Optical and Digital Image Processing

Fundamentals and Applications
18
Linking Analog and Digital Image Processing
Leonid P. Yaroslavsky

18.1
Introduction

Informational optics is the branch of optics that deals with optical information processing, imaging, and optical measurements. The history of informational optics from ancient magnifying glasses through Galileo's telescope and van Leeuwenhoek's microscope to the huge diversity of modern optical imaging and measuring devices is a history of creating and perfecting optical imaging and measuring devices and mastering new bands of waves and new kinds of radiation. A revolutionary stage in the evolution of informational optics was the marriage of optics and electronics in the second half of the twentieth century. Nowadays, informational optics is reaching its maturity. It has gone digital. Digital computers and signal processors are becoming an inherent integral part of optical imaging and measuring devices. Here are only a few examples. Digital cameras are replacing analog photographic cameras and completely eliminating photographic alchemy; a new generation of computerized synthetic aperture telescopes and optical interferometers (see Chapter 14) is helping astronomers study stars in hundred times finer detail than is possible with classical analog optics; computer-controlled adaptive optics and digital video processing allow perfect sharp imaging through turbulent media with super-resolution exceeding the resolving power of the sensors (see Chapter 27); nowadays, optical and holographic interferometers widely make use of digital processors to achieve higher accuracy, versatility, and informational throughput; computerized tomographic synthesis and digital image processing methods have revolutionized medical imaging and nondestructive testing industry.

There are three new major qualities that have been brought into optical systems by digital computers and processors. These are as follows:

- The most substantial advantage of digital computers as compared to analog electronic and optical information processing devices is their flexibility and adaptability: no hardware modifications are necessary to reprogram digital computers to solving different tasks. This feature also makes digital computers an
ideal vehicle for processing optical signals since they can adapt rapidly and easily to varying signals, tasks, and end-user requirements.

- Digital computers integrated into optical information processing systems enable them to perform any operations needed in addition to elementwise and integral signal transformations, such as spatial and temporal Fourier analysis, signal convolution, and correlation that are characteristic for analog optics. This removes the major limitation of analog optical information processing.

- Acquiring and processing quantitative data contained in optical signals and connecting optical systems to other informational systems and networks is most natural when data are handled in a digital form. In the same way as money is the general equivalent in economics, digital signals are the general equivalent in information handling. Digital signals that represent optical ones are, so to say, purified information carried by optical signals deprived of their physical carrier. Digital signals, thanks to their universal nature, are also the ideal means for integrating different information systems.

As always, there is a trade-off between good and bad features. The fundamental limitation of signal processing in computers is the speed of computations. What optics does in parallel and with the speed of light, computers perform as a sequence of very simple logical operations with binary digits, which is fundamentally slower whatever the speed of these operations is. Obviously, the optimal design of image processing systems requires appropriate combination of analog and digital processing using the advantages and taking into consideration the limitations of both.

The marriage of analog electro-optical and digital processing requires appropriate linking analog and digital signals and transforms. In this chapter, we address several aspects of this fundamental issue specifically for digital imaging. In Section 18.2, we outline basic principles and review methods of digital representation of images and imaging transforms, such as convolution and Fourier integral transforms. In Section 18.3, we discuss applying these principles to the optimization of methods for building, in computers, continuous image models for image resampling, image recovery and super-resolution from sparse data, differentiating, and integrating. In Section 18.4, digital-to-analog conversion in digital holography is illustrated by the results of analysis of its influence on image reconstruction of physical parameters of devices used for recording computer-generated hologram.

18.2
How Should One Build Discrete Representation of Images and Transforms?

In informational optics, physical reality can be considered a continuum, whereas computers have only a finite number of states. How can one represent physical reality of optical signals and transforms in computers? The answer to this question is discretization.
18.2 How Should One Build Discrete Representation of Images and Transforms?

Signal Discretization

Signal discretization is converting continuous signals into discrete ones represented by a finite set of numbers. In principle, there might be many different ways of signal discretization. However, our technological tradition and all technical devices that are used at present for such a conversion implement a method that can mathematically be modeled as signal approximation by its expansion over a finite set of basis functions:

\[ a(x) \approx \sum_{k=0}^{N-1} \alpha_k \psi_k(x); \quad \alpha_k = \int x a(x) \psi_k^d(x) \, dx \]  

where \( a(x) \) is the image signal as a function of spatial coordinate \( x \), \( \{ \psi_k^d(x) \} \), and \( \{ \psi_k(x) \} \) are sets of basis functions used for signal discretization and their reconstruction from the discrete representation, correspondingly, \( k \) is the basis function index, and \( N \) is the number of discrete representation coefficients \( \{\alpha_k\} \) of the signal.

For different bases, signal approximation accuracy for a given \( N \) might be different. Obviously, the discretization bases that provide better approximation accuracy for a given \( N \) are preferable. However, the accuracy of signal recovery from its discrete representation is not the only issue in selecting discretization and reconstruction basis functions. Another issue is that of complexity of generating and implementing discretization and reconstruction basis functions and computing the signal representation coefficients.

In principle, discretization and reconstruction basis functions can be implemented in a form of prefabricated templates in the image discretization and display devices (or, in digital processing, in computer memory). However, in practice, the number \( N \) of the required template functions is most frequently very large (for instance, on the order of \( 10^6 \) to \( 10^7 \) for modern digital cameras). This is why signal discretization and reconstruction basis functions are usually generated (in hardware or software) by means of certain modifications of a unique “mother” function. Three families of such basis functions are best known: shift (convolutional) basis functions, scale (multiplicative) basis functions, and wavelets.

Shift (convolutional) basis functions, usually called sampling functions, implement the simplest method for generating basis functions from one “mother” function, that of spatial translation (coordinate shift) of a “mother” function \( \psi_0(x) \), such that

\[ \psi_k(x) = \psi_0[x - (k + u)\Delta x]; \quad \psi_k^r(x) = \psi_0^r[x - (k + u)\Delta x] \]  

and signal representation coefficients are samples of signal convolution with sampling basis functions:

\[ \alpha_k = \int_{-\infty}^{\infty} a(\xi) \psi_k^{(i)}(x - \xi) \, d\xi \bigg|_{x = (k + u)\Delta x} \]
where \( \Delta x \) is an elementary shift interval called the sampling interval and \((u^0, u^1)\) are analogous (not necessarily integer) shift parameters, which, together with the sample index \( k \), determine positions of sampling points in the signal coordinate system.

If the number of signal samples is unlimited, the best signal approximation, in terms of the minimum of mean square approximation error, is achieved when sinc-functions

\[
sinc \left( \frac{\pi [x - (k + u)\Delta x]}{\Delta x} \right) = \frac{\sin \left( \frac{\pi [x - (k + u)\Delta x]}{\Delta x} \right)}{\pi [x - (k + u)\Delta x]/\Delta x}
\]

are used as sampling and reconstruction basis functions. The signal mean square reconstruction error is, in this case, minimal possible for shift basis functions and it is equal to the energy of signal Fourier spectrum components outside the frequency interval \([-1/2\Delta x, 1/2\Delta x]\) called the sampling baseband. In the idealistic case of bandlimited functions, whose spectrum vanishes outside this interval, the approximation error is zero. These statements constitute the meaning of the classical sampling theorem.

For real-life signals that are not bandlimited and are sampled with a finite number of samples using sampling bases other than physically not realizable sinc-functions, the approximation error is never zero. It is caused by distortions of signal spectra within the baseband \([-1/2\Delta x, 1/2\Delta x]\) and by aliasing signal spectral components outside the baseband [1].

In 2D and higher dimensions, the signal translation parameter is a vector rather than a scalar. Therefore, in higher dimensions, in addition to the amount of translation, there is an additional degree of freedom for generating basis functions – the selection of directions, in which translations are implemented. The most simple and widespread is the rectangular sampling grid, which corresponds to rectangular signal basebands, although hexagonal sampling grids or sampling grids in skewed coordinates may be, in some cases, more appropriate in terms of minimization of the signal approximation error.

Scale (multiplicative) basis functions are built by means of scaling argument of a “mother” function proportionally to the function index \( k \):

\[
\varphi_k(x) = \varphi_0(kx)
\]

Such scaling basis functions are also called Fourier kernels. Typical examples are sinusoidal functions \( \{\cos(2\pi kx/X)\} \) and \( \{\sin(2\pi kx/X)\} \), where \( X \) is a finite interval, within which signals are approximated. Signal expansions over sets of sinusoidal basis functions are versions of the Fourier series expansion on finite intervals. Mathematical treatment of the Fourier series expansion is much simplified if pairs of cosine and sine functions are replaced by complex exponential functions \( \{\exp(i2\pi kx/X) = \cos(i2\pi kx/X) + i\sin(i2\pi kx/X)\} \), \( i = \sqrt{-1} \). With them, the Fourier series expansion takes the form

\[
a(x) = \sum_k \alpha_k \exp(i2\pi kx/X), \text{ with } \alpha_k = \frac{1}{X} \int_{-X/2}^{X/2} a(x) \exp(-i2\pi kx/X) dx
\]
18.2 How Should One Build Discrete Representation of Images and Transforms?

The basis of exponential functions \( \{ \exp(i2\pi k x/X) \} \) can be treated in another way as well – as having been generated by means of multiplying the “mother” function:

\[
\exp(i2\pi k x/X) = \prod_{i=1}^{k} \exp(i2\pi x/X) = \exp(i\pi \sum_{i=1}^{k} 2x/X) \quad (18.7)
\]

This is why such bases are also called multiplicative bases. Another example of family of basis functions built on this principle is the family of Walsh functions [1].

A distinctive feature of shift (convolution) bases is that, for them, signal representation coefficients depend on signal values in the vicinity of the corresponding sampling point, and therefore they carry local information about signals. In contrast to them, signal discrete representation for scale (multiplicative) bases is “global” because signal representation coefficients depend on the entire signal they represent. It is frequently useful to have a combination of these two features in the signal discrete representation. This is achieved with wavelet basis functions built using both shifting and scaling of “mother” functions. At present, numerous versions of wavelet bases are known (see Chapter 7 for a more detailed description of wavelets).

18.2.2 Imaging Transforms in the Mirror of Digital Computers

The most frequently used mathematical models of image formation in optical and holographic systems are convolutional and Fourier integral transforms. For digital processing, one has to find their appropriate representation in computers. At first glance, obtaining discrete representation of analog transforms is a trivial task: one has only to replace integrals by integral sums of integrand values in sampling points. However, this solution that originates from traditions of classical numerical mathematics, based on signal Taylor expansion approximation, disregards physical characteristics of signal sampling and reconstruction devices and, owing to this, cannot properly treat the problem of accuracy of the discrete representation. The rigorous approach to discrete representation of analog signal transforms is based on the following consistency and mutual correspondence principles [1]:

- Discrete representation of signal transformations should parallel that of signals.
- Discrete transformation corresponds to its analog prototype if both act to transform identical input signals into identical output signals.
- Digital processors incorporated into optical information systems should be regarded and treated together with signal discretization and reconstruction devices as integrated analog units and should be specified and characterized in terms of equivalent analog transformations.

Tables 18.1–18.3 summarize discrete representations of such most widely used imaging transforms as the convolution and Fourier transforms. These representations are built for different sampling conditions on the assumption that analog signals are discretized using sampling basis functions. Detailed derivations can be found in Refs [1–3].
Table 18.1 Convolution integral and digital filters (in 1D denotations).

The convolution integral of a signal $a(x)$ with shift invariant kernel $h(x)$

$$b(x) = \int_{-\infty}^{\infty} a(\xi) h(x - \xi)d\xi$$

Digital filter for samples $\{a_k\}$ and $\{b_k\}$ of input and output signals

$$b_k = \sum_{n=0}^{N_h-1} h_n a_{k-n}$$

$N_h$ is the number of nonzero samples of $h_n$

Discrete impulse response $\{h_n\}$ of the digital filter for input and output signal sampling bases $\phi_0^{(0)}(x)$ and $\phi_0^{(r)}(x)$ and sampling interval $\Delta x$

$$h_n = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h[x - \xi - n\Delta x] \psi^{(r)}(\xi) \phi^{(0)}(x) dx d\xi$$

As one can see in Tables 18.2 and 18.3, taking into consideration different possible sampling conditions leads to extension of the canonical discrete Fourier transform (DFT) to shifted DFT, discrete cosine transform (DCT), scaled DFT, and rotated DFT. Shifted, scaled, and rotated DFTs contain analog shift, scale, and rotation angle parameters that determine the geometrical position of the image, its spectrum sampling grids in analog image and its spectrum coordinate systems. Their presence enables efficient computational means for image perfect and arbitrary translation, rescaling with arbitrary scale factor and rotation, which is particularly useful in template matching with subpixel accuracy, in numerical reconstruction of color holograms, and in many other applications [3, 4].

DCT is a special case of shifted DFT with shift parameters $(1/2, 0)$ of virtual signals, obtained from original ones by means of padding them with their mirror reflection from their borders. Such a symmetrization eliminates signal discontinuities at its borders, which otherwise appear due to periodical replication of the signal that corresponds to working in the domain of DFTs. This radically improves the energy compaction property of the transform. DCT is well known for its application in image compression. Not less important is its application for boundary effect free fast digital convolution [1, 3].

Similar representations that maintain correspondence with analog transforms can be derived for the integral Fresnel transform [3] used for the numerical reconstruction of holograms and synthesis of computer-generated holograms (CGHs).

18.2.3 Characterization of Discrete Transforms in Terms of Equivalent Analog Transforms

Characterization of discrete transforms in terms of equivalent continuous transforms is necessary for proper design and performance analysis of digital image processing systems. In this section, we address this issue for digital filters and DFTs.
18.2 How Should One Build Discrete Representation of Images and Transforms?  

### Table 18.2 Discrete representations of 1D integral Fourier transform.

<table>
<thead>
<tr>
<th>Description</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Direct and inverse integral Fourier transforms of a signal</strong> ( a(x) )</td>
<td>[ a(f) = \int_{-\infty}^{\infty} a(x) \exp(i2\pi fx) , dx ] [ a(x) = \int_{-\infty}^{\infty} a(f) \exp(-i2\pi fx) , df ]</td>
</tr>
<tr>
<td><strong>Direct and inverse canonical discrete Fourier transforms (DFTs)</strong></td>
<td>[ a_r = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a_k \exp \left( \frac{i2\pi kr}{N} \right) ] [ a_k = \frac{1}{\sqrt{N}} \sum_{r=0}^{N-1} a_r \exp \left( -\frac{i2\pi kr}{N} \right) ]</td>
</tr>
<tr>
<td><strong>Sampling conditions:</strong></td>
<td></td>
</tr>
<tr>
<td>( \bullet ) Signal and signal sampling device coordinate systems are identical as also those of the signal spectrum and the assumed signal spectrum sampling device.</td>
<td></td>
</tr>
<tr>
<td>( \bullet ) Signal samples ( {a_k} ) as well as samples ( {k_r} ) of its Fourier spectrum are positioned in such a way that samples with indices ( k = 0 ) and ( r = 0 ) are taken in signal and spectrum coordinates at points ( x = 0 ) and ( f = 0 ), respectively.</td>
<td></td>
</tr>
<tr>
<td>( \bullet ) Signal and its Fourier spectrum sampling intervals ( \Delta x ) and ( \Delta f ) satisfy the “cardinal” sampling relationship ( \Delta x = 1/N\Delta f ) associated with the signal baseband ( N\Delta f ).</td>
<td></td>
</tr>
<tr>
<td><strong>Direct and inverse shifted DFTs (SDFT((u, v))):</strong></td>
<td></td>
</tr>
<tr>
<td>( a_r^u = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a_k \exp \left[ i2\pi \frac{(k+u)(r+v)}{N} \right] ] [ a_k = \frac{1}{\sqrt{N}} \sum_{r=0}^{N-1} a_r^u \exp \left[ -i2\pi \frac{(k+u)(r+v)}{N} \right] ]</td>
<td></td>
</tr>
<tr>
<td><strong>Sampling conditions:</strong></td>
<td></td>
</tr>
<tr>
<td>( \bullet ) Signal and signal sampling device coordinate systems as well as, correspondingly, those of signal spectrum and of the assumed signal spectrum discretization device, are shifted with respect to each other in such a way that signal sample ( {a_0} ) and, correspondingly, sample ( {k_0} ) of its Fourier spectrum are taken in signal and spectrum coordinates at points ( x = -u\Delta x ) and ( f = -v\Delta f ).</td>
<td></td>
</tr>
<tr>
<td>( \bullet ) Signal “cardinal” sampling: ( \Delta x = 1/N\Delta f ).</td>
<td></td>
</tr>
<tr>
<td><strong>Direct and inverse discrete cosine transform (DCT):</strong></td>
<td></td>
</tr>
<tr>
<td>( a_r = \frac{2}{\sqrt{N}} \sum_{k=0}^{N-1} a_k \cos \left( \frac{\pi}{N} \frac{k+1/2}{r} \right) ] [ a_k = \frac{1}{\sqrt{N}} \left[ a_0 + \sum_{r=1}^{N-1} a_r \cos \left( \frac{\pi}{N} \frac{k+1/2}{r} \right) \right] ]</td>
<td></td>
</tr>
<tr>
<td><strong>Sampling conditions:</strong></td>
<td></td>
</tr>
<tr>
<td>( \bullet ) Special case of SDFT for sampling grid shift parameters: ( u = 1/2; v = 0 ).</td>
<td></td>
</tr>
<tr>
<td>( \bullet ) Analog signal of final length is, before sampling, artificially padded with its mirror copy to form a symmetrical sampled signal of double the length: ( {a_k} = {a_{2N-1-k}} ).</td>
<td></td>
</tr>
<tr>
<td><strong>Direct and inverse scaled shifted DFTs [ScSDFT((u,v; \sigma))]</strong></td>
<td></td>
</tr>
<tr>
<td>[ a_r^{u,v} = \sum_{k=0}^{N-1} a_k \exp \left[ i2\pi \frac{(k+u)(r+v)}{\sigma N} \right] ]</td>
<td></td>
</tr>
<tr>
<td><strong>Sampling conditions:</strong></td>
<td></td>
</tr>
<tr>
<td>( \bullet ) Sampling rate is ( \sigma ) times the cardinal rate: ( \Delta x = 1/\sigma N\Delta f )</td>
<td></td>
</tr>
<tr>
<td>( \bullet ) Sampling shift parameters: ( u, v \neq 0 )</td>
<td></td>
</tr>
</tbody>
</table>

#### 18.2.3.1 Point Spread Function and Frequency Response of a Continuous Filter Equivalent to a Given Digital Filter

A linear filter is fully characterized by its point spread function (PSF), or by its Fourier transform, the filter frequency response. Given the discrete PSF \( \{h_n\} \) of a digital filter, one can, in accordance to the above formulated consistency and mutual correspondence principles, find the overall PSF \( h_{eq}(x, \xi) \) of an analog filter...
Table 18.3 Discrete representations of 2D integral Fourier transform.

<table>
<thead>
<tr>
<th>2D direct and inverse integral Fourier transforms of a signal ( a(x, y) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha(f, p) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a(x, y) \exp[i 2\pi (fx + py)] dx dy )</td>
</tr>
<tr>
<td>( a(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \alpha(f, p) \exp[-i 2\pi (fx + py)] df dp )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2D separable direct and inverse canonical DFTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_{k,l} = \frac{1}{\sqrt{N_1 N_2}} \sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} a_{r,s} \exp[-i 2\pi \left( \frac{kr}{N_1} + \frac{ls}{N_2} \right)] )</td>
</tr>
<tr>
<td>( a_{r,s} = \frac{1}{\sqrt{N_1 N_2}} \sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} a_{k,l} \exp[-i 2\pi \left( \frac{kr}{N_1} + \frac{ls}{N_2} \right)] )</td>
</tr>
</tbody>
</table>

Sampling conditions:
- Sampling in a rectangular sampling grid with cardinal sampling rates \( \Delta x = 1/N_1 \Delta f_x \), \( \Delta y = 1/N_2 \Delta f_y \).
- Zero sampling grid shift parameters.

Scaled shifted DFTs
| \( \sigma_{\alpha_{r,s}}^u \sigma_{\alpha_{r,s}}^v \sigma_{\alpha_{r,s}}^p \sigma_{\alpha_{r,s}}^q = \sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} a_{k,l} \exp[2 \pi \left( \frac{(k + u)(r + v)}{\sigma_1 N_1} + \frac{(l + p)(s + q)}{\sigma_2 N_2} \right)] \) |

Sampling conditions:
- Sampling in a rectangular sampling grid. Sampling rates \( \Delta x = 1/\sigma_1 N_1 \Delta f_x \); \( \Delta y = 1/\sigma_2 N_2 \Delta f_y \).
- Nonzero sampling grid shift parameters \( (u, v) \) and \( (p, q) \).

Rotated and scaled DFTs
| \( \sigma_{\alpha_{r,s}}^x \sigma_{\alpha_{r,s}}^y \sigma_{\alpha_{r,s}}^\theta \alpha_{r,s}^p = \sqrt{\frac{N_1 N_2}{N_1 N_2 - 1}} \sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} a_{k,l} \exp[2 \pi \left( \frac{\tilde{k} + \tilde{r}}{\sigma N} \cos \theta - \frac{\tilde{s} \sin \theta}{\sigma N} \right)] \) |

Sampling conditions:
- Sampling in a rectangular sampling grid in a rotated coordinate system
| \( \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \) with \( \theta \) as a rotation angle.
- Sampling rates \( \Delta x = 1/\sigma N \Delta f_x \); \( \Delta y = 1/\sigma N \Delta f_y \); nonzero sampling grid shift parameters \( (u, v) \) and \( (p, q) \).

Equation (18.8)

\[
\hat{h}_{eq}(x, \xi) = \sum_{k=0}^{N_h-1} \sum_{m=0}^{N_h-1} h_n \varphi^{(r)}(x - n \Delta x) \varphi^{(l)}(\xi - (k - n) \Delta x)
\]

where \( N_h \) is the number of samples of the filter output signal \( \{b_k\} \) involved in reconstruction, from its discrete representation, of analog output signal \( b(x) \) of the equivalent analog filter, \( N_h \) is the number of nonzero samples of the digital filter PSF, and \( \Delta x \) is the signal sampling interval [1].

It is more convenient to design digital filters and to characterize and analyze their equivalent continuous filters through the overall frequency response \( H_{eq}(f, p) \) of
the filter found as the Fourier transform of its overall PSF:

\[
H_{eq}(f, p) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_{eq}(x, \xi) \exp \left[ i2\pi \left( f x - p \xi \right) \right] dx \, d\xi \\
= \left[ \sum_{k=0}^{N_h-1} \exp \left[ i2\pi (f - p)k \Delta x \right]\right] \int_{-\infty}^{\infty} \varphi^{(r)}(x) \exp (i2\pi fx) dx \\
\times \int_{-\infty}^{\infty} \varphi^{(p)}(\xi) \exp (-i2\pi px) d\xi \times \left[ \sum_{m=0}^{N_h-1} h_m \exp (i2\pi pm \Delta x) \right]
\] (18.9)

This expression contains four multiplicands:

\[
SV(f, p) = \sum_{k=0}^{N_h-1} \exp \left[ i2\pi (f - p)k \Delta x \right] \tag{18.10a}
\]

\[
\Phi^{(r)}(f) = \int_{-\infty}^{\infty} \varphi^{(r)}(x) \exp (i2\pi fx) dx \tag{18.10b}
\]

\[
\Phi^{(p)}(p) = \int_{-\infty}^{\infty} \varphi^{(p)}(x) \exp (i2\pi px) dx \tag{18.10c}
\]

\[
CFrR(p) = \sum_{n=0}^{N_h-1} h_n \exp (i2\pi pn \Delta x) \tag{18.10d}
\]

The term \(SV(f, p)\) reflects the fact that the digital filters defined in Table 18.1 and obtained as a discrete representation of the convolution integral with shift invariant kernel are shift variant because a finite number \(N_h\) of samples \({b_k}\) of the filter output signal is involved in the reconstruction of the filter analog output signal \(b(x)\). This discrepancy vanishes in the limit when \(N_h \to \infty\).

The terms \(\Phi^{(r)}(.)\) and \(\Phi^{(p)}(.)\) are frequency responses of signal sampling and reconstruction devices, respectively.

The term \(CFrR(p)\) is the continuous frequency response of the digital filter. Being a Fourier series, it is a periodical function in the frequency domain. If the reconstruction and sampling basis functions of the signal are ideal sinc-functions, the sampling and reconstruction devices act as ideal low-pass filters that remove all but one of its periods. In reality, this is not the case, and therefore one should anticipate aliasing effects in the convolution results similar to those in signal reconstruction from its sampled representation.

It is shown [1, 3] that continuous frequency response of the digital filter \(CFrR(p)\) can be represented via coefficients \(\{\eta_r\}\) of DFT of the filter discrete PSF \(\{h_n\}\):

\[
\eta_r = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} h_n \exp \left[ i2\pi \frac{n - (N - 1)/2}{N} r \right] \tag{18.11}
\]
as

$$C_{FrR}(p) \propto \sum_{r=0}^{N-1} \eta_r \text{sincd} [N; \pi (p N \Delta x - r)]$$

(18.12)

where

$$\text{sincd}(N; x) = \frac{\sin x}{N \sin(x/N)}$$

(18.13)

is the discrete sinc-function and $N$ is the number of samples of the filter input signal.

These results lead to the following important conclusions:

- Given signal sampling and reconstruction devices and the number $N_b$ of samples used for reconstruction of analog signals from their samples, the overall frequency response $H_{eq}(f, p)$ of the digital filter is fully defined by DFT coefficients $\{\eta_r\}$ of the discrete PSF $\{h_n\}$ of the filter.
- These coefficients are its samples taken within the baseband with a sampling interval $1/N \Delta x$.

These conclusions constitute the base for the design of digital filters with a desired analog frequency response.

This has been illustrated in Section 18.3.

18.2.3.2 **Point Spread Function of the Discrete Fourier Analysis**

An immediate application of DFTs is discrete Fourier spectral analysis performed by applying an appropriate version of DFTs to signal samples. In this process, the computer, together with the signal sampling device, plays the role of a sampling device for the Fourier spectra of the signal; spectral samples obtained from signal samples are regarded as samples of Fourier spectra of the analog signal.

The ultimate characteristic of sampling devices is their PSF. For spectral analysis, the PSF $h_{DFA}(f, r)$ of spectral sampling, which defines the resolving power of the discrete Fourier analysis (DFA), can be found from the relationship:

$$\alpha_r = \int_{-\infty}^{\infty} \alpha(f) h_{DFA}(f, r) df$$  

(18.14)

where $\alpha(f)$ is the Fourier spectrum of signal $a(x)$ and $\{\alpha_r\}$ are its samples.

Replacing in Eq. (18.14) samples $\{\alpha_r\}$ of the Fourier spectrum of the analog signal by their expression as scaled DFTs of signal samples and linking the latter with the signal and its Fourier spectrum, one can obtain the PSF of the discrete spectrum analysis defined by the equation [1, 3]:

$$h_{DFA}(f, r) = N \text{sinc} \left[ N; \left( \frac{r}{\sigma} - f N \Delta x \right) \right] \Phi_s(f)$$  

(18.15)

where $\sigma$ is a sampling scale parameter, which, together with the number $N$ of signal samples, links the signal sampling interval $\Delta x$ with that of $\Delta f$ of its spectrum.
Equation (18.15) has a clear physical interpretation:

- Discrete spectral analysis can be treated as the sampling, with a discrete sinc-function, of the spectrum of an analog signal, masked by the sampling device frequency response $\Phi_s(f)$.
- When $\sigma = 1$, spectrum samples are taken with a cardinal sampling interval $\Delta f_{\text{card}} = 1/N \Delta x$. When $\sigma > 1$, the signal spectrum is "oversampled"; its samples are taken with sampling interval $\Delta f_{\text{card}}/\sigma < \Delta f_{\text{card}}$ and are sinc-interpolated versions of the "cardinal" samples that correspond to the case $\sigma = 1$. When $\sigma < 1$, the spectrum is "undersampled."

Figure 18.1a,b illustrate the resolving power of the DFA. Figure 18.1a shows spectra of continuous sinusoidal signals with frequencies 64/256 and 65/256 (in fraction of the width of the signal baseband) and sampled to 256 samples as well as the spectrum of the sum of these two signals. Figure 18.1b shows the same for signals with normalized frequencies 64/256 and 65.5/256. One can conclude from these figures that, in DFA, sinusoidal signals can be reliably resolved if the difference of their frequencies is larger than about $1.5/N \Delta x$.

![Figure 18.1](image-url)

**Figure 18.1** Analog spectra of two periodical signals with close frequencies and their mixture computed by means of the discrete Fourier analysis: (a) signals with frequencies 64/256 and 65/256 are not resolved; (b) signals with frequencies 64/256 and 65.5/256 are resolved in the mixture.
Building Continuous Image Models

When working with sampled images in computers, one frequently needs to return to their analog originals. Typical applications that require reconstruction of analog image models are image resampling, image restoration from sparse samples, and differentiation and integration, to name a few. In this section, we illustrate the use of the above concepts for these applications.

18.3.1
Discrete Sinc Interpolation: The Gold Standard for Image Resampling

Image resampling assumes reconstruction of a continuous approximation of the original nonsampled image by means of interpolation of available image samples to obtain samples in between the given ones. In some applications, for instance in computer graphics and print art, simple numerical interpolation methods, such as nearest neighbor or linear (bilinear) interpolations, can provide satisfactory results. In applications, such as optical metrology, that are more demanding in terms of the interpolation accuracy, higher order spline interpolation methods are frequently recommended. However, all these methods are not perfectly accurate and introduce signal distortions in addition to those caused by the primary image sampling.

The discrete signal interpolation method that is capable, within limitations defined by the given finite number of signal samples, of securing reconstruction of analog images without adding any additional interpolation errors is discrete sinc interpolation [4].

A perfect resampling filter can be designed as a discrete representation of the ideal analog shift operator with frequency response \( H(f) = \exp(i2\pi \delta x f) \), where \( \delta x \) is the analog signal shift in signal coordinate \( x \) and \( f \) is the frequency. According to Eq. (18.12), samples \( \{\eta_n = \exp(i2\pi n \delta x / N \Delta x)\} \) of this frequency response in the baseband \( N \Delta x \), defined by the sampling interval \( \Delta x \) and the number of signal samples \( N \), are examples of DFT of the filter discrete PSF \( \{h_n\} \). These samples define the discrete sinc-interpolator with discrete PSF \( \{h_n = \text{sinc} \{N, \pi [n - (N - 1)/2 - \delta x / \Delta x]\}\} \) [4].

Discrete sinc interpolation, by definition, preserves samples of the spectrum of the analog signal in its baseband. All the other convolutional interpolation methods with PSF, other than the discrete sinc-function distort signal spectrum, therefore introduce interpolation error. This means that discrete sinc interpolation can be regarded the “gold standard” of discrete signal interpolation. Discrete sinc interpolation is also competitive with other less-accurate interpolation methods in terms of the computational complexity, thanks to its implementation through fast Fourier or DCT transforms [4]. The latter implementation is the most recommended because it is virtually completely free of oscillation artifacts at signal borders characteristic for implementation through FFT.

Perfect interpolation capability of the discrete sinc interpolation in comparison to other numerical interpolation methods is illustrated in Figure 18.2, which shows
the results of rotation of a test image (Figure 18.2a) through $3 \times 360^\circ$ in 60 equal steps, using two most popular methods of image resampling, bilinear and bicubic ones (Figure 18.2b,c), and that of discrete-sinc interpolation (Figure 18.2d) [4].

Images shown in the figure clearly show that, after 60 rotations though $18^\circ$ each, bilinear and bicubic interpolation methods completely destroy the readability of the text, while discrete sinc-interpolated rotated image is virtually unchanged and is not distinguishable from the original one. Further comparison data can be found in Ref. [4].

Bilinear and bicubic interpolations are spline interpolations of the first and second order. The higher the spline order, the higher is the interpolation accuracy, and the higher the computational complexity of spline interpolation. Note that the higher the spline order, the closer their PSFs approximate the discrete sinc-function.

18.3.1.1 Signal Recovery from Sparse or Nonuniformly Sampled Data

In many applications, sampled image data are collected in an irregular fashion or are partly lost or unavailable. In these cases, it is required to restore missing data and to convert irregularly sampled signals into regularly sampled ones. This problem can be treated and solved as a least square approximation task in a framework of the discrete sampling theorem for “bandlimited” discrete signals
that have a limited number of nonzero transform coefficients in a certain transform domain [5]. We illustrate possible applications of this signal recovery technique through two examples.

One of the attractive potential applications is image super-resolution from multiple digital video frames with chaotic pixel displacements due to atmospheric turbulence, camera instability, or similar random factors [6]. By means of elastic registration of the sequence of frames of the same scene, one can determine, for each image frame and with subpixel accuracy, pixel displacements caused by random acquisition factors. Using these data, a synthetic fused image can be generated by placing pixels from all available turbulent video frames in their proper positions on the correspondingly denser sampling grid according to their found displacements. In this process, some pixel positions on the denser sampling grid will remain unoccupied, when a limited number of image frames is fused. These missing pixels can then be recovered assuming that the high-resolution image is bandlimited in a domain of certain transform. In many applications, including the discussed one, using DCT as such transform is advisable owing to the excellent energy compaction capability of DCT.

The super-resolution from multiple chaotically sampled video frames is illustrated in Figure 18.3, which shows one low-resolution frame (a); an image fused from 15 frames (b); and a result of iterative restoration (c) achieved after 50 iterations [6]. Image band limitation was set in this experiment twice the baseband of raw low-resolution images.

Yet another application that the discussed sparse data recovery technique can find in computed tomography, where it frequently happens that a substantial part of slices, which surrounds the studied body slice, is a priori known to be an empty field. This means that slice projections (sinograms) are “bandlimited” functions in the domain of the Radon transform. Therefore, whatever number of projections is available, a certain number of additional projections, commensurable, according to the discrete sampling theorem, with the size of the slice empty zone, can be obtained and the corresponding resolution increase in the reconstructed images.

![Figure 18.3](image-url) Iterative image interpolation in the super-resolution process: (a) a low-resolution frame; (b) an image fused by the elastic image registration from 15 frames; and (c) a result of bandlimited restoration of image (b).
Figure 18.4 Super-resolution in computed tomography: (a) a set of initial projections supplemented with the same number of presumably lost projections (shown dark) to double the number of projections; initial guesses of the supplemented projections are set to zero; (b) image reconstructed from initially available projections; (c) a result of restoration of missing projections; and (d) an image reconstructed from the restored double set of projections.

can be achieved. Figure 18.4 illustrates such super-resolution by means of recovery of a missing 50% of the projections achieved using the fact that, by segmentation of the restored image, it was found that the outer 55% of the image area was empty.

18.3.2 Image Numerical Differentiation and Integration

Signal numerical differentiation and integration are operations that are defined for continuous functions and require measuring infinitesimal increments or decrements of signals and their arguments. Therefore, numerical computing signal derivatives and integrals assume one or another method of building analog models of signals specified by their samples through explicit or implicit interpolation between available signal samples.

Because differentiation and integration are shift invariant linear operations, methods of computing signal derivatives and integrals from their samples can...
18 Linking Analog and Digital Image Processing

be conveniently designed, implemented, and compared through their frequency responses.

Frequency responses of ideal analog differentiation and integration filters are, correspondingly $H_{\text{diff}} = -i2\pi f$ and $H_{\text{int}} = i/2\pi f$, where $f$ is the signal frequency. Therefore, according to Eq. (18.12), samples of frequency responses of perfect digital differentiating and integrating filters must be set correspondingly to {
\[ \eta_{\text{diff}} = \frac{-i2\pi r}{N} \text{ and } \eta_{\text{int}} = \frac{iN}{2\pi r} \]}
We refer to these filters as the differentiation ramp filter and the DFT-based integration filter, respectively. Being defined in the frequency domain, these filters can be efficiently implemented by means of FFT-type fast transforms. In order to diminish boundary effects, filtering should be carried out using the DCT-based convolution algorithm [3, 4]. One can show that such numerical differentiation and integration imply the discrete sinc interpolation of signals.

In classical numerical mathematics, a common approach to numerical computing of signal derivatives and integrals is based on a Taylor series signal expansion. This approach results in algorithms implemented through discrete convolution of the signal in the signal domain. The following simplest differentiating kernels \{h^\text{diff}_{n}\} of two and five samples are recommended in manuals on numerical methods, such as [7]: \{h^\text{diff}_{D1}\} = (-1, 1) and \{h^\text{diff}_{D2}\} = (-1/12, 8/12, 0, -8/12, 1/12), which we refer to as numerical differentiating methods D1 and D2, respectively.

The best known numerical integration methods are the Newton–Cotes quadrature rules [7]. The three first rules are the trapezoidal, the Simpson, and the 3/8-Simpson ones. Another frequently used alternative is cubic spline integration [8]. In these integration methods, a linear, a quadratic, a cubic, and a cubic spline interpolation, respectively, are assumed between the sampled data.

Frequency responses of the above-described perfect and conventional numerical differentiation and integration methods are presented for comparison in Figures 18.5 and 18.6, respectively. One can see from these figures that conventional numerical differentiation and integration methods entail certain, and sometimes very substantial, distortions of spectral contents of the signal at high frequencies. In particular, numerical differentiation methods D1 and D2 attenuate signals at high frequencies. Among integration methods, Simpson and 3/8-Simpson integration methods, being slightly more accurate than the trapezoidal method at middle frequencies, tend to generate substantial artifacts if signals contain higher frequencies. Frequency response of the 3/8-Simpson rule tends to infinity at a frequency that is two-third of the maximum frequency, and frequency response of the simple Simpson rule has almost the same tendency for the maximal frequency in the baseband. This means that round-off computation errors and noise that might be present in the input data will be overamplified by Simpson and 3/8-Simpson integration filters in these frequencies.

Figure 18.7 presents results of the comparison of differentiation errors for different numerical differentiation methods applied to pseudorandom test signals of different bandwidths [4]. They convincingly evidence that D1 and D2 methods
Figure 18.5 Absolute values of frequency responses of three described numerical differentiation filters: D1, D2, and ramp filter implemented using DCT.

Figure 18.6 Absolute values, in a logarithmic scale, of frequency responses of described numerical integration filters: trapezoidal, Simpson, 3/8 Simpson, cubic spline, and DFT-based ones.
Figure 18.7 Standard deviations of differentiation errors, in fraction of standard deviation of the derivative, averaged over 100 realizations of pseudorandom test signals for D1, D2, and DCT-ramp differentiation methods as a function of test signal bandwidth (in fractions of the baseband defined by the signal sampling rate).

provide reasonably good differentiation accuracy only for signals with bandwidth less than 0.1–0.2 of the baseband, which means that a substantial signal oversampling is required in order to maintain acceptable differentiation error. But even for such signals, standard deviation for the normalized error is much lower for the DCT-ramp method. For signals with broader bandwidth, this method outperforms other methods, in terms of the differentiation accuracy, by at least 3–4 orders of magnitude.

Similar comparison and conclusions can be made for different numerical integration methods [4, 9].

18.4

Digital-to-Analog Conversion in Digital Holography. Case Study: Reconstruction of Kinoform

CGHs have numerous applications as spatial filters for optical information processing, as diffractive wavefront correcting elements for large optical objectives and telescope mirrors, beam-forming elements (for instance, for laser tweezers), laser focusers, deflectors, beam splitters, and multipliers, and for information display. Being generated in computers as sets of numbers, CGHs are then
recorded on a physical medium to be used in analog optical setups. Methods of this digital-to-analog conversion have their impact on optical characteristics of CGHs. In this section we demonstrate how this impact can be studied by means of an example of optical reconstruction of kinoform.

Kinoform is one of the simplest hologram-encoding methods for phase-modulating spatial light modulators (SLMs). In this method, amplitude components of samples of the computed holograms are forcibly set to a constant and only their phase components are preserved and used for recording on phase-modulating SLM. These samples of kinoform are recorded, one by one, on resolution cells of the SLM arranged in a rectangular grid, as illustrated in Figure 18.8.

The resulting recorded hologram can be represented mathematically as

$$\Gamma(\xi, \eta) = w(\xi, \eta) \sum_r \sum_s \Gamma_{r,s} h_{\text{rec}}(\xi - \xi_0 - r\Delta\xi, \eta - \eta_0 - s\Delta\eta)$$  \hspace{1cm} (18.16)

where \( \{\Gamma_{r,s}\} \) are samples of the computed hologram with \( \{r, s\} \) as 2D sample indices; \((\xi, \eta)\) are physical coordinates on the hologram-recording SLM; \((\Delta\xi, \Delta\eta)\) are sampling intervals of the rectangular sampling grid along coordinates; \((\xi_0, \eta_0)\) are coordinates of the first sample of the hologram in the reconstruction setup; \(h_{\text{rec}}(\xi, \eta)\) are hologram-recording device aperture functions, and \(w(\xi, \eta)\) is a hologram window function that defines the physical dimensions of the recorded hologram: \(0 \leq w(\xi, \eta) \leq 1\), when \((\xi, \eta)\) belongs to the hologram area and \(w(\xi, \eta) = 0\), otherwise.

At the reconstruction stage, this hologram is subjected to analog optical transformation in an optical setup to reconstruct an image. Assume that the computer-generated kinoform is designed for image reconstruction in the far diffraction zone or in the focal plane of a lens, which corresponds to an optical Fourier transform. In this case, the reconstruction image can be represented as

$$A_{\text{restr}}(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Gamma(\xi, \eta) \exp \left(-i2\pi \frac{x\xi + y\eta}{\lambda Z} \right) d\xi d\eta$$  \hspace{1cm} (18.17)

Figure 18.8 Definitions related to recorded physical computer-generated holograms.
where \( \lambda \) is the wavelength of reconstructing coherent illumination and \( Z \) is the hologram-to-object distance or the focal length of the lens.

Samples \( \{ \Gamma_{r,s} \} \) of the kinoform are linked with samples of the input image \( \{ A_{k,l}^{(k)} \} \) through, in general, shifted DFT:

\[
\Gamma_{r,s} = \sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} A_{k,l}^{(k)} \exp \left\{ i2\pi \left( \frac{(k+u)(r+p)}{N_1} + \frac{(l+v)(s+q)}{N_2} \right) \right\} \tag{18.18}
\]

where \( \{ k, l \} \) are pixel indices, \( \{ N_1, N_2 \} \) are input image dimensions, and \( (u, v) \) and \( (p, q) \) are shift parameters associated with shifts of samples of the input image with respect to its coordinate system and shifts of kinoform samples with respect to the coordinate system of the optical setup. Note that the input image \( \{ A_{k,l}^{(k)} \} \) should be, before the synthesis of kinoform, appropriately preprocessed to secure minimal distortions of the reconstructed image caused by neglecting amplitude components of samples of computed hologram (for details see, for instance, Ref. [2]).

Substituting Eq. (18.16) in Eq. (18.17) and replacing, in the former, samples of kinoform \( \{ \Gamma_{r,s} \} \) by their expression (Eq. (18.18)) through input image samples \( \{ A_{k,l}^{(k)} \} \), one can obtain, after completing corresponding integrations and summations, that the image reconstructed by the kinoform in the optical setup, is described, for shift parameters \( p = q = 0 \) and \( \xi_0 = \eta_0 = 0 \), by the following expression [2]:

\[
A_{\text{recon}}(x,y) = \sum_{do_x=-\infty}^{\infty} \sum_{do_y=-\infty}^{\infty} \sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} \left( \frac{\lambda Z}{N_1} \Delta \xi \right) \left( \frac{\lambda Z}{N_2} \Delta \eta \right) A_{k,l}^{(k)} \times H_{\text{rec}} \left[ \left( \frac{k+u}{N_1} - do_x \right) \frac{\lambda Z}{N_1} \Delta \xi \right] \times W \left( x - \frac{k+u \lambda Z}{N_1} \Delta \xi - do_x \frac{\lambda Z}{N_1} \Delta \xi \right) \times W \left( x - \frac{l+v \lambda Z}{N_2} \Delta \eta - do_y \frac{\lambda Z}{N_2} \Delta \eta \right) \tag{18.19}
\]

where

\[
W(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi(\xi, \eta) \exp \left( -i2\pi \frac{\xi \bar{\xi} + \eta \bar{\eta}}{\lambda Z} \right) d\xi d\eta \tag{18.20}
\]

is the Fourier transform of the hologram window function \( \psi(\xi, \eta) \) and

\[
H_{\text{rec}}(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_{\text{rec}}(\xi, \eta) \exp \left[ -i2\pi \frac{\xi \bar{\xi} + \eta \bar{\eta}}{\lambda Z} \right] d\xi d\eta \tag{18.21}
\]

is the Fourier transform of the hologram-recording aperture \( h_{\text{rec}}(\xi, \eta) \), or frequency response of the hologram-recording device.

Equation (18.19) has a clear physical interpretation illustrated in Figure 18.9:

- The object wavefront is reconstructed in a number of diffraction orders \( \{ do_x, do_y \} \).
18.5 Conclusion

Optimal design of image processing systems is always hybrid with appropriately selected sharing of functions between analog, optical, and digital components of the system. This requires characterization and design of digital image processing in terms of the hybrid processing. In this article, we addressed the following aspects of this problem:

- discrete representation of analog convolution and Fourier transforms for sampled signals, which maintains mutual correspondence with analog and discrete transform;
- characterization of digital filtering and discrete Fourier analysis in terms of corresponding analog operations;

- In each particular diffraction order \((d_{0x}, d_{0y})\), the reconstructed wavefront is a result of the interpolation of samples \(A_{k}^{(0)}\) of the object wavefront used for synthesis of kinoform, with an interpolation kernel \(W(x, y)\), which is the Fourier transform of the recorded hologram window function \(w(\xi, \eta)\); samples of this object wavefront are weighted by samples of the frequency response \(H_{rec}(x, y)\) of the hologram-recording device.

This is how physical parameters of the SLM used for recording computer-generated kinoform and those of the optical setup used for reconstructing images affect reconstructed images.

Figure 18.9  (a) Computer simulation of optical reconstruction of computer-generated kinoform; numbers are diffraction order indices \(\{d_{0x}, d_{0y}\}\). (b) Reconstruction masking function \(H_{rec}(x, y)\) for a rectangular hologram-recording aperture of size \(\Delta \xi \times \Delta \eta\).
• building continuous image models from their sampled representation on examples of image resampling with perfect discrete sinc interpolation, image recovery from sparse data, and signal differentiation and integration;
• digital-to-analog conversion in computer-generated holography (with an example of optical image reconstruction of computer-generated kinoforms).

References


