Robust Text Reading in Natural Scene Images

Phase II

Stroke Width Transform (SWT)

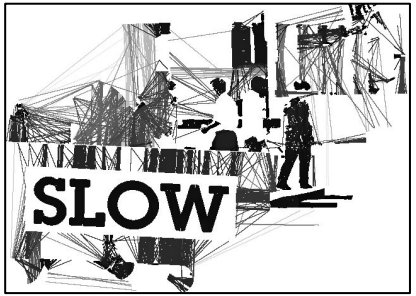
- SWT Algorithm :
  - Define stroke as a contiguous part that forms a band of nearly constant width
  - Compute edges in the image using Canny edge detector
  - Follow along the direction of gradient till the Stroke boundary is reached
  - Store the computed SWT values and assign the median SWT value to the stroke
- Find the letter candidates by comparing the SWT value ratios
- Group the letters into sentences by clustering them into chains

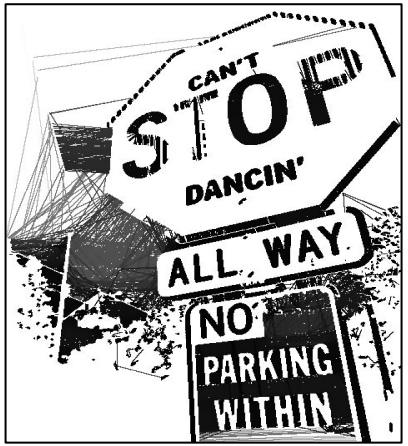
Detection of Extremal Regions (MSER)

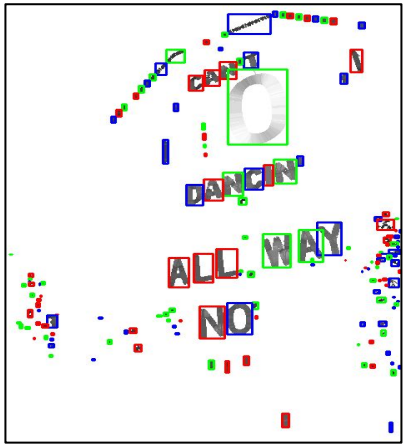
- Detect candidate text regions using MSER
- Remove non-text regions based on basic Geometric properties
- Remove non-text regions based on Stroke Width variation
- Merge text regions for final detection result

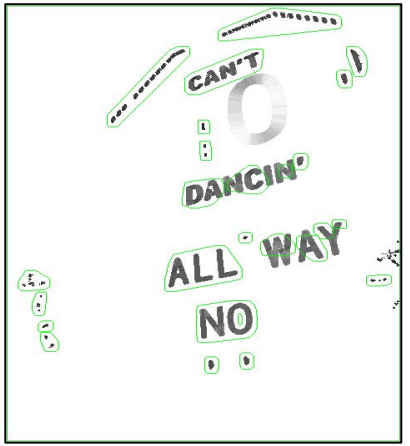
Qualitative Results - SWT

Original Image | Stroke Width Transform | Letter Candidates | Sentence Chaining

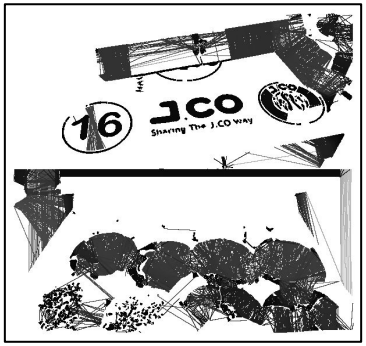


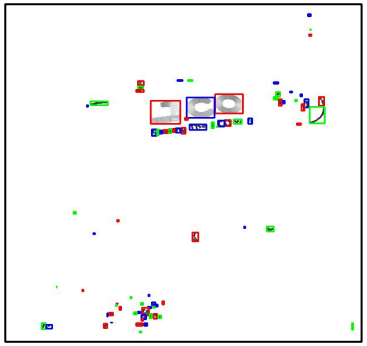


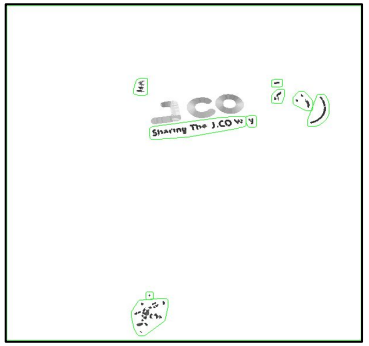


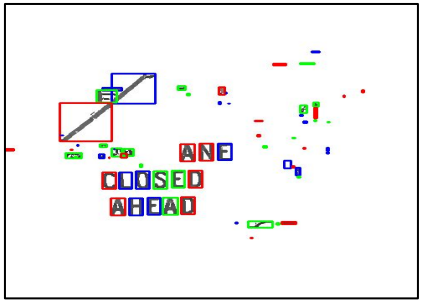


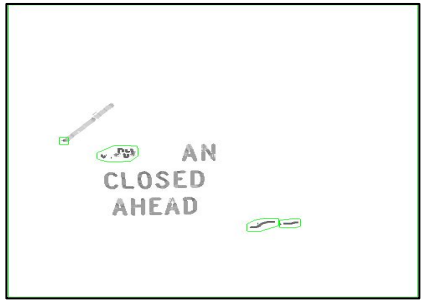
Original Image | Stroke Width Transform | Letter Candidates | Sentence Chaining







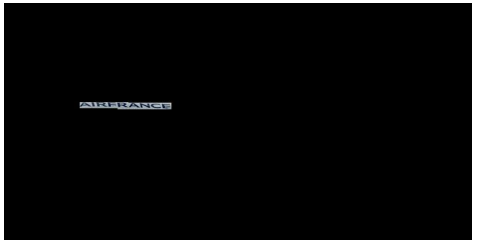


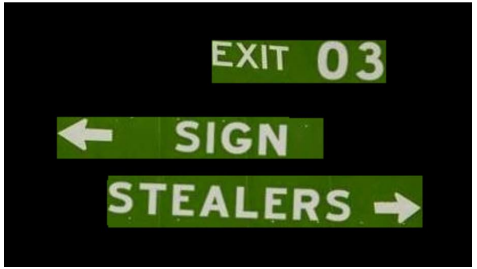


Qualitative Results - MSER

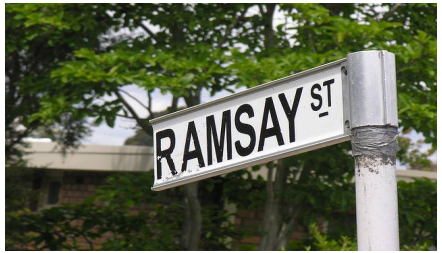
Original Image

Text Detection

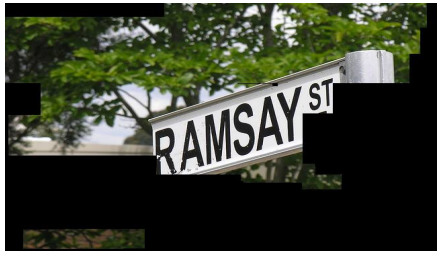




Original Image



Text Detection



Method Comparison

Advantages

- Stroke Width Transform
  - Invariant to scale, occlusions and affine transformations
  - Robust to noise after some preprocessing
  - Proper demarcation between textual and non-textual regions

- Detection of Extremal Regions
  - Invariant to scale, occlusions, illumination and affine/homographic transformations
  - Robust to noise after some preprocessing
  - Erroneous textual regions can be removed using proper Canny thresholding
  - Uses very less memory as compared to standard MSER algorithm
  - High detection rate due to usage of multiple image channels

Disadvantages

- Stroke Width Transform
  - Not invariant to homographic transformations
  - Additional parameter denoting darkness/lightness of text wrt background is needed
  - Improper results on text having non-uniform width

- Detection of Extremal Regions
  - Absence of an effective text candidate construction algorithm

Relevant Links

- Github Repository
  - Contains updated code and documentation
  - goo.gl/shjHj3
- Google Drive
  - Contains text detection results
  - goo.gl/kPuumA