

CHAPTER 8

ANALYSIS OF NONSTATIONARY AND MULTICOMPONENT SIGNALS

A stationary (or homogeneous) signal is one that possesses the same statistical measures for all time, or at least over the duration of observation. We have seen in the preceding chapters that most biomedical signals, being manifestations of dynamic systems and pathophysiological processes, are *nonstationary* (or heterogeneous): Figure 3.3 shows that the variance of the speech signal used as an example varies with time; Figure 3.4 shows that the spectrum or frequency content of the speech signal also varies considerably over its duration. Figures 6.12 and 6.13 show that the spectrum of a heart sound signal or PCG varies from systole to diastole and could vary in between the two events as well.

When the characteristics of a signal being studied vary considerably over the duration of interest, measures and transforms computed over the entire duration do not carry useful information: They gloss over the dynamics of the signal. A single PSD computed from a long EMG, PCG, VAG, or speech record is of no practical value. The PSD does not provide information on time localization of the frequency components of the signal. We addressed this concern in PCG signal analysis in Section 6.3.5 by segmenting the PCG into its systolic and diastolic parts by using the ECG and carotid pulse signals as timing references. But how would we be able to handle the situation when murmurs are present in systole and diastole, and when we

need to analyze the spectra of the murmurs without the contributions of S1 and S2? How could one perform segmentation of a PCG cycle into separate portions with S1, S2, and murmurs (if present)?

Furthermore, the EEG signal changes its nature in terms of rhythms, waves, transients, and spindles for which no independent references are available (see Section 1.2.6). In fact, the EEG represents a conglomeration of a number of mental and physiological processes going on in the brain at any given instant of time. The VAG signal has nonstationary characteristics related to the cartilage surfaces that come into contact depending upon the activity performed, and no other source of information can assist in identifying time instants when the signal properties change (see Section 1.2.14). Indeed, a VAG signal contains no specific events that may be identified as such, but is a concatenation of nonspecific vibrations (with, perhaps, the exception of clicks). Would we able to extend the application of the well-established signal analysis techniques that we have studied so far to such nonstationary signals?

In addition to being nonstationary, several biomedical signals possess multiple components from the same source as well as other sources that may be active at the same time. Cohen [381] defines a multicomponent signal as one that has delineated concentrations of power or energy in the time-frequency (TF) plane, with each such portion possibly related to a component of the signal; see Boashash [382] for related discussions. The speech signal is considered to be an example of a multicomponent signal: The formant structure of voiced speech indicates the presence of multiple resonances that relate to different components (see Section 1.2.12). Extending the same interpretation, we could consider EEG, PCG, and VAG signals to be multicomponent signals. In addition to the presence of multiple delineated components in the TF plane, biomedical signals could also possess multiple components arising from different sources that are active at the same time. A VAG signal could possibly contain simultaneous contributions from multiple sources of sound or vibration in the same knee joint due to different types of cartilage pathology (see Section 8.2.3). The ECG of an expectant mother recorded from the surface of her abdomen contains a mixture of the maternal as well as fetal ECGs (see Section 3.3.5). Such mixtures are also multicomponent signals with multiple sources. Given such composite signals that are also nonstationary, how can we separate them into their components for further detailed analysis?

8.1 Problem Statement

Develop methods to study the dynamic characteristics of nonstationary and multicomponent biomedical signals. Propose schemes to apply the well-established Fourier transform and AR-modeling techniques to analyze and parameterize nonstationary and multicomponent signals.

The case studies presented in the following section provide the motivation for the present study from the perspective of a few representative biomedical signals. Approaches to solve the stated problem are presented in the sections to follow. This chapter presents several techniques for segmentation-based analysis of nonstationary

signals. In addition, the chapter includes brief introductions to wavelets and joint TF analysis of nonstationary and multicomponent signals without segmentation.

8.2 Illustration of the Problem with Case Studies

8.2.1 Heart sounds and murmurs

We noted in Section 6.3.5 that the spectral contents of S1 and S2 are different due to the different states of contraction or relaxation of the ventricular muscles and the differences in their blood content during the corresponding cardiac phases. In the normal case, the QRS in the ECG signal and the dicrotic notch in the carotid pulse signal may be used to split the PCG into S1 and S2, and separate PSDs may be obtained for the signal parts as illustrated in Section 6.3.5. However, when a PCG signal contains murmurs in systole and/or diastole and possibly valve opening snaps (see Figure 6.13), it may be desirable to split the signal further.

Iwata et al. [378] applied AR modeling to PCG signals by breaking the signal into fixed segments of 25 *ms* duration (see Section 7.9). While this approach may be satisfactory, it raises questions on optimality. What should be the window duration? Is it necessary to break the intervals between S1 and S2 into multiple segments? Would it not be more efficient to compute a single AR model for the entire durations of each of S1, S2, systolic murmur, and diastolic murmur — that is, a total of only four models? It is conceivable that each model would be more accurate as all available signal samples would be used to estimate the required ACF if the signal were to be segmented adaptively as mentioned above.

8.2.2 EEG rhythms and waves

The scalp EEG represents a combination of the multifarious activities of many small zones of the cortical surface beneath each electrode. The signal changes its characteristics in relation to mental tasks, external stimuli, and physiological processes. As we have noted in Section 1.2.6 and observed in Figure 1.38, a visual stimulus blocks the alpha rhythm; slower waves become prominent as the subject goes to deeper stages of sleep (see Section 6.7); and patients with epilepsy may exhibit sharp spikes and trains of spike-and-wave complexes (see Sections 1.2.6 and 4.4.3). Description of an EEG record, as outlined in Sections 1.2.6, 4.2.4, and 4.4.3, requires the identification of several types of waves and rhythms. This suggests that the signal may first have to be broken into segments, each possessing certain properties that remain the same for the duration of the segment. Each segment may then be described in terms of its characteristic features.

8.2.3 Articular cartilage damage and knee joint vibrations

Movement of the knee joint consists of coupled translation and rotation. The configuration of the patella is such that some portion of the articular surface is in contact with

the femur throughout knee flexion and to almost full extension (see Section 1.2.14 and Figure 1.52). Goodfellow et al. [383] demonstrated that initial patellofemoral engagement occurs at approximately 20° of flexion involving both the medial and lateral facets. Figure 8.1 shows the patellar contact areas at different joint angles. As the knee is flexed, the patellofemoral contact area moves progressively upward, involving both the medial and lateral facets. At 90° of flexion, the band of contact engages the upper pole of the patella. The odd facet does not articulate with the lateral margin of the medial femoral condyle until about $120^\circ - 135^\circ$ of knee flexion.

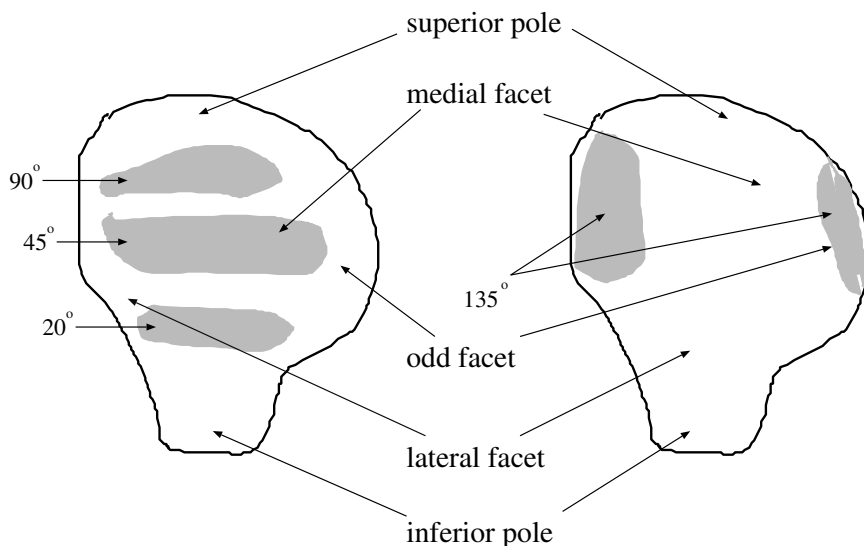


Figure 8.1 Contact areas of the patella with the femur during patellofemoral articulation. Adapted with permission from S. Krishnan, Adaptive Signal Processing Techniques for Analysis of Knee Joint Vibroarthrographic Signals, Ph.D. Thesis, University of Calgary, 1999.

Articular cartilage is composed of a solid matrix and synovial fluid [384]; it has no nerves, blood vessels, or lymphatics and is nourished by the synovial fluid covering its free surface. During articulation, friction between the bones is reduced as a result of the lubrication provided by the viscous synovial fluid [87,90]. The material properties of articular cartilage and cartilage thickness are variable not only from joint to joint but also within the same joint. In cases of abnormal cartilage, alterations in the matrix structure, such as increased hydration, disruption of the collagen fibrillar network, and disaggregation or loss of proteoglycans occur. As the compositional and biomechanical properties of abnormal articular cartilage continue to deteriorate, substance loss eventually occurs. This may be focal or diffuse, restricted to superficial fraying and fibrillation, or partial-thickness loss to full-thickness loss. In some cases, focal swelling or blistering of the cartilage may be seen before there is fraying of the articular surface [385].

Chondromalacia patella (soft cartilage of the patella) is a condition in which there is degeneration of patellar cartilage, often associated with anterior knee pain. Exposed subchondral bone and surface fibrillation of the articular cartilage are evident on the posterior patellar surface in chondromalacia patella [386]. Chondromalacia patella is usually graded in terms of the severity of the lesions [387,388] as follows:

- *Grade I:* Softening, cracking, and blistering, but no loss of articular cartilage.
- *Grade II:* Damage is moderate and there is some loss of cartilage.
- *Grade III:* Severe damage of fibrocartilage has occurred but bone is not exposed.
- *Grade IV:* The cartilage is eroded and the subchondral bone is exposed.

Osteoarthritis is a degenerative joint disease that involves specific changes to bone in addition to cartilage. In the late stages of osteoarthritis, there is full-thickness articular cartilage degeneration and exposed bone. Other structural changes include fibrous changes to the synovium, joint capsule thickening, and further alterations to the bone such as osteophyte formation [389]. Higher-grade chondromalacia may be categorized as osteoarthritis.

The menisci are subject to vertical compression, horizontal distraction, and rotary and shearing forces of varying degrees in the course of normal activities [390]. Advance of the aging process in both articular cartilage and fibrocartilage causes progressive liability to horizontal cleavage lesion [390].

The semiinvasive procedure of *arthroscopy* (fiber-optic inspection of joint surfaces, usually under general anesthesia) is often used for diagnosis of cartilage pathology. Through an arthroscope, the surgeon can usually see the patellofemoral joint, the femoral condyles, the tibial plateau (menisci), the anterior cruciate ligament, and the medial and lateral synovial spaces. Arthroscopy has emerged as the “gold standard” for relatively low-risk assessment of joint surfaces in order to determine the prognosis and treatment for a variety of conditions. Figure 8.2 shows the different stages of chondromalacia patella as viewed during arthroscopy.

Abnormal structures and surfaces in the knee joint are more likely to generate sound during extension and flexion movements than normal structures. Softened articular cartilage in chondromalacia patella — as well as cracks, fissures, or thickened areas in osteoarthritis — almost certainly increase the friction between the articular surfaces and are, therefore, likely to increase the sounds emitted during normal joint movement [92,391]. Injury to the menisci in the form of tearing causes irregularity in shape and disruption to normal joint movement, and may produce sharp clicking sounds during normal knee movement [92,97,392].

It is obvious from this discussion (and the related introduction in Section 1.2.14) that the VAG signal is a nonstationary and multicomponent signal. Different aspects of the articulating surfaces come into contact at different joint angles; their quality in terms of lubrication and functional integrity could vary from one position to another. Inspection of the VAG signals and their spectrograms illustrated in Sections 3.9.3 and 3.14 reveals that the nature of a VAG signal changes significantly over the duration of the signal. As no prior or independent information is available about changes

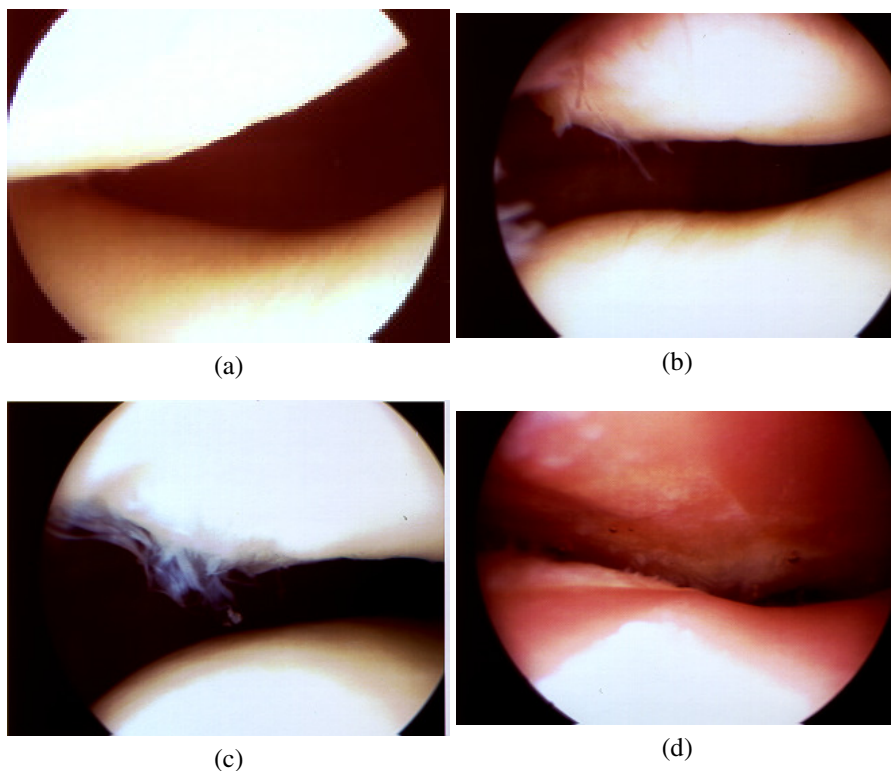


Figure 8.2 Arthroscopic views of the patellofemoral joint. (a) Normal cartilage surfaces. (b) Chondromalacia Grade II at the patella. (c) Chondromalacia Grade III at the patella. (d) Chondromalacia Grade IV at the patella and the femur; the bones are exposed. The under-surface of patella is at the top and the femoral condyle is at the bottom. Figure courtesy: G.D. Bell, Sport Medicine Centre, University of Calgary.

in the knee-joint structures that could lead to vibrations, adaptive segmentation of the VAG signal is desirable before it may be analyzed using the methods we have studied to this point in the present book. Illustrations of adaptive segmentation of VAG signals are provided in Sections 8.6.1 and 8.6.2. In addition, methods for analysis of VAG signals without segmentation are described in Section 8.14.

8.3 Time variant Systems

The linear system model represented by Equation 7.12 is a time-invariant system: The coefficients a_k and b_l of the system do not change with time, and consequently, the poles and zeros of the system stay fixed for all time. A nonstationary (or dynamic) system will possess coefficients that do vary with time: We saw in Sections 3.9.2

and 3.9.3 that the coefficient or tap-weight vectors of the adaptive LMS and RLS filters are expressed as functions of time. (*Note:* The Wiener filter described in Section 3.8, once optimized for a given set of signal and noise statistics, is a time-invariant filter.) Since the coefficients of an LMS or RLS filter vary with time, so do the transfer function and the frequency response of the filter. It follows that the impulse response of such a system also varies with time.

Let us consider an all-pole filter for the sake of simplicity; the filter characteristics are determined by the positions of the poles (except for a gain factor). If the poles are expressed in terms of their polar coordinates, their angles correspond to (resonance) frequencies and their radii are related to the associated bandwidths. We may, therefore, characterize time-variant or nonstationary systems and signals by describing their pole positions in the complex z -plane — or, equivalently, the related frequencies and bandwidths — as functions of time. A description of the variation or the modulation of the pole parameters over time can thus capture the nonstationary or dynamic nature of a time-variant system or signal. Variations in the gain factor also lead to nonstationarities in the signal produced by the system. Appel and v. Brandt [393, 394] describe the simulation of different types of nonstationary behavior of signals and systems.

In the general case of a nonstationary system that is an AR process, we may modify Equation 7.17 to indicate that the model coefficients are functions of time:

$$\tilde{y}(n) = - \sum_{k=1}^P a_k(n)y(n-k). \quad (8.1)$$

Methods related to the Kalman filter or the least-squares approach may be used to analyze such a system [18, 148, 395–397] (not considered in this book). Time-varying AR and ARMA modeling techniques have been applied to analyze the EEG [398], EGG [69], PCG [399], and HRV [400] signals; the application of time-varying analysis techniques to HRV signals is discussed in Section 8.11.

8.3.1 Characterization of nonstationary signals and dynamic systems

The output of a time-variant or dynamic system will be a nonstationary signal. The system may be characterized in terms of its time-variant model coefficients, transfer function, or related parameters derived thereof. Various short-time statistical measures computed over moving windows may be used to characterize a nonstationary signal; the measures may also be used to test for stationarity, or lack thereof, of a signal.

- **Mean:** The short-time mean represents the average or DC level of the signal in the analysis window. Variation of the mean from one window to another is usually an indication of the presence of a wandering baseline or low-frequency artifact, as in the case of the ECG signal in Figure 3.6. Clearly, the signal in Figure 3.6 is nonstationary in the mean. However, the mean is not an important measure in most signals, and it is typically blocked at the data-acquisition stage

via capacitive coupling and/or a highpass filter. Furthermore, since a DC level carries no sound or vibration information, its removal is of no consequence in the analysis of signals such as heart sounds, speech, VAG, and the VMG.

- **Variance:** Figure 3.3 illustrates the short-time variance for a speech signal. It is evident that the variance is high in regions of high signal variability (swings or excursions) about the mean, as in the case of the vowels or voiced-speech segments in the signal. The variance is low in the regions related to the fricatives or unvoiced-speech segments in the signal where the amplitude swing is small, in spite of their high-frequency nature. Since the mean of the signal is zero, the variance is equal to the MS value and represents the average power level in the corresponding signal windows. Although variations in the power level of speech signals may be useful in making voiced/ unvoiced/ silence decision, the parameter does not bear much information and provides a limited representation of the general statistical variability of signal characteristics. Regardless of the interpretation of the parameter, it is seen that the speech signal in Figure 3.3 is nonstationary in its variance (and the related measures of SD , MS , and RMS). From the discussion in Section 1.2.12, it is also clear that the vocal-tract system producing the speech signal is a dynamic system with time-varying configuration and filtering characteristics.
- **Measures of activity:** We have studied several measures of activity that indicate the “busy-ness” of the given signal, such as turning points, ZCR , and turns count, in Chapters 3 and 5. (The term “activity” has several connotations in the literature; see Chapter 5.) The short-time count of turning points is plotted in Figure 3.1 for a speech signal: It is evident that the signal is more active or busy in the periods related to the fricatives than those related to the vowels (a trend that is the opposite of that in the short-time variance of the same signal shown in Figure 3.3). The short-time turns count plot of the EMG signal in Figure 5.9 indicates the rising level of complexity of the signal with the level of breathing (inspiration). Although turning points, ZCR , and turns count are not among the traditional statistical measures derived from PDFs, they characterize signal variability and complexity in different ways. Both the examples cited above illustrate variation of the parameters measured over the duration of the corresponding signals: The signals are, therefore, nonstationary in terms of the number of turning points or the turns count.
- **ACF:** The ACF was defined in Section 3.2.1 in general as $\phi_{xx}(t_1, t_1 + \tau) = E[x(t_1)x(t_1 + \tau)]$. In Section 3.2.4, one of the conditions for (wide-sense or second-order) stationarity was defined as the ACF being independent of time, that is, $\phi_{xx}(t_1, t_1 + \tau) = \phi_{xx}(\tau)$. A nonstationary signal will not satisfy this condition and will have an ACF that varies with time. Since the ACF is based on the expectation of pairs of signal samples separated by a certain time difference or shift, it is a more general measure of signal variability than the variance and related measures. Note that the ACF for zero lag is the MS value of the signal.

One faces limitations in computing the ACF of short-time segments of a signal to investigate (non)stationarity: The shorter the analysis window, the shorter the maximum lag up to which the ACF may be estimated reliably. Regardless, the short-time ACF may be used to track nonstationarities in a signal. If the signal is the result of a dynamic AR system, the system parameters may be derived from the ACF (see Section 7.5).

- **PSD:** The PSD and ACF of a signal are interrelated by the Fourier transform. Therefore, a signal that is (non)stationary in its ACF is also (non)stationary in its PSD. However, the PSD is easier to interpret than the ACF, as we have seen in Chapter 6. The spectrogram of the speech signal in Figure 3.4 indicates significant variations in the short-time PSD of the signal: The speech signal is clearly nonstationary in its PSD (and ACF). The spectrograms of VAG signals in Sections 3.9.3 and 3.14 illustrate the nonstationary nature of VAG signals.
- **Higher-order statistics:** A major limitation of signal analysis using the ACF (or equivalently the PSD) is that the phase information is lost. The importance of phase in signals is discussed by Oppenheim and Lim [401]. Various conditions under which a signal may be reconstructed from its magnitude spectrum only or from its phase spectrum only are described by Hayes et al. [402] and Oppenheim and Lim [401]. Analysis based only upon the ACF cannot be applied to signals that are of mixed phase (that is, not minimum phase), that are the result of nonlinear systems, or that follow a PDF other than a Gaussian [403].

The general n^{th} -order moment of a random signal $x(t)$ at the instant t_1 is defined as [148, 403, 404]

$$m_x^n(t_1, t_1 + \tau_1, t_1 + \tau_2, \dots, t_1 + \tau_{n-1}) = E[x(t_1) x(t_1 + \tau_1) x(t_1 + \tau_2) \cdots x(t_1 + \tau_{n-1})], \quad (8.2)$$

where $\tau_1, \tau_2, \dots, \tau_{n-1}$ are various shifts. It is evident that the ACF is a special case of the above with $n = 2$, that is, the ACF is the second-order moment.

A set of parameters known as cumulants may be related to the moments as follows: The second-order and third-order cumulants are equal to the corresponding moments. The fourth-order cumulant is related to the fourth-order moment as [148, 403, 404]

$$\begin{aligned} c_x^4(t_1, t_1 + \tau_1, t_1 + \tau_2, t_1 + \tau_3) &= m_x^4(t_1, t_1 + \tau_1, t_1 + \tau_2, t_1 + \tau_3) \\ &- m_x^2(t_1, t_1 + \tau_1) m_x^2(t_1 + \tau_2, t_1 + \tau_3) \\ &- m_x^2(t_1, t_1 + \tau_2) m_x^2(t_1 + \tau_3, t_1 + \tau_1) \\ &- m_x^2(t_1, t_1 + \tau_3) m_x^2(t_1 + \tau_1, t_1 + \tau_2). \end{aligned} \quad (8.3)$$

The Fourier transforms of the cumulants provide the corresponding higher-order spectra or polyspectra (with as many frequency variables as the order minus

one). The Fourier transforms of the second-order, third-order, and fourth-order cumulants are known as the power spectrum (PSD), bispectrum, and trispectrum, respectively. A Gaussian process possesses only first-order and second-order statistics: Moments and spectra of order higher than two are zero. Higher-order moments, cumulants, and spectra may be used to characterize nonlinear, mixed-phase, and non-Gaussian signals [148, 403, 404]. Variations over time of such measures may be used to detect the related types of nonstationarity.

- **System parameters:** When a time-varying model of the system producing the signal is available in terms of its coefficients, such as $a_k(n)$ in Equation 8.1, we may follow or track changes in the coefficients over time. Significant changes in the model parameters indicate corresponding changes in the output signal.

The following sections present several examples of application of the concepts and notions listed above.

8.4 Fixed Segmentation

Given a nonstationary signal, the simplest approach to break it into quasistationary segments would be to consider small windows of fixed duration. Given a signal $x(i)$ for $i = 0, 1, 2, \dots, N-1$, we could consider a fixed segment duration of M samples, with $M \ll N$, and break the signal into $K = N/M$ parts as

$$x_k(n) = x[n + (k-1)M], \quad 0 \leq n \leq M-1, \quad 1 \leq k \leq K. \quad (8.4)$$

With the assumption that the signal does not change its characteristics to any significant extent within the duration corresponding to M samples (or $\frac{M}{f_s}$ s), each segment may be considered to be quasistationary.

Note that the segmentation here is similar to that in the Bartlett and Welch procedures described in Sections 6.3.2 and 6.3.3. However, we will not be averaging the spectra over the segments now, but will be treating them as separate entities. The signal processing techniques we have studied so far may then be applied to analyze each segment separately.

8.4.1 The short time Fourier transform

Once a given signal has been segmented into quasistationary parts $x_k(n)$ as in Equation 8.4, we may compute the Fourier spectrum for each segment as

$$X_k(\omega) = \sum_{n=0}^{M-1} x_k(n) \exp(-j\omega n). \quad (8.5)$$

The array of spectra $X_k(\omega)$ for $k = 1, 2, \dots, K$ will describe the time-varying spectral characteristics of the signal.

Segmentation of the given signal as above may be interpreted as the application of a moving window to the signal. The k^{th} segment $x_k(n)$ may be expressed as the multiplication of the signal $x(n)$ with a window function $w(n)$ positioned at the beginning of the segment as

$$x_k(n) = x(n) w[n - (k - 1)M], \quad 1 \leq k \leq K, \quad (8.6)$$

where

$$w(n) = \begin{cases} 1 & \text{for } 0 \leq n \leq M - 1, \\ 0 & \text{otherwise.} \end{cases} \quad (8.7)$$

Figure 8.3 (a) illustrates the PCG of a patient with systolic murmur and opening snap of the mitral valve, with a moving rectangular analysis window of duration 64 *ms* superimposed on the signal at three different instants of time. The duration of each window is 64 samples, equal to 64 *ms* with $f_s = 1 \text{ kHz}$. The three windows have been positioned approximately over the S1, systolic murmur, and S2 events in the signal. Figure 8.3 (b) shows the log PSDs of the signal segments extracted by the three analysis windows. It is seen that the PSDs differ significantly, with the second window displaying the largest amount of high-frequency power due to the murmur. The third window displays more medium-frequency content than the first. It is clear that the PCG signal is nonstationary in the PSD.

In general, the window may be positioned at any time instant m , and the resulting segment may be expressed as $x(n)w(n - m)$. We need to state how the window is moved or advanced from one segment to another; in the extreme situation, we may advance the window one sample at a time, in which case adjacent windows would have an overlap of $(M - 1)$ samples. We may then compute the Fourier transform of every segment as

$$X(m, \omega) = \sum_{n=0}^{M-1} [x(n) w(n - m)] \exp(-j\omega n). \quad (8.8)$$

In the case when both the time and frequency variables are continuous, we may write the expression above in a more readily understandable form as

$$X(\tau, \omega) = \int_{-\infty}^{\infty} [x(t) w(t - \tau)] \exp(-j\omega t) dt. \quad (8.9)$$

The spectrum is now expressed not only as a function of frequency ω , but also as a function of time τ . Although the limits of the integral have been stated as $(-\infty, \infty)$, the finite duration of the window placed at time τ performs segmentation of the signal as desired.

The spectral representation of the signal as a function of time in Equations 8.8 and 8.9 is known as a *time-frequency distribution* or TFD [405–408]. In practice, only the squared magnitudes of the expressions in Equations 8.8 and 8.9 are used for the display of TFDs. Because the Fourier transform is applied, in the procedure

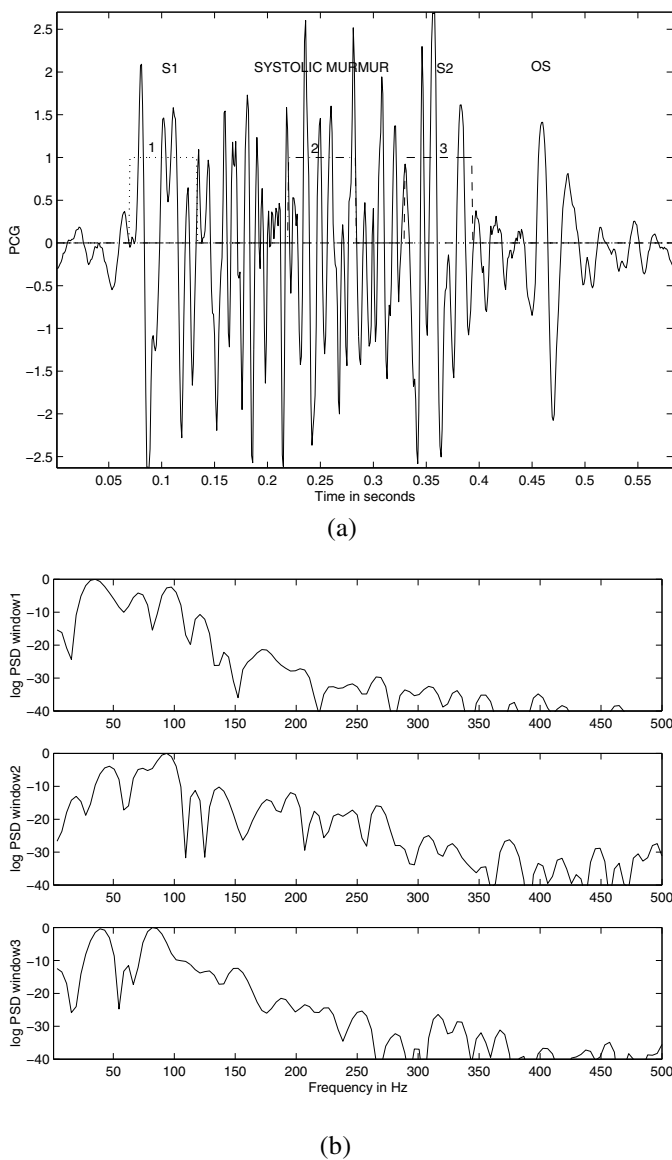


Figure 8.3 (a) PCG signal of a patient (female, 14 months) with systolic murmur and opening snap (OS) of the mitral valve. Three short-time analysis windows are superimposed, each one being a rectangular window of duration 64 ms. (b) Log PSDs of the three windowed signal segments. Each DFT was computed with zero-padding to a total length 256 samples. $f_s = 1$ kHz. See also Figures 6.13 and 8.4.

above, to short windows of the signal in time, the result is known as the *short-time Fourier transform* or STFT of the signal. The method of analysis of a nonstationary signal in short windows is, in general, known as *short-time analysis*. The magnitude of the STFT (squared and/or with the logarithmic operation if desired) is known as the *spectrogram* of the signal.

Figure 8.4 illustrates the spectrogram of the PCG signal of a patient with systolic murmur and opening snap of the mitral valve: The signal and the window parameters are the same as in Figure 8.3, but now the spectra are plotted for every window position with a displacement of 32 *ms*. The relatively high-frequency nature of the murmur as compared to S1 and S2 is clearly evident in the spectrogram.

We have previously encountered spectrograms of speech and VAG signals: Refer to Figure 3.4 and Sections 3.9.3 and 3.14. More examples of spectrograms are provided at the end of this section and later in this chapter.

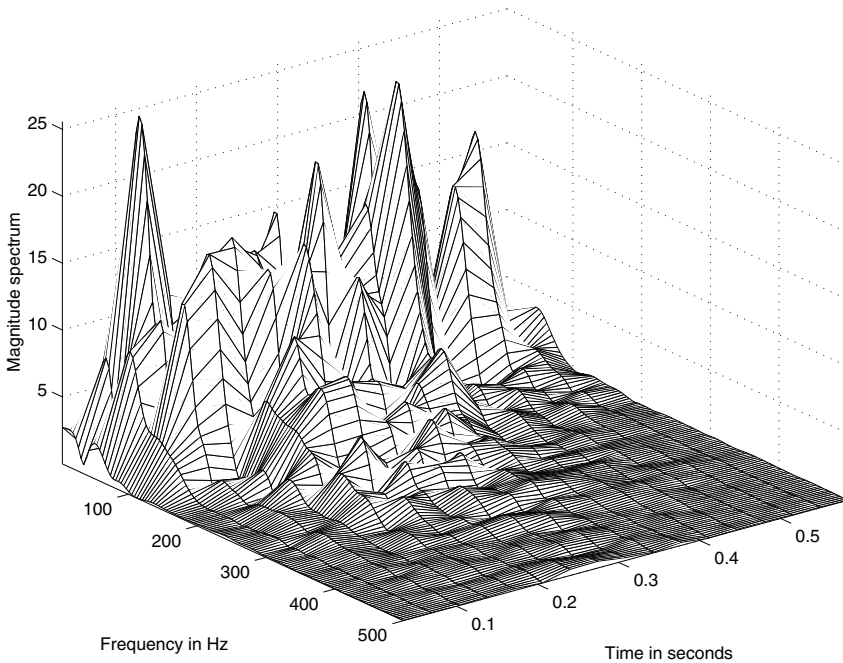


Figure 8.4 Spectrogram of the PCG signal of a patient (female, 14 months) with systolic murmur and opening snap of the mitral valve, computed with a moving short-time analysis window of duration 64 samples (64 *ms* with $f_s = 1 \text{ kHz}$), with the window advance interval being 32 samples. Each DFT was computed with zero-padding to a total length 256 samples. $f_s = 1 \text{ kHz}$. See also Figures 6.13 and 8.3.

8.4.2 Considerations in short time analysis

Short-time analysis of signals could be computationally expensive. In the case of the STFT, the Fourier transform has to be computed for each segment of the signal. In practice, there should be no need to compute the Fourier transform for every possible window position, that is, for every m in Equation 8.8. We could advance the analysis window by M samples, in which case adjacent windows will not overlap. It is common practice to advance the analysis window by $\frac{M}{2}$ samples, in which case adjacent windows will overlap for $\frac{M}{2}$ samples; some overlap is desirable in order to maintain continuity in the STFT or TFD computed.

An important question arises regarding the duration of the analysis window M to be used. The window should be short enough to ensure that the segment is stationary, but long enough to permit meaningful analysis and adequate representation of low-frequency components. We have seen in Section 6.4 that a short window possesses a wide main lobe in its frequency response. Since the given signal is multiplied in the time domain with the analysis window, the spectrum of the signal gets convolved with the spectral response of the window in the frequency domain. Convolution in the frequency domain with a function having a large main lobe leads to significant blurring and loss of spectral resolution.

The limitation imposed by the use of a window is related to the uncertainty principle or time–bandwidth product, expressed as [406]

$$\Delta t \times \Delta \omega \geq \frac{1}{2}, \quad (8.10)$$

where

$$(\Delta t)^2 = \int_{-\infty}^{\infty} (t - \bar{t})^2 |x(t)|^2 dt, \quad (8.11)$$

$$\bar{t} = \int_{-\infty}^{\infty} t |x(t)|^2 dt, \quad (8.12)$$

$$(\Delta \omega)^2 = \int_{-\infty}^{\infty} (\omega - \bar{\omega})^2 |X(\omega)|^2 d\omega, \quad (8.13)$$

$$\bar{\omega} = \int_{-\infty}^{\infty} \omega |X(\omega)|^2 d\omega, \quad (8.14)$$

and Δt and $\Delta \omega$ represent the time extent (duration) and frequency extent (bandwidth) of the signal $x(t)$ and its Fourier transform $X(\omega)$, respectively. The gist of the limitation stated above is that both a signal and its Fourier transform cannot simultaneously be made arbitrarily narrow. The effect of this limitation on the STFT and TFD-based analysis is that we cannot simultaneously obtain arbitrarily high resolution along both the time and frequency axes.

At the extremes, a continuous-time signal $x(t)$ provides infinite time resolution but no frequency resolution: The value of the signal is known at every instant of time t , but nothing is known about the frequency content of the signal. Conversely, the PSD $S_{xx}(f)$ provides infinite frequency resolution but no time resolution: The

overall strength of sinusoids at every frequency f present in the signal over all time t is known, but nothing is known about where exactly in time a given frequency component begins or ends. (The phase spectrum contains this information but cannot be readily interpreted and used for the purposes of this discussion.)

In the case of sampled signals and spectra, the sampling intervals Δt in the time domain and Δf in the frequency domain will be finite and will be limited by Heisenberg's inequality as stated above. Increasing the time resolution of the STFT by making the analysis window short in duration will compromise frequency resolution; on the other hand, increasing the window duration will lead to a loss in time resolution.

In general, the window function $w(n)$ included in Equation 8.8 need not be a rectangle: any of the window functions listed in Section 6.3.3 may be used. Once a window is chosen, the joint time and frequency resolution is the same over the entire TF plane.

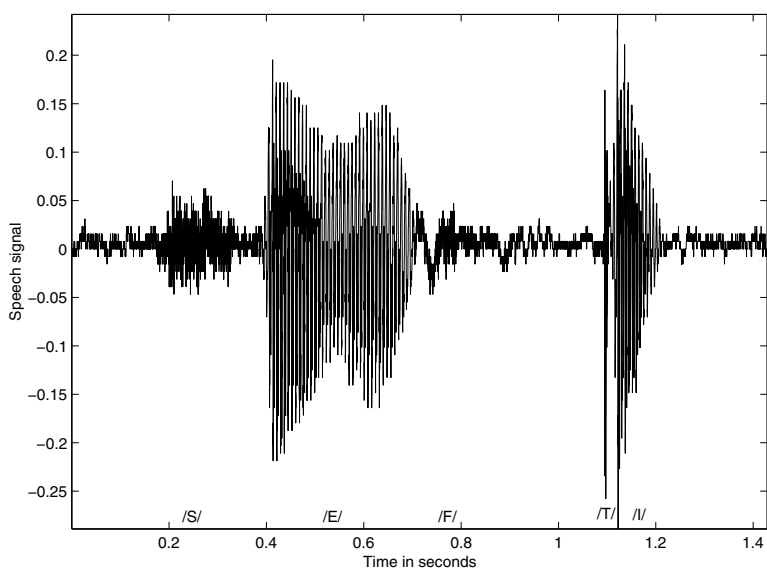
The STFT expression in Equation 8.8 indicated the placement of a causal analysis window beginning at the time instant of reference m in the argument of the STFT. It is also common practice to use a symmetrical noncausal window defined for $-\frac{M}{2} \leq n \leq \frac{M}{2}$, in which case the reference point of the analysis window would be the center of the window.

See Allen and Rabiner [409], Portnoff [410], and Rabiner and Schafer [79] for discussions on short-time Fourier analysis.

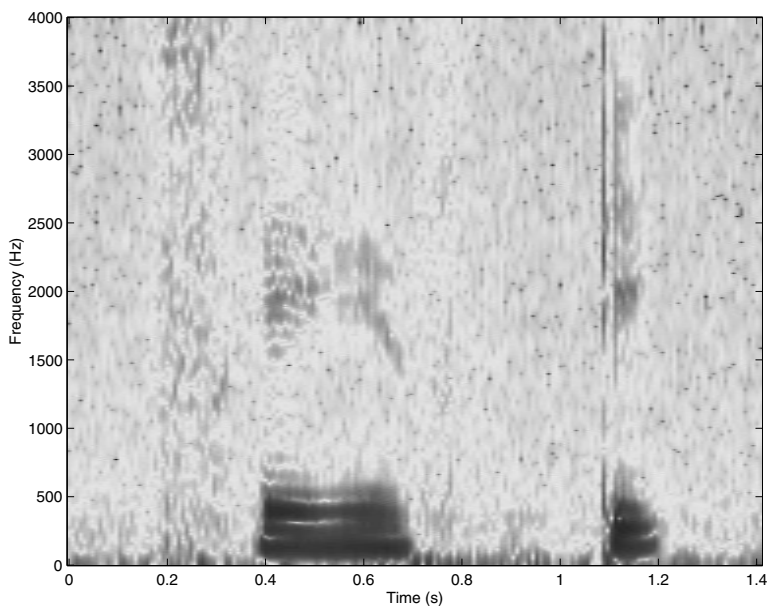
Illustration of application: Spectrograms of the speech signal in Figure 1.50 with different window parameters are provided in Figures 8.5 and 8.6. The spectrograms are shown here as gray-scale images, with the darkness at each point being proportional to the log PSD for the corresponding temporal analysis window position and frequency coordinate. It is evident that increasing the length of the analysis window provides better frequency resolution (the definition or clarity of the frequency components) while at the same time reducing the temporal resolution (that is, causing smearing in the temporal dimension). Decreasing the window length causes the reverse effects. The spectrogram in Figure 8.5 (b) with the analysis window duration being 16 ms clearly illustrates the high-frequency (broadband) nature of the fricatives; the transient and broadband nature of the plosive /T/ is also clearly shown. The same features are not clearly depicted by the spectrogram in Figure 8.6 (b) where the analysis window is fairly long (128 ms); however, the formant structure of the voiced-speech components (the vowels) is clearly depicted. The formant structure of the voiced-speech components is not clearly visible in the spectrogram in Figure 8.5 (b).

8.5 Adaptive Segmentation

One of the limitations of short-time analysis lies with the use of a fixed window duration. A signal may remain stationary for a certain duration of time much longer than the window duration chosen, and yet the signal would be broken into many segments over such a duration. Conversely, a signal may change its characteristics

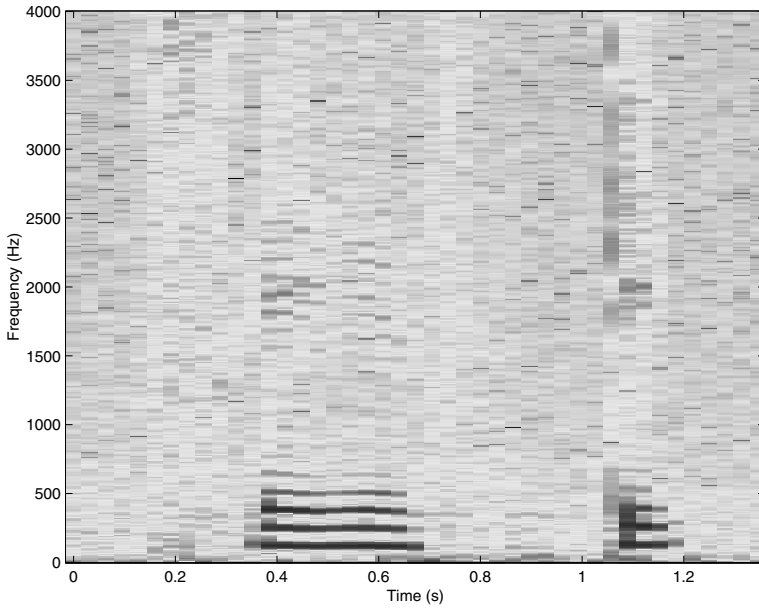


(a)

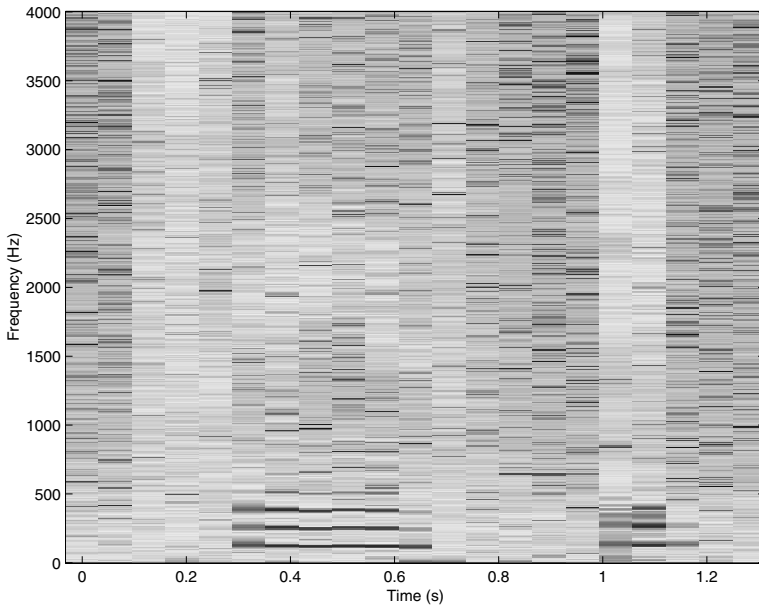


(b)

Figure 8.5 (a) Time-domain speech signal of the word “safety” uttered by a male speaker. (The signal is also illustrated in Figures 1.50, 3.1, and 3.3.) (b) Spectrogram (log PSD) of the signal computed with a moving short-time analysis window of duration 16 *ms* (128 samples with $f_s = 8$ *kHz*), with the window advance interval being 8 *ms*.



(a)



(b)

Figure 8.6 Spectrograms (log PSD) of the speech signal in Figure 8.5: (a) with a moving window of duration 64 *ms* (512 samples with $f_s = 8$ *kHz*) and window advance interval of 32 *ms*, and (b) with a moving window of duration 128 *ms* (1,024 samples) and window advance interval of 64 *ms*.

within the duration of the fixed window: Short-time analysis cannot guarantee stationarity of the signal over even the relatively short duration of the analysis window used. It would be desirable to adapt the analysis window to changes in the given signal, allowing the window to be as long as possible while the signal remains stationary, and to start a new window at the exact instant when the signal or the related system changes its characteristics.

Problem: *Propose adaptive methods to break a nonstationary signal into quasi-stationary segments of variable duration.*

Solution: We saw in Section 7.5 that a signal may be represented or modeled as a linear combination of a small number of past values of the signal, subject to a small error of prediction. It then follows that if a signal were to change its behavior, it would no longer be predictable from its preceding samples as they would correspond to the previous state of the time-variant system generating the nonstationary signal. Therefore, we could expect a large jump in the prediction error at instants of time when the signal changes in its characteristics. Furthermore, the AR-model parameters represent the system generating the signal and provide the poles of the system. If the system were to change in terms of the locations of its poles (related to its resonance frequencies), the same model would no longer hold: A new model would have to be initiated at such instants of change. This suggests that we could estimate AR models on a short-time basis and monitor the model parameters from segment to segment: A significant change in the model parameters would indicate a point of change in the signal. (We have seen in Section 7.9 how a similar approach was used by Iwata et al. [379] to detect S1 and S2 in PCGs.) Adjacent segments that have the same or similar model parameters could be concatenated to form longer segments in a subsequent stage of processing. As the AR model provides several parameters and may be interpreted in several ways (see Section 7.5.2), tracking the behavior of the model over a moving analysis window may be accomplished in many ways. The following sections provide the details of a few approaches for adaptive segmentation based upon the notions stated above.

8.5.1 Spectral error measure

Bodenstein and Praetorius [197] and Praetorius et al. [411] used the all-pole LP or AR model (see Section 7.5) for adaptive segmentation of EEG signals into quasistationary segments and also for further feature extraction. They made the following observations about the application of AR modeling to EEG signals:

- *Time domain:* The present value of the prediction error indicates the instantaneous degree of “unexpectedness” in the signal.
- *Autocorrelation domain:* The prediction error is decorrelated.
- *Spectral domain:* The prediction error being white noise, the AR model yields an all-pole representation of the signal spectrum, which is particularly suitable for the modeling of resonance.

These properties are useful for

- detection and elimination of transients;
- segmentation of the EEG into quasistationary segments; and
- feature extraction and pattern recognition (diagnosis).

Ferber [412] provided a description of nonstationarities in the EEG and suggested a few approaches to treat the same.

Analysis of spectral change: Let the PSD of the given nonstationary signal be $S(0, \omega)$ at zero time and $S(t, \omega)$ at time t . The *spectral error* of $S(t, \omega)$ with respect to $S(0, \omega)$ may be taken to be dependent upon the difference between the corresponding log PSDs — that is, to be proportional to $\{\log[S(t, \omega)] - \log[S(0, \omega)]\}$ or to be proportional to $\log \left[\frac{S(t, \omega)}{S(0, \omega)} \right]$. Consider the state when an AR model has been adapted to the signal spectrum $S(0, \omega)$ at zero time. If we pass the signal at time t through the AR model, the prediction error will have an instantaneous spectrum given by

$$S_e(\omega) = \frac{S(t, \omega)}{S(0, \omega)}, \quad (8.15)$$

which is similar to the spectral ratio in Equation 7.50. Thus, the problem of comparing two arbitrary PSDs of a nonstationary signal at two different instants of time may be expressed as testing $S_e(\omega)$ for deviation from a uniform PSD.

Let $a_R(k)$, $k = 1, 2, \dots, P$, represent the reference AR model. When the current signal $y(n)$ is passed through the filter represented by the AR model, we obtain the prediction error

$$e(n) = \sum_{k=0}^P a_R(k) y(n - k), \quad (8.16)$$

with the prediction model defined as in Equation 7.17, the prediction error as given in Equation 7.18, and $a_R(0) = 1$. The error indicates the deviation of the current signal from the previously computed model. Consider the integral

$$\varepsilon = \int_{-\infty}^{\infty} [1 - S_e(\omega)]^2 d\omega, \quad (8.17)$$

where $S_e(\omega)$ is the PSD of the prediction error. Ideally, when the AR model has been optimized for the signal on hand, the prediction error is expected to have a uniform PSD. However, if the signal is nonstationary, some changes would have occurred in the spectral characteristics of the signal, which would be reflected in the PSD of the error. If $\phi_e(k)$ is the ACF corresponding to $S_e(\omega)$, the latter is given by the Fourier transform of the former. However, since both functions are real and even, we have

$$S_e(\omega) = \phi_e(0) + 2 \sum_{k=1}^{\infty} \phi_e(k) \cos(2\pi\omega k). \quad (8.18)$$

Then,

$$\varepsilon = \int_{-\infty}^{\infty} \left[1 - \phi_e(0) - 2 \sum_{k=1}^{\infty} \phi_e(k) \cos(2\pi\omega k) \right]^2 d\omega. \quad (8.19)$$

Due to the orthonormality of the trigonometric functions, we get

$$\varepsilon = [1 - \phi_e(0)]^2 + 2 \sum_{k=1}^{\infty} \phi_e^2(k). \quad (8.20)$$

In practice, the summation may be performed up to some lag, M . Bodenstein and Praetorius [197] recommended normalization of the error measure by division by $\phi_e^2(0)$, leading to the *spectral error measure (SEM)*

$$SEM = \left[\frac{1}{\phi_e(0)} - 1 \right]^2 + 2 \sum_{k=1}^M \left[\frac{\phi_e(k)}{\phi_e(0)} \right]^2. \quad (8.21)$$

Here, the first term represents the change in the total power of the prediction error; the second term depends upon the change in spectral shape only. Note that the prediction error is expected to have a uniform (flat) PSD as long as the signal remains stationary with respect to the AR model designed. The *SEM* was shown to vary significantly in response to changes in the spectral characteristics of EEG signals and to be useful in breaking the signals into quasistationary parts. Figure 8.7 shows the general scheme of EEG segmentation by using the *SEM*.

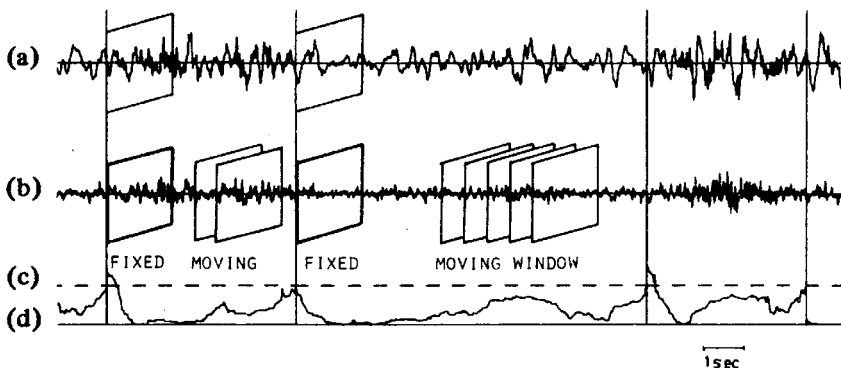


Figure 8.7 Adaptive segmentation of EEG signals via use of *SEM*. (a) Original EEG signal. The rectangular window at the beginning of each adaptive segment indicates the signal window to which the AR model has been optimized. (b) Prediction error. The initial ACF of the error is computed over the fixed window; the running ACF of the error is computed over the moving window. (c) Segmentation threshold. (d) *SEM*. The vertical lines represent the segmentation boundaries. Reproduced with permission from G. Bodenstein and H.M. Praetorius, Feature extraction from the electroencephalogram by adaptive segmentation, *Proceedings of the IEEE*, 65(5):642–652, 1977. ©IEEE.

Algorithm for adaptive segmentation [197]: Let $n = 0$ represent the starting point of analysis where the first reference or fixed analysis window is placed for each adaptive segment, as in Figure 8.7 (a). $(N + P)$ samples of the signal $y(n)$ should be available prior to the arbitrarily designated origin at $n = 0$, where $(2N + 1)$ is the size of the analysis window and P is the order of the AR model to be used.

1. Using the signal samples $y(-N)$ to $y(N)$, compute the signal ACF up to lag P .
2. Derive the corresponding AR model of order P .
3. Using the signal values $y(-N - P)$ to $y(n + N)$, compute the prediction error $e(-N)$ to $e(n + N)$ and compute the running short-time ACF $\phi_e(n, m)$ of the prediction error as

$$\phi_e(n, m) = \frac{1}{2N + 1} \sum_{k=-N}^{N-m} e(n + k) e(n + k + m). \quad (8.22)$$

Note that the ACF now has two indices: the first index n indicates the position of the short-time analysis window, and the second index m indicates the lag for which the ACF is computed.

4. Calculate $\phi_e(0, m)$ for $m = 0, 1, \dots, M$. This represents the fixed window at the beginning of each adaptive segment in Figure 8.7 (a).

Perform the following three steps for each data point:

5. Compute $\phi_e(n, m)$ for the moving window [see Figure 8.7 (b)] by the recursive relationship

$$\begin{aligned} \phi_e(n, m) &= \phi_e(n - 1, m) + e(n + N) e(n + N - m) \\ &\quad - e(n - N - 1) e(n - N - 1 - m). \end{aligned} \quad (8.23)$$

This represents the moving window in Figure 8.7 (b).

6. Compute the SEM at time n as

$$SEM(n) = \left[\frac{\phi_e(0, 0)}{\phi_e(n, 0)} - 1 \right]^2 + 2 \sum_{k=1}^M \left[\frac{\phi_e(n, k)}{\phi_e(n, 0)} \right]^2, \quad (8.24)$$

where $\phi_e(0, 0)$ accounts for the fact that the signal may have an arbitrary power level.

7. Test if $SEM(n) > Th_1$, where Th_1 is a threshold.

If the condition is not satisfied, increase n by 1 and return to Step 5.

If the condition is satisfied, a segment boundary has been detected at time n , as indicated by the vertical lines in Figure 8.7. Reset the procedure by the following step:

8. Shift the time axis by substituting $(n + k)$ with $(k - N)$ and start the procedure again with Step 1.

In the investigations of Bodenstein and Praetorius [197], SEM demonstrated sharp jumps as transients of duration less than 100 *ms* entered and left the moving

analysis window of duration 2 s ($2N + 1 = 101$ samples with $f_s = 50$ Hz). Such jumps could lead to inappropriate segmentation, especially with burst-suppression type EEG episodes as illustrated in Figure 8.8. To overcome this problem, it was suggested that the prediction error $e(n)$ be limited (clipped) by a threshold Th_2 as

$$e(n) = \begin{cases} e(n), & \text{if } |e(n)| < Th_2, \\ \text{sgn}[e(n)] Th_2, & \text{if } |e(n)| \geq Th_2. \end{cases} \quad (8.25)$$

The threshold Th_2 is shown by the dashed lines in Figure 8.8 (c). The *SEM* computed from the clipped $e(n)$ is shown in Figure 8.8 (d), which, when checked against the original threshold Th_1 , will yield the correct segmentation boundary. The signal reconstructed from the clipped prediction error is shown in Figure 8.8 (e), which shows that the clipping procedure has suppressed the effect of the transient without affecting the rest of the signal.

In spite of the clipping procedure as in Equation 8.25, it was indicated by Bodenstein and Praetorius [197] that the procedure was more sensitive than desired and caused false alarms. To further limit the effects of random fluctuations in the prediction error, a smoothed version $e_s(n)$ of the squared prediction error was computed as

$$e_s(n) = e^2(n-1) + 2e^2(n) + e^2(n+1) \quad (8.26)$$

for those samples of $e(n)$ that satisfied the condition $|e(n)| > Th_2$. Another threshold Th_3 was applied to $e_s(n)$, and the triplet $\{y(n-1), y(n), y(n+1)\}$ was considered to be a part of a transient only if $e_s(n) > Th_3$. The procedure of Bodenstein and Praetorius combines adaptive segmentation of EEG signals with transient detection as the two tasks are interrelated.

Illustration of application: Figure 8.9 shows the EEG signal of a child in sleep stage I, superimposed with 14 Hz spindles. The *SEM* and its components are also shown in the figure. The vertical lines indicate the segment boundaries detected. Bodenstein et al. [413] and Creutzfeldt et al. [414] describe further extension of the approach to computerized pattern classification of EEG signals including clustering of similar segments and labeling of the types of activity found in an EEG record.

The *SEM* method was applied by Tavathia et al. [93] for adaptive segmentation of VAG signals. It was indicated that each segment could be characterized by the frequency of the most-dominant pole obtained via AR modeling and the spectral power ratio $E_{40:120}$ as per Equation 6.40; however, no classification experiments were performed. More examples of application of the *SEM* technique are presented in Sections 8.5.4 and 8.9.

8.5.2 ACF distance

Michael and Houchin [415] proposed a method comparable to that of Bodenstein and Praetorius [197], but based upon a simpler scheme using the ACF. It should be noted that the AR-model coefficients are derived from the ACF and that the spectra used to compute *SEM* are related to the corresponding ACFs by the Fourier transform.

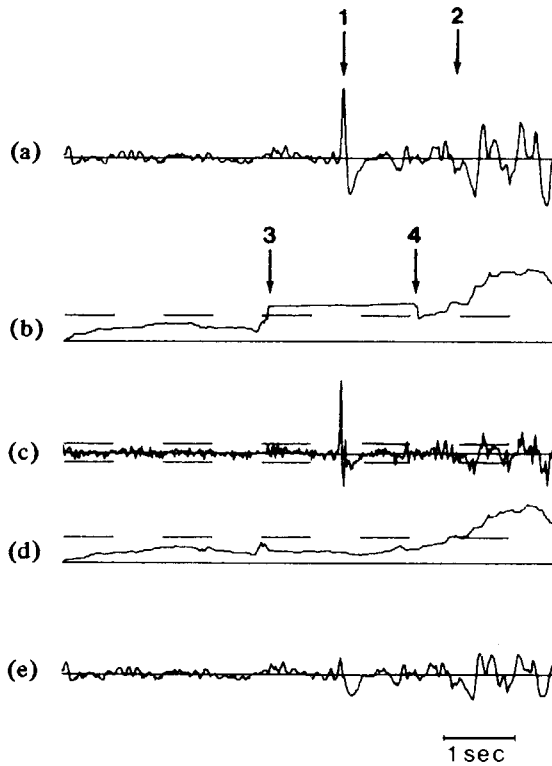


Figure 8.8 Elimination of transients by clipping the prediction error. (a) Original EEG signal of the burst-suppression type. The sharp wave marked by the arrow 1 is followed by the onset of a burst marked by the arrow 2. (b) *SEM* showing sudden jumps at points indicated by the arrows 3 and 4 as the sharp wave enters and leaves the analysis window. (c) Clipping of the prediction error with threshold Th_2 . (d) *SEM* after clipping the prediction error. The dashed line represents the threshold Th_1 . (e) Signal reconstructed from the clipped prediction error. Reproduced with permission from G. Bodenstein and H.M. Praetorius, Feature extraction from the electroencephalogram by adaptive segmentation, *Proceedings of the IEEE*, 65(5):642–652, 1977. ©IEEE.

However, direct use of the ACF removes the assumption made in AR modeling that the signal is the result of an AR process.

In the method of Michael and Houchin, the ACF is treated as a statistical measure of the given signal, and significant variations in the ACF are used to detect nonstationarity. A reference window is extracted at the beginning of each scan, and the given signal (EEG) is observed through a moving window. The duration of the window has to be chosen such that it is shorter than the shortest expected quasistationary segment of the given signal, but long enough to characterize the lowest frequency present. If the difference between the signal's statistics (ACF) in the moving win-

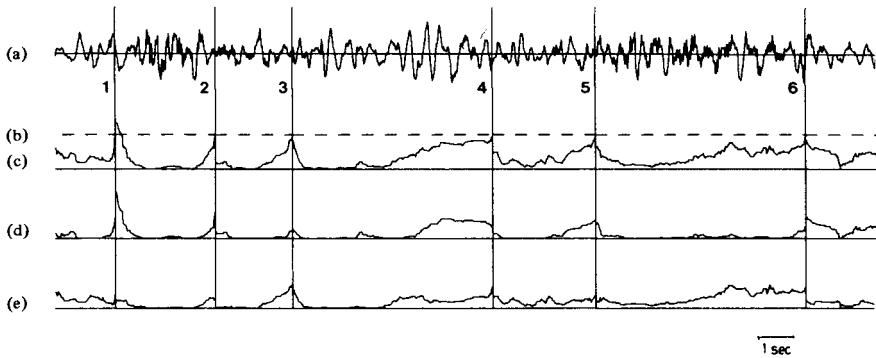


Figure 8.9 Use of *SEM* to segment an EEG signal. (a) Original EEG signal of a child in sleep stage I with superimposed 14 *Hz* spindles. (b) Segmentation threshold. (c) *SEM*. (d) Deviation in prediction error power. (e) Deviation in prediction error spectral shape. The vertical lines represent the segmentation boundaries. Reproduced with permission from G. Bodenstein and H.M. Praetorius, Feature extraction from the electroencephalogram by adaptive segmentation, *Proceedings of the IEEE*, 65(5):642–652, 1977. ©IEEE.

dow and the reference window is significant, a segment boundary is drawn, and the procedure is restarted.

Let $\phi_R(k)$ be the ACF of the reference window at the beginning of a new segmentation step, where k is the lag or delay. Let $\phi_T(n, k)$ be the ACF of the test window positioned at time instant n . Given that the ACF for zero lag is the power of the signal, Michael and Houchin computed a normalized power distance $d_P(n)$ between the ACFs as (see also Appel and v. Brandt [394])

$$d_P(n) = \frac{|\sqrt{\phi_T(n, 0)} - \sqrt{\phi_R(0)}|}{\min\{\sqrt{\phi_T(n, 0)}, \sqrt{\phi_R(0)}\}}. \quad (8.27)$$

A spectral distance $d_F(n)$ was computed using the ACF coefficients only up to lag q as

$$d_F(n) = \frac{\sum_{k=1}^q |\phi_T(n, k) - \phi_R(k)|}{0.5 + \sum_{k=1}^q \min\{\sqrt{\phi_T(n, k)}, \sqrt{\phi_R(k)}\}}. \quad (8.28)$$

The lag limit q was set as the lower value of the lags at which the ACFs changed from positive to negative values for the first time. The net ACF distance $d(n)$ was then computed as

$$d(n) = \frac{d_P(n)}{Th_P} + \frac{d_F(n)}{Th_F}, \quad (8.29)$$

where Th_P and Th_F are thresholds. The condition $d(n) > 1$ was considered to represent a significant change in the ACF, and used to mark a segment boundary.

Due to the use of a moving window of finite size, the true boundary or point of change in the signal characteristics will lie within the last test window before a segment boundary is triggered. Michael and Houchin used a linear interpolation procedure based upon the steepness of the ACF distance measure to correct for such a

displacement. Barlow et al. [416] provide illustrations of application of the method to clinical EEGs. Their work includes clustering of similar segments based upon mean amplitude and mean frequency measures, “dendrograms” to illustrate the clustering of segments [417], and labeling of the various states found in an EEG record. Illustrations of application of the ACF method are provided in Section 8.5.4.

8.5.3 The generalized likelihood ratio

Appel and v. Brandt [393] proposed the generalized likelihood ratio (GLR) method, which uses a reference window that is continuously grown as long as no new boundary is marked. The test window is a sliding window of constant duration as in the case of the *SEM* and ACF methods. Figure 8.10 illustrates the windows used. The advantage of the growing reference window is that it contains the maximum amount of information available from the beginning of the new segment to the current instant. Three different data sets are defined: the growing reference window, the sliding test window, and a pooled window formed by concatenating the two. Distance measures are then derived using AR-model prediction errors computed for the three data sets.

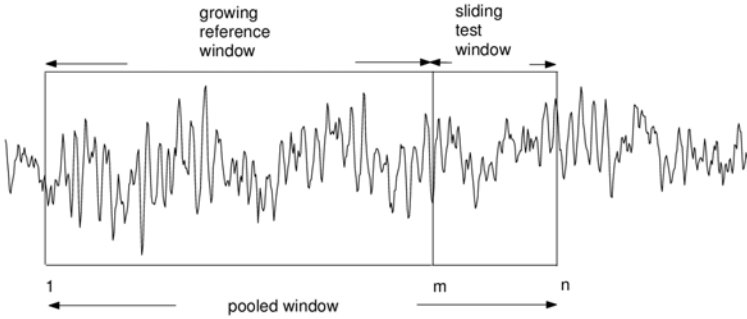


Figure 8.10 The growing reference window, the sliding test window, and the pooled window used in the GLR method for adaptive segmentation.

Let $\varepsilon(m : n)$ represent the prediction error energy (TSE ε as in Equation 7.19) within an arbitrary data set or window with boundaries m and n . The maximum log likelihood measure $H(m : n)$ for the window is defined as

$$H(m : n) = (n - m + 1) \ln \left[\frac{\varepsilon(m : n)}{(n - m + 1)} \right]. \quad (8.30)$$

Three measures are computed for the three data sets described above as $H(1 : m - 1)$ for the growing reference window, $H(m : n)$ for the test window, and $H(1 : n)$ for the composite or pooled window. Here, the reference window is denoted as commencing from the time instant or sample 1, m is the last sample of the growing reference window, and the current test window spans the duration from m to the current time instant n (see Figure 8.10). The GLR distance measure is defined as

$$d(n) = H(1 : n) - [H(1 : m - 1) + H(m : n)]. \quad (8.31)$$

Here, the first quantity represents the TSE if the test window is added to the growing reference window; the second quantity represents the TSE of the reference window grown so far; and the third quantity represents the TSE in modeling the test window itself. The measure $d(n)$ answers the question, “How much is the increase in the TSE if we add the test window to the growing reference window?”

Appel and v. Brandt [393] and Cohen [343] provide more details on the GLR. The GLR distance is a measure of the statistical similarity of the reference and test data sequences, with the assumption that their AR-model coefficients have a normal (Gaussian) distribution. The GLR distance is also a measure of the loss of information caused if no segment boundary is drawn at the position of the test window, that is, if it is assumed that the null hypothesis that the two sequences are similar is true.

Appel and v. Brandt [393] discuss issues related to the choice of the parameters involved in the GLR method, including the AR-model order, the test window length, and the threshold, on the GLR distance measure. The GLR method was also used by Willsky and Jones [418] to detect abrupt changes (sporadic anomalies and failures) in the variables of stochastic linear systems, as well as by Basseville and Benveniste [419] for segmentation of nonstationary signals (see also Cohen [343]). Illustrations of application of the GLR method are provided in Section 8.5.4.

8.5.4 Comparative analysis of the ACF, *SEM*, and GLR methods

Appel and v. Brandt [394] performed a comparative analysis of the performance of the ACF, *SEM*, and GLR methods of adaptive segmentation using synthesized signals as well as EEG signals. A simple two-pole system was used as the basis to simulate nonstationary signals. The gain, pole radius, and pole angle were individually varied back and forth between two sets of values. Several outputs of the dynamic system were computed with random signals (Gaussian-distributed white noise) as input. The signals were processed by the ACF, *SEM*, and GLR methods for adaptive segmentation. The variability of the segment boundaries detected for various realizations of the nonstationary (random) output signals for the same sequences of system parameters was analyzed.

Figure 8.11 shows the results related to variations in the angles of the poles, that is, in the resonance frequency of the system. The angle of the pole in the upper-half of the z -plane was changed from 20° to 40° and back at samples 200 and 400; the conjugate pole was also varied accordingly. The same changes were repeated at samples 700 and 800. The upper panel in the figure shows the pole positions and the related PSDs. The middle panel illustrates one sample of the 200 test signals generated: The higher-frequency characteristics of the signal related to the shifted pole positioned at 40° are evident over the intervals 200 – 400 and 700 – 800 samples. The lower panel illustrates the variability in the detected segment indices (dotted curve) and the estimated segment boundary positions (solid curves) for the three methods over 200 realizations of the test signals. (The true segment indices and boundaries are 1 : 200, 2 : 400, 3 : 700, and 4 : 800; ideally, the curves should exhibit steps at the points of change.) It is evident that the GLR method has provided the most consistent and accurate segmentation results, although at the price of increased computational load.

The *SEM* method has performed better than the ACF method, the latter showing the poorest results.

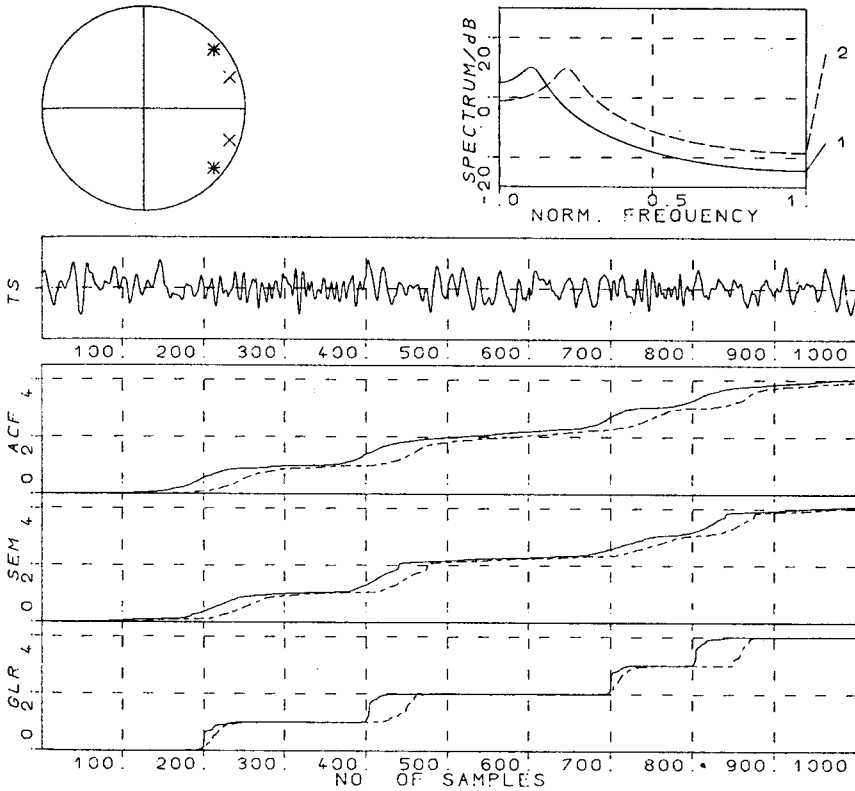


Figure 8.11 Comparative analysis of the ACF, *SEM*, and GLR methods for adaptive segmentation with the pole angle varied. Upper panel: pole positions and the related PSDs. *Note:* Norm. Frequency is normalized frequency such that the maximum frequency present in the sampled signal is unity. Middle panel: sample test signal; TS = time series. Lower panel: variability in the detected segment indices (dotted curve) and the estimated segment boundary positions (solid curves) for the three methods over 200 realizations of the test signals. See the text for more details. Reproduced with permission from U. Appel and A. v. Brandt, A comparative analysis of three sequential time series segmentation algorithms, *Signal Processing*, 6:45–60, 1984. ©Elsevier Science Publishers B.V. (North Holland).

Figure 8.12 shows the results related to variations in the distance of the poles from the origin, that is, in the bandwidth of the resonance frequency of the system. The distance of the poles from the origin was changed from 0.7 to 0.9 and back at samples 200 and 400. The same changes were repeated at samples 700 and 800. The PSDs display the increased prominence of the spectral peak when the poles are pushed toward the unit circle. The ACF method has not performed well in recognizing the

nonstationarities of this type in the test signals. The GLR method has performed better than the ACF method in segmentation.

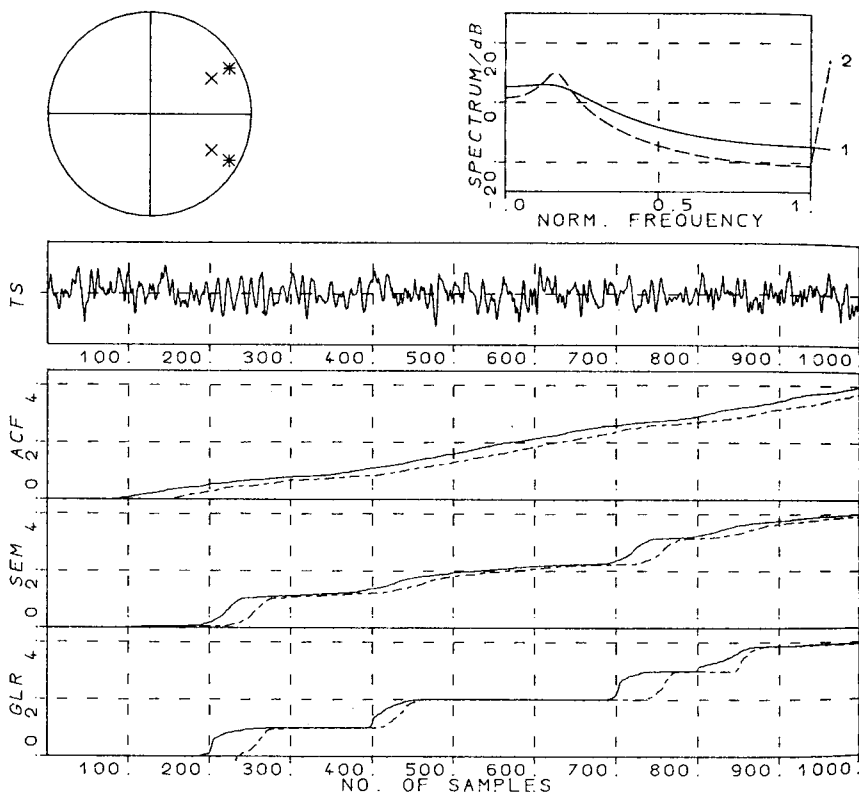


Figure 8.12 Comparative analysis of the ACF, *SEM*, and GLR methods for adaptive segmentation with the pole radius varied. Upper panel: pole positions and the related PSDs. *Note:* Norm. Frequency is normalized frequency such that the maximum frequency present in the sampled signal is unity. Middle panel: sample test signal; TS = time series. Lower panel: variability in the detected segment indices (dotted curve) and the estimated segment boundary positions (solid curves) for the three methods over 200 realizations of the test signals. See the text for more details. Reproduced with permission from U. Appel and A. v. Brandt, A comparative analysis of three sequential time series segmentation algorithms, *Signal Processing*, 6:45–60, 1984. ©Elsevier Science Publishers B.V. (North Holland).

Figure 8.13 shows the results of application of the three methods to an EEG signal. Although the exact locations where the signal changes its characteristics are not known for the EEG signal, the boundaries indicated by the GLR method appear to be the most accurate. It may be desirable in real-life applications to err on the side of superfluous segmentation; a subsequent clustering step could merge adjacent segments with similar model parameters.

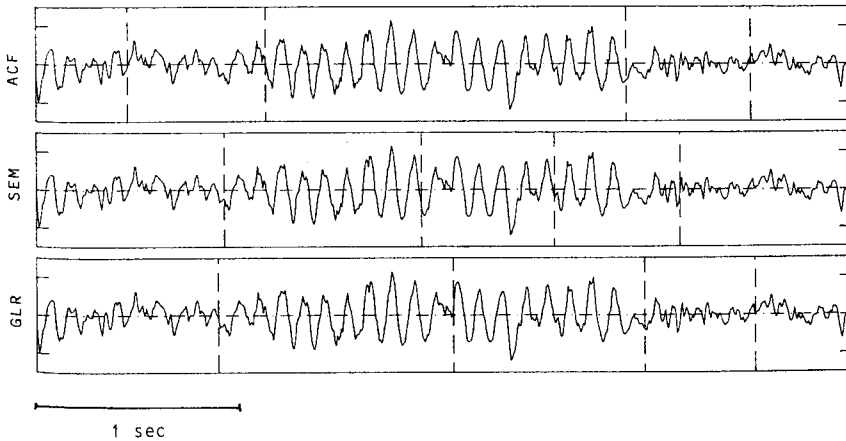


Figure 8.13 Comparative analysis of the ACF, *SEM*, and GLR methods for adaptive segmentation of an EEG signal. Reproduced with permission from U. Appel and A. v. Brandt, A comparative analysis of three sequential time series segmentation algorithms, *Signal Processing*, 6:45–60, 1984. ©Elsevier Science Publishers B.V. (North Holland).

8.6 Use of Adaptive Filters for Segmentation

We saw in Sections 3.9.2 and 3.9.3 that the coefficient (tap-weight) vectors of the adaptive LMS and RLS filters are expressed as functions of time. The filters adapt to changes in the statistics of the primary and reference signals. Could we, therefore, use the tap-weight vector $\mathbf{w}(n)$ to detect nonstationarities in a signal?

Problem: Investigate the potential use of the RLS adaptive filter for adaptive segmentation of nonstationary signals.

Solution: When we have only one signal to work with — the signal that is to be segmented — the question arises as to how we may provide two inputs, namely, the primary and reference signals, to the adaptive filter. If we assume that the signal to be segmented (applied at the primary input) was generated by an AR system, then we may provide the same signal with a delay as the reference input to the adaptive filter. The delay is to be set such that the reference input at a given instant of time is uncorrelated with the primary input; the delay may also be set on the basis of the order of the filter. (It is also possible to apply white noise at the reference input.) In essence, the adaptive filter then acts the role of an adaptive AR model. The filter tap-weight vector is continually adapted to changes in the statistics (ACF) of the input signal. The output represents the prediction error. Significant changes in the tap-weight vector or the prediction error may be used to mark points of prominent nonstationarities in the signal. Figure 8.14 shows a signal-flow diagram of the adaptive filter as described above; the filter structure is only slightly different from that in Figure 3.91.

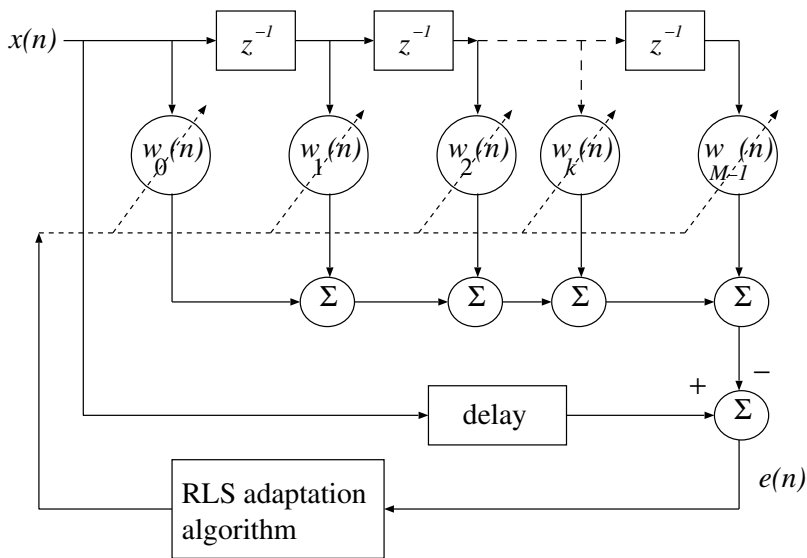


Figure 8.14 Adaptive RLS filter for segmentation of nonstationary signals. See also Figure 3.91.

8.6.1 Monitoring the RLS filter

The RLS filter as in Figure 8.14 attempts to predict the current signal sample from the available knowledge of the previous samples stored in the filter's memory units. If a large change occurs in the signal, the prediction error exhibits a correspondingly large value. In response, the adaptive filter's tap-weight vector is modified by the RLS algorithm.

Moussavi et al. [94] applied the RLS filter for segmentation of VAG signals. The order of the filter was set to be 5 in order to be low enough to detect transient changes and also to provide fast convergence. The forgetting factor was defined as $\lambda = 0.98$ so that the filter may be assumed to operate in an almost-stationary situation. The delay between the input and the reference input was set to be 7 samples (which corresponds to 3.5 ms with $f_s = 2$ kHz).

The adaptive segmentation algorithm of Moussavi et al. is as follows:

1. Initialize the RLS algorithm.
2. Find the squared Euclidean distance between the current tap-weight vector $\mathbf{w}(n)$ and the preceding vector $\mathbf{w}(n-1)$ as

$$\Delta(n) = |\mathbf{w}(n) - \mathbf{w}(n-1)|^2. \quad (8.32)$$

3. After computing $\Delta(n)$ for all samples of the signal available (in off-line processing), compute the *SD* of the $\Delta(n)$ values. Define a threshold as three times the *SD*.

4. Label all samples n for which $\Delta(n)$ exceeds the threshold as primary segment boundaries.
5. Compute the primary segment lengths (durations) as the differences between successive primary segment boundaries. Reject all primary segment boundaries that result in segment duration less than a preset minimum (defined in the work of Moussavi et al. [94] as 120 samples or 60 *ms*, corresponding to a knee-joint angle range of approximately 4°).
6. Mark the remaining boundary points as the final segment boundaries.

The main advantage of the RLS method is that there are no explicit reference and test windows as in the case of the ACF, *SEM*, and GLR methods. The RLS method computes a new filter tap-weight vector at each sample of the incoming signal. The method was found to perform well in the detection of gradual changes as well as sudden variations in VAG signals.

Illustration of application: Figures 8.15 and 8.16 illustrate the segmentation of the VAG signals of a normal subject and a patient with arthroscopically confirmed cartilage pathology, respectively. The figures also illustrate the spectrograms of the two signals. While the segmentation of the abnormal signal in Figure 8.16 may appear to be superfluous at first sight, close inspection of the corresponding spectrogram indicates that the spectral characteristics of the signal do change within short intervals. It is evident that the RLS method has detected the different types of nonstationarity present in the signals. Moussavi et al. [94] tested the method with 46 VAG signals and observed that the segmentation boundaries agreed well with the nature of the joint sounds heard via auscultation with a stethoscope as well as with the spectral changes observed in the spectrograms of the signals.

8.6.2 The RLS lattice filter

In order to apply the RLS method for adaptive segmentation in a nonstationary environment, it is necessary to solve the least-squares problem recursively and rapidly. The *recursive least-squares lattice* (RLSL) algorithm is well suited for such purposes. Since the RLSL method uses a lattice filter, and is based upon forward and backward prediction and time-varying reflection coefficients, it is necessary to define some of the related procedures.

Forward and backward prediction: Let us rewrite Equation 7.17 related to LP or AR modeling as

$$\tilde{y}(n) = - \sum_{k=1}^M a_{M,k} y(n-k), \quad (8.33)$$

with the inclusion of the order of the model M as a subscript for the model coefficients a_k . In this procedure, M past samples of the signal, $y(n-1), y(n-2), \dots, y(n-M)$, are used in a linear combination to predict the current sample

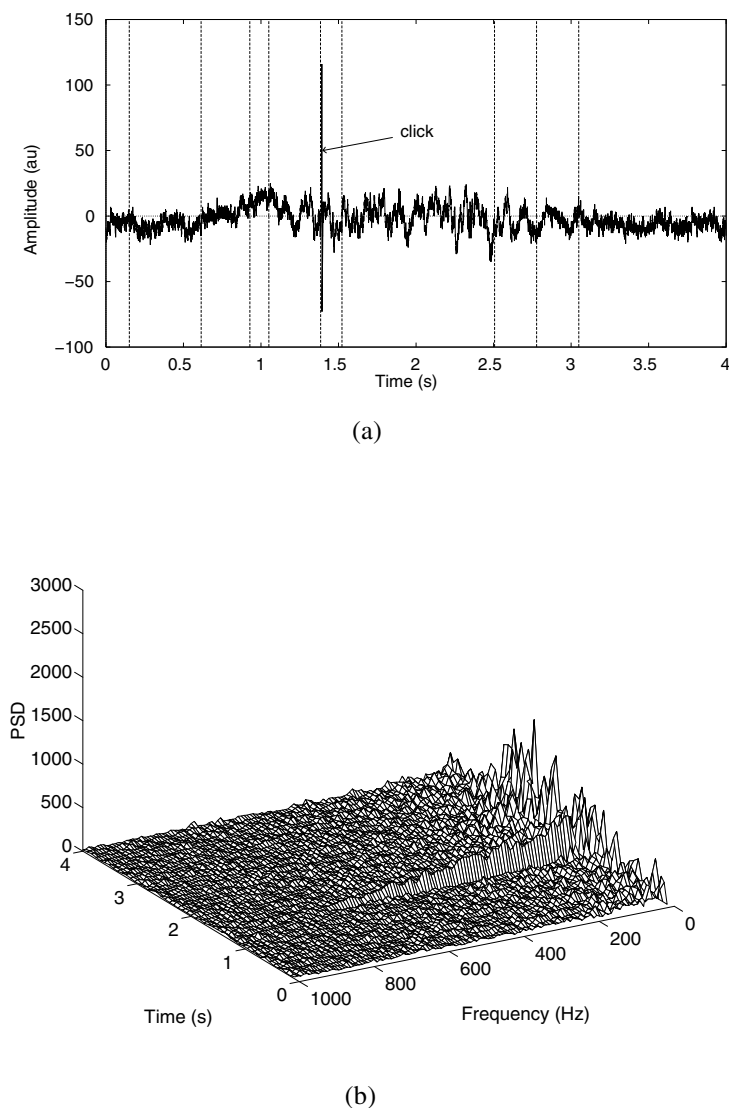
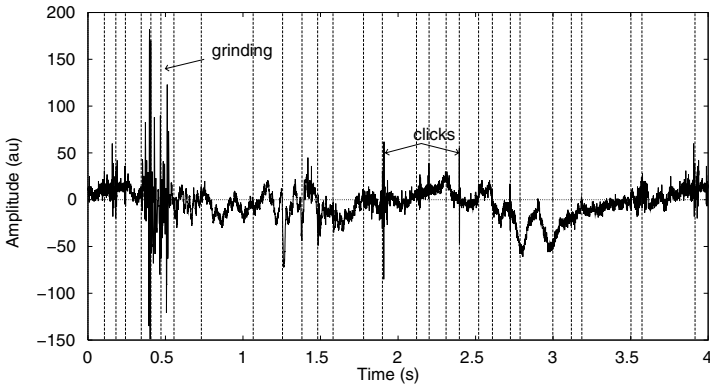
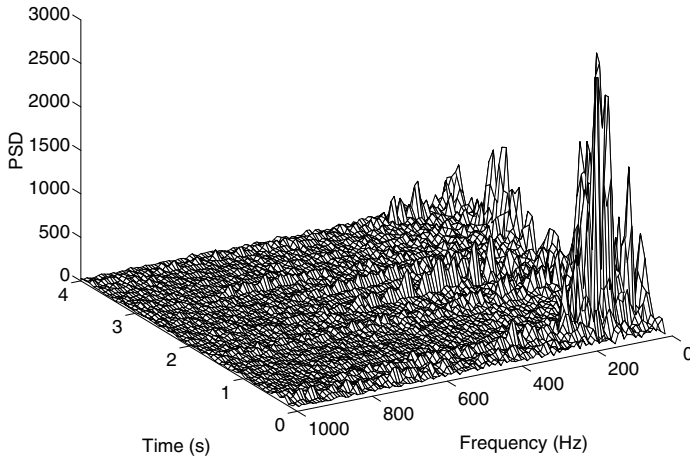


Figure 8.15 (a) Segmentation of the VAG signal of a normal subject using the RLS method. A click heard in auscultation of the knee joint is labeled. (b) Spectrogram (STFT) of the signal. Reproduced with permission from Z.M.K. Moussavi, R.M. Rangayyan, G.D. Bell, C.B. Frank, K.O. Ladly, and Y.T. Zhang, Screening of vibroarthrographic signals via adaptive segmentation and linear prediction modeling, *IEEE Transactions on Biomedical Engineering*, 43(1):15–23, 1996. ©IEEE.



(a)



(b)

Figure 8.16 (a) Segmentation of the VAG signal of a subject with cartilage pathology using the RLS method. Clicking and grinding sounds heard during auscultation of the knee joint are labeled. (b) Spectrogram (STFT) of the signal. Reproduced with permission from Z.M.K. Moussavi, R.M. Rangayyan, G.D. Bell, C.B. Frank, K.O. Ladly, and Y.T. Zhang, Screening of vibroarthrographic signals via adaptive segmentation and linear prediction modeling, *IEEE Transactions on Biomedical Engineering*, 43(1):15–23, 1996. ©IEEE.

$y(n)$ in the *forward* direction. The *forward prediction error* is

$$e_{M,f}(n) = y(n) - \tilde{y}(n) = \sum_{k=0}^M a_{M,k} y(n-k), \quad (8.34)$$

with $a_{M,0} = 1$. This equation is a restatement of Equation 7.18 with the inclusion of the order of the model M as a subscript for the error e as well as the subscript f to indicate that the prediction is being performed in the forward direction.

The term *backward prediction* refers to the estimation of $y(n-M)$ from the samples $y(n), y(n-1), \dots, y(n-M+1)$ as

$$\tilde{y}(n-M) = - \sum_{k=0}^{M-1} a_{M,k}^{\#} y(n-k), \quad (8.35)$$

where $a_{M,k}^{\#}$ are the backward prediction coefficients. Application of the least-squares method described in Section 7.5 for a stationary signal leads to the result

$$a_{M,k}^{\#} = a_{M,M-k}, \quad k = 0, 1, 2, \dots, M, \quad (8.36)$$

that is, the backward prediction coefficients are the same as the forward prediction coefficients, but in reverse order [148]. The *backward prediction error* is, therefore, given by

$$\begin{aligned} e_{M,b}(n) &= y(n-M) - \tilde{y}(n-M) \\ &= \sum_{k=0}^M a_{M,k}^{\#} y(n-k) = \sum_{k=0}^M a_{M,M-k} y(n-k). \end{aligned} \quad (8.37)$$

The Burg-lattice method: The Burg-lattice method [148] is based on minimizing the sum of the squared forward and backward prediction errors. Assuming that the input $y(n)$ is ergodic, the *performance index* ξ_m is given by

$$\xi_m = \sum_{n=m+1}^N [e_{m,f}^2(n) + e_{m,b}^2(n)], \quad (8.38)$$

where $e_{m,f}(n)$ is the forward prediction error and $e_{m,b}(n)$ is the backward prediction error, with the model order m being recursively updated as $m = 1, 2, \dots, M$. The length of the available block of data is N samples.

If we use the Levinson–Durbin method to estimate the forward prediction coefficients, we get (see Section 7.5 and Equation 7.38)

$$a_{m,k} = a_{m-1,k} + \gamma_m a_{m-1,m-k}, \quad (8.39)$$

where γ_m is the reflection coefficient for order m . Similarly, for the case of backward prediction, we get

$$a_{m,m-k} = a_{m-1,m-k} + \gamma_m a_{m-1,m}, \quad (8.40)$$

including the substitution $a_{m,k}^{\#} = a_{m,m-k}$.

Combining the relationships in Equations 8.34, 8.37, 8.39, and 8.40 leads to the lattice structure for computation of the forward and backward prediction errors, where the two prediction error series are interrelated recursively as [148]

$$e_{m,f}(n) = e_{m-1,f}(n) + \gamma_m e_{m-1,b}(n-1) \quad (8.41)$$

and

$$e_{m,b}(n) = e_{m-1,b}(n-1) + \gamma_m e_{m-1,f}(n). \quad (8.42)$$

(All coefficients are assumed to be real-valued in this derivation; Haykin [148] allows for all coefficients to be complex-valued.) Figure 8.17 illustrates a basic unit of the lattice structure that performs the recursive operations in Equations 8.41 and 8.42. The reflection coefficient γ_m may be chosen so as to minimize the performance index given in Equation 8.38, that is, by setting

$$\frac{\partial \xi_m}{\partial \gamma_m} = 2 \sum_{n=m+1}^N \left[e_{m,f}(n) \frac{\partial e_{m,f}(n)}{\partial \gamma_m} + e_{m,b}(n) \frac{\partial e_{m,b}(n)}{\partial \gamma_m} \right] = 0. \quad (8.43)$$

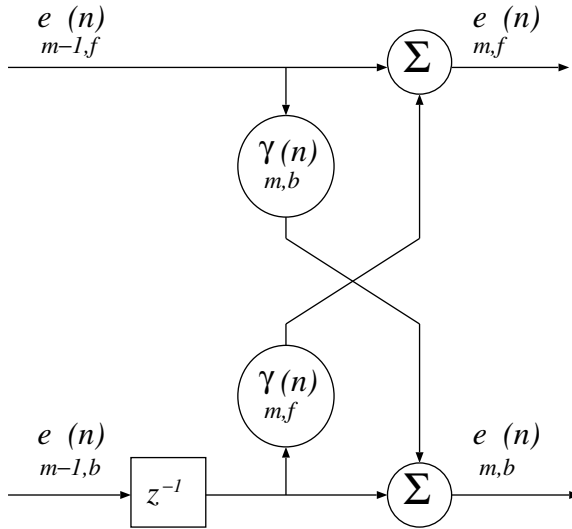


Figure 8.17 Basic unit of the lattice structure that performs the recursive operations in Equations 8.41 and 8.42 as well as the recursive operations in Equations 8.52 and 8.53. In the case of the former, due to the stationarity of the processes involved, the forward and backward reflection coefficients are the same and are independent of time (γ_m). Adapted with permission from S. Krishnan, Adaptive Signal Processing Techniques for Analysis of Knee Joint Vibroarthrographic Signals, Ph.D. Thesis, University of Calgary, 1999.

Partial differentiation of Equations 8.41 and 8.42 with respect to γ_m yields

$$\frac{\partial e_{m,f}(n)}{\partial \gamma_m} = e_{m-1,b}(n-1) \quad (8.44)$$

and

$$\frac{\partial e_{m,b}(n)}{\partial \gamma_m} = e_{m-1,f}(n). \quad (8.45)$$

Substituting the results above in Equation 8.43, we get

$$\sum_{n=m+1}^N [e_{m,f}(n) e_{m-1,b}(n-1) + e_{m,b}(n) e_{m-1,f}(n)] = 0. \quad (8.46)$$

Substituting Equations 8.41 and 8.42 in Equation 8.46, we get

$$\begin{aligned} \sum_{n=m+1}^N & [\{e_{m-1,f}(n) + \gamma_m e_{m-1,b}(n-1)\} e_{m-1,b}(n-1) \\ & + \{e_{m-1,b}(n-1) + \gamma_m e_{m-1,f}(n)\} e_{m-1,f}(n)] = 0. \end{aligned} \quad (8.47)$$

The reflection coefficients γ_m can then be calculated as

$$\gamma_m = -2 \frac{\sum_{n=m+1}^N e_{m-1,f}(n) e_{m-1,b}(n-1)}{\sum_{n=m+1}^N [e_{m-1,f}^2(n) + e_{m-1,b}^2(n-1)]}. \quad (8.48)$$

The magnitudes of the reflection coefficients are less than unity. The Burg formula always yields a minimum-phase design for the lattice predictor.

The prediction coefficients or the AR-model parameters can be computed from the reflection coefficients by using the relationship in Equation 8.39. The order m is updated recursively as $m = 1, 2, \dots, M$, with $a_{m,0} = 1$, and $a_{m-1,k} = 0$ for $k > m-1$. From Equation 8.39 and Figure 8.17, it can be observed that the AR coefficients can be computed for any model order by simply adding one or more lattice stages without affecting the earlier computations for lower orders. This is one of the main advantages of the Burg-lattice AR-modeling algorithm, especially in situations where the order of the system being modeled is not known in advance.

RLSL algorithm for adaptive segmentation: A general schematic representation of the RLSL filter structure is given in Figure 8.18. Two levels of updating are used in the RLSL algorithm:

1. *Order-update:* This involves updating the forward prediction error $e_{m,f}(n)$, the backward prediction error $e_{m,b}(n)$, the forward prediction error power $\varepsilon_{m,f}(n)$, and the backward prediction error power $\varepsilon_{m,b}(n)$. Here, m indicates the model order and n indicates the time instant.
2. *Time-update:* This involves time-updating of the parameters that ensure adaptation, including the forward reflection coefficients $\gamma_{m,f}(n)$ and backward reflection coefficients $\gamma_{m,b}(n)$. Note that, in the general nonstationary environment, $\gamma_{m,f}(n) \neq \gamma_{m,b}(n)$.

Order-updating and time-updating together enable the RLSL algorithm to achieve fast convergence and excellent tracking capability.

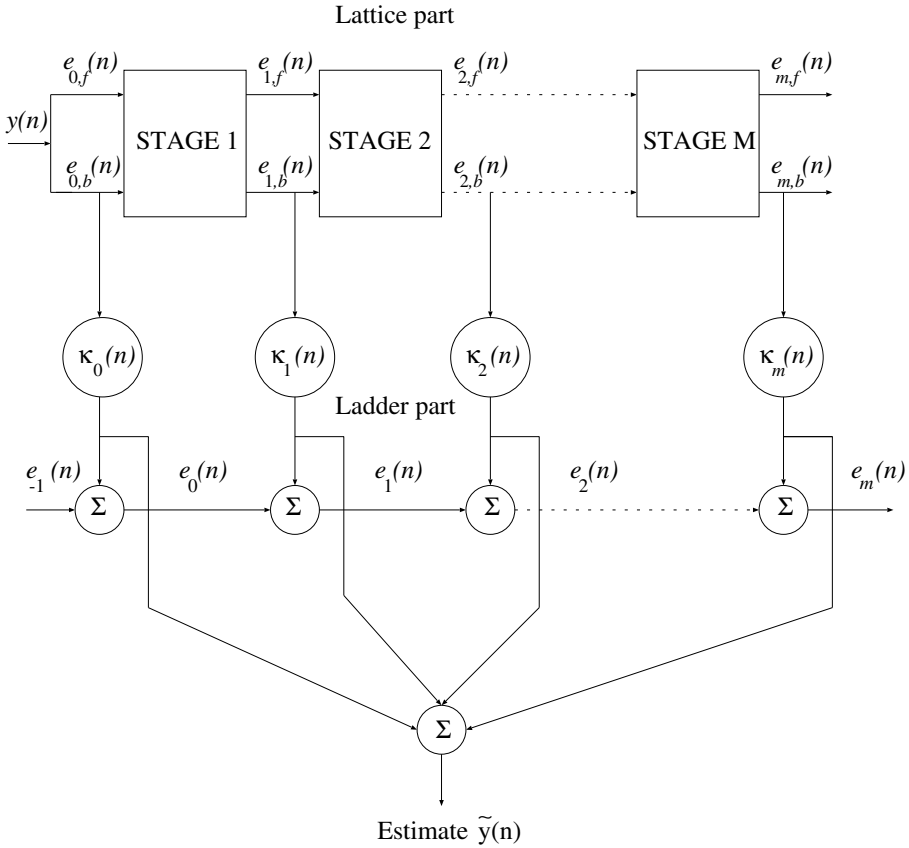


Figure 8.18 General schematic representation of the RLSL filter structure for adaptive segmentation of nonstationary signals. Adapted with permission from S. Krishnan, Adaptive Signal Processing Techniques for Analysis of Knee Joint Vibroarthrographic Signals, Ph.D. Thesis, University of Calgary, 1999.

The RLSL algorithm can be expressed in three stages [148, 185, 187]:

1. *Initialization of the algorithm and lattice for filter order M* : The parameters of the algorithm are initialized at $n = 0$ and for each order $m = 1, 2, \dots, M$ by setting the forward prediction error power $\varepsilon_{m-1,f}(0)$ and the backward prediction error power $\varepsilon_{m-1,b}(0)$ equal to a small positive constant; the forward reflection coefficients $\gamma_{m,f}(0) = 0$; the backward reflection coefficients $\gamma_{m,b}(0) = 0$; the conversion factor $\gamma_{0,c}(0) = 1$; and an auxiliary variable $\Delta_{m-1}(0) = 0$.

For each time instant $n \geq 1$, the following zeroth-order variables are generated: the forward prediction error $e_{0,f}(n)$ equal to the data input $y(n)$; the backward

prediction error $e_{0,b}(n) = y(n)$; $\varepsilon_{0,f}(n) = \varepsilon_{0,b}(n) = \lambda \varepsilon_{0,f}(n-1) + |y(n)|^2$, where λ is the forgetting factor, and $\gamma_{0,c}(n) = 1$.

The variables involved in joint process estimation, for each order $m = 0, 1, \dots, M$ at time $n = 0$, are initialized by setting the scalar $\rho_m(0) = 0$, and for each instant $n \geq 1$ the zeroth-order variable of *a priori* estimation error $e_0 = d(n)$, where $d(n)$ is the desired response of the system.

2. *Prediction part of the RLSL algorithm:* For $n = 1, 2, \dots, N_s$, where N_s is the number of signal samples available, the various order-updates are computed in the sequence $m = 1, 2, \dots, M$, where M is the final order of the least squares predictor, as follows:

$$\Delta_{m-1}(n) = \lambda \Delta_{m-1}(n-1) + \frac{e_{m-1,b}(n-1) e_{m-1,f}(n)}{\gamma_{m-1,c}(n-1)}, \quad (8.49)$$

where $\Delta_{m-1}(n)$ is the cross-correlation between the delayed backward prediction error $e_{m-1,b}(n-1)$ and the forward prediction error $e_{m-1,f}(n)$ of the lattice filter.

The forward reflection coefficient $\gamma_{m,f}(n)$ is then updated as

$$\gamma_{m,f}(n) = -\frac{\Delta_{m-1}(n)}{\varepsilon_{m-1,b}(n)}. \quad (8.50)$$

Similarly, the backward reflection coefficient is updated as

$$\gamma_{m,b}(n) = -\frac{\Delta_{m-1}(n)}{\varepsilon_{m-1,f}(n-1)}. \quad (8.51)$$

In general, $\varepsilon_{m-1,f}(n)$ and $\varepsilon_{m-1,b}(n-1)$ are unequal, so that in the RLSL algorithm, unlike in the Burg algorithm described earlier in this section, we have $\gamma_{m,f}(n) \neq \gamma_{m,b}(n)$.

From the lattice structure as described earlier in the context of Equations 8.41 and 8.42 and depicted in Figure 8.17, and noting that the reflection coefficients $\gamma_{m,f}(n)$ and $\gamma_{m,b}(n)$ are now different and time-variant parameters, we can write the order-update recursion of the forward prediction error as (see Figure 8.17)

$$e_{m,f}(n) = e_{m-1,f}(n) + \gamma_{m,f}(n) e_{m-1,b}(n-1), \quad (8.52)$$

and the order-update recursion of the backward prediction error as

$$e_{m,b}(n) = e_{m-1,b}(n-1) + \gamma_{m,b}(n) e_{m-1,f}(n). \quad (8.53)$$

The prediction error powers are updated as

$$\varepsilon_{m,f}(n) = \varepsilon_{m-1,f}(n) + \gamma_{m,f}(n) \Delta_{m-1}(n) \quad (8.54)$$

and

$$\varepsilon_{m,b}(n) = \varepsilon_{m-1,b}(n-1) + \gamma_{m,b}(n) \Delta_{m-1}(n). \quad (8.55)$$

The conversion factor $\gamma_{m,c}(n-1)$ is updated as

$$\gamma_{m,c}(n) = \gamma_{m-1,c}(n) - \frac{e_{m-1,b}^2(n)}{\varepsilon_{m-1,b}(n)}. \quad (8.56)$$

The equations in this step constitute the basic order-update recursions for the RLSL predictor. The recursions generate two sequences of prediction errors: the forward prediction error and the backward prediction error. The two error sequences play key roles in the recursive solution of the linear least-squares problem.

3. *Filtering part of the RLSL algorithm:* For $n = 1, 2, \dots, N_s$, the various order-updates are computed in the sequence $m = 0, 1, \dots, M$ as follows:

$$\rho_m(n) = \lambda \rho_m(n-1) + \frac{e_{m,b}(n)}{\gamma_{m,c}(n)} e_{m-1}(n). \quad (8.57)$$

The regression coefficients $\kappa_m(n)$ of the joint process estimator are defined in terms of the scalar $\rho_m(n)$ as

$$\kappa_m(n) = \frac{\rho_m(n)}{\varepsilon_{m,b}(n)}. \quad (8.58)$$

The order-update recursion of the *a posteriori* estimation error $e_m(n)$ is then given as

$$e_m(n) = e_{m-1}(n) - \kappa_m(n) e_{m,b}(n). \quad (8.59)$$

The dynamics of the input signal, that is, the statistical changes occurring in the signal, are reflected in the lattice filter parameters. Parameters such as the reflection coefficients (γ_f and γ_b) and the *MS* value of the estimation error {that is, $E[e_m^2(n)]$ } may therefore be used to monitor the statistical changes.

The conversion factor γ_c that appears in the algorithm can be used as a good statistical detection measure of the “unexpectedness” of the recent data samples. As long as the data belong to the same distribution, the variable γ_c will be near unity. If the recent data samples belong to a different distribution, γ_c will tend to fall from unity. This will cause the factor $\frac{1}{\gamma_c}$ appearing in the time-update formula (Equation 8.49) to be large, which leads to abrupt changes in the lattice parameters. The quantities γ_c , $\frac{1}{\gamma_c}$, or $\frac{1}{1-\gamma_c}$ may be used for fast tracking of changes in the input data and to test for segment boundaries in a nonstationary environment.

Illustration of application: The advantage in using the RLSL filter for segmentation of VAG signals is that the statistical changes in the signals are well reflected in the filter parameters, and hence segment boundaries can be detected by monitoring any one of the filter parameters, such as the MSE, conversion factor, or the reflection coefficients. Krishnan et al. [95] and Krishnan [185] used the conversion factor (γ_c)

to monitor statistical changes in VAG signals. In a stationary environment, γ_c starts with a low initial value and remains small during the early part of the initialization period. After a few iterations, γ_c begins to increase rapidly toward the final value of unity. In the case of nonstationary signals, such as VAG, γ_c will fall from its steady-state value of unity whenever a change occurs in the statistics of the signal. This can be used in segmenting VAG signals into quasistationary components. The segmentation procedure proposed by Krishnan et al. [95] and Krishnan [185] is summarized as follows:

1. The VAG signal is passed twice through the segmentation filter: The first pass is used to allow the filter to converge, and the second pass is used to test the γ_c value at each sample against a threshold value for the detection of segment boundaries.
2. Whenever γ_c at a particular sample during the second pass is less than the threshold, a primary segment boundary is marked.
3. If the difference between two successive primary segment boundaries is less than the minimum desired segment length (120 samples in the work of Krishnan et al.), the later of the two boundaries is deleted.

Figures 8.19 and 8.20 show the results of application of the RLSL segmentation method to two VAG signals. Plots of $\gamma_c(n)$ are also included in the figures. It may be observed that the value of $\gamma_c(n)$ drops whenever there is a significant change in the characteristics of the signal. Whereas the direct application of a threshold on $\gamma_c(n)$ would result in superfluous segmentation, inclusion of the condition on the minimum segment length that is meaningful in the application is seen to provide practically useful segmentation. The number of segments was observed to be, on the average, eight segments per VAG signal. Signals of patients with cartilage pathology were observed to result in more segments than normal signals.

An advantage of the RLSL method of adaptive segmentation is that a fixed threshold may be used; Krishnan et al. [95] found a fixed threshold value of 0.9985 to give good segmentation results with VAG signals. The adaptive segmentation procedure was found to provide segments that agreed well with manual segmentation based upon auscultation and/or arthroscopy. Adaptive analysis of VAG signals is described further in Section 9.12.

8.7 Wavelets and Time frequency Analysis

Biomedical signals are often mixtures of multiple components with time-varying properties; to complicate matters, such components may have varying durations and overlap one another. For example, PCG signals could include not only S1 and S2 but also murmurs and opening snaps of valves. Furthermore, S1 and S2 are themselves made up of multiple overlapping components, murmurs are noise-like high-frequency sounds, and opening snaps are transients. We have seen that VAG signals

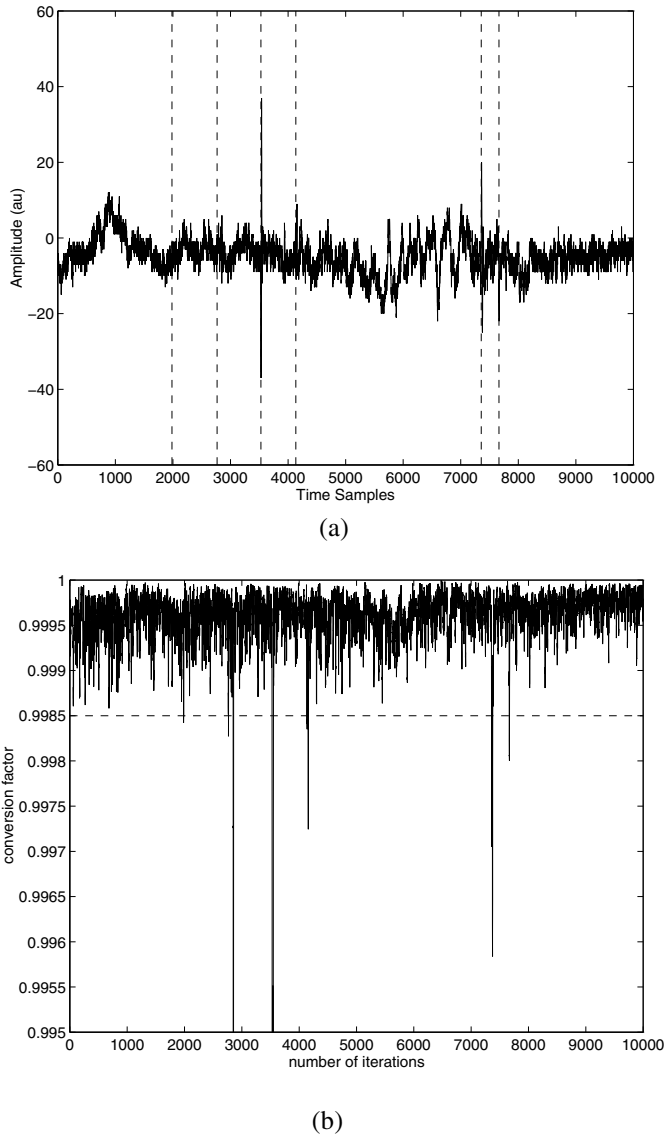


Figure 8.19 (a) VAG signal of a normal subject with the final segment boundaries given by the RLSL method shown by vertical dashed lines. (b) Plot of the conversion factor $\gamma_c(n)$; the horizontal dashed line represents the fixed threshold used to detect segment boundaries. The duration of the signal is 5 s, with $f_s = 2 \text{ kHz}$. Reproduced with permission from S. Krishnan, R.M. Rangayyan, G.D. Bell, C.B. Frank, and K.O. Ladly, Adaptive filtering, modelling, and classification of knee joint vibroarthrographic signals for non-invasive diagnosis of articular cartilage pathology, *Medical and Biological Engineering and Computing*, 35(6):677–684, 1997. ©IFMBE.

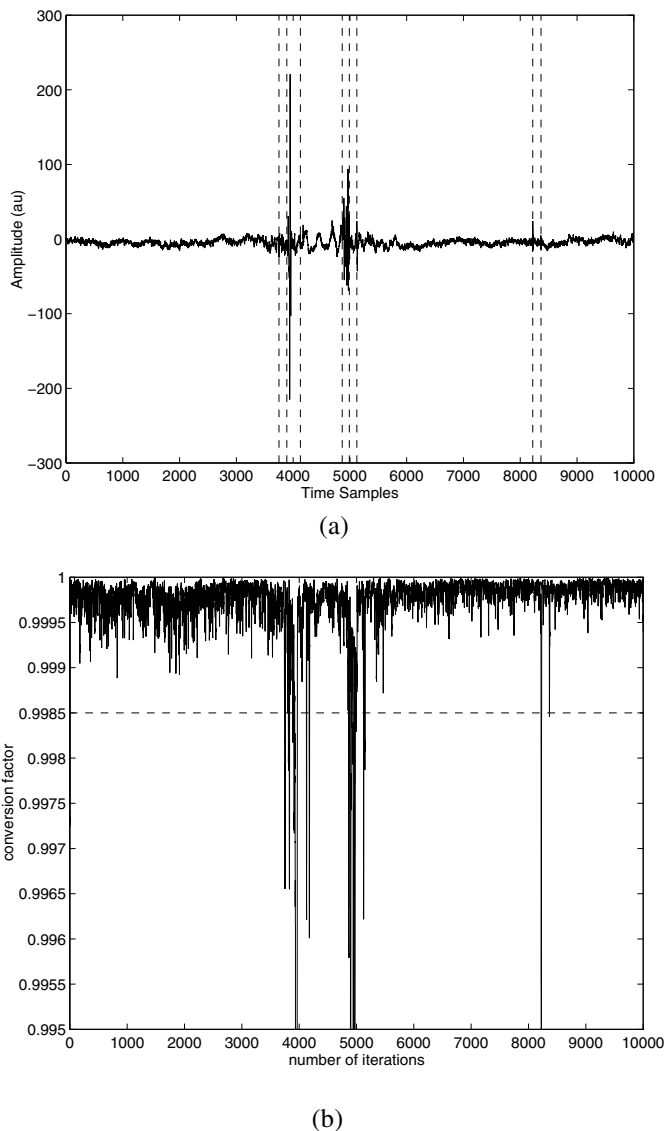


Figure 8.20 (a) VAG signal of a subject with cartilage pathology, with the final segment boundaries given by the RLSL method shown by vertical dashed lines. (b) Plot of the conversion factor $\gamma_c(n)$; the horizontal dashed line represents the fixed threshold used to detect segment boundaries. The duration of the signal is 5 s, with $f_s = 2 \text{ kHz}$. Reproduced with permission from S. Krishnan, R.M. Rangayyan, G.D. Bell, C.B. Frank, and K.O. Ladly, Adaptive filtering, modelling, and classification of knee joint vibroarthrographic signals for non-invasive diagnosis of articular cartilage pathology, *Medical and Biological Engineering and Computing*, 35(6):677–684, 1997. ©IFMBE.

could have multiple components related to various types and stages of deterioration of knee-joint surfaces causing sustained grinding sounds or transient clicks. When analyzing such signals, it would be desirable to obtain a joint distribution of their temporal and frequency-domain characteristics. The STFT and the various segmentation-based approaches described in the preceding sections of the present chapter are approaches that provide information of this nature, but impose several limitations.

A basic limitation of Fourier analysis is that the sinusoidal signals used as the basis functions are of infinite duration: The integral (see Equation 3.86) or summation (see Equation 3.75) used to compute the projection of the given signal on to each of the basis functions is defined over all time (or the entire duration of the given signal). Therefore, the properties of the given signal are averaged over its duration. Such global averaging provides good frequency resolution but no temporal resolution: The procedure obscures transients and components or events that exist over short durations of time. The STFT provides some improvement in temporal resolution by using short windows along the time axis but suffers from related limitations as described in Sections 8.4.1 and 8.4.2. Another limitation of Fourier analysis is that the basis function is limited to sinusoids; this limitation applies also to the STFT.

Problem: How can we overcome the limitations of Fourier analysis and study the time-varying characteristics of nonstationary and multicomponent signals with improved accuracy and flexibility? How can we tailor the tool or procedure of analysis to adapt to the particular nature of a given signal?

Solution: The answers to the questions stated above are provided by wavelets and methods for joint TF analysis that were developed over the past few decades and have gained substantial popularity in the analysis of biomedical signals [286, 406, 408, 420–429]. TFDs may be used for representation and analysis of nonstationary signals; furthermore, TF and other decomposition methods may be used for parameterization of signals and feature extraction for classification. The following sections provide brief descriptions of wavelet analysis, TFDs, and signal decomposition methods with illustrations of their application.

8.7.1 Approximation of a signal using wavelets

The analysis of a signal using a transform may be viewed as decomposition of the given signal in the form of a weighted combination of the basis functions that define the transform (see Equation 3.75). For example, in Fourier analysis, the given signal is expressed as a weighted combination of mutually orthonormal sinusoids (see Equation 3.76). If only some of the basis functions are retained and others are rejected (or given selective weighting), we obtain an approximation of the original signal, commonly referred to as a filtered version of the original signal. We will now generalize this concept and extend its definition and application.

In linear approximation, the given signal is projected over M mutually orthogonal basis functions that are predetermined or chosen a priori. A linear approximation of a signal $x(t)$ may be written as

$$\tilde{x}(t) = \sum_{m=0}^{M-1} \langle x, \psi_m \rangle \psi_m(t), \quad (8.60)$$

where $\langle x, \psi_m \rangle$ denotes the projection, inner product, or dot product of $x(t)$ with one of the M orthogonal basis vectors $\psi_m(t)$; see also Equation 3.80. An optimal linear approximation is provided by the Karhunen–Loève transform (KLT) [430]; see Section 8.8.1.

An adaptive approximation may be achieved by tailoring the M orthogonal basis functions to the properties of the given signal. The use of signal-adaptive basis functions leads to nonlinear decomposition, which may be expressed as

$$\tilde{x}(t) = \sum_{m \in B_M} \langle x, \psi_m \rangle \psi_m(t), \quad (8.61)$$

where B_M denotes a group of basis functions from a dictionary that provides the first M inner-product values $\langle x, \psi_m \rangle$ arranged in decreasing order. The functions in B_M are chosen as functions that correlate best with $x(t)$ and may be interpreted as the main features of $x(t)$. Instead of sinusoids as in the Fourier transform, one could define a set of wavelets that facilitate analysis of the characteristics of a certain type of signals (see Section 8.7.6 for examples of wavelets).

Filtering may be performed using wavelets by approximating the given signal with a small number of nonzero wavelet coefficients. If $w_m = \langle x, \psi_m \rangle$ denotes the wavelet coefficients, hard thresholding is performed as

$$\tilde{w}_m = \begin{cases} w_m & \text{if } |w_m| \geq T, \\ 0 & \text{if } |w_m| < T, \end{cases} \quad (8.62)$$

where T denotes a threshold value. Soft thresholding is given by

$$\tilde{w}_m = \begin{cases} \text{sgn}(w_m)(|w_m| - T) & \text{if } |w_m| \geq T, \\ 0 & \text{if } |w_m| < T. \end{cases} \quad (8.63)$$

The filtered signal (or the approximation) is given by

$$\tilde{x}(t) = \sum_{m \in B_M} \tilde{w}_m \psi_m(t). \quad (8.64)$$

The wavelet transform provides an approximation as above, where the basis functions or wavelets are obtained by dilating and translating a prototype function, $\psi(t)$, also known as the mother wavelet, written as

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right), \quad (8.65)$$

where s denotes the dilation or scale parameter and τ is the translation or shift parameter. The parameter τ represents the position on the time axis where the wavelet ψ is placed; the notion of scale represented by the parameter s is related to the concept of frequency or a band of frequencies. Formally, the continuous wavelet transform (CWT) is defined as

$$X(\tau, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t) \psi^*\left(\frac{t-\tau}{s}\right) dt. \quad (8.66)$$

The integral in the CWT equation represents the projection of the given signal $x(t)$ on to the scaled wavelet with a shift so as to be positioned at $t = \tau$. The coefficient $X(\tau, s)$ resulting from the projection indicates the commonality between the two functions or the extent of the wavelet that is present in the signal being analyzed. The CWT equation also represents analysis of the signal using several wavelets so as to derive the corresponding projections or coefficients $X(\tau, s)$.

In practice, the CWT may be implemented with suitable discretization of the (τ, s) space. Typically, the time axis is discretized to the same sampling interval as that of the signal being processed. Due to substantial overlap between the shifted and scaled wavelets used, the CWT results in a redundant representation of the signal being analyzed. The (τ, s) space is referred to as the time–scale space and facilitates TF analysis if the relationship between the scaled wavelets and their spectral characteristics are taken into account (see Section 8.7.6).

Functions to be used as wavelets are required to possess certain properties. The property of finite energy is expressed as

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty. \quad (8.67)$$

The admissibility condition is stated as

$$C_\psi = \int_0^\infty \frac{|\Psi(f)|^2}{f} df < \infty, \quad (8.68)$$

where $\Psi(f)$ is the Fourier transform of $\psi(t)$ and C_ψ is known as the admissibility constant. If the wavelet is complex, its Fourier transform should be real and be zero for negative frequencies.

The inverse CWT is defined as

$$x(t) = \frac{1}{C_\psi} \int_{\tau=-\infty}^{\infty} \int_{s=0}^{\infty} X(\tau, s) \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right) d\tau \frac{ds}{s^2}. \quad (8.69)$$

The integral in the inverse CWT represents reconstruction or synthesis of the signal $x(t)$ using a weighted combination of scaled and shifted wavelets $\psi_{\tau,s}(\tau)$; each wavelet is weighted by the corresponding CWT coefficient $X(\tau, s)$.

The DWT is defined as

$$X(m, n) = \int_{-\infty}^{\infty} x(t) \psi_{m,n}^*(t) dt \quad (8.70)$$

with

$$\psi_{m,n}(t) = \frac{1}{\sqrt{s_0^m}} \psi \left(\frac{t - n \tau_0 s_0^m}{s_0^m} \right), \quad (8.71)$$

where the integers m and n control dilation and translation of the wavelet on a discrete grid; the parameters s_0 and τ_0 are the step parameters for dilation and translation, respectively. (In the equation given above, only the wavelet parameters have been shown with discrete variables; if the input signal has been sampled, the functions x and ψ need to be changed to discrete-time notation and the integral should be changed to summation.) The dyadic grid is obtained by setting $s_0 = 2$ and $\tau_0 = 1$; then, we have

$$\psi_{m,n}(t) = \frac{1}{\sqrt{2^m}} \psi (2^{-m}t - n). \quad (8.72)$$

The use of an appropriate grid for the DWT reduces the redundancy of representation as compared to the CWT. (The critically sampled DWT is a shift-variant transform; see Bradley [431] for related discussions.) The wavelet coefficients $X(m, n)$ obtained as above may be subjected to thresholding or selection procedures and an approximation of the signal may be reconstructed as

$$\tilde{x}(t) = \sum_m \sum_n \tilde{X}(m, n) \psi_{m,n}(t), \quad (8.73)$$

where $\tilde{}$ indicates the result of approximation or filtering and the range of summation depends upon the choices made and the application.

The wavelet approach permits simultaneous analysis of the local characteristics of a given nonstationary signal in both the temporal and spectral domains. The method facilitates the analysis of short-duration high-frequency signal components as well as long-duration low-frequency components. A multicomponent signal may be analyzed using multiple wavelets that match the expected characteristics of the various components so as to extract measures related to their properties in separate channels. Several types of wavelets and related procedures have been described in the literature [286, 406, 408, 420–429]; the presentation in this chapter is limited to a few concepts that can be related to the remaining topics in the chapter. The use of appropriate wavelets can assist in the identification and analysis of transient, aperiodic, multicomponent, and nonstationary signals.

The approach of selecting the best basis among a dictionary of bases by minimizing a cost function is known as the method of wavelet packets (WP) [432]. The WP approach uses a family of orthogonal bases that include different types of localized

TF functions (also known as TF atoms). The WP approach decomposes the given signal into TF atoms that are adapted to the TF structures expected or known to be present in the signal. The WP coefficients may then be used to obtain filtered or approximate versions of the signal; the coefficients may also be used in the form of a compact representation of the signal for further analysis or classification.

8.7.2 Signal decomposition using the Matching Pursuit algorithm

Matching Pursuit (MP) is a generic iterative signal approximation framework, whose properties depend on the type of dictionary of TF atoms used. In the MP algorithm [424], the given signal $x(t)$ is projected on to a dictionary of TF atoms obtained by scaling, translating, and modulating a window function $g(t)$ as

$$x(t) = \sum_{n=0}^{\infty} a_n g_{\gamma_n}(t), \quad (8.74)$$

where

$$g_{\gamma_n}(t) = \frac{1}{\sqrt{s_n}} g\left(\frac{t - \tau_n}{s_n}\right) \exp[j(2\pi f_n t + \phi_n)]. \quad (8.75)$$

Here, a_n is an expansion coefficient, the scale factor s_n is used to control the width of the window function, the parameter τ_n controls temporal placement, and the parameters f_n and ϕ_n are the frequency and phase of the exponential function, respectively. The subscript or index γ_n represents the set of parameters $(s_n, \tau_n, f_n, \phi_n)$. If the window is a Gaussian function, that is,

$$g(t) = 2^{\frac{1}{4}} \exp(-\pi t^2), \quad (8.76)$$

the TF atoms are known as Gabor atoms.

In the MP algorithm, the given signal is iteratively projected on to a dictionary of TF atoms. The first projection decomposes the signal into two parts:

$$x(t) = \langle x, g_{\gamma_0} \rangle g_{\gamma_0}(t) + R^1 x(t), \quad (8.77)$$

where $\langle x, g_{\gamma_0} \rangle$ denotes the inner product or projection of $x(t)$ on to the first TF atom $g_{\gamma_0}(t)$; the term $R^1 x(t)$ represents the residue after approximating $x(t)$ using $g_{\gamma_0}(t)$. The process of decomposition is continued by projecting the residue on to the subsequent functions in the dictionary, and after M iterations, we have

$$x(t) = \sum_{n=0}^{M-1} \langle R^n x, g_{\gamma_n} \rangle g_{\gamma_n}(t) + R^M x(t), \quad (8.78)$$

with $R^0 x(t) = x(t)$.

The iterative decomposition process may be stopped after using a prespecified number M of the TF atoms or until the energy of the residue $R^M x(t)$ decreases below a certain value. A decay parameter, defined as [424]

$$\lambda(m) = \sqrt{1 - \frac{\|R^m x\|^2}{\|R^{m-1} x\|^2}}, \quad (8.79)$$

may also be used; the process is continued until the decay parameter does not reduce substantially any further. The selected atoms or components may be considered to represent the coherent structures present in the signal with respect to the dictionary used. The residue may be assumed to be due to random noise. The approximate signal reconstructed using M selected (coherent) structures is given by

$$\tilde{x}(t) = \sum_{n=0}^{M-1} \langle R^n x, g_{\gamma_n} \rangle g_{\gamma_n}(t), \quad (8.80)$$

which may be taken to be a filtered version of the original signal. As in the case of wavelets, the coefficients $a_n = \langle R^n x, g_{\gamma_n} \rangle$, $n = 0, 1, 2, \dots, M-1$, may be used, along with the set of corresponding parameters, as a compact representation of the signal for further analysis or classification.

8.7.3 Empirical mode decomposition

Empirical mode decomposition (EMD) [433–437] is a decomposition procedure with which a multicomponent and nonstationary signal can be decomposed into a small number of simpler signals referred to as intrinsic mode functions (IMFs). IMFs may be analyzed using the Hilbert transform and other methods to derive TFDs [433]. The decomposition procedure is based on the local characteristics of the given signal; it is a data-driven algorithm that does not use predetermined basis functions.

An IMF is required to have the following properties [433, 435]:

- Over the duration of an IMF, the number of extrema and the number of zero-crossings must be equal or differ at most by one.
- At any location, the mean of the envelopes defined by the local maxima and the local minima is zero.

An IMF is not restricted to be a narrowband signal; it can be modulated in both amplitude and frequency and also be nonstationary [433].

IMFs are extracted level by level [433, 434]. First, high-frequency local oscillations riding on the corresponding low-frequency parts of the given signal are extracted. Then, the next level of high-frequency local oscillations in the residual of the signal are extracted. This iterative procedure is continued until no complete oscillation can be identified in the residual, which is considered to be the trend in the original signal; then, the final residual is a monotonic function.

In practice, EMD is implemented through a sifting process that detects and uses the local extrema. The standard EMD algorithm works as follows [433–437]:

1. Find the locations of all of the extrema in the given signal $x(n)$.
2. Interpolate between the minima to obtain the lower signal envelope, $x_{\min}(n)$; do the same with the maxima to obtain $x_{\max}(n)$. The cubic spline function is recommended for interpolation.
3. Compute the mean $x_m(n) = [x_{\min}(n) + x_{\max}(n)]/2$.
4. Subtract the mean obtained in the previous step from the signal to obtain the oscillatory mode signal $s(n) = x(n) - x_m(n)$.
5. If $s(n)$ meets the criteria for an IMF, define $c(n) = s(n)$ as an IMF; otherwise, set $x(n) = s(n)$ and repeat the process from Step 1.
6. Initially, the procedure given above results in the first IMF. To get the remaining IMFs, define the residual as $x(n) - c(n)$ and repeat the procedure.

If $c(n)$ is an IMF obtained as above, let

$$c_a(n) = c(n) + j c_H(n), \quad (8.81)$$

where $c_H(n)$ is the Hilbert transform of $c(n)$ and $c_a(n)$ is the related analytic signal. Let

$$c_a(n) = a(n) \exp[j \theta(n)], \quad (8.82)$$

with

$$a(n) = \sqrt{c^2(n) + c_H^2(n)} \quad (8.83)$$

and

$$\theta(n) = \arctan \left[\frac{c_H(n)}{c(n)} \right]. \quad (8.84)$$

Then, according to one of the definitions used, the instantaneous frequency $\omega(n)$ is given by the derivative of $\theta(n)$ with respect to time. Analysis as above can be performed on all IMFs and the results combined to perform TFD analysis of the original signal. See Section 5.5.1, Section 5.5.3, and Huang et al. [433] for related discussions.

The original signal may be expressed using the IMFs as

$$x(n) = \text{real} \left\{ \sum_i a_i(n) \exp[j \theta_i(n)] \right\}, \quad (8.85)$$

where the index i indicates the i^{th} IMF and the summation is over all IMFs. The expression provided above gives both the amplitude and the frequency of each component as functions of time; the residual component is neglected. With EMD and

expansion of the given signal into its IMFs, the amplitude and the frequency modulation in the original signal are separated: this facilitates TFD analysis.

Figure 8.21 shows an example of application of EMD to a VAG signal. The interpretation of IMFs is not straightforward because of the data-dependent nature of the components. In the example, the first three IMFs may be considered to be high-frequency or small-scale noise and not used in further analysis. The IMFs numbered 4 to 7 may be considered to contain most of the useful information. The remaining components may represent low-frequency or large-scale variations and trends that are not of interest.

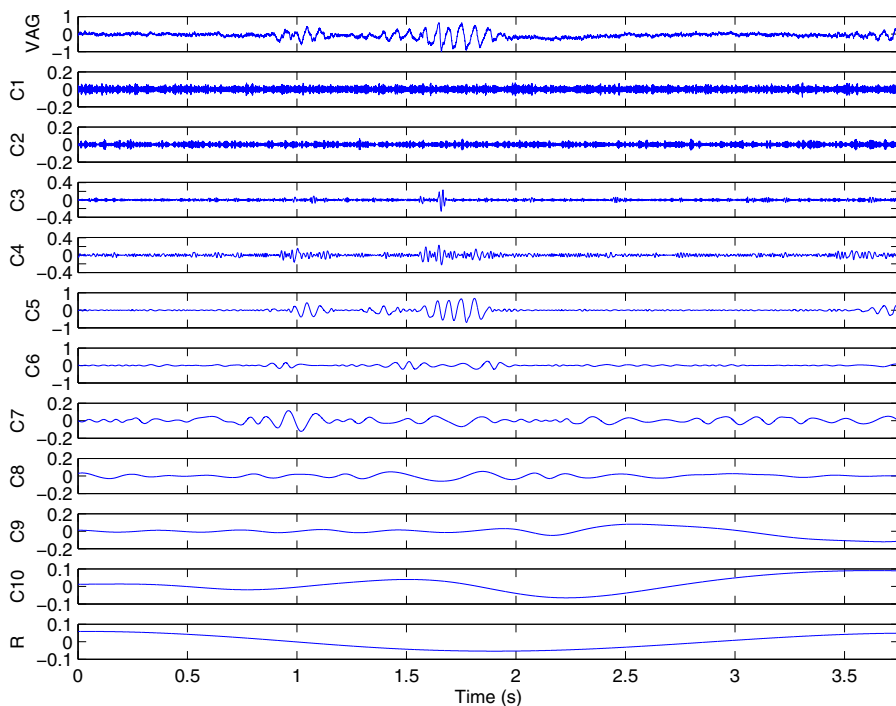


Figure 8.21 A VAG signal and 10 IMFs obtained via EMD. The last plot labeled “R” is the residual component. The number of IMFs was fixed at 10; the residual signal is not monotonic and the EMD procedure could be continued further. Figure courtesy of Yunfeng Wu, Xiamen University, China.

Wu et al. [438] applied a modified method known as ensemble EMD and detrended fluctuation analysis for removal of artifacts in knee-joint VAG signals. They found that EMD can improve the quality of VAG signals and that the results obtained were superior to those achieved via decomposition by the MP algorithm.

Echeverria et al. [437] applied EMD to analyze HRV data. The results were said to facilitate isolation of at least four components with overlapping and dynamic frequency excursions localized in the currently recognized spectral bands of autonomic

modulation. In addition, EMD was observed to permit the separation of the very-low-frequency (VLF) content as well.

Kaleem et al. [439] used EMD to analyze pathological speech signals. Instead of using sustained vowels, EMD was used to decompose randomly chosen portions of speech signals into IMFs, which were then analyzed to extract temporal and spectral features. Instantaneous features were used to capture discriminative information in signals that were suspected to be hidden at local time scales. Using a database consisting of continuous speech from 51 normal speakers and 161 speakers with speech pathology, a classification accuracy of 95.7% was obtained.

8.7.4 TFDs and their characteristics

If $x(t)$ is a signal and $\text{TFD}(t, \omega)$ is its TFD, the important criteria that the TFD is expected to satisfy are listed in the following paragraphs [406, 408].

Total energy:

$$\int \int \text{TFD}(t, \omega) dt d\omega = \int |x(t)|^2 dt = \int |X(\omega)|^2 d\omega, \quad (8.86)$$

where $X(\omega)$ is the Fourier transform of $x(t)$. This criterion indicates that, at particular values of t and ω , $\text{TFD}(t, \omega)$ gives the fractional energy of the signal; thus, $\text{TFD}(t, \omega)$ may be viewed as a joint TF density function.

Invariance: The TFD should be invariant to any linear shift in time and frequency; it is also expected that the TFD is scale-invariant.

Nonnegativity: $\text{TFD}(t, \omega) \geq 0$ for all t and ω . This criterion helps in treating the TFD as a density function.

Marginals: The instantaneous energy of the signal is given by

$$\int \text{TFD}(t, \omega) d\omega = |x(t)|^2. \quad (8.87)$$

Integration along the time axis gives the PSD of the signal as

$$\int \text{TFD}(t, \omega) dt = |X(\omega)|^2. \quad (8.88)$$

Expectation values:

$$E[g(t, \omega)] = \int \int g(t, \omega) \text{TFD}(t, \omega) dt d\omega. \quad (8.89)$$

The equation given above expresses a generalized moment of a TFD. The function $g(t, \omega)$ needs to be chosen according to the desired moment. The instantaneous mean frequency (IMF) is given by the time-varying first moment of the TFD along the frequency axis as

$$E_t[\omega] = \frac{1}{|x(t)|^2} \int \omega \text{TFD}(t, \omega) d\omega. \quad (8.90)$$

The group delay is given by the frequency-varying first moment of the TFD along the time axis as

$$E_{\omega} [t] = \frac{1}{|X(\omega)|^2} \int t \text{TFD}(t, \omega) dt. \quad (8.91)$$

See Boashash [382] for discussions on analysis of the instantaneous frequency of nonstationary signals.

Finite support: If $x(t)$ is zero at $t = t_1$, then $\text{TFD}(t_1, \omega)$ should be zero. Similarly, if $X(\omega)$ is zero at $\omega = \omega_1$, then $\text{TFD}(t, \omega_1)$ should be zero.

The simplest TFD is the spectrogram, which is obtained by taking the squared magnitude of the STFT defined in Equation 8.9.

The Wigner–Ville distribution (WVD) of a signal $x(t)$ is given by [406]

$$\text{WVD}(t, \omega) = \int_{-\infty}^{\infty} x(t + \tau/2) x^*(t - \tau/2) \exp(-j\omega\tau) d\tau. \quad (8.92)$$

A drawback of the WVD is that, in the case of multicomponent signals, cross-terms (referred to as “ghost frequencies”) are generated in the TFD that could lead to misinterpretation.

The Cohen’s class of generalized TFD (GTFD) is given by

$$\begin{aligned} \text{GTFD}(t, \omega) &= \frac{1}{4\pi^2} \int \int \int x(u + \tau/2) x^*(u - \tau/2) \Phi(\theta, \tau) \\ &\quad \times \exp(-j\theta t - j\tau\omega + j\theta u) du d\tau d\theta, \end{aligned} \quad (8.93)$$

where $\Phi(\theta, \tau)$ is the TF transformation kernel. The kernel acts as a lowpass filter and minimizes cross-terms. The selection of the kernel, which may be signal-independent or signal-dependent, determines the properties of the resulting distribution [405, 406].

Considering a signal-independent kernel, cross-terms may be minimized by 2D filtering of the WVD; the smoothed version of the WVD (SWVD) can be expressed as

$$\text{SWVD}(t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(u, \Omega) \text{WVD}(t - u, \omega - \Omega) du d\Omega, \quad (8.94)$$

where $\phi(u, \Omega)$ is a kernel function with lowpass-filter characteristics. The kernel function may be expressed as

$$\phi(u, \Omega) = g(u)H(\Omega), \quad (8.95)$$

where $g(u)$ and $H(\Omega)$ are smoothing functions or windows in the time and frequency domains, respectively. The smoothing windows are independent of the signal; Gaussian functions are commonly used as smoothing windows. The resulting

TFD, known as the smoothed pseudo-WVD (SPWVD) [440] is expressed as

$$\text{SPWVD}(t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(u) H(\Omega) \text{WVD}(t-u, \omega-\Omega) du d\Omega. \quad (8.96)$$

Smoothing windows suppress cross-terms but smear localized components, leading to less accurate TF localization of signal components as compared to the WVD. A reassignment method was proposed by Auger and Flandrin [441] to improve TF localization in smoothed TFDs, such as SPWVDs. In the reassignment method, the window is moved from a point (t, ω) to the center of energy, $(\bar{t}, \bar{\omega})$, of the TFD. The center of energy is given by the group delay and the mean frequency computed as the mean values of the WVD:

$$\bar{t}(t, \omega) = t - \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u g(u) H(\Omega) \text{WVD}(t-u, \omega-\Omega) du d\Omega}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(u) H(\Omega) \text{WVD}(t-u, \omega-\Omega) du d\Omega}, \quad (8.97)$$

$$\bar{\omega}(t, \omega) = \omega - \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Omega g(u) H(\Omega) \text{WVD}(t-u, \omega-\Omega) du d\Omega}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(u) H(\Omega) \text{WVD}(t-u, \omega-\Omega) du d\Omega}. \quad (8.98)$$

The modified TFD, known as the reassigned SPWVD (or RSPWVD), can be expressed as

$$\begin{aligned} \text{RSPWVD}(t', \omega') &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{SPWVD}(t, \omega) \\ &\times \delta[t' - \bar{t}(t, \omega)] \delta[\omega' - \bar{\omega}(t, \omega)] dt d\omega. \end{aligned} \quad (8.99)$$

8.7.5 Decomposition based adaptive TFD

Krishnan et al. [273] presented the design of an adaptive TFD based on signal decomposition for the analysis of VAG signals; a schematic representation of their procedure is shown in Figure 8.22. In this method, the given signal is first decomposed into components of a specified mathematical representation; the MP algorithm described in Section 8.7.2 was used for this purpose. By knowing the components of the signal, the interaction between them can be established and used to remove or prevent cross-terms.

The TFD based on signal decomposition using the MP algorithm (or MPTFD) is obtained by taking the WVD of the TF atoms used to represent the signal as in Equation 8.80 and is given by [424]

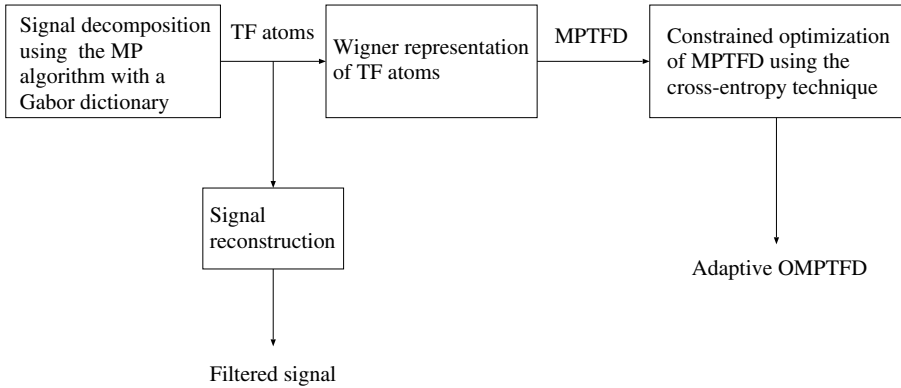


Figure 8.22 Schematic representation of the procedure to construct a decomposition-based adaptive TFD. Adapted with permission from S. Krishnan, R.M. Rangayyan, G.D. Bell, and C.B. Frank, Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology, *IEEE Transactions on Biomedical Engineering*, 47(6):773–783, June 2000. ©IEEE.

$$\begin{aligned}
 W(t, \omega) &= \sum_{n=0}^{M-1} |\langle R^n x, g_{\gamma_n} \rangle|^2 W g_{\gamma_n}(t, \omega) \\
 &+ \sum_{n=0}^{M-1} \sum_{\substack{m=0 \\ m \neq n}}^{M-1} \langle R^n x, g_{\gamma_n} \rangle \langle R^m x, g_{\gamma_m} \rangle^* W_{[g_{\gamma_n}, g_{\gamma_m}]}(t, \omega),
 \end{aligned} \tag{8.100}$$

where $W g_{\gamma_n}(t, \omega)$ is the WVD of the Gaussian window function. The double sum corresponds to the cross-terms of the WVD indicated by $W_{[g_{\gamma_n}, g_{\gamma_m}]}(t, \omega)$ and should be rejected in order to obtain a cross-term-free distribution of $x(t)$ in the TF plane. Thus, only the first term is retained, and the resulting TFD is given by

$$W'(t, \omega) = \sum_{n=0}^{M-1} |\langle R^n x, g_{\gamma_n} \rangle|^2 W g_{\gamma_n}(t, \omega). \tag{8.101}$$

Because the MPTFD is constructed using the WVD of TF atoms, most of the properties of the WVD are applicable to the MPTFD, except the marginals.

The MPTFD may be modified to satisfy the marginal requirements while preserving its other important characteristics. One way to optimize the MPTFD is by applying the cross-entropy minimization method [442–444]. Cross-entropy minimization is a general method of inference about an unknown PDF when there exists a prior estimate of the PDF and new information is available in the form of constraints on its expected values. The optimized MPTFD (or OMPTFD), denoted by $M(t, \omega)$, should satisfy the condition on the marginals, stated as

$$\int M(t, \omega) d\omega = |x(t)|^2 = m(t), \quad (8.102)$$

$$\int M(t, \omega) dt = |X(\omega)|^2 = m(\omega). \quad (8.103)$$

The two conditions given above may be treated as constraints for optimization. The desired $M(t, \omega)$ may be obtained from $W'(t, \omega)$, which is taken to be a prior estimate of the density, by minimizing the cross-entropy between them, given by

$$H(M, W') = \int \int M(t, \omega) \log \left(\frac{M(t, \omega)}{W'(t, \omega)} \right) dt d\omega. \quad (8.104)$$

Considering the marginals, the OMPTFD may be written as [443]

$$M(t, \omega) = W'(t, \omega) \exp\{-[\alpha_0(t) + \beta_0(\omega)]\}, \quad (8.105)$$

where α and β are Lagrange multipliers that may be determined using the constraints in an iterative algorithm as proposed by Loughlin et al. [444].

At the first iteration, let

$$M_1(t, \omega) = W'(t, \omega) \exp[-\alpha_0(t)]. \quad (8.106)$$

Imposing the constraint related to the time marginal given by Equation 8.102 on Equation 8.106, we obtain

$$\alpha_0(t) = \ln \left(\frac{m'(t)}{m(t)} \right), \quad (8.107)$$

where $m(t)$ is the desired time marginal and $m'(t)$ is the time marginal estimated from $W'(t, \omega)$. Then,

$$M_1(t, \omega) = W'(t, \omega) \frac{m(t)}{m'(t)}. \quad (8.108)$$

To have the desired frequency marginal $m(\omega)$, let

$$M_2(t, \omega) = M_1(t, \omega) \exp[-\beta_0(\omega)]. \quad (8.109)$$

By imposing the constraint on the frequency marginal given by Equation 8.103 on Equation 8.109, we obtain

$$\beta_0(\omega) = \ln \left(\frac{m'(\omega)}{m(\omega)} \right), \quad (8.110)$$

where $m(\omega)$ is the desired frequency marginal and $m'(\omega)$ is the frequency marginal estimated from $W'(t, \omega)$. Now, we have

$$M_2(t, \omega) = M_1(t, \omega) \frac{m(\omega)}{m'(\omega)}. \quad (8.111)$$

At this point, $M_2(t, \omega)$ may be altered and need not necessarily give the desired time marginal. Successive iteration could overcome this problem and modify the TFD to get closer to $M(t, \omega)$ [444]. This follows from the fact that the cross-entropy between the desired TFD and the estimated TFD decreases with the number of iterations [443, 444]. Because the iterative procedure is started with a positive distribution, $W'(t, \omega)$, the TFD at the n^{th} iteration, $M_n(t, \omega)$, is guaranteed to be a positive distribution.

8.7.6 Illustrations of application

Some of the commonly used wavelets are the Mexican hat (or sombrero), Morlet, Daubechies, and symmlet functions [286, 406, 408, 420–429]. The Mexican hat function is defined as the negative of the second derivative of a Gaussian, and is given by

$$\psi(t) = (1 - t^2) \exp(-0.5 t^2). \quad (8.112)$$

Figure 8.23 shows the wavelet defined above for three scales of 1, 2, and 4 (top to bottom), scaled according to Equation 8.72. It is evident that, as the scale value increases, the duration of the wavelet is increased and its frequency bandwidth is reduced; thus, there exists an inverse relationship between scale and frequency.

The Morlet function is defined as

$$\psi(t) = \frac{1}{\pi^{1/4}} [\exp(j \omega_0 t) - \exp(-0.5 \omega_0^2)] \exp(-0.5 t^2), \quad (8.113)$$

where ω_0 is the central frequency of the frequency response of the mother wavelet. The term $\exp(-0.5 \omega_0^2)$ is negligible for $\omega_0 > 5$ and may be ignored. Figures 8.24 and 8.25 show Morlet wavelets obtained using the real part of Equation 8.113 for three scales of 1, 2, and 4 and $\omega_0 = 2$ and 6 *radians/s*, respectively; the wavelets have been scaled according to Equation 8.72. The Morlet wavelets in Figure 8.24 are similar in appearance to the Mexican hat wavelets in Figure 8.23.

Figure 8.26 shows a noisy ECG signal over three cardiac cycles, sampled at 200 *Hz*. Figure 8.27 shows the STFT of the ECG signal using a window width of 32 samples, which was moved in steps of 16 samples. It is evident from the STFT that the spectrum of the signal changes substantially over its duration, with increased

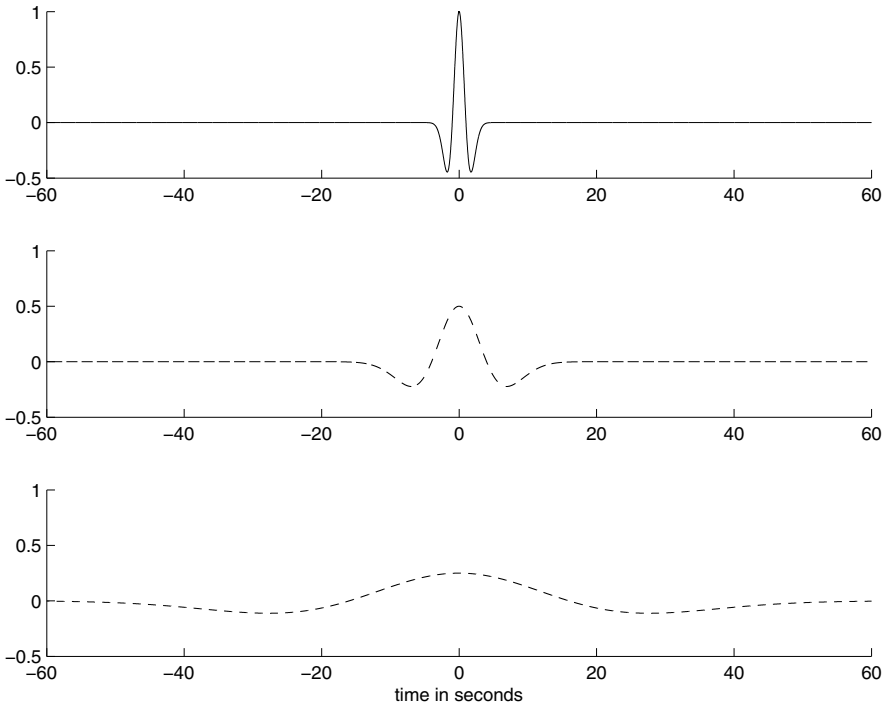


Figure 8.23 Top to bottom: The Mexican hat wavelet in Equation 8.112 for three scales of 1, 2, and 4. The wavelets have been scaled according to Equation 8.72.

high-frequency content during the intervals corresponding to the QRS complexes and predominantly low-frequency composition in periods related to the P and T waves.

In order to analyze the ECG signal using wavelets, wavelets were prepared as per Equation 8.113 (using only the real part) with $\omega_0 = 200 \text{ radians/s}$ and by changing the Gaussian term to $\exp(-5000 t^2)$. The parameters were selected so as to create wavelets that were comparable to typical QRS waves in ECG signals. Figure 8.28 shows the wavelets for four scales of 1, 2, 4, and 8. Figure 8.29 shows the ECG signal (top) and the results of the CWT for the four scales used (labeled correspondingly as wt1, wt2, wt4, and wt8 in the figure). In each case, the wavelet was translated by one sample per step. It is evident that the wavelets of smaller scale result in higher coefficients corresponding to the sharp QRS complexes, whereas the wavelets of larger scale provide larger coefficients at the locations of the slower T waves. One may, if desired, compute the CWT for several values of the scale parameter and compose the result for display as an image, known as the scalogram (see Figure 8.30).

The last plot in Figure 8.29 (labeled as wr) shows an approximation of the ECG signal reconstructed using only the coefficients for scales of 2, 4, and 8; that is, the result for the smallest scale of 1 was discarded. To obtain the reconstructed signal,

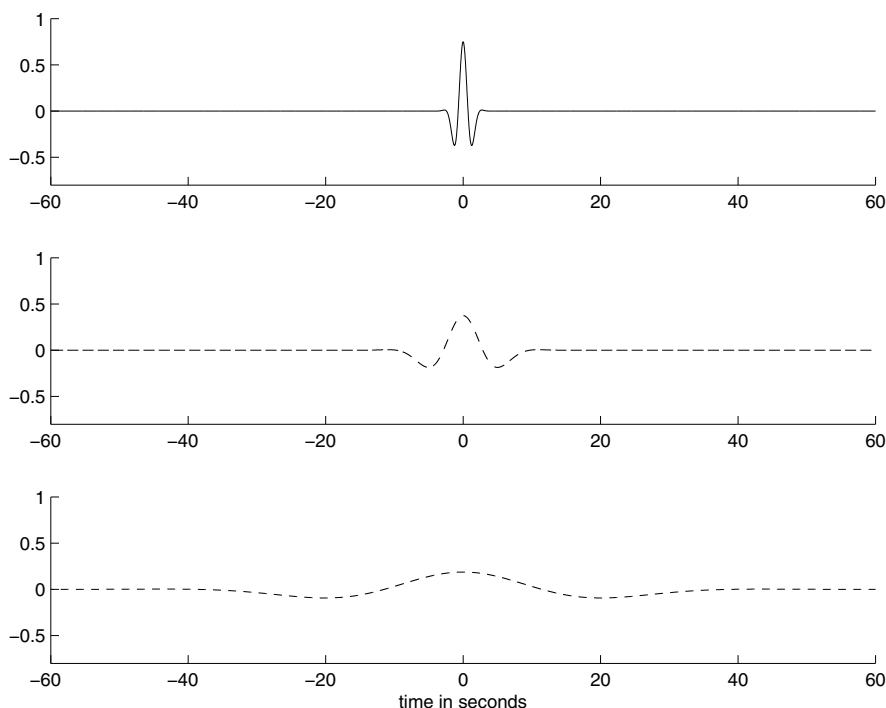


Figure 8.24 Top to bottom: The Morlet wavelet in Equation 8.113 for three scales of 1, 2, and 4, and $\omega_0 = 2$ radians/s. Only the real part has been used and the wavelets have been scaled according to Equation 8.72.

the wavelet transform for each scale was convolved with the corresponding wavelet according to Equation 8.69; however, due to the redundancy of the CWT, a reconstruction may be obtained by simply adding the results of the wavelet transform for the selected scales [428, 429]. The result of approximation is seen to retain well the P and QRS waves with most of the small-scale noise removed; however, the T waves have not been reproduced well. This is a simplified illustration of the CWT with only four scales used; in practice, one may have to use a large number of scales depending upon the nature of the signal and the desired result.

Representation and analysis of a given signal as shown above may be viewed as an application of matched filters, correlation filters, and multiscale filter banks. The CWT given in Equation 8.66 represents a convolution operation between the given signal and a wavelet. When a symmetric wavelet is used, convolution is the same as correlation. The CWT operation is then comparable to matched filtering (see Section 4.6.1). Thus, the CWT may be interpreted as seeking a match between a collection of wavelets and the given signal: the illustration in Figure 8.29 demonstrates this aspect of wavelet analysis.

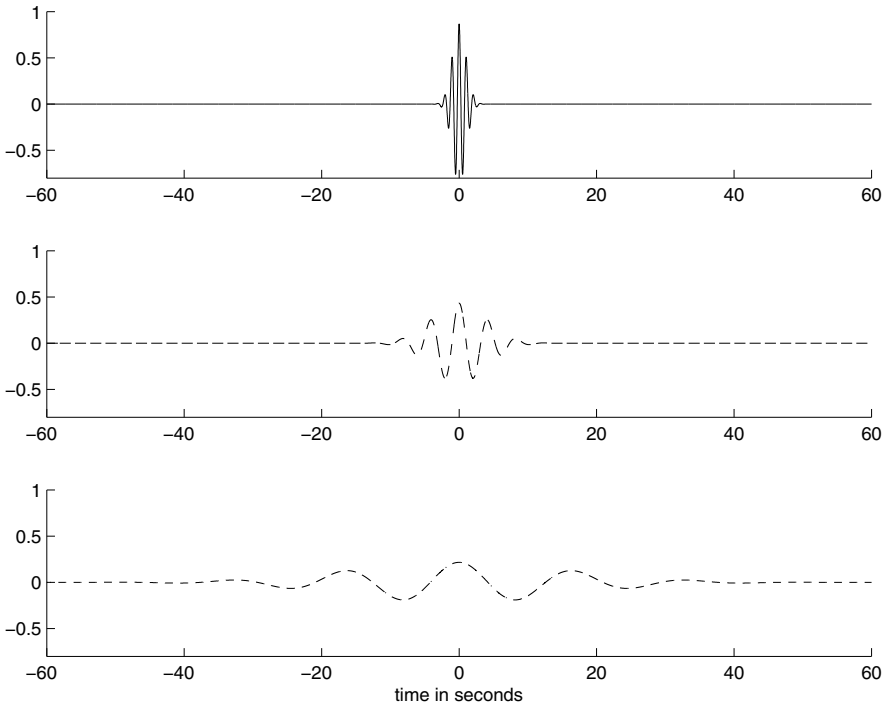


Figure 8.25 Top to bottom: The Morlet wavelet in Equation 8.113 for three scales of 1, 2, and 4, and $\omega_0 = 6$ radians/s. Only the real part has been used and the wavelets have been scaled according to Equation 8.72.

Further insight into the interpretation of wavelet analysis as an application of multiscale filter banks can be gained by examining the PSDs of the wavelets used. Figure 8.31 shows the normalized PSDs of the four wavelets in Figure 8.28; it is evident that the central frequency of the frequency response of the filter corresponding to each wavelet shifts towards lower values as the scale is increased. When several wavelets are used over a range of scales, a number of filtered versions of the given signal are obtained, as illustrated in Figure 8.29. The coefficients obtained using the CWT or DWT may also be used for further analysis of signals such as the ECG for various purposes, including detection of waves and classification [420,421,427].

In the work of Krishnan and Rangayyan [192], several methods were applied to remove noise in multicomponent and nonstationary synthetic signals as well as VAG signals. Gaussian random noise was added to the synthetic signals such that the resulting signals had SNR of 10 dB and 0 dB. The symmlet 4 wavelet [430] was used for wavelet-based filtering. A soft threshold ($T = 0.5$) was applied to the wavelet coefficients. In the case of the WP method, the best basis was selected by using the Schur concavity cost function [432], and the filtered version was obtained by applying the soft thresholding method ($T = 0.5$) to the WP coefficients. Gaussian

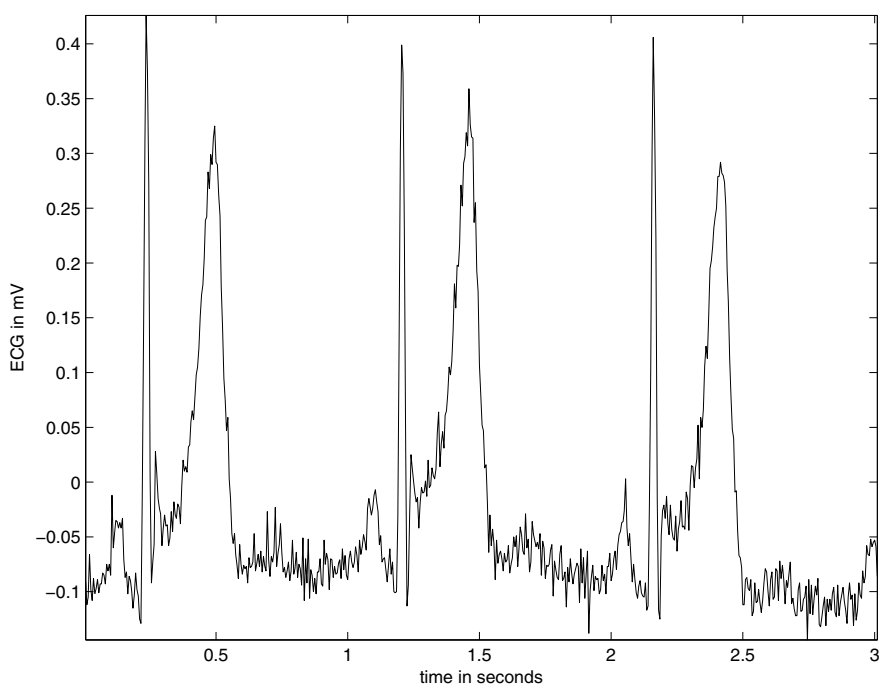


Figure 8.26 A noisy ECG signal over three cardiac cycles or beats.

functions were used for the MP method. The stopping criterion was based on the decay parameter as given by Equation 8.79.

Figure 8.32 (a) shows a synthetic signal and part (b) of the same figure shows the signal with noise added such that $SNR = 0$ dB. The filtered versions of the signal using the wavelet method, the WP method, and the MP method are shown in Figures 8.33 and 8.34. Visual comparison indicates that the MP result has preserved most of the important characteristics of the signal — in particular, the transient component.

The WP method may lead to better results if the threshold is selected in an optimal manner. The MP method locally optimizes the choice of the decomposition function for the signal residue at each stage. In the case of a multicomponent signal where different types of structures are located at different times but in the same frequency band, there may be no WP basis that is well adapted to all of them. MP is a translation-invariant method if a translation-invariant dictionary such as a Gabor dictionary is used. See Krishnan and Rangayyan [192] for more details and quantitative evaluation of the results.

The VAG signal of a subject with cartilage pathology is shown in Figure 8.35 (a), its filtered version using MP is shown in part (b) of the same figure, and the difference between the original signal and the filtered result is shown in Figure 8.36. From

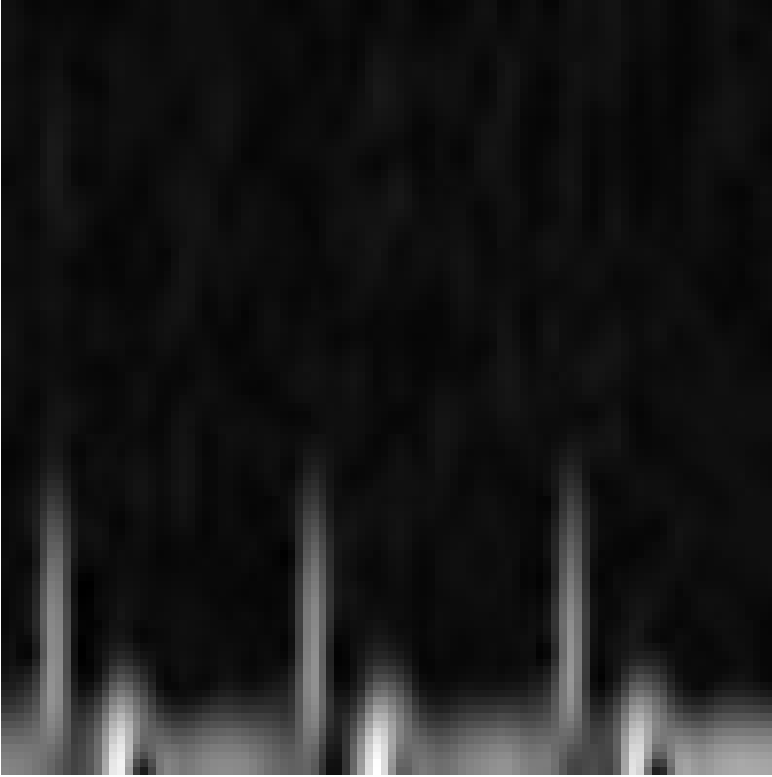


Figure 8.27 STFT of the ECG signal in Figure 8.26. The horizontal axis represents the time axis from 0 to 3 s (left to right) and the vertical axis represents frequency from 0 to 100 Hz (bottom to top).

Figure 8.36, it is evident that a substantial amount of random noise has been removed from the original signal.

The TFD of the original VAG signal in Figure 8.35 computed using the STFT is shown in Figure 8.37 (a); the TFD of the MP-filtered version is shown in part (b) of the same figure. The two TFDs were computed using the same STFT parameters. Tonal and FM components are seen more clearly in the TFD of the filtered signal than that of the original signal. For demonstrations of the use of MPTFDs for feature identification and classification of VAG signals, see Krishnan et al. [273] and Rangayyan and Krishnan [445]; see also Section 8.14. See Kovács et al. [446, 447] for illustrations of MP-based analysis of fetal PCG signals.

In the work of Krishnan [448], the synthetic signal shown in Figure 8.38 was prepared by combining overlapping chirp, impulse, and sinusoidal FM components; part (b) of the same figure shows the ideal TFD of the signal (known by design). Figure 8.39 shows a noisy version of the signal with $SNR = 10$ dB and its RSP-WVD. Even though the basic nature of the ideal TFD is visible in the RSPWVD, a

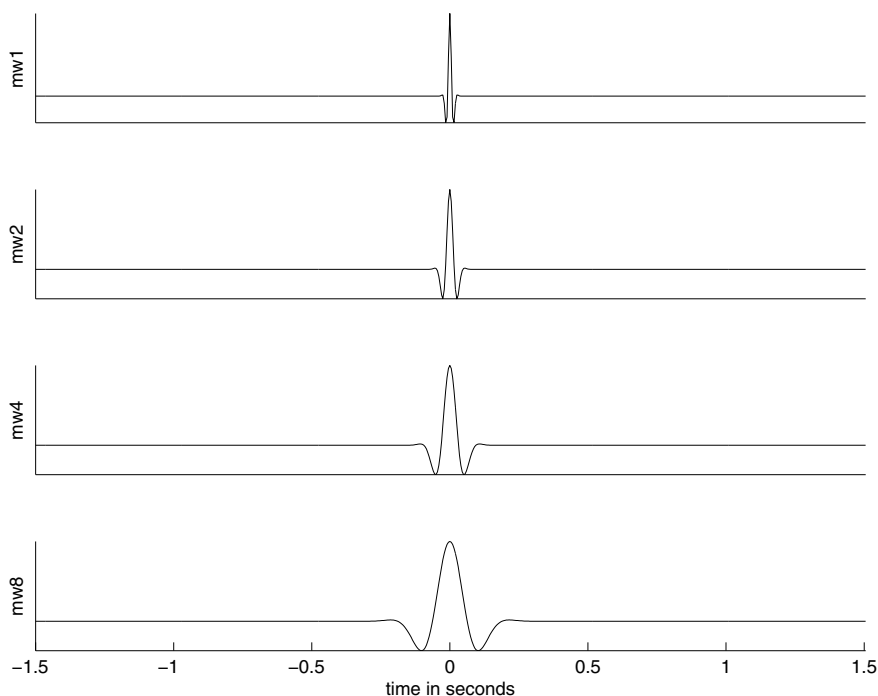


Figure 8.28 Wavelets to analyze the ECG signal in Figure 8.26. The wavelets were prepared using Equation 8.113 with $\omega_0 = 200$ radians/s and by changing the Gaussian term to $\exp(-5000 t^2)$, and are shown for four scales of 1, 2, 4, and 8. The amplitude details have been suppressed and all signals have been mapped to the full range of the ordinate available for clearer display.

large amount of interference is present. Figure 8.40 shows the MPTFD of the noisy signal, which clearly demonstrates the various components present in the signal and their TF characteristics without degrading interference.

See Section 8.14 for further illustrations and discussion of the application of TFD analysis to VAG signals. See Section 8.12 for a discussion on the application of the DWT to analyze EEG signals.

8.8 Separation of Mixtures of Signals

In studies of biomedical signals, it is common to acquire multiple channels of signals from several sources, some of which could provide mutually independent information while others could be correlated and cause redundancy. As shown in Chapter 2, studies of the cardiac, nervous, and neuromuscular systems are often conducted with multiple channels of signals, some of which could be inherently electrical while others could be basically of mechanical, acoustic, and other types, although all are

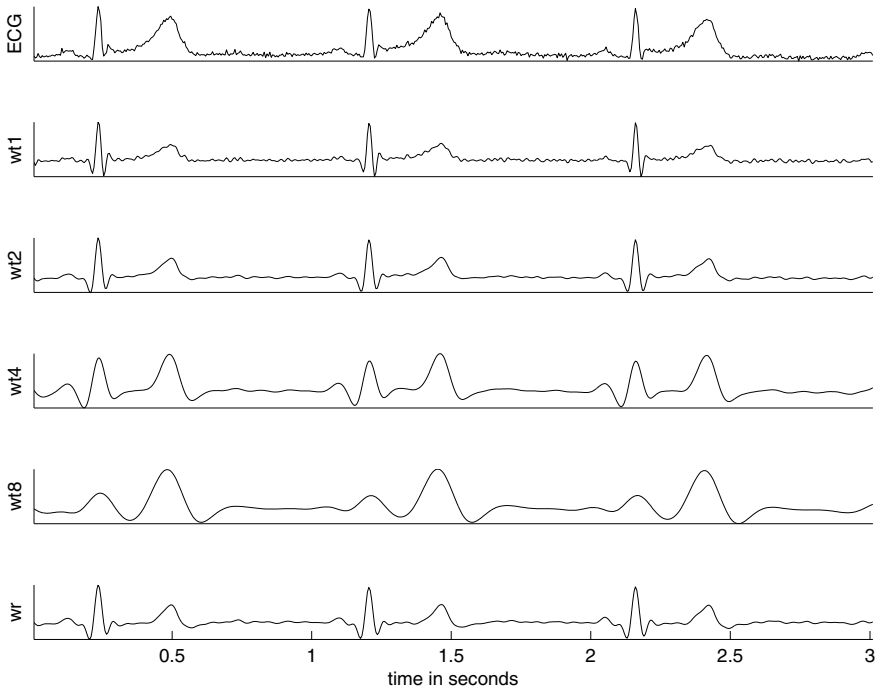


Figure 8.29 Results of analysis of the ECG signal in Figure 8.26 using the wavelets shown in Figure 8.28. Top to bottom: the original ECG signal; results of the CWT for the four scales used (wt1, wt2, wt4, and wt8); and approximation of the signal (wr) reconstructed using only wt2, wt4, and wt8. The amplitude details have been suppressed and all signals have been mapped to the full range of the ordinate available for clearer display.

acquired as electrical signals via suitable instrumentation. Regardless of the nature of the signals, cross-talk and cross-contamination can occur at their sources, along their paths of propagation, and at the sites of their acquisition. When analyzing such collections of multichannel signals, questions arise regarding their interrelationships, the potential for existence of correlation, the presence of redundancy in the data, and the quality of the information extracted. It would be desirable to remove redundant and correlated information for the sake of efficiency and extract information related to each source without contamination from others that are active at the same time. These considerations lead to the need to separate mixtures of signals into their components. Depending upon the assumptions made and the techniques used, one can extract uncorrelated components via principal component analysis (PCA) or statistically independent components via methods of independent component analysis (ICA) and BSS [65,449–463].

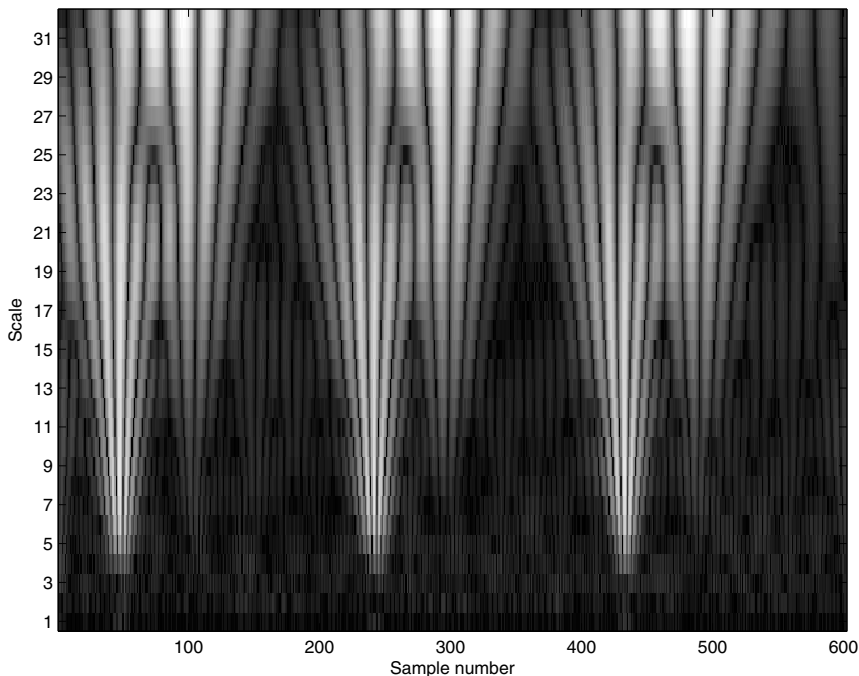


Figure 8.30 CWT of the ECG signal in Figure 8.26 using Morlet wavelets. The horizontal axis represents the time axis from 0 to 3 s sampled at 200 Hz and the vertical axis represents scale from 1 to 32.

8.8.1 Principal component analysis

In order to facilitate efficient representation of collections of signals, measurements, and features for pattern analysis and classification, it is desirable to prepare a compact set of the data with no redundancy or mutual correlation, and to keep the dimension of the result as small as possible. An approach to reduce the dimension of a data vector is PCA, which facilitates the selection of a compact set of uncorrelated components. The PCA method, which is related to the KLT [18, 175, 222], is described briefly in the following paragraphs.

Consider a set of signal samples, measurements, or features represented as a vector \mathbf{y} , of size $K \times 1$. The vector \mathbf{y} may be represented without error by a deterministic linear transformation of the form

$$\mathbf{y} = \mathbf{W} \mathbf{x} = \sum_{k=1}^K x_k \mathbf{W}_k, \quad (8.114)$$

$$\mathbf{W} = [\mathbf{W}_1 \mathbf{W}_2 \cdots \mathbf{W}_K], \quad (8.115)$$

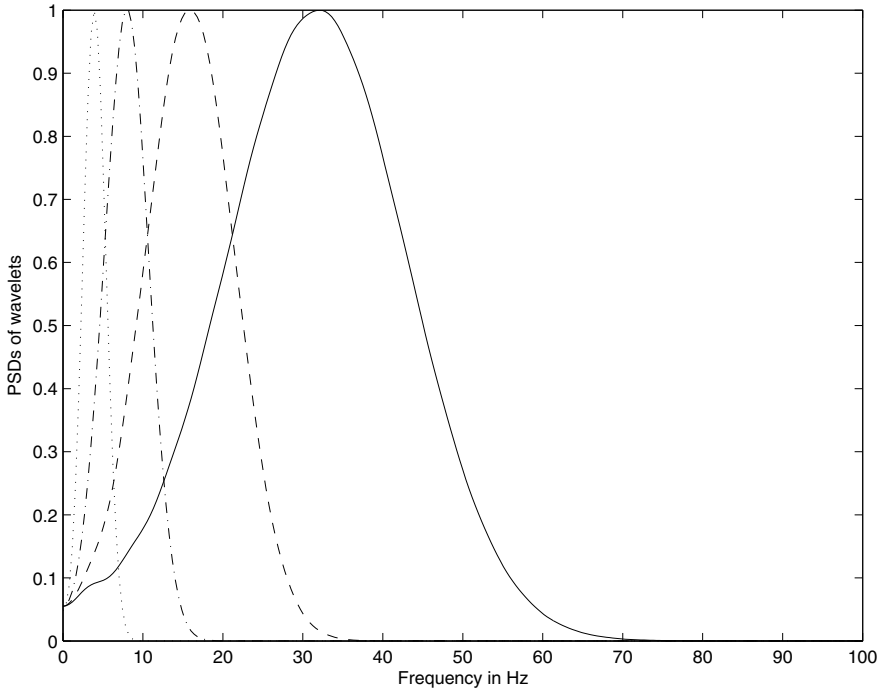


Figure 8.31 Normalized PSDs of the four wavelets in Figure 8.28. The shift towards lower values in the central frequency of the filter corresponding to each wavelet is evident as the scale is increased from 1 (solid line), to 2 (dashed line), 4 (dot–dash line), and 8 (dotted line).

where $|\mathbf{W}| \neq 0$, and \mathbf{W}_k are $K \times 1$ column vectors that make up the $K \times K$ matrix \mathbf{W} . The matrix \mathbf{W} should be formulated such that the vector \mathbf{x} , also of size $K \times 1$, facilitates a reduced and efficient representation of the original vector \mathbf{y} .

The matrix \mathbf{W} may be viewed as being composed of K linearly independent column vectors in the K -dimensional space containing \mathbf{y} . Let \mathbf{W} be orthonormal, that is,

$$\mathbf{W}_k^T \mathbf{W}_l = \begin{cases} 1, & k = l, \\ 0, & k \neq l. \end{cases} \quad (8.116)$$

Then,

$$\mathbf{W}^T \mathbf{W} = \mathbf{I} \text{ or } \mathbf{W}^{-1} = \mathbf{W}^T. \quad (8.117)$$

The column vectors of \mathbf{W} form a set of orthonormal basis vectors of a linear transformation. The inverse relationship is given by

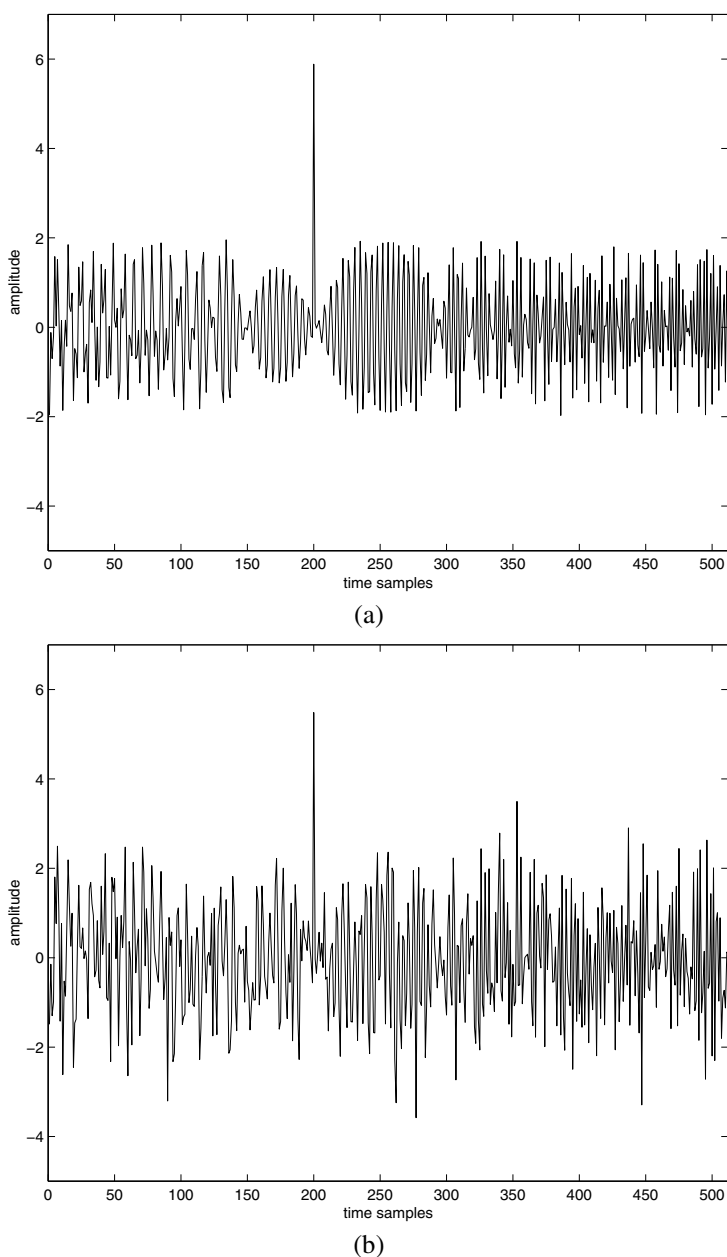


Figure 8.32 (a) Multicomponent and nonstationary synthetic signal with a linear frequency modulation (FM), a nonlinear FM, and a transient component. (b) The signal with noise ($SNR = 0$ dB). Reproduced from S. Krishnan and R. M. Rangayyan, Automatic de-noising of knee-joint vibration signals using adaptive time-frequency representations, *Medical and Biological Engineering and Computing*, 38(1):2–8, 2000, with kind permission from Springer Science+Business Media B.V. ©IFMBE and Springer.

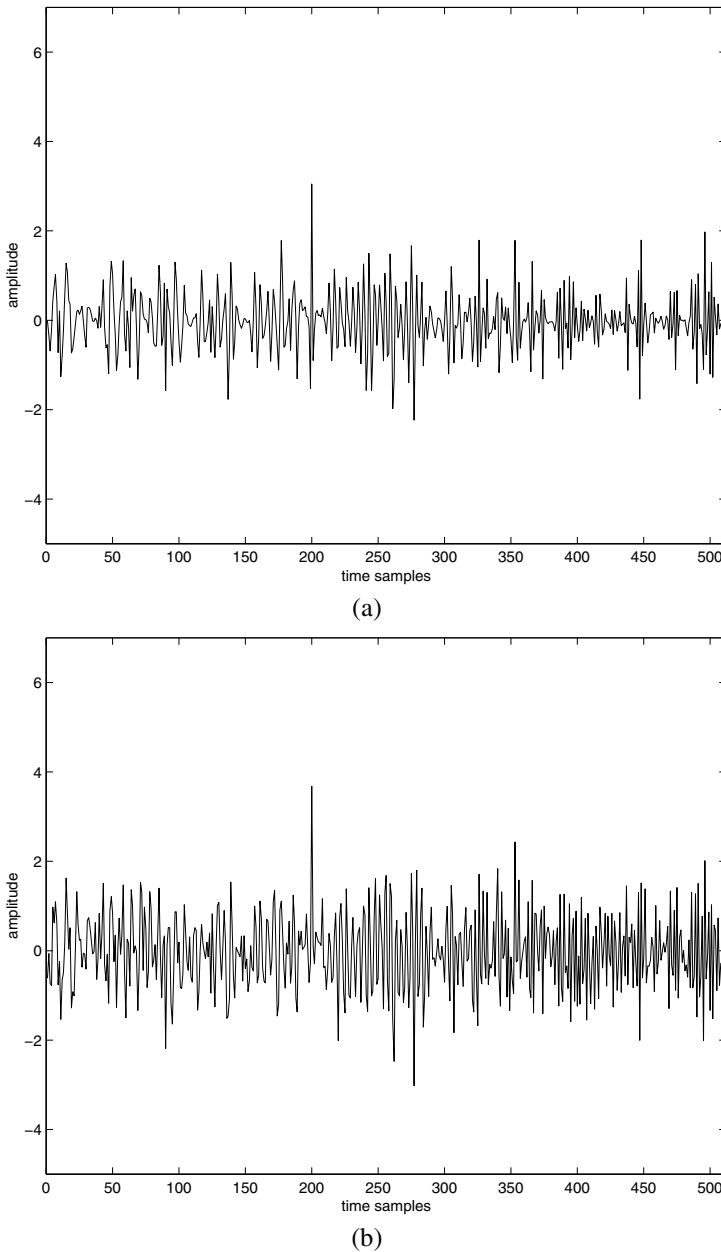


Figure 8.33 Filtered version of the noisy signal in Figure 8.32 (b) using (a) wavelets (symmlet 4) and (b) the WP method. Reproduced from S. Krishnan and R. M. Rangayyan, Automatic de-noising of knee-joint vibration signals using adaptive time-frequency representations, *Medical and Biological Engineering and Computing*, 38(1):2–8, 2000, with kind permission from Springer Science+Business Media B.V. ©IFMBE and Springer.

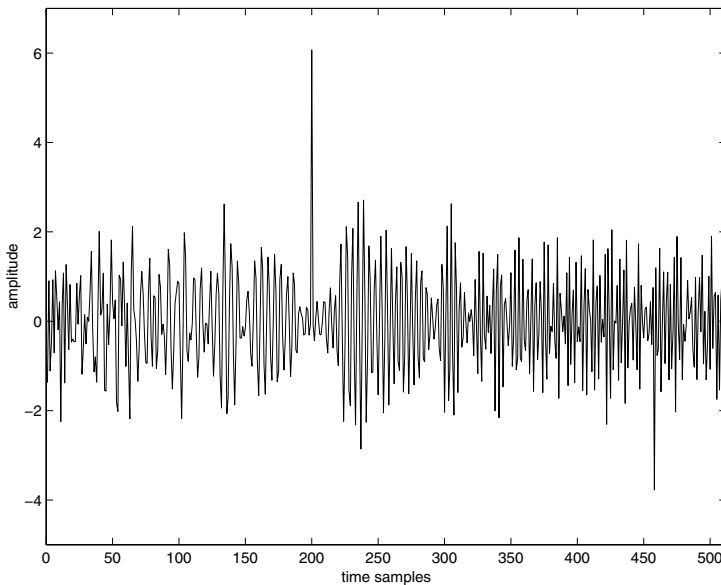


Figure 8.34 Filtered version of the noisy signal in Figure 8.32 (b) using the MP method. Reproduced from S. Krishnan and R. M. Rangayyan, Automatic de-noising of knee-joint vibration signals using adaptive time-frequency representations, *Medical and Biological Engineering and Computing*, 38(1):2–8, 2000, with kind permission from Springer Science+Business Media B.V. ©IFMBE and Springer.

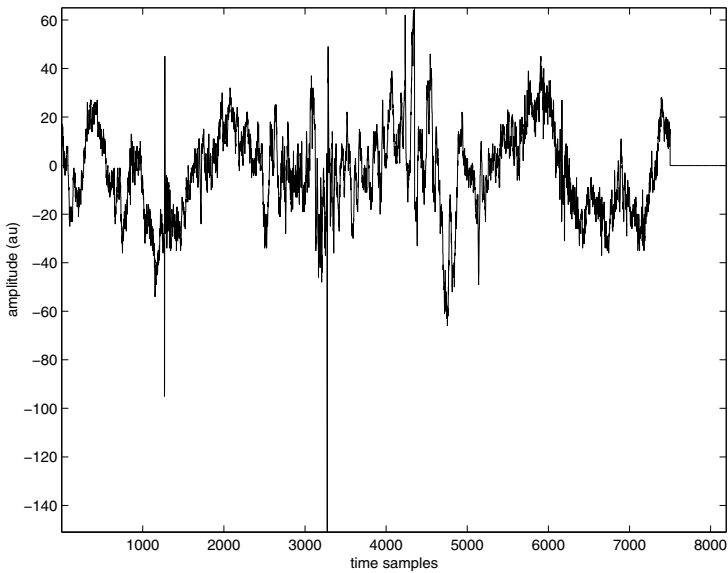
$$\mathbf{x} = \mathbf{W}^{-1} \mathbf{y} = \sum_{k=1}^K y_k \mathbf{W}_k^{-1}. \quad (8.118)$$

In the representation formulated above, each component of \mathbf{x} contributes to the representation of \mathbf{y} and \mathbf{W}_k^{-1} is the k^{th} $K \times 1$ column vector of the $K \times K$ matrix \mathbf{W}^{-1} . With the design of \mathbf{W} as a linear and reversible transformation, \mathbf{x} provides a complete or lossless representation of \mathbf{y} if all of its K elements are used.

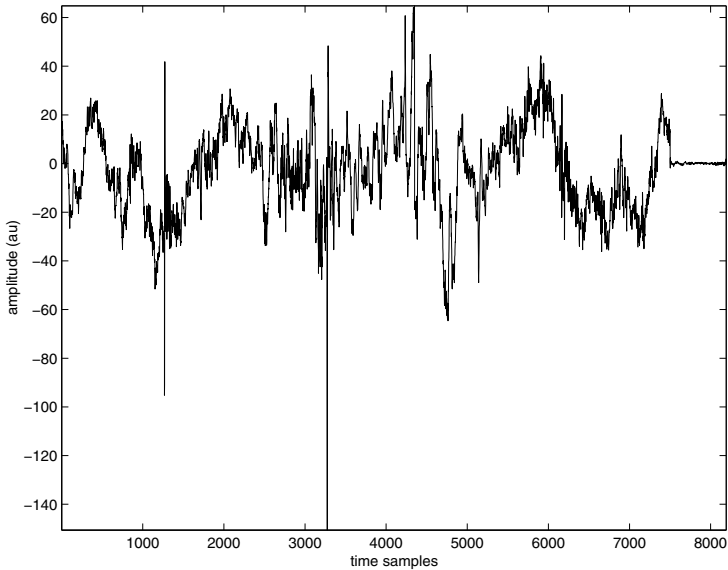
To achieve compact and efficient representation of the original vector \mathbf{y} by extracting the most significant information contained, we may choose to use only $L < K$ components of \mathbf{x} . The omitted components of \mathbf{x} may be replaced with other values \tilde{x}_k , $k = L + 1, L + 2, \dots, K$. We now have an approximation of \mathbf{y} , represented as

$$\tilde{\mathbf{y}} = \sum_{k=1}^L x_k \mathbf{W}_k + \sum_{k=L+1}^K \tilde{x}_k \mathbf{W}_k. \quad (8.119)$$

The error of approximation is



(a)



(b)

Figure 8.35 (a) VAG signal of a subject with cartilage pathology and (b) MP-filtered version of the signal. Reproduced from S. Krishnan and R. M. Rangayyan, Automatic de-noising of knee-joint vibration signals using adaptive time-frequency representations, *Medical and Biological Engineering and Computing*, 38(1):2–8, 2000, with kind permission from Springer Science+Business Media B.V. ©IFMBE and Springer.

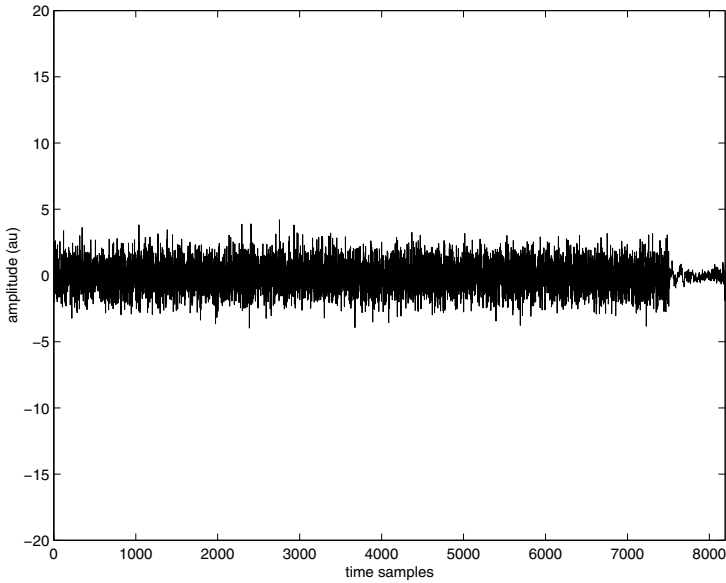


Figure 8.36 Difference between the original VAG signal in Figure 8.35 (a) and the MP-filtered version in Figure 8.35 (b). Reproduced from S. Krishnan and R. M. Rangayyan, Automatic de-noising of knee-joint vibration signals using adaptive time-frequency representations, *Medical and Biological Engineering and Computing*, 38(1):2–8, 2000, with kind permission from Springer Science+Business Media B.V. ©IFMBE and Springer.

$$\boldsymbol{\varepsilon} = \mathbf{y} - \tilde{\mathbf{y}} = \sum_{k=L+1}^K (x_k - \tilde{x}_k) \mathbf{W}_k. \quad (8.120)$$

The MSE is

$$\begin{aligned} \overline{\boldsymbol{\varepsilon}^2} &= E[\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}] \\ &= E\left[\sum_{k=L+1}^K \sum_{l=L+1}^K (x_k - \tilde{x}_k)(x_l - \tilde{x}_l) \mathbf{W}_k^T \mathbf{W}_l \right] \\ &= \sum_{k=L+1}^K E[(x_k - \tilde{x}_k)^2]. \end{aligned} \quad (8.121)$$

The last step above is due to the orthonormality of \mathbf{W} .

Taking the derivative of the MSE with respect to \tilde{x}_k and setting the result to zero, we get

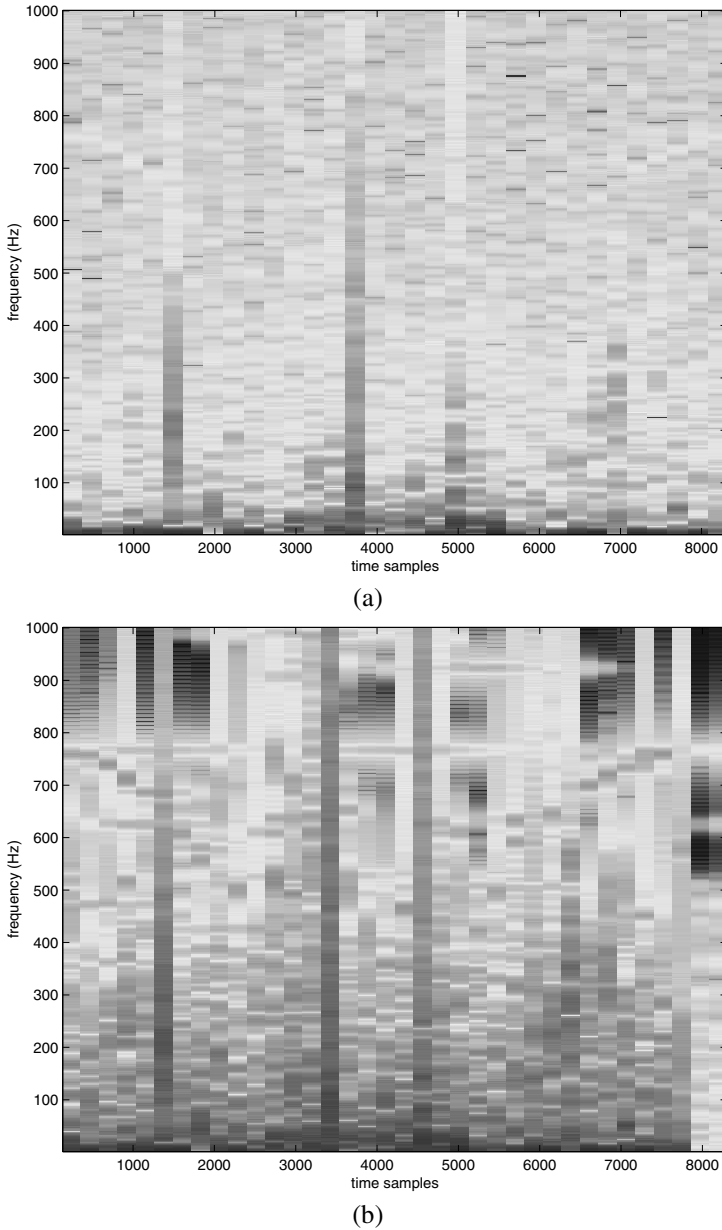


Figure 8.37 (a) TFD of the VAG signal in Figure 8.35 (a) and (b) the TFD of the MP-filtered VAG signal in Figure 8.35 (b) computed using the STFT. $f_s = 2,000 \text{ Hz}$. Reproduced from S. Krishnan and R. M. Rangayyan, Automatic de-noising of knee-joint vibration signals using adaptive time-frequency representations, *Medical and Biological Engineering and Computing*, 38(1):2–8, 2000, with kind permission from Springer Science+Business Media B.V. ©IFMBE and Springer.

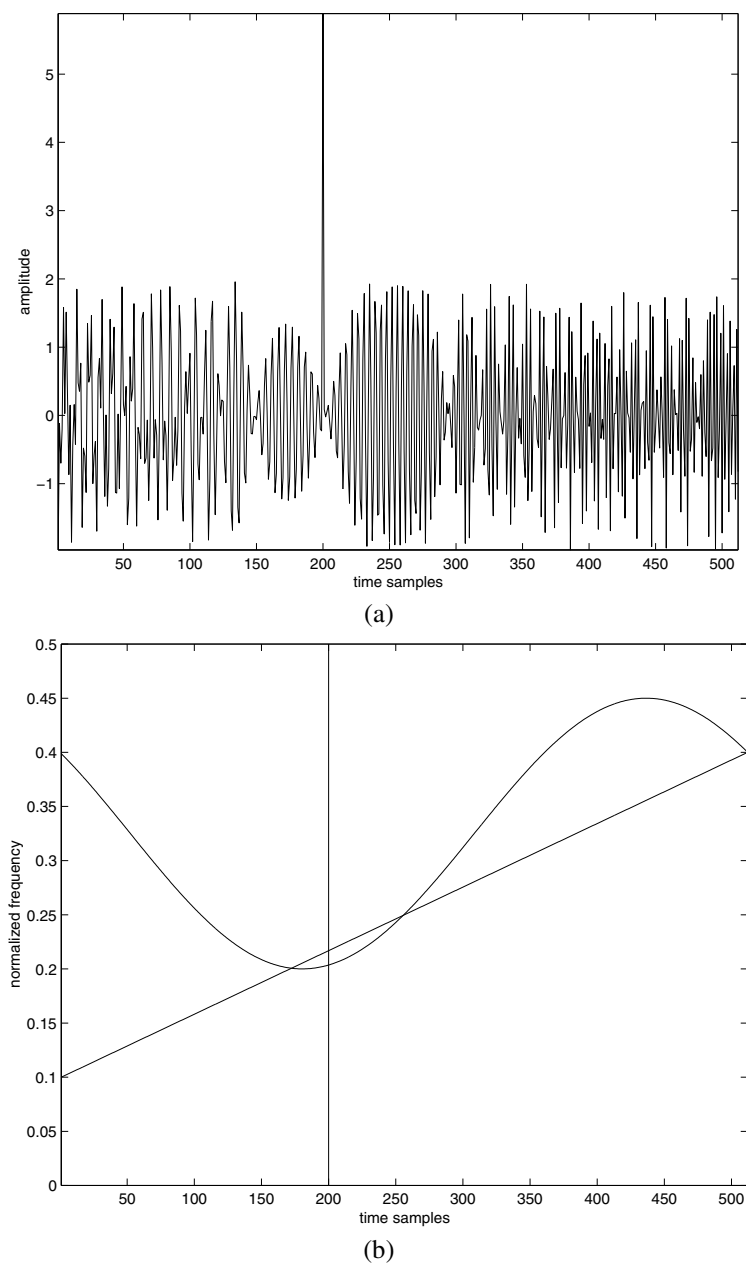


Figure 8.38 (a) Multicomponent and nonstationary synthetic signal consisting of overlapping chirp, impulse, and sinusoidal FM components and (b) its ideal TFD. Reproduced with permission from S. Krishnan, *Adaptive Signal Processing Techniques for Analysis of Knee Joint Vibroarthrographic Signals*, Ph.D. Thesis, University of Calgary, 1999.

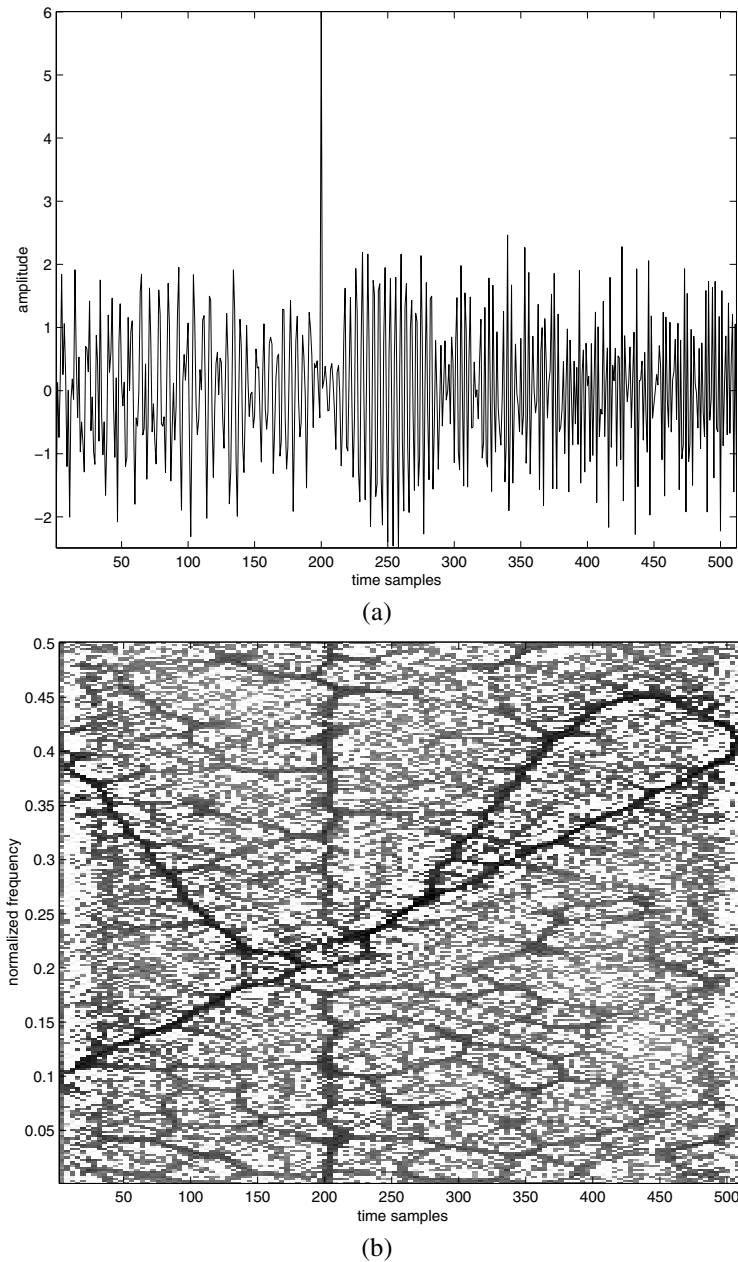


Figure 8.39 (a) The signal in Figure 8.38 (a) with random noise added such that $SNR = 10$ dB. (b) RSPWVD of the noisy signal. Reproduced with permission from S. Krishnan, Adaptive Signal Processing Techniques for Analysis of Knee Joint Vibroarthrographic Signals, Ph.D. Thesis, University of Calgary, 1999.

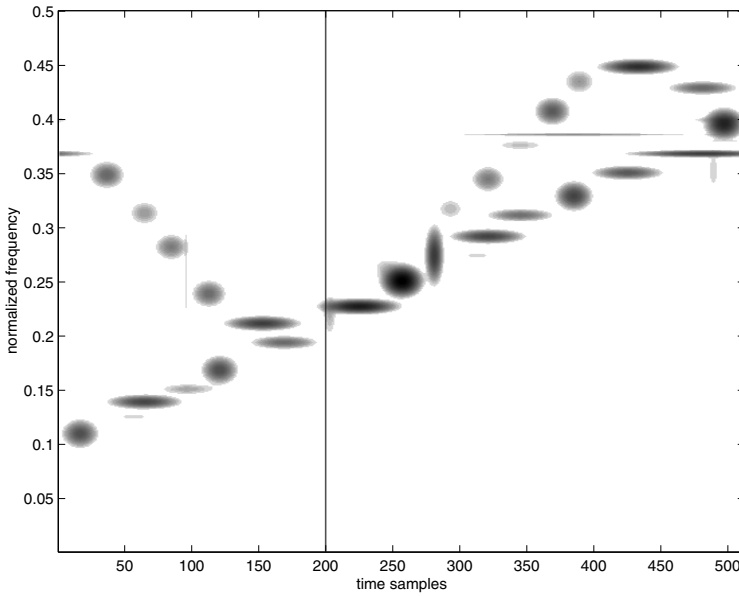


Figure 8.40 MPTFD of the noisy signal in Figure 8.39 (a). Reproduced with permission from S. Krishnan, Adaptive Signal Processing Techniques for Analysis of Knee Joint Vibroarthrographic Signals, Ph.D. Thesis, University of Calgary, 1999.

$$\frac{\partial \overline{\epsilon^2}}{\partial \tilde{x}_k} = -2 E[(x_k - \tilde{x}_k)] = 0. \quad (8.122)$$

The optimal MMSE choice for \tilde{x}_k is, therefore,

$$\tilde{x}_k = E[x_k] = \bar{x}_k = \mathbf{W}_k^{-1} E[y_k], \quad k = L + 1, L + 2, \dots, K. \quad (8.123)$$

This result indicates that the omitted components are replaced by their means estimated for the given population of the vectors \mathbf{y} . The MMSE is given by

$$\begin{aligned} \overline{\epsilon^2}_{\min} &= \sum_{k=L+1}^K E[(x_k - \bar{x}_k)^2] \\ &= \sum_{k=L+1}^K E[(\mathbf{W}_k^{-1} y_k - \mathbf{W}_k^{-1} \bar{y}_k)(\mathbf{W}_k^{-1} y_k - \mathbf{W}_k^{-1} \bar{y}_k)^T] \\ &= \sum_{k=L+1}^K E[\mathbf{W}_k^{-1} (y_k - \bar{y}_k)(y_k - \bar{y}_k)^T \{\mathbf{W}_k^{-1}\}^T] \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=L+1}^K E[\mathbf{W}_k^{-1}(y_k - \bar{y}_k)(y_k - \bar{y}_k)^T \mathbf{W}_k] \\
&= \sum_{k=L+1}^K \mathbf{W}_k^T \boldsymbol{\sigma}_y \mathbf{W}_k,
\end{aligned} \tag{8.124}$$

where $\boldsymbol{\sigma}_y = E[(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})^T]$ is the $K \times K$ covariance matrix of \mathbf{y} .

If the basis vectors \mathbf{W}_k are selected as the eigenvectors of $\boldsymbol{\sigma}_y$, that is,

$$\boldsymbol{\sigma}_y \mathbf{W}_k = \lambda_k \mathbf{W}_k, \tag{8.125}$$

where

$$\lambda_k = \mathbf{W}_k^T \boldsymbol{\sigma}_y \mathbf{W}_k \tag{8.126}$$

are the corresponding eigenvalues, then we have

$$\bar{\varepsilon}_{\min}^2 = \sum_{k=L+1}^K \lambda_k. \tag{8.127}$$

Therefore, the MSE may be minimized by ordering the eigenvectors such that the corresponding eigenvalues are arranged in decreasing order, that is, $\lambda_1 > \lambda_2 > \dots > \lambda_K$. Then, if a component x_k of \mathbf{x} is replaced by $\tilde{x}_k = \bar{x}_k$, the MSE increases by λ_k . By replacing the components of \mathbf{x} corresponding to the eigenvalues at the lower end of the list by their respective mean values, the MSE is kept at its minimum for a chosen number of components L .

From the procedure and the properties described above, it is evident that the components of \mathbf{x} are mutually uncorrelated:

$$\boldsymbol{\sigma}_x = \mathbf{W}^T \boldsymbol{\sigma}_y \mathbf{W} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_K \end{bmatrix} = \boldsymbol{\Lambda}, \tag{8.128}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix with the eigenvalues λ_k placed along its diagonal. Because the eigenvalues λ_k are equal to the variances of x_k , a selection of the larger eigenvalues implies the selection of the transformed components with the higher variance or information content across the ensemble of the data vectors considered.

Illustration of application: In a study reported by Mesin et al. [457], surface EMG signals were recorded from the biceps brachii muscle. The electrodes were placed 5 mm apart and two channels of EMG were acquired from approximately the same group of muscle fibers. Therefore, the two EMG signals, shown in Figure 8.41, parts (a) and (b), are expected to be highly correlated. The correlation between the two EMG signals is demonstrated in part (c) of the figure, which also

shows the directions of the two principal components (eigenvectors). The two principal components of the EMG signals are shown in parts (d) and (e) of the figure; whereas the first component is close in appearance to the two EMG signals, the second component could be considered to be random noise. By design, the two principal components are uncorrelated. Due to the high degree of correlation between the two input signals, there is considerable redundancy present in them. PCA has extracted most of the power in the signals within the first principal component, which may be considered to be an adequate and redundancy-free representation of both input channels. In this manner, PCA can achieve data reduction and efficient representation of information extracted from multiple measurements.

PCA is commonly used to achieve reduction in the dimension of feature vectors to facilitate efficient pattern classification.

8.8.2 Independent component analysis

Let us consider the situation where we have an observation with K channels of signals, labeled as $y_1(n), y_2(n), \dots, y_K(n)$. The observation at one instant of time may be expressed as the $K \times 1$ vector $\mathbf{y}(n) = [y_1(n), y_2(n), \dots, y_K(n)]^T$. Suppose that the observed signals are being produced by L sources $x_1(n), x_2(n), \dots, x_L(n)$ that are mutually statistically independent. Let us also assume that we are able to gather more than the required number of observations to determine the sources, that is, the system is overdetermined, with $K \geq L$. Then, letting the $L \times 1$ vector $\mathbf{x}(n) = [x_1(n), x_2(n), \dots, x_L(n)]^T$ represent the values of the sources at the instant of time n , we have

$$\mathbf{y}(n) = \mathbf{M} \mathbf{x}(n) + \eta(n), \quad (8.129)$$

where η is random noise that is assumed to be independent of \mathbf{y} and \mathbf{x} . This model relates the K observed values in $\mathbf{y}(n)$ at the instant of time n to the corresponding values of the L sources in $\mathbf{x}(n)$ through the $K \times L$ mixing matrix \mathbf{M} . The problem in ICA and BSS is to derive the values of \mathbf{x} from the given observations \mathbf{y} without knowing \mathbf{M} (hence the term “blind” in BSS). We may assume that there is no noise in the observed data or let η be one of the sources contained in \mathbf{x} , and simplify Equation 8.129 to

$$\mathbf{y}(n) = \mathbf{M} \mathbf{x}(n). \quad (8.130)$$

PCA, as described in Section 8.8.1, facilitates the decomposition of a data vector into uncorrelated components. If the sources of the data or signal components have Gaussian PDFs, the property of being uncorrelated implies that the components are statistically independent; this, however, is not always true. PCA also does not assume any mixing matrix, and characterizes the redundancy in the input using its covariance matrix. Thus, PCA is based on second-order statistics.

ICA achieves separation of the sources by assuming them to be statistically independent and not only being uncorrelated as in PCA. Various measures of statistical

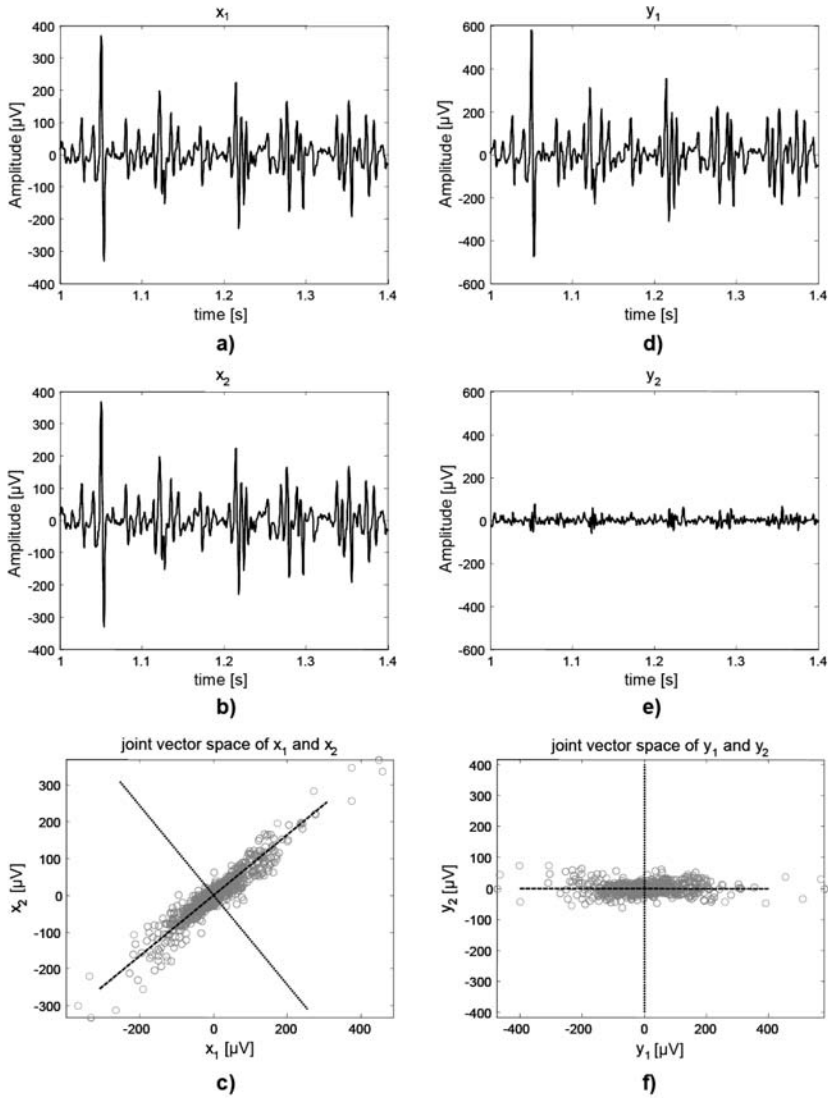


Figure 8.41 Illustration of PCA of correlated EMG signals. (a) and (b) Two channels of EMG signals from the biceps brachii muscle. (c) Directions of the principal components and demonstration of the correlation between the two EMG channels. (d) and (e) Principal components extracted via PCA. (f) Joint vector space of the principal components. Reproduced with permission from L. Mesin, A. Holobar, and R. Merletti, Blind source separation: Application to biomedical signals, Chapter 15 in *Advanced Methods of Biomedical Signal Processing*, IEEE and Wiley, 2011, Edited by S. Cerutti and C. Marchesi, pp. 379–409. ©IEEE.

independence can be used in optimization procedures to estimate a separation or “un-mixing” matrix \mathbf{W} (of size $L \times K$) such that estimates of the original source signals can be obtained as

$$\tilde{\mathbf{x}}(n) = \mathbf{W} \mathbf{y}(n). \quad (8.131)$$

The central limit theorem [5] states that the PDF of a linear combination of several statistically independent random variables tends towards a Gaussian PDF regardless of the PDFs of the individual variables. Therefore, the mixed observed signal \mathbf{y} in Equation 8.130 is likely to have a PDF that is close to a Gaussian regardless of the PDFs of the sources of the input signals \mathbf{x} . One approach to achieve BSS is to employ means to minimize the Gaussian characteristics of the reconstructed or estimated signals $\tilde{\mathbf{x}}$. The kurtosis excess (see Section 3.2.1) of a Gaussian PDF is zero: This property may be used to construct a function $F(\tilde{\mathbf{x}})$ that can be used in an iterative optimization procedure to estimate the separation matrix as [457]

$$\tilde{\mathbf{W}}_{(k+1)} = \tilde{\mathbf{W}}_k - \mu_k \nabla F(\tilde{\mathbf{x}}), \quad (8.132)$$

where k is the iteration number, μ is the learning rate, and ∇ is a gradient measure. This procedure is a gradient descent approach for optimization (see Section 3.9.2). Several other algorithms exist for ICA and BSS [65, 449–463]; some BSS methods use PCA as a preprocessing step.

Illustration of application: Zarzoso and Nandi [190] performed comparative analysis of BSS with adaptive filtering (see Section 3.9.1) for the purpose of extraction of the fetal ECG from maternal ECG recordings. Figure 8.42 shows the input ECG signals used in their work, including eight channels of the maternal ECG obtained using abdominal and thoracic surface electrodes. The result of BSS is shown in Figure 8.43. Although BSS separates the mixed input signals and provides independent components, the method, by itself, does not identify the various sources. However, knowing that the fetal heart rate is typically higher than the maternal heart rate, one can determine, by visual inspection, that the fifth and seventh traces from the top in Figure 8.43 represent the extracted fetal ECG over the interval of the recordings with 24 beats. Similarly, the first, second, and fourth traces can be taken to be representations of the maternal ECG (with 14 beats over the same interval) without interference from the fetal ECG. The third, sixth, and eighth traces do not show any QRS complexes with a regular rhythm and, hence, could be considered to be noise. If the derivation of the fetal heart rate is the goal of the recording, it can be easily achieved by using one of the extracted fetal ECG channels. The maternal heart rate may also be estimated from one of the extracted maternal ECG channels without interference from the fetal ECG.

It should be observed that BSS could lead to multiple representations of a single source in the resulting output channels. In the present example, the electrical activity of the maternal heart is shown in three output channels and that of the fetal heart is shown in two channels; the user is required to analyze the results and select appropriate channels for further analysis. (Indeed, all of the eight input channels demonstrate various views of the maternal ECG, with some of them including the fetal ECG as

well.) Zarzoso and Nandi [190] showed that BSS performs better than adaptive filtering in the extraction of fetal ECG from maternal ECG signals. See Callaerts et al. [464], Vanderschoot et al. [465], and De Lathauwer et al. [453] for other methods for the same application.

8.9 Application: Adaptive Segmentation of EEG Signals

Problem: *Propose a method for parametric representation of nonstationary EEG signals.*

Solution: Bodenstein and Praetorius [197] applied their adaptive segmentation procedure based upon the *SEM* (see Section 8.5.1) for representation and analysis of EEG signals with the following propositions.

1. An EEG signal consists of quasistationary segments upon which transients may be superimposed.
2. A segment is specified by its time of occurrence, duration, and PSD (represented by its AR-model coefficients). A transient is specified by its time of occurrence and a set of graph elements (or directly by its samples).
3. An EEG signal consists of a finite number of recurrent states.

It should be noted that, whereas the adaptive segments have variable length, each adaptive segment is represented by the same number of AR-model coefficients. The number of parameters is, therefore, independent of segment duration, which is convenient when pattern classification techniques are applied to the segments. Since the AR model is computed once at the beginning of each segment and some prediction error is permitted in the moving analysis window, the initial AR model may not adequately represent the entire adaptive segment. A new model may be computed using the signal samples over the entire duration of each adaptive segment. Instead, Bodenstein and Praetorius maintained the initial AR model of order P of each adaptive segment, and an additional *corrective predictor* of order M was derived for each adaptive segment using the ACF of the prediction error which is computed and readily available in the segmentation procedure. Each adaptive segment was then represented by the $(P + M)$ AR-model coefficients, the associated prediction error *RMS* values, and the segment length. The PSD of the segment may be derived from the two sets of AR-model coefficients.

With the EEG signals bandpass filtered to the range $1 - 25$ Hz and sampled at 50 Hz in the work of Bodenstein and Praetorius [197], the ACF window length was set to be 2 s with $2N + 1 = 101$ samples. Bodenstein and Praetorius used the rule of thumb that the AR-model order should be at least twice the number of expected resonances in the PSD of the signal. Short segments of EEG signals rarely demonstrate more than two spectral peaks, which suggests that an AR-model order of $P = 5$ should be adequate. Regardless, Bodenstein and Praetorius used $P = 8$, which met the Akaike criterion as well (see Section 7.5.2). The order of the ACF of the prediction error and the associated corrective predictor was set to a low value of

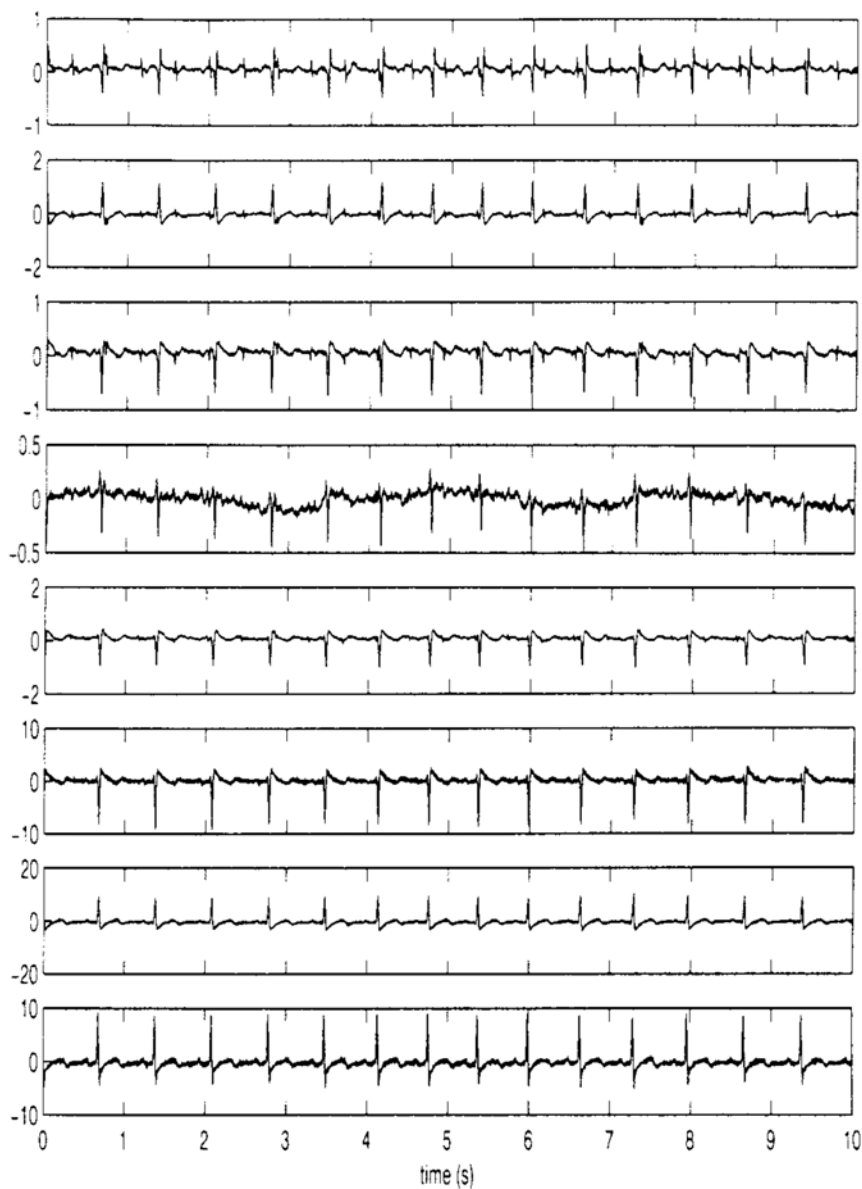


Figure 8.42 Eight channels of maternal ECG signals using abdominal and thoracic surface electrodes. Reproduced with permission from V. Zarzoso and A.K. Nandi, Noninvasive fetal electrocardiogram extraction: blind separation versus adaptive noise cancellation, *IEEE Transactions on Biomedical Engineering*, 48(1):12–18, 2001. ©IEEE.

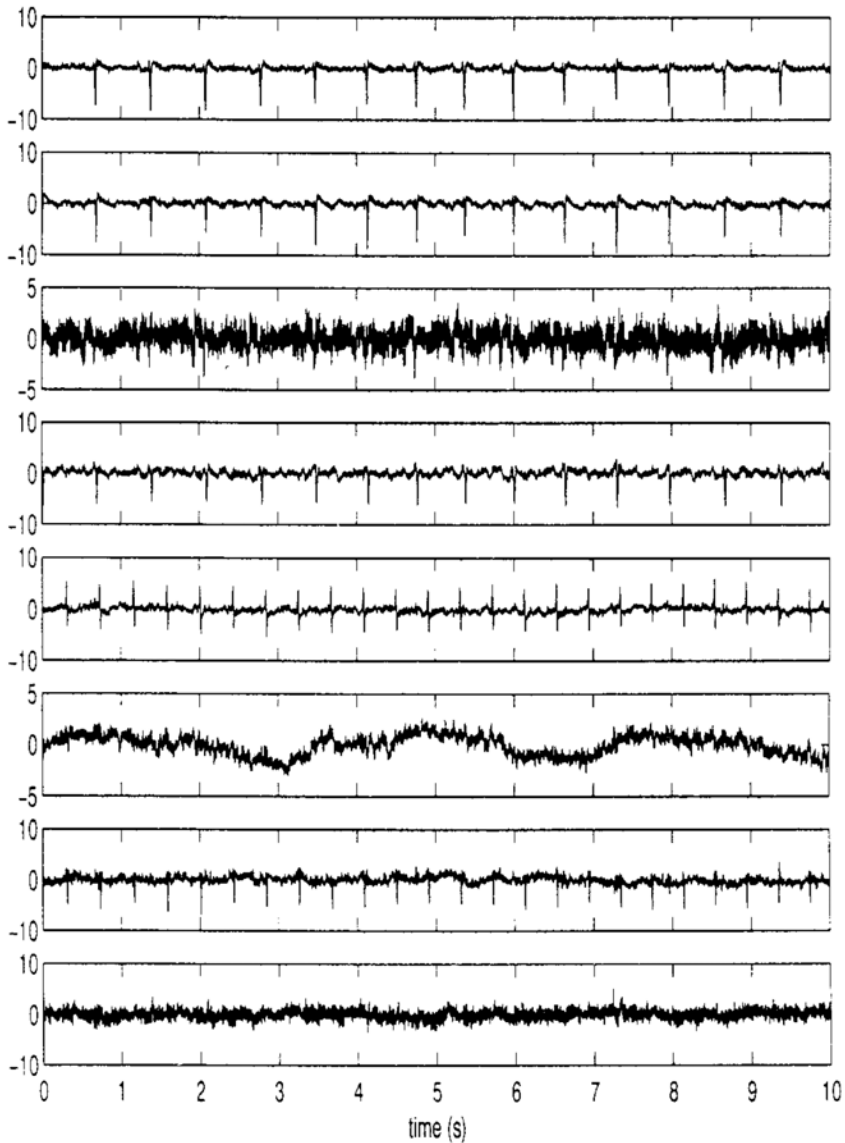


Figure 8.43 Results of BSS showing eight independent components. The fifth and seventh traces from the top show the extracted fetal ECG with 24 beats over the 10-s interval of the recordings. The first, second, and fourth traces show the maternal ECG with 14 beats over the same interval. The remaining traces represent noise. Reproduced with permission from V. Zarzoso and A.K. Nandi, Noninvasive fetal electrocardiogram extraction: blind separation versus adaptive noise cancellation, *IEEE Transactions on Biomedical Engineering*, 48(1):12–18, 2001. ©IEEE.

$M = 3$, allowing for one spectral peak (the error should ideally have a flat PSD). The thresholds were defined as $Th_1 = 0.5$ (empirical) and $Th_2 = 2.5\sigma$, where σ is the *RMS* value of the prediction error (see Section 8.5.1). The range of $20\sigma^2$ to $40\sigma^2$ was recommended for Th_3 . A transitional delay of 25 samples was allowed between each segmentation boundary and the starting point of the following fixed window to prevent the inclusion of the spectral components of one segment into the following segment.

Figure 8.44 shows a few examples of adaptive segmentation of EEG signals. A clustering procedure was included to remove spurious boundaries, some examples of which may be seen in Figure 8.44 (d); neighboring segments with similar parameters were merged in a subsequent step. Visual inspection of the results indicates that most of the adaptive segments are stationary (that is, they have the same appearance) over their durations. It is worth noting that the longest segment in Figure 8.44 (d) of duration 16 s or 800 samples is represented by just 12 parameters.

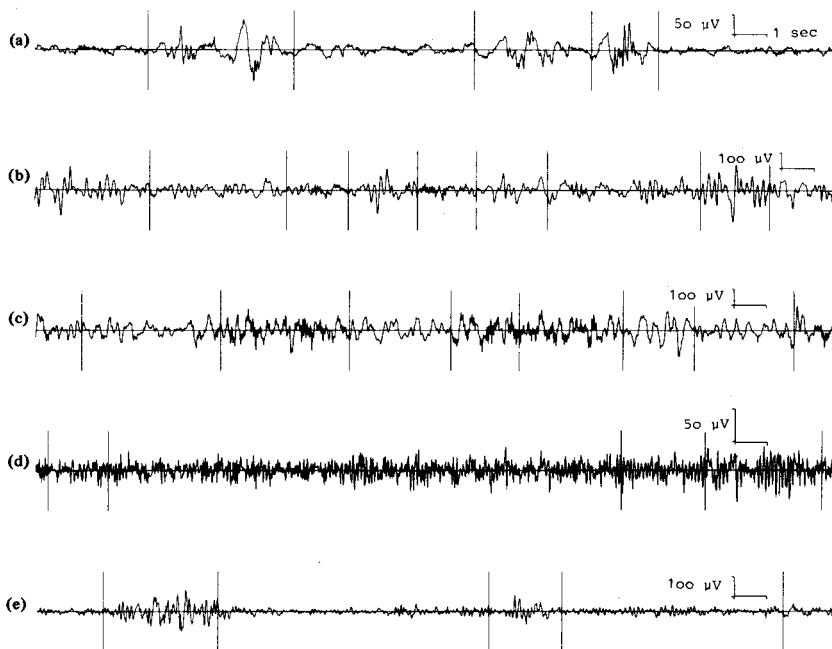


Figure 8.44 Examples of segmentation of EEG signals. (a) Newborn in non-REM sleep. REM = rapid eye movement. (b) Child of age 7 years in sleep stage I. (c) Child of age 8 years in sleep stage III. (d) Alpha rhythm of an adult. (e) EEG of an adult with paroxysms. The vertical lines represent the segmentation boundaries. Reproduced with permission from G. Bodenstein and H.M. Praetorius, Feature extraction from the electroencephalogram by adaptive segmentation, *Proceedings of the IEEE*, 65(5):642–652, 1977. ©IEEE.

Figure 8.45 shows examples of detection of transients in two contralateral channels of the EEG of a patient with epilepsy. The EEG signal between seizures (inter-

ictal periods) is expected to exhibit a large number of sharp waves. The length of the arrows shown in the figure was made proportional to the cumulated suprathreshold part of the squared prediction error in order to indicate how pronounced the event was regarded to be by the algorithm.

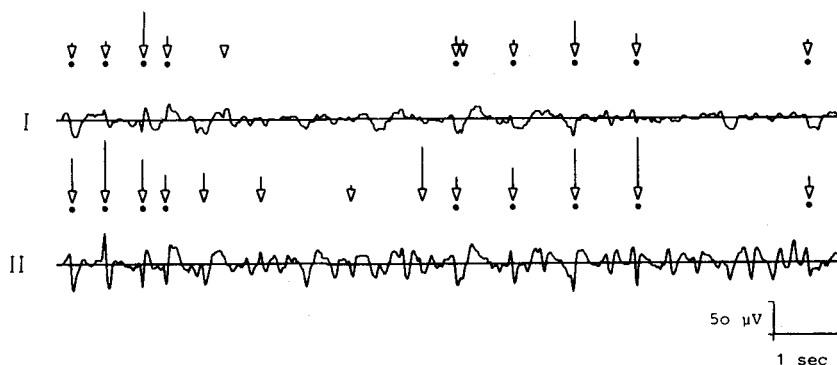


Figure 8.45 Example of detection of transients in the EEG signal of a patient with epilepsy. The signals shown are from contralateral channels between seizures (interictal period). The longer the arrow, the more pronounced the transient detected at the corresponding time instant. Transients detected simultaneously in the two contralateral channels are marked with dots. Reproduced with permission from G. Bodenstein and H.M. Praetorius, Feature extraction from the electroencephalogram by adaptive segmentation, *Proceedings of the IEEE*, 65(5):642–652, 1977. ©IEEE.

The method was further extended to parallel analysis of multichannel EEG signals by Bodenstein et al. [413] and Creutzfeldt et al. [414]. Procedures were proposed for computerized pattern classification and labeling of EEG signals, including clustering of similar segments and state diagrams indicating the sequence of the types of activity found in an EEG record. Figure 8.46 illustrates the record produced by the application of the procedure to two channels of an EEG signal. Typical EEG segments belonging to the four clusters detected in the signal are shown on the left and right sides of the upper portion of the figure. Each signal segment is labeled with the frequencies (FRQ, in Hz) and amplitudes (AMP, in μV) of the resonances detected using an eighth-order AR model. The central column of the upper portion of the figure illustrates the PSDs of the corresponding segments on the left (solid line) and right (dashed line). The middle portion of the figure provides the state diagram, indicating the transitions between the four states (represented by the four clusters of the EEG segments) detected in the two channels of the signal. The states represent: 1, background; 2, eyes open; 3, paroxysm; and 4, epileptiform spike-and-wave complexes. The values on the right of the state diagram give the percentage of the total duration of the signal for which the EEG was in the corresponding states. The bottom portion of the figure illustrates singular events, that is, segments that could not be grouped with any of the four clusters. It was indicated that the segments of most EEG signals could be clustered into at most five states, and that the summarized

record as illustrated in Figure 8.46 could assist clinicians in analyzing lengthy EEG records in an efficient manner.

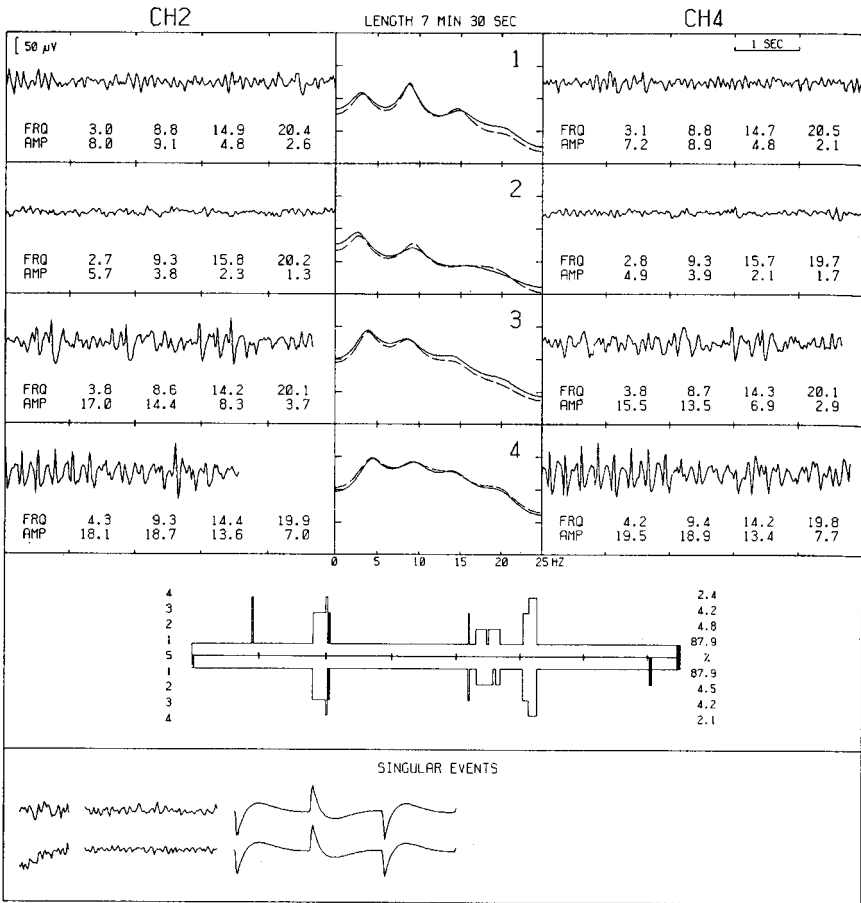


Figure 8.46 Example of application of segmentation and pattern analysis to the EEG signal of a patient with epileptiform activity. Refer to the text for details. Reproduced with permission from G. Bodenstein, W. Schneider, and C.V.D. Malsburg, Computerized EEG pattern classification by adaptive segmentation and probability-density-function classification. Description of the method, *Computers in Biology and Medicine*, 15(5):297–313, 1985. ©Elsevier Science.

8.10 Application: Adaptive Segmentation of PCG Signals

We have noted several times that the PCG signal is nonstationary. Let us now assess the feasibility of adaptive segmentation of PCG signals using the RLSL method (see Section 8.6.2), with no other signal being used as a reference.

Figure 8.47 illustrates the results of segmentation of the PCG signal of a normal subject. The top trace shows the PCG signal over three cardiac cycles; the segment boundaries detected are indicated by the vertical dotted lines as well as by the triangular markers on the time axis. The second trace illustrates a plot of the conversion factor γ_c : The conversion factor drops from unity whenever there is a change in the signal's characteristics, in particular at the boundaries of S1 and S2. A threshold of 0.995 (indicated by the horizontal line overlaid on the second trace) applied to γ_c and a condition imposing a minimum segment length of 50 samples (50 *ms*) were used to obtain the segment boundaries. The third and fourth traces illustrate the ECG and carotid pulse signals of the subject acquired simultaneously with the PCG. The segment boundaries obtained by the RLSL method agree well with the readily noticeable S1 and S2 boundaries as well as the QRS and dicrotic notch positions. (See also Sections 1.2.9, 2.3, and 4.10 for related details.)

Figure 8.48 illustrates the results of adaptive segmentation of the PCG signal of a subject with systolic murmur due to aortic stenosis. The results in this case, however, are not as clear or as easy to interpret as in the preceding case. The method has identified the beginning of S1 and S2; furthermore, the split nature of S2 has been identified by an additional segment boundary within each S2. However, the method has not reliably identified the boundaries between the episodes of S1 and systolic murmur illustrated: The condition on the minimum segment length has affected the placement of the segment boundary after the beginning of S1. Use of other conditions on γ_c may provide better segmentation results.

8.11 Application: Time varying Analysis of HRV

The heart rate is controlled by the ANS and the CNS: The vagal and sympathetic activities lead to a decrease or increase, respectively, in the heart rate (see Section 1.2.5). We saw in Sections 2.2.4 and 7.8 how respiration affects heart rate and how Fourier analysis may be extended to analyze HRV. When heart rate data such as beat-to-beat *RR* intervals are collected over long periods of time (several hours), the signal could be expected to be nonstationary.

Bianchi et al. [400] extended AR-modeling techniques for time-variant PSD analysis of HRV data in order to study transient episodes related to ischemic attacks. The prediction error was weighted with a forgetting factor, and a time-varying AR model was derived. The RLS algorithm was used to update the AR-model coefficients at every *RR* interval sample (every cardiac cycle). The AR coefficients were then used to compute a time-varying PSD. The following frequency bands were indicated to be of interest in the analysis of *RR* interval PSDs: VLF band in the range 0 – 0.03 *Hz* related to humoral and thermoregulatory factors; low-frequency (LF)

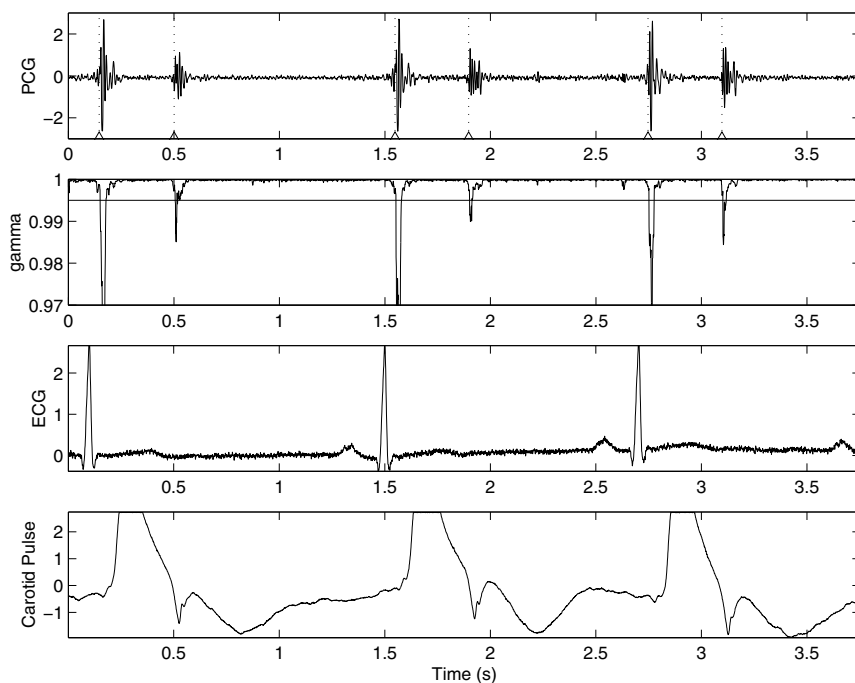


Figure 8.47 Adaptive segmentation of the PCG signal of a normal subject using the RLSL method. Top to bottom: PCG signal (the vertical dotted lines and triangular markers represent the segmentation boundaries); conversion factor γ_c (the horizontal line is the threshold used); ECG; carotid pulse (clipped due to saturation).

band in the range $0.03 - 0.15 \text{ Hz}$ related to sympathetic activity; high-frequency (HF) band in the range $0.18 - 0.4 \text{ Hz}$ related to respiration and vagal activity.

(Note: According to the standards of measurement, physiological interpretation, and clinical use of HRV published by the Task Force of the European Society of Cardiology and the North American Society of Pacing Electrophysiology [466], the frequency bands of HRV are slightly different from those given above and are: VLF, $\leq 0.04 \text{ Hz}$; LF, $0.04 - 0.15 \text{ Hz}$; and HF, $0.15 - 0.4 \text{ Hz}$.)

Figure 8.49 shows an RR interval series including an ischemic episode (delineated by B for beginning and E for ending points, respectively). Figure 8.50 shows the time-varying PSD in the form of a spectrogram. Figure 8.51 shows a segment of RR interval data and a few measures derived from the data.

Some of the important observations made by Bianchi et al. (and illustrated by the spectrogram in Figure 8.50 and the parameters in Figure 8.51) are as follows:

- There is an increase in LF power about $1.5 - 2 \text{ minutes}$ before an ischemic event.
- The RR variance decreases as an episode begins.

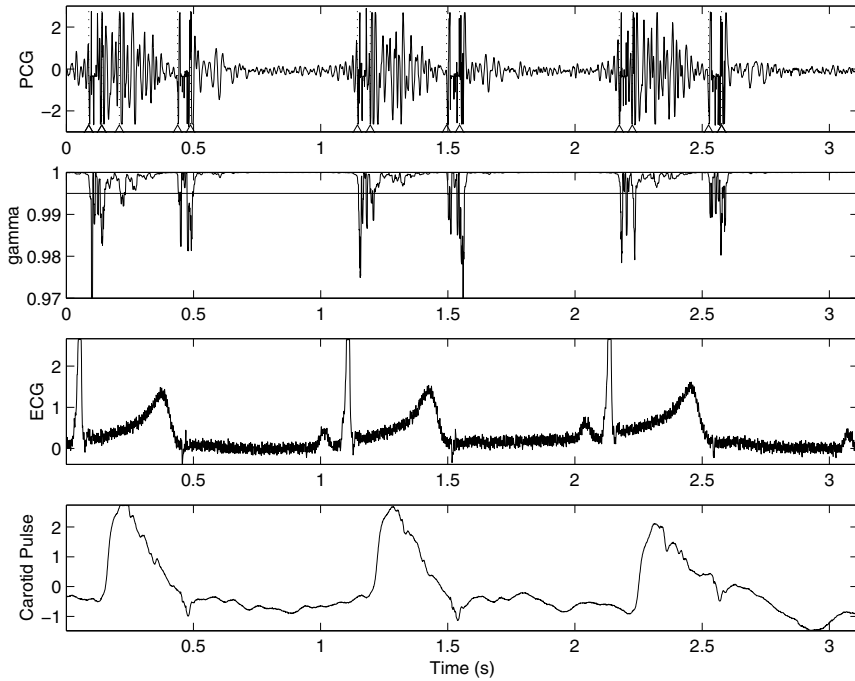


Figure 8.48 Adaptive segmentation of the PCG signal of a subject (female, 11 years) with systolic murmur due to aortic stenosis. Top to bottom: PCG signal (the vertical lines and triangular markers represent the segmentation boundaries); conversion factor γ_c (the horizontal line is the threshold used); ECG; carotid pulse.

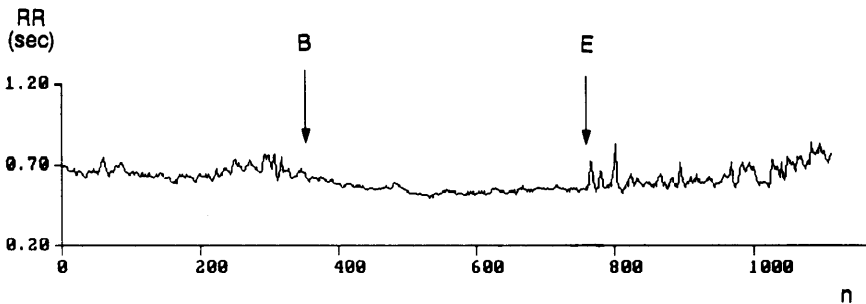


Figure 8.49 *RR* interval series including an ischemic episode. B denotes the beginning and E denotes the end of the episode. Reproduced with permission from A.M. Bianchi, L. Mainardi, E. Petrucci, M.G. Signorini, M. Mainardi, and S. Cerutti, Time-variant power spectrum analysis for the detection of transient episodes in HRV signal, *IEEE Transactions on Biomedical Engineering*, 40(2):136–144, 1993. ©IEEE.

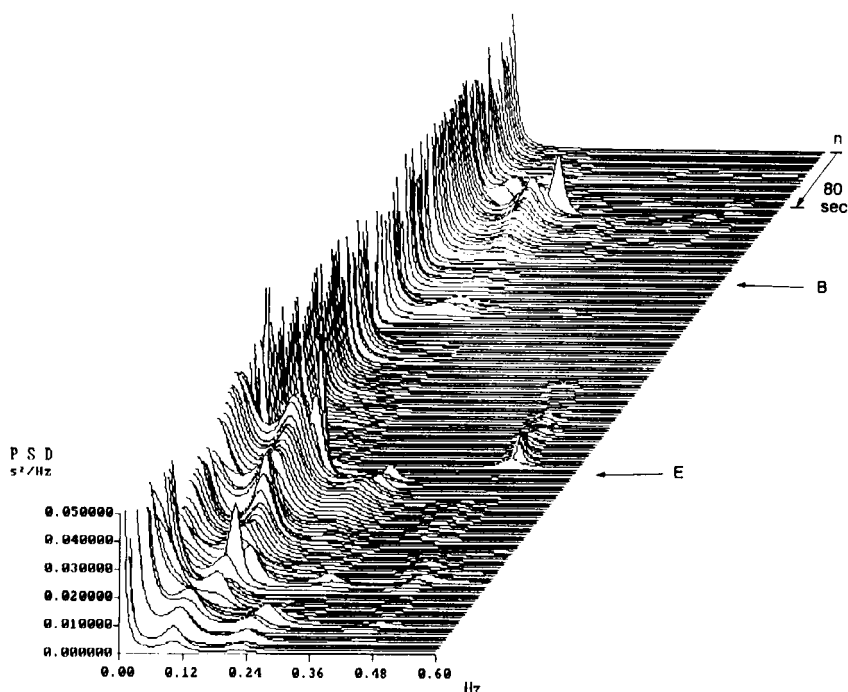


Figure 8.50 Spectrogram of the *RR* interval series in Figure 8.49. Time progresses from the top to the bottom. B denotes the beginning and E denotes the end of an ischemic episode. Reproduced with permission from A.M. Bianchi, L. Mainardi, E. Petrucci, M.G. Signorini, M. Mainardi, and S. Cerutti, Time-variant power spectrum analysis for the detection of transient episodes in HRV signal, *IEEE Transactions on Biomedical Engineering*, 40(2):136–144, 1993. ©IEEE.

- There is a predominant rise in LF power at the end of an ischemic episode.
- A small HF component appears toward the end of an episode.
- Early activation of an LF component precedes tachycardia and ST displacement in the ECG that are generally indicative of the onset of an ischemic episode.
- The results suggest an arousal of the sympathetic system before an acute ischemic attack.

Time-varying AR-modeling techniques have also been applied for the analysis of EEG signals [398]. Time-varying ARMA modeling techniques have been applied to analyze EGG signals [69].

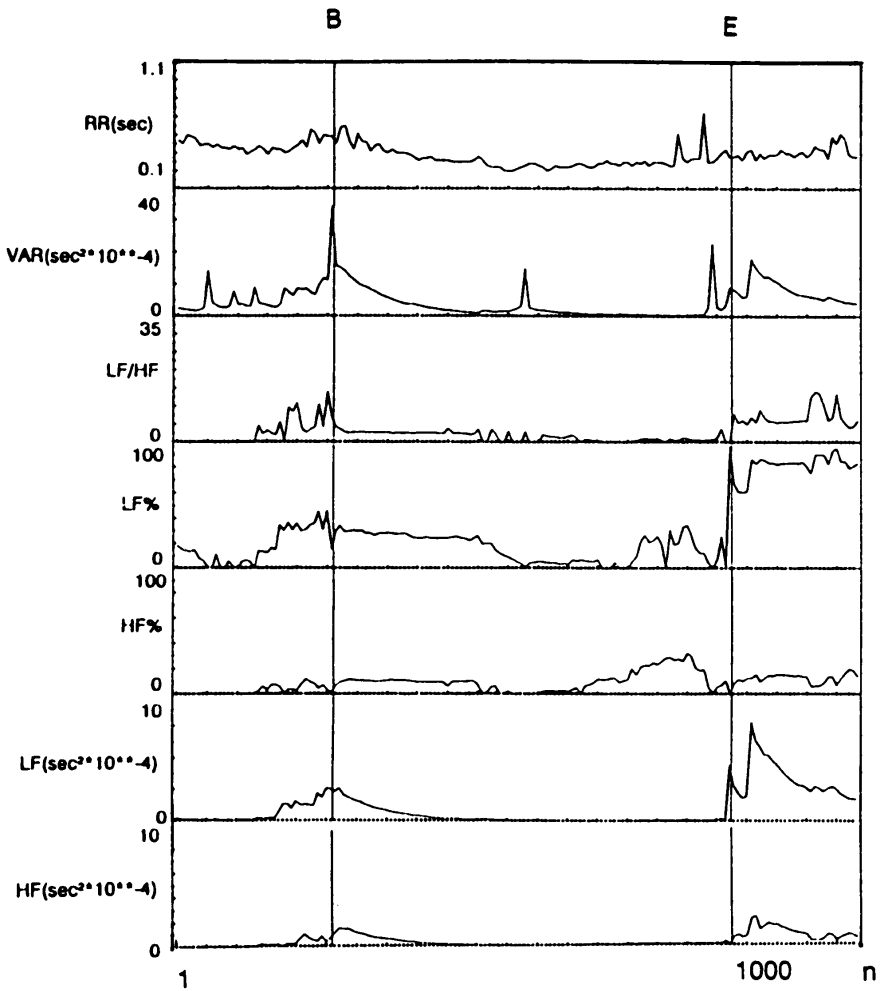


Figure 8.51 Top to bottom: RR interval series including an ischemic episode; variance; low-frequency (LF) to high-frequency (HF) power ratio; percentage of LF power; percentage of HF power; LF power; and HF power. B denotes the beginning and E denotes the end of the episode. Reproduced with permission from A.M. Bianchi, L. Mainardi, E. Petrucci, M.G. Signorini, M. Mainardi, and S. Cerutti, Time-variant power spectrum analysis for the detection of transient episodes in HRV signal, *IEEE Transactions on Biomedical Engineering*, 40(2):136–144, 1993. ©IEEE.

8.12 Application: Detection of Epileptic Seizures in EEG Signals

Zhou et al. [467] proposed measures of lacunarity and fluctuation intensity (FI) of DWT coefficients of EEG signals for the detection of epileptic seizures. EEG signals were subjected to DWT-based decomposition with the Daubechies 4 wavelet and five scales, and wavelet coefficients at scales 3, 4, and 5 were selected for further processing. The selection of the scales mentioned above was based on the observation that seizure signals in intracranial EEG signals usually occur in the range of 3 to 29 Hz .

Noting that ictal EEG segments typically display larger fluctuations than interictal segments, the feature FI was defined to measure the intensity of the fluctuations in an EEG signal as

$$FI(s) = \frac{1}{N} \sum_{i=1}^{N-1} |d(i+1) - d(i)|, \quad (8.133)$$

where N is the number of DWT coefficients $d(i)$ for scale s . It was demonstrated that the values of FI during seizures were typically greater than those during other periods.

Lacunarity is a scale-dependent measure of heterogeneity of a signal or image, and quantifies the presence of gaps [275, 468]. If $p(m, l)$ represents the probability of data samples whose amplitude is equal to m , l represents the length of an epoch, and $[A, B]$ is the range of values of the data samples, it follows that

$$\sum_{m=A}^B p(m, l) = 1. \quad (8.134)$$

With

$$M_1(l) = \sum_{m=A}^B m p(m, l) \quad (8.135)$$

and

$$M_2(l) = \sum_{m=A}^B m^2 p(m, l), \quad (8.136)$$

the measure of lacunarity was obtained as

$$\lambda(l) = \frac{M_2(l) - [M_1(l)]^2}{[M_1(l)]^2}. \quad (8.137)$$

Zhou et al. [467] computed lacunarity values of the EEG wavelet coefficients at scales 3, 4, and 5. It was demonstrated that the values of lacunarity during seizures were typically lower than those during other periods.

Including further postprocessing and classification steps, Zhou et al. [467] analyzed a data set of 289.14 h of intracranial EEG data recorded from 21 patients, and obtained a sensitivity of 96.25% in the detection of epileptic seizures with a false detection rate of 0.13/ h and a mean delay time of 13.8 s . See Zhou et al. [467] for

a comparative analysis of their results with those obtained in other recent works on the detection of epileptic seizures in EEG signals.

8.13 Application: Analysis of Crying Sounds of Infants

Várallyay [114] developed techniques to analyze time-varying spectral patterns in crying sounds of infants. It was observed that during a crying episode, the volume, pitch, and tone of the signal vary over time. The varying patterns of the pitch (or fundamental frequency) was referred to as the *cry melody*. Just as the melody of normal speech expresses the mood and intention of a speaker, the melody of crying is expected to relate to the physical and physiological state of the infant. Cry melodies are expected to possess characteristic patterns for various states, such as pain, hunger, discomfort, and boredom. It was also noted that, whereas a person with normal hearing can control the sound or speech produced, a person with hearing impairment is unable to exercise such control due to the lack of auditory feedback. Hirschberg [469, 470] observed high-pitched cry melodies with fundamental frequencies in the range of 1,000 to 2,000 Hz in the case of infants with dysphonia, as compared to the range of 400 to 500 Hz for cries of healthy infants.

The spectral structure of a cry segment is expected to be regular, containing a fundamental frequency and its harmonics. Dynamic variations of the fundamental frequency characterize the cry melody. In order to follow the variations of the fundamental frequency over time, Várallyay [114] divided cry segments into short-time windows of duration 40 ms . The fundamental frequencies of several infant cries that were analyzed were found to vary between 200 and 1,000 Hz , with the most common values being between 300 and 600 Hz . In order to follow variations of the fundamental frequency, three fundamental units of cry melody patterns were defined as falling (-1), flat (0), and rising ($+1$). Cry melodies were coded as sequences of these fundamental patterns.

Figure 8.52 shows the temporal and spectral variations in a cry signal. The top-most trace shows the time-varying amplitude of five cry segments. The second trace shows the spectrogram in which the time-varying patterns of the fundamental frequency and its harmonics (that is, the cry melody) are clearly seen. The two traces at the bottom of the figure show the cry melody as a function of time; the lowest trace has five guidelines superimposed on the cry melody to facilitate its interpretation over predefined frequency ranges.

Várallyay [114] analyzed 2,762 crying sounds recorded from 316 infants. The mean value of the fundamental frequency of the melodies was higher for infants with hearing impairment (425.51 ± 78.10 Hz) than for the control group (408.74 ± 64.77 Hz).

8.14 Application: Adaptive Time frequency Analysis of VAG Signals

Krishnan et al. [273] proposed several methods to derive time-varying parameters from TFDs for the analysis and classification of VAG signals. Specifically, four

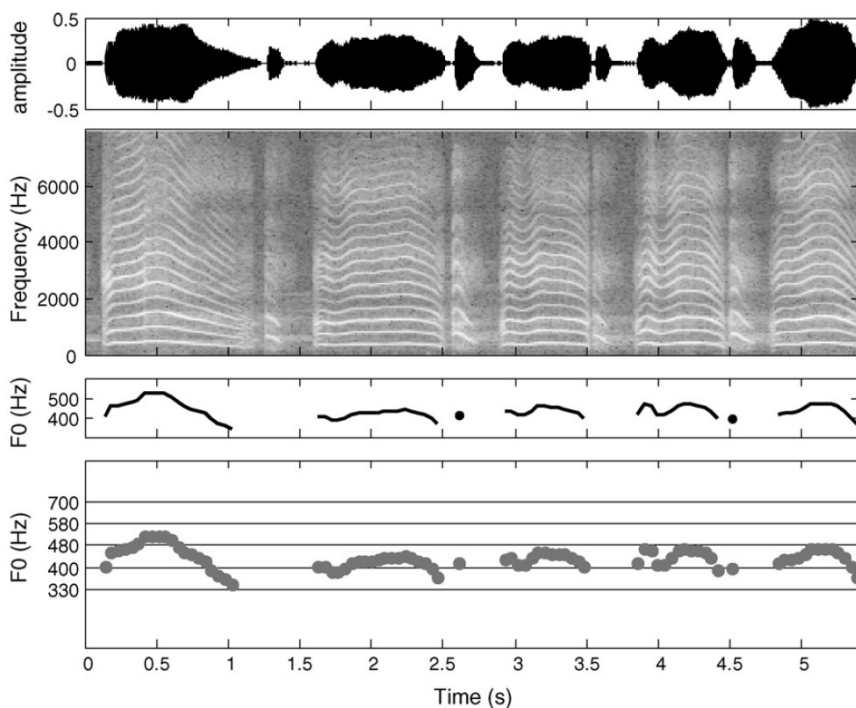


Figure 8.52 Top to bottom: Time-varying amplitude of five cry segments; spectrogram; cry melody as a function of time; the previous item with five guidelines superimposed. Reprinted with permission from G. Várallyay Jr., The melody of crying, *International Journal of Pediatric Otorhinolaryngology*, 71(11):1699–1708, 2007, ©Elsevier.

TF features were derived from OMPTFDs (see Section 8.7.5) of VAG signals: the energy parameter (EP), the energy spread parameter (ESP), the frequency parameter (FP), and the frequency spread parameter (FSP). EP was computed as the mean of the OMPTFD $M(t, \omega)$ along each time slice, which gives a representation of energy variation with time as

$$\text{EP}(t) = \frac{1}{\omega_m} \sum_{\omega=0}^{\omega_m} M(t, \omega), \quad (8.138)$$

where ω_m is the maximum frequency present in the signal. Signals generated by pathological knees are expected to be highly time-variant (nonstationary) because of differences in cartilage roughness and nonuniformity. Thus, $\text{EP}(t)$ of an abnormal VAG signal is expected to show large variations over time. Although $\text{EP}(t)$ may be computed directly from the signal, its derivation as above from the OMPTFD includes the benefits of adaptive filtering because only the coherent structures in the signal are retained (see Section 8.7.2).

The parameter ESP, computed as the SD of $M(t, \omega)$ along each time slice, measures the spread of energy over frequency for each time slice as

$$\text{ESP}(t) = \left[\frac{1}{\omega_m} \sum_{\omega=0}^{\omega_m} [M(t, \omega) - \text{EP}(t)]^2 \right]^{\frac{1}{2}}. \quad (8.139)$$

ESP could serve as a feature suitable for the analysis of multicomponent signals. Abnormal VAG signals generated as a result of friction between rough cartilage surfaces may have more components because of the nonuniformity of the surfaces, and hence a higher ESP than normal VAG signals.

The parameter FP, also known as IMF, is given by the first moment of $M(t, \omega)$ along each time slice, as

$$\text{FP}(t) = \frac{\sum_{\omega=0}^{\omega_m} \omega M(t, \omega)}{\sum_{\omega=0}^{\omega_m} M(t, \omega)}. \quad (8.140)$$

FP characterizes the frequency dynamics of the signal. Movement of the knee during VAG signal acquisition may cause linear or nonlinear FM components, with the modulation index depending on the state of lubrication, stiffness, and roughness of the cartilage surfaces [445]. Pathological knees have less lubricated and rougher cartilage surfaces than normal knees, and hence the variations in FP for pathological knees are expected to be different from those for normal knees.

The parameter FSP is given by the second central moment of $M(t, \omega)$ along each time slice as

$$\text{FSP}(t) = \left[\frac{\sum_{\omega=0}^{\omega_m} [\omega - \text{FP}(t)]^2 M(t, \omega)}{\sum_{\omega=0}^{\omega_m} M(t, \omega)} \right]^{\frac{1}{2}}. \quad (8.141)$$

FSP gives the spread of frequency about the mean frequency for each time instant. The spread of frequency at an instant of time could arise as a result of AM components. AM effects may occur in VAG signals and may be dependent on the quality and intensity of the sound produced due to joint vibration. FSP could serve as a good feature to identify noisy knees.

In order to facilitate the derivation of a global decision on each VAG signal, the mean and SD of the arrays of each of the four parameters were calculated. The mean and SD characterize the central tendency and dispersion of the nonstationary parameters.

The VAG signal of a normal subject is shown in Figure 8.53. The signal demonstrates a normal noisy knee with a click and grinding sound heard during auscultation. The OMPTFD of the signal is shown in part (b) of the same figure. The TFD was constructed using only the coherent structures of the signal and the number of TF atoms used was 588. Optimization of the marginals was achieved in two iterations. The click and grinding sound are shown as high-frequency activities in the TFD and have been represented with good TF localization.

Figure 8.54 shows the VAG signal of a subject with chondromalacia patella. Grinding sound was heard during auscultation. The OMPTFD shown in part (b)

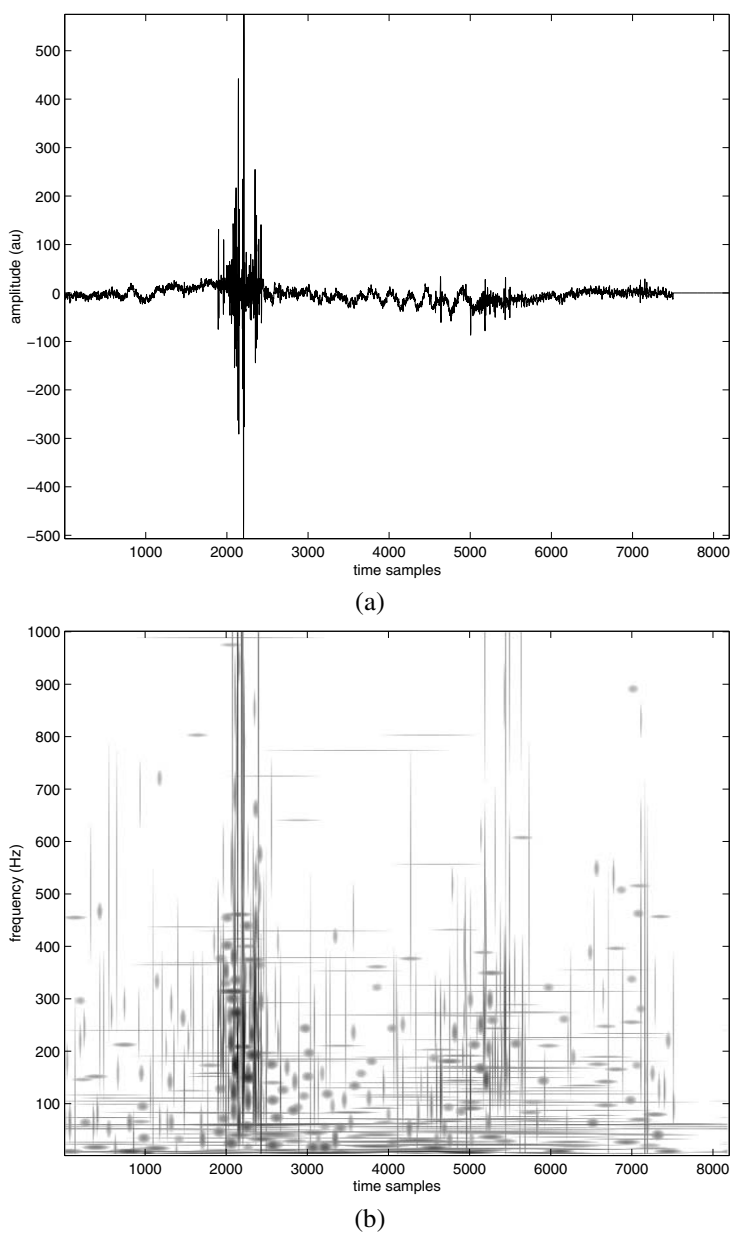


Figure 8.53 (a) VAG signal of a normal subject. A click and grinding sound were heard during auscultation of the knee. au: arbitrary acceleration units. Sampling rate = 2 kHz . (b) OMPTFD of the signal. Reproduced with permission from S. Krishnan, R.M. Rangayyan, G.D. Bell, and C.B. Frank, Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology, *IEEE Transactions on Biomedical Engineering*, 47(6):773–783, June 2000. ©IEEE.

of the same figure was constructed using 803 atoms, and the marginals were optimized in two iterations. The OMPTFD indicates two linear FM components (chirps) in the high-frequency range of 500 Hz to 800 Hz . Several tonal components below 400 Hz can also be identified in the TFD.

Figures 8.55 and 8.56 show the EP, ESP, FP, and FSP waveforms derived from the OMPTFD of the normal VAG signal shown in Figure 8.53. Figures 8.57 and 8.58 show the EP, ESP, FP, and FSP waveforms derived from the OMPTFD of the abnormal VAG signal shown in Figure 8.54. It may be readily seen that the parameters of the abnormal VAG signal have larger variations over time as compared to the corresponding parameters of the normal VAG signal. Whereas visual comparative analysis of the TFDs of the normal and abnormal signals could be difficult, several differences are readily observable in the plots of the related parameters.

The EP and ESP features are dependent on the data acquisition protocol and related parameters, in particular, the gain of the amplifier and the power of the VAG signal. To overcome this problem, the values of the CV of EP and ESP were computed and used with the mean and SD of FP and FSP for pattern classification. Using a database consisting of 71 VAG signals, with 51 normal and 20 abnormal signals restricted to chondromalacia patella, an overall accuracy of 77.5% and area under the ROC curve of 0.75 were achieved. The frequency-related parameters were observed to play a major role in the classification task. See Rangayyan and Krishnan [445] for descriptions of methods to detect different types of FM components in TFDs.

It should be noted that, in the analysis of signals using AR coefficients, dominant poles, and other parameters as described in the preceding sections of the present chapter, each signal is divided into a variable number of quasistationary segments. In TF analysis, each signal in its entirety is represented by a small number of features and a global decision is made on each signal, rather than the segment-by-segment decision made in the other approaches mentioned. The TF method does not require labeling of signal segments as normal or abnormal, and it eliminates the need to estimate the joint angle corresponding to the pathology as observed during arthroscopy. See Krishnan et al. [273] and Kim et al. [471] for further discussion on related topics.

8.15 Remarks

We have now reached the stage where we have extended the application of a number of signal processing, modeling, and analysis techniques to nonstationary and multicomponent biomedical signals. Fixed or adaptive segmentation of the signals into quasistationary segments is one approach to facilitate the analysis of such signals using traditional techniques, and we studied several approaches for segmentation. Adaptive segmentation facilitates not only the identification of distinct and separate events at unknown time instants in the given signal, but also the characterization of events of variable duration using the same number of parameters. We have also studied advanced and recently developed techniques such as wavelets and TFDs that permit the analysis of nonstationary and multicomponent biomedical signals without segmentation, and BSS techniques that can help in separating mixed and overlapping

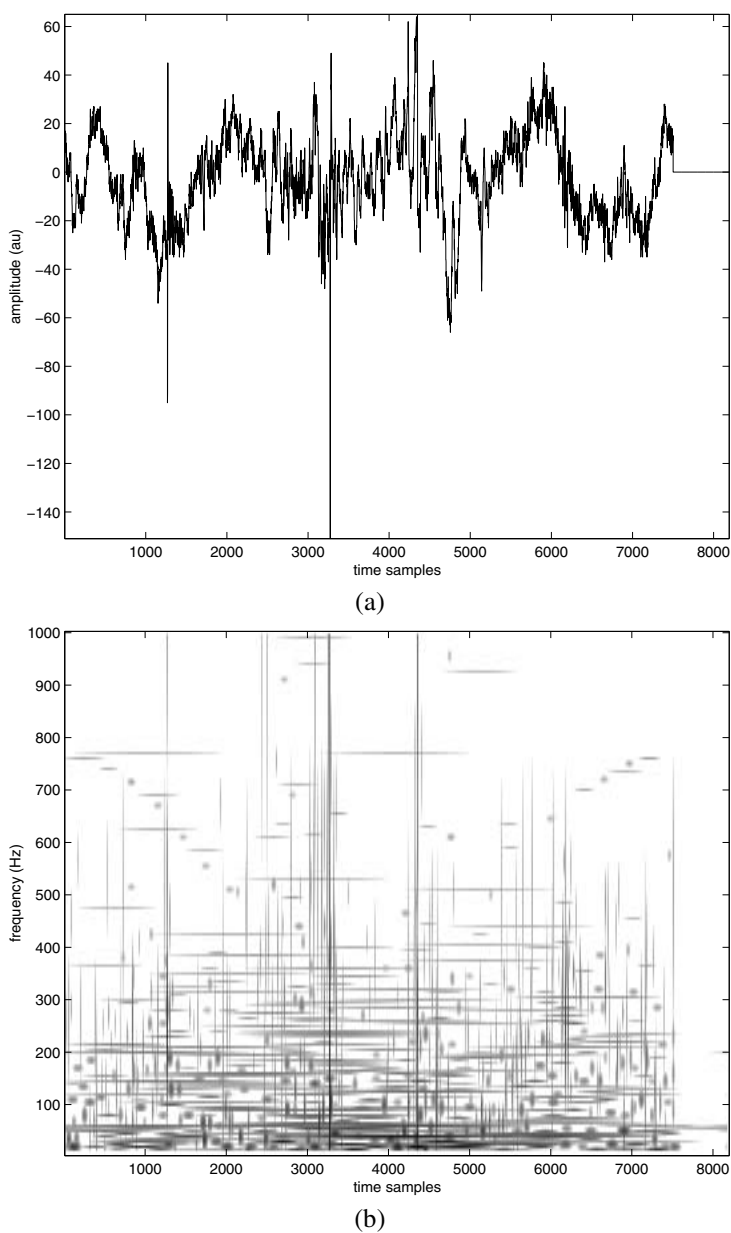
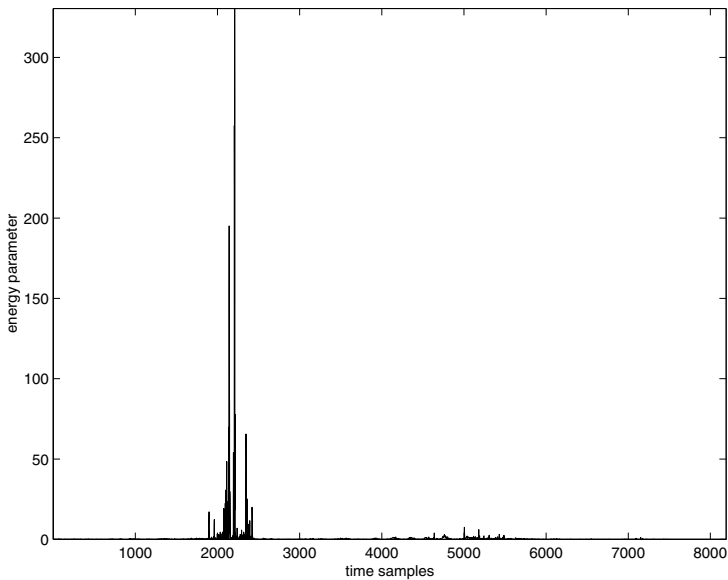
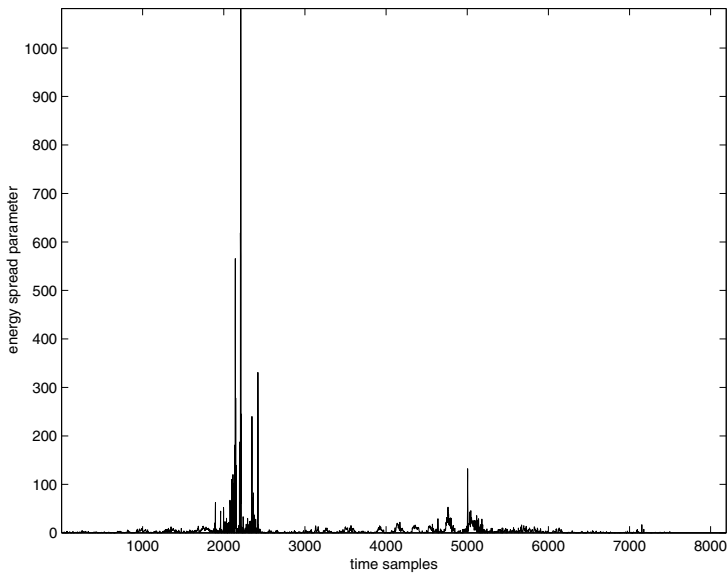


Figure 8.54 (a) VAG signal of a subject with knee-joint pathology. Grinding sound was heard during auscultation of the knee. au: arbitrary acceleration units. Sampling rate = 2 kHz. (b) OMPTFD of the signal. Reproduced with permission from S. Krishnan, R.M. Rangayyan, G.D. Bell, and C.B. Frank, Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology, *IEEE Transactions on Biomedical Engineering*, 47(6):773–783, June 2000. ©IEEE.



(a)



(b)

Figure 8.55 (a) EP and (b) ESP obtained from the OMPTFD of the normal VAG signal in Figure 8.53. Reproduced with permission from S. Krishnan, R.M. Rangayyan, G.D. Bell, and C.B. Frank, Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology, *IEEE Transactions on Biomedical Engineering*, 47(6):773–783, June 2000. ©IEEE.

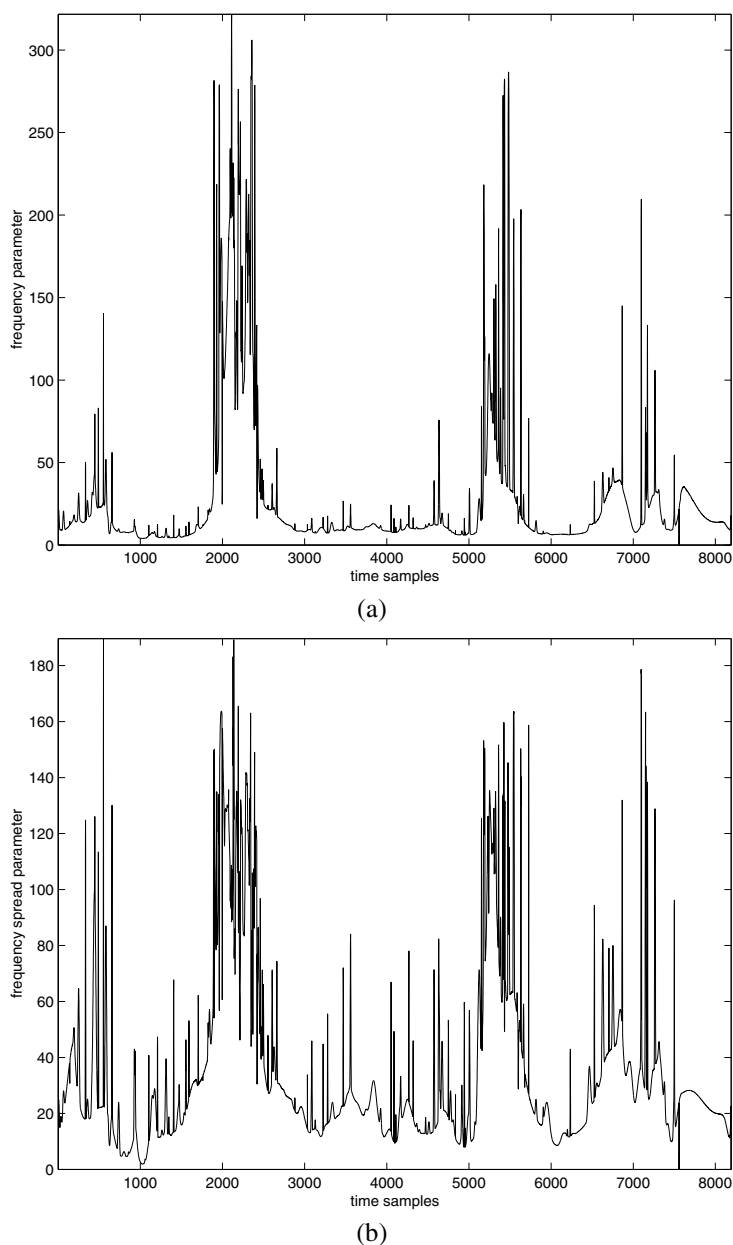


Figure 8.56 (a) FP and (b) FSP obtained from the OMPTFD of the normal VAG signal in Figure 8.53. Reproduced with permission from S. Krishnan, R.M. Rangayyan, G.D. Bell, and C.B. Frank, Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology, *IEEE Transactions on Biomedical Engineering*, 47(6):773–783, June 2000. ©IEEE.

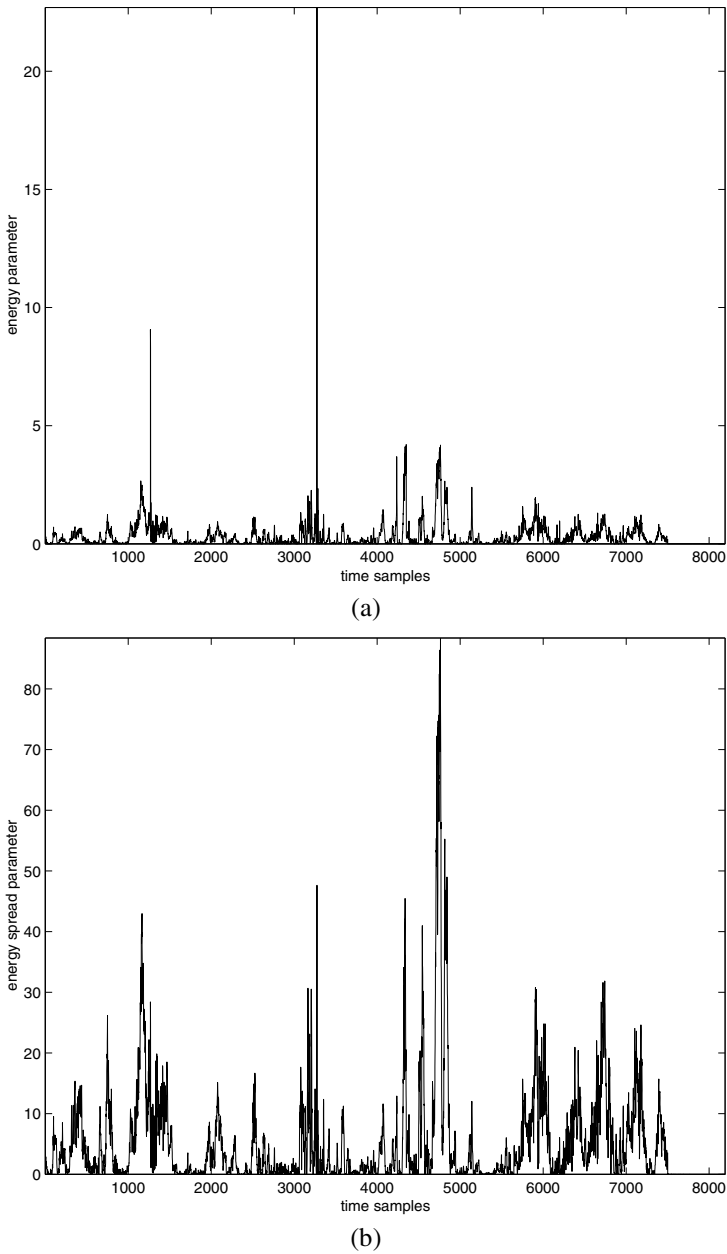


Figure 8.57 (a) EP and (b) ESP obtained from the OMPTFD of the abnormal VAG signal in Figure 8.54. Reproduced with permission from S. Krishnan, R.M. Rangayyan, G.D. Bell, and C.B. Frank, Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology, *IEEE Transactions on Biomedical Engineering*, 47(6):773–783, June 2000. ©IEEE.

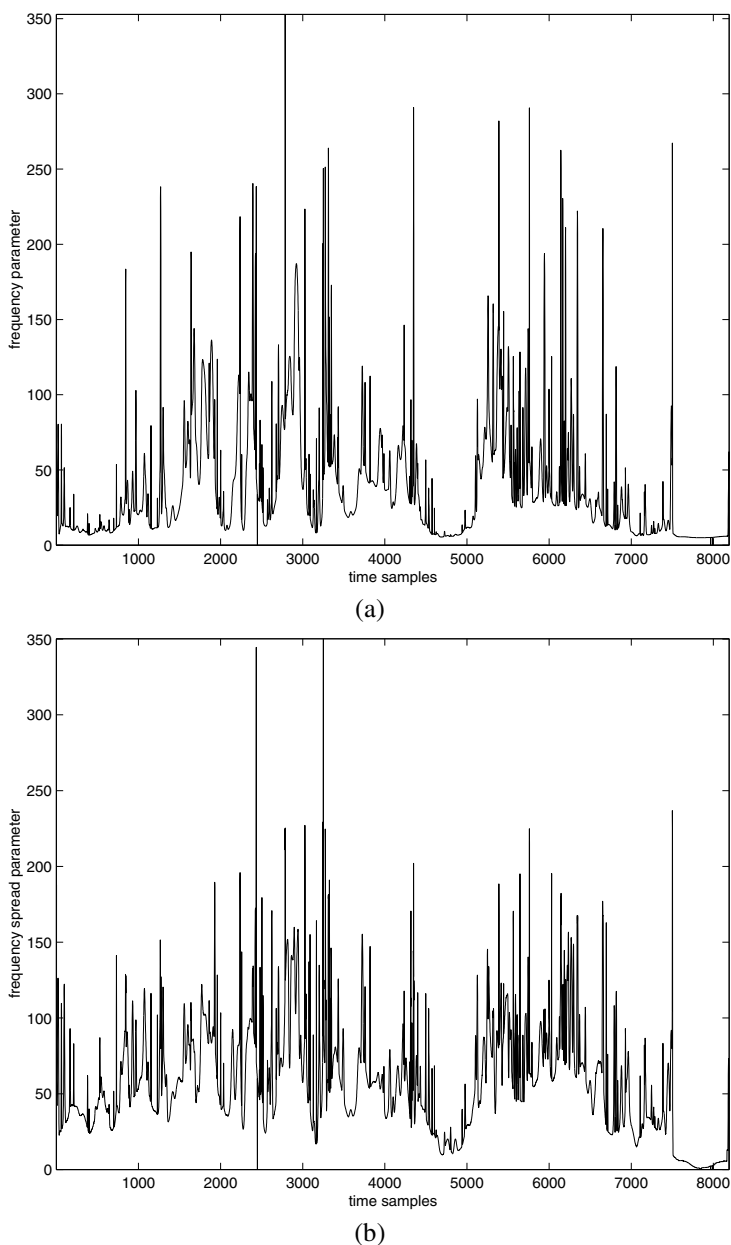


Figure 8.58 (a) FP and (b) FSP obtained from the OMPTFD of the abnormal VAG signal in Figure 8.54. Reproduced with permission from S. Krishnan, R.M. Rangayyan, G.D. Bell, and C.B. Frank, Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology, *IEEE Transactions on Biomedical Engineering*, 47(6):773–783, June 2000. ©IEEE.

components. The results obtained using the techniques studied in the present chapter provide advantages in pattern classification (see Chapter 9) as well as efficient representation and analysis of signals.

8.16 Study Questions and Problems

1. Describe the characteristics of PCG signals that would make them nonstationary. Propose signal processing strategies to break a PCG signal into quasistationary segments.
2. Discuss features of the EEG that make the signal nonstationary. Propose signal processing strategies to detect each type of nonstationarity and to break an EEG signal into quasistationary segments.
3. Investigate features of the EMG that make the signal nonstationary. Propose signal processing strategies to track the time-varying characteristics of the signal. Under what conditions can the signal be partitioned into quasistationary segments? What are the physiological features that you would be able to derive from each segment?
4. Propose an algorithm to perform the segmentation of PCG signals into four parts per cardiac cycle as: (a) the first heart sound; (b) systolic murmur, if present; (c) the second heart sound; and (d) diastolic murmur, if present. If your proposal includes the use of other signals, explain the need and rationale for the use of such signals. Explain the relationship between events in your reference signals and the events of interest in the PCG signal. Provide sketches of typical signals and the results of your methods to illustrate your procedures. Document your procedures using a flowchart or an algorithmic listing. Give at least three nontrivial equations representing important steps in your procedures.

8.17 Laboratory Exercises and Projects

Note: Data files related to the exercises are available at the site

<http://people.ucalgary.ca/~ranga/enel563>

1. The speech signal of the word “safety” is given in the file `safety.wav`. You may use the program `safety.m` to read the data. Explore the use of short-time statistics such as *ZCR* and *RMS* values for segmentation of the signal. Study the effect of the duration of the short-time analysis window on the trends in the parameters computed and on segmentation.
2. The files `pec1.dat`, `pec22.dat`, `pec33.dat`, and `pec52.dat` give the PCG, ECG, and carotid pulse signals of two normal subjects and two patients with systolic murmur. You may use the program `plotpec.m` to read the data. Explore the use of short-time *ZCR*, *RMS*, and AR-model coefficients for segmentation of the signals. Evaluate the segment boundaries obtained in relation to the events in the PCG signals as well as the corresponding events in the ECG and carotid pulse channels.