Text-background Decomposition Using Adaptive Boundary Clustering

Presented by Ungsumalee Suttapakti
Faculty of Information Technology,
King Mongkut’s Institute of Technology Ladkrabang,
Thailand

K. Woraratpanya, U. Suttapakti, K. Pasupa, P. Boonchukusol, T. Titijaroonroj, R. Hokking, Y. Kuroki and Y. Kato,
Outline

• Introduction
• Proposed methods
• Experimental results
• Conclusion
What does it mean?
What does it mean?

Input image:

Output text:

Translator:
What does it mean?

Input image:

Output text:
อาคารที่จอดรถยนต์

Translator:
Car Parking
Introduction

Text localization

Character recognition

Editable Text

Output text:
อาคารที่จอดรถยนต์

Translator:
car parking
I) Structure of language

Upper level

Middle level

Lower level

Thai text

English text
II) A variety of text fonts, sizes, styles, colors, and orientations

III) Complex background

IV) Lighting condition
Multilayer Separation [1]

- Based on intensity value of the reduced-color image

Color-based partition [2]
- This algorithm detects text based on connected pixel with similar colors.

Boundary Clustering (BC) [3]

- The color feature of BC method is averaged from all edge pixels of each feature boundary.
- A number of clusters is fixed, 5 clusters.

Long Computing Time

Fast Boundary Clustering (FBC) [4]

- FBC method extracts color features based on a pixel boundary.
- A number of clusters is fixed, 5 clusters.

Faster than

Problems: 1) The fixed number of clusters [3,4] is unsuitable, 5 clusters.

Text is separated into different layers

Cannot detect
Problems: 2) The extracted color features based on a pixel boundary [4] cannot exactly represent the object when foreground and background are non-uniform colors.

A failure of text-background decomposition
Objectives: To improve the precision of text localization and the accuracy rate of character recognition
Proposed Methods

Scene images → Proposed Methods

Text-background decomposition

Adaptive Boundary Clustering (ABC) method → n-Point Boundary Clustering (n-PBC) method

Text localization

Modified connected component analysis (MCCA) [4] → PCA combined with predefined pattern [5]

Character recognition

Recognized character


• Introduction
• Proposed methods
• Experimental results
• Conclusion
• Adaptive Boundary Clustering (ABC) method
• n-Point Boundary Clustering (n-PBC) method
To find the suitable number of clusters of each scene image

Scene image → Text-background decomposition → A number of clusters = ? → Text localization → Character recognition

Editable text
Adaptive boundary clustering based on color quantization

- Color quantization is a process that reduces the number of distinct colors used in a scene image.

\[ k = n + 1 \]
\[ = 2 + 1 \]
\[ = 3 \]

- \( k \) is the number of clusters.
- \( n \) is the number of bins whose frequency is greater than the mean of the histogram.

- The diagram illustrates the quantization process for the colors Yellow, Magenta, Cyan, Red, Green, Blue, White, and Black, with their respective RGB values and frequency distribution.
Adaptive Boundary Clustering (ABC) method

- Three clusters are suitable.

All text in the same layer.  

Background Boundaries
The extracted color features based on an average of n-point boundary
(n = 1, 2, 3, ..., 8)
The feature of each object edge is extracted by determining the maximum color difference between two opposite pixels on a $3 \times 3$ mask.

Example; $n = 3$

\[ V_i = [R_h, G_h, B_h, R_l, G_l, B_l, C_y]_i \]

$[R_h, G_h, B_h]$ are pixel color values of the highest intensity. $[R_l, G_l, B_l]$ are pixel color values of the lowest intensity. $[C_y]$ is y-coordinate of the centroid of boundary.
• Introduction
• Proposed methods
• **Experimental results**
• Conclusion
Datasets

- A uniform color
- Lighting condition
- Complex background
Proposed Methods

Modified connected component analysis (MCCA) [1]

Scene images

Text-background decomposition

Text localization

Modified connected component analysis (MCCA) [1]

PCA combined with predefined pattern [2]

Evaluation: Self-entropy

Precision and recall

Recognition rate

Targets: lower, better

higher, better

higher, better

Evaluation: Computing time

Targets: lower, better

lower, better

lower, better

Recognized character
Experimental results

**Self-entropy**

![Bar chart showing self-entropy for different comparative methods with annotations]

**Recall**

![Bar chart showing recall for different comparative methods with annotations]

**Precision**

![Bar chart showing precision for different comparative methods with annotations]

**Recognition rate**

![Bar chart showing recognition rate for different comparative methods with annotations]
Experimental results

Computing time

Comparitive Methods

- Text-background Decomposition
- Text localization
- Character Recognition
The ABC method achieves the very low self-entropy, that is the good decomposition of text and background.

The ABC method helps increase the precision of text localization and improves the accuracy rate of character recognition, when compared to the conventional methods.
Thank you for your attention
Evaluation

• **Self-entropy**

\[ E_s = -\sum_{i=1}^{k} \left( P(x_i) \log P(x_i) + P(y_i) \log P(y_i) \right) \]

where \( P(x_i) \) is the ratio of the number of text boundaries to the total number of boundaries.
\( P(y_i) \) is the ratio of the number of non-text boundaries to the total number of boundaries and \( k \) denotes the number of clusters.

The low self-entropy implies the purity of text-background decomposition.
Evaluation

\[
\text{precision} = \frac{tp}{tp + fp} = \frac{5}{5 + 2} = 0.71
\]

\[
\text{recall} = \frac{tp}{tp + fn} = \frac{5}{5 + 1} = 0.83
\]

<table>
<thead>
<tr>
<th>Actual class</th>
<th>Predicted class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text</td>
<td>( tp = 5 )</td>
</tr>
<tr>
<td>Non-text</td>
<td>( fp = 1 )</td>
</tr>
</tbody>
</table>
Text Localization and Character Recognition

- A diagram of text localization and character recognition.
Text-background Decomposition

Modified Body of Text Detection
A failure of text-background decomposition of FBC method

<table>
<thead>
<tr>
<th>A failure of text-background decomposition of FBC method</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Image of text-background decomposition failure of FBC method" /></td>
</tr>
</tbody>
</table>

A success of text-background decomposition of n-PBC method

<table>
<thead>
<tr>
<th>A success of text-background decomposition of n-PBC method</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image2" alt="Image of text-background decomposition success of n-PBC method" /></td>
</tr>
</tbody>
</table>