Grands Relevés et Observatoires Virtuels
Ecole CNRS de Goutelas XXIV (2001)
Edité par D. Egret et J.-L. Halbwachs

Astronomical Data Analysis : Perspectives from Multiscale Analysis

Jean-Luc Starck

DAPNIA/SEI-SAP, CEA-Saclay, F-91191 Gif sur Yvette, France

Abstract. Wavelets have been used extensively for several years now in astronomy for many purposes, ranging from data filtering and deconvolution, to star and galaxy detection or cosmic ray removal. We review in this paper a range of methods and applications. New recent methods, such the ridgelets and the curvelets are also described, and we show their interests when the data present anisotropic features.

Keywords

Wavelet, ridgelet, curvelet, multiscale methods, detection, compression, multi channel, filtering, deconvolution.

1. Introduction

The wavelet transform has been extensively used in astronomical data analysis during the last ten years. A quick search with ADS shows that around 500 papers contain the keyword "Wavelet" in their abstract, and all astrophysical domains were concerned, from the sun study to the CMB analysis :

- Sun : Active region oscillations (Ireland et al., 1999; Blanco et al., 1999), solar cycle length variations determination (Fligge et al., 1999), feature extraction from solar images (Irbah et al., 1999), Velocity fluctuations (Lawrence et al., 1999).
- Solar system : Asteroidal resonant motion (Michtchenko and Nesvorny, 1996), classification of asteroids (Bendjoya, 1993), Saturn and Uranus ring analysis (Bendjoya et al., 1993; Petit and Bendjoya, 1996)
- Star studies : Ca II feature detection in magnetically active stars (Soon et al., 1999), variable star research (Szatmary et al., 1996).
- Interstellar medium : large scale extinction maps of giant molecular clouds using optical star counts (Cambrésy, 1999), fractal structure analysis in molecular clouds (Andersson and Andersson, 1993).

- Planetary nebula detection : confirmation of the detection of a faint planetary nebula around IN Com (Brosch and Hoffman, 1999), evidence for extended high energy gamma-ray emission from the Rosette/Monoceros Region (Jaffe et al., 1997).
- Galaxy : evidence for a Galactic gamma-ray halo (Dixon et al., 1998).
- QSO : QSO brightness fluctuations (Schild, 1999), detecting the non-Gaussian spectrum of QSO Ly$_{a}lpha$ absorption line distribution (Pando and Fang, 1998).
- Gamma-ray Burst : GRB detection (Kolaczyk, 1997; Norris et al., 1994) and GRB analysis (Greene et al., 1997; Walker et al., 2000).
- black hole : periodic oscillation detection (Steiman-Cameron et al., 1997; Scargle, 1997)
- galaxies : Starburst detection (Hecquet et al., 1995), galaxies count (Aussel et al., 1999; Damiani et al., 1998), morphological of galaxies (Weistrop et al., 1996; Kriessler et al., 1998), multifractal character of the galaxy distribution (Martinez et al., 1993).
- galaxy cluster : sub-structure detection (Pierre and Starck, 1998; Krywult et al., 1999; Arnaud et al., 2000), hierarchical clustering (Pando et al., 1998a), distribution of superclusters of galaxies (Kalinkov et al., 1998).
- Cosmic Microwave Background : evidence for scale-scale correlations in the Cosmic Microwave Background radiation in COBE data (Pando et al., 1998b), large-scale CMB non-Gaussian statistics (Popa, 1998; Aghanim et al., 2001), massive CMB data set analysis (Gorski, 1998)
- Cosmology : Comparing simulated cosmological scenarios with observations (Lega et al., 1996), cosmic velocity fields analysis (Rauzy et al., 1993).

This large success of the wavelet transform (WT) is due to the fact that astronomical data presents generally complex hierarchical structures, often described as fractals. Using multiscale approaches such as the wavelet transform (WT), an image can be decomposed into components at different scales, and the WT is therefore well-adapted to astronomical data study.

The following section presents the different WT algorithms which can be used. In section 3., we discuss how noise, which is always present in astronomical images, is managed. In section 4., we review some wavelet based applications. The following section presents the differents WT algorithms which can be used. In section 3., we discuss how noise, which is always present in astronomical images, is managed. Other sections describe how filtering, deconvolution, detection, and astronomical

image compression are actually performed. New recent multiscale methods, such the ridgelets and the curvelets are describe in section 6., and we show their interests when the data present anisotropic features.

2. The Wavelet Transform

There are many WT algorithms (Mallat, 1998; Starck et al., 1998a). The most well-known are :

- The (bi-) orthogonal wavelet transform. This wavelet transform (Mallat, 1989), often referred to as the Fast Wavelet Transform (FWT), is certainly the most widely used among available discrete wavelet transform algorithms. It is a non-redundant representation of the information. An introduction to this type of transform can be found in (Daubechies, 1992). The famous Haar transform belongs to this class. Using the FWT, a signal s can be decomposed by :

$$s(l) = \sum_k c_{J,k}\phi_{J,l}(k) + \sum_k \sum_{j=1}^{J} \psi_{j,l}(k) w_{j,k} \quad (1)$$

  with $\phi_{j,l}(x) = 2^{-j}\phi(2^{-j}x - l)$ and $\psi_{j,l}(x) = 2^{-j}\psi(2^{-j}x - l)$, where ϕ and ψ are respectively the scaling function and the wavelet function. J is the number of resolutions used in the decomposition, w_j the wavelet (or details) coefficients at scale j, and c_J is a coarse or smooth version of the original signal s.
- The Feauveau wavelet transform. Feauveau (Feauveau, 1990) introduced quincunx analysis. This analysis is not dyadic and allows an image decomposition with a resolution factor equal to $\sqrt{2}$.
- The à trous algorithm : the wavelet transform of an image by this algorithm produces, at each scale j, a set $\{w_j\}$. This has the same number of pixels as the input data set. The original data c_0 can be expressed as the sum of all the wavelet scales and the smoothed array c_p by $c_0 = c_J + \sum_{j=1}^{J} w_j$ and a pixel at position k can be expressed also as the sum of all the wavelet coefficients at this position, plus the smoothed array : $c_{0,k} = c_{J,k} + \sum_{j=1}^{J} w_{j,k}$.

The multiresolution median transform is an alternative to the wavelet transform for multiscale decomposition. The median transform is nonlinear (so it is not a wavelet transform, but represents an image in a similar way to the à trous algorithm), and offers advantages for robust smoothing (i.e. the effects of outlier pixel values are mitigated). The multiresolution median transform (Starck et al., 1996) consists of a series of smoothings of the input image, with successively broader kernels. Each successive smoothing provides a new resolution scale. The multiresolution coefficient values constructed from differencing images at successive

resolution scales are not necessarily of zero mean, and so the potential artifact-creation difficulties related to this aspect of wavelet transforms do not arise. For integer input image values, this transform can be carried out in integer arithmetic only which may lead to computational savings. As in the case of the à trous algorithm, the original image can be expressed by a sum of the scales and the smoothed array.

All these methods have advantages and drawbacks. Following the content of the image, and the nature of the noise, each of these transforms can be considered as optimal. In astronomical images, there are generally no edges, and objects are relatively diffuse. This is the reason why an isotropic or symmetric analysis produces better results. Furthermore, for the most usual applications (detection, filtering, deconvolution, etc.), undersampling leads to severe artifacts which can be easily avoided by non-orthogonal transforms such the à trous algorithm. For these reasons, the FWT is rarely used in the astronomical domain.

3. Significant wavelet coefficients

Astronomical data are always contaminated by a noise, and it is important to detect the wavelet coefficients which are "significant", i.e. the wavelet coefficients which have an absolute value too large to be due to noise. We defined the multiresolution M^D of the data set D by :

$$M^D_{j,k} = \begin{cases} 1 & \text{if } w_{j,k} \text{ is significant} \\ 0 & \text{if } w_{j,k} \text{ is not significant} \end{cases} \tag{2}$$

where j is the scale, k the pixel position, and $w_{j,k}$ the wavelet coefficient of D at scale j and at position k. We need now to determine when a wavelet coefficient is significant. For Gaussian noise, it is easy to derive an estimation of the noise standard deviation σ_j at scale j from the noise standard deviation, which can be evaluated with good accuracy in an automated way (Starck and Murtagh, 1998). To detect the significant wavelet coefficients, it suffices to compare the wavelet coefficients $w_{j,k}$ to a threshold level t_j. t_j is generally taken equal to $k\sigma_j$, and k is chosen between 3 and 5. The value of 3 corresponds to a probability of false detection of 0.27%. If $w_{j,k}$ is small, then it is not significant and could be due to noise. If $w_{j,k}$ is large, it is significant :

$$\begin{array}{ll} \text{if } \mid w_{j,k} \mid \geq t_j & \text{then } w_{j,k} \text{ is significant} \\ \text{if } \mid w_{j,k} \mid < t_j & \text{then } w_{j,k} \text{ is not significant} \end{array} \tag{3}$$

Other thresholding approaches have been proposed, like the *universal threshold* (Donoho and Johnstone, 1993), or the SURE method (Coifman and Donoho, 1995), but they generally do not produce as good results as the k-sigma method.

When the noise is not Gaussian, other strategies may be used :

- **Poisson noise :** if the noise in the data I is Poisson, the transformation (Anscombe, 1948) $t(I) = 2\sqrt{I + \frac{3}{8}}$ acts as if the data arose from a Gaussian white noise model, with $\sigma = 1$, under the assumption that the mean value of I is sufficiently large. However, this transform has some limits and it has been shown that it cannot be applied for data with less than 20 photons per pixel. So for X-ray or gamma ray data, other solutions have to be chosen, which manage the case of a reduced number of events or photons under assumptions of Poisson statistics
- **Gaussian + Poisson noise :** the generalization of variance stabilization (Murtagh et al., 1995) is :

$$t(I(x,y)) = \frac{2}{\alpha}\sqrt{\alpha I(x,y) + \frac{3}{8}\alpha^2 + \sigma^2 - \alpha g}$$

where α is the gain of the detector, and g and σ are the mean and the standard deviation of the read-out noise.

- **Poisson noise with few events using the à trous transform :** if a wavelet coefficient $w_{j,k}$ is due to noise, it can be considered as a realization of the sum $\sum_{l \in L} n_l$ of independent random variables with the same distribution as that of the wavelet function (n_k being the number of events used for the calculation of $w_{j,k}$). This allows comparison of the wavelet coefficients of the data with the values which can be taken by the sum of n independent variables.
  The distribution of one event in wavelet space is then directly given by the histogram H_1 of the wavelet ψ. As we consider independent events, the distribution of a coefficient w_n (note the changed subscripting for w, for convenience) related to n events is given by n autoconvolutions of H_1 :

$$H_n = H_1 \otimes H_1 \otimes ... \otimes H_1 \tag{4}$$

  For a large number of events, H_n converges to a Gaussian.

- **Poisson noise with few events using the Haar Transform :** the Haar transform presents the advantage of its simplicity for modeling Poisson noise. A wavelet coefficient is only the difference of two values which both follow a Poisson distribution. The probability density function of an unnormalized wavelet coefficient which is given by (Jammal and Bijaoui, 1999) :

$$p(w_j = \nu) = e^{-2^{2j}\lambda} I_\nu(2^{2j}\lambda) \tag{5}$$

  where $I_\nu(x)$ is the modified Bessel function of integer order ν. For a given false detection rate α. The threshold t_j can be derived from

this pdf. Kolaczyk (Kolaczyk, 1997; Kolaczyk and Dixon, 2000) has derived thresholds which can be directly used : for 1D data, and using a normalized Haar transform (L2-normalization), he proposed the to use a detection level for the scale j equal to :

$$t_j = 2^{\frac{-(j+2)}{2}}(2\log(n_l + \sqrt{(4\log n_l)^2 + 8\log n_l \lambda_l}) \quad (6)$$

where $n_l = 2^{\log_2 n - 1}$, n being the number of samples, and λ_l the background rate per bin in n_l, and in 2D

$$t_j = 2^{-(j+1)}[z^2_{\alpha/2} + \sqrt{z^4_{\alpha/2} + 4\lambda_j z^2_{\alpha/2}}] \quad (7)$$

where α is the false detection rate, $\lambda_j = 2^{2j}\lambda$ is the background rate over $n_j = 2^{2j}$ pixels, and $z\alpha/2$ is the point under the Gaussian density function for which there falls $\alpha/2$ mass in the tails beyond each $z\alpha/2$. More details can be found in (Kolaczyk and Dixon, 2000).

- **Model and Simulation :** simulations can be used for deriving the probability that a wavelet coefficient is not due to the noise (Escalera et al., 1992). Modeling a sky image (i.e. uniform distribution and Poisson noise) allows determination of the wavelet coefficient distribution and derivation of a detection threshold. For substructure detection in a cluster, the large structure of the cluster must be first modeled, or otherwise noise photons related to the large scale structure will introduce false detections at lower scales. If we have a physical model, Monte Carlo simulations can also be used (Escalera and Mazure, 1992; Grebenev et al., 1995), but this approach requires a long computation time, and the detections will always be model-dependent. Damiani et al. (Damiani et al., 1997), and also Freeman et al. (Freeman et al., 1996) propose to calculate the background from the data in order to derive the fluctuations due to the noise in the wavelet scales.
- **Root mean square map** The root mean square (RMS) map $R_\sigma(x,y)$ is often part of the data, For each wavelet coefficient $w_j(x,y)$ of R, the exact standard deviation $\sigma_j(x,y)$ has to be calculated from the root mean square map $R_\sigma(x,y)$ (Starck et al., 1999).

  A wavelet coefficient $w_j(x,y)$ is obtained by the correlation product between the image R and a function g_j :

$$w_j(x,y) = \sum_k \sum_l R(x,y) g_j(x+k, y+l) \quad (8)$$

  Then we have :

$$\sigma^2_j(x,y) = \sum_k \sum_l R^2_\sigma(x,y) g^2_j(x+k, y+l) \quad (9)$$

In the case of the à trous algorithm, the coefficients $g_j(x,y)$ are not known exactly, but they can be computed by taking the wavelet transform of a Dirac function (w^δ, in our notation) :

$$g_j(x,y) = w_j^\delta(x,y) \tag{10}$$

Then the map σ_j^2 is calculated by correlating the square of the wavelet scale j of w^δ with $R_\sigma^2(x,y)$.

- **Non-stationary additive noise :** the noise is assumed to be locally Gaussian. So we must consider one noise standard deviation per pixel, and treat the problem as previously. The $R_\sigma(x,y)$ map can be obtained by taking the standard deviation in a box around each pixel.
- **Non-stationary multiplicative noise :** the image is first log-transformed. Then the transformed image is treated as an image with non-stationary additive noise.
- **Stationary correlated noise :** the noise is stationary, but correlated. This noise modeling requires a noise map, containing a realization of the noise. The threshold at a scale j S_j is found by computing the wavelet transform of the noise map, and using the histogram of S_j to derive the noise probability density function, pdf, of S_j.
- **Undefined stationary noise :** in this case, the data can be treated as for the Gaussian case, but the noise standard deviation σ_j at scale j is calculated independently at each scale. Two methods can be used :
  1. σ_j can be derived from a k-sigma clipping method applied at scale j.
  2. the absolute median deviations (MAD) can be used as an estimator of the noise standard deviation :

$$\sigma_j = \text{median}(\mid w_j \mid)/0.6745 \tag{11}$$

- **Undefined noise :** the standard deviation is estimated for each wavelet coefficient, by considering a box around it, and the calculation of σ is done in the same way as for non-stationary additive noise. The latter determines a map of variances for the image, and then derives the variances for the wavelet coefficients. "Undefined noise" does not assume additivity of the noise, and so calculates the noise from local variance in the resolution scales.

4. Wavelet based Methods in Astronomical Data Processing

4.1 Filtering

The most used filtering method is the hard thresholding, which consists of setting to 0 all wavelet coefficients which have an absolute value lower than a threshold t_j

$$\tilde{w}_{j,k} = \begin{cases} w_{j,k} & \text{if } w_{j,k} > t_j \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

We define the function $\mathcal{T}$ as the function which set to zero all wavelet coefficients outside a given multiresolution support M :

$$\mathcal{F}(M, x) = c_{J,k} + \sum_{j=1}^{J} M_{j,k} w_{j,k} \tag{13}$$

where $c_{J,k}$ and $w_{j,k}$ are obtained from the à trous wavelet transform of x. The filtered version $\tilde{s}$ of the input signal s is obtained by $\tilde{s} = \mathcal{F}(M^s, s)$, M^s being the multiresolution support of s. This solution can be refined by the following iterative scheme :

$$\tilde{s}^{n+1}(k) = s^n(k) + \mathcal{F}(M^s, r^n) \tag{14}$$

where $r^n = s - \tilde{s}^n$. This algorithm allows us to constraint the residual to have a zero value inside the the multiresolution support of s (Starck et al., 1998a). For astronomical image filtering, iterating improves significantly the results, especially for the photometry.

4.2 Image deconvolution

Observed data Y in the physical sciences are generally corrupted by noise, which is often additive and which follows in many cases a Gaussian distribution, a Poisson distribution, or a combination of both. Using Bayes' theorem to evaluate the probability of the realization of the original signal X, knowing the data Y, we have

$$Prob(X|Y) = \frac{Prob(Y|X).Prob(X)}{Prob(Y)} \tag{15}$$

$Prob(Y|X)$ is the conditional probability of getting the data Y given an original signal X, i.e. it represents the distribution of the noise.

The denominator in equation (15) is independent of X and is considered as a constant (stationary noise). $Prob(X)$ is the a priori distribution of the solution X. In the absence of any information on the solution X except its positivity, a possible course of action is to derive the probability of X from its entropy. Several definitions of entropy has been proposed, and the main ones are :

- Burg (Burg, 1978) : $H_b(X) = -\sum_{pixels} \ln(X)$
- Frieden (Frieden, 1978) : $H_f(X) = -\sum_{pixels} X \ln(X)$
- Gull and Skilling (Gull and Skilling, 1991) : $H_g(X) = \sum_{pixels} X - M - X \ln(X|M)$

Each of these entropies can be used, and they correspond to different probability distributions that one can associate with an image (Narayan and Nityananda, 1986). The last definition of the entropy above has the advantage of having a zero maximum when X equals the model M, usually taken as a flat image. All of these entropy measures are negative, and maximum when the image is flat. They are negative because an offset term is omitted which has no impact on the final solution. The fact that we consider that a signal has maximum information value when it is flat is evidently a curious way to measure information (at least from a physical point of view).

It was shown in (Narayan and Nityananda, 1986) that results vary strongly with the background level, and that these entropy functions produce poor results for negative structures, i.e. structures under the background level (absorption area in an image, absorption band in a spectrum, etc.), and compact structures in the signal. The Gull and Skilling entropy gives rise to the difficulty of estimating a model. Furthermore it was shown in (Bontekoe et al., 1994) that the solution is dependent on this choice.

Many studies (Weir, 1991; Weir, 1992; Bontekoe et al., 1994; Pantin and Starck, 1996) have been carried out in order to improve the functional to be minimized. But the question which should be raised is : what is a good entropy measure for signal restoration ?

In (Starck et al., 1998b), the benchmark properties for a good "physical" definition of entropy were discussed. Assuming that a signal X is the sum of several components : $X = S + B + N$, where S is the signal of interest, B the background, and N the noise, we proposed that the following criteria should be verified :

1. The information in a flat signal is zero ($S = 0$, $N = 0$ and $B = \text{Cst}$).
2. The amount of information in a signal is independent of the background (i.e., $H(X)$ is independent of B).
3. The amount of information is dependent on the noise (i.e., $H(X)$ is dependent on N). A given signal X does not furnish the same information in the different cases where the noise N is high or small.
4. The entropy must work in the same way for a pixel which has a value $B + \epsilon$, and for a pixel which has a value $B - \epsilon$. $H(X)$ must be a function of the absolute value of S instead of S.
5. The amount of information is dependent on the correlation in the signal. If the signal S presents large features above the noise, it

contains a lot of information. By generating a new set of data from S, by randomly taking the pixel values in S, the large features will evidently disappear, and this new signal will contain less information. But the pixel values will be the same as in S.

The Burg and Frieden entropy functions do not verify any of these criteria, and the Skilling one verifies only point 2. Using the wavelet transform, it has been shown (Starck et al., 1998b; Starck and Murtagh, 1999; Starck et al., 2001b) that an entropy function verifying all cited properties can be obtained, which produces very good results.

4.3 Interferometric Image Deconvolution

In interferometric imaging, measurements are carried out in Fourier space but the (u, v) plane is not completely covered. The image, called the dirty map, is obtained by a simple inverse Fourier transform of the data and the PSF, called the dirty beam, by an inverse Fourier transform of the (u, v) plane coverage. The presence of secondary lobes in the dirty beam creates very serious artifacts in the dirty map and a deconvolution is necessary. By applying the CLEAN method (Högbom, 1974) at each scale of the wavelet transform using the FFT, we can localize significant structures, and an iterative reconstruction algorithm allows solutions to be found which satisfy the positivity constraint, and the constraint of fidelity to measurements (i.e. at each measured $V_m(u, v)$ $+/-$ $\Delta_m(u, v)$, we require that the solution O satisfies $\mid \hat{O}(u, v) - V_m(u, v) \mid < \Delta_m(u, v)$). More details can be found in (Starck et al., 1994).

4.4 Object detection

Using the à trous algorithm, an image I can be expressed as the sum of all the wavelet scales and the smoothed array c_J by the expression

$$I(k, l) = c_{J,k,l} + \sum_{j=1}^{J} w_{j,k,l}. \tag{16}$$

Hence, we have a *multiscale pixel representation*, i.e. each pixel of the input image is associated to a set of pixels of the multiscale transform. A further step is to consider a *multiscale object representation*, which would associate to an object contained in the data, a volume in the multiscale transform. Such a representation obviously depends on the kind of image we need to analyze. A Multiscale Vision Model (MVM) has been developed (Bijaoui and Rué, 1995) for astronomical data. Using the MVM, an image I can be decomposed, from its wavelet transform, into a set of components :

$$I(k, l) = \sum_{i=1}^{N_o} O_i(k, l) + B(k, l) + N(k, l) \tag{17}$$

where N_o is the number of object, O_i are the objects contained in the data (stars galaxies, etc), B is the background image, and N is the noise. Furthermore, it has been shown (Starck et al., 2000) that a deconvolution can be introduced in this decomposition, and the set of components verifies :

$$I(k,l) = \sum_{i=1}^{N_o} (P_i * O_i)(k,l) + B(k,l) + N(k,l) \tag{18}$$

where P_i is the Point Spread Function (PSF) associated to the object i. We have therefore an elegant solution for the deconvolution with a spatially variant PSF.

4.5 Astronomical Image Compression

The principle of the compression method is to select the information we want to keep, by using the Pyramidal Median Transform (Starck et al., 1996), and to code this information without any loss. Thus the first phase searches for the minimum set of quantized multiresolution coefficients which produce an image of "high quality". Criteria which can be retained for estimating the quality of a compression method in astronomy are :

1. Visual aspect.
2. Signal to noise ratio.
3. Photometry (i.e. integrated pixel values in a given object).
4. Astrometry (i.e. object positions).
5. Detection of real and faint objects.
6. Objects morphology.

Lost information cannot be recovered, so if we do not accept any loss, we have to compress what we take as noise too, and the compression ratio will be low (3 or 4 only). The PMT compression method employed involves the following sequence of operations :

1. Determination of the significant coefficients.
2. Determination of the quantized multiresolution coefficients which give the filtered image.
3. Coding of each resolution level using the Huang-Bijaoui method (Huang and Bijaoui, 1991). This consists of quadtree-coding each image, followed by Huffman-coding the quadtree representation. There is no information lost during this phase.
4. Compression of the noise if this is wished.
5. Decompression consists of reconstituting the noise-filtered image (plus the compressed noise if this was specified).

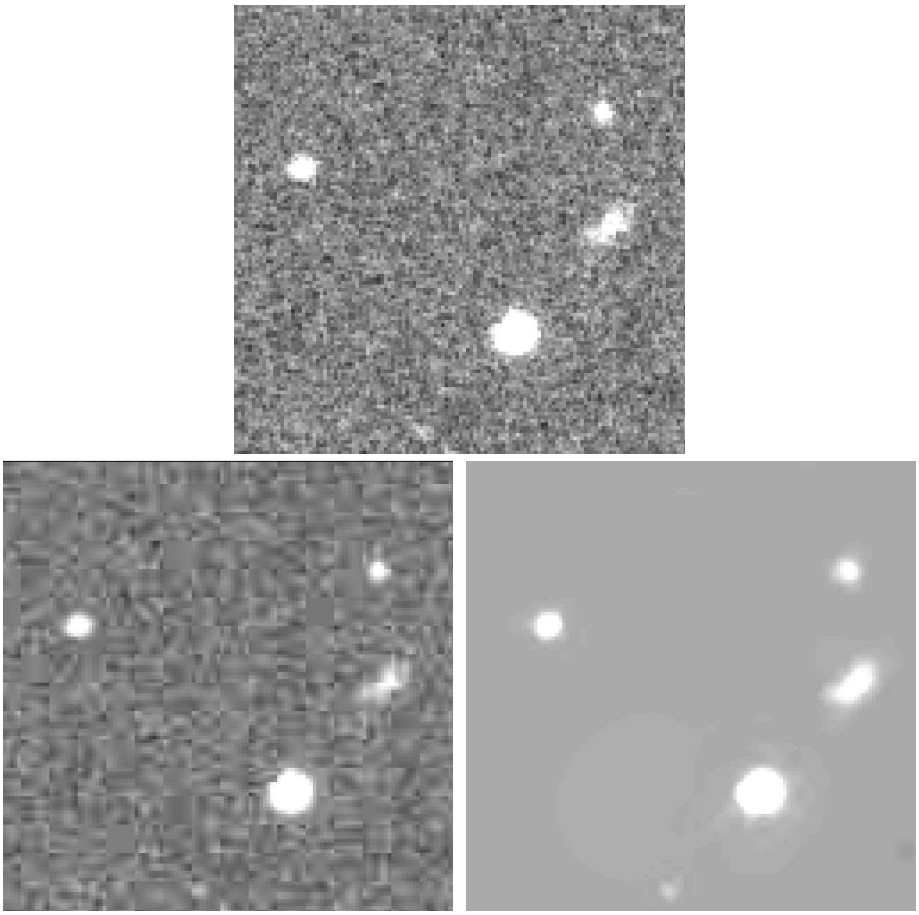


Figure 1.: *Top, original image, bottom left, JPEG compressed image at 40 :1, and bottom right, PMT compressed image at 260 :1.*

The decompression is carried out scale by scale, starting from the low resolution, so it is not necessary to decompress the entire file if one is just interested in having a look at the image. Noise is decompressed and added at the end, if this is wanted. See the full description of the method in (Starck et al., 1996; Murtagh et al., 1998).

Various studies have been carried out in order to determine the best compression method for astronomical images, especially in the framework of the ALADIN project (Bonnarel et al., 1999). It has been shown, using a sample of nearly 2000 stars from an ESO Schmidt plate centered on the globular cluster M5, that the PMT can give compression ratios of up to 5 times the maximum ratio obtained from other methods (Louys et al., 1999), such JPEG or wavelet based methods.

Figure 1 top shows a subimage of a Schmidt photographic plate of the region of M5 (numbered ESO 7992v). This photographic plate was digitized using the MAMA ("Machine Automatique à mesurer pour l'Astronomie") machine of Paris Observatory (Paris, France).
Figures 1 bottom left and right allow the visual quality of both the JPEG (compression ratio 40 :1) and PMT (compression ratio 260 :1) to be compared.

5. Multichannel Data

The challenge for multichannel data restoration is to have a data representation which takes into account at the same time both the spatial and the spectral (or temporal) correlation.

The Wavelet-Karhunen-Loève Transform (WT-KLT) (Starck and Querre, 2001) allows us to decorrelate first the data in the spatial domain using the WT and afterwards in the spectral domain using the KLT. From a noise modeling in the WT-KLT space, the data can be filtered in an efficient way.

Considering a vector $D = d_1, ..., d_L$ of L signals or images of dimension N (i.e. N pixels per image), the WT-KLT algorithm has the following steps :

1. Calculate the wavelet transform $W^{(l)}$ of each input data set d_l.
2. For each band j of the wavelet transform, calculate the correlation matrix $C^{(j)}$ relative to the vector $x_j = \left\{W_j^{(1)}, W_j^{(2)}, ..., W_j^{(L)}\right\}$, where $W_j^{(l)}$ represents the band j of the wavelet transform $W^{(l)}$ of d_l.
3. For each band j, we diagonalize the matrix $C^{(j)}$ and build the transform matrix A_j from the eigenvectors of $C^{(j)}$.
4. For each band j and each position k, we apply the matrix A_j to the vector $x_{j,k} = \{w_{1,j,k}, w_{2,j,k}, ..., w_{L,j,k}\}$:
$$y_{j,k} = A_j x_{j,k} \tag{19}$$
5. The WT-KLT coefficients $c_{l,j,k}$ are derived from $y_{j,k}$ by $c_{l,j,k} = y_{j,k}(l)$. The l index in the transformed coefficients no longer represents the observation number, but instead the eigenvector number. $l = 1$ indicates the main eigenvector while $l = L$ indicates the last one.

Figure 2 shows the flowchart of the WT-KLT transform. Figure 3 shows a simulation. We have created a dataset of 18 frames, each of them containing a source at the same position, but at different intensity levels.

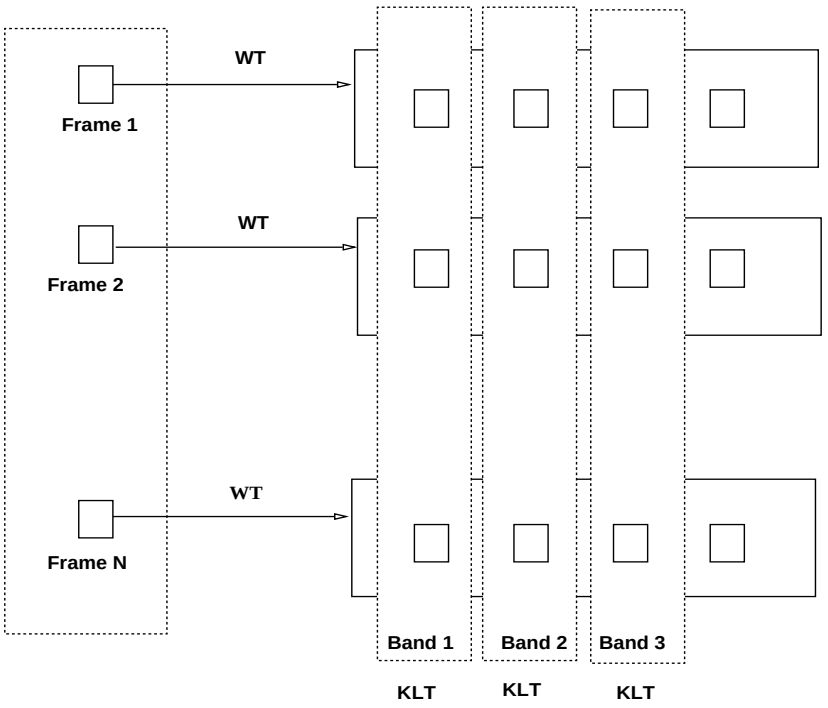
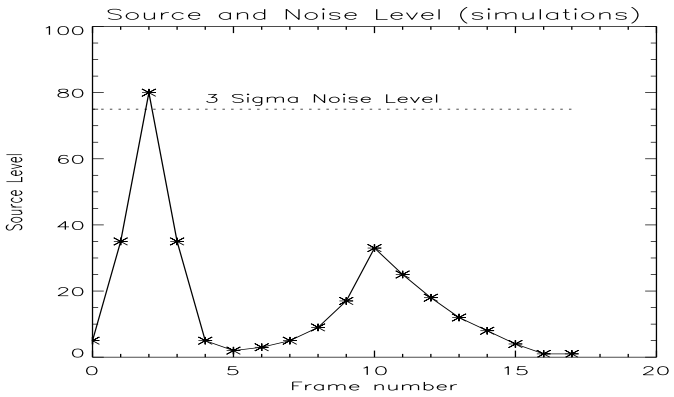


Figure 2.: *WT-KLT transform flowchart. Each frame of the input dataset is first wavelet transformed, and a principal component analysis is applied at each resolution level.*

The source is a small Gaussian, and the source SNR is defined as the ratio between the maximum of the source and the noise standard deviation. Figure 3 (top) shows the evolution of the SNR in the 18 frames. The maximum level of the source is plotted versus the frame number. We see that the source intensity first increases in the first three frames until it reaches a level equals to three times the noise standard deviation. Then it decreases and the source almost disappears around frame five, increases again until frame 10, and disappears completely in the last frames. Frames two and ten are shown in Figure 3 (middle) left and right. The source SNR ratio is respectively three and one. We see that the source is undetectable by eye in the frame 10. Figure 3 bottom left and right shows respectively the filtered frame ten by the wavelet transform and the WT-KLT. The WT detects only noise, while the WT-KLT clearly identifies the source.



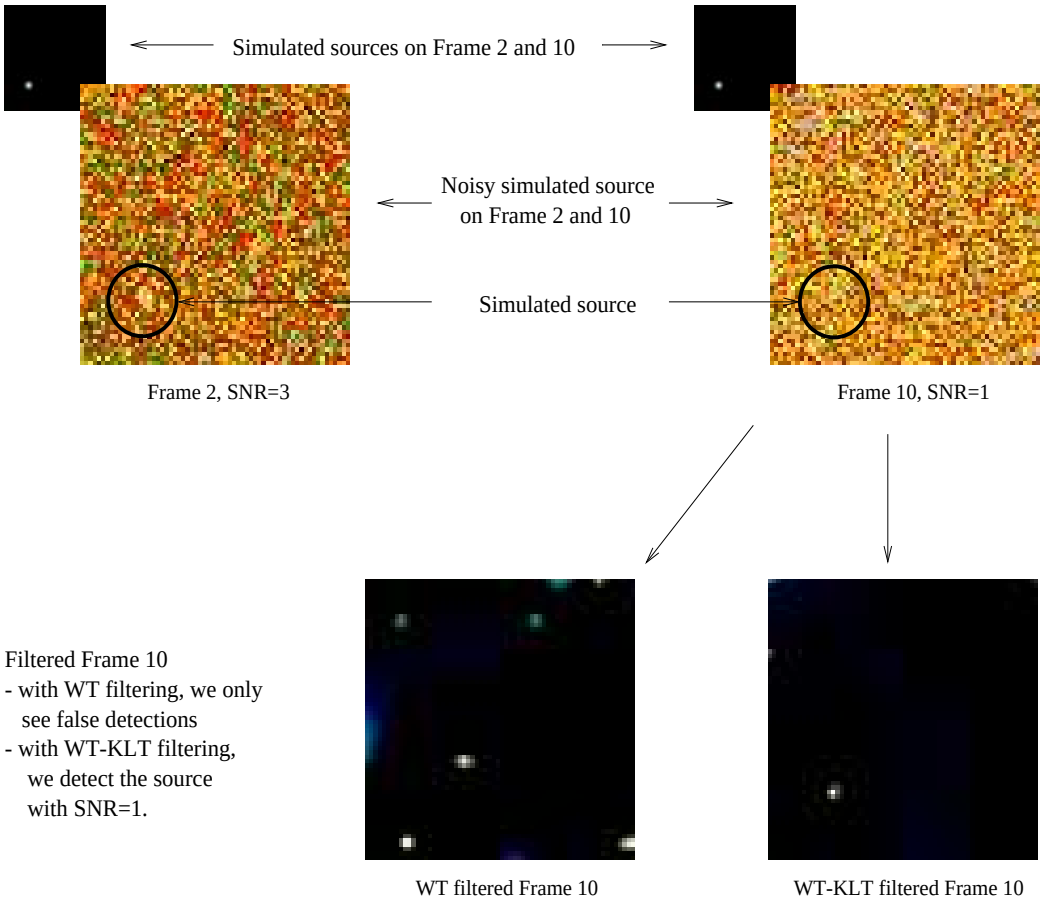


Figure 3.: *Simulation : the dataset is composed of 18 frames. Each of them contains a source (small Gaussian) at the same position, but at different intensity levels. Top, plot of the source maximum value versus the frame number. Middle, frames 2 and 10, and bottom, filtered version of the frame 10 by the wavelet transform and wavelet Karhunen-Loève transform.*

6. The Ridgelet Transform

The application of the à trous algorithm has lead to impressive results, compared to previous methods, for data restoration and object detection. As it was discussed before, this wavelet transform is well adapted to the analysis of isotropic features. However, all features included in 2D and 3D astronomical data set are not isotropic (filaments, elongated galaxies, planetary images, arclet, ...). The FWT may be better than the à trous for such data set, but still presents some limitations which may impact in some applications. Indeed, if the FWT performs better than the FFT to represent edges in an image, it is still not optimal. There is only a fixed number of directional elements independent of scales, and there is no highly anisotropic elements. For instance, the Haar 2D wavelet transform is optimal to find features with a ratio $length/width = 2$, and a horizontal,vertical, or diagonal orientation. This problem have lead to the development of other multiscale representations, like the ridgelet (Candès and Donoho, 1999a) or the curvelet transform (Donoho and Duncan, 2000).

The two-dimensional continuous ridgelet transform in $\mathbf{R}^2$ can be defined as follows (Candès and Donoho, 1999a). We pick a smooth univariate function $\psi : \mathbf{R} \rightarrow \mathbf{R}$ with sufficient decay and satisfying the admissibility condition

$$\int |\hat{\psi}(\xi)|^2/|\xi|^2 \, d\xi < \infty, \tag{20}$$

which holds if, say, ψ has a vanishing mean $\int \psi(t)dt = 0$. We will suppose that ψ is normalized so that $\int |\hat{\psi}(\xi)|^2 \xi^{-2} d\xi = 1$.

For each $a > 0$, each $b \in \mathbf{R}$ and each $\theta \in [0, 2\pi)$, we define the bivariate *ridgelet* $\psi_{a,b,\theta} : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ by

$$\psi_{a,b,\theta}(x) = a^{-1/2} \cdot \psi((x_1 \cos\theta + x_2 \sin\theta - b)/a); \tag{21}$$

Figure 6. (upper left) shows an example ridgelet function. Figure 6. upper right, and bottom left shows the the same function after rotation and rescaling. This function is constant along lines $x_1 \cos\theta + x_2 \sin\theta = const.$ Transverse to these ridges it is a wavelet (see figure 6. bottom right).

Given an integrable bivariate function $f(x)$, we define its ridgelet coefficients by

$$\mathcal{R}_f(a, b, \theta) = \int \overline{\psi}_{a,b,\theta}(x) f(x) dx.$$

We have the exact reconstruction formula

$$f(x) = \int_0^{2\pi} \int_{-\infty}^{\infty} \int_0^{\infty} \mathcal{R}_f(a, b, \theta) \psi_{a,b,\theta}(x) \frac{da}{a^3} db \frac{d\theta}{4\pi} \tag{22}$$

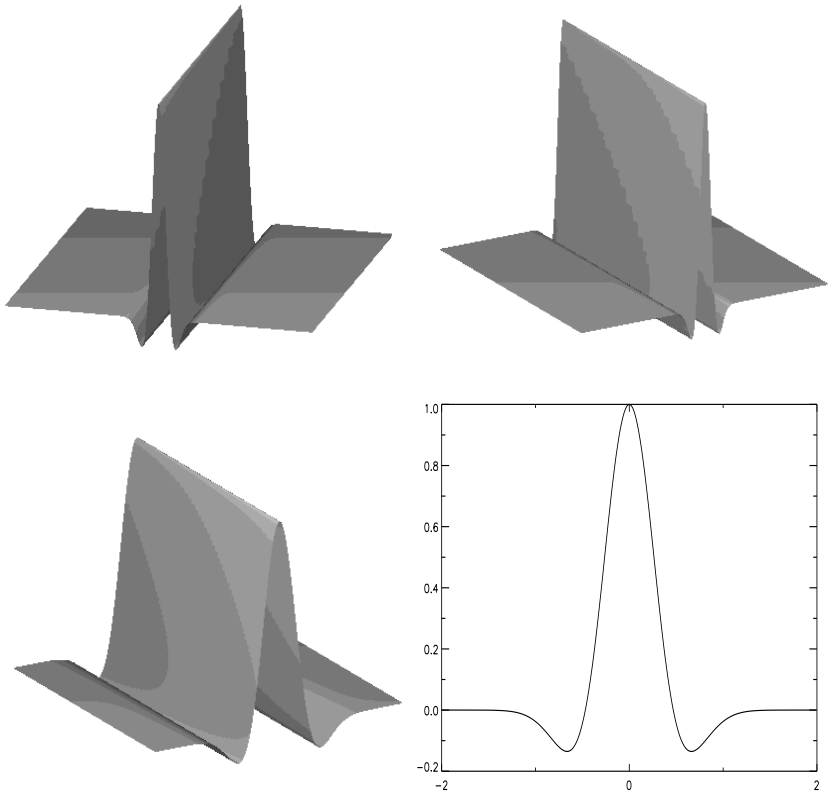


Figure 4.: *Example of ridgelet function.*

valid a.e. for functions which are both integrable and square integrable.

It has been shown (Candès and Donoho, 1999a) that the ridgelet transform is precisely the application of a 1-dimensional wavelet transform to the slices of the Radon transform where the angular variable θ is constant and t is varying. More details on the implementation of the digital ridgelet transform can be found in (Starck et al., 2001a).

Local Ridgelet Transform

The ridgelet transform is optimal to find only lines of the size of the image. To detect line segments, a partitioning must be introduced (Candès, 1998). The image is decomposed into smoothly overlapping blocks of sidelength b pixels in such a way that the overlap between two vertically adjacent blocks is a rectangular array of size b by $b/2$; we use overlap to avoid blocking artifacts. For an n by n image, we count $2n/b$ such blocks in each direction.

The partitioning introduces redundancy, as a pixel belongs to 4 neighboring blocks. We present two competing strategies to perform the analysis and synthesis :

1. The block values are weighted (analysis) in such a way that the co-addition of all blocks reproduce exactly the original pixel value (synthesis).
2. The block values are those of the image pixel values (analysis) but are weighted when the image is reconstructed (synthesis).

Experiments have shown that the second approach leads to better results. We calculate a pixel value, $f(i,j)$ from its four corresponding block values of half-size $\ell = b/2$, namely, $B_1(i_1,j_1)$, $B_2(i_2,j_1)$, $B_3(i_1,j_2)$ and $B_4(i_2,j_2)$ with $i_1, j_1 > b/2$ and $i_2 = i_1 - \ell, j_2 = j_1 - \ell$, in the following way :

$$\begin{aligned} f_1 &= w(i_2/\ell)B_1(i_1,j_1) + w(1-i_2/\ell)B_2(i_2,j_1) \\ f_2 &= w(i_2/\ell)B_3(i_1,j_2) + w(1-i_2/\ell)B_4(i_2,j_2) \\ f(i,j) &= w(j_2/\ell)f_1 + w(1-j_2/\ell)f_2 \end{aligned} \tag{23}$$

with $w(x) = \cos^2(\pi x/2)$. Of course, one might select any other smooth, nonincreasing function satisfying, $w(0) = 1$, $w(1) = 0$, $w'(0) = 0$ and obeying the symmetry property $w(x) + w(1-x) = 1$.

It is worth mentioning that the spatial partitioning introduces a redundancy factor equal to 4.

Examples

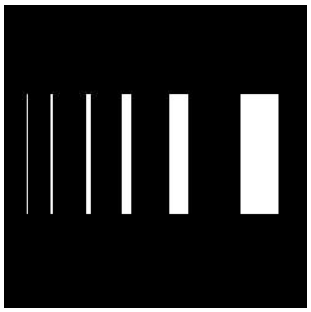


Figure 5.: *Simulated image*

Figure 5 shows a simulated image. It contains six line segments, with a height of 100 pixels, and a width respectively equals to 1,2,4,8,16,32. Pixel values on the lines are 1, and 0 outside. Four noisy images are shown in Figure 6 top. The noise standard deviation is respectively equals to 1,2,5 and 10. The noisy images have been filtered using the ridgelet

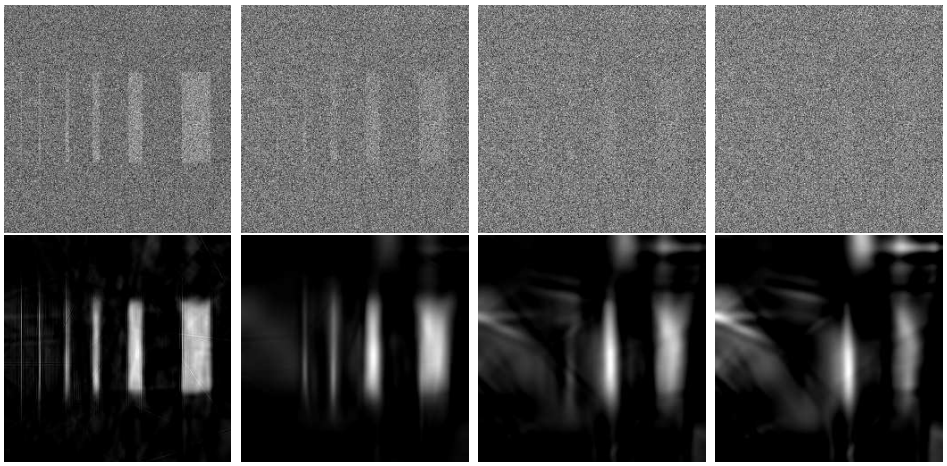


Figure 6.: *Top, simulated image, noisy image (sigma=1), and ridgelet filtered image. Middle, noisy images with a noise standard deviation respectively equals to 2,5 and 10. Bottom, ridgelet filtered of the previous images.*

transform, and are displayed in Figure 6 bottom. This experiment shows the ability of the ridgelet transform to detect feature at a very low signal to noise ratio (the SNR=0.1 in the case of the noisy image with the noise standard deviation equals to 10).

The Curvelet Transform

The curvelet transform, proposed by Donoho (Donoho and Duncan, 2000; Candès and Donoho, 1999b), open us the possibility to analyze an image with different block sizes, but with a single transform. The idea is to first decompose the image into a set of wavelet bands, and to analyze each band by a local ridgelet transform. The block size can be changed at each scale level. Roughly speaking, different levels of the multiscale ridgelet pyramid are used to represent different subbands of a filter bank output. The "à trous" subband filtering algorithm is especially well-adapted to the needs of the digital curvelet transform. The algorithm outputs $J+1$ subband arrays of size $n \times n$. The sidelength of the localizing windows is doubled *at every other* dyadic subband, hence maintaining the fundamental property of the curvelet transform which says that elements of length about $2^{-j/2}$ serve for the analysis and synthesis of the j-th subband $[2^j, 2^{j+1}]$. Note also that the coarse description of the image c_J is not processed.

This implementation of the curvelet transform is also redundant. The redundancy factor is equal to $16J + 1$ whenever J scales are employed.

Finally, the method enjoys exact reconstruction and stability, because these invertibility holds for each element of the processing chain. More detail can be found in (Starck et al., 2001a).

7. Combined Transform

If 2D or 3D astronomical data set may contain anisotropic features, they certainly will also contains isotropic ones. Hence, a perfect multiscale decomposition should benefit of both the à trous algorithm advantages and that of the ridgelet transform as well. More generally, we can imagine that we have $\mathcal{T}_1, ..., \mathcal{T}_{N_t}$ transform operators, each one being optimal to detect one kind of structure. A solution α is obtained by minimizing a functional of the form :

$$J(\alpha) = \| s - \sum_{k=0}^{N_t} \mathcal{T}_k^{-1} \alpha_k \|_2^2 + \lambda \sum_k \| \alpha_k \|_0 \tag{24}$$

where s is the original signal, and α_k are the coefficient obtained with the transform $\mathcal{T}_k$.

An algorithm to perform such a minimization has been presented in (Starck, 2001). It consists in hard thresholding the residual successively on the different bases.

1. Initialize $L_{\max}$ and the number of iterations Ni. For noise filtering, estimate the noise standard deviation σ, and set $L_{\min} = k$. Otherwise, set $\sigma = 1$ and $L_{\min} = 0$.
2. Set $\delta_\lambda = \frac{L_{\max} - L_{\min}}{Ni}$, $\lambda = L_{\max}$, and all coefficients α_k to 0.
3. While $\lambda >= L_{\min}$ do
4. for k = 1, .., N_t do
   - Calculate the residual $R = s - \sum_k \mathcal{T}_k^{-1} \alpha_k$.
   - Calculate the transform $\mathcal{T}_k$ of the residual : $r_k = \mathcal{T}_k R$.
   - For all coefficients $r_{k,i}$ do
     - Update the coefficients : if $\alpha_{k,i} \neq 0$ or $| r_{k,i} | > \lambda\sigma$ then $\alpha_{k,i} = \alpha_{k,i} + r_{k,i}$.
5. $\lambda = \lambda - \delta_\lambda$, and goto 6.

For an exact representation of the data, k must be set to 0. Choosing $k > 0$ introduces a filtering. If a single transform is used, it corresponds to the standard $k\sigma$ hard thresholding.

7.1 Example 1 : Simulation

Figure 7 illustrates the result in the case where the input image contains only lines and Gaussians. In this experiment, we have initialized $L_{\max}$ to 20, and δ to 2 (10 iterations). Two transform operators were used,

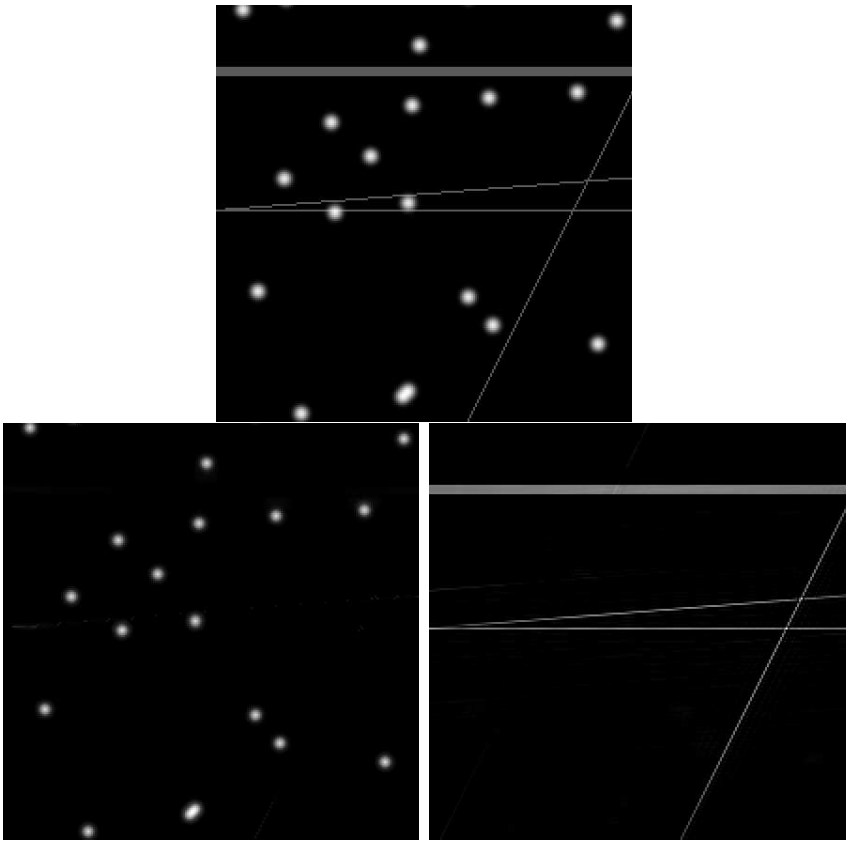


Figure 7.: *Top, original image containing lines and Gaussians. Botton left, reconstructed image for the à trous wavelet coefficient, bottom right, reconstructed image from the ridgelet coefficients.*

the à trous wavelet transform and the ridgelet transform. The first is well adapted to the detection of Gaussian due to the isotropy of the wavelet function (Starck et al., 1998a), while the second is optimal to represent lines (Candès and Donoho, 1999a). Figure 7 top, bottom left, and bottom right represents respectively the original image, the reconstructed image from the à trous wavelet coefficient, and the reconstructed image from the ridgelet coefficient. The addition of both reconstructed images reproduces the original one.

7.2 Example 2 : Elongated - point like object

Figure 8 shows the result of a decomposition of a spiral galaxy (NGC2997). This image (figure 8 top left) contains many compact structures (stars and HII region), more or less isotropic, and large scale elonga-

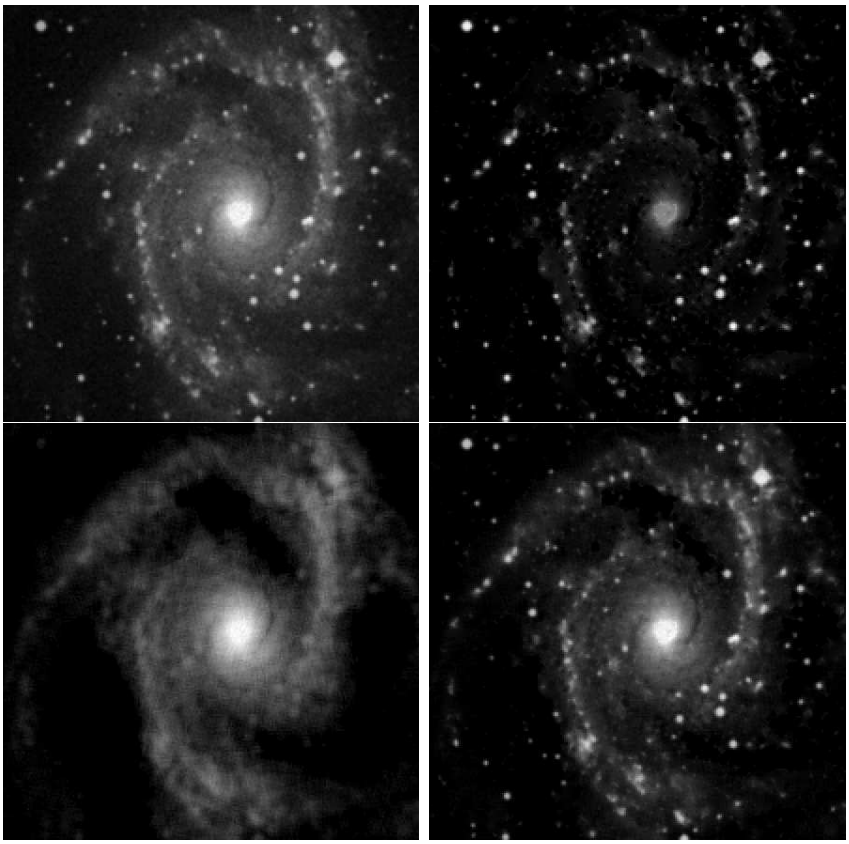


Figure 8.: *Top left, galaxy NGC2997, top right reconstructed image from the à trous wavelet coefficients, bottom left, reconstruction from the ridgelet coefficients, and bottom right addition of both reconstructed images.*

ted features (NGC2997 spiral part). Compact objects are well represented by isotropic wavelets, and the elongated features are better represented by a ridgelet basis. In order to benefit of the optimal data representation of both transforms, the image has been decomposed on both the à trous wavelet transform and on the ridgelet transform by using the combined transform method. When the functional is minimized, we get two images, and their coaddition is the filtered version of the original image. The reconstructions from the à trous coefficient, and from the ridgelet the ridgelet coefficient can be seen in figure 8 top right and bottom left. The addition of both images is presented in figure 8 bottom right.

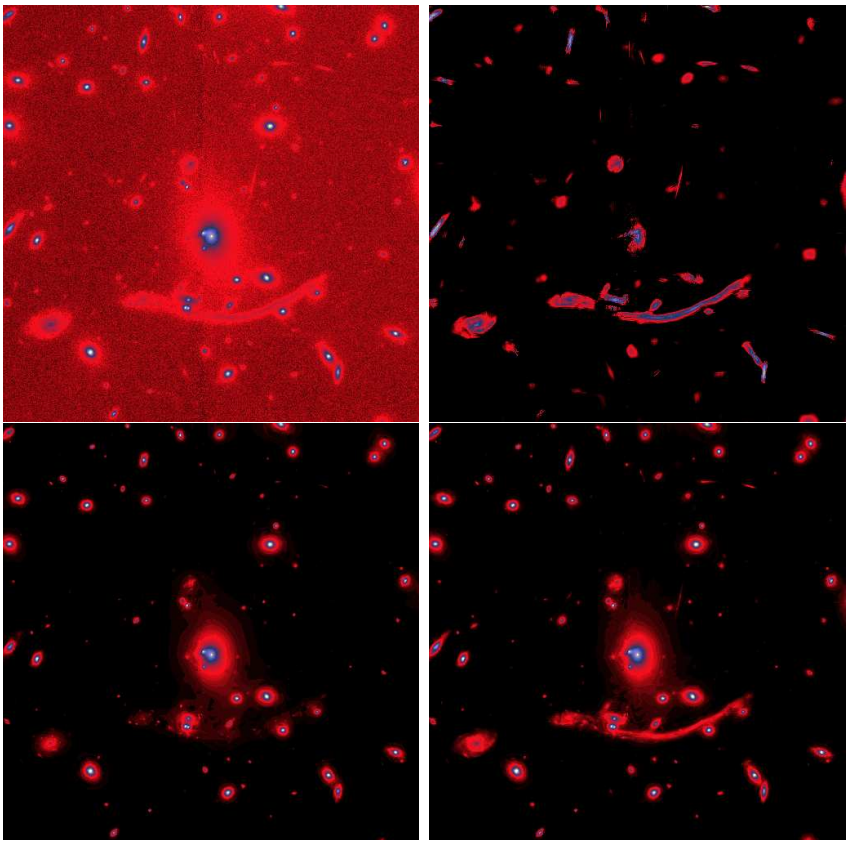


Figure 9.: *Top left, HST image of A370, top right coadded image from the reconstructions from the ridgelet and the curvelet coefficients, bottom left reconstruction from the à trous wavelet coefficients, and bottom right addition of the three reconstructed images.*

7.3 Example 3 : A370

Figure 9 upper left shows the HST A370 image. It contains many anisotropic features such the gravitational arc, and the arclets. The image has been decomposed using three transforms : the ridgelet transform, the curvelet transform, and the à trous wavelet transform. Three images have then been reconstructed from the coefficients of the three basis. Figure 9 upper right shows the coaddition of the ridgelet and curvelet reconstructed images. The à trous reconstructed image is displayed in Figure 9 lower left, and the coaddition of the three images can be seen in Figure 9 lower right. The gravitational arc and the arclets are all represented in the ridgelet and the curvelet basis, while all isotropic features are better represented in the wavelet basis.

References

Aghanim, N., Forni, O., and Bouchet, F. R. : 2001, *Astronomy and Astrophysics* **365**, 341

Andersson, W. D. L. and Andersson, B. G. : 1993, *American Astronomical Society Meeting* **183**, 1604+

Anscombe, F. : 1948, *Biometrika* **15**, 246

Arnaud, M., Maurogordato, S., Slezak, E., and Rho, J. : 2000, A&A **355**, 461

Aussel, H., Elbaz, D., Cesarsky, C., and Starck, J. : 1999, *Astronomy and Astrophysics* **342**, 313

Bendjoya, P. : 1993, *Astronomy and Astrophysics, Suppl. Ser.* **102**, 25+

Bendjoya, P., Petit, J. M., and Spahn, F. : 1993, *Icarus* **105**, 385+

Bijaoui, A. and Rué, F. : 1995, *Signal Processing* **46**, 229

Blanco, S., Bocchialini, K., Costa, A., Domenech, G., Rovira, M., and Vial, J. C. : 1999, *Solar Physical Journal* **186**, 281

Bonnarel, F., Fernique, P., Genova, F., Bartlett, J. G., Bienaymé, O., Egret, D., Florsch, J., Ziaeepour, H., and Louys, M. : 1999, *Astronomical Data Analysis Software and Systems VIII, ASP Conference Series, Vol. 172. Ed. David M. Mehringer, Raymond L. Plante, and Douglas A. Roberts. ISBN : 1-886733-94-5 (1999), p. 229.* **8**, 229

Bontekoe, T., Koper, E., and Kester, D. : 1994, *Astronomy and Astrophysics.* **294**, 1037

Brosch, N. and Hoffman, Y. : 1999, *MNRAS* **305**, 241

Burg, J. : 1978, Annual Meeting International Society Exploratory Geophysics, Reprinted in Modern Spectral Analysis, D.G. Childers, ed., IEEE Press, New York, 34–41

Cambrésy, L. : 1999, *Astronomy and Astrophysics* **345**, 965

Candès, E. : 1998, *Ph.D. thesis*, Stanford University

Candès, E. and Donoho, D. : 1999a, *Phil. trans ; R. Soc. Lond. A* **357**, 2495

Candès, E. J. and Donoho, D. L. : 1999b, in A. Cohen, C. Rabut, and L. Schumaker (eds.), *Curve and Surface Fitting : Saint-Malo 1999*, Vanderbilt University Press, Nashville, TN

Coifman, R. and Donoho, D. : 1995, in A. Antoniadis and G. Oppenheim (eds.), *Wavelets and Statistics*, pp 125–150, Springer-Verlag, New York

Damiani, F., Maggio, A., Micela, G., and Sciortino, S. : 1997, *Astrophysical Journal* **483**, 350+

Damiani, F., Sciortino, S., and Micela, G. : 1998, *Astronomische Nachrichten* **319**, 78+

Daubechies, I. : 1992, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia

Dixon, D. D., Hartmann, D. H., Kolaczyk, E. D., Samimi, J., Diehl, R., Kanbach, G., Mayer-Hasselwander, H., and Strong, A. W. : 1998, *New Astronomy* **3**, 539

Donoho, D. and Duncan, M. : 2000, in H. Szu, M. Vetterli, W. Campbell, and J. Buss (eds.), *Proc. Aerosense 2000, Wavelet Applications VII*, Vol. 4056, pp 12–29, SPIE, Bellingham Washington

Donoho, D. and Johnstone, I. : 1993, *Ideal spatial adaptation by wavelet shrinkage*, Technical Report 400, Stanford University

Escalera, E. and Mazure, A. : 1992, *Astrophysical Journal* **388**, 23

Escalera, E., Slezak, E., and Mazure, A. : 1992, *Astronomy and Astrophysics* **264**, 379

Feauveau, J. : 1990, *Ph.D. thesis*, Université Paris Sud

Fligge, M., Solanki, S. K., and Beer, J. : 1999, *Astronomy and Astrophysics* **346**, 313

Freeman, P. E., Kashyap, V., Rosner, R., Nichol, R., Holden, B., and Lamb, D. Q. : 1996, *ASP Conf. Ser. 101 : Astronomical Data Analysis Software and Systems V* **5**, 163+

Frieden, B. : 1978, *Image Enhancement and Restoration*, Springer-Verlag, Berlin

Gorski, K. M. : 1998, in *Evolution of Large-Scale Structure : From Recombination to Garching*, pp E49–+

Grebenev, S. A., Forman, W., Jones, C., and Murray, S. : 1995, *Astrophysical Journal* **445**, 607

Greene, J. E., Norris, J. P., and Bonnell, J. T. : 1997, *American Astronomical Society Meeting* **191**, 4805+

Gull, S. and Skilling, J. : 1991, *MEMSYS5 Quantified Maximum Entropy User's Manual*

Hecquet, J., Augarde, R., Coupinot, G., and Auriere, M. : 1995, *Astronomy and Astrophysics* **298**, 726+

Högbom, J. : 1974, *Astronomy and Astrophysics Supplement Series* **15**, 417

Huang, L. and Bijaoui, A. : 1991, *Experimental Astronomy* **1**, 311

Irbah, A., Bouzaria, M., Lakhal, L., Moussaoui, R., Borgnino, J., Laclare, F., and Delmas, C. : 1999, *Solar Physical Journal* **185**, 255

Ireland, J., Walsh, R. W., Harrison, R. A., and Priest, E. R. : 1999, *Astronomy and Astrophysics* **347**, 355

Jaffe, T. R., Bhattacharya, D., Dixon, D. D., and Zych, A. D. : 1997, *Astrophysical Journal Letter* **484**, L129

Jammal, G. and Bijaoui, A. : 1999, in *SPIE's 44th Annual Meeting, Wavelet Applications in Signal and Image Processing VII*, Vol. 3813 842-9, p. 842, SPIE-Int. Soc. Opt. Eng

Kalinkov, M., Valtchanov, I., and Kuneva, I. : 1998, *Astronomy and Astrophysics* **331**, 838

Kolaczyk, E. and Dixon, D. : 2000, *Astrophysical Journal* **534(1)**, 490

Kolaczyk, E. D. : 1997, *Astrophysical Journal* **483**, 340+

Kriessler, J. R., Han, E. H., Odewahn, S. C., and Beers, T. C. : 1998, *American Astronomical Society Meeting* **193**, 3820+

Krywult, J., MacGillivray, H. T., and Flin, P. : 1999, *Astronomy and Astrophysics* **351**, 883

Lawrence, J. K., Cadavid, A. C., and Ruzmaikin, A. A. : 1999, *Astrophysical Journal* **513**, 506

Lega, E., Bijaoui, A., Alimi, J. M., and Scholl, H. : 1996, *Astronomy and Astrophysics* **309**, 23

Louys, M., Starck, J.-L., Mei, S., Bonnarel, F., and Murtagh, F. : 1999, *Astronomy and Astrophysics, Suppl. Ser.* **136**, 579

Mallat, S. : 1989, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11(7)**, 674

Mallat, S. : 1998, *A Wavelet Tour of Signal Processing*, Academic Press

Martinez, V. J., Paredes, S., and Saar, E. : 1993, *MNRAS* **260**, 365

Michtchenko, T. A. and Nesvorny, D. : 1996, *Astronomy and Astrophysics* **313**, 674

Murtagh, F., Starck, J., and Bijaoui, A. : 1995, *Astronomy and Astrophysics, Suppl. Ser.* **112**, 179

Murtagh, F., Starck, J., and Louys, M. : 1998, *International Journal of Imaging Systems and Technology* **9**, 38

Narayan, R. and Nityananda, R. : 1986, *Ann. Rev. Astron. Astrophys.* **24**, 127

Norris, J. P., Nemiroff, R. J., Scargle, J. D., Kouveliotou, C., Fishman, G. J., Meegan, C. A., Paciesas, W. S., and Bonnel, J. T. : 1994, *Astrophysical Journal* **424**, 540

Pando, J. and Fang, L. : 1998, *Astronomy and Astrophysics* **340**, 335

Pando, J., Lipa, P., Greiner, M., and Fang, L. : 1998a, *Astrophysical Journal* **496**, 9+

Pando, J., Valls-Gabaud, D., and Fang, L. : 1998b, *American Astronomical Society Meeting* **193**, 3901+

Pantin, E. and Starck, J. : 1996, *Astronomy and Astrophysics, Suppl. Ser.* **315**, 575

Petit, J. M. and Bendjoya, P. : 1996, in *ASP Conf. Ser. 107 : Completing the Inventory of the Solar System*, pp 137–146

Pierre, M. and Starck, J. : 1998, *Astronomy and Astrophysics* **128(3)**, 801

Popa, L. : 1998, *New Astronomy* **3**, 563

Rauzy, S., Lachieze-Rey, M., and Henriksen, R. N. : 1993, *Astronomy and Astrophysics* **273**, 357+

Scargle, J. D. : 1997, in *Astronomical Time Series*, pp 1+

Schild, R. : 1999, *Astrophysical Journal* **514**, 598

Soon, W., Frick, P., and Baliunas, S. : 1999, *Astrophysical Journal Letter* **510**, L135

Starck, J., Aussel, H., Elbaz, D., Fadda, D., and Cesarsky, C. : 1999, *Astronomy and Astrophysics, Suppl. Ser.* **138**, 365

Starck, J., Bijaoui, A., Lopez, B., and Perrier, C. : 1994, *Astronomy and Astrophysics* **283**, 349

Starck, J., Bijaoui, A., Vatchanov, I., and Murtagh, F. : 2000, *Astronomy and Astrophysics, Suppl. Ser.* **147**, 139

Starck, J., Candès, E., and Donoho, D. : 2001a, *IEEE Transactions on Image Processing*, submitted

Starck, J. and Murtagh, F. : 1998, *Publications of the Astronomical Society of the Pacific* **110(744)**, 193

Starck, J. and Murtagh, F. : 1999, *Signal Processing* **76(2)**, 147

Starck, J., Murtagh, F., and Bijaoui, A. : 1998a, *Image Processing and Data Analysis : The Multiscale Approach*, Cambridge University Press, Cambridge (GB)

Starck, J., Murtagh, F., and Gastaud, R. : 1998b, S*pecial Issue on Multirate Systems, Filter Banks, Wavelets, and Applications of IEEE Transactions on CAS II* 45(8)

Starck, J., Murtagh, F., Pirenne, B., and Albrecht, M. : 1996, *Publications of the Astronomical Society of the Pacific* **108**, 446

Starck, J., Murtagh, F., Querre, P., and Bonnarel, F. : 2001b, *Astronomy and Astrophysics* **368**, 730

Starck, J. and Querre, P. : 2001, *Signal Processing*, in press

Starck, J.-L. : 2001, in T. Barth, T. Chan, and R. Haimes (eds.), *Advanced Multiscale and Multiresolution Methods*, Springer-Verlag

Steiman-Cameron, T. Y., Scargle, J. D., Imamura, J. N., and Middleditch, J. : 1997, *Astrophysical Journal* **487**, 396+

Szatmary, K., Gal, J., and Kiss, L. L. : 1996, *Astronomy and Astrophysics* **308**, 791

Walker, K. C., Schaefer, B. E., and Fenimore, E. E. : 2000, *Astrophysical Journal* **537**, 264

Weir, N. : 1991, in *3rd ESO/ST-ECF Data Analysis Workshop*

Weir, N. : 1992, in D. Worral, C. Biemesderfer, and J. Barnes (eds.), *Astronomical Data Analysis Software and System 1*, pp 186–190, Astronomical Society of the Pacific

Weistrop, D., Harvey, V., Pitanzo, D., Cruzen, S., Rogers, P., Beaver, M., and Pierce, D. : 1996, *American Astronomical Society Meeting* **189**, 12216+