



European Ph.D. defense

Communauté
d'Agglomération de
La Rochelle

Segmentation and indexation of complex objects in comic book images

Christophe Rigaud
December 11th, 2014

Co-supervised by:

Jean-Christophe Burie¹
Dimosthenis Karatzas²
Jean-Marc Ogier¹

Comic books

Introduction

“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”

Hiroaki Tobita, Sony CSL Interaction Laboratory, 2014

Comic books

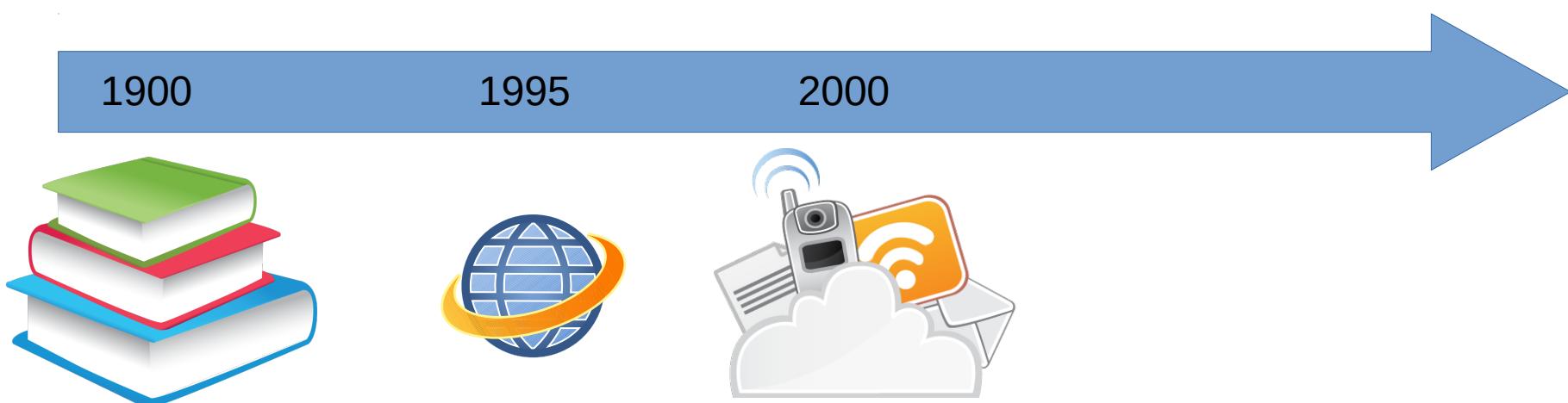
Introduction

“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”

Hiroaki Tobita, Sony CSL Interaction Laboratory, 2014



Comic books

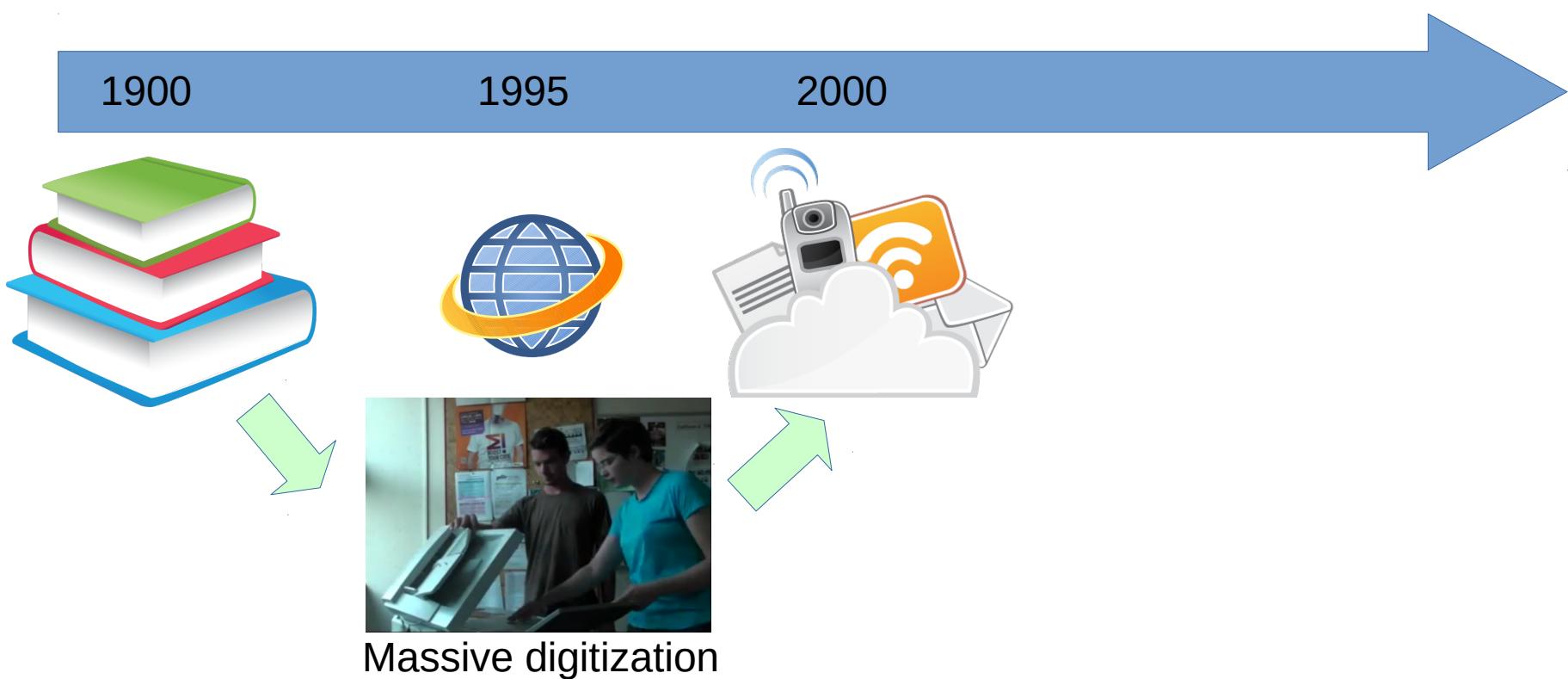
Introduction

“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”

Hiroaki Tobita, Sony CSL Interaction Laboratory, 2014



Comic books

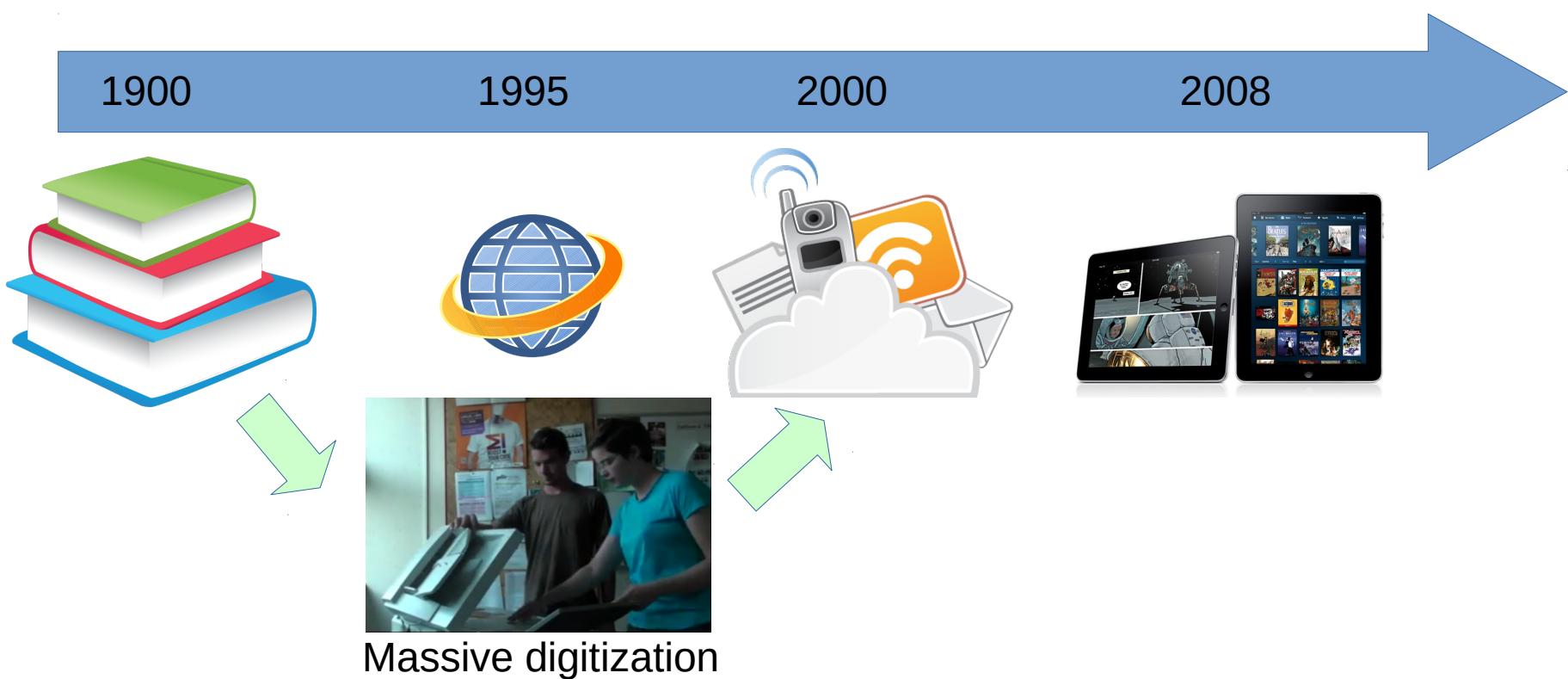
Introduction

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

Wikipédia, 2014

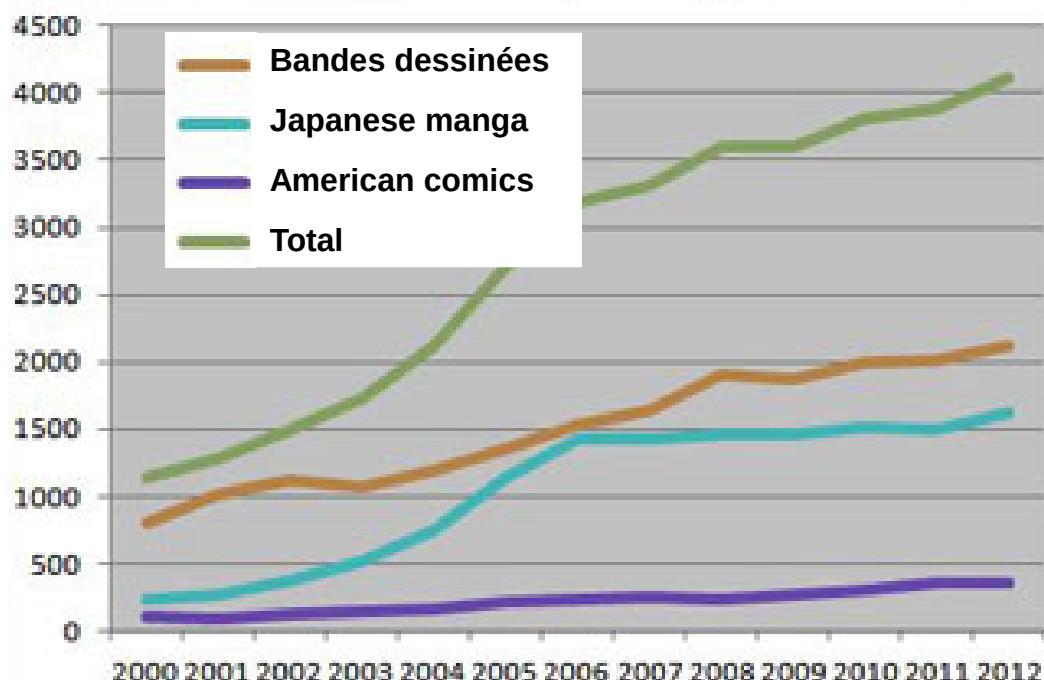
“One of the most popular and familiar forms of graphic content”

Hiroaki Tobita, Sony CSL Interaction Laboratory, 2014



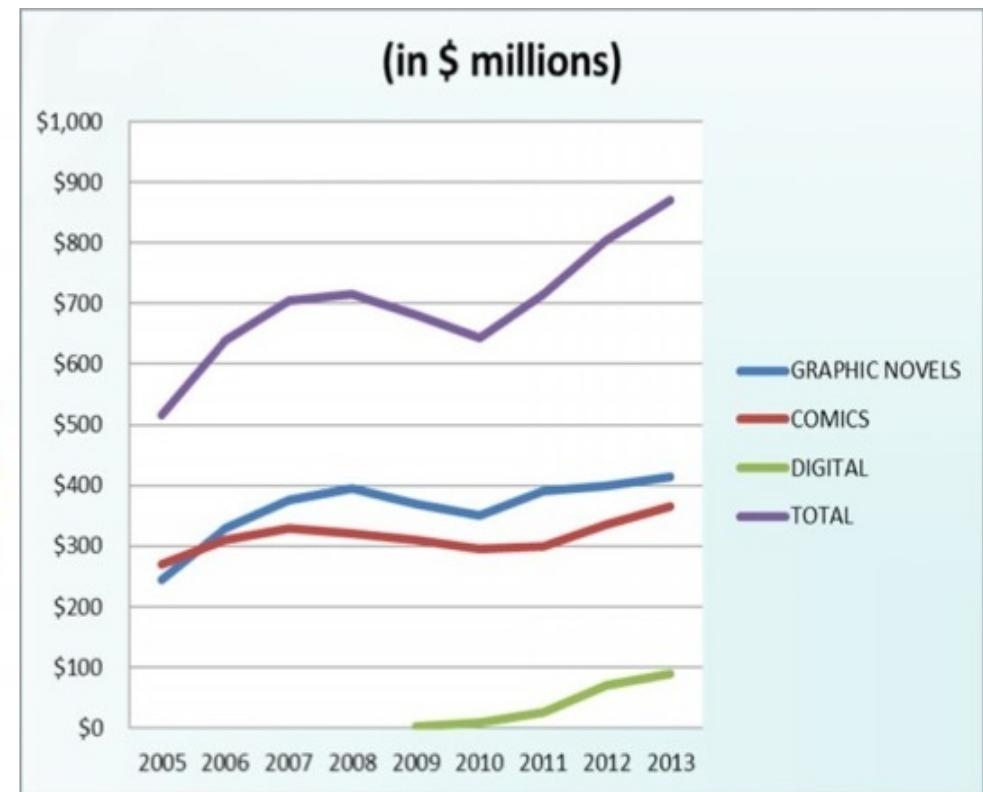
Comic books

Introduction



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

- eBDthèque project (since 2011)
 - Add value to **digitized comics** using the **new technologies**
 - Content extraction (thesis of Christophe Rigaud)
 - Knowledge representation (thesis of Clément Guérin)
 - Public founding CPER 2007-2013
 - 2 Ph.D. students, 1 engineer, 1 post doc, 6 professors (L3i)
- Scientific challenges
 - Mixed contents of a **graphical** and **textual** nature
 - Combination of the difficulties of **free-form** and **complex background** documents
 - Recent field of research with a **largely unknown**
- Objectives
 - Propose **generic approaches** able to **retrieve** as many **elements** as possible from **any comic book image**
 - Provide a **first dataset** and **ground truth**

- eBDthèque project (since 2011)
 - Add value to **digitized comics** using the **new technologies**
 - Content extraction (thesis of Christophe Rigaud)
 - Knowledge representation (thesis of Clément Guérin)
 - Public founding CPER 2007-2013
 - 2 Ph.D. students, 1 engineer, 1 post doc, 6 professors (L3i)
- Scientific challenges
 - Mixed contents of a **graphical** and **textual** nature
 - Combination of the difficulties of **free-form** and **complex background** documents
 - Recent and **unexplored** field of research
- Objectives
 - Propose **generic approaches** able to **retrieve** as many **elements** as possible from **any comic book image**
 - Provide a **first dataset** and **ground truth**

- eBDthèque project (since 2011)
 - Add value to **digitized comics** using the **new technologies**
 - Content extraction (thesis of Christophe Rigaud)
 - Knowledge representation (thesis of Clément Guérin)
 - Public founding CPER 2007-2013
 - 2 Ph.D. students, 1 engineer, 1 post doc, 6 professors (L3i)
- Scientific challenges
 - Mixed contents of a **graphical** and **textual** nature
 - Combination of the difficulties of **free-form** and **complex background** documents
 - Recent and **unexplored** field of research
- Objectives
 - Propose **generic approaches** able to **retrieve** as many **elements** as possible from **any comic book image**
 - Provide a **first dataset** and **ground truth**

- Panel extraction
- Balloon extraction
- Text extraction & recognition
- Comic character extraction
- Conclusions



Pencil drawing. Image credits: Le cycle des bulles, Christophe Rigaud, 2012

Panel extraction

Background

- Challenges

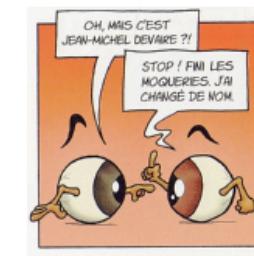
- Diversity of styles (gutter, implicit)
- Semi-structured layout

- Panel extraction

- White line cut [Chung07]
- Recursive X-Y cut [Eunjung07]
- Density gradient [Tanaka07]
- Connected-components [Arai10, Pang14]
- Polygon detection [Li14a]
- Corners and line segments [Stommel12]

- Conclusions

- Specific approaches not appropriate for other type of comics
- Remaining difficulties for non-rectangle and implicit panels
- Copyrighted images (not shareable)



Panel extraction

Background

- Challenges

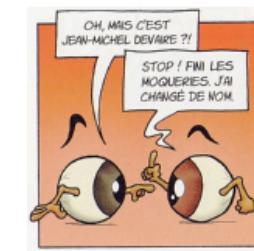
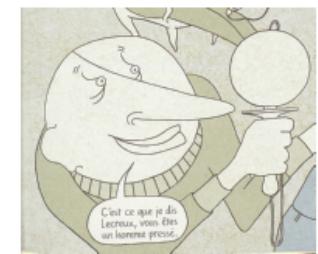
- Diversity of styles (gutter, implicit)
- Semi-structured layout

- Panel extraction

- White line cut [Chung07]
- Recursive X-Y cut [Eunjung07]
- Density gradient [Tanaka07]
- Connected-components [Arai10, Pang14]
- Polygon detection [Li14a]
- Corners and line segments [Stommel12]

- Conclusions

- Specific approaches not appropriate for other type of comics
- Remaining difficulties for non-rectangle and implicit panels
- Copyrighted images (not shareable)



Balloon extraction

- Challenges

- Difference between **shape** and **contour**
- **Implicit balloon positions**
- **Semantics** related to text

- Extraction

- **Connected-components** [Arai11, Ho12]

- Conclusions

- **Closed** balloon with text inside
- Several **unexplored** fields (e.g. **implicit** balloon positions, balloon, **classification**, **tail detection**)

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

Balloon extraction

- Challenges

- Difference between **shape** and **contour**
- **Implicit balloon positions**
- **Semantics** related to text

- Extraction

- **Connected-components** [Arai11, Ho12]

- Conclusions

- **Closed** balloon with text inside
- Several **unexplored** fields (e.g. **implicit** balloon positions, balloon, classification, tail detection)

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

Text extraction & recognition

Background

- Challenges
 - Non-standard fonts
 - Multi-script/orientation/scale
 - Complex background (sound effects)
 - Hyphenation, voluntary spelling mistakes
 - Extraction
 - Sliding Concentric Windows + SVM [Su11]
 - Connected-components [Ho12, Pang14]
 - SVM and Bayesian classifier [Li14b]
 - Recognition
 - OCR trained for a specific comics font [Ponsard12]



Text extraction & recognition

Background

- Challenges
 - Non-standard fonts
 - Multi-script/orientation/scale
 - Complex background (sound effects)
 - Hyphenation, voluntary spelling mistakes
 - Extraction
 - Sliding Concentric Windows + SVM [Su11]
 - Connected-components [Ho12, Pang14]
 - SVM and Bayesian classifier [Li14b]
 - Recognition
 - OCR trained for a specific comics font [Ponsard12]



- Conclusions
 - Speech text only (from balloons)
 - Captions and sound effects unexplored
 - Text recognition very poor

Comic character extraction

Background

- Challenges

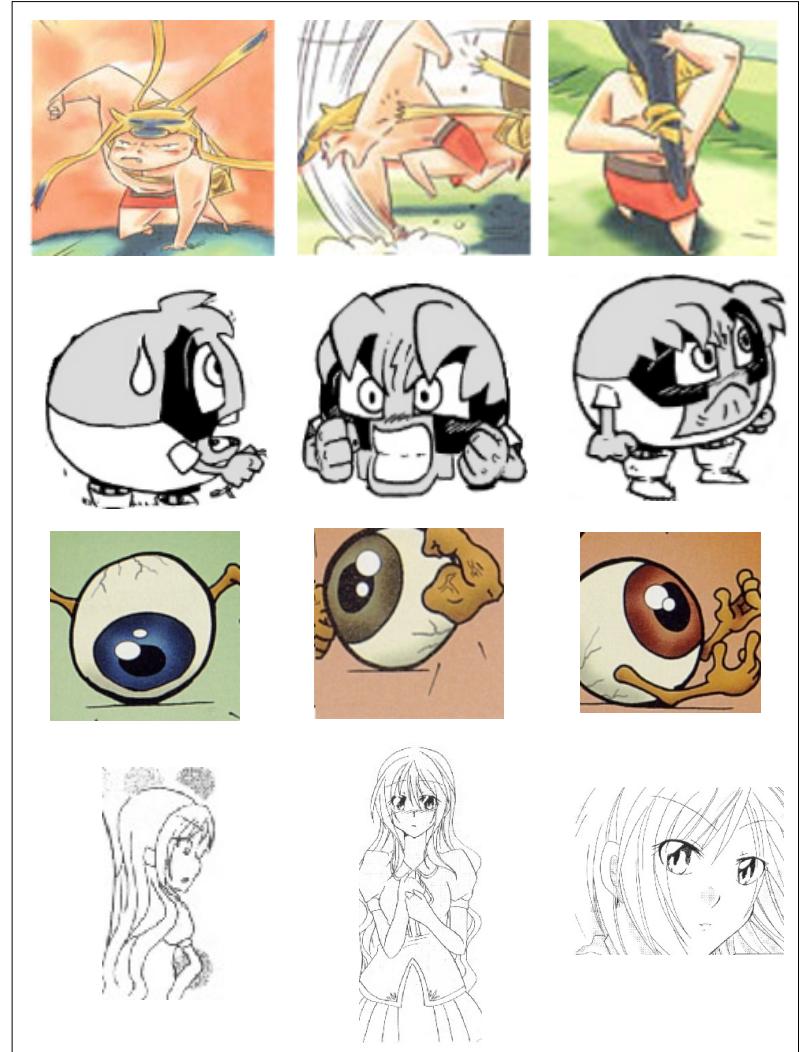
- Hand-drawn, stroke-based
- Intra/inter class variability
- Scale, deformation, posture, occlusion

- Extraction & recognition

- Manga faces [Cheung08, Sun10, Kohei12]
- Cartoons [Khan12]

- Conclusions

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



Comic character extraction

Background

- Challenges

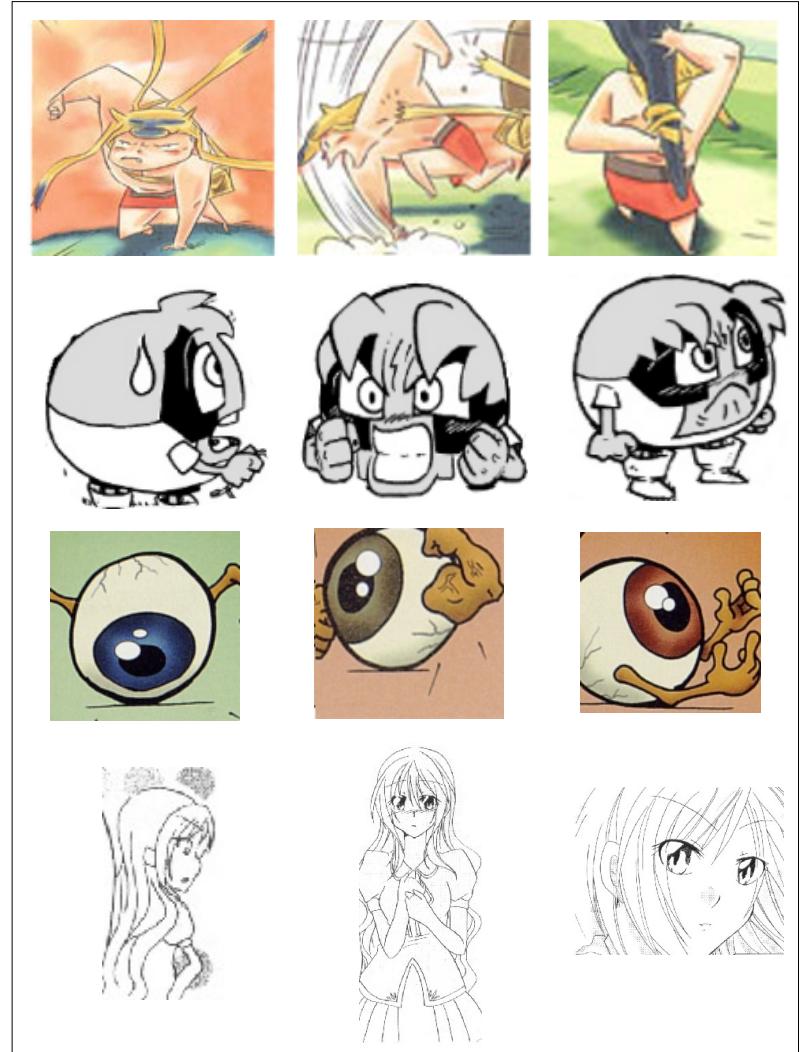
- Hand-drawn, stroke-based
- Intra/inter class variability
- Scale, deformation, posture, occlusion

- Extraction & recognition

- Manga faces [Cheung08, Sun10, Kohei12]
- Cartoons [Khan12]

- Conclusions

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



Conclusions

Background

Element	Process type	Status
Panel	Localisation	
	Classification	
Balloon	Localisation	
	Classification	
	Tail detection	
Text	Localisation	
	Recognition	
Comic character	Localisation	
	Identification	
	Face/pose	
Context	Inter-element link	
	Situation retrieval	
	Timestamps	
Dataset	Localisation	
	Semantic	

Solved
 Advanced
 Medium
 Early stage
 Unexplored

- Introduction
- Sequential approach
- (Independent approach)
- Knowledge-driven approach



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

- **Introduction**
- Sequential approach
- (Independent approach)
- Knowledge-driven approach



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Introduction

- Objective: cover the widest possible scope of study

Contributions

Introduction

- Objective: cover the widest possible scope of study

1) Creation of heterogeneous dataset

- 100 mixed pages from 20 albums
- Franco-Belgium “bandes dessinées”, American comics and Japanese manga
- From 1905 to 2012, paper and webcomics
- Rights holder permissions agreement

Contributions

Bibliographic annotations



PAGE (100)
Collection: Chilling Tales
Album: 17 Geo
Editor: Youthful Magazines
Drawer: Matt Fox
Writer: Matt Fox
Language: English
Page number: 16
Release date: 1953

Visual and semantic annotations

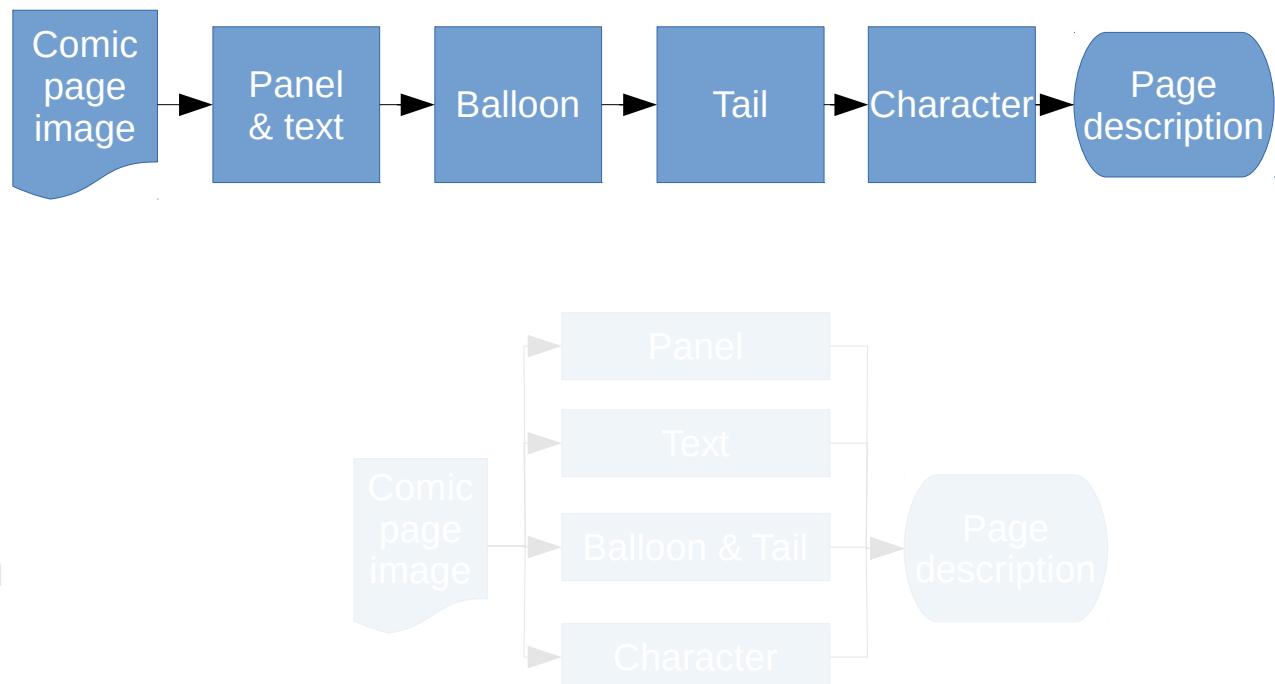


PANEL (850)
Rank: 1
BALLOON (1092)
Rank: 2
Shape: Oval
Tail direction: South-West
TEXT LINE (4691)
Text: « STARK RAVING »
CHARACTER (1550)
LinkedToBalloon: 2

Introduction

- Objective: cover the widest possible scope of study
 - 1) Creation of heterogeneous dataset
 - 2) Three approaches
- Content-driven
 - Sequential approach
 - Similar to literature
 - Intuitive
 - Sensible to error propagation
 - Independent approach
 - Avoid error propagation
- Knowledge-driven
 - Knowledge-driven approach
 - Based on domain knowledge
 - Retrieve context

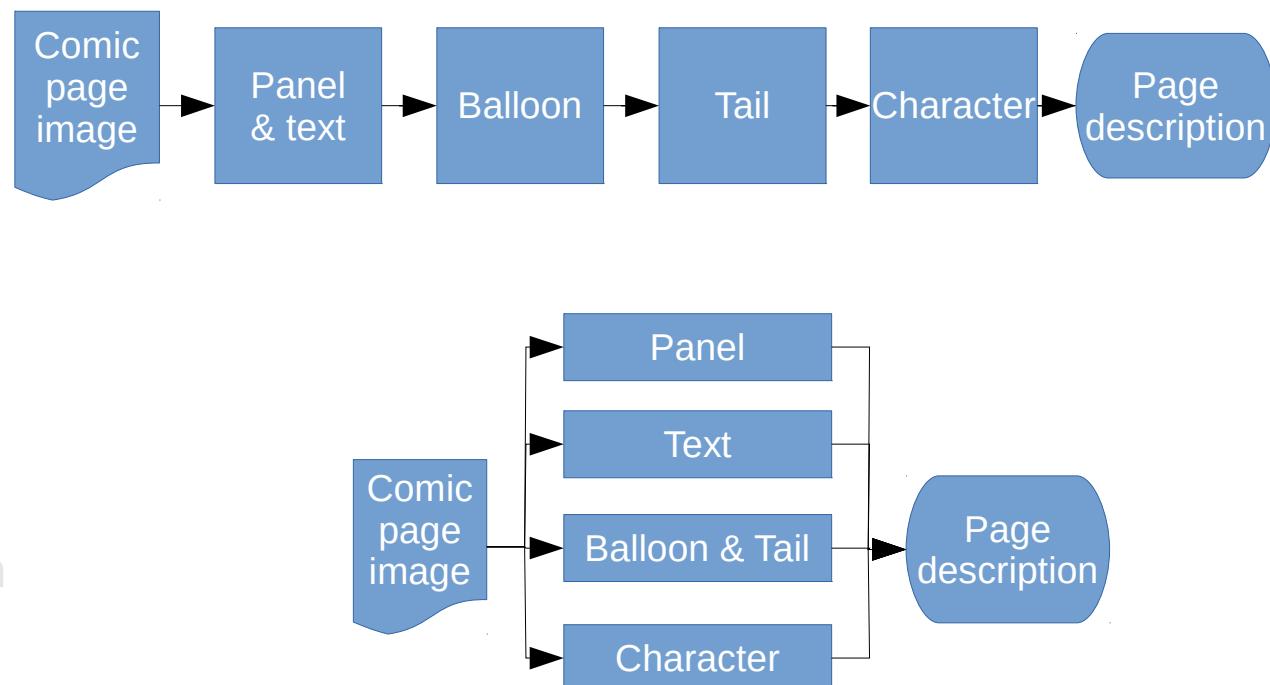
Contributions



Introduction

- Objective: cover the widest possible scope of study
 - 1) Creation of heterogeneous dataset
 - 2) Three approaches
- Content-driven
 - Sequential approach
 - Similar to literature
 - Intuitive
 - Sensible to error propagation
 - Independent approach
 - Avoid error propagation
- Knowledge-driven
 - Knowledge-driven approach
 - Based on domain knowledge
 - Retrieve context

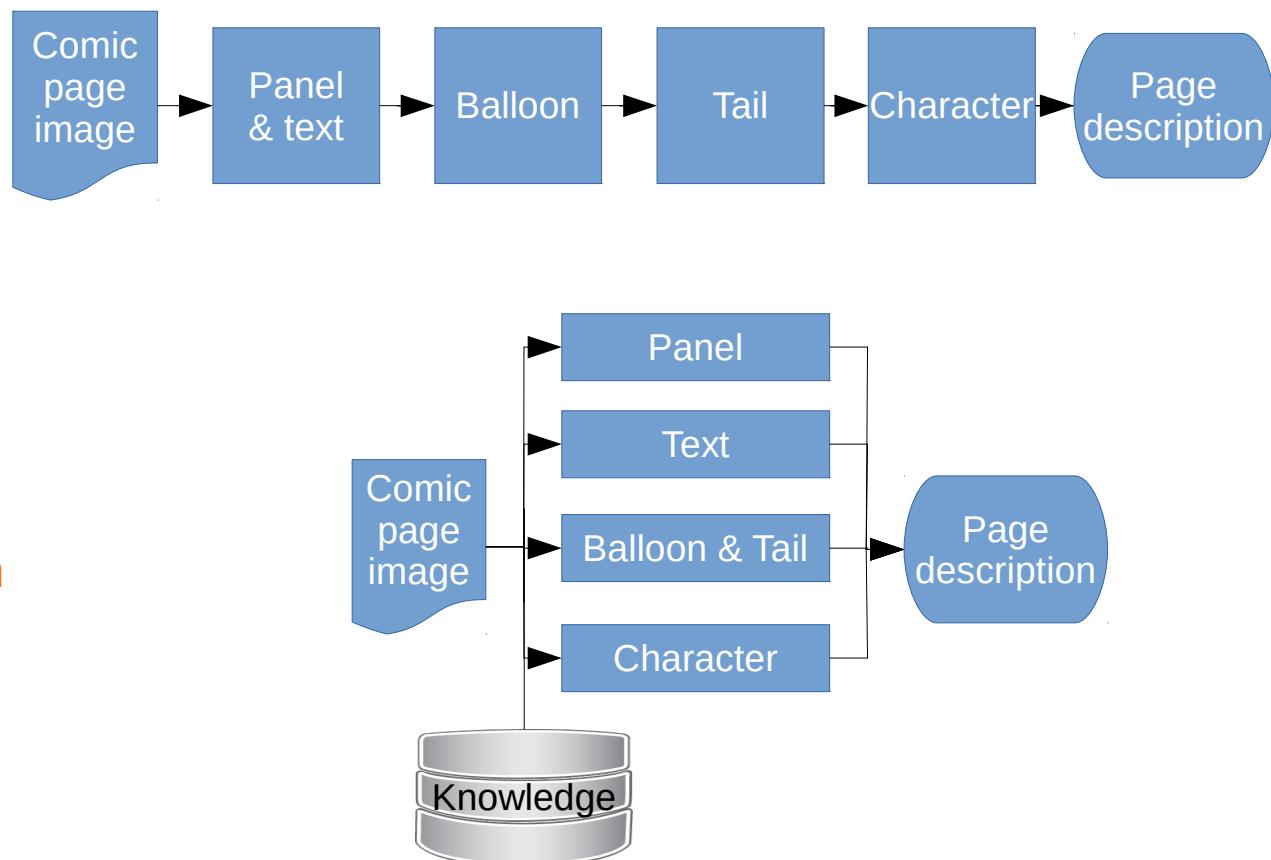
Contributions



Introduction

- Objective: cover the widest possible scope of study
 - 1) Creation of heterogeneous dataset
 - 2) Three approaches
- Content-driven
 - Sequential approach
 - Similar to literature
 - Intuitive
 - Sensible to error propagation
 - Independent approach
 - Avoid error propagation
- Knowledge-driven
 - Knowledge-driven approach
 - Based on domain knowledge
 - Retrieve context

Contributions



- Introduction
- Sequential approach
- (Independent approach)
- Knowledge-driven approach



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

- Introduction
- Sequential approach
 - Panel & text extraction
 - Balloon extraction
 - Tail extraction
 - Comic character extraction
- (Independent approach)
- Knowledge-driven approach



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

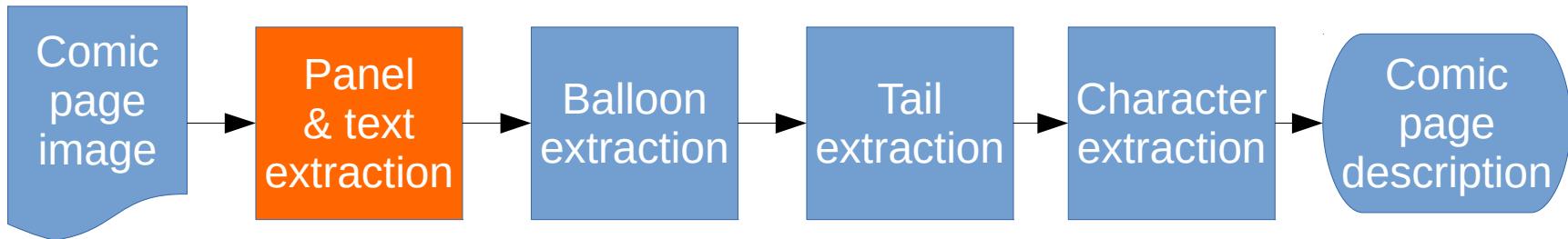
- Introduction
- Sequential approach
 - Panel & text extraction
 - Balloon extraction
 - Tail extraction
 - Comic character extraction
- (Independent approach)
- Knowledge-driven approach



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Panel & text extraction

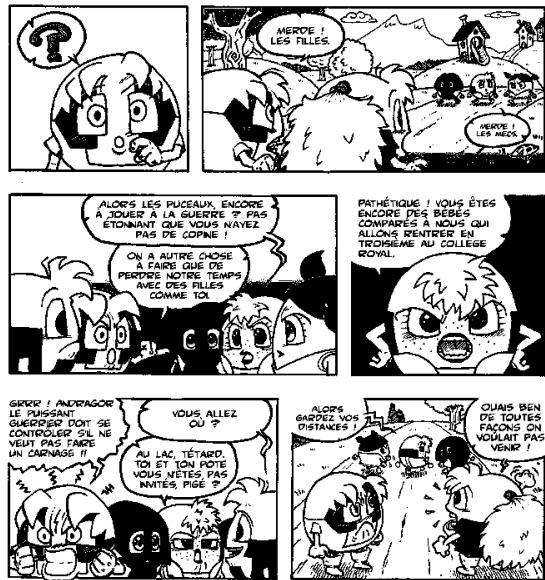
Contributions
Sequential approach



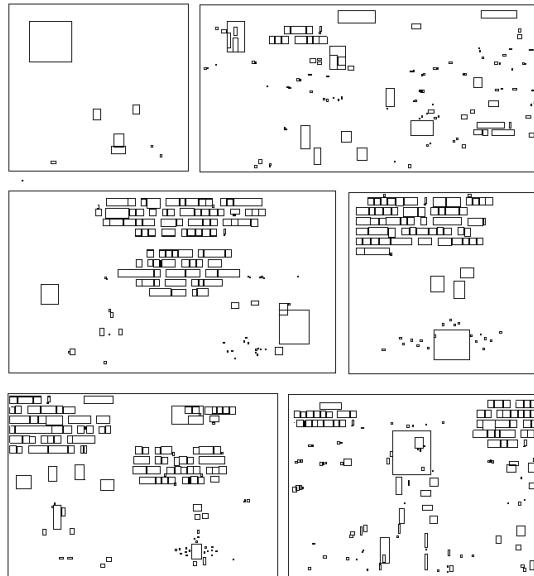
- Literature
 - Panel with frame, separated by **gutters** or black **line**
 - Text located **inside** balloons
- Contribution
 - **Simultaneous panel and text** extraction from binary image
 - Consider **implicit** and **non-rectangle** panels
 - **Location-independent** text extraction

Panel & text extraction

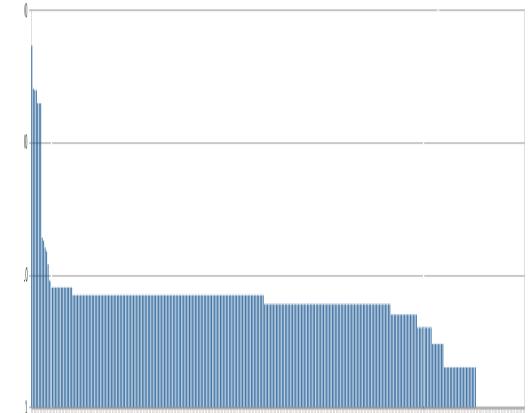
Contributions Sequential approach



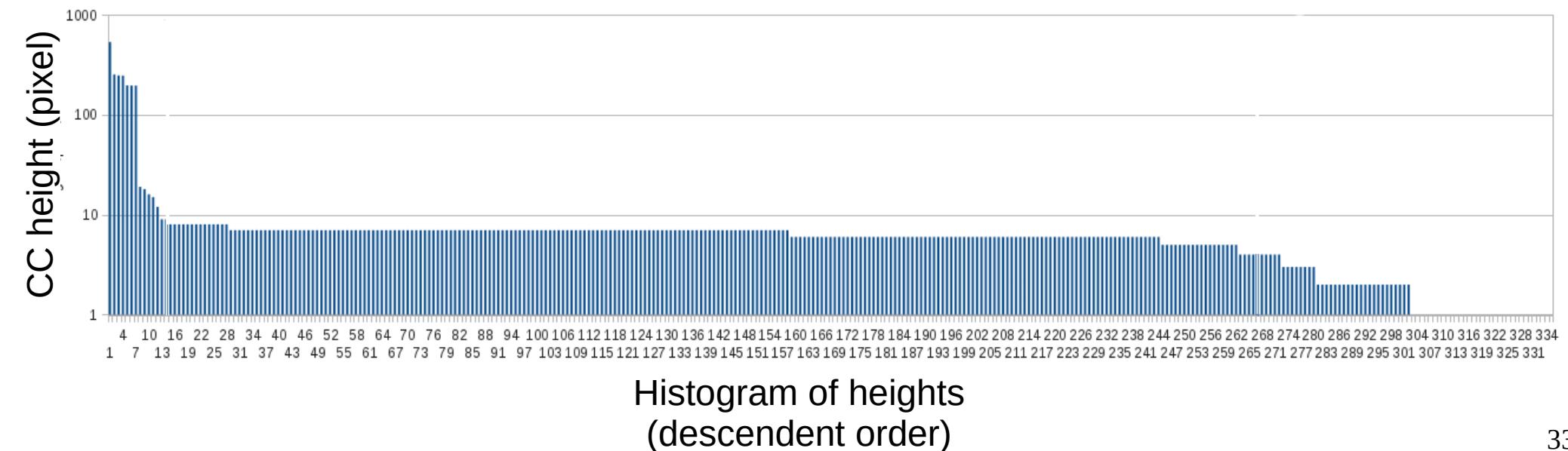
Binary image



Black connected-component (CC) bounding boxes



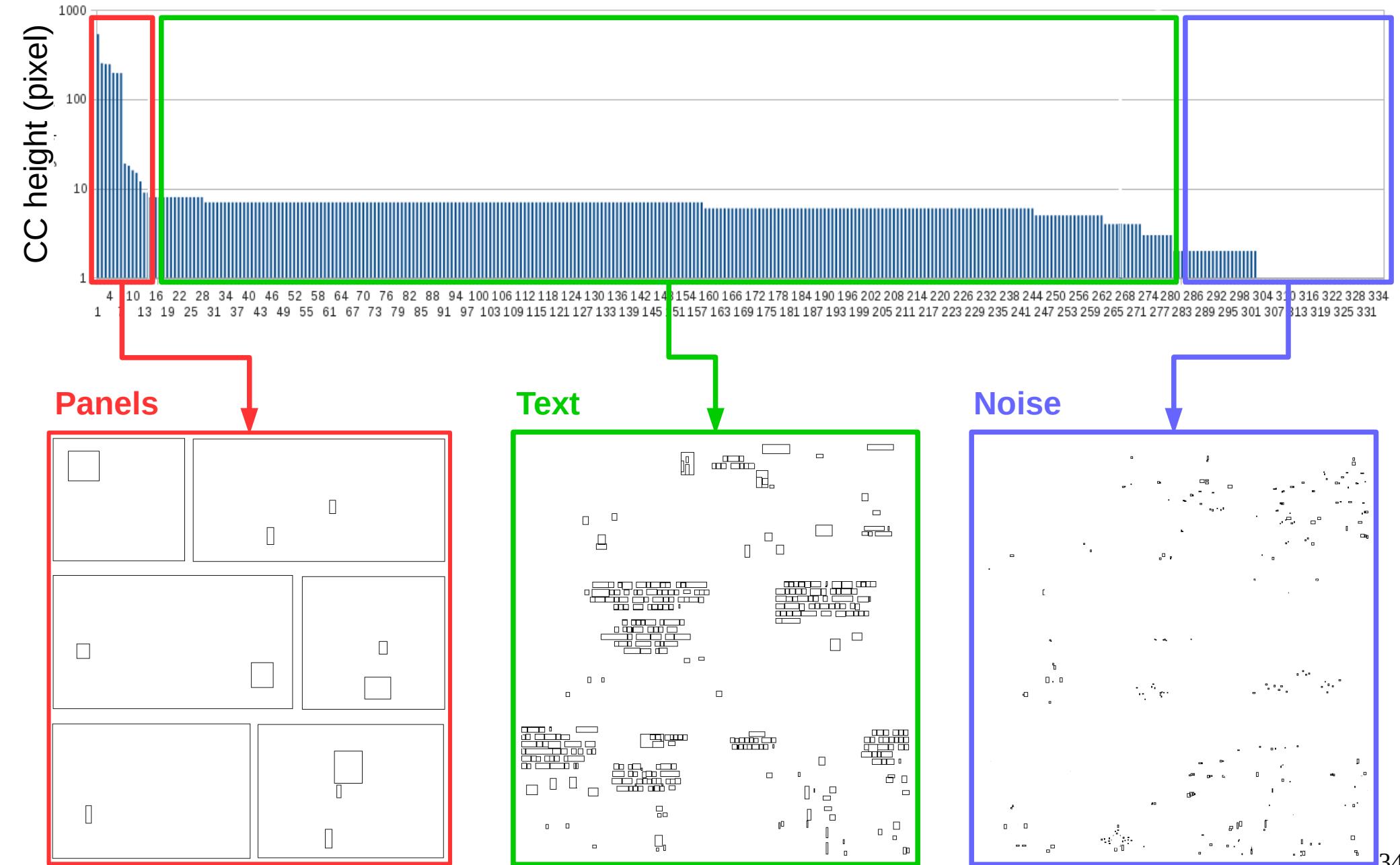
Histogram of heights of CC



Panel & text extraction

Contributions Sequential approach

K-means clustering ($k=3$)



Panel & text extraction: results

Contributions Sequential approach

W CE JOUR-LÀ, COURANT
SUR LA PLAGE IL NE SE
DOUTAIT PAS QUE SA
VIE ALLAIT BASCULER...

OK, POUR MOI, LA PRISE EST BONNE !
TU PEUX REMETTRE STEVE DANS NE
LA CHAMBRE DE LA SOEUR DE KID.
ON N'EN A PLUS BESOIN.

ONE EVENING LEANING OVER THE TAFFRAIL,
I OBSERVED A STRANGE CLOUD THAT STRUCK
A CHILL TO MY HEART.
N

... ?!
BATTRE DES
BRAS COMME LES
MOUETTES ?... TU
VEUX DIRE VOLER
COMME DES
MOUETTES, NON
?
SE

TODO

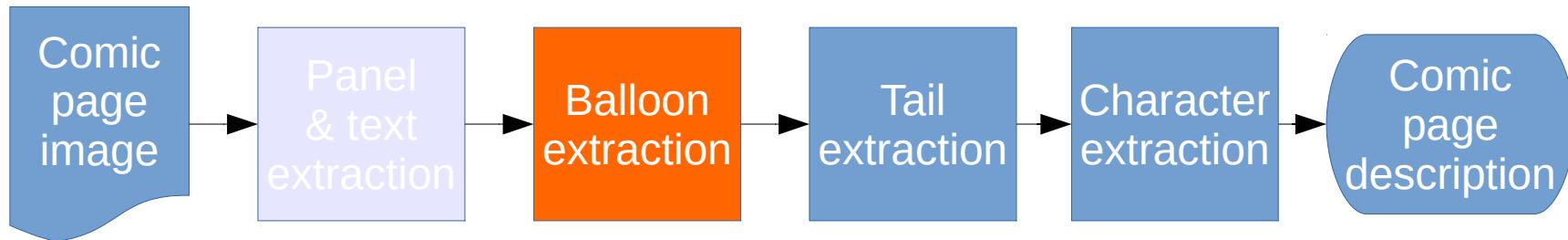
- Introduction
- Sequential approach
 - Panel & text extraction
 - **Balloon extraction**
 - Tail extraction
 - Comic character extraction
- (Independent approach)
- Knowledge-driven approach



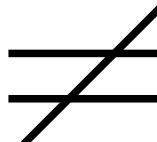
Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Balloon extraction

Contributions
Sequential approach



Regular balloon

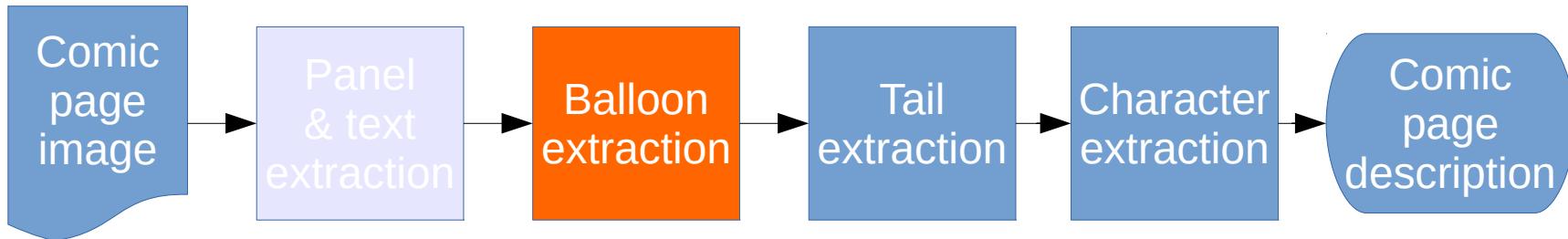


Implicit balloon

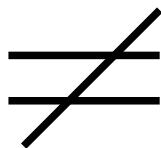
- Literature
 - Top-down approaches: extract white blobs and then text inside
 - Limited to regular balloons
- Contribution
 - Bottom-up approaches: extract text and then surrounding balloons
 - Appropriate for regular and implicit balloons

Balloon extraction

Contributions
Sequential approach



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
 - Improvement of regular and a first approach for implicit balloon extractions

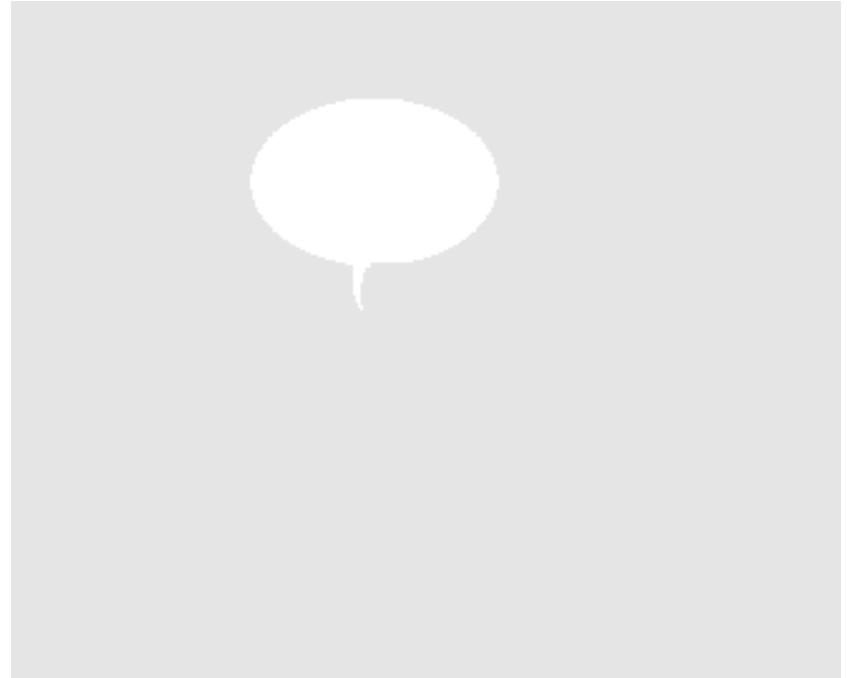
Balloon extraction: regular

Contributions
Sequential approach

- Assumptions
 - Panels and text block positions are known
 - Regular balloons contain centred text
- Proposition → structural analysis
 - Extract closed contours that fully include centred text



Original image



Expected result

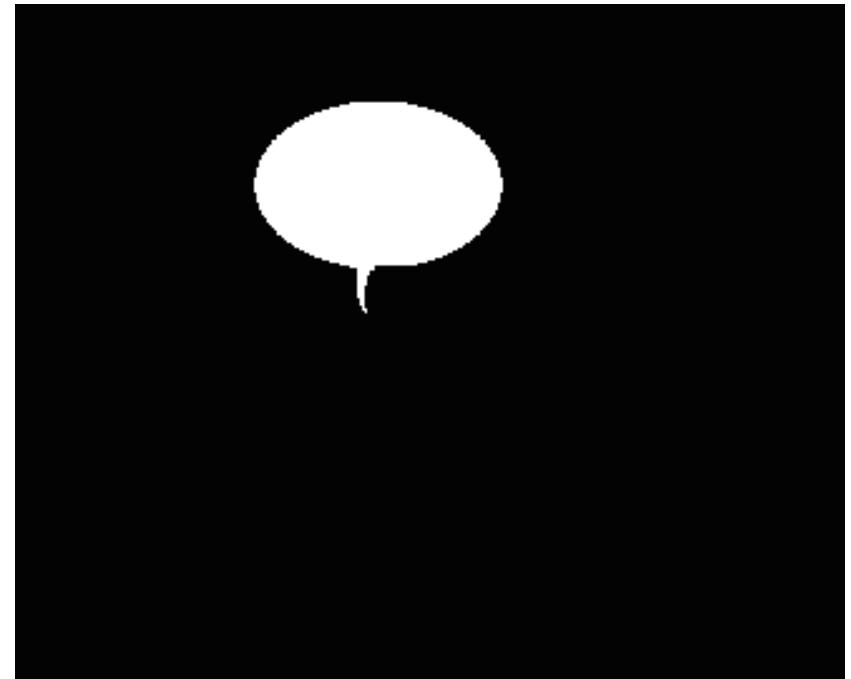
Balloon extraction: regular

Contributions
Sequential approach

- Assumptions
 - Panels and text block positions are known
 - Regular balloons contain centred text
- Proposition → structural analysis
 - Extract closed contours that fully include centred text



Original image



Expected result

Balloon extraction: regular

Contributions
Sequential approach



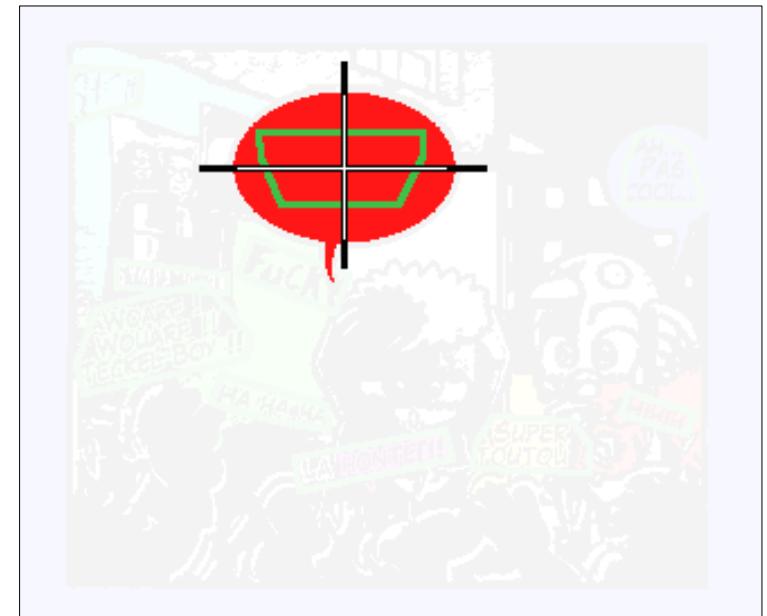
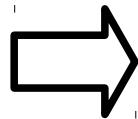
Original image



Text block positions (green)



Regions including text blocks (coloured)

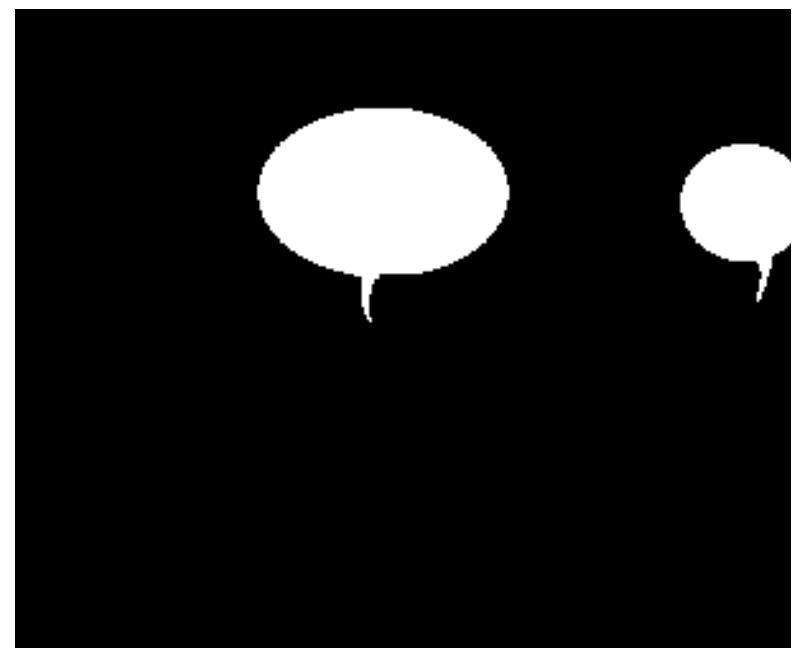
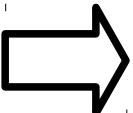


Regions including aligned text blocks

Balloon extraction: implicit

Contributions
Sequential approach

- Assumptions
 - Panel and text blocks positions are known
 - Implicit balloons contain centred text
- Proposition
 - Extract implicit balloons from text regions by inflating a deformable contour
 - Adaptation of active contour model (snake)

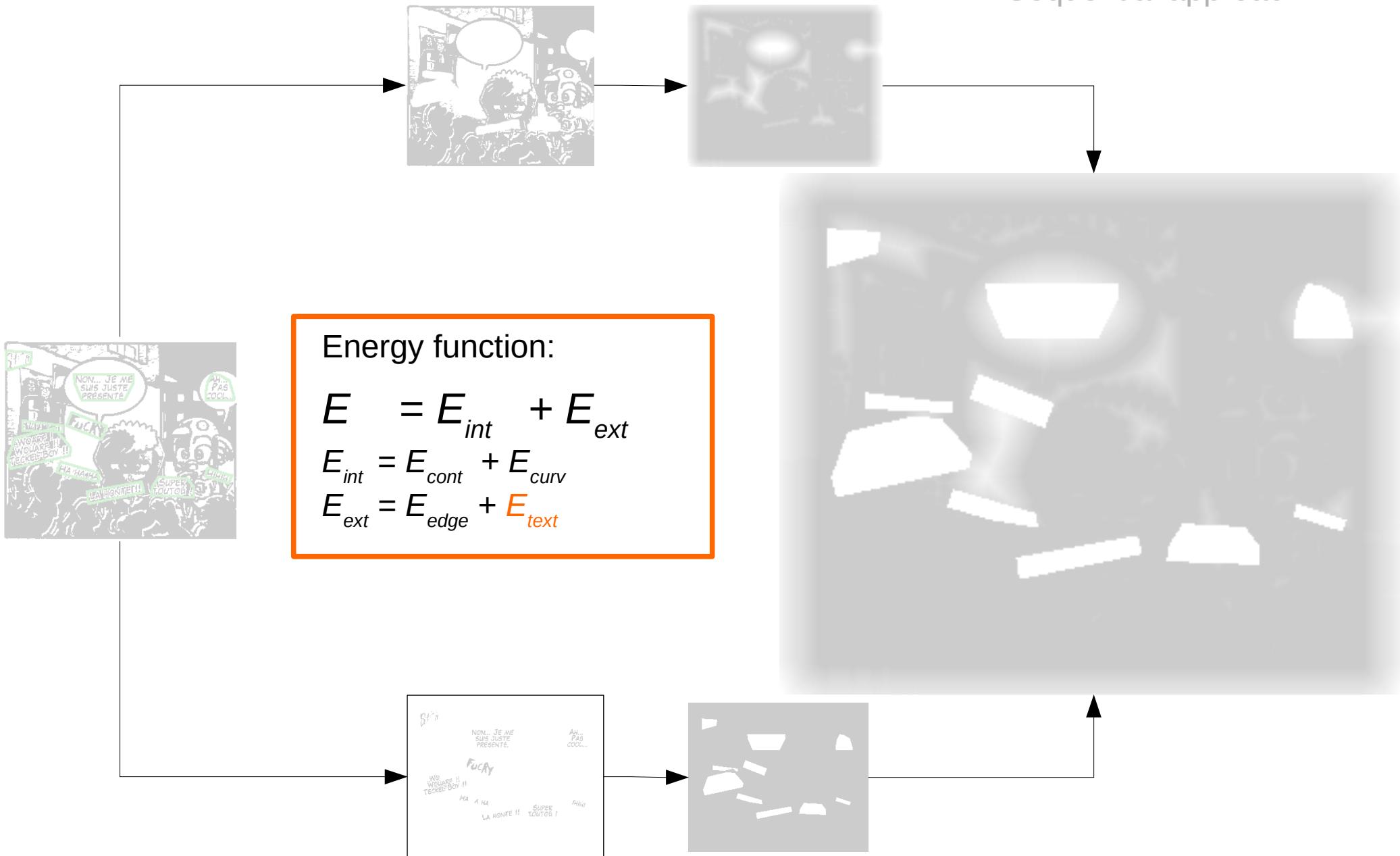


Original image and text locations

Expected result

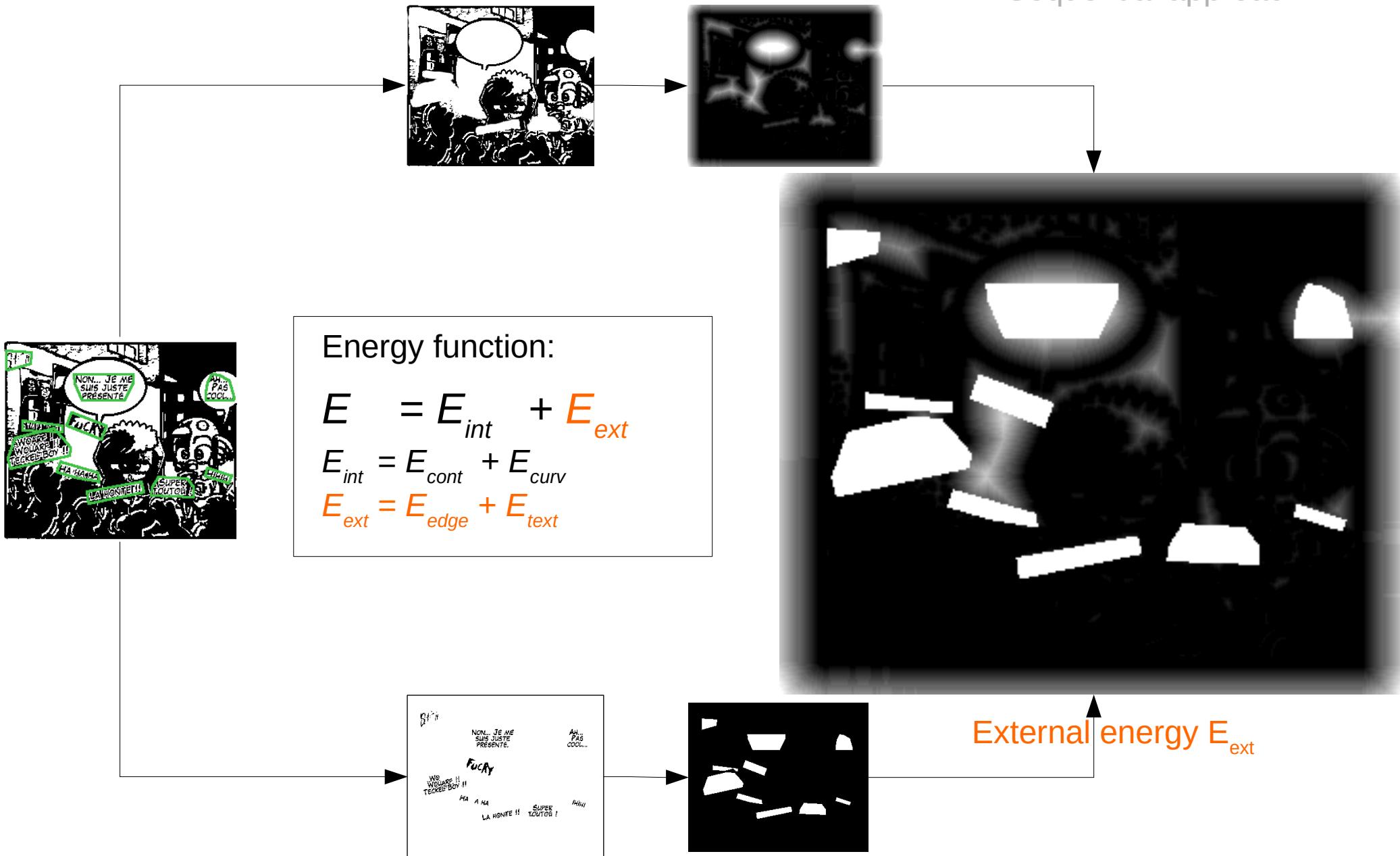
Balloon extraction: implicit

Contributions
Sequential approach



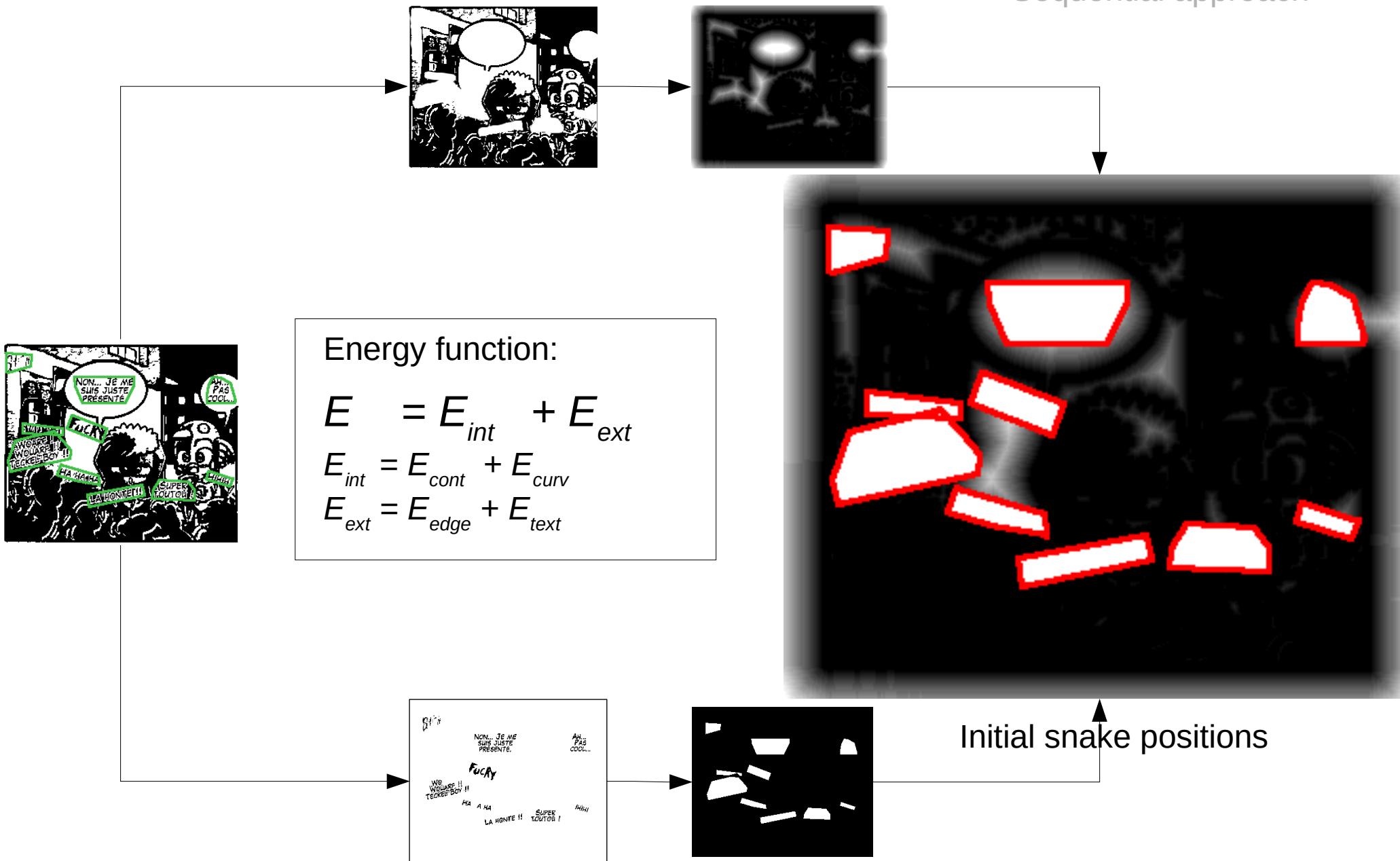
Balloon extraction: implicit

Contributions
Sequential approach



Balloon extraction: implicit

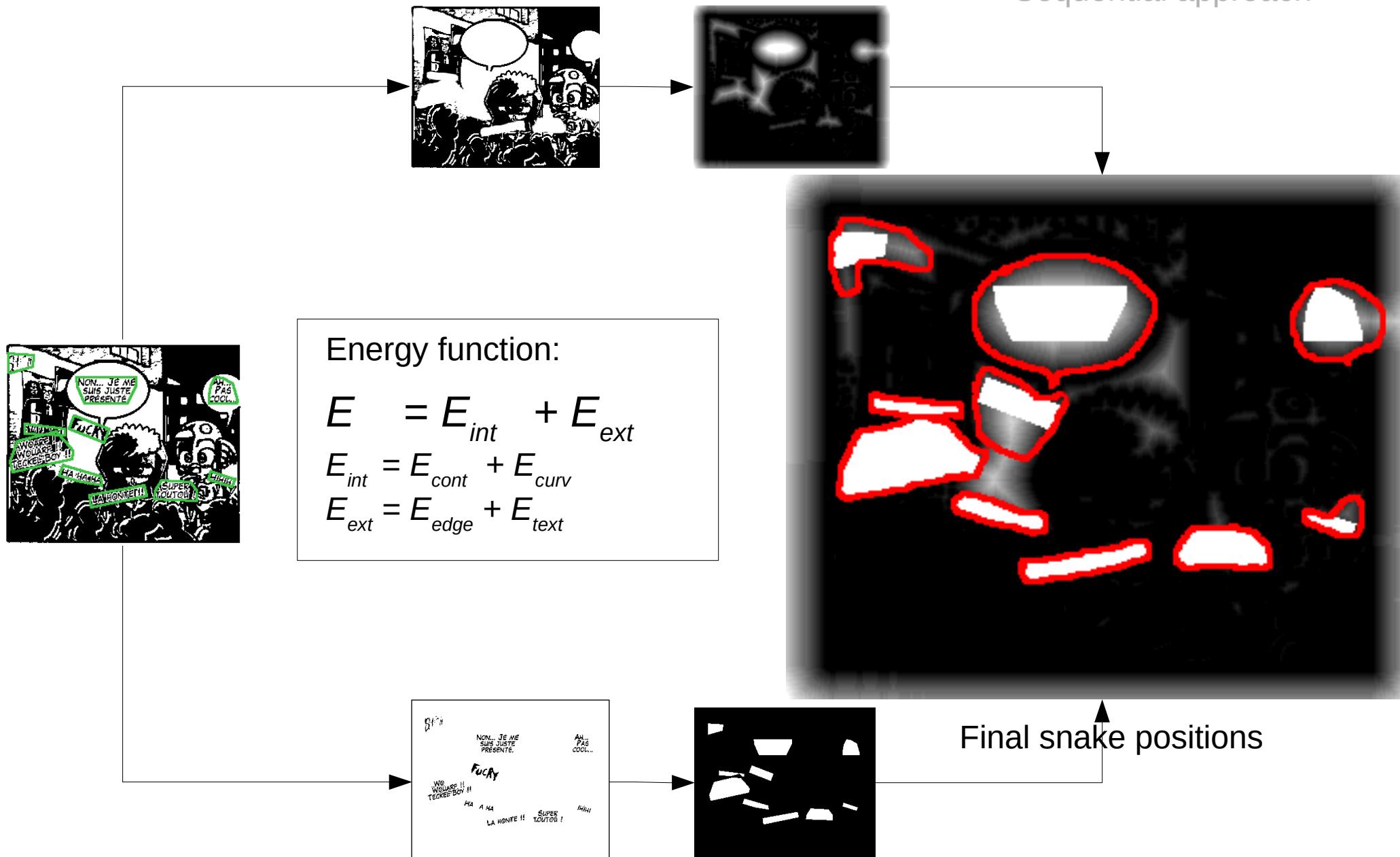
Contributions
Sequential approach



Initial snake positions

Balloon extraction: implicit

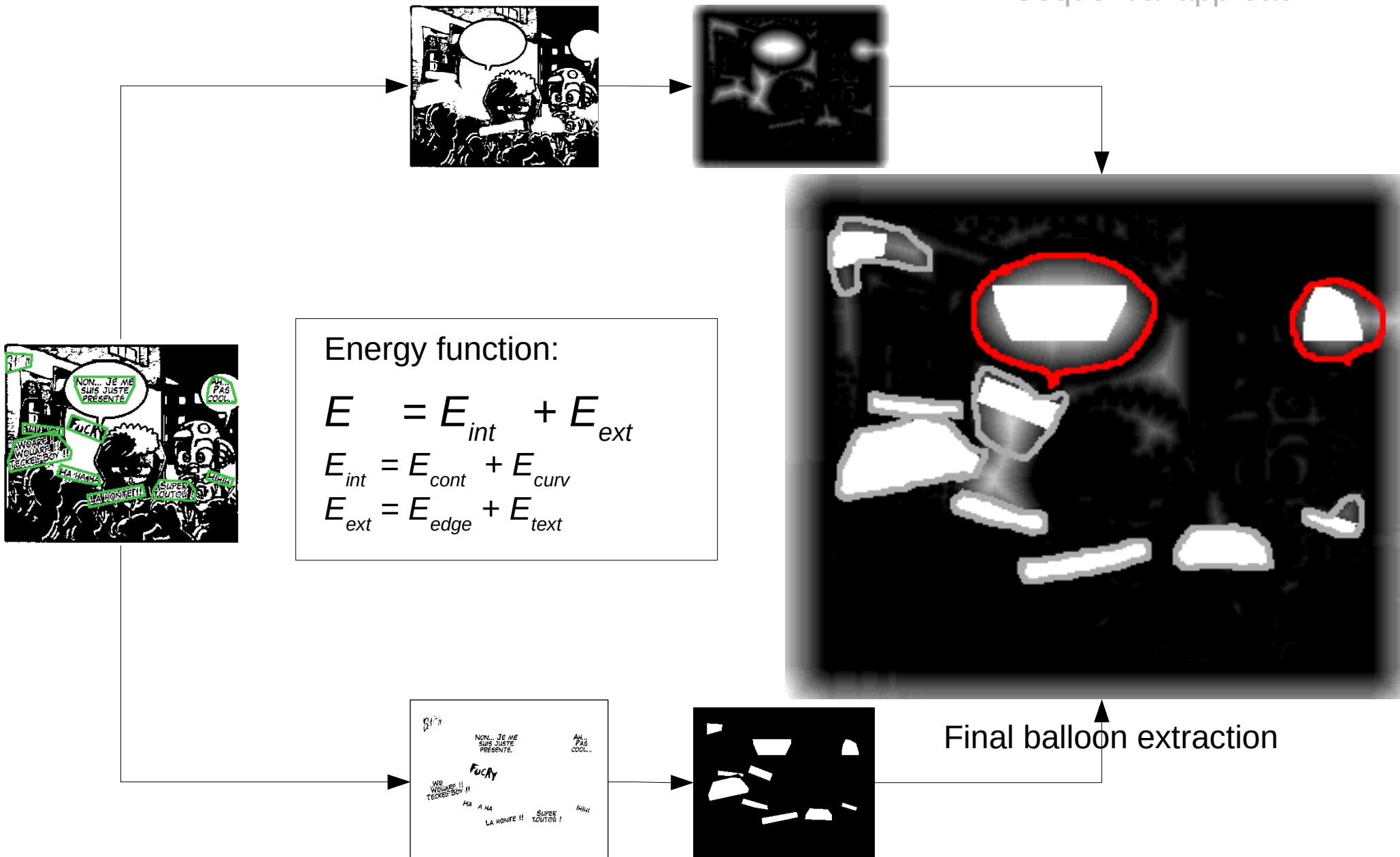
Contributions
Sequential approach



The snake is attracted to the “dark side”

Balloon extraction: implicit

Contributions
Sequential approach



The snake is attracted to the “dark side”

Balloon extraction: results

Contributions
Sequential approach

Regular/implicit balloon extraction
examples

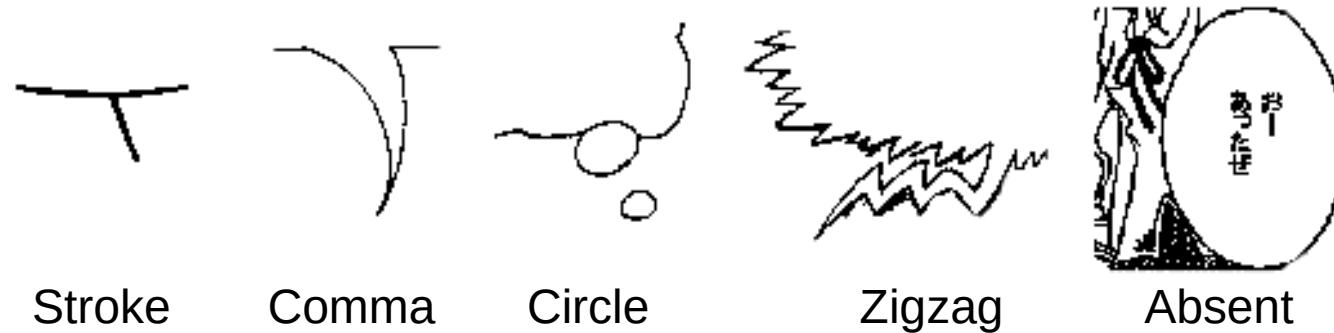
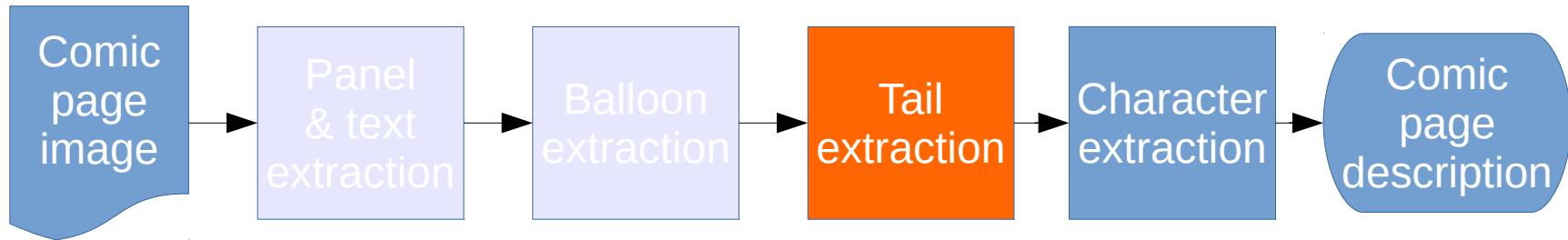
- Introduction
- Sequential approach
 - Panel & text extraction
 - Balloon extraction
 - Tail extraction
 - Comic character extraction
- (Independent approach)
- Knowledge-driven approach



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Tail extraction

Contributions
Sequential approach

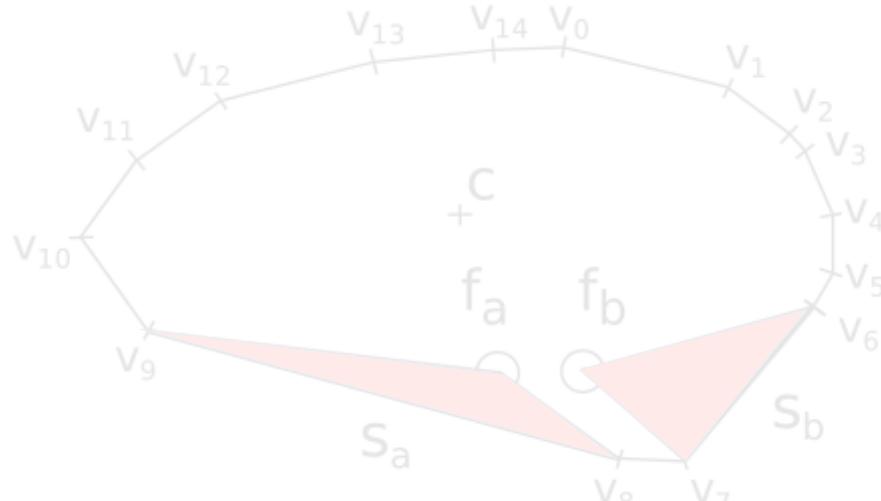


- Literature
 - First time studied in document image analysis
- Objectives
 - Extraction of tail tip position and direction
 - Focus on comma, zigzag and absent types

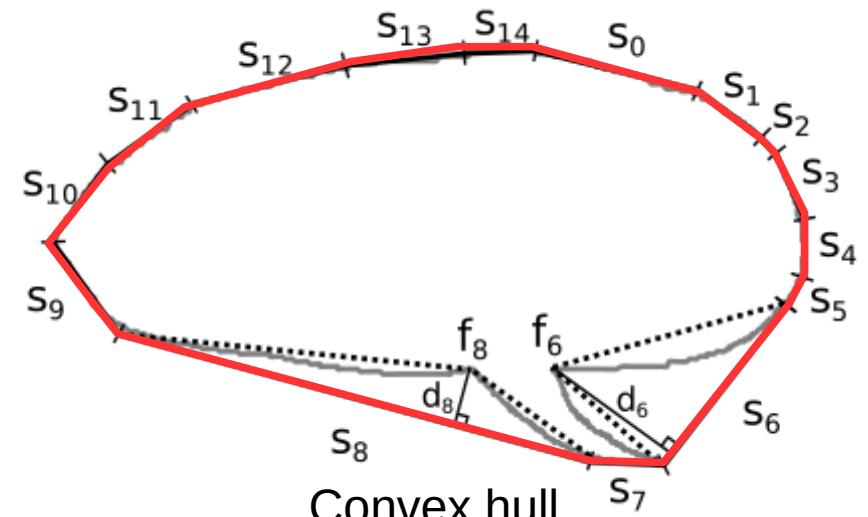
Tail extraction: tip position



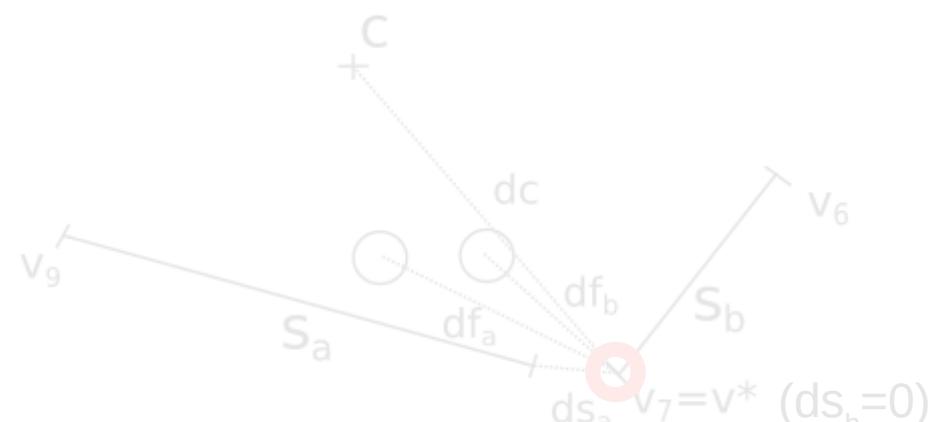
Balloon contour



Two biggest
convexity defects



Convex hull



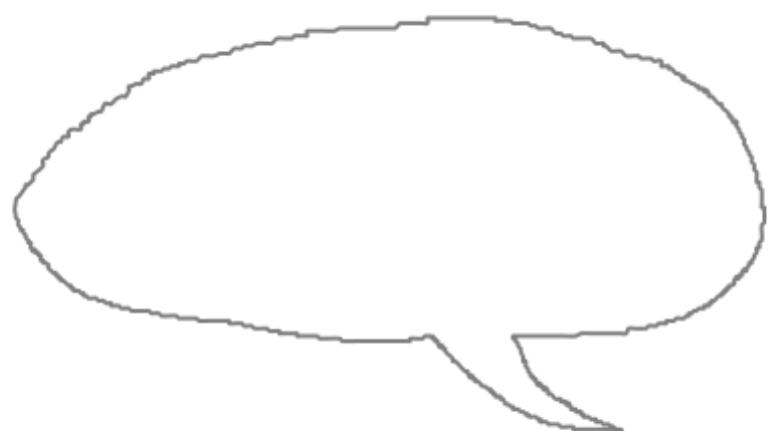
Tail tip position

Optimal vertex selection:

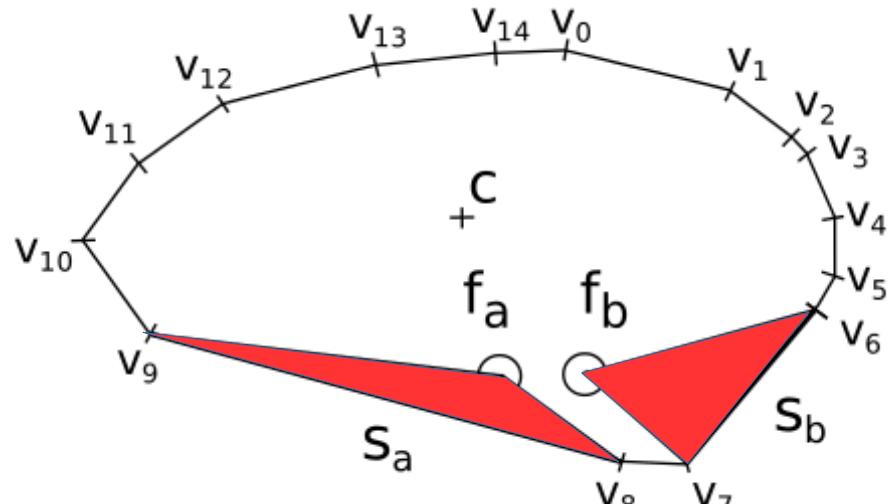
$$v^* = \operatorname{argmax}(\max(dc + df_a + df_b) + \min(ds_a + ds_b))$$

Tail extraction: tip position

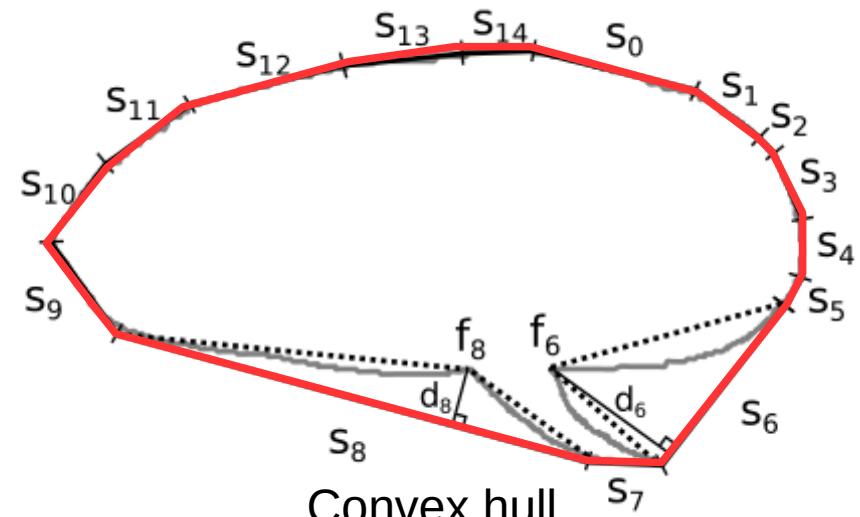
Contributions
Sequential approach



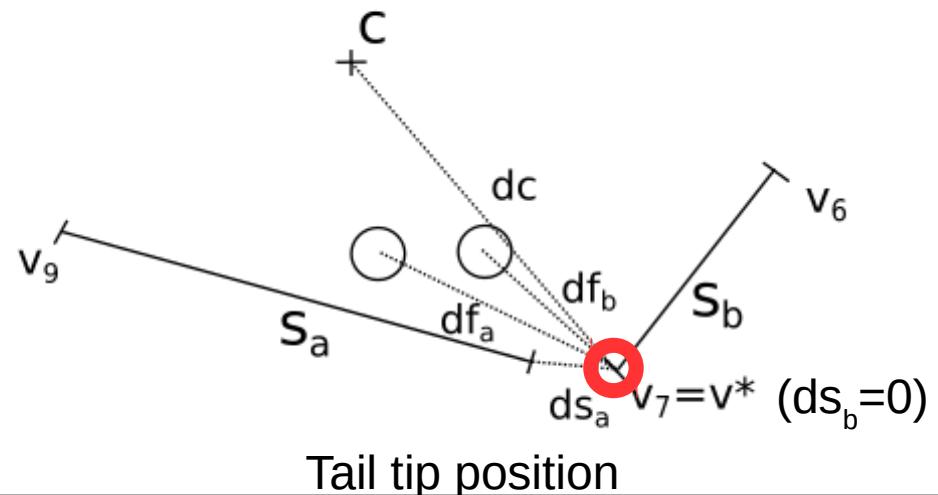
Balloon contour



Two biggest
convexity defects



Convex hull



Tail tip position

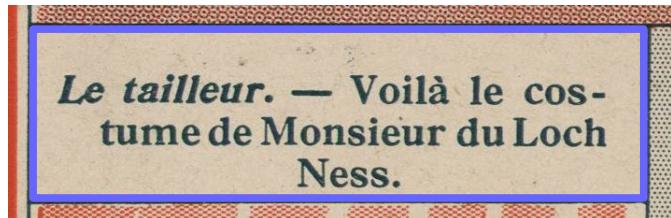
Optimal vertex selection:

$$v^* = \operatorname{argmax}(\max(dc + df_a + df_b) + \min(ds_a + ds_b))$$

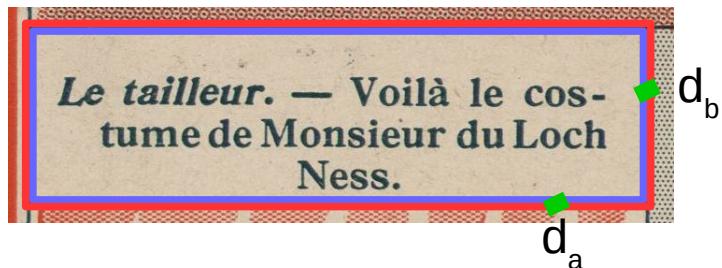
Tail extraction: confidence value

Balloon
contour (blue)

Balloon 1



Convex hull
(red)



Confidence

$$C_{tail} = \frac{(d_a + d_b)/2}{meanBalloonSize}$$

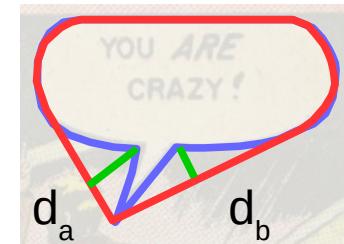
$$C_{tail} = 0.0$$

Presence of tail

NO

Contributions
Sequential approach

Balloon 2



$$C_{tail} = 0.73$$

YES (>0)

Tail extraction: tail direction

Contributions
Sequential approach

- Definition
 - Vector starting from “background” to “external edge” tail tip positions
- Approach
 - Extract **external edge**
 - Find **external edge tail tip coordinates**
 - Define the **tail direction** (N, NE, E, SE, S, SW, W, NW)



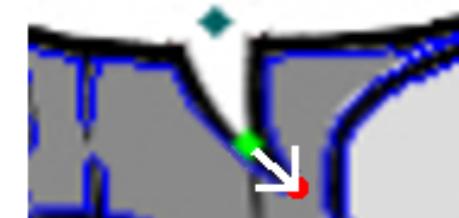
Background tail tip
(green) and
external edge (blue)



Closest point on
external edge
(red)



Farthest point
from origin and tip
(red)



Direction from tip
to farthest point
(white arrow)

Tail extraction: results

Experiments

W CE JOUR-LÀ, COURANT
SUR LA PLAGE IL NE SE
DOUTAIT PAS QUE SA
VIE ALLAIT BASCULER...

ONE EVENING LEANING OVER THE TAFFRAIL,
I OBSERVED A STRANGE CLOUD THAT STRUCK
A CHILL TO MY HEART.

OK, POUR MOI, LA PRISE EST BONNE !
TU PEUX REMETTRE STEVE DANS NE
LA CHAMBRE DE LA SOEUR DE KID.
ON N'EN A PLUS BESOIN.

... ?!
BATTRE DES
BRAS COMME LES
MOUETTES ?... TU
VEUX DIRE VOLER
COMME DES
MOUETTES, NON
?
SE

TODO

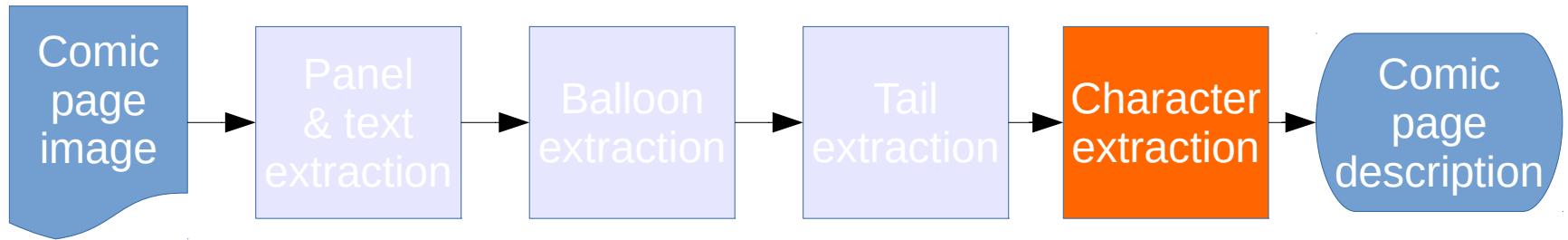
- Introduction
- Sequential approach
 - Panel & text extraction
 - Balloon extraction
 - Tail extraction
 - Comic character extraction
- (Independent approach)
- Knowledge-driven approach



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Comic character extraction

Contributions
Sequential approach

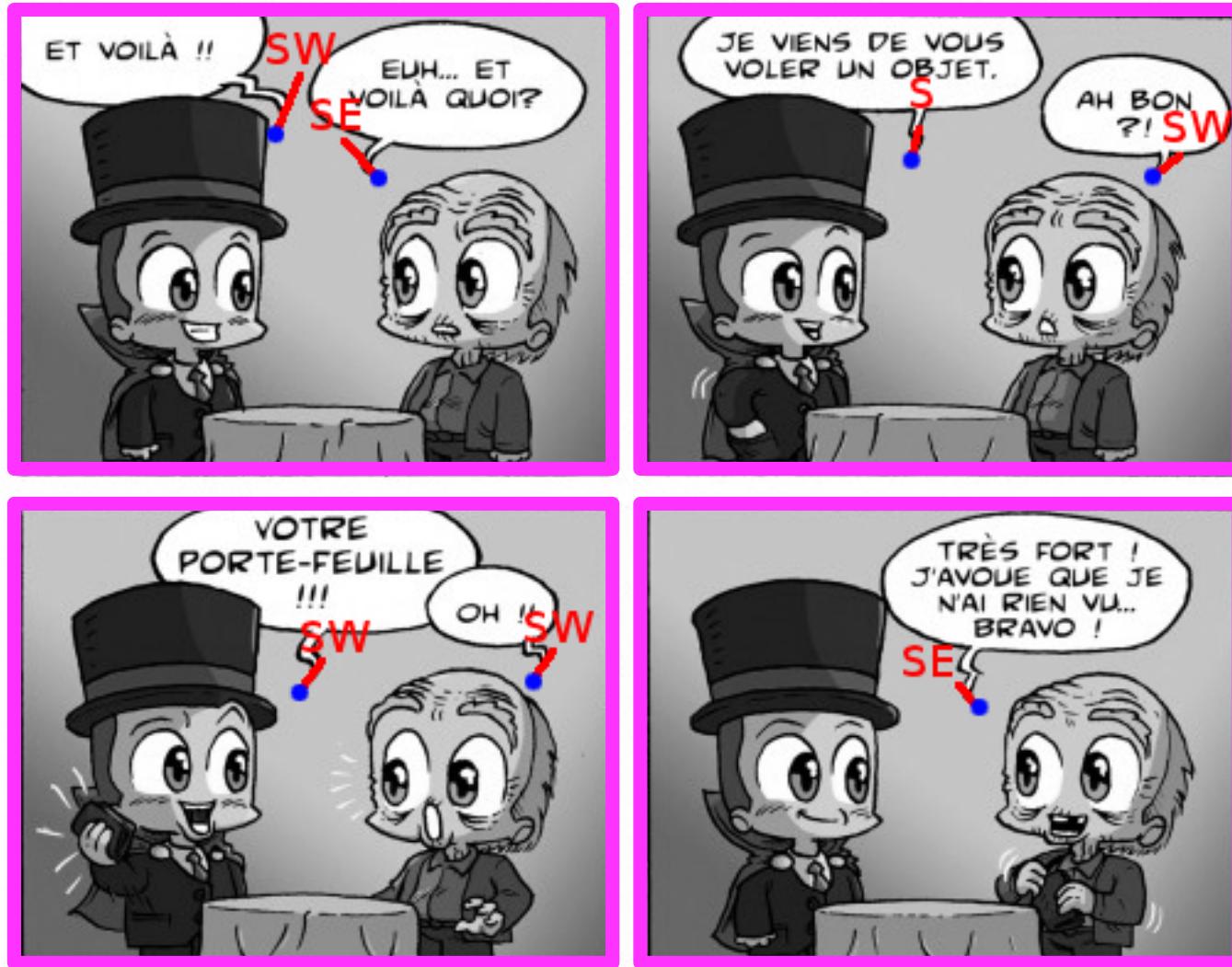


- Literature
 - Supervised approaches for **manga** and **cartoon** characters
 - No public **dataset** (copyright issues)
- Challenges
 - Variety of **styles** of comic books
 - **Intra** and **extra** class **variations** of each character **instance** (e.g. position, scale, pose, occlusion and human-like, invented)
- Objective
 - **Unsupervised** and **generic** approach for **all styles** of comic books

Comic character extraction

Contributions
Sequential approach

Panels + Tails = ?



Comic character extraction

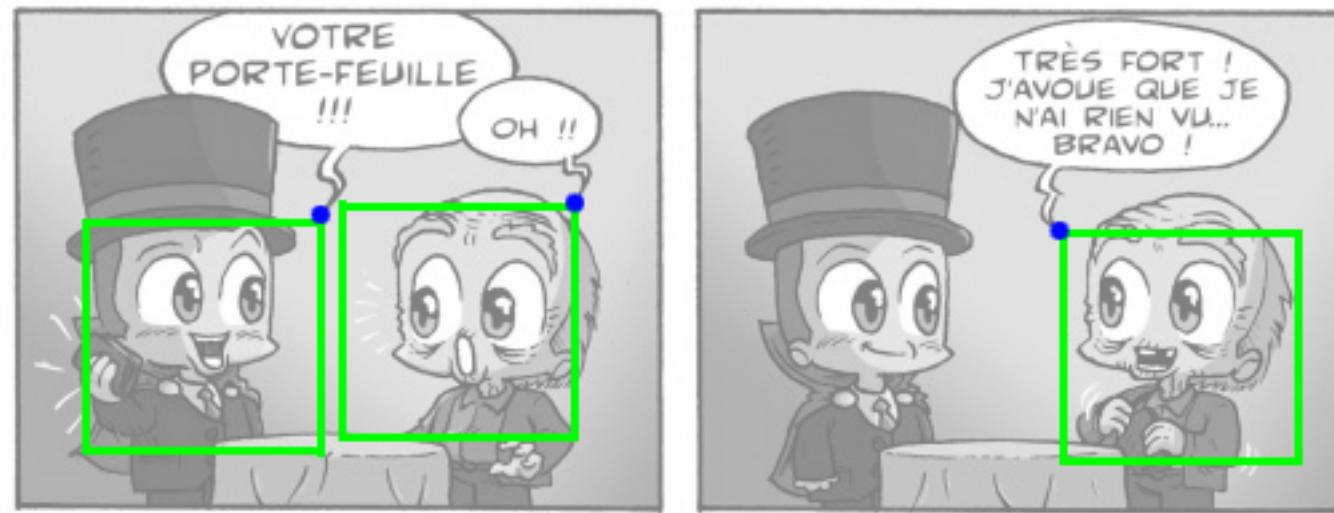
Contributions
Sequential approach

Panels + Tails = Comic character ROIs

Large ROI



Small ROI



Comic character extraction: results

Contributions
Sequential approach

ADD results here

- Introduction
- Sequential approach
- (Independent approach)
- Knowledge-driven approach
 - Introduction
 - Knowledge representation
 - Processing sequence



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Introduction

- High level image description
- Framework for comics understanding
- Independent element extraction
- Increase overall precision
- Collaboration with Clément Guérin

Contributions

Knowledge-driven approach

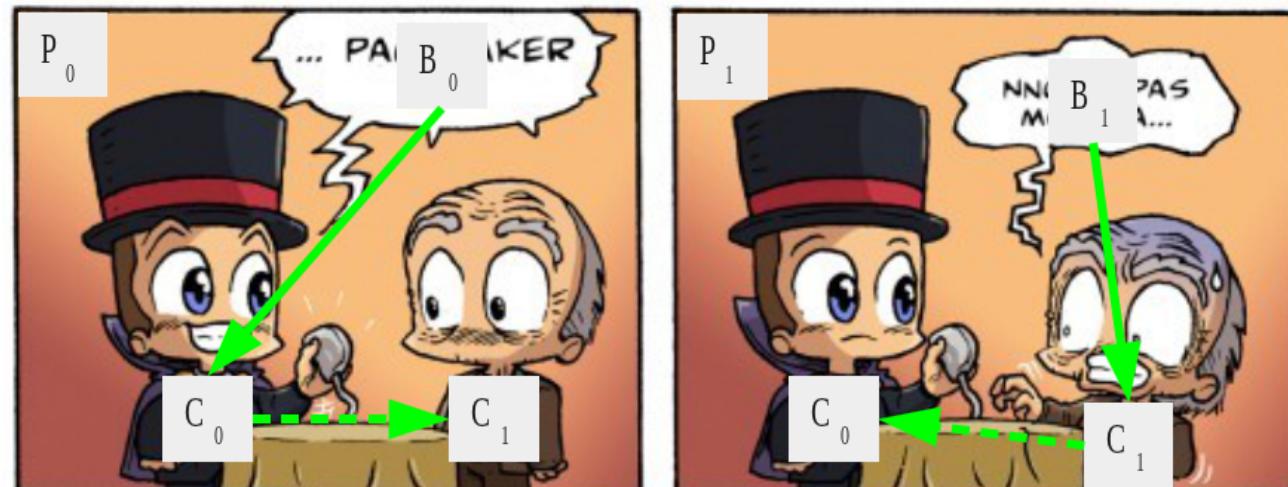
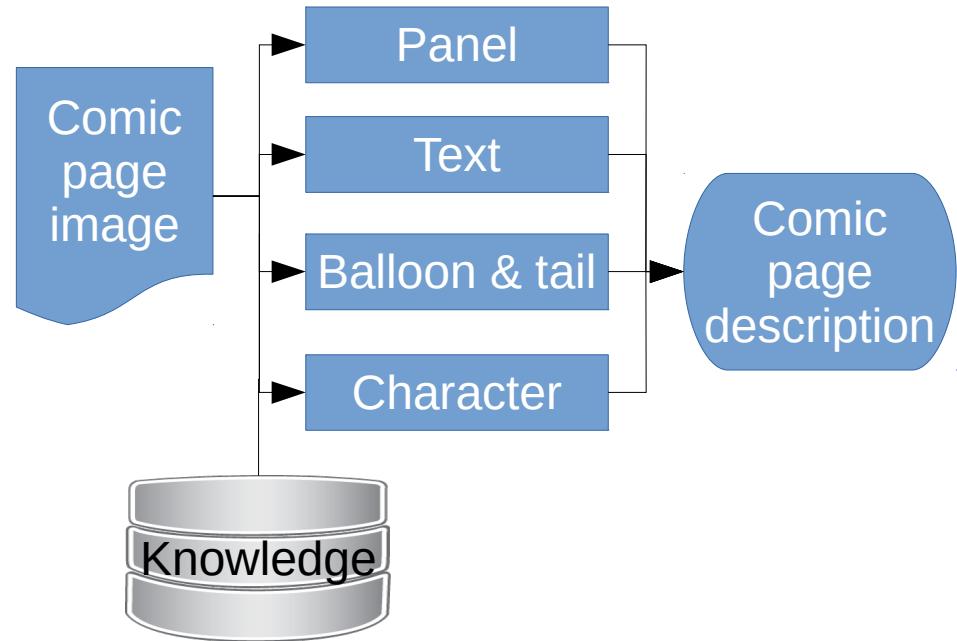
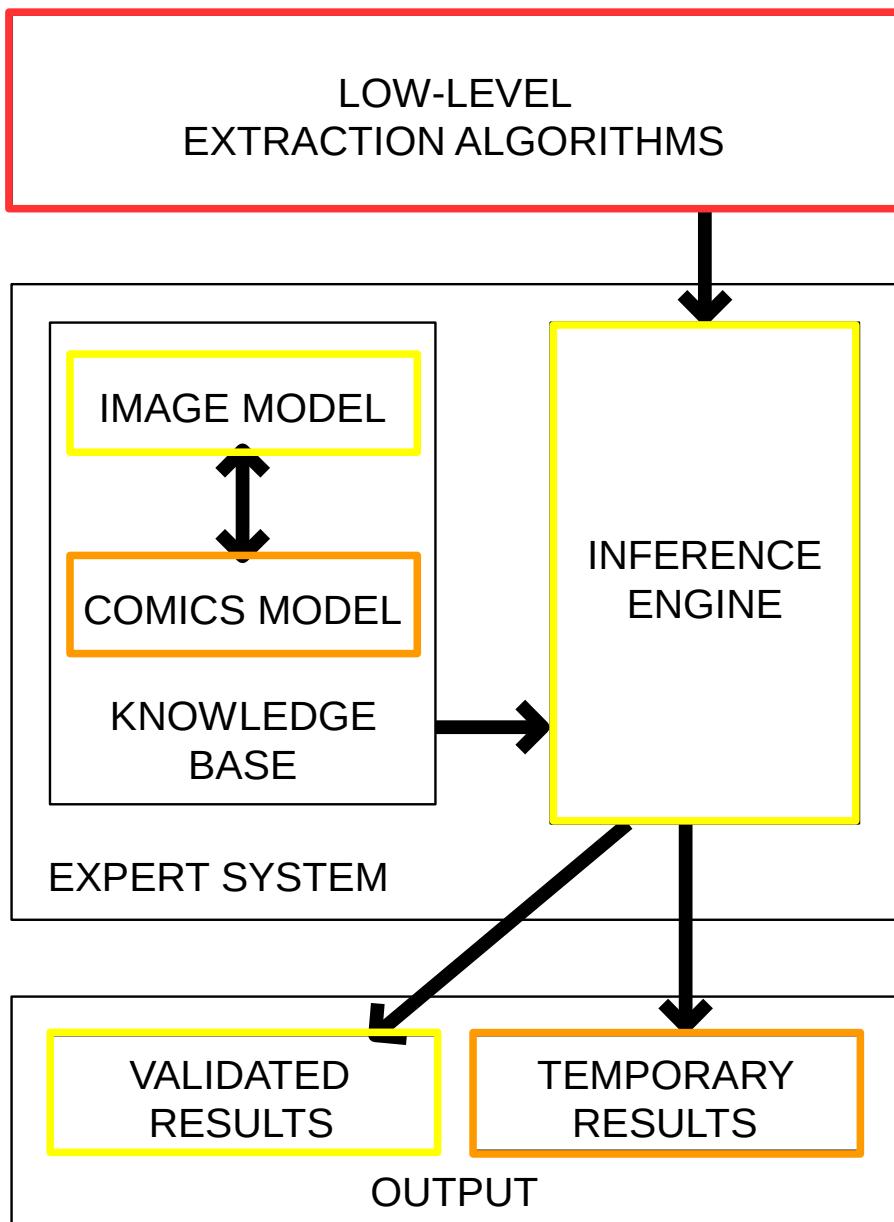


Illustration of high level description

Knowledge representation



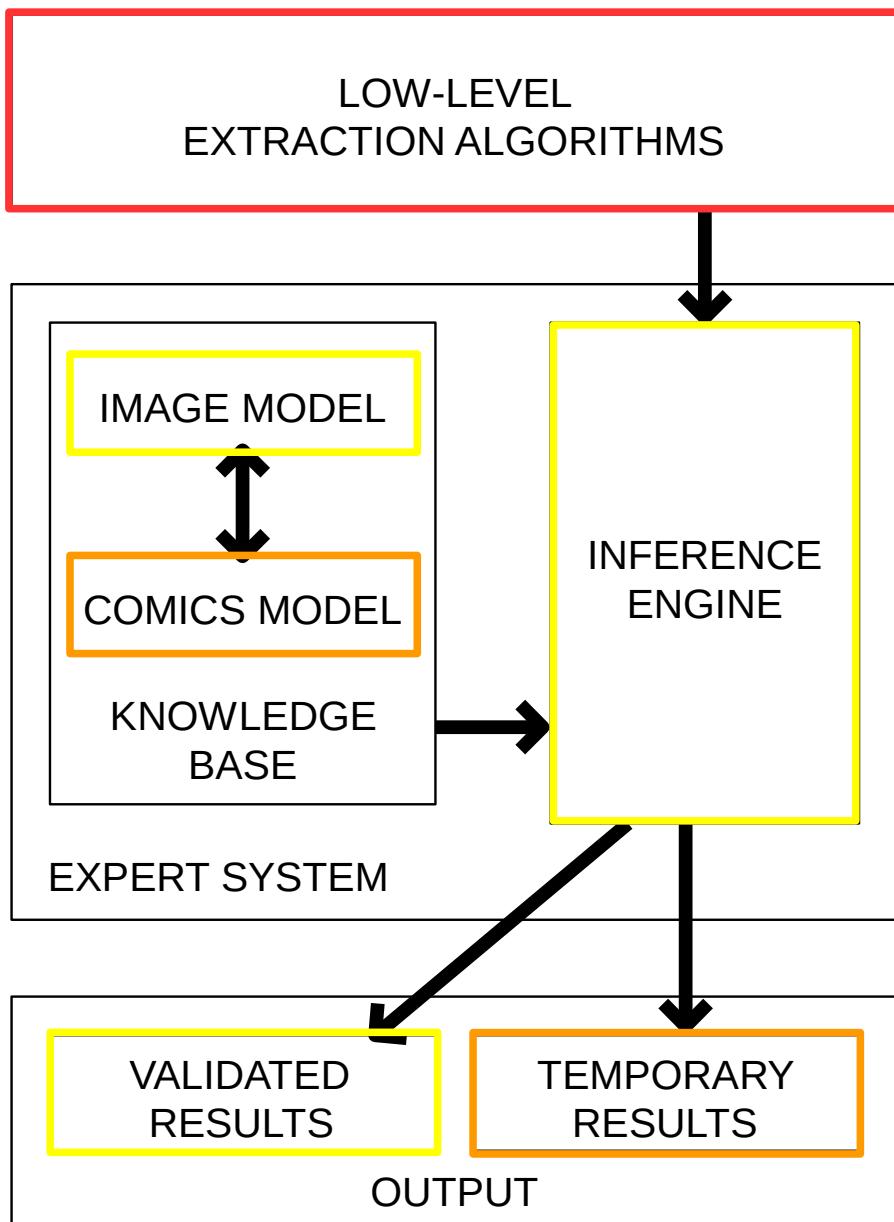
Contributions

Knowledge-driven approach

- Rigaud's thesis
- Collaboration
- Guérin's thesis

- Image model
 - Physical support
 - Regions of interest
- Comics model
 - Validations
 - A panel **P** is related to one page
 - A balloon **B** is related to one panel and may have a tail **Q**
 - A character **C** is related to one panel
 - A text line **T** is related to one balloon
 - Inferences
 - **B + Q + T => speech balloon SB**
 - **SB + T => speech text ST**
 - **SB + C => speaking character SC**

Knowledge representation



Contributions

Knowledge-driven approach

- Rigaud's thesis
- Collaboration
- Guérin's thesis

- Image model
 - Physical support
 - Regions of interest
- Comics model
 - Validations
 - A panel **P** is related to one page
 - A balloon **B** is related to one panel and may have a tail **Q**
 - A character **C** is related to one panel
 - A text line **T** is related to one balloon
 - Inferences
 - **B + Q + T => speech balloon SB**
 - **SB + T => speech text ST**
 - **SB + C => speaking character SC**

Processing sequence

Contributions

Knowledge-driven approach

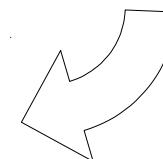


Formulate hypotheses



Validate hypotheses

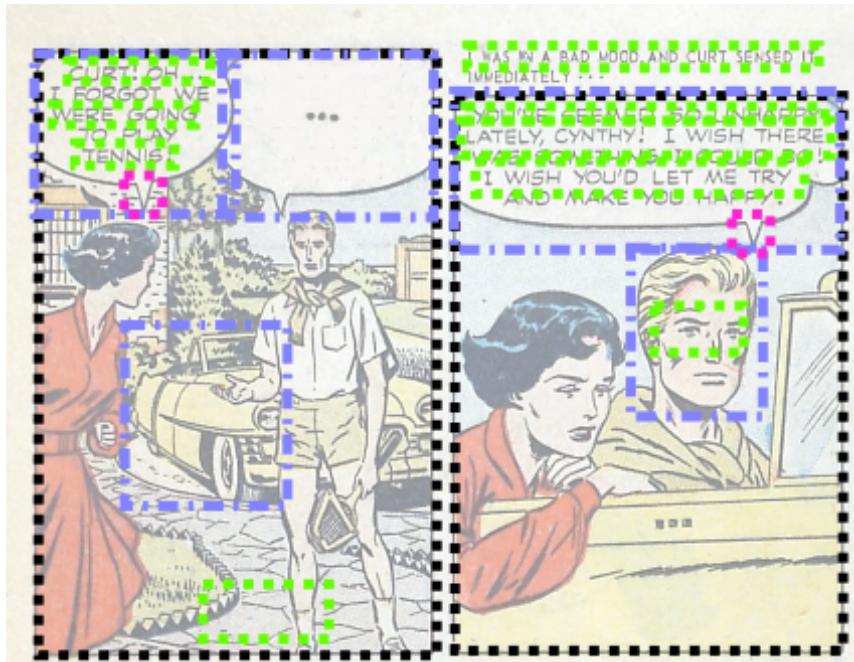
Infer new information



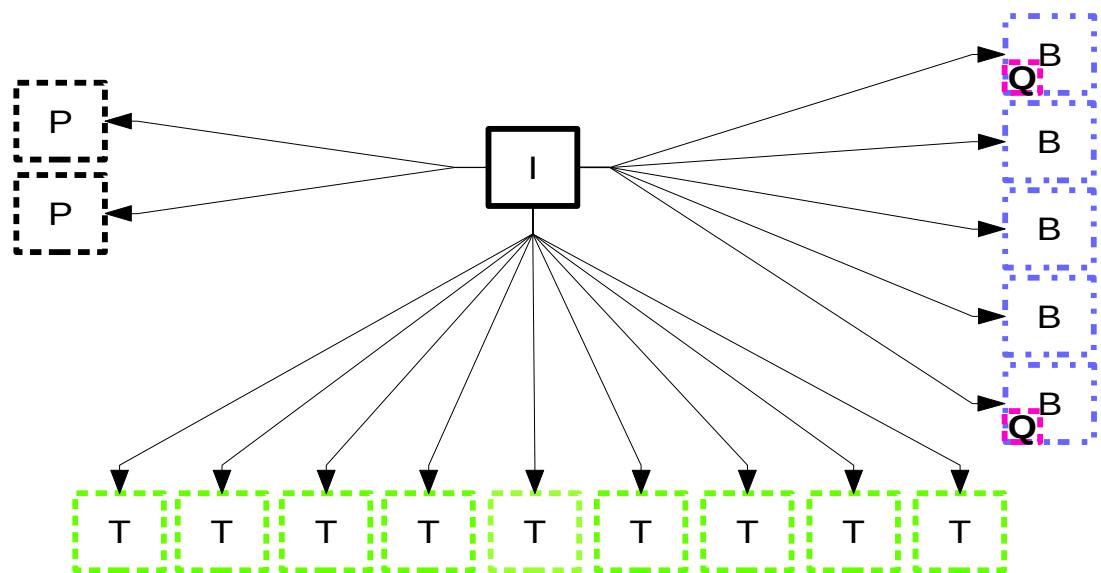
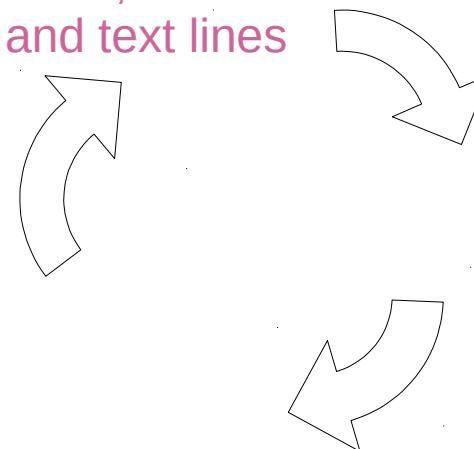
- Iteration 1
 - Step 1: hypotheses of **simple element** positions
 - Step 2: validation of the positions
 - Step 3: inference a new information
- Iteration 2
 - Step 1: hypotheses of more **complex elements**
 - Step 2: validation of the positions
 - Step 3: inference a new information
 - ...

Processing sequence

Contributions
Knowledge-driven approach

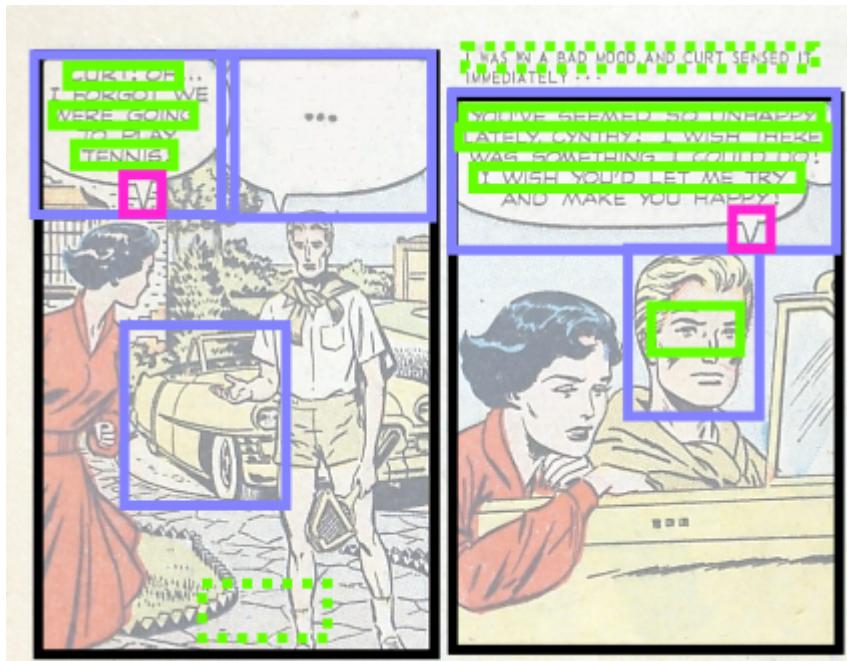


Hypotheses of
panels, balloons
and text lines

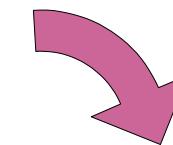
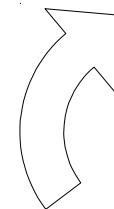


Processing sequence

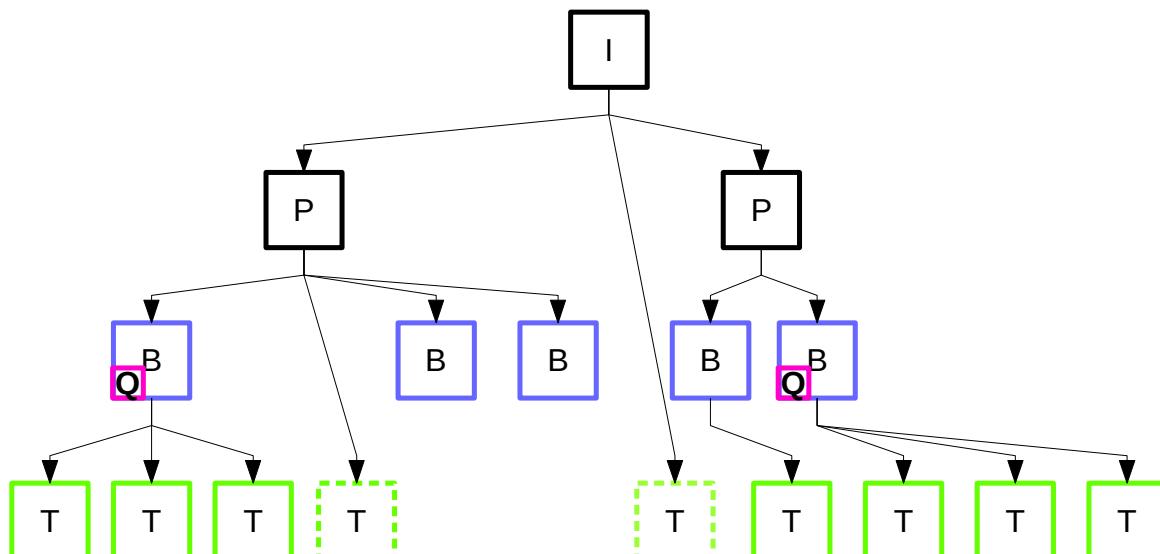
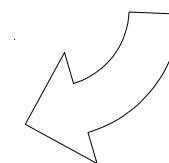
Contributions
Knowledge-driven approach



Hypotheses of
panels, balloons
and text lines



Validation of the
hypotheses

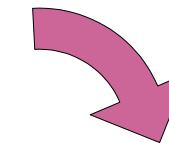
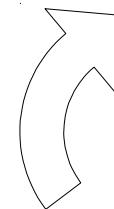


Processing sequence



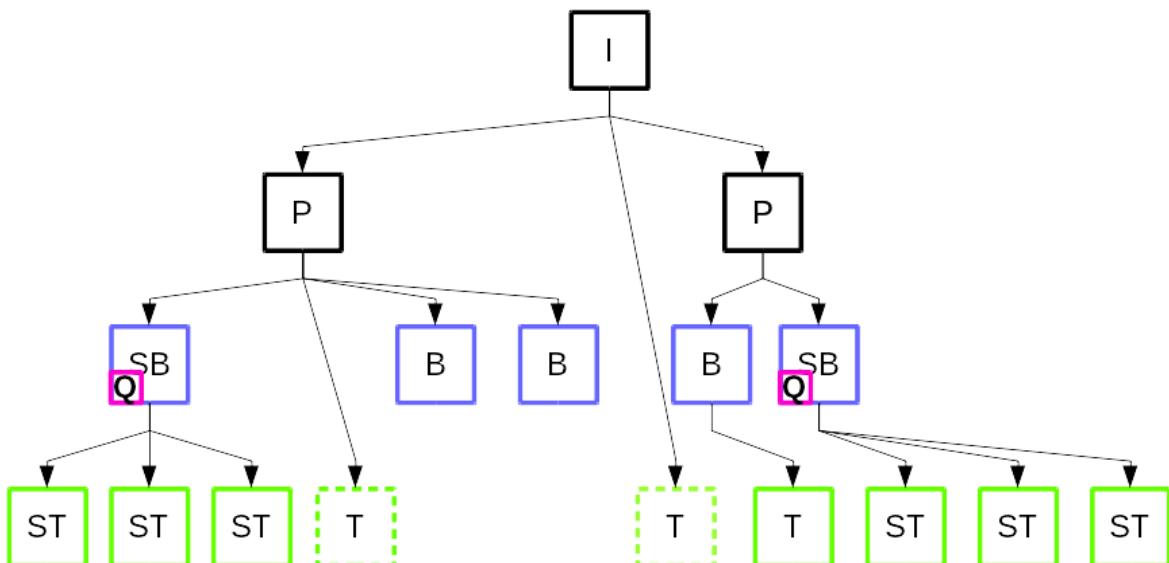
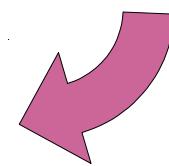
Contributions
Knowledge-driven approach

Hypotheses of panels, balloons and text lines

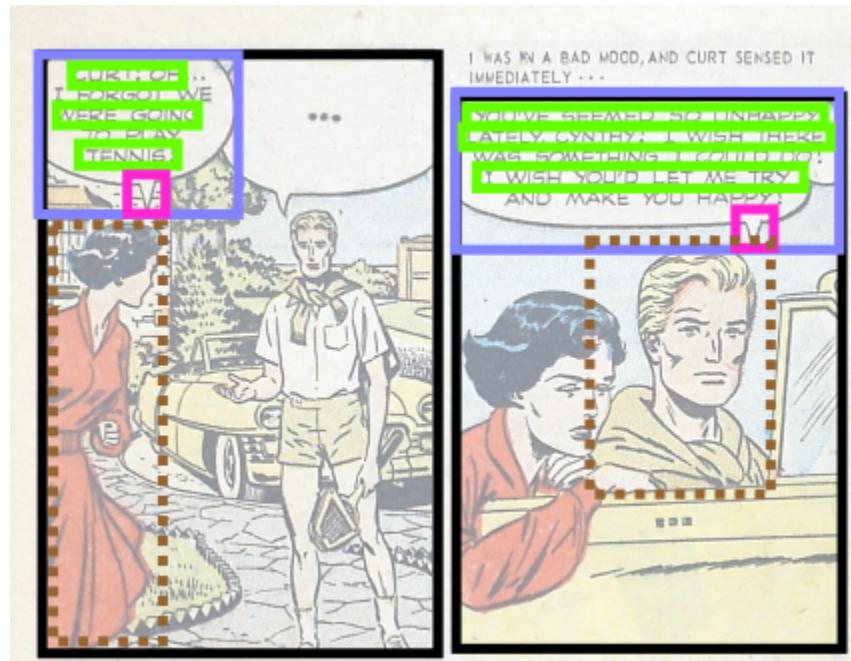


Validation of the hypotheses

Inferences of specific types



Processing sequence



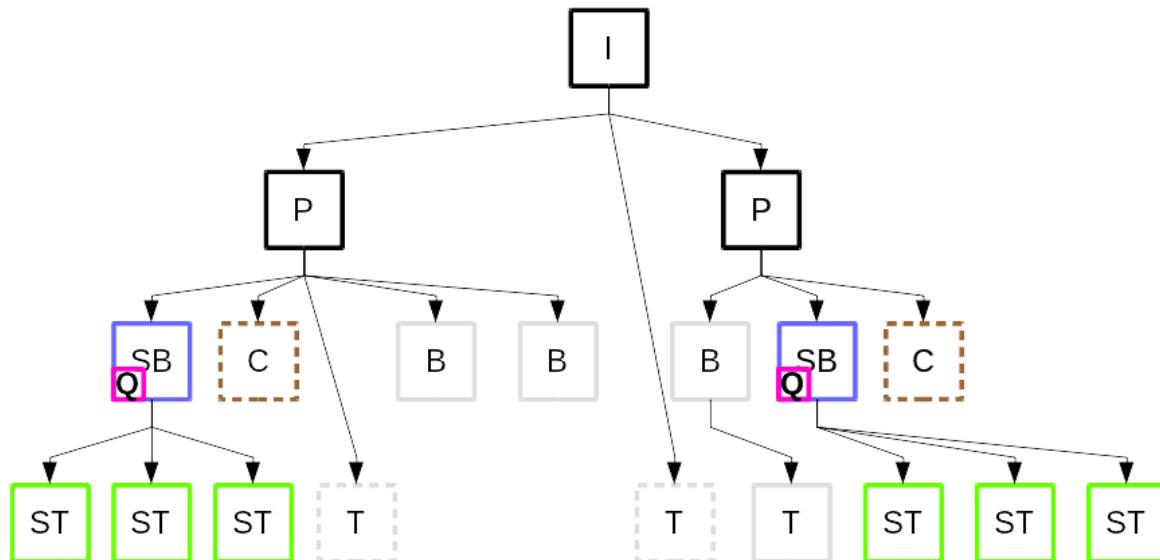
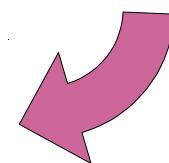
Contributions
Knowledge-driven approach

Hypotheses of
comic characters

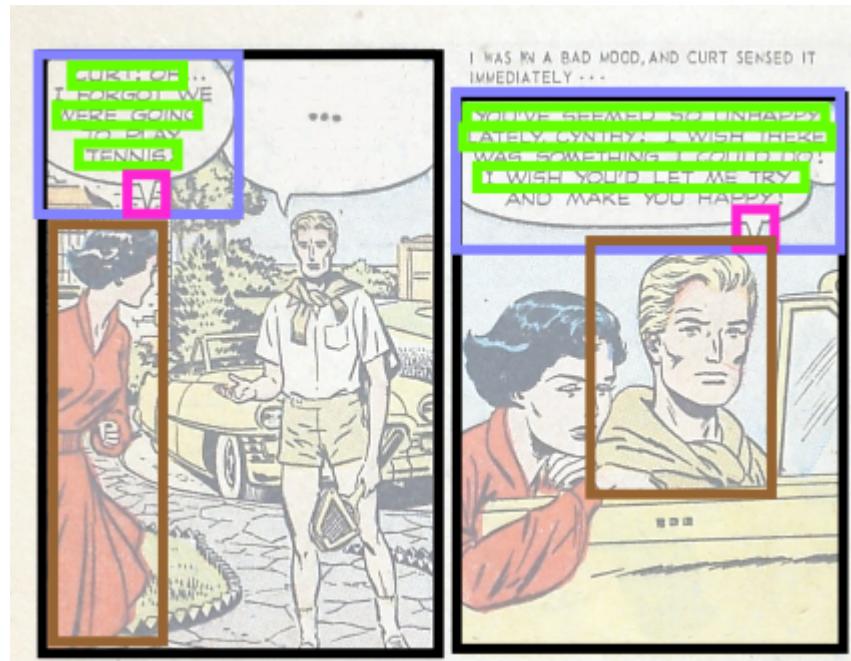


Validation of the
hypotheses

Inferences of
specific types

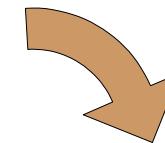


Processing sequence



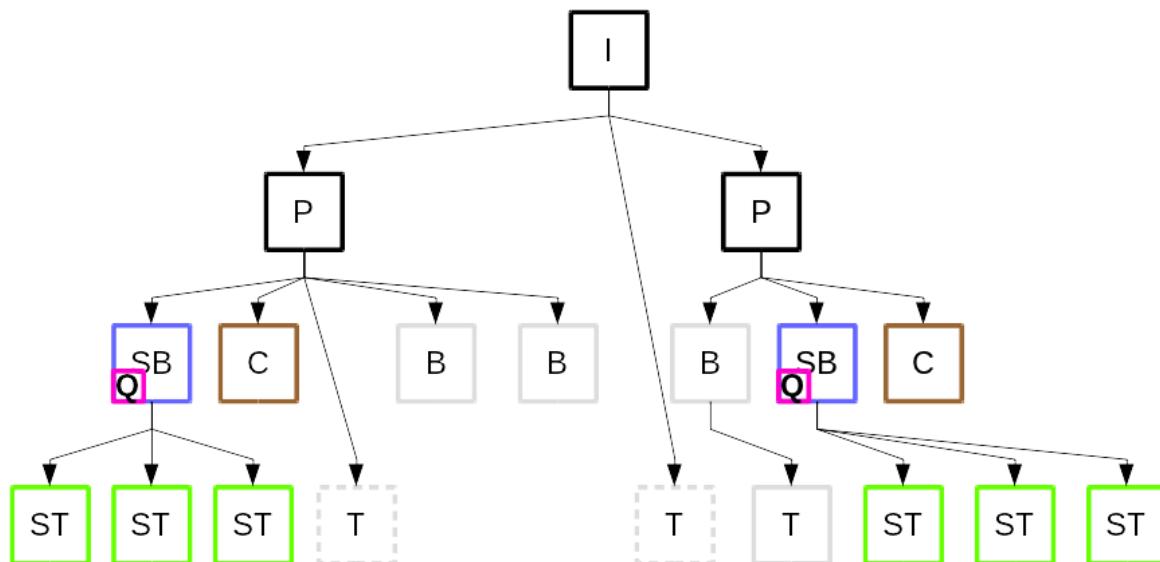
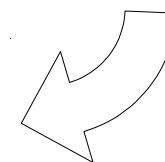
Contributions
Knowledge-driven approach

Hypotheses of
comic characters



Validation of the
hypotheses

Inferences of
specific types

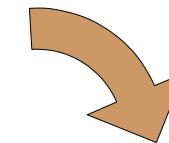
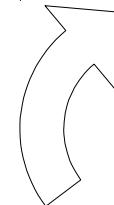


Processing sequence



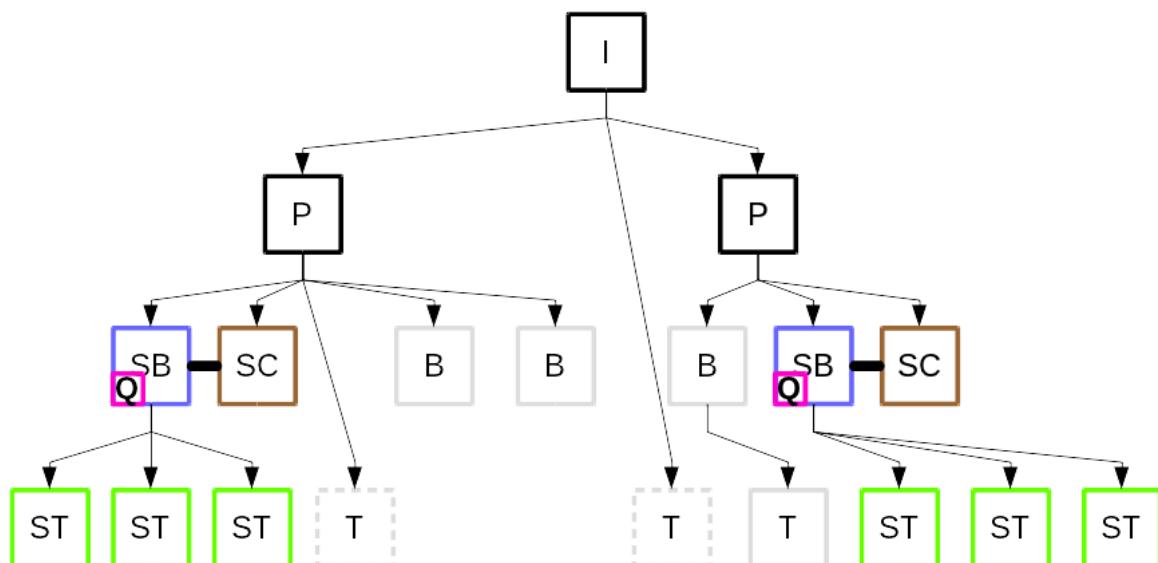
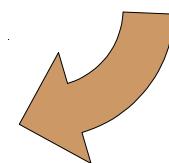
Contributions
Knowledge-driven approach

Hypotheses of
comic characters



Validation of the
hypotheses

Inferences of
specific types
+ semantic links



Results

- Performances

Contributions

Knowledge-driven approach

TODO

- Evaluations
- Overall contribution



Lettering. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Evaluations

Experiments

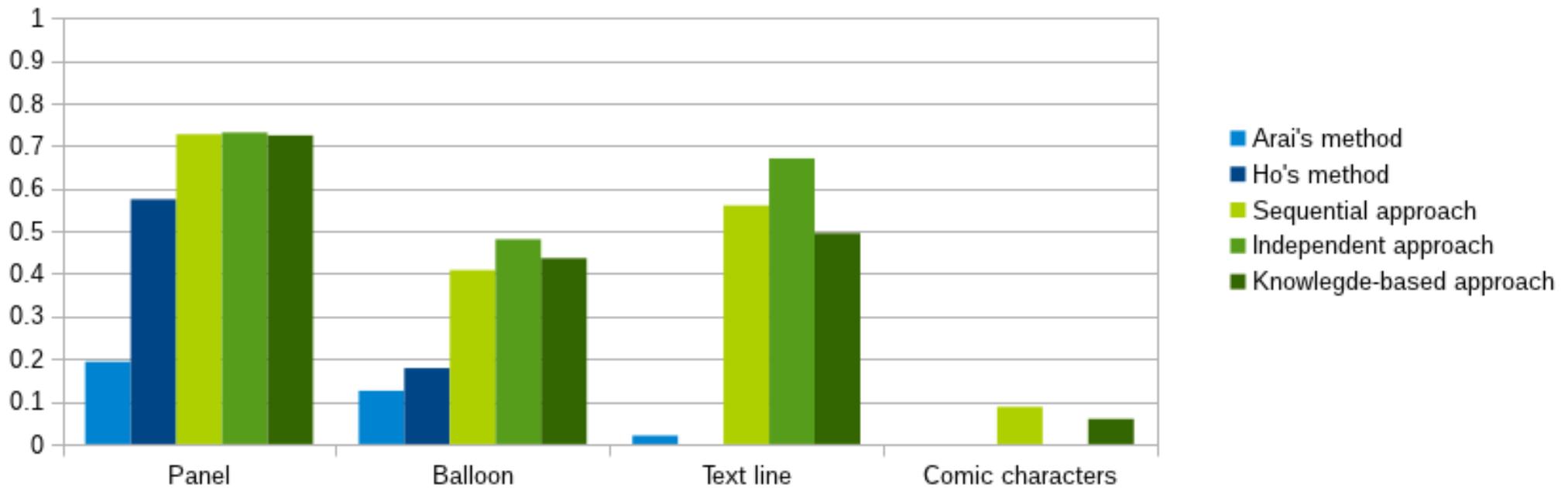
B_p = predicted region

B_{gt} = ground truth region

B_p is valid if $a_0 > 0.5$

$$a_0 = \frac{\text{area}(B_p \cup B_{gt})}{\text{area}(B_p \cap B_{gt})}$$

Extraction results on the eBDtheque dataset (F-score)



Overall contribution

Experiments

Element	Process type	Before	After
Panel	Localisation		
	Classification		
Balloon	Localisation		
	Classification		
	Tail detection		
Text	Localisation		
	Recognition		
Comic character	Localisation		
	Identification		
	Face/pose		
Context	Inter-element link		
	Situation retrieval		
	Timestamps		
Dataset	Localisation		
	Semantic		

Solved
 Advanced
 Medium
 Early stage
 Unexplored

- Global conclusions
- Global perspectives
- Publications



Lettering. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Global conclusions

Conclusion

- Reached objectives
 - Efficient **panel**, **balloon**, **text** and **tail** extraction methods
 - First approaches for **comic character** extraction and **context retrieval**
 - Public **dataset** and **ground truth** (<http://ebdtheque.univ-lr.fr>)
- Publications
 - 1 journal, 2 book series , 4 conferences, 5 workshops (3 national)
 - 6 local **seminars**
- Research impacts
 - **L3i** is now a **main actor** of comic book analysis in Europe
 - New Ph.D. **thesis** started in 2013 (Nam Le Thanh)
 - **Dataset** used by international peers (Germany, India, China, Japan)
 - **National projects** (PIA BigData Actialuna/LIP6, ANR EXPION 2015)
 - **International project** on manga analysis (PHC-SAKURA with Japan)

- Content extraction
 - Consider overlapping panel extraction
 - Investigate text recognition
 - Improve implicit balloon extraction and evaluation
 - Extract and identify non-speaking comic characters
- Content understanding
 - Situation retrieval (e.g. landscape, outdoor, night)
 - Action recognition (e.g. running, driving, dreaming)
 - Interaction retrieval (e.g. balloon said by/to)
 - Labelling from text analysis (e.g. auto tagging)
- Dataset
 - Increase the number of pages
 - Add more annotation (e.g. panel situation, character names and roles)
 - Annotate multi-parts of comic characters

Publications

Conclusion

JOURNAL

Christophe Rigaud, Clément Guérin, Dimosthenis Karatzas, Jean-Christophe Burie and Jean-Marc Ogier. “**Knowledge-driven understanding of images in comic books**”. International Journal on Document Analysis and Recognition (**IJDAR**), **2015** (accepted with minor reviews).

BOOK SERIES

Christophe Rigaud, Dimosthenis Karatzas, Jean-Christophe Burie and Jean-Marc Ogier. “**Adaptive contour classification of comics speech balloons**”. In Graphic Recognition. New Trends and Challenges. Lecture Notes in Computer Science (**LNCS**), Vol. 8746, **2014**.

Christophe Rigaud, Norbert Tsopze, Jean-Christophe Burie and Jean-Marc Ogier. “**Robust frame and text extraction from comic books**”. In Graphic Recognition. New Trends and Challenges. Lecture Notes in Computer Science (**LNCS**), Vol. 7423, pp. 129-138, **2013**.

Publications

Conclusion

CONFERENCES

Christophe Rigaud, Dimosthenis Karatzas, Jean-Christophe Burie and Jean-Marc Ogier. “**Color descriptor for content-based drawing retrieval**”. In the Proceedings of the 11th IAPR International Workshop on Document Analysis Systems (**DAS**), pp. 267-271 , Tours, France, April, **2014**.

Christophe Rigaud, Dimosthenis Karatzas, Joost Van de Weijer, Jean-Christophe Burie and Jean-Marc Ogier. “**An active contour model for speech balloon detection in comics**”. In the Proceedings of the 12th International Conference on Document Analysis and Recognition (**ICDAR**), pp. 1240-1244, Washington DC, USA, August, **2013**.

Clément Guérin, Christophe Rigaud, Antoine Mercier, Farid Ammar-Boudjelal, Karelle Bertet, Alain Bouju, Jean-Christophe Burie, Georges Louis, Jean-Marc Ogier and Arnaud Revel. “**eBDtheque: a representative database of comics**”. In the Proceedings of the 12th International Conference on Document Analysis and Recognition (**ICDAR**), pp. 1145-1149, Washington DC, USA, August, **2013**.

Christophe Rigaud, Dimosthenis Karatzas, Joost Van de Weijer, Jean-Christophe Burie and Jean-Marc Ogier. “**Automatic Text Localisation in Scanned Comic Books**”. In the Proceedings of the 8th International Conference on Computer Vision Theory and Applications (**VISAPP**), pp. 814-819, Barcelona, Spain, February, **2013**.

Publications

Conclusion

WORKSHOPS

Clément Guérin, Christophe Rigaud, Karell Bertet, Jean-Christophe Burie, Arnaud Revel and Jean-Marc Ogier. “**Réduction de l'espace de recherche pour les personnages de bandes dessinées**”. In the Proceedings of the 19ème congrès national sur la Reconnaissance de Formes et l'Intelligence Artificielle (**RFIA**), Rouen, France, July, **2014**.

Christophe Rigaud, and Clément Guérin. “**Localisation contextuelle des personnages de bandes dessinées**”. In the Proceedings of the 13ème Colloque International Francophone sur l'Ecrit et le Document (**CIFED**), pp. 367–370, Nancy, France, March **2014**.

Christophe Rigaud, Dimosthenis Karatzas, Jean-Christophe Burie and Jean-Marc Ogier. “**Speech balloon contour classification in comics**”. Proceedings of the 10th International Workshop on Graphics RECognition (**GREC**), pp. 23-25, Bethlehem, USA, August, **2013**.

Hoang Nam Ho, Christophe Rigaud, Jean-Christophe Burie and Jean-Marc Ogier. “**Redundant structure detection in attributed adjacency graphs for character detection in comics books**”. In the Proceedings of the 10th IAPR International Workshop on Graphics RECognition (**GREC**), pp. 109-113, Bethlehem, PA, USA, August, **2013**.

Christophe Rigaud, Norbert Tsopze, Jean-Christophe Burie and Jean-Marc Ogier. “**Extraction robuste des cases et du texte de bandes dessinées**”. In the Proceedings of the 10ème Colloque International Francophone sur l'Ecrit et le Document (**CIFED**), pp. 349-360, Bordeaux, France, March **2012**.

References

Conclusion

- [Arai10] Kohei Arai and Herman Tolle. **Method for automatic e-comic scene frame extraction for reading comic on mobile devices.** In Seventh International Conference on Information Technology: New Generations, ITNG '10, pages 370–375, Washington, DC, USA, 2010. IEEE Computer Society.
- [Cheung08] S.C.S. Cheung, City University of Hong Kong. Run Run Shaw Library, and City University of Hong Kong. **Face Detection and Face Recognition of Human-like Characters in Comics.** Outstanding academic papers by students. Run Run Shaw Library, City University of Hong Kong, 2008.
- [Chung07] ChungHo Chan, Howard Leung, and Taku Komura. **Automatic panel extraction of color comic images.** In HoraceH.-S. Ip, OscarC. Au, Howard Leung, Ming-Ting Sun, Wei-Ying Ma, and Shi-Min Hu, editors, Advances in Multimedia Information Processing - PCM 2007, volume 4810 of Lecture Notes in Computer Science, pages 775–784. Springer Berlin Heidelberg, 2007.
- [Eunjung07] Eunjung Han, Kirak Kim, HwangKyu Yang, and Keechul Jung. **Frame segmentation used mlp-based x-y recursive for mobile cartoon content.** In Proceedings of the 12th international conference on Human-computer interaction: intelligent multimodal interaction environments, HCI'07, pages 872–881, Berlin, Heidelberg, 2007. Springer.
- [Ho12] Anh Khoi Ngo Ho, Jean-Christophe Burie, and Jean-Marc Ogier. **Panel and speech balloon extraction from comic books.** In 2012 10th IAPR International Workshop on Document Analysis Systems, pages 424–428. IEEE, 2012.

References

Conclusion

- [Khan12] Fahad Shahbaz Khan, Muhammad Anwer Rao, Joost van de Weijer, Andrew D. Bagdanov, Maria Vanrell, and Antonio Lopez. **Color attributes for object detection**. In 25th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012), 2012.
- [Li14a] Luyuan Li, Yongtao Wang, Zhi Tang, and Liangcai Gao. **Automatic comic page segmentation based on polygon detection**. Multimedia Tools Applications, 171–197, 2014, Kluwer Academic Publishers.
- [Li14b] Luyuan Li, Yongtao Wang, Zhi Tang, Xiaoqing Lu, and Liangcai Gao. **Unsupervised speech text localization in comic images**. In Proceedings of International Conference on Document Analysis and Recognition (ICDAR), pages 1190–1194, Aug 2013
- [Pang14] Xufang Pang, Ying Cao, Rynson W.H. Lau, and Antoni B. Chan. **A robust panel extraction method for manga**. In Proceedings of the ACM International Conference on Multimedia, MM '14, pages 1125–1128, New York, NY, USA, 2014.
- [Ponsard12] Christophe Ponsard, Ravi Ramdoyal, and Daniel Dziamski. **An ocr-enabled digital comic books viewer**. In Computers Helping People with Special Needs, pages 471–478. Springer, 2012.
- [Stommel12] Martin Stommel, Lena I Merhej, and Marion G Müller. **Segmentation-free detection of comic panels**. In Computer Vision and Graphics, pages 633–640, 2012.

References

Conclusion

- [Su11] Chung-Yuan Su, Ray-I Chang, and Jen-Chang Liu. **Recognizing text elements for svg comic compression and its novel applications**. In Proceedings of International Conference on Document Analysis and Recognition (ICDAR), pages 1329–1333, Washington, DC, USA, 2011.
- [Sun10] Weihan Sun and Koichi Kise. **Similar partial copy detection of line drawings using a cascade classifier and feature matching**. In Hiroshi Sako, Katrin Franke, and Shuji Saitoh, editors, ICWF, volume 6540 of Lecture Notes in Computer Science, pages 126–137. Springer, 2010.
- [Tanaka07] Takamasa Tanaka, Kenji Shoji, Fubito Toyama, and Juichi Miyamichi. **Layout analysis of tree-structured scene frames in comic images**. In IJCAI’07, pages 2885–2890, 2007.



GRACIAS ARIGATO SHUKURIA JUSPAXAR TASHAKKUR ATU YAQHANYELAY MAAKE GOZAIMASHITA EFCHARISTO LAH GRAZIE MEHRBANI PALDIES BOLZIN THANK YOU BIYAN SHUKRIA MERCI

SPASSIBO NUHUN CHALTU WABEEJA MAITEKA HUI GUI EKOJU SIKOMO MANETAI
SNACHALHYA DHANYABAAD ANNA UNALCHEESH DENKAUJA HEACHALHYA MERSI SPASIBO
MERASTAWHY SANCO GAEJTHO AGUYJE FAKAAUE
BAINKA KOMAPSUMNIDA ATTO
TAVTAPUCH MEDAWAGSE LAH
BANKA
TINGKI
HATUR
EKOJU
SIKOMO
MANETAI
MINMONCHAR



<https://github.com/crigaud/thesis/tree/master/presentation>
<http://www.christophe-rigaud.com>

Complementary information

History of comics art

- Pre-history: painting of animals and hunters in caves [Marx, 2007]
- 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 to social media sites and mobile devices



Pre-history and ancient Egypt



Rodolphe Töpffer, Histoire de Monsieur Cryptogame (1830)



Ted McCall, Robin Hood And Company (1946)



Tezuka Osamu, Manga Classroom (1953)



Marion Montaigne, Tu mourras moins bête (2013)

GT validation

Knowledge-driven analysis detail

- Comic character region refinement

Production to interpretation

Background

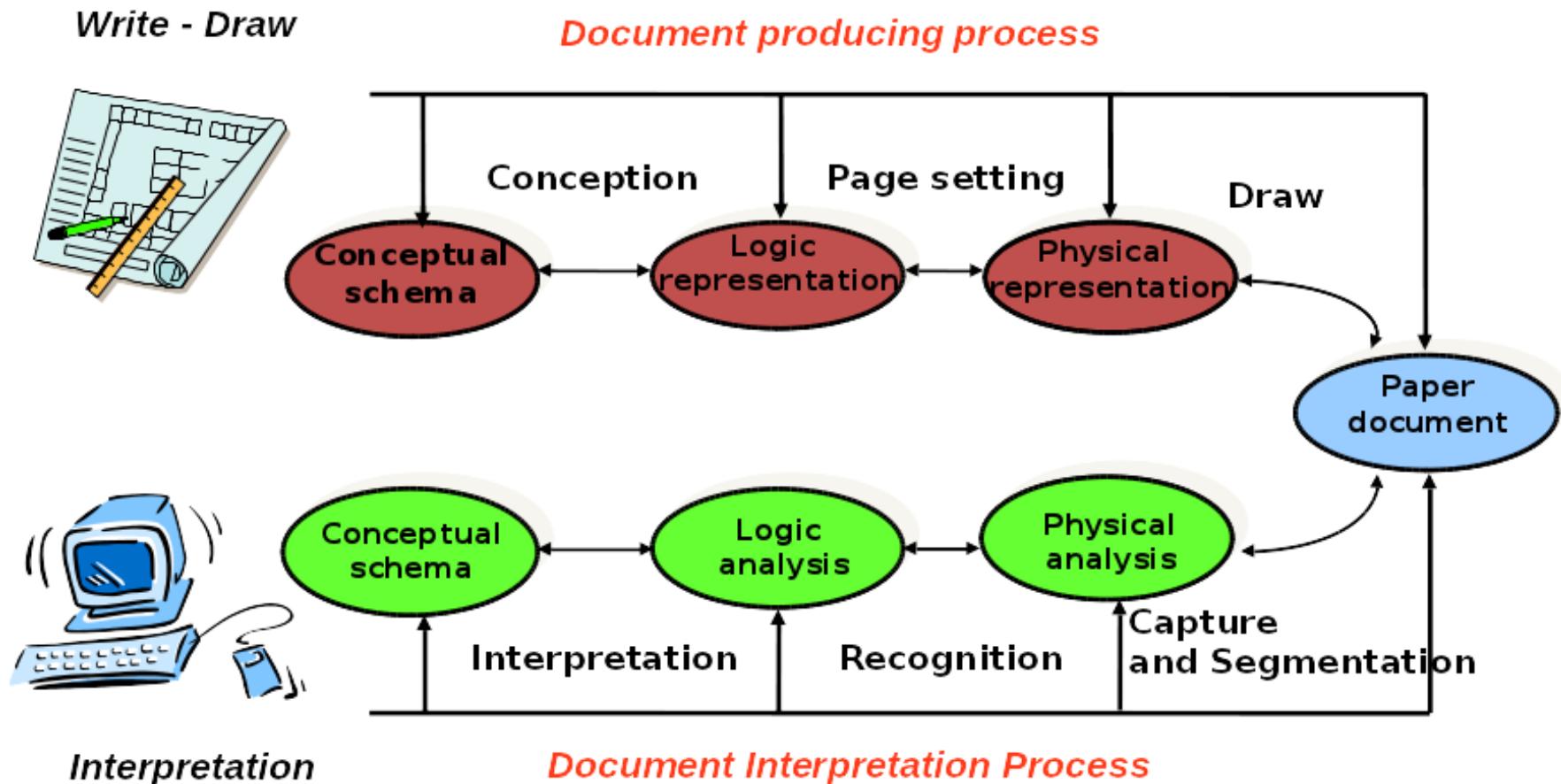


Image source: Handbook of Document Image Processing and Recognition. Springer, 2014

Element	Process type	Analysis method	Method	Publication	Before	After	Δ
Panel	Localisation	Histogram	S	CIFED'12 LNCS'13	60%	70%	+10%
		Topology	I, K	IJDAR'15			
Balloon	Localisation	Topology	S	-	20%	50%	+30%
		Region	I, K	-			
		Edge	S	ICDAR'13			
	Classification	Contour	S, I, K	GREC'13 LNCS'14	0%	80%	+80%
Text	Localisation	Topology	S		5%	60%	+55%
		Structure	I, K	VISAPP'13			
	Recognition	-	-	-	10%	10%	0%
Comic characters	Localisation	Graph	-	GREC'13 (2nd)	5%	10%	+5%
		Context	S, K	CIFED'14 RFIA'14 (2nd)			
		Spotting	Colour	DAS'14			
	Identification	-	-	-	10%	10%	0%
	Face/pose	-	-	-	0%	0%	0%
	Inter-element	Proximity	S, K	CIFED'14 RFIA'14 (2nd) IJDAR'15	0%	10%	+10%

Background



Thinking about the scenario. Image credits:
Wissam Shekhani, 2010



European Ph.D. defense

Segmentation and indexation of complex objects in comic book images

Christophe Rigaud
December 11th, 2014

Co-supervised by:

Jean-Christophe Burie¹
Dimosthenis Karatzas²
Jean-Marc Ogier¹

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



Thank you Mr. the president, thanks to all the other members for being in the jury of my Ph.D. defence.

Thank you to the public and online people for .

I'm Christophe Rigaud, Ph.D. student in computer science and document image analysis.

Today I will present my work about “segmentation and indexation of complex objects in comic book images”.

This work have been co-supervised by JCB and JMO from L3i and DK from CVC

SEGMENTATION

In computer vision, image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as superpixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze.

INDEXATION

“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”
Hiroaki Tobita, Sony CSL Interaction Laboratory, 2014

2

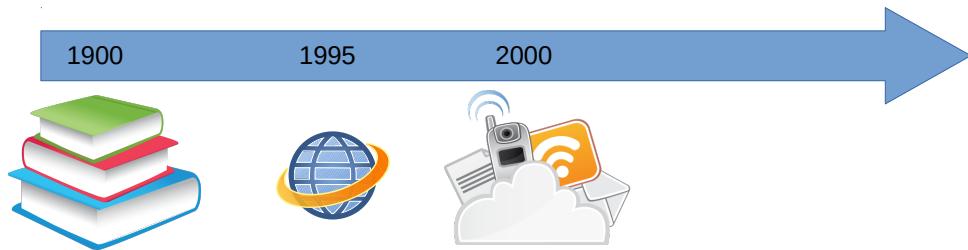
- Pre-history: painting of animals and hunters in caves [Marx, 2007]
- 1846: Rodolphe Töpffer, the inventor of the “bandes dessinées”
- 1930s: magazine-style comic books production in the US
- 1950s: massive production of manga in Japan (Osamu Tezuka)
- 1971: the term of ninth art is attributed to comics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

Comic books

Introduction

“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”
Hiroaki Tobita, Sony CSL Interaction Laboratory, 2014

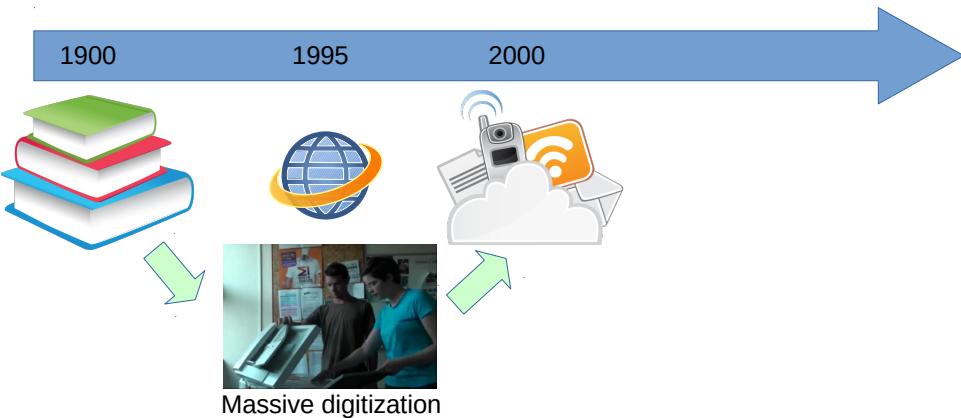


Comic books

Introduction

“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”
Hiroaki Tobita, Sony CSL Interaction Laboratory, 2014

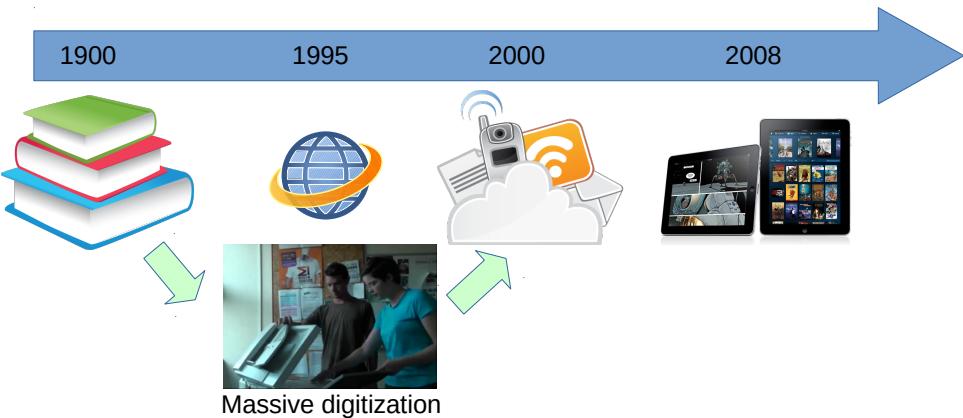


Comic books

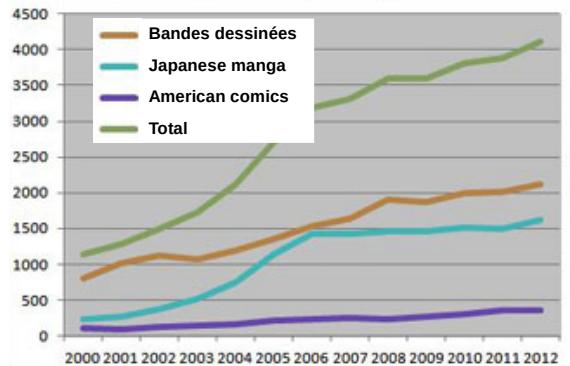
Introduction

“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”
Hiroaki Tobita, Sony CSL Interaction Laboratory, 2014

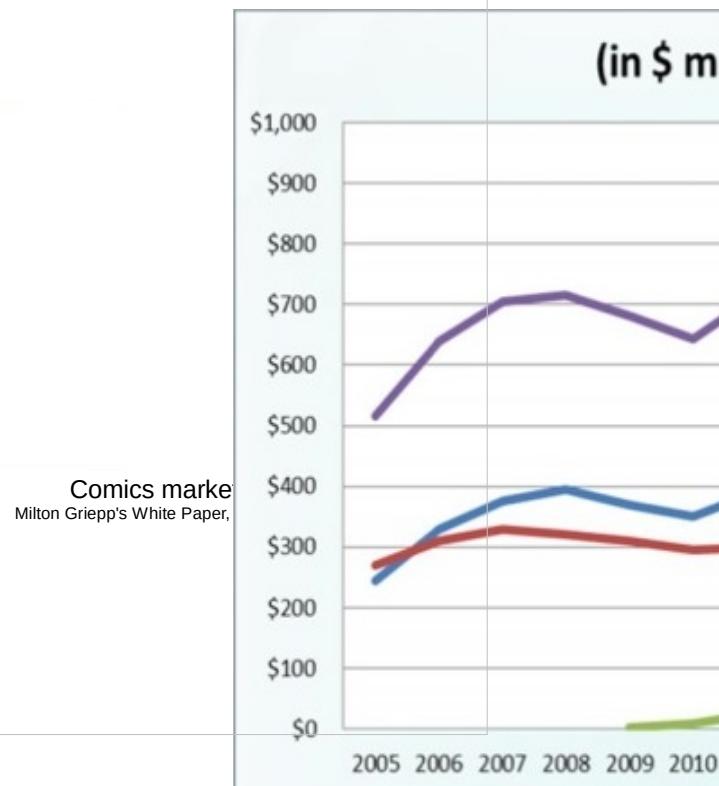


Comic books



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

Introduction



(in \$ m)

Comics market
Milton Griepp's White Paper,

Context of the thesis

Introduction

- eBDthèque project (since 2011)
 - Add value to digitized comics using the new technologies
 - Content extraction (thesis of Christophe Rigaud)
 - Knowledge representation (thesis of Clément Guérin)
 - Public funding CPER 2007-2013
 - 2 Ph.D. students, 1 engineer, 1 post doc, 6 professors (L3I)
- Scientific challenges
 - Mixed contents of a graphical and textual nature
 - Combination of the difficulties of free-form and complex background documents
 - Recent field of research with a largely unknown
- Objectives
 - Propose generic approaches able to retrieve as many elements as possible from any comic book image
 - Provide a first dataset and ground truth

7

Challenges:

- no regular structure allowing to extract easily the layout and to predict the location of text and graphics
- authors are entirely free in their layout choices (few conventions widely adopted by comic book's authors in order to avoid the reader to be confused)

Context of the thesis

Introduction

- eBDthèque project (since 2011)
 - Add value to digitized comics using the new technologies
 - Content extraction (thesis of Christophe Rigaud)
 - Knowledge representation (thesis of Clément Guérin)
 - Public funding CPER 2007-2013
 - 2 Ph.D. students, 1 engineer, 1 post doc, 6 professors (L3I)
- Scientific challenges
 - Mixed contents of a graphical and textual nature
 - Combination of the difficulties of free-form and complex background documents
 - Recent and unexplored field of research
- Objectives
 - Propose generic approaches able to retrieve as many elements as possible from any comic book image
 - Provide a first dataset and ground truth

8

Challenges:

- no regular structure allowing to extract easily the layout and to predict the location of text and graphics
- authors are entirely free in their layout choices (few conventions widely adopted by comic book's authors in order to avoid the reader to be confused)

Context of the thesis

Introduction

- eBDthèque project (since 2011)
 - Add value to digitized comics using the new technologies
 - Content extraction (thesis of Christophe Rigaud)
 - Knowledge representation (thesis of Clément Guérin)
 - Public funding CPER 2007-2013
 - 2 Ph.D. students, 1 engineer, 1 post doc, 6 professors (L3I)
- Scientific challenges
 - Mixed contents of a graphical and textual nature
 - Combination of the difficulties of free-form and complex background documents
 - Recent and unexplored field of research
- Objectives
 - Propose generic approaches able to retrieve as many elements as possible from any comic book image
 - Provide a first dataset and ground truth

9

Origin:

- master thesis of NGO HO Anh Khoi

Challenges:

- no regular structure allowing to extract easily the layout and to predict the location of text and graphics
- authors are entirely free in their layout choices (few conventions widely adopted by comic book's authors in order to avoid the reader to be confused)

- Panel extraction
- Balloon extraction
- Text extraction & recognition
- Comic character extraction
- Conclusions



Pencil drawing. Image credits: Le cycle des bulles, Christophe Rigaud, 2012

Announce the plan here

Panel extraction

Background

- Challenges
 - Diversity of styles (gutter, implicit)
 - Semi-structured layout
- Panel extraction
 - White line cut [Chung07]
 - Recursive X-Y cut [Eunjung07]
 - Density gradient [Tanaka07]
 - Connected-components [Arai10, Pang14]
 - Polygon detection [Li14a]
 - Corners and line segments [Stommel12]
- Conclusions
 - Specific approaches not appropriate for other type of comics
 - Remaining difficulties for non-rectangle and implicit panels
 - Copyrighted images (not shareable)



13

Insist on panel border style differences

Panel extraction

Background

- Challenges
 - Diversity of styles (gutter, implicit)
 - Semi-structured layout
- Panel extraction
 - White line cut [Chung07]
 - Recursive X-Y cut [Eunjung07]
 - Density gradient [Tanaka07]
 - Connected-components [Arai10, Pang14]
 - Polygon detection [Li14a]
 - Corners and line segments [Stommel12]
- Conclusions
 - Specific approaches not appropriate for other type of comics
 - Remaining difficulties for non-rectangle and implicit panels
 - Copyrighted images (not shareable)



14

- Stommel12
 - Able to extract gutter-free and overlapped panels
 - Output only rectangle regions

Balloon extraction

- Challenges
 - Difference between **shape** and **contour**
 - Implicit balloon positions
 - Semantics related to text
- Extraction
 - Connected-components [Arai11, Ho12]
- Conclusions
 - Closed balloon with text inside
 - Several unexplored fields (e.g. implicit balloon positions, balloon, classification, tail detection)

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

15

Intro on contour/shape difference

Arai [8] proposed a blob detection method based on connected component detection with four filtering rules applied to manga analysis.

The rules are based on blob minimum size, white pixel occurrence, inclusion of vertical straight lines and width to length ratio (Figure 2.6).

Another connected component approach proposed by Ho [68] uses HSV colour space to make a first selection of bright blobs and then consider as balloons the blobs with a ratio between the text area and the blob bounding box higher than sixty per cent.

Implicit balloons: not studied yet but active contour model initialised from text seems to be a good approach

Balloon extraction

- Challenges
 - Difference between **shape** and **contour**
 - Implicit balloon positions
 - Semantics related to text
- Extraction
 - Connected-components [Arai11, Ho12]
- Conclusions
 - Closed balloon with text inside
 - Several unexplored fields (e.g. implicit balloon positions, balloon, classification, tail detection)

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

16

Intro on contour/shape difference

Arai [8] proposed a blob detection method based on connected component detection with four filtering rules applied to manga analysis.

The rules are based on blob minimum size, white pixel occurrence, inclusion of vertical straight lines and width to length ratio (Figure 2.6).

Another connected component approach proposed by Ho [68] uses HSV colour space to make a first selection of bright blobs and then consider as balloons the blobs with a ratio between the text area and the blob bounding box higher than sixty per cent.

Implicit balloons: not studied yet but active contour model initialised from text seems to be a good approach

Text extraction & recognition

Background

- Challenges
 - Non-standard fonts
 - Multi-script/orientation/scale
 - Complex background (sound effects)
 - Hyphenation, voluntary spelling mistakes
- Extraction
 - Sliding Concentric Windows + SVM [Su11]
 - Connected-components [Ho12, Pang14]
 - SVM and Bayesian classifier [Li14b]
- Recognition
 - OCR trained for a specific comics font [Ponsard12]



- Conclusions
 - Speech text (from balloons)
 - Captions and sound effects unexplored
 - Text recognition very poor

Text extraction & recognition

Background

- Challenges
 - Non-standard fonts
 - Multi-script/orientation/scale
 - Complex background (sound effects)
 - Hyphenation, voluntary spelling mistakes
- Extraction
 - Sliding Concentric Windows + SVM [Su11]
 - Connected-components [Ho12, Pang14]
 - SVM and Bayesian classifier [Li14b]
- Recognition
 - OCR trained for a specific comics font [Ponsard12]



- Conclusions
 - Speech text only (from balloons)
 - Captions and sound effects unexplored
 - Text recognition very poor

18

Mention onomatopoeia

Mention that scene text methods may be appropriate for complex background text (not tested yet) but not optimal for speech text (our first intention).

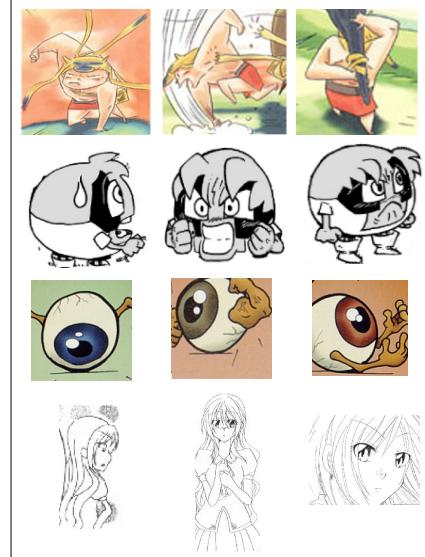
It is like document-like text inside scene-like complex background

2 in 1 is not obvious, otherwise scene text method would perform better document analysis methods.

Comic character extraction

Background

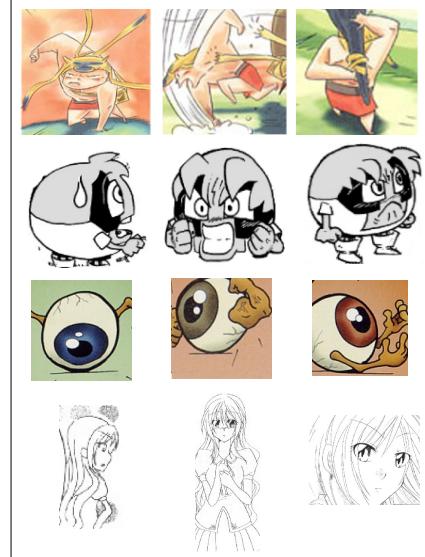
- Challenges
 - Hand-drawn, stroke-based
 - Intra/inter class variability
 - Scale, deformation, posture, occlusion
- Extraction & recognition
 - Manga faces [Cheung08, Sun10, Kohei12]
 - Cartoons [Khan12]
- Conclusions
 - Preliminary results
 - Complex and versatile structure
 - Contains most of the interesting information



Comic character extraction

Background

- Challenges
 - Hand-drawn, stroke-based
 - Intra/inter class variability
 - Scale, deformation, posture, occlusion
- Extraction & recognition
 - Manga faces [Cheung08, Sun10, Kohei12]
 - Cartoons [Khan12]
- Conclusions
 - Preliminary results
 - Complex and versatile structure
 - Contains most of the interesting information



Conclusions

Background

Element	Process type	Status
Panel	Localisation	
	Classification	
Balloon	Localisation	
	Classification	
	Tail detection	
Text	Localisation	
	Recognition	
Comic character	Localisation	
	Identification	
	Face/pose	
Context	Inter-element link	
	Situation retrieval	
	Timestamps	
Dataset	Localisation	
	Semantic	

 Solved
 Advanced
 Medium
 Early stage
 Unexplored

21

- From my interpretation, no common data for quantitative comparison.
- Panel extraction is often associated with timestamps retrieval, that is why they are both the most advanced researches.
- A lot of open issues...

- Introduction
- Sequential approach
- (Independent approach)
- Knowledge-driven approach



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

- **Introduction**
- Sequential approach
- (Independent approach)
- Knowledge-driven approach



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Introduction

- Objective: cover the widest possible scope of study

Contributions

24

Because there are a lot of open issues, we preferred to create an holistic framework in order to tackle most of them instead of one in particular.

Introduction

- Objective: cover the widest possible scope of study

1) Creation of heterogeneous dataset

- 100 mixed pages from 20 albums
- Franco-Belgium "bandes dessinées", American comics and Japanese manga
- From 1905 to 2012, paper and webcomics
- Rights holder permissions agreement

Contributions

Bibliographic annotations



PAGE	(100)
Collection:	Chilling Tales
Album:	17 Geo
Editor:	Youthful Magazines
Drawer:	Matt Fox
Writer:	Matt Fox
Language:	English
Page number:	16
Release date:	1953

Visual and semantic annotations



PANEL	(850)
Rank:	1
BALLOON	(1092)
Rank:	2
Shape:	Oval
Tail direction:	South-West
TEXT LINE	(4691)
Text:	« STARK RAVING »
CHARACTER	(1550)
LinkedToBalloon:	2

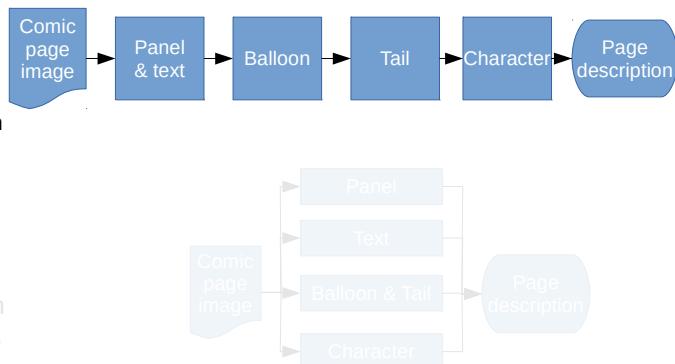
25

Because there are a lot of open issues, we preferred to create an holistic framework in order to tackle most of them instead of one in particular.

Introduction

- Objective: cover the widest possible scope of study
 - 1) Creation of heterogeneous dataset
 - 2) Three approaches
- Content-driven
 - Sequential approach
 - Similar to literature
 - Intuitive
 - Sensible to error propagation
 - Independent approach
 - Avoid error propagation
- Knowledge-driven
 - Knowledge-driven approach
 - Based on domain knowledge
 - Retrieve context

Contributions



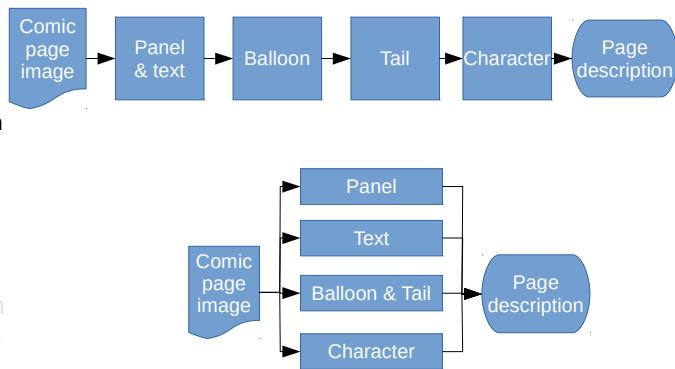
26

Because there are a lot of open issues, we preferred to create an holistic framework in order to tackle most of them instead of one in particular.

Introduction

- Objective: cover the widest possible scope of study
 - 1) Creation of heterogeneous dataset
 - 2) Three approaches
- Content-driven
 - Sequential approach
 - Similar to literature
 - Intuitive
 - Sensible to error propagation
 - Independent approach
 - Avoid error propagation
- Knowledge-driven
 - Knowledge-driven approach
 - Based on domain knowledge
 - Retrieve context

Contributions



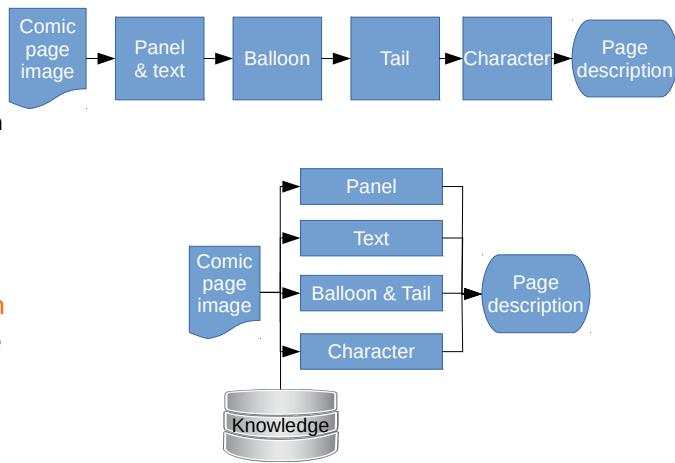
27

Explain why these three approaches

Introduction

- Objective: cover the widest possible scope of study
 - 1) Creation of heterogeneous dataset
 - 2) Three approaches
- Content-driven
 - Sequential approach
 - Similar to literature
 - Intuitive
 - Sensible to error propagation
 - Independent approach
 - Avoid error propagation
- Knowledge-driven
 - Knowledge-driven approach
 - Based on domain knowledge
 - Retrieve context

Contributions



28

Explain why these three approaches

- Introduction
- Sequential approach
 - (Independent approach)
 - Knowledge-driven approach



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

- Introduction
- Sequential approach
 - Panel & text extraction
 - Balloon extraction
 - Tail extraction
 - Comic character extraction
- (Independent approach)
- Knowledge-driven approach



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

- Introduction
- Sequential approach
 - Panel & text extraction
 - Balloon extraction
 - Tail extraction
 - Comic character extraction
- (Independent approach)
- Knowledge-driven approach

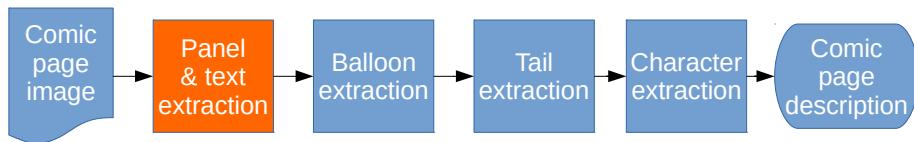


Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Panel & text extraction

Contributions

Sequential approach



- Literature
 - Panel with frame, separated by **gutters** or black line
 - Text located **inside** balloons
- Contribution
 - **Simultaneous panel and text** extraction from binary image
 - Consider **implicit** and **non-rectangle** panels
 - **Location-independent** text extraction

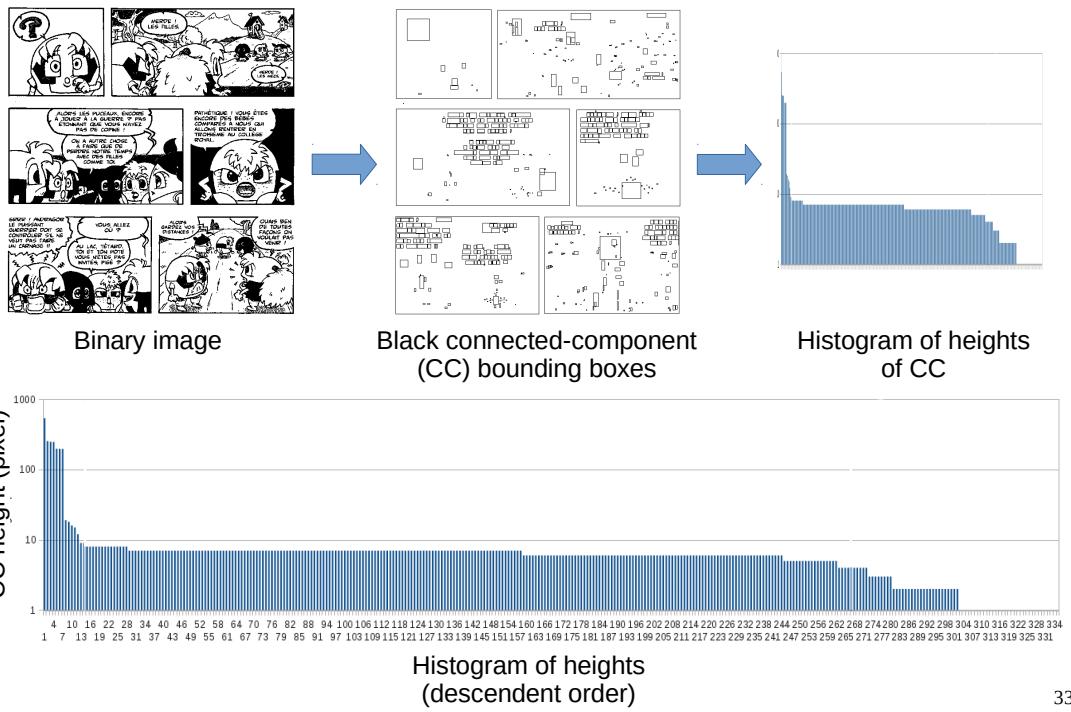
32

Justify connected-component analysis because it is a simple approach that have been used until recently for panel and text extraction

Panel & text extraction

Contributions

Sequential approach



33

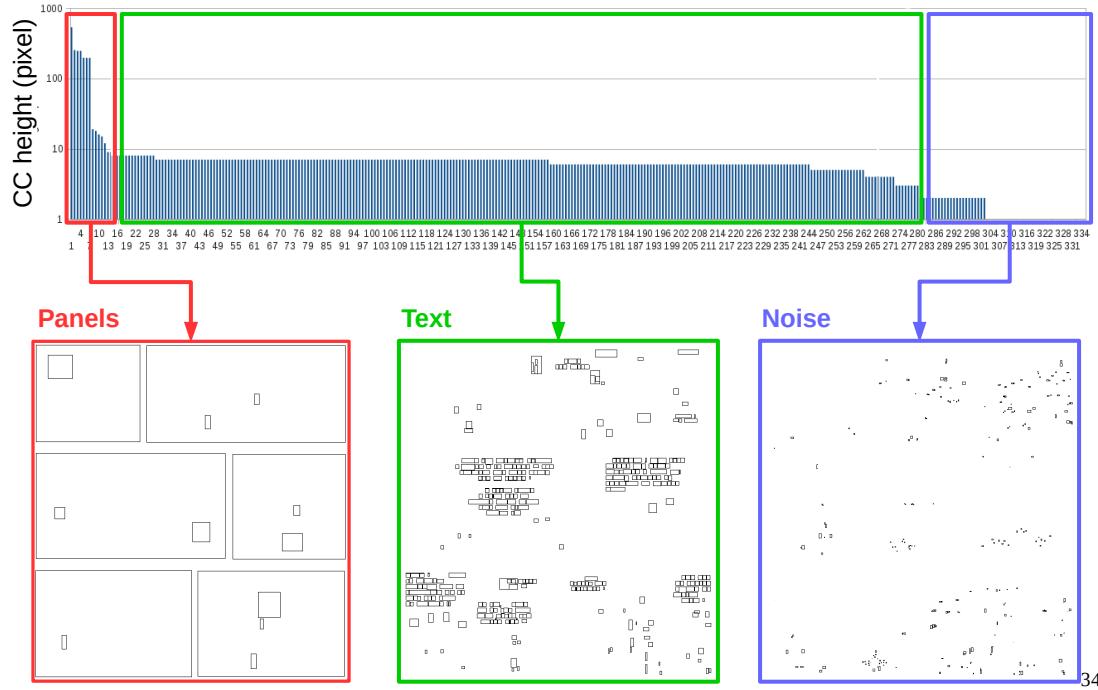
Justify connected-component analysis because it is a simple approach that have been used until recently for panel and text extraction

Panel & text extraction

Contributions

Sequential approach

K-means clustering (k=3)



Panel & text extraction: results

Contributions Sequential approach



TODO

35

- Panel & text extraction
 - Simple and **fast**
 - **Disconnected panels only**

- Introduction
- Sequential approach
 - Panel & text extraction
 - **Balloon extraction**
 - Tail extraction
 - Comic character extraction
- (Independent approach)
- Knowledge-driven approach

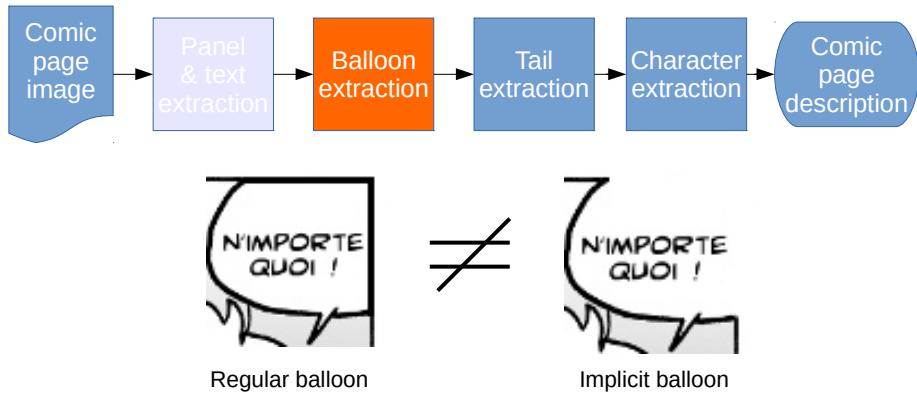


Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Balloon extraction

Contributions

Sequential approach

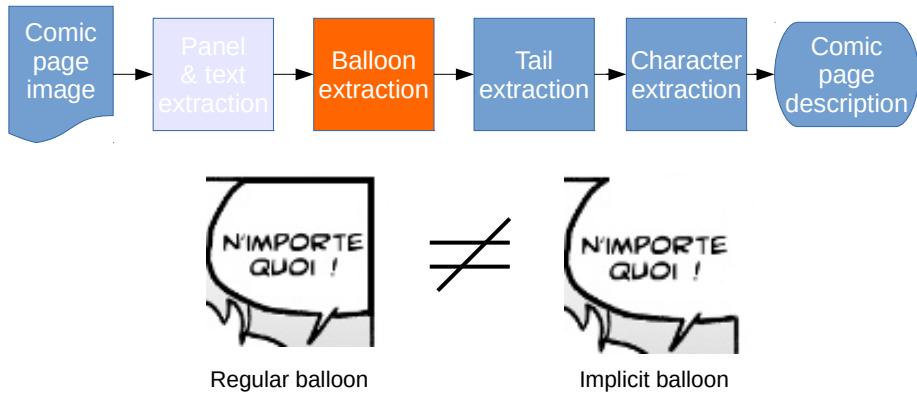


- Literature
 - Top-down approaches: extract white blobs and then text inside
 - Limited to regular balloons
- Contribution
 - Bottom-up approaches: extract text and then surrounding balloons
 - Appropriate for regular and implicit balloons

Balloon extraction

Contributions

Sequential approach



- 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
 - Improvement of regular and a first approach for implicit balloon extractions

Balloon extraction: regular

Contributions

Sequential approach

- Assumptions
 - Panels and text block positions **are known**
 - Regular balloons **contain centred text**
- Proposition → structural analysis
 - Extract **closed contours** that **fully include centred text**



Original image



Expected result

39

Justify contour extraction:

- Commonly used for balloon extraction
- Pixel level segmentation

“We use connected-component extraction because it is a simple and common approach for regular balloon extraction”

Balloon extraction: regular

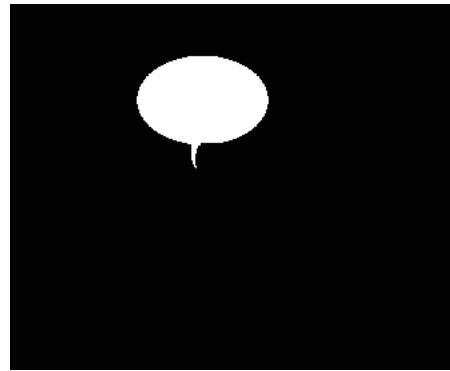
Contributions

Sequential approach

- Assumptions
 - Panels and text block positions are known
 - Regular balloons contain centred text
- Proposition → structural analysis
 - Extract closed contours that fully include centred text



Original image



Expected result

Balloon extraction: regular



Original image

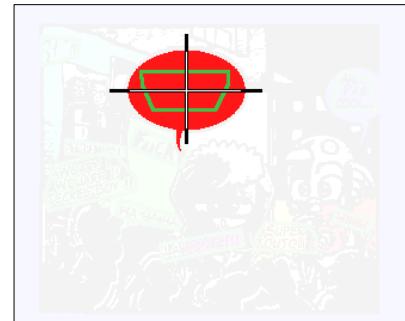
Contributions Sequential approach



Text block positions (green)



Regions including text blocks (coloured)



Regions including aligned text blocks

Balloon extraction: implicit

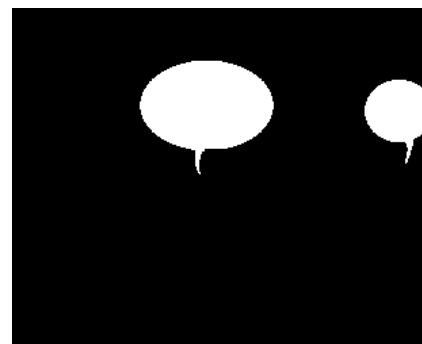
Contributions

Sequential approach

- Assumptions
 - Panel and text blocks positions are known
 - Implicit balloons contain centred text
- Proposition
 - Extract implicit balloons from text regions by inflating a deformable contour
 - Adaptation of active contour model (snake)



Original image and text locations



Expected result

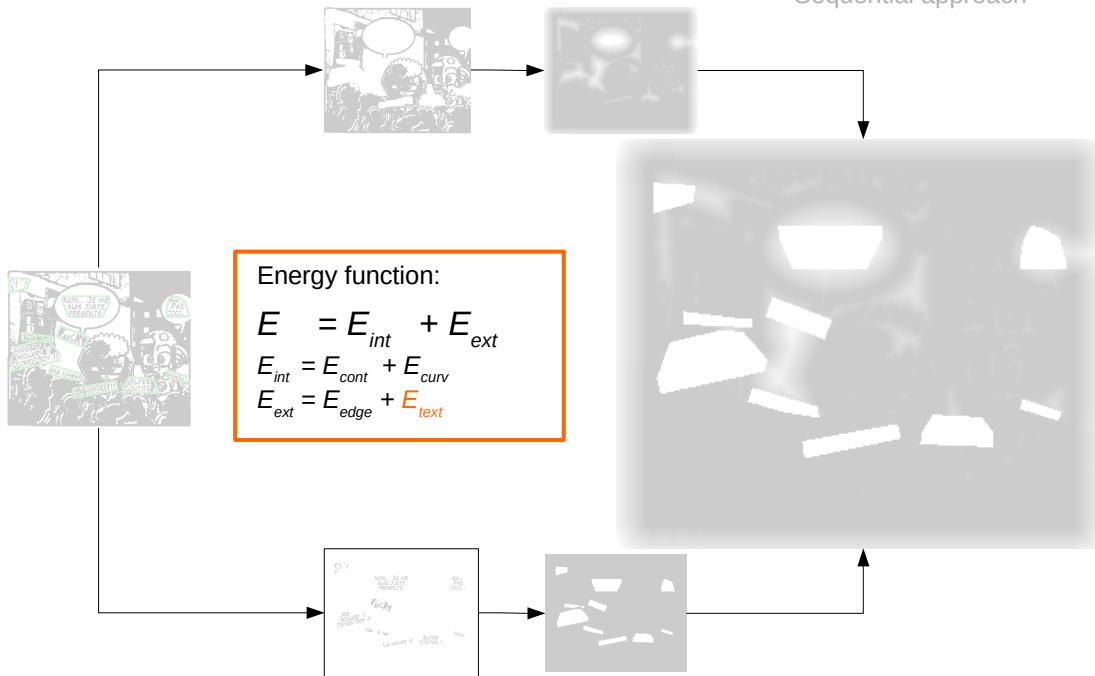
42

Justify the choice of active contour:

- Deformable to fit balloon contour
- Balloon-like shape given from internal energies

Balloon extraction: implicit

Contributions
Sequential approach

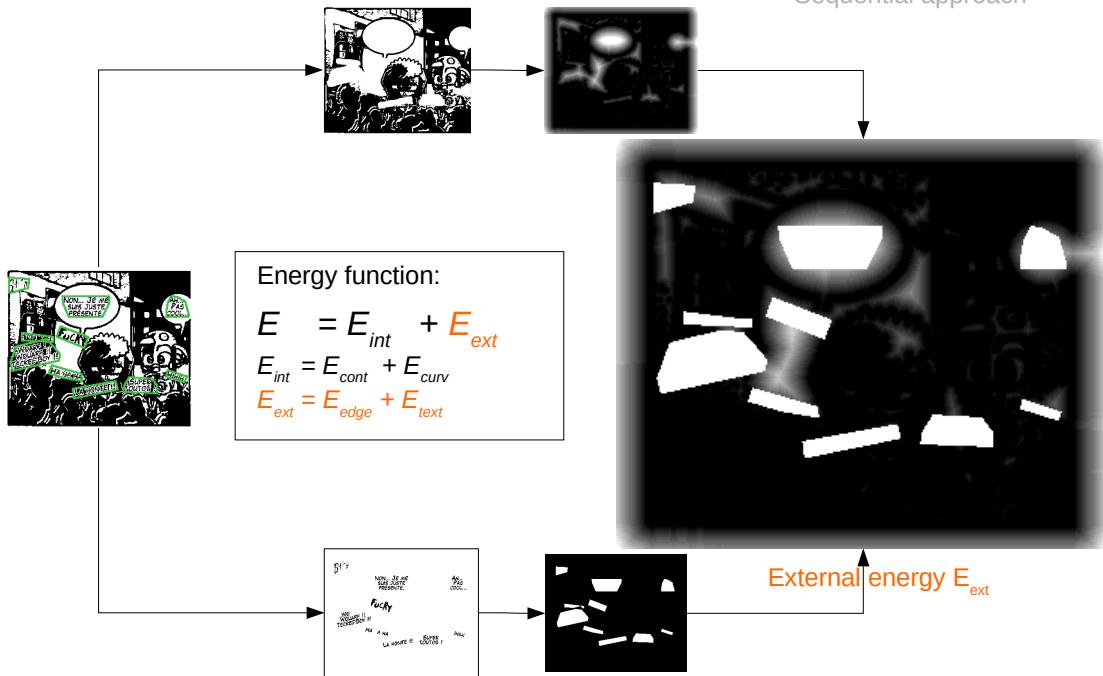


43

$E_{int} = \text{curvature and continuity}$
Our contribution: E_{text} from domain knowledge

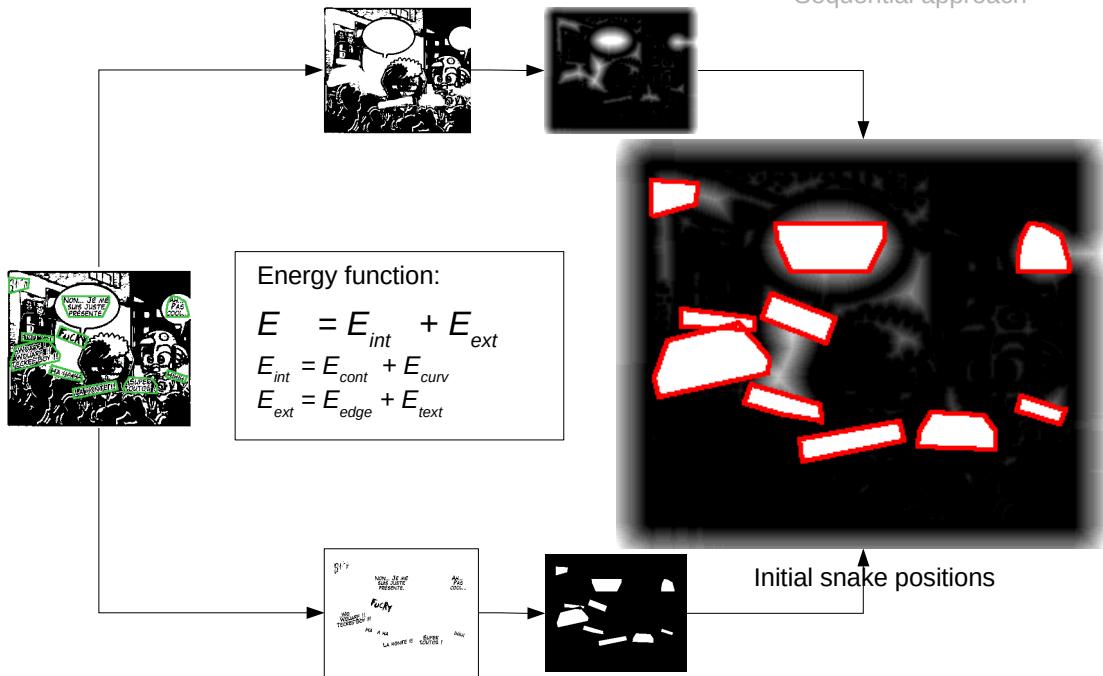
Balloon extraction: implicit

Contributions
Sequential approach



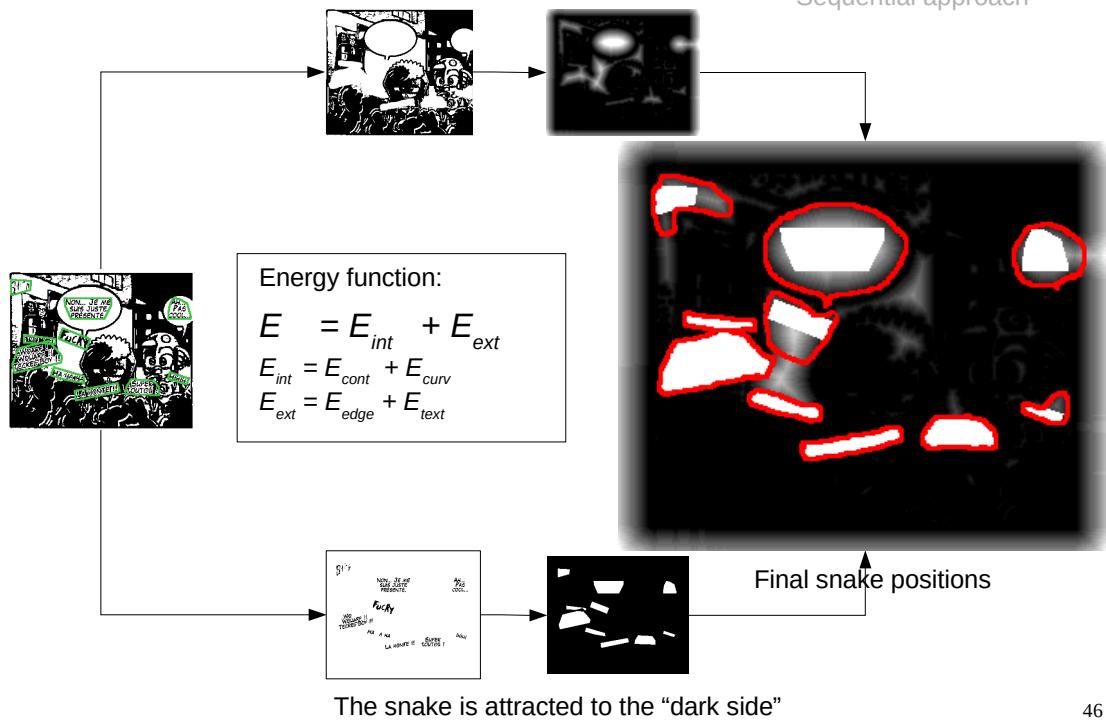
Balloon extraction: implicit

Contributions
Sequential approach



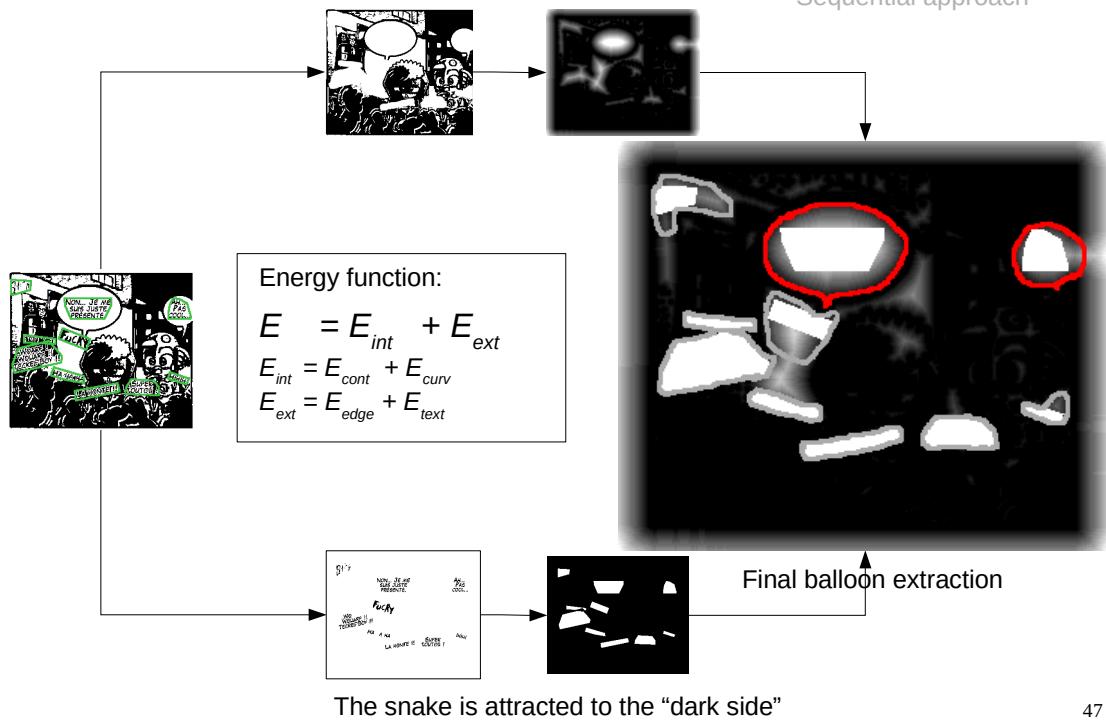
Balloon extraction: implicit

Contributions
Sequential approach



Balloon extraction: implicit

Contributions
Sequential approach



47

Valid detection if:

- snake did grow (diff between init/final)
- initial text block remains centred

Balloon extraction: results

Contributions

Sequential approach

Regular/implicit balloon extraction
examples

48

- Balloon extraction
 - Accurate for closed balloons
 - Implicit balloons would require more **contextual** information

- Introduction
- Sequential approach
 - Panel & text extraction
 - Balloon extraction
 - Tail extraction
 - Comic character extraction
- (Independent approach)
- Knowledge-driven approach

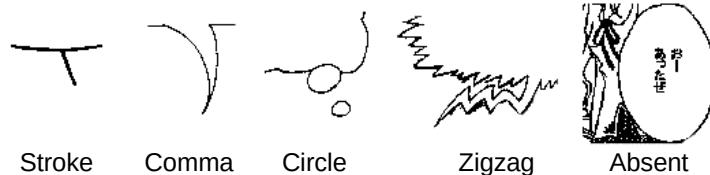
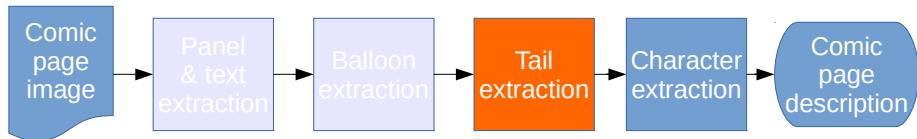


Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Tail extraction

Contributions

Sequential approach



- Literature
 - First time studied in document image analysis
- Objectives
 - Extraction of tail tip position and direction
 - Focus on comma, zigzag and absent types

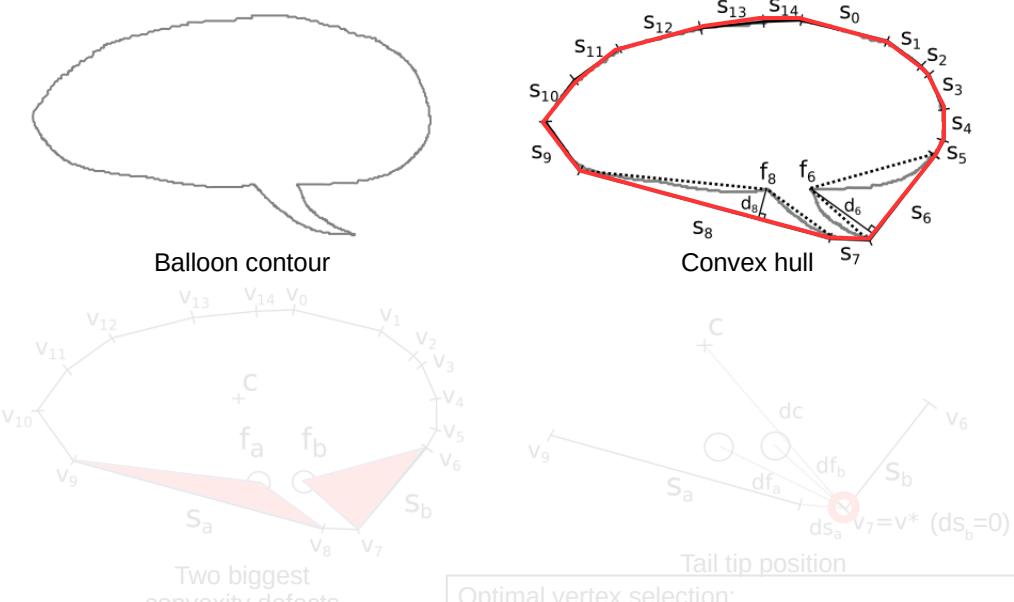
50

- Definitions
 - Tail: a pointer that indicates where is the emitter (speaker) of the balloon
 - Tail tip: the extremity of the tail, the closest part to the emitter
 - Tail direction: direction pointed by the tail

Tail extraction: tip position

Contributions

Sequential approach



52

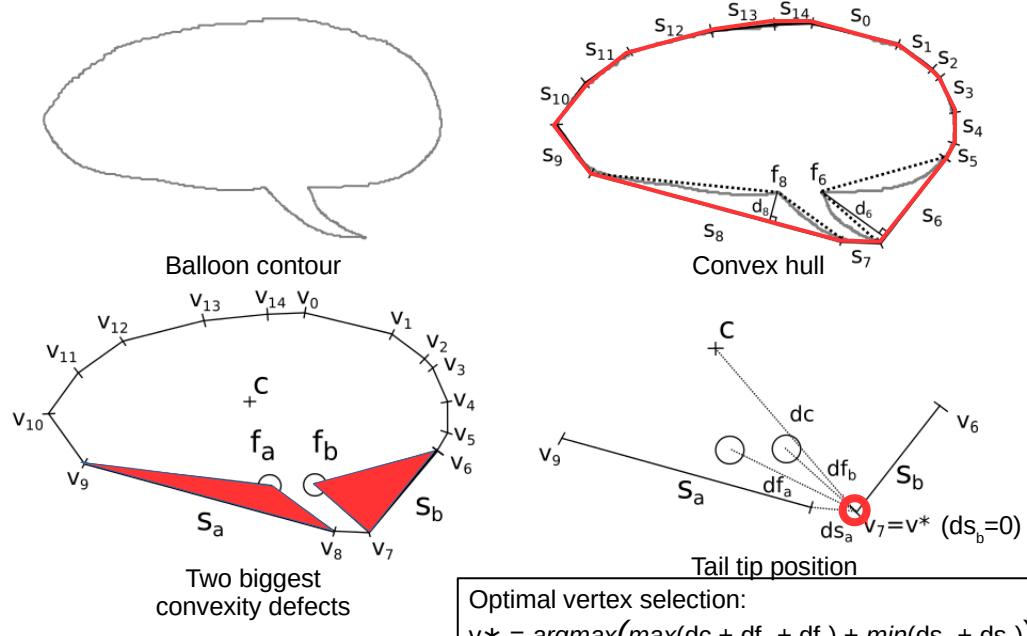
Justify contour analysis:

- Tail is in the extension of the contour
- Local deformation going outside of the balloon (convexity defect)

Tail extraction: tip position

Contributions

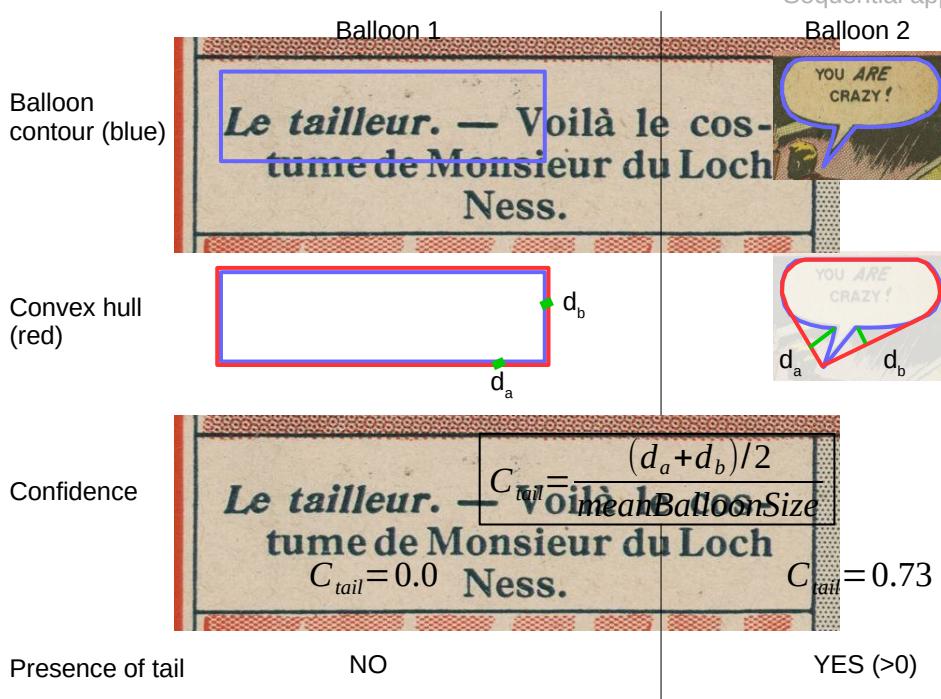
Sequential approach



Tail extraction: confidence value

Contributions

Sequential approach

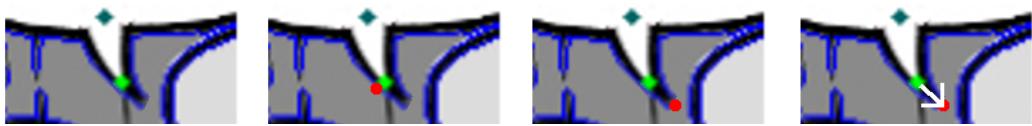


Tail extraction: tail direction

Contributions

Sequential approach

- Definition
 - Vector starting from “background” to “external edge” tail tip positions
- Approach
 - Extract external edge
 - Find external edge tail tip coordinates
 - Define the tail direction (N, NE, E, SE, S, SW, W, NW)



Background tail tip
(green) and
external edge (blue)

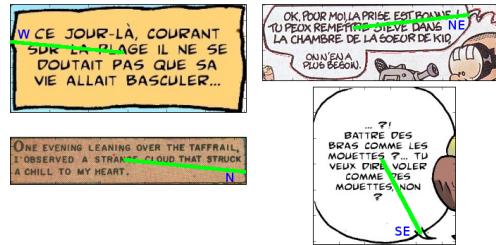
Closest point on
external edge
(red)

Farthest point
from origin and tip
(red)

Direction from tip
to farthest point
(white arrow)

Tail extraction: results

Experiments



TODO

56

- Tail extraction
 - Relies on the **balloon contour** extraction
 - Limited to **certain types** of tail

- Introduction
- Sequential approach
 - Panel & text extraction
 - Balloon extraction
 - Tail extraction
 - Comic character extraction
- (Independent approach)
- Knowledge-driven approach

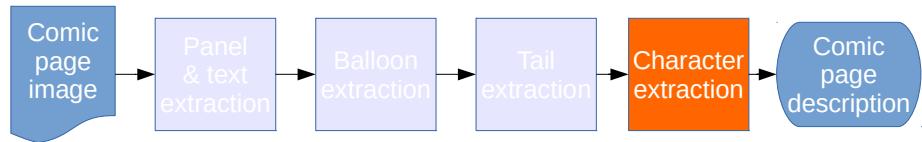


Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Comic character extraction

Contributions

Sequential approach



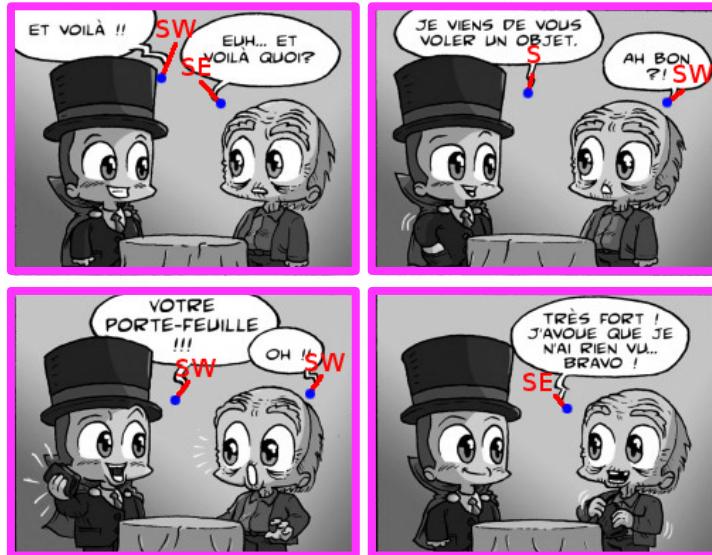
- Literature
 - Supervised approaches for **manga** and **cartoon** characters
 - No public **dataset** (copyright issues)
- Challenges
 - Variety of **styles** of comic books
 - Intra and extra class **variations** of each character **instance** (e.g. position, scale, pose, occlusion and human-like, invented)
- Objective
 - Unsupervised and **generic** approach for **all styles** of comic books

Comic character extraction

Contributions

Sequential approach

Panels + Tails = ?



Comic character extraction

Contributions

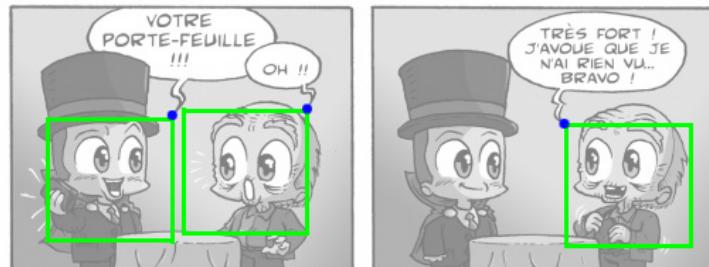
Sequential approach

Panels + Tails = Comic character ROIs

Large ROI



Small ROI



ADD results here

- Conclusion
 - Preliminary results
 - Highly relies on the quality of the tail tip and direction (error propagation)
 - Only for “speaking” characters
 - Implicitly retrieves the relationship between balloons and characters

- Introduction
- Sequential approach
- (Independent approach)
- Knowledge-driven approach
 - Introduction
 - Knowledge representation
 - Processing sequence



Inking. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Introduction

- High level image description
- Framework for comics understanding
- Independent element extraction
- Increase overall precision
- Collaboration with Clément Guérin

Contributions

Knowledge-driven approach

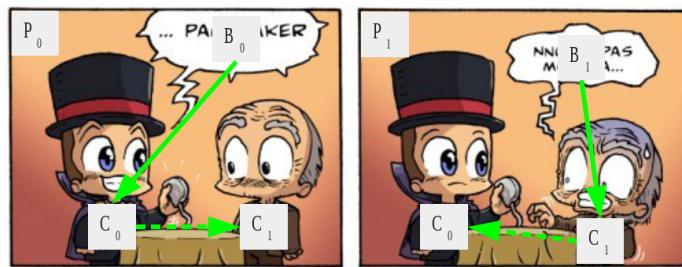
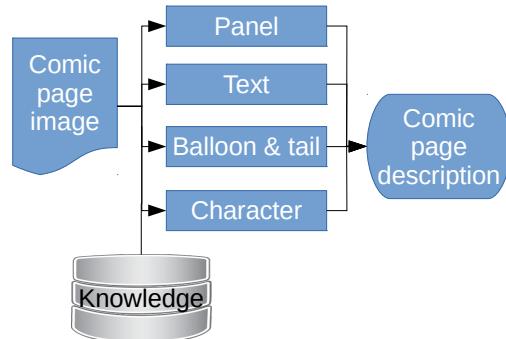


Illustration of high level description

63

Our intention is to increase the overall precision of the extractions and (unfortunately) not recall (because of removal only)

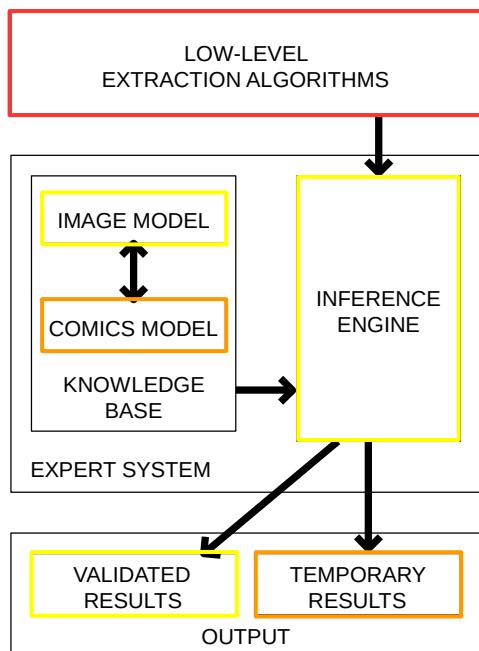
Knowledge representation

Contributions

Knowledge-driven approach

- Rigaud's thesis
- Collaboration
- Guérin's thesis

- Image model
 - Physical support
 - Regions of interest
- Comics model
 - Validations
 - A **panel P** is related to one page
 - A **balloon B** is related to one panel and may have a **tail Q**
 - A **character C** is related to one panel
 - A **text line T** is related to one balloon
 - Inferences
 - **B + Q + T => speech balloon SB**
 - **SB + T => speech text ST**
 - **SB + C => speaking character SC**

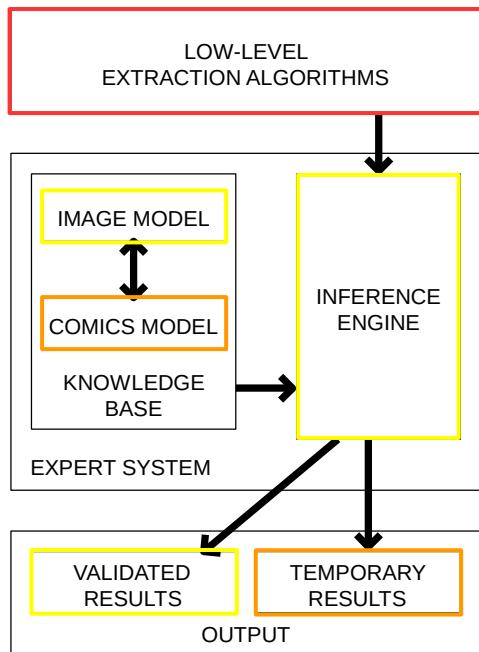


Knowledge representation

Contributions
Knowledge-driven approach

- Rigaud's thesis
- Collaboration
- Guérin's thesis

- Image model
 - Physical support
 - Regions of interest
- Comics model
 - Validations
 - A **panel P** is related to one page
 - A **balloon B** is related to one panel and may have a **tail Q**
 - A **character C** is related to one panel
 - A **text line T** is related to one balloon
 - Inferences
 - **B + Q + T => speech balloon SB**
 - **SB + T => speech text ST**
 - **SB + C => speaking character SC**



Processing sequence



Contributions
Knowledge-driven approach

Formulate hypotheses

Infer new information

- Iteration 1

- Step 1: hypotheses of simple element positions

- Step 2: validation of the positions

- Step 3: inference a new information

- Iteration 2

- Step 1: hypotheses of more complex elements

- Step 2: validation of the positions

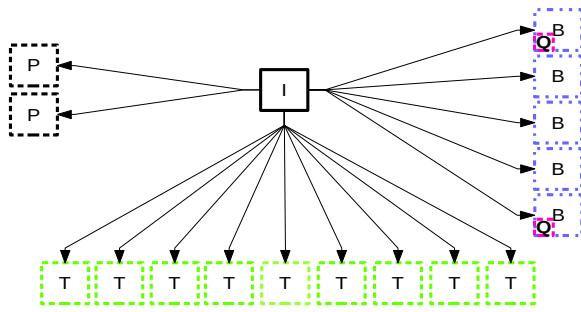
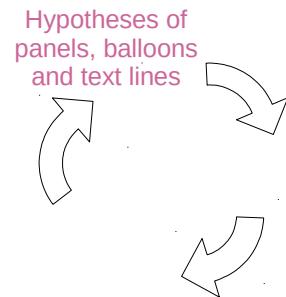
- Step 3: inference a new information

- ...



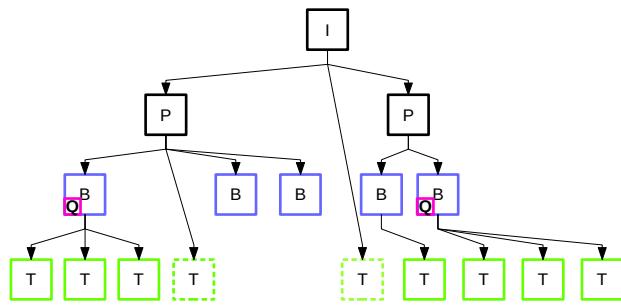
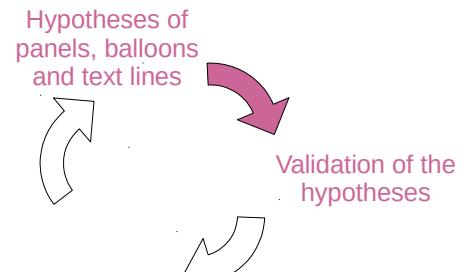
Processing sequence

Contributions
Knowledge-driven approach



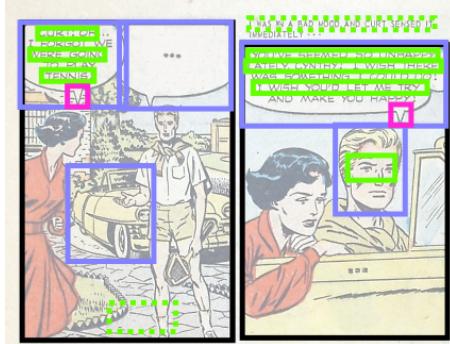
Processing sequence

Contributions
Knowledge-driven approach



Processing sequence

Contributions
Knowledge-driven approach

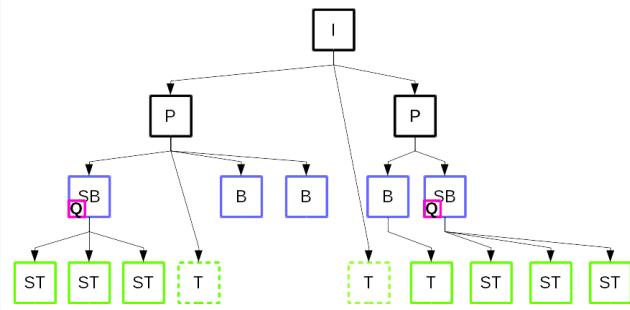


Hypotheses of
panels, balloons
and text lines



Validation of the
hypotheses

Inferences of
specific types

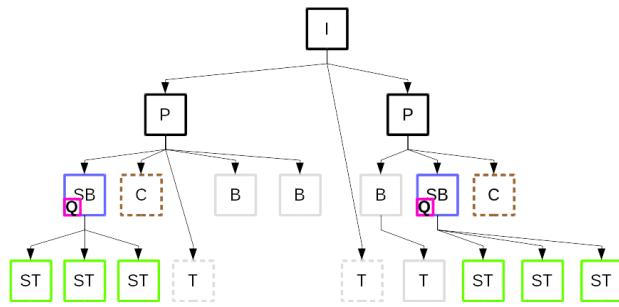
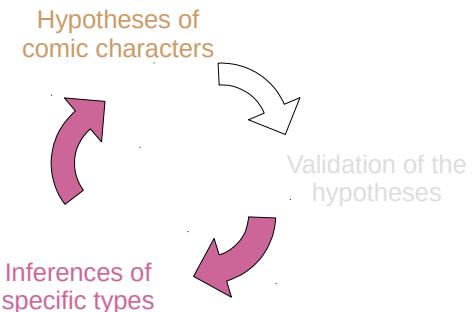


Processing sequence



Contributions

Knowledge-driven approach

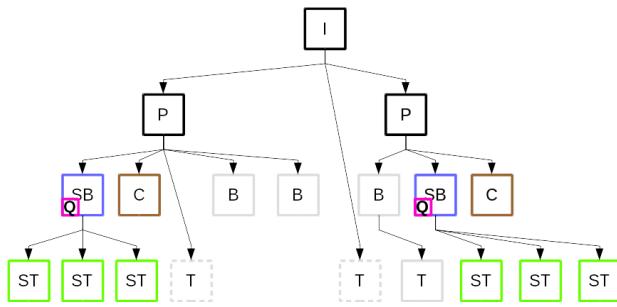
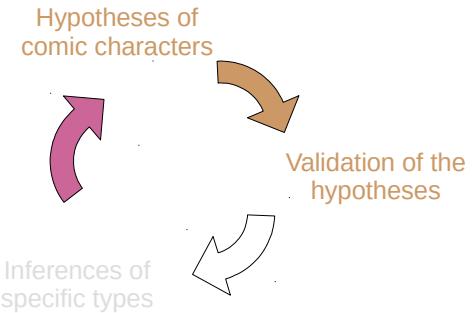


70

Processing sequence

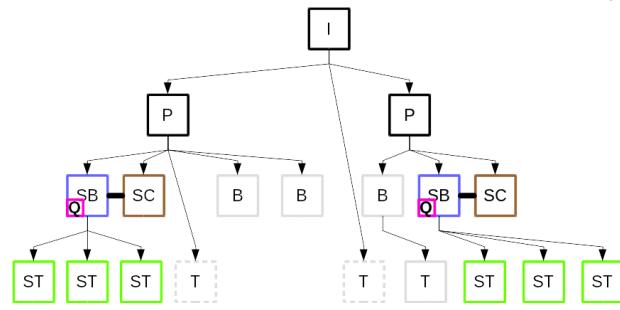
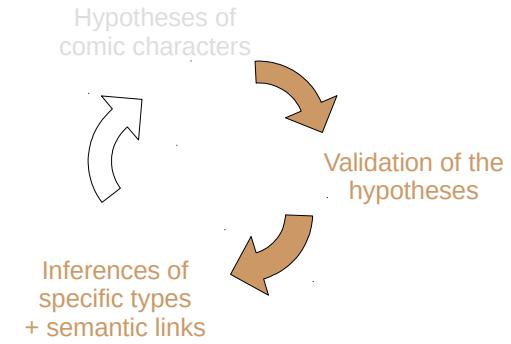


Contributions
Knowledge-driven approach



Processing sequence

Contributions
Knowledge-driven approach



Results

- Performances

Contributions

Knowledge-driven approach

TODO

73

- Contributions
 - Generic, unsupervised and expandable framework for comics understanding (pending acceptance IJDAR)
 - Formalization of image and comic book domains
- Limitations
 - Validation process (deletion only)

- Evaluations
- Overall contribution



Lettering, Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Evaluations

Experiments

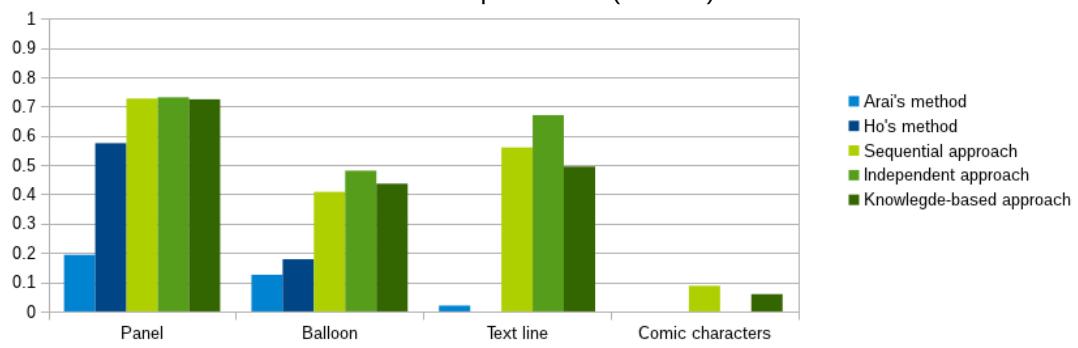
B_p = predicted region

B_{gt} = ground truth region

B_p is valid if $a_0 > 0.5$

$$a_0 = \frac{\text{area}(B_p \cup B_{gt})}{\text{area}(B_p \cap B_{gt})}$$

Extraction results on the eBDtheque dataset (F-score)



76

In blue method from the literature, in green our contributions.

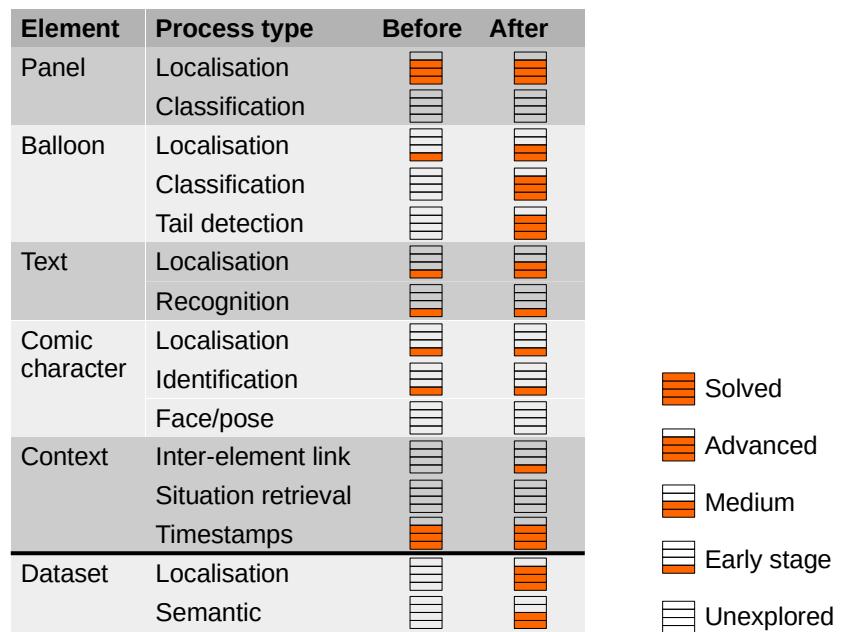
Detailed results can be found in the thesis.

Low comic character

- last step in the sequence
- speaking only (over all in GT)

Overall contribution

Experiments



- Global conclusions
- Global perspectives
- Publications



Lettering. Image credits: Le cycle des bulles,
Christophe Rigaud, 2012

Global conclusions

Conclusion

- Reached objectives
 - Efficient panel, balloon, text and tail extraction methods
 - First approaches for comic character extraction and context retrieval
 - Public dataset and ground truth (<http://ebdtheque.univ-lr.fr>)
- Publications
 - 1 journal, 2 book series , 4 conferences, 5 workshops (3 national)
 - 6 local seminars
- Research impacts
 - L3i is now a main actor of comic book analysis in Europe
 - New Ph.D. thesis started in 2013 (Nam Le Thanh)
 - Dataset used by international peers (Germany, India, China, Japan)
 - National projects (PIA BigData Actialuna/LIP6, ANR EXPION 2015)
 - International project on manga analysis (PHC-SAKURA with Japan)

79

Use “the last but not least” if needed

PHC = Partenariats Hubert Curien (Hubert Curien Partnerships)

Global perspectives

Conclusion

- Content extraction
 - Consider overlapping panel extraction
 - Investigate text recognition
 - Improve implicit balloon extraction and evaluation
 - Extract and identify non-speaking comic characters
- Content understanding
 - Situation retrieval (e.g. landscape, outdoor, night)
 - Action recognition (e.g. running, driving, dreaming)
 - Interaction retrieval (e.g. balloon said by/to)
 - Labelling from text analysis (e.g. auto tagging)
- Dataset
 - Increase the number of pages
 - Add more annotation (e.g. panel situation, character names and roles)
 - Annotate multi-parts of comic characters

Publications

Conclusion

JOURNAL

Christophe Rigaud, Clément Guérin, Dimosthenis Karatzas, Jean-Christophe Burie and Jean-Marc Ogier. “**Knowledge-driven understanding of images in comic books**”. International Journal on Document Analysis and Recognition (**IJDAR**), **2015** (accepted with minor reviews).

BOOK SERIES

Christophe Rigaud, Dimosthenis Karatzas, Jean-Christophe Burie and Jean-Marc Ogier. “**Adaptive contour classification of comics speech balloons**”. In Graphic Recognition. New Trends and Challenges. Lecture Notes in Computer Science (**LNCS**), Vol. 8746, **2014**.

Christophe Rigaud, Norbert Tsopze, Jean-Christophe Burie and Jean-Marc Ogier. “**Robust frame and text extraction from comic books**”. In Graphic Recognition. New Trends and Challenges. Lecture Notes in Computer Science (**LNCS**), Vol. 7423, pp. 129-138, **2013**.

Publications

Conclusion

CONFERENCES

Christophe Rigaud, Dimosthenis Karatzas, Jean-Christophe Burie and Jean-Marc Ogier. "**Color descriptor for content-based drawing retrieval**". In the Proceedings of the 11th IAPR International Workshop on Document Analysis Systems (**DAS**), pp. 267-271 , Tours, France, April, **2014**.

Christophe Rigaud, Dimosthenis Karatzas, Joost Van de Weijer, Jean-Christophe Burie and Jean-Marc Ogier. "**An active contour model for speech balloon detection in comics**". In the Proceedings of the 12th International Conference on Document Analysis and Recognition (**ICDAR**), pp. 1240-1244, Washington DC, USA, August, **2013**.

Clément Guérin, Christophe Rigaud, Antoine Mercier, Farid Ammar-Boudjelal, Karelle Bertet, Alain Bouju, Jean-Christophe Burie, Georges Louis, Jean-Marc Ogier and Arnaud Revel. "**eBDtheque: a representative database of comics**". In the Proceedings of the 12th International Conference on Document Analysis and Recognition (**ICDAR**), pp. 1145-1149, Washington DC, USA, August, **2013**.

Christophe Rigaud, Dimosthenis Karatzas, Joost Van de Weijer, Jean-Christophe Burie and Jean-Marc Ogier. "**Automatic Text Localisation in Scanned Comic Books**". In the Proceedings of the 8th International Conference on Computer Vision Theory and Applications (**VISAPP**), pp. 814-819, Barcelona, Spain, February, **2013**.

Publications

Conclusion

WORKSHOPS

Clément Guérin, Christophe Rigaud, Karel Bertet, Jean-Christophe Burie, Arnaud Revel and Jean-Marc Ogier. "**Réduction de l'espace de recherche pour les personnages de bandes dessinées**". In the Proceedings of the 19ème congrès national sur la Reconnaissance de Formes et l'Intelligence Artificielle (**RFIA**), Rouen, France, July, **2014**.

Christophe Rigaud, and Clément Guérin. "**Localisation contextuelle des personnages de bandes dessinées**". In the Proceedings of the 13ème Colloque International Francophone sur l'Ecrit et le Document (**CIFED**), pp. 367–370, Nancy, France, March **2014**.

Christophe Rigaud, Dimosthenis Karatzas, Jean-Christophe Burie and Jean-Marc Ogier. "**Speech balloon contour classification in comics**". Proceedings of the 10th International Workshop on Graphics RECognition (**GREC**), pp. 23-25, Bethlehem, USA, August, **2013**.

Hoang Nam Ho, Christophe Rigaud, Jean-Christophe Burie and Jean-Marc Ogier. "**Redundant structure detection in attributed adjacency graphs for character detection in comics books**". In the Proceedings of the 10th IAPR International Workshop on Graphics RECognition (**GREC**), pp. 109-113, Bethlehem, PA, USA, August, **2013**.

Christophe Rigaud, Norbert Tsopze, Jean-Christophe Burie and Jean-Marc Ogier. "**Extraction robuste des cases et du texte de bandes dessinées**". In the Proceedings of the 10ème Colloque International Francophone sur l'Ecrit et le Document (**CIFED**), pp. 349-360, Bordeaux, France, March **2012**.

References

Conclusion

[Arai10] Kohei Arai and Herman Tolle. **Method for automatic e-comic scene frame extraction for reading comic on mobile devices**. In Seventh International Conference on Information Technology: New Generations, ITNG '10, pages 370–375, Washington, DC, USA, 2010. IEEE Computer Society.

[Cheung08] S.C.S. Cheung, City University of Hong Kong. Run Run Shaw Library, and City University of Hong Kong. **Face Detection and Face Recognition of Human-like Characters in Comics**. Outstanding academic papers by students. Run Run Shaw Library, City University of Hong Kong, 2008.

[Chung07] ChungHo Chan, Howard Leung, and Taku Komura. **Automatic panel extraction of color comic images**. In HoraceH.-S. Ip, OscarC. Au, Howard Leung, Ming-Ting Sun, Wei-Ying Ma, and Shi-Min Hu, editors, Advances in Multimedia Information Processing - PCM 2007, volume 4810 of Lecture Notes in Computer Science, pages 775–784. Springer Berlin Heidelberg, 2007.

[Eunjung07] Eunjung Han, Kirak Kim, HwangKyu Yang, and Keechul Jung. **Frame segmentation used mlp-based x-y recursive for mobile cartoon content**. In Proceedings of the 12th international conference on Human-computer interaction: intelligent multimodal interaction environments, HCI'07, pages 872–881, Berlin, Heidelberg, 2007. Springer.

[Ho12] Anh Khoi Ngo Ho, Jean-Christophe Burie, and Jean-Marc Ogier. **Panel and speech balloon extraction from comic books**. In 2012 10th IAPR International Workshop on Document Analysis Systems, pages 424–428. IEEE, 2012.

84

Here are the references I mentioned throughout this presentation.

References

Conclusion

- [Khan12] Fahad Shahbaz Khan, Muhammad Anwer Rao, Joost van de Weijer, Andrew D. Bagdanov, Maria Vanrell, and Antonio Lopez. **Color attributes for object detection**. In 25th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012), 2012.
- [Li14a] Luyuan Li, Yongtao Wang, Zhi Tang, and Liangcai Gao. **Automatic comic page segmentation based on polygon detection**. Multimedia Tools Applications, 171–197, 2014, Kluwer Academic Publishers.
- [Li14b] Luyuan Li, Yongtao Wang, Zhi Tang, Xiaoqing Lu, and Liangcai Gao. **Unsupervised speech text localization in comic images**. In Proceedings of International Conference on Document Analysis and Recognition (ICDAR), pages 1190–1194, Aug 2013
- [Pang14] Xufang Pang, Ying Cao, Rynson W.H. Lau, and Antoni B. Chan. **A robust panel extraction method for manga**. In Proceedings of the ACM International Conference on Multimedia, MM '14, pages 1125–1128, New York, NY, USA, 2014.
- [Ponsard12] Christophe Ponsard, Ravi Ramdoyal, and Daniel Dziamski. **An ocr-enabled digital comic books viewer**. In Computers Helping People with Special Needs, pages 471–478. Springer, 2012.
- [Stommel12] Martin Stommel, Lena I Merhej, and Marion G Müller. **Segmentation-free detection of comic panels**. In Computer Vision and Graphics, pages 633–640, 2012.

85

Here are the references I mentioned throughout this presentation.

References

Conclusion

[Su11] Chung-Yuan Su, Ray-I Chang, and Jen-Chang Liu. **Recognizing text elements for svg comic compression and its novel applications**. In Proceedings of International Conference on Document Analysis and Recognition (ICDAR), pages 1329–1333, Washington, DC, USA, 2011.

[Sun10] Weihan Sun and Koichi Kise. **Similar partial copy detection of line drawings using a cascade classifier and feature matching**. In Hiroshi Sako, Katrin Franke, and Shuji Saitoh, editors, ICWF, volume 6540 of Lecture Notes in Computer Science, pages 126–137. Springer, 2010.

[Tanaka07] Takamasa Tanaka, Kenji Shoji, Fubito Toyama, and Juichi Miyamichi. **Layout analysis of tree-structured scene frames in comic images**. In IJCAI'07, pages 2885–2890, 2007.



Image credits: @carenaifaro



<https://github.com/crigaud/thesis/tree/master/presentation>
<http://www.christophe-rigaud.com>



- Authors, publishers, CIBDI
- Founding: Communauté d'Agglomération de La Rochelle
- Time for demo on easy case with comic character ROI computation?

Complementary information

History of comics art

- Pre-history: painting of animals and hunters in caves [Marx, 2007]
- 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 to social media sites and mobile devices



Pre-history and ancient Egypt



Rodolphe Töpffer, Histoire de Monsieur Cryptogame (1830)

Ted McCall, Robin Hood And Company (1946)



Tezuka Osamu, Manga Classroom (1953)

Marion Montaigne, Tu mourras moins bête (2013)

89



- Pre-history
- Ancient Egypt
- Rodolphe Töpffer is a Francophone Swiss artist
- Francis Lacassin, auteur du livre Pour un neuvième art, la bande dessinée¹⁷, mais ce livre ne date que de 1971, 7 ans après les articles de Spirou

GT validation

Knowledge-driven analysis detail

- Comic character region refinement

Production to interpretation

Background

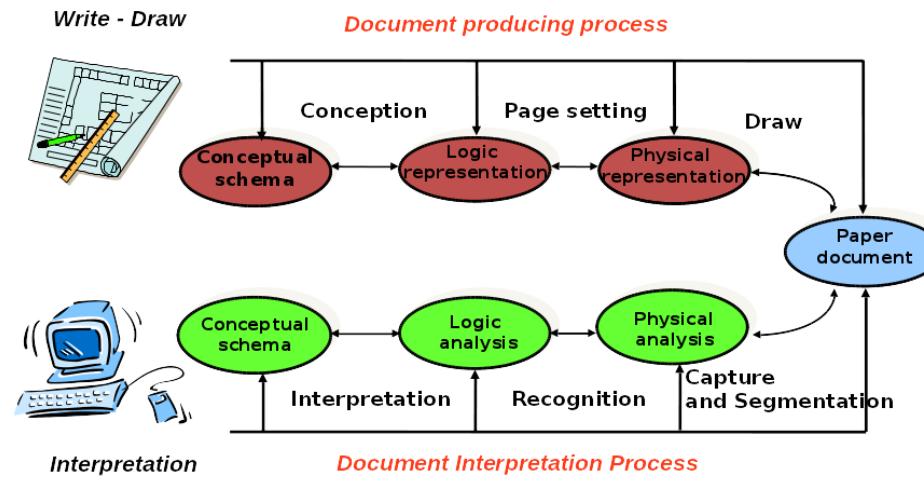


Image source: Handbook of Document Image Processing and Recognition. Springer, 2014

92

Low-level: pre-processing, feature extraction
Middle-level: classification, recognition
High-level: interpretation, understanding

Element	Process type	Analysis method	Method	Publication	Before	After	Δ
Panel	Localisation	Histogram	S	CIFED'12 LNCS'13	60%	70%	+10%
		Topology	I, K	IJDAR'15			
Balloon	Localisation	Topology	S	-	20%	50%	+30%
		Region	I, K	-			
		Edge	S	ICDAR'13			
	Classification	Contour	S, I, K	GREC'13 LNCS'14	0%	80%	+80%
Text	Localisation	Topology	S	-	5%	60%	+55%
		Structure	I, K	VISAPP'13			
	Recognition	-	-	-	10%	10%	0%
Comic characters	Localisation	Graph	-	GREC'13 (2nd)	5%	10%	+5%
		Context	S, K	CIFED'14 RFIA'14 (2nd)			
	Spotting	Colour	I	DAS'14			
	Identification	-	-	-	10%	10%	0%
	Face/pose	-	-	-	0%	0%	0%
Context	Inter-element	Proximity	S, K	CIFED'14 RFIA'14 (2nd) IJDAR'15	0%	10%	+10%

Background

Thinking about the scenario. Image credits:
Wissam Shekhani, 2010

