Investigation of Human Identification using Two-Lead Electrocardiogram (ECG) Signals

Can Ye[1,2], Miguel Tavares Coimbra[2] and B.V.K. Vijaya Kumar[1]

[1] *Department of Electrical & Computer Engineering, Carnegie Mellon University, USA*
[2] *Instituto de Telecomunicações, Faculdade de Ciências da Universidade do Porto, Portugal*

Email: *cany@ece.cmu.edu, mcoimbra@dcc.fc.up.pt, kumar@ece.cmu.edu*

Abstract -- In this paper, we investigate the applicability of Electrocardiogram (ECG) signals for human identification. Wavelet Transform (WT) and Independent Component Analysis (ICA) methods are applied to extract morphological features that appear to offer excellent discrimination among subjects. The proposed method is aimed at the two-lead ECG configuration that is routinely used in long-term continuous monitoring of heart activity. The information from the two ECG leads is fused to achieve improved subject identification. The proposed method was tested on three public ECG databases, namely, MIT-BIH Arrhythmias Database [1], MIT-BIH Normal Sinus Rhythm Database [2] and Long-Term ST Database [3], in order to evaluate the proposed subject identification method on normal ECG signals as well as ECG signals with arrhythmias. Excellent rank-1 recognition rates (as high as 99.6%) were achieved based on single heartbeats. The proposed method exhibits good identification accuracies not just with the normal ECG signals, but also in the presence of various arrhythmias. This work adds to the growing evidence that ECG signals can be useful for human identification.

Key Words: Electrocardiogram (ECG), Biometrics, Wavelet Transform, Independent Component Analysis

I. Introduction

Biometrics-based human identification is receiving increasing research attention. Traditional biometrics modalities include fingerprint, iris and face. However, these traditional recognition technologies suffer from some limitations, such as vulnerability to spoofing, limited standoff range and the need for subject cooperation [4]. Other emerging biometric modalities are being investigated to address these limitations and complement traditional biometrics. Electrocardiogram (ECG) signal is one of them.

ECG signals have been extensively studied and used in clinical diagnosis for many years. It has only recently been suggested that ECG signals could be used as a biometric for human identity recognition [5]-[9]. One of the advantages of ECG signals as biometrics is that it provides an automatic liveness indication, significantly decreasing the possibility of spoofing. Also, in some recent efforts [10][11], ECG sensor-equipped jackets are being proposed to monitor human stress levels and automatically identifying a subject wearing such a jacket can ensure that the health data being monitored from that jacket corresponds to the correct individual, without resorting to cumbersome login solutions. Most of the ECG biometrics work reported made use of only one lead and ignored the other leads that may contain additional information [6]-[8]. On the other hand, some ECG-based biometrics research was based on the conventional 12-lead configuration [5][9], which is impractical in conditions where convenience and portability are important. In this work, we target the two-lead ECG configuration, which is routinely used in Holter Monitoring [12] that is widely accepted as the setting for long-term continuous heart monitoring using portable devices. The main goal of this paper is to investigate the identification accuracy (IDA) that can be achieved using two-lead ECG in different scenarios.

In previous ECG biometrics studies [5]-[9], the sizes of the subject pools varied from 5 to 21 subjects and almost all the work focused on the ECG signals collected from healthy subjects. In this paper, we propose to characterize single heartbeats using a set of morphological features that appear to not only provide an excellent representation of between-individual difference, but also robust to the presence of cardiac arrhythmias. Besides, the proposed method was tested on three different public ECG databases, consisting of 18, 47 and 65 subjects respectively, larger than in the previous studies.

The rest of this paper is organized as follows. We briefly introduce the theoretical framework in Section II and describe the details of the databases and the experimental methodology in Section III. Results are presented and discussed in Section IV. Finally, conclusions are provided in Section V.

II. Theoretical Framework

A. Electrocardiogram (ECG) Fundamentals

Electrocardiogram (ECG) recordings are quantitative measures of the electrical activity of a person's heart over time [13]. Figure 1 shows the characteristic ECG waveform in one heart cycle. It basically consists of P wave, QRS wave and T wave. The QRS wave is usually called QRS composite or QRS complex and it contains the most significant information in that heartbeat. For example, physicians would classify the heartbeat into normal or arrhythmias according to the shape of QRS complex. On the other hand, it is believed that the QRS complex could also provide characteristic information that is useful for identifying subjects. Hence the QRS complex is the part that should be emphasized in our analysis. The location of the R peak indicates the occurrence of a heartbeat. The interval between R peaks of two successive heartbeats is defined as RR interval, which is a measure of heart rhythm.
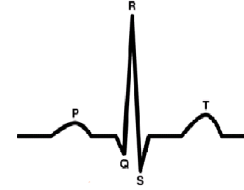

Figure 1: A schematic ECG waveform corresponding to one heart cycle (adapted from [13]).

Different parts of one heartbeat cycle are caused by different heart activities. The P wave is a result of atrial depolarization and reflects atrial contraction that pumps the blood into ventricles. The QRS wave is caused by ventricular contraction that pumps the blood to the lung and the rest of the body, whereas the T wave represents ventricular repolarization that is a necessary recovery process.

The two-lead ECG configuration is routinely used in the Holter monitoring that enables the recording of the heart activity of a subject continuously over long time with portable devices [12]. The three public ECG databases used in this study also adopted the setting of the Holter monitoring and the two-lead configuration. The ECG signals from the two leads are essentially two observations of the same physiological activity from two different perspectives.

The two leads are typically placed such that their orientations are orthogonal to each other, which suggests that the two leads may offer complementary information. Although more complicated fusion strategies are possible, we use a very simple fusion strategy. If the classification results from two leads are not identical, the heartbeat is rejected (i.e., no subject classification decision is made) since we believe that the information of the heartbeat is not reliable enough for the classifier to make a correct decision. As a result, the classification confidence is improved.

B. Wavelets and Multi-Resolution Analysis

Since we propose to use wavelet transform for extracting features, we provide a short summary of the relevant material. For any function $f(x) \in L^2(R)$, wavelet function $\Psi(x)$ and the corresponding scaling function $\varphi(x)$, we can define its wavelet series expansion as

$$f(x) = \sum_k c_{j_0}(k)\varphi_{j_0,k}(x) + \sum_{j=j_0}^{\infty} \sum_k d_j(k)\Psi_{j,k}(x) \quad (1)$$

where j_0 represents an arbitrary starting scale, j is the index of any scale higher than j_0. Usually, $c_{j_0}(k)$ and $d_j(k)$ are referred to as approximation coefficients and detail coefficients. The first sum provides a coarse approximation of $f(x)$ at scale j_0 and the second sum contains the details of the signal. For each higher scale $j \geq j_0$ in the second sum, a finer resolution is added to the approximation to provide increased detail.

The multi-resolution analysis nature of discrete wavelet transform (DWT) is suitable for characterizing the energy distribution of non-stationary signals, such as ECG signals. We choose Daubechies wavelets of order 8 due to their similarity with most characteristic QRS waveforms. For the MIT ECG database, the sampling frequency of ECG signals is 360 Hz, indicating that the highest possible frequency presented in the digitized ECG signals is 180 Hz. It has been shown that most of the energy of QRS complex is in the range between 0.5 and 40 Hz [14]. After applying the 4-level wavelet decomposition, this frequency range corresponds to the detail coefficients at level 3 and 4 as well as the approximation coefficients at level 4 that are selected as the wavelet features to characterize the shape of the heartbeat waveform.

C. Independent Component Analysis

Another tool used for feature extraction is Independent Component Analysis (ICA). ICA is a method originally proposed to solve the blind source separation (BSS) problem, which aims to separate mixture signals into a set of underlying independent sources [15].

ICA assumes that underlying sources are statistically independent and non-Gaussian. ICA can be formulated as

$$x = A \cdot s \quad (2)$$

where x represents the N observed signals, A is referred as the mixture matrix and s contains the M sources. Each observation is modeled as a linear combination of

underlying sources. The underlying independent components are estimated via maximizing some metric quantifying the statistical independence.

The use of ICA for ECG signal analysis has been justified in [16][17]. Atrial activity (AA) and ventricular activity (VA) are generated by independent physiological sources; both of them exhibit non-Gaussian distribution given the sub-Gaussian statistical character of AA as opposed to the super-Gaussian behavior of VA. Following the literature [11], 18 ICA bases are trained using all heartbeats in the training dataset in each case that was investigated.

Figure 2 shows the 18 ICA bases derived for the MIT-Arrhythmias Database and the histograms of corresponding ICA coefficients. From the magnitudes of ICA coefficients, we can see that heart beat is mostly represented by components 2, 4, 7 (circled in red). These three ICA components seem to be well matched to VA and AA, considering that VA presents high values within QRS complex and low values in the rest of the cardiac cycle whereas AA could be modeled as a sawtooth signal consisting of a sinusoid with several harmonics [16][17]. In this sense, the information of the physiological activity underlying the ECG wave could be extracted, which is characteristic of a subject.
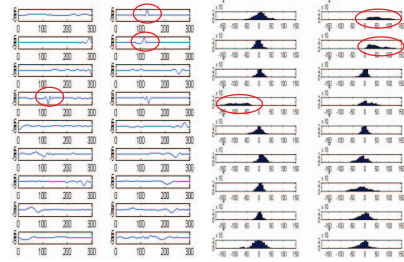

Figure 2: The ICA bases and the distributions of corresponding coefficients. (a) Eighteen ICA bases (for each figure, the x axis represents the sample index in one segment and y axis represents the amplitude); (b) Histograms of eighteen ICA coefficients.

D. RR Interval Features

Features such as wavelet features and ICA features are morphological features since they are extracted from a waveform of a single heartbeat. In contrast to morphological features, dynamic features are introduced to characterize the rhythm near a heartbeat. In recent efforts [11], it has been shown that dynamic features derived from ECG signals are useful to detect and classify arrhythmias. In this study, we want to investigate the effectiveness of these dynamic features for human identification. Following previous work [18], we selected

four dynamic features, namely, previous RR interval, post RR interval, local RR interval and average RR interval. The RR intervals refer to the separation between R peaks. The previous RR interval is defined as the interval between the given heartbeat and the previous heartbeat while the post RR interval refers to the interval between the given beat and the following beat. Local RR interval is defined as the average of ten RR intervals centered at the given beat. Finally, the average RR interval is determined by averaging the RR intervals within a 5-min period. Figure 3 shows four RR interval features for one particular heartbeat (with a red circle).

Figure 3: The previous RR, post RR, local RR and average RR interval features for one heartbeat

III. Methodology

A. ECG Databases

The ECG data used for algorithm evaluation was obtained from MIT-BIH Arrhythmia Database [1], MIT-BIH Normal Sinus Rhythm Database [2] and Long-Term ST Database [3]. These three ECG databases allow us to investigate four different scenarios, namely, short-term and long-term normal ECG, and short-term and long-term ECG with arrhythmias.

The MIT-BIH Arrhythmias Database (MITDB) contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, collected from 47 subjects. The subjects are 25 men with ages ranging from 32 to 89 years and 22 women with ages from 23 to 89 years (Records 201 and 202 come from the same male subject). This database is typically used as the benchmark for the arrhythmia detection and classification. In the current work, the database is used to study the performance of the proposed subject identification algorithm in the scenario of short-term ECG with arrhythmias. All 48 30-minute Holter recordings were included and each was divided into six 5-minute intervals and the first 5-minute interval of each record was used for training and the rest for testing.

The MIT-BIH Normal Sinus Rhythm Database (NSRDB) includes 18 long-term ECG recordings collected from 18 different subjects, 5 men, aged 26 to 45 and 13 women, aged 20 to 50. Subjects were found to have no significant arrhythmias. The database is used to study two different scenarios: short-term normal ECG and long-term normal ECG. For the short-term study, the first

Table 1: Detailed description of the four scenarios being investigated

	Short-term Normal ECG	Short-term ECG with Arrhythmias	Long-term Normal ECG	Long-term ECG with Arrhythmias
Database	NSRDB	MITDB	NSRDB	LTSTDB
Number of subjects	18	47	18	65
Number of records	18	48	18	68
Length of each record	23~24 hrs	30 mins	23~24 hrs	21~24 hrs
Length of each interval	5 mins	5 mins	5 mins	5 mins
Period between intervals	15 mins	0	1 hr 55 mins	1 hr 55 mins
Number of intervals in each record	6	6	11	11
Training Data	First interval of each record	First interval of each record	First interval of each record	First interval of each record
Testing Data	Remaining intervals	Remaining intervals	Remaining intervals	Remaining intervals

two hours of each recording were used and the first five minutes of every 20 minutes was extracted, resulting in six 5-minute intervals for each subject. The first 5-minute interval of each was used for training. For the long-term study, as the record length varies, we used the first 22 hours of each record and the first 5 minutes of every two hours was extracted such that there are eleven 5-minute intervals for each record and the first 5 minutes of each was utilized as the training data.

The Long-Term ST Database (LTSTDB) contains 86 Holter recordings from 80 human subjects. The individual recordings vary between 21 and 24 hours in duration, and contain two or three ECG lead signals. The 68 two-lead recordings are selected, with four from the same subject. In our work, the database is used to study for the scenario of long-term ECG with arrhythmias. The first 21 hours of each record were used and the first 5 minutes of every two hours were extracted so that there are eleven 5-minute intervals that have been extracted for each record. The first 5-minute interval of each record was used as the training set and the rest were used for evaluation.

The information for each scenario that was investigated is summarized in Table 1. All three databases are well-annotated, which means there is an annotation file associated with each recording, providing the information such as the location of each R peak and the type of each heartbeat, i.e., normal or exhibiting certain type of arrhythmia. The annotations of R peak locations would be used for the segmentation of heartbeats.

B. Data Preprocessing

The sampling frequency of the data from the MIT-Arrhythmia Database is 360 Hz whereas the sampling frequency of the MIT-BIH Normal Sinus Rhythm Database is 128 Hz and that of Long-Term ST Database is 250 Hz. Therefore, the data from the MIT Normal Sinus Rhythm and Long-Term ST Database were first re-

sampled so that all the digitized data in the analysis has the same sampling frequency of 360 Hz.

The data was then preprocessed to correct the baseline wander and then filtered by a band-pass filter [13] to remove high-frequency and low-frequency artifacts. The data was subsequently segmented based on the annotation regarding the location of the R peaks. Using the physician-provided R peak locations will allow us to evaluate how well subject identification can be performed using ECG signals, without worrying about how accurately R peak needs to be located. A sample size of 300 (0.83 seconds) was used, consisting of 100 samples before the R peak and 200 samples after the R peak. This appears to be sufficient to capture most if not all of the information from a particular heart cycle.

C. Feature Extraction

Daubechies wavelet of order 8 was used for wavelet analysis. Wavelets coefficients of D3, D4 and A4 were extracted to obtain a total number of 118 wavelet features for each heartbeat segment. In addition, all heartbeats in the training dataset were used to train 18 ICA components in each scenario using the FastICA [19] algorithm. The 18 ICA components were then used to calculate 18 ICA coefficients for each heartbeat. The two types of features were concatenated and principal component analysis (PCA) was employed to reduce the feature dimensionality from 136 to 26, which accounts for approximately 99% variance.

D. SVM Classification

A support vector machine (SVM) with a Gaussian radial basis function (RBF) kernel was chosen as the classification tool. At first, model parameters (e.g., the penalty parameter for the SVM and the width parameter for the kernel) were selected using the 10-fold cross

validation in a grid search scheme. Afterwards, a SVM classifier was trained based on the training data. The trained SVM classifier was then utilized to evaluate testing data. All SVM algorithms are implemented using the well-known LIBSVM [20] package.

E. Information Fusion

Same procedure was separately applied to the data from both leads (namely, the upper lead and the lower lead) so that there are two classification results for each heartbeat. Heartbeat segments for which the two leads give different classification results are rejected, i.e., no classification decision is made for such segments. This decision-based information fusion improves our classification confidence. The subject identity is finally determined either based on the classification result of single heartbeat or using the majority voting among N consecutive consistently classified heartbeats.

IV. Results and Discussion

As mentioned in Section III, the proposed methodology is evaluated in four scenarios to test its robustness against time interval between training and testing and to the presence of heart arrhythmias. For each scenario, the following three different settings are investigated: (I) Overall Analysis; (II) Single Heartbeat Interval Analysis; (III) Multiple Heartbeats Interval Analysis. The classifier is trained based on the first 5-minute intervals in each setting. In the overall analysis, the classifier is used to evaluate the entire testing dataset. In the single heartbeat interval analysis, the classifier is applied to test on different 5-minute intervals respectively. In other words, it is used to analyze the algorithmic performance decomposed over time. The subject identity is decided using the classification result of single heartbeat and the strategy is also adopted in the overall analysis. In the multiple heartbeats interval analysis, the basic configuration is identical to single heartbeat interval analysis except that the subject identity is determined by the majority voting between multiple consecutive

consistently classified heartbeats. In addition, the algorithmic performance of including the four dynamic features has been investigated and the results are presented to show that they are inappropriate for subject identification.

A. Overall Analysis

In this setting, for each scenario, the classifier is trained based on the first 5-minute intervals and used to evaluate on the rest data (i.e., entire test dataset). The performance for each scenario is summarized in Table 2, which displays the number of heartbeats used for testing, the rank-1 identification accuracy (IDA) on the first and the second lead, the final accuracy by fusing the classification results of both leads, and the percentage of heartbeats that were rejected. It shows that the decision-based fusion improved the performance and boosted the classification confidence significantly. Given the subsequent 5-minute intervals are 15 minutes apart in scenario I, contiguous in scenario II and 1 hours 55 minutes apart in scenarios III and IV, we can infer that the more separation between subsequent intervals, the lower single-lead accuracy and the higher percentage of rejected heartbeats. Although the single-lead performance seems better and the percentage of rejections is lower in long-term normal ECG signals than long-term arrhythmia ECG signals, one must be careful with this comparison given the large difference between the number of subjects in these scenarios (Table 2). However, what we can definitely observe is that considerably high recognition rates have been achieved with the fusion between two ECG leads no matter whether it is short-term or long-term and with arrhythmias or without.

B. Single Heartbeat Interval Analysis

In this setting, for each scenario, the classifier is trained based on the first 5-minute intervals and utilized to test on different subsequent 5-minute intervals respectively. The subject identity is decided based on the classification result of single heartbeat (i.e., Rank-1 Identification Accuracy).
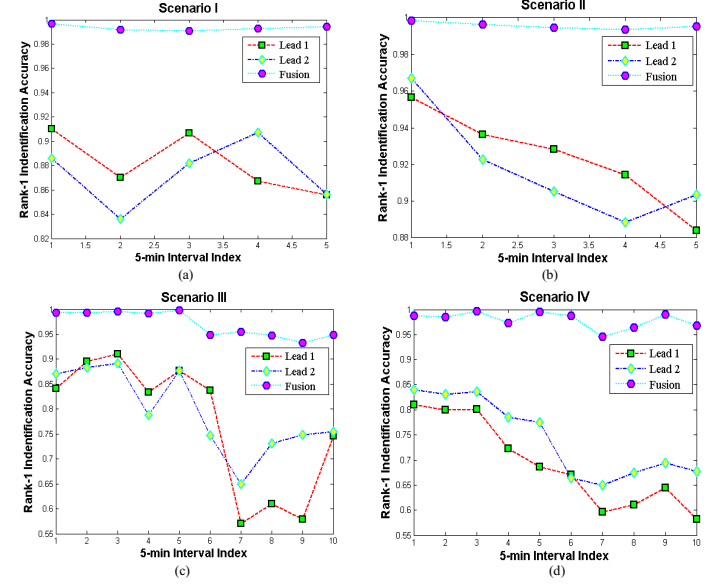
Table 2: Overall analysis performance summary: single-lead performance, two-leads fusion performance and percentage of rejected heartbeats

Scenario	Test Size	Upper Lead Accuracy	Lower Lead Accuracy	Final Accuracy	Percentage of Rejections
Short-term normal ECG	36369	88.20%	87.36%	99.33%	16.89%
Short-term ECG with Arrhythmias	91751	92.34%	91.72%	99.57%	12.26%
Long-term normal ECG	62563	78.00%	79.94%	97.54%	28.71%
Long-term ECG with Arrhythmias	232769	70.04%	74.93%	98.11%	37.09%

Figure 4: Rank-1 Indentification Accuracy of the Single Heartbeat Interval Analysis for the four scenarios

The setting is used to analyze the algorithmic robustness of the proposed approach against time. As shown in Figure 4, the single-lead performance generally tends to decrease with time. However, the decision-based fusion still enables us to attain and preserve reasonably good final IDA, ranging from 95% to 100%.

C. Multiple Heartbeats Interval Analysis

The basic configuration is identical to the single Heartbeat Interval Analysis except that the subject identity is determined based on the majority voting among N consecutive consistently classified heartbeats. Intuitively, the IDA would increase by voting among more consecutive consistently classified heartbeats.

The N was varied over all odd integers from 1 to 31. For each scenario, we studied the performance on all testing data as well as on three testing intervals, that is, the first testing interval, a middle testing interval and the last interval. From Figure 5, it appears that N = 21 would be a reasonable choice that would give us an overall IDA above 98% in all cases investigated.

D. Dynamic features

The effectiveness of the previously mentioned four dynamic features for human identification was investigated. The overall IDAs from including the dynamic features are presented in Table 3. Compared with Table 2, the performance deteriorated significantly after including dynamic features, indicating that the dynamic features proposed for arrhythmia classification do not seem beneficial for human identification. These dynamic features are sensitive to heartbeat rhythm and seem to work well for arrhythmia classification [11]. It is not clear why these dynamic features do not contribute to individual identification. We expect heart rhythm to vary significantly as a subject conducts different activities and is in different emotional states (e.g. sitting, walking, staying calm or being anxious), and these results seem to hint that this variation does not follow an individual personalized pattern. More studies with specific experiments are needed to better understand this issue.
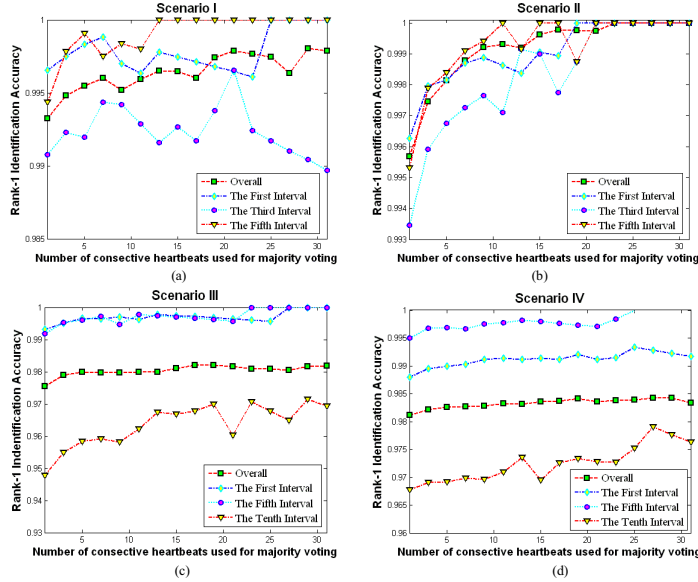
Figure 5: The Rank-1 Identification Accuracy of the Multiple Heartbeats Interval Analysis for the four scenarios

Table 3: Overall Analysis (including Dynamic Features)

Scenario	Test Size	Upper Lead Accuracy	Lower Lead Accuracy	Final Accuracy	Percentage of Rejections
Short-term normal ECG	36369	76.34%	77.67%	89.24%	28.91%
Short-term ECG with Arrhythmias	91751	81.01%	79.33%	90.28%	25.64%
Long-term normal ECG	62563	70.31%	69.49%	86.54%	39.47%
Long-term ECG with Arrhythmias	232769	61.63%	62.28%	87.31%	43.09%

V. Conclusions

In this paper, we investigated the use of two-lead ECG signals for human identification. Wavelet Transform and Independent Component Analysis are applied to each single lead signal to extract morphological information. The information from two leads is fused by rejecting the heartbeat segments that are inconsistently classified between two leads. This decision-based fusion significantly enhanced the identification accuracy. The subject identity is finally determined based on the majority voting among multiple consecutive consistently classified heartbeats. The methodology has been validated over three public ECG databases and substantially high rank-1 IDAs (as high as 99.6%) are achieved whether in the short-term or the long-term and with the presence of arrhythmia or not. The result demonstrates the great potential of ECG signals and the proposed method in the biometrics system.

ACKNOWLEDGMENTS

REFERENCES

[1] MIT-BIH Arrhythmia Database (Cited in May 2010): http://physionet.org/physiobank/database/mitdb/

[2] MIT-BIH Normal Sinus Rhythm Database (Cited in May 2010): http://physionet.org/physiobank/database/nsrdb/

[3] Long-Term ST Database (Cited in May 2010): http://physionet.org/physiobank/database/ltstdb/

[4] N.V. Boulgouris, N.P. Konstantinos and E.M. Tzanakou, *Biomtrics:Theory, Methods and Applications,* Wiley IEEE, 2009

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

[6] T.W. Shen, W.J. Tompkins, Y.H. Hu, "One-lead ECG for identity verification," *IEEE EMBS*, vol. 1, no., pp. 62-63, 2002.

[7] S. A. Israel, W. T. Scruggs, W. J. Worek, and J.M. Irvine, "Fusing face and ECG for personal identification", *Applied Imagery Pattern Recognition Workshop*, vol., no., pp. 226-231, 2003.

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

[9] F. Agrafioti and D. Hatzinakos, "Fusion of ECG sources for human identification," *IEEE ISCCSP*, vol., no., pp.1542-1547, 2008.

[10] Vital Responder Project Website (Cited at May 2010): http://ribosome.ieeta.pt/vitalresponder/

[11] C. Ye, M. T. Coimbra and B.V.K. Vijaya Kumar, "Arrhythmia Detection and Classification using Morphological and Dynamic Features of ECG Signals," Submitted to *IEEE EMBS*, 2010.

[12] Holter Monitor – Texas Heart Institute (Cited at May 2010): http://www.texasheart.org/HIC/Topics/Diag/diholt.cfm

[13] K. Robert and E.C. Colleen, *Basis and Treatment of Cardiac Arrhythmias*, 1st ed., Springer, New York, 2006.

[14] N.V. Thakor, J.G. Webster and W.J. Tompkins, "Estimation of QRS Complex Power Spectra for Design of a QRS Filter," *IEEE Trans. Biomed. Eng.*, vol. 31, no. 11, pp. 702-706, 1984.

[15] A. Hyvärinen, J. Karhunen and E. Oja, *Independent Component Analysis*, 1st ed., Wiley Interscience, 2001.

[16] A. Bollmann, N. K. Nanuru, K. K. McTeague, P.F. Walter, D.B. DeLurgio and J.J. Langberg, "Frequency analysis of human atrial fibrillation using the surface electrocardiogram and its response to ibutilide," *Amer. J. Cardio.,* vol. 81, no. 12, pp. 1439–1445, 1998.

[17] J.J. Rieta, F. Castells, C. Sanchez, V. Zarzoso and J. Millet, "Atrial activity extraction for atrial fibrillation analysis using blind source separation," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1176-1186, 2004.

[18] D.C. Philip, M. O'Dwyer and R.B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1196-1206, 2004.

[19] A. Hyvärinen and E.Oja, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis", *IEEE Trans. Neural Networks*, vol. 10, no. 3, pp. 626-634, 1999. Software available at http://www.cis.hut.fi/projects/ica/fastica/.

[20] C.C. Chang and J.L. Chih, "LIBSVM: a library for support vector machines", 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.