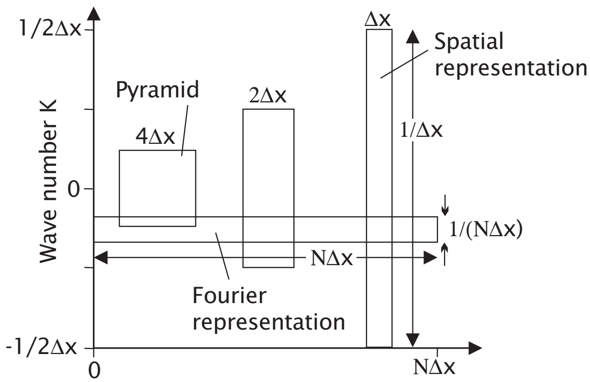

Figure 8.12: *Illustration of the interdependence of resolution in the spatial and wave-number domain in one dimension. Representations in the space domain, the wave-number domain, and the space/wave-number domain (2 planes of pyramid with half and quarter resolution) are shown.*

8.8 Scale of signals

8.8.1 Basics

In Sections 8.5 and 8.7 the representation of images in the spatial and wave-number domain were discussed. If an image is represented in the *spatial domain*, we do not have any information at all about the wave numbers contained at a point in the image. We know the position with an accuracy of the lattice constant Δx, but the local wave number at this position may be anywhere in the range of the possible wave numbers from $-1/(2\Delta x)$ to $1/(2\Delta x)$ (Fig. 8.12).

In the *wave-number domain* we have the reverse case. Each pixel in this domain represents one wave number with the highest wave-number resolution possible for the given image size, which is $-1/(N\Delta x)$ for an image with N pixels in each coordinate. But any positional information is lost, as one point in the wave-number space represents a periodic structure that is spread over the whole image (Fig. 8.12). Thus, the position uncertainty is the linear dimension of the image $N\Delta x$. In this section we will revisit both representations under the perspective of how to generate a multiscale representation of an image.

The foregoing discussion shows that the representations of an image in either the spatial or wave-number domain constitute two opposite extremes. Although the understanding of both domains is essential for any type of signal processing, the representation in either of these domains is inadequate to analyze objects in images.

In the wave-number representation the spatial structures from various independent objects are mixed up because the extracted periodic structures cover the whole image. In the spatial representation we have no information about the spatial structures contained in an object, we just know the local pixel gray values.

What we thus really need is a type of joint representation that allows for a separation into different wave-number ranges (scales) but still preserves as much spatial resolution as possible. Such a representation is called a multiscale or *multiresolution representation*.

The limits of the joint spatial/wave-number resolution are given by the *uncertainty relation* discussed in Section 8.6.3. It states that the product of the resolutions in the spatial and wave-number domain cannot be beyond a certain threshold. This is exactly what we observed already in the spatial and wave-number domains. However, besides these two domains any other combination of resolutions that meets the uncertainty relation can be chosen. Thus the resolution in wave numbers, that is, the distinction of various scales in an image, can be set to any value with a corresponding spatial resolution (Fig. 8.12). As the uncertainty relation gives only the lower limit of the joint resolution, it is important to devise efficient data structures that approach this limit.

In the last two decades a number of various concepts have been developed for multiresolution signal processing. Some trace back to the early roots of signal processing. This includes various techniques to filter signals for certain scale ranges such as the windowed Fourier transform, Gabor filters, polar separable quadrature filters, and filters steerable in scale (Section 8.8).

Some of these techniques are directly suitable to compute a *local wave number* that reflects the dominant scale in a local neighborhood. Multigrid image structures in the form of pyramids are another early and efficient multiresolution [12]. More recent developments are the *scale space* (Section 8.8) and *wavelets* [13, 14].

Although all of these techniques seem to be quite different at first glance, this it not the case. They have much in common; they merely look at the question of multiresolutional signal representation from a different point of view. Thus an important issue in this chapter is to work out the relations between the various approaches.

An early account on multiresolution imaging was given by Rosenfeld [15]. The standard work on linear scale space theory is by Lindeberg [16] (see also CVA2 [Chapter 11]), and nonlinear scale space theory is treated by Weickert [17] (see also Chapter 12).

8.8.2 Windowed Fourier transform

One way to a multiresolutional signal representation starts with the Fourier transform. If the Fourier transform is applied only to a section of the image and this section is moved around through the whole image, then a joint spatial/wave-number resolution is achieved. The spatial resolution is given by the size of the window and due to the uncertainty relation (Section 8.6.3), the wave-number resolution is reduced by the ratio of the image size to the window size. The window function $w(\boldsymbol{x})$ must not be a box function. Generally, a useful window function has a maximum at the origin, is even and isotropic, and decreases monotonically with increasing distance from the origin. This approach to a joint space/wave-number representation is the *windowed Fourier transform*. It is defined by

$$\hat{g}(\boldsymbol{x}, \boldsymbol{k}_0) = \int_{-\infty}^{\infty} g(\boldsymbol{x}')w(\boldsymbol{x}' - \boldsymbol{x}) \exp(-2\pi \mathrm{i}\boldsymbol{k}_0\boldsymbol{x}')\, \mathrm{d}x'^2 \tag{8.91}$$

The integral in Eq. (8.91) looks almost like a convolution integral (Section 8.6.3). To convert it into a convolution integral we make use of the fact that the window function is even ($w(-\boldsymbol{k}) = w(\boldsymbol{k})$) and rearrange the second part of Eq. (8.91):

$$\begin{aligned} & w(\boldsymbol{x}' - \boldsymbol{x}) \exp(-2\pi \mathrm{i}\boldsymbol{k}_0\boldsymbol{x}') = \\ & w(\boldsymbol{x} - \boldsymbol{x}') \exp(2\pi \mathrm{i}\boldsymbol{k}_0(\boldsymbol{x} - \boldsymbol{x}')) \exp(-2\pi \mathrm{i}\boldsymbol{k}_0\boldsymbol{x}) \end{aligned}$$

Then we can write Eq. (8.91) as a convolution:

$$\hat{g}(\boldsymbol{x}, \boldsymbol{k}_0) = [g(\boldsymbol{x}) * w(\boldsymbol{x}) \exp(2\pi \mathrm{i}\boldsymbol{k}_0\boldsymbol{x})]\ \exp(-2\pi \mathrm{i}\boldsymbol{k}_0\boldsymbol{x}) \tag{8.92}$$

This means that the local Fourier transform corresponds to a convolution with the complex convolution kernel $w(\boldsymbol{x}) \exp(2\pi \mathrm{i}\boldsymbol{k}_0\boldsymbol{x})$ except for a phase factor $\exp(-2\pi \mathrm{i}\boldsymbol{k}_0\boldsymbol{x})$. Using the *shift theorem* (Table 8.4), the transfer function of the convolution kernel can be computed to be

$$w(\boldsymbol{x}) \exp(2\pi \mathrm{i}\boldsymbol{k}_0\boldsymbol{x}) \Longleftrightarrow \hat{w}(\boldsymbol{k} - \boldsymbol{k}_0) \tag{8.93}$$

This means that the convolution kernel is a *bandpass filter* with a peak wave number of $\boldsymbol{k}_0$. The width of the bandpass is inversely proportional to the width of the window function. In this way, the spatial and wave-number resolutions are interrelated to each other. As an example, we take a Gaussian window function

$$w(\boldsymbol{x}) = \frac{1}{\sigma^D} \exp\left(-\pi \frac{|\boldsymbol{x}|^2}{\sigma^2}\right) \Longleftrightarrow \hat{w}(\boldsymbol{k}) = \exp\left(-\pi \frac{|\boldsymbol{k}|^2}{\sigma^{-2}}\right) \tag{8.94}$$

The Gaussian window function reaches the theoretical limit set by the uncertainty relation and is thus an optimal choice; a better wave-number resolution cannot be achieved with a given spatial resolution.

The windowed Fourier transform Equation (8.91) delivers a complex filter response. This has the advantage that both the phase and the amplitude of a bandpass-filtered signal are retrieved.

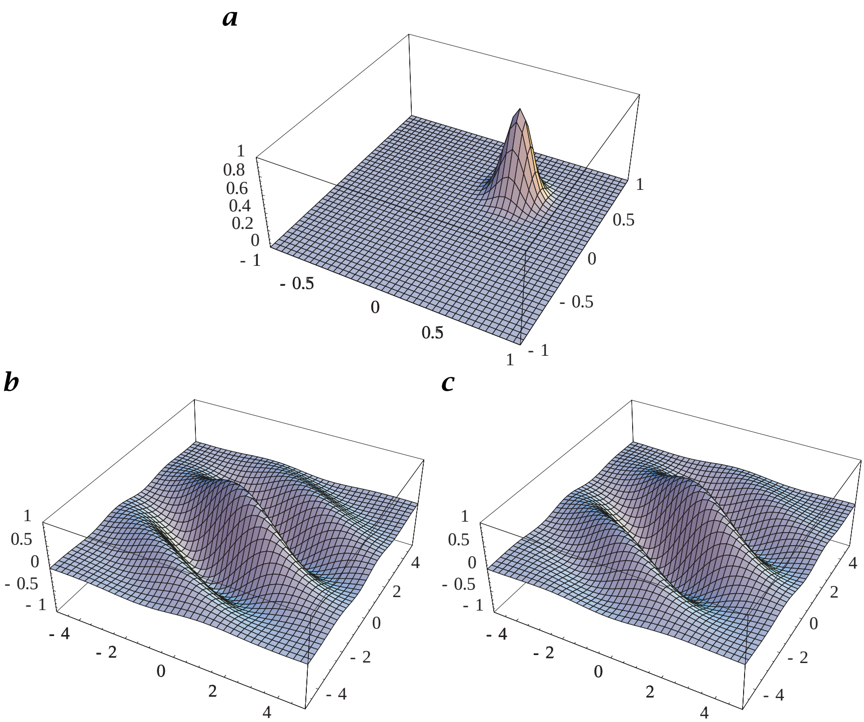

Figure 8.13: ***a*** *Transfer function (Eq. (8.95));* ***b*** *even; and* ***c*** *odd part of the filter mask (Eq. (8.97)) of a Gabor filter.*

8.8.3 Gabor filter

Definition. A *Gabor filter* is a bandpass filter that selects a certain wavelength range around the center wavelength $\boldsymbol{k}_0$ using the Gaussian function. The Gabor filter is very similar to the windowed Fourier transform if the latter is used with a Gaussian window function. The transfer function of the Gabor filter is real but asymmetric and defined as

$$\hat{G}(\boldsymbol{k}) = \exp\left(-\pi|\boldsymbol{k} - \boldsymbol{k}_0)|^2\sigma_x^2\right) \tag{8.95}$$

From this equation it is obvious that a Gabor filter is only a useful bandpass filter if it does not include the origin, that is, it is $\hat{G}(\mathbf{0}) = 0$. This condition is met in good approximation if $|\boldsymbol{k}_0|\sigma_x > 3$.

The filter mask (point spread function) of these filters can be computed easily with the shift theorem (Table 8.4):

$$G(\boldsymbol{x}) = \frac{1}{\sigma^D}\exp(2\pi \mathrm{i}\boldsymbol{k}_0\boldsymbol{x})\exp\left(-\frac{\pi|\boldsymbol{x}|^2}{\sigma_x^2}\right) \tag{8.96}$$

The complex filter mask can be split into an even real and an odd imaginary part:

$$
\begin{aligned}
G_+(\boldsymbol{x}) &= \frac{1}{\sigma^D}\cos(\boldsymbol{k}_0\boldsymbol{x})\exp\left(-\frac{\pi|\boldsymbol{x}|^2}{\sigma_x^2}\right) \\
G_-(\boldsymbol{x}) &= \frac{1}{\sigma^D}\sin(\boldsymbol{k}_0\boldsymbol{x})\exp\left(-\frac{\pi|\boldsymbol{x}|^2}{\sigma_x^2}\right)
\end{aligned}
\quad (8.97)
$$

Quadrature filters and analytic signals. Gabor filters are examples of quadrature filters. This general class of filters generates a special type of signal known as the *analytic signal* from a real-valued signal.

It is the easiest way to introduce the quadrature filter with the complex form of its transfer function. Essentially, the transfer function of a D-dimensional quadrature filter is zero for one half-space of the Fourier domain parted by the hyperplane $\boldsymbol{k}^T\bar{\boldsymbol{n}} = 0$:

$$
\hat{q}(\boldsymbol{k}) = \begin{cases} 2h(\boldsymbol{k}) & \boldsymbol{k}^T\bar{\boldsymbol{n}} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (8.98)
$$

where $h(\boldsymbol{k})$ is a real-valued function. Equation (8.98) can be separated into an even and odd function:

$$
\begin{aligned}
\hat{q}_+(\boldsymbol{k}) &= (\hat{q}(\boldsymbol{k}) + \hat{q}(-\boldsymbol{k}))/2 \\
\hat{q}_-(\boldsymbol{k}) &= (\hat{q}(\boldsymbol{k}) - \hat{q}(-\boldsymbol{k}))/2
\end{aligned}
\quad (8.99)
$$

The relation between the even and odd part of the signal response can be described by the *Hilbert transform*:

$$
\hat{q}_-(\boldsymbol{k}) = \mathrm{i}\,\mathrm{sgn}(\boldsymbol{k}^T\bar{\boldsymbol{n}})\hat{q}_+(\boldsymbol{k}) \Longleftrightarrow q_-(\boldsymbol{x}) = \frac{\mathrm{i}}{\pi}\int_{-\infty}^{\infty}\frac{q_+(\boldsymbol{x}')}{(\boldsymbol{x}'-\boldsymbol{x})^T\bar{\boldsymbol{n}}}\,\mathrm{d}^D\boldsymbol{x}' \quad (8.100)
$$

The even and odd part of a quadrature filter can be combined into a complex-valued signal by

$$
q_A = q_+ - \mathrm{i}q_- \quad (8.101)
$$

From Eq. (8.100) we can then see that this combination is consistent with the definition of the transfer function of the quadrature filter in Eq. (8.98).

The basic characteristic of the analytic filter is that its even and odd part have the *same* magnitude of the transfer function but that one is even and real and the other is odd and imaginary. Thus the filter responses of the even and odd part are shifted in phase by 90°. Thus the even part is cosine-like and the odd part is sine-like—as can be seen

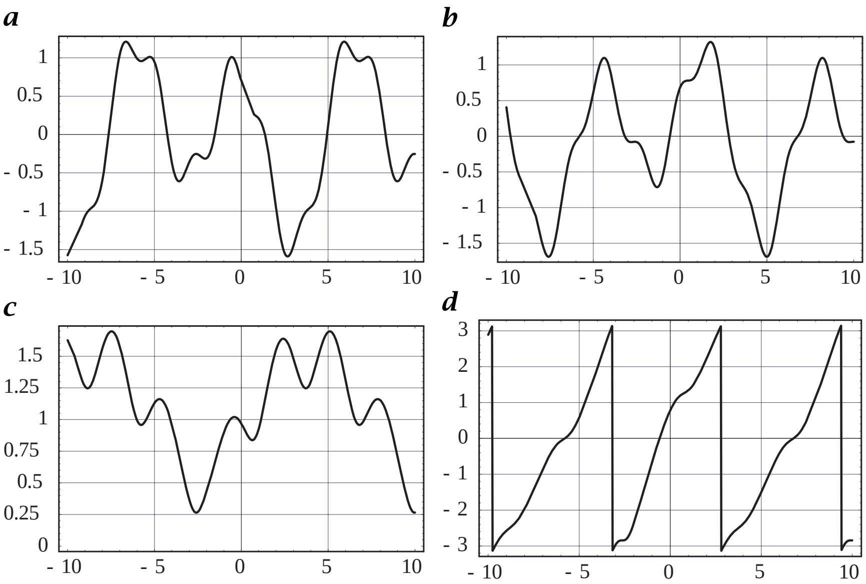

Figure 8.14: *Representation of a filtered 1-D signal as an analytic signal: Signal filtered with* ***a*** *the even and* ***b*** *the odd part of a quadrature filter;* ***c*** *amplitude; and* ***d*** *phase signal.*

from the Gabor filter (Fig. 8.13b and c)—and they are shifted in phase by 90° (Fig. 8.14).

Although the transfer function of the analytic filter is real, it results in a complex signal because it is asymmetric. For a real signal no information is lost by suppressing the negative wave numbers. They can be reconstructed as the Fourier transform of a real signal is Hermitian (Section 8.6.3).

The analytic signal can be regarded as just another representation of a real signal with two important properties. The magnitude of the analytic signal gives the *local amplitude* (Fig. 8.14c)

$$|q_A|^2 = q_+^2 + q_-^2 \tag{8.102}$$

and the argument the *local phase* (Fig. 8.14d)

$$\arg(\mathcal{A}) = \arctan\left(\frac{-\mathcal{H}}{\mathcal{I}}\right) \tag{8.103}$$

While the concept of the analytic signal works with any type of 1-D signal, it must be used with much more care in higher-dimensional signals. These problems are related to the fact that an analytical signal cannot be defined for all wave numbers that lie on the hyperplane defined by $\boldsymbol{k}^T\bar{\boldsymbol{n}} = 0$ partitioning the Fourier domain in two half-spaces.

For these wave numbers the odd part of the quadrature filter is zero. Thus it is not possible to compute the local amplitude nor the local phase of the signal. This problem can only be avoided if the transfer function of the quadrature filter is zero at the hyperplane. For a phase definition in two dimensions that does not show these restrictions, see CVA3 [Chapter 10].

8.8.4 Local wave number

The key to determining the local wave number is the *phase* of the signal. As an introduction we discuss a simple example and consider the 1-D periodic signal $g(x) = g_0 \cos(kx)$. The argument of the cosine function is known as the phase $\phi(x) = kx$ of the periodic signal. This is a linear function of the position and the wave number. Thus, we obtain the wave number of the periodic signal by computing the first-order spatial derivative of the phase signal

$$\frac{\partial \phi(x)}{\partial x} = k \tag{8.104}$$

These simple considerations emphasize the significant role of the phase in signal processing.

Local wave number from phase gradients. In order to determine the local wave number, we need to compute just the first spatial derivative of the phase signal. This derivative has to be applied in the same direction as the Hilbert or quadrature filter. The phase is given by

$$\phi(\boldsymbol{x}) = \arctan\left(\frac{-g_+(\boldsymbol{x})}{g_-(\boldsymbol{x})}\right) \tag{8.105}$$

Direct computation of the partial derivatives from Eq. (8.105) is not advisable, however, because of the inherent discontinuities in the phase signal. A phase computed with the inverse tangent restricts the phase to the main interval $[-\pi, \pi[$ and thus inevitably leads to a wrapping of the phase signal from π to $-\pi$ with the corresponding discontinuities.

As pointed out by Fleet [18], this problem can be avoided by computing the phase gradient directly from the gradients of $q_+(\boldsymbol{x})$ and $q_-(\boldsymbol{x})$:

$$\begin{aligned} k_p &= \frac{\partial \phi(\boldsymbol{x})}{\partial x_p} \\ &= \frac{\partial}{\partial x_p} \arctan(-q_+(\boldsymbol{x})/q_-(\boldsymbol{x})) \\ &= \frac{1}{q_+^2(\boldsymbol{x}) + q_-^2(\boldsymbol{x})} \left(\frac{\partial q_+(\boldsymbol{x})}{\partial x_p} q_-(\boldsymbol{x}) - \frac{\partial q_-(\boldsymbol{x})}{\partial x_p} q_+(\boldsymbol{x})\right) \end{aligned} \tag{8.106}$$

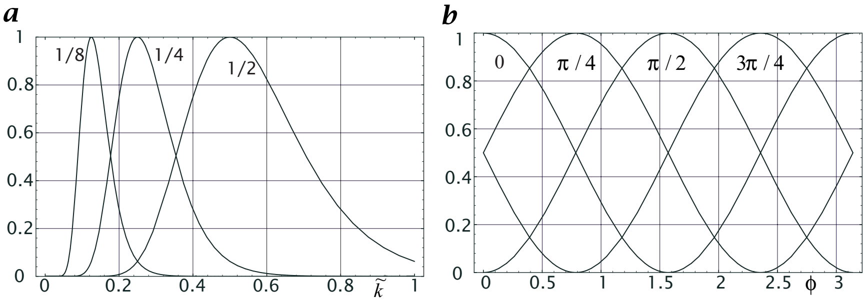

Figure 8.15: ***a*** *Radial and* ***b*** *angular part of quadrature filter according to Eq. (8.107) with $l = 2$ and $B = 2$ in different directions and with different peak wave numbers.*

This formulation of the phase gradient also eliminates the need for using trigonometric functions to compute the phase signal and is, therefore, significantly faster.

Local wave number from filter ratios. With *polar separable* quadrature filters ($\hat{r}(k)\hat{d}(\phi)$) as introduced by Knutsson [19] another scheme for computation of the local scale is possible. These classes of filters are defined by

$$
\begin{aligned}
\hat{r}(k) &= \exp\left[-\frac{(\ln k - \ln k_0)^2}{(B/2)^2 \ln 2}\right] \\
\hat{d}(\phi) &= \begin{cases} \cos^{2l}(\phi - \phi_k) & |\phi - \phi_k| < \pi/2 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
\tag{8.107}
$$

In this equation, the complex notation for quadrature filters is used as introduced at the beginning of this section. The filter is directed into the angle ϕ_k.

The filter is continuous, as the cosine function is zero in the partition plane for the two half-spaces ($|\phi - \phi_k| = \pi/2$). The constant k_0 denotes the peak wave number. The constant B determines the half-width of the wave number in number of octaves and l the angular resolution of the filter. In a logarithmic wave-number scale, the filter has the shape of a Gaussian function. Therefore, the radial part has a *lognormal* shape. Figure 8.15 shows the radial and angular part of the transfer function.

The *lognormal* form of the radial part of the quadrature filter sets is the key for a direct estimate of the *local wave number* of a narrowband signal. According to Eq. (8.107), we can write the radial part of the transfer function as

$$
\hat{r}_l(k) = \exp\left[-\frac{(\ln k - \ln k_l)^2}{2\sigma^2 \ln 2}\right]
\tag{8.108}
$$

We examine the ratio of the output of two different radial center frequencies k_1 and k_2 and obtain:

$$\begin{aligned}
\frac{\hat{r}_2}{\hat{r}_1} &= \exp\left[-\frac{(\ln k - \ln k_2)^2 - (\ln k - \ln k_1)^2}{2\sigma^2 \ln 2}\right] \\
&= \exp\left[\frac{2(\ln k_2 - \ln k_1)\ln k + \ln^2 k_2 - \ln^2 k_1}{2\sigma^2 \ln 2}\right] \\
&= \exp\left[\frac{(\ln k_2 - \ln k_1)[\ln k - 1/2(\ln k_2 + \ln k_1)]}{\sigma^2 \ln 2}\right] \\
&= \exp\left[\frac{\ln(k/\sqrt{k_2 k_1})\ln(k_2/k_1)}{\sigma^2 \ln 2}\right] \\
&= \left(\frac{k}{\sqrt{k_1 k_2}}\right)^{\ln(k_2/k_1)/(\sigma^2 \ln 2)}
\end{aligned}$$

Generally, the ratio of two different radial filters is directly related to the local wave number. The relation becomes particularly simple if the exponent in the last expression is one. This is the case, for example, if the wave-number ratio of the two filters is two ($k_2/k_1 = 2$ and $\sigma = 1$). Then

$$\frac{\hat{r}_2}{\hat{r}_1} = \frac{k}{\sqrt{k_1 k_2}} \tag{8.109}$$

8.9 Scale space and diffusion

As we have seen with the example of the windowed Fourier transform in the previous section, the introduction of a characteristic *scale* adds a new coordinate to the representation of image data. Besides the spatial resolution, we have a new parameter that characterizes the current resolution level of the image data. The scale parameter is denoted by ξ. A data structure that consists of a sequence of images with different resolutions is known as a *scale space*; we write $g(\boldsymbol{x}, \xi)$ to indicate the scale space of the image $g(\boldsymbol{x})$. Such a sequence of images can be generated by repeated convolution with an appropriate smoothing filter kernel.

This section is considered a brief introduction into scale spaces. For an authoritative monograph on scale spaces, see Lindeberg [16].

8.9.1 General properties of a scale space

In this section, we discuss some general conditions that must be met by a filter kernel generating a scale space. We will discuss two basic requirements. First, new details must not be added with increasing scale parameter. From the perspective of information theory, we may