7

Image Processing: Discrete Images

In the previous chapter we explored linear, shift-invariant systems in the continuous two-dimensional domain. In practice, we deal with images that are both limited in extent and sampled at discrete points. The results developed so far have to be specialized, extended, and modified to be useful in this domain. Also, a few new aspects appear that must be treated carefully.

The sampling theorem tells us under what circumstances a discrete set of samples can accurately represent a continuous image. We also learn what happens when the conditions for the application of this result are not met. This has significant implications for the design of imaging systems.

Methods requiring transformation to the frequency domain have become popular, in part because of algorithms that permit the rapid computation of the discrete Fourier transform. Care has to be taken, however, since these methods assume that the signal is periodic. We discuss how this requirement can be met and what happens when the assumption does not apply.

7.1 Finite Image Size

In practice, images are always of finite size. Consider a rectangular image of width W and height H. Then the integrals in the Fourier transform no longer need to be taken to infinity:

$$F(u,v) = \int_{-H/2}^{H/2} \int_{-W/2}^{W/2} f(x,y) e^{-i(ux+vy)}\, dx\, dy.$$

Curiously, we do not need to know $F(u,v)$ for all frequencies in order to reconstruct $f(x,y)$. Knowing that $f(x,y) = 0$ for $|x| > W/2$ and $|y| > H/2$ provides a strong constraint. Put another way, there is a lot less information in a function that is nonzero only over a finite part of the image plane than in one that is not.

To see this, consider the image plane tiled with copies of the image. That is, extend the image in a doubly periodic fashion into a function

$$\widetilde{f}(x,y) = \begin{cases} f(x,y), & \text{for } |x| \le W/2 \text{ and } |y| \le H/2; \\ f(x-kW, y-lH), & \text{for } |x| > W/2 \text{ or } |y| > H/2, \end{cases}$$

where

$$k = \left\lfloor \frac{x + W/2}{W} \right\rfloor \qquad \text{and} \qquad l = \left\lfloor \frac{y + H/2}{H} \right\rfloor.$$

Here $\lfloor x \rfloor$ is the largest integer that is not larger than x. The Fourier transform of the repeated image is

$$\begin{aligned} \widetilde{F}(u,v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \widetilde{f}(x,y) e^{-i(ux+vy)}\, dx\, dy \\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \int_{-H/2}^{H/2} \int_{-W/2}^{W/2} f(x,y)\, e^{-i(u(x-kW)+v(y-lH))}\, dx\, dy \\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} e^{iukW} e^{ivlH} F(u,v). \end{aligned}$$

It is shown in exercise 7-1, using suitable convergence factors, that

$$\sum_{k=-\infty}^{\infty} e^{ikx} = 2\pi \sum_{k=-\infty}^{\infty} \delta(x - 2\pi k).$$

Thus

$$\widetilde{F}(u,v) = 4\pi^2 \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(uW - 2\pi k)\, \delta(vH - 2\pi l)\, F(u,v)$$

$$= 4\pi^2 \frac{1}{WH} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta\left(u - \frac{2\pi}{W}k\right) \delta\left(v - \frac{2\pi}{H}l\right) F(u,v),$$

from which we see that $\widetilde{F}(u,v)$ is zero except at a discrete set of frequencies,

$$(u,v) = \left(\frac{2\pi}{W}k, \frac{2\pi}{H}l\right).$$

Thus, to find $\widetilde{f}(x,y)$ we only need to know $F(u,v)$ at these frequencies. But $f(x,y)$ can be obtained from $\widetilde{f}(x,y)$ by just "cutting out" the piece for which $|x| < W/2$ and $|y| < H/2$. So we only need to know

$$F_{kl} = F\left(\frac{2\pi}{W}k, \frac{2\pi}{H}l\right)$$

for all k and l to recover $f(x,y)$. This is a countable set of numbers.

Note that the transform of a periodic function is discrete. The inverse transform can be expressed in the form of a series, since

$$\widetilde{f}(x,y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u,v) e^{+i(ux+vy)}\, dx\, dy$$

$$= \frac{1}{WH} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta\left(u - \frac{2\pi}{W}k\right) \delta\left(v - \frac{2\pi}{H}l\right)$$

$$\times F(u,v) e^{+i(ux+vy)} dx dy$$

$$= \frac{1}{WH} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F_{kl}\, e^{2\pi i (\frac{k}{W}x + \frac{l}{H}y)}.$$

Another way to look at this is to consider $f(x,y)$ a windowed version of some $\widetilde{f}(x,y)$, where $\widetilde{f}(x,y) = f(x,y)$ within the window. That is,

$$f(x,y) = \widetilde{f}(x,y)\, w(x,y),$$

where the window function $w(x,y)$ is defined as

$$w(x,y) = \begin{cases} 1, & \text{for } |x| \le W/2 \text{ and } |y| \le H/2; \\ 0, & \text{for } |x| > W/2 \text{ or } |y| > H/2. \end{cases}$$

The transform of $f(x, y)$ is then just the convolution of the transform of $\widetilde{f}(x, y)$ with the transform of $w(x, y)$. The latter is

$$WH \, \frac{\sin(uW/2)}{uW/2} \, \frac{\sin(vH/2)}{vH/2}.$$

The Fourier transform of $f(x, y)$ is thus a highly smoothed version of the transform of $\widetilde{f}(x, y)$. We shall see later that such a filtered function can be fully specified by suitably chosen samples. The function at points other than the given sample points can easily be found by interpolation from the given samples.

7.2 Discrete Image Sampling

When the image is digitized, the brightness is known only at a discrete set of locations. We can think of the result as defined by a discrete grid of impulses,

$$f(x, y) = wh \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_{kl} \, \delta(x - kw, y - lh),$$

where w and h are the horizontal and vertical sampling intervals, respectively. The Fourier transform now becomes

$$\begin{aligned} F(u, v) &= wh \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_{kl} \, \delta(x - kw, y - lh) e^{-i(ux+vy)} \, dx \, dy \\ &= wh \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_{kl} \, e^{-i(ukw+vlh)}. \end{aligned}$$

This is a periodic function. The period in u is $2\pi/w$ and that in v is $2\pi/h$. Thus a discrete function transforms into a periodic one. This means that we can forget the part of $F(u, v)$ for $|u| > \pi/w$ and $|v| > \pi/h$. We do not need it to recover $f(x, y)$.

It is of interest to recover the inverse transform of a function that is equal to $F(u, v)$ in this region and zero outside:

$$\widetilde{F}(u, v) = \begin{cases} F(u, v), & \text{for } |u| \le \pi/w \text{ and } |v| \le \pi/h; \\ 0, & \text{for } |u| > \pi/w \text{ or } |v| > \pi/h. \end{cases}$$

The inverse transform is

$$\widetilde{f}(x, y) = \frac{1}{4\pi^2} \int_{-\pi/h}^{\pi/h} \int_{-\pi/w}^{\pi/w} F(u, v) e^{+i(ux+vy)} \, dx \, dy.$$

This function is defined for all x and y, but we are particularly interested in its values at the grid points $(x, y) = (kw, lh)$. We can write the function as

$$\begin{aligned}
\widetilde{f}(x,y) &= \frac{wh}{4\pi^2} \int_{-\pi/h}^{\pi/h} \int_{-\pi/w}^{\pi/w} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_{kl}\, e^{-i(ukw+vlh)} e^{+i(ux+vy)}\, du\, dv \\
&= \frac{wh}{4\pi^2} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_{kl} \int_{-\pi/h}^{\pi/h} \int_{-\pi/w}^{\pi/w} e^{i(u(x-kw)+v(y-lh))}\, du\, dv \\
&= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_{kl} \frac{\sin\bigl(\pi(x/w-k)\bigr)}{\pi(x/w-k)} \frac{\sin\bigl(\pi(y/h-l)\bigr)}{\pi(y/h-l)}.
\end{aligned}$$

At $(x, y) = (kw, lh)$ the above reduces to f_{kl}. Between grid points, $\widetilde{f}(x, y)$ is interpolated using a kernel that is the product of a $\sin(x)/x$ term and a $\sin(y)/y$ term.

Another way to look at this is to consider the function created by multiplying $f(x, y)$ by the sampling grid:

$$g(x,y) = wh \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(x-kw, y-lh).$$

The Fourier transform of the result is $1/(4\pi^2)$ times the convolution of the transform of $f(x, y)$ with the transform of the sampling grid. The latter is

$$4\pi^2 \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta\left(u - \frac{2\pi k}{w}, v - \frac{2\pi l}{h}\right),$$

so that the Fourier transform of $f(x, y)$ times $g(x, y)$ is a sampled version of the transform $F(u, v)$ of $f(x, y)$, namely

$$\sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F\left(\frac{2\pi}{w}k, \frac{2\pi}{h}l\right) \delta\left(u - \frac{2\pi k}{w}, v - \frac{2\pi l}{h}\right).$$

7.3 The Sampling Theorem

From the foregoing discussion we see that a function that is bandlimited is fully specified by samples on a regular grid. This result is known as the *sampling theorem.* If $F(u, v) = 0$ for $|u| > \pi/w$ or $|v| > \pi/h$, then $f(x, y)$

can be recovered from the set $f(kw, lh)$ for all integers k and l. In fact, we have an explicit interpolation formula,

$$f(x, y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_{kl} \frac{\sin\bigl(\pi(x/w - k)\bigr)}{\pi(x/w - k)} \frac{\sin\bigl(\pi(y/h - l)\bigr)}{\pi(y/h - l)},$$

that does not involve the Fourier transform.

This is an important result because it justifies sampling the image. No information is lost, provided that the function sampled is "smooth" enough, that is, provided that it is bandlimited. The required sampling interval is also specified. If only frequencies less than B occur, the sampling interval can be as large as $\delta = \pi/B$. Stated in a different way, the sampling interval should be less than $\lambda/2$ when λ is the wavelength of the highest frequency present. If δ is the sampling interval, the result can be expressed in terms of the *Nyquist frequency*, π/δ. The signal can contain frequencies only up to the Nyquist frequency if it is to be faithfully reconstructed from samples.

The sampling theorem makes clear the dangers of applying this method to functions that are not limited in this way. Information is lost, and the original function cannot be recovered. What happens specifically is that higher frequencies, when sampled, look no different than frequencies within the acceptable interval. It is as if their frequencies were "folded back" at the frequency π/B. This is also called *aliasing*, since a wave of frequency $\omega > B$ produces the same samples as one of frequency $2B - \omega$.

One of the advantages of sampling an image with a sensor that has a finite area also becomes clear now. If each sensor element has a response function $r(x, y)$, the image is effectively convolved with $r(x, y)$ before sampling. This amounts to multiplication of the transform by the Fourier transform of $r(x, y)$. This smoothing operation will have the effect of attenuating the higher frequencies.

Suppose, for example, that rectangular sensors of width w and height h are tightly packed on the rectangular grid

$$r(x, y) = \begin{cases} 1, & \text{for } |x| \le w/2 \text{ and } |y| \le h/2; \\ 0, & \text{for } |x| > w/2 \text{ or } |y| > h/2. \end{cases}$$

Then

$$R(u, v) = wh \frac{\sin(uw/2)}{uw/2} \frac{\sin(vh/2)}{vh/2}.$$

This transfer function becomes zero for $u = \pm(2\pi/w)$ and $v = \pm(2\pi/h)$. This is twice the maximum frequency allowed by the sampling theorem.

While this filter does pass some of the higher frequencies, it at least attenuates them significantly. To do better, adjacent sensing areas would have to overlap, and each sensing element would have to have sensitivity falling off toward its edge. Another way to achieve the desired effect is to have the imaging system itself introduce blurring of the right magnitude in order to lowpass filter the image sufficiently. The overall point-spread function is then equal to the convolution of the optical system point-spread function with the sensing element response function $r(x, y)$.

Ideally the sensor spacing should match the resolution of the optical elements. If they are too far apart, the conditions of the sampling theorem are violated. It is wasteful, on the other hand, to pack them too closely, since no new information is picked up by the extra sensing elements. The human visual system, for example, appears to have a reasonable match between resolution and sensor spacing, at least for intermediate opening of the pupil.

Filtering after sampling, of course, does no good. The damage has already been done. It is certainly possible to suppress higher frequencies within the acceptable band, but even higher frequencies in the original image have already been aliased down to lower frequencies in the sampled image. One cannot recover from such errors.

Bandlimited functions are "smooth" because their higher derivatives are limited in amplitude. For example, the transform of

$$f''(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

is

$$F''(u, v) = -(u^2 + v^2)F(u, v).$$

So if $F(u, v) = 0$ for $|u| > B$ or $|v| > B$, then

$$|F''(u, v)| \leq B^2 \, |F(u, v)| \, .$$

Now the power in the signals $f(x, y)$ and $f''(x, y)$ is given by

$$P = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^2(x, y) \, dx \, dy = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |F(u, v)|^2 \, du \, dv$$

and

$$P'' = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(f''(x, y)\right)^2 dx \, dy = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |F''(u, v)|^2 \, du \, dv.$$

Therefore $P'' \leq B^4 P$. This places a constraint on how rapidly $f(x, y)$ can fluctuate.

So far we have assumed that the image is sampled at points lying on a rectangular grid. It turns out that there are some advantages to other sampling schemes. We can use hexagonal picture cells lying on a triangular grid, for example. Fewer samples are required using this tessellation when the image is limited to frequencies $\rho < B$, say. Such a circularly symmetric cutoff in the frequency domain is more natural than the rectangular one, for which the rectangular tessellation is well suited. The difference between the two tessellations comes from the different regions in the transform space that can be recovered without aliasing. For a square tessellation this region is square. To fit all of the disk $\rho < R$ into the square, its side must be of length 2ρ. For the hexagonal tessellation this region is hexagonal. To fit all of the disk $\rho < R$ into the hexagon, however, the maximum cross section need only be $\sqrt{3}\rho$.

7.4 The Discrete Fourier Transform

A discrete image has a periodic transform. If we think of the image as part of a periodic infinite image, then the transform is also discrete. Thus both the image and its transform are periodic and discrete. Both are fully defined by a finite number of values. If the image is specified by the values f_{kl} of $f(x, y)$ at points (kw, lh) for $k = 0, 1, \ldots, M-1$ and $l = 0, 1, \ldots, N-1$, the transform can be written as

$$F_{mn} = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} f_{kl}\, e^{-\pi i (\frac{km}{M} + \frac{ln}{N})},$$

and the inverse transform as

$$f_{kl} = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} F_{mn}\, e^{+\pi i (\frac{km}{M} + \frac{ln}{N})},$$

where again we have near symmetry in the definitions of the forward and inverse transforms. There is never any question about the existence of the transforms, since they consist of finite sums of finite values.

Note that this transform is based on an image that is doubly periodic. Unless the brightness of the left edge of the image happens to match that of the right edge, there will be a step discontinuity in brightness where adjacent copies of the image touch. Even if the image itself is very smooth, this discontinuity will introduce some high-frequency components into the transform. There are a number of ways to suppress this undesirable effect. One is to flip copies of the image over sideways before attaching them on the left and on the right. Similarly, copies of the image can be flipped over top to bottom before being attached at the top and the bottom. The brightness of the resulting doubly periodic function is continuous across the image border, but the odd derivatives of brightness are still discontinuous. Spurious spectral components will result, but at least these are typically smaller than those due to discontinuities in brightness itself.

The resulting function is periodic, though the period is twice as long as it was earlier when we simply replicated the image unchanged. On the other hand, the function is even in both x and y, so that only cosinusoidal components can occur. The discrete transform obtained, called the *cosine transform*, is explored further in exercise 7-6.

Another way to ameliorate effects of potential discontinuities at the image borders is to modulate the image by multiplying it by a function that drops to zero on the border. Adjacent copies will then automatically match. The modulation function itself should vary smoothly, so that it will not introduce spurious effects. Such a function is often called a *window function* since it provides us with a look through a weighted window into the potentially infinite image.

An example of a simple window function is

$$\frac{1}{2}\left(1+\cos\left(2\pi\frac{x}{W}\right)\right)\,\frac{1}{2}\left(1+\cos\left(2\pi\frac{y}{H}\right)\right)=\cos^2\left(\pi\frac{x}{W}\right)\,\cos^2\left(\pi\frac{y}{H}\right).$$

We saw earlier that, just as convolution in the image domain is equivalent to multiplication in the frequency domain, so multiplication in the image domain is equivalent to convolution in the frequency domain. Thus the transform of the windowed image is $4\pi^2$ times the transform of the original image convolved with the transform of the window function. Thus the transform is smeared out somewhat by windowing. The transform of the above window function is

$$\frac{1}{2}\left(\frac{1}{2}\delta\Big(\frac{2\pi}{W}-u\Big)+\delta(u)+\frac{1}{2}\delta\Big(\frac{2\pi}{W}+u\Big)\right)$$
$$\times\,\frac{1}{2}\left(\frac{1}{2}\delta\Big(\frac{2\pi}{H}-v\Big)+\delta(v)+\frac{1}{2}\delta\Big(\frac{2\pi}{H}+v\Big)\right);$$

each value in the transform is a local weighted average over a 3×3 neighborhood. The center is multiplied by the weight $1/4$, the four edge-adjacent neighbors by $1/8$, and the four neighbors on the corners by $1/16$. This convolutional weighting scheme can be represented by the following stencil:

$$\frac{1}{16}\begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

A *stencil* is a pattern of weights used in computing a convolution, arranged in such a way as to suggest the spatial relationships between the places where the weights are applied.

One of the reasons for the attention given the discrete Fourier transform is that an algorithm has been discovered for computing it efficiently. The obvious implementation, in which each term f_{kl} is computed separately, requires MN multiplications for each of MN results, that is, M^2N^2 multiplications overall. The Fast Fourier Transform (FFT) takes only $4MN \log_2 MN$ multiplications to compute all results by clever sharing of intermediate terms. This makes it reasonable to compute convolutions by Fourier transformation, multiplication, and inverse transformation. Modern developments in parallel hardware, however, often favor direct convolution methods.

We now turn to a consideration of noise in images. The discrete Fourier transform of an image that is just noise will of course depend on the particular values at each picture cell. Can we say something in general about the transform? We are interested in the expected values of each of the transformed numbers. We have

$$F_{mn} = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} f_{kl}\, e^{-\pi i\left(\frac{km}{M} + \frac{ln}{N}\right)}.$$

Here, each f_{kl} is a random value independent of the others. For the moment, allow f_{kl} to take on complex values. Note that

$$f_{kl}\, e^{-\pi i\left(\frac{km}{M} + \frac{ln}{N}\right)}$$

is also a random value, with random phase, provided that f_{kl} has random phase. Thus each value F_{mn} is obtained by adding MN independent random complex numbers.

Now, if we further assume that f_{kl} has a normal distribution with zero mean and standard deviation σ, we can conclude that F_{mn} also has zero mean. The standard deviation will be $\sqrt{MN}\,\sigma$ since the standard deviation of the sum of the MN values is $\sqrt{MN}$ times their individual standard deviation. It can further be shown that the resulting values are independent. Thus the transform has properties identical to those of the original image (except for a scale factor); that is, the transform is a set of independent random numbers of mean zero and standard deviation $\sqrt{MN}\,\sigma$.

7.5 Circular Convolution

The interest in the discrete Fourier transform stems in large part from the fact that convolutions can be computed by multiplying Fourier transforms. Note, however, that there are problems at the edges of the window and the filter function. It is assumed that both are periodic. Thus near the left edge of the image the output will be affected somewhat by what appears near the right edge of the image.

This may be undesirable. An alternative is to extend the image with a border of zeros. The border should be as wide as the *support* of the filter, that is, the region over which the point-spread function of the filter is nonzero. Many filters do not have finite support and have to be artificially truncated. This includes the Gaussian blob, one of our favorites.

Adding a border of zeros has its own drawbacks since the output near the border will be affected by the image extension. This is not surprising if we consider the available image to be part of an unknown larger image. We have to guess at how the image can be extended. The part of the image affected this way is a border of width equal to the support of the filter. Only the part inside this border is trustworthy. If this is all that will be used, there is no point, of course, in adding a border of zeros outside the original image.

In summary, we have seen how imaging systems can be characterized by their point spread function or their modulation-transfer function. Shortcomings in such systems can be analyzed by studying the resulting point-spread function. The sampling theorem allows us to rationally match image sensors to image-forming systems. The optimal filtering methods allow us to design systems for extracting signals of interest in the presence of noise. Convolutional methods will also be useful in edge detection.

7.6 Some Useful Rules

Some of the results we have developed can be summarized in a short table:

Spatial domain	Frequency domain
Periodic	Discrete
Symmetric	Real
Sum of two functions	Sum of two transforms
Convolution of two functions	Product of two transforms
Periodic sampling	Periodic copies

The implications go both ways in this table. Thus, for example, the transform of a periodic function is discrete, and a function with a discrete transform is periodic. Furthermore, the columns can be relabeled in inverted order. That is, the transform of a discrete function is periodic, and a function with a periodic transform is discrete. There are many more such helpful relationships.

When the scale is compressed in one domain, it is expanded proportionally in the other, so that the product of corresponding measures of size is constant. For example, the product of the width of a Gaussian in the spatial domain and the width of its transform in the frequency domain is a constant.

A function has *finite support* when it is nonzero only over a bounded region. It can be shown that the transform of a function with finite support cannot also have finite support.

7.7 References

Many of the basic references on image processing were given at the end of the previous chapter. Hamming first presented fast methods for computing the discrete Fourier transform in *Numerical Methods for Scientists and Engineers* [1962]. Somehow his method was overlooked, to be rediscovered much later, as he observes dryly in another excellent book, *Digital Filters* [1977, 1983]. He treated relatively short vectors since he was interested in calculations that could be done by hand. It may not have been obvious to others how his technique generalized to input vectors of arbitrary lengths. *Fast Algorithms for Digital Signal Processing* by Blahut [1985] summarizes what is known about fast algorithms for the discrete Fourier transform and related problems.

The sampling theorem gives the conditions under which discrete samples contain enough information to fully describe a continuous image. Mersereau [1979] shows the advantage of hexagonal image tessellations for representing two-dimensional signals. Recovering the continuous image requires interpolation and convolution methods. A discrete image can be enlarged or reduced by sampling an interpolated version of the original. Hou & Andrews [1978] discuss one approach to this. Ahmed, Natarjan, & Rao [1974] discuss the discrete cosine transform, while Chen, Smith, & Fralick [1977] describe fast ways of computing it.

Image restoration has been one of the main applications of the results of the field of image processing. This is still an active area, as shown by the recent paper by Ramakrishna, Mullick, & Rathore [1985].

7.8 Exercises

7-1 Show that

$$\sum_{k=-\infty}^{\infty} e^{ikx} = 2\pi \sum_{k=-\infty}^{\infty} \delta(x - 2\pi k).$$

Hint: A Gaussian makes a suitable convergence factor for use in evaluating this sum.

7-2 A discrete two-dimensional system is characterized by the point-spread function $\{h_{ij}\}$. The output $\{g_{ij}\}$ is computed from the input $\{f_{ij}\}$ according to the rule

$$f_{i,j} = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_{i-k,j-l}\, h_{k,l}.$$

Compute the modulation-transfer function. Hint: Is the modulation-transfer function discrete? Is it periodic?

7-3 A discrete two-dimensional system performs a convolution, as in the previous exercise. Now suppose that the input $\{f_{ij}\}$ is just noise. That is, each of the f_{ij} is an independent random variable with zero mean and with variance σ^2.

(a) What are the mean and variance of the output g_{ij}?

(b) What point-spread function satisfying the constraint

$$\sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h_{kl} = 1$$

minimizes the noise in the output? Hint: You may want to use a Lagrange multiplier to enforce the constraint in the minimization (see the appendix).

(c) What point-spread function satisfying the constraint

$$\sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} (k^2 + l^2) h_{kl} = 1$$

minimizes the noise in the output? Warning: This part is harder than the rest of the problem.

7-4 Consider the discrete approximation of the Laplacian given by convolution with the following pattern of weights:

$$\frac{1}{6\epsilon^2}\begin{array}{|c|c|c|}\hline 1 & 4 & 1 \\ \hline 4 & -20 & 4 \\ \hline 1 & 4 & 1 \\ \hline \end{array}$$

where ϵ is the spacing between picture cells. Write this weighting scheme as the sum of nine impulse functions. Find the Fourier transform and show that near the origin it tends to $-(u^2 + v^2)$ as $\epsilon \to 0$.

7-5 Consider the discrete approximation of the Laplacian used in the previous exercise. Apply this convolutional weighting scheme to a Taylor series expansion of the image brightness about the central point. Show that the result is independent of the constant and the linear terms. Show that it computes the expected combination of the second-order terms. Also work out the lowest-order error terms.

7-6 Suppose that an image $f(x, y)$ lies in the region $0 \le x \le W$, $0 \le y \le H$. To avoid discontinuities in brightness at the borders, extend this by mirror imaging. For example, let $\widetilde{f}(x, y) = f(2W - x, y)$ for $W \le x \le 2W$, while $\widetilde{f}(x, y) = f(x, 2H - y)$ for $H \le y \le 2H$.

(a) Show that the extended image has period $2W$ in the x-direction and period $2H$ in the y-direction. Also show that the extended image is an even function of both x and y.

(b) Show that the extended image can be represented by the cosine series

$$\widetilde{f}(x, y) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} C_{kl} \cos\left(\frac{x}{W}\pi k\right) \cos\left(\frac{y}{H}\pi l\right).$$

(c) Show that the coefficients of the cosine series can be found using

$$C_{ij} = \frac{4}{WH} \int_0^{+H} \int_0^{+W} f(x, y) \cos\left(\frac{x}{W}\pi i\right) \cos\left(\frac{y}{H}\pi j\right)\, dx\, dy,$$

for $i \ne 0$ and $j \ne 0$. Find the corresponding result for the cases where $i = 0$ or $j = 0$.

(d) How can the above method be modified to deal with an image that lies in the region $-W/2 \le x \le W/2$, $-H/2 \le y \le H/2$?

(e) Is it possible to replicate a hexagonal image region in a way that will ensure continuity in brightness across the image borders?

7-7 Show that, at least formally,

$$1 + \left(\frac{\sigma^2}{2}\nabla^2\right) + \frac{1}{2!}\left(\frac{\sigma^2}{2}\nabla^2\right)^2 + \frac{1}{3!}\left(\frac{\sigma^2}{2}\nabla^2\right)^3 + \cdots$$

is the inverse of convolution with the Gaussian

$$\frac{1}{2\pi\sigma^2}e^{-\frac{1}{2}\frac{x^2+y^2}{\sigma^2}}.$$

Hint: Expand $e^{\frac{1}{2}\sigma^2\omega^2}$ in a Taylor series.

Suppose that the first n terms of the series above are to be used to partially undo the blurring introduced by convolution with a Gaussian. Let the blurred image be $f(x, y)$, while the partially deblurred version is $g(x, y)$. Show that $g(x, y)$ can be found using the iterative scheme

$$g^0 = f,$$

$$g^{k+1} = f + \frac{1}{n-k}\left(\frac{\sigma^2}{2}\nabla^2\right) g^k \qquad \text{for } k = 0, 1, \ldots, n-1.$$

7-8 An image can be smoothed by means of a simple 2×2 averaging filter. The point-spread function of this filter can be written

$$h_{ij} = \begin{cases} 1/4, & \text{for } i = 0,\ 1,\ j = 0, 1; \\ 0, & \text{otherwise.} \end{cases}$$

Consider the function

$$h'_{ij} = \begin{cases} 1, & \text{for } i = -1, 0,\ j = -1, 0; \\ \text{even}\big(|x + 1/2| - 1/2\big)\,\text{even}\big(|y + 1/2| - 1/2\big), & \text{otherwise,} \end{cases}$$

where

$$\text{even}(z) = \begin{cases} +1, & \text{when } z \text{ is even;} \\ -1, & \text{when } z \text{ is odd.} \end{cases}$$

Show that h'_{ij} is a convolutional inverse to h_{ij}. (Note that, except near the axes, h'_{ij} is just a checkerboard of $+1$s and -1s.)

Discuss the difficulties we would encounter if we attempted to use this result in practice. Suggest a suitable window function that can be employed to attenuate the convolutional inverse for large arguments i and j. How would convolution of a filtered image with the product of this window function and h'_{ij} differ from convolution with the "exact" inverse?

7-9 Here we establish a useful correspondence between power series and discrete point-spread functions. We also show a way to approximate continuous Gaussian filters using a discrete filter. Consider two polynomials $f(x)$ and $g(x)$.

(a) Show how the set coefficients of the product polynomial $f(x)g(x)$ can be obtained by discrete convolution of the sets of coefficients of the individual polynomials $f(x)$ and $g(x)$. Extend this to power series that include negative powers of x.

This result can be extended to two dimensions. That is, the coefficients of the power series $f(x,y)g(x,y)$ can be obtained by two-dimensional discrete convolution of the coefficients of the power series $f(x,y)$ and $g(x,y)$. Note that associativity and commutativity of multiplication of power series follows from the corresponding properties of discrete convolution. We now consider a particular example.

(b) Show how convolution with a filter whose weights are given by the pattern

$$\frac{1}{16}\begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

can be accomplished by repeated convolution with a filter that has the pattern

$$\frac{1}{4}\begin{array}{|c|c|}\hline 1 & 1 \\ \hline 1 & 1 \\ \hline \end{array}$$

Show the weighting scheme that results when we convolve one more time with the 2×2 pattern shown.

(c) Show that the modulation-transfer function of a discrete convolutional filter can be obtained by substituting e^{-iuw} for x and e^{-ivh} for y in the corresponding power series, where w and h are the horizontal and vertical sampling intervals, respectively.

(d) What is the Fourier transform of the convolutional filter corresponding to the polynomial $1 + x + y + xy$?

A polynomial of particular interest is obtained by expanding $(1+x)^n$:

$$(1+x)^n = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n-1}x^{n-1} + \binom{n}{n}x^n,$$

where the *binomial coefficients* are

$$\binom{n}{r} = \frac{n!}{(n-r)!\,r!}.$$

(e) Suppose that we convolve the filter in part (d) with itself n times to obtain the two-dimensional binomial distribution whose general term is

$$\binom{n}{k}\binom{n}{l}x^k y^l \qquad \text{for } 0 \le x \le n,\ 0 \le y \le n.$$

Show that the magnitude of the modulation-transfer function is

$$4^n \cos^n(\tfrac{1}{2}uw)\cos^n(\tfrac{1}{2}vh).$$

How can you remove the scale factor 4^n? How can you arrange for the modulation-transfer function to be real, at least when n is odd?

(f) The binomial distribution of order n approximates a Gaussian with mean $(n-1)/2$ and variance 2^{n-1}. The amplitude of this Gaussian is about 2^{n-1}. The modulation-transfer function also approximates a Gaussian. Find the amplitude, mean, and variance of the Gaussian approximated by

$$4^n \cos^n(\tfrac{1}{2}uw)\cos^n(\tfrac{1}{2}vh).$$

Hint: Find the first few terms in the Taylor series expansion of the two functions being compared and match coefficients.

The above analysis is useful when we wish to approximate a Gaussian with a finite, discrete filter. This is a better approach than simply truncating the continuous Gaussian after it becomes "small" enough. The latter approach introduces spurious high-frequency components due to the arbitrary cutoff.

7-10 Suppose that samples are taken not only of image brightness but also of the first partial derivatives of brightness. Show that fewer samples are required to capture all the information about a bandlimited image than are necessary when only brightness is sampled. Show how to recover the bandlimited image from its samples. Warning: This is not an easy problem.

7-11 How would you interpolate a bandlimited function from its samples on a hexagonal grid? Hint: What is the inverse Fourier transform of a function that is one inside a hexagonal region in the frequency domain and zero elsewhere?