Chapter 8
Noise Reduction

Oleg Tischenko and Christoph Hoeschen

In this chapter, we will describe the theoretical background of noise reduction in medical imaging as well as give some examples of noise reduction methods. To do so, we start with a fundamental description of digital image generation in medical imaging, since we will only focus on digital images and noise reduction by means of digital image processing. In the next part, we will discuss the corresponding processing in general before we describe the approaches typically used mainly based on linear filtering and some new approaches based on nonlinear approaches.

8.1 Idealized Model of Digital Image Generation

The simplified model of an imaging system can be characterized by the image space $\mathbf{D}$, which can be assumed to be two dimensional without a loss of generality, and impulse response φ. Let $\mathbf{D} \subset \mathbb{R}^2$ be squared area of the size $D \times D$, and let φ be a continuous function of fast decay whose Fourier transform $\hat{\varphi}(\omega_x, \omega_y)$ is supported inside the circle $\omega_x^2 + \omega_y^2 = \Omega_0^2$, i.e., $\hat{\varphi}(\omega_1, \omega_2) = 0$ if $\omega_1^2 + \omega_2^2 > \Omega_0^2$. The output $F : \mathbf{D} \to \mathbb{R}$ of the imaging system is modeled as a convolution

$$F = f * \varphi, \tag{8.1}$$

where $f : \mathbf{D} \to \mathbb{R}$ is the input signal. It is supposed that f is integrable on $\mathbf{D}$. Then F can be represented by the Fourier series

O. Tischenko (✉) • C. Hoeschen
Research Unit Medical Radiation Physics and Diagnostics, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany
e-mail: oleg.tischenko@helmholtz-muenchen.de

$$F(x, y) = \sum_{m,k \in \mathbb{Z}} \hat{F}_{m,k} e^{i\delta(xm+yk)}, \tag{8.2}$$

where $\hat{F}_{m,k} = D^{-2}\hat{f}(\delta m, \delta k)\hat{\varphi}(\delta m, \delta k)$ and

$$\delta = 2\pi/D. \tag{8.3}$$

Since the impulse response φ is band limited with the bandwidth Ω_0, the series (8.2) can be truncated:

$$F(x, y) = \sum_{-N \leq m,k \leq N} \hat{F}_{m,k} e^{i(xm+yk)\delta} \tag{8.4}$$

with

$$N = \left\lceil \frac{D}{\Delta} \right\rceil \tag{8.5}$$

and

$$\Delta = \frac{2\pi}{\Omega_0} \tag{8.6}$$

That is, the output F is a trigonometric polynomial and, consequently, can be uniquely recovered from its samples (see, e.g., [1]) measured over the rectangular grid $\{(x_m, y_k) | x_m = m\Delta, y_k = k\Delta\}_{-N \leq m,k \leq N}$. Hence, under certain idealizing assumptions, the continuous image can be identified with its samples.

We refer to the matrix $\{F_{m,k}\}_{-N \leq m,k \leq N}$ as a digital image of the input signal f. In the following, we call the number N the *resolution limit* of the imaging system, meaning that $N\delta$ is the maximal resolution that can be achieved with this system. Clearly, for the case considered here, the value $N\delta$ coincides with the Nyquist frequency of the imaging system.[1]

Finally, we point out that extending the matrix $\{F_{j,n}\}$ periodically with the period $M \times M$, $M = 2N + 1$, we can write

$$\hat{F}_{l,m} = \frac{1}{M^2} \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} F_{j,k} e^{-2\pi i(jl+km)/M}, \tag{8.7}$$

which is the standard notation for the discrete 2D Fourier transform.

---

[1] It is not always the case in real systems.

8.2 Image Processing

In practice, we do not have to deal with the image F but with its corrupted version I

$$I = F + \eta, \tag{8.8}$$

where η is the random value referred to as system noise. For an X-ray imaging system, another substantial noise component, which is due to the Poisson statistics and scattering of the X-ray quanta, is referred to as quantum noise (see, e.g., Chap. 9 of [2]). In view of the fact that the quantum noise is an inherit feature of the input signal f, the representation (8.8) can be replaced by

$$I = (g + q) * \varphi + \eta, \tag{8.9}$$

where q is the input quantum noise and g is the *desired* signal. One of the most important problems of the image processing is to recover g from the measured image I. It is clear, that this can be done to some degree of accuracy only. The best possible approximation of g in terms of the mean squared error is given by

$$Ag(x, y) = \sum_{-N \leq k,l \leq N} \hat{g}_{k,l} e^{i(xk+yl)\delta}, \tag{8.10}$$

where, at the right, there is the Fourier series of g truncated up to the resolution limit of the given imaging system. Leaving aside details, we mention that the linear problem of recovering unknowns $\hat{g}_{k,l}$ from I is sometimes referred to as the Wiener filtering.[2]

As it follows from (8.9), the problem of recovering Ag from I can be looked at as the one consisting of two distinct tasks: the suppression of the noisy component $q * \varphi + \eta$ (*noise reduction*) and the deconvolution of blurred image $g * \varphi$ (*de-blurring*). In the following, we use this discrimination in two tasks and describe some approaches used for the reduction of noise in the image.

8.3 Linear Filtering

The general strategy of this approach is to separate in the spectral description of the medical image the frequencies corresponding mainly to signals and those mainly corresponding to noise. Afterwards it is tried to suppress the power of those spectral components which are related to noise. This is done via appropriate weighting of

---

[2] Norbert Wiener had stated and solved this problem under special conditions for stationary stochastic time series (see [3]).

the frequency components of image I. The result of the weighting is the image wI, given by

$$wI(x, y) = D^{-2} \sum_{-N \leq l,m \leq N} \hat{w}_{l,m} \hat{I}_{l,m} \, e^{i(xl+ym)\delta}. \tag{8.11}$$

The set $\{\hat{w}_{l,m}\}_{-N \leq l, m \leq N}$ is called transfer matrix. The elements of transfer matrix are samples of the Fourier transform of the filter w, in terms of which the right-hand side of (8.11) can be written as the 2D convolution

$$wI = I * w. \tag{8.12}$$

In the following, we call function $w \in L_2(\mathbb{R}^2)$ *smoothing function* if it is continuous and

$$\hat{w}(0) = \int_{\mathbb{R}^2} w(x, y) dx dy \neq 0. \tag{8.13}$$

Due to (8.12) the kernel w acts upon the image I as a local weighted averaging, which is uniform for the whole image and is independent from its current space context. Linear filtering is perfect if the power spectrum of the image I can be separated in two disjoint regions, the one of which contains the spectral component of the useful signal and, the other, of the noise. We illustrate this apparent property in Fig. 8.1, where in the middle, there is an image corrupted with additive noise, the spectral power of which is concentrated outside the circle $\omega_x^2 + \omega_y^2 = \Omega^2$. On the contrary, the spectral power of the useful signal (left) is concentrated inside this circle. The right image is the convolution of the corrupted image with the ideal filter, i.e., the filter, the transfer function of which is

$$\hat{w}(\omega) = \begin{cases} 1, & |\omega| \leq \Omega, \\ 0, & \text{otherwise.} \end{cases} \tag{8.14}$$

Clearly, in this case, the original image and the denoised one are identical. In fact, if the power spectrum of the useful signal is concentrated inside a bounded region of the frequency domain and the signal-to-noise ratio is big enough in this region, then the linear filtering is still excellent independently from the spectral distribution of the noise. The corresponding example is illustrated in Fig. 8.2. On the left of this figure, there is the same image as in Fig. 8.1 with added white noise and, on the right, the result of the filtering with Gaussian filter g_Ω, the transfer function of which is

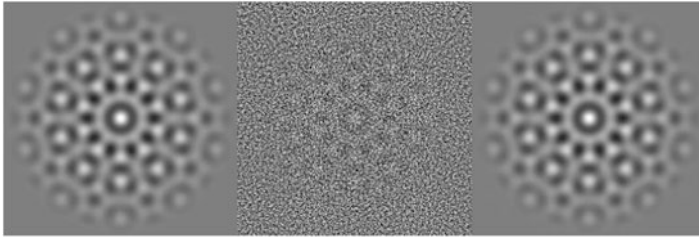$$\hat{g}_\Omega(\omega) = e^{-\frac{1}{2}\left(\frac{|\omega|}{\Omega}\right)^2}. \tag{8.15}$$

Fig. 8.1 Linear filtering of the band-limited image. *Left*: the image the power spectrum of which is inside the circle B. *Middle*: the same image with added noise, the power spectrum of which is outside the circle B. *Right*: convolution of the middle image with the ideal filter given by (8.14)

Fig. 8.2 Linear filtering of the band-limited image. *Left*: image from Fig. 8.1 with added white noise. *Right*: the convolution with Gaussian kernel
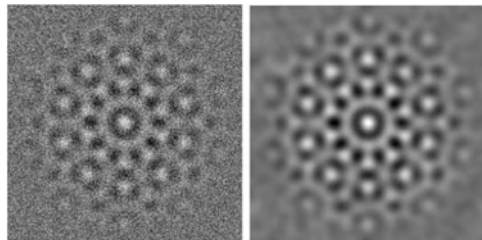
Thus, the following short conclusion can characterize the linear filtering: it works well if applied to regular signals, i.e., such signals the power spectrum of which is concentrated within a bounded region of low frequencies and it vanishes fast outside this region. The filtering operation in this case is reduced to the suppression of high-frequency components of the image. If the Fourier transform of an image decays slowly as $|\omega| \to \infty$, the linear filtering is not efficient anymore. Since the spectral power of such images is high also in the high-frequency sub-band, suppressing within this sub-band affects the sharpness of edges and destroys fine details of the image. This results in overall blurring of the image (see Fig. 8.3).

There is a variety of filters and corresponding smoothing kernels which are available. However, all rely on the same principle and show similar behavior.

8.4   Adaptive Nonlinear Filtering

In contrast to methods based on linear filtering, many advanced noise reduction techniques assume a control over the local amount of smoothing. Appropriately choosing a smoothing kernel at the current positions, it is possible to avoid unnecessary blurring in regions of high variation of the useful signal. Usually in practice the desired kernel is chosen from a predefined family of kernels. As an example we consider the family $\{g_\sigma, \sigma \geq 0\}$, where
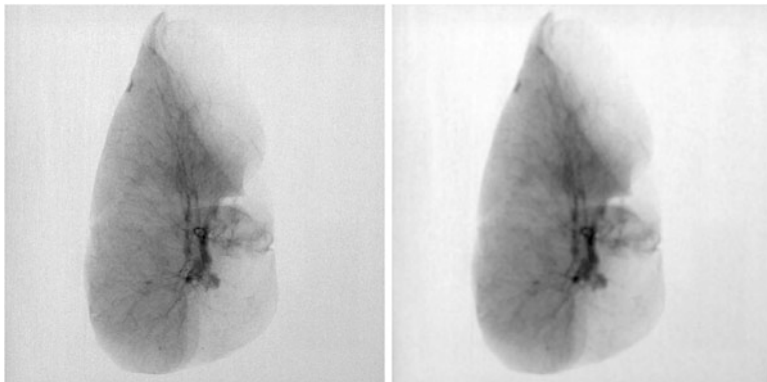
Fig. 8.3 Blurring effect of the linear filtering. *Left*: X-ray image of a human lung specimen. *Right*: convolution with a Gaussian kernel

$$g_\sigma(x,y) = \sigma^{-2}g\left\{\frac{x}{\sigma},\frac{y}{\sigma}\right\} \tag{8.16}$$

and the *generating kernel g* is a smoothing kernel [see definition above, (8.13)]. It can easily be seen that for $\sigma > 0$, $\hat{g}_\sigma(0) = 1$, that is, g_σ is also a smoothing kernel. Besides, $\lim_{\sigma \to 0} F * g_\sigma(x,y) = F(x,y)$, which implies that $\lim_{\sigma \to 0} g_\sigma = \delta$, where δ is a Dirac delta function. These properties allow to construct the image gF defined by

$$gF(x,y) = \int F(t,s)g_{\sigma(x,y)}(x-t,y-s)dtds, \tag{8.17}$$

where we determine σ depending on current position (x,y). The challenging task is to establish a rule according to which the value σ can be adapted to the current context of the image. Usually, σ is required to change as a function of some *decisive characteristic* such as the gradient, the Laplacian, the curvature, or some other local characteristic of the image. An example of such filtering with the Gaussian function

$$g(x,y) = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}} \tag{8.18}$$

as a generating kernel, and the squared deviation $\xi(x,y)$ as a local decisive characteristic, is given in Fig. 8.4. Particularly for this example, the current value of σ was set to

$$\sigma(x,y) = h\left(C;\frac{\xi_r(x,y)}{\xi_{\max}}\right), \tag{8.19}$$
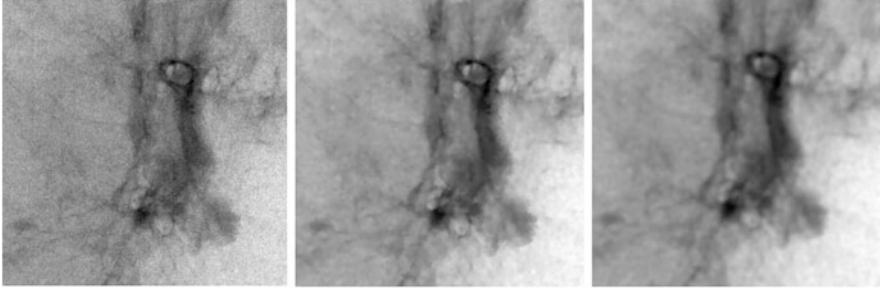
Fig. 8.4 Filtering with adaptive bandwidth Gaussian kernel. *Left*: a patch of the X-ray image shown in Fig. 8.3; *middle*: the result of the filtering by given tuning parameters r and $C = 5r$; *right*: linear filtering with Gaussian kernel by $\sigma = r$

where

$$h(C; t) = \frac{r}{1 + Ch_1(t)} \tag{8.20}$$

and the function $h_1 : [0, 1] \to [0, 1]$ is monotonically increasing from zero to one. For the local squared deviation, we have

$$\xi_r^2(x, y) = \frac{1}{2\pi r^2} \int_{B_r(x,y)} |F(t, s) - \bar{F}(x, y)|^2 dt ds, \tag{8.21}$$

$$\xi_{max} = \max_{(x,y) \in \mathbf{D}} \xi_r(x, y), \tag{8.22}$$

where the ball $B_r(x, y)$ is a ball of radius r located in (x, y); $\bar{F}(x, y)$ is the local mean value of F inside the ball.

The smoothing considered in this example incorporates two tuning parameters, which are the constants C and r. From (8.19) it follows that σ varies between r and $r/(C + 1)$, and the type of this variation depends on the type of the growth of the function h_1 defined in (8.20).

Another class of smoothing techniques is the so-called sigma filtering. Here we use smoothing kernels of the following general form:

$$\varphi_{I;x,y}(t, s) = C(x, y)w_\tau(I(t, s) - I(x, y))g_\sigma(x - t, y - s), \tag{8.23}$$

where (x, y) is the current *update position*; g_σ and w_τ are the smoothing kernels, the effective width of which are σ and τ, respectively; and the local constant $C(x, y)$ is determined from the condition

$$\int_{\mathbb{R}^2} \varphi_{I;x, y}(t, s) dt ds = 1. \tag{8.24}$$

Fig. 8.5 Sigma filtering. Three iterations of applying sigma filter to the patch of the lung phantom

Because of two kernels acting both in the space and the *intensity* domains, sigma filtering is sometimes referred to as *bilateral* filtering (see [4]). As it follows from (8.23), the sigma filter is represented by a family of kernels which depends on two scale parameters σ and τ. The parameter σ determines the radius of the ball $B(x, y)$ centered in (x, y). Using the parameter τ, it has to be decided which values $I(x - t, y - s)$ inside the ball $B(x, y)$ have to be weighted down. Normally, these are values which deviate from the value $I(x, y)$. In other words, the kernel w generates the local *averaging mask* that allocates those points within the ball $B(x, y)$ where image values are close to $I(x, y)$.

Bilateral filters are especially efficient if applied iteratively. Denoting the filtering operation with F, the iterative chain can formally be represented as $F_{\sigma_n \tau_n} F_{\sigma_{n-1} \tau_{n-1}} \ldots F_{\sigma_1 \tau_1}$.

In [5], it has been shown that the best result by an iterative sigma filter can be achieved if $\sigma_1 > \ldots > \sigma_{n-1} > \sigma_n$ and $\tau_1 < \ldots < \tau_{n-1} < \tau_n$. It can be shown that such a sequence of filtered images converges to the image which is piecewise constant. As an example, in Fig. 8.5, there are three iterations of bilateral filtering where w and g are both Gaussian.

8.5 Wavelet-Based Nonlinear Filtering

A quite different principle for spatially adaptive noise reduction is based on the reconstruction from selected wavelet coefficients. For the sake of completeness, before describing the related techniques, we will review main points of the theory of the wavelet transform referring mainly to [6] and [7]. In doing so, we will use the terminology of the theory of functions, and therefore, we start by stating formal notations accepted in this theory and used in the text.

C_m, $m \geq 1$, is a space whose elements are m times continuously differentiable functions; the space C_0 is the space of continuous functions, and C_{-1}, the space of piecewise continuous functions.

Let us have $j \in \mathbb{Z}$ and $m \in \mathbb{N}$. As *spline space* S_m^j, we will call the set of all such functions $f \in C_{m-2}$, the restriction of which to the interval $[2^j k, 2^j (k+1)]$ is the algebraic polynomial of order $m - 1$; the points $2^j k$ are called knots of the spline space S_m^j.

A space with an inner product $\langle \cdot, \cdot \rangle$ is called Hilbert space. In the following, the norm of element $a \in H$ is defined by

$$\|a\|_H = \sqrt{\langle a, a \rangle}. \tag{8.25}$$

$L_p(T)$ is a space of functions satisfying $\int_T |f(x)|^p < \infty$. The space $L_2(T)$ is a Hilbert space with the inner product

$$\langle a, b \rangle = \int_T a(x) b(x)^* \mathrm{d}x, \tag{8.26}$$

where b^* is a complex conjugate of b.

ℓ_p is a set of all infinite sequences $\{a_k\}_{k \in \mathbb{Z}}$ such that $\sum_k |a_k|^p < \infty$. The space ℓ_2 is a Hilbert space with the inner product

$$\langle a, b \rangle = \sum_k a_k b_k^*. \tag{8.27}$$

Span $\{g_n\}_n$ is the minimal space *spanned* on the family $\{g_n\}_n$, i.e., for any $a \in \ell_2$,

$$\sum_n a_n g_n = f \in \mathrm{span}\{g_n\}_n. \tag{8.28}$$

Operators that map a Hilbert space E into Hilbert space G will be denoted as $\mathbf{H} : E \to G$. With $\mathbf{H}^*$, we denote the adjoint of $\mathbf{H}$, i.e., such that $\langle \mathbf{H}/f, g \rangle_G = \langle f, \mathbf{H}^* g \rangle_E$ for any $f \in E$ and any $g \in G$. The operator is self-adjoint, if $\mathbf{H} = \mathbf{H}^*$.

A wavelet is a function $\psi \in L_2(\mathbb{R})$ that necessarily satisfies the condition

$$\int_{\mathbb{R}} \psi(x) \mathrm{d}x = \hat{\psi}(0) = 0. \tag{8.29}$$

Everywhere in the text, it will be supposed that ψ is real and normalized, i.e., $\|\psi\|_{L_2} = 1$. Wavelets are "famous" for being well localized both in the space domain and in the frequency domain. The localization properties of wavelets are usually expressed in terms of a so-called *space-frequency window*. This is a rectangle $\sigma \times \hat{\sigma}$ in the space-frequency plane with σ and $\hat{\sigma}$ defined by

$$\sigma^2 = \int_{\mathbb{R}} (x - \bar{x}) |\psi(x)|^2 \mathrm{d}x, \quad \bar{x} = \int_{\mathbb{R}} x |\psi(x)|^2 \mathrm{d}x, \tag{8.30}$$

$$\hat{\sigma}^2 = \pi^{-1} \int_0^\infty (\omega - \bar{\omega})^2 |\hat{\psi}(\omega)|^2 dx, \quad \bar{\omega} = \pi^{-1} \int_0^\infty \omega |\hat{\psi}(\omega)|^2 dx. \tag{8.31}$$

The point $(\bar{x}, \bar{\omega})$ is the center of the space-frequency window of the wavelet. Since ψ is real, the modulus of its Fourier transform is even, and this is why the integration in (8.31) is made over the positive half axis only.

Wavelet ψ is called the *mother wavelet* for the family $\{\psi_{s,u}\}_{s>0,u\in\mathbb{R}}$, where the function

$$\psi_{s,u}(x) = s^{-1/2}\psi\left(\frac{x-u}{s}\right) \tag{8.32}$$

is the dilated and translated version of the mother wavelet ψ. As an example, let us consider the family generated by so-called Mexican hat wavelet

$$\lambda(x) = \frac{2}{\sqrt{3}}\pi^{-1/4}(1 - x^2)e^{-\frac{x^2}{2}}. \tag{8.33}$$

In Fig. 8.6 wavelets $\lambda_{1,0}$ and $\lambda_{5,0}$ of this family (left) as well as their Fourier transforms (right) are shown.

Figure 8.7 shows space-frequency windows of these wavelets. One observes that the bigger is s, the wider is the effective width of the wavelet in the space domain and the better is its spectral resolution. In contrary, we obtain better space resolution by smaller s. Since $\hat{\sigma}_s = \hat{\sigma}/s$ and $\sigma_s = s\sigma$, the area of the space-frequency window does not change.

The bivariate function

$$wf(s, u) = \langle f, \psi_{s,u} \rangle, \quad s, u \in \mathbb{R} \tag{8.34}$$

is called a continuous wavelet transform of the function f. Since

$$\langle \hat{f}, \hat{\psi}_{s,u} \rangle = 2\pi \langle f, \psi_{s,u} \rangle, \tag{8.35}$$

the wavelet coefficient $\langle f, \psi_{s,u} \rangle$ by fixed s, u relates to the local contribution made by f within the space-frequency window of the wavelet $\psi_{s,u}$.

Besides the space-frequency window, another important characteristic of a wavelet is the number of vanishing moments. A wavelet ψ is said to have m vanishing moments if $\int x^k\psi(x)dx = 0$ for $0 \le k < m$. Since

$$\frac{d^n}{d\omega^n}\hat{\psi}(\omega) = \int (-ix)^n\psi(x)e^{-i\omega x}dx, \tag{8.36}$$

the number of vanishing moments of ψ is equal to the number of zeros of its Fourier transform at $\omega = 0$. Using this property, it can be shown that a compactly supported
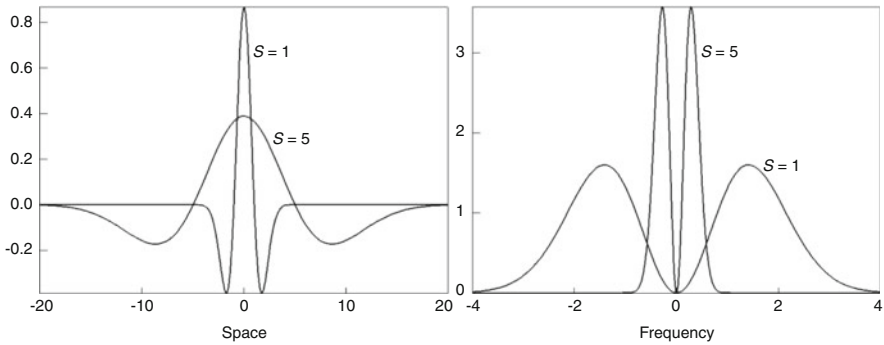
Fig. 8.6 Mexican hat wavelet. Wavelets $\lambda_{s,0}$ for $s = 1, 5$ (*left*) and their Fourier transforms (*right*)
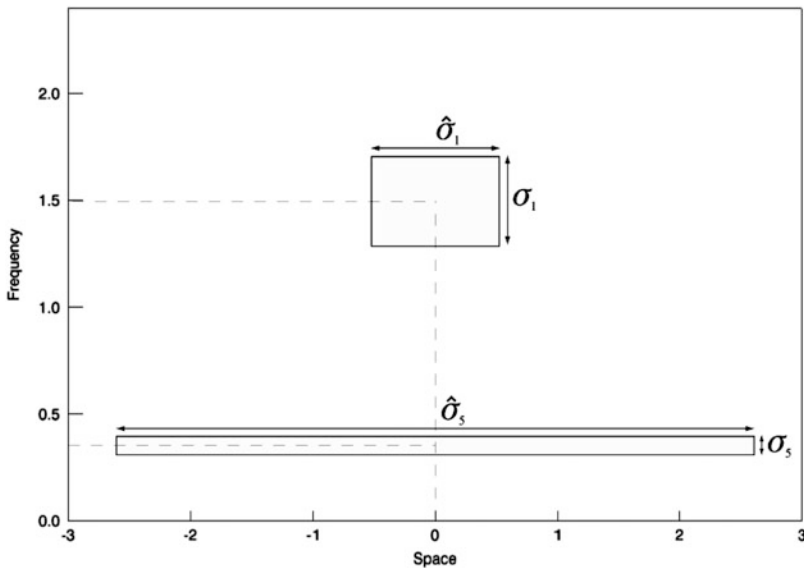

Fig. 8.7 Space-frequency windows of $\lambda_{s,0}$ for $s = 1, 5$

wavelet ψ has m vanishing moments if and only if there exists a compactly supported smoothing function θ such that

$$\psi(t) = (-1)^n \frac{d^n \theta(t)}{dt^n} \tag{8.37}$$

(see Chap. 6 of [7] for details); the corresponding wavelet transform of f is consequently

$$Wf(s, t) = s^{n+1/2} \frac{d^n}{dt^n} (f * \bar{\theta}_s), \tag{8.38}$$

where $\bar{\theta}_s(t) = \theta_s(-t)$ and $\theta_s(t) = s^{-1}\theta(t/s)$. Hence, the number of vanishing moments is crucial in those applications where the local regularity of the signal has to be measured.

The identity

$$f(x) = C_\psi^{-1} \int_0^\infty \int_{-\infty}^\infty \langle f, \psi_{s,u} \rangle \psi_{s,u}(x) \mathrm{d}u \frac{\mathrm{d}s}{s^2} \tag{8.39}$$

with

$$C_\psi = \int_0^\infty \frac{|\hat{\psi}(\omega)|^2}{\omega} \mathrm{d}\omega \tag{8.40}$$

is valid and makes sense for any $f \in L_2\ \mathbb{R}$ if $C_\psi < \infty$. Therefore, the wavelet ψ is called admissible if $C_\psi < \infty$. Condition (8.29) formulated above is necessary for ψ to be admissible. However, assuming that ψ is "reasonable," that is, ψ is such that it makes sense in the practice, this condition is also sufficient.

There are different possibilities to choose wavelets in 2D. For example, this can be a wavelet $\psi \in L_2\ (\mathbb{R}^2)$ that is radially symmetric. Since the wavelet in this case depends effectively on one variable only, the situation is essentially the same as in the 1D case: the 2D wavelet transform is the function that depends on two parameters, the shift (which is now 2D vector) and the dilation. The reconstruction formula is in this case

$$f(\mathbf{r}) \sim \int_0^\infty \int_{\mathbb{R}^2} \langle f, \psi_{s,\mathbf{u}} \rangle \psi_{s,\mathbf{u}}(\mathbf{r}) \mathrm{d}\mathbf{u} \frac{\mathrm{d}\mathbf{s}}{\mathbf{s}^3}. \tag{8.41}$$

It is also possible to choose a 2D wavelet that is not radially symmetric. Such wavelets are often called oriented wavelets. The wavelet transform in this case is a function that depends on three parameters, dilation, translation, and rotation, and the corresponding reconstruction formula is

$$f(\mathbf{r}) \sim \int_0^{2\pi} \int_{\mathbb{R}^2} \int_0^\infty \langle f, \psi_{s,\mathbf{u},\theta,} \rangle \psi_{s,\mathbf{u},\theta}(\mathbf{r}) \mathrm{d}\mathbf{u} \frac{\mathrm{d}\mathbf{s}}{\mathbf{s}^3} \mathrm{d}\theta. \tag{8.42}$$

In practice, one often applies separable 2D wavelets, i.e., a wavelet of the form $\varphi(t)\psi(s)$, where both φ and ψ are functions from $L_2(\mathbb{R})$, and at least the one of which is a wavelet.

The continuous wavelet transform provides an extremely redundant description of the function f. In fact, even a discrete subset of $\{Wf(s,u)\}_{s,u}$ can be used to recover functions exactly. A related discrete wavelet transform is $Wf_{m,n} = \langle f, \psi_{m,n} \rangle$ where

$$\psi_{m,n}(x) = a^{-m/2}\psi(a^{-m}x - nb) \tag{8.43}$$

for some fixed positive a, b.

For a wide range of choices for ψ, a and b discrete families of wavelets constitute *frames* of the functional space $L_2(\mathbb{R})$. The family $\{g_n\}_n$ is called a frame of the Hilbert space H if there exist constants $0 < A \le B < \infty$ such that

$$A\|f\|_H^2 \le \sum_n |\langle f, g_n \rangle|^2 \le B\|f\|_H^2 \tag{8.44}$$

for any $f \in H$. The equivalency $\langle f, g_n \rangle \equiv 0 \iff f = 0$ that follows from this definition implies that frames are complete in H, which means that any function in H can be approximated by linear combination of frame elements with any degree of accuracy. The frame is called redundant if its elements are linearly dependent.

In general, recovering a function from the frame coefficients is equivalent to the inversion of a self-adjoint operator $\mathbf{H} : H \to H$ that relates to the frame $\{g_n\}_n$ of H via

$$\mathbf{H}f = \sum_n \langle f, g_n \rangle g_n. \tag{8.45}$$

It is easy to see that the frame condition (8.44) is equivalent to the bounding condition

$$A\|f\|^2 \le |\langle \mathbf{H}f, f \rangle| \le B\|f\|^2 \tag{8.46}$$

that holds for any $f \in H$. This means that the inverse $\mathbf{H}^{-1}$ exists and is stable. Applying $\mathbf{H}^{-1}$ to both sides of (8.45) yields

$$f = \sum_n \langle f, g_n \rangle \tilde{g}_n \tag{8.47}$$

with

$$\tilde{g}_n = \mathbf{H}^{-1} g_n. \tag{8.48}$$

The family $\{\tilde{g}_n\}_n$ is called dual frame. Indeed, let us introduce the operator $\tilde{\mathbf{H}}$ that relates to $\{\tilde{g}_n\}_n$ via

$$\tilde{\mathbf{H}}f = \sum_n \langle f, \tilde{g}_n \rangle \tilde{g}_n. \tag{8.49}$$

One directly verifies that $\tilde{\mathbf{H}} = \mathbf{H}^{-1}$, and since $\mathbf{H}^{-1}$ is bounded by B^{-1} and A^{-1}, $\{\tilde{g}_n\}$ satisfies the frame bounding conditions with the bounds B^{-1} and A^{-1}. Moreover, the dual of $\{\tilde{g}_n\}_n$ is $\{g_n\}_n$, and therefore,

$$f = \sum_n \langle f, \tilde{g}_n \rangle g_n. \tag{8.50}$$

The inverse of $\mathbf{H}$ can be determined by means of the auxiliary operator $\mathbf{R} = \mathbf{I} - 2(A+B)^{-1}\mathbf{H}$ where $\mathbf{I}$ is the identity operator. The operator $\mathbf{R}$ is self-adjoint and satisfies the bounding condition

$$|\langle \mathbf{R}f, f \rangle| \leq \frac{B-A}{A+B} \|f\|_{L^2}^2 \leq \|f\|_{L^2}^2 \tag{8.51}$$

which means that $\|\mathbf{R}\| < 1$. Then,

$$\mathbf{H}^{-1} = 2(A+B)^{-1}(I-\mathbf{R})^{-1} = 2(A+B)^{-1}\sum_{k=0}^{\infty}\mathbf{R}^k. \tag{8.52}$$

In practice, one truncates the series (8.52) up to some N and uses the approximation

$$\tilde{g}_n \approx \tilde{g}_n^N = 2(A+B)^{-1}\sum_{k=0}^{N}\mathbf{R}^k g_n. \tag{8.53}$$

It is known (see Chap. 3 of [6]) that

$$\left\| f - \sum_n \langle f, g_n \rangle \tilde{g}_n^N \right\|_{L^2} \leq \left(\frac{B-A}{A+B}\right)^{N+1} \|f\|_{L^2}. \tag{8.54}$$

That is, the norm in (8.54) converges exponentially to zero at a rate depending on the value $(B-A)/(A+B)$. If $A \approx B$, one can truncate the series (8.52) up to zeroth term avoiding the computation of the dual frame and still have a high quality of reconstruction of arbitrary function.

If $A = B$, the frame is called *tight*. In this case, we can rearrange the identity

$$\sum_n |\langle f + g, g_n \rangle|^2 = A\|f + g\|_H^2 \tag{8.55}$$

to the identity

$$\left\langle f - A^{-1}\sum_n \langle f, g_n \rangle g_n, h \right\rangle = 0 \tag{8.56}$$

that holds for any $h \in H$ and any $f \in H$, and as a result, obtain the inversion formula

$$f = A^{-1} \sum_n \langle f, g_n \rangle g_n \tag{8.57}$$

that holds for any $f \in H$ at least in the weak sense.[3] The simplicity of (8.57) makes tight frames especially interesting in practice. Note that in spite of the similarity between (8.57) and a decomposition in an orthonormal basis, tight frames can be redundant. In the following, we show that a tight frame is either redundant or orthonormal. Indeed, setting $f = g_k$ in (8.56), we obtain that for any k, $\sum_n \alpha_n g_n = 0$ where

$$\alpha_n = \begin{cases} A^{-1} \langle g_k, g_k \rangle - 1, & n = k, \\ A^{-1} \langle g_n, g_k \rangle, & n \neq k.. \end{cases} \tag{8.58}$$

This means that if $\{g_n\}$ are linearly independent, then $\langle g_n, g_k \rangle = A\delta_{n,k}$, i.e., $\{g_n\}_n$ is orthonormal. On the other side, if $\|g\|_H^2 = A$ for all n, then

$$\sum_n |\langle g_k, g_n \rangle|^2 = A^2 + \sum_{n \neq k} |\langle g_k, g_n \rangle|^2 = A^2, \tag{8.59}$$

which is possible only if $\langle g_k, g_n \rangle = A\delta_{k,n}$. It follows that if for some n, $\|g_n\|_H^2 = C \neq A$, then the tight frame is redundant and $C \leq A$. For a tight frame, elements of which are of the same norm, the value A/C, where $\|g_n\|_H^2 = C$, can be interpreted as a measure of the redundancy of the frame.

Redundant tight frames can be used for noise reduction. Suppose that the signal f is measured over a redundant frame $\{g_n\}_n$, and let the measured coefficients f_n be contaminated with noise, i.e., $f_n = \langle f, g_n \rangle + q_n$, where q is a random variable. Let $\mathbf{U} : H \to \ell_2$ be the *frame operator*

$$(\mathbf{U}f)_n = \langle f, g_n \rangle. \tag{8.60}$$

Since the frame is redundant, there exists nontrivial $c \in \ell_2$, so that $\sum_n c_n g_n = 0$. In other words, the orthogonal complement of Im $\mathbf{U}$ in ℓ_2, where we denote the image of the operator $\mathbf{U}$ as Im $\mathbf{U}$, is not empty. Let $\mathbf{P} : \ell_2 \to \ell_2$ be the orthoprojector onto Im $\mathbf{U}$. Then, $\mathbf{P}(\mathbf{U}f + q) = \mathbf{U}f + \mathbf{P}q$. Decomposing $q = q_1 + q_2$, where $q_1 \in$ Im, and $q_2 \in (\text{Im } \mathbf{U})^\perp$, where $(\text{Im } \mathbf{U})^\perp$ is the orthogonal complement of Im $\mathbf{U}$ in ℓ_2, we obtain that $\mathbf{P}q = q_1$, which in turn implies that $\|\mathbf{P}q\|_{\ell_2} \leq \|q\|_{\ell_2}$. The following result is due to [7, Chap. 5]: let $\|g_n\|_{L_2}^2 = C$ for all n, and let q be a zero-mean white noise of variance σ_q^2. Then,

---

[3] The equality $f = g$ holds in the weak sense if $\langle f, h \rangle = \langle g, h \rangle$ for any $h \in H$.
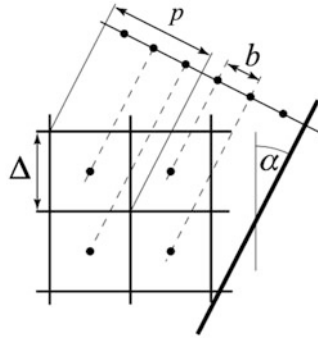
Fig. 8.8 Idealized setup for the determination of the MTF of a digital detector. Four-cell patch of the detector matrix and the edge line (*bold*) tilted at an angle α relative to the matrix. Centers of detector cells are projected onto the edge profile

$$\sigma_{\mathbf{P}q}^2 \leq \sigma_q^2 \frac{C}{A}. \tag{8.61}$$

This approach can be applied, e.g., in the edge-based technique used for the determination of the pre-sampling *modulation transfer function* (MTF) of digital detectors. The setup used within the framework of this technique is depicted in Fig. 8.8, left.

The samples of the edge profile can be thought of to be the coefficients measured over the set $\{\mathrm{Sinc}(\Omega(x - bn))\}_n$, which is the frame of the space of band-limited functions. Usually, the edge profile is highly oversampled. As a consequence, the frame is redundant, and the redundancy factor A/C is very high. Let E be the smallest subspace of ℓ_2 such that $\mathrm{Im}\,\mathbf{U} \subset E$. Then, due to (8.61), the orthogonal projection onto E is supposed to reduce the noisy fraction of the edge profile significantly.

The linear independent frames constitute a special class of families called *Riesz basis*: the family $\{g_n\}$ is called Riesz basis of H if $H = \overline{\mathrm{span}\{g_n\}}$ and there exist constants $A > 0$ and $B < \infty$ such that

$$A \sum_n |c_n|^2 \leq \left\| \sum_n c_n g_n \right\|^2 \leq B \sum_n |c_n|^2 \tag{8.62}$$

for any $c \in \ell_2$. Note that the equivalence $\sum_n c_n g_n = 0 \;\Leftrightarrow\; c = 0$ that follows from (8.62) implies that elements of the Riesz basis are linearly independent.

For the Riesz basis constituted by integer translates of the function $g \in L_2(\mathbb{R})$, the condition (8.62) can be written as

$$A \leq \sum_n |\hat{g}(\omega + 2\pi n)|^2 \leq B, \quad \omega \in \mathbb{R}. \tag{8.63}$$

Since

$$\langle g, g(\cdot - k)\rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{g}(\omega)|^2 e^{i\omega k} d\omega = \frac{1}{2\pi} \int_{0}^{2\pi} \sum_{n} |\hat{g}(\omega + 2\pi n)|^2 e^{i\omega k} d\omega, \quad (8.64)$$

one concludes that translates of g are orthonormal if and only if

$$\sum_{n} |\hat{g}(\omega + 2\pi n)|^2 \equiv 1. \tag{8.65}$$

In a similar way, it is verified that $\{g(\cdot - n)\}$ and $\{\tilde{g}(\cdot - n)\}$ are biorthogonal, i.e., $\langle g(\cdot - n), \tilde{g}(\cdot - k)\rangle = \delta_{n,k}$, if and only if

$$\sum_{n} \hat{g}(\omega + 2\pi n)\hat{\tilde{g}}(\omega + 2\pi n)^* \equiv 1. \tag{8.66}$$

Returning to wavelets, we accent that most known wavelet bases in $L_2(\mathbb{R})$ have been constructed by means of so-called *multi-resolution analysis* (MRA) formulated first in [8]. The key role within this approach is played by the so-called *scaling functions*: function $\varphi \in L_2(\mathbb{R})$ is called a scaling function if for any fixed $j \in \mathbb{Z}$ the family $\{\varphi_{j,n}\}_n$,

$$\varphi_{j,n}(x) = 2^{-j/2}(2^{-j}x - n), \tag{8.67}$$

is the Riesz basis of the subspace $V_j \subset L_2(\mathbb{R})$, and

(i)
$$V_j \subset V_{j-1}.$$

(ii)
$$f(\cdot) \in V_j \iff f(2\cdot) \in V_{j-1}.$$

(iii)
$$\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L_2(\mathbb{R}), \quad \bigcap_{j\in\mathbb{Z}} V_j = \{0\}.$$

Three necessary conditions for φ to be a scaling function are:

(1) The Fourier transform $\hat{\varphi}$ must satisfy (8.63).
(2) It satisfies the *two-scale relation*

$$\varphi = \sum_{n} h_n \varphi_{-1,n}. \tag{8.68}$$

(3) $$\hat{\varphi}(0) \neq 0.$$

In particular from (3), it follows that φ is a smoothing function. Therefore, V_j is called approximation, and the ladder $\ldots \subset V_{j+1} \subset V_j \subset V_{j-1} \subset \ldots$ as a multi-resolution approximation of $L_2(\mathbb{R})$.

Suppose that the Fourier transform of the scaling function φ satisfies (8.65), i.e., its integer translates are orthonormal. Let W_0 be the orthogonal compliment of V_0 in V_{-1}, that is, $V_{-1} = V_0 \oplus W_0$. Then there exists $\psi \in L_2(\mathbb{R})$, the integer translates of which constitute a basis of W_0. Moreover, ψ is necessarily a wavelet, the Fourier transform of which satisfies (8.65). In order to show this, let us rewrite the two-scale relation (8.68) in the Fourier domain:

$$\hat{\varphi}(\omega) = 2^{-1/2}\hat{h}(\omega/2)\hat{\varphi}(\omega/2), \tag{8.69}$$

where

$$\hat{h}(\omega) = \sum_n h_n \mathrm{e}^{-\mathrm{i}n\omega} \tag{8.70}$$

is 2π periodic. The orthonormality of $\varphi(\cdot - n)$ implies the condition

$$\left|\hat{h}(\omega)\right|^2 + \left|\hat{h}(\omega + \pi)\right|^2 \equiv 2, \tag{8.71}$$

which can be checked using (8.65). For any $f \in W_0$, the inclusion $f \in V_{-1}$ takes place, and consequently, there exists $a \in \ell_2$, so that $f = \sum_n a_n \varphi_{-1,n}$, or in the Fourier domain

$$\hat{f}(\omega) = 2^{-1/2}\hat{a}(\omega/2)\hat{\varphi}(\omega/2), \tag{8.72}$$

where $\hat{a}(\omega) = \sum_n a_n \mathrm{e}^{-\mathrm{i}\omega n}$ is 2π periodic. The condition $W_0 \perp V_0$ implies that $\langle f, \varphi(\cdot - k) \rangle = 0$ for any $k \in \mathbb{Z}$, which in turn leads to the identity

$$\hat{a}(\omega)\hat{h}(\omega)^* + \hat{a}(\omega + \pi)\hat{h}(\omega + \pi)^* \equiv 0. \tag{8.73}$$

The possible solution of (8.73) is

$$\hat{a}(\omega) = \hat{v}(2\omega)\hat{h}(\omega + \pi)^*, \tag{8.74}$$

where $\hat{v}$ is any 2π-periodic function. Substitution of (8.74) into (8.72) yields

$$\hat{f}(\omega) = 2^{-1/2}\hat{v}(\omega)\hat{h}(\omega/2 + \pi)^*\hat{\varphi}(\omega/2), \tag{8.75}$$

or equivalently

$$f(x) = \sum_n v_n \psi(x - n), \tag{8.76}$$

where ψ is the function, the Fourier transform of which is

$$\hat{\psi}(\omega) = 2^{-1/2} \hat{h}(\omega/2 + \pi)^* \hat{\varphi}(\omega/2). \tag{8.77}$$

Hence, we have shown that there exists ψ, translates of which constitute the basis in W_0. In order to see that ψ is a wavelet, note first that from (8.69), it follows that $\hat{h}(0) = \sqrt{2}$. Together with (8.71), this implies $\hat{h}(\pi) = 0$. Using this in (8.77) and taking into account that $\hat{\varphi}(0) \neq 0$ yields $\hat{\psi}(0) = 0$. That is, ψ is a wavelet (see condition (8.29)). The orthonormality of its translates can be checked by means of (8.65). Substituting (8.77) into (8.65), we obtain that

$$\sum_n |\hat{\psi}(\omega + 2\pi n)|^2 = 2^{-1} \sum_n |\hat{h}(\omega/2 + \pi n + \pi)^* \hat{\varphi}(\omega/2 + \pi n)|^2. \tag{8.78}$$

Representing the sum on the right-hand side of (8.78) by two sums over odd and even indices, using (8.71) and accounting for the 2π periodicity of $\hat{h}$, we obtain

$$\sum_k |\hat{\psi}(\omega + \pi n)|^2 \equiv 1, \tag{8.79}$$

that is, $\{\psi(\cdot - n)\}_n$ is an orthonormal basis of W_0. This basis is not unique. Any λ defined by

$$\hat{\lambda}(\omega) = \hat{\alpha}(\omega) \hat{\psi}(\omega), \tag{8.80}$$

where $\hat{\alpha}$ is 2π periodic and $|\hat{\alpha}(\omega)| \equiv 1$, is a wavelet whose integer translates constitute the orthonormal basis in W_0. Indeed, $\hat{\lambda}(0) = 0$; moreover,

$$\sum_n |\hat{\lambda}(\omega + 2\pi n)|^2 = |\hat{\alpha}(\omega)|^2 \sum_n |\hat{\psi}(\omega + 2\pi n)|^2 \equiv 1. \tag{8.81}$$

In fact, any two functions, integer translates of which constitute orthogonal bases in W_0, relate to each other via (8.80) (see Chap. 8 of [6]). Therefore,

$$\hat{\psi}(\omega) = 2^{-1/2} \hat{\alpha}(\omega) \hat{h}(\omega/2 + \pi)^* \hat{\varphi}(\omega/2), \tag{8.82}$$

where $\hat{\alpha}$ is 2π periodic and $|\hat{\alpha}(\omega)| \equiv 1$, characterizes all possible wavelets whose translates constitute the orthonormal basis in W_0. The choice of $\hat{\alpha}$ substantiates the wavelet. Usually, one sets $\hat{\alpha}(\omega) = e^{-i\omega/2}$. Then

$$\hat{\psi}(2\omega) = 2^{-1/2}\hat{g}(\omega)\hat{\varphi}(\omega), \tag{8.83}$$

where

$$\hat{g}(\omega) = \hat{h}(\omega + \pi)^* e^{-i\omega} \tag{8.84}$$

is the transfer function of coefficients of the decomposition of ψ in the basis of V_{-1}, that is,

$$\psi = \sum_n g_n \varphi_{-1,n}, \tag{8.85}$$

with

$$g_n = (-1)^n h_{1-n}. \tag{8.86}$$

From (i) and (ii) of MRA, it follows that the decomposition

$$V_{j-1} = V_j \oplus W_j \tag{8.87}$$

is valid for any j. Together with (iii), this yields

$$L_2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j \tag{8.88}$$

Since $\{\psi_{j,k}\}_k$ with

$$\psi_{j,k} = 2^{-j/2}\psi(2^j x - k) \tag{8.89}$$

is the orthonormal basis of W_j, the whole family $\{\psi_{j,k}\}_{j,k}$ is the orthonormal basis of $L_2(\mathbb{R})$.

In order to construct the orthonormal wavelet basis in $L_2(\mathbb{R}^2)$, we first notice that the family $\{\varphi(\cdot - l)\varphi(\cdot - m)\}_{l,m}$ is the orthonormal basis of $V_0^2 = V_0 \otimes V_0$. It is also easy to see that for any j, the family $\{2^{-j}\varphi(2^{-j} \cdot -n)\varphi(2^{-j} \cdot -k)\}_{n,k}$ is the orthonormal basis of $V_j^2 = V_j \otimes V_j$ and $V_j^2 \subset V_{j-1}^2$. Using the fact that the set of all functions of view $f(x)g(y)$ is dense in $L_2(\mathbb{R}^2)$, we conclude that subspaces V_j^2 constitute multi-resolution approximation of $L_2(\mathbb{R}^2)$ in the sense (i)–(iii). Then, representing V_{j-1}^2 by the orthogonal sum

$$V_{j-1}^2 = V_j^2 \oplus W_j^2 \tag{8.90}$$

and recalling that $V_{j-1} = V_j \oplus W_j$, we obtain

$$(V_j \otimes V_j) \oplus (V_j \otimes W_j \oplus W_j \otimes V_j \oplus W_j \otimes W_j) = (V_j \otimes V_j) \oplus W_j^2, \qquad (8.91)$$

which yields

$$W_j^2 = (V_j \otimes W_j) \oplus (W_j \otimes V_j) \oplus (W_j \otimes W_j). \qquad (8.92)$$

Taking into account that $L_2(\mathbb{R}^2) = \underset{j}{\oplus} W_j^2$, we conclude that the family $\{\psi_{j,k}^1, \psi_{j,k}^2, \psi_{j,k}^3\}_{j\in\mathbb{Z}, k\in\mathbb{Z}^2}$, where $\psi_{j,k}^i = 2^{-j}\psi^i(2^{-j}x - k_1, 2^{-j}y - k_2)$, and

$$\psi^1(x, y) = \varphi(x)\psi(y), \qquad (8.93)$$

$$\psi^2(x, y) = \psi(x)\varphi(y), \qquad (8.94)$$

$$\psi^3(x, y) = \psi(x)\psi(y), \qquad (8.95)$$

constitutes the orthonormal basis in $L_2(\mathbb{R}^2)$.

MRA allows to fulfill the wavelet transformation of f by means of the cascade filter bank algorithm. We describe this algorithm for 1D. The extension to 2D is straightforward.

Rescaling (8.68) and (8.85) to $\varphi_{j,k} \sum_n h_{n-2k}\varphi_{j-1,n}$ and to $\psi_{j,k} \sum_n g_{n-2k}\varphi_{j-1,n}$, respectively, and denoting $f_{j,k} = \langle f, \varphi_{j,k} \rangle, d_{j,k} = \langle f, \psi_{j,k} \rangle$, one obtains the filter bank decomposition at step j:

$$f_{j,k} = \sum_n f_{j-1,n} h_{n-2k},$$
$$d_{j,k} = \sum_n f_{j-1,n} g_{n-2k}. \qquad (8.96)$$

On the other hand, denoting with $\mathbf{P}A$ the operator of the orthogonal projection onto the subspace A, we can write

$$\mathbf{P}V_{j-1} = \mathbf{P}V_j + \mathbf{P}W_j. \qquad (8.97)$$

Therefore, for any $f \in L_2(\mathbb{R})$,

$$\sum_n \langle f, \varphi_{j-1,n} \rangle \varphi_{j-1,n} = \sum_n \langle f, \varphi_{j,n} \rangle \varphi_{j,n} + \sum_n \langle f, \psi_{j-1,n} \rangle \psi_{j,n}. \qquad (8.98)$$

Building the inner product of both sides of (8.98) with $\varphi_{j-1,k}$ yields the reconstruction formula

$$f_{j-1,k} = \sum_n f_{j,n} h_{k-2n} + \sum_n d_{j,n} g_{k-2n} \tag{8.99}$$

which is fulfilled recursively down to $j = 0$.

Let us consider an example of constructing orthonormal bases with box splines $\mathcal{B}_m$. Per definition, $\mathcal{B}_m$ is the $m - 1$ times convolution of $\mathcal{B}_1$ with itself, and

$$\mathcal{B}_1(x) = \begin{cases} 1 \text{ if } 0 \leq x < 1, \\ 0 \text{ otherwise.} \end{cases} \tag{8.100}$$

For $\mathcal{B}_m$, it is known that its integer translates constitute Riesz basis of $V_0^m = \overline{L_2(\mathbb{R}) \cap S_m^0}$ with the Riesz-bounds A_m and B_m

$$A_m = 2\pi^{-2m}(2^{2m} - 1) \sum_{k=1}^{\infty} k^{-2m}, \ B_m = 1 \tag{8.101}$$

(see, e.g., Chap. 4 of [9]). The two-scale relation for B_m is known to be

$$\mathcal{B}_m(x) = \sum_{n=0}^{m} 2^{-m+1/2} \binom{m}{n} \mathcal{B}_m(2x - n). \tag{8.102}$$

Since $\hat{\mathcal{B}}_m(\omega) = \left(\hat{\mathcal{B}}_1(\omega)\right)^{m-1}$ and

$$\hat{\mathcal{B}}_1(\omega) = \frac{\sin(\omega/2)}{\omega/2} e^{-i\omega/2}, \tag{8.103}$$

we obtain that $\hat{\mathcal{B}}_m(0) = 1$. That is, $\mathcal{B}_m$ is a scaling function that generates a multi-resolution approximation $\ldots \subset V_j^m \subset V_{j-1}^m \subset \ldots$, where $V_j^m = \overline{L_2(\mathbb{R}) \cap S_m^j}$.

By $m = 1$, the coefficients of the two-scale relation (8.102) are $h_0 = h_1 = 2^{-1/2}$. The wavelet associated with $\mathcal{B}_1$, constructed with the help of (8.85) and (8.86), is

$$\psi(x) = \mathcal{B}_2(2x) - \mathcal{B}_1(2x - 1) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1/2, \\ -1 & \text{if } 1/2 \leq x \leq 1, \\ 0, & \text{otherwise.} \end{cases} \tag{8.104}$$

This is the well-known Haar wavelet. Since Riesz-bounds $A_1 = B_1 = 1$, the translates $\mathcal{B}_1(\cdot - n)$ are orthonormal (see above (8.63) and (8.65), and consequently so are the translates $\psi(\cdot - n)$. The whole family $\{2^{-j/2}\psi(2^{-j} \cdot -n)\}_{j,n}$ is the orthonormal basis of $L_2(\mathbb{R})$.

As it follows from (8.101), the lower Riesz-bound A_m decays with m. Therefore, $A_m < 1$ for any $m > 1$, and as a consequence, the integer translates of $\mathcal{B}_m$ are not orthonormal. Applying to $\mathcal{B}_m$ the orthogonalization trick

$$\hat{\mathcal{B}}_m^{\#}(\omega) = \hat{\mathcal{B}}_m(\omega) \left(\sum_n |\hat{\mathcal{B}}_m(\omega + 2\pi n)|^2 \right)^{-1/2}, \quad \omega \in R, \tag{8.105}$$

one obtains functions $\mathcal{B}_m^{\#}$ which satisfy (8.65), and as a consequence, their integer translates are orthonormal. However, these functions are no more compactly supported, and neither are the associated wavelets called Battle-Lemarié wavelets.

In practical applications, numerically advantageous compactly supported wavelets are preferable. In applications where the perceptual assessment of the signal is important, e.g., medical imaging, the symmetry of wavelets is another important feature of preference: applying the asymmetric wavelets can yield asymmetric errors,[4] and it is known that the human visual system is less tolerant to asymmetric errors than to symmetric ones. Additionally, symmetric or antisymmetric wavelets are more efficient while treating boundaries of the compactly supported signal, such as an image. For such applications, the basis of $L_2(T)$, where $T \subset \mathbb{R}^n$ is supposed to be the support of the signal, has to be constructed. Normally this is done by modifying wavelets of the basis of $L_2(\mathbb{R}^n)$: wavelets, the support of which is entirely inside T are not changed; the wavelets, the support of which is entirely outside T are skipped; and those whose support overlaps the boundaries of T are modified. There are different ways to modify the boundary wavelets (see, e.g., Chap. 7 of [7]). For symmetric or antisymmetric wavelets, boundary wavelets are simply folded back away from boundaries. The decomposition in the so-obtained basis is equivalent to the decomposition of the signal that is symmetrically extended beyond T. An example of such extension for 1D signal supported on $T = [a, b]$ is given in Fig. 8.9.

It is well known (see, e.g., [6]) that constructing the orthonormal basis from compactly supported wavelet is possible only if the wavelet is asymmetric. The exception is the discontinuous Haar wavelet. For any continuous compactly supported wavelet, the symmetry and orthonormality are inconsistent. But compactly supported symmetric or antisymmetric wavelets can be used to construct biorthogonal bases of $L_2(\mathbb{R})$.

For the following we accept (without proof) several important facts about compactly supported scaling functions. Let h be the filter the transfer function of which is the trigonometric polynomial

$$\hat{h}(\omega) = \sum_{n=N_1}^{N_2} h_n e^{-i\omega n}. \tag{8.106}$$

In other words, h is a finite impulse response (FIR) filter, i.e., the filter finitely many taps of which are not zero. Let in addition

$$\hat{h}(0) = \sqrt{2} \quad \text{and} \quad \hat{h}(\pi) = 0. \tag{8.107}$$

---

[4] That is, errors that are described with asymmetric density distribution.
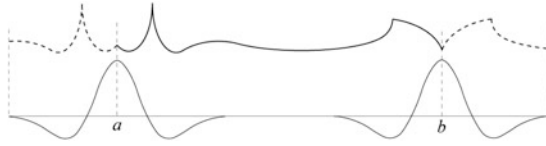
Fig. 8.9 To the boundary problem of the decomposition of the signal supported on the limited interval. The signal (*bold*) and its extension (*dashed*); two boundary wavelets located at most extreme positions

Then φ, the Fourier transform of which is

$$\hat{\varphi}(\omega) = \prod_{j=1}^{\infty} \frac{\hat{h}(2^{-j}\omega)}{\sqrt{2}}, \tag{8.108}$$

is a compactly supported square integrable scaling function. The function defined by (8.108) will be referred to as the one related to h. A pair of FIR low-pass filters h and $\tilde{h}$, both satisfying (8.107), are called dual if the families $\{\varphi(\cdot - n)\}_n$ and $\{\tilde{\varphi}(\cdot - n)\}_n$, where φ and $\tilde{\varphi}$ are related scaling functions, are biorthogonal. The necessary condition for h and $\tilde{h}$ to be dual is

$$\hat{h}(\omega)^{\bullet}\hat{\tilde{h}}(\omega) + \hat{h}(\omega + \pi)^{\bullet}\hat{\tilde{h}}(\omega + \pi) \equiv 2 \tag{8.109}$$

(for sufficient conditions, see [10]).

The multi-resolution analysis can be reformulated in terms of biorthogonal bases. Let h and $\tilde{h}$ be dual filters, and let the related scaling functions generate multi-scale approximations of $L_2(\mathbb{R})$

$$\ldots \subset V_j \subset V_{j-1} \subset \ldots \tag{8.110}$$

and

$$\ldots \subset \tilde{V}_j \subset \tilde{V}_{j-1} \subset \ldots \tag{8.111}$$

respectively.

Denote with W_j the subspace that completes the subspace V_j in V_{j-1} and with $\tilde{W}_j$ the complement subspace of $\tilde{V}_j$. Apparently $W_0 \perp \tilde{V}_0$ and $\tilde{W}_0 \perp V_0$. In the similar way as we did in the orthonormal case, but using (8.66) instead of (8.65), it is possible to show that there exist biorthogonal Riesz bases of W_0 and $\tilde{W}_0$ that are constituted with integer translates of wavelets ψ and $\tilde{\psi}$, the Fourier transform of which are

$$\hat{\psi}(\omega) = 2^{-1/2}\hat{\alpha}(\omega)\hat{\tilde{h}}(\omega/2 + \pi)^*\hat{\varphi}(\omega/2), \tag{8.112}$$

$$\hat{\tilde{\psi}}(\omega) = 2^{-1/2}\hat{\beta}(\omega)\hat{h}(\omega/2 + \pi)^*\hat{\tilde{\varphi}}(\omega/2), \tag{8.113}$$

with $\hat{\alpha}$ and $\hat{\beta}$ being 2π periodic, $|\hat{\alpha}(\omega)| \equiv 1$, $|\hat{\beta}(\omega)| \equiv 1$, and $\hat{\alpha}(\omega)\hat{\beta}(\omega)^* \equiv 1$. The families $\{2^{-j/2}\psi(2^{-j} \cdot -n)\}_{j,n}$ and $\{2^{-j/2}\tilde{\psi}(2^{-j} \cdot -n)\}_{j,n}$ are biorthogonal Riesz bases of $L_2(\mathbb{R})$.

For $\hat{\alpha}(\omega) = e^{-i\omega/2}$, the equalities (8.112) and (8.113) in the space domain are

$$\psi = \sum_n g_n \varphi_{-1,n} \; and \; \tilde{\psi} = \sum_n \tilde{g}_n \tilde{\varphi}_{-1,n} \tag{8.114}$$

with

$$\begin{aligned} g_n &= \langle \psi, \tilde{\varphi}_{-1,n} \rangle = (-1)^{n+1} \tilde{h}_{1-n} \quad \text{and} \\ \tilde{g}_n &= \langle \tilde{\psi}, \varphi_{-1,n} \rangle = (-1)^{n+1} h_{1-n}. \end{aligned} \tag{8.115}$$

The extension to the 2D case is made in a similar way as we did it in the orthonormal case. One defines three wavelets ψ^1, ψ^2, and ψ^3 exactly as in (8.93)–(8.95), which in this case generate the Riesz basis of $L_2(\mathbb{R}^2)$. The wavelets $\tilde{\psi}^1(x,y) = \tilde{\varphi}(x)\tilde{\psi}(y)$, $\tilde{\psi}^2(x,y) = \tilde{\psi}(x)\tilde{\varphi}(y)$, and $\tilde{\psi}^3(x,y) = \tilde{\psi}(x)\tilde{\psi}(y)$ generate the dual Riesz basis of $L_2(\mathbb{R}^2)$.

In the same way as in the orthonormal case, the two-scale relation on the one hand and the relations (8.114) on the other imply the fast filter bank algorithm of the wavelet transform of f:

$$\langle f, \varphi_{j,k} \rangle = \sum_m \langle f, \varphi_{j-1,m} \rangle h_{m-2k}, \tag{8.116}$$

$$\langle f, \psi_{j,k} \rangle = \sum_m \langle f, \psi_{j-1,m} \rangle g_{m-2k}. \tag{8.117}$$

In order to obtain the reconstruction formula, let us introduce the operator

$$\mathbf{P}_V f = \sum_k \langle f, \lambda_k \rangle \tilde{\lambda}_k, \tag{8.118}$$

where $\{\lambda_k\}_k$, $\{\tilde{\lambda}_k\}_k$ are biorthonormal Riesz bases of the subspace V and $\tilde{V}$, respectively. The operator $\mathbf{P}_V$ is an orthoprojector. Since the families $\{\varphi_{j,n}, \psi_{j,n}\}_n$ and $\{\tilde{\varphi}_{j,n}, \tilde{\psi}_{j,n}\}_n$ are biorthogonal Riesz bases of V_{j-1} and $\tilde{V}_{j-1}$, respectively, we obtain that $\mathbf{P}V_{j-1} = \mathbf{P}V_j + \mathbf{P}W_j$, that is, for any $f \in L_2(\mathbb{R})$, the identity

$$\sum_n \langle f, \varphi_{j-1,n} \rangle \tilde{\varphi}_{j-1,n} = \sum_n \langle f, \varphi_{j,n} \rangle \tilde{\varphi}_{j,n} + \sum_n \langle f, \psi_{j,n} \rangle \tilde{\psi}_{j,n} \tag{8.119}$$

is valid. The inner product of (8.119) with $\varphi_{j-1,k}$ yields the reconstruction formula

$$\langle f, \varphi_{j-1,k} \rangle = \sum_n \langle f, \varphi_{j,n} \rangle \tilde{h}_{k-2n} + \sum_n \langle f, \psi_{j,n} \rangle \tilde{g}_{k-2n}. \qquad (8.120)$$

We conclude our review by considering the redundant dyadic wavelet transform that is defined by

$$Wf_j(u) = \langle f, \psi_j(\cdot - u) \rangle, \ u \in \mathbb{R}, \ j \in \mathbb{Z}, \qquad (8.121)$$

where $\psi_j(x) = 2^{-j/2}\psi(2^{-j}x)$. In a similar way as we did before, let us introduce the self-adjoint operator $\mathbf{H} : L_2(\mathbb{R}) \rightarrow L_2(\mathbb{R})$, defined by

$$\mathbf{H}f(t) = \sum_j 2^{-j} \int Wf_j(u)\psi_j(t - u)\mathrm{d}u. \qquad (8.122)$$

If there exist constants $0 < A \leq B < \infty$ such that the bounding condition

$$A\|f\|^2 \leq \langle \mathbf{H}f, f \rangle \leq B\|f\|^2 \qquad (8.123)$$

holds for all $f \in L_2(\mathbb{R})$, then the operator $\mathbf{H}$ is invertible with the inverse $\mathbf{H}^{-1}$ bounded by B^{-1} and A^{-1}. The sufficient condition for $\mathbf{H}$ to satisfy (8.123) is

$$A \leq \sum_j \left|\hat{\psi}(2^j\omega)\right|^2 \leq B. \qquad (8.124)$$

In order to see this, the inequality (8.124) has to be multiplied by $|\hat{f}(\omega)|^2$ and integrated over ω. Then, using the representation

$$\widehat{Wf_j}(\omega) = 2^{j/2}\hat{\psi}(2^j\omega)^*\hat{f}(\omega) \qquad (8.125)$$

and taking into account the Parseval's identity, one obtains

$$A\|f\|^2_{L_2} \leq \sum_j 2^{-j}\|Wf_j\|^2_{L_2} \leq B\|f\|^2_{L_2}, \qquad (8.126)$$

which is another equivalent notation of (8.123).

Applying $\mathbf{H}^{-1}$ to both sides of (8.122) yields

$$f(t) = \sum_j 2^{-j} \int Wf_j(u)\tilde{\psi}_j(t - u)\mathrm{d}u, \qquad (8.127)$$

where $\tilde{\psi} = \mathbf{H}^{-1}\psi$ is called a dual wavelet. In order to determine $\mathbf{H}^{-1}$, one has to notice that

$$\widehat{\mathbf{H}f}(\omega) = \hat{f}(\omega) \sum_j |\hat{\psi}(2^j\omega)|^2, \tag{8.128}$$

that is, in the Fourier domain, the operator $\mathbf{H}$ is associated with the operator $\hat{\mathbf{H}}$, which is simply the multiplicator $\sum_j |\hat{\psi}(2^j\omega)|^2$. Therefore,

$$\hat{\mathbf{H}}^{-1} = \frac{1}{\sum\limits_j |\hat{\psi}(2^j\omega)|^2}, \tag{8.129}$$

and consequently,

$$\hat{\tilde{\psi}}(\omega) = \hat{\mathbf{H}}^{-1}\hat{\psi}(\omega) = \frac{\hat{\psi}(\omega)}{\sum\limits_j |\hat{\psi}(2^j\omega)|^2} \tag{8.130}$$

which makes sense due to (8.124). In fact, any $\tilde{\psi}$ satisfying

$$\sum_j \hat{\tilde{\psi}}(2^j\omega)\hat{\psi}(2^j\omega)^* = 1, \omega \in \mathbb{R}\backslash\{0\}, \tag{8.131}$$

satisfies also (8.127). This follows from the fact that in the Fourier domain, the relation (8.127) is represented by

$$\hat{f}(\omega) = \hat{f}(\omega) \sum_j \hat{\tilde{\psi}}(2^j\omega)\hat{\psi}(2^j\omega)^*. \tag{8.132}$$

Let h and g be a pair of FIR filters, and let $\hat{h}(0) = \sqrt{2}$. Let φ be the scaling function that relates to h via (8.108). Then, apparently,

$$\hat{\varphi}(2\omega) = 2^{-1/2}\hat{h}(\omega)\hat{\varphi}(\omega). \tag{8.133}$$

The Fourier transform of the corresponding wavelet is defined by

$$\hat{\psi}(2\omega) = 2^{-1/2}\hat{g}(\omega)\hat{\varphi}(\omega). \tag{8.134}$$

Let $\tilde{h}$ and $\tilde{g}$ be, in turn, another pair of FIR filters, and $\tilde{\varphi}$ and $\tilde{\psi}$ are the related scaling function and the corresponding wavelet. Since both φ and $\tilde{\varphi}$ are compactly supported, they are from $L_1(\mathbb{R}) \cap L_2(\mathbb{R})$. Therefore, their Fourier transforms are continuous. This fact can be used to show that the condition

$$\hat{\tilde{h}}(\omega)\hat{h}(\omega)^* + \hat{\tilde{g}}(\omega)\hat{g}(\omega)^* = 2 \tag{8.135}$$

is sufficient for ψ and $\tilde{\psi}$ to satisfy (8.131).

Hence, given h and g, one has a certain freedom in choosing filters $\tilde{h}$ and $\tilde{g}$. For example, one can set $\tilde{h} = h$. Then from (8.135), we obtain that

$$\hat{\tilde{g}}(\omega) = \frac{2 - |\hat{h}(\omega)|^2}{\hat{g}(\omega)^*}. \tag{8.136}$$

The relation (8.136) imposes certain restrictions on g. Namely, zeros of $\hat{g}$ must coincide with zeros of $2 - |\hat{h}(\omega)|^2$ and be of the corresponding order. In particular at $\omega = 0$, there must be one zero of $\hat{g}$. Since the number of zeros of $\hat{g}$ at $\omega = 0$ coincides with the number of zeros of $\hat{\psi}$ at $\omega = 0$, it follows that the corresponding wavelet ψ must have one vanishing moment.

A described strategy was applied in [11] for the construction of the so-called *Mallat-wavelet*. Redundant wavelet family generated with this wavelet has been proven to be efficient for the reconstruction of the image from the wavelet coefficients measured over multi-scale edges. The starting point while constructing the wavelet in [11] was to set

$$\varphi(x) = \mathcal{B}_3(x + 1), \tag{8.137}$$

where $\mathcal{B}_3$ is the box spline of degree 2. The Fourier transform of φ is

$$\hat{\varphi}(\omega) = \left(\frac{\sin(\omega/2)}{\omega/2}\right)^3 e^{-i\omega/2}. \tag{8.138}$$

We know that $\mathcal{B}_3$ satisfies the two-scale relation (see (8.102)), and so consequently does φ. That is, there exists a filter h such that $\varphi = \sum_n h_n \varphi_{-1,n}$. The transfer function of h is yielded from (8.133):

$$\hat{h}(\omega) = \sqrt{2}\frac{\hat{\varphi}(2\omega)}{\hat{\varphi}(\omega)} = \sqrt{2}\cos^3\frac{\omega}{2}e^{-i\omega/2}. \tag{8.139}$$

Let us turn our attention to the fact, that the Mallat-wavelet was constructed using the additional requirement to be efficient for the detection of edges. From (8.134) and the fact that both g and h are FIR filter, it follows that ψ is compactly supported. Since ψ defined by (8.136) must have one vanishing moment, then due to (8.37), there must exist compactly supported smoothing function θ such that

$$\psi = -\theta'. \tag{8.140}$$

Due to (8.38), we obtain then that $Wf_j(u) = 2^j \frac{d}{dt} f * \bar{\theta}_j(t)$, so that local maxima of $|Wf_j(u)|$ are exactly local maxima of $|d/dx(f * \bar{\theta}_j)|$. The latter are edge points within

Canny's approach of an edge detection [12]. Hence, the desired wavelet is determined by the choice of θ. For example, one can set

$$\theta = \varphi. \tag{8.141}$$

However, this choice leads to a piecewise linear wavelet. Since the regularity of the wavelet is crucial for the quality of reconstruction, in [11], θ was set to

$$\theta(x) = 2^{-1} \mathcal{B}_4(2x + 1). \tag{8.142}$$

The Fourier transform of the wavelet $\psi = -\theta'$, referred to as the Mallat-wavelet, is

$$\hat{\psi}(\omega) = -i\frac{\omega}{4} \left(\frac{\sin(\omega/4)}{\omega/4} \right)^4 e^{-i\omega/2}. \tag{8.143}$$

The transfer function of the corresponding filter g, obtained using (8.134) and (8.138), is

$$\hat{g}(\omega) = -i\sqrt{2} \sin\frac{\omega}{2} e^{-i\omega/2}. \tag{8.144}$$

Finally using (8.136), for the transfer function of the reconstruction filter $\tilde{g}$, one obtains

$$\hat{\tilde{g}}(\omega) = -i\sqrt{2} \sin\frac{\omega}{2} \left(\sum_{k=0}^{3} \cos^{2k}\frac{\omega}{2} \right) e^{-i\omega/2}. \tag{8.145}$$

Using pairs of filters h, g and $\tilde{h}$, $\tilde{g}$, a fast dyadic wavelet transform can be calculated with a filter bank algorithm called *algorithme à trous* [13]. The corresponding filter bank algorithm in 2D is

$$f_{j+1}(m, n) = f_j * \bar{h}_j\bar{h}_j(m, n), \tag{8.146}$$

$$d_{j+1}^1(m, n) = f_j * \bar{g}_j\delta(m, n), \tag{8.147}$$

$$d_{j+1}^2(m, n) = f_j * \delta\bar{g}_j(m, n), \tag{8.148}$$

where $\alpha\beta(m, n)$ is a separable two-dimensional filter, i.e., $\alpha\beta(m, n) = \alpha(m)\beta(n)$; $\delta(n)$ is a discrete Dirac; and α_j is the filter obtained from α by inserting $2^j - 1$ zeros (*trous*) between its samples. The reconstruction is made recursively by means of

$$f_j(m, n) = f_{j+1} * h_jh_j(m, n) + d_{j+1}^1 * \tilde{g}_j\tilde{q}_j(m, n) + d_{j+1}^2 * \tilde{q}_j\tilde{g}_j(m, n), \tag{8.149}$$

Fig. 8.10 Analysis (*left*) and synthesis (*right*) steps of the filter bank filtration
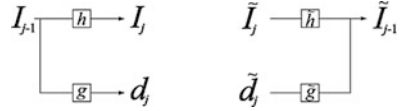
where $\tilde{g}$ is defined by (8.136), and for $\tilde{q}$, we have

$$\hat{\tilde{q}}(\omega) = \frac{1 + |\hat{h}(\omega)|^2}{2} \tag{8.150}$$

(see [11]).

We have described several of the most often applied wavelet decomposition-reconstruction frameworks. All of them can be implemented with fast filter bank algorithms, the common principal scheme of which is depicted in Fig. 8.10.

In the following, we give some heuristic considerations concerning the question of how these schemes can be used for the reduction of noise. As one can see in Fig. 8.10, during the analysis step, the coarse approximation I_{j-1} is further decomposed in the coarser approximation I_j and the details image d_j that contains the wavelet coefficients $\langle I, \psi_{j,k} \rangle$. Due to the localization properties of wavelets, one can change the wavelet coefficients within some region R of the image d_j, not bothering much about what effect this would have for those structures of the synthesized image $\tilde{I}_{j-1}$ which lie apart from R. If, for example, we want to reproduce the region R of the image I_j in the image $\tilde{I}_{j-1}$, we should suppress all details of d_j located in R, that is, the corresponding wavelet coefficients have to be set to zero. In order to suppress noisy details of the image I_{j-1}, we need to know which detail is due to noise.

There are pretty different scenarios to decide whether (or not) the given wavelet coefficient is due to noise. We mention an approach developed in [14]. The idea lying in the background of the method proposed there is to extract *relevant* details by means of comparative analysis of two or several images of the same object. It will be not quite far from the truth to say, that the comparative analysis is what the radiologist does while investigating the image. Evaluating the X-ray image, the observer automatically extracts features which he qualifies as *relevant*. Speaking figuratively, the skilled observer *correlates* a real image with an imaginary one that was created by his visual system. Motivated by the wish to facilitate the daily task of the radiologist, the authors have decided to apply the "correlation approach" in clinical practice. As soon as we have two images of the same object, we can define as relevant those details, which are present in both of them. The other details are qualified as noise and suppressed. In [14] the results of the method applied to quasi identical as well as to those which are slightly deformed relative to each other were reported. The technical aspect of the method is as follows. The both images, say A and B, are decomposed using the dyadic wavelet frame generated by the Mallat mother wavelet. The reason for the choice of this wavelet is the shift-invariance of the related wavelet transform. After the decomposition, details of

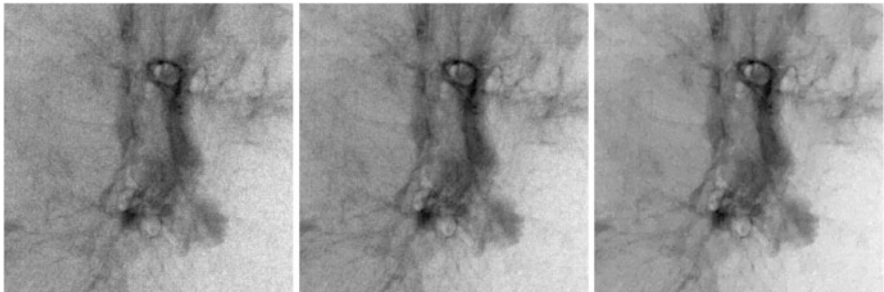**Fig. 8.11** Weighting masks appearing as a result of comparative analysis of details images
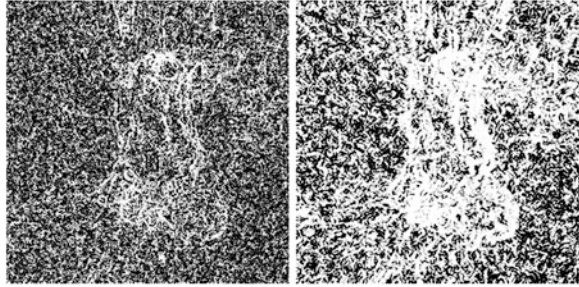

Fig. 8.12 Denoising of the lung phantom. *Left*: one of the two images of the same object. *Middle*: the sum of both images. *Right*: the image reconstructed from details weighted with masks shown in Fig. 8.11

image A and image B are correlated with each other at each resolution level. The result of the correlation is the *weighting matrix*, each element of which expresses a measure of similarity between the detail of image A and the detail of image B at the corresponding location. Denoting the measure of similarity with $w_{m,k}^j$, where j is the resolution level and (m, k) the location within the image, we have

$$0 \leq w_{m,k}^j \leq 1. \tag{8.151}$$

If $w_{m,k}^j = 1$, details of A and B at the level j and location (m, n) coincide; the case $w_{m,k}^j = 0$ means that the corresponding detail is due to noise; the intermediate cases should point at the portion of the noise at the location. Here we demonstrate the functionality of the method applying it to images of a lung phantom, the patch from the one of which is shown on the left of Fig. 8.4. Figure 8.11 shows weighting matrices of the first two resolution levels.

During the following step, elements of detail images are weighted by elements of corresponding weighting matrices. The final image is reconstructed from weighted coefficients. The result of the reconstruction is shown on the right-hand side of Fig. 8.12.

In spite of the seemingly disadvantage expressed by the necessity to have two or more images, the method provides a reduction of noise with minimal risk to lose useful information. In practice, this method could be applied to two images each acquired at half doses, for example, by means of detector based on CCD camera.

In this chapter, we described the mathematical foundations of image denoising and its practical implications. The basic principles are also valid for three-dimensional imaging and can in general be performed in all types of clinical diagnostic imaging. In nuclear medical imaging, the consequences of image denoising have to be studied in very much detail due to the fact that in nuclear medical imaging, the signal-to-noise ratio of the raw data is very poor to avoid high doses of radiation to the patient. This implies that unreflected use of denoising methods particularly also those based on compressing information can result in mainly reducing relevant image structures together with or even instead of the noise. Which kind of noise reduction can be used without deteriorating necessary image information has to be studied for each single application.

References

1. Zygmund A (1959) Trigonometric series, chapter X, vol 2. Cambridge University, Cambridge
2. Oppelt A (2005) Imaging systems for medical diagnostics: fundamentals, technical solutions, and applications for systems applying ionizing radiation, nuclear magnetic resonance, and ultrasound. Wiley-VCH Verlag GmbH, New York, 2/E
3. Wiener N (1949) Extrapolation, interpolation, and smoothing of stationary time series. Wiley, New York
4. Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. In: ICCV'98: proceedings of the sixth international conference on computer vision. IEEE Computer Society, Washington, DC, p 839
5. Aurich V, Weule J (1995) Non-linear Gaussian filters performing edge preserving diffusion. In: Proceedings of the 7th DAGM-Symposium, Bielefeld, Springer, pp 538–545
6. Daubechies I (1992) Ten lectures on wavelets. SIAM, Philadelphia, PA
7. Mallat S (1999) A wavelet tour of signal processing. Academic, San Diego, CA
8. Mallat S (1989) A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 11(7):674–693
9. Chui CK (1992) Introduction to wavelets. Academic, London
10. Cohen A, Daubechies I, Feauveau JC (1992) Orthonormal bases of compactly supported wavelets. Commun Pure Appl Math 45:485–560
11. Mallat S, Zhong S (1992) Characterization of signals from multiscale edges. IEEE Trans Pattern Anal Mach Intell 4(7):70–732
12. Canny J (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 36:961–1005
13. Holschneider M, Kronland-Martinet R, Morlet J, Tchamitchian P (1989) A real time algorithm for signal analysis with help of the wavelet transform. In: Combes JM, Grossmann A, Tchamitchian P (eds) Wavelets: time frequency methods and phase space. Springer, Berlin, pp 286–297
14. Tischenko O, Hoeschen C, Buhr E (2005) Reduction of anatomical noise in medical X-ray images. Radiat Prot Dosimetry 4(Nos ß3):69–74