Auto-windowed Super-virtual Interferometry via Machine Learning: A Strategy of First-arrival Traveltime Automatic Picking for Noisy Seismic Data

Kai Lu and Shihang Feng*
King Abdullah University of Science and Technology

SUMMARY

Supervirtual interferometry (SVI) was developed to significantly enhance the signal-to-noise ratio of noisy first arrivals. However, a time window must be specified that contains these first arrivals, and the window should be no wider than several times the dominant period of the source wavelet. The accurate specification of this window is very challenging for noisy data and involves manual picking. To overcome this problem, we propose to automatically pick these windows via machine learning methods. Convolutional neural network (CNN) and density-based spatial clustering of applications with noise (DBSCAN) are used to distinguish first-arrival signals completely buried in noise. Numerical tests validate that this method can accurately specify the correct window as well as that of a human interpreter. The benefit is an automatic means for picking first-arrival traveltimes in noisy traces from a large 3D data set.

INTRODUCTION

First arrival traveltime tomography (Bishop et al., 1985) is one of the most important applications for velocity model building. However, manual picking the first breaks can be very labor-intensive. To solve this problem, many auto-picking methods have been proposed (Allen, 1978; Coppens, 1985; Zhang. et al., 2003; Song et al., 2010; Chen, 2018), yet these methods become less robust for data with high noise levels. This disadvantage is becoming more intolerable since wider source-receiver offsets are needed for imaging deeper targets.

One way to increase the performance of auto-picking methods is to enhance the signal-to-noise ratio (SNR) of the data. This can be accomplished with super-virtual interferometry (Bharadwaj et al., 2011; Mallinson et al., 2011), which is a robust and effective SNR enhancer for far-offset head waves. The general idea of SVI is to construct and stack super-virtual refractions to improve the SNR of coherent refractions. Figure 1 describes the workflow for creating super-virtual refractions in far-offset traces, where the SNR is significantly improved. One of the requirements of SVI is that a time window must be accurately specified that is no wider than about twice the period of the dominant wavelet and it contains the hidden first arrivals. Accurate specification becomes more challenging with noisier data and often involves manual editing.

To overcome manual specification of the time window, we propose a noise-resistant automatic windowing strategy for SVI with machine learning. Results with noisy data suggest that they are at least equal in effectiveness to manual specification of the correct time window.

In the following, I will first explain the methodology, which includes two steps: classification and window parameter determination. Two machine learning methods: convolutional neural network (CNN) and density-based spatial clustering of applications with noise (DBSCAN) are applied for the classification. After that, I present several examples of automatic window picking for seismic data with different noise levels. The conclusions are drawn at the end.

MACHINE LEARNING METHODOLOGY FOR AUTOMATIC WINDOW SELECTION

We automatically pick the temporal windows for first arrivals using two steps: 1) signal/noise classification and 2) window parameter determination. In the first step, subimages are extracted from the seismic data display and each subimage is classified as either signal (coherent events) or random noise by the machine learning algorithm. For the purpose of more robust classification, the subimages are pre-processed to be more noise-resistant before they are used as inputs for the machine learning classifiers. In the second step, the window parameters are determined by the time and velocity information obtained from the subimages that are classified as signals at early traveltime.

Signal/Noise Classification

Pre-processing
To better distinguish the signals from the noise, we identify the coherent features of the seismic events. A local slant-stack is applied to each subimage with different scanning velocities. The stacking velocity that gives the highest energy is used to create the signal in this local window (see Figure 2).

Classification by CNN

A convolutional neural network (CNN) is a class of deep, feed-forward artificial neural networks, which uses a variation of multilayer perceptrons designed to require minimal preprocessing (Krizhevsky et al., 2012; LeCun et al., 2015). CNNs have attracted a lot of attention recently, because of its robustness in many fields like image and video recognition, recommender systems and natural language processing.

In our application, CNN is designed as a binary classifier to distinguish signals from noise for the subimages after local slant stacking. To train or invert the CNN filters, the training and validation sets are prepared where the the signal examples are picked from subimages with identifiable first arrivals, and the noise examples are randomly selected from the subimages before the traveltimes of first arrivals. To avoid biasing the CNN weights, we choose the same number of signal and noise examples for training. It is not recommended to pick

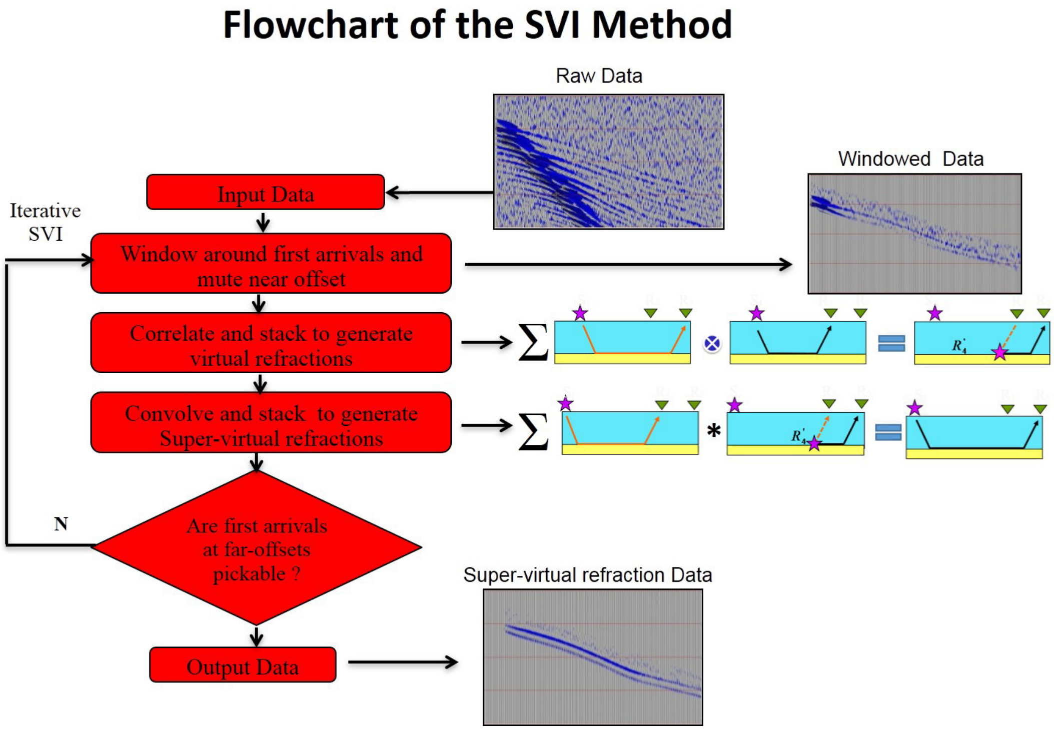

Figure 1: Diagram outlining the main steps for super-virtual interferometry. The reconstructed first arrivals are significantly enhanced compared to those in the original data.

very weak first arrivals as the training examples, as it might lead CNN to misidentify noise as signals in the testing set.

Classification by DBSCAN

Density-based spatial clustering of applications with noise (DBSCAN) is a data clustering algorithm (Martin et al., 1996). It is a density-based clustering algorithm: given a set of points in some space, it groups together points that are closely packed together (points with many nearby neighbors), and marks the outlier points that lie alone in low-density regions (whose nearest neighbors are too far away).

By masking the pixels with low absolute amplitudes in the slant-stacked data, the points in the areas where the energy is focused will form a dense cluster, which can be recognized by the DBSCAN algorithm. There are only a few of these clusters for the subimage with coherent events since such events will only be focused at the correct apparent velocity in the slant stack result. On the other hand, for the subimage with random noise, the energy in the slant-stacked data is dispersive, so that there can be many clusters randomly distributed. This phenomenon provides a simple way to distinguish between signals and noises based on the number of clusters as shown in Figure 3.

Window parameter determination

A classification map of signals and noise is obtained after the above step. The subimages classified as signals at early times delineate the approximate arrival times of the first arrivals. In order to properly select the windows that isolate the first arrivals, the center times and the slopes of the local windows (we assume a local window with constant dipping angle and temporal width) need to be determined. We search the center point of the energy focus zone in the slant-stacked data. The horizontal coordinate of this point gives the apparent velocity and the vertical coordinate is the center time of the window.

NUMERICAL RESULTS

A synthetic test is conducted by applying the proposed method to auto-window an undulating event with three different SNRs of -9.3 dB, -16.8 dB and -19.0 dB, respectively. Here the SNRs are much lower than 0 dB, because the proposed method is for data with high noise levels. Subimages with a size of 91 pixels along the temporal axis (approximately 2 periods) by 10 traces along the offset axis are extracted from the data examples, and the steps between adjacent subimages are 20 pixels along the time axis and 5 traces along the offset axis. Both CNN and DBSCAN algorithms are tested for classification.

In the CNN classification, the subimages are used as input patches. The CNN architecture consists of two convolutional layers and two fully-connected layers. At each convolutional layer, there are 8 convolutional filters with size of 2×2. We used the softmax function to give the possibility of each subimage to be signal or noise.

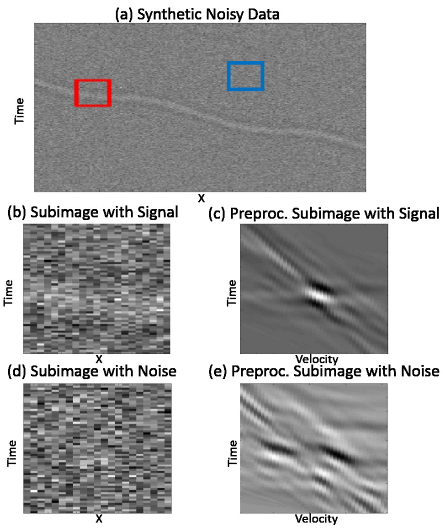

Figure 2: Illustration of pre-processing results for classification. (a) A noisy data example, in which subimages with the signal (red) and another with the noise (blue) are picked. The subimages with signal (b) and noise (d) are pre-processed as shown in the panels (c) and (e), where each trace is a slant-stacked trace associated with the optimal stacking velocity. (c) is obviously distinguished from (e) as its energy is well focused.

The training and validation sets include 120 signal examples and 120 noise examples picked from the data with SNRs of -9.3 dB and -16.8 dB. Examples are not picked from the most noisy data to avoid the classification system confusing extremely weak signals with noise. 80% of the examples are used for training and the rest for validation. The accuracy of the training and validation sets are 98.33% and 97.5%, respectively. Figures 4 (d)$\sim$(f) show the CNN classification results, where the yellow pixels represent the subimages identified as signals, and the rest are noise.

In the DBSCAN classification, the masking threshold is set to be 0.5 (points with amplitudes smaller than 0.5 after normalization are muted). The subimages with the number of clusters less than 4 are classified as signals. Figures 4 (g)$\sim$(i) show the classification results with DBSCAN.

Both methods perform well in the first two examples with relatively higher SNRs. In the last data example, where the event is barely noticeable, both methods suffer from strong noise, suggesting that the noise tolerance level of the proposed classification method does not exceed that of a human interpreter. For comparison, CNN beats the DBSCAN method for extremely noisy cases.

Figure 5 (c) shows the final result after the window parameter determination step for data example 2, where the weak event is correctly windowed.

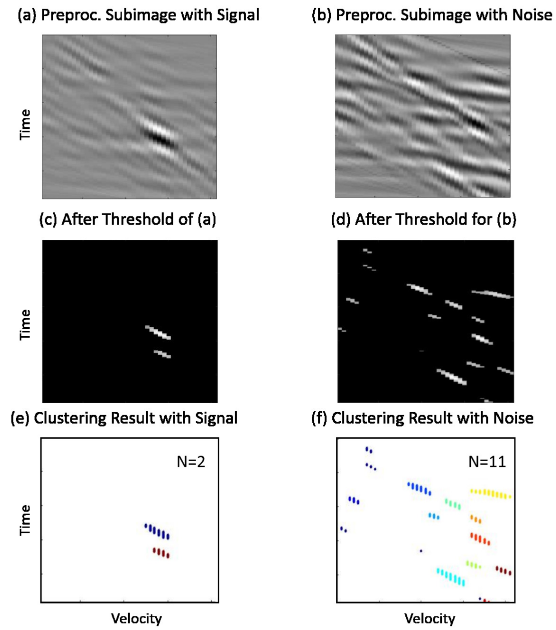

Figure 3: Illustration of classifying subimages by DBSCAN. (a) and (b) are pre-processed subimages with signals and noise, respectively. All the pixels with amplitudes below the threshold are muted as shown in (c) and (d). (e) and (f) demonstrate the clustering results, where the number of clusters is small for signals but much larger for noise.

CONCLUSIONS

We propose a strategy for automatically picking the first breaks of noisy seismic arrivals by enhancing the SNR with auto-windowed SVI. We design the auto-windowing process as two steps: subimage classification and window parameter determination. The subimage classification can be fulfilled by either of the two machine learning methods: the supervised CNN method and an unsupervised learning method (DBSCAN). The numerical tests show that both CNN and DBSCAN performances are satisfactory when the SNR is not lower than -17 dB. Our results showed that CNN beats DBSCAN for data with extremely low SNR. However, DBSCAN, as an unsupervised learning method, requires no prior labels or training, we suggest using DBSCAN in most cases, unless the data are heavily polluted by noise.

The limitation of the proposed first arrival auto-windowing method is that it assumes that the additive noise is random. If this assumption is not satisfied, the coherent noise appearing at earlier times than first arrivals need to be masked or filtered. For very low SNRs, in which some parts of the event can not be identified (such as the cases in Figures 4(f) and (i)), interpolation can be an option.

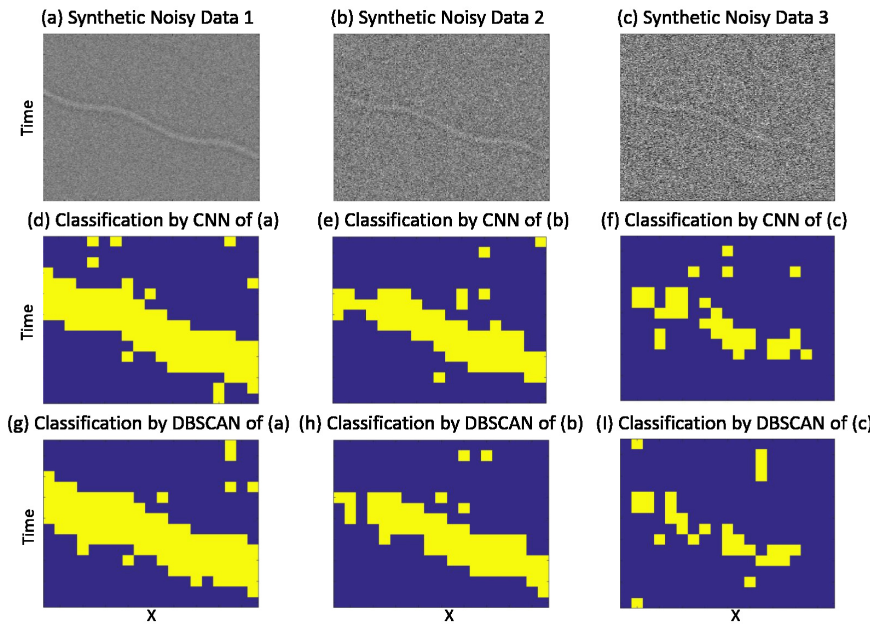

Figure 4: The classification results for three data examples with different SNRs: (a) -9.3 dB, (b) -16.8 dB and (c) -19.0 dB. (d) (f) and (g) (I) are classification results from CNN and DBSCAN, respectively, where the yellow pixels represent the subimages identified as signals and the blue pixels represent the subimages identified as noise.

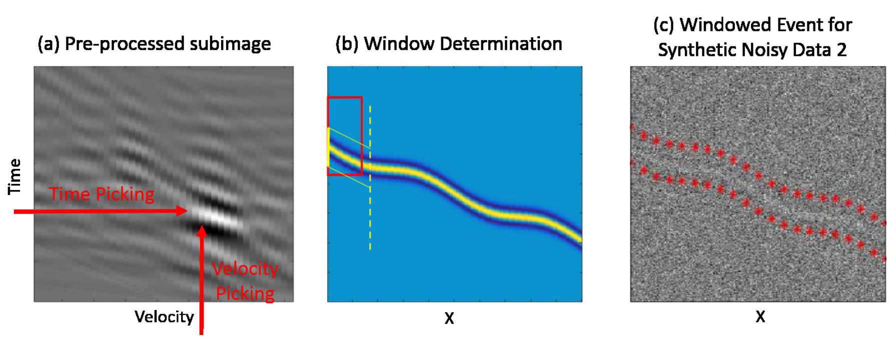

Figure 5: Diagram illustrating how to determine the window for a subimage identified as signal. The relative temporal position and velocity picked in (a) determines the central time of the window (the most left vertical yellow bar) and the window slope in (b), where the red box represents the subimage position. (c) shows the windows (marked by the red dots) automatically picked for the synthetic data with a SNR of -16.8 dB.

ACKNOWLEDGMENTS

We thank the CSIM members for supporting this research, and we also thank the High Performance Computational Center and IT support at KAUST.

REFERENCES

Allen, R., 1978, Automatic earthquake recognition and timing from single traces: Bull. seism. Soc. Am., **68**, no. 5, 1521–1532.

Bharadwaj, P., G. T. Schuster, and I. Mallinson, 2011, Theory of supervirtual refraction interferometry: Geophysical Journal International, **288**, 263–273.

Bishop, T., K. Bube, R. Cutler, R. Langan, J. Love, P.and Resnick, R. Shuey, D. Spindler, and H. Wyld, 1985, Tomographic determination of velocity and depth in laterally varying media: The Leading Edge, **30**, 546–550.

Chen, Y., 2018, Automatic microseismic event picking via unsupervised machine learning: Geophysical Journal International, **212**, no. 1, 88–102.

Coppens, F., 1985, First arrival picking on common-offset trace collections for automatic estimation of static corrections: Geophys. Prospect., **33**, no. 8, 1212–1231.

Krizhevsky, A., I. Sutskever, and G. E. Hinton, 2012, Imagenet classification with deep convolutional neural networks: NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems, **1**, 1097–1105.

LeCun, Y., Y. Bengio, and G. Hinton, 2015, Deep learning: Nature, **521**, 436–444.

Mallinson, I., P. Bharadwaj, G. Schuster, and H. Jakubowicz, 2011, Enhanced refractor imaging by supervirtual interferometry: The Leading Edge, **30**, 546–550.

Martin, E., K. Hans-Peter, S. Jrg, and X. Xu, 1996, A density-based algorithm for discovering clusters in large spatial databases with noise: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), 226–231.

Song, F., H. Kuleli, M. Toksoz, E. Ay, and H. Zhang, 2010, An improved method for hydrofracture-induced microseismic event detection and phase picking: Geophysics, **75**, no. 6, A47–A52.

Zhang., H., C. Thurber, and C. Rowe, 2003, Automatic p-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings: Bull. seism. Soc. Am., **93**, no. 5, 1904–1912.