A SIMPLE DESIGN OF SPARSE SIGNAL REPRESENTATIONS USING OVERLAPPING FRAMES

Karl Skretting, John Håkon Husøy, and Sven Ole Aase

Stavanger University College, Department of Electrical and Computer Engineering
P. O. Box 2557 Ullandhaug, N-4091 Stavanger, Norway
Phone: +47 51 83 20 16, Fax: +47 51 83 17 50, E-mail: karl.skretting@tn.his.no

ABSTRACT

The use of *frames* and *matching pursuits* for signal representation are receiving increased attention due to their perceived potential in various signal processing applications. Good design algorithms for *block oriented* frames have recently been published. Viewing these block oriented frames as generalizations of block oriented transforms, it is natural to seek corresponding generalizations of critically sampled filter banks leading to *overlapping frames*. Here we show that a large class of overlapping frames can be decomposed into a critically sampled orthogonal filter bank, – that can be chosen prior to the design, and a block oriented frame. Based on this we show that excellently performing overlapping frames can be designed using the already established and simple theory for the design of block oriented frames.

1. INTRODUCTION

Block oriented transforms are widely used in signal processing. A signal, $\mathbf{x}$, is divided into L blocks of length N, where each block is represented by a column vector denoted by $\mathbf{x}_l$. The forward transform, which we for notational convenience denote by $\mathbf{T}^{-1}$, is used to compute the transform coefficients, denoted $\mathbf{y}_l$, for each block through what is commonly called the analysis equation, $\mathbf{y}_l = \mathbf{T}^{-1}\mathbf{x}_l$. The reconstructed signal vector is then given by the corresponding synthesis equation, where the tilde is used to indicate the possibility of approximated quantities,

$$\tilde{\mathbf{x}}_l = \mathbf{T}\,\tilde{\mathbf{y}}_l = \sum_{n=1}^{N} \tilde{y}_l(n)\mathbf{t}_n. \quad (1)$$

The *synthesis vectors*, denoted $\{\mathbf{t}_n\}_{n=1}^{N}$, are the columns of the matrix $\mathbf{T}$. In the case of common transforms, such as the Discrete Cosine Transform (DCT) and the Karhunen-Loeve Transform (KLT), these synthesis vectors form an orthogonal basis for $\mathbb{R}^N$. The reconstructed signal is built up as a linear combination of these synthesis vectors. Allowing for the possibility of having more than N terms in the linear combination of Equation 1, say K terms, the collection of K vectors will be denoted as $\{\mathbf{f}_k\}_{k=1}^{K}$. Interpreting these vectors, collectively referred to as a *frame*, as columns of an $N \times K$ matrix $\mathbf{F}$ we have a more general situation than that of Equation 1.

In many signal processing applications approximating (or representing) a signal vector through a linear combination of *a small number of* vectors selected from a predefined set of vectors is important, for example in compression, feature extraction and modeling. These are referred to as sparse representations, and the sparseness factor is defined by the ratio of the number of vectors used in a given expansion to the number of signal samples in the original vector. If the vector set forms an orthogonal basis the synthesis coefficients, $\{\tilde{y}_l(n)\}$, are computed using the analysis equation, whereas if the vector set is a frame, practical solutions employ *vector selection* algorithms such as Matching Pursuit (MP) [1], Orthogonal Matching Pursuit (OMP) [2], and Fast Orthogonal Matching Pursuit (FOMP) [3]. Even though the use of frames has been reported by several authors, the development of procedures for their design is still in its infancy. In fact most authors use ad-hoc frame designs. Early attempts at design of optimal frames, given a sparseness constraint, are given in [4, 5]. Note that so far we have talked only about *block oriented* sparse representations.

The first attempt at designing frames where the frame vectors are allowed to overlap, referred to as *overlapping frames*, is reported in [6]. While producing good results, this method for designing overlapping frames is both computationally and conceptually more demanding than block oriented frame design. Here we propose a simple and elegant strategy for designing overlapping frames while keeping the attractive conceptual and computational aspects of the design procedure for block oriented frames. This is obtained at no cost in performance. In this paper we present some necessary background on frame design before we focus on the new design algorithm. This is followed by some examples and a discussion.

2. FRAME DESIGN

If we use a frame, $\mathbf{F}$, of size $N \times K$, $K \geq N$, instead of a transform of size $N \times N$, the expansion corresponding to Equation 1 would be

$$\tilde{\mathbf{x}}_l = \mathbf{F}\,\mathbf{w}_l = \sum_{k=1}^{K} w_l(k)\mathbf{f}_k, \quad (2)$$

where we have replaced the transform coefficients by weights, denoted $\mathbf{w}_l$. The synthesis equation for several adjacent blocks can be written as $\tilde{\mathbf{x}} = \mathcal{F}\,\mathbf{w}$ or

$$\begin{bmatrix} \vdots \\ \tilde{\mathbf{x}}_l \\ \tilde{\mathbf{x}}_{l+1} \\ \tilde{\mathbf{x}}_{l+2} \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddots & & & \\ & \boxed{\mathbf{F}} & & \\ & & \boxed{\mathbf{F}} & \\ & & & \boxed{\mathbf{F}} \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ \mathbf{w}_l \\ \mathbf{w}_{l+1} \\ \mathbf{w}_{l+2} \\ \vdots \end{bmatrix}. \quad (3)$$

The block diagonal structure of the large frame, $\mathcal{F}$, is evident from Equation 3, values outside the boxes are zero.

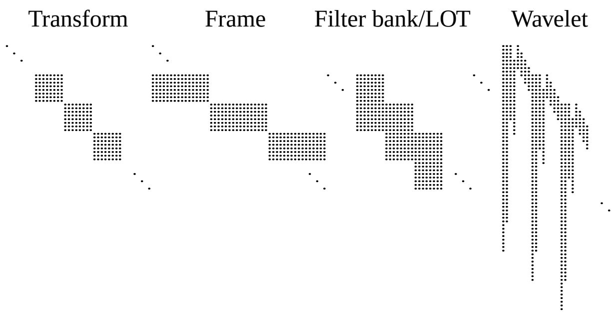

Fig. 1. The support structure of 4 different synthesis matrices $\mathcal{F}$. Each column of dots represent one synthesis vector.

The sparsity of the representation is expressed by the *sparseness factor*

$$S = \frac{\text{number of non-zero coefficients in } \mathbf{w}}{\text{number of samples in } \mathbf{x}}. \tag{4}$$

We point out that this is a global definition; for each block a larger or a smaller number of vectors than the average, SN, may be selected. The use of a global sparseness constraint makes the vector selection procedure more computationally demanding, but the benefit is that we get a better global approximation quality.

2.1. Block oriented frames

The optimal frame will depend on the target sparseness factor and the class of signals we want to represent. The problem of finding the optimal frame, $\mathbf{F}$, for a given class of signals and a given sparseness factor, represented by a training signal, $\mathbf{x}$, was treated in [5] and can briefly be summarized as follows: The training vectors, $\{\mathbf{x}_l\}$, are collected as columns into a matrix $\mathbf{X}$, and the weights, $\{\mathbf{w}_l\}$, into a matrix $\mathbf{W}$ in the same manner. The algorithm starts with a user supplied initial frame $\mathbf{F}^{(0)}$ and then improves it by iteratively repeating two main steps:

1. $\mathbf{W}^{(m)}$ is found by vector selection using frame $\mathbf{F}^{(m)}$,
2. $\mathbf{F}^{(m+1)}$ is found from $\mathbf{X}$ and $\mathbf{W}^{(m)}$.

m is the iteration number. The second step is optimal in the sense that it finds the $\mathbf{F}$ that minimizes the object function

$$J = J(\mathbf{F}) = \|\mathbf{x} - \tilde{\mathbf{x}}\|^2 = \|\mathbf{x} - \mathcal{F}\mathbf{w}\|^2 \tag{5}$$

when $\mathbf{W}$ is given. The norm used is $\|\mathbf{A}\|^2 = \text{trace}(\mathbf{A}^T\mathbf{A})$. More details on this algorithm can be found in [5].

2.2. Overlapping frames

In many applications, for example signal compression, feature extraction and noise suppression, it has been demonstrated that filter banks and wavelets perform better than transforms. This is the motivation for trying a synthesis system that uses overlapping vectors. The hypothesis is that the overlapping frame will better approximate a target vector than a block oriented frame when the sparseness of the expansion is fixed.

To obtain overlapping synthesis vectors the block diagonal structure of $\mathcal{F}$ in Equation 3 should be replaced by a band diagonal structure as in Equation 6.

$$\mathcal{F} = \begin{bmatrix} \ddots & \mathbf{F}_1 & & & \\ \ddots & \vdots & \mathbf{F}_1 & & \\ \ddots & \mathbf{F}_P & \vdots & \mathbf{F}_1 & \ddots \\ & & \mathbf{F}_P & \vdots & \ddots \\ & & & \mathbf{F}_P & \ddots \end{bmatrix}, \qquad \mathbf{F} = \begin{bmatrix} \mathbf{F}_1 \\ \vdots \\ \mathbf{F}_P \end{bmatrix}. \tag{6}$$

The synthesis vectors are the columns of $\mathcal{F}$ or $\mathbf{F}$. $\mathbf{F}$ can be partioned into P submatrices, $\{\mathbf{F}_p\}_{p=1}^P$ each of size $N \times K$. P is the overlap factor. The two rightmost portions of Figure 1 show the support structure of the synthesis matrix, $\mathcal{F}$, corresponding to a (critically sampled) uniform FIR filter bank / Lapped Orthogonal Transform (LOT) and a (critically sampled) wavelet type synthesis filter bank. The design problem for this situation, leading to a computationally and conceptually demanding algorithm, has been treated in [6].

In the next section we show that overlapping frames can be designed using the design theory developed for block oriented frames. The proposed method shows excellent results in spite of greatly simplified design and use. This method is inspired by the method used by Malvar & Staelin to design a signal adapted LOT [7].

3. THE PROPOSED METHOD

In arriving at the proposed method we posed the following question: Given a desired overlapping frame structure as shown in Equation 6, is it possible to decompose it into the product of one structure as in Equation 3 (block oriented frame) and another matrix? If so, could this other matrix be fixed, predefined prior to the design process, while the design effort is spent on the block structured part of the decomposition?

It is easily verified that setting $\mathcal{F} = \mathcal{G}\mathcal{H}$, that is

$$\begin{bmatrix} \ddots & & \\ & \boxed{\mathbf{F}} & \\ & & \boxed{\mathbf{F}} \\ & & & \ddots \end{bmatrix} = \begin{bmatrix} \ddots & & \\ & \boxed{\mathbf{G}} & \\ & & \boxed{\mathbf{G}} \\ & & & \ddots \end{bmatrix} \begin{bmatrix} \ddots & & \\ & \boxed{\mathbf{H}} & \\ & & \boxed{\mathbf{H}} \\ & & & \ddots \end{bmatrix} \tag{7}$$

with $\mathcal{G}$ as in Equation 6 (but with $\mathbf{G}$ of size $NP \times N$) and $\mathcal{H}$ as in Equation 3 ($\mathbf{H}$ of size $N \times K$) gives the overlapping frame $\mathcal{F}$ with the desired structure. Note that the structure of the first matrix, $\mathcal{G}$, corresponds to the synthesis matrix of a critically sampled FIR synthesis filter bank. The constituent matrices of $\mathcal{F}$, the $\mathbf{F}$ matrices, are each of size $NP \times K$ and defined by

$$\mathbf{F} = \mathbf{G}\mathbf{H} = \begin{bmatrix} \mathbf{G}_1 \\ \vdots \\ \mathbf{G}_P \end{bmatrix} \mathbf{H} = \begin{bmatrix} \mathbf{G}_1\mathbf{H} \\ \vdots \\ \mathbf{G}_P\mathbf{H} \end{bmatrix}. \tag{8}$$

The signal representation is now $\tilde{\mathbf{x}} = \mathcal{F}\,\mathbf{w} = \mathcal{G}\mathcal{H}\,\mathbf{w}$.

For a given class of signals, specified by a large vector $\mathbf{x}$ containing an appropriate training set of signal segments, the task of designing $\mathcal{F}$ can be divided into two parts: selecting a reasonable

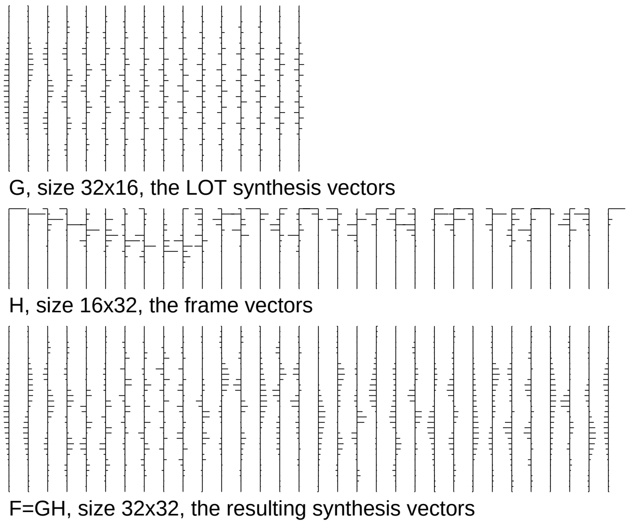

Fig. 2. The synthesis matrix $\mathbf{F}$ is made up by the matrix product of $\mathbf{G}$ and $\mathbf{H}$. We notice that each column of $\mathbf{F}$ is a linear combination of the columns of $\mathbf{G}$, the coefficients are given by the corresponding column of $\mathbf{H}$.

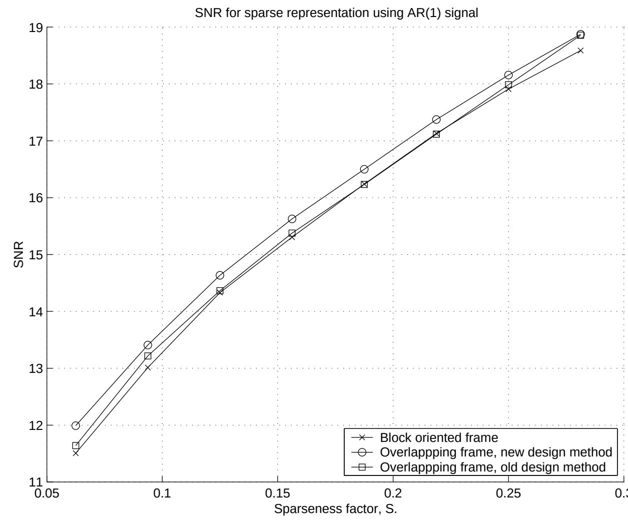

Fig. 3. Approximation quality, measured by Signal to Noise Ratio (SNR), plotted as a function of sparseness factor. Here we compare the results on an AR(1) signal for a block oriented frame (like $\mathbf{H}$ in Figure 2) designed with the method of Engan [5], ($\times$), the results of an overlapping frame (like $\mathbf{F}$ in Figure 2) designed with the method of Aase et al. [6] ($\Box$), and with the proposed method ($\circ$).

$\mathcal{G}$, which we then keep fixed, and finding a $\mathcal{H}$ (or equivalently its constituent matrices $\mathbf{H}$) using the method of Section 2.1. The object function for the second step in Section 2.1, Equation 5, will now be $J = J(\mathbf{H}) = \|\mathbf{x} - \mathcal{G}\mathcal{H}\mathbf{w}\|$.

Suppose that the columns of $\mathcal{G}$'s constituent matrices, $\mathbf{G}$, are chosen as the synthesis vectors (filter responses) of an *orthogonal* perfect reconstruction filter bank, then $\mathcal{G}^{-1} = \mathcal{G}^T$ and the norm is conserved, $\|\mathbf{x}\| = \|\mathcal{G}\mathbf{x}\| = \|\mathcal{G}^{-1}\mathbf{x}\|$. This implies that
$J = \|\mathbf{x} - \mathcal{G}\mathcal{H}\mathbf{w}\| = \|\mathcal{G}^{-1}(\mathbf{x} - \mathcal{G}\mathcal{H}\mathbf{w})\| = \|\mathcal{G}^T\mathbf{x} - \mathcal{H}\mathbf{w}\|$,
and we can design $\mathcal{H}$ in exactly the same manner as we design block oriented expansions in [5]. The only difference is that we use $(\mathcal{G}^T\mathbf{x})$ rather than $\mathbf{x}$ as the training signal. That is, we do the approximation in the coefficient domain rather than in the signal domain. In Figure 2 we illustrate the results of this process when $\mathbf{G}$ was selected as the synthesis vectors of a 32 tap 16 channel LOT taken from [7].

In practical use the most important improvement of this method compared with the method of [6], is when using the vector selection algorithm in the first step of Section 2.1. Vector selection algorithms are block oriented, and the complexity increases dramatically as the block size increases. In [6] we had to adapt the vector selection algorithm to the overlapping frame by using larger blocks. With the proposed method, we may use any block oriented vector selection algorithm directly, here we used FOMP [3]. The synthesis vectors (i.e. the columns) of the submatrices of $\mathcal{F}$, ($\mathbf{F}$), are still orthogonal to each other (as they are in $\mathcal{G}$). This is what makes it possible to use block wise vector selection algorithms.

The total synthesis system, specified by $\mathbf{F} = \mathbf{GH}$, has one fixed part and one part with free variables. In Figure 2, $\mathbf{F}$ has a total of 1024 parameters, but only the 512 in $\mathbf{H}$ are free variables. We may ask if this reduction in degrees of freedom is important. To answer this we compared the synthesis system designed using the proposed method, with a similar 32×32 overlapping frame having all variables free, designed using the method in [6]. We found that the overlapping frame designed with the present method performs best, see Figure 3. This illustrates that what we loose in degrees of freedom, for the tested case at least, is more than compensated for by what we win in the vector selection step. This indicate that just having more free variables in an optimization problem does not necessarily give a better solution if the optimization algorithm can not handle the increase of free variables in a good way.

We should also point out that the synthesis vectors we design are tied to the choice of orthogonal filter bank, $\mathcal{G}$. The columns of $\mathbf{F}$, the synthesis vectors of length PN, will be in the N dimensional subspace of $\mathbb{R}^{PN}$ spanned by the N columns of $\mathbf{G}$. To summarize: the main idea of this design method is that we may design an overlapping frame by selecting an appropriate orthogonal filter bank, $\mathbf{G}$, and design the frame $\mathbf{H}$ using established design procedures for block oriented frames.

4. DESIGN EXAMPLES AND DISCUSSION

The framework presented can be used to design many different synthesis systems, $\mathcal{F} = \mathcal{G}\mathcal{H}$. The structure will be decided by the choice of $\mathbf{G}$ ($\mathcal{G}$) and the size $(N \times K)$ of $\mathbf{H}$, where N depends on $\mathbf{G}$ but K may be selected freely. Given the structure, and the target sparseness factor, the values of $\mathbf{H}$ will be adjusted iteratively during the design. Since the design process converges towards a local optimum (if it converges at all), the values will depend on the initial frame, $\mathbf{F}^{(0)}$ or more precisely $\mathbf{H}^{(0)}$. We have chosen to present designs of 32 frames: 2 signal sources $\times$ 4 different $\mathcal{G}$s $\times$ 4 sparseness factors.

We used two different signal sources for training: a Gaussian AR(1) signal with $\rho = 0.95$ and an electrocardiogram (ECG) signal from the MIT arrhythmia database [8]: the MIT100 signal starting from the first sample. When we tested the designed filter banks in sparse representation experiments we used another AR(1) signal also with $\rho = 0.95$ and the MIT100 signal starting at sam-

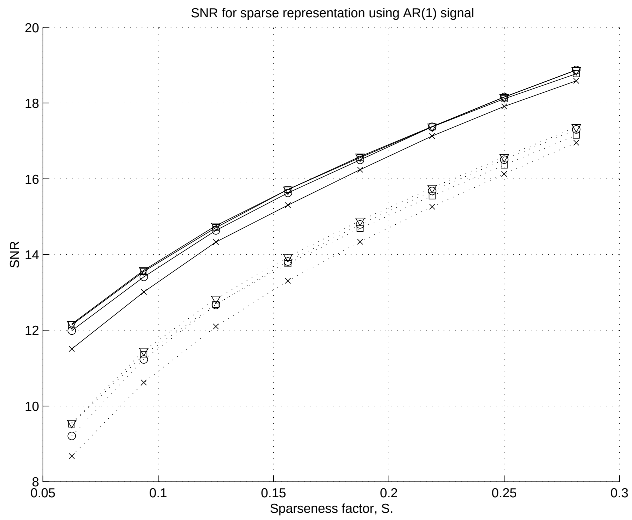

Fig. 4. Approximation quality, measured by SNR, plotted as a function of sparseness factor. Here we have the AR(1) signal. The solid lines are for the designed synthesis systems, using $\mathbf{G}$ as LOT ($\circ$), ELT ($\triangledown$) and wavelet ($\square$), and the block oriented frame ($\times$). The dotted lines are when we just keep the largest coefficients of traditional (orthogonal) decomposition methods, DCT ($\times$), LOT ($\circ$), ELT ($\triangledown$) and wavelet ($\square$).

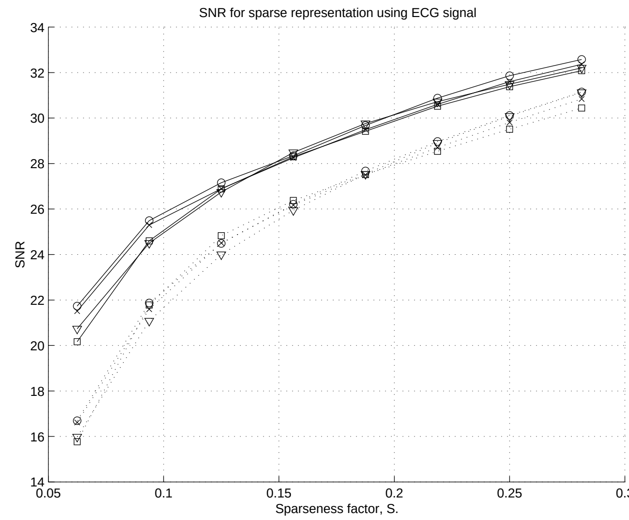

Fig. 5. As in Figure 4, but an ECG signal.

ple 130000, that is, we did not test on the same signals as in the design phase. In both cases we used 102400 signal samples in training and testing.

The orthogonal filter banks we used, i.e. the $\mathbf{G}$s, were the 16 channel, 32 tap LOT [7], the 16 channel, 64 tap Extended Lapped Transform (ELT) [9], and a 4 level dyadic filter bank using the Daubechies wavelet filters of length 12, (Matlab, db6 in wfilters.m). For the sake of comparison we also designed block oriented frames using the method presented in [5], having the fourth $\mathcal{G}$ as the identity matrix. The designed frame, $\mathbf{H}$, was of size 16×32, thus overcomplete by a factor of 2. During design we used four different target sparseness factors.

Testing was done by representing the test signals using the designed frames and some few sparseness factors. In Figure 4 we have plotted the results for the AR(1) signal. Comparing the solid lines with the dotted lines we see the improvement we get by using a frame to make the sparse representation rather than just keep the largest coefficients after an orthogonal transform or filter bank. The improvement is approximately 2 dB. Comparing the marks ($\circ$, $\triangledown$ and $\square$) to the mark ($\times$) we see the improvement we get when we use overlapping synthesis vectors rather than block oriented synthesis vectors. We see that the ELT and wavelet, which have the largest overlap factors, do marginally better than the LOT, but they all do better than the DCT. Figure 5 is similar to Figure 4 but here we have used the ECG signal. The benefit of overlapping synthesis vectors is not as obvious here, the scheme based on the LOT is perhaps marginally better than the rest on an overall assessment.

5. CONCLUSION

We have demonstrated that the simple block oriented frame design method in [5] may be used also for design of overlapping frames, filter banks with overlapping synthesis vectors. It seems reasonable to conclude that for signal classes where filter banks outperform transforms, overlapping frames will outperform block oriented frames, and the improvment will be of approximately the same magnitude.

6. REFERENCES

[1] S. G. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[2] G. Davis, *Adaptive Nonlinear Approximations*, Ph.D. thesis, New York University, Sept. 1994.

[3] M. Gharavi-Alkhansari and T. S. Huang, "A fast orthogonal matching pursuit algorithm," in *Proc. ICASSP '98*, Seattle, USA, May 1998, pp. 1389–1392.

[4] K. Engan, S. O. Aase, and J. H. Husøy, "Designing frames for matching pursuit algorithms," in *Proc. ICASSP '98*, Seattle, USA, May 1998, pp. 1817–1820.

[5] K. Engan, S. O. Aase, and J. H. Husøy, "Method of optimal directions for frame design," in *Proc. ICASSP '99*, Phoenix, USA, Mar. 1999, pp. 2443–2446.

[6] S. O. Aase, K. Skretting, J. H. Husøy, and K. Engan, "Design of signal expansions for sparse representation," in *Proc. ICASSP 2000*, Istanbul, Turkey, June 2000, pp. 105–108.

[7] H. S. Malvar and D. H. Staelin, "The LOT: Transform coding without blocking effects," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 4, pp. 553–559, Apr. 1989.

[8] Massachusetts Institute of Technology, *The MIT-BIH Arrhythmia Database CD-ROM*, MIT, 2 edition, 1992.

[9] H. S. Malvar, "Extended lapped transforms: Fast algorithms and applications," in *Proc. ICASSP '91*, Toronto, Canada, May 1991, pp. 1797–1800.