

Imagerie multispectrale et Image Fusion Application en Astronomie

A.BIJAOUTI¹

¹UMR 6202 CASSIOPÉE, CNRS, Observatoire de la Côte d'Azur, BP 4229, 06304 Nice Cedex 4

Albert.Bijaoui@oca.eu

Résumé - L'analyse d'un ensemble d'images astronomiques obtenues dans différents canaux spectraux afin d'en extraire tous les objets identifiables et de les décomposer spectralement est l'un des problèmes essentiels pour l'observation astronomique actuelle. L'extraction des sources s'effectue à partir d'une image fusion construite dans l'espace en ondelettes. Cette image est ensuite décomposée en un ensemble de fonctions d'échelle avec des positions et des échelles déterminées (*pyrels*). L'identification des *pyrels* à considérer est faite via les maxima d'une transformée en ondelettes. Un algorithme de segmentation est appliqué aux *pyrels* afin de pouvoir les associer à des objets distincts, dont la reconstruction est immédiate. À partir de la carte des *pyrels* sur l'image fusion, les caractéristiques des objets sont déterminées dans chaque canal spectral. Une séparation en composantes spectrales est obtenue grâce à un ensemble de distributions spectrales. L'application à des images astronomiques a montré la pertinence de ce nouveau modèle de vision pour cette catégorie d'images.

Abstract – Today the same sky field is observed with different telescopes in different spectral bands. The resulting images have to be analyzed in order to extract, to measure and to classify the detectable objects. This multispectral analysis constitutes one of the main problems of the astronomical data analysis. In order to get consistent detections for one band to the others the objects are extracted from a fusion image built in the wavelet space. The wavelet local maxima, in location and in scale are detected and an image is restored from discrete scaling functions (so-called *pyrels*) at this same locations and scales. Their amplitudes are fitted by a classical Least Mean Squares estimator. A labelling is done on the *pyrels*, taking into the connected fields built by their projection on a binary image. At one label corresponds a set of *pyrels* which allows one to construct a specific image. An object and its image for each spectral band are so associated to each label. Taking into account known spectral distribution for typical spectral classes, these images can be decomposed into weighted spectral classes. The application to astronomical images showed the model pertinence for multispectral analysis in this context.

1. Introduction.

Aujourd'hui le Ciel est observé par des instruments très variés, permettant l'analyse des différentes sources avec un grand nombre de bandes spectrales, couvrant toutes les longueurs d'onde, des rayons gamma aux ondes radios. Dans le cadre du développement de l'Observatoire Virtuel [1], des outils comme Aladin [2] permettent de fournir à un astronome un grand ensemble d'images relatives à une même région dans ces différents domaines de longueur d'onde. Même dans le cas d'un faible nombre de bandes, la forme des sources variant beaucoup avec la longueur d'onde, il n'est pas facile de mettre en correspondance les objets identifiés dans chaque image, afin d'en obtenir des mesures cohérentes et de les analyser spectralement.

Dans le cadre de l'ACI *Masse de Données en Astronomie* [3] nous avons examiné la construction d'un outil devant permettre :

- d'optimiser la détection des objets faibles ;
- d'obtenir sur les objets des mesures cohérentes d'une bande à l'autre ;
- d'identifier les classes spectrales ;
- de cartographier les classes.

Pour résoudre ce problème, nous proposons une nouvelle méthode basée sur le concept d'image de fusion. Dans cette communication nous passerons en revue les différentes étapes de cette méthode :

- Construction de l'image fusion ;
- Extraction des masques des objets de cette image ;
- Reconstruction des images des objets pour chaque bande ;
- Identification des classes spectrales ;
- Décomposition de l'image en sources spectrales ;
- Décomposition des objets en sources spectrales.

2. L'image fusion

L'analyse séparée des images conduit à la nécessité de mettre en correspondance les objets identifiés. Compte tenu de l'aspect diffus d'une partie d'entre eux, ceci est très difficile à réaliser à partir des données extraites pour chaque bande. La co-addition d'images est une pratique courante en astronomie pour améliorer la détection des objets faibles. La combinaison de sources diffuses observées dans différents canaux spectraux peut aussi permettre de détecter des objets plus faibles, mais leur

forme variant avec la longueur d'onde, des informations peuvent être perdues par cette opération.

L'approche quadratique, dite chi-2 [4], a été mise en œuvre sous différentes formes [5]. Nous avons examiné plusieurs variantes, en particulier en se basant sur des images débruitées avec la transformation en ondelettes. Un gain significatif dans la détection des sources a été ainsi obtenu [5].

Après analyse des images, il nous est apparu que le meilleur résultat provenait d'une fusion quadratique dans un espace en ondelettes. La méthode obtenue est proche de Scheunders et De Backer [6]. Les images discrètes $V(k,l,n)$, $k \in (1,K)$, $l \in (1,L)$, $n \in (1,N)$, sont décomposées avec un algorithme à trous [7] en coefficients en ondelettes discrètes $w(i,k,l,n)$, $i \in (1,I)$. Les variances $\sigma^2(i,n)$ des coefficients à chaque échelle, pour chaque image, sont estimées sur la base d'un bruit gaussien blanc discret. L'existence d'une information dans l'ensemble des images au point (k,l) et pour l'échelle i est caractérisée par le résidu correspondant :

$$R(i,k,l) = \sum_{n=1,N} \frac{w^2(i,k,l,n)}{\sigma^2(i,n)} \quad (1)$$

$R(i,k,l)$ étant toujours positif, il ne peut s'agir d'un coefficient d'une transformée en ondelettes. Une pseudo transformée en ondelettes est construite avec pour module la racine carrée de cette quantité. Son signe est celui de la combinaison linéaire

$$C(i,k,l) = \sum_{n=1,N} \frac{w(i,k,l,n)}{\sigma(i,n)} \quad (2)$$

Soit :

$$W(i,k,l) = R(i,k,l) \text{Signe}(C(i,k,l)) \quad (3)$$

Les coefficients $W(i,k,l)$ peuvent permettre de reconstruire une image fusion $V_f(k,l)$ dans l'espace des pixels. Compte tenu de la redondance, la transformée de cette image n'est pas exactement $W(i,k,l)$ mais $W_f(i,k,l)$. L'inversion de $W_f(i,k,l)$ conduit aussi à $V_f(k,l)$. C'est la raison pour laquelle nous avons utilisé l'expression pseudo transformée. $W_f(i,k,l)$ est la projection orthogonale de $W(i,k,l)$ dans l'espace de la transformée en ondelettes.

L'extraction des objets peut s'effectuer à partir de l'image de fusion $V_f(k,l)$, mais il est naturel de poursuivre l'analyse en effectuant cette opération directement sur les coefficients, évitant une éventuelle perte d'information dans la projection.

3. L'extraction des objets

La méthode proposée dérive de celle décrite dans [8]. Les objets sont représentés par une somme de composantes appelées *pyrels*, correspondant à des fonctions d'échelle discrètes $H(i_j, i, k, l)$ au point (k_j, l_j) et à l'échelle i_j , j étant l'indice du pyrel. Cette appellation *pyrel* résulte de sa construction par l'algorithme à trous résultant de l'équation de dilatation :

$$H(i+1, k, l) = \sum_{n,m} h(n, m) H(i, k+2^n, l+2^m) \quad (4)$$

Nous posons $H(0,k,l)=0$ partout sauf $H(0,1,1)=1$. Le filtre $h(n,m)$ est choisi séparable en lignes et en colonnes, les coefficients étant les coefficients binomiaux de degré 4, centrés sur l'indice 0 ($C_4^{n+2}/16$, $n \in (-2, +2)$).

Pour identifier les pyrels nécessaires à la reconstruction des objets, on localise les maxima en échelle et en position des $W(i,k,l)$. À chaque maximum on associe un pyrel $c(j)H(i_j, k-k_j, l-l_j)$. Le triplé (i_j, k_j, l_j) provient de la localisation du maximum local, l'amplitude $c(j)$ dérive, en première approximation, du coefficient $W(i_j, k_j, l_j)$. Une correction des amplitudes des pyrels est faite avec un critère de minimisation de résidu dans l'espace en ondelettes.

Une seule étape ne suffit pas. L'écart entre les coefficients initiaux et ceux de la transformée en ondelettes de l'image reconstruite est traité itérativement jusqu'à obtention d'écarts tous statistiquement non significatifs.

Lorsqu'un critère d'arrêt est atteint, on dispose d'un ensemble de pyrels $H(i_j, k-k_j, l-l_j)$, $j \in (1,J)$, auxquels est associé un vecteur de coefficients $c(j)$. Cet ensemble permet de reconstruire une image $V_r(k,l)$ de transformée en ondelettes $W_r(i,k,l)$ telle que dans les échelles considérées il n'existe aucun coefficient qui ne diffère de manière significative des coefficients $W(i,k,l)$. On peut remarquer que toute image différant de l'image $V_r(k,l)$ par une fonction $F(k,l)$, telle que sa transformée en ondelettes n'a aucun coefficient significatif dans la gamme d'échelle considérée, est aussi valide. L'algorithme fournit donc une image possible, parmi toute une classe. Il est construit de manière à rechercher le minimum de pyrels, mais ce critère de parcimonie n'est pas prouvé.

À partir des positions des pyrels, on construit une image $T(k,l)$ telle que $T(k,l)=0$ si aucun pyrel n'a pour localisation (k,l) et $T(k,l)=1$ dans le cas contraire. Un algorithme de croissance de région permet d'attribuer une même étiquette à tous les pixels d'un même domaine connexe de $T(k,l)$.

On postule ensuite que tous les pyrels (i_j, k_j, l_j) associés à la même étiquette correspondent au même objet astronomique. Si $m \in (1,M)$ désigne un étiquette, c'est-à-dire un objet donné, l'ensemble de ses pyrels est désigné par J_m . L'image de la source m dans l'image Fusion est donc :

$$V_m(k,l) = \sum_j c(j) H(i_j, k-k_j, l-l_j) \quad (5)$$

où $j \in J_m$.

4. Image dans chaque bande

En fait, reconstruire l'image d'un objet dans l'image de fusion n'a pas grand intérêt scientifique. L'objectif étant de déterminer l'image des objets dans toutes les bandes spectrales, les pyrels déterminés sur l'image fusion, et la décomposition en objets correspondante, doivent être exploitées pour chaque bande.

Les coefficients $c(j,n)$ de l'ensemble des pyrels de l'image n sont déterminés par moindres carrés à partir de

la transformée en ondelettes de cette image. Comme pour l'image fusion, l'image $V_r(k,l,n)$ reconstruite à partir des pyrels peut différer de l'image observée $V(k,l,n)$ d'un fond ne donnant pas de coefficients significatifs. Mais elle peut aussi différer d'une fonction conduisant à des coefficients significatifs, tels que dans la fusion leur faible valeur a été noyée dans le bruit des autres images. Le fait que la détection se soit faite sur la fusion n'a pas éliminée l'information ne pouvant être détectée que sur une seule image. L'analyse des images des différences permet ainsi de retrouver les spécificités dans une image.

L'image $V_r(k,l,n)$ est reconstruite avec la relation :

$$V_r(k,l,n) = \sum_j c(j,n)H(i_j, k-k_j, l-l_j) \quad (6)$$

où $j \in (1, J)$.

Grâce à la segmentation en pyrels la reconstruction individuelle des objets pour chaque bande s'effectue avec la relation :

$$V_r(k,l,n) = \sum_j c(j,n)H(i_j, k-k_j, l-l_j) \quad (7)$$

où $j \in J_m$.

5. Décomposition spectrale

En raison des processus physiques à l'origine de l'émission d'énergie et de toutes les interactions jusqu'à leur détection, les distributions spectrales des sources célestes occupent un continuum multidimensionnel de classes. De ce fait le nombre de classes spectrales augmente exponentiellement avec le nombre de bandes considérées. Une nouvelle voie descriptive a été développée en se basant sur la séparation en éléments purs [9]. Avec un algorithme d'analyse des nuées, par exemple du type k-means [10], on commence par partitionner l'espace des pixels. Le choix du nombre de classes est empirique, mais il est généralement plus grand que le nombre d'images. Un algorithme du type *matching pursuit* [11] permet d'obtenir des cartes de pourcentages des classes.

En fait, au lieu d'analyser les images, ce sont les coefficients des pyrels $c(j,n)$ qui sont décomposés en tenant compte des distributions spectrales :

$$c(j,n) = \sum_p s(j,p)a(p,n) \quad (8)$$

où $p \in (1, P)$ désigne un élément pur et $a(p,n)$ la distribution spectrale dans la bande n pour l'élément p . $s(j,p)$ représente l'amplitude de l'élément pur p pour le pyrel j .

Le nombre de classes P pouvant être plus grand que le nombre de bandes N , la détermination des amplitudes est un problème mal posé qui est résolu avec un algorithme du type *matching pursuit*.

L'image source p complète peut être reconstruite par la relation :

$$S(k,l,p) = \sum_j s(j,p)H(i_j, k-k_j, l-l_j) \quad (9)$$

où $j \in (1, J)$. On peut aussi reconstruire la part correspondant à l'objet m pour la source p :

$$S_m(k,l,p) = \sum_j s(j,p)H(i_j, k-k_j, l-l_j) \quad (10)$$

où $j \in J_m$.

6. Exemple d'application

Plusieurs ensembles d'images observées et simulées ont servis pour le développement de cette méthode. Nous l'illustrons avec les résultats obtenus avec une série d'images obtenues avec la caméra à grand champ du télescope de 2m20 de l'ESO. Le champ a été observé en 6 bandes. On a sélectionné une zone de 256x256 pixels pour les figures.

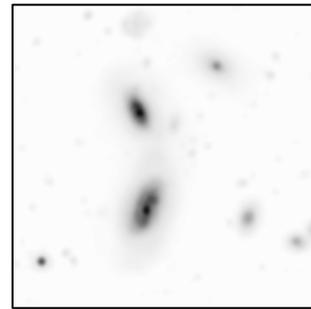


FIG. 1 : Image de la zone sélectionnée en filtre B.

Sur la figure 1 nous avons tracé l'image obtenue dans le filtre B. Sur la Figure 2 nous avons représenté l'image chi-2 brute.

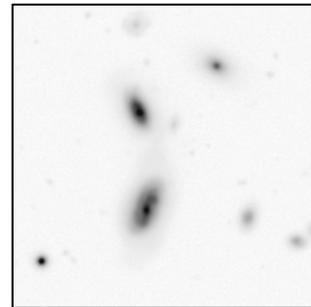


FIG. 2 : Image chi-2 brute

L'image reconstruite à partir de l'ensemble des pyrels pour l'image dans le filtre B est représentée figure 3. On remarque divers artefacts dans la reconstruction, indiquant les limites de la représentation à partir de l'image fusion.

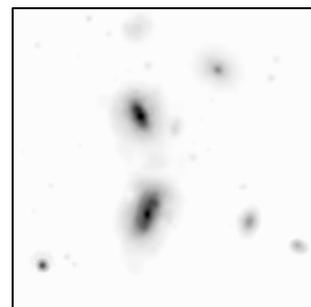


FIG. 3 : Image reconstruite à partir des pyrels en B.

Sur la figure 4, nous avons représenté l'image des étiquettes associées aux objets déterminés à partir de la fusion.

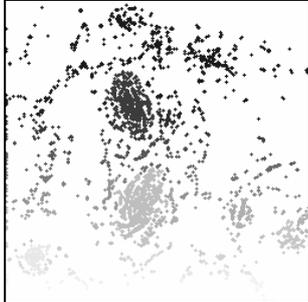


FIG. 4 : Image des étiquettes des maxima

Enfin sur la figure 5, nous avons représenté l'image d'un des objets reconstruits à partir des pyrels dans l'image fusion.

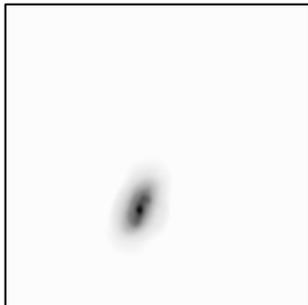


FIG.5 : Reconstruction de l'image d'une des galaxies dans l'image fusion

7. Conclusion

En résumé nous avons construit un outil permettant :

- De construire dans l'espace en ondelettes une image fusion tenant compte au mieux de l'information pertinente sur l'ensemble des bandes ;
- D'analyser directement la pseudo transformée obtenue avec un algorithme basé sur la détection itérative des maxima locaux de cet espace. L'image est reconstruite sous la forme d'une somme de pyrels identifiés par les maxima ;
- De regrouper les pyrels en objets, avec un algorithme de croissance de régions basé sur les positions des pyrels ;
- De déterminer par moindres carrés sur les transformées en ondelettes les coefficients des pyrels de toutes les bandes ;
- De reconstruire l'image de chaque objet dans chaque bande ;
- D'identifier les classes spectrales de pixels par un algorithme de type k-means ;
- De décomposer les valeurs des pixels ou les coefficients des pyrels en coefficients de sources spectrales, tenant compte des distributions spectrales ;
- De reconstruire les images des sources spectrales pour l'ensemble des images ;
- De reconstruire les images des sources pour chaque objet.

De nombreuses améliorations de ce modèle d'analyse d'images multibandes pourraient résulter de modifications de chacune des étapes de traitement. Parmi celles-ci, l'introduction de transformations pyramidales et non isotropes en ondelettes constitue l'une des voies les plus évidentes.

Le logiciel d'analyse multibande exploitant cette méthode est en cours de finalisation. Il a été développé dans le cadre de l'ACI *Masse de Données en Astronomie*, pilotée par F.Genova. Ce logiciel sera mis en oeuvre sur le réseau internet dans le cadre de l'Observatoire Virtuel développé par la communauté astronomique internationale.

Références

- [1] D.Egret et F.Genova, *Federation and Fusion of Astronomical Information: Standards and Tools for the Virtual Observatory*. Towards an International Virtual Observatory, P.J.Quinn et K.M.Gorski eds, pp. 94-105, ESO, 2004.
- [2] <http://aladin.u-strasbg.fr/aladin.gml>
- [3] <http://cdsweb.u-strasbg.fr/MDA/index.html>
- [4] A.S.Szalay, A.J.Connolly et G.P.Szokoly. *Simultaneous multicolor detection of faint galaxies in the Hubble Deep Field*. The Astronomical Journal, 117, pp.68--74, 1999.
- [5] A.Bijaoui, A.Guennec, C.Benoist, E.Slezak. *Analysis of multiband astronomical images using multiscale tools*, dans Wavelet applications in industrial processing III, ed. F.Truchetet et O.Laligant, SPIE 6001, pp.01-1::01-12, 2005.
- [6] P.Scheunders, S. De Backer. *Fusion and Merging of Multispectral images using multiscale fundamental forms*. Journal of Optical Society of America, A., 18, pp.2468--2477, 2001.
- [7] M.Holschneider, R.Kronland-Martinet, J.Morlet, P.Tchamitchian. *A Real-Time Algorithm for Signal Analysis with the Help of the Wavelet Transform*. Wavelets: Time-Frequency Methods and Phase-Space, pp.286--297. Springer Berlin, 1989.
- [8] A.Bijaoui. *Représentation des images via les maxima en ondelettes. Application à l'extraction des objets*. GRETSI'07, 2007.
- [9] A.C.Bateson, G.P.Asner et C.A.Wessman. *Endmember bundles: a new approach to incorporating endmember variability into spectral mixture analysis*. IEEE Trans. On GeoSci. And Remote sensing, 38, 1083-1094, 2000.
- [10] J.B. MacQueen. *Some Methods for classification and Analysis of Multivariate Observations*, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability", Berkeley, University of California Press, 1, pp.281-297, 1967.
- [11] S.Mallat, Z.Zhang. *Matching pursuit with time-frequency dictionaries*. IEEE Trans Signal Proc., 41, pp. 3397—3415, 1993.