A Framework for Filter Design Emphasizing Multiscale Feature Preservation

Gheorghe Craciun[1], David Thompson[2], Raghu Machiraju[3], and Ming Jiang[4]

[1,3,4]*Department of Computer and Information Science*
Ohio State University, Columbus OH 43210, USA
gcraciun@math.ohio-state.edu, raghu@cis.ohio-state.edu, jiang@cis.ohio-state.edu

[2]*Engineering Research Center for Computational Systems*
Mississippi State University, Mississippi State MS 39762, USA
dst@erc.msstate.edu

Abstract

In this paper we develop a filter design framework emphasizing feature preservation. We are particularly interested in multiscale filters that can be used in wavelet transforms for large datasets generated by computational fluid dynamics simulations. High-fidelity wavelet transforms can facilitate the accurate mining of scientific data. However, it is important that the salient characteristics of the original features be preserved under the transformation. Our effort is different from classical filter design approaches which focus solely on performance in the frequency domain. In particular, we present a set of filter design axioms that ensure certain feature characteristics are preserved and that the resulting filter corresponds to a wavelet transform admitting in-place implementation. Three standard filters, corresponding to the Haar, linear, and cubic lifting wavelets, are shown to violate at least one of the criteria related to feature preservation. We also demonstrate how the axioms can be used to design a simple feature-centric filter. Results are included that demonstrate the feature-preservation characteristics of each filter.

1. Introduction

Large-scale computational fluid dynamics simulations of physical phenomena produce data of unprecedented size (terabyte and petabyte range). Unfortunately, development of appropriate data management and visualization techniques has not kept pace with the growth in size and complexity of such datasets. One paradigm of large-scale visualization is to browse regions containing significant features of the dataset while accessing only the data needed to reconstruct these regions. Such an effort is akin to mining or discovering spatially and temporally correlated data. The cornerstone of an approach of this type is a representational scheme that facilitates ranked access to macroscopic features in the dataset [1, 2, 3]. In this approach, a feature-detection algorithm is used to identify and rank contextually significant features directly in the wavelet domain.

In [1, 2, 3], the linear lifting scheme [4] was used for compressing components of a vector field. The work reported here grew out of our efforts to analyze the implementation of the lifting scheme and design new transforms that more ardently preserve features in discrete flow fields. The rate-distortion characteristics of many wavelet transforms do not bode well for feature preservation [3]. However, it was unclear as to what distortions the wavelet transform wrought on the data. It is therefore useful to evaluate the effect of the wavelet transform in terms of processes that alter the "shape" of the data, i.e., features. Additionally, for very large datasets it is necessary that data-mining or feature detection be performed in the compressed domain. In this context, it is essential that the wavelet transform preserve significant features in the data set.

It is well known that wavelets can efficiently approximate smooth data [5] and produce efficient compression schemes. To suitably preserve edges in scalar image fields, several linear and non-linear or data-dependent schemes have been proposed [6, 7, 8, 9]. In particular, Zhou [10] utilizes Essentially Non-Oscillatory (ENO) reconstructions [11] of the data so that fewer high frequency coefficients are created.

Techniques employed in the study of partial differential equations (PDEs) have been extensively utilized to define the multiscale behavior of feature detection algorithms [12, 13, 14, 15] for images. Typically, the time variable in an evolutionary PDE is taken to represent a scale parameter. In vision and image processing applications, edges can be thought of as discontinuities. These techniques are used to enhance interregion boundaries and smooth intraregion variations. It should be noted that linear PDEs are not completely successful in enhancing boundaries while eliminating noise. Discrete models of the diffusion equation with a nonlinear conductance based on gradient information have proven to be particularly useful for these applications [12, 15].

In this paper we define a framework for the analysis and design of of multiscale filters through a variational characterization and a multiscale PDE formulation. Given the need for efficient compression and processing, we consider only linear transforms at this time. We suggest that the methods proposed here can be used in conjunction with frequency-based methods to design multiscale linear wavelet filters. A result of our three-fold characterization is a set of axioms that can be used to analyze and eventually design filters. Filters that satisfy our axioms will be more likely to preserve features in a linear wavelet space and enable high-fidelity mining of scientific datasets. Additionally, we seek to design filters corresponding to wavelets that can be implemented as a sequence of lifting steps [4].

Our axiomatic filter design resembles the work of Weickert et al. [15] as well as that of Alvarez et al. [16]. Although these efforts yield similar sets of axioms, their frameworks are different since the domain of interest is limited to images populated with strong discontinuities such as edges. In our application, however, not all regions of strong gradients correspond to discontinuities. In fact, features with strong gradients such as expansions and boundary layers should not be treated as discontinuities.

Our paper is structured as follows. We first describe the general linear filter. Next, we formalize our ideas regarding feature preservation. We then present a set of filter design axioms. Representative forms of the lifting scheme are analyzed using our framework, including a new scheme that satisfies our criteria. Suitable examples are provided to support our analysis and claims. We stress that our current efforts should be considered as a "work in progress."

2. General Linear Filter

We begin by defining a discrete, scalar quantity $s_{j,l}$ on an equally-spaced mesh $x_{j,l} = l\Delta x_j$ for $l = 0, ..., 2N$ with N being a positive integer. We seek a multiscale approximation to $s_{j,l}$ on a second equally-spaced mesh, $x_{j-1,l} = l\Delta x_{j-1}$ for $l = 0, ..., N$ with $\Delta x_{j-1} = 2\Delta x_j$, that preserves certain characteristics of the original scalar field. We denote this approximation as $s_{j-1,l}$.

We now consider a general linear filter of the form

$$s_{j-1,l} = \sum_{k=-m}^{+n} a_k s_{j,2l+k} \tag{1}$$

where m and n are positive integers and the a_k are constants that are independent of the data. The a_k are composite coefficients that represent the combined effects of a wavelet transform implemented as a filter. The discrete moments of the filter are given by

$$\alpha_q = \sum_{k=-m}^{+n} k^q a_k \quad . \tag{2}$$

After the filter is applied, the data is subsampled to define the space $x_{j-1,l}$.

It can be shown [17] that the application of the linear filter defined in (1) can be thought of as the evolution of the solution to the partial differential equation

$$\begin{aligned} \frac{\partial s}{\partial \tau} &= \alpha_1 \frac{\Delta x_j}{\Delta \tau} \frac{\partial s}{\partial x} + \frac{1}{2!} \frac{\Delta x_j^2}{\Delta \tau} \left(\alpha_2 - \alpha_1^2\right) \frac{\partial^2 s}{\partial x^2} + \frac{1}{3!} \frac{\Delta x_j^3}{\Delta \tau} \left(\alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^3\right) \frac{\partial^3 s}{\partial x^3} \\ &+ \frac{1}{4!} \frac{\Delta x_j^4}{\Delta \tau} \left(\alpha_4 - 4\alpha_1\alpha_3 - 3\alpha_2 + 12\alpha_1^2\alpha_2 - 6\alpha_1^4\right) \frac{\partial^4 s}{\partial x^4} + O\left(\frac{\Delta x_j^5}{\Delta \tau}\right) \end{aligned} \tag{3}$$

with the initial data given as the original discrete data $s_{j,l}$ provided $\sum_k a_k = 1$. Note that the evolution of the scalar s in (3) depends explicitly on the sampling rate Δx_j. In this context, filter design can be thought of as the design of a discrete approximation to this PDE which more naturally relates to preservation of spatial features than a frequency-based approach. A similar spatial filter analysis was also described in [18] where the filter performance was described in terms of spatial criteria by examining the non-zero discrete moments, i.e., the α_q in (2).

The frequency response or amplification factor of the filter is given by

$$G(\beta) = \sum_{k=-m}^{+n} a_k e^{ik\beta} \tag{4}$$

where the amplification factor represents the response for the frequency β. The magnitude of $G(\beta)$ measures the amplitude of a unit Fourier coefficient upon application of the filter and the phase of $G(\beta)$ measures the phase shift that occurs after application of the filter.

3. Feature Preservation

It is now appropriate to define what we mean by feature preservation. In this context, feature preservation implies that the "location", "shape", and "strength" of features are unchanged after the application of the general filter (1). Of course, differences naturally occur due to the change in resolution between $x_{j,l}$ and $x_{j-1,l}$. We now state these ideas in more concrete terms.

- The "location" of a feature is simply its position within the domain. As discussed in [17, 19], odd order derivative terms in the evolutionary PDE correspond to a convection or translation of features in the domain. If the filter is symmetric, $a_k = a_{-k}$ for all k, the coefficients multiplying the convective terms are identically zero and no translation of the data occurs.
- The "shape" of a feature can be described in terms of regions of monotone variation in the data. For the "shape" to be preserved, the linear filter should not introduce new extrema. This condition is expressed in [12] as the "causality condition." This condition can be imposed by ensuring that the linear transform is Total Variation Diminishing (TVD) after the data is subsampled [19], i.e., $a_k + a_{k+1} \geq 0$ for all k provided $\sum_k a_k = 1$.
- The "strength" of a feature can be described in terms of the changes in the data. For the strength to be preserved, the linear filter should not accentuate or diminish local extrema. This condition can be related to the frequency response (4) of the filter [19]. Our evolutionary PDE framework (3) characterizes changes in feature strength in terms of the even order derivative terms which represent the diffusive tendencies of the filter.

4. Filter Design Axioms

Having defined feature preservation, we now enumerate a list of filter design axioms. We want to formulate a set of requirements on the coefficients a_k that can guide the design of a wavelet transform associated with the coefficients a_k in addition to preserving features. We define the *restricted transfer operator* T as follows: if the length of the convolution product of the sequence of coefficients a_k with itself is N, then T is the $(N-2) \times (N-2)$ matrix obtained from double shifts of this convolution product, times 2 (see [20]). Let us impose the following requirements on the coefficients a_k:

$(R1)$ $\sum_k a_k = 1$

$(R2)$ $\sum_k (-1)^k k^j a_k = 0,$ for $j = 0, 1, ..., p-1,$ and some $p \geq 1$

$(R3)$ the restricted transfer operator T has one eigenvalue $\lambda = 1$, and all other eigenvalues have $|\lambda| < 1$

$(R4)$ $a_k + a_{k+1} \geq 0$ for all k

$(R5)$ $a_k = a_{-k}$ for all k

$(R6)$ if a_{-n} is the first nonzero coefficient, then the polynomials $a_{-n}+$ $a_{-n+2}z + a_{-n+4}z^2 + ...$ and $a_{-n+1} + a_{-n+3}z + a_{-n+5}z^2 + ...$ are relatively prime.

$(R7)$ between all filters with the desired properties, the filter given by the coefficients a_k minimizes the L^2 distance to the *sinc* filter.

(5)

Axioms $(R1)$, $(R4)$, and $(R5)$ are related to the feature preservation properties of the filter. Axiom $(R2)$ dictates the performance of the frequency response at $\beta = \pi$. Axioms $(R3)$ and $(R6)$ ensure that the proposed filter is the low pass filter of a wavelet transform which can be implemented as a series of lifting steps. Axiom $(R7)$ minimizes blurring and aliasing. Our three-fold framework provides for a complete specification of coefficients.

Theorem. *Requirements (R1)-(R7) are necessary and sufficient conditions for the following properties to hold:*

(a) *(Convergence of the cascade algorithm, see [20])*
The iteration $\phi^{(i+1)}(t) = \sum_k 2a_k\phi^{(i)}(2t-k)$, *where* $\phi^{(0)}$ *is a box function, converges in* L^2.

(b) *(Accuracy of approximation of order p)*
The error estimate for a function $f(t)$ *of class* C^p *at scale* $\Delta t = 2^{-j}$ *is of the form* $C(\Delta t)^p|f^{(p)}(t)|$.

(c) *(Total variation diminishing from finer to coarser scales)*
$TV(s_{j-1}) \leq TV(s_j)$.

(d) *(Zero phase shift from finer to coarser scale)*
In the evolutionary PDE (3), all coefficients multiplying even order derivatives are zero.

(e) *(Lifting scheme implementation, see [21])*
There exists complementary high-pass filter, and the associated wavelet transform admits in-place implementation using the lifting scheme.

(f) *(Average grey level invariance)*
The average of the data is unchanged when passing from finer to coarser scales.

(g) *(Preservation of low frequencies)*
The moment of order 0 is 1, and the moment of order 1 is 0.

(h) *(Optimality)*
Between all filters with the desired properties, the filter given by the coefficients a_k *minimizes the* L^2 *distance in the frequency domain to the ideal brick wall filter.*

Properties $(a)-(h)$ are related to $(R1)-(R7)$ as follows: (a) is equivalent to $(R3)$, (b) is equivalent to $(R2)$, (c) is implied by $(R1)$ and $(R4)$, (d) is implied by $(R5)$, (e) is equivalent to $(R6)$, (f) is equivalent to $(R1)$, (g) is implied by $(R1)$ and $(R5)$, and (h) is equivalent to $(R7)$.

4.1 Examples

Notice that requirements $(R1)-(R7)$ are in terms of the coefficients a_k. Therefore, it is possible to use all or some of these requirements, together with conditions derived from the properties of the data with which we work, in order to design wavelets that are optimal for feature preservation.

For example, suppose we are looking for the shortest filter that, for $p=1$, satisfies requirements $(R1)-(R6)$. Due to $(R5)$ the length of the filter will have to be odd. Due to $(R2)$ the length of the filter cannot be 1, so it has to be at least 3. Then, by using $(R1)$ and $(R2)$ we get $a_{-1} = 1/4$, $a_0 = 1/2$, $a_1 = 1/4$.

If we now check $(R3)$, $(R4)$, and $(R6)$ we will notice that they also hold for this choice of coefficients. Let us illustrate this by looking at condition $(R3)$. The convolution product of $(1/4,\ 1/2, 1/4)$ with itself is $(1/16,\ 1/4,\ 3/8,\ 1/4,\ 1/16)$, $N = 5$, and the corresponding restricted transfer operator has eigenvalues $1,\ 1/2,\ 1/4$. Therefore the TVD filter coefficients $(1/4,\ 1/2,\ 1/4)$ are the only solution to $(R1) - (R6)$ for filter lengths less than 5.

As another example, suppose we are looking for filters of length 5 that satisfy $(R1) - (R6)$ for the largest value of p possible. By using $(R1)$, $(R2)$, and $(R5)$ we find that there is no solution for $p = 5$, and for $p = 4$ there is an unique solution $a_{-2} = 1/16$, $a_{-1} = 1/4$, $a_0 = 3/8$, $a_1 = 1/4$, $a_2 = 1/16$. These filter coefficients are the optimal solution to $(R1) - (R6)$ if we are looking at filters of length less than 7 and we want to maximize the accuracy of approximation for smooth input. If we relax the condition on p, or if we look at filters of length 7 or more, there will be room for a design strategy that allows us to approach the ideal brick wall filter while satisfying other functional and variational criteria. Additionally, the ideal brick wall filter (a *sinc* function in the spatial domain) will not satisfy the TVD condition. Thus, any filter design strategy will seek a compromise between the ideal frequency behavior and the feature preservation properties of TVD filters. Given that this is a work-in-progress, we use these axioms to analyze existing wavelet filtering schemes rather than design new ones. We show our results in the next section.

5. Preliminary Results

We now demonstrate some of the concepts presented in the previous sections. The lifting scheme consists of three steps: split, predict, and update [4]. By combining the three steps, a single expression in the form of (1) can be derived for the updated values. Three basic forms of lifting described in [4] can be defined using the coefficients given in Table 1. We also consider the simple $(1/4,\ 1/2,\ 1/4)$ TVD filter derived in the previous section. By inspection, all four schemes given in Table 1 satisfy the partition of unity $(R1)$. Of the four schemes considered here, only the versions using the Haar and TVD wavelets satisfy the TVD constraint $(R4)$.

The implications of the failure of the different filters to satisfy the axioms given in Section 4 are investigated by applying the schemes to the data used in [10]. We have assumed periodicity so that boundary effects can be ignored. Figure 1 shows three levels of transform with the original data consisting of 128 discrete, equally-spaced values. Here, one application of the transform consists of applying the filter (1) using the selected coefficient values and then downsampling the data. From Figure 1, it is apparent that only results obtained using Haar lifting and TVD lifting may be characterized as TVD. New extrema (wiggles) are created by both linear and cubic lifting for the two square wave signals. This behavior is caused by the violation of $(R4)$ and is also observed to a much lesser degree for the triangle wave. The lack of symmetry in the results for the Haar wavelet is caused by the violation of $(R5)$. Since the filter is not symmetric, the coefficients of the convective terms in the evolutionary PDE are nonzero and the square wave translates to the left. Notice there is some asymmetry for the linear, cubic, and TVD filters for the middle square wave. This behavior is caused by the fact that the the wavelength of this feature is of insufficient length to be captured symmetrically by the second and third downsamplings. The dissipative nature of the symmeteric TVD filter is observed in the reduction in amplitudes of the triangle wave and the middle square wave and represents the primary undesirable characteristic of this filter. This behavior is caused by the fact that the frequency response of the $(1/4,\ 1/2,\ 1/4)$ filter departs from the ideal brick wall filter $(R7)$. Our long-term goal is to develop linear filters that improve upon this behavior.

6. Conclusions

In this paper we defined a framework for the analysis and design of multiscale filters through a variational characterization and a multiscale PDE formulation. Included in this framework are a set of axioms that can be used to design filters that preserve certain characteristics of the data—namely the position, shape, and strength of features. We showed that three standard forms of the lifting scheme violate at least one of the criteria and that data filtered using these schemes exhibited undesirable characteristics. We further demonstrated that the simple $(1/4,\ 1/2,\ 1/4)$ filter satisfies these criteria and shows potential as a component of a linear TVD wavelet. We suggest that the methods proposed here can be used in conjunction with frequency-based methods to design multiscale linear wavelet filters. We plan to utilize these techniques to develop advanced wavelets with feature-preserving qualities.

7. Acknowledgments

This work is partially funded by the NSF under the Large Data and Scientific Software Visualization Program (ACI-9982344), an NSF Early Career Award (ACI-9734483), and the Mississippi State University Engineering Research Center for Computational Systems.

References

[1] R. Machiraju, J. Fowler, D. Thompson, W. Schroeder, and B. Soni. EVITA - Efficient visualization and interrogation of tera-scale data. *Data Mining for Scientific and Engineering Applications (to appear)*, 2001.

[2] R. Machiraju, B. Nakshatrala, and D. Thompson. Feature-based ranked representation of vector fields. *Proc. of Dagstuhl 2000 Seminar on Scientific Visualization (submitted)*, 2001.

[3] B. Nakshatrala. Feature-based embedded representation of vector fields. Master's thesis, Mississippi State University, December 1999.

[4] W. Sweldens. The lifting scheme: A construction of second generation wavelets. *SIAM J. Math. Anal.*, 29(2):511–546, 1997.

[5] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, 1992.

[6] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1988.

[7] D. Donoho. Denoising by soft thresholding. *IEEE Trans. on Inf. Theory*, 41:613–627, 1995.

[8] D. Donoho, I. Daubechies, R. DeVore, and M. Vetterli. Adapting to unknown smoothness via wavelet shrinkage. *J. Amer. Stat. Assoc.*, 90:1200–1224, 1995.

[9] P. Claypoole, G. Davis, W. Sweldens, and R. Baraniuk. Nonlinear wavelet transforms for image coding. *Asilomar Conference on Systems, Signals and Computers, 1997 (preprint)*.

[10] H.-M. Zhou. *Wavelet transforms and pde techniques in image compression*. PhD thesis, University of California Los Angeles, 2000.

[11] A. Harten, E. Enquist, S. Osher, and S. Chakravarthy. Uniformly high order essentially non-oscillatory scheme, III. *J. Comp. Phys.*, 71:231–303, 1987.

[12] P. Perona and J. Malik. Scale-space diffusion and edge detection using anisotropic diffusion. *IEEE Trans. Pat. Anal. Int.*, 12:629–639, 1990.

[13] Special Issue on Partial Differential Equations and Geometry-Driven Diffusion. *IEEE Trans. Image Proc.*, 7(3), March 1998.

[14] B. Kimia, A. Tannenbaum, and S. W. Zucker. Towards a computational theory of shape. *Lecture Notes in Computer Science*, 427, 1990.

[15] J. Weickert, B. M. ter Haar Romeny, and M. A. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Trans. Image Proc.*, 7(3):398–410, March 1998.

[16] L. Alvarez, F. Guichard, P. L. Lions, and J. M. Morel. Axioms and fundamental equations of image processing. *Arch. Rat. Mech. Anal.*, 16(9):199–257, 1993.

[17] D. Thompson, R. Machiraju, G. Craciun, and M. Jiang. Feature preservation properties of linear wavelet transforms. *Proc. of YES2000 (submitted)*, 2001.

[18] T. Möeller, R. Machiraju, K. Mueller, and R. Yagel. Evaluation and design of optimal filters using a Taylor series expansion. *IEEE Trans. Vis. and Graphics*, 3(2):184-199, June 1997.

[19] D. Thompson, R Machiraju, G. Craciun, and M. Jiang. Multiscale Feature Preservation in Discrete Flow Fields. *in preparation*, 2001.

[20] G. Strang and T. Nguyen. *Wavelets and Filter Banks.* Wellesley-Cambridge Press, 1996.

[21] W. Sweldens and I. Daubechies. Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl.*, 29(2):511–546, June 1997.

Table 1: Coefficient values for selected versions of the lifting scheme

Haar	$a_0 = \frac{1}{2}, a_{+1} = \frac{1}{2}$
Linear	$a_0 = \frac{3}{4}, a_{\pm 1} = \frac{1}{4}, a_{\pm 2} = -\frac{1}{8}$
Cubic	$a_0 = \frac{348}{512}, a_{\pm 1} = \frac{144}{512}, a_{\pm 2} = -\frac{63}{512}, a_{\pm 3} = -\frac{16}{512},$
	$a_{\pm 4} = \frac{18}{512}, a_{\pm 5} = 0, a_{\pm 6} = -\frac{1}{512}$
Symmetric TVD	$a_0 = \frac{1}{2}, a_{\pm 1} = \frac{1}{4}$

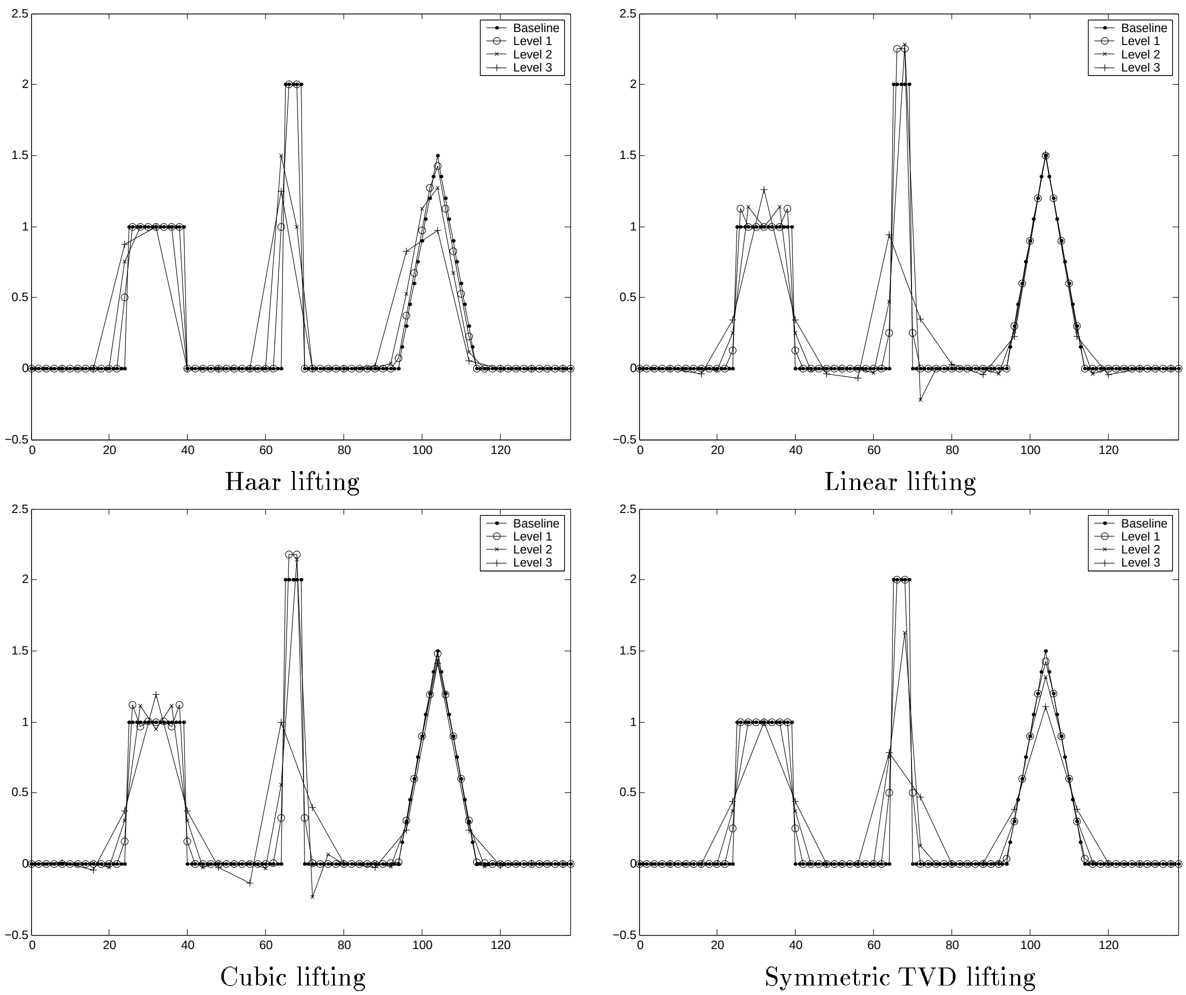

Figure 1: Three levels of lifting for the data of [10]–Haar, linear, cubic, and symmetric TVD