92-10

Time-Frequency and Time-Scale Analysis

Sponsored by

The IEEE Signal Processing Society

Organized by

October 4 – 6, 1992
Victoria, BC, Canada

92TH0478-8

THE PARAMETRIC WAVELET MAXIMA REPRESENTATION

Zeev Berman, John S. Baras, and Carlos Berenstein

Systems Research Center, University of Maryland , College Park, MD-20742.

Abstract

A general signal description, called an inherently bounded Adaptive Quasi Linear Representation (AQLR), motivated by two important examples, namely, the wavelet maxima representation, and the wavelet zero-crossings representation, is introduced. This paper presents a new reconstruction scheme based on the minimization of an appropriate cost function. The convergence of two described algorithms is guaranteed for all inherently bounded AQLR. As a consequence, we describe possible modifications in the basic multiscale maxima representations.

INTRODUCTION

An interesting and promising approach to signal representation is to make explicit important features in the data. The first example, taught in elementary calculus, is a "sketch" of a function based on extrema of a signal and possibly of its first few derivatives. The second instance, widely used in computer vision, is an edge representation of an image. If the size of expected features is a priori unknown, the need for a multiscale analysis is apparent. Therefore, it is not surprising that multiscale sharp variation points (edges) are meaningful features for many signals, and they have been applied, for example, in edge detection [4, 9], signal compression [8], pattern matching [7], detection of transient signals [5, 6] and speech analysis [10].

S. Mallat in [7, 8] (the last joint with S. Zhong) introduced zero-crossings and extrema of the wavelet transform as a multiscale edge representation. Two important advantages of this method are low algorithmic complexity and flexibility in choosing the basic filter. Moreover, [7] and [8] propose reconstruction procedures and show accurate numerical reconstruction results from zero-crossings and maxima representations. In [7, 8], as in many other works in this area, the basic algorithms were developed using continuous variables. The continuous approach gives an excellent background to motivate and justify the use of either local extrema or zero-crossings as important signal features. Unfortunately, in the continuous framework, analytic tools to investigate the information content of the representation are not yet available. The knowledge about properties of the representations is mainly based on empirical reconstruction results.

The main goal of our undergoing research is to understand, analyze, and generalize the numerical reconstruction results from the wavelet maxima and zero-crossings representations, as described in [8, 7]. This objective leads to the discrete and finite data assumption. The first observation is that the structure of the wavelet transform is not essential for the analysis and can be generalized to any linear filter bank. The precise definitions of the multiscale maxima (zero-crossings) have been introduced in [2]. Since reconstruction sets (the family of signals having the same representation) of maxima and zero-crossings representations have a similar structure, a general form called Adaptive Quasi Linear Representation (AQLR) is introduced.

Using this framework, we have shown [3] that, in general, neither the wavelet maxima representation nor the wavelet zero-crossings representation is unique. In other words, for any discrete dyadic wavelet maxima (zero-crossings) representation there exists a sequence (an appropriate sinusoid) which has a nonunique representation. Due to this result, we consider the reconstruction from the wavelet maxima representation in a general set-up of point-to-set maps.

The next investigated subject was stability of the representation. This issue is of great importance because there are many known examples of unstable zero-crossings representations. In order to improve the stability properties Mallat has included additional sums in the standard zero-crossings representation and together with Zhong, they have introduced the wavelet maxima representation. Indeed, they have reported very good numerical results. It turns out that stability is closely related to boundedness of the reconstruction set. By introducing the idea of the inherently bounded AQLR, we were able to prove stability results. For a general perturbation, global BIBO stability can be shown. For a

0-7803-0805-0/92$3.00©1992 IEEE

special case, where perturbations are limited to the continuous part of the representation, a Lipschitz condition is satisfied (for details see [3] or [2]).

One of the most important practical problems is the need for an effective reconstruction scheme. Mallat and Zhong [8] and Mallat [7] have used an algorithm based on alternate projections. In this paper, a new reconstruction scheme, defined for a general inherently bounded AQLR, is proposed. It is based on the minimization of the appropriate cost function which is zero on the reconstruction set and it is positive otherwise. Thus the reconstruction can be achieved by any minimization technique. In particular, the convergence of two algorithms is shown: the first is based on the integration of the gradient of the cost function and can be implemented by analog hardware; the second is a standard steepest descent algorithm which is used in digital simulations.

Having developed the framework of the inherently bounded AQLR, we consider the question how to generalize the basic wavelet maxima representation in order to trade off the quality of the representation with the amount of information required to describe the representation. The result is the parametric wavelet maxima representation which has the ability to add or delete information from the basic representation. In turns out that the related reconstruction algorithms are very similar to those corresponding to the basic wavelet maxima representation.

PREVIOUS RESULTS

This section describes, in the discrete context, the representations proposed in [8] and extends it to a multiscale case .

Loosely speaking, a multiscale maxima representation is based on a linear filter bank followed by one-level maxima representations. A filter bank will be denoted as a set of $J+1$ linear operators $\{W_1,\ldots,W_J,S_J\}$. In the sequel, this filter bank is assumed to be complete in the sense that $\{W_1f,\ldots,W_Jf,S_Jf\}$ is a unique representation of a signal f.

In this work signals are interpreted as real, finite sequences. One-level maxima representation of a sequence $\{f(k)\}_{k=0}^{N-1}$ consists of local extrema points (indices at which local extrema occur) of the sequence f and the values of a sequence f at these points. Precise definitions, proofs, and many additional details can be found in [1].

A multiscale maxima representation consists of one-level maxima representations of signals W_jf ($j=1,2,\ldots,J$). In addition, one entire sequence S_Jf is allowed to be a part of this representation. If $\{W_1,\ldots,W_J,S_J\}$ describes a wavelet decomposition (see [8] for details) then the corresponding multiscale maxima representation is called the wavelet maxima representation.

For a given multiscale maxima representation Rf, the corresponding reconstruction set Γ is defined as a set of all signals having the same multiscale maxima representation, namely $\Gamma \triangleq \{x : Rx = Rf\}$.

In general, any multiscale maxima representations can be cast into the form $Rf=\{Vf,Tf\}$. Vf consists of sets of points from $\{0,1,\ldots,N-1\}$ and T, for a fixed Vf, is a linear operator.

It turns out that the maxima representation implies constrains on Γ (having a local extremum at a given point) which do not appear directly in Rf. These constraints may be described in a common structure.

Definition 1 $Rf=\{Vf,Tf\}$ *is called an Adaptive Quasi Linear Representation (AQLR) if there exists a linear operator* C *and a sequence* c *such that:* $x\in\Gamma$ *if and only if* $Tx=Tf$ *and* $Cx>c$. C,c *may depend on* Vf, *but they must be independent of* Tf.

Observe that, since every equality can be replaced by two inequalities, the closure of the reconstruction set $\overline{\Gamma}$ can be written as: $\overline{\Gamma}=\{x : \mathbf{B}x\geq \mathbf{b}\}$ for a given $p\times N$ matrix $\mathbf{B}$ and a p-dimensional vector $\mathbf{b}$.

The following definition and proposition are essential for stability and reconstruction results.

Definition 2 *An AQRL is called inherently bounded if there exists a real* $K>0$ *such that*

$$x\in\Gamma \Rightarrow \|x\|\leq K\|Tf\|.$$

Proposition 1
Any multiscale maxima (zero-crossings) representation in an inherently bounded AQLR.

THEORY FOR RECONSTRUCTION

In this section we assume that an arbitrary inherently bounded Adaptive Quasi Linear Representation (AQLR) is given. A reconstruction algorithm is defined as a procedure to find an arbitrary element x belonging to the closure of the reconstruction set, $\overline{\Gamma}$. As mentioned earlier, we propose a reconstruction algorithm based on an appropriate potential function $v(x)$.

The function $v(x)$ is derived from the representation $\overline{\Gamma}=\{x : \mathbf{B}x\geq \mathbf{b}\}$ in the following way.

$$v(x) \triangleq \sum_{i=1}^{p} \rho(\mathbf{B}x-\mathbf{b})_i \tag{1}$$

where $(\mathbf{B}x-\mathbf{b})_i$ denotes the i-th component of the vector $\mathbf{B}x-\mathbf{b}$. The function $\rho(\cdot)$ is defined by:

$$\rho(\xi) \triangleq \begin{cases} \xi^2 & \text{if } \xi\leq 0 \\ 0 & \text{otherwise} \end{cases}.$$

In words, $v(x)$ is calculated by scanning all p constraints which define the closure of the reconstruction set, and summing up penalties for every constraint. For a particular constraint, the penalty is zero if the constraint is satisfied. If the constraint is violated then a quadratic term is used as the penalty. It is easy to verify that the nonnegative cost function $v(x)$ is equal to zero if and only if $x \in \overline{\Gamma}$.

Thus reconstruction can be implemented by any minimization of the cost function $v(x)$. Observe that $\rho(\xi)$ is continuously differentiable. Therefore $v(x)$ is continuous and continuously differentiable. Moreover, it can be shown (see [1]) that $v(x)$ has no local extrema outside $\overline{\Gamma}$, namely $\nabla v(x) = 0$ if and only if $x \in \overline{\Gamma}$.

Since, usually, p is a large number, we are mostly interested in iterative minimization procedures. To show convergence of such schemes, the following boundedness property is required.

Lemma 1 *For all $K_v > 0$ there exists $K_x > 0$ such that $v(x) \leq K_v$ implies $\|x\| \leq K_x$*

Using Lemma 1 and La Salle's Theorem, we are able to prove the following result [1].

Theorem 1 *For all $x(0)$, the solution of*

$$\dot{x}(t) = -\nabla\left(v\left(x(t)\right)\right). \tag{2}$$

will approach $\overline{\Gamma}$ as $t \to \infty$, namely the distance between $x(t)$ and $\overline{\Gamma}$ converges to zero as $t \to \infty$.

Theorem 1 enables us to use a very fast analog-hardware implementation to reconstruct signals. However, before acquiring a costly and not flexible hardware, an ability to perform digital simulations is required. The following theorem defines a steepest descent algorithm, based on the cost function $v(x)$, and states its convergence.

Theorem 2 *For any x_0, we define the sequence $\{x_k\}$.*

$$x_{k+1} = x_k - \alpha_k \cdot \nabla v(x_k) \quad k = 0, 1, 2, \ldots \tag{3}$$

where α_k is a nonnegative scalar minimizing $v(x_k - \alpha_k \cdot \nabla v(x_k))$. Then x_k approaches $\overline{\Gamma}$ as $k \to \infty$.

PRACTICAL IMPLEMENTATION

In order to implement the reconstruction algorithms described in the previous section one needs to calculate the cost function $v(x)$ and its gradient $\nabla v(x)$. A direct calculation, based on the system of inequalities $\mathbf{B}x \geq \mathbf{b}$, may yield unnecessary high complexity, related to the use of the "large" matrix $\mathbf{B}$. In this section, using the structure of the multiscale maxima representation, an efficient algorithm to calculate the cost function $v(x)$, and its gradient $\nabla v(x)$ is described.

The cost function $v(x)$ is calculated by taking into account all conditions that x should satisfy in order to belong to the closure of the reconstruction set, $\overline{\Gamma}$. For the multiscale representation, these conditions can be clustered according to the different scales. To be more specific, let us consider the multiscale maxima representation Rf composed of $R_j f$, one-level maxima representations of $W_j f$ $(j = 1, 2, \ldots, J, J+1)$ [1]. Then the cost function $v(x)$ can be written as

$$v(x) = \sum_{j=1}^{J+1} \nu(W_j x, R_j f), \tag{4}$$

where $\nu(W_j x, R_j f)$ is called the local cost function and describes how well the sequence $W_j x$ matches the one-level maxima representation $R_j f$. It is calculated from all constraints implied by the one-level maxima representation of $W_j f$.

It turns out, that the algorithmic complexity of calculating the cost function does not exceed the algorithmic complexity of calculating the multiscale decomposition. A similar statement is true for the calculation of the cost function gradient, $\nabla v(x)$.

Equation (4) yields

$$\nabla v(x) = \sum_{j=1}^{J+1} \mathbf{W}_j' \cdot \nu_y(W_j x, R_j f). \tag{5}$$

where $\mathbf{W}_j'$ is the transpose of the matrix corresponding to the operator W_j. The local gradient $\nu_y(W_j x, R_j f)$ is a column vector consisting of derivatives of $\nu(\cdot,\cdot)$ with respect to components of the first argument.

The gradient calculation consists of four steps: calculate the decomposition $\{W_j x\}_{j=1}^{J+1}$, calculate local gradients $\nu_y(W_j x, R_j f)$, calculate $\mathbf{W}_j' \cdot \nu_y(W_j x, R_j f)$, sum up results for $j = 1, 2, \ldots, J+1$.

GENERALIZATIONS

As a side benefit of the above mathematical analysis, structural attributes required to attain the described stability and reconstruction characteristics have been well understood. This knowledge enables us to introduce many modifications while preserving the desired properties within the framework of inherently bounded AQLR's.

Our main objective is to create a structure allowing a trade-off between the amount of required information and the reconstruction quality.

[1] $R_{J+1} f$ denotes here $S_J f$ and $W_{J+1} x$ is used instead of $S_J x$

In the wavelet transform case, $S_J f$ is a significantly blurred version of f and the whole sequence $\{S_J f(k)\}_{k=0}^{N-1}$ appears to contain redundant information. A version of $S_J f$, downsampled at rate Δ, is defined as follows: $S_J^{\Delta} f \triangleq \{S_J f(k)\}_{k=0,\Delta,2\Delta,\ldots}$ It turns out, that the wavelet maxima representation with downsampled $S_J f$, for $\Delta = 2^j$ $j = 0, 1, 2, \ldots, J$, is an inherently bounded AQLR.

When one considers the amount of information required to describe a given representation, perhaps the most important issue is the arithmetic precision in which the values $W_j f(k)$, at extreme points, are described. Observe that, even if raws of T are linearly independent, due to additional constraints $Cx \geq c$, approximate values of Tx may lead to a representation with an empty reconstruction set !

We propose to overcome this problem by including quantization as a part of the representation and the reconstruction. The main idea is that we replace a precise sample $W_j f(k)$ by an approximation interval, say with a center $(W_j f(k))_q$ and length $2q$. Then instead of requiring $W_j x(k) = (W_j f(k))_q$, we require

$$(W_j f(k))_q - q \leq W_j x(k) \leq (W_j f(k))_q + q.$$

Using this approach we preserve the structure of the inherently bounded AQLR. Moreover, the algorithm complexity of the reconstruction is preserved as well.

The following example has been obtained using the cubic spline wavelet with $N = 256$, $J = 5$, and $f(k) = \sin\left(\frac{6\pi k}{256}\right) \cdot e^{-\frac{k}{128}}$. The abscissa of Figure 1 gives the number of iteration of a descent reconstruction algorithm. The ordinate describes the noise (reconstruction error) to signal ratio. Continuous line presents the behavior of the reconstruction from the basic maxima representation. Dashed line describes the reconstruction from the representation $R^{\Delta,q} f$, in which, first S_J was reduced by the factor 32 and then quantization was performed using interval length $2q = 0.02$ corresponding to 8-bit digital representation.

Form the graph and from many additional examples studied in [1] we may conclude that it is very easy, robust, and fast to get results corresponding to noise to signal ratio of $\approx 10^{-2}$. Moreover, the reconstructed signals, even for a high noise to signal ratio, appear to be very " similar " to the original one.

Conclusions The structure of the inherently bounded AQLR is a framework to develop, analyze, and test many representation and reconstruction algorithms. Further study, more application oriented, is required to develop methods to choose appropriate modifications and their parameters.

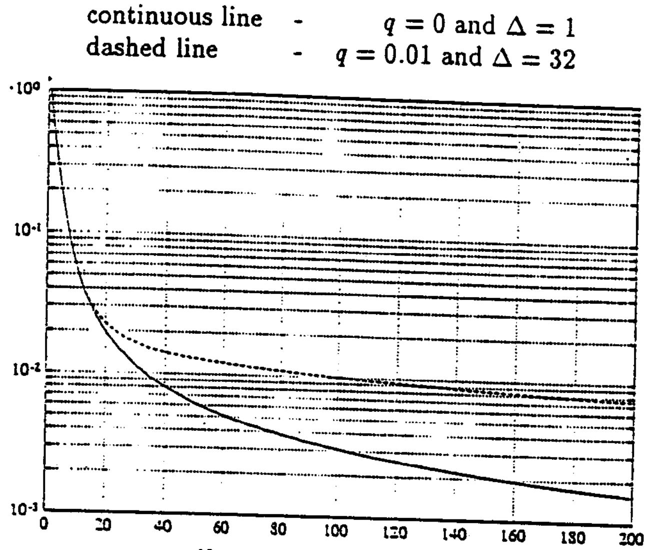

Figure 1: $\frac{N}{S}$ for reconstruction from $R^{\Delta,q} f$

References

[1] Z. Berman. *Generalizations and properties of the multiscale maxima and zero-crossings representations.* Technical Report Ph.D 92-9, University of Maryland, System Research Center, August 1992.

[2] Z. Berman and J. S. Baras. *A Theory of Adaptive Quasi Linear Representations.* Technical Report TR 92-24, University of Maryland, System Research Center, January 1992.

[3] Z. Berman and J. S. Baras. Uniqueness and stability of discrete zero-crossings and maxima wavelet representations. In *Proc. Sixth European Signal Processing Conference*, Brussels, 1992.

[4] J. Canny. A comptutational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8():679–698, 1986.

[5] M. Frish and H. Messer. The use of the wavelet transform in the detection of an unknown transient signal. *IEEE Trans. on Information Theory*, 38(2):892–897, March 1992.

[6] S. Kadambe and G. F. Boudreaux-Bartels. Application of the wavelet transform for pitch detection of speech signals. *IEEE Trans. on Information Theory*, 38(2):917–924, March 1992.

[7] S. Mallat. Zero-crossing of a wavelet transform. *IEEE Trans. on Information Theory*, 37(4):1019–1033, July 1991.

[8] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. Courant Institute of Mathematical Sciences, Technical Report 592, November 1991, to appear in IEEE Trans. on PAMI.

[9] D. Marr and E. Hildreth. Theory of edge detection. *Proc. of the Royal Society of London*, 207():187–217, 1980.

[10] X. Yang, K. Wang, and S. A. Shamma. Auditory representation of acoustic signals. *IEEE Trans. on Information Theory*, 38(2):824–839, March 1992.