ECG Biometric Identification Using Wavelet Analysis Coupled with Probabilistic Random Forest

Robin Tan
Department of Electrical and Computer Engineering
Portland State University
Portland, Oregon, USA
Robin.Tan98@gmail.com

Marek Perkowski
Department of Electrical and Computer Engineering
Portland State University
Portland, Oregon, USA
h8mp@pdx.edu

***Abstract*—A novel algorithm is proposed in this study for improving the accuracy and robustness of human biometric identification using electrocardiograms (ECG) from mobile devices. The algorithm combines the advantages of both fiducial and non-fiducial ECG features and implements a fully automated, two-stage cascaded classification system using wavelet analysis coupled with probabilistic random forest machine learning. The proposed algorithm achieves a high identification accuracy of 99.43% for the MIT-BIH Arrhythmia database, 99.98% for the MIT-BIH Normal Sinus Rhythm database, 100% for the ECG data acquired from an ECG sensor integrated into a mobile phone, and 98.79% for the PhysioNet Human-ID database acquired from multiple tests within a 6-month span. These results demonstrate the effectiveness and robustness of the proposed algorithm for biometric identification, hence supporting its practicality in applications such as remote healthcare and cloud data security.**

Keywords—ECG; Biometric Recognition; Random Forest; Wavelet Analysis; Wavelet Coefficients; Cloud Data Security; Remote Health Care

I. INTRODUCTION

The rapid advancement of the integration of medical sensors into mobile devices has facilitated the real-time monitoring of one's physiological conditions, such as electrocardiograms (ECG), heart rate, respiration rate, oxygen saturation, blood pressure, body temperature, body impedance, blood sugar, and motion artifacts [1-2]. This new technology enables an imminent transition to a remote healthcare system where people can real-time monitor their health and wellness during work, fitness, travel, and even patient care at home while remotely sending information to medical centers for diagnosis and treatment [3]. To ensure the reliability and effectiveness of such a telemedicine system, an accurate subject identification is critical since medical doctors must ensure that the data received remotely are associated with the correct patient, as assigning the right medical treatment to the right patient is of vital importance.

Biometric recognition uniquely identifies individuals based on their features derived from physiological or behavioral characteristics. Conventional biometric recognition techniques utilize retinal structures, fingerprints, face and voice for access control. However, each of these techniques exhibits weaknesses associated with confidentiality and robustness against spoofing attacks [4]. As today's social networking system is experiencing rapid growth, personal data and internet security become more and more of a pressing concern. To meet the demand for a highly secure remote data communication and remote access control, ECG signals have been studied over the last decade as a promising biometric trait for subject identification because they are nearly impossible to forge. And more recently, ECG signals are readily available from mobile devices with no added cost.

To date, considerable efforts have been made to develop various algorithms using ECG for biometric recognition [5-8]. The ECG signal feature extraction methods can be generally categorized as fiducial and non-fiducial based. The fiducial based method relies on an accurate detection of fiducial ECG landmarks to achieve a high identification accuracy. The disadvantage is that the onset and offset fiducial ECG landmarks are very susceptible to noise and baseline drifting, causing errors in feature extraction. The non-fiducial based method analyzes the entire ECG complex after each heartbeat segmentation and alignment, using time and frequency analysis such as Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) [9-12]. The advantage of using non-fiducial features is that it obviates the need for rigorous detection of fiducial points. However, the identification accuracy based on the entire ECG waveform varies depending on the choice of classification methods such as k-nearest neighbor, linear discriminant analysis (LDA) and cross-correlation. In addition, it was reported that the identification accuracy tends to worsen and the computation load increases significantly as the subject size becomes too large [5].

Despite the aforementioned progress, numerous challenges remain to be resolved before ECG signals can be used for practical applications with mobile devices. To overcome the difficulties of reliable feature extraction when mobile ECG signals suffer from environmental impairments, a novel algorithm is proposed that utilizes both fiducial and non-fiducial ECG features, and combines probabilistic random forest with wavelet distance measurement to achieve a high identification accuracy while maintaining a low computation load.

II. PROPOSED ECG IDENTIFICATION SYSTEM

The ECG measures the electrical activity of the heart by placing electrodes on the surface of body. Fig. 1 shows a typical ECG P-QRS-T complex, where P, Q, R, S, T, P_{on}, P_{off}, QRS_{on}, QRS_{off}, T_{on}, and T_{off} points are referred to as ECG fiducial landmarks according to the standard cardiographic nomenclature [13], while T_{R-R} represents the ECG heart rate.

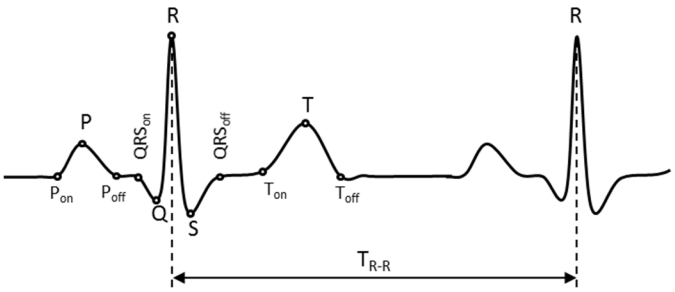

Fig. 1. Typical ECG P-QRS-T complex and fiducial points

Unlike the conventional 12-lead ECG clinical setup, the ECG signals measured from 1-lead mobile devices suffer worse signal-to-noise ratio, baseline artifacts caused by human respiration or movement of fingers while contacting electrode pads, as well as power line interference. These impairments lead to an inaccurate feature extraction due to the detection errors around ECG onset and offset fiducial points (P_{on}, P_{off}, QRS_{on}, QRS_{off}, T_{on} and T_{off}) as they are most susceptible to noise and baseline wandering. To overcome those difficulties, Fig. 2 presents the proposed two-stage identification algorithm for human biometric identification using a mobile ECG.

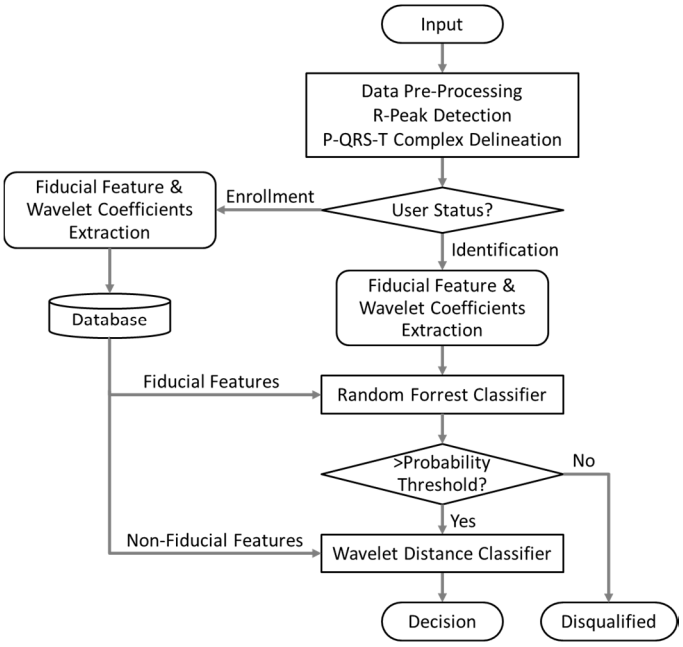

Fig. 2. Proposed algorithm for ECG subject identification

This classification system operates in two stages: the enrollment phase and the identification phase. During the enrollment phase, the system receives the ECG data from each individual. After data pre-processing and R-peak detection, the system computes its ECG fiducial features based on limited characteristic points, as well as its non-fiducial features using wavelet coefficients, and stores the feature set of each individual into a database. During the identification phase, the ECG signal of an unknown subject is received. The system again extracts the ECG fiducial features from this unknown subject to be applied to the random forest classifier. Meanwhile, the system derives the non-fiducial features based on wavelet analysis and compares them with the stored wavelet coefficients of a few candidate subjects probabilistically selected from the random forest classifier. In comparison to the methods from previous literature, the proposed algorithm pertains the following unique advantages:

1) *Extracts ECG features based on P, R, T peaks and Q, S valleys only, eliminating potential detection errors caused by using onset and offset points.*
2) *Uses the probabilistic scheme of the random forest classifier to obtain the probability of the unknown subject being the i^{th} subject in the database (i = 1,2…N, where N is the total number of subjects registered in the database). The algorithm selects a few candidate subjects whose probability must be higher than an optimized threshold and sends the subjects' data to the subsequent template matching classifier for further decision.*
3) *For better accuracy the wavelet coefficients are obtained from an optimized range of wavlet decomposition based on DWT.*
4) *Uses 1-to-S template matching instead of 1-to-N (where S<<N), thus maintaining a low computation load and eliminating previously reported performance degradation caused by excessive subject size.*

To evaluate the robustness of the proposed algorithm, the following four datasets were collected:

- ECG data from 30 subjects measured from a mobile phone. This database evaluates the real-life applicability of the algorithm using actual mobile ECG data.
- MIT-BIH Arrhythmia database from PhysioNet [14]. All R-peaks have already been well annotated by medical staff. It also allows the algorithm to test classification accuracy using ECG data from patients with abnormal heart conditions.
- MIT-BIH Normal Sinus Rhythm database from PhysioNet. This database tests the performance for people with normal heart rhythm.
- ECG-ID database from PhysioNet. Data was collected from 89 subjects ranging from 13-75 years of age, with each subject recorded the ECG data 2-20 times over a period of 6 months.

III. METHODOLOGY

A. Data Pre-Processing and R-peak Detection

The received raw ECG data is first applied to a bandpass filter to eliminate baseline wandering, high frequency noise and power line interference. Using Fast Fourier Transform (FFT), the bandwidth of the band-pass filter is set to 2Hz ~ 50Hz, where most of the ECG signal energy is located. The filtered signal in frequency domain is then converted into time domain using Inverse Fast Fourier Transform (IFFT) before it is applied for R-peak detection.

Accurate R-peak detection is critical to ensure that each P-QRS-T complex is correctly delineated. The commonly accepted and well-known algorithm is the one proposed by Pan and Tompkins [15-16], where the input ECG signal goes through a series of signal processing procedures such as differentiation, squaring, and moving windows integration, followed by a decision with an adaptive threshold. To make a compromise between R-peak detection accuracy and implementation complexity, a simple peak detection algorithm is applied with a moving window filter to identify the local maxima. An adaptive threshold similar to that used in Pan and Tompkins algorithm is then implemented to filter out the false R peaks (i.e. the T-wave). The adaptive threshold helps to improve the R-peak detection accuracy and true positive rate, as it effectively deals with the ECG R-peak amplitude variations over time. Using the MIT-BIH arrhythmia database as an example, the R-peak detection performance is compared between the Pan and Tompkins algorithm and this algorithm under the same simulation environment. Three performance metrics are used: accuracy, true positive rate, and computational time:

$$\text{Accuracy (\%)} = \frac{\text{TP}}{(\text{TP+FN})} * 100 \quad (1)$$

$$\text{True Positive Rate (\%)} = \frac{\text{TP}}{(\text{TP+FP})} * 100 \quad (2)$$

Here, the true positive (TP), false positive (FP), and false negative (FN) are defined in Table I.

TABLE I. TP, FP AND FN DEFINITION

Annotation	R-peak (+)	Non-peak (-)
R-peak (+)	TP	FN
Non-peak (-)	FP	TN

The performance results of the two algorithms are given in Table II, which shows that the R-peak detection algorithm can achieve a similar performance as the Pan and Tompkins algorithm, but the computation time is significantly reduced.

TABLE II. R-PEAK DETECTION PERFORMANCE COMPARISON

Methods	Accuracy (%)	True Positive Rate (%)	Computational Time (s)
Pan-Tompkins	99.56	99.01	2.83
This Work	98.70	99.46	0.53

B. ECG Complex Delineation

Once the R-peak of the ECG signal is correctly located, the corresponding P-QRS-T complex is delineated through time windowing for each detected R-peak. To further improve the feature extraction accuracy, some of the P-QRS-T outliers are removed by calculating the Pearson correlation coefficients for each P-QRS-T complex against the mean complex calculated from a 3D array. The distribution of the correlation coefficients is examined, as shown in Fig. 3. After some experimentation, a threshold is set to (μ-0.5*σ), where μ is the mean value of the correlation coefficient probability density function and σ is the standard deviation. If the correlation coefficient of a P-QRS-T complex falls below the threshold, it is considered as an outlier and is removed.

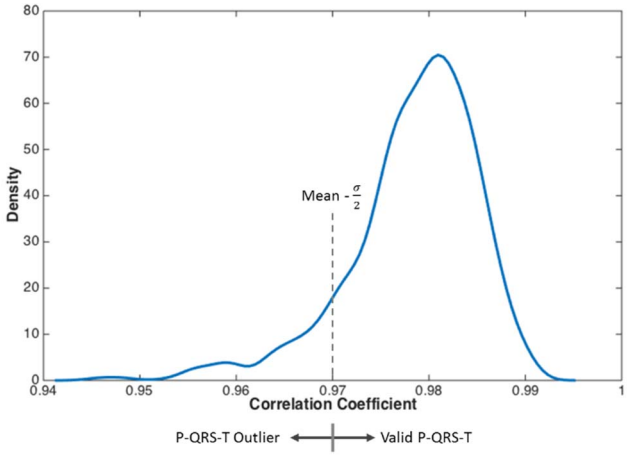

Fig. 3. Probability density of the correlation coefficients

After P-QRS-T delineation, an L number of complexes are averaged to further improve the signal to noise ratio, where L varies for each database depending on the total number of ECG complexes recorded for each subject.

C. Fiducial Feature Extractions

As mentioned earlier, it is difficult to detect the ECG P-wave, QRS-complex and T-wave onset and offset boundaries due to noise and baseline variations over time. Using the proposed two-stage classifiers, the intention is to utilize only a limited number of fiducial characteristic points, such as P, R, T peaks and Q, S valleys for feature extraction, which can be detected by a local maximum/minimum search within a defined physical region. For performance comparison, the results from the random forest classifier are also investigated using only Q, R, S fiducial points, as well as using 9 fiducial points including P_{on}, P_{off}, T_{on}, T_{off} points. The onset and offset fiducial points are detected using a simple triangle optimization method described in [9]. The fiducial features are then extracted based on relative temporal, amplitude and angle attributes, as illustrated in Fig. 4.

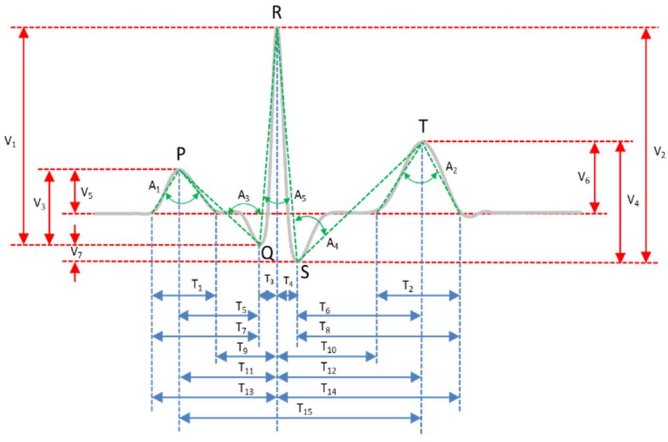

Fig. 4. Feature extraction based on fiducial points

Table III summaries the extracted fiducial features for three cases mentioned above. Each case is applied to the random forest classifier for performance comparison.

TABLE III. SUMMARY OF EXTRACTED FIDUCIAL FEATURES

Fiducial Points	Temporal	Amplitude	Angle	Total Features
P, Q, R, S, T, P_{on}, P_{off}, T_{on}, T_{off}	T_1 to T_{15}	V_1, V_2, V_3, V_4, V_5, V_6, V_7	A_1 to A_6	28
P, Q, R, S, T	T_3 to T_6, T_{11},T_{12}, T_{15}	V_1, V_2, V_3, V_4, V_7	A_3 to A_5	17
Q, R, S	T_3, T_4	V_1, V_2, V_7	A_5	6

It should be noted that both amplitude and angle attributes do not change much with one's heart rate. But the Q-T temporal interval attribute varies with respect to the heart rate (i.e., the sporting). Therefore, the Q-T temporal is corrected according to the Framingham formula in [17]:

$$QT_{corrected} = QT + 0.154\ (1\text{-}T_{R\text{-}R}) \tag{3}$$

Where $T_{R\text{-}R}$ is the time interval between adjacent R peaks. In this way, our proposed system can work for people with varied heart rates.

D. Random Forest Classifier

After feature extraction, the following data matrix is generated and applied to random forest classifier:

$$A_{MK}^{N} = \begin{bmatrix} a_{11}^{1} & \cdots & a_{1K}^{1} \\ \vdots & \ddots & \vdots \\ a_{M1}^{1} & \cdots & a_{MK}^{1} \end{bmatrix} \begin{bmatrix} a_{11}^{2} & \cdots & a_{1K}^{2} \\ \vdots & \ddots & \vdots \\ a_{M1}^{2} & \cdots & a_{MK}^{2} \end{bmatrix} \vdots \vdots \begin{bmatrix} a_{11}^{N} & \cdots & a_{1K}^{N} \\ \vdots & \ddots & \vdots \\ a_{M1}^{N} & \cdots & a_{MK}^{N} \end{bmatrix} \tag{4}$$

K = numbe of features; M = number of complexes; N = number of subjects

Here, *N* represents the number of subjects in the database, *K* represents the number of features extracted as given in Table III and *M* represents the number of complexes for each subject after removing P-QRS-T outliers and averaging. Out of the total number of complexes for each subject, 67% randomly chosen complexes are used for training in the random forest model, while the remaining 33% are used for testing.

A random forest is composed of a combination of decision tree predictors. Each decision tree (*t)* uses bootstrapping technique to resample the training datasets and to generate a vector of classification probabilities p_t^i for the unknown input ECG with respect to the i^{th} subject stored in the database. The sum of the ensemble output from all decision trees is obtained by:

$$P^i = \frac{1}{T}\sum_{t=1}^{T} p_t^i,\ \ i=1, 2, \ldots N \tag{5}$$

Where *T* is the total number of decision trees used in the random forest algorithm and N is the total number of subjects stored in the database. Only a few candidate subjects *S*, whose probability is above an optimized probability threshold P_{th}, will be applied to the subsequent *1-to-S* template matching classifier for further decision. After random forest, most of the classified probabilities P^i are close to zero and only few (1~3) are bigger than 0.1. Therefore, the selection of P_{th} can be varied from 0.1 to 0.2, based on the data used. In this study, the P_{th} is set to be 0.15. In addition, our preliminary results indicated that final classification results are insensitive to the selection of P_{th} ranged from 0.1 to 0.2.

E. 1-to-S Template Matching using Wavelet Coefficients

Wavelet analysis provides a useful method for decomposing the original time series data. The wavelet coefficients at each level of decomposition give detailed information for their corresponding frequency sub-bands, particularly, providing higher temporal resolution at higher frequency sub-band than lower frequency sub-band. The unique characteristics of the wavelet analysis offers a strong benefit in analyzing the detailed ECG timing waveform at a multi-scale resolution associated with the frequency range of each sub-band. In this study, the wavelet coefficients of the ECG signal are derived using DWT at each level of decomposition for template matching classification. The relationship between the ECG signal sampling rate (F_s), and the p^{th} level decomposed sub-band frequency range (F_p) is derived as:

$$\frac{F_s}{2^{p+1}} \le F_p \le \frac{F_s}{2^p} \tag{6}$$

Taking the MIT-BIH arrhythmia database (F_s = 360Hz) as an example, Table IV shows the frequency range (F_p) and the number of wavelet coefficients when p=1 to 5. Because the ECG frequency spectrum is located between 2Hz~50Hz, it is expected that the optimized wavelet coefficients should be obtained from D2 to D5 in this example.

TABLE IV. FREQUENCY SUB-BAND & NUMBER OF WAVELET COEFFICIENTS

Decomposition Level (D_p)	Decomposed Sub-band Frequency Range Fp (Hz)	Number of Wavelet Coefficients
D1	90 to 180	144
D2	45 to 90	72
D3	22.5 to 45	36
D4	11.25 to 22.5	18
D5	6.75 to 11.25	9
A5	<6.75	9

Fig. 5 illustrates the time-domain waveforms of the original ECG signal and its decomposed waveforms for levels D1 to D5 and A5.

Unlike a conventional *1-to-N* template-matching classifier where the feature vector of an unknown subject is compared against that of all N subjects in the enrolled database, the proposed algorithm only needs to implement a *1-to-S* template-matching based on wavelet coefficients, where S is the reduced number of candidates from the random forest classifier based on their probabilities. Assuming the stored wavelet coefficient vector for the i^{th} subject is represented as $\left(x_1^i, x_2^i \ldots\ x_P^i\right)$, and an

unknown subject j has a feature vector $(x_1^j, x_2^j, \ldots x_P^j)$, where p is the level of decomposition ($p = 1, 2 \ldots P$), the template-matching classifier calculates the wavelet distance by:

$$WDIST(i) = \sum_{p=1}^{P} \frac{\|x_p^i - x_p^j\|}{max(|x_p^i|)} \quad (7)$$

The smallest *WDIST(i)* (*i=1..S)* indicates the final identified subject.

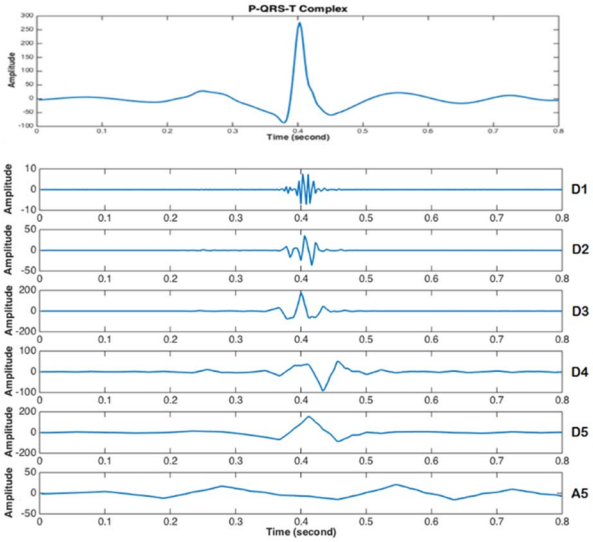

Fig. 5. ECG complex decomposed time waveform

IV. RESULTS

To evaluate the proposed algorithm in this study, it is assumed that each ECG complex represents one subject. The identification accuracy is defined as:

$$Accuracy = \frac{\#\ of\ correctly\ identified\ ECG\ complexes}{\#\ of\ total\ ECG\ complexes} \quad (8)$$

For model parameter optimization, the identification accuracy is first evaluated over T number of trees used in the random forest classifier and various Daubechies wavelet functions used in DWT. Fig. 6 illustrates the accuracy vs number of trees in the random forest classifier, and Fig. 7 illustrates the accuracy vs the Daubechies function from dB3 to dB8. According to these results, T = 100 and the dB3 functions are chosen in the final algorithm.

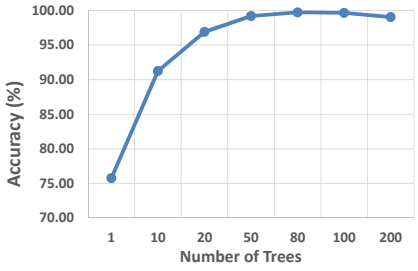

Fig. 6. Accuracy vs number of random forest trees

To obtain the best overall system performance, the machine learning classifier and wavelet distance classifier programmatically optimize their parameters and combinations of feature sets.

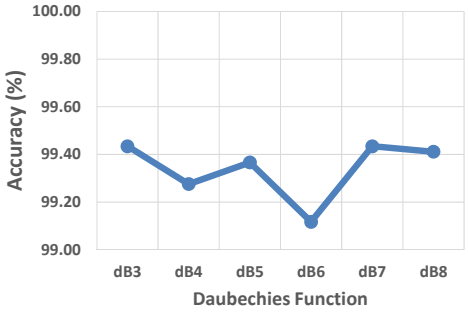

Fig. 7. Accuracy vs Daubechies functions

A. *Identification Accuracy using Random Forest Only*

Table V shows the accuracy of using only a random forest classifier when features are extracted from various numbers of fiducial points:

TABLE V. ACCURACY OF USING RANDOM FOREST ONLY

Algorithm	Dataset			
	Mobile	MIT-Arrhythmia	MIT-Normal	Human ID
RF-3	99.02	98.98	98.39	95.02
RF-5	99.67	99.77	99.98	93.54
RF-9	99.67	99.32	100	90.71

Here, RF-3 represents the results of using random forest with only 6 features from Q, R, S points; RF-5 represents the results using 17 features from P, Q, R, S, T points; RF-9 represents the results using 27 features from P_{on}, P_{off}, P, Q, R, S, T, T_{on}, T_{off} points. On average, RF-5 gives a better identification accuracy among the four datasets. Therefore, RF-5 is used in the proposed algorithm.

B. *Identification Accuracy Using 1-to-N Template Matching Only*

The identification accuracy of the template matching classifier is investigated using wavelet coefficients obtained from different levels of wavelet decomposition. The results are shown in Table VI.

TABLE VI. ACCURACY OF USING WAVELET DISTANCE ONLY

Algorithm	Dataset			
	Mobile	MIT-Arrhythmia	MIT-Normal	Human ID
WDIST1-5	99.67	98.33	98.89	91.79
WDIST2-5	99.35	98.62	95.31	96.23
WDIST3-5	99.02	98.57	91.1	98.52

Here, WDIST1-5 represents the accuracy using wavelet coefficients from decomposition levels D1 to D5; WDIST2-5 represents wavelet coefficients from D2 to D5; WDIST3-5

represents wavelet coefficients from D3 to D5. In general, WDIST2-5 gives better performance among the four datasets and therefore is used in the proposed hybrid classification system.

C. Identification Accuracy using Proposed Algorithm

Using the proposed algorithm, Table VII gives the identification accuracy when the random forest classifier (RF-5) is coupled with the wavelet distance measurement classifier (WDIST2-5). For performance comparison, the results from each single stage classifier are also presented.

TABLE VII. ACCURACY USING COMBINED RANDOM FOREST AND 1-TO-S TEMPLATE MATCHING

Algorithm	Dataset			
	Mobile (%)	MIT-Arrhythmia (%)	MIT-Normal (%)	Human ID (%)
WDIST2-5	99.35	98.62	95.31	96.23
RF-5	99.67	99.77	99.98	93.54
Proposed Method	100	99.43	99.98	98.79

From Table VII, it is clearly indicated that the proposed two-stage classification algorithm achieves an overall higher identification accuracy than each individual classifier. The amount of improvement depends on the quality of ECG data acquired. For example, when the ECG data is measured multiple times over a 6-month span, the proposed algorithm achieves an accuracy of 98.79%, better than 93.54% from a single random forest classifier and 96.23% from a single 1-to-N template matching classifier. In this study, the classification accuracy is demonstrated assuming that one ECG complex represents one subject. In reality, the test subject could easily record multiple heartbeats in one instance to be used for identification. Under this assumption, the subject identification rate would perform better than what has been simulated in this study according to the binomial theorem.

V. CONCLUSION

The newly proposed subject identification system demonstrates that using a probabilistic approach combining random forest and wavelet distance measure provides a better identification accuracy than that obtained from each single classifier. The proposed algorithm exhibits remarkable identification accuracies for ECG signals obtained from mobile devices, ECG signals of patients with heart conditions, as well as ECG signals acquired multiple times over a long term span. Due to the use of the two-stage classifier, the proposed algorithm allows to use only a limited number of fiducial features, eliminating the complexity and instability of the detection of ECG onset and offset points. Using a 1-to-S template matching instead of 1-to-N (S<<N), the computational load is significantly reduced. More importantly, the proposed solution avoids the potential performance degradation as the subject size N becomes too large. The results demonstrate the accuracy and robustness of the proposed algorithm, hence affirm the possibility to use ECG as a reliable biometric modality for subject recognition using wearable devices.

REFERENCES

[1] Patel, S., Park, H., Bonato, P., Chan, L., and Rodgers, M. (2012). A review of wearable sensors and systems with application in rehabilitation. Journal of Neuroengineering and Rehabilitation, 9:21.

[2] Boulos, M. N. K., Wheeler, S., Tavares, C., & Jones, R. (2011). How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. Biomedical engineering online, 10(1), 1.

[3] Jain, A., Bolle, R., & Pankanti, S. (Eds.). (2006). Biometrics: personal identification in networked society (Vol. 479). Springer Science & Business Media.

[4] Biggio, B., Akhtar, Z., Fumera, G., Marcialis, G.L., and Roli, F.: "Security Evaluation of Biometric Authentication Systems under Real Spoofing Attacks", IET Biometrics, doi: 10.1049/iet-bmt.2011.0012, March 2012.

[5] Odinaka, I., Lai, P. H., Kaplan, A. D., O'Sullivan, J. A., Sirevaag, E. J., & Rohrbaugh, J. W. (2012). ECG biometric recognition: A comparative analysis. IEEE Transactions on Information Forensics and Security, 7(6), 1812-1824.4.

[6] Nawal, M., & Purohit, G. N. (2014). ECG based human authentication: A review. Int. J. Emerg. Eng. Res. Technol, 2(3), 178-185.

[7] Mazomenos, E.B., Biswas, D., Acharyya, A., Chen, T., Maharatna, K., Rosengarten, J., Morgan, J. and Curzen, N., 2013. A low-complexity ECG feature extraction algorithm for mobile healthcare applications. IEEE journal of biomedical and health informatics, 17(2), pp.459-469.

[8] Maglaveras, N., Stamkopoulos, T., Diamantaras, K., Pappas, C., & Strintzis, M. (1998). ECG pattern recognition and classification using non-linear transformations and neural networks: a review. International journal of medical informatics, 52(1), 191-208.

[9] Plataniotis, K. N., Hatzinakos, D., & Lee, J. K. (2006, September). ECG biometric recognition without fiducial detection. In 2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference (pp. 1-6). IEEE.

[10] Y.Wang, F. Agrafioti, D. Hatzinakos, and K. N. Plataniotis (2008). Analysis of human electrocardiogram for biometric recognition. Journal on Advances in Signal Processing, Article ID 148658.

[11] Daamouche, A., Hamami, L., Alajlan, N., & Melgani, F. (2012). A wavelet optimization approach for ECG signal classification. Biomedical Signal Processing and Control, 7(4), 342-349.

[12] Chan, A. D., Hamdy, M. M., Badre, A., & Badee, V. (2008). Wavelet distance measure for person identification using electrocardiograms. IEEE transactions on instrumentation and measurement, 57(2), 248-253.

[13] Fratini, A., Sansone, M., Bifulco, P., and Cesarelli, M.: "Individual Identification via Electrocardiogram Analysis", Biomedical Engineering Online, 14:78, doi: 10.1186/s129380150072y, August 2015.

[14] The MIT-BIH Database, http://www.physionet.org/physiobank/database/nsrdb/

[15] Pan, J., Tompkins, W.J., "A Real-Time QRS Detection Algorithm", IEEE Transactions on Biomedical Engineering, Vol BME-32, Issue 3, pp. 230 – 236, March, 1985.

[16] Amin Hashim, Chia Yee Ooi, "Comparative Study of Electrocardiogram QRS Complex Detection Algorithm on Field Programmable Gate Array platform" 2014 IEEE Conference on Boimedical Engineering and Science, pp241-246.

[17] "QT Interval", LifeInTheFastLine, http://lifeinthefastlane.com/ecg-library/basics/qt_interval/