Automated ambient noise processing applied to fiber optic seismic acquisition (DAS)

Fantine Huot, Eileen R. Martin and Biondo Biondi

ABSTRACT

Distributed acoustic sensing (DAS) is an emerging technology used to record seismic data that employs fiber optic cables as a probing system. Since September 2016, a DAS array has been deployed beneath Stanford campus in the existing fiber optic telecommunication conduits. Because we can so easily use our telecomm infrastructure for continuous, dense, seismic acquisition, data collected in such a manner will go to waste unless we significantly automate ambient noise processing. Herein we present relevant data features for exploratory data analysis and identify coherent noise sources which inhibit reliable extraction of useful signals. We then train a convolutional neural network for detecting traffic noise and selectively filter it out to generate ambient seismic noise fields that are suitable for interferometry purposes. Further, we use Markov decision processes to reconstruct the array geometry from the data, which gives us the potential to extend this type of acquisition to other existing fiber optic networks.

INTRODUCTION

By measuring the speed of seismic waves propagating in the Earth's near-surface, we can image the top tens to hundreds of meters of the subsurface. These seismic velocity images can be interpreted to evaluate earthquake or landslide risk, to find sinkholes or tunnels, or to track near-surface changes related to drilling activities.

By cross-correlating noise recorded at a selected receiver with noise recorded by all other receivers in an array, we can extract signals mimicking an active seismic survey, using the virtual source function method (Shapiro and Campillo, 2004; Lin et al., 2008; Wapenaar et al., 2010). Thus, when active sources are too costly or logistically prohibitive, passive seismic can be a good option for near-surface imaging. However, the theory is limited by the assumption of homogeneous uncorrelated sources (Wapenaar et al., 2010) and non-ideal sources can cause artifacts in extracted velocities.

Distributed acoustic sensing (DAS) is a new acquisition technology being increasingly adopted in the energy industry for microseismic monitoring (Webster et al., 2013), hydraulic fracture monitoring (Bakku, 2015), CO_2 sequestration observation (Daley et al., 2013), and time-lapse imaging of reservoirs (Mateeva et al., 2013). DAS probes a fiber-optic cable with a laser interrogator unit (IU) to repurpose that fiber as a series of strain sensors. By running fibers in existing telecommunication conduits, easy, on-demand, repeatable seismic studies (even in urban areas) will soon be a reality.

However, no matter how long we average cross-correlations, coherent repeating noises, even weak ones, can lead to artifacts in virtual-source response estimates (Martin et al., 2016), but it is currently time consuming and subjective for an expert to comb through days of passive data and decide how to filter out certain noises. To this end, herein we use an unsupervised learning approach to quickly explore the ambient noise field (Huot et al., 2017; Martin et al., 2018) and show how to use interpreted results to define a targeted noise filtering process using a convolutional neural network to remove traffic noise from the recorded wave field.

Mapping the recorded data to its spatial coordinates is also a costly and time-consuming task, requiring tap tests and active surveys. For the Stanford DAS array it took several weeks of iterating between performing tap tests and manually inspecting the data to assign locations to channels (Martin et al., 2017), and to this day, the assigned geometry may not accurately reflect some of the small features deviating from straight line paths. This manual process is not scalable, and would be a potential bottleneck in broader deployment of urban DAS arrays. In the last section, we present a methodology combining a convolutional neural network and Markov decision processes, to automatically retrieve the array geometry from the data. This approach would allow us to extend this type of acquisition to other existing fiber optic networks.

THE STANFORD DAS ARRAY

We present this methodology in the context of a case study: a figure-eight-shaped array of 2.4 km of fiber optics lying loosely in existing telecommunications conduits underneath the Stanford University campus (Figure 1). The array detects a wide variety of seismic noise sources that do not conform to the ideals of existing ambient noise theory: it sits in a seismically active region, 20 km from the Pacific ocean, 7 km from the San Francisco bay, with major highways on either side, a variety of roads with differing levels of traffic near the fiber, regular quarry blasts within 15 km, plumbing and HVAC systems throughout the site, multiple construction sites near the array, and foot and bicycle traffic throughout. With over 600 channels continuously recording 50 samples per second since September 2016, manual inspection of most data is infeasible, making automation tools critical to extracting the full value from the data.

In this study, we present data from a subset of 300 channels. The array generates contiguous time series that conveniently lend themselves to image processing (Figure 2). However, from a seismological perspective, the data are poorly coupled with the ground as the fiber is laid in narrow (10-15 cm wide) conduits buried under-

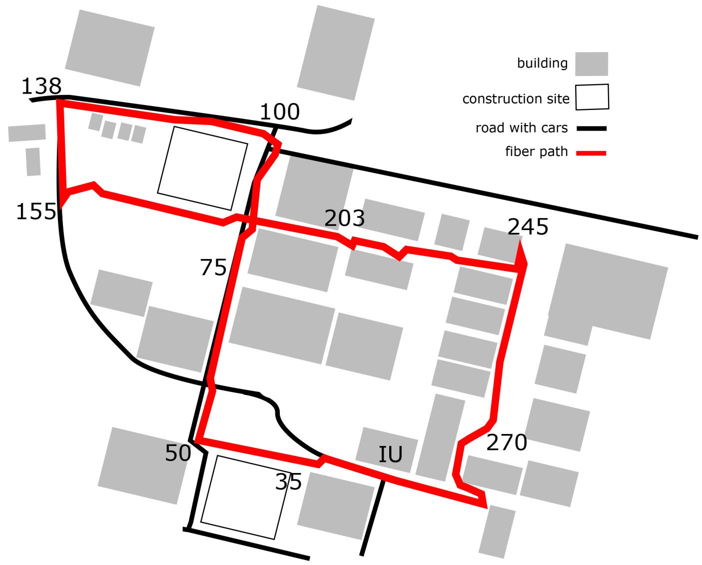


Figure 1: The fiber optic DAS array, 600 meters along its widest direction, follows a figure-eight path forming two rectangular arrays under the Stanford University campus. [**NR**]

ground, and a handful of channels must pass manholes between adjacent conduits. The channels in manholes are mostly freely hanging. We preprocessed the data by detrending and bandpass filtering (0.5 and 20 Hz).

AUTOMATICALLY IDENTIFYING DIFFERENT TYPES OF SEISMIC NOISE

As manual inspection of large, complex data volumes is infeasible, seismic interpretation requires new processing tools for event detection, signal classification and data visualization. Machine learning techniques have recently been introduced for some seismic applications to automate decisions and speed up processing. Self-organizing maps (SOM), an unsupervised learning technique, have been used to detect volcano-tectonic and rockfall events from long-term background variations (Köhler et al., 2010). Support vector machines (SVM), a supervised learning method, have been used for earth dam and levee health monitoring and automatic detection of anomalous events (Fisher et al., 2016). Data mining algorithms have been adapted for computationally efficient earthquake detection by similarity search on large scale datasets (Yoon et al., 2015).

To identify a variety of common wavefield patterns, we performed unsupervised learning on a subset of the DAS data following the methodology introduced by Huot et al. (2017) and Martin et al. (2018). We used clustering algorithms on seven days of data to capture the daily variations in the noise field. Once the main types of patterns are identified and have been grouped into clusters, the trained algorithm can be

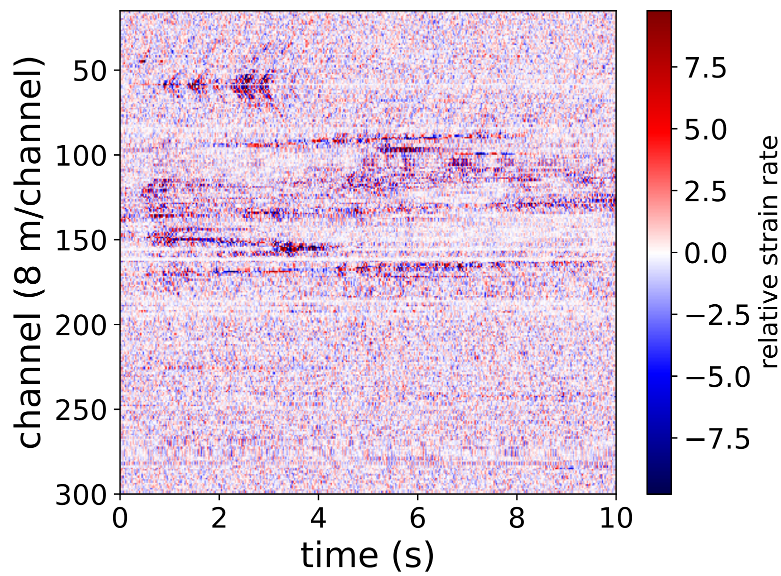


Figure 2: Data recorded by the Stanford DAS array after detrending and bandpass filtering (0.5 and 20 Hz). Ambient noise in urban areas is far from white or spatially uniform. In just ten seconds of noise collected around noon local time we can see significant variation throughout the array. [**ER**]

applied to new data to classify them within the identified clusters. This approach allows us to automate data exploration with the aim of speeding up the overall ambient noise workflow and reducing human bias. Seismic noise source identification constitutes the first step toward selective noise removal and preprocessing for interferometry purposes.

We computed data features over a week of continuous data that capture the temporal and spatial variations of the signal using continuous wavelet transforms (CWT). The CWT is commonly used in pattern recognition, as it has the ability to decompose complex patterns into elementary forms (Mallat, 2008). The CWT measures the similarity between a signal and an analyzing wavelet by comparing the input signal to shifted and compressed or stretched versions of the wavelet.

Common clustering methods for wavelet domain time series include k-means, agglomerative clustering and self-organizing maps (Liao, 2005; Köhler et al., 2010). For computational efficiency, we selected the mini-batch optimization implementation for k-means clustering (Sculley, 2010). We experimented with different numbers of clusters and empirically settled on four main clusters (Figure 4), as more clusters merely yielded subdivisions of these main clusters that were not always more physically interpretable.

By examining cluster counts over time and examining the corresponding CWT coefficients, it appeared that the largest cluster (blue in Figure 4) corresponded to ambient background noise. The small red cluster showed diurnal trends and appeared

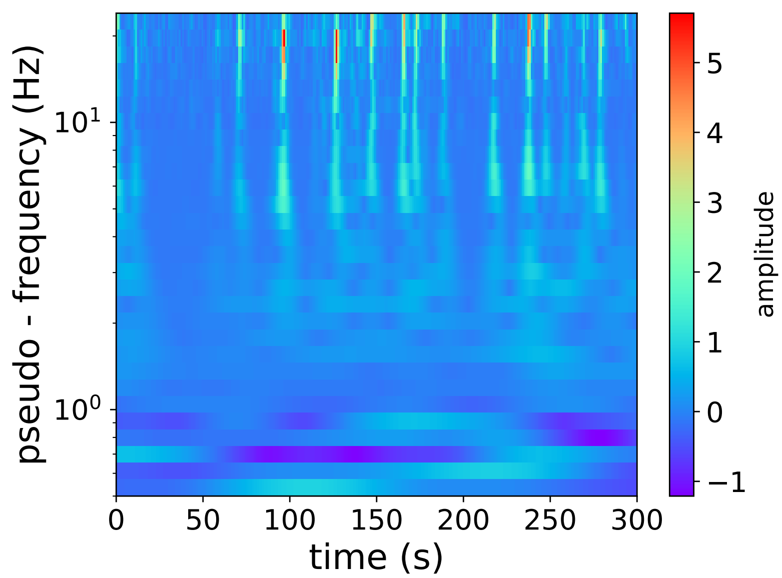


Figure 3: Continuous wavelet transform (CWT) coefficients computed over time, on 5 minutes of data. We can distinguish car arrivals as high amplitude streaks over the CWT coefficients. [**CR**]

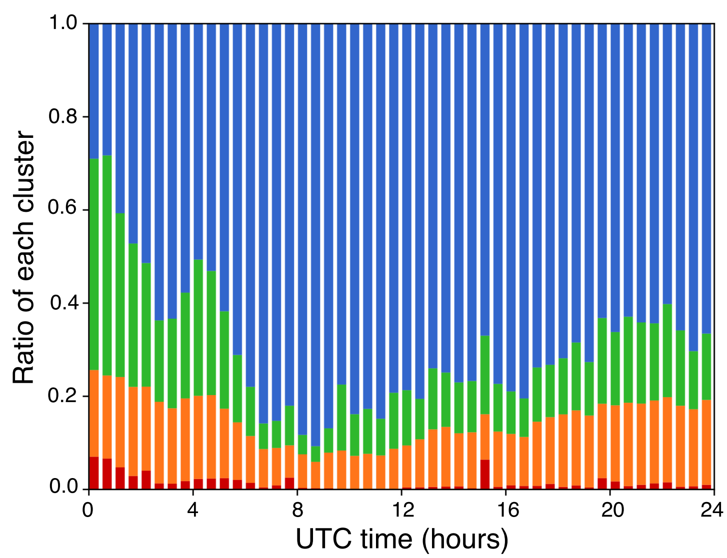
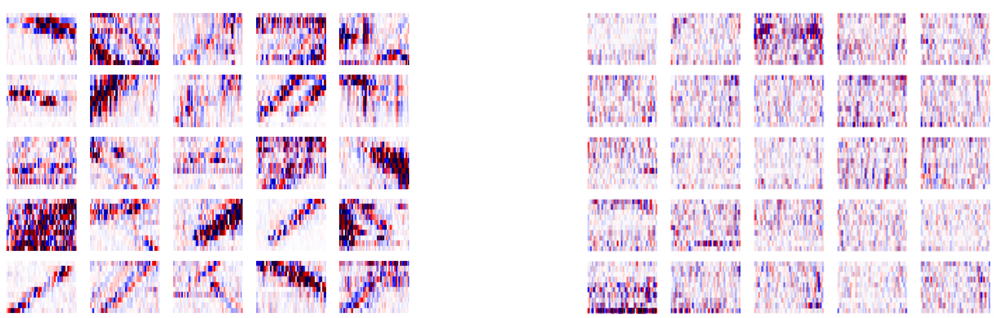


Figure 4: The different types of identified clusters averaged over the full DAS array for one day. The largest cluster, plotted in blue, corresponds to ambient background noise. The small red cluster shows diurnal trends and appears mostly on roads open to cars, and appears to be due to vehicle noise. [**CR**]

mostly on main roads open to cars, so it was interpreted as nearby vehicle noise. The orange cluster behaved similarly to red but yielded more diffuse CWT coefficients and was more spread in space. We interpreted it as noise associated with vehicle traffic but not necessarily showing the particular space-time pattern of nearby, individual cars. The green cluster is another source of coherent noise with diurnal trends, possibly linked to construction noise.

TRAFFIC NOISE DETECTION USING A NEURAL NETWORK

The methodology described in the previous section allows us to automatically detect traffic noise in the ambient noise field. However, calculating wavelet attributes can be computationally expensive and is not suited for real-time detection on streaming data. Therefore, we proceeded to build a neural network for detecting traffic noise. While training neural networks can be computationally involved, they perform remarkably fast at run-time and are hence well-suited for event detection problems.



(a) (b)

Figure 5: (a) Examples of data windows containing car traffic noise. (b) Examples of data windows of background ambient noise. [**CR**]

We divided the data into detection windows of 10 channels by 10 seconds, and labeled each window as ambient or traffic noise using the methodology from the previous section (Figure 5). The windows were downsampled along the time axis, resulting in arrays of 10×50.

All the windows were shuffled to avoid any time bias and normalized with the same means and standard deviations. The data sets were then arranged as follows: 50,000 detection windows for the training data, and 10,000 windows for the testing data.

A basic fully-connected 2-layer softmax classifier achieved poor accuracy on this data set (55%). In order to leverage the grid-like topology of the DAS data, we designed a small convolutional neural network:

1. 5×10 convolution layer with rectified linear unit (ReLu) activation function
2. Max pooling 1×2
3. 5×5 convolution layer with ReLu activation function
4. Max pooling 2×2
5. Fully connected layer with ReLu activation function
6. Dropout with a probability of 0.5
7. Fully connected output layer

The network architecture is presented in Figure 6. A description of the different components of the convolutional network is provided in Huot (2018).

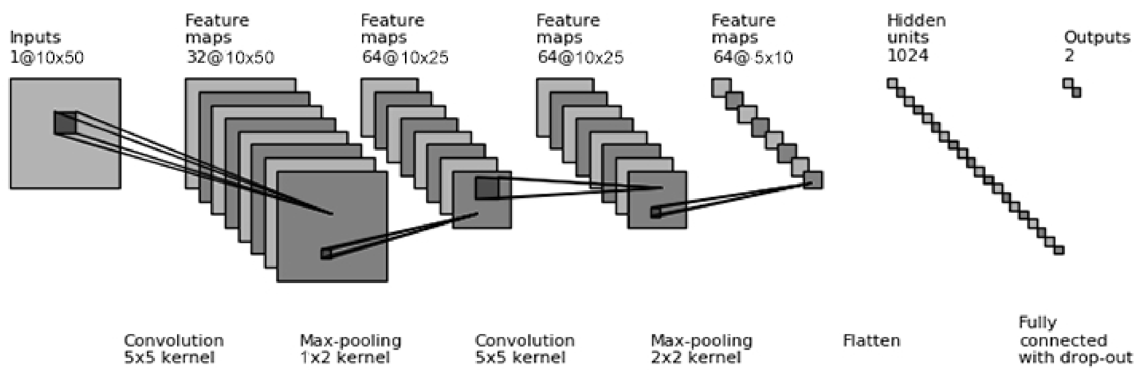


Figure 6: We used a small convolutional neural network classifier to detect traffic noise in the DAS data. [**ER**]

This classifier achieved 99.4% accuracy. Out of 5000 cars, 38 obtained a probability score less than 90%, 59 less than 95%. Figure 7 presents some examples of car windows to which the network gave lower probabilities. These examples correspond mostly to extremely faint cars or very noisy data portions.

When comparing with existing literature, other seismic event detectors (Fisher et al., 2016; Rubin et al., 2012; Ruano et al., 2014; Lan et al., 2005) reported accuracies ranging between 74% and 94% with traditional geophones and using other detection methodologies. In particular, (Lan et al., 2005) focuses on moving ground targets and achieves about 80% accuracy.

SELECTIVE FILTERING WITH CONTINUOUS WAVELET TRANSFORMS

We then used the classifier to build an automated targeted noise filtering process to remove traffic noise from the recorded wave field. For each window with detected

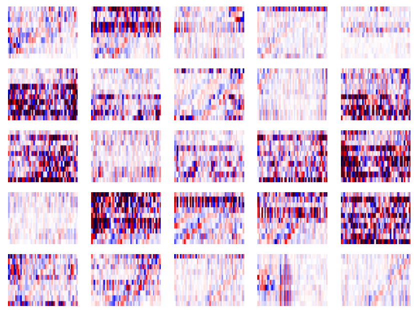


Figure 7: Examples of data windows containing traffic noise to which the network gave lower classification probabilities. [**CR**]

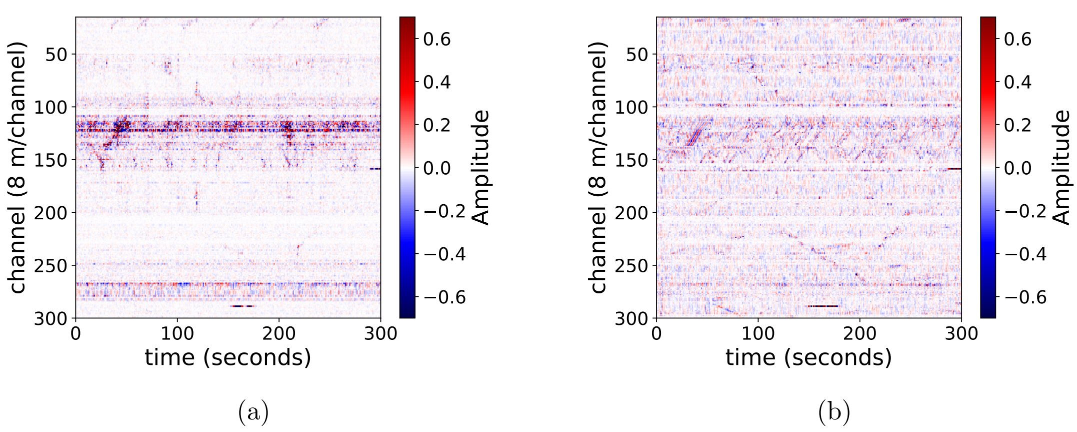


Figure 8: A comparison of data over five minutes of data with many examples of cars (a) before and (b) after muting out detected cars. [**CR**]

traffic noise, we mute out the high amplitude scales among the CWT scales computed over time. We then performed inverse CWT over time to reconstruct the signal. A comparison of data with many examples of cars before and after muting out detected cars (Figure 8) shows that not all types of noise were removed, but the impact of cars was reduced.

GEOMETRY MAPPING USING MARKOV DECISION PROCESSES

Mapping the recorded data to its spatial coordinates is a costly and time-consuming task, requiring tap tests and active surveys. This manual process is not scalable, and would be a potential bottleneck in broader deployment of urban DAS arrays. In this section, we explore the possibility of retrieving the array's geometry directly from the recorded data.

Using the geometry information inferred by tap tests as labels, we trained a convolutional neural network to determine whether each portion of data corresponds to a straight portion of the fiber, a turn, or a manhole. We used the first half of sensors as training data and the second half as test data. We trained the neural network over a day of continuous data, divided into overlapping windows of 5 channels over 20 seconds. The labels were discretized, with turning angles ranging from 0 to 180° divided into 10° bins, and an additional label for manholes. We used the VGG-net architecture and pre-trained weights as convolutional neural network classifier (Simonyan and Zisserman, 2014), and mapped each portion of the data to a normalized probability of being a straight line, a turn or a manhole. While the accuracy obtained on the test data was fairly low (73%), the misclassification was often due to incorrect turning angle.

With uncertain labels, geometry mapping is a challenging task, and the number of possible geometry mappings increases. Moreoever, the angle labels for the classifier only range from 0 to 180° because there is no way of telling directly from the data whether the fiber is turning left or right. To overcome this problem, we built a Markov decision process to reconstruct the array geometry.

Markov decision processes (MDP) provide a mathematical framework for modeling decision making in situations where outcomes are partly random and partly under control of a decision maker (Bellman, 1957). They are used in a wide area of disciplines, including robotics, automatic control, and artificial intelligence.

More precisely, a Markov decision process is a discrete time stochastic control process. At each step, the process is in some state s, and the decision maker may choose any action a that is available in state s. The process responds at the next time step by randomly moving into a new state s', and giving the decision maker a corresponding reward $R_a(s, s')$. The probability that the process moves into its new state s' is influenced by the chosen action. Specifically, it is given by the state transition function $P_a(s, s')$. Thus, the next state s' depends on the current state s and the decision maker's action a.

For our application, we used an MDP to map the geometry of the array, starting at the first receiver channel, and determining the relative positions of all the contiguous channels until reaching the last channel. The successive states in the MDP represent the channels along the fiber. We constrained the mapping by defining a guiding path: a list of coordinates defining an initial guess as to where the fiber lies, with a certain error margin. We then defined the MDP as follows:

- States: Coordinates (x, y) of current position, and (dx, dy) of current direction
- Starting state: initial position, initial direction
- Actions: Go straight, turn, or manhole
- Transition probabilities: Probabilities computed by the aforementioned convolutional neural network classifier. When turning, we set an equal probability for turning left or right. We added a small exploration probability (0.1) to all turn directions, and normalized all the probabilities.
- Reward function: We considered several options for defining the reward function. If the total number of turns and manholes is known, we can define the reward function as a metric of how close the total number of turns and manholes is to their actual number. If we know that the fiber follows a loop pattern, we can define the reward function as the distance between the computed position of the last channel and the initial position. In our case, we defined the reward function as the mean square error distance between the computed path and the initial guiding path.
- End state: We defined to end states: the MDP reaches the last channel, or the path went out of the bounds of the error margin around the guiding path.

We solved the MDP using the value iteration algorithm (Bellman, 1957). An example of one the realizations is presented in Figure 9. While there still remains uncertainty in the mapping process, this methodology generates approximate mappings at low cost, since human input is limited to defining the guiding path. The proposed solution using a convolutional neural network and an MDP can compute the mapping of the channels within a minute, while it originally took several weeks of manual labor, iterating between performing tap tests and manually inspecting the data (Martin et al., 2017), to assign locations to the Stanford DAS array channels. Moreover, to this day, the manually assigned geometry may not accurately reflect some of the small features deviating from straight line paths. For future applications, we plan on adding additional functionalities to the MDP, such as the possibility of constraining the mapping process further by defining which receiver channel corresponds to manholes and sharp turns, in order to adapt the proposed algorithm to DAS arrays for which this information is available.

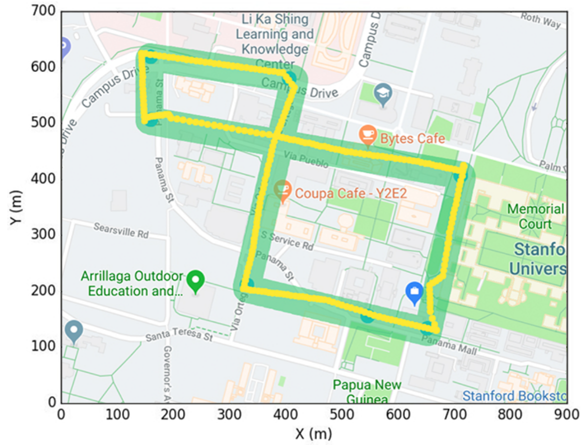


Figure 9: While there still remains uncertainty in the mapping process, the Markov decision process generates approximate mappings at low cost (yellow), since human input is limited to defining the guiding path (green). Background image courtesy of Google Maps. [**CR**]

DISCUSSION AND CONCLUSIONS

We employed unsupervised and supervised learning techniques to characterize seismic data recorded by a fiber optics array deployed underneath the Stanford University campus. We used wavelet attributes for exploratory data analysis, used clustering algorithms to distinguish between different types of noise, automatically separating noise generated by cars from incoherent background noise without requiring any information related to the geometry of the array. We then trained a convolutional neural network classifier to speed up the detection process and build a automated targeted filtering workflow to generate wavefields that are suitable for interferometry. We demonstrated it was possible to retrieve an approximate geometry of the array directly from the data using Markov decision processes, with little human input. This approach opens up the possibility of extending this type of acquisition to other existing fiber optic networks.

ACKNOWLEDGEMENTS

We would like to thank OptaSense for their support of the DAS interrogator unit. We thank the Stanford Center for Computational Earth and Environmental Sciences for computing resources, and Stanford IT and School of Earth Sciences IT for ongoing support of the DAS array.

REFERENCES

Bakku, S. K., 2015, Fracture characterization from seismic measurements in a borehole: PhD thesis, Massachusetts Institute of Technology.

Bellman, R., 1957, A markovian decision process: Journal of Mathematics and Mechanics, 679–684.

Daley, T. M., B. M. Freifeld, J. Ajo-Franklin, S. Dou, R. Pevzner, V. Shulakova, S. Kashikar, D. E. Miller, J. Goetz, J. Henninges, et al., 2013, Field testing of fiber-optic distributed acoustic sensing (das) for subsurface seismic monitoring: The Leading Edge, **32**, 699–706.

Fisher, W. D., T. K. Camp, and V. V. Krzhizhanovskaya, 2016, Anomaly detection in earth dam and levee passive seismic data using support vector machines and automatic feature selection: Journal of Computational Science.

Huot, F., 2018, Convolutional neural networks explained: SEP-Report, **172**.

Huot, F., Y. Ma, R. Cieplicki, E. Martin, and B. Biondi, 2017, Automatic noise exploration in urban areas, *in* SEG Technical Program Expanded Abstracts 2017: Society of Exploration Geophysicists, 5027–5032.

Köhler, A., M. Ohrnberger, and F. Scherbaum, 2010, Unsupervised pattern recognition in continuous seismic wavefield records using self-organizing maps: Geophysical Journal International, **182**, 1619–1630.

Lan, J., S. Nahavandi, T. Lan, and Y. Yin, 2005, Recognition of moving ground targets by measuring and processing seismic signal: Measurement, **37**, 189–199.

Liao, T. W., 2005, Clustering of time series dataa survey: Pattern recognition, **38**, 1857–1874.

Lin, F.-C., M. P. Moschetti, and M. H. Ritzwoller, 2008, Surface wave tomography of the western united states from ambient seismic noise: Rayleigh and love wave phase velocity maps: Geophysical Journal International, **173**, 281–298.

Mallat, S., 2008, A wavelet tour of signal processing: the sparse way: Academic press.

Martin, E., N. Lindsey, S. Dou, J. Ajo-Franklin, T. Daley, B. Freifeld, M. Robertson, C. Ulrich, A. Wagner, K. Bjella, et al., 2016, Interferometry of a roadside das array in fairbanks, ak: Presented at the 2016 SEG International Exposition and Annual Meeting, Society of Exploration Geophysicists.

Martin, E. R., C. M. Castillo, S. Cole, P. S. Sawasdee, S. Yuan, R. Clapp, M. Karrenbach, and B. L. Biondi, 2017, Seismic monitoring leveraging existing telecom infrastructure at the sdasa: Active, passive, and ambient-noise analysis: The Leading Edge, **36**, 1025–1031.

Martin, E. R., F. Huot, Y. Ma, R. Cieplicki, S. Cole, M. Karrenbach, and B. L. Biondi, 2018, A seismic shift in scalable acquisition demands new processing: Fiber-optic seismic signal retrieval in urban areas with unsupervised learning for coherent noise removal: IEEE Signal Processing Magazine, **35**, 31–40.

Mateeva, A., J. Lopez, J. Mestayer, P. Wills, B. Cox, D. Kiyashchenko, Z. Yang, W. Berlang, R. Detomo, and S. Grandi, 2013, Distributed acoustic sensing for reservoir monitoring with vsp: The Leading Edge, **32**, 1278–1283.

Ruano, A. E., G. Madureira, O. Barros, H. R. Khosravani, M. G. Ruano, and P. M. Ferreira, 2014, Seismic detection using support vector machines: Neurocomputing,

135, 273–283.

Rubin, M. J., T. Camp, A. Van Herwijnen, and J. Schweizer, 2012, Automatically detecting avalanche events in passive seismic data: Machine Learning and Applications (ICMLA), 2012 11th International Conference on, IEEE, 13–20.

Sculley, D., 2010, Web-scale k-means clustering: Proceedings of the 19th international conference on World wide web, ACM, 1177–1178.

Shapiro, N. M., and M. Campillo, 2004, Emergence of broadband rayleigh waves from correlations of the ambient seismic noise: Geophysical Research Letters, **31**.

Simonyan, K., and A. Zisserman, 2014, Very deep convolutional networks for large-scale image recognition: arXiv preprint arXiv:1409.1556.

Wapenaar, K., D. Draganov, R. Snieder, X. Campman, and A. Verdel, 2010, Tutorial on seismic interferometry: Part 1basic principles and applicationstutorial on interferometry: Part 1: Geophysics, **75**, 75A195–75A209.

Webster, P., J. Wall, C. Perkins, M. Molenaar, et al., 2013, Micro-seismic detection using distributed acoustic sensing: Presented at the 2013 SEG Annual Meeting, Society of Exploration Geophysicists.

Yoon, C. E., O. OReilly, K. J. Bergen, and G. C. Beroza, 2015, Earthquake detection through computationally efficient similarity search: Science advances, **1**, e1501057.