Towards a Wavelet Information Theory

H.M. de Oliveira, IEEE Member

Federal University of Pernambuco - UFPE, Signal Processing Group,
C.P. 7800, 50.711-970, Recife - PE, Brazil E-mail: hmo@ufpe.br

WORKSHP BRAZILIAN IT CHAPTER 06/07/Nov. 2003 Recife-PE

Abstract —

- ✓*New reading for wavelets, based on the De Broglie principle*
- ✓*A continuous wavelet is associated to a probability density*
- ✓*Entropy of a wavelet: Shannon Jumarie or Renyi*
- ✓*Entropy conservation principle*
- ✓*Shannon entropy for a few standard wavelet families*
- ✓*Wavelet mutual information between a signal and an analysing wavelet*
- ✓*Information provided by a multiresolution analysis*

1. Preliminaries

Entropy (Greek: *en*+*trope*=in+turning) is one of the most fundamental concepts of Science. Entropy: origin of the life or the future of the universe.

German Chemistry W. Ostwald:

> "*Energy is the queen of the world, and entropy is her shadow*!"

Von Neumman to Shannon:

> "*it will give you a great edge in debates because nobody really knows what entropy is anyway*."

A first question: "*Is there possible to associate an entropy measure to a wavelet*?"

Jumarie introduced entropy associated to a given continuous differentiable function $f:\Omega\rightarrow\Re$ as

$$H(f(.);\Omega):=\frac{\int_{\Omega}|f'(x)|.\log\left(|f'(x)|\right)dx}{\int_{\Omega}|f'(x)|dx}.$$

An option for measuring entropy of a continuous differentiable wavelet $y:\Re\rightarrow\Re$:

$$H(y;\Re):=\frac{\int_{-\infty}^{+\infty}|y'(t)|.\log\left(|y'(t)|\right)dt}{\int_{-\infty}^{+\infty}|y'(t)|dt}.$$

Max Born footnote for Schrödinger equation applied in wavelet framework:

wavelets can behave as some kind of corpuscle of a true random nature.

✓Associate a PDF $p_t(y) := y^2(t)$ to every continuous wavelet $y(t)$.

Fourier transform is isometric, ==>

✓Another probability density function $p_f(y) := \frac{1}{2p} \left|\Psi(w)\right|^2$.

It can be stated that:

"*wavelets are to corpuscle as wave-functions are to particles*."

After unveiling probabilistic properties associated to wavelets:

Entropy associated to a wavelet as a measure of the disorder of a signal.

The entropy of a random variable:

Shannon differential entropy of a r.v. X with density $p(x)$ is

$$H(X) := -\int_{-\infty}^{+\infty} p(x) \log p(x) dx \, .$$

<u>Definition 1</u>. (Shannon entropy of a wavelet).

The time entropy, $H_t(y)$, of a continuous wavelet $y(.)$ is defined by

$$H_t(y) := -\int_{-\infty}^{+\infty} y^2(t) \log_2\left(y^2(t)\right) dt \quad \textit{shannons};$$

Analogously, the frequency entropy, $H_f(y)$, of a continuous wavelet $y(.)$ is defined by

$$H_f(y) := -\int_{-\infty}^{+\infty} \frac{1}{2p} \left|\Psi(w)\right|^2 \log_2 \frac{1}{2p} \left|\Psi(w)\right|^2 dw \quad \textit{shannons}. \ ❑$$

The entropy gives information on the spreading of the wavelet, i.e., it furnishes a "localising measure" of the corpuscle in a particular domain (time or frequency).

REMARK:

Shannon entropy of a wavelet is exactly the Jumarie entropy of the PDF associated to the density $y^{2}(t)$, i.e., $H_t(y) = H(P(.);\Re)$.

"Non-shannonian" measures such as the Renyi entropy of order *s*>0:

$$H_t(y/s) := \frac{1}{s-1} ln\left(\int_{-\infty}^{+\infty} y^{2s}(t')dt' \right),$$

which hold $\lim_{s \to 1} H_t(y/s) = H_t(y)$.

Wavelets × entropy:

1999, Quian Quiroga, Rosso and co-worker: a disorder measure in Neuroscience.

The wavelet entropy, S_{WT}, was then defined by $S_{WT} := -\sum_j P_j \, ln(P_j)$.

✓Electroencephalogram analysis [ROSS 2001, QUIQUI 2001, YOR *et al*. 2002].

✓Astronomy (solar activity) [SELL 2003].

2. On The Shannon Entropy of Continuous Wavelets

Effect of scaling or shift:

<u>Proposition 1.</u> *Given a continuous mother wavelet* $y(.)$ *with time entropy* $H_t(y)$ *and frequency entropy* $H_f(y)$, *the entropy of a daughter wavelet can be computed by*

$$H_t(y_{a,b}) = H_t(y) + log_2|a| \text{ and } H_f(y_{a,b}) = H_f(y) - log_2|a| .$$

Among all supportly compacted wavelets, the highest time-entropy is achieved by the Haar wavelet

(Compatible with the fact that maximum entropy of a discrete random variable is achieved by a uniform distribution)

<u>Proposition 2.</u> *The time-entropy of any wavelet of compact support is upper bounded by* $H_t(y) \leq log_2(length(sup\ p(y)))$ *and the bound is only met by the Haar wavelet.* ❑

Corollary 1. *The time Shannon entropy of* dBN *wavelet is bounded by*

$$H_t(dBN) \le log_2(2N-1).$$ ❑

This result can equally be translated into the frequency domain

Corollary 2. *The frequency Shannon entropy of the* deO *wavelet* [deO et al. 2003] *is upper bounded by* $H_f(deO) \le log_2(p + 3pa).$ ❑

✓Some amount of uncertainty/information in both domains (matching with the Gabor uncertainty principle)

In the probabilistic interpretation offered here, even knowing $y(.)$ and $Y(.)$, it do exists some non-zero uncertainty in both time and frequency domain.

✓The global entropy would be partially due to the (inherent) time-uncertainty and partially due to the (inherent) frequency-uncertainty.

<u>Definition 2</u>. (Global entropy of a continuous wavelet).

The global entropy of a wavelet $y(.)$ *is defined by* $H_y := H_t(y) + H_f(y)$. ❑

Every daughter wavelet has the same global entropy of the mother wavelet, some sort of conservation principle.

<u>Corollary 3.</u> *The global entropy is preserved within the same wavelet family* $\{y_{a,b}(t)\}_{a\neq 0, b\in\Re}$ *so we are able now to find a unique entropy value associate to a wavelet basis*. ❑

Since the two densities $y^{2}(t)$ and $\frac{1}{2p}\left|\Psi(w)\right|^{2}$ are related via $y(t)\leftrightarrow\Psi(w)$, the time-frequency relationship is disclosed on the "partition" of the total uncertainty.

The concept of isoresolution:

"*Are there wavelets that achieve the same entropy in both domain*?"

A class of signals that hold special and nice-looking properties are the invariant signals under Fourier transform.

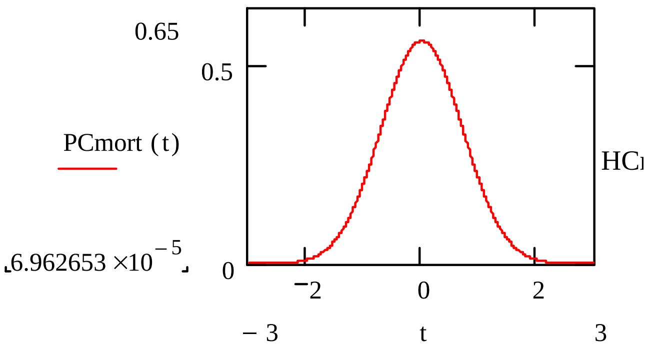
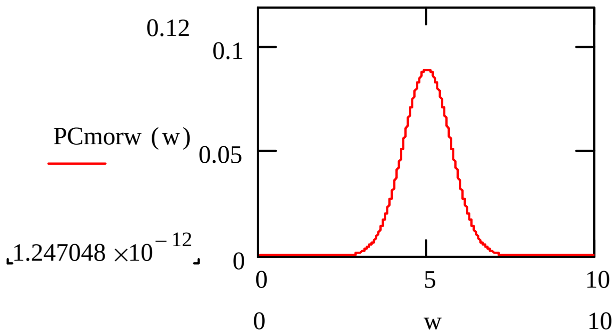
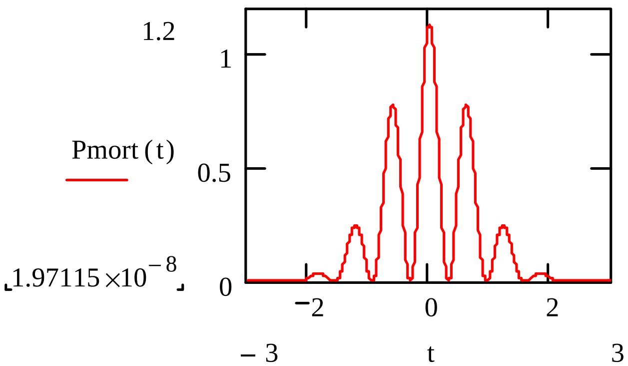
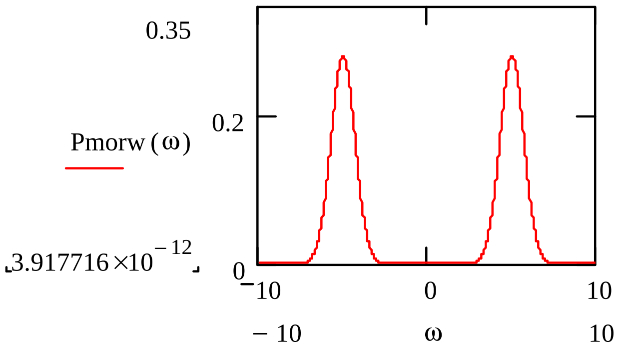
<u>Proposition 3</u>. Possible eigenvalues of the Fourier transform operator are the four roots of the unit (±1,±*j*) times $\sqrt{2p}$. ❑

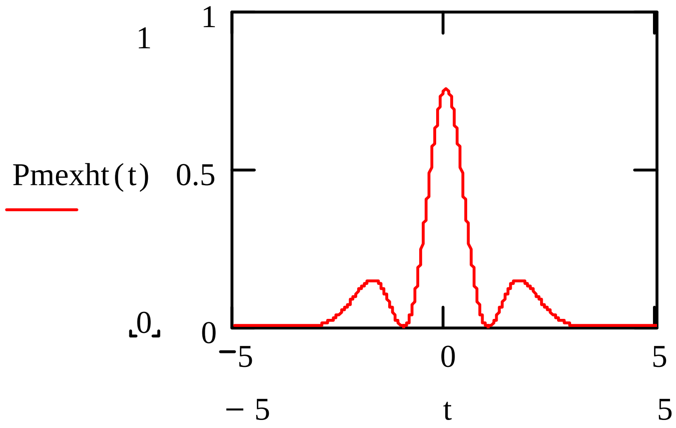
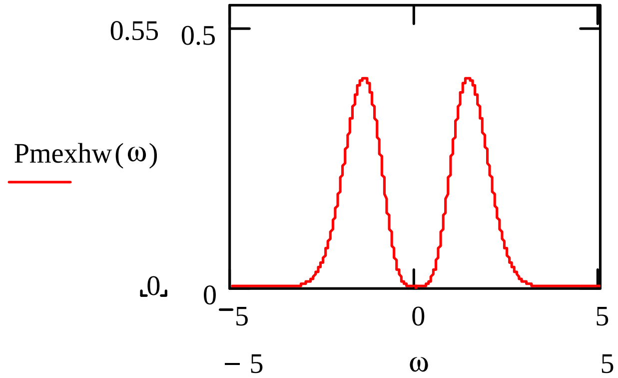
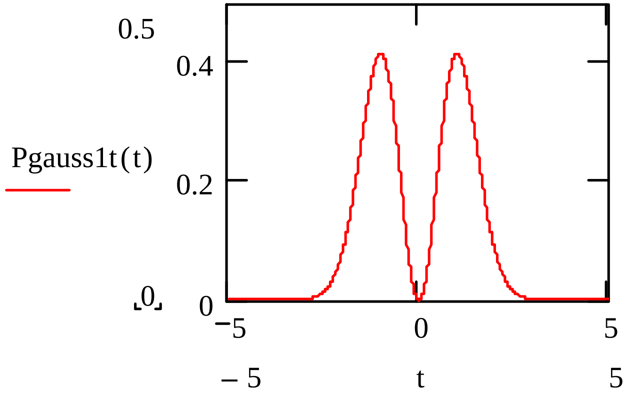
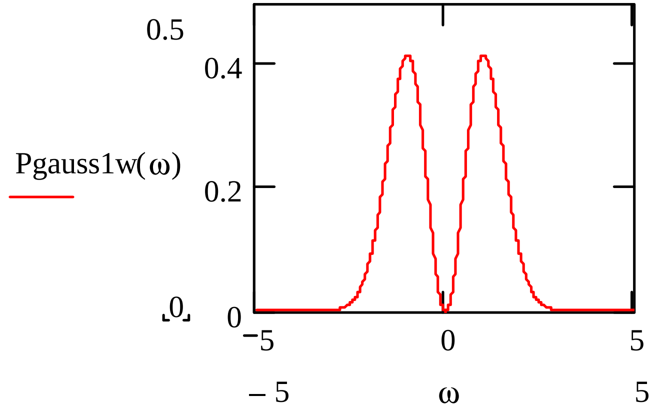
<u>Proposition 4</u>. *Isoresolution wavelets hold* $H_t(y) = H_f(y)$.

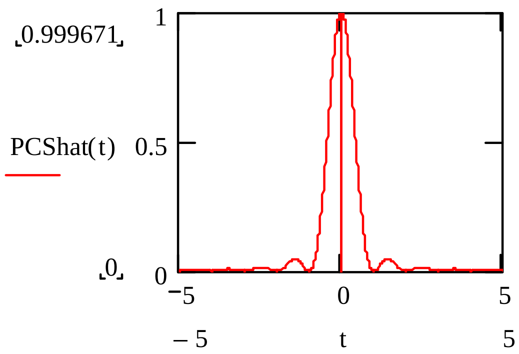
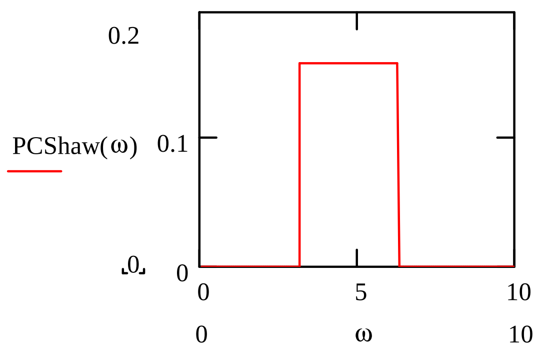
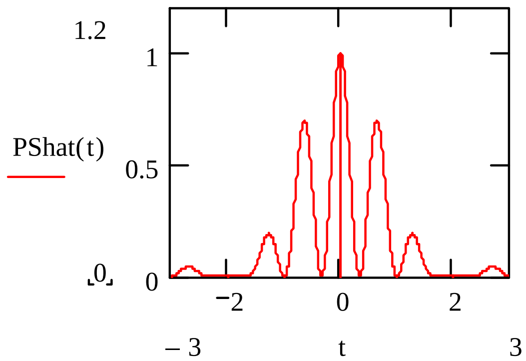
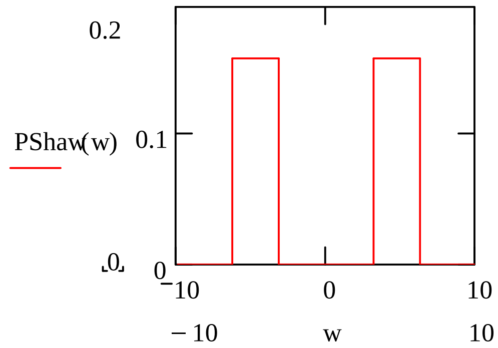
Plots of both the time and frequency density functions associated to wavelets

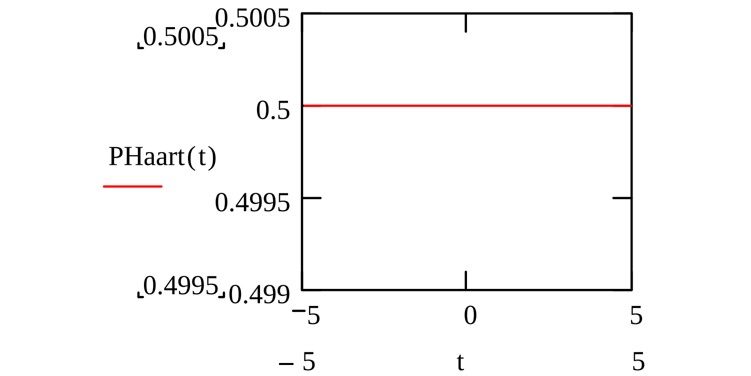
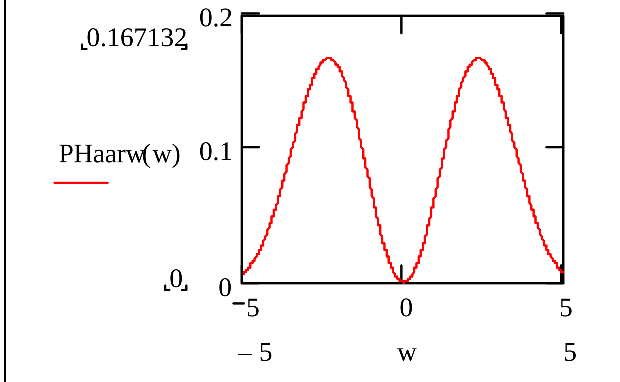
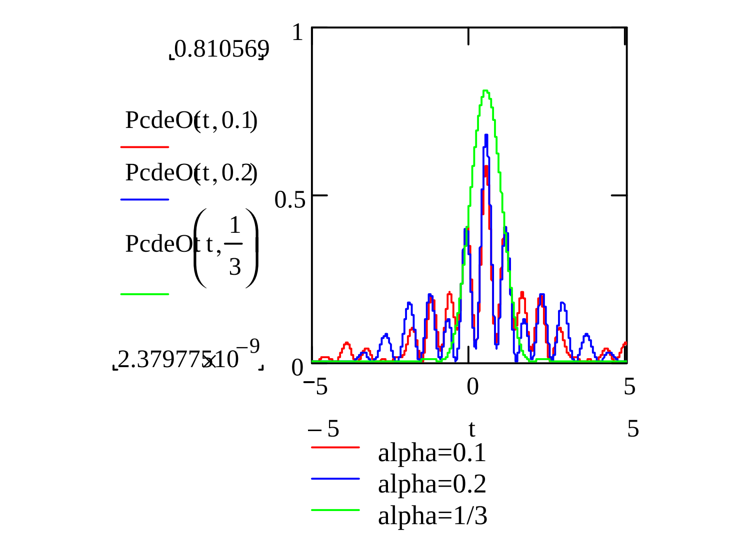
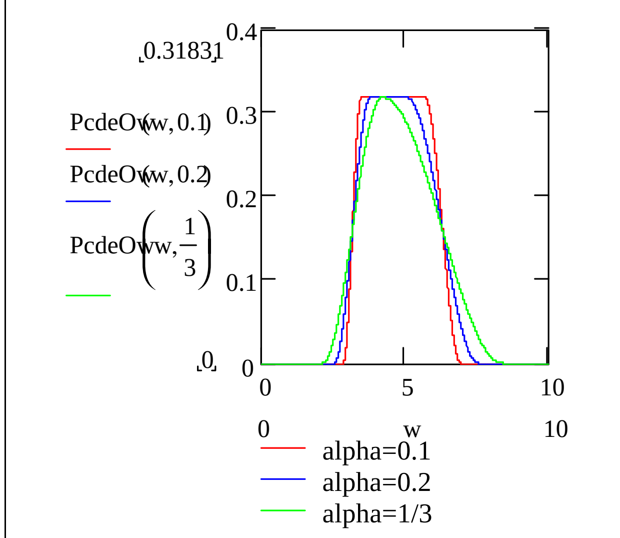
Complex Morlet (CMor)	$\psi\text{Cmor}(t) := \dfrac{e^{-j\cdot 5\cdot t}\cdot e^{\frac{-t^2}{2}}}{\sqrt[4]{\pi}}$	Morlet (Mor)	$\dfrac{\sqrt{2}\,cos(5t)e^{\frac{-t^2}{2}}}{\sqrt[4]{p}}$
Sombrero (mexh)	$\dfrac{2}{\sqrt{3}}\left(t^2-1\right)\dfrac{e^{\frac{-t^2}{2}}}{\sqrt[4]{p}}$	Gaussian 1 (gauss1)	$\dfrac{\sqrt{2}te^{\frac{-t^2}{2}}}{\sqrt[4]{p}}$
C Shannon (Sinc)	$Sinc(t)e^{-j2pt}$	Shannon (Sha)	$Sinc\left(\dfrac{t}{2}\right)cos\left(\dfrac{3pt}{2}\right)$
Haar (Haar)	$\begin{cases}\dfrac{1}{\sqrt{2}}sgn(t) & if \quad \|t\|<1 \\ 0 & otherwise.\end{cases}$	de Oliveira (CdeO)	$\Re ey(t,a) := s(t-0.5,a)$ $\Im my(t,a) := \ddot{s}(t-0.5,a)$

Figure 1. Time and frequency pdf associated to some common wavelets.

wavelet	Time pdf	Frequency pdf
Cmor (a)		
Mor (b)		

Mexh (c)		
Gauss1 (d)		

<table>
<tr>
<td>Sinc
(e)</td>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td>Shan
(f)</td>
<td>
</td>
<td>
</td>
</tr>
</table>

Haar (g)		
CdeO (h)		

CMorlet and gauss1 are invariant wavelet, hence achieve isoresolution

⇨ balanced time and frequency entropy (Proposition 3).

Table 2. Entropy of some wavelets: Time entropy, frequency entropy, and area of a wavelet cell in the joint *t-f* plane and global entropy.

Wavelet y	Time entropy $H_t(y)$	Frequency entropy $H_f(y)$	Product $H_t(y). H_f(y)$	Wavelet entropy H_y
CMor	1.547096	1.547096	2.393506	3.094191
Mor	1.104425	2.547095	2.813075	3.651520
Mexh	1.715098	1.988567	3.410587	3.703665
CSha	2.221052	1.651383	3.667807	3.872435
Gauss1	1.937147	1.937147	3.752538	3.874293
Sha	1.768634	2.651665	4.689824	4.420299
CdeO a=				
0.1	3.045824	1.818698	5.539434	4.864522
0.2	2.915276	1.985887	5.789408	4.901163
Haar	1.000000	3.985653	3.985653	4.985653

(All information content measured in shannon units).

The closed expression (Morlet) $H_t(y_{CMorlet}) = H_f(y_{CMorlet}) = log_2\left(\sqrt{pe}\right)$.

<u>Conjecture</u>: minimum entropy H_y is achieved by a double Gaussian distribution (time and frequency domain), which will be in accordance with the concept of "logon" established by Gabor.

Apparently, two wavelets are particularly significant: Morlet wavelet and Haar wavelet. *Inquiringly, these were exactly pioneer wavelets!*

3. Cross Density and Cross-Entropy of Wavelets

Lemma 7. (Gibbs inequality for wavelets)

Given two continuous wavelets $\{y_i(t)\}_{i=1,2}$ *, then*

$$-\int_{-\infty}^{+\infty} y_1^2(t)\log_2\left(y_1^2(t)\right)dt \leq -\int_{-\infty}^{+\infty} y_1^2(t)\log_2\left(y_2^2(t)\right)dt$$

with equality if and only if $|y_1(t)|=|y_2(t)|$ $(\forall t)$. ❑

A simple upper bound on the entropy of a wavelet can be easily derived.

Proposition 8. *The global entropy of a wavelet* $y(.) \leftrightarrow Y(.)$ *is upper bounded by*

$$H_y \leq -\int_{-\infty}^{+\infty} y^2(z)\log_2\left(\frac{1}{2p}\left|\Psi(z)\right|^2\right)dz - \int_{-\infty}^{+\infty}\frac{1}{2p}\left|\Psi(z)\right|^2 \log_2\left(y^2(z)\right)dz$$

with equality if and only if $y(.)$ *is an invariant wavelet.* ❑

Kullback-Leiber distance can assess the distance between two wavelets.

<u>**Definition 3**</u>. (Wavelet distance). *Given two continuous wavelets* $\{y_i(t)\}_{i=1,2}$ *compactly supported on Supp(*$y_i(t)$*), i*=1,2 *respectively, the Kullback discriminant can evaluate the distance between the two associated densities according to*

$$D_1(y_1,y_2) := H(y_1 \| y_2) = 2\int_{Supp(y_1)} y_1^2(t).\log_2\left(\frac{y_1(t)}{y_2(l_{1,2}.t)}\right)dt$$

, where $l_{1,2}$ *is the ratio between the length of the support of the two wavelets* $y_1(t)$ and $y_2(t)$. ❑

4. By Way of a Wavelet Information Theory

Questions such as "*How many information unity an analysing wavelet provides on a continuous signal*?" could then be addressed.

Evoking the definition of the mutual information between two continuous random variables *X* and *Y*, we recognise that the key for creating such a measure is the definition of some joint probability measure.

We imply that the energy conservation and the resolution of the identity can be used so as to define a joint density between a signal and a wavelet in the *wavelet domain*.

<u>Definition 4</u>. (Joint signal-wavelet probability density). *Let* $f \in L^2(\Re)$ *be a signal of energy* $E := \int_{-\infty}^{+\infty} f^2(t)dt$. *Given a mother continuous wavelet* $y(.)$ *with admissibility constant* $c_y < +\infty$, *the joint density between f and a version* $y_{a,b}(t)$, $a \neq 0$ *can be defined as*

$$p(a,b) := \frac{|CWT(a,b)|^2}{E.c_y\,.|a|^2}, \quad a \neq 0. \quad \square$$

$p(a,b)$ is a joint probability density (between the signal and the analysing wavelet) over the scale-translation domain.

<u>Definition 5.</u> (Mutual signal-wavelet information). *The wavelet mutual information* $I(f(t),\{y_{a,b}(t)\})$ *between the signal and analysing wavelet can be found in a way somewhat similar the classical definition* $I(X;Y)=H(X)-H(X|Y)$ *after using Bayes' rule*:

$$I(f(t),\{y_{a,b}(t)\}) := \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{|CWT(a,b)|^2}{E.c_y\,.|a|^2} \log_2\left(\frac{E.c_y\,.|CWT(a,b)|^2}{\int_{-\infty}^{+\infty}|CWT(a,b)|^2\,db \int_{-\infty}^{+\infty}|CWT(a,b)|^2\,\frac{da}{a^2}} \right) dadb \qquad \square$$

This allows for computing the amount of information provided by the decomposition of a signal f using a wavelet y.

Continuous time wavelet series (CTWS) can be used instead of the CWT.

✓ dyadic case, $\left\{ y_{n,m} = \frac{1}{\sqrt{2^m}} y\left(2^{-m} t - n\right) \right\}_{n,m \in Z}$.

✓

✓Wavelet coefficients (details)

$$w_{n,m} := < f, y_{n,m} > = \int_{-\infty}^{+\infty} f(t) y^{*}_{n,m}(t) dt \text{ , } n,m \in \mathrm{Z},\ n \neq 0.$$

For orthogonal wavelets, the set $\{y_{n,m}(t)\}_{n\in Z, m\in Z, t\in \Re}$ is a basis for $L^2(\Re)$

$$f(t) = l.i.m. \sum_{n\in Z} \sum_{m\in Z} w_{n,m} y_{n,m}(t) \, ,$$

$$\sum_{n\in Z} \sum_{m\in Z} | w_{n,m} |^2 = \int_{-\infty}^{+\infty} f^2(t) dt = E$$

The mutual signal-wavelet information can be derived from the homogeneous wavelet expansion as

$$I(f, \{y\}) = \sum_{n\in Z} \sum_{m\in Z} \frac{| w_{n,m} |^2}{E} \log_2 \frac{E. | w_{n,m} |^2}{\sum_{m'\in Z} | w_{n,m'} |^2 \sum_{n'\in Z} | w_{n',m} |^2} . \; ❑$$

Mallat multiresolution approach

Suppose that $\{V_j\}$ constitutes the closed approximation subspaces of a multiresolution analysis (MRA) of $L^2(\Re)$ and let $\{W_j\}$ denote the sequence of orthogonal complementary (wavelet) spaces.

For any given level *j*, $clos\left(V_j \oplus W_j \oplus W_{j-1} \oplus \ldots\right) = L^2(\Re)$, where *clos* denote the closure union and $\oplus$ indicate "orthogonal sums" of subspaces.

An MRA expansion can be written in terms of scaling functions (j) and wavelet functions (y) as

$$f(t) = l.i.m. \sum_{k \in Z} v_{k,J} j_{k,J}(t) + \sum_{j=-\infty}^{J} \sum_{k \in Z} w_{k,j} y_{k,j}(t)$$, and

$$\sum_{k \in Z} | v_{k,J} |^2 + \sum_{j=1}^{J} \sum_{k \in Z} | w_{k,j} |^2 = \int_{-\infty}^{+\infty} f^2(t) dt = E$$,

where $\boldsymbol{v}_J = \{v_{k,J,}\}_{k \in Z}$ are the *approximation coefficients* and $\boldsymbol{w}_j = \{w_{k,j}\}_{k \in Z}$. $j=1,2,\ldots,J$ are the *detail coefficients* of the MRA.

The mutual signal-wavelet information of an MRA can be computed from the inhomogeneous wavelet expansion by

$$I\left(f,\{j\ ,y\ \}\right)=\sum_{k\in Z}\sum_{j=1}^{J}\frac{\left(|\,v_{k,J}\,|^{2}/J+|\,w_{k,j}\,|^{2}\right)}{E}.log_{2}\frac{E.\left(|\,v_{k,J}\,|^{2}/J+|\,w_{k,j}\,|^{2}\right)}{\left(\sum_{k\in Z}\left(|\,v_{k,J}\,|^{2}+|\,w_{k,j}\,|^{2}\right)\right).\left\{|\,v_{k,J}\,|^{2}+\sum_{j=1}^{J}|\,w_{k,j}\,|^{2}\right\}}.$$

Computations are in progress...

5. Concluding Remarks

✓This paper intends to regard wavelet as an information processing technique.

✓Mimicking the Quantum Mechanics, each wavelet is associated to a PDF

✓Shannon entropy and Renyi entropy of a wavelet were defined.

✓It was derived some sort of entropy conservation principle.

✓The core of a continuous wavelet information theory is fulfilled by defining wavelet mutual information.

- ✓ **The emphasis of this paper was on conveying the chief ideas behind the WIT as opposed to presenting a formal mathematical development or applications.**

- ✓ We hope this attempt to present the underlying philosophy of WIT may help users navigate the ocean of wavelets.

Acknowledgements: *The author is obliged to Professor Valdemar C. da Rocha Jr (head of the Brazilian IEEE-IT-Chapter–Brazil Council), who was permanently fine at creating an exceptional atmosphere of enthusiastic, inspiring interest in IT.*