GEOPHYSICS, VOL. 81, NO. 6 (NOVEMBER-DECEMBER 2016); P. A17–A21, 3 FIGS., 1 TABLE.
10.1190/GEO2016-0270.1

Restricted model domain time Radon transforms

Juan I. Sabbione[1] and Mauricio D. Sacchi[1]

ABSTRACT

The coefficients that synthesize seismic data via the hyperbolic Radon transform (HRT) are estimated by solving a linear-inverse problem. In the classical HRT, the computational cost of the inverse problem is proportional to the size of the data and the number of Radon coefficients. We have developed a strategy that significantly speeds up the implementation of time-domain HRTs. For this purpose, we have defined a restricted model space of coefficients applying hard thresholding to an initial low-resolution Radon gather. Then, an iterative solver that operated on the restricted model space was used to estimate the group of coefficients that synthesized the data. The method is illustrated with synthetic data and tested with a marine data example.

INTRODUCTION

The Radon transform has received a lot of attention over the past four decades as a flexible tool for processing exploration seismology data. Applications of the Radon transform span a diverse set of problems, such as velocity analysis (Thorson and Claerbout, 1985), multiple suppression (Hampson, 1986; Foster and Mosher, 1992), near-offset reconstruction (Kabir and Verschuur, 1995), antialiasing for reverse time migration (Wang and Nimsaila, 2014), separation of simultaneous sources (Trad et al., 2012; Ibrahim and Sacchi, 2014), diffraction enhancement (Bansal and Imhof, 2005), deghosting (Wang et al., 2014), and noise suppression of microseismic records (Forghani-Arani et al., 2013; Sabbione et al., 2015).

Radon transforms can be divided into two categories: time-invariant and time-variant Radon transforms. The linear and parabolic Radon transforms belong to the category of time-invariant transforms (Hampson, 1986). Time-invariant transforms are implemented via fast solvers that exploit the special Toeplitz structure of the Radon operator in the frequency-space domain (Beylkin, 1987; Kostov, 1990; Sacchi and Porsani, 1999). On the other hand, time-variant transforms are commonly computed in the time domain via iterative solvers devised for large linear-inverse problems (Thorson and Claerbout, 1985). Time-variant transforms include the hyperbolic Radon transform (HRT) and the apex-shifted HRT, and they are relatively slow in comparison with parabolic transforms. This explains the popularity of the parabolic Radon transforms for multiple attenuation. There are situations, however, where multiples cannot be modeled by the parabolic Radon transform. For instance, the parabolic approximation fails to provide sufficient accuracy to model long-offset data. Moreover, the NMO correction that is applied prior to the parabolic Radon transform stretches the wavelet and thus deteriorates the quality of the data. In these particular situations, one might prefer to adopt the HRT. The computational cost of the HRT could become a problem when processing a large number of common-midpoint (CMP) gathers or when it is used as part of 3D multiple prediction techniques (van Dedem and Verschuur, 2005). Hence, different methods have been proposed to improve the computational efficiency of the HRT, some of them in close relation with the present work (Yilmaz, 1989; de Bazelaire et al., 1991; Liu and Sacchi, 2002; Ng and Perz, 2004; Wang et al., 2009; Hu et al., 2013; Nikitin et al., 2016).

In this paper, we discuss an easy-to-apply strategy to compute the HRT quickly and efficiently. The proposed method uses the adjoint operator to identify the subset of the Radon domain that is needed to represent the data. An on-the-fly computation of the forward and adjoint operators on the restricted model space leads to a very fast implementation of the conjugate-gradient method. We also utilize the amplitude of the Radon coefficients of the active subset to compute model-space regularization weights for the linear inversion.

METHOD

Preliminaries

A common-shot or a common-receiver gather can be represented by a superposition of apex-shifted hyperbolas in the t-x domain as follows:

Manuscript received by the Editor 27 May 2016; revised manuscript received 1 July 2016; published online 20 September 2016.
[1]University of Alberta, Department of Physics, Edmonton, Alberta, Canada. E-mail: sabbione@ualberta.ca; msacchi@ualberta.ca.

$$d(t,x) = \int_a \int_v m\left(\tau = \sqrt{t^2 - \frac{(x-a)^2}{v^2}}, v, a\right) dv da, \quad (1)$$

where x is the source-receiver distance, τ indicates intercept traveltime, v is the velocity of the hyperbola, a denotes the apex, and $m(\tau, v, a)$ are the Radon coefficients. Equation 1 is often called the forward Radon transform. Similarly, one can define an adjoint transform that maps the data into the τ, v, a domain as follows:

$$m_{\text{adj}}(\tau, v, a) = \int_x d\left(t = \sqrt{\tau^2 + \frac{(x-a)^2}{v^2}}, x\right) dx. \quad (2)$$

Equations 1 and 2 are numerically computed by replacing the integrals by summations over discretized t, x, τ, v, and a variables. Then, the estimation of $m(\tau, v, a)$ from the data $d(t, x)$ entails solving a discrete linear-inverse problem (Thorson and Claerbout, 1985).

We simplify the notation by expressing the discretized version of equations 1 and 2 in matrix-vector form as follows:

$$\mathbf{d} = \mathbf{L}\mathbf{m}, \quad (3)$$

$$\mathbf{m}_{\text{adj}} = \mathbf{L}^T\mathbf{d}, \quad (4)$$

where $\mathbf{d}$ is a vector of size $N \times 1$ representing the data ($N = N_x \times N_t$). Similarly, $\mathbf{m}$ and $\mathbf{m}_{\text{adj}}$ are vectors of size $M \times 1$ representing the Radon coefficients ($M = N_t \times N_v \times N_a$). In realistic scenarios, one must consider the presence of observational noise. Therefore, the vector of Radon coefficients must be estimated from the expression $\mathbf{d} = \mathbf{L}\mathbf{m} + \mathbf{e}$, where $\mathbf{e}$ denotes noise. A common practice is to estimate $\mathbf{m}$ by minimizing a cost function composed of a quadratic misfit plus a regularization term as follows:

$$J(\mathbf{m}) = \|\mathbf{L}\mathbf{m} - \mathbf{d}\|_2^2 + \mu\mathcal{R}(\mathbf{m}). \quad (5)$$

The scalar μ is the tradeoff parameter that weights the relative importance of the misfit function versus the regularization term $\mathcal{R}(\mathbf{m})$.

The classical least-squares solution of equation 5 adopts a quadratic regularization term: $\mathcal{R}(\mathbf{m}) = \|\mathbf{m}\|_2^2$. When $\mathcal{R}(\mathbf{m})$ is a quadratic form, the cost function J is minimized via the method of conjugate gradients (CG). High-resolution Radon transforms, on the other hand, adopt either a Cauchy criterion or an l_1-norm to impose sparsity on the distribution of the Radon coefficients. In these cases, the cost function J is minimized via iterative reweighted least-squares (Cauchy regularization; Sacchi and Ulrych, 1995) or iterative solvers such as FISTA (l_1-regularization; Lu, 2013). One can also use the information contained in $\mathbf{m}_{\text{adj}}$ to generate model weights for the regularization term in equation 5 by adopting $\mathcal{R}(\mathbf{m}) = \|\mathbf{W}\mathbf{m}\|_2^2$ with $\mathbf{W} = \text{diag}(|\mathbf{L}^T\mathbf{d} + \epsilon|^{-1})$, where $|.|$ denotes the absolute value of each element of its argument, the operator diag(.) maps a vector into a diagonal matrix, and ϵ is a small value to avoid division by zero. In other words, the absolute values of the adjoint coefficients $m_{\text{adj}}(\tau, v, a)$ are used to improve the focusing power of the transform. This is equivalent to one iteration of the high-resolution Radon transform in Sacchi and Ulrych (1995) and Trad et al. (2003). In all cases, the computational cost of minimizing J is primarily associated to the cost of the forward and adjoint operators.

Restricted model space hyperbolic Radon transform

The matrix-vector formulation (equations 3 and 4) permits us to adopt the language of linear algebra and linear-inverse problems. However, we stress that our algorithm never forms matrices in explicit form. In other words, the forward and adjoint operators are computed by a series of nested loops on the variables x, τ, v, and a. A naive implementation of the Radon forward and adjoint operators has computational complexity of $\mathcal{O}(N_\tau N_x N_v N_a)$ for the case, where $N_t = N_\tau$. If we consider data $\mathbf{d}$ of fixed size $N_t \times N_x$, it is clear that the cost of applying the forward Radon operator and its adjoint is controlled by N_τ, N_v and N_a; that is to say, by the size of the Radon domain.

In the proposed algorithm, we restrict the size of the model domain by determining the triplets τ, v, a that correspond to the coefficients needed to fit the data. According to equation 2, the coefficients $m_{\text{adj}}(\tau, v, a)$ will be relatively large when integrating over reflections and small when integrating over random zero-mean noise. After computing $m_{\text{adj}}(\tau, v, a)$ using equation 2, we define a restricted Radon space $\mathbb{A}$ of active coefficients via the following expression:

$$\mathbb{A} = \left\{(\tau, v, a) : \frac{1}{N_x}|m_{\text{adj}}(\tau, v, a)| > T\right\}, \quad (6)$$

where T is the threshold. We consider a gather that has been normalized to unity to adopt a parameter T that satisfies $0 < T < 1$. The restricted operators $\mathbf{L}_{\mathbb{A}}$ and $\mathbf{L}_{\mathbb{A}}^T$ are then used by our iterative solver to compute the Radon coefficients

$$\tilde{m} = \underset{\mathbf{m}_{\mathbb{A}}}{\text{argmin}}[\|\mathbf{L}_{\mathbb{A}}\mathbf{m}_{\mathbb{A}} - \mathbf{d}\|_2^2 + \mu\|\mathbf{W}_{\mathbb{A}}\mathbf{m}_{\mathbb{A}}\|_2^2], \quad (7)$$

where we have selected a model-weighted quadratic regularization term. Given that now there is no risk of dividing by zero, we set $\mathbf{W}_{\mathbb{A}} = \text{diag}(|\mathbf{L}_{\mathbb{A}}^T\mathbf{d}|^{-1})$. The new cost function is minimized via the method of CG (Hestenes and Stiefel, 1952).

We have termed this algorithm restricted domain hyperbolic Radon transform (RHRT) to differentiate it from the classical HRT that uses the complete domain of Radon coefficients to model the data. Our algorithm has similarities to the methods presented by Liu and Sacchi (2002), Ng and Perz (2004), and Wang et al. (2009). Liu and Sacchi (2002) use an erosion/deflation algorithm, where the active set of coefficients are updated in each iteration of the CG algorithm. As a result, the cost function of the problem is continuously changed throughout iterations of the CG method and its convergence can be negatively affected. Another disadvantage of this algorithm is that extra computational cost is introduced in each iteration. Ng and Perz (2004) propose to consider the energy stacked along each moveout parameter to prioritize their estimation sequence via a Gauss-Seidel method. Wang et al. (2009) develop a greedy algorithm that iteratively builds the Radon domain using the residuals given by the data modeled in each step. The latter two methods involve an outer loop with certain number of iterations, thus introducing an extra parameter. Our algorithm, in contrast, applies CG directly to the restricted domain operators throughout all of the iterations. This guarantees that the cost function of the prob-

lem does not change during the optimization process. Furthermore, if the parameter T is properly chosen, the HRT solution that uses all the coefficients is very similar to the solution obtained via the RHRT.

SYNTHETIC TESTS

To illustrate the proposed method, we have generated a noisy synthetic common-shot gather with three reflections. The data consist of 51 traces separated by $\Delta x = 20$ m, with sampling interval $\Delta t = 4$ ms. A Ricker wavelet with peak frequency $f_0 = 20$ Hz models the seismic source. Band-limited zero-mean random noise was added to the data (signal-to-noise ratio = 2). The new method was applied to model a noise-free version of the data. For the Radon domain, the intercept ranges from $\tau = 0$ to 1.2 s, the velocity from $v = 1000$ to 3200 m/s, and the apex from $a = -300$ to 300 m.

A series of tests with varying parameters (N_v, N_a) were used to analyze the performance of the RHRT and compare it against the classical HRT. For all the tests, we adopted $\Delta\tau = 4$ ms and a threshold $T = 0.1$ to restrict the domain for the RHRT. In all cases, the trade-off was set ad hoc to $\mu = 100$, so as to be large enough to help in the denoising of the data without hurting the signal. Both algorithms were interrupted when the cost function change between iteration was less than 0.01%, in agreement with the stopping criterion adopted by Ibrahim and Sacchi (2014). Table 1 summarizes the test results comparing both methods. We show the computing times, number of iterations of the CG, and the final value of the cost function for the HRT (using all the coefficients) and RHRT. Additionally, we provide the percentage $P_{\mathbb{A}}$ of the total Radon domain coefficients that were used by the RHRT for the inversion. The tests correspond to a code written in Julia using tools from the Seismic.jl package (Stanton and Sacchi, 2016). In general, we observe that RHRT selects less than 5% of the Radon domain coefficients to reach a solution that properly honors the data, and is approximately 20 times faster than the HRT. The relatively small difference between methods in the value reached by the cost function suggests that the Radon coefficients included in $\mathbb{A}$ are sufficient to represent the data.

Table 1. Computational time of the HRT and RHRT for the synthetic tests. C denotes the computing time required to estimate the Radon coefficients and model the data, I is the number of iterations used by CG to converge to the solution, and $J(\tilde{m})$ is the final value of the cost function. In the last column, $P_{\mathbb{A}}$ denotes the percentage of the total number of coefficients $(N_\tau \times N_v \times N_a)$ used by RHRT.

		HRT			RHRT			
N_v	N_a	C (s)	I	$J(\tilde{m})$	C (s)	I	$J(\tilde{m})$	$P_{\mathbb{A}}$
26	25	8.4	16	214.3	0.6	18	228.8	4.1
26	41	19.9	25	232.8	1.2	23	252.1	4.5
26	61	31.7	27	211.6	1.7	25	232.3	4.5
45	25	20.9	25	233.6	1.1	22	251.6	4.4
45	41	36.3	26	161.2	1.4	23	176.9	3.4
45	61	54.5	27	241.2	2.9	27	266.2	4.6
81	25	41.5	27	209.7	2.0	25	232.0	4.1
81	41	71.7	29	213.7	3.8	29	238.1	4.7
81	61	133.7	37	234.5	5.2	29	267.9	4.6

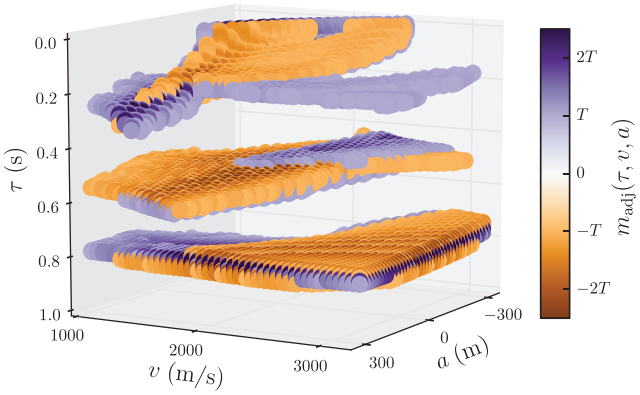

Figure 1. Active Radon space of coefficients $\mathbb{A}$ computed via amplitude thresholding of the adjoint operator coefficients $m_{\text{adj}}(\tau, v, a)$.

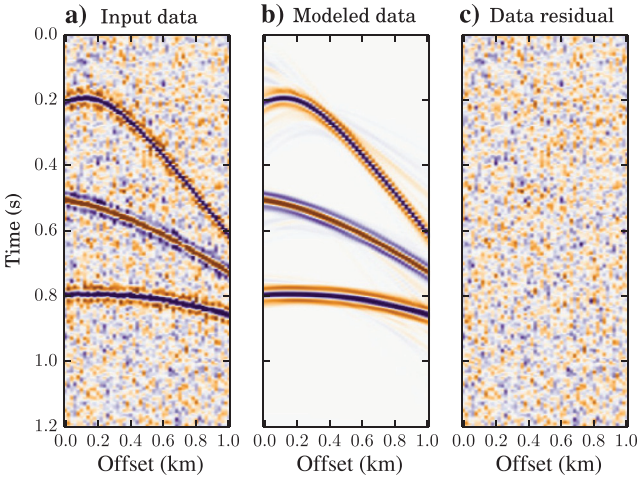

Figure 2. Synthetic data example: (a) Input noisy data. (b) Modeled data via the RHRT. (c) Data residual of panels (a and b).

To gain more insight into the proposed method, we show the results that were obtained with $N_v = 45$ and $N_a = 61$. Figure 1 displays the Radon coefficients that form the subset $\mathbb{A}$ used to invert the data. The colorbar scale was clipped to facilitate the visualization of the plot. The noisy input, the modeled data, and the data residual are shown in Figure 2 using the same color scale. The figure demonstrates that the data were properly modeled with the restricted domain $\mathbb{A}$ and the hyperbolic events were preserved in the inversion.

REAL-DATA EXAMPLE

We applied our algorithm to suppress the free-surface multiples in a marine data set from the Gulf of Mexico. The data consist of 810 shots with 183 traces per shot. The distance between sources and the distance between receivers is 26.67 m, with the nearest offset at −20.72 and the farthest offset at −4874.67 m. The data were sampled at 4 ms and the recording time was 7 s.

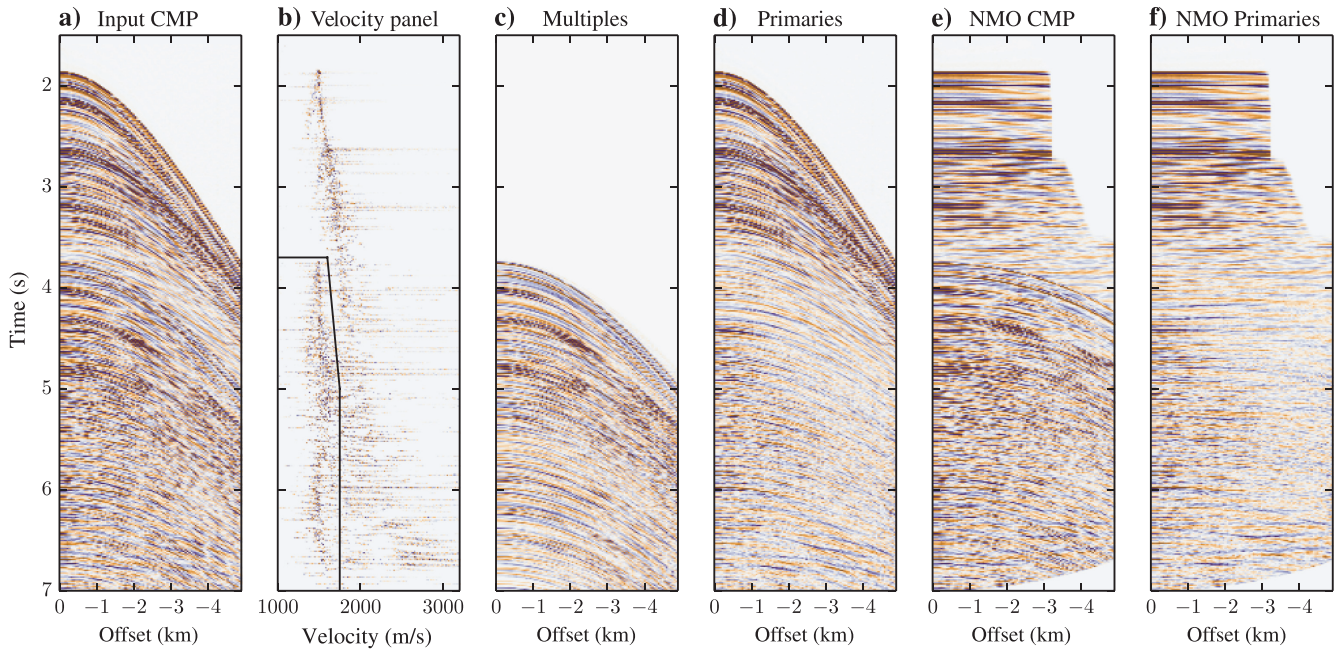

Figure 3. Multiple attenuation example. (a) Input CMP gather. (b) Velocity panel and subregion used to model the multiples. (c) Multiples modeled from panel (b). (d) Primary reflections obtained by subtracting the multiples from the input CMP. (e) Input CMP after NMO correction. (f) Primaries after NMO correction.

Figure 3 shows demultiple results of a CMP gather severely contaminated with multiples. We defined τ from 1.5 to 7 s with $\Delta\tau = 4$ ms, and used a fine sampling of $\Delta v = 5$ m/s with $N_v = 441$ velocities varying from $v = 1000$ to 3200 m/s. We set the apex to $a = 0$. According to equations 1 and 2, this is a special case of our algorithm, in which the method reduces to the HRT. To restrict the Radon domain, we set the threshold parameter to $T = 0.03$. A low value of T is preferred to account for low-energy reflections. According to equation 6, $T = 0.03$ means that we are only discarding those Radon coefficients that stack less than 3% of the maximum possible energy. However, this low value leads to using only approximately 17% (one sixth) of the Radon coefficients for the inversion. Thus, the computational cost of the RHRT is one sixth the cost of the HRT. In this example, we preferred to favor the data fitting over the regularization term to properly model all the multiples for their removal. Thus, we used a relatively small trade-off parameter $\mu = 0.1$. The multiples are indicated with a polygon in the velocity panel showed in Figure 3b. We isolated that subregion of the velocity gather and model them via equation 1. Then, we subtracted the modeled multiples shown in Figure 3c from the input CMP gather to form the multiple-free CMP gather (Figure 3d). Finally, we show the input CMP gather and the output demultiple CMP gather after NMO correction in Figure 3e and 3f, respectively. These figures demonstrate that the RHRT successfully removed the multiples from the data. The proposed method suggests that one should not use all the coefficients of the Radon domain to model the data.

CONCLUSION

This paper provides a strategy to model seismic data with time HRTs using model parameters finely sampled without an excessive computational cost. The method is based on a restriction of the Radon coefficients that are used to model the data. This leads to an algorithm that reduces the computational cost of time domain Radon transforms by approximately one order. We identify the active set of coefficients that stack reflections using the adjoint operator followed by amplitude hard thresholding. The adjoint coefficients, that are computed only once, also provide the weights for the regularization term of the cost function. This method improves the usage of time-domain HRT to perform multiple removal, denoising, interpolation, and other applications.

ACKNOWLEDGMENTS

The authors are grateful to the sponsors of the Signal Analysis and Imaging Group at the University of Alberta. This research was also supported by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Bansal, R., and M. G. Imhof, 2005, Diffraction enhancement in prestack seismic data: Geophysics, **70**, no. 3, V73–V79, doi: 10.1190/1.1926577.
Beylkin, G., 1987, Discrete Radon transform: IEEE Transactions on Acoustics, Speech and Signal Processing, **35**, 162–172.
de Bazelaire, E., S. Desaunay, and J. R. Viallix, 1991, The hyperbolic Radon transform: 53rd Annual International Conference and Exhibition, EAGE, Extended Abstracts, A028.
Forghani-Arani, F., M. Willis, S. S. Haines, M. Batzle, J. Behura, and M. Davidson, 2013, An effective noise-suppression technique for surface microseismic data: Geophysics, **78**, no. 6, KS85–KS95, doi: 10.1190/GEO2012-0502.1.
Foster, D. J., and C. C. Mosher, 1992, Suppression of multiple reflections using the Radon transform: Geophysics, **57**, 386–395, doi: 10.1190/1.1443253.
Hampson, D., 1986, Inverse velocity stacking for multiple elimination: Journal of Canadian Society of Exploration Geophysics, **22**, 44–55.

Hestenes, M., and E. Stiefel, 1952, Methods of conjugate gradients for solving linear systems: Journal of Research of the National Bureau of Standards, **49**, 409–436, doi: 10.6028/jres.049.044.
Hu, J., S. Fomel, L. Demanet, and L. Ying, 2013, A fast butterfly algorithm for generalized Radon transforms: Geophysics, **78**, no. 4, U41–U51, doi: 10.1190/geo2012-0240.1.
Ibrahim, A., and M. D. Sacchi, 2014, Simultaneous source separation using a robust Radon transform: Geophysics, **79**, no. 1, V1–V11, doi: 10.1190/geo2013-0168.1.
Kabir, M. N., and D. Verschuur, 1995, Restoration of missing offsets by parabolic Radon transform: Geophysical Prospecting, **43**, 347–368, doi: 10.1111/gpr.1995.43.issue-3.
Kostov, C., 1990, Toeplitz structure in slant-stack inversion: 60th Annual International Meeting, SEG, Expanded Abstracts, 1618–1621.
Liu, Y., and M. D. Sacchi, 2002, De-multiple via a fast least-squares hyperbolic Radon transform: 72nd Annual International Meeting, SEG, Expanded Abstracts, 2182–2185.
Lu, W., 2013, An accelerated sparse time-invariant Radon transform in the mixed frequency-time domain based on iterative 2D model shrinkage: Geophysics, **78**, no. 4, V147–V155, doi: 10.1190/geo2012-0439.1.
Ng, M., and M. Perz, 2004, High resolution Radon transform in the t-x domain using intelligent prioritization of the Gauss-Seidel estimation sequence: 74th Annual International Meeting, SEG, Expanded Abstracts, 2160–2163.
Nikitin, V. V., F. Andersson, M. Carlsson, and A. A. Duchkov, 2016, Fast hyperbolic Radon transform represented as convolutions in log-polar coordinates: ArXiv e-prints.
Sabbione, J. I., M. D. Sacchi, and D. R. Velis, 2015, Radon transform-based microseismic event detection and signal-to-noise ratio enhancement: Journal of Applied Geophysics, **113**, 51–63, doi: 10.1016/j.jappgeo.2014.12.008.
Sacchi, M. D., and M. Porsani, 1999, Fast high resolution parabolic Radon transform: 69th Annual International Meeting, SEG, Expanded Abstracts, 1477–1480.
Sacchi, M. D., and T. J. Ulrych, 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, **60**, 1169–1177, doi: 10.1190/1.1443845.
Stanton, A., and M. D. Sacchi, 2016, Efficient geophysical research in Julia: CSEG GeoConvention 2016, 1–3.
Thorson, J. R., and J. F. Claerbout, 1985, Velocity-stack and slant-stack stochastic inversion: Geophysics, **50**, 2727–2741, doi: 10.1190/1.1441893.
Trad, D., R. Siliqi, G. Poole, and J. Boelle, 2012, Fast and robust deblending using apex shifted Radon transform: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-0703.1.
Trad, D., T. Ulrych, and M. Sacchi, 2003, Latest views of the sparse Radon transform: Geophysics, **68**, 386–399, doi: 10.1190/1.1543224.
van Dedem, E. J., and D. J. Verschuur, 2005, 3D surface-related multiple prediction: A sparse inversion approach: Geophysics, **70**, no. 3, V31–V43, doi: 10.1190/1.1925752.
Wang, J., M. Ng, and M. Perz, 2009, Fast high-resolution Radon transforms by greedy least-squares method: 79th Annual International Meeting, SEG, Expanded Abstracts, 3128–3132.
Wang, P., and K. Nimsaila, 2014, Fast progressive sparse Tau-P transform for regularization of spatially aliased seismic data: 84th Annual International Meeting, SEG, Expanded Abstracts, 3589–3593.
Wang, P., S. Ray, and K. Nimsaila, 2014, 3D joint deghosting and crossline interpolation for marine single-component streamer data: 84th Annual International Meeting, SEG, Expanded Abstracts, 3594–3598.
Yilmaz, O., 1989, Velocity-stack processing: Geophysical Prospecting, **37**, 357–382, doi: 10.1111/j.1365-2478.1989.tb02211.x.