

Restauration d'images dans le cas d'un modèle Poisson-Gamma-Gaussien

Céline THEYS, Henri LANTÉRI

Laboratoire Universitaire d'Astrophysique de Nice
UMR6525, CNRS/Université de Nice Sophia-Antipolis
Céline.Theys@unice.fr, Henri.Lanteri@unice.fr

Résumé – Le sujet de ce papier est la restauration d'images pour un modèle Poisson Gamma Gaussien, modèle le plus adéquat pour des images astrophysiques faible flux acquises avec des caméras L3CCD. Après avoir développé la distribution de probabilité pour les données L3CCD, un algorithme itératif de déconvolution est construit pour estimer au sens du MAP l'objet observé. Des résultats de simulation mettent en évidence l'intérêt des capteurs L3CCD pour l'imagerie très faible flux.

Abstract – The subject of this paper is the restoration of Poisson Gamma Gaussian images, the most realistic model for images acquired with L3CCD cameras. Once the probability density of L3CCD data has been derived, an iterative algorithm for MAP estimation of the object is developed. Some simulation results point out the interest of using L3CCD cameras for very low intensity images.

1 Introduction

Le problème traité dans ce papier est la restauration d'images astrophysiques acquises avec des capteurs L3CCD (Low Light Level CCD), une technologie très récente de capteurs CCD. Jusqu'à présent, l'acquisition de données astrophysiques se faisaient grâce à des capteurs CCD introduisant un bruit de lecture non négligeable dans le cas de données faible flux. La motivation principale et communément admise pour l'utilisation de capteurs L3CCD est l'amplification importante des données avant lecture, permettant de négliger le bruit de lecture. Cependant, des travaux récents ont mis en évidence le caractère stochastique du gain d'amplification, nécessitant une étude approfondie de son effet dans le cas de données avant amplification très faibles. Nous proposons dans ce papier de donner un modèle probabiliste complet des données L3CCD et de développer un algorithme de reconstruction pour un tel modèle. Un résultat de simulation permettra de comparer une reconstruction CCD versus L3CCD pour des données très faible flux.

2 Modèle Poisson-Gamma-Gaussien de l'image

On utilisera des lettres capitales pour des tableaux $N \times N$, des caractères gras pour les vecteurs $N \times 1$, l'indice i pour le pixel i de l'image, ordonnée lexicographiquement.

Les différentes étapes du processus physique conduisant aux données L3CCD peuvent être schématisées comme suit:

$$\mathbf{x} \xrightarrow{\text{convolution}} H\mathbf{x} \xrightarrow{\text{Poisson}} \mathbf{n} \xrightarrow{\text{Gamma}} \mathbf{y} \xrightarrow{\text{Gaussien}} \mathbf{z} \quad (1)$$

avec:

$$\mathbf{z} = \mathbf{y} + \mathbf{b}, \quad \mathbf{y} \sim \mathcal{G}(\mathbf{n}, 1/G), \quad \mathbf{n} \sim \mathcal{P}(H\mathbf{x}), \quad \mathbf{b} \sim \mathcal{N}(0, \sigma^2) \quad (2)$$

L'étape de convolution est la modélisation de l'observation de la lumière issue du ciel se propageant à travers l'atmosphère et observée par un télescope. Cette première transformation est notée $H\mathbf{x}$ sous forme matricielle où H est une matrice Toeplitz bloc Toeplitz. Chaque ligne de H comporte les valeurs "dépliées" du tableau de la *PSF* (Point Spread Function) de l'ensemble télescope/atmosphère, on se placera toujours dans le cas où l'on peut calculer ce produit de convolution par une transformée de Fourier discrète.

Les deux étapes suivantes sont la photo conversion réalisée par le capteur puis l'amplification stochastique liée au capteur L3CCD. En effet la structure d'un capteur L3CCD est la même que celle d'un capteur CCD mais avec un nombre de registre étendu, à chaque étage, un électron produit un électron supplémentaire avec une probabilité p , p est très faible mais le nombre de registres N est très grand, ce qui donne un gain moyen G pour le capteur égal à $(1-p)^N$, il est très difficile d'obtenir la loi exacte des processus d'avalanche, type "branching processes" et pour l'instant des recherches sont en cours. Basden et al. [3] ont montré l'adéquation par simulations numériques entre les données issues d'un capteur L3CCD et une loi Gamma de paramètres n et $1/G$. Nous nous proposons d'utiliser ce résultat dans la suite.

Enfin un bruit de lecture Gaussien s'ajoute (*read out noise*), indépendant du processus de Poisson, de moyenne nulle et dont la valeur numérique de la variance est typique du capteur L3CCD considéré. Comme dit en introduction ce bruit est négligé aussi bien pour l'acquisition CCD que pour l'acquisition L3CCD ce que nous avons fait dans [7].

Nous nous proposons de l'inclure ici dans le modèle.

La densité de probabilité de z_i conditionnellement à \mathbf{x} est alors:

$$p(z_i|\mathbf{x}) = \sum_{y_i} p(z_i|y_i)p(y_i|\mathbf{x}) \quad (3)$$

$p(y_i|\mathbf{x})$ se décompose suivant:

$$\begin{aligned} p(y_i|\mathbf{x}) &= \sum_n p(y_i|n)p(n|\mathbf{x}) \\ &= \sum_{n \geq 0} p(y_i|n)p(n|\mathbf{x}) + p(y_i|n=0)p(n=0|\mathbf{x}) \end{aligned} \quad (4)$$

soit en considérant les relations 1 et 2:

$$\exp(-(H\mathbf{x})_i) \exp(-y_i/G) \sum_{n=1}^{\infty} \frac{(H\mathbf{x})_i^n}{n!} \frac{y_i^{n-1}}{\Gamma(n)G^n} + \exp(-H\mathbf{x})_i \delta_{y_i,0}. \quad (6)$$

Le premier terme de l'éq. 6 est l'approximation proposée dans [2], le second terme est rajouté pour compléter la densité en $n_i = 0$, c'est à dire pas de photons dans le pixel i avant avalanche, cas non modélisable par une loi Gamma. Dans ce cas, il est évident qu'il n'y aura pas de photons en sortie, soit $y_i = 0$, c'est donc la probabilité d'avoir $n_i = 0$ à la sortie du processus de Poisson soit $\exp(-H\mathbf{x})_i$. Il est important de noter que la loi Gamma est une loi continue et que la variable aléatoire en sortie du processus d'avalanche est une variable discrète, il faut donc calculer le coefficient de normalisation de la loi Gamma, on aboutit à une expression probablement non exhaustive qui est:

$$\sum_{y \geq 0} \exp(-y/G) \frac{y^{n-1}}{\Gamma(n)G^n} = \frac{1}{\Gamma(n)G^n} \text{Polylog}(1-n, \exp(-1/G)) \quad (7)$$

où $\text{Polylog}(x, y) = \sum_{k=1}^{\infty} \frac{y^k}{k^x}$, cette expression n'a pas à notre connaissance de simplification pour des variables z réelles et entières. Cependant des simulations numériques donnent 1 quelque soient les valeurs de $n \geq 1$ et G entiers réels positifs, nous considèrerons dans la suite que la densité Gamma discrète telle que nous l'utilisons est normée.

La probabilité de la donnée z_i conditionnellement à y_i est donnée par la densité de probabilité Gaussienne de moyenne y_i et de variance σ^2 et finalement:

$$\begin{aligned} p(z_i|\mathbf{x}) &= \frac{1}{\sigma\sqrt{2\pi}} \exp(-(H\mathbf{x})_i) \sum_{n=1}^{\infty} \frac{(H\mathbf{x})_i^n}{n!\Gamma(n)G^n} \\ &\quad \left(\sum_{y_i} \exp\left(-\frac{(z_i - y_i)^2}{2\sigma^2}\right) y_i^{n-1} \exp(-y_i/G) + \right. \\ &\quad \left. \exp(-H\mathbf{x})_i \delta_{y_i,0} \right) \end{aligned} \quad (8)$$

Pour donner une idée du comportement de cette loi que nous noterons en abrégé PGG, on considère le cas simple où $H = I$, la matrice identité, c. à d. le cas où la valeur de z_i dépend uniquement de la valeur de x_i and $p(z_i|\mathbf{x}) = p(z_i|x_i)$. Des paramètres physiques réalistes ont été donnés pour G et σ soit 400 et 25 respectivement. La log vraisemblance $\log p(z_i|x_i)$ est tracée sur la figure 2. La

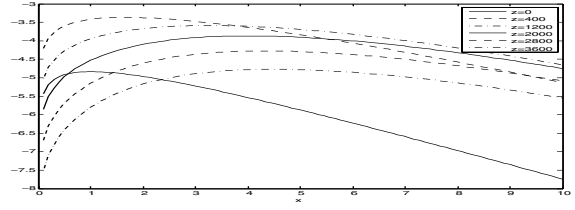


Figure 1: $\log p(z_i|x_i)$ pour $H = I$

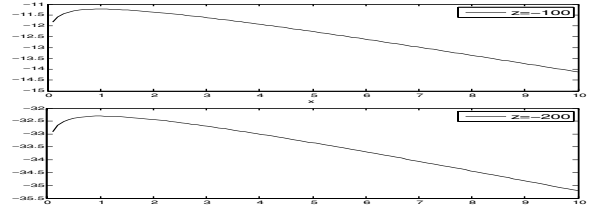


Figure 2: $\log p(z_i|x_i)$ pour $H = I$

figure de gauche pour des valeurs de z_i positives, celle de droite pour deux valeurs négatives de z_i illustrant deux "tendances":

- Pour des valeurs de z_i supérieures à G , correspondant à des valeurs les plus probables de x_i supérieures à 1, prépondérance du phénomène d'amplification.
- Pour des valeurs de z_i inférieures à G , correspondant à une valeur la plus probable de x_i égale à 1, prépondérance du bruit additif Gaussien.

3 Algorithme de déconvolution

Le problème est de retrouver \mathbf{x} à partir de \mathbf{y} , c'est un problème mal posé car la matrice H est mal conditionnée, ce qui est le résultat du fait que l'observation de \mathbf{x} à travers le télescope a supprimé l'information ou tout du moins l'a fortement diminué en dehors de la bande passante du télescope. Dans ce cas une solution du type maximum de vraisemblance, si on peut la calculer, n'est pas satisfaisante et une approche classique est de régulariser la solution en ajoutant un a priori de douceur sur la solution et de faire une estimation au sens du MAP (Maximum A Posteriori), de plus l'utilisation d'un algorithme itératif de descente (dans ce cas on minimisera l'inverse du log de la densité a posteriori noté J) est impératif pour inclure des contraintes d'inégalité sur la solution, en l'occurrence ici la contrainte de positivité sur \mathbf{x} et la contrainte de conservation de l'énergie, i.e $\sum_i x_i = \sum_i y_i$.

La fonction J à minimiser s'écrit alors:

$$J(\mathbf{x}, \gamma) = J_1(\mathbf{x}) + \gamma J_2(\mathbf{x}). \quad (9)$$

Où $J_1(\mathbf{x}) = -\sum_i \log p(z_i|\mathbf{x})$ est le terme d'attache aux données, $J_2(\mathbf{x})$ le terme d'a priori et γ , le facteur de régularisation.

En écrivant le lagrangien correspondant à la minimisation de $J(\mathbf{x})$ sous la contrainte $\mathbf{x} \geq 0$, [5, 6], on obtient la

condition de Kuhn Tucker:

$$[\nabla J(\mathbf{x}^*, \gamma)]_i g_i(\mathbf{x}^*) = 0 \quad \forall i \quad (10)$$

Où \mathbf{x}^* est la valeur qui minimise $J(\mathbf{x}, \gamma)$ et g toute fonction qui exprime la contrainte $x \geq 0$, i.e $g(x) \geq 0$ quand $x \geq 0$, contrainte inactive, et $g(x) = 0$ quand $x = 0$, le choix le plus simple consiste à prendre $g(x) = x$. On propose d'utiliser alors un algorithme de minimisation basé sur les propriétés des méthodes de descente, on peut alors écrire l'équation de réactualisation de \hat{x}_i :

$$x_i^{k+1} = x_i^k + \alpha_i^k x_i^k (-\nabla J(\mathbf{x}^k, \gamma))_i \quad (11)$$

L'estimé initial $\hat{\mathbf{x}}^0$ est choisi de manière à satisfaire les contraintes, $\alpha_i^k > 0$ est le facteur de relaxation choisi de manière à garantir la convergence, k est le numéro de l'itération, $-\nabla J(\mathbf{x}^k, \gamma)_i$ est la direction de descente, la multiplication par x_i^k résulte des conditions de Kuhn Tucker, voir plus haut. Un schéma itératif de ce type converge vers le minimum global de $J(\mathbf{x})$ si ce dernier est convexe et si le pas de descente est choisi de manière à minimiser J . La recherche de pas se fera suivant une méthode économique de recherche de pas (Armijo, [1]).

Il est important de noter qu'il faudrait faire une étude de la convexité du critère J , ce qui implique le calcul du hessien de J , calcul complexe. Dans les exemples choisis dans les simulations, on peut cependant observer que J ne possède qu'un minimum, en tout cas sur l'espace considéré; la démonstration fera l'objet d'un autre papier.

Si l'on rajoute la contrainte de conservation de l'énergie en appelant P l'opérateur de projection sur l'espace des solutions de même énergie, on a alors:

$$x_i^{k+1} = P(x_i^{k+1}) \quad (12)$$

La convergence est maintenue si l'opérateur P est un opérateur de contraction, [8], ce qui est le cas ici. La conservation de l'énergie se fera donc à chaque itération par projection sur l'espace des solutions de même énergie, i.e en normalisant la reconstruction à $\sum_i y_i$.

Le gradient de J_1 pour le pixel i , calculé à partir de (8) est:

$$(\nabla J_1(\mathbf{x}))_i = \sum_j (h_{ji} - h_{ji} r_j), \quad (13)$$

avec

$$r_j = \frac{p_j}{q_j}, \quad (14)$$

$$p_j = \exp(-H\mathbf{x})_j \sum_{y_i} \left(\exp\left(-\frac{(z_j - y_i)^2}{2\sigma^2}\right) - \frac{y_i}{G} \right) \sum_n \frac{n(H\mathbf{x})_j^{n-1} y_i^{n-1}}{n! \Gamma(n) G^n} \quad (15)$$

et

$$q_j = \exp(-H\mathbf{x})_j \sum_{y_i} \left(\exp\left(-\frac{(z_j - y_i)^2}{2\sigma^2}\right) - \frac{y_i}{G} \right) \sum_n \frac{(H\mathbf{x})_j^{n-1} y_i^{n-1}}{n! \Gamma(n) G^n} + \exp(-H\mathbf{x})_j \exp\left(-\frac{(z_j)^2}{2\sigma^2}\right) \delta_{y_i,0} \quad (16)$$

On remarque que lorsque $\sigma^2 \rightarrow 0$, on retrouve l'expression du gradient sans bruit de lecture, [7]. A ce jour, des expressions analytiques pour les séries p_j et q_j sont à l'étude,

on peut tout de même vérifier aisément qu'elles convergent, des simulations numériques ont permis de vérifier qu'une sommation sur y_i allant jusqu'à $\max(z_i)$ est suffisante. On se propose d'utiliser un terme de régularisation J_2 de type $L1$ soit:

$$J_2(\mathbf{x}) = \sum_{i=1}^N |x_i - p_i| \quad (17)$$

où p_i est une constante. Ce terme d'a priori présente l'avantage d'être moins sensible aux valeurs aberrantes des données, donc plus robuste au bruit.

Le gradient de J_2 pour le pixel i s'écrit:

$$(\nabla J_2(\mathbf{x}))_i = \sum_{i=1}^N \text{sgn}(x_i - p_i), \forall x_i \neq p_i \quad (18)$$

où la fonction $\text{sgn}(x)$ vaut 1 pour $x > 0$, -1 pour $x < 0$. Le problème de la non différentiabilité de J_2 en $x_i = p_i$ peut être contourné, par exemple, en "faisant un raccord" $L2_i$ en $x_i = p_i$ pour rendre la fonction différentiable en tout point.

4 Application à des données astrophysiques faible flux

Cet algorithme est illustré sur une image provenant du télescope spatial Hubble (HST), fig. 3(a). Ces données ont été convoluées avec une réponse physiquement réaliste d'un télescope terrestre incluant l'effet de la turbulence atmosphérique, fig. 3(b) donnant l'image convoluée, fig. 3(d). La figure 4 compare la reconstruction d'une image CCD avec celle d'une image L3CCD dans le cas de données très faible flux, soit 1000 photons pour toute l'image. Figs. 4(a) et 4(c) sont respectivement les données bruitées CCD et la meilleure reconstruction (sans régularisation) au sens de la distance euclidienne entre l'objet réel \mathbf{x} et l'objet reconstruit $\hat{\mathbf{x}}^k$, la reconstruction étant réalisée avec l'algorithme proposé dans [4]. Figs. 4(b) et 4(d) sont respectivement les données bruitées L3CCD et la meilleure reconstruction obtenue grâce à l'algorithme proposé ici. La fig. 5 donne la moyenne et la variance de l'erreur d'estimation en fonction du nombre de photons dans l'image pour 3 algorithmes non régularisés. Les résultats RL sont ceux obtenus en utilisant le standard algorithme de Richardson Lucy sur des données CCD, les résultats GP sont ceux obtenus en utilisant l'algorithme prenant en compte le bruit Gaussien des données CCD, enfin les résultats GGP sont ceux obtenus avec l'algorithme proposé ici sur des données LLLCCD. On peut remarquer, entre autre, que les résultats sont très distincts pour que pour une image très faible flux, on notera la très faible variance de l'erreur GGP pour 1000 photons. Il est également important de signaler que le critère d'erreur quadratique est un grossier critère quantitatif et que si l'on regarde les images correspondant à des erreurs apparemment assez proches, il y a une différence qualitative très importante.

Enfin la fig. 6 compare les objets reconstruits pour l'algorithme régularisé L1 ou non régularisé. Les fig 6(a) et 6(b) sont les données respectivement avant et après

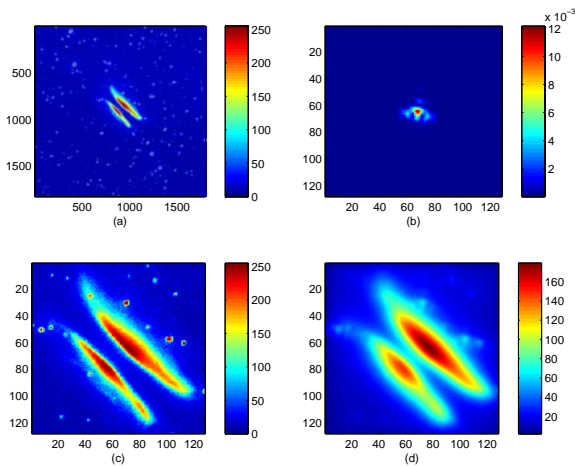


Figure 3: (a) Image HST, (b) PSF, (c) objet sélectionné, (d) résultat de la convolution

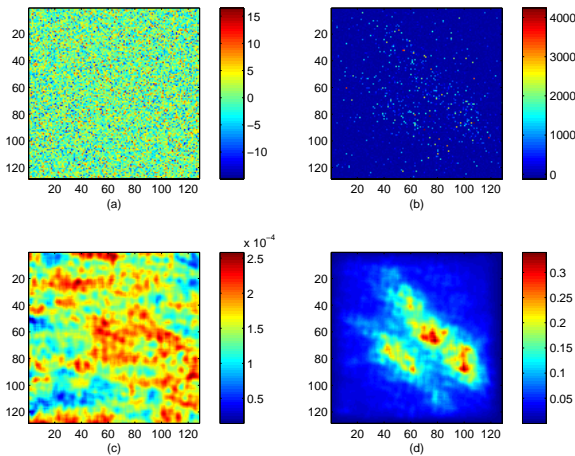


Figure 4: (a) Données CCD, 1000 photons, $\sigma^2 = 4$ (b) Données LLLCCD, 1000 photons, $\sigma^2 = 625$ (c) Meilleure reconstruction CCD, (d) Meilleure reconstruction LLL-CCD

amplification Gamma, la fig. 6(c) est la meilleure reconstruction sans régularisation et la fig. 6(d) la meilleure reconstruction avec régularisation, obtenue en réglant “à la main” le paramètre de régularisation. On voit nettement l’intérêt de la régularisation L1 sur cet exemple, qui permet d’affiner de manière significative le résultat.

5 Conclusion

Nous avons développé un algorithme de reconstruction d’images pour des capteurs L3CCD, les résultats de simulation montrent que l’on peut faire de l’imagerie avec ces nouveaux capteurs et leur intérêt non contestable pour l’imagerie très faible flux, de plus l’utilisation d’un terme de régularisation de type L1 améliore nettement la reconstruction dans cette situation “très bruitée”.

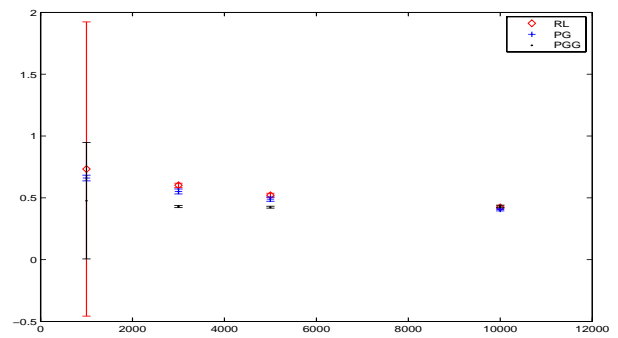


Figure 5: Moyenne et variance de l’erreur quadratique moyenne sur 50 réalisations

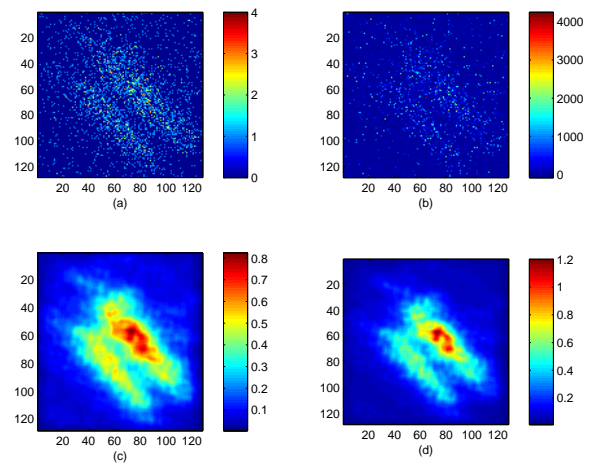


Figure 6: (a) Données avant amplification - 3000 photons (b) Données après amplification (c) $\gamma = 0$, $\epsilon = 0.44653$ (d) $\gamma = 0.3$, $\epsilon = 0.42363$

References

- [1] L. Armijo. Minimization of functions having continuous derivatives. *Pacific Journal of Mathematics*, (16):1-3, 1966.
- [2] A. G. Basden and C. A. Haniff. Low light level ccds and visibility parameter estimation. *Mon. Not. R. Astron. Soc.*, 347:1187-1197, 2004.
- [3] A. G. Basden, C. A. Haniff, and C. D. Mackay. Photon counting strategies with low-light-level ccds. *Mon. Not. R. Astron. Soc.*, (345):985-991, 2003.
- [4] H. Lanteri and C. Theys. Restoration of astrophysical images - the case of poisson data with additive gaussian noise. *Eurasip Journal on Applied Signal Processing - Special Issue on Applications of Signal Processing in Astrophysics and Cosmology*, (15):2500-2513, 2005.
- [5] D. G. Luenberger. *Introduction to linear and non linear programming*. Addison Wesley, Reading, Massachusetts, 1973.
- [6] M. Minoux. *Programmation mathématique - Théorie et algorithmes*, volume 1. Collection technique et scientifique des télécommunications, dunod edition, 1983.
- [7] C. Theys and H. Lanteri. Restoration of astrophysical images - the case of poisson gamma data. In *EUSIPCO*, 2006.
- [8] V.T. Tom, T.F. Quattieri, M.H. Hayes, and J.H. McClellan. Convergence of iterative non-expansive signal reconstruction algorithms. *IEEE Transactions on Acoustics Speech and Signal Processing*, 29:1052-1058, 1981.