



Extraction de Formes Idéalisées de Galaxies en Imagerie

E. Bertin^{1,2}, A. Baillard¹, E. Mellier^{1,2}, V. de Lapparent¹, M. Arroyo³, M. Campedel³, H. Maitre³, R. Dehak⁴, T. Géraud⁴, O. Ricou⁴, P. Fouqué⁵, J.-F. Leborgne⁵, R. Pelló⁵, S. Arnouts⁶, O. Ilbert⁶, L. Tasca⁶, Ph. Prugniel⁷, G. Paturel⁷, A. Refregier⁸

¹Institut d'Astrophysique de Paris, ²Observatoire de Paris, ³Laboratoire TCI de l'ENST (Paris), ⁴Laboratoire de Recherche De l'EPITA (Paris), ⁵Observatoire Midi-Pyrénées, ⁶Laboratoire d'Astronomie de Marseille, ⁷Centre de Recherche Astronomique de Lyon, ⁸Service d'Astrophysique du CEA (Saclay)

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 cœur de questions fondamentales de l'astrophysique moderne (morphogénèse, 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. Le projet EFIGI associe des laboratoires ayant des expertises complémentaires (apprentissage automatique, calcul distribué et optimisation du côté STIC, traitement/analyse des images du ciel et morphologie des galaxies du côté astronomique) autour de données panchromatiques sol et spatial, afin de créer un système générique et performant d'analyse de la morphologie des galaxies.

La morphologie des galaxies

Depuis les années 30, les astronomes ont pris l'habitude de classer les galaxies proches en fonction de leur position sur une échelle morphologique dite "séquence de Hubble". Cependant cette représentation ne traduit pas certaines caractéristiques importantes du point de vue physique, comme par exemple la présence de bandes de poussières ou de spots de formation stellaire; elle semble par ailleurs devenir caduque pour les galaxies les plus lointaines (c'est à dire plus dans leur passé lointain). Un jeu de 18 attributs morphologiques en plus du type *T* de Hubble a donc été retenu par les astronomes d'EFIGI.

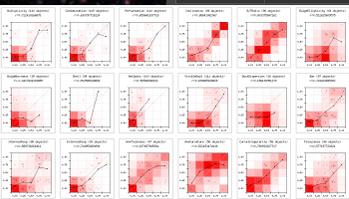
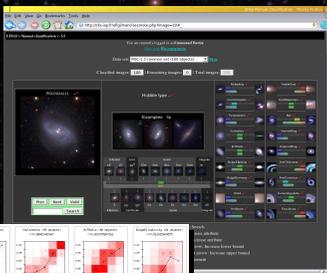


Les données de référence

Alors que l'on recense plus d'un million de galaxies dont les images sont bien résolues depuis le sol, il n'existe pas de large base de données homogène contenant un jeu étendu d'informations morphologiques détaillées.

L'acquisition d'un jeu homogène d'images multi-canaux dans le relevé SDSS de 4500 galaxies bien résolues (2To au total) du catalogue PGC (Paturel et al. 2003) et leur prétraitement en vue d'un étiquetage dans des conditions contrôlées ont donc été parmi les premières priorités d'EFIGI.

L'étiquetage lui-même est réalisé par 9 astronomes au moyen d'une interface web développée au sein d'EFIGI. Un jeu commun permet d'estimer et de corriger les variations interpersonnelles.



Références

- Baillard et al., 2006, in *Astronomical Data Analysis Software and Systems XV*, ASP Conf. Series 351, 236
- Bertin, 2005, in *Traitement du Signal et des Images GRETSI XX*, Presses Univ. de Louvain, 299
- Géraud, 2005, in *ISMM VII, Computational Imaging and Vision Series*, Springer, 105
- Ilbert et al., 2006, *Astronomy & Astrophysics* 457, 841
- Paturel et al., 2003, *Astronomy & Astrophysics* 412, 45

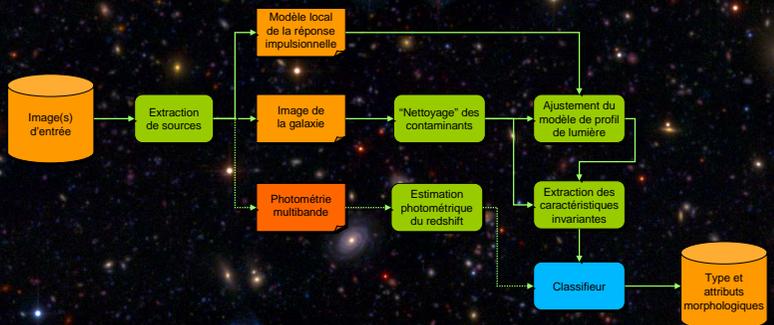
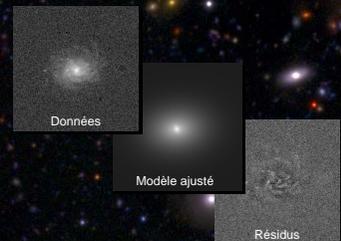


Schéma de fonctionnement global

- La galaxie est extraite de l'image d'entrée, ainsi qu'un modèle de la réponse impulsionnelle locale (Bertin 2006).
- L'image de la galaxie est "nettoyée" par application d'opérateurs de morphologie mathématique (Géraud 2005, Baillard et al. 2006).
- Un modèle de Sersic (plus éventuellement un modèle de disque exponentiel lorsque le rapport S/B est suffisant) convolué avec la réponse impulsionnelle locale est ajusté au profil de lumière.
- A partir des paramètres ajustés l'image est rendue invariante par translation, rotation et changement d'échelle, d'amplitude et de chiralité. Un vecteur de dimensions réduites est extrait par décomposition sur une base de primitives.
- Dans le cas de données multibandes, les informations photométriques permettent d'estimer le décalage spectral de la galaxie (Ilbert et al. 2006).
- Un classifieur à base de SVM identifie les attributs présents dans l'image.

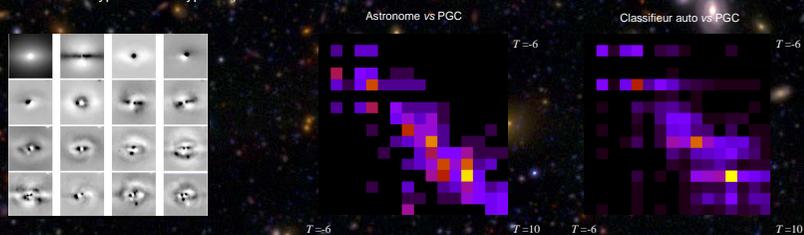
Modélisation du profil de lumière

Une galaxie peut être décomposée en une composante centrale de *bulbe* (dominée par des populations d'étoiles âgées) et une composante plus étendue de *disque* (constituée de populations plus jeunes). La mesure de l'enveloppe de brillance de ces composantes fournit aux astronomes des informations précises sur les dimensions et l'orientation de la galaxie, ainsi qu'une première estimation du type morphologique. Les 8 à 12 paramètres du modèle non-linéaire sont ajustés via la méthode Levenberg-Marquardt, avec prise en compte de la réponse impulsionnelle locale. Le défi présent est de réduire le temps de calcul de quelques secondes à quelques dizaines de millisecondes par galaxie.



Classification automatique du type de Hubble

Le prototype de classifieur actuel est basé sur une machine à vecteur de support (SVM) avec fonctions noyaux gaussiennes; solution qui s'est révélée le plus performante eu-égard à la taille limitée de l'échantillon d'apprentissage (~10³ galaxies). La base de fonctions ou primitives sur laquelle sont projetées les images réduites fait toujours l'objet d'études et de comparaisons. Nous présentons ci-dessous les résultats obtenus avec une base de Karhunen-Loève pour la détermination du type de Hubble. Les matrices de confusion indiquent des performances de classification comparables à celles d'un astronome spécialiste, avec un taux de succès de ~75% à 2 types de Hubble près (c'est-à-dire par exemple l'écart entre le type Sa et le type Sb).



Implantation

- L'implantation finale sur ferme de calcul sera accessible sous forme de webservice respectant les protocoles et formats d'échange de données de l'Observatoire Astronomique Virtuel (<http://www.france-vo.org>).