WAVELET DOMAIN ASTRONOMICAL MULTIBAND IMAGE FUSION AND RESTORATION USING MARKOV QUADTREE AND COPULAS

Farid Flitti[1], *Christophe Collet*[1] *and Eric Slezak*[2]

[1] Université Strasbourg I, LSIIT - UMR 7005 - France
[2] Observatoire de la Côte d'Azur - UMR 6162 - France
[1] name@lsiit.u-strasbg.fr, [2] slezak@obs-nice.fr

Key words : Multiband image fusion, copulas, Markov quadtree, restoration, wavelet.

1. INTRODUCTION

Fusion of multiband images is of great interest in several applications like astronomy, remote sensing and medecine. It allows to obtain an efficient summary of the whole multiband information in a single scene. Obviously, fusion is more difficult for noisy observations. On one hand multiscale analysis is a popular choice in recent fusion research [1]. On the other hand, wavelet framework is very well adapted for denoising task [2]. Thus wavelet domain seems a quite appropriate for noisy image fusion. Recently, an efficient wavelet Markov modeling was introduced in [3] capturing interscale and spacial wavelet coefficient correlations. In this paper we use a more general Markovian framework, modeling not only spatial and interscale dependencies as the existent models do but also interband correlation for multiband image joint fusion and denoising. Moreover, the multidimensional likelihood is modeled using the copulas theory [4] which allows us to use any kind of marginal densities with a given interband correlation.

The proposed approach is summarized in Fig.1. We first carry out a decimated multiresolution transform using pyramidal algorithm with one wavelet [5] for each band (section 2). In a second step, we combine all pyramidal representations obtained in the previous step in a single Multiband-Multiresolution Pyramid (MMP) by considering detail coefficients at same space-scale location as a unique vector. In a third step, the MMP is segmented in two-classes using a vectorial hidden Markov quadtree (section 3) to separate significant wavelet coefficients from those associated with the noise. This modeling is very useful since the selection relies now not only on the sole coefficient magnitude [2] but also takes into account its neighbors : in space, in scale and with wavelength. This classification scheme produces a multiresolution binary mask highlighting significant wavelet coefficients. In the fourth step the significant coefficients of all pyramids are fused in single one using different rules explained in section 4. The final step consists in the recon-

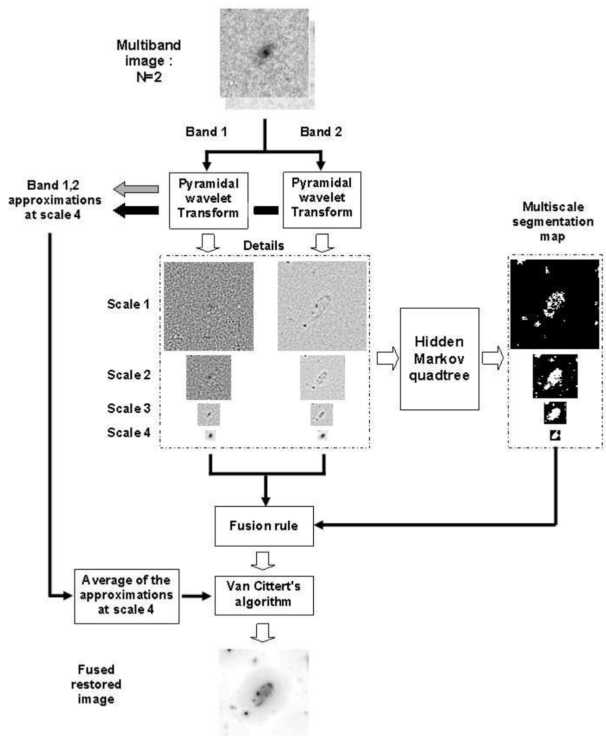

Fig. 1. Fusion-restoration algorithm illustrated for a bi-band image. A new fusion-restoration scheme for multispectral data, operating as follows. A pyramidal wavelet transform analyzes the $N = 2$ spectral bands (on the top). This leads to a multiresolution pyramid of wavelet coefficients for each band, up to scale $M = 4$. Then, all wavelet pyramids are combined to carry out two-class multiresolution Markovian segmentation map (on the right). This segmentation map masks small coefficients at different scales. The remainder coefficients are fused using an appropriate rule. The result with the average of coarsest approximations feed an iterative reconstruction procedure to give a unique fused restored image.

This research is supported by the French government (ACI/MDA project: Masses de Données en Astronomie : http://cdsweb.u-strasbg.fr/MDA/mda.html).

struction of the fused image using the iterative Van Cittert's algorithm [5].

2. THE PYRAMIDAL ALGORITHM WITH ONE WAVELET

The most usual tool for multiscale analysis of real images is the orthogonal/bi-orthogonal wavelets transform [6]. This transformation detects singularities across given directions. However, generally astronomical objects are diffuse and exhibit smooth edges in scales. This is why isotropic wavelet transforms which do not privilege any direction are preferred [5]. The pyramidal algorithm with one wavelet is an isotropic transform obtained by adapting the well known Laplacian pyramid of Burt and Adelson [7, 5].

Let $f(k)$ a 1D discrete signal. Its successive approximations are obtained by repeatedly applying a low-pass filtering followed by a down-sampling by a factor 2. Thus the approximation $c_{j+1}(k)$ at scale $j+1$ is given by :

$$c_{j+1}(k) = \sum_{l} h(l-2k) c_j(l), \quad (1)$$

where $c_j(k)$ is the approximation at scale j and h a normalized, symmetric low-pass filter verifying the equal contribution constraint[1] [7]. This procedure is initialized by setting $c_0(k) = f(k)$.

The detail information $w_j(k)$ lost from resolution j to resolution $j+1$ is computed as the difference between the approximation at scale j and the undecimated approximation at scale j as follows :

$$w_{j+1}(k) = c_j(k) - \tilde{c}_{j+1}(k) \quad (2)$$

where $\tilde{c}_{j+1}(k) = \sum_l h(l-k) c_j(l)$.

The reconstruction procedure is the same one as the Laplacian pyramid. Given details and approximation at scale $j+1$, the approximation at scale j is computed by:

$$c_j(k) = w_{j+1}(k) - \sum_{l} h(k-2l) c_{j+1}(l). \quad (3)$$

However, this reconstruction is not exact [5] but can be approached iteratively with the Van Citter's algorithm (Sect. 5).

For an image F (*i.e.;* 2D discrete signal), one operates separately with the same filter h on the rows and then on the columns [5]. Thus we obtain a pyramid W where each plan W_j corresponds to a given resolution.

For a multiband image $\mathcal{F}$ with N bands, a wavelet decomposition is carried out for each band b separately leading to a multiresolution pyramids $\mathcal{W}^b$, $b = 1, \cdots, N$. These N pyramids are combined in unique Multiband-Multiresolution Pyramid (MMP) $\mathcal{W}$ by considering details coefficients, $\mathcal{W}_j^1(k)$, $\cdots, \mathcal{W}_j^N(k)$, for space location k at scale j as a components of an unique vector $\mathcal{W}_j(k)$.

The (MMP) $\mathcal{W}$ feeds a vectorial hidden Markov quadtree which generates a two classes segmentation map separating significant coefficients from those associated with the noise.

[1] Burt and Adelson used a filter of length 5. If $h(0) = a$, $h(-1) = h(1) = b$, and $h(-2) = h(2) = c$ then the equal contribution requires $a + 2c = 2b$.

3. HIDDEN MARKOV QUADTREE MODEL

Let a quadtree $G = (S, L)$ be a graph composed of a set S of nodes and a set L of edges as illustrated in Fig. 2. Each node s apart from the *root* r has a unique direct predecessor, its *parent* s^-, on the path to the root. Each node s, apart from the terminal ones, the *leaves*, has four *children* s^+. The set of nodes S can be partitioned into *scales*, $S = S^0 \cup S^1 \ldots \cup S^R$, according to the path length from the *leaves* to the root . Thus, $S^R = \{r\}$, S^n involves 4^{R-n} sites, and S^0 is the finest scale formed by the leaves. This graph has the same hierarchical structure as the wavelet pyramid $\mathcal{W}$. Then each space-scale location (k, j) in $\mathcal{W}$ can be associated with a given site s in the G. In the sequel we note $\mathcal{W}_j(k) = \mathcal{W}_s$.

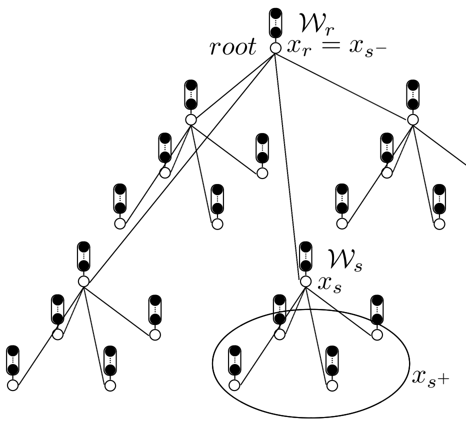

Fig. 2. Example of a dependency graph corresponding to a quadtree structure on a 4×4 lattice. White circles represent labels and black circles represent multiband observations $\mathcal{W}_s$, $s \in S$. Each node s has a unique *parent* s^-, and four *children* s^+.

Let $X = (X_s)_{s \in S}$ and $\mathcal{W} = (\mathcal{W}_s)_{s \in S}$ two stochastic processes indexed on S corresponding to the hidden states and the wavelet coefficients respectively. Each X_s takes one of two values: 0 if $\mathcal{W}_s$ is non significant and 1 otherwise. Conditionally to its hidden state X_s, $\mathcal{W}_s \in \mathbb{R}^N$ is independent from all the other nodes. Also, $\forall s \in S^n$, $\forall n \in \{0, ..., R\}$, $P(\mathcal{W}_s | x_s = i) \triangleq f_i^n(\mathcal{W}_s)$, captures the likelihood [2] of the data $\mathcal{W}_s$ given the hidden state $x_s = i$. This multidimensional likelihood has been modeled using the copulas theory as explained in the next subsection.

The process X is supposed to be Markovian in scale[3]. Moreover, given its *parent*, each hidden state is supposed to be independent from all its ascendants. From these assumptions, it can be easily seen that the joint distribution $P(x, \mathcal{W})$ can be expressed as [8, 9]:

$$P(x, \mathcal{W}) = P(x_r) \prod_{s \neq r} P(x_s | x_{s^-}) \prod_{s \in S} P(\mathcal{W}_s | x_s) \quad (4)$$

[2] In this paper the observations are introduced on M scales in the quadtree, For the other scale the likelihood equals 1: for $n > M$, $f_i^n(\mathcal{W}_s) = 1 \forall i$.

[3] To simplify notation, we will denote the probability $P(X = x)$ as $P(x)$.

where $P(x_r)$ stands for the *a priori*, probability at the root and $P(x_s|x_{s^-})$ is the parent-to-child transition probability.

The expression of $P(x, \mathcal{W})$ allows to estimate exactly and efficiently $P(x_s = i|\mathcal{W})$ for all nodes $s \in S$ by two passes on the quadtree. The segmentation label map is finally given by the Maximum *a Posteriori* Marginal criterion (MPM) as follows: $\hat{x}_s = \arg\max_{i\in\{1,2\}} P(x_s = i|\mathcal{W})$.

3.1. Copulas for likelihood computation

The basis of the copulas theory is Sklar's Theorem [4] which asserts the existence of a function C , called copula and defined on $[0, 1]^N$, binding the joint cumulative distribution function $F_i(\mathcal{W}_s^1, \cdots, \mathcal{W}_s^N)$ to the marginal cumulative distribution functions $F_i^{[1]}(\mathcal{W}_s^1), \cdots, F_i^{[N]}(\mathcal{W}_s^N)$ as follows:

$$F_i(\mathcal{W}_s^1, \cdots, \mathcal{W}_s^N) = C(F_i^{[1]}(\mathcal{W}_s^1), \cdots, F_i^{[N]}(\mathcal{W}_s^N)) \quad (5)$$

If the marginals $F_i^{[1]}, \cdots, F_i^{[N]}$ are continuous, then C is unique. Moreover, if C is differentiable it is possible to define a copula density as [4]:

$$f_i(\mathcal{W}_s^1, \cdots, \mathcal{W}_s^N) = f_i^{[1]}(\mathcal{W}_s^1) \times \cdots \times f_i^{[N]}(\mathcal{W}_s^N) \times c(F_i^{[1]}(\mathcal{W}_s^1), \cdots, F_i^{[N]}(\mathcal{W}_s^N)) \quad (6)$$

where $f_i^{[j]}(\mathcal{W}_s^j)$ is the probability density function corresponding to $F_i^{[j]}(\mathcal{W}_s^j)$ and $c = \partial C/(\partial F_i^{[1]}, \cdots, \partial F_i^{[N]})$ is the copula density. We use Gaussian copula C_G which density is given by [10, 11]:

$$\forall \mathbf{t} = (t^1, \cdots, t^N) \in \mathbb{R}^N : \quad c_G(\mathbf{t}) = |R_i|^{-\frac{1}{2}} \exp\left[-\frac{\tilde{\mathbf{t}}^T(R_i^{-1} - I)\tilde{\mathbf{t}}}{2}\right] \quad (7)$$

where $\tilde{\mathbf{t}} = (\Phi^{-1}(t^1), \cdots, \Phi^{-1}(t^N))^T$ with $\Phi(.)$ the standard Gaussian cumulative distribution, R_i is the inter-band correlation matrix within class i and I the $N \times N$ identity matrix.

To model non-Gaussian multivariate densities, we use Eq. 6 with a Gaussian copula density (Eq. 7) and Generalized Gaussian marginal densities [12] each one characterized by three parameters namely the mean, the standard deviation and the shape parameter. This modeling allows us to cover Super-Gaussian (shape parameter < 2), Gaussian (shape parameter $= 2$) and Sub-Gaussian (shape parameter > 2) densities.

3.2. Model parameter estimation

The quadtree defined in section 3 with a likelihood model described in subsection 3.1 needs a parameter estimation procedure to be unsupervised. This can be easily achieved using the segmental K-means algorithm [13] which involves iteration of two fundamental steps : segmentation and optimization. We start from an initial parameter set. The marginal *a posteriori* in each site is then computed and maximized to infer the hidden states. Using this intermediate labeling, a new parameter set is computed by maximizing the likelihood to the current model. This two steps are iterated until convergence.

Practically, optimization step is operated as follows : given a realization of X, one can :

1) count the number of sites assigned to each class at the bottom of the quadtree, to estimate the *a priori* of the classes;

2) count the number of (*parent*,*child*) node pairs for each combination $(i, j)_{i,j=1,\cdots,2}$ to estimate the transition probability.

Then, for the likelihood parameter estimation of a given class i at a given scale, we use the i-labeled observations at this scale to estimate marginals densities parameter as described in [12]. In the same way, the interband correlation matrix within each class is estimated efficiently [10].

4. FUSION RULES

When significant wavelet coefficients are selected by the classification process, we may use one rule among the followings:

1. $\forall s \in S^n : W_s^{fused} = \sum_{i=1}^N x_s \mathcal{W}_s^i$
2. $\forall s \in S^n : W_s^{fused} = \frac{\sum_{i=1}^N \sigma_i^n x_s \mathcal{W}_s^i}{\sum_{i=1}^N \sigma_i^n x_s}$, σ_i^n being the standard deviation of the i^{th} marginal of the likelihood associated with class 1 at scale n.
3. $\forall s \in S^n : W_s^{fused} = \max_i \mathcal{W}_s^i$ if $x_s = 1, 0$ otherwise.

Then, the approximation of the fused image is computed by averaging approximations of all bands.

5. RECONSTRUCTION PROCEDURE

The reconstruction using the fused wavelet coefficients W^{fused} can not be done using the Eq.3. Indeed, the structure W^{fused} does not correspond to a smooth image since all non significant coefficients are put to zero before fusion. We seek instead a smooth solution $\hat{F}^{fused}$ which minimizes $\parallel (W^{fused} - O(\hat{F}^{fused})) \parallel$ where O is the wavelet transform operator. In practice we use the Van Cittert's algorithm [5] :

- Initialisation : $p = 0$
  $\hat{F}^{[p]} = O^{-1}(W^{fused})$
  where O^{-1} is the reconstruction operator Eq. 3.

1. $W_r^{[p]} = (W^{[p]} - O(\hat{F}^{[p]})) \odot X$
   $F_r^{[p]} = O^{-1}(W_r^{[p]})$
   $\odot$ being the term by term multiplication and X the binary map obtained in section 3.
2. $\hat{F}^{[p+1]} = \hat{F}^{[p]} + F_r^{[p]}$
3. if the residual $\parallel W_r^{[p]} \parallel$ is under a given threshold value then stop, else $p = p + 1$ and goto 1.

When the algorithm converges the restored image is given by $\hat{F}^{fused} = \hat{F}^{[p+1]}$.

6. RESULTS AND CONCLUSION

The technique described above was tested on real astronomical multiband images of high-z galaxies from the Hubble Deep Field observed with the Hubble Space Telescope at six wavelengths, from the rest-frame FUV to I band (Fig.3). We perform a wavelet transform on 4 levels for each band using the pyramidal algorithm. The final restored fused image using different fusion rules are presented in Fig.4. Rules 1 and 2 clearly outperform rule 3 and the simple band averaging. The fused image summarizes the main properties of the object in a single de-noised picture and highlighting the galaxy global structure.

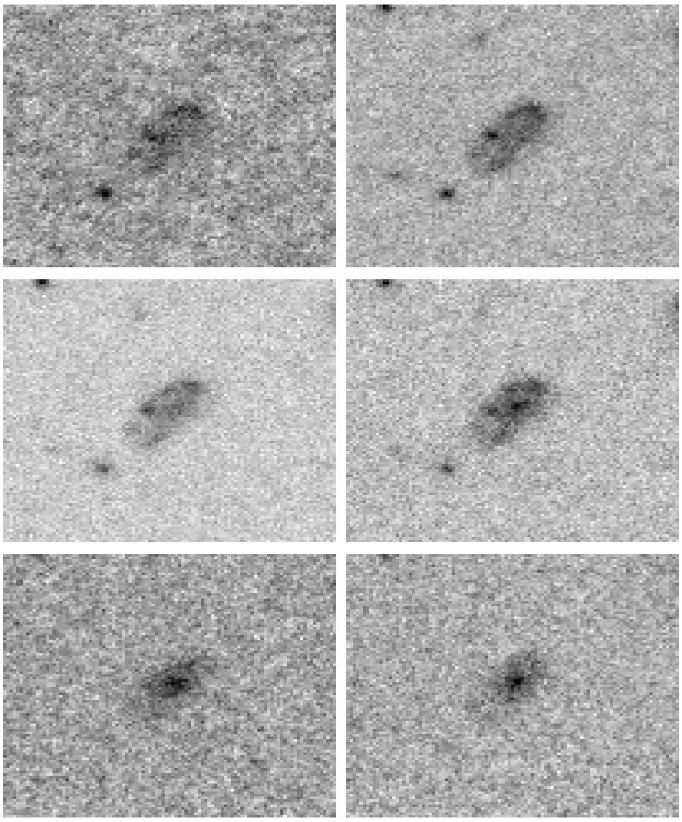

Fig. 3. Original multiband image (in inverse video): hdf4-378. 6 bands corresponding to the wavelengths (in nm): 300, 450, 606, 814, 1100 and 1600.

7. REFERENCES

[1] Z. Zhang and R. S. Blum, "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," *Proceedings of the IEEE*, vol. 87, no. 8, pp. 1315 – 1326, August 1999.

[2] D.L.Donoho and I.M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, September 1994.

[3] M.S. Crouse, R.D. Nowak, and R.G. Baraniuk, "Wavelet-based statistical signal processing using hidden markov models," *IEEE Trans. Image Processing*, vol. 46, no. 4, April 1998.

[4] R. B. Nelsen, *An introduction to copulas*, Lecture Notes in Statistics. Springer, New York, 1998.

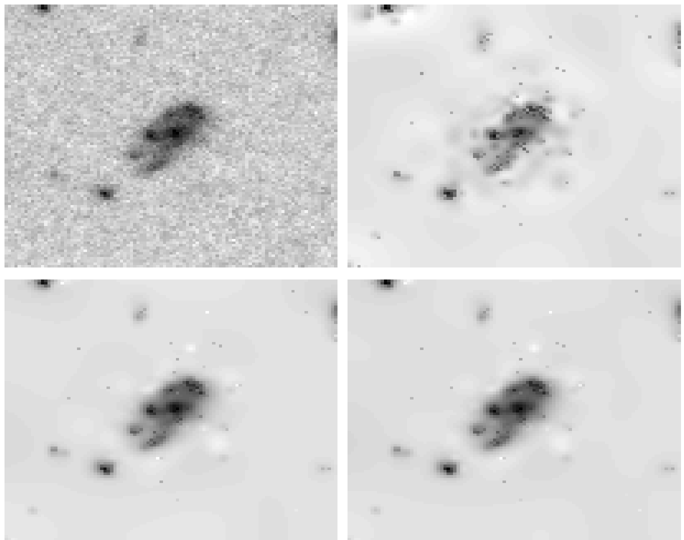

Fig. 4. Fusion result (in inverse video) of the multiband image of Fig.3 using different rules. Top left : a simple averaging of the bands. Top right : proposed technique with rule 3 Sect. 4. Down left : proposed technique with rule 1 Sect. 4. Down right : proposed technique with rule 2 Sect. 4.

[5] J.-L. Starck, F. Murtagh, and A. Bijaoui, *Image Processing and Data Analysis: The Multiscale Approach*, Cambridge University Press, 1998.

[6] S. Mallat, *A wavelet tour of signal processing*, Academic Press, 1998.

[7] P. J. Burt and E. H. Adelson, "The laplacian pyramid as a compact image code," *IEEE Trans. Communications*, vol. 31, no. 4, pp. 532–540, April 1983.

[8] J.-M. Laferté, P. Pérez, and F. Heitz, "Discrete markov image modeling and inference on the quad-tree," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 390–404, March 2000.

[9] M. Luettgen, *Image Processing with Multiscale Stochastic Models*, Phd thesis, MIT Laboratory of Information and Decision Systems, May 1993.

[10] K. K. Phoon, "Application of fractile correlations and copulas to non-gaussian random vectors," in *Proceedings, Second International ASRANet(Network for Integrating Structural Safety, Risk, and Reliability) Colloquium*, Barcelona, Spain, July 2004.

[11] N. Brunel and W. Pieczynski, "Unsupervised signal restoration using copulas and pairwise Markov chains," Saint-Loui Missouri, 2003, IEEE Workshop on Statistical Signal Prrocessing.

[12] J.-N Provost, Ch.Collet, P. Rostaing, P. Pérez, and P. Bouthemy, "Hierarchical markovian segmentation of multispectral images for the reconstruction of water depth maps," *Computer Vision and Image Understanding*, vol. 93, no. 2, pp. 155–174, February 2004.

[13] B. H. Juang and L. R. Rabiner, "The segmental k-means algorithm for estimating parameters of hidden markov models," *IEEE Trans. On ASSP*, vol. 38, no. 9, pp. 1639 1641, September 1990.