Shear-wave imaging from traffic noise using seismic interferometry by cross-coherence

Nori Nakata\textsuperscript{1,2}, Roel Snieder\textsuperscript{2}, Takeshi Tsuji\textsuperscript{1}, Ken Larner\textsuperscript{2}, & Toshifumi Matsuoka\textsuperscript{1}

\textsuperscript{1}Department of Urban Management, Kyoto University
\textsuperscript{2}Center for Wave Phenomena, Colorado School of Mines

ABSTRACT

We apply the "cross-coherence method" to seismic interferometry of traffic noise, which originated from roads and railways, to retrieve both body waves and surface waves. Our preferred algorithm in the presence of highly variable and strong additive random noise uses cross-coherence, which uses normalization by the spectral amplitude of each of the traces, rather than cross-correlation or deconvolution. This normalization suppresses the influence of additive noise and overcomes problems resulting from amplitude variations among input traces. By using only the phase information and ignoring amplitude information, the method effectively removes the source signature from the extracted response and yields a stable structural reconstruction even in the presence of strong noise. This algorithm is particularly effective where the relative amplitude among the original traces is highly variable from trace to trace. We use the extracted reflected shear waves from the traffic noise data to construct a stacked and migrated image, and use the extracted surface waves (Love waves) to estimate the shear velocity as a function of depth. This shear-velocity profile agrees well with the interval velocity obtained from the normal moveout of the reflected shear waves constructed by seismic interferometry. These results are useful for a wide range of situations applicable in both geophysics and civil engineering.

Key words: noise, passive seismic, shear wave, near surface, coherence

1 INTRODUCTION

The main purpose of seismic interferometry is to construct a Green’s function between two geophones, hydrophones, or accelerometers through data processing of signals generated by earthquakes, microtremors, cultural noise, or artificial seismic sources. Green’s function extraction can be derived from normal modes (Lobkis and Weaver, 2001), representation theorems (Wapenaar, 2004; Wapenaar and Fokkema, 2006), principle of time reversal (Roux and Fink, 2003), and stationary phase analysis (Snieder et al., 2006). After such processing, one geophone serves as a (virtual) source for waves recorded by other receivers, which leads to a pseudo shot gather for many receivers, without using an active source.

Although the first application of seismic interferometry was based on cross-coherence (Aki, 1957), the first applied algorithm in seismic interferometry that found wide application is based on cross-correlation (Claerbout, 1968; Wapenaar, 2003; Bakulin and Calvert, 2004; Schuster et al., 2004). Another proposed algorithm is based on deconvolution. In this method, the source signal is removed by means of spectral division. The mathematical theory of deconvolution interferometry has been derived by (Vasconcelos and Snieder, 2008a), and the method has been applied to field data (Snieder and Şafak, 2006; Vasconcelos and Snieder, 2008b; Vasconcelos et al., 2008). A multidimensional deconvolution method has been formulated for seismic interferometry (Wapenaar et al., 2008; @warning Citation ‘2008wapenaar’ on page 139 undefined).

The various methods each have both advantages
and disadvantages (Table 1 of Snieder et al., 2009). Cross-correlation, for example, is stable but needs estimation of the power spectrum of the noise source, and deconvolution is potentially unstable and needing regularization but does not require estimation of the source spectrum. We should choose the method that best suits the data; to date, however, it has not been clear how these different methods behave when applied to data contaminated with highly variable and strong additive noise.

In this study we analyze the use of cross-coherence. This approach, used in seismology (e.g., Aki, 1957; Bendat and Piersol, 2000; Chávez-Garí and Luzón, 2005), calculates the cross-correlation of traces normalized by their spectral amplitudes in the frequency-domain. Thus, the method uses the phase of each trace, ignoring amplitude information, for suppressing the influence of additive noise and handling irregular input amplitudes. (Bensen et al., 2007) show examples of normalization techniques applied in seismic interferometry.

Extracting surface waves from cross-correlation of ambient noise is by now an established technique (e.g., Campillo and Paul, 2003; Shapiro et al., 2005). In contrast, the extraction of body waves has been proven to be much more difficult. Extracting body waves by cross-correlation has, however, been accomplished in some studies. Examples include extraction of P waves using distributed sources (e.g., Roux et al., 2005; Hohl and Mateeva, 2006; Draganov et al., 2007; Gerstoft et al., 2008; Draganov et al., 2009; Zhang et al., 2009) and of S waves using localized noise sources (e.g., O’Connell, 2007; Miyazawa et al., 2008). This paper presents data processing of field data dominated by strong and highly variable traffic noise as well as incoherent additive noise. We first present the basic equations of cross-correlation, deconvolution, and cross-coherence interferometry. We further demonstrate the merits of cross-coherence interferometry applied to traffic noise data for retrieval of surface waves and reflected shear waves. Because the interferometry is based on the transverse horizontal component, the extracted surface waves consist of Love waves. Finally, we explain why the cross-coherence is particularly suitable to extract the approximated Green’s function from this type of noisy data.

2 EQUATIONS OF INTERFEROMETRY

Consider the wavefield $u(r, s)$ excited at $s$ and received at $r$. In this work we use a frequency-domain formulation for all data processing. Ignoring additive noise, the wavefield can be described as the multiplication of a source wavelet and a Green’s function:

$$u(r, s) = W(s)G(r, s),$$

where $W(s)$ is the source wavelet and $G(r, s)$ is the Green’s function.

2.1 Cross-correlation and deconvolution

Let us first review the cross-correlation and deconvolution methods (Snieder et al., 2006; Vasconcelos and Snieder, 2008a) in order to compare them with the cross-coherence method. The cross-correlation of wavefields recorded at locations $r_A$ and $r_B$ is

$$C_{AB} = u(r_A, s)u^*(r_B, s) = |W(s)|^2 G(r_A, s)G^*(r_B, s),$$

where the asterisk denotes complex conjugate. Integrating this equation over a closed surface $\partial V$, which is a part of the earth’s free surface and an arbitrarily shaped surface at depth that includes all sources, gives

$$\int_{\partial V} C_{AB} ds = \langle |W(s)|^2 \rangle \int_{\partial V} G(r_A, s)G^*(r_B, s) ds,$$

where $\langle |W(s)|^2 \rangle$ is the average of the power spectra for the source wavelets. Because the integral\(\int_{\partial V} G(r_A, s)G^*(r_B, s) ds\) in equation 3 approximates $G(r_A, r_B)$ (Wapenaar and Fokkema, 2006), this gives the approximated Green’s function between the two receivers.

In contrast, the basic equation of deconvolution interferometry is

$$D_{AB} = \frac{u(r_A, s)}{u(r_B, s)} = \frac{G(r_A, s)}{G(r_B, s)} = \frac{G(r_A, s)G^*(r_B, s)}{|G(r_B, s)|^2}.$$

Deconvolution removes the influence of the source wavelet $W(s)$. Because of the absolute value in the denominator, the phase of $D_{AB}$ is determined by the numerator $G(r_A, s)G^*(r_B, s)$ in the last term of equation 4, hence the deconvolution gives the same phase as the cross-correlation method (equation 2). The deconvolution method can also be used to extract the impulse response (Vasconcelos and Snieder, 2008a) when integrating over sources located on a closed surface $\partial V$. Because of the spectral division, the result is independent of the source signature. The method thus can deal with data generated by long and complicated source signals.

2.2 Cross-coherence

The cross-coherence $H_{AB}$ is defined in the frequency-domain as

$$H_{AB} = \frac{u(r_A, s)u^*(r_B, s)}{|u(r_A, s)||u(r_B, s)|}.$$
robust information than do either cross-correlation or deconvolution. We can rewrite equation 5 as

\[ H_{AB} = \frac{G(r_A,s)G^*(r_B,s)}{|G(r_A,s)||G(r_B,s)|} \]  

because cross-coherence cancels the source wavelet term \( W(s) \) by division, as does deconvolution. Integrating equation 6 over a closed surface \( \partial V \) containing all sources gives

\[ \int_{\partial V} H_{AB} ds = \int_{\partial V} \frac{G(r_A,s)G^*(r_B,s)}{|G(r_A,s)||G(r_B,s)|} ds; \] 

hence it provides the phase of the approximated Green’s function between two receivers, but the amplitude is not preserved in this cross-coherence approach. To clarify characteristics of each approach, (Nakata, 2010) shows Taylor expansions of equations 2, 4, and 5 for small-amplitude scattered waves.

When \( r_A = r_B \) in equations 5-7, the right-hand side is equal to 1, which corresponds to the Dirac delta function \( \delta(t) \) in the time-domain. This means that the field extracted by cross-coherence satisfies a so-called \textit{clamped boundary condition} at \( r_0 \); the same boundary condition occurs in deconvolution interferometry (Vasconcelos and Snieder, 2008a; Snieder et al., 2009) but not in cross-correlation interferometry.

3 FIELD DATA PROCESSING

3.1 Data acquisition and pseudo shot gathers

We apply the cross-coherence method to traffic noise data acquired in Gunma, Japan. An aerial photograph of the observation site is shown in Figure 1a. The survey line (blue line in Figure 1a) is quasi-linear, paralleling the river; several roads and railways were crossed by or parallel to the line (shown by solid arrows in Figure 1a). In order to retrieve body waves correctly, it is necessary to have sources at the stationary phase points from every combination of receivers and reflectors (Snieder et al., 2006; Forghani and Snieder, 2010). It might appear that the sparsity of roads and railways does not provide an adequate illumination, but, as shown in Figure 2, the number of stationary points increases dramatically when one considers multiples that are reflected from different reflectors. The length of the survey line is about 2180 m, with single-horizontal-component geophones oriented orthogonal to the survey line at 10-m intervals aimed at obtaining subsurface structure from shear-wave data. The data were stored in 1200 time windows of 30-s duration, at a 4-ms sampling interval. Figure 1b shows an example of a noise record along the entire line. Higher levels of traffic noise originated from roads and railways that cross the survey line at receiver numbers 20, 50, 180, and 200. The source wavelets for the traffic noise had wide-ranging and complex frequency spectra because much traffic ran continuously with differing characteristics attributable to varying speed or weight of the vehicles. Frequency analysis reveals that the most energetic part of traffic noise is in the range 12-16 Hz (Figure 1c).

Figures 3a-3c compare the pseudo shot gathers derived from cross-correlation, deconvolution, and cross-coherence. Because we used transverse geophones, these shot gathers are dominated by shear waves. For each of the three different operations, the data acquired at receiver point 60 is used as the reference trace. The interferometric data from the 1200 records are stacked; and no other filter is applied to the records to construct these interferometry profiles. In the cross-correlation result (Figure 3a), ringing noise is dominant because of the periodic characteristic of the source wavelets of the traffic noise, and amplitude levels are particularly high at positions near the traffic noise sources (receiver num-
Figure 3. Virtual source gathers generated by (a) cross-correlation, (b) deconvolution, and (c) cross-coherence. The pseudo source point of these gathers is at receiver number 60. We applied no filter to these displayed data. (d) Detail showing hyperbolic events of panel (c). The main hyperbolic events are highlighted in transparent yellow. A bandpass filter and f-k filter have been applied to the data in panel (d). The data gaps in panel (d) are caused by the removal of incoherent traces.

Figure 4. Constant-velocity NMO-corrected CMP gathers at the CMP number 60. (a) Leftmost panel is the original gather, and the other three panels are gathers corrected using the NMO velocity shown below each panel. The highlighted areas show the time intervals within which the RMS velocity best flattens the local reflection. (b) Detailed views of the CMP gathers highlighting the reflection event at about 0.4 s.

3.2 Reflection profiles from body waves

We perform seismic reflection data processing using the CMP stack method and apply time migration to the data sets from each of the three methods. CMPs are numbered by location of the receivers along the acquisition line, consistent with the numbering in Figure 1a. After generating pseudo shot gathers by each interferometry method, we apply identical steps of bandpass filtering (5-35 Hz), f-k filtering (to reject surface waves with velocities outside the range 250-2500 m/s), NMO correction by the root-mean-square (RMS) velocity obtained from cross-coherence, CMP stack, time migration, and depth conversion. We determine the stacking velocity by performing constant-velocity stacks on NMO-corrected CMP gathers (Yilmaz et al., 2001; Stucchi and Mazzotti, 2009).

Figure 4 displays a CMP gather at the CMP number 60 corrected with three different constant values of moveout velocity; highlighted areas in Figure 4a show the time intervals used for determining the stacking velocity. As shown in Figure 4b, the reflection arrival at around 0.43 s is flattened for a RMS velocity between 680-880 m/s. Given the noise level in this figure, it is difficult to estimate the stacking velocity with great accuracy. The panels in Figure 4b and the corresponding range of RMS velocity (680-880 m/s) suggest an uncertainty of perhaps 100 m/s. The interval velocity structure obtained from the RMS velocity function estimated by cross-coherence interferometry is shown with the black line in Figure 5d. Figures 6a-6c show the migrated depth sections using all pseudo shot records derived through cross-correlation, deconvolution, and
Figure 6. Subsurface structure obtained by CMP stack, time migration, and depth conversion using reflected waves obtained by (a) cross-correlation, (b) deconvolution, and (c) cross-coherence interferometry. Panel (d) is generated from active-source data. A bandpass filter between 5 and 35 Hz has been applied to all the data before imaging.

cross-coherence, respectively. We applied the same shear wave stacking velocity function to the recorded active-shot data.

In the image obtained from cross-correlation (Figure 6a), anomalously strong waves dominate the image. These are caused by strong vibrations generated by the crossing traffic. This result agrees with that of (Hohl and Mateeva, 2006) whose image is also contaminated by local strong waves induced by rig activity. The image in figure 6a displays a marked periodicity, this is due to the narrow-band character of the noise sources (trucks and trains). As shown in figure 3a, the cross-correlation does not compensate for the narrow-band properties of the noise sources. The image retrieved by deconvolution (Figure 6b) is also noisy, although the amplitude is less variable from one location to another. The image in figure 6b is very noisy and incoherent. This is due to the fact that the virtual source gathers obtained by deconvolution interferometry (e.g., figure 3b) show few coherent arrivals. Cross-coherence interferometry gives by far the clearest image of the three methods (Figure 6c). Because cross-coherence interferometry flattens the power spectrum via the normalization in frequency domain, this type of interferometry gives an image that contains a much larger range of spatial wavenumbers than the image obtained from cross-correlation (Figure 6a).

For comparison, we show in Figure 6d a conventional stacked and migrated reflection seismic section using 224 transverse active seismic sources, which are at approximately 10-m intervals along the receiver line and recorded by transverse-component geophones along the same line. Since for practical reasons it is not possible to deploy sources close to the receiver line in the active shot experiments, the structure shallower than about 50 m is not imaged well in Figure 6d. In contrast, shallow reflections are evident in the seismic section of Figure 6c obtained from cross-coherence interferometry. Although we apply the same bandpass filter, the images in Figures 6c and 6d have different depth resolution because of the different frequency content of the sources. Images from reflected shear waves are usually noisier than P-wave images, and the images obtained from traffic noise (Figure 6c) and from active shots (Figure 6d) are both contaminated with noise. Yet both images show coherent layered structures with a region of large reflectivity between receivers 80 and 140.

Delineation of structures shallower than 300 m in Figure 6c is useful not only for geophysical exploration, e.g., static corrections and near-surface tomography, but also for ground-motion prediction by seismic monitoring in earthquake disaster prevention and basement surveys in civil-engineering applications. Although trace-to-trace amplitudes are not preserved in the cross-coherence method, the method can still be used for the delineation of underground structures.
3.3 S-wave velocity from surface waves

We apply the cross-coherence interferometry technique to shear-wave velocity estimation from the retrieved Love waves. The virtual shot record from cross-coherence interferometry (Figure 5a) clearly displays surface waves. Figures 5b and 5c show the frequency-dependent phase velocity and dispersion curve estimated from the pseudo shot record at receiver number 190.

We pick the phase velocity of the fundamental mode Love wave (blue crosses in Figure 5c) and use these measurements to invert for the shear velocity as a function of depth using a genetic algorithm (Saito and Kabasawa, 1993; Yamanaka and Ishida, 1995; Hayashi, 2008). The inverted shear velocity model is shown by the red line in Figure 5c. The shear-wave interval-velocity distribution down to around 300 m obtained from the phase velocity measurements is shown by the red line in Figure 5d. This shear-velocity profile agrees well with the interval-velocity profile obtained from NMO correction of the reflected shear waves (black line in Figure 5d). The shallow part of the shear-velocity profile from the surface-wave analysis resolves several layers, with gradually changing velocity. This shear-wave interval velocity, obtained by cross-coherence interferometry of traffic noise, is useful for estimating and monitoring ground soil strength.

4 DISCUSSION - ERROR PROPAGATION

In this section, we compare the statistical properties of cross-coherence with those of cross-correlation and deconvolution, and show theoretically why cross-coherence is preferable in the data applications shown in this work.

4.1 Correction of amplitude variation among traces

Consider the processing of data whose amplitudes vary trace by trace as a result of variations in source strength and differences in positioning or sensitivity of receivers. Ideally, the sensitivity is the same for all receivers, but in practice this is not the case because of variations in ground coupling and local topography. The equations of the three methods are

\[ C_{AA} = u(r_A,s)u^*(r_A,s) \]  \( (8) \)
\[ D_{AA} = \frac{u(r_A,s)u^*(r_A,s)}{|u(r_A,s)|^2} \]  \( (9) \)
\[ D_{A\alpha} = \frac{u(r_A,s)u^*(r_A,s)}{|u(r_A,s)|^2} \]  \( (10) \)
\[ H_{AA} = \frac{u(r_A,s)u^*(r_A,s)}{|u(r_A,s)||u(r_{\alpha},s)|} \]  \( (11) \)

where two receivers are at \( r_A \) and \( r_\alpha \). Deconvolution interferometry is asymmetric in that the amplitude of the extracted signal changes when we exchange the pseudo source and receiver points: this method thus has two different forms. Consider the case where receiver \( r_A \) records the average amplitude of all receivers and \( r_\alpha \) records an anomalously large amplitude; the recorded motion at receiver \( r_A \) is \( u(r_A,s) = G(r_A,s) \), and the motion recorded at \( r_\alpha \) is \( u(r_\alpha,s) = RG(r_\alpha,s) \) with \( R \gg 1 \), an amplification factor. The amplitude of the signals extracted with cross-correlation and cross-coherence does not change by exchanging \( \alpha \) and \( A \).

Let us compare the amplitudes among equations 8 - 11. In equation 8, \( C_{AA} = RG(r_A,s)G^*(r_\alpha,s) \), and the amplitude of \( C_{AA} \) is thus amplified by a factor \( R \). Similarly, the amplitudes of \( D_{AA} \) and \( D_{A\alpha} \) are multiplied by \( 1/R \) and \( R \), respectively. As mentioned above, the amplitudes of \( C_{AA}, D_{AA} \), and \( D_{A\alpha} \) differ from the average amplitude. Accordingly, in an analysis based on cross-correlation or deconvolution that includes the anomalous receiver \( \alpha \), the amplitude of the extracted response is unbalanced, thus requiring the additional task of removing these variations. The amplitude of \( H_{AA} \) is 1, independent of \( R \). That is, cross-coherence removes the influence of amplitude variations and achieves a stable
amplitude without separate processing to normalize the amplitude of traces constructed by interferometry.

In the deconvolution approach here, some level of white noise has to be added to prevent numerical instability. If we choose a regularization parameter that is too large, the regularized deconvolution reduces to cross-correlation. If, however, the regularization parameter is too small, the amplitude variation between traces remains large. In cross-coherence this problem does not arise because the numerator and denominator are both small when the spectral amplitude is small.

4.2 The influence of additive noise

When the data are contaminated by additive random noise $N(r)$ with zero mean, the wavefield $u(r_A, s)$ includes a noise term $N(r_A)$,

$$u(r_A, s) = W(s)G(r_A, s) + N(r_A).$$

(12)

For simplicity, let us set the source signature $|W(s)| = 1$. This additive noise might be caused by microtremors, electric noise in the equipment, and human activities. Henceforth, we abbreviate $G(r_A, s)$ as $G_A$, $N(r_A, s)$ as $N_A$, and $W(s)$ as $W$. In Appendix A, we insert equation 12 into equations 2, 4, and 5 and expand in the small quantity $|N|/|G| < 1$.

To investigate the influence of additive noise, we take the ensemble average to estimate the mean and variance of equations A5 - A7 (Appendix A). Because the ensemble average of random noise is assumed to vanish, the ensemble average of the convolution of $G$ and $N$ also vanishes. As shown in the equations A5 - A7, the ensemble average values of the cross-correlation, deconvolution, and cross-coherence in the presence of additive noise are thus given by

$$\langle C_{AB}\rangle = G_AG_B^*$$

(13)

$$\langle D_{AB}\rangle = \frac{G_AG_B^*}{|G_B|^2}$$

(14)

$$\langle H_{AB}\rangle = \left(1 - \frac{1}{4} \frac{|N_A|^2}{|G_A|^2} - \frac{1}{4} \frac{|N_B|^2}{|G_B|^2}\right) \frac{G_AG_B^*}{|G_A||G_B|^2}$$

(15)

in which "$\langle \rangle$" indicates an ensemble average. Their variances are

$$\sigma_C^2 = |G_B|^2 \sigma_{N_A}^2 + |G_A|^2 \sigma_{N_B}^2$$

(16)

$$\sigma_D^2 = \frac{1}{|G_B|^2} \sigma_{N_A}^2 + \frac{1}{|G_B|^4} \sigma_{N_B}^2$$

(17)

$$\sigma_H^2 = \frac{1}{2|G_A|^2} \sigma_{N_A}^2 + \frac{1}{2|G_B|^2} \sigma_{N_B}^2.$$  

(18)

Here $\sigma_N$ denotes standard deviation of the additive noise. Note that the noise does not bias the ensemble average of the cross-correlation and the deconvolution (equations 13 and 14), but according to expression 15 it does lead to a bias in the cross-coherence. Therefore, when we stack many times to mimic an ensemble average, the influence of noise remains as a bias in cross-coherence. Since, however, the cross-coherence does not preserve amplitude even in the absence of noise, the multiplicative bias in equation 15 is of little concern in practice.

It is difficult to compare the variances in equations 16, 17, and 18 because they express the variance in different quantities (cross-correlation, deconvolution, and cross-coherence, respectively). Instead, we consider the normalized standard deviations:

$$\frac{\sigma_C}{|C_{AB}|} = \sqrt{\frac{\sigma_{N_A}^2}{|G_A|^2} + \frac{\sigma_{N_B}^2}{|G_B|^2}}$$

(19)

$$\frac{\sigma_D}{|D_{AB}|} = \sqrt{\frac{\sigma_{N_A}^2}{|G_A|^2} + \frac{\sigma_{N_B}^2}{|G_B|^2}}$$

(20)

$$\frac{\sigma_H}{|H_{AB}|} = \frac{\sqrt{\sigma_{N_A}^2}}{2|G_A|^2} + \frac{\sqrt{\sigma_{N_B}^2}}{2|G_B|^2}.$$  

(21)

From equations 19 - 21, the normalized standard deviations are related as follows:

$$\frac{\sigma_H}{|H_{AB}|} = \frac{1}{\sqrt{2}} \frac{\sigma_C}{|C_{AB}|} = \frac{1}{\sqrt{2}} \frac{\sigma_D}{|D_{AB}|}.$$  

(22)

As shown in equation 22, the relative uncertainty in the of cross-coherence is about 70% of that of the other methods.

In summary, additive random noise causes an inconsequential bias in the cross-coherence, but the relative statistical uncertainty in the cross-coherence is reduced by a factor $1/\sqrt{2}$ ($\approx 70\%$) compared with that of cross-correlation and deconvolution. Thus, in addition to treating the problem of anomalous trace amplitudes, cross-coherence is more stable in the presence of noise.

We study the influence of additive random noise by a numerical example of synthetic data generated by a two-dimensional acoustic finite-difference time-

Figure 7. Two horizontal constant-density layers model, with an interface at 2500-m depth. The velocities of layers are 1500 m/s and 2000 m/s, respectively. Two receivers, which are shown with triangles, are positioned at 900-m depth and at lateral positions $x = 1500$ m and 3500 m. Dashed and thick arrows denote created direct and reflected wavefield by interferometry, respectively. The sources, which are represented by stars, are distributed in a horizontal line at 500-m depth, ranging from $x = 500$ m to 4500 m, in increments of 50 m.
domain method with a model consisting of two horizontal constant-density layers (Figure 7). The virtual source sections, obtained from the noise-contaminated traces and shown in Figure 8, display the direct arrival, which is represented with the dashed arrow in Figure 7 at 1.3 s, and the reflected wave, which is depicted by the thick arrow in Figure 7 at 2.5 s. The leftmost trace in each plot is noise-free so the signal-to-noise ratio is infinite. The signal-to-noise ratio (calculated from the maximum amplitude of direct arrival) of the second trace from the left is 20, which means that we added 5% random noise, while that of the third trace is 19, and that of the fourth trace is 18. The amount of noise gradually increases until the rightmost trace whose signal-to-noise ratio is 1. Figures 8a, 8b, and 8c show the wavefields retrieved from cross-correlation, deconvolution, and cross-coherence, respectively. For low signal-to-noise ratios, the direct and reflected waves are buried in ambient noise in both the cross-correlation and deconvolution results (Figures 8a and 8b). In contrast, cross-coherence interferometry (Figure 8c) reduces the influence of ambient noise.

## 5 CONCLUSION

In our study, cross-coherence interferometry provided the clearest pseudo shot gathers generated from highly variable and strong traffic noise and retrieved both reflected shear waves and Love waves. Since we used recordings of the transverse motion, this procedure yielded virtual source gathers for shear waves. The imprint of the source signature and amplitude variations between receivers is suppressed in the virtual source gathers obtained from cross-coherence. They provide shear-wave images obtained by migrating virtual source data that agree to a large extent with those obtained with active sources. Moreover, the images obtained by active sources lack the shallow structures seen in the image obtained from the cross-coherence of traffic noise because the active sources were placed at a distance from the survey line. The virtual source sections obtained from cross-coherence exhibit both surface waves and body waves; we used the surface-wave data to carry out dispersion measurements of the fundamental mode Love waves and obtained a shear-wave velocity profile that agrees with the interval-velocity function calculated by the stacking velocity profile. Corresponding with early studies, it is easier to extract surface waves than body waves.

Compared with the standard deviation of the virtual source sections obtained from cross-correlation and deconvolution, the relative statistical uncertainty in cross-coherence is 70% as large. In contrast to cross-correlation and deconvolution, additive noise leads to a multiplicative bias in virtual source signals based on cross-coherence. Since the cross-coherence method does not conserve trace-to-trace amplitude, this multiplicative bias is of little concern. Because of the normalization employed, the method overcoming amplitude variations among traces. In any case, because the amplitude information is lost, cross-coherence interferometry is inappropriate for data analysis that exploits amplitude information, such as the measurement of reflection coefficients, amplitude versus offset, and attenuation. Cross-coherence is particularly suitable for data that are noisy, vary in amplitude among traces, or have long and complex source wavelets. The primary target of cross-coherence interferometry here is the estimation of shear-wave velocity from surface waves and the shape of subsurface structures obtained from reflected body waves. By using the transverse-component of the ground motion for cross-coherence, we obtain a shear-velocity profile and a shear-wave image of the subsurface. This information is useful for various applications such as static corrections, near-surface tomography, ground motion prediction for earthquake disaster prevention, mon-
onitoring the ground soil strength, and basement surveys for civil engineering.

6 ACKNOWLEDGMENTS

We are grateful to Kyosuke Onishi of Akita University, Toshiyuki Kurahashi of Public Works Research Institute, and Takao Aizawa of Suncoh Consultants for the acquisition of the seismic data and Tatsunori Ikeda of Kyoto University for the technical support in surface wave analysis. We also thank Akihisa Takahashi of JGI, Inc., Shohei Minato of Kyoto University, the editor, and three anonymous reviewers for suggestions, corrections, and discussions. Norimitsu Nakata is grateful for support from the Japan Society for the Promotion of Science (JSPS: 22-5857).

REFERENCES


Bendat, J., and A. Piersol, 2000, Random data: analysis and measurement procedures: John Wiley & Sons.


Miyazawa, M., R. Snieder, and A. Venkataaraman, 2008, Application of seismic interferometry to extract P- and S-wave propagation and observation of shear-wave splitting from noise data at Cold Lake, Alberta, Canada: Geophysics, 73, D35–D40.


Wapenaar, K., E. Slob, and R. Snieder, 2008, Seismic and
APPENDIX A: ERROR PROPAGATION

By substituting equation 12 into equations 2, 4, and 5, and assuming that $|W(s)| = 1$, we obtain the following expressions for cross-correlation, deconvolution, and cross-coherence using wavefields that include random noise with zero mean:

\[
C_{AB} = (WG_A + N_A)(W^*G_B^* + N_B^*) \tag{A1}
\]
\[
D_{AB} = \frac{WG_A + N_A}{WG_B + N_B} \tag{A2}
\]
\[
H_{AB} = \frac{(WG_A + N_A)(W^*G_B^* + N_B^*)}{|WG_A + N_A||WG_B + N_B|} \tag{A3}
\]

Below, we assume that the additive noise at different locations is uncorrelated; hence

\[
\langle |N_A|^2 \rangle = \sigma_{N_A}^2, \quad \langle |N_B|^2 \rangle = \sigma_{N_B}^2, \quad \langle N_A N_B^* \rangle = 0.
\]

When $\psi$ represents the phase,

\[
N = |N|e^{i\psi}; \tag{A4}
\]

then under the assumption that the amplitude and phase are uncorrelated, the ensemble average of $N^2$ is

\[
\langle N^2 \rangle = \langle |N|^2 e^{2i\psi} \rangle = \langle |N|^2 \rangle \langle e^{2i\psi} \rangle.
\]

We assume that the phase has a uniform distribution, thus $\langle e^{2i\psi} \rangle = 0$; hence

\[
\langle N_A^2 \rangle = \langle N_B^2 \rangle = 0.
\]

We further assume that the level of additive noise is small ($|N|/|G| < 1$) and expand equations A1 - A3 in $|N|/|G|$. Ignoring noise terms higher than second-order in $|N|/|G|$, gives

\[
C_{AB} = G_A G_B^* + W^*N_A G_B^* + WG_A N_B^* + N_A N_B^* \tag{A5}
\]
\[
D_{AB} = \frac{G_AG_B^*}{|G_B|^2} + \frac{W^*N_AG_B^*}{|G_B|^2} - \frac{W^*G_AG_B^*N_BG_B^*}{|G_B|^4} - \frac{(W^*)^2N_AG_BN_BG_B^*}{|G_B|^4} + \frac{(W^*)^2G_AG_B^*N_BG_B^*}{|G_B|^4} \tag{A6}
\]
\[
H_{AB} = \left(1 - \frac{1}{4} \frac{|N_A|^2}{|G_A|^2} - \frac{1}{4} \frac{|N_B|^2}{|G_B|^2} \right) \frac{G_AG_B^*}{|G_B|^2} \frac{G_AG_B^*}{|G_B|^2} + \frac{\frac{1}{2} |W^*N_AG_B^*}{|G_A||G_B|} + \frac{\frac{1}{4} |W^*N_AG_B^*}{|G_A||G_B|} + \frac{1}{4} \frac{1}{|G_A||G_B|}
\]
\[
- \frac{1}{2} \frac{1}{|G_A||G_B|} - \frac{1}{2} \frac{1}{|G_A||G_B|} - \frac{1}{8} \frac{|W^*|^2N_AG_B^*N_AG_A^*}{|G_A|^4} - \frac{1}{8} \frac{|W^*|^2N_AG_BN_BG_B^*}{|G_A|^4} - \frac{1}{8} \frac{1}{|G_A||G_B|^2}
\]
\[
- \frac{3}{8} \frac{1}{|G_A|^4} - \frac{3}{8} \frac{1}{|G_A|^4} - \frac{3}{8} \frac{1}{|G_A|^4} - \frac{3}{8} \frac{1}{|G_A|^4} + \frac{1}{4} \frac{1}{|G_A|^4} + \frac{1}{8} \frac{1}{|G_A|^4} \tag{A7}
\]

Taking expectation values gives the mean and variance of equations 13 - 18.