Development of a time-frequency representation for acoustic detection of buried objects

Vincent Valeau, a) James Sabatier, b) R. Daniel Costley, c) and Ning Xiang d)

National Center for Physical Acoustics, Coliseum Drive, The University of Mississippi, University, Mississippi 38677

(Received 13 September 2002; revised 6 August 2004; accepted 6 August 2004)

A recently developed apparatus permits the detection of buried objects, such as landmines, by remotely sensing the variations of ground vibration that occur over the buried object with a laser Doppler vibrometer (LDV), when the ground is insonified by means of acoustic-to-seismic coupling. As it is currently implemented, the LDV scans individual points on the ground. The technique shows much promise, but it is slow when compared to some other techniques. This work demonstrates that mines can be detected as the LDV beam moves continuously across the ground, by using an optimal time-frequency representation of the LDV signals. This improvement has the potential to signiﬁcantly increase scanning speeds. The vibrometer output signal is analyzed by means of time-frequency representations, which exhibit characteristic acoustic “signatures” when scanning over buried objects. The most efﬁcient representation appears to be the smoothed spectrogram weighted by the time-frequency coherence function. It detects the searched signal energy enhancements while ﬁltering out most features due to speckle noise. Detection is then efﬁciently achieved by searching simultaneous extrema of the marginals and moments of this representation. Experiments show that scanning speeds up to 3.6 km/h can be achieved for successful detection of buried landmines in outdoor ground. © 2004 Acoustical Society of America.

PACS numbers: 43.60.Qv, 43.60.Gk, 43.60.Rw [JCB]

Pages: 2984–2995

I. INTRODUCTION

Over the last few years, a new acoustic technique to detect buried objects has been developed.1 Its first application is the detection of landmines,2–4 as landmines have proliferated in many of the world’s conﬂicts. Current conventional detectors rely on detecting the metallic parts of landmines, causing many manufacturers to focus their efforts in building nonmetallic landmines. Therefore, there is a need for new techniques aimed at detecting nonmetallic objects. Among those, this new acoustic technique appears to be very successful in terms of detection probability and false alarm rate.5

This technique relies on the acoustic-to-seismic (A/S) coupling phenomenon. The physics of airborne sound penetrating into the ground has been well described.6–10 Outdoor ground is a particulate material containing air-filled pores. It permits the coupling of energy from an airborne acoustic wave into waves propagating through both its frame and through the fluid contained in its pores. A buried object (sometimes referred to as “target” in this study) presents a different acoustic impedance to waves traveling through the ground because it lacks porosity. Recently, Donskoy11 has shown that the acoustic compliance of a mine is much greater than that of the soil media. As a result, the buried object scatters the incident acoustic energy. This yields a local enhancement of the A/S coupled motion in the soil over a buried object that can be detected by measuring the ground surface vibration.

The ﬁrst successful experiments using a laser Doppler vibrometer (LDV) to detect A/S coupled ground motion12 led to the development of the actual landmine detector prototype, which has been proven efﬁcient over the past several years.2–4 Additionally, modeling13–17 has shown the inﬂuence of buried objects on the A/S coupled surface vibration. In its current conﬁguration, the landmine detection technique consists of ﬁrst insonifying the ground by means of A/S coupling. A LDV remotely measures the ground vibration at discrete points along a virtual grid projected on the ground, with the laser beam going from one point to the next one. A map of ground vibration is constructed from the LDV measurements made over this grid and makes possible the detection of local variations of the vibration induced by buried objects. Because outdoor measurements are often conducted under hostile experimental conditions and because the roughness of natural ground causes the LDV signals to be corrupted with speckle noise, time averages have to be performed for each LDV measurement. Consequently, the time required for data acquisition can be incompatible with real applications.

To overcome this problem, the present study proposes to assess the feasibility of detecting buried objects by sweeping the laser beam across the ground. Spatial variations of ground vibrations along a straight line are then detected by
measures, on the other hand, are better suited for such applications.

Performing suitable and efficient processing of the LDV signal is then compulsory, and the main contribution of this paper concerns this problem. The signal to be analyzed is a superposition of sine waves with time-varying amplitudes and phases, corrupted by many optical noise artifacts. Its representation in the time-frequency plane exhibits characteristic features accounting for energy increases when scanning above a target, referred to in the following as a "mine signature." A mine signature can thus be detected by using time-frequency representations.\(^1\) Time-frequency is a "natural" approach to this problem as the time variable is immediately interpretable as a spatial position and frequency contains the clues for target presence. Adaptations of the classical spectrograms\(^2\) are then achieved in order to detect the typical changes of signal spectral content that are associated with the presence of a target. In the case of high scanning speed, the rarely used notion of time-frequency coherence appears to be of great practical interest to filter out nonacoustic contributions in this application.

Section II presents the acoustic detection technique. The proposed sweeping LDV scanning is then described in Sec. III. Section IV discusses the signal processing tools needed to process the LDV signals; it starts from the well-known spectrogram and presents how an efficient tool can be developed from it, given the physical phenomena that must be detected. Section V is a synthesis of the results and system performances achieved in this study in terms of scanning speed, target depth, and ground nature.

II. ACOUSTIC DETECTION OF BURIED OBJECTS

A. Physical background

When airborne sound is present on the surface of a soil, air contained within the connected pores in the soil is caused to oscillate in and out of the pores. In this way, the energy from the vibrating air above the soil surface couples with the air in the soil pores. This process is both dependent on the frequency of the sound waves and on the properties of the soil. Hence, the penetration of the sound waves is strong if the air permeability is high; conversely, low air permeability leads to weak coupling between vibrations of air above the surface and air in the pores.\(^3\) Through momentum transfer at the air–soil boundary and viscous friction at the pore walls, some energy is also transferred to the soil frame. This transfer of energy can be successfully described by Biot's model.\(^4\)\(^5\)

The Biot equations of motion include two compressional wave solutions that propagate simultaneously in both fluid and solid and have different propagation constants.\(^6\) The waves are referred to as the waves of the first and second kind, or simply as "fast" and "slow" waves. The fast wave is relatively unattenuated, nondispersive, and most of the wave's energy is contained in the matrix displacement. Conversely, the slow wave is highly attenuated with depth and very dispersive and most of the energy is in the fluid. The key aspect of the physics of air-filled solids as used in the acoustic technology for mine detection is that air-borne sound is preferentially coupled into the soil through the slow wave, as was shown experimentally in Ref. 10. Although most of the energy of the slow wave is contained in the fluid displacement, it can be observed by sensing the matrix displacement or velocity. This can be achieved, for example, by measuring the vibrational velocity of the surface.

The speed of this slow wave is much slower than the sound speed in air, so the slow wave is refracted towards the normal. It seems then highly plausible that it scatters from targets buried in soils (or even from inhomogeneities of soil properties) and reflects back to the surface, which locally increases the vibrational velocity of the surface. Kargl\(^7\) modeled the scattering of the slow wave with sea mines buried in sediment. Recent works\(^8\)\(^9\) aim at modeling the physics of scattering from buried objects in air-filled porous materials, but it is clearly observed experimentally by sensing the surface velocity of a soil acoustically excited by A/S coupling.\(^1\)

B. Experimental setup

Soil is acoustically excited with two loudspeakers generating airborne sound waves directed towards the ground at sound pressure levels of about 100 dB, C-weighted. The resulting A/S-coupled vibration is remotely estimated by using a LDV (PSV 200 manufactured Polytec P I, Inc.). The LDV emits a laser beam onto the vibrating surface of the ground area under test. The surface vibration causes a Doppler shift of the reflected light, which is sensed by a photo detector. The output of the photo detector is a frequency-modulated signal whose modulating signal is the vibrating surface velocity. After demodulation, the obtained signal is a voltage proportional to the instantaneous velocity of the point where the laser beam strikes the ground.

A set of LDV measurements is carried out at discrete points of a virtual grid projected onto the surface. Typically, the grid contains 16×16 points, evenly spaced over a square meter. In such an arrangement, the spatial resolution is approximately 6 cm. This spatial resolution can be adjusted as a function of the expected size of the buried object. In the usual configuration, the excitation signal is a pseudo-random noise covering a frequency range between 80 and 300 Hz. Experience has shown that this band contains frequencies for which buried objects present are most visible. The measured vibrational velocity is collected, Fourier-transformed, and averaged in the complex frequency domain. More details about the detection and the associated processing may be found in Ref. 1 or 3. An image is constructed showing the mean ground velocity amplitude within a given frequency band. The presence of a mine is indicated by an area of local velocity enhancement in this frequency band. Figure 1, taken from Ref. 2, shows an example: a VS 2.2 plastic mine is buried at a depth of 7.5 cm. In Fig. 1(a) the color dots show the 16×16 LDV measurement locations with integrated velocity values in the frequency band between 130 and 160 Hz.
(color scale is given). In Fig. 1(b) a color map is derived from Fig. 1(a) after interpolation and spatial filtering. The presence of the buried mine is clearly visible in this example. The grid covers an area of one square meter.

As mentioned in Sec. I, outdoor conditions and natural ground roughness greatly increase the noise level. It is then that the processing of LDV signals with bad signal-to-noise ratios can be critical. A single-point measurement requires the laser beam to be kept still as the signals are collected so that averages can be computed and the velocity spectrum estimated. The required time for the laser beam to be kept at the same point strongly depends on experimental conditions such as ground surface, wind, etc. After completion the laser beam is moved to the next point. Scanning along a 1-m-long line can take from about 20 s to more than a minute, and scanning a whole square meter can take up to 20 min. This procedure gives excellent results in terms of detection probability and false alarm rate, but its main drawback is that it is time-consuming and the time required for data acquisition is not compatible with real applications.

III. CONTINUOUSLY SCANNING LDV

A. Principle

In order to reduce the time required for detection, the possibility of using a continuously scanning LDV has been investigated. This is achieved by sweeping the laser beam over the ground with a constant velocity. In other words, the ground velocity variations are measured by means of a single “sweeping” measurement along a straight line parallel to an axis $\vec{x}$, as plotted in Fig. 1(a). For this last example (see Sec. II), the velocity map of a square meter can then be obtained by sweeping the laser beam along 16 parallel lines instead of $16 \times 16$ measurements with the grid method. This new procedure is a good candidate to be integrated in future configurations, as the whole apparatus will be mounted on a vehicle moving at a steady speed. Hence, the applicability of this continuous LDV scan and the subsequent signal processing are of great importance for the integration of this landmine detection technique in real applications.

The experiment, which was performed in order to assess the feasibility of such an approach, is sketched in Fig. 2. A mirror, whose position is controlled with a function generator, drives the laser beam position. Hence, it is assumed that the position of the impact of the laser beam on the ground is proportional to the control signal. In this study, LDV scans are carried out along 1-m-long lines. A buried object is located under the trajectory of the beam. In order to repeat the experience several times, the laser beam is swept back and forth over the mine; the beam position is thus controlled by a
saw-tooth signal, whose frequency will be referred to as the "sweep rate." This back-and-forth sweeping scan gives some indication about the repeatability of the detection. The sweep rate is proportional to the velocity of the laser beam position on the ground: as the swept distance in one period is usually 2 m, a sweep rate of 0.1 Hz corresponds to a laser beam speed (or "scanning speed") of 0.2 m/s, or 720 m/h.

The airborne acoustic waves which couple into the ground are generated by loud speakers: the sound pressure level is 100 dBC in the vicinity of the scanned area. The acoustic excitation is a superposition of single frequency sine waves, whose number is usually four in the presented results. These frequencies are located in the low-frequency range 80–300 Hz, which is the range for which scattering of sound waves by buried objects is most visible. Former experiments showed that this frequency range is mostly dependent upon the resonance of the target free air acoustical admittance (i.e., the ratio of the velocity of the target top plate over the incident acoustic pressure). It seems to be independent of the target size and mostly controlled by the mechanical properties that contribute to the target’s acoustical compliance.

Because of A/S coupling, the measured ground vibration is a superimposition of the generated airborne sine waves. Above a target, the measurement signal exhibits not only an increase in amplitude at these frequencies, but also phase changes. Therefore, changes in signal energy and phase at these frequencies as the laser beam sweeps the ground indicates the presence of a buried object. Some nonlinearities such as those observed by Donskoy have also been observed, but their levels are generally 40 dB below the linear frequency components. Hence the nonlinear components levels are much nearer to the velocity noise floor of current LDV devices and are not expected to influence the measurements presented in this study.

In all the experiments reported here, the buried object was a cylindrical anti-tank (AT) mine with a diameter of about 30 cm. Experiments to determine whether smaller objects can be detected with this technique, such as anti-personal mines (with diameters smaller than 10 cm), are presently in progress. All experiments have been implemented outdoors in dirt roads, gravel roads, and natural ground (i.e., a lane of natural loess soil). All of the mines were buried at least one year before these measurements were made. The signal processing methods will be evaluated for different situations, and the influence of the following parameters is studied: most importantly the speed of the laser beam, but also the type of ground and its surface, and the object depth.

B. Measurement noise

In recent studies, spatially continuous LDV scans have been used in different applications for analyzing vibrating structures in order to detect damage or to measure vibrational mode shapes. The detection of temporal changes in signal characteristics was not of interest, as in this study. Moreover, the context of the present study is very different as measurements are performed outdoors on "real" grounds, away from the well-controlled environment of a laboratory. These uncompromising conditions imply that noise will be an important limiting factor.

Indeed, the type of optical noise, known as "speckle noise," is critical in this study. Light back-scattered from the ground is composed of a population of wavelets with random relative phases called "speckles." When the ground surface is rough the population of speckles contributing to the reflected light changes with time, causing random phase changes, or speckle noise, in the photo detector output. Rougher ground corresponds to higher speckle noise levels. The problems increase when the laser beam is swept over the surface: when the speed of the beam over the ground increases, the speckle population changes faster, and the speckle noise level increases. Because of speckle noise, there is a limit in the sweep rate (or scanning speed) above which detection is impossible because the noise level is too high. Other sources of noise are photo detector shot-noise and vibrations of the LDV platform (due to wind, for instance). Platform vibrations will result in an apparent velocity not related to the ground motion, introducing spurious information. However, these noise sources are likely to be minor compared to speckle noise. In this study all noise sources are assumed to be additive components of the photo detector output (i.e., the measured vibrational velocity).

LDV measurement noise can be clearly observed by considering Fig. 3. The laser beam has been swept back and forth over a mine buried at 5-cm depth in the loess soil which has been acoustically excited via A/S coupling. The excitation frequencies were 105, 125, 170, and 205 Hz. The resulting ground velocity amplitude at these frequencies is of the order of a hundred μm/s, which is similar to other measurements. The power density spectrum (PDS) of the output velocity has been computed for sweep rates of 0.1, 0.2, and 0.4 Hz, and 200 estimates of the PDS have been frequency averaged to achieve this measurement. The four single frequencies are fairly visible for all sweep rates, but the noise floor level increases substantially with the scanning...
A. Need for time-frequency analysis

The acoustic excitation is a superposition of single-frequency sine waves (see Sec. III A) whose frequencies are noted as \( f_i \). Hence the instantaneous velocity of the ground surface is a sum of four sine waves. In order to detect the presence of a target, it is intended to estimate the variation of the spectral density of this velocity along a straight line (axis \( x \), see Fig. 1). \( v(x,t) \) is the two-variable function describing the ground velocity at position \( x \) and time \( t \):

\[
v(x,t) = \sum_{q=1}^{N} A_q(x) \cos(2\pi f_q t + \varphi_q(t)).
\]

The spatial variations \( A_q(x) \) along the \( x \)-axis permit the detection of the presence of buried objects. These functions can have many forms, but it is generally observed that a buried object induces a local maximum of \( A_q(x) \), which permits its detection. The aim is to estimate the functions \( A_q(x) \) in a single sweeping LDV measurement along the axis \( x \). The sweeping velocity in m/s of the laser is noted as \( \lambda \). At time \( t \), the laser beam is scanning the ground at velocity \( v(x = \lambda t, t) \). Hence, by sweeping the laser beam over the ground, a signal \( s(t) \) is obtained which is, after demodulation,

\[
s(t) = \sum_{q=1}^{N} A_q(t) \cos(2\pi f_q t + \varphi_q(t)) + b(t),
\]

with \( t = x / \lambda \).

Noise \( b(t) \) is an additive noise (see Sec. III B). The spatial velocity variations \( A_q(x) \) can be simply retrieved from the amplitude modulations functions \( A_q(t) \) by the rescaling operation \( x = \lambda t \).

\( \hat{s}(t) \) is now constructed from the real-valued \( s(t) \) and its Hilbert transform\(^{20} \) as the imaginary part to obtain the so-called analytic function. We assume that \( A_q(x = \lambda t) \) is a slowly varying function compared to \( \cos(2\pi f_q t + \varphi_q(x = \lambda t)) \), which is achieved, for objects of the size of AT mines, for sweeping velocities \( \lambda \) of less than 20 m/s. This corresponds to a maximum sweep rate of 10 Hz. Under this condition, the frequency content of the amplitude modulation functions \( A_q(t) \) are located at sufficiently low frequencies so as not to spill over the domain containing the frequencies \( f_q \).

This is achieved, the Hilbert transform \( H[s(t)] \) gives exactly the imaginary part of the analytical signal \( \hat{s}(t) \) of \( s(t) \),\(^{23} \) which yields

\[
\hat{s}(t) = s(t) + jH[s(t)] = \sum_{q=1}^{N} A_q(t) e^{j\varphi_q(t)} e^{-j2\pi f_q t} + \hat{b}(t),
\]

where \( \hat{b}(t) \) represents the complex version of noise \( b(t) \). The term \( A_q(t) e^{j\varphi_q(t)} \) is the time-varying complex envelope for the component at frequency \( f_q \) for the deterministic part of signal \( \hat{s}(t) \). \( A_q(t) \) is called the instantaneous amplitude of the single tone \( f_q \) component. It is exactly the amplitude of the vibrational velocity component at frequency \( f_q \) and at position \( x = \lambda t \). The signal \( \hat{s}(t) \) is strongly nonstationary due to amplitude modulation and violent noise fluctuations, due essentially to burst noise. Time-frequency analysis techniques are signal-processing tools, which are widely used to analyze nonstationary signals. Hence, in order to estimate the “time-varying” spectrum of the signal \( \hat{s}(t) \), we wish to have a complex time-frequency representation \( S(t,f) \) which would be ideally a

\[
S(t,f) = \sum_{q=1}^{4} A_q(t) e^{j\varphi_q(t)} \delta(f-f_q) + B(t,f),
\]

with \( B(t,f) \) being a time-frequency representation of the noise measurement \( b(t) \). Ultimately, the idea is to retrieve, for the excitation frequency \( f_q \), the spatial variations \( A_q(x) \) and \( \varphi_q(x) \) of the velocity amplitude and phase, respectively, from the modulus and phase angle of \( S(t,f_q) \) in the exposed time-frequency representations.

The next part deals with how to generate efficient time-frequency representations for detecting the spatial variations.

FIG. 4. LDV output signal for scanning over a SIM30 type mine buried in a dirt road (5 cm depth) at scanning speed 0.2 m/s. (a) Temporal signal in which 1 m was scanned in 5 s. (b) Contour plot of smoothed spectrogram. The mine signature appears around \( t = 3 \) and 8.5 s.
The question of detecting and using optimally local phase variations $\varphi_q(t)$ is addressed in Secs. IV C and IV D. Section IV E provides tools for a simple automation of the detection from the signal time-frequency representation.

### B. Use of smoothed spectrograms

The most popular estimator of “time-varying” spectrum is the short-time Fourier transform (STFT). Let us first define the STFT of signal $s(t)$ in the discrete time domain:

$$\text{STFT}_N^s(t_i, f_i) = \frac{1}{\sqrt{2N+1}} \sum_{\tau=-N}^{N} \hat{s}(\tau) h_N(\tau-t_i) e^{-j2\pi f_i \tau}, \quad (5)$$

in which $t_i$ and $f_i$ are discrete time and frequency, $\tau$ is a discrete lag variable, and $h_N$ is the short-time window, whose length is $(2N+1)$. This relation describes the discrete Fourier transform of a portion of signal $\hat{s}$, centered on instant $t_i$. The window length $(2N+1)$ has to be short enough so that local properties of signal $\hat{s}(\tau)$ do not change “too much” within this duration. It means that the variations of the complex envelope $A_q(t) e^{j\varphi_q(t)}$ have to remain stationary enough along the window length. Meanwhile, the frequency spread of each peak, linked to the time-frequency spread obtainable with the STFT in terms of uncertainty, must remain small enough so that each frequency peak stays distinct from the other peaks in the time-frequency plane. The spectrogram is defined as the squared modulus of the STFT:

$$\text{SP}_N^s(t_i, f_i) = |\text{STFT}_N^s(t_i, f_i)|^2, \quad (6)$$

For analyzing noisy signals, the spectrogram can give high variance estimates of the time-varying spectrum. In order to improve variance estimates, averages may be used. For this purpose, $(2M+1)$ calculations of $\text{SP}_N^s(t_i + \eta \times K, f_i)$, taken at instants $(t_i + \eta \times K)$ around instant $t_i$, are used to perform an averaged estimation of $\text{SP}_N^s(t_i, f_i)$. The subsequent “locally averaged” version of the spectrogram is called the smoothed spectrogram. The smoothed spectrogram used in this study is adapted from the one proposed by Martin and Flindrin and can be written

$$\text{SP}_{N,M}^s(t_i, f_i) = \sum_{\eta=-\infty}^{\eta=\infty} g_M(\eta) \text{SP}_N^s(t_i + \eta \times K, f_i). \quad (7)$$

This estimator of the time-varying spectrum is used in this study to get an estimation of the variation of the ground surface velocity amplitude. As an example, a contour plot of the smoothed spectrogram of the signal presented in Fig. 4(a) (see Sec. III B, a SIM30 mine buried in a dirt road) is presented in Fig. 4(b). Time is on the horizontal axis and frequency is on the vertical axis. The ground is acoustically excited at frequencies 95, 120, 145, and 170 Hz. The energy at those frequencies is clearly enhanced when the scanning beam passes above the mine, as the scattering of the acoustic wave by the target locally leads to an increase of the ground surface velocity. This set of horizontal features associated with the passage of the laser beam over a mine is the “mine signature” allowing its detection. It is also noticed that during the second passage of the beam over the mine, some vertical features appear. These features are the time-frequency representations of burst noise: they are wide-band and usually have a short duration. Burst noise is troublesome because it can mask the mine signature or, for some longer bursts, can be interpreted as an ambiguous signature. It is also interesting to notice that the occurrence of burst noise is not repeatable; the laser beam passes twice over the mine, but burst noise occurs only once.

When using smoothed spectrograms, a proper adjusting of the windows $h_N$ and $g_M$ and of the increment $K$ is necessary and crucial. The window $h_N$ is 4096 points long ($N = 4096$), which, at a sampling frequency of 25 kHz, allows a frequency resolution of 6 Hz. Window $g_M$ smoothes the data in the frequency domain. Both windows $h_N$ and $g_M$ are rectangular windows in the presented results. For calculating the smoothed spectrogram at time $t_i$, 20 averages are performed (i.e., $M = 20$). The critical parameter is the increment $K$, which determines the amount of smoothing: it sets up the time length over which the smoothing is achieved (i.e., $2MK$), a duration that will be referred as the smoothing length in the following. Introducing the parameter $K$ allows for an adjustment of the smoothing length while keeping constant the computational load ($M$ constant). An equivalent spatial smoothing length is also defined, and is equal to $(2MK/s_f)$, where $s$ is the speed of the scanning beam and $f_s$ is the sampling frequency (25 kHz in this paper).

Figure 5 illustrates the influence of the smoothing length. A VS-1.6 mine (diameter 24 cm) is buried at 5-cm depth in the loess soil. The soil is excited at frequencies of 105, 125, 170, and 205 Hz. The laser beam scans above the mine about instant $t = 6.2$ s, and moves at scanning speed 0.4 m/s (sweep rate 0.2 Hz). The spatial smoothing length is set to 1 cm (top, $K = 1500$), 16 cm (middle, $K = 2500$), and 64 cm (bottom, $K = 10000$). On the right, the contour plots of the smoothed spectrograms (the equivalent spatial scale is on the top of the graphs) are depicted, and for a better visualization of the smoothing effect, the corresponding three-dimensional representation is plotted on the left. The mine signature appears about instant $t = 6$ s, and the equivalent size of the signature is about 20 cm, which reasonably approaches the mine size. An additional feature appears around instant $t = 7$ s at the higher frequencies, and is supposedly due to local variations of the soil acoustical properties. This kind of event is referred as “clutter,” and can be mixed with a mine signature, although here its apparent size (about 10 mm) is much less than the size of the searched object.

It is noticed that a smoothing length of 16 cm (middle) gives better results than a 1 cm length (top), as noise is substantially filtered out. In more hostile cases (i.e., for higher scanning speeds), a smoothing length of the order of a few cm does not allow the mine to be detected, while the mine is detected for smoothing lengths between 15 and 25 cm. On the other hand, a smoothing length that is too large (bottom) results in a loss of resolution power of the ground velocity variations: here the effect is the loss of the mine signature, as it mixes with clutter. Therefore, a tradeoff is necessary between variance reduction and the detection of refined velocity variations. Practically, a smoothing length ranging from $D/2$ and $D$ gives best results, $D$ being the equivalent diameter of the searched object.
Let us emphasize here that in the results presented, the spatial resolution has been set to about 2.5 cm. This spatial resolution is set by the choice of the increment \( [t_{i+1} - t_i] = \Delta t \), and corresponds to \( \lambda \Delta t \); it is the distance between two consecutive estimation points. The resolution power for the estimation of the spatial velocity variations is a function of both the spatial resolution and the smoothing length.

C. Use of phase information: The time-frequency coherence

As proposed in Eq. (4), we wish to find a complex representation of the signal, which would estimate both variations of amplitude \( A_q(x) \) and phase \( \varphi_q(x) \) of the ground acoustic velocity. The estimation of \( A_q(x) \) was the concern of the last section. For obtaining the phase, a reference signal is required. In this study, the reference signal \( r(t) \) is the excitation signal sent to the loudspeakers: this signal provides the phase reference, being the sum of single-tones with constant amplitudes and phases. Conversely, the analyzed signal \( s(t) \) is a sum of single tones with time-varying amplitudes and phases. In order to obtain the evolution of the phase \( \varphi_q(t) \), one has to evaluate the complex cross-spectrum between \( r(t) \) and \( s(t) \). In fact, due to the nonstationary nature of \( s(t) \), this cross-spectrum is a time-varying cross-spectrum. Hence, the following estimator is proposed to obtain a time-frequency representation of this time-varying cross-spectrum, noted \( \text{CSP}_{N,M}^r(t_i,f_i) \):

\[
\text{CSP}_{N,M}^r(t_i,f_i) = \sum_{\eta = -\infty}^{\infty} g_M(\eta) \text{STFT}^r_N(t_i + \eta \times K f_i)^* \times \text{STFT}^r_N(t_i + \eta \times K f_i), \tag{8}
\]

where \( * \) denotes the complex conjugate. This estimator is a direct extension to the cross-spectrum of the estimator of the smoothed spectrogram written in Eq. (7). The phase variations \( \varphi_q(t_i) \) at discrete time \( t_i \) and acoustic excitation frequency \( f_q \) are estimated directly from the phase of this time-varying cross-spectrum estimator:

\[
\varphi_q(t_i) = \text{Arg}(\text{CSP}_{N,M}^r(t_i,f_q)), \tag{9}
\]

where function \( \text{Arg}(...) \) denotes the phase angle of a complex quantity. From Eq. (8), it is also possible to propose an estimator of the time-varying coherence (or “time-frequency coherence”) that we may write as

\[
C_{N,M}^r(t_i,f_i) = \frac{|\text{CSP}_{N,M}^r(t_i,f_i)|^2}{\text{SP}_{N,M}^r(t_i,f_i)/\text{SP}_{N,M}^r(t_i,f_i)} \tag{10}
\]

A theoretical approach to the notions of time-frequency cross-spectrum and time-frequency coherence may be found in Ref. 26. Although those quantities were mathematically derived and justified, very few practical applications\(^2\) have been carried out with the time-frequency coherence so far. In practice the calculation of this coherence function requires the simultaneous acquisition of both the input signal of the sound source \( r(t) \) and the LDV output signal \( s(t) \).

In order to estimate the spatial variations of the ground velocity phase, the phase angle of the time-varying cross-spectrum is calculated. It is then possible to observe, for a given excitation frequency \( f_q \), the temporal (or equivalently spatial) evolution of the ground velocity phase as the beam scans over a buried object. An example is shown in Fig. 6. It considers the same signal as in Fig. 4 resulting from the detection of a mine buried in a dirt road. At the lowest acoustic excitation frequency \( f_1 = 95 \text{ Hz} \), the squared amplitude estimate \( A_1(t)^2 \) (i.e., smoothed spectrogram at frequency \( f_1 \)), phase estimate \( \varphi_1(t) \) (from the phase angle of the cross-spectrum at \( f_1 \)), and coherence time-variations \( C_{N,M}^r(t_i,f_1) \) are considered. They are respectively depicted in Figs. 6(a)–(c). In Fig. 6(a), the two “bumps” of the velocity amplitude around instant 3 and 8.5 s are the consequence of scanning over the target. In Fig. 6(b), we notice that those local enhancements of the velocity energy \( A_1(t)^2 \) coincide with a local stabilization of the corresponding phase \( \varphi_1(t) \). In other
words, when the beam scans off-target locations, the estimated velocity phase varies rapidly with time (or position), but this phase remains quite steady over the buried objects. This phenomenon is very characteristic, and the spatial stability of phase is a good clue for detecting the presence of a mine. Now, the coherence is a good indicator of the stability in time of the phase of a single tone, and this justifies its use, as demonstrated in Fig. 6. Coherence at frequency \( f_1 \) is close to one when scanning over the mine and drops down when scanning off-mine.

D. Coherence-weighted spectrogram

Another characteristic example can be found in Fig. 7. A scan (sweep rate 0.1 Hz, scanning speed 0.2 m/s) is made over a mine buried in the loess soil at 5 cm depth, and the smoothed spectrogram is represented in Fig. 7(a). Some wide-band features appear also, for example, around instants 9.3 and 10.5 s, and are the consequence of burst noise. As in Fig. 6, the quantities \( A_3(t) \), \( \varphi_3(t) \), and \( C^r_{N,M}(t_i, f_3) \) are plotted for frequency \( f_3 = 145 \) Hz in Figs. 7(b), (c), and (d), respectively. In Fig. 7(b), it can be seen that noise burst enhances the estimated velocity more than the mine itself: a burst of noise could then be mixed with a mine if considering only the spectrogram at one frequency and not in the whole time-frequency plane, where mine signature has a very characteristic shape. Now, as noticed in the former example, the mine locally stabilizes the velocity phase, as anywhere else, and, especially when burst noise occurs, the phase varies randomly [Fig. 7(c)]. This results in a very low coherence when scanning outside of the mine vicinity [Fig. 7(d)]. As burst noise results in bumps of the estimated velocity energy but also in a low coherence, it is suggested that an efficient time-frequency representation of the signal may be the coherence-weighted spectrogram that may be defined as

\[
CWS\!SP^r_{N,M}(t_i, f_i) = SP^r_{N,M}(t_i, f_i) \times C^r_{N,M}(t_i, f_i). \tag{11}
\]

This coherence-weighted spectrogram is plotted in Fig. 7(e). As a result, the mine signature appears with an enhanced contrast, but the artifacts due to burst noise are mostly removed.

Another example may be found in Fig. 8. The same scan as Fig. 4 (dirt road) is repeated at a higher sweep rate of 0.3 Hz (scanning speed around 2.2 km/h). In 10 s, the laser beam comes back and forth three times over the target. The smoothed spectrogram, time-frequency coherence, and coherence-weighted spectrogram are plotted respectively in Figs. 8(a)–(c). Mine signatures hardly appear in the spectrogram, as speckle noise causes a bad signal-to-noise ratio. The time-frequency coherence does not show the mine signature, because the contrast of coherence between off- and on-target locations is weak. On the other hand, the coherence-weighted spectrogram allows for detection each time the beam passes over the mine, and appears as a successful tool for analyzing the data. Interestingly, about time 2.3 s, the mine signature occurs at the same time as a violent burst noise, which results in masking the signature. One could think that such portions of the signal, corresponding to violent burst noise, consist essentially in noise and do not carry any acoustic information, but this is probably wrong as the mine signature is reconstituted after coherence-weighting. The coherence-weighting is basically a sort of filtering which acts at all frequencies. It removes contributions that are of an optical nature, more especially burst noise, and leaves in the spectrogram most of the energy which is induced from acoustic effect.
FIG. 7. Scan of a mine buried at 5 cm depth in loess soil (scanning speed 0.2 m/s). (a) Contour plot of smoothed spectrogram. (b) Smoothed spectrogram represented at frequency $f_2 = 145$ Hz. (c) Angle phase of the time varying cross-spectrum at frequency $f_2$. (d) Time-frequency coherence represented at frequency $f_2$. (e) Contour plot of coherence-weighted spectrogram.

One can conclude that the spectrogram takes exclusively into account variations of instantaneous amplitude of each single tone. Time-frequency coherence is mostly governed by phase stability but is less sensitive to an increase of energy; hence it does not appear as a representation with enough contrast of the signal to deliver “good” mine signatures. As for the coherence-weighted spectrogram, it combines both advantages of spectrogram and coherence, as it allows, in a single representation, to take into account both local increase of velocity amplitude and local stability of velocity phase. So far, the coherence-weighted spectrogram is the time-frequency representation that reconstitutes the mine signatures with the best quality in terms of shape and contrast.

It can be seen that an event in the time-frequency plane has to be considered under shape-based criteria to be validated as a mine signature. Now, in order to make an easier automation of mine detection, one can consider the marginal and moments of the representation: they provide one-variable functions (time functions) that can be more manageable than the time-frequency representation itself.

E. Use of marginal and moments

Let $D^i(t,f)$ be a real and positive time-frequency representation of the signal $s(t)$. The time marginal of this distribution is obtained by integrating the distribution over frequency for a given time:

$$P(t) = \int D^i(t,f) 2\pi df.$$  \hspace{1cm} (12)

The quantity $P(t)$ provides an estimation of the instantaneous energy of the signal, i.e., $\sum_{n=1}^{N} |A^2_n(t)|$, although it does not give an exact estimate if the distribution is a spectrogram.\textsuperscript{20} The instantaneous energy estimate is of great practical interest as it provides, for position $x = \lambda t$, an estimation of the amount of energy contained in the ground surface velocity. This quantity should be at the maximum when the beam scans over a buried object.

The local frequency $\langle f \rangle$, is the average frequency of the distribution for a given time $t$, or first conditional moment of the distribution:

$$\langle f \rangle = \frac{1}{P(t)} \int f D^i(t,f) df.$$  \hspace{1cm} (13)

This quantity is often called instantaneous frequency signal. In this case, as the signal is a sum of several single tones components, this notion is not of straightforward physical interpretation, but the first conditional moment is used in the calculation of the function

$$\sigma^2(t) = \frac{1}{P(t)} \int (f - \langle f \rangle)^2 D^i(t,f) df.$$  \hspace{1cm} (14)

This quantity provides an estimation of the instantaneous bandwidth: it represents the local spread in frequency for a given time $t$, and it will be seen in the following that this quantity, coupled with the instantaneous energy estimate, provides a robust clue for detecting the presence of a buried object. In this study, these two quantities are calculated from the coherence-weighted spectrogram.

Figure 9 provides an example of the use of the estimates of instantaneous energy and bandwidth. The same buried mine as in Fig. 4 is scanned again, but at a sweep rate of 0.48 Hz (about 3.6 km/h scanning speed). The vertical bold solid lines indicate the approximate instants for the passages of the beam over the target (about every second). The spectrogram and coherence-weighted spectrograms are plotted in Figs. 9(a) and (b). As expected, most of noncoherent signal struc-
tures are removed by coherence-weighting the spectrogram. Although the mine signature appears clearly most of the time, some signatures are somewhat ambiguous. The time marginal of the coherence-weighted spectrogram. (c) Time marginal of the coherence-weighted spectrogram. (d) First conditional moment of the coherence-weighted spectrogram.

In conclusion, a simple method for automatic recognition of the mine signature is the detection of the simultaneous occurrence of time marginal maxima and first conditional moment minima of the coherence-weighted spectrogram. For ambiguous cases, a more complicated shape-based criteria aiming at recognizing a signature in the time-frequency plane may lead to a final decision about mine presence.

V. PERFORMANCES AND LIMITATIONS OF THE METHOD

In order to assess the performance of the sweeping scan procedure for acoustic detection of landmines, measurements have been carried out on different types of grounds: a dirt road, a gravel road, and a lane of natural loess soil. Mines were buried at 5-cm depth, except for a set of experiments where the depth was 12.5 cm. The mine was always detected with sweep rates of 0.2 Hz and lower, which corresponds to a scanning speed of 1.4 km/h. Detection depends on the ground nature, the mine depth, and the ground surface nature.

Detection is easier in roads than in loess soil, as the acoustical properties of the soil are more variable in natural ground. Those variations can accompany zones of higher vibrational velocity, or so-called clutter, and can be interpreted as a buried object when scanning the ground only in one dimension (see Sec. IV B). This clutter can be recognized by constructing a two-dimensional image of the ground velocity, which can be achieved by scanning the ground with several parallel beams. When scanning grounds such as gravel or dirt roads, mines were detected with sweep rates up to 0.4 Hz (2.8 km/h) for mines buried at 5-cm depth. Detection can be achieved at even higher speeds in more favorable conditions (see Fig. 9 with a detection at 3.6 km/h scanning speed). For measurements over a deep mine (12.5 cm), the detection was not regularly achieved for speeds higher than 1.4 km/h. Indeed, the scattering of acoustic waves by buried objects gets weaker when the buried object is deeper. Hence, the contrast of velocity amplitude between off and on-target locations is weaker for deep mines than shallow mines. This is very clearly shown in Fig. 10. Two mines of the same type (M19)
were buried in the same gravel road, at depths of 5 and 12.5 cm. The laser beam moved back and forth over the location of the mine, and the spectrogram of the signal is given for the shallow (a) and deep (b) mines, for a scanning speed of 2.8 km/h. The peak energy is higher for the shallow mine than for the deep mine, which confirms that the contrast between on- and off-target locations is diminished when the depth is increased, making the detection more difficult.

The detection technique has been tested on various types of surfaces, including gravel roads, dirt roads, and natural soil. Even for a very rough surface such as a gravel road, mines can be reliably detected. This is remarkable as rough surfaces provide strong speckle noise at high scanning speeds (see, for example, Fig. 10 showing a scanning over a gravel road at 2.8 km/h scanning speed). A set of measurements investigating the influence of diverse elements on the surface has been carried out over a mine (VS16 type) buried at 5 cm depth in loess soil. Those elements were a thick and a thin layer of loose grass; a patch of grass just above the mine; and some pebbles or a big stone placed on the laser path. Figure 11 represents the smoothed spectrograms of the measured signals prior to coherence-filtering, in order to observe the effects of those objects placed on the laser beam path. Figure 11(a) shows the mine’s acoustic signature at 2.8 km/h scanning speed. Figure 11(b) shows the spectrogram obtained after a thick patch of grass (dimension are about a fourth of the object size) was placed above the mine; the components of the mine signature do not appear distinctly because of a broadband speckle noise burst due to the grass.

In Fig. 11(c), a thin layer of grass is scattered all along the laser path. In Fig. 11(d), a big stone (about 5 cm in diameter) was placed about 10 cm beside the mine (about t = 1.6 s for spectrogram). The stone produced a burst of speckle noise in the spectrogram. In all cases, after applying some coherence-filtering to the spectrograms of Fig. 11, the mine signature is reconstituted and most speckle artifacts are removed. In these experiments, only a thick layer of loose grass was able to prevent the detection.

VI. CONCLUSIONS

The study presented in this paper is concerned with the processing of the signal coming from a recently developed acoustic detector. This apparatus detects buried objects, such as landmines, by remotely sensing the variations of ground vibration that occur over a buried object with a LDV, when the ground is insonified by means of acoustic-to-seismic coupling. It is demonstrated in this paper that it is possible, with suitable signal processing, to detect landmines buried outdoors by using a continuous LDV scan. This step is vital in the development of the detector, for future integration to a moving vehicle, and for reaching scanning speeds compatible with real applications.

Time-frequency representations of collected signals deliver the characteristic “signatures” that permit the detection of temporal changes of energy and phase occurring when the laser beam scans over a buried object. Indeed, the spectrum temporal variations contain the clues for detection, and time is directly interpretable in terms of spatial position. Energy enhancements are detected using smoothed spectrograms with adequate smoothing length. Experimental data show that vibration phase is relatively stable over buried objects and unsteady elsewhere. Thus time-frequency coherence exhibits maxima when scanning over objects; it also filters out efficiently most speckle noise contributions. Finally, the coherence-weighted spectrogram appears to deliver the most contrasted signature by fusing together the searched clues in terms of energy and phase variations. Moreover, the time marginal and first conditional moment of this representation provides one-variable functions whose extrema indicate efficiently the presence of a mine, and thus make possible a quite simple automation of the detection process.

In the tests performed in this study, detection is systematically achieved in natural ground and roads for mines buried at 5-cm depth at 2.8 km/h scanning speeds. This speed can reach 3.6 km/h in favorable conditions, and has to be decreased to 1.4 km/h for deep mines (12.5 cm depth).

At this time, a new detector is being developed consisting of multiple parallel LDV beams that scan simultaneously. A two-dimensional vibration map, similar to that in Fig. 1, will then be constructed from a single sweeping measurement, by using the signal processing method presented in this paper.

ACKNOWLEDGMENTS

This work is supported by the U.S. Army Communications-Electronics Command, Night Vision and Electronics Sensors Directorate. The authors thank Ronald Craig for his dedication and expertise in collecting the data for this paper. Any opinions, findings and conclusions, or
recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of U.S. Army Communications-Electronics Command.


