Signal Processing

Prince Singh[1], Erik Forseth[1], Oğuz Tanrikulu[2], Ed Tricker[3]

Abstract
Signal processing is a subject whose importance and potential may sometimes be overlooked. We often hear about advances in familiar domains such as computing, communications, and artificial intelligence, but it is signal processing which lies at the heart of these fields, and which facilitates many other cutting-edge research endeavors and everyday technologies. The objective of this article is to shed light on this discipline by touching upon the historical developments, providing a qualitative overview of the techniques involved, and elaborating on relevant practical applications. The article concludes with a light discussion of the applications of signal processing to systematic trading, where it is uniquely well-suited to the analysis of financial time series.

Keywords
Signal Processing; Machine learning; Systematic Trading

[1] *Quantitative Research Analyst*
[2] *Director, Quantitative Research*
[3] *Chief Investment Officer of Quantitative Strategies*

1. Introduction

Signal Processing is the science behind our digital life.

– IEEE Signal Processing Society.

Signal processing (SP) is a branch of electrical engineering that plays an indispensable role in powering our modern (digital) world—enabling nearly all of the technologies (e.g., radios, computers, videos, cellular phones) that we use and rely on in our everyday lives. SP played a central role in the *Digital Revolution*[1] which marked the onset of today's constantly-evolving *Information Age*, highlighting the importance of this discipline and its necessity in fundamental scientific advances. An unknown field to many, the term "Signal Processing" is often misconstrued. SP does not refer to the *transmission* of signals[2] via telephone lines or via radio waves, but rather SP refers to the set of mathematical techniques and algorithms developed for analyzing and altering signals to meet task-driven applications (e.g., improving signal quality, or capturing information in a measured signal). SP judiciously interacts with three domains, illustrated in Figure 1, to facilitate the data acquisition to interpretation process.

The development of modern (digital) SP essentially began during World War II, when a number of researchers contributed towards a mathematical theory of signals and noise, notably *Norbert Wiener*[3]. Prompted by Wiener's emphasis on the statistical nature of communication, in 1948 *Claude E. Shannon*[4] made landmark contributions for quantifying the reliable transmission of information over imperfect communication channels (like phone lines or wireless networks). However, it was not until the development of integrated-circuit technologies—interconnected electronic components on a single wafer of silicon or other semiconductor—and the subsequent proliferation of computers in the 1960s and onwards, that Shannon's work became widespread and influenced a generation of communication engineers.

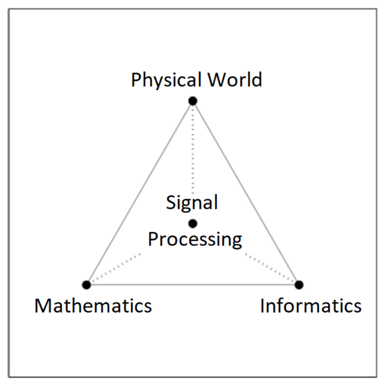

Figure 1. SP is at the intersection of three main domains: 1) signals are produced from the *Physical* world, 2) the performance of perceived signals are evaluated via *Mathematics*, and 3) the associated algorithms are efficiently implemented via *Informatics*

[1] a.k.a. the Third Industrial Revolution: 1960s-1970s

[2] Functions of time (represented as waves) that convey information about some phenomenon that can be stored in a digitized format, e.g., sound, images, sensor measurements, electrocardiogram (EKG) recordings, textual data.

[3] An MIT faculty member who worked on the theory of Brownian motion, prediction for stationary time-series, and became known in the wider scientific community for *Cybernetics*—the scientific study of how humans, animals and machines control and communicate with each other.

[4] Another MIT faculty member, who conceived and laid the foundations for *Information theory*. Shannon made contributions to the problem of electrical switching (i.e. logically manipulating binary digits: 0's and 1's)—the nervous system of digital computing. His work has also been fundamental in developing the electronic communications networks that now envelop the Earth.

The colossal increase in computing power since the 1960s [Fig. 2] has rendered SP applicable to a variety of economically-relevant fields. Applications include wireless communication (phones, radars, satellites), health (e.g. localizing epileptic sources in the brain, echography, Magnetic-Resonance Imaging), transportation systems (assimilating data from noisy sensors deployed in robots, self-driving vehicles, aviation, smart grids, smart cities), and finance (forecasting the movements of asset prices, evolving financial portfolios).

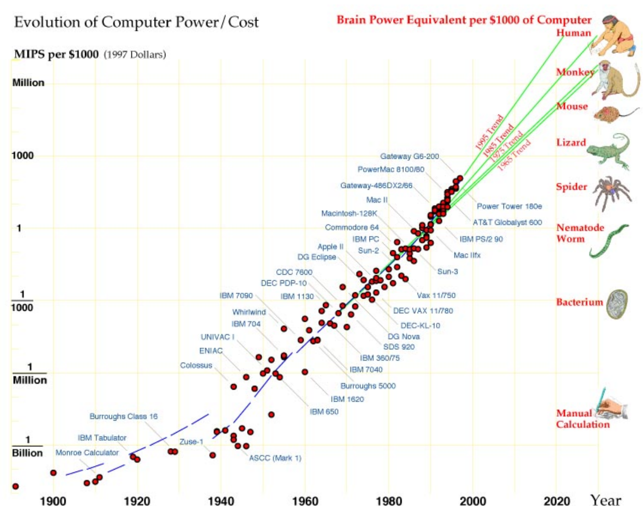

Figure 2. The evolution of computational power-to-cost ratio: the power-per-cost of computing technologies has been steadily increasing by a factor of about 1000 every 20 years. [Bostrom (2003)]

The remainder of this article is organized as follows. We first present the essentials ingredients of SP, from which some of the widely used modern SP methods have been built. Next, we review the prevailing techniques and present the results of a simulation which illustrates a practically-meaningful application. We conclude with a brief discussion of the implications for investment management.

2. Elements of Signal Processing (SP)

Here we qualitatively discuss a couple of the basic ingredients and mathematical preliminaries for SP.

The Fourier Transform

The roots of SP arguably begin with *Joseph Fourier*. Fourier proposed a set of mathematical techniques—including the *Fourier Transform* (FT)—for representing and working with signals in the *frequency-domain*. That is, he developed a way to decompose signals into mixtures of fundamental, periodic components, each of which oscillates at some fixed rate (or frequency). This representation allows for a simplification of complicated mathematics, and greatly facilitates the understanding of many intricate phenomena arising in physics and engineering. However, it was not until 1965 that the FT became computationally tractable for most practical tasks, when *James Cooley* and *John Tuckey* proposed the *Fast Fourier Transform* (FFT)[5]—an algorithm that immensely reduced the calculation cost of the FT. The FFT algorithm was an opportune development, as mass production of integrated circuits had recently begun. Fourier methods became ubiquitous thereafter. This is an example of the concurrent development of theory and rapid growth in computing power, which enabled many previously unimaginable feats. Examples include the live television broadcast of the first steps on the moon in 1969, production of the Computerized Tomography scanning (CT scan) device in 1971, and the development of the celebrated *Kalman Filter*, which solved problems in missile and aerospace tracking and guidance, radar, sonar, etc.

Nyquist Sampling Theorem (NST)

Another cornerstone of SP is the Nyquist Sampling Theorem (NST), which establishes a fundamental bridge between *physically-derived* continuous-time signals (referred to as "analog signals") and *computationally-tractable* discrete-time signals (referred to as "digital signals"). The importance of this fundamental connection is hard to overstate, especially because manipulating *digitized* signals is much faster and more efficient compared with operations on traditional *analog* signals. It was known that an analog signal could be re-constructed from a finite digitized representation when the analog signal is effectively *band-limited*; that is, when it does not contain certain frequency components [Nyquist (1928)]. However, it remained to be shown that the analog signal could be constructed *perfectly* (i.e., without any loss of information) and *uniquely* from the digitized counterparts. This gap and other fundamental principles of Information Theory were established in [Shannon (1948a,b, 1949)].

3. Signal Processing (SP) Techniques

The goal here is threefold: 1) to present an overview of the widely used SP techniques, 2) to discuss a practically-meaningful application of the methods introduced, and 3) to discuss ties between SP and another area of active research, Machine Learning (ML).

Filters

A *filter* originally referred to a physical device that selected certain frequencies (or a range of frequencies) from an analog signal while suppressing others. However, upon the advent of the digital-era in the 1960s, the term *digital-filter* came to refer to any of a class of computer algorithms which perform mathematical operations on digital (discrete-time) signals in order to meet user-defined signal specifications. Many digital filters employ the efficient FFT algorithm (discussed in the

[5]The FFT was rated, by the IEEE society, to be one of the top 10 algorithms developed in the 20th century.

previous section) in order to identify the *frequency spectrum*[6] of a signal, which can then be manipulated in various ways. There also exists a broad class of algorithms which fall under the umbrella of *adaptive filtering*, which can be used for applications such as *system identification* and *control*. These methods are extremely useful for inferring the properties of signals which are corrupted by noise and/or which are time-varying.

The design and implementation of digital filters poses many practical challenges, and they continue to be a topic of active research. Their importance is apparent from their omnipresence in everything from common electronics to cutting-edge AI technologies.

Denoising

In the real world, physical signals are always corrupted by some amount of noise. An important application of SP involves *denoising* applications—attenuation of noise in order to reveal some "true" underlying information or dynamics. Denoising is vital in all manner of applications, from cell phone communication to scientific experiments. For example, 2015 saw the first detection of a gravitational wave signal produced by the merger of a pair of black holes. This was a landmark event, confirming predictions that Einstein had made a century prior and inaugurating a new era in observational astronomy and astrophysics. These detections are extremely subtle, and would not be remotely possible without the application of denoising and template-matching techniques from SP.

Some denoising operations can be quite simple, including *smoothing* operations[7]. These correspond to very simple *low-pass filters*, which block the high-frequency components of a signal while letting the low-frequency content "pass" with little or no modification. More sophisticated denoising approaches include *energy-transfer* filters that move undesired noise components into (or split them across) various frequency regimes, or the widely used *Kalman Filter* which facilitated trajectory estimation for the Apollo program.

Prediction

Many of the applications described before have been concerned with cleaning or transforming data in useful ways. However, SP methods have also been developed for *prediction* problems. Indeed, the field of *Adaptive Signal Processing* [Haykin (2013)] is concerned with the development of digital filters with predictive capabilities. In this case the filter is a recursive algorithm with a feedback loop which allows it to learn (in a sequential fashion) from data in order to minimize the error between the filter's output and some specified target. There is virtually no difference between these kinds of models and what's now referred to as *machine learning* (ML), except that ML commonly refers to a set of newer methods which have come into favor in recent decades for various reasons.

[6]The representation of a signal waveform as a (possibly infinite) sum of periodic (sinusoidal) functions—each with different magnitude and frequency.

[7]Smoothing refers to taking a group of adjacent points in the original data and performing an *averaging* procedure, thereby eliminating unimportant high-frequecy artifacts and capturing the important patterns in the data.

A Simple Experiment

Let's consider a simple simulation to make some of these ideas concrete. We'll apply the aforementioned techniques to denoise a *chirp* signal, which has important applications in sonar, radar, and spread-spectrum communications.

At a discrete time-step k with frequency $f(k)$, a chirp signal is defined as the superposition of a periodic (cosine) signal $\cos(f(k))$ and a white noise signal $w(k)$–$\cos(f(k)) + w(k)$–such that the two signals are uncorrelated (i.e., lack a deterministic relationship) with one another. The goal here is to *predict* the underlying clean periodic signal at the subsequent time-step $\cos(f(k+1))$ (i.e., first subplot of Figure 3) from the noisy corrupted signal (i.e., middle subplot of Figure 3).

The underlying prediction problem is exacerbated by the fact that the time-varying frequency $f(k)$ is unknown. Nevertheless, the ASP scheme touched upon in the previous subsection can be applied to the raw noisy data in order to reliably predict the subsequent values of the periodic signal with reduced noise levels, as illustrated in the last subplot of Figure 3. Further, it is crucial to note that the success of this procedure rests upon the fact that the *coherent* periodic (cosine) signal we are trying to predict is uncorrelated to the white noise signal that we aim to cancel off.

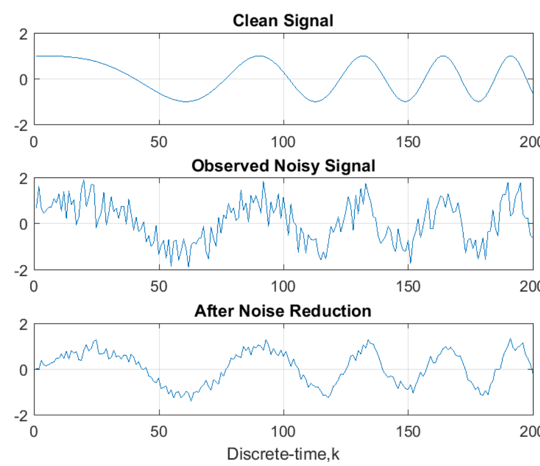

Figure 3. Denoising the chirp signal.

Relationship to Machine Learning (ML)

Machine Learning (ML) describes the computerized implementation of statistical predictive modeling techniques for inferring relationships in data. The increasing availability of cheap computational power has lead to ML enjoying immense success in a variety of crucial applications, among them credit card fraud detection, the control of autonomous robots (cars, drones), stock market analysis, and tumor detection.

Recently, many ML techniques have been applied to various problems in SP, and vice versa, blurring the lines between the two disciplines and enabling exciting applications. The two following examples give a flavor of the limitless possibilities:

- *Re-creating hand movements from imagination (to assist people with limited mobility):* SP de-noises brain signals and analyzes patterns in these signals. ML then attempts to distinguish different signal patterns (e.g., if the person is imagining vs. simply resting) and provide commands to the actuating device.

- *Real-time automatic speech recognition and translation (to bridge language barriers):* SP extracts relevant patterns in audio signals from a sender, then ML recognizes the patterns (using models built on historical data and experience) and makes appropriate (language-specific) recommendations for the recipient.

These examples illustrate a common two-step process: 1) SP applies rigorous techniques to identify meaningful information (or *features*) in signals (i.e., a process referred to as *feature extraction*), and 2) ML algorithms process these information-rich features in order to make forecasts or decisions about unseen data. It is worth noting that the quality of features often largely dictates the success of an ML algorithm; even state-of-the-art modeling techniques may be unable to make up for shallow or irrelevant input features.

Of course, the relationship between SP and ML is even deeper than that illustrated by this generic pipeline. Many of the mathematical techniques originally developed for SP have found important applications in ML, and the recent explosion of popularity in ML has lead to fundamental research with implications for SP.

4. Applications to Systematic Trading

Financial time series form an interesting subset of time series datasets, and pose a number of special challenges for practitioners hoping to investigate or model their behavior. They are typically expensive to acquire, and tend to be relatively small (there are many fewer data points of daily S&P closing prices than there are, say, images of cats on the internet). Moreover, they're *messy*. They often suffer from small signal-to-noise ratios, and are largely nonstationary—that is, the generating distributions of the datasets are not constant through time. These limitations imply that great care must be taken when applying various prediction techniques, e.g., traditional time-series forecasting tools and ML algorithms, as there is an increased risk of overfitting to random noise and spurious correlations. Furthermore, the majority of ML algorithms are not inherently designed to cope with sequential data. They can be used for time series prediction when the effects are stationary, and may even be useful in non-stationary settings when applied in an appropriate rolling fashion, but—with some exceptions—they typically do not take advantage of any ordering of the data.

It should therefore come as no surprise that SP methods comprise an extremely useful set of tools for this domain. As we've seen, they are naturally suited to handling sequential data, especially very noisy sequences. They provide approaches for transforming and representing time series in enlightening ways. Adaptive SP methods have been developed specifically for time series with non-stationarities. Finally, just like other statistical learning algorithms, prediction models in SP can be *regularized* to prevent overfitting. Of course, we are not proposing that these tools offer any kind of magic solution. Rather, we simply argue that they are at least as useful as more fashionable techniques, and often underappreciated by the investing public.

References

N. Bostrom. When machines outsmart humans. *Futures*, 35 (7):759–764, 2003.

S. O. Haykin. *Adaptive Filter Theory*. Prentice-Hall, fifth edition, 2013.

H. Nyquist. Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, 47(2):617–644, 1928.

C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948a.

C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(4):623–656, 1948b.

C. E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, 1949.

Legal Disclaimer