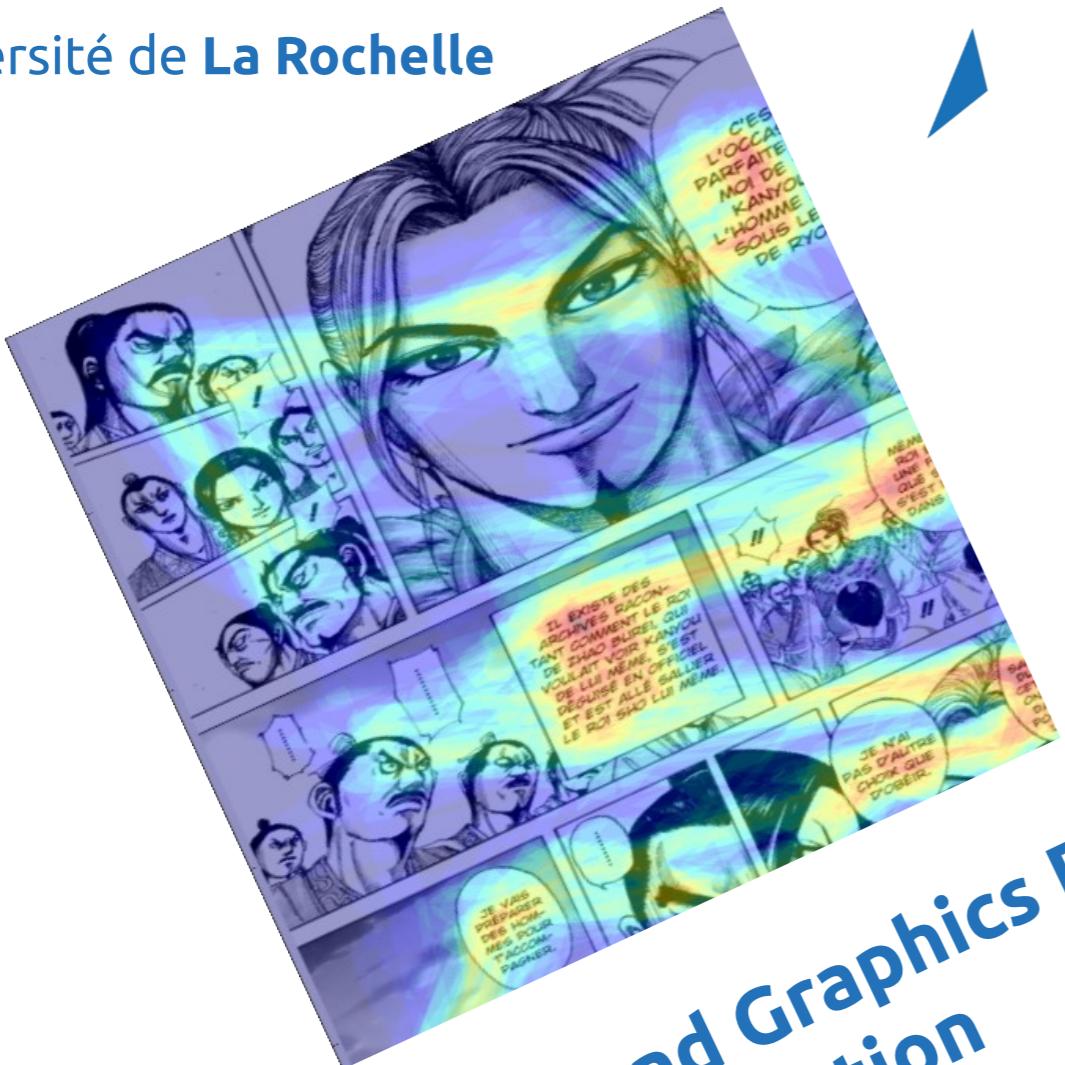


Semi-automatic Text and Graphics Extraction of Manga Using Eye Tracking Information

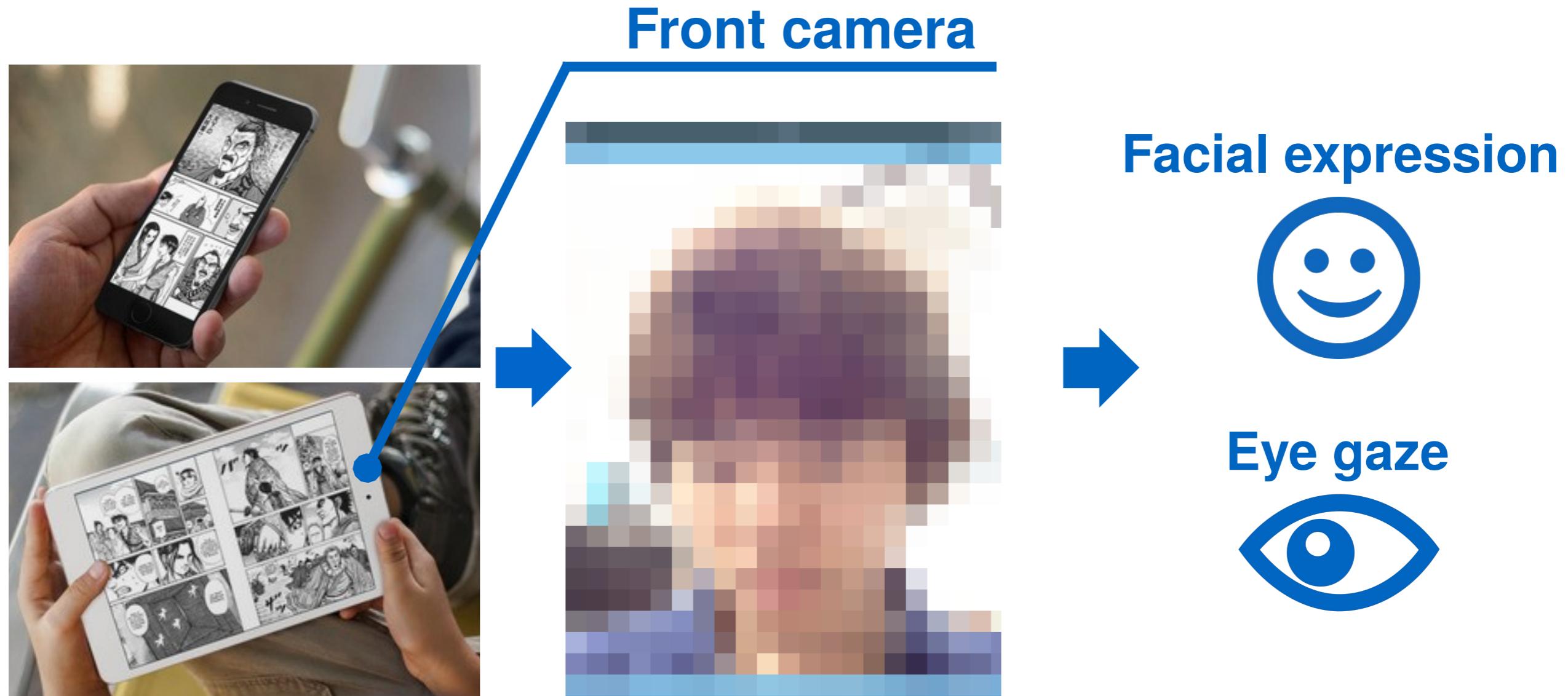
Christophe Rigaud*, Thanh-Nam Le*, Shoya Ishimaru°,
J.-C. Burie*, J.-M. Ogier*, Motoi Iwata°, Koichi Kise°



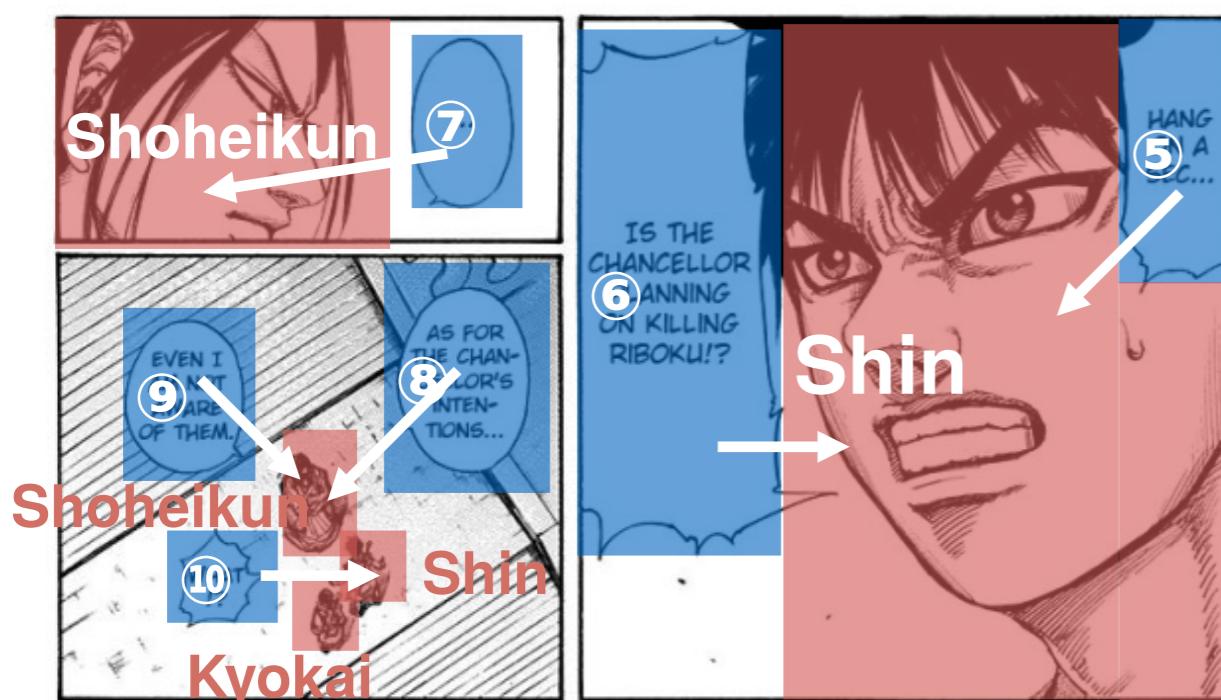
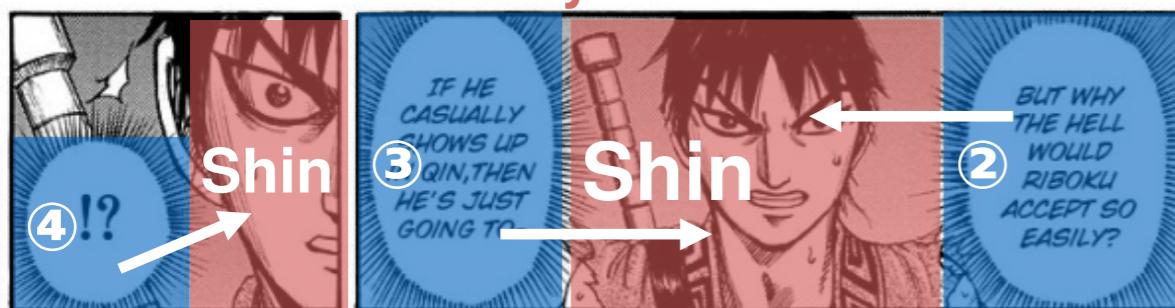
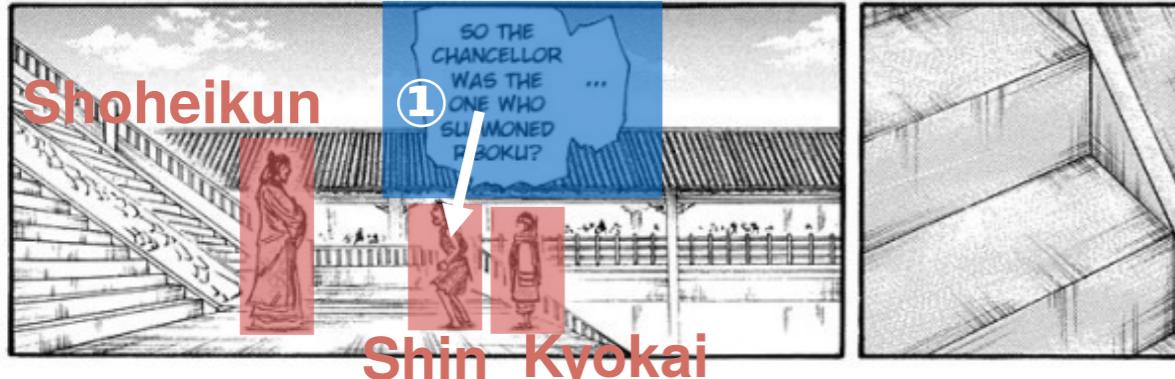
Summary

- Introduction
- Related work
- Proposed approach
- Evaluation
- Conclusion

Why do we focus on manga-reading behavior?

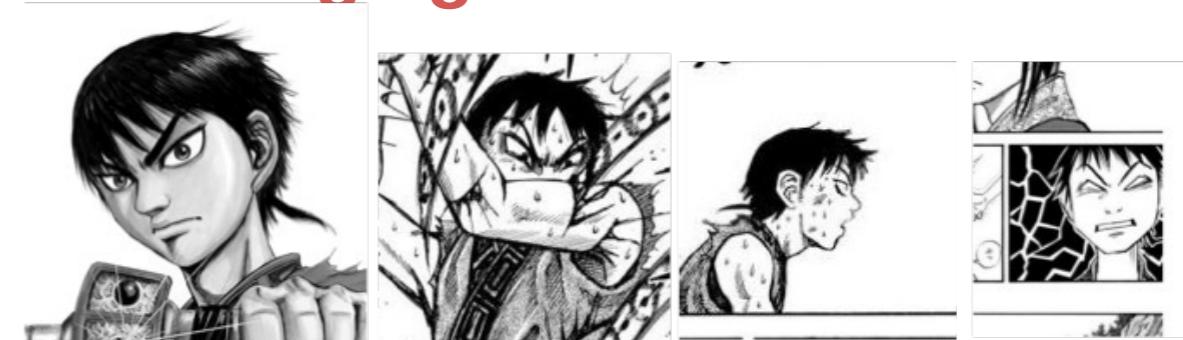


Manga layout analysis

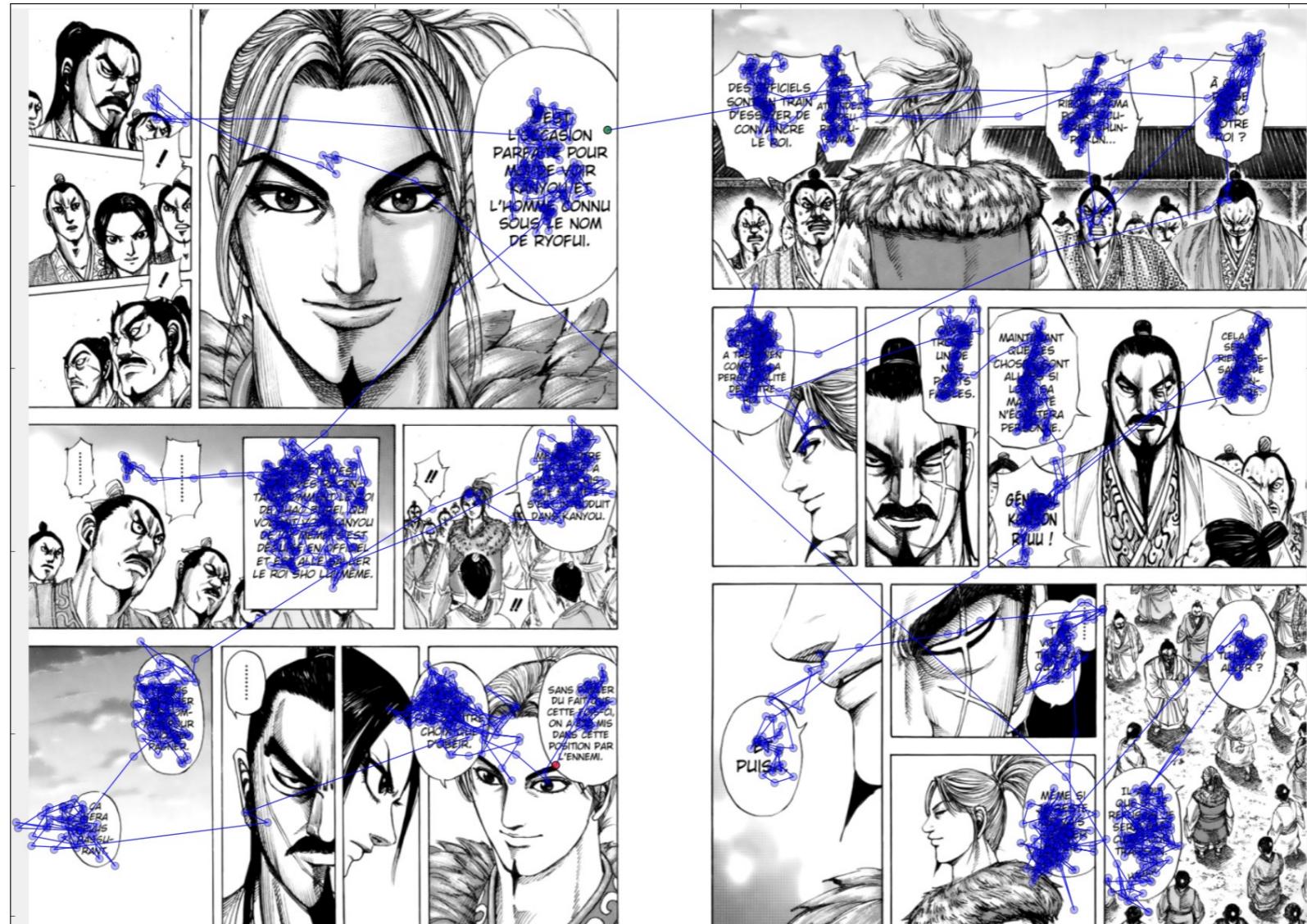


Problem

Sequential art image layout analysis
is challenging



Manga layout analysis using eye tracker



Raw eye gaze position (42Hz)

Can we separate text from graphics by analyzing eye movement?

Related works

Related works

Document recognition using eye tracker - Kunze *et al.* [1]

Saccade direction count/mean/variance

Decision tree classifier

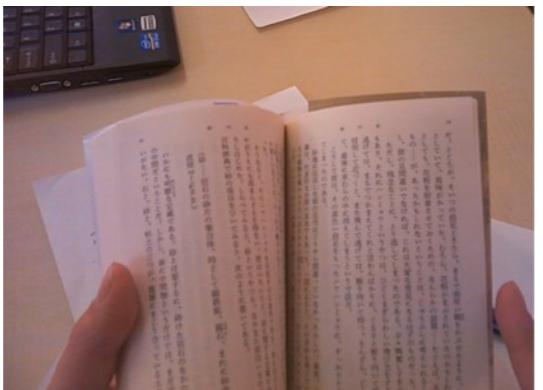
SMI eye tracking glasses (30Hz)



text book



novel



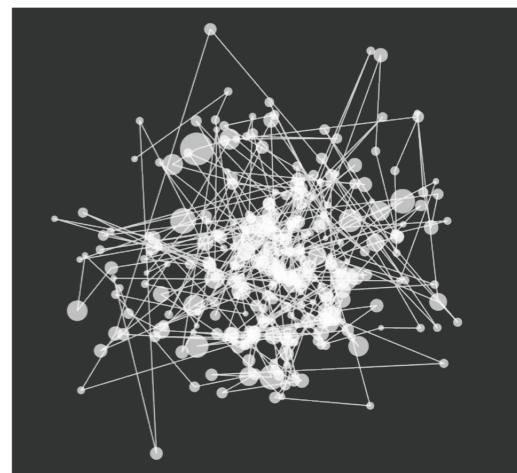
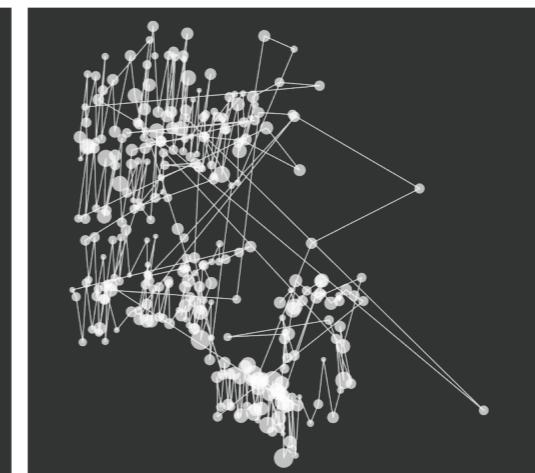
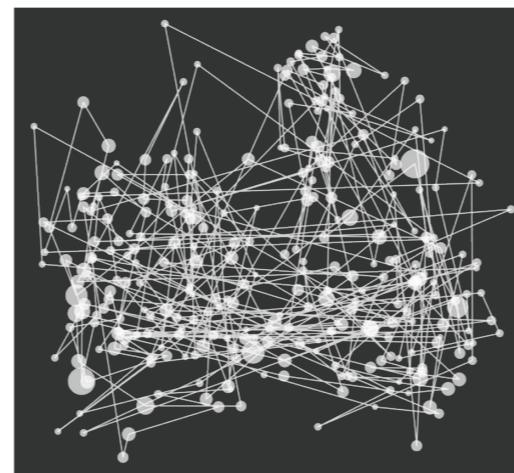
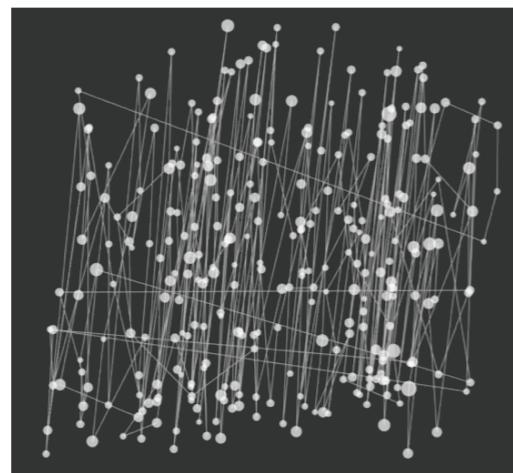
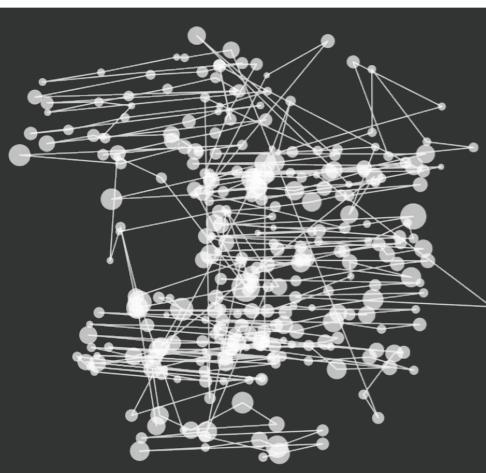
magazine



news paper



manga



- [1] K. Kunze, Y. Utsumi, Y. Shiga, K. Kise, and A. Bulling, “I know what you are reading: recognition of document types using mobile eye tracking,” in *proceedings of 17th annual international symposium on International symposium on wearable computers, (ISWC)*, pp. 113—116, 2013.
- [2] C. Rigaud, J.-C. Burie, and J.-M. Ogier, “Text-independent speech balloon segmentation for comics and manga,” in *proceedings of the 11th IAPR International Workshop on Graphics Recognition (GREG)*, Nancy, France, 2015.
- [3] M. Iwata, A. Ito, and K. Kise, “A study to achieve manga character retrieval method for manga images” in *proceedings of 11th IAPR International workshop on Document Analysis System (DAS)*, pp. 309–313, 2014.

Related works

Document recognition using eye tracker - Kunze *et al.* [1]

Saccade direction count/mean/variance

Decision tree classifier

SMI eye tracking glasses (30Hz)



text book



novel



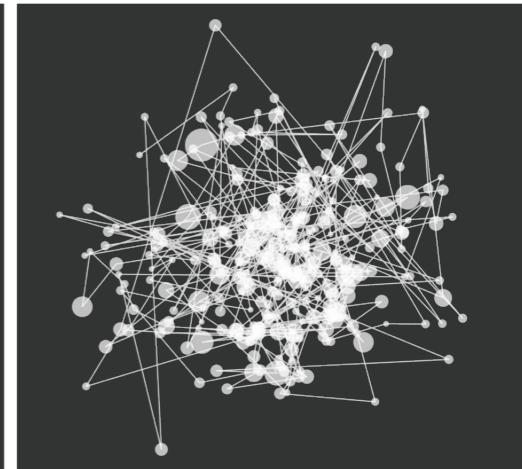
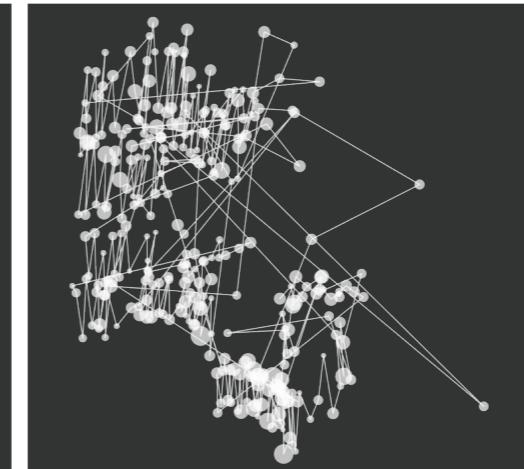
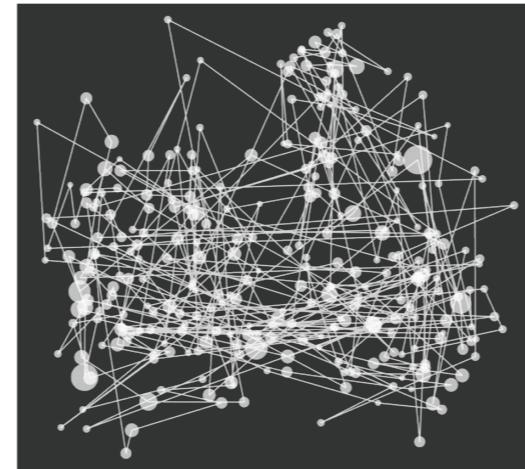
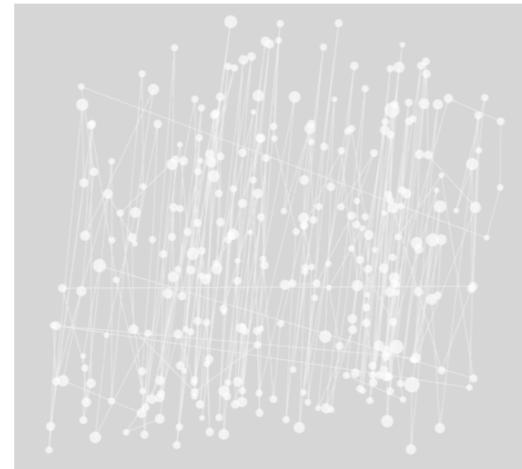
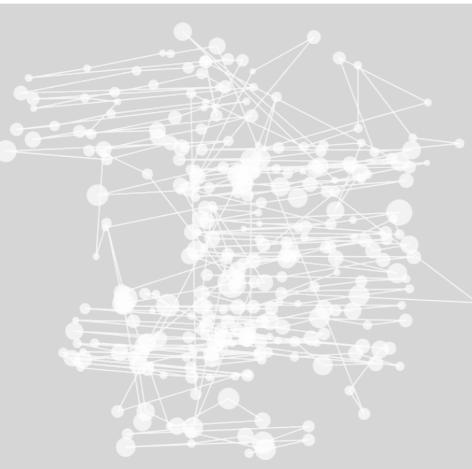
magazine



news paper



manga



- [1] K. Kunze, Y. Utsumi, Y. Shiga, K. Kise, and A. Bulling, “I know what you are reading: recognition of document types using mobile eye tracking,” in *proceedings of 17th annual international symposium on International symposium on wearable computers, (ISWC)*, pp. 113—116, 2013.
- [2] C. Rigaud, J.-C. Burie, and J.-M. Ogier, “Text-independent speech balloon segmentation for comics and manga,” in *proceedings of the 11th IAPR International Workshop on Graphics Recognition (GREG)*, Nancy, France, 2015.
- [3] M. Iwata, A. Ito, and K. Kise, “A study to achieve manga character retrieval method for manga images” in *proceedings of 11th IAPR International workshop on Document Analysis System (DAS)*, pp. 309–313, 2014.

Related works

Document recognition using eye tracker - Kunze *et al.* [1]

Saccade direction count/mean/variance

Decision tree classifier

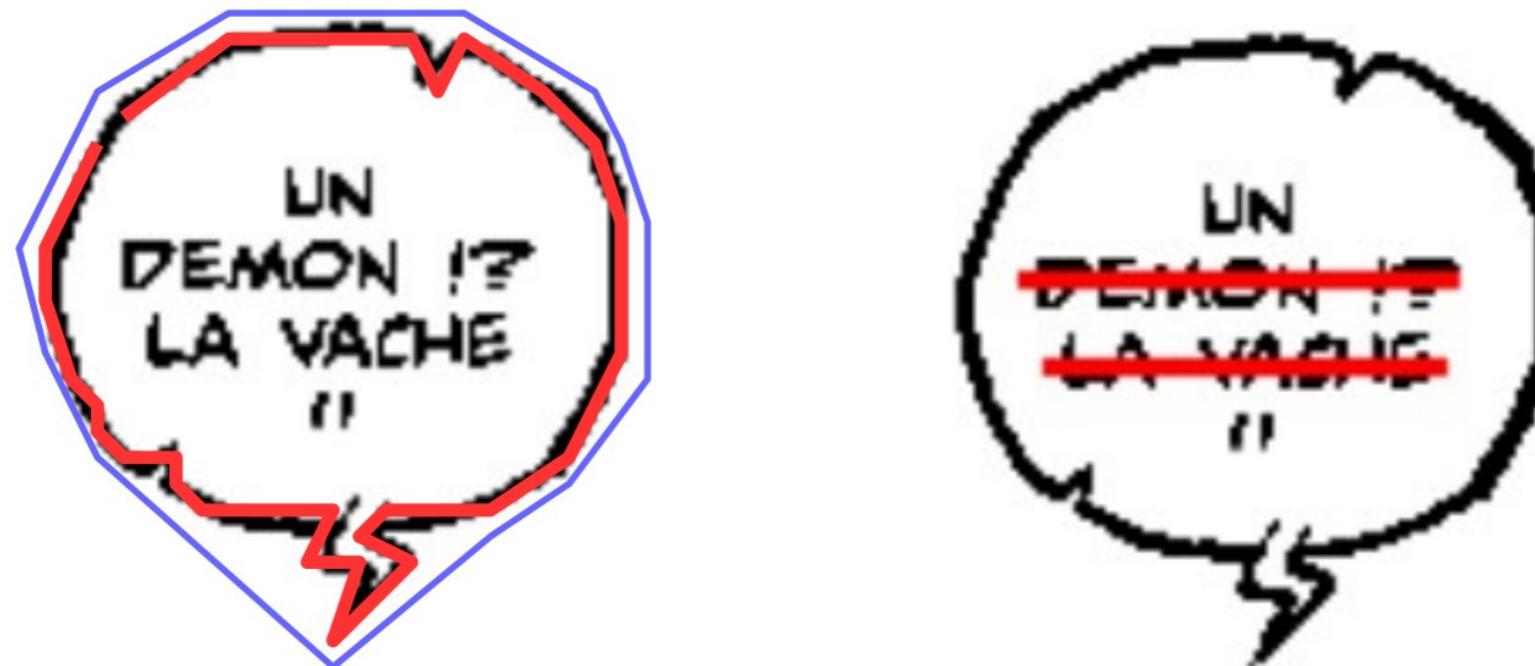
SMI eye tracking glasses (30Hz)

Speech balloon segmentation - Rigaud *et al.* [2]

Connected-component analysis

Overall convexity and content alignment

Limited to closed balloons



[1] K. Kunze, Y. Utsumi, Y. Shiga, K. Kise, and A. Bulling, "I know what you are reading: recognition of document types using mobile eye tracking," in proceedings of 17th annual international symposium on International symposium on wearable computers, (ISWC), pp. 113—116, 2013.

[2] C. Rigaud, J.-C. Burie, and J.-M. Ogier, "Text-independent speech balloon segmentation for comics and manga," in proceedings of the 11th IAPR International Workshop on Graphics Recognition (GREC), Nancy, France, 2015.

[3] M. Iwata, A. Ito, and K. Kise, "A study to achieve manga character retrieval method for manga images" in proceedings of 11th IAPR International workshop on Document Analysis System (DAS), pp. 309–313, 2014.

Related works

Document recognition using eye tracker - Kunze *et al.* [1]

Saccade direction count/mean/variance

Decision tree classifier

SMI eye tracking glasses (30Hz)

Speech balloon segmentation - Rigaud *et al.* [2]

Connected-component analysis

Overall convexity and content alignment

Limited to closed balloons

Manga character detection - Iwata *et al.* [3]

MSER + HOG

Few manually labeled data are required



[1] K. Kunze, Y. Utsumi, Y. Shiga, K. Kise, and A. Bulling, "I know what you are reading: recognition of document types using mobile eye tracking," in proceedings of 17th annual international symposium on International symposium on wearable computers, (ISWC), pp. 113—116, 2013.

[2] C. Rigaud, J.-C. Burie, and J.-M. Ogier, "Text-independent speech balloon segmentation for comics and manga," in proceedings of the 11th IAPR International Workshop on Graphics Recognition (GREC), Nancy, France, 2015.

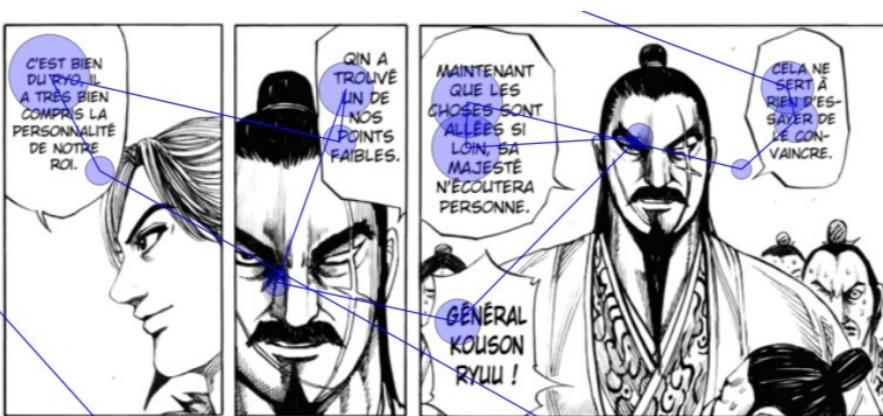
[3] M. Iwata, A. Ito, and K. Kise, "A study to achieve manga character retrieval method for manga images" in proceedings of 11th IAPR International workshop on Document Analysis System (DAS), pp. 309–313, 2014.

Proposed approach

Overview of Proposed Method



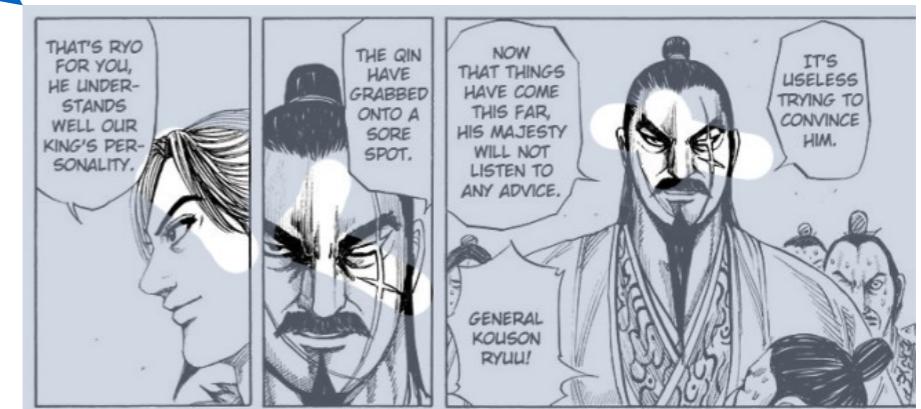
Fixation-saccade detection



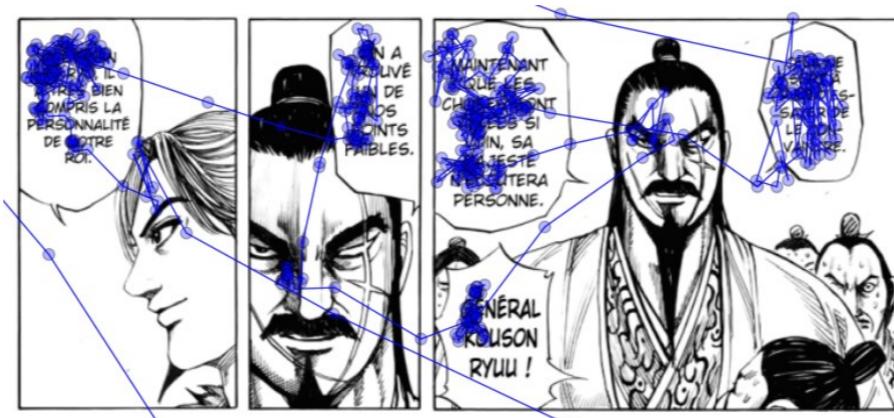
Text region extraction



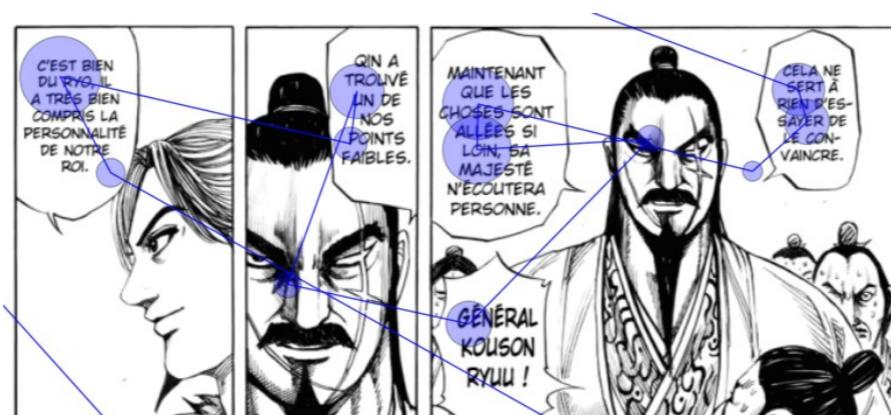
Graphics region extraction



Overview of Proposed Method



Fixation-saccade detection



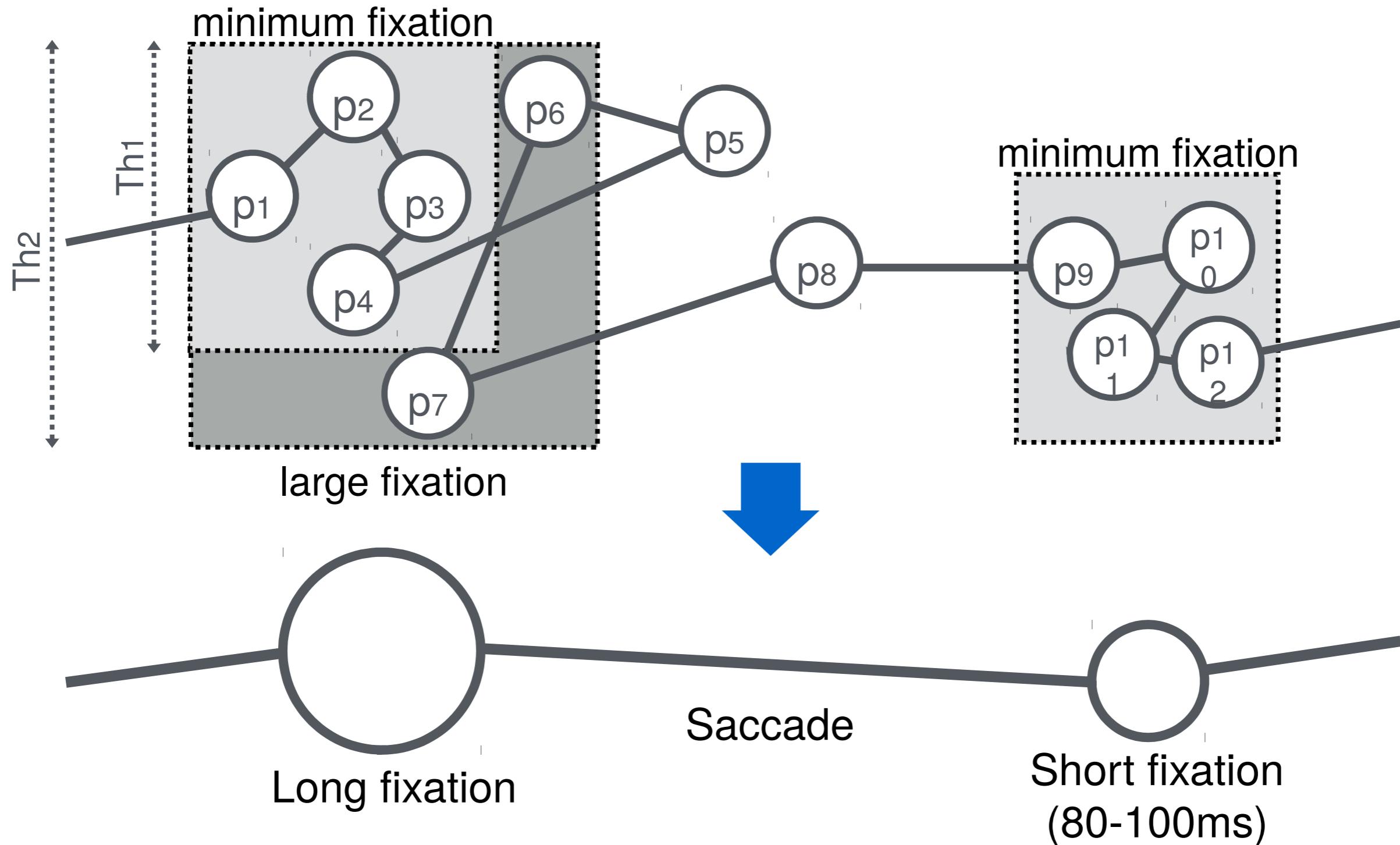
Text region
extraction



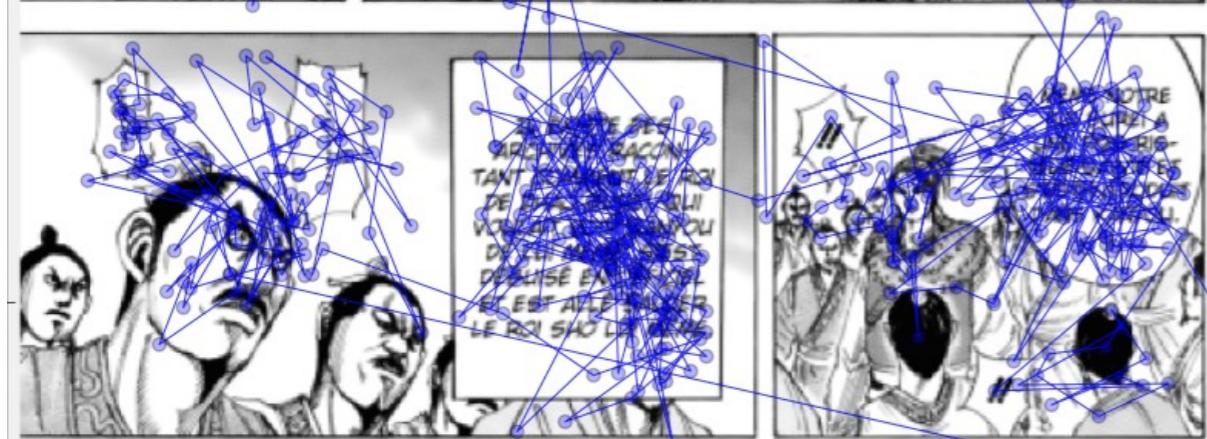
Graphics region
extraction



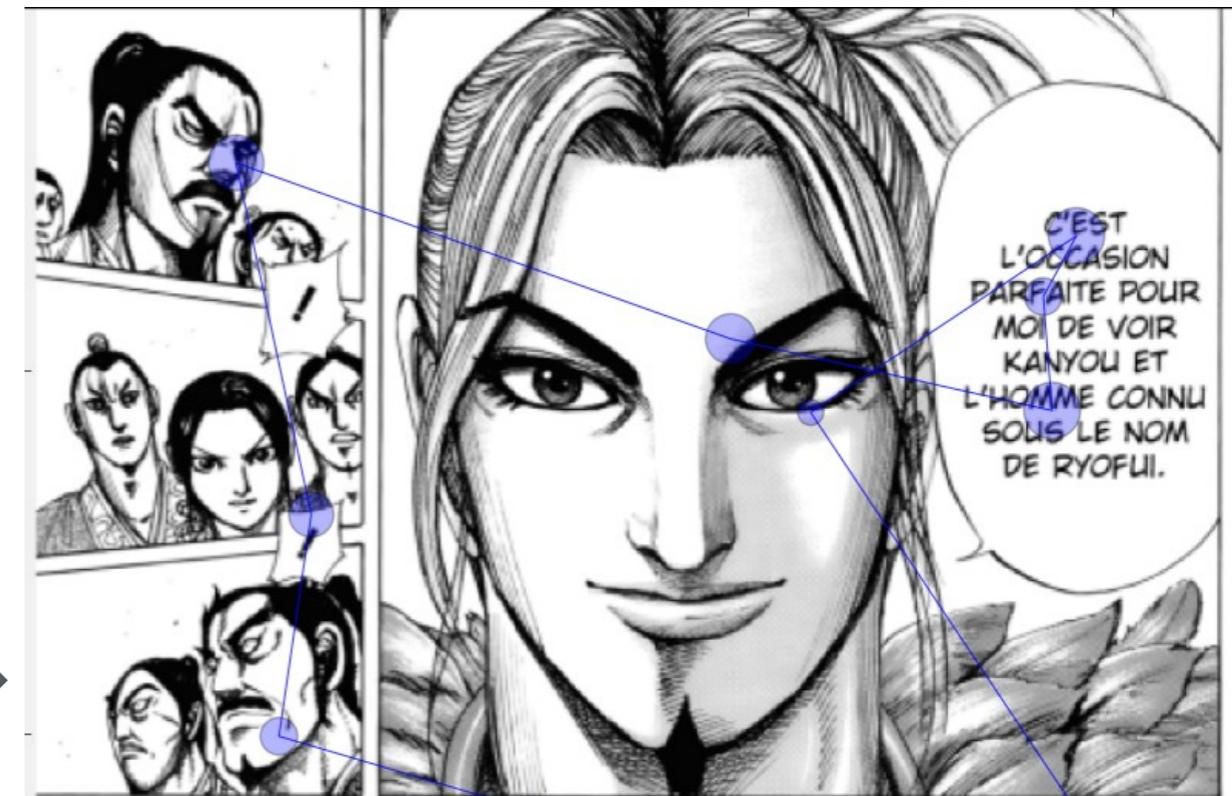
Fixation-saccade detection [4]



Fixation-saccade detection output

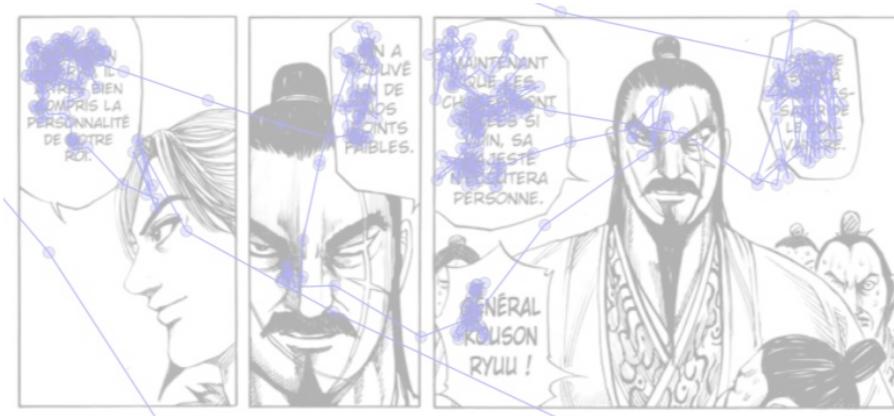


Raw eye gaze



Fixation-saccade detection
(Th1 = 50px, Th2 = 80px)

Overview of Proposed Method



↓ Fixation-saccade detection



Text region extraction



Graphics region extraction



Text region extraction

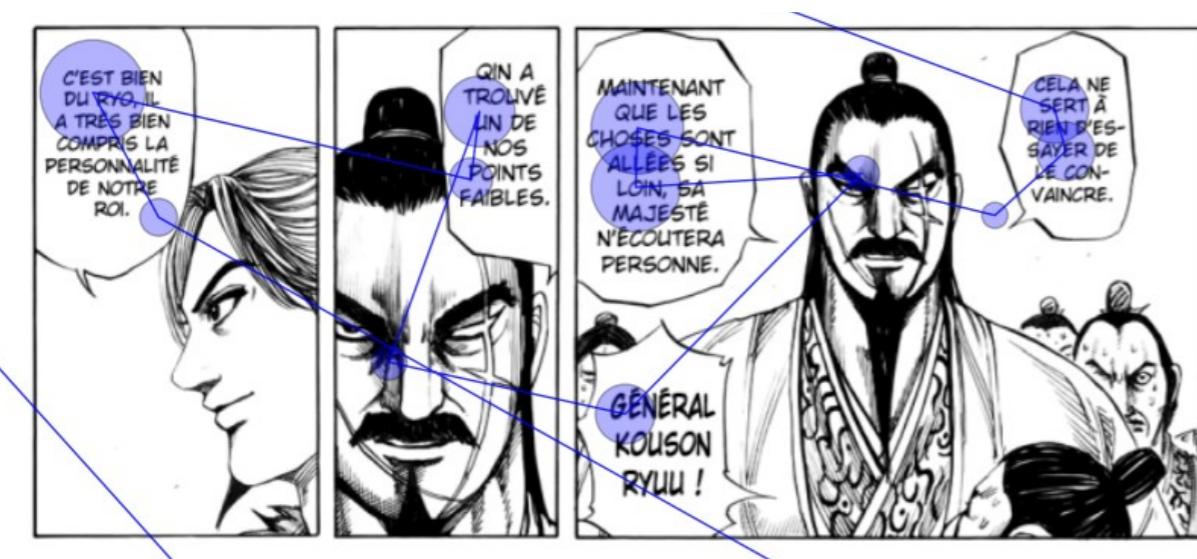
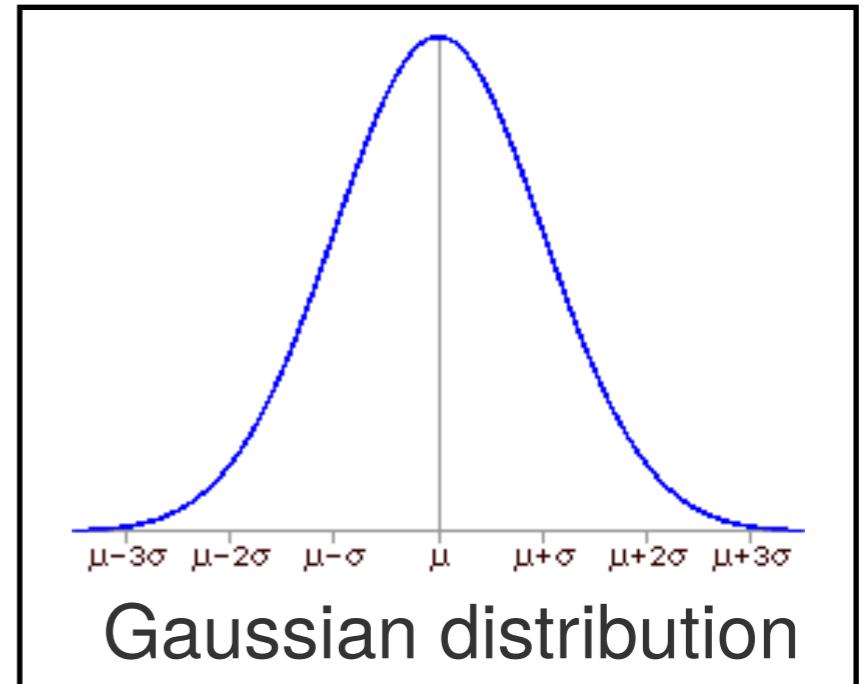
1. Calculating each reader's heat map

$$X(x, y) = d * N((C_x, C_y), \sigma^2)$$

d : fixation duration

$N(\mu, \sigma^2)$: Gaussian distribution

$C_x C_y$: fixation coordinate



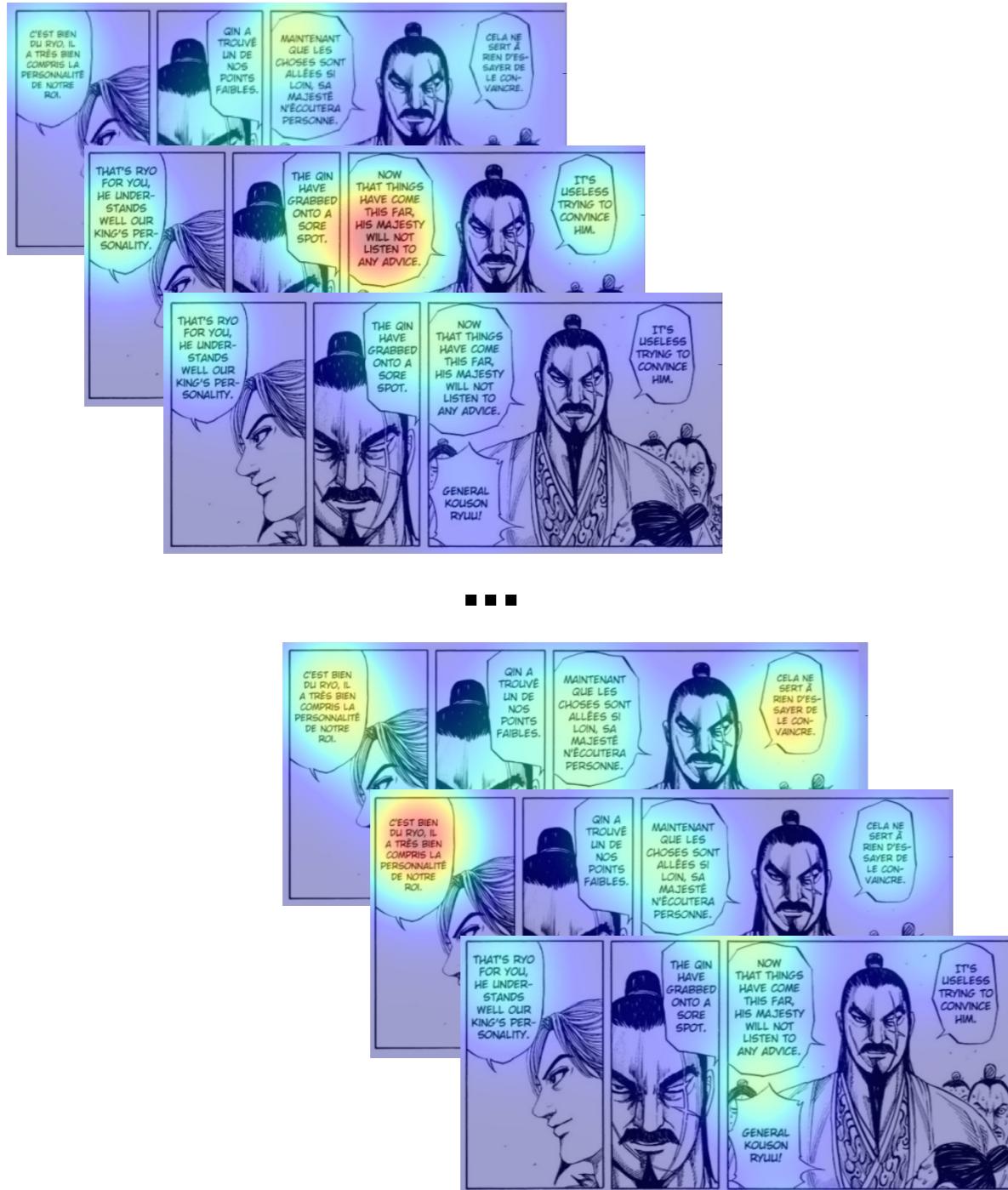
Fixations and saccades



Fixation duration heat map
($\sigma^2 = 50$)

Text region extraction

2. Overlaying multiple readers' data



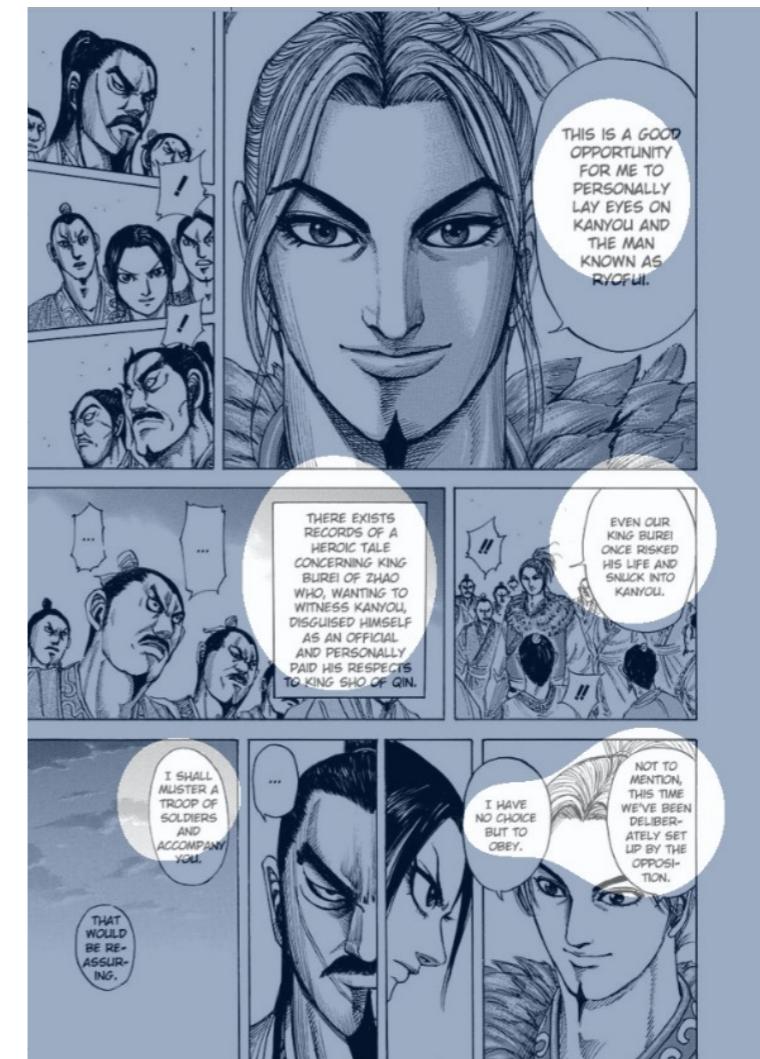
Average of multiple readers



Text region extraction

3. Calculating binary map from fixation heat map

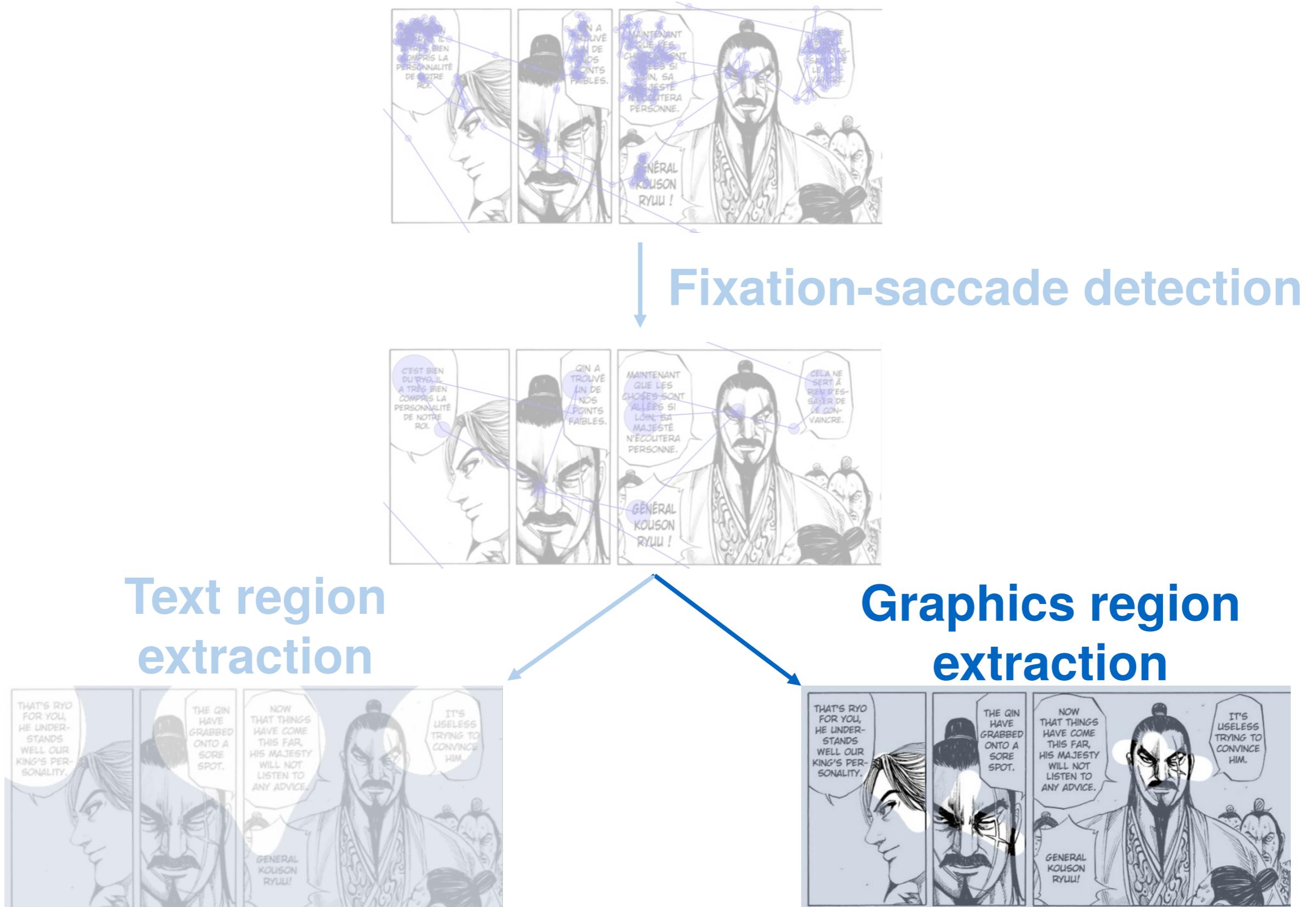
$$f(x, y) = \begin{cases} 1 & X(x, y) \geq \text{median}(X) \\ 0 & X(x, y) < \text{median}(X) \end{cases}$$



Fixation heat map

Text regions mask

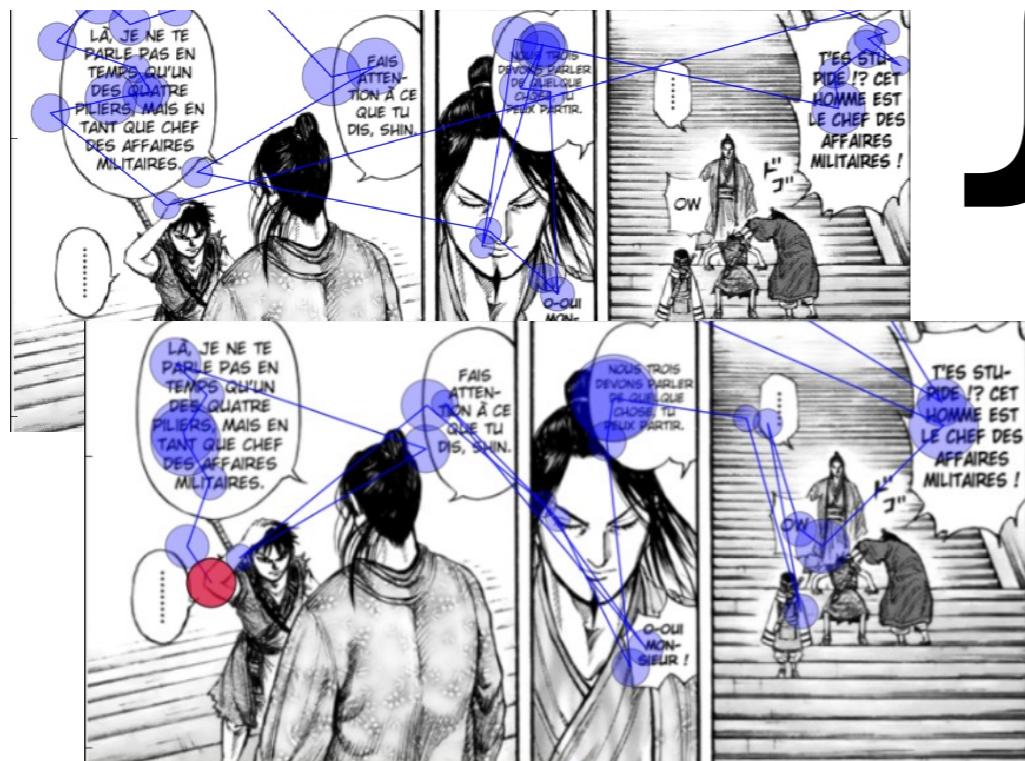
Overview of Proposed Method



Graphics region extraction



...

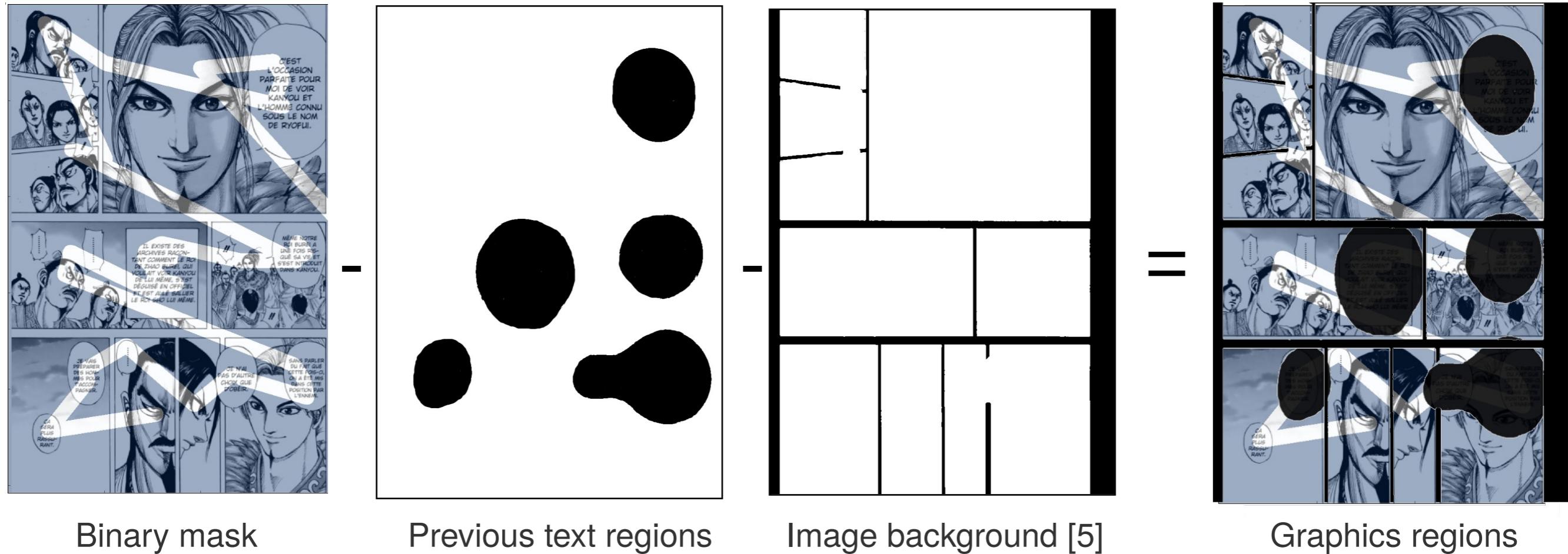


Average of multiple readers



Binary mask

Graphics region extraction



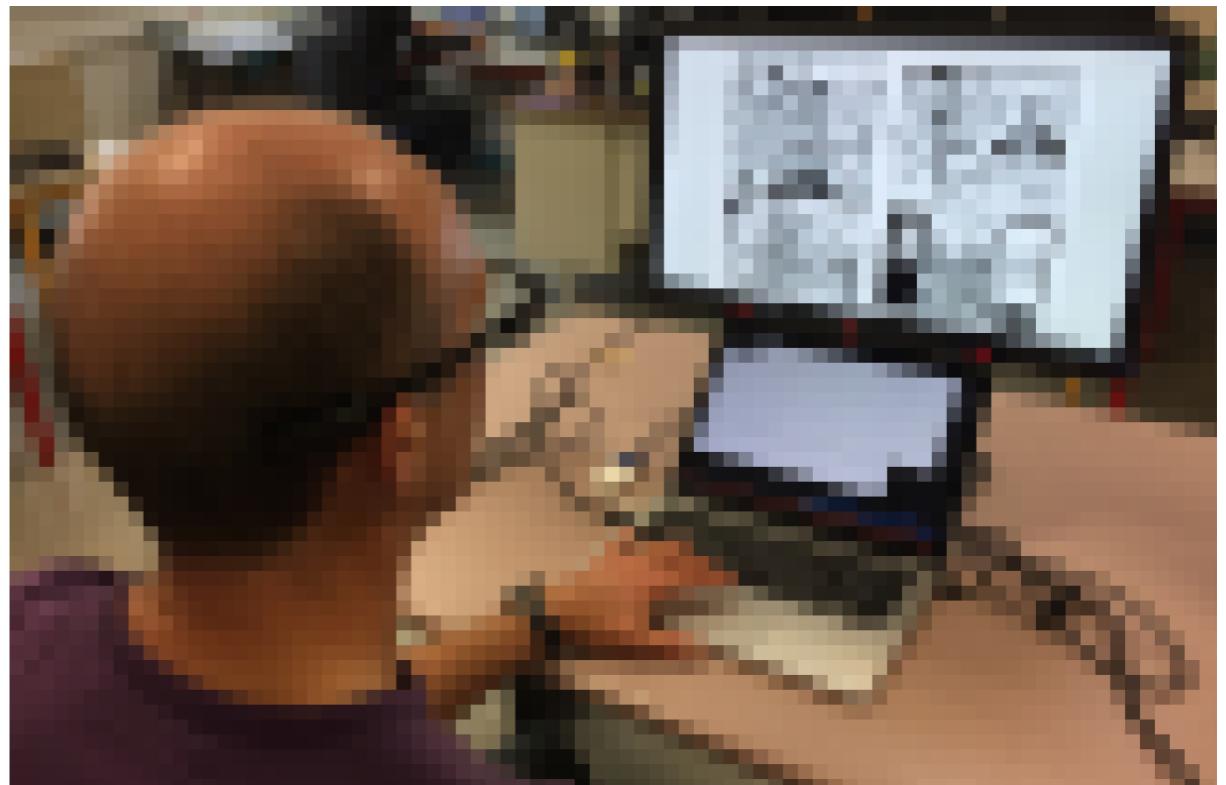
[5] C. Rigaud, C. Gurin, D. Karatzas, J.-C. Burie, and J.-M. Ogier, "Knowledge-driven understanding of images in comic books," International Journal on Document Analysis and Recognition (IJDAR), vol. 18, no. 3, pp. 199–221, 2015.
©Kingdom, Vol17, published by Shueisya, 2009.

Evaluation

Experimental Setup

Participants:

- 25 people (14-45 years old)
- 8 nationalities
- Eye tracker Tobii eyeX (99€)
- Sampling rate is 42Hz

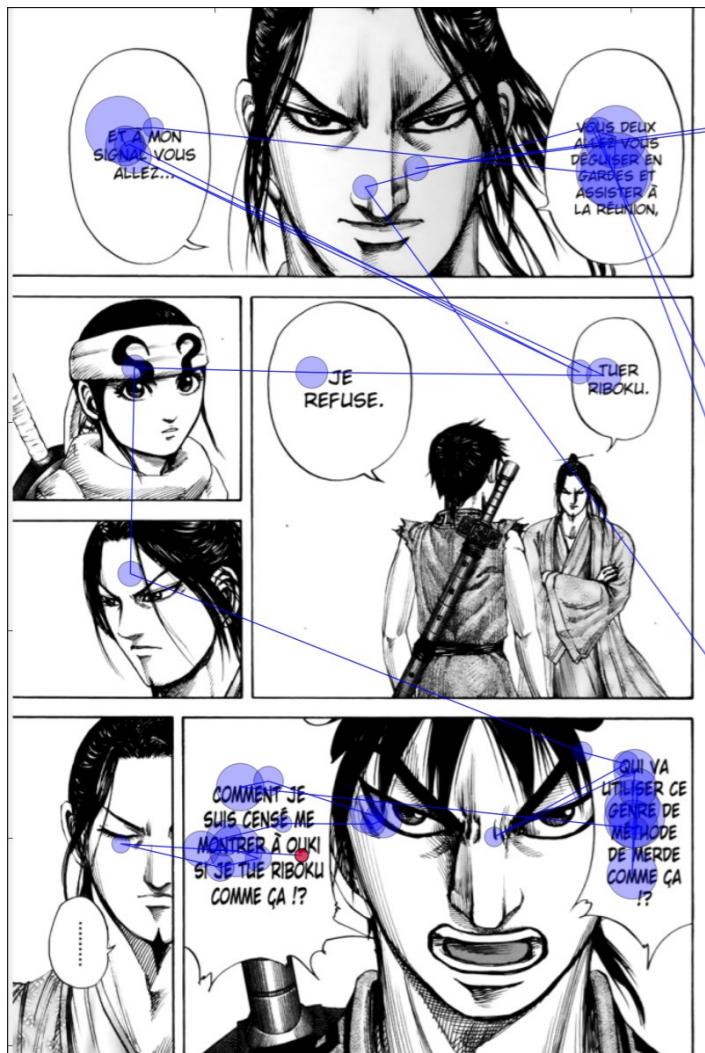


Documents:

- Kingdom vol. 17 chap. 175 (20 pages)
- 139 speech balloons
- 124 main comic character instances
- Two language versions: English and French

Text Extraction Evaluation

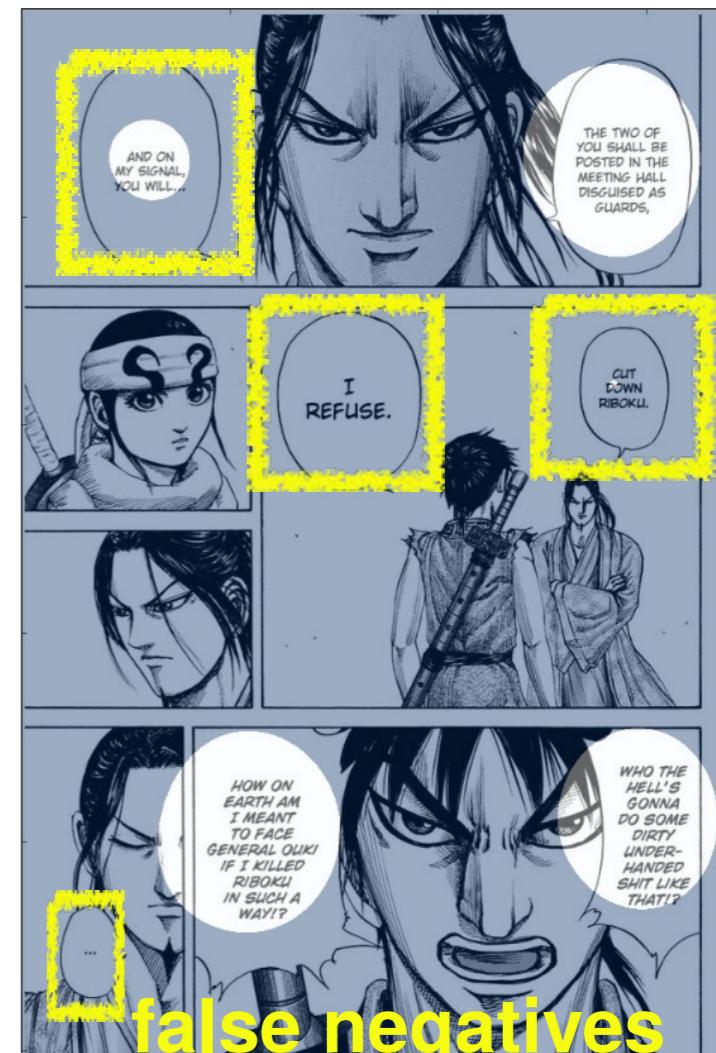
	<i>Recall</i>	<i>Precision</i>	<i>F-measure</i>
<i>Rigaud et al. (image processing based)</i>	74.10	60.95	66.88
<i>Proposed (eye gaze based)</i>	68.84	95.95	80.17



A single participant



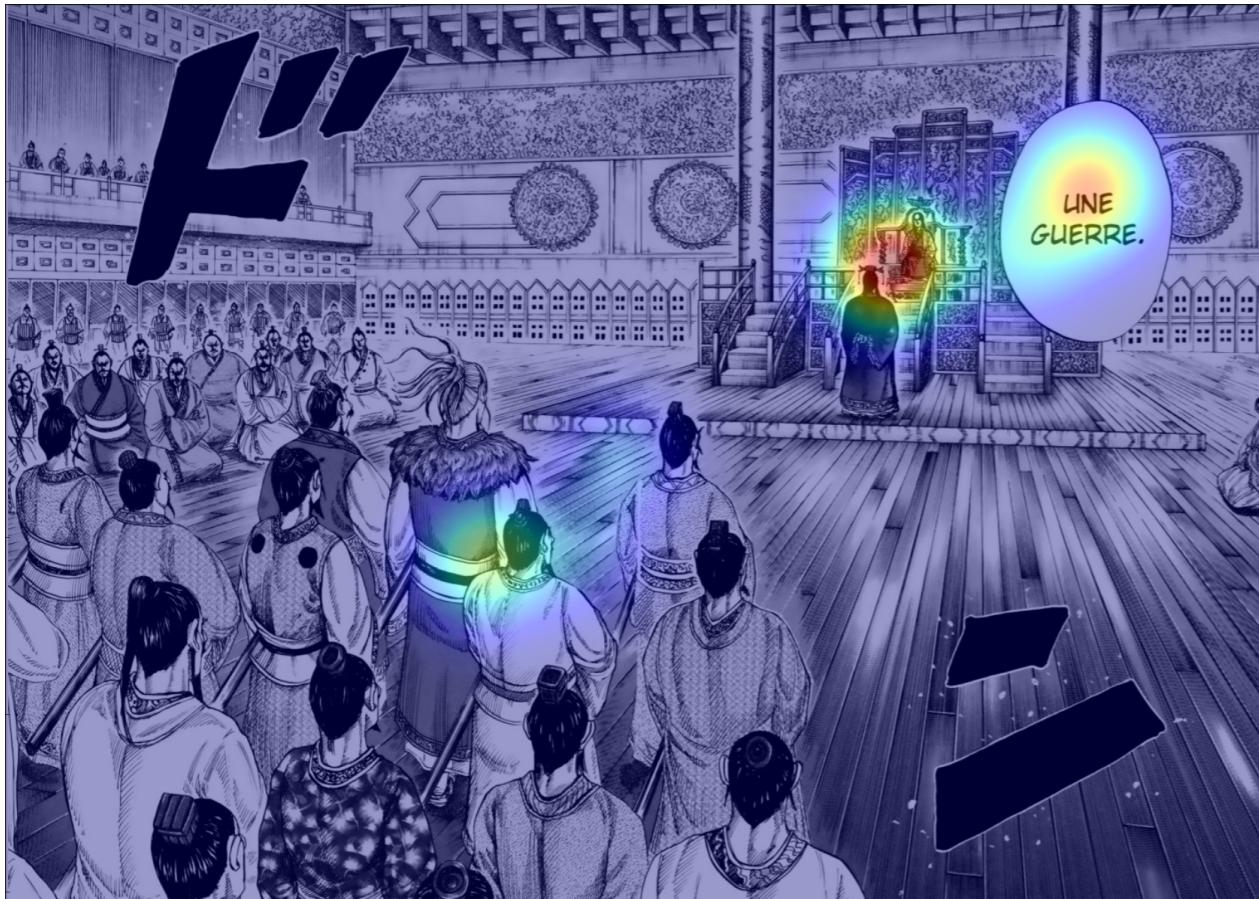
All participant fixations



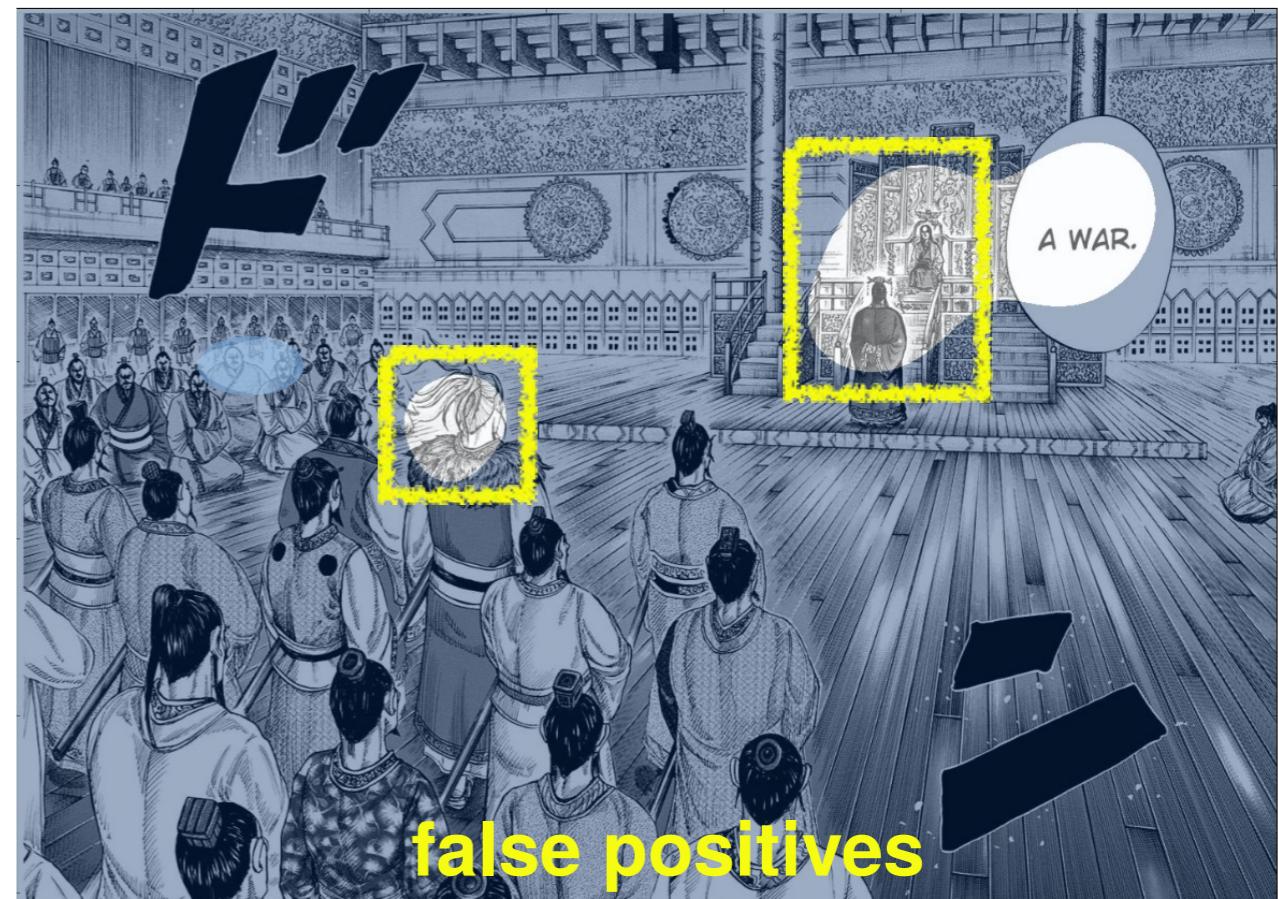
Corresponding mask

Text Extraction Evaluation

	Recall	Precision	F-measure
Rigaud et al. (image processing based)	74.10	60.95	66.88
Proposed (eye gaze based)	68.84	95.95	80.17



All participant fixations

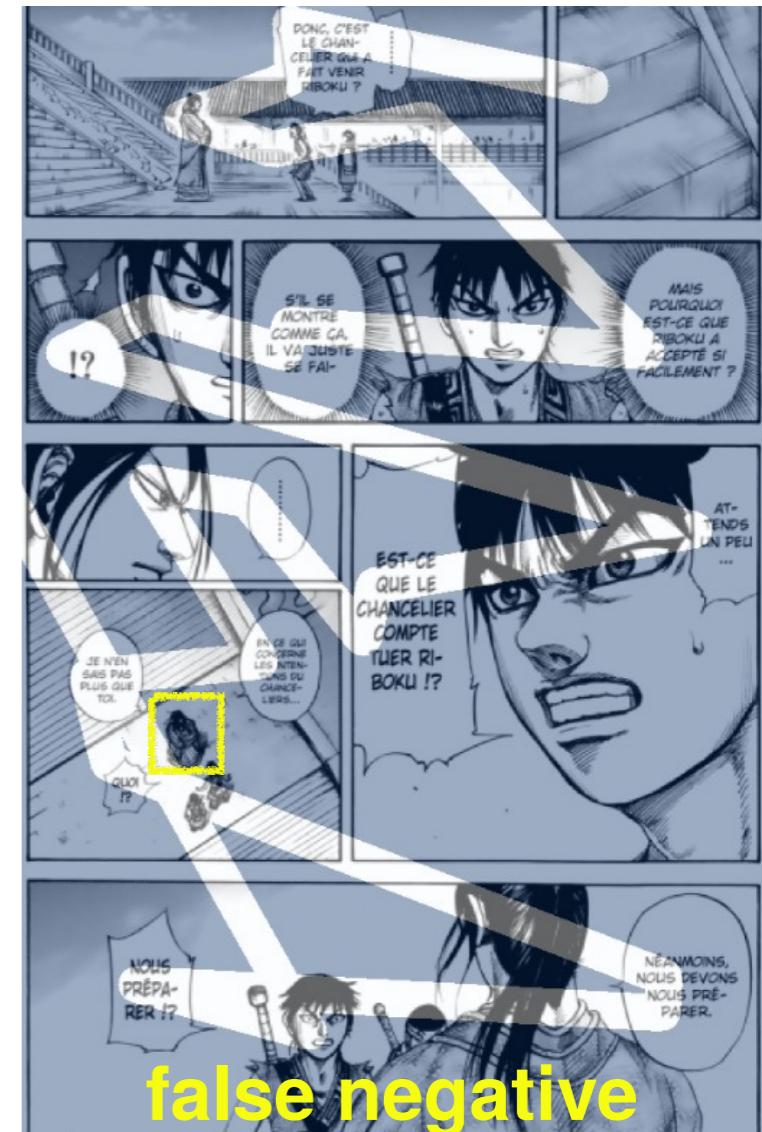


false positives

Corresponding mask

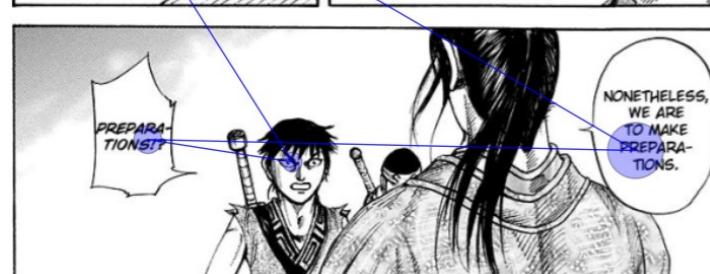
Graphics Extraction Evaluation

	<i>Recall</i>	<i>Precision</i>	<i>F-measure</i>
<i>Iwata (image processing based)</i>	54.90	72.72	62.57
<i>Proposed (eye gaze based)</i>	95.96	66.11	78.28



Graphics Extraction Evaluation

	Recall	Precision	F-measure
<i>Iwata (image processing based)</i>	54.90	72.72	62.57
<i>Proposed (eye gaze based)</i>	95.96	66.11	78.28



A single participant



All participant saccade paths



false positives

Corresponding mask

Conclusion

Conclusion

Conclusion

- Semi-automatic extraction methods for text and graphics extraction
- Potential of eye gaze movement for manga analysis.

Future work

- Use short fixations to split saccades and refine graphics ROI
- Link text and graphic regions
- Capture first user's eye movement during editorial process

Université de La Rochelle



More info:

<https://github.com/crigaud/publication/tree/master/2016/DAS/>

www.christophe-rigaud.com