13th IAPR International Conference on Document Analysis and Recognition - ICDAR 2015 – Nancy, France

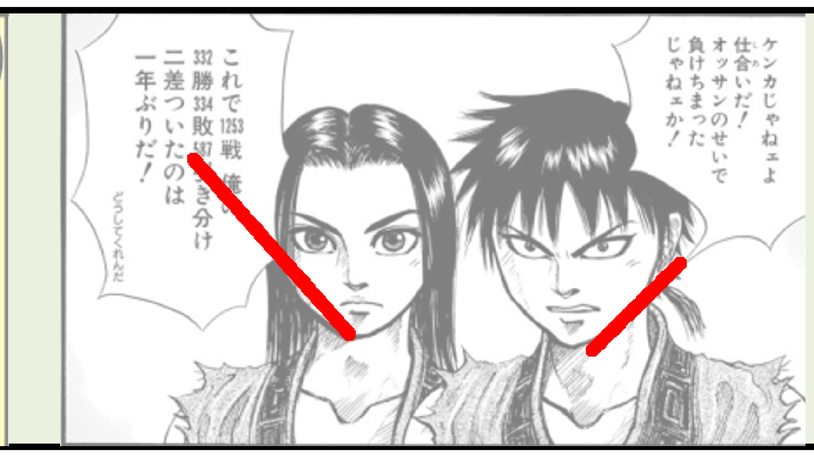
Speech balloon and speaker association for comics and manga understanding

Christophe Rigaud[1], Nam Le Thanh[1], J.-C. Burie[1], J.-M. Ogier[1], Motoi Iwata[2], Eiki Imazu[2], Koichi Kise[2]
[1]Laboratory L3i, University of La Rochelle, Avenue Michel Crépeau 17042 La Rochelle, France
[2]Graduate School of Engineering, Osaka Prefecture University, 1-1 Gakuen-cho, Nakaku, Sakai, Osaka 599-8531, Japan
{christophe.rigaud, tle05 jcburie, jmogier }@univ-lr.fr, {iwata, imazu_e, kise}@cs.osakafu-u.ac.jp

Presentation

Context
- Important cultural heritage
- Scanned and digital-born comics and manga
- Unsupervised approach

Applications
- Character's interaction and situation analysis
- Text-to-speech with adaptive voice and tone

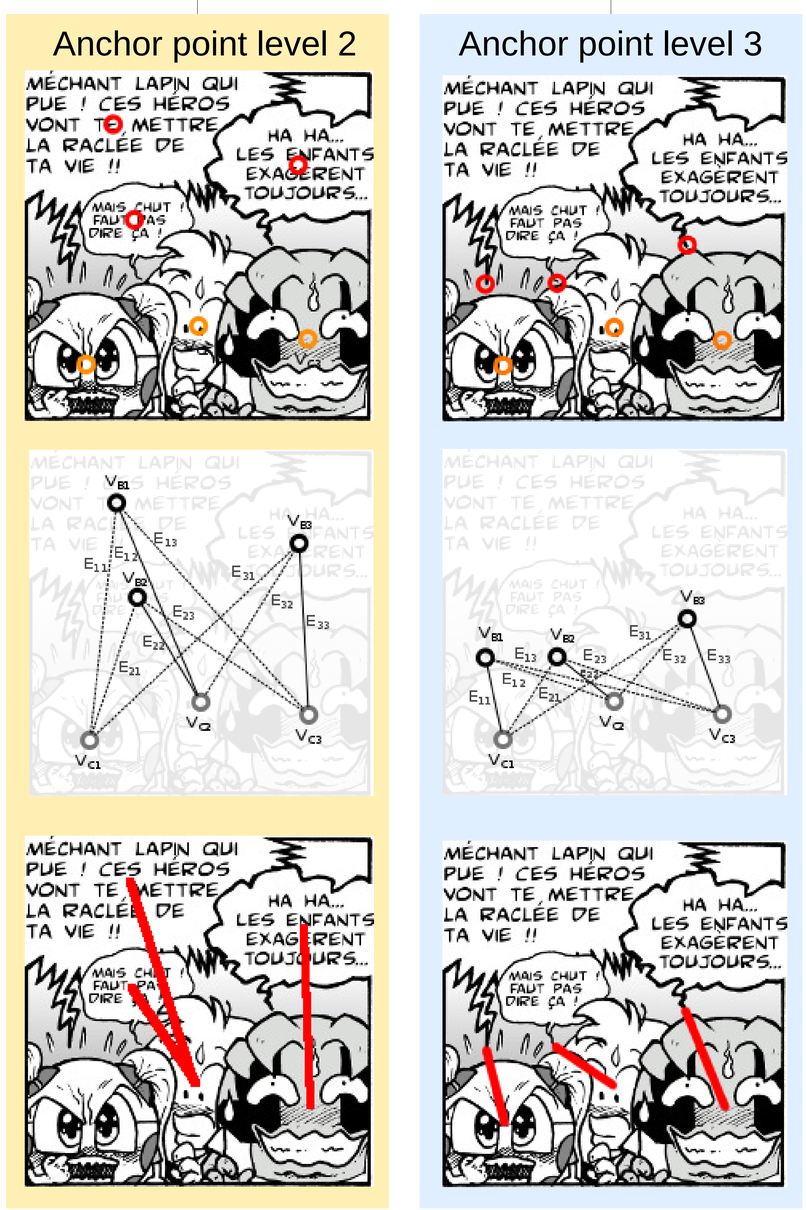


Contributions

Association method

- Prerequisites: anchor points of balloons and characters
- Geometric graph from all anchor point combination
- Optimal 2-tuples considering Euclidean distance (*L**)

$$L^*(V_{Bi}, V_{Cj}) = \text{ArgMin}(E_{ij})$$



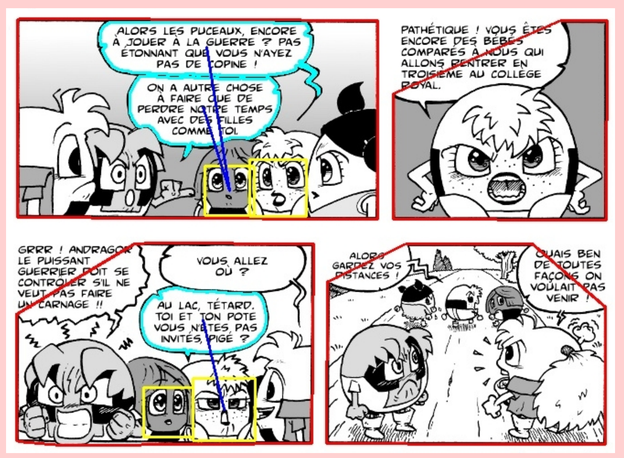
Anchor points

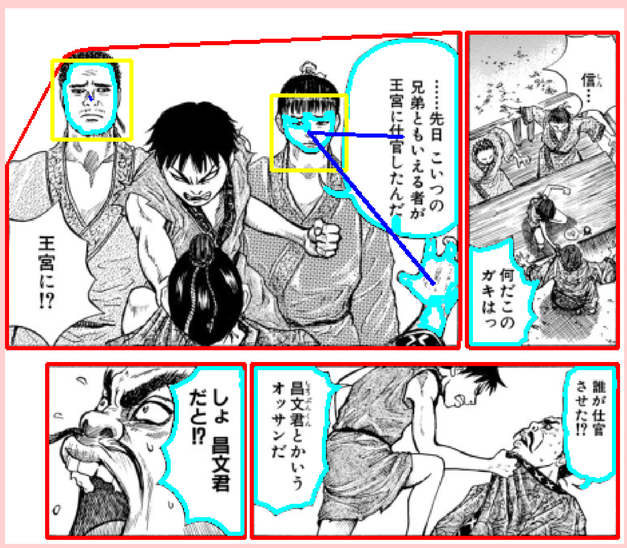
- Related to previously extracted elements (e.g. balloons, characters)
- From coarse (level 1) to fine (level 4) and more

Level	Speech balloon V_B	Comic character V_C
1	Bounding box center	Bounding box center
2	Balloon centroid	Character centroid
3	Tail tip position	Face center [1]
4	Tail position and direction	Mouth center

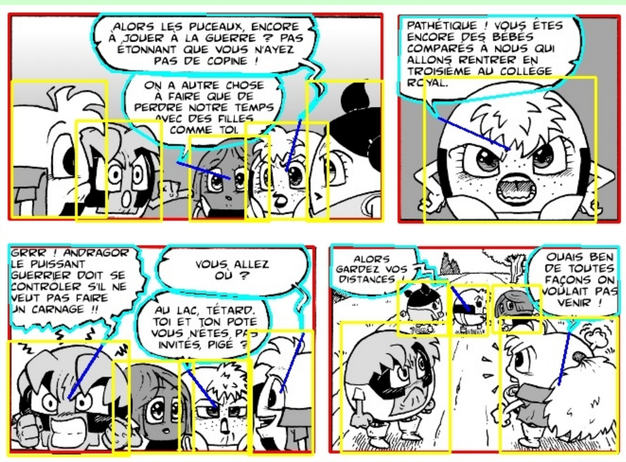
Evaluation

Prerequisites extracted automatically





Prerequisites loaded from GT



Anchors	eBDtheque[2]	Kingdom
Level 2	78.58%	76.35%
Level 3	93.32%	87.74%

Anchors	eBDtheque	Kingdom
Level 2	4.33%	18.60%
Level 3	18.01%	19.41%

References

[1] W. Sun and K. Kise, “Similar partial copy recognition for line drawings using concentric multi-region histograms of oriented gradients,” in Proceedings of the IAPR Conference on Machine Vision Applications, ser. MVA2011, Nara Japan, 2011.
[2] C. Guérin, C. Rigaud, A. Mercier, and al., “ebdtheque: a representative database of comics,” in Proceedings of International Conference on Document Analysis and Recognition (ICDAR), Washington DC, 2013.

Supported by the Higher Education and Research Ministry and the French foreign affair office.

Conclusion

Contribution
- One-to-one speech balloon/comic character link
- Robust against missing panels
- Adaptive to different level of information (tail)

Perspectives
- Consider tail direction
- Handle out-of-panel speakers and multi-speakers

Source: https://github.com/crigaud/publication/tree/master/2015/ICDAR/speech_balloon_and_speaker_association_for_comics_and_manga_understanding