Chapter 9

Discrete Sampling of Continuous Images

Fourier transform pairs. Image Blurring vs Image Sharpening revisited. Sampling theory in a nutshell. Nonlinearities in the human visual system.

9.1 Introduction

We're now in a position to put several of the previous lectures together, to get a really important result: the Shannon sampling theorem. In words: how is it possible to go from a continuous signal to a sampling of it? (Think about the distinct receptors in your retina; the integers on a music CD, or the ommatidia from Limulus.) It's almost magic, because (as we saw last time) there is a real sense in which continuous signals live in infinite-dimensional spaces, while sampled ones live in finite-dimensional ones! In the course of understanding this, we'll revisit our question from two lectures ago: can you undo blurring by sharpening?

Our analysis also provides a new (and quite central) role for the δ-function, because it suggests using it as part of the sampling process – in fact, we'll use a 'train' of them to get a sequence of samples.

To start, let's examine some Fourier transforms of certain functions.

9.2 Fourier Transform Pairs

To build further intuition regarding Fourier Transforms, it is important to note a few pairs of temporal (or 1D) functions and their transforms in the frequency domain; these are shown in Fig. 9.1.

The first of these is the δ-function. Recall that this had "infinite" mass distributed across "zero" width on the real number line, such that it's integral evaluated to 1. Also remember that the product of a delta function with another 'test' function, $f(x)$,

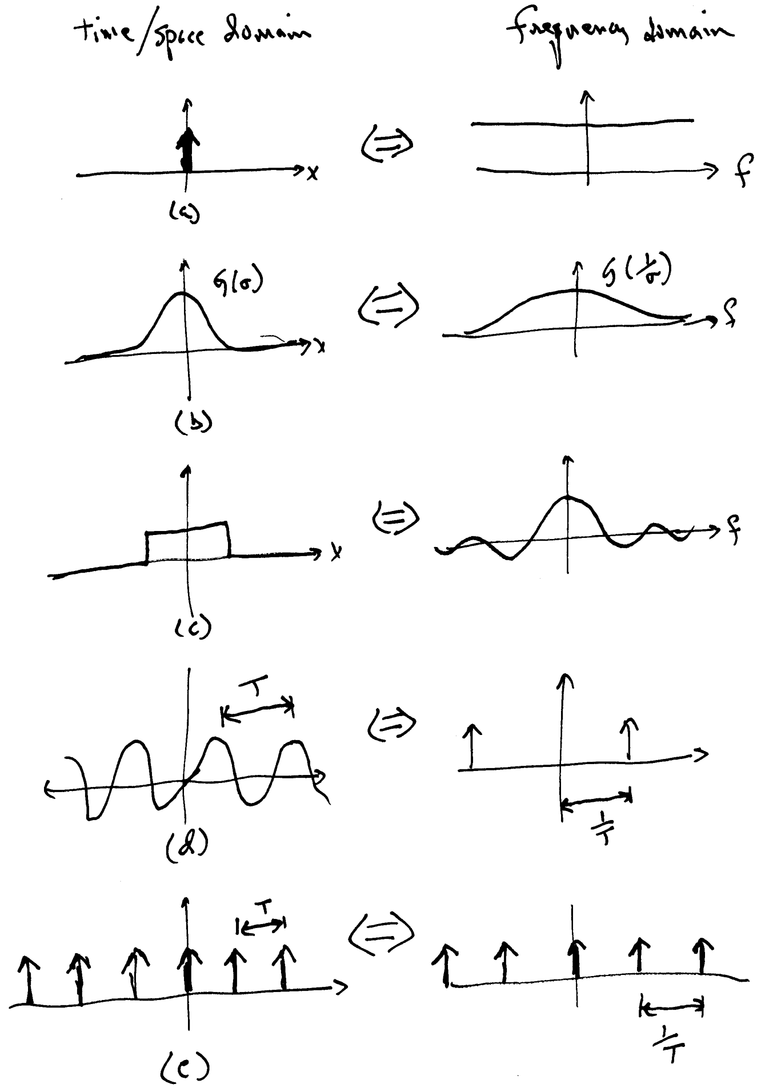


Figure 9.1: Fourier transform pairs. Signals in the time domain are shown on the left and the frequency domain on the right. The Dirac $\delta(x)$-function transforms to a constant across all frequencies, which is another expression of its physical impossibility. Taking a value from the limiting process that we used in deriving it, we note that a Gaussian in time transforms to a Gaussian in frequency. Squaring off the sides of the Gaussian to a rectangular pulse, or boxcar function adds ripples in frequency; it's pair is the sinc function. The sine, of course becomes a δ in frequency (negative frequency also shown for technical reasons) and the comb function, or a train of δ's, transforms to a train of δ's. Note, in particular, how the period between them is inverted; this will be important in the next lecture.

evaluated to the value of the test function at 0 (that is, f(0). This sifting property will become imporant as a basis for sampling, shortly.

We had earlier talked about the δ-function being unrealizable in the limit, and it's Fourier transform (FT) illustrates this in a new light. To get something as concentrated as a delta function, there are always additional "error" terms that would have to be made up, and these never get smaller. Thus, its FT is a constant. (It's an easy exercise to work this out for yourself.)

Next we have the Gaussian in space. That is, while it's concentrated in position, it's not totally concentrated like the delta function. Its FT is also a Gaussian, which you might have guessed from our experiments in the previous lecture – remember, the Gaussian was a model for image blur. In the frequency domain we can think about noise differently. Imagine a "clean" image of a face and add (pointwise addition) some 'salt and pepper noise' (that is, some isolated bright and dark values to the pixel intensities.) Since the noise is indepentent across position, while the image is not (nearby intensities are not unrelated to one another!) then the noise will appear as additional high-frequency components. Gaussian blur will remove this. Previously we thought about this removal as an averaging process across position; now we can think of it as filtering (reducing) the high-frequency content of the image.

In case you've forgotten, we can write the Gaussian (centered at the origin) as

$$G(x) = \frac{1}{2\pi\sigma^2}\exp(-\frac{x^2}{2\sigma^2})$$

where σ is a parameter that controls the spread. Note the inverse relationship between the space and the frequency domains. We'll have lots more to say about this Gaussian in future lectures.

If you square off the Gaussian you get a pulse, and this 'squaring' of the corners has an important effect: it introduces ripples in frequency. The RECT(X) pulse has the (normalized) sinc function ($sin(\pi x)/(\pi x)$) as its FT pair. Like the Gaussian, there's a reciprical relationship here – as the rect function gets more concentrated the sinc spreads out. Conceptually: concentrating more in space leads to an increased support requirement in frequency. Can you guess the limit, e.g. when the width of the rect approaches 0? (You might find a hint in this table of FT pairs.) Notice how interesting it is to compare this sinc function, with its zero-crossings, to the constant developed above for the delta function.

The next example is the 'easiest' function in frequency: a pure frequency is a sinusoid in space. Actually, there's a bit of a technicality here because Euler's identity comes in again – have a look back at Fig. 8.13(d). Now we have to turn it around, yielding:

$$
\begin{aligned}
\cos x = & \quad \text{Real part}\{e^{ix}\} & = \frac{e^{ix} + e^{-ix}}{2} \\
\sin x = & \quad \text{Imaginary part}\{e^{ix}\} & = \frac{e^{ix} - e^{-ix}}{2i}.
\end{aligned}
$$

These explain the technical point about why there are two delta functions, one in the positive frequency range and the other in the "negative frequency" side; we won't worry any further about these technicalities here, though.

Finally, we have the COMB FUNCTION, or a train of Dirac delta's:

$$
\mathbf{comb}(t) = \sum_{k=-\infty}^{k=\infty} \delta(t - kT)
$$

which will be incredibly important when we do sampling. It's FT is another comb function, this time in frequency with, again, a reciprical relationship to spacing.

It's interesting to compare the comb example to the single impulse with which we began. That, you remember, had a constant representation in the frequency domain – now we have a series of delta functions that span frequency space. Here's the huge difference: the comb function is periodic, the delta function is not. This comb function will clearly play a central role in the sampling theorem because it is the mechanism for "sifting" out the periodic samples.

9.3 Image Deblurring

With these Fourier pairs in mind, we can now return to the question posed two lectures ago: what happens when an image blurring operation is composed with an image sharpening operation, in either order. In particular, we asked whether:

Blur (Sharpen (Image)) $\stackrel{?}{=}$ Sharpen (Blur (Image)) $\stackrel{?}{=}$ Image

We're now in a position to address this a little more seriously. To be concrete, assume that the blur is implemented by a Gaussian convolution. This is an interesting problem because the blur of the lens for a "pretty good" human eye is about $\sigma = 2$ minutes of arc when the pupil is about 3 mm dilation. It also comes up in practical applications, such as astronomy and CT scanners.

Let $\mathcal{G}$ denote the Gaussian blurring kernal and $\mathcal{I}$ denote the image, so that the question becomes:

$$
\mathcal{G} * \mathcal{H} * \mathcal{I} \stackrel{?}{=} \mathcal{H} * \mathcal{G} * \mathcal{I} \stackrel{?}{=} \mathcal{I} \tag{9.1}
$$

where we have introduced the symbol $\mathcal{H}$ for the deblurring operation. A quick glance at the above shows the equation will work provided $\mathcal{H} = \mathcal{G}^{-1}$, and we have:

$$
\mathcal{G} * \mathcal{G}^{-1} = \mathcal{G}^{-1} * \mathcal{G} = \mathbf{1}
$$

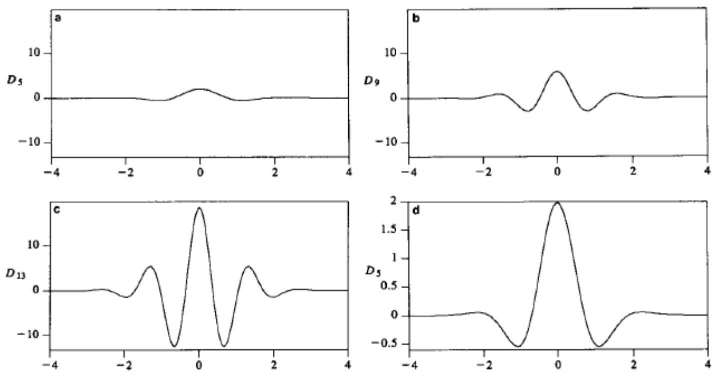


Figure 9.2: Gaussian deblurring kernels for an image that can be represented as a (a) 5-th, (b) 9-th, and (c) 13-th order polynomial. (d) rescaled version of (a), to emphasize its structure. Notice how it resembles the lateral inhibitory kernel for Limulus linearis. Figures from Hummel.

where **1** is the identity operator. Thus the question becomes: does an inverse to the Gaussian blur operator exist? The identities above give us a hint: since convolution in space is multiplication in frequency, we would need something like (1/Gaussian) for this inverse operator. But while the Gaussian drops off in frequency reasonably fast, 1/(a small number) can be huge! So the inverse to the Gaussian operator doesn't exist as a general convolution operator.

However, all is not lost. If we could restrict the space of images to those that were "very well behaved," for example could be modeled by a low-order polynomial (or a sum of low spatial frequencies), then forms for this operator can be found; see Fig. 9.2. Their application is shown in Fig. 9.3.

9.4 Sampling Theory

We are at the point where we shall develop the Shannon theory of sampling for continuous functions. The analysis makes direct use of the Fourier material and the (trains of) delta functions because we move from the spatial (or temportal) domain to the frequency domain and back again several times. The discussion follows Fig. 9.5 so you should refer to this while reading the text.

We begin with a smooth function $f(x)$, in space,, with a Fourier transform that is *bandlimited* to the maximum frequency component w. The sampling operation itself employs a sequence of *delta* functions, because (please remember) these have the property of sifting out the value at the point where its argument is 0. To get a sequence of samples, then, we use a sequence of delta functions, which is called a *comb* function $s(x)$; it is shown in space and in frequency. Notice, in particular, that

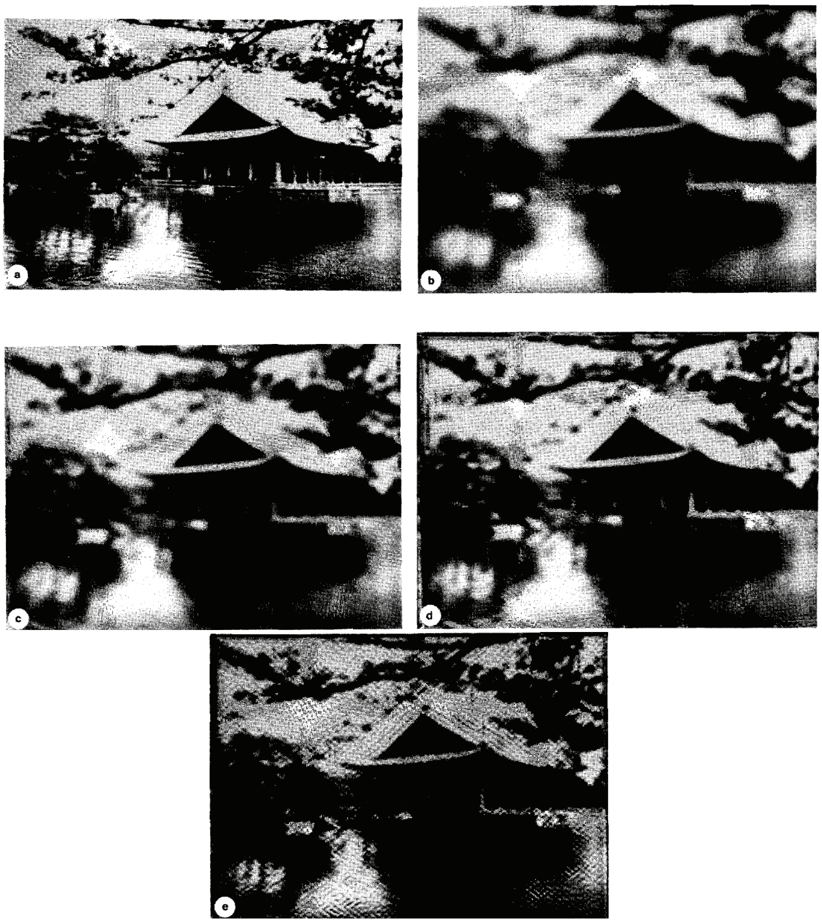


Figure 9.3: Illustration of Gaussian deblurring.

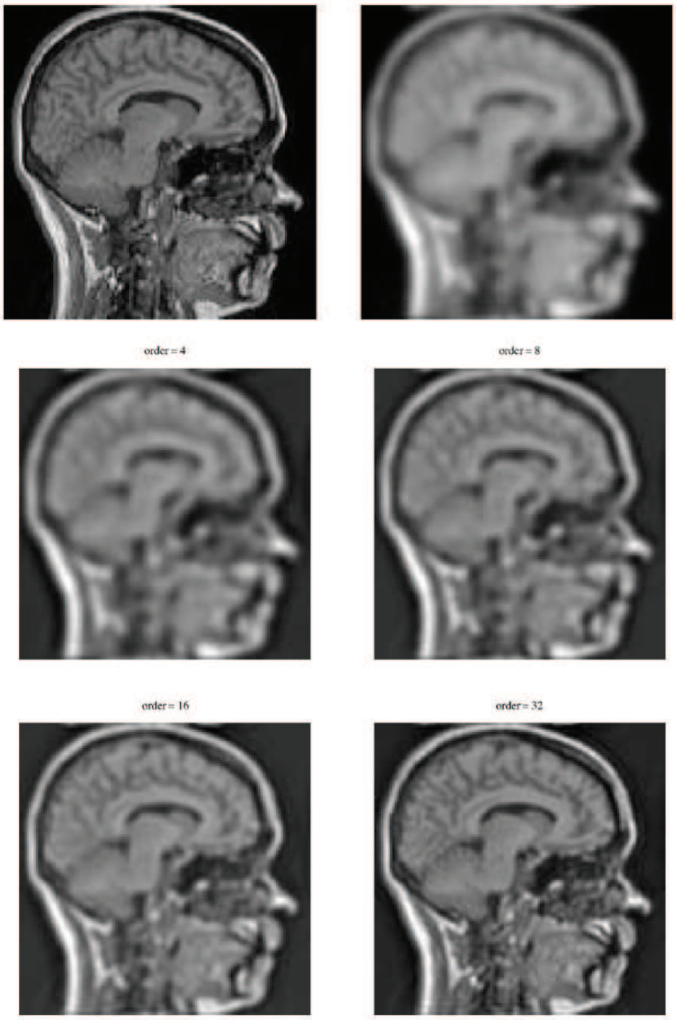


Figure 9.4: Another illustration of Gaussian deblurring for a CT image (top) Original (128 x 128) pixels and blurred ($\sigma = 2$) images. (bottom) different deblurring kernels; the lower right is N = 32.

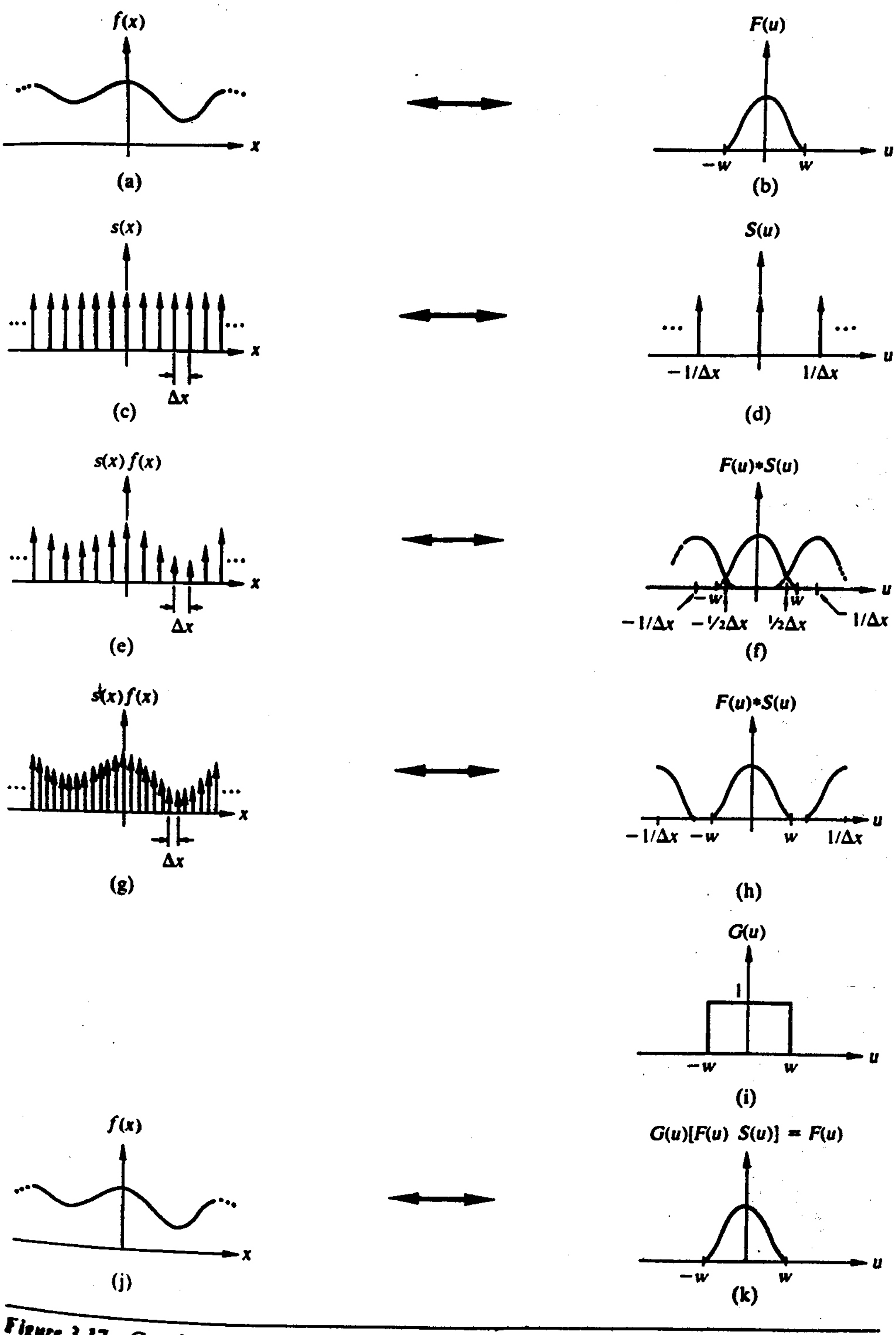


Figure 3.17 Graphic development of sampling concepts.

Figure 9.5: Illustration of the sampling theorem. The left side shows computations in the spatial domain x and the right side in frequency u. See text for discussion.

the Fourier transform of the comb function in space is a comb function in frequency $S(u)$, but the spacing of the delta functions is inverted; delta's space Δx apart in space implies frequency spacing of $(\Delta x)^{-1}$. In other words, closer in space implies further apart in frequency.

Now, the actual sampling operation is taken as the product of the comb function in space with the function being sampled. Since a product in one domain implies convolution in the other (for linear systems, such as this one!), we can readily compute the Fourier transform of the sampled function as the convolution. This reveals an important point: if the samples are too close in space, then the frequency content of the original function will overlap in frequency; when these terms add up a distortion is introduced called "aliasing". (Examples of aliasing will be shown shortly, in the discussion of subsampling.)

To spread out the samples in the frequency domain we have to sample more frequently in the spatial domain. The natural requirement, of course, is to sample just as frequently as required to avoid aliasing, which is twice the maximum frequency component w. (This is the *Nyquist criterion.*) Note that the factor of two comes from the inclusion of "negative" frequencies in the top panel of the figure. This yields the sequence of sampled values of the original function at a rate sufficent to recover it exactly. Sampling even more frequently wouldn't add anything, because the support in the frequency domain would just spread out further.

Finally, if we wish to recover the continuous signal, then a *box filter* in frequency selects just the component of the transform at the origin, whose inverse Fourier transform is exactly the original signal. Multiplying by this rect filter in frequency cancels all of the irrelevant copies of the 'bumps.'

A realization of this final, reconstruction step would be to convolve the sampled data against a sinc function (the inverse FT of the rect). For this reason the sinc function is often called a reconstruction filter.

Recall that we discussed sampling of audio signals in class at about 44KHz; or twice the frequency limit of the cochlea in our ears.

9.5 Images as a function of space and time

We now consider (moving) images; i.e., images that vary as a function of time, $I(x, y, t)$. These now require three coordinates; see Fig. 9.6. When sampled, you can think of this as a "stack" (in time) of images. What is the structure that should emerge from an analysis of time-varying imagery?

The discussion of the sampling theorem focussed on how we could sample in time, for 1-D acoustic signals, or in 2D space, for images. But now we've got to do both: sample a time varying image so that, for each instant in time it's a 2D spatially-sampled image; or for a given point it's a 1D time sampled sequence. Of course, putting these together yields the "3D" (that is, 2D x 1D) cube.

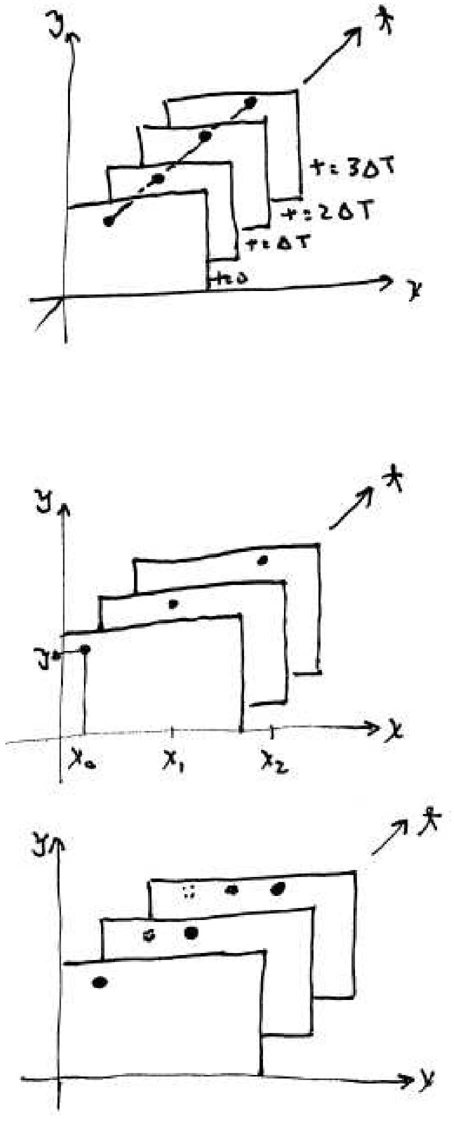


Figure 9.6: Time-varying images can be thought of as a kind of image stack, in which (sampled or continuous) images are arrayed along the temporal axis. (top) Images at a spacing of ΔT along the temporal axis are shown. A single point fixed in position in the image is also shown near the top; notice how this point traces out a curve through space-time that is a straight line parallel to the t-axis and orthogonal to the image coordinate (x, y)-plane. (middle) Movement of the point in a direction parallel to the x-axis is shown; there is no movement component in the y-direction. (bottom) In general, spatio-temporal filters will cause some change in the images such as smearing in the spatial and temporal dimensions. Finite response time for photoreceptors is one source of such motion blur.

Figure 9.7: Motion Smear. Image taken with an integration time of 125 msec. Notice how much blur there is; how elongated objects are in the direction of blur. Why don't we see this blur when viewing the world? Figure from Burr, Nature, 1980.

Reconstruction from sampled images in time gives rise to APPARENT MOTION, or the appearance of continuous motion from a sequence (in time) of static frames. In general, then, we are going to have to understand spatio-temporal reconstruction filters.

9.5.1 Motion Smear

The photoreceptors in our eyes have an integration time that allows for the summation of photon impacts. (Photons arrive with Poisson distribution?) This is a clear example of "averaging" to reduce noise, but it illustrates the tradoff that arises: averaging also blurs the image; see Fig. 9.7.

As we've been discussing, it is natural to postulate "filters" that are the 'inverse' of these blur filters to anhililate their effects. One way to do this is to estimate what the deblurring filter would look like and then apply it to such motion-smeared images. A psychophysical attempt to do so my Michael Morgan revealed the filter in Fig. 9.8(top) which is very similar to what we saw for the fly.

Most curious are some other experiments by David Burr. Upon short duration a moving dot appears as a streak; upon longer duration as a moving dot Fig. 9.8(bottom). How can this change in interpretation be explained?

Artists have attempted to depict the feeling of movement and motion for millenia. In the Chauvet cave there are examples of bison with 8 legs, likely depicting their movment. It is fascinating to observe about 30,000 years of history in Fig. 9.9.

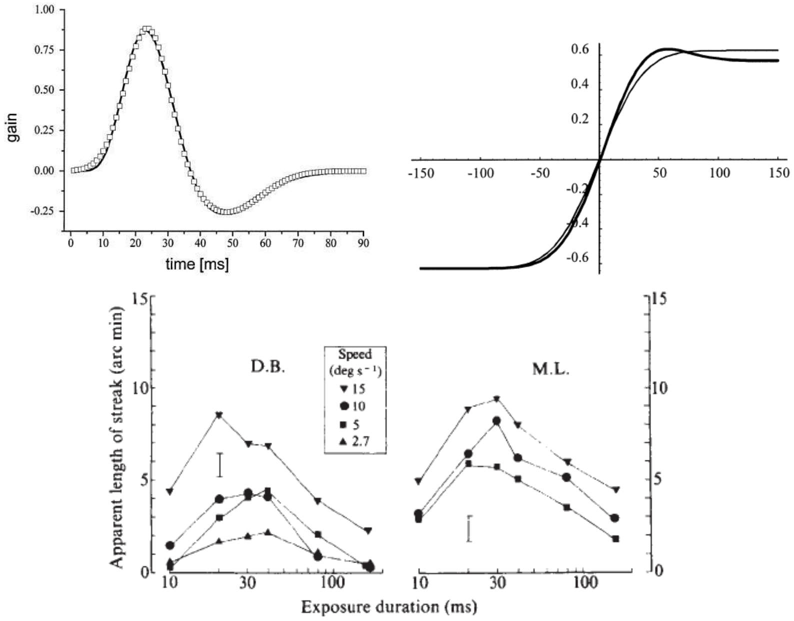


Figure 9.8: Attempts to estimate the sharpening filter. (top) Attempt based on estimating moving sine gratings. Think of each sinusoid as the profile of a blurred edge. Now, as such gratings move, the appearance is a sharpening of the edge – a transformation of the sinusoid toward a square wave grating. The filter estimated is shown at left, and the result of applying this to a blurred edge on right. While this can be made to work to some extent, there is a "best" frequency and contrast for each such filter. However we do not see these (artificial) peaks in deblurring performance, which raises a question of how it is done. Clearly non-linear techniques are required. Figures from morgan (bottom) Attempt based on moving dots. The x-axis is the time that a dot lives on the screen; the y-axis is the apparent length of the streak. Notice that this hits a peak around 30 msec. Most curiously, if the dot lives for longer than about 65 msec it appears not as a streak but as a moving dot. Figure from Burr.

Put some discussion of the movement "filter" in Photoshop here.

9.6 Subsampling an Image

Suppose you have an image that is too large, for example an image of a product that you would like to display in a 'thumbnail' spot. How should this be done?

The first thought is that you can simply choose every other row or column – or every tenth one, if you want to make it smaller. But this doesn't work – see Fig. 9.10. It creates lots of aliased structure, as is especially annoying in the building and checkerboard examples.

The solution is to remember the sampling theorem. If it's already been sampled at the Nyquist rate, then ignoring some samples will mean that it's now being sampled too slowly. The solution: first blur the image, say by convolving with a low-pass (or averaging) filter to reduce the frequency content, then resample. It's common practice to use a Gaussian averaging filter to do this (convolve in space; multiply in frequency). Doing this repeatedly - blur - subsample -blur- subsample - blur - subsample until only a single pixel remains creates a GAUSSIAN IMAGE PYRAMID.

You can also sample incorrectly in time – just look up the WAGON WHEEL ILLUSION on the web. What do you think happens if you sample incorrectly in space and in time?

9.7 Mild Non-linearities

Thus far this lecture has been about a linear system. However, when we introduced the lateral inhibitory network in more realistic terms, please recall θ, the non-linear parameter introduced in eq. 5.5; to repeat this equation:

$$F_i = e_i - \sum_{j \in neigh(i)} \alpha_{i,j} \lfloor e_j - \theta \rfloor . \tag{9.2}$$

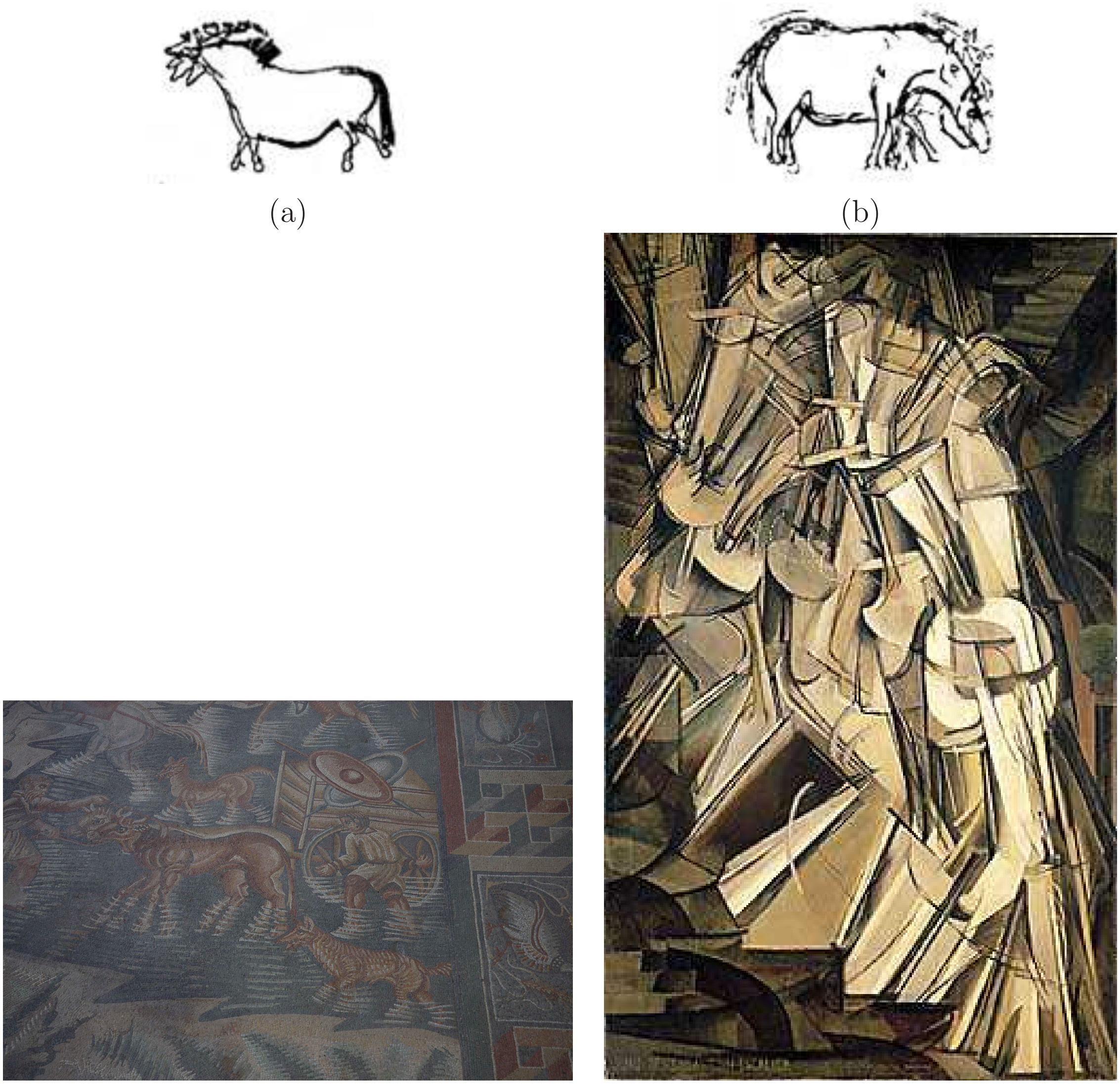
It is interesting to explore whether a small value for θ actually will effect performance; and whether it will be apparent in the MTF; See Fig. 9.11 and compare against Fig. 9.12.

It is clear from this that small values of θ appear to have little effect; while larger ones do lead to quite different behaviors.

This would be a good place to put linear vs. square nonlinearities.

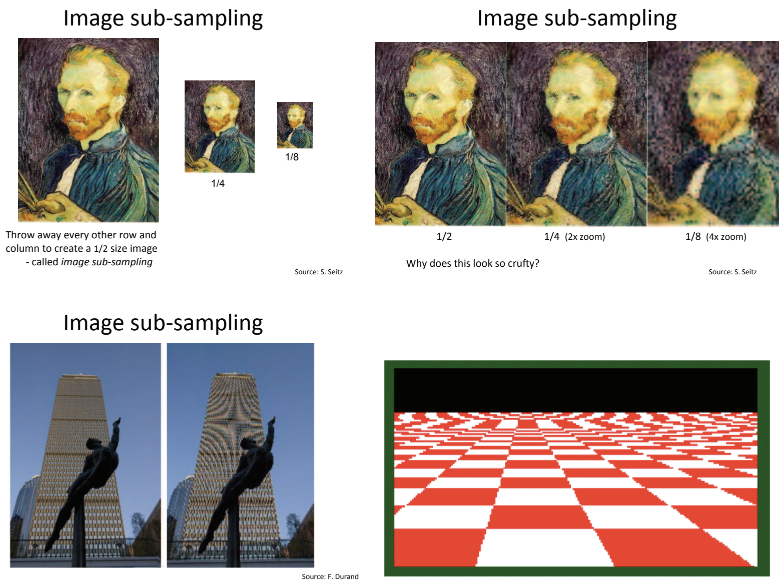
9.8 Frequency filtering, Block Sampling, and Art

The removal of certain frequency bands can be very disruptive to the visual information; see Fig. 9.13.

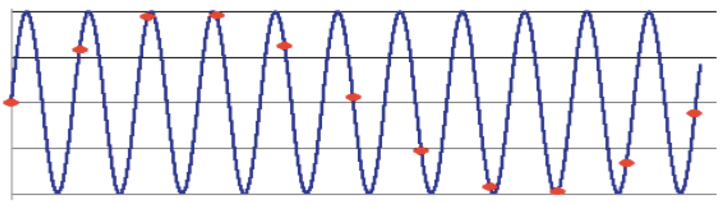


(a) (b) (c) (d)

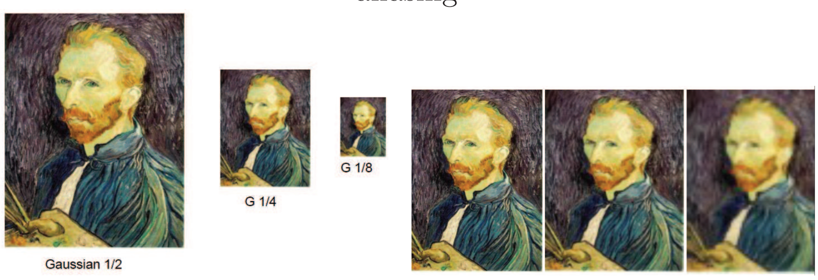
Figure 9.9: Attempts to depict movement in art. Prehistoric artists split movement into multiple images and then superimposed them. (a) Example from Lascaux and (b) LaMarche. From the article: www.bradshawfoundation.com/inora/divers_43_1.html (c) Mosaics in Sicily depicting movement in the Villa Romano del Tellaro, 3 - 4 century; see I Mosaici del Tellaro Lusso e cultura nel Sud-Est della Sicilia di G.Voza (d) About two millenia later Marcel Duchamp's Nude descendant un escalier n 2,1912, Phila Museum of Art. figure from wikipedia.



subsample rows and columns



aliasing



blur then subsample

Figure 9.10: Subsampling an image illustrates the importance of understanding the frequency content in an image. (top) Simply taking every fourth or eigth or whatever rows leads to an awkard effect – aliasing – see middle cartoon. (bottom) prefiltering with a low-pass filter to reduce frequency content and then subsampling works much better. Figures from S. Seitz, F. Durand and L. Zhang

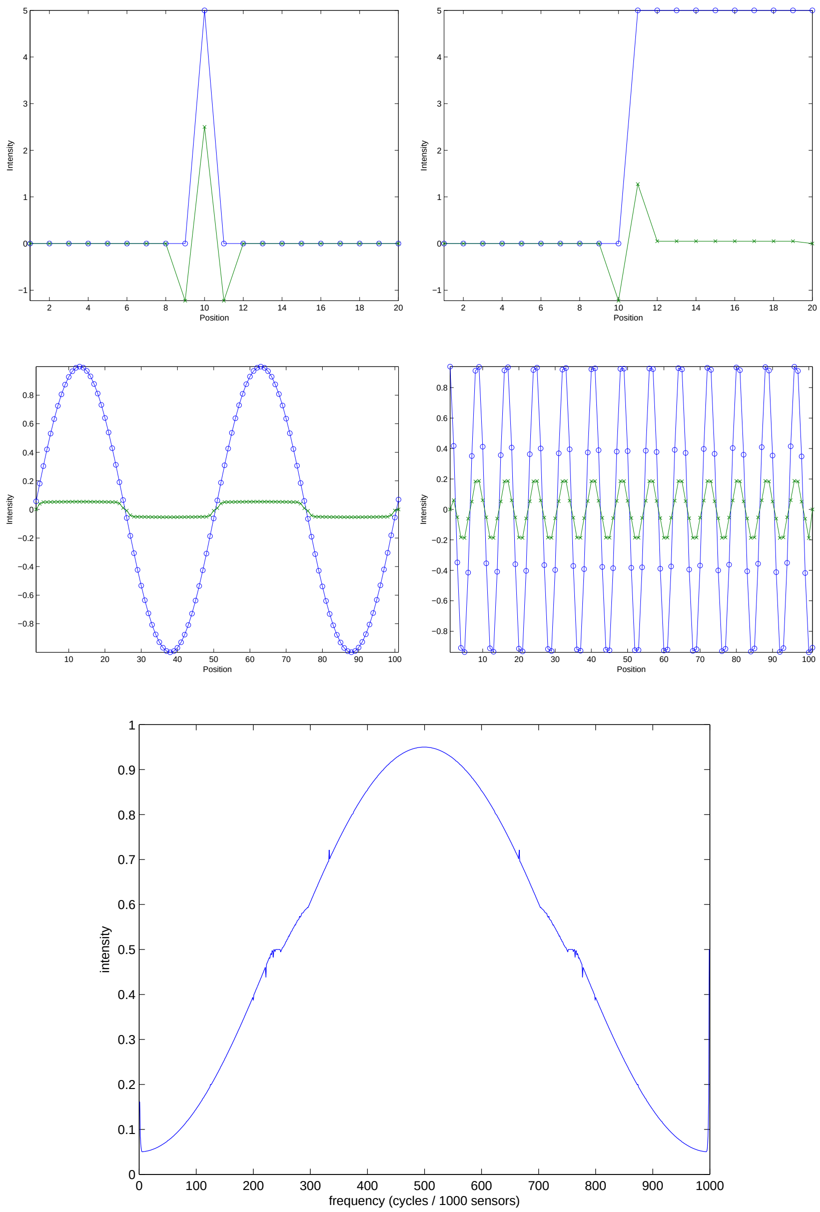


Figure 9.11: Illustration of the response of a lateral inhibitory network with $\theta = 0.1$.

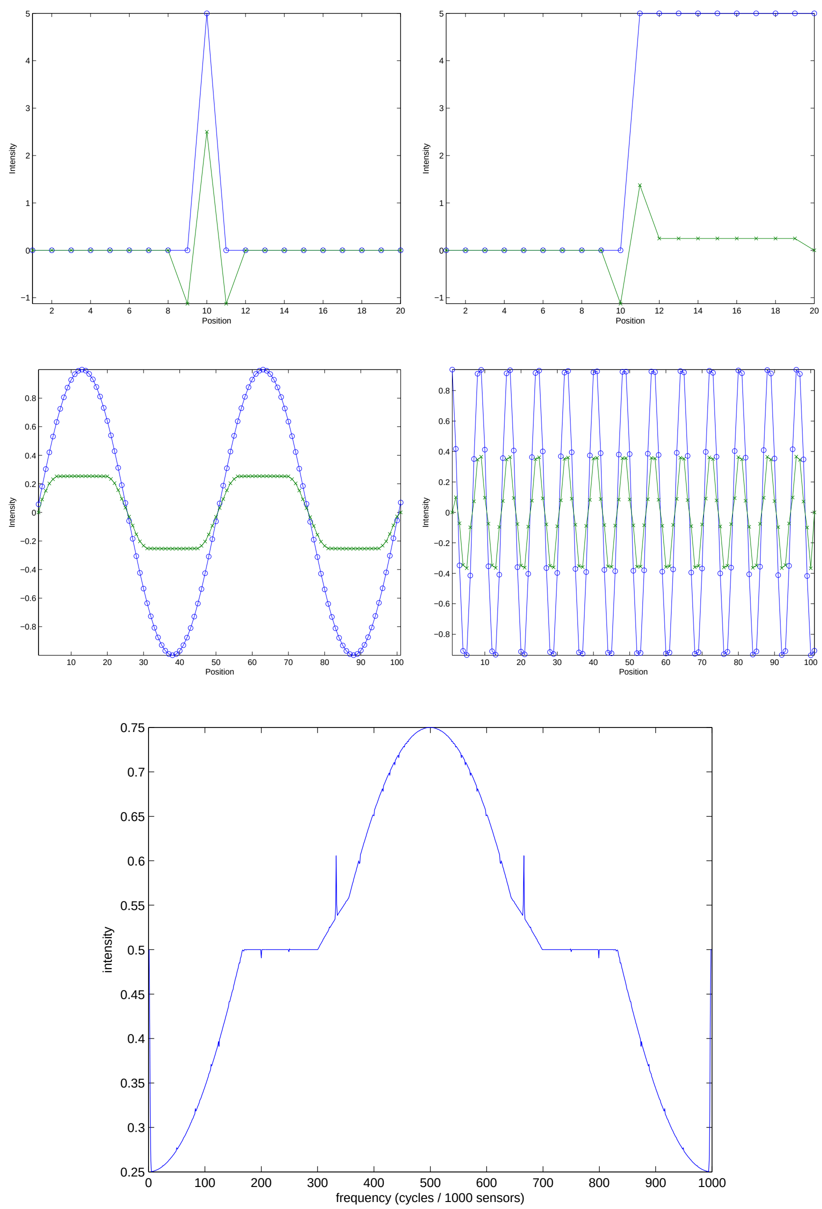


Figure 9.12: Illustration of the response of a lateral inhibitory network with $\theta = 0.5$.

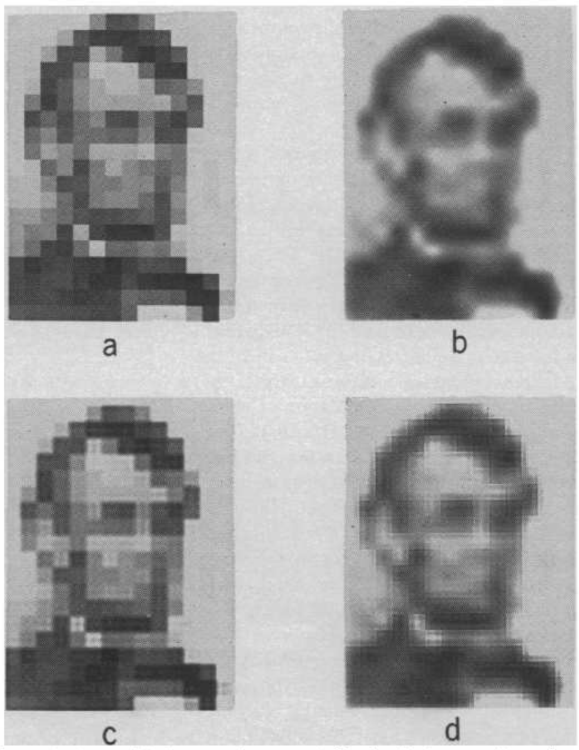


Fig. 1. (a) Computer-processed block portrait. A high-resolution photograph was scanned with a flying spot, digitized, and stored on magnetic tape. A digital computer averaged the brightness values in each of 15 by 20 small square domains. Each of these squares consists of 25 by 25 samples. New, average brightness values were quantized linearly into 16 levels (4 bits), written on digital tape, and converted to analog signals, which were printed on a Facsimile receiver. The print was cropped to 14 by 19 blocks; this portrait is thus represented by 14 $\times$ 19 $\times$ 4 = 1064 bits. (b) The block portrait of (a) was low-pass filtered just above the highest spatial frequency, w, of the picture information present. Since the intact block portrait was 20 blocks high, $w = 10$ cycles. The picture in (b) has been low-pass filtered to $1.26w = 12.6$ cycles. (c) Block portrait of (a) low-pass filtered at $4.06w = 40.6$ cycles. (d) Block portrait of (a) filtered to eliminate the frequency band from $1.22w$ to $3.94w$; that is, all spatial frequencies from 12.2 to 39.4 cycles, roughly the two octaves just above the picture information, have been removed.

Figure 9.13: Famous Harmon-Julesz portrait of Lincoln.Squint your eyes while looking at the different images – what effect does it have? why?

Photography, of course, relies on the fact that we are unable to resolve the grains of the photographic film process, as illustrated in Fig. 9.14. The same basic idea can be seen in works by many artists, including the modern realist Chuck Close; Fig. 9.15, who exploit its expressive qualities.

Unsharp Masking

UNSHARP MASKING is the process of making an image appear sharper by first producing a blurred version of it, and then subtracting this from the original. The result is an intermediate image that resembles an accentuated version of the original, where image intensity changes are magnified. Ideally the accents are placed around the sharpest intensity changes—edges—in the image. The final enhanced result is obtained by adding pixelwise this edge image to the original; see Fig. 9.16.
Can you explain this process given our background in image filtering?

9.9 Non-Linearities Rule

Should we expect the visual system to be linear? Let's consider this at two different levels.

9.9.1 Orientation-based Inhibition

Cells often show no response to a grating at an orientation orthogonal to their preferred one. However, when this grating is superimposed upon another one at the preferred orientation, we have cross-orientation inhibition. This would not be expected if the circuit was simply trying to build up a representation of the image.

It also suggests we start considering *networks of neurons*, rather than individual ones. What function might such networks be implementing, if the task is not one of image reconstruction?

See Shapley/Lennie review.

9.9.2 Non-linear combinations of opaque objects

Projection of opaque objects is highly non-linear. Various completion phenomena.

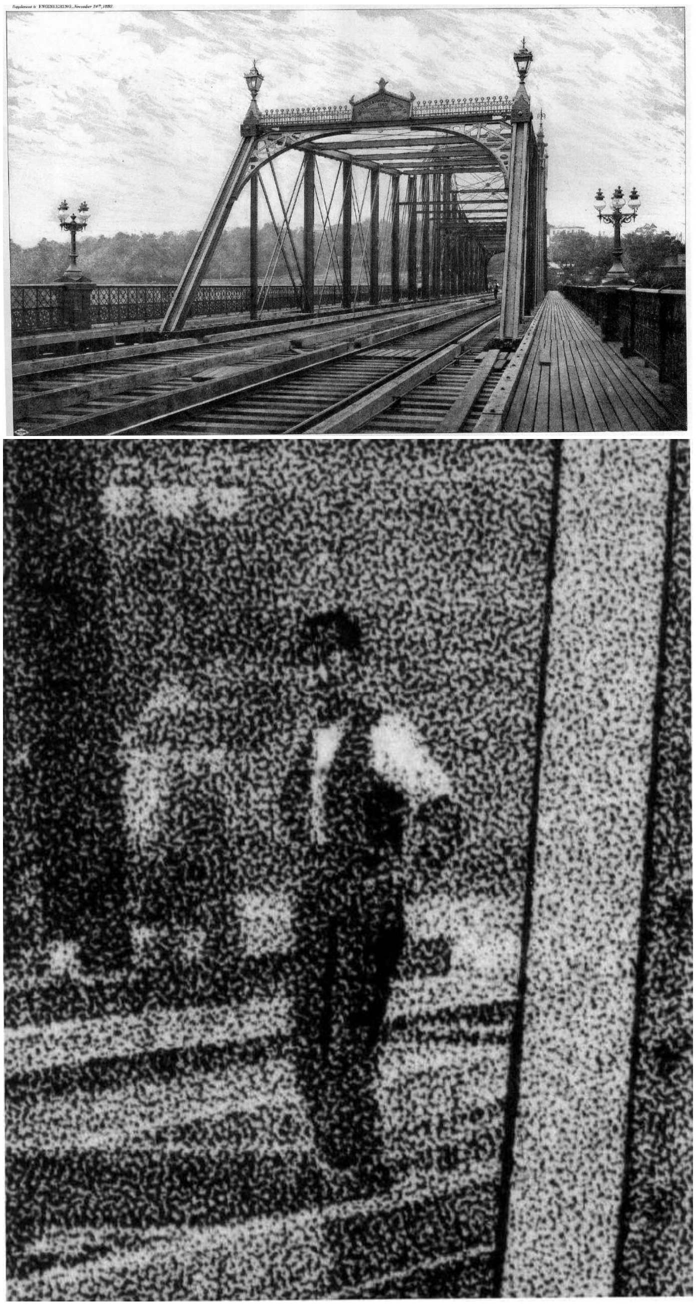


Figure 9.14: The photographic process involves a chemical reaction that deposits grains of silver on a surface. At normal resolutions these grains are below our visual acuity and appear as continuous tonal variations. However, when enlarged (bottom) the grain structure is clear. Example from Benson, The Printed Picture.

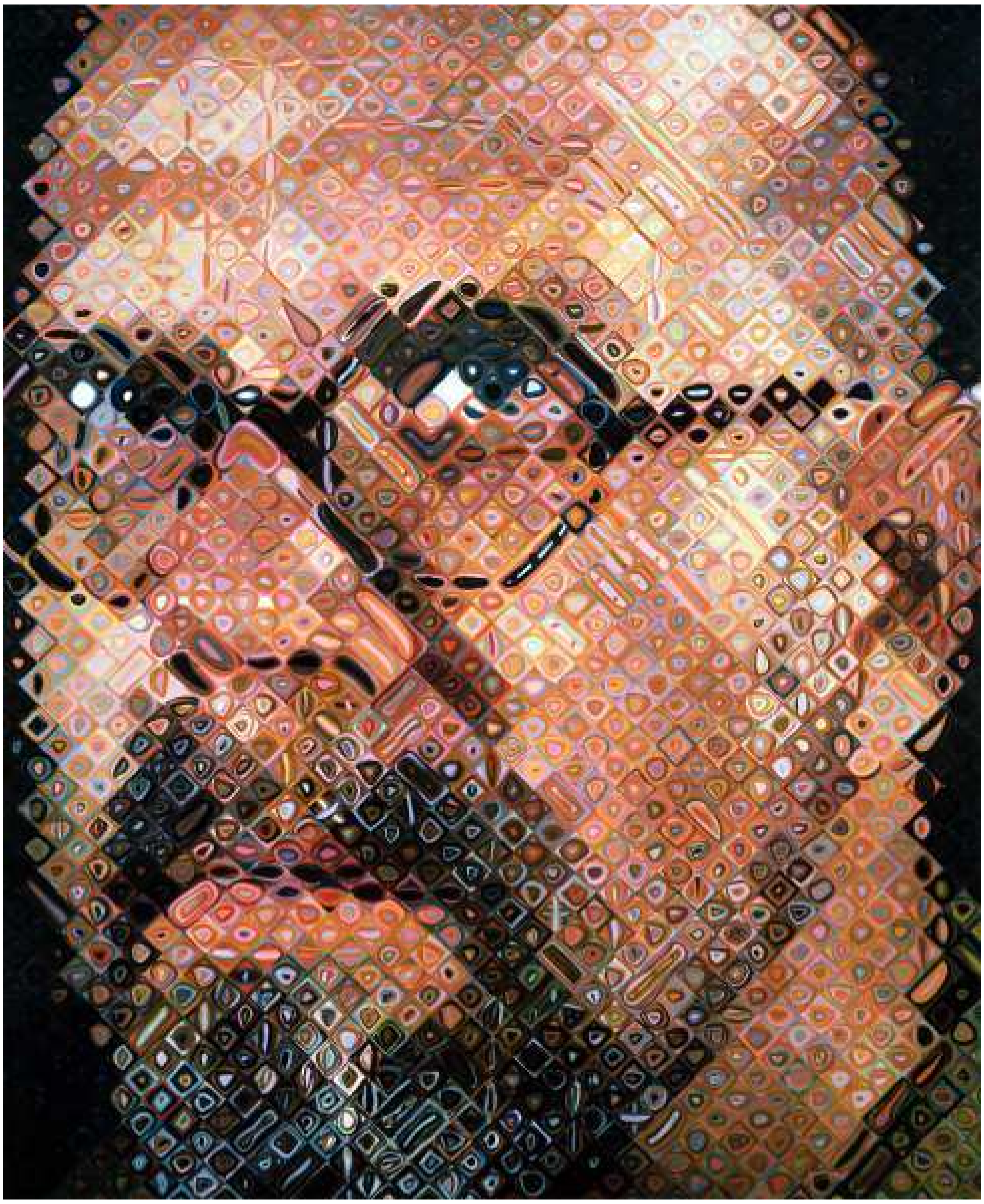


Figure 9.15: Chuck Close, self portrait, 1997. Oil on canvas. MoMA, New York. View the painting from different distances, and by squinting your eyes; find it on the web to see it better color.

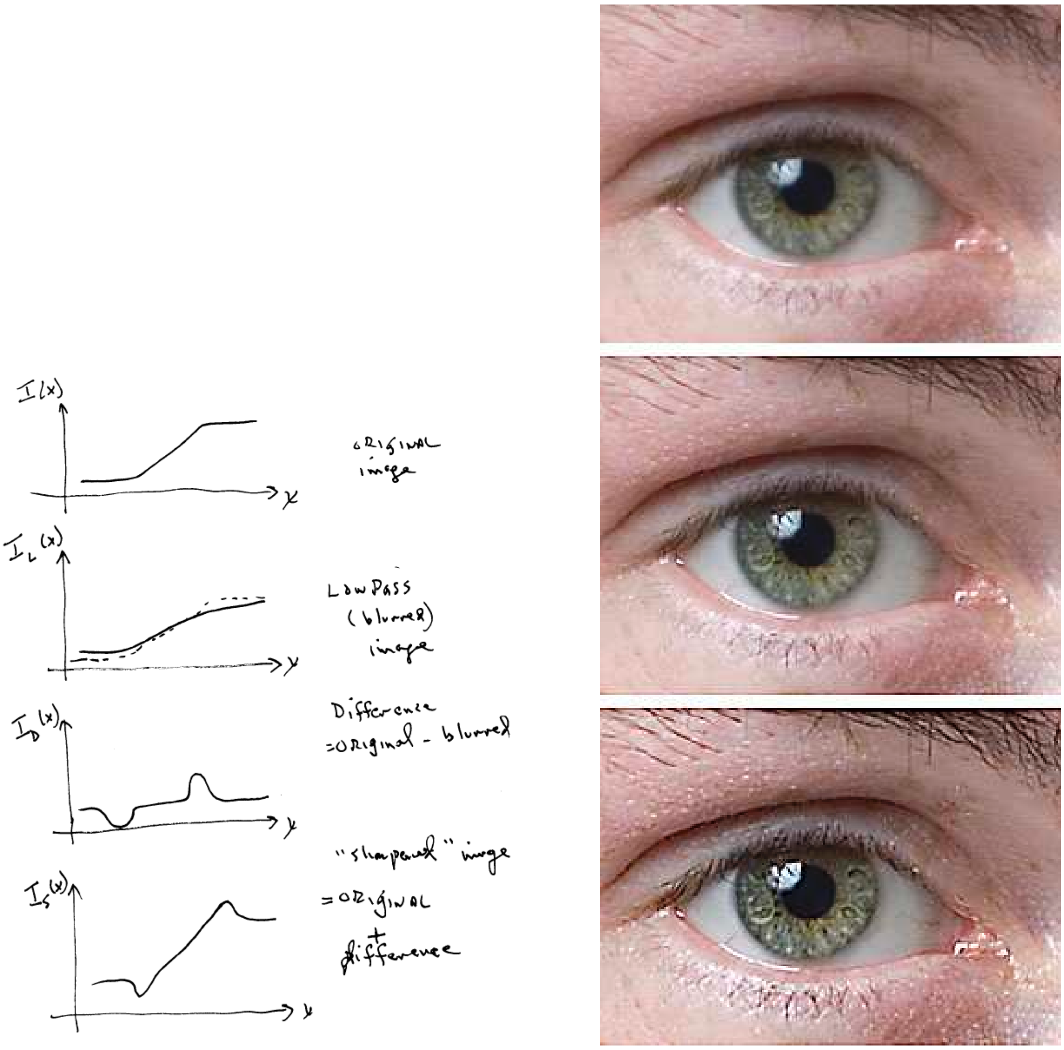


Figure 9.16: Unsharp masking. (**left**) A graph of the image intensities along a scan line in the neighborhood of a dark to bright (blurry) transition. When blurred even further by e.g. Gaussian convolution the difference from the original is clear; subracting the blurred image from the original results in an "edge image" which, when added to the original image, provides an "edge-enhanced" version. (**right**)Top - original image; middle - enhanced image by unsharp masking; bottom - overly enhanced image by adding too much of the "edge image"; notice the distortions. example from Wikipedia.

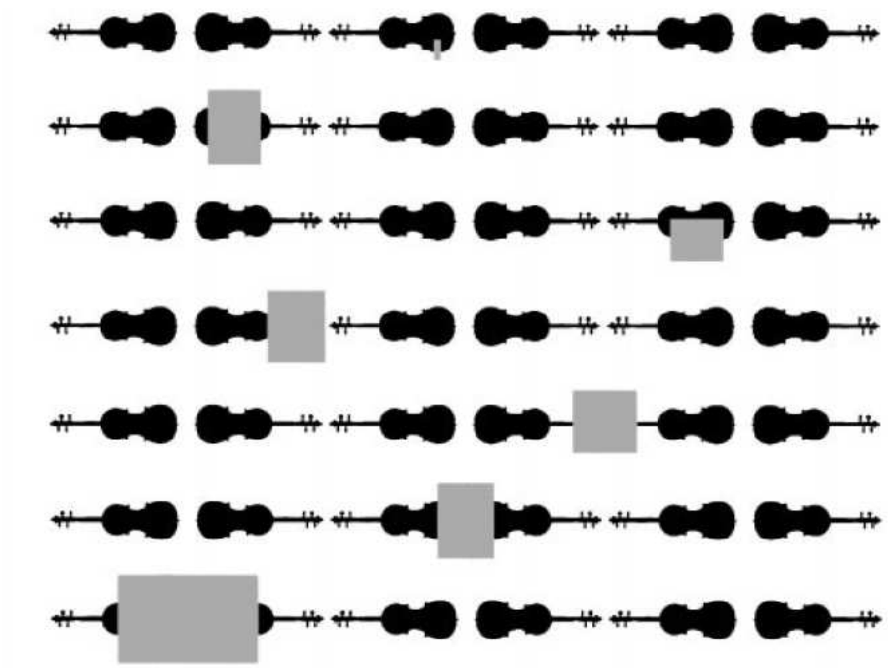


Figure 9.17: Illustration of non-linearities in visual processing. Our world contains occluders. Figure after Kanizsa.

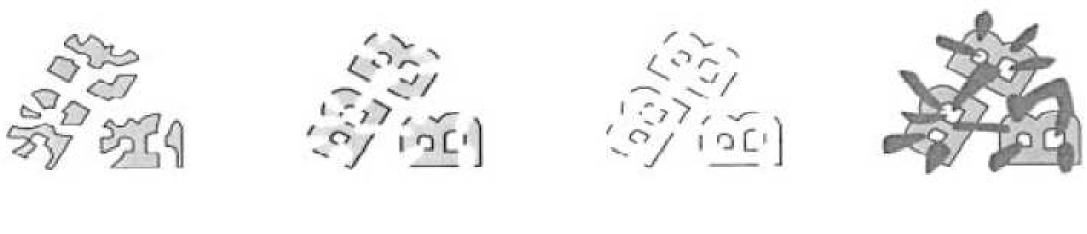


Figure 9.18: Further illustrations of non-linearities. Notice how the presence or absence of different portions of the boundary around colored regions effects the percept. Are all of these brightness effects actually in the image? After Bregman's B's.

9.10 Two summary points

Two points emerged in these last few lectures that we'll now carry forward in surprising ways. First, we note that a sinusoidal input to a linear system is an output sinusoid of the same frequency although, in general it will be changed in amplitude. Second, we started off with an infinite dimensional input and ended up with a finite dimensional representation of it that contained everything, provided the signal was limited in its frequency support.

These two ideas will be related in the next lecture.

9.11 Notes

There are many books and websites on Fourier Series and the Fourier Transform. A classic is: Bracewell, R. N. (1986), The Fourier Transform and Its Applications (revised ed.), McGraw-Hill; 1st ed. 1965, 2nd ed. 1978. You might find it fascinating to glance through **Fourier Analysis** by T. W. Korner, Cambridge U.P.

The sampling theorem illustration comes from Carleson but can be found in all books on signal processing.

For a visual perception perspective on this material, see the classic: T. Cornsweet, **Visual Perception**, Academic Press, 1970 (highly recommended, since it also has a terrific discussion of Mach bands, etc.)

or the more recent B. Wandell, **Foundations of Vision**, Sinauer, 1995.

Lincoln's face comes from Harmon, L. D. and Julesz, B. (1973). Masking in Visual Recognition: Effects of Two-Dimensional Filtered Noise. *Science*, 1973, 180:11941197

We'll be looking at much more regarding non-linearities, but for now Bregman's B's comes from: Al Bregman, **Auditory Scene Analysis**, MITPress.

The violins are after G. Kanizsa, **Organization in Vision: essays on Gestalt perception**, publisher = Praeger, city = New York, year = 1979,

(There are few more fascinating and motivating books on vision than Kanizsa! We shall touch on Gestalt phenomena later in this class, and much more in the next semester.)

Classical reference on image processing: W Pratt

Stephane Mallat on wavelets.

Deblurring examples from Hummel, Kimia, Zucker, Deblurring Gaussian blur, Computer Vision, Graphics, And Image Processing 38, 66-80 (1987).

CT example from Front-End Vision and Multi-Scale Image Analysis by B. ter haar Romeny, Springer, 2003.

Some references on the wagon wheel controversy: http://en.wikipedia.org/wiki/Wagon-wheel_effect

Some references on whether neural processing is discrete or continuous in time: VanRullen R, Koch C. Is perception discrete or continuous? Trends Cogn Sci. 2003;7:207213

Burr, 1980 D.C. Burr, Motion smear. Nature, 284 (1980), pp. 164165.

H.B. Barlow, Temporal and spatial summation in human vision at different background intensities. Journal of Physiology (London), 141 (1958), pp. 337350.

D.C. Burr, J. Ross and M.C. Morrone, Seeing objects in motion, Proceedings of the Royal Society of London B, 227 (1986), pp. 249265

Linear mechanisms can produce motion sharpening Ari K. Paakkonen , Michael J. Morgan, Vision Research 41 (2001) 27712777