

Une première approche, développée par l'IAP, consiste à utiliser une décomposition selon une base calculée à partir d'une analyse en composante principale (ACP) d'un sous-ensemble d'images issues du catalogue RC3 du SDSS<sup>7</sup>. Ils utilisent ensuite un perceptron multicouches (PMC) pour classifier les galaxies.

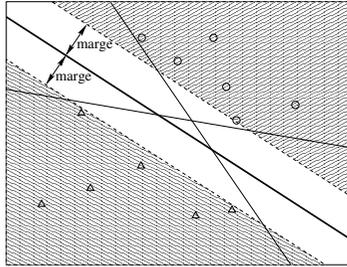


FIG. 2 – Le meilleur séparateur de la méthode SVM

Nous avons évalué une autre approche fondée sur les séparateurs à vaste marge (SVM) proposés par Vapnik et Chervonenkis pour la classification des galaxies. Pour comparer l'approche neuronale avec le SVM, nous avons utilisé la même représentation (la même base ACP) et le même protocole d'évaluation. Comme le PMC, les méthodes SVM cherchent un hyperplan (séparateur linéaire) dans l'espace des caractéristiques qui permet de séparer les classes. La différence consiste au fait que les méthodes SVM déterminent le meilleur séparateur là où le PMC

cherche une solution au problème.

Si vous examinez la figure 2, vous remarquez que tous les hyperplans (des droites dans ce cas) peuvent distinguer les triangles des cercles, mais la droite en gras offre un avantage supplémentaire qui est de maximiser la marge qui sépare les exemples de l'hyperplan (d'où le nom des SVMs). Intuitivement ceci revient à chercher l'hyperplan le plus sûr. En effet, supposant qu'un exemple soit mal défini une petite variation ne modifiera pas sa classification si sa distance à l'hyperplan est grande.

Dans le cas où le séparateur linéaire n'est pas suffisant, on utilise une transformation non linéaire de l'espace vers un espace de dimension supérieure, voire infinie, afin de pouvoir séparer les classes. La distance entre deux points de cet espace est estimée à l'aide d'une fonction particulière appelée *noyau* directement dans l'espace original, ceci permet d'utiliser des espaces de dimension infinie sans alourdir les temps de calcul.

Nous avons testé plusieurs fonctions noyaux et les meilleurs résultats ont été obtenus avec un noyau RBF. En comparaison avec le MLP, les résultats sont plus fiables et restent stables entre l'ensemble d'apprentissage et l'ensemble de test. L'étape suivante consiste à valider ces résultats sur une base plus importante.

Une deuxième idée est actuellement en cours de développement, elle utilise une décomposition en base analyse en composantes indépendantes (ACI) pour trouver la meilleure représentation des caractéristiques de ces galaxies.

## En bref

**Publications** Elles sont toujours disponibles sur notre site, sur <http://publis.lrde.epita.fr/>.

- **Probabilistic abstraction for model checking: an approach based on property testing** par Sophie Laplante, Richard Lassaingne, Frederic Magniez, Sylvain Peyronnet et Michel de Rougemont accepté à *ACM TOCL*. - 21 novembre
- **C-Transformers: A Framework to Write C Transformations** par Alexandre Borghi, Valentin David et Akim Demaille accepté à *Crossroads on Programming Languages, the Association for Computing Machinery Student Ma-*

*gazine*. - 16 octobre

**Séminaire LRDE** Les étudiants CSI 2006 ont présenté leurs travaux sur Transformers, Olena & Traitement des Images et Vaucanson & Automates. - 16 et 17 novembre

**Vaucanson 0.7.2**<sup>8</sup> Cette nouvelle version corrige des bogues (dans le système XML et les démos) et met à jour le projet pour les versions 4.0 du compilateur GNU C++ et 9 du compilateur Intel. - 9 novembre

**Les nouveaux venus** Le labo accueille en thèse Alban Linard et Sébastien Hémon. - octobre



# L'air de rien N° 1

L'aléastriel du Laboratoire de Recherche et de Développement de l'EPITA<sup>1</sup>

Numéro 1.2, Décembre 2005

## Edito

**Pourquoi L'air de rien**  
par Akim Demaille

En diffusion de l'information, le push est bien plus efficace que le pull. Aussi, ce bulletin aléastriel a pour but de pallier une communication quasi inexistante de notre laboratoire et, de votre point de

vue, elle a pour vocation que vous nous connaissiez mieux. À bien regarder, il ne s'agit que d'une compilation synthétique (et rapide à pondre) de ce que nous diffusons déjà sur [notre site web](#)<sup>2</sup>, et en particulier [nos news](#)<sup>3</sup>.

En espérant qu'elle vous intéresse...

## Focus

**Jérôme Darbon, Athlète informatique**  
par Akim Demaille



Urbi et Orbi<sup>4</sup>.

Les années passant, il est devenu un ami proche, indispensable aux bonnes heures de « Chez Louise », le guide nocturne pour les oiseaux de passage au merle moqueur, à l'heure où C-sport. Tard dans la nuit son avenir s'y dessinait : d'abord un DEA (comme on disait à l'époque), puis une thèse. Mais où, et avec qui ?

Jérôme, ce fut d'abord ce grand étudiant de la salle d'en face, à l'époque où le LRDE avait vue, ouïe et odorat sur le périph, à l'emplacement actuel de la SM 25. Toujours actif, il s'armait souvent d'un peu de craie, et au coin d'un tableau nous expliquait sa vision du monde, sa perception du projet

Après avoir réussi un DEA théorique parmi les plus difficiles, il nous rejoindra au LRDE, en qualité de thésard, sous la direction de Théo. D'abord assez ancré dans la pratique, Jérôme s'est libéré et s'est enfin révélé : un éplucheur de séries, un marathonien des champs de Markov, un minimiseur de contraintes avec un maximum d'énergie, un steppeur d'algorithmes mathématiques, un descendant de plus grande pente, un artiste abstrait sur tableau blanc, mais aussi un auteur à succès : pas moins de 12 articles de conférences internationales et 3 de revue !

Cette vie-là s'est achevée vendredi 7 octobre lorsque son jury de thèse faisait de lui un docteur. Tout en préparant son départ, en cherchant un laboratoire d'accueil, Jérôme trouvera encore le temps de travailler sur plusieurs publications simultanément, entre autres avec Sylvain Peyronnet, Igor Ciril et Michaël Chekroun. À son départ nous perdons une source sûre de publications, mais aussi et surtout un ami.

Salut l'athlète !

<sup>1</sup>L'air de rien, <http://publis.lrde.epita.fr/LrdeBulletin>.

<sup>2</sup>notre site web, <http://www.lrde.epita.fr>.

<sup>3</sup>nos news, <http://www.lrde.epita.fr/Lrde/News>.

<sup>4</sup>Urbi et Orbi, <http://projects.lrde.epita.fr/UrbiEtOrbi>.

<sup>7</sup>SDSS, <http://www.sdss.org/>.

<sup>8</sup>Vaucanson 0.7.2, <http://vaucanson.lrde.epita.fr/Vaucanson072>.

## Au fil du temps

### Publi{cité, -cation}

La traduction française du best-seller "C++ Coding Standards"<sup>1</sup> vient de sortir, éditée par Pearson Education, sous le titre "Standards de programmation C++"<sup>5</sup>. Ce livre est particulièrement conseillé aux étudiants qui ont commencé l'apprentissage du C++ et qui souhaitent, dès le début, adopter de bonnes habitudes de programmation. À noter : le relecteur et validateur technique de ce livre est Théo.

<sup>1</sup> C++ Coding Standards : 101 Rules, Guidelines, and Best Practices, Herb Sutter and Andrei Alexandrescu, C++ in Depth Series, Addison-Wesley, 2004.

### Ça bouge !

Cette section a pour but de présenter les changements dans la vie des anciens du LRDE. Nous ne parlerons donc pas de ceux qui continuent leur thèse (même si leurs travaux avancent on ne peut pas parler de "changements") mais de ceux qui en commencent une ou qui changent de labo.

### Les nouveaux thésards

**Nicolas Burrus, 2004.** Détection d'événements visuels saillants avec opérateurs rapides de réduction d'information. *Laboratoire d'InfoRmatique en Images et Systèmes d'information (LIRIS; INSA de Lyon, CNRS, Universités de Lyon, Ecole Centrale de Lyon) et École Nationale Supérieure de Techniques Avancées (ENSTA; Unité Électronique-Informatique) et Université Pierre et Marie Curie (Laboratoire d'Informatique de Paris VI, Pôle Intelligence Artificielle)* avec Jean-Michel Jolion (LIRIS), Thierry Bernard (ENSTA).

**Valentin David, 2005.** Transformation de C++ pour optimisations spécifiques au domaine basées sur les spécifications dans le cadre du calcul scientifique. *University of Bergen (Mathematics and Natural Science Faculty, Computer Science Department), Norvège* avec Magne Haveraaen.

**David Lesage, 2004.** Vasculaire en Imagerie Médicale à l'aide de morphologie mathématique et des ensembles de niveaux; application aux artères coronaires et à la vasculature pulmonaire. *Ecole Nationale Supérieure des Télécommunications (ENST; Département Traitement du Signal et des Images) et Siemens Corporate Research (SCR; Imaging and Visualization Department), Princeton NJ, USA* avec Isabelle Bloch (ENST), Elsa Angelini (ENST) et Gareth Funka-Lea (SCR).

**Francis Maes, 2004.** Apprentissage structuré : application à la transformation de documents. *Université Pierre et Marie Curie (Laboratoire d'Informatique de Paris VI, Pôle Intelligence Artificielle)* avec Patrick Gallinari.

**Jean-Baptiste Mouret, 2004.** Génération par évolution artificielle de contrôleurs neuronaux à dynamiques multiples; application au contrôle d'un robot à ailes battantes. *Université Pierre et Marie Curie (Laboratoire d'Informatique de Paris VI, équipe AnimatLab)* avec Jean-Arcady Meyer et Stéphane Doncieux.

### Les docteurs

On n'en a pas beaucoup parce qu'on est encore un jeune labo (pourvu que ça dure...). On y tient donc beaucoup et on suit leurs premiers battements d'ailes.

**Alexis Angelidis, 2001.** Il vient d'obtenir le titre de Docteur de l'Université d'Otago (*Computer Graphics and Vision Research Laboratory, New Zealand*). Il est actuellement en séjour post-doctoral d'un an à l'Université de Toronto (*Department of Computer Science, Ontario, Canada*). Alexis est l'auteur du logo du LRDE.

**Jérôme Darbon** a soutenu sa thèse intitulée *Composants logiciels et algorithmes de minimisation exacte d'énergies dédiés au traitement des images*. Pour plus d'information sur le docteur, voir le focus.

### Jeu informatique : "histoire d'O"

Dans le programme C ci-dessous :

```
struct O;
typedef struct {} O;
O mina(O);
O main(O) {}
```

se cache une unique erreur; vous avez 1 minute, sans l'aide d'un compilateur, pour la trouver. Idem pour ce programme C++ :

```
struct O {};
O main(O) { O O = O; }
```

## EFIGI, la tête dans les galaxies

<http://www.efigi.org/>  
par Olivier Ricou et Réda Dehak



EXTRACTION DE FORMES  
IDÉALISÉES DE GALAXIES  
EN IMAGERIE

Le problème est relativement simple. Si on regarde une photo du ciel, [une bonne photo](#)<sup>6</sup>, on y voit des milliards de galaxies comme la nôtre, enfin pas exactement comme la nôtre et c'est là le but du projet : reconnaître de quel type sont les galaxies. Comme elles sont bien trop nombreuses, il faut écrire un programme qui les classe automatiquement suivant leur type, ce type pouvant être celui de Hubble, cf [figure 1](#), mais si on peut être plus précis, c'est encore mieux.

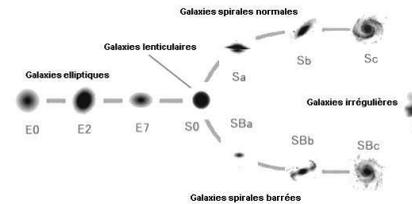


FIG. 1 – Classification de Hubble

Pour faire la classification des galaxies de l'univers, restons modeste, l'Institut d'Astrophysique de Paris, l'IAP, a monté un projet sur trois ans avec d'autres laboratoires d'astronomie mais aussi avec deux laboratoires d'informatique, le [Laboratoire de Traitement et Communication de l'Information de l'ENST-Paris](#) et le LRDE.

Notre travail porte sur quatre parties :

1. le nettoyage des images dont Théo s'occupe, cf l'article de Théo dans le prochain numéro,
2. participer à l'extraction des images de caractéristiques décrivant les galaxies (taille, nombre de bras, ratio bulbe/taille...),
3. la classification en type de galaxies suivant les

caractéristiques extraites à l'étape précédente, travail sur lequel Réda se concentre, voir son article ci-dessous,

4. enfin pour Olivier, la parallélisation et la distribution de tout ce travail sur des clusters car, petit détail, la prochaine génération de télescope sortira plus de 10 TO de données par nuit, qui bien sûr doivent être traitées en 24 heures.

Ce projet ambitieux est un véritable défi pour le LRDE qui devra trouver les méthodes pour traiter ces teraoctets de données en temps réel.

### Évaluation d'un algorithme de classification automatique des galaxies fondé sur les SVMs

Le but principal du projet EFIGI est de fournir une classification morphologique des galaxies. Pour cette première année, les principales pistes explorées se sont limitées à la classification de Hubble, ce qui consiste à déterminer le type de la galaxie parmi les 17 classes proposées par Hubble, cf [figure 1](#).

Une des contraintes de ce projet est de développer une méthode fiable et indépendante du type de l'image. On retrouve dans ce domaine deux types d'images : les images récupérées sur un satellite (exemple HST : Hubble Space Telescope) ou bien des images issues d'un télescope terrestre. La deuxième catégorie d'images souffre d'un flou introduit par la présence de l'atmosphère terrestre, mais reste une solution plus économique que la première, d'où l'intérêt d'avoir une méthode stable dans les deux cas.

Pour s'affranchir de ce problème présent sur les images terrestres, l'approche choisie s'est orientée vers la représentation des images de galaxies dans une base orthonormale. Cette base est censée représenter les caractéristiques principales qui peuvent discriminer les différentes catégories de galaxies. En plus de la compression, cette approche a le mérite d'être utilisable pour les deux cas car il suffit de convoluer les vecteurs de la base par la réponse impulsionnelle du filtre correspondant à l'atmosphère terrestre (l'effet flou) pour l'utiliser avec des images terrestres. En pratique, cette réponse impulsionnelle peut être estimée à partir des étoiles présentes sur l'image.

<sup>5</sup>Standards de programmation C++, <http://www.amazon.fr/exec/obidos/ASIN/2744071447/>.

<sup>6</sup>une bonne photo, [http://terapix.iap.fr/cpl/Press\\_Release/gri\\_d1.jpg](http://terapix.iap.fr/cpl/Press_Release/gri_d1.jpg).