Signal Processing 91 (2011) 2793–2809

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Resonance-based signal decomposition: A new sparsity-enabled signal analysis method ☆

Ivan W. Selesnick

Polytechnic Institute of New York University, 6 Metrotech Center, Brooklyn, NY 11201, USA

ARTICLE INFO

Article history:
Received 14 January 2010
Received in revised form
17 August 2010
Accepted 29 October 2010
Available online 24 November 2010

Keywords:
Sparse signal representation
Constant-Q transform
Wavelet transform
Morphological component analysis

ABSTRACT

Numerous signals arising from physiological and physical processes, in addition to being non-stationary, are moreover a mixture of sustained oscillations and non-oscillatory transients that are difficult to disentangle by linear methods. Examples of such signals include speech, biomedical, and geophysical signals. Therefore, this paper describes a new nonlinear signal analysis method based on signal *resonance*, rather than on frequency or scale, as provided by the Fourier and wavelet transforms. This method expresses a signal as the sum of a 'high-resonance' and a 'low-resonance' component—a high-resonance component being a signal consisting of multiple simultaneous sustained oscillations; a low-resonance component being a signal consisting of non-oscillatory transients of unspecified shape and duration. The resonance-based signal decomposition algorithm presented in this paper utilizes sparse signal representations, morphological component analysis, and constant-Q (wavelet) transforms with adjustable *Q*-factor.

1. Introduction

Frequency-based analysis and filtering are fundamental tools in signal processing. However, frequency and time–frequency analysis are not productively applied to all signals indiscriminately; they are effective for signals that are substantially oscillatory or periodic in nature. Signals that are piecewise smooth, defined primarily by their transients (singularities), are more fruitfully represented, analyzed and processed in the time domain or wavelet domain; for example, scan-lines of a photographic image, recordings of eye movements, evoked response potentials, neurological spike trains, etc.

However, many complex signals arising from physiological and physical processes are not only non-stationary but also exhibit a mixture of oscillatory and non-oscillatory transient behaviours. For example, speech, biomedical (EEG, phonocardiograms, etc.), and geophysical signals (ocean wave-height data, etc.) all possess both sustained oscillatory behaviour and transients. EEG signals contain rhythmic oscillations (alpha and beta waves, etc.) but they also contain transients due to measurement artifacts and non-rhythmic brain activity. Ocean wave-height data measures the superposition of ocean waves that have travelled many 100's of miles [97], but weather events through which the waves travel induce disruptions to the oscillatory behaviour. Indeed, signals obtained by measuring physiological and geophysical systems often consist of sustained oscillations and transient phenomena that are difficult to disentangle by linear methods.

In order to advance the representation, analysis, and processing of complex non-stationary signals, we describe a new nonlinear signal analysis method based not on frequency or scale, as provided by the Fourier and wavelet transforms, but on *resonance*. This method expresses a signal as the sum of a 'high-resonance' and a 'low-resonance' component. By a high-resonance component, we mean a signal consisting of multiple simultaneous sustained oscillations. In contrast, by a low-resonance component, we mean a signal consisting of non-oscillatory transients of unspecified shape and duration.

Aspects of this work have been presented in two earlier conference papers [84,85].

☆ This work is supported in part by NSF under grant CCF-1018020.
E-mail address: selesi@poly.edu

0165-1684/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2010.10.018

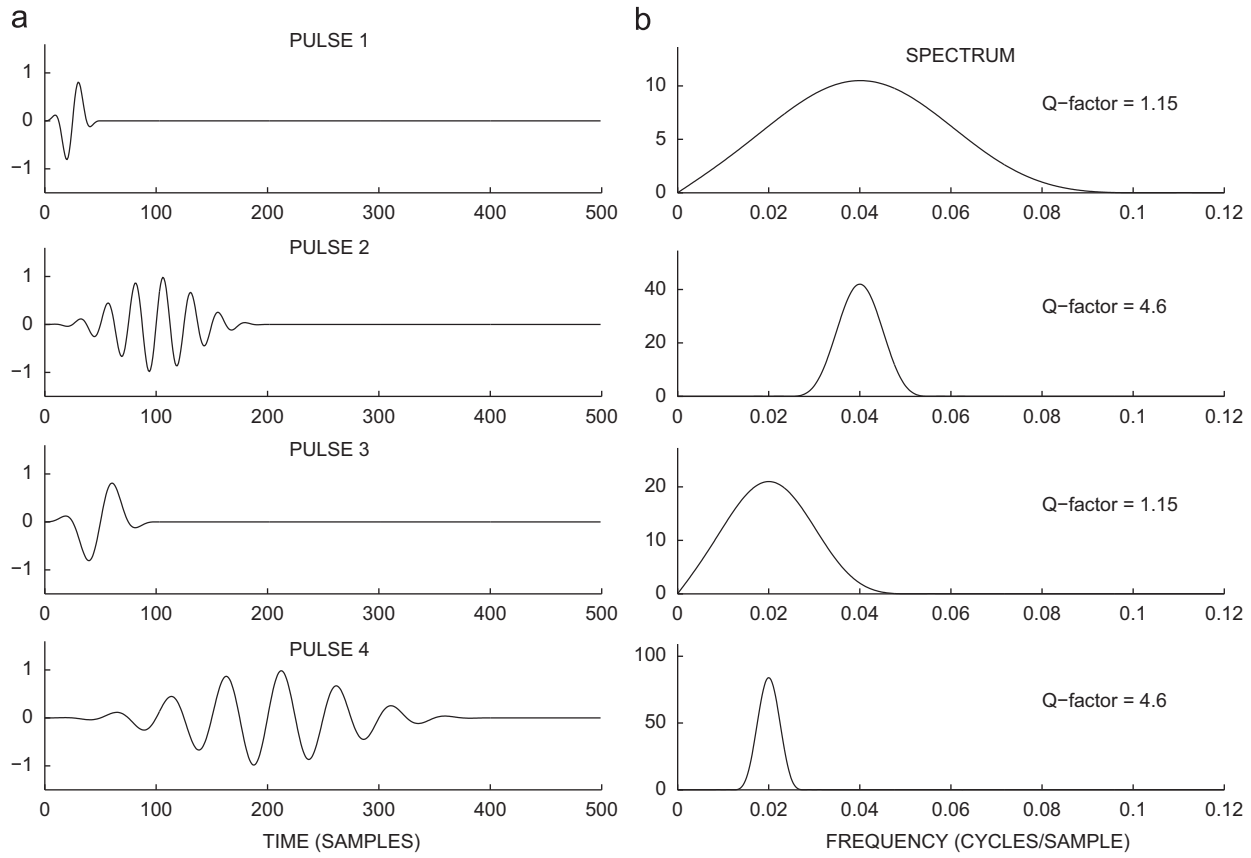

Fig. 1. The resonance of an isolated pulse can be quantified by its Q-factor, defined as the ratio of its center frequency to its bandwidth. Pulses 1 and 3, essentially a single cycle in duration, are low-resonance pulses. Pulses 2 and 4, whose oscillations are more sustained, are high-resonance pulses. A low Q-factor wavelet transform (for example, the classical dyadic wavelet transform) is suitable for the efficient representation of pulses 1 and 3. The efficient representation of pulses 2 and 4 calls for a wavelet transform with higher Q-factor. (a) Signals. (b) Spectra.

2. Signal resonance

Fig. 1 illustrates the concept of signal resonance. Pulses 1 and 3 in Fig. 1 each consist of essentially a single cycle of a sine wave. We classify both pulses as low-resonance signals because they do not exhibit sustained oscillatory behaviour. Note that these pulses are time-scaled versions of one another. Time-scaling a pulse does not effect its degree of resonance. Clearly, a low-resonance pulse may be either a high frequency signal (pulse 1) or a low frequency signal (pulse 3). Low-resonance pulses are not restricted to any single band of frequencies. Therefore, *the low-resonance component of a signal cannot be extracted from the signal by frequency-based filtering.*

We classify pulses 2 and 4 in Fig. 1 as high-resonance signals because they exhibit sustained[1] oscillations. Both pulses consist of about five cycles of a sine wave multiplied by a bell-shaped function (specifically, a Blackman window). As above, both pulses are time-scaled versions of one another, and they have the same degree of resonance. Likewise, a high-resonance pulse may be either a high frequency signal (pulse 2) or a low frequency signal (pulse 4), and therefore, as above, *the high-resonance component of a signal cannot be extracted from the signal by frequency-based filtering.*

2.1. Resonance-based signal decomposition

Resonance-based signal decomposition, as we present it, should be able to (approximately) separate pulses 1 and 2 in Fig. 1, even when they overlap in time. To illustrate the results of the resonance-based signal decomposition algorithm (detailed below) we apply it to the synthetic test signal in Fig. 2. The test signal consists of six pulses of three frequencies and two levels of resonance. The goal is to separate the test signal into a high- and a low-resonance component. The computed high- and low-resonance components, produced using the algorithm, are illustrated in Fig. 2a. The algorithm also produces a residual signal to allow for the presence of a stochastic (noise) component. The test signal is equal to the sum of the three signals: the high- and low-resonance components and the residual signal. (The amplitude of the residual signal can be controlled by parameters in the decomposition algorithm.)

[1] We comment that pulses 2 and 4 in Fig. 1 are high-resonance signals only in comparison with pulses 1 and 3. Depending on the type of signal being analyzed, we may wish to classify all four pulses in Fig. 1 as low-resonance pulses, and to classify as high-resonance only those pulses having many more oscillations than any of these four pulses (say 20 or 50 cycles instead of 5).

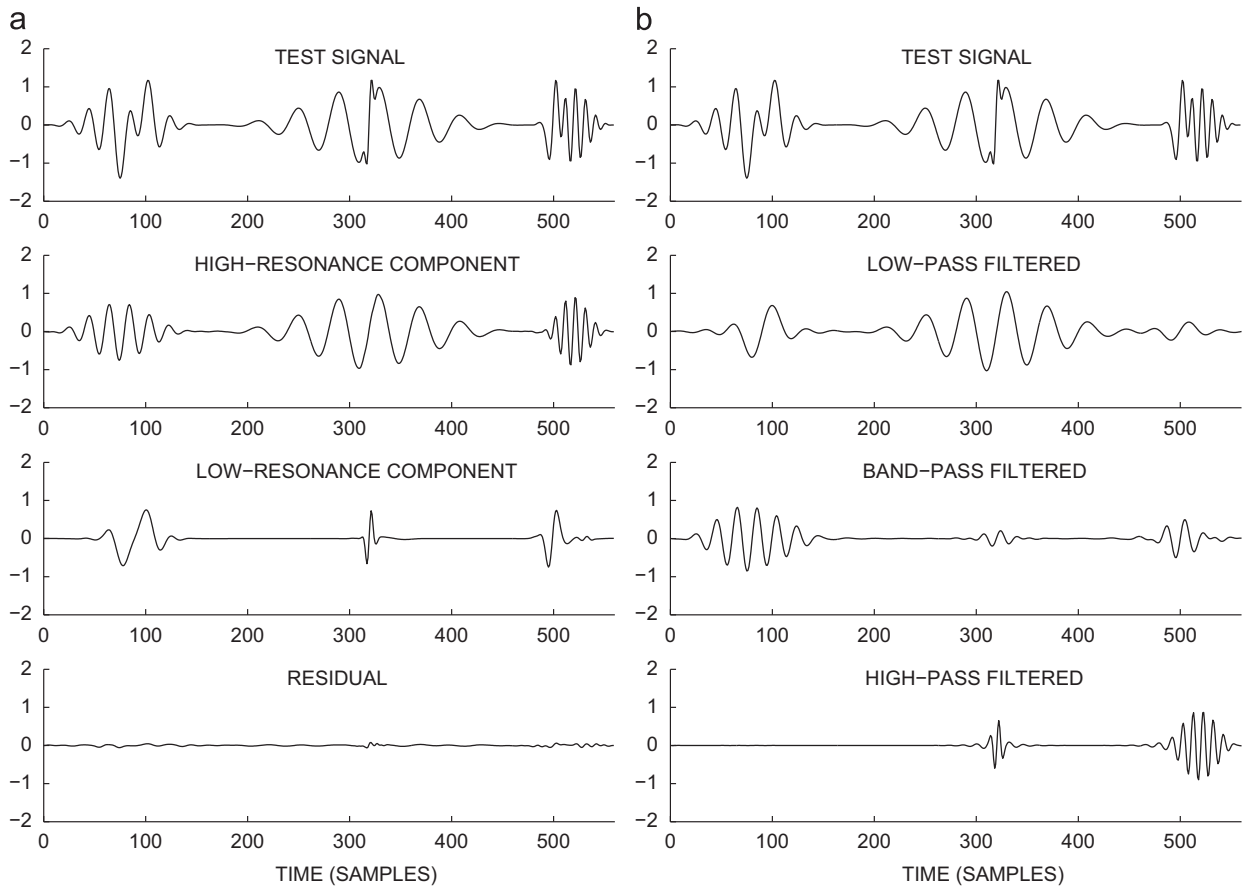

Fig. 2. Resonance- and frequency-based filtering. (a) Decomposition of a test signal into high- and low-resonance components. The high-resonance signal component is sparsely represented using a high Q-factor RADWT. Similarly, the low-resonance signal component is sparsely represented using a low Q-factor RADWT. (b) Decomposition of a test signal into low, mid, and high frequency components using LTI discrete-time filters. (a) Resonance-based decomposition. (b) Frequency-based filtering.

Note that linear time-invariant (LTI) filtering will be unable to yield the separation illustrated in Fig. 2a because the three frequencies present in the high-resonance component are the same three frequencies present in the low-resonance component. The pulses in the high-resonance component differ from those in the low-resonance components not in their frequency, but by the degree to which their oscillations are sustained.

Of course, LTI filtering can separate the test signal into low, mid, and high frequencies. Using low-pass, band-pass, and high-pass LTI filters, we obtain the frequency-based decomposition of the test signal into frequency components, as illustrated in Fig. 2b for comparison.

The frequency-based decomposition of a signal, as illustrated in Fig. 2b, depends partly on the characteristics of the frequency-selective filters utilized: transition-band widths, pass-band and stop-band deviations, phase response, etc. Likewise, the resonance-based decomposition of a signal will also depend on algorithm parameters.

2.2. Resonance-based signal decomposition must be nonlinear

Resonance-based signal decomposition and filtering, as presented here, cannot be achieved by any *linear* filtering scheme, as illustrated in Fig. 3. Each row in Fig. 3 illustrates the (hypothetical and ideal) decomposition of a signal into low- and high-resonance components. The first six signals are low-resonance signals and therefore the low-resonance components are the signals themselves (and the high-resonance components are identically zero). The last signal is a high-resonance signal and therefore the high-resonance component is the signal itself (and the low-resonance component is identically zero).

As illustrated in Fig. 3, neither the low- nor high-resonance component of a signal satisfies the superposition property. The high-resonance signal illustrated in the bottom left panel is exactly the *sum* of the six low-resonance signals illustrated above it. If the resonance-components of a signal were linear functions of the signal, then the low- and high-resonance components in the bottom row of Fig. 3 should be the sum of the components above them. But that is not the case, and therefore the proposed resonance-based signal decomposition is necessarily a nonlinear function of the signal under analysis.

2.3. Can resonance-based signal decomposition be well defined?

Clearly, the separation of a signal into low- and high-resonance components may be ill-defined. If we classify

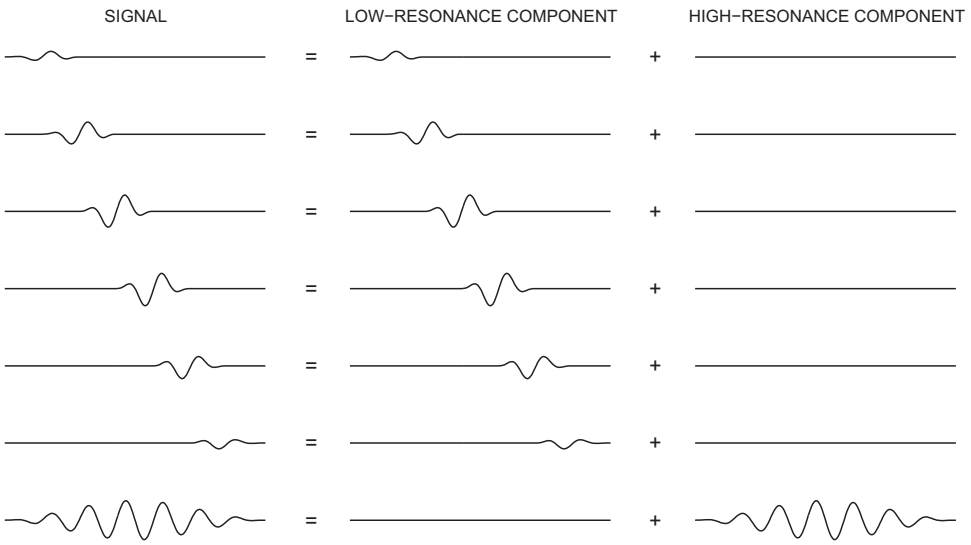

Fig. 3. Resonance-based signal decomposition must be nonlinear: the signal in the bottom left panel is the sum of the signals above it; however, the low-resonance component of the sum is not the sum of the low-resonance components. The same is true for the high-resonance component. Neither the low- nor high-resonance components satisfy the superposition property.

pulses 1 and 3 (consisting of roughly one cycle) in Fig. 1 as low-resonance signals, and pulses 2 and 4 (consisting of roughly five cycles) as high-resonance signals, then how should we classify a pulse consisting of three cycles? Likewise, if a signal consists of several such pulses of indeterminate resonance, then how should its low- and high-resonance components be defined?

It is not initially clear how the resonance of a generic signal should be defined, let alone how a generic signal can be separated into low- and high-resonance components. In contrast, frequency-based filtering is straightforward to define: a low-pass filter preserves (annihilates) sinusoids oscillating with frequencies less than (greater than) the filter's cut-off frequency. The frequency response function together with the linearity of the filter, fully determine the input–output behaviour of the filter.

Consequently, it may appear that the concept of resonance-based signal decomposition is vague, imprecise, and ambiguous. However, such a decomposition can be well defined, albeit indirectly, by formulating it as the solution to an appropriately chosen optimization problem. (The resonance-based decomposition illustrated in Fig. 2a was computed via numerical minimization of the cost function (1) below.) As such, the resonance-components of a signal depend on the specific cost function, and the exact decomposition can be adjusted by varying a set of parameters defining the cost function.

The resonance-based signal decomposition we present is therefore a nonlinear function of the signal, computed numerically by an iterative optimization algorithm. In contrast, frequency-based filtering can be written in closed form using the convolution integral (or sum). Resonance-based decomposition is necessarily nonlinear and numerical, while frequency-based decomposition is linear and analytic.

2.4. Quality-factor and constant-Q bases

While defining the resonance of a generic signal may be problematic, the resonance of an isolated pulse can be quantified by its quality-factor, or Q-factor, defined as the ratio of its center frequency to its bandwidth; this quantity is well known in filter design, control, and the physics of dynamical systems.

The Q-factor of a pulse reflects its degree of resonance as illustrated in Fig. 1. The more oscillatory cycles comprising a pulse, the higher is its Q-factor. The first two pulses illustrated in Fig. 1 oscillate with the same frequency, 0.04 cycles/sample; but the second pulse exhibits oscillations that are more sustained and accordingly it has a higher Q-factor (4-times higher). The second two pulses illustrated in Fig. 1 each oscillate at a frequency of 0.02 cycles/sample and have the same Q-factors, respectively, as the first two pulses. Note that the Q-factor of a pulse, as illustrated in Fig. 1, essentially counts the number of oscillations (cycles) the pulse consists of.

The method described below for computing the high- and low-resonance components of a signal is based on the efficient representation of these two signal components using two suitably designed bases. The efficient representation of the high-resonance signal component calls for a basis ideally comprised entirely high-resonance (high Q-factor) functions; such a basis can be obtained from a single high Q-factor pulse by translating and time-scaling it. The functions in such a basis will all have the same Q-factor. Similarly, for the efficient representation of the low-resonance signal component we should utilize a basis comprised entirely of low-resonance (low Q-factor) functions; which can likewise be obtained from a single low Q-factor pulse through translation and time-scaling. We therefore need two 'constant-Q' bases—one characterized by a high Q-factor, the other characterized by

a low Q-factor. Bases obtained from a single pulse through translation and time-scaling are well known as wavelet bases, the generating pulse being known as the 'wavelet'.

The best known and most widely used constant-Q basis, the dyadic wavelet basis [21], has a very low Q-factor. Indeed, the effectiveness of the (dyadic) discrete wavelet transform stems from its ability to provide relatively sparse representations of piecewise smooth signals, that is, of low-resonance signals. The dyadic wavelet transform is applied much less frequently to oscillatory (high-resonance) signals such as speech and audio because it does not provide particularly efficient representations of these signals.

The need for high Q-factor constant-Q transforms may be questioned; indeed, speech and audio signals are usually analyzed and processed using constant-bandwidth transforms (for example, the MPEG 2/4 AAC codec uses the MDCT switching between 128 and 1024 frequency bands; speech enhancement often utilizes the STFT). Although constant-bandwidth analysis can be implemented with high computational efficiency using the FFT, and although it serves as a key component of numerous audio coders, it does not provide the constant-Q analysis needed for resonance-based signal decomposition.

Constant-Q frequency analysis has been of interest in acoustics and signal processing for many years. This interest is partly inspired by biology—the characteristics of the human and other mammalian auditory systems have been extensively studied; and it has been established that the cochlea possesses a near constant-Q property. Specifically, the cochlea can be modeled as a bank of highly overlapping band-pass filters having constant Q-factors above a species-dependent frequency. (The human cochlea is approximately constant-Q over 500 Hz, and tends towards constant bandwidth below that frequency.) Several parametric models have been proposed for these *auditory filter banks*, including the Gammatone and Gammachirp filter banks [44,50,98] which are designed to be consistent with psychoacoustic measurements.

3. Methods

3.1. Rational-dilation wavelet transform (RADWT)

The pursuit of easily invertible constant-Q discrete-time transforms naturally leads to discrete wavelet transforms (WTs) based on rational dilation factors [4,5,59] and to perfect reconstruction filter banks based on rational sampling factors [9,10,62,106]. However, critically sampled filter banks based on rational sampling factors are substantially constrained and the filter bank design methods used for the dyadic case cannot be used. Due to the difficulty of the design problem, few solutions have been proposed for the rational-dilation case.

Motivated by the need for high Q-factor constant-Q (wavelet) transforms for the sparse representation of high-resonance signals, we recently developed a new rational-dilation wavelet transform [6] that is fully discrete, easily invertible, energy preserving, approximately shift-invariant, and which provides the user the ability to adjust the Q-factor. The new wavelet transform can be used for high Q-factor analysis or the same low Q-factor analysis as the widely used dyadic wavelet transform. While the transform is not critically sampled, it can be implemented with modest redundancy (e.g., 3-times overcomplete, depending on parameters). Furthermore, the inverse filter bank is the mirror image of the analysis filter bank, so that the transform is 'self-inverting' (it implements a 'tight' frame rather than an orthonormal basis), which facilitates its use for sparse signal representation.

The rational-dilation wavelet transform (RADWT) introduced in [6] is based on the filter bank (FB) illustrated in Fig. 4. When the integers p, q, and s in Fig. 4 are chosen so that the FB is overcomplete, we provide in [6] a set of filters for this multirate filter bank achieving the perfect reconstruction property, good time–frequency localization, and high regularity. The Q-factor of the wavelet transform, obtained when the FB is iterated on it low-pass branch, depends on the parameters p, q, and s. Instead of being based on integer dilations, the RADWT is based on a rational dilation (q/p) between 1 and 2. Setting the dilation factor close to 1, and $s>1$, gives a WT with analysis/synthesis functions (wavelets) having a high Q-factor. Setting $s=1$, one obtains a WT with a low Q-factor like the dyadic DWT. The non-uniform frequency decomposition and the associated wavelet are illustrated in Fig. 5 for two cases: a low Q-factor and a high Q-factor transform.

3.2. Sparsity-based signal decomposition

We define high-resonance signals as those signals that are efficiently (sparsely) represented by a *high* Q-factor constant-Q transform (or 'high-Q transform') such as the RADWT with appropriately chosen parameters p, q, and s (as in Fig. 5b). This definition of a high-resonance component is therefore relative to a specified constant-Q transform. Similarly, we define low-resonance signals as those signals efficiently represented by a *low* Q-factor constant-Q transform (or 'low-Q transform') such as the conventional dyadic DWT or the RADWT (as in Fig. 5a). Note that a high-resonance signal will not be efficiently represented with a low-Q transform and likewise a low-resonance signal will not be efficiently represented with a high-Q transform. Therefore, *the efficiency (sparsity) of a signal representation with respect to low-Q and high-Q transforms can be employed as a means by which to achieve resonance-based signal decomposition.*

The nonlinear separation of a signal into components defined by distinct behaviours has been addressed in several publications. For example, Refs. [2,3,101,102] propose algorithms following ideas of Meyer [74] for the decomposition of an image into oscillatory and bounded variation components. A general framework for nonlinear signal decomposition based on sparse representations has

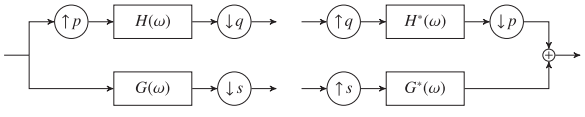

Fig. 4. Analysis and synthesis filter banks for the implementation of the rational-dilation wavelet transform (RADWT). The dilation factor is q/p and the redundancy is $(s\,(1-p/q))^{-1}$ assuming iteration of the filter bank on its low-pass (upper) branch ad infinitum.

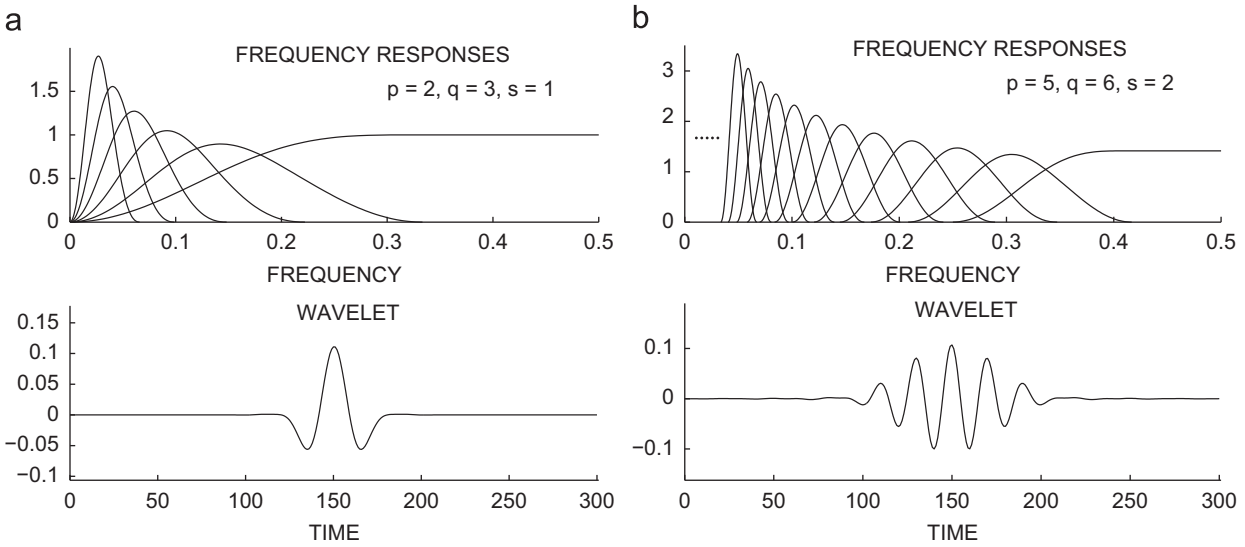

Fig. 5. Rational-dilation wavelet transforms (RADWT): frequency responses and wavelet. (a) Low Q-factor RADWT using $p=2$, $q=3$, $s=1$. The wavelet is approximately the Mexican hat function. (b) High Q-factor RADWT using $p=5$, $q=6$, $s=2$. The dilation factor is 1.2, much closer to 1 than the dyadic wavelet transform. The RADWTs in both (a) and (b) have the same redundancy: they are both 3-times overcomplete.

been described in several papers [33,36,51,94,95]. In order for this approach, called 'morphological component analysis' (MCA), to be successful, the respective transforms must have a low coherence (the analysis/synthesis functions of each transform should have low correlation with the analysis/synthesis functions of the other transform), a condition satisfied by low-Q and high-Q transforms; see (2) below.

Given an observed signal $\mathbf{x}=\mathbf{x}_1+\mathbf{x}_2$, with $\mathbf{x},\mathbf{x}_1,\mathbf{x}_2 \in \mathbb{R}^N$, the goal of MCA is to estimate/determine $\mathbf{x}_1$ and $\mathbf{x}_2$ individually. Assuming that $\mathbf{x}_1$ and $\mathbf{x}_2$ can be sparsely represented in bases (or frames) $\mathbf{S}_1$ and $\mathbf{S}_2$, respectively, they can be estimated by minimizing the objective function,

$$J(\mathbf{w}_1,\mathbf{w}_2) = \|\mathbf{x}-\mathbf{S}_1\mathbf{w}_1-\mathbf{S}_2\mathbf{w}_2\|_2^2+\lambda_1\|\mathbf{w}_1\|_1+\lambda_2\|\mathbf{w}_2\|_1 \qquad (1)$$

with respect to $\mathbf{w}_1$ and $\mathbf{w}_2$. Then MCA provides the estimates $\hat{\mathbf{x}}_1=\mathbf{S}_1\mathbf{w}_1$ and $\hat{\mathbf{x}}_2=\mathbf{S}_2\mathbf{w}_2$.

The effectiveness of MCA for certain image processing problems (image in-painting, interpolation, etc.) has been well demonstrated, especially with the curvelet transform, the 2D DCT, and 2D wavelet transforms in the role of $\mathbf{S}_1$ and $\mathbf{S}_2$ [33,37,35,95]. A variant of this approach is shown to be effective for the separation of ventricular and atrial components in an ECG signal in [27], where the respective representations are adapted to ventricular and atrial activity, respectively.

For resonance-based signal decomposition, we propose to use low-Q and high-Q RADWTs in the role of $\mathbf{S}_1$ and $\mathbf{S}_2$. Then $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$ obtained by minimizing (1), will serve as the extracted low- and high-resonance signal components. For example, the resonance-based decomposition illustrated in Fig. 2 was obtained by minimizing (1) with $\lambda_1=\lambda_2=0.2$ where $\mathbf{S}_1$ and $\mathbf{S}_2$ are the two RADWTs illustrated in Fig. 5.

More general forms of MCA allow the sparsity measures for $\mathbf{x}_1$ and $\mathbf{x}_2$ in (1) to be different from each other. Furthermore, prior information can be utilized in the objective function to further improve the achievable component separability [28]. Additionally, the data fidelity term need not be an ℓ_2 norm, and other sparsity priors besides the ℓ_1-norm can be used, etc. We use the ℓ_1-norm here because it promotes sparsity while being convex function.

3.2.1. *Convexity and the* ℓ_1-norm

Casting resonance-based signal decomposition as a *convex* optimization problem as in (1) facilitates the computation of the resonance components. Here, we use the ℓ_1-norm in (1) because it makes the objective function convex. Although an ℓ_p-norm with $0 \le p < 1$ in (1) promotes sparsity in the solution more aggressively than the ℓ_1-norm, the objective function J will not be convex and the solution will therefore be more difficult to obtain—we can generally find a solution that is only locally optimal, and it will depend on the particular optimization algorithm utilized, and on the way it is initialized.

3.2.2. *Coherence and the RADWT*

In order for morphological component analysis to be successful at decomposing a signal $\mathbf{x}$ into components $\mathbf{x}_1$ and $\mathbf{x}_2$, it is important that the two utilized transforms, $\mathbf{S}_1$ and $\mathbf{S}_2$, have a low mutual coherence. That is, the synthesis functions (columns) of transform $\mathbf{S}_1$ should have minimal correlation with the synthesis functions (columns) of transform $\mathbf{S}_2$. Although some pairwise correlations can be zero (some columns of $\mathbf{S}_1$ can be orthogonal to some columns of $\mathbf{S}_2$), it is impossible that they all have zero correlation.

When MCA is performed using two wavelet transforms, characterized by wavelets $\psi_1(t)$ and $\psi_2(t)$, respectively, it is therefore necessary that the translations and dilations of $\psi_1(t)$ and $\psi_2(t)$ have a small inner product for all dilations and translations. Denote the maximum inner product as $\rho_{\max}(Q_1,Q_2)$ where Q_i is the Q-factor of wavelet $\psi_i(t)$. We assume that $Q_2 > Q_1$ in the following. To evaluate these inner products, consider a simplified case where the wavelets $\psi_i(t)$ are ideal band-pass functions with Fourier

transforms given by

$$\Psi_i(f) \coloneqq \begin{cases} \sqrt{Q_i/f_i}, & f_i - f_i/(2Q_i) < f < f_i + f_i/(2Q_i) \\ 0 & \text{otherwise} \end{cases}$$

as illustrated in Fig. 6. The band-pass functions (single-sided in Fig. 6 for convenience) are normalized to have unit energy, $\int |\Psi_i(f)|^2 df = 1$. In this case the inner products can be defined in the frequency domain,

$$\rho(f_1, f_2) \coloneqq \int \Psi_1(f)\Psi_2(f)\, df,$$

and the maximum inner product can be written as

$$\rho_{\max}(Q_1, Q_2) \coloneqq \max_{f_1, f_2} \rho(f_1, f_2).$$

The inner product, $\rho(f_1, f_2)$, is given explicitly in the equation on the top of the next page. The maximum value of $\rho(f_1, f_2)$ occurs when $f_2 = f_1(2+1/Q_1)/(2+1/Q_2)$ and is given by

$$\rho_{\max}(Q_1, Q_2) = \sqrt{\frac{Q_1 + 1/2}{Q_2 + 1/2}}, \quad Q_2 > Q_1. \tag{2}$$

Eq. (2) shows how the maximum inner product depends on the Q-factors of the two wavelet transforms. For MCA to be successful, $\rho_{\max}$ should be substantially less than 1. If Q_2 is only slightly greater than Q_1 then the maximum inner product between the two wavelet transforms is near 1 and the result of MCA may be poor (both components $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$ may be similar to $\mathbf{x}$). On the other hand, if $Q_1 = 1$ and $Q_2 = 5.5$, then $\rho_{\max} = 0.5$. Further increasing Q_2, further decreases $\rho_{\max}$. Therefore, in order to ensure the reliability and accuracy of resonance-based signal decomposition using MCA with low-Q and high-Q RADWTs, it is advantageous that the two RADWTs be chosen so as to minimize their coherence; that is, the high-Q RADWT should be designed so that its Q-factor is sufficiently greater than the Q-factor of the low-Q RADWT. However, if this Q-factor is too high, then it may not be well matched to the oscillatory behaviour expected in the high-resonance component, and accordingly the high-Q RADWT may not provide an efficient representation, thereby degrading the results of MCA. The two Q-factors should be chosen so as to (i) roughly reflect the expected behaviour of the two components, yet on the other hand, so as to (ii) minimize $\rho_{\max}$. The two Q-factors should therefore depend to some degree on the signal under analysis:

$$\rho(f_1, f_2) = \begin{cases} 0, & f_2 \le f_1\left(\dfrac{1-1/(2Q_1)}{1+1/(2Q_2)}\right), \\ \sqrt{\dfrac{Q_1\, Q_1}{f_1\, f_2}}\left(f_2\left(1+\dfrac{1}{2Q_2}\right) - f_1\left(1-\dfrac{1}{2Q_2}\right)\right), & f_1\left(\dfrac{2-1/Q_1}{2+1/Q_2}\right) \le f_2 \le f_1\left(\dfrac{2-1/Q_1}{2-1/Q_2}\right), \\ \sqrt{\dfrac{f_2 Q_1}{f_1\, Q_2}}, & f_1\left(\dfrac{2-1/Q_1}{2-1/Q_2}\right) \le f_2 \le f_1\left(\dfrac{2+1/Q_1}{2+1/Q_2}\right), \\ \sqrt{\dfrac{Q_1\, Q_1}{f_1\, f_2}}\left(f_1\left(1+\dfrac{1}{2Q_2}\right) - f_2\left(1-\dfrac{1}{2Q_2}\right)\right), & f_1\left(\dfrac{2+1/Q_1}{2+1/Q_2}\right) \le f_2 \le f_1\left(\dfrac{2+1/Q_1}{2-1/Q_2}\right), \\ 0, & f_1\left(\dfrac{2+1/Q_1}{2-1/Q_2}\right) \le f_2 \end{cases}$$

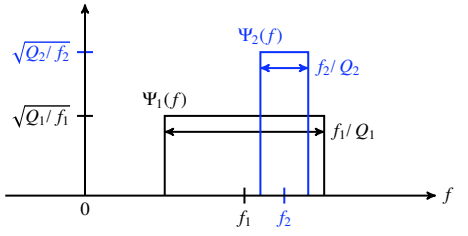

Fig. 6. For reliable resonance-based decomposition, the inner product between the low-Q and high-Q wavelets should be small for all dilations and translations. The computation of the maximum inner product is simplified by assuming the wavelets are ideal band-pass functions and expressing the inner product in the frequency domain.

For the RADWT described in [6], we do not have a formula for $\rho_{\max}$. However, it can be computed numerically. Table 1 reports $\rho_{\max}$ for several specific cases. As reflected in the table, increasing the higher Q-factor causes a decrease in $\rho_{\max}$.

3.3. Split augmented Lagrangian shrinkage algorithm (SALSA)

The proposed framework for resonance-based signal decomposition requires the minimization of the objective function (1). Although this function is convex, its minimization can be difficult due to (i) the non-differentiability of the ℓ_1-norm and (ii) the large number of variables (if each of the two transforms are 3-times overcomplete, then the

Table 1
The coherence, $\rho_{\max}$, between the low-Q RADWT with parameters $p_1 = 2$, $q_1 = 3$, $s_1 = 1$, (Q-factor = 1) and the high-Q RADWT with parameters p_2, q_2, s_2, for several higher-Q RADWTs (Q-factor given in table).

p_2	q_2	s_2	Q_2	$\rho_{\max}$
5	6	2	3	0.723
6	7	3	5	0.582
7	8	3	5	0.573
8	9	3	5	0.572
9	10	3	5	0.581
11	12	4	7	0.484

number of unknowns is 6-times the length of the signal $\mathbf{x}$). Due to the important role of sparsity-promoting objective functions such as (1) in the formulation of recent signal processing methods (including 'compressive sensing' [29]), several algorithms have recently been proposed to minimize this type of objective function. An important early algorithm is the iterated soft-thresholding algorithm (ISTA) developed in [22,40] (this algorithm appeared earlier in the optimization literature, as noted in [18]). However, ISTA converges slowly for some problems and faster algorithms have since been developed, for example [7,8,17,32,39,42,105]. A survey of the literature related to the minimization of convex functions arising in signal processing, including minimization problems with constraints, is given in [18]. In particular, FISTA [7] has a quadratic rate of convergence instead of the linear rate of ISTA; yet it is almost as simple an algorithm as ISTA. We have found SALSA particularly effective for resonance-based decomposition (and likely for MCA in general), as it solves a sequence of ℓ_2-norm regularized problems which for MCA can be solved easily (provided the two transforms are tight frames, as they are here.)

The split augmented Lagrangian shrinkage algorithm (SALSA), developed in [41,1], is based on casting the minimization problem

$$\min_{\mathbf{w}} \quad f_1(\mathbf{w})+f_2(\mathbf{w}) \tag{3}$$

as

$$\min_{\mathbf{u},\mathbf{w}} \quad f_1(\mathbf{u})+f_2(\mathbf{w}) \tag{4}$$

such that $\mathbf{u}=\mathbf{w}$

which is minimized by the alternating split augmented Lagrangian algorithm:

$$\mathbf{u}^{(k+1)} = \operatorname{argmin}_{\mathbf{u}} f_1(\mathbf{u})+\mu\,\|\mathbf{u}-\mathbf{w}^{(k)}-\mathbf{d}^{(k)}\|_2^2, \tag{5}$$

$$\mathbf{w}^{(k+1)} = \operatorname{argmin}_{\mathbf{w}} f_2(\mathbf{w})+\mu\,\|\mathbf{u}^{(k+1)}-\mathbf{w}-\mathbf{d}^{(k)}\|_2^2, \tag{6}$$

$$\mathbf{d}^{(k+1)} = \mathbf{d}^{(k)}-\mathbf{u}^{(k+1)}+\mathbf{w}^{(k+1)}, \tag{7}$$

where k is the iteration index and μ is a user-specified scalar parameter. Each iteration calls for the solution of an ℓ_2-regularized inverse problem, which is often itself a challenge for large-scale problems. However, for resonance-based signal decomposition as we formulate it, the relevant ℓ_2 problem can be solved easily, as described in the following.

In order to specialize SALSA to the MCA problem (1), define

$$f_1(\mathbf{u}) = \|\mathbf{x}-\mathbf{H}\mathbf{u}\|_2^2, \quad f_2(\mathbf{w}) = \lambda_1\|\mathbf{w}_1\|_1+\lambda_2\|\mathbf{w}_2\|_1,$$

and

$$\mathbf{H}=[\mathbf{S}_1\ \ \mathbf{S}_2], \quad \mathbf{u}=\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix}, \quad \mathbf{w}=\begin{bmatrix}\mathbf{w}_1\\ \mathbf{w}_2\end{bmatrix}.$$

Then (5)–(7) gives the iterative algorithm:

$$\mathbf{u}^{(k+1)} = \operatorname{argmin}_{\mathbf{u}} \|\mathbf{x}-\mathbf{H}\mathbf{u}\|_2^2+\mu\|\mathbf{u}-\mathbf{w}^{(k)}-\mathbf{d}^{(k)}\|_2^2, \tag{8}$$

$$\mathbf{w}^{(k+1)} = \operatorname{argmin}_{\mathbf{w}} \lambda_1\|\mathbf{w}_1\|_1+\lambda_2\|\mathbf{w}_2\|_1+\mu\|\mathbf{u}^{(k+1)}-\mathbf{w}-\mathbf{d}^{(k)}\|_2^2, \tag{9}$$

$$\mathbf{d}^{(k+1)} = \mathbf{d}^{(k)}-\mathbf{u}^{(k+1)}+\mathbf{w}^{(k+1)}, \tag{10}$$

where k is the index of iteration. The parameter μ needs to be selected by the user; see [41] for details. We have used $\mu=0.5\lambda$ in the MCA experiments below.

Note that (8) is an ℓ_2 problem and therefore the minimization in (8) can be expressed straightforwardly:

$$\mathbf{u}^{(k+1)} = (\mathbf{H}^t\mathbf{H}+\mu\mathbf{I})^{-1}(\mathbf{H}^t\mathbf{x}+\mu(\mathbf{w}^{(k)}+\mathbf{d}^{(k)})).$$

Using $\mathbf{S}_1\mathbf{S}_1^t=\mathbf{I}$ and $\mathbf{S}_2\mathbf{S}_2^t=\mathbf{I}$ (because the RADWT is a tight frame) and the matrix inverse lemma, we can write

$$(\mathbf{H}^t\mathbf{H}+\mu\mathbf{I})^{-1} = \frac{1}{\mu}\,\mathbf{I}-\frac{1}{\mu(\mu+2)}\begin{bmatrix}\mathbf{S}_1^t\\ \mathbf{S}_2^t\end{bmatrix}[\mathbf{S}_1\ \ \mathbf{S}_2].$$

Also, note that (9) is an ℓ_1-norm regularized denoising problem and therefore the minimization in (9) is given by soft-thresholding. Therefore, SALSA for the MCA problem (1) is

$$\mathbf{b}_i^{(k)} = \mathbf{S}_i^t\mathbf{x}+\mu\,(\mathbf{w}_i^{(k)}+\mathbf{d}_i^{(k)}), \quad i=1,2, \tag{11}$$

$$\mathbf{c}^{(k)} = \mathbf{S}_1\mathbf{b}_1^{(k)}+\mathbf{S}_2\mathbf{b}_2^{(k)}, \tag{12}$$

$$\mathbf{u}_i^{(k+1)} = \frac{1}{\mu}\mathbf{b}_i^{(k)}-\frac{1}{\mu(\mu+2)}\,\mathbf{S}_i^t\mathbf{c}^{(k)}, \quad i=1,2, \tag{13}$$

$$\mathbf{w}_i^{(k+1)} = \mathit{soft}\left(\mathbf{u}_i^{(k+1)}-\mathbf{d}_i^{(k)}, \frac{\lambda_i}{2\mu}\right), \quad i=1,2, \tag{14}$$

$$\mathbf{d}_i^{(k+1)} = \mathbf{d}_i^{(k)}-\mathbf{u}_i^{(k+1)}+\mathbf{w}_i^{(k+1)}, \quad i=1,2, \tag{15}$$

where $\mathit{soft}(x,T)$ is the soft-threshold rule with threshold T, $\mathit{soft}(x,T)=x\max(0,1-T/|x|)$.

To illustrate the convergence of ISTA and SALSA, we apply 100 iterations of each algorithm to minimize the objective function $J(\mathbf{w}_1,\mathbf{w}_2)$ in (1) where $\mathbf{x}$ is the test signal illustrated in Fig. 2a, and where the two transforms are the RADWTs illustrated in Figs. 5a and b. The decay of the objective function (1) is illustrated in Fig. 7 for both ISTA

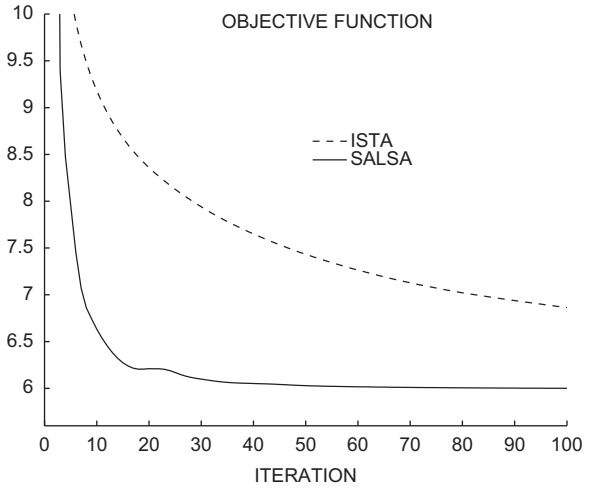

Fig. 7. Reduction of objective function during the first 100 iterations. SALSA converges faster than ISTA.

and SALSA. The signal decomposition obtained with 100 iterations of SALSA is illustrated in Fig. 2a. The signal decomposition obtained with of 100 iterations of ISTA (not-shown) is inferior because ISTA requires many more iterations to converge than SALSA.

4. Example: resonance-selective nonlinear band-pass filtering

In the study of multi-resonance component signals, the analysis of the frequency (oscillatory) content of a signal is sometimes of primary interest; for example, the extraction of alpha rhythms from EEG signals, sinusoidal modeling of speech signals, and spectral analysis of ocean wave-height data. We note that the extraction of alpha rhythms from EEG signals is generally and most simply performed by conventional LTI filtering using a band-pass filter designed to pass 8–12 Hz. However, transients in the EEG signal, which are not considered part of the alpha rhythm, can manifest themselves in the filtered signal as oscillations in the 8–12 Hz band. Therefore, even if no alpha rhythm is present in the EEG signal of interest, the filtered signal may exhibit alpha oscillations. This behaviour has become a point of discussion in the evaluation of methods for investigating the neural mechanism for the generation of certain evoked response potentials (ERPs) in EEG signals [82,108,109].

To illustrate the applicability of resonance-based signal decomposition, we demonstrate that it offers a potential alleviation of this phenomenon, namely the appearance of oscillations at frequency f in a band-pass filtered signal when there are no sustained oscillations at this frequency in the signal being filtered, as illustrated in Figs. 8 and 9. Fig. 8a illustrates a discrete-time test signal consisting of a sinusoidal pulse oscillating with frequency 0.1 cycles/sample and a biphasic transient pulse. The two band-pass filters, illustrated in Fig. 8b, are tuned to 0.07 and 0.1 cycles/sample, respectively. The test signal is filtered with each of the two filters to obtain two output signals, illustrated in Fig. 8c and d. Note that the output signal produced by 'Filter 1' exhibits oscillations at a frequency of 0.07 cycles/sample even though the test signal contains no sustained oscillations at that frequency. Of course, this phenomenon is a basic fact of LTI filtering, yet it can nevertheless impede the interpretation of band-pass filtered signals as noted in [108] and may lead one to the conclusion that the test signal contains more (and stronger) oscillations at this frequency than it actually does.

The band-pass filters in Fig. 8 can be understood to perform frequency analysis of the signal as a whole, whereas we wish to apply frequency analysis only to the 'part' of the signal on which it is appropriate to apply frequency analysis—namely, the part of the signal consisting of sustained oscillations. Resonance-based signal decomposition offers an opportunity to achieve such *resonance-selective* frequency-based filtering. Specifically, we can apply resonance-based decomposition to the test signal and subsequently filter the high-resonance (oscillatory) component with conventional LTI band-pass filters. Applying resonance-based decomposition to the test signal in Fig. 8a yields the high- and low-resonance components illustrated in Fig. 9a and b. Filtering the high-resonance component in Fig. 9a with each of the two band-pass filters in Fig. 8b produces the two output signals illustrated in Fig. 9c and d. Observe that the oscillations in the output of 'Filter 1' are substantially attenuated as compared with that of Fig. 8c. This output signal, being near zero, reflects the fact that the test signal does not contain sustained

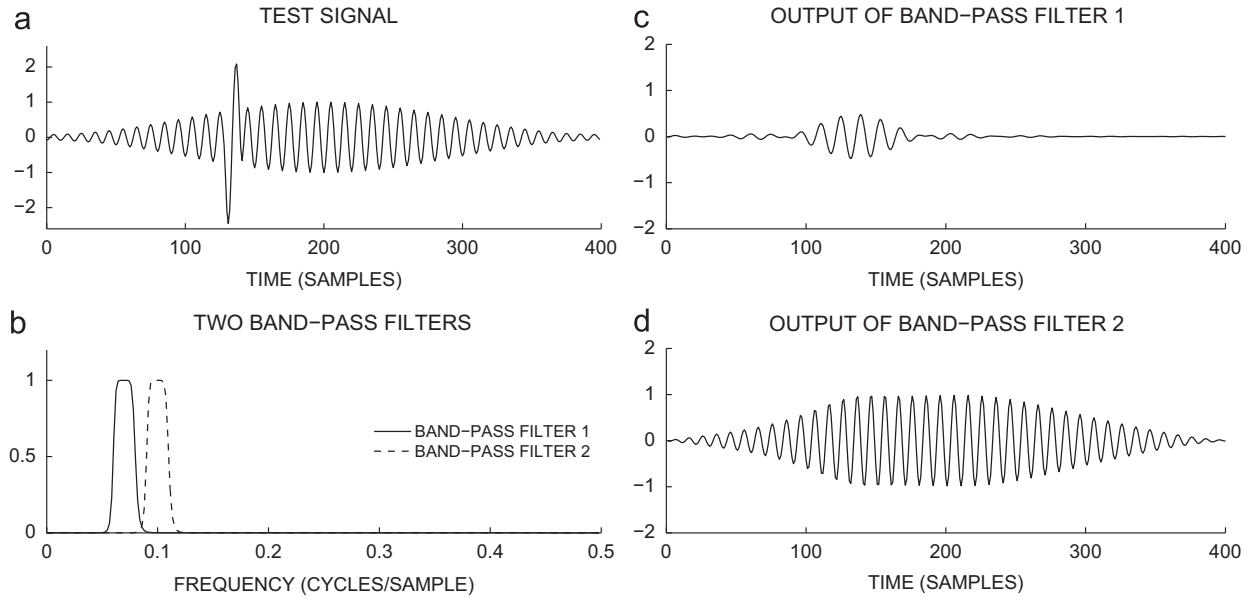

Fig. 8. LTI band-pass filtering. The test signal (a) consists of a sinusoidal pulse of frequency 0.1 cycles/sample and a transient. Band-pass filters 1 and 2 in (b) are tuned to the frequencies 0.07 and 0.10 cycles/second, respectively. The output signals, obtained by filtering the test signal with each of the two band-pass filters, are shown in (c) and (d). The output of band-pass filter 1, illustrated in (c), contains oscillations due to the transient in the test signal. Moreover, the transient oscillations in (c) have a frequency of 0.07 Hz even though the test signal (a) contains no sustained oscillatory behaviour at this frequency.

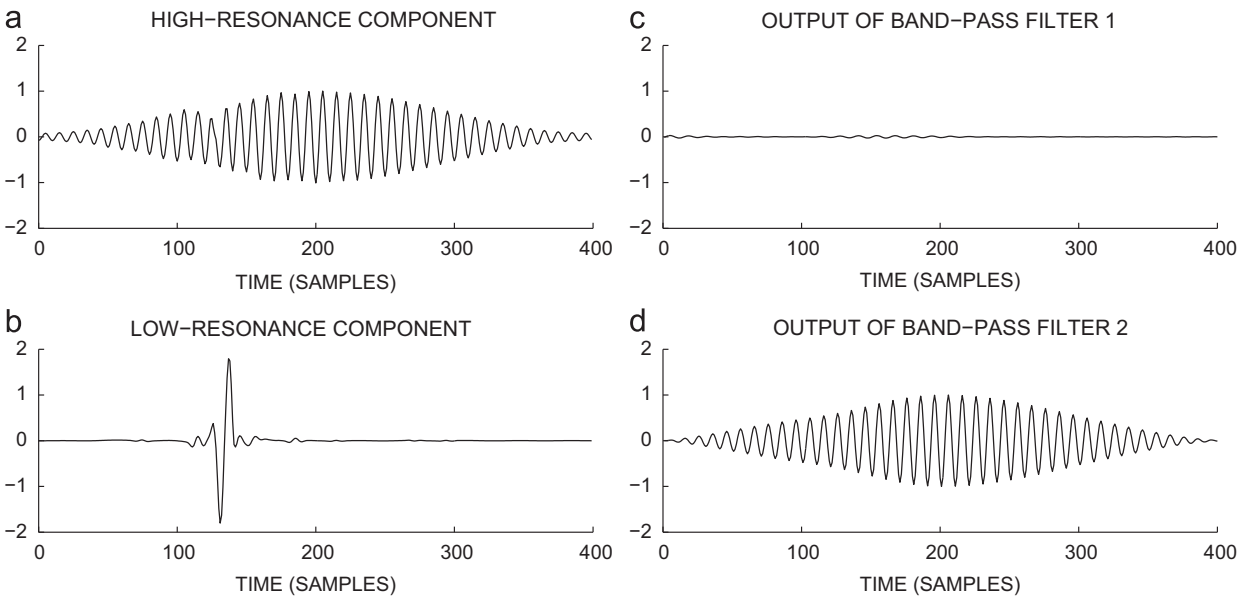

Fig. 9. Resonance-based decomposition and band-pass filtering. When resonance-based analysis method is applied to the test signal in Fig. 8a, it yields the high- and low-resonance components illustrated in (a) and (b). The output signals, obtained by filtering the high-resonance component (a) with each of the two band-pass filters shown in Fig. 8b, are illustrated in (c) and (d). The transient oscillations in (c) are substantially reduced compared to Fig. 8c.

oscillations at the frequency 0.07 cycles/sample. Similarly, the output of 'Filter 2' in Fig. 9d maintains the shape of the sinusoidal pulse more accurately as compared with the output illustrated in Fig. 8d.

This example illustrates the potential of resonance-based decomposition method to overcome the limitations of frequency-selective linear filters. By separating the signal into high- and low-resonance components (which requires nonlinear processing) and subsequently performing conventional LTI frequency-selective filters, we can utilize conventional band-pass filters, while reducing the ringing artifacts due to transients in the signal of interest. In this way, we can achieve nonlinear band-pass filtering that is robust (insensitive) to transients. Furthermore, the method uses a very generic model to separate transients and oscillatory behaviour—no template is required for the shape of the transient; the separation is based only on sparsity in low-Q and high-Q transforms.

Note that the resonance-based decomposition in Fig. 9 is not perfect; largely because the transient pulse is not itself in the low-Q basis. The resonance decomposition in Fig. 9 was obtained using the low-Q and high-Q RADWTs illustrated in Fig. 5. From Table 1, $\rho_{\max} = 0.72$. The result can be improved by using a higher Q-factor for the high-Q RADWT or a more aggressively sparsity promoting (non-convex) regularizer in (1).

5. Example: resonance-based decomposition of a speech signal

To further illustrate how resonance-based signal decomposition can aid the analysis of non-stationary signals, consider the speech signal illustrated in Fig. 10. Fig. 10 illustrates a 150 ms segment of a speech signal ("I'm" spoken by an adult male) in which a vowel-consonant transition is visible. The high- and low-resonance components obtained by minimizing (1) are illustrated in Fig. 10b and c. The high-resonance component captures the sustained oscillations in the speech signal while the low-resonance component captures a sequence of isolated impulses corresponding to the glottal pulses produced by the vibration of the vocal folds during voiced speech. Therefore, although the original speech signal in Fig. 10a is largely oscillatory, it is not a purely high-resonance signal in the sense of our definition: its resonance-decomposition yields a non-negligible low-resonance component. The decomposition was obtained using the high-Q RADWT with parameters: $p=8$, $q=9$, $s=3$, with 38 levels; and using the low-Q RADWT with parameters: $p=2$, $q=3$, $s=1$, with 12 levels (as illustrated in Fig. 5a). From Table 1, $\rho_{\max} = 0.57$.

Note that neither the low- nor high-resonance components shown in Fig. 10 are concentrated in a specific frequency band. Indeed, the high-resonance component consists of low- and high-frequency oscillations; and the low-resonance component consists of a set of impulses and therefore has a broad frequency-spectrum. The frequency spectra of the original speech signal and the resonance components, computed using the middle 50 ms (50–100 ms) of the speech waveform, illustrated in Fig. 11, show that the energy of each resonance component is widely distributed in frequency and that their frequency-spectra overlap.

The low-resonance component illustrated in Fig. 10 resembles the excitation signal in source-filter model based LPC or the cepstrum, etc. [80]. However, resonance-based decomposition uses no such source-filter model; requires no estimation of the pitch period; nor requires that the pitch period be approximately constant over several pitch periods. The decomposition does not depend on any speech model, implicit or explicit; its only model is the sparsity of the resonance components in high-Q and low-Q transforms.

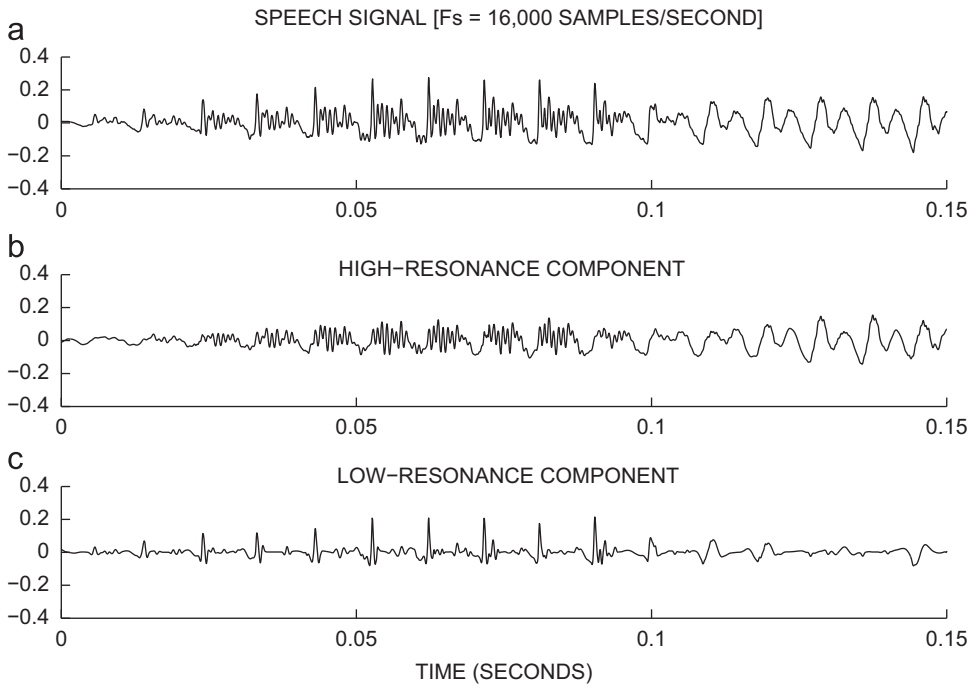

Fig. 10. Decomposition of a speech signal ("I'm") into high- and low-resonance components. The high-resonance component (b) contains the sustained oscillations present in the speech signal, while the low-resonance component (c) contains non-oscillatory transients. (The residual is not shown.)

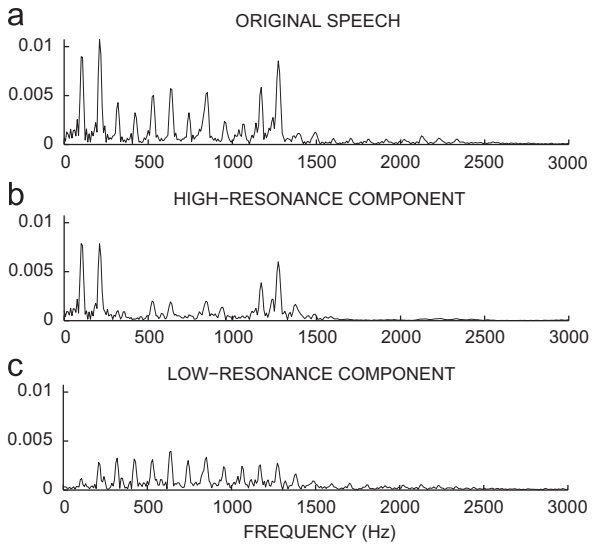

Fig. 11. Frequency spectra of the speech signal in Fig. 10 and of the extracted high- and low-resonance components. The spectra are computed using the 50 ms segment from 0.05 to 0.10 s. The energy of each resonance component is widely distributed in frequency and their frequency-spectra overlap. (a) Original speech, (b) high-resonance component and (c) low-resonance component.

Therefore, although it does not yield a generative model for speech as does a source-filter model, resonance-based decomposition is not adversely affected by rapidly variations in the pitch period to which some other techniques are sensitive.

Table 2
Comparison of prediction error for AR modeling. The table lists e_{hr}/e_{orig} where e_{orig} is the prediction error (σ) for the original speech signal, and e_{hr} is likewise the prediction error for the high-resonance component, illustrated in Fig. 10. The model order is denoted by p. Each method is labeled by its Matlab function name.

Method	e_{hr}/e_{orig}				
	p=6	p=7	p=8	p=9	p=10
`burg`	0.147	0.119	0.082	0.070	0.047
`arcov`	0.146	0.119	0.082	0.070	0.047
`armcov`	0.147	0.119	0.082	0.070	0.047
`aryule`	0.401	0.387	0.417	0.416	0.421
`lpc`	0.401	0.387	0.417	0.416	0.421

We also note that the high-resonance component appears more amenable to sinusoidal modeling [73] than does the original speech signal. Sinusoidal modeling, a method for representing speech as a sum of time-varying sinusoids, is useful in speech coding and manipulation (pitch scaling, voice morphing, etc.). However, impulses are not efficiently represented as a sum of sinusoids and their presence degrades the effectiveness of sinusoidal modeling. Because the high-resonance component is largely free of impulses and transients, sinusoidal modeling can be expected to be especially effective when applied to it.

To make a preliminary quantification of the compressibility or predictability of the high-resonance component in comparison with the original speech signal, we applied AR modeling to both signals using the methods and model orders listed in Table 2. In each case, the prediction error

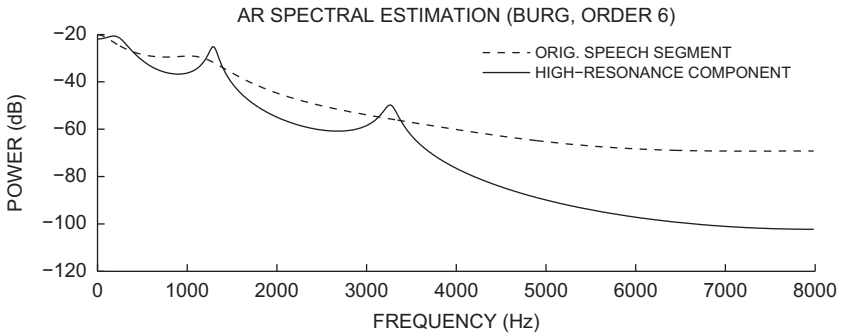

Fig. 12. AR spectral estimation using the Burg method with model order $p=6$, for the speech signal and its high-resonance component, shown in Fig. 10.

(standard deviation, σ) was computed, denoted as e_{orig} and e_{hr} for the original speech signal and high-resonance component, respectively. It was found that e_{hr} is substantially less than e_{orig}. Table 2 lists the ratio, $e_{\mathrm{hr}}/e_{\mathrm{orig}}$, for each method and model order. For example, using the Burg AR method with model order $p=6$, the prediction error of the high-resonance component is 14.7% that of the original speech signal. The values in Table 2 suggest that the high-resonance component is substantially more predictable than the original speech signal, at least when an AR model is used for the prediction. Fig. 12 illustrates the estimated power spectral density using the Burg method with $p=6$; the formants are more clearly defined for the high-resonance component, as expected from the signal waveforms in Fig. 10. Therefore, the high-resonance component may facilitate formant tracking with fine temporal resolution.

As suggested by its more distinctive peaks in Fig. 12, the high-resonance component can accentuate the speech formants, the relative spacing of which characterize distinct vowels, etc. Meanwhile, the low-resonant component can accentuate the harmonics, which characterize the pitch of the speaker. In this case, coding only the high-resonance component may serve as a method for more efficient speech coding. Given that the speech waveform carries both the identity of the speaker and the spoken message, it has been postulated that 'who' and 'what' are conveyed by the auditory system to the cortex along separate sensory pathways [64]. Resonance-based decomposition may similarly provide a separation along these lines.

A principal tool for the analysis of speech is the spectrogram; according to Pitton et al., "Practically every aspect of speech communication has greatly benefited from time–frequency (TF) analysis" [77]. However, the pursuit of alternatives to the spectrogram has led to the development of a variety of adaptive nonlinear time–frequency distributions with improved resolution properties [16,45,47,48,88]. For certain signals (multicomponent AM/FM signals, or more generally, 'high-resonance' signals) these powerful TF techniques reveal the signal's time-varying frequency characteristics that cannot be seen from the time-domain waveform itself. However, for other signals, the information of interest is more easily ascertained from the time-domain waveform (for example, inter-pulse intervals of a neural spike sequence, or more generally, 'low-resonance' signals). For signals consisting of a mixture of simultaneous high- and low-resonance components, existing time–frequency analysis techniques might be more effectively applied to the high-resonance component than to the original signal.

6. Further remarks

Parameter selection. As the minimization of the objective function in (1) depends on the parameters λ_1 and λ_2, the resulting decomposition can be tuned by varying these parameters. The relative values of λ_1 and λ_2 influence the energy of the two components: for example, with λ_1 fixed, increasing λ_2 will decrease the energy of $\hat{\mathbf{x}}_2$ and increase the energy of $\hat{\mathbf{x}}_1$. The values of λ_i also influence the energy of the residual: increasing both λ_1 and λ_2 will decrease the energy of both components and increase the energy of the residual. For the examples in this paper, the parameters were manually selected by visually inspecting the induced components. An appropriate balance between λ_1 and λ_2 must be struck. Automatic selection of the parameters λ_i could be performed using hyper-parameter estimation procedures, as described in [15] for example. For image decomposition into texture and structure components, Ref. [3] describes a method for parameter selection based on the assumption that these two components are not correlated; a similar approach may be useful here.

The selection of the Q-factors of the two constant-Q transforms also influences the result. It appears reasonable to use $Q \approx 1$ for the low-Q transform (like the dyadic wavelet transform) in order to obtain a sparse representation of the non-oscillatory component; and that the high Q-factor should be set depending on the oscillatory behaviour of signal in question. However, the decomposition result does not appear to be as sensitive to the Q-factors as it is to the λ_i.

Additionally, the parameters λ_i and Q_i could be selected based on optimizing appropriate performance measurements, along the lines of [103] for example.

Why not an ℓ_2-norm *penalty?* If the ℓ_2-norm is used in the penalty term of (1),

$$J(\mathbf{w}_1,\mathbf{w}_2) = \|\mathbf{x}-\mathbf{S}_1\mathbf{w}_1-\mathbf{S}_2\mathbf{w}_2\|_2^2+\lambda_1\|\mathbf{w}_1\|_2^2+\lambda_2\|\mathbf{w}_2\|_2^2, \qquad (16)$$

then, using $\mathbf{S}_1\mathbf{S}_1^t = \mathbf{S}_2\mathbf{S}_2^t = \mathbf{I}$, the minimizing $\mathbf{w}_1$ and $\mathbf{w}_2$ can be found in closed form,

$$\mathbf{w}_1 = \frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_1\lambda_2}\mathbf{S}_1^t\mathbf{x},$$

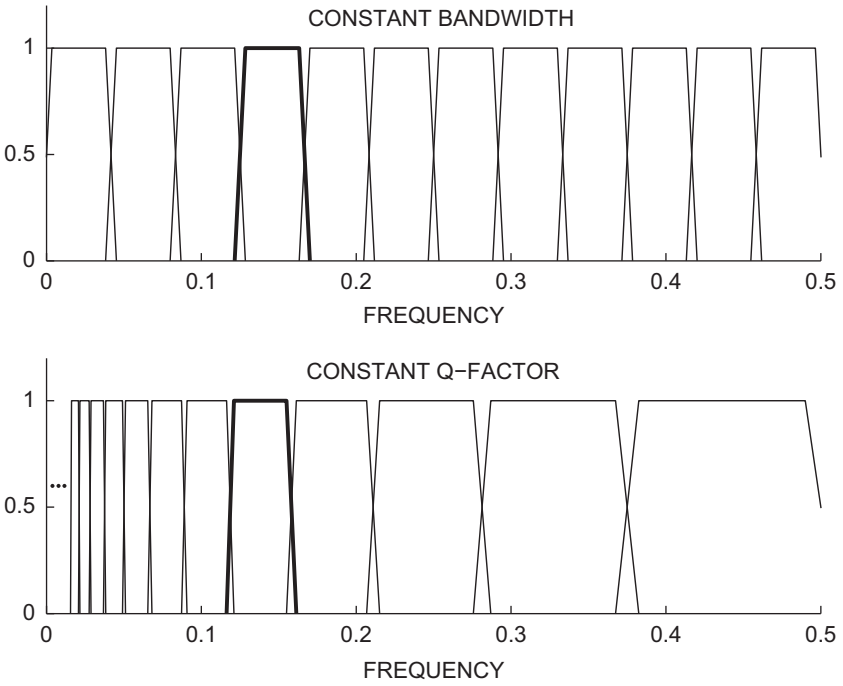

Fig. 13. A constant-bandwidth and a constant-Q decomposition may have analysis functions with similar frequency support.

$$\mathbf{w}_2 = \frac{\lambda_2}{\lambda_1+\lambda_2+\lambda_1\lambda_2}\mathbf{S}_2^t\mathbf{x},$$

and the estimated components, $\hat{\mathbf{x}}_1 = \mathbf{S}_1\mathbf{w}_1$ and $\hat{\mathbf{x}}_2 = \mathbf{S}_2\mathbf{w}_2$, are given by

$$\hat{\mathbf{x}}_1 = \frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_1\lambda_2}\mathbf{x},$$

$$\hat{\mathbf{x}}_2 = \frac{\lambda_2}{\lambda_1+\lambda_2+\lambda_1\lambda_2}\mathbf{x}.$$

That is, both $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$ are simply scaled versions of $\mathbf{x}$. Although the objective function (16) admits a closed form minimizer, it does not lead to any decomposition whatsoever.

More than two resonance components. The same approach can be used with more than two resonance components, although we have not explored it here. The main issue is that the use of more transforms will generally reduce the incoherence between the transforms which diminishes the likelihood of obtaining distinct components. However, if a constant-Q transform with sufficiently high Q-factor is used to define a third resonance component, then the incoherence will be maintained and the performance of MCA should not be compromised. Specifically, Q_3 should be chosen sufficient large so that $\rho_{\max}(Q_2,Q_3) \le \rho_{\max}(Q_1,Q_2)$. Assuming ideal band-pass wavelets, (2) can be used to calculate the minimum Q_3.

MCA using constant-Q and constant-BW transforms. For the separation of oscillatory and transient signals using sparse signal representations, it is quite natural to utilize a short-time Fourier transform (or similar constant-bandwidth transform, like the MDCT) and a wavelet transform (with low-Q, like the dyadic WT). This approach is presented in [23,24,75,96] which illustrate excellent separation results. However, depending on the parameters (window length, etc.), a constant bandwidth and a constant Q-factor decomposition may have analysis functions with similar frequency support, as illustrated in Fig. 13. (The pass-bands between 0.1 and 0.2 have significant overlap, as indicated in the figure.) In this case, the two transforms will have a high coherence because the maximum inner product between analysis functions of the two transforms will be close to unity, which can degrade the achievable component separation, depending on the signal being analyzed.

On the other hand, as noted above, two constant-Q transforms with markedly different Q-factors will have low coherence because no analysis functions from the two decompositions will have similar frequency support. Likewise, two constant-bandwidth transforms with markedly different bandwidths will also have low coherence and are therefore also suitable for MCA-based signal decomposition. This calls for short and long windows for the two constant-BW transforms, respectively. This approach is presented in [25,38,57].

7. Related work and concepts

Related decompositions: The problem of decomposing a given signal into oscillatory and non-oscillatory components has received a fair amount of attention, especially for speech and audio processing; however, previous approaches, not being based on signal resonance, are different from the method presented here. Speech and musical sounds are often modeled as consisting of three components: a quasi-sinusoidal (oscillatory) component, a component consisting of temporal transients, and a stochastic (noise-like) component. For speech and audio processing (pitch- or time-scaling, morphing, enhancement, de-reverberation, coding, synthesis of speech

and other sounds) it is often useful that these three components, each of which are psycho-acoustically significant, be modeled separately. Early work on this topic decomposed sounds using a 'deterministic+stochastic' model [87], a 'sine+noise' model [86], or a 'harmonic+noise' model [66]. Subsequent work decomposed sounds using a 'sine+transient +noise' model [67,68,100] where a transient (deterministic, non-oscillatory) component is introduced. Related works consider a 'periodic+aperiodic' model [19,107] and pitch-synchronous models [34,78]. These multi-component models follow, elaborate and enhance, in a sense, sinusoidal modeling of speech [73,80]. Although these methods decompose signals into oscillatory, non-oscillatory, and residual components, as does the proposed method, they estimate and extract the oscillatory component using constant-bandwidth transforms rather than constant-Q transforms, and hence they do not perform resonance-based signal decomposition.

More recent methods for decomposing signals into oscillatory and transient components are based on sparse signal representations [24,25,38,75] as is the method described here; however, these methods are again not based on signal resonance as they use constant-bandwidth transforms as one or both of the two transforms. The dyadic (low Q-factor) wavelet transform and the modified discrete cosine transform (MDCT) are used in [24,75], while over-complete modulated complex lapped transform (MCLT) dictionaries are utilized in [25]. Note that the MDCT and MCLT are constant-bandwidth, rather than constant-Q, transforms. A Bayesian approach is described in [38] and is illustrated using an MDCT basis with a long window for the tonal component and an MDCT with a short window for the transient component. Along these lines, the use of MDCT and wavelet transforms for speech enhancement was discussed in [96]. Ref. [23] develops a new algorithm: the molecular matching pursuit algorithm. An example of tonal/transient separation in audio is also given in [57] which introduces the time–frequency jigsaw puzzle.

Another class of methods utilizes low Q-factor wavelet transforms for the detection, estimation and characterization of transients in various (often noisy) signals [20,79]; however, these methods differ from resonance-based signal decomposition in that they do not explicitly account for, or model, the presence of a highly oscillatory (high-resonance) component. These methods (for example, [20,79]) utilize wavelet transforms having low Q-factors (the dyadic wavelet, etc.) due to the ability of such wavelet transforms to efficiently represent the transient (low-resonance) component of a signal. Along these lines, a novel approach to transient separation is the time-scale method introduced in [13] which exploits the multi-scale characterization of transients.

The decomposition of a signal into a set of pulses of various time–frequency characteristics using matching pursuits [72] and related algorithms has been effective in a number of applications; however, these methods also differ from the proposed method in that they usually utilize sets, or dictionaries, of functions that are far more overcomplete than the constant-Q transforms utilized here, and the use of these techniques for resonance-based signal analysis has not been explored. The approximation of EEG signals by Gabor functions using matching pursuits (MP) has been applied to the analysis of sleep EEG [71], evoked potentials [90], epileptic activities, and several other research and clinical problems in EEG signal analysis [30,31], as well as to the estimation of otoacoustic emissions [58].

Empirical mode decomposition: Empirical mode decomposition (EMD) is another nonlinear signal decomposition [46,55]. One goal of EMD is to extract AM/FM components from a multicomponent signal. EMD decomposes a signal into components, called 'intrinsic mode functions' (IMFs), which are approximately AM/FM functions. While EMD can yield similar results as signal-resonance decomposition for certain signals, the objectives and algorithms of the two approaches are quite different.

Constant-Q transforms and non-uniform frequency decomposition: The biological prevalence of constant-Q analysis led to the study of constant-Q transforms for signal processing starting in the 1970s [49,53,76,110] and continuing until the present. In fact, the continuous wavelet transform was effectively described in [110]. The calculation and use of constant-Q transforms for the analysis of musical notes is described in [11,12]. However, continuous-time integral transforms are highly overcomplete and not always easily inverted [43], yet some solutions are presented in [54,56,69,70]. More recent papers have drawn on the theory of perfect-reconstruction critically sampled filter banks to design discrete-time transforms with non-uniform frequency analysis. These transforms, mostly based on wavelet-packets [104] and therefore easily invertible, have been applied to audio coding [91,93], speech enhancement [14,81,89], and speech quality evaluation [60]. Wavelet-packet transforms have been designed to approximate the critical-bands of the human auditory system [61]; however, they cannot achieve the constant-Q property exactly [65]. As mentioned in Section 3.1, earlier discrete-time rational-dilation filter banks and wavelet transforms have been presented in [9,10,62,106]. Other approaches to the design of approximately constant-Q filter banks are described in [26,52,83,92,99]; some of these methods use frequency warping or multiple voices and have approximate perfect reconstruction.

8. Conclusion

Signal 'resonance', being an attribute distinct and independent to that of frequency, provides a dimension complementary to frequency for the analysis of signals. While frequency components are straightforwardly defined and can be obtained by linear filtering, resonance components are more difficult to define and procedures to obtain resonance components are necessarily nonlinear. Related decompositions ('sines + transients + noise', etc.) have been proposed for speech and audio processing, but are based partly are wholly on constant-bandwidth transforms and therefore do not provide a resonance-based decomposition.

This paper describes an algorithm for resonance-based decomposition that relies on techniques for sparse signal representations and on morphological component analysis (MCA). The algorithm uses a rational-dilation wavelet transform (RADWT) for the sparse representation of each resonance component. The RADWT is a self-inverting fully discrete transform which is important for the SALSA

algorithm as described. We expect that other (near) constant-Q transforms may be used in place of the RADWT.

We note that for real signals there will rarely be a totally unambiguous distinction between the low and high resonance components; consequently it is expected that some low-resonance behaviour appears in the high-resonance component and vice versa. Additional modeling, for example by 'structured-sparsity' or context modeling of the two components (e.g. [23,63]), may improve the separation capability of the approach.

Acknowledgements

The author thanks Ivan Bodis-Wollner and Hans von Gizycki of the State University of New York, Downstate Medical Center, for enlightening discussions regarding the EEG. The author also thanks Gerald Schuller of the Technical University of Ilmenau, Chin-Tuan Tan of the New York University, School of Medicine, and the anonymous reviewers for valuable suggestions and corrections.

References

[1] M.V. Afonso, J.M. Bioucas-Dias, M.A.T. Figueiredo, Fast image recovery using variable splitting and constrained optimization, IEEE Trans. Image Process. 19 (9) (2010) 2345–2356. doi:10.1109/TIP.2010.2047910.
[2] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, A. Chambolle, Image decomposition into a bounded variation component and an oscillating component, J. Math. Imag. Vis. 22 (2005) 71–88.
[3] J.-F. Aujol, G. Gilboa, T. Chan, S.J. Osher, Structure-texture image decomposition—modeling, algorithms, and parameter selection, Int. J. Comput. Vis. 67 (2006) 111–136.
[4] P. Auscher, Wavelet bases for $L^2(R)$ with rational dilation factor, in: M.B. Ruskai et al. (Ed.), Wavelets and their Applications, Jones and Barlett, Boston, 1992.
[5] A. Baussard, F. Nicolier, F. Truchetet, Rational multiresolution analysis and fast wavelet transform: application to wavelet shrinkage denoising, Signal Processing 84 (10) (2004) 1735–1747.
[6] I. Bayram, I.W. Selesnick, Frequency-domain design of overcomplete rational-dilation wavelet transforms, IEEE Trans. Signal Process. 57 (8) (2009) 2957–2972.
[7] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imag. Sci. 2 (1) (2009) 183–202.
[8] J.M. Bioucas-Dias, M.A.T. Figueiredo, A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration, IEEE Trans. Image Process. 16 (12) (2007) 2992–3004.
[9] T. Blu, Iterated filter banks with rational rate changes—connection with discrete wavelet transforms, IEEE Trans. Signal Process. 41 (12) (1993) 3232–3244.
[10] T. Blu, A new design algorithm for two-band orthonormal rational filter banks and orthonormal rational wavelets, IEEE Trans. Signal Process. 46 (6) (1998) 1494–1504.
[11] J.C. Brown, Calculation of a constant Q spectral transform, J. Acoust. Soc. Am. 89 (1) (1991) 425–434.
[12] J.C. Brown, M.S. Puckette, An efficient algorithm for the calculation of a constant Q transform, J. Acoust. Soc. Am. 92 (5) (1992) 2698–2701.
[13] V. Bruni, D. Vitulano, Transients detection in the time-scale domain, in: Proceedings of the 3rd International Conference on Image and Signal Processing (ICISP), Lecture Notes in Computer Science, vol. 5099, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 254–262.
[14] B. Carnero, A. Drygajlo, Perceptual speech coding and enhancement using frame-synchronized fast wavelet packet transform algorithms, IEEE Trans. Signal Process. 47 (6) (1999) 1622–1635.
[15] L. Chaâri, J.-C. Pesquet, J.-Y. Tourneret, Ph. Ciuiu, A. Benazza-Benyahia, A hierarchical Bayesian model for frame representation. IEEE Trans. Signal Process. 58 (11) (2010) 5560–5571. doi:10.1109/TSP.2010.2055562.
[16] L. Cohen, Time Frequency Analysis: Theory and Applications, Prentice-Hall, 1994.
[17] P.L. Combettes, J.-C. Pesquet, Proximal thresholding algorithm for minimization over orthonormal bases, SIAM J. Optim. 18 (4) (2007) 1351–1376.
[18] P.L. Combettes, J.-C. Pesquet, Proximal splitting methods in signal processing, in: H.H. Bauschke, R. Burachik, P.L. Combettes, V. Elser, D.R. Luke, H. Wolkowicz (Eds.), Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Springer-Verlag, New York, 2010.
[19] C. d'Alessandro, V. Darsinos, B. Yegnanarayana, Effectiveness of a periodic and aperiodic decomposition method for analysis of voice sources, IEEE Trans. Speech Audio Process. 6 (1) (1998) 12–23.
[20] C.E. D'Attellis, S.I. Isaacson, R.O. Sirne, Detection of epileptic events in electroencephalograms using wavelet analysis, Ann. Biomed. Eng. 25 (1997) 286–293.
[21] I. Daubechies, Ten Lectures On Wavelets, SIAM, 1992.
[22] I. Daubechies, M. Defriese, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Commun. Pure Appl. Math. LVII (2004) 1413–1457.
[23] L. Daudet, Sparse and structured decompositions of signals with the molecular matching pursuit, IEEE Trans. Audio Speech Lang. Proc. 14 (5) (2006) 1808–1816.
[24] L. Daudet, B. Torrésani, Hybrid representations for audiophonic signal encoding, Signal Processing 82 (11) (2002) 1595–1617.
[25] M.E. Davies, L. Daudet, Sparse audio representations using the MCLT, Signal Processing 86 (3) (2006) 457–470.
[26] F.C.C.B. Diniz, I. Koethe, S.L. Netto, L.W.P. Biscainho, High-selectivity filter banks for spectral analysis of music signals, EURASIP Journal on Advances in Signal Processing (2007) 12, doi: 10.1155/2007/94704 (2007:Article ID 94704).
[27] O. Divorra Escoda, L. Granai, M. Lemay, J.M. Hernandez, P. Vandergheynst, J.-M. Vesin, Ventricular and atrial activity estimation through sparse ECG signal decompositions, in: Proceedings of the IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP), 2006.
[28] O. Divorra Escoda, L. Granai, P. Vandergheynst, On the use of a priori information for sparse signal approximations, IEEE Trans. Signal Process. 54 (9) (2006) 3468–3482.
[29] D. Donoho, Compressed sensing, IEEE Trans. Inform. Theory 52 (4) (2006) 1289–1306.
[30] P. Durka, From wavelets to adaptive approximations: time–frequency parametrization of EEG, BioMed. Eng. OnLine 2 (1) (2003) 1.
[31] P. Durka, On the methodological unification in electroencephalography, BioMed. Eng. OnLine 4 (1) (2005) 15.
[32] M. Elad, B. Matalon, M. Zibulevsky, Coordinate and subspace optimization methods for linear least squares with non-quadratic regularization, J. Appl. Comput. Harm. Anal. 23 (2007) 346–367.
[33] M. Elad, J. Starck, P. Querre, D. Donoho, Simultaneous cartoon and texture image inpainting using morphological component analysis MCA, J. Appl. Comput. Harm. Anal. 19 (3) (2005) 340–358.
[34] G. Evangelista, S. Cavaliere, Discrete frequency warped wavelets: theory and applications, IEEE Trans. Signal Process. 46 (4) (1998) 874–885.
[35] J.M. Fadili, J.-L. Starck, M. Elad, D.L. Donoho, MCALab: reproducible research in signal and image decomposition and inpainting, Comput. Sci. Eng. 12 (2009) 44–63.
[36] M.J. Fadili, J.-L. Starck, J. Bobin, Y. Moudden, Image decomposition and separation using sparse representations: an overview, Proc. IEEE 98 (6) (2010) 983–994. doi:10.1109/JPROC.2009.2024776.
[37] M.J. Fadili, J.-L. Starck, F. Murtagh, Inpainting and zooming using sparse representations, Comput. J. 52 (1) (2009) 64–79 Published online on July 29, 2007.
[38] C. Févotte, L. Daudet, S.J. Godsill, B. Torrésani, Sparse regression with structured priors: Application to audio denoising, in: Proceedings of the IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP), vol. 3, 2006.
[39] M. Figueiredo, J. Bioucas-Dias, R. Nowak, Majorization-minimization algorithms for wavelet-based image restoration, IEEE Trans. Image Process. 16 (12) (2007) 2980–2991.
[40] M. Figueiredo, R. Nowak, An EM algorithm for wavelet-based image restoration, IEEE Trans. Image Process. 12 (8) (2003) 906–916.
[41] M.A.T. Figueiredo, J.M. Bioucas-Dias, M.V. Afonso, Fast frame-based image deconvolution using variable splitting and constrained optimization, in: IEEE Workshop on Statistical Signal Processing, August 31–September 3, 2009.
[42] M.A.T. Figueiredo, R.D. Nowak, S.J. Wright, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems, IEEE J. Sel. Top. Signal Process. 1 (4) (2007) 586–598.

[43] D. FitzGerald, M. Cranitch, M.T. Cychowski, Towards an inverse constant Q transform, in: 120th AES Convention, Paris, May 2006.
[44] J. L Flanagan, Models for approximating basilar membrane displacement, Bell Syst. Tech. J. 39 (1960) 1163–1191.
[45] P. Flandrin, Time–Frequency/Time-Scale Analysis, Academic Press, 1999.
[46] P. Flandrin, P. Goncalvès, Empirical mode decompositions as data-driven wavelet-like expansions, Int. J. Wavelets Multires. Inf. Process. 2 (4) (2004) 477–496.
[47] S.A. Fulop, Phonetic applications of the time-corrected instantaneous frequency spectrogram, Phonetica 64 (2007) 237–262.
[48] S.A. Fulop, K. Fitz, A spectrogram for the twenty-first century, Acoust. Today 2 (3) (2006) 26–33.
[49] G. Gambardella, A contribution to the theory of short-time spectral analysis with nonuniform bandwidth filters, IEEE Trans. Circuit Theory 18 (4) (1971) 455–460.
[50] B.R. Glasberg, B.C.J. Moore, Derivation of auditory filter shapes from notched-noise data, Hear. Res. 47 (1990) 103–138.
[51] L. Granai, P. Vandergheynst, Sparse decomposition over multi-component redundant dictionaries, in: IEEE Workshop on Multimedia Signal Processing (MMSP), 2004.
[52] T. Gulzow, T. Ludwig, U. Heute, Spectral-subtraction speech enhancement in multirate systems with and without non-uniform and adaptive bandwidths, Signal Processing 83 (8) (2003) 1613–1631.
[53] H. Helms, Power spectra obtained from exponentially increasing spacings of sampling positions and frequencies, IEEE Trans. Acoust. Speech Signal Proc. 24 (1) (1976) 63–71.
[54] V. Hohmann, Frequency analysis and synthesis using a gammatone filterbank, Acta Acustica United with Acustica 88 (3) (2002) 433–442.
[55] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.-C. Yen, C.C. Tung, H.H. Liu, The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. R. Soc. London A 454 (1971) 903–995 (March 8, 1998).
[56] T. Irino, H. Kawahara, Signal reconstruction from modified auditory wavelet transform, IEEE Trans. Signal Process. 41 (12) (1993) 3549–3554.
[57] F. Jaillet, B. Torrésani, Time–frequency jigsaw puzzle: adaptive multiwindow and multilayered Gabor representations, Int. J. Wavelets Multires. Inf. Process. 5 (2) (2007) 293–316.
[58] W.W. Jedrzejczak, K.J. Blinowska, W. Konopka, A. Grzanka, P.J. Durka, Identification of otoacoustic emissions components by means of adaptive approximations, J. Acoust. Soc. Am. 115 (5) (2004) 2148–2158.
[59] B.D. Johnson, Stable filtering schemes with rational dilations, J. Fourier Anal. Appl. 13 (5) (2007) 607–621.
[60] A. Karmakar, A. Kumar, R.K. Patney, A multiresolution model of auditory excitation pattern and its application to objective evaluation of perceived speech quality, IEEE Trans. Audio Speech Lang. Proc. 14 (6) (2006) 1912–1923.
[61] A. Karmakar, A. Kumar, R.K. Patney, Design of optimal wavelet packet trees based on auditory perception criterion, IEEE Signal Process. Lett. 14 (4) (2007) 240–243.
[62] J. Kovačević, M. Vetterli, Perfect reconstruction filter banks with rational sampling factors, IEEE Trans. Signal Process. 41 (6) (1993) 2047–2066.
[63] M. Kowalski, B. Torrésani, Sparsity and persistence: mixed norms provide simple signal models with dependent coefficients, Signal Image Video Process. 3 (3) (2009) 251–264.
[64] N. Kraus, T. Nicol, Brainstem origins for cortical 'what' and 'where' pathways in the auditory system, Trends Neurosci. 28 (4) (2005) 176–181.
[65] F. Kurth, M. Clausen, Filter bank tree and m-band wavelet packet algorithms in audio signal processing, IEEE Trans. Signal Process. 47 (2) (1999) 549–554.
[66] J. Laroche, Y. Stylianou, E. Moulines, HNS: Speech modification based on a harmonic+noise model, in: Proceedings of the IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP), vol. 2, April 1993.
[67] N. Laurenti, G. De Poli, D. Montagner, A nonlinear method for stochastic spectrum estimation in the modeling of musical sounds, IEEE Trans. Audio Speech Lang. Proc. 15 (2) (2007) 531–541.
[68] S.N. Levine, J.O. Smith III, A sines+transients+noise audio representation for data compression and time/pitch scale modifications, in: Proceedings of the 105th Audio Engineering Society Convention, 1998.
[69] Q. Li, An auditory-based transform for audio signal processing, in: IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (WASPAA), 2009.
[70] L. Lin, W.H. Holmes, E. Ambikairajah, Auditory filter bank inversion, in: Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), vol. 2, May 2001.
[71] U. Malinowska, P.J. Durka, K.J. Blinowska, W. Szelenberger, A. Wakarow, Micro- and macrostructure of sleep EEG, IEEE Eng. Med. Bio. Mag. 25 (4) (2006) 26–31.
[72] S. Mallat, Z. Zhang, Matching pursuits with time–frequency dictionaries, IEEE Trans. Signal Process. 41 (1993) 3397–3415.
[73] R. McAulay, T. Quatieri, Speech analysis/synthesis based on a sinusoidal representation, IEEE Trans. Acoust. Speech Signal Process. 34 (4) (1986) 744–754.
[74] Y. Meyer, Oscillating Patterns in Image Processing and Nonlinear Evolution Equations, University Lecture series American Mathematical Society, vol. 22, 2001.
[75] S. Molla, B. Torrésani, An hybrid audio coding scheme using hidden Markov models of waveforms, J. Appl. Comput. Harm. Anal. 18 (2) (2005) 137–166.
[76] T. Petersen, S. Boll, Critical band analysis-synthesis, IEEE Trans. Acoust. Speech Signal Process. 31 (3) (1983) 656–663.
[77] J.W. Pitton, K. Wang, B.-H. Juang, Time–frequency analysis and auditory modeling for automatic recognition of speech, Proc. IEEE 84 (9) (1996) 1199–1215.
[78] P. Polotti, G. Evangelista, Fractal additive synthesis, IEEE Signal Process. Mag. 24 (2) (2007) 105–115.
[79] L.-S. Pon, M. Sun, R.J. Sclabassi, Adaptive separation of background activity and transient phenomenon in epileptic EEG using mathematical morphology and wavelet transforms, in: Proceedings of the IEEE Annual EMBS Conference, 2000.
[80] T. Quatieri, Discrete-Time Speech Signal Processing: Principles and Practice, Prentice-Hall, 2001.
[81] Y. Ren, M.T. Johnson, J. Tao, Perceptually motivated wavelet packet transform for bioacoustic signal enhancement, J. Acoust. Soc. Am. 124 (2008) 316–327.
[82] P. Sausenga, W. Klimescha, W.R. Grubera, S. Hanslmayrb, R. Freunbergera, M. Doppelmayra, Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion, Neuroscience 146 (4) (2007) 1435–1444.
[83] C. Schörkhuber, A. Klapuri, Constant-Q transform toolbox for music processing, in: 7th Sound and Music Conference, July 2010.
[84] I.W. Selesnick, A new sparsity-enabled signal separation method based on signal resonance, in: Proceedings of the IEEE International Conference Acoustics, Speech, Signal Processing (ICASSP), 2010.
[85] I.W. Selesnick, I. Bayram, Oscillatory+transient signal decomposition using overcomplete rational-dilation wavelet transforms, in: Proceedings of SPIE, vol. 7446 (Wavelets XIII), 2009.
[86] X. Serra, Musical sound modeling with sinusoids plus noise, in: C. Roads, S.T. Pope, A. Picialli, G. De Poli (Eds.), Musical Signal Processing, Swets & Zeitlinger1997, pp. 91–122.
[87] X. Serra, J. Smith, Spectral modeling synthesis: a sound analysis synthesis system based on a deterministic plus stochastic decomposition, Comput. Music J. 14 (1990) 12–24.
[88] I. Shafi, J. Ahmad, S.I. Shah, F.M. Kashif, Techniques to obtain good resolution and concentrated time–frequency distributions: a review. EURASIP J. Adv. Signal Proc. 2009 (2009).
[89] Y. Shao, C.-H. Chang, A generalized time–frequency subtraction method for robust speech enhancement based on wavelet filter banks modeling of human auditory system, IEEE Trans. Syst. Man Cybern. Part B 37 (4) (2007) 877–889.
[90] C. Sieluzycki, R. Konig, A. Matysiak, R. Kus, D. Ircha, P.J. Durka, Single-trial evoked brain responses modeled by multivariate matching pursuit, Trans. Biomed. Eng. 56 (1) (2009) 74–82.
[91] D. Sinha, A.H. Tewfik, Low bit rate transparent audio compression using adapted wavelets, IEEE Trans. Signal Process. 41 (12) (1993) 3463–3479.
[92] J.O. Smith, J.S. Abel, Bark and ERB bilinear transforms, IEEE Trans. Speech Audio Process. 7 (6) (1999) 697–708.
[93] P. Srinivasan, L.H. Jamieson, High-quality audio compression using an adaptive wavelet packet decomposition and psychoacoustic modeling, IEEE Trans. Signal Process. 46 (4) (1998) 1085–1093.
[94] J.-L. Starck, M. Elad, D. Donoho, Image decomposition via the combination of sparse representation and a variational approach, IEEE Trans. Image Process. 14 (10) (2005).
[95] J.-L. Starck, Y. Moudden, J. Bobina, M. Elad, D.L. Donoho, Morphological component analysis, in: Proceedings of SPIE, vol. 5914 (Wavelets XI), 2005.
[96] C. Tantibundhit, J.R. Boston, C.C. Li, J.D. Durrant, S. Shaiman, K. Kovacyk, A. El-Jaroudi, New signal decomposition method based speech enhancement, Signal Processing 87 (11) (2007) 2607–2628.

[97] The Open University. Waves, Tides and Shallow-Water Processes, Butterworth-Heinemann, 2000.
[98] M. Unoki, T. Irino, B. Glasberg, B.C.J. Moore, R.D. Patterson, Comparison of the roex and gammachirp filters as representations of the auditory filter, J. Acoust. Soc. Am. 120 (3) (2006) 1474–1492.
[99] P. Vary, An adaptive filter-bank equalizer for speech enhancement, Signal Processing 86 (6) (2006) 1206–1214.
[100] T.S. Verma, T.H.Y. Meng, Time scale modification using a sines+transients+noise signal model, in: Proceedings of the Digital Audio Effect Workshop (DAFX), 1998, pp. 49-52.
[101] L.A. Vese, S. Osher, Modeling textures with total variation minimization and oscillating patterns in image processing, J. Sci. Comput. 15 (2003) 553–572.
[102] L.A. Vese, S. Osher, Image denoising and decomposition with total variation minimization and oscillatory functions, J. Math. Imag. Vis. 20 (2004) 7–18.
[103] E. Vincent, R. Gribonval, C. Févotte, Performance measurement in blind audio source separation, IEEE Trans. Audio Speech Lang. Proc. 14 (4) (2006) 1462–1469.
[104] M.V. Wickerhauser, Adapted Wavelet Analysis from Theory to Software, AK Peters Ltd, 1994.
[105] S.J. Wright, R.D. Nowak, M.A.T. Figueiredo, Sparse reconstruction by separable approximation, IEEE Trans. Signal Process. 57 (7) (2009) 2479–2493.
[106] X.M. Xie, S.C. Chan, T.I. Yuk, Design of perfect-reconstruction nonuniform recombination filter banks with flexible rational sampling factors, IEEE Trans. Circuits Systems I 52 (9) (2005) 1965–1981.
[107] B. Yegnanarayana, C. d'Alessandro, V. Darsinos, An iterative algorithm for decomposition of speech signals into periodic and aperiodic components, IEEE Trans. Speech Audio Process. 6 (1) (1998) 1–11.
[108] N. Yeung, R. Bogacz, C.B. Holroyd, J.D. Cohen, Detection of synchronized oscillations in the electroencephalogram: an evaluation of methods, Psychophysiology 41 (2004) 822–832.
[109] N. Yeung, R. Bogacz, C.B. Holroyd, S. Nieuwenhuis, J.D. Cohen, Theta phase resetting and the error related negativity, Psychophysiology 44 (2007) 39–49.
[110] J. Youngberg, S. Boll, Constant-Q signal analysis and synthesis, in: Proceedings of the IEEE International Conference on Acoust., Speech, Signal Processing (ICASSP), vol. 3, April 1978, pp. 375–378.