Linear Digital Imaging System Fidelity Analysis

Steve Park
College of William & Mary
Department of Computer Science
May 30, 1989

In this paper, the combined effects of image gathering, sampling, and reconstruction are analyzed in terms of image fidelity. The analysis is based upon a standard end-to-end linear system model which is sufficiently general so that the results apply to most line-scan and sensor-array imaging systems. Shift-variant sampling effects are accounted for with an expected value analysis based upon the use of a fixed deterministic input scene which is randomly shifted (mathematically) relative to the sampling grid. This random sample-scene phase approach has been used successfully by the author and associates in several previous related papers [1]–[4].

Formulation

The end-to-end linear model upon which the results of this paper are based is characterized by three independent system components, an input scene $f(x, y)$, an image gathering point spread function $h(x, y)$, and an image reconstruction point spread function $r(x, y)$. All three of these components are referenced to a common orthogonal spatial coordinate system (x, y) normalized so that the sampling interval in both directions is unity. That is, sampling occurs at the integer coordinates (m, n). Because of this normalizing convention, when the model is analyzed in the Fourier domain, the associated spatial frequencies (μ, ν) have units of cycles/pixel and the Nyquist (folding) frequency is 0.5.

For notational convenience two other components are introduced, the pre-sampling image $g(x, y)$, and the reconstructed image $f'(x, y)$. The end-to-end model that relates the input scene f to the output reconstructed image f' is then

$$f(x,y) \overset{*h}{\longrightarrow} g(x,y) \overset{\text{sample}}{\longrightarrow} g(m,n) \overset{*r}{\longrightarrow} f'(x,y) \tag{1}$$

where the $*$ operator denotes 2-d spatial convolution and $g(m, n)$ is $g(x, y)$ sampling onto the pixel grid. This model is the basis for all the analysis that follows, and, consequently, the results of this paper are applicable to the fidelity analysis of any sampled imaging system whose performance is characterized by the equation

$$f'(x,y) = \big[[f(x,y) * h(x,y)]\text{comb}(x,y)\big] * r(x,y) \tag{2a}$$

where

$$\text{comb}(x,y) = \sum_m \sum_n \delta(x - m, y - n) \tag{2b}$$

is the conventional 2-d comb function which accounts for sampling.

3

Image Fidelity

A variety of metrics have been advocated to measure how well one image matches another. These metrics include the 1-norm

$$\|f - g\|_1 = \int_x \int_y |f(x,y) - g(x,y)| \, dx \, dy$$

the common RMS 2-norm

$$\|f - g\| = \left(\int_x \int_y |f(x,y) - g(x,y)|^2 \, dx \, dy \right)^{1/2}$$

which generalizes for $p \neq 2$ to the p-norm

$$\|f - g\|_p = \left(\int_x \int_y |f(x,y) - g(x,y)|^p \, dx \, dy \right)^{1/p}$$

and which approaches the ∞-norm

$$\|f - q\|_\infty = \max_{x,y} |f(x,y) - g(x,y)|$$

in the limit as $p \to \infty$. Of these, the RMS norm $\|f - g\|$ is far and away the most common, presumably because it lends itself so well to mathematical analysis.

The RMS norm *squared*

$$\|f - g\|^2 = \int_x \int_y |f(x,y) - g(x,y)|^2 \, dx \, dy \tag{3}$$

is a measure of image fidelity [5]. Specifically, the conventional definition of fidelity is

$$\text{fidelity} = 1 - \frac{\|f - g\|^2}{\|f\|^2} \tag{4}$$

The primary purpose of this paper is to illustrate how the method of sample-scene phase averaging can be used to derive expressions for the three fundamental "fidelity loss" metrics

$$\|f - g\|^2 \quad \text{and} \quad \|g - f'\|^2 \quad \text{and} \quad \|f - f'\|^2.$$

The first of these metrics is a measure of *image blur*, the common loss of high spatial frequencies caused, for example, by defocus [5]. The second is *sampling and reconstruction blur*, the loss of fidelity caused by sampling (aliasing) and imperfect reconstruction [1]. The third, and most important, metric is the *end-to-end blur*, the net loss of fidelity caused by the combined effects of image gathering, sampling and reconstruction [6], [7]. Each of these fidelity loss terms will be analyzed in order, beginning with image blur.

4

Image Blur

The conventional *continuous-continuous* model of image formation (image gathering) is that the process is both linear and shift-invariant. That is, f and g are related by a convolution as

$$g(x,y) = \int_{x'} \int_{y'} h(x-x', y-y')f(x',y')dx'\, dy' \tag{5a}$$

where $h(x,y)$ is the image gathering point spread function (PSF) conventionally normalized so that

$$\int_x \int_y h(x,y)dx\, dy = 1. \tag{5b}$$

This model is much more easily understood when expressed in the spatial frequency (μ, ν) domain as

$$\hat{g}(\mu,\nu) = \hat{h}(\mu,\nu)\hat{f}(\mu,\nu) \tag{6a}$$

where

$$\hat{g}(\nu,\mu) = \int_x \int_y g(x,y)\exp(-2\pi i(x\mu+y\nu))dx\, dy \tag{6b}$$

is the Fourier transform of g and the transforms $\hat{h}$, $\hat{f}$ are defined analogously.

It is well known that the PSF h typically acts as a low-pass filter. As a result, g is a blurred copy of f and the extent of this image blur is

$$\|f-g\|^2 = \int_x \int_y |f(x,y)-g(x,y)|^2 dx\, dy \tag{7a}$$

which can be rewritten, using the energy (Parseval's) theorem, as

$$\|f-g\|^2 = \int_\mu \int_\nu |\hat{f}(\mu,\nu)-\hat{g}(\mu,\nu)|^2 d\mu\, d\nu. \tag{7b}$$

However, from equation (6a), this last equation can be written as

$$\|f-g\|^2 = \int_\mu \int_\nu |1-\hat{h}(\mu,\nu)|^2|\hat{f}(\mu,\nu)|^2 d\mu\, d\nu. \tag{7c}$$

Note that if some metric other than the $\|\cdot\|^2$ norm were used, the energy theorem would not be applicable and the corresponding easy transition from a spatial domain integral to a corresponding frequency domain integral would not be possible. As the following discussion illustrates, this easy transition is a powerful argument in favor of the squared RMS metric. That is, the insight provided by equation (7c) is profound.

- Both terms in the integral are non-negative. Therefore,

$$\|f-g\|^2 = 0 \iff |1-\hat{h}(\mu,\nu)|^2|\hat{f}(\mu,\nu)|^2 = 0 \quad \text{for } all \ (\mu,\nu).$$

5

- Image blur is significant $\iff$ the scene has significant energy $|\hat{f}(\mu, \nu)|^2$ at spatial frequencies (μ, ν) where the optical transfer function (OTF) $\hat{h}(\mu, \nu)$ is significantly different from 1.

- Although the scene energy tends to decrease rapidly with increasing spatial frequency, most "natural" scenes have energy at *all* spatial frequencies. That is, natural scenes are not band-limited.

- The OTF typically decreases smoothly in magnitude from 1 at low spatial frequencies to 0 at high frequencies. Thus image blur is caused by a suppression of moderate to high spatial frequencies.

All these observation are well-known. However, the point is that they follow immediately by inspection of the frequency domain integral equation for $\|f - g\|^2$. This observation is the motivation for a search to find analogous equations for $\|g - f'\|^2$ and $\|f - f'\|^2$.

Sampling

The conventional *continuous-discrete-continuous* (end-to-end) model of image gathering, sampling and reconstruction is the convolution equation

$$f'(x, y) = \sum_m \sum_n r(x - m, y - n) g(m, n) \tag{8a}$$

where f' is the (continuous) reconstructed image and (as before) $g = h * f$. The (discrete-to-continuous) reconstruction process is conventionally assumed to be both linear and shift-invariant. It is therefore completely characterized by the reconstruction point spread function r conventionally normalized so that

$$\int_x \int_y r(x, y) dy \, dy = 1. \tag{8b}$$

This PSF can be thought of as the (continuous) output corresponding to a (discrete) sampled input which is 1 at the origin ($m = n = 0$) of the sampling grid and 0 at all other grid points. The reconstruction function is a low-pass filter which accounts for the combined effects of all post-sampling operations such as resampling and display.

The (continuous-to-discrete) sampling process is linear. However,

sampling is **not** a shift-invariant process.

That is, sampling causes the end-to-end system to be shift-variant. This sample-scene phase dependence complicates the end-to-end analysis significantly. For example, the end-to-end fidelity loss expression that one would write by analogy with equation (7c) is

$$\|f - f'\|^2 \neq \int_\mu \int_\nu |1 - \hat{h}(\mu, \nu) \hat{r}(\mu, \nu)|^2 |\hat{f}(\mu, \nu)|^2 d\mu \, d\nu.$$

However (except in special cases) this equation is *not* correct.

Although the end-to-end model is not shift-invariant, it can be demonstrated that by using sample-scene phase averaging the metrics $\|f - f'\|^2$ and $\|g - f'\|^2$ can be written as

$$\|f - f'\|^2 = \int_\mu \int_\nu [\text{non-negative}] \, |\hat{f}(\mu, \nu)|^2 \, d\mu \, d\nu \tag{9a}$$

and

$$\|g - f'\|^2 = \int_\mu \int_\nu [\text{non-negative}] \, |\hat{g}(\mu, \nu)|^2 \, d\mu \, d\nu. \tag{9b}$$

Sample-Scene Phase Averaging

As first established in references [1] and [2], sample-scene phase averaging consists of the following steps.

- Fix the sampling grid.

- Shift the scene a random amount (u, v) relative to the fixed sampling grid

$$f(x, y) \rightarrow f(x - u, y - v).$$

- Calculate (in the frequency domain) the corresponding shifted pre-sampling image

$$g(x, y) \rightarrow g(x - u, y - v)$$

and reconstructed image

$$f'(x, y) \rightarrow f'(x, y; u, v).$$

- Assume that the random u and v shifts are independently and uniformly distributed between 0 and 1.

- Calculate (in the frequency domain) the expected values

$$E\left[\|f - f'\|^2\right] = \int_0^1 \int_0^1 \|f - f'\|^2 \, du \, dv$$

and

$$E\left[\|g - f'\|^2\right] = \int_0^1 \int_0^1 \|g - f'\|^2 \, du \, dv.$$

- Observe that the image blur is independent of the sample-scene phase so that

$$E\left[\|f - g\|^2\right] = \|f - g\|^2.$$

The results of this process are expected value equations consistent with $(9a)$ and $(9b)$.

Sampling and Reconstruction Blur

By using sample-scene phase averaging, it can be shown that

$$E\left[\|g - f'\|^2\right] = \int_\mu \int_\nu \left[|1 - \hat{r}(\mu, \nu)|^2 + \sum_m \sum_n{}' |\hat{r}(\mu - m, \nu - n)|^2\right] |\hat{g}(\mu, \nu)|^2 \, d\mu \, d\nu \quad (10)$$

where the double summation is over all $(m, n) \neq (0, 0)$. However, an algebraically equivalent representation provides more insight into the fidelity loss associated with sampling and reconstruction. That is

$$E\left[\|g - f'\|^2\right] = \epsilon_s^2 + \epsilon_r^2 \tag{11a}$$

where

$$\epsilon_s^2 = \int_\mu \int_\nu \left[\sum_m \sum_n{}' |\hat{g}(\mu - m, \nu - n)|^2\right] |\hat{r}(\mu, \nu)|^2 d\mu \, d\nu \tag{11b}$$

and

$$\epsilon_r^2 = \int_\mu \int_\nu |1 - \hat{r}(\mu, \nu)|^2 |\hat{g}(\mu, \nu)|^2 \, d\mu \, d\nu. \tag{11c}$$

These two terms can be interpreted as follows [1].

- The term ϵ_s^2 accounts for aliasing caused by undersampling; it measures the loss of fidelity caused by the folding of significant image energy $|\hat{g}(\mu, \nu)|^2$ beyond the Nyquist frequency into those (low) frequencies where the reconstruction filter response $\hat{r}(\mu, \nu)$ is not 0. Moreover

$$\epsilon_s^2 = 0 \iff \left[\sum_m \sum_n{}' |\hat{g}(\mu - m, \nu - n)|^2\right] |\hat{r}(\mu, \nu)|^2 = 0 \qquad \text{for all } (\mu, \nu).$$

- The term ϵ_r^2 accounts for imperfect reconstruction; it measures the loss of fidelity caused by the presence of significant image energy at those (high) frequencies where $\hat{r}(\mu, \nu)$ is not 1. Moreover

$$\epsilon_r^2 = 0 \iff |1 - \hat{r}(\mu, \nu)|^2 |\hat{g}(\mu, \nu)|^2 = 0 \qquad \text{for all } (\mu, \nu).$$

- If it were possible to produce a truly band-limited and sufficiently sampled image g, and if the reconstruction function was then taken to be $r(x, y) = \text{sinc}(x)\text{sinc}(y)$ then these two terms would be 0. (This is the sampling theorem.)

End-To-End Blur

In a similar manner, by using sample-scene phase averaging it can be shown that $E\left[\|f - f'\|^2\right]$ is

$$\int_\mu \int_\nu [\text{non-negative}] |\hat{f}(\mu, \nu)|^2 \, d\mu \, d\nu \tag{12a}$$

where the [non-negative] term is

$$\left[|1 - \hat{h}(\mu,\nu)\hat{r}(\mu,\nu)|^2 + |\hat{h}(\mu,\nu)|^2 \sum_m \sum_n {}' |\hat{r}(\mu - m, \nu - n)|^2 \right] \qquad (12b)$$

and again the summation is over all $(m,n) \neq (0,0)$. Also, as before, an algebraically equivalent representation provides more insight into the end-to-end fidelity loss. That is

$$E\left[\|f - f'\|^2\right] = \epsilon_s^2 + \epsilon_e^2 \qquad (13a)$$

where ϵ_s^2 is the sampling (aliasing) term defined previously and

$$\epsilon_e^2 = \int_\mu \int_\nu |1 - \hat{h}(\mu,\nu)\hat{r}(\mu,\nu)|^2 |\hat{f}(\mu,\nu)|^2 \, d\mu \, d\nu. \qquad (13b)$$

This new term can be interpreted as follows.

- It accounts for the end-to-end loss of fidelity caused by significant scene energy at (mid to high) frequencies where the cascaded response, $\hat{h}(\mu,\nu)\hat{r}(\mu,\nu)$ is not 1. Moreover,

$$\epsilon_e^2 = 0 \iff |1 - \hat{h}(\mu,\nu)\hat{r}(\mu,\nu)|^2 |\hat{f}(\mu,\nu)|^2 = 0 \qquad \text{for all } (\mu,\nu).$$

- It measures how well the reconstruction filter $\hat{r}$ is able to "deblur" (restore) those spatial frequencies which were suppressed prior to sampling by the image gathering OTF $\hat{h}$.

There is an inevitable trade-off here. For a fixed scene f and sampling grid, any attempt to decrease ϵ_e^2 by modifying $\hat{h}$ and $\hat{r}$ will result in an increase in ϵ_s^2 and conversely.

Fidelity Loss Budget

All of the previous analysis can be summarized in a *fidelity loss budget* given by the three sample-scene phase averaged metrics

$$E\left[\|f - g\|^2\right] = \epsilon_i^2 \qquad (14a)$$

$$E\left[\|g - f'\|^2\right] = \epsilon_s^2 + \epsilon_r^2 \qquad (14b)$$

$$E\left[\|f - f'\|^2\right] = \epsilon_s^2 + \epsilon_e^2 \qquad (14c)$$

where

$$\epsilon_i^2 = \int_\mu \int_\nu |1 - \hat{h}(\mu,\nu)|^2 \, |\hat{f}(\mu,\nu)|^2 \, d\mu \, d\nu \qquad (14d)$$

$$\epsilon_s^2 = \int_\mu \int_\nu |\hat{h}(\mu,\nu)|^2 \left[\sum_m \sum_n {}' |\hat{r}(\mu - m, \nu - n)|^2 \right] |\hat{f}(\mu,\nu)|^2 \, d\mu \, d\nu \qquad (14e)$$

$$\epsilon_r^2 = \int_\mu \int_\nu |\hat{h}(\mu,\nu)|^2 \, |1 - \hat{r}(\mu,\nu)|^2 \, |\hat{f}(\mu,\nu)|^2 \, d\mu \, d\nu \qquad (14f)$$

$$\epsilon_e^2 = \int_\mu \int_\nu |1 - \hat{h}(\mu,\nu)\hat{r}(\mu,\nu)|^2 \, |\hat{f}(\mu,\nu)|^2 \, d\mu \, d\nu. \qquad (14g)$$

The four ϵ^2 terms can be easily calculated via numerical integration. All that is required is a knowledge of the scene energy $|\hat{f}(\mu,\nu)|^2$, image gathering OTF $\hat{h}(\mu,\nu)$ and reconstruction filter $\hat{r}(\mu,\nu)$—and ready access to a computer with a fast CPU and sufficient memory.

The four ϵ^2 terms are all interrelated and any attempt to minimize one must be carefully weighted against the potential increase of the others. Trade-off studies like this are the stuff of digital imaging system design.

References

[1] S.K. Park and R.A. Schowengerdt
Image Sampling, Reconstruction, and the Effect of Sample-Scene Phasing
Applied Optics, 21, 3142–3151, 1982.

[2] S.K. Park and R.A. Schowengerdt
Image Reconstruction by Parametric Cubic Convolution
Computer Vision, Graphics and Image Processing, 23, 258–272, 1983.

[3] R.A. Schowengerdt, S.K. Park and R. Gray
Topics in the Two-dimensional Sampling and Reconstruction of Images
International Journal of Remote Sensing, 5, 2, 333–347, 1984.

[4] S.K. Park, R.A. Schowengerdt and M.A. Kaczynski
Modulation-Transfer-Function Analysis for Sampled Image Systems
Applied Optics, 23, 2571–2582, 1984.

[5] E.H. Linfoot
Quality Evaluations of Optical Systems
Optica Acta, 5, 1–14, 1958.

[6] F.O. Huck, C.L. Fales, N. Haylo, R.W. Samms and K. Stacy
Image Gathering and Processing: Information and Fidelity
Journal of the Optical Society of America, A2, 1644–1666, 1985.

[7] F.O. Huck, C.L. Fales, J.A. McCormick and S.K. Park
Image-Gathering System Design for Information and Fidelity
Journal of the Optical Society of America, A5, 285–299, 1988.

Bibliography

[A] P.B. Fellgett and E.H. Linfoot
On the Assessment of Optical Images
Philosophical Transactions of the Royal Society of London 247, 369–407, 1955.

[B] F.O. Huck and S.K. Park
Optical-Mechanical Line-Scan Imaging Process: Its Information Capacity and Efficiency
Applied Optics, 14, 2508–2520, 1975.

[C] F.O. Huck, N. Haylo and S.K. Park
Aliasing and Blurring in 2-D Sampled Imagery
Applied Optics, 19, 2174–2181, 1980.