

# Analyse Structurale en Interprétation de Documents: Un Bref Survol

Delalandre Mathieu<sup>1</sup>, Trupin Eric<sup>1</sup>, Ogier Jean-Marc<sup>2</sup>

<sup>1</sup>Laboratoire PSI, Université de Rouen, 76 821 Mont Saint Aignan, France

<sup>2</sup>Laboratoire L3I, Université de La Rochelle, 17042 La Rochelle, France

<sup>1</sup>{mathieu.delalandre;eric.trupin}@univ-rouen.fr ; <sup>2</sup>jmogier@univ-lr.fr

## Résumé

*L'interprétation d'images de documents est un domaine de recherche qui regroupe trois grandes applications : l'écriture manuscrite, les documents graphiques, et les documents structurés. Un des champs d'investigation de ce domaine est la reconnaissance de forme structurale. Celle-ci a pour but de reconnaître les formes à partir de leur représentation sous forme de graphes. Appliquée à l'image, elle se décompose en deux étapes : l'étape d'analyse structurale extrait les graphes de l'image et l'étape de reconnaissance exploite ces graphes extraits. Nous présentons dans cet article, à travers une étude bibliographique, un bref survol de l'analyse structurale. Puis, nous dressons un bilan comparatif des méthodes d'analyse structurale étudiées et dégageons quelques thématiques de recherche, en particulier celle concernant la combinaison des méthodes.*

## Abstract

*The documents interpretation is a domain research which covers three main applications: the handwriting, the graphic documents, and the structured documents. One of studies in this domain is the structural pattern recognition. It recognizes the patterns according to their description graphs. Applied to images, it uses two steps: the structural analysis step extracts the image graphs and the recognition step exploits these extracted graphs. We present in this article a primer on the structural analysis. Then, we present a comparison between the structural analysis methods, and give some current research topics, especially the methods combination.*

## 1. Introduction

L'interprétation d'images de documents est un domaine de recherche qui regroupe trois grandes applications : l'écriture manuscrite [65], les documents graphiques (documents techniques [57], symboles [38]), et les documents structurés [45]. L'interprétation de documents est un problème qui relève des techniques de l'intelligence artificielle faisant appel à trois entités différentes [29] : un système de contrôle, une librairie de reconnaissance de forme dédiée majoritairement aux images de documents<sup>1</sup>, et une base de connaissances. Ces trois entités sont fortement liées, et plusieurs travaux à la croisée de celles-ci ont été réalisés ces 15 dernières années [14].

Cet article aborde plus précisément la reconnaissance de forme. Un processus de reconnaissance de forme est classiquement décomposé en deux étapes [30]. La première étape est une étape de traitement d'images qui a deux objectifs : le pré traitement des images, et l'extraction de caractéristiques décrivant les formes des images. Dans la suite de cet article, nous emploierons, comme [40], le terme 'analyse' pour qualifier l'étape d'extraction de caractéristiques. La seconde étape est l'étape de 'reconnaissance' qui exploite les graphes extraits durant l'étape d'analyse. Ce terme de reconnaissance est relativement restrictif car cette étape peut avoir d'autres applications comme l'apprentissage [43], l'indexation [17], la recherche de zone d'intérêt [16], la structuration et la correction des données [66], etc. Il existe deux grandes approches de reconnaissance de forme : l'approche statistique & connexionniste [28], et l'approche syntaxique & structurale [60]<sup>2</sup>. De ces deux approches, cet article porte sur la reconnaissance de forme structurale. Celle-ci a pour but de reconnaître les formes à partir de leurs représentations sous forme de graphes (extraits durant l'étape d'analyse). De nombreuses formes se prêtent à des descriptions structurales en interprétation de documents comme : les formulaires [3], les caractères orientaux [32], et les parties graphiques [38] [57].

---

<sup>1</sup>Nous omettrons dans la suite de l'article la mention 'dédiée aux images de documents'.

<sup>2</sup>Pour alléger le texte, nous emploierons dans la suite de l'article le terme structurel (le), pour syntaxique & structurel (le).

L'étape de reconnaissance structurelle est un problème d'exploitation de graphes que l'on peut découper en deux grandes familles d'approches: les approches à base de parcours de graphes [23], et les approches à base de grammaire de graphes [5]. Les premières parcourent les graphes dans le but d'en extraire des informations, par exemple pour appairer les graphes extraits avec des graphes modèles. Les secondes appliquent un ensemble de règles de transformation (la grammaire) afin de faire converger (ou corriger) les graphes candidats vers les graphes modèles. Un problème d'exploitation de graphes dépend de deux critères principaux : exploitation de graphes/sous-graphes, exploitation de graphes exacts/inexactes. Un sous-graphe est un sous-ensemble de nœuds et d'arcs appartenant à un graphe de taille plus importante. Lorsque les graphes candidats traités sont strictement équivalents aux graphes modèles, on dit que le problème de graphes est exact. Cependant en interprétation de documents, les formes sont souvent connectées à d'autres éléments, et les graphes extraits sont souvent de taille importante et bruités. Ainsi c'est souvent un problème de sous-graphes inexactes [39], appartenant à des graphes de larges tailles.

L'étape d'analyse extrait (ou construit) les graphes de l'image. Nous nommons simplement cette étape, dans cet article, analyse structurelle. Il existe globalement deux approches d'analyse structurelle, l'analyse locale et l'analyse globale<sup>3</sup>. La frontière entre ces deux analyses est la composante connexe. L'analyse locale extrait les objets du document en décomposant les composantes connexes, et l'analyse globale en les groupant. La Figure 1 illustre la différence entre ces deux analyses. L'analyse locale (bas gauche) décompose une composante connexe (haut) en arcs, jonctions et vecteurs. L'analyse globale (bas droit) groupe trois composantes connexes selon des contraintes de distance.

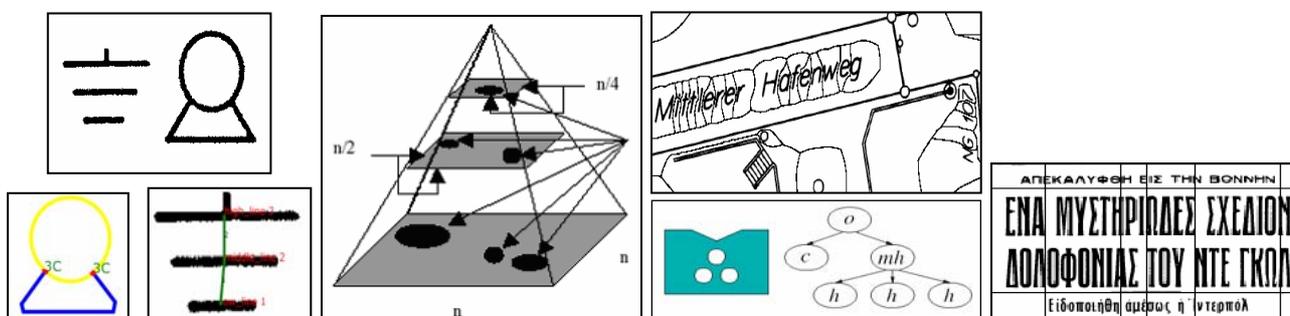


Figure 1 : Globale et locale    Figure 2 : Multi résolution    Figure 3 : Analyse directe    Figure 4 : Analyse indirecte

Cet article aborde le domaine de l'analyse structurelle. Nous proposons ici un 'bref survol' de ce domaine. Dans la suite de l'article, nous présentons dans la section 2 l'analyse structurelle globale. Dans la section 3, nous présentons l'analyse structurelle locale. A l'issue de chacune de ces deux présentations, nous dressons un bilan comparatif des méthodes d'analyse structurelle étudiées, et dégageons quelques thématiques de recherche actuelles. Dans la section 4, nous abordons plus spécifiquement la thématique de la combinaison des méthodes. Enfin, dans la section 5, nous concluons.

## 2. Analyse globale

### 2.1. Introduction

L'analyse globale exploite les relations spatiales entre les objets du document. L'analyse globale est très généralement utilisée pour l'interprétation de documents structurés (formulaires [3], chaînes de caractères [41], formules mathématiques [34]). On l'utilise également pour l'analyse de documents graphiques (plans techniques [7], symboles [15]). L'analyse globale extrait différents objets du document. Ces objets se déclinent selon différents niveaux. Cette notion de niveau est très 'subjective', celle-ci étant fonction de la granularité d'exploration du document. Le plus petit objet est la composante connexe, et le plus haut objet est le document lui-même. Entre ces deux extrêmes, le type d'objet est fonction du type de document : symbole, mot, paragraphe, etc. Le problème est donc le choix de la granularité d'exploration dans le document, l'analyse multi résolution [59] illustre ce principe (Figure 2). Il existe deux approches d'analyse globale, l'analyse indirecte et directe. L'analyse indirecte extrait les objets à partir des composantes connexes. L'analyse directe extrait directement les objets de l'image. Nous présentons ces deux approches dans la section suivante (section 2.2).

<sup>3</sup> Cette classification existe également en analyse statistique [47].

## 2.2. Analyse indirecte et directe

L'analyse indirecte exploite les informations sur les composantes connexes pour l'extraction des objets. Deux méthodes sont couramment utilisées, le marquage des composantes connexes, et les diagrammes de Voronoi. Les systèmes basés sur le marquage des composantes connexes [41] [4] [15] utilisent des contraintes spatiales (distance Euclidienne, alignement, etc.) de façon à grouper les composantes connexes (Figure 1 (bas droit)). Les systèmes basés sur l'utilisation des diagrammes de Voronoi [7] [34] exploitent les informations de voisinage entre les composantes connexes, obtenues à partir des diagrammes de Voronoi [7] (Figure 3 (haut)).

L'analyse directe extrait directement des objets de plus haut niveau. Elle tente donc de localiser directement la structure de ces objets dans l'image. Elle exploite des méthodes très diverses, car le type de traitement est fonction de l'objet à extraire. Ces objets peuvent être génériques ou fortement dédiés au type de document analysé. [22] propose par exemple un découpage en zones rectangulaires de l'image, pour extraire les différents éléments des pages de journaux (titres, paragraphes, etc.) (Figure 4). [3] exploite l'algorithme RLSA<sup>4</sup> pour la recherche de mots sur les documents à caractères imprimés.

## 2.3. Conclusion, analyse globale

Nous comparons dans le Tableau 1 les méthodes d'analyse globale en fonction de trois critères : le critère 'Segmentation' renseigne sur la capacité des méthodes à extraire correctement les objets du document, le critère 'Topologie' renseigne sur la capacité des méthodes à exprimer les relations spatiales entre les objets du document, et le critère 'Sémantique' renseigne sur l'apport sémantique (ou ajout de sens) des méthodes.

Les méthodes indirectes segmentent exactement les objets : les graphes construits contiennent les informations de chacune des composantes connexes. De plus, ces méthodes ne sont pas sensibles aux problèmes d'orientations [15] [34]. Les méthodes directes effectuent déjà un groupement des composantes connexes durant l'étape d'analyse, la segmentation peut donc être inexacte. De plus, certaines méthodes employées peuvent être très sensibles à l'orientation [22]. L'analyse de la topologie, dans le cas direct, dépend des méthodes employées. En effet, les systèmes basés sur les diagrammes de Voronoi sont plus aptes à extraire les relations de voisinage [34], et ceux utilisant le marquage de composantes connexes les relations d'inclusions et la présence d'occlusions [4] (Figure 3 (bas)). Pour les méthodes indirectes, cette capacité peut différer: [22] par exemple exprime les relations de voisinage entre ses objets, car ceux-ci sont agencés sous forme de pavage). Les objets extraits par les méthodes indirectes ont une faible sémantique (ils doivent être enrichi [15]), ceux-ci ne décrivent que des composantes connexes. Ce n'est pas le cas des méthodes directes qui labellisent déjà les objets (mots [3]) (paragraphes, images, titres, etc. [22]), mais cette labellisation est approximative et donc inexacte.

**Tableau 1 : Comparaison des méthodes globales**

	<b>Segmentation</b>	<b>Topologie</b>	<b>Sémantique</b>
<b>Indirect (Voronoi)</b>	Exacte/ Invariante	Voisinage	Faible/Exacte
<b>Indirect (MMC)</b>	Exacte/ Invariante	Inclusion	Faible/Exacte
<b>Direct</b>	Inexacte/~	~	Forte/Inexacte

Les thématiques de recherche actuelles de l'analyse globale concernent ses interactions avec l'ensemble du système d'interprétation. [3] exploite par exemple une grammaire afin de construire des paragraphes à partir de l'information des mots extraits par l'algorithme RLSA, ce qui fournit le même niveau de données que [22]. [21] réalise le même type de traitement à des documents peu structurés à l'aide d'une grammaire probabiliste. Dans [15] nous utilisons les informations d'une classification statistique afin de guider notre processus d'analyse globale. Nous exploitons également une approche stratégique afin de combiner nos méthodes. [51] exploite une méthode de classification dynamique suite à l'extraction de ses paragraphes, afin de les labelliser.

<sup>4</sup> Run Length Smoothing Algorithm

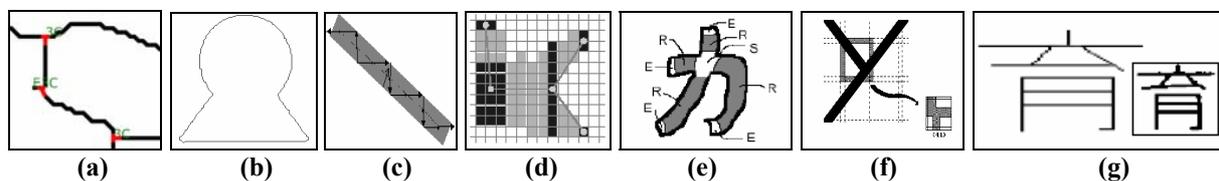
### 3. Analyse locale

#### 3.1. Introduction

L'analyse locale emploie différentes approches issues de l'analyse de caractères (latins [62] et asiatiques [20]), et de l'analyse d'éléments graphiques (documents techniques [37], symboles [38], etc.). Comme l'analyse globale, l'analyse locale extrait différents objets du document fonction de la granularité d'exploration. Le plus petit objet est le pixel, et le plus haut objet la composante connexe. Nous proposons ici une présentation de l'analyse locale en quatre étapes : extraction du graphe d'objets, approximation des objets, construction d'objets haut-niveaux, et correction du graphe d'objets. Nous présentons chacune des quatre étapes dans les sections 3.2, 3.3, 3.4, et 3.5.

#### 3.2. Extraction du graphe d'objets

L'étape d'extraction du graphe d'objets décompose les composantes connexes en graphes d'objets. Nous avons répertorié sept familles de méthodes que nous illustrons dans la Figure 5 : les méthodes à base de, squelettisation (a), détection de contours (b), parcours des formes (c), décomposition en plages (d), décomposition en régions (e), décomposition en mailles (f), et segmentation d'objets (g).



**Figure 5 : (a) Squelettisation (b) Détection de contours (c) Parcours des formes (d) Décomposition en plages (e) Décomposition en régions (f) Décomposition en mailles (g) Segmentation d'objets**

Les méthodes à base de squelettisation sont les plus utilisées. Elles se décomposent en deux étapes. La première étape extrait les images de squelettes [31]. Il existe deux grandes familles de squelettiseurs : par amincissements successifs et par transformées de distances [1]. La seconde étape analyse les images de squelettes afin d'extraire des graphes de pixels [61]. Elle utilise différentes méthodes, basées sur l'analyse de la connexité [61] [49], ou la destruction des pixels 3-connexes et plus [32] [16]. La Figure 5 (a) donne un exemple de graphe de pixels extrait [16].

Les méthodes à base de détection de contours sont également couramment utilisées. Il existe deux catégories de méthodes [63]. Les premières utilisent des représentations intermédiaires sous forme d'images des contours. Elles sont basées sur des méthodes de morphologie mathématique [24] ou de test de voisinages [16]. Comme les méthodes à base de squelettisation, elles utilisent une seconde étape afin d'extraire les chaînes de pixels des contours. Les secondes extraient les chaînes de pixels des contours de l'image initiale. Elles utilisent des méthodes de suivi de contours [1], ou de marquage de composantes connexes [16]. Leur avantage est de pouvoir exprimer les relations d'inclusions entre les chaînes, et de sélectionner les contours internes et/ou externes aux formes [16]. La Figure 5 (b) donne un exemple de chaîne de contours extérieurs extraite de la Figure 1 (haut) [16].

Les méthodes à base de parcours des formes analysent directement les images sans utiliser de représentations intermédiaires. Elles sont basées sur l'utilisation d'éléments structurants pour parcourir les formes, de type pixellaire [18] [54], ou surfacique (cercle [13], élément gaussien [68]). Elles produisent des graphes d'objets géométriques (cercles et vecteurs [54]), ou de pixels [13]. Ceci est fonction du modèle de parcours adopté (linéaire [18], circulaire [54]), et de la progression de l'élément structurant dans la forme (continu [13], par pas [18]). La procédure d'analyse peut se diviser en deux : l'analyse des lignes et des jonctions. Dans les deux cas, l'élément structurant employé peut être de type pixellaire [18] [54], de type surfacique [13] [68], ou la méthode peut être hybride [48]. La Figure 5 (c) donne un exemple de parcours de ligne avec un élément structurant de type pixellaire [18].

Les méthodes à base de décomposition en plages sont utilisées dans plusieurs domaines, comme la reconnaissance de symboles [8] et la reconnaissance de l'écriture manuscrite chinoise [19] [67]. Une plage est une séquence de pixels noirs, horizontale ou verticale. Partant de cette définition, [8] construit le « Mixed Run Graph », un graphe de plages verticales et horizontales (Figure 5 (d)). Les plages 2-connexes successives sont concaténées par des arcs dans le graphe, et les plages verticales et horizontales juxtaposées sont assimilées à des nœuds jonctions.

Les méthodes à base de décomposition en régions sont peu utilisées. Le principe est de décomposer la composante connexe en un ensemble de régions. [11] et [9] calculent pour cela des informations d'orientations pour chacun des pixels de l'image avec les pixels contours de la forme. Ils peuvent ainsi par la suite rechercher les pics des directions majoritaires pour chacun des pixels. Il détermine ainsi si un pixel appartient à une région ligne, extrémité, ou jonction (Figure 5 (e)). Dans [16] nous proposons une méthode de décomposition en régions basée sur une technique d'agrégation par vagues. Les ruptures ou les blocages des vagues déterminent les frontières entre les régions. Dans une seconde passe les graphes de régions obtenus sont analysés afin de construire les régions 'jonction' et 'ligne' des formes.

Les méthodes à base de décomposition en mailles n'ont pour le moment été utilisées que pour des applications de vectorisation [35] [64]. L'image est d'abord découpée en un pavage de zones régulières : les mailles. Par la suite, les mailles sont reconnues selon une librairie de mailles (Figure 5 (f)) [64]. La carte de mailles reconnues est ensuite analysée de façon à construire les relations structurelles entre les mailles.

Les méthodes à base de segmentation d'objets sont couramment utilisées en vision [44] ainsi qu'en interprétation de documents [58]. Le principe est de segmenter directement des objets de l'image comme les lignes, les arcs, et les ellipses. Ces méthodes emploient des transformées mathématiques afin de 'transposer' les images dans un nouvel espace de représentation. Celui-ci est exploité afin de rechercher les objets en fonction de leurs modèles mathématiques. [58] extrait ainsi les lignes diagonales, verticales, et horizontales pour la reconnaissance de l'écriture manuscrite chinoise (Figure 5 (g)). Certains travaux portent sur la segmentation directe des jonctions (jonction en 'T', en 'X', etc.) [12]. Plusieurs méthodes existent, parmi elles la transformée de Hough [44] et les filtres de Gabor [12] sont couramment utilisés. Si la méthode employée ne permet pas la segmentation des jonctions, celles-ci sont reconstruites par analyse classique des croisements géométriques [42], et/ou par recherche des proximités entre les extrémités des objets segmentés [58].

### 3.3. Approximation mathématique

Durant cette étape, les objets extraits précédemment sont approximés mathématiquement. Différents objets peuvent être construits comme les arcs, les vecteurs, et les courbes (Figure 1 (bas gauche)). La particularité de cette étape est l'hétérogénéité des données d'entrée acceptées : points (pixels), vecteurs, courbes, et cercles. On peut très bien approximer des vecteurs en cercle(s), une courbe en cercle(s), etc. Les graphes de pixels et de vecteurs (Figure 5 (a), (b), (c)) sont couramment utilisés. Les graphes de plages et de régions (Figure 5 (d), (e)) sont également exploitables car on peut en extraire les squelettes et les contours [19] [16]. Les graphes de mailles et d'objets segmentés (Figure 5 (f), (g)) peuvent être utilisés, mais leurs objets sont déjà fortement approximés, ce qui en limite l'intérêt. En pratique, les graphes et chaînes de pixels sont les plus utilisés pour la vectorisation [37] [61]. Un algorithme permettant la combinaison d'algorithmes d'approximation, ainsi qu'un état de l'art, peuvent être trouvés dans [52]. De même, un algorithme permettant d'extraire des indices contextuels sur la qualité des données à approximer (de façon à guider le système dans le choix de l'algorithme d'approximation) peut être trouvé dans [53].

### 3.4. Construction d'objets haut-niveaux

Après l'étape d'approximation mathématique, certains systèmes construisent à partir des objets bas-niveaux obtenus (vecteurs, arcs, courbes) des objets de plus haut niveau comme les cercles, les triangles, les parallélogrammes, etc. Ces objets peuvent être construits par exemple à partir d'images de squelettes [27], ou de contours [33] [10] [69] (on parle d'appariement de contours). L'appariement de contours est généralement utilisé pour reconstruire les jonctions des formes [33]. Durant cette construction d'objets haut-niveaux, des contraintes de positions sur les objets bas-niveaux sont testées : cette étape de construction ne se limite donc pas à un simple procédé de factorisation de graphe. A l'issue de cette construction, une nouvelle représentation peut être créée par l'analyse des relations structurelles entre les objets haut-niveaux [50]. La Figure 6 (a) illustre un cas de reconstruction de cercle à partir d'une image de squelette [27]. La Figure 6 (b) donne un exemple de reconstruction des contours vectorisés en un graphe de triangles [69].

### 3.5. Correction du graphe d'objets

Certains systèmes analysent les graphes d'objets extraits de façon à les corriger. [55] distingue ainsi les systèmes avec corrections 'two steps' et sans correction 'one step'. Ces corrections peuvent se faire en terme de structure des graphes d'objets (ajout/suppression d'objets), ou en terme d'attributs des objets des graphes. Ces méthodes de corrections font intervenir des étapes de traitements d'images, elles ne se limitent donc pas à de la manipulation de graphes. Les règles de corrections exploitées peuvent être figées dans les procédures, ou paramétrables de façon externe. Ces corrections peuvent intervenir sur différents types de données dans (ou entre) chacune des étapes 1, 2, et 3 (sections 3.2, 3.3, et 3.4).

Sur le graphe de pixels plusieurs types de corrections sont exploitables comme : la fusion et l'ébarbulage [16] (Figure 6 (c)), et la fusion des jonctions [36] (Figure 6 (d)). Ces traitements sont également réalisables sur un graphe de vecteurs [18], dans ce cas la correction des nœuds par prolongement vectoriel est également utilisée [26] (Figure 6 (e)). [64] corrige son graphe de mailles par décomposition et la fusion des mailles. [19] corrige son graphe de plages avec des procédures de fusion/suppression des plages segmentées/isolées. [50] corrige son graphe d'objets haut-niveaux par procédure de fusion.

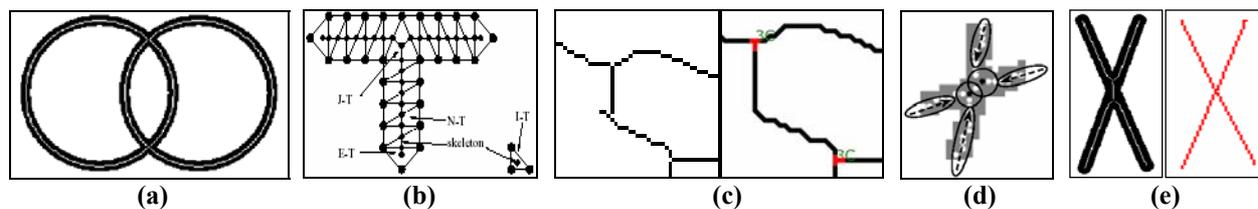


Figure 6 : (a) Construction cercles (b) Appariement contours (c) (d) Correction squelette (e) Correction vectorielle

### 3.6. Conclusion, analyse locale

Nous comparons dans le Tableau 2 les atouts et handicaps des méthodes de l'étape d'extraction du graphe d'objets (section 3.2). Les comparaisons des méthodes des étapes d'approximation mathématique (section 3.3), de construction d'objets haut-niveaux (section 3.4) et de correction du graphe d'objets (section 3.5) sortent du cadre de cet article : nous invitons le lecteur à se reporter respectivement à [52], [46], et [6]. Nous comparons les méthodes de l'étape d'extraction du graphe d'objets en fonction de sept critères : le critère 'Jonction' renseigne sur la capacité des méthodes à détecter les jonctions des formes, le critère 'Morphologie' renseigne sur la capacité des méthodes à analyser des formes hétérogènes, le critère 'Invariance' renseigne sur la capacité des méthodes à résister aux changements d'orientations et d'échelles, le critère 'Sensibilité' renseigne sur la capacité des méthodes à résister aux bruits, le critère 'Sémantique' renseigne sur l'apport sémantique (ou ajout de sens) des méthodes, le critère 'Réversibilité' renseigne sur la capacité des méthodes à restituer les images analysées<sup>5</sup>, le critère 'Complexité' renseigne sur la complexité algorithmique des méthodes.

Les méthodes à base de squelettisation sont invariantes [61], mais ne permettent que l'analyse que des formes linéaires [50] et sont sensibles aux bruits (en particulier sur les zones de jonctions [61]). Les méthodes à base de détection de contours permettent l'analyse de tous types de formes et sont réversibles [24]. Leur défaut est la non détection des jonctions, que l'on peut cependant reconstruire via l'analyse structurale des objets haut-niveaux [33]. Les méthodes à base de parcours des formes permettent une bonne détection des jonctions. De plus, elles exportent des données vectorielles et sont de faibles complexités algorithmiques [54]. Cependant, elles présentent des difficultés lors des variations d'épaisseurs des formes [48]. Les méthodes à base de décomposition en plages [8] permettent une bonne détection des jonctions, l'analyse de tous types de formes, et la réversibilité. Elles sont cependant sensibles à l'orientation et complexes, de part la nature, l'encodage, et la structuration des plages. Les méthodes à base de décomposition en régions permettent également une bonne détection des jonctions, l'analyse de tous types de formes, et la réversibilité. Cependant, les régions brutes sont pauvres d'information (elles doivent être enrichies [16]), et complexes à manipuler. Les méthodes à base de décomposition en mailles permettent une bonne détection des jonctions, l'exportation d'un graphe d'objets divers, et sont de faibles complexités. Cependant ces méthodes sont très sensibles au positionnement initial du pavage de mailles (invariance, sensibilité), et à l'adéquation 'bibliothèque de mailles'/'formes analysées' (morphologie, réversibilité). Les méthodes à base d'objets segmentés extraient des objets géométriques et sont très robustes aux bruits [56]. Cependant les transformées utilisées sont complexes [44], approximativement les formes [56] (réversibilité, morphologie), sont limitées à certaines orientations [58] (invariance), et permettent difficilement la détection des jonctions [12].

<sup>5</sup> Equivalent « structurel » de la propriété mathématique de complétude utilisée en analyse statistique [40]

**Tableau 2 : Comparaison des méthodes locales**

	<b>Atouts</b>	<b>Handicaps</b>
<b>Squelettisation</b>	Invariance	Morphologie, Sensibilité, Complexité
<b>Contours</b>	Morphologie, Invariance, Réversibilité	Jonction
<b>Parcours</b>	Jonction, Sémantique, Complexité	Morphologie
<b>Plages</b>	Jonction, Morphologie, Réversibilité	Invariance, Complexité
<b>Régions</b>	Jonction, Morphologie, Réversibilité	Sémantique, Complexité
<b>Mailles</b>	Jonction, Sémantique, Complexité	Morphologie, Invariance, Sensibilité, Réversibilité
<b>Segmentation</b>	Sensibilité, Sémantique	Jonction, Morphologie, Invariance, Réversibilité, Complexité

Comme pour l'analyse globale, les thématiques de recherche actuelles de l'analyse locale concernent les interactions avec l'ensemble du système d'interprétation. De nombreux travaux ont recours à des étapes de reconnaissance afin d'assister l'analyse locale dans sa construction/correction des graphes, et cela à différents niveaux [2] [58] [66] [8]. D'autres travaux utilisent des approches stratégiques, afin de faciliter la tâche d'interprétation [54], ou de combiner les méthodes [46] [16]. Enfin, certains travaux abordent le problème de la segmentation/reconnaissance par des approches perceptives [68], ou multi agents [25].

#### 4. Combinaison des méthodes

Nous avons présenté et comparé (dans les sections 2 et 3) différentes méthodes d'analyse locale et globale, et dégagé quelques thématiques de recherche actuelles. Une thématique commune est la combinaison des méthodes. Les méthodes combinées peuvent être locales, locales/globales, et globales. Nous proposons ici une classification de cette combinaison en trois catégories : comparative, hybride, et coopérative. La combinaison comparative analyse les formes afin d'extraire différents graphes. Ces graphes sont ensuite comparés durant l'étape de reconnaissance. [46] compare des graphes de vecteurs issus des squelettes et des contours. Dans [15] nous comparons des graphes d'occlusions et de squelette. La combinaison hybride analyse les formes de façon à extraire des graphes hybrides. Ces graphes sont alors des combinaisons de deux (ou plusieurs) méthodes. De plus, les méthodes globales [15] [34] et locales [8] [9] [16] à base d'objets régions permettent d'utiliser une approche statistico-structurale [15]. [4] extrait des graphes de composantes connexes et d'occlusions enrichis des informations locales de la concavité des contours des formes (Figure 3 (bas)). Dans [15] nous utilisons une approche statistico structurelle pour la reconnaissance de graphes d'occlusions. La combinaison coopérative enchaîne les étapes d'analyse afin de simplifier itérativement la complexité du processus dans son ensemble. [54] simplifie son analyse locale en procédant du plus simple au plus complexe. Dans [15] nous simplifions notre analyse globale par l'utilisation d'une analyse locale. Les combinaisons comparative et hybride permettent une multi représentations des formes (ou modèle de graphe adopté). La pluralité de la multi représentation est obtenue de part les combinaisons « locale, locale/globale, globale » et/ou « statistico structurelle ». Des travaux sur la construction, l'exploitation, et l'analyse des multi représentations constituent certainement une perspective de recherche importante de l'analyse structurelle.

#### 5. Conclusion

Dans cet article, nous avons présenté un 'bref survol' de l'analyse structurelle en interprétation de documents. Cette analyse a pour but d'extraire les graphes de l'image, qui sont reconnus par la suite durant l'étape de reconnaissance. Cette analyse se décline selon différents niveaux de granularité d'exploration du document, localement et globalement aux composantes connexes. Une de ses thématiques de recherche est la combinaison des méthodes. Cette combinaison permet la multi représentation des formes, une perspective de recherche importante de l'analyse structurelle.

#### 6. Références

- [1] S. Ablameyko, T.P. Pridmore, "Machine Interpretation of Line Drawing Images", *Springer-Verlag*, 2000.
- [2] M. Ahmed, R. Ward, "A Rotation Invariant Rule Based Thinning Algorithm for Character Recognition", *Pattern Analysis and Machine Intelligence (PAMI)*, 24(12) : 1672-1678, 2002.

- [3] O. Altamura, F. Esposito, D. Malerba, "Transforming Paper Documents Into XML Format with Wisdom++", *International Journal on Document Analysis And Recognition (IJ DAR)*, 3(2) : 175-198, 2001.
- [4] O. El Badawy, M. Kamel, "Shape Representation using Concavity Graphs", *International Conference on Pattern Recognition (ICPR)*, 2002.
- [5] D. Blostein, H. Fahmy, A. Grbavec, "Issues in the Practical Use of Graph Rewriting", *Graphics Recognition (GREC)*, 1995.
- [6] E. Bodansky, A. Gribov, M. Pilouk, "Post-processing of Lines Obtained by Raster-to-Vector Conversion", *Graphics Recognition (GREC)*, 2001.
- [7] M. Burge, G. Monagan, "Using the Voronoi Tessellation for Grouping Words and Multi Part Symbols in Document", *Vision Geometry IV*, 1995.
- [8] M. Burge, W.G. Kropatsh, "A Minimal Line Property Preserving Representation of Line Images", *Structural and Syntactical Pattern Recognition (SSPR)*, 1998.
- [9] R. Cao, C.L. Tan, "A Model of Stroke Extraction from Chinese Character Images", *International Conference on Pattern Recognition (ICPR)*, 2000.
- [10] F. Chang, Y.C. Lu, T. Palvidis, "Feature Analysis Using Line Sweep Thinning Algorithm", *Pattern Analysis and Machine Intelligence (PAMI)*, 21(2) : 145-158, 1999.
- [11] Y.S. Chen, "Segmentation and Association Among Lines and Junctions for a Line Image", *Pattern Recognition (PR)*, 27(9) : 1135-1157, 1994.
- [12] J. Chen, Y. Sato, S. Tamura, "Orientation Space Filtering for Multiple Orientation Line Segmentation", *Pattern Analysis and Machine Intelligence (PAMI)*, 22(5) : 417-429, 2000.
- [13] J. Chiang, S. Tue, "A New Algorithm for Line Image Vectorization", *Pattern Recognition (PR)*, 31(10) : 1541-1549, 1998.
- [14] D. Crevier, R. Lepage, "Knowledge-Based Image Understanding Systems: A Survey", *Computer Vision and Image Understanding (CVIU)*, 67(2), 161-185, 1997.
- [15] M. Delalandre, S. Nicolas, E. Trupin, J.M. Ogier, "Reconnaissance de Symboles par Approche Structurale Globale-Locale, Basée sur l'Utilisation de Scénarios, et Exploitant une Représentation XML des Données", *International Conference on Image and Signal Processing (ICISP)*, 2003.
- [16] M. Delalandre, Y. Saidali, J.M. Ogier, E. Trupin, "Adaptable Vectorisation System Based on Strategic Knowledge and XML Representation Use", *Graphics Recognition (GREC)*, 2003.
- [17] D. Doermann, "The Indexing and Retrieval Document, a Survey", *Technical Report CS-TR-3876, University of Maryland Computer Science Department, USA*, 1998.
- [18] D. Dori, "Sparse Pixel Vectorisation : An Algorithm and its Performance Evaluation", *Pattern Analysis and Machine Intelligence (PAMI)*, 21(3) : 202-215, 1999.
- [19] K.C. Fan, W.H. Wu, "A Run Length Coding Based Approach to Stroke Extraction of Chinese Characters", *Pattern Recognition (PR)*, 33(11) : 1881-1895, 2000.
- [20] J. Fan, "Off-line Optical Character Recognition for Printed Chinese Character-A Survey", *Technical Report, University of Colombia, USA*, 2002.
- [21] N. Gauthier, T. Artieres, "Grammaires Probabilistes Appliquées à la Segmentation de Documents Manuscrits en ligne", *Colloque International Francophone sur l'Ecrit et le Document (CIFED)*, 2002.
- [22] K. Hadjar, O. Hitz, R. Ingold, "Newspaper Page Decomposition using a Split and Merge Approach", *International Conference on Document Analysis And Recognition (ICDAR)*, 2001.
- [23] E. Hancock, R. Wilson, "Graph-Based Methods for Vision: A Yorkist Manifesto", *Structural and Syntactical Pattern Recognition (SSPR)*, 2002.

- [24] Y.M.Y Hasan, L.J Karan, "Morphological Reversible Contour Representation", *Pattern Analysis and Machine Intelligence (PAMI)*, 22(3) : 227-239, 2000.
- [25] T.C. Henderson, L. Swaminathan, "Agent Based Engineering Drawing Analysis", *Symposium on Document Image Understanding Technology (SDIUT)*, 2003.
- [26] X. Hilaire and K. Tombre, "Improving the Accuracy of Skeleton-Based Vectorisation", *Graphics Recognition (GREC)*, 2001.
- [27] X. Hilaire, "Ranvec and the Arc Segmentation Contest", *Graphics Recognition (GREC)*, 2001.
- [28] A.K. Jain, R.P.W. Duin, J. Mao, "Statistical Pattern Recognition : a Review", *Pattern Analysis and Machine Intelligence (PAMI)*, 22(1) : 4-37, 2000.
- [29] J.M. Jolion, "Les Systèmes de Vision", *Hermès Science Publications*, 2001.
- [30] R. Kasturi, L. O'Gorman, V. Govindaraju, "Document Image Analysis : A Primer", *Sadhana*, 27(1) : 3-22, 2002.
- [31] L. Lam, C.Y. Suen, "An Evaluation of Parallel Thinning Algorithms for Character Recognition", *Pattern Analysis and Machine Intelligence (PAMI)*, 17(9) : 914-919, 1995.
- [32] K.K. Lau, P.C. Yuen, Y.Y. Tang, "Stroke Extraction and Stroke Sequence Estimation On Signatures", *International Conference on Pattern Recognition (ICDAR)*, 2002.
- [33] C. Lee, B. Wu, "A Chinese Character Stroke Extraction Algorithm Based on Contour Information", *Pattern Recognition (PR)*, 31(6) : 651-653, 1998.
- [34] I. Leplumey, Ch. Queguiner, "Un Graphe de Voisinage Basé sur l'Utilisation des Distances Discrètes", *Colloque International Francophone sur l'Écrit et le Document (CIFED)*, 2000.
- [35] X. Lin, S. Shimotsuji, M. Mihoh, T. Sakai, "Efficient Diagram Understanding with Characteristic Pattern Detection", *Computer Vision Graphics and Image Processing*, 30 : 84-106, 1985.
- [36] F. Lin, X. Tang, "Off-line Handwritten Chinese Character Stroke Extraction", *International Conference on Pattern Recognition (ICPR)*, 2002.
- [37] W. Liu, D. Dori, "From Raster to Vectors : Extracting Visual Information from Line Drawings", *Pattern Analysis and Applications (PAA)*, 2(2) : 10-21, 1999.
- [38] J. Lladós, E. Valveny, G. Sánchez, E. Martí, "Symbol Recognition : Current Advances and Perspectives", *Graphics Recognition (GREC)*, 2001.
- [39] J. Lladós, E. Martí, J.J. Villanueva, "Symbol Recognition by Error Subgraph Matching Between Region Adjacency Graphs", *Pattern Analysis and Machine Intelligence (PAMI)*, 23(10) : 1137-1143, 2001.
- [40] S. Locarnic, "A Survey of Shape Analysis Techniques", *Pattern Recognition (PR)*, 31(8) : 983-1001, 1998.
- [41] P.K. Loo, C.L. Tan, "Detection of Word Group Based on Irregular Pyramid", *International Conference on Document Analysis And Recognition (ICDAR)*, 2001.
- [42] K. Loudon, "Mastering Algorithms with C", *Erolyles Editions*, 2000.
- [43] Y. Luo, W. Liu, "Engineering Drawings Recognition Using a Case-based Approach", *International Conference on Document Analysis And Recognition (ICDAR)*, 2003.
- [44] J. Matas, C. Galambos, J. Kittler, "Progressive Probabilistic Hough Transform for Line Detection", *Computer Vision and Pattern Recognition (CVPR)*, 1999.
- [45] G. Nagy, "Twenty Years of Document Image Analysis in PAMI", *Pattern Analysis and Machine Intelligence (PAMI)*, 22(1) : 38-62, 2000.
- [46] Y. Nakajima, S. Mori, S. Takegami, S. Sato, "Global Methods for Stroke Segmentation", *International Journal Document Analysis and Recognition (IJ DAR)*, 2 : 19-23, 1999.

- [47] J. Neumann, H. Samet, A. Soffer, "Integration of Local and Global Shape Analysis for Logo Classification", *International Workshop on Visual Form (IWVF)*, 2001.
- [48] J.M. Ogier, C. Olivier, Y. Lecourtier, "Extraction of Roads from Digitized Maps", *European Signal Processing Conference (EUSIPCO)*, 1992.
- [49] D.V. Popel, "Compact Graph Model of Handwritten Images: Integration into Authentication and Recognition", *Structural and Syntactical Pattern Recognition (SSPR)*, 2002.
- [50] J.Y. Ramel, N. Vincent, H. Emptoz, "A Structural Representation for Understanding Line-Drawing Images", *International Journal on Document Analysis And Recognition (IJ DAR)*, 3 : 58-66, 2000.
- [51] L. Robadey, O. Hitz, R. Ingold, "A Pattern-Based Method for Document Structure Recognition", *Document Layout Interpretation and its Applications (DLIA)*, 2001.
- [52] P.L. Rosin, A.W. West, "Nonparametric Segmentation of Curves Into Various Representations", *Pattern Analysis and Machine Intelligence (PAMI)*, 17(12) : 1140-1153, 1995.
- [53] P.L. Rosin, "Techniques for Assessing Polygonal Approximation of Curves", *Pattern Analysis and Machine Intelligence (PAMI)*, 19(6) : 659-666, 1997.
- [54] J. Song, F. Su, C. Tai, S. Cai, "An Object-Oriented Progressive-Simplification based Vectorisation System for Engineering Drawings: Model, Algorithm and Performance", *Pattern Analysis and Machine Intelligence (PAMI)*, 24(8) : 1048-1060, 2002.
- [55] J. Song, M. Cai, M.R. Lyu, S. Cai, "Graphics Recognition from Binary Images : One Step or Two Steps", *International Conference on Pattern Recognition (ICPR)*, 2002.
- [56] J. Song, M. Cai, M.R. Lyu, S. Cai, "A New Approach for Line Recognition in Large-Size Images Using Hough Transform", *International Conference on Pattern Recognition (ICPR)*, 2002.
- [57] J. Song, M.R. Lyu, M. Cai, and S. Cai, "Graphic Object Recognition from Binary Images: a Survey and an Integrated Paradigm", *Transactions on Systems, Man and Cybernetics, part C : Applications and Reviews (TSMCC)*, under review.
- [58] Y.M. Su, J.F. Wang, "A Learning Process to the Identification of Feature Points on Chinese Characters", *International Conference on Pattern Recognition (ICPR)*, 2002.
- [59] C.L. Tan, Z. Zhang, "Text Block Segmentation Using Pyramid Structure", *Document Recognition and Retrieval*, 2001.
- [60] K. Tombre, "Structural and Syntactic Methods in Line Drawing Analysis : To Which Extent do they Work ? ", *Structural and Syntactical Pattern Recognition (SSPR)*, 1996.
- [61] K. Tombre, C. Ah-Soon, P. Dosch, G. Masini, S. Tabbone, "Stable and Robust Vectorization : How to Make the Right Choices", *Graphics Recognition (GREC)*, 1999.
- [62] O. D. Trier, A. K. Jain, T. Taxt, "Features Extraction Methods for Character Recognition – A Survey", *Pattern Recognition (PR)*, 29(4) : 641-662, 1996.
- [63] M.J. Turner, N.E. Wiseman, "Efficient Lossless Image Contour Coding", *Computer Graphics Forum*, 15(2) : 107-118, 1996.
- [64] P. Vaxivière, K. Tombre, "Subsampling : A Structural Approach to Technical Document Vectorisation", *Syntactical and Structural Pattern Recognition (SSPR)*, 1995.
- [65] A. Vinciarelli, "A Survey on Off-Line Cursive Word Recognition", *Pattern Recognition (PR)*, 35(7) : 1443-1446, 2002.
- [66] M. Weindorf, "Structure Based Interpretation of Unstructured Vector Maps", *Graphics Recognition (GREC)*, 2001.
- [67] H. Xue, "Building Skeletal Graphs for Structural Feature Extraction on Handwriting Images", *International Conference on Document Analysis And Recognition (ICDAR)*, 2001.
- [68] S. Yoon, G. Kim, Y. Choi, Y. Lee, "New Paradigm for Segmentation and Recognition", *Graphics Recognition (GREC)*, 2001.
- [69] J.J. Zou, H. Yan, "Vectorization of Cartoon Drawings", *Visual Information*, 2000.