

# Artifact Detection in Electrodermal Activity using Sparse Recovery

Authors: Malia Kelsey, Richard Vincent Palumbo, Alberto Urbaneja, Murat Akcakaya, Jeannie Huang, Ian R. Kleckner, Lisa Feldman Barrett, Karen S. Quigley, Ervin Sejdic, and Matthew S. Goodwin

## ABSTRACT

Electrodermal Activity (EDA) – a peripheral index of sympathetic nervous system activity - is a primary measure used in psychophysiology. EDA is widely accepted as an indicator of physiological arousal, and it has been shown to reveal when psychologically novel events occur. Traditionally, EDA data is collected in controlled laboratory experiments. However, recent developments in wireless biosensing have led to an increase in out-of-lab studies. This transition to ambulatory data collection has introduced challenges. In particular, artifacts such as wearer motion, changes in temperature, and electrical interference can be misidentified as true EDA responses. The inability to distinguish artifact from signal hinders analyses of ambulatory EDA data. Though manual procedures for identifying and removing EDA artifacts exist, they are time consuming – which is problematic for the types of longitudinal data sets represented in modern ambulatory studies. This manuscript presents a novel technique to automatically identify and remove artifacts in EDA data using curve fitting and sparse recovery methods. Our method was evaluated using labeled data to determine the accuracy of artifact identification. Procedures, results, conclusions, and future directions are presented.

## 1. INTRODUCTION

Electrodermal Activity (EDA), a peripheral measure of sympathetic nervous system activity, is a widely accepted indicator of autonomic and psychological arousal [1], [2]. EDA measures changes in electrical potential on the skin surface due to eccrine sweat gland activity. EDA has long been used to study a variety of physiological topics, including stress, emotion, depression, anxiety, attention, and information processing [3], [4], [5], [6]. Although EDA has been widely used in the psychological literature, relatively little work has directly addressed detection of artifacts and noise in EDA signals. Traditionally, robust artifact detection has not been necessary due to the low prevalence of noise and artifacts in laboratory collected data [4]. However, recent developments in wireless biosensing has increased the number of long-term, ambulatory studies [4], [7], [8]; increasing the need for a robust, automatic way to detect and remove artifacts and noise from signals.

EDA signals are built up of three components that are often separated for analysis: skin conductance level (SCL); skin conductance response (SCR); and artifacts. SCL refers to slowly fluctuating tonic levels. These are typically used to evaluate general trends in activity and levels of activation. SCL can vary widely, typically between 2-20  $\mu\text{S}$ , within and between individuals due to environmental and personal factors [1], [2]. Given this variability, it is not uncommon to remove the tonic response during between-person analyses. SCR refers to quick, high-frequency responses superimposed on the tonic level. They can be characterized by rise time, amplitude, and half recovery time. In healthy adults, rise time is usually between 1 and 3 seconds, amplitude often varies (a minimum is commonly set between .01 and .05  $\mu\text{S}$ ), and half recovery time is typically between 2 and 10 seconds [2]. Figure 1 shows the typical shape and parameters used to describe an SCR. The third component of EDA data is artifact. These responses can be caused by many factors, including physical movement, environmental factors such as changes in ambient temperature, and electrical noise [9]. Artifacts can be challenging to identifying because they are often similar in shape and phase to SCRs, making identifying and distinguishing between artifacts and SCRs a challenging and manually intensive endeavor.

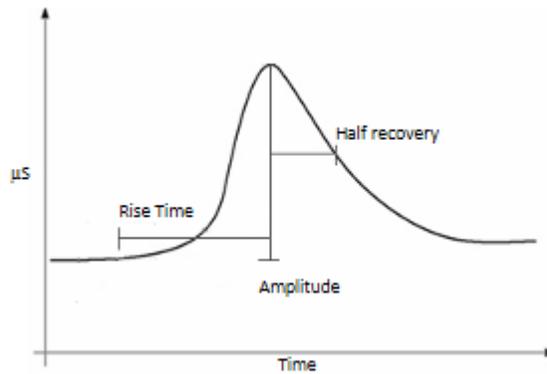


Figure 1: Illustration of typical SCR shape

### 1.1 Current Methods

Current approaches to removing artifacts from EDA data involve applying exponential smoothing [5], low pass filtering [3], [10], [11] or manual visual inspection procedures. Though each of these methods work well when applied to data collected in a lab, they may not extend well to ambulatory data wherein the data collection environment is less controlled. Exponential smoothing and low pass filtering are widely used in the literature, but typically they only smooth out low amplitude artifacts and are less effective on higher magnitude responses which are more common in naturalistic ambulatory settings [2]. Additionally, filter cutoff frequencies are based only loosely on prior knowledge of typical characteristics of SCR shape, and vary widely study to study (1 and 5 Hz). The cutoff frequency ultimately chosen for a study is specific to that particular study, making generalization difficult [9]. Manual inspection is prone to varying interpretations, particularly given the similarities between SCRs and artifacts. Additionally, many ambulatory studies collect data for hours or days, making manual inspection infeasible [4]. One technique used to overcome some of the challenges posed by similarities in SCRs and artifacts is to collect EDA data from multiple sites (i.e., both wrists). In this scenario, only SCRs simultaneously identified in EDA signals from bilateral sites are considered as valid, disregarding the rest as artifacts. However, this still requires manual inspection and collection from multiple sites which may not be feasible or tolerated in ambulatory studies [6], [10].

A recent study applied machine learning techniques to automatically identify EDA artifacts [9]. It was found that Support Vector Machines (SVM) with a radial basis function was successfully able to distinguish between clean and noisy sections of data with test accuracies of about 96%. While this method addresses the issue of missing high magnitude artifacts, it performs classification by sectioning the data. This can introduce additional issues, including: (1) if there are multiple responses in a section (including artifacts and SCRs) there is no way to classify them individually; (2) an SCR or artifact could be missed if split between two sections; and (3) noisy data may still require manual cleaning.

Our previous work proposed the use of a sparse recovery algorithm, specifically orthogonal matching pursuit (OMP), to address the need for automatic SCR and artifact detection [12]. Sparse recovery is a technique that can be used to estimate a signal by linearly adding columns from a dictionary of predefined waveforms. Sparse recovery algorithms aim to solve for  $\gamma$  using  $D\gamma = x$  where  $D$  is the dictionary,  $x$  is the original signal, and  $\gamma$  is a coefficient matrix which gives the weight of the dictionary columns that can be used to estimate the original signal [13]. Dictionaries are typically represented through a matrix where each column, called atoms, represent a specific waveform. Dictionaries are designed to be fat matrices, meaning a single exact solution to the above equation does not exist. Instead, greedy approaches are used to solve for  $\gamma$  as an approximation [13]. The most popular greedy approaches fall under the category of Matching Pursuit (MP) algorithms, one of which is the OMP algorithm. Our previous work showed that sparse recovery could be used to successfully identify SCRs in lab collected data, but did not apply OMP to artifact identification. The current work aims to extend that methodology to identify artifacts.

## 1.2 Our Contributions

In our previous work, we showed that an OMP methodology could be successfully used to identify SCRs in lab collected EDA signals with 69% accuracy. Additionally, we showed that SCRs and artifacts were separable after fitting both with the Bateman equation [12]. Using similar methodologies to our prior work, our contributions in this manuscript are 2-fold:

- 1) We show the effect that different filters have on SCRs and artifacts, specifically focusing on maintaining the shape of SCRs while removing artifacts.
- 2) We demonstrate that artifacts can be successfully identified using sparse recovery with high accuracy.

In this study, we analyzed 10 EDA signals from a pediatric pain study that collected data in natural clinical settings [8], [14]. Using expert human labels as the ground truth our proposed method allowed us to label SCRs and artifacts with an average accuracy of 85%. Additionally, an investigation of filters showed that using between a .35 Hz and .5 Hz low pass filter gave the best ability to distinguish between SCRs and artifacts.

## 2. METHODS

### 2.1 Experimental Data

Data used in this study was collected as part of a larger study evaluating pain in youth after surgery. Ambulatory data was collected using wireless biosensors worn on both wrists by youth after laparoscopic appendectomy. Trials were run at three intervals: within 24 of surgery; within 48 hours of surgery; and after clinical resolution of pain. For this work, an expert rater (author A.U trained by author R.P) completed labeling of SCRs and artifacts was completed for 5 randomly selected files. Each file contained data collected from both the right and left wrists. To label each file, an algorithm was used to identify possible artifacts and SCRs in the signals. Algorithm-identified artifact and SCRs were then re-evaluated by the trained expert. For both the right and left wrist data, potential responses were labeled an SCR, an artifact, or an SCR plus artifact by the expert. Since the original algorithm was not designed to capture every response in the signal, the labels in each signal provide information for initial testing, but may not create a label for all responses in a signal. The 10 test signals generated 264 labeled SCRs, 70 SCRs plus artifacts, and 43 artifacts.

### 2.2 Optimal Filtering Analysis

To test the potential of various filters to successfully remove artifacts of various magnitudes, three different low pass filters with cut off frequencies of .35, .5, and 1 Hz were tested. The cutoff frequencies were chosen to match filters used in the literature as well as prior knowledge of SCR rise time. In healthy adults, SCR rise time is typically between 1 – 3 seconds, corresponding to frequencies between .35 – 1 Hz. Any components changing faster than 1 second were therefore assumed to be noise, and were removed before further analysis. Filter performance was evaluated using expert labels provided for each signal as the gold standard for comparison. For each labeled artifact, filter performance was determined using visual inspection and by calculating the mean square error (MSE) and peak amplitude difference between the filtered artifact and original artifact. This performance comparison was repeated for labeled SCRs and SCR plus artifacts.

Differential performance of the 3 different filters was assessed using two-sided Wilcoxon rank sum tests to determine the MSE and amplitude differences across the different filters. Two-side rank sum tests allowed us to determine whether performance of the 3 filters were statistically the same or not. The rank sum test is a nonparametric rank test that assesses equality of the median for two populations. The null hypothesis of the rank sum test assumes that the medians of the two populations are the same, and the alternative hypothesis assumes that the two population medians are different [15].

### 2.3 Artifact Detection with Sparse Recovery

To identify labeled SCRs and artifacts in the signals, sparse recovery was applied using the same methodology introduced in our previous work. This section aims to provide a brief overview of that methodology. The interested reader is directed to [12] for further details.

*Preprocessing and phasic estimation:* Each signal was preprocessed using optimal filtering (see section 2.2). Signals were then downsampled to provide the lowest possible sampling frequency while avoiding aliasing. Downsampling significantly decreases the run time required to complete sparse recovery, and previous work showed no benefits to using higher sampling frequencies [12]. After preprocessing, tonic level was estimated by finding local minima throughout the signal, then interpolating through the minima using a linear fit. A local minimum was identified if it was greater than 1 second from the previous minimum, and the directional derivative was greater than zero on both right and left edges. Finally, the phasic component of the signal was estimated by subtracting the tonic estimate from the full EDA signal.

*Dictionary Creation:* To create a dictionary for response identification, two different dictionaries were built and tested. The first contained atoms representing artifacts only. The second contained atoms representing both artifacts and SCRs. Each atom in the dictionary was selected based on fitting labeled artifacts and SCRs using the Bateman equation, shown in (1).

$$F(t) = e^{-\frac{t}{\tau_1}} + e^{-\frac{t}{\tau_2}} \quad (1)$$

In (1),  $t$  is the time over which the SCR or artifact is present and  $\tau_1$  and  $\tau_2$  control the overall shape of the response. To build each dictionary, different  $\tau_1$  and  $\tau_2$  pairs were used to represent different shapes of SCRs and artifacts common in EDA data. The most common parameters were chosen through the use of bivariate histograms.

*Sparse recovery:* The OMP algorithm employed in this methodology, termed batch orthogonal matching pursuit (BOMP) [16], is solved using the optimization show in (2)

$$\underline{\gamma} = \underset{\underline{\gamma}}{\text{Argmin}} \left\| x - D\underline{\gamma} \right\|_2^2 \quad \text{Subject To} \quad \left\| \underline{\gamma} \right\|_0 \leq K \quad (2)$$

In (2),  $\gamma$  is the estimated coefficient vector,  $x$  is the phasic estimate to be analyzed,  $D$  is the dictionary, and  $K$  is the sparsity constraint. The BOMP algorithm was chosen over a traditional OMP algorithm due to the fact that the BOMP algorithm reduces computational complexity, and therefore analysis time, by introducing a Cholesky factorization [16]. Another difference between traditional OMP methods and BOMP is that BOMP uses sparsity as the stopping criteria for iterations, which further enforces sparsity on the estimates. Using sparsity as the stopping criteria makes  $K$  into a system parameter that needs to be optimized for each signal. To accomplish this, the BOMP analysis was run over a range of  $K$  values, 10 to 200 for each signal, and the best value was chosen based on the calculated performance measures. Figure 2 summarizes the algorithm used to complete BOMP.

1	Input: Dictionary $D$ , signal $x$ , target sparsity $K$
2	Initialize: Set $I := ()$ , $L := [1]$ , $r := x$ , $\gamma := 0$ , $\alpha := D^T x$ , $n := 1$
3	<b>while</b> sparsity $< K$
4	$\hat{k} := \text{argmax}  d_k^T r $
5	<b>if</b> $n > 1$
6	$w := \text{Solve for } w \{Lw = D_I^T d_{\hat{k}}\}$
7	$L := \begin{bmatrix} L & 0 \\ w^T & \sqrt{1 - w^T w} \end{bmatrix}$
8	<b>end</b>
9	$I := (I, \hat{k})$
10	$\gamma_I := \text{solve for } c \{LL^T c = \alpha_I\}$
11	$r = x - D_I \gamma_I$
12	$n = n + 1$
13	<b>end</b>
14	Output: Sparse representation $\gamma$ such that $x \approx D\gamma$

Figure 2: Pseudocode for the BOMP algorithm

*Performance measures:* Performance was evaluated based on sparse recovery's ability to correctly identify the expert labels provided for each signal introduced in section 2.1. To evaluate the performance of the artifact only dictionary, each labeled SCR and artifact was evaluated as follows: a correctly identified artifact was counted as a true positive, a labeled artifact not identified during analysis was counted as a false negative, a labeled SCR identified as an artifact was counted as a false positive, and a labeled SCR not identified during analysis was counted as a true negative. Using these counts, accuracy was calculated. To evaluate performance when using the SCR and artifact dictionary, each SCR and artifact was evaluated as follows: for each labeled SCR or artifact, a true positive was counted if our methodology correctly identified the response, a false positive was counted if a response was labeled as an artifact but identified as an SCR or vice versa, and a false negative was counted if nothing was identified in a section of data that was labeled as having an artifact or SCR. As no data was labeled as no response, there were no true negatives. Accuracy was again calculated using performance counts.

### 3. RESULTS AND DISCUSSION

#### 3.1 Optimal Filtering Analysis

We tested the effectiveness of 3 low pass filters to remove artifacts without impacting the shape of SCRs. The filters had cutoff frequencies of: .35; .5; and 1 Hz. Results from comparative analyses of the MSE of each SCR and artifact showed no clear difference between the filters. That is, no statistical differences were found between the average MSE values of the 3 filters. Additionally, no statistical differences were found between the MSE for each labeled SCR using the .35 Hz and .5 Hz filters or between the .5 Hz and 1 Hz. However, a statistical difference was identified between the .35 Hz and 1 Hz filters ( $p = 2.29 * 10^{-4}$ ). A similar analysis of the MSE and average amplitude difference for artifacts and SCRs plus artifact produced similar results, indicating no statistical differences between filters. Table 1 shows average MSE and difference in amplitude calculated across all SCRs and artifacts for each filter.

Table 1: Average mean square error and difference in amplitude for each label type and filter

	<b>.35 Hz LP filter</b>	<b>.5 Hz LP Filter</b>	<b>1 Hz LP Filter</b>
<b><i>SCR</i></b>			
Average MSE	0.025	0.017	0.012
Average Amplitude difference	0.122	0.103	0.092
<b><i>Artifact</i></b>			
Average MSE	0.122	0.675	0.410
Average Amplitude difference	0.022	0.220	0.173
<b><i>SCR plus artifact</i></b>			
Average MSE	0.122	0.027	0.019
Average Amplitude difference	0.022	0.132	0.118

Visual inspection showed more distinction between the 3 filters, though no filter completely removed high magnitude artifacts. Additionally, all three filters caused artifacts to be smoothed out, resulting in artifacts looking more like SCRs, and none significantly changed labeled SCRs. After filtering, high magnitude artifacts looked like SCRs when filtered

with 1 Hz, but were more successfully removed by the .35 Hz and .5 Hz filter. Figure 3a shows an example artifact plotted using unfiltered data and 1 Hz filtered data. The same artifact is shown in Figure 3b, which shows unfiltered data and .35 Hz filter data. It can be seen that the .35 Hz filter mostly removes the artifact, while the 1 Hz filter still shows a response.

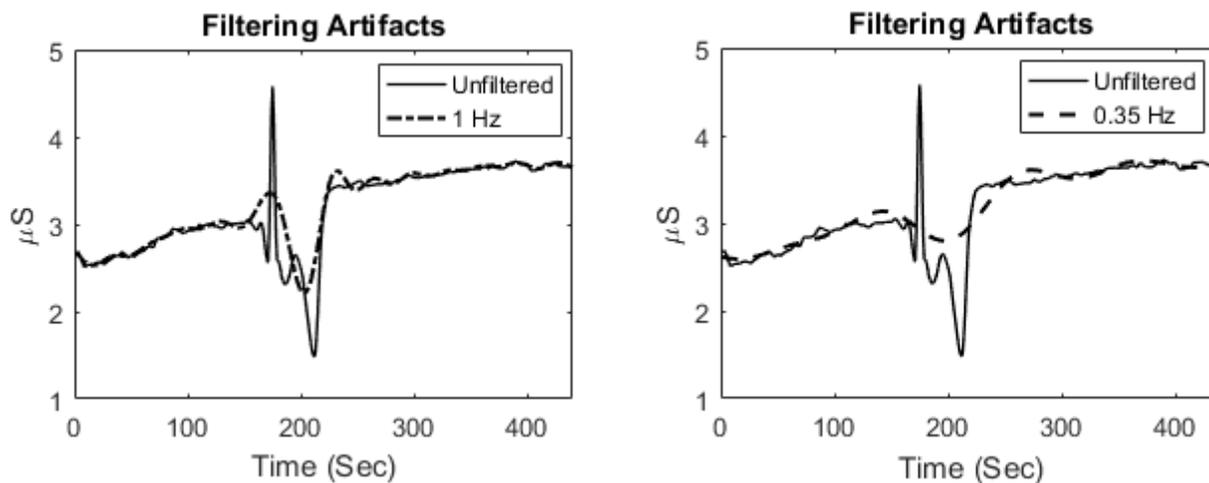


Figure 3a: Investigating unfiltered data and 1 Hz filtered data still shows a response. 3b: Investigating unfiltered data and .35 Hz filter shows that the artifact is mostly removed.

Visual inspection of the labeled SCRs before and after filtering also suggested that either .35 Hz or .5 Hz filter would be an optimal cutoff frequency. The 1 Hz filter tended to add extra “bumps” to the signal, as some higher frequency components are allowed through. Due to this, the 1 Hz filter tended to make single SCRs appear as multiple SCRs. This affect is shown in Figure 4 below. Again, the .35 Hz and .5 Hz filters performed similarly and did not have this effect.

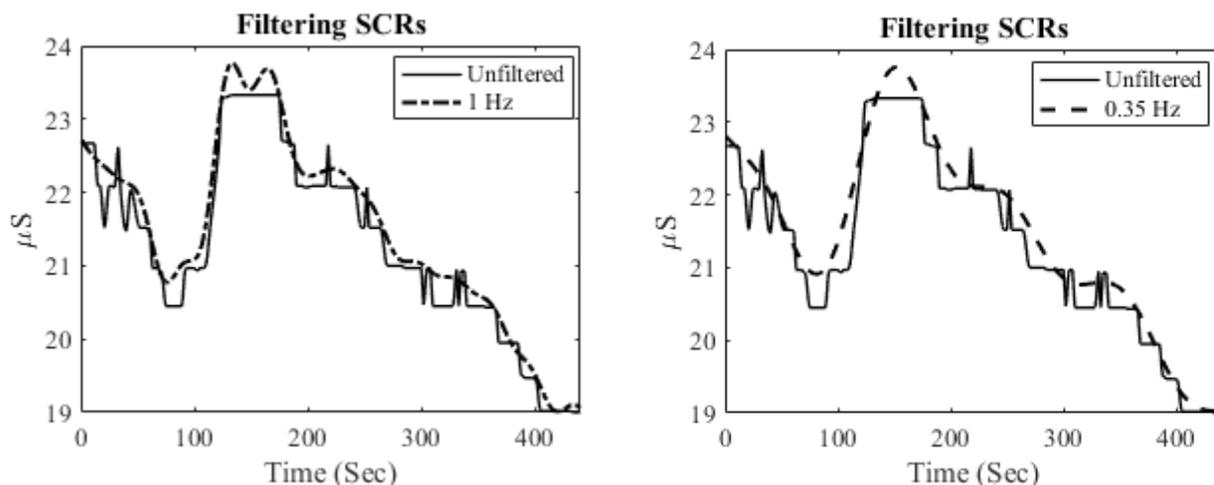


Figure 4a: SCR unfiltered data and 1 Hz filtering. 4b: SCR unfiltered data and .35 Hz filtering.

Finally, the SCR plus artifact labels were processed with all 3 filters. Results indicated similar issues, suggesting .35 Hz or .5 Hz filters are optimal.

### 3.2 Sparse Recovery with BOMP

Once an optimal filter was selected (either .35 Hz or .5 Hz), each signal was downsampled and the tonic and phasic components estimated and separated. Before BOMP could be applied, a dictionary containing appropriated basis vectors had to be created, as described in section 2.3. The  $\tau_1$  and  $\tau_2$  values used for SCRs were determined in our previous work, and represent SCRs with a range of SCR lengths between 5 and 30 seconds [12]. The parameter pairs used to represent artifacts were determined after fitting labeled artifacts with (1) and determining the most common  $\tau_1$  and  $\tau_2$  values. This lead to 3 parameter pairs, resulting in the three artifact shapes illustrated in Figure 5.

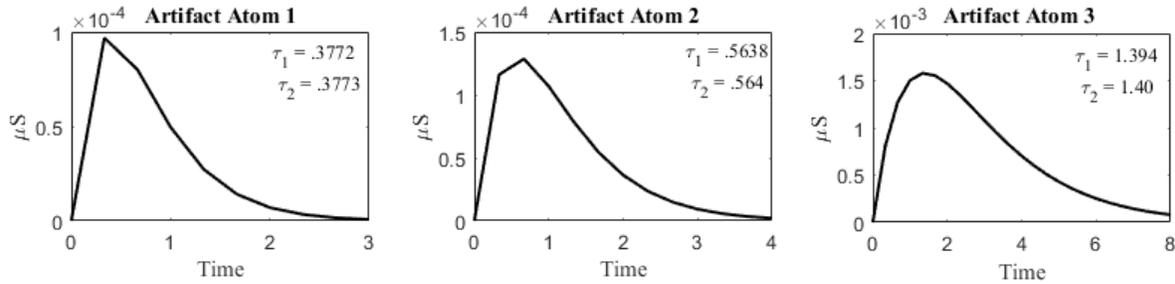


Figure 5: Artifact atoms used to create EDA-specific atoms for dictionary creation

Using the artifact only dictionary, BOMP was unable to successfully label artifacts, and produced mostly false positive identifications. The artifact only dictionaries' poor performance is not surprising given similarities between artifacts and SCRs. Additionally, OMP algorithms will mostly fit to high magnitude responses in the signal, making it more likely that SCRs will be fit before artifacts are considered.

However, the dictionary containing atoms for both SCRs and artifacts resulted in better results, with most files achieving accuracies over 80%. Additionally, the best accuracy for 60% of the files came from the .5 Hz filter, and the best accuracy for the other 40% came from the .35 Hz filter. This matches what was seen when investigating the different filters. The best accuracy achieved for each file and filter is shown below in Table 2, with the maximum accuracy for each file shown in bold.

Table 2: Accuracy achieved using BOMP by file and filter

File	0.35 Hz	0.5 Hz	1 Hz
1	85.19%	<b>96.30%</b>	81.48%
2	53.17%	<b>82.61%</b>	82.61%
3	73.68%	<b>78.95%</b>	63.16%
4	82.14%	<b>82.14%</b>	64.29%
5	64.86%	<b>70.27%</b>	54.05%
6	<b>75.00%</b>	72.92%	60.42%
7	85.19%	<b>92.60%</b>	62.96%
8	<b>80.00%</b>	73.33%	70.00%
9	<b>88.46%</b>	69.23%	80.77%
10	<b>86.67%</b>	66.67%	80.00%

#### 4. CONCLUSION

This work showed that an OMP methodology can be successfully applied to identify artifacts in EDA signals with accuracies around 80%. Using an OMP approach to simultaneously identify SCRs and artifacts also removes the need for artifacts to be removed separately, since a single analysis can be performed. Additionally, tests of different filtering cutoff frequencies indicate that, for ambulatory data, lower cutoff frequencies will be better able to remove artifacts while maintaining separability between artifact and SCR shapes. In the future, we hope to extend this study to a larger test set and complete more rigorous comparisons with other methods. Additionally, up to this point our analysis has required the removal of the tonic level before OMP is run. However, the tonic level includes different information that could be valuable to researchers. Investigating the addition of tonic atoms into the dictionary may allow for more robust analysis to be completed.

#### ACKNOWLEDGMENT

Murat Akcakaya was supported by Air Force Office of Scientific Research (AFOSR) under award number FA9550-16-1-0386.

#### REFERENCES

- [1] W. Boucsein, D. C. Fowles, S. Grimnes, G. Ben-Shakhar, W. T. Roth, M. E. Dawson and D. L. Filion, "Publication Recommendations for Electrodermal Measurements," *Psychophysical Psychophysiology*, vol. 49, no. 8, pp. 1017-1034, 2012.
- [2] J. T. Cacioppo, L. G. Tassinary and G. Berntson, *Handbook of Psychophysiology*, Cambridge: Cambridge Up, 2007.
- [3] A. Sano and R. W. Picard, "Stress Recognition using Wearable Sensors and Mobile Phones," in *Humaine Association Conference on Affective Computing and Intelligent Interaction (IEEE)*, 2013.
- [4] S. Doberenz, W. T. Roth, E. Wolburgh, N. I. Maslowski and S. Kim, "Methodological considerations in ambulatory skin conductance monitoring," *International Journal of Psychophysiology*, vol. 80, no. 2, pp. 87-95, 2011.
- [5] J. Hernandez, R. R. Morris and R. W. Picard, "Call Center Stress Recognition with Person-specific Models," Springer, 2011, pp. 125-134.
- [6] H. Storm, A. Fremming, S. Odegaard, O. G. Martinsen and L. Morkid, "The development of a software program for analyzing spontaneous and externally elicited skin conductance changes in infants and adults," *Clinical Neurophysiology*, vol. 111, no. 10, pp. 1889-1898, 2000.
- [7] R. Hoehn-Saric, D. R. McLeod, F. Funderburk and P. Kowalski, "Somatic Symptoms and Physiologic Responses in Generalized Anxiety Disorder and Panic Disorder," *Arch Gen Psychiatry*, vol. 61, 2004.
- [8] M. Goodwin, M. Moghadamfalahi, R. V. Palumbo, K. Sikka, D. Diaz, M. Bartlett, K. Craig and J. Huang, "Electrodermal activity correlates with self-reported acute pain scores in youth following surgery," in *Pediatric Academic Societies*, Baltimore, MD, 2016.
- [9] S. Taylor, N. Jaques, W. C. Chen, S. Fedor, A. Sano and R. Picard, "Automatic Identification of Artifacts in Electrodermal Activity Data," in *Engineering in Medicine and Biology Society (IEEE)*, 2015.

- [10] R. Kocielnik, N. Sidorova, F. M. Maggi, M. Ouwerkerk and J. H. Westerink, "Smart Technologies for Long-Term Stress Monitoring at Work," in *IEEE International Symposium on Computer-Based Medical Systems*, 2013.
- [11] S. Schmidt and H. Walach, "Electrodermal Activity (EDA) - State-of-the-art Measurement and Techniques for Parapsychological Purposes," *Journal of Sarapsychology*, vol. 64, pp. 139-163, 2000.
- [12] M. Kelsey, M. Akcakaya, I. R. Kleckner, R. V. Palumbo, L. F. Barrett, K. S. Quigley and M. S. Goodwin, "Applications of Sparse Recovery and Dictionary Learning to Enhance Analysis of Electrodermal Activity Data," *Biomedical Signal Processing and Control*, Under Review.
- [13] G. Rath and A. Sahoo, "A comparative Study of some Greedy Pursuit Algorithms for Sparse Approximation," in *Signal Processing Conference (IEEE)*, 2009.
- [14] K. Sikka, A. A. Ahmed, D. Diaz, M. S. Goodwin, K. D. Craig, M. S. Bartlett and J. S. Huang, "Automated assesment of children's postoperative pain using computer vision," *Pediatrics*, vol. 1, no. 136, pp. 124-131, 2015.
- [15] J. D. Gibbons and S. Chakraborti, "Nonparametric Statistical Inference," in *International Encyclopedia of Statistical Science*, Springer Berlin Heidelberg, 2011, pp. 977-979.
- [16] R. Rubinstein, M. Zibulevsky and M. Elad, "Efficient Implementation of the K-SVD Algorithm using Batch Orthogonal Matching Pursuit," *CS Technion*, vol. 40, no. 8, pp. 1-15, 2008.
- [17] D. M. Alexander, C. Trengrove, P. Johnston, T. Cooper, J. August and E. Gordon, "Separating individual skin conductance responses in a short interstimulus-interval paradigm," *Journal of Neuroscience Methods*, vol. 146, no. 1, pp. 116-123, 2005.
- [18] M. Benedek and C. Kaernbach, "Decomposition of skin conductance data by means of nonnegative deconvolution," *Psychophysiology*, vol. 47, no. 4, pp. 647-658, 2010.
- [19] M. Benedek and C. Kaernback, "A Continuous Measure of Phasic Electrodermal Activity," *Journal of Neuroscience Methods*, vol. 190, no. 1, pp. 80-91, 2010.
- [20] D. R. Bach, G. Flandin, K. J. Friston and R. J. Dolan, "Modelling event-related skin conductance responses," *International Journal of Psychophysiology*, vol. 75, no. 3, pp. 349-356, 2010.
- [21] D. R. Bach, J. Daunizeau, K. J. Friston and R. J. Dolan, "Dynamic Causal Modelling of Anticipatory Skin Conductance Responses," *Biological Psychology*, vol. 85, pp. 163-170, 2010.
- [22] D. R. Bach, K. J. Friston and R. J. Dolan, "Analytic measures for quantification of arousal from spontaneous skin conductance fluctuations," *Journal of Psychoophysiology*, vol. 76, pp. 52-55, 2010.
- [23] D. R. Bach, K. J. Friston and R. J. Dolan, "An improved algorithm for model-based analysis of evoked skin conductance responses," *Biological Psychology*, vol. 94, pp. 490-497, 2013.
- [24] D. R. Bach, "A Head-to-head comparison of SCRalyze and Ledalab; Two model-based methods for skin conductance Analysis," *Biological Psychology*, vol. 103, pp. 63-68, 2014.

- [25] D. R. Bach, G. Flandin, K. J. Friston and R. J. Dolan, "Time-series analysis for rapid event-related skin conductance analysis," *Journal of Neuroscience Methods*, vol. 184, no. 2, pp. 224-234, 2009.
- [26] T. Chaspari, A. Tsiartas, L. I. Stein, S. A. Cermak and S. S. Narayanan, "Sparse Representation of Electrodermal Activity With Knowledge-Driven Dictionaries," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 3, pp. 960-971, 2015.
- [27] R. Edelberg, "Mechanisms of Electrodermal adaptations for locomotion, manipulation, or defense," *Progress in Physiological Psychology*, vol. 8, pp. 155-209, 1973.
- [28] D. C. Fowles, M. J. Christie, R. Edelberg, W. W. Grings, D. T. Lykken and P. H. Venables, "Publication Recommendations for Electrodermal Measurements," *Psychophysiology*, vol. 49, pp. 1017-1034, 2012.
- [29] M. S. Goodwin, W. Velicer and S. Intille, "Telemetric Monitoring in the Behavior Sciences," *Behavioral Research Methods*, vol. 40, pp. 328-341, 2008.
- [30] S. R. Green, P. A. Kragel, M. E. Fecteau and K. S. LaBar, "Development and Validation of an Unsupervised Scoring System (Autonomate) for Skin Conductance Response Analysis," *Journal of Psychophysiology*, vol. 91, pp. 186-193, 2014.
- [31] C. Kappeler-Setz, F. Gravenhorst, J. Schumm, B. Arnrich and G. Troster, "Towards long term monitoring of electrodermal activity in daily life," *Personal and ubiquitous computing*, vol. 17, no. 2, pp. 261-271, 2013.
- [32] M. Kelsey, A. Dallal, S. Eldeeb, M. Akcakaya, I. Kleckner, C. Gerard, K. S. Quigley and M. S. Goodwin, "Dictionary Learning and Sparse Recovery for Electrodermal Activity Analysis," in *SPIE Commercial Scientific Sensing and Imaging*, Baltimore, 2016.
- [33] C. L. Lim, C. Rennie, R. J. Barry, H. Bahramali, I. Lazzaro, B. Manor and E. Gordon, "Decomposing Skin conductance into tonic and phasic components," *Journal of Psychophysiology*, vol. 25, pp. 97-109, 1997.
- [34] V. Xia, N. Jaques, S. Taylor, S. Fedor and R. Picard, "Active Learning for Electrodermal Activity Classification," in *Signal Processing in Medicine and Biology Symposium*, 2016.