1 Doctoral consortium application

1.1 General Information

Name surname:	Christophe Rigaud
University:	Laboratory L3i, University of La Rochelle, France
Supervisors:	Jean-Christophe Burie, Dimosthenis Karatzas and Jean-Marc Ogier
Thesis title:	Segmentation and indexation of complex object: application to the detection and recognition of complex objects in comics books
Date:	October 1^{st}, 2011 to September 30^{st}, 2014

1.2 Research plan

1.2.1 Overview

L3i lab initiated the e-BDthèque project to develop digital comics books under the CPER 2008-2013 founds. Nowadays, many large scale digitalization processes of comics are considered at national and international scale. From one hand, the digitalization of the comics rise new needs of indexing tools allowing to browse a large amount of data. With such tool, it could be possible to find specific drawings, animate objects or characters, analyse specific strip sequence, etc... On the other hand, several company feedbacks shows that there is no tool allowing to automatically extract comics content from digitized books such as frames, speech balloons and text. The implementation of reading scenario for new devices as tablet and smartphones is made manually and take an terrible time. The development of automatic content extraction tools is relevant for full-text search, frame per frame reading, automatic translation, text-to-speech, etc...

1.2.2 Contributions

Comics content many heterogeneous elements that is better to treat separately. We decided to split our work into several steps. First, we defined a dataset and its ground truth in order to evaluate our work later on and to share with the community, make new cooperation and propose new solutions for comics in a meantime. Second, we extract frames, text, balloons in order to focus on the much more complex elements in a limited areas: we call them "complex objects" (e.g. character, monument, vehicle). Third, we index the extracted elements to allow information retrieval and evaluation.

Dataset

We built the first dataset of comics in the literature. It took almost one year to define which type of comics we needed, get an agreement for copyright authorizations from comics authors and publishers, develop a specific tool to make the ground truth, define a protocol and hire people to make the ground truth. This ground truth will be published in ICDAR'13 poster session and is now available at *http://ebdtheque.univ-lr.fr*.

Frame segmentation

Frame segmentation has been mainly studied for reading comics on mobile device in order to display them frame by frame on a small screen. Our work concerns the indexing of a huge amount of albums that raises new issues in terms of variety of format, resolution and content. We published a new comics frame extraction method[1].

Text localization

To solve the particular problems which are provoked by the combination of complex background and unstructured documents, we have published a new text localization method called "Automatic Text Localisation in Scanned Comic Books"[2]. The evaluation part shows that we overpass other method combination found in the literature by reaching 76% recall and precision for text line localization in comics.

Balloon detection

We will present at ICDAR'13 poster session an active contour based method to accurately localize open and closed speech balloon in comics book. The proposed approach relies on text detection and prior knowledge to fit balloon contour at different resolutions. The evaluation shows 92.3% recall and 94.4% precision using ground truth text and 83.4% recall and 55.5% precision using the automatic text detector presented above. Balloon classification is also in the pipeline (GREC'13), it consist in analysing the balloon contour variations and classifying as "dialogue", "exclamation" or "though" according to "smooth", "zigzag" and "wavy" contour types.

1.2.3 Future work

We now have detected all the recurrent feature that compose comics but for a complete comics content understanding we need to focus on remaining panel contents which is graphical drawing. Those drawings are everything except text and balloons (e.g. characters, buildings, objects). Segmenting such "non-rigid" objects is a real challenge because they are position, rotation, scale, colour and occlusion variants. We plan to split this work into two tasks. First using CBIR techniques based on local feature, colour and spatial information to retrieve similar objects in all the frames of the same album given an image example. Second we will try to automatically detect the more redundant objects (hopefully the hero of the comics), without previous learning, by analysing frame content redundancy throughout pages and albums. So far, we investigated redundant object detection by adjacent sub-graph matching (GREC'13) and a first work on high apparition frequency local feature matching applied on comics has been carried out ("Specific Comic Character Detection Using Local Feature Matching", ICDAR'13).

[1] C. Rigaud, N. Tsopze, J-C. Burie and J-M Ogier, Frame and text extraction from comic books, Graphics Recognition. New Trends and Challenges, volume 7423 of Lecture Notes in Computer Science (LNCS), pp. 129-138. Springer Berlin Heidelberg, 2013.

[2] C. Rigaud, D. Karatzas, J. Van de Weijer, J-C Burie and J-M Ogier. Automatic text localisation in scanned comic books. International Conference on Computer Vision Theory and Applications (VISAPP), pp. 814-819. SCITEPRESS Digital Library, 2013.