

Réduction de l'espace de recherche pour les personnages de bandes dessinées

Clément Guérin¹ Christophe Rigaud^{1,2} Karell Bertet¹
Jean-Christophe Burie¹ Arnaud Revel¹ Jean-Marc Ogier¹

¹Laboratoire L3i, Université de la Rochelle
Avenue Michel Crépeau 17042 La Rochelle, France

²Computer Vision Center, Universitat Autònoma de Barcelona
AE-08193 Bellaterra (Barcelona), Spain

{cguerin, crigaud, kbertet, jcburie, arevel, jmogier}@univ-lr.fr

Résumé

Les bandes dessinées représentent un patrimoine culturel important dans de nombreux pays et leur numérisation massive offre la possibilité d'effectuer des recherches dans le contenu des images. À ce jour, ce sont principalement les structures des pages et leurs contenus textuels qui ont été étudiés, peu de travaux portent sur le contenu graphique. Nous proposons de nous appuyer sur des éléments déjà étudiés tels que la position des cases et des bulles, pour réduire l'espace de recherche et localiser les personnages en fonction de la queue des bulles. L'évaluation de nos différentes contributions à partir de la base eBDtheque montre un taux de détection des queues de bulle de 81.2%, de localisation des personnages allant jusqu'à 85% et un gain d'espace de recherche de plus de 50%.

Mots Clef

recherche contextuelle, analyse de document, personnage de bandes dessinées.

Abstract

Comic strips represent an important piece of cultural heritage in numerous countries and the large-scale digitization offers the possibility of making searches directly in the image content. So far, it is essentially the page structures and the textual contents that have been studied, few works concern the graphic content. We suggest to rely on simple elements, already studied in the literature, such as the position of frames and balloons, to localize more complex elements, such as the comic book heroes, with a higher accuracy. The evaluation of our different contributions from the eBDtheque dataset shows a tail detection accuracy of 81.2%, a character localization accuracy up to 85% and a search space reduction of more than 50%.

Keywords

contextual search, document analysis, comics characters.

1 Introduction

La bande dessinée, bien que représentant une part importante de l'héritage culturel de nombreux pays, notamment en Europe de l'ouest, au Japon et aux États-Unis, n'a pas encore bénéficié du même niveau d'attention que la musique, le cinéma ou la littérature quant à son adaptation au format numérique. Les œuvres anciennes numérisées pourraient être valorisées en utilisant des outils de recherche d'information avancés pour retrouver, par exemple, une séquence spécifique ou l'évolution d'un personnage au fil du temps. Néanmoins, le processus de conversion et d'adaptation n'est pas aussi simple que celui mis en place pour la publication numérique d'un film ou d'un roman. La bande dessinée diffère de ces derniers dans le fait que le média lui-même est intimement lié au médium utilisé. Elle est définie comme "une suite d'images picturales et fixes volontairement juxtaposées en séquences" par l'ouvrage de référence en la matière écrit par McCloud [1]. Elle diffère cependant du cinéma sur la forme et le positionnement spatial de ces images. Là où pour ce dernier les images sont toutes de tailles identiques et où chacune vient remplacer la précédente, les cases de bande dessinée sont de tailles variables et organisées spatialement dans un espace limité. À ces deux particularités s'ajoute le fait que le lecteur a la possibilité de voir d'un coup d'œil toutes les cases d'une même page et n'est donc pas contraint de parcourir l'histoire dans un ordre donné. Ces spécificités sont autant d'outils au service de l'auteur lui permettant de mettre en scène son histoire. Dès lors, changer le support de l'œuvre, la surface de lecture ou la manière de séquencer les cases, implique une modification de la mise en scène qui peut, dans certains cas, être néfaste au récit.

L'idéal serait donc de comprendre et d'identifier automatiquement les procédés de mise en scène employés par les auteurs sur le format papier, afin de permettre leur restitution sous une forme adaptée au support sur lequel l'œuvre est lue. Cela passe d'abord par une analyse de la page afin de détecter les différents éléments la composant, puis par une reconstruction du récit en replaçant les éléments ex-

traits dans l'ordre narratif, la compréhension du rôle de chaque case, l'association des bulles aux cases auxquelles elles appartiennent, l'identification des personnages locuteurs et ceux à qui ils s'adressent, etc.

Différents travaux ont été menés dans le cadre du projet eBDtheque¹ (L3i), en analyse d'images et en représentation des connaissances, afin d'extraire et de manipuler les objets visuels élémentaires des bandes dessinées.

L'objectif de cet article est de présenter une nouvelle méthode visant à faciliter la localisation et l'extraction des personnages. Nous nous concentrons dans un premier temps sur les personnages intervenant dans le récit par l'émission d'une bulle de dialogue. Nous partons du postulat que les cases et les bulles (phylactères) ont été correctement détectés par un traitement antérieur.

La structure de l'article s'articule de la manière suivante : la section 2 introduit le contexte de cette étude, la section 3 détaille notre proposition pour localiser les personnages au sein des cases d'une bande dessinée. Viennent ensuite les expérimentations et leurs évaluations dans la section 4, suivies d'une discussion et de la conclusion.

2 Analyse de bandes dessinées

Des travaux antérieurs ont été successivement menés afin d'être en mesure d'extraire automatiquement, par l'analyse des images, les éléments caractéristiques de la bande dessinée, à savoir les cases, les bulles, le texte et les personnages. Ces travaux sont brièvement présentés dans les trois sous-sections suivantes.

D'autre part, nous avons développé une ontologie (cf. figure 1) permettant de structurer les données, les informations et les connaissances extraites lors de ces étapes d'analyse de l'image afin de pouvoir les exploiter. La connaissance du domaine de la bande dessinée y est modélisée, accompagnée de concepts issus du milieu de l'image, tels que les notions de *région d'intérêt* ou encore d'*extracteur*. Une description détaillée est disponible dans [2].

2.1 Extraction et séquençage des cases

L'extraction automatique des cases a principalement été étudiée pour permettre la détection et le classement automatique des cases selon leurs sens de lecture. Ce besoin est né d'une volonté d'automatiser la lecture case par case sur des appareils mobiles dont la taille des écrans n'est pas suffisamment grande pour lire confortablement une page au format A4 sans l'utilisation de fonctionnalités supplémentaires telles que le zoom ou le défilement.

Plusieurs techniques de séparation du fond et des cases ont été proposées comme le montre [3]. Les études les plus récentes sont basées sur un découpage suivant la détection de lignes [4], les composantes connexes [5], la morphologie mathématique [6], les lignes de partage des eaux (watershed) [7] ou encore la détection de zones d'intérêt [8] telles que les angles et les segments de ligne. Ces travaux

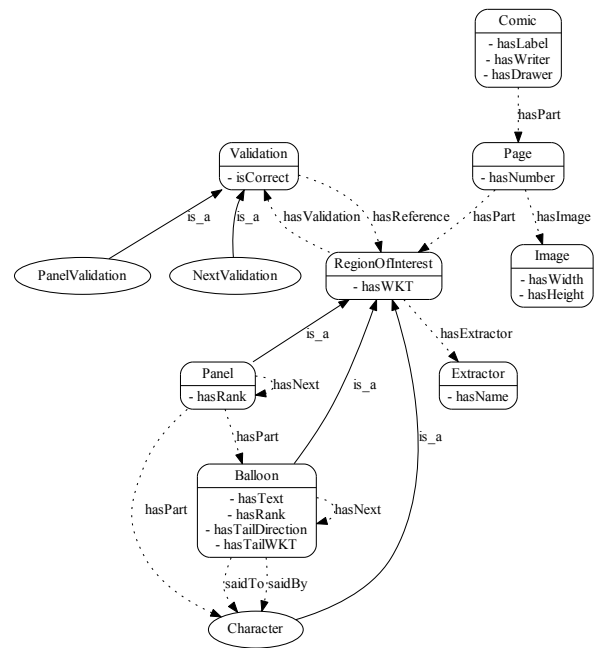


FIGURE 1 – Extrait de la hiérarchie de concepts (en flèches pleines) de notre ontologie et leurs relations (en flèches pointillées).

ont montré des résultats intéressants autant sur les bandes dessinées de type manga que de type européen.

Certains travaux s'intéressant à l'extraction des cases ont aussi étudié l'ordre de lecture des cases [7]. Des études très récentes, menées en psychologie, ont été jusqu'à théoriser l'approche utilisée par les lecteurs pour parcourir une page de bande dessinée afin de comprendre les mécanismes mis en œuvre lors de la lecture [9]. Un modèle développé au sein de notre équipe, nous a permis de traiter et d'interpréter les cas volontairement ambigus où deux cases ne sont pas parfaitement disjointes, ou quand une case est incluse dans une autre [2]. Nous avons proposé une déduction automatique de la séquence de lecture en triant les cases en fonction de leurs positions respectives. La page est alors représentée par un graphe orienté sans cycle (Directed Acyclic Graph ou DAG) où les nœuds représentent les cases et les arcs les transitions possibles entre ces cases. Le calcul de différents tris topologiques de ce dernier permet d'en extraire l'ensemble des séquences de lecture possible. Le DAG est construit à partir d'un ensemble de relations spatiales topologiques et d'orientation entre les cases d'une même page. Ces relations nous permettent de déterminer si une case *A* est à gauche, à droite, au dessus ou au dessous d'une case *B*. *A* sera lu en premier si elle est au dessus de *B* ou sur la même ligne et à gauche de *B*. Lorsque deux cases sont disjointes, la position, l'une par rapport à l'autre, du polygone représentant leur contour est étudiée. La priorité de lecture est donnée à celle se situant au des-

1. <http://ebdtheque.univ-lr.fr>

sus de l'autre puis, si les polygones possèdent une intersection en ordonnées, à celle se situant à gauche de l'autre. Lorsque deux cases possèdent une intersection en abscisses et en ordonnées, c'est à dire lorsqu'une partie de leur surface est commune, seul leur centre de gravité est utilisé. Si l'écart entre les deux centres est plus important en ordonnées qu'en abscisses, les deux cases sont classées comme étant l'une en dessous de l'autre. Dans le cas contraire, nous considérons que l'une est à gauche de l'autre. La règle utilisée pour choisir l'ordre de lecture est alors la même que pour deux cases disjointes. Le cas *A est en dessous de B* est illustré par l'équation 1 où $DC(A, B)$ et $PO(A, B)$ signifient respectivement que les éléments *A* et *B* sont disjointes ou qu'ils possèdent une intersection. y_{X_i} représente l'ordonnée de la position d'un point de la case *X*, y_{X_c} représente l'ordonnée de son barycentre. Δy_c et Δx_c correspondent respectivement à la différence en ordonnées et en abscisses des barycentres des cases *A* et *B*.

$$Down(A, B) \Leftrightarrow (DC(A, B) \wedge \forall (y_{A_i}, y_{B_i}), y_{B_i} > y_{A_i}) \quad (1)$$

$$\vee (PO(A, B) \wedge y_{B_c} > y_{A_c} \wedge (\Delta y_c > \Delta x_c))$$

Enfin, lorsqu'une case est incluse dans une autre, l'indéterminisme souhaité par l'auteur est conservé et les différents choix possibles sont intégrés dans le modèle. Les différents cas de figure sont illustrés par la figure 2.

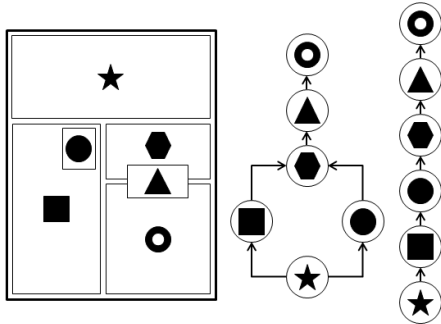


FIGURE 2 – Pour une configuration de cases donnée, le DAG et un tri topologique possible (e.g. sens de lecture).

2.2 Extraction et séquençement des bulles

L'extraction des bulles est un élément clef de l'analyse de bandes dessinées. En effet, les bulles mettent en relations les éléments graphiques et textuels les uns avec les autres et leur style contribuent à l'identité visuelle de l'album. Il existe différents types de bulles que l'on peut caractériser de *régulières*, *déformées* et *ouvertes* (voir fig. 3). Les méthodes d'extractions de la littérature sont basées sur la détection de régions [10, 11], associées à des règles de filtrage. Une méthode proposée par notre équipe et basée sur les contours actifs, a été spécialement développée pour détecter les bulles de type ouvertes [12].

Le séquençement des bulles peut s'effectuer d'une manière similaire à celle utilisée lors du séquençement des cases, à ceci près que l'espace de recherche n'est plus la page en-

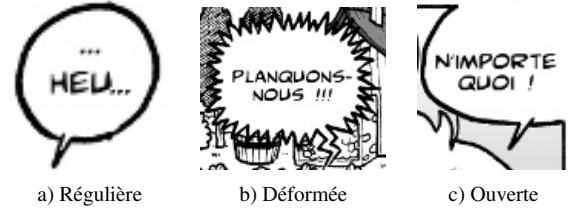


FIGURE 3 – Exemples de types de bulles. Crédits image [13].

tière mais est réduit à la case. Les relations spatiales topologiques et d'orientation entre deux bulles sont autant d'informations utilisables pour déterminer, grâce à l'application de règles spatiales et d'un tri topologique, l'ordre de lecture. L'ordre établi en amont sur les cases de la page peut alors être utilisé pour lier la dernière bulle de la case *n* à la première bulle de la case *n + 1*, recomposant ainsi les dialogues d'une page entière.

2.3 Extraction et séquençement des personnages

L'extraction et le séquençement des personnages vont de pair avec l'extraction et le séquençement des bulles. En effet, si l'on est capable de trouver l'ordre dans lequel sont lues les bulles, nous pouvons en déduire l'ordre dans lequel parlent les personnages. Pour cela, il faut dans un premier temps détecter les personnages, une opération non triviale en raison de la diversité de forme et de style inhérente au monde de la bande dessinée (cf. [1], chapitre 2).

L'extraction des humains, et notamment des visages, dans les images de scènes réelles est une thématique de recherche très active depuis plusieurs décennies. En revanche, seules quelques études récentes se sont intéressées aux personnages de bande dessinée. Elles montrent que les techniques appliquées aux images de scènes réelles ne sont pas les plus adaptées à l'analyse de dessins. Ces études traitent de la détection de fraude dans les mangas [14], de la reconnaissance de cartoons [15], de la génération automatique de BD à partir d'un film [16, 17] ou encore de la détection automatique des personnages principaux avec des approches basées sur la redondance de structures [18, 19]. Nous proposons une approche originale non pas pour détecter directement les personnages mais pour réduire l'espace de recherche dans lesquels ils pourraient se trouver. L'objectif visé à travers la définition de ces régions d'intérêt est de 1) réduire en amont le temps de calcul nécessaire à la détection des personnages, 2) limiter le risque d'erreurs potentielles en réduisant la surface de recherche.

3 Méthode de localisation contextuelle des personnages

Dans cette section nous présentons comment localiser les personnages dans une page de bande dessinée en fonction de l'organisation des cases et des bulles, déjà étudiées dans

la littérature. L'objectif est de définir des régions d'intérêts contenant au moins un personnage, en fonction de la position des bulles et de la direction des queues dans chaque case. La détection des cases et des bulles de dialogues ayant déjà été étudiée, nous proposons une méthode de détection de la position et direction des queues de bulles afin de définir les régions d'intérêts des personnages.

3.1 Extraction de la queue des bulles

Les queues des bulles indiquent au lecteur les relations entre des ensembles d'éléments textuels ou graphiques contenus dans les bulles et les personnages ou autres objets présents, ou non, dans les cases. Dans le cas où une bulle possède une queue, elle est généralement représentée par une discontinuité sur son contour en direction du personnage concerné (voir fig. 4). Nous limiterons cette étude aux types de queue pouvant être considérés comme un prolongement de la région intérieure de la bulle à savoir les types "virgule" et "zigzag". Nous nous focalisons sur ces deux cas car ce sont les plus fréquemment utilisés et que les autres types, de part leur différences, nécessitent une étude bien plus spécifiques.

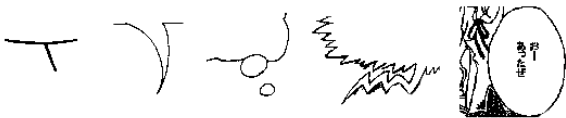


FIGURE 4 – Exemples de types de queues de bulles. De gauche à droite : "trait", "virgule", "cercle", "zigzag" et "absente".

Afin de localiser plus précisément les personnages par rapport aux bulles, nous avons besoin de connaître la position de la queue, ainsi que la direction vers laquelle elle pointe. Pour cela nous définissons un vecteur \vec{be} où b représente les coordonnées du point du début de la queue et e son extrémité. Notons que le point b ne se trouve pas nécessairement à l'endroit où la queue est rattachée à la bulle. En effet, dans le cas où la queue change de direction plusieurs fois avant d'atteindre e , le point b devra être au plus proche de l'extrémité e afin d'indiquer une direction pertinente correspondant à la direction du dernier segment de la queue.

Dans un premier temps, nous détectons la position de l'extrémité e de la queue sur le contour de la bulle en analysant la distance Euclidienne entre son contour original o et son contour lissé. Le contour original o est créé à partir de la distance, en nombre de pixels, qui sépare le barycentre de la bulle et chaque point de son contour (voir fig. 5).

Le lissage du contour nous permet de mesurer les irrégularités du contour dont celles créées par présence de la queue. Le point du contour original o où l'écart est maximal définit le point d'extrémité e de la queue. Le point b est défini à équidistance entre les deux premiers points de contact des signaux o et s de part et d'autre du pic maximum e (voir fig. 5). Lorsque l'on reporte le point b dans l'image, il se situe au sein de la queue à proximité de son extrémité et

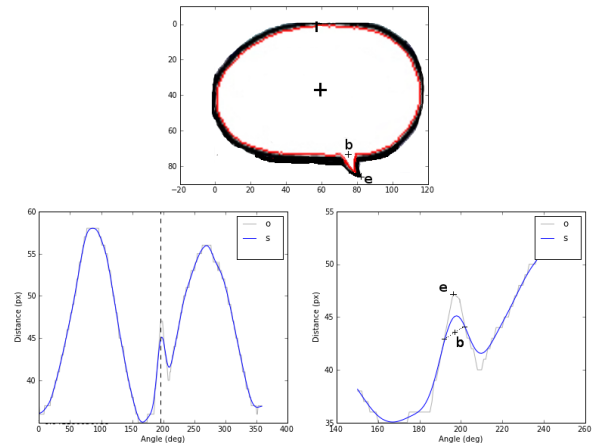


FIGURE 5 – Exemple d'une bulle (en haut) et de son contour original o , du contour lissé s (à gauche) et de la détection des points e et b de la queue (à droite). La mesure du contour o correspond, sur la bulle, à la distance entre son barycentre (croix noire) et le contour de la bulle sur 360° à partir du repère noir (en haut de la bulle) et dans le sens antihoraire.

permet de définir le vecteur de direction \vec{be} .

Dans le cas de queues courtes et régulières, la différence entre les signaux o et s est telle que le point b indique également le point de rattachement de la queue à la bulle comme sur la figure 5.

3.2 Localisation des personnages

L'objectif est ici de définir une sous-partie de la case où pourrait, potentiellement, être dessiné un ou plusieurs personnages. Les informations dont nous disposons pour baser notre réflexion sur une page donnée sont la position des cases, la position des bulles et le vecteur directeur de leur queue. Il faut donc, dans un premier temps, déterminer à quelle case appartient chaque bulle afin de délimiter l'espace de recherche des personnages à la case puis, dans un second temps, réduire la région d'intérêt en fonction des vecteurs directeurs des queues présentes des les cases.

Dans la grande majorité des cas, la bulle est spatialement incluse dans la case. Une simple vérification d'inclusion du polygone de la bulle dans celui de la case suffit alors pour définir une relation d'appartenance scénaristique entre les deux. Cette relation est formalisée dans notre modèle par la propriété *hasPart* entre une instance de *Case* et une instance de *Bulle* (voir fig. 1). Dans le cas où une bulle possède une intersection avec plusieurs cases, l'extrémité de sa queue fait foi, affiliant la bulle à la case contenant sa queue. Le vecteur directeur de la queue d'une bulle, calculé section 3.1, nous donne une indication sur la position probable du personnage dans la case, par rapport à la bulle. Nous avons alors la possibilité de déterminer la région d'intérêt de différentes manières selon la direction pointée par ce vecteur. Nous considérons en effet les huit directions cardinales (Nord, Nord-Ouest, Ouest, Sud-Ouest, Sud, Sud-Est, Est, Nord-Est) et échantillons le cercle trigonométrique

en huit angles égaux à $\pi/4$ afin de classifier la direction du vecteur. Dans le cas où le vecteur pointe vers le coin d'une case, la région d'intérêt s'étendra d'un point de départ p au coin de la case correspondant. Dans le cas où le vecteur pointe vers le bord d'une case, la région d'intérêt couvre toute la case, depuis p , vers la direction correspondante. Dans tous les cas, le point de référence p peut être positionné, au choix, à l'extrémité de la queue ou bien au centre de la bulle comme illustré sur la figure 6.



a) Extrémité de la queue b) Centre de la bulle

FIGURE 6 – Différentes possibilités de définition de la région d'intérêt avec une direction Sud-Est. Crédits image [13].

4 Évaluations

Afin d'évaluer nos différentes contributions, nous avons besoin d'un ensemble de données de références concernant la position et l'orientation des bulles ainsi que la position des personnages. À notre connaissance, il n'existe pas à ce jour de tels jeux de données publics. Nous nous sommes donc basés sur la base d'images publiques de bandes dessinées eBDtheque [20] auquel nous avons ajouté les éléments de vérité terrain nécessaires à cette étude. Ces éléments complémentaires sont détaillés sections 4.1 et 4.2.

4.1 Position et orientation des queues

Cette section évalue la détection de la position et de la direction des queues proposée section 3.1. Pour effectuer cette évaluation nous avons utilisé un sous ensemble de 373 bulles fermées de la base eBDtheque que nous avons segmentées manuellement au niveau pixel en considérant la région intérieure de la bulle. Ce sous ensemble comporte 71 bulles ne possédant pas de queue, nous avons donc évalué notre méthode sur les 302 bulles comportant une queue.

Positionnement. L'évaluation du positionnement des queues consiste à vérifier la détection de son extrémité e vis à vis du contour de la bulle. Ce point est considéré comme correct s'il a été détecté entre le point d'intersection de la bulle et de la queue, et l'extrémité réelle de la queue.

Les résultats ont montré un taux de bonnes détections de 81.2% pour les 302 bulles testées. Parmi les erreurs de détection, 14% étaient dues à une détection des queues dans les angles de bulles partiellement rectangulaires ou dans les pics sur le contour des bulles (voir fig. 7). Les 4.8% d'erreurs restants sont principalement liés à une erreur de mesure de distance entre le centre de la bulle et son bord, comme on peut le voir sur la bulle en bas à droite sur la fi-

gure 7 où l'extrémité de la queue est plus proche du centre de la bulle que son point de rattachement à la bulle.

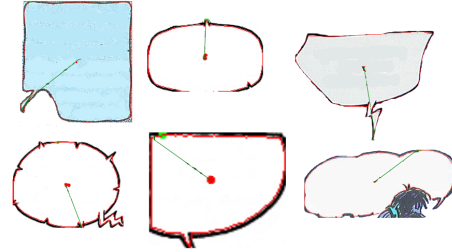


FIGURE 7 – Exemples de bonnes (première ligne) et mauvaises détections (seconde ligne) de positions de queues de bulles. Le trait partant du barycentre indique où a été détecté la queue sur le contour.

Orientation. L'orientation de la queue des bulles est liée à la détection de l'extrémité e et du calcul de son point de départ b . Nous avons évalué visuellement l'orientation de la queue suivant les huit directions cardinales pour les bulles où la position de la queue avait été correctement détectée (251 bulles). L'analyse des résultats montre un taux de bonne détection de l'orientation de 87.6% dont 8% des erreurs sont liés à une imprécision due à la continuité de la queue au-delà de la région détectée. Ce type d'erreur est illustré figure 8 où l'on voit que la détection au niveau pixel de la bulle a été effectuée sur le contour intérieur, la partie blanche, alors que l'extrémité de la queue, indiquant une direction différente, est matérialisée par le contour extérieur, le trait noir. La direction de la queue est matérialisée par la direction du vecteur \vec{be} .

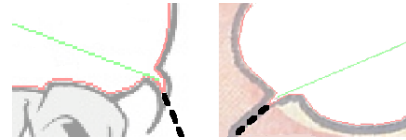


FIGURE 8 – Exemples de mauvaise (gauche) et bonne (droite) détections de l'orientation de queues de bulles (pointillés).

4.2 Localisation des personnages

Afin d'évaluer la détection des personnages, nous avons enrichi la base eBDtheque avec la segmentation des personnages. Le protocole utilisé est le suivant : nous avons enregistré les coordonnées des boîtes englobantes de chaque personnage émettant au moins une bulle au cours de l'album. Cette sélection a pour but de ne conserver que les personnages principaux (ceux qui ont une influence directe sur l'histoire) et d'ignorer les figurants. Certaines parties de personnages (ex : mains, pieds) ont été ignorées dans certaines postures afin de maximiser la surface occupée par le personnage dans sa boîte englobante.

Puisque notre ambition ici est de détecter la zone de présence des personnages émettant une bulle de dialogue, nous avons également réduit la base eBDtheque aux 373 bulles segmentées au niveau pixel. Nous avons conservé les 71

bulles ne possédant pas de queue dans le jeu de données car nous n'étions pas en mesure de les retirer en amont de manière automatique. Elles sont donc intégrées au corpus et nous en tiendrons compte lors de l'analyse des résultats. Ces 373 bulles sont réparties entre 60 pages de 21 albums différents, plus précisément dans 218 cases contenant 497 personnages (locuteur ou non).

Le tableau 1 présente les différentes performances de notre méthode dans les cas suivants :

- (1) Recouvrement oui/non à partir du centre de la bulle. La direction utilisée pour localiser le personnage est celle inscrite dans les métadonnées de la bulle, entrées manuellement dans la base eBDtheque.
- (2) Recouvrement oui/non à partir de l'extrémité de la queue. La direction utilisée est celle calculée automatiquement à partir de la détection de la queue.

Nous considérons qu'il y a recouvrement à partir du moment où la zone identifiée recouvre un personnage de la vérité terrain d'au moins 80%. Nous n'avons pas appliqué de seuil sur la précision de ce recouvrement, l'essentiel étant que le personnage soit inclus dans la zone identifiée. Le taux de réduction moyen de l'espace de recherche (aire cumulée de toutes les cases divisée par l'aire cumulée des personnages validés) est indiqué dans la colonne *A* du tableau 1. Le rappel et la précision de nos résultats sont respectivement indiqués dans les colonnes *R* et *P*.

TABLE 1 – Performances de la méthode proposée.

	<i>R</i> (%)	<i>P</i> (%)	<i>A</i> (%)
(1)	65,2	60,6	57,5
(2)	64	46,4	55,4

Nous pouvons voir que dans les deux cas, 64 à 65% des personnages sont recouverts par une zone d'intérêt. Le ratio moyen de personnages effectifs recouverts par une même zone d'intérêt est de 1,8. Ce ratio s'explique par le choix qui a été fait d'étendre ces régions d'intérêt jusqu'aux bords des cases ce qui augmente la possibilité d'englober plus d'un personnage. La précision des détections est en revanche plus hétérogène, avec un score 60% dans le cas (1) contre 46% dans le cas (2). Cela est dû à la prise en compte, ou non, des bulles de narration, sans queue et donc reliées à aucun personnage. Dans le cas (1), l'information est contenue dans la métadonnée *tailDirection* et la recherche d'un personnage peut donc être ignorée au besoin, chose qu'il est impossible de faire dans le cas (2) car nous n'avons pas d'information a priori sur le type de bulle. C'est ce qui explique la différence de 71 unités entre le nombre d'éléments détectés dans les deux cas.

Il faut toutefois noter que parmi les 497 personnages apparaissant dans les cases analysées, certains ne sont potentiellement pas détectables par notre méthode, car n'émettant pas de bulle de dialogue. En partant du principe qu'une bulle n'est reliée qu'à un personnage, et en omettant les cas où un personnage prononce plusieurs bulles, le nombre de personnages à détecter peut théoriquement être ramené au

nombre de bulles du corpus, c'est à dire 373. Le taux de rappel atteint alors 85% sachant qu'une partie des personnages détectés appartient à l'ensemble des non détectables (non locuteurs). Cette situation s'explique par le fait que notre méthode définit une région étendue autour des locuteurs qui peut aussi recouvrir des non locuteurs se trouvant à proximité. Enfin, il est à noter une réduction de l'espace de recherche de plus de 55% dans les deux cas étudiés.

5 Discussion

La méthode présentée permet de localiser les locuteurs en fonction de la queue des bulles, cependant, il existe des bandes dessinées où les bulles n'ont pas de queue. Dans ce cas, une approche différente basée sur la proximité des bulles et des personnages pourrait être étudiée.

La sélection de la position de la queue est effectuée à partir de l'écart maximum entre le contour réel et son lissage. Une des limites de cette technique apparaît dans le cas où plusieurs points se trouvent à une distance du même ordre de grandeur que le maximum. Le choix de ce dernier devient alors moins pertinent. Une amélioration possible serait d'extraire plusieurs positions de queue et de laisser le système d'inférence choisir la plus appropriée en fonction du contexte (e.g. bordure de case, autres bulles).

Nous pourrions également affiner les résultats en exploitant plus en profondeur, et dans un second temps, les relations spatiales entre les éléments au sein d'une même page. Il est par exemple peu probable que deux personnages occupent le même espace dans l'image. Les intersections des régions d'intérêt calculées peuvent alors être mises à profit afin d'affiner la précision de ces zones autour de personnages individuels et ainsi faire baisser le taux de sursegmentation. De même, lorsqu'une bulle est adjacente à un bord de l'image et que sa queue pointe vers ce même bord, il est envisageable de considérer que le personnage émetteur n'apparaisse pas dans cette case mais dans une case adjacente.

6 Conclusion

Nous avons présenté une méthode permettant de localiser avec précision les personnages émettant des bulles de dialogue au sein de cases de bande dessinée. Les expérimentations menées ont montré la capacité du système à localiser les personnages liés à une bulle et ainsi à réduire le temps de traitement nécessaire à l'identification ultérieure de ces personnages. Des pistes ont également été proposées pour affiner les résultats, et ce à différents niveaux du système. Nos futurs travaux porteront sur la combinaison de cette méthode avec une méthode de détection de régions d'intérêts redondantes, pour permettre une extraction effective des personnages.

7 Remerciements

Ce travail est réalisé dans le cadre de l'action eBDtheque du Contrat de plan État-Région (CPER) avec le soutien financier de la région Poitou-Charentes, de conseil général

de Charente Maritime et de la communauté d'agglomération de La Rochelle.

Références

- [1] S. McCloud, *Understanding comics*. William Morrow Paperbacks, 1994.
- [2] N. Tsopze, C. Guérin, K. Bertet, and A. Revel, "Ontologies et relations spatiales dans la lecture d'une bande dessinée," in *Proceedings of Ingénierie des Connaissances 2012*, 2012. [Online]. Available : <http://www.springerlink.com/index/C2002217522352P1.pdf>
- [3] Y. In, T. Oie, M. Higuchi, S. Kawasaki, A. Koike, and H. Murakami, "Fast frame decomposition and sorting by contour tracing for mobile phone comic images," *International journal of systems applications, engineering and development*, vol. 5, no. 2, pp. 216–223, 2011.
- [4] L. Li, Y. Wang, Z. Tang, and L. Gao, "Automatic comic page segmentation based on polygon detection," *Multimedia Tools and Applications*, vol. 11042, pp. 1–27, 2012. [Online]. Available : <http://dx.doi.org/10.1007/s11042-012-1241-7>
- [5] C. Rigaud, N. Tsopze, J.-C. Burie, and J.-M. Ogier, "Extraction robuste des cases et du texte de bandes dessinées," in *Proceedings of Colloque International Francophone sur l'Ecrit et le Document (CIFED)*, 2012, pp. 349–360.
- [6] A. K. N. Ho, J.-C. Burie, and J.-M. 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] C. Ponsard, R. Ramdoyal, and D. Dziamski, "An ocr-enabled digital comic books viewer," in *Computers Helping People with Special Needs*. Springer, 2012, pp. 471–478.
- [8] M. Stommel, L. I. Merhej, and M. G. Müller, "Segmentation-free detection of comic panels," in *Computer Vision and Graphics*. Springer, 2012, pp. 633–640.
- [9] N. Cohn, "Navigating comics : an empirical and theoretical approach to strategies of reading comic page layouts," *Frontiers in psychology*, vol. 4, 2013.
- [10] K. Arai and H. Tolle, "Method for real time text extraction of digital manga comic," *International Journal of Image Processing (IJIP)*, vol. 4, no. 6, pp. 669–676, 2011.
- [11] A. K. N. Ho, J.-C. Burie, and J.-M. Ogier, "Panel and Speech Balloon Extraction from Comic Books," in *2012 10th IAPR International Workshop on Document Analysis Systems*. Ieee, Mar. 2012, pp. 424–428.
- [12] C. Rigaud, D. Karatzas, J. Van de Weijer, J.-C. Burie, and J.-M. Ogier, "An active contour model for speech balloon detection in comics," in *Proceedings of the 12th International Conference on Document Analysis and Recognition (ICDAR)*. IEEE, 2013.
- [13] Cyb, *Bubblegôm Gôm vol. 1, pp. 3*. Goven, France : Studio Cyborga, 2009, vol. 1.
- [14] W. Sun and K. Kise, "Detection of exact and similar partial copies for copyright protection of manga," *International Journal on Document Analysis and Recognition (IJ DAR)*, vol. 16, no. 4, pp. 331–349, 2013. [Online]. Available : <http://dx.doi.org/10.1007/s10032-013-0199-y>
- [15] F. S. Khan, M. A. Rao, J. van de Weijer, A. D. Bagdanov, M. Vanrell, and A. Lopez, "Color attributes for object detection," in *Twenty-Fifth IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012)*, 2012.
- [16] M. Wang, R. Hong, X.-T. Yuan, S. Yan, and T.-S. Chua, "Movie2comics : Towards a lively video content presentation." *IEEE Transactions on Multimedia*, vol. 14, no. 3-2, pp. 858–870, 2012.
- [17] T. Sawada, M. Toyoura, and X. Mao, "Film comic generation with eye tracking," in *Advances in Multimedia Modeling*. Springer, 2013, pp. 467–478.
- [18] W. Sun, K. Kise, J.-C. Burie, and J.-M. Ogier, "Specific comic character detection using local feature matching," in *Proceedings of International Conference on Document Analysis and Recognition (ICDAR 2013)*, Washington, USA, 2013.
- [19] H. N. Ho, C. Rigaud, J.-C. Burie, and J.-M. 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.
- [20] C. Guérin, C. Rigaud, A. Mercier, F. Ammar-Boudjelal, K. Bertet, A. Bouju, J.-C. Burie, G. Louis, J.-M. Ogier, and A. Revel, "ebdtheque : a representative database of comics," in *12th International Conference on Document Analysis and Recognition (ICDAR)*, 2013.