

Rapport d'enquête

État des lieux de la production informatique de l'équipe IMéDoc.

Antoine MERCIER

Laboratoire d'Informatique, Images et Interactions
Université de La Rochelle

Responsables :

Rémy Mullet

Karell Bertet

Muriel Visani

26 janvier 2009

Table des matières

Introduction	2
Les Projets scientifiques	3
Marvel	4
Alpage	5
TOpViSION	6
MaDoNne & NaviDoMass	7
Les Projets industriels	9
SAG - Ridoret	10
RecoNomad	11
Activités scientifiques	12
Vision à travers les milieux opaques	13
Minimum Spanning Tree	14
Textures Dynamiques	15
Treillis	16

Introduction

Une des priorités principales du laboratoire L3i est de capitaliser les ressources produites par ses membres au cours de leur recherche et développement. Le choix a été fait de commencer par l'équipe IMéDoc car elle est l'une des deux plus anciennes équipes ; l'autre étant gérée par une autre ingénieure. Cet état des lieux sera divisé selon les projets rattachés à l'équipe. Pour chacun, un bref descriptif du projet sera fourni avant de rendre compte de ce qui existe, ce qui est capitalisable et enfin, les perspectives à plus ou moins long terme. Enfin, nous essayerons de conclure sur les grandes tendances qui se dégagent de cette enquête.

Les Projets scientifiques

Le laboratoire est tourné vers la constitution de consortium dans une dynamique de projets. Aujourd'hui, la politique scientifique est résolument tournée vers les projets labellisés (ANR, PCRD, Eureka, ...) qui constituent le coeur de l'activité actuelle des projets scientifiques.

Marvel

“Multimodal Analysis of Recorded Video for E-Learning”

Le responsable de ce projet Jean-Marc Ogier.

Description du projet

Ce projet a pour but l'extraction et le raisonnement sémantique sur plusieurs médias simultanément. Lorsqu'on tente d'analyser une seule modalité, image, vidéo, parole ou texte, à la fois, ce problème est difficile. La recherche dans l'extraction de sémantique en étudiant qu'une seule modalité semble atteindre aujourd'hui ses limites. C'est pourquoi une nouvelle voie de recherche intéressante commence à voir le jour pour combler le fossé entre ces sources (ou modalités) et leurs sémantiques associées. Il s'agit de coupler l'analyse de plusieurs modalités (visuelle, auditive, textuelle, web ou autres) afin d'obtenir une complémentarité d'informations entre toutes ces sources et pour déclencher une synergie dans l'analyse des différents indices sémantiques inclus dans chaque modalité.

État des lieux

La thèse de Thomas Martin arrivant à son terme, les développements logiciels en rapport avec ce projet n'évolueront plus beaucoup à court terme. Pour ses travaux, Thomas a utilisé diverses bibliothèques comme LtiLib, libSVM, boost, Qt, OpenCV, etc ... Dans ses travaux, il a créé un OCR avec analyse lexicale (modèles n-grammes) utilisant des graphes pour représenter les graphèmes, un système de suivi de personne dans une vidéo. La majeure partie des développements a été effectuée en C++ et quelques petites parties en Shell ou Perl.

Conclusion

Une démonstration peut être réalisée en utilisant les programmes en ligne de commande codés par Thomas. Ceux-ci fournissent principalement des graphes sous la forme d'une image bitmap représentant les propositions de l'OCR pour un mot donné. Par ailleurs, il a aussi pris le soin de créer une passerelle entre les bibliothèques “LtiLib” et “OpenCV”. Globalement, le code devrait être réutilisable car Thomas a pris le soin de respecter des méthodes de conception logicielle utilisées dans la bibliothèque “LtiLib”.

Des vidéos sont disponibles pour montrer l'efficacité du suivi de personne. Un léger traitement sera nécessaire sur celles-ci afin de les rendre facilement lisibles mais ceci ne devrait pas prendre plus de quelques heures.

Nous pouvons donc dire que ce projet donne lieu à des éléments capitalisables rapidement, tant au niveau logiciel et code source qu'au niveau démonstration vidéo.

Alpage

“Analyse diachronique de l’espace urbain Parisien : approche Géomatique”

Le responsable de ce projet est Jean-Marc Ogier.

Description du projet

Ce projet se positionne dans le cadre d’une valorisation du patrimoine Français, plus précisément, une mise en exergue des plans cadastraux de la ville de Paris issus de l’atlas VASSEROT du 19ème siècle. Afin de rendre exploitable l’information spatiale contenue dans les documents cadastraux, il est nécessaire d’individualiser informatiquement des objets cohérents en l’occurrence les rues, les îlots, les parcelles pour ensuite insérer ces éléments dans un système d’information géographique (SIG). Ces étapes d’extraction de connaissance s’appuient sur les avancées de domaines scientifiques tel que la reconnaissance de formes, la vectorisation d’objets linéaires, le traitement d’images couleur : les espaces de représentation, l’analyse d’histogrammes couleur 3D, les invariants colorimétriques et les signatures dynamiques pour le suivi spatial et/ ou temporel d’objets.

État des lieux

À l’heure actuelle, le principal contributeur à ce projet est Romain Raveaux dans le cadre de sa thèse de doctorat. Dans le cadre de son travail, il a réalisé divers petits programmes avec interface graphique. Il s’agit d’outils scientifiques, ceux-ci ne sont pas exempts de bugs. Par ailleurs, la manipulation de ceux-ci n’est pas forcément évidente au premier abord.

Conclusion

Ses programmes constituent une base sérieuse de code potentiellement réutilisable en Java et R. Pour l’élaboration de ceux-ci, plusieurs bibliothèques ont été utilisées. Par exemple, pour pouvoir utiliser du code R dans ses programmes en Java, il a du utiliser une bibliothèque permettant de faire des appels à un serveur R afin d’effectuer certains type de calculs. Toutefois, une phase d’ingénierie serait nécessaire pour étudier quelles sont les bibliothèques absolument nécessaires parmi celles qu’il utilise. En effet, dans la longue liste apparaissent bon nombre de bibliothèque de traitement de analyse d’image. Il faudrait donc étudier si toutes sont absolument nécessaire ou si leur utilisation ne peut pas être compensée par une autre déjà employée.

TOpViSION

“Tests OPérationnels de Vidéos Sous-marines pour l’Identification d’Objets Nuisibles”
Le responsable de ce projet est Jean-Christophe Burie.

Description du projet

Dans le cadre de la réponse à l’appel à propositions TECHNOVISION initié par le Ministère de la recherche et le Ministère de la défense, Thales Underwater Systems propose de mener une campagne d’évaluation des algorithmes de traitement de séquences vidéos pour des fonctions de détection, localisation, caractérisation et identification automatiques d’objets sous-marins nuisibles (polluants, dangereux, voire létaux) par caméra mobile. Le L3i a développé des algorithmes pour les fonction de détection, localisation et la caractérisation.

État des lieux

Ce projet est terminé et donne lieu à un ensemble de bouts de code écrits par des étudiants en MatLab. Tous ces bouts de code constituent une chaîne de traitement complète.

Conclusion

Un démonstrateur exécutable n’existe pas mais devrait être réalisable étant donné que toutes les parties de la chaîne de traitement sont en notre possession. Ceux-ci sont étudiés pour traiter des images sous-marines et, par conséquent, sont un peu particulier. Cependant, ils devraient être réutilisable sous réserve de problème de propriété industrielle.

MaDoNne & NaviDoMass

“MASse de DONnées issues de la Numérisation du patrimoiNE”

“NAVigation In DOcument MASSes”

Le responsable de ces projets est Jean-Marc Ogier.

Description des projets

MaDoNne

Ce projet a pour mission de mettre en valeur différents biens du patrimoine international et plus particulièrement les ouvrages, les collections d’images et autres documents iconographiques. A court terme, ces nombreux documents constitueront une source gigantesque d’information (masse de données). L’objectif de ce projet est de contribuer à la réalisation de systèmes d’indexation d’images de document du patrimoine. Ce projet s’inscrit ainsi dans la volonté actuelle de préserver le patrimoine culturel et scientifique et d’assurer au plus grand nombre l’accès à ce patrimoine.

NaviDoMass

Faisant suite au projet Madonne, le but général de ce projet est de construire un framework de valorisation du patrimoine. Ce système vise à préserver et à fournir un accès aux fonds documentaires des bibliothèques, des musées et des archives publiques. NaviDoMass est fondé sur quatre grands principes : « N’importe où » (un accès global), « N’importe qui » (publique et multi-langues), « N’importe quand » et « N’importe quel média » (différents modes de transmission de l’information, portail Web, Interface Smartphone, réimpressions de documents. . .).

L’ensemble du procédé tente de solutionner le problème de la navigation dans des masses d’images de documents, contenant du texte, des lettrines, des illustrations et des schémas. Tout le défi consiste alors à exécuter une retro-conversion de documents en donnant une interprétation sémantique à des amas de pixels inertes.

État des lieux

Sur ces deux projets qui se sont succédés, quatre doctorants (dont trois devenus docteurs depuis) ont apporté d’importantes contributions.

Stéphanie Guillas a écrit toute une chaîne d’apprentissage et reconnaissance de symboles bruités utilisant les treillis de gallois ainsi que d’autres types de classifieurs pour comparaison. Cette chaîne de traitement a été réécrite par la suite dans le cadre d’un projet de recherche et le tout a été intégré au sein d’un seul programme appelé “NaviGaLa”. En plus de ce travail, S Guillas a développé quelques classifieurs et autres outils dans le but de faire des comparatifs pour étayer sa thèse.

La première version de cette chaîne de traitement a été écrite en partie en C et en partie en Java. Cette version n’était pas intuitive d’utilisation et nécessitait d’exécuter plusieurs programmes et de manipuler des fichiers texte pour faire la liaison entre eux. La deuxième version, NaviGaLa, permet une utilisation de l’ensemble de la chaîne au sein d’un unique programme entièrement écrit en Java et qui utilise quelques bibliothèques

libres optimisées pour les calculs. Cette version permet des export en xml et dot dans informations calculées et une seconde version a été commencé afin de gérer l'interface multi-langue cependant, elle n'est pas fonctionnelle en l'état. Après concertation avec Karell Bertet, une réécriture de ce programme (et éventuellement découpage en plusieurs programmes) est à envisager dans le cadre de projets étudiants ou de stages de Master.

Nicholas Journet ...

Surapong Utama ...

Enfin, Mickaël Coustaty a travaillé sur la mise au point d'une signature basée sur la structure des symboles à reconnaître (et non des calculs statistiques) en utilisant des treillis de gallois. Depuis un an maintenant, il travaille aussi sur le projet NaviDo-Mass dans l'indexation d'images de documents anciens par le contenu en prenant comme contexte les lettrines en un premier temps.

Ces développements sont en Java pour la signature structurale. Celle-ci est sous la forme d'un exécutable Java en ligne de commande qui serait intégrable dans NaviGaLa sans trop de difficultés. Son travail sur les lettrines est actuellement en MatLab pour la partie "tests scientifiques" et en C# pour l'interface graphique finale. Cette dernière est codé en utilisant la bibliothèque "SharpVision" développée conjointement avec des membres de l'équipe Imagin.

Conclusion

Avec ces deux projets très liés de par leur sujet d'étude, nous avons beaucoup d'éléments a priori capitalisable rapidement. Cependant, de par les nuances entre chacune des contributions, il va falloir étudier sérieusement les problèmes d'interopérabilité. En effet, certains des développements ont été effectué en C/C++, d'autres en Java, MatLab ou encore C#. Tous ces langages ne sont pas réputés pour être facilement "mélangeables". Cette diversité risque donc de ralentir la capitalisation et va mettre en évidence la nécessité de fixer des limites quand aux langages préconisés et les bibliothèques indiquées.

Les Projets industriels

Les projets industriels constituent une activité de recherche, développement et même de transfert en juste équilibre avec les recherches plus fondamentales.

SAG - Ridoret

Les responsables sont Jean-Marc Ogier & Jean-Christophe Burie.

Description de la contribution

Le projet s'inscrit dans le cadre du contrat de collaboration de recherche établi entre la société S.A.G. Groupe Ridoret et le laboratoire L3i « Informatique Image Interaction » de l'université de La Rochelle.

Le but est de créer une application basée sur de l'analyse et de la synthèse d'images dans le but de permettre à des clients de simuler le changement de leurs menuiseries à partir de photographies numériques. Les problématiques scientifiques abordées concernent donc l'estimation des paramètres relatifs aux menuiseries (taille, surface, ...) à changer en fonction d'informations liées à l'angle de vue et la distance de la photographie, et également la mise en œuvre d'une synthèse d'image réaliste du changement de menuiserie suivant le choix du client (couleur, texture, ...) , permettant à ce dernier de visualiser de façon autonome le rendu de ce changement. Cet outil sera mis en ligne sur un site internet dédié, permettant ainsi de simuler l'aspect d'une façade ou d'une vue intérieure d'une habitation après travaux de rénovation des menuiseries extérieures.

État des lieux

Lors de la phase de faisabilité, un étudiant de M2 a réalisé pendant son stage un démonstrateur. Cette application a été conçue en langage C avec les bibliothèques de traitement d'images OpenCV.

Conclusion

En terme de démonstration, une version exécutable est disponible au sein du laboratoire cependant, du point de vue du code source, celui-ci est lié au Groupe Ridoret de par la notion de propriété industrielle résultante du partenariat.

RecoNomad

Le responsable de ce projet est Jean-Marc Ogier.

Description du projet

Le projet vise à développer un ensemble de bibliothèques qui seront vendues par le partenaire industriel avec une tablette de saisie électronique. Ces bibliothèques devront être capables de pouvoir déterminer quel type de formulaire a été rempli sur la dite tablette, de segmenter les informations saisies et reconnaître les caractères écrits.

La première année du projet devait traiter des caractères pré-casés et est arrivé à son terme avec une reconnaissance des lettres majuscules, minuscules et des chiffres.

La deuxième année, débutant en janvier avec l'arrivée de la nouvelle équipe, devra traiter des saisies de caractères liés ainsi que de l'amélioration de la chaîne globale.

État des lieux

Ce projet apporte donc au L3i des outils de reconnaissances de structure de documents, des outils de segmentation de documents ainsi qu'un puissant ICR basé sur un mélange de caractéristiques "en-ligne" et "hors-ligne". Dans le cadre de ce projet, des bases d'apprentissage de caractères manuscrit ont été acquises par le laboratoire et un démonstrateur technique a été développé afin d'aider au debug et à la démonstration des compétences du laboratoire. Une refonte de ce démonstrateur est envisageable afin d'homogénéiser sa structure. D'autres petits outils avec interface graphique ont été mis au point afin d'aider à l'apprentissage et au debug des bibliothèques. Les bibliothèques ont été développées en C++ en utilisant quelques bibliothèques du monde libre comme OpenCV, SQLite3, boost, LipiToolKit. Les interfaces graphiques ont été développées en Java ou en C/C++. La liaison entre les interfaces Java et les bibliothèques C++ ont été fait au travers de petits exécutable en ligne de commande mais des JMI sont envisageable. Par ailleurs, pour les besoins de l'industriel, une encapsulation C# a été faite par dessus la structure C++ donc une réécriture des interfaces graphiques en C# pourrait être envisageable aussi. Cependant, comme beaucoup d'affichage se fait au travers d'une bibliothèque SVG open source en Java, cette conversion risque d'être difficile.

Conclusion

Ce projet apporte beaucoup d'éléments capitalisable dès maintenant. Des démonstrations ont déjà été faite à plusieurs reprises. Il faudrait principalement harmoniser l'ensemble des outils graphiques, s'assurer de l'absence de bugs et enfin, faire des tutoriels pour que tout un chacun puisse savoir comment s'en servir. Il n'y a pas de démonstration vidéo disponible mais des versions exécutable le sont à l'heure actuelle.

Activités scientifiques

En dehors des projets scientifiques, les membres travaillent sur des sujets divers seuls, avec des doctorants ou avec des étudiants en projet de master.

Dans cette partie nous détaillerons donc ces quelques contributions.

Vision à travers les milieux opaques

Les responsables de ces activités sont Alain Gaugue et Jamal Khamlichi.

Description du projet

Ce projet vise à développer un système composé d'une partie matérielle et d'une partie logicielle fortement liées permettant d'obtenir une vision au travers de divers milieux opaques. Ces milieux opaques peuvent être aussi bien des cloisons en plaques de plâtre standard comme des murs en béton de plusieurs dizaines de centimètres d'épaisseur.

Alors que la partie matérielle est assimilable à un radar, la partie logicielle, qui nous intéresse, est à l'heure actuelle un ensemble de petits programmes écrits en C++ ou MatLab. Afin d'exécuter la chaîne de traitement, ceux-ci doivent être appelés les uns après les autres.

État des lieux

Étant donné la nature de ce projet, il est compréhensible que la majeure partie du code source qui a été écrit est spécifique et dédié à l'utilisation conjointe avec le matériel fabriqué par les membres de l'équipe. Une partie du code reste néanmoins une contribution de Michel Ménard et constitue une petite bibliothèque de traitements d'images adapté au contexte. Cette partie de code écrite en MatLab est, a priori, réutilisable avec l'accord de M Ménard. Le reste du code est écrit principalement en C++ et utilise, pour le moment, les bibliothèques "MFC" de Microsoft pour la gestion de l'affichage des fenêtres. À côté de cela, diverses bibliothèques sont utilisées et les principales sont "Qt" de TrollTech et "OpenGL". Il est prévu de se détacher de la bibliothèque Microsoft en utilisant à la place le système de fenêtrage de la bibliothèque "Qt". Par ailleurs, Jamal Khamlichi essaye le plus possible de transcoder le reste du code MatLab en code C++ afin de l'intégrer au code existant et accélérer les calculs. À terme, ils espèrent avoir un code entièrement écrit en C++ et n'utilisant que des bibliothèques accessibles librement et multiplateformes.

Conclusion

À l'heure actuelle, il n'existe pas de démonstrateur logiciel exécutable ni de vidéo de démonstration. D'après les membres de l'équipe, un démonstrateur logiciel peut être réalisé relativement rapidement (estimation : 2 semaines à temps plein). Pour ce qui est de la démonstration vidéo, celle-ci est réalisable très rapidement aussi et la condition principale pour sa réalisation étant la caméra et la personne filmant la démonstration et réalisant le montage.

Minimum Spanning Tree

Le responsable est Patrick Franco.

Description de la contribution

Il s'agit du fruit d'un travail de recherche effectué par P Franco sur la notion des Arbres de Recouvrement Minimaux.

État des lieux

Cette contribution est entièrement réalisé par P Franco en MatLab. Une interface graphique est disponible dans cet environnement.

Un problème de confidentialité apparaît car, pour réaliser ce travail, P Franco a du signer un accord de non divulgation auprès de l'ENS Lyon pour un module de manipulation de graphe écrit en MatLab. Étant donné que le dit module est un "simple" fichier ".m", le code source est directement accessible. D'après les explications de P Franco, si un module équivalent était réalisé au sein du laboratoire (projet de master, stage, etc ...), nous pourrions nous détacher de ce problème de non divulgation.

Conclusion

Une réflexion est à envisager par rapport à ce module et voir avec les contributions disponibles dans le laboratoire si il n'y a pas moyen de régler ce souci de non divulgation. En terme de démonstration, une version "exécutable" est disponible auprès de P Franco sous la forme de quelques fichiers MatLab mais avec comme réserve la clause pré-citée. Toutefois, il est envisageable de réaliser une version compilée de ce programme. Pour cela, il faudra étudier la licence de MatLab acquise par le laboratoire pour voir s'il faut investir ou pas pour pouvoir compiler une programme écrit en MatLab.

Textures Dynamiques

Le responsable est Michel Ménard.

Description de la contribution

Il s'agit du fruit d'un travail de recherche effectué par Michel Ménard et repris dans le cadre des thèses de Sloven Dubois et Mathieu Lugiez sous son encadrement ainsi que celui de Renaud Péteri. Ces travaux sont focalisés sur les textures dynamiques. D'un côté, M Lugiez met au point des méthodes d'extraction de celles-ci dans des séquences d'images tandis que de l'autre, S Dubois met au point des méthodes de caractérisation de façon à pouvoir par la suite les détecter plus facilement lors d'indexation de vidéos.

État des lieux

Mathieu Lugiez a, à l'heure actuelle, des vidéos de démonstrations assez probantes sur les résultats de ses recherches générées par un programme écrit en MatLab de sa conception. Celui-ci est d'ors et déjà capitalisable avec certaines petites restrictions concernant quelques fonctions qu'il faudrait réécrire afin de se détacher de problème de propriété intellectuelle. Ce travail de réécriture ne devrait pas être trop difficile.

Sloven Dubois a des programmes écrit en MatLab pouvant être utilisés à des fins de démonstration. Ceux-ci exploitent quelques bibliothèques de tierce partie dont certaines sont écrites en C.

Pour tous les deux, le code est lisible et relativement documenté ce qui le rend réutilisable assez facilement.

Conclusion

Cette équipe dirigée par M Ménard a des éléments capitalisables à très court terme aussi bien en terme d'exécutable qu'en terme de vidéos. La demande a été faite afin de récupérer quelques unes de ces vidéos rapidement dans le but de les mettre à disposition sur le site du laboratoire. Les éventuels problèmes de propriété intellectuelle devraient être résolus facilement.

Treillis

Le responsable est Karell Bertet.

Description de la contribution

Il s'agit du travail effectué dans le cadre de la thèse de Stéphanie Guillas encadrée par Karell Bertet.

État des lieux

Karell Bertet travaille sur les treillis depuis plusieurs années et écrit plusieurs algorithmes sur cette problématique. Récemment, elle a aussi encadré Stéphanie Guillas dans le cadre de sa thèse portant sur la classification de symboles bruités en utilisant des treillis. De ce fait, un socle important de code source existe tout autour de la problématique des treillis. De plus, des étudiants de Master ont contribué à des développements dans ce domaine sous la tutelle de S Guillas et K Bertet et, notamment, pour l'implémentation des algorithmes écrit dans le cadre de leur travaux respectifs.

Actuellement, il existe globalement des bouts de codes permettant de gérer les tables binaires, les règles et les treillis eux-même cependant, K Bertet est en cours de développement d'une bibliothèque Java rassemblant tous ces bouts de code. Celle-ci permettra en même temps d'assainir ceux-ci, de les faire interagir les un avec les autres et enfin de développer des petites interfaces permettant leur manipulation de tests.

De ce fait, le code Java sera lisible, documenté et réutilisable.

Conclusion

K Bertet estime que d'ici juin 2009 ces outils seront disponibles et utilisables.