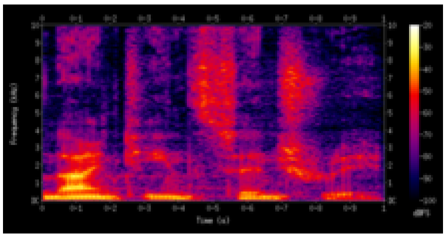
Spectrogram

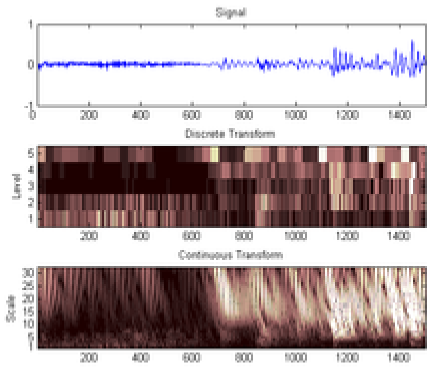

Spectrogram of the spoken words "nineteenth century". Frequencies are shown increasing up the vertical axis, and time on the horizontal axis. The legend to the right shows that the color intensity increases with the density.

A **spectrogram** is a visual representation of the spectrum of frequencies of a signal as it varies with time. When applied to an audio signal, spectrograms are sometimes called **sonographs**, **voiceprints**, or **voicegrams**. When the data is represented in a 3D plot they may be called **waterfalls**.

Spectrograms are used extensively in the fields of music, sonar, radar, and speech processing,[1] seismology, and others. Spectrograms of audio can be used to identify spoken words phonetically, and to analyse the various calls of animals.

A spectrogram can be generated by an optical spectrometer, a bank of band-pass filters, by Fourier transform or by a wavelet transform (in which case it is also known as a **scaleogram** or **scalogram**).[2]

A spectrogram is usually depicted as a heat map, i.e., as an image with the intensity shown by varying the colour or brightness.



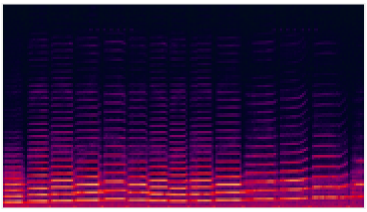
Scaleograms from the DWT and CWT for an audio sample

Contents

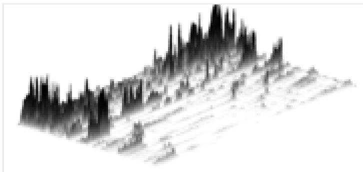
Format

A common format is a graph with two geometric dimensions: one axis represents time, and the other axis represents frequency; a third dimension indicating the amplitude of a particular frequency at a particular time is represented by the intensity or color of each point in the image.

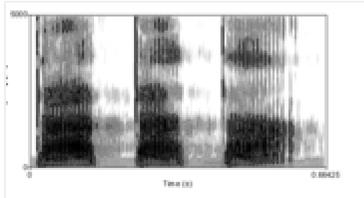
There are many variations of format: sometimes the vertical and horizontal axes are switched, so time runs up and down; sometimes as a waterfall plot where the amplitude is represented by height of a 3D surface instead of color or intensity. The frequency and amplitude axes can be either linear or logarithmic, depending on what the graph is being used for. Audio would usually be represented with a logarithmic amplitude axis (probably in decibels, or dB), and frequency would be linear to emphasize harmonic relationships, or logarithmic to emphasize musical, tonal relationships.



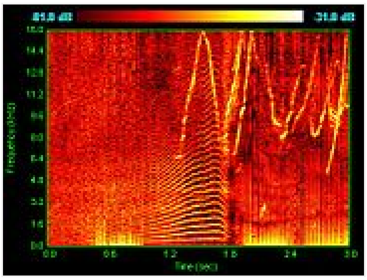
Spectrogram of this recording of a violin playing. Note the harmonics occurring at whole-number multiples of the fundamental frequency.



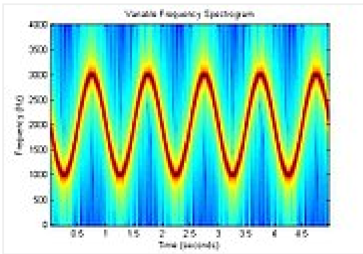
3D surface spectrogram of a part from a music piece.



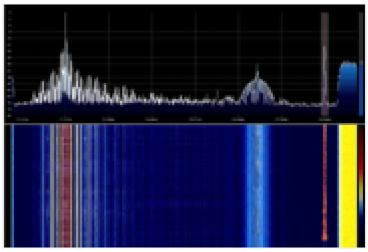
Spectrogram of a male voice saying 'ta ta ta'.



Spectrogram of dolphin vocalizations; chirps, clicks and harmonizing are visible as inverted Vs, vertical lines and horizontal striations respectively.

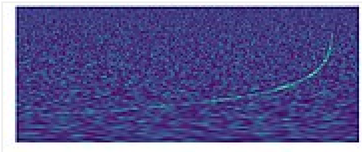


Spectrogram of an FM signal. In this case the signal frequency is modulated with a sinusoidal frequency vs. time profile.

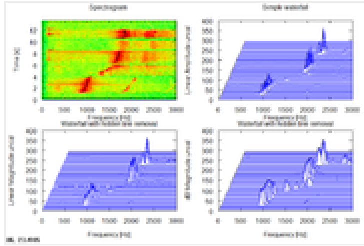


Spectrum above and waterfall (Spectrogram) below of an 8MHz wide PAL-I Television signal.

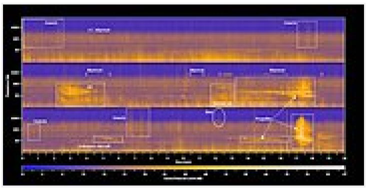
Spectrogram of great tit song.



Spectrogram of a gravitational wave (GW170817).



Spectrogram and waterfalls of 3 whistled notes.



Spectrogram of the soundscape ecology of Mount Rainier National Park, with the sounds of different creatures and aircraft highlighted

Generation

Spectrograms of light may be created directly using an optical spectrometer over time.

Spectrograms may be created from a time-domain signal in one of two ways: approximated as a filterbank that results from a series of band-pass filters (this was the only way before the advent of modern digital signal processing), or calculated from the time signal using the Fourier transform. These two methods actually form two different time–frequency representations, but are equivalent under some conditions.

The bandpass filters method usually uses analog processing to divide the input signal into frequency bands; the magnitude of each filter's output controls a transducer that records the spectrogram as an image on paper.[3]

Creating a spectrogram using the FFT is a digital process. Digitally sampled data, in the time domain, is broken up into chunks, which usually overlap, and Fourier transformed to calculate the magnitude of the frequency spectrum for each chunk. Each chunk then corresponds to a vertical line in the image; a measurement of magnitude versus frequency for a specific moment in time (the midpoint of the chunk). These spectrums or time plots are then "laid side by side" to form the image or a three-dimensional surface,[4] or slightly overlapped in various ways, i.e. windowing. This process essentially corresponds to computing the squared magnitude of the short-time Fourier transform (STFT) of the signal $s(t)$ — that is, for a window width ω, $\mathbf{spectrogram}(t,\omega) = |\mathbf{STFT}(t,\omega)|^2$.[5]

Limitations and resynthesis

From the formula above, it appears that a spectrogram contains no information about the exact, or even approximate, phase of the signal that it represents. For this reason, it is not possible to reverse the process and generate a copy of the original signal from a spectrogram, though in situations where the exact initial phase is unimportant it may be possible to generate a useful approximation of the original signal. The Analysis & Resynthesis Sound Spectrograph[6] is an example of a computer program that attempts to do this. The Pattern Playback was an early speech synthesizer, designed at Haskins Laboratories in the late 1940s, that converted pictures of the acoustic patterns of speech (spectrograms) back into sound.

In fact, there is some phase information in the spectrogram, but it appears in another form, as time delay (or group delay) which is the dual of the Instantaneous Frequency.

The size and shape of the analysis window can be varied. A smaller (shorter) window will produce more accurate results in timing, at the expense of precision of frequency representation. A larger (longer) window will provide a more precise frequency representation, at the expense of precision in timing representation. This is an instance of the Heisenberg uncertainty principle, that the product of the precision in two conjugate variables is greater than or equal to a constant (B*T>=1 in the usual notation).[7]

Applications

- Early analog spectrograms were applied to a wide range of areas including the study of bird calls (such as that of the great tit), with current research continuing using modern digital equipment[8] and applied to all animal sounds. Contemporary use of the digital spectrogram is especially useful for studying frequency modulation (FM) in animal calls. Specifically, the distinguishing characteristics of FM chirps, broadband clicks, and social harmonizing are most easily visualized with the spectrogram.
- Spectrograms are useful in assisting in overcoming speech deficits and in speech training for the portion of the population that is profoundly deaf[9]
- The studies of phonetics and speech synthesis are often facilitated through the use of spectrograms.[10][11]
- In deep learning-based speech synthesis, spectrogram (or spectrogram in mel scale) is first predicted by a seq2seq model, then the spectrogram is fed to a neural vocoder to derive the synthesized raw waveform.
- By reversing the process of producing a spectrogram, it is possible to create a signal whose spectrogram is an arbitrary image. This technique can be used to hide a picture in a piece of audio and has been employed by several electronic music artists.[12] See also steganography.
- Some modern music is created using spectrograms as an intermediate medium; changing the intensity of different frequencies over time, or even creating new ones, by drawing them and then inverse transforming. See Audio timescale-pitch modification and Phase vocoder.
- Spectrograms can be used to analyze the results of passing a test signal through a signal processor such as a filter in order to check its performance.[13]
- High definition spectrograms are used in the development of RF and microwave systems[14]
- Spectrograms are now used to display scattering parameters measured with vector network analyzers[15]
- The US Geological Survey and the IRIS Consortium provide near real-time spectrogram displays for monitoring seismic stations[16][17]
- Spectrograms can be used with recurrent neural networks for speech recognition.[18]

See also

- Acoustic signature
- Chromagram
- Generalized spectrogram
- List of unexplained sounds
- Reassignment method
- Short-time Fourier transform
- Spectral music
- Spectrometer
- Spectrum
- Strobe tuner
- Time-frequency representation
- Waterfall plot
- Wavelet transform

References

1. JL Flanagan, Speech Analysis, Synthesis and Perception, Springer- Verlag, New York, 1972
2. Sejdic, E.; Djurovic, I.; Stankovic, L. (August 2008). "Quantitative Performance Analysis of Scalogram as Instantaneous Frequency Estimator". *IEEE Transactions on Signal Processing*. **56** (8): 3837–3845. doi:10.1109/TSP.2008.924856 (https://doi.org/10.1109%2FTSP.2008.924856). ISSN 1053-587X (https://www.worldcat.org/issn/1053-587X).
3. "Spectrograph" (https://www.sfu.ca/sonic-studio/handbook/Spectrograph.html). *www.sfu.ca*. Retrieved 7 April 2018.
4. "Spectrograms" (https://ccrma.stanford.edu/~jos/mdft/Spectrograms.html). *ccrma.stanford.edu*. Retrieved 7 April 2018.
5. "STFT Spectrograms VI – NI LabVIEW 8.6 Help" (http://zone.ni.com/reference/en-XX/help/371361E-01/lvanls/stft_spectrogram_core/#details). *zone.ni.com*. Retrieved 7 April 2018.
6. "The Analysis & Resynthesis Sound Spectrograph" (http://arss.sourceforge.net). *arss.sourceforge.net*. Retrieved 7 April 2018.
7. http://fourier.eng.hmc.edu/e161/lectures/fourier/node2.html
8. "BIRD SONGS AND CALLS WITH SPECTROGRAMS (SONOGRAMS) OF SOUTHERN TUSCANY (Toscana – Italy)" (http://www.birdsongs.it/index.asp). *www.birdsongs.it*. Retrieved 7 April 2018.
9. Saunders, Frank A.; Hill, William A.; Franklin, Barbara (1 December 1981). "A wearable tactile sensory aid for profoundly deaf children". *Journal of Medical Systems*. **5** (4): 265–270. doi:10.1007/BF02222144 (https://doi.org/10.1007%2FBF02222144). PMID 7320662 (https://pubmed.ncbi.nlm.nih.gov/7320662).
10. "Spectrogram Reading" (https://web.archive.org/web/19990427185722/http://cslu.cse.ogi.edu/tutordemos/SpectrogramReading/spectrogram_reading.html). *ogi.edu*. Archived from the original (http://cslu.cse.ogi.edu/tutordemos/SpectrogramReading/spectrogram_reading.html) on 27 April 1999. Retrieved 7 April 2018.
11. "Praat: doing Phonetics by Computer" (http://www.fon.hum.uva.nl/praat/). *www.fon.hum.uva.nl*. Retrieved 7 April 2018.
12. "The Aphex Face – bastwood" (http://www.bastwood.com/aphex.php). *www.bastwood.com*. Retrieved 7 April 2018.
13. "SRC Comparisons" (http://src.infinitewave.ca). *src.infinitewave.ca*. Retrieved 7 April 2018.
14. "constantwave.com – constantwave Resources and Information" (http://www.constantwave.com/gallery.aspx). *www.constantwave.com*. Retrieved 7 April 2018.
15. "Spectrograms for vector network analyzers" (https://web.archive.org/web/20120810020043/http://www.constantwave.com/spectro_vna.aspx). Archived from the original (http://www.constantwave.com/spectro_vna.aspx) on 2012-08-10.
16. "Real-time Spectrogram Displays" (https://earthquake.usgs.gov/monitoring/spectrograms/24hr/). *earthquake.usgs.gov*. Retrieved 7 April 2018.

17. "IRIS: MUSTANG: Noise-Spectrogram: Docs: v. 1: Help" (https://service.iris.edu/mustang/noise-spectrogram/docs/1/help/).
18. Geitgey, Adam (2016-12-24). "Machine Learning is Fun Part 6: How to do Speech Recognition with Deep Learning" (https://medium.com/@ageitgey/machine-learning-is-fun-part-6-how-to-do-speech-recognition-with-deep-learning-28293c162f7a). *Medium*. Retrieved 2018-03-21.

External links

- See an online spectrogram of speech or other sounds captured by your computer's microphone. (https://auditoryneuroscience.com/acoustics/spectrogram)
- Generating a tone sequence whose spectrogram matches an arbitrary text, online (http://www.audiocheck.net/audiocheck_spectrotyper.php)
- Further information on creating a signal whose spectrogram is an arbitrary image (https://web.archive.org/web/20110725231858/http://devrand.org/show_item.html?item=64&page=Project)
- Article describing the development of a software spectrogram (http://kdenlive.org/users/granjow/introducing-scopes-audio-spectrum-and-spectrogram)
- History of spectrograms & development of instrumentation (http://www.spectrogramsforspeech.com/background/history-of-spectrograms/)
- How to identify the words in a spectrogram (http://home.cc.umanitoba.ca/~robh/howto.html) from a linguistic professor's *Monthly Mystery Spectrogram* publication.