Heart Rate Monitoring from Wrist-Type Photoplethysmographic (PPG) Signals During Intensive Physical Exercise

Zhilin Zhang
Samsung Research America – Dallas
1301 E. Lookout Dr., Richardson, TX 75082, USA
zhilinzhang@ieee.org

***Abstract*—Heart rate monitoring from wrist-type photoplethysmographic (PPG) signals during subjects' intensive exercise is a difficult problem, since the PPG signals are contaminated by extremely strong motion artifacts caused by subjects' hand movements. In this work, we formulate the heart rate estimation problem as a sparse signal recovery problem, and use a sparse signal recovery algorithm to calculate high-resolution power spectra of PPG signals, from which heart rates are estimated by selecting corresponding spectrum peaks. To facilitate the use of sparse signal recovery, we propose using bandpass filtering, singular spectrum analysis, and temporal difference operation to partially remove motion artifacts and sparsify PPG spectra. The proposed method was tested on PPG recordings from 10 subjects who were fast running at the peak speed of 15km/hour. The results showed that the averaged absolute estimation error was only 2.56 Beats/Minute, or 1.94% error compared to ground-truth heart rates from simultaneously recorded ECG.**

***Index Terms*—Photoplethysmographic (PPG) Signals , Singular Spectrum Analysis (SSA), Sparse Signal Recovery, Heart Rate Estimation, Wearable Computing**

I. Introduction

Heartbeat rate monitoring during fitness is a key feature in many modern wearable devices such as Samsung Gear Fit. These devices generally record photoplethysmographic (PPG) signals [1] from wearers' wrist, and then estimate heart rates from the PPG signals in real time.

However, PPG signals are vulnerable to motion artifacts, which is adverse to heart rate monitoring during fitness (see Figure 1 for example). Many signal processing techniques have been proposed to remove motion artifact (MA) from raw PPG signals.

One technique is independent component analysis (ICA). For example, Kim *et al*. [2] suggested to use a basic ICA algorithm and block interleaving to remove MA. Krishnan *et al*. [3] proposed using frequency-domain based ICA technique to remove MA. However, this technique has some limitations. One limitation is that the key assumption in basic ICA algorithms, namely statistical independence or uncorrelation, is not hold in PPG applications [4]. Thus MA removal is not satisfactory especially when MA is strong.

Another popular technique is adaptive noise cancelation (ANC) [5], [6]. For example, Ram *et al*. [5] proposed using ANC to remove MA, where the reference signal was constructed from fast fourier transform (FFT), singular value decomposition, or ICA of the artifact-contaminated PPG signal. However, one limitation in this method is that the artifact-removal performance of ANC is sensitive to the reference signal, while reconstructing qualified reference signals is extremely difficult when subjects are exercising.

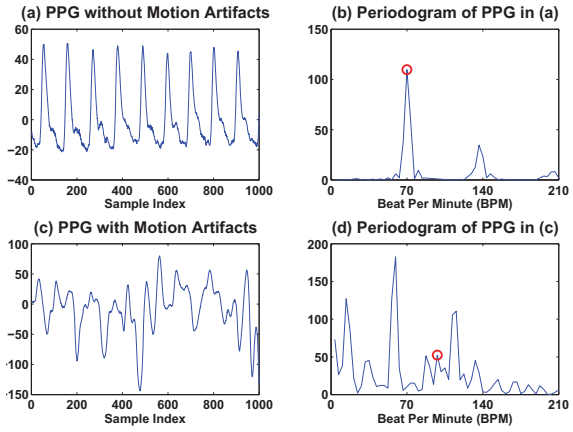

Fig. 1. Comparison between an artifact-free PPG signal and an artifact-contaminated PPG signal. Plot (a) and Plot (b) show the waveform and the spectrum (calculated by Periodogram) of a PPG signal without motion artifacts, respectively. Plot (c) and Plot (d) show the waveform and the spectrum (calculated by Periodogram) of a PPG signal with strong motion artifacts, respectively. The circles in Plot (b) and Plot (d) indicate the spectrum peaks corresponding to the heart rate (calculated from simultaneously recorded ECG). From Plot (d) we see selecting the spectrum peak corresponding to the heart rate is difficult.

It is worth noting that most of these techniques were proposed for the scenarios when small motion movements were incurred, such as finger movements [3], [5], [6], walking [6], or slowly running (with the speed less than 8 km/hour) [7], [8]. Due to the limitations stated above, these proposed techniques may not be effective when subjects perform intensive physical exercise such as fast running.

This work proposes an approach to estimate heart rates when subjects perform intensive physical exercise. We first formulate the heart rate estimation problem into a sparse signal reconstruction problem. To facilitate the use of sparse signal reconstruction algorithms, singular spectrum analysis (SSA) and temporal difference operation are proposed to partially

978-1-4799-7088-9/14/$31.00 ©2014 IEEE

remove motion artifacts. The proposed approach was evaluated on PPG recordings from 10 subjects who were running at the peak speed of 15-17 km/hour on a treadmill. Results showed that the framework has high performance with only 1.94% error compared to ground-truth heart rates calculated from simultaneously recorded ECG.

II. Sparse Signal Recovery (SSR)

Sparse signal recovery (SSR) [9]–[11] is a focus of signal processing and information theory in recent years. The basic SSR model can be expressed as follows,

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \mathbf{v}. \tag{1}$$

Here $\mathbf{\Phi}$ is a known basis matrix of the size $M \times N$. The vector $\mathbf{y}$ is an observed signal of the size $M \times 1$. $\mathbf{x}$ is an unknown solution vector which is assumed to be sparse (i.e. most elements of $\mathbf{x}$ are zero while only a few elements are nonzero). $\mathbf{v}$ is an unknown noise vector. The goal is to find the sparsest vector $\mathbf{x}$ based on $\mathbf{y}$ and $\mathbf{\Phi}$, namely,

$$\widehat{\mathbf{x}} \leftarrow \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2^2 + \lambda g(\mathbf{x}) \tag{2}$$

where λ is a regularization parameter, and $g(\mathbf{x})$ is a penalty function encouraging the sparsity of the solution $\widehat{\mathbf{x}}$. Some popular penalty functions are the ℓ_1-norm function, i.e. $\|\mathbf{x}\|_1$ [12], or the ℓ_p-norm function ($0 < p < 1$), i.e. $\|\mathbf{x}\|_p$ [9].

It is interesting to see that SSR can be used in spectrum estimation [9], [13]. Let $\mathbf{y} \in \mathbb{R}^{M\times 1}$ be a real-valued signal. Constructing the matrix $\mathbf{\Phi} \in \mathbb{C}^{M\times N}$ such that its (m,n)-th element $\mathbf{\Phi}_{m,n}$ is given by

$$\mathbf{\Phi}_{m,n} = e^{j\frac{2\pi}{N}mn},\ m = 0, \cdots, M-1;\ n = 0, \cdots, N-1 \tag{3}$$

the solution (2) leads to a sparse spectrum of the signal $\mathbf{y}$, i.e. $\mathbf{s} \in \mathbb{R}^{N\times 1}$ with the n-th element given by $s_n = |\widehat{x}_n|^2$ ($\widehat{x}_n$ is the n-th element of $\widehat{\mathbf{x}}$).

The SSR-based spectrum estimation has a number of advantages over traditional spectrum estimation methods [9], [13]. Compared to nonparametric spectrum estimation methods such as the FFT-based Periodogram, the SSR-based spectrum features high spectrum resolution, low estimation variance, increased robustness; compared to line spectral estimation methods such as the MUSIC algorithm, the SSR-based spectrum estimation does not require model selection and improved estimation performance.

Note that a key assumption in SSR is that $\mathbf{x}$ is sparse. If not, the performance of SSR algorithms is seriously degraded. When complicated intensive hand movements occur during PPG recordings, the PPG signal contains strong MA and the spectrum coefficients are not sparse. In this situation, SSR algorithms perform poorly. Thus, removing MA and sparsifying the spectrum is crucial to the success of SSR. This is why we need to use bandpass filtering and singular spectrum analysis before SSR.

III. Singular Spectrum Analysis

Singular spectrum analysis (SSA) [14] is a signal decomposition method, which decomposes a time series into oscillatory components and noise. SSA includes four steps, namely *Embedding*, *Singular Value Decomposition (SVD)*, *Grouping*, and *Reconstruction*.

In the Embedding Step, a time series $\mathbf{y} \triangleq [y_1, \cdots, y_M]^T$ is mapped into an $L \times K$ matrix ($K = M - L + 1, L < M/2$), called L-trajectory matrix,

$$\mathbf{Y} \triangleq \begin{bmatrix} y_1 & y_2 & \cdots & y_K \\ y_2 & y_3 & \cdots & y_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & y_M \end{bmatrix}. \tag{4}$$

In the SVD Step, the L-trajectory matrix is decomposed by SVD as follows,

$$\mathbf{Y} = \sum_{i=1}^{d} \mathbf{Y}_i, \quad d \triangleq \min\{L, K\} \tag{5}$$

where $\mathbf{Y}_i = \sigma_i \mathbf{u}_i \mathbf{v}_i^T$, and $\sigma_i, \mathbf{u}_i, \mathbf{v}_i$ are the ith singular value, the corresponding left-singular vector and the corresponding right-singular vector, respectively.

In the Grouping Step, the d rank-one matrix $\mathbf{Y}_i$ are assigned into g groups, namely the set of indices $\{1, \cdots, d\}$ is partitioned into g disjoint subsets $\{I_1, \cdots, I_g\}$ ($g \le d$) and

$$\mathbf{Y} = \sum_{p=1}^{g} \mathbf{Y}_{I_p}, \tag{6}$$

with $\mathbf{Y}_{I_p} = \sum_{t\in I_p} \mathbf{Y}_t$. The rank-one matrices in each group $\mathbf{Y}_{I_p}$ generally satisfy some common characteristics (such as their corresponding oscillatory components after reconstruction have the same frequency or exhibit harmonic relation).

In the Reconstruction Step, each $\mathbf{Y}_{I_p}$ is used to reconstruct a time series $\widetilde{\mathbf{y}}_p$ with the length M by a so-called diagonal averaging procedure. Thus the original signal $\mathbf{y}$ is decomposed into g time series, i.e.

$$\mathbf{y} = \sum_{p=1}^{g} \widetilde{\mathbf{y}}_p. \tag{7}$$

For details, the reader is referred to the literature [14].

SSA is widely used to extract trends and periodic components from time series. In our proposed method, it is used to remove MA components whose dominant frequencies are close to the dominant frequencies in simultaneously recorded acceleration signals. Details will be given in the following.

IV. Proposed Approach

Our proposed approach is based on a single-channel PPG signal and simultaneously recorded acceleration data. In our experiments the acceleration signal is a three-channel signal. A time window of 8 seconds is sliding on the PPG signal and the acceleration signal, with incremental step of 2 seconds. The proposed approach estimates heart rates in each time window. Since the time window is relative long and the incremental

TABLE I
ESTIMATION ERRORS ON ALL TEN SUBJECTS' RECORDINGS.

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Subject 8	Subject 9	Subject 10
Error1 (BPM)	3.09	1.37	2.51	3.58	7.19	2.21	1.60	1.22	1.11	1.79
Error2 (%)	2.82%	0.98%	2.06%	2.20%	5.42%	1.72%	1.14%	1.01 %	0.91%	1.16%

step is relatively short, it is expected that the heart rates in two successive time windows are very close.

Each part of the approach is described below in order.

A. Bandpass Filtering

In a given time window, the recorded raw PPG signal and the acceleration signal are first band-pass filtered with the cut-off frequency of 0.4 Hz and 7 Hz [1]. This preprocessing removes lots of noise and MA outside of the frequency band of interest. This bandpass filtering also sparsifies the spectrum coefficients when using the SSR-based spectrum method later.

B. SSA

After bandpass filtering, SSA is used to remove MA components in the frequency band from 0.4 Hz to 7 Hz. This can further sparsify the spectrum coefficients when using the SSR-based spectrum method.

First, for each channel of acceleration data in the time window, the spectrum is calculated using the Periodogram algorithm, from which dominant frequencies in the spectrum are determined. Here using the Periodogram algorithm instead of using the SSR-based spectrum method is to save computational resources. The dominant frequencies are selected as the frequencies corresponding to the spectrum peaks with magnitudes larger than 50% of the maximum magnitude value in the spectrum. Denote by $\mathcal{F}_{\rm acc}$ the collection of frequency bin indexes which correspond to the dominant frequencies of acceleration data in the three channels.

Then we can perform SSA on the PPG signal. After SVD, the grouping is automatically finished by clustering singular values as described in [14, pp. 66]. Then the original PPG signal is decomposed into a number of time series. Next, we throw away those time series whose dominant frequencies completely lie in $\mathcal{F}_{\rm acc}$. Finally, using the remained time series we can reconstruct a cleansed PPG signal.

Note that the dominant frequencies of acceleration data may contain some frequencies close to the frequency of heart rate. Thus the above procedure may remove the PPG. Hence some modifications to the above procedure are needed. Denote by $f_{\rm prev}$ the frequency bin index of the estimated heart rate frequency in the previous time window. We exclude $\{f_{\rm prev}-\Delta,\cdots,f_{\rm prev}-1,f_{\rm prev},f_{\rm prev}+1,\cdots,f_{\rm prev}+\Delta\}$ from $\mathcal{F}_{\rm acc}$, obtaining a refined collection of frequency bin indexes, denoted by $\widetilde{\mathcal{F}}_{\rm acc}$. Δ is a small positive integer. In our experiment we set $\Delta=4$. Then SSA is performed using the refined $\widetilde{\mathcal{F}}_{\rm acc}$. The parameter L in SSA was set to 400 in our experiment.

The cleansed PPG signal is denoted by $\mathbf{y}_{\rm SSA}$.

[1]Note that the heart rate generally ranges from 40 BPM to 210 BPM, corresponding to 0.6 Hz to 3.5 Hz.

C. Temporal Difference Operation

For a periodic time series $h=[h(1),h(2),\cdots,h(M)]$ with the fundamental frequency f_0, its first-order difference, defined as $h'\triangleq[h(2)-h(1),h(3)-h(2),\cdots,h(M)-h(M-1)]$, maintains the fundamental frequency and the harmonic frequencies. The second-order difference of h, i.e. the first-order difference of h', also maintains the fundamental frequency and the harmonic frequencies. As long as k is not large, the spectrum of the k-th difference of the periodic time series always significantly exhibits the fundamental frequency and its harmonic frequencies. In contrast, this is not observed from a non-periodic time series.

Note that an artifact-free PPG signal is approximately periodic in short time, while MA is generally not periodic (except to the situation when only hand swing occurs). Therefore, we calculate the k-th difference of the cleansed PPG signal $\mathbf{y}_{\rm SSA}$. In our experiments we calculated the second-order difference. The resulting time series of difference is denoted by $\mathbf{y}_{\rm diff}$.

After this step, the spectrum peak corresponding to the heart rate and harmonic spectrum peaks are more prominent in the spectrum of $\mathbf{y}_{\rm diff}$ than in the spectrum of $\mathbf{y}_{\rm SSA}$.

D. SSR by Using Sparse Bayesian Learning

The SSR-based spectrum estimation is performed on $\mathbf{y}_{\rm diff}$. There are many SSR algorithms. In our experiments sparse Bayesian learning (SBL) [15], [16] was used, since SBL algorithms have been known for their high performance even when the basis matrix $\mathbf{\Phi}$ is highly coherent. Particularly we chose the BSBL-BO algorithm proposed in [15] with block size set to 1, due to its good balance between speed and performance. Although the original BSBL-BO algorithm is based on real-valued signals, it can be used for complex cases by making minor modifications in the same way as in [16].

In our experiment we set the parameter N in the model (1) to 2048. The resulting spectrum is denoted by $\mathbf{s}$.

E. Frequency Bin Selection

Denote by $f_{\rm prev}$ the frequency bin index corresponding to the estimated heart rate in the previous time window. We set a search range for the fundamental frequency of heart rate in the spectrum $\mathbf{s}$, denoted by $R_0=[f_{\rm prev}-\Delta,\cdots,f_{\rm prev}+\Delta]$. Also, we set another search range for the first-order harmonic frequency of the heart rate, denoted by $R_1=[2(f_{\rm prev}-\Delta-1)+1,\cdots,2(f_{\rm prev}+\Delta-1)+1]$ [2]. In the spectrum, we select at most two highest peaks in each search range. Denote the frequency bin indexes of the two peaks in R_0 by f_1^0 and f_2^0, and the frequency bin indexes of the two peaks in R_1 by f_1^1 and f_2^1.

[2]Remind that in the spectrum, the first frequency bin corresponds to 0 Hz.

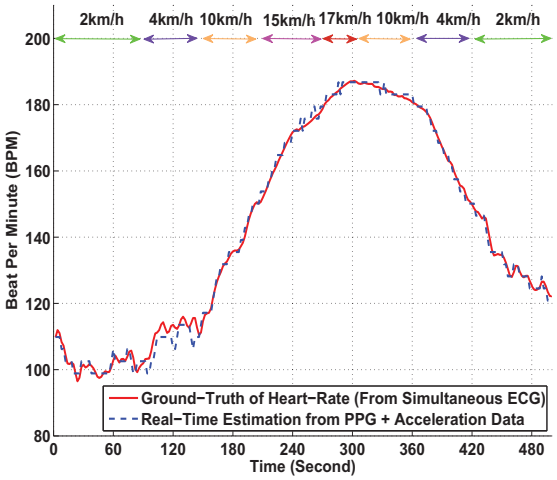

Fig. 2. Estimation result on one subject's recordings. The averaged absolute estimation error was 1.37 BPM, and the averaged error percentage was 0.98%. The speed at each duration is indicated on the top of the figure.

If there exists a peak-pair $(f_i^0, f_j^1)(i \in \{1,2\}, j \in \{1,2\})$ which holds a harmonic relation, then f_i^0 corresponds to the heart rate. If there is no such a peak-pair (this can happen in the presence of strong MA), we select $\widehat{f}$ as follows

$$\widehat{f} \leftarrow \arg\min_f \{|f - f_{\text{prev}}|\} \tag{8}$$

where $f \in \{f_1^0, f_2^0, \frac{f_1^1-1}{2}+1, \frac{f_2^1-1}{2}+1\}$. And $\widehat{f}$ corresponds to the heart rate. Here the expressions $f_1^1 - 1$ and $f_2^1 - 1$ are due to the same reason stated in Footnote 2.

F. Remarks

The proposed approach needs an initial stage, in which wearers are required to reduce hand motions as much as possible for several seconds (2 or 3 seconds are enough). During this short initial stage, the approach can estimate the initial heart rate by simply looking at the highest peak in the spectrum of a PPG signal. Without this initial stage, there is no way to estimate the heart rate, since the spectrum of an MA-contaminated PPG signal contains lots of peaks and there is no prior information to help find the correct spectrum peak.

V. Experimental Results

We simultaneously recorded PPG signals, three-axis acceleration signals, and ECG signals from 10 male subjects with age from 18 to 33. For each subject, the PPG signal was recorded from his wrist using a pulse oximeter with a green LED (wavelength: 609nm). The sensor and a three-axis accelerometer were built in a wrist-band. A single-channel ECG signal was recorded simultaneously from the chest of each subject using a wet ECG sensor. All signals were sampled at 125 Hz and sent to a nearby computer via bluetooth.

During data recording, each subject performed a series of exercise on a treadmill: first walked at the speed of 2 km/hour for 1 minute, then fast walked at the speed of 4 km/hour for 1 minute, and then ran at the speed of 10 km/hour for 1 minute, next fast ran at the speed of 15km/hour - 17km/hour for 1 to 1.5 minutes, and then again ran at the speed of 10 km/hour for 1 minute, and finally walked at the speed of 4 km/hour for 1 minute. The subjects were asked to purposely use the hand with the wrist-band to pull clothes, wipe sweat on forehead, and push buttons on the treadmill, in addition to freely swing.

Using the PPG signal and the simultaneously recorded acceleration signals, our proposed approach reported estimated heart rate every two seconds.

The ground-truth of the heart rate in each time window was calculated manually from the simultaneously recorded ECG signal. That is, given a time window, we counted the number of cardiac cycles Q and the duration T (in seconds), and then calculated the heart rate as $60Q/T$ (in BPM).

Denote the heart rate ground-truth in the i-th time window by $\text{BPM}_{\text{true}}(i)$, and denote the estimated heart rate by our approach by $\text{BPM}_{\text{est}}(i)$. The averaged absolute estimation error is defined as $\text{Error1} = \frac{1}{W}\sum_{i=1}^{W}\left|\text{BPM}_{\text{est}}(i) - \text{BPM}_{\text{true}}(i)\right|$, where W is the total number of time windows. Similarly, we define another performance index called averaged error percentage by $\text{Error2} = \frac{1}{W}\sum_{i=1}^{W}\frac{\left|\text{BPM}_{\text{est}}(i)-\text{BPM}_{\text{true}}(i)\right|}{\text{BPM}_{\text{true}}(i)}$.

Figure 2 shows the estimation result on one subject's recordings. One can see the estimated heart rate was very close to the ground-truth in each time window. The estimation results on the 10 subjects' recordings are listed in Table I. Averaged across the 10 subjects, the absolute estimation error was 2.56 BPM and the error percentage was 1.94%.

VI. Conclusion

This work studied the problem of heart rate estimation from wrist-type PPG signals when subjects are performing intensive physical exercise. A new method based on singular spectrum analysis was proposed to remove motion artifact in raw PPG signals. And a new spectrum estimation method based on sparse signal recovery was proposed to overcome the limitations of traditional spectrum estimation algorithms. Experimental results on recordings from 10 subjects showed that the proposed method has superior performance with 1.94% error compared to ground-truth heart rates from simultaneously recorded ECG. Note that the errors are much lower than the estimation errors reported in literature [7], [8].

Acknowledgement

The author is deeply grateful to Dr. Benyuan Liu for his help in data collection. This work was supported in part by Samsung Research America–Dallas. Any opinions, findings, and conclusions expressed in this work are those of the author and do not reflect the views of the funding organization.

References

[1] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological measurement*, vol. 28, no. 3, pp. R1–R39, 2007.

[2] B. S. Kim and S. K. Yoo, "Motion artifact reduction in photoplethysmography using independent component analysis," *Biomedical Engineering, IEEE Transactions on*, vol. 53, no. 3, pp. 566–568, 2006.

[3] R. Krishnan, B. Natarajan, and S. Warren, "Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 8, pp. 1867–1876, 2010.

[4] J. Yao and S. Warren, "A short study to assess the potential of independent component analysis for motion artifact separation in wearable pulse oximeter signals," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, 2005, pp. 3585–3588.

[5] M. Ram, K. V. Madhav, E. H. Krishna, N. R. Komalla, and K. A. Reddy, "A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter," *Instrumentation and Measurement, IEEE Transactions on*, vol. 61, no. 5, pp. 1445–1457, 2012.

[6] R. Yousefi, M. Nourani, S. Ostadabbas, and I. Panahi, "A motion-tolerant adaptive algorithm for wearable photoplethysmographic biosensors," 2014.

[7] H. Fukushima, H. Kawanaka, M. S. Bhuiyan, and K. Oguri, "Estimating heart rate using wrist-type photoplethysmography and acceleration sensor while running," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 2012, pp. 2901–2904.

[8] S. M. López, R. Giannetti, M. L. Dotor, J. P. Silveira, D. Golmayo, F. Miguel-Tobal, A. Bilbao, M. Galindo Canales, P. Martín Escudero *et al.*, "Heuristic algorithm for photoplethysmographic heart rate tracking during maximal exercise test," *Journal of Medical and Biological Engineering*, vol. 32, no. 3, pp. 181–188, 2012.

[9] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm," *IEEE Trans. on Signal Processing*, vol. 45, no. 3, pp. 600–616, 1997.

[10] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[11] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer, 2010.

[12] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *J. R. Statist. Soc. B*, vol. 58, no. 1, pp. 267–288, 1996.

[13] M. F. Duarte and R. G. Baraniuk, "Spectral compressive sensing," *Applied and Computational Harmonic Analysis*, vol. 35, no. 1, pp. 111–129, 2013.

[14] N. Golyandina, V. Nekrutkin, and A. A. Zhigljavsky, *Analysis of time series structure: SSA and related techniques*. CRC Press, 2001.

[15] Z. Zhang and B. D. Rao, "Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation," *IEEE Trans. on Signal Processing*, vol. 61, no. 8, pp. 2009–2015, 2013.

[16] D. Wipf and B. Rao, "An empirical bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3704–3716, 2007.