On the use of time-scale representations for the analysis of seismic signals

Antoine Roueff, Jocelyn Chanussot, and Jérôme Mars
Laboratoire des Images et des Signaux, Grenoble, France

ABSTRACT
The presented study, based on the continuous wavelet transform, introduce new algorithms which perform different kinds of separation processing depending on the nature of the seismic data. When dealing with a one dimensional trace, we propose a segmentation of its time-scale representation, which leads to the automatic detection and separation of the different waves present in the trace. This algorithm can be applied to a whole seismic profile containing several traces, by tracking the segmentation features in the time-scale image sequence from one trace to the next. The resulting separation algorithm is efficient as long as the patterns of the different waves do not overlap in the time-scale plane. Afterwards, the purpose is to take into account the redundancy of information in more dimensional data to increase the separation possibilities in presence of interference. In the case of vectorial sensors, we use the polarization information to separate the different waves using phase shifts, rotations, and amplifications. At last, in the case of linear array data, we use the propagation velocity information to separate propagating waves which cross. For this problem in particular, we focus on the case where waves are dispersives and we propose a new time-scale representation which enable the estimation of the wave dispersion from a small array of sensors.

Keywords: time-scale plane, seismic processing, surface wave extractions, segmentation, polarization, dispersion.

1. INTRODUCTION

The estimation of the subsoil characteristics and physical properties is a very important task in many various fields (geology, risk evaluation before a building construction, petroleum exploration). Seismic prospecting mainly consists in generating very low amplitude artificial sources at predetermined times and positions. After propagation in the subsoil, the induced seismic disturbances are recorded by a set of sensors regularly placed on the ground. By analyzing the recorded signals, different waves can be identified (surface waves, reflected waves...). Figure 1 presents an example of a real seismic data and its recording schema. On the profile image (on (b)), the horizontal axis is the time recording, and on the vertical axis is the number of sensor which can be seen as a distance axis since the sensors are linearly and equally spaced.

The main goal of this study is the extraction of the surface waves from the data. This analysis has two applications. Firstly, in geology, the parameters of the dispersive waves make it possible to invert for the shear velocity as a function of depth. Secondly, in geophysics, those energetic surface waves are considered as noise, so their extraction improve greatly the Signal to Noise Ratio (SNR). Most of the times, the different waves recorded by a given sensor overlap, both in the time and frequency domains. Therefore, there is a strong need for another representation where the different waves have disjoint supports and can thus be separated. The continuous wavelet transform turned out to be extremely well suited for this task.

In this paper, we present an overview of some recent developments which use the continuous wavelet transform (CWT) to describe the content of the data. The CWT of a given signal $s$ is:

$$\text{CWT}_s(b,a) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} s(t) \psi^*(t) \left(\frac{t-b}{a}\right) dt,$$

Further author information, send E-mails to: antoine.roueff@lis.inpg.fr, jocelyn.chanussot@lis.inpg.fr, jerome.mars@lis.inpg.fr, Address: 38402 Saint Martin d’Hères Cédex France CNRS, UMR 5083, OSUG
where $\psi$ the wavelet, $b$ is a delay argument which can be interpreted as the time argument, and $a$ is the scale argument, it is inversely proportional to the frequency.

Thus, the description of $s$ is given by the presentation of the squared modulus of the two dimensional function $CWT(a, b)$. The induced image is called a scalogram. The presented methods could be adapted to other time-frequency representations. However, compared with other time-frequency transformations, the advantages of the CWT is its linearity which results in the absence of interference cross-term and its dyadic pavement of the time-frequency plane which enable a good time-frequency resolution. This adaptative resolution is actually the reason of the definition of the CWT by Morlet in 1982. In our approach, we did not use the discrete wavelet transform because the absence of redundancy in the induced time-frequency image prevent us from a detailed description. In the geophysical literature it appears that the main advantage of the discrete transform is the calculation speed which is a crucial property when dealing with large amount of data.

For all the applications, we use the Morlet wavelet. It is very similar to a natural seismic wavelet and is therefore appropriate for the analysis of seismic data. Furthermore, reaching the Heisenberg Gabor bound, it offers a perfect trade-off between the spectral and the temporal resolutions. In addition, in order to separate the amplitude information from the phase information, we use the Morlet wavelet in its analytic complex form:

$$\psi(t) = (\pi t_0)^{-1/4} \exp \left[ -\frac{1}{2} \left( \frac{t}{t_0} \right)^2 + 2i\pi \nu_0 t \right].$$

where $\nu_0$ is the central frequency of the passband filter $\psi$, and $t_0$ is the linked scale (i.e. frequency) of the mother wavelet.

In the development are presented different approaches in order to perform the wave separation depending of the nature of the data. Section 2 presents a segmentation of the time-frequency image which enable the automatic separation of the different waves present in a seismic trace. Applied to the different traces of a whole seismic profile, the resulting algorithm is efficient as long as the patterns of the different waves do not overlap in the time-frequency plane. Then, in the next sections the purpose is to use some redundancy of information in more dimensional data in order to increase the separation possibilities in this difficult case where wave patterns interference in the time-scale plane. In section 3 we introduce the use of the polarization information in order to separate waves in the case of vectorial sensors (sensors which record vibrations in different directions). Finally in section 4, we propose another alternative which is, in the case of linear array data, to use the velocity information to separate the crossing waves. For this problem, we focus on the case where waves are dispersives and we propose a new time-scale representation which enable the estimation of the wave dispersion from a small array of sensors.
2. WAVES SEPARATION BY TIME-FREQUENCY FILTERING

2.1. Time-Frequency Filtering

In this section, we assume that the studied signal \( s \) is the sum of several waves \( w_k \) \( (s(t) = \sum_{k=1}^{N_w} w_k(t)) \), where \( N_w \) is the number of wave present in \( s \). Assuming the different patterns of the different waves do not overlap in the time-frequency plane, then different possibilities exist in order to separate the waves.\(^9\) We decide to use the algorithm of Nguyen.\(^4\) This algorithm estimate one wave \( w \) in three steps:

- Computation of the continuous wavelet transform \( CWT_s(a,b) \).
- Definition of the pass region \( R(a,b) \) in which the wave \( w \) is.
- Estimation of the inverse continuous wavelet transform (ICWT) of \( CWT_s(a,b).R(a,b) \).

\[ \text{(a) Initial trace in the time-scale plane } |CWT_s|. \]
\[ \text{(b) Pass region } R(a,b). \]
\[ \text{(c) Initial trace } s \text{ in plain, and estimated wave } \hat{w} \text{ in dashed.} \]
\[ \text{(d) Scalogram of the estimated wave } |CWT_{\hat{w}}|. \]
\[ \text{(e) Label image.} \]

**Figure 2.** Extraction of a wave \( w \) from a real trace \( s \) by time-frequency filtering.

This algorithm is illustrated on figure 2. From the initial signal \( s \) presented on (a), we extract the wave \( w \) which is located in the pass region presented in black on figure (b). On (d) is presented the CWT of the estimated wave \( \hat{w} \); one can check that the estimation is similar to what we could expect (see figure (a) and (d)). On (c) is presented the result in time. A similar procedure can be done for each wave. Then, in order to automatize the mask selection of each wave in the time-frequency image, we propose a segmentation of the scalogram image. This will lead to a label image (see figure 2(e)). That is developed in the next section.

2.2. Partition of the Time-Scale Plane

In order to make a partition of the time-frequency image, we choose to use the segmentation algorithm called the watershed algorithm.\(^{10}\) This tool is designed to clusterize all pixels which are connected to the same local
minima. The one dimensional illustration of this algorithm presented on figure 3 shows that the segmentation process can be seen as an immersion process. Each region grows leaving from one local minimum which is considered as a seed, and when two regions are about to mix, a dam is build in order to separate them. This algorithm can be extended in two dimensions and thus can be applied to the negative scalogram image, such as it leads to the definition of different regions located around the local maxima of the scalogram. The application of the algorithm on the previous real trace is presented on figure 2(e). Finally, using the different patterns defined by the segmentation, each wave can be extracted.

2.3. Application to a Seismic Profile
To apply the previous tool to a profile containing several traces, we need to segment all the time-scale image of the different traces. The stacking of the different scalograms is presented on figure 4. On the foreground is the time-scale image of the last trace of the profile (number 47 on figure 1), and on the background is the time-scale image of the first trace. The image is presented in perspective, thus on the horizontal axis is time, on vertical axis is frequency and on the depth axis is the distance.

![Figure 3. Illustration in four steps of the immersion algorithm.](image)

![Figure 4. Representation of the profile in the volume time-frequency-distance](image)

It is necessary to segment each of these scalograms. In order to follow the information from one image (i.e. sensor) to the next, we designed an algorithm which tracks the seeds of segmentation from one (labelled)
scalogram to the next (unlabelled) scalogram. It is based on the spatial continuity. This tracking enable to group the labels along the sensors. From a practical point of view, it is crucial since it enables to automatize the algorithm in the distance dimension.

2.4. Results and Discussion

Our unsupervised algorithm is applied on the real seismic profile of figure 1. This leads to the four profile presented on figure 5. As a result, the four waves have been separated efficiently. The white area which are present on the first sensors for the reflected wave, and the fast Rayleigh wave are due to the fact that on the first sensors, the patterns of the waves overlap in the time-frequency plane. This example shows the limitation of this one dimensional filtering, when waves are not separated in the time-frequency plane, a time-frequency filtering cannot separate them.

![Graphs showing four types of waves](image)

*Figure 5. Results of the separation of the four waves by time-frequency filtering.*

If we make an evaluation of this method, it appears that, except for the initial step where the user select the seeds to identify the different waves he want to extract, the rest of the algorithm is automatic. The segmentation allows to automatize the design of the time-frequency pass region, and the seeds tracking enable to link the labels along the sensors. This automation is an important practical advantage. In addition, compared with standard methods such as two dimensional Fourier filtering which introduce artifacts because of the aliasing and the evolution of the wave along the sensors, our method does not suffer from these problem because the filtering is done on each trace separately. However, the main issue that remains is that when the patterns of the waves overlap in the time-frequency plane, we cannot separate them. To solve this issue, we make in the next section some more complex models which enable to get another degree of freedom to separate the waves.
3. POLARIZATION FILTERING

3.1. Context and Description of the Algorithm

In this section, we consider sensors which record vibrations in two directions (for example horizontal and vertical). Let $s_h$ and $s_v$ be both signals recorded at a two component sensor. In presence of one wave, the model is:

$$
\begin{bmatrix}
    s_h^a(t) \\
    s_v^a(t)
\end{bmatrix} = \frac{1}{\sqrt{1 + \rho^2}} \begin{bmatrix} 1 \rho e^{i\theta} \end{bmatrix} w^a(t),
$$

where $\rho$ and $\theta$ are the polarization parameters and $w$ the wave. Note that $s^a$ denotes the analytic signals of $s$.

This model assumes that the recorded signals on each component are identical, close to one phase shift $\theta$ and one amplification $\rho$. The use of analytic forms allows to model the phase shift with a simple multiplication by a complex exponential $e^{i\theta}$. This model implies that polarization of the wave depends neither on time nor on frequency. Figure 6 presents three examples of polarization:

- on (a) and (d) is the most general case where $\rho \neq 0$ and $\theta \neq 0$. Signals are not in phase. Polarization is called elliptic.
- on (b) and (e) is the linear case where $\rho \neq 0$ and $\theta = 0$. Signals are in phase.
- on (c) and (f) is the case $\rho = 0$: the wave is projected on one component.

The plots of the function $t \mapsto (s_h(t), s_v(t))$ presented on figures 6(d), (e), (f) are called Lissajous plots. They allow to characterize visually the polarization.

![Figures](image1.png)

**Figure 6.** Examples of three kinds of polarization. On (a), (b), (c) are presented in solid $s_h(t)$, and in dash $s_v(t)$. On (d), (e), (f) are presented Lissajous plots.

Figure 6 shows that leaving from the general case ($\rho \neq 0$, $\theta \neq 0$), it is possible to come back to the case where the wave is projected on one component ($\rho = 0$) by applying a phase shift and a rotation. The phase shift is $-\theta$ and the angle of rotation is $\arctan 1/\rho$. Then, in presence of two waves, it is possible to separate both waves by projecting the first wave on one component and the second wave on the other component. This operation is applied using a combination of phase shifts, rotations, and amplifications. The name of this filter is the oblique polarization filter. The only condition of success of the separation is that the model of equation (3) stands
for both waves and that the polarization parameters of both waves are different. It does not require both waves to be separated in the time-frequency plane.

An illustration of this filter is presented on figure 7. Two components are presented. Each component is the sum of two waves: the first wave is around the sample 60, and the second wave is around is sample 210. The algorithm enables to separate perfectly both waves. The only difficulty of this method is that it requires to know the polarization parameters of each wave. For this example, since waves have different time support, the polarization parameters can easily be found by inter-correlation in the time area where one wave is present, or it can be read from the Lissajous plots. However, when both waves get closer to each other, and have the same time-support, the estimation is more difficult. So, in spite of the separation possibilities, since the polarization estimation is usually performed in time, when both waves have the same time support, this method cannot be used in practice to solve the separation problem.

Then, in order to improve the estimation of the polarization parameters, and thus increase the possibilities of the oblique polarization filter, we introduce a method which performs the estimation of the polarization parameters, not in time but in the time-frequency plane, where there are more possibilities to find an area where waves do not overlap. As an estimation of the polarization parameters, we propose to use the cross product between the linear time-frequency representations of $s_h$ and $s_v$ noted $TFR_h$ and $TFR_v$. As a transformation, one can use either the short time Fourier transform (STFT) or the CWT. This kind of approach has similarly been considered in the geology community by Keilis-Borok. We note $C$ the cross product (which corresponds to the inter-scalogram if the CWT is used and the inter-spectrogram is the STFT is used):

$$C(t, \nu) = \frac{TFR_v(t)TFR_h^*(t, \nu)}{|TFR_h(t, \nu)|^2},$$

(4)

Then, we define as a $\theta$ image in the time-frequency plane:

$$\Theta(t, \nu) = \angle(TFR_v, TFR_h)(t, \nu),$$

(5)

and as a $\rho$ image in the time-frequency plane:

$$R(t, \nu) = \frac{|TFR_v(t, \nu)|}{|TFR_h(t, \nu)|}.$$  

(6)

$\Theta(t, \nu)$ and $R(t, \nu)$ give some relevant description of the polarization of the data in the time-frequency plane. When this information is added to the amplitude information from module images, it allows us to characterize a region where the contribution of the wave is paramount, and where $\theta$ and $\rho$ can be estimated.
3.2. Results and Discussion

To illustrate the algorithm we present a synthetic example. We consider two signals \( s_h \) and \( s_v \), which are the sum of two waves plus one small Gaussian noise. Both waves are linear chirps whose envelope is Gaussian. But one has an increasing frequency modulation and the other one has a decreasing frequency modulation. The Signal to Noise Ratio (SNR) between the chirps and the Gaussian added noise is 8 dB. In order to fit the resolution of the chirps, the time-frequency representation used is the Short Time Fourier Transform (STFT). The time representation of each component is impossible to interpret because both waves are in the same time interval. On spectrograms (presented figure 8(a) and (e)), the patterns of both waves can be recognized, but their separation by time-frequency filtering is only partial because both patterns overlap in the center of the image.

The \( \Theta(t,\nu) \) and \( R(t,\nu) \) images are respectively presented on figure 8(b) and (f). On both images, for each wave, a region corresponds to the coordinates where the amplitude of one wave is bigger than the other wave and than the noise. This region is the top left and bottom right for the chirp with decreasing frequency modulation, and the top right and bottom left for the other wave. In those regions, \( \Theta(t,\nu) \) and \( R(t,\nu) \) are respectively the polarization parameters \( \theta \) and \( \rho \) of both waves. On the rest of those images, which is the borders and the center of the image, values oscillate because of the uncorrelation from one component to the other. Thus the polarization parameters of each wave can be estimated and at the output, we get figure 8(c), (d), (g) and (h). The noise that remain in the estimations is due to the white noise.

![Figure 8](image-url)

**Figure 8.** Representation of the sum of two chirps on two components.

In this section, we introduced new polarization representations which give complementary information to module images. We presented how those tools can be inserted in the application of the oblique polarization filter which performs an efficient separation of two waves. Compared with the approach of section 2 which operates on each component separately, we improve the separation performances since we are able to separate two waves overlapping partially in the time-frequency plane. Necessary assumptions for the efficiency of the algorithm are that polarization is constant in the time-frequency plane (which is usually checked) and that for both waves, it exists an area where the wave is paramount compared to other waves and noise. In the next section, we present, in case of scalar sensors, another alternative to separate crossing waves for array data.
4. DISPERSION ESTIMATION FROM A LINEAR ARRAY

4.1. Context of the Study

In this section, we consider to have a seismic profile of several traces which contains several waves which cross each other. In order to improve the performances of the algorithm of section 2, we need to use the redundancy of information among the different traces like it is done by the two dimensional Fourier transform (2DFT). In addition, in order to avoid the 2DFT filtering problems such as variability of the velocity along the sensor and aliasing, which results in the introduction of artifacts, we decide to compensate the propagation of one of the wave such as it is lined up in the profile, and then sum all the traces to estimate the corresponding wavelet. Remember that we focus on the extraction of the surface wave, the estimation of its propagation is a difficult task because this wave is dispersive: this implies that the different harmonics of the wave propagate at different velocities. Therefore, the filter which models its propagation is not a simple delay. We present the two classical methods of dispersion estimation and then present our new method which takes advantage of both classical ones.

The first set of classical methods is based on the analysis of each trace in the time-frequency plane. The transfer function between the source and the sensor is noted $H$. We assume there is no attenuation during the propagation, so the module of the Fourier transform ($FT$) of $h$ (noted $H(\nu)$) is equal to 1 ($H(\nu) = e^{i\phi_H(\nu)}$, where $\phi_H(\nu)$ is the phase of $H(\nu)$). If at the source, all the harmonics of the wave are localized at the same instant $t_0$, the transfer function $H(\nu)$ can be written as:

$$H(\nu) = e^{i(-2\pi \int_{t_0}^{t_0} (\tau_g(f) - t_0) df + \phi_H(0))},$$

where $\tau_g(f)$ is the group delay of the wave.

In equation (7), the unknown function to be estimated is $\tau_g(f)$. Different methods are available to estimate it depending on the time-frequency representation used. The delay $\tau_g(f) - t_0$ at each frequency is estimated as the duration between the origin $t_0$ and the ridge of the pattern of the wave on the time-frequency plane. The major issue of this 1D technique is that in presence of several waves with close time-frequency location, there may be some overlapping phenomena between the patterns of the different waves, so that the ridges of the different components may be impossible to characterize.

The second set of classical methods is based on the presentation of the whole profile in a frequency-velocity image. We consider a synthetic dispersive wave presented figure 9(a). The dispersion between sensor $n$ and sensor $n+1$ is characterized by the transfer function noted $h_n$. Noting $w_n$ the dispersive wave at sensor $n$, we have $w_{n+1} = w_n * h_n$, where $*$ stands for the convolution operator. The dispersion is assumed to be spatially stationary along the $N$ sensors (i.e. $h_n = h$), and the module $|H(\nu)|$ to be equal to one. In presence of one dispersive wave, the trace number $n$ is equal to the wave convoluted $n$ times to $h$ (i.e. $p(n, t) = FT_{-1}^{-1}|W(f)e^{i\phi_H(f)}|(t)$, where $W(f)$ the temporal Fourier transform ($FT_t$) of $w(t)$). So the 2DFT of $p$ leads to:

$$P(\nu_1, \nu_2) = W(\nu_2)\delta(\nu_2 - \frac{\phi_H(\nu_1)}{2\pi}).$$

So to estimate $\phi_H(\nu_1)$, we can follow the ridge of the pattern on the $f - k$ spectrum (see figure 9(d)). This kind of method can be described in two steps. First, it analyzes the profile at the frequency $\nu_1$ by computing the temporal Fourier transform at $\nu_1$. This leads to the profile presented in figure 9(b) in time-space (real part) and 9(e) in $f - k$ space. The second step is to find the correction which lines up the sinusoids in figure 9(b), which is equivalent in $f-k$ representation to move the point on the spatial frequency axes (vertical) on figure 9(e) (see the arrow). To choose the best correction, the criteria is the maximization of the energy of the summation of all traces. It leads to a frequency-velocity representation. Compared to the previous method, we gain the velocity resolution but we lose the time resolution.

Then, our purpose is to introduce the time resolution in the frequency-velocity analysis of the profile. So, to make a tradeoff between time and frequency resolution, to analyze the profile at $\nu_1$, instead of convolving the profile with the complex exponential $e^{-2\pi j \omega t_1}$, we convolve it with a short time duration sine wave localized around $\nu_1$ noted $\psi_{\nu_1}$. The increase of the time resolution, and its consequence: the decrease of the frequency resolution, are shown by comparing respectively figure 9(b) and 9(c), and figure 9(e) and 9(f).
Considering that the bandwidth of the wavelet is small enough compared to the derivative of the curvature of the ridge on figure 9(d), we can assume that the pattern on figure 9(e) is a segment. This means that the propagation of the filtered profile $p_{\psi}$ can be approximated by a delay followed by a phase shift (see both arrows on figure 9(c)). So to compensate the dispersion, we need to find the good correction (delay and phase shift) and in order to get also the time resolution: the time arrival of the wave on the last sensor, for each time arrival on the last sensor we look for the maximum of amplitude of the summation of all traces. This ends with an amplitude function of time, where at the maxima are located the different waves, and the corresponding delay and phase shift arguments leads to $\phi_H(\nu_t)$.

Applying a similar procedure for each frequency $\nu_t$ means that we apply a bunch of bandpass filters $\Psi_{\nu_t}(f)$ on the profile and then for each one, look for the correction which maximizes the summation of all traces. This may take time. However, because the filtering operation is linear, we can apply the bunch of filters after the correction and the summation of the traces, and thus estimate the best correction for all frequency simultaneously. This means that for each delay and phase shift correction, we apply a linear time-frequency representation (TFR) to the sum signal to visualize where in the time-frequency plane the summation magnifies the amplitude. This leads to a stack of time-frequency representations. This volume can be interpreted as a time-frequency-velocity volume. Then, in order to visualize it, we project this 3D information on the time-frequency plane. To do this, for each (time, frequency) location, we look for the maximum of amplitude among all the time-frequency images. This is illustrated on figure 10, where on the right we get the amplitude image and both corresponding argument correction images: delay and phase shift. Using those two argument we get for each time-frequency location an estimator of $\phi_H$. The best estimator is located on the ridge of the amplitude image.

### 4.2. Results and Discussion

To illustrate the different methods, we consider a synthetic profile with four propagating waves (see figure 11(a)). When the number of sensors is fifty, the f-k image (see figure 11(b)) is still useful, even if the pattern of different waves overlap on the image. However, when the number of sensors fall to ten (we kept traces from 31 to 40), those frequency-velocity representations are no more understandable (see figure 11(c) and (d)). This is because the spatial frequency resolution which is inversely proportional to the sensor number ($\Delta\nu_n = \frac{1}{N}$) is bad. This is a major problem in practice, since holding the spatial stationarity condition may require to consider only a few traces from the whole profile.17
The other classical method which is considering the CWT of each trace would not be efficient neither because the patterns of dispersive and non dispersive waves overlap in the time-frequency plane. Then, when using our method, we use the information a priori that the dispersive waves that we are aiming to visualize have a positive velocity, so only positive delay corrections $\tau_k$ have been applied. Thus, when estimating the maximum of amplitude among all corrections, only the dispersive waves are present: the velocity information enabled to separate the dispersive waves from the non dispersive waves. In addition, thanks to the time resolution, on the resulting figure 11(e), both dispersive waves are separated, they do not overlap in the time-frequency plane. Thus our method enables to estimate the dispersion of each dispersive wave. In order to check the estimation result for one synthetic wave, we plot on figure the $f - k$ spectrum before and after the correction of the dispersion on the wave alone. The image (plot figure 11(g) and (h)) shows that the dispersion has been perfectly compensated.
5. CONCLUSION

In this paper, we showed that the continuous wavelet transform is a performing tool to describe the content of seismic data. Adapting this tool to the nature of the data, we proposed different robust and semi-automatic methods in order to separate the wave. Those method illustrated on synthetic data have also been validated on real data in\textsuperscript{18} and\textsuperscript{19}. Of course in a more operational context, it is clear that some techniques will be necessary in order to adapt the time calculation to the amount of the data. In addition, in order to assess the efficiency of the presented methods, it will be necessary to apply them on other data.

In the literature, many publications tried to introduce the best amplitude image to describe the data. In this study, we introduce new time-frequency representation which describe other information: polarization and velocity. As a perspective is would be interesting to design fusion algorithm in order link those different kinds of information.

REFERENCES