

WIKIPEDIA

Digital signal processing

Digital signal processing (DSP) is the use of digital processing, such as by computers or more specialized digital signal processors, to perform a wide variety of signal processing operations. The digital signals processed in this manner are a sequence of numbers that represent samples of a continuous variable in a domain such as time, space, or frequency. In digital electronics, a digital signal is represented as a pulse train,^{[1][2]} which is typically generated by the switching of a transistor.^[3]

Digital signal processing and analog signal processing are subfields of signal processing. DSP applications include audio and speech processing, sonar, radar and other sensor array processing, spectral density estimation, statistical signal processing, digital image processing, data compression, video coding, audio coding, image compression, signal processing for telecommunications, control systems, biomedical engineering, and seismology, among others.

DSP can involve linear or nonlinear operations. Nonlinear signal processing is closely related to nonlinear system identification^[4] and can be implemented in the time, frequency, and spatio-temporal domains.

The application of digital computation to signal processing allows for many advantages over analog processing in many applications, such as error detection and correction in transmission as well as data compression.^[5] Digital signal processing is also fundamental to digital technology, such as digital telecommunication and wireless communications.^[6] DSP is applicable to both streaming data and static (stored) data.

Contents

Signal sampling

Domains

- Time and space domains

- Frequency domain

- Z-plane analysis

- Autoregression analysis

- Time-frequency analysis

- Wavelet

- Empirical mode decomposition

Implementation

Applications

Techniques

Related fields

References

Further reading

Signal sampling

To digitally analyze and manipulate an analog signal, it must be digitized with an analog-to-digital converter (ADC).^[7] Sampling is usually carried out in two stages, discretization and quantization. Discretization means that the signal is divided into equal intervals of time, and each interval is represented by a single measurement of amplitude. Quantization means each amplitude measurement is approximated by a value from a finite set. Rounding real numbers to integers is an example.

The Nyquist–Shannon sampling theorem states that a signal can be exactly reconstructed from its samples if the sampling frequency is greater than twice the highest frequency component in the signal. In practice, the sampling frequency is often significantly higher than this.^[8]

Theoretical DSP analyses and derivations are typically performed on discrete-time signal models with no amplitude inaccuracies (quantization error), "created" by the abstract process of sampling. Numerical methods require a quantized signal, such as those produced by an ADC. The processed result might be a frequency spectrum or a set of statistics. But often it is another quantized signal that is converted back to analog form by a digital-to-analog converter (DAC).

Domains

In DSP, engineers usually study digital signals in one of the following domains: time domain (one-dimensional signals), spatial domain (multidimensional signals), frequency domain, and wavelet domains. They choose the domain in which to process a signal by making an informed assumption (or by trying different possibilities) as to which domain best represents the essential characteristics of the signal and the processing to be applied to it. A sequence of samples from a measuring device produces a temporal or spatial domain representation, whereas a discrete Fourier transform produces the frequency domain representation.

Time and space domains

Time domain refers to the analysis of signals with respect to time. Similarly, space domain refers to the analysis of signals with respect to position, e.g., pixel location for the case of image processing.

The most common processing approach in the time or space domain is enhancement of the input signal through a method called filtering. Digital filtering generally consists of some linear transformation of a number of surrounding samples around the current sample of the input or output signal. The surrounding samples may be identified with respect to time or space. The output of a linear digital filter to any given input may be calculated by convolving the input signal with an impulse response.

Frequency domain

Signals are converted from time or space domain to the frequency domain usually through use of the Fourier transform. The Fourier transform converts the time or space information to a magnitude and phase component of each frequency. With some applications, how the phase varies with frequency can be a significant consideration. Where phase is unimportant, often the Fourier transform is converted to the power spectrum, which is the magnitude of each frequency component squared.

The most common purpose for analysis of signals in the frequency domain is analysis of signal properties. The engineer can study the spectrum to determine which frequencies are present in the input signal and which are missing. Frequency domain analysis is also called *spectrum-* or *spectral analysis*.

Filtering, particularly in non-realtime work can also be achieved in the frequency domain, applying the filter and then converting back to the time domain. This can be an efficient implementation and can give essentially any filter response including excellent approximations to brickwall filters.

There are some commonly used frequency domain transformations. For example, the cepstrum converts a signal to the frequency domain through Fourier transform, takes the logarithm, then applies another Fourier transform. This emphasizes the harmonic structure of the original spectrum.

Z-plane analysis

Digital filters come in both IIR and FIR types. Whereas FIR filters are always stable, IIR filters have feedback loops that may become unstable and oscillate. The Z-transform provides a tool for analyzing stability issues of digital IIR filters. It is analogous to the Laplace transform, which is used to design and analyze analog IIR filters.

Autoregression analysis

A signal is represented as linear combination of its previous samples. Coefficients of the combination are called autoregression coefficients. This method has higher frequency resolution and can process shorter signals compared to the Fourier transform.^[9] Prony's method can be used to estimate phases, amplitudes, initial phases and decays of the components of signal.^{[10][9]} Components are assumed to be complex decaying exponents.^{[10][9]}

Time-frequency analysis

A time-frequency representation of signal can capture both temporal evolution and frequency structure of analyzed signal. Temporal and frequency resolution are limited by the principle of uncertainty and the tradeoff is adjusted by the width of analysis window. Linear techniques such as Short-time Fourier transform, wavelet transform, filter bank,^[11] non-linear (e.g., Wigner — Ville transform^[10]) and autoregressive methods (e.g. segmented Prony method)^{[10][12][13]} are used for representation of signal on the time-frequency plane. Non-linear and segmented Prony methods can provide higher resolution, but may produce undesirable artefacts. Time-frequency analysis is usually used for analysis of non-stationary signals. For example, methods of fundamental frequency estimation, such as RAPT and PEFAC^[14] are based on windowed spectral analysis.

Wavelet

In numerical analysis and functional analysis, a discrete wavelet transform is any wavelet transform for which the wavelets are discretely sampled. As with other wavelet transforms, a key advantage it has over Fourier transforms is temporal resolution: it captures both frequency *and* location information. The accuracy of the joint time-frequency resolution is limited by the uncertainty principle of time-frequency.

Empirical mode decomposition

Empirical mode decomposition is based on decomposition signal into intrinsic mode functions (IMF). IMFs are quasiharmonical oscillations that are extracted from the signal.^[15]

Implementation

DSP algorithms may be run on general-purpose computers and digital signal processors. DSP algorithms are also implemented on purpose-built hardware such as application-specific integrated circuit (ASICs). Additional technologies for digital signal processing include more powerful general purpose microprocessors, graphics processing units, field-programmable gate arrays (FPGAs), digital signal controllers (mostly for industrial applications such as motor control), and stream processors.^[16]

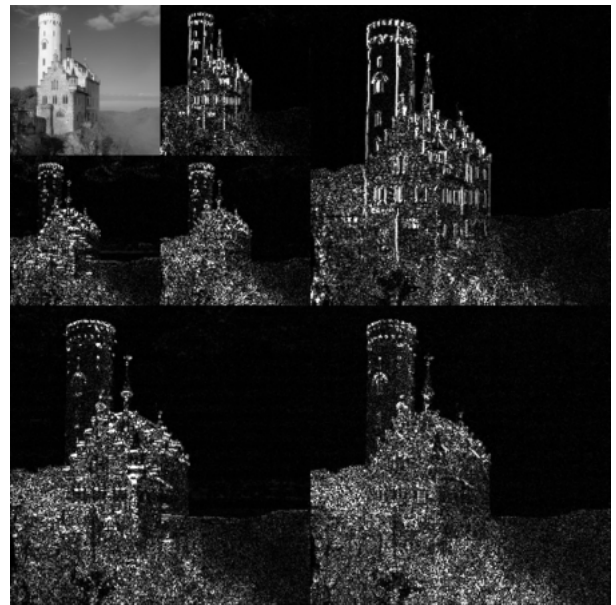
For systems that do not have a real-time computing requirement and the signal data (either input or output) exists in data files, processing may be done economically with a general-purpose computer. This is essentially no different from any other data processing, except DSP mathematical techniques (such as the DCT and FFT) are used, and the sampled data is usually assumed to be uniformly sampled in time or space. An example of such an application is processing digital photographs with software such as Photoshop.

When the application requirement is real-time, DSP is often implemented using specialized or dedicated processors or microprocessors, sometimes using multiple processors or multiple processing cores. These may process data using fixed-point arithmetic or floating point. For more demanding applications FPGAs may be used.^[17] For the most demanding applications or high-volume products, ASICs might be designed specifically for the application.

Applications

General application areas for DSP include

- Audio signal processing
- Audio data compression e.g. MP3
- Video data compression
- Computer graphics
- Digital image processing
- Photo manipulation
- Speech processing
- Speech recognition
- Data transmission
- Radar
- Sonar
- Financial signal processing
- Economic forecasting
- Seismology
- Biomedicine
- Weather forecasting



An example of the 2D discrete wavelet transform that is used in JPEG2000. The original image is high-pass filtered, yielding the three large images, each describing local changes in brightness (details) in the original image. It is then low-pass filtered and downsampled, yielding an approximation image; this image is high-pass filtered to produce the three smaller detail images, and low-pass filtered to produce the final approximation image in the upper-left.