Sensor data fusion

Optimal Sensor Placement for Space–Time Potential Mapping and Data Fusion

Rui Zhu[1], Bing Yao[1], Hui Yang[1*], and Fabio Leonelli[2]

[1]Complex System Monitoring, Modeling and Analysis Laboratory, Pennsylvania State University, University Park, PA 16802 USA
[2]Cardiac Electrophysiology Lab, James A. Haley Veterans' Hospital, Tampa, FL 33620 USA
*Senior Member, IEEE

Manuscript received September 4, 2018; revised October 31, 2018; accepted November 20, 2018. Date of publication December 3, 2018; date of current version January 16, 2019.

Abstract—Current ECG imaging (ECGi) systems deploy a large number of ECG sensors to provide the high-resolution body surface potential mapping (BSPM). The availability of BSPM was shown to substantially improve the early detection of life-threatening heart disease. However, most existing ECGi systems employ an approximately uniform distribution of hundreds of ECG sensors on the body surface. Very little has been done to investigate the optimal sensor placement for BSPM. In this article, we propose a new optimal sensing strategy to search the optimal number and locations of the sensors. First, we develop a greedy algorithm to sequentially place ECG sensor on the body surface, which will maximize the information gain at each step. Second, we leverage the available BSPM data to develop a spatiotemporal model of cardiac electrical activity. Third, we study the algorithmic convergence and stopping criteria by evaluating diminishing return of the placement of two sequential ECG sensors. Experimental results show that the optimal sensing strategy with 30 sensors yields large R^2 statistics (>97%) for BSPM during the P, QRS, and T waves, as well as an average R^2 statistics of 97.71% for 12-lead ECG, and 99.44% for 3-lead VCG. The proposed methodology has strong potentials to help further improve the design of ECGi systems.

Index Terms—Body sensor network (BSN), electrocardiogram (ECG) imaging, greedy heuristics, location optimization, optimal sensor placement.

I. INTRODUCTION

Body sensor networks (BSNs) have emerged as a key technology to improve the quality of life by wearable sensing and health monitoring. Technological advances in nanomaterials and miniaturization of electronic devices make it possible to deploy a large number of sensors in BSN systems [1]. Indeed, new electrocardiogram imaging (ECGi) systems distribute a network of ECG sensors over the entire torso, thereby providing high-resolution body surface potential mapping (BSPM) signals [2]. Note that traditional ECG systems measure cardiac electrical activity with a limited number of leads and provide only time-domain views of ECG signals [3], [4]. ECGi systems shown in Fig. 1, however, enable the provision of a complete picture of spatiotemporal cardiac dynamics on the body surface.

Availability of spatiotemporal BSPM signals has been shown to facilitate mathematical reconstruction of the underlying cardiac electrical dynamics from spatiotemporal distribution of BSPMs over the entire torso. This, in turn, substantially improves the early detection of life-threatening cardiac disorders. ECGi offers an unprecedented opportunity to observe subjects with high risks of heart diseases beyond the confines of, often, high-end healthcare settings. Current ECGi systems involve a network of hundreds of sensors that are approximately uniformly distributed on the body surface (see Fig. 1), providing a high-resolution BSPM. A significant challenge lies in the use of a large number of sensors, which impact the user experience of ECGi systems to some extent. For acquisition of BSPM, the questions are whether the number of sensors can be reduced and how to better allocate these sensors on the body surface? The answers to these questions remain sketchy. Very little has been done to investigate sensor-placement problems that is, however, critical to improve the design of ECGi systems.

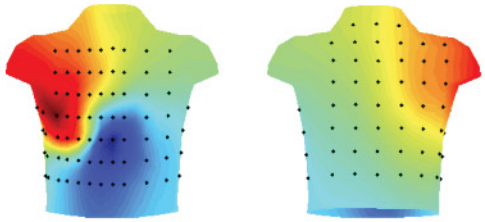

Fig. 1. Body area sensor network that uses 120 electrodes (i.e., black dots) [5] with an approximately uniform distribution on the body surface for BSPM. (a) Front view. (b) Back view.

In this article, we develop a new optimal sensor-placement strategy for the design of ECGi systems to capture a complete picture of spatiotemporal dynamics in cardiac electrical activity. The present investigation provides a viable solution that uses a sparse set of ECG sensors to realize high-resolution ECGi systems. In summary, our contributions in this paper are as follows.

1) Optimal placement of ECG sensors is introduced to improve the design of ECGi systems, rather than the approximately uniform placement used in current practice.

2) This investigation develops a spatiotemporal model of cardiac electrical activity on the body surface, as opposed to the traditional time-domain analysis of ECG signals.

3) We propose sequential design and diminishing return to optimally strike a balance between the number of ECG sensors and the

Corresponding author: Hui Yang (e-mail: huy25@psu.edu).
Associate Editor: Chun-An Chou.
Digital Object Identifier 10.1109/LSENS.2018.2884205

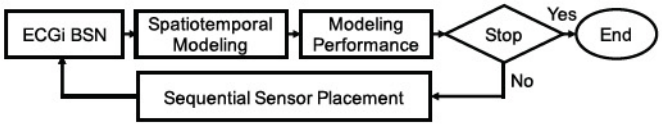

Fig. 2. Flowchart of the proposed optimal sensing strategy.

quality of BSPM data acquisition, as opposed to a large number of sensors (e.g., >100 electrodes) approximately uniformly distributed in existing ECGi systems.

II. RESEARCH METHODOLOGY

This article is aimed at improving the design of ECGi systems with optimal sensor placement. We are concerned about the quality of data acquisition (i.e., BSPM, ECG, and VCG) provided by ECGi systems if the number of sensors is reduced. As shown in Fig. 2, we propose to sequentially place ECG sensor on the body surface to maximize information gain in the spatiotemporal model of BSPM. If the model performance meets the satisfactory precision, the algorithm ends. If not, we continue to add more sensors to the ECGi system. The spatiotemporal model is then updated with the newly generated data and next iteration starts.

A. Spatiotemporal Modeling

The BSPM signals on the body surface are highly dynamic and varying with respect to time. Before the spatiotemporal modeling, we first introduce the kernel-weighted regression [6] to model the spatial distribution of BSPM at a specific time point as follows:

$$\Phi(s) = \sum_{i=1}^{I} \kappa_i(s) \cdot f_i^T(s) \cdot \omega_i + \tau(s) \tag{1}$$

where s represents the location on the body surface, $\Phi(s)$ is the BSPM value at location s, I denotes the total number of sensors, $f_i(s)$ is the set of basis functions $(f_{i1}(s), f_{i2}(s), \ldots, f_{ip}(s))^T$, where p is the number of basis functions, $\omega_i = (\omega_{i1}, \omega_{i2}, \ldots, \omega_{ip})^T$ is the model parameter with dimensionality $I \times p$, and $\tau(s)$ is the noise. We choose the basis functions as $f_i(s) = (1, x, y)^T$ in this investigation, where x and y are the coordinates of location s, but this choice does not preclude others to use a different form to model more complex structured spatial data. The kernel function $\kappa_i(s)$ is defined as follows:

$$\kappa_i(s) \propto |\Sigma_i|^{-\frac{1}{2}} exp\left\{\frac{(s-\mu_i)^T \Sigma_i^{-1}(s-\mu_i)}{2}\right\} \tag{2}$$

where μ_i is the center, and Σ_i is the covariance function of the ith kernel. In general, the selection of kernel function depends on the application [7]. As the Gaussian function is flexible to represent ECG, it is therefore chosen for BSPM data representation.

Let us denote the BSPM data at N locations on the body surface $\Phi = (\Phi(s_1), \ldots, \Phi(s_N))^T$, the vector of kernel weights $\kappa_i = (\kappa_i(s_1), \ldots, \kappa_i(s_N))^T$, and the ith basis functions $\eta_i = (f_{i1}(s), f_{i2}(s), \ldots, f_{ip}(s))^T$. If we define $\zeta = [\text{diag}(\kappa_1)\eta_1, \ldots, \text{diag}(\kappa_I)\eta_I]$ and $\omega = (\omega_1, \ldots, \omega_I)^T$, then (1) can be rewritten in a simplified matrix form as follows:

$$\Phi = \zeta\omega + \tau. \tag{3}$$

Based on the availability of BSPM data, the model parameter ω can be directly estimated by solving (3). However, this kernel regression model in (3) only captures the spatial pattern at a given time point. In order to model the time-varying BSPM signals, there is an urgent need to further add the component of temporal evolution to capture the spatiotemporal patterns of $\Phi(s, t)$.

The temporal evolution is modeled by making parameter ω_t change over time $t = 1, 2, \ldots, T$. At each sensor location, we obtain the observational data of BSPM and, at the same time, place a kernel function κ_i. As such, the spatiotemporal model is formulated to represent the dynamically evolving BSPM data as follows:

$$\Phi(s, t) = \sum_{i=1}^{I} \kappa_i(s, \theta_i) \cdot f_i^T(s) \cdot \omega_i(t) + \tau \tag{4}$$

where $\omega_i(t)$ is the time-varying model parameter, and $\theta_i = (\mu_i, \Sigma_i)$ gives the sensor location μ_i and covariance function Σ_i of the kernel function. Given an unobserved location s^* and a certain time point t^*, the BSPM value is estimated as $\hat{\Phi}(s^*, t^*) = \zeta_t^* \omega_t$. Now, the objective function is to identify a parsimonious set of sensors that achieve a satisfactory modeling performance of BSPM data as follows:

$$\arg\min_{I, \theta_i} \left[\|\Phi(s, t) - \sum_{i=1}^{I} \kappa_i(s, \theta_i) \cdot f_i^T(s) \cdot \omega_i(t)\|, \theta_i \right]. \tag{5}$$

B. Lazy Greedy Algorithm

As the search space of potential sensor locations is large, classic greedy algorithms were found to be computationally demanding. Therefore, we propose a lazy greedy algorithm to optimize the location of each sensor placement. Every time when a sensor θ from the search space is selected to join the set A, we compute its marginal benefit as $\delta(\theta) = E(A) - E(A \cup \theta)$, where $E(A) = \|\Phi - \sum_{i=1}^{N} \kappa_i(\theta_i) f_i^T \omega_i\|$. We then construct a priority heap structure that keeps track of marginal benefit of each sensor in the search space Θ. For each iteration, we only update marginal benefit for the top sensor θ^* in the heap as $\delta(\theta^*) = E(A) - E(A \cup \theta^*)$. If the updated marginal benefit of θ^* remains to be the largest, θ^* will be included in the selected set A as the next sensor. Otherwise, the priority queue of sensors is sorted again according to their marginal benefits, and the benefit of next top sensor in the heap is reevaluated. This avoids the evaluation of marginal benefits for all the remaining sensors in the heap to determine next optimal location. As such, the algorithm for optimal sensor placement is computationally efficient.

C. Diminishing Return

We evaluate the algorithmic convergence and stopping criteria in the optimal placement of ECG sensors with the use of diminishing returns. Adding a new sensor θ_{i+1} increases the coverage of ECGi BSN and decreases the error, i.e., $E(A \cup \{\theta_{i+1}\}) \leq E(A)$. Interestingly, there exists a turning point of diminishing returns if we keep increasing the number of sensors. In other words, adding a new sensor θ_{i+1} to a small set A_s yields more returns than a large set A_l, i.e., $A_l \supseteq A_s$: $E(A_s) - E(A_s \cup \{\theta_{i+1}\}) \geq E(A_l) - E(A_l \cup \{\theta_{i+1}\})$.

Therefore, we investigate a parsimonious set of sensors by sequentially maximizing marginal benefits, $\arg\max_{\theta_i \in \Omega \setminus A_{i-1}} E(A_{i-1}) - E(A_{i-1} \cup \{\theta_i\})$. Once the marginal benefit is smaller than a stopping criteria ε, the algorithm converges at the turning point of diminishing returns and we will then stop adding the sensors. The proposed methodology is evaluated and validated with experimental studies, which are detailed in Section III.

III. EXPERIMENTAL DESIGN AND RESULTS

In this article, we evaluate and validate the proposed methodology using real-world BSPM data from PhysioBank [8]. This ECGi system

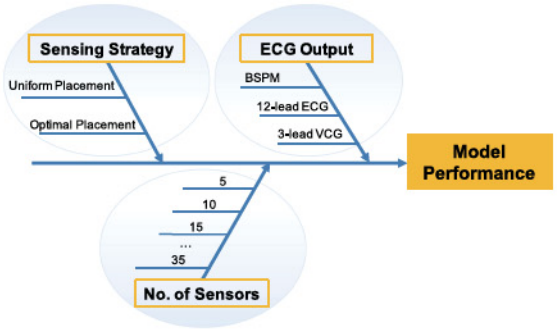

Fig. 3. Design of benchmark experiments.

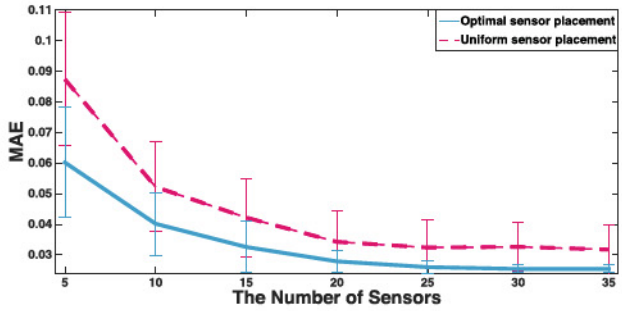

Fig. 4. Average MAE with standard deviation of model performances.

is the most commonly used setting with an approximately uniform distribution of 120 sensors. We hereby refer it as the *golden standard* for benchmark experiments. Note that ECGi systems allows for the collection of BSPM signals, which include both 12-lead ECG and 3-lead VCG. Therefore, as shown in Fig. 3, we design a three-way layout experiment to evaluate the performance of optimal sensor placement by comparing the ECGi outputs (i.e., BSPM, 12-lead ECG, and 3-lead VCG) with 1) uniform placement of the same number of sensors; and 2) the *golden standard* system. The first factor group is optimal versus uniform placement. The second factor is to vary the number of sensors from 5 to 35 in both uniform and optimal sensor-placement strategies, and then compare the third factor group of ECGi outputs from both strategies with the *golden standard* system.

A. Variation of Model Performance With Respect to the Number of Placed Sensors

We benchmark the performances of optimal and uniform placements of ECG sensors by evaluating average mean absolute error (MAE) and standard deviation of both ECGi systems, as follows:

$$\text{MAE} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}|\Phi(s_i,t) - \hat{\Phi}(s_i,t)| \tag{6}$$

where $\hat{\Phi}(s_i,t)$ represents the model performance, and i denotes the index of mapping sites on the body surface, $i = 1, 2, \ldots, N$. In other words, if we increase the number of ECG sensors from 5 to 35 in optimal and uniform placements, the differences in BSPM data quality are benchmarked with the *golden standard* system. The experiments are replicated on four different human subjects and, therefore, yield the standard deviation. Fig. 4 shows the variations of average MAE and standard deviation of optimal and uniform placements of ECG sensors. On the one hand, as the number of sensors increases from 5 to 35, the average MAE of optimal ECGi decreases monotonically from 0.0602 to 0.0254, and the standard deviation becomes significantly smaller. On the other hand, the average MAE of uniform ECGi descends from 0.0874 to 0.0318, and the standard deviation does not get smaller. Optimal placement of ECG sensors consistently yields a better performance than that by uniform placement with the increment in the number of sensors. The advantage of the proposed optimal sensor placement is mainly due to spatiotemporal modeling of real-world BSPM that help account for spatiotemporal patterns of electrical potentials rather than uniformly distributing sensors on the body surface.

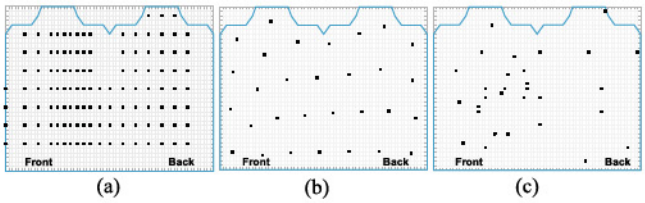

Fig. 5. (a) *Golden standard*. (b) Uniform ECGi BSN (30 sensors). (c) Optimal ECGi BSN (30 sensors).

B. Algorithmic Convergence and Diminishing Return

The optimal placement of ECG sensors aims to achieve a satisfactory measurement performance with a parsimonious set of sensors. In the experiment, we sequentially add sensors to ECGi BSN and place sensors at optimal locations determined by the lazy greedy algorithm. We use the "diminishing return" as the stopping criteria to compute when the marginal benefit $\delta(\theta^*) = E(A) - E(A \cup \theta^*) \leq \varepsilon$, where ε is set as 5×10^{-4}. In other words, if the MAE discrepancy between two subsequent iterations is less than $\varepsilon = 5 \times 10^{-4}$, the sequential addition of sensors will be stopped. Note that the stopping criterion is met when there are 30 sensors in the optimal ECGi system. In other words, optimal placement of 30 sensors is sufficient to achieve a satisfactory level of BSPM data quality as in the *golden standard* ECGi system. Fig. 5 shows the locations of ECG sensors on the body surface in the *golden standard* ECGi system, uniform placement of 30 sensors, and optimal placement of 30 sensors. It may be noted that most sensors in optimal ECGi tend to cluster around the location of heart on the front body. As shown in Fig. 5 (a) and (c), the number of ECG sensors is significantly reduced from 120 to 30 using the optimal sensing strategy, but still maintaining the quality of outputs (i.e., BSPM, ECG, and VCG). The benchmark of output data quality will now be detailed in Section C.

C. Comparison of ECGi Outputs From ECGi BSNs—Optimal, Uniform, and Golden Standard ECGi Systems

We use R^2 statistic to quantify the proportion of BSPM (or ECG, VCG) variations in the *golden standard* that can be explained by the optimal or uniform sensor placements. The R^2 statistic is defined as follows:

$$R^2 = \left(1 - \frac{T}{N}\cdot\frac{\sum_{t=1}^{T}[\Phi(t) - \hat{\Phi}(t)]^2}{\sum_{t=1}^{T}[\Phi(t) - \bar{\Phi}(t)]^2}\right) \times 100 \tag{7}$$

where $\Phi(t)$ are the outputs (i.e., BSPM, ECG, and VCG) from the *golden standard* ECGi system, T is the total number of time points, and N is the number of mapping sites. In this study, T is 1000, and N is set to 352 as in the golden standard ECGi system.

Fig. 6 (a) shows BSPMs at P wave measured with optimal and *golden standard* ECGi systems, respectively. Optimal ECGi yields an R^2 statistic of 98.35%. Note that if the R^2 statistic is 100%, then two systems do not have any differences in terms of measurement performance. Fig. 6 (b) demonstrates BSPM at QRS wave measured

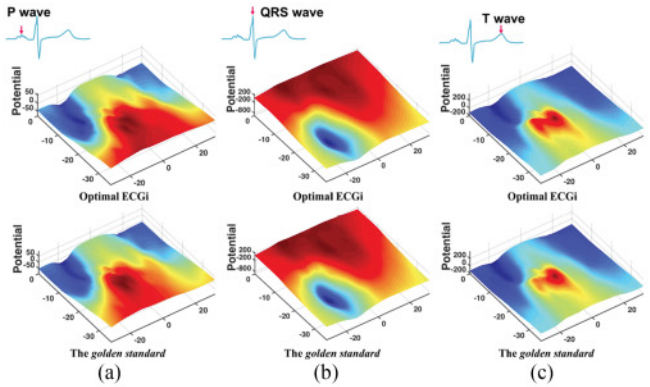

Fig. 6. BSPM measured by optimal ECGi BSN (top figure) and the golden standard (bottom figure) at (a) P wave ($R^2 = 98.35\%$), (b) QRS wave ($R^2 = 99.29\%$), and (c) T wave ($R^2 = 97.18$).

Table 1. R^2 Statistics of the Derived 12-Lead ECG from BSPM.

Leads	Optimal Sensor Placement			Uniform Sensor Placement		
	20	25	30	35	40	45
I	87.34	98.76	99.87	88.99	97.61	99.23
II	68.42	85.21	99.53	98.42	98.62	99.29
III	69.49	92.32	99.78	94.24	99.64	99.69
aVR	88.55	91.49	99.59	93.07	96.16	98.71
aVL	76.17	95.96	99.86	91.56	99.09	99.59
aVF	65.87	88.57	99.67	96.71	99.63	99.64
V1	87.52	93.46	93.78	76.98	77.18	89.85
V2	80.73	92.43	96.39	94.30	98.19	98.40
V3	89.51	93.85	98.12	29.70	71.93	90.25
V4	11.35	53.03	89.42	18.00	39.78	75.08
V5	70.24	88.59	98.10	86.72	89.78	95.94
V6	95.00	95.88	98.35	93.43	95.21	99.46
Total	890.19	1069.55	1172.47	962.11	1062.82	1138.12
Mean	74.18	89.13	97.71	80.18	88.57	94.84
Max	95.00	98.76	99.87	98.42	99.64	99.69
Min	11.35	53.03	89.42	18.00	39.78	75.08

with optimal ECGi BSN and the real-world measurement of BSPM at QRS wave with the *golden standard* ECGi system. Note that 99.29% of variations in BSPM measured by the *golden standard* system can be described by the optimal ECGi system. BSPMs at T wave measured by optimal and the *golden standard* ECGi systems are shown in Fig. 6(c). Optimal ECGi yields an R^2 statistic of 97.18% that means a good fit into the BSPM data measured by the *golden standard*. The large R^2 values (>97% for P, QRS, and T waves) statistically show that optimal ECGi is effective and efficient in capturing variations contained in BSPM with a parsimonious set of 30 sensors.

In this investigation, we derive the 12-lead ECG signals from optimal ECGi, uniform ECGi, as well as the *golden standard* ECGi systems. Table 1 shows the quantitative benchmark results between optimal ECGi, uniform ECGi, and the *golden standard* ECGi systems. The optimal ECGi yields mean R^2 statistics of 74.18% with 20 sensors, 89.13% with 25 sensors, and 97.71% with 30 sensors. It may also be noted that R^2 statistics increase for all 12 leads as more sensors are added, but the percentage of improvement is different because the lead locations are not the same. A mean R^2 statistics of 97.71% shows that the optimal ECGi with 30 sensors gives the 12-lead ECG as good as the *golden standard* ECGi system. Table 1 also shows the uniform ECGi yield's mean R^2 statistics of 80.18% with 35 sensors, 88.57% with 40 sensors, and 94.84% with 45 sensors. Note that we increase the number of sensors in uniform ECGi to reach comparable performance as the optimal ECGi. The mean R^2 is 80.18% for the uniform ECGi with 35 sensors, which is around 9% lower than that for the optimal ECGi with 25 sensors. When the number of sensors increases to 45,

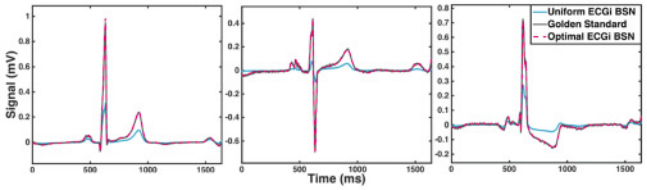

Fig. 7. Three-lead VCG signals from uniform and optimal ECGi (30 sensors), and the *golden standard* ECGi system. (a) X. (b) Y. (c) Z.

the mean R^2 reaches 94.84% for uniform ECGi, while the mean R^2 of optimal ECGi is 97.71%. Table 1 shows that the optimal placement of ECG sensors with a smaller set of sensors yields better results in 12-lead ECG signals than those of the uniform ECGi.

In addition, we benchmark the performance of three ECGi systems using the 3-lead VCG derived from BSPM signals. Fig. 7 shows the 3-lead VCG signals from uniform and optimal ECGi (30 sensors), and the *golden standard* ECGi system. Note that optimal ECGi produces nearly the same waveform as the golden standard. This is evident by comparing the goodness-of-fit of VCG signals in Fig. 7. As 12-lead ECG and 3-lead VCG are widely used by physicians for clinical decision making, optimal ECGi is effective due to the smaller set of sensors to produce BSPM signals, and the provision of commonly used ECG signals accurately.

IV. CONCLUSION

In this article, we propose an optimal sensing strategy for the placement of ECG sensors on the body surface to capture spatiotemporal cardiac electrical dynamics. Specifically, we integrate the sequential optimization of sensor placement with spatiotemporal modeling to improve the design of ECGi systems. The algorithmic convergence and stopping criteria are investigated by evaluating diminishing return of the placement of two sequential ECG sensors. Experimental results show that the optimal placement strategy with 30 sensors yields large R^2 statistics (>97%) for BSPM during the P, QRS, and T waves, as well as an average R^2 statistics of 97.71% for 12-lead ECG, 99.44% for 3-lead VCG.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation under Grant CMMI-1646660, and in part by the Harold and Inge Marcus Career Professorship.

REFERENCES

[1] B. H. Calhoun *et al.*, "Body sensor networks: A holistic approach from silicon to users," *Proc. IEEE*, vol. 100, no. 1, pp. 91–106, Jan. 2012.

[2] Y. Rudy and J. E. Burnes, "Noninvasive electrocardiographic imaging," *Ann. Noninvasive Electrocardiol.*, vol. 4, no. 3, pp. 340–359, 1999.

[3] H. Yang, S. T. Bukkapatnam, and R. Komanduri, "Spatiotemporal representation of cardiac vectorcardiogram (VCG) signals," *BioMed. Eng. OnLine*, vol. 11, no. 2, pp. 16–30, 2012.

[4] G. Liu and H. Yang, "Multiscale adaptive basis function modeling of spatiotemporal vectorcardiogram signals," *IEEE J. Biomed Health Inform.*, vol. 17, no. 2, pp. 484–492, Mar. 2013.

[5] B. Yao, R. Zhu, and H. Yang, "Characterizing the location and extent of myocardial infarctions with inverse ECG modeling and spatiotemporal regularization," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 5, pp. 1445–1455, Sep. 2018.

[6] Y. Chen and H. Yang, "Sparse modeling and recursive prediction of space-time dynamics in stochastic sensor networks," *IEEE Trans. Automat. Sci. Eng.*, vol. 13, no. 1, pp. 215–226, Jan. 2016.

[7] H. Yang, S. Bukkapatnam, and R. Komanduri, "Nonlinear adaptive wavelet analysis of electrocardiogram signals," vol. 76, Aug. 2007, Art. no. 026214.

[8] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.