1st International Workshop on coMics ANalysis, Processing and Understanding – MANPU 2016 – Cancun, Mexico

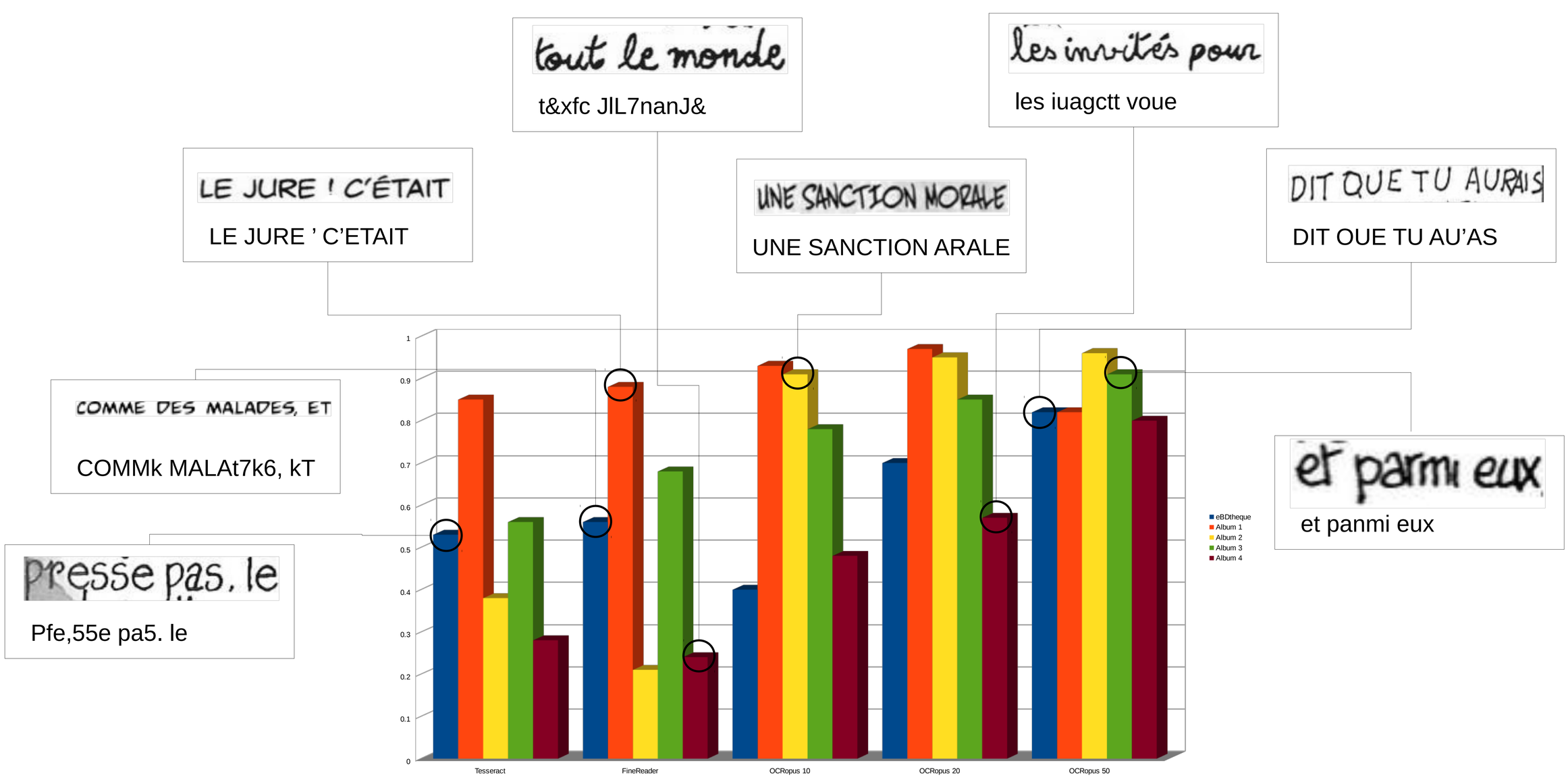
Toward speech text recognition for comic books

Christophe Rigaud, Srikanta Pal, Jean-Christophe Burie, Jean-Marc Ogier
Laboratory L3i, University of La Rochelle, Avenue Michel Crépeau 17042 La Rochelle, France
{christophe.rigaud, srikanta.pal, jcburie, jmogier}@univ-lr.fr

Challenges

- Mix of typewritten, handwritten and free form text
- Uniform or complex background
- Text deformation and overlapping
- Multiple writing styles
- Non-dictionary words and hyphenation

Contributions



Sting similarity comparison for Tesseract, ABBY FineReader and OCRopus with 3 different amount of training data

Normalized Levenshtein distance

s_1 = ça devient vite

s_2 = eu deient vute

$L = len(s_1) + len(s_2) = 15 + 14 = 29$

$dist(s_1, s_2) = 4$

string similarity $= \frac{(L - dist(s_1, s_2))}{L} = 0.86$

Impact of the amount of training data for OCRopus

	Percentage of training data		
	10%	20%	50%
eBDtheque	0.40	0.70	0.82
Album 1	0.93	0.97	0.99
Album 2	0.91	0.95	0.96
Album 3	0.78	0.85	0.91
Album 4	0.48	0.57	0.80

Datasets

eBDtheque (scientific dataset)
- ~20 albums, 3537 text lines in French (subset)
- http://ebdtheque.univ-lr.fr/

Sequencity (online comics library)
- 4 albums, 2000 text lines in French (subset)
- http://sequencity.com/

Conclusion

Contribution
- Comparison of 5 different OCR/scenario
- Minimum amount of required training data

Perspectives
- Train OCRopus on other complex writing styles
- Build a generic comics-like text recognizer

Online: https://github.com/crigaud/publication/tree/master/2016/MANPU/toward_speech_text_recognition_for_comic_books/