Understanding of comic images from pixel to semantics

A dissertation submitted by **Christophe Rigaud** at Universitat Autònoma de Barcelona to fulfil the degree of **Doctor of Philosophy**.

Bellaterra, ???, 2014

UNIVERSITÉ DE LA ROCHELLE

ÉCOLE DOCTORALE S2IM

Laboratoire Informatique, Image et Interaction (L3i)

THÈSE présentée par :

Christophe RIGAUD

soutenue le : **[XX mois en lettres 2014]**
pour obtenir le grade de : **Docteur de l'université de La Rochelle**
Discipline : **informatique et applications**

Understanding of comic images from pixel to semantics

[Sous titre éventuel]

JURY :

Prénom NOM	Professeur, Université xxxxxx, Président du jury
Prénom NOM	Directeur de recherche CNRS, Université xxxx, Rapporteur
Prénom NOM	Professeur, Université xxxxxxx, Rapporteur
Prénom NOM	Professeur, Université de La Rochelle, Directeur de thèse
Prénom NOM	Professeur, Université xxxxxxxxxxx
Prénom NOM	Maître de conférences, Université xxxxxxxxxxxx
Prénom NOM	Titre, établissement

Director	**Prof. Dr. Jean-Christophe Burie** Laboratoire Informatique, Image et Interaction Université de La Rochelle
Co-Directors	**Dr. Dimosthenis Karatzas** Centre de Visió per Computador Universitat Autònoma de Barcelona **Prof. Dr. Jean-Marc Ogier** Laboratoire Informatique, Image et Interaction Université de La Rochelle
Thesis committee	**?** ? ? **?** ? ? **?** ? ? **?** ? ? **?** ? ?
European evaluators	**?** ? ? **?** ? ?

This document was typeset by the author using LaTeX 2$_\varepsilon$.

The research described in this book was carried out at the Laboratoire Informatique, Image et Interaction, Universtité de La Rochelle and at the Centre de Visió per Computador, Universitat Autònoma de Barcelona.

ISBN: XXX

Printed by Ediciones Gráficas Rey, S.L.

To my parents...

An idea that is not dangerous is
unworthy of being called as an idea at all
Oscar Wilde (1854 - 1900)

Live as if you were to die tomorrow.
Learn as if you were to live forever
Mahatma Gandhi (1869 - 1948)

Acknowledgement

- Supervisors + L3i + ULR, CVC (DAG group) + UAB
- eBDtheque project manager and participant, Antoine Mercier for his programming expertise and the development of ToonShop (annotation tool and demonstrator)
- eBDtheque dataset material contributors (see acknowledgement section of eBDtheque ICDAR'13 publication)
- Anh Koi for his first work on the comics at L3i
- Jean-Baptiste Fasquel, Laurent Hardouin et Jean-Louis Boimond, François Chapeau-Blondeau, David Rousseau (engineering school and research master, first works of image processing)
- Family

Abstract

TODO

Resumen

TODO

Resum

TODO

Contents

Acknowledgement **i**

Abstract **iii**

Resumen **v**

Resum **vii**

1 Introduction **1**
1.1 The evolution of the "bandes dessinée" 1
1.2 Motivations 2
1.3 Objectives and contributions 4
1.4 Outlines 6

2 State-of-the-art **7**
2.1 Panel extraction and layout analysis 7
2.2 Balloon segmentation and tail detection 8
2.3 Text extraction and recognition 9
2.4 Comic character detection 9
2.5 Holistic understanding 10
2.6 Existing applications 11
2.7 Conclusion 12

3 Dataset and ground truth **13**
3.1 Introduction 13
3.2 Dataset description 14
3.2.1 Term of use 15
3.3 Ground truth construction 16
3.3.1 Visual annotation 16
3.3.2 Semantic annotation 19
3.3.3 File structure 20
3.4 Ground truth quality assessment 23
3.5 Conclusions 26

4 Panel extraction **27**
4.1 Introduction . . . 27
4.2 Methodology (contribution) . . . 27
4.3 Experimental results . . . 27
4.4 Conclusions . . . 27

5 Balloon segmentation and classification **29**
5.1 Introduction . . . 29
5.2 Methodology . . . 29
5.2.1 Localization . . . 29
5.2.2 Segmentation & classification . . . 29
5.2.3 Tail detection and description . . . 30
5.3 Experimental results . . . 30
5.4 Conclusions . . . 30

6 Text localization and recognition **31**
6.1 Introduction . . . 31
6.2 Methodology . . . 31
6.2.1 Text localization . . . 31
6.2.2 Text recognition . . . 31
6.3 Experimental results . . . 31
6.4 Conclusions . . . 31

7 Comic character spotting and clustering **33**
7.1 Introduction . . . 33
7.2 Methodology . . . 33
7.2.1 Detection and spotting . . . 33
7.2.2 Clustering . . . 33
7.3 Experimental results . . . 33
7.4 Conclusions . . . 33

8 System for document understanding **35**
8.1 Introduction . . . 35
8.2 Methodology . . . 35
8.2.1 Framework . . . 35
8.2.2 Application to comics . . . 35
8.3 Experimental results . . . 36
8.4 Conclusions . . . 36

9 Conclusions **37**
9.1 Summary and contributions . . . 37
9.2 Future perspectives . . . 37

Bibliography **41**

List of Tables

List of Figures

2.1 Flow diagram of comic balloon detection using comic blob extraction method proposed by Arai [6]. 8
2.2 Examples of comics character postures. This example shows deformation, pose, rotation, translation and occlusion variations. Image credits: Prunelle, la fille du cyclope, Vicky Portail-Kernel and Cédric Kernel, Ankama, 2010. 10

3.1 Distribution of the number of elements per text lines. 15
3.2 Panel annotation 17
3.3 Speech balloon contour annotation 18
3.4 Text line location annotation 18
3.5 Comic character position annotation 19
3.6 Balloon shapes 20
3.7 Annotation rendering in a browser 21
3.8 Distance to the mean position 24
3.9 Error measurement image 25

Chapter 1

Introduction

1.1 The evolution of the "bandes dessinée"

Setting/positioning the scenario, place comics in the addressed context, origin of comics/BD until new usages in 2014 created by new technologies. [54, p. 215]

- The origin of BD, comics, Manga, other
- see Clement's intro
- collectif PANIC book
- Évolution de la bande dessinée en France depuis quinze ans `http://ejournals.library.ualberta.ca/index.php/af/article/download/21313/16112`
- `http://fr.wikipedia.org/wiki/Bande_dessinee_en_ligne`
- `http://en.wikipedia.org/wiki/Comics`
- European Comic Art `http://journals.berghahnbooks.com/eca/`
- Webcomics `http://www.phdcomics.com` and `http://xkcd.com`
- Le manga: Une synthèse de référence qui éclaire en image (Eyrolles, 2013) `http://books.google.fr/books?id=3MqgAgAAQBAJ&printsec=frontcover&v=onepage&q&f=false`
- `http://en.wikipedia.org/wiki/Manga`
- Etat présent BANDE DESSINEE STUDIES LAURENCE GROVE UNIVERSITY OF GLASGOW `http://fs.oxfordjournals.org/content/68/1/78.full.pdf+html`
- Définition de la bande dessinée interactive `Bande_dessinee_interactive_Tony_Rageul-part1.pdf`

- 1.1.5 La bande dessinée et le numérique (intro thèse Clément)
- IDPF (ePub3 http://www.figoblog.org/node/2014 and http://idpf.org/idpf-comics-manga-workshop-paris and http://idpf.org/digital-book-
- Public domain http://digitalcomicmuseum.com/
- Thesis of Julien Falgas and Cohn
- Terminology of comics ~/Biblio_LINK/Bandes_dessin\unhbox\voidb@x\bgroup\let\unhbox\voidb@x\setbox\@tempboxa\hbox{e\global\mathchardef\accent@spacefactor\spacefactor}\accent19e\egroup\spacefactor\accent@spacefactores/WritingforAnimation,Comics,andGames,chapter6
- Evolution of the story book ~/Biblio_LINK/Bandes_dessin\unhbox\voidb@x\bgroup\let\unhbox\voidb@x\setbox\@tempboxa\hbox{e\global\mathchardef\accent@spacefactor\spacefactor}\accent19e\egroup\spacefactor\accent@spacefactores2007_WritingforAnimation,Comics,andGames_CHAPTER5_History_Evolutionofthecomicbook
- De la page à l'écran 2010_Boudissa_La_Bande_dessinee_entre_page_et_ecran.pdf
- Link between time and space Bandes_dessinees/2012_THESIS_Cortsen_ComicsasAssemblage_HowSpatio-TemporalityinComicsisConstructed.pdf
- Check /home/crigau02/Bureau/PhD/News/BD
- Define comics vocabulary (see Clement's thesis and master student report) for panel, balloon, text, character

S'il est difficile de définir avec précision la bande dessinée, c'est qu'elle se situe précisément au carrefour de plusieurs moyens d'expression artistique: l'art graphique, l'art cinématographique et la littérature. Elle est tout à la fois dessin, cinéma, écriture, se conjuguant entre eux pour former un art nouveau, doté d'un ensemble de moyens d'expressions extrêmement complet et varié [...] [21]

Generic part: mixed content document or semi-structured documents

1.2 Motivations

The needs? numbers? impact on the society?

- http://fr.wikipedia.org/wiki/Bande_dessinee#Aspects_.C3.A9conomiques
- La lecture des bandes dessinées en france http://www.culturecommunication.gouv.fr/Politiques-ministerielles/Etudes-et-statistiques/Les-publications/Collections-de-synthese/Culture-etudes-2007-2014/

La-lecture-de-bandes-dessinees-CE-2012-2 and 2012_Evans_Lalecture
unhbox\voidb@x\bgroup\let\unhbox\voidb@x\setbox\@tempboxa\
hbox{e\global\mathchardef\accent@spacefactor\spacefactor}\
accent19e\egroup\spacefactor\accent@spacefactores_CE-2012-2.
pdf

- Japanese manga market http://www.animenewsnetwork.com/news/2013-12-0
top-selling-manga-in-japan-by-series/201413
- The Comics Chronicles (US market) is a free resource for academic research: http://www.comichron.com/ and http://comicsbeat.com/category/sales-charts/
- See motivations in PhD/Publications/Doctoral_school_L3i_first_year_phd_report
- IGS-CP (http://www.igs-cp.fr), a content extraction companies working for digital comics promoting (e.g. espritBD, alterComics), ITEsoft
- Present the solution: eBDtheque project from L3i and its founding supports (region, Europe)
- Present the actors of eBDtheque project, especially Clement's work topic

Comics or "bande dessinée" represents an important part of the cultural heritage of many countries, especially in the US [38, 78], western Europe (particularly France and Belgium) [67], and Japan [7]. Unfortunately, they did not yet received the same level of attention as music, cinema or literature about their adaptation to the digital format. Indeed, while the latter entered to the common digital uses via the proliferation of music services and video on demand and quasi-systematic output of books in paper and electronic book formats, millions of works from the imagination of the comics authors still struggled to find an echo on the side of digitized world. Ancient works could be reused with information technology to explore digital libraries [8], assist translators [10], augmented reading [40, 75], enhancing accessibility [11, 65], story analysis, advertising etc. Nevertheless, the process of conversion and adaptation is not as simple and straightforward as the one established for the digital publication of films and novels. The comics differ from the latter in that the media itself is intimately linked to the medium. Indeed, a film can be decomposed into a series images plus a soundtrack. Just watching these images in the right order and at the right frame rate allow to reconstruct the initial content, regardless of the medium. In the same way a novel is ultimately a sequence of words. Reading these words in the correct order, on paper or on a screen does not change neither the content nor the artistic dimension of the work. The "bande dessinée" is defined as juxtaposed sequences of image by McCloud [54] and Thomas [89].

However, it differs from the films on the form and the spatial positioning of the images. Where the latter pictures are all of equal size and each new image replaces the previous one, comic panels vary in size and spatial organisation in a limited space (paper sheet). These two features, added to the fact that the reader has the

opportunity to see all the boxes of a same page, but not those of the next page, are tools at the service of the author to stage the story. Therefore, changing the medium, the reading surface format or the sequence order involves a modification of the staging that may in some cases be detrimental to the story.

Several companies offer printed to digital format conversion services for comics, to facilitate reading on mobile devices. However, the conversion process is both tedious because done by hand, and simplistic as it is too often reduced to the successive display of panel interspersed with user selected transitions. The ideal would be to understand the process used by authors to draw the paper version of the comics and automatically change it in a form adapted to the medium in which the work is read (e.g. smartphone, web page, 3D book). It starts with an analysis of the digitalized paper page to extract the different components (e.g. panel, balloon, text, comic character) and their relations (e.g. read before, said by, addressed to). Once this initial work is done, it is necessary to reconstruct the story by placing the extracted elements in the initial order to keep the story coherent.

Why is this scientifically relevant?

The analysis of comic images by computer is particularly challenging because they are semi-structured document with mixed content. Comic document are at the intersection between unstructured (e.g. teaching board [63], free-form document [20]) and complex background (e.g. advertising poster [15], real scene [23,61,96]) images which are nowadays active fields of research for the community. Being at the intersection of several fields of research increase the complexity of the problem. This is one of the reason why the analysis of comics is a recent (in the document analysis history) and not solved field of research.

1.3 Objectives and contributions

Comics contain many heterogeneous elements that are hard to process in once. Our objective is to process them separately, from simple to complex, in order to progressively build a complete comics understanding system. First, we construct a public dataset of comic images and the corresponding ground truth in order to evaluate our work and to give to the community the opportunity to work on identical data in order to make comparable and reproducible research. Second, we focus on extraction panels, texts, balloons in order to facilitate the extraction of more complex elements by focussing on a region of interest instead of the whole image. Third, we combine the extraction processes to reconstruct the context and the relations between the elements. The dataset and this last work are the result of a collaboration with Clément Guérin, Ph.D. student working on the same project.

To meet the above objectives, we have made the following contributions in this thesis.

1) Dataset and ground truth: The eBDtheque dataset is the first publicly avail-

able[1] dataset and ground truth of comic images. Such dataset is important for the community to make comparable, reproductive and growing research. The dataset consist in a mixture comic images coming from different albums with the goal of being as representative as possible of the comics diversity. The database consists of a hundred pages of various comic book albums including Franco-Belgian "bande dessinée", American comics and Japanese mangas. The ground truth contains the spacial position of panels, balloons and text lines, comic characters and their associated semantic annotations, that appear in the images. Also, bibliographic information are given such for each image. This work is presented in Chapter 3.

2) Panel extraction: Comics are mixed content documents that require different techniques to extract different element. The first particularity of comics is the sequence of panels that we extract using connected component classification and filtering in Chapter 4.

3) Balloon detection: Balloons or bubbles are key elements in comics, they link graphical and textual elements and are part of the comics style. They can have various shapes (e.g. oval, rectangular) and contours (e.g. smooth, wavy, spiky, absent). In this work we propose a closed balloon extractor based on the analysis of the blob content, a active contour model to extract open balloons from text line position. Both methods are described in Chapter 5 along with balloon contour classification (smooth, wavy, zigzag) and a tail detector.

4) Text localization: Text is of different nature in comics, there are sound effects (onomatopoeias), graphic text (illustration), speech text (dialogues) and narrative text (captions). Speech text represents the majority of the text present in comics [REF?], we propose a adaptive binarisation process from a Minimum Connected Component Thresholding followed by a text/graphic separation based on contrast ratio and then a text line grouping algorithm. Finally, an OCR system filters out non text region. This work is explained in Chapter 6

5) Comic character detection: Unsupervised comic character extraction is a difficult task as soon as we aim to process heterogeneous comic styles using the same algorithm. In this context, learning-based approaches are not reliable unless if we have enough data to train on all comic styles. We first propose a query by example approach that asks the user to select a part of the object he is looking for in one comics image and the system spots other occurrences everywhere in all the pages of the comics album, assuming that they have been digitized under the same conditions. Second, we go one step further by refining the comic character location according to the contextual elements (e.g. panel contents, speech balloon position, tail direction) given a region of interest. These works are presented in Chapter 7.

6) Comics understanding: Making a computer understand the complete story of a comics is a really challenging task, especially because it is even hard for human sometimes. Putting comics domain knowledge in an ontology-based framework

[1]Dataset website: `http://ebdtheque.univ-lr.fr`

enable to interact between image processing and semantic information in order to progressively understand the content of a document. This work is introduced in Chapter 8.

1.4 Outlines

The rest of the thesis is organized as follows:

- Chapter 2 presents a detailed review of the state-of-the-art methods for the analysis of comic images. This chapter details several image processing methods in the four first subsections and then we review the holistic understanding systems that have been applied to document analysis so far and the more advanced application on the market.
- Chapter 3 presents the dataset and ground truth of comic images that we provided to the community. The idea of this dataset consists in a mixture images coming from different albums with the goal of being as representative as possible of the comics diversity. It also describe the indexation structure in a Scalable Vector Graphics (SVG) format.
- Chapter 4 introduces a fast panel extraction method by classifying the connected components into three classes for panel, text and noise. This method can also be used for text/graphic separation.
- Chapter 5 introduces both closed and non closed balloon localisation and segmentation methods together with contour classification and tail detection and description.
- Chapter 6 addresses the difficulty of text localization and recognition in comics. It propose speech text localisation method using connected components alignment and neighbourhood similarity to form text lines. Text recognition is also addressed and preliminary results are detailed.
- Chapter 7 presents two methods for comic character detection. The first method overcomes the difficulties of such non rigid object detection thanks to the proposed descriptor invariant to scale, object deformation, translation and rotation transformations. The second method takes profit of the contextual information to define the region of interest of comic characters.
- Chapter 8 presents a system that combines low and high level processing to build a scalable system of comics image understanding reaching the level of semantic interaction between elements. This work is a collaboration with Clément Guérin, Ph.D. student working on data mining applied to comic documents.
- Chapter 9 concludes the thesis and defines the future direction of comics document analysis.

Chapter 2

State-of-the-art

Comics images are mixed content documents that are processed differently depending on the purpose. The involved techniques can vary a lot between panel, balloon, text and comic character extractions. The contents of different nature are related to each other to produce a story. Treating each of the content separately has a limit that can be exceeded in a holistic understanding approach by reaching a higher level of semantic. We review specific comics image content processing, holistic understanding and the existing state-of-the-art applications related to comics.

2.1 Panel extraction and layout analysis

Panel extraction and ordering has been mainly studied for panel to panel reading. The need is increasing since the first generation of mobile devices with small screens in colour or B&W. In fact, people want to continue reading their favourite comics or manga on the way, without caring kilos of books. Printed comics require tedious work to be manually scanned and split into screen size parts small enough to avoid zooming and scrolling.

Several techniques have been proposed to automatically extract panels as [39], assuming that panels are small enough elements to be comfortably read on mobile devices. They are based on white line cutting algorithm [12, 22, 46], recursive X-Y cut [31] or gradient [88]. Those methods do not consider empty area [39] and border free panels. These issues have been corrected by connected component approaches [5] but they are sensible to region that sometimes connect several panels and increase the detection error rate. Another approach based on morphological mathematics and growing region [33] can remove such connecting elements but also remove information on the panel border. After the region segmentation step, heuristic filtering is often applied [6, 33] to classify panel region according to the size ratio according to the page size, which depends on the page format. More recently, new methods have shown interesting results for manga and European comics with different background

colours. They are based on watershed [66], line segmentation using Canny operator and polygon detection [47], region of interest detection [79] such as corners [91] and line segments. Panel retargeting have been addressed for manga by Matsui [53].

Page layout analysis have been studied to calculate the reading order of the panels. The page layout influences the reader at choosing pathway [17], nevertheless few studies [4,28,65] demonstrated the possibility of calculating such Z-path (left-to-right and down) or right-to-left (e.g. Arabic, Japanese) and down [48,91] according to the position of the panels.

2.2 Balloon segmentation and tail detection

Balloons or bubbles are the visual unit that conveys dialogue, either spoken or thought. Balloons have developed into a more-or-less oval shape, with a pointer or tail to indicate to which character they belong. There are many specialized forms of balloons, either traditional or invented [52]. Few work about balloon detection have been done until now and mainly closed speech balloon have been studied. Arai [6] 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 (see figure 2.1). Another connected component approach proposed by Ho [34] uses HSV color 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 percent.

From our knowledge, the analysis of open balloons, balloon contour and tails have not been studied before, we propose a first approaches in Chapter 5. Those extra information about the balloons are also important for a broad understanding of the dialogues and emotions [60].

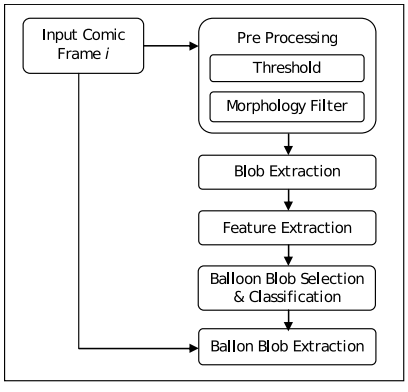

Figure 2.1: Flow diagram of comic balloon detection using comic blob extraction method proposed by Arai [6].

2.3 Text extraction and recognition

Text extraction and recognition have attracted of lot a attention in document analysis such as newspaper, administrative documents, cheques, maps, floor plans and engineering drawings. Techniques from real scene text recognition can sometimes be transposed to comics images, especially the ones applied to car plate recognition [3] because the text is in a salient and contrasted area with a complex background around it such as speech balloon in comics.

Few works concern text extraction in comics speech balloons, bottom-up approaches use connected component which often relies on the segmentation step [66]. Su [80] uses Sliding Concentric Windows as text/graphic separation and then apply mathematical morphology and an SVM classifier to classify text from non-text components. Li [49] proposed an unsupervised speech text localization for comics that trains a Bayesian classifier on aligned connected component and then detect the rest of the text using the classifier for text/non text separation.

Top-down approaches starting from balloons (white blobs) detection followed by mathematical morphology operations have been proposed by Arai [6], Yamada [97] and Sundaresan [85]. From our knowledge, there are no work concerning graphic sounds (onomatopoeia) and illustrative text extraction.

Text recognition applied to comics is really a challenging because it includes most of the difficulties from text recognition in document analysis domain if we consider all the type of text that composed the comics. From typewritten to handwritten and free-form text in uniform to complex background including image noise, text deformation and overlapping. Nevertheless, Ponsard [66] solved a sub part of the problem by focusing on speech text of a single font and language for which an OCR system is trained for.

2.4 Comic character detection

Human detection in computer vision field has been largely studied during the past decades, mainly based on grayscale image and gradients. Color information is rarely relevant for human detection because of clothing and skin color difference. In videos, moving regions are often used as region of interest.

Although, comics are often a reproduction of human life situations, it is a domain where we can not directly apply human detection based methods. The main difference is that comic characters (e.g. protagonist, hero) are hand drawn and therefore more variant in terms of deformation, shape and appearance than real life humans (see fig. 2.2). In colored comics, the color information gives the identity of characters and plays a main role for character spotting with speech balloon positions. The simplistic character design allows for easy identification and representation and goes along with how human process visual information [1, 16, 57]. This is a big difference compared to natural images and human detection, since comics are designed in a way that the

information (e.g where are the comic characters, who is talking, where is it going) can be quickly found.

Figure 2.2: Examples of comics character postures. This example shows deformation, pose, rotation, translation and occlusion variations. Image credits: Prunelle, la fille du cyclope, Vicky Portail-Kernel and Cédric Kernel, Ankama, 2010.

Recent studies have been published for partial manga copy detection [82], mainly based on shape information because of the absence of color information. This work has been extended to Manga copyright infringement protection [83] using a faces [93] as region of interest and concentric HOG [19] as region descriptor. This work shows good results for manga part retrieval which is a subset of the whole comics world.

Colored comics may be compared to cartoon images sequence, for which a first work based on HOG, SVM and color attributes has been published in 2012 [41]. Preliminary work about cartoon and comics faces recognition have been carried out by Kohei [87] and Cheung [13]. More recently, graph theory has been used to find redundant color structure in order to automatically localize the most frequent color group apparition and label them as main characters [35]. The thesis of TA [86] (section 1.1 and 1.2) gives a good overview about stroke-based image analysis similar to comics, it defines the issues of poor information, occlusion, deformation, inter-class and intra-class variations, scale, spatial and/or temporal relations and structured data. TA [86] also mentions that "story board scene understanding still remains an open problem, few results are available in the literature about stroke images".

Another work uses HOG descriptor with redundant information classification to also find the most frequent elements [84]. Both graph and descriptor based methods need an image preprocessing step to remove irrelevant redundant elements such as text and speech balloons.

Other interesting approaches tries to automatize comic generation [90, 95] using image cartoonification and script analysis.

2.5 Holistic understanding

One of the original goals of image or graphical document analysis was to fully understand the content of any image [44]. This requires solving several sub-tasks simultaneously, for instance region detection, labeling of meaningful regions and semantic understanding using layout analysis. In the past, researchers have developed classi-

fiers for tackling each of these sub-tasks independently [51]. However, these sub-tasks can help each other. For instance, in a comics page, if we know the panel positions, then we can make a better guess at the location of the comics characters (they are usually inside the panels). However, it is not easy to combine different related sub-tasks together. Previous works concern real scene image analysis [9], retrieval [72] and understanding [25,45], medical image annotation using description logic and inference engine [36], object-based image retrieval [59, 71] (between keyword-based and query-by-example) and image interpretation [37, 62] that mentions the importance of using topological information, distances, directional relative position and much complex relations such as "between" and "surround" and "among". Also, Geographic Object-Based Image Analysis (GEOBIA) [9] makes extensive use of ontologies to interpret maps.

Recently, comics book images have been also considered. An ontology of comics has been proposed from a philosophical approach [58], a semantic annotation tool [32] makes use of previous knowledge and consistency information to suggest new knowledge to the user in a interactive way. Spatial inferences have been used to infer the comic books reading order, for panel in the page and balloon in the panel [28]. In [72], they highlight the benefit of using contextual information of simple object to build more complex ones.

2.6 Existing applications

Comic books are a graphic art form combining text and images to tell a story. Comics are now used in a wide variety of styles, not only on paper (e.g., magazines, newspapers, TV show) but also as electronic content (smartphone apps, e-books, websites). Comics are one of the most popular and familiar forms of graphic content. People read comics easily and learn many things, so even children can learn about cultures and trends, among other things, through comics even unconsciously.

From this background, we can divide comic-based computer systems into two categories. One type would be using comics to represent complex information such as online communication in a form of comics [42] and generating video or story log summaries [2, 73, 92] (computer graphics domain). Also, it is possible to listen to manga [74] which have been recorded by people reading the story or to use mobile app for automatic translation [26,77] (requires user text selection). These systems are useful to add value to the content while making it more funny and interesting. The other category concerns the comic design by enabling novices to create comics interactively using a computer for augmenting an individual user's memory [81], turning photo albums into a comic [14,64], making comic-like video [68], making collaborative comics [50] and exchanging rich message [70].

2.7 Conclusion

To conclude the literature review, some of the challenges of comics image analysis can be highlighted from the above state-of-the-art reviews. First, comics image suffer from noise as any other document image processing. It comes from hardware processes (e.g. drawing techniques, printing, digitization) and software (e.g. image compression). So efficiently handling noise is crucial for image analysis and understanding. Knowing the design process of the comics creation helps for image denoising. Second, according to the results of the reviewed methods, we can order them by level of difficulty, from the simplest to the hardest: panel, balloon, text, comic character and holistic understanding of an image or album (add comparison table about genericity, processing time, complexity???).

Third, most of the work in the literature use different copyrighted images that they are not allowed to share publicly and do not share their code neither. This is a key issue for researchers which can not share, reproduce and compare results on identical data in a collaborative way. This is one of the reason that retains comics analysis to progress as fast as other field of research of document analysis.

In the next chapter, we are going to present the first comics image dataset and ground truth publicly available for researchers. It is a selection of hundred pages from more than twenty different albums from America, Japan and Europe. The construction of the dataset and ground truth and a quality assessment test are detailed.

Chapter 3

Dataset and ground truth

In this chapter we propose and detail the construction of the first publicly available comics image dataset and ground truth. The comics image were selected with the agreement of consenting authors and publishers that had the objective to foster innovation in this domain trough academic research. The ground truth is defined in accordance to existing formalism in order to fulfil the needs of a large amount of researchers related to comics material.It integrates low and high level information such as the spatial position of the elements in the image, their semantic links and also bibliographic information.

3.1 Introduction

With the analysis and processing of data comes the need of the output results evaluation. Traditionally, this evaluation is made by validating the results of an algorithm with a ground truth that represents what an ideal output should be [24, 27, 76]. Ideally, such a ground truth is made publicly available so anyone can challenge his own algorithm to the community [43]. This can be applied to any kind of results from image segmentation to classification or information retrieval. Being in need of comic books material and an associated ground truth to evaluate our work, we noticed that there is not such dataset publicly available for scientific purpose. Therefore, we decided to gather the first publicly available comic books dataset in association with several comic books authors and publishers and to build up the corresponding ground truth according to document analysis and understanding concerns.

It took almost one year to define which type of comics to use, meet and convince comics authors and publishers, get copyright authorizations for the scientific community, develop a specific annotation tool and finally to hire people to do the ground truth.

A selection of hundred comic pages were annotated in one day by twenty vol-

unteers affiliated to the L3i lab. In order to provide a common basis for evaluating research work, the ground truth have been published in 2013 [30] and made available to the scientific community via the project website [1]. It has been enhanced in 2014 by adding semantic information to the already annotated elements.

The content of the dataset, the ground truth construction protocol and its quality assessment are detailed in the next section.

3.2 Dataset description

Scott McCloud defined comics as "juxtaposed pictorial and other images in deliberate sequence, intended to convey information and/or to produce an aesthetic response in the viewer" [55]. This definition is intentionally broad enough to encompass the spectrum of the majority of works produced so far. The dataset composition should reflects this heterogeneity to give everyone the opportunity to compare its algorithms to a globally representative dataset of the comics world. We contacted authors with different comic styles and have selected a corpus of one hundred images, representing single or double comics page.

The images were partly processed by the French company A3DNum (`http://www.a3dnum.fr`) which was commissioned to digitize 14 albums. Among all the files, scanned at a resolution of 300 dots per inch and encoded in uncompressed Portable Network Graphic (PNG) format, we used 46 pages to integrate the eBDtheque corpus. The remaining 54 images were selected from webcomics, public domain comics[2] and unpublished artwork with different styles from 72 to 300 dots per inch. We encoded all the images of the eBDtheque dataset in Joint Photographic Experts Group (JPEG) format with a lossy compression to facilitate file exchange.

Hereafter we describe the characteristics of the selected albums and their content.

Albums published between 1905 and 2012. 29 pages were published before 1953 and 71 after 2000. Quality paper, colour saturation and textures related to printing technique changes can vary a lot from one page to another. The artworks are mainly from France (81%), United States (13%) and Japan (6%). Their styles varies from classical Franco-Belgium "bandes dessinées" to Japanese manga through webcomic and American comics.

Pages themselves have very diverse characteristics. Among all, 72 are printed in colours and according to the authors and periods, there are a majority of the tint areas, watercolours and hand-coloured areas. Among the remaining 28, 16 have are greyscale and 12 are simply black and white. One album has two versions of each page, one in colour and the other one in black and white. We have integrated an

[1] http://ebdtheque.univ-lr.fr

[2] http://digitalcomicmuseum.com

examples of each of them in order to allow performance comparison of algorithms on the same graphic style by using colour information or not. Five of the 100 images are double page, others are single page and 20% are not A4 format.

Panels contained in the pages are of various shapes. Although most of them are bounded by a black line, a significant proportion has at least one part of the panel which is indistinguishable from the background of the page (frame less panel). Two pages consist only of frame less panels, the visual delimitation uses background contrast difference between the panel and image. Nine images contain overlapping panels, twelve contain only panels without border and several has panels connected by an straddle object.

Balloons also contain a great diversity. Some of them are completely surrounded by a black stroke, some partially and others not at all. They have a bright background with a rectangular, oval or non geometric shape with "smooth", "wavy" or "spiky" contour in general. Most of them has a tail pointing towards the speaker, but some do not. There is text without any surrounding balloons on 33 images of the corpus.

Text is either typewritten (61% of the image) or handwritten, mainly upper-case. The text lines contains 12 elements in average (see figure 3.1) and there are more than hundred text lines that are composed by only one letter which are punctuation or single letter words such as "I" or "A".

Most pages are from French artworks, where the text is written in French. Only 13 pages contain English text and 6 images are in Japanese. Onomatopoeia appears in 18 pages.

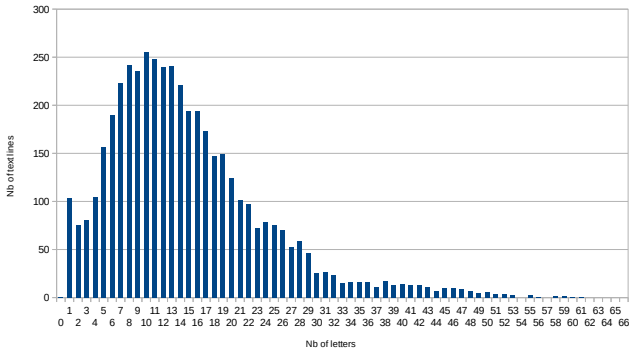

Figure 3.1: Distribution of the number of elements per text lines.

Comic characters or protagonist are specific to each album. They all have eye, harm and leg but at least 50% are not humanoid, depending on the interpretation.

3.2.1 Term of use

We obtain the minimum rights for sharing and publishing image material from the right holders but we had to make sure the user accept it before using the data. In collaboration with the intellectual property department of the University of La Rochelle, we established the following:

In order to use this database, you must firstly agree to the terms. You may not use the database if you don't accept the terms. The use of this database is limited to scientific and non-commercial purpose only, in the computer science domain. For instance, you are allowed to split the images, through the use of segmentation algorithms. You can also use pieces of this database to illustrate your research in publications and presentations. Any other use case must be validated by our service. If you do agree to be bound by all of these Terms of Use, please fill and email the request form and then use the login and password provided to download the selected version below.

The concerned request form requires the identity, affiliation, address and intended use of the person who wishes to use data.

3.3 Ground truth construction

In order to cover a wide range of possible research matters, it has been decided to extract three different type of objects from the corpus: text lines, balloons and panels. We decided to do this first ground truth by drawing horizontal bounding boxes as close as possible from the feature and including all its pixels. We chose this level of granularity in order to limit the subjectiveness of the person making the annotation.

The comics art is extremely heterogeneous and our dataset voluntarily integrates albums that can be classified as unconventional. This leaves room for interpretation on the form which increase annotation variations by different people and decrease the uniformity of the ground truth. This precision level is used in several, widely used datasets [24, 98].

Hereafter we detail the two levels of annotation (visual and semantic) that form the ground truth and how they are indexed in a file. The combination of visual and semantic annotation provides the advantage of making this ground truth relevant for document analysis and semantic evaluation which are both part of the comics understanding process.

3.3.1 Visual annotation

The first annotation consist in defining the spacial region where elements are located in the image. We describe hare how the visual annotation have been performed for the panels, balloons, texts and comic characters.

Panels The frame or panels are defined as an image area, generally rectangular, representing a single scene in the story. There is always at least one panel per page, the entire page region can be used as panel if necessary. When a panel has a black border, the bounding box is placed as close as possible to its frame. Sometimes, images have not been scanned perfectly horizontally, it is then impossible to have an horizontal bounding box sticking exactly to the border. When the panel border is partially absent or suggested by the neighbourhood, the bounding box just defines the content of the panel. In all cases, the other elements (balloon, text, drawings) extending from the frame are truncated, see figure 3.2.

Figure 3.2: Example of three panel annotations. The bounding box (transparent red) is defined without taking into account of non panel elements in all cases.

Balloons We define a balloon (phylactery or bubble) as the region of an image including one or several lines of text, graphically defined by an identifiable physical boundary or suggested by the presence of an arrow pointing to the speaker (the tail). Although rare, empty balloons (not containing lines of text) are also annotated if they are clearly identifiable by their shape or the tail representation. Pixel level annotation follows the contour of the balloon (see figure 3.3c), while bounding box annotation does not consider the contour of phylactery and truncates the tail. Sometimes it crosses the entire panel and generates an unrepresentative position of the desired balloon (see figure 3.3a). When the balloon is not closed (e.g. open contour) the annotated contour has to stick as close as possible to the contained text (see figure 3.3b). Note, the first version of the ground truth (2013) have been defined at bounding box level ignoring the tail and the second version (2014) at pixel level following the contour variations and the tail region.

Text lines The text lines are defined as a sequence of text characters aligned in the same direction (see figure **??**a). This definition encompasses both speech texts and narrative text, often located inside balloon, onomatopoeia (graphic sound) that are written or drawn directly in the panel without particular container. Comics are static graphics, the expression of emotions of a comic character is the joint action of drawing and text, sometimes in the form of a single punctuation symbol. For instance, an exclamation mark for surprise or a question mark for a misunderstanding. These isolated symbols convey information and are segmented as text line as well (see

Figure 3.3: Example of balloon clipping: a) using bounding box excluding the tail, b) bounding box of non closed balloon, c) pixel level contour annotation.

figure **??**b). Similarly, we have chosen to include in this category the illustrative text, such as a road sign or storefront (see figure **??**c). Although at the boundary between text and graphic, these elements are still invariably read by the reader and their annotation is potentially interesting for multiple purposes, including story and scene analysis.

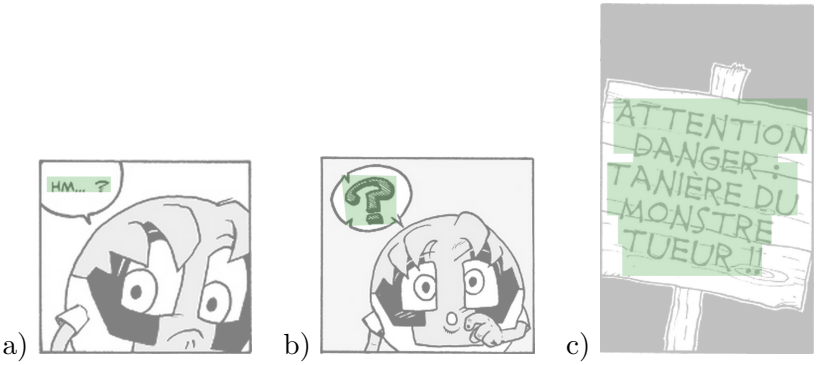

Figure 3.4: Different examples of text line position annotation: a) a classical text line in a speech balloon, b) a unique symbol, c) illustrative text, non horizontal.

Comic characters The comic characters position have been included in the second version of the ground truth only (2014). The concept of "character" may have different interpretations when used for comics and must be specified. Characters in a comic have not necessarily a human-like, or even living beings appearance. Even so, it would be appropriate to annotate every human being appearing in a box while

some are nothing but a part of the scenery. Therefore, we have chosen to limit the annotation to the comic characters that emits at least one speech balloon in the album (minimal impact in the story). Their bounding box has been defined to maximize the region occupied by the comic character inside the box region. Therefore, some parts of the character such as arms or legs, are clipped sometimes (see figure **??**).

Figure 3.5: Example of comic character annotation: sniper's arms are not included in the bounding box in order to maximize the region occupied by the sniper in its bounding box. The two snipers and the two characters in the car are annotated because they emit a speech balloon in a different panel.

3.3.2 Semantic annotation

This second level of annotation complete each spacial region with additional information about its semantic. Also, the image itself is annotated with extra information about its origin (e.g. album, collection, author, publisher).

Images The image, often assimilated as pages, has been annotated with bibliographical information, so that anyone using this ground truth is free to get is own paper copy of the comic books for extra uses. The first annotation is the page number (`pageNumber`) then the comic book title, from which the page has been picked up, and its release date (`albumTitle, releaseDate`), the series it belongs to (`collectionTitle`), the authors and editor names (`writerName, drawerName, editorName`) and, finally, the website and/or ISBN (`website, ISBN`). The album title is not mandatory for webcomics. Structural information about the page content has been added as well, such as resolution (`resolution`), reading direction (`readingDirection`), main language of the text (`language`) and single or double page information (`doublePage`).

Panels The panes are annotated with a `rank` metadata which stand for its position in the reading sequence. The first panel to be read on a given page has its rank property set to 1, while the last one is set to n, where n is the number of panels in the page.

Balloons Balloons are also annotated with a `rank` property that defines their reading order relatively to the image because balloons are not always included in panels. For a page containing m balloons, the first balloon's rank will be 1 and the last will be m. A second information concerns the `shape` of the balloon. This feature conveys an information about how the contained text is spoken (tone). The type of `shape` is given from the following list {`smooth, wavy, spiky, suggested`} as pictured in figure 3.6. Finally, the tail tip position (extremity) and its pointing direction have been added into the second version of the ground truth. There are given through the `tailTip` and `tailDirection` properties. The possible values of the direction are reduced to the eight cardinal directions plus a ninth additional value for the lack of tail: {`N, NE, E, SE, S, SW, W, NW, none`}. In the second version of the ground truth (2014), we added the identifier of the comic character which is emitting the balloon `idCharacter`.

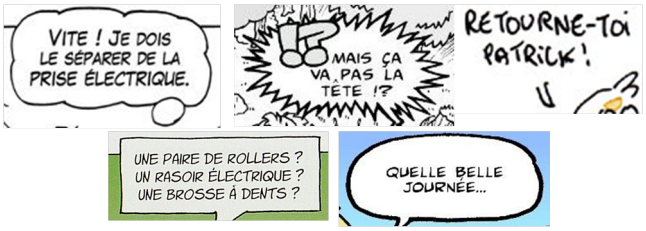

Figure 3.6: Les différents types de contour de bulle, de gauche à droite et de haut en bas : nuage, hérissé, suggéré, rectangle, ovale.

Text lines The text lines are associated with their transcription and the identifier of the corresponding balloon is added in `idBalloon` if the text line is included in a balloon.

Comic characters The comic character are identified by `idCharacter` in order to be easily referred.

3.3.3 File structure

The ground truth file structure have been thought according to comics related formalism such as Comics Markup Language (ComicsML) [56], Comic Book Markup Lan-

guage (CBML) [94], Periodical Comics[3], A Comics Ontology [69], Advanced Comic Book Format (ACBF)[4] and the Grand Comics Database (GCD)[5]. See the Ph.D. thesis of Guérin [29] for an extended review.

As we wanted to keep the ground truth file system simple and easy to share, visual and semantic annotations about a given page are gathered in a single full-text file following the specifications of Scalable Vector Graphics (SVG). Besides being an open-standard developed by the World Wide Web Consortium (W3C) since 1999, the SVG format fulfils two essential needs for this database.

First, using a recent Internet browser or your favourite image viewer, it provides a simple, fast and elegant way to display the visual annotation of any desired object over a comic book page using layers. No need to install software such as Matlab, Adobe Illustrator or equivalent open source to visualize the ground truth information.

It is XML-based vector image format that allows to display an animate the annotated region, stored as polygon object in the SVG file, as desired using the Cascading Style Sheets (CSS) properties, see figure 3.7.

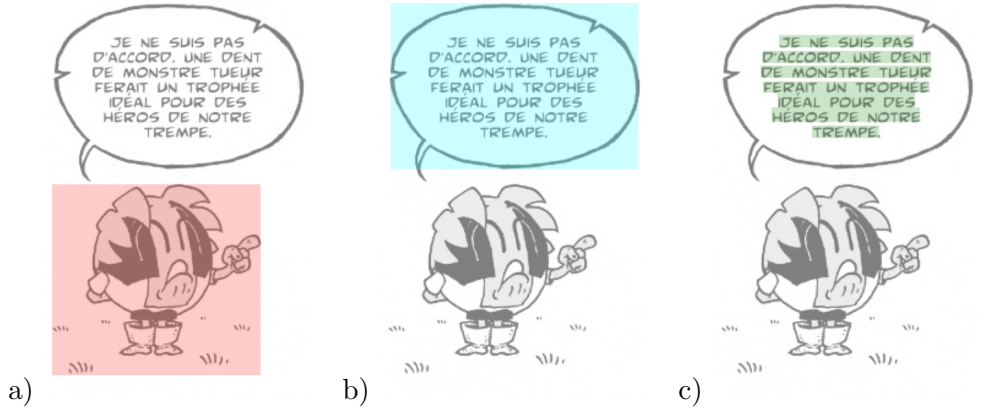

Figure 3.7: Example of rendering for each class of element. For example, red for panels (a), cyan for balloons (b) and green for text (c). The opacity is set to 50% to allow seeing the corresponding image by transparency.

Each layer can be displayed or not in order to enhance the clearness of the annotations when browsing the database. Secondly, SVG being a XLM-based language, it makes the integration of semantic annotation very easy via the use of the predefined `metadata` element.

One ground truth file contains the complete description of one comics image. There is no hierarchical link between pages from a same comic book. Following

[3] `http://www.w3.org/wiki/WebSchemas/PeriodicalsComics`
[4] `https://launchpad.net/acbf`
[5] `http://www.comics.org`

the basic XML encoding information, a SVG file starts with a root <svg> element containing the title of the document, <title>, and five <svg> children with different class attributes. These contain annotations collected on five types of elements which are the page, panels, balloons, text lines and comic characters. The type of element in a tag is specified by its class attribute. The first tag, class = ``Page'' contains description on the image and has two daughters. The first one, image has several attributes which specifies a link to the image file in the dataset xlink: href and two specifying the width and height of the image. The second, metadata, contains bibliographic information about the album and page properties described in 3.3.2. The four following <svg> siblings, <svg class=``Panel''>, <svg class=``Balloon''>, <svg class=``Line''> and <svg class=``Character''> respectively contain the annotations about panels, balloons, text lines and comic characters. They all contain SVG <polygon> elements with a list of five or more points in a point attribute that define the position of the bounding box corners or the pixel-level contour. Note that the last point is always equal the first one to "close" the polygon according to the SVG specifications. Those points are used by the viewer to draw polygons with the corresponding CSS style. Each <polygon> has a <metadata> child to store information on the corresponding polygon, according to the attributes list described in 3.3.2.

An example of ground truth file is given figure 3.1.

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<svg>
  <title>CYB_BUBBLEGOM_T01_005</title>
  <svg class="Page">
    <image
        x="0"
        y="0"
        width="750"
        height="1060"
        href="CYB_BUBBLEGOM_T01_005.jpg"
    />
    <metadata
        collectionTitle="Bubblegom_Gom"
        editorName="Studio_Cyborga"
        doublePage="false"
        website="http://bubblegom.over-blog.com"
        albumTitle="La_Legende_des_Yaouanks"
        drawerName="Cyborg_07"
        language="french"
        resolution="300"
        ISBN="979-10-90655-01-0"
        readingDirection="leftToRight"
        writerName="Cyborg_07"
        releaseDate="2009"
        pageNumber="5"
```

```
    />
  </svg>
  <svg class="Panel">
    <polygon points="53,95 268,95 268,292 53,292 53,95">
      <metadata rank="1" />
    </polygon>
    ...
  </svg>
  <svg class="Balloon">
    <polygon points="61,103 143,103 143,172 61,172 61,103">
      <metadata
        idBalloon="B01"
        idCharacter="C01"
        shape="smooth"
        tailDirection="SE"
        rank="1"
      />
    </polygon>
    ...
  </svg>
  <svg class="Line">
    <polygon points="373,121 432,121 432,132 373,132 373,121">
      <metadata idBalloon="B01">
        LIKE YOU.
      </metadata>
    </polygon>
    ...
  </svg>
  <svg class="Character">
    <polygon points="84,153 261,153 261,298 84,298 84,153" />
        <metadata idCharacter="C01" />
    ...
  </svg>
</svg>
```

Listing 3.1: Example of SVG file of the eBDtheque ground truth

- Comic Book Markup Language `http://digitalhumanities.org/dhq/vol/6/1/000117/000117.html`

3.4 Ground truth quality assessment

When several persons are involved in the creation of a graphical ground truth, it is very difficult to obtain a perfectly homogeneous segmentation. Indeed, it could

vary from one person to another because each person has a different sensitivity at reading comics and at integrating instructions. Therefore, in addition to the package of pages he was in charge of, each participant has been asked to annotate the panels of a extra page. This extra page was the same for everybody and was chosen for its graphical components heterogeneity. It contained ten panels from which, four were full-framed, five half-framed and one was frame less. We defined an acceptable error for the position of a corner given by several persons. The images of dataset being of different definitions, using a percentage of the page size makes more sense than using a specific number of pixels. We set this percentage p at 0.5% of the page height and width in x and y. Given the definition of the test image of 750x1060 pixels, this makes a delta of +/- 5 pixels in y axis and +/- 4 pixels in x axis.

We asked to each one of the twenty involved persons to draw the four points bounding box of the panels ignoring text area. A mean position from the twenty different values has been calculated for each of them. Then, the distance of each point to its mean value is computed. Figure 3.8 shows the amount of corners for a distance, centred on zero.

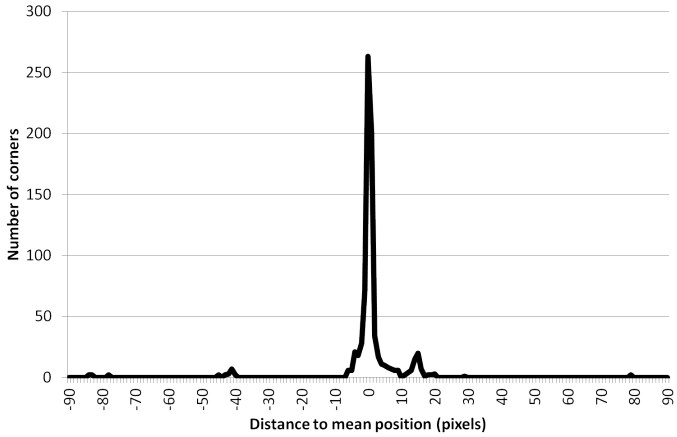

Figure 3.8: Number of corners for a given standard deviation value. This has been calculated on y axis and x axis and produces similar plot.

Given the threshold $p = 0.5$, 87.5% of pointed corners can be considered as being homogeneous over the group of labelling people. The overall mean standard deviation on this page reaches 0,15% (1.13 pixels) for the width, and 0.12% (1.28 pixels) for the height. The two bumps, at -40 and 15, are related to the missegmentation of 13 of the 80 panels. Indeed, instructions have been misunderstood by some people who included text area outside of the panels or missed some panel's parts. Figure 3.9 shows the difference between areas labelled as a panel by at least one person and areas labelled as a panel by every participant. However, such mistakes have been manually corrected before publishing the ground truth.

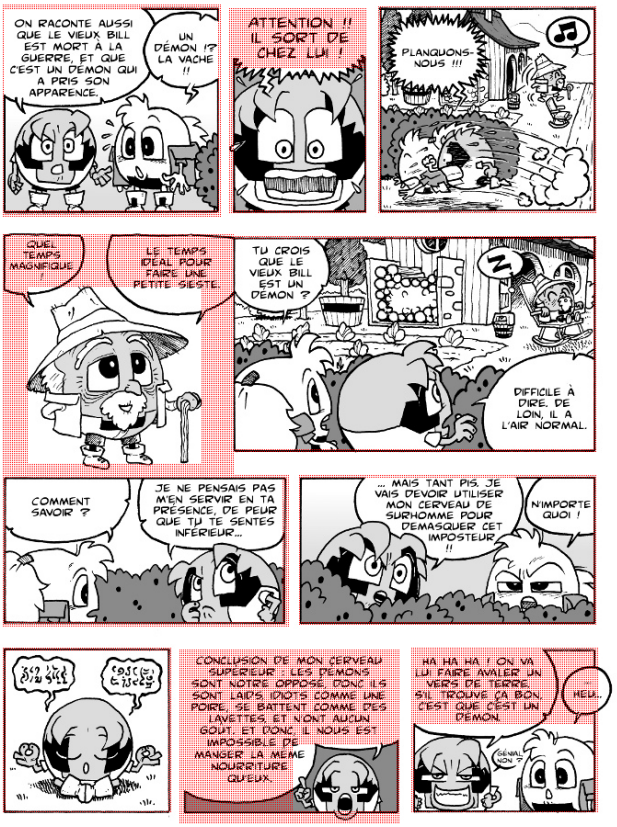

Figure 3.9: Error measurement page. Red hatched areas are the difference between areas labelled as panels by at least one person, and areas labelled by everybody. Image credit: [18].

Even though the error criterion has only been estimated on panels, it is reasonable to extend it to balloons and text lines as well. Indeed, the segmentation protocol being quite similar for all features (bounding box as close as possible to the drawing), the observed standard deviation of panel corner positions has no reason to be different from balloons and text lines. The pixel-level balloon and the comic characters visual annotation have been carried out by a single person, the homogeneity is only subject to the regularity of the person over time and is, therefore, difficult to assess quantitatively.

3.5 Conclusions

We presented the eBDtheque, a representative database of comics, which is composed by comics images, visual (spatial) and semantic annotations. The one hundred image corpus has been introduced as well as its ground truth construction and quality assessment protocols.

The material have been published in 2013 [30] and made available to the scientific community via the project website [6].

At the time this thesis is written, five requests from India, China and Japan have been received to use this dataset and ground truth.

[6] http://ebdtheque.univ-lr.fr

Chapter 4

Panel extraction

4.1 Introduction

4.2 Methodology (contribution)

- Compute a confidence value between 0 and 1 for each panel: the inverse of overlapping percentage with other panels. (maximal when no overlap). This can replace the topological filtering by considering as correct the panel with a confidence ¿ 0%
- `/PhD/publication/2013/LNCS/robust_frame_and_text_extraction_from_comic_books/paper`
- `/PhD/Publications/CIFED_2012`

4.3 Experimental results

4.4 Conclusions

Chapter 5

Balloon segmentation and classification

5.1 Introduction

5.2 Methodology

5.2.1 Localization

5.2.2 Segmentation & classification

- Balloon type study: The Rise and Reason of Comics and Graphic Literature: Critical Essays on the Form chapter 4 `http://afflatus.ucd.ie/Papers/comics%202010.pdf` or `/Biblio_LINK/Bandes_dessin\unhbox\voidb@x\bgroup\let\unhbox\voidb@x\setbox\@tempboxa\hbox{e\global\mathchardef\accent@spacefactor\spacefactor}\accent19e\egroup\spacefactor\accent@spacefactores/2010_Forceville_Balloonics:TheVisualsofBalloonsinComics.pdf` and `http://books.google.fr/books?id=8yXWG0efa_8C&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false`
- Closed balloon published in IJDAR 2014 (comics understanding)
- Open balloon published in ICDAR 2013 `2013_Rigaud_An_active_contour_model_for_speech_balloon_detection_in_comics.pdf`
- Classification published in GREC'13 + LNCS'14 (not published yet) `/home/crigau02/Bureau/PhD/research/balloon_classification/output/publication/GREC/paper`

5.2.3 Tail detection and description

5.3 Experimental results

5.4 Conclusions

Chapter 6

Text localization and recognition

6.1 Introduction

6.2 Methodology

6.2.1 Text localization

6.2.2 Text recognition

- Without training results of journal publication
- Detail and complete master's student work

6.3 Experimental results

6.4 Conclusions

Chapter 7

Comic character spotting and clustering

7.1 Introduction

7.2 Methodology

7.2.1 Detection and spotting

- Insist on the key point of DAS'14 publication: the colour selection method for the object descriptor.

7.2.2 Clustering

7.3 Experimental results

7.4 Conclusions

Chapter 8

System for document understanding

In literature, there are methods that formulate...

8.1 Introduction

- From our knowledge, there is no framework for document understanding in the literature that infers new knowledge iteratively and without user interaction (unsupervised).

8.2 Methodology

8.2.1 Framework

8.2.2 Application to comics

$$SW_X = \lim_{n\to\infty} \sum_{k=0}^{n} \lambda^k W_{X_k} \tag{8.1}$$

where

$$W_{X_k} = W_X^k \tag{8.2}$$

Here λ is a weighting factor to discount the longer walks, as they often contain redundant or repeated information. In this chapter we always choose $\lambda = \frac{1}{a}$, where $a = \min(\Delta^+(W_X), \Delta^-(W_X))$. Here $\Delta^+(W_X)$, $\Delta^-(W_X)$ are respectively the maximum outward and inward degree of W_X.

8.3 Experimental results

8.4 Conclusions

Chapter 9

Conclusions

9.1 Summary and contributions

9.2 Future perspectives

List of Publications

This dissertation has led to the following communications:

Journal Papers

- Christophe Rigaud, Clément Guérin, Dimosthenis Karatzas, Jean-Christophe Burie and Jean-Marc Ogier. "Comics understanding". In ??? (pending acceptance).

Book Chapters

- 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, Vol. ???, pp. ????, 2015.

- 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, Vol. 7423, pp. 129-138, 2013.

Conference Contributions

- 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), pp. ???, Rouen, France, July, 2014.

- 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. ???, Tours, France, April, 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.

- Clément Guérin, Christophe Rigaud, Antoine Mercier, Farid Ammar-Boudjelal, Karell 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. "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.

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

- 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.

Bibliography

[1] Hafiz Aziz AHMAD, Shinichi KOYAMA, and Haruo HIBINO. Impacts of manga on indonesian readers self-efficacy and behavior intentions to imitate its visuals. *Bulletin of JSSD*, 59(3), 2012. 9

[2] Tiago Alves, Ana Simões, Rui Figueiredo, Marco Vala, Ana Paiva, and Ruth Aylett. So tell me what happened: Turning agent-based interactive drama into comics. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 3*, AAMAS '08, pages 1269–1272, Richland, SC, 2008. International Foundation for Autonomous Agents and Multiagent Systems. 11

[3] C-N E. Anagnostopoulos, Ioannis E Anagnostopoulos, Ioannis D Psoroulas, Vassilis Loumos, and Eleftherios Kayafas. License plate recognition from still images and video sequences: A survey. *Intelligent Transportation Systems, IEEE Transactions on*, 9(3):377–391, 2008. 9

[4] Kohei Arai and Herman Tolle. Automatic e-comic content adaptation. *International Journal of Ubiquitous Computing (IJUC)*, 1(1):1–11, 2010. 8

[5] 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. 7

[6] Kohei Arai and Herman Tolle. Method for real time text extraction of digital manga comic. *International Journal of Image Processing (IJIP)*, 4(6):669–676, 2011. xiii, 7, 8, 9

[7] Japan Book Publishers Association. An introduction to publishing in japan 2012-2013, 2012. 3

[8] M. Back, R. Gold, A. Balsamo, M. Chow, M. Gorbet, S. Harrison, D. MacDonald, and S. Minnerman. Designing innovative reading experiences for a museum exhibition. *Computer*, 34(1):80–87, Jan 2001. 3

[9] Thomas Blaschke, Geoffrey J. Hay, Maggi Kelly, Stefan Lang, Peter Hofmann, Elisabeth Addink, Raul Queiroz Feitosa, Freek van der Meer, Harald van der Werff, Frieke van Coillie, and Dirk Tiede. Geographic object-based image analysis

– towards a new paradigm. {*ISPRS*} *Journal of Photogrammetry and Remote Sensing*, 87(0):180 – 191, 2014. 11

[10] Michał Borodo. Multimodality, translation and comics. *Perspectives*, pages 1–20, 2014. 3

[11] DC Brandon. Graphic novels and comics for the visually impaired explored in award-winning paper, 2014. 3

[12] 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. 7

[13] S.C.S. Cheung, City University of Hong Kong. Run Run Shaw Library, and City University of Hong Kong. Department of Computer Science. *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. 10

[14] Wei-Ta Chu and Chia-Hsiang Yu. Optimized speech balloon placement for automatic comics generation. In *Proceedings of the 3rd ACM International Workshop on Interactive Multimedia on Mobile and Portable Devices*, IMMPD '13, pages 1–6, New York, NY, USA, 2013. ACM. 11

[15] Antonio Clavelli and Dimosthenis Karatzas. Text segmentation in colour posters from the spanish civil war era. In *Proceedings of the 2009 10th International Conference on Document Analysis and Recognition*, ICDAR '09, pages 181–185, Washington, DC, USA, 2009. IEEE Computer Society. 4

[16] Neil Cohn. The limits of time and transitions: Challenges to theories of sequential image comprehension. *Studies in Comics*, 1(1):127–147, 2010. 9

[17] Neil Cohn. Navigating comics: an empirical and theoretical approach to strategies of reading comic page layouts. *Frontiers in psychology*, 4(April):186, 2013. 8

[18] Cyb. *Bubblegôm Gôm*. Studio Cyborga, Goven, France, 2009. 25

[19] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In Cordelia Schmid, Stefano Soatto, and Carlo Tomasi, editors, *International Conference on Computer Vision & Pattern Recognition*, volume 2, pages 886–893, INRIA Rhône-Alpes, ZIRST-655, av. de l'Europe, Montbonnot-38334, June 2005. 10

[20] Adrien Delaye and Cheng-Lin Liu. Multi-class segmentation of free-form online documents with tree conditional random fields. *International Journal on Document Analysis and Recognition (IJDAR)*, pages 1–17, 2014. 4

[21] B. Duc. *L'art de la B.D.: Du scénario à la réalisation graphique, tout sur la création des bandes dessinées*. Editions Glénat, 1997. 2

[22] Richard O. Duda and Peter E. Hart. Use of the hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15:11–15, January 1972. 7

[23] B. Epshtein, E. Ofek, and Y. Wexler. Detecting text in natural scenes with stroke width transform. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2963 –2970, june 2010. 4

[24] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, 88(2):303–338, June 2010. 13, 16

[25] S. Fidler, Jian Yao, and R. Urtasun. Describing the scene as a whole: Joint object detection, scene classification and semantic segmentation. *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 0:702–709, 2012. 11

[26] Google. Ocr manga reader, 2014. 11

[27] G. Griffin, A. Holub, and P. Perona. Caltech-256 object category dataset. Technical Report 7694, California Institute of Technology, 2007. 13

[28] Clément Guérin. Ontologies and spatial relations applied to comic books reading. In *PhD Symposium of Knowledge Engineering and Knowledge Management (EKAW)*, Galway, Ireland, 2012. 8, 11

[29] Clément Guérin. *Représentation des données complexes pour une plateforme de recherche d'information interactive*. PhD thesis, Université de La Rochelle, Technoforum - 23, avenue Albert Einstein BP 33060 - 17031 La Rochelle - France, 2014. Part of the same project: eBDtheque. 21

[30] Clément Guérin, Christophe Rigaud, Antoine Mercier, and al. ebdtheque: a representative database of comics. In *Proceedings of International Conference on Document Analysis and Recognition (ICDAR)*, Washington DC, 2013. 13, 26

[31] 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-Verlag. 7

[32] Alice Hermann, Sébastien Ferré, and Mireille Ducassé. Guided semantic annotation of comic panels with sewelis. In *EKAW*, volume 7603 of *Lecture Notes in Computer Science*, pages 430–433. Springer, 2012. 11

[33] Anh Khoi Ngo Ho, Jean-Christophe Burie, and Jean-Marc Ogier. Comics page structure analysis based on automatic panel extraction. In *GREC 2011, Nineth IAPR International Workshop on Graphics Recognition*, Seoul, Korea, September, 15-16 2011. 7

[34] 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, March 2012. 8

[35] 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 *Proceedings of the 10th IAPR International Workshop on Graphics Recognition (GREC)*, Bethlehem, PA, USA, 2013. 10

[36] B. Hu, S. Dasmahapatra, P. Lewis, and N. Shadbolt. Ontology-based medical image annotation with description logics. In *Tools with Artificial Intelligence, 2003. Proceedings of the 15th Int. Conf. on Tools with Artificial Intelligence (ICTAI'03)*, 2003. 11

[37] Céline Hudelot, Jamal Atif, and Isabelle Bloch. Fuzzy spatial relation ontology for image interpretation. *Fuzzy Sets and Systems*, 159(15):1929 – 1951, 2008. From Knowledge Representation to Information Processing and Management Selected papers from the French Fuzzy Days (LFA 2006). 11

[38] IBISWorld. Comic book publishing in the us: Market research report, 2013. 3

[39] Yusuke In, Takashi Oie, Masakazu Higuchi, Shuji Kawasaki, Atushi Koike, and Hitomi Murakami. Fast frame decomposition and sorting by contour tracing for mobile phone comic images. *Internatinal journal of systems applications, engineering and development*, 5(2):216–223, 2011. 7

[40] Raulet Jérémy and Boyer Vincent. Comics reading: An automatic script generation. In *Proceedings of the 21st International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG)*, pages 88–96, 2013. 3

[41] Fahad Shahbaz Khan, Muhammad Anwer Rao, Joost van de Weijer, Andrew D. Bagdanov, Maria Vanrell, and Antonio Lopez. Color attributes for object detection. In *Twenty-Fifth IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012)*, 2012. 10

[42] David Kurlander, Tim Skelly, and David Salesin. Comic chat. *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques - SIGGRAPH '96*, 96:225–236, 1996. 11

[43] Bart Lamiroy, Daniel Lopresti, Hank Korth, and Jeff Heflin. How Carefully Designed Open Resource Sharing Can Help and Expand Document Analysis Research. In Christian Viard-Gaudin Gady Agam, editor, *Document Recognition and Retrieval XVIII - DRR 2011*, volume 7874, San Francisco, United States, January 2011. SPIE, SPIE. 13

[44] Bart Lamiroy and Jean-Marc Ogier. Analysis and interpretation of graphical documents. In David Doermann and Karl Tombre, editors, *Handbook of Document Image Processing and Recognition*. Springer, 2014. 10

[45] Congcong Li, Adarsh Kowdle, Ashutosh Saxena, and Tsuhan Chen. Toward holistic scene understanding: Feedback enabled cascaded classification models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(7):1394–1408, 2012. 11

[46] Luyuan Li, Yongtao Wang, Zhi Tang, and Liangcai Gao. Automatic comic page segmentation based on polygon detection. *Multimedia Tools and Applications*, 11042:1–27, 2012. 7

[47] Luyuan Li, Yongtao Wang, Zhi Tang, and Liangcai Gao. Automatic comic page segmentation based on polygon detection. *Multimedia Tools Appl.*, 69(1):171–197, 2014. 8

[48] Luyuan Li, Yongtao Wang, Zhi Tang, and Dong Liu. Comic image understanding based on polygon detection. *Proc. SPIE*, 8658:86580B–86580B–11, 2013. 8

[49] Luyuan Li, Yongtao Wang, Zhi Tang, Xiaoqing Lu, and Liangcai Gao. Unsupervised speech text localization in comic images. In *12th International Conference on Document Analysis and Recognition (ICDAR)*, pages 1190–1194, Aug 2013. 9

[50] Ricardo Lopes, Tiago Cardoso, Nelson Silva, and ManuelJ. Fonseca. Calligraphic shortcuts for comics creation. In Andreas Butz, Brian Fisher, Marc Christie, Antonio Krüger, Patrick Olivier, and Roberto Therón, editors, *Smart Graphics*, volume 5531 of *Lecture Notes in Computer Science*, pages 223–232. Springer Berlin Heidelberg, 2009. 11

[51] Song Mao, Azriel Rosenfeld, and Tapas Kanungo. Document structure analysis algorithms: a literature survey. In Tapas Kanungo, Elisa H. Barney Smith, Jianying Hu, and Paul B. Kantor, editors, *DRR*, volume 5010 of *SPIE Proceedings*, pages 197–207. SPIE, 2003. 11

[52] Christy Marx. *Writing for Animation, Comics, and Games*. Focal Press, 2006. 8

[53] Yusuke Matsui, Toshihiko Yamasaki, and Kiyoharu Aizawa. Interactive Manga retargeting. In *ACM SIGGRAPH 2011 Posters on - SIGGRAPH '11*, page 1, New York, New York, USA, 2011. ACM Press. 8

[54] S. McCloud. *Understanding comics*. William Morrow Paperbacks, 1994. 1, 3

[55] S. McCloud. *Understanding comics*. William Morrow Paperbacks, 1994. 14

[56] Jason McIntosh. ComicsML, 2011. 20

[57] Stuart Medley. Discerning pictures: How we look at and understand images in comics. *Studies in Comics*, 1(1):53–70, 2010. 9

[58] Aaron Meskin and Roy T. Cook. *The Art of Comics: A Philosophical Approach - Ch 2 The Ontology of Comics*, volume 1, pages 218–285. John Wiley and Sons, New York, USA, 2011. 11

[59] Vasileios Mezaris, Ioannis Kompatsiaris, and Michael G. Strintzis. An ontology approach to object-based image retrieval. In *In Proc. IEEE Int. Conf. on Image Processing (ICIP03*, pages 511–514, 2003. 11

[60] G.S. Millidge. *Comic Book Design*. Watson-Guptill Publications, 2009. 8

[61] L. Neumann and J. Matas. Real-time scene text localization and recognition. *Computer Vision and Pattern Recognition*, pages 1485–1490, 2012. 4

[62] J.M. Ogier, R. Mullot, J. Labiche, and Y. Lecourtier. Semantic coherency: the basis of an image interpretation device-application to the cadastral map interpretation. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 30(2):322–338, Apr 2000. 11

[63] Daniel M. Oliveira and Rafael D. Lins. Generalizing tableau to any color of teaching boards. In *Proceedings of the 2010 20th International Conference on Pattern Recognition*, ICPR '10, pages 2411–2414, Washington, DC, USA, 2010. IEEE Computer Society. 4

[64] Plasq. Comic life 3, 2014. 11

[65] Christophe Ponsard. Enhancing the accessibility for all of digital comic books. *e-Minds*, 1(5), 2009. 3, 8

[66] 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. 8, 9

[67] Gilles Ratier. 2013 : l'année de la décélération, 2013. 3

[68] J. Raulet and V. Boyer. A sketch-based interface to script comics reading. In *SIGGRAPH Asia 2011 Sketches*, SA '11, pages 3:1–3:2, New York, NY, USA, 2011. ACM. 11

[69] Paul Rissen. A Comics Ontology, 2012. 20

[70] Antti Salovaara. Appropriation of a mms-based comic creator: From system functionalities to resources for action. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 1117–1126, New York, NY, USA, 2007. ACM. 11

[71] Sohail Sarwar, Zia Ul Qayyum, and Saqib Majeed. Ontology based image retrieval framework using qualitative semantic image descriptions. *Procedia Computer Science*, 22(0):285 – 294, 2013. 17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems - {KES2013}. 11

[72] Eugenio Di Sciascio, Francesco M. Donini, and Marina Mongiello. Structured knowledge representation for image retrieval. *CoRR*, abs/1109.1498, 2011. 11

[73] Ariel Shamir, Michael Rubinstein, and Tomer Levinboim. Generating comics from 3d interactive computer graphics. *IEEE Comput. Graph. Appl.*, 26(3):53–61, May 2006. 11

[74] Shueisha. Vomic, 2014. 11

[75] Siddharth Singh, Adrian David Cheok, Guo Loong Ng, and Farzam Farbiz. 3d augmented reality comic book and notes for children using mobile phones. In *Proceedings of the 2004 Conference on Interaction Design and Children: Building a Community*, IDC '04, pages 149–150, New York, NY, USA, 2004. ACM. 3

[76] A.F. Smeaton, P. Over, and W. Kraaij. Evaluation campaigns and trecvid. In *Proceedings of the 8th ACM international workshop on Multimedia information retrieval*, pages 321–330. ACM, 2006. 13

[77] Sourceforge. Capture2text, 2014. 11

[78] Alan Stewart. A brief history of the american comic book industry, 2000. 3

[79] Martin Stommel, Lena I Merhej, and Marion G Müller. Segmentation-free detection of comic panels. In *Computer Vision and Graphics*, pages 633–640. Springer, 2012. 8

[80] Chung-Yuan Su, Ray-I Chang, and Jen-Chang Liu. Recognizing text elements for svg comic compression and its novel applications. In *Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR)*, ICDAR '11, pages 1329–1333, Washington, DC, USA, 2011. IEEE Computer Society. 9

[81] Yasuyuki Sumi, Ryuuki Sakamoto, Keiko Nakao, and Kenji Mase. Comicdiary: Representing individual experiences in a comics style. In *Ubicomp*, pages 16–32, 2002. 11

[82] 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. 10

[83] Weihan Sun and Koichi Kise. Detection of exact and similar partial copies for copyright protection of manga. *International Journal on Document Analysis and Recognition (IJDAR)*, 16(4):331–349, 2013. 10

[84] Weihan Sun, Koichi Kise, Jean-Christophe Burie, and Jean-Marc Ogier. Specific comic character detection using local feature matching. In *Proceedings of International Conference on Document Analysis and Recognition (ICDAR 2013)*, Washington, USA, 2013. 10

[85] M. Sundaresan and S. Ranjini. Text extraction from digital english comic image using two blobs extraction method. In *Pattern Recognition, Informatics and Medical Engineering (PRIME), 2012 International Conference on*, pages 449–452, March 2012. 9

[86] Anh-Phuong TA. *Inexact Graph matching techniques: Application to Object detection and Human action recognition.* PhD thesis, INSA de Lyon, 20, rue Albert Einstein, 69621 Villeurbanne Cedex, France, 2010. 10

[87] Kohei Takayama, Henry Johan, and Tomoyuki Nishita. Face detection and face recognition of cartoon characters using feature extraction. In *Image Electronics and Visual Computing Workshop (IEVC'12)*, Kuching, Malaysia, 2012. 10

[88] 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. 7

[89] Evan Thomas. *Invisible Art, Invisible Planes, Invisible People.* Multicultural Comics: From Zap to Blue Beetle. University of Texas Press, 2010. 3

[90] Hiroaki Tobita. Comic engine: Interactive system for creating and browsing comic books with attention cuing. In *Proceedings of the International Conference on Advanced Visual Interfaces*, AVI '10, pages 281–288, New York, NY, USA, 2010. ACM. 10

[91] Chia-Jung Tsai, Chih-Yuan Yao, Pei-ying Chiang, Yu-Chi Lai, Ming-Te Chi, Hung-Kuo Chu, Yu-Shiang Wong, and Yu-Shuen Wang. Adaptive manga relayout on mobile device. In *ACM SIGGRAPH 2013 Posters*, SIGGRAPH '13, pages 57:1–57:1, New York, NY, USA, 2013. ACM. 8

[92] Shingo Uchihashi, Jonathan Foote, Andreas Girgensohn, and John Boreczky. Video manga: Generating semantically meaningful video summaries. In *Proceedings of the Seventh ACM International Conference on Multimedia (Part 1)*, MULTIMEDIA '99, pages 383–392, New York, NY, USA, 1999. ACM. 11

[93] Paul Viola and Michael J Jones. Robust real-time face detection. *International journal of computer vision*, 57(2):137–154, 2004. 10

[94] John A Walsh. Comic Book Markup Language : An Introduction and Rationale. *Digital Humanities Quarterly (DHQ)*, 6(1):1–50, 2012. 20

[95] Meng Wang, Richang Hong, Xiao-Tong Yuan, Shuicheng Yan, and Tat-Seng Chua. Movie2comics: Towards a lively video content presentation. *IEEE Transactions on Multimedia*, 14(3-2):858–870, 2012. 10

[96] J.J. Weinman, E. Learned-Miller, and A.R. Hanson. Scene text recognition using similarity and a lexicon with sparse belief propagation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(10):1733 –1746, 2009. 4

[97] Masashi Yamada, Rahmat Budiarto, Mamoru Endo, and Shinya Miyazaki. Comic image decomposition for reading comics on cellular phones. *IEICE Transactions*, 87-D(6):1370–1376, 2004. 9

[98] B. Yao, X. Yang, and S.C. Zhu. Introduction to a large-scale general purpose ground truth database: methodology, annotation tool and benchmarks. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 169–183. Springer, 2007. 16

[Titre de la thèse (en français)]

Résumé : (1700 caractères max.)

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Proin volutpat ipsum id purus ultrices lobortis. Maecenas ornare enim quis eros. Nunc eget mauris ut quam malesuada mattis. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae ; Integer vel tellus. Nam rutrum, purus non sodales rhoncus, quam magna imperdiet eros, sit amet euismod justo metus at orci. Suspendisse neque turpis, feugiat interdum, faucibus vel, aliquet quis, risus. Etiam est elit, eleifend a, consequat sit amet, scelerisque nec, odio. Quisque id odio quis libero iaculis tincidunt. Sed non mi. Morbi aliquam commodo nibh. Integer justo purus, pulvinar a, suscipit vel, iaculis a, justo. Morbi ut orci. Maecenas fringilla orci. Phasellus auctor, enim vitae tempus egestas, justo mi cursus sem, vel blandit leo turpis vitae quam. Etiam sit amet felis vitae eros ornare porttitor.
Proin orci ligula, vehicula non, ultrices at, ultrices ut, massa. Vestibulum ac est. Curabitur at erat. Mauris gravida. Praesent vestibulum. Curabitur eget orci ac massa cursus condimentum. Integer sapien dui, ultricies eget, dapibus a, dapibus id, mauris. Curabitur felis velit, aliquam at, aliquet in, iaculis vitae, velit. Nunc lobortis magna id ligula. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae ; Integer congue ultrices mi. Isdem diebus Apollinaris Domitiani gener, paulo ante agens palatii Caesaris curam, ad Mesopotamiam missus a socero per militares numeros immodice scrutabatur, an quaedam altiora meditantis iam Galli secreta susceperint scripta, qui conpertis Antiochiae gestis per minorem Armeniam lapsus Constantinopolim petit.

Mots clés : mot 1, mot 2, ...

[Titre de la thèse (en anglais)]

Summary : (1700 caractères max.)

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Proin volutpat ipsum id purus ultrices lobortis. Maecenas ornare enim quis eros. Nunc eget mauris ut quam malesuada mattis. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae ; Integer vel tellus. Nam rutrum, purus non sodales rhoncus, quam magna imperdiet eros, sit amet euismod justo metus at orci. Suspendisse neque turpis, feugiat interdum, faucibus vel, aliquet quis, risus. Etiam est elit, eleifend a, consequat sit amet, scelerisque nec, odio. Quisque id odio quis libero iaculis tincidunt. Sed non mi. Morbi aliquam commodo nibh. Integer justo purus, pulvinar a, suscipit vel, iaculis a, justo. Morbi ut orci. Maecenas fringilla orci. Phasellus auctor, enim vitae tempus egestas, justo mi cursus sem, vel blandit leo turpis vitae quam. Etiam sit amet felis vitae eros ornare porttitor.
Proin orci ligula, vehicula non, ultrices at, ultrices ut, massa. Vestibulum ac est. Curabitur at erat. Mauris gravida. Praesent vestibulum. Curabitur eget orci ac massa cursus condimentum. Integer sapien dui, ultricies eget, dapibus a, dapibus id, mauris. Curabitur felis velit, aliquam at, aliquet in, iaculis vitae, velit. Nunc lobortis magna id ligula. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae ; Integer congue ultrices mi. Isdem diebus Apollinaris Domitiani gener, paulo ante agens palatii Caesaris curam, ad Mesopotamiam missus a socero per militares numeros immodice scrutabatur, an quaedam altiora meditantis iam Galli secreta susceperint scripta, qui conpertis Antiochiae gestis per minorem Armeniam lapsus Constantinopolim petit.

Keywords : word 1, word 2,...

SIGLE DU LABORATOIRE

[logo partenaire] **[Nom et adresse du laboratoire]** [logo partenaire]

17xxx LA ROCHELLE