Generating Antialiased Images at Low Sampling Densities

Don P. Mitchell
AT&T Bell Laboratories
Murray Hill, New Jersey 07974

ABSTRACT

Ray tracing produces point samples of an image from a 3-D model. Constructing an antialiased digital picture from point samples is difficult without resorting to extremely high sampling densities. This paper describes a program that focuses on that problem. While it is impossible to totally eliminate aliasing, it has been shown that nonuniform sampling yields aliasing that is less conspicuous to the observer. An algorithm is presented for fast generation of nonuniform sampling patterns that are optimal in some sense. Some regions of an image may require extra sampling to avoid strong aliasing. Deciding where to do extra sampling can be guided by knowledge of how the eye perceives noise as a function of contrast and color. Finally, to generate the digital picture, the image must be reconstructed from the samples and resampled at the display pixel rate. The nonuniformity of the samples complicates this process, and a new nonuniform reconstruction filter is presented which solves this problem efficiently. This paper was presented in SIGGRAPH 87.

CR Categories and Subject Descriptions: I.3.3 [Computer Graphics]: Picture/Image Generation

General Terms: Algorithms

Additional Keywords and Phrases: Adaptive Sampling, Antialiasing, Filtering, Noise Perception, Nonuniform Sampling, Ray Tracing, Reconstruction

1. Introduction

While ray tracing is a straightforward technique for image synthesis, it has in some ways proven to be one of the most perverse. On the one hand, impressive effects of shading, reflection, and refraction can be simulated by ray tracing [WHI80], and it remains one of the simplest methods for evaluating constructive solid geometry models [ROT82]. On the other hand, ray tracing is expensive, and because it is based on point sampling it is especially prone to aliasing problems.

This paper describes a method that uses the sampling ability of ray tracing more effectively than alternative techniques. The ray tracer itself is viewed as a black box capable of supplying samples of a two-dimensional image signal (by casting rays into a scene). A separate antialiasing module decides where to take samples, how densely to sample each region of the image, and how to construct a digital picture from all the sample information. This cleanly isolates the aliasing problem from all other issues. By not considering the inner workings of the sample source, some useful information may be ignored, but the resulting program is very general. In our laboratory, four different ray tracers currently use the prototype antialiasing module.

The module is composed of three principal stages. First, the entire image is sampled at a relatively low density. Then, on the basis of these initial sample values, some regions of the image may be sampled further at a higher density to reduce noise. Finally, in the reconstruction phase, all of the samples of the image are used to generate the pixel values for a displayed image.
Of the many issues involved in optimizing ray tracing, this paper will focus on reducing sampling density while still producing an image of high quality. The approach has been to consider the perception of noise and operate near its threshold. While perception of noise is a complex subject, basic facts learned from work in color-television encoding and image noise measurement can be applied to this problem [LIM77, SAK77].

By varying position and density of the sampling, a more optimal sampling pattern can be obtained. For example, certain nonuniform sampling patterns have been shown to scatter aliasing into high-frequency random noise [YEL83]. Deciding where to concentrate sampling (supersample) can be guided by estimation of the noise visibility as a function of contrast and color.

2. The Preliminary Sampling Pattern

The choice of sampling strategy is now known to be closely related to the aliasing problem. The history of this problem began with Crow's realization that the "jaggies" in synthetic pictures were an example of aliasing, a well-known phenomenon of signal processing [CRO77]. According to the Sampling Theorem [SHA49], aliasing can only be reduced by either filtering out the high-frequency components of the image (i.e., removing everything above the Nyquist frequency) or by increasing the sampling density to an appropriate level. Both of these techniques have been used to combat aliasing in ray tracers [AMA84, WHI80], but neither is completely satisfactory. Prefiltering has been done only approximately and for scenes built up from a restricted class of primitive objects. The alternative of increasing the sampling density means a corresponding increase in computing cost.

An entirely different approach to the problem is suggested by Yellott's work [YEL83]: when aliasing cannot be removed, it can still be made less conspicuous by sampling nonuniformly. Uniform sampling tends to produce highly visible forms of noise because regular structure in the image can "beat" with the sampling pattern producing Moiré patterns and "jaggies". With nonuniform sampling, aliasing noise can be random and structureless. Yellott realized this by studying the layout of photoreceptors in the retina of the eye. This concept was first applied to image synthesis by researchers at Lucasfilm and has since been explored by a number of others [COO84, KAJ84, ABR85, DIP85, LEE85, MIT85, COO86, KAJ86].

Unfortunately, nonuniform sampling of nonbandlimited signals is poorly understood. For the most part, only problems that are very remote to the image-synthesis problem have been studied by signal processing experts (e.g., the extraction of statistics from nonbandlimited stationary noise [SHA69]). In the application to graphics, several important questions remain unanswered. In particular, the quality of images produced by nonuniform sampling has been questioned [ABR85], and systems which definitely do produce high-quality images [COO84, LEE85] use rather high lower bounds on sampling densities (16 and 8 samples per pixel, respectively).

For fixed-density (i.e., nonadaptive) sampling, two common forms of nonuniform sampling have been discussed. Yellott [YEL83] recommends using the distribution he observed in retinal cells, a Poisson-disk distribution (a Poisson distribution with a minimum-distance constraint between points). Except for a few experimental pictures [DIP85], I know of no rendering system that uses this pattern. Instead, the more easily generated jitter pattern is commonly used [COO86]: a uniform sampling pattern with a random offset added to each point. The random offset of the jitter demonstrated in this paper is uniformly distributed, so the process is equivalent to choosing one random point within each pixel tile.

Theoretical evidence in favor of the Poisson-disk sampling pattern was presented by Yellott [YEL83]. He proposed that the least-conspicuous form of aliasing would be produced if the spectrum of the sampling pattern had two properties. First, the spectrum should be noisy and lack any concentrated spikes of energy. If present, such spikes could yield coherent aliasing in the same way that uniform sampling can. Second, the spectrum should have a deficiency of low-frequency energy. This causes aliasing noise to be concentrated in higher, less conspicuous frequencies. These conditions will be referred to as the blue-noise criteria. Figure 1 shows an instance of such a sampling pattern (generated by the dart-throwing algorithm described below) and its Fourier transform.
Studies have shown that the eye is most sensitive to noise in intermediate frequencies [SAK77]. While frequencies up to 60 cycles per degree can be visible, the maximum response to noise is at about 4.5 cycles per degree. Taking advantage of this response is also how ordered dither improves the quality of halftoned images [LIM69]. By adding high-frequency noise to an image before quantizing, Limb discovered that annoying contour effects were masked. Although more noise was added to the image, the results were far more pleasing when viewed at a reasonable distance.

Figure 2 shows the pattern and spectrum of a jitter process. The sampling pattern is clumpier, or more "granular", and the spectrum shows a smaller noise-free region around the origin. This implies that noise with lower frequencies (i.e., larger features) would be generated by this type of sampling.

Figures 3 through 5 give a more direct visual demonstration of uniform, jitter, and Poisson-disk sampling, respectively. For these figures, a test pattern has been designed specifically to generate aliasing. It is given by the following simple formula:

\[
\left\lfloor \frac{100x}{x+y} \right\rfloor \mod 2
\]

An image is reconstructed from these samples using the weighted-average filter (described in section 4). The resulting pictures show the character of the aliasing noise clearly. For each sampling method, the average sampling density is the same, 160 by 160 samples. As expected, uniform sampling (Figure 3) yields large Moiré patterns, which are visible even at a distance. Comparing Figures 4 and 5, both lack Moiré patterns, but jitter sampling (Figure 4) has introduced significantly more "grain noise" into the picture. The pits and bumps along edges are much less severe when Poisson-disk sampling (Figure 5) is used. Note that Figures 3 through 5 were originally 160 by 160 pixels, but the images have been enlarged by digital resampling to 512 by 512 so that details of the aliasing noise can be seen.

What has made jitter sampling attractive is undoubtedly its simplicity. Poisson-disk sampling is somewhat more complicated. Several point processes could be referred to as "Poisson-disk" [RIP77], but by strict definition, a true Poisson-disk process is realized by generating entire patterns with Poisson statistics and then throwing them away until one meets the minimum-distance criterion.

A more practical point distribution is realized by a "dart-throwing" algorithm [MIT85, COO86], in which points are generated randomly with uniform distribution over the area being filled (this by itself closely approximates a Poisson distribution). Each new point is rejected if it falls within a certain distance of any previously chosen points, otherwise it is added to the pattern, and the process continues. This process is repeated until no new points can be added to the pattern. This is an expensive algorithm (several hours of VAX 780 time to generate 1024-point pattern), which is only practical for precomputing a small pattern with periodic boundary conditions to be permanently stored and replicated over the entire picture. As mentioned above, this pattern satisfies the blue-noise criteria.

I have found that Poisson-disk samples can be generated cheaply on the fly by an algorithm inspired by the Floyd-Steinberg halftoning algorithm [FLO75]. To generate a pattern with an average of one sample per pixel, the algorithm selects points from a uniform two-dimensional grid four times finer than the display pixel lattice (i.e., there are 16 grid points per pixel area). A small amount of random jitter can be added to these sample positions to dislocate them from the grid.

As the algorithm proceeds, the grid points are operated on in a scanline order. Each point on the grid is associated with a diffusion value \(D_{ij}\), which is computed from previously computed diffusion values and a noise source \(R\):

\[
T = \frac{4D_{i-1,j} + D_{i-1,j-1} + 2D_{i,j-1} + D_{i+1,j-1}}{8} + R
\]
The noise source $R$ is a uniform random value in a range of $1/16$ - $1/64$ to $1/16 + 1/64$. This will ensure that about one out of 16 grid points will be selected as a sampling point and provides enough fluctuation to prevent orderly structures from appearing in the pattern. The diffusion values (defined below) correspond to the diffusing quantization error in the Floyd-Steinberg algorithm.

The temporary value $T$ is used to decide whether to select a grid point as a sampling position:

$$\text{SELECT} = \begin{cases} 
0 & \text{if } T < 0.5 \\
1 & \text{Otherwise}
\end{cases}$$

(3)

If SELECT is 1, the grid point is a sampling position. Finally, a local diffusion value is computed as:

$$D_{i,j} = T - \text{SELECT}$$

(4)

Only two scan lines of $D$ values must be stored to execute this algorithm. Equation 2 assumes a scanning from left to right. A more isotropic pattern is produced if scanning sweeps back and forth. Figure 6 shows the pattern and spectrum produced by this algorithm. The error-diffusion weights in (2) have been chosen experimentally to produce a sampling pattern that satisfies the blue-noise criteria. In addition, these weights do not require any multiplications in an efficient implementation. This method of generating Poisson-disk patterns is called the point-diffusion algorithm.

3. Adaptive Sampling

Nonuniform sampling by itself does not eliminate aliasing; it merely changes the characteristics of aliasing to make it less noticeable. However, when a region of an image contains an edge or some other type of high-frequency pattern, the sampling density must be increased to get a genuine reduction of aliasing noise.

Adaptive sampling is not a new idea. Whitted describes an adaptive sampling algorithm based around uniform sampling [WHI80]. Two basic types of adaptive nonuniform sampling have been described in the literature. One approach allows sampling density to vary as a function of local image variance [LEE85, KAJ86]. An example of another approach is in [COO86] where two sampling densities are used, a regular pattern for most areas and a higher-density pattern for troublesome spots. For low densities, a two-level strategy is appealing because it allows the optimal blue-noise pattern to be used easily. In the prototype antialiasing module, the entire image is sampled in a base pattern with an average of one sample per pixel. Additional supersampling is typically done at a density of 4 to 9 samples per pixel. At 4 samples per pixel, aliasing noise is barely visible, and at 9 samples per pixel, the image looks perfect. Much higher supersampling densities may be required if "distributed ray tracing" is being done.

Having chosen to use two levels of sampling, where should the high density sampling be invoked? To begin, the image is divided into small supersampling cells containing eight or nine of the low-density samples, and a decision is made whether to sample each cell at a higher density. In the current version of the antialiasing module, these cells are square.

Theoretically, it is not possible to infallibly detect the presence of high frequency regions in an image by sampling at such a low rate. The success of this algorithm depends on additional assumptions about the structure of typical images. For example, this method is likely to detect edges, and it is known that edges are a common type of region that needs high-density sampling. On the other hand, coarse sampling can miss minute isolated features; special action would have to be taken to ensure that they will be sampled. Objects which are going to appear small in the scene could be surrounded by a larger invisible bounding surface [WHI80], but this cannot be applied to small features resulting from shadows, reflections or highlights.

Given the set of eight or nine samples in a supersampling cell, some quantitative measure of variation in
signal is needed to decide whether supersampling should be done. The variance of the sample values can be used to indicate high frequency [LEE85]; however, variance is a poor measure of visual perception of local variation. The nonlinear response of the eye to rapid variations in light intensity is more closely modeled by contrast [CAE81]:

\[ C = \frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{max}} + I_{\text{min}}} \]  

(5)

Each sample value actually consists of three separate intensities for red, green, and blue. These could be combined into a luminance value before contrasts are computed, but a safer method is to compute three separate contrasts for red, green, and blue. These three contrasts are each tested against separate thresholds, and supersampling is done if any one is higher than threshold. In a prototype system, red, green, and blue thresholds are set to 0.4, 0.3, and 0.6 respectively.

This test is most sensitive to green in accordance with the human eye’s known response to noise as a function of color. Because green-sensitive cone cells are far more common in the human retina (and thus green is sampled more densely), sensitivity to green-colored noise is twice as great as to red-colored noise and four times greater than to blue-colored noise. This knowledge has been used in the encoding of color television transmission to make the most economical use of bandwidth [LIM77].

Figure 7 shows an image generated by a prototype antialiasing module combined with a basic ray tracer. Figure 8 indicates with white where supersampling had been selected. Note that many edges were not supersampled because the contrast was not judged to be sufficient to produce conspicuous noise. In the thin shadow of the tripod’s leg (seen between the footpads) a couple small gaps can be seen. This is a failure of the high-frequency detection (as discussed above), but overall detection is remarkably good for such coarse sampling (one base sample per pixel). This small flaw can be removed by doubling the density of the base sample, and thus doubling the cost of the picture; a typical dilemma when rendering is entrenched in point sampling.

4. Reconstruction from Nonuniform Samples

A digital picture is a uniform array of pixel values. These must be computed from the nonuniform sample values obtained by the methods described above. Ideally, this is done by constructing a surface that interpolates the sample values and can be resampled at the uniform pixel locations. At the same time, this surface must be bandlimited so that the resampling process does not itself introduce aliasing. This bandlimited interpolation is known as reconstruction in the signal-processing domain.

Little practical information exists for dealing with nonuniform samples. Yen discusses the reconstruction problem and presents solutions for several one-dimensional cases [YEN56]. Unfortunately, these formulas—as Yen points out—are too unstable to be of real use. Many ad hoc algorithms exist for "smooth" interpolation of nonuniform data (Franke gives a good review of this subject. [FRA82]), but none of these algorithms are suited for the problem at hand. In general, they are not meant to deal with such a large number of points (on the order of a million samples must be processed in a typical digital color picture), and they are not designed with any digital filtering characteristics in mind.

If the samples had been made uniformly, the problem would be quite a bit simpler. Reconstruction can then be done by convolving the samples with an appropriate filter \(k\). In one dimension, this is:

\[ f(x) = G \sum_{n=-\infty}^{\infty} f(x_n)k(x - x_n) \]  

(6)

where \(f(x_n)\) are the sample values. It is important to note that a constant factor \(G\) is needed for normalization. This is inversely proportional to sampling density. The summation is infinite, but in practice, filter
If any filter could be used in the reconstruction, its design would be a complex problem. In addition to well-understood quantitative criteria such as bandpass and leakage characteristics, filters for images must deal with poorly understood subjective criteria. Ringing and blurring effects must be traded off [BRO69], and because the filter is two-dimensional, there are conflicts between isotropy and interpolation [PET62].

High-frequency leakage is a particularly important filter characteristic because aliasing can result during resampling if the reconstruction from the nonuniform samples is not properly bandlimited. Figures 9 and 10 demonstrate this fact. Each of these figures has been sampled nonuniformly 100 times per pixel. Figure 9 is reconstructed with a box filter, and some faint aliasing can be seen near the origin. This aliasing is reduced in Figure 10 where a better filter has been used (a sinc filter with a Hamming window).

When samples are nonuniform, there are many added complications, and the convolution in (6) is no longer valid. If samples clump together in a small region, (6) would make that area too bright; and similarly, regions where the samples were sparse would be too dark. This unwanted effect of sample-density fluctuation is called grain noise.

If the sampling pattern is not too clumpy, grain noise can easily be handled by turning the filter in (6) into a weighted-average filter. In one dimension this would be:

\[ f(x) = \frac{\sum_{n=\infty}^{\infty} \sum_{k} k(x - x_n) f(x_n)}{\sum_{n=\infty}^{\infty} k(x - x_n)} \] (7)

This filter has been used in a number of implementations of nonuniform sampling [COO86, DIP85], but unfortunately, it does not handle extreme variation in local sampling density. An example of this failure can be seen in Figure 11. In this 3D plot, height represents the gray level in a 2D image. This figure has been generated by sampling a step function nonuniformly. In a narrow region about the step, sampling is 16 times more dense than in the rest, simulating the effects of adaptive sampling. The surface shown is a weighted-average reconstruction from those samples.

Simple weighted average fails because it allows dense clumps of supersamples to overwhelm the influence of near-by base samples, as demonstrated by a simple thought experiment: Imagine a white circle being rendered against a black background with simple box filtering (i.e. area averaging). In one pixel, the edge of the circle may barely graze a corner. However, that tiny corner may be covered with many supersamples, while only one base sample at the opposite corner of the pixel lies on the black background. If all the samples in the pixel are averaged, the whole pixel will be nearly white, possibly creating a visible bump on the edge.

A practical solution, used in the prototype antialiasing module, is a multi-stage filter. Weighted-average filters are repeatedly applied with ever-narrowing low-pass cutoff until the proper bandwidth for the display is reached (i.e. until frequencies above the Nyquist frequency for the display are eliminated). This filter operates on a discrete two-dimensional grid much finer than the display pixels; 16 grid points per pixel area. For efficiency, each individual stage is a box filter.

Figure 12 shows the result of applying the first stage of this filter. For each grid square, all samples lying within it are added up and the result divided by the number of samples. This is equivalent to filtering with a weighted-average box filter one grid square in width (one fourth the width of a display pixel).

Figure 13 shows the second stage after applying a box filter twice as wide (i.e. with half the frequency bandwidth) as the first. This half-pixel filter is applied twice in the prototype antialiasing module to
improve the filter characteristics.

Figure 14 shows the result of applying the final stage of the filter. A weighted-average box filter one full pixel wide has been applied to complete a smooth surface.

There are three relevant properties of this filter that make it successful. First, it is quite simple and efficient. The evaluation of the multi-stage filter dominates the computing cost of the antialiasing module, but because it is made up of box filters and because fixed-point arithmetic is used in the filter evaluation, only about two minutes of CPU time (on a VAX 8650) are used for a 640-by-480 picture. That time is more or less constant for all pictures, and it is almost always dwarfed by the time spent in the ray tracer computing sample values for a picture (which is typically 30 minutes to several hours).

Secondly, the multistage filter deals well with highly variable sampling density. This filter is a generalization of the reconstruction used by Whitted [WHI80] in his adaptive sampling scheme. Dense clusters of supersamples are averaged and normalized locally before being combined with nearby base-sample values.

Finally, the low-pass filter characteristics are adequate. The stages of the filter can be described by the convolution of four box filters:

\[
k(x) = \text{box}(x) * \text{box}(2x) * \text{box}(2x) * \text{box}(4x)
\]

This results in a piece-wise cubic filter. As shown below, the multi-stage filter has much less leakage of signal above the low-pass cutoff than a simple box filter. In the following figures, the sampling frequency is 1, and the Nyquist frequency is 0.5.
This filter fits into a general scheme for nonuniform reconstruction proposed by Netravali [NET86]. He suggests that the interpolation and bandwidth properties of the reconstruction be separated. First, the nonuniform samples can be interpolated with any simple algorithm that does not attenuate too much signal below the display bandwidth. In a second stage, the surface resulting from the first step is then filtered in a more sophisticated manner to remove frequencies above the display bandwidth. The multi-stage filter accomplishes this, but in a sequence of gradual steps.

5. Conclusions

Constructing a digital picture from a point-sampling function (e.g., a ray tracer) is complicated by the problem of aliasing. Nonuniform sampling is now recognized as a method for making aliasing noise less conspicuous to the viewer. Experiments with test patterns indicate that sampling patterns that meet the blue-noise criteria of Yellott produce results that are superior to those produced by sampling at the same density with jitter. The point-diffusion algorithm for generating such blue-noise sampling patterns is presented.

Some regions of an image may contain a large high-frequency component and require local supersampling to reduce the intensity of aliasing noise. Rather than sample many times in one pixel area, a larger neighborhood of low-density base samples are used to detect high frequency, and this entire neighborhood is supersampled if necessary. This allows a lower overall sampling density to be used at the expense of less-precise targeting of supersampling.

Ideally, the decision of whether or not to supersample should be based on whether or not aliasing would be visible if supersampling were not done. A step in that direction is to measure local variation in sample intensity by a function reflecting the eye’s non-linear response to contrast. In color images, red, green, and blue contrasts are calculated and compared against three separate thresholds. The red, green, and blue thresholds can be adjusted to account for the eye’s highly variable sensitivity to noise as a function of color.

Having gathered all samples, the final step in producing a digital picture is to reconstruct and resample at the display pixel rate. The nonuniformity of the samples and the sudden changes in sampling density (on the boarders of supersampled regions) make reconstruction difficult. A multi-stage filtering algorithm is presented that solves these problems efficiently.

6. Acknowledgements

I would like to thank Jim Kaiser, Henry Landau, John Limb, and Arun Netravali for many educational discussions about image and signal processing. Tom Duff wrote the program which produced figures 11 through 14.
References


