33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Reconstruction of ECG Signals in Presence of Corruption

Gartheeban Ganeshapillai, Jessica F. Liu, and John Guttag

Abstract— **We present an approach to identifying and reconstructing corrupted regions in a multi-parameter physiological signal. The method, which uses information in correlated signals, is specifically designed to preserve clinically significant aspects of the signals.**

We use template matching to jointly segment the multi-parameter signal, morphological dissimilarity to estimate the quality of the signal segment, similarity search using features on a database of templates to find the closest match, and time-warping to reconstruct the corrupted segment with the matching template.

In experiments carried out on the MIT-BIH Arrhythmia Database, a two-parameter database with many clinically significant arrhythmias, our method improved the classification accuracy of the beat type by more than 7 times on a signal corrupted with white Gaussian noise, and increased the similarity to the original signal, as measured by the normalized residual distance, by more than 2.5 times.

I. INTRODUCTION

A modern Intensive Care Unit (ICU) employs multiple bedside monitors to track the state of the patients. They allow continuous monitoring of a patient, and inform medical staff of changes in the status of the patient. Automated analysis systems are typically used to analyze these signals in real-time. These systems critically depend on continuous uninterrupted real-time monitoring of the physiological signals such as electrocardiogram (ECG), arterial blood pressure (ABP), and photo-plethysmogram (PPG). Unfortunately, these signals are often severely corrupted by noise, artifacts, and missing data, which can result in a high incidence of missed detections and false alarms [1], [2], [3].

In this paper, we address the problem of identifying the corrupted regions in a multi-parameter signal, which is a set of multiple time-aligned synchronized correlated signals, and reconstructing them in a clinically useful way using the information available in correlated signals.

There are existing methods that exploit the information available in the correlated channels of a multi-parameter physiological signal to assist automated medical systems to produce results that are more reliable. For example, researchers have tried to fuse information from various ECG channels, and other signals to robustly estimate the heart rate [1], [4], [5]. Fusion of multiple signals often requires signal quality estimation, and researchers have developed several measures [6]: ECGSQI [4] provides the Signal Quality Estimates (SQE) of ECG signals, ABPSQI [7], [4] provides the SQE of ABP signals, and Hjorth parameters [8] are used to identify abnormal PPG pulses [5].

There have been studies [9] on the noise reduction on ECG signals using bandwidth filters to improve the SNR on the P, Q, and T waves without damaging the QRS complex [9]. Researchers have also attempted to identify the morphological features of ECG signals with added noise and abnormalities [10].

The problem of reconstructing a corrupted multi-parameter physiological signal was formally posed in the 11th annual PhysioNet/CinC challenge. Data collected from the MIMIC II project that contains records of ECG, APB, and PPG was used for the contest [11].

In our work, we focus on reconstructing the corrupted signals in clinically useful way, so that the automated systems that depend on these signals can produce results that are more reliable. In experiments carried out on the MIT-BIH Arrhythmia Database, a two-parameter database with many clinically significant arrhythmias, our method improved the classification accuracy of the beat type by more than 7 times on a signal corrupted with white Gaussian noise, and increased the similarity to the original signal, as measured by the normalized residual distance, by more than 2.5 times.

The organization of this paper is as follows. In Section II, we present our method and provide the mathematical framework of our work. In Section III, we discuss the measures of performance used to evaluate our method, and present the results of a series of tests in which comparisons are made using each of the performance measures. Finally, in Section IV, we discuss our work.

II. METHOD

We consider a multi-parameter signal represented by a matrix $\mathbf{S}_{n\times m}$, where each column represents a single signal (e.g., ECG) and each row represents a point in time. There are m synchronous correlated single parameter signals in $\mathbf{S}$. Each cell $s_{i,j}$ contains one sample. We assume that all the signals are sampled at the same rate. Our goal is to identify the corrupted regions, and estimate the actual sample values on that region.

We first identify the segment boundaries of the multi-parameter signal in the presence of significant amounts of transient corruption spanning multiple columns and rows of the matrix $\mathbf{S}$ using the method in [12].

We use a template, a short multi-parameter signal, and match it with a sliding window of the multi-parameter signal. The template is regularly updated to reflect the time evolution of the signal. The initial template is derived from an archived signal. We continuously extract non-overlapping

This work was supported by Quanta Computer, Inc.

G. Gartheeban (garthee@mit.edu), J. F. Liu (jessiliu@mit.edu), and J. Guttag (guttag@mit.edu), are with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

windows from $\mathbf{S}$, and identify the boundary in the window by finding the prefix of the window that most closely matches the template. The matching is done using weighted time warping (WTW) to minimize the weighted morphological dissimilarity across all the parameters. The warped distance between two signals gives the morphological dissimilarity. The weight represents the estimated quality of a single parameter signal in the multi-parameter signal. The signal quality estimated from the morphological dissimilarity is used to find the corrupted regions.

For reconstruction, we use a database of templates. Here, a template is a segment of the multi-parameter signal that was chosen from previously seen regions that were believed to be free of signal corruption. When we come across segments of high signal quality, we add them to the database; thus, we learn new morphologies.

The method is based on finding the closest match (template) to the corrupted segment from the database. We time-warp the template to fit the corrupted segment's interval, and replace the corrupted segment with the result. The closest match is found using the dynamic time warping (DTW) [13] cost. As a preliminary step, we represent the segments by features. This has the dual advantages of providing a level of abstraction that preserves clinically relevant information and speeding up the matching.

Goal : Let $\mathbf{S} \in \Re^{n\times 2}$ be a multi-parameter time series consisting of two single parameter physiological signals. The goal is to identify the corrupted segments $\{\mathbf{U}_i\}$, and reconstruct the samples in those segments.

Procedure : First, we detrend the signal, and remove baseline wander using a low pass filter[1]. Then, using a template $\mathbf{Z}_{\ell\times 2} = \{Z_j \in \Re^{\ell}\}_2$, we segment $\mathbf{S}$ into a set of quasiperiodic units, $\mathbf{S}_{[p_i,p_{i.+1})}$, that correspond to single heartbeats [12]. $\mathbf{S}_{[p_i,p_j)}$ denotes the window in the target sequence $\mathbf{S}$ from time $t = p_i$ to $t = p_j - 1$. We next run the reconstruction algorithm starting at the first segment $\mathbf{U}_1$, continuously evaluating the SQE of each segment and reconstructing those segments with an SQE below the threshold ζ_{low}. We add new segments to our database if their SQE is above the theshold ζ_{high}. This process is iterated over each of the segments. In the experiments reported later, we set ζ_{low} and ζ_{high} to 0.5 and 0.6 respectively.

An iteration : We start each iteration with a segment $\mathbf{U}_i = \mathbf{S}_{[p_i,p_{i.+1})}$ from $\mathbf{S}$. Using morphological dissimilarity, we determine whether the segment requires reconstruction.

If the SQE is below a threshold $q_i < \zeta_{low}$, we proceed with the reconstruction process. First, we build the feature representation F_i of the segment. The signal $\mathbf{S}$ is a 2-parameter signal. Hence, $\mathbf{U}_i$ contains two signals, and F_i is the joint representation of the both. We search the database, using F_i as the key, and find the top 20 matches. We find the best match on this set using the DTW distance (c_i) between the corruption-free channel of the segment $\mathbf{U}_i$ and the corresponding channel in the top matches. If the cost of the match is above a threshold $c_i > \kappa$, we abort the reconstruction process on the current segment, and continue to the next segment. Otherwise, we use the best match as a template for reconstructing the corrupted signal. We time-warp the channel V_i^a from the matching template $\mathbf{V}_i$, with the corruption-free channel U_i^a from the segment $\mathbf{U}_i$ to obtain the alignment w. Then, we replace the corrupted channel U_i^b of the current segment $\mathbf{U}_i$ with V_i^b*, which is obtained from the template $\mathbf{V}_i$ by warping V_i^b using the alignment w.

If the signal doesn't need reconstruction, i.e, if the SQE is above a threshold ($q_i > \zeta_{high}$), we build the feature representation F_i of the segment, and add the entry ($\mathbf{F}_i \Rightarrow \mathbf{U}_i$) to the database. If the length of the database exceeds a given limit, we evict the least recently used entry.

[1] http://www.mit.edu/~gari/CODE/FILTERS/

TABLE I

SET OF FEATURES THAT ARE USED TO REPRESENT A SEGMENT.

Feature	Description
$f_1 - f_4$	Pre, first-half, second-half, and post R-R intervals
f_5	Square root of the total energy
f_6 .. f_{15}	The fraction of the energy in the k^{th} section
f_{16}	Kurtosis of the sample values
f_{17}	DTW distance between the signal in the segment, and the median of the same signal
f_{18} .. f_{27}	DTW of k^{th} subsequence
f_{28}	Fraction of spectral energy in the QRS complex of the first signal in the segment
f_{29}	The maximum sample value
f_{30}	The minimum sample value

A. Feature Representation

By representing segments as feature vectors, we both achieve a level of abstraction that highlights physiologically significant aspects of the segments and greatly speeds up the matching process.

In reconstruction, we search the database for the closest match to the current segment. Since we want to do it in real-time with a large database of templates, we need to do this quickly. The use of features decreases the time complexity of a database lookup from $O(n.\ell^2)$ to $O(\ell^2)$, where n is the length of the database, and ℓ is the length of the sample.

Since the segments are usually of different lengths, a direct comparison function, such as Euclidean distance, is not suitable. On the other hand, variable length metrics such as DTW and longest common sub-sequence (LCSS) are typically of quadratic time complexity. Our feature representation represents a segment with a vector of fixed length c, hence two sequences can be compared in $O(c)$ time. The features also help avoid over-fitting.

Table I lists the set of features in the feature representation. The first row of the table contains the features related to R-R intervals. The remaining rows contain features of an individual segment. We choose this feature set to capture the morphological attributes and ECG specific characteristics.

B. Reconstruction

We want to reconstruct the corrupted channel U_i^b of the current segment $\mathbf{U}_i$ with the corresponding channel V_i^b from the replacement candidate $\mathbf{V}_i$.

We first verify the correctness of the match found. We accept the reconstruction only if the cost of the match c_i is less than a threshold. If the cost c_i is greater, we flag the segment $\mathbf{U}_i$ so that automated systems could avoid producing false alarms in those regions.

Since the length of the current segment $\mathbf{U}_i$, and the length of the candidate found (template) $\mathbf{V}_i$ are typically unequal, we next time-warp the template with the current segment. Time-warping is done by finding the optimal alignment $\phi(k)$ between the corruption-free channel of the current segment U_i^a and the corresponding channel of the template V_i^a (Equations 1-2).

$$\phi(k) = (\phi_1(k), \phi_2(k)), 1 \leq k \leq K \tag{1}$$

$$C_\phi(V_i^a, U_i^a) = \sum_{k=1}^{K} d(V_i^a[\phi_1(k)], U_i^a[\phi_2(k)]) \tag{2}$$

$$C(V_i^a, U_i^a) = \min_{\phi} C_\phi(V_i^a, U_i^a) \tag{3}$$

We then replace each sample of the corrupted channel $U_i^b[x]$ with the time-warped sample $V_i^b[x*]$, which is obtained from the median of the samples with which it is aligned.

$$x* = median(\phi_2(k)), 1 \leq k \leq K \text{ and } \phi_1(k) = x \tag{4}$$

III. Experimental Results

In our experiments, we use the multi-parameter ECG data from MIT-BIH Arrhythmia Database at Physionet.org [14]. The database has 48 ECG waveform records; each contains two channels and is 30 minutes long. The recordings were selected to include a variety of clinically significant arrhythmias. This helps us evaluate the robustness of our method. We use the 39 records from this set that are relatively free of significant corruption.

We add synthetic corruption to one channel, and then evaluate our method by quantifying the effectiveness of the reconstruction on this corrupted data.

We use the following criteria for comparison.

1) **Q1: Similarity:** We measure the similarity between the reconstructed data (S^b*), and the original uncorrupted data (S^b) by measuring the Euclidean residual distance r of the reconstructed data.

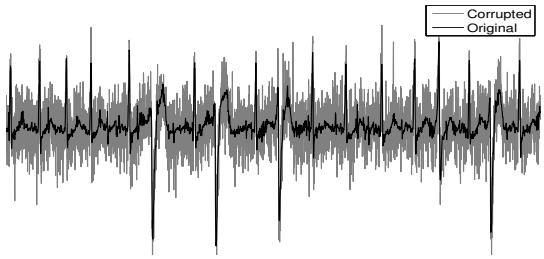
$$r = \sqrt{\frac{\sum_k^n (S^b*[k] - S^b[k])^2}{n \times \sigma_S^2}} \tag{5}$$

We normalize the Euclidean distance to make it comparable across the records. Residual distance is commonly used to measure the error in time-series analysis [15]. In addition, CinC challenge used $1 - r^2$ to evaluate the submissions [11].

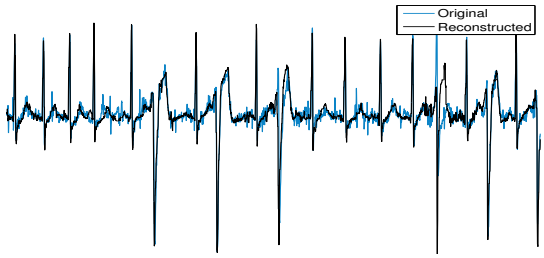
2) **Q2: Reproducibility:** Our goal is to enable the automated analysis systems produce more reliable results. Hence, we test our method's ability to improve the classification accuracy of a clinically relevant task. We run a widely used Premature Ventricular Contraction (PVC) detector[2] on the original data (S^b), the artificially corrupted data (S^b#) and the reconstructed data (S^b*), and record their agreements. If the PVCs are detected within 150 ms on two signals, we consider it an agreement. We quantify the ability to preserve the clinically relevant events by counting the disagreements. The number of disagreements $n_{\text{disagreement}}$ is evaluated between the original data (S^b), and the artificially corrupted data (S^b#), and between the original data (S^b), and the reconstructed data (S^b*). The disagreement Δ is finally expressed in terms of the fraction between the total number of disagreements $n_{\text{disagreement}}$, and the total number of beats n_{beats} in the region.

$$\Delta = n_{\text{disagreement}} / n_{\text{beats}} \tag{6}$$

[2] http://www.eplimited.com/software.htm



(a) Corrupted Vs Original



(b) Reconstructed Vs Original

Fig. 1. First channel of record 200 from MIT-BIH Arrhythmia database. The full segment of the channel is corrupted with AWGN at SNR 0dB (a), and reconstructed using our method (b).

TABLE II
Experiment 1

		Q1		Q2	
	PVC	$r_{S^b\#}$	r_{S^b*}	$\Delta_{S^b\#}$	Δ_{S^b}
Median	4	1.01	0.39	0.09	0
Average	47.64	1.01	0.40	0.14	0.02

A. Experiment 1: Effectiveness of Reconstruction

We build our database from the first 80% of each record and corrupt the last 20% of the first channel with the additive white gaussian noise (AWGN) at 0dB SNR.

Table II summarizes the results. It shows that our method reduces the residual distance (Q_1) by 2.5 times for a signal corrupted at SNR 0dB. Further, on average, it was able to improve the classification accuracy (Q_2) by more than seven fold. As an example, Fig. 1(b), shows the reconstruction on Record 200.

TABLE III
EXPERIMENT 2

SNR	r_{S^b*}	Δ_{S^b}
10 dB	0.410	0.031
0 dB	0.401	0.021
-10 dB	0.402	0.022

TABLE IV
EXPERIMENT 3

	r_{S^b*}	Δ_{S^b}
AWGN	0.410	0.031
EM	0.36	0.023
MA	0.19	0.003
BW	0.05	0.001

B. Experiment 2 : Different SNR levels

Again, we build our database from the first 80% of each record. We corrupt the last 20% of the first channel with AWGN at SNR levels of 10 dB, 0 dB, and -10 dB.

Table III summarizes the average disagreement (Δ), and the residual distance (r) for the reconstructed signal (S^b*). Somewhat surprisingly, we get the worst performance at the highest signal to noise ratio, and at low SNR levels the performance does not deteriorate with decreasing signal quality. The relatively poor performance at an SNR of 10 dB can be attributed to the fact that our algorithm makes a binary decision to reconstruct the signal or leave it as it is. This results in poor performance, because when the signal is only mildly corrupted our algorithm chooses to not attempt to reconstruct it.

C. Experiment 3: Simulated real-world corruptions

We alter the first 20% of the first channel with the following types of corruptions at SNR = 10 dB: Additive White Gaussian Noise (AWGN), Electromagnetic Interference (EM), Muscle Artifact (MA), and Baseline Wander (BW). We use MIT-BIH Noise Stress Test Database[3] and *nstdbgen*[4] to generate the non-Gaussian noise.

Table IV summarizes the average disagreement (Δ), and the residual distance (r) for the reconstructed signal (S^b*). We achieve the best performance for Baseline Wander. The worst performance was observed for AWGN, and EM noise.

IV. DISCUSSION

We presented a method for reconstructing a corrupted signal in a multi-parameter physiological signal using the information available in a correlated signal.

Using data from the MIT-BIH Arrhythmia Database, we conducted a series of experiments to test the effectiveness of our method. Our evaluation criteria were normalized residual distance and classification accuracy. For AWGN, our method improved the classification accuracy by more than 7 times, and increased the similarity to the original signal, as measured by the normalized residual distance by 2.5 times. Our algorithm is faster than real-time when it is run on a standard computer. We use feature vectors to achieve a running time that is independent of the size of the database.

Our method is useful only when the signals are time-aligned synchronized correlated and quasi-periodic. Further, we also require that at each point in time at least one of the correlated signals to be free of corruption.

While we have tested our method only on ECG data, we believe that it should be useful in other multi-signal settings in which one or more signals are corrupted and at least one of the correlated signals is transiently uncorrupted [16], [17]. Going forward, we plan to test our algorithm on a database containing simultaneous recordings of ECG, ABP, PPG, and CVP.

[3] http://www.physionet.org/physiobank/database/nstdb

[4] http://www.physionet.org/physiotools/wag/nst-1.htm

REFERENCES

[1] A. Aboukhalil, L. Nielsen, M. Saeed, R. Mark, and G. Clifford, "Reducing false alarm rates for critical arrhythmias using the arterial blood pressure waveform," *Journal of biomedical informatics*, vol. 41, no. 3, pp. 442–451, 2008.

[2] T. Heldt, B. Long, G. Verghese, P. Szolovits, and R. Mark, "Integrating Data, Models, and Reasoning in Critical Care," *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, pp. 350–353, 2006.

[3] H. Cao, L. Eshelman, N. Chbat, L. Nielsen, B. Gross, and M. Saeed, "Predicting ICU hemodynamic instability using continuous multiparameter trends," *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pp. 3803–3806, 2008.

[4] Q. Li, R. G. Mark, and G. D. Clifford, "Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter." *Physiological measurement*, vol. 29, no. 1, pp. 15–32, Jan. 2008.

[5] A. Deshmane, "False Arrhythmia Alarm Suppression Using ECG, ABP, and Photoplethysmogram," *M.S. thesis, Dept. Elect. Eng. Comp. Sci., MIT, Cambridge, MA*, 2009.

[6] M. Oefinger, "Monitoring transient repolarization segment morphology deviations in Mouse ECG," *Ph.D. dissertation, Dept. Elect. Eng. Comp. Sci., MIT, Cambridge, MA*, 2006.

[7] W. Zong, G. Moody, and R. Mark, "Reduction of false blood pressure alarms by use of electrocardiogram blood pressure relationships," *Computers in Cardiology 1999*, pp. 305–308, 1999.

[8] B. Hjorth, "The physical significance of time domain descriptors in EEG analysis." *Electroencephalography and clinical neurophysiology*, vol. 34, no. 3, pp. 321–325, Mar. 1973.

[9] H. Lee and Z. Z. Bien, "Estimation of instantaneous bandwidth and reconstruction of noisy ECG signal measured by a 24-hour continuous healthcare system for the elderly and people with disabilities," in *ICOST'10: Proceedings of the Aging friendly technology for health and independence, and 8th international conference on Smart homes and health telematics*. Springer-Verlag, jun 2010.

[10] F. Sufi and I. Khalil, "A New Feature Detection Mechanism and Its Application in Secured ECG Transmission with Noise Masking," *Journal of Medical Systems*, vol. 33, no. 2, pp. 121–132, jun 2008.

[11] G. Moody, "The physionet/computing in cardiology challenge 2010: Mind the gap," *Computing in Cardiology*, pp. 305–308, 2010.

[12] G. Gartheeban and J. Guttag, "Weighted Time Warping for Temporal Segmentation of Multi-parameter Physiological Signals," *Biostec*, 2011.

[13] C. Myers, L. Rabiner, and A. Rosenberg, "Performance tradeoffs in dynamic time warping algorithms for isolated word recognition," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 28, no. 6, pp. 623–635, 1980.

[14] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000, circulation Electronic Pages: http://circ.ahajournals.org/cgi/content/full/101/23/e215.

[15] "Structural Health Monitoring With Autoregressive Support Vector Machines," *Journal of Vibration and Acoustics*, vol. 131, pp. 1–9, jun 2009.

[16] A. Hartmann, "Reconstruction of missing cardiovascular signals using adaptive filtering," *Computing in Cardiology*, pp. 321–324, 2010.

[17] I. Silva, "PhysioNet 2010 Challenge: A robust multi-channel adaptive filtering approach to the estimation of physiological recordings," *Computing in Cardiology*, pp. 313–316, 2010.