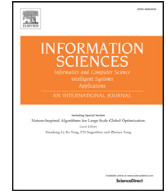


Contents lists available at [ScienceDirect](#)

Information Sciences

journal homepage: www.elsevier.com/locate/ins

An ontology-based framework for the automated analysis and interpretation of comic books' images



Clément Guérin*, Christophe Rigaud, Karell Bertet, Arnaud Revel

L3i – University of La Rochelle, 23 avenue Albert Einstein La Rochelle 17031 France

ARTICLE INFO

Article history:

Received 13 April 2016

Revised 8 August 2016

Accepted 20 October 2016

Available online 20 October 2016

Keywords:

Comic books

Images

Complex data

Knowledge representation

Ontologies

Spatial reasoning

ABSTRACT

Since the beginning of the twenty-first century, the cultural industry has been through a massive and historical mutation induced by the rise of digital technologies. The comic books industry keeps looking for the right solution and has not yet produced anything as convincing as the music or movie have. A lot of energy has been spent to transfer printed material to digital supports so far. The specificities of those supports are not always exploited at the best of their capabilities, while they could potentially be used to create new reading conventions. In spite of the needs induced by the large amount of data created since the beginning of the comics history, content indexing has been left behind. It is indeed quite a challenge to index such a composition of textual and visual information. While a growing number of researchers are working on comic books' image analysis from a low-level point of view, only a few are tackling the issue of representing the content at a high semantic level. We propose in this article a framework to handle the content of a comic book, to support the automatic extraction of its visual components and to formalize the semantic of the domain's codes. We tested our framework over two applications: 1) the unsupervised content discovery of comic books' images, 2) its capabilities to handle complex layouts and to produce a respectful browsing experience to the digital comics reader.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Comic books are a unique art form which popularity grew increasingly from its beginning in the 19th century to the present day. Carrying an important legacy, the 9th art created its own narrative codes, mixing like no other art text and drawings to tell a story. Those very specific codes, evolving over decades, are intimately bounded to their physical support. That is probably one of the main reasons why the comic book's industry struggles to find a proper and efficient way on the digital market, like the movies, music and regular books industries did years ago. There are plenty of applications that would emerge from a smart digitization of comic books.

The enhancement of the small screen reading experience is an obvious example. Spotting panels, speech balloons and so on, is an essential feature to anyone who wants to provide a smooth and nice reading experience. Speech synthesis of written text lines in combination with the spotting of the talking characters would lead to a very immersive experience that could help small children to apprehend comic books. Those pieces of text could be automatically translated in any required

* Corresponding author.

E-mail addresses: clement.guerin@univ-lr.fr, clement.guerin87@gmx.com (C. Guérin), christophe.rigaud@univ-lr.fr (C. Rigaud), karell.bertet@univ-lr.fr (K. Bertet), arnaud.revel@univ-lr.fr (A. Revel).

language if needed. The recognition of what is going on in a given panel, when and where the story takes place, what real or fictional characters are involved in it, could bring to the reader additional information that would help her or him to grasp the whole meaning of what she or he is reading. Some issues about old comic books preservation tend to rise as well, as the pieces produced at the early stages of the art's history are getting older every passing day. The world produced a large amount of heterogeneous comic books that now needs to be archived in some sort of database. Only information such as title, year and author are commonly used as intent for the queries, while an insightful content description would be very helpful to the archivists as well.

However, such use cases got stuck with a few scientific locks. How to extract the visual elements that compose a comic book, e.g. the panels, balloons, text and so on, is one of those locks. Then, this extracted content needs to be organized in a way that can show the original meaning of the author during the creation, in order to be enhanced, queried, adapted, with respect to that meaning. We started a wide study a few years ago to discuss these two issues. We described how we tackled the challenge of visual elements extraction in a previous publication [41]. The present article details the results of the latter, namely how we created a framework able to handle extracted comic books visual elements in a meaningful way.

We organized this article as follows. The first section presents several comic books formalisms from the literature. We present as well the state of the art on the use of formalized knowledge for image analysis, especially when it comes to discover what is the meaning of the extracted content. The second section introduces, through a complete and detailed presentation, our own proposition, including two ontologies, as a framework to handle the content of comic books' images. Then we discuss our conceptualization choices and test them over a public dataset. We also present some experimental results on two applications: a new panels ordering method and an iterative process to discover the content of a page. We discuss the results in the last section that concludes this article.

2. State of the art

This section introduces the different studies that came as a source of inspiration for our own work. We begin with presenting the major comic books' formalization endeavors from the last few years. We go through their pros and cons and highlight what is missing to meet the aim of our study.

Then we present a short survey of how semantics formalization showed up as an increasingly promising tool to enhance computer vision. We focused our analysis on how they can help narrowing the semantic gap and of what use they can be to supervise the image analysis process.

2.1. Multiple formalisms for comic books

If one wants to fully extract and analyze a comic books' content, it is necessary to know how that content is organized and the nature of the different relations that exist between its elements.

While a growing number of studies came up lately in linguistics [2] and psychology [12] communities, the computer science community, as we know of, conducted a very few works on the comic books language conceptualization. Besides the *Periodical Comics* web-schema proposal [49] and the *Grand Comics Database* platform [25], both focusing on the publishing metadata dimension, we could only find a very small amount of endeavors on the content formalization. Not all of them have been the subject of an academic publication.

In 2001, *ComicsML*, an XML formalism to describe the content of a comic books plate has been proposed [29]. The authors developed it in the perspective of becoming a standard for web-comics publishing which was barely emerging then. Indeed, web-comics, back then and still now, are often published as raw images, straight to the author's web page, with a varying frequency. Through its wish of standardization, the real ambition of *ComicsML* was to generate new usages, especially with the development of reading tools exploiting the language specificities. It provides means to describe a series of web-comics published online, from the bibliographical information to the kind of balloons drawn in the panels and the font used for the text lines. Its syntax is based on XML, it uses tags and implies a hierarchical structure of the described elements. The root node, *comic*, describes the series that the web-comics belongs to. Its children, *strip*, are the several issues of the series, i.e. the published images. These tags embed the bibliographical metadata about the author and the publication date. The *panel* tags, children of *strip*, describes the panels' content. It includes the kinds of balloons (thought, dialogue, etc.), the text typography (bold, italic, etc.) and the involved characters. The authors dropped the support of the language after version 0.3. A similar approach, but specifically adapted for manga, can be found in [35].

A second initiative, *CBML* (Comic Books Markup Language), came out a few years later with the work of Walsh [50]. Also based on an XML syntax, its philosophy is quite similar to *ComicsML* without being exclusively used for web-comics annotation. The main difference with *ComicsML* lies in the fact that *CBML* is based on the "Text Encoding Initiative P5: Guidelines for Electronic Text Encoding and Interchange" [46]. The *TEI* is a framework, proposed by the homonym consortium, to encode textual documents and illustrated documents especially. Its purpose is to make information retrieval in large encoded datasets easier. *CBML* extends the *TEI* vocabulary with some comic book's notions (panels, balloons, etc.), while reusing, as much as possible, the existing encoding.

Pastierovic proposed a third formalism, called *Advanced Comic Book Format* or *ACBF* [37]. It focuses on the encoding of digital comic books. Also XML-based, it aims to enhance the reading experience on computers. Its authors define it as the comic books archive file formats, such as CBR and CBZ, evolution, as it adds metadata on several elements. It includes the

Download English Version:

<https://daneshyari.com/en/article/4944862>

Download Persian Version:

<https://daneshyari.com/article/4944862>

[Daneshyari.com](https://daneshyari.com)