



# Text-independent speech balloon segmentation for comics and manga



# Summary

- Context
- Speech balloons
- Proposed segmentation
- Dataset
- Evaluation
- Conclusion



<http://www.tumblr.com>

# Context

## Project eBDtheque

- Presentation
  - Born in L3i (2011)
  - Collaboration with French startup (PIA iiBD) and Japanese lab (PHC-SAKURA)
  - <http://l3i.univ-larochelle.fr/eBDtheque>
- Comic books
  - Cultural heritage
  - Need to be valorized by the new technologies
- Objective: comic book image understanding
  - Augmented reading experience
  - Information retrieval (e.g. semantic query, full text search)



# Speech balloons

## Previous work

- Arai [5]: blob size, vertical white lines, width to length ratio
- Ho [6]: bright blob, text region > 60%
- Rigaud [7]: inflate an active contour around text regions

## Challenges



[5] K. Arai and H. Tolle, "Method for real time text extraction of digital manga comic," IJIP, vol. 4, no. 6, pp. 669–676, 2011

[6] A. K. N. Ho, J.-C. Burie, and J.-M. Ogier, "Panel and Speech Balloon Extraction from Comic Books," DAS, pp. 424–428, 2012

[7] C. Rigaud et al.. "An active contour model for speech balloon detection in comics" ICDAR, pp. 1240-1244, 2013

# Proposed segmentation 1/3

Adaptive thresholding and connected-component (CC) extraction

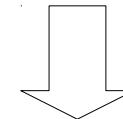


# Proposed segmentation 2/3

CC alignment



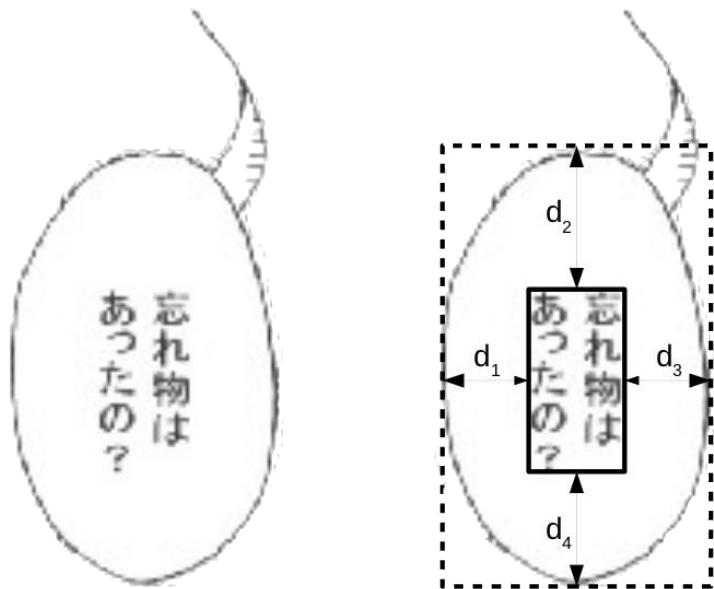
- (a) 0/18 aligned, start (b) 7/18 aligned, 1 line, continue (c) 14/18 aligned, 2 lines, continue (d) 16/18 aligned, 3 lines, stop



Alignment confidence  
 $cAlign = 16/18 = 88.9\%$

# Proposed segmentation 3/3

CC centering



$$hCenter = 1 - \frac{|d_1 - d_3|}{d_1 + d_3}$$

$$vCenter = 1 - \frac{|d_2 - d_4|}{d_2 + d_4}$$

$$\text{Global confidence } C = \frac{1}{2} * cAlign + \frac{1}{4} * hCenter + \frac{1}{4} * vCenter$$

# Dataset

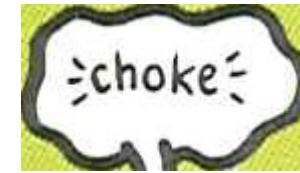
<http://ebdtheque.univ-lr.fr>

- eBDtheque dataset
  - 1092 balloons (84.5% are closed)
  - French album (46%), webcomics (37%)
  - American comics (11%)
  - Japanese manga (6%)

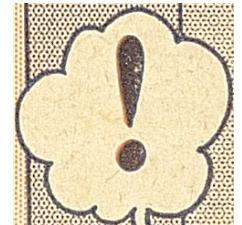


GRRR!

RIQUEUR MILITAIRE  
LOGIQUE IMPARABLE  
RÉFLEXION CLAIRVOYANTE



FUYONS!



STAY! LOOK AHEAD!  
A STAG IN THE CLEARING!  
'TIS SENT US BY DAME  
FORTUNE HERSELF!  
QUIET!

GRRR ! ANDRAGOR  
LE PUISSANT  
GUERRIER DOIT SE  
CONTROLLER S'IL NE  
VEUT PAS FAIRE  
UN CARNAGE !!

PLANQUONS-  
NOUS !!!



# Evaluation

- Pixel-level recall (R), precision (P), F-measure ( $F_1$ )

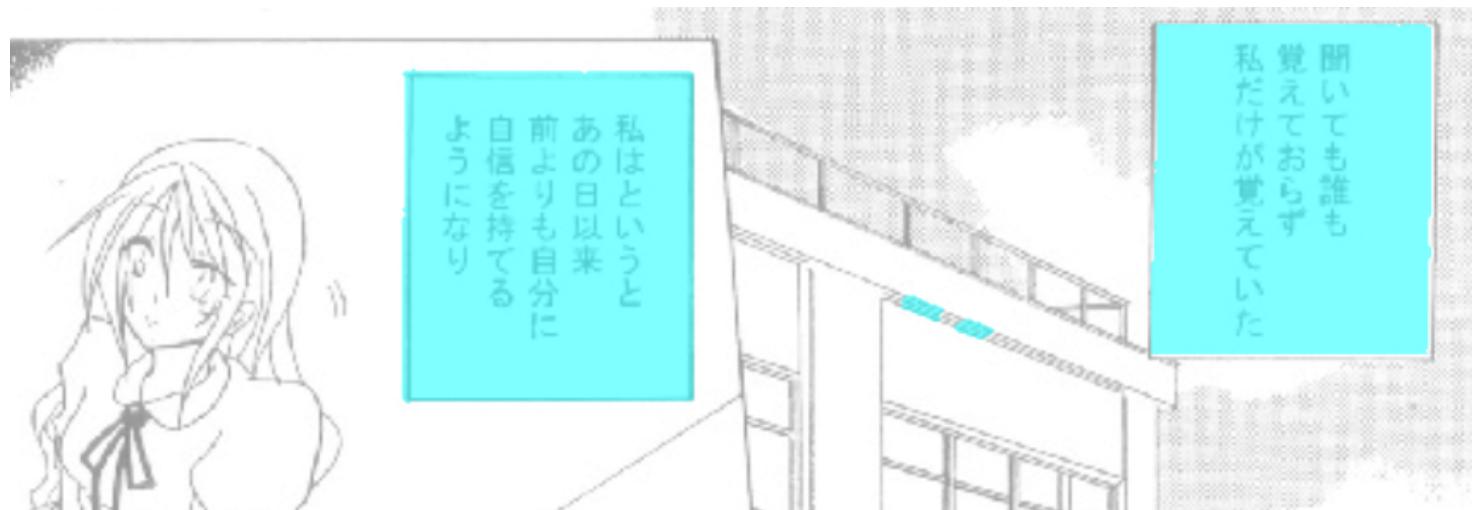
Method	R	P	$F_1$
Arai [5]	18.70	23.14	20.69
Ho [6]	14.78	32.37	20.30
Rigaud [7]	<b>69.81</b>	32.83	44.66
Proposed	62.92	<b>62.27</b>	<b>63.59</b>

[5] K. Arai and H. Tolle, "Method for real time text extraction of digital manga comic," IJIP, vol. 4, no. 6, pp. 669–676, 2011

[6] A. K. N. Ho, J.-C. Burie, and J.-M. Ogier, "Panel and Speech Balloon Extraction from Comic Books," DAS, pp. 424–428, 2012

[7] C. Rigaud et al.. "An active contour model for speech balloon detection in comics" ICDAR, pp. 1240-1244, 2013

# Sample results



# Conclusion

- Speech balloon segmentation for text/graphics association
- Topological and spatial position relationship
- Simple and generic approach
- Promising performance
- Next:
  - Add more scanned and digital-born manga
  - Handle open balloons by analysing image background content

**THANK YOU VERY MUCH**

*Image credits: <http://ebdtheque.univ-lr.fr/database/#Acknowledgment>*