

SEISMIC IMAGING USING COMPLEX WAVELETS

Mark Miller, Nick Kingsbury

Signal Processing Group
Dept. of Engineering, University of Cambridge, UK
Email: mam59@cam.ac.uk, ngk@eng.cam.ac.uk

Richard Hobbs

Dept. of Earth Sciences
University of Durham, UK
Email: r.w.hobbs@durham.ac.uk

ABSTRACT

The aim of seismic imaging is to reconstruct subsurface reflectivity from scattered acoustic data. In standard reconstruction techniques, the reflectivity model parameters are usually defined as a grid of point scatterers over the area or volume of the subsurface to be imaged. We propose an approach to subsurface imaging using the Dual Tree Complex Wavelet Transform (DT-CWT) as a basis for the reflectivity. This basis is used in conjunction with an iterative optimization which frames the problem as a linearized inverse scattering problem. We demonstrate the method on a marine seismic data set acquired over the Gippsland Basin near Australia. The technique is shown to reduce noise and processing artifacts while preserving discontinuities. It is likely to be particularly useful in cases where the acquired data is incomplete.

1. INTRODUCTION

The objective of marine seismic imaging is to reconstruct subsurface reflectivity from scattered acoustic data generally observed near the ocean surface. This can be achieved by iteratively minimizing a cost function consisting of a data matching term and a regularization term. In seismic processing literature the procedure is usually termed least-squares migration (LSM).

LSM has been shown to be effective in optimizing the reconstruction of subsurface reflectivity, particularly in cases of incomplete data [1], [2]. There are several cases where a data set may be incomplete: finite recording apertures, coarse source/receiver distribution, gaps between recording lines for 3D imaging and poor subsurface illumination caused by irregular ray coverage due to strong lateral variations in velocity.

Choosing a basis that is appropriate for the statistical characteristics of a particular class of input signal is a classical problem. Linear transforms have proven to be popular due to their simplicity and mathematical tractability. Reflectivity is usually defined on a grid of point reflectors. We propose using a complex wavelet basis for the reflectivity. The Q-shift version of the Dual Tree Complex Wavelet Transform (DT-CWT) is chosen as the wavelet basis for its key advantages compared to other wavelet transforms [3]. These are summarized as follows:

Shift invariance: Eliminates aliasing defects.

Basis vectors form an almost tight frame: Allows use of the transform in the efficient Conjugate Gradient Descent (CGD) algorithm.

Good directional selectivity: The DT-CWT has six directionally selective subbands in 2 dimensions and 28 in 3 dimensions.

Limited redundancy: Redundancy is independent of the number of scales and limited to 2^n for n dimensions.

Efficient computation: Filters are separable and computation is less than 2^n times that of the simple DWT for n dimensions.

Perfect reconstruction.

Wavelet bases tend to decorrelate or diagonalize a range of non-stationary signals. This has led to extensive use of wavelet bases in the area of information coding and compression. In LSM, diagonalization of the model space allows a more accurate and practical representation of prior information about the model parameters. This prior information is incorporated in the regularization term of the cost function that is minimized.

Using more sophisticated regularization becomes more important for missing or undersampled data problems or when increased resolution is required in the reflectivity model to be reconstructed. Alternatively, relaxing sampling requirements can reduce the cost of data acquisition.

Wavelet bases have been successfully applied to other linearized inverse problems [4] including linear inversion with application to well logging [5], [6]. Herrmann employed curvelet/contourlet transforms for seismic imaging similar to that described here but took a minimax-style approach [7]. The use of a complex wavelet transform with similar properties to the DT-CWT in seismic processing was suggested in [8]. De Rivaz and Kingsbury used the DT-CWT in a similar manner to here for image restoration in [9].

2. THEORY

Using the high-frequency single-scattering Born approximation the seismic data can be modeled using a generalized Radon transform:

$$\mathbf{d} = \mathbf{L}\mathbf{m} + \mathbf{n} \quad (1)$$

In (1) \mathbf{d} is the observed scattered acoustic data which is a function of time and acoustic source and receiver positions. \mathbf{m} is the reflectivity on a regular grid in 2 or 3 dimensions and \mathbf{L} is the linearized seismic modeling operator mapping from one to the other [1]. Observed data not modeled by the linear operator is modeled by additive noise \mathbf{n} assumed to be Gaussian distributed with covariance \mathbf{C}_n . Thus, (1) is an inverse problem where the objective is to recover the reflectivity \mathbf{m} from the observed data \mathbf{d} .

Expressing the reflectivity using wavelet basis functions by substituting $\mathbf{m} = \mathbf{P}\mathbf{w}$ in (1) results in equation (2).

$$\mathbf{d} = \mathbf{L}\mathbf{P}\mathbf{w} + \mathbf{n} \quad (2)$$

In (2) \mathbf{w} are the lexicographically ordered wavelet coefficients whose real and imaginary parts are treated as separate variables. Starting with equation (2), we derive in a Bayesian framework a cost function which when minimized provides a *maximum a posteriori* (MAP) estimate for the wavelet coefficients \mathbf{w} . The conditional probability for the data given the wavelet coefficients is:

$$p_{\mathbf{d}|\mathbf{w}}(\mathbf{d}|\mathbf{w}) = \frac{1}{(2\pi|\mathbf{C}_n|)^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{LP}\mathbf{w} - \mathbf{d})^T \mathbf{C}_n^{-1} (\mathbf{LP}\mathbf{w} - \mathbf{d}) \right\} \quad (3)$$

The MAP estimate for the coefficients is presented in equation (4) and manipulated into a more convenient form using Bayes theorem.

$$\begin{aligned} \mathbf{w}_{MAP} &= \arg \max_{\mathbf{w}} \{ p_{\mathbf{w}|\mathbf{d}}(\mathbf{w}|\mathbf{d}) \} \\ &= \arg \max_{\mathbf{w}} \{ p_{\mathbf{d}|\mathbf{w}}(\mathbf{d}|\mathbf{w}) p_{\mathbf{w}}(\mathbf{w}) \} \\ &= \arg \min_{\mathbf{w}} \{ -\log (p_{\mathbf{d}|\mathbf{w}}(\mathbf{d}|\mathbf{w}) p_{\mathbf{w}}(\mathbf{w})) \} \\ &= \arg \min_{\mathbf{w}} \{ (\mathbf{LP}\mathbf{w} - \mathbf{d})^T \mathbf{C}_n^{-1} (\mathbf{LP}\mathbf{w} - \mathbf{d}) + f(\mathbf{w}) \} \end{aligned} \quad (4)$$

where $f(\mathbf{w}) = -\log (p_{\mathbf{w}}(\mathbf{w}))$ is the regularization term. Our prior model for the wavelet coefficients assumes the real and imaginary parts have independent Gaussian distributions with zero mean and known variance. Thus $f(\mathbf{w})$ becomes:

$$f(\mathbf{w}) = \mathbf{w}^T \mathbf{C}_w^{-1} \mathbf{w} \quad (5)$$

and we obtain an estimate for the wavelet coefficients by minimizing:

$$\mathbf{E}_w(\mathbf{w}) = (\mathbf{LP}\mathbf{w} - \mathbf{d})^T \mathbf{C}_n^{-1} (\mathbf{LP}\mathbf{w} - \mathbf{d}) + \mathbf{w}^T \mathbf{C}_w^{-1} \mathbf{w} \quad (6)$$

We compare this to the case where a wavelet basis is not used. The relevant cost function is:

$$\mathbf{E}_m(\mathbf{m}) = (\mathbf{Lm} - \mathbf{d})^T \mathbf{C}_n^{-1} (\mathbf{Lm} - \mathbf{d}) + f(\mathbf{m}) \quad (7)$$

but a representative $f(\mathbf{m})$ is now more difficult to define.

Note that traditional migration in its most basic form is equivalent to estimating the reflectivity by applying the transpose of the forward modeling operator to the observed data. This is equivalent to one iteration of a conjugate gradient or steepest descent minimization procedure like that described here.

3. ALGORITHM

We now turn to the implementation of the algorithm. To speed up convergence and allow estimation of the coefficient variances, we start with a non-iterative estimate for the reflectivity. If the available estimate $\tilde{\mathbf{m}}_0$ is not scaled correctly, it is multiplied by λ that minimizes $\mathbf{E}_m(\lambda\tilde{\mathbf{m}}_0)$ with $f(\mathbf{m}) = 0$, so that $\mathbf{m}_0 = \lambda\tilde{\mathbf{m}}_0$. This is equivalent to performing one iteration of steepest descent minimization of (7) with $\tilde{\mathbf{m}}_0$ as the search direction and no regularization or preconditioning.

Following [9] the coefficient variances are allowed to vary within each subband but are the same for the real and imaginary parts of a given coefficient. These variances are estimated from the

forward DT-CWT of the initial reflectivity estimate $\mathbf{w}_0 = \mathbf{P}^T \mathbf{m}_0$ as $\sigma_i^2 = 0.5|w_i|^2$ where w_i is the corresponding complex wavelet coefficient in \mathbf{w}_0 .

The algorithm is initialized with \mathbf{w}_0 and the energy function (6) is minimized using a preconditioned CGD algorithm optimized to use the minimum number of forward and transpose seismic modeling operators per iteration (one of each), since these operations comprise the bulk of the computational effort - see appendix. A block diagram illustrating the algorithm implementation is displayed in figure 1. Figure 1(a) shows the preprocessing as detailed in this section, while figure 1(b) contains the optimized conjugate gradient loop, where the steepest descent direction is updated at each iteration rather than recalculated from scratch.

After sufficient iterations the reflectivity estimate is found by applying the inverse wavelet transform to the current coefficient estimate:

$$\hat{\mathbf{m}} = \mathbf{P}\hat{\mathbf{w}}_k \quad (8)$$

Note that the matrices \mathbf{P} and \mathbf{P}^T required in the conjugate gradient algorithm are not implemented as matrix multiplications but using much faster wavelet decomposition and reconstruction algorithms.

An area where the use of a wavelet basis makes the inversion more difficult is the preconditioning of the system. To speed up convergence, CGD algorithms used to solve linear inversion problems are preconditioned by scaling the variables prior to minimization. In its usual form the diagonal elements of the Hessian of the cost function $\nabla^2 \mathbf{E}$ need to be calculated, so that they can be scaled to unity by the preconditioning. For the non-wavelet cost function (7), ignoring the regularization term ($f(\mathbf{m}) = 0$), the Hessian is:

$$\nabla^2 \mathbf{E} = \mathbf{L}^T \mathbf{C}_n^{-1} \mathbf{L} \quad (9)$$

However, for wavelet based inversion this becomes:

$$\nabla^2 \mathbf{E} = \mathbf{P}^T \mathbf{L}^T \mathbf{C}_n^{-1} \mathbf{L} \mathbf{P} \quad (10)$$

Calculating the diagonal elements of (10) is more difficult. Fortunately, for preconditioning the exact values are not required and a rough estimate is sufficient. The computational effort required to obtain these values can be significantly reduced by using heavy interpolation of the wavelet subbands, taking advantage of the small support of most of the wavelet basis functions and by decimation of the trace locations for large data sets.

4. RESULTS AND CONCLUSIONS

The data used to generate the results presented here are from a 3-D survey acquired in 2001 over the offshore Gippsland Basin near Australia. A single shot/streamer line was selected from the data volume to test the complex wavelet least-squares migration algorithm in 2 dimensions. From this 14216 traces (lines of data at particular source/receiver locations) with source/receiver mid-points along a length of 2km above the area to be imaged are used. The modeling operator \mathbf{L} is implemented as a Kirchhoff demigration operator in 2.5 dimensions (3D geometrical spreading with reflectors homogeneous in one dimension) [10].

The noise covariance matrix is estimated directly from the trace data. It consists of two components. The first, which models background noise incoherent scatterers, was picked roughly by hand, decaying with time and slowly with source/receiver offset (the distance between the two). The second is proportional to a windowed average of the data values squared along each trace.

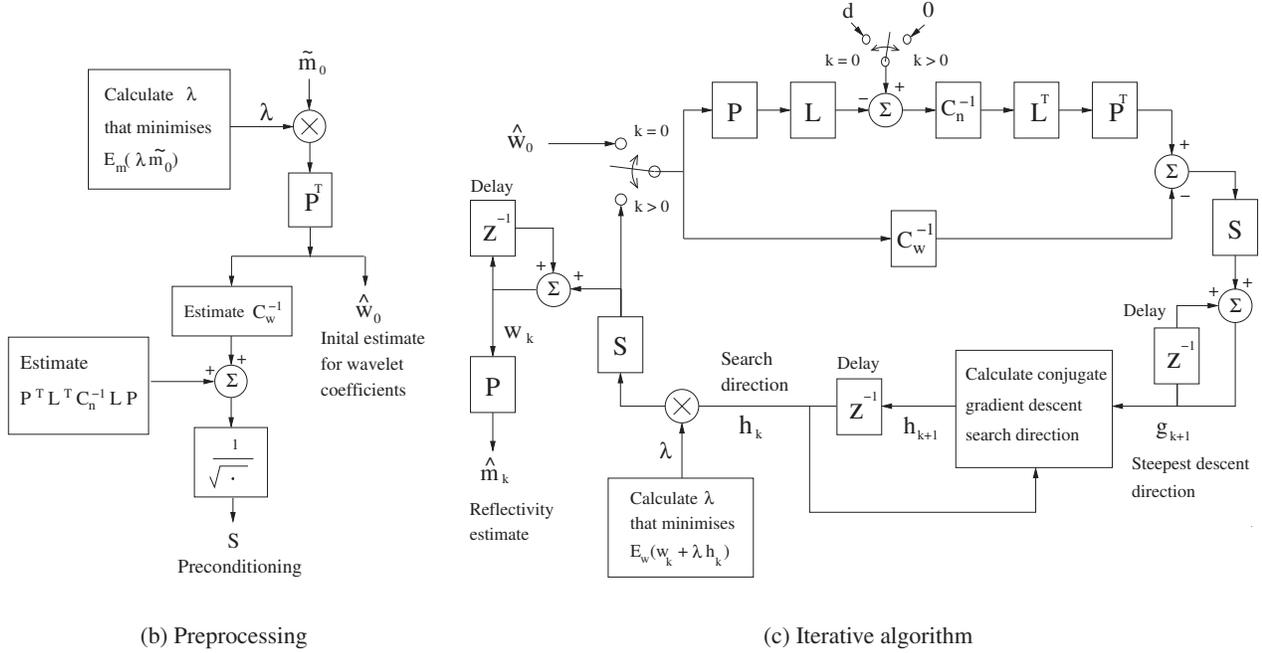


Fig. 1. Block diagram of the complex wavelet seismic imaging system

This component models the modeling error which is assumed to be roughly proportional to the data and the stronger reverberations near strong reflectors.

For our results we used the Kirchhoff based estimate $\tilde{\mathbf{m}}_0 = \mathbf{L}^T \mathbf{C}_n^{-1} \mathbf{d}$ to initialize the algorithm. We demonstrate the algorithm's ability to cope with undersampled data by using a reduced data set where the traces are decimated by a factor of 8. In the absence of any ground truth data, we compare the result to the Kirchhoff based estimate obtained using the full data set. We also compare the result to that obtained with non-wavelet least-squares migration using the cost function in (7) with the regularization defined as $f(\mathbf{m}) = \mathbf{m}^T \mathbf{C}_m^{-1} \mathbf{m}$, where $\mathbf{C}_m = \sigma_m^2 \mathbf{I}$ as suggested in [11]. σ_m^2 is the variance of the reflectivity model parameters, which we have estimated as the sample variance of the initial Kirchhoff estimate.

Figure 2 displays the results for an approximately 1250m wide by 500m deep target area. Figure 2(a) contains the Kirchhoff estimate for the decimated data used to initialize the complex wavelet system. The Kirchhoff reconstruction using the full data set in figure 2(c) suppresses much although not all of the noise seen in the reconstruction using the decimated data. The result for the complex wavelet system in figure 2(b) is seen to significantly suppress noise and reduce so called migration artifacts, such as that seen in figure 2(a) at a horizontal position of about 450m and at depth of about 1700m.

The non-wavelet iterative algorithm (figure 2(d)) is seen to somewhat reduce the migration artifacts of the Kirchhoff reconstruction but does not suppress noise as well as the complex wavelet system. The non-wavelet algorithm diverges in just a few iterations if no regularization is employed. The results for the complex wavelet system presented here are stable *i.e.* continuing for more iterations brings increasingly smaller updates to the estimate and the reflectivity estimate does not diverge.

5. ACKNOWLEDGEMENTS

The Authors acknowledge the support of ExxonMobil, who provided the data from the Gippsland Basin marine seismic survey. We are also grateful for financial support from the New Zealand Vice Chancellors Committee, Canterbury University, New Zealand and Trinity College, Cambridge, UK.

6. APPENDIX

The Seismic Modelling Operator \mathbf{L} and its Transpose \mathbf{L}^T

We briefly discuss the computational aspects of applying the operator \mathbf{L} and its transpose in the following types of calculation:

$$\mathbf{d} = \mathbf{L}\mathbf{m} \quad \text{and} \quad \hat{\mathbf{m}} = \mathbf{L}^T \mathbf{d}$$

Assuming that the size of the (2D) model \mathbf{m} is $N_x \times N_z$ and there are N_p recorded waveforms (traces) and N_t time samples per trace so that \mathbf{d} is of length $N_p N_t$, then direct multiplication by \mathbf{L} or \mathbf{L}^T would require approximately $N_x N_z N_p N_t$ operations. This is likely to be unacceptably large. Fortunately \mathbf{L} may be regarded as the cascade of a sparse matrix multiplication and a relatively efficient convolution process, summarized as follows:

a) Each point in the model \mathbf{m} excites each recorded waveform with a band-limited impulse of short duration, at a point determined by the total propagation delay from source to receiver via that point in \mathbf{m} . In applying \mathbf{L} these short pulses may be added to the traces in turn, and only affect a proportion α of the total traces in which the path delay lies within the range of delays being measured in \mathbf{d} . Interpolated pulse waveforms are used to synthesize fractional sample delays and the pulses are N_s samples long. Each pulse is scaled by an appropriate attenuation coefficient, dependent

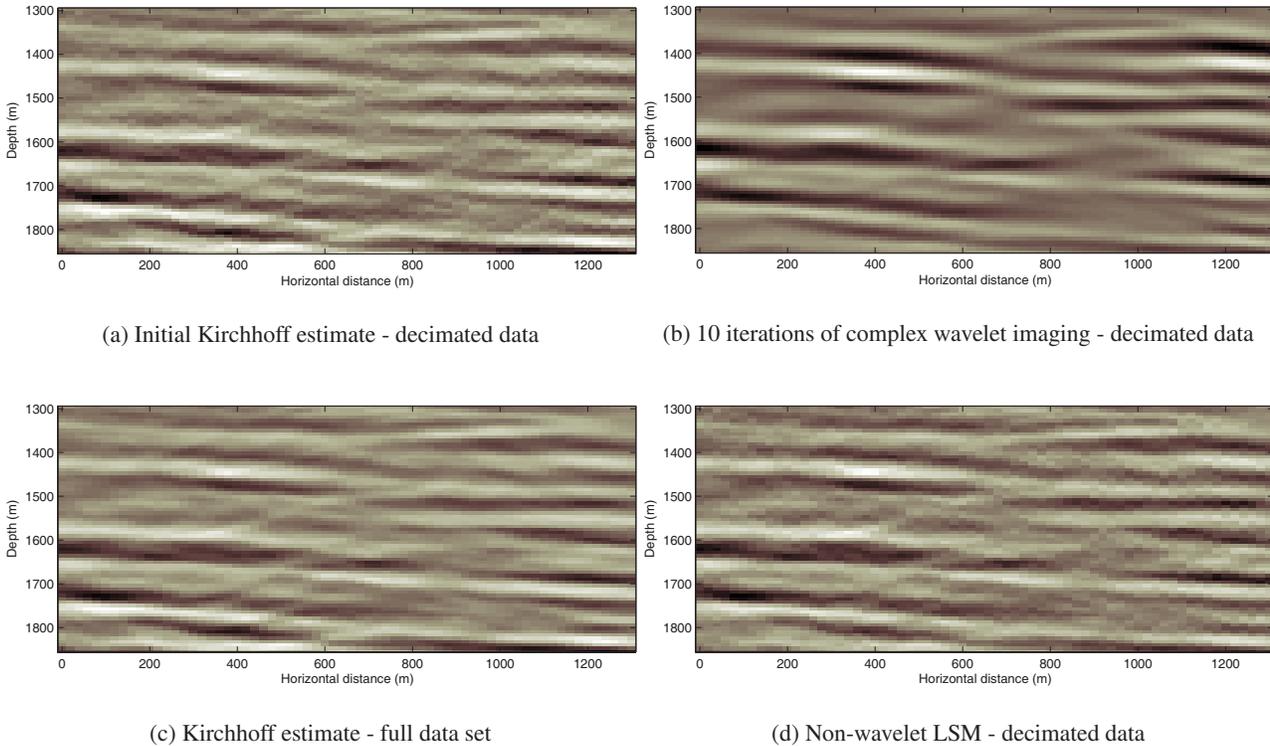


Fig. 2. Results of the complex wavelet seismic imaging system

on the path. This all requires approximately $\alpha N_x N_z N_s N_p$ operations.

b) Each trace is convolved with the source basis function (combined 'source wavelet' and 'wavelet shaping factor' in seismic processing terminology). This is typically performed in the frequency domain and requires of order $N_p N_t \log_2(N_t)$ operations, significantly less processing than step (a).

Hence the total processing is of order $N_s N_x N_z N_p$ operations. Note that this is much greater than the DT-CWT of the model, which is approximately $100 N_x N_z$ operations in 2D. Applying the adjoint operator \mathbf{L}^T to \mathbf{d} requires about the same amount of computation as applying the forward operator to \mathbf{m} , and is essentially step (b) with the filters reversed, followed by step (a) but going from the trace domain to the model domain.

7. REFERENCES

- [1] T. Nemeth, C. Wu, and G. T. Schuster, "Least-squares migration of incomplete reflection data," *Geophysics*, vol. 64, no. 1, pp. 208–221, 1999.
- [2] B. Duquet, K. J. Marfurt, and J. A. Dellinger, "Kirchhoff modeling, inversion for reflectivity, and subsurface illumination," *Geophysics*, vol. 65, no. 4, pp. 1195–1209, 2000.
- [3] N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Journal of Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 234–253, May 2001.
- [4] E. L. Miller and A. S. Willsky, "A multiscale, statistically-based inversion scheme for the linearized, inverse scattering problem," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 2, pp. 346–357, 1996.
- [5] X. Li, M. D. Sacchi, and T. J. Ulrych, "Wavelet transform inversion with prior scale information," *Geophysics*, vol. 61, no. 5, pp. 1379–1385, 1996.
- [6] J. A. Kane and F. J. Herrmann, "Wavelet domain linear inversion with application to well logging," in *71st Ann. Internat. Meeting, Soc of Expl. Geophys., Expanded Abstracts*, San Antonio, Tx, Sep 2001, pp. 694–697.
- [7] F. J. Herrmann, "Multifractional splines: application to seismic imaging," in *Proceedings of SPIE*, San Diego, Ca, Nov 2003, vol. 5207, pp. 240–258.
- [8] F. C. A. Fernandes, *Shift-Insensitive, Complex Wavelet Transforms with Controllable Redundancy*, Ph.D. thesis, Rice University, Electrical and Computer Engineering, Aug 2001.
- [9] P. de Rivaz and N. Kingsbury, "Bayesian image deconvolution and denoising using complex wavelets," in *IEEE International Conference on Image Processing*, Oct 2001, vol. 2, pp. 273–276.
- [10] N. Bleistein, J. K. Cohen, and F. G. Hagin, "Two and one-half dimensional born inversion with an arbitrary reference," *Geophysics*, vol. 52, no. 1, pp. 26–36, 1987.
- [11] A. J. W. Duijndam, A. W. F. Volker, and P. M. Zwartjes, "Reconstruction as an efficient alternative for least squares migration," in *70th Ann. Internat. Meeting, Soc of Expl. Geophys., Expanded Abstracts*, Calgary, Canada, Aug 2000, pp. 1012–1015.