out using PSNR (Peak Signal to Noise Ratio). The PSNR value of OTSU combined with sobel edge detection is higher than canny edge detection technique.
- ❖ Sobel Edge Detection takes lesser computational time than canny edge detection

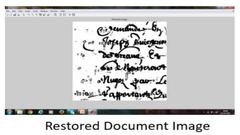




Sub-binarized Image



Classification Image



Restored Document Image

Future Work

- ❖ Various other global thresholding techniques such as adaptive thresholding can be used with sobel edge detection to obtain better results.
- ❖ The obtained output can be tested for better results using other techniques such as F-Measure , MSE .