Astronomical Data Analysis Software and Systems IV
ASP Conference Series, Vol. 77, 1995
R. A. Shaw, H. E. Payne, and J. J. E. Hayes, eds.

Multiresolution and Astronomical Image Processing

J.-L. Starck

CEA, DSM/DAPNIA, CE-Saclay, F-91191 Gif-sur-Yvette Cedex, France

F. Murtagh[1]

Space Telescope–European Coordinating Facility, European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748 Garching, Germany

A. Bijaoui

Observatoire de la Côte d'Azur, B.P. 229, F-06304 Nice Cedex 4, France

Abstract. We present several wavelet transform algorithms and their applications in astronomical image processing (restoration, object detection, compression, etc.).

1. The Discrete Wavelet Transform

1.1. Mallat's Transform

Extensive literature exists on the wavelet transform and its application (Chui 1992; Daubechies 1992; Meyer 1989). A discrete wavelet transform approach can be obtained from multiresolution analysis (Mallat 1989). Multiresolution analysis results from the embedded subsets generated by interpolations at different scales. A function $f(x)$ is projected at each step j onto the subset V_j. This projection is defined by the scalar product $c_j(k)$ of $f(x)$ with the scaling function $\phi(x)$ which is dilated and translated:

$$c_j(k) = < f(x), 2^{-j}\phi(2^{-j}x - k) > . \tag{1}$$

$\phi(x)$ is a scaling function which has the property

$$\frac{1}{2}\phi\left(\frac{x}{2}\right) = \sum_n h(n)\phi(x-n). \tag{2}$$

Equation 2 permits the set $c_{j+1}(k)$ to be computed directly from $c_j(k)$. If we start from the set $c_0(k)$, we compute all the sets $c_j(k)$, with $j > 0$, without directly computing any other scalar product:

$$c_{j+1}(k) = \sum_n h(n-2k)c_j(n). \tag{3}$$

[1]Affiliated to Astrophysics Division, Space Science Department, European Space Agency

At each step, the number of scalar products is divided by 2; the signal is smoothed and information is lost. The remaining information can be restored using the complementary subspace W_{j+1} of V_{j+1} in V_j. This subspace can be generated by a suitable wavelet function $\psi(x)$ with translation and dilation:

$$\frac{1}{2}\psi\left(\frac{x}{2}\right) = \sum_n g(n)\phi(x-n). \tag{4}$$

We compute the scalar products $< f(x), 2^{-(j+1)}\psi(2^{-(j+1)}x - k) >$ with

$$w_{j+1}(k) = \sum_n g(n-2k)c_j(n). \tag{5}$$

With this analysis, we have built the first part of a filter bank. In order to restore the original data, Mallat uses the properties of orthogonal wavelets, but the theory has been generalized to a large class of filters (Cohen, Daubechies, & Feauveau 1992). Filters $\tilde{h}$ and $\tilde{g}$, the conjugates of h and g, have been introduced (Daubechies 1992), and the restoration is performed with

$$c_j(k) = 2\sum_l [c_{j+1}(l)\tilde{h}(k+2l) + w_{j+1}(l)\tilde{g}(k+2l)]. \tag{6}$$

In order to get an exact restoration, two conditions are required for the conjugate filters: (1) the De-aliasing condition:

$$\hat{h}(\nu+\frac{1}{2})\hat{\tilde{h}}(\nu) + \hat{g}(\nu+\frac{1}{2})\hat{\tilde{g}}(\nu) = 0, \tag{7}$$

and (2) exact restoration:

$$\hat{h}(\nu)\hat{\tilde{h}}(\nu) + \hat{g}(\nu)\hat{\tilde{g}}(\nu) = 1. \tag{8}$$

Many sets of filters have been proposed, especially for coding. It has been shown (Daubechies 1992) that the choice of these filters must be guided by the regularity of the scaling and the wavelet functions. The complexity is proportional to N. The algorithm provides a pyramid of N elements. The 2D algorithm is based on separable variables leading to x and y directions being prioritized. This implies a non-isotropic analysis, which often runs counter to physical patterns.

1.2. Feauveau's Transform

Feauveau (1990) introduced the quiconx analysis according to Adelson (Adelson, Simoncelli, & Hingorani 1987). This analysis is not dyadic and allows an image decomposition with a resolution factor equal to $\sqrt{2}$. By this method, we have one wavelet image at each scale, and not three as with the previous method.

1.3. Wavelet Transform with the FFT

If $\phi(\nu)$ has a cut-off frequency, the discrete filter $h(n)$ used to compute the coefficients at a given resolution from the previous one (Starck & Bijaoui 1994), is $\hat{h}(\nu) = \hat{\phi}(2\nu)/\hat{\phi}(\nu)$. Therefore, between two resolutions, the cut-off frequency is divided by 2, allowing us to reduce the number of samples. The information lost in the filtering is given by the discrete filter $g(n)$ defined by $\hat{g}(\nu) = \hat{\psi}(2\nu)/\hat{\phi}(\nu)$. The signal is reconstructed with the two filters $\tilde{h}$, $\tilde{g}$:

$$\hat{\tilde{h}}(\nu) = \frac{\hat{h}^*(\nu)}{\mid \hat{h}(\nu) \mid^2 + \mid \hat{g}(\nu) \mid^2} \text{ and } \hat{\tilde{g}}(\nu) = \frac{\hat{g}^*(\nu)}{\mid \hat{h}(\nu) \mid^2 + \mid \hat{g}(\nu) \mid^2}. \tag{9}$$

The described algorithm is easily adapted to image processing. As opposed to classical multiresolution analysis, it is a decomposition with isotropic scaling and wavelet functions. The frequency band is also reduced by a factor 2 at each step. Applying the sampling theorem, we can build a pyramid of $N + N/2 + \ldots + 1 = 2N$ elements. For image analysis the number of elements is $4N^2/3$. The increase in storage is small.

1.4. À Trous Algorithm

The "*à trous*" algorithm is a powerful algorithm for the following reasons: (1) the computational requirement is reasonable, (2) the algorithm is easy to program, (3) in two dimensions, the transform is practically isotropic, (4) we can use compact scaling functions, (5) the reconstruction algorithm is trivial, (6) the transform is known at each pixel, allowing detection without any error, and without interpolation, (7) we can follow the evolution of the transform from one scale to the next, and (8) invariance under translation is completely verified.

Details of the algorithm are given in Bijaoui, Starck, & Murtagh (1994) and Shensa (1992). The wavelet transform of an image by this algorithm produces, at each scale j, a set $\{w_j\}$ which we will call a wavelet plane throughout the following discussion. This has the same number of pixels as the image. The original image c_0 can be expressed as the sum over wavelet planes and the smoothed array c_p

$$c_0(x, y) = c_p(x, y) + \sum_{j=1}^{p} w_j(x, y) \tag{10}$$

1.5. Pyramidal Algorithm

The pyramidal algorithm (Starck 1993) is derived from the "*à trous*" algorithm. At each scale, w_{j+1} is obtained by computing the difference between c_j and $F.c_j$ (where F denotes a linear filtering). But c_{j+1} is equal to $U.F.c_j$ (U being the undersampling or decimating operator), and not equal to $F.c_j$ (we first undersample $F.c_j$ by keeping one pixel out of two). This transform produces data similar to those obtained by the Burt & Adelson (1983) pyramidal Laplacian algorithm, but it is done with one wavelet. Burt's algorithm needs three wavelets and is therefore not isotropic (Bijaoui 1991). The number of elements is $4N^2/3$, as in the case of the transform using the FFT, but this method has the advantage to allow all computation to be carried out in direct space. An iterative algorithm is necessary for an exact restoration.

1.6. Problems Related to the Wavelet Transform

Anisotropic Wavelet. The 2-dimensional extension of Mallat's algorithm leads to a wavelet transform with three wavelet functions (three wavelet coefficient sub-images at each scale). An isotropic wavelet seems more appropriate, specially in astronomical imaging where objects are often isotropic (e.g., stars).

Invariance by Translation. Mallat's and Feauveau's methods provide a remarkable framework to code a signal, and especially an image, with a pyramidal set of values. But, contrary to the continuous wavelet transform, these analyses are not covariant under translation. At a given scale, we derive a decimated number of wavelet coefficients. We cannot restore the intermediate values without using the approximation at this scale and the wavelet coefficients at smaller scales. Since the multiresolution analysis is based on scaling functions without cut-off frequency, the application of the Shannon interpolation theorem is not possible. The interpolation of the wavelet coefficients can only be done after reconstruction and shift. This is unimportant for a signal coding that does not modify the data, but is not the same if we want to analyze or restore an image.

Scale Separation. If the image I we want to analyze is the convolution product of an object O by a point spread function (PSF) ($I = P * O$), we have

$$\hat{W}^{(I)}(a, u, v) = \sqrt{a}\hat{\psi}^*(au, av)\hat{I}(u, v), \tag{11}$$

where $W^{(z)}$ are the wavelet coefficients of z, and a is the scale parameter. We deduce that

$$\begin{aligned} \hat{W}^{(I)}(a, u, v) &= \sqrt{a}\hat{\psi}^*(au, av)\hat{P}(u, v)\hat{O}(u, v) & (12) \\ &= \hat{O}(u, v)\hat{W}^{(P)}(a, u, v). & (13) \end{aligned}$$

We can directly analyze the object from the wavelet coefficients of the image. But due to decimation effects in Mallat's, Feauveau's, and pyramidal methods, this equation fails. The wavelet transform using the FFT (which decimates using Shannon's theorem) and the "*à trous*" algorithm (which does not decimate) are the only ones which respect the scale separation property.

Negative Values. By definition, the wavelet coefficient mean is null. Every time we have a positive structure at a scale, we have negative values surrounding it. These negative values often create artifacts during the restoration process, or complicate the analysis. For instance, if we threshold small values (noise, insignificant structures, etc.) in the wavelet transform, and if we reconstruct the image at the full resolution, the structure's flux will be modified. Furthermore, if an object is very high, the negative values will be important, too, and will lead to false structure detections.

Point Objects. We often have bright point objects in astronomical imaging (stars, cosmic ray hits, etc.) The convolution of a Dirac function by the wavelet function is equal to the wavelet function, so at each scale, and at each point source, we will have the wavelet. Cosmic rays can pollute all the scales of the wavelet transform.

1.7. Conclusion

There is no ideal wavelet transform algorithm—the selection will depend on the application. The two last problems cannot be solved by the wavelet transform, and they lead to other multiresolution tools which we now present.

2. Other Multiresolution Approaches

The search for new multiresolution tools was motivated by problems related to the wavelet transform. We would prefer that a point structure (represented in one pixel in the image) is present only at the first scale, and that a positive structure in the image not create negative values in the multiresolution space. We will see how such an algorithm can be found, using morphological filters such as the median filter.

2.1. Multiresolution from the Median

By modifying the "à trous" algorithm, we easily obtain the desired transform. The algorithm becomes:

1. We define a mask M_t with a size t ($t = 2 * s + 1$) (for instance a square mask with a size 3×3, implying that $t = 3$).
2. We initialize j to 0, and we start from data c_0.
3. med being the filtering median function, we calculate $c_{j+1} = med(M_t, f_j)$ and median coefficients at scale j by: $w_{j+1} = c_j - c_{j+1}$
4. We double the mask size: $s = 2 * s$ and $j = j + 1$.
5. If j is less than the number of scales we want, return to 3.

The reconstruction is carried out by a simple addition of all the scales:

$$c_0 = c_p + \sum_j w_j \tag{14}$$

Such an algorithm has several advantages: (1) the transform can be carried out with integer values, (2) energy in structures at each scale is real, and is not modified by the treatment, (3) structure contours are better respected, (4) there are no negative values around positive structures, and (5) the algorithm can easily be modified to work at intermediate scales (in order to have a non-dyadic analysis). We just multiply the size s at the step 4 by a coefficient different from 2 (if we want half resolution, we multiply by 1.5).

However this algorithm has a big disadvantage: computation time. The mask increases by two at each scale, and we have to sort a number of values which increases considerably. To solve this problem, we introduce decimation.

2.2. Pyramidal Multi-Median Transform

We reduce the number of samples (pixels) by keeping one pixel out of every two (in each dimension) at each scale. The algorithm is:

1. We define a mask M_t with a size t ($t = 2 * s + 1$).
2. We initialize j to 0, and we start from data c_0.
3. Calculate $m_{j+1} = med(M_t, f_j)$ and median coefficients at scale j as $w_{j+1} = c_j - m_{j+1}$
4. We calculate c_{j+1} from m_{j+1} by decimating.
5. Set $j = j + 1$. If j is less than the number of scales we want, return to 3.

In this transform, the mask size is constant, and there is no time computation problem anymore. However, the reconstruction is not exact, and an iterative procedure has to be introduced during the transformation or at the reconstruction. If we choose the iterative transform, we have:

1. Initialize i to 0, R^i to the image, and multiresolution coefficients $w_j^{(I)}$ to 0.
2. Compute the pyramidal multi-median transform of $R^{(i)}$; we have $w_j^{(R)}$.
3. $w_j^{(I)} = w_j^{(I)} + w_j^{(R)}$.
4. Reconstruct $\tilde{I}$ from $w_j^{(I)}$. B-spline interpolation is used.
5. Compute $R^{i+1} = I - \tilde{I}$.
6. Set $i = i + 1$, and return to 3.

R^{i+1} in step 6 tends towards a null image (cf. Van Cittert 1931). We have found that 4 or 5 iterations are sufficient.

2.3. Conclusion

Such a transform does not replace the wavelet transform, but complements it. When images contain cosmic ray hits, for instance, the use of the pyramidal multi-median transform avoids cosmic ray hits affecting all scales. Other morphological tools can be used to perform a similar transform, such as opening (N erosions followed by N dilations). However, results with the median filter were superior.

3. Significant Coefficients

Images generally contain noise (Gaussian or Poisson, usually) and hence the wavelet coefficients are noisy, too. In most applications, it is necessary to know whether a coefficient is due to signal or to noise. The wavelet transform yields a set of resolution-related views of the input image. A wavelet image plane at level j has coefficients given by $w_j(x,y)$. If we obtain the distribution of the coefficient $w_j(x,y)$ for each plane, based on the noise, we can introduce a statistical significance test for this coefficient.

Given stationary Gaussian noise, it suffices to compare $w_j(x,y)$ to $k\sigma_j$, where σ_j is the standard deviation of wavelet plane j, and k is often chosen as 3. If $w_j(x,y)$ is small it is not significant, and could be due to noise. If $w_j(x,y)$ is large, it is significant. If the noise in the data I is Poisson, the transform $T(I(x,y)) = 2\sqrt{I(x,y)+3/8}$ acts as if the data arose from the Gaussian white noise model with unit standard deviation (Anscombe 1948). Taking the wavelet transform of $T(I)$, $w_j^{(I)}(x,y)$ will be significant if $w_j^{(T(I))}(x,y)$ is above a given threshold. (The superscript on the wavelet coefficients indicates the image to which the wavelet transform was applied.)

The appropriate value of σ_j in the succession of wavelet planes is assessed from the standard deviation of the noise σ_I in the original image and from study of the noise in the wavelet space. Details of this study can be found in Starck & Murtagh (1994).

4. Applications

4.1. Multiresolution Support

We will say that a multiresolution support (Starck, Bijaoui, & Murtagh 1994) of an image describes in a logical or boolean way whether an image I contains information at a given scale j and at a given position (x,y). If $M^{(I)}(j,x,y) = 1$ (or *true*), then I contains information at scale j and at the position (x,y). M depends on several parameters: (1) the input image, (2) the algorithm used for the multiresolution decomposition, (3) the noise, and (4) all additional constraints we want the support to satisfy.

Such a support results from the data, the treatment (noise estimation, etc.), and from knowledge on our part of the objects contained in the data (size of objects, linearity, etc.). In the most general case, *a priori* information is not available to us.

The multiresolution support of an image is computed in several steps:

1. Compute the wavelet (or other multiresolution) transform of the image.
2. Binarization of each scale leads to the multiresolution support. We have:

$$M(j,x,y) = \begin{cases} 1 & \text{if } w_j(x,y) \text{ is significant} \\ 0 & \text{if } w_j(x,y) \text{ is not significant} \end{cases} \tag{15}$$

3. *A priori* knowledge can be introduced by modifying the support.

The last step depends on the knowledge we have of our images. For instance, if we know there is no interesting object smaller or larger than a given size in our image, we can suppress, in the support, anything which is due to that kind of object. This can often be done conveniently by the use of mathematical morphology. Or we can add star positions from a catalog in order to improve a restoration. In the most general setting, we have no information to add to the multiresolution support.

The multiresolution support indicates where the information is, and allows all information we have about the image (data, noise, catalog information, etc.) to be integrated. It can then be used to visualize or to further process the data. The fact that we can add constraints and laws to the support makes this a very convenient data structure for introducing information into the treatment.

4.2. Visualization

The visualization of the wavelet transform allows us to have better knowledge of the structure of the image. The "*à trous*" algorithm is particularly efficient, because each scale has the same size as the original image, and can be visualized, processed, and compared very easily. See Starck (1993) and Bijaoui, Starck, & Murtagh (1994) which use various ways to visualize a wavelet transform.

4.3. Detection

In general in astronomy, it is not the image which is of interest but rather the objects in the image! The multiresolution support provides all of the information needed to demarcate and then label the objects. Varying background is automatically catered for by the multiresolution transform (equation 14 or its equivalent in the case of a pyramidal transform). This is a very real advantage, since alternative approaches must devote considerable processing attention to adaptive or other methods in order to allow for variable backgrounds or superimposed objects.

Multiresolution provides much information on the scale of the objects. Thus foreground objects (which may not be of interest, e.g., cosmic ray hits) as well as background objects (e.g., faint galaxies) can be found—assuming an appropriate wavelet (or transform kernel) and processing chain—at different resolution scales. It is not difficult to determine all objects, in this way, and then to retain only those which are of immediate interest. Examples of the use of the multiresolution approach to object detection can be found in Murtagh, Zeilinger, Starck, & Bijaoui (1994).

4.4. Quality Criteria

Objective comparison of two images is often necessary, for instance when we want to evaluate the restoration quality. Very few quantitative parameters can be extracted. The correlation between the original image and the restored one provides a classical criterion. Another way to compare two pictures is to determine the mean-square error. These criteria, however, are not sufficient. They give no information about the resulting resolution. A comprehensive criterion must take into account the resolution. We can compute for each dyadic scale the correlation coefficient and the quadratic error between the wavelet transforms of the original and the restored images. This allows us to compare, for each resolution, the restoration quality (Starck & Bijaoui 1994).

4.5. Filtering

An easy way to filter an image is to keep only the significant wavelet coefficients and to reconstruct an image (Starck & Bijaoui 1994; Donoho 1992). We have shown (Starck, Bijaoui, & Murtagh 1994) that, with few iterations and by using the multiresolution support, we can have excellent results.

4.6. Deconvolution

The object-image relation is often given by:

$$I = O * P + N \tag{16}$$

O is the observed object, I is the obtained image, P is the point spread function (PSF) of the imaging system, and N is an additive noise. We want to determine O knowing I and P, the main difficulties being the existence of: (1) a cut-off frequency of the PSF, and (2) the additive noise (e.g., Cornwell 1988). The multiresolution approach (Starck & Murtagh 1994) allows the noise to be controlled during the deconvolution process, and leads to a regularization of this ill-posed problem. The Poisson noise case and the use of a multiresolution support framework have been treated in (Murtagh, Starck, & Bijaoui 1994).

4.7. Interferometric Deconvolution

In interferometric imaging, measurements are carried out in the Fourier space but the (u, v) plane is not completely covered. The image ("dirty map"), is obtained by a simple inverse Fourier transform of the data, and the PSF ("dirty beam"), by an inverse Fourier transform of the (u, v) plane coverage. The presence of secondary lobes in the dirty beam creates very important 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. An iterative reconstruction algorithm allows solutions to be found which satisfy the positivity constraint, and the fidelity to measurements constraint (i.e., at each measured $V_m(u, v) \pm \Delta_m(u, v)$, 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 & Bijaoui (1994) and Starck, Bijaoui, Lopez, & Perrier (1994).

4.8. Compression

Several studies have used the orthogonal wavelet transform to compress astronomical data (Press 1991; White 1994). Using the multiresolution support should improve the quality of the decompressed image. Results presented in Starck, Murtagh, & Louys (1994) show how good results can be.

5. Conclusion

The wavelet transform, and more generally the multiresolution transform, provides a powerful and versatile framework for astronomical image processing. A data structure is created which allows such objectives as the following to be handled: visualization, object detection, filtering, deconvolution, and compression. Studies in all of these areas have been cited here, and in many cases the results obtained are considerably better than traditional approaches.

References

Adelson, E. H., Simoncelli, E., & Hingorani, R. 1987, SPIE Visual Communication and Image Processing II, 845, 50

Anscombe, F. J. 1948, Biometrika, 15, 246

Bijaoui, A. 1991, Ondelettes et Paquets d'Ondes, Inria, Rocquencourt

Bijaoui, A., Starck, J. L., & Murtagh, F. 1994, Traitement du Signal, 3, 11

Burt, P. J., & Adelson A. E. 1983, IEEE Trans. on Communications, 31, 532

Cornwell, T. J. 1988, Proc. NATO Advanced Study Institute on Diffraction-Limited Imaging with Very Large Telescopes, Cargèse, 273

Chui, C. H. 1992, Wavelet Analysis and its Application (New York, Academic Press)

Cohen, A., Daubechies, I., & Feauveau, J. C. 1992, Comm. Pur. Appl. Math., 45, 485

Daubechies, I. 1992, Ten Lectures on Wavelets, Philadelphia.

Donoho, D. L. 1992, Proc. Progress in Wavelet Analysis and Applications (Toulouse, Ed. Frontières)

Feauveau, J. C. 1990, Thesis, University Paris Sud.

Högbom, J. A. 1974, A&AS, 15, 417

Mallat, S. 1989, IEEE Trans. on Pattern Analysis and Machine Intelligence, 11, 574

Meyer, Y. 1989, Wavelets, ed. J. M. Combes et al., (Berlin, Springer Verlag), 21

Murtagh, F., Starck, J. L., & Bijaoui, A. 1994, A&A, submitted

Murtagh, F., Zeilinger, W., , Starck, J. L., & Bijaoui, A. 1994, this volume, p. **??**

Press, W. H. 1991, in Astronomical Data Analysis Software and Systems I, ASP Conf. Ser., Vol. 25, eds. D. M. Worrall, C. Biemesderfer, & J. Barnes (San Francisco, ASP), p. 25

Shensa, M. J. 1992, Proc. IEEE Transactions on Signal Processing, 40, 2464

Starck, J. L. 1993, The Wavelet Transform, MIDAS Manual (Garching, ESO)

Starck, J. L., & Bijaoui, A. 1994, Signal Processing, 35, 195

Starck, J. L., Bijaoui, A., Lopez, A., & Perrier, C. 1994, A&A, 283, 349

Starck, J. L., & Bijaoui, A. 1994, J. Opt. Soc. Am. A, 11, No. 4

Starck, J. L., & Murtagh, F. 1994, A&A, 288, 343

Starck, J. L., Bijaoui, A., & Murtagh, F., 1994, CVIP: Graphical Models and Image Processing, submitted

Starck, J. L., Murtagh, F., & Louys, M., 1994, this volume, p. **??**

Van Cittert, P. H. 1931, Physik, Z., 69, 298

White, R. L. 1993, in Space and Earth Science Data Compression Workshop, ed. James C. Tilton, NASA Conference Publication 3183, p. 117