ACÚSTICA 98

USO DE LA TRANSFORMADA WAVELET EN APLICACIONES DE ACÚSTICA MUSICAL

Enric Guaus Térmens, Jordi-Joan Garcia Garcia, Jose Manuel Garcia de la fuente

Departament d'Acústica
Escola d'Enginyeria La Salle, Universitat Ramon Llull
Pg. Bonanova 8, 08022 Barcelona, Catalunya
Tel: +34 932902427 / Fax: +34 932902416
e-mail: eguaus@salleURL.edu

SUMMARY

In former times, the study of instrument dynamical behaviour was divided in two parts: The Frequency Domain and the Time Domain. The common choice for combining both, is the use of time-frequency representations, so called sonograms. But this technique has a drawback: the incertainty principle. When a fine resolution in the time domain is required, then resolution in the frequency domain is lost, and viceversa. The workaround is to use the Wavelet transform.

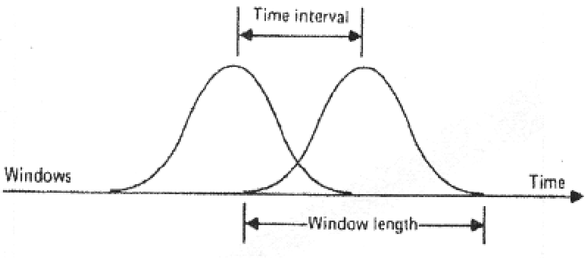
PRINCIPIOS MATEMÁTICOS

La STDFT: Principio de Incertidumbre

Un sonograma es una representación tiempo-frecuencia de una señal cualquiera. Para señales estacionarias o señales cuya variación en el tiempo sea relativamente lenta es una herramienta de cálculo potentísima, puesto que podemos ver la evolución de las características frecuenciales a lo largo del tiempo. La base de su funcionamieto es la siguiente:

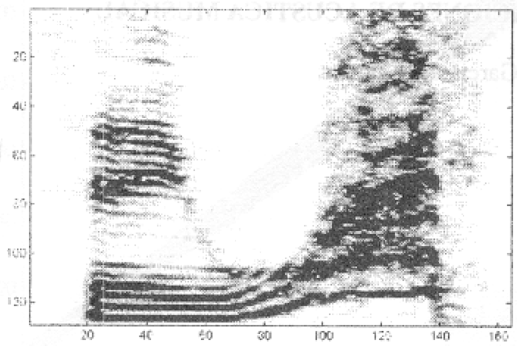
Partimos de una secuencia discretizada x[N], donde N es un entero que indica la longitud. Se divide esta secuencia en fragmentos, se enventanea y se aplica el algoritmo de la FFT. Con las distintas FFTs de los distintos fragmentos vamos construyendo nuestro sonograma, ya sea en parte real e imaginaria o en módulo y fase. El tipo de ventana escogida (Hanning, Hamming, Blackman...) dependerá de las características de Leakage y Picked-fence que estemos dispuestos a aceptar. Por otra parte, a fin de asegurar la continuidad temporal, los fragmentos escogidos serán solapados ligeramente (en un porcentage de overlap que nos determina el intérvalo de tiempo entre ventana y ventana). Finalmente, la longitud de la ventana será la misma que la cada fragmento.

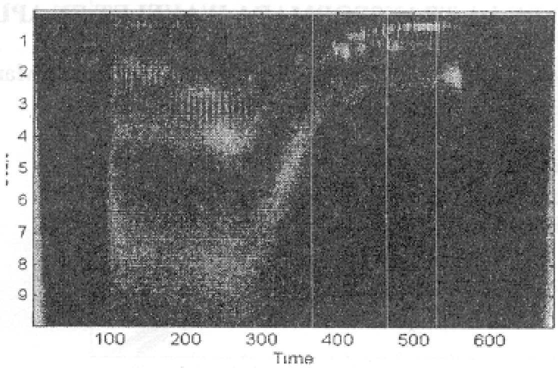
Hay dos parámetros con los que podemos jugar: El intérvalo de tiempo entre ventanas y su longitud (igual a la longitud de la FFT). Con el tiempo entre ventanas podremos influir sobre la resolución temporal. Si aplicamos un overlap muy bajo perderemos los fenómenos ocurridos en el intervalo de tiempo donde las ventanas tienen mínima amplitud, mientras que si aplicamos un overlap muy elevado haremos un promediado de



distintos intervalos de tiempo para una misma muestra, con lo que quedará un sonograma más suavizado. Tenemos un compromiso. Típicamente, se usan valores de overlap=50%. Con la longitud de la ventana (igual al nº puntos para calcular la FFT) influiremos sobre la resolución frecuencial. Si escogemos un número muy elevado de puntos, tendremos una elevada resolución frecuencial (el ancho de banda de la ventana es inversamente proporcional a su longitud) pero perderemos resolución temporal ya que el tiempo entre ventanas aumenta, y si decidimos aumentar el overlap conseguiremos un sonograma aun más suavizado en el tiempo. Si escogemos un número de puntos pequeño, aunque aumentemos la resolució temporal perdemos resolución frecuencial.

Este fenómeno es conocido como principio de incertidumbre. La Transformada Wavelet elimina el principio de incertidumbre mediante el uso de ventanas adaptativas para cada frecuencia.





La Transformada Wavelet

La transformada de Fourier de corta duración es la convolución de una señal con senos de distinta frecuéncia.

$$F(w,t) = \int f(s) g(s-t) e^{-jwt} ds$$

donde g(s) es la ventana aplicada.

El resultado de esta convolución es la misma señal pero representada en una nueva base de funciones. El problema de este tipo de base es que la anchura de la ventana es constante para cada frecuéncia. Tal y como hemos visto anteriormente, nos iteresa una ventana ancha para bajas frecuencias y una ventana estrecha para altas frecuencias. Para ello, nos aprovechamos de la propiedad de escalado de la TF:

$$\psi(t) \leftrightarrow \Psi(w) \Rightarrow \frac{1}{\sqrt{a}} \psi\left(\frac{t}{a}\right) \leftrightarrow \sqrt{a}\,\Psi(aw)$$

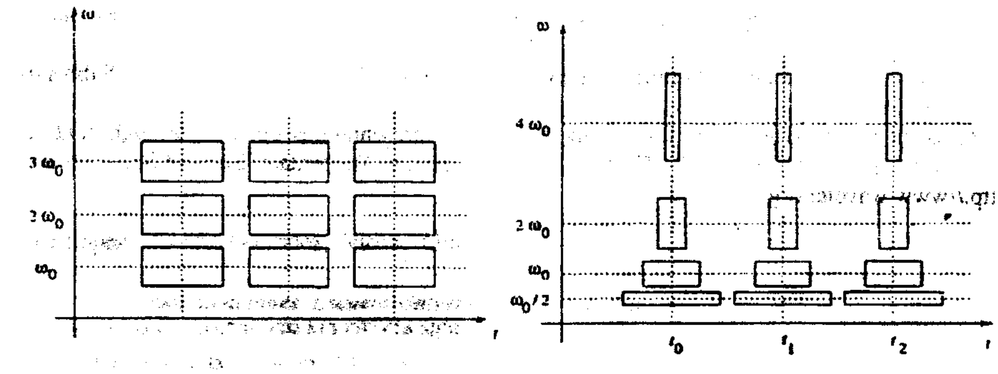
donde una contracción en un dominio implica la expansión en el otro. Ahora, podemos pensar en construir una nueva base de funciones a partir de dilataciones y translaciones de una función Wavelet madre:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right)$$

consiguiendo un nuevo cambio de base: La Transformada Wavelet W(a,b). La wavelet madre más usada es la ventana Morlet. Los parámetros de dilatación y translación tienen el mismo papel que la *w* y *t* de la TF. La fórmula general de la TW es la siguiente:

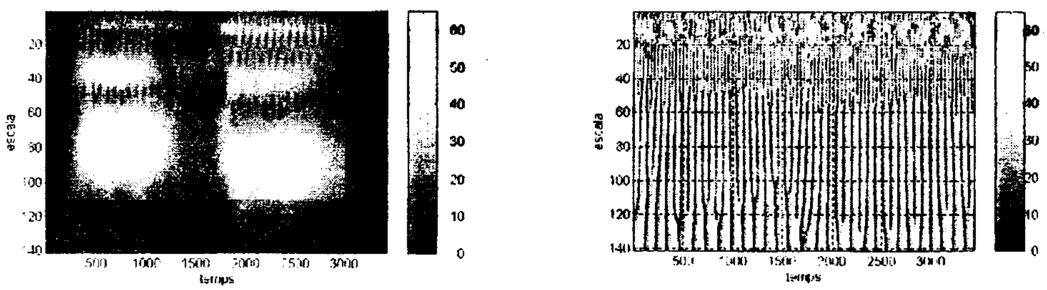
$$W(a,b) = \int_{-\infty}^{+\infty} \psi_{a,b}(t) f(t) dt$$

Si comparamos el plano tiempo-frecuéncia de la TF y la TW comprenderemos como actuan la dilatación y translación de la ventana.



APLICACIÓN: SÍNTESIS ADITIVA

La síntesis aditiva és una técnica para crear un sonido mediante la superposición de los senos sintonizados a la frecuencia de unos harmónicos, con su amplitud, fase y amortiguamiento. En este apartado trataremos de hacer síntesis aditiva pero con la TW, en vez de la TF.



Como se ve en la figura, en una representación tiempo-frecuencia mediante la TW, tenemos información de amplitud (modificada por la ventana) y de fase para cada escala (la relación entre escala y frecuencia se deduce fácilmente en función de la ventana utilizada). Para la síntesis aditiva necesitamos información de frecuencia, amplitud, fase y amortiguamiento de cada harmónico. Puesto que tenemos la evolución temporal, la información del amortiguamiento es redundante. Así pues, tenemos los datos necesários para nuestra síntesis.

Que ventajas presenta trabajar con la TW en vez de la TF? Evidentemente, tenemos una mayor resolución tiempo-frecuencia. Pero por otra parte, la fase de la TW nos da información de la frecuencia instantanea de la señal: si tenemos localizados a que escala pertenecen los máximos del valor absoluto de la TW para cada instante de tiempo, el valor de la señal sintetizada (ya en el dominio temporal) será una suma de senos cuyo argumento es directamente el valor de fase correspondiente a esos máximos. Recordamos que los máximos de amplitud corresponden a los harmónicos.

Pero no todo son ventajas: El coste computacional de la TW es mucho más elevado que en el caso de la TF. Para minimizar este defecto existe el llamado Análisis Multiresolución que se basa en una descomposición de la señal mediante filtrado y otro tipo de procesado adicional de la señal. Pero ese ya es otro tema.

CONCLUSIONES

Hemos presentado una herramienta de cálculo que puede mejorar claramente las posibilidades de otras herramientas más usadas. Puede que sea más complicada (o menos intuitiva) y que requiera un coste computacional más elevado, pero no es problema. Una representación tiempo-escala de la TW es como una representación tiempo-frecuéncia de la TF, pero el eje frecuencial está en escala logarítmica, con lo que podemos representar octavas en lugar de bandas constantes. Y en cuanto al coste computacional, sinceramente, no creemos que sea un factor crítico. Algunos analizadores de espectros de última generación permiten ya la visualización tanto en sonogramas como en escalogramas.

REFERENCIAS

1. Joan Serrat, Felipe Lumbreras: Introducció a les Wavelets i les seves aplicacions en processament d'imatges. Abstract
2. R.A.Haddad, A.N.Akansu: Multiresolution Signal Descomposition. Transforms, Subbands Wavelets. Academic Press, 1992
3. A. Grossmann, R. Kroland-Martinet: Time-and-scale representations obtaines through continuous wavelet transforms" Signal Processing IV: Theories and Applications. Amsterdam: Elsevier Science Publishers
4. http://www.wavelet.org