# Suppressing Migration Image Artifacts Using a Support Vector Machine Method

**Yuqing Chen*, Yunsong Huang*, Lianjie Huang***

* *Los Alamos National Laboratory, Geophysics Group, MS D452, Los Alamos, NM 87545, USA*

*Email: yuqing.chen@kaust.edu.sa*

(July 7, 2018)

**Suppressing Migration Artifacts by SVM**

Running head: ***Suppressing Migration Image Artifacts***

## ABSTRACT

Least-squares reverse-time migration (LSRTM) can produce images with more balanced amplitudes and fewer artifacts compared to reverse-time migration. However, conventional LSRTM might not be able to suppress significant image artifacts caused by a sparse data acquisition geometry because of incomplete cancelation of migration"smiles". This problem becomes even more severe in the case of migration using vertical seismic profiling data. These artifacts not only slow down the convergence of LSRTM, but also obscure the identification of the actual geology. To alleviate this problem, we develop a new method to effectively suppress migration image artifacts using a support vector machine method. The method employs a weighted stacking of common-shot migration images, where the weights preserve the signal while suppressing image artifacts. We calculate the weights using a supervised support vector machine method. The weights approach one for valid image points and are close to zero for artifact points. We form a migration image using a weighted stacking across the shot-domain common-image gathers. We verify the effectiveness of our new method, denoted as SVM filtering, using numerical tests on both synthetic and field data to

produce migration images with better signal-to-noise ratios (S/N) and fewer aliasing artifacts compared to the conventional LSRTM method.

# INTRODUCTION

Least-squares reverse-time migration (LSRTM) produces images with higher resolution, more balanced amplitudes and fewer artifacts compared to reverse-time migration (Nemeth et al., 1999; Duquet et al., 2000; Plessix and Mulder, 2004; Tang, 2009; Dai et al., 2012; Dutta and Schuster, 2014; Dai et al., 2015; Dutta, 2017). However, conventional LSRTM may not able to suppress artifacts caused by inadequate consideration of physics in both the modeling and adjoint operators, errors in the velocity model, and sparsity in the source-receiver acquisition geometry. In these cases, the migration "smiles" cannot be completely canceled. These artifacts affect the convergence rates of LSRTM, deteriorate the amplitude versus offset analysis, and interfere with the ability of the interpreter to identify important geological features of interests. This problem becomes more severe for migration of vertical seismic profiling (VSP) data because of small numbers of sources and receivers in VSP surveys compared to those in surface seismic surveys.

To alleviate this problem, Cabrales-Vargas et al. (2013) developed a preconditioned least-squares migration (LSM) method with an additional roughness penalty term in the conventional LSM misfit function. The method better preserves migration image amplitudes and further reduce aliasing image artifacts than the conventional LSM. Dutta and Schuster (2015) used a sparse LSRTM method with the seislet transform as a change of basis for the reflectivity. Along with a dip-constrained preconditioner, this method can produce an image with more meaningful structural updates with each iteration. Lin et al. (2016) incorporated an $L_p$-norm based compressive-sensing term in the misfit function of LSRTM to improve migration imaging of sparse seismic data. Dutta et al. (2017) introduced a local Radon-based preconditioning LSRTM method to alleviate the problems of low S/N data and image aliasing. For the case of VSP, Yu et al. (2008) implemented a stereographic local beam imaging scheme to attenuate migration noise caused by limited acquisition apertures in VSP data. Zhou et al. (2010) used a local-angle-domain correlation imaging condition to attenuate migration artifacts.

Prestack LSRTM (Schuster, 2017) applies the iterative LSM separately to each shot gather to obtain $n$ prestack migration images, where $n$ is the number of shots. Then all $n$ images are stacked together to yield the final migration image. The advantage of this approach is that, rather than simply stacking, one can first pre-process each shot-domain common-image gather (CIG) for better stacking using such as a trim statics applied to CIGs (Huang et al., 2014).

In this paper, we employ a supervised support vector machine method to suppress migration image artifacts in prestack LSRTM images. We first use the supervised support vector machine (SVM) method to compute a weight for each point in the shot-domain CIG. The weight approaches one for a real image point, and approaches zero for an artifact point. The weights preserve the real image components and suppress the artifacts. With those weights, we then carry out a weighted stacking across all shot-domain CIGs to form the final migration image. We use both synthetic and field seismic data to verify the effectiveness of our new method. In our synthetic tests, we mainly focus on reducing migration image artifacts caused by a sparse source-receiver geometry.

This paper is organized as the following. After the introduction, we introduce the theory of prestack LSRTM using a supervised SVM method to suppress migration image artifacts. We then present numerical results in both the synthetic- and field-data examples, discuss the limitations of our new method, and give conclusions.

# THEORY

## Prestack Least-squares reverse-time migration

Prestack least-squares reverse-time migration (Schuster, 2017) assumes each shot gather $\mathbf{d}_i$ is associated with its own reflectivity model $\mathbf{m}_i$. The forward modeling problem can be written as

$$\begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_n \end{pmatrix} = \begin{bmatrix} \mathbf{L}_1 & & & \\ & \mathbf{L}_2 & & \\ & & \ddots & \\ & & & \mathbf{L}_n \end{bmatrix} \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \vdots \\ \mathbf{m}_n \end{pmatrix}, \tag{1}$$

where $\mathbf{L}_i$ represents a linear modeling operator associated with reflectivity model $\mathbf{m}_i$, and $n$ is the number of shot gathers. Equation (1) can be rewritten as

$$\mathbf{d} = \widehat{\mathbf{L}}\widehat{\mathbf{m}}. \tag{2}$$

The migration image $\widehat{\mathbf{m}}^{mig}$ is obtained after applying migration operator $\hat{\mathbf{L}}^T$ to the observed data,

$$\begin{pmatrix} \mathbf{m}_1^{mig} \\ \mathbf{m}_2^{mig} \\ \vdots \\ \mathbf{m}_n^{mig} \end{pmatrix} = \begin{bmatrix} \mathbf{L}_1^T & & & \\ & \mathbf{L}_2^T & & \\ & & \ddots & \\ & & & \mathbf{L}_n^T \end{bmatrix} \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_n \end{pmatrix}, \tag{3}$$

which can be shortened as

$$\widehat{\mathbf{m}}^{mig} = \widehat{\mathbf{L}}^T\mathbf{d}. \tag{4}$$

Therefore, the prestack LSRTM misfit function is defined as

$$\epsilon(\widehat{\mathbf{m}}) = \frac{1}{2}\sum_{i=1}^{n} ||\widehat{\mathbf{L}}\widehat{\mathbf{m}} - \mathbf{d}||^2 + R, \tag{5}$$

where $R$ is the regularization term used to balance the amplitude between $\mathbf{m}_i$ and $\mathbf{m}_{i+1}$, which is controlled by the normalization parameter $\lambda_i$

$$R = \frac{1}{2}\|\frac{\mathbf{m}_{i+1}}{\lambda_{i+1}} - \frac{\mathbf{m}_i}{\lambda_i}\|. \tag{6}$$

In summary, prestack LSRTM applies iterative LSM to each shot gather to obtain $n$ prestack migration images, and then stacks all $n$ images together to produce the final migration image. The advantage of prestack LSRTM is that, when there are velocity errors and migration artifacts, rather than simply stacking shot-domain CIGs, the method allows us to preform some pre-processing to improve stacking. For example, one can apply trim statics to CIGs to align migrated reflections with one another.

## Supervised Support Vector Machine

Support vector machine (SVM) is a supervised machine learning method that classifies binary data into two classes (Cortes and Vapnik, 1995). The supervised learning problem of classification uses a set of training data consisting of $N$ training samples

$$(\mathbf{x}^{(i)}, y^{(i)}), \tag{7}$$

where $i = 1, 2, ..., N$, $\mathbf{x}^{(i)}$ is the $i^{th}$ feature vector with dimension $D \times 1$ and $y^{(i)} = 1$ or $0$ is the binary target data value for classification. $D$ indicates the number of feature for each feature vector. For a set of two-dimensional feature vectors $\mathbf{x}^{(i)}$, the goal of SVM is to determine the dashed line that has the greatest separation from the green and blue dots in Figure 1. The optimal dashed line has the fattest margin width and is equidistant from the solid lines that are parallel to one another. This dashed line, also known as the

decision line, is mathematically defined by the normal vector $\mathbf{w}$ and the intercept $b$, where any vector $\mathbf{x}$ on the decision line satisfies $w_1x_1 + w_2x_2 + b = 0$. The red and blue dots intersecting with the black solid line are known as the support vectors. These support vectors define the points with the closest perpendicular distance to the decision line. The decision lines in Figure 1a and 1b are also able to separate the two classes of data, but they are less optimal compared to that in Figure 1c. The margin thicknesses of the decision lines in Figure 1a and 1b are thinner than the one in Figure 1c. The thinner the margin thickness, the more likely it is to misclassify the data subject to errors in the feature vector $\mathbf{x}^{(i)}$.

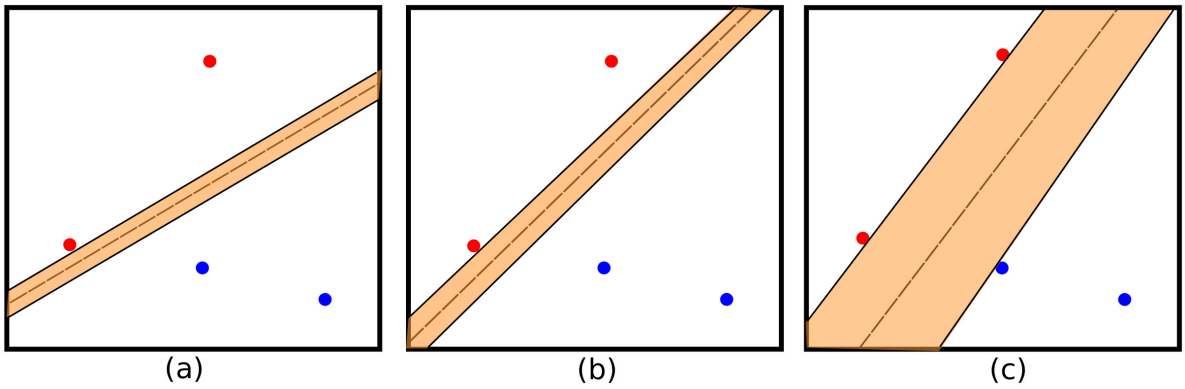


Figure 1: Two classes of data with different decision lines, where the red and blue dots represent different classes. Even though the decision boundaries in (a) and (b) cleanly separate the two datasets, they are not optimal separators compared to the decision boundary in (c) which has the fattest thickness margin.

Once the normal vector $\mathbf{w}$ and the intercept $b$ are determined after sufficient training, the SVM classifies new points $\tilde{\mathbf{x}}^{(i)}$ that are out of the training set. Figure 1 depicts samples in a two-dimensional case. In a more general case where the samples are in a $D$-dimensional space, SVM predicts a $D$-1 dimension hyperplane to separate different classes of data. The samples $\mathbf{x}$ on this hyperplane satisfies

$$\mathbf{w} \cdot \mathbf{x}^{(i)} = w_1x_1 + w_2x_2 + w_3x_3 + ... + w_Dx_D + b = 0, \tag{8}$$

where $\mathbf{w}$ is determined by solving a minimization problem

$$arg\min_{\mathbf{w},b} d = \frac{1}{2}\|\mathbf{w}\|^2,$$
$$\text{subject to } y^{(n)}(\mathbf{w}^T \cdot \mathbf{x}^{(n)} + b) \geq 1. \tag{9}$$

Solving the minimization problem determines the optimal line with the fattest thickness of margin. Inequality (9) ensures that $\mathbf{w}$ cleanly separates the training set $\mathbf{x}^{(i)}$. A numerical solution to the above minimization problem obtained using, e.g. quadratic programming (Wright and Nocedal, 1999), results in the fattest margin as shown in Figure 1.

## Workflow of Using SVM to Suppress Migration Image Artifacts in Prestack LSRTM

We use SVM to classify between the real image components and artifact points for prestack LSRTM CIGs in the shot domain. Figure 2a shows a prestack CIG containing both the real image components and migration image artifacts. For a real image component, SVM predicts a weight value close to 1. On the other hand, the weight approaches 0 for an artifact point. We then carry out a weighted stacking across all shot-domain CIGs to produce the final migration image. The weights preserve the real image components and suppress the artifact points. The detailed workflow is described below.

(1) We define three features in the shot-domain CIGs: Amplitude $A$, local dip angle $\theta$ and coherency $S$. We obtain the amplitude feature using the envelope of each trace in a CIG gather as shown in Figure 2b. To estimate the local dip and coherency, we define a local window ($win_x$ and $win_z$) centered at the $i^{th}$ pixel, and apply the Radon transform to the image within this local window to determine the local dip information at this local image point. Similarly, we estimate the coherency

features using the semblance analysis within the local window,

$$S(iz,is) = \frac{1}{win_z+1} \frac{\sum_{i=iz-N_z}^{iz+N_z} \Big( \sum_{j=is-N_x}^{is+N_x} CIG(i,j) \Big)^2}{\sum_{i=iz-N_z}^{iz+N_z} \sum_{j=is-N_x}^{is+N_x} CIG(i,j)^2}, \tag{10}$$

where $N_z = win_z/2$ and $N_x = win_x/2$ is the half length of the local window along the vertical and horizontal directions, respectively. For notational brevity, the offset coordinate $x$ is silent. Repeating this procedure for all points in the CIG gives the dip angle and coherency features in Figures 2c and 2d, respectively.

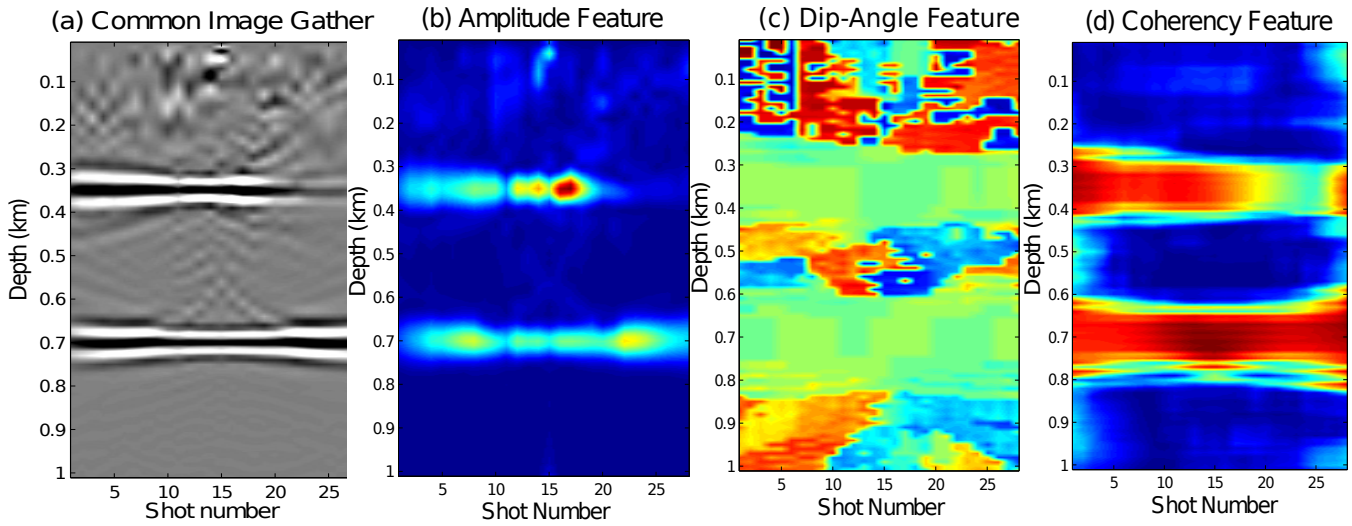


Figure 2: The original shot-domain common-image gather (a) together with its corresponding amplitude (b), dip angle (c) and coherency (d) features, respectively.

(2) Building the training set. We define the training set as $(\mathbf{x}_{train}^{(i)}, y_{train}^{(i)}) = (A^{(i)}, \theta^{(i)}, S^{(i)}, y^{(i)})$ for about $5\%$ picked points in the CIG shown in Figure 3a. The red and blue dots indicate the picked real image components and artifact points, respectively. Each point in this training set is manually classified as either signal $y^{(i)} = 1$ or noise $y^{(i)} = 0$. Figure 4a shows the training set in the feature

domain, where the effective signal points and artifact points are cleanly separable from one another. We then use the SVM procedure to determine the $3 \times 1$ vector $\mathbf{w} = (w_1, w_2, w_3)$ and the intercept $b$ that cleanly separates the labeled signal from noise. This vector and the intercept map a surface in feature space as shown in Figure 4b, where the colored plane defines the decision plane in the 3D feature space.

(3) Once the normal vector $\mathbf{w}$ and the intercept $b$ are determined, we compute $\mathbf{w} \cdot \mathbf{x} + b$ for all points in all CIGs, and output image weights as shown in Figure 3b for Figure 2a. These weights preserve the real image components in Figure 2a while suppressing the artifact points, as depicted in Figure 3c.

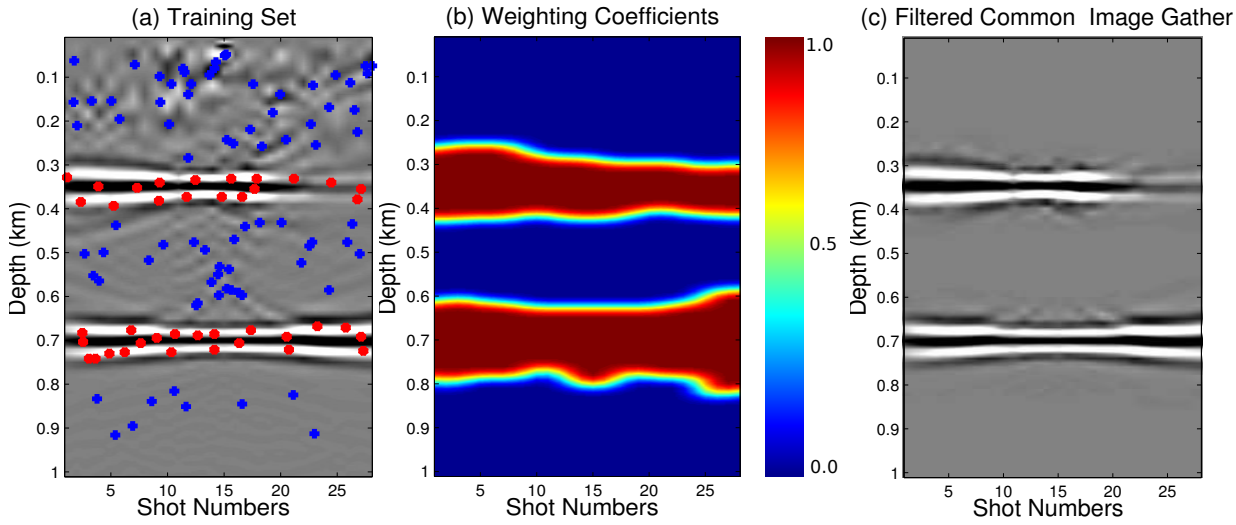


Figure 3: (a) The shot-domain common-image gather with picked points. The red and blue dots represent the picked real image components and artifact points, respectively, for building the training set. (b) The computed weights for all points in (a). (c) The weighted common-image gather.

We compare our method with the preconditioning approach as follows. By analyzing the features, each point in a CIG can map into a high dimensional space, where the real image components and artifacts are cleanly separable from one another. This idea is very similar to many other preconditioning LSRTM methods, such as LSRTM with Radon transform- or seislet transform-based preconditioning methods (Dutta

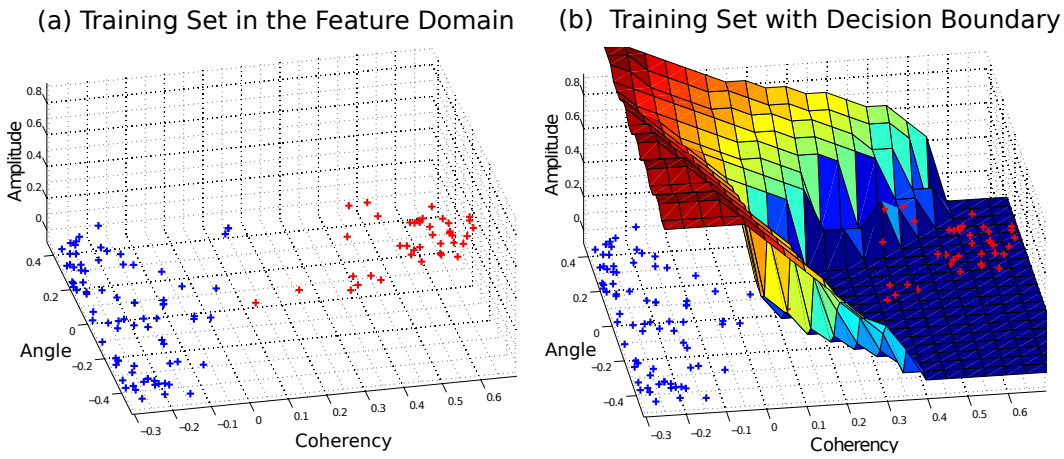


Figure 4: (a) The 3D plot of the training set in the 3D feature domain, where the red and blue dots indicates the real image component and artifact points, respectively. The real image component points and the artifact points are separable in this 3D feature domain. (b) The real image component points and the artifact points in the training set are well separated by a colored hyperplane calculated using the SVM.

and Schuster, 2015; Dutta et al., 2017). They all transform the aliasing image into the other domain where the real image components and artifacts are well separable, enabling us to preserve the real image components while suppressing the artifacts. In our method, we use high-dimensional feature space to separate the real image component points from artifact points, and use SVM to suppress the artifacts and preserve the real image components. One of the advantages of our method is that the additional computational cost for using SVM is trivial compared to the other methods.

Theoretically, the more features used, the higher dimension is the feature space, and in principle we can more easy separate the real image components from the artifacts. Nevertheless, this method is based on the assumption that all these features are independent. To analyze the redundancy of the features, one can use the principal component analysis (PCA) (Wold et al., 1987).

## NUMERICAL RESULTS

We now test the effectiveness of our SVM filtering of migration artifacts in synthetic surface seismic data and VSP data. We then apply SVM filtering to field seismic data acquired at Kevin Dome in north central Montana. The migration results in the same figure are compared using the same colorbar scale.

### Three-Layer Model

We use a three-layer model to illustrate the artifact-reduction capability of our SVM filtering method. We generate synthetic seismic data using the reflectivity model in Figure 5 with a homogeneous background velocity of 2000 m/s. There are 20 shots evenly spaced at a distance of 100 m, where each shot is recorded using 40 receivers separated at an interval of 50 m on the surface.

Figures 6a and 6b compare the prestack LSRTM image before and after suppressing artifacts using SVM filtering, respectively, for this sparse dataset. It is evident from Figure 6b that SVM filtering reduces the migration artifacts and preserves the real image components compared to the original prestack LSRTM image in Figure 6a. In this numerical example, we only used $5\%$ of the points of a selected CIG at $x = 1$ km for training to determine $\mathbf{w}$ and $\mathbf{b}$. We then used the trained $\mathbf{w}$ and $\mathbf{b}$ vectors to classify every image point for all CIGs.

We next validate the capability of our SVM-based method for VSP data. Migration of VSP data suffers more from aliasing artifacts compared to surface seismic data because of the limited acquisition aperture in VSP. We place 50 receivers with an interval of 20 m at depths ranging from 10 m to 990 m in a vertical well located at x = 1 km. It can be clearly observed that Figure 7b is free from aliasing artifacts compared to Figure 7a since the weights computed using our SVM-based method efficiently suppress the artifacts and preserve the real image components. Because the characteristics of CIGs change dramatically along the horizontal direction, CIGs at $x$ = 0.3 km, $x = 0.9$ km and $x = 1.7$ km are used for training to compute the

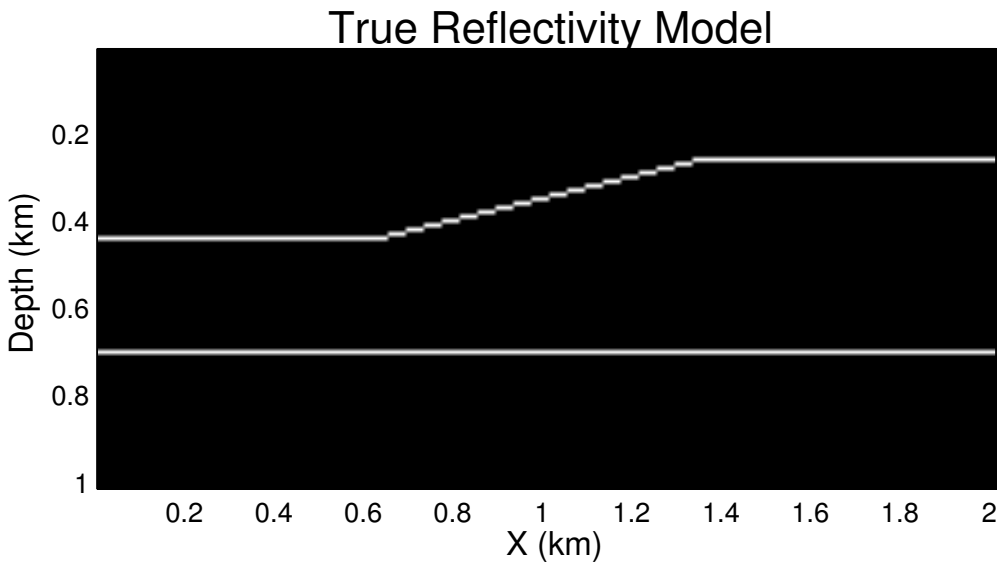


Figure 5: The true reflectivity model.

weights of the CIGs in the range of $0 \sim 0.4$ km, $0.4 \sim 1.6$ km and $1.6 \sim 2.0$ km, respectively.

### Marmousi Model

We use a portion of the Marmousi model to test the effectiveness of SVM filtering. The synthetic data are generated for the upper-right region of the Marmousi model with a size of $200 \times 312$ grid points. Figure 8 is the true velocity model used to generate synthetic seismic data. The true velocity model is smoothed with a $5 \times 5$ window for migration imaging. A 4-8 time-domain acoustic finite-difference method is used for both modeling and migration. The source function is a Ricker wavelet with a peak frequency of 20 Hz. The synthetic data are computed for 31 sources and 62 receivers on the surface with a 100 m source interval and 50 m receiver interval.

The conventional prestack LSRTM image in Figure 9a contains severe migration image artifacts caused by the sparse data acquisition geometry. Figure 9b shows an improved image where these artifacts are

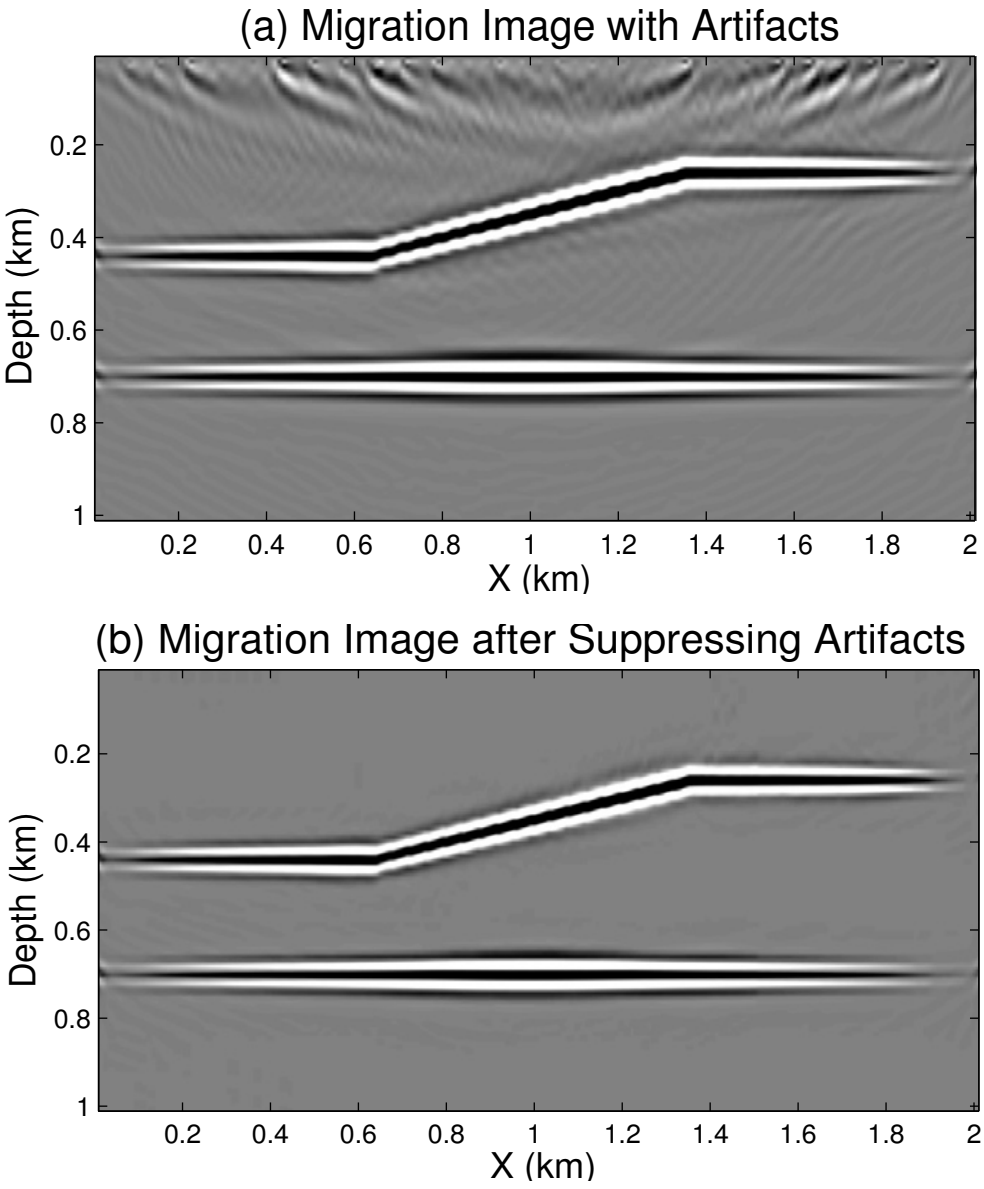


Figure 6: (a) The prestack LSRTM image contains significant aliasing artifacts caused by the sparse acquisition geometry. (b) The corresponding image produced using our SVM-based method to suppress artifacts in (a).

largely suppressed using our SVM filtering. In this model, we use the CIGs at $x$ = 0.75 km and $x$ = 2.25 km to compute the weights of the CIGs in the range of $0 \sim 1.5$ km and $1.5 \sim 3.12$ km, respectively. We use only $3\%$ points of a CIG for training, which is trivial in computation cost compared to some other preconditioning

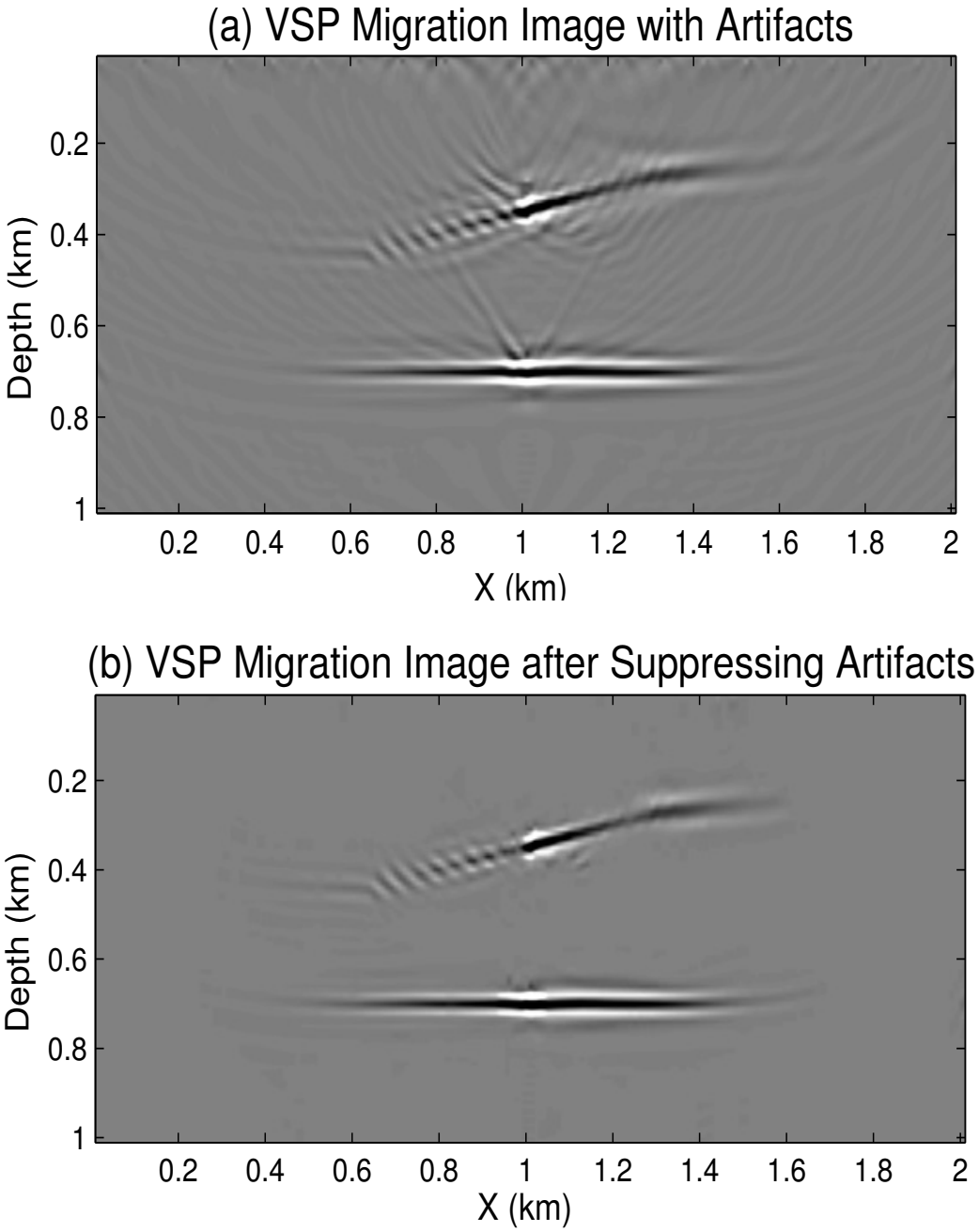


Figure 7: (a) The prestack LSRTM image of the VSP data. (b) The corresponding image after SVM filtering where the image artifacts are mostly suppressed.

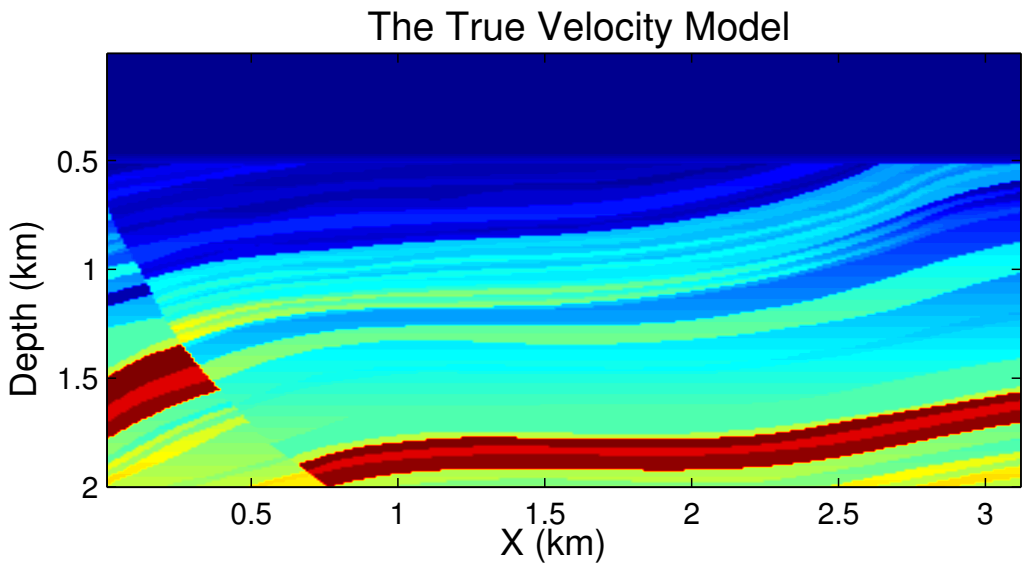


Figure 8: The true velocity model in the upper-region region of the Marmousi model.

method. Figure 10 displays the magnified views of these images, showing significant artifact reduction.

For the case of VSP, we place 100 receivers at depths from $z = 10$ m to $z = 1990$ m in a borehole with a receiver interval of 20 m. The receiver well is located at $x = 1.56$ km. Figure 11b shows the prestack LSRTM image after SVM filtering. Most artifacts are removed compared to the conventional prestack LSRTM result in Figure 11a. In particular, the image between $x = 1.3$ km to $x = 2.3$ km in Figure 11a contains significant of high-dip angle artifacts that interfere with horizontal reflectors, which are successfully removed with preservation of the horizontal reflections. Magnified views of these images in Figures 12 and 13 illustrate the artifact reduction more clearly. The magnified events in the red box (Figure 12) still contain some artifacts around $x$ = 1.4 km, which can likely be removed with a larger training set.

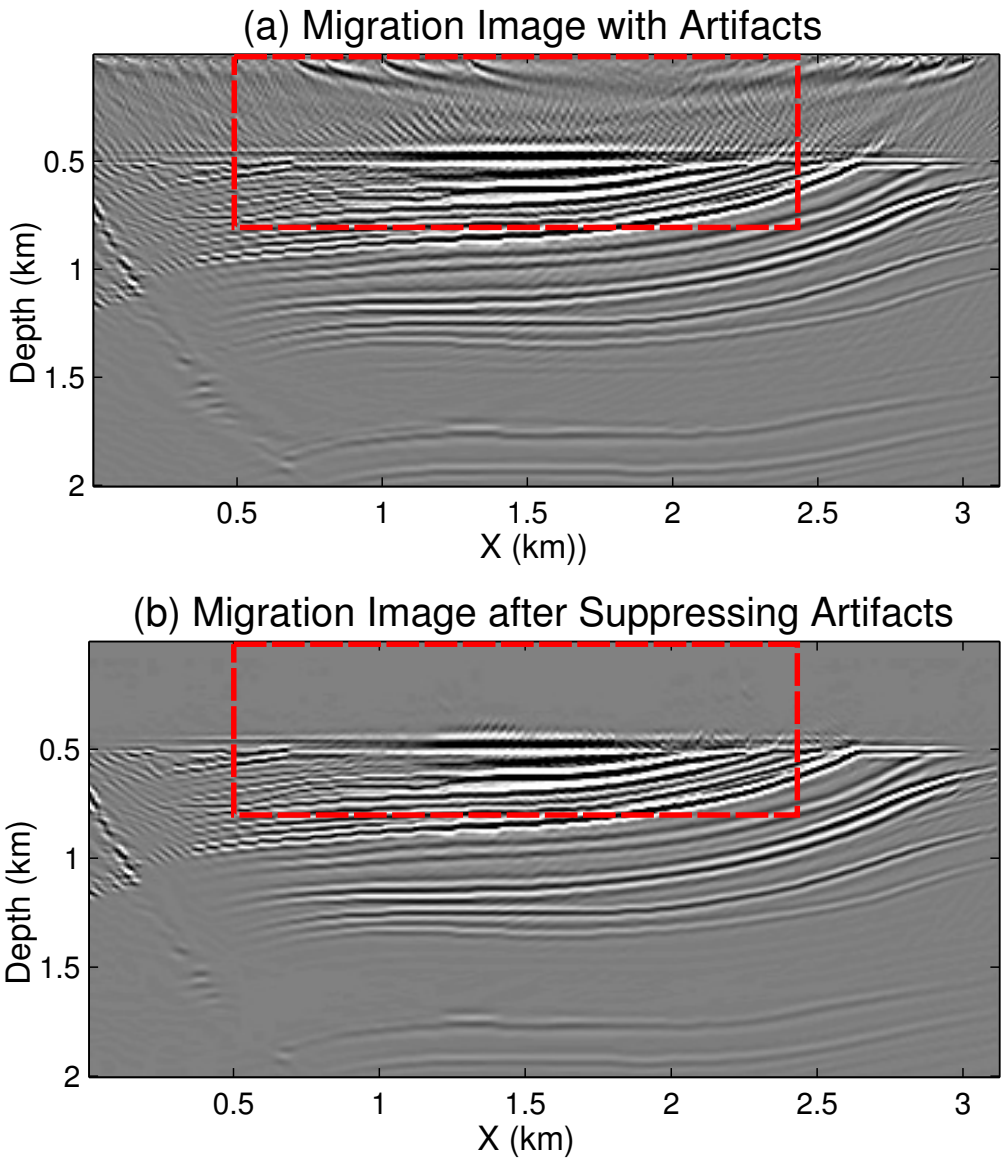


Figure 9: (a) The prestack LSRTM image contains significant aliasing artifacts caused by the sparse acquisition geometry. (b) The corresponding image produced using our SVM-based method to suppress artifacts in (a).

**SEG/EAGE Salt Model**

Migration images computed for data generated from the SEG/EAGE salt model are now SVM filtered. 33 shot gathers are computed using the true velocity model shown in Figure 14. At the depth of 10 m below the

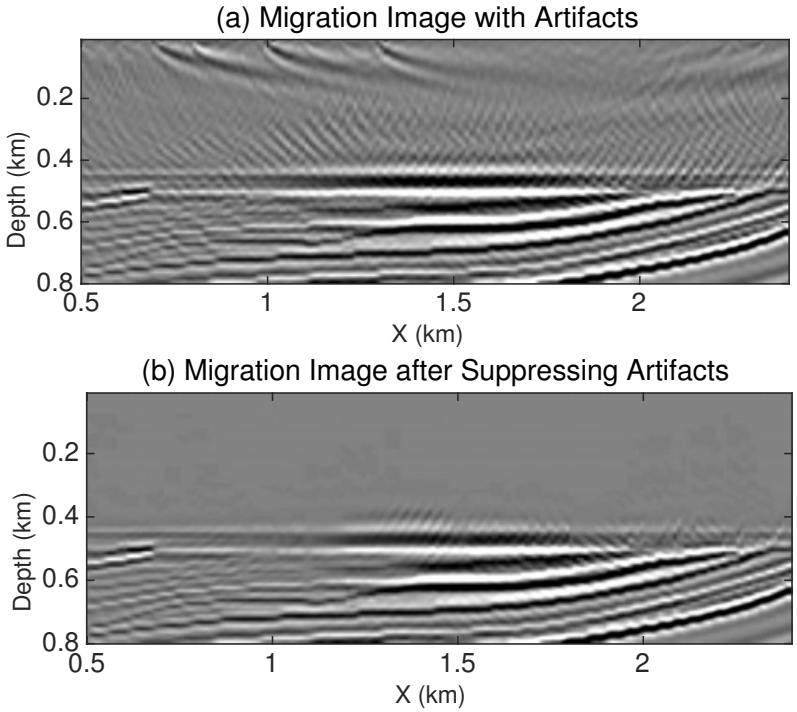


Figure 10: Magnified views of the red boxes in Figure 9. The image in (b) obtained using our SVM-based method contains significant fewer artifacts than the conventional pre-stack LSRTM image in (a).

free surface, the sparse shot spacing is 200 m and the receivers are uniformly distributed at a spacing of 30 m. The conventional prestack LSRTM image in Figure 15a contains severe aliasing artifacts above the salt body. Figure 15b shows an much improved image where these artifacts are removed using SVM filtering. The salt boundary in Figure 15b is much cleaner compared to the conventional prestack LSRTM image in Figure 15a. Figures 16 and 17 depicts a comparison of the magnified views of the red and black boxes in Figure 15, showing significant artifact suppression. In this model, we select the CIGs at $x = 1.5$ km, $x = 3.5$ km and $x = 5.5$ km to compute the weights of the nearby CIGs.

For the VSP test, we place 105 receivers at depths from $z = 10$ m to $z = 2090$ m in a borehole with a receiver interval of 20 m. The receiver well is located at $x = 5.4$ km. Figures 18a and 18b show the images before and after SVM filtering. Figure 18a depicts the conventional prestack LSRTM image which suffers

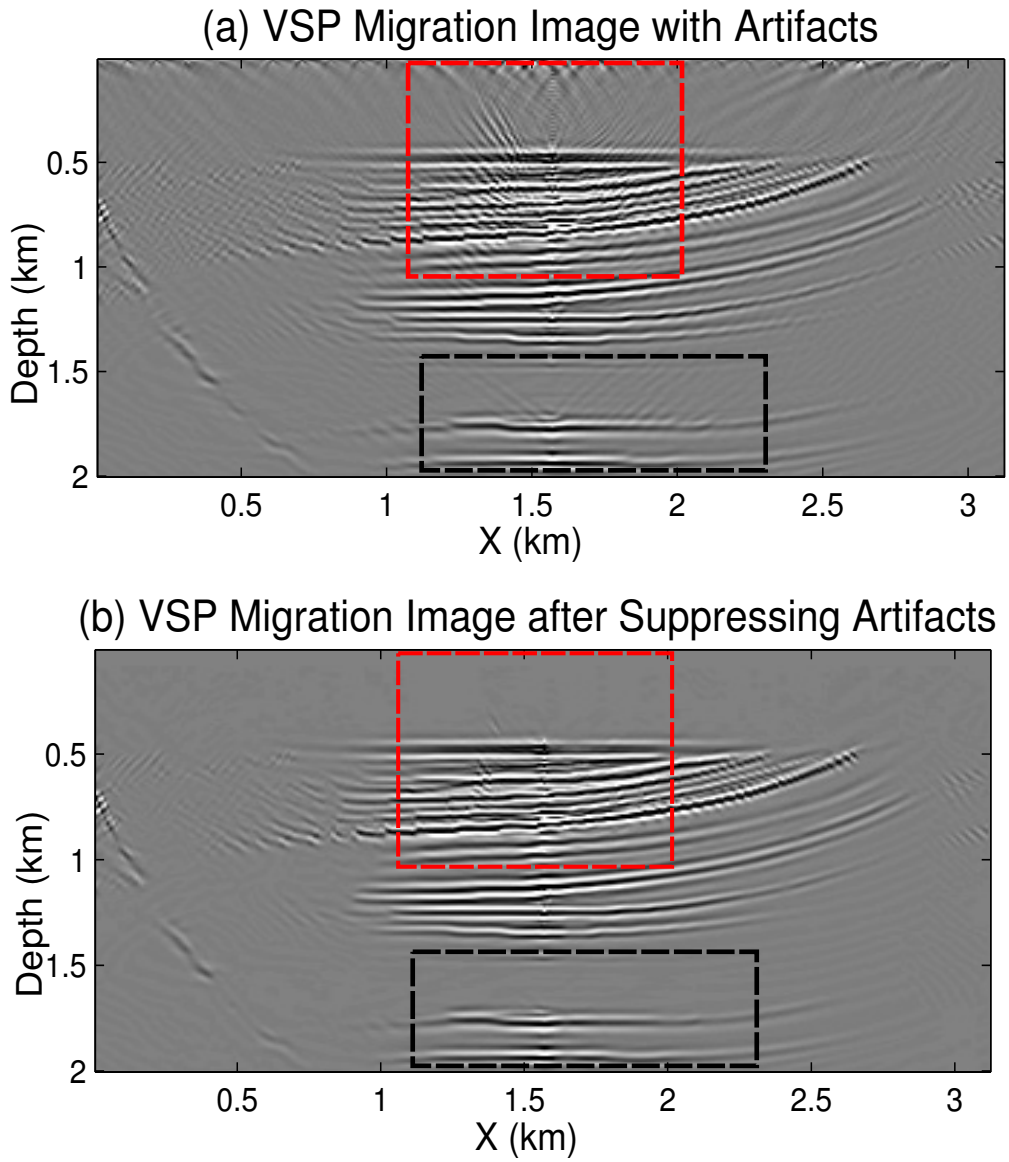


Figure 11: (a) The prestack LSRTM image of the VSP data. (b) The corresponding image after SVM filtering.

from strong aliasing artifacts especially near the well location because of the sparse geometry. However, these artifacts are significantly alleviated in Figures 18b and 19b using SVM filtering.

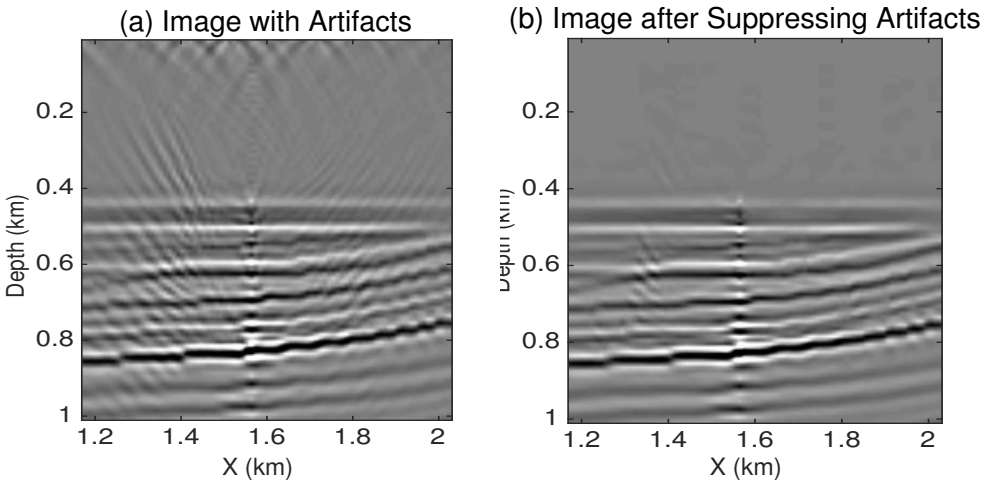


Figure 12: Magnified views of the red boxes in Figure 11. The image in (b) obtained using SVM filtering contains significantly fewer artifacts than the conventional pre-stack LSRTM image in (a).

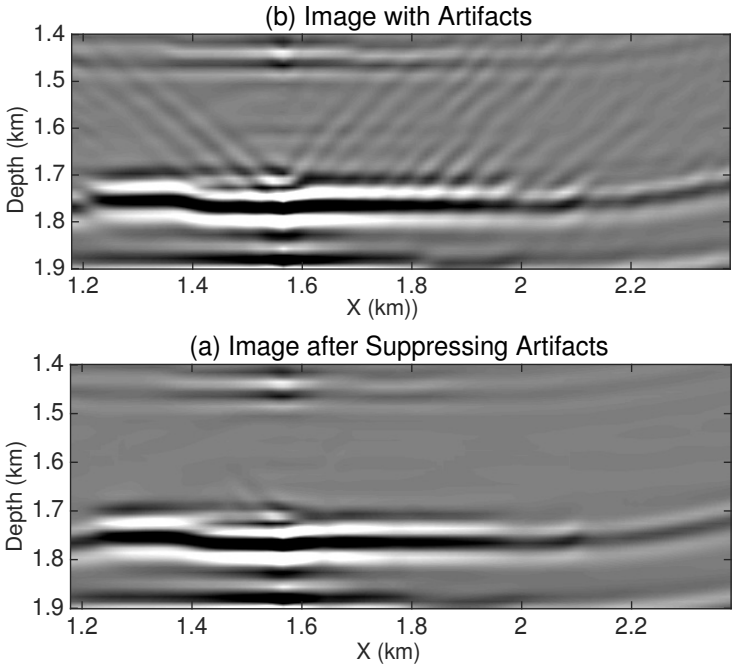


Figure 13: Magnified views of the black boxes in Figure 11. The image in (b) obtained using SVM filtering contains significantly fewer artifacts than the conventional pre-stack LSRTM image in (a).

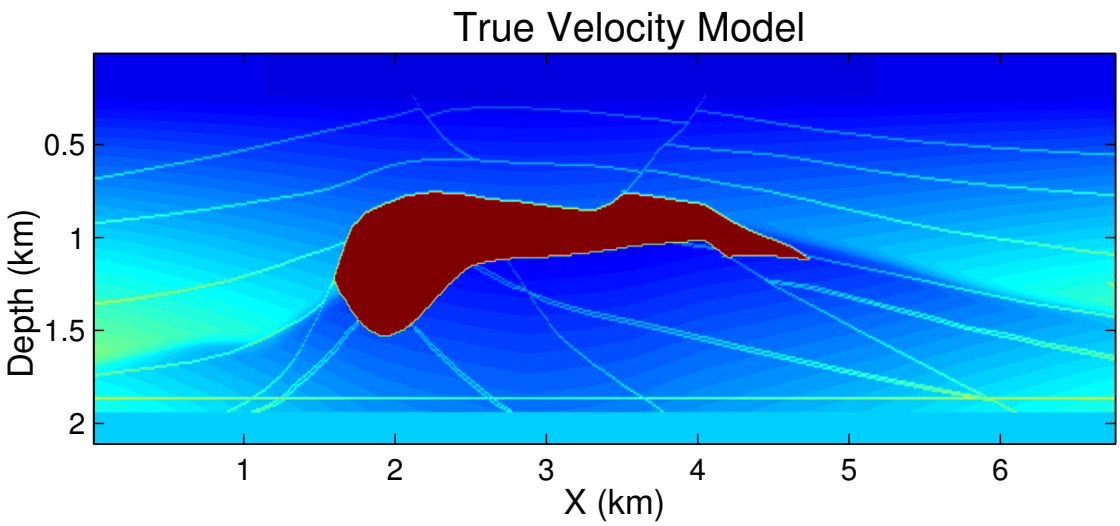


Figure 14: The true velocity model.

**Kevin Dome Field Seismic Data**

As a final example, we test the capability of SVM filtering using the Kevin Dome field data set. These data consist of 24 common-shot gathers with a shot interval around 270 m. The numbers of receivers for each shot range from 56 to 111 with a interval around 33 m. The migration velocity in Figure 20 is estimated using a plane-wave migration velocity analysis method (Guo and Schuster, 2017).

Figure 21a shows the conventional prestack LSRTM image with severe artifacts particularly at the very shallow part above $z = 0.5$ km. These artifacts result from the sparse acquisition geometry. Figure 22 displays three shot-domain CIGs at $x = 0.7$ km, $x = 3.5$ km, and $x = 5.6$ km, respectively. The red boxes in Figure 22 show that the shallow artifacts have very similar characteristics with the real image component characterized by nearly flat reflectors with strong amplitudes. Therefore, the three features used for the previous data examples are no longer sufficient for separating the real image components from the artifacts. A unique feature of this data is that there is poor spatial coherency in the stacked image. Therefore we

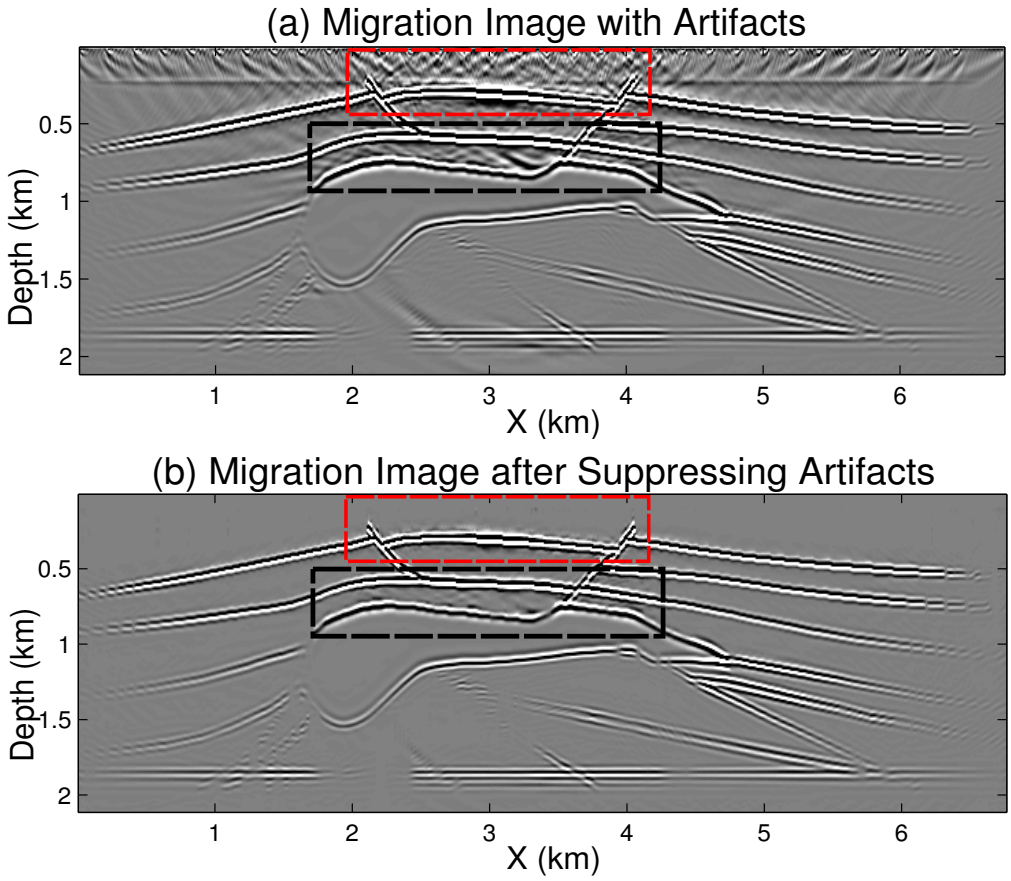


Figure 15: (a) The prestack LSRTM image contains significant aliasing artifacts caused by the sparse acquisition geometry. (b) The corresponding image produced using our SVM-based method to suppress artifacts in (a).

include spatial coherency of the data as the fourth feature to help distinguish the real image component from the artifacts in the shot-domain CIGs. Figure 21b shows an improved image where these artifacts are largely removed using SVM filtering with four features. Figures 23 and 24 compare the magnified views of the red and black boxed in Figure 21, showing a clearly improvement in arifact reduction.

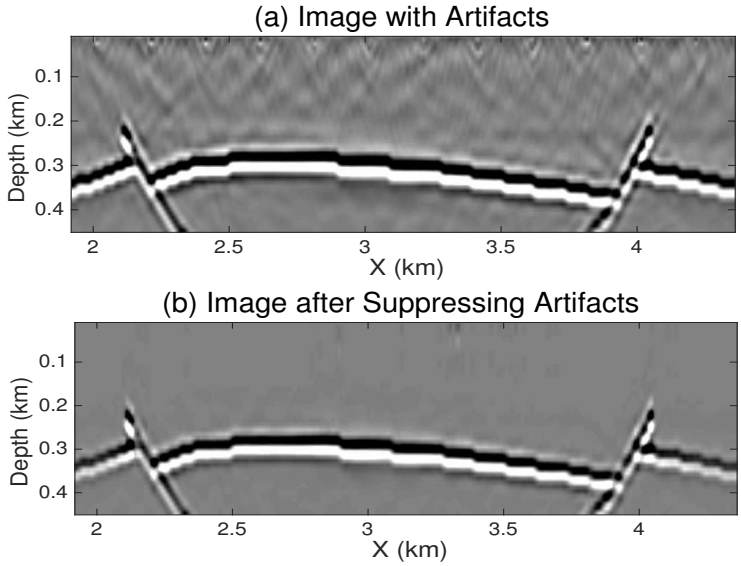


Figure 16: Magnified views of the red boxes in Figure 15. The SVM filtered image in (b) contains fewer artifacts than the conventional pre-stack LSRTM image in (a).

## CONCLUSIONS

We present a SVM filtering method for suppressing migration image artifacts caused by the sparse data acquisition geometry. Its effectiveness is validated using both synthetic and field seismic data. SVM filtering method employs the features of coherency, amplitude, and dipping angle from selected shot-domain common-image gathers to distinguish the real image components from artifacts. For each migration image, only a few common-image gathers are used for training to compute the SVM weights for all common-image gathers. The method uses only $3\% \sim 10\%$ of the points in selected common-image gathers for training. For our example, fewer than $1\%$ of the total CIGs were used for trainig. The additional computational cost of SVM filtering is trivial compared to some other preconditioning LSRTM methods.

The limitation of our method is that its accuracy relies on an accurate migration velocity model. However, this problem can be alleviated by combining our method with the trim statics method (Huang et al.,

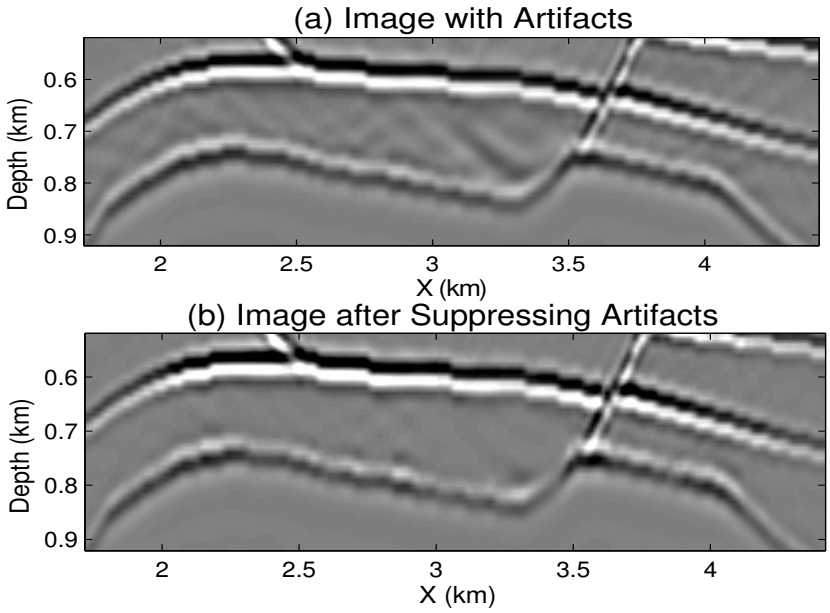


Figure 17: Magnified views of the black boxes in Figure 21. The SVM filtered image in (b) contains fewer artifacts than the conventional pre-stack LSRTM image in (a).

2014), which can first automatically align the events in common-image gathers for certain velocity errors before SVM filtering the artifacts in the common-image gathers.

## ACKNOWLEDGMENTS

This work was supported by U.S. Department of Energy through contract DE-AC52-06NA25396 to Los Alamos National Laboratory (LANL). We thank Zongcai Feng for providing the Kevin Dome migration velocity model. Yuqing Chen would like to thank King Abdullah University of Science and Technology (KAUST) for funding his graduate studies. The computation was performed using super-computers of LANL's Institutional Computing Program.

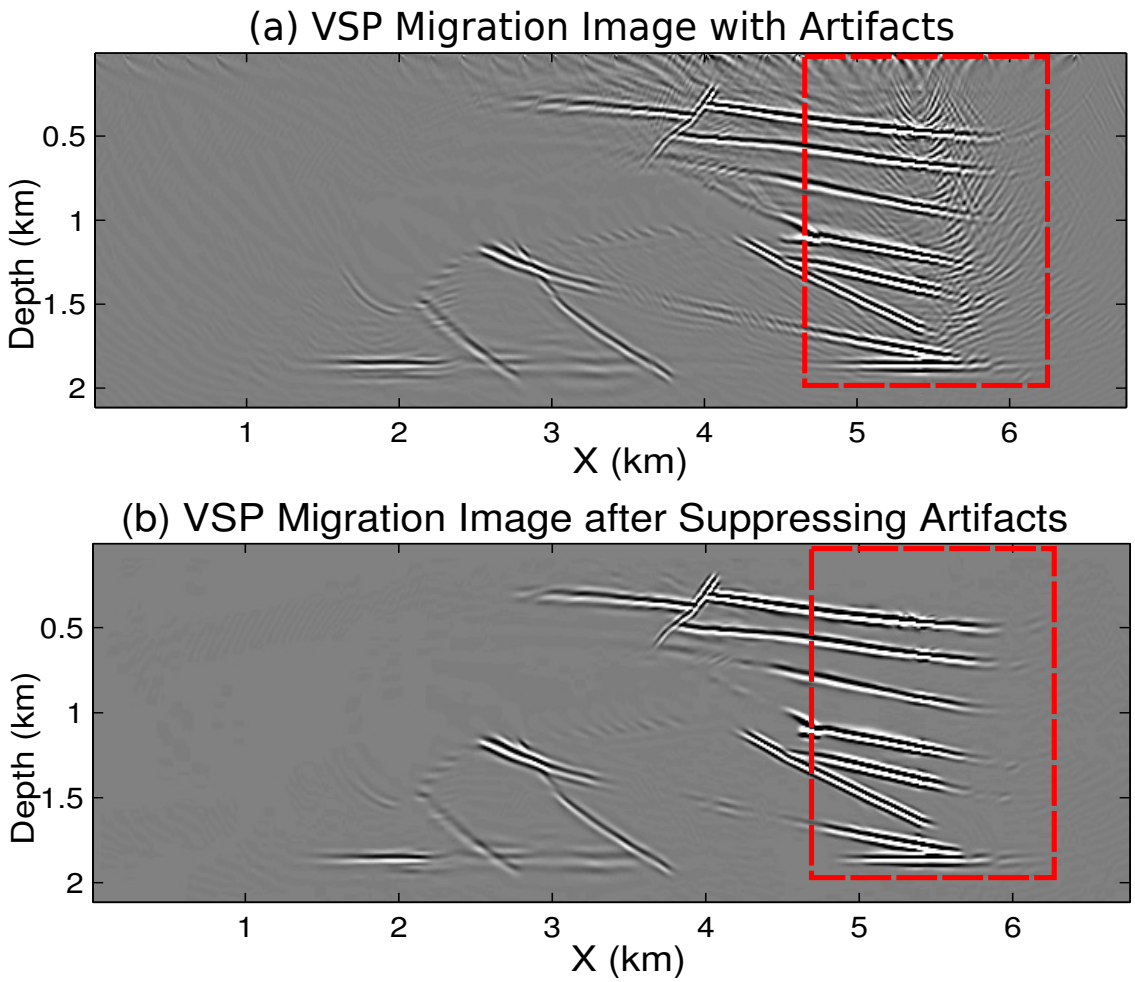


Figure 18: (a) The prestack LSRTM image of the VSP data. (b) The corresponding image after SVM filtering.

## REFERENCES

Cabrales-Vargas, A., K. J. Marfurt, et al., 2013, Amplitude-preserving imaging of aliased data using pre-conditioned kirchhoff least-squares depth migration.: SEG Annual Meeting, 3726–3730.

Cortes, C. and V. Vapnik, 1995, Support-vector networks: Machine learning, **20**, 273–297.

Dai, W., P. Fowler, and G. T. Schuster, 2012, Multi-source least-squares reverse time migration: Geophysical Prospecting, **60**, 681–695.

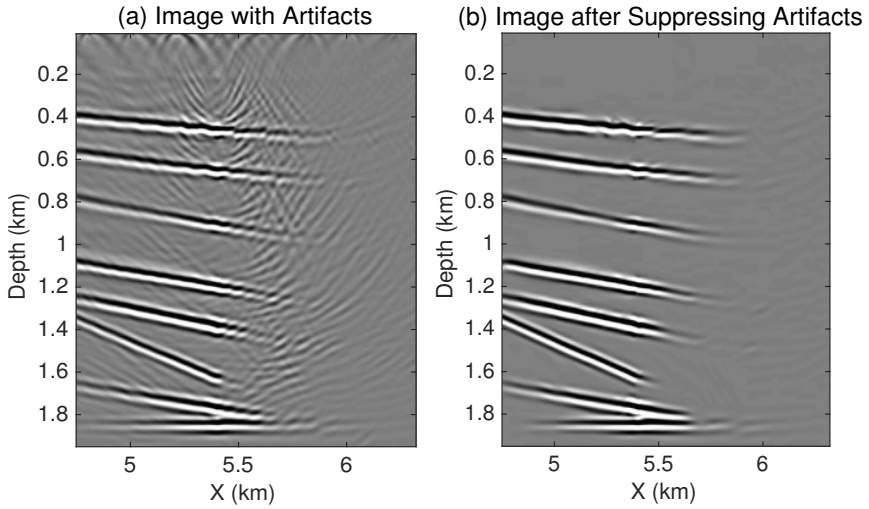


Figure 19: Magnified views of the red boxes in Figure 18. The SVM filtered image in (b) contains fewer artifacts than the conventional pre-stack LSRTM image in (a).

Dai, W., Z. Xu, and R. Coates, 2015, Least-squares reverse-time migration for visco-acoustic media: SEG Expanded Abstracts, 3387–3391.

Duquet, B., K. J. Marfurt, and J. A. Dellinger, 2000, Kirchhoff modeling, inversion for reflectivity, and subsurface illumination: Geophysics, **65**, 1195–1209.

Dutta, G., 2017, Sparse least-squares reverse time migration using seislets: Journal of Applied Geophysics, **136**, 142–155.

Dutta, G., M. Giboli, C. Agut, P. Williamson, and G. T. Schuster, 2017, Least-squares reverse time migration with local radon-based preconditioning: Geophysics, **82**, S75–S84.

Dutta, G. and G. T. Schuster, 2014, Attenuation compensation for least-squares reverse time migration using the viscoacoustic-wave equation: Geophysics, **79**, S251–S262.

———, 2015, Sparse least-squares reverse time migration using seislets: SEG Technical Program Expanded Abstracts 2015.

Guo, B. and G. T. Schuster, 2017, Wave-equation migration velocity analysis using plane-wave common-

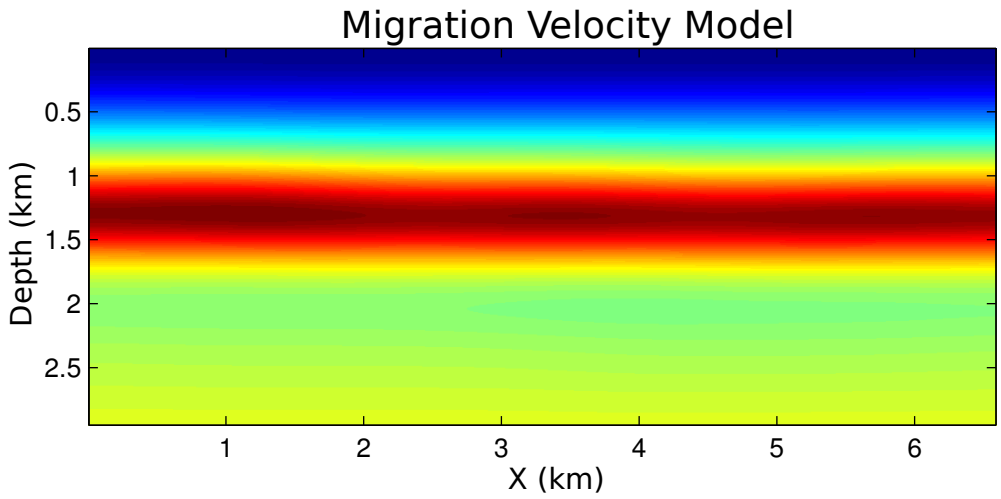


Figure 20: The migration velocity model inverted using a plane-wave migration velocity analysis method.

image gathers: Geophysics, **82**, S327–S340.

Huang, Y., X. Wang, G. T. Schuster, et al., 2014, Non-local means filter for trim statics: SEG Annual Meeting, 3925–3929.

Lin, Y., L. Huang, J. Queen, J. Moore, and E. Majer, 2016, Least-squares reverse-time migration with compressive sensing for sparse seismic data: Workshop on Geothermal Engineering.

Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221.

Plessix, R.-E. and W. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophysical Journal International, **157**, 975–987.

Schuster, G., 2017, Seismic inversion: Society of Exploration Geophysicists.

Tang, Y., 2009, Target-oriented wave-equation least-squares migration/inversion with phase-encoded hessian: Geophysics, **74**, WCA95–WCA107.

Wold, S., K. Esbensen, and P. Geladi, 1987, Principal component analysis: Chemometrics and intelligent laboratory systems, **2**, 37–52.

Wright, S. and J. Nocedal, 1999, Numerical optimization: Springer Science, **35**, 7.

Yu, J., B. Hornby, et al., 2008, A strategy for attenuating vsp migration artifacts: local beam migration: SEG Annual Meeting, 3385–3389.

Zhou, Y.-H., J.-H. Gao, B.-L. Wang, and Y.-Y. He, 2010, Migration scheme for imaging offset vsp data within local phase space: Applied Geophysics, **7**, 31–40.

## LIST OF FIGURES

1 Two classes of data with different decision lines, where the red and blue dots represent different classes. Even though the decision boundaries in (a) and (b) cleanly separate the two datasets, they are not optimal separators compared to the decision boundary in (c) which has the fattest thickness margin.

2 The original shot-domain common-image gather (a) together with its corresponding amplitude (b), dip angle (c) and coherency (d) features, respectively.

3 (a) The shot-domain common-image gather with picked points. The red and blue dots represent the picked real image components and artifact points, respectively, for building the training set. (b) The computed weights for all points in (a). (c) The weighted common-image gather.

4 (a) The 3D plot of the training set in the 3D feature domain, where the red and blue dots indicates the real image component and artifact points, respectively. The real image component points and the artifact points are separable in this 3D feature domain. (b) The real image component points and the artifact points in the training set are well separated by a colored hyperplane calculated using the SVM.

5 The true reflectivity model.

6 (a) The prestack LSRTM image contains significant aliasing artifacts caused by the sparse acquisition geometry. (b) The corresponding image produced using our SVM-based method to suppress artifacts in (a).

7 (a) The prestack LSRTM image of the VSP data. (b) The corresponding image after SVM filtering where the image artifacts are mostly suppressed.

8 The true velocity model in the upper-region region of the Marmousi model.

9 (a) The prestack LSRTM image contains significant aliasing artifacts caused by the sparse acquisition geometry. (b) The corresponding image produced using our SVM-based method to suppress artifacts in (a).

10 Magnified views of the red boxes in Figure 9. The image in (b) obtained using our SVM-based

method contains significant fewer artifacts than the conventional pre-stack LSRTM image in (a).

11 (a) The prestack LSRTM image of the VSP data. (b) The corresponding image after SVM filtering.

12 Magnified views of the red boxes in Figure 11. The image in (b) obtained using SVM filtering contains significantly fewer artifacts than the conventional pre-stack LSRTM image in (a).

13 Magnified views of the black boxes in Figure 11. The image in (b) obtained using SVM filtering contains significantly fewer artifacts than the conventional pre-stack LSRTM image in (a).

14 The true velocity model.

15 (a) The prestack LSRTM image contains significant aliasing artifacts caused by the sparse acquisition geometry. (b) The corresponding image produced using our SVM-based method to suppress artifacts in (a).

16 Magnified views of the red boxes in Figure 15. The SVM filtered image in (b) contains fewer artifacts than the conventional pre-stack LSRTM image in (a).

17 Magnified views of the black boxes in Figure 21. The SVM filtered image in (b) contains fewer artifacts than the conventional pre-stack LSRTM image in (a).

18 (a) The prestack LSRTM image of the VSP data. (b) The corresponding image after SVM filtering.

19 Magnified views of the red boxes in Figure 18. The SVM filtered image in (b) contains fewer artifacts than the conventional pre-stack LSRTM image in (a).

20 The migration velocity model inverted using a plane-wave migration velocity analysis method.

21 (a) The prestack LSRTM image contains significant aliasing artifacts caused by the sparse acquisition geometry. (b) The corresponding image after SVM filtering.

22 The shot domain common image gather at (a) x = 0.7 km, (b) x = 3.5 km and x = 5.6 km, respectively. The shallower artifacts emphasized by the red boxes shows a similar characters compared with real image component.

23 Magnified views of the red boxes in Figure 21. The image in (b) obtained using SVM filtering contains significant fewer artifacts than the conventional pre-stack LSRTM image in (a).

24 Magnified views of the black boxes in Figure 21. The image in (b) obtained using SVM filtering contains significant fewer artifacts than the conventional pre-stack LSRTM image in (a).

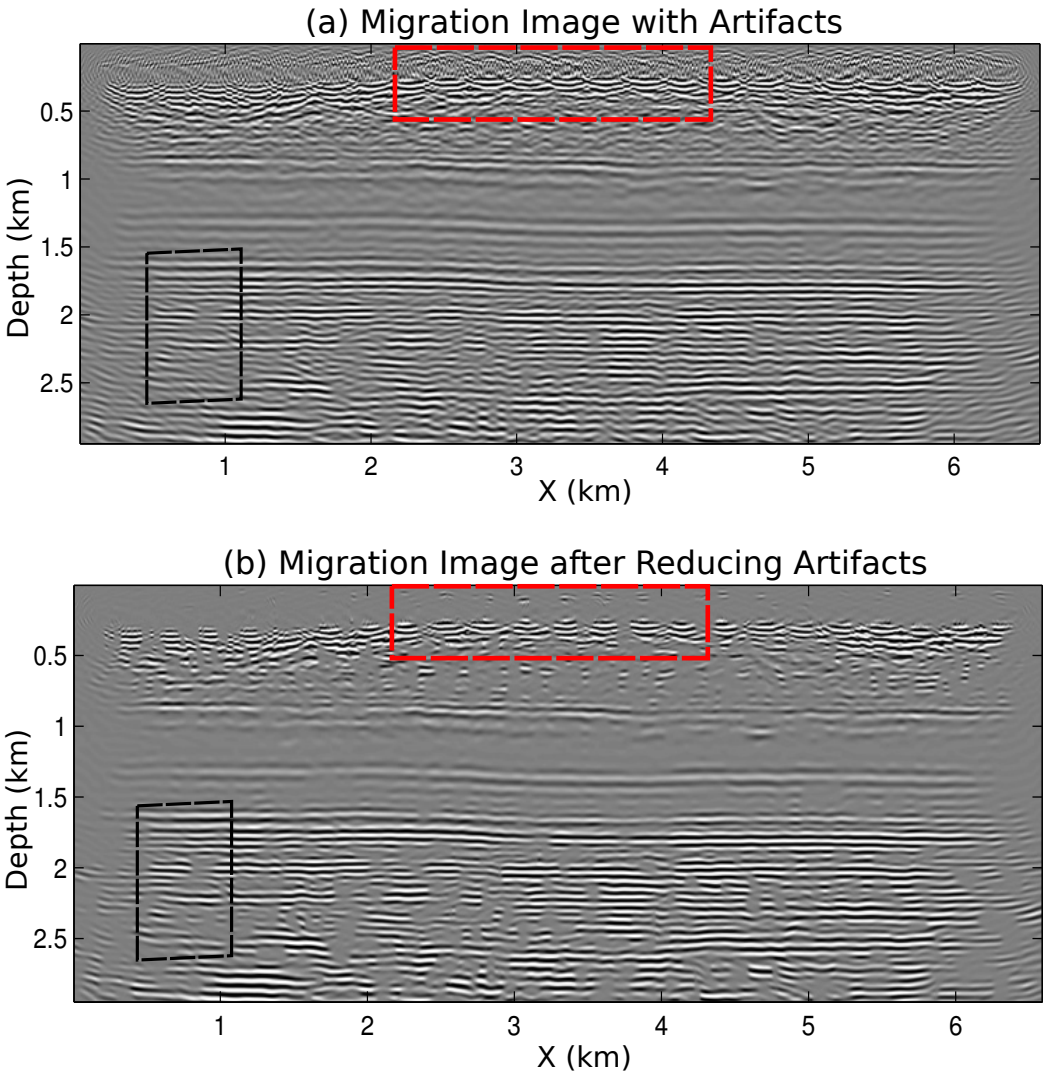


Figure 21: (a) The prestack LSRTM image contains significant aliasing artifacts caused by the sparse acquisition geometry. (b) The corresponding image after SVM filtering.

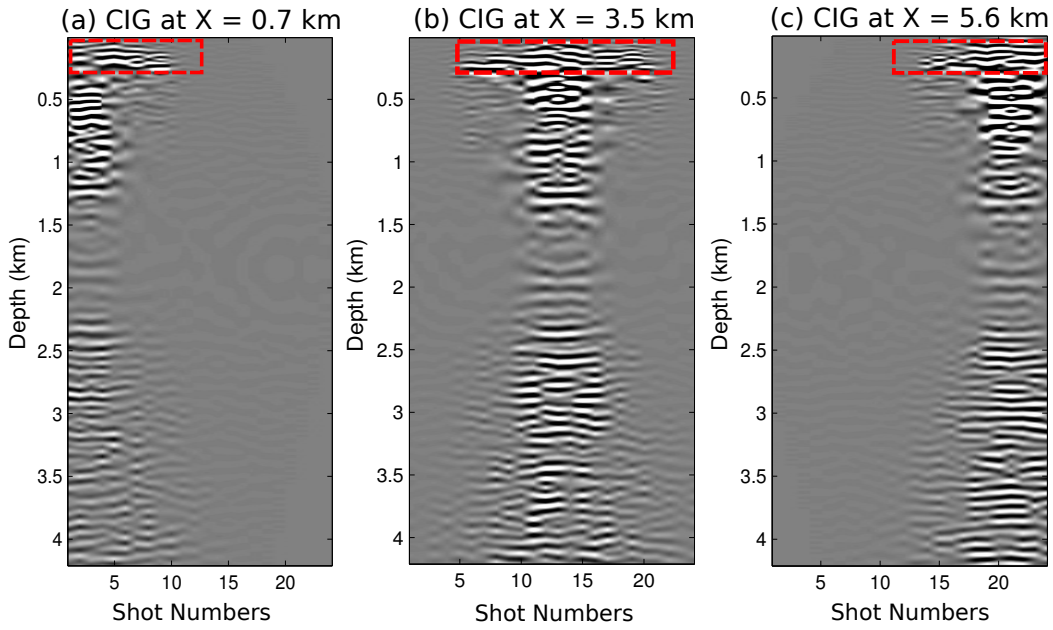


Figure 22: The shot domain common image gather at (a) x = 0.7 km, (b) x = 3.5 km and x = 5.6 km, respectively. The shallower artifacts emphasized by the red boxes shows a similar characters compared with real image component.

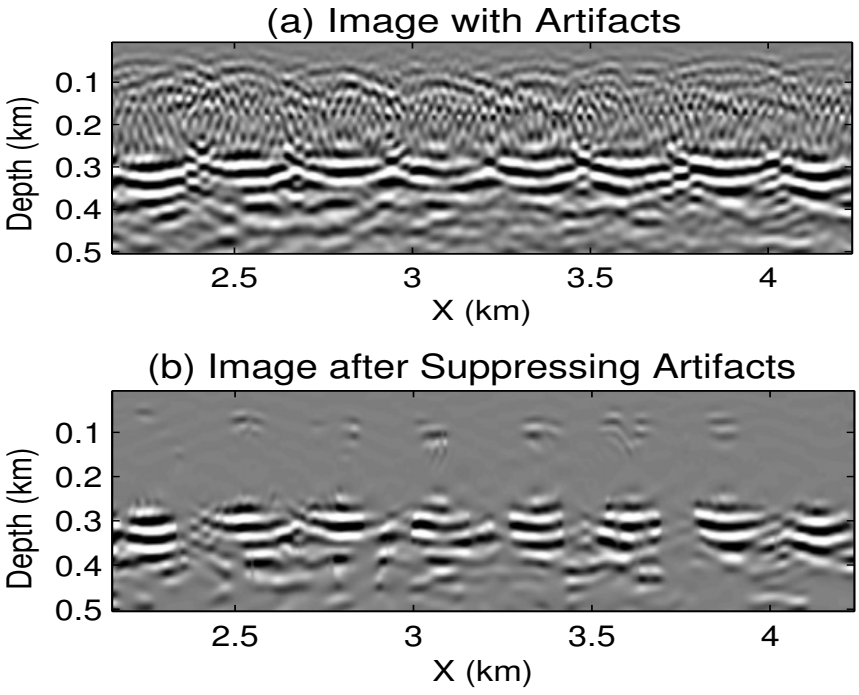


Figure 23: Magnified views of the red boxes in Figure 21. The image in (b) obtained using SVM filtering contains significant fewer artifacts than the conventional pre-stack LSRTM image in (a).

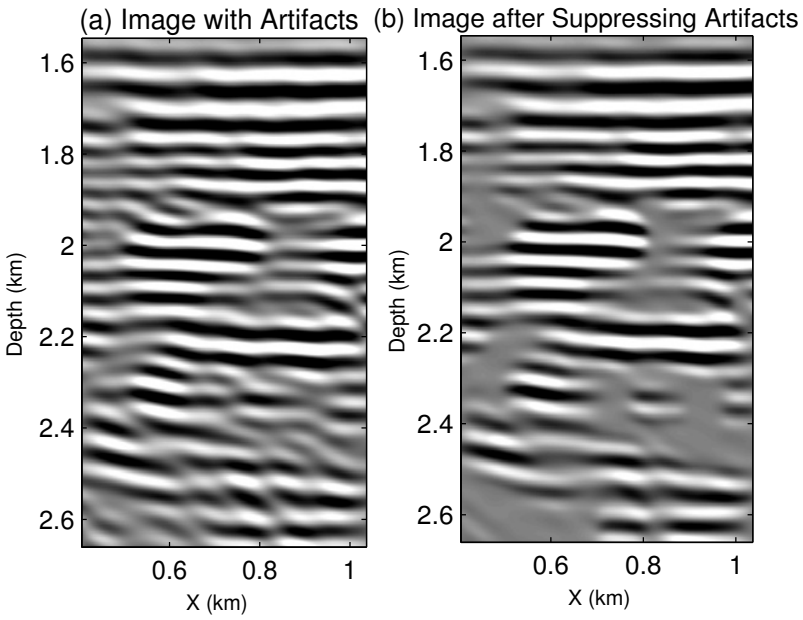


Figure 24: Magnified views of the black boxes in Figure 21. The image in (b) obtained using SVM filtering contains significant fewer artifacts than the conventional pre-stack LSRTM image in (a).