Wim Sweldens

Wavelets and the Lifting Scheme: A 5 Minute Tour

In this paper, we give a brief introductory tour to the lifting scheme, an new method to construct wavelets. We show its advantages over classical constructions and give pointers to the literature.

1. Introduction

The purpose of this paper is to give a very short, "nuts and bolts", introduction to the *lifting scheme* and provide references for further reading. The lifting scheme is a new method for constructing biorthogonal wavelets. The main difference with classical constructions [1-3] is that it does not rely on the Fourier transform. This way lifting can be used to construct *second generation wavelets*, wavelets which are not necessarily translates and dilates of one function. The latter we refer to as *first generation wavelets*.

In the case of first generation wavelets, the lifting scheme will never come up with wavelets which somehow could not be found by the Cohen-Daubechies-Feauveau machinery [2]. Nevertheless, it has the following advantages:

1. It allows a faster implementation of the wavelet transform. Traditionally, the fast wavelet transform is calculated with a two-band subband transform scheme, see Figure 1. In each step the signal is split into a high pass and low pass band and then subsampled. Recursion occurs on the low pass band. The lifting scheme makes optimal use of similarities between the high and low pass filters to speed up the calculation. The number of flops can be reduced by a factor of two.
2. The lifting scheme allows a fully in-place calculation of the wavelet transform. In other words, no auxiliary memory is needed and the original signal (image) can be replaced with its wavelet transform.
3. In the classical case, it is not immediately clear that the inverse wavelet transform actually is the inverse of the forward transform. Only with the Fourier transform one can convince oneself of the perfect reconstruction property. With the lifting scheme, the inverse wavelet transform can immediately be found by undoing the operations of the forward transform. In practise, this comes down to simply changing each $+$ into a $-$ and vice versa.
4. The lifting scheme is a very natural way to introduce wavelets in a classroom. Indeed, since it does not rely on the Fourier transform, the properties of the wavelets and the wavelet transform do not appear as somehow "magical" to students without a strong Fourier background. In fact, this paper originated from typing up lecture notes.

Secondly, the lifting scheme can be used in situations where no Fourier transform is available. Typical examples are:

1. *Wavelets on bounded domains*: The construction of wavelets on general possibly non-smooth domains is needed in applications such as data segmentation and the solution of partial differential equations. A special case is the construction of wavelets on an interval, which is needed to transform finite length signals without introducing artifacts at the boundaries.
2. *Wavelets on curves and surfaces*: To analyze data that live on curves or surfaces or to solve equations on curves or surfaces, one needs wavelets intrinsically defined on these manifolds, independent of parametrization.
3. *Weighted wavelets*: Diagonalization of differential operators and weighted approximation require a basis adapted to weighted measures. Wavelets biorthogonal with respect to a weighted inner product are needed.
4. *Wavelets and irregular sampling*: Many real life problems require basis functions and transforms adapted to irregularly sampled data.

It is obvious that wavelets adapted to these setting cannot be formed by translation and dilation. The Fourier transform can thus no longer be used as a construction tool. The lifting scheme provides an alternative.

The basic idea behind the lifting scheme is very simple. It starts with a trivial wavelet, the "Lazy wavelet"; a function which essentially doesn't do a thing, but which has the formal properties of a wavelet. The lifting scheme then gradually builds a new wavelet, with improved properties, by adding in new basis functions. This is the

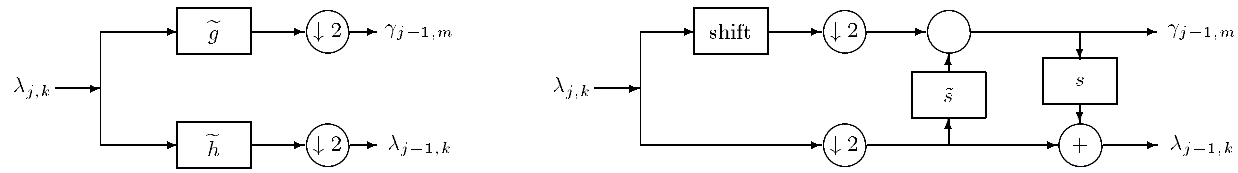

Figure 1: *The fast wavelet transform: classical implementation (left) and lifting scheme (right). The lifting scheme first does a Lazy wavelet transform, then calculates the $\gamma_{j-1,m}$, and finally lifts the $\lambda_{j-1,k}$.*

inspiration behind the name "lifting scheme". Instead of diving into the theory and firing off a sequence of formulas, we give a simple example which illustrates the above mentioned properties.

2. A simple example

Suppose we sample a signal $f(t)$ with sampling distance 1. We denote the original samples as $\lambda_{0,k} = f(k)$ for $k \in \mathbb{Z}$. We would like to "decorrelate" this signal. In other words, we would like to see if it is possible to capture the information contained in this signal with fewer coefficients, i.e. coefficients with a larger sampling distance. A more compact representation is needed in applications such as data compression. Maybe it will not be possible to *exactly* represent the signal with fewer coefficients but instead find an *approximation* within an acceptable error bound. We thus want to have precise control over the information which is lost by using fewer coefficients. Obviously, we would like the difference between the original and approximated signal to be small.

We can reduce the number of coefficients, by simply subsampling the even samples. This way we obtain a new sequence given by

$$\lambda_{-1,k} := \lambda_{0,2k} \quad \text{for} \quad k \in \mathbb{Z}\,. \tag{1}$$

We would like to have an idea on how much information was lost. In other words, which extra information (if any) do we need to recover the original $\{\lambda_{0,k}\}$ from the $\{\lambda_{-1,k}\}$. We will use coefficients $\{\gamma_{-1,k}\}$ to encode this difference and refer to them as *wavelet coefficients*. Many different choices are possible, and depending on the statistical behavior of the signal, one will be better then the other. Better means smaller wavelet coefficients.

The most naive, trivial choice would be to say that the lost information is simply contained in the odd coefficients, $\gamma_{-1,k} := \lambda_{0,2k+1}$ for $k \in \mathbb{Z}$. This choice corresponds to the Lazy wavelet. Indeed, we haven't done very much except for subsampling the signal in even and odd samples. Obviously, this will not really decorrelate the signal. The wavelet coefficients are only small in case the odd samples are small and there is no reason whatsoever why this should be the case. But, bear with us, and you will see why the Lazy wavelet is useful.

Let us try to find a more elaborate scheme to recover the original samples $\{\lambda_{0,k}\}$ from the subsampled coefficients $\{\lambda_{-1,k}\}$. The even samples $\{\lambda_{0,2k}\}$ can immediately be found as $\lambda_{0,2k} := \lambda_{-1,k}$. But maybe we can also find, or at least predict, the odd samples based on the $\{\lambda_{-1,k}\}$. Assuming some correlation amongst neighboring samples, we suggest to predict an odd sample $\lambda_{0,2k+1}$ as the average of its two (even) neighbors: $\lambda_{-1,k}$ and $\lambda_{-1,k+1}$. This need not be exact, the wavelet coefficient therefore encodes the difference between the exact sample and its predicted value:

$$\gamma_{-1,k} := \lambda_{0,2k+1} - 1/2\,(\lambda_{-1,k} + \lambda_{-1,k+1})\,. \tag{2}$$

If the signal is somehow correlated, the majority of the wavelet coefficients is small. The ones that fall below the error threshold can simply be ignored. This way one obtains more compact representations. If the original signal is piecewise linear between the even samples, all wavelet coefficients are zero.

We are quite happy with these wavelet coefficients. They encode the detail needed to go from the $\{\lambda_{-1,k}\}$ coefficients to the $\{\lambda_{0,k}\}$. Essentially they measure to which extend the original signal *fails to be linear*. Their expected value is small. In terms of frequency content, the wavelet coefficients capture the high frequencies present in the original signal while the $\{\lambda_{-1,k}\}$ somehow capture the low frequencies. In principle we could now simply iterate this scheme. Starting from the $\{\lambda_{-1,k}\}$, we could find coarser level coefficients $\{\lambda_{-2,k}\}$ (by subsampling) and $\{\gamma_{-2,k}\}$ (as the failure to be linear).

However, we are not very pleased with the choice of the $\{\lambda_{-1,k}\}$. The reason is the following. Suppose we are given $2^n + 1$ original samples $\{\lambda_{0,k} \mid 0 \le k \le 2^n\}$. We could apply our scheme n times thus obtaining $\{\gamma_{j,k} \mid -n \le j \le -1, 0 \le k < 2^{n+j}\}$ and two (coarsest level) coefficients $\lambda_{-n,0}$ and $\lambda_{-n,1}$. These are the first ($\lambda_{-n,0} = \lambda_{0,0}$) and the last ($\lambda_{-n,1} = \lambda_{0,2^n}$) original sample. This introduces considerable aliasing. We would like the

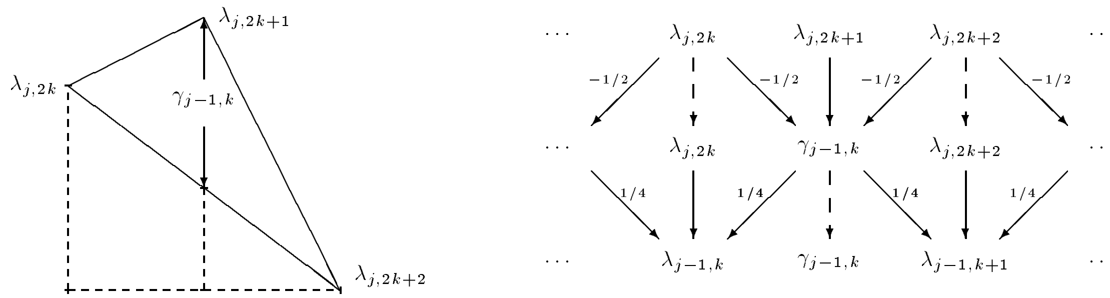

Figure 2: *Left: the wavelet coefficient is the failure to be linear. Right: The lifting scheme: First calculate the wavelet coefficients* $\gamma_{j-1,m}$ *and then use them to lift the* $\lambda_{j-1,k}$.

average of the $\lambda_{j,k}$ coefficients to be the same on each level, or $\sum_k \lambda_{-1,k} = 1/2 \sum_k \lambda_{0,k}$. The current subsampling obviously does not have this property. We will obtain it by a second step: lifting the $\lambda_{-1,k}$ *with the help of* the wavelet coefficients $\gamma_{-1,k}$. Again we use the neighboring wavelet coefficients. A moments thought yields that the correct choice is:

$$\lambda_{-1,k} \mathrel{+}= 1/4\,(\gamma_{-1,k-1} + \gamma_{-1,k})\,. \tag{3}$$

The wavelet transform on each level now consists of two stages: first calculate the wavelet coefficients as the failure to be linear (2), secondly lift the subsampled coefficients (1) with the help of these wavelet coefficients (3). See Figure 2 for a scheme. The inverse transform can immediately be found: first undo the lifting ($-=$ instead of $+=$in (3)), then add in the failure to be linear with the linear prediction: $\lambda_{0,2k+1} := \gamma_{-1,k} + 1/2\,(\lambda_{-1,k} + \lambda_{-1,k+1})$.

In-place calculation can be done as follows. Assume the original samples are stored in a vector $v[k]$. Each coefficient $\lambda_{j,k}$ or $\gamma_{j,k}$ is stored in location $v[2^{-j}\,k]$. The Lazy wavelet transform is then immediate. All other operations can be done with $+=$ or $-=$ operations, see Figure 2.

3. Formal description

The algorithm of the previous section can be casted into a formal description with the use of basis functions. Let Λ be the classical "Hat" function: $\Lambda(x) = \max\{0, 1-|x|\}$. Define the piecewise linear approximations on each level as

$$P_j(x) = \sum_k \lambda_{j,k}\,\Lambda_{j,k}(x) \quad \text{with} \quad \Lambda_{j,k}(x) = \Lambda(2^j x - k)\,.$$

The difference between two approximations can now be encoded with the help of the wavelets,

$$P_j(x) - P_{j-1}(x) = \sum_k \gamma_{j,k}\,\psi_{j,k}(x) \quad \text{with} \quad \psi_{j,k}(x) = \psi(2^j x - k)\,.$$

In case the $\lambda_{j,k}$ coefficients are not lifted, the correct choice for the *wavelet* is: $\psi(x) = \Lambda(2x-1)$. With lifting, the wavelet is given by

$$\psi(x) = \Lambda(2x-1) - 1/4\,\Lambda(x) - 1/4\,\Lambda(x+1)\,. \tag{4}$$

This way of constructing a wavelet is typical for the lifting scheme. A new wavelet $\psi(x)$ is found as an old wavelet $\Lambda(2x-1)$ combined with two Hat functions on the *same* level, $\Lambda(x)$ and $\Lambda(x+1)$. This opposed to the classical case were a wavelet is constructed as a linear combination of Hat functions on the next *finer* level, namely $\Lambda(2x-k)$ with $k \in \mathbb{Z}$. We can now write the original signal as

$$P_0(x) = P_{-n}(x) + \sum_{j=-n}^{-1} \sum_k \gamma_{j,k}\,\psi_{j,k}\,.$$

The $\gamma_{j,k}$ are calculated with the algorithm described in the previous section. An approximation using fewer coefficients can now be obtained by simply omitting the wavelet coefficients below a certain threshold. The wavelet (4) is constructed so that its integral is zero and thus $\int_{\mathbb{R}} P_0(x)\,dx = \int_{\mathbb{R}} P_j(x)\,dx$.

For those familiar with the work in [2,3], the wavelet presented here is the $(2,2)$ biorthogonal wavelet. In the classical case the $\{\lambda_{-1,k}\}$ coefficients are found as the convolution of the $\{\lambda_{0,k}\}$ coefficients with the filter $\widetilde{h} = \{-1/8, 1/4, 3/4, 1/4, -1/8\}$. This step would take 6 flops per coefficients while lifting only needs 3.

4. Generalizations and pointers for further reading

Many generalizations of the wavelet from the previous section are possible. We mention here a few. First of all one can use a higher order scheme to predict the odd indexed values from the even. For example, one can predict $\lambda_{j,2k+1}$ based on the cubic polynomial through the values $\lambda_{j,2k-2}, \lambda_{j,2k}, \lambda_{j,2k+2}, \lambda_{j,2k+4}$. In this case the wavelets coefficients are all zero in case the original signal is cubic. Also higher order (N) polynomial interpolation is possible. These predictions can be computed quickly with the Neville algorithm. Secondly, one can use fancier schemes for the lifting of the $\lambda_{j,k}$ coefficients. For example one can assure that not only the integral but the first $\widetilde{N}$ moments of the approximations are preserved: $\int_{\mathbb{R}} x^p P_j(x)\, dx = \int_{\mathbb{R}} x^p P_{j-1}(x)\, dx$ for $0 \leq p < \widetilde{N}$. Consequently the first $\widetilde{N}$ moments of the wavelet are zero. The algorithm above has $\widetilde{N} = N = 2$.

The lifting scheme also allows to adapt the transform in a natural way to finite length signals (intervals). For example, the predicted value now is based on the linear interpolation of the $\lambda_{j,2k}$ values left and right. In case no value on the left exist, one can take two values on the right and use linear extrapolation. Same for higher order schemes and also for the lifting of the $\lambda_{j,k}$. A similar philosophy can be used to adapt wavelets to domains, weighted measures, curves and surfaces, and irregular samples, but this is beyond the scope of this introductory tour.

The theory behind the construction of first and second generation wavelets with the help of lifting is described respectively in [7] and [8]. Wavelets on a sphere are constructed with the lifting scheme in [5]. These spherical wavelets are used for spherical image processing in [6]. In [4], the generalization of Haar wavelets to the second generation case is presented. They form a perfect example to start lifting. A "nuts and bolts" introduction to the construction of second generation wavelets is the topic of [9]. There concrete examples concerning interval constructions, irregular samples, and weighted measures are given. The technical reports from South Carolina are available as Postscript files through anonymous ftp to `ftp.math.sc.edu`, in the directories `/pub/imi_94` and `/pub/imi_95`.

Acknowledgements

The author is Senior Research Assistant of the National Fund of Scientific Research Belgium (NFWO). Part of this work was done at the University of South Carolina with support from NSF EPSCoR Grant EHR 9108772 and DARPA Grant AFOSR F49620-93-1-0083. Special thanks to Peter Schröder for many inspiring and stimulating discussions and to Geert Uytterhoeven for carefully reading the manuscript.

5. References

1 Chui, C. K.: An Introduction to Wavelets. Academic Press (1992).

2 Cohen, A.,Daubechies, I., Feauveau, J.: Bi-orthogonal bases of compactly supported wavelets; Comm. Pure Appl. Math., **45** (1992), 485–560.

3 Daubechies, I.:Ten Lectures on Wavelets. SIAM (1992).

4 Girardi, M., Sweldens, W.: A new class of unbalanced Haar wavelets that form an unconditional basis for L_p on general measure spaces. Technical Report 1995:2, Industrial Mathematics Initiative, Department of Mathematics, University of South Carolina (1995)

5 Schröder, P., Sweldens, W.: Spherical wavelets: Efficiently representing functions on the sphere; Computer Graphics, (SIGGRAPH '95 Proceedings) (1995).

6 Schröder, P., Sweldens, W.: Spherical wavelets: Texture processing. Technical Report 1995:4, Industrial Mathematics Initiative, Department of Mathematics, University of South Carolina (1995).

7 Sweldens, W.: The lifting scheme: A custom-design construction of biorthogonal wavelets. Technical Report 1994:7, Industrial Mathematics Initiative, Department of Mathematics, University of South Carolina (1994).

8 Sweldens, W.: The lifting scheme: A construction of second generation wavelets. Technical Report 1995:6, Industrial Mathematics Initiative, Department of Mathematics, University of South Carolina (1995).

9 Sweldens, W., Schröder, P.: Building your own wavelets at home. Technical Report 1995:5, Industrial Mathematics Initiative, Department of Mathematics, University of South Carolina (1995).

Address: Wim Sweldens, Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, B 3001 Heverlee, Belgium, `wim.sweldens@cs.kuleuven.ac.be`.