



UNIVERSITÉ DE LA ROCHELLE

ÉCOLE DOCTORALE S2IM

Laboratoire Informatique, Image et Interaction (L3i)

THÈSE présentée par :

Christophe RIGAUD

soutenue le : **11 décembre 2014**

pour obtenir le grade de : **Docteur de l'université de La Rochelle**

Discipline : **informatique et applications**

**Segmentation et indexation d'objets complexes dans les
images de bandes dessinées**

JURY :

Bart LAMIROY

Maître de conférences, HDR, Université de Lorraine (France), Examinateur, Président du jury

Simone MARINAI

Professeur associé, Université de Florence (Italie), Rapporteur

Apostolos ANTONACOPOULOS

Professeur associé, Université de Salford (Grande-Bretagne), Rapporteur

Jean-Philippe DOMENGER

Professeur, Université de Bordeaux (France), Examinateur

Nicholas JOURNET

Maître de conférences, Université de Bordeaux (France), Examinateur

Jean-Christophe BURIE

Professeur, Université de La Rochelle (France), Directeur de thèse

Dimosthenis KARATZAS

Professeur associé, Université Autonome de Barcelone (Espagne), Encadrant de thèse

Jean-Marc OGIER

Professeur, Université de La Rochelle (France), Encadrant de thèse

Segmentation et indexation d'objets complexes dans les images de bandes dessinées

Résumé :

Dans ce manuscrit de thèse, nous détaillons et illustrons les différents défis scientifiques liés à l'analyse automatique d'images de bandes dessinées, de manière à donner au lecteur tous les éléments concernant les dernières avancées scientifiques en la matière ainsi que les verrous scientifiques actuels.

Nous proposons trois approches pour l'analyse d'image de bandes dessinées. La première approche est dite "séquentielle" car le contenu de l'image est décrit progressivement et de manière intuitive. Dans cette approche, les extractions se succèdent, en commençant par les plus simples comme que les cases, le texte et les bulles qui servent ensuite à guider l'extraction d'éléments plus complexes tels que la queue des bulles et les personnages au sein des cases. La seconde approche propose des extractions indépendantes les unes des autres de manière à éviter la propagation d'erreur due aux traitements successifs. D'autres éléments tels que la classification du type de bulle et la reconnaissance de texte y sont aussi abordés. La troisième approche introduit un système fondé sur une base de connaissance *à priori* du contenu des images de bandes dessinées. Ce système permet de construire une description sémantique de l'image, dirigée par les modèles de connaissances. Il combine les avantages des deux approches précédentes et permet une description sémantique de haut niveau pouvant inclure des informations telles que l'ordre de lecture, la sémantique des bulles, les relations entre les bulles et leurs locuteurs ainsi que les interactions entre les personnages.

Mots clés : traitement d'images, reconnaissance de formes, analyse de documents, compréhension de bandes dessinées.

Segmentation and indexation of complex objects in comic book images

Summary:

In this thesis, we review, highlight and illustrate the challenges related to comic book image analysis in order to give to the reader a good overview about the last research progress in this field and the current issues.

We propose three different approaches for comic book image analysis that are composed by several processing. The first approach is called "sequential" because the image content is described in an intuitive way, from simple to complex elements using previously extracted elements to guide further processing. Simple elements such as panel text and balloon are extracted first, followed by the balloon tail and then the comic character position in the panel. The second approach addresses independent information extraction to recover the main drawback of the first approach, the error propagation. This second method is called "independent" because it is composed by several specific extractors for each elements of the image without any dependence between them. Extra processing such as balloon type classification and text recognition are also covered. The third approach introduces a knowledge-driven and scalable system of comics image understanding. This system called "expert system" is composed by an inference engine and two models, one for comics domain and another one for image processing, stored in an ontology. This expert system combines the benefits of the two first approaches and enables high level semantic description such as the reading order of panels and text, the relations between the speech balloons and their speakers and the comic character identification.

Keywords: image processing, graphics recognition, document analysis, comics understanding.