Time-frequency analysis with the continuous wavelet transform

W. Christopher Lang and Kyle Forinash
Natural Sciences Division, Indiana University Southeast, New Albany, Indiana 47150

(Received 20 June 1997; accepted 6 February 1998)

The continuous wavelet transform can be used to produce spectrograms which show the frequency content of sounds (or other signals) as a function of time in a manner analogous to sheet music. While this technique is commonly used in the engineering community for signal analysis, the physics community has, in our opinion, remained relatively unaware of this development. Indeed, some find the very notion of frequency as a function of time troublesome. Here spectrograms will be displayed for familiar sounds whose pitches change with time, demonstrating the usefulness of the continuous wavelet transform. © *1998 American Association of Physics Teachers.*

Imagine the following engineering problem: Develop software which from recorded music will produce correct sheet music notation for that music. Thus if the note ''C'' is heard (a pitch of 262 Hz) for one-fourth of a second, followed by the note ''A'' (440 Hz) for another quarter-second, the software would produce a plotting indicating 262 Hz for $0\leqslant t\leqslant 0.25$ and 440 Hz for $0.25\leqslant t\leqslant 0.5$. We may assume the input music would be presented in numerical form. (Thus, for example, a monaural sound sampled at 9000 Hz might be given as 9000 eight-bit numbers per second.) How then would our software accomplish this task?

The most basic technique for determining the frequency distribution of a signal $f(t)$ is the Fourier transform. This is given by the familiar integral transform $\hat{f}(\omega)=\int_{-\infty}^{\infty}f(t)e^{-it\omega}\,dt$. If we wish to determine what pitches or frequencies were audible during the time interval $0.25\leqslant t\leqslant 0.5$, we could perhaps compute the Fourier transform of f restricted to that time interval. That is, we could simply compute that integral from $t=0.25$ to $t=0.5$ instead of $t=-\infty$ to ∞. This will work after a fashion (although we cannot expect the Fourier transform to display too narrow a peak at ω equal to 440 Hz, since the integral is not computed over very many cycles). More generally, we could select a value h (representing perhaps one-half of the duration of the shortest notes of the music) and compute for each t and ω the integral

$$T_h f(t,\omega)=\int_{t-h}^{t+h} f(\tau)e^{-i\tau\omega}\,d\tau.$$

So $T_h f(t,\omega)$ would represent in some sense the energy of the signal at frequency ω in the neighborhood of time t. This transform is known as the ''short-time Fourier transform,'' and it has been the traditional technique in signal analysis for tracking pitches or frequencies as they change over time.

There is, however, an important limitation with the short-time Fourier transform: In restricting the integral to the interval $t-h$ to $t+h$, we are in effect chopping the signal sharply at those times. This has the effect of introducing lots of high-frequency ''information'' into the transform; we get a Fourier transform that is more spread-out than necessary. We can mitigate this problem to some extent by introducing a smooth ''window function'' g. We could arrange g to be nonzero on the interval $[-h,h]$ but be (close to) zero outside that interval. The Gaussian function $g(t)=(1/h\sqrt{2\pi})e^{-t^2/2h^2}$ would serve the purpose. We would then for each t and ω compute

$$Tf(\omega,t)=\int_{-\infty}^{\infty} g(\tau-t)f(\tau)e^{-i\omega\tau}\,d\tau.$$

Plotting $|Tf(\omega,t)|$ for each (t,ω) (as a density plotting) would produce a spectrogram which would show (at least roughly) the frequency content of the signal as a function of time. It could actually produce recognizable ''music notation'' for recorded music. This transform is known as the Gabor transform, after Dennis Gabor, who introduced it in the 1940s.

The Gabor transform has a subtle limitation which the continuous wavelet transform will be introduced to address. The limitation is this: The ''width'' of the window function g is constant. A narrow window (h small) will localize higher pitches both in frequency and time nicely, but lower pitches will be ''blurry'' in frequency. A wider window (h larger) will determine lower pitches (bass notes, say) better, but the higher pitches will be ''blurry'' in time. See Fig. 1(a) and (b) for plottings of an artificial signal which demonstrate this effect.

In the early 1980s, Morlet and Grossman modified the Gabor transform to produce the continuous wavelet transform. The idea is this: Change the width of the window function according to the pitch of the note being considered. The new transform could be given by

$$Wf(t,\omega)=\int_{-\infty}^{\infty} f(\tau)g((\tau-t)\omega)e^{-i\omega\tau}\,d\tau.$$

Notice the insertion of the factor ω into the window function. Actually, the continuous wavelet transform is more properly given as

$$Wf(t,\omega)=\int_{-\infty}^{\infty} f(\tau)w((\tau-t)\omega)\omega\,d\tau, \tag{1}$$

where now the ''wavelet'' function w combines the complex exponential and the Gaussian; w is typically given as

$$w(t)=\frac{1}{h\sqrt{2\pi}}\,e^{it}e^{-t^2/2h^2}. \tag{2}$$

This is called a wavelet function because it is oscillatory but has the limited support provided by the window function g. [The extra factor of ω in integral (1) keeps the ''mass'' $\int|w(t\omega)\omega|dt$ of the wavelet constant as it is dilated, as ω increases; otherwise the spectrogram will ''fade out'' at higher values of ω.] The parameter h can be adjusted to tune the transform to be more sensitive to frequency or more sen-

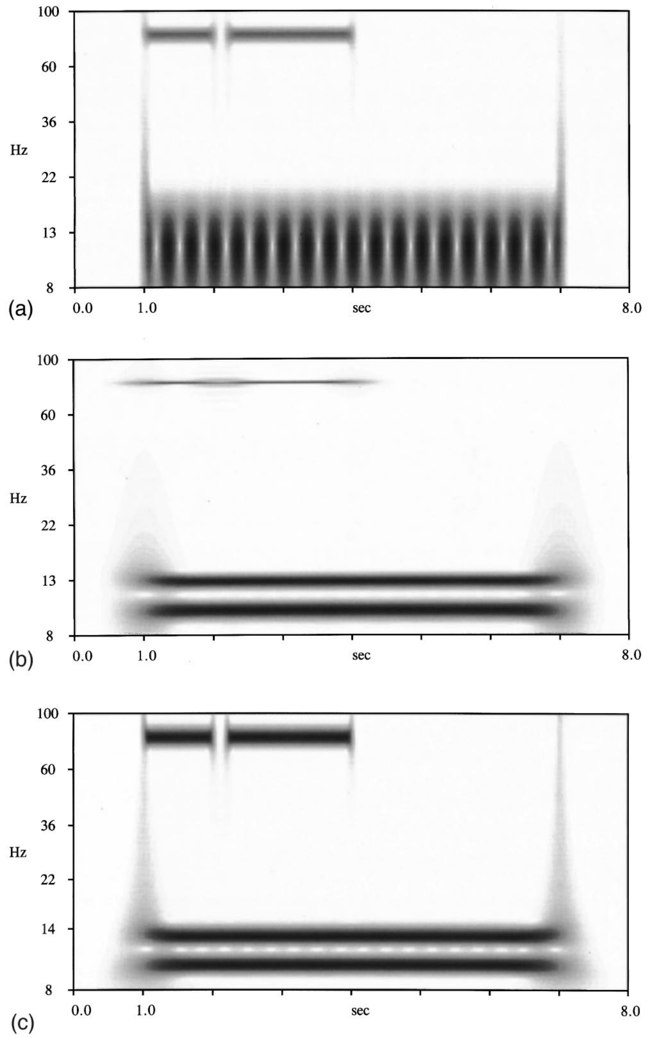

Fig. 1. Spectrograms of an artificial sinusoidal signal consisting of an interrupted 80-Hz pure tone superimposed over pure tones of 10 and 13 Hz. (a) shows a short-time Fourier (Gabor) transform spectrogram with a narrow window ($h=0.05$ s); (b) shows the same type of spectrogram with a wide window ($h=0.3$ s); (c) shows a continuous wavelet transform spectrogram which resolves both time and frequency well.

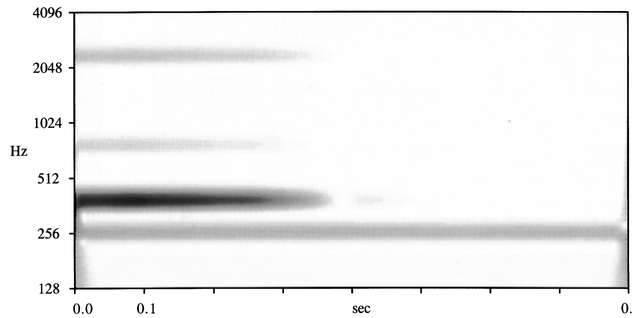

Fig. 2. The spectrogram of the sound of two tuning forks, one producing a 256-Hz signal, the other producing a 384-Hz signal; the second tuning fork was pulled away from the microphone after about 0.35 s.

sitive to time; h plays the same role as it does in the short-time Fourier transform described above.

See Fig. 1(c) for a spectrogram made using this transform which shows good frequency and time localization for the artificial signal shown in Fig. 1(a) and (b). Often, the scalogram is plotted with the frequency axis logarithmic, so that equal octaves in frequency occupy the same vertical width in the scalogram (an octave is an interval of frequency from a given frequency to twice that frequency, just as in music).

Here, we provide the spectrograms of several example signals using the above transform to give an idea of the usefulness of the method. These were computed using C++ by the authors; each pixel of the image represents a value of $Wf(t,\omega)$ for a certain (t,ω); the value of that integral was computed as a simple inner product. (The authors will provide pseudocode to any interested reader.)

Figure 1 shows spectrograms of an artificial signal. The signal consists of several tones; one tone is an 80-Hz signal which lasts from 1 to 4 s with a brief 0.2 s interruption at time 2 s. The other tones last from 1 to 7 s and have frequencies of 10 and 13 Hz, respectively. Figure 1(a) shows a short-time Fourier (Gabor) transform with a relatively narrow window width; time is well-localized but the two lower frequency tones are not resolved. Figure 1(b) shows the same type of transform but with a wider window; the two low frequencies are now resolved but now the interruption in the higher-frequency term is not resolved. [The values of h used in Fig. 1(a) and (b) are, respectively, 0.05 and 0.3 s.] Figure 1(c) shows a continuous wavelet transform where both time and frequency are well-localized. Note the vertical bars on the ends of the notes as they appear in the spectrograms reflect the sharp cut-off and cut-on of the tones—sharp edges in a signal imply higher-frequency content.

Figures 2–4 show several sound samples, each taken at 9000 samples per second for 0.8 s using Vernier's ULI A/D board and microphone.[1] Figure 2 shows the spectrogram of the sound of two tuning forks, one producing the note ''C'' or *do* (256 Hz), the other producing the note ''G'' or *so* (384 Hz); the second tuning fork was pulled away from the microphone after about 0.35 s. Figure 3 shows the spectrogram of a whistle that is rising in pitch. The breaks in the plotting are caused by changes in volume of the whistle, perhaps due to the process of reshaping the mouth during the whistle. Figure 4 shows the spectrogram of one of the authors singing the notes *do ré mi*.

There is a fundamental limitation on how well frequency and time can both be determined using this method. This amounts to a type of uncertainty principle inherent in both the Fourier transform and the wavelet transform. This theorem says that a signal and its Fourier transform cannot both have small support. More precisely, for a function $\psi(t)$ with

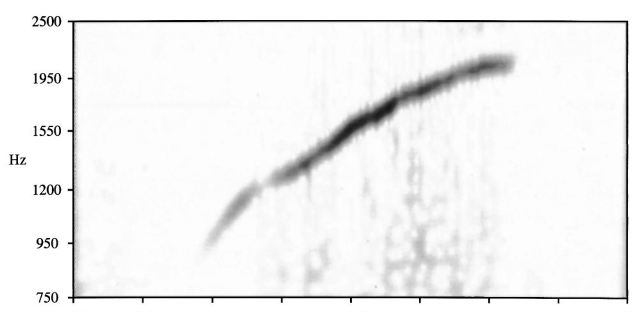

Fig. 3. The spectrogram of the sound of a whistle of rising pitch.

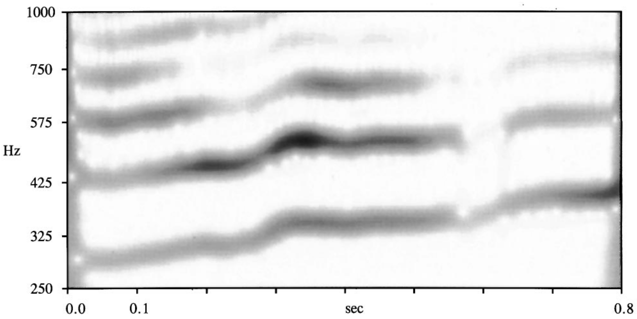

Fig. 4. The spectrogram of one of the authors singing the notes *do ré mi*. Harmonics corresponding to the different pitches can be clearly seen.

norm 1 ($\int|\psi(t)|^2\, dt=1$), define the center of the function to be the number $t'=\int t|\psi(t)|^2\, dt$, and define the width of the function to be

$$\Delta_\psi=\left(\int (t-t')^2|\psi(t)|^2\, dt\right)^{1/2}.$$

(The center of ψ is the expected value of $|\psi|^2$ in the sense of probability theory, and the width of ψ is the variation of $|\psi|^2$.) The uncertainty principle states $\Delta_\psi \Delta_{\hat{\psi}} \geqslant 1/2$. The product in this inequality reaches its minimum value (of 1/2) exactly when the signal is a Gaussian. This explains why the Morlet Gaussian wavelet works well at time and frequency localization. (See Ref. 2 for an elementary proof and discussion of the uncertainty principle in the context of wavelets; Ref. 3 is a general mathematical reference on continuous wavelet transforms.)

In quantum mechanics, the probability density for the position of a particle is given by $|\psi(\mathbf{r},t)|^2$, where the position wave function $\psi(\mathbf{r},t)$ is a function of position $\mathbf{r}$ and time t. The momentum probability density is then given by $|\Pi(\rho,t)|^2$ where the momentum waveform $\Pi(\rho,t)$ is the Fourier transform (over the space variable $\mathbf{r}$) of the position waveform. The uncertainty principle inherent in the Fourier transform then becomes the familiar statement that the momentum and position of a particle cannot both be determined to arbitrary precision. (See Ref. 4.)

The question of quantum uncertainty was treated by Wigner and led eventually in the 1940s to the Ville–Wigner distribution of time and frequency (Ref. 5 or Ref. 6), an early attempt to express frequency as a function of time. The distribution $Vf(t,\omega)$ for a signal $f(t)$ is intended to have the following properties.

(1) If $Vf(t,\omega)$ is the Wigner distribution for a signal $f(t)$, then $Vf(t-t_0,\omega)$ is the Wigner distribution for the signal $f(t-t_0)$; and $Vf(t,\omega-\omega_0)$ is the Wigner distribution for the signal $f(t)e^{-i\omega t}$ (the Wigner distribution is ''time-frequency shift covariant'').

(2) The Wigner distribution for the signal $f(t)=h^{-1/2}\times\exp(i\omega_0 t)g((t-t_0)/h)$ [where $g(t)=\pi^{-1/4}e^{-t^2/2}$] is $Vf(t,\omega)=2\exp((t-t_0)^2/h^2)\exp(-h^2(\omega-\omega_0))$. Here $f(t)$ represents a signal concentrated both in time near t_0 and in frequency near ω_0; the Wigner distribution is a Gaussian with respect to time and to frequency, concentrated at (t_0,ω_0).

(3) $1/2\pi\int_{-\infty}^{\infty} Vf(t,\omega)d\omega=|f(t)|^2$ [this serves as a Plancheral formula; the idea is that at each t, $Vf(t,\omega)$ is an instantaneous Fourier transform].

(4) $\int_{-\infty}^{\infty} Vf(t,\omega)dt=|\hat{f}(\omega)|^2$.

(5) $1/2\pi\int Vf(t,\omega)W_g(t,\omega)dtd\omega=|\int f(t)g^*(t)dt|^2$ (this is Moyal's formula, a sort of Parseval formula).

It turns out that the Wigner distribution is given by

$$Vf(t,\omega)=\int_{-\infty}^{\infty} f(t+\tau/2)f^*(t-\tau/2)e^{-i\omega\tau}\, d\tau;$$

the properties above are easy to verify. [For example, the integral in (3) follows from the fact that $1/2\pi\int\hat{h}(\omega)d\omega=h(0)$, where h is taken to be the function $h(\tau)=f(t+\tau/2)f^*(t-\tau/2)$.]

But the Wigner distribution has limitations for use in analyzing signals. These include expense [knowledge of the entire signal is required to compute $Vf(t,\omega)$ for each t, and there is no ''fast'' algorithm available], and the fact that a spectrogram based on the Wigner distribution [a plotting of $Vf(t,\omega)$ over the (t,ω) plane] will show ''interference'' artifacts [the sum of two Gaussian ''notes'' as in (2) above will result in a Wigner distribution spectrogram which shows ''noise'' in regions where there should be none].

It should be noted that other transforms or wavelets besides (1) or (2) may be chosen. Indeed, the discrete wavelet transform is widely used because of its algorithmic properties.

In discrete wavelet transforms, only specific frequencies are considered; typically, a wavelet ψ is chosen and instead of considering $\psi((\tau-t)\omega)$ for all (t,ω) in a certain rectangle, the discrete set of wavelets $\psi_{jk}(t)=\psi(2^{-j}t-k)$ is used (here j and k are integers). (So only the angular frequencies 2^{-j} are dealt with.) It is possible to choose ψ such that the resulting set forms an orthogonal basis of $L^2(\mathbf{R})$, and also such that the wavelet is nicely limited in time and in frequency. Then a function f can be expressed as a ''wavelet series'' $f=\Sigma_{jk}c_{jk}\psi_{jk}$, where $c_{jk}=\int f^*\psi_{jk}\, dx$, much like a Fourier series. For certain wavelets ψ, there are very simple, fast algorithms for computing the wavelet coefficients c_{jk}. These algorithms rely on the fact that each coefficient is always a linear combination of other coefficients, so integrals do not have to be computed for every coefficient; a wavelet ψ with smaller support will result in simpler relations among the coefficients, which makes algorithms simpler and faster.

The speed of these algorithms is a primary reason that wavelets have proven so valuable for applications. One key application is image compression. The coefficients of a Fourier series of a digital signal can be used to reconstruct the original signal. Since most of the coefficients are typically close to zero, they can be coded in a manner which takes less computer memory; this has proven to be a useful way to store images in a compact way (the popular JPEG format uses this technique). Because of their sensitivity to the details of an image, some discrete wavelet transforms have shown their value in image and video compression when employed in a similar manner; this is an area of active research and development.

An overview and discussion of wavelets and many of their applications and the question of which wavelet choice is appropriate for a particular application is given by Ref. 5. Other expository works include Refs. 7–9. Reference 10 gives a lucid account of the mathematical theory and development of discrete wavelet bases. Recent articles dealing with discrete wavelet bases include Refs. 11 and 12; the latter includes suggestions for other applications in physics. There is an extremely large literature on wavelets, involving many variations in their construction and application. The *Wavelet Digest* published on the World Wide Web (see Ref. 13) contains links to bibliographies and other information on wavelets; commercial software packages are available for programs such as MATHEMATICA and MATLAB for persons wishing to experiment with wavelets.

ACKNOWLEDGMENTS

The authors would like to express their gratitude for the very helpful criticisms of the referees.

[1]Vernier Software, 8565 S. W. Beaverton-Hillsdale Hwy., Portland, OR 97225-2429.

[2]Charles K. Chui, *An Introduction to Wavelets* (Academic, Boston, 1992), pp. 54–59.

[3]M. Holschneider, *Wavelets, An Analysis Tool* (Clarendon, Oxford, 1995).

[4]Albert Messiah, *Quantum Mechanics, Vol. 1*, translated by G. M. Temmer (North-Holland, Amsterdam, 1970), Vol. I, p. 50.

[5]Yves Meyer, *Wavelets: Algorithms and Applications* (translated and revised by Robert D. Ryan) (SIAM, Philadelphia, 1993).

[6]Leon Cohen, *Time-Frequency Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1995).

[7]Barbara Burke Hubbard, *The World According to Wavelets: The Story of a Mathematical Technique in the Making* (Peters, Wellesley, MA, 1998), 2nd ed.

[8]Gerald Kaiser, *A Friendly Guide to Wavelets* (Birkhäuser, Boston, 1994).

[9]Eugenio Hernandez and Guido L. Weiss, *A First Course in Wavelets (Studies in Advanced Mathematics)* (CRC, Boca Raton, FL, 1996).

[10]Ingrid Daubechies, *Ten Lectures on Wavelets, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 61* (SIAM, Philadelphia, 1992).

[11]O. V. Vasilyev, D. A. Yuen, and S. Paolucci, ''Solving PDEs Using Wavelets,'' Comput. Phys. **11** (5), 429–435 (1997).

[12]George D. J. Phillies, ''Wavelets: A new alternative to Fourier transforms,'' Comput. Phys. **10** (3), 247–252 (1996).

[13]The URL of the *Wavelet Digest* is http://www.wavelet.org/wavelet/index.html.