



ÉCOLE NATIONALE SUPÉRIEURE DES TÉLÉCOMMUNICATIONS

RAPPORT DE STAGE

Modélisation du profil de lumière des galaxies

05/07/2006 au 05/11/2006



Marta ARROYO

Sous la responsabilité de M. Henri MAÎTRE

Remerciements

Je tiens à remercier tout d'abord Henri Maître, mon maître de stage au sein de l'École Nationale Supérieure des Télécommunications (ENST) de Paris, non seulement pour m'avoir donné la chance de travailler dans son équipe mais aussi pour sa pédagogie et ses conseils tout au long de mon stage.

Je remercie plus particulièrement Emmanuel Bertin, astronome à l'Institut d'Astrophysique de Paris pour sa disponibilité et son aide précieuse sur le traitement des images de galaxies.

Je remercie également Saïd Ladjal, enseignant chercheur au département TSI de l'ENST, pour son aide au développement de l'implémentation logicielle de mon projet.

La présente note vise également à souligner le soutien indéfectible de mes collègues de laboratoire qui ont fourni de leur temps et de leurs efforts pour rendre possibles le développement de mon stage et mon intégration dans la culture française ainsi que l'apprentissage de la langue et qui ont participé au bon déroulement de mon stage.

Résumé

Nous proposons un algorithme d'ajustement 2D conçu dans le but d'extraire les composantes structurales des galaxies. 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.

Nous discutons les contraintes propres posées par ce type de modélisation et présentons notre implémentation logicielle, dont nous décrivons toutes les étapes suivies.

Nous présentons des résultats préliminaires obtenus à partir d'un sous échantillon de 200 galaxies du *Principal Galaxies Catalog* (PGC). Les performances de la modélisation sont analysées visuellement et sur des galaxies créées artificiellement.

Enfin nous discutons des diverses améliorations envisagées surtout dans le domaine du temps de calcul et la possibilité d'introduire d'autres modèles dans l'ajustement du profil pour mesurer des caractéristiques morphologiques plus compliquées.

Mots clés : Galaxies, morphologie, décomposition bulbe+disque, loi de Sérsic, modélisation du profil.

Table des matières

1	Introduction	4
1.1	Problématique	4
1.2	Le projet EFIGI	5
1.3	État de l’art	6
1.3.1	La séquence de Hubble	6
1.3.2	Modélisation non-linéaire du profil des galaxies	8
2	Méthodologie	10
2.1	Le profil de Sérsic	10
2.2	Seuillage	12
2.2.1	Méthode des courbes de niveau	12
2.2.2	Méthode de l’histogramme	13
2.3	Calcul des paramètres de position	15
2.4	L’algorithme de Levenberg-Marquardt	17
3	Base d’images	19
3.1	Le format FITS	21
3.2	Nettoyage des images	21
4	Résultats	23
5	Conclusions et perspectives	28

Table des figures

1.1	Caption for LOF	7
2.1	Images de la décomposition bulbe+disque selon la méthode des courbes de niveau.	13
2.2	Histogramme de la galaxie PGC0001621	14
2.3	Seuillage de la galaxie PGC0001621 par la méthode de Rosin	15
3.1	Images de la galaxie PGC60330 dans les 5 bandes du SDSS. .	19
3.2	Exemples de galaxies retenues.	20
3.3	Exemples de galaxies rejetées.	20
3.4	Galaxie PGC0010060 avant et après le processus de nettoyage	22
4.1	Résultats obtenus pour des galaxies réelles PGC0003830 ((a), (b) et (c)) et PGC0004415 ((d), (e) et (f)).	24
4.2	Profil de lumière des galaxies réelles et modèle approché . . .	24
4.3	Résultats obtenus pour des galaxies générées artificiellement .	25
4.4	Profil de lumière des galaxies générées artificiellement	25
4.5	Histogramme de l'erreur sur une centaine de galaxies réelles .	27
4.6	Histogramme du temps de calcul	27

Chapitre 1

Introduction

La classification morphologique des galaxies est un centre d'intérêt de l'astronomie extragalactique depuis de nombreuses années car l'analyse de la morphologie des galaxies donne des informations précieuses pour comprendre les processus d'évolution de ces objets et, par extension, de l'univers. Les galaxies sont des ensembles autogravitants en trois dimensions composés principalement d'étoiles, de gaz et de poussière, mais aussi de matière noire, invisible. Ces constituants sont répartis selon des proportions et des organisations variables d'une galaxie à une autre.

On distingue des structures communes dans toutes les galaxies, une composante sphéroïdale qui est la partie centrale; et un disque, la partie extérieure, dans laquelle on peut trouver des structures cohérentes ou non, traduisant des ondes de densité au sein de la galaxie. Le bulbe est la composante sphéroïdale des galaxies spirales, il a un profil dit "de Vaucouleurs", c'est-à-dire qu'il suit une loi radiale en $\exp(-r^{1/4})$. Les composantes sphéroïdales sont généralement constituées d'étoiles âgées, plus rouges que notre Soleil et pauvres en gaz et en poussière. Le disque est fortement aplati et son profil de brillance radial est exponentiellement décroissant. La population stellaire des disques est jeune, souvent plus bleue que notre Soleil et riche en gaz et en poussière (sauf pour les galaxies lenticulaires).

Le rapport entre la composante sphéroïdale (appelé bulbe pour les galaxies ayant un disque) et le disque peut varier de façon importante, donnant un panel de profils très différents. Observées depuis la Terre, elles sont semi-transparentes dans le spectre visible et se présentent sous différents angles de position et d'inclinaison ce qui contribue à la variation de leur brillance de surface.

1.1 Problématique

Les sources ponctuelles (indistinguables de la réponse impulsionnelle) comme les étoiles sont relativement simples à cataloguer : position, flux et

éventuellement 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 2^e ordre. À ceci trois raisons essentielles, qui sont autant de défis :

- Le temps de calcul : les volumes de données à analyser imposent un temps moyen de calcul par source ne pouvant excéder une dizaine de millisecondes.
- 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.
- 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 maîtrisés. 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 à 1/1000.

En conséquence, les catalogues des grands relevés actuels ne contiennent aucune description morphologique sophistiquée des galaxies. L'objectif de ce stage est de collaborer à la solution de ce problème, en proposant un système générique et performant de description et classification morphologiques des galaxies détectées sur les images astronomiques.

1.2 Le projet EFIGI

Ce stage s'inscrit dans le cadre d'un projet ACI "masses de données" de 3 ans nommé EFIGI (Extraction de Formes Idéalisées de Galaxies en Imagerie), financé par Fond National pour la Science et coordonné par TERAPIX. Ce projet implique ainsi la collaboration du LTCI (Laboratoire de Traitement et Communication de l'Information du CNRS et GET/Télécom Paris) et les astrophysiciens des observatoires de Paris, Marseille et Toulouse.

Le Projet EFIGI se propose d'aborder à la fois les aspects théoriques et algorithmiques de la problématique expliquée dans la section précédente et de développer une solution robuste pour mesurer en nombre la morphologie des galaxies présentes dans les grands relevés d'imagerie.

1.3 État de l'art

Étant donnée que la morphologie des galaxies est fortement liée à leurs propriétés physiques, il est rapidement apparu nécessaire de déterminer une façon de les distinguer qui ait une signification physique.

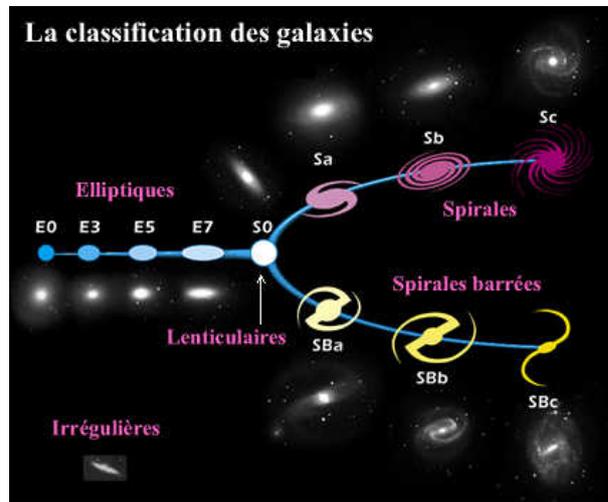
Les galaxies peuvent présenter diverses apparences mais, dans la grande majorité des cas, elles semblent pouvoir être rapportées dans une poignée de morphologies typiques. C'est sur ce constat qu'en 1926, Edwin Hubble (1889-1956) proposa la première classification largement utilisée parmi d'autres classifications plus exhaustives et plus objectives qui ont été créées postérieurement.

1.3.1 La séquence de Hubble

La séquence de Hubble (voir figure 1.3.1) distingue trois groupes principaux des galaxies auxquelles vient se rajouter un quatrième groupe :

- Les galaxies **elliptiques** sont assez communes. Elles ont une forme d'ellipsoïde à deux ou trois axes et sont classées avec la lettre E suivie d'un numéro indiquant leur ellipticité, de E0 pour les plus sphériques, à E7 pour les plus aplaties. Leurs régions centrales sont plus brillantes et leur éclat diminue régulièrement du centre vers l'extérieur.
- Les galaxies **spirales** se composent d'une grande concentration d'étoiles, nommée bulbe, qui définit leur région centrale. Ce bulbe est prolongé par un disque plat et étalé sur le plan équatorial, et dans lequel se développent dans deux directions opposées à partir du bulbe, deux bras spiraux. On distingue deux groupes différents de galaxies spirales : les galaxies spirales normales (notées S), dans lesquelles les bras partent du bulbe, et les galaxies spirales barrées (notées SB), dont le bulbe est traversé par une barre à l'extrémité de laquelle se détachent les bras spiraux. Chaque catégorie de galaxies spirales se subdivise en trois types selon l'ouverture des bras.
- Les galaxies **lenticulaires** sont des galaxies elliptiques très aplaties au bulbe très lumineux. Elles sont considérées comme une groupe de transition entre les galaxies spirales et les galaxies elliptiques.
- Les galaxies **irrégulières** n'ont pas de forme définie. Elles ne possèdent aucun axe de symétrie et sont souvent de faible masse. Ces galaxies représentent environ 2 ou 3% des galaxies répertoriées.

Malgré l'utilité du système de Hubble dans les relations entre les morphologies des galaxies, l'évolution galactique et les populations stellaires, il existe des imperfections. En particulier, la classification de Hubble est subjective dans la mesure où il n'y a pas de frontières clairement définies entre les différentes classes morphologiques et qu'elles reposent sur une classification manuelle effectuée par des experts. Par ailleurs, il existe des galaxies qui ne s'adaptent pas à la séquence de Hubble.

FIG. 1.1 – La classification morphologique des galaxies selon Hubble ^a

^a<http://jrosu.club.fr/univers/galaxies/classification/classification.htm>

Gérard de Vaucouleurs a proposé des types supplémentaires pour assurer la transition entre les galaxies spirales et irrégulières, les galaxies supergéantes, les spirales à distorsion ovale ou faiblement barrées, les galaxies qui présentent un anneau autour du noyau d'où partent les bras et celles dont les bras partent directement du noyau. Le tableau 1.1 donne les codes numériques pour les types morphologiques dans le système révisé de Vaucouleurs.

Code	type	Code	type	Code	type
-6	<i>cE</i>	0	<i>S0/a</i>	6	<i>Scd</i>
-5	<i>E0</i>	1	<i>Sa</i>	7	<i>Sd</i>
-4	<i>E⁺</i>	2	<i>Sab</i>	8	<i>Sdm</i>
-3	<i>S0⁻</i>	3	<i>Sb</i>	9	<i>Sm</i>
-2	<i>S0⁰</i>	4	<i>Sbc</i>	10	<i>Im</i>
-1	<i>S0⁺</i>	5	<i>Sc</i>	11	<i>cI</i>

TAB. 1.1 – Codes numériques pour les types morphologiques de Vaucouleurs

Des travaux plus récents ont été menés dans l'optique d'un système de classification objectif et automatique. Bien entendu, la classification manuelle des images à faible résolution mène malheureusement, à une grande différence entre les classificateurs humains. De plus, l'arrivée de bases de données astronomiques considérables requiert une classification automatique.

1.3.2 Modélisation non-linéaire du profil des galaxies

Historiquement, les efforts de description morphologique des galaxies se sont principalement portés sur la modélisation des profils des objets. En 1948 de Vaucouleurs montra que la majorité des galaxies elliptiques ont une distribution de lumière suivant une loi exponentielle $\exp -r^{1/4}$ tandis qu'en 1970 Freeman ajouta au modèle un disque exponentiel pour améliorer les résultats de Vaucouleurs.

La plupart des travaux empiriques effectués plus tard suivant les résultats de Vaucouleurs et Freeman ont montré qu'une galaxie peut être décomposée en 2 composantes axisymétriques de bulbe (loi de "de Vaucouleurs") et de disque (loi exponentielle tronquée), chacune décrite par plusieurs paramètres. En effet, les résultats obtenus d'après la modélisation des profils des galaxies ont mené à une notable amélioration de la compréhension de l'évolution des galaxies ainsi de leur formation. Ces techniques, utilisant la décomposition bulbe+disque des galaxies, ont été incluses dans la catégorie de classificateurs paramétriques avec, parmi d'autres, la déconvolution radiale multi-gaussienne.

Malgré cela, les modèles sont évidemment une simple approximation car la majorité des galaxies présentent plus de deux composantes, à savoir des bras spiraux, des barres ... c'est-à-dire, des profils plus complexes. L'ajustement du profil des galaxies par des modèles représente une approche prometteuse puisque après l'analyse soignée des paramètres obtenus, on peut trouver des informations utiles et plus complexes sur la structure des galaxies. Par exemple, une déviation bien marquée entre l'angle de rotation du bulbe et celui du disque peut indiquer la présence d'une barre dans la galaxie.

Il existe deux types d'ajustements : 1D et 2D, le premier est fréquemment utilisé parce qu'il donne des résultats suffisants pour certaines études, et de plus il est simple à implémenter. Par contre, il n'est pas capable d'ajuster les galaxies qui présentent un désalignement entre le centre du bulbe et le centre du disque. À cause de ces problèmes, les techniques d'ajustement 2D ont été développées dans les programmes du type GIM2D [1] et GALFIT [3], elles fournissent des résultats plus précis et des modèles détaillés des galaxies bien échantillonnées mais, en revanche, le temps de calcul est aussi plus élevé. C'est pourquoi la recherche des algorithmes plus rapides et fiables est importante ainsi comme les nouvelles techniques qui mélangent les deux types d'ajustements pour sélectionner les meilleures caractéristiques de chacun. Cela peut être remarqué dans les programmes du type GASPHOT [4].

Dans tous ces cas, l'ajustement est évidemment non-linéaire et réalisé par minimisation globale. À chaque pas de calcul, une convolution du modèle par la réponse impulsionnelle locale est nécessaire. 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.

Chapitre 2

Méthodologie

La méthode développée pendant ce stage peut être globalement divisée en trois étapes consécutives :

1. Seuillage.
2. Calcul des conditions initiales du modèle.
 - Calcul des paramètres de position.
3. Ajustement du profil de la galaxie grâce à l'algorithme de Levenberg-Marquardt.

Les étapes antérieures à l'ajustement sont décisives pour l'obtention de résultats corrects, et cela avec un temps de calcul raisonnable.

Dans les sections suivantes nous aborderons chacune des étapes dans le processus en analysant leurs aspects théoriques.

2.1 Le profil de Sérsic

Nous avons introduit dans la section précédente la décomposition des galaxies en 2 composantes axisymétriques de bulbe (loi de “de Vaucouleurs”) et de disque (loi exponentielle tronquée), chacune décrite par plusieurs paramètres.

Cette approche est convenable car les profils de brillance de la surface des bulbes sont bien modélisés par une loi de Vaucouleurs, $\exp -r^{1/4}$, une loi de Sérsic, $\exp -r^{1/n}$, ou une loi exponentielle, alors que les disques suivent une loi purement exponentielle de manière assez exacte.

Nous avons utilisé un modèle avec une seule composante suivant la loi de Sérsic (1968),

$$I(r) = I_e \cdot e^{-k \left(\left(\frac{r}{r_e} \right)^{1/n} - 1 \right)} \quad (2.1)$$

où r_e représente le rayon effectif de la galaxie, I_e est la brillance de surface au rayon effectif, n exprime l'indice de Sérsic et k est un paramètre qui dépend étroitement de l'indice de Sérsic à condition que r_e enferme la

moitié de la luminosité totale de l'objet. Pour le paramètre k , nous utilisons l'approximation donnée par Ciotti (1991) :

$$k \approx 2n - \frac{1}{3} + \frac{4}{405n} + \frac{46}{25515n^2} + \frac{131}{1148175n^3} \quad (2.2)$$

Cette expression est valide sur l'intervalle d'intérêt physique avec une erreur relative inférieure à 10^{-6} de n . Particulièrement, $k = 7.676$ pour $n = 4$ et $k = 1.676$ pour $n = 1$ (lois de Vaucouleurs et exponentielle respectivement).

Il faut remarquer que la loi de Vaucouleurs correspond à l'indice de Sérsic $n = 4$, le profil gaussien et modélisé avec $n = 0.5$.

Ensuite, nous ajoutons les dépendances avec la relation entre les axes de telle sorte que nous puissions ajuster non seulement les modèles des galaxies circulaires mais aussi ceux des galaxies elliptiques. Nous supposons d'abord que les axes sont alignés avec les axes de coordonnées. Ainsi pour la coordonnée radiale, on obtient :

$$\left(\frac{r}{r_e}\right) = \sqrt{\left(\frac{x - x_c}{A}\right)^2 + \left(\frac{y - y_c}{B}\right)^2} \quad (2.3)$$

avec x_c et y_c les coordonnées du barycentre de la galaxie, A , la longueur du demi-grand axe et B , la longueur du demi-petit axe de l'ellipse quand le rayon de celle-ci est le rayon effectif, r_e .

De plus, il faut prendre en considération le possible angle de rotation entre l'axe principal de l'ellipse qui modélise la galaxie et l'axe de coordonnées ; en appliquant la matrice de rotation, on obtient :

$$\begin{pmatrix} x_r \\ y_r \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \cdot \begin{pmatrix} x - x_c \\ y - y_c \end{pmatrix}$$

Autrement exprimé :

$$x_r = (x - x_c) \cdot \cos(\theta) - (y - y_c) \cdot \sin(\theta) \quad (2.4)$$

$$y_r = (x - x_c) \cdot \sin(\theta) + (y - y_c) \cdot \cos(\theta) \quad (2.5)$$

Finalement, nous ajoutons une constante, b , qui ajuste le niveau moyen du fond de ciel pour éviter que notre modèle soit biaisé, et on remplace les équations (2.3), (2.4) et (2.5) dans (2.1) pour obtenir l'expression analytique.

On a ainsi :

$$I(r) = b + I_e e^{-k \left(\left(\left(\frac{(x-x_c) \cdot \cos(\theta) - (y-y_c) \cdot \sin(\theta)}{A} \right)^2 + \left(\frac{(x-x_c) \cdot \sin(\theta) + (y-y_c) \cdot \cos(\theta)}{B} \right)^2 \right)^{1/2n} - 1 \right)} \quad (2.6)$$

On peut conclure que pour l'ajustement complet du modèle de Sérsic dans une implémentation 2D on a besoin des huit paramètres libres suivants : x_c , y_c , n , θ , I_e , A , B et b . Par conséquent, notre but principal par la suite sera la recherche de ces paramètres, grâce auxquels on obtiendra un ajustement efficace et robuste du profil de brillance par un modèle de bulbe selon la loi de Sérsic présentée antérieurement.

2.2 Seuillage

Le problème du seuillage est délicat. Il a comme but principal la fabrication d'un masque de pixels permettant de mesurer les moments du deuxième ordre qui vont permettre à continuation d'initialiser les paramètres du modèle à des valeurs raisonnables avant itération. En effet, le seuillage a une influence importante sur les conditions initiales et par conséquent sur l'effectivité du programme. C'est pourquoi on implémente deux méthodes pour le calcul du seuil. La première repose sur les courbes de niveau et elle nous fournit des informations supplémentaires sur les caractéristiques de la galaxie. Pour commencer, nous pouvons non seulement détecter le bulbe de la galaxie mais tous les paramètres associés à la décomposition bulbe+disque, en particulier le rapport entre l'intensité du bulbe et celui de l'ensemble de la galaxie.

En revanche, cette première méthode n'est pas assez performante pour la détection de la galaxie elle-même dans le bruit du fond de ciel, surtout si notre image est très bruitée. Par conséquent, nous utilisons une deuxième méthode basée sur l'histogramme qui nous assurera le bon seuil.

Les sections suivantes détaillent les deux méthodes utilisées dans notre algorithme.

2.2.1 Méthode des courbes de niveau

La méthode des courbes de niveau repose sur le principe de Helmholtz qui est utilisé pour la reconnaissance de structures dans une image sans aucune information *a priori* [9].

Selon le principe de Helmholtz, une structure géométrique est significative si le nombre des occurrences est petit dans une situation aléatoire, c'est-à-dire, si les structures géométriques sont considérées comme de grandes déviations de l'aspect aléatoire. Pourtant, au lieu de chercher des objets le plus similaires possible du modèle *a priori* comme on le ferait selon le critère du Maximum a Posteriori (MAP), on considère un modèle simple, à savoir une distribution aléatoire uniforme, et les objets sont définis comme de grandes déviations de ce modèle générique, donc, on peut dire qu'il s'agit d'une méthode de Minimum a Posteriori avec un modèle de bruit aléatoire uniforme.

Notre but est de chercher une frontière entre l'objet et le fond de ciel, puis une frontière entre la région centrale (dominée par le bulbe) et le disque de la galaxie. Il se trouve que l'on peut parvenir à notre objectif en cherchant des courbes fermées assez longues, avec un contraste assez grand et qui s'ajustent bien avec la géométrie de l'image, à savoir, orthogonales au gradient de l'image dans tous ses points. En conséquence, l'événement géométrique à considérer est : un grand contraste sur une courbe de niveau de l'image.

Ensuite, nous calculons le nombre d'apparitions espérées de tels événements,

autrement dit, le nombre de fausses alarmes permises. Cela nous donnera non seulement la longueur minimale mais aussi le contraste minimal nécessaire pour que la courbe de niveau soit significative.

La courbe de niveau frontière entre le bulbe et le disque est définie comme la courbe fermée la plus significative de l'image avec une quantité maximale de fausses alarmes $N_{FA} = 1$. Il se trouve que l'on peut détecter plusieurs courbes de niveau maximales qui vérifient la condition du nombre maximal de fausses alarmes, donc on diminuera itérativement le nombre de fausses alarmes permises jusqu'à ne plus obtenir qu'une seule courbe de niveau. Cela sera la frontière qui sépare la nuage du noyau de la galaxie.

D'autre part, nous obtenons le seuil qui sépare la galaxie du fond de ciel en fixant un nombre maximal de fausses alarmes assez élevé pour détecter toutes les courbes de niveau avec une certaine significativité et nous choisissons la courbe fermée la plus longue.

La figure 2.1 illustre les résultats obtenus après l'application de la méthode des courbes de niveau.

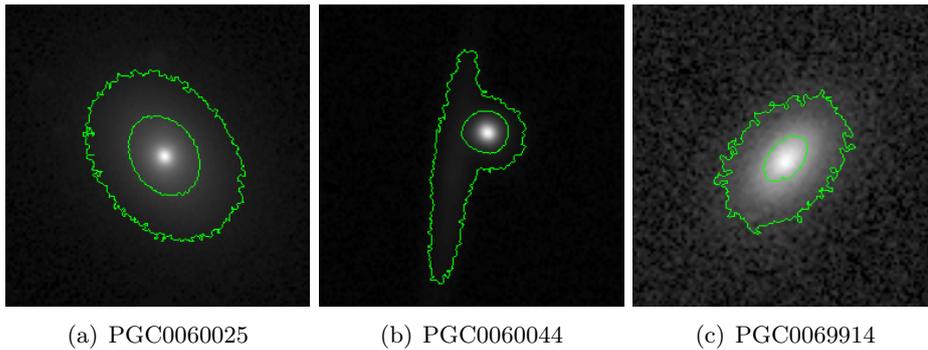


FIG. 2.1 – Images de la décomposition bulbe+disque selon la méthode des courbes de niveau.

2.2.2 Méthode de l'histogramme

La méthode de l'histogramme est utilisée dans notre chaîne d'ajustement des profils de galaxies comme une aide supplémentaire pour la recherche de l'objet dans l'image. Malgré l'efficacité de la méthode des courbes de niveau, développée dans la section précédente, nous n'obtenons pas des résultats corrects dans la totalité des cas en nous limitant au calcul du seuil entre la galaxie et le fond de ciel, donc nous avons besoin d'une autre méthode auxiliaire qui nous fournira le bon seuil quand la méthode des courbes de niveau échoue.

On sait que l'information issue des histogrammes a été historiquement utilisée en traitement d'images pour effectuer des tâches très diverses. Le seuillage des images, qui est l'un des traitements des images les plus courants,

se fait souvent par des techniques de ce type, mais pour notre problème, il rencontre des difficultés particulières.

En effet, l'histogramme typique d'une galaxie, présente un seul pic placé vers les faibles niveaux de gris qui correspond au fond de ciel comme l'on peut le voir sur la figure 2.2.

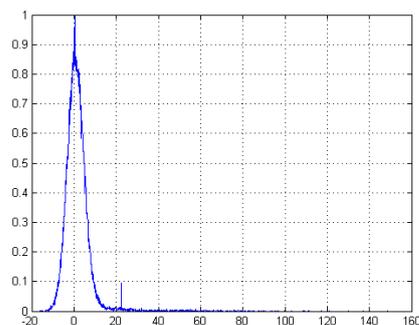


FIG. 2.2 – Histogramme de la galaxie PGC0001621

Étant donné que l'algorithme décrit dans cette section est un algorithme auxiliaire, nous cherchons à implémenter une méthode simple avec des performances acceptables avec les histogrammes qui présentent un seul mode, c'est-à-dire, les histogrammes avec un seul pic. C'est pourquoi nous allons implémenter l'algorithme proposé par Rosin [10], caractérisée par sa simplicité.

La méthode prend comme hypothèse qu'il y a une population dominante dans l'image qui produit le pic principal et qui est placée vers les niveaux bas de l'histogramme alors que la population secondaire est placée vers les niveaux hauts de l'échelle de gris.

Ensuite, on trace une droite qui relie le pic principal provoqué par la population principale à la queue de l'histogramme et on choisit le seuil comme le niveau de gris qui maximise la distance perpendiculaire entre l'histogramme et la droite. La figure 2.3 illustre la procédure avec une image d'une galaxie.

Pour conclure, il faut remarquer que l'algorithme repose sur quelques hypothèses de base qui doivent être vérifiées par l'histogramme de l'image, à savoir :

1. Il faut savoir si les valeurs de gris présentes dans la population principale sont plus grandes ou plus petites que celles présentes dans la population secondaire pour adapter la méthode.
2. Le pic principal de l'histogramme a une base détectable qui correspond au seuil.
3. Pour plus de simplicité, on considère aussi que la population secondaire est assez éloignée de la principale et qu'elle ne présente pas de pics significatifs dans l'histogramme.

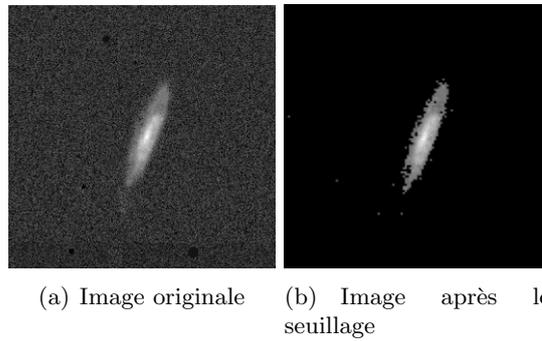


FIG. 2.3 – Seuillage de la galaxie PGC0001621 par la méthode de Rosin

Les histogrammes des images des galaxies vérifient les hypothèses précédentes, donc la méthode est applicable dans la chaîne d'analyse que l'on cherche à implémenter.

2.3 Calcul des paramètres de position

Le calcul des paramètres de position est basé sur des calculs simples et bien connus, notamment les moments du premier et du deuxième ordre. On fait le calcul des moments de l'image après le seuillage et on obtient les conditions initiales que l'on fournira ensuite à l'algorithme de Levenberg-Marquardt pour achever la modélisation du profil de lumière.

Le barycentre

Les deux premiers paramètres à calculer correspondent aux coordonnées du barycentre de l'image, c'est-à-dire, aux coordonnées du centre de la galaxie. Nous faisons tout simplement le calcul des moments du premier ordre de l'image obtenue après le seuillage.

$$x_c = \frac{\sum_{i \in S} I_i x_i}{\sum_{i \in S} I_i}, \quad (2.7)$$

$$y_c = \frac{\sum_{i \in S} I_i y_i}{\sum_{i \in S} I_i} \quad (2.8)$$

où S représente les pixels de l'image au dessus le seuil.

Moments de deuxième ordre

Les moments du deuxième ordre centrés sont adaptés à la mesure de l'extension du profil d'une source. Ils sont calculés à partir des équations suivantes :

$$\overline{x^2} = \frac{\sum_{i \in S} I_i (x_i - x_c)^2}{\sum_{i \in S} I_i}, \quad (2.9)$$

$$\overline{y^2} = \frac{\sum_{i \in S} I_i (y_i - y_c)^2}{\sum_{i \in S} I_i}, \quad (2.10)$$

$$\overline{xy} = \frac{\sum_{i \in S} I_i (x_i - x_c)(y_i - y_c)}{\sum_{i \in S} I_i} \quad (2.11)$$

À partir des moments du deuxième ordre nous pouvons calculer les paramètres qui décrivent la forme de la galaxie, à savoir l'angle de rotation ainsi que la longueur du grand axe et du petit axe. Ces paramètres nous permettent de décrire les galaxies comme des ellipses.

On calcule d'abord le système de coordonnées de référence tourné d'un angle θ de la façon suivante :

$$\overline{x_\theta^2} = \overline{x^2} \cos^2 \theta + \overline{y^2} \sin^2 \theta - 2\overline{xy} \cos \theta \sin \theta, \quad (2.12)$$

$$\overline{y_\theta^2} = \overline{x^2} \sin^2 \theta + \overline{y^2} \cos^2 \theta + 2\overline{xy} \cos \theta \sin \theta, \quad (2.13)$$

$$\overline{x_\theta y_\theta} = \overline{x^2} \sin \theta \cos \theta - \overline{y^2} \cos \theta \sin \theta + \overline{xy} (\cos^2 \theta - \sin^2 \theta) \quad (2.14)$$

Ensuite, on calcule l'angle de rotation comme celui qui présente la variance minimale (ou maximale) selon l'axe x_θ donc on cherche à résoudre l'équation :

$$\left. \frac{\partial \overline{x_\theta^2}}{\partial \theta} \right|_{\theta_0} = 0 \quad (2.15)$$

Le résultat conduit à l'expression suivante :

$$2 \cos \theta_0 \sin \theta_0 (\overline{y^2} - \overline{x^2}) + 2 (\cos^2 \theta_0 - \sin^2 \theta_0) \overline{xy} = 0 \quad (2.16)$$

Il en découle donc que si $\overline{x^2} \neq \overline{y^2}$ on peut calculer l'angle de rotation de la galaxie de manière directe :

$$\theta_0 = \frac{1}{2} \arctan \frac{2\overline{xy}}{\overline{x^2} - \overline{y^2}} \quad (2.17)$$

L'angle de rotation est défini comme celui qui maximise la valeur de $\overline{x_\theta^2}$, c'est pourquoi la solution à l'équation 2.17 choisie est celle qui a le même signe que la covariance \overline{xy} .

Pour conclure, nous calculons la longueur du grand axe ainsi que celle du petit axe de façon directe à partir des moments de deuxième ordre en utilisant les équations 2.13, 2.14 et 2.14. On a ainsi :

$$A^2 = \frac{\overline{x^2} + \overline{y^2}}{2} + \sqrt{\left(\frac{\overline{x^2} - \overline{y^2}}{2}\right)^2 + \overline{xy}^2}, \quad (2.18)$$

$$B^2 = \frac{\overline{x^2} + \overline{y^2}}{2} - \sqrt{\left(\frac{\overline{x^2} - \overline{y^2}}{2}\right)^2 + \overline{xy}^2} \quad (2.19)$$

2.4 L'algorithme de Levenberg-Marquardt

Le moteur de calcul utilisé pour achever la modélisation du profil de lumière des galaxies une fois que l'on a obtenu les conditions initiales est l'algorithme de Levenberg-Marquardt.

L'algorithme de Levenberg-Marquardt est notamment utilisée sur les problèmes d'optimisation non-linéaire d'échelle moyenne. Il fournit des améliorations par rapport à l'algorithme du gradient car il estime la courbure de la surface d'erreur en utilisant des informations de deuxième ordre.

On cherche à minimiser notre fonction χ^2 qui est représentée par la formule suivante :

$$\chi^2(\vec{a}) = \sum_{x=1}^{nx} \sum_{y=1}^{ny} \left(\frac{im_{x,y} - f_{x,y}(\vec{a})}{\sigma_{x,y}} \right)^2 \quad (2.20)$$

où $im_{x,y}$ est la valeur du pixel (x, y) dans l'image à modéliser de dimension $(nx \times ny)$ et $f_{x,y}(\vec{a})$ représente le modèle de Sérsic au pixel (x, y) (voir équation 2.6). Le vecteur \vec{a} est formé par tous les paramètres à estimer, c'est-à-dire, $\vec{a} = [x_c, y_c, I_e, n, \theta, A, B, b]$. Par la suite, on va représenter la fonction de coût en chaque pixel de la façon suivante :

$$J(\vec{a}) = (im_{x,y} - f_{x,y}(\vec{a}))^2 \quad (2.21)$$

Notre problème est non-linéaire, cependant on peut faire une approximation par un modèle linéaire et convertir la fonction de coût en une fonction quadratique si l'on se trouve près du minimum. On obtient ainsi :

$$\hat{f}_{x,y}(\vec{a}) = f_{x,y}(\vec{a}_0) + (\vec{a} - \vec{a}_0)^T \nabla f_{x,y}(\vec{a}_0) \quad (2.22)$$

Ensuite, on calcule l'expression de la fonction de coût en supposant que le modèle est $\hat{f}_{x,y}(\vec{a})$ ainsi comme l'expression du gradient :

$$\hat{J}(\vec{a}) = (im_{x,y} - \hat{f}_{x,y}(\vec{a}))^2 \quad (2.23)$$

$$\nabla \hat{J}(\vec{a}) = 2(\hat{f}_{x,y}(\vec{a}) - im_{x,y}) \nabla \hat{f}_{x,y}(\vec{a}) \quad (2.24)$$

En remplaçant l'expression de $\hat{f}_{x,y}(\vec{a})$ dans l'équation 2.24 on obtient :

$$\nabla \hat{J}(\vec{a}) = 2(\mathbf{H}(\vec{a} - \vec{a}_0) + \nabla J(\vec{a}_0)) \quad (2.25)$$

où \mathbf{H} est la matrice hessienne définie comme :

$$\mathbf{H} = \nabla f_{x,y}(\vec{a}_0) \nabla f_{x,y}(\vec{a}_0)^T$$

Reprenant donc l'expression de la fonction χ^2 on trouve que son gradient s'exprime selon l'équation suivante :

$$\nabla \chi^2(\vec{a}) = \sum_{x=1}^{nx} \sum_{y=1}^{ny} \frac{2}{\sigma_{x,y}^2} (\mathbf{H}(\vec{a} - \vec{a}_0) + \nabla J(\vec{a}_0)) \quad (2.26)$$

Autrement exprimé :

$$\nabla \chi^2(\vec{a}) = \nabla \chi^2(\vec{a}_0) + \sum_{x=1}^{nx} \sum_{y=1}^{ny} \frac{2\mathbf{H}(\vec{a} - \vec{a}_0)}{\sigma_{x,y}^2} \quad (2.27)$$

Pour conclure, on égale à zéro l'équation 2.25 et on obtient le vecteur de paramètres optimal qui résout notre problème d'optimisation non-linéaire pour la modélisation des galaxies.

$$\nabla \hat{J}(\vec{a}) = 0 \Rightarrow \vec{a}_{opt} = -\mathbf{H}^{-1} \nabla J(\vec{a}_0) + \vec{a}_0 \quad (2.28)$$

Une fois que l'on a obtenu l'expression analytique de l'algorithme on peut chercher à implémenter l'algorithme itératif de Levenberg-Marquardt. On introduit ainsi le paramètre λ qui règle le comportement de l'algorithme selon que la fonction de coût augmente ou diminue. Si λ diminue la formule prise par l'algorithme est celle qu'on a développée précédemment et si, par contre, λ augmente, l'algorithme prend la formule du gradient. On a donc :

$$\vec{a}_{i+1} = \vec{a}_i - (\mathbf{H} + \lambda \text{diag}[\mathbf{H}])^{-1} \nabla J(\vec{a}_0) \quad (2.29)$$

En définitive, le processus suivi par l'algorithme Levenberg-Marquardt pour achever la minimisation de la fonction χ^2 est le suivant :

1. Calcul de la fonction χ^2
2. On prend une valeur modeste pour λ , par exemple $\lambda = 0.001$
3. On calcule le nouveau vecteur de paramètres et la fonction χ^2 associée.
4. Si la fonction χ^2 a grandi on augmente la valeur de λ d'un facteur 10.
5. Si la fonction χ^2 est plus petite qu'avant on diminue la valeur de λ par un facteur 10.
6. On retourne au (3) et on continue jusqu'à la convergence des résultats.

Chapitre 3

Base d'images

Nous avons travaillé sur des images extraites du catalogue *Principal Galaxies Catalog RC3/PGC*¹, observées avec le *Sloan Digital Sky Survey (SDSS) release 2*². L'IAP les a mise à notre disposition, ainsi que les types morphologiques et types de Hubble associés (voir tableau 1.1).

Pour chaque galaxie PGC, dans chaque bande, 9 images ont été obtenues directement du serveur SDSS DAS dont le niveau de fond de ciel a été automatiquement soustrait. Les images des galaxies ont été rééchantillonnées et redimensionnées à 255x255 (en utilisant SWarp avec l'interpolation de Lanczos3). Le centre de la galaxie est défini par des coordonnées J2000, récupérées du service web SESAME³.

Notre choix s'est porté sur ce programme d'observation car il est homogène, récent et de bonne qualité, ce choix est également justifié parce que les objets de ce catalogue sont limités en diamètre. De cette façon, l'échantillon doit représenter équitablement toutes les classes de galaxies, ce qui n'aurait pas été le cas avec une sélection selon la brillance de surface par exemple.

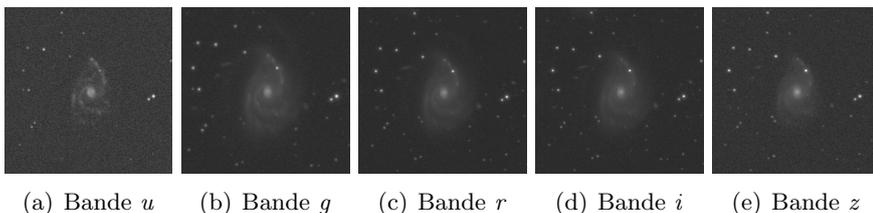


FIG. 3.1 – Images de la galaxie PGC60330 dans les 5 bandes du SDSS.

Toutes les images dans les bandes *g* (bleu-vert), *r* (orange) et *i* (très proche infrarouge) ont été vérifiées à l'oeil, et environ 300 galaxies ont été

¹Paturel et al., 2003, *Astronomy & Astrophysics* 412, 45

²<http://www.sdss.org/>

³CDS, Strasbourg, France

effacées de la base car elles étaient incomplètes dans ces bandes. Après toutes les étapes suivies on avait encore 4462 galaxies dans la base d'images. Finalement il faut remarquer que les images appartenant aux bandes u (ultraviolet) et z (proche infrarouge) n'ont pas été vérifiées visuellement donc quelques galaxies peuvent rester incomplètes dans ces bandes là.

Nous nous sommes concentrés sur les galaxies de la bande g car cette bande est similaire a celle des plaques photographiques sur lesquelles la séquence de Hubble a été élaborée. En effet, les éléments qui composent les galaxies émettent et absorbent différemment la lumière aux longueurs d'onde différentes ce qui peut être vérifié dans la figure 3.1. Plus spécifiquement nous avons travaillé sur des images correspondant aux galaxies appartenant aux classes elliptiques, lenticulaires et spirales. La figure 3.2 illustre une sélection de galaxies utilisées dans notre travail alors que la figure 3.3 montre une ensemble de galaxies rejetées.

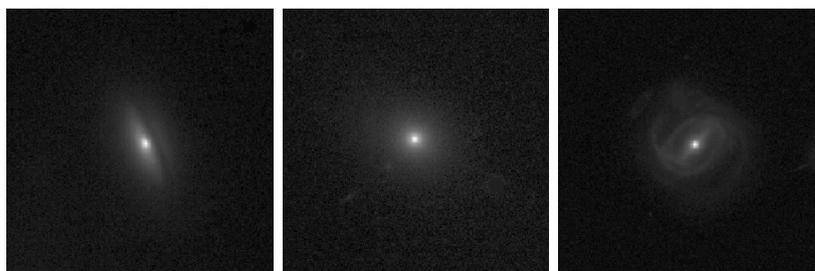


FIG. 3.2 – Exemples de galaxies retenues.

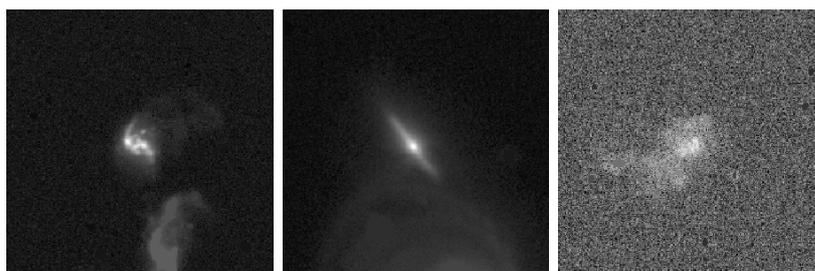


FIG. 3.3 – Exemples de galaxies rejetées.

L'IAP nous a également fourni une version d'images des galaxies avec une bonne résolution, variées, générées artificiellement. Ces simulations avaient un niveau de bruit réaliste (gaussien + poissonnien en régime gaussien), mais une réponse impulsionnelle de largeur négligeable. Certaines galaxies avaient des profils simples, à savoir bulbe de Sérsic tout simplement, d'autres étaient la somme d'un bulbe de Sérsic et d'un disque.

3.1 Le format FITS

Le format des images dans la base de données est nommé FITS ou Flexible Image Transfer System [15]. Il s'agit du format de fichiers le plus communément utilisé en astronomie. FITS est souvent utilisé pour sauvegarder aussi d'autres données, comme le spectre, des listes de photons, des cubes de données, et bien d'autres choses encore.

Un fichier FITS est normalement formé par une séquence de *Header Data Units (HDU)* qui peut être suivie par des registres spéciaux. La première partie de chaque HDU est formée par un texte ASCII qui contient des mots clés auxquels sont associées des valeurs qui indiquent la taille, le format et la structure des données qui suivent. Ces paramètres sont écrits de telle sorte qu'ils sont interprétables pour les humains et les ordinateurs.

Les données sont représentées par vecteurs composés soit de valeurs entières, soit de flottants de 32-bit ou 64-bit, ce qui sera spécifié dans la HDU principal.

Un fichier FITS peut contenir plusieurs extensions, et chacune de celles-ci peut contenir des données. Par exemple, il est possible de sauvegarder dans le même fichier FITS des images à la fois dans le domaine des rayons X et dans celui de l'infrarouge.

3.2 Nettoyage des images

Le premier objectif à atteindre est le nettoyage des images. L'IAP nous a fourni des images qui ont été soumises à cette première étape, la technique décrite a été mise au point et appliquée par Anthony Baillard[5]. Le but du nettoyage est l'élimination des objets parasites présents dans l'image, c'est-à-dire, que l'on a retiré certaines composantes à partir des résultats d'une segmentation.

La morphologie mathématique est un outil d'extraction de composantes d'image qui est utile pour la représentation et la description. C'est pourquoi la technique utilisée dans la réalisation du nettoyage repose sur la représentation sous la forme d'un arbre en utilisant des filtres morphologiques sur un tel arbre. De façon générale, la plupart des filtres morphologiques sont basés sur de simples opérations d'érosion et de dilatation.

Tout d'abord on construit l'arbre des composantes de l'image et pour compléter le nettoyage, on applique les opérateurs basiques suivants :

1. Ouverture d'aire avec une taille donnée pour repérer les zones claires.
2. Chapeau haut-de-forme blanc (*White top-hat*) pour conserver uniquement les pixels de l'ouverture d'aire.
3. Seuillage pour créer un masque qui contient uniquement les points les plus lumineux.
4. Étiquetage par 4-connexité pour obtenir des composantes distinctes.

5. Suppression de la composante centrale correspondant au cœur de la galaxie.
6. Dilatation par 8-connexité pour récupérer une valeur de remplacement pour chaque composante étiquetée.
7. Minimum local sur l'image d'origine pour niveler les objets parasites en fonction des composantes dilatées.

La figure 3.4 illustre les résultats du nettoyage sur une exemple de galaxie, à partir desquelles nous pouvons commencer à travailler. On voit que quelques étoiles fortes situées en avant de la galaxie ont été éliminées.

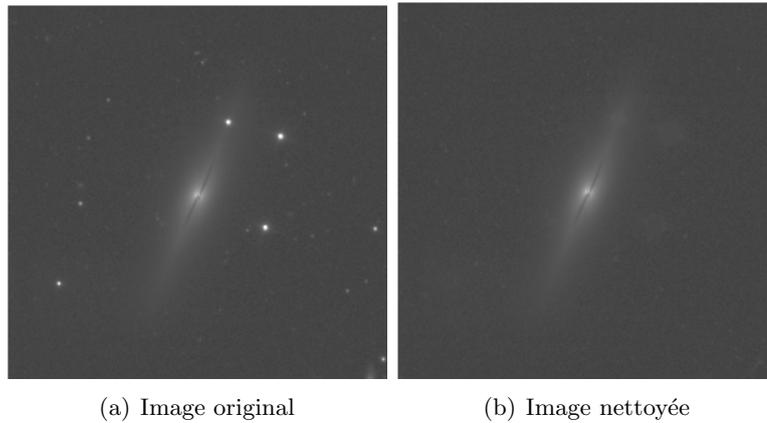


FIG. 3.4 – Galaxie PGC0010060 avant et après le processus de nettoyage

Chapitre 4

Résultats

Ce chapitre détaille les résultats de l'ajustement du profil des galaxies de notre base d'images, obtenus à partir des méthodes expliquées dans la section 2.

Les résultats de la modélisation sont assez performants, on obtient des paramètres qui modélisent les galaxies en quelques secondes de manière satisfaisante et fiable. Les figures suivantes illustrent plusieurs exemples des résultats obtenus lors de la modélisation du profil de lumière avec un modèle de Sérsic sur des galaxies.

Nous remarquons que les images des résidus qui montrent la différence entre l'image original et le modèle ne contiennent que du bruit et ponctuellement des objets plus complexes de la géométrie de la galaxie, à savoir des barres ou des bras spiraux.

De plus, si nous analysons le profil de lumière sur l'axe principal de la galaxie nous notons que l'ajustement du profil est bien meilleur vers le centre de la galaxie car le profil réel du bulbe est bien modélisé par la fonction de Sérsic tandis que les ailes suivent des fonctions exponentielles que nous n'utilisons pas dans notre algorithme. En effet, lorsque nous examinons les résultats pour les images des galaxies créées artificiellement et qui n'ont qu'une seule composante de bulbe, nous remarquons que l'ajustement est très performant (voir figure 4.3 et 4.4).

La valeur de l'indice de Sérsic apparaît normalement sous-estimé car nous ne considérons pas la réponse impulsionnelle. La résolution des images astronomiques est limitée par la turbulence atmosphérique qui présente une MTF (*Modulation Transfer Function*) qui est approximativement gaussienne.

$$MTF(\|\vec{k}\|) \propto e^{-\|\vec{k}\|^{5/3}} \quad (4.1)$$

C'est pourquoi après la convolution du profil de Sérsic avec la réponse impulsionnelle on obtient un profil dont la partie centrale (le pic de luminosité) prend une forme gaussienne, les ailes restant dominées par le profil de Sérsic et c'est cette image là que l'on cherche à modéliser sans prendre en compte

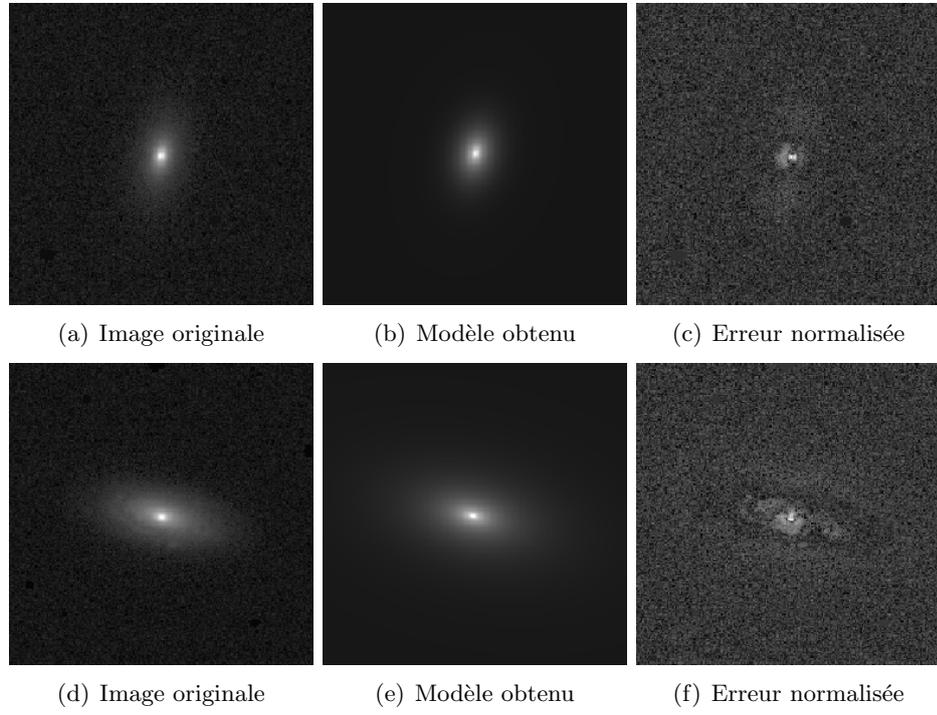


FIG. 4.1 – Résultats obtenus pour des galaxies réelles PGC0003830 ((a), (b) et (c)) et PGC0004415 ((d), (e) et (f)).

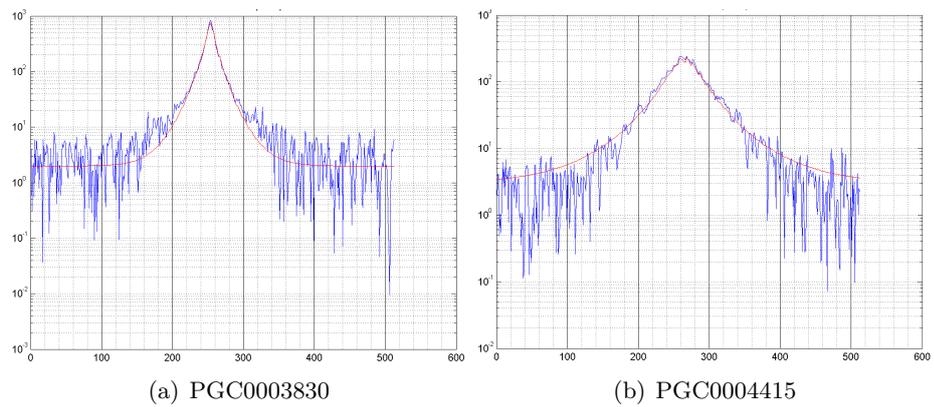


FIG. 4.2 – Profil de lumière des galaxies réelles et modèle approché

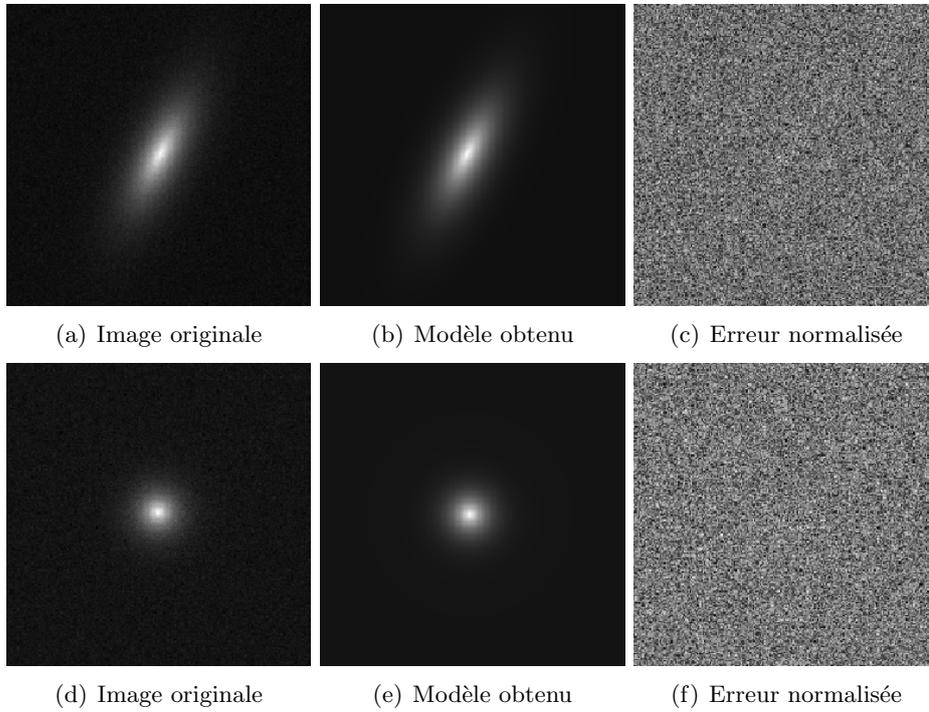


FIG. 4.3 – Résultats obtenus pour des galaxies générées artificiellement

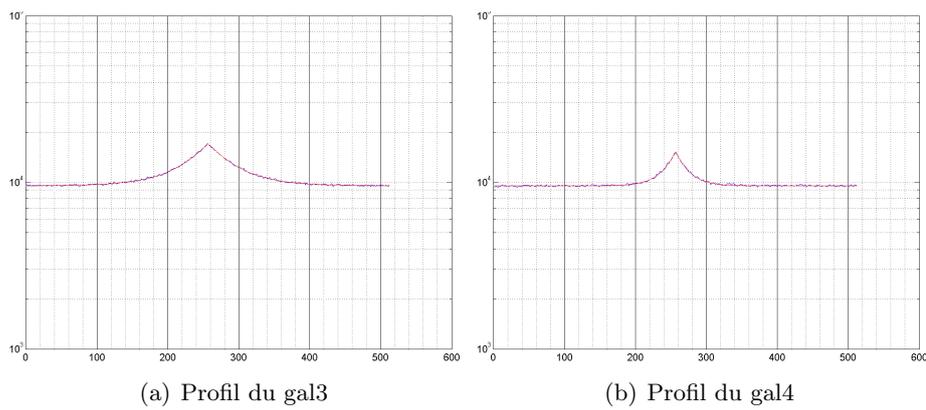


FIG. 4.4 – Profil de lumière des galaxies générées artificiellement

la PSF, du coup on obtient des indices de Sérsic assez bas et l'erreur est notamment plus forte si le profil de la galaxie est proche du profil théorique de Vaucouleurs ($n = 4$).

Nous pourrions adapter l'algorithme pour obtenir un indice de Sérsic plus correct en donnant un poids plus important aux ailes du profil. On calculerait dans l'expression 2.20 la variance autour de chaque pixel comme une valeur proportionnelle à celle de l'image pour le même pixel après la soustraction du fond de ciel, on aurait ainsi :

$$\sigma_{x,y} \propto (im_{x,y} - b) \quad (4.2)$$

La plupart des galaxies analysées présentent une erreur relative moyenne inférieure à 0.5, typiquement environ à 0.20 considérant seulement les pixels de l'image au dessus du seuil calculé dans l'étape de seuillage. La figure 4.5 illustre l'histogramme de l'erreur obtenu après l'analyse d'une centaine des images de galaxies réelles.

En ce qui concerne les images mal modélisées, nous avons remarqué que la plupart d'entre elles présentent des caractéristiques spéciales, c'est-à-dire que ce sont soit des galaxies avec un noyau déplacé, soit avec plusieurs noyaux ou des particularités similaires.

En outre, une mauvaise modélisation a eu lieu pour des images mal nettoyées, contenant des restes d'autres galaxies voisines, l'algorithme a interprété ces résidus comme des composantes de la galaxie à analyser et donc il a calculé de faux paramètres.

Pour conclure, la figure 4.6 montre la distribution des temps de calcul que l'algorithme a pris pour chaque image. Le temps de calcul suit une loi exponentielle décroissante et dans la majorité des cas, ce temps est inférieur à une minute comme on peut l'apprécier sur la figure. Il faut remarquer que l'algorithme a été développé sous MATLAB donc il peut être optimisé en écrivant le code dans un autre langage de programmation, à savoir C ou C++. Ensuite, le temps de calcul rapporté est obtenu à partir des simulations sous une machine dont le processeur avait une fréquence d'horloge de 1.73 GHz et une mémoire RAM de 0.99 GB.

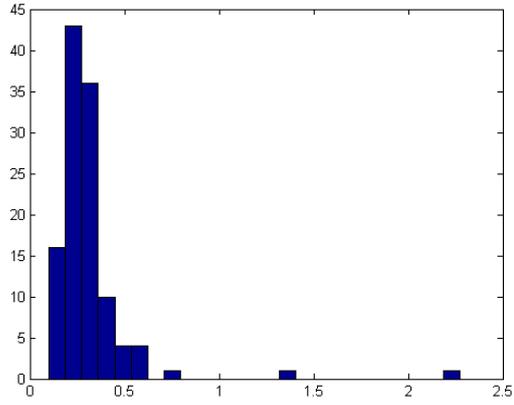


FIG. 4.5 – Histogramme de l’erreur sur une centaine de galaxies réelles

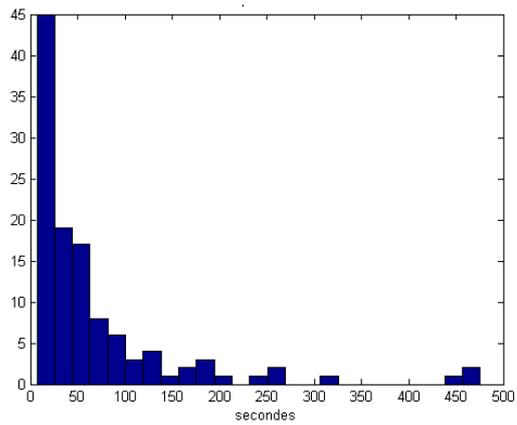


FIG. 4.6 – Histogramme du temps de calcul

Chapitre 5

Conclusions et perspectives

À l'issue de ce travail, nous obtenons un système capable de modéliser les galaxies. Nous nous limitons à réaliser l'ajustement du profil selon un modèle du bulbe de Sérsic sachant que cela peut rester insuffisant à cause de la complexité du profil que présentent la majorité des galaxies, à savoir anneaux, barres, bras spiraux ...

Nous avons étudié tout le processus qui nous mène jusqu'à la génération du modèle final du profil de la galaxie, en gérant le temps de calcul, l'adaptation des conditions initiales et la méthode du seuillage qui reste étroitement liée à l'amélioration du temps de calcul. Par ailleurs, nous avons proposé l'application d'une nouvelle méthode, reposant sur la théorie des courbes de niveau, qui semble très utile dans la décomposition de la galaxie en nuage et noyau ainsi que dans la détection d'éléments plus complexes présents dans son profil de brillance.

Pour chacune des images présentes dans la base de données, nous avons fait la décomposition de la galaxie en ses composantes, nuage et noyau, à partir de ces résultats, nous avons calculé les conditions initiales pour lancer le programme qui grâce à l'algorithme de Levenberg-Marquardt nous donne les paramètres qui décrivent la galaxie analysée. Nous avons calculé la fonction à minimiser, χ^2 , ensuite, nous avons analysé les résultats au niveau de l'ajustement et du temps mis dans le calcul qui décrivent la performance de chacune des techniques.

Par ailleurs le temps de calcul suppose une contrainte à remarquer car les bases d'images sont de plus en plus grandes et le processus d'ajustement devrait calculer plusieurs galaxies par minute pour que cela reste assez puissant sans perte de robustesse. Selon les résultats obtenus dans la section 4 nous avons pu remarquer que l'exactitude des conditions initiales est un point très important dans l'algorithme ainsi que les bornes utilisés pour les paramètres. Le travail dans ce domaine pourrait donc améliorer l'efficacité de notre algorithme, le précalcul d'un maximum de paramètres et son utilisation pour obtenir les conditions initiales données au programme qui est

responsable de l'ajustement nous mèneront à des résultats prometteurs. Par exemple, l'utilisation de la relation B/T pourrait sûrement nous aider puisqu'elle est très corrélée avec l'indice de Sérsic et, en conséquence, avec le type de la galaxie.

Une autre approche sur le problème du temps de calcul à proposer serait la génération des modèles numériques à partir desquels on pourrait calculer les ajustements et obtenir les paramètres des galaxies. En simplifiant les calculs des modèles réalisés de manière analytique lors de ce stage, on pourrait obtenir un temps de calcul environ cinq fois inférieur au temps actuel.

Dans la suite de l'étude nous nous proposons de travailler sur l'ajustement des modèles plus complexes ajoutant la possibilité de la détection de barres, anneaux et bras spiraux, la méthode des courbes de niveau proposée dans la section 2.2.1 semble intéressante dans ce but. La décomposition de la galaxie en bulbe et disque obtenue à partir de cette méthode sera utile pour le calcul des différences entre les angles de rotation du nuage et du noyau qui nous donneront un premier indice pour trouver une barre et la conception ultérieure d'un algorithme de détection de celle-ci. En outre, l'analyse des résultats de l'ajustement et particulièrement de l'image des résidus obtenus conduira à des résultats sûrement intéressants sur la détection des bras spiraux et permettra le développement d'un algorithme d'analyse des caractéristiques de ceux-ci.

Un point important pour les prochains travaux est que nous ne devons pas oublier que dans cette première étape nous utilisons uniquement des galaxies bien résolues tandis que nous traiterons par la suite des images dégradées de façon plus significative par la réponse impulsionnelle (PSF). Cela impliquera une augmentation du temps de calcul dû à la convolution par la réponse impulsionnelle à chaque itération, que l'introduction de la PSF implique.

Bibliographie

- [1] Simard, L., Willmer, Ch., Vogt, N., Sarajedini, V., Phillips, A., Weiner, B., Koo, D., Im, M., Illingworth, G., Faber, S. M. *The DEEP Groth Strip Survey II. Hubble Space Telescope Structural Parameters of Galaxies in the Groth Strip*, The Astrophysical Journal, 2002.
- [2] Allen, P.D., Driver, S.P., Graham, A.W., Cameron, E., Liske J., De Propris, R. *The Millennium Galaxy Catalogue : bulge-disc decomposition of 10 095 nearby galaxies*, MNRAS, 371, 2, 2006.
- [3] Peng, C. Y., Ho, L. C., Impey, Ch. D., Rix, H. *Detailed Structural Decomposition of Galaxy Images*, The Astrophysical Journal, 2002.
- [4] Pignatelli, E., Fasano G., Cassata P. *GASPHOT : a tool for Galaxy Automatic Surface PHOTometry*, Astronomy and astrophysics, 2006.
- [5] Baillard, A. *Projet EFIGI : Classification automatique de galaxies*, 2005.
- [6] Liénoù, M. *Classification morphologique des galaxies*, 2005.
- [7] De la Calleja J., Fuentes O. *Machine learning and image analysis for morphological galaxy classification*, Monthly notices of the Royal Astronomical Society, 2004.
- [8] Ciotti, L. *Stellar systems following the $R^{1/m}$ luminosity law*, Astronomy and Astrophysics, 1991.
- [9] Desolneux, A., Moisan, L., Morel, J.M. *Edge detection by the Helmholtz principle*, Journal of Mathematical Imaging and Vision, 2001.
- [10] Rosin, P. L. *Unimodal Thresholding*, Pattern Recognition, 2001.
- [11] Bloch, I., Gousseau, Y., Maître, H., Matignon, D., Pesquet-Popescu, B., Schmitt, F., Sigelle, M., Tupin, F. *Le traitement des images*, Département Traitement du Signal et de l'image, école Nationale Supérieure des Télécommunications, 2004-2005.
- [12] Pratt, W. K. *Digital Image Processing*, Addison Wesley, 2001.
- [13] Roweis S. *Levenberg-Marquardt Optimization*.
- [14] Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery B. P. *Numerical Recipes in C*, Cambridge University Press, 2002.

- [15] Wells, D. C., Greisen, E. W., and Harten, R. H. *FITS : A Flexible Image Transport System*, Astronomy and Astrophysics Supplement Series, 1981.