Sampling and Reconstruction

Lecture #9: Tuesday, 18 Feb. 2003
Lecturer: Greg Humphreys
Scribe: Rui Wang
Reviewer: Nolan Goodnight

1 Aliasing

In a sense, all rendering has to do is integration. Real-world imaging systems convert a continuous image to a discrete image by integrating over some sensor, such as photoreceptor, CCD array, Light Sensitive Crystals. Remember the response of sensor is a multi-dimensional integration:

$$\int_T \iint_\Omega \iint_A L(x,\omega,t)\ P(x)\ S(t)\cos\theta\ dA\ d\omega\ dt$$

Computer can only record discrete data, as a result, to represent continuous signal, sampling is necessary: divide continuous signal into small regions and associate a number with each region. For example, we can choose a value for (x, ω, t) and the function value of it is a sample. But in real world, sensors don't sample, instead, they integrate over the sampling region. To display, the discrete signal is converted back to continuous signal. For example, in a CRT (Fig.1), each phosphor glows in some distribution, like a gaussian in area. Usually this is a two-step process involving a DAC which 'samples and holds'. This process of converting discrete signal to continuous signal is called "reconstruction'. The two processes can cause artifacts. Sampling artifacts are also called 'aliasing', as opposed to reconstruction artifacts. To distinguish, sampling artifacts are also called 'pre-aliasing' and reconstruction artifacts 'post-aliasing'. Any approach to prevent these artifacts is called 'anti-aliasing'. A good way to understand, analyze and eliminate these artifacts is through Fourier analysis.

2 Sampling

2.1 Uniform Sampling

In the spatial domain, given a function $f(x)$, and a bunch of sample locations x_i, sampling of $f(x)$ can be done just by plugging in each x_i and evaluate. To get a uniform sampling, we can multiply $f(x)$ by a Shah function (Fig.2)

$$|||_T(x) = \sum_{n=-\infty}^{+\infty} \delta(x - nT)$$

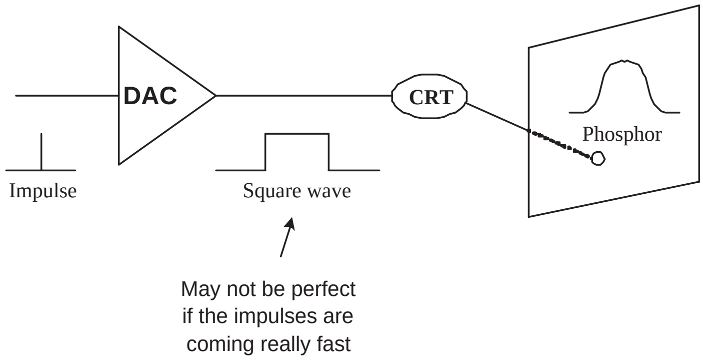

Figure 1: Typical CRT

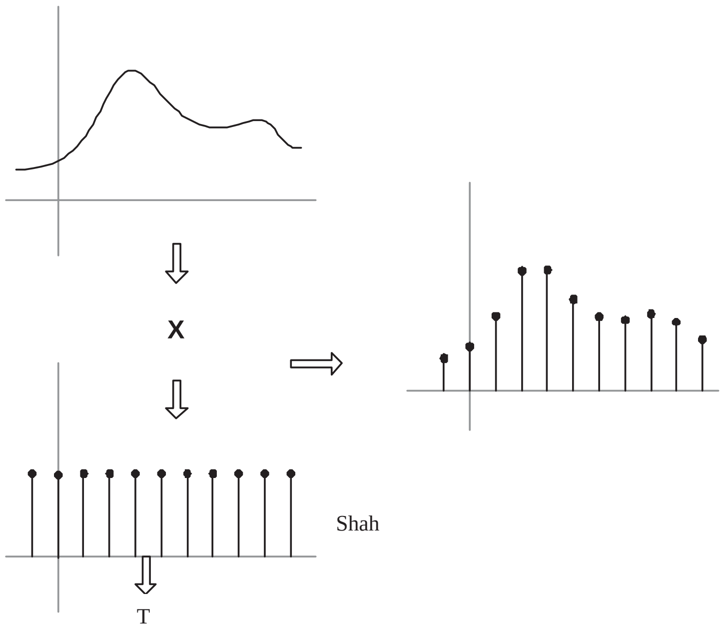

Figure 2: Uniform sampling in spatial domain

In the frequency domain, the spectrum of the Shah function is another shah function, with reciprocal period T:

$$|||(\omega) = \frac{2\pi}{T} \sum_{n=-\infty}^{+\infty} \delta(\omega - n\frac{2\pi}{T})$$

The reciprocal law means that if we sample too far away in the spatial domain, the impulses of its Fourier transformed Shah function in the frequency domain will be too close to each other. Because point-wise multiplication in the spatial domain is equal to

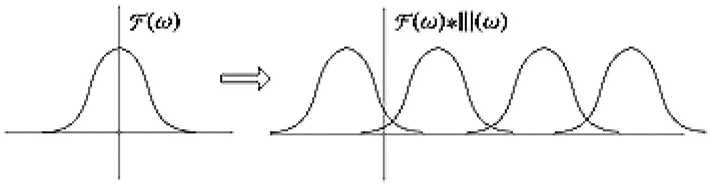

Figure 3: Uniform sampling in frequency domain

convolution in the frequency domain, so uniform sampling of $f(x)$ in the spatial domain is the same with convolving its spectrum with the transformed Shah function in frequency domain. And this essentially means laying down multiple copies of the spectrum of $f(x)$ (Fig.3) Notice that if the period of the transformed Shah function is too small (which means the sampling step is too far away), the convolved signal spectrum will overlap each other, this is where aliasing comes from.

2.2 Reconstruction

To reconstruct the original signal, in frequency domain, we can take $|||(\omega) \bigotimes \mathcal{F}(\omega)$ (essentially the spectrum of the sampled signal), multiply it with a box filter (low pass filter) to obtain a copy of $\mathcal{F}(\omega)$, then apply a reverse Fourier transform to get the original signal $f(x)$. However, this does not guarantee reconstruction of original signal. The problem is, if the copies of $\mathcal{F}(\omega)$ overlaps each other (like in Fig.3), high frequencies of $\mathcal{F}(\omega)$ will be added to low frequencies and thus there is no way to obtain an exact copy of $\mathcal{F}(\omega)$, hence the aliasing happens.

The Fourier transform of a Box filter is $\mathcal{F}^{-1}(Box) = sinc(x) = \frac{\sin(\pi x)}{\pi x}$, so Multiplying $\mathcal{F}(\omega)$ with a box filter in frequency domain is the same with convolving the sampled signal with *sinc* filter in spatial domain. Notice the *sinc* is zero at all sample plints except the one that it's convolving around. So the *sinc* function is called a perfect low-pass filter. In fact any function that is 1 at the origin and zero at all integers (or sample points, more precisely) will act as a low-pass filter for reconstructing a discrete signal. In conclusion, to correctly reconstruct a signal by uniform sampling it, the spectrum of $f(x)$ (or $\mathcal{F}(\omega)$ must fit entirely within the low-pass filter, which is a box in frequency domain.

2.3 The Sampling Theorem

The Sampling Theorem is discovered by Claude Shannon (1949), it says: a signal that can be reconstructed from its samples without loss of information if the sampling frequency is more than twice above the highest frequency contained in the original signal. This critical sampling frequency is called the Nyquist Frequency. According to the sampling theorem, if a signal is not band-limited, meaning its highest frequency is infinite, there is

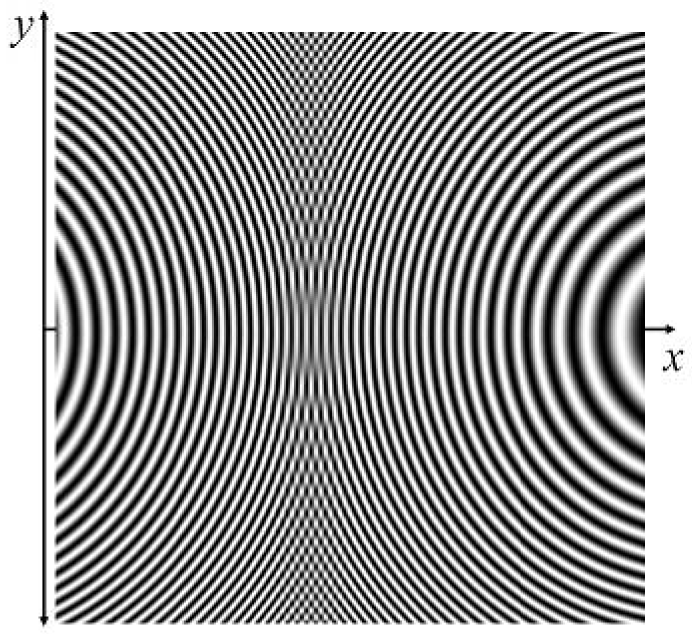

Figure 4: The Zone Plate

no way it can be reconstructed without loss of information. Unfortunately, most signals we deal with in Computer Graphics are not band-limited, such as

- Box, or generally, polygon, which has sharp corners (changes)
- Polynomial, or generally, any signal that going to infinity
- Any signal that has compact support in the spatial domain.

One way to check if a function is band-limited is to think if it can be represented as the sum of finite number of sin and cos functions. The fact that signals with compact support (in spatial domain) are not band-limited is the reason why aliasing is so prevalent in graphics.

2.4 Under Sampling

Under sampling happens when we sampling a signal with frequency less than Nyquist frequency. The sampled signal, in the frequency domain, will appear to be laid copies of $\mathcal{F}(\omega)$, and with some overlapping tails. This is called 'aliases', because high frequencies are masquerades as low frequencies. A good example of showing aliases is the zone plate (Fig.4): the zone plate function is $\sin(x^2 + y^2)$, or $\sin(r^2)$, as it gets farther from the origin, the frequency going to infinity. The function is supposed to look like a bunch of rings, but in Fig.4 rings appear on the right, and those are aliases.

2.5 Reconstruction Filters

Although *sinc* function is perfect low-pass filter, it's problematic: they have infinite extent; also, *sinc* function introduces 'ringing' at discontinuities which are perceptually objectionable. Some experiments use cubic reconstruction filter, but since it's not perfect low-pass filter, it causes post-aliasing called moiré patterns.

Mitchell explored the space of possibilities with cubic polynomial filters and derived a two-parameter filter family:

$$K(x) = \frac{1}{6}\begin{cases} (12-9B-6C)|x|^3 + (-18+12B+6C)|x|^2 + (6-2B) & \text{if } |x| < 1 \\ (-B-6C)|x|^3 + (6B+30C)|x|^2 + (-12B-48C)|x| + (8B+24C) & \text{if } 1 \le |x| \le 2 \\ 0 & \text{otherwise} \end{cases}$$

The formula is derived by first considering a piecewise cubic function that looks similar to a *sinc* function:

$$K(x) = \begin{cases} P|x|^3 + Q|x|^2 + R|x| + S & \text{if } |x| < 1 \\ T|x|^3 + U|x|^2 + V|x| + W & \text{if } 1 \le |x| \le 2 \\ 0 & \text{otherwise} \end{cases}$$

then by adding constrains that the function must be C^0 and C^1 continuous at $x = 0, 1, 2$, and also it must be normalized such that $\sum_{n=-\infty}^{+\infty} K(x-n) = 1$.
Certain parameters of (B, C) prove to correspond to well-know filters:

- (0,1): B-spline [positive everywhere]
- (0,$\frac{1}{2}$): Catnull-Rom (good approximation to a *sinc*

Although the formula defines a two-parameter family of filters, Mitchell recommends a 1D sub-family along the line $B + 2C = 1$ for numerical fit reasons.