Sampling, Aliasing, and the Blur Filter

Douglas A. Kerr, P.E.

Issue 3
November 14, 2003

ABSTRACT

When an electrical signal (such as an audio signal), or a two-dimensional photographic image, is converted to digital form, sampling is involved. When a signal or image is represented through its samples, a phenomenon of distortion called aliasing can occur. Here we discuss the basic concepts of sampling, aliasing, and the use of an antialiasing filter. Finally we discuss the "blur filter", an implementation of the antialiasing filter concept in a digital camera.

Sampling of an audio signal

When we convert an audio signal to digital form, we usually begin by capturing the instantaneous value of the signal at regular intervals, a process called *sampling*. The remainder of the signal is discarded.

Nevertheless, if the rate at which we take samples is above twice the highest frequency contained in the signal, then the entire original signal can be perfectly reconstructed from the series of sample values alone.[1] This is the principle expressed by the famous Nyquist-Shannon sampling theorem.[2]

Looking at it from the other direction, if we sample at a rate of R samples per second (*the sampling rate*), we will completely capture any signal containing component frequencies up to (but not including) $R/2$ Hz. That frequency is called the *Nyquist frequency*.

Suppose we present for sampling a signal containing components above the Nyquist frequency. Will they just be left behind in the sampling and reconstruction process? No, something more harmful occurs. Each such component will be replaced in the reconstructed signal by a spurious component. Its frequency lies below the Nyquist

---

[1] "Perfectly" assumes that each sample perfectly reflects the value of the signal at the instant of sampling.

[2] Harry Nyquist expressed it in 1928, and it was rigorously proven by Claude Shannon in 1959. Although Shannon has long been recognized for his gigantic contributions to the entire field of information theory, his role in the sampling theorem has only recently come to be widely mentioned.

frequency by the same amount that the original “out of bounds” component was above it.

For example, suppose we have established a sampling rate of 8000 samples per second. This will accommodate (and tolerate) all frequency components below 4000 Hz. If the signal being sampled has a component (likely one of no interest to us) at 4900 Hz, the reconstructed signal will contain nothing above 4000 Hz, but will have a spurious component at 3100 Hz.

In effect the maverick component travels under a false identity within the train of samples, having the same representation there as a component within the legitimate frequency range of the system. For this reason this phenomenon is called *aliasing*,[3] and its impact is called *aliasing distortion*.

In most applications, we cannot just assume that a signal about to be sampled contains no components at or above the Nyquist frequency for the sampling rate we have adopted in the system. We must assure this by first passing the signal through a low-pass filter, one which removes any components at or above the Nyquist frequency. This is often called an *anti-aliasing filter*.

Sampling images

In digital photography, the pattern of image brightness along one row of pixel sensors[4] is equivalent to our audio signal, and the capturing of that pattern by regularly-spaced pixel sensors is an example of sampling. Here, the sampling rate is expressed not in samples per second but rather in terms of samples (pixels) per millimeter. The concept of the Nyquist frequency also pertains here, in this case also expressed in cycles per millimeter. (The frequencies in this case are not in terms of time—*temporal frequency*—as in the case of an electrical signal, but rather in terms of distance—*spatial frequency*.)

If the image on the sensor array contains detail so fine that, looked at in terms of spatial frequency, it contains components at or above the

---

[3] This term has unfortunately also come to be applied to other anomalies in signal and image processing to which it does not so aptly apply!

[4] This would be equally true along one column of the sensors, or along any diagonal path through a set of sensors. In reality, the process proceeds two-dimensionally, but fully recognizing this introduces mathematical complications that would only interfere with our purpose here!

Nyquist frequency corresponding to the pixel sensor spacing, then the image as reconstructed from the pixel data will contain spurious components—the result of aliasing. To prevent this corruption of the reconstructed image, we must filter out such overly-high-frequency components—overly fine detail—in the original image.

Modulation transfer function (MTF)

To establish the next principle in a familiar context, let us return to the world of audio signals.

A system handling an audio signal (such as an amplifier or a filter) has a frequency response curve. This is a plot of how much relative attenuation the system affords to signal components at different frequencies. In an audio amplifier, we ordinarily seek to have this curve "flat" over the range of audio frequencies in which we are interested.

On the other hand, in a filter we have a response curve which is intentionally non-uniform. The antialiasing filter we discussed above, for example, has a response curve that drops sharply as we approach the Nyquist frequency.

An imaging system, such as a lens, has a comparable curve, called its modulation transfer function (MTF). This is a plot of how variations in brightness along a path across an image are attenuated, compared to the variations in the scene itself, at different spatial frequencies. (Higher spatial frequencies correspond to "finer" detail.) This approach treats the imaging system as a filter. "Blurring", the result of imperfect focus, lens aberrations, diffraction effects, and the like, is a manifestation of decline in the MTF at spatial frequencies of interest to us.

The MTF plot of an imaging system is a much more revealing indication of performance than the simplistic measure of resolution in terms of lines resolved per millimeter, and is widely used in connection with optical system design.

The blur filter

Returning to our need to avert aliasing in the case of the digital camera sensor array, we can eliminate (spatial) frequencies in the image above the Nyquist frequency by introducing into the optical path an element whose MTF drops fairly sharply as we approach the Nyquist frequency—an element which intentionally (horrors!) introduces a kind of blur. Such a "blur filter" is normally found between the lens of a digital camera and its sensor array—commonly just in front of the array..

The degree of blur this element introduces does not cause any substantial degradation to the viewed image. Keep in mind that the image already has limited resolution owing to the finite number of pixels used to represent it. The limit to resolution caused by the blur filter should be comparable to that limit.

Digital anti-aliasing filters

The provision of the anti-aliasing filter can be a big pain in system design, as it is a purely analog device. The cost of one capacitor needed in the implementation of such a filter can equal the cost of thousands of transistors in our digital integrated circuit. We can mitigate that with the use of an active filter, but that involves linear transistors, with many times the cost (and size) of the transistors in our digital circuits. Thus, the very first thing we encounter in our digital transmission system can contribute the major portion of its overall size and cost.

Isn't is possible to perform the anti-alias filter process instead with a digital filter?

Yes indeed, although with an ironic twist (which we'll hear about at the end). To do so, we must initially sample the signal at a rate many times higher than the sample rate we actually need for our "output" digital representation. (The process is often spoken of as "oversampling".) By doing so, the initial sampling process can embrace, without aliasing, the "irrelevant" higher-frequency components of our original signal—those that would have been beyond the Nyquist frequency of our basic sampling rate, but are within the Nyquist frequency of our increased sampling rate.

We then employ classical digital filtering techniques to obtain a sampled representation (at the high sampling rate) of a signal containing no components at or above the Nyquist frequency for our basic sampling rate. This is the anti-aliasing filter process with respect to the "real" sampling, at the basic sampling rate, which comes next.

Sampling this signal at the basic sampling actually involves only choosing those samples of its current form that fall at the "basic" sampling interval and discarding the rest—"resampling by decimation".

The ironic twist is that, even for our higher initial sampling rate, there is the potential that there could be components in the original signal at or above its Nyquist frequency (much higher than the one for our basic sampling rate, of course) and thus of the possibility of aliasing at this point in the process.

Thus, we still have to have an anti-aliasing low-pass filter at the very beginning of the process. And it has to be an analog filter. But because the Nyquist frequency for this first sampling stage is quite high, the analog components needed to make such a filter can be small in value (capacitance and/or inductance), and accordingly small in size and cost. Thus we have mitigated the cost impact of the anti-aliasing filter—but not eliminated it!

#