Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Attention,
toute proposition doit faire l'objet d'un résumé enregistré
avant le 22 mars 2004
directement sur le site web de l'ACI :
http://acimd.liris.cnrs.fr

Chaque dossier doit être

- **déposé au plus tard le** $\boxed{\textbf{29 mars 2004}}$ **sur le site web de l'ACI**[1] **:**

  http://acimd.liris.cnrs.fr

- **envoyé par voie postale avec les signatures requises**

  avant le 04 avril 2004
  (cachet de la poste faisant foi)

  à

Ministère délégué à la Recherche et aux Nouvelles Technologies
Direction de la Recherche
Cellule ACI
ACI Masses de données
1, rue Descartes
75231 Paris cedex 05

[1] En cas de difficulté, une soumission par courrier électronique est possible à l'adresse acimd@liris.cnrs.fr

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

I - FICHE D'IDENTITÉ DU PROJET

Nom du Projet : *(maximum 20 caractères)*

EFIGI

Titre du Projet : *(maximum 3 lignes)*

Extraction de Formes Idéalisées de Galaxies en Imagerie

Type du Projet[2] :

Projet de recherche	Projet de recherche multi-thématiques	Projet de recherche avec infrastructure	Projet associé à l'imagerie spatiale	Autre
X				

Durée du projet[3] : 36 mois

Description courte du Projet : *(une demi-page maximum)*

Les grands relevés d'imagerie astronomique accumulent aujourd'hui des quantités de données de l'ordre de 10^{13} pixels, et 10^{9} sources ; essentiellement des galaxies. Seul un catalogage systématique et automatique des ces sources permet d'exploiter scientifiquement leur contenu. En particulier, les statistiques concernant les formes des galaxies sont au coeur de questions fondamentales de l'astrophysique moderne (morphogenèse, marqueurs d'évolution, d'interactions, ou encore de distorsions produites par des lentilles gravitationnelles), Malheureusement, les limitations algorithmiques et matérielles des solutions actuelles excluent des catalogues des grands relevés les informations morphologiques de haut niveau. Même les solutions "classiques" existantes, basées sur des ajustements de profils, sont environ 100 fois trop lentes par rapport aux débits typiques d'acquisition des données. Notre projet est d'associer des laboratoires ayant des expertises complémentaires (apprentissage automatique, calcul paralléle et optimisation de code du coté STIC, traitement/analyse des images du ciel et morphologie des galaxies du coté astronomique) autour de données panchromatiques sol (MEGACAM/WIRCAM) et spatial (GALEX), afin de créer un système générique et performant d'analyse de haut niveau de la morphologie des galaxies.

Coordinateur du projet :

Nom	Prénom	Laboratoire (sigle éventuel et nom complet)
Bertin	**Emmanuel**	**Institut d'Astrophysique de Paris (IAP)**

Organisme de rattachement financier pour le présent projet :

CNRS

Équipes ou laboratoires partenaires (nom complet et éventuellement sigle)[4]

[2]Cocher la case correspondante au type du projet soumis.

[3]La durée d'un projet ne peut excéder 36 mois. Des demandes de projets d'une durée plus courte devront être particulièrement argumentées.

Institut d'Astrophysique de Paris (IAP)
Laboratoire Traitement et Communication de l'Information (LTCI) UMR 5141 - Dept TSI - GET/Telecom Paris
Laboratoire de Recherche et Développement de l'EPITA (LRDE)
Laboratoire d'Astronomie de Marseille (LAM)
Laboratoire d'Astrophysique de l'Observatoire Midi-Pyrénées (LAT)
Centre de Recherche Astronomique de Lyon (CRAL)
Service d'Astrophysique, CEA/Saclay (SAp)

[4] Insérer autant de lignes que nécessaire.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Informations de cadrage du projet :

Durée : **36** mois

Moyens demandés dans le cadre de l'ACI via le Fonds National de la Science (montants en Euros TTC) :

Equipement	**50k**
Fonctionnement	**266k**
Dépenses de personnels (CDD)	**120k**
Total	**316k**

II - PRÉSENTATION DÉTAILLÉE DU PROJET

A - IDENTIFICATION DU COORDINATEUR ET DES AUTRES PARTENAIRES DU PROJET :

A1 - Coordinateur du Projet :

Un unique coordinateur doit être désigné par les partenaires.

M. ou Mme. Prénom Nom[5]	M. Emmanuel Bertin
Fonction[5]	Astronome-adjoint
Laboratoire (Nom complet et sigle le cas échéant)[5]	Institut d'Astrophysique de Paris et LERMA/Observatoire de Paris
Adresse[5]	98bis, bd Arago, F-75014 Paris
Téléphone[5]	01 44 32 81 51
Fax	01 44 32 80 01
Mél[5]	bertin@iap.fr

[5] Champ obligatoire.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

A2- Équipes ou laboratoires partenaires du Projet [6] :

Identification de l'équipe ou du laboratoire

Équipe ou Laboratoire	Institut d'Astrophysique de Paris (IAP)
Adresse	98bis bd Arago, F-75014 Paris

Organisme de rattachement financier de l'équipe pour le présent projet

CNRS Paris-A UMR 7095

Responsable du projet au sein de l'équipe ou du laboratoire

M. ou Mme. Prénom Nom	Mr Emmanuel Bertin
Fonction	Astronome-Adjoint
Téléphone	01 44 32 81 51
Fax	01 44 32 80 01
Mél	bertin@iap.fr

Membres de l'équipe participant au projet (y compris le responsable)

Nom	Prénom	Poste statutaire	% du temps de recherche consacré au projet
Bertin	Emmanuel	Astronome-adjoint	50
Mellier	Yannick	Astronome	30
McCracken	Henri-Joy	Astronome-adjoint	40
de Lapparent	Valérie	C.R.1 CNRS	30
van Waerbeke	Ludovic	C.R.1 CNRS	20
Domisse	Laurent	Ingénieur d'étude	40
Magnard	Fréderic	Ingénieur de recherche	50
Missonnier	Gilles	Ingénieur de recherche	40
Dantel-Fort	Mireille	Ingénieur d'étude	30

Références :

Pour chaque (enseignant-)chercheur participant, liste de 3 à 5 publications, logiciels ou brevets les plus significatifs, en relation avec la thématique du projet.

Bertin E, 1994, "Classification of astronomical images with a neural network" *Astrophysics and Space Science* **217**, 49

Bertin E., Arnouts S., 1996, "SExtractor : Software for source extraction", *Astronomy &Astrophysics Supplements* **117**, 393

Bertin E, 1997, "Self Organizing Maps and Imaging Surveys", in *The Impact of Large Scale Near-IR Sky Surveys*, Kluwer Dordrecht, 221

Bertin E., 1997, "Galaxy evolution at low redshift ? I. Optical counts", *Astronomy & Astrophysics* **317**, 43

Bertin E., 1999, "SExtractor 2.1, User's manual", IAP

Bertin E., 2001, "Mining Pixels : The Extraction and Classification of Astronomical Sources", in *Mining the Sky*, Springer-Verlag, Heidelberg

Bertin E., Thion A., Mellier Y., van Waerbeke L., 2001, "Weak Lensing Analysis with SExtractor", in *Gravitational Lensing : Recent Progress and Future Goals*, ASP Conference Proceedings, Vol. **237**. Brainerd T.G. & Kochanek C.S. eds., 365

[6] Une fiche doit être remplie pour chaque laboratoire ou équipe partenaire.

Bonnet H., Mellier Y., 1995, “Statistical analysis of gravitational shear in the extended periphery of rich clusters of galaxies”, *Astronomy & Astrophysics* **303**, 331

Cabanac R.A., de Lapparent‘V., Hickson‘P., 2002, “Classification and redshift estimation by principal component analysis”, *Astronomy & Astrophysics* **389**, 1090

de Lapparent V., Galaz G., Bardelli S., Arnouts S., 2003, “The ESO-Sculptor Survey : Luminosity functions of galaxies per spectral type at redshifts 0.1-0.5”, *Astronomy & Astrophysics* **404**, 831

de Lapparent V., 2003, “Critical analysis of the luminosity functions per galaxy type measured from redshift surveys”, *Astronomy and Astrophysics* **408**, 845

“How accurately can we measure weak gravitational shear ?” Erben T., Van Waerbeke L., Bertin E., Mellier Y., Schneider P., 2001, *Astronomy &Astrophysics* **366**, 717

McCracken H.J., Shanks T., Metcalfe N., Fong R., Campos A., 2000, “Galaxy clustering in the Herschel Deep Field”, *Monthly Notices of the Royal Astronomical Society*, **318**, 913

McCracken H.J., Le Fèvre O., Brodwin M., Foucaud S., Lilly S.J., Crampton D., Mellier Y., 2001, *Astronomy & Astrophysics* **376**, 756

McCracken H.J., Radovich M., Bertin E., Mellier Y., Dantel-Fort M., Le Fèvre O., Cuillandre J.C., Gwyn S., Foucaud S., Zamorani G., 2003, “The VIRMOS deep imaging survey. II : CFH12K BVRI optical data for the 0226-04 deep field”, *Astronomy & Astrophysics* **410, 17**

Mellier Y., 1999, “Probing the Universe with Weak Lensing”, *Annual Review of Astronomy and Astrophysics* **37**, 127

Mellier Y., van Waerbeke L., Bertin E., Tereno I., Schneider P., Bernardeau F., Erben T., 2003, “Prospects for weak lensing/cosmic shear with VLTs”, *SPIE* **4834**, 223

Mellier Y. ; van Waerbeke L., Bertin E., Tereno I., Bernardeau F., 2003, “Wide-field cosmic shear surveys”, *SPIE* **4847**, 112

Mellier Y., Bertin E., Missonnier G., Didelon P., Radovich M., Domisse L., Morin B., Dantel-Fort M., 2002, “TERAPIX Report for the Project Review”, IAP

Roche M., Bracco C., Aime C., Lantéri H., Mellier Y., 2003, “Reduction of ringing effects by an improved deconvolution algorithm. Application to A370 CFHT image of gravitational arcs”, *Astronomy & Astrophysics* **409, 387**

van Waerbeke L., Mellier Y., Erben T., Cuillandre J.-C., Bernardeau F., Maoli R., Bertin E., McCracken H.-J., Le Fèvre O., Fort B., Dantel-Fort M., Jain B., Schneider P., 2000, “Detection of correlated galaxy ellipticities on CFHT data : first evidence for gravitational lensing by large-scale structures”, *Astronomy & Astrophysics* **358**, 30

van Waerbeke L., Mellier Y., Pelló R., Pen U.-L., McCracken H.J., Jain B., 2003, “Likelihood analysis of cosmic shear on simulated and VIRMOS-DESCART data”, *Astronomy & Astrophysics* **393**, 369

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Identification de l'équipe ou du laboratoire

Équipe ou Laboratoire	LTCI UMR 5141 GET/Télécom Paris
Adresse	46, rue Barrault F-75013 Paris

Organisme de rattachement financier de l'équipe pour le présent projet

CNRS Paris-A UMR 5141

Responsable du projet au sein de l'équipe ou du laboratoire

M. ou Mme. Prénom Nom	M. Henri Maître
Fonction	Professeur
Téléphone	01 45 81 76 55
Fax	01 45 81 37 94
Mél	henri.maitre@enst.fr

Membres de l'équipe participant au projet (y compris le responsable)

Nom	Prénom	Poste statutaire	% du temps de recherche consacré au projet
Maître	Henri	Professeur	20
Moulines	Eric	Professeur	10
Cardoso	Jean-François	D.R. CNRS	20
Gousseau	Yann	C.R. CNRS	20
Cappé	Olivier	C.R. CNRS	20
Campedel	Marine	Post-doc	30

Références :

Pour chaque (enseignant-)chercheur participant, liste de 3 à 5 publications, logiciels ou brevets les plus significatifs, en relation avec la thématique du projet.

M. Campedel et E. Moulines *Classification et sélection automatique de caractéristiques de textures* SFC 04, Brest 2004 (soumis)

M. Campedel et E. Moulines *Modélisation de textures* Conférence d'Apprentissage CAp 2004, Montpellier, (soumis)

R. Douc, E. Moulines et T. Ryden *Asymptotic Properties of the Maximum Likelihood Estimator in Autoregressive Models with Markov Regime* Annals of Statististics, 2004 (à paraitre)

J-F. Cardoso, J. Delabrouille and G. Patanchon *Independent Component Analysis of the Cosmic Microwave Background* International Symposium on Independent Component Analysis and Blind Source Separation (ICA), Nara, Japan, apr 2003, pp. 1111-1116

H. Maître *Le traitement des images* Hermès, Traité IC2, Paris, France 2003

O. Cappé *A Bayesian Approach for Simultaneous Segmentation and Classification of Count Data* IEEE Transactions on Signal Processing feb 2002 vol. 50, n° 2, pp. 400-410

P. Muse, F. Sur, F. Cao et Y. Gousseau *Unsupervised thresholds for shape matching* Proceedings of the IEEE International Conference on Image Processing Barcelona, Spain, sep 2003

Y. Gousseau *Comparaison de la composition de deux images, et application a la recherche automatique* GRETSI 2003 Paris, France sep 2003

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Identification de l'équipe ou du laboratoire

Équipe ou Laboratoire	Laboratoire de Recherche et Développement de l'EPITA (LRDE)
Adresse	14-16 rue Voltaire, F-94276 Le Kremlin-Bicêtre Cedex, France

Organisme de rattachement financier de l'équipe pour le présent projet

Association loi 1901 ("Association EPITA", 14-16 rue Voltaire, 94270 Le Kremlin-Bicêtre, SIRET 443 220 223 00012, APE/NAF 803Z)

Responsable du projet au sein de l'équipe ou du laboratoire

M. ou Mme. Prénom Nom	M. Thierry Géraud
Fonction	Maître de conférence
Téléphone	01 53 14 59 47
Fax	01 53 14 59 22
Mél	thierry.geraud@lrde.epita.fr

Membres de l'équipe participant au projet (y compris le responsable)

Nom	Prénom	Poste statutaire	% du temps de recherche consacré au projet
Géraud	Thierry	Maître de conférence	20
Demaille	Akim	Maître de conférence	10
Dehak	Réda	Maître de conférence	30 (années 1 et 2)
Ricou	Olivier	Maître de conférence	30 (année 3)

Références :

Pour chaque (enseignant-)chercheur participant, liste de 3 à 5 publications, logiciels ou brevets les plus significatifs, en relation avec la thématique du projet.

Nicolas Burrus, Alexandre Duret-Lutz, Thierry Géraud, David Lesage and Raphaël Poss. *A Static C++ Object-Oriented Programming (SCOOP) Paradigm Mixing Benefits of Traditional OOP and Generic Programming* in the Proceedings of the Workshop on Multiple Paradigm with OO Languages (MPOOL'03), Anaheim, CA Oct. 2003.

Thierry Géraud, Yoann Fabre, Alexandre Duret-Lutz, Dimitri Papadopoulos-Orfanos, and Jean-Franc,ois Mangin *Obtaining Genericity for Image Processing and Pattern Recognition Algorithms* in the proceedings of the International Conference on Pattern Recognition (ICPR'2000), IEEE Computer Society, vol. 4, pages 816-819, Barcelona, Spain, September 2000

Heru Xue, Thierry Géraud and Alexandre Duret-Lutz. Multi-band Segmentation using Morphological Clustering and Fusion. In the proceedings of the IEEE International Conference on Image Processing (ICIP 2003), Barcelona, Sep. 2003

Jérôme Darbon, Thierry Géraud and Alexandre Duret-Lutz. Generic Implementation of Morphological Image Operators. In the proceedings of the International Symposium on Mathematical Morphology VI (ISMM'2002), pages 175-184, Sydney, Australia, April 2002

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Identification de l'équipe ou du laboratoire

Équipe ou Laboratoire	Laboratoire d'Astronomie de Marseille (LAM)
Adresse	Traverse du Siphon - Les trois Lucs, BP8-13376 Marseille Cedex 12

Organisme de rattachement financier de l'équipe pour le présent projet

CNRS - UMR 6110

Responsable du projet au sein de l'équipe ou du laboratoire

M. ou Mme. Prénom Nom	M. Stéphane Arnouts
Fonction	C.R.1 CNRS
Téléphone	04 91 05 59 93
Fax	04 91 66 18 55
Mél	Stephane.Arnouts@oamp.fr

Membres de l'équipe participant au projet (y compris le responsable)

Nom	Prénom	Poste statutaire	% du temps de recherche consacré au projet
Arnouts	Stéphane	C.R.1 CNRS	30
Millard	Bruno	C.R.1 CNRS	10
Donas	José	C.R.1 CNRS	20
Treyer	Marie-Agnès	C.R.1 CNRS	10

Références :

Pour chaque (enseignant-)chercheur participant, liste de 3 à 5 publications, logiciels ou brevets les plus significatifs, en relation avec la thématique du projet.

Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, "Measuring and modelling the redshift evolution of clustering : the Hubble Deep Field North", *Monthly Notices of the Royal Astronomical Society* **310**, 540

Arnouts S., Vandame B., Benoist C., Groenewegen M.A.T., da Costa L., Schirmer M., Mignani R.P., Slijkhuis R., Hatziminaoglou E., Hook R., Madejsky R., Rité C., Wicenec A., 2001, *Astronomy & Astrophysics* **379**, 740

Arnouts S., Moscardini L., Vanzella E., Colombi S., Cristiani S., Fontana A., Giallongo'E., Matarrese S., Saracco P., 2002, "Measuring the redshift evolution of clustering : the Hubble Deep Field South", *Monthly Notices of the Royal Astronomical Society* **329**, 355

Donas J., 2001, "Ultraviolet Observations of Galaxies", *Astrophysics and Space Science Supplement* **277**, 325

Martin C. et al., 2003, "The Galaxy Evolution Explorer", *SPIE* **4854**, 336

Milliard B., Donas J., Laget M., Armand C., Vuillemin A., 1992, "Galaxy counts at ultraviolet wavelengths (2000 Å)", *Astronomy and Astrophysics*, **257**, 24

Milliard B., Donas J., Buat V., Deharveng J.-M., 2001, "UV measurements of the local SFR", *American Astronomical Society* **198**, 888

Schiminovich D., Arnouts S., Milliard B., 2003, "The GALEX Deep Fields : UV-selected galaxies at z=1", *American Astronomical Society Meeting* **203**, 103

Treyer M.-A., Ellis R.-S., Milliard B., Donas J., Bridges T.J., 1998, "An ultraviolet-selected galaxy redshift survey : new estimates of the local star formation rate", *Monthly Notices of the Royal Astronomical Society*, **300**, 303

Treyer M.-A., 2001, "Star-formation in the local universe" *New Astronomy Reviews* **45**, 351

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Identification de l'équipe ou du laboratoire

Équipe ou Laboratoire	Laboratoire d'Astrophysique de l'Observatoire Midi-Pyrénées
Adresse	14, avenue Edouard Belin - F-31400 Toulouse

Organisme de rattachement financier de l'équipe pour le présent projet

CNRS UMR 5572

Responsable du projet au sein de l'équipe ou du laboratoire

M. ou Mme. Prénom Nom	M. Pascal Fouqué
Fonction	Astronome
Téléphone	05 61 33 27 86
Fax	05 61 33 28 40
Mél	pascal.fouque@ast.obs-mip.fr

Membres de l'équipe participant au projet (y compris le responsable)

Nom	Prénom	Poste statutaire	% du temps de recherche consacré au projet
Fouqué	Pascal	Astronome	30
Pelló	Roser	Astronome-adjointe	20
Le Borgne	Jean-François	Astronome	10

Références :

Pour chaque (enseignant-)chercheur participant, liste de 3 à 5 publications, logiciels ou brevets les plus significatifs, en relation avec la thématique du projet.

Bolzonella M., Pelló R., Maccagni D., 2002, "Luminosity functions beyond the spectroscopic limit. I. Method and near-infrared LFs in the HDF-N and HDF-S", *Astronomy and Astrophysics* **395**, 443

de Vaucouleurs G., de Vaucouleurs A., Corwin H.G.Jr., Buta R.J., Paturel G., Fouqué P., 1991, "Third Reference Catalogue of Bright Galaxies", Volumes 1-3, Springer-Verlag

Fouqué P., Solanes J.M., Sanchis T., Balkowski C., 2001, "Structure, mass and distance of the Virgo cluster from a Tolman-Bondi model" *Astronomy & Astrophysics* **375**, 770

Le Borgne J.F., Vilchez-Gomez R., 1993, "An Optical Identification of Radio Sources in the Field of the Cluster of Galaxies Abell 2218", *Astronomy & Astrophysics* **271**, 425

Pelló R. ; Bolzonella'M., Lemoine-Busserolle M., Le Borgne J.-F., 2001 "Photometric redshifts in the CFH Legacy Survey", *Canada-France-Hawaii Telescope Information Bulletin* **43**, 7

Pelló R., Schaerer D., Richard J., Le Borgne J.-F., Kneib J.-P., 2004, "ISAAC/VLT observations of a lensed galaxy at z = 10.0", *Astronomy & Astrophysics* **416**, 35

Vauglin I., Rousseau J., Paturel G., Borsenberger J., Epchtein N., Fouqué P., Kimeswenger S., Le Bertre'T., Mamon G.A., 2002, "Serendipitous detection of galaxies behind the Milky Way from the DENIS survey", *Astronomy & Astrophysics*, **387**, 1

Vauglin I., Paturel G., Borsenberger J., Fouqué P., Epchtein N., Kimeswenger S., Tiphène D., Lanoix P., Courtois H., 1999, "First DENIS I-band extragalactic catalog" *Astronomy & Astrophysics Supplement* **135**, 133

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Identification de l'équipe ou du laboratoire

Équipe ou Laboratoire	Centre de Recherche Astronomique de Lyon (CRAL)
Adresse	9, avenue Charles André, F-69561 Saint Genis-Laval Cedex

Organisme de rattachement financier de l'équipe pour le présent projet

CNRS UMR 5574

Responsable du projet au sein de l'équipe ou du laboratoire

M. ou Mme. Prénom Nom	M. Georges Paturel
Fonction	Astronome
Téléphone	04 78 86 85 45
Fax	04 78 86 83 86
Mél	patu@obs.univ-lyon1.fr

Membres de l'équipe participant au projet (y compris le responsable)

Nom	Prénom	Poste statutaire	% du temps de recherche consacré au projet
Paturel	Georges	Astronome	20
Prugniel	Philippe	Astronome	10
Chilingarian	Igor	Visiteur	40
Theureau	Gilles	Astronome-adjoint	10
Petit	Chantal	Tech. CNRS	40
Tajahmady	Françoise	I.R. CNRS	10

Références :

Pour chaque (enseignant-)chercheur participant, liste de 3 à 5 publications, logiciels ou brevets les plus significatifs, en relation avec la thématique du projet.

de Vaucouleurs G., de Vaucouleurs A., Corwin H.G.Jr., Buta R.J., Paturel G., Fouqué P., 1991, "Third Reference Catalogue of Bright Galaxies", Volumes 1-3, Springer-Verlag

Paturel G., Fang Y., Petit C., Garnier R., Rousseau J., 2000, "An image database. III. Automatic extraction for millions of galaxies", *Astronomy & Astrophysics Supplement* **146**, 19

Paturel G. ; Petit C., Prugniel P., Theureau G., Rousseau J., Brouty M., Dubois P., Cambrésy L., 2003, "HYPERLEDA. I. Identification and designation of galaxies", *Astronomy & Astrophysics* **412**, 45

Prugniel P., Heraudeau P., 1998, "Total magnitude, radius, colour indices, colour gradients and photometric type of galaxies", *Astronomy & Astrophysics Supplement* **128**, 299

Prugniel P., Maubon G., Simien F., 2001, "The formation of galaxy bulges : Spectrophotometric constraints", *Astronomy & Astrophysics* **366**, 68

Prugniel P., Chilingarian I., Flores H., Guibert J., Jegouzo I., Royer F., Tajahmady F., Theureau G., Vétois J., 2003, "MIGALE : A Multiparametric Virtual Instrument to Study Galaxy Evolution", *SF2A-2003*, EdP-Sciences, Conference Series, 283

Prugniel P., 1985-2003, logiciel "Plein pot", `http ://www-obs.univ-lyon1.fr/hypercat/pleinpot/pleinpot.html` Theureau G., Hanski M., Teerikorpi P., Bottinelli L., Ekholm T., Gouguenheim L., Paturel G., 1997, "Kinematics of the local universe. IV. Type dependence in the diameter Tully-Fisher relation and implications on the mass-luminosity structure", *Astronomy & Astrophysics* **319**, 435

Theureau G., 1998, "Kinematics of the local universe. VI. B-band Tully-Fisher relation and mean surface brightness", *Astronomy & Astrophysics* **331**, 1

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Identification de l'équipe ou du laboratoire

Équipe ou Laboratoire	Service d'Astrophysique du Commissariat à l'Energie Atomique (SAp)
Adresse	CEA Saclay, Orme des Merisiers, F-91191 Gif-sur-Yvette Cedex

Organisme de rattachement financier de l'équipe pour le présent projet

CEA/DAPNIA

Responsable du projet au sein de l'équipe ou du laboratoire

M. ou Mme. Prénom Nom	M. Alexandre Refregier
Fonction	Ingénieur-chercheur CEA
Téléphone	01 69 08 34 92
Fax	01 69 08 65 77
Mél	arefreg@discovery.saclay.cea.fr

Membres de l'équipe participant au projet (y compris le responsable)

Nom	Prénom	Poste statutaire	% du temps de recherche consacré au projet
Refregier	Alexandre	Ingénieur-chercheur CEA	20

Références :

Pour chaque (enseignant-)chercheur participant, liste de 3 à 5 publications, logiciels ou brevets les plus significatifs, en relation avec la thématique du projet.

Heavens A., Refregier A., Heymans C., 2000, "Intrinsic correlation of galaxy shapes : implications for weak lensing measurements", *Monthly Notices of the Royal Astronomical Society* **319**, 649

Massey R., Refregier A., Conselice C.J., Bacon D.J., 2004, "Image simulation with shapelets" *Monthly Notices of the Royal Astronomical Society* **348**, 214

Refregier A., Bacon D., 2003, "Shapelets - II. A method for weak lensing measurements", *Monthly Notice of the Royal Astronomical Society* **338**, 48

Refregier A., 2003, "Shapelets - I. A method for image analysis", *Monthly Notice of the Royal Astronomical Society* **338**, 35

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

B - DESCRIPTION DU PROJET

B1 – Objectifs et contexte :

On précisera, en particulier, les verrous scientifiques et technologiques à dépasser, l'état de l'art ainsi que les projets concurrents ou similaires connus dans le contexte national et international, en particulier ceux auxquels les équipes du projet participent.

En termes de masses de données, l'astronomie moderne est largement dominée par les images des relevés effectués à partir de caméras électroniques à grand champ. Ces dernières ont, depuis la fin des années 90, avantageusement remplacé les plaques photographiques en usage pendant près d'un siècle. Elles offrent une réponse en flux plus linéaire et plus homogène, un meilleur piqué, et une sensibilité décuplée. Ce développement s'accompagne logiquement d'un effort important au niveau du traitement de données. Ainsi pour un seul instrument le volume de pixels à traiter et analyser dépasse actuellement les 50 Toctets. Par ailleurs d'une génération d'instruments à l'autre le gain en *débit* de données excède le progrès en puissance de calcul par processeur (loi de Moore), et plus encore celui du débit des entrées-sorties durant la même période de temps, ce qui rend désormais nécessaire une approche distribuée à la fois en calcul et en stockage.

Après calibration, l'exploitation scientifique des données se fait essentiellement au travers d'un catalogue des sources (astres) détectées automatiquement dans les images. Des statistiques exhaustives effectuées sur des attributs de ces astres (position dans le ciel, flux, orientation, etc.) sont issues une grande partie de nos connaissances actuelles en cosmologie (géométrie et densité à grande échelle de l'Univers, matière et énergie sombre, évolution des galaxies) et en structure galactique (populations d'étoiles, dynamique de la Galaxie). Du point de vue de l'astronome, le catalogue idéal est une forme de version compressée des images qui préserve et condense tout le contenu scientifique exploitable.

Les sources ponctuelles (indistinguables de la réponse impulsionnelle) comme les étoiles sont relativement simples à cataloguer : position, flux et eventuellement couleur suffisent à caractériser entièrement leur contribution aux images. Cependant les objets résolus, constitués à une écrasante majorité de galaxies, posent bien d'autres problèmes. Les paramètres de formes extraits des sources des relevés actuels se résument le plus souvent à des rapports d'axes et angles de position issus des moments du 2ème ordre. A ceci trois raisons essentielles, qui sont autant de défis :

1. Le temps de calcul : les volumes de données à analyser (typiquement 10^9 sources) imposent un temps moyen de calcul par source ne pouvant excéder une dizaine de millisecondes.
2. Les complications introduites par la convolution des images des galaxies par la réponse impulsionnelle ; depuis le sol une grande partie des galaxies détectées ont un rayon d'échelle équivalent à celui de la réponse impulsionnelle. La mesure de paramètres "déconvolués" doit donc fréquemment faire face à des problèmes de dégénérescence, qui ne peuvent être résolus que par l'application d'*a priori* précis sur la morphologie des sources.
3. La nécessité de maîtriser les erreurs de mesure systématiques : les mesures morphologiques effectuées sur les sources ne sont exploitables scientifiquement que si les biais éventuels qui les affectent peuvent être parfaitement maitrisés. A titre

d'exemple, certains objectifs scientifiques comme la mesure du cisaillement gravitationnel aux très grandes échelles à partir de l'élongation des profils, nécessitent des erreurs systématiques inférieures au 1/1000.

En conséquence, les catalogues des grands relevés actuels ne contiennent aucune description morphologique sophistiquée des galaxies. **Notre objectif est de combler cette lacune, en proposant un système générique et performant de description et classification morphologiques des galaxies détectées sur les images astronomiques**. Les statistiques de formes des galaxies sont en effet au cœur de questions fondamentales de l'astrophysique moderne ; citons en particulier

- La morphogenèse des galaxies : comment se développent les différents types constituant la séquence de Hubble, leur évolution au sein des amas en fonction du décalage spectral et de la densité locale, triaxalité des profils tridimensionnels ou sous-structures internes, fréquence et distribution en taille des barres et des "coquilles" internes de galaxies, fonctions de luminosité par type.
- L'évolution des composantes de populations stellaire des galaxies, à travers l'histoire de la formation stellaire, et les interactions gravitationnelles éventuelles.
- Les collisions entre galaxies, et leur dépendance avec le décalage spectral et l'environnement local.
- Les distorsions gravitationnelles sur la ligne de visée : recherche automatique d'arcs gravitationnels signalant des concentrations de matière noire, cartographie des grandes structures en avant-plan, corrélation matière-noire et lumière ("biais") ou matière noire des halos et type morphologique.

Historiquement, les efforts de description morphologique des galaxies se sont principalement portés sur la modélisation des profils des objets. Dans les programmes du type GIM2D (Simard 1998), une galaxie est décomposée en 2 composantes axisymétriques de bulbe ("loi de de Vaucouleurs") et de disque (loi exponentielle tronquée), chacune décrite par 4 à 5 paramètres, pour un total de 10 paramètres par objet. L'ajustement est évidemment non-linéaire et réalisé par minimisation globale. A chaque pas de calcul, une reconvolution du modèle par la réponse impulsionnelle locale est nécessaire. Pour les objets compacts, il est important de limiter le modèle en bande passante spatiale, ou de travailler sur une grille convenablement sur-échantillonnée. Une majorité de galaxies des relevés profonds est très bruitée et/ou mal résolue, et les solutions sont sujettes à des dégénérescences importantes. D'où l'usage d'algorithmes lents mais sachant éviter les minima locaux, du type Metropolis (Metropolis 1953). La conséquence est un temps moyen de calcul de une dizaine à quelques dizaines de secondes par galaxie sur un processeur 32bits à 2GHz (Simard 2004, communication privée), soit environ 1000 fois trop élevé pour nos spécifications (ou 100 fois pour une ferme de 10 processeurs). Au moyen de multiples artifices (calcul par secteur angulaire de pixels, minimisation abrégée), nos collègues américains du projet concurrent SDSS parviennent à ajuster des profils simplifiés avec un temps moyen de calcul de quelques dizaines de millisecondes par galaxie (Lupton 2004, communication privée), mais la qualité des solutions obtenues reste à estimer, surtout pour les relevés profonds comme le CFHTLS.

En complément de ces ajustements de profils symétriques, d'autres paramètres empiriques tels que coefficient de Gini, indices d'"asymétrie", de "concentration", (voir par exemple Bershady et al. 2000, Conselice et al. 2000, Abraham et al. 2003) sont évalués, mais ils sont difficilement utilisables pour des expériences sol en raison notamment de la difficulté à contrôler les effets de la réponse impulsionnelle. Il en va de même des paramètres de contour ou isophotaux employés dans les tentatives passées de classification morphologique par apprentissage supervisé (Storrie-Lombardi et al. 1992, Naim et al. 1995, Owens et al. 1996, Odewahn et al. 1996, Bazell & Aha 2001).

Conscients de ce problème, un certain nombre d'auteurs ont proposé plus récemment une décomposition *linéaire* des images de galaxies sur des bases de fonctions : composantes principales (Bertin 2001) ou polynômes de Gauss-Hermitte (Bernstein & Jarvis 2002, Refregier 2003), les “shapelettes”. Ces dernières ont une expression analytique simple et facilitent l’approximation des transformations géométriques élémentaires (décalage, anamorphose, rotation, ...) sous forme de combinaisons d’opérateurs ascendants et descendants. En ajustant les fonctions de base convoluées par la réponse impulsionnelle locale, il est possible de déconvoluer paramétriquement les images. Toutefois une solution basée uniquement sur une maximisation de la vraisemblance est susceptible de dégénérescences, en particulier dans les directions azimuthales. Une partie des efforts actuels en ce domaine se porte sur l’inclusion d’*a priori* réalistes permettant de lever les dégénérescences les plus sévères. Une solution par apprentissage automatique est une approche possible.

La décomposition en shapelettes permet de représenter fidèlement l’image d’une galaxie modéremment résolue en quelques dizaines de coefficients, ce qui correspond à une réduction de sa dimensionnalité d’un facteur ≈ 50. Cette “compression”, très rapide (il s’agit d’un ajustement linéaire), doit permettre d’accélérer considérablement un travail de description et de classification morphologique en aval. Les coefficients de shapelettes, ainsi que les composantes principales, sont bien plus aisées à manipuler que les coefficients d’ondelettes des modèles de vision multi-échelle envisagés dans les années 90 (Rué & Bijaoui 1996), même s’ils ne gèrent pas une dynamique d’échelles aussi vaste (moins d’une décade pour quelques dizaines de coefficients). Les premiers travaux publiés dans cette direction, basés sur un modèle à mixture Gaussienne (Kelly & McKay 2004), ou un Perceptron multicouche (de la Calleja & Fuentes 2004) laissent présager d’excellentes possibilités. Compte-tenu de la forte non-linéarité et de la complexité de la transformation des coordonnées shapelettes vers les paramètres physiques et caractéristiques morphologiques, là-aussi une approche par apprentissage automatique apparaît la plus appropriée.

Notre projet s’inscrit dans un contexte international de compétition sur l’exploitation des grands relevés d’imagerie numérique. Autant, sinon plus encore que la réalisation des caméras elles-même, l’analyse des données en détermine l’impact scientifique. Mais en matière d’analyse morphologique automatique, on ne peut pas réellement citer de projet concurrent du nôtre, dans le sens où, à notre connaissance, aucun effort d’envergure n’a été porté sur le sujet. A l’exception de GIM2D (mesure de profils de galaxies, Simard 1998), les réalisations disponibles dans ce domaine sont des prototypes expérimentaux destinés à tester un concept ou à la rigueur pour un usage ponctuel sur des données spécifiques.

Parmi les projets connexes auxquels nos équipes participent l’on peut recenser

– AstroWISE[7] (associé au projet OmegaCam du télescope VST, principalement Pays-Bas-Allemagne-Italie) : des membres de l’équipe TERAPIX sont associés dans le cadre d’un developpement conjoint de certains logiciels pour la visualisation et le contrôle-qualité (logiciels Panorapix et QualityFITS développés à TERAPIX).
– AVO (Astronomical Virtual Observatory) : le projet europée d’observatoire virtuel, destiné à homogénéïser l’échange des données et des méthodes de traitement au sein de la communauté astronomique. TERAPIX fournit entre autres le moteur d’extraction de sources de l’AVO.
– MDA (Masses de Données en Astronomie) : ce projet a permis de financer la moitié du stage d’Anthony Baillard à TERAPIX (identification automatique de défauts étendus sur les images astronomiques).

Parmi les projets de grands relevés d’imagerie auxquels sont associés des efforts de réduction et d’analyse de données citons

– SDSS[8] (Sloan Digital Sky Survey) : projet américano-japonais en cours de complétion, incluant une partie photométrique visible avec une caméra multi-bande opérant en mode transit sur un télescope dédié.
– Vista[9] (Visible and Infrared Telescope for Astronomy) : caméra proche-infrarouge

[7] `http ://www.astro-wise.org/`
[8] `http ://www.sdss.org`
[9] `http ://www.vista.ac.uk`

grand-champ et télescope dédié (principalement Royaume-Uni) destinés au site de Paranal au Chili.
– Pan-Starrs[10], projet américain d'un ensemble de télescopes dédiés et leurs caméras grand-champ opérant de concert à Hawaii. Après le SDSS, représente problablement l'expérience incluant les développements logiciels le plus significatifs.
– LSST[11] (Large-aperture Synoptic Survey Telescope) : projet américain (non-encore financé) d'un télescope de 8m équipé d'une caméra à très grand-champ. Compte-tenu des débits attendus (10 Toctets/nuit), l'analyse des données représente un défi majeur.

B2 – Description du projet : (5 à 10 pages)

Entre autres, le caractère innovant du projet (concepts, technologies, expériences ...) devra être explicité et la valeur ajoutée des coopérations entre les différentes équipes sera discutée.

Notre objectif est de mettre en place un système générique de description et de classification morphologique des images de sources extragalactiques (galaxies) au moyen, notamment, de techniques d'apprentissage automatique. Le projet réunit laboratoires d'astronomie/astrophysique et laboratoires des STIC autour d'un projet qui de par le monde n'a jusqu'ici surtout fait l'objet que d'études exploratoires, et cantonnées essentiellement à la communauté astronomique, comme l'attestent les articles cités dans la bibliographie ci-dessous. L'avalanche récente de masses de données d'imagerie astronomique grand-champ a été le déclencheur pour attaquer le problème de la classification morphologique des galaxies de façon à la fois pragmatique et rigoureuse, en joignant les compétences de spécialistes en analyse d'images, reconnaissance de formes, apprentissage automatique, optimisation d'algorithmes, traitement de données en astronomie et physique des galaxies.

Le système proposé est constitué d'un paquetage logiciel, dont la partie "production", très optimisée, doit pouvoir opérer sur un ensemble distribué de machines, éventuellement hétérogène. Portabilité, distributivité et optimisation sont donc trois exigences fondamentales dans cette partie du projet, pour lesquelles le LRD de l'EPITA apporte son expertise. En matière d'entrées-sorties, l'analyse morphologique d'une galaxie requiert, au plus, l'expédition des pixels d'une vignette centrée sur l'objet et de la réponse impulsionnelle locale vers le noeud de calcul, soit une à quelques dizaines de kilo-octets, et la réception des paramètres morphologiques, soit quelques dizaines d'octets. A quelques centaines de galaxies par secondes au maximum, le système est peu sensible aux latences et requiert une bande passante totale raisonnable, compatible même avec un calcul distribué à l'échelle de plusieurs instituts.

Le logiciel de modèlisation de la réponse impulsionnelle locale, a été développé à TERAPIX (Bertin 2004), et il est déjà opérationnel depuis plusieurs années. Sa qualité et sa robustesse ont été évaluées à maintes reprises (voir par exemple Kalirai et al. 2001). La modélisation prend en compte les réponses impulsionnelles hautement variables et les images sous-échantillonnées. TERAPIX fournit également le logiciel de segmentation d'images et d'extraction de sources produisant les petites vignettes centrées autour de chaque galaxie à analyser (SExtractor : Bertin & Arnouts 1996).

La technique précise de réduction de dimensionnalité retenue (ACP, ACI, shapelettes, combinaison d'entre-elles ou autres) doit être soigneusement évaluée. Le LTCI (GET/Télécom Paris) et le Service d'Astrophysique du CEA/Saclay (A. Refregier est l'inventeur des shapelettes) apportent leurs expertises en ce domaine. Afin de pouvoir être utilisable de façon robuste dans un pipeline de traitement d'image astronomique, il faut que la décomposition des images prenne en compte la qualité relative des pixels des galaxies. TERAPIX fournit les "cartes de poids" associées. A noter que le "pesa-

[10] `http ://panstarrs.ifa.hawaii.edu/public/index.html`
[11] `http ://www.lsst.org`

ge" des pixels affectés de défauts optiques étendus fait déjà l'objet d'un stage utilisant l'apprentissage automatique, par un étudiant EPITA travaillant pour TERAPIX à l'IAP (Baillard 2004, en préparation).

Le paquetage logiciel doit également contenir un module d'apprentissage automatique, parallélisé (mémoire partagée), et optimisé afin d'accélérer la mise-au-point des extracteurs de forme et du classifieur. Le choix est du ressort du LTCI dont l'expertise en ce domaine n'est plus à démontrer. La méthodologie de sélection des descripteurs, non-redondants, pertinents, tirera profit des abondants travaux qui lui ont été consacrés depuis 10 ans (Blum et al. 1997, Guyon et al. 2003), et dans le domaine de l'image (Tieu et al. 2004), en particulier. Comme nous disposerons d'excellents ensembles d'apprentissage, les techniques de *wrappers* couplées à des machines à vecteurs de support (SVM) semblent bien adaptées puisqu'elles s'appuient sur la pertinence des caractéristiques au regard des performances du système final. Nous évaluerons également les possibilités de prendre en compte les réponses impulsionnelles au sein du classifieur lui-même, soit en associant cette information aux descripteurs, soit en agissant directement au niveau de *clusters* intermédiaires.

Du point de vue matériel, le module d'apprentissage sera destiné à être exécuté sur une plateforme performante, multi-processeurs, et possédant une quantité de mémoire suffisante pour stocker les échantillons d'entrainement en mode "batch".

Du côté astronomique, une étape préalable du projet concerne le choix des attributs morphologiques que l'on souhaite voir figurer dans la description produite par l'extracteur de formes. Certains sont des "classiques" déjà répertoriés dans des catalogues historiques comme le RC2[12] (de Vaucouleurs et al. 1976) : présence de bras, d'une barre, d'anneaux, d'un noyau compact, aspect inhabituel... Il est important que l'extracteur de formes produise des grandeurs analogiques pour chaque attribut (par exemple, une échelle de 0 à 5), afin que des corrections de systématiques puissent être apportées après apprentissage. Les équipes du CRA de Lyon (Georges Paturel, Philippe Prugniel) et du Laboratoire d'Astrophysique de l'Observatoire Midi-Pyrénées (P. Fouqué) regroupent parmi les meilleurs experts français en analyse morphologique des galaxies. Elles prendront en charge la fabrication des échantillons de test et de validation, avec des critères de sélection rigoureux au travers de la base Hyperleda (Paturel et al. 2003), et de la cross-identification avec les images numériques des relevés visible SDSS[13] (Strauss et al. 2002) et éventuellement infrarouge 2MASS[14] (Jarrett 2000). Les images de ces galaxies sont bien résolues ; l'étalement de la réponse impulsionnelle y est négligeable aux échelles analysées. Elles sont donc reconvoluables et réinjectables dans des simulations d'images, en variant les conditions d'observation, les échelles, les angles de position, et la contamination par les sources voisines. En outre, des "variations anatomiques" peuvent y être simulées en agissant dans l'espace des shapelettes (Massey et al. 2004). L'équipe TERAPIX dispose de l'outil de simulation nécessaire, SkyMaker[15] pour générer à partir de ces prototypes des images sol et spatial réalistes (voir par exemple Erben et al. 2001). Il faudra prévoir cependant l'inclusion des shapelettes ainsi que la parallélisation dans SkyMaker.

D'autres attributs morphologiques plus "exotiques", mais d'une grande importance astrophysique nécessiteront un travail de classification visuelle, comme les déformations (mirages) par les lentilles gravitationnelles par exemple. En plus des simulations, TERAPIX dispose de plusieurs dizaines de Toctets d'images MEGACAM dans lesquelles ces échantillons pourront être puisés de manière semi-empirique (plusieurs membres de l'équipe inspectent quotidiennement les images dans le cadre de la production). Les données publiques à haute résolution du télescope spatial Hubble seront également mises à contribution (en particulier les relevés Hubble Deep et Ultra Deep).

La dernière étape du projet, avant son exploitation en astrophysique, est la validation

[12] voir `http ://vizier.u-strasbg.fr/cgi-bin/getCatFile ?VII/155/mtypes.tex`

[13] `http ://www.sdss.org/`

[14] `http ://www.ipac.caltech.edu/2mass/`

[15] `http ://terapix.iap.fr/cplt/oldSite/soft/skymaker/`

statistique des résultats et la calibration du système : correction d'effets sytématiques résiduels, parfois subtils, parfois saillants, sur les paramètres issus de l'extracteur de formes et la classification morphologique. Ces effets ont de multiples origines : dépendance avec la qualité d'image, l'inclinaison de l'axe de la galaxie par rapport à la ligne de visée, détectabilité variable selon le type morphologique (dépendance avec la brillance de surface), contamination par des compagnons, des étoiles d'avant-plan, ou des défauts optiques,... La majeure partie de ce travail se fera au moyen de simulations d'images. Grâce à l'expertise de l'IAP (Valérie de Lapparent) dans ce domaine, la corrélation avec le type spectroscopique devrait permettre de lever certaines ambiguités en apportant un éclairage physique aux résultats (Galaz & de Lapparent 1998, de Lapparent 2003) pour les objets pour lesquels un spectre est disponible.

Un point essentiel est la dépendence de la classification morphologique avec la bande spectrale dans laquelle est observée la galaxie : plus on va vers le bleu, plus les structures des disques (bras spiraux, régions de formation stellaire, bande de poussières) deviennent proéminentes, et plus la galaxie semble irrégulière. Une approche panchromatique de la morphologie est donc indispensable. Outre les 5 bandes en visible offertes par les images du SDSS (galaxies proches) et MEGACAM (galaxies lointaines), nous disposerons d'image proche-infrarouge des relevés 2MASS et DeNIS (galaxies proches) et de la caméra WIRCAM (galaxies lointaines). Les images WIRCAM seront traitées à TERAPIX, et l'équipe de Toulouse (Jean-François Le Borgne, Roser Pelló), responsable scientifique de l'expérience, apportera son expertise dans ce domaine.

En raison de l'expansion de l'Univers, la lumière des objets distants est émise à des longueurs d'onde plus bleues que celles reçues. Ainsi, une caméra comme MEGACAM observe dans le visible la morphologie ultra-violette des galaxies lointaines. Pour pouvoir comparer les formes (passées) des galaxies lointaines avec leurs formes actuelles, il faut donc à la fois disposer d'un échantillon de galaxies proches en ultra-violet, et connaître au moins grossièrement les distances individuelles des objets lointains. L'équipe GALEX du Laboratoire d'Astronomie de Marseille (Stéphane Arnouts, Bruno Millard) fournira les images en ultra-violet moyen, calibrées par leurs soins, de plusieurs milliers de galaxies proches issues du relevé MIS, pour lesquels on disposera des images visibles du SDSS. Les décalages spectraux (distances) des galaxies, seront quant-à-eux obtenus grâce aux logiciels développés à l'Observatoire Midi-Pyrénées (Bolzonello et al. 2002) et à Marseille (Arnouts et al. 1999).

L'aboutissement du projet sera la mise-à-disponibilité par internet pour l'ensemble de la communauté :

– D'un paquetage logiciel "libre" totalement portable (source sous forme `.tar.gz`, et binaires Linux `.rpm`/`.deb`). Le paquetage contiendra à la fois le module d'apprentissage parallélisé, le module de calcul distribué "serveur de morphologie/extracteur de formes", et la documentation attenante.
– du ou des fichiers de paramètres du classifieur issus de l'apprentissage automatique après calibration et validation, et destinés à être utlisés par l'extracteur de formes.
– des données images et morphologie issues du travail de collecte, d'homogénéisation et de classification visuelle des astronomes pour entrainer et valider le module d'apprentissage du classifieur.

Le LRDE et TERAPIX ont une longue expérience de distribution et maintenance de logiciels "libres"[16] [17]. Akim Demaille du LRDE est notamment le mainteneur d'`autoconf`, le logiciel d'auto-configuration utilisé dans la plupart des paquetages logiciels portables modernes. TERAPIX délivre et maintient déjà une dizaine de logiciels libres portables largement utilisés au niveau mondial par la communauté astronomique pour le traitement, la calibration et l'analyse des images.

[16]http ://www.lrde.epita.fr/cgi-bin/twiki/view/Projects/WebHome
[17]http ://terapix.iap.fr/rubrique.php ?id_rubrique=34

Enfin il faut souligner les aspects d'enseignement et de formation contenus dans le projet EFIGI. En complément du poste CDD d'ingénieur de recherche de 24 mois demandé, la portée scientifique d'EFIGI et sa visibilité nous inciteront à proposer une thèse BDI en co-tutelle LTCI et IAP pour accompagner le projet. Nous avons également identifié 6 stages répartis à la fois sur les versants astronomie et STIC, à différentes étapes de réalisation de l'ensemble du programme (voir ci-dessous).

B3 – Résultats attendus :

On détaillera l'échéancier des résultats et réalisations intermédiaires et finaux attendus. On précisera les risques scientifiques qui seront pris. On discutera de l'impact potentiel de ce projet sur les scènes européenne et internationale.

L'échéancier provisionnel est présenté Fig. 1. Le rectangle violet représente la durée du CDD de 24 mois et l'éventail des activités auxquelles il participe.

Etapes *complétées* à 3 mois, avec indication des partenaires les plus concernés :
- Pré-selection des attributs morphologiques et définition des spécifications — tous les partenaires.
- Constitution d'un premier pré-échantillon sommaire de galaxies bien résolues — TERAPIX.
- 1ère Mise-à-jour du logiciel de simulation SkyMaker : parallélisation et support des vignettes d'entrée (avec angle de position et échelle variables) — TERAPIX + LRDE.

à 6 mois :
- 2ème Mise-à-jour du logiciel de simulation SkyMaker : support d'une base de fonctions arbitraire — TERAPIX + LRDE.

à 9 mois :
- Constitution de l'échantillon d'images de "galaxies larges" bien résolues — CRAL (stagiaire) + LAT).
- Revue de projet à 6 mois sur létude en cours de la déconvolution paramétrique — LTCI + SAp + TERAPIX. Si résultats peu encourageants, basculement sur correction *a posteriori* de la réponse impulsionnelle.

à 15 mois :
- Extension GALEX de l'échantillon de galaxies bien résolues — LAM (stagiaire) + CRAL + LAT.
- Constitution de l'échantillon d'images de "galaxies faibles" mal résolues — TERAPIX.
- Choix de la base de fonctions finalisé — SAp + LTCI + TERAPIX.
- Principe de correction de réponse impulsionnelle finalisé — LTCI + SAp + TERAPIX.
- Revue de projet à 12 mois du module d'apprentissage du classifieur — LTCI + TERAPIX.

à 18 mois :
- Prototype du module d'apprentissage du classifieur — LTCI.
- Serveur de morphologie — LRDE (stagiaire).

à 21 mois :
- Extension WIRCAM de l'échantillon d'images de "galaxies faibles" mal résolues — LAT (stagiaire).
- Optimisation du serveur de morphologie — LRDE.

à 24 mois :
- Parallélisation et optimisation du module d'apprentissage du classifieur — LRDE (stagiaire).

à 30 mois :
- Apprentissage, test et validation du classifieur — TERAPIX (stagiaire) + LTCI.
- Packaging de la suite logicielle — LRDE + TERAPIX.
- Version bêta délivrée à la communauté astronomique.

à 33 mois :
- Calibration du classifieur — IAP.
- Finalisation de la documentation de la suite logicielle — tous les partenaires.
- Mise en ligne du paquetage EFIGI.

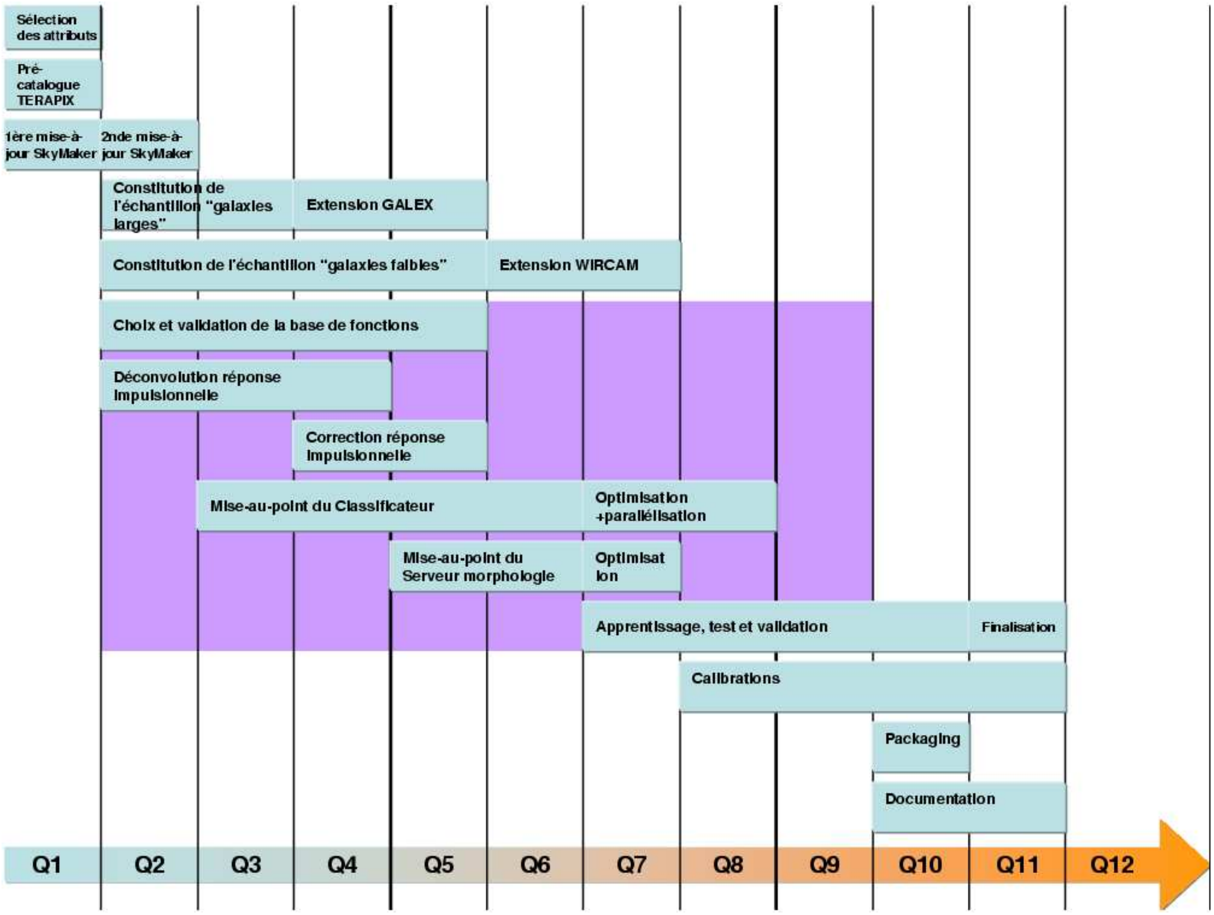

FIG. 1 – Echéancier du projet.

Les retombées de ce projet sur les scènes européenne et internationale sont doubles :

– Au niveau européen d'abord, le concept de "serveur de morphologie" pourra être très facilement intégré, en tant que méthode, au concept AVO de l'Observatoire Virtuel. Il en va de même pour la description morphologique des données de l'échantillon d'apprentissage. TERAPIX, participant AVO, veillera durant toutes les phases de développement d'EFIGI, à la conformité des choix pris avec les recommendations d'AVO.
– Au niveau mondial, la large disponibilité de cet outil d'un type inédit, devrait fournir pour la première fois à la communauté astronomique un ensemble complet, performant et surtout standard d'analyse morphologique des masses de données d'images. La relative insensibilité d'EFIGI aux caractéristiques précises des images devrait en particulier être bénéficiaire aux travaux impliquant des données de qualité hétérogène.

B4 – Summary (in English) : (1 to 2 pages)

Le Conseil Scientifique pourra solliciter des experts non francophones auxquels sera envoyé l'ensemble des documents. Le présent résumé, entièrement rédigé en anglais, visera à fournir une présentation synthétique de l'ensemble du projet.

In terms of total amount of data, modern astronomy is dominated by survey imaging data produced by wide-field digital cameras. After two decades of development, these cameras, constructed from mosaics of many individual charge-coupled devices, can now cover an area equivalent or larger with respect to their photographic predecessors. Moreover, unlike photographic plates, digital cameras have excellent sensitivity over a wide wavelength range and respond linearly to the incident light. Rather than directly using this vast quantity of pixel data, statistical observational cosmology is mostly carried out using catalogues in which galaxies are described only by positions, brightnesses and colours, and perhaps one or two simple shape parameters such as ellipticity or concentration index ; a vast amount of useful information is therefore discarded. The morphology of galaxies can provide a key insight into their formation histories but until now, extracting useful morphological information for a large (10^9) numbers of galaxies has been extremely challenging due to both computational and algorithmic difficulties ; existing tools are all either too slow or provide output insufficiently detailed to fulfill our goal of a complete description of each galaxy using a dozen or so parameters. We propose to address both the computational and algorithmic aspects of this problem and to develop a robust and scalable solution to measure galaxy morphologies for large numbers of objects in very large imaging surveys.

Solving this problem will require the skills of signal processing specialists, computer scientists and astronomers. In algorithmic terms, several promising avenues will be developed, either machine learning techniques trained on local data sets of large, well-resolved galaxies, or complementary approaches which attempt to reduce the dimensionality of galaxy images by describing them in terms of a set of discrete parameters. TERAPIX, thanks to its involvement in several generations of wide-field imaging surveys, has at its disposal several Terabytes of data suitable for performing this task. We will also address systematic errors, selection effects and band-shifting effects which have plagued this work in the past by carrying out extensive simulations and using data spanning from the UV (GALEX) to the infrared (WIRCAM). Following the "open source" philosophy, we will provide the tools we develop free of charge for use by the entire astronomical community (both LRDE and TERAPIX have a long experience in this field). Our software will be designed to operate in a distributed computing environment so that more processing power can easily be added to deal with even larger surveys. The group at EPITA already has considerable expertise in distributed computing solutions.

The combination of expertise at LCTI, LRDE and TERAPIX provides an unique opportunity to make a substantial contribution to the problem of determining in a robust way the morphological properties of galaxies in very large imaging surveys. Although several large surveys similar in size or larger than those managed at TERAPIX currently exist or are in planning (such as SDSS, Pan-Starss), none has yet addressed the difficult task of morphological classification in a complete and reliable manner ; that is our objective.

Note : la plupart des papiers récents sont récupérables sous forme de *preprint* en Postscript et PDF à `http ://arxiv.org/form/astro-ph?MULTI=form+interface`

Références

[1] Abraham R.G., van den Bergh S., Nair P., 2003 *The Astrophysical Journal* **588**, 218

[2] Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, *Monthly Notices of the Royal Astronomical Society* **310**, 540

[3] Bazell D., Aha D.W., 2001, *The Astrophysical Journal*, **548**, 219

[4] Bershady M.A., Jangren A, Conselice C.J., 2000, *The Astronomical Journal*, **119**, 2645

[5] Bernstein G.M., Jarvis M., 2002, *The Astronomical Journal* **123**, 583

[6] Bertin E., Arnouts S., 1996, *Astronomy &Astrophysics Supplements* **117**, 393

[7] Bertin E., 1999, *SExtractor 2.1, User's manual*, IAP

[8] Bertin E., Thion A., Mellier Y., van Waerbeke L., 2001, in *Gravitational Lensing : Recent Progress and Future Goals*, ASP Conference Proceedings, Vol. **237**. Brainerd T.G. & Kochanek C.S. eds., 365

[9] Bertin E., 2004, in *Mathematical Challenges in Astronomical Imaging*, IPAM workshop `http ://www.ipam.ucla.edu/publications/ai2004/ai2004_4209.ppt`

[10] Blum A.L., Langley P., 1997, "Selection of Relevant Features and Examples in Machine", Elsevier Science P. **97**, 273

[11] Bolzonella M., Pelló R., Maccagni D., 2002, *Astronomy and Astrophysics* **395**, 443

[12] Burrus N., Duret-Lutz A., Géraud T., Lesage D., Poss R., 2003, *Proceedings of the Workshop on Multiple Paradigm with OO Languages (MPOOL'03)*

[13] de la Calleja J., Fuentes O., 2004, *Monthly Notices of the Royal Astronomical Society* **349**, 87

[14] Cardoso J.-F., Delabrouille J., Patanchon G., 2003, in *International Symposium on Independent Component Analysis and Blind Source Separation (ICA)*, Nara, Japan, 1111

[15] Christopher J.C., Bershady M.A., 2000, *The Astrophysical Journal*, **529**, 886

[16] de Lapparent V., 2003, "Critical analysis of the luminosity functions per galaxy type measured from redshift surveys", *Astronomy and Astrophysics* **408**, 845

[17] de Vaucouleurs G., de Vaucouleurs A., Corwin H.G., 1976, *Second Reference Catalogue of bright galaxies (RC2)*, University of Texas Press, Austin

[18] de Vaucouleurs G., de Vaucouleurs A., Corwin H.G., Buta R.J., Paturel G. and Fouqué P., 1991, *Third Reference Catalogue of bright galaxies (RC3)*, Springer-Verlag, New-York

[19] Erben T., Van Waerbeke L., Bertin E., Mellier Y., Schneider P., 2001, *Astronomy &Astrophysics* **366**, 717

[20] Galaz G., de Lapparent V., 1998, "The ESO-Sculptor Survey : spectral classification of galaxies with $z < 0.5$", *Astronomy and Astrophysics* **332**, 459

[21] Guyon I., Elisseeff A., 2003, "An introduction to variable and feature selection", *J. of Machine Learning Research* **3**, 1157

[22] Jarrett T.H., 2000, *Publications of the Astronomical Society of the Pacific*, **112**, 774 et `http ://www.ipac.caltech.edu/2mass/gallery/galmorph/2mass_galmorp.html`

[23] Kalirai J.S., Richer H.B., Fahlman G.G., Cuillandre J.-C., Ventura P., D'Antona F., Bertin E., Marconi G., Durrell P.R., 2001, *The Astronomical Journal* **122**, 257

[24] Kelly B.C., McKay T.A., 2004, *The Astronomical Journal*, **127**, 625

[25] Massey R., Refregier A., Conselice C.J., Bacon D.J., 2004, *Monthly Notices of the Royal Astronomical Society* **348**, 214

[26] Metropolis N. et al., 1953, *Journal of Chemical Physics*, **21**, 1087

[27] Naim A., Lahav O., Sodré L. Jr, Storrie-Lombardi M.C., 1995, *Monthly Notices of the Royal Astronomical Society* bf 275, 567

[28] Odewahn S.C., Windhorst R.A., Driver S.P., Keel W.C., 1996, *Astrophysical Journal Letters* **472**, L13

[29] Owens E.A., Griffiths R.E., Ratnatunga K.U., 1996, *Monthly Notices of the Royal Astronomical Society* **281**, 153

[30] Paturel G., Petit C., Prugniel P., Theureau G., Rousseau J., Brouty M., Dubois P., Cambrésy L., 2003, *Astronomy and Astrophysics*, **412**, 45

[31] Refregier A., 2003, *Monthly Notice of the Royal Astronomical Society* **338**, 35

[32] Rué F., Bijaoui A., 1996, *Vistas in Astronomy* **40**, 495

[33] Simard L., 1998, in *Astronomical Data Analysis Software and Systems VII*, A.S.P. Conference Series, Vol. **145**, Albrecht R., Hook R.N., Bushouse H.A. eds., 108

[34] Storrie-Lombardi M. C., Lahav O., Sodré L. Jr, Storrie-Lombardi L.J., 1992, *Monthly Notices of the Royal Astronomical Society* **259**, 8

[35] Strauss M.A., Weinberg D.H., Lupton R.H. et al., 2002, *The Astronomical Journal* **124**, 1810

[36] Szalay A.S., Connolly A.J., Szokoly G.P., 1999, *The Astronomical Journal* **117**, 68

[37] Tieu K., Viola P., 2004, "Boosting Image Retrieval", *Int. J. Comp. Vision* **56**, 17

[38] Wolberg G., 1992, *Digital Image Warping* (IEEE Computer Society Press)

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

C – MOYENS FINANCIERS ET HUMAINS DEMANDÉS PAR CHAQUE ÉQUIPE[18]

Comme indiqué dans les tableaux ci-dessous, on distinguera

- *les financements via le Fonds National pour la Science qui peuvent inclure*
  - *du fonctionnement*
  - *de l'équipement*
  - *des mois de personnel temporaire (CDD) pour un montant ne pouvant excéder 50% du financement total attribué. La durée du ou des contrat(s) prévus, qui ne peuvent excéder 24 mois chacun, sera précisée.*
- *les moyens demandés aux organismes de recherche qui peuvent inclure*
  - *des postes de post-doc*
  - *des demandes de délégation ou détachement pour des enseignants-chercheurs*
  - *des accueils de chercheurs étrangers*
- *les demandes d'allocations de recherche*

Les diverses possibilités concernant l'attribution de moyens pour recruter ou accueillir des personnels seront globalement très limitées pour l'ensemble des ACI. Leurs demandes devront donc être particulièrement justifiées. Si les bénéficiaires de ces demandes sont connus ou pressentis, les CV correspondants seront joints à la présente demande.

Dans le cas des moyens alloués par les organismes, il n'est pas nécessaire de préciser à quel organisme (CNRS ou INRIA) ces moyens seront demandés, sauf cas particulier à expliciter. Ces moyens seront en effet répartis globalement au niveau de l'ACI, en tenant compte bien sûr des règles et contraintes propres à chaque organisme.

On présentera une justification scientifique des moyens demandés pour chacune des équipes impliquées dans le projet.

[18]Une fiche C doit être remplie pour chaque laboratoire ou équipe partenaire.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

C1 - Demandes effectuées dans le cadre de l'ACI pour le présent projet :

Nom de l'équipe ou du laboratoire : TERAPIX (IAP)

Moyens demandés dans le cadre de la présente ACI (en K euros TTC) :

Dans les tables suivantes, les missions portent le label (V), le divers fonctionnement le label (D) et les stagiaires le label (S).

Le salaire du CDD est donné en TTC et représente le coût moyen d'un IR2.

- 1 mini-cluster de 12 processeurs, type quadriprocesseurs, 8 Go de mémoire (30 Keuros),
- 11 missions à 1 keuro chacune,
- Un CDD IR en développement logiciel pendant 2 ans (IR2 cout moyen 60 Keuro/an).
- Un stagiaire de 6 mois (2006) — que les écoles EPITA ou Telecom-Paris peuvent fournir.

Financements via le Fonds National de la Science :

	2004	2005	2006	Total
Équipement	-	30	-	30
Fonctionnement (dont CDD décrits ci-dessous)	3(V)+1(D) +15(CDD)	5(V)+1(D) +60(CDD)	5(V)+1(D)+10(S) +45(CDD)	146
Total / année	19	96	61	176

Dépenses de personnels (CDD)[19] :

Nature de l'emploi (post-doc, ingénieur, assistant-ingénieur,...)	Ingénieur IR2 en développement logiciel
Durée de l'emploi (en mois)[20]	24
Coût total de l'emploi	120 Keuros

Financements via les organismes de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre de post-docs (préciser pour chaque demande la durée en mois)[21]				
Nombre d'accueils de chercheurs étrangers (préciser pour chaque demande la durée en mois)				
Nombre d'acceuils en délégations ou détachements[22]				

[19] Un tableau doit être rempli pour chaque demande de CDD.
[20] Doit être inférieure à 24 mois.
[21] Sauf demande argumentée, la durée d'un contrat de type post-doc ne pourra excéder 12 mois.
[22] Certaines des demandes déjà faites pour 2003-2004 pourront être attribuées au titre de l'ACI.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Allocations de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre d'allocations de recherche débutant en :				

Justifications scientifiques de l'ensemble des demandes : La justification de la demande est la suivante :

- Les voyages sont pour le suivi du projet, les conseils et la synchronisation des tâches du PI (E. Bertin) et de H.J. McCracken et Y. Mellier sur les sites de Toulouse, Lyon et Marseille.
- Mini-cluster : pour les apprentissages parallélisés en mode batch (voir section B2). F. Magnard en sera le responsable informatique.
- Le stagiaire sera affecté au travail d'apprentissage et de validation du classifieur. Le LTCI d'un part et le LDRE, par le biais de l'école EPITA, d'autre part, vont fournir un réservoir de stagiaires dont la formation correspond exactement à nos besoins : nous insistons particulièrement sur le rôle de formation de notre projet.
- Le CDD, de niveau I.R.2, sera affecté aux tâches de développement et de validation des logiciels sur le site TERAPIX. Son travail comportera des phases de prototypage, de mise-au-point et de test du classifieur et du serveur de morphologie sur les échantillons de galaxies à la fois "étendus" et "faibles" (voir texte), et enfin de la documentation du code.

En ce qui concerne les missions demandées pour les équipes de Lyon, Marseille et Toulouse, il s'agit pour l'essentiel des 3 grandes revues annuelles : un Kick-off meeting et deux revues annuelles, une un an après le Kick off, l'autre deux ans après. S'ajoutent ensuite les ateliers ciblés :

- Un concernant la construction de l'échantillon de galaxies larges
- Un concernant la construction de l'échantillon de galaxies faibles
- Deux concernant la base de fonctions
- Deux concernant la déconvolution et la correction de la réponse impulsionnelle
- Deux concernant la mise au point du classifieur
- Un concernant la mise au point du serveur morphologique
- Deux concernant l'apprentissage et la validation du serveur.

Avec les trois grandes réunions, nous envisageons donc 15 meetings sur trois ans. Dans la mesure du possible, ils seront regroupés.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Nom de l'équipe ou du laboratoire : LTCI

Moyens demandés dans le cadre de la présente ACI (en K euros TTC) : Les moyens se composent de :

- 8 missions à 1 keuro chacune
- 1 stagiaire de 6 mois (2005)
- A noter que le LTCI utilisera les machines demandées par TERAPIX

Financements via le Fonds National de la Science :

	2004	2005	2006	Total
Équipement	-	-	-	0
Fonctionnement (dont CDD décrits ci-dessous)	2(V)+1(D)	3(V)+1(D)+10(S)	3(V)+1(D)	21
Total / année	3	14	4	21

Dépenses de personnels (CDD)[23] :

Nature de l'emploi (post-doc, ingénieur, assistant-ingénieur,...)	
Durée de l'emploi (en mois)[24]	
Coût total de l'emploi	

Financements via les organismes de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre de post-docs (préciser pour chaque demande la durée en mois)[25]				
Nombre d'accueils de chercheurs étrangers (préciser pour chaque demande la durée en mois)				
Nombre d'acceuils en délégations ou détachements[26]				

[23] Un tableau doit être rempli pour chaque demande de CDD.
[24] Doit être inférieure à 24 mois.
[25] Sauf demande argumentée, la durée d'un contrat de type post-doc ne pourra excéder 12 mois.
[26] Certaines des demandes déjà faites pour 2003-2004 pourront être attribuées au titre de l'ACI.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Allocations de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre d'allocations de recherche débutant en :				

Justifications scientifiques de l'ensemble des demandes : La justification de la demande est la suivante :

- Les voyages serviront à fournir les explications détaillées des besoins du classifieur et ses spécifications aux partenaires chargés de construire les échantillons d'apprentissage et de définir les attributs.
- Le stagiaire sera affecté au serveur de morphologie du classifieur.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Nom de l'équipe ou du laboratoire : LRDE

Moyens demandés dans le cadre de la présente ACI (en K euros TTC) : Les moyens se composent de :

- 1 machine type quadriprocesseurs (10 Keuros).
- 8 missions à 1 keuro chacune
- Un stagiaire de 6 mois (2005).

Financements via le Fonds National de la Science :

	2004	2005	2006	Total
Équipement	-	10	-	10
Fonctionnement (dont CDD décrits ci-dessous)	2(V)+1(D)	3(V)+1(D)+10(S)	5(V)+1(D)	21
Total / année	3	24	4	31

Dépenses de personnels (CDD)[27] :

Nature de l'emploi (post-doc, ingénieur, assistant-ingénieur,...)	
Durée de l'emploi (en mois)[28]	
Coût total de l'emploi	

Financements via les organismes de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre de post-docs (préciser pour chaque demande la durée en mois)[29]				
Nombre d'accueils de chercheurs étrangers (préciser pour chaque demande la durée en mois)				
Nombre d'acceuils en délégations ou détachements[30]				

[27] Un tableau doit être rempli pour chaque demande de CDD.
[28] Doit être inférieure à 24 mois.
[29] Sauf demande argumentée, la durée d'un contrat de type post-doc ne pourra excéder 12 mois.
[30] Certaines des demandes déjà faites pour 2003-2004 pourront être attribuées au titre de l'ACI.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Allocations de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre d'allocations de recherche débutant en :				

Justifications scientifiques de l'ensemble des demandes : La justification de la demande est la suivante :

- Les voyages serviront à fournir les explications détaillées du rôle du serveur de morphologie et de ses spécifications aux partenaires astronomes utilisateurs du système.
- Le stagiaire sera affecté au travail de développement du serveur de morphologie.
- La machine quadri-processeur possède les mêmes performances que celle installée à TERAPIX mais est utilisée de manière indépendante pour le développement en local.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Nom de l'équipe ou du laboratoire : LAM Marseille

Moyens demandés dans le cadre de la présente ACI (en K euros TTC) : Les moyens se composent de :
- 13 missions à 1 keuro chacune
- Un stagiaire de 6 mois (2005).

Financements via le Fonds National de la Science :

	2004	2005	2006	Total
Équipement	-	-	-	0
Fonctionnement (dont CDD décrits ci-dessous)	3(V)+1(D)	5(V)+1(D)+10(S)	5(V)+1(D)	26
Total / année	4	16	6	26

Dépenses de personnels (CDD)[31] :

Nature de l'emploi (post-doc, ingénieur, assistant-ingénieur,...)	
Durée de l'emploi (en mois)[32]	
Coût total de l'emploi	

Financements via les organismes de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre de post-docs (préciser pour chaque demande la durée en mois)[33]				
Nombre d'accueils de chercheurs étrangers (préciser pour chaque demande la durée en mois)				
Nombre d'acceuils en délégations ou détachements[34]				

[31] Un tableau doit être rempli pour chaque demande de CDD.
[32] Doit être inférieure à 24 mois.
[33] Sauf demande argumentée, la durée d'un contrat de type post-doc ne pourra excéder 12 mois.
[34] Certaines des demandes déjà faites pour 2003-2004 pourront être attribuées au titre de l'ACI.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Allocations de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre d'allocations de recherche débutant en :				

Justifications scientifiques de l'ensemble des demandes : La justification de la demande est la suivante :

– Les voyages servivont à la participation aux Kick-off meeting et aux revues annuelles (à Paris) ainsi qu'ła participation aux ateliers ciblés sur les tâches du LAM (GALEX, décalages spectraux photométrique) et aux réunions d'expertise (classifieur, correction de réponse impulsionnelle des données GALEX, apprentissage).

– Le stagiaire sera affecté aux tâches de construction de l'extension GALEX des échantillons de base de galaxies bien résolues, ainsi qu'à la mise en place des catalogues utilisant la morphologie et la distribution spectrale d'énergie des galaxies (UV, optique, infra-rouge).

Action Concertée Incitative
MASSES DE DONNÉES
Descriptif complet du projet

Nom de l'équipe ou du laboratoire : LAT Toulouse

Moyens demandés dans le cadre de la présente ACI (en K euros TTC) : Les moyens se composent de :

- 1 machine type quadriprocesseurs pour apprentissage en local et classification combinée photo-z/Morphologie (10 Keuros).
- 13 missions à 1 keuro chacune
- Un stagiaire de 6 mois (2005)

Financements via le Fonds National de la Science :

	2004	2005	2006	Total
Équipement	-	10	-	10
Fonctionnement (dont CDD décrits ci-dessous)	3(V)+1(D)	5(V)+1(D)+10(S)	5(V)+1(D)	26
Total / année	4	26	6	36

Dépenses de personnels (CDD)[35] :

Nature de l'emploi (post-doc, ingénieur, assistant-ingénieur,...)	
Durée de l'emploi (en mois)[36]	
Coût total de l'emploi	

Financements via les organismes de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre de post-docs (préciser pour chaque demande la durée en mois)[37]				
Nombre d'accueils de chercheurs étrangers (préciser pour chaque demande la durée en mois)				
Nombre d'acceuils en délégations ou détachements[38]				

[35] Un tableau doit être rempli pour chaque demande de CDD.
[36] Doit être inférieure à 24 mois.
[37] Sauf demande argumentée, la durée d'un contrat de type post-doc ne pourra excéder 12 mois.
[38] Certaines des demandes déjà faites pour 2003-2004 pourront être attribuées au titre de l'ACI.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Allocations de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre d'allocations de recherche débutant en :				

Justifications scientifiques de l'ensemble des demandes : La justification de la demande est la suivante :

– Les voyages servivont à la participation aux Kick-off meeting et aux revues annuelles (à Paris) ainsi qu'la participation aux ateliers ciblés sur les tâches du LAT (WIRCAM, décalages spectraux photométriques) et aux réunion d'expertise (base de fonctions, déconvolution/correction de la réponse impulsionnelle des données WIRCAM, apprentissage des données multi-longeur d'onde).
– Le stagiaire sera affecté aux tâches de construction des échantillons de base de galaxies faibles et lointaines, à l'apprentissage en local et au calcul des décalages spectraux photométriques.
– La machine servira à l'apprentissage en local conjointement au calcul des décalages spectraux photométriques.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

Nom de l'équipe ou du laboratoire : Obs. Lyon/CRAL

Moyens demandés dans le cadre de la présente ACI (en K euros TTC) : Les moyens se composent de :
- 13 missions à 1 keuro chacune
- Un stagiaire de 6 mois (2005).

Financements via le Fonds National de la Science :

	2004	2005	2006	Total
Équipement	-	-	-	0
Fonctionnement (dont CDD décrits ci-dessous)	3(V)+1(D)	5(V)+1(D)+10(S)	5(V)+1(D)	26
Total / année	4	16	6	26

Dépenses de personnels (CDD)[39] :

Nature de l'emploi (post-doc, ingénieur, assistant-ingénieur,...)	
Durée de l'emploi (en mois)[40]	
Coût total de l'emploi	

Financements via les organismes de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre de post-docs (préciser pour chaque demande la durée en mois)[41]				
Nombre d'accueils de chercheurs étrangers (préciser pour chaque demande la durée en mois)				
Nombre d'acceuils en délégations ou détachements[42]				

[39] Un tableau doit être rempli pour chaque demande de CDD.
[40] Doit être inférieure à 24 mois.
[41] Sauf demande argumentée, la durée d'un contrat de type post-doc ne pourra excéder 12 mois.
[42] Certaines des demandes déjà faites pour 2003-2004 pourront être attribuées au titre de l'ACI.

Allocations de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre d'allocations de recherche débutant en :				

Justifications scientifiques de l'ensemble des demandes : La justification de la demande est la suivante :

- Les voyages servivont à la participation aux Kick-off meeting et aux revues annuelles (à Paris) ainsi qu'à la participation aux ateliers ciblés sur les tâches du CRAL et aux réunion d'expertise (classificateur, apprentissage).
- Le stagiaire sera affecté à la construction des échantillons de galaxies proches.

C2 - Autres soutiens financiers apportés au projet :

On mentionnera les autres actions relatives au projet dans lesquelles l'équipe ou le laboratoire est engagé (projets européens, RNRT, RNTL, autres ACI, . . .).

Le projet s'inscrit dans un cadre plus vaste visant à développer et à proposer à la communauté astronomique européenne des outils d'exploitation scientifique des images astronomiques produits par les grands relevés multi-longeur d'onde sol et spatial. TERAPIX a été créé pour produire ces outils, pour délivrer des images calibrées des caméras MEGACAM et WIRCAM et pour aider la communauté française à exploiter efficacement toutes ces données. Le projet EFIGI utilisera donc des ressources fournies par l'INSU/SDU, le Programme National de Cosmologie, et l'IAP. Il s'agit notamment de 3 astronomes, 3 ingénieurs permanents et d'un CCD INSU.

Par ailleurs, TERAPIX est un noeud des réseaux européens RTD du 5ième PCRD AstroWise et AVO. Ces deux contrats financent deux ingénieurs CDD. L'un est spécialisé dans le développement logiciel pour l'analyse des données des caméras MEGACM du CFHT et OMEGACAM du VST ; l'autre est spécialisé dans le développement de base de données et d'outils AVO. Ces ingénieurs prendront part aux développement d'EFIGI.

D - RÉCAPITULATIF GLOBAL DES DEMANDES DU PROJET :

Financements via le Fonds National de la Science :

	2004	2005	2006	Total
Équipement	0	50	0	50
Fonctionnement (dont CDD décrits ci-dessous)	37	122	107	266
Total / année	37	192	87	316

Dépenses de personnels (CDD) :

Nature de l'emploi (post-doc, ingénieur, assistant-ingénieur,...)	Ingénieur IR2 en développement logiciel.
Durée de l'emploi (en mois)	24
Coût total de l'emploi	120

Financements via les organismes de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre de post-docs (préciser pour chaque demande la durée en mois)[43]				
Nombre d'accueils de chercheurs étrangers (préciser pour chaque demande la durée en mois)				
Nombre d'acceuils en délégations ou détachements				

Allocations de recherche :

	2003–2004	2004–2005	2005–2006	Total
Nombre d'allocations de recherche débutant en :				

[43] Sauf demande argumentée, la durée d'un contrat de type post-doc ne pourra excéder 12 mois.

Action Concertée Incitative

MASSES DE DONNÉES

Descriptif complet du projet

E - ENGAGEMENT DU COORDINATEUR DU PROJET :

La présente page ne sera remplie que dans la version sous forme papier.

Je soussigné, **Emmanuel Bertin**, coordinateur du projet **EFIGI**, m'engage dans l'hypothèse où le présent projet serait retenu à :

- fournir un rapport d'évaluation à mi-parcours permettant au Conseil Scientifique d'apprécier l'avancement des travaux et la coopération des équipes participantes,
- un rapport à la fin de l'exécution du projet,
- maintenir régulièrement une page web résumant l'ensemble des activités du projet.

Signature du coordinateur du projet :

Visa du Directeur du Laboratoire ou de l'Unité de Recherche auquel appartient le coordinateur du projet :