Robust ECG Biometrics by Fusing Temporal and Cepstral Information

Ming Li, Shrikanth Narayanan
Signal Analysis and Interpretation Laboratory, University of Southern California
mingli@usc.edu, shri@sipi.usc.edu

***Abstract*—The use of vital signs as a biometric is a potentially viable approach in a variety of application scenarios such as security and personalized health care. In this paper, a novel robust Electrocardiogram (ECG) biometric algorithm based on both temporal and cepstral information is proposed. First, in the time domain, after pre-processing and normalization, each heartbeat of the ECG signal is modeled by Hermite polynomial expansion (HPE) and support vector machine (SVM). Second, in the homomorphic domain, cepstral features are extracted from the ECG signals and modeled by Gaussian mixture modeling (GMM). In the GMM framework, heteroscedastic linear discriminant analysis and GMM supervector kernel is used to perform feature dimension reduction and discriminative modeling, respectively. Finally, fusion of both temporal and cepstral system outcomes at the score level is used to improve the overall performance. Experiment results show that the proposed hybrid approach achieves 98.3% accuracy and 0.5% equal error rate on the MIT-BIH Normal Sinus Rhythm Database.**

***Keywords*-biometrics; Electrocardiogram; cepstral features; hermite polynomial expansion;**

I. Introduction

The ECG signal is an emerging novel behavior biometric for human identification. Individual differences in the heart structure, such as chest geometry, position, and size, manifest unique characteristics in their ECG signals which can be used as a biometric trait. Furthermore, the ECG signal can provide real-time liveness feedback. In many existing mobile health monitoring or body area sensor network systems, ECG is an important diagnostic tool for physiological measure, and thus subject verification based on the ECG signal itself can be useful for personalized health service. One disadvantage is that the ECG signal inherently varies at different heartbeats of the same subject due to variations in fitness, physical and emotional states as well as variabilities caused by sensor position changes and long term baseline shifts. Therefore, most previous works have used a sequence of heartbeats in rest condition to model healthy individuals. Compared to systems using ECG signals from multiple leads [1], promising results based on a single lead ECG signal have also been proposed in [2]–[6]. Thus, in this paper, our focus is on ECG biometrics for healthy people from a single lead sequential signal under rest condition.

With single lead ECG signals, most methods rely on the reduced feature sets derived from ECG characteristic points and subsequently supervised classification. Fiducial points based features are widely used in [2]–[5]. Moreover, since the fiducial points extraction algorithm is not robust for all possible types of ECG traces and since only a certain percentage of the population can be successfully enrolled [6], the principal component analysis (PCA) based approach [5], [6] was introduced to model each heartbeat at the appearance level. However, model scoring on a long sequence of heartbeats requires feature extraction and supervised classification for each heartbeat which becomes computationally expensive. In this paper, we consider robust alternatives for ECG signal representation and modeling.

In the time domain, after ECG signal pre-processing, we use Hermite polynomial expansion (HPE) to capture the ECG shape characteristics and adopt the SVM to perform supervised classification. Hermite polynomials are a classical orthogonal polynomial sequence which have been successfully used to describe the ECG signal [7] and thus can be employed to model the morphological differences for each heartbeat. Furthermore, the HPE projection matrix and linear kernel SVM scoring functions are combined together into a model vector, which makes the HPE feature extraction and sequential scoring just one simple inner product for computational efficiency.

In most previous works [3]–[6] and the proposed time domain system, the unit for modeling is each normalized heartbeat; thus, peak detection, heartbeat segmentation, and normalization are required as a pre-processing step due to the inherent heartbeat variability. If the pre-processing is not accurate or robust, this front end error can accumulate to influence the subsequent feature extraction and modeling steps. Moreover, the computational cost of this pre-processing step is also high. Therefore, algorithms that do not require this kind of pre-processing are preferred. Only a few works have addressed this problem. In [5], long term (5-seconds window) autocorrelation in conjunction with discrete cosine transform is adopted for feature extraction, while nearest neighbor matching is used as the classifier. The pre-processing step is avoided, but since the ECG signal within the long processing window is non-stationary, it may negatively influence the system performance. More recently, we adopted cepstral features extracted from single lead ECG signals combined with GMM to perform physical activity recognition task and demonstrated promising results [8]. Cepstral features, which reside in the homomorphic signal representation realm (inverse of log spectrum) are especially conducive for mitigating convolutional effects. Moreover,

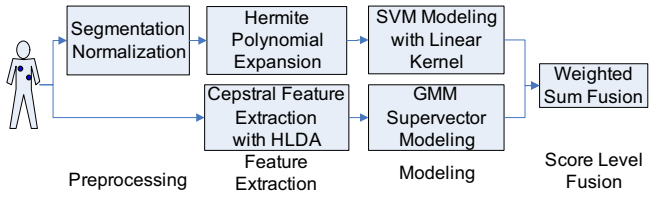

Figure 1. The proposed system overview

ECG cepstral feature calculation does not require heartbeat segmentation and normalization. This inspired us to explore the potential of using GMM to model cepstral domain ECG features from the viewpoint of robust biometrics. In the GMM framework, Heteroscedastic Linear Discriminant Analysis (HLDA) [9] and GMM Supervector (GSV) kernel [10] are used to reduce the cepstral feature dimension and to perform discriminative modeling, respectively. The system overview is shown in Figure 1. By fusing both temporal and cepstral system outcomes at the score level, the biometric system performance is further improved.

The paper organization is as follows: Sections II-IV describe the proposed methods, Section V provides the experimental results and Section VI is the conclusion.

II. SVM SYSTEM BASED ON TEMPORAL FEATURES

A. Signal pre-processing

The length of each heartbeat is different due to inherent heart rate variability. For mapping each heartbeat of the ECG waveform signal into a fixed length feature vector, normalization is performed for temporal feature extraction. In the ECG pre-processing step, each heartbeat waveform is normalized to the same time scale and amplitude scale by using standard methods in the ECG toolbox [6], [11].

B. Hermite polynomial expansion

In this work, the Hermite Polynomial Expansion is used to map the ECG shapes into feature coefficients which are further modeled by SVM. Hermite polynomials are classical orthogonal polynomial sequence representations and have been successfully used to describe the ECG signals as they successfully exploit existing similarities between the shapes of Hermite basis functions and the ECG waveforms [7]. Let us denote the ECG curve vector and polynomial order by $a(t)$ and L respectively. After pre-processing, each $a(t)$ has the same time scale length. The Hermite polynomial expansion can be written as follows [7]:

$$a(t) = \sum_{n=0}^{L-1} c_n \phi_n(t,\delta) \qquad t \in [-M, M] \tag{1}$$

where c_n $(n = 0, \cdots, L-1)$ are the HPE coefficients and $\phi_n(t,\delta)$ is the Hermite basis function defined as

$$\phi_n(t,\delta) = \frac{1}{\sqrt{\delta 2^n n! \sqrt{\pi}}} e^{-t^2/2\delta^2} H_n(t/\delta) \tag{2}$$

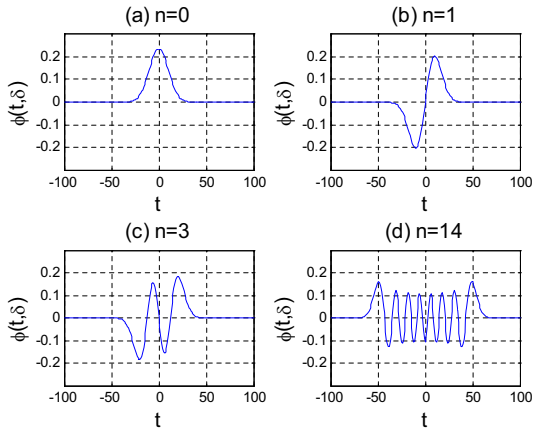

Figure 2. Hermite basis functions with $\delta = 10$ and $M = 100$.

The functions $H_n(t/\delta)$ are the physicists' Hermite polynomials which are defined recursively by:

$$H_0(t) = 1, H_1(t) = 2t \tag{3}$$

$$H_n(t) = 2tH_{n-1}(t) - 2(n-1)H_{n-2}(t) \tag{4}$$

Figure 2 presents several Hermite basis functions for different orders, where it can be observed that the higher the order, the higher is its frequency of changes within the time domain, and thus resulting in a better capability for capturing the morphological details. The HPE basis functions can be denoted by a $(2M+1)\times L$ matrix $H = [\underline{\phi}_0 \underline{\phi}_1 \cdots \underline{\phi}_{L-1}]$, and then the expansion coefficients $\underline{C} = [c_0 c_1 \cdots c_{L-1}]^T$ can be obtained by pseudo-inverse to minimize the sum squared error E:

$$\begin{aligned} E &= \left\| a(t) - \sum_{n=0}^{L-1} c_n \phi_n(t,\delta) \right\|_2^2 = \left\| \underline{a} - H\underline{C} \right\|_2^2 \\ \underline{C} &= (H^T \cdot H)^{-1} \cdot H^T \cdot \underline{a} \end{aligned} \tag{5}$$

C. SVM Classification

In this work, the linear kernel is employed for SVM modeling. Thus, if we arbitrarily add one dummy dimension with value 1 at the head of each SVM input feature vector, the scoring function on a sequence of ECG heartbeats $\underline{A} = \{\underline{a}_i\}_{i=1\cdots N_a}$ is defined as follows:

$$\begin{aligned} f(\{\underline{A}\}) &= w^t \cdot \frac{1}{N_a} \sum_{i=1}^{N_a} \underline{C}(\underline{a}_i) \\ &= w^t \cdot \frac{1}{N_a} \sum_{i=1}^{N_a} ((H^T \cdot H)^{-1} \cdot H^T \cdot \underline{a}_i) \\ &= \{w^t \cdot (H^T \cdot H)^{-1} \cdot H^T\} \cdot \{\frac{1}{N_a} \sum_{i=1}^{N_a} \underline{a}_i\} \\ &= W^t \cdot \{\frac{1}{N_a} \sum_{i=1}^{N_a} \underline{a}_i\} \end{aligned} \tag{6}$$

Where $\underline{C}(\underline{a}_i)$ is the HPE coefficients of $\underline{a}_i$ and w is the collapsed linear kernel SVM scoring model which is defined in [12]. Therefore, by collapsing all the support vectors and the HPE matrix down into a single model vector W, the scoring function of a target model on a sequence of heartbeats can be calculated by just an inner product between the model vector and the averaged heart beats which makes this framework very computationally efficient.

III. GMM SYSTEM BASED ON CEPSTRAL FEATURES

A. Cepstral feature extraction

The ECG signal has quasi-periodic characteristics as a result of the convolution between the excitation (heart beat rate) and system response (ECG waveform shapes) [11]. Thus, in this work, we use cepstral features [8], [13] that allow us to separate such convolutive effects by simple linear filtering to model the frequency information of the native ECG. Furthermore, linear frequency bands are used for simplicity. Cepstral mean substraction (CMS) and cepstral variance normalization (CVN) are adopted for achieving measurement robustness. Finally, HLDA [9] is adopted to perform dimensionality reduction.

B. GMM and GSV modeling

GMM is used to model the cepstral features of the ECG signals. In this biometric task, each subject is represented by a GMM. Since the training data for each subject is too limited to train a good GMM, the Universal Background Model (UBM) using a Maximum A Posteriori (MAP) adaptation approach [13] is used to model each individual in a supervised manner while combating data sparsity.

Recently, a combination of both GMM and SVM was successfully applied by using GMM supervector kernels in speaker verification and language identification tasks [10]. The GSV approach in this work consists of support vector machines with GMM supervectors, which is a concatenation of the GMM mean vectors, as input features for the ECG biometric applications. A linear kernel based upon an approximation to KL divergence between two GMM models was derived in [10]. The GMM supervector can be considered as a mapping between a segment of sensor signal and a high-dimensional SVM input vector.

In our work, for each segment of time, a GMM is adapted from the UBM by MAP adaptation; the GMMs were modeled with diagonal covariance matrices and only the means of GMMs were adapted. The linear kernel is defined as the corresponding inner product [10]:

$$K(\mu, \hat{\mu}) = \sum_{i=1}^{M} (\sqrt{p_i}\Sigma_i^{-\frac{1}{2}}\mu_i)(\sqrt{p_i}\Sigma_i^{-\frac{1}{2}}\hat{\mu}_i)^t \qquad (7)$$

where p_i and $\sum_i$ are the i^{th} UBM mixture weights and diagonal covariance matrix; μ_i corresponds to the mean of the i^{th} Gaussian component in this GMM. Since this kernel is linear, we can apply the model compaction technique mentioned in [10], [12]. Therefore we only have to compute a single inner product between the target model and the GMM supervector to obtain a score.

IV. SCORE LEVEL FUSION

Let there be K input ECG biometric subsystems, each acting on a specific feature set, where the k^{th} system outputs its own normalized log-likelihood vector $l_k(x_t)$ for every trial. The fused log-likelihood vector is:

$$\acute{l}(x_t) = \sum_{k=1}^{K} \alpha_k l_k(x_t) \qquad (8)$$

The weight α_k is determined by logistic regression based on the training data [14].

V. EXPERIMENT RESULTS

To evaluate the performance of the proposed methods, we conducted our experiments on the MIT-BIH Normal Sinus Rhythm Database [15]. It contains 18 ECG recordings from different subjects collected at the Arrhythmia laboratory of Boston's Beth Israel Hospital. The subjects included in the database did not exhibit significant arrhythmias. The sample frequency of each ECG recording is 128Hz. Since the duration of each recording is more than 20 hours, we simply used the data from hour 1 to hour 2 as the training data and the data from hour 2 to hour 3 as the testing data. The experimental results in this paper are based on a 10 seconds testing duration. The empirically-chosen HPE order, M, δ, and the GMM gaussian components number are 60, 100, 10, and 32, respectively. The configuration of cepstral feature extraction used in this paper is 32 cepstral coefficients, 32 frequency bands and 500 milliseconds window with 50% overlap with first order delta which is reported to have good performance in [8]. HLDA is employed to reduce the cepstral feature dimension from 64 to 40. Before cepstral feature extraction, a 0.05Hz - 40Hz FIR bandpass filter was applied to reduce both high frequency noises and low frequency baseline fluctuations. Both close-set identification accuracy P_c and Equal Error Rate (EER) of Detection Error Tradeoff curves are used to evaluate the proposed methods.

In Table I and Figure 3, the results of cepstral feature based GMM framework including HLDA and GSV algorithms are shown. By using HLDA to perform dimension reduction, the results improved from 93.96% to 94.78%. Furthermore, GSV modeling which combines both GMM and SVM yielded the best cepstral domain result of 95.90%. Therefore, the experiment results demonstrate that an ECG based biometric method without the pre-processing steps of heartbeat segmentation and fiducial detection is effective.

Table I also shows the results of the SVM system based on HPE coefficients with linear kernel. The proposed temporal HPE features in conjunction with SVM achieved 98.11% accuracy. Compared to the result of 89.46% for the PCA

Table I
PERFORMANCE IN P_c AND EER.

ID	Pre-Process	Methods	P_c (%)	EER (%)
1	no	Cepstral: GMM	93.96	4.30
2	no	Cepstral: GMM+HLDA	94.78	4.05
3	no	Cepstral: GSV+HLDA	95.90	2.5
4	yes	Temporal: HPE+SVM	98.11	0.55
5	yes	Fusion of system (3+4)	98.26	0.50
6	yes	EigenPulse: PCA [6]	89.46	13.2

ID and Pre-Process denote System ID (in Figure 3) and ECG pre-processing (peak detection, heartbeat segmentation and normalization), respectively.

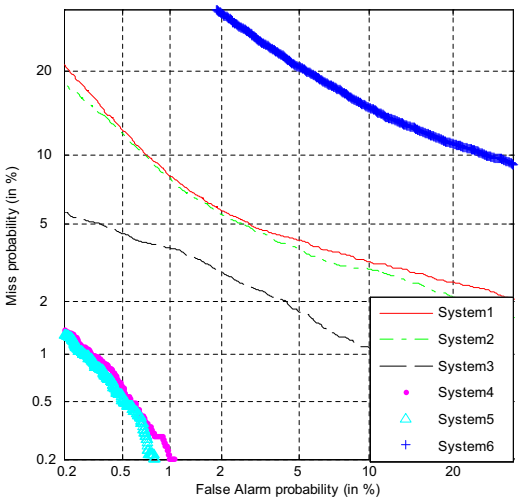

Figure 3. Performance in DET curves

approach in [6] (In our implementation, the first 50 eigenvectors corresponding to the largest eigenvalues are used), the proposed discriminative SVM modeling based on HPE features achieved significant improvement. When fusing both temporal and cepstral domain system outcomes together at the score level, the performance is further improved.

VI. CONCLUSION

In this work, a novel ECG biometrics algorithm is proposed. In the time domain, HPE and SVM are used to efficiently model the intra-heartbeat patterns of different individuals with linear kernel. In the frequency domain, without the need of heartbeat segmentation and normalization, cepstral feature extraction is combined with GMMs to directly model the short time characteristics of the ECG signal. Cepstral features provide a natural way for minimizing convolutive effects by linear filtering. And by fusing both temporal and cepstral information together, the overall biometric system performance is improved. Future work would include validating the performance over a much larger population of test subjects, comparing/combining with other biometrics modalities, and investigating the robustness against a variety of physical and emotional state variabilities.

REFERENCES

[1] G. Wübbeler, M. Stavridis, D. Kreiseler, R. Bousseljot, and C. Elster, "Verification of humans using the electrocardiogram," *Pattern Recognition Letters*, vol. 28, no. 10, pp. 1172–1175, 2007.

[2] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG analysis: a new approach in human identification," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 3, pp. 808–812, 2001.

[3] S. Israel, J. Irvine, A. Cheng, M. Wiederhold, and B. Wiederhold, "ECG to identify individuals," *Pattern Recognition*, vol. 38, no. 1, pp. 133–142, 2005.

[4] J. Irvine and S. Israel, "A Sequential Procedure for Individual Identity Verification Using ECG," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, 2009.

[5] Y. Wang, F. Agrafioti, D. Hatzinakos, and K. Plataniotis, "Analysis of human electrocardiogram for biometric recognition," *EURASIP journal on Advances in Signal Processing*, vol. 1, pp. 1–6, 2008.

[6] J. Irvine, S. Israel, W. Todd Scruggs, and W. Worek, "eigenPulse: Robust human identification from cardiovascular function," *Pattern Recognition*, vol. 41, no. 11, pp. 3427–3435, 2008.

[7] L. Sornmo, P. Borjesson, M. Nygards, and O. Pahlm, "A method for evaluation of QRS shape features using a mathematical model for the ECG," *IEEE Transactions on Biomedical Engineering*, pp. 713–717, 1981.

[8] M. Li, V. Rozgić, G. Thatte, S. Lee, A. Emken, M. Annavaram, U. Mitra, D. Spruijt-Metz, and S. Narayanan, "Multimodal physical activity recognition by fusing temporal and cepstral information," *to appear IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 4, 2010.

[9] N. Kumar and A. Andreou, "Heteroscedastic discriminant analysis and reduced rank HMMs for improved speech recognition," *Speech communication*, vol. 26, no. 4, pp. 283–298, 1998.

[10] W. Campbell, D. Sturim, D. Reynolds, and A. Solomonoff, "SVM based speaker verification using a GMM supervector kernel and NAP variability compensation," in *Proc. ICASSP*, vol. 1, 2006.

[11] G. Clifford, F. Azuaje, and P. McSharry, *Advanced methods and tools for ECG data analysis*. Artech House, 2006.

[12] W. Campbell, J. Campbell, D. Reynolds, E. Singer, and P. Torres-Carrasquillo, "Support vector machines for speaker and language recognition," *Computer Speech & Language*, vol. 20, no. 2-3, pp. 210–229, 2006.

[13] D. Reynolds, T. Quatieri, and R. Dunn, "Speaker verification using adapted Gaussian mixture models," *Digital Signal Processing*, vol. 10, pp. 19–41, 2000.

[14] A. Ross, K. Nandakumar, and A. Jain, *Handbook of multibiometrics*. Springer, 2006.

[15] "MIT-BIH Normal Sinus Rhythm Database," *http://www.physionet.org/physiobank/database*.