Text-background Decomposition Using Adaptive Boundary Clustering

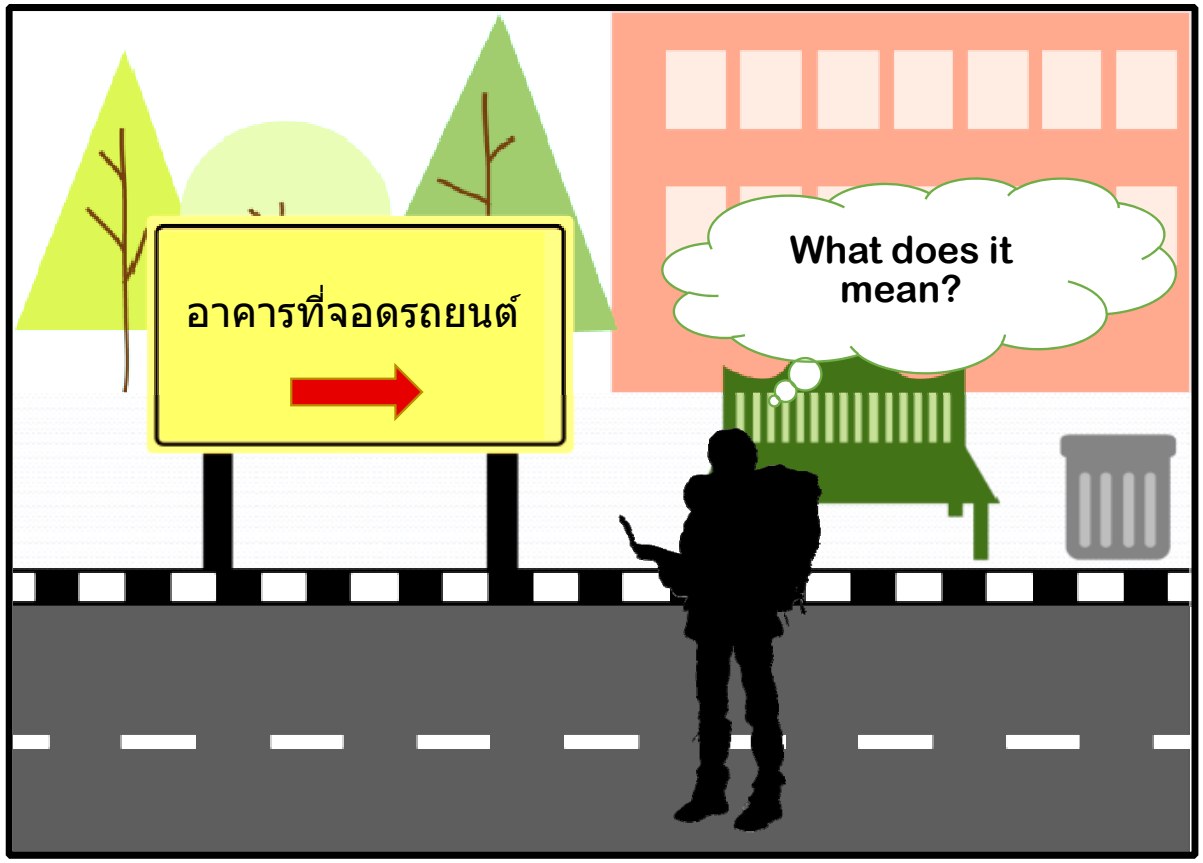
Presented by Ungsumalee Suttapakti
Faculty of Information Technology,
King Mongkut's Institute of Technology Ladkrabang,
Thailand

K. Woraratpanya, U. Suttapakti, K. Pasupa, P. Boonchukusol, T. Titijaroonroj, R. Hokking, Y. Kuroki and Y. Kato, "Text-Background Decomposition for Thai Text Localization and Recognition in Natural Scenes," the 6th International Conference on Information Technology and Electrical Engineering, October 7-8, 2014, pp. 138-143.

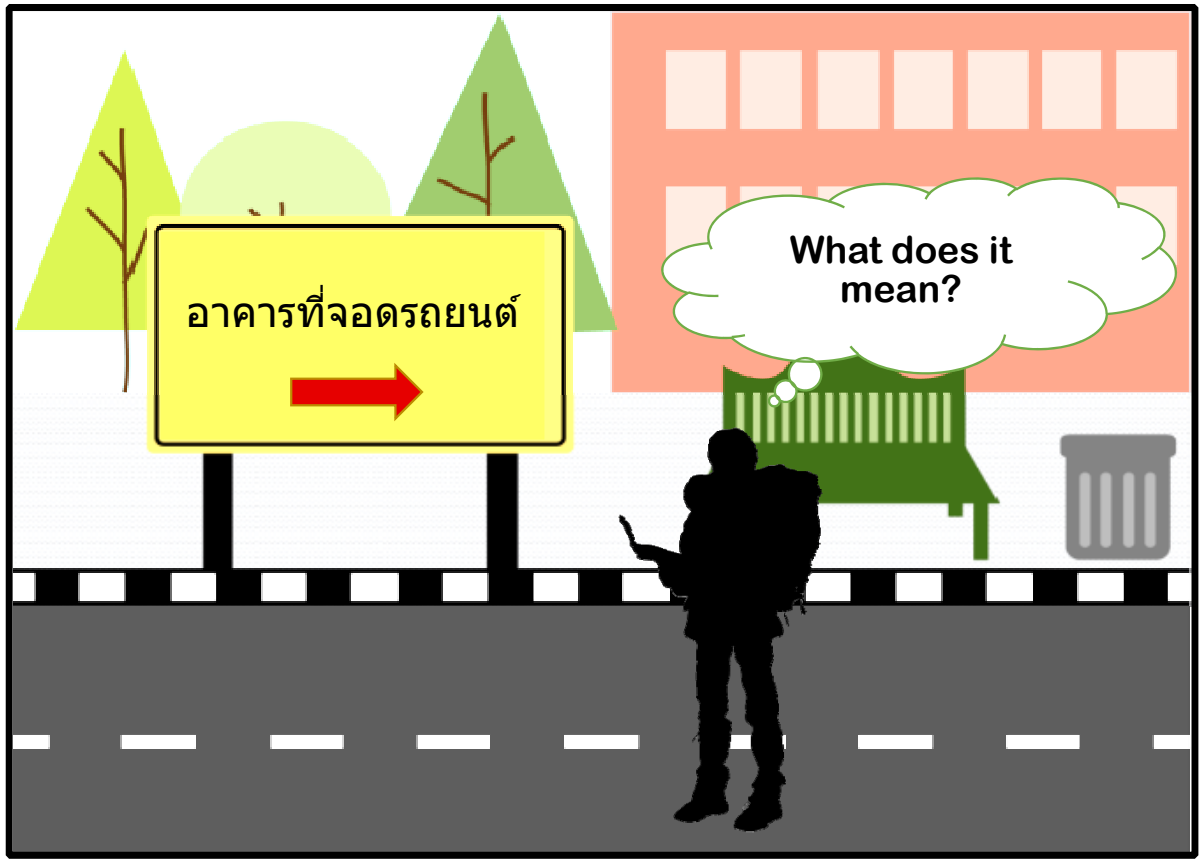
Outline

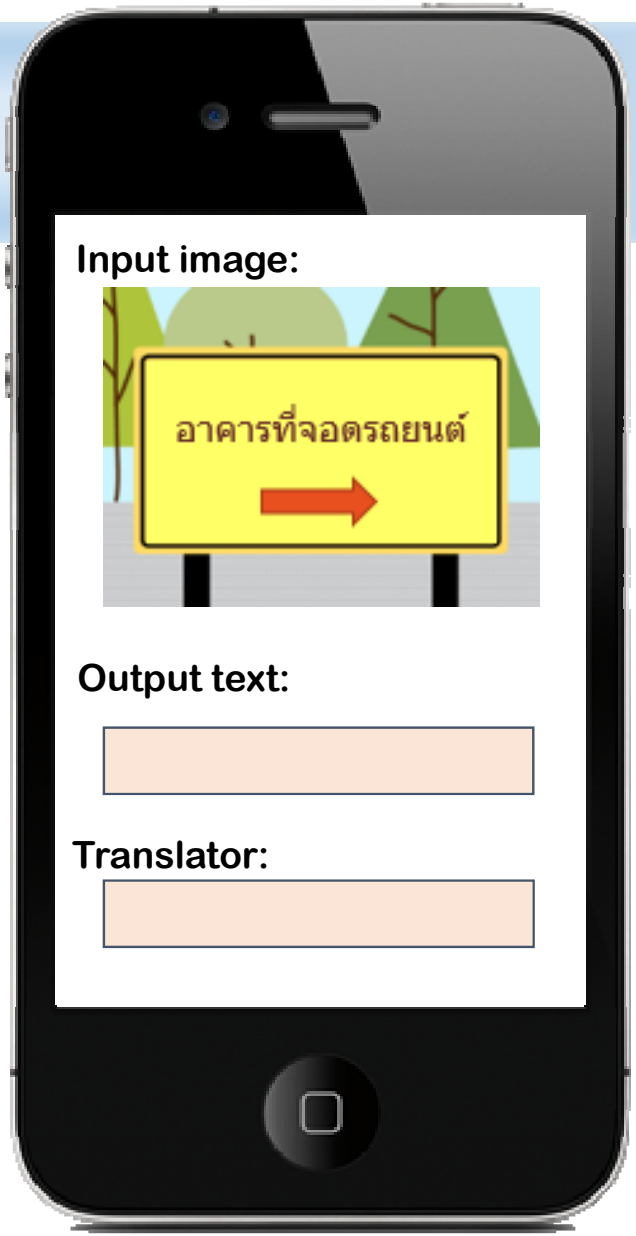
- Introduction
- Proposed methods
- Experimental results
- Conclusion

Introduction

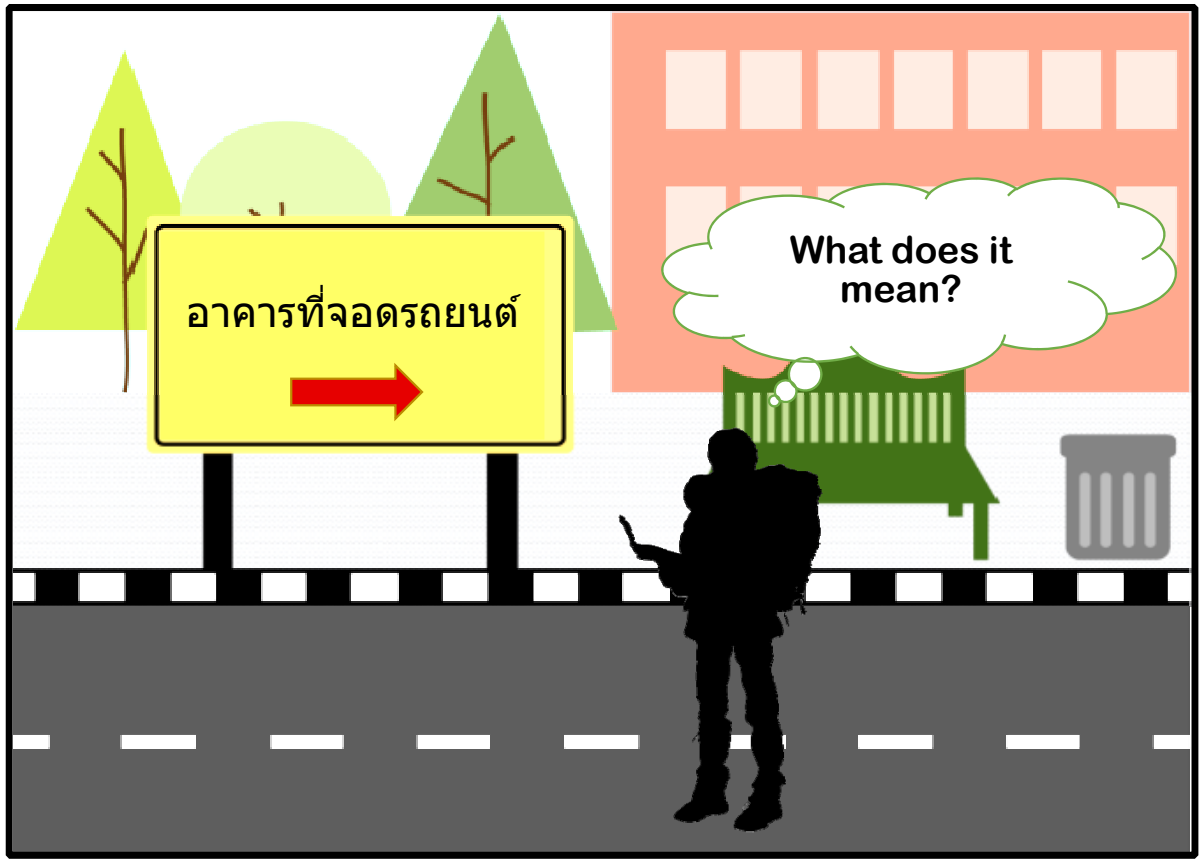


Introduction

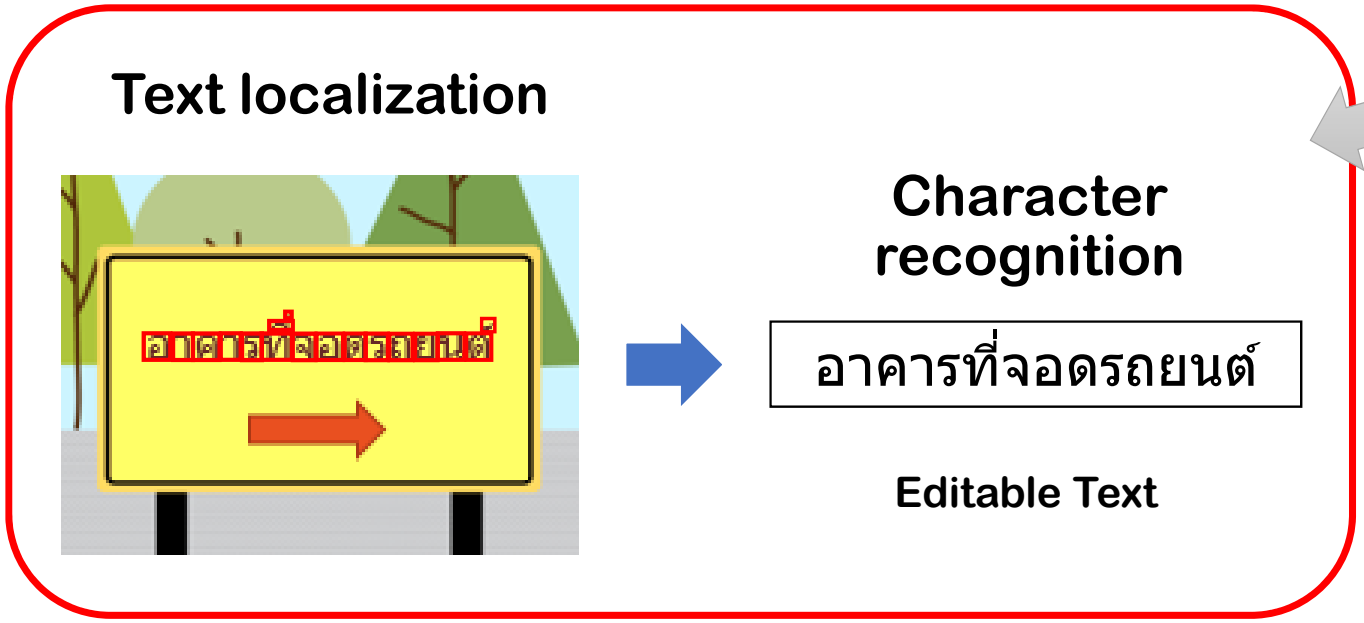




Introduction



Introduction



Challenges

I) Structure of language

Upper level

Middle level

Lower level

Thai text

English text

Challenges

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

III) Complex background

IV) Lighting condition

Related work

Multilayer Separation [1]

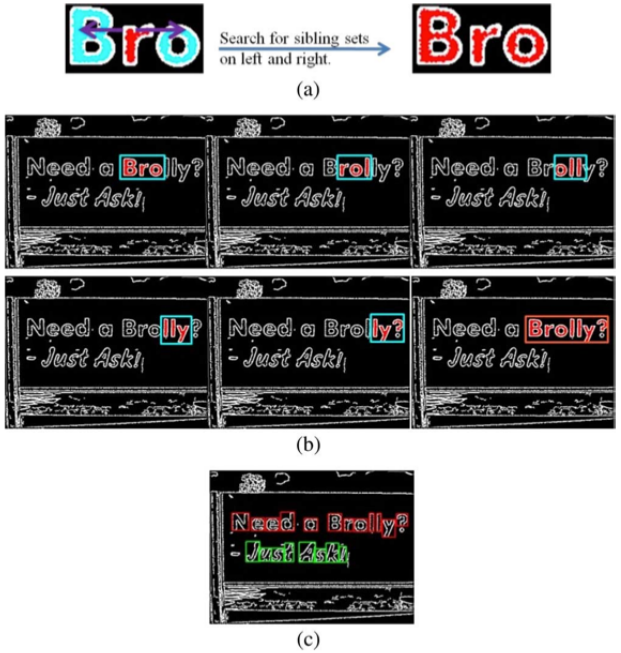
- Based on intensity value of the reduced-color image

[1] Kumar, M., Kim, Y. C., and Lee, G. S., "Text Detection using Multilayer Separation in Real Scene Images," Proc. 10th IEEE International Conference on Computer and Information Technology, pp. 1413-1417,2010.

Related work

Color-based partition [2]

- This algorithm detects text based on connected pixel with similar colors.



Cannot completely detect Thai text, especially vowels and tonal marks

Many detected non-text components

[2] Yi, C. and Tian, Y., "Text String Detection From Natural Scenes by Structure-Based Partition and Grouping," IEEE Trans. on Image Process., vol. 20, no. 9, pp. 2594-2605. September 2011.

Related work

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

5 clusters

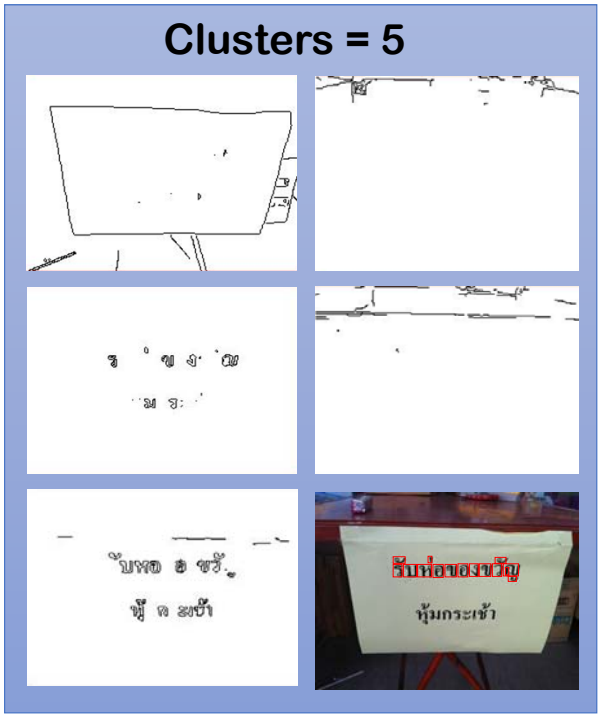
[3] Yi, C. and Tian, Y., "Localizing Text in Scene Images by Boundary Clustering, Stroke Segmentation, and String Fragment Classification," IEEE Trans. on Image Process., vol. 21, no. 9, pp. 4256-4268. September 2012.

Related work

Fast Boundary Clustering (FBC) [4]

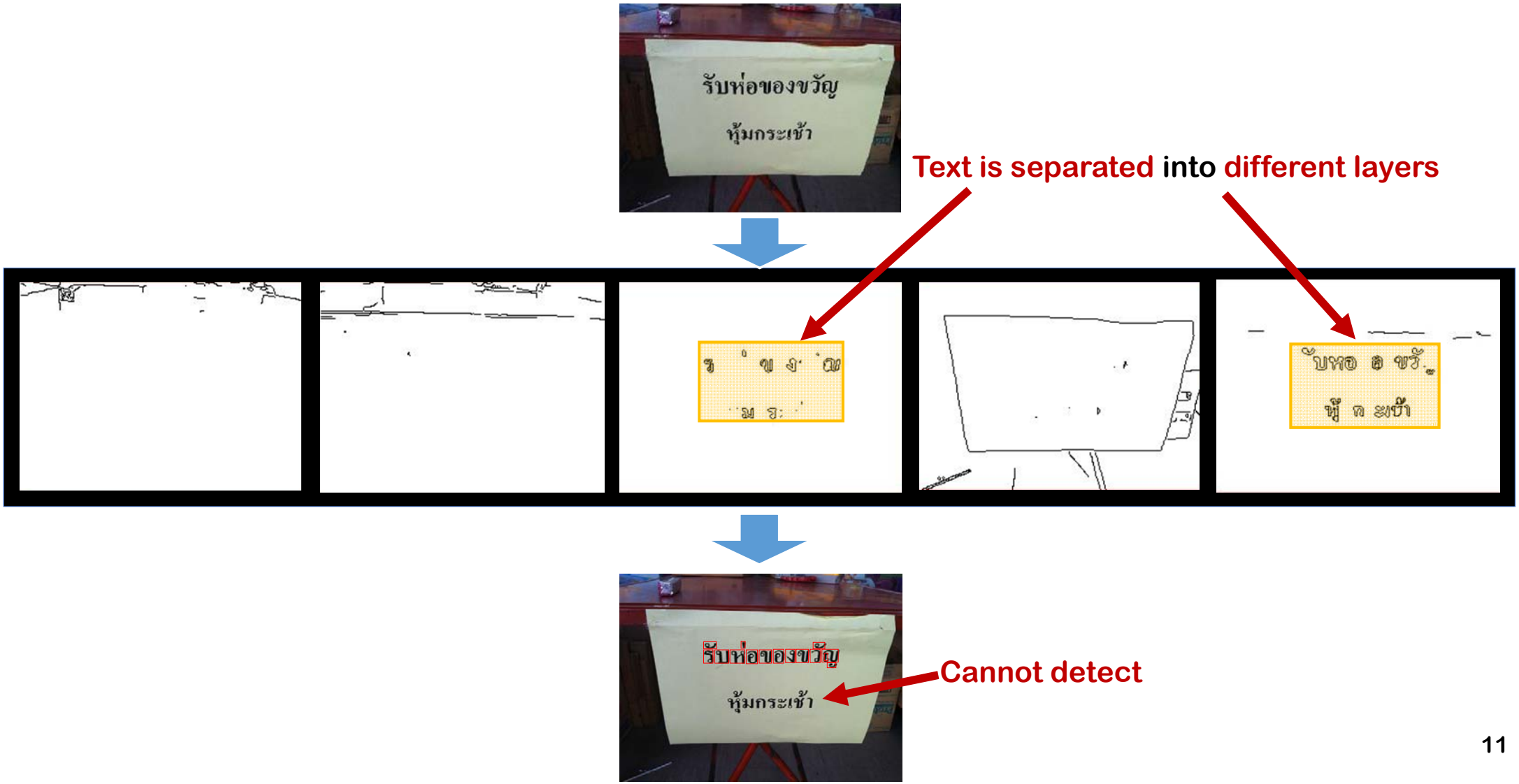
- FBC method extracts color features based on a pixel boundary.
- A number of clusters is fixed, 5 clusters.

Faster than



[4] K. Woraratpanya, P. Boonchukusol, Y. Kuroki, and Y. Kato, "Improved Thai Text Detection from Natural Scenes," Proc. 2013 International Conference on Information Technology and Electrical Engineering, pp. 137-142, 2013.

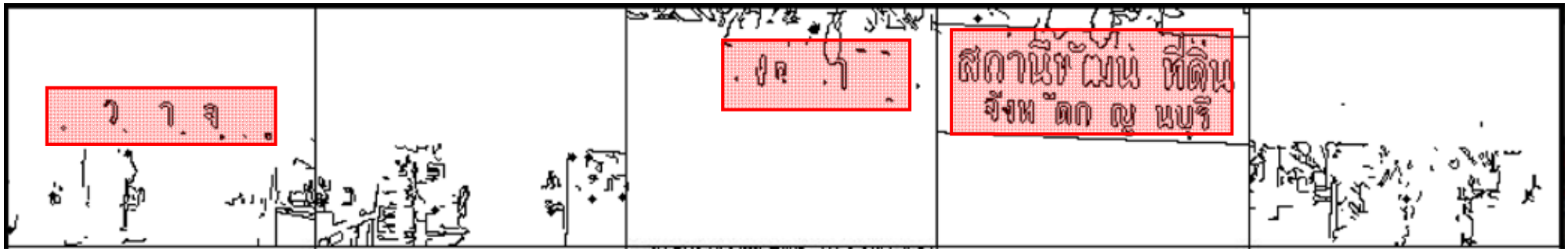
Problems: 1) The fixed number of clusters [3,4] is unsuitable, 5 clusters.



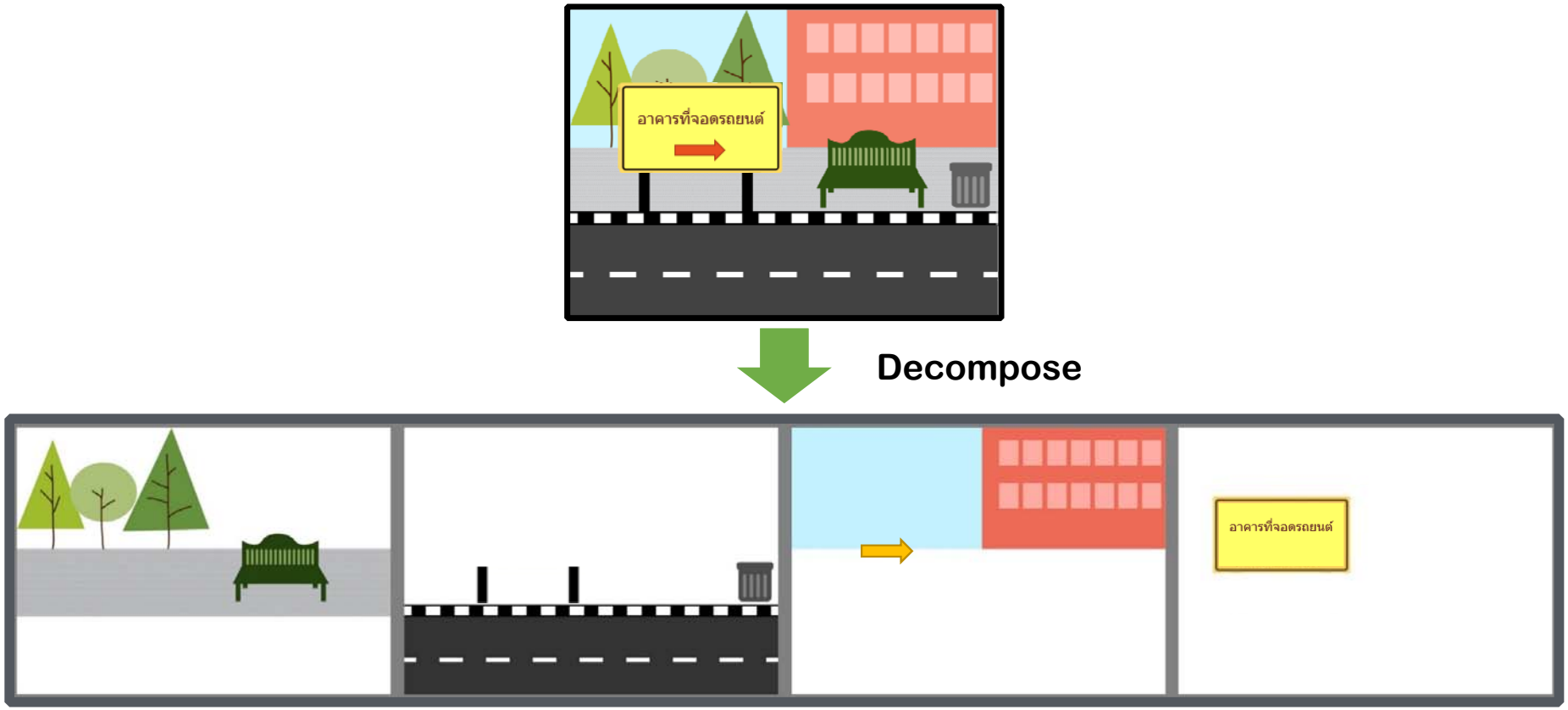
Problems: 2) The extracted color features based on a pixel boundary [4] cannot exactly represent the object when foreground and background are non-uniform colors.

Shadow

A failure of text-background decomposition

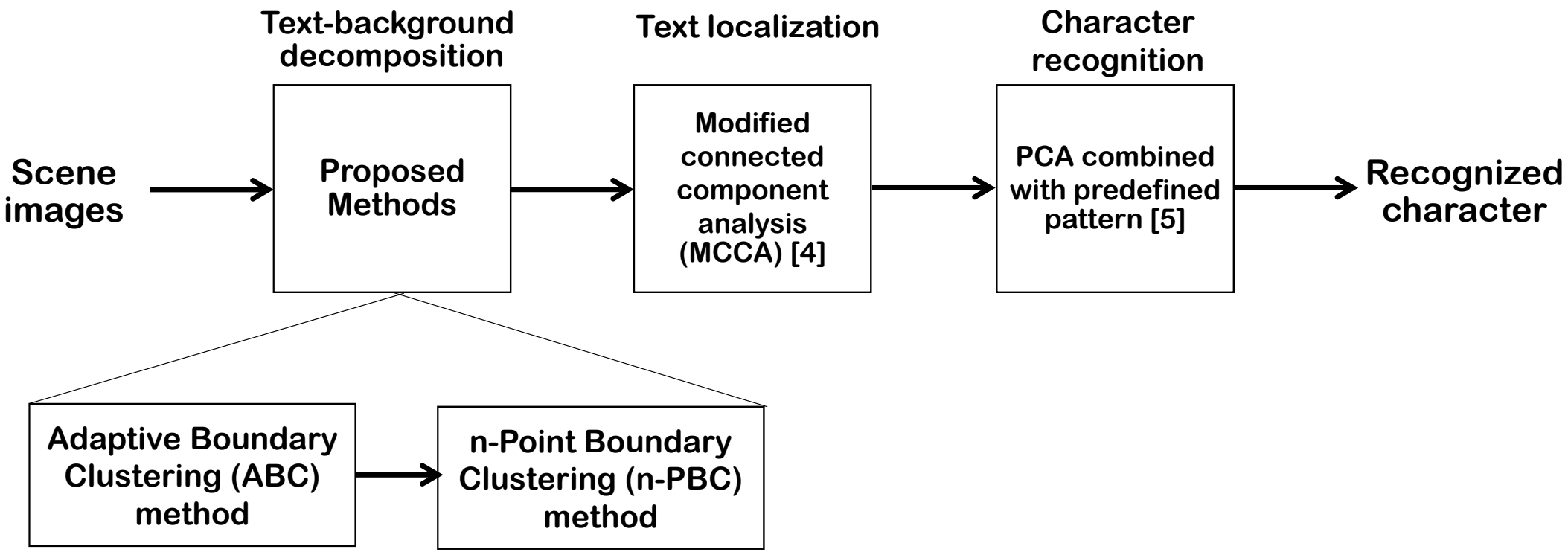


Text-background decomposition



Objectives: To improve the precision of text localization and the accuracy rate of character recognition

Overall framework



[4] K. Woraratpanya, P. Boonchukusol, Y. Kuroki, and Y. Kato, "Improved Thai Text Detection from Natural Scenes," Proc. 2013 International Conference on Information Technology and Electrical Engineering, pp. 137-142, 2013.
[5] U. Suttapakti, K. Woraratpanya, and K. Pasupa, "Font Descriptor Construction for Printed Thai Character Recognition," International Conference on Machine Vision Applications, pp. 45-48, 2013.

Outline

- Introduction
- Proposed methods
- Experimental results
- Conclusion

Proposed methods

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

Adaptive Boundary Clustering (ABC) 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

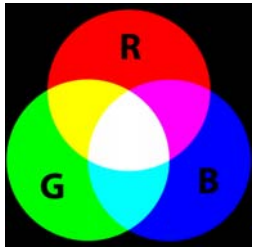
รับห่อของขวัญ
หุ้มกระเช้า

Editable text

Adaptive Boundary Clustering (ABC) method

Adaptive boundary clustering based on color quantization

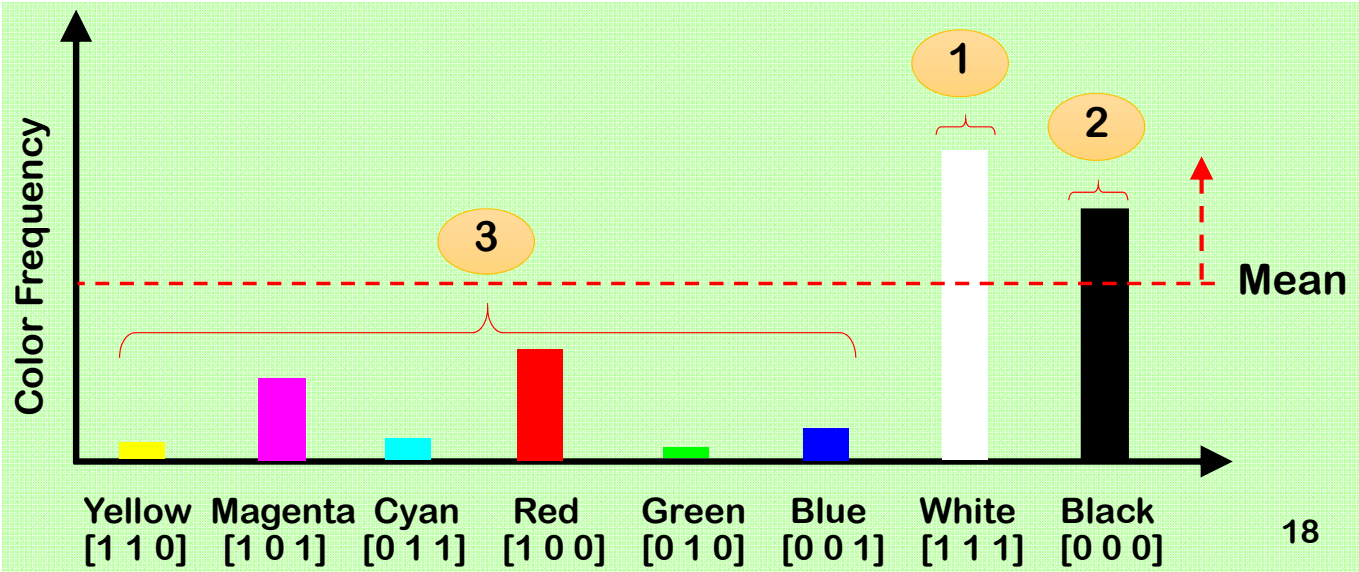
- Color quantization is a process that reduces the number of distinct colors used in a scene image.



$$k = n+1$$
$$= 2+1$$
$$= 3$$

k = the number of clusters.
n = the number of bins whose frequency is greater than mean of histogram.

Quantization

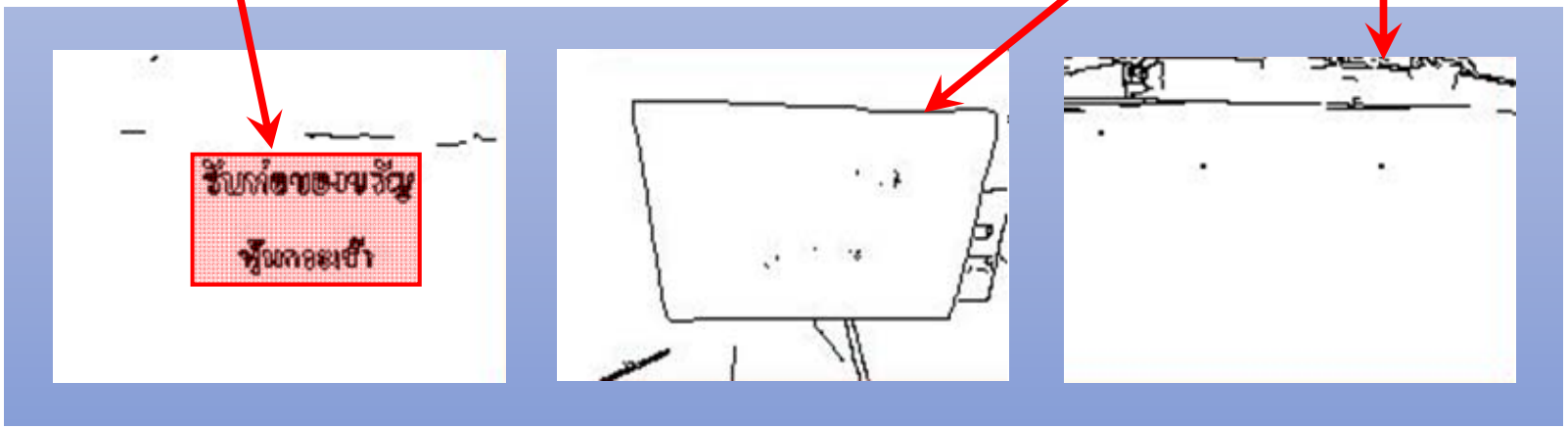


Adaptive Boundary Clustering (ABC) method

-Three clusters are suitable.

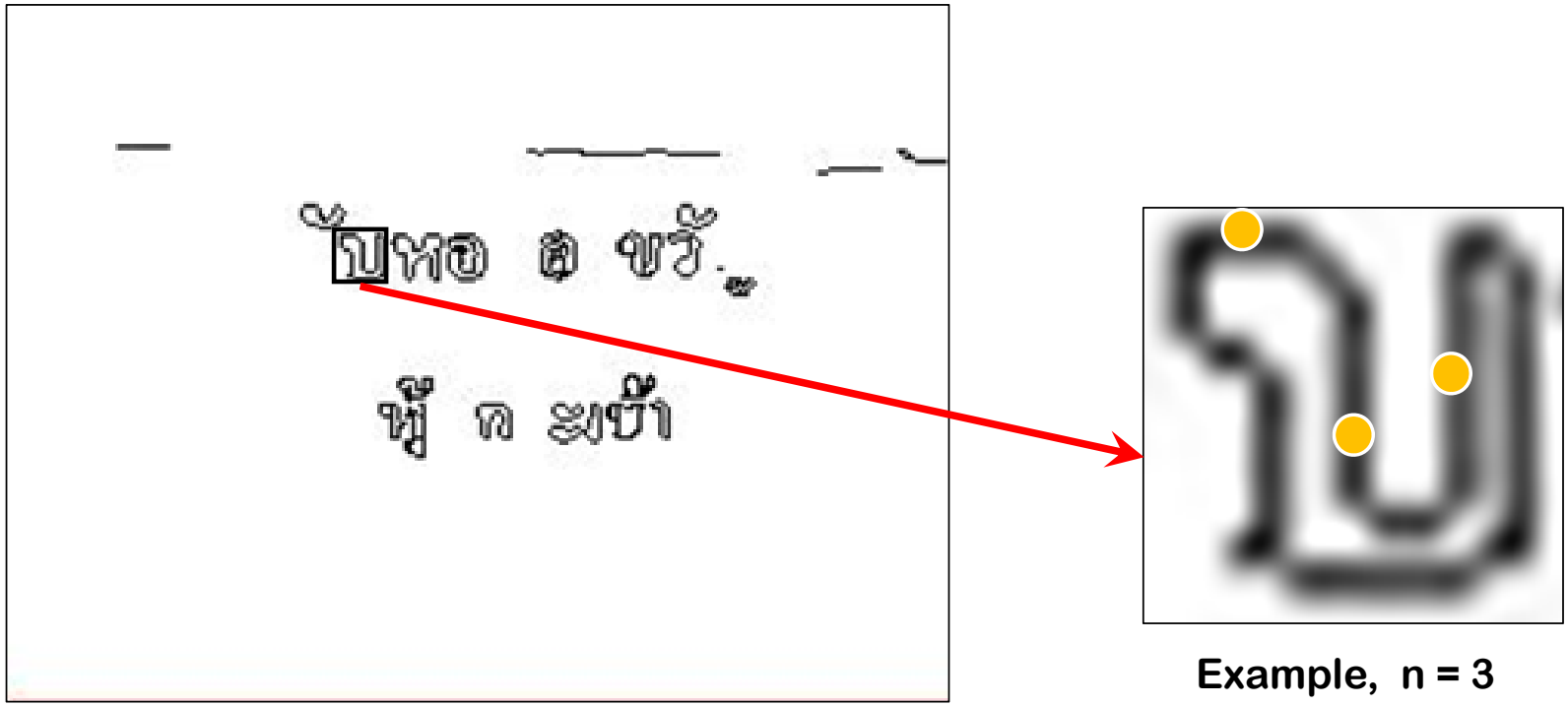
All text in the same layer.

Background Boundaries



n-Point Boundary Clustering (n-PBC) method

- The extracted color features based on an average of n-point boundary (n = 1, 2, 3, …, 8)

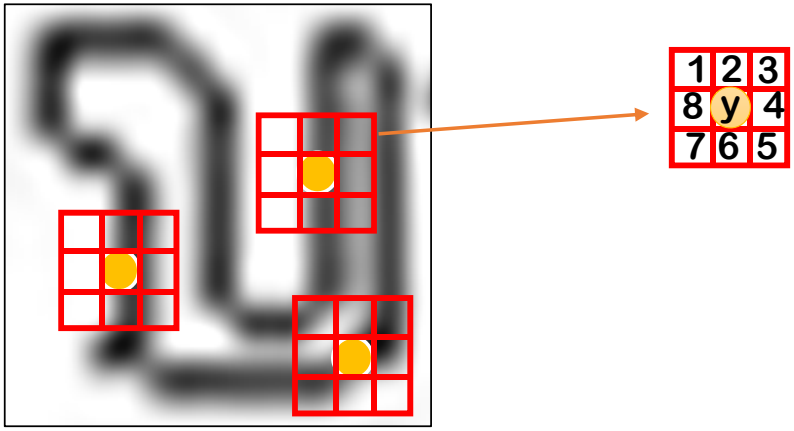


Example, n = 3

Feature vector extraction

- The feature of each object edge is extracted by determining the maximum color difference between two opposite pixels on a 3×3 mask.

Example; n = 3



An object edge

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

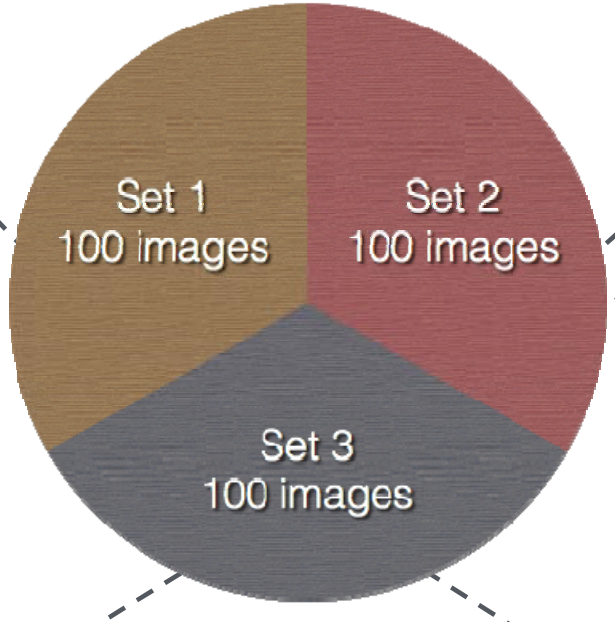
$[R_h, G_h, B_h]$ are pixel color values of the highest intensity.
$[R_l, G_l, B_l]$, are pixel color values of the lowest intensity.
$[C_y]$ is y-coordinate of the centroid of boundary.

Outline

- Introduction
- Proposed methods
- Experimental results
- Conclusion

Datasets

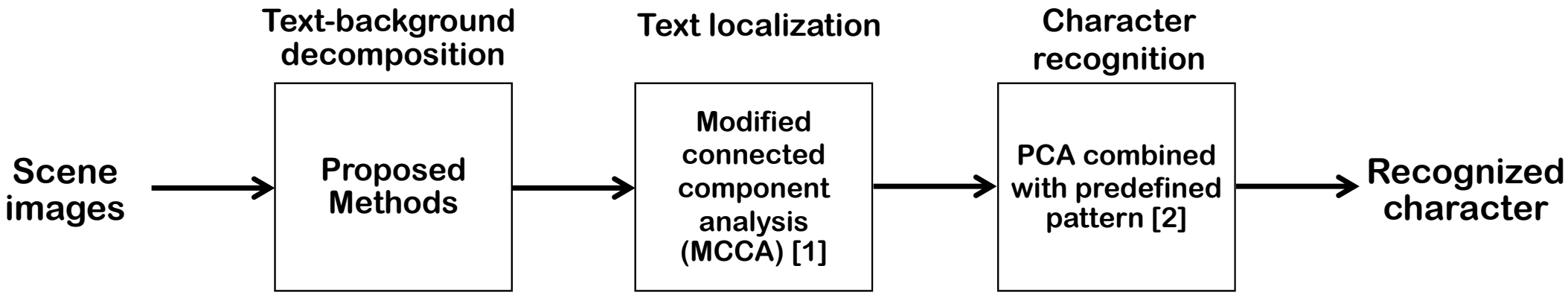
A uniform color



Complex background

Lighting condition

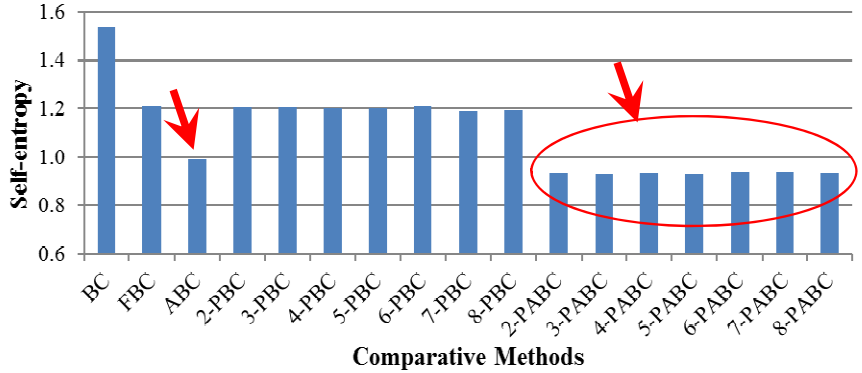
Evaluation



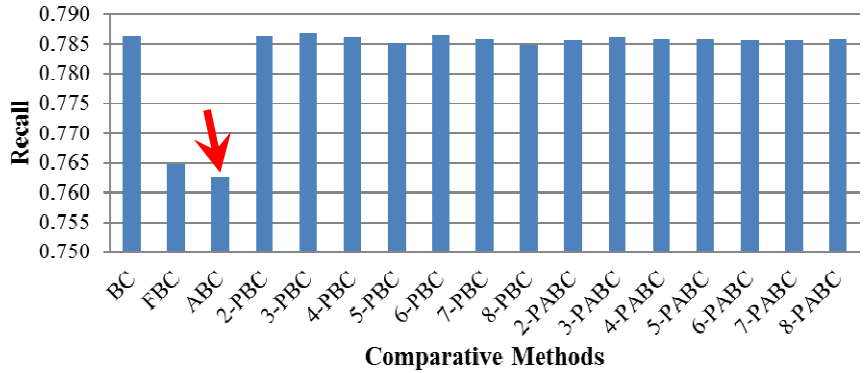
Evaluation:	Self-entropy	Precision and recall	Recognition rate
Targets:	lower, better	higher, better	higher, better
Evaluation:	Computing time	Computing time	Computing time
Targets:	lower, better	lower, better	lower, better

Experimental results

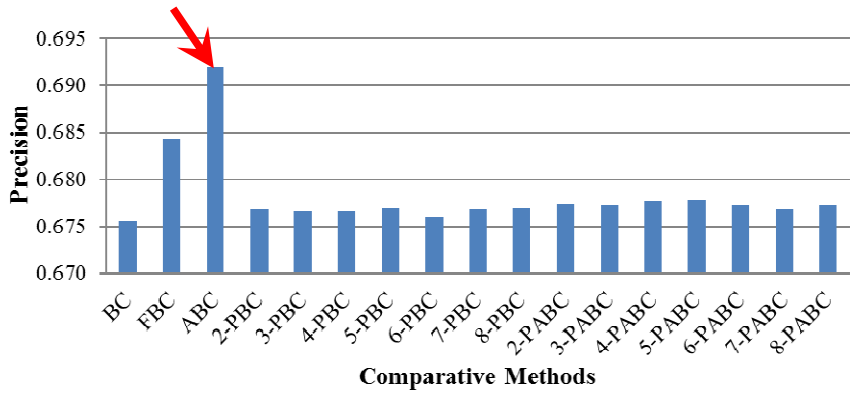
Self-entropy



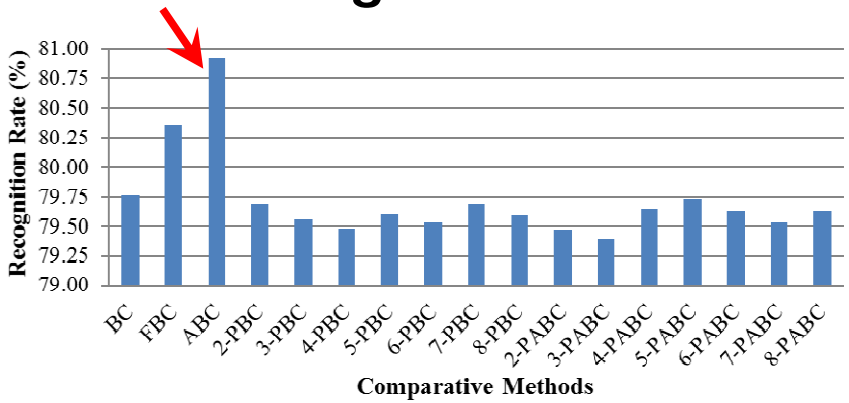
Recall



Precision

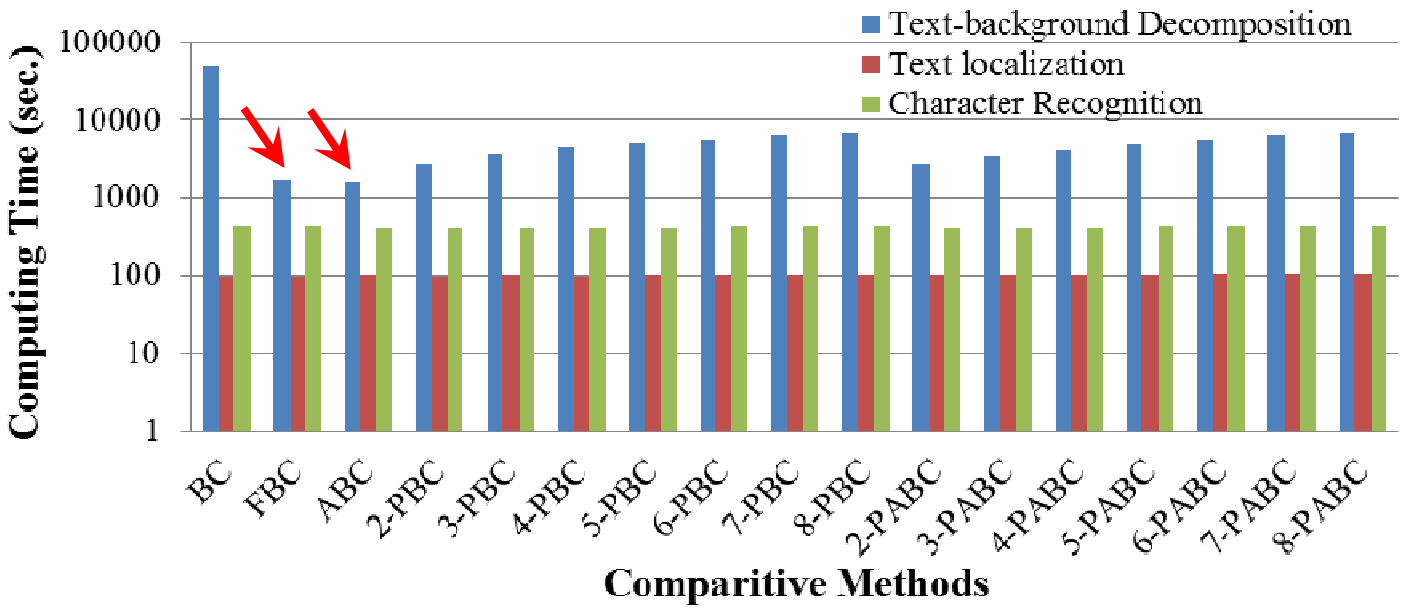


Recognition rate



Experimental results

Computing time



Conclusion

- 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.

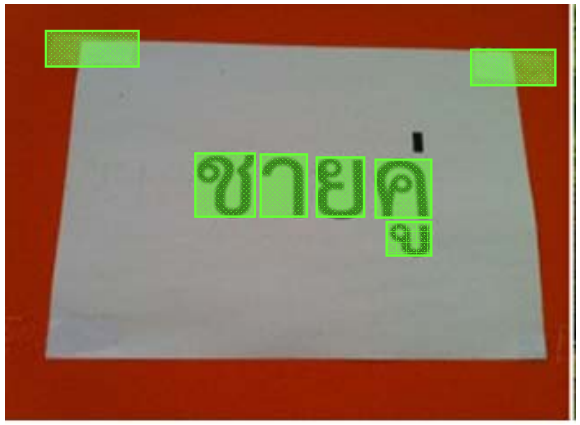
$P(y_i)$ is the ratio of the number of non-text boundaries to the total number of boundaries and k denotes the number of clusters.

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

Evaluation

$$precision = \frac{tp}{tp + fp} = \frac{5}{5+2} = 0.71$$

$$recall = \frac{tp}{tp + fn} = \frac{5}{5+1} = 0.83$$



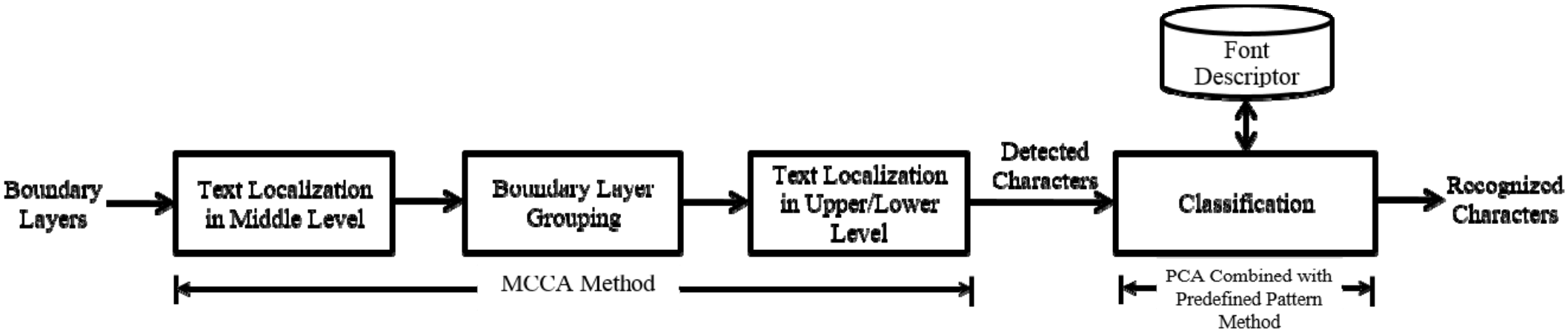
Predicted class

Actual class

Classes	Text	Non-text
Text	$tp = 5$	$fn = 1$
Non-text	$fp = 1$	$tn = 0$

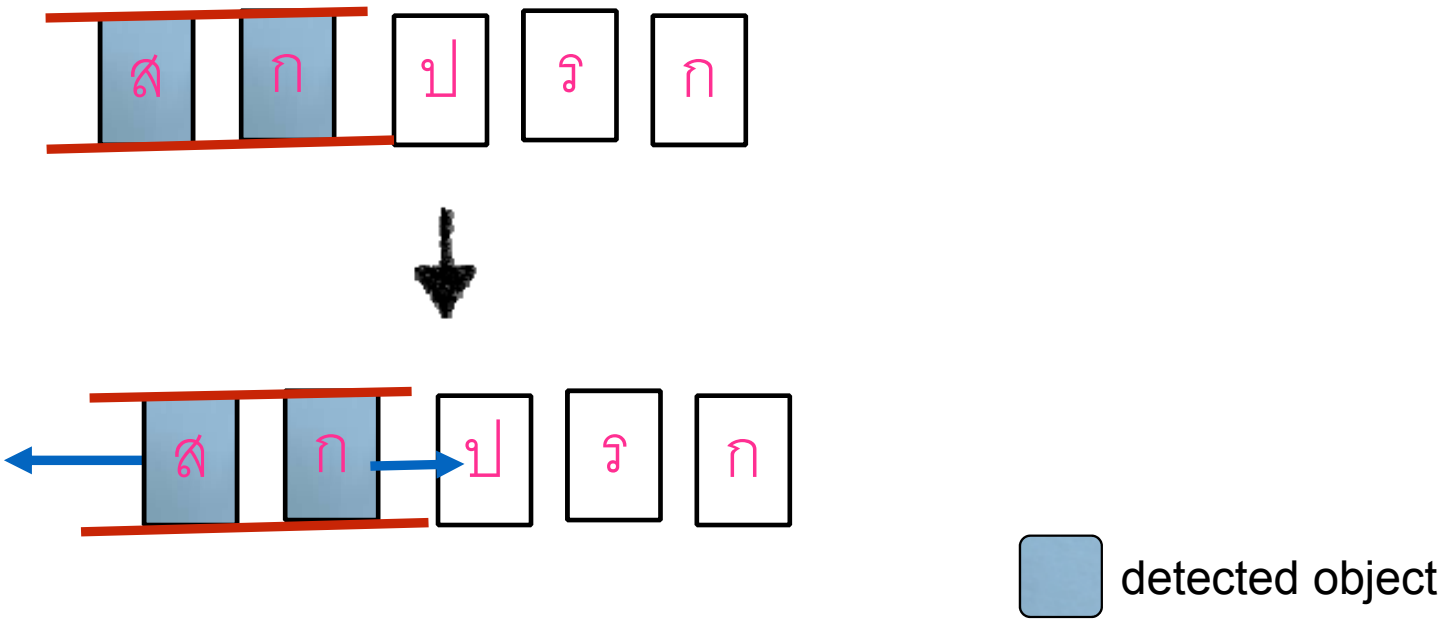
Text Localization and Character Recognition

- A diagram of text localization and character recognition.



Text-background Decomposition

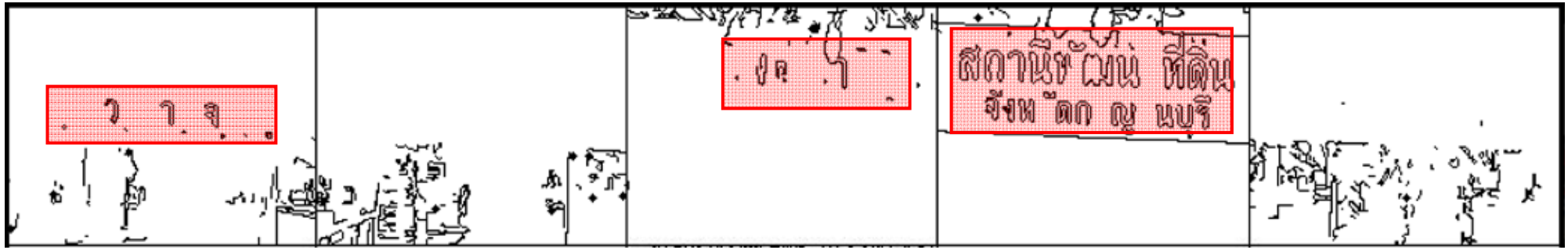
Modified Body of Text Detection



Scene image

Comparative FBC and n-PBC method

A failure of text-background decomposition of FBC method



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