The convolution theorem for the continuous wavelet transform

Antonio F. Pérez-Rendón, Rafael Robles
Dpto. Applied Mathematics I, University of Seville, Spain

Abstract

We study the application of the continuous wavelet transform to perform signal filtering processes. We first show that the convolution and correlation of two wavelet functions satisfy the required admissibility and regularity conditions. By using these new wavelet functions to analyze both convolutions and correlations, respectively, we derive convolution and correlation theorems for the continuous wavelet transform and show them to be similar to that of other joint spatial/spatial–frequency or time/frequency representations. We then investigate the effect of multiplying the continuous wavelet transform of a given signal by a related transfer function and show how to perform spatially variant filtering operations in the wavelet domain. Finally, we present numerical examples showing the usefulness of applying the convolution theorem for the continuous wavelet transform to perform signal restoration in the presence of additive noise.

Keywords: Continuous wavelet transform; Convolution; Correlation; Space varying filtering; Signal restoration

1. Introduction

The received signal in a large number of signal processing systems is modeled as a convolution between the input signal and the impulse response of the system with additive noise. Signal restoration (or deconvolution) problems have been mainly studied under the assumption that the input signal and noise are stationary (though the terms stationary and non-stationary have some precise meanings in stochastic processes, in this work we use these terms in a qualitative sense, i.e., a stationary signal is a signal whose properties do not change much with time or through space), and conventional estimation algorithms such as Wiener or Kalman filters can be directly applied. However, the problem of restoration of non-stationary signals such as seismic data, acoustic signals, mechanical vibration, wireless communication, radar and sonar signals, etc., has recently attracted the attention of signal processing researchers. Properties of joint time–frequency representations have allowed to find optimal deconvolution filters for non-stationary signal restoration [4,14].

The wavelet transform is of interest for the analysis of non-stationary signals because it provides an alternative to classical linear time–frequency representations with better time and frequency localization properties [32]. On the other hand, the role of the wavelet transformation is to remove the non-stationary properties of the involved signals; consequently, the conventional estimation algorithms for stationary signal processing can be employed in each scale of the wavelet domain [5]. For these reasons, several
Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{f}(w)$</td>
<td>Fourier transform of function $f(x)$</td>
</tr>
<tr>
<td>$E_f(x, w; \phi)$</td>
<td>Generalized joint distribution of function $f$ with kernel $\phi$</td>
</tr>
<tr>
<td>$f \ast (a, b)$</td>
<td>Wavelet transform of function $f$ with wavelet $\psi$</td>
</tr>
<tr>
<td>$L^2(\mathbb{R})$</td>
<td>Set of all measurable functions that are square integrable</td>
</tr>
<tr>
<td>$f^*(x)$</td>
<td>Complex conjugate of function $f(x)$</td>
</tr>
<tr>
<td>$|\cdot|_p$</td>
<td>$L^p$ norm defined in Lebesgue space $L^p(\mathbb{R})$</td>
</tr>
<tr>
<td>$(f \otimes g)(x)$</td>
<td>Convolution of functions $f(x)$ and $g(x)$</td>
</tr>
<tr>
<td>$(f \ast g)(x)$</td>
<td>Correlation of functions $f(x)$ and $g(x)$</td>
</tr>
<tr>
<td>$(f \otimes_1 g)(x, y)$</td>
<td>Convolution in the first variable of functions $f$ and $g$</td>
</tr>
<tr>
<td>$(f \otimes_2 g)(x)$</td>
<td>Convolution in the second variable of functions $f$ and $g$</td>
</tr>
<tr>
<td>$(f \ast_1 g)(x, y)$</td>
<td>Correlation in the first variable of functions $f$ and $g$</td>
</tr>
<tr>
<td>$(f \ast_2 g)(x)$</td>
<td>Correlation in the second variable of functions $f$ and $g$</td>
</tr>
<tr>
<td>$(f \otimes \otimes g)(x)$</td>
<td>Generalized convolution of functions $f(x)$ and $g(x)$</td>
</tr>
<tr>
<td>$M^\psi_q$</td>
<td>$q$-order moment of wavelet $\psi(x)$</td>
</tr>
<tr>
<td>$\mathcal{F}_2{\tilde{f}^\psi}(a, w)$</td>
<td>Fourier transform of $\tilde{f}^\psi(a, b)$ in the second variable $(b \leftrightarrow w)$</td>
</tr>
</tbody>
</table>

Algorithms have been developed in order to perform signal restoration in the wavelet domain [1,5,17,23,40,43], but the lack of a convolution theorem for the continuous or discrete wavelet transform prevents us from finding optimal deconvolution filters in the wavelet domain. Therefore, we think that the establishing of such convolution theorems, besides their theoretical importance from a mathematical point of view, have interesting applications in signal processing problems such as signal restoration.

Filtering for signals and images is at present extensively used in different areas of signal and image processing such as image enhancement, image restoration, signal coding and image analysis [3]. Traditionally, most optical data processing systems only perform space invariant processing operations, mathematically described in terms of convolutions, which are easily implemented through the Fourier transform (FT). The convolution theorem for this one states that the FT of the convolution of two signals is the product of their respective FT, that is

$$g(x) = \int_{-\infty}^{+\infty} dt \ f(t) \ h(x - t) \Rightarrow$$

$$\tilde{g}(w) = \tilde{f}(w) \ \tilde{h}(w). \quad (1)$$

Therefore, the filtering operation is generally reduced to multiply the FT of the signal by the transfer function of the filter, and to make the inverse Fourier transformation of the product, which provides a filtered version of the signal.

The FT performs a global analysis of a given signal or image, and it is especially suited for stationary signals. Nevertheless, in many cases of practical interest the signal under investigation is non-stationary and has a different frequency content at different instants or localizations. One problem with the application of the FT to such non-stationary signals is that high-frequency terms appear and they are not localized, but are added everywhere.

On the other hand, signals can also be affected by non-stationary processes; for instance, optical defocus may introduce a spatially variant blur in the signal. Space or time-variant filtering [39] is mathematically described as a generalized convolution or a Fredholm integral equation of the first kind:

$$g(x) = \int_{-\infty}^{+\infty} dt \ f(t) \ h(x, t). \quad (2)$$

The convolution theorem no longer applies for these space or time-varying filters and thus the Fourier domain is less useful for their application [22].
Therefore, we come to the conclusion that spatially variant signals and processes can be better characterized by joint time–frequency or spatial/spatial-frequency representations [8,13,28].

Wigner [41] introduced a bilinear distribution as a joint representation of the phase space in Quantum Mechanics. Later, Ville [37] derived the same distribution, nowadays called the Wigner–Ville distribution, in the field of signal processing. Another way to obtain a joint representation is through the complex spectrogram, that can be expressed as a windowed FT, first introduced by Gabor [10]. Both Wigner–Ville and Gabor distributions belong to the Cohen class of bilinear distributions [7], in which each member is obtained by introducing a particular kernel \( \phi \) in the generalized distribution defined as

\[
E_f(x, w; \phi) = \int \int \int_{\mathbb{R}^3} \d\rho \, \d x \, \d \xi \, \phi(\xi, x) \times f(\rho + x/2) \, f^*(\rho - x/2) \times \exp\{2\pi i(\xi x - w\alpha - \xi\rho)\}. \tag{3}
\]

There exist convolution theorems for the Wigner–Ville distribution [6] as well as for other joint representations belonging to Cohen class [26]. Such theorems state that the joint representation of the convolution of two signals is the convolution in the time or space variable, at every fixed frequency, of the respective joint representations, that is

\[
E_{\hat{g}}(x, w; \phi) = \int_{-\infty}^{+\infty} dt \, E_f(t, w; \phi) \, E_h(x - t, w; \phi). \tag{4}
\]

Joint representations also allow us to perform space variant filterings on a given function by multiplying its joint representation by a function \( T(x, w) \), which can be regarded as a space-variant transfer function, and taking the inverse transformation. This process has been carried out with the complex spectrogram [34] and the Wigner–Ville distribution [2,11,34].

The continuous wavelet transform (CWT) [12,15] stands for a scale-space representation and is an effective way to analyze non-stationary signals and images. Typical applications are the detection and characterization of singularities [25], fractal analysis [21,38], pattern recognition [20], noise reduction [18], image processing [27], geophysics [9], aeromagnetic processing [31] and the analysis of biomedical [42] and meteorological [35] images. The existence of fast algorithms for the calculus of the CWT [24,36] guarantees its use for solving real problems.

The CWT is a scale-space representation of a given function which may be regarded as a joint representation by identifying the scale with a frequency ratio [15]. This fact introduces an adaptability in the analysis which joint representations lack. Nevertheless, there is no convolution theorem for the CWT, and this limits its usability in image and signal processing.

We have found in the literature some works relative to convolution related filtering and CWT, which are summarized below.

In [19], given a function \( g(x) \) as the convolution of another function \( f(x) \) and a filter \( h(x) \), let \( \hat{g}(a, b) \) and \( \hat{f}(a, b) \) denote the CWT with the same wavelet of \( g(x) \) and \( f(x) \), respectively. Then, it is shown that

\[
\hat{g}(a, b) = \int_{-\infty}^{+\infty} dt \, \hat{f}(a, t) \, h(b - t). \tag{5}
\]

This means that the CWT of the convolution between some function and a filter is the convolution in the time or space variable, at every fixed scale, of the CWT of the function and the very filter.

In [33], given the convolution \( g(x) \) as above, let \( \hat{f}(a, b) \) denote the CWT of the function \( f(x) \) with wavelet \( \psi(x) \) and let \( \hat{h}(a, b) \) denote the CWT of the filter \( h(x) \) with wavelet \( \chi(x) = \psi(-x) \). Then, it is shown that

\[
g(x) = \frac{1}{C} \int_{0}^{+\infty} \frac{da}{a^2} \left( \int_{-\infty}^{+\infty} \d\xi \, \hat{f}(a, \xi) \hat{h}(a, x - \xi) \right), \tag{6}
\]

where the normalization constant \( C_{\psi} \) is defined in (21).

That is, the convolution (instead of its CWT) between some function and a filter is the sum over all scales of the convolution in the time or space variable, at every fixed scale, of their respective CWT, but with different wavelets.

A different approach is given in [16], where wavelet-based filters, known as scale filters, are generated as follows. Let \( \hat{f} \) denote the CWT of a given function with wavelet \( \psi \) and let \( \hat{h}(w) \) denote the FT of the filter \( h(x) \); then, a scale filter \( v(a) \) is
defined as
\[ \tilde{h}(w) = \sum_a v(a) \hat{\psi}^*(aw) \]  
(7)
and is calculated via linear squares methods. Then, it is shown that
\[ g(x) = \sum_a v(a) \hat{f}(a,x). \]  
(8)

A similar process has been used to construct space varying scale filters \(^{29,30}\).

Let us remark that neither result can be regarded as a convolution theorem, because this must formulate the convolution equation in the wavelet domain entirely, i.e., it must show a relationship between the CWT of two functions and their convolution. Nevertheless, these works point out the existence of such a convolution theorem similar to that for joint representations as shown in (4) and the possibility to use different wavelets for every involved function. This possibility is deeply investigated in Section 3.

This paper is organized as follows. In Section 2 we define some known topics about Fourier and continuous wavelet transforms and introduce the notation we will use in following sections. In Section 3 we derive the convolution and correlation theorems for the continuous wavelet transform and show the effect of multiplying this by some transfer function. In Section 4 we initially show the validity of the theorems proved in previous section and then apply them to a computer simulation of a signal restoration problem; comparisons with Fourier based deconvolution are provided showing that the signal restoration method proposed in this paper yields better restorations. Finally, we summarize the main conclusions of this work in Section 5.

\[ \|f\|_p = \left( \int_{\mathbb{R}^n} |\hat{f}(\vec{a})|^p \right)^{1/p} < \infty. \]  
(9)

It is well known that \( L^p(\mathbb{R}) \) is a Banach space with norm \( \| \cdot \|_p \) and \( L^2(\mathbb{R}) \) is a Hilbert space with inner product \( \langle f, g \rangle = \int_\mathbb{R} \hat{f}(x) g^*(x) \), where \( (\cdot)^* \) denotes complex conjugate.

The FT of a function \( f \in L^2(\mathbb{R}) \) is defined as
\[ \tilde{f}(w) = \int_{-\infty}^{\infty} f(x) \exp\{-2\pi iwx\} \]  
(10)
which also belongs to \( L^2(\mathbb{R}) \). The reconstruction formula is given by
\[ f(x) = \int_{-\infty}^{\infty} dw \hat{f}(w) \exp\{2\pi iwx\}. \]  
(11)

If \( f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \), then \( \tilde{f}(w) \) is a bounded function in \( L^2(\mathbb{R}) \), i.e., there exists a positive real number \( M \) such that \( |\tilde{f}(w)| \leq M \) for all \( w \in \mathbb{R} \).

If \( f \) and \( h \) are complex-valued functions defined on \( \mathbb{R} \), their convolution \( f \otimes h \) is defined as a new function:
\[ g(x) = (f \otimes h)(x) = \int_{-\infty}^{\infty} dt f(t) h(x-t) \]  
(12)
provided that the integral exists. If \( f \in L^2(\mathbb{R}) \) and \( h \in L^1(\mathbb{R}) \), then such an integral exists almost everywhere and \( f \otimes h \in L^2(\mathbb{R}) \) with \( \| f \otimes h \|_2 = \| f \|_2 \cdot \| h \|_1 \).

If \( f \) and \( h \) are complex-valued functions defined on \( \mathbb{R}^2 \), it is possible to convolve them in just one variable as follows:
\[ \begin{align*}
g_1(x, y) &= (f \otimes_1 h)(x, y) \\
&= \int_{-\infty}^{\infty} dt f(t, y) h(x-t, y), \quad (13) \\
g_2(x, y) &= (f \otimes_2 h)(x, y) \\
&= \int_{-\infty}^{\infty} dt f(x, t) h(x, y-t), \quad (14) 
\end{align*} \]
provided that both integrals exist. If \( f \in L^2(\mathbb{R}^2) \) and \( h \in L^1(\mathbb{R}^2) \), then such integrals exist almost everywhere and \( g_1, g_2 \in L^2(\mathbb{R}^2) \).

An operator quite related to the convolution is the so-called correlation operator. If \( f \) and \( h \) are complex-valued functions defined on \( \mathbb{R} \), their correlation \( f \ast h \) is defined as a new function
\[ g(x) = (f \ast h)(x) = \int_{-\infty}^{\infty} dt f(t) h(t+x), \]  
(15)
provided that the integral exists. If \( f \in L^2(\mathbb{R}) \) and \( h \in L^1(\mathbb{R}) \), then such an integral exists almost everywhere and \( f \ast h \in L^2(\mathbb{R}) \) with \( \| f \ast h \|_2 = \| f \|_2 \cdot \| h \|_1 \).
When \( f \equiv h \), \( f \ast f \) is known as the autocorrelation of \( f \), while otherwise \( f \ast h \) is known as the cross-correlation between \( f \) and \( h \).

If \( f \) and \( h \) are complex-valued functions defined on \( \mathbb{R}^2 \), it is possible to correlate them in just one variable as follows:

\[
g_1(x, y) = (f \ast h)(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} dt f(t, y) h(t + x, y), \tag{16}
\]

\[
g_2(x, y) = (f \ast h)(x, y) = \int_{-\infty}^{+\infty} dt f(t, x) h(t + y), \tag{17}
\]

provided that both integrals exist. If \( f \in L^2(\mathbb{R}^2) \) and \( h \in L^1(\mathbb{R}^2) \), then such integrals exist almost everywhere and \( g_1, g_2 \in L^2(\mathbb{R}^2) \).

Given two functions \( f \in L^2(\mathbb{R}) \) and \( h \in L^2(\mathbb{R}^2) \), their generalized convolution is defined as the Fredholm integral equation of the first kind:

\[
g(x) = (f \otimes h)(x) = \int_{-\infty}^{+\infty} dt f(t) h(x, t), \tag{18}
\]

which is also known as a space-variant filtering in optics and engineering literature \([2, 11]\), as opposed to the space invariant filtering given by the convolution defined in (12).

Given a fixed function \( \psi \in L^2(\mathbb{R}) \), called wavelet, consider its translations and dilations defined as

\[
\psi_{ab}(x) = \frac{1}{|a|^{1/2}} \left( \frac{x - b}{a} \right) \tag{19}
\]

with \( a, b \in \mathbb{R} \) and \( a \neq 0 \). These functions are scaled so that their \( L^2(\mathbb{R}) \) norms are independent of \( a \). The CWT of a function \( f \in L^2(\mathbb{R}) \) with wavelet \( \psi \) is now defined as

\[
\hat{f}(a, b) = \int_{-\infty}^{+\infty} dx f(x) \psi^*_a(x) = \frac{1}{|a|^{1/2}} \int_{-\infty}^{+\infty} dx f(x) \psi^* \left( \frac{x - b}{a} \right). \tag{20}
\]

A wavelet function \( \psi(x) \) is said to be admissible if

\[
0 < C_\psi = \int_{-\infty}^{+\infty} dw \frac{|\hat{\psi}(w)|^2}{|w|} < \infty. \tag{21}
\]

If this is the case, the CWT \( \hat{f}(a, b) \) is invertible on its range, and an inverse transform is given by the relation

\[
f(x) = \frac{1}{C} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{da db}{a^2} \hat{f}(a, b) \psi_{ab}(x), \tag{22}
\]

where the integral over \( a \) is understood to exclude \( a = 0 \) (integrate over \( 0 < a \leq |a| < \infty \), then take the limit \( \epsilon \to 0 \)).

The most important properties of wavelets are admissibility (21) and regularity conditions. As can be seen in (20), the wavelet transform of a one-dimensional function is two-dimensional and for most practical applications this is not a desirable property. Therefore, one imposes some additional conditions on the wavelet functions in order to make the wavelet transform decrease quickly with scale. These are the regularity conditions and they state that the wavelet function should have some smoothness and concentration in both time and frequency domains. Regularity is a quite complex topic which is related to the concept of vanishing moments; the number of vanishing moments required depends heavily on the application. The moments of a wavelet \( \psi \in L^2(\mathbb{R}) \) are defined as

\[
M_\psi^\alpha = \int_{-\infty}^{+\infty} dt t^\alpha \psi(t). \tag{23}
\]

Let \( \mathcal{L} \) be the Hilbert space of all measurable functions \( H(a, b) \) that are square integrable with respect to the weight function \( |a|^{-2} \), and let \( \mathcal{F} \) be the space of all wavelet transforms with respect to a fixed wavelet function \( \psi(x) \). It has been shown [15] that \( \mathcal{F} \) is a proper subspace of \( \mathcal{L} \); that is, not every function \( H(a, b) \in \mathcal{L} \) is the CWT of some function \( f \in L^2(\mathbb{R}) \). If we apply the reconstruction formula (22) to such a function \( H(a, b) \):

\[
h(x) = \frac{1}{C} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{da db}{a^2} H(a, b) \psi_{ab}(x), \tag{24}
\]

then the function \( h(x) \) belongs to \( L^2(\mathbb{R}) \) and it is the unique function such that for any other function \( f \in L^2(\mathbb{R}) \) it is \( \|H(a, b) - \hat{f}(a, b)\|_\psi > \|H(a, b) - \hat{f}(a, b)\|_\psi \). The function \( h(x) \) is known as the least-squares approximation in \( L^2(\mathbb{R}) \) to the inverse transform of \( H(a, b) \).

It is important to note that in (20) and (22) the wavelet functions are not specified. This is a difference between the CWT and the FT, or other transforms. The
theory of wavelet transforms deals with the general properties of the wavelets and the wavelet transforms only. The choice of the wavelet function depends on the particular application. For space–frequency analysis, we need a wavelet that is optimally localized in terms of both spatial width and frequency bandwidth. For smooth signals, we generally want a wavelet that is itself smooth and therefore has good frequency localization. In contrast, signals that contain discontinuities are better analyzed using wavelets with good spatial localization to accurately map rapid changes in the signal. There are several other useful properties that the wavelet function can possess, as its orthogonality, complexity, width or shape.

3. The convolution theorem

In [33] it is pointed out the possibility to use different wavelet functions for the signal and the filter. Assuming that we have selected the best wavelets for both, what is then the best choice for the convolution or correlation? It is somewhat natural to consider the convolution or correlation of the respective wavelets provided to satisfy the admissibility condition (21). Calculate now the moment \( \rho \) of wavelet \( \psi_l \) and \( \psi_k \) satisfy the admissibility condition (21). Calculate now the moment \( \rho \) of wavelet \( \psi_l \), defined as

\[
M^3_{\rho} = \int_{-\infty}^{+\infty} dt \, t^\rho \, \psi_l(t).
\]

By inserting (12) for \( \psi(t) \) into above equation and changing the order of integration, we find

\[
M^3_{\rho} = \int_{-\infty}^{+\infty} dx \, \psi_l(x) \int_{-\infty}^{+\infty} dt \, t^\rho \, \psi_l(t - x).
\]

If we perform the change of variables \( y = t - x \) into the inner integral and expand \( (x + y)^\rho \) according to Newton binomial expansion, we arrive at

\[
M^3_{\rho} = \sum_{k=0}^{\rho} \binom{\rho}{k} \int_{-\infty}^{+\infty} dx \, x^k \, \psi_l(x)
\]

\[
\times \int_{-\infty}^{+\infty} dy \, y^{\rho-k} \, \psi_l(y),
\]

that is

\[
M^3_{\rho} = \sum_{k=0}^{\rho} \binom{\rho}{k} \, M^1_{\rho-k} \, M^2_{\rho-k}.
\]

An analogous reasoning provides

\[
M^4_{\rho} = \sum_{k=0}^{\rho} (-1)^k \binom{\rho}{k} \, M^1_{\rho-k} \, M^2_{\rho-k}.
\]

Let be \( \rho \leq N_1 + N_2 \). If \( k \leq N_1 \), then \( M^4_{\rho-k} = 0 \) because \( \psi_l \) has \( N_1 \) vanishing moments; otherwise, \( \rho - k \leq N_1 + N_2 - N_1 = N_2 \), so that \( M^4_{\rho-k} = 0 \) because \( \psi_k \) has \( N_2 \) vanishing moments. Therefore, all addends

\[ |\tilde{\psi}_3(w)| = |\tilde{\psi}_4(w)| = |\tilde{\psi}_1(w)| |\tilde{\psi}_2(w)|. \] (28)

Therefore, we get

\[
0 < C_3 = C_4 = \int_{-\infty}^{+\infty} dw \, \frac{\tilde{r}_1(w)^2 |\tilde{\psi}_2(w)|^2}{|w|}.
\]

Since \( \psi_k \in L^1(\mathbb{R}) \), its FT \( \tilde{\psi}_2(w) \) is a bounded function, so that there exists a positive real number \( M \) such that \( |\tilde{\psi}_2(w)| \leq M \) for all \( w \in \mathbb{R} \). Then, we arrive at:

\[
0 < C_3 = C_4 \leq M^2 \int_{-\infty}^{+\infty} dw \, \frac{\tilde{r}_1(w)^2}{|w|} = M^2 C_1 < \infty.
\]

Therefore, the new defined wavelets \( \psi_l \) and \( \psi_k \) satisfy the admissibility condition (21). Calculate now the moment \( \rho \) of wavelet \( \psi_l \), defined as

\[ C_i = \int_{-\infty}^{+\infty} dw \, \frac{\tilde{r}_i(w)^2}{|w|} \]

for \( i = 1,\ldots,4 \). Since \( \psi_l(x) \) and \( \psi_k(x) \) are admissible wavelets, then according to (21) we know that \( 0 < C_1, C_2 < \infty \). Taking into account some properties of the FT, we have

\[ \tilde{\psi}_3(w) = \tilde{\psi}_1(w) \tilde{\psi}_2(w), \]

\[ \tilde{\psi}_4(w) = \tilde{\psi}_1(w) \tilde{\psi}_2^*(w), \]

from which we can find

\[ |\tilde{\psi}_3(w)| = |\tilde{\psi}_4(w)| = |\tilde{\psi}_1(w)| |\tilde{\psi}_2(w)|. \] (28)
The correlation of two signals is an operation quite related to the convolution. In fact, the correlation theorem for the FT is quite similar to the convolution theorem; the only difference is the complex conjugate that appears in the FT of the filter, that is, \( \hat{g}(w) = \hat{f}(w) \hat{h}^* (w) \). Therefore, it is hoped that the CWT satisfies a correlation theorem similar to the previous convolution theorem.

**Theorem 3** (Wavelet correlation theorem). Let \( f \in L^2(\mathbb{R}) \) and \( \psi_b \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \) be two admissible wavelets, and let \( \hat{f}_{\psi_b} \) and \( \hat{h}_{\psi_b} \) denote the CWT of two functions \( f \in L^2(\mathbb{R}) \) and \( h \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \) with wavelets \( \psi_f \) and \( \psi_h \), respectively. If \( g = (f \ast h) \) and \( \psi_g = (\psi_f \otimes \psi_h) \), then

\[
\hat{g}_{\psi_b}(a,b) = \frac{1}{|a|^{1/2}} (\hat{f} \ast_2 \hat{h}_{\psi_b}) (a,-b). 
\]  

**Proof.** The CWT of \( g(x) \), given by (20), may be rewritten as

\[
\hat{g}_{\psi_b}(a,b) = \frac{1}{|a|^{1/2}} \int \int_{\mathbb{R}^3} \, dx \, dt \, dy \, f(t) \, h(x-t)
\times \psi_f^*(y) \psi_b^* \left( \frac{x-b}{a} - y \right).
\]  

Performing the change of variables \( x = t + x \) and \( \beta = b + ay - t \), (37) may be rewritten as

\[
\hat{g}_{\psi_b}(a,b) = \frac{1}{|a|^{3/2}} \int_{-\infty}^{+\infty} \, d\beta \int_{-\infty}^{+\infty} \, dt \, f(t)
\times \psi_f^* \left( \frac{t - (\beta - b)}{a} \right)
\times \int_{-\infty}^{+\infty} \, dx \, h(x) \psi_b^* \left( \frac{x-\beta}{a} \right),
\]  

from which it is easy to find

\[
\hat{g}_{\psi_b}(a,b) = \frac{1}{|a|^{1/2}} \int_{-\infty}^{+\infty} \, d\beta \hat{f} \hat{h}^* (a,\beta)
= \frac{1}{|a|^{1/2}} \hat{f} \ast_2 \hat{h}_{\psi_b}(a,b).
\]  

The correlation theorem has been analyzed with two different wavelet functions and their convolution has been analyzed with the convolution of the chosen wavelets, what is the relation between the three obtained CWT? As it was pointed out in Section 1, such a relationship is quite similar to the convolution theorem for joint representations shown in (4).
According to above theorems, the CWT allows us to perform any linear filtering. Although this transform is far too redundant to yield fast algorithms, the main feature of Theorems 2 and 3 is that the convolution and correlation operators apply at every scale independently of each other, and this fact allows us to deeply understand the effect of the filter on every scale of the given signal. This property may be very useful in deconvolution related problems, where it will be possible to design a different algorithm to every scale, depending on the properties of the filter at such a scale and the condition number of the corresponding inverse problem.

Despite of previous considerations, the main drawback of Theorems 2 and 3 is that the convolution and correlation operators do not disappear in the wavelet domain, and this implies no time reduction in the application of such operators with respect to the spatial domain. This is not the case with the FT. The convolution and correlation theorems for the FT are powerful tools to perform space invariant filtering processes because they reduce these processes to simply multiply the FT of the signal by the transfer function associated with the filter.

It would be desirable to perform some kind of filtering by simply multiplying the CWT by a transfer function; for joint representations this product performs a spatially variant filtering on the given signal. The similarity between the convolution theorems for joint representations and CWT leads to think about the possibility to perform such space varying filtering operations in the wavelet domain.

Let \( \hat{f} \) denote the CWT of a function \( f \in L^2(\mathbb{R}) \) with an admissible wavelet \( \psi \in L^2(\mathbb{R}) \) and let \( H \in L^2(\mathbb{R}^2) \). Define \( G(a,b) = \hat{f} \hat{\psi}(a,b)H(a,b) \), which may not be the CWT of any function in \( L^2(\mathbb{R}) \); the least square approximation to such a function is obtained by applying the reconstruction formula (22) to \( G(a,b) \). Therefore,

\[
g(x) = \frac{1}{C} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{da \, db}{a^2} \hat{f}(a,b) \\
\times H(a,b) \psi \left( \frac{x - b}{a} \right)
\]

(44)

with \( C_\psi \) defined in (21). By inserting the definition of \( \hat{f} \hat{\psi}(a,b) \) according to (20) into above equation and changing the order of integration, we find

\[
g(x) = \frac{1}{C} \int_{-\infty}^{+\infty} dt \, f(t) \\
\times \left( \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{da \, db}{|a|^{5/2}} H(a,b) \\
\times \psi^* \left( \frac{t - b}{a} \right) \left( \frac{x - b}{a} \right) \right).
\]

(45)

If we define

\[
h(x,t) = \frac{1}{C} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{da \, db}{|a|^{5/2}} H(a,b) \\
\times \psi^* \left( \frac{t - b}{a} \right) \left( \frac{x - b}{a} \right)
\]

(46)

provided the above integral to converge, then \( g(x) = (f \otimes \otimes h)(x) \) is the function in \( L^2(\mathbb{R}) \) whose CWT with wavelet \( \psi \) is as close as possible to the product \( \hat{f}(a,b)H(a,b) \).

Although we have not found an inverse formula for obtaining the transfer function \( H(a,b) \) which corresponds to a given kernel \( h(x,t) \), this result raises an interesting question. Is it possible to find useful filters such that their Fourier transfer functions cannot be well approximated by rational functions but their action is diagonal in the wavelet domain? We think that this question might be of interest for the signal processing community, so that future work may relay on it.

4. Numerical examples

In this section, two numerical examples are given to illustrate (i) the validity of the theorems proved in previous section and (ii) the usability of such theorems in signal restoration problems in the presence of additive noise. In both examples, the test signals have been analyzed with the first and second derivatives of the well–known normal distribution with zero mean and unit variance, given by

\[
\psi_1(x) = \sqrt{\frac{2}{\sqrt{\pi}}} x \exp \left\{ -\frac{x^2}{2} \right\}
\]

(47)

and

\[
\psi_2(x) = \frac{2}{\sqrt{3\sqrt{\pi}}} (1 - x^2) \exp \left\{ -\frac{x^2}{2} \right\}
\]

(48)
respectively. It can be shown that both wavelets satisfy the admissibility condition (21) with coefficients $C_1 = 44.5466$ and $C_2 = 29.6977$, the wavelet $\psi_1(x)$ has a unique vanishing moment and the wavelet $\psi_2(x)$ has two vanishing moments.

In the following examples we have also employed the convolution $\psi_3 = \psi_1 \otimes \psi_2$ and the correlation $\psi_4 = \psi_1 \ast \psi_2$ of the previous wavelets, whose analytical expressions are given by

$$\psi_3(x) = \frac{1}{2\sqrt{6}} x(x^2 - 6) \exp \left\{ -\frac{x^2}{4} \right\}$$

(49)

and $\psi_4(x) = -\psi_2(x)$ because of the symmetry properties of $\psi_1(x)$ and $\psi_2(x)$. Both new wavelet functions also satisfy the admissibility condition (21) with coefficients $C_3 = C_4 = 52.6379$ and have three vanishing moments, in accordance with Theorem 1. All four wavelet functions are depicted in Fig. 1.

Example 1. Consider a test signal defined as

$$f(x) = \exp \left\{ -\frac{(x - 1)^2}{6} \right\} \cos(3(x - 1))$$

(50)

whose plot is shown in Fig. 2(a). The validity of the wavelet correlation theorem is established by calculating the autocorrelation of the test signal in the wavelet domain using Theorem 3 and comparing the result with the autocorrelation obtained using the correlation theorem for the FT.

The generated test signal is analyzed with wavelets $\psi_1(x)$ and $\psi_2(x)$ (all integrals have been numerically evaluated using the midpoint quadrature rule). According to Theorem 3, both CWT are correlated at every fixed scale in order to obtain the CWT of the autocorrelation of the test signal with wavelet $\psi_4(x)$. Finally, such an autocorrelation is obtained by applying the reconstruction formula (22) to its CWT.
This autocorrelation of the test signal is shown in Fig. 2(b) together with the autocorrelation obtained in the Fourier domain. Both autocorrelations are identical, which constitutes an empirical confirmation of the validity of the wavelet correlation theorem proved in this work.

Example 2. In this example, we present results that demonstrate the use and applicability of the wavelet convolution theorem in order to restore degraded signals and compare them with the restoration obtained working in the Fourier domain. Consider now a non-stationary test signal defined as:

\[
f(x) = \begin{cases} 
0 & \text{if } x \leq 2, \\
x - 2 & \text{if } 2 < x \leq 3, \\
1 & \text{if } 3 < x \leq 4, \\
5 - x & \text{if } 4 < x \leq 6, \\
-1 & \text{if } 6 < x \leq 7, \\
x - 8 & \text{if } 7 < x \leq 8, \\
2 \exp \left\{ \frac{(x - 15)^2}{10} \right\} & \text{if } x \geq 8 \\
\times [\cos((x - 15)^2) - a] & \text{if } x \geq 8 
\end{cases}
\] (51)

whose plot is shown in Fig. 3(a). A value \( a = 0.2337977 \) has been chosen because in such a case the test signal verifies \( \int_{-\infty}^{+\infty} dx f(x) = 0 \). This signal has been blurred with a gaussian filter \( h(x) \) with standard deviation \( \sigma_h = 0.15 \) and random white noise \( n(x) \) of zero mean and standard deviation \( \sigma_n = 0.2 \) has been added; Fig. 3(b) shows the degraded signal \( g(x) = (f \otimes h)(x) + n(x) \).

We have first deconvolved this degraded signal with a Wiener-type filter defined in the Fourier domain; if \( \hat{h}(w) \) and \( \hat{g}(w) \) are the FT of the gaussian blur and the degraded signal, respectively, then the FT of the restored signal is given by

\[
\hat{f}^\prime(w) = \frac{\hat{g}(w)}{\hat{h}(w)^2 + \alpha N|\hat{h}(w)|^2},
\] (52)

where \( N \) is the maximum value of \( |\hat{h}(w)|^2 \) and \( \alpha \) is a regularization parameter. Fig. 4(a) shows the restored signal \( f^\prime(x) \) for \( \alpha = 0.001 \).

Now, we have deconvolved the degraded signal in the wavelet domain taking into account Theorem 2.

The gaussian blur has been analyzed with wavelet \( \psi_h(x) \) and the degraded signal with wavelet \( \psi_g(x) \), so that the restored signal must be synthesized with wavelet \( \psi_f(x) \). Given the CWT \( \hat{f}^\psi(a, b) \) of some function \( f(x) \), let \( \mathcal{F}_2 \{ \hat{f}^\psi \} (a, w) \) be its FT in the second variable \( (b \leftrightarrow w \text{ are a Fourier pair}) \), that is

\[
\mathcal{F}_2 \{ \hat{f}^\psi \} (a, w) = \int_{-\infty}^{+\infty} db \hat{f}^\psi(a, b) \exp\{-2\pi i bw\}. \] (53)

Deconvolution is then performed at every fixed scale by applying the same Wiener-type filter than before, that is

\[
\mathcal{F}_2 \{ \hat{f}^\psi \} (a, w) = \mathcal{F}_2 \{ \hat{g}^\psi \} (a, w) \times \frac{(\mathcal{F}_2 \{ \hat{h} \} (a, w))^*}{|\mathcal{F}_2 \{ \hat{h} \} (a, w)|^2 + \alpha N|\mathcal{F}_2 \{ \hat{h} \} (a, w)|^2}, \] (54)
where $N$ is the maximum value of $|\mathcal{F}_2\{\hat{h}(a,w)\}|^2$. Fig. 4(b) shows the restored signal $f'(x)$ for the same value of the regularization parameter than before, $\alpha = 0.001$. In order to obtain a quantitative measure of the quality of the restorations, define the percentage error of a restoration $f'(x)$ as

$$\text{error} = 100 \left( \frac{\int_{-\infty}^{+\infty} dx (f(x) - f'(x))^2}{\int_{-\infty}^{+\infty} dx f(x)^2} \right)^{1/2}.$$  

(55)

Restored signal in the Fourier domain (Fig. 4(a)) has error = 5.7%, while the restoration in the wavelet domain (Fig. 4(b)) has error = 3.3%. These results show the usefulness of Theorem 2 in order to perform signal restoration in the presence of noise, although the employed deconvolution filter may be improved because it is not an optimal one, since it would be possible to use a different regularization parameter for every scale; this remains an open question for future research.

5. Concluding remarks

In this work we have studied the application of the CWT to perform signal filtering processes. We have first shown that the convolution and correlation of two wavelet functions satisfy the required admissibility and regularity conditions, and then we have derived convolution and correlation theorems for the CWT, which are similar to that of other joint time–frequency representations. The main feature of such theorems is that the convolution and correlation operators apply at every scale independently of each other, and this fact allows us to deeply understand the effect of the filter on every scale of the given signal. We have then shown that multiplication of the CWT of a given signal by some transfer function is equivalent to perform a spatially variant filter on such a signal; it remains for future work the search for useful filters that can be better defined in the wavelet domain than in the Fourier domain.
We have also shown results from two numerical examples that have illustrated the validity of the theorems proved in this work and their usability in signal restoration problems in the presence of additive noise. Comparisons between restorations of a non-stationary test signal in Fourier and wavelet domains have demonstrated the benefit of using the wavelet convolution theorem for performing signal deconvolution. It remains an open question the search for an optimal deconvolution filter defined in the wavelet domain.

Results proved in this paper show a close parallelism between convolution related filtering properties of joint time–frequency representations and CWT and may extend the range of applications of the later one. Future work may rely on the search for analogous convolution and correlation theorems for the discrete wavelet transform; the main drawbacks are that scaling and wavelet functions are not translation invariants and must satisfy more restrictive properties so that their convolutions and correlations are no longer usable ones except in particular cases.

Acknowledgements

The authors would like to thank Professor Alberto Márquez for his support and for helpful comments which have improved the development of this work. The authors would also like to appreciate the anonymous reviewers for their valuable suggestions, which greatly improved this paper.

References


