



## European Ph.D. defense

# Segmentation and indexation of complex objects in comic book images

Christophe Rigaud  
December 11<sup>th</sup>, 2014

Jointly supervised by:

Jean-Christophe Burie<sup>1</sup>  
Dimosthenis Karatzas<sup>2</sup>  
Jean-Marc Ogier<sup>1</sup>

# 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

*"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

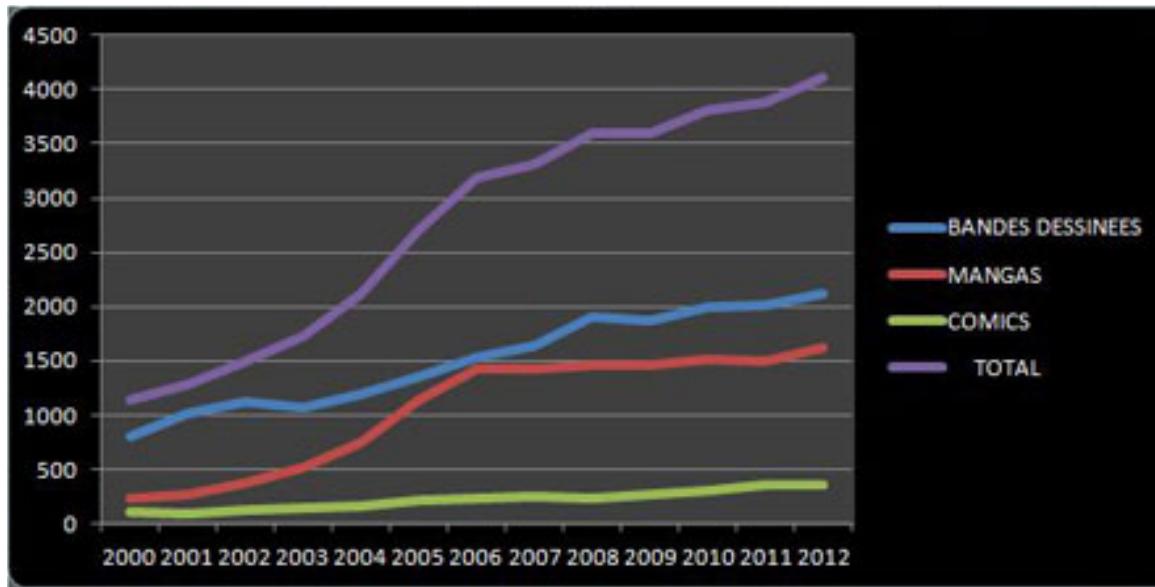
*"a visual medium used to express ideas via images, often combined with text or visual information"* Wikipédia, 2014

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

- 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

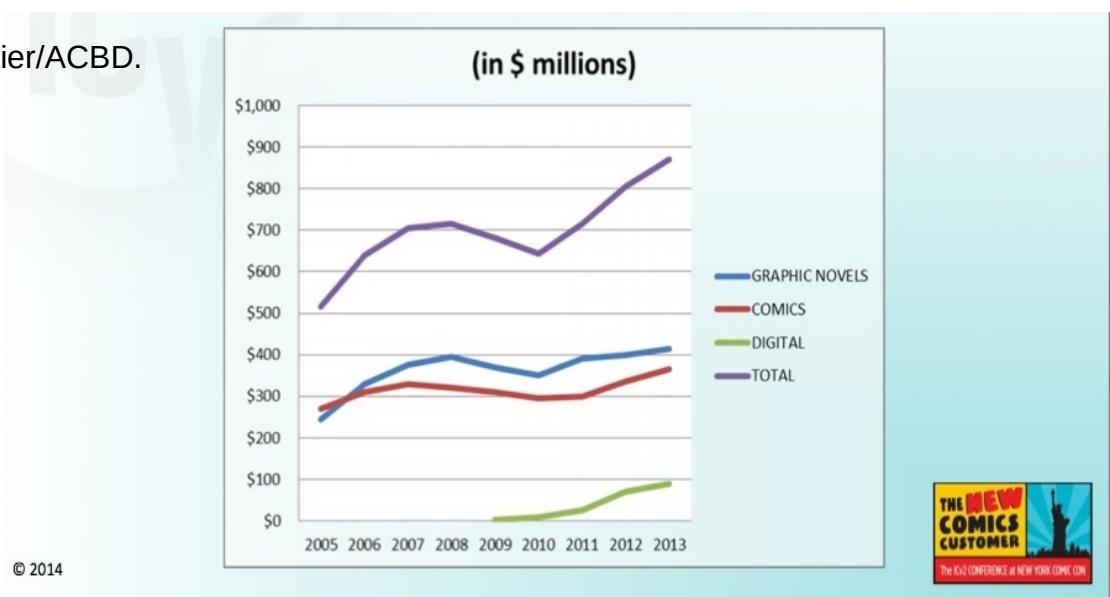
# Market place

## Introduction



Francophone comics production

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



© 2014

Comics market in the US  
Milton Griepp's White Paper, ICv2 Conference 2014

# Comics project (eBDtheque)

## Introduction

- 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 funding
    - 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

## Introduction

- 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

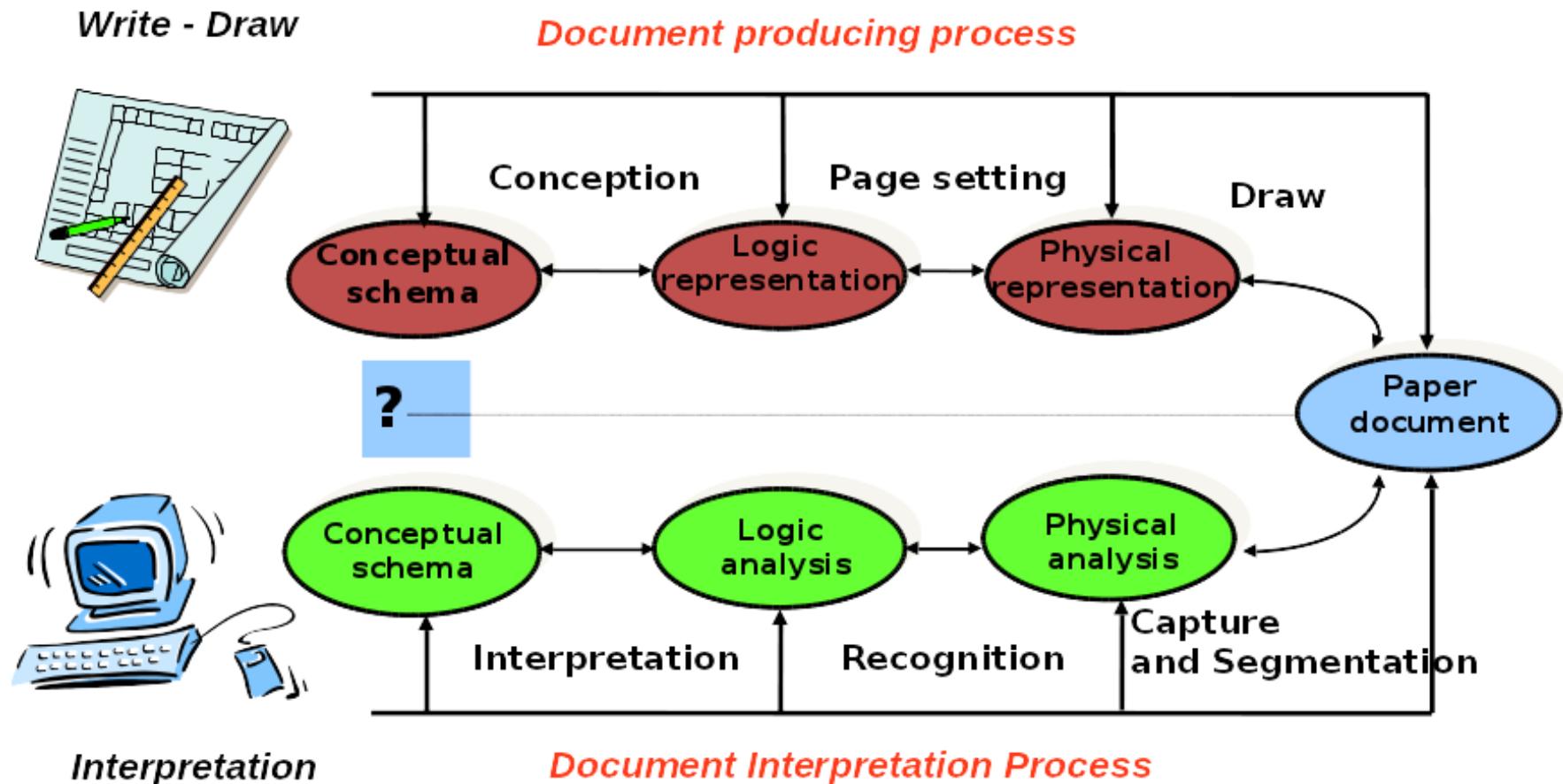
- Introduction
- Document analysis
- State of the art of comics analysis
- Contributions
- Experimentations
- Conclusions

# 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

# Document analysis

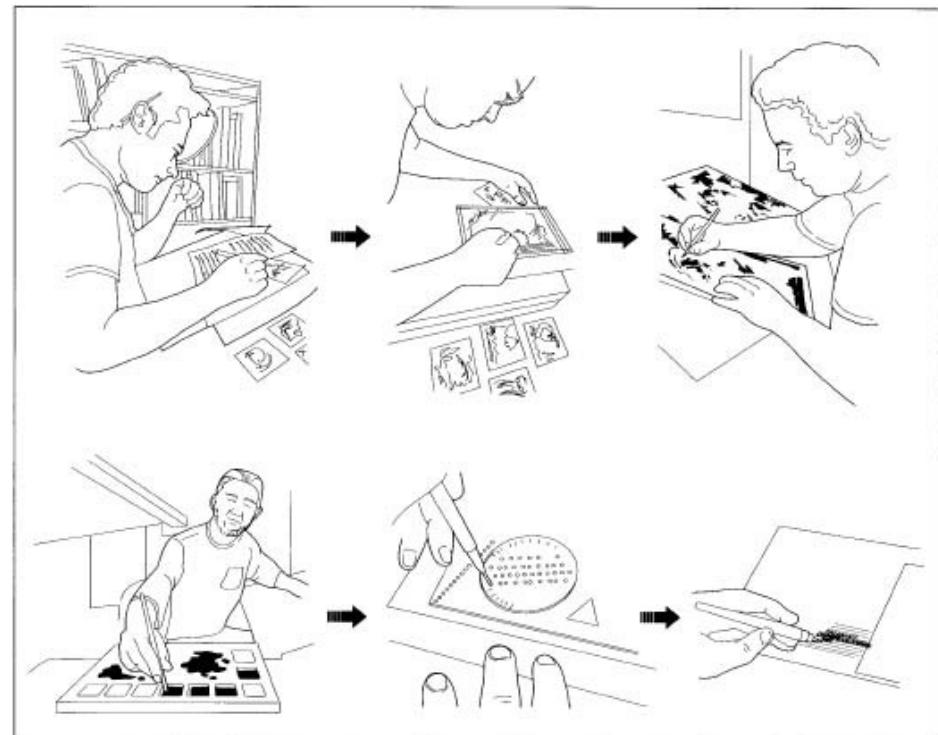


Handbook of Document Image Processing and Recognition. Springer, 2014

# Comic books production

## Document analysis

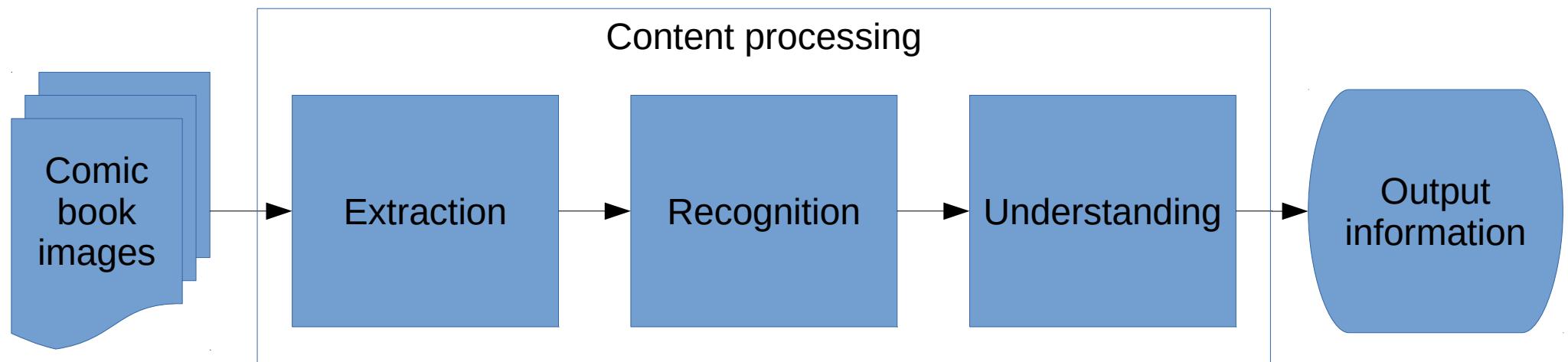
- 1) Synopsis and scenario
- 2) Pencil drawing
- 3) Inking
- 4) Flatting and colouring
- 5) Lettering and sound effects



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

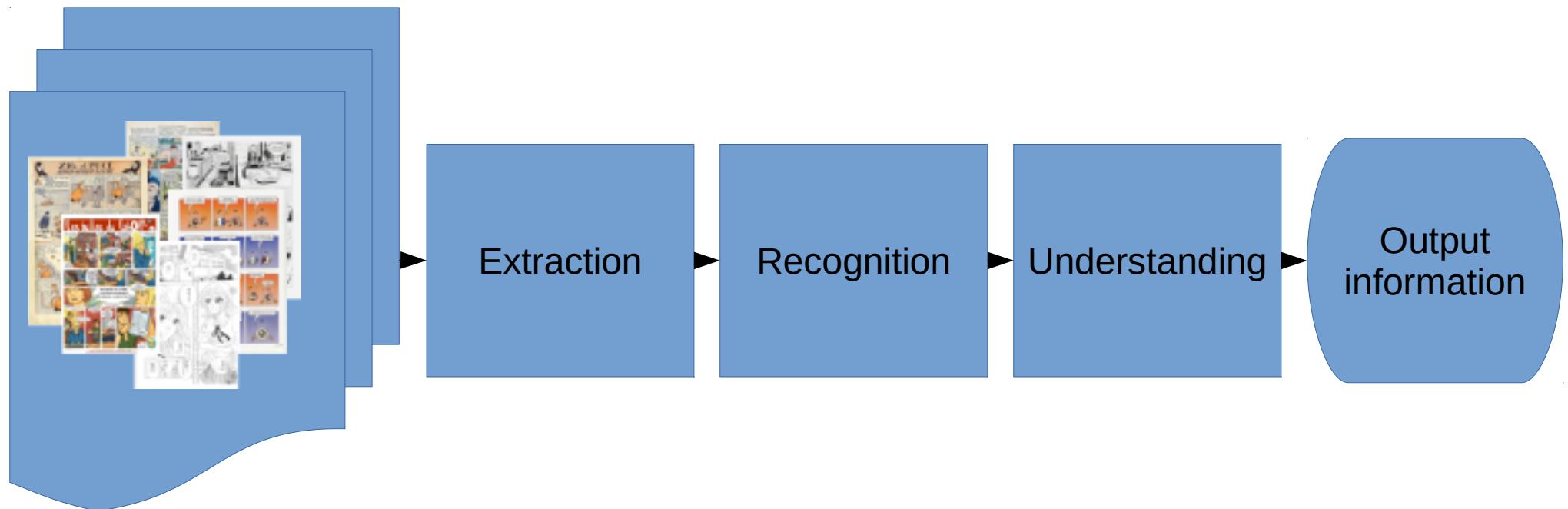
# Comic books interpretation

# Document analysis



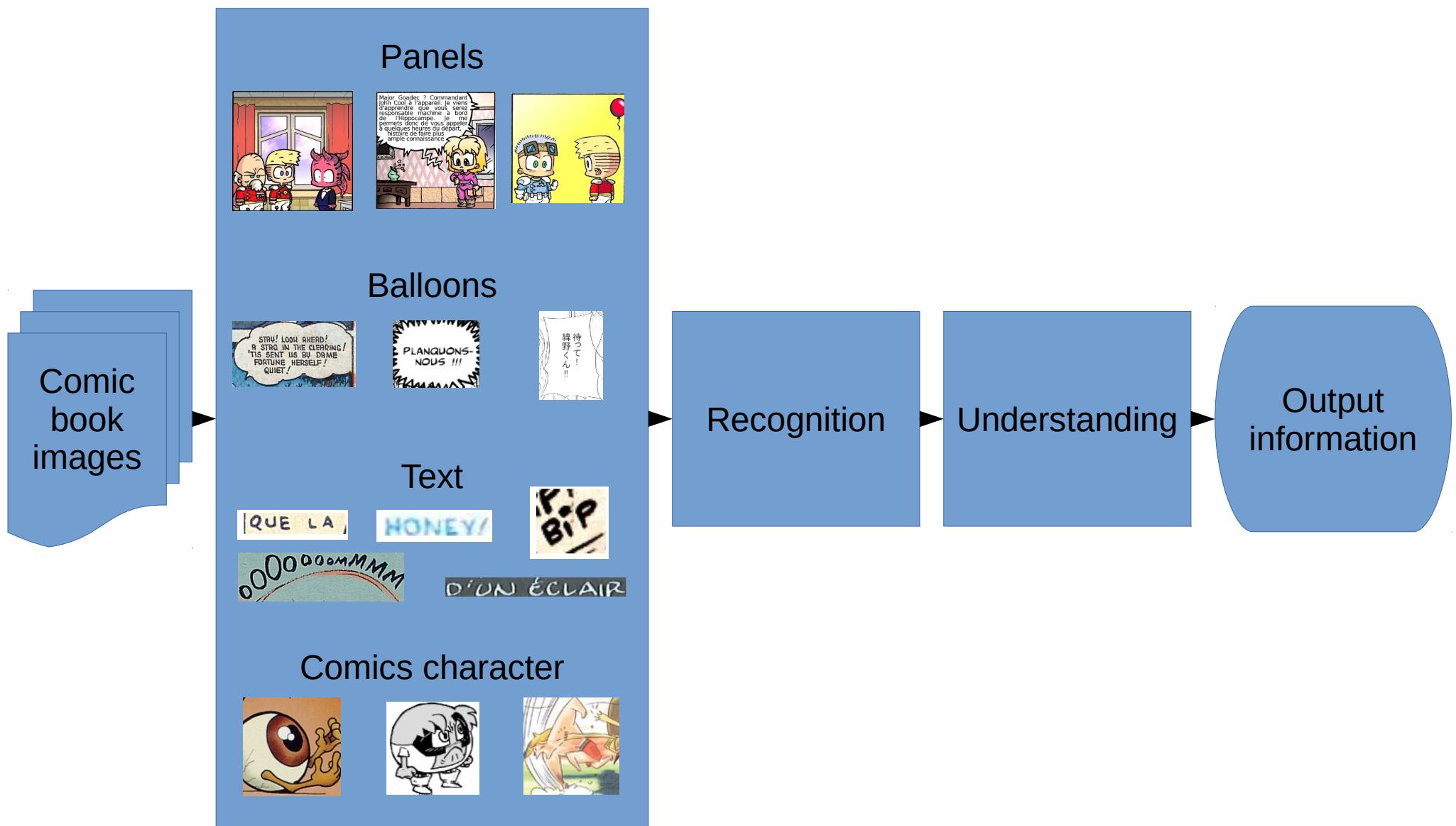
# Comic books interpretation

# Document analysis



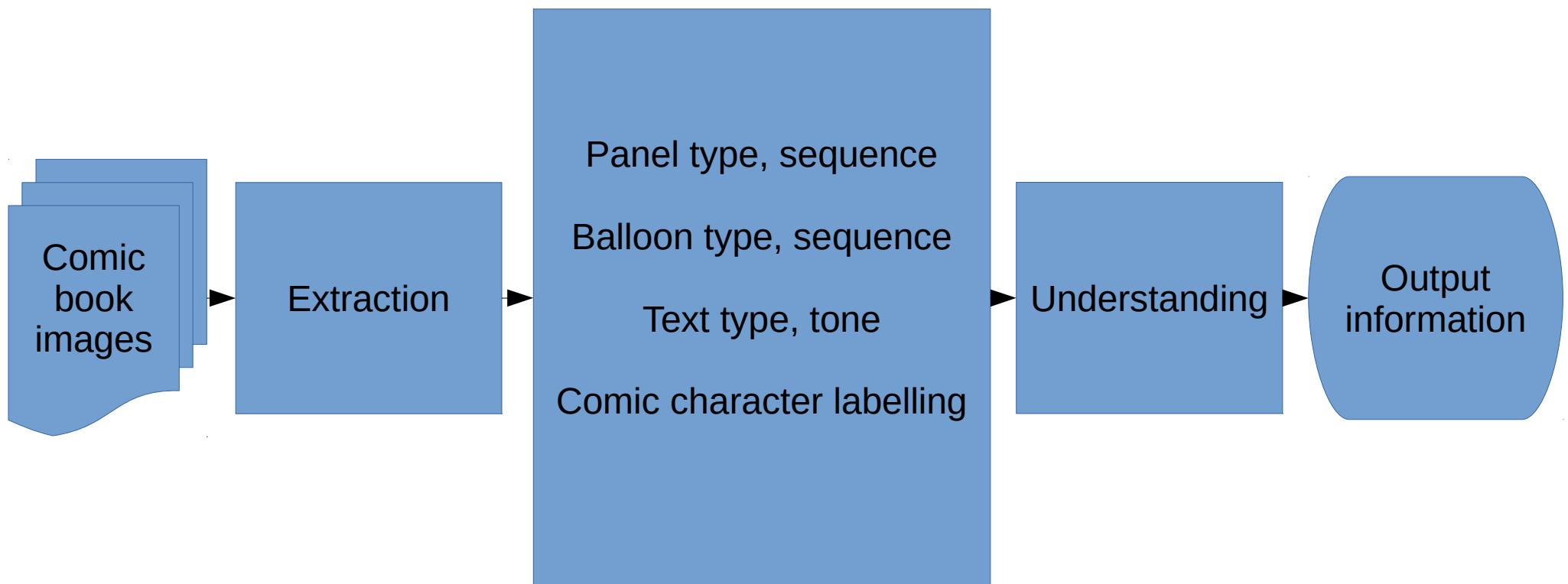
# Comic books interpretation

# Document analysis



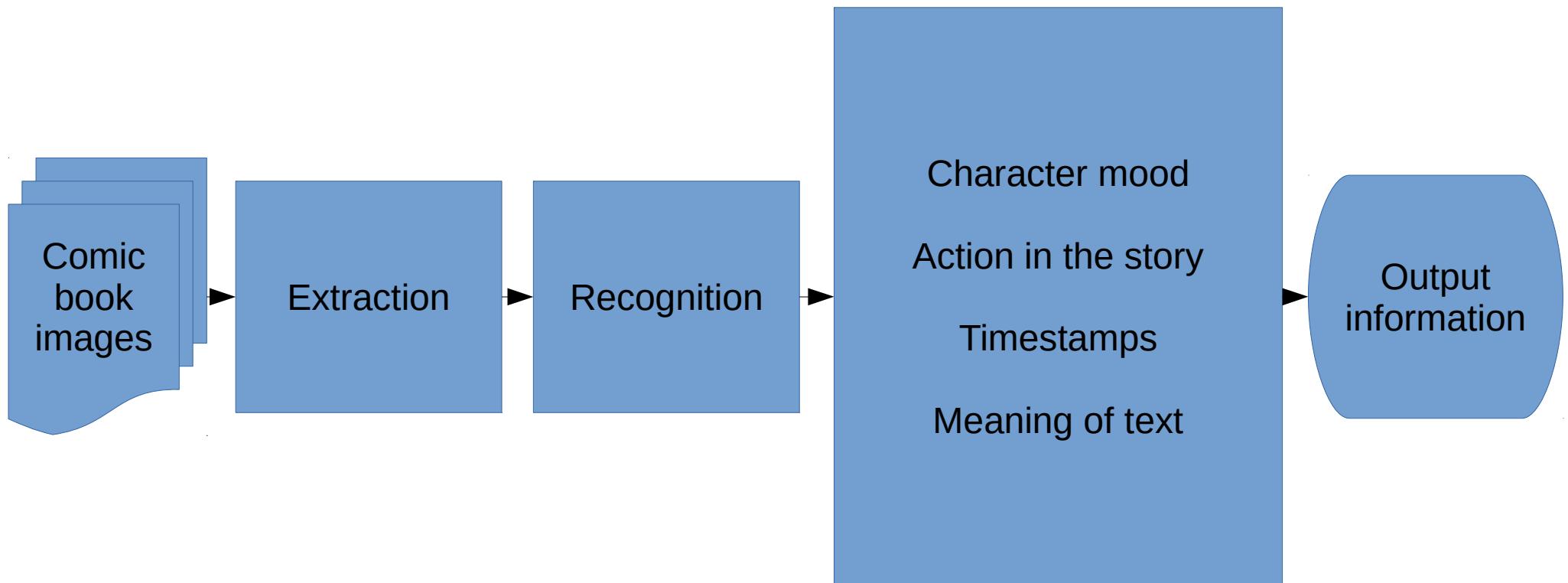
# Comic books interpretation

# Document analysis

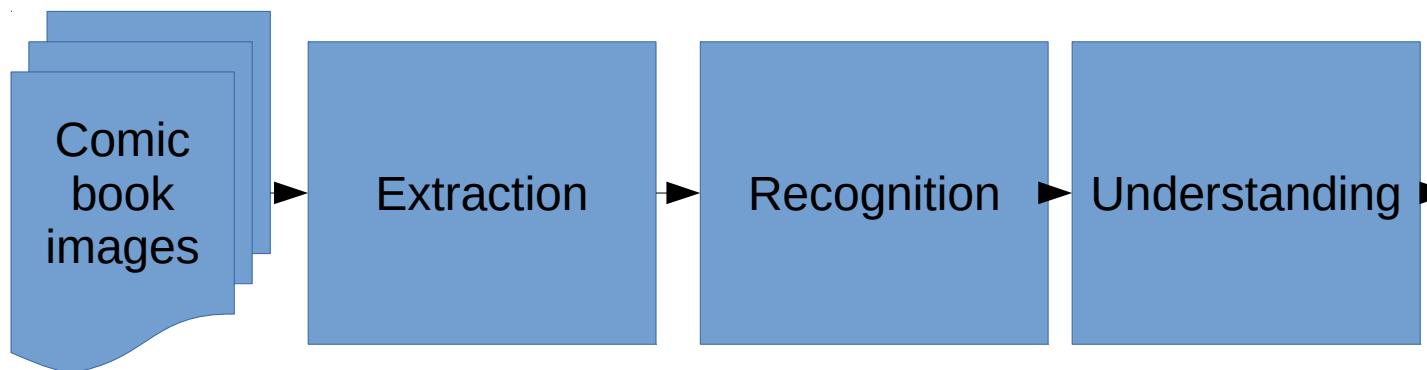


# Comic books interpretation

# Document analysis



# Comic books interpretation



# Document analysis

– IMAGE 1 –

**Contains** 2 panels, 2 **different** characters and 2 balloons

\* PANEL 1 \*

Contains 2 characters  
**saying** 2 balloons

Character 1's **name** is "Bob"  
Bob is **often represented** in blue and dark green

Character 2's **name** is "Tom"  
Tom is **always represented** in yellow, purple and grey

Bob is **shouting** to Tom:  
"Be quiet!!!"

Tom **answers quietly** "not yet"

\* PANEL 2 \*

**Contains** 2 characters,  
same as panel 1, **not talking**  
Bob and Tom are **fighting**

# 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
  - Text extraction and recognition
  - Comic character detection and recognition
  - Holistic understanding
- Contributions
- Experimentations
- Conclusions

# Panel and layout analysis

# State of the art

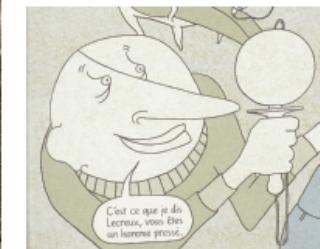
- Challenges
    - Diversity of comic books
    - Semi-structured layout
  - Panel extraction
    - White line cut
    - Recursive X-Y cut algorithm
    - Gradient
    - Connected-components
    - Polygon detection
    - Corners and line segments
  - Layout
    - Reading order (Z-path)
  - Conclusions
    - Problem solved for common ..
    - Remaining difficulties are for ..
    - No approach tested over all ..



# Panel and layout analysis

# State of the art

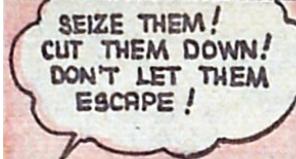
- Challenges
    - Diversity of comic books
    - Semi-structured layout
  - Panel extraction
    - White line cut
    - Recursive X-Y cut algorithm
    - Gradient
    - Connected-components
    - Polygon detection
    - Corners and line segments
  - Layout
    - Reading order (Z-path)
  - Conclusions
    - Problem solved for common
    - Remaining difficulties are for
    - No approach tested over all



# 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

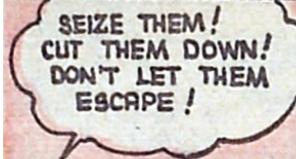
## State of the art

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

## State of the art

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



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

# Comic character extraction

State of the art

- Challenges

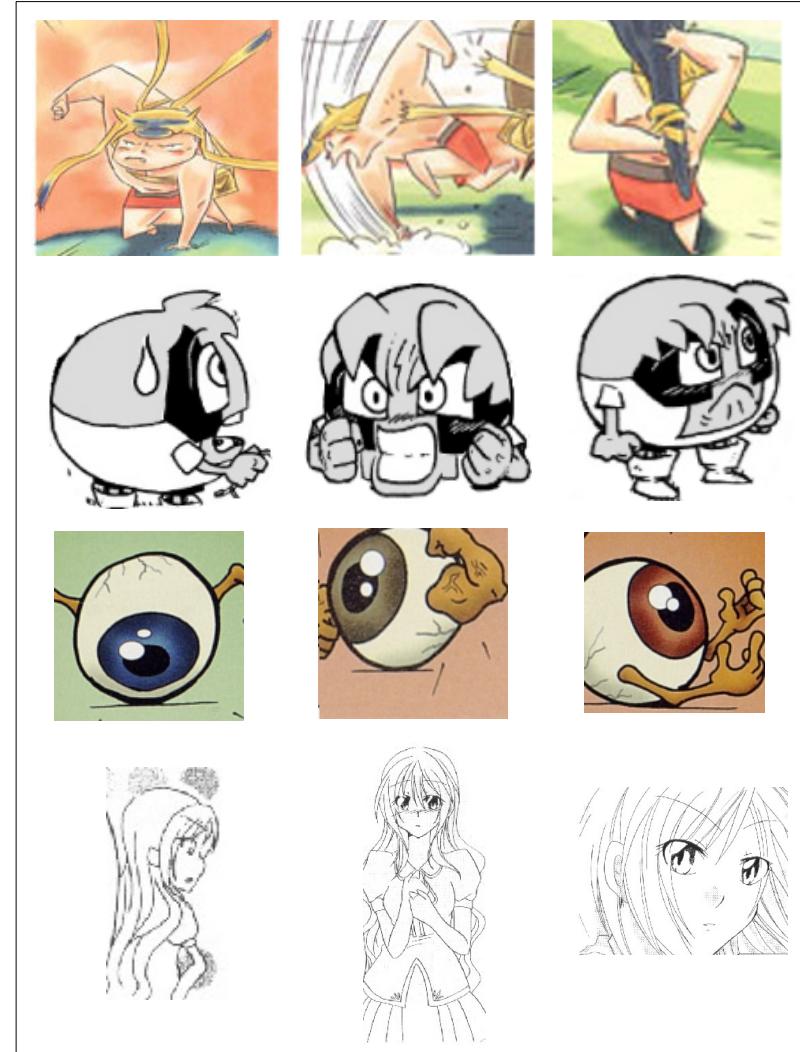
- Hand-drawn, stroke-based
- Intra/inter class variability
- Scale, deformation, posture, occlusion

- Extraction & recognition

- Manga faces [Cheung2008, Sun2010, Kohei2012]
- Cartoons [Khan2012]

- Conclusions

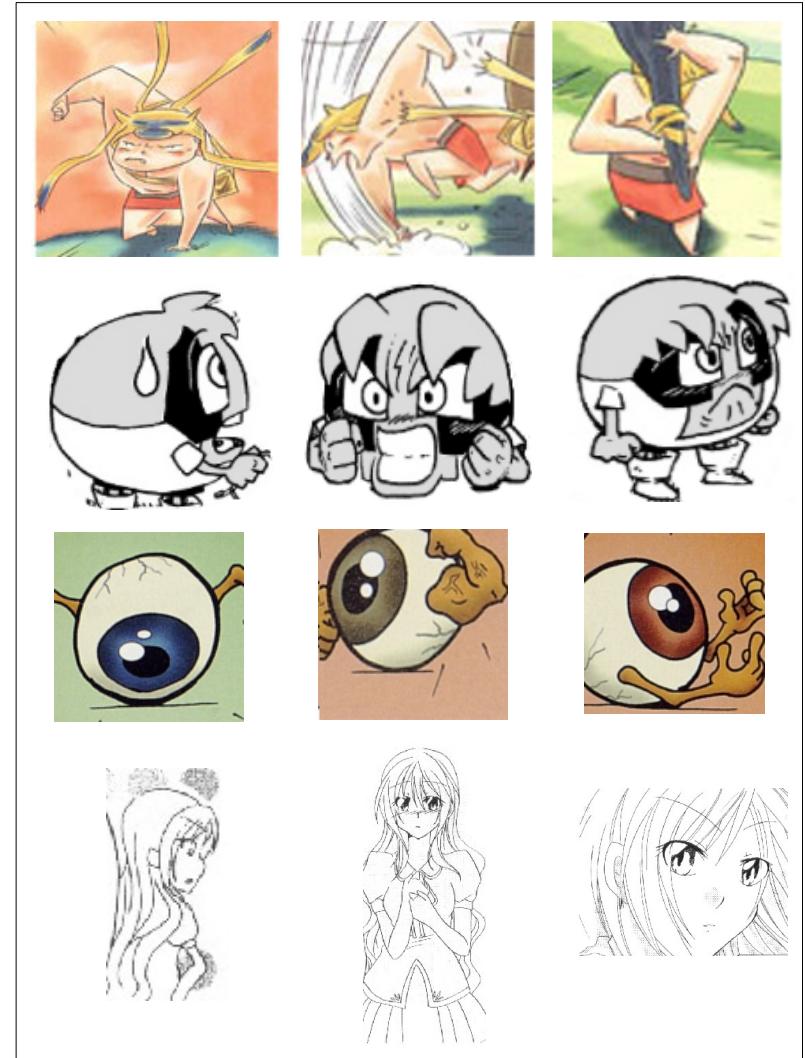
- Preliminary results
- Complex and versatile structure
- Contains most of the interesting information



# Comic character extraction

State of the art

- Challenges
  - Hand-drawn, stroke-based
  - Intra/inter class variability
  - Scale, deformation, posture, occlusion
- Extraction & recognition
  - Manga faces [Cheung2008, Sun2010, Kohei2012]
  - Cartoons [Khan2012]
- Conclusions
  - Preliminary results
  - Complex and versatile structure
  - Contains most of the interesting information



# Outlines

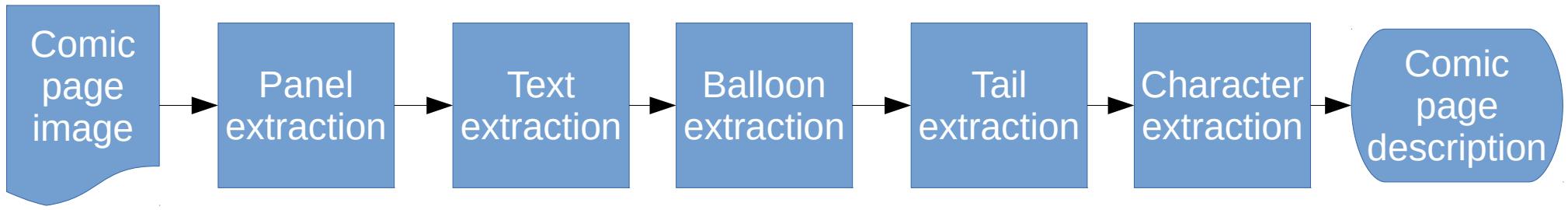
- Introduction
- Document analysis
- State of the art of comics analysis
- Contributions
- Experimentations
- Conclusions

# Outlines

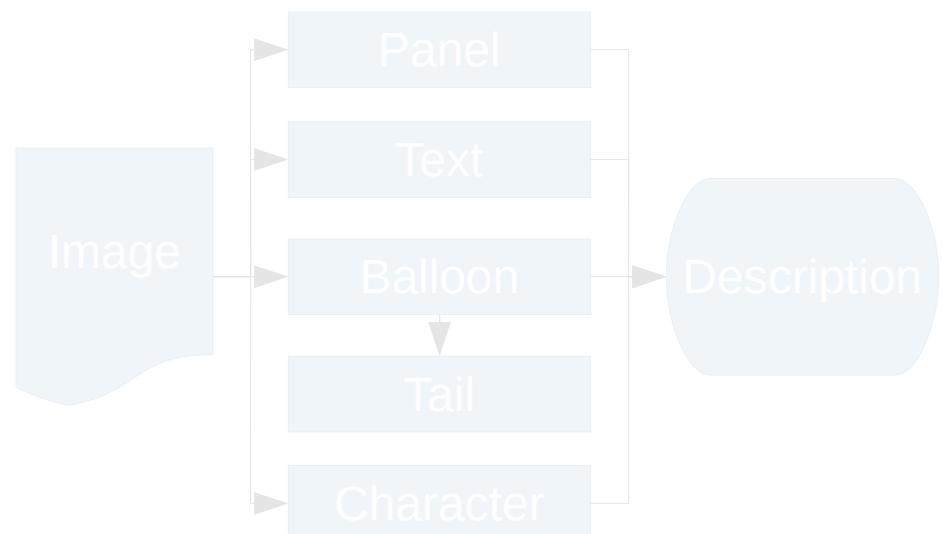
- Introduction
- Document analysis
- State of the art of comics analysis
- Contributions
  - Introduction
  - Sequential approach
  - Independent approach
  - Knowledge-driven approach
  - Conclusion
- Experimentations
- Conclusions

# Three different approaches for comics analysis

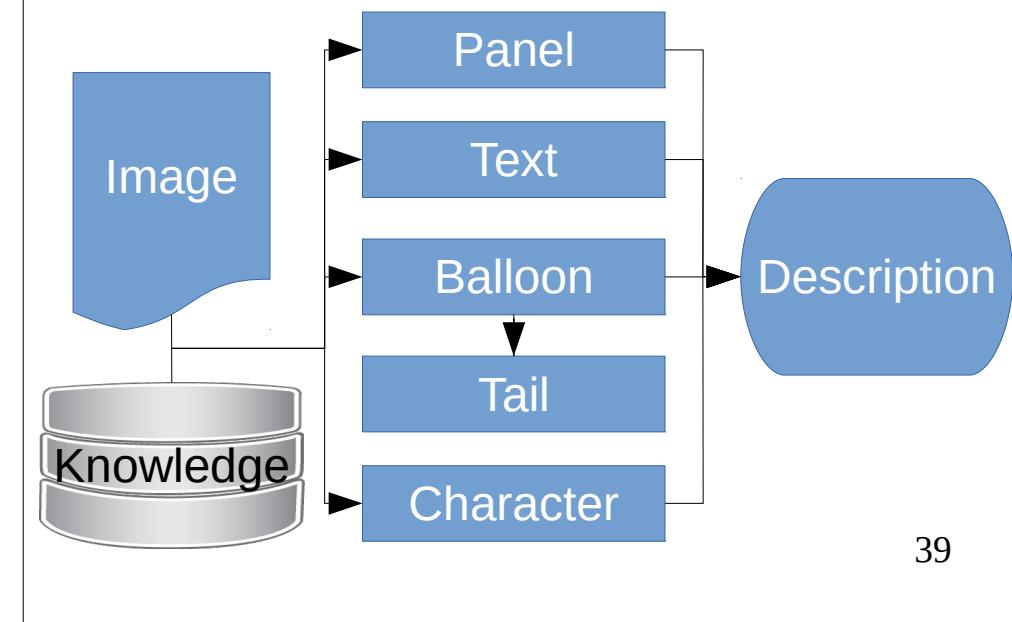
## Sequential



## Independent

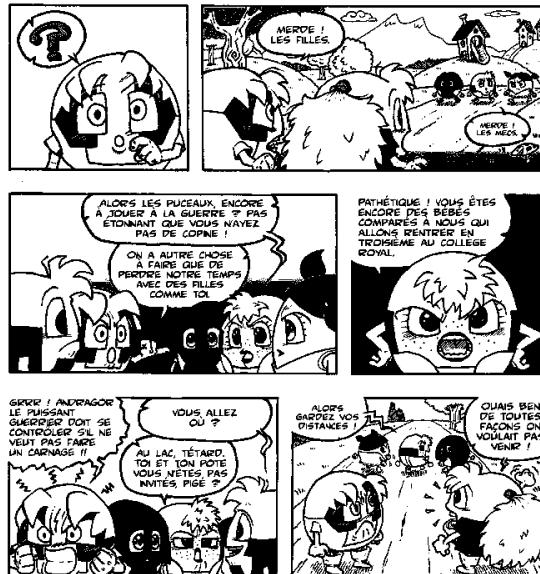
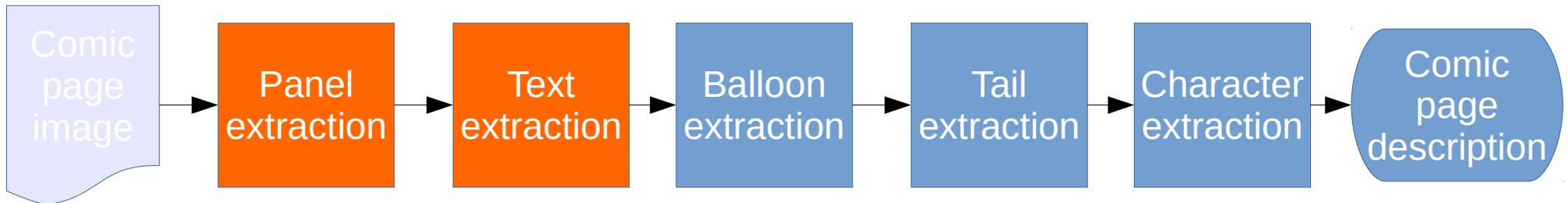


## Knowledge-driven

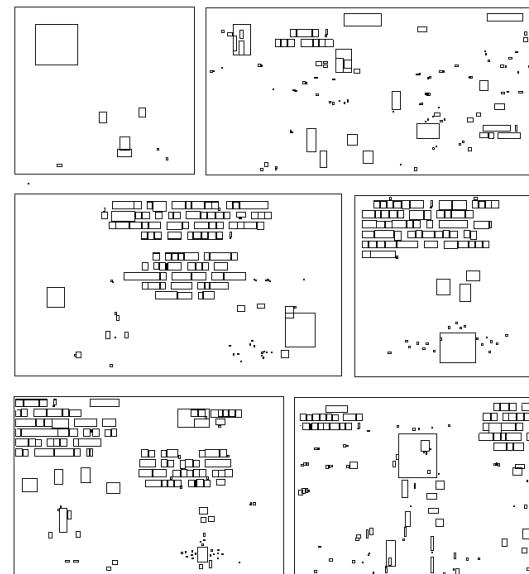


# Panel and text extraction

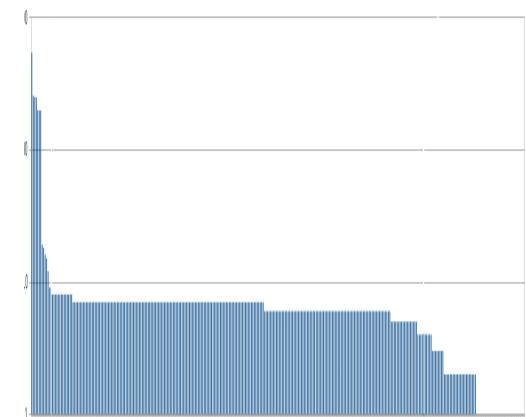
Sequential



Binary image



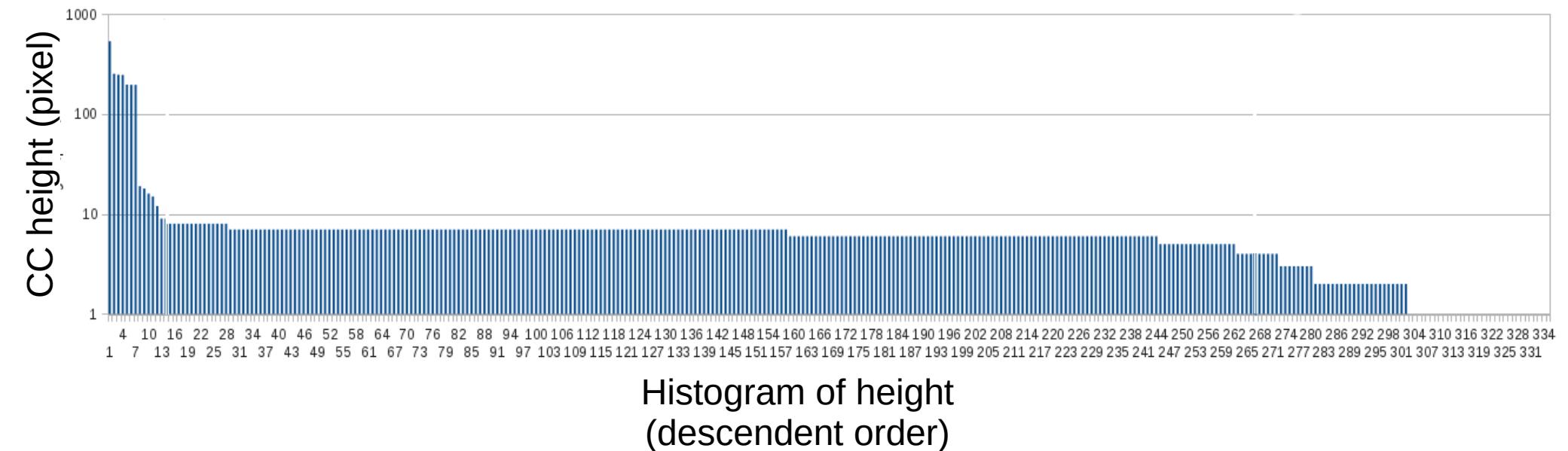
Connected-component (CC) bounding boxes



Histogram of heights of CC

# Panel and text extraction

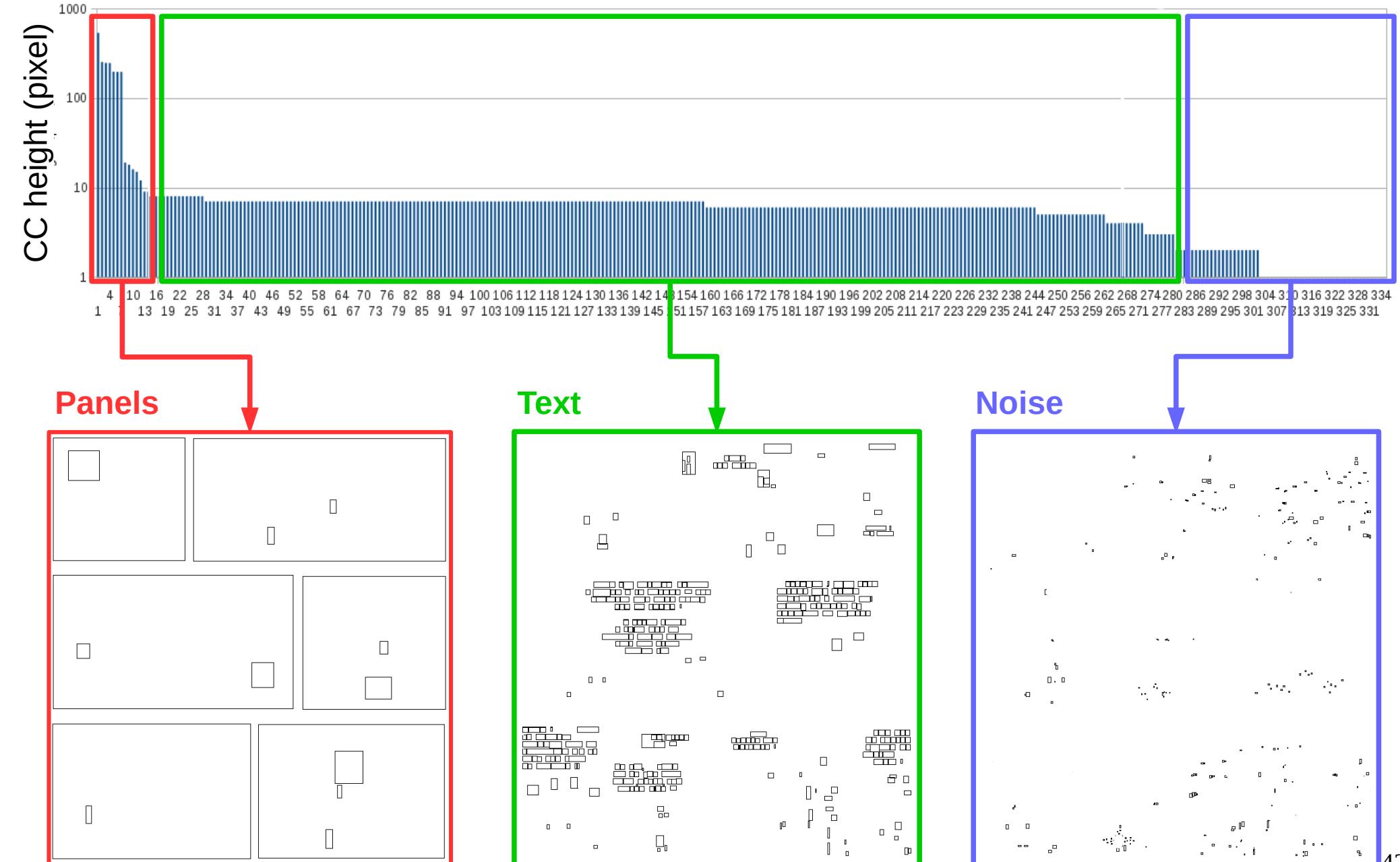
Sequential



# Panel and text extraction

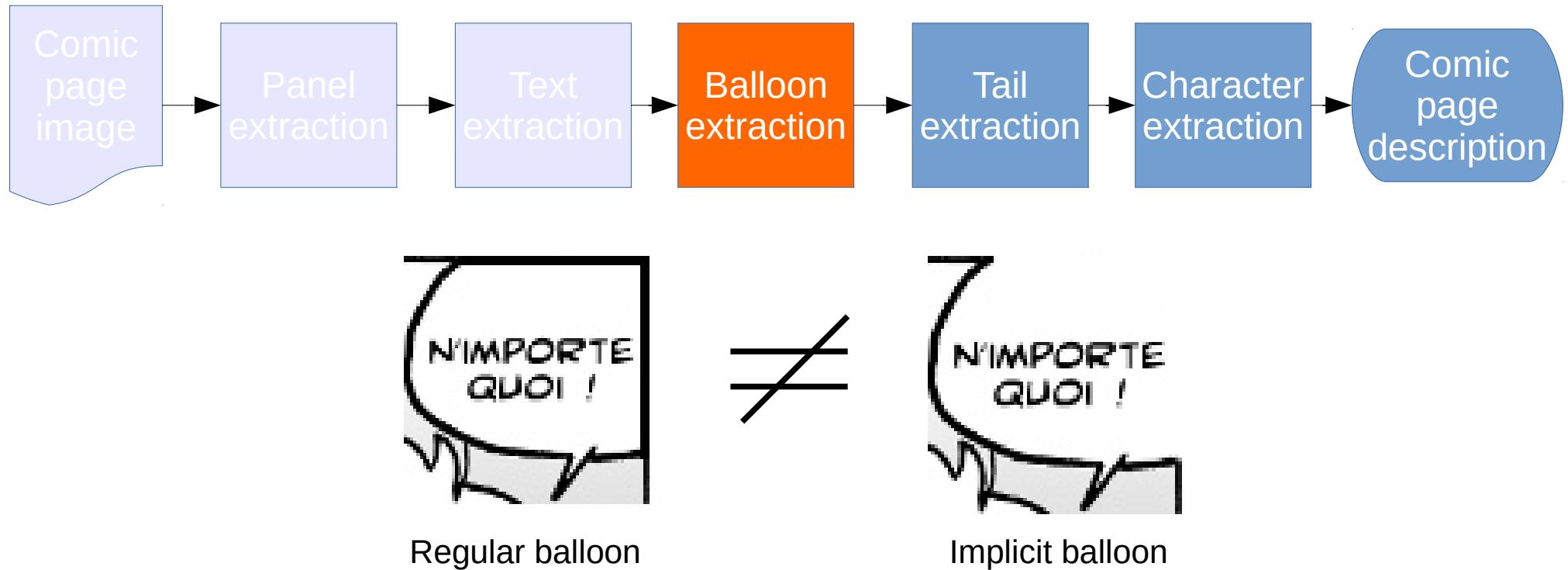
Sequential

K-means clustering (k=3)



# Balloon extraction

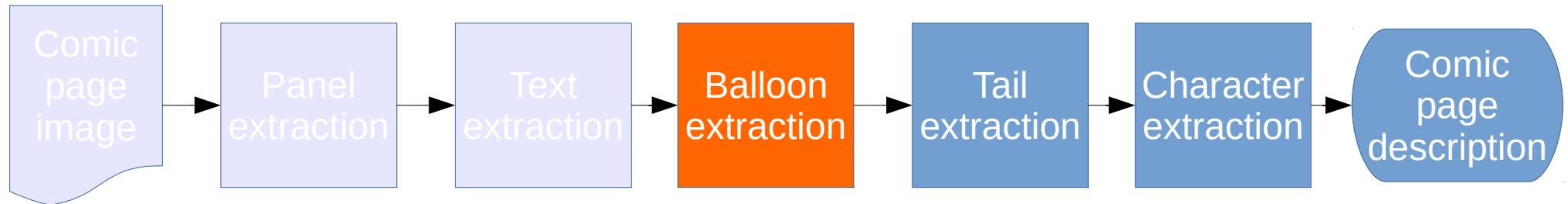
Sequential



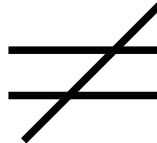
- Literature
  - Top-down approaches: extract white blobs and then text inside
  - Limited to regular balloons
- Contribution
  - Bottom-up approaches: extract text and then surrounding balloons
  - Appropriate for regular and implicit balloons

# Balloon extraction

Sequential



Regular balloon



Implicit balloon

- Literature
  - Top-down approaches: extract white blobs and then text inside
  - Limited to regular balloons
- Contribution
  - Bottom-up approaches: extract text and then surrounding balloons
  - Appropriate for regular and implicit balloons

# Regular balloon extraction

Sequential

- Assumptions
  - Panels and text block positions are known
  - Balloons contain text
  - Text is fully contained and centred in balloons
- Proposition → structural analysis
  - Extract closed contours that includes centred text

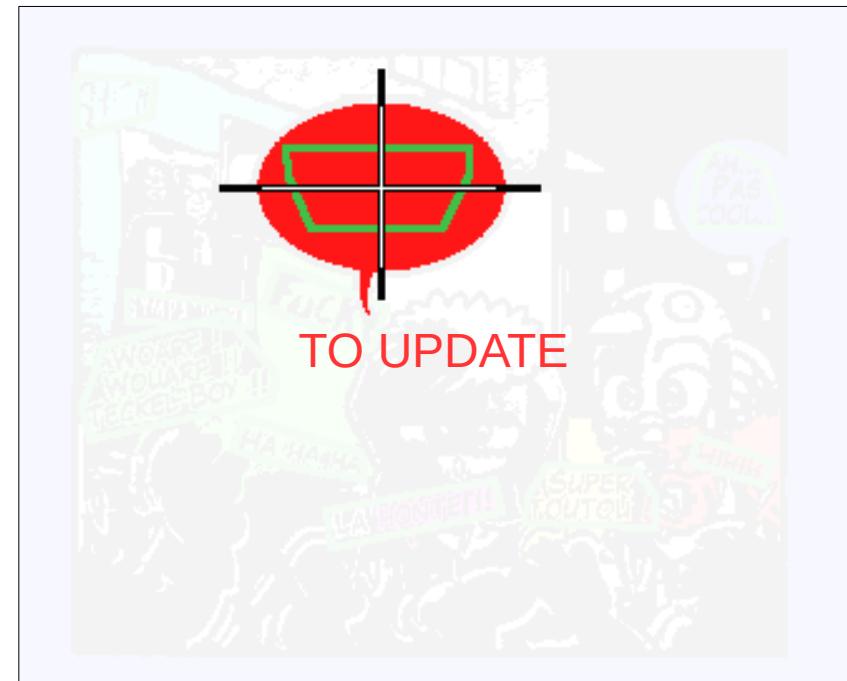
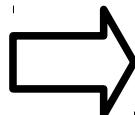
# Regular balloon extraction

Sequential

- Assumptions
  - Panels and text block positions are known
  - Balloons contain text
  - Text is fully contained and centred in balloons
- Proposition → structural analysis
  - Extract closed contours that includes centred text



Original image



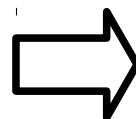
Result of regular balloon extraction

# Regular balloon extraction

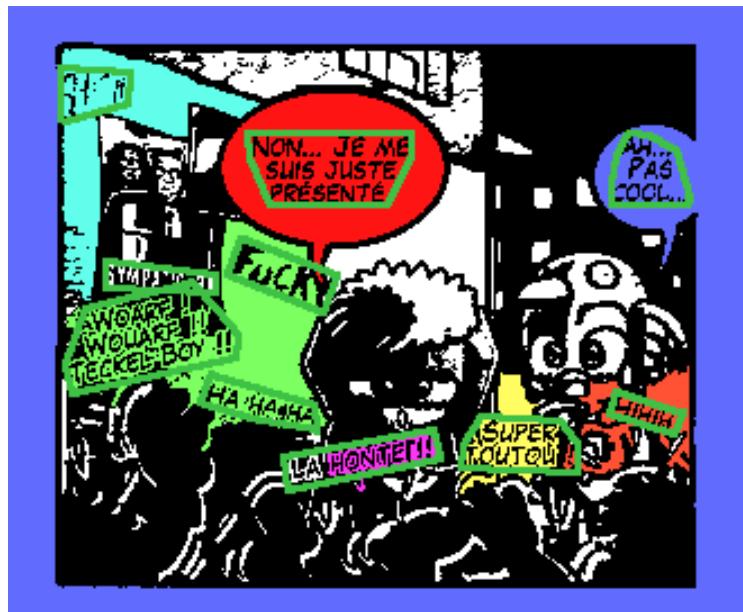
Sequential



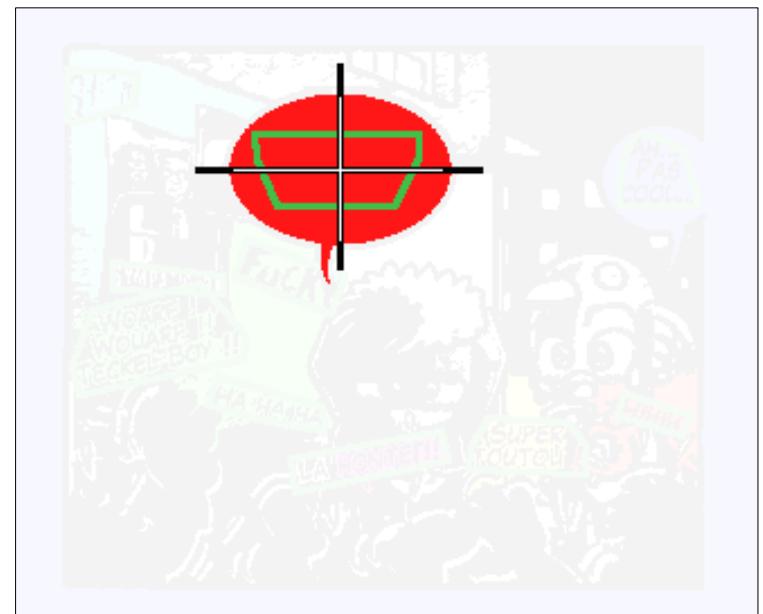
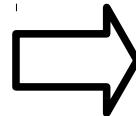
Original image



Text block positions (green)



Regions including text blocks (colours)

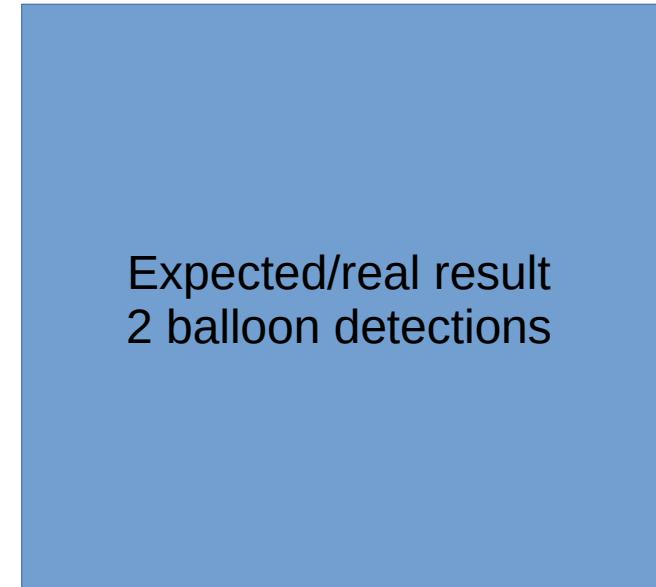
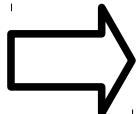


Regions including aligned text block

# Implicit balloon extraction

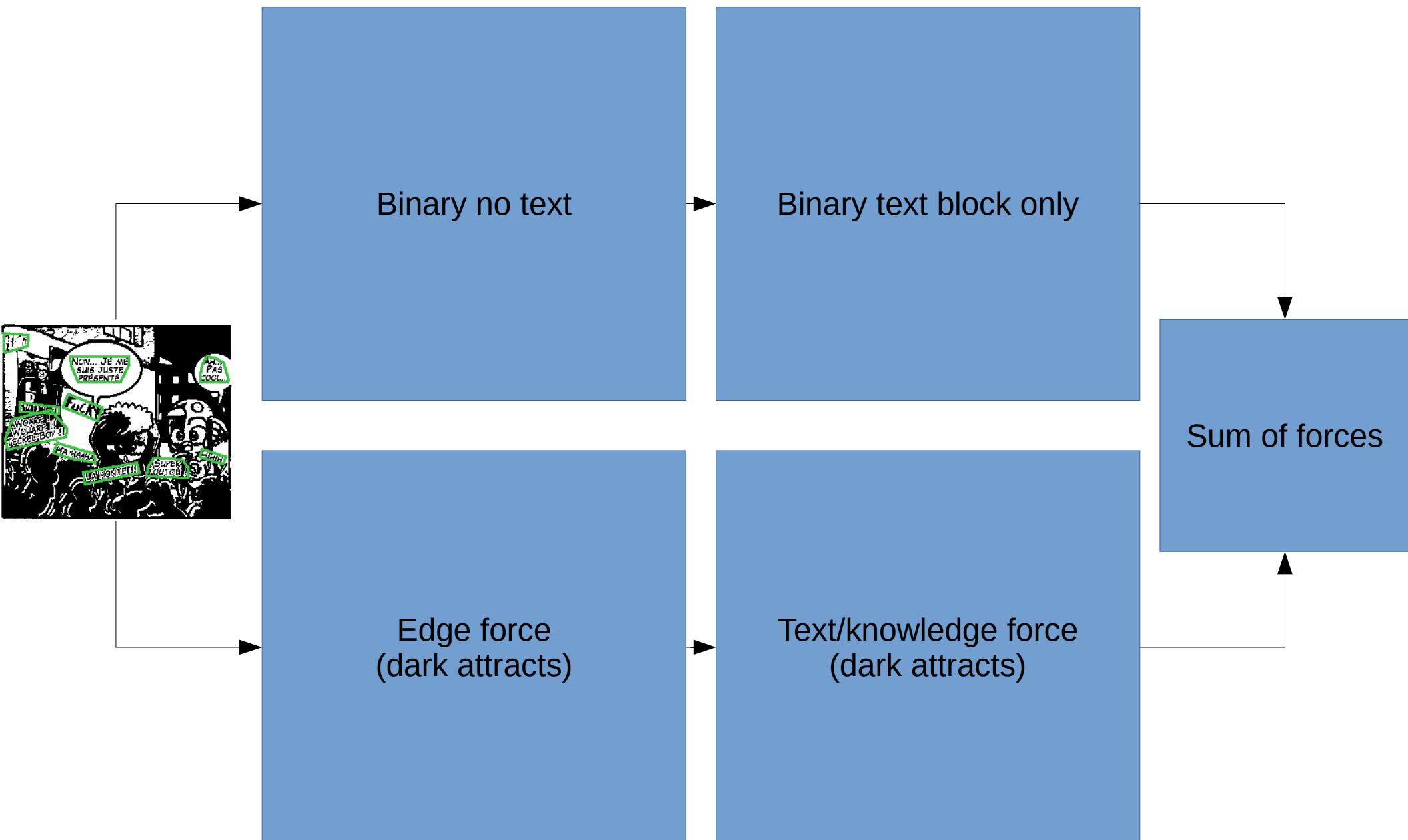
Sequential

- Assumptions
  - Panels and text block positions are known
  - Balloons contain text
  - Text is fully contained and centred in balloons
- Proposition → active contour algorithm
  - Extract implicit contours from text regions



# Implicit balloon extraction

# Sequential



- General conclusion
  - Relies on the performance of text extraction
  - New top-down approach
- Regular balloons
  - Accurate and precise pixel-level extraction
  - New top-down approach
  - Relies on the performance of text extraction
- Implicit balloons
  - First time studied
  - Time consuming
  - Appropriate for regular balloons as well
  - Over-detection