

Analyse des documents graphiques :  
une approche par reconstruction d'objets

15 juin 2007

# Annexe C : La Librairie PSI

## Introduction

Les travaux de cette thèse ont démarré dans le cadre d'une collaboration au consortium DocMining. Ce dernier a regroupé différents partenaires de recherche et industriel : Perception Systèmes Information (PSI), Laboratoire Informatique Image et Interaction (L3i), Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA), Document Image and Voice Analysis (DIVA) et France Télécom (FT). Le but de DocMining était de proposer un environnement logiciel permettant la construction et mise en oeuvre de scénarios de traitement de document. Il a abouti à la plate forme DocMining qui a donné lieu à différentes publications<sup>1</sup> [Clavier 03], [Adam 04a] et [Adam 04b]. Dans le cadre de cette collaboration, il nous a été demandé de réaliser une librairie de traitement d'images et de reconnaissance de formes regroupant des travaux du laboratoire PSI : la librairie PSI (PSILib). Cette librairie a alors été exploitée dans le système DocMining, mais aussi dans d'autres projets par la suite : interface de conception de scénarios [Saidali 02], interface d'apprentissage [Adam 01], TP de traitement d'images<sup>2</sup>, ... Via la conception de la PSILib, nous avons rejoint certaines initiatives existantes en matière d'intégration logicielle de travaux de recherche : Qgar<sup>3</sup>, Library of Efficient Data Types and Algorithms (LEDA)<sup>4</sup>, Intel Computer Vision Library (ICVL)<sup>5</sup>, ...

La PSILib se compose d'environ 80 000 lignes de code et est exploitable sur les systèmes Windows et Linux. Elle regroupe différents opérateurs Java et natifs (C et C++) répartis au sein de différents projets informatiques. Les opérateurs natifs sont intégrés au sein de la PSILib via les JNI<sup>6</sup>. La (quasi) totalité de ces opérateurs génèrent leurs résultats au format XML et dérivés (SVG, XGMML, ...) [Michard 00]. La documentation de la PSILib est générée à l'aide de l'outil Javadoc, elle décrit chacun des opérateurs avec des exemples de résultats. La figure suivante donne une capture d'écran de la page d'accueil de la documentation de la PSILib.

---

<sup>1</sup>Nous citons ici celles qui nous semblent les plus représentatives.

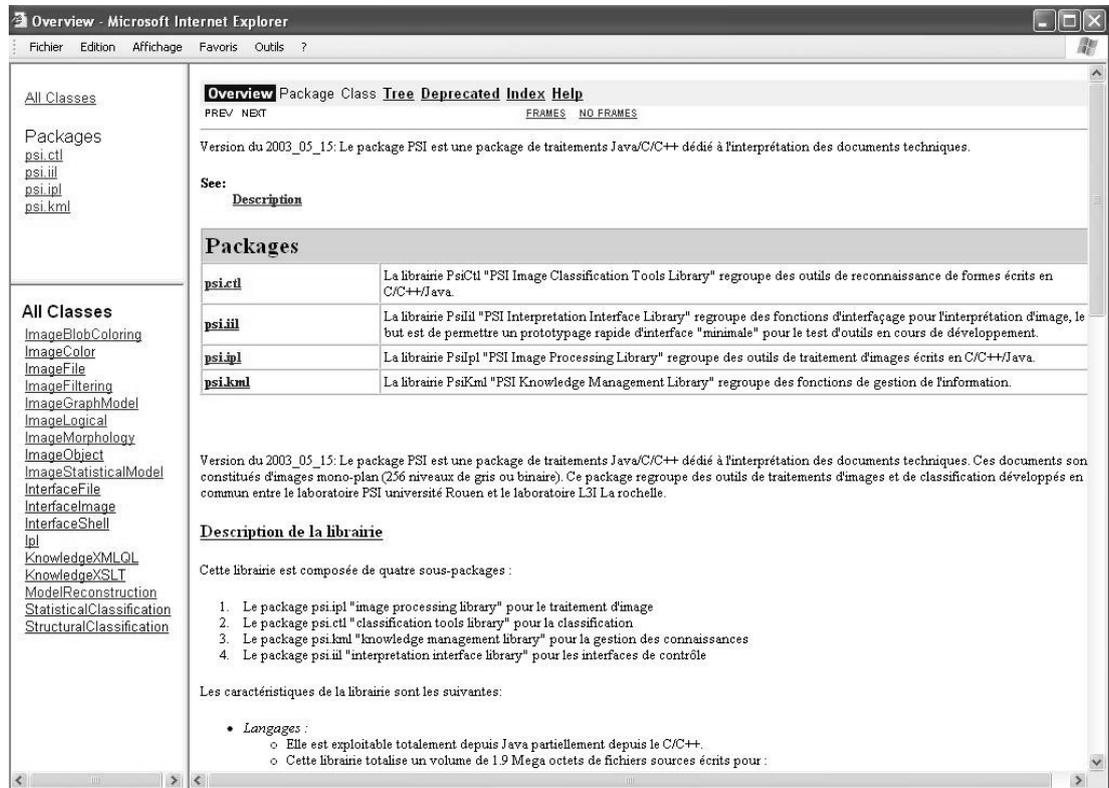
<sup>2</sup><http://mathieu.delalandre.free.fr/teachings/tiapp.doc>

<sup>3</sup><http://www.qgar.org/>

<sup>4</sup><http://www.cs.sunysb.edu/>

<sup>5</sup><http://www.intel.com/>

<sup>6</sup>Java Native Interface [Descartes 00]



Page d'accueil de la documentation de la PSILib

La PSILib est composée de deux sous-librairies principales : la librairie de traitement d'images et celle de classification. Nous présentons chacune d'entre elles dans la suite de cette Annexe.

## Librairie de traitement d'images

La librairie de traitement d'images est composée de huit familles d'opérateurs :

1. Opérateurs de filtrage
2. Opérateurs de binarisation
3. Opérateurs logiques
4. Opérateurs morphologiques

5. Opérateurs de marquage
6. Opérateurs statistiques
7. Opérateurs structurels
8. Opérateurs d'approximation mathématique

Les quatre premières familles d'opérateurs utilisent des méthodes usuelles de traitement d'images. Les opérateurs de filtrage (1) sont basés sur l'utilisation de la JAI<sup>7</sup>. Ils mettent en oeuvre des filtres moyenneur, médian et gradient à partir d'images en niveaux de gris [Parker 96]<sup>8</sup>. Les opérateurs de binarisation (2) sont basés, eux, sur les travaux de [Lassaulzais 97]<sup>8</sup>. Ils emploient deux méthodes de la littérature : [Otsu 79] et [Kittler 86]. Ceux logiques (3) sont également exploités de la JAI<sup>7</sup>. Ils mettent en oeuvre des opérations booléennes comme le *non*, *ou*, *et*, *ou exclusif*, ... Enfin, les opérateurs morphologiques (4) utilisent des opérations de morphologie mathématique<sup>9</sup> telles que l'érosion, la dilatation, la fermeture et l'ouverture. Nous utilisons également en complément une méthode de détection morphologique des contours par test de voisinage [Rosenfeld 82]<sup>10</sup>.

Les opérateurs de marquage (5) sont basés sur nos travaux présentés dans de ce manuscrit<sup>11</sup>. Il permettent différentes applications comme le marquage de composantes connexes, l'extraction d'occlusions, le filtrage de composantes et l'évaluation du bruit impulsionnel. Nous exploitons en complément quatre opérateurs statistiques (6) pour l'extraction des attributs géométriques, moments de Zernike, invariants de Fourier-Mellin et sondes circulaires. Ils permettent d'extraire des caractéristiques décrivant les composantes connexes de façon invariante aux changements d'orientation et parfois d'échelle. Nous extrayons les attributs géométriques (surface, périmètre, compacité) à l'aide de nos opérateurs combinés de marquage de composantes connexes et de détection de contours<sup>12</sup>. Les extractions des moments de Zernike, invariants de Fourier-Mellin et sondes circulaires sont basées sur les travaux de [Adam 01]<sup>8</sup>. Elles reposent sur des principes différents :

**Moments de Zernike** Calcul de polynômes orthogonaux constituant un espace dans lequel on projette la forme.

**Sondes Circulaires** Calcul des valeurs angulaires obtenues à partir d'un ensemble de cercles concentriques.

**Invariants de Fourier Mellin** Développement de la transformée de Fourier-Mellin d'une fonction en coordonnées polaires correspondant à la forme.

---

<sup>7</sup>Java Advanced Imaging <http://www.sun.com/>

<sup>8</sup>Nous reportons le lecteur aux références citées pour plus de précisions.

<sup>9</sup>Nous reportons le lecteur page ?? sur ces méthodes

<sup>10</sup>Nous reportons le lecteur page ?? pour une présentation de cette méthode.

<sup>11</sup>Nous reportons le lecteur aux chapitres ?? et ??.

<sup>12</sup>Nous reportons le lecteur page ??.

Nous utilisons également une famille d'opérateurs structurels (7). Ceux-ci permettent d'extraire des primitives graphiques structurées via des graphes. Ils sont basés sur les différentes approches région, contour et squelette. Les opérateurs région reposent sur nos travaux<sup>13</sup> présentés dans ce manuscrit. Ils permettent l'extraction de graphe de voisinage, d'inclusion et sous contraintes de distance. Celui contour est basé sur notre méthode de détection de contours également présentée dans ce manuscrit<sup>13</sup>. Enfin, nous utilisons deux niveaux d'opérateurs pour notre approche squelette. Tout d'abord nous exploitons deux opérateurs de squelettisation basés sur des algorithmes de la littérature : [Taconet 90] et [Baja 94]. Ces derniers sont exploités des implémentations respectives<sup>14</sup> de [Lassaulzais 97] et [Adam 98]. Nous utilisons en complément notre opérateur<sup>13</sup> de construction du graphe de squelette afin de structurer les images squelettisées.

Nous utilisons enfin différents opérateurs d'approximation mathématique à partir des résultats des opérateurs squelette et contour. Des opérateurs sont tout d'abord utilisés pour l'approximation des vecteurs, arcs et courbes des listes de points. Ces derniers sont exploités des travaux de [Hameau 02] et [Perin 04]. [Perin 04] utilise des algorithmes de la littérature pour la polygonalisation : la corde [Ramer 72] suivie d'un post-traitement split & merge [Pavlidis 74]. [Hameau 02] utilise la variance des angles entre vecteurs deux à deux pour la détection des arcs de cercle. Finalement, [Perin 04] utilise la méthode de [Liao 90] pour l'approximation de courbes par des polynômes de Bernstein-Bezier. Nous utilisons un dernier opérateur d'approximation pour l'appariement de contours polygonalisés. Celui-ci est basé sur les travaux de [Hamidou 04]. Il utilise tout d'abord des critères d'appariement entre vecteurs inspirés des travaux de [Ramel 00]. Il emploie alors un algorithme de décision floue pour l'appariement des vecteurs en quadrilatères.

## Librairie de classification

Une seconde librairie est employée dans la PSILib dite de classification. Celle-ci est composée de deux familles d'opérateurs pour la classification statistique et structurelle.

Nous utilisons une première famille d'opérateurs pour la classification statistique. Ceux-ci sont basés sur l'utilisation d'un algorithme de la littérature : celui des k plus proches voisins [Fix 51]. Ce dernier est implémenté des travaux de [Adam 01]<sup>14</sup>. Il permet l'utilisation de différentes distances (Hamming, Euclidienne, ...) et la normalisation des primitives extraites. Nous l'utilisons au sein de deux opérateurs : un de classification et un de recherche de similarité. Le premier calcule les distances entre le vecteur candidat et tout ceux de la base d'apprentissage dans le but d'obtenir un résultat de reconnaissance. Le second les calcule entre le vecteur candidat et tout ceux extraits afin de rechercher des formes similaires.

---

<sup>13</sup>Nous reportons le au chapitre ??.

<sup>14</sup>Nous reportons le lecteur aux références citées pour plus de précisions.

Nous utilisons une deuxième famille d'opérateurs pour la classification structurelle. Ces derniers sont exploités<sup>15</sup> des travaux de [Héroux 00] et [Nicolas 02]. Ils permettent à la fois l'appariement de graphes et de sous-graphes.

Concernant l'appariement de graphes nous utilisons un premier opérateur basé sur les travaux de [Héroux 00]. Celui-ci calcule des taux de similarité entre graphes basés sur le recouvrement entre un graphe candidat ( $G_1$ ) et un graphe modèle ( $G_2$ ) (équation (1)). Ce recouvrement correspond à leur plus grand sous-graphe commun ( $G_c$ ) (équation (2)). Deux taux de similarité peuvent être calculés selon le nombre d'éléments communs sur les noeuds ( $\delta_n$ ) et sur les arcs ( $\delta_a$ ) (équation (3)). Le taux global est obtenu par calcul de la variance entre ( $\delta_n$ ) et ( $\delta_a$ ). Le résultat de la classification correspond au label du graphe modèle de taux de similarité minimum.

$$G_1 = (\{n_0, \dots, n_{n1}\}, \{a_0, \dots, a_{a1}\}) \quad G_2 = (\{n_0, \dots, n_{n2}\}, \{a_0, \dots, a_{a2}\}) \quad (1)$$

$$G_c = (\{n_0, \dots, n_{nc}\}, \{a_0, \dots, a_{ac}\}) \quad (2)$$

$$\delta_n(G_1, G_2) = \frac{n1 \times n2}{nc^2} - 1 \quad \delta_a(G_1, G_2) = \frac{a1 \times a2}{ac^2} - 1 \quad (3)$$

Notre opérateur d'appariement de sous-graphes est basé lui sur les travaux de [Nicolas 02]. Ils emploient différentes méthodes de la littérature : algorithme  $A^*$  [Nilsson 80], Ullman [Ullman 76], VF [Cordella 01], ... Cet opérateur emploie également une méthode de factorisation des sous-graphes détectés. Le but est alors de factoriser itérativement les sous-graphes reconnus d'un graphe plus large, afin de produire un graphe des résultats de reconnaissance.

---

<sup>15</sup>Nous reportons le lecteur aux références citées pour plus de précisions.

# Bibliographie

- [Adam 98] S. Adam. Interprétation de documents du réseau téléphonique, approche multi-spécialistes. Rapport de Master, Laboratoire PSI, Université de Rouen, France, 1998.
- [Adam 01] S. Adam. *Interprétation de Documents Techniques, des Outils à leur Intégration dans un Système à Base de Connaissances*. Thèse de Doctorat, Université de Rouen, France, 2001.
- [Adam 04a] E. Clavier S. Adam, P. Héroux, M. Rigamonti & J.M. Ogier. *Docmining : Une Plate-Forme de Conception de Systèmes d'Analyse de Documents*. In Colloque International Francophone sur l'Écrit et le Document (CIFED), pages 97–102, 2004.
- [Adam 04b] S. Adam & al. *Docmining : Une plate-forme de conception de systèmes d'analyse de documents*. In Colloque International sur l'Écrit et le Document (CIFED), pages 97–102, 2004.
- [Baja 94] G.S. Di Baja. *Well-Shaped, Stable, and Reversible Skeletons from the (3,4)-Distance Transform*. Journal of Visual Communication and Image Representation, vol. 5, no. 1, pages 107–115, 1994.
- [Clavier 03] E. Clavier, P. Heroux, J. Gardes & E. Trupin. *Ground-Truth Production and Benchmarking Scenarios Creation With DocMining*. In Workshop on Document Layout Interpretation and its Applications (DLIA), pages 31–35, 2003.
- [Cordella 01] L.P. Cordella, P. Foggia, C. Sansone & M. Vento. *An Improved Algorithm for Matching Large Graphs*. In Workshop on Graph-based Representation (GBR), pages 149–159, 2001.
- [Descartes 00] A. Descartes. Java native methods. Editions O'Reilly, ISBN : 1565923456, 2000.
- [Fix 51] E. Fix & J. Hodges. *Discriminatory Analysis : Nonparametric Discrimination : Consistency Properties*. Rapport technique 11, USAF School of Aviation Medicine, Randolph Field, USA, 1951.
- [Hameau 02] A. Hameau & G. Renault. Réalisation d'un outil de recherche d'arcs de cercle. Rapport de Master, Laboratoire PSI, Université de Rouen, France, 2002.
- [Hamidou 04] O. Hamidou. Contour matching. Rapport de Master, Laboratoire PSI, Université de Rouen, France, 2004.
- [Héroux 00] P. Héroux, S Diana, E. Trupin & Y. Lecourtier. *A Structural Classification for Retrospective Conversion of Documents*. In Conference on Structural and Syntactical Pattern Recognition (SSPR), volume 1876 of *Lecture Notes in Computer Science (LNCS)*, pages 154–162, 2000.

- [Kittler 86] J. Kittler & J. Illingworth. *Minimum Error Thresholding*. Pattern Recognition (PR), vol. 19, no. 1, pages 41–47, 1986.
- [Lassaulzais 97] A. Lassaulzais. Contribution au développement d’un système de vectorisation et d’interprétation de plans de réseaux. Rapport de Master, Laboratoire PSI, Université de Rouen, France, 1997.
- [Liao 90] C. W Liao & J. S. Huang. *Stroke Segmentation by Bernstein-Bezier Curve Fitting*. Pattern Recognition (PR), vol. 23, no. 5, pages 475–484, 1990.
- [Michard 00] A. Michard. Xml langage et application. Editions Eyrolles, 2 edition, ISBN : 2-212-09206-7, 2000.
- [Nicolas 02] S. Nicolas. Vers une bibliothèque pour la manipulation et la factorisation de sous-graphes. Rapport de Master, Laboratoire PSI, Université de Rouen, France, 2002.
- [Nilsson 80] J. Nilsson. Principles of artificial intelligence. Morgan Kaufmann Editor, ISBN : 0-935382-01-1, 1980.
- [Otsu 79] N. Otsu. *A Threshold Selection Method from Gray-Level Histograms*. Transactions on Systems, Man and Cybernetics (TSMC), vol. 9, no. 1, pages 62–66, 1979.
- [Parker 96] J.R. Parker. Algorithms for image processing and computer vision. Paperback Editions, ISBN : 0471140562, 1996.
- [Pavlidis 74] T. Pavlidis & S. L. Horowitz. *Segmentation of Plane Curves*. Transactions on Computers (TC), vol. 23, pages 860–870, 1974.
- [Perin 04] S. Perin. Formalisation et opérationnalisation de connaissances graphiques pour l’interopérabilité en analyse d’image de document. Rapport de Master, Laboratoire PSI, Université de Rouen, France, 2004.
- [Ramel 00] J.Y. Ramel, N. Vincent & H. Emptoz. *A Structural Representation for Understanding Line-Drawing Images*. International Journal on Document Analysis and Recognition (IJ DAR), vol. 3, pages 58–66, 2000.
- [Ramer 72] V. Ramer. *An Iterative Procedure for the Polygonal Approximation of Plane Curves*. Computer Vision Graphics and Image Processing (CVGIP), vol. 1, pages 244–246, 1972.
- [Rosenfeld 82] A. Rosenfeld & A.C. Kak. Digital picture processing. Academic Press, 2 edition, ISBN : 0125973020, 1982.
- [Saidali 02] Y. Saidali. *Modélisation et Acquisition Incrémentale de Connaissances Traiteurs d’Images*. Thèse de Doctorat, Université de Rouen, France, 2002.
- [Taconet 90] B. Taconet, A. Zahour, S. Zhang & A. Faure. *Deux Algorithmes de Squelettisation*. In Colloque sur la Reconnaissance Automatique de l’Ecriture (RAE), volume 68 of *BIGRE*, pages 68–78, 1990.
- [Ullman 76] J.R. Ullman. *An Algorithm for Subgraph Isomorphism*. Journal Association for Computer Machinery, vol. 23, pages 31–42, 1976.

# Table des figures

Page d'accueil de la documentation de la PSILib . . . . . 1

# Table des matières

<b>Annexe C : La Librairie PSI</b>	<b>0</b>
Introduction . . . . .	0
Librairie de traitement d'images . . . . .	1
Librairie de classification . . . . .	3
<b>Bibliographie</b>	<b>5</b>
<b>Table des figures</b>	<b>7</b>