# ON THE ESTIMATION OF TIME-DELAYS FOR IMAGE SOURCE METHOD

*A. Drira\*, A.O. Boudraa, L. Guillon and A. Komaty*

IRENav, Ecole Navale, BCRM Brest, CC 600, 29240 Brest Cedex 9, France.
{achraf.drira, boudra, guillon, ali.komaty}@ecole-navale.fr

## ABSTRACT

The image source method is based on the study of signals reflected by a sedimentary seabed. It produces an estimation of the geoacoustic parameters of the stratified sediments. The key step of this method is the detection and estimation of the time delays of the signals reflected by the interfaces between the sedimentary strata. The Teager-Kaiser (TK) operator is known as a reliable way to extract from the recorded signal the time delays of the reflected signals. Due to its sensitivity to discontinuities and excellent time resolution, this operator is a good tool to detect and estimate the time delays. We investigate whether the TK operator associated with time-frequency approaches (spectrogram, wavelets) could be of interest to the detection of time delays and perform better than the TK operator applied directly on the reflected signals. The different methods are presented and compared first on synthetic signals. The results show that the use of wavelet transform improve the quality of the time-delays estimation. The algorithm is finally applied to real signals acquired in an acoustic tank experiment and the estimated parameters are very close to the ground-truth.

*Index Terms*— Image method, Medium impulse response, Geoacoustic inversion, Time-delay, Teager-Kaiser energy operator.

## 1. INTRODUCTION

The image source model (ISM) has become a well established analysis tool in many fields of acoustics and engineering [1]. This model states that the reflections of sound, from acoustic source, can be viewed as direct signals coming from *virtual sources*, placed symmetrically on the far side, of each wall, as mirror image sources [2]. Originally, the ISM was used to compute the room impulse response (IR) from an acoustic source to a receiver in a room by simulation [2]. The interest of the ISM relies on its conceptual simplicity and to its easy algorithmic implementation, which make the ISM-based algorithms relatively straightforward to implement [1]. Consequently, the ISM technique has been used as a basis principle for a wide range of purposes such as prediction of sound propagation, noise control, speech intelligibility and transmission, and architectural modeling and design [1].

Recently a geoacoustic inversion strategy in underwater acoustic based on the ISM has been introduced and validated on both synthetic and real data [3], [4]. More precisely, the acoustic IR or media IR (MIR) models the propagation of the sound from an acoustic source to a hydrophone. This MIR is a sum of delayed and attenuated Dirac impulses which correspond to the direct path from the source to the hydrophone and to all the reflections from the seafloor and the layer interfaces. A crucial step of the ISM inversion strategy is the estimation of the time delays which correspond to the positions of the peaks associated to the different reflections. In [4] Teager-Kaiser (TK) energy operator [5] has been used for peak detection due to its excellent time resolution and fast computation. An important aspect of this operator is that it amplifies discontinuities and sudden changes in amplitude while the smooth transitions between the samples are reduced. The peaks of the TK energy operator output correspond to the time delays. Even interesting inversion results are obtained, the performances of the inversion method can be affected in very noisy environments. Due to sensitivity of the TK energy operator to noise of high level, spurious peaks (reflections) can be produced which provide wrong time-delay estimation. This is also the case in presence of impulsive noise as encountered in underwater acoustic. A solution to this problem can be the denoising of the received signals on the hydrophones. However the denoising probably can make the expected denoising effects unsatisfying due to over-denoising or under-denoising. The over denoising can affect the temporal structure of arrivals and under denoising does not reduce the number of false reflections. For more robustness to noise, in this paper the detection of peaks is done in time-scale (time-frequency) domain where the TK energy operator is applied to the coefficients at each scale. This is essentially motivated by the fact that sub-band signals derived, for example, from a wavelet decomposition are narrow band where the TK energy operator is more efficient for peak detection than on the original signal which is wideband.

## 2. ISM-INVERSION

The ISM-inversion is a recently developed technique for geoacoustic inversion [3], which gives the local sound speed profile with a good resolution and low computational cost

compared to conventional approaches [6]. The inversion is based on the analysis of an acoustic wave emitted by a broadband source and reflected by a layered seabed [4] where a single recording of the wave reflected at the bottom on a hydrophone array is used (Fig. 1a). Under the Born approximation [3], the reflected signal can be modeled as a sum of contributions from image sources which are symmetrical to the actual source relative to the seabed layers (Fig. 1b). The locations of these image sources are related to the Sound Speed Profile (SSP) of these layers. Thus, this structure can be reversed by detecting the image sources. The first step of the ISM-inversion is the detection and the location of image sources in a homogeneous medium (Fig. 1b). The location of the source images depends on the knowledge of the geometry of the array, which allows us to determine the two key inversion parameters: the time delay $(t)$ and the arrival angles $\theta$ of each source image $i$ to each hydrophone of the antenna. With these two parameters, the thickness and sound speed of each layer are found recursively, from the first one to the deepest one, with formula derived from the Snell-Descartes law of refraction.
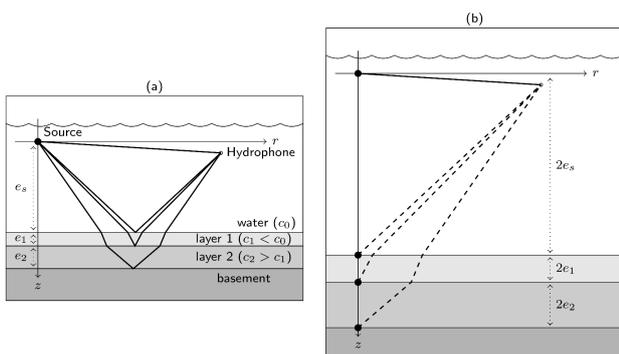


**Fig. 1**. (a) Reflection of a spherical source on a stratified background under the Born approximation. (b) System equivalent to image sources taking into account the refraction.
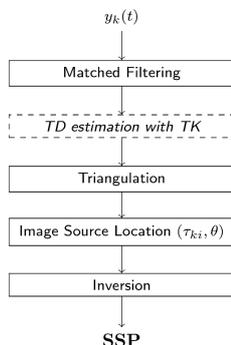


**Fig. 2**. Flowchart diagram of the ISM-inversion.

## 3. MODELING

The acoustic response of the underwater channel can be described by a family of IR (or MIR) which model the channel between a source and a hydrophone. The multi-reflection phenomenon causes a signal to arrive at a hydrophone via different paths with different time delays [7]. This effect can be modeled as the output of a finite impulse filter (FIR) with a MIR, corresponding to the $k^{th}$ hydrophone, given by:

$$h_k(t) = \sum_{i=1}^{L} \alpha_{k,i} \delta(t - \tau_{k,i}) \qquad (1)$$

where $\delta(t)$ denotes the Delta function, $\alpha_{k,i}$ and $\tau_{k,i}$ represent respectively the attenuation coefficient and the propagation time (time delay) of the $i^{th}$ path. $L$ is the FIR length which is the equivalent of the paths identified here as the number of layer interfaces. The direct path corresponds to the index $i = 1$ in equation (1) and the reflections correspond to $i \neq 1$. Using the MIR, we can find the signal received by the $k^{th}$ hydrophone as a convolution between the wideband signal $x(t)$, emitted by the source, and the MIR [8]:

$$\begin{aligned} y_k(t) &= \int x(s) h_k(t - s) ds + n(t) \\ y_k(t) &= \sum_{i=1}^{L} \alpha_{k,i} x(t - \tau_{k,i}) + n(t) \end{aligned} \qquad (2)$$

Note that the ISM inversion is limited to the first-order reflections [3], [4].

## 4. TK OPERATOR IN TIME-FREQUENCY DOMAIN

It has recently been shown the interest of time-frequency representations (TFRs) for peak detection or source separation [9]. We investigate in this work the interest to combine the TK energy operator and the resolution of a TFR such spectrogram or scalogram for peak detection and particularly in noisy environments. One advantage of these two TFRs is they are free of cross terms.

### 4.1. STFT-CWT

The STFT is a TFR widely used for studying non-stationary signals. It is obtained by applying the Fourier transform by a fixed-sized, moving window to input signal. The STFT gives a 2D spectra for time-frequency analysis and is defined by

$$Y_s(t, f) = \int_{-\infty}^{+\infty} y(t) w_s(t - \tau) e^{-j2\pi ft} dt \qquad (3)$$

where $w_s(.)$ is sliding window of a fixed length. The energy density surface of the STFT, called spectrogram, is given by

$$S(t, f) = |Y_s(t, f)|^2 \qquad (4)$$

As the STFT, the continuous wavelet transform (CWT) is a well established tool for time-frequency signal analysis. It allows high localization in time of high frequency signal features, unlike the STFT [10]. The CWT has gained attractiveness in representing and preserving the signal energy in the presence of noise. The CWT of a signal $y(t)$ is defined as:

$$T(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} y(t)\Phi^*\left(\frac{t-b}{a}\right) dt \qquad (5)$$

where $\Phi^*(t)$ is the complex conjugate of the analyzing wavelet function $\Phi(t)$. $T(a,b)$ represents the wavelet coefficient, $a$ and $b$ are the dilation and location parameters respectively. The contribution to the energy of $y(t)$ at the specific $a$ scale and $b$ locations is given by the 2-D wavelet energy density function known as the scalogram:

$$E_w(a,b) = |T(a,b)|^2 \qquad (6)$$

The relative contribution to the total energy of $y(t)$ at $(a,b)$ is given by the variance of the wavelet defined for the CWT as the scale-dependent energy distribution by

$$E(a) = \frac{1}{C_g} \int_{-\infty}^{+\infty} E_w(a,b)db \qquad (7)$$

where $C_g$ is the admissibility constant. In practice $E(a)$ is used to determine the dominant energetic scales within $y(t)$.

### 4.2. TK energy operator

The TK energy operator is a nonlinear energy tracking operator and its output to a given signal, $y(t)$, approximates the actual physical energy required to produce $y(t)$. In continuous time form of the TK energy operator given by:

$$\Psi[y(t)] = \left(\frac{dy(t)}{dt}\right)^2 - y(t)\frac{d^2y(t)}{dt^2} \qquad (8)$$

It is evident from equation 8 that this operator is expected to suppress slowly varying parts of a signal and highlight abruptly changing parts of this signal.

### 4.3. TK-CWT

In this strategy, the peaks are detected by application of the operator on the dominant scale $\Psi[E_w(\tilde{a},b)]$ instead on the noisy signal itself $\Psi[y(t)]$ where $\tilde{a} = \arg\max_a[E(a)]$.

### 4.4. TK-Spectrogram

The spectrogram is considered as an image in which the peaks are detected. The application of the TK energy operator on the spectrogram can be viewed as an enhancement operation of this image representation and thus improves the robustness of the peak detection. More precisely, the operator is applied

along the rows (time) and columns (frequency) of $X(t,f)$ and the filtering results summed as follows:

$$TS(t,f) = \Psi_t[S(t,f)] + \Psi_f[S(t,f)] \qquad (9)$$

where $\Psi_l[.]$ is the operator applied along to the $l$ component. The created time-frequency mask for time-delay estimation is calculated as follows:

$$M(t,f) := \begin{cases} 1 & \text{if} \quad TS(t,f) \geq \texttt{th} \\ 0 & \text{otherwise} \end{cases} \qquad (10)$$

where $\texttt{th} = \texttt{mean}[\texttt{mean}(TS(t,f))]$ is an automatic threshold used to isolate the most prominent peaks that are considered as the reflections [4].

## 5. RESULTS

Three methods are compared, namely, TK applied directly on $y(t)$, TK-CWT and TK-Spectrogram on synthetic and real data. The flowchart diagram of the ISM-inversion is shown in figure 2. The CWT is computed using the Mexican hat mother wavelet and the output of the TK energy operator of the dominant scale of the signal calculated, $\Psi[E_w(\tilde{a},b)]$.

### 5.1. Synthetic data

Synthetic data used are simulated using numerical evolution of the Sommerfeld integral and under the Born approximation. A broadband source, 12 m over the seafloor composed of $N = 2$ fluid layers is used. The array is horizontal made of 15 hydrophones at the source height as the source and 24 m away from it. An example of synthetic signal corrupted with a white Gaussian noise, $n(t)$, with SNR set to -9dB is shown in figure 3. The associated scale-dependent energy distribution, $E(a)$, is illustrated by figure 4. The peak of $E(a)$ is located at $a = 8$ which corresponds to the dominant energetic scale. Thus the peak detection is performed on this scale. The TK energy operator is applied to component of scale $a = 8$ and the results are illustrated in figure 5. This figure shows that the TK-CWT reduces the number of false reflections compared to the TK method where the operator is applied directly on the $y(t)$ signal. For each SNR and each scale, $a$ the relative error between the true time delay and the estimated one by the method TK-CWT is shown in figure 6. This figure shows that the relative errors depend on the scales and the better estimates are obtained for the dominant energetic scales ($a \geq 8$) and this confirms the findings of figure 4. Similarly, results of TK-Spectrogram are shown in figure 7. As shown in this figure, as the window size of the STFT decreases, the more precise the time-delay estimates become. These two figures show that the relative errors are very small even for low SNR values showing the interest of the combining effect of the TK and a TFR for peak detection in noisy environments.
Figure 8 shows the spectrogram of $y(t)$ with SNR=-3dB and the associated extracted mask $M(t,f)$. The peaks are well evidenced on the mask (Fig. 8).
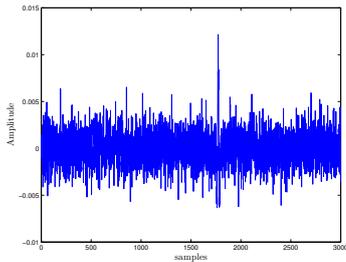
**Fig. 3**. Received signal $y(t)$ on the first hydrophone.
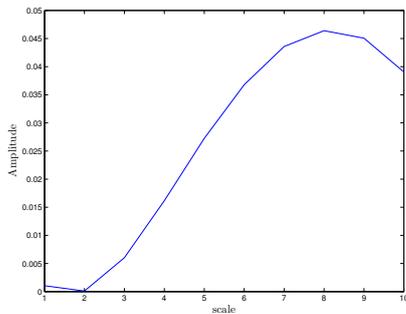


**Fig. 4**. Plot of $E(a)$ of the different sub-band (scales) $a$ extracted from $y(t)$.
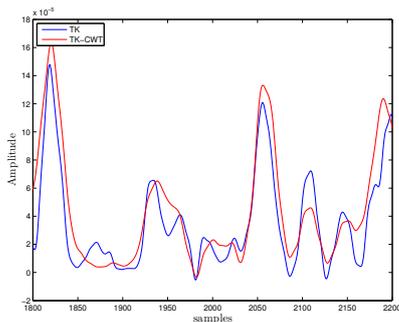


**Fig. 5**. TK applied to the received signal $y(t)$ and to its component of scale $a = 8$.
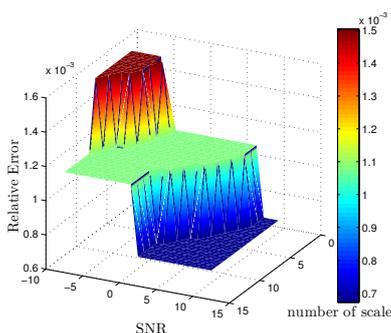


**Fig. 6**. Relative errors of time-delays estimated by TK-CWT.

## 5.2. Real data

To investigate the feasibility of the proposed method and compare the performance of different time delays detection on real
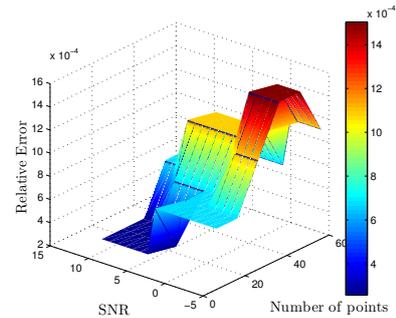


**Fig. 7**. Relative errors of time-delays estimated by TK-Spectrogram.
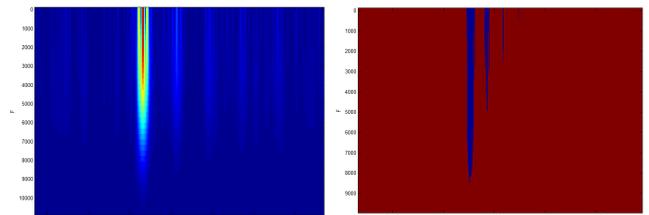


(a) Spectrogram of $y(t)$.　　　(b) Extracted masque $M(t, f)$.

**Fig. 8**. TK-Spectrogram.

data, experiments have been carried out in water tank. The principle is to reproduce in a controlled environment a configuration similar to what one might have in the sea: a point source, a hydrophone, an array and a layered seafloor. The advantage is to benefit from an ideal laminated floor whose characteristics are known. The signals are recorded with a single hydrophone whose position is controlled. A synthetic array is then reconstituted by moving the hydrophone horizontally. The dimensions of the tank used for the experiment are 2 m wide, 3 m long and 1.5 m in height. The choice of hydrophone was a spherical piezoelectric type, 2 cm in diameter. As for the receiver, a needle transducer Precision Acoustics was chosen. The laminated medium is formed of two plates consisting of different materials. Their dimensions are 50 cm wide, 50 cm long and of different thicknesses (Table 1). The first plate is made of Teflon, and the second is made of Resin. Based on these parameters, one can equate the Teflon plate to a layer of sediment and the resin plate to a solid core. The ISM-inversion consists in analysis the time delays of a pulse signal emitted by a point source. In order to increase the precision of the time delays detection, the signal sent from the equipment used is a period of sinus at frequency of 150 kHz with a duration of 6.6 e-006s. An example of a signal recorded by the hydrophone is shown in figure 3. The IR of the system composed of plates is then obtained after a Wiener deconvolution:

$$H(f) = \frac{Y(f)X^*(f)}{|X(f)|^2 + \epsilon} W(f) \qquad (11)$$
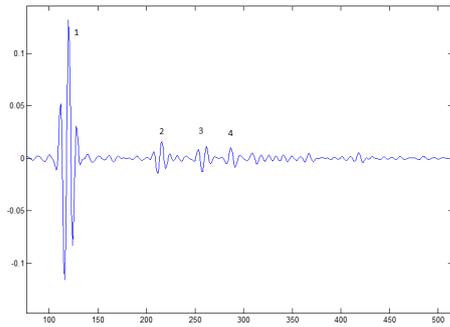
**Fig. 9**. Signal received after Wiener deconvolution. 1: Direct 2: first reflexion 3: reflection between the two plates. 4: reflection between the second plate and the seafloor.

where $Y(f)$ and $X(t)$ are the Fourier transforms of $y(t)$ and $x(t)$ respectively. $W(f)$ is a bandwidth Gaussian spectrum of a weighting window and $\epsilon$ is a constant used to limit the influence of noise on the areas of the spectrum where there is little energy. This constant is determined empirically. The signal obtained after the Wiener deconvolution (Eq. 11) in shown in figure 9. In addition, the temporal resolution is good enough to visualize three reflections on the panel system interfaces (arrivals 2, 3, and 4).

| Experiment | Teflon | Resin |
|---|---|---|
| SS (m/s) | $1433 \pm 10$ | 2025 |
| Density | $2178 \pm 45$ | $1928 \pm 45$ |
| Thickness (m) | 4.35 | 4.47 |

**Table 1**. Geoacoustic parameters of the two plates.

| Method | Parameters | Water | Layer 1 | Layer 2 |
|---|---|---|---|---|
| TK-CWT | SS (m/s) | 1475.4 | 1438.9 | 2028.2 |
| | Thickness (m) | 0.161 | 0.053 | 0.045 |
| TK | SS (m/s) | 1475.4 | 1439.5 | 2036.6 |
| | Thickness (m) | 0.161 | 0.0533 | 0.0474 |

**Table 2**. Geoacoustic parameters for real data estimated by TK and TK-CWT (SS: Sound Speed).

The SSP obtained with the ISM using the new approach (TK-CWT) to find the time delay are presented in Table 2. The results are compared with the results found with the TK applied directly on the signals. The comparison is very satisfactory. In particular, for the TK-CWT approach, the sound speeds are very close to the ones measured directly at 1 MHz : the difference is below 0.6% . The estimation of the thicknesses is less accurate for the two methods but still satisfactory for the second layer. The problem of thickness estimation of the first layer should be address in a future work.

## 6. CONCLUSIONS

In this paper, we present a study of TFRs (CWT and Spectrogram) associated with TK operator to detect and estimate time-delays in the context of a geoacoustic inversion method. The results obtained on both synthetic and real signals lead to the following conclusions. TK-CWT is effective to find the desired peaks, even for low SNR (- 9 dB), and it reduces the number of false alarms. The supplementary operations are fast enough to not affect the total computational cost. TK-Spectrogram is efficient for SNR above -4 dB and it involves an increase in computational cost. For the three algorithms (TK, TK-CWT, TK-Spectrogram), the threshold is computed automatically from the data themselves. The SSP inversion performed on real data acquired in a tank experiment are very satisfactory compared to the ground truth.

## REFERENCES

[1] E.A. Lehmann and A. Johansson, "Prediction of energy decay in room impulse responses simulated with an image-source model," *J. Acoust. Soc. Am.,* vol. 124, pp. 269-277, 2008.

[2] J. Allen and D. Berkeley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.,* vol. 64, no. 4, pp. 943-950, 1979.

[3] S. Pinson and L. Guillon, "Sound speed profile characterization by the image source method," *J. Acoust. Soc. Am.,* vol. 128, no. 4, pp. 1685-1693, 2010.

[4] A. Drira, L. Guillon and A.O. Boudraa, "Image source detection for geoacoustic inversion by the Teager-Kaiser energy operator," *J. Acoust. Soc. Am.,* vol. 135, no. 6, pp. 258-264, 2014.

[5] J.F. Kaiser, "Some useful properties of Teager's energy operators," *Proc. ICASSP,* vol. 3, pp. 149-152, 1993.

[6] Ch. Holland and J. Osler, "High resolution geoacoustic inversion in shallow water: A joint time and frequency domain technique," *J. Acoust. Soc. Am.,* vol. 107, no. 3, pp. 1263-1279, 2010.

[7] J. Smith and B. Friedlander *Adaptive Multipath Delay estimation. Reprinted in Coherence and Time Delay Estimation.* IEEE Press Book, 1993.

[8] O. Ocal, I. Domanic and M. Vetterli, "Source localization and tracking in non-convex rooms," *Proc. ICASSP,* pp. 1429-1433, 2014.

[9] O. Yilmaz and S. Richard, "Blind separation of speech mixtures via time-frequency masks," *IEEE Trans. Sig. Proc.,* vol. 52, no. 7, pp. 1830-1847, 2004.

[10] P.S. Addison, J. Walker and R.C. Guido, "Time-frequency analysis of biosignals," *IEEE Eng. Med. Biol. Mag.,* vol. 28, no. 55, pp. 14-29, 2009.