OPTIMAL WAVELETS FOR SIGNAL DECOMPOSITION AND THE EXISTENCE OF SCALE-LIMITED SIGNALS

J. E. Odegard R. A. Gopinath C. S. Burrus

Department of Electrical and Computer Engineering,
Rice University, Houston, TX-77251

ABSTRACT

Wavelet methods give a flexible alternative to Fourier methods in non-stationary signal analysis. The concept of *band-limitedness* plays a fundamental role in Fourier analysis. Since wavelet theory replaces *frequency* with *scale*, a natural question is whether there exists a useful concept of *scale-limitedness*. Obvious definitions of scale-limitedness are too restrictive, in that there would be few or no useful scale-limited signals. This paper introduces a viable definition for scale-limited signals, and shows that the class is rich enough to include bandlimited signals, and impulse trains, among others. Moreover, for a wide choice of criteria, we show how to design the optimal wavelet for representing a given signal, and how to design *robust* wavelets that optimally represent certain classes of signals.

Also Technical Report Rice University, CML TR91-07

1. INTRODUCTION

The recognition and existence of band limited signals plays a fundamental role in sampling theory and Fourier analysis. However, the wavelet transform (WT) or time-scale decomposition does not tie so nicely with the concept of band limited. We will in this paper explore this lack of correspondence and introduce a notion of scale-limited signals. Since we can not speak of "the" wavelet transform, the recognition of scale limited signals will be with respect to a given wavelet. A closely related problem is the design of wavelet(s) for best approximation of a class of signals. In particular we will show how bandlimited signals can be approximated up to a desired error with a finite number of scales. This paper is limit to consider the compactly supported wavelets [1] of multiplicity M [2].

2. MULTIPLICITY M WAVELET TIGHT FRAMES

Let h_0 be a *scaling vector* of length N. Then there exists a *scaling function* $\psi_0(t)$,

$$\psi_0(t) = \sqrt{M}\sum_{k=0}^{N-1} h_0(k)\psi_0(Mt-k), \tag{1}$$

wavelet vectors $h_l(k)$, $l = 1, 2, \ldots M-1$ of length N such that the $M-1$ *wavelets* defined by

$$\psi_l(t) = \sqrt{M}\sum_{k=0}^{N-1} h_l(k)\psi_0(Mt-k) \text{ for } l = 0, 1, \ldots M-1. \tag{2}$$

gives rise to a wavelet tight frame (Gopinath and Burrus [2]) defined by

$$\psi_{l,j,k}(t) = M^{j/2}\psi_l(M^j t - k). \tag{3}$$

In most cases the set, $\{\psi_{l,j,k}\}$, forms an orthonormal basis for $L^2(\mathbb{R})$. Then according to Mallat's multiresolution analysis [3], the functions $\{\psi_{l,j,k}\}$ spans spaces $W_{l,j}$ and

$$W_{0,j+1} = \bigoplus_{l=0}^{M-1} W_{l,j}. \tag{4}$$

Given some signal $f(t)$, the discrete wavelet expansion onto the wavelets ψ_l of the signal is given by

$$f(t) = \sum_{l=1}^{M-1}\sum_{j,k} Wf(l,j,k)\psi_{l,j,k}(t) \tag{5}$$

where $Wf(l,j,k)$ are the wavelet expansion coefficients given by

$$Wf(l,j,k) = \langle f, \psi_{l,j,k}\rangle . \tag{6}$$

Theoretically both summations over j, k are infinite; however, for practical reasons finite limits on j, k are necessary. The question one then need to answer is: can j be restricted to cover a finite range?

Assume for now that the signal $f(t)$ resides on some fine scale J, we trivially have

$$\begin{aligned} f(t) &= \sum_k Wf(0,J,k)\psi_{0,J,k}(t) \\ &= \sum_{l=1}^{M-1}\sum_k\sum_{j=-\infty}^{J-1} Wf(l,j,k)\psi_{l,j,k}(t) \end{aligned} \tag{7}$$

where $Wf(0,J,k)$ is the scaling expansion coefficients at scale J. We also observe that if $f(t)$ is a finite duration signal, the summation over k is finite. Combining both the scaling expansion and the wavelet expansion we get

$$\begin{aligned} f(t) &= \sum_k Wf(0,J_{min},k)\psi_{0,J_{min},k}(t) \\ &+ \sum_{l=1}^{M-1}\sum_k\sum_{j=J_{min}}^{J-1} Wf(l,j,k)\psi_{l,j,k}(t) \end{aligned} \tag{8}$$

where $J_{min} \leq J$. In the above equation the lower limit J_{min} is generally of no concern since the scaling function at scale J_{min} will capture the remaining information about $f(t)$ not represented by the wavelet expansion coefficients.

However, it is of great importance to accurately be able to estimate the "correct" J. Furthermore, we would also like to know how to chose (if possible) the smallest possible J such that the above equation still holds.

It is realistic to assume that the signal energy is small for sufficiently large frequency and hence will have vanishing wavelet coefficients for some level j greater than J. A natural question to pose based on the above observation is whether scale limited signals exist. That is, does there exist signals where J is finite? Answering this question will remove some of the ambiguity of implementing Eqn. 5 on a digital computer since finite limits on j then can be justified.

3. SCALE LIMITED

Consider first the obvious definition of scale limited signals

Definition 1 *Given a signal $f(t) \in L^2(\mathbb{R})$ we shall say that f is scale limited to J with respect to the wavelet ψ iff*

$$Wf(l,j,k) = 0 \quad \forall j > J, \quad \forall\ l,k. \tag{9}$$

That is, the wavelet coefficients vanishes for scales $j > J$ for all l, k.

Based on this definition we see that scale limited signals will be with respect to a particular wavelet. Hence, signals scale limited with respect to a wavelet ψ will in general not be scale limited with respect to any other wavelet $\psi\prime$.

Definition 1 is very restrictive with respect to knowledge of the origin of the signals, and in general translates of $f(t)$ is not a part of the class. In fact, it is generally well known that a simple shift of f results in vastly different distribution of the energy in the wavelet expansion coefficients (Mallat [4] and Gopinath [5]). Hence, one measurement on f might indicate a scale limited signal whereas a simple shift could change J, or even in the worst case make J tend to infinity.

In general, for signal processing, the origin is not very important and we would like signals and all it's translates to be in the class in analogy with the definition of bandlimitedness. Based on this we define signals as being strictly scale limited if they satisfy the following definition.

Definition 2 *Given a signal $f(t) \in L^2(\mathbb{R})$ we shall say that f is strictly scale limited to J with respect to the wavelet ψ iff*

$$Wf_T(l,j,k) = 0 \quad \forall j > J, \quad \forall\ l,k. \tag{10}$$

That is, the wavelet coefficients vanishes for scales $j > J$ for all l, k and translates T.

However, if Definition 2 is the desired framework for analysis of scale limitedness we immediately have the following trivial lemma.

Lemma 1 *The class of strictly scale limited signals (i.e., signals satisfying Definition 2) contains only the trivial zero signal.*

Proof: To show that in general only the zero signal satisfy the definition consider a signal $f(t) \in L^2(\mathbb{R})$ and an arbitrary translation $f(t-T)$. Now compute the contribution of $f(t-T)$ at an arbitrary scale J.

$$\int f(t-T)\psi(M^J t)dt = \int \hat{f}(\omega)e^{-j\omega T}\hat{\psi}(M^{-J}\omega)d\omega$$

From this we observe that there exist a T for which the above expression is non-zero for any J except if the $\mathcal{FT}\{f\}$ is the zero signal. (Gopinath, Odegard and Burrus [6]). □

The definitions of scale limited and strictly scale limited given above would in general restrict the class of interesting signals. Hence, a less restrictive definition of scale limited is formulated based on a least square error principle. Incorporating a tolerance error in the scale limited definition gives us more flexibility and hence will potentially be easier to achieve. Using these ideas we propose the following definition of essentially scale limited.

Definition 3 *Given a signal $f(t) \in L^2(\mathbb{R})$ we will say that f is essentially scale limited to J with respect to the wavelet ψ iff*

$$\frac{\sum_l \sum_k \sum_{j=J}^{\infty} \left| Wf_T(l,j,k) \right|^2}{\|f\|_2^2} \le \varepsilon \tag{11}$$

where $f_T(t) = f(t-T)$. That is, the total energy of the wavelet coefficients are "acceptably" small for scales $j > J$ for all translates T.

This alternate definition does not ensure a true scale limited signal as defined in Definition 1. However, it does guarantee an upper bound on the error energy, incurred by assuming that the signal is scale limited to J.

In particular we will show that bandlimited signals and impulse trains will satisfy Eqn. 11, and we will also introduce a *wavelet sampling thorem.*

4. DESIGN OF OPTIMAL WAVELETS

An important problem in wavelet analysis is to find the best wavelet that approximates a given signal in some norm. Another related problem, especially when trying to study a class of signals, is to find the wavelet that minimize the worst case approximation error over all signals in a class. This wavelet, if it exist, will be referred to as the "robust wavelet". The former problem has been addressed by Tewfik and Jorgensen [7]. In the paper they derive upper bonds on the L^2 and L^∞ approximation error of a given signal up to any desired finite scale. The problem of obtaining robust wavelets was considered by Gopinath [8]. In this case the error can be modeled as the output of an error operator acting on the function, hence robustness is achieved by minimizing the induced norm of the operator. We will show how implicit expressions for the error operator norms are obtained, and in the case of bandlimited we obtain explicit expressions for both the optimal and robust wavelets for the cases $L^{p,p}(\mathbb{R})$ and $L^{p,1}(\mathbb{R})$ (where $L^{p,p}$ indicates the L^p to L^p induced operator norm).

From Eqn. 3 we have

$$\psi_{0,j,k}(t) = M^{j/2}\psi_0(M^j t - k) \tag{12}$$

and if we only consider signals $f(t) \in L^1_{loc}(\mathbb{R})$ the J^{th} level scaling expansion coefficients are obtained by the inner product in Eqn. 6 as

$$a_J(k) = \langle f(t), \psi_{0,J,k}(t)\rangle \tag{13}$$

which in the Fourier domain is equivalent to

$$\hat{a}_J\left(\frac{\omega}{M^J}\right) = M^{-J/2}\sum_k \hat{f}(\omega + 2\pi M^J k)\psi_0^*\left(\frac{\omega + 2\pi M^J k}{M^J}\right). \tag{14}$$

If we let $Pf(t)$ denote the projection of $f(t)$ onto $W_{0,J}$, the span of $\{\psi_{0,J,k}\}$ the best L^2 approximation of f in $W_{0,J}$ is

$$Pf(t) = \sum_k a_J(k)\psi_{0,J,k}(t). \tag{15}$$

Equivalently in the Fourier domain

$$\widehat{Pf}(\omega) = M^{J/2} \hat{a}_J\left(\frac{\omega}{M^J}\right) \hat{\psi}_0\left(\frac{\omega}{M^J}\right). \qquad (16)$$

If we denote by $Qf(t)$ the approximation error, then $Pf \perp Qf$ and we have

$$Qf(t) = f(t) - \sum_k a_J(k)\psi_{0,J,k}(t) \qquad (17)$$

or equivalently in the transform domain using Eqn. 14

$$\widehat{Qf}(\omega) = \hat{f}(\omega)\left(1 - \left|\hat{\psi}_0\left(\frac{\omega}{M^J}\right)\right|^2\right) - \hat{\psi}_0\left(\frac{\omega}{M^J}\right) \times \sum_{k\neq 0} \hat{f}(\omega + 2\pi M^J k)\hat{\psi}_0^*\left(\frac{\omega + 2\pi M^J k}{M^J}\right). \qquad (18)$$

Hence, for a fixed J, the approximation error, which by a closer look is related to the maximum ε as defined in Definition 3, depends on the choice of ψ_0. For multiplicity M wavelets tight frames of compact support ψ_0 is uniquely determined by a finite length CQF filter h_0.

Hence, for a given $f(t)$, fixed J and possibly a fixed N an important problem is to find ψ_0 such that Qf is minimized. A directly related problem is to find the ψ_0 such that the induced operator norm on Qf is minimized assuming appropriate norms on the input space. In particular we will show how we can obtain an upper bound on Qf when f is a bandlimited signal.

4.1. L^p to L^p induced operator norm

Consider the L^p norm of $\widehat{Qf}$ which is easily obtained from Eqn. 18. To simplify notation define

$$\hat{g}(\omega) = \left\{ \left|\left(1 - \left|\hat{\psi}_0\left(\frac{\omega}{M^J}\right)\right|^2\right)\right|^p + \left|\hat{\psi}_0\left(\frac{\omega}{M^J}\right)\right|^p \sum_{k\neq 0}\left|\hat{\psi}_0^*\left(\frac{\omega + 2\pi M^J k}{M^J}\right)\right|^p \right\}. \qquad (19)$$

Assuming that f is bandlimited to $M^{J-1}\pi$ the L^p norm of $\widehat{Qf}$ is given by

$$\begin{aligned}\left\|\widehat{Qf}\right\|_p^p &= \int_{\mathbb{R}} \left|\widehat{Qf}(\omega)\right|^p d\omega \\ &= \frac{1}{2\pi}\int_{-M^{J-1}\pi}^{M^{J-1}\pi} \left|\hat{f}(\omega)\right|^p \hat{g}(\omega) d\omega. \qquad (20)\end{aligned}$$

Implementing Eqn. 20, using a nonlinear numerical optimization routine, an optimal wavelet can be obtained that minimize the p^{th} norm of the Fourier domain error for a particular input signal bandlimited to $M^{J-1}\pi$. Moreover, if we in Eqn. 20 pull out the sup of $|\hat{g}|^p$ we get

$$\left\|\widehat{Qf}\right\|_p^p \leq \left\|\hat{f}(\omega)\right\|_p^p \sup_{0\leq\omega\leq M^{J-1}\pi} \hat{g}(\omega). \qquad (21)$$

Then by Eqn. 21 we can numerically solve for a scaling function ψ_0 which is optimal, with respect to the p^{th} norm of the error, for a class of signals bandlimited to $M^{J-1}\pi$. In particular from Eqn. 18 with $p = 2$ using the identity

$$\sum_k \left|\hat{\psi}_0(\omega + 2\pi k)\right|^2 = 1 \qquad (22)$$

we have

$$\left\|\widehat{Qf}\right\|_2^2 = \frac{1}{2\pi}\int_{-M^{J-1}\pi}^{M^{J-1}\pi} \left|\hat{f}(\omega)\right|^2 \left(1 - \left|\hat{\psi}_0\left(\frac{\omega}{M^J}\right)\right|^2\right) d\omega.$$

For the *sinc* wavelet we see that the error is identically zero. Hence, we would expect the optimal wavelet to approximate the *sinc* wavelet.

Equivalent equalities and inequalities can be derived for $\widehat{Pf}$ (Gopinath [8]) and in particular for $p = 2$

$$\left\|\widehat{Pf}\right\|_2^2 = \frac{1}{2\pi}\int_{-M^{J-1}\pi}^{M^{J-1}\pi} \left|\hat{f}(\omega)\right|^2 \left|\hat{\psi}_0\left(\frac{\omega}{M^J}\right)\right|^2 d\omega \qquad (23)$$

and we notice as expected that

$$\left\|\widehat{Pf}\right\|_2^2 + \left\|\widehat{Qf}\right\|_2^2 = \left\|\hat{f}\right\|_2^2.$$

Hence, for each J there exist a class of bandlimited signals $f(t)$ such that the approximation error $\widehat{Qf}$ is bounded above. As a result we trivially observe that bandlimited signals fall in to the class of functions defined in Definition 3. Furthermore, if f is an impulse train we have form Eqn. 18 that

$$\begin{aligned}\widehat{Qf}(\omega) &= \hat{f}(\omega)\Bigg[\left(1 - \left|\hat{\psi}_0\left(\frac{\omega}{M^J}\right)\right|^2\right) \\ &- \hat{\psi}_0\left(\frac{\omega}{M^J}\right)\sum_{k\neq 0}\hat{\psi}_0^*\left(\frac{\omega + 2\pi M^J k}{M^J}\right)\Bigg]. \qquad (24)\end{aligned}$$

Now using this in Eqn. 20 we have

$$\left\|\widehat{Qf}\right\|_p^p = \frac{1}{2\pi}\int_{\mathbb{R}} \left|\hat{f}(\omega)\right|^p \hat{g}(\omega) d\omega. \qquad (25)$$

and trivially all impulse trains are included by the definition of essentially scale limited signals.

Tewfik and Jorgensen [7] provided upper bounds for design of optimal wavelets for representation of a signal. We can do better than that, and in the special case of $L^{2,2}(\mathbb{R})$ an explicit formula can be obtained for the error. This is suited for numerical optimization to obtain the best wavelet for representing a given signal. Furthermore, for the case of $L^{p,p}(\mathbb{R})$, bounds are derived for design of robust wavelets.

Another interesting consequence of the analysis is the following *wavelet sampling theorem.* Shannon's sampling theorem states that signals bandlimited to $[0, M^{J-1}\pi)$ are uniquely determined by the samples $f(M^{J-1}k)$. From Eqn. 23 we can obtain the following interesting theorem:

Theorem 1 *Let $\hat{f}(\omega) = 0$ for $|\omega| \geq M^{J-1}\pi$. Then, $f(t)$ is uniquely determined by the scaling expansion coefficients at level J. That is, there exists a function $g_{\psi_0}(t)$ such that,*

$$f(t) = \sum_k a_J(k) g_{\psi_0}(t - k) \qquad (26)$$

Proof: See [8] □

Notice that, the theorem states that for a bandlimited signal, knowing Pf (equivalent to knowing $a_J(k)$), which is *not* bandlimited, is adequate.

4.2. L^p to L^1 induced operator norm

If required we could also look at the L^1 norm of $\widehat{Qf}$. So consider again Eqn. 18 and we have the frequency domain L^1 norm for f bandlimited to $M^{J-1}\pi$ given by

$$\begin{aligned}\left\|\widehat{Qf}\right\|_1 &= \int_{\mathbb{R}} \left|\widehat{Qf}(\omega)\right| d\omega \\ &= \int_{-M^{J-1}\pi}^{M^{J-1}\pi} \left|\hat{f}(\omega)\right| \hat{g}(\omega) d\omega \qquad (27)\end{aligned}$$

where $\hat{g}$ is defined as in Eqn. 19 with $p = 1$. By Hölders inequality we have

$$\left\|\widehat{Qf}\right\|_1 \leq \left\|\hat{f}\right\|_p \|\hat{g}\|_q . \qquad (28)$$

From Eqns. 27 and 28 we can as mentioned above numerically solve for the optimal and robust wavelets expansion, but in this case with respect to the L^1 norm of the transform domain error.

5. CONCLUSION

We have shown that the most obvious way of defining scale limited leads to a class of trivial signal of no interest. However, by considering signals that are essentially scale limited we are able to include signals such a impulse trains and bandlimited hence making it a useful definition form a practical point of view. Reconstruction of signals that fall in to this class has the property that they are guaranteed a predetermined "small" upper bound on the error. We also obtain a form of sampling theorem for wavelet analysis that we call the *wavelet sampling theorem.*

Moreover, we have developed a framework for numerically designing optimal and robust wavelets for signal expansion of essentially scale limited signals. In the case of $L^{2,2}(\mathbb{R})$ induced operator norm explicit bounds are given while in the more general case of $L^{p,p}$ and $L^{p,1}$ upper bounds are given.

Acknowledgments

This work was supported by AFOSR under grant number 90-0334 which was funded by DARPA.

REFERENCES

[1] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure Applied Math.*, 41:909–996, November 1988.

[2] R. A. Gopinath and C. S. Burrus. Wavelets and filterbanks. Technical Report CML TR91-20, Computational Mathematics Laboratory, Rice University, Houston, TX-77251, 1991. To appear in: "Wavelets - A Tutorial in theory and Applications" edited by C. Chui, published by Academic Press 1992.

[3] S. G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(7):674–693, July 1989.

[4] S. G. Mallat. Zero-crossings of a wavelet transform. *IEEE Trans. Inform. Theory*, 37(4):1019–1033, July 1991.

[5] R. A. Gopinath. The wavelet transforms and time-scale analysis of signals. Master's thesis, Rice University, Houston, TX-77251, 1990.

[6] R. Gopinath, J. E. Odegard, and C. S. Burrus. On the correlation structure of the scaling function and wavelets. Technical Report CML TR91-19, Computational Mathematics Laboratory, Rice University, Houston, TX-77251, 1991.

[7] A. H. Tewfik and P. E. Jorgensen. On the optimal choice of a wavelet for signal representation. Preprint, 1991.

[8] R. A. Gopinath and C. S. Burrus. On the optimal and robust wavelet representation of signals and the wavelet sampling theorem. Technical Report CML TR91-18, Computational Mathematics Laboratory, Rice University, Houston, TX-77251, 1991.