Chapter 2
Intelligent Digital Signal Processing and Feature Extraction Methods

János Szalai and Ferenc Emil Mózes

Abstract Intelligent systems comprise a large variety of applications, including ones based on signal processing. This field benefits from considerable popularity, especially with recent advances in artificial intelligence, improving existing processing methods and providing robust and scalable solutions to existing and new problems. This chapter builds on well-known signal processing techniques, such as the short-time Fourier and wavelet transform, and introduces the concept of instantaneous frequency along with implementation details. Applications featuring the presented methods are discussed in an attempt to show how intelligent systems and signal processing can work together. Examples that highlight the cooperation between signal analysis and fuzzy c-means clustering, neural networks and support vector machines are being presented.

Keywords Signal processing · Signal-adaptive processing · Frequency domain transforms · Instantaneous frequency · Fuzzy c-means · Fuzzy systems · Support vector machine · Frequency analysis · Time-frequency analysis · Neural networks

2.1 Introduction

Intelligent systems on their own are not always enough to handle complex tasks. In these cases the preprocessing of the signals is a necessary step. Often the preprocessing algorithms not only convert the initial data into a more advantageous format, but they are also capable of realizing feature extraction on the data. This aspect

J. Szalai (✉)
Technical University of Cluj Napoca, 128-130 21 December 1989 Boulevard, 400604 Cluj Napoca, Romania
e-mail: szalai@gmail.com

F.E. Mózes
Petru Maior University of Târgu Mures, 1 Nicolae Iorga Street, 540088 Târgu Mures, Romania
e-mail: mozes.ferenc.emil@gmail.com

R. Kountchev and K. Nakamatsu (eds.), *New Approaches in Intelligent Image Analysis*, Intelligent Systems Reference Library 108,
DOI 10.1007/978-3-319-32192-9_2

is equally important as the concept of intelligent systems themselves, for no classification can succeed if the features of the input data are not emphasized correctly beforehand.

The goal of the chapter is to present well-known signal processing methods and the way these can be combined with intelligent systems in order to create powerful feature extraction techniques. In order to achieve this, several case studies are presented to illustrate the power of hybrid systems. The main emphasis is on instantaneous time-frequency analysis, since it is proven to be a powerful method in several technical and scientific areas. The authors' contributions to the computation of the instantaneous frequency and application of the empirical mode decomposition are also presented, highlighting the limitations of the existing methods and showing at the same time a possible approach on T-wave peak detection in electrocardiograms.

Classical signal processing methods have been widely used in different fields of engineering and natural sciences in order to highlight meaningful information underlying in a wide variety of signals. In this chapter we aim to present not only the best known signal processing methods, but also ones that proved to be the most useful. The oldest and most utilized method is the Fourier transform, which has been applied in several domains of scientific data processing, but it has very strong limitations due to the constraints it imposes on the analyzed data. Then the short-time Fourier transform and the wavelet transform are presented, as they provide both temporal and frequency information as opposed to the Fourier transform. These methods form the basis of most applications nowadays, as they offer the possibility of time-frequency analysis of signals. Finally, the Hilbert-Huang transform is presented as a novel signal processing method, which introduces the concept of instantaneous frequency that can be determined for every time point, making it possible to have a deeper look into different phenomena.

The combinations of these methods with intelligent systems are described in the second part of the chapter. Several applications are presented where fuzzy classifiers, support vector machines and artificial neural networks are used for decision making. Interconnecting these intelligent methods with signal processing will result in hybrid intelligent systems capable of solving computationally difficult problems.

2.2 The Fourier Transform

It is a well known fact that periodic functions can be expanded into Fourier series using weighted sums of sines and cosines. In the real world, however, most of the physical phenomena can't be treated as periodical occurrences. For these shapes there exists the Fourier transform which is an integral taken over the whole definition domain. It is thus assumed that the function is represented on the whole real axis. Applying the Fourier transform to real-world signals makes the spectral analysis of phenomena possible, offering more information than it would be available in the time domain.

The definition of the Fourier transform is given by Eq. 2.1 for any integrable function $f : \mathbb{R} \rightarrow \mathbb{C}$ [1]

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-2\pi i t\omega} dt \tag{2.1}$$

where t is the time variable and ω is the frequency variable of the Fourier plane.

The Fourier transform describes the original function (or signal) as a bijection, thus the original function can be fully recovered if its Fourier transform is known. This process is achieved by applying the inverse Fourier transform described by Eq. 2.2.

$$f(t) = \int_{-\infty}^{\infty} F(\omega) e^{2\pi i t\omega} d\omega \tag{2.2}$$

The Fourier transform is built upon the Fourier series and in a straightforward manner on the decomposition of functions into sinusoidal basis functions. This can be easily proven using Euler's formula:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)(\cos(-2\pi t\omega) + i \sin(-2\pi t\omega)) dt \tag{2.3}$$

However, most of the practical applications are not dealing with continuous domain signals and functions. Instead they use digitally sampled signals. For these signals an adapted version of the Fourier transform can be used, called the discrete Fourier transform (DFT), defined by Eq. 2.4 for a time series x_n, where n is the sample number, k is the wave number and N is the total number of samples.

$$X_k = \sum_{n=0}^{N-1} x_n e^{\frac{-2\pi i k n}{N}} \tag{2.4}$$

Naturally, this also has an inverse transform described by Eq. 2.5.

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{\frac{2\pi i k n}{N}} \tag{2.5}$$

Whenever possible, the fast Fourier transform (FFT) is used instead of the discrete Fourier transform, mostly based on performance and execution time considerations [2]. The most efficient way to use the FFT is to have signals with number of samples equal to some power of two. This has to do with the way the FFT algorithm is constructed. Further details can be found in [3, 4]. The computational complexity that can be achieved this way is $O(n \log n)$ as opposed to the DFT's $O(n^2)$.

The Fourier transform considers the analyzed signal in its full length, it is not using an analyzing window, thus sometimes identifies false frequency components that in reality aren't present in the signal. This explains why the best results are achieved for full periods of periodic signals.

2.2.1 Application of the Fourier Transform

The FFT is used in various fields, such as multimedia [5], optical [6], seismological [7], spectroscopy [8] or magnetic resonance signal processing [9]. In this part an application in the field of magnetic resonance imaging (MRI) is going to be presented briefly.

Magnetic resonance imaging produces images of the human body by exciting the hydrogen (^{1}H) nuclei with radio frequency pulses and then measuring the radio frequencies emitted by these nuclei as they recover to their initial energy state. Localization of different frequencies is done by modifying the main magnetic field using imaging gradients along the axes of the imaged plane. The frequencies emitted by different nuclei are equal to their precessional frequencies [9]. By measuring the electrical current induced in a receiver coil by the emitted RF frequencies the Fourier-space (or k-space) of the image is constructed using frequencies relative to the position of imaging gradients. The measured spatial frequency spectrum is then transformed to space domain (image domain) using the inverse Fourier transform.

This is an example of using the inverse Fourier transform in more than one dimension. Equations 2.6–2.9 describe the two dimensional continuous Fourier transform, the two dimensional continuous inverse Fourier transform, the two dimensional discrete Fourier transform and respectively, the two dimensional discrete inverse Fourier transform.

$$F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-2\pi i(ux+vy)}dxdy \tag{2.6}$$

$$f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)e^{2\pi i(ux+vy)}dudv \tag{2.7}$$

In Eqs. 2.6 and 2.7 x and y are space variables, u and v are spectral variables.

$$X_{k,l} = \frac{1}{\sqrt{MN}}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} x_{n,m}e^{-2\pi i\left(\frac{mk}{M}+\frac{nl}{N}\right)} \tag{2.8}$$

$$x_{n,m} = \frac{1}{\sqrt{MN}}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} X_{k,l}e^{2\pi i\left(\frac{mk}{M}+\frac{nl}{N}\right)} \tag{2.9}$$

In Eqs. 2.8 and 2.9 n and m are discrete space variables, k and l are discrete wave numbers.

Figure 2.1 shows the equivalence between 2D spectral space and 2D image space.

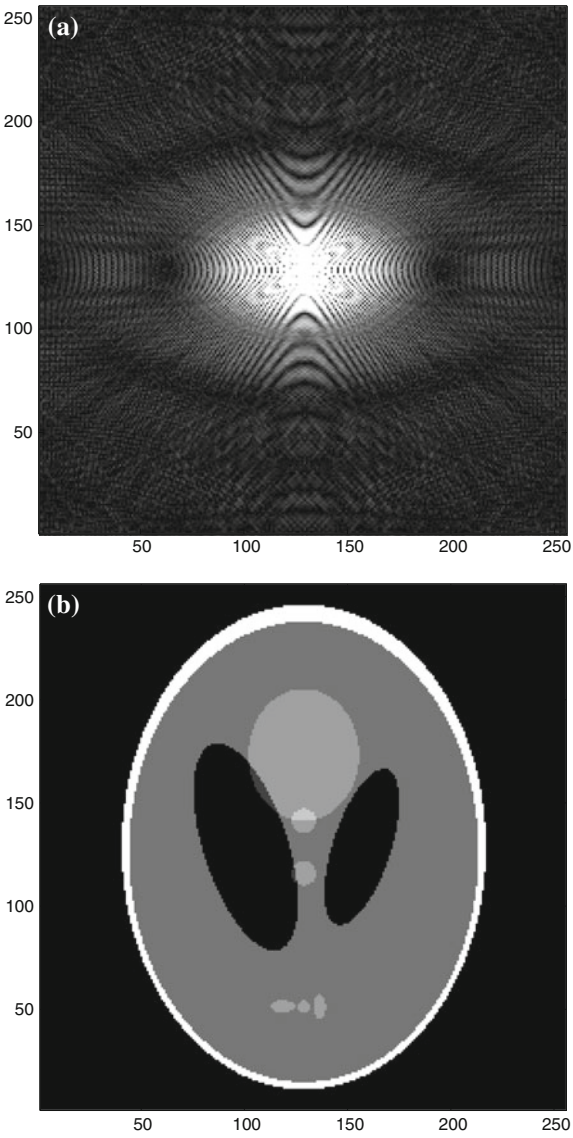

Fig. 2.1 Equivalence between k-space and image space. **a** k-space representation of the Shepp-Logan head phantom. **b** Reconstructed image of the Shepp-Logan head phantom

2.3 The Short-Time Fourier Transform

The Fourier transform does not reveal any temporal information about the frequency components present in the signal, making it impossible to locate them in specific applications. The Fourier transform "looks" at the analyzed signal during its whole time span and identifies frequency components as if they were present during the whole signal. Obviously, this is not the case in most of the applications. A way to introduce temporal information in the Fourier transform is to apply it on the signal using a sliding window with a constant width [1]. Sliding this window over the whole length of the signal will offer the possibility to get both spectral and temporal information about the signal. Equation 2.10 defines the short-time Fourier transform, where w is the windowing function, t is the time variable, ω is the frequency variable and τ is the time variable of the spectrogram.

$$F(\omega,\tau)=\int_{-\infty}^{\infty} f(t)e^{-2\pi i\omega t}w(t-\tau)dt \tag{2.10}$$

Similarly to the Fourier transform, the short time Fourier transform is also invertible. The original signal can be recovered using Eq. 2.11.

$$f(t)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\omega,\tau)e^{2\pi i\omega t}d\omega d\tau \tag{2.11}$$

The discrete short-time Fourier transform is described by Eq. 2.12, where similarly to the discrete Fourier transform, n is the discrete sample number of the time series x_n and k is the wave number. The time variable of spectrogram is represented by m. Inverting the discrete short-time Fourier transform is not as simple as the continuous one, it is heavily based on knowledge about the window function and the overlap between successive windows.

$$X_{m,k}=\sum_{n=0}^{\infty} x_n w_{n-m}e^{-\frac{2\pi ink}{N}} \tag{2.12}$$

A very important element of the short time Fourier transform is the windowing function. The way the windowing function and its parameters are chosen, will affect the spectrum produced by the transform. Hann and Hamming windows are often used due to their favorable frequency responses. The most important parameters of a window function are its length and overlap—the resolution of the short time Fourier transform in both the temporal and spectral domain is influenced by these parameters. As we reduce the length of the window, temporal localization gets better in detriment of spectral resolution. If the window width is increased, the temporal resolution decreases but the frequency resolution increases. This

phenomenon is strongly related to Heisenberg's uncertainty principle and is expressed by the inequality described with Eq. 2.13 [10].

$$\sigma_t \sigma_f \geq \frac{1}{4\pi} \tag{2.13}$$

Here, σ_t and σ_f represent the standard deviations in time and frequency, respectively, and their product is bounded. The windowing function that can offer maximal resolution in both domains is the Gaussian window [11]. This limitation on resolution is the major drawback of the transform, making it unsuitable for situations where precise time and frequency information are both essential, e.g. in some electrophysiological applications.

Another drawback of the short time Fourier transform is that it can become numerically unstable [11], i.e. for small perturbations in the initial data set the output would be significantly different. Using analyzing windows is necessary to introduce temporal localization of frequency components of the signal, but there is a drawback to this. The shape of the windowing function will determine the amount of false frequency components in the spectrum and it also influences the amplitude of the spectral components. This is why there exists such a wide range of window functions to accommodate all needs.

2.3.1 Application of the Short-Time Fourier Transform

A wide range of applications exist for the short-time Fourier transform in the audio signal processing domain. However, this transform can also be used in image processing.

The short-time Fourier transform can be used to enhance fingerprint images, as it is described in Chikkerur's paper [12]. The authors present an enhancement technique based on frequency domain filtering of the fingerprint image. The short-time Fourier transform is utilized first to get orientation and frequency maps of the original image. Just like the Fourier transform, the short-time Fourier transform can also be extended to higher dimensions. Equation 2.14 presents the 2D form of this transform.

$$X(\tau_1, \tau_2, \omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x, y) \overline{W}(x - \tau_1, y - \tau_2) e^{-j(\omega_1 x + \omega_2 y)} dxdy \tag{2.14}$$

where $\bar{W}$ is the complex conjugate of a window function, x and y are spatial variables, ω_1 and ω_2 are frequency variables and τ_1 and τ_2 are time variables of the two-dimensional spectrogram. In the case of this application, the window function is a raised cosine. After the short-time Fourier transform was carried out, the whole image can be modeled as a surface wave function:

$$I(x,y) = A(\cos(2\pi r(x\cos\theta + y\sin\theta))) \tag{2.15}$$

Here r and θ are the frequency and the ridge orientation in the image and A is a constant amplitude. They can be deduced from the short-time Fourier transform and they are considered random variables as they are defined by probability density functions. Thus the joint probability density function is defined by:

$$p(r,\theta) = \frac{|F(r,\theta)|^2}{\int_r \int_\theta |F(r,\theta)|^2} \tag{2.16}$$

where F represents the Fourier spectrum in polar form.

The marginal density functions for the frequency and the orientation are described by Eqs. 2.17 and 2.18.

$$p(r) = \int_\theta p(r,\theta)d\theta \tag{2.17}$$

$$p(\theta) = \int_r p(r,\theta)dr \tag{2.18}$$

Then the ridge orientation is computed as the expected value of the orientation variable:

$$E(\theta) = \frac{1}{2}\tan^{-1}\frac{\int_\theta p(\theta)\ \sin(2\theta)d\theta}{\int_\theta p(\theta)\ \cos(2\theta)d\theta} \tag{2.19}$$

This average is further smoothened using a Gaussian smoothing kernel.

The ridge frequency image is obtained by calculating the expected value of the frequency variable:

$$E(r) = \int_r p(r)rdr \tag{2.20}$$

This frequency image is also smoothened, but an isotropic diffusion smoothing is used to avoid errors being propagated from the edges of the image towards the middle of it.

In the next step an energy map is determined, used for differentiating areas which do not contain ridges and thus have very low energy from the short-time Fourier transform, from areas of interest. This energy map is then used as a basis of thresholding in order to get two different regions of the image. Equation 2.21 gives the definition of this energy based region mask.

$$E(x,y) = \log\left(\int_r \int_\theta |F(r,\theta)|^2\right) \tag{2.21}$$

To further reduce the discontinuities due to the block processing method, a coherence image is produced, which takes into consideration the level of orientation matching around individual points, as described by Eq. 2.22, (x_0, y_0) being the central point and (x_i, y_i) are the points overlapped by the W window. The values of this map will be high when the orientation of a block is similar to neighboring blocks' orientation.

$$C(x_0, y_0) = \frac{\sum_{(i,j)\in W} \cos\left(\theta(x_0, y_0) - \theta\left(x_i, y_j\right)\right)}{W \times W} \tag{2.22}$$

The actual image quality enhancement is then produced by applying algorithms described by Sherlock and Monro in [13].

2.4 The Wavelet Transform

The fixed width of the window used in the short time Fourier transform and the limited resolution of both the spectral and temporal domain are key reasons why the STFT cannot be used in many of the applications demanding time-frequency analysis. For example, non-stationary signal analysis depends heavily on determining what frequency components are present in the signal at a certain moment in time as well as on the possibility to search the signal for the occurrences of certain frequency components.

A new transform method was developed to face all these problems, called the wavelet transform. By definition, the wavelet transform is also an integral transform, using windows to slide over the analyzed signal in order to obtain time-frequency information from it [14, 15]. The major difference between this transform and the short time Fourier transform is that the wavelet transform uses windows which have variable width and amplitude, allowing the transform to analyze every bit of the signal. There are many different analyzing windows and they are called wavelets. Every wavelet is generally characterized by two parameters: scale (a) and translation (b). Equation 2.23 describes the continuous wavelet transform where a represents the scaling factor, b is the translation factor, ψ is a continuous function ($\psi \in L^2(\mathbb{R})$) with $\bar{\psi}$ being its complex conjugate and x is the analyzed signal in the time domain. ψ is also called the mother wavelet.

$$X_{a,b} = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\overline{\psi}\left(\frac{t-b}{a}\right) dt \tag{2.23}$$

Mother wavelets have zero mean and their square norm is one, as presented by Eqs. 2.24 and 2.25.

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \tag{2.24}$$

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1 \tag{2.25}$$

Although applications exist for the continuous wavelet transform (mostly analyzing theoretical aspects, see [16]), most of the problems necessitate a discrete wavelet transform. This can be derived from the continuous transform by quantization and without loosing any redundant information. Equation 2.26 defines the discrete scaling (s) and translation (u) factors used for the definition of wavelet functions.

$$s = 2^{-j}, u = k2^{-j}; j, k \in \mathbb{Z} \tag{2.26}$$

The result of substituting these variables in the integral transform is described by Eq. 2.27.

$$X(j,k) = 2^{j/2} \int_{-\infty}^{\infty} x(t)\psi\left(2^j t - k\right) dt \tag{2.27}$$

By discretizing the analyzed signal function, the integral transform becomes a sum as presented by Eq. 2.28.

$$X_{j,k} \approx 2^{j/2} \sum_n x_n \psi\left(2^j n - k\right) \tag{2.28}$$

The time-frequency resolution from wavelet decomposition point of view can be represented by the Heisenberg rectangle, where time and frequency is spread proportional to the scaling factor s and $\frac{1}{s}$. With the variation of s the two parameters of the rectangle, i.e. height and width, change accordingly but with a constraint stipulating that the area remains the same, as illustrated in Fig. 2.2. In most cases for a multiscale analysis, a scaling function φ is introduced. The relationship between the two functions φ and ψ is presented below by Eq. 2.29.

$$\left|\hat{\phi}(\omega)\right|^2 = \int_1^{\infty} \left|\hat{\psi}(s\omega)\right|^2 \frac{ds}{s} \tag{2.29}$$

With this notation the frequency analysis of the signal x at the scale of s is computed as follows:

$$X(u,s) = \langle x(t), \varphi_s(t-u) \rangle \tag{2.30}$$

Here $\langle \, . \, \rangle$ denotes the inner product of two functions and φ_s is given by:

$$\varphi_s(t) = \frac{1}{\sqrt{s}} \varphi\left(\frac{t}{s}\right) \tag{2.31}$$

In practice, the wavelet transform is computed for a defined number of scales (2^j). The low frequency component $W(u, 2^j)$ is often called the DC component of the

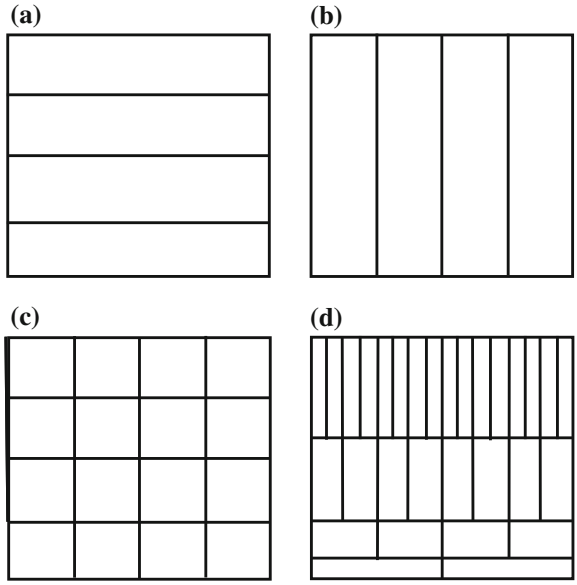

Fig. 2.2 Heisenberg rectangles for different transforms: **a** continuous domain (equal windows in the time direction, no frequency information); **b** Fourier transform (equal windows in the amplitude direction, no time information); **c** short-time Fourier transform (equal windows in both time and frequency direction); **d** Wavelet transform (rectangles with equal area but varying width and height)

signal. Wavelet coefficients W have a length of $\frac{N}{2^J}$ as the largest depth is bounded by the signal length.

For implementation purposes a set of conjugate mirror filters h and g is constructed using the scaling function φ and the wavelet function ψ. Their definitions are given by Eqs. 2.32 and 2.33.

$$h(n) = \langle \frac{1}{\sqrt{2}} \varphi\left(\frac{t}{2}\right), \varphi(t-n) \rangle \tag{2.32}$$

$$g(n) = \langle \frac{1}{\sqrt{2}} \psi\left(\frac{t}{2}\right), \varphi(t-n) \rangle \tag{2.33}$$

These filter functions have to satisfy the following conditions (here k denotes a filter function):

$$\left|\hat{k}(\omega)\right|^2 + \left|\hat{k}(\omega+\pi)\right|^2 = 2 \tag{2.34}$$

and

$$\hat{k}(0) = 2. \tag{2.35}$$

The discrete orthogonal wavelet decomposition can be calculated by applying the filter functions on the signal recursively. The two functions separate the signal into low and high frequency domains were h is a low-pass filter and g ia a high-pass filter, as presented in Fig. 2.3, while Fig. 2.4 shows the whole decomposition and

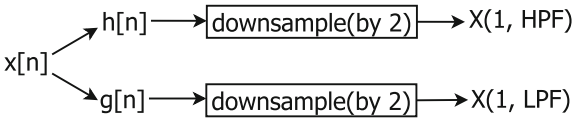

Fig. 2.3 One level of wavelet decomposition. LPF denotes a low-pass filter while HPF stands for a high-pass filter. Further decompositions are usually carried out using the results of the high pass filter

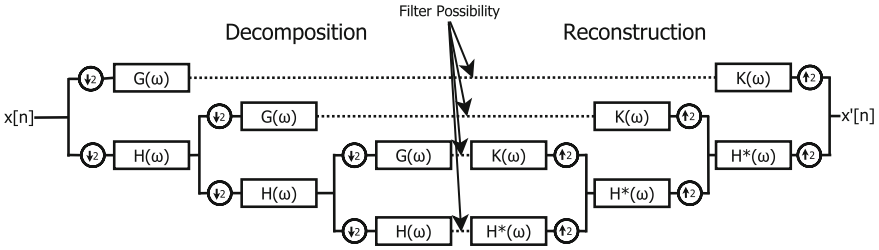

Fig. 2.4 Full wavelet decomposition and reconstruction scheme. Notice the down-sampler blocks on the decomposition side and the up-samplers on the reconstruction side

reconstruction scheme, highlighting the places of eventual filters. A detailed proof of this concept can be found in [15]. Another, more general approach can be achieved by using bi-orthogonal decomposition and reconstruction filters. It allows for a larger room for analyzing, modifying and filtering even multidimensional data.

Having a pair of wavelet function ψ and a reconstruction function χ, the decomposition and reconstruction is a straightforward implementation with H, G and K filters satisfying the Eq. 2.36 [17].

$$\begin{aligned}
\hat{\varphi}(2\omega) &= e^{-i\omega s} H(\omega)\hat{\varphi}(\omega) \\
\hat{\psi}(2\omega) &= e^{-i\omega s} G(\omega)\hat{\psi}(\omega) \\
\hat{\chi}(2\omega) &= e^{i\omega s} K(\omega)\hat{\chi}(\omega) \\
|H(\omega)|^2 + G(\omega)K(\omega) &= 1
\end{aligned} \tag{2.36}$$

As it turns out, filtering a signal is equal to applying a simple scaling function after the transformation. By modifying the wavelet coefficients after the decomposition, we can attenuate or amplify different frequency bands according to the resolution. This can be done even for a defined time-period depending on the scaling function.

2.4.1 Application of the Wavelet Transform

One of the most well-known examples of using the wavelet transform is the JPEG-2000 image compression method. The widely used JPEG image compression

format traditionally uses the discrete cosine transform and has limitations both in compression efficiency and the quality of edge preserving. In order to overcome these shortcomings, the JPEG-2000 standard was developed, which uses the discrete wavelet transform in order to decompose the image and then compress it.

As a first step, the image is cut in rectangular tiles and all the following operations are executed on each tile separately. Each tile is transformed using the one dimensional discrete wavelet transform, so 4 results are obtained: one with low-resolution rows and columns, one with low resolution rows and high resolution columns, one with high resolution rows and low resolution columns and one with high resolution rows and columns. This decomposition is then repeated a number of times on the low resolution image block. The decomposition process is called dyadic decomposition. Both lossless and lossy decompositions are possible: when using integer wavelet functions, the result is lossless decomposition, while using real-valued wavelet functions, the result is a lossy decomposition. In both cases the decomposition is done by applying low-pass and high-pass filters built up using wavelet functions. On the reconstruction side quadrature mirror filter pairs [1] of the decomposition side filters are used.

After the decomposition, the obtained coefficient values are quantized, i.e. this is the step where information loss can occur. When using lossless compression, though, there is no quantization, or the quantization step is 1.0. Following quantization, entropy encoding of the bit planes is carried out. After this procedure is done, the image is ready to be stored [18].

2.5 The Hilbert-Huang Transform

2.5.1 Introducing the Instantaneous Frequency

All of the transform methods presented up until now use the classical definition of frequency, i.e. the number of repetitions of a phenomenon in unit time. This means that intra-wave oscillations are totally disregarded when analyzing non-stationary signals. Already in the 1990s it was recognized that there are serious consequences of not taking into consideration these intra-wave changes in frequency.

The basic idea behind the instantaneous frequency is very simple. For example, if we consider a uniform circular motion, then the projection of this movement on one of the x or y axes will result in one of the Fourier series, more accurately a single sine or cosine wave. In this case the instantaneous frequency is constant. The speed of the motion is described as the derivative of the angular position and it is called angular velocity. This angular velocity is exactly the definition of the instantaneous frequency. If the motion is uniform, i.e. has a constant speed, the angular velocity will be constant, thus the frequency will be constant in each time point. However, if the motion is non-uniform, the sinusoid will have time-variable periods. This way, the definition of the instantaneous frequency becomes a generalization of the classical frequency model.

From now on, we can treat every signal as a sum of non-uniform circular motions with varying amplitude where the derivative of the angular position is the instantaneous frequency. Analyzing data from this perspective exceeds the breakdown capability of the Fourier transform as we get an absolute time depending spectral overview.

It was, however, not until recently when the computation of the instantaneous frequency was made easily achievable and Huang [19] contributed significantly to this by developing the Hilbert-Huang transform. Although the new definition of frequency was also known before Huang [20], there were no straightforward algorithms for its computation.

2.5.2 *Computing the Instantaneous Frequency*

There are many ways of computing the instantaneous frequency of a non-stationary signal. The basic idea is to decompose the original signal into mono-components which have only one oscillation per period and then computing the instantaneous frequency of these components by determining the quadrature of the component and computing the derivative of the phase of the two sub-signals. As we are talking about empirical methods to decompose a signal into mono-components, there is a high risk to end up with elements without any link to the original physical phenomenon. A minimal change in the signal form will result in a totally different decomposition.

The Hilbert-Huang transform is one of the modalities to calculate the instantaneous frequency. It consists of two steps:

1. Determine the mono-components of the original signal. This step uses the empirical mode decomposition algorithm
2. Compute the instantaneous frequency with the Hilbert transform for every mono-component.

The empirical mode decomposition (EMD) is an iterative algorithm that extracts mono-components called intrinsic mode functions (IMF) from the original signal. These IMFs hold two important properties:

1. They have the same number of local maxima and minima, or their number differs at most in one;
2. Their upper and lower envelope averages to zero.

Because it is an iterative algorithm, the exact number of final IMFs is unknown beforehand. The algorithm is described by listing 2.1. The mono-components or intrinsic mode functions can be described by the general equation of a non-uniform, variable-amplitude circular motion:

Algorithm 2.1: Pseudocode of the EMD algorithm

```
s ← getSignal();
ε ← 10^−5;
k ← 1;
while !isMonotonic(s) do
    SD_k ← 1;
    h_{k,0} ← s;
    i ← 0;
    while SD_k > ε do
        max ← getLocalMaxima(s);
        min ← getLocalMinima(s);
        upperEnvelope ← getSpline(maxima);
        lowerEnvelope ← getSpline(minima);
        m_i ← (upperEnvelope + lowerEnvelope) / 2;
        h_{k,i} ← h_{k,i−1} − m_i;
        SD_k ← Σ(h_{k,i−1} − h_{k,i})^2 / Σ h_{k,i−1}^2;
    c_k ← h_{k,i}; // the k-th IMF
    s ← s − c_k;
    k ← k + 1;
r ← s; // the residue
```

$$c(t) = A(t)\cos(\varphi(t)) \tag{2.37}$$

where $c(t)$ is the IMF, $A(t)$ describes the time-variable amplitude part and $\varphi(t)$ represents the time-variable phase.

Due to their fundamental properties, it is easy to find their quadrature and the Hilbert transform does exactly that. Taking the IMF $c(t)$, the corresponding quadrature $q(t)$ is given by Eq. 2.38.

$$q(t) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{c(\tau)}{t-\tau}d\tau \tag{2.38}$$

Then the phase difference between them is computed:

$$\varphi(t) = \arctan\frac{q(t)}{c(t)} \tag{2.39}$$

Then the instantaneous frequency of the IMF will be:

$$\omega(t) = \frac{d\varphi(t)}{dt} \tag{2.40}$$

Although this transform is a powerful tool for calculating the instantaneous frequency of an IMF, there are some limitations imposed by the Bedrosian and Nuttal theorems. According to the Bedrosian theorem [21, 22], the IMFs should satisfy the condition described by Eq. 2.41 in order for the analytical signal provided by the Hilbert transform to make sense.

$$H\{A(t)\cos(\varphi(t))\} = A(t)H\{\cos(\varphi(t))\} \tag{2.41}$$

This condition states that the IMFs should be band limited, i.e. the spectrum of the time-variable amplitude and the one of the variable phase should not overlap [23]. Failure to conform to this requirement will result in the confusion of amplitude frequencies in the instantaneous frequency of the IMF.

The Nuttall theorem [24] further states that the Hilbert transform of $\cos(\varphi(t))$ is not $\sin(\varphi(t))$ for any arbitrary $\varphi(t)$. Therefore, the result of the Hilbert transform is not necessarily equal to the true quadrature of the IMF. Because there are no further restrictions on what an IMF can be, except the ones presented before, using the Hilbert transform to get the instantaneous frequency of an IMF is considered unsafe. In practice, the instability of the Hilbert transform can be easily spotted in regions where the instantaneous frequency becomes negative or a sudden peak appears in a region where there are no sudden frequency changes in the IMF.

A more accurate way to compute the instantaneous frequency of an IMF is to use the empirical AM-FM decomposition algorithm [23]. This algorithm is also an iterative one, just like the EMD and it is presented in listing 2.2. The aim of this algorithm is to separate the two parts of the IMF, i.e. the amplitude part from the phase part. The model used by this method is that of modulated signals: the amplitude part is considered to be the modulator signal and the phase part the carrier; this is considered the AM modulation in the IMF. The carrier is also modulated in frequency; this constitutes the FM modulation in the IMF. The decomposition results in two signals: $AM(t)$, the amplitude part and $FM(t)$, the frequency part.

Algorithm 2.2: Pseudocode of the empirical AM-FM decomposition

```
1 s ← getImf();
2 s_0 ← s;
3 k ← 1;
4 while s_k ∈ [−1;1] do
5     e ← getSplineEnvelope(|s_{k−1}|);
6     s_k ← s_{k−1}/e;
7 FM ← s_k;
8 AM ← s/FM;
```

The separation process implies successive approximations of the amplitude part by cubic spline functions through the local maxima of the IMF signal. After the two signals are separated, the quadrature of the frequency part is computed:

$$Q(t) = \pm\sqrt{1 - FM^2(t)} \tag{2.42}$$

From here, getting the instantaneous frequency of the IMF matches the Hilbert transform-based method.

It is important to note that this method satisfies the conditions imposed by both the Bedrosian and the Nuttall theorems. However, the method is not free of disadvantages. The cubic spline envelopes often cross the IMF, meaning that they are not always greater or equal than the signal, which is a necessary condition for the algorithm to converge in finite steps. Instead, the method will never converge and a frequency part belonging to the $[-1; 1]$ interval will never be reached. This problem and a possible solution to it is addressed in [25].

First, the intervals where the spline crosses the IMF are identified. These intervals are characterized by an entry and an exit point. The spline is approximated with a line segment over these intervals and the maximum distance between this approximation and the IMF is sought. To do this, the segment and the IMF are rotated to the horizontal axis and the maximum of the IMF is identified. After an inverse geometric transform a new spline is defined, which contains the maximal distance point but not the previous extremum point, thus not intersecting the IMF anymore. These steps are repeated for each interval, until the spline only touches the IMF but does not cross it. The correction mechanism is illustrated in Fig. 2.5 and it is described as a pseudocode in listing 2.3.

Algorithm 2.3: Pseudocode of the empirical AM-FM decomposition

```
1  s ← getImf();
2  s_0 ← s;
3  k ← 1;
4  while s_k ∉ [−1;1] do
5      e ← getSplineEnvelope(|s_{k−1}|);
6      intervalList ← getCrossingIntervals(s_{k−1}, e);
7      while !isempty(intervalList) do
8          p ← getMaximumDistancePoint(s_{k−1}, intervalList_i);
9          e ← getSplineEnvelope(|s_{k−1}|, p);
10         intervalList ← getCrossingIntervals(s_{k−1}, e);
11     s_k ← s_{k−1}/e;
12 FM ← s_k;
13 AM ← s/FM;
```

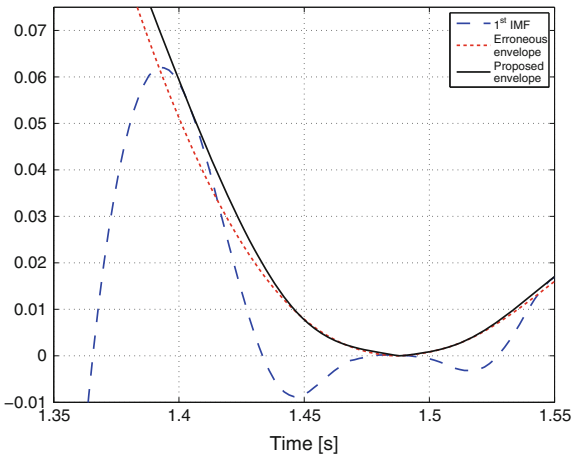

Fig. 2.5 A slice created by the intersection of the spline and the IMF and the corrected spline

2.5.3 Application of the Hilbert-Huang Transform

2.5.3.1 Detecting the Third and Fourth Heart Sounds in Phonocardiograms

The Hilbert-Huang transform is an ideal tool for analyzing the heart activity since it offers frequency information for every sample point. In the following, a detection of the third and fourth heart sound will be presented through the usage of HHT. The method is described in [26]. These sounds (S3, S4) represent an abnormal activity of the heart indicating failure during diastolic period. Processing sound measurements, i.e. identifying S3, S4 is a noninvasive way to detect early myocardial ischemia. Deficient S3 is related with ventricle activity or problematic blood flow in the rapid filling phase. The fourth heart sound, S4 is located before the first heart sound and is a sign of forceful contractions in an effort to overcome a rigid ventricle. The method localizes the presence of S3, S4 hearts sounds using time-frequency distribution and K-means clustering in electronic stethoscope recordings. The proposed method can be divided into three stages. The first step preprocesses the signal, the second one applies the Hilbert-Huang transform to get a time-magnitude-frequency decomposition followed by a clustering and recognition procedure.

The role of preprocessing is to eliminate high frequency noise due to recording artifacts by a low-pass, finite impulse response filter. Then an envelope is produced using the Hilbert transform noted as $x_{envelope}[n]$ (Eq. 2.43).

$$x_{envelope}(n) = |x(n) + jH\{x(n)\}| \tag{2.43}$$

This envelope is then normalized (Eq. 2.44) to fit the interval $[-1;1]$ and segmented into systolic (S1-S2) and diastolic (S2-S1) periods.

$$x_{norm} = \frac{x_{envelope}(n)}{\max(x_{envelope}(n))} \tag{2.44}$$

The detection of these terms is based on the Shannon energy as it shows a high rate of noise rejection property, suppressing low amplitude components. Because of the main components S1 and S2 will stand out, detection and, therefore, segmentation is much easier. Equation 2.45 defines how the Shannon energy is computed.

$$SE(n) = -x^2_{norm}(n)\log(x^2_{norm}(n)) \tag{2.45}$$

A threshold of 70 % of the maximum Shannon energy is set to differentiate S1 and S2 from other components. Naturally, a higher limit will result in a better filtering but with the risk of missing some S2 sounds. Recognition is built upon the following rules to ensure a correct result:

- If two peaks higher than the threshold are detected within 50 ms, the one with lower energy is eliminated.
- For every interval between the peaks, an interval with shorter length than the previous interval is denoted as a systolic period, while the other one is a diastolic period. The uncertain intervals are annotated.
- For those uncertain intervals, a secondary threshold is set to find S1 or S2 which probably have not been recognized.

As a second step the Hilbert-Huang transformation is applied to the nonlinear signal, extracting the instantaneous frequencies. It is an iterative empirical procedure resulting in a series of intrinsic mode functions and therefore a sum of instantaneous frequencies is calculated for a single time instance (Hilbert-Huang spectrum).

Having a time-frequency map allows for a cluster figure based on pairs of instantaneous frequencies and their amplitudes. Correlating in-scope cluster points with the time scale will reveal the position of S3 and S4 sounds. For the cluster graph the instantaneous frequency with the highest magnitude is selected from the sum of instantaneous frequencies for a particular time instance. Only one with the highest impact factor is selected based on the magnitude. The number of points depends on the resolution of the Hilbert spectrum. These pairs originate from the diastolic interval as it was segmented in the second part of the method. The K-means algorithm was used to divide pairs into different groups. Three separate clusters have been formed depending on the frequency and magnitude distribution. The one of interest is called the abnormal group with the highest frequencies or magnitudes. Usually this group has the fewest of points and their projections on the time axis is periodical. If the points appear right before S1 we found the S4 sound but if they appear periodically after S2 we can identify S3 elements. For those possibly missing components, an iterative method was applied to enhance accuracy.

Records from the Cardiac Auscultatory Recording Database (CARD) of Johns Hopkins University were used to test the proposed method. As a result, 13 signals were processed. The method with iterative recognition, achieved a 90.3 % identification rate for S3, 9.6 % were missed and 9.6 % were marked falsely. For S4, 94.4 % were detected correctly and 5.5 % were missed; 16 % were false positive. Sensitivity for S3 and S4 was 90.4 % and 94.5 %, respectively. Precision was 90.4 and 85.5 %.

Still, the existence of artifacts during diastolic period, such as diastolic murmur or noise produced by the electronic stethoscope, would contribute to misjudgments. Components with low amplitude make the separation of S3 and S4 from background noise difficult. The proposed method aims to detect early heart diseases, such as left ventricle dysfunction, congenital heart failure or myocardial ischemia.

2.5.3.2 Identifying T Waves in ECG Signals

The empirical mode decomposition can also be used in ECG signal analysis. Identifying T waves in an ECG is of special interest, since its morphology can suggest different heart conditions. Finding the positions of these waves is often done manually, thus an automated detector is much needed. In [27] we propose a method that can solve this problem for a wide range of ECG sources.

The general idea behind the method is to identify and remove QRS complexes from the ECG signal, because these are present in most of the IMFs owing to their large amplitude and wide frequency spectrum. A method described in [28] is employed for this task. According to this method, after the ECG is decomposed in IMFs, the Shannon energy of the sum of the first three IMFs is computed, as described by Eqs. 2.46 and 2.47.

$$c(t) = \sum_{k=1}^{3} IMF_i(t) \tag{2.46}$$

$$E_s(t) = -c(t)^2 \log(c(t)^2) \tag{2.47}$$

Then, using a threshold, the Shannon energy (E_s) is filtered to eliminate background noise. The threshold is determined by Eq. 2.48 where N is the number of data points in the Shannon energy.

$$T = \frac{1}{N} \sum_{i=1}^{N} E_s(i) \tag{2.48}$$

The threshold filter substitutes each value less than the threshold with zero and keeps all the other values. Non-zero intervals correspond to the QRS complex intervals of the ECG signal. After identifying each of these intervals, the corresponding places in

the original ECG are exchanged with a line segment and the signal is decomposed again using EMD.

In the newly obtained IMF set the 5th, 6th and 7th will contain waves that construct the T wave, thus they are summed. Finally, the peaks in this newly calculated sum appearing right after the positions of QRS complexes are marked as T waves in the original signal. Figure 2.6 presents the steps of the identification process.

Signals from PhysioNet's QT database [29] were used to test the algorithm. This testing resulted in a 95.99 % positive predictability and 99.96 % sensitivity when

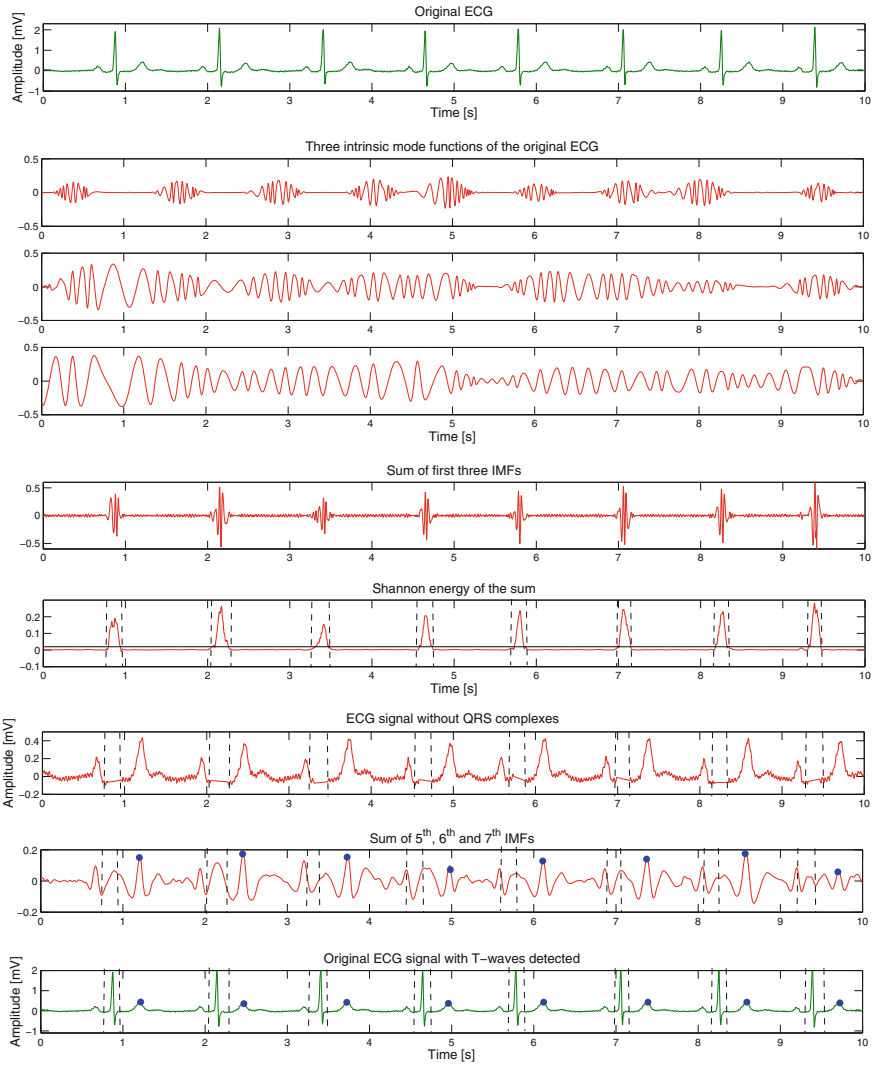

Fig. 2.6 Graphical representation of the T wave identification process

different ECG signals were presented to it. Noise tolerance was also tested by feeding an ECG signal with different amplitudes of noise to the method. An 87.71 % positive predictability was obtained for a signal to noise ratio as low as 12 dB. The corresponding sensitivity was of 100 %.

2.6 Hybrid Signal Processing Systems

2.6.1 The Discrete Wavelet Transform and Fuzzy C-Means Clustering

Medical signal processing represents an essential part of a medical decision making system, however it is able to provide only signal conditioning and feature extraction without offering any information on the diseases that are reflected in the acquired signal. Getting information about any condition is basically a classification problem, and thus it can be addressed by a variety of artificial intelligence methods. Therefore, it makes sense to combine artificial intelligence methods with well-known signal processing methods to assist a medical decision making system.

A classical but also very important application area of medical decision making systems is electrocardiogram (ECG) interpretation. ECG signal processing is a much discussed topic. There are different ways of processing these signals depending on what features are being sought for. Thus, there are methods using classical Fourier analysis [30], time frequency analysis by making use of the wavelet transform [31], the Hilbert-Huang transform [32, 33], the Wigner-Ville distribution [34] or by using a multivariate signal analysis such as the independent component analysis [35] or the principal component analysis [36]. Support vector machine based analysis has also gained popularity in the past years [37, 38].

A very good example of combining a classical signal processing method with an intelligent classification method is presented in [39] and describes a method of ECG signal classification using a combination of the wavelet transform and fuzzy c-means clustering. The main goal of said paper is to describe a way to implement the classification system on a mobile device.

The presented approach makes use of the discrete wavelet transform, which provides discrete wavelet coefficients that describe the signal. To achieve these coefficients, the finite-length signal is convolved with filters that separate the signal into low and high frequency sub-signals. The signal can be reconstructed with the use of quadrature mirror filters from the coefficient sets.

The artificial intelligence algorithm used in this article is the fuzzy c-means clustering. The idea of this method is to do a fuzzy partitioning of the data into classes. In fact, the algorithm can be reduced to a constrained optimization problem. The cost function to be optimized is:

$$J_\beta(z) = \sum_{i=1}^{N} \sum_{j=1}^{c} A_{ij}^{\beta} \left\| x_i - z_j \right\|^2 \tag{2.49}$$

In this equation, J_β is the cost function, β is the level of fuzziness, chosen to be 1.5 in this specific application, N is the number of data vectors, c represents the number of clusters, A_{ij} is the grade of membership of data vector i to cluster j, x_i is the ith data vector and z_j is the center of the jth cluster. The grades of membership are computed as Euclidean distances of data vectors from the center of the cluster. The detailed fuzzy c-means algorithm is described in [40].

Figure 2.7 presents the components of the diagnosing system. The ECG signals collected by body sensors are first preprocessed, since they are contaminated by both low and high frequency noise. The baseline drift represents a low frequency component in the signal, while thermal noise is responsible for the high frequency noise. Both artifacts are removed using the discrete wavelet transform, applying the Daubechies-9 wavelet for the signal decomposition. By carefully combining some of the signal components and then reconstructing the signal, the noises can be filtered out. The signals were also normalized to lay between −1 mV and 1 mV. This normalization is also convenient if at a later stage support for different body sensors will be included in the overall system. This step was necessary because in the testing phase the ECG signals from the MIT-BIH database were used, which were recorded by a large variety of devices from all over the world.

The next stage consisted of extracting different features of the signals. In total 6 characteristics were chosen, as follows: the length of a heartbeat period, called the RR period, the length of the QRS complex relative to the RR length, the length of the RS interval relative to the QRS length, the period of the ST segment relative to the RR length, the amplitude of the QRS complex and the amplitude of the T wave. Measuring these parameter is done also by using the DWT and combining components so that the QRS complex and the P-T-U waves could be separated. An ECG time-series was thus characterized by a feature vector having 6 elements corresponding to the measurements of the 6 aforementioned parameters.

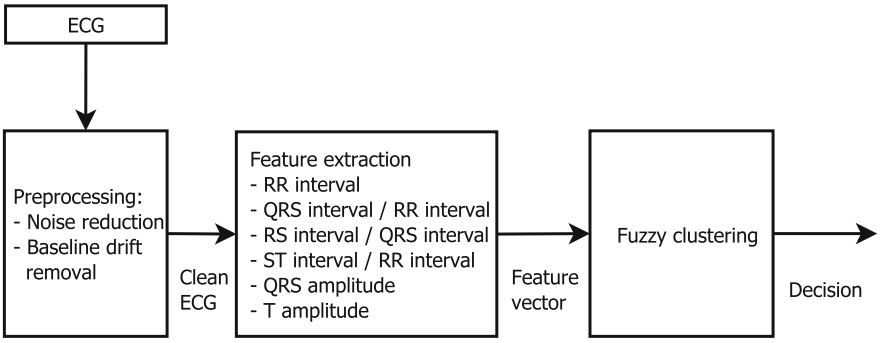

Fig. 2.7 The general architecture of the classification system

In the last stage of the classification algorithm the feature vectors are separated into clusters by the fuzzy c-means algorithm, using a fuzziness number $\beta = 1.5$ and 3 clusters. Two ECG data sets from the MIT-BIH database were used. The first one contained 25 ECG signals while the other one had 23 ECGs. Portions with a length of 10 s were selected from each signal and their feature vectors were fed to the fuzzy c-means clustering algorithm. Feature vectors corresponding to normal ECGs were classified into one cluster and abnormal ones in the other two clusters. There were, however mismatches too. 3 out of 25 ECGs from the first group and 6 out of 23 ECGs from the second group were misclassified.

The diagnostic system was implemented on a mobile device running Windows CE having an ARM9 processor. The wavelet decomposition and fuzzy c-means clustering were implemented in the Matlab environment and then C++ code was generated using Code Generator. According to the article, a nearly real-time execution was made possible on the mobile device.

2.6.2 Automatic Sleep Stage Classification

The correct diagnosis of sleep stages is crucial when it comes to identifying and treating possible sleep apnea, insomnia or narcolepsy conditions. Even today much of this work is done visually by experts, based on polysomnogram (PSG) techniques described by the American Academy of Sleep Medicine (AASM) [41] or by the guidelines defined by Rechtschaffen's and Kales's [42]. Basically, most specialists try to identify 6 repeating sleep stages from electroencephalogram (EEG) measurements combined with electrooculograms (EOG) and electromyogram (EMG) signals to enhance accuracy. The initial stage is characterized by being awake (Awa) followed by S1, transition between wakefulness and sleep. S2 is considered to be the baseline; it may consist of 45–55 % of the entire sleep duration. Stage three and four (S3, S4) represent the recovery mode of the body and mind also known as deep-sleep period followed by the rapid eye movement (REM) stage. From this perspective, two major phases can be distinguished when it comes to sleep, REM and non-REM (NREM), where S1, S2, S3 and S4 are sub-divisions of the NREM phase. These stages repeat themselves 4–5 times during one night.

The EMG activity is almost missing during REM phase. This helps to differentiate S1 from REM using EMG measurements. Similarly EOG is useful when it comes to eye movement detection in S1 and REM sleep. In the following, different methods will be presented, all developed to make the evaluation of sleep signals as automatic as possible. The task to design such a system usually takes into consideration the amount of data and the complexity of classification algorithms. These two parameters greatly influence the overall behavior of the end result. For instance, multichannel EEG devices represent a drawback to patient comfort and in ambulatory environment compared to single channel devices. Therefore recent studies aim to develop methods that use only one measurement to detect sleep stages [43–49]. Another challenge represents the right choice and combination of different

feature extraction methods working together with pattern recognition systems. A great variety has been proposed over the years with different and surprising outcomes. For instance Hidden Markov Models (HMM), fuzzy classifiers, different types of artificial neural networks (ANN) using feature vectors based on power spectral densities (PSD), wavelet decompositions or visibility graphs (VG). Still, it is difficult to achieve a higher accuracy like the ones produced by experts using manual techniques. However, in [50] several approaches are proposed for the discussed problem through a particular neural network, financial forecast based method, non-smooth optimization problem and frequency domain analysis. The ANN approach suggests the usage of a time-delay neural network (TDNN) with the property of reducing input data size, an important element when it comes to processing EEG, EMG and EOG simultaneously. A particular output from such a network depends not only on the input but it takes into consideration a range of the previous input values. The benefit of this particularity makes it possible to reduce the input-volume as previous information is already stored in the network. In this study the network was configured with 1 input layer, 3 hidden layers and 1 output layer. Normalized PSG variables were fed as input: direct EEG, EMG and EOG measurements. The output layer consists of 6 nodes corresponding to the 6 sleep stages. As a result, the network was able to obtain a classification accuracy of 76.15 % from thousands of test cases.

A similar model proposed by [51] uses a multi-layer perceptron (MLP), a neural network designed for classification. The network had 5 input nodes and again 6 output nodes for the 6 sleep-stages. Electrode distribution followed the 10–20 system but only the C3-A2 left-right and C4-A1 right-left measurements were used. A well known fact in EEG signal processing is related to the frequency bands that reflect the mind's different states. There are usually 5 frequency bands related to mental activity: δ (<4 Hz), θ (4–7 Hz), α (8–12 Hz), σ (13–16 Hz) and β (>16 Hz). Usually the short time Fourier transform is used to gain an overview in the time-frequency domain. Due to the non-linearly of these signals, the power spectra is used in this case. The proposed method calculates the relative spectral power (RSP) for a window of 30 s, equal to the band spectral power (BSP) divided to the total spectral power (TSP).

$$RSP_i = \frac{BSP_i}{TSP}, i \in \{\delta, \theta, \alpha, \sigma, \beta\} \tag{2.50}$$

After training the network and finding the optimal structure the method was able to recognize sleep states with an accuracy of 76 %. The final conclusion states that Awa, S2, S4 and REM stages are easily recognized, however S1 is confused with S2 and REM, being the hardest stage to identify. By adding the second derivative of the EEG signal to the input vector the overall performance of the system did not improve but remained in the 76–77 % range.

ANN combined with wavelet packet coefficients was suggested in [52] where they used a 3-layer feed forward perceptron with 12 input nodes, 1 hidden layer (with 8 neurons for best results) and 1 output layer (with 4 nodes) as a

classifier. The goal of the perceptron was to distinguish between sleep stages Awa, S1 + REM, S2 and SWS also known as deep-sleep stage. Adaptive learning rate has been applied to avoid stagnation on the error surface. This translates to an adaptive learning step solution minimizing the learning period. For testing purposes PhysioBank's EEG Database was used, more precisely Pz-Oz bipolar recordings. Wavelet packets had been chosen as a feature extraction method, allowing for a finer frequency resolution. After fine-tuning the packet tree, the structure shown in Fig. 2.8 was defined as the transformation method supplying feature vectors to the perceptron. The subbands represent the following EEG frequency bands:

1. Delta—0.39–3.13 Hz
2. Theta—3.13–8.46 Hz
3. Alpha—8.46–10.93 Hz
4. Spindle—10.93–15.63 Hz
5. Beta1—15.63–21.88 Hz
6. Beta2—21.88–37.50 Hz

However, the 6 wavelet coefficient groups were further refined to a 5 element classification series as a statistical time-frequency distribution representation. The first element holds the mean quadratic value for each of the 6 bands. The second element is the total energy, followed by multiple elements calculated as the ratio of different energy bands (alpha, delta and theta). The fourth and fifth elements are the mean of the absolute values and the standard deviation of the coefficients in each sub-band. The 5 element series fed to the MLP resulted in a very high classification rate, indicating that the method could discriminate between Awa, S1 + REM, S2 and SWS with a specificity of 94.4 $\pm$ 4.5 %, a sensitivity of 84.2 $\pm$ 3.9 % and an accuracy of 93.0 $\pm$ 4.0 %.

Another feature extraction method uses financial forecasting to predict sleep stages based on the hypothesis that a variable from the sleep measurement depends on the previous values. It relies on conditional probability to foretell the next stage. Raw date is mapped into a 5 symbol series depending on preset derivative thresholds. The new values are tagged as follows: BI (big increase), SI (small

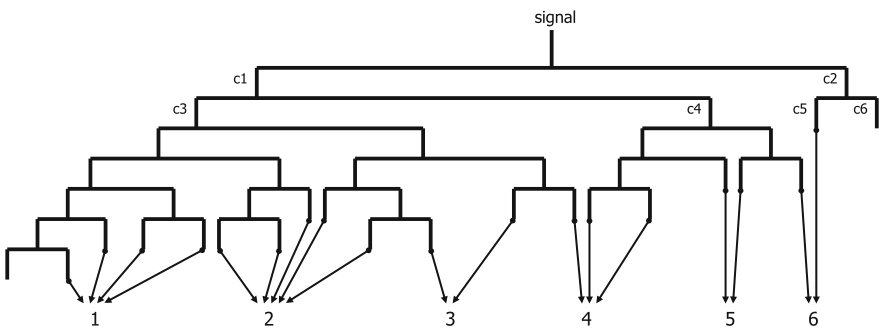

Fig. 2.8 Wavelet packet transform and selected subbands

increase), N (no change), SD (small decrease) and BD (big decrease). After the tagging the method looks for particular sequences like BI, BI, BI, SI, N, SD and SI defining one of the sleep stages. As a result, this approach finds it hard to classify S2 correctly; on the other hand its simplicity is a great advantage with an overall accuracy of above 70 % compared again to expert classification structures.

Non-smoothing optimization is a relatively new extraction method having its roots in the field of signal processing. The main idea behind it is simple, where the algorithm tries to fit a sum of two sine waves on the raw EEG. The deviation from the sine waves represents the information itself. In this case the amplitude is a linear function resulting in a more adequate curve fitting procedure then with the ordinary scalar value. This way sudden changes can be tracked maintaining vital information of the signal. The sum is composed of two sine curves, a low frequency component tracking the baseline wandering and a high frequency component being able to adapt to all kind of shapes.

As a summary to the sleep-stage classification methods we can add that even manual scoring done by experts show differences. Comparisons between results can score below 80 %. This means that the actual automatic sleep stage classification methods are as reliably as the experts.

2.6.3 The Hilbert-Huang Transform and Support Vector Machines

In recent years the car industry has produced a whole new generation of vehicles, more reliable, safer, and with build-in-intelligence to recognize and anticipate different type of engine failures. Wang et al. present in [53] a method of failure detection based on the Hilbert-Huang transform and support vector machines. Based on the failure source, the engine fault diagnosis (EFD) technology defines 4 types of analysis methods: engine performance detection, lubricating oil analysis, vibration based methods, and noise based diagnostic methods. The vibration and noise based EFD methods normally extract the failure features from the vibration or noise signals of a running engine and make decisions based on diagnostic results by using pattern recognition algorithms. Usually such a system is composed of two larger elements. One is responsible for processing the measured signals and supplying useful features, called a feature extractor and the second element is a classifier. The classifier's role is to separate the different engine failures into categories based on feature types.

Noises from a car's engine are non-stationary under working conditions, for this reason the Hilbert-Huang transform presents itself as an ideal candidate supplying time variant mono-components related to circular motions. For the classification algorithm support vector machine (SVM) [54] has been used as a multi-category classifier based on "one against one" voting decision method. Such a model will define an optimal hyper-plane which separates fault types by mapping one group of

feature vectors (belonging to a particular engine fault) on one side of the plane and other groups on different sides thus categorizing them geometrically. Basically, for n categories $\frac{n(n-1)}{2}$ SVM models are constructed from a training database for every two differing category. A measurement sample will pass every model and in the end it will be ranked upon the received votes. The final category is the one with the largest number of votes.

For the analysis it is presumed that in fault conditions, particular engine noises will reflect in the amplitude and frequency domain and certain failures will shift the energy quantum from one frequency band to another. Therefore the energy pattern of the intrinsic mode function is considered to be at the heart of the proposed feature extraction method.

The process is performed in 4 steps. First the sound is filtered to remove unwanted noise, and then the IMFs are calculated together with the residual element. Usually the residual and particular IMFs are discarded as they carry almost no energy. After selecting the IMFs, a correlation coefficient is calculated between the original signal and every IMF with Eq. 2.51.

$$\rho_{S,c_i} = \left| \frac{E[(c_i(t) - \mu_{c_i})(S(t) - \mu_S)]}{\sigma_{c_i}\sigma_S} \right| \tag{2.51}$$

Here S represents the original measurement, c_i are the IMFs while μ_{c_i}, σ_{c_i} and μ_S, σ_S are the mean values and the standard deviations of these signals. Every correlation coefficient falls between 0 and 1, where 1 means that the two signals are identical and 0 means that they are totally different. A large value means that the two signals have much in common, giving the method a chance to choose between the IMFs. Usually mono-components with a low correlation coefficient are discarded, therefore only IMFs with relevant information content are processed further. The third step computes the energy moment of the selected IMFs, denoted with E_i.

$$E_i = \sum_{k=1}^{n} \left[(k\Delta t)|c_i(k\Delta t)|^2 \right] \tag{2.52}$$

As a last step a feature vector is constructed from the energy moments.

$$T_E = [E_1, E_2, \ldots, E_7] \tag{2.53}$$

These vectors also describe the energy distribution among the IMFs and reflect their change through time. To further enhance the vector's efficiency to track changes, the marginal spectrum of the IMFS is added to them as the maximum amplitude A_0 and the corresponding instantaneous frequency f_0:

$$T_E = [E_1, E_2, \ldots, E_7, A_0, f_0] \tag{2.54}$$

The differences between features are then used to distinguish between the states of the engine, including failures and normal regimes, making it possible to diagnose engine faults.

The proposed method was tested using a sample car with a fault generator reproducing valves, pistons, rods, crankshaft, timing belt and structure vibrations. A microphone was placed over the engine to record its operation at 2500 rotations per minute. The simulated 7 engine faults are: normal state without faults, a break or short in the circuit of throttle threshold sensor, break or short in the circuit of ignition coil, a break or short in the circuit of the Hall sensor, basic set errors in the throttle control unit, defect in the circuit of the first cylinder injector and defect in the circuit of the third cylinder injector. Every state was recorded 20 times, totaling in 140 signals which were used to train and validate the SVM models through 7 and 9 dimensional feature vectors. An iterative training algorithm has been used to train the SVM models by minimizing an error function. Based on the error function's form, two model types can be distinguished: C-SVM and nu-SVM. However results show that the C-SVM model outperforms the nu-SVM regardless of the dimensions of the feature vectors.

For testing purposes 70 measurements were used to evaluate the proposed method. Measurements included faulty and normal behavior and both 9 and 7 dimensional feature vectors were used to see which model is best suited for EFD. For a total of 56 engine states the 7 dimensional feature vector model classified correctly 80 % of the states, where the one using a 9 dimensional model was able to predict 91.43 % of it right. The difference between the accuracy can be acknowledged to the extra information carried by A_0 and f_0. The outcome is greatly influenced by the model's ability to distinguish between states, however faults 4 and 5 are sometimes confused with state 1 and fault 2. This can be attributed to the fact that faults 2 and 5 represent throttle problems which are hard to distinguish only by engine noise and fault 4 is a failure of the Hall sensor which has almost no effect on the acoustic properties of the engine, easily mistaken with the normal state 1. As a conclusion, the paper states that the noise based HHT SVM method can't be applied to all types of engine failures. Furthermore to enhance precision, a 5 state classifier has been proposed (omitting faults 4 and 5) which resulted in 96 % accuracy. To validate this new model a second set of measurements has been acquired resulting in a test set of 50 states. Again, the model recognized every state with an accuracy of 94–96 %. Thus, the HHT based 9 dimensional feature vectors together with multi-class SVMs for pattern recognition can be used to design noise based EFD with accuracy above 90 %. The approach itself can be extended to be used in different fields of engineering, i.e. machinery diagnostics, speech and image recognition.

2.7 Conclusions

We have presented both classical and new signal processing and artificial intelligence methods in the context of complex digital signal processing and feature extraction. Signal processing methods were revisited in a historical order, from the Fourier transform to the Hilbert-Huang transform, covering the whole interval of frequency and time-frequency analysis.

The Fourier and discrete Fourier transform are the most basic ones still being in use today. They are limited to offering information only on the frequency components but none on their time localization. In contrast, the short-time Fourier transform is capable of offering a basic insight into the time scale of frequency distributions, however, it is limited by the maximum achievable resolution in both time and frequency domains.

The wavelet transform comes to avoid this limitation of resolution by offering variable width analyzing windows, which can adapt to the time scales present in the analyzed signal. Still, it is the Hilbert-Huang transform which introduces the instantaneous frequency and offers information about the signal in each time point.

These methods by themselves are capable to decompose signals in different ways, but are not able to draw conclusions about them. Thus, integrating them in a hybrid system with artificial intelligence methods offers a robust solution to many signal processing problems. We have presented several hybrid systems having various uses: a combination between the wavelet transform and the fuzzy c-means clustering to aid the differentiation of ECG patterns on mobile devices. Another hybrid system consists of using the Hilbert-Huang transform and support vector machines together to perform engine-fault detection. The last presented method combines the power of wavelet packet decomposition and artificial neural networks in order to detect sleep stages, based on EEG, EMG and EOG signals.

In conclusion, the combination of intelligent methods with classical, well-known signal processing transforms yield robust feature extraction systems with applications in different technical and scientific fields, as the presented case studies have shown.

References

1. Kaiser, G.: A Friendly Guide to Wavelets. Birkhäuser (1994)
2. Cooley, J.W., Tukey, J.W.: An algorithm for the machine calculation of complex Fourier series. Math. Comput. **19**, 297–297 (1965)
3. White, S.: A simple FFT butterfly arithmetic unit. IEEE Trans. Circuits Syst. **28**, 352–355 (1981)
4. Johnson, S.G., Frigo, M.: A Modified split-radix FFT with fewer arithmetic operations. IEEE Trans. Signal Process. **55**, 111–119 (2007)
5. Megas, D., Serra-Ruiz, J., Fallahpour, M.: Efficient self-synchronised blind audio watermarking system based on time domain and FFT amplitude modification. Signal Process. **90**, 3078–3092 (2010)

6. Hillerkuss, D., et al.: Simple all-optical FFT scheme enabling Tbit/s real-time signal processing. Opt. Express **18**, 9324–9340 (2010)
7. Zhong, R., Huang, M.: Winkler model for dynamic response of composite caisson–piles foundations: seismic response. Soil Dyn. Earthquake Eng. **66**, 241–251 (2014)
8. Carbonaro, M., Nucara, A.: Secondary structure of food proteins by Fourier transform spectroscopy in the mid-infrared region. Amino Acids **38**, 679–690 (2010)
9. McRobbie, D.W., Moore, E.A., Graves, M.J., Prince, M.R.: MRI from Picture to Proton, 2nd edn. Cambridge University Press (2007)
10. Gabor, D.: Theory of communication. Part 1: the analysis of information. J. Inst. Electr. Eng. Part III: Radio Commun. Eng. **93**, 429–441 (1946)
11. Allen, R.L., Mills, D.: Signal Analysis: Time, Frequency, Scale, and Structure. Wiley, IEEE Press (2004)
12. Chikkerur, S., Cartwright, A.N., Govindaraju, V.: Fingerprint enhancement using STFT analysis. Pattern Recogn. **40**, 198–211 (2007)
13. Sherlock, B.G.: Fingerprint enhancement by directional Fourier filtering. IEEE Proc. Vision, Image, Signal Process. **141**, 87 (1994)
14. Mallat, S. Peyre, G.: A Wavelet Tour of Signal Processing: The Sparse Way, 3rd edn. Academic Press (2009)
15. Daubechies, I.: Ten lectures on wavelets. Soc. Ind. Appl. Math. (1992)
16. Rucka, M., Wilde, K.: Application of continuous wavelet transform in vibration based damage detection method for beams and plates. J. Sound Vib. **297**, 536–550 (2006)
17. Mallat, S., Zhong, S.: Characterization of signals from multi-scale edges. IEEE Pattern Anal. Mach. Intell. **14**, 710–732 (1992)
18. Rabbani, M., Joshi, R.: An overview of the JPEG2000 still image compression standard. Signal Process. Image Commun. 3–48 (2002)
19. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In: Proceedings of the Royal Society London A, pp. 903–995 (1998)
20. Földvári, R.: Generalized instantaneous amplitude and frequency functions and their application for pitch frequency determination. J. Circuits, Syst. Comput. (1995)
21. Bedrosian, E.: A product theorem for hilbert transforms. Technical report, United States Air Force (1962)
22. Xu, Y., Yan, D.: The Bedrosian identity for the Hilbert transform of product functions. Proc. Am. Math. Soc. **134**, 2719–2728 (2006)
23. Huang, N.E., Wu, Z., Long, S.R., Arnold, K.C., Chen, X., Blank, K.: On instantaneous frequency. Adv. Adapt. Data Anal. **1**, 177–229 (2009)
24. Bedrosian, E., Nuttall, A.H.: On the quadrature approximation to the Hilbert transform of modulated signals. Proc. IEEE **54**, 1458–1459 (1966)
25. Szalai, J., Mozes, F.E.: An improved AM-FM decomposition method for computing the instantaneous frequency of non-stationary signals. In: Proceedings of the 2nd IFAC Workshop on Convergence of Information Technologies and Control Methods with Power Systems, pp. 75–79, May 2013
26. Tseng, Y.L., Ko, P.Y., Jaw, F.S.: Detection of the third and fourth heart sounds using Hilbert-Huang transform. BioMed. Eng. OnLine **11**, 8 (2012)
27. Szalai, J., Mozes, F.E.: T-Wave Detection Using the Empirical Mode Decomposition. Scientific Bulletin of "Petru Maior" University of Tirgu-Mures, **11**, 53–56 (2014)
28. Taouli, B.-R.F., S A.: Detection of QRS complexes in ECG signals based on Empirical Mode Decomposition (2011)
29. Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation, **101**, e215–e220, circulation Electronic Pages: http://circ.ahajournals.org/cgi/content/full/101/23/e215. (2000) PMID:1085218; doi:10.1161/01.CIR.101.23.e215

30. Sadhukhan, D., Mitra, M.: ECG noise reduction using Fourier coefficient suppression. In: International Conference on Control, Instrumentation, Energy and Communication, pp. 142–146 (2014)
31. Martinez, J.P., Almeida, R., Olmos, S., Rocha, A.P., Laguna, P.: A wavelet-based ECG delineator: evaluation on standard databases. IEEE Trans. Bio-med. Eng. **51**, 570–581 (2004)
32. Huang, Z., Chen, Y., Pan, M.: Time-frequency characterization of atrial fibrillation from surface ECG based on Hilbert-Huang transform. J. Med. Eng. Technol. **31**, 381–389 (2009)
33. Anas, E.M.A., Lee, S.Y., Hasan, M.K.: Exploiting correlation of ECG with certain EMD functions for discrimination of ventricular fibrillation. Comput. Biol. Med. **41**, 110–114 (2011)
34. Chouvarda, I., Maglaveras, N., Boufidou, A., Mohlas, S., Louridas, G.: Wigner-Ville analysis and classification of electrocardiograms during thrombolysis. Med. Biol. Eng. Comput. **41**, 609–617 (2003)
35. Zhu, Y., Shayan, A., Zhang, W., Chen, T.L., Jung, T.-P., Duann, J.-R., Makeig, S., Cheng, C.-K.: Analyzing high-density ECG signals using ICA. IEEE Trans. Bio-med. Eng. **55**, 2528–2537 (2008)
36. Martis, R.J., Acharya, U.R., Mandana, K.M., Ray, A.K., Chakraborty, C.: Application of principal component analysis to ECG signals for automated diagnosis of cardiac health. Expert Syst. Appl. **39**, 11792–11800 (2012)
37. Park, J., Pedrycz, W., Jeon, M.: Ischemia episode detection in ECG using kernel density estimation, support vector machine and feature selection. Biomed. Eng. Online **11**, 30 (2012)
38. Bakul, G., Tiwary, U.S.: Automated risk identification of myocardial infarction using Relative Frequency Band Coefficient (RFBC) features from ECG. Open Biomed. Eng. J. **4**, 217–222 (2010)
39. Tseng, T.-E., Peng, C.-Y., Chang, M.-W., Yen, J.-Y., Lee, C.-K., Huang, T.-S.: Novel approach to fuzzy-wavelet ECG signal analysis for a mobile device. J. Med. Syst. 71–81 (2010)
40. Bezdek, J.C., Ehrlich, R., Full, W.: FCM: the fuzzy c-means clustering algorithm. Comput. Geosci. **10**, 191–203 (1984)
41. Iber, C., Ancoli-Israel, S., Chesson, A., Quan, F.: The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification. American Academy of Sleep Medicine (2007)
42. Rechtschaffen, A., Kales, A.: A Manual Of Standardized Terminology, Techniques and Scoring Systems for Sleep Stages of Human Subjects. Washington DC Public Health Service (1968)
43. Ronzhina, M., Janousek, O., Kolarova, J., Novakova, J., Honzik, P., Provaznik, I.: Sleep scoring using artificial neural networks. Sleep Med. Rev. **16**, 251–263 (2012)
44. Flexer, A., Gruber, G., Dorffner, G.: A reliable probabilistic sleep stager based on a single EEG signal. Artif. Intell. Med. **33**, 199–207 (2005)
45. Berthomier, C., Drouot, X., Herman-Stoica, M., Berthomier, P., Prado, J., Bokar-Thire, D., Benoit, O., Mattout, J., D'ortho, M.P.: Automatic analysis of single-channel sleep EEG: validation in healthy individuals. Sleep **30**, 1587–1595 (2007)
46. Hsu, Y.L., Yang, Y.T., Wang, J.S., Hsu, C.Y.: Automatic sleep stage recurrent neural classifier using energy features of EEG signals. Neurocomputing **104**, 105–114 (2013)
47. Liang, S.F., Kuo, C.E., Hu, Y.H., Pan, Y.H., Wang, Y.H.: Automatic stage scoring of single-channel sleep EEG by using multiscale entropy and autoregressive models. IEEE Trans. Instrum. Meas. **61**, 1649–1657 (2012)
48. Fraiwan, L., Lweesy, K., Khasawneh, N., Wenz, H., Dickhaus, H.: Automated sleep stage identification system based on time-frequency analysis of a single EEG channel and random forest classifier. Comput. Methods Programs Biomed. **108**, 10–19 (2012)
49. Jo, H.G., Park, J.Y., Lee, C.K., An, S.K., Yoo, S.K.: Genetic fuzzy classifier for sleep stage identification. Comput. Biol. Med. **40**, 629–634 (2010)
50. Sukhorukova, N., et al.: Automatic sleep stage identification: difficulties and possible solutions. In: Proceedings of the 4th Australasian Workshop on Health Informatics and Knowledge Management, pp. 39–44 (2010)

51. Kerkeni, N., Alexandre, F., Bedoui, M.H., Bougrain, L., Dogui, M.: (2005) Neuronal spectral analysis of EEG and expert knowledge integration for automatic classification of sleep stages. CoRR. arXiv:0510083
52. Ebrahimi, F., Mikaeili, M., Estrada, E., Nazeran, H.: Automatic sleep stage classification based on EEG signals by using neural networks and wavelet packet coefficients. In: Conference Proceedings: … Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2008, pp. 1151–1154 (2008)
53. Wang, Y.S., Ma, Q.H., Zhu, Q., Liu, X.T., Zhao, L.H.: An intelligent approach for engine fault diagnosis based on Hilbert-Huang transform and support vector machine. Appl. Acoust. 1–9 (2014)
54. Boser, B.E., Guyon, I.M., Vapnik, V.N.: A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory—COLT'92, New York, USA, pp. 144–152. ACM Press, July 1992

http://www.springer.com/978-3-319-32190-5

New Approaches in Intelligent Image Analysis
Techniques, Methodologies and Applications
Kountchev, R.; Nakamatsu, K. (Eds.)
2016, XX, 373 p. 157 illus., 119 illus. in color.,
Hardcover
ISBN: 978-3-319-32190-5