

Utilisation d'images de référence pour la compression d'images Application à la mission GAIA

Albert BIJAOU¹, Yves BOBICHON², Benjamin MASSART¹, François MIGNARD¹, Julien PARNAUDEAU¹

¹Département CERGA, UMR CNRS 6527, Observatoire de la Côte d'Azur, B.P.4220 06304 Cedex 4, France

²Alcatel-Space Industries, BP. 99 - 06156 Cannes-la-Bocca Cedex, France

bijaoui@obs-nice.fr, Yves.Bobichon@space.alcatel.fr,
ben_massart@yahoo.fr, mignard@obs-azur.fr, juparnaudeau@airfrance.fr

Résumé – La mission GAIA de l'Agence Spatiale Européenne devra mesurer tous les astres jusqu'à la magnitude 20 avec une précision de plusieurs ordres de grandeur meilleure que celle obtenue avec le satellite Hipparcos. Le débit d'information étant totalement incompatible une transmission complète des données au sol, un taux de compression très élevé est nécessaire. La compression des images sans *a priori* à l'aide de la transformation de Haar conduit à un taux de compression très insuffisant. La compression à partir des images de synthèse formées par les objets connus permet un facteur de compression adéquat. Cela devrait permettre la réalisation d'une carte du ciel avec une résolution d'au moins un ordre de grandeur meilleure que celle des cartes disponibles aujourd'hui.

Abstract – The GAIA mission of the European Space Agency is planned to measure all the celestial sources up to magnitude 20th with an accuracy many orders above the one obtained with the Hipparcos satellite. The information rate is incompatible with a full transmission of the data to the ground. Thus a very large compression factor is necessary. An optimal image compression using the Haar transform, without prior knowledge led to an insufficient factor. Using synthetic images which simulated the observations the compression factor reached the compatibility with the transmission rate. This would allow the astronomers to build from the GAIA mission a sky map with a resolution increased by a factor at least ten.

1 La mission GAIA

Le satellite HIPPARCOS de l'ESA a été la première mission spatiale permettant la mesure précise des positions, des distances et des mouvements propres des étoiles. Le catalogue final de 120000 objets a été publié en 1997 [3]. De très nombreux résultats astrophysiques ont été obtenus grâce au gain significatif en précision de ce catalogue par rapport aux mesures antérieures.

La mission GAIA, également de l'ESA, devra mesurer les positions et les magnitudes de tous les astres jusqu'à la magnitude 20 avec une précision de plusieurs ordres de grandeur meilleure que celle obtenue avec HIPPARCOS [2]. Le satellite sera composé de 3 télescopes. Au foyer de deux d'entre eux 250 CCD de 2780×2150 pixels seront installés. Le débit d'information sera totalement incompatible avec la transmission au sol de toutes les valeurs des pixels des images. Une première réduction massive de l'information consistera dans la sélection des régions à mesurer. Une compression efficace de ces données permettra d'obtenir un débit compatible avec la transmission au sol. Quatre détecteurs CCD formant l'*Astro Sky Mapper* (ASM) devront assurer la détection des objets qui seront mesurés ensuite plusieurs fois sur les CCD du plan focal.

L'ensemble du ciel sera observé une centaine de fois par l'ASM. Son débit brut correspond à 20Gbs, ce qui est très loin de la bande passante disponible. Cette infor-

mation n'est pas très utile pour l'amélioration des mesures des objets détectés, mais ce seront les seules données qui concerneront le ciel dans son ensemble. Leur transmission pourrait permettre la construction d'une carte du ciel à la résolution minimum de 0,1 arcsec. Pour l'envisager, il est indispensable de compresser ces images avec un facteur très important.

2 Simulation des images.

Pour étudier les possibilités de compression optimale, nous avons développé une procédure complète de simulation des images.

Nous extrayons une image de fond assez représentative dans la collection des images du télescope spatial Hubble (HST). Les étoiles sont éliminées à l'aide d'un filtre de la médiane. L'image résiduelle est convoluée avec un fonction d'étalement de l'image (PSF) moyenne de GAIA. Ceci représente le fond caractéristique de l'image à traiter. Plusieurs fonds de ce type sont créés pour tester leur influence sur le taux de compression.

Un catalogue représentatif d'étoiles est fabriqué par tirage aléatoire, en tenant compte des fonctions de répartition selon le type d'étoiles, la magnitude et la couleur. Une PSF est calculée pour chaque étoile en tenant compte de son spectre, la tache d'Airy des télescopes variant avec la longueur d'onde. On tient compte de la traînée de la lec-

ture des CCD. Une image de synthèse est fabriquée par superposition des PSF aux positions des étoiles en tenant compte de leur éclat. Cette image est additionnée à l'image du fond.

L'image obtenue représente l'espérance du signal reçu. A l'aide d'un algorithme de simulation de bruit de Poisson, cette image est transformée en un compte de photons par pixel du CCD. Un bruit gaussien uniforme est ajouté pour tenir compte du bruit de lecture et du bruit thermique. Quelques traces dues au rayonnement cosmique sont ajoutées pour obtenir l'image finale. Sur la figure 1 une image représentative des simulations est tracée. Le fond correspond à la galaxie M100 et le catalogue d'étoiles provient de l'analyse de l'amas globulaire 47 Tucanae. Les données initiales proviennent d'observations du HST. La richesse de l'image obtenue correspond aux cas extrêmes des observations avec la mission GAIA.

3 La compression des images sans *a priori*.

Compte tenu de la très grande dynamique du signal qui existe dans les images à traiter, il est indispensable de transformer la valeur de l'intensité mesurée v dans un pixel donné de manière à obtenir une variable t de variance constante. Nous appliquons une transformation d'Anscombe généralisée à la somme d'un bruit de Poisson avec un bruit gaussien [6] :

$$t = \frac{2}{\alpha} \sqrt{\alpha v + \frac{3}{8}\alpha^2 + \sigma^2 - \alpha g} \quad (1)$$

g est la moyenne du bruit gaussien et σ^2 sa variance. α est le facteur de conversion du nombre de photons en signal mesuré ($v = \alpha n$ sans le bruit gaussien).

L'image résultante est transformée avec la transformation de Haar. Cette transformation est bien adaptée à la compression des images astronomiques. Le relevé complet du ciel (*Digital Sky Survey*) est accessible par internet grâce à ce mode de compression [8].

Un seuillage est effectué pour sélectionner les coefficients les plus significatifs de l'image [1]. Le seuil T_i à l'échelle i est de la forme $k - \sigma$, où σ est l'écart-type du bruit. Après application de la transformation d'Anscombe généralisée, σ vaut 1. Le facteur k est lié à la probabilité de détection d'un coefficient supérieur à la valeur mesurée, dans l'hypothèse d'une image localement uniforme. Si k est trop petit (inférieur à 3) de nombreuses détections ne sont dues qu'au bruit. Si k est trop élevé (supérieur à 5), le seuillage élimine trop de valeurs significatives. Nous avons choisi une valeur typique de 3 par la suite. De nombreux artefacts subsistent, ils pourront être éliminés dans les traitements ultérieurs.

Les données à transmettre sont codées par plan de bits selon une méthode hiérarchique dissymétrique à 4-bits [4] permettant d'obtenir un taux proche de l'entropie. Le taux de compression est le rapport est le nombre de bits obtenus à partir de l'image codée à la sortie du détecteur sur le nombre total de bits nécessaire pour reconstruire l'image.

La restauration directe des images introduit un effet de

bloc très gênant, lié à la mise à zéro de coefficients en ondelettes. Nous avons introduit une méthode d'inversion basée sur une régularisation avec la contrainte de Tikhonov de minimisation des gradients à chaque échelle [1]. Une contrainte de positivité est aussi appliquée. Ceci permet d'améliorer sensiblement la qualité des résultats.

Le taux de compression obtenu en moyenne sur les images simulées est de l'ordre de 50 [7]. Ce taux reste très insuffisant pour proposer la transmission de ces données afin de construire une carte du ciel.

4 Compression avec une carte préalable.

La quasi totalité des sources célestes ont des positions et des magnitudes ne variant que faiblement. Il est parfaitement possible d'avoir avant la mission un catalogue presque complet des objets à mesurer. Par conséquent il est possible de construire à tout instant une image réaliste de la scène vue par le satellite. Cette image est construite selon la procédure décrite dans la section 2. Aucun bruit bien sûr n'est introduit à la fin de la procédure.

Sur cette image on applique les mêmes transformations que celles appliquées aux images observées : transformation d'Anscombe et transformation de Haar. On effectue la différence entre les coefficients résultant des deux séries de transformation, puis on seuille cette valeur à $3 - \sigma$.

De nombreuses simulations ont été réalisées pour étudier l'influence de plusieurs paramètres sur le taux de compression :

- La précision de la connaissance de la carte ;
- L'influence du fond sous-jacent ;
- Les variations des objets du catalogue.

Si un taux de compression moyen de 520 a été obtenu avec une détection à $3 - \sigma$, ce facteur peut se réduire à 80 dans le rare cas d'un champ d'un amas globulaire très riche ou d'un galaxie brillante.

L'étude a considéré toutes les sources connues à la magnitude 24,5. La réduction à la magnitude 18 de la connaissance du champ réduit fortement le facteur de compression. De même la méconnaissance de la couleur des étoiles conduit à une dégradation sensible du facteur de compression en raison des variations de la fonction d'étalement des images avec la longueur d'onde.

Un autre facteur de réduction du taux de compression réside dans les erreurs induites par les variations de position des objets. Même si le catalogue initial a des positions d'étoiles suffisamment précises, au cours de la mission les étoiles se déplacent en raison de leur mouvement propre et de leur parallaxe. Si on ne tient pas compte de ces modifications de position, le facteur de compression est réduit à une centaine, ce qui est très insuffisant.

Les variations lumineuses des astres induisent également une réduction du taux de compression. Dans la plupart des cas les variations sont périodiques et bien prédictibles. Plusieurs objets présentent des variations chaotiques parfaitement imprédictibles, mais ces objets sont rares et leur prise en compte n'affectera pas sensiblement la statistique du taux de compression des images.

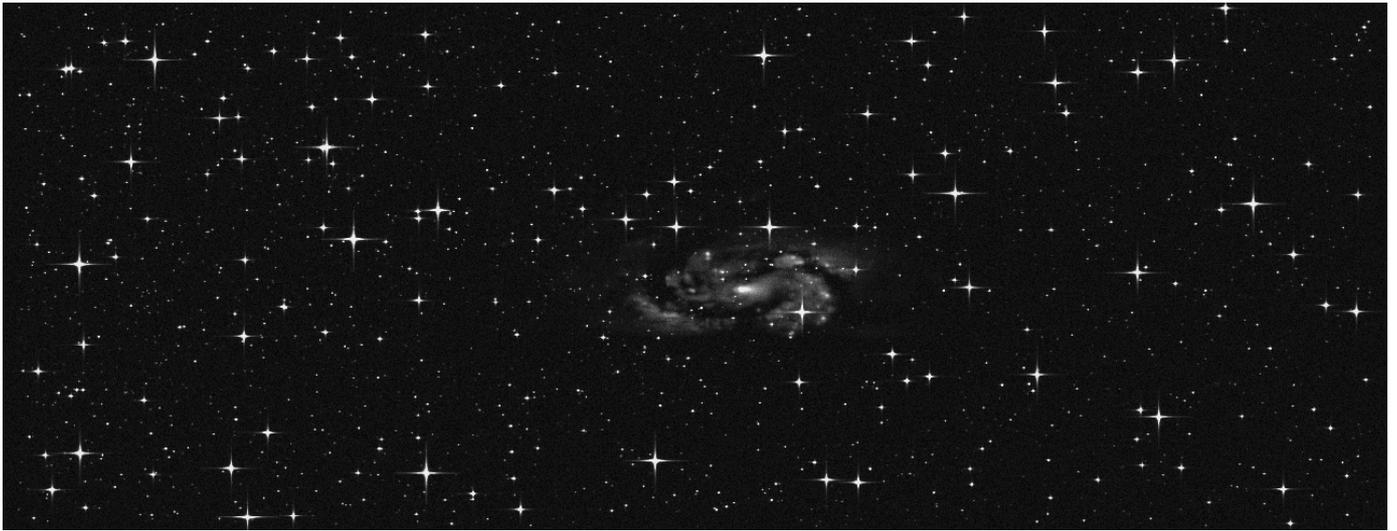


FIG. 1 – Un exemple d’image simulée pour l’*Astro Sky Mapper* de GAIA.

Enfin quelques objets non préalablement identifiés peuvent être détectés. Il pourra s’agir par exemple de supernovae ou d’astéroïdes. Leur rareté est telle que le taux moyen de compression ne sera pas modifié.

5 Utilisation d’un catalogue actualisé.

La compression en exploitant une carte préalable est donc limité par la connaissance que l’on a *a priori* des objets de l’image. La construction d’une cartographie du ciel avant le lancement de la mission n’augmente que d’une manière faible le taux de compression par rapport à une compression directe de l’image.

Les CCD du plan focal de la mission GAIA ont pour tâche d’obtenir sur tous les objets détectés les mesures les plus précises, celles la même dont on a besoin pour obtenir le facteur de compression maximal. La construction de la carte doit donc tenir compte de ces mesures. Dans ces conditions le facteur de compression de 520 peut être atteint dans les champs usuels du ciel. Il descend à 420 dans les champs les plus denses testés.

Le débit est alors à peine plus grand de celui qui serait lié à une image d’un bruit gaussien, avec un seuil à $3 - \sigma$, pour laquelle nous avons trouvé un facteur de compression d’environ 600. La comparaison entre les facteurs de compression d’une image observée avec une image de bruit nous donne une idée du faible taux d’information réelle transmise par les images. Si on souhaite accroître le taux de compression de manière significative il faut augmenter le seuil. Dans le tableau 1 nous avons indiqué les variations du facteur de compression avec le seuil à partir de 1000 expériences numériques. Nous avons aussi indiqué le rapport signal sur bruit moyen.

Il apparaît ainsi qu’avec un seuillage à seulement $3,5 - \sigma$ le débit devient compatible avec la bande passante, le rapport signal sur bruit n’étant pas pour autant trop dégradé. Jusqu’à un seuillage à $5 - \sigma$, la cartographie obtenue reste

TAB. 1 – Évolution du taux de compression en fonction du seuil. *Seuil* est le seuil considéré pour les coefficients en ondelettes. *Taux* est le taux de compression moyen obtenu sur un millier d’images simulées. *RSB* est le rapport signal sur bruit moyen, exprimé en décibels, obtenu après restauration régularisée avec une contrainte sur le gradient.

<i>Seuil</i>	<i>Taux</i>	<i>RSB</i>
$3 - \sigma$	520	29.90
$3,5 - \sigma$	1260	29.77
$4 - \sigma$	2076	29.74
$5 - \sigma$	2938	29.72

d’un grand intérêt scientifique.

Une procédure de compression avec une carte actualisée à bord peut donc être mise en œuvre à bord du satellite GAIA. Dans cette procédure les autres observations focales servent à construire à bord des images synthétiques de chaque champ observé, seules les informations sur les objets non mesurés devant être transmises (figure 2) [5].

6 Conclusion.

Nous avons montré dans cette communication que, grâce à l’utilisation de cartes modifiées dynamiquement pendant la mission, il est possible d’atteindre un taux de compression permettant la réalisation d’une carte du ciel avec une résolution d’au moins un ordre de grandeur meilleure que celle des cartes étendues disponibles aujourd’hui.

Sur un balayage du ciel, la qualité de l’image restaurée ne sera pas peut être pas assez précise pour une exploitation des objets non identifiés parallèlement par l’instrumentation focale, mais chaque zone du ciel sera vue plus d’une centaine de fois lors de la mission. Ceci permettra donc en théorie d’avoir un gain d’une dizaine en rapport signal sur bruit.

Un programme spécifique de décompression des images pourrait alors être mis en œuvre afin de tenir compte de toutes les variations mesurées au fur et à mesure de la

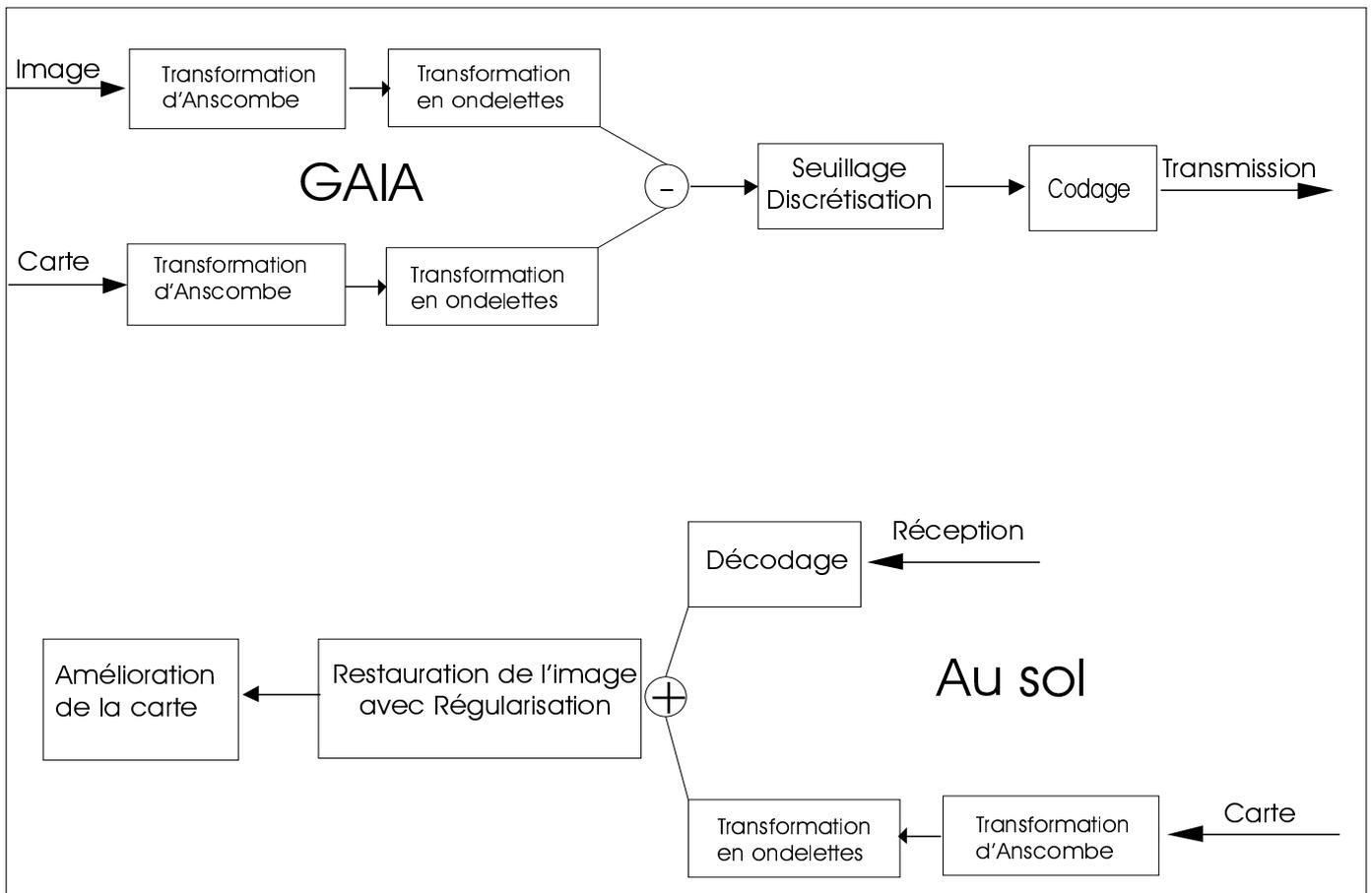


FIG. 2 – Traitement proposé pour la compression des données de l’*Astro Sky Mapper* de GAIA.

mission. L’attitude du satellite sur le ciel étant différente d’une observation à l’autre, des problèmes très délicats d’échantillonnage sont aussi posés.

Dans l’hypothèse où cette cartographie systématique du ciel pourra être prise en compte dans la mission GAIA, l’exploitation globale des données comprimées conduira à restaurer une image dont la résolution sera très supérieure aux grands atlas astronomiques actuels ou ceux qui dont la publication est prévue dans la présente décennie.

Références

- [1] Y. Bobichon, A. Bijaoui. A regularized image restoration algorithm for lossy compression in astronomy. *Experimental Astronomy*, 7, pp.239-255, 1997.
- [2] ESA. GAIA, Composition, Formation and Evolution of the Galaxy, ESA-SCI 4, 2000.
- [3] ESA. The HIPPARCOS and TYCHO catalogues. SP-1200, vols. 1-17, 1997.
- [4] L. Huang, A. Bijaoui. Astronomical Image data compression by morphological transformations. *Experimental Astronomy* 1 pp.311-327 1991.
- [5] B. Massart. Compression et Restauration d’images pour la mission GAIA de l’Agence Spatiale Européenne. *Rapport de stage* DEA d’électronique de l’Université de Nantes, 2000
- [6] F. Murtagh, J.L. Starck, A. Bijaoui. Image restoration with noise suppression using the Wavelet Transform II. *Astron. Astrophys. sup. ser.*, 112, pp. 179-189, 1995.
- [7] J. Parnaudeau. Étude de la compression d’images pour la mission spatiale GAIA. *Rapport de stage*. Institut Supérieur de Modélisation et de leurs Applications. Aubière, 1999.
- [8] R. White, M. Postman, M. Lattanzi. Compression of the guide star digitised Schmidt plates. *Digitised sky surveys II*, pp. 167-175, ed. H. Mac Gillivray, Kluwer, 1992.