Correction par bloc des transitoires de la caméra infrarouge ISOPHOT C-100 avec un modèle non linéaire dissymétrique

Alain COULAIS[1], Jean-François GIOVANNELLI[2], Johann SÉE[2,1], Fanny BALLEUX[1], Alain ABERGEL[1]

[1]Institut d'Astrophysique Spatiale, Bâtiment 121, Université Paris XI, 91405 Orsay cedex, France

[2]Laboratoire des Signaux et Systèmes (CNRS-Supélec-UPS)
Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France

Alain.Coulais@ias.fr, Jean-Francois.Giovannelli@lss.supelec.fr, Alain.Abergel@ias.fr

Résumé – Les photo-détecteurs utilisés en astronomie infrarouge à très bas niveau de flux présentent généralement une rémanence lors d'un changement d'illumination. Ce phénomène transitoire est non linéaire et dissymétrique selon que le flux entrant croît ou décroît. Pour le détecteur ISOPHOT C-100 embarqué sur le satellite ISO, nous avons montré qu'un modèle semi-empirique permet de décrire avec une bonne précision (de l'ordre de quelques pour-cent) ces transitoires. Une méthode d'inversion permet de corriger ces transitoires. Les limitations de l'approche sont détaillées.

Abstract – Low level photo detectors used in infrared astronomy do not response instantaneously after a change of the incident illumination. These transient processes are non linear, and upward steps and downward steps are not symmetrical. We show that a semi-empirical model allows to describe with a good accuracy the transient processes for the C-100 detector of ISOPHOT aboard the ISO satellite. One inversion method allows to correct the transient processes. Limitations are described.

1 Position du problème

Le ciel observé en astronomie infrarouge (IR) autour de 100 μm se caractérise par des niveaux de flux très bas (typiquement 10^{-9} fois le flux d'une scène IR terrestre). A ces niveaux, les détecteurs utilisés sont généralement des semi-conducteurs extrinsèques refroidis à quelques Kelvins. La réponse de ces détecteurs IR à un changement de flux n'est pas instantanée. Un phénomène transitoire se produit jusqu'à stabilisation sur la nouvelle valeur. S'il n'est pas corrigé, ce transitoire introduit des distortions dans les images reconstruites, équivalentes à un biais. La modélisation puis l'inversion de ce transitoire permet d'éviter ces distorsions, de réduire le temps d'observation pour une même précision et, en mode imagerie par balayage, de réduire les artefacts. Ce problème se rencontre sur la majorité des détecteurs IR comme ceux des satellites IRAS et ISO [1].

Dans [3], nous nous intéressions au problème de la correction des transitoires du détecteur LW de la caméra CAM [2] du satellite ISO [11]. Il a été montré qu'un modèle non-linéaire issu de la physique du solide [8] permet de décrire avec une excellente précision ce transitoire, aussi bien en montée qu'en descente, lorsque l'éclairement du détecteur est uniforme. Nous avions proposé une méthode d'inversion instant après instant liée aux particularités du modèle. Alors que l'erreur systématique sans correction du transitoire pouvait atteindre jusqu'à 40 %, la correction permet d'atteindre une précision de l'ordre du % sous éclairement uniforme.

Nous nous intéressons ici à la caméra PHOT C-100 [13] d'ISO. Cette caméra est constituée de 3×3 pixels; elle est conçue pour observer le rayonnement infrarouge lointain, dans un domaine s'étendant de $\sim$70 à $\sim$100 μm. La forme du transitoire de C-100 est nettement différente de celle du transitoire CAM, tout en étant aussi non-linéaire et non-symétrique. À ce jour, aucune correction du transitoire C-100 n'est faite en routine lors du traitement transformant les observations brutes en données utilisables scientifiquement. Pourtant, en fonction de la structure du ciel et de la taille des blocs (même pointage sur le ciel), l'erreur sur la valeur moyenne par bloc sans correction du transitoire va de quelques % à $\sim$30 %. Le modèle choisi comme la méthode de correction sont totalement différents de ceux étudiés dans [3].

Dans cet article, nous justifions le modèle choisi en Section 2 puis nous décrivons en Section 3 la méthode d'inversion adaptée au modèle. Les limitations connues de l'approche sont rassemblées en Section 4.

2 Transitoires C-100

2.1 Description

Le ciel observé en IR est composé d'une structure à très grande échelle (la lumière zodiacale) sur laquelle se superposent des structures à plus petite échelle spatiale, allant jusqu'aux sources ponctuelles par rapport à la résolution instrumentale. Un des modes d'observation le plus classique consiste à observer durant quelques poses élémentaires (typiquement de quelques dizaines à quelques centaines, qui forment ainsi un bloc) dans une direction fixe, puis à changer de direction de pointage. Le rayonnement observé est ainsi constant par morceaux. Des transitoires se produisent après chaque changement de pointage lorsque le niveau de flux du ciel change (*cf.* Fig. 1).

L'étude des transitoires [16] en montée et en descente pour

1. ISO : http://www.iso.vilspa.esa.es/

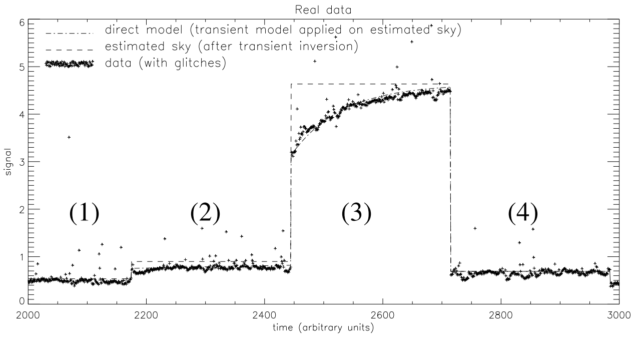

FIG. 1: *Exemple de transitoire réel, forte marche de flux incident. Le premier bloc (t entre ∼2000 et 2175) semble stable. Le second bloc (t entre ∼2175 et 2450) présente un long transitoire sur cette petite marche de flux. Le troisième bloc (t entre ∼2450 et 2710) présente un transitoire rapide et presque stabilisé en fin de bloc. Comme la constante de temps est proportionnelle à λ/J^∞, la stabilisation est d'autant plus rapide que la marche de flux J^∞ est grande. D'autre part, la redescente (bloc 4, après t ∼2710) apparaît ici instantanée.*

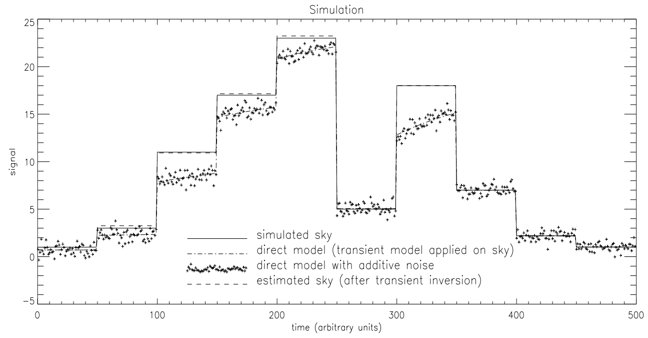

FIG. 2: *Exemple de correction de transitoire sur des simulations. Le signal d'entrée est constitué d'une série de paliers, correspondant aux différents pointages du télescope embarqué par le satellite. Le temps de repointage entre deux positions fixes est négligé. Ce sont ces paliers que l'on cherche à reconstruire. On applique le modèle direct sur ce signal d'entré. Un bruit gaussien est ajouté à la sortie du modèle direct. Par rapport aux observations, cette simulation correspond à un fort niveau de bruit et à des blocs de taille standard (la taille des blocs peut varier de 8 à 256, par multiple de 2). Excepté pour les marches 50-100 et 200-250, la correction se confond parfaitement avec le signal d'entrée.*

chacun des 9 pixels du détecteur a mis en évidence plusieurs propriétés (reproductibles) du détecteur (Fig. 1). Les transitoires montants présentent (*i*) un saut instantané de l'ordre de $\beta = 60$ % de la valeur à atteindre J^∞ et (*ii*) un temps de stabilisation en première approximation inversement proportionnel à la différence d'illumination. Les transitoires descendants sont, en première approximation, instantanés. Au contraire du transitoire de CAM, le transitoire montant ne présente pas de point d'inflexion et le descendant est quasi-instantanée.

Tout ceci indique que le modèle décrivant les transitoires de ce détecteur doit être non-linéaire pour les montées, et qu'il doit de plus distinguer les montées des descentes (dissymétrie).

2.2 Modèle utilisé

Après étude des transitoires C-100 sur des marches de flux nettes avec des blocs longs et des temps de pose courts [16], la comparaison avec les modèles analytiques (par ex. [7]) n'a pas été concluante.

Nous avons finalement utilisé un modèle semi-empirique non linéaire distinguant les montées des descentes. Pour les transitoires montants, ce modèle est néanmoins inspiré du modèle physique[2] proposé par Fouks dans [7]. Malheureusement, ce dernier ne décrit pas les fronts descendants.

Le modèle décrit le passage d'un niveau stabilisé J^∞_{n-1} à un nouveau niveau J^∞_n à l'instant $t = t_n$ selon le comportement suivant pour $t \in [t_n, t_{n+1}[$, on a $J_n(t) =$

$$\begin{cases} J^\infty_n & (\mathcal{M}a) \\ J^\infty_n - (1-\beta)(J^\infty_n - J^\infty_{n-1})\exp(-J^\infty_n(t-t_n)/\lambda) & (\mathcal{M}b) \end{cases}$$

2. En fonction des ordres de grandeur des différents paramètres majeurs du détecteur (dopage du substrat et des contacts, tension et orientation du champ par rapport à l'éclairage, distance inter-contacts, ...) différents modèles de transitoire peuvent être déduits selon les approximations introduites pour les équations physiques décrivant ces photo-détecteurs [7]. Néanmoins, la fabrication de ce type de détecteurs IR reste délicate, certains paramètres sont mal contrôlés et de nombreuses autres causes de biais existent (conditions d'illumination du détecteur, qualité de l'électronique, stabilité de la température de fonctionnement (cryogénie à quelques Kelvins), linéarité et traitement des rampes (*cf.* Sect. 4.2), ...). Au final, l'accord des modèles aux réponses expérimentales est souvent limité, d'où le recours fréquent à des modèles empiriques, surtout pour les détecteurs de technologie Ge:Ga comme C-100.

le cas (*a*) décrit les fronts descendants ($J^\infty_n < J^\infty_{n-1}$) et le cas (*b*) décrit les fronts montants ($J^\infty_n > J^\infty_{n-1}$). β est l'amplitude de la réponse instantanée qui suit le changement de flux montant et λ une constante (assimilée à une constante de temps). Le modèle utilisé est toutefois simplifié par rapport aux propriétés observées (et décrites en Sect. 2.1) : on utilise une constante de temps en J^∞_n/λ au lieu de $(J^\infty_n - J^\infty_{n-1})/\lambda$.

Un exemple de données synthétiques est présenté en Fig. 2.

3 Inversion des transitoires

Ce paragraphe est consacré à la correction des transitoires, c'est à dire l'inversion du modèle $\mathcal{M}$. Les blocs de données sont au nombre de N et indicés par n ($n = 1, \ldots, N$). Les données mesurées sont M échantillons des $J_n(t)$ regroupés dans le vecteur $\boldsymbol{J} \in \mathbb{R}^M$. Les flux inconnus, au nombre de N (un par bloc), sont notés comme précédemment J^∞_n, $n = 1, \ldots, N$. Ils sont regroupés dans le vecteur $\boldsymbol{J}^\infty \in \mathbb{R}^N$.

L'inversion, *i.e.* l'estimation $\boldsymbol{J}^\infty$ à partir de $\boldsymbol{J}$, est réalisée par moindres carrés. Il s'agit de mesurer la distance quadratique globale entre les données effectivement mesurées $\boldsymbol{J}$ et les données sorties du modèle à inverser $\mathcal{M}(\boldsymbol{J}^\infty)$:

$$Q_{\mathrm{MC}}(\boldsymbol{J}^\infty) = \| \, \boldsymbol{J} - \mathcal{M}(\boldsymbol{J}^\infty) \, \|^2, \tag{1}$$

et de choisir comme << meilleure valeur >> de l'objet inconnu $\boldsymbol{J}^\infty$ celui qui minimise cette distance :

$$\widehat{\boldsymbol{J}}^\infty_{\mathrm{MC}} = \underset{\mathbb{R}_+}{\arg\min}\, Q_{\mathrm{MC}}(\boldsymbol{J}^\infty)\,. \tag{2}$$

On exploite ainsi les deux types d'informations disponibles : le flux inconnu est positif et constant par morceaux.

L'optimisation du critère Q_{MC} est le point délicat, essentiellement à cause de la non linéarité et de la dissymétrie du modèle $\mathcal{M}$. Q_{MC} possède cependant une structure particulière visible sur l'expression de $\mathcal{M}$: à l'intérieur du bloc n seule interviennent les deux valeurs J_n^∞ et J_{n-1}^∞. Cette particularité permet décrire Q_{MC} sous la forme :

$$Q_{\mathrm{MC}}(\boldsymbol{J}^\infty) = \sum_{n=2}^{N-1} Q_n(J_n^\infty, J_{n-1}^\infty), \tag{3}$$

faisant ainsi apparaître une structure << couplée >>, typique des structures de chaîne de Markov [14]. La minimisation globale est alors possible sur une grille discrète, arbitrairement fine, de K valeurs du flux entrant grâce à un algorithme de programmation dynamique de type Viterbi [6].

Il s'agit d'un algorithme récursif sur n. Dans un premier temps, il calcule le critère pour tous les couples de valeurs de J_1^∞ et J_2^∞ de la grille. Il en extrait alors le minimum par rapport à J_1^∞, pour chaque valeur de J_2^∞ et conserve les K minimum et les K minimiseurs. L'étape suivante consiste à introduire J_3^∞. A partir des résultat de la première étape il détermine le minimum par rapport à J_2^∞ pour toutes les valeurs de J_3^∞. Ainsi, de proche en proche il réalise l'optimisation jusqu'à J_N^∞. La dernière opération consiste alors à revenir de proche en proche pour retrouver les minimiseurs J_{N-1}^∞, J_{N-2}^∞ jusqu'à J_1^∞.

La grille peut être adaptée pour raffiner l'optimisation. L'optimum peut également être localement raffiné par une technique de descente selon le gradient.

Un exemple de correction du transitoire sur des données synthétiques est présenté en Fig. 2. Malgré une amplitude de bruit gaussien nettement supérieur à ce qui est observé, la reconstruction est à mieux que le pour-cent même lorsque plusieurs marches montantes se succèdent. Un exemple de correction de données réelles, avec superposition du transitoire modélisé et de la correction (constante par bloc) est donné en Fig. 1.

4 Limitations et travaux en cours

4.1 Paramètres du modèle

Le modèle choisi pour décrire les transitoires est caractérisé par deux paramètres, β et λ, supposés constants pour un pixel donné. Dans un premier temps, il s'agit d'estimer ces paramètres et de vérifier leur stabilité durant les $\sim$ 800 orbites de la mission ISO. Un problème majeur apparaît alors : le changement de direction de pointage du ciel (changement de bloc) n'est pas instantané.

Même si les données inter-bloc peuvent être décrites par le modèle, l'absence de stabilisation du flux incident durant l'inter-bloc complique fortement l'ajustement des paramètres (β, λ) à partir de ces marches de flux. Pour la correction, en première approximation, les N mesures $J_n(t)$ d'un inter-bloc peuvent être découpées en N blocs de longueur élémentaire. Ceci augmenterait très fortement le temps de calcul de l'inversion.

A ce jour, il a été choisi de supprimer les inter-blocs. Sous cette hypothèse restrictive, pour un même pixel et différentes marches de flux, la dispersion du paramètre β de saut est de l'ordre de 5 %, et celle de λ de 10 %. De plus, il apparaît clairement que (*i*) le nombre de marches exploitables pour caractériser les transitoire est insuffisant et (*ii*) les glitches modifient les transitoires (*cf.* Fig. 1 et Sect. 4.2).

4.2 Traitement des rampes et influence des glitches

Les données temporelles $J(t)$ exploitées ici résultent d'un pré-traitement des mesures élémentaires du détecteur. Une partie du rayonnement IR incident est transformé en paires électrons/trous dans le détecteur. Les électrons vont s'accumuler dans une capacité, dont la charge est linéaire en première approximation. Cette charge est lue de manière non-destructive plusieurs fois (typiquement 8 à 256 pour PHOT) entre deux remises à zéro (R.A.Z.) de la charge de la capacité. Les données temporelles $J(t)$ sont les pentes de ces rampes. Lorsque l'échantillonnage de la rampe est suffisant, et lorsque le bruit sur cette rampe est faible et gaussien, l'estimation de la pente est bonne. Plusieurs effets compliquent le calcul effectif des pentes. Pour l'ensemble des détecteurs ISO, il a été observé que, juste après la R.A.Z, la charge n'est pas linéaire [9, 19, 17]. De plus, lors d'observations scientifiques, de nombreuses contraintes (dont le temps total d'observation et le risque de saturer le détecteur) fixent la durée de la rampe et la fréquence d'échantillonnage sans que ces valeurs soient nécessairement optimales pour le calcul de la pente. Enfin, de temps en temps, des particules énergétiques interagissent avec le substrat du détecteur et donnent des glitches, excès subit de charges dans la capacité. Ces glitches apparaissent comme des données aberrantes de forte amplitude sur les données réelles de la Fig. 1, souvent suivis de déviation à l'évolution en tendance du transitoire. Ce phénomène est d'autant plus marqué que les temps d'intégration sont longs. En effet, la probabilité de glitches multiples augmente, et les méthodes standards de traitement des rampes deviennent alors moins efficaces [9, 19, 17].

Dans le cas particulier de PHOT C-100, l'erreur principale dans le traitement des rampes vient des glitches [9]. Ce problème, déjà sensible sur les données réelles de la Fig. 1, devient très gênant sur des rampes longues et rend imprécise l'inversion sur des blocs à petit nombre de points (parce que la forme du transitoire change).

Un traitement plus robuste pourra être fondé sur une modélisation plus fine du bruit et de sa composante glitch : événements relativement rares, forte amplitude, positivité, *etc.*.. De tels traitements pourront être fondés sur des statistiques robustes, normes L_1, traitement des données abbérantes [10, 18, 1].

4.3 Observations de calibration

Les observations scientifiques sont encadrées d'observations de calibration qui permettent de convertir les observations en unités physiques. Ces calibrations sont aussi affectées par les transitoires. Ce travail a mis en évidence que ces calibrations sont le plus souvent trop courtes et avec un mauvais rapport signal sur bruit. Et la méconnaissance de l'état précédent la première calibration empêche en théorie l'emploi de la méthode de correction des transitoire pour corriger cette première calibration.

Au total, on estime une imprécision pouvant aller jusqu'à 20 % sur la valeur de ces calibrations. Il sera nécessaire de confirmer que cette imprécision domine ou non sur les autres erreurs.

4.4 Bilan des limitations

Ces travaux en cours ont pour but de préciser les limites du modèle en valeur relative sur les transitoires mais aussi du traitement global en valeur absolu (précision photométrique, unités physiques). Ces limitations montrent clairement que le traitement des transitoires n'a de sens que si les autres points critiques du traitement sont maîtrisés. Néanmoins, comme cela avait déjà été le cas pour PHOT-S [8], la modélisation et la correction des transitoires est un excellent moyen de faire ressortir d'autres problèmes instrumentaux.

5 Conclusion et Prospectives

Des progrès significatifs ont été fait pour la description et l'inversion de la réponse temporelle d'un détecteur IR du satellite ISO. Un modèle semi-empirique, non-linéaire et dissymétrique s'est montré adapté. Une méthode d'inversion basée sur ce modèle et prenant en compte le caractère constant par morceaux et positif du flux inconnu a été développée.

Cette étude a permis de corriger en bonne partie les distortions introduites par les phénomènes de transitoires. Après cette correction, il semble maintenant, dans l'état actuel des connaissances, que ce soit les mesures de calibration qui limitent la qualité des résultats. Ceci a une conséquence majeure pour les stratégies d'observation et de calibration des instruments similaires en préparation [3].

Le satellite ISO embarquait quatre modules (CAM, PHOT, SWS et LWS) totalisant 16 détecteurs. Tous ces détecteurs sont affectés, à des degrés divers, par des effets de transitoire. Ces transitoires restent le phénomène induisant la plus grande erreur systématique dans les observations. Souvent, un premier traitement des transitoires permet de mettre en évidence d'autres problèmes instrumentaux (optique, couplage, dérives, ...). En éliminant les détecteurs les plus problématiques ou présentant un intérêt scientifique désormais secondaire (grâce aux progrès des observations depuis le sol), avec les avancées déjà faites [15, 12, 4, 5], il reste 4 à 7 détecteurs d'ISO dont les observations scientifiques bénéficieraient d'un tel travail : la recherche d'un modèle de transitoire et le développement d'une méthode d'inversion adaptée.

Remerciements

A. Coulais remercie (1) le CNES, le CNRS, le SAp-CEA et l'IAS pour leur support financier, (2) J.-F. Giovannelli pour son aide indispensable pour la partie inversion, (3) G. Lagache et B. Stepnik pour leur aide lors de la préparation des données PHOT et (4) ses stagiaires J. Sée et F. Balleux pour leurs travaux rassemblés ici.

3. D'après les test au sol, les détecteurs IR des prochains satellites IR (SIRTF (`http://sirtf.caltech.edu`) en 2002, ASTRO-F (`http://www.ir.isas.ac.jp/index.html`) en 2004) présentent le même type de problèmes de transitoires (et de rampes, *cf.* Sect. 4.2). que les détecteurs embarqués sur ISO.

Références

[1] Black, M. J. et A. Rangarajan : 1996, *Int. J. Computer Vision*, Vol. 19, no. 1, pp. 57–91

[2] Cesarsky, C. J. et al. : 1996, *Astron. Astrophys.* **315**, L32

[3] Coulais, A. et A. Abergel : 1999, Vol. T-1, pp. 147–150, GRETSI'99

[4] Coulais, A. et A. Abergel : 2000, Astron. Astrophys. Suppl. Ser., Vol. 141, pp. 533–544

[5] Coulais, A., Fouks, B. I., Giovannelli, J.-F., Abergel, A., et Sée, J.: 2000, in M. Strojnik et B. Andresen (eds.), *Infrared Spaceborne Remote Sensing VIII*, Vol. 4131, pp. 205–217, SPIE, San Diego, USA

[6] Forney, G. D. : 1973, *Proc. IEEE*, Vol. 61, no. 3, pp. 268–278

[7] Fouks, B. I. : 1992, in *ESA SP-356 Photon Detectors for Space Instrumentation*, pp. 167–174

[8] Fouks, B. I. et J. Schubert : 1995, *Proc. of SPIE* **2475**, 487

[9] Gabriel, C. et J. A. Acosta-Pulido : 2000, Exp. Astronomy, Vol. 10, pp. 319–331

[10] Huber, P. J. : 1981, *Robust Statistics*, John Wiley, New York, NY

[11] Kessler, M. F. et al.: 1996, *Astron. Astrophys.* **315**, L27

[12] Kester, D. : 1999, *The impact of memory effects correction on SWS data*, Technical report, SRON, Netherlands

[13] Lemke, D. et al. : 1996, *Astron. Astrophys. Suppl. Ser.* **315**, L64

[14] Rabiner, L. R. et B. H. Juang : 1986, *IEEE ASSP Mag.*, pp. 4–16

[15] Schubert, J., Roth, G., Wolf, J., Lemke, D., et Fouks, B. I. : 1994, *Proc. of SPIE* **2268**, 283

[16] Sée , J. : 1999, *Correction des effets mémoire de l'instrument ISOPHOT C-100*, tech. rep., Stage de magistère

[17] Swinyard, B. et al. : 2000, Exp. Astronomy, Vol. 10, pp. 157–176

[18] Walter, E. et L. Pronzato : 1994, *Identification de modèles paramétriques à partir de données expérimentales*, Masson, Paris, France

[19] Wieprecht, E., E. Wiezorrek, et L. Haser : 2000, Exp. Astronomy, Vol. 10, pp. 199–214