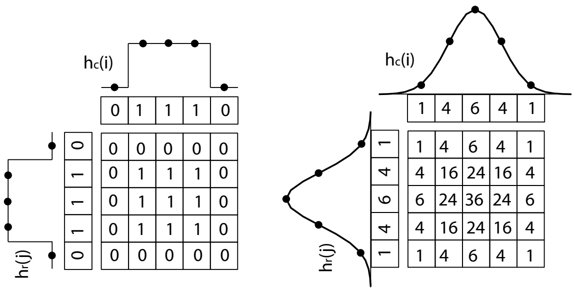


**Fig. 9.41** Examples of 2D separable convolution masks rectangular and Gaussian function

With modern multiprocessors and dedicated processing systems (example pipeline architectures available on PC video cards) convolutions can be achieved with reasonable calculation times. In implementing the convolution, you will also have to take into account the time of access to the mass storage devices to access the input image (which can be large, even of several gigabytes) and to save the output image in processing. In the convolution equation, the already processed $g(i, j)$ pixels are not involved during the convolution process. This implies that the convolution produces as a result of a new image $g(i, j)$ that must be saved in a memory area separate from the input image $f$. In the hypothesis of working on sequential processor, the necessary memory could still be optimized, saving the processed pixels of $r$ lines above the *i-th* line being processed.

An alternative approach is to temporarily save in a buffer $(r + 1)$ lines of the input image being processed, while the pixel under consideration $(i, j)$ is saved in the same position as the input image. In the case of a separable mask, while the calculation time is optimized, on the other hand, additional memory space is required to manage the intermediate result of the first 1D convolution that must be reused by applying the second 1D convolution.

## 9.11 Filtering in the Frequency Domain

Previously we introduced the usefulness of studying the spatial structures of an image in the frequency domain that more effectively describe the periodic spatial structures present in the image itself. To switch from the spatial domain to the frequency domain, there are several operators that are normally called transformation operators or simply transformed. Such transforms, e.g., that of Fourier which is the best known, when applied to the images, they decompose it from the gray level structures of the spatial domain to the components in fundamental frequencies in the frequency domain.

Each frequency component is expressed through a phase and modulus value. The inverse transform converts a structured image into frequencies, reconstructing the original spatial structures of the image backward. The complete treatment of the Fourier transform and the other transforms are described in the following chapters.

The next paragraph briefly describes the discrete Fourier transform (DFT) for digital filtering aspects in the frequency domain.

### 9.11.1 Discrete Fourier Transform DFT

The DFT is applied to an image $f(k, l)$ with a finite number of elements $N \times M$, obtained from a sampling process at regular intervals, and associates to $f(k, l)$ the matrix $F(u, v)$ of dimensions $N \times M$ given by

$$F(u, v) = \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} f(k, l) B(k, l; u, v) \tag{9.85}$$

where $F(u, v)$ represent the coefficients of the transform (or Fourier image), and $B(k, l; u, v)$ indicates the images forming the base of the frequency space identified by the $u$-$v$ system (variables in the frequency domain) each with dimension $k \times l$. In essence, the $F(u, v)$ coefficients of the transform represent the projections of the image $f(k, l)$ on the bases. These coefficients indicate quantitatively the degree of similarity of the image with respect to the bases $B$. The transformation process quantifies the decomposition of the input image $f(k, l)$ in the weighted sum of the base images, where the coefficients $F(u, v)$ are precisely the weights. The Eq. (9.85) can also be interpreted considering that the value of each point $F(u, v)$ of the *Fourier image* is obtained by multiplying the spatial image $f(k, l)$ with the corresponding base image $B(k, l; u, v)$ and adding the result.

The values of the frequencies near the origin of the system $(u, v)$ are called *low frequencies* while those farthest from the origin are called *high frequencies*. $F(u, v)$ is a continuous and complex functions. The input image $f(k, l)$ can be reconstructed (re-transformed) in the spatial domain through the coefficients of the transform $F(u, v)$ with the equation of the inverse Fourier transform, that is,

$$f(k, l) = F^{-1}(F(u, v)) = \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} F(u, v) B^{-1}(k, l; u, v) \tag{9.86}$$

The Fourier transform applied to an image produces a Fourier coefficient matrix of the same image size that fully represents the original image. The latter can be reconstructed with the inverse Fourier transform (9.86) which does not cause any loss of information. As we will see later, it is possible to manipulate the pixels of the image $f(k, l)$ in the spatial domain and see how it changes in the frequency domain (or spectral), or vice versa to modify the Fourier coefficients $F(u, v)$ in the spectral domain and see how the original image is modified after reconstruction.

The basic images of the transformation are represented by sine and cosine functions and the transformation of the image $f(k, l)$ is given by

$$F(u,v) = \frac{1}{\sqrt{NM}} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} f(k,l) \cdot \left[\cos\left(2\pi\left(\frac{uk}{N} + \frac{vl}{M}\right)\right) + j \sin\left(2\pi\left(\frac{uk}{N} + \frac{vl}{M}\right)\right)\right] \tag{9.87}$$

in which the variables $(u, v)$ represent the spatial frequencies. The function $F(u, v)$ represents the frequency content of the image $f(k, l)$, which is complex and periodic in both $u$ and $v$ with period $2\pi$. The cosine represents the real part and the sine is the complex part, thus obtaining the general expression

$$F(u, v) = R_e(u, v) + jI_m(u, v) \tag{9.88}$$

#### 9.11.1.1 Magnitude, Phase Angle, and Power Spectrum

The real component $R_e(u, v)$ and the imaginary component $I_m(u, v)$ of the complex coefficients $F(u, v)$ constituting the Fourier image do not have useful informative content. A more effective representation is obtained by representing each complex coefficient $F(u, v)$ through its *magnitude* $|F(u, v)|$ and *phase* $\Phi(u, v)$. The *spectral magnitude (or amplitude)* is defined by

$$|F(u, v)| = \sqrt{R_e^2(u, v) + I_m^2(u, v)} \tag{9.89}$$

which specifies how much of the intensities (magnitudes) of the base images are present in the input image, while the information about the orientation and shifts of the object are encoded by the *phase angle* given by

$$\Phi(u, v) = \tan^{-1}\left[\frac{I_m(u, v)}{R_e(u, v)}\right] \tag{9.90}$$

The Fourier transform can be written in terms of its magnitude and phase

$$F(u, v) = R_e(u, v) + jI_m(u, v) = |F(u, v)|\, e^{j\Phi(u,v)}$$

The *power spectrum or spectral density* $P(u, v)$ of an image is defined as follows:

$$P(u, v) = |F(u, v)|^2 = R_e^2(u, v) + I_m^2(u, v) \tag{9.91}$$

The Fourier transform (9.87) is useful to represent it also in a complex exponential form by applying to the trigonometric form the Euler relation $e^{jx} = \cos x + j \sin x$, when rewritten becomes

$$F(u, v) = \frac{1}{\sqrt{NM}} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} f(k, l) \cdot e^{-2\pi j\left[u\frac{k}{N} + v\frac{l}{M}\right]} \tag{9.92}$$

and the inverse DFT is

$$f(k, l) = \frac{1}{NM} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} F(u, v) e^{2\pi j\left[k\frac{u}{N} + l\frac{v}{M}\right]} \tag{9.93}$$

In physical reality a 2D signal (image) is obtained by the finite superposition of sinusoidal components $e^{j2\pi(uk+vl)}$ for real values of $(u, v)$ frequencies. The 2D DFT, being applied to a function with limited support (it is also said with limited band) or a sample image with a finite number of elements, constitutes a particular form of the continuous Fourier transform. The DFT realizes with Eqs. (9.92) and (9.93) a bijection and linear correspondence between spatial domain represented by the image $f(k,l)$ and spectral domain $F(u,v)$. In other words, fixed the Fourier bases, given the input image $f(k,l)$ can be obtained with the (9.92) the spectral information given by $F(u,v)$ (also called the *spectrum* of $f$), and vice versa, starting from the frequency domain $F(u,v)$ we can reconstruct the original image with the inverse transform given by Eq. (9.93).

#### 9.11.1.2 DFT in Image Processing

In general, DFT is a complex function. Therefore, the results of the transform $F(u,v)$, a matrix of complex coefficients, can be visualized through the decomposition of the latter in terms of the three matrices, respectively, of the *magnitude, phase angle, and spectral power*. For an image with dimensions $M \times N$, these quantities are defined with matrices of the same dimensions of the image. In the Fourier domain, the sinusoidal components that make up the image are characterized by these magnitudes that encode the information in terms of spatial frequencies, amplitude, and phase.

The *magnitude* represents the information associated with the contrast or rather the variation (modulation) of intensity in the spatial domain at the various frequencies. The *phase* represents how the sinusoidal components are translated from the origin. All sinusoidal components of which the image is composed are encoded in terms of magnitude and phase in the Fourier domain for all spatial frequencies in the discrete range for $u = 0, 1, 2, \ldots, M-1$ and $v = 0, 1, 2, \ldots, N-1$. In other words, the DFT encodes all sinusoidal components from the zero frequency to the maximum possible frequency (according to the Nyquist frequency) depending on the spatial resolution possible for the image. In correspondence of $(u,v) = (0,0)$ there is no modulation, in fact, from the Eq. (9.92) we have that $F(0,0)$ is proportional to the average value of the image intensity

$$F(0,0) = \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} f(k,l) \tag{9.94}$$

which is referred to as the DC component (also called a continuous or constant component) of the DFT. The DC component represents the sum of the input values $f$, to give it the meaning of the average often in the Fourier transform is inserted the term $\frac{1}{M \cdot N}$. A term with DC zero would mean an image with an average brightness of zero, which would mean the alternating sinusoid between positive and negative values in the input image. In the context of images represented by real functions, the DC component has positive values. It is also highlighted that the DFT of a real function generates a real part equal $F(u,v) = F(-u,-v)$ and an odd imaginary part. It follows that the DFT of a real function is conjugated symmetric $F^*(u,v) = F(-u,-v)$ (also called Hermite function), with the resulting spectrum a function

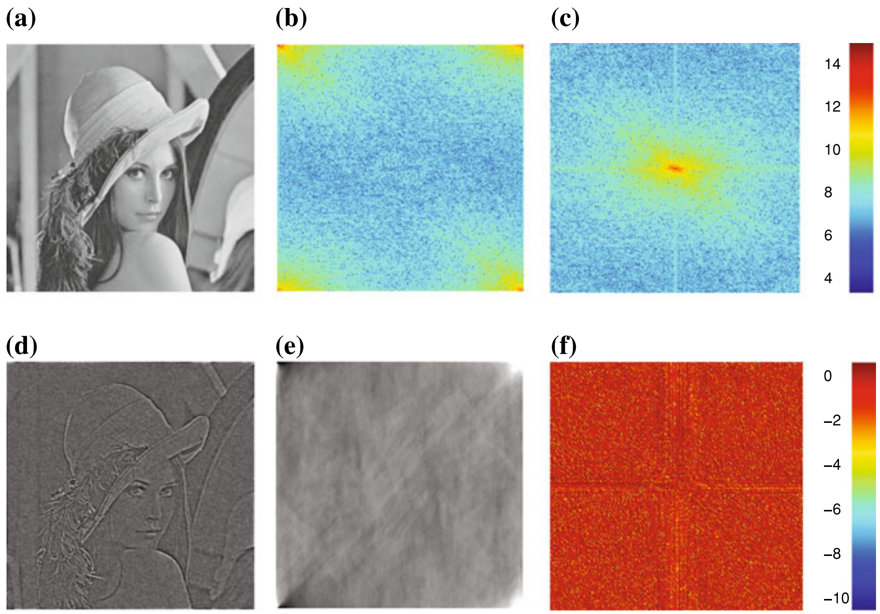


**Fig. 9.42** The discrete Fourier transform applied to the **a** gray level image; **b** the spectrum not translated; **c** the spectrum translated with the DC component at the center of the matrix; **d** the original image **a** reconstructed with the inverse DFT using only the phase matrix; **e** the original image reconstructed using only the spectrum; and **f** the phase angle of the image

equal to the origin, i.e., $|F(u, v)| = |F(-u, -v)|$ while the phase has odd symmetry with $\Phi(u, v) = -\Phi(-u, -v)$.

For better visualization of the magnitude and the power spectrum, considering that the values of the spectrum are very high starting from the DC component and then being very compressed in the high frequencies, it is convenient to modify the dynamic range of amplitudes with a logarithmic law, that is,

$$F_L(u, v) = c \cdot \log\left[1 + |F(u, v)|\right] \tag{9.95}$$

where the constant $c$ is used to scale the variability range of $F(u, v)$ with the dynamic range of the monitor (normally between 0 and 255).

Figure 9.42 shows the results of the DFT applied to an image. In Fig. 9.42b the spectrum is displayed after the logarithmic transformation has been applied, the Eq. (9.95). It is observed that the higher values of the spectrum are at the four angles near the origin of the transform (the DC component, i.e., $F(0, 0)$ results in an angle).

A peculiarity of the DFT is to be periodic.[1] The sinusoidal components of which an image is composed are repeated indefinitely, and therefore, the image of the DFT repeats indefinitely with period $M$ and $N$, respectively, along the axes $u$ e $v$

$$F(u, v) = F(u + M, v) = F(u, v + N) = F(u + M, v + N) \tag{9.96}$$

It is shown by the Fourier analysis that the DFT of a matrix of $M \times N$ dimensions of real data (the case of images) generates a data matrix of the transform of the same dimensions and indexing $u = 0, 1, 2 \ldots, M - 1$ and $v = 0, 1, 2 \ldots, N - 1$ but with the following meaning of the frequencies. The one relative to $(u, v) = (0, 0)$ represents the DC term given by (9.94). The frequencies indexed by $(u, v) = (1, 1), (2, 2), \ldots, (M/2 - 1, N/2 - 1)$ are considered positive while the negative ones are relative to the interval from $(u, v) = (M/2 + 1, N/2 + 1, \ldots, (N - 1)$. At the frequency $(u, v) = (M/2, N/2)$ applying Eq. (9.92) the value of the transform is

$$\begin{aligned} F\left(\frac{M}{2}, \frac{N}{2}\right) &= \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} f(k, l) \cdot e^{-2\pi j\left[\frac{Mk}{2M} + \frac{Nl}{2N}\right]} \\ &= \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} f(k, l) \cdot e^{-\pi j[k+l]} = \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} (-1)^{k+l} \cdot f(k, l) \end{aligned} \tag{9.97}$$

which can be considered as an alternative to the DC term sum of the input values.

This term is real and constitutes the intersection point of the four parts in which the spectrum is divided at the frequency $(u, v) = (M/2, N/2)$. Similarly, the symmetry property and the relationship of complex conjugates between the symmetric components with respect to the frequency $(u, v) = M/2, N/2)$ can be verified

$$\begin{gathered} F\left(\tfrac{M}{2} + 1, \tfrac{N}{2} + 1\right) = F^*\left(\tfrac{M}{2} - 1, \tfrac{N}{2} - 1\right) \\ F\left(\tfrac{M}{2} + 2, \tfrac{N}{2} + 2\right) = F^*\left(\tfrac{M}{2} - 2, \tfrac{N}{2} - 2\right) \\ \cdots \\ F(1, 1) = F^*(M - 1, N - 1) \end{gathered} \tag{9.98}$$

It can be concluded that for real input data, all the information of the spectrum resides in $(u, v) = (0, 0)$ corresponding to the DC component, in the components indexed by $u = 1, 2, \ldots, M/2-2, M/2-1$ and $v = 1, 2, \ldots, N/2-2, N/2-1$, and to the component indexed by $(u, v) = (M/2, N/2)$. All the other components, complex conjugate of the previous ones, indexed by $u = M/2+1, M/2+2, \ldots, M-2, M-1$ and $v = N/2+1, N/2+2, \ldots, N-2, N-1$ are not used because they do not give an added value.

[1]The periodicity property, given by the Eq. (9.96), is demonstrated by applying the equation of the Fourier transform (9.92) in the point $(u + M, v + N)$ and simplifying we get $F(u + M, v + N) = F(u, v)$.

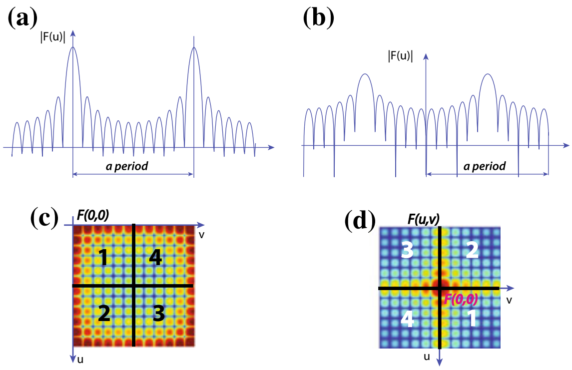


**Fig. 9.43** Graphical representation of the periodicity and symmetry properties of the DFT with an infinite number of replicas of the spectrum: **a** The DFT of the 1D rectangle function with the nontranslated spectrum; **b** with the spectrum translated with the term DC centered; **c** the DFT of the 2D rectangle function with the nontranslated spectrum indicated by the central rectangle; **d** the DFT translated into the spatial domain with the exponential factor $(-1)^{k+l}$ with the aim of reallocating the spectrum centered with respect to the DC term

The properties of periodicity Eq. (9.96) and of conjugated symmetry Eq. (9.98) of the DFT inform us that in a period (interval $[0 : M-1, 0 : N-1]$) in the spectral domain there are four quadrants in which are located the samples indexed, respectively, by $u = 1, 2, \ldots, M/2-1$ and $u = M/2+1, M/2+2, \ldots, M-1$ and by $v = 1, 2, \ldots, N/2-1$ and $v = N/2+1, N/2+2, \ldots, N-1$, and the replicates of related complex conjugate samples.

Figure 9.43 displays the situation better in the 1D and 2D context. As highlighted above all the information of the real data transform is available in an entire period that can be better used with a translation into the location $(u, v) = (M/2, N/2)$ of the current spectrum origin, coinciding with the DC component, as suggested by Eq. (9.97). It should be noted that the inverse transform of the DFT regenerates the periodic input function $f(k, l)$ with the period inherited from the DFT.

From the translation property of the DFT it is also known that a displacement of $(\triangle u, \triangle v)$ in the frequency domain produces

$$F(u, v) \Longleftrightarrow f(k, l) \Longrightarrow F(u - \triangle u, v - \triangle v) \Longleftrightarrow e^{2\pi j\left[\triangle u \frac{k}{M} + \triangle v \frac{l}{N}\right]} f(k, l), \quad (9.99)$$

that is, translation introduces an exponential multiplicative factor in the spatial domain. The reallocation of the spectrum with the displacement of the origin in the location $(M/2, N/2)$, in the central point of the spectral matrix is obtained as follows:

$$F(u - M/2, v - N/2) \Longleftrightarrow f(k, l) \cdot e^{\pi j[k+l]} = f(k, l) \cdot (-1)^{k+l} \qquad (9.100)$$

From these relationships we can, therefore, reallocate the spectrum by deriving the following:

$$F(u - M/2, v - N/2) = \mathcal{F}[f(k,l) \cdot (-1)^{k+l}] \tag{9.101}$$

which suggests carrying out the translation in the spatial domain by multiplying each input datum by the factor $(-1)^{k+l}$ before the transform or by operating after the transformation with a direct translation in the frequency domain. In Fig. 9.42c is shown the spectrum of Fig. 9.42b translated with the term DC at the center of the spectral matrix. It is observed that the translation does not modify the content of the spectrum (magnitude invariance with the translation of the input data in the spatial domain).

This is also highlighted by (9.99) where the coefficient of the exponential term is 1. Also remembering that

$$e^{2\pi j\left[\triangle u\frac{k}{N}+\triangle v\frac{l}{M}\right]} = \cos\left(2\pi j\left[\triangle u\frac{k}{M} + \triangle v\frac{l}{N}\right]\right) + j\sin\left(2\pi j\left[\triangle u\frac{k}{M} + \triangle v\frac{l}{N}\right]\right) \tag{9.102}$$

and it is instead noted that the phase is altered in proportion to the degree of translation $(\triangle u, \triangle v)$.

The effects of the rotation of an image in the frequency domain is better explained by operating in polar coordinates given by the following expressions:

$$k = r \cdot \cos\theta \qquad l = r \cdot \sin\theta \qquad u = \omega\cos\phi \qquad v = \omega\sin\phi \tag{9.103}$$

Expressing the Fourier pair with $f(r,\theta)$ and $F(\omega,\phi)$, and considering a rotation of the image $f(k,l)$ of $\theta_0$, the Fourier pair results

$$f(r, \theta + \theta_0) \Longleftrightarrow F(\omega, \phi + \theta_0) \tag{9.104}$$

from which it emerges that with the rotation of $f(k,l)$ of an angle $\theta_0$ in the spatial domain there is an identical rotation of the power spectrum $F(u,v)$ and vice versa a rotation in the spectral domain corresponds to the same rotation in the spatial domain (see Fig. 9.44). From previous considerations we can highlight that, in the case of images, the information content of the power spectrum, through the magnitude of sinusoidal components, captures the intensity levels (levels of gray, color, ...), while the phase spectrum, through the shift values relative to the origin of the sinusoids, captures the information associated with the image morphology, orientation, and movement of the objects.

In Fig. 9.42d we can see the reconstruction of the original image ($a$) using the inverse transform only the phase angle matrix displayed in Fig. 9.42f. Despite the lack of intensity information, the entire image is reconstructed with all the original morphological information corresponding to the person's face. Therefore, without information on the phase, the spatial structure of the image is completely devastated to such an extent that it makes it impossible to recognize the objects present. Instead, the reconstruction of the original image using only the magnitude matrix is shown in Fig. 9.42e which highlights the total absence of the morphological structures of the image and how it was expected presents the spatial distribution of the intensity

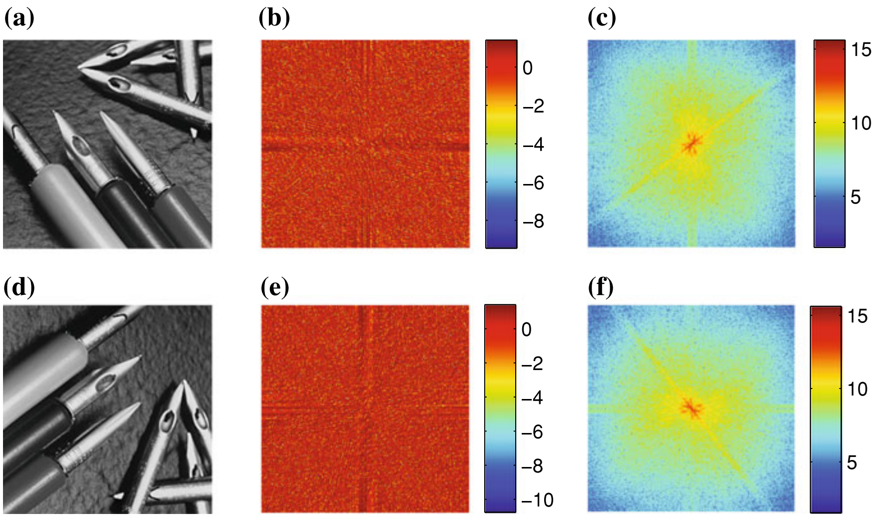


**Fig. 9.44** The discrete Fourier transform applied to the **a** gray level image; **b** the phase spectrum of the image (**a**); **c** the power spectrum of the image (**a**); **d** the original image (**a**) rotated by 45° (**e**) the phase spectrum of the image (**d**) rotated which is modified; and **f** the power spectrum of the rotated image which is identical to that of the original image

of the source image. With DFT and Inverse DFT (IDFT) it is possible to reconstruct complex images by adequately combining the phase angle matrices obtained from different images useful to create particular effect images.

From the analysis of the Fourier transform, it follows that in image processing applications, operating in the frequency domain, the filtering algorithms can only modify the spectrum of the amplitude while that of the phase cannot be modified.

#### 9.11.1.3 Separability of the DFT

With reference to Eq. (9.92), the exponential $e^{-2\pi j\left[u\frac{k}{N}+v\frac{l}{M}\right]}$ can be factored by rewriting it in the following form:

$$\begin{aligned} F(u,v) &= \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}\left(\frac{1}{\sqrt{M}}\sum_{l=0}^{M-1} f(k,l)\cdot e^{-2\pi j v\frac{k}{M}}\right)e^{-2\pi j u\frac{l}{N}} \\ &= \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} F(k,v)e^{-2\pi j u\frac{l}{N}} \end{aligned} \tag{9.105}$$

with the expression in large brackets indicating the 1D transform F (k, v). Therefore, the 2D DFT transform is realized through 2 1D transforms: *first a 1D transform is performed on each column of the input image obtaining the intermediate result*

*F(k,v) and then the 1D transform is performed on each row of the intermediate result F(k,v).*

#### 9.11.1.4 Fast Fourier Transform (FFT)

To optimize the computational load of the DFT transform several algorithms have been developed that minimize redundant operations through the DFT separability property, symmetry, and intelligent data reorganization. The most common algorithm is Cooley-Tukey [1,3] which coined the term FFT (*Fast Fourier Trasnsform*). An image of size $N \times N$ would require a computational complexity of $\mathcal{O}(N^4)$ operations to obtain the DFT transform. The FFT algorithm reduces the computational load to $\mathcal{O}(N^2 logN)$. With the separability of the DFT, the FFT transform is performed directly as two 1D transforms. Given the various applications of DFT in the field of image processing, the analysis of audio signals, the analysis of data in the field of physics and chemistry, the firmware versions of the FFT algorithm has been implemented.

### 9.11.2 Frequency Response of Linear System

Previously we described the linear spatially invariant (LSI) systems for which the principle of superimposition is valid: *the response to a linear combination of signals is the linear combination of responses to individual signals (like step or ramp signals)*. We know that the behavior of such systems is completely defined by the response of the system to the pulse function $\delta(x)$. In fact, if $\mathcal{O}$ is the linear operator of an LSI and $h(x)$ is the response $h(x) = \mathcal{O}\{\delta(x)\}$ of the system to the pulse $\delta(x)$, then the response $g(x) = \mathcal{O}\{f(x)\}$ to any input $f(x)$ is given by the *convolution* $g = f * h$.

> What happens to an LSI system if it is stressed by a frequency signal represented as a complex exponential of the type $f(x) = e^{jux}$?

$$e^{jux} \quad \Longrightarrow \quad H(u) \cdot e^{jux}$$

It can be observed that the LSI system produces in response an analogous signal of a complex exponential type with the same frequency $u$ and only changed in amplitude by the factor $H(u)$. By applying the theory of linear systems we can demonstrate this property. In fact, if we consider an LSI system with impulse response $h(x)$, we know that for every input $f(x) = e^{jux}$ the response of the system $g(x)$ is determined by (9.70) the convolution integral, as follows:

$$g(x) = h(x) * f(x) = \int_{-\infty}^{\infty} h(\tau) \cdot f(x-\tau)\mathrm{d}\tau$$
$$= \int_{-\infty}^{\infty} h(\tau) \cdot e^{ju(x-\tau)}\mathrm{d}\tau$$
$$= e^{\mathrm{jux}} \int_{-\infty}^{\infty} h(\tau) \cdot e^{-ju\tau)}\mathrm{d}\tau$$

If the integral of the last member is defined, the signal at the output of the system is in the form

$$g(x) = H(u) \cdot e^{jux}$$

with

$$H(u) = \int_{-\infty}^{\infty} h(\tau) \cdot e^{-ju\tau)}\mathrm{d}\tau \tag{9.106}$$

where the function $H(u)$, the frequency response, is called the *Transfer Function* of the system. It is noted that the function $H(u)$ given by the (9.106) corresponds to the Fourier transform, i.e., it represents the spectrum of the impulse response $h(x)$. The extension to 2D signals $f(x, y)$ with a 2D transfer function $H(u, v)$ is immediate by virtue of the DFT Eq. (9.92). This remarkable result justifies the interest of the Fourier transform for the analysis of LSI systems.

At this point we can be interested to understand what is the frequency response $G(u, v)$ of the system knowing the transfer function $H(u, v)$ and the Fourier transform $F(u, v)$ of the 2D input signal $f(x, y)$.

The *answer to this question* is given by the *Convolution Theorem*.

### 9.11.3 Convolution Theorem

Given two functions $f(x)$ and $h(x)$, note the corresponding Fourier transforms $F(u)$ and $H(u)$, we know that the spatial convolution between the two functions is given by

$$h(x) * f(x) = \int_{-\infty}^{\infty} h(\tau) \cdot f(x-\tau)\mathrm{d}\tau$$

If with $\mathcal{F}$ we indicate the Fourier transform operator applied to the convolution $h(x) * f(x)$ we obtain

$$
\begin{aligned}
\mathcal{F}[h(x) * f(x)] &= \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} h(\tau) \cdot f(x-\tau) \mathrm{d}\tau \right] e^{-jux} \mathrm{d}x \\
&= \int_{-\infty}^{\infty} h(\tau) \int_{-\infty}^{\infty} f(x-\tau) e^{-jux} \mathrm{d}x \mathrm{d}\tau \\
&= \int_{-\infty}^{\infty} h(\tau) e^{-ju\tau} \cdot F(u) \mathrm{d}\tau \\
&= F(u) \int_{-\infty}^{\infty} h(\tau) e^{-ju\tau} \mathrm{d}\tau \\
&= F(u) \cdot H(u)
\end{aligned} \tag{9.107}
$$

where the first expression represents the Fourier transform of $h(x) * f(x)$ (shown in square brackets), in the second expression the integrals are exchanged, in the third one the spatial translational property of the transform is applied, while in the fourth, it is observed that the integral corresponds to the Fourier transform of the function $h(x)$, that is, the Eq. (9.106).

With the convolution theorem it has been shown that

$$
\mathcal{F}[h(x) * f(x)] = H(u) \cdot F(u) \tag{9.108}
$$

and it can be affirmed that

*the convolution between two functions in the spatial domain corresponds to the simple multiplication of the functions in the frequency domain.*

By applying the inverse Fourier transform to the product $H(u) \cdot F(u)$ we get back the convolution of the functions in the spatial domain

$$
\mathcal{F}^{-1}[H(u) \cdot F(u)] = h(x) * f(x) \tag{9.109}
$$

The *convolution theorem* allows the study of the input/output relationship of an LSI system to be realized in the spectral domain thanks to the transformed and antitrasformed pair (Eqs. 9.108 and 9.109) rather than in the spatial domain through the *convolution* that we know is computationally complex and less intuitive than simple multiplication in the frequency domain. The convolution theorem also responds to the question posed at the conclusion of the previous paragraph on the prediction of the $G(u)$ response of an LSI system when stimulated in input by a sinusoidal signal $f(x)$ known the transfer function $H(u)$ of the system.

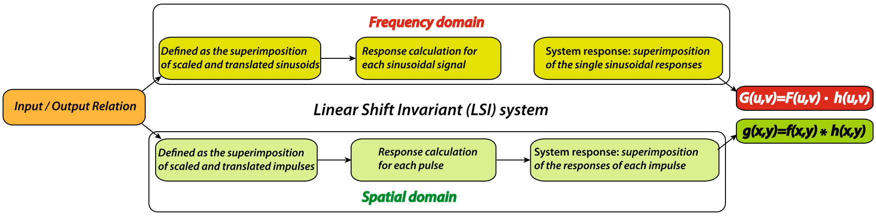


**Fig. 9.45** Operative domains of an LSI system: spatial and frequency domain

Being able to obtain the Fourier transform $F(u)$ of the input signal $f(x)$, the output of the LSI system, results

$$F(u) \quad \overset{H(u)}{\Longrightarrow} \quad G(u) = H(u) \cdot F(u)$$

The convolution theorem asserts, in other words, that an LSI system can combine in the spatial domain the signals $f(x)$ and $h(x)$ in such a way that the frequency components of $F(u)$ are scaled by the frequency components of $H(u)$ or vice versa. Figure 9.45 graphically schematizes the behavior and response of an LSI system when it is stimulated by signals in the spatial domain and in the frequency domain.

#### 9.11.3.1 Frequency Convolution

The frequency convolution property states that the convolution in the frequency domain corresponds to multiplication in the spatial domain

$$F(u) * H(u) = \int_{-\infty}^{\infty} F(\tau) \cdot H(u-\tau)d\tau = \mathcal{F}[f(x) \cdot h(x)] \tag{9.110}$$

derived by virtue of the *duality (symmetry)* property, i.e. $\mathcal{F}[F(x)] = f(-u)$ of the Fourier transform. We can say that: a function $f(x)$ is *modulated* by another function $h(x)$ if they are multiplied in spatial domain.

#### 9.11.3.2 Extension to 2D Signals

The extension of the previous formulas for 2D signals is immediate and using the same symbolism we rewrite the equations of the convolution theorem in the 2D spatial domain

$$g(x, y) = f(x, y) * h(x, y) \tag{9.111}$$

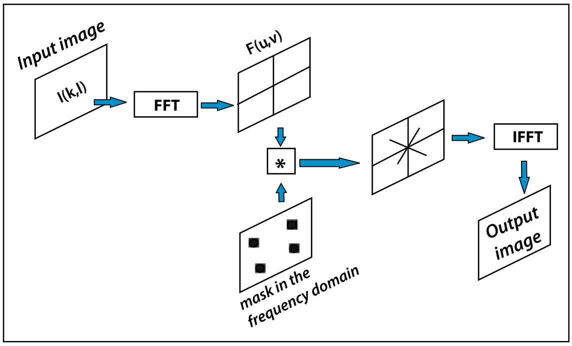


**Fig. 9.46** Filtering process in the frequency domain using the FFT/IFFT Fourier transform pair

and

$$G(u, v) = F(u, v) \cdot H(u, v) \tag{9.112}$$

and the convolution in the frequency domain

$$F(u, v) * H(u, v) = \mathcal{F}[f(x, y) \cdot h(x, y)] \tag{9.113}$$

considering $G$, $F$, and $H$ (transfer function) the Fourier transforms of the images, respectively, $g$ and $f$, and of the impulse response function $h$.

#### 9.11.3.3 Selection of the Filtering Domain

What is the criterion for deciding whether to use a filtering operator in the spatial domain or in the frequency domain?
The filtering, in the frequency domain, is very selective, allowing you to remove, attenuate or amplify specific frequency components or frequency bands of the input signal by designing the mask $H(u, v)$ adequately. We can see that the *filtering* term is more appropriate by operating in the frequency domain, as well as the *convolution mask* term, that recalls the meaning of masking (filtering) some frequencies of the input signal.

In Fig. 9.46 the whole filtering process is highlighted: *convolution in the frequency domain*. The input image $f(i, j)$ (normally with integers, gray levels) is transformed by FFT in the frequency domain (complex numbers). A frequency mask is generated for the purpose of removing some of them. This is obtained, for example, by setting the frequencies to be eliminated in the mask to zero, while they are set to 1 at the frequencies that must be preserved.

The convolution process in the frequency domain easily eliminates a specific frequency component $u_1$, $v_1$ of the input image. In fact, for the (9.112) the response spectrum of the system $G(u, v)$ is realized by multiplying the spectrum of the input image $F(u, v)$ with the spectrum of the mask $H(u, v)$ which takes zero value in $(u_1, v_1)$, and therefore, this output frequency is canceled

$$G(u_1, v_1) = F(u_1, v_1) \cdot H(u_1, v_1) = F(u_1, v_1) \cdot 0 = 0$$

The output image $G(u, v)$ obtained in the frequency domain is later transformed back into the spatial domain through the IFFT (inverse FFT). The filtering in the frequency domain can be more advantageous than the spatial one, especially in the case of images with additive noise of a periodic (not random) nature easily describable in the frequency domain. In Fig. 9.46, high-intensity point areas represent in the frequency domain the spatial frequencies in the noise band. By means of an appropriate mask, it is possible to eliminate the frequencies represented by these point areas and applying the IFFT the source image is reconstructed with the attenuated noise.

It should be remembered that this filter also eliminates the structures of the image that correspond to the same frequency band eliminated and this often explains the nonperfect reconstruction of the source image. Figure 9.47 illustrates an application of the functional diagram represented in Fig. 9.46 with the aim of eliminating/attenuating the periodic noise present in an image operating in the frequency domain. From the spectrum analysis, it is possible to design an ad hoc filtering mask to selectively cut the frequencies associated with the noise. We will return on the topic when we discuss more generally the argument of restoration of images using the Fourier approach and in particular notch filters.

Basically, the spatial convolution is computationally expensive especially if the convolution mask is large. In the frequency domain, the use of the FFT and IFFT transforms considerably reduces the computational load and is also suitable for real-time applications using specialized hardware, in particular, in the processing of signals and images. In the artificial vision different processes are not traceable to linear spatially invariant (LSI) operators. This limits very much the use of FFT and IFFT. When an image elaboration process can be modeled or approximated to a LSI system, the filtering masks are normally small $(3 \times 3)$ or $(5 \times 5)$ and consequently, the spatial convolution is convenient as an alternative to using the FFT. This is the reason for the widespread use of linear filters. Finally, we remember the limitations of the use of FFT for rounding errors due to numerical computation, stability of the inverse filter, multiplicity of solutions obtained, and limitations of the domain of gray values.

After defining the theory and implementative aspects of a linear operator, based essentially on the convolution process in spatial domain (9.112) and in the frequency domain (9.113), we are now able to illustrate in the following paragraphs two categories of *local operators*: *smoothing* (leveling gray levels or color) and *edging* (extraction or enhancement of contours).