Features Extraction For Text Detection and Localization

- By H.A. Driss, S. ELFKIHI, A. JILBAB

- Presented By

Riti Gupta

CS 297

INTRODUCTION

- Discuss text detection in images
  - Localize text without recognizing individual characters
  - Challenges : large variations (font, complex background, size)

- Texture features and connected component analysis used to detect text in images

- Basic algorithm used:
  - Detect all the edges (edge-detection algorithms)
  - Preserve important structural features using sharp contrasts
  - Texture information used to differentiate text from non-text regions
    - Gradient magnitude as texture characteristic used
  - Verify regions to reduce false alarms and generate valid text regions

GENERATION OF CANDIDATE TEXT REGIONS

- Edge detection
  - Use canny edge detection algorithm (Fig. 1)

Figure1. Some inputs (color images) and outputs of the edge detection algorithm.

GENERATION OF CANDIDATE TEXT REGIONS (COND.)

- Closed contour selection (Fig. 2)
  - Separate text regions from non-text regions
  - Analyze each connected component in image
    - Might include non-text objects

Figure2. result of a Closed contour selection.

TEXT LOCALIZATION

- Identify non-text regions in closed contours
  - Use gradient magnitude as a texture characteristic
  - Create a new image containing gradient magnitude of regions
  - Gradient magnitude computed as

$$Ig(i,j) = \sqrt{\left(\frac{dI(i,j)}{dx}\right)^2 + \left(\frac{dI(i,j)}{dy}\right)^2} \quad . \qquad (1)$$

where I(i,j) is a grey level of a corresponding pixel in a closed contour image I_c .

  - For each region Rt, the edge intensity is computed as

$$EI_{Rt} = \sum_{(i,j)\in Rt} Ig(i,j) \quad . \qquad (2)$$

Where Ig (i,j) is the magnitude of the intensity gradient at location(i,j)

  - Region with high edge density considered as text region

TEXT LOCALIZATION

- Classify based on a threshold T

$$T = \alpha \sum_{EI_{Rt} \in Ig} \frac{EI_{Rt}}{N} \qquad (3)$$

where α is weighing factor (0.8), N = number of regions

Figure3. Some inputs and outputs of the proposed algorithm.

HOW TO AVOID FALSE ALARMS IN TEXT REGION

- Classify based on a threshold T

$$T = \alpha \sum_{EI_{Rt} \in Ig} \frac{EI_{Rt}}{N} \qquad (3)$$

where α is weighing factor (0.8), N = number of regions

- Consider the component as character if following rule is met

$$0.1 \le \frac{W_i}{H_i} \le 2 \quad ; \quad 50 \le W_i \times H_i . \qquad (4)$$

where W_i and Hi are width and height of text region i.

OTHER PARAMETERS TO AVOID FALSE ALARMS

- Consider geometric spatial relationship of components.

- Vicinity rules for two neighboring regions(xi; yi;Wi;Hi) and (xj; yj;Wj;Hj)
  - Ensure that components are neighbors in horizontal direction
    - $dist(xi, xj) \leq 0.2max(Hi,Hj)$
  - Ensure that components belong to the same text line
    - $dist(yi, yj) \leq 2max(Wi,Wj)$
  - Ensure that components have similar height
    - $0.5 \leq Hi/Hj \leq 2$

RESULTS

- Evaluation done on ICDAR 2003 529 images text locating database.
- Text of various size and font sizes could be located.
- As per Recall – Precision curve, Precision rate of 0.95 observed for 0.89 Recall rate.
- Standard measure of quality f = 0.92.