European Ph.D. defense

Segmentation and indexation of complex objects in comic book images

Christophe Rigaud
December 11th, 2014

Jointly supervised by:

Jean-Christophe Burie[1]
Dimosthenis Karatzas[2]
Jean-Marc Ogier[1]

[1]L3i – Université de La Rochelle
[2]CVC - Universitat Autònoma de Barcelona

Outlines

- Introduction
- Document image analysis
- State of the art of comics analysis
- Contributions
- Experimentations
- Conclusions

Outlines

- Introduction
  - Comic books
  - History of comics art
  - Market place
  - Comics project
  - Objectives of the thesis
- Document image analysis
- State of the art of comics analysis
- Contributions
- Experimentations
- Conclusions

Comic books

“juxtaposed pictorial and other images in deliberate sequence, intended to convey information and/or to produce an aesthetic response in the viewer” Mc Cloud, 1993

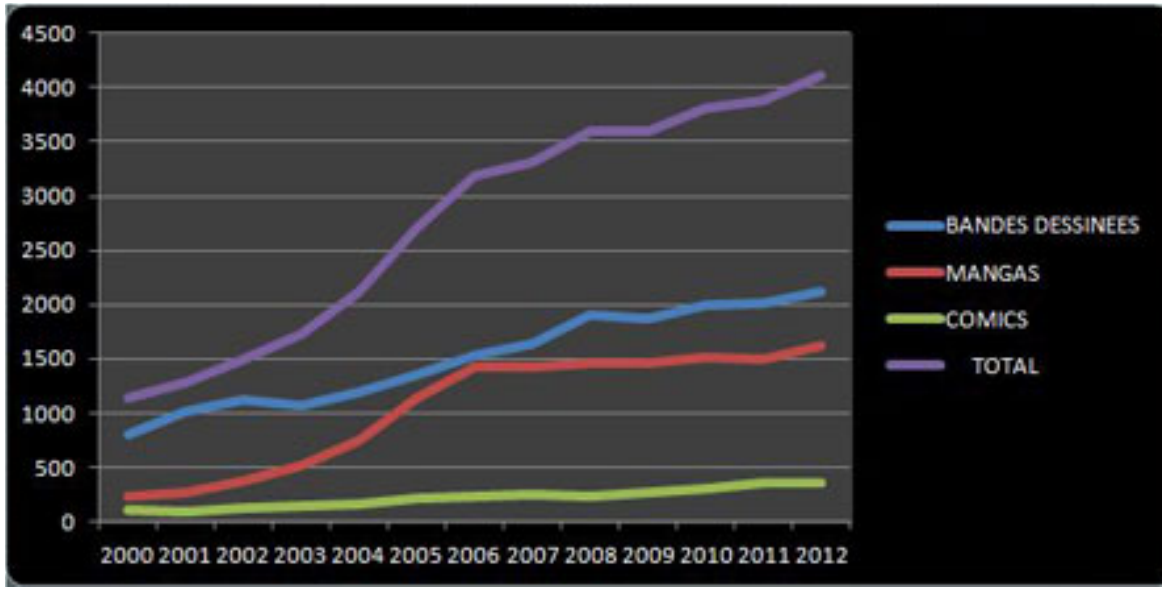
“*a visual medium used to express ideas via images, often combined with text or visual information*” Wikipédia, 2014

- One of the most popular and familiar forms of graphic content
-

History of comics art

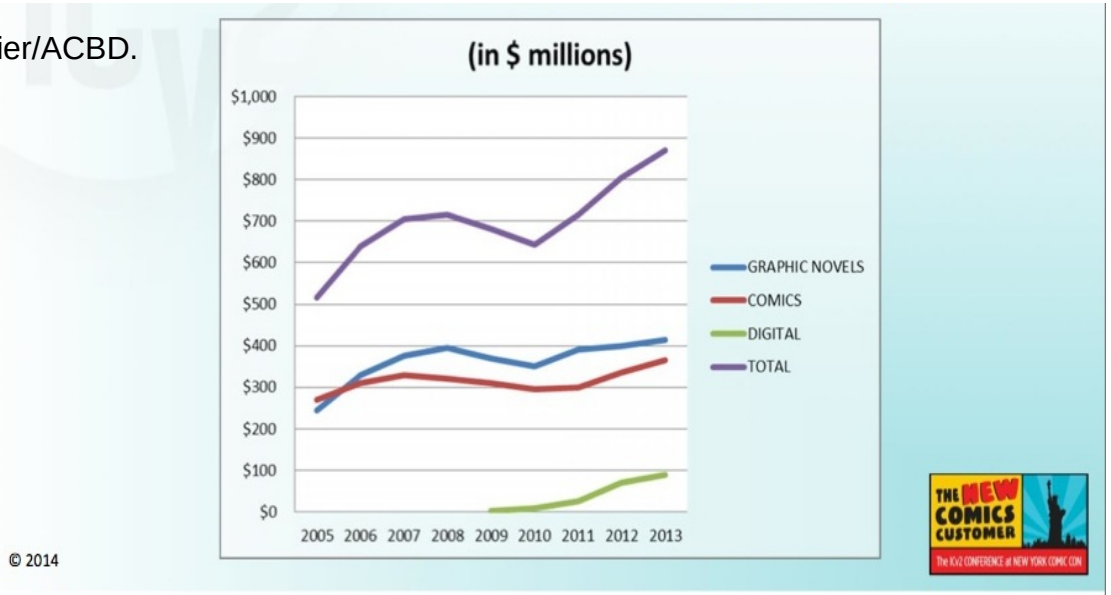
- Prehistory: painting of animals and hunters in caves
- 1846: Rodolphe Töpffer, the inventor of the “bandes dessinées”
- 1930s: magazine-style comic books production in the US
- 1950s: massive production of manga in Japan (Osamu Tezuka)
- 1971: the term of ninth art is attributed to comics art (Francis Lacassin)
- 1996: explosion of the Internet bubble and webcomics
- 2007: adaptation ofto social media sites and mobile devices

Market place



Francophone comics production

Infographie (c) L'Agence BD d'après les chiffres de Gilles Ratier/ACBD.



Comics market in the US

Milton Griepp's White Paper, ICv2 Conference 2014

Comics project (eBDtheque)

- What?
  - Add value to paper-based comics using the new technologies
- Why?
  - Answer to a real need from librarians, advertisers and readers
  - Allows text/image search, reflowable documents, augmented reading and translation assistance
- How?
  - Extracting content of digitalized comic books (e.g. panels, balloons, text, comic characters)
  - Retrieving the semantic of the elements (e.g. read before, said by, thought by, addressed to)

- Who?
  - Supported by L3i lab
  - 2 Ph.D. students, 6 professors, 1 engineer (one year) and 1 post doc (one year)
  - Public founding
    - CPER 2007-2013 (State-Region Project Contract)
    - PHC-Sakura 2014-2015 (France/Japan Bilateral Joint Research Project)
    - PIA-iiBD 2015-2017 (future investment project with French company/labs)

Objectives of this thesis

- Propose generic methods for content extraction of digitalized comic books
- Indexation of content in order to be browsable and exchangeable???
- Duration 36 months
- Challenges:
  - Recent field of research with a largely unknown
  - The documents are semi-unstructured, free-form and with complex background

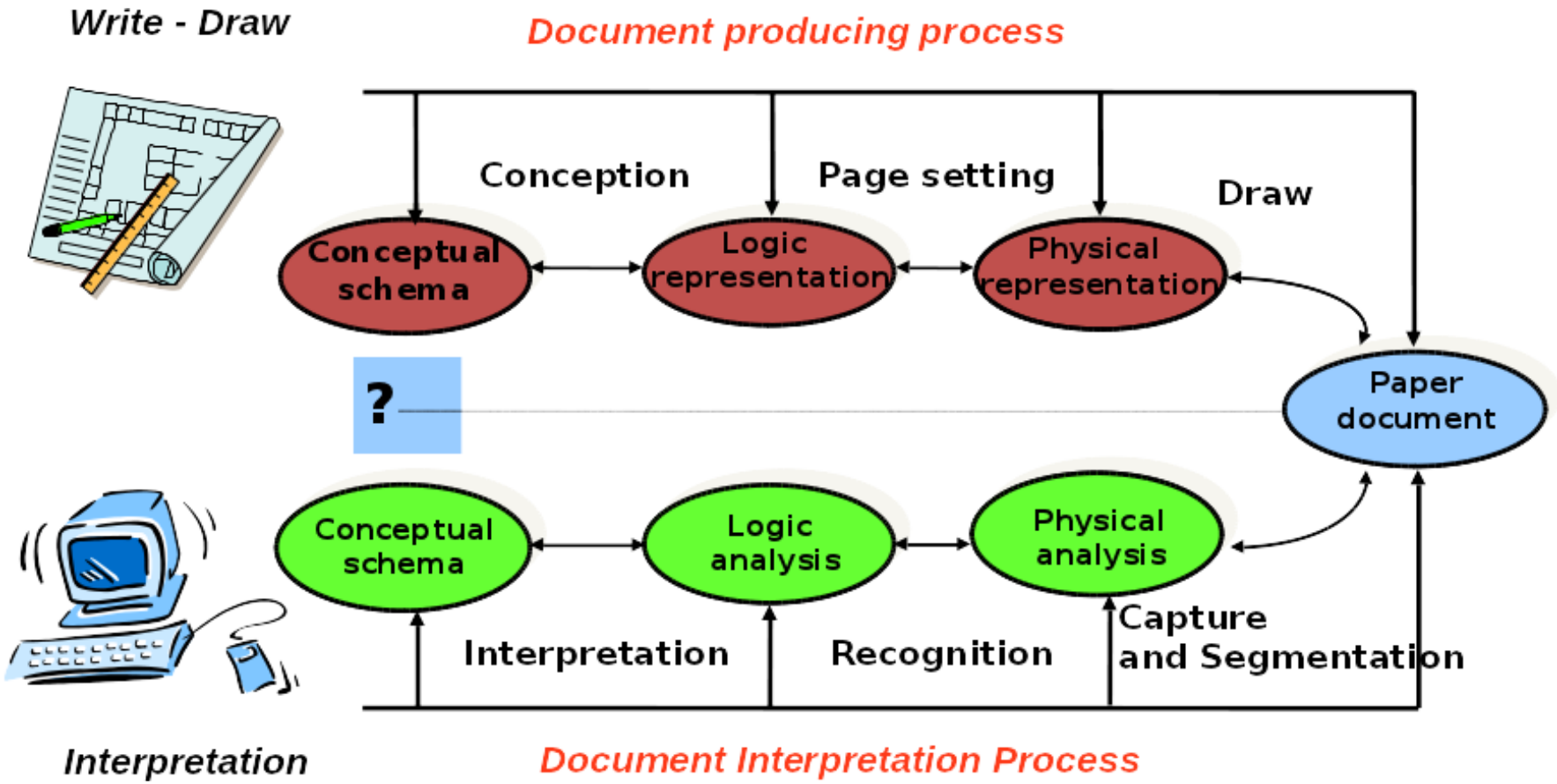
Outlines

- Introduction
- Document analysis
- State of the art of comics analysis
- Contributions
- Experimentations
- Conclusions

Outlines

- Introduction
- Document analysis
  - Conception to interpretation
  - Comic books production
  - Comic books interpretation
  - (Document type comparison)
- State of the art of comics analysis
- Contributions
- Experimentations
- Conclusions

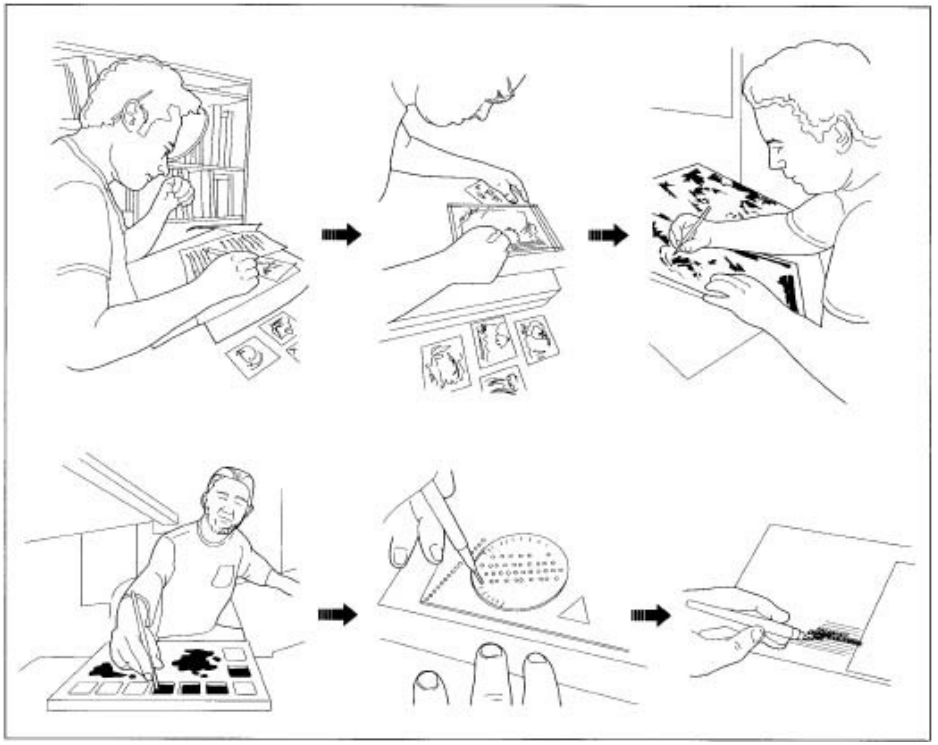
Conception to interpretation



Handbook of Document Image Processing and Recognition. Springer, 2014

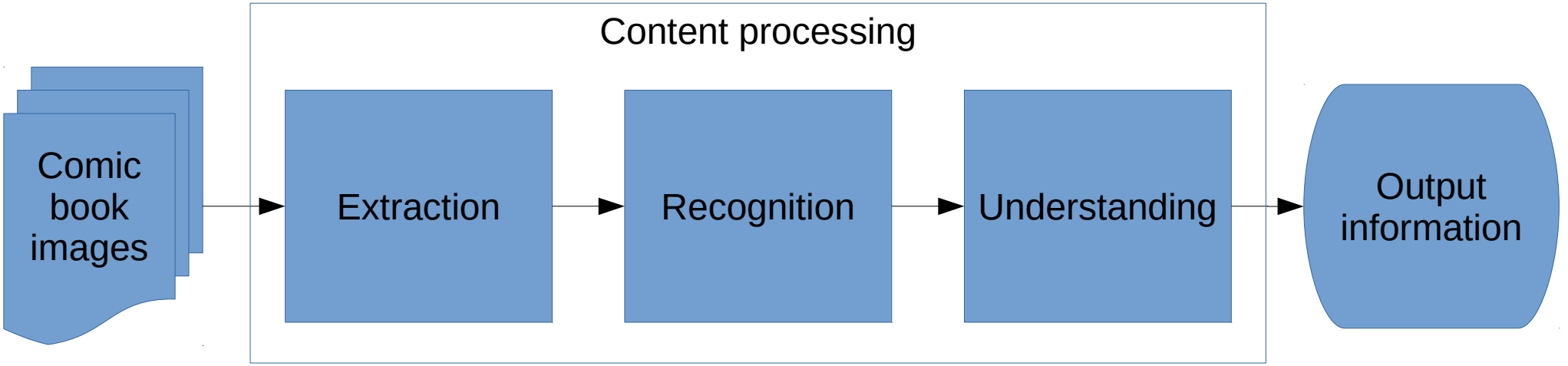
Comic books production

1) Synopsis and scenario
2) Pencil drawing
3) Inking
4) Flatting and colouring
5) Lettering and sound effects



http://www.madehow.com/Volume-6/Comic-Book.html

Comic books interpretation



Comic books interpretation

Extraction → Recognition → Understanding → Output information

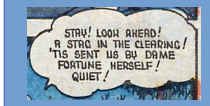
Comic books interpretation

Document analysis

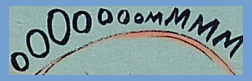
Comic book images

Panels

Balloons



Text



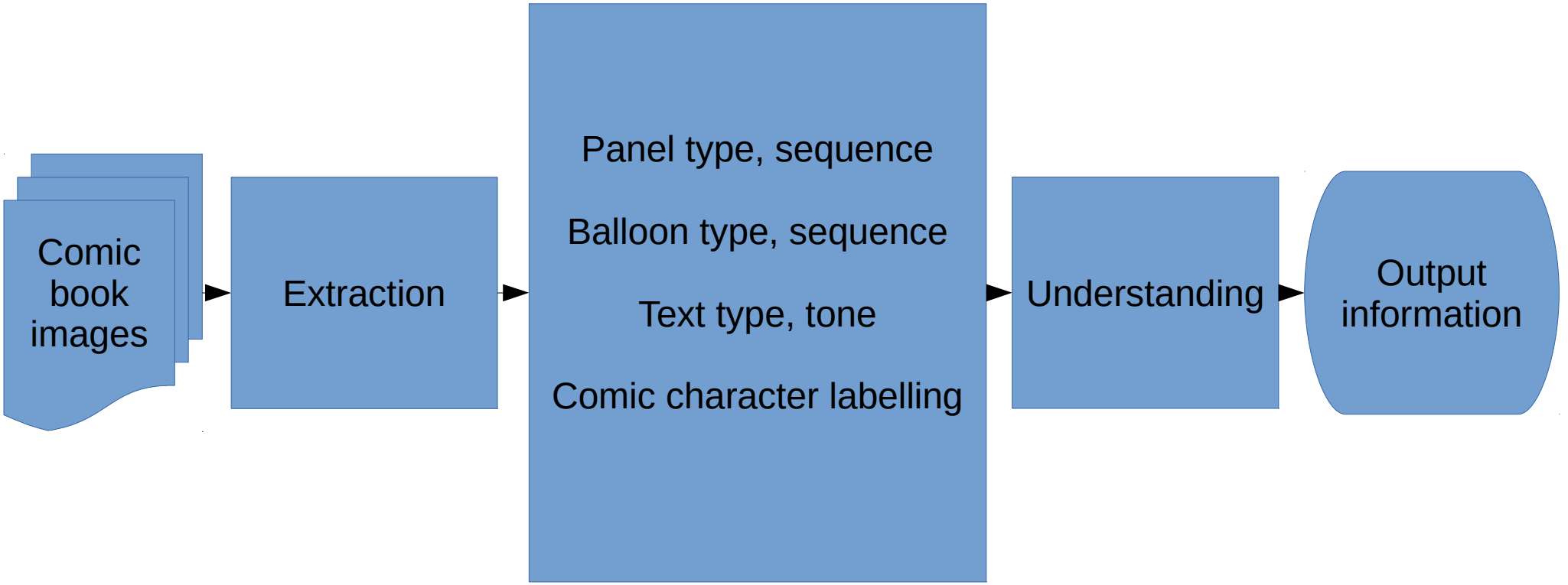
Comics character

Recognition

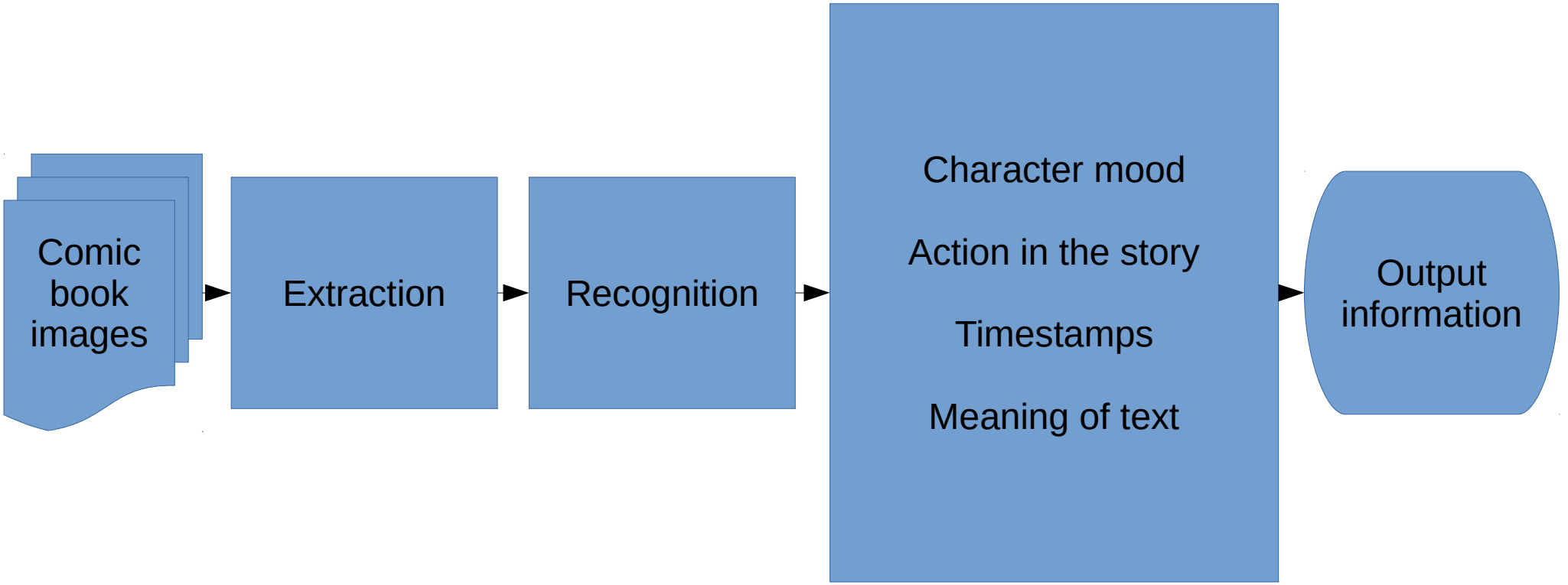
Understanding

Output information

Comic books interpretation

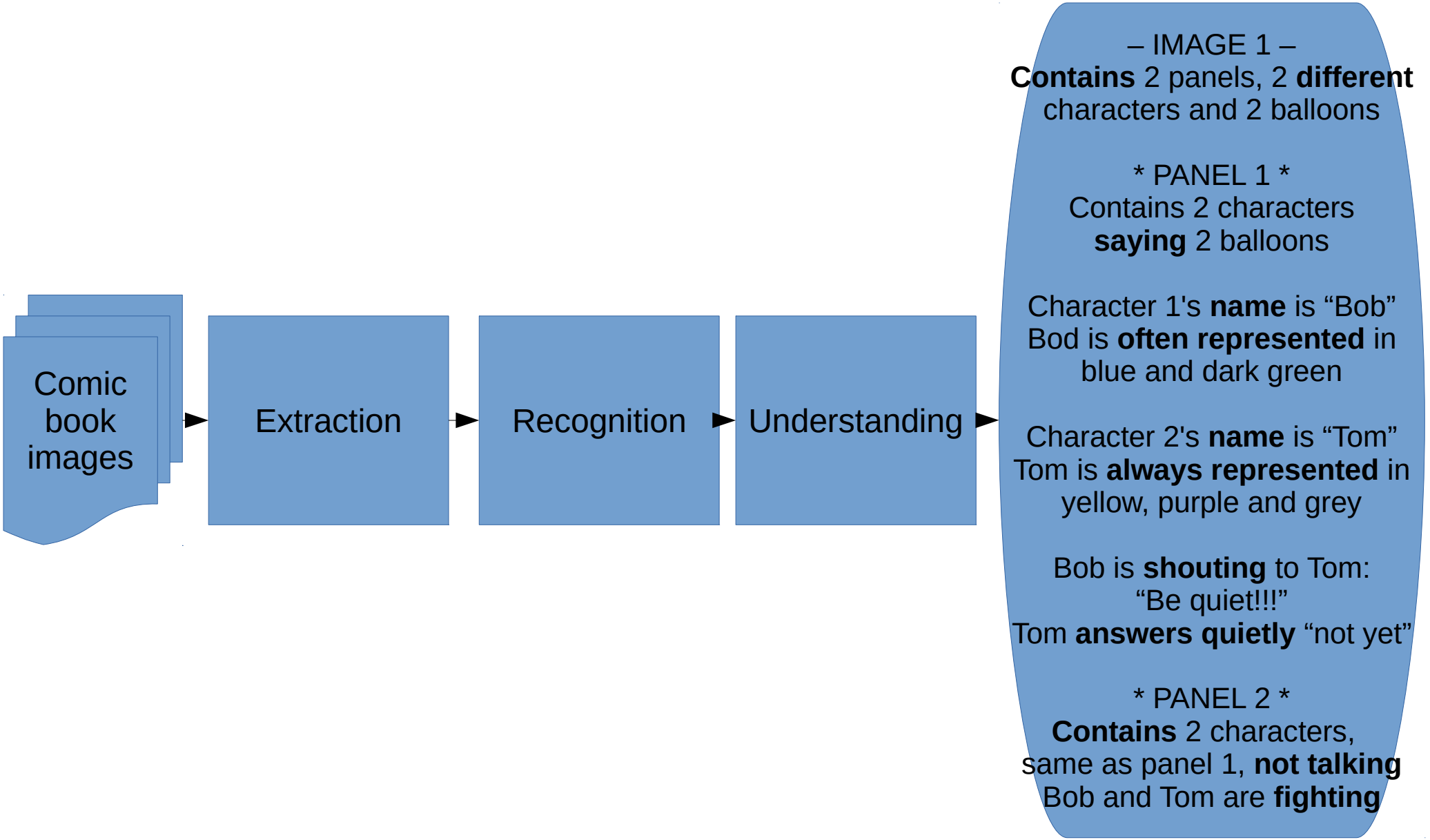


Comic books interpretation



Comic books interpretation

Document analysis



Outlines

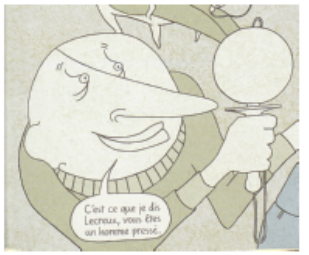
- Introduction
- Document analysis
- State of the art of comics analysis
- Contributions
- Experimentations
- Conclusions

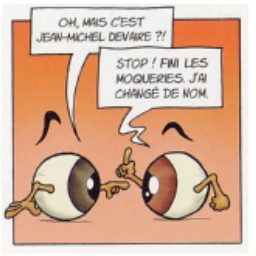
Outlines

- Introduction
- Document analysis
- State of the art of comics analysis
  - Panel extraction and layout analysis
  - Balloons analysis (extraction, classification, tail)
  - Text extraction and recognition
  - Comic character detection and recognition
  - Holistic understanding
- Contributions
- Experimentations
- Conclusions

Panel and layout analysis

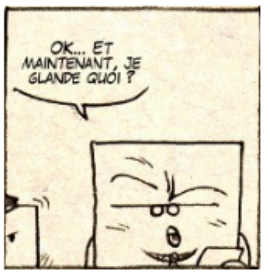
- Challenges
  - Diversity of comic books
  - Uncommon layout
- Panel extraction
  - White line cut
  - Recursive X-Y cut algorithm
  - Gradient
  - Connected-components
  - Polygon detection
  - Corners and line segments
- Layout
  - Reading order (Z-path)

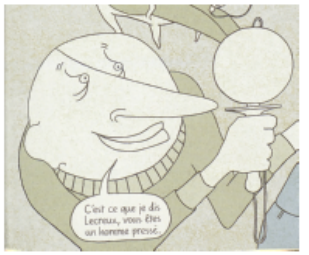


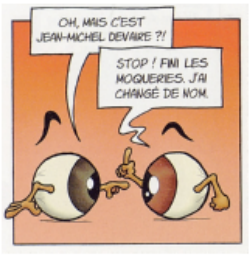


Panel and layout analysis

- Challenges
  - Diversity of comic books
  - Uncommon layout
- Panel extraction
  - White line cut
  - Recursive X-Y cut algorithm
  - Gradient
  - Connected-components
  - Polygon detection
  - Corners and line segments
- Layout
  - Reading order (Z-path)







- Conclusions
  - Problem solved for common manga and European comics if treated separately
  - Remaining difficulties are for connected, nested and implicit panels
  - No approach tested over all comics styles, no dataset, no reproductive results

Balloon analysis

- Challenges
  - Diversity of balloons intra/inter comics
  - Implicit balloons
  - Interface between text and graphics
- Extraction
  - Shape vs contour
  - Blob detection [Arai 2011, Ho 2012]
- Classification
  - Speech tone information (contour)
- Tail detection
  - Indicate the position of the emitter

Image	Shape	Contour
	Oval	Smooth
	Rectangle	Smooth
	Oval	Wavy
	Oval	Spiky
	Oval / implicit	Smooth / Implicit

Balloon analysis

- Challenges
  - Diversity of balloons intra/inter comics
  - Implicit balloons
  - Interface between text and graphics
- Extraction
  - Shape vs contour
  - Blob detection [Arai 2011, Ho 2012]
- Classification
  - Speech tone information (contour)
- Tail detection
  - Indicate the position of the emitter
- Conclusions
  - Closed balloon solved (sequential)
  - Implicit, classification and tail were not explored

Image	Shape	Contour
	Oval	Smooth
	Rectangle	Smooth
	Oval	Wavy
	Oval	Spiky
	Oval / implicit	Smooth / Implicit

Text extraction and recognition

State of the art

- Challenges
  - Non-standard fonts
  - Multi-script/orientation/scale
  - Complex background (sound effects)
  - Short length, hyphenation
  - Voluntary spelling mistakes
- Extraction
  - Scene text localization
  - Connected-components approach
  - SVM and Bayesian classifier
  - Sound effects have not been investigated yet
- Recognition
  - At is early stage
  - OCR trained for a specific comics font

Text extraction and recognition

State of the art

- Challenges
  - Non-standard fonts
  - Multi-script/orientation/scale
  - Complex background (sound effects)
  - Short length, hyphenation
  - Voluntary spelling mistakes
- Extraction
  - Scene text localization
  - Connected-components approach
  - SVM and Bayesian classifier
  - Sound effects have not been investigated yet
- Recognition
  - At is early stage
  - OCR trained for a specific comics font

- Conclusions
  - Speech text studied but not solved
  - Captions and sound effects unexplored
  - Text recognition not usable yet
  - (Next: automatic font learning?)

Comic character extraction