Ultrasonics 52 (2012) 351–363

Contents lists available at SciVerse ScienceDirect

Ultrasonics

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

Sparse signal representation and its applications in ultrasonic NDE

Guang-Ming Zhang [a,*], Cheng-Zhong Zhang [b], David M. Harvey [a]

[a] General Engineering Research Institute, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, United Kingdom
[b] Department of Computer Engineering, Nanhai Campus, South China Normal University, Foshan, Guangdong 528225, China

A R T I C L E I N F O

Article history:
Received 23 March 2010
Received in revised form 29 September 2011
Accepted 1 October 2011
Available online 10 October 2011

Keywords:
Sparse signal representation
Overcomplete dictionary
Ultrasonic NDE
Ultrasonic signal processing

A B S T R A C T

Many sparse signal representation (SSR) algorithms have been developed in the past decade. The advantages of SSR such as compact representations and super resolution lead to the state of the art performance of SSR for processing ultrasonic non-destructive evaluation (NDE) signals. Choosing a suitable SSR algorithm and designing an appropriate overcomplete dictionary is a key for success. After a brief review of sparse signal representation methods and the design of overcomplete dictionaries, this paper addresses the recent accomplishments of SSR for processing ultrasonic NDE signals. The advantages and limitations of SSR algorithms and various overcomplete dictionaries widely-used in ultrasonic NDE applications are explored in depth. Their performance improvement compared to conventional signal processing methods in many applications such as ultrasonic flaw detection and noise suppression, echo separation and echo estimation, and ultrasonic imaging is investigated. The challenging issues met in practical ultrasonic NDE applications for example the design of a good dictionary are discussed. Representative experimental results are presented for demonstration.

Contents

* Corresponding author. Tel.: +44 (0)151 231 2385; fax: +44 (0)151 231 2158.
E-mail address: g.zhang@ljmu.ac.uk (G.-M. Zhang).

0041-624X/$ - see front matter
doi:10.1016/j.ultras.2011.10.001

1. Introduction

Ultrasonic echoes reflected from homogeneities or discontinuities in tested materials contain information pertaining to the location, size, and characteristics of defects, along with the material and geometry of the sample under evaluation. The accurate detection, location and sizing of the defects are limited by the ability to precisely estimate the information contained in the ultrasonic signals obtained during an inspection. Signal processing plays a crucial role in ultrasonic non-destructive evaluation (NDE).

In ultrasonic NDE, the received signals are often contaminated by noise originating from both the measurement system and test sample. In many applications, the noise is generally assumed to be an uncorrelated Gaussian random variable, with zero mean and a band limited power spectral density function [1]. However, for highly-scattering materials, there is another type of noise called grain noise, which is produced by the microstructures/grain boundaries of the inspected materials. Each grain behaves like a scattering centre, producing an echo that is isolated or superimposed with other echoes coming from other grains which can hide the echoes produced by a possible defect. Grain noise will exhibit a small-scale correlation and can be erroneously associated to defects. Noise places a fundamental limit on the detection of small defects and the accuracy of measurement. Various signal processing techniques have been developed for noise reduction and enhancement of detected echoes. Since the ultrasonic signal is usually a time and frequency limited transient signal in ultrasonic flaw detection, time–frequency analysis techniques, for example wavelet transform, are more appropriate to process ultrasonic signals.

Moreover, a small defect is sometimes masked by the echo from a larger nearby reflector, and it is difficult to detect, particularly for ultrasonic inspection of layered structures. When the layer thickness is less than or comparable to the wavelength of ultrasound, the reflected echoes from the front and the back surface of the layer overlap. The traditional way to meet the challenge is to increase the transducer frequency. However, in practical applications it is often not possible because high frequencies have high attenuation, resulting in low penetration. To discover the internal structure of the test sample as precisely as possible, various signal processing techniques for example deconvolution [2] have been explored to improve the resolution without increasing the transducer frequency (thus without sacrificing the penetration).

n NDE industry, the guided wave ultrasound is quite common for defect detection and structural health monitoring (SHM). For example, the guided Lamb waves are widely used for quantitative identification of damage in composite structures [3]. For long range inspection of pipes, guided longitudinal and torsional modes are often used [4]. Various transducers such as angled piezoelectric wedge transducers, comb transducers, and electromagnetic acoustic transducers and phased array have been used to generate guided waves. Signal processing is a crucial aspect in any guided wave ultrasound to extract information about damage type and severity from the measured signal. A signal-processing technique should be able to isolate from the measured signal the time and frequency centers associated with scattered waves from the damage and identify their modes [5]. The signal-processing approach should also be robust to noise in the guided wave signals.

Laser ultrasonics is another popular technique in NDE industry. For example, laser generated multimode Lamb waves have been widely used for defect detection and material characterization in composite structures and plates, where the dispersion curve is often used to indentify defects or characterize materials. Therefore, one key signal processing issue is to accurately resolve the dispersion relationships from the received Lamb waves. Time–frequency representations of Lamb waves have been used to analyse multimode Lamb waves effectively localize multiple closely spaced Lamb modes in both time and frequency [6].

Mathematically, for pulse-echo mode ultrasonic testing, a signal $y(t)$ recorded from materials is a linear combination of echoes $s_i(t)$ reflected from different interfaces in the sample being examined, i.e., $y(t) = \sum_{i=1}^{m} s_i(t) + \xi(t)$, where $\xi(t)$ accounts for the noise originating from the measurement system and materials. The recorded signal is usually simplified as a convolution form by assuming that the reflected echoes are the time-shifted, amplitude-scaled replicas of a reference echo,

$$y(t) = \phi(t) * \left\{ \sum_{i=1}^{m} c_i \delta(t - \tau_i) \right\} + \xi, \tag{1}$$

where the bracketed term denotes the reflectivity function, and $\phi(t)$ is the impulse response. Notice that this model by no means represents all the measurements in NDE, for example, this model does not hold for flaw detection from metals that present microstructure scattering. Rather it can only represent echoes from close interfaces where frequency dependent attenuation is negligible, or echoes in homogeneous medium like water. In general, the ultrasonic pulse, as it propagates through the medium changes its shape. Although sometimes this change can be negligible, in many applications such as the guided wave ultrasound, a more generic model where each individual echoes may have unique shape, is required

$$y(t) = \sum_{i=1}^{m} c_i \phi_i(t) + \xi(t), \tag{2}$$

where $\phi_i(t)$ is the incident pulse impinged to the ith interface/defect. As a result, ultrasonic signal representation is formulated as the problem of *blind source separation* that separates a set of linear mixtures onto a number of unknown source signals, inferring both the reflectivity function c_i, which can be considered as the virtual ultrasonic sources, and the ultrasonic incident pulses ϕ_i from the observed signal y.

Ultrasonic NDE signals are often a class of very special signals comprised of limited echoes. Thus, the reflectivity function c_i has a sparse distribution, i.e., only a limited number of randomly located samples have non-zero values. The sparse constraint has been exploited in some conventional ultrasonic signal processing techniques, for example sparse deconvolution (SD) and model-based echo estimation. Many sparse decovolution methods based on Bernoulli-Gaussian modelling and MAP estimate have been developed in ultrasonic NDE such as Iterated Window Maximization [7] and using genetic optimization [8]. These SD methods have mainly been limited to deconvolve ultrasonic NDE signals with time-invariant pulses. However, the ultrasonic incident pulse often changes considerably as it passes through the medium due to dispersive attenuation in many applications. In general, the exact nature of the changes is poorly known. Thus the recorded traces could be both sparse and time-varying. Adaptive deconvolution algorithms have been designed to cope with time-varying pulses [9]. The majority are autocorrelation based solutions that are incapable of estimating the phase of the pulse. These typically resort to a questionable minimum phase assumption. Bernoulli-Gaussian modelling and generalized maximum likelihood solutions have been shown to provide precise identification of nonminimum phase pulses and high-resolution deconvolution results [10]. The time-varying pulses usually require either a parametric model for the pulse variations [11,12] or some sort of time gating. In [11,12], minimum entropy deconvolution methods were developed to deconvolve ultrasonic pulse-echo signals acquired from attenuative multilayered media. In [13], a sparse deconvolution method based on the Iterated Window Maximization was proposed to deal with this kind of ultrasonic NDE signals where the pulse is slowly time-varying. A similar method [14,15] has been developed for sparse deconvolution of ultrasonic NDE signals with time-varying pulses. A more comprehensive review on sparse deconvolution can be found in [16]. In addition, there is a line of model-based echo estimation techniques that provide sparse representations of ultrasonic signals via parameter estimation. These techniques represent an ultrasonic signal in terms of parametric models (such as real Gabor function in [17], Gaussian chirplet in [18]), and have been used for ultrasonic echo estimation and data compression. Notice that these conventional techniques specifically developed for ultrasonic NDE signals, are not sparse overcomplete representations as which will be reviewed in this paper.

Obviously, given a recorded signal y, if the accurate signal representation of the generic model (2) can be obtained, the challenging ultrasonic signal processing issues above-mentioned can be better-solved. Sparse signal representation (SSR) is an emerging technique that computes a signal approximation in the model (2) by exploiting the sparse constraint. The SSR technique is summarized as follows: *Given a signal $y \in R^N$ and an overcomplete dictionary D we seek a sparse representation of y*. An overcomplete dictionary is a collection of elementary signals called atoms such that the number of atoms exceeds the dimension of the signal space. Assume that D consists of L atoms and is overcomplete, that is, $D = \{\phi_i\}_{i=1}^{L}$ and $N < L$, and that the atoms are also N-dimensional and have unit norm, that is, $\phi_i \in D$, $\phi_i \in R^N$ and $\|\phi_i\|_2 = 1$. The SSR technique is to seek a sparse vector $r \in R^L$ satisfying the relationship:

$$y = Dr + \varepsilon, \tag{3}$$

ϵ is an error term.

Comparing Eqs. (2) and (3), it can be seen that the introduction of SSR gave rise to the ability to address the signal representation problem as a direct sparse decomposition technique over overcomplete dictionaries. Although under an overcomplete dictionary the decomposition of a signal is underdetermined, recent research shows that in many applications this can offer great advantages compared to the conventional signal processing methods. One is that there is greater flexibility in capturing structure in the data [19]. Instead of a small set of general basis vectors, there is a larger set of more specialized atoms such that relatively few are required to represent any particular signal. These can form more compact representations, because each atom can describe a significant amount of structure in the data. The second is super-resolution [20]. We can obtain a resolution of sparse objects that is much higher than that possible with traditional methods. An additional advantage is that overcomplete representations increase stability of the representation in response to small perturbations of the signal. In addition, the redundant representations have the desired shift invariance property [21].

These advantages of SSR are of benefit to the interpretation of an ultrasonic NDE signal for the above-mentioned challenging issues such as pulse detection and noise suppression, echo separation and echo estimation, resolution improvement, and multi-mode guided wave separation. For example, if the dictionary atoms are designed to match the ultrasonic flaw echoes well, $\sum_{i=1}^{M} x_i(t)$ will be packed into very few significant coefficients while the noise energy will spread out around the whole time–frequency domain due to its mismatch with dictionary atoms. This will greatly facilitate the application of denoising methods, for instance wavelet thresholding [22].

The sparse vector r in the model (2) is computed by solving:

$$\min \|r\|_0 \text{ subject to } y = Dr \tag{4a}$$

or

$$\min \|r\|_0 \text{ subject to } \|y - Dr\|_2 \leqslant \varepsilon, \tag{4b}$$

where r is the sparse representation of y, and $\|r\|_0$ is the l_0-norm of r, defined to be the number of nonzero entries of r.

Solving the sparse representation problem (4) is NP-hard. This has led to considerable effort being put into the development of many sub-optimal strategies. Commonly used strategies are often based on greedy approximations, l_1 and l_p relaxations where $\|r\|_0$ is replaced by $\|r\|_1$ and $\|r\|_p$ for $p < 1$, and iterative shrinkage. For greedy approximations, many algorithms have been devised over the years, including matching pursuit (MP) [23], orthogonal matching pursuit (OMP) [24], gradient pursuit [25], subspace pursuit [26], and their variants [27–29]. The greedy approaches are often relatively fast, and have therefore been used extensively in practical applications. Under certain conditions these sub-optimal methods are guaranteed to optimize the problem (4) [30]. The l_1 relaxation is used in algorithms such as basis pursuit (BP) [20], greedy basis pursuit (GBP) [31], the least absolute shrinkage and selection operator (LASSO) [32], and other l_1 optimization approaches [33,34]. l_p relaxation is used in algorithms such as the focal underdetermined system solver (FOCUSS) [35] and Bayesian approaches such as sparse Bayesian learning (SBL) [36]. For the iterative shrinkage strategy, an initial set of coefficients is obtained and then iteratively modified with shrinkage to achieve sparse representation. The algorithms such as iterative hard thresholding algorithm [37], iterative projection-based noise shaping [38], and local competition [39] fall into this category. The iterative thresholding algorithms operate directly on the l_0 regularized cost functions. For the iterative thresholding algorithms, the non-convexity of the optimization problem means that these algorithms are only guaranteed to find local minima and good initialization is of paramount importance. The main advantage is the computational efficiency. These algorithms can be used to improve the results calculated with other methods such as MP [37] and dual-tree discrete wavelet transform [38].

2. Greedy approximations

Greedy approximations make a sequence of locally optimal choices in an effort to determine a globally optimal solution by approximating the observed signal y iteratively.

2.1. Matching pursuit

MP is currently the most popular algorithm for computing sparse signal representations using an overcomplete dictionary

MP computes a signal representation by greedily constructing successive approximations to the signal, $y^{(0)}, y^{(1)}, y^{(2)}, \ldots$, by orthogonal projections on atoms of D. MP begins by setting an initial approximation $y^{(0)} = 0$ and residue $R^{(0)} = y$. At stage k, it identifies the dictionary atom that best correlates with the residue by solving an optimization problem

$$\phi_k = \arg\max_{\phi \in D} \|R^{(k-1)}, \phi\|. \tag{5}$$

Then it calculates a new approximation and a new reside by $y^{(k)} = y^{(k-1)} + r_k\phi_k$ and $R^{(k)} = y - y^{(k)}$, where r_k is computed by the L^2-inner product of $R^{(k-1)}y$ and ϕ_k, i.e., $r_k = < R^{(k-1)}, \phi_k >$. After m iterations, a matching pursuit decomposes the signal y into $y = \sum_{i=1}^{m} r_i\phi_i + R^{(m)}$.

When the dictionary is an orthogonal basis, the approximation $y^{(m)}$ is always an optimal m-term representation of the signal. For general dictionaries, the norm of the residual converges exponentially when the signal space is finite dimensional [24].

The atom ϕ_i selected at each iteration by the MP algorithm is not in general orthogonal to the previously selected atoms $\{\phi_i\}_{0 \leqslant i < k}$. The approximation derived from the matching pursuit can be refined by orthogonalizing the directions of projection. OMP adds a least-square minimization to each step of MP to obtain the best approximation over the atoms that have already been chosen. This revision significantly improves the behavior of the algorithm. However, for m iterations, MP is m-times faster than OMP. For small m, which is often the case in ultrasonic signal processing applications, MP and OMP give roughly equivalent approximations [24].

2.2. Support matching pursuit

Support matching pursuit is a variant of MP, proposed recently for resolving the individual echoes that compose an ultrasonic signal [40]. SMP based on a dictionary of Gabor functions that serve as prototypes for ultrasonic echoes, searches for sparse approximations in which each ultrasonic echo is approximated by a single prototype atom possessing a clear physical interpretation. SMP operates iteratively, improving the approximation at each iteration by adding a single atom to the solution set compromising between reducing the energy of the resulting residue $R^{(k)}$ and its "robust support" defined as

$$||\boldsymbol{R}^{(\boldsymbol{k})}||_{\boldsymbol{q}} = \sum_{i=1}^{\boldsymbol{N}} |\boldsymbol{R}^{(\boldsymbol{k})}(\boldsymbol{t_i})|^{\boldsymbol{q}}, \text{ for } \boldsymbol{0} \ll \boldsymbol{q} \ll \boldsymbol{1}, \tag{6}$$

which for $q \approx \mathbf{0}$, gives the number of t_i's with $\boldsymbol{R}(^{\boldsymbol{k}})(\boldsymbol{t_i}) \neq \mathbf{0}$ that coincide with the theoretical support [41]. SMP is initialized the same way as MP by setting an initial approximation $y^{(0)} = 0$ and residue $R^{(0)} = y$. At iteration k, firstly calculate the correlation coefficients $ck = D^T R^{(k-1)}$ between the dictionary D and the current residue $R^{(k-1)}$; then select all atoms corresponding to an absolute correlation coefficient above a predefined threshold $\boldsymbol{T_k}$:$\boldsymbol{_k} = \{\boldsymbol{p}:|\boldsymbol{c_k}(\boldsymbol{p})| > \boldsymbol{T_k}\}$ where $\boldsymbol{c_k}(\boldsymbol{p})$ is the p element of the vector c_k. Choosing an atom from this candidate set ensures a significant reduction in $||\boldsymbol{R}(^{\boldsymbol{k}})||_{\mathbf{2}}$. The next residual signal for each atom in P_k is calculated as $\boldsymbol{R}_{\boldsymbol{p}}^{(\boldsymbol{k})} = \boldsymbol{R}^{(\boldsymbol{k}-1} - \boldsymbol{c_k}(\boldsymbol{p})\boldsymbol{d_p}$, where $\boldsymbol{p} \in \boldsymbol{P_k}$ and dp is the pth column of dictionary D. Then choose the atom resulting in a residue with lowest "Robust support": $\hat{\boldsymbol{p}}(\boldsymbol{k}) = min_{\boldsymbol{p}}||\boldsymbol{R}_{\boldsymbol{p}}^{(\boldsymbol{k})}||_{\boldsymbol{q}} = min_{\boldsymbol{p}}\left\{\sum_{i=1}^{N}|\boldsymbol{R}_{\boldsymbol{p}}^{(\boldsymbol{k})}(\boldsymbol{t_i})|^{\boldsymbol{q}}\right\}$. Finally update the residue as $\boldsymbol{R}^{(\boldsymbol{k})} = \boldsymbol{R}_{\hat{\boldsymbol{p}}(\boldsymbol{k})}^{(\boldsymbol{k})}$. Continue the iteration until the residue energy $||\boldsymbol{R}(^{\boldsymbol{k}})||_{\mathbf{2}} < \varepsilon$.

2.3. Other variants

Many variants of MP have been developed in the literature. Some effort aims at improving the computation efficiency: in [20], an approach based on adaptive biorthogonalisation techniques is proposed to improve the computation efficiency of OMP. In [29], Gradient Pursuit is developed that approximate OMP with computational requirements more akin to MP. Other effort is to adapt the MP to a class of particular signals. For instance, Interpolation Matching Pursuit (IMP) [42] is a variant of MP, which is specifically designed for the analysis of complex ultrasonic signals so that overlapped echoes can be extracted accurately and efficiently based upon a dictionary of small size.

2.4. Applications in ultrasonic NDE

MP and its variants have found increasing interest in the ultrasonic NDE community in the last decade. The first application area of MP is on ultrasonic flaw detection and noise suppression [42–46]. As the ultrasonic signal is usually a time- and frequency-limited transient signal in ultrasonic flaw detection, time–frequency analysis techniques are more appropriate to process ultrasonic NDE signals. Wavelet transform has been widely used for ultrasonic flaw detection, leading to a state-of-the-art denoising performance. Due to the advantages of SSR mentioned in Section 1, MP has potential of outperforming the wavelet transform. In [43], a fast and efficient MP-based algorithm was proposed, which is very efficient eliminating noise and increasing the visibility of ultrasonic flaw signals. The algorithm utilizes time-shifted Morlet functions as dictionary elements because they are well matched with the ultrasonic pulse echoes obtained from the transducer used in the experiments. Numerical results showed SNR (signal-to-noise ratio) improvements of about 19 dBs with low computational cost. It has been possible to enhance flaw echoes in highly scattering materials ($SNR_{in} < 0$ dB), even when two adjacent flaw echoes appear in the same ultrasonic signal. In [45], an artificial fish swarm optimized MP was proposed to detect weak signals from noisy ultrasonic NDE signals. In [46], matching-pursuit with wavelet-packet dictionaries was proposed for improvement of ultrasonic flaw detection during ultrasonic NDE.

The second application area of MP is for ultrasonic echo separation and echo estimation, particularly in processing dispersive waves [47–50]. The success of the guided-wave damage inspection technology depends not only on the generation and measurement of desired waveforms but also on the signal processing of the measured waves, but less attention has been paid to the latter. In [47], a two-stage matching pursuit approach based on a Gabor dictionary was proposed, which consists of the following: rough approximations with a set of predetermined parameters characterizing the Gabor pulse, and fine adjustments of the parameters by optimization. The parameters estimated from measured longitudinal elastic waves are then directly used to assess not only the location but also the size of a crack in a rod. In [48], a correlation filtering-based matching pursuit over an orthogonal wavelet dictionary was applied to Lamb wave signals acquired from delaminated carbon-fibre/epoxy composite materials for identifying the location and size of delamination. In [49], a high-resolution match pursuit was proposed to detect flaw echoes close to the material surface. In [50], a chirplet matching pursuit based method was developed for guided wave pulse echo-based structural health monitoring. The research results as shown in Fig. 1 showed that the chirplet matching pursuit can separate overlapping multimodal reflections and estimate the time–frequency centres, the modes and individual energy, which are used to locate and characterize defects. Fig. 1a shows the difference signal of guided waves obtained from an aluminium plate structure between the pristine and damaged cases. From Fig. 1b, it can be seen that the spectrogram is incapable of resolving the overlapping S_0 mode reflections from two defects C_1 and C_2 while as shown in Fig. 1c, the chirplet matching pursuit resolves the overlapping S_0 mode reflections as well as the S_0 and A_0 mode reflections from the boundary.

The third application area is for super-resolution ultrasonic imaging. In the ultrasonic testing of layered structures such as modern microelectronic packages and composite materials, one of the key challenges is axial resolution for delamination and cracks at closely-spaced interfaces [51]. Taking the advantages of accurate echo separation and echo estimation, the sparse signal representation techniques have been integrated into a conventional acoustic micro imaging (AMI) system, resulting in super-resolution imaging methods in recent research [52–54]. In [54], holographic-like three-dimensional acoustic imaging was developed for micro-NDE of microelectronics. It was implemented by stacking all the interface slices together to locate and identify hidden defects. Matching pursuit based acoustic time–frequency domain imaging was proposed to overcome the wavelength limit of axial resolution so that ultra-thin slices are generated. Experimental results of imaging a stacked die package with delaminations are presented in Fig. 2 to demonstrate the performance of MP based AMI. Due to that the interface echo is interfered by the neighbouring echoes as seen in Fig. 2a so that the conventional AMI image is contaminated as shown in Fig. 2c. Fig. 2d shows the AMI image obtained using the MP based AMI technique. From Fig. 2, it can be observed that although the echoes overlap, MP based AMI still

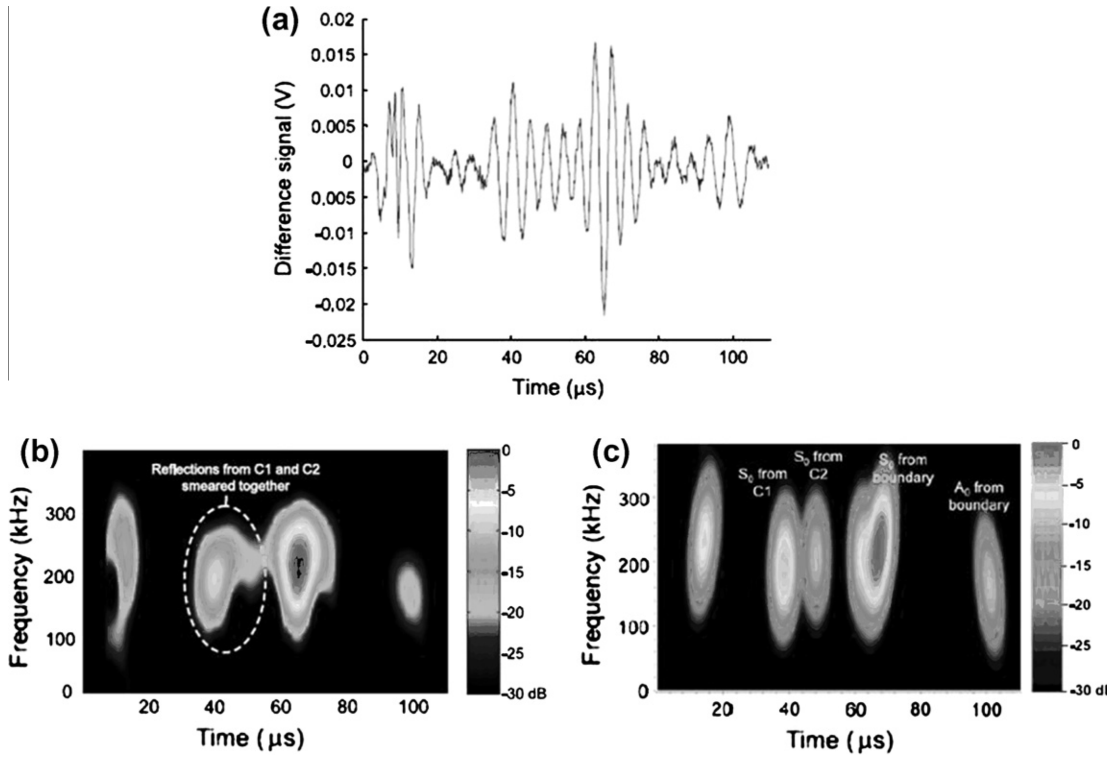

Fig. 1. Chirplet matching pursuits for guided wave based structural health monitoring. (a) Difference signal between pristine and 'damaged' states; (b) spectrogram of the signal in (a); (c) interference-free WVD of constituent chirplet atoms for the signal in (a). (From Fig. 9 of Ref. [50].)

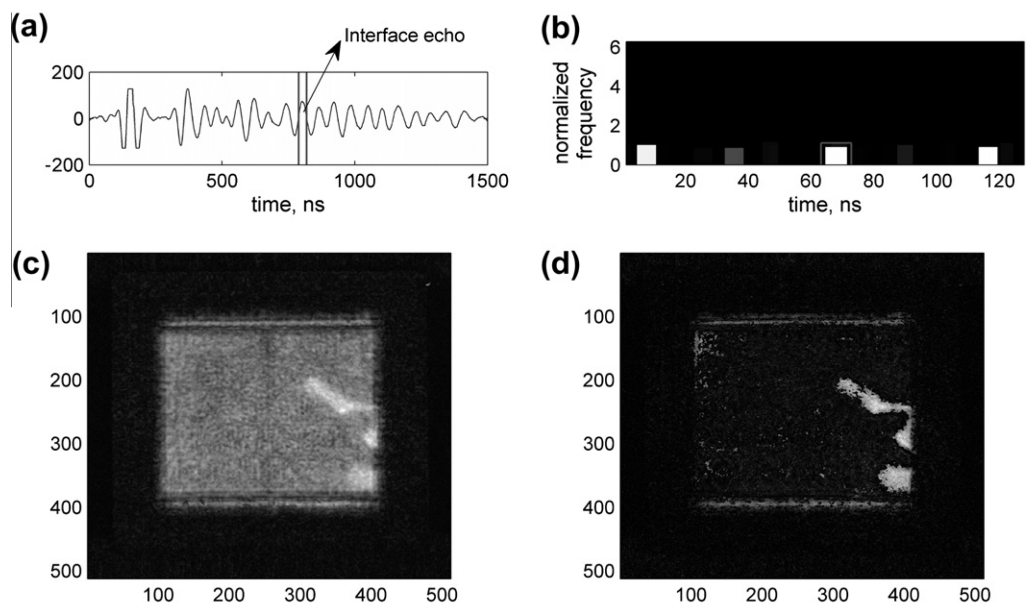

Fig. 2. Acoustic micro imaging of a stacked die package using a 50 MHz transducer. (a) An A-scan signal from the stacked die package; (b) the sparse representation of an ultrasonic signal segment obtained by the MP method; (c) conventional AMI; (d) MP based AMI. The white areas are delaminations.

generates a very sharp and clear image about the defects previously only thought possible with higher frequency transducers such as 230 MHz.

The fourth application area is for ultrasonic echo/parameter estimation [40,42,55–58]. The estimated echoes or the parameters are then used for subsequent analysis, for example time delay estimation. In [40], SMP was developed to achieve accurate approximations of measured ultrasonic echoes. The echoes are represented by a set of intuitive parameters (amplitude, scale, arrive time, center frequency, and phase), which enable the estimation of physical parameters of the investigated sample. In [42], IMP was used in structural health monitoring applications to decompose ultrasonic signals to locate and identify discrete echoes embedded in complex signals. IMP consists of the selection of a coarse set of basis functions, the search method for finding the best matching basis function, and interpolation of the basis function parameters to achieve high resolution. It interprets, the matching basis functions as characteristic wavelets. Changes in parameters of these wavelets are related to

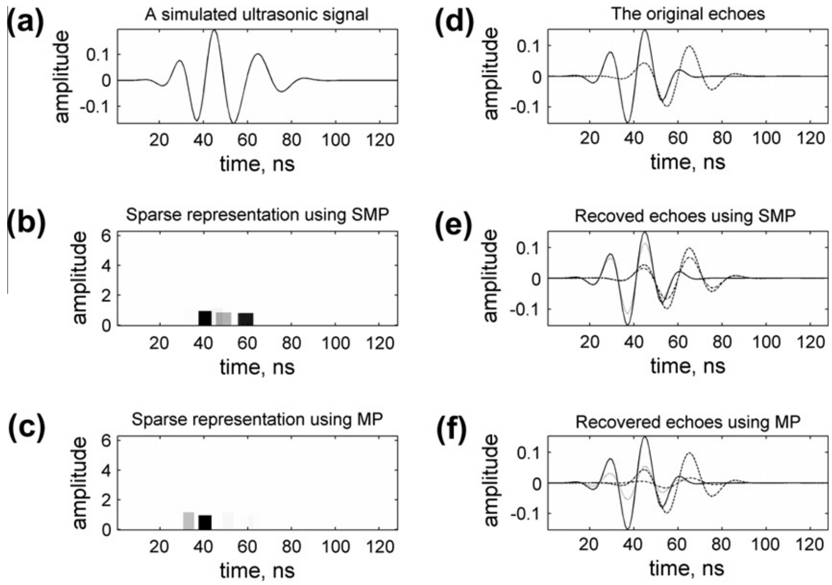

Fig. 3. Echo separation and estimation of a simulated ultrasonic signal with two echoes heavily overlapping using MP and SMP. The solid line: the original first echo; the dashed line: the original second echo; the dotted line: the estimated first echo; the dash-dot line: the estimated second echo.

changes in the structure. In [55], an echo windowing technique embedded with model-based echo estimation was developed to estimate the envelope and instantaneous phase. In [56], MP was used to design a matched filter.

A general drawback of greedy approximations is its greediness. Because of its greediness, MP (or OMP) initially may select an atom that is not part of the optimal sparse representation; as a result, many of the subsequent atoms selected by MP simply compensate for the poor initial selection. This shortcoming motivated the development of BP and GBP. OMP never selects the same atom twice because the residue is orthogonal to the atoms that have been chosen. In consequence, the residues of OMP converge strongly to zero and that the number of iterations required for convergence is less than or equal to the dictionary dimension. However, echoes in an ultrasonic NDE signal may not be orthogonal.

It is worth noting that IMP and SMP are two very promising methods specifically tailored for the decomposition of ultrasonic signals. For IMP, a Gabor dictionary is obtained, which is small but sufficient by adaptively choosing the frequency and scale ranges of the dictionary based upon the spectrum of the signal to be decomposed. The resolution of the parameter space is significantly improved by defining the search in scale and frequency from a coarse grid to a fine grid, and then interpolating both the fine grid and the time shift parameter. SMP selects an atom at each iteration by reducing both the residue energy and the residue support in time. The atom selection criterion utilizes the time localization nature of ultrasonic echoes, which causes portions of a multi-echo ultrasonic signal to be composed mainly from a single echo. This leads to accurate approximations in which each echo is characterized by a set of physical parameters that represent the composing ultrasonic echoes. Fig. 3 presents a simulated result to demonstrate the performance improvement of SMP over MP. The ultrasonic signal shown in Fig. 3a is generated by superimposing two measure ultrasonic pulses so that we can adjust the space between the two pulses. The ultrasonic signal representations obtained using SMP and MP are presented respectively in Fig. 3b and c. The darkness of the time–frequency image increases with the energy value, and each time–frequency atom is represented by a Heisenberg. The first and second echoes in Fig. 3e and f are recovered using the coefficients located in the center position of the corresponding echo. From Fig. 3, it can be seen SMP outperforms MP, particularly when the echoes are heavily overlapped. Fig. 4 (from Fig. 6 of Ref. [40]) presents an comparison result of MP and SMP on a real ultrasonic signal measured from a sample composed of two disks of Ultem attached by an adhesive layer. Fig. 4a shows the MP approximation of a representative signal measured from the 0.3-mm adhesive thickness step. It is clear that MP approximated the measured signal fairly well, however this approximation is effected using three composing Gabor atoms as shown in Fig. 4b. Investigating the time arrivals of those atoms revealed that those atoms did not match any of the adhesive interfaces, indicating that they did not approximate any of the reflected echoes. Fig. 4c shows the approximation obtained by the SMP method, which indicates that the SMP method also yielded an accurate approximation of the measured signal. This approximation is effected using only two atoms (a sparser result than MP). Moreover, the difference in time arrival between the two composing atoms closely matched the 0.3 mm thickness of the adhesive layer, indicating that these atoms were good approximations to the reflected echoes.

3. l_1 and l_p *relaxations*

3.1. Basis pursuit

It has been shown recently that if the signal y is "sparse enough" with respect to a dictionary D, then the l_1 relaxation in fact leads to the exact solution of the sparse signal representation problem (4) [30]. BP is a l_1 relaxation method designed for problem (4a) by solving the problem

$$\min \|r\|_1 \text{ subject to } y = Dr. \tag{7}$$

This is a convex optimization problem, and the global optimum can be found for real-valued data. Chen et al. [20] describe two algorithms for BP, BP-simplex and BP-interior, which are the well-known simplex and interior-point methods of linear programming applied to signal representation. All the techniques are computationally intensive.

The BP decomposition can be adapted to the case of noisy signals, i.e., the problem (4b) by modifying the optimization (7) to the following unconstrained form:

$$\min \|y - Dr\|_2^2 + \lambda \|r\|_1. \tag{8}$$

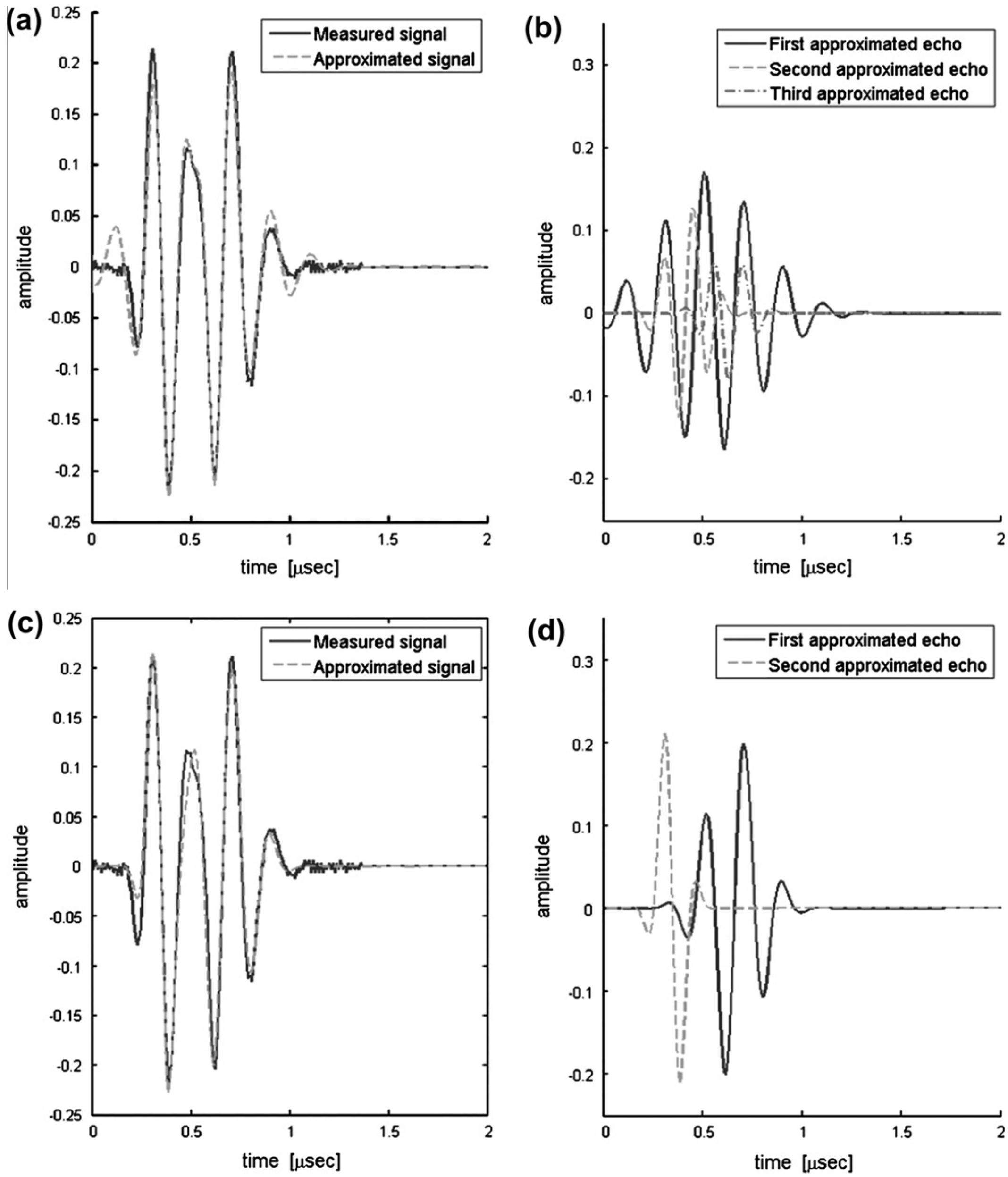

Fig. 4. An example of the approximation performed by MP and SMP on an ultrasonic signal retrieved from an adhesive layer of 0.3 mm thickness. The measured signal (solid line) along with the approximated signal (dashed line) are presented in (a) and (c) for the MP and SMP, respectively. The individual echoes, composing the approximated signal, are presented in (b) and (d) for MP and SMP, respectively. (From Fig. 6 of Ref. [40].)

The first term forces the residue $y - Dr$ to be small, and the l_1-term enforces sparsity of the representation. The parameter λ controls the tradeoff between the sparsity of the representation r and the residue norm. This optimization is again a convex optimization problem and can be readily handled by quadratic programming for real data. In [59], a second-order cone (SOC) programming method is used to solve the problem (8).

3.2. FOCUSS

FOCUSS is a class of methods for solving the sparse signal representation problem with a more general l_p relaxation instead of l_1 [35]. FOCUSS typically maintains a value of $p \in [0, 1)$. However, for $p < 1$, the cost function is non-convex, and the convergence to global minima is not guaranteed. The discussion in the Section VI of Ref. [60] indicates that the best results are obtained for p close to 1, whereas the convergence is also slowest for $p \approx 1$.

3.3. Sparse Baysian learning

Another sparse signal representation method with l_p relaxation is SBL [36]. In contrast to FOCUSS, which assumes a fixed prior p, SBL estimates a flexibly parameterized prior from the signal y. The parametric form of the prior is given by a multivariate Gaussian distribution, i.e., $p(r) = \prod_{i=1}^{L}(2\pi\alpha_i)^{-1/2}\exp(-r_i^2/2\alpha_i)$, where $\alpha = [\alpha_1, \ldots, \alpha_L]^T$ is a vector of L hyper-parameters controlling the prior variance of each r_i. These hyper-parameters are then estimated from the signal y.

3.4. Other variants

GBP [31] is the recent method for computing sparse representations. GBP is an algorithm for BP: it minimizes the l_1-norm of the representation r. However, unlike standard linear programming methods for BP, GBP proceeds much like MP, building up the

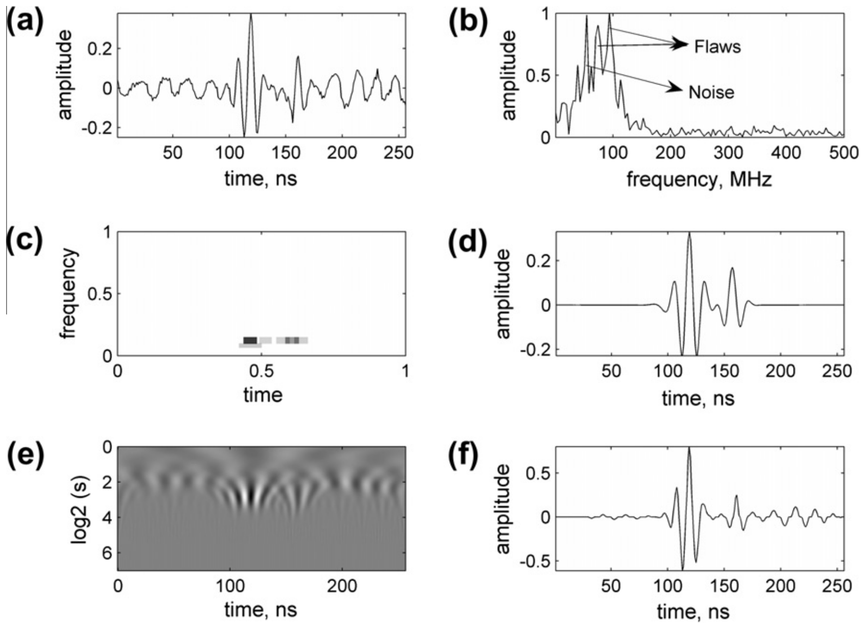

Fig. 5. Pulse detection and noise suppression using SBL and WTSP. (a) The noisy ultrasonic signal obtained with correlated noise; (b) its power spectrum; (c) the sparse representation of (a); (d) the noise reduction output using SBL with a customized Gabor dictionary; (e) the time–frequency representation obtained by CWT; (f) the WTSP output.

representation by iteratively selecting atoms. GBP selects an atom based on computational geometry by searching for the intersection between the signal and the convex hull of the dictionary. The GBP algorithm includes a step of discarding atoms that it has already selected at each stage. This is crucial, as it allows GBP to overcome the "mistakes" that MP can take in atom selection when compared to BP. While GBP returns the signal representation with the minimum l_1-norm, and thus GBP enjoys the theoretical benefits of BP, the greedy strategy of GBP leads to computational gains when compared to standard linear programming methods. In addition, a number of l_1 optimization approaches for sparse signal representation such as Gradient Projection [61] and LASSO [32] and its applications in blind source separation were developed [28]. Gradient projection algorithms are often significantly faster than competing methods, however their performance tends degrade as the regularization term is de-emphasized. According to our knowledge, none of these variants has been applied to ultrasonic NDE yet because sparse signal representation techniques have evolved very rapidly in the last decade although GBP has a potential to produce a good performance for ultrasonic NDE signal processing.

3.5. Applications in ultrasonic NDE

Basis Pursuit has been applied in ultrasonic NDE of highly-scattering materials for enhancing flaw detection [62]. However, a few drawbacks limit its applications. First, the computational cost of the basis pursuit. In particular, when decomposing an ultrasonic signal using BP over an overcomplete dictionary without fast algorithms such as a Gabor dictionary, the computational burden is very heavy. Second, the structure errors. In [36], the performance of BP, FOCUSS and SBL is discussed. BP guarantees convergence to the global minimum. However, the global minimum of the cost function does not necessarily coincide with the sparsest solution to problem (4). This misalignment is referred to as *structural errors*. In [63], a study was carried out to compare the MP, BP, OMP, and stagewise OMP algorithms over an overcomplete Gabor dictionary for ultrasonic imaging. Their results showed that stagewise OMP performs best overall in extracting the specific echo, since this algorithm is precise and fast. In [64], experimental results showed that MP with a Gabor dictionary outperforms the BP with wavelet dictionaries for echo separation and echo estimation.

Overcomplete ICA, FOCUSS and SBL, especially with learned overcomplete dictionaries are recent techniques. These new techniques are very promising techniques for suppressing both additive Gaussian noise and correlated noise, which have been demonstrated in [22]. Because of the high compact essence of sparse representations, the significant ultrasonic echoes are packed into a few significant coefficients, and noise energy is scattered all over the dictionary atoms, generating insignificant coefficients. Hence, the pruning operation and the thresholding operation [65] could be even more efficient than in the wavelet transform domain. Notice that the theresholding operation is optimal for white noise [66], but it can fail if the noise is correlated. Fortunately, unlike the wavelet transform where daughter wavelets are generated from the mother wavelet by dilation and shift operations, for the SSR techniques the dictionary atoms could be designed to match the significant ultrasonic echoes maximally and mismatch the correlated noise, relieving the impact of correlated noise on pruning and thresholding operations. Fig. 5 presents an example of denosing an ultrasonic signal with grain noise using SBL with a customized Gabor dictionary. Fig. 5d is obtained using the techniques proposed in Ref. [22]. The corresponding wavelet transform signal processor (WTSP) [65] processing result of the ultrasonic signal is presented as well in Fig. 5e and f for comparison. In WTSP, the ultrasonic signal is decomposed into the time–frequency domain by using the continuous wavelet transform (CWT), and *pruning and thresholding* operations are applied to the decomposed coefficients. The processed coefficients are then used to reconstruct the signal using the inverse CWT.

In [67], a study was carried out to compare the performance of overcomplete ICA, FOCUSS, and SBL for ultrasonic signal representations. An example result is presented in Fig. 6 (from Fig. 3 of Ref. [67]). The efficiency of ultrasonic signal representations were evaluated in terms of the different criteria that can be used to measure its performance for different applications, such as waveform estimation, echo detection, echo location and C-scan imaging. The results showed that the FOCUSS algorithm performs best overall. In [52,53], overcomplete ICA and SBL with learned overcomplete dictionaries were

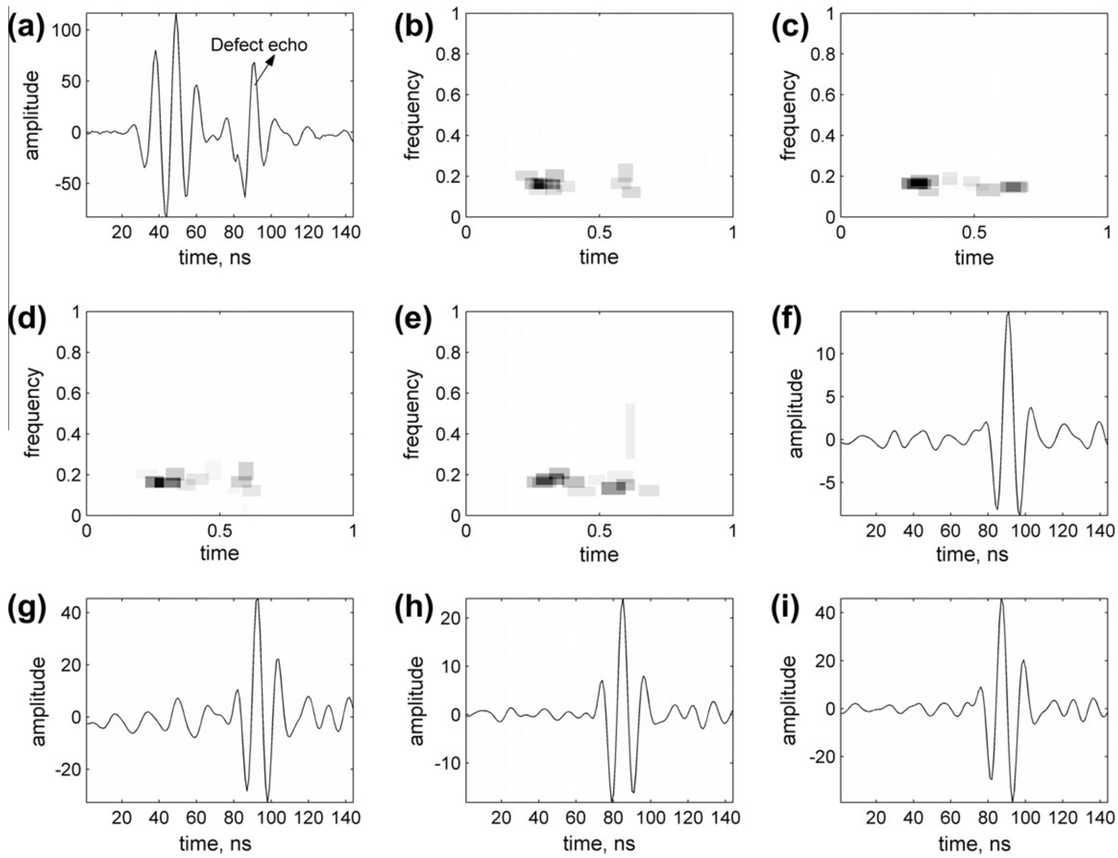

Fig. 6. Ultrasonic signal representations of a measured A-scan. (a) The test A-scan acquired from a flip-chip package using a 230 MHz transducer; (b) the sparse representation accomplished by the overcomplete ICA + dictionary 1; (c) by FOCUSS + dictionary 2; (d) by SBL + dictionary 1; (e) by SBL + dictionary 2; (f) the recovered defect echo from (b); (g) from (c); (h) from (d); (i) from (e). Dictionary 1: a 2*x*-overcomplete dictionary learned using the FOCUSS-CNDL algorithm; Dictionary 2: another 2*x*-overcomplete dictionary learned using the overcomplete ICA algorithm. (From Fig. 3 of Ref. [67].)

developed to improve the axial resolution of ultrasonic imaging. Experimental results showed the superior performance of sparse signal representation based ultrasonic imaging techniques compared with that of conventional time-domain ultrasonic imaging and frequency-domain ultrasonic imaging techniques.

Other applications such as sparse deconvolution have been attempted as well in the literature. Using learning overcomplete representations to deconvolve sparse ultrasonic NDE traces with time-varying pulses was attempted in [68]. In [69], a modified FOCUSS based sparse deconvolution method was developed, in which matching pursuit was used to estimate the ultrasonic pulse.

FOCUSS has many local minima. As $p \to 0$, the correlation between the global minimum of the cost function and the sparsest solution to the problem (4) approaches certainty since we are effectively now performing l_0-norm minimization. While there are no structural errors in this scenario, FOCUSS converges to suboptimal local minimum termed *convergence errors*. For SBL there are no structural errors with fewer convergence errors than FOCUSS in their experiments [36]. In addition, like BP, the FOCUSS and SBL have been adapted to handle both noiseless and noisy conditions.

The performance of the sparse representation algorithms mentioned above including MP is not guaranteed in general and only under certain conditions can they be shown to optimize the problem (4). Many theoretical studies [30,70–73] have been carried out to develop the conditions under which the problem of minimizing the l_0-norm of the representation coefficients of a signal in an overcomplete basis is equivalent to the problem of minimizing their l_1-norm and l_p-norm, or greedy approximations. These conditions depend on the behavior of the dictionary D, the spasity of the desired solution r, and the nature of the input signal y. If these conditions are not satisfied, the ultrasonic signal representation performance is unstable which has been observed in our previous work [52–54,64,67]. This drawback will have a big negative impact on the use of these algorithms to practical NDE applications. The first solution to tackle this issue is to design an appropriate dictionary which we will talk about in the next section. The second solution is to choose a suitable sparse signal representation algorithm according to the nature of the input signal. The impact of reflection coefficient prior distribution is studied experimentally in [53].

In addition, for practical applications the observed signals are usually noisy signals. In general, the problem of finding sparse representations is unstable in the presence of noise. It has been shown that stable recovery of sparse signal representations in the presence of noise under a combination of sufficient sparsity and favorable structure of the overcomplete system are possible [74,75]. This is observed in our previous research, for example, when the SNR of an ultrasonic NDE signal is extremely low, decomposing the noisy signal using SBL is not stable [22]. In that paper, although the two flaw echoes might still be detected, some noise might also be detected wrongly as flaw echoes. When SNR < −5 dB in our experiments, small echoes might be lost. The unreliable results are likely contributed by the following as well: One is the convergence errors existing in sparse decomposition using SBL. When

Table 1
Computation efficiency of various SSR algorithms.

Computation time (s)				
MP	SMP	Overcomplete ICA	FOCUSS	SBL
0.0156	0.8906	6.5625	0.8125	1.5000

SNR decreases the convergence errors increase. Indeed, this has been observed by Wipf and Rao [37]. Another reason is that in SBL, the noise level is estimated from the observed data inside the SBL algorithm, and very low SNR likely leads to a big estimation error. There are two possible remedies that could alleviate the problem. One is to use the BP denoising algorithm to replace SBL because there is no convergence error for BP. The dictionary design including the dictionary size is, however, vital for BP to avoid a big structural error as it determines the sparsity conditions under which minimizing the l_0-norm is equivalent to minimizing their l_1-norm. Second is to design a dictionary to maximally match the dictionary atoms and the ultrasonic echoes while mismatching the noise.

3.6. Computation efficiency

The computation complexity of each algorithm can be found from the corresponding references for example the complexity of MP and OMP is discussed in [24]. The complexity may be a factor in choice of the algorithm for commercial applications. In order to let the readers have an intuitive understanding, in Table 1, the computation complexity of a few SSR algorithms widely-used in ultrasonic NDE signal processing is presented experimentally by processing a measured ultrasonic signal with 144 samples in length using our Pentium 4 computer and MATLAB codes. The signal was measured from a stacked die packages using a 230 MHz transducer on a commercial AMI system. For MP and SMP, 64 iterations which are normally enough for most of the applications are used. A learned 2 × overcomplete dictionary used in Ref. [22] is used for all the processing.

In addition, the overcomplete ICA, FOCUSS, and SBL algorithms require large computer memory, especially when processing a signal with a length bigger than 256 samples. So they are not suitable to process a long signal.

4. Selection of overcomplete dictionary in ultrasonic NDE

It has been shown that the behavior of the overcomplete dictionary has a great impact on the performance [30] of the SSR methods. If the dictionary is orthonormal, the sparse signal representation admits a straightforward algorithm. Early computational techniques for sparse representation concentrated on specific dictionaries. For example the best orthogonal basis (BOB) algorithm calculates sparse representations over wavelet packet and cosine packet dictionaries. Although BOB frequently produces good results, it does not offer any guarantees on the quality of the sparse representation in terms of minimizing the l_0-norm of the representation. In [30], the coherent parameter and cumulative coherence function are introduced to quantify the behavior of a dictionary. The coherent parameter is defined as the maximum absolute inner product between two distinct atoms, i.e.,

$$\mu \stackrel{\text{def}}{=} \max_{j=k} |\langle \phi_j, \phi_k \rangle| \tag{9}$$

The cumulative coherence function measures the maximum total coherence between a fixed atom and a collection of other atoms. For a positive integer m, the cumulative coherence function is defined as:

$$\mu_1(m) \stackrel{\text{def}}{=} \max_{|\Lambda|=m} \max_{\varphi} \sum_{\Lambda} |\langle \varphi, \phi_i \rangle|, \tag{10}$$

where the φ ranges over the atoms indexed by Ω/Λ, and Ω is the index set of the dictionary. Informally, a dictionary is said to be incoherent when μ is small. When the cumulative coherence of a dictionary grows slowly, the dictionary is said to be quasi-incoherent. It argues that OMP is a good approximation algorithm for the sparse problem only over a quasi-incoherent dictionary [30,71].

4.1. Model based overcomplete dictionaries

There are several ways to design an overcomplete dictionary. Since a lot of overcomplete dictionaries have been developed such as Gabor dictionaries and wavelet packet dictionaries, one way is to tailor or choose one from the existing structured dictionaries. For ultrasonic NDE signals, Gabor dictionaries are often chosen because the ultrasonic signal is usually a broadband pulse modulated at the center frequency of the transducer, and is usually modelled as a Gabor function. For example, in [22] an overcomplete multiscale Gabor dictionary was customized for ultrasonic flaw detection and noise suppression. The dictionary consisted of several fixed scale critically sampled cosine Gabor bases. Each atom was defined by the parameters $\gamma = (s, u, v)$ as [22]:

$$g_\gamma = (A/\sqrt{s})\exp^{-\pi(t-u)^2/s^2} \cos(v(t-u)), \tag{11}$$

where s is the scale of the function, $s \in \{lms{:}ds{:}rms\}$; u its translation, $u \in \{0{:}du{:}N-1\}$; v its frequency modulation, $v \in \{lmv{:}dv{:}rmv\}$; and constant factor $A/\sqrt{s}$ normalize g_γ. The parameters $\gamma = (s, u, v)$ was discretized as follows: $ds = s_r/10$, $du = 2$, $dv = v_r/10$, $lms = 6ds$, $rms = s_r$, $lmv = 6dv$, and $rmv = v_r$, where the reference parameters $\gamma_r = (s_r, u_r, v_r)$ were obtained by approximating a reference ultrasonic echo using the Gabor model. The dictionary design including the dictionary size and the sampling scheme is very important as they affect the coherence and cumulative coherence of the dictionary, thus the quality of the sparse representations. The sampling scheme also determines the matching degree between the dictionary atoms and ultrasonic echoes. The smaller the discretization intervals, the better the atoms match the ultrasonic echoes. On the other hand, the smaller the discretization intervals, two atoms look more alike so that the dictionary may be more coherent. Moreover, the dictionary size affects the computation time and memory as well. The dictionary design should be customized for a particular application.

In [64], the 'standard' Gabor dictionary used in [23] was compared with various wavelet packet dictionaries and cosine packet dictionary, experimental results showed that the Gabor dictionary improve the SSR performance including the accuracy and robustness significantly for echo separation and echo estimation. In fact, in most of the applications mentioned in Section 2, a Gabor dictionary tailored to their specific applications is used.

In addition to the frequency modulated real Gaussian functions, complex Gabor functions [42], redundant morlet wavelets, and chirplet functions [76] are also used as the basis functions in ultrasonic NDE applications. We can also combine dictionaries to make a bigger, more expressive dictionary.

4.2. Learning an overcomplete dictionary

Although the Gabor dictionary was well-matched to ultrasonic NDE signals in the experiments, the performance is expected to be further improved using learned basis vectors that are adapted to the signal statistics of the inputs. The real ultrasonic echoes inevitably deviate from the idealized Gabor model in Eq. (11). The adapted basis vectors could represent the activities of incident pulses ϕ_i in Eq. (2) much closer than the existing model, and then the reflection coefficients c_i could be inferred more accurately. Therefore, another way to design a dictionary is to learn an adapted overcomplete dictionary for a class of specific signals. Using learned basis vectors instead of fixed basis vectors can further improve the quality of the sparse representation, because each atom can describe a significant amount of structure in the input signal. Several algorithms have been developed in the community of sparse signal representation to learn an overcomplete dictionary, such as overcomplete independent component analysis (ICA) [19], focal underdetermined system solver-based column normalized dictionary learning algorithm (FOCUSS-CNDL) [77], and K-SVD (singular value decomposition) [78].

4.2.1. Overcomplete ICA

The overcomplete ICA dictionary learning algorithm contains two steps: inferring the sparse representation r and learning the overcomplete dictionary D. Each iteration consists of the computation of $\hat{r}$, followed by a dictionary update.

A probabilistic approach to estimating r is based on finding the maximum *a posteriori* value of r:

$$\hat{r} = \max_r p(r|D,y) = \max_r p(y|D,r)p(r). \tag{12}$$

By assuming a Laplacian prior on r and that r_i are mutually independent, the representations r are estimated using a modified conjugate gradient optimization in the overcomplete ICA algorithm.

Learning the dictionary is done using learning algorithm based on the maximum likelihood (ML) principle with log-likelihood of the data, $\log p(y|D)$. This likelihood is approximated by fitting a multivariate Gaussian around r as follows

$$\log p(y|r) \approx \frac{N}{2}\log\frac{\lambda}{2\pi} + \frac{L}{2}\log(2\pi) + \log p(r) - \frac{\lambda}{2}(y - Dr)^2 - \frac{1}{2}\log\det H, \tag{13}$$

where $\lambda = 1/\sigma^2$, and H is the Hessian of the log posterior at r, $H = \lambda D^T D - \nabla\nabla \log p(r)$. The dictionary is updated by performing gradient ascent on the likelihood approximation of Eq. (13). The learning rule is

$$\Delta D \propto DD^T \frac{\partial}{\partial D}\log p(y|D) \approx -D\left(z\hat{r}^T + \mathrm{I}\right), \tag{14}$$

where $z(\hat{r}_i) = \partial \log p(\hat{r}_i)/\partial \hat{r}_i$ is called the score function, and I is the identity matrix. The pre-factor DD^T produces the natural gradient extension, which speeds convergence.

4.2.2. FOCUSS-CNDL

Similarly to the overcomplete ICA, each iteration of the FOCUSS-CNDL algorithm also includes two steps: a sparse basis selection step and a dictionary learning step. The sparse basis selection is done by FOCUSS, and the dictionary learning D-update step uses gradient descent. The algorithm is summarized as follows. Giving a set of training data $Y = (y_1, \ldots, y_k)$, for each vector y_k the corresponding sparse representation vector r_k is updated using one iteration of the FOCUSS algorithm:

$$\hat{r}_k = \Pi^{-1}(\hat{r}_k)\hat{D}^T\left(\lambda_k I + \hat{D}\Pi^{-1}(\hat{r}_k)\hat{D}^T\right)^{-1} y_k, \tag{15}$$

where $\Pi^{-1}(\hat{r}_k) = diag(|\hat{r}_k[i]|^{2-p})$ and $\lambda_k = \beta(\hat{r}_k)$ is the regulation parameter. After updating the K representation vectors r_k, $k = 1, \ldots, K$, the dictionary D is re-estimated by

$$\hat{D} = \hat{D} - \gamma(\delta\hat{D} - tr(\hat{D}^T\delta\hat{D})\hat{D}), \tag{16}$$

where $\delta\hat{D} = \hat{D}\sum_{\hat{r}\hat{r}} - \sum_{y\hat{r}}$, $\sum_{\hat{r}\hat{r}} = \frac{1}{K}\sum_{k=1}^{K}\hat{r}_k\hat{r}_k^T$, $\sum_{y\hat{r}} = \frac{1}{K}\sum_{k=1}^{K}y_k\hat{r}_k^T$, and $\gamma > 0$ is the learning rate parameter. After each update of the dictionary $\widehat{D}$, each basis vector ϕ_i is renormalized to $\phi_i = \phi_i/(\sqrt{L}||\phi_i||)$.

4.2.3. K-SVD

The K-SVD is the latest algorithm, generalizing the K-means to learn an overcomplete dictionary that best suits a set of given signals. This algorithm, starting by initializing the dictionary $D^{(0)} \in R^{N \times L}$ with l_2 normalized atoms, consists of two steps: inferring sparse representations and updating the dictionary. Giving a set of training data $Y = (y_1, \ldots, y_k)$, for each vector y_k any pursuit algorithms are used to compute the sparse representation vectors r_k. For each atom ϕ_m, $m = 1, \ldots, L$ in $\widehat{D}$, it is updated as follows: (1) Define the group of training data that use this atom, $\omega_m\{i|1 \ll i \ll K, r_k(i) \neq 0\}$; (2) Compute the overall representation error matrix, E_m, by $E_m = Y - \sum_{j\neq m}\phi_j r_j$; (3) Restrict E_m by choosing only the columns corresponding to ω_m, and obtain E_m^R; (4) Apply SVD decomposition $E_m^R = U\Delta V^T$. Choose the updated dictionary column ϕ_m to be the first column of U. Update the representation vector r_k to be the first column of V multiplied by $\Delta(1,1)$. We repeat the two steps until convergence.

4.2.4. Applications in ultrasonic NDE

In [52], the overcomplete ICA dictionary learning algorithm was used to learn a number of dictionaries with different degrees of overcompleteness. In [53], various dictionaries were learned using the FOCUSS-CNDL algorithms with different reflection coefficient prior distribution parameter p and different degrees of overcompleteness. For ultrasonic NDE, a universal overcomplete dictionary could be learned from the AMI signal samples that are acquired by different transducers and for different types of packages. However, the dictionary size should be very big in order to matching the dictionary atoms with various echoes. The dictionary performance can be improved by learning a unique overcomplete dictionary for each transducer and certain type of test samples. The details to learn a dictionary from a training data set consisted of A-scans can be found in [52,53]. These learned dictionaries were then used to investigate the performance and impact of different dictionaries on ultrasonic signal representations. The efficiency of ultrasonic signal representations was evaluated in terms of the different criteria that can be used to measure its performance for different ultrasonic NDE applications, such as waveform estimation, echo detection, echo location and C-scan imaging. The performance improvement using the learned dictionaries instead of Gabor dictionaries was studied in [52], and great impact of the overcomplete dictionary on the performance of SSR in ultrasonic NDE was observed. Experimental results in [53] showed that the FOCUSS-CNDL algorithm can further improve the performance over the overcomplete ICA dictionary learning algorithm. The effect of the overcompleteness degrees of learned dictionaries were investigated in [52,53] as well. Experimental results showed that there is an optimal dictionary size for a specific ultrasonic testing application.

5. Discussions

The advantages of SSR in ultrasonic NDE have been discussed in comparison to the conventional ultrasonic signal processing techniques in Sections 2 and 3. However, as a technology recently introduced to ultrasonic NDE, to realize its potential in practical applications, many challenging issues still need to be addressed in the future:

(1) *Resolution*. Although a resolution much higher than that possible with traditional methods can be achieved by the SSR techniques, in practical NDE applications for a chosen SSR algorithm with a given overcomplete dictionary for example MP, how close two ultrasonic echoes can be resolved consistently and accurately. Research has been attempted to address this issue [40,42,52,53,67]. Further research effort is still required.

(2) *Reliability and stability*. ALL the SSR algorithms are sub-optimal. Their performance is not guaranteed in general and only under certain conditions can they be shown to optimize the problem (4). Although many theoretical studies [30,70–73] have been carried out to develop the conditions under which the problem of minimizing the l_0-norm of the representation coefficients of a signal in an overcomplete dictionary is equivalent to the problem of minimizing their l_1-norm and l_p-norm, or greedy approximations. These conditions depend on the behavior of the dictionary D, the spasity of the desired solution r, and the nature of the input signal y. In practical applications, when we apply these SSR algorithms to ultrasonic signals, how to judge if these conditions are satisfied or not is a big challenge. If not, the decomposition may be unreliable and unstable which has been observed in our previous research. Reliability and stability are crucial in ultrasonic testing in order to avoid false alarm. For a given SSR algorithm, how to design a dictionary to satisfy these conditions is another challenge.

(3) *Dictionary customization*. The behavior of the dictionary has a great impact to the performance of SSR. The challenge is how to design or learn an optimal dictionary not only to avoid the reliability and stability issue but also fit in a practical ultrasonic NDE application well? For example, for a model-based overcomplete dictionary, how to optimally design the model, the dictionary size, and the sampling scheme?

(4) *Parameter determination*. Our experiments shows that for many SSR algorithms, their performance is sensitive to some parameters, for example, SMP is sensitive to the parameter q in Eq. (6) and FOCUSS is sensitive to p. For a given ultrasonic NDE applications, how to determine these parameters needs further research.

(5) Noise impact on the quality and robustness of SSR.

6. Conclusions

This paper presents a review on the up-to-date development of sparse signal representation techniques, which will benefit researchers and engineers in the ultrasonic NDE community. Their applications in ultrasonic NDE are summarized. The advantages of sparse signal representation techniques compared to conventional signal processing methods are analyzed. The limitations of various sparse representation algorithms in ultrasonic NDE applications are addressed. The challenging issues of applying sparse signal representation techniques to ultrasonic NDE are discussed.

Acknowledgements

This work has been supported by The Engineering and Physical Sciences Research Council (EPSRC) of UK. Thanks to anonymous reviewers' comments that helped us improve this paper.

References

[1] S.P. Neal, D.O. Thompson, The measurement and analysis of acoustic noise as a random variable in ultrasonic nondestructive evaluation, J. Acoust. Soc. Am. 86 (1989) 94–101.
[2] A. Yamani, M. Bettayeb, L. Ghouti, High-order spectra-based deconvolution of ultrasonic NDT signals for defect identification, Ultrasonics 35 (1997) 525–531.
[3] Z. Su, L. Ye, Y. Lu, Guided lamb waves for identification of damage in composite structures: a review, J. Sound Vib. 295 (2006) 753–780.
[4] J.L. Rose, Y. Cho, M.J. Avioli, Next generation guided wave health monitoring for long range inspection of pipes, J. Loss Prev. Process Ind. 22 (2009) 1010–1015.
[5] A. Raghavan, C.E.S. Cesnik, Review of guided wave structural health monitoring, Shock Vibr. Digest 39 (2007) 91–114.
[6] H. Kuttig, M. Niethammer, S. Hurlebaus, L.J. Jacobs, Model-based analysis of dispersion curves using chirplets, J. Acoust. Soc. Am. 119 (4) (2006) 2122–2130.
[7] K.F. Kaaresen, Evaluation and applications of the iterated window maximization method for sparse deconvolution, IEEE Trans. Signal Process. 46 (1998) 609–624.
[8] T. Olofsson, T. Stepinski, Maximum a posteriori deconvolution of sparse ultrasonic signals using genetic optimization, Ultrasonics 7 (1999) 423–432.
[9] B. Widrow, S.D. Stearns, Adaptive Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1985.
[10] J.M. Mendel, Maximum-Likelihood Deconvolution: A Journey into Model-based Signal Processing, Springer-Verlag, New York, 1990.
[11] T. Olofsson, T. Stepinski, Minimum entropy deconvolution of pulse-echo signals acquired from attenuative layered media, J. Acoust. Soc. Am. 109 (2001) 2831–2839.
[12] T. Olofsson, Semi-sparse deconvolution robust to uncertainties in the impulse response, Ultrasonics 42 (2004) 969–975.
[13] K.F. Kaaresen, E. Bolviken, Blind deconvolution of ultrasonic traces accounting for pulse variance, IEEE Trans. Ultrason. Ferroelect., Freq. Control 46 (1999) 564–573.
[14] R. Demirli, J. Saniie, Model-based estimation of ultrasonic echoes, Part I: analysis and algorithms, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 48 (3) (2001) 787–802.
[15] R. Demirli, J. Saniie, Model-based estimation of ultrasonic echoes, Part II: Nondestructive evaluation applications, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 48 (3) (2001) 803–811.
[16] G.-M. Zhang, D.M. Harvey, Contemporary ultrasonic signal processing approaches for nondestructive evaluation of multilayered structures, Nondestructive Testing and Evaluation, 2011, doi:10.1080/10589759.2011.577428.
[17] G. Cardoso, J. Saniie, Ultrasonic data compression via parameter estimation, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 52 (2) (2005) 313–325.
[18] Y. Lu, R. Demirli, G. Cardoso, J. Saniie, A successive parameter estimation algorithm for chirplet signal decomposition, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 53 (11) (2006) 2121–2131.
[19] M.S. Lewicki, T.J. Sejnowski, Learning overcomplete representations, Neural Comput. 12 (2000) 337–365.
[20] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, SIAM J. Sci. Comput. 20 (1999) 33–61.
[21] R. Coifman, D.L. Donoho, Translation Invariant De-noising, Wavelets and Statistics, Lecture Notes in Statistics, Springer-Verlag, New York, 1995, pp. 125–150.
[22] G.-M. Zhang, D.M. Harvey, D.R. Braden, Signal denoising and ultrasonic flaw detection via overcomplete and sparse representations, J. Acoust. Soc. Am. 124 (2008) 2963–2972.
[23] S. Mallat, Z. Zhang, Matching pursuits with time-frequency dictionaries', IEEE Trans. Signal Process. 41 (1993) 3397–3415.
[24] G. Davis, S. Mallat, M. Avellaneda, Greedy adaptive approximation, J. Constr. Approx. 13 (1997) 57–98.
[25] T. Blumensath, M.E. Davies, Gradient pursuit, IEEE Trans. Signal Process. 56 (2008) 2370–2382.
[26] D. Needell, J.A. Tropp, CoSaMP: iterative signal recovery from incomplete and inaccurate samples, Appl. Comput. Harmon. Anal. 26 (2009) 201–321.
[27] M. Andrle, L.R. Neira, A swapping-based refinement of orthogonal matching pursuit strategies, Signal Process. 86 (2006) 480–495.
[28] L. Daudet, Sparse and structured decompositions of signals with the molecular matching pursuit, IEEE Trans. Speech Audio Process. 14 (2006) 1808–1816.
[29] R. Gribonval, E. Bacry, Harmonic decompositions of audio signals with matching pursuit, IEEE Trans. Signal Process. 51 (2003) 101–111.
[30] J.A. Trop, Greed is good: algorithmic results for sparse approximation, IEEE Trans. Inf. Theory 50 (2004) 2231–2242.
[31] P.S. Huggins, Steven W. Zucker, Greedy basis pursuit, IEEE Trans. Signal Process. 55 (2007) 3760–3772.
[32] R. Tibshirani, Regression shrinkage and selection via the Lasso, J. Roy. Statist. Soc. Ser. B 58 (1996) 267–288.
[33] S.J. Wright, R.F. Nowak, M.A.T. Figueiredo, Sparse reconstruction by separable approximation, IEEE Trans. Signal Process. 57 (7) (2009) 2479–2493.
[34] Y. Li, S. Amari, A. Cichochi, et al., Underdetermined blind source separation based on sparse representation, IEEE Trans. Signal Process. 54 (2006) 423–437.
[35] I. Gorodnitsky, B.D. Rao, Sparse signal reconstruction from limited data using FOCUSS: a re-weighed minimum norm algorithm, IEEE Trans. Signal Process. 45 (1997) 600–616.
[36] D.P. Wipf, B.D. Rao, Sparse Bayesian learning for basis selection, IEEE Trans. Signal Process. 52 (2004) 2153–2164.
[37] T. Blumensath, M.E. Davies, Iterative thresholding for sparse approximations, J. Fourier Anal. Appl. 14 (2008) 629–654.
[38] T.H. Reeves, N.G. Kingsbury, Overcomplete image coding using iterative projection-based noise shaping, in: Proceedings of International Conference on Image Processing, New York, USA, 2002, pp. 597–600.
[39] S. Fischer, G. Cristobal, R. Redondo, Sparse overcomplete Gabor wavelet representation based on local competitions, IEEE Trans. Image Process. 15 (2) (2006) 265–272.
[40] E. Mor, A. Azoulay, M. Aladjem, A matching pursuit method for approximating overlapping ultrasonic echoes, IEEE Trans. Ultrason., Ferroelectr. Freq. Control 57 (9) (2010) 1996–2004.
[41] N. Hurley, S. Rickard, Comparing measures of sparsity, IEEE Trans. Inf. Theory 55 (10) (2009) 4723–4741.

[42] Y. Lu, J.E. Michaels, Numerical implementation of matching pursuit for the analysis of complex ultrasonic signals, IEEE Trans. Ultrason., Ferroelectr. Freq. Control 55 (1) (2008) 173–182.
[43] N. Ruiz-Reyes, P. Vera-Candeas, J. Curpian-Alonso, R. Mata-Campos, J.C. Cuevas-Martínez, New matching pursuit-based algorithm for SNR improvement in ultrasonic NDT, NDT&E Int. 38 (2005) 453–458.
[44] N. Ruiz, P. Vera, J. Curpian, et al., Matching pursuit-based signal processing method to improve ultrasonic flaw detection in NDT applications, Electron. Lett. 39 (2003) 413–414.
[45] A.L. Qi, H. W. Ma, T. Liu, A weak signal detection method based on artificial fish swarm optimized matching pursuit, CSIE, in: 2009 WRI World Congress on Computer Science and Information Engineering, vol. 6, 2009, pp. 185–189.
[46] G. Yang, Q. Zhang, P. Que, Matching pusrsuit based adaptive wavelet packet atomic decomposition applied in ultrasonic inspection, Russ. J. Nondestr. Test. 43 (1) (2007) 62–68.
[47] J.C. Hong, K.H. Sun, Y.Y. Kim, The matching pursuit approach based on the modulated Gaussian pulse for efficient guided-wave damage inspection, Smart Mater. Struct. 14 (2005) 548–560.
[48] F. Li, Z. Su, Y. Lin, G. Meng, A correlation filtering-based matching pursuit (CF-MP) for damage identification using lamb waves, Smart Mater. Struct. 15 (2006) 1585–1594.
[49] N. Ruiz, P. Vera, J. Curpian, et al., High-resolution pursuit for detecting flaw echoes close to the material surface in ultrasonic NDT, NDT&E Int. 39 (2006) 487–492.
[50] A. Raghavan, C. Cesnik, Guided-wave signal processing using chirplet matching pursuits and mode correlation for structural health monitoring, Smart Mater. Struct. 16 (2007) 355–366.
[51] R. Dias, R. Goruganthu, et al., Assembly analytical forum analytical tool roadmap white paper, in: International SEMATECH, 2005.
[52] G.-M. Zhang, D.M. Harvey, D.R. Braden, An improved acoustic microimaging technique with learning overcomplete representation, J. Acoust. Soc. Am. 118 (2005) 3706–3720.
[53] G.-M. Zhang, D.M. Harvey, D.R. Braden, Adaptive sparse representations of ultrasonic signals for acoustic microimaging, J. Acoust. Soc. Am. 120 (2006) 862–869.
[54] Guang-Ming Zhang, David M. Harvey, David R. Burton, Micro-nondestructive evaluation of microelectronics using three-dimensional acoustic imaging, Appl. Phys. Lett. 98(8) (2011) 094102–094102-3
[55] R. Demirli, J. Saniie, An efficient sparse signal decomposition technique for ultrasonic signal analysis using envelop and instantaneous phase, in: Proceedings of the IEEE International Ultrasonic Symposium, 2008, pp. 1503–1507.
[56] W. Liang, P. Que, H. Lei, L. Chen, Matching pursuit for decomposition and approximation of ultrasonic pulse-echo wavelet and its application in ultrasonic nondestructive evaluation, Rev. Sci. Instrum. 79 (7) (2008) 075105.
[57] G.-M. Zhang, D.M. Harvey, D.R. Braden, Microelectronic package characterisation using scanning acoustic microscopy, NDT&E Int. 40 (2007) 609–617.
[58] Q. Zhang, G. Yang, P. Que, Ultrasonic signals processing base on parameters estimation, Russ. J. Nondestr. Test. 45 (1) (2009) 61–66.
[59] D. Malioutov, M. Cetin, A.S. Willsky, A sparse signal reconstruction perspective for source localization with sensor arrays, IEEE Trans. Signal Process. 53 (2005) 3010–3022.
[60] B.D. Rao, K. Kreurz-Delgado, An affine scaling methodology for the best basis selection, IEEE Trans. Signal Process. 47 (1999) 187–200.
[61] M.A.T. Figueirredo, R.D. Nowak, S.J. Wright, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems, IEEE J. Select. Top. Signal Process. 1 (4) (2007) 586–597.
[62] G. Zhang, S. Zhang, Y. Wang, Application of adaptive time-frequency decomposition in ultrasonic NDE of highly-scattering materials, Ultrasonics 38 (2000) 961–964.
[63] R. Mohammadi, A. Mahloojifar, Resolution improvement of scanning acoustic microscopy using sparse signal representation, J. Signal Process. Syst. 54 (1–3) (2009) 15–24.
[64] G.-M. Zhang, D.M. Harvey, D.R. Braden, Advanced acoustic microimaging using sparse signal representation for the evaluation of microelectronic packages, IEEE Trans. Adv. Packing 29 (2) (2006) 271–283.
[65] A. Abbate, J. Koay, J. Frankel, S. Schroeder, P. Das, Signal detection and noise suppression using a wavelet transform signal processor: application to ultrasonic flaw detection, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 44 (1997) 14–26.
[66] D.L. Donoho, De-noising by soft thresholding, IEEE Trans. Inf. Theory 41 (1995) 613–627.
[67] G.-M. Zhang, D.M. Harvey, D.R. Braden, Effect of sparse basis selection on ultrasonic signal representation, Ultrasonics 45 (1–4) (2006) 82–91.
[68] G.-M. Zhang, D.M. Harvey, D.R. Braden, Sparse deconvolution of ultrasonic NDE traces – a preliminary study, in: IEEE International Ultrasonics Symposium, Beijing, China, 2008.
[69] L. Wei, Z. Huang, P. Que, Sparse deconvoltion method for improving the time-resolution of ultrasonic NDE signals, NDT&E Int. 42 (2009) 430–434.
[70] D.L. Donoho, M. Elad, Optimally sparse representation in general (nonorthogonal) dictionaries via l^1 minimization, PNAS 100 (5) (2003) 2197–2202.
[71] J.A. Tropp, Just relax: convex programming methods for identifying sparse signals in noise, IEEE Trans. Inf. Theory 52 (3) (2006) 1030–1049.
[72] M. Aharon, M. Elad, A.M. Bruckstein, On the uniqueness of overcomplete dictionaries, and a practical way to retrieve them, Linear Algebra Appl. 416 (2006) 48–67.
[73] D.M. Malioutov, A Sparse Signal Reconstruction Perspective for Source Localization With Sensor Arrays, Master Thesis, MIT, 2003.
[74] E.J. Candes, J.K. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Commun. Pur. Appl. Math. 59 (8) (2006) 1207–1223.
[75] D.L. Donoho, M. Elad, V.N. Temlyakov, Stable reconvery of sparse overcomplete representations in the presence of noise, IEEE Trans. Inf. Theory 52 (1) (2006) 6–18.
[76] J.C. Hong, K.H. Sun, Y.Y. Kim, Waveguide damage detection by the matching pursuit approach employing the dispersion-based chirp functions, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 53 (3) (2006) 592–605.
[77] K. Kreurz-Delgado, J.F. Murray, B.D. Rao, Dictionary learning algorithms for sparse representation, Neural Comput. 15 (2003) 349–396.
[78] M. Abaron, M. Elad, A. Bruckstein, K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation, IEEE Trans. Signal Process. 54 (2006) 4311–4314.