1/C Kelly Nelan

A Study of Shannon's Sampling Theory

12 Week Paper

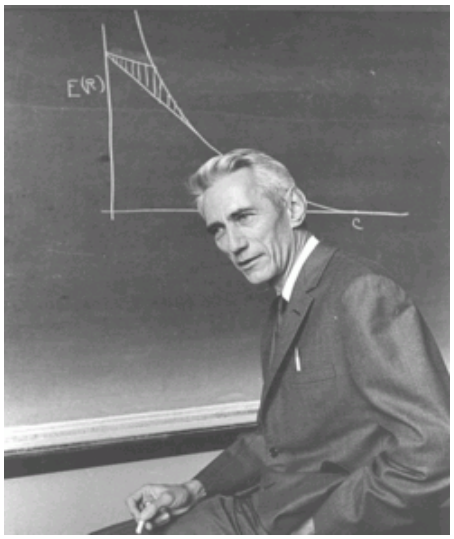
SM472

16FEB2007

In this paper I will explain the components that make up Sampling Theory, the man behind the theory, and the different types of sampling that make communication possible. Sampling Theory is known as one of the most important mathematical techniques used in communication engineering and information theory. It has become very useful in physics and engineering, such as in signal analysis, image processing, radar, sonar, acoustics, optics, holography, meteorology, oceanography, crystallography, physical chemistry, medical imaging and many more. Sampling Theory can be useful in any manner in which functions need to be reconstructed from sampled data. [1]

Historical Information on C.E. Shannon

The mathematical minds behind sampling theory can be traced back to many great mathematicians, such as Poisson, Borel, Hadamard, and E.T. Whittaker. Although it is unknown who originally discovered Sampling Theory, we do know that the results from these older mathematicians were rediscovered by C. E. Shannon in 1940 by using information theory and communication engineering. [1]



(http://www.windoweb.it/edpstory_new/foto_storia_inventori/ep_shannon_f1.gif)

Claude Elwood Shannon is commonly known as the "founding father of the electronic communications age." He was born in Gaylord, Michigan on April 30, 1916. During the first sixteen years of his life, Shannon was always mechanically inclined, building model planes, a radio-controlled model boat and a telegraphy system constructed of two barbed wires around a nearby pasture. He also earned money fixing radios for a local department store. [3]

In 1932 he entered the University of Michigan and earned the degrees of Bachelor of Science in Electrical Engineering and Bachelor of Science in Mathematics. His interest in these two subjects continued throughout the remainder of his life. After graduation he got the position of research assistant in the Department of Electrical Engineering at the Massachusetts Institute of Technology. His work at the time specialized in the Bush differential analyzer, the most advanced calculating machine at that time. It solved by analog means differential equations of up to the sixth degree. The machine was so large and complex that at some times, up to four assistants would be needed to crank in the functions. [3]

Shannon furthered his career by joining AT&T Bell Telephones in New Jersey in 1941 and stayed there until 1972 as a research mathematician. It is here that he made his many advances in information technology and communication systems. Shannon's Information Theory showed that the basic elements of any general communications system include [4]:

1. a transmitting device that tranforms the information or message into a suitable form for transmission over a medium

2. the medium over which the message is transmitted
3. a receiving device which decodes the message back into some form of the original signal
4. the destination or recipient of the message
5. a source of noise from either interference or distortion

Shannon also noted many important quantities yielded from the generalized communication system, including [4]:

1. the rate at which information is produced at the source
2. the capacity of the channel for handling information
3. the average amount of information in a message

It is from this information that Shannon developed his sampling theory and applied it to information theory. Shannon was looking at how the quantities stated above could transform the original message, and he explored in what ways the message would be impossible to reconstruct.

The Sampling Theory states that

If a function of time is limited to the band from 0 to W cycles per second, it is completely determined by giving its ordinates at a series of discrete points spaced 1/2W seconds apart in the manner indicated by the following result: If f(t) has no frequencies over W cycles per second, then

$$f(t) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2W}\right)\frac{\sin \pi(2Wt - n)}{\pi(2Wt - n))} \qquad [1]$$

In other words, this means that it is possible to reconstruct a signal from samples if the signal is band-limited and the sampling frequency is greater than twice the frequency of the original signal. One important principle for this theorem is to note that all the information contained in a signal is also contained in the sampled values taken at equidistantly spaced instances. The minimum rate at which the signal needs to be

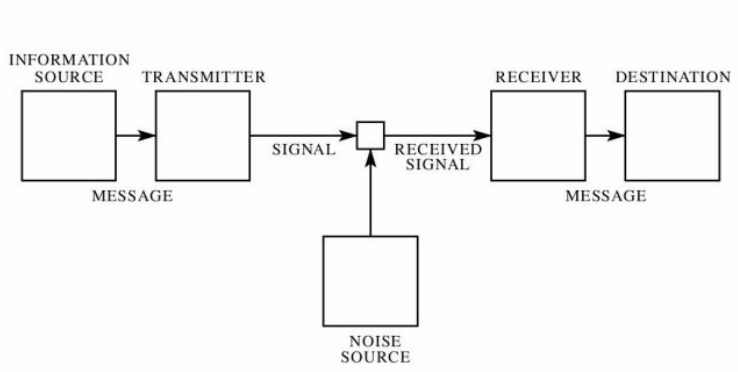
sampled in order to reconstruct the simple depends upon the knowledge of the frequency bound. This minimum rate is also known as the Nyquist rate, named after H. Nyquist, who was the first person to point out the importance of having a minimum rate in connection with telegraphy. [1]

Terminology

Signal – a continuous (analog) signal, distinguishable from a discrete (digital) signal.

Ex: May represent the voltage difference at time t between two points in an electrical circuit

A **communication system** consists mainly of a transmitter, communication channel (medium), and a receiver. The purpose of this system is to send a message, which may consist of written or spoken words, pictures, sound, or etc, from the transmitter to the receiver. The transmitter changes the message into a signal that can be sent through the medium (ex: wires, atmosphere) to the receiver, which will then turn the signal back into the original message. [1] The following picture is an example of a general communication system:



(http://alumni.media.mit.edu/~carsonr/phd_proposal/figures/shannon-schematic.jpg)

In the process of transmitting a signal, the signal may aquire some alterations, such as static, sound, or distortions in shape, by the time it reaches the receiver. These

alterations fall under two categories: distortion and noise. **Distortion** is a fixed operation applied to a signal, and therefore, in theory, can be undone by applying the correct inverse operation. **Noise** involves statistical and unpredictable perturbation which sometimes cannot be corrected. [1]

A **distortionless transmission** of a signal to a receiver means that the exact shape of the input signal is reproduced at the output, regardless of whether there is a change in amplitude of a time-delay. If the input signal is denoted by f(t) and the output by g(t), a distortionless signal can be represented by

$$g(t) = Lf(t) = Af(t - t_0)$$

where L is a linear, time-invariant operator. [1]

Taking the Fourier transform of both sides of this equation, we get

$$G(\omega) = H(\omega)F(\omega)$$

where F, G are the Fourier transforms of f, g respectively and $H(\omega) = Ae^{it_0\omega}$. Here $H(\omega)$ is called the system transfer function, or the system function. It's inverse Fourier transform h(t) is called the impulse response to the system. [1]

Processing a signal means that we are operating on it in some fashion, in order to change its shape, configuration and properties or to extract some useful information. Usually this operation is required to be invertible. Sometimes for practical and economical reasons, only some data extracted from the signal are transmitted and are used at the receiver to reconstruct the original signal. This is why we use sampling theory. [1]

In electrical engineering a **filter** is a circuit or system that has some frequency selective mechanism. Before signal is sampled it must be filtered. Theoretically, the

maximum frequency is half of the sampling frequency, but in practice we must use a higher sampling rate due to non-ideal filters. The ideal filters usually come in four categories: low pass, high pass, band pass or band stop. The system transfer functions of these ideal filters are as follows:

Low pass filter $$H(\omega)=\begin{cases} Ae^{it_0\omega}, for|\omega|\le\omega_1 \\ 0, otherwise \end{cases}$$

Where ω_1 is the cut-off frequency;

High pass filter $$H(\omega)=\begin{cases} Ae^{it_0\omega}, for|\omega|\ge\omega_2 \\ 0, otherwise \end{cases}$$

Band pass filter $$H(\omega)=\begin{cases} Ae^{it_0\omega}, for\omega_2\le|\omega|\le\omega_1 \\ 0, otherwise \end{cases}$$

Band stop filter $$H(\omega)=\begin{cases} 0, for\omega_2\le|\omega|\le\omega_1 \\ Ae^{it_0\omega}, otherwise \end{cases}$$

[1]

A function *f* is called **band limited** if its Fourier transform $\hat{f}$ is 0 outside of a finite interval [-L,L].

Periodic sampling is the process of representing a continuous signal with a sequence of discrete data values. With regards to sampling, the main concern is to ensure that the sampling is fast enough to ensure that the information content is preserved. [3]

Proof

I will now attempt to prove a version of the Sampling Theorem from James S. Walker's Fast Fourier Transforms. His theorem states:

Suppose that f is band limited. If $\hat{f}$ is 0 outside of [-L,L], then

$$f(x) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2L}\right) \sin c(2Lx - n)$$

Proof: Since $\hat{f}$ is 0 outside of [-L,L], we can periodically extend it. We will use $\hat{f}_P$ to denote the periodic extension of $\hat{f}$ with a period equal to 2L.

$$\hat{f}_P(u) = \sum_{n=-\infty}^{\infty} \hat{f}(u - 2nL)$$

Using Poisson summation, the Fourier series for $\hat{f}_P$ is

$$\hat{f}_P(u) = \sum_{n=-\infty}^{\infty} \frac{1}{2L} \hat{\hat{f}}\left(\frac{n}{2L}\right) e^{in\pi u/L}$$

Since $\hat{\hat{f}} = f(-x)$

$$\hat{f}_P(u) = \sum_{n=-\infty}^{\infty} \frac{1}{2L} f\left(\frac{-n}{2L}\right) e^{in\pi u/L} \qquad (*)$$

We will now multiply the previous equation by the window function in Sampling Theory, W(u), which satisfies

$$W(u) = \begin{Bmatrix} 1, |u| \le L \\ 0, |u| > L \end{Bmatrix}.$$

For this window, we have

$$\hat{f}_P(u)W(u) = \hat{f}(u)W(u) = \hat{f}(u).$$

Multiplying both sides of (*) by W(u), we get

$$\hat{f}(u) = \sum_{n=-\infty}^{\infty} \frac{1}{2L} f\left(\frac{-n}{2L}\right) W(u) e^{in\pi u/L}$$

Next we will multiply both sides by $e^{i2\pi ux}$ and integrate with respect to u from –L to L, giving the following

$$\int_{-L}^{L} \hat{f}(u) e^{i2\pi ux} du = \sum_{n=-\infty}^{\infty} f\left(\frac{-n}{2L}\right) \frac{1}{2L} \int_{-L}^{L} W(u) e^{i2\pi\left(\frac{x+n}{2L}\right)u} du$$

Let $S(x) = \frac{1}{2L} \int_{-L}^{L} W(u) e^{i2\pi ux} du$, and substitute this into the previous equation.

$$\int_{-L}^{L} \hat{f}(u) e^{i2\pi ux} du = \sum_{n=-\infty}^{\infty} f\left(\frac{-n}{2L}\right) S\left(x + \frac{n}{2L}\right)$$

Since $\hat{f}$ is 0 outside of [-L,L], we have

$$\int_{-L}^{L} \hat{f}(u) e^{i2\pi ux} du = \int_{-\infty}^{\infty} \hat{f}(u) e^{i2\pi ux} du = f(x)$$

by Fourier inversion, which says $f(x) = \int_{-\infty}^{\infty} \hat{f}(y) e^{2\pi i x y} dy$. From this we get

$$f(x) = \sum_{n=-\infty}^{\infty} f\left(\frac{-n}{2L}\right) S\left(x + \frac{n}{2L}\right)$$

Subsitute –n in place of n.

$$f(x) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2L}\right) S\left(x - \frac{n}{2L}\right)$$

This describes a very general sampling theory, where the sampled values are

$$\left\{ f\left(\frac{n}{2L}\right) \right\}_{n=-\infty}^{\infty}$$

Putting the window function back in S, we get

$$S(x) = \frac{1}{2L} \int_{-L}^{L} e^{i 2\pi u x} du = \sin c(2Lx)$$

This brings us back to our original equation in the theorem.

$$f(x) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2L}\right) \sin c\left(2Lx - n\right)$$

q.e.d. [7]

Errors and Aliasing

Several different errors exist and have been intensively studied in Sampling Theory. Some of these errors are as follows:

1) The truncation error is the error that results when only a finite number N samples are used instead of the infinitely many samples needed for the signal reconstruction. The object function cannot be reconstructed exactly because there is insufficient information. [2] The truncation error $(T_n f)(t)$ or $T_n(t)$, for short, can be controlled by imposing some extra conditions on f besides being band-limited. The first estimated value for the truncation error is credited to A. Jerr, Tsybakov, and V. Iakovlev. This estimation can be written as:

$$|T_N(t)| \le \frac{\sqrt{2}}{\pi} E \left| \sin\left(\frac{\pi t}{\Delta t}\right) \right| \sqrt{\frac{T\Delta t}{(T^2 - t^2)}}, T > 0$$

Where $-T \le t \le T, 0 < \Delta t, (1/\sigma)$ and E is the total energy of the signal which is given by

$$E = \int_{-\sigma}^{\sigma} |F(\omega)|^2 \, d\omega \quad [1]$$

2) The amplitude error, $A_e f$, arises if the exact sampled values $f(t_k)$ are not accurately known, but we know some approximation of the values, such as $\tilde{f}(t_k)$, differing from $f(t_k)$, by not more than e, are known:

$$(A_e f)(t) = \sum_{k=-\infty}^{\infty} \left[f(t_k) - \tilde{f}(t_k) \right] \frac{\sin \sigma(t - t_k)}{\sigma(t - t_k)} \quad [1]$$

3) The time-jitter error, $J_e f$, is caused by sampling at instants $\tilde{t}_k = t_k + \gamma_k$, which differ from the Nyquist sampling instants:

$$(J_e f)(t) = \sum_{k=-\infty}^{\infty} \left[f(t_k) - f(\tilde{t}_k) \right] \frac{\sin \sigma(t - t_k)}{\sigma(t - t_k)} \quad [1]$$

4) The aliasing error, $R_\sigma f$, as stated earlier, is the result when the band-limitedness conditions are violated. Therefore an aliasing error occurs during under-sampling.

$$(R_\sigma f)(t) = f(t) - \sum_{k=-\infty}^{\infty} f\left(\frac{k\pi}{\sigma}\right) \frac{\sin(\sigma t - k\pi)}{(\sigma t - k\pi)} \quad [1]$$

All four of these error types can be combined into the form:

$$(Ef)(t) = f(t) - \sum_{k=-N}^{N} \tilde{f}(t_k + \gamma_k) \frac{\sin(\sigma t - k\pi)}{(\sigma t - k\pi)} \quad [1]$$

References

1. Zayed, Ahmed I., Advances in Shannon's Sampling Theory, CRC Press, 1993.
2. Collins, Graham P. , "Claude E. Shannon: Founder of Information Theory", http://www.sciam.com/print_version.cfm?articleID=000745C4-9E66-1DA5-815A809EC5880000.
3. Sloane, N.J.A, Wyner, A.D., "Biography of Claude Elwood Shannon",
4. http://www.research.att.com/~njas/doc/shannonbio.html.
5. http://www2.egr.uh.edu/~glover/applets/Sampling/Sampling.html
6. http://members.aol.com/ajaynejr/nyquist.htm
7. Walker, James S., Fast Fourier Transforms, CRC Press, 1996.