Analysis of QRS Patterns in 15-Lead ECG for Person Verification

Vessela Tzvetanova Krasteva, Irena Ilieva Jekova and Roger Abächerli

Abstract – **This paper presents a method for evaluation of similarity and difference scores of QRS patterns in 15-lead ECG for the aims of person verification. An ECG database with 316 healthy subjects, including two records per subject taken >1year apart is used to simulate the real case scenario. Discriminant analysis estimates the best specificity/sensitivity for limb+chest leads (92.9/92.1%), lower for limb leads (92.1/89.6%), and the top-scored single leads: aVR (84.6/84%), II (83.8/83.5%), I (81.2/80.2%).**

Keywords – **ECG Biometrics, multilead ECG scoring, QRS patterns, Discriminant analysis, person verification.**

I. INTRODUCTION

The analysis of the electrocardiogram (ECG) as a biometric tool has been started about a decade ago in the context of two typical scenarios for application:

1) Person verification (one-to-one scenario): the ECG of the tested subject is compared to previously recorded ECG with known identity (ID). The tested person is either verified or rejected.
2) Person identification (one-to-many scenario): the ECG of the subject under identity examination is compared to previously recorded set of ECGs in a specific database. The tested person is identified as a subject with unique ID among all in the database.

Two general methods could be distinguished:

1) using measurements after fiducial points detection;
2) analyzing the overall ECG waveform morphology.

At first, the fiducial based approaches are applied. The earliest work involves 12 uncorrelated diagnostic features of P-QRS-T amplitudes and durations [1]. The inter-subject heartbeat similarities are studied via Principle Component analysis score plots. The authors report 100% identification accuracy (IDA) over a database with 20 subjects. Other authors employ 15 temporal features of the P-QRS-T segment into a set of discriminant functions [2]. They report IDA in the range from 97% to 100% over 29 subjects under various stress conditions. A two-step identification method involves temporal and amplitude measurements based on fiducial points detection together with appearance based features that capture the heartbeat patterns [3]. This combined approach provides 100% IDA when tested over 31 healthy subjects: 18 with a single ECG record [4] and 13 with more than one ECG record [5].

Fiducial independent approaches have been developed since 2006. Person identification via autocorrelation (AC) and discrete cosine transform of windowed ECG reports 100% IDA over a database with 14 subjects [6]. Another study also utilizes AC of 5s ECG for person identification and verification [7]. Classification of AC functions via discriminant analysis achieves 96.2% IDA, 87% and 99% verification sensitivity and specificity, reported for a joint dataset [4,8] and 13 healthy subjects with more than one ECG record [5]. The maximal correlation coefficient of a single-lead and 12-lead ECG is reported to provide 91.4% and 100% IDA over a database with 11 subjects [9]. Another effective method calculates the two-dimensional heart vector formed by the limb ECG leads and its first and second derivatives, reporting 98.1% IDA and 97.2% verification accuracy by a distance based approach over 74 subjects [10]. The processing of a normalized QRS complex via Multilyer perceptron provides 96.1% IDA over a database with 30 healthy subjects [11]. Recently, a human ECG identification system has been announced based on ECG decomposition in a number of intrinsic mode functions combined with Welch spectral analysis for extraction of significant heartbeat features [12]. The classification with the K-Nearest Neighbors provides 95.6% IDA over a joint dataset with 108 subjects having one ECG record with ST-segment changes [13,14] and 12 healthy subjects with more than one ECG record [5].

Majority of the cited methods are tested with small-sized ECG databases [1,2,6,8,11] or track intra-subject changes of ECG characteristics measured in very short distanced temporal intervals [2,3,7,11,12]. This might bias the reported high identification/verification accuracy from the real case scenario.

This works aims to compare inter-subject QRS patterns of 15-lead ECG and to rate leads by similarity and difference scores via Discriminant analysis for the purpose of person verification. The use of a large sized ECG-database with two different records per subject taken >1year apart aims at an unbiased accuracy report.

II. ECG DATABASE

The ECG database is collected in the period 2004-2009, including 316 patients at the Emergency Department of the University Hospital Basel. The ECGs are acquired via SCHILLER CS-200 Excellence device with 500Hz sampling rate, 2.5μV resolution.

The database has the following content:

- Includes subjects with a healthy cardiac status, 143 man, 173 woman, aged from 18 to 89 years;
- Includes two 10s resting ECG recordings per subject taken at different times distanced from 1 to 2 years.
- All ECG recordings have a high quality signal in 15 ECG leads – limb (I, II, III, aVR, aVL, aVF), chest (V1-V6), synthesized orthogonal (X, Y, Z).

V. Krasteva, I. Jekova are with the Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences, Acad. G. Bonchev str., bl. 105, 1113 Sofia, Bulgaria,
e-mail: vessika@biomed.bas.bg, irena@biomed.bas.bg
R. Abächerli is with SCHILLER AG, Altgasse 68, CH-6341 Baar, Switzerland, e-mail: Roger.Abaecherli@schiller.ch

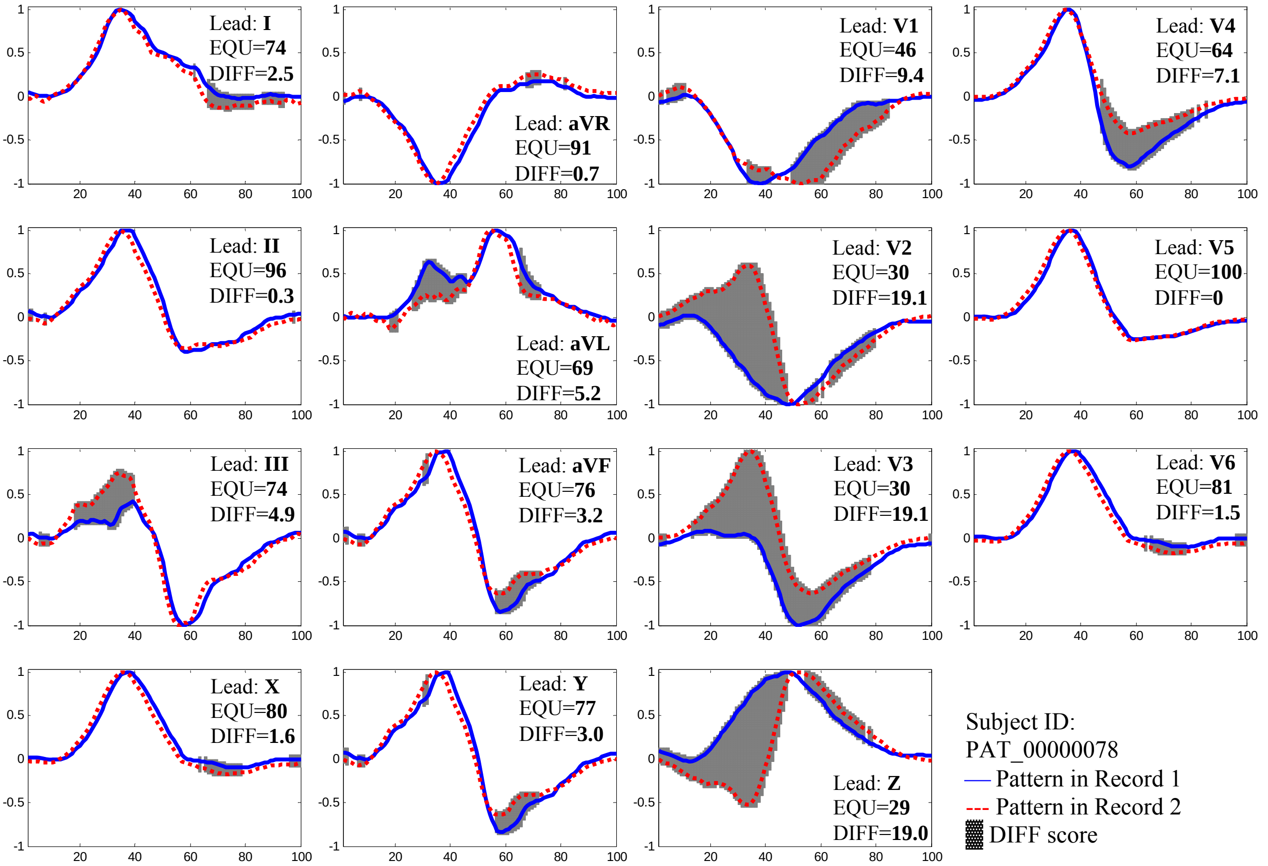

Fig. 1. Comparison of QRS patterns in all 15-leads extracted from two recordings of the same subject.

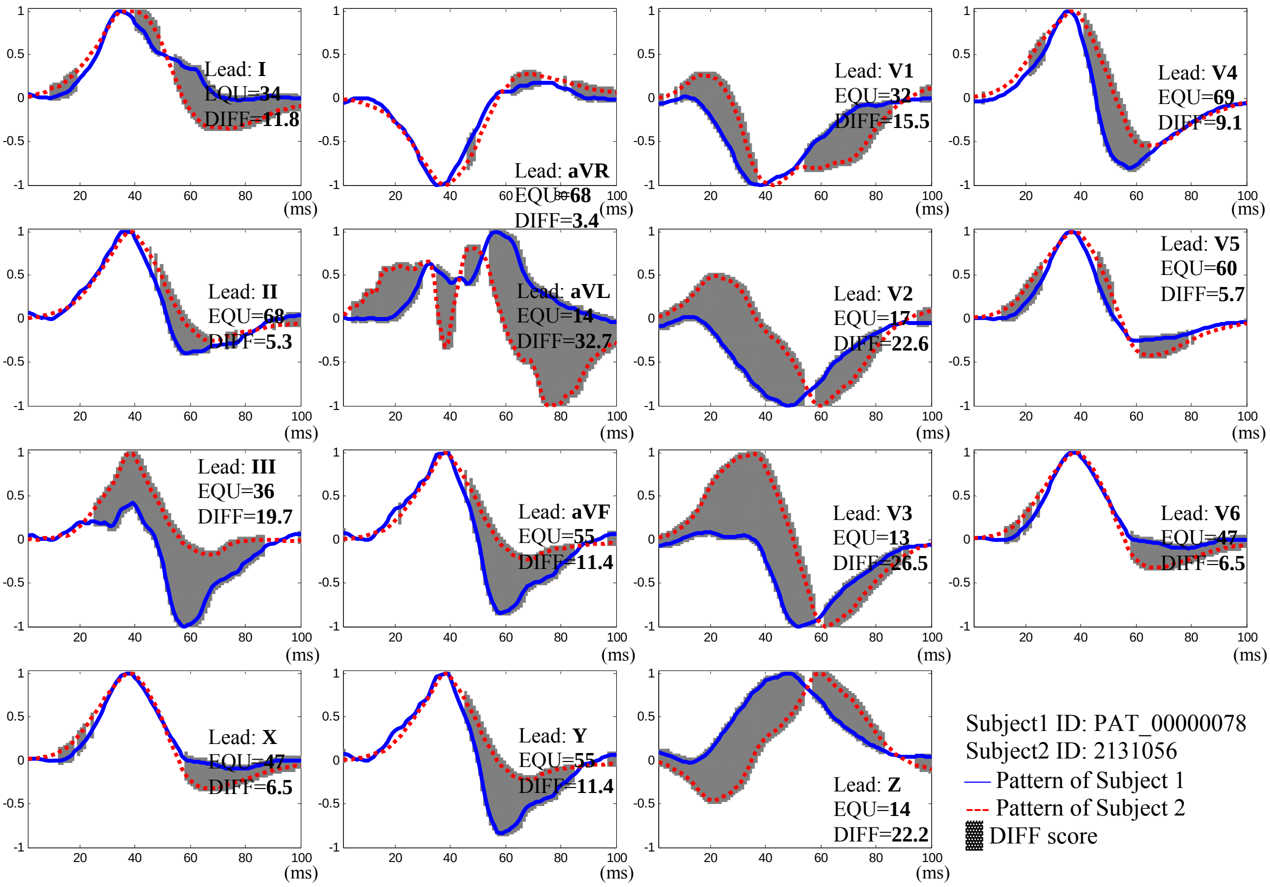

Fig. 2. Comparison of QRS patterns in all 15-leads extracted from the recordings of two different subjects.

III. Method

All 10s 15-lead ECG recordings are pre-filtered in a diagnostic bandwidth 0.05–75 Hz. The embedded CS-200 QRS detector is applied to locate RR-intervals, which are then fed to a baseline correction for zeroing of the mean amplitude at the P-wave beginning and the T-wave end. An improved signal-to-noise averaged PQRST patterns are then calculated for each lead.

The method considers QRS patterns extracted from the averaged PQRST patterns during cardiac depolarization, aligned for all leads within a window of 30ms before and 70ms after the R-peak of Lead I. The pattern of each lead is drawn in a normalized 2D space with x-axis [0 to 100] ms; y-axis [-1 to 1], considering y-axis normalization towards the maximal lead amplitude to avoid the influence of intra-subject and inter-lead ECG amplitude differences. The patterns of two different recordings are compared when the respective leads are overlapped in the normalized 2D space and are scored in respect of normalized piecewise equality (EQU score) and difference (DIFF score), where:

- EQU=100%, DIFF=0% corresponds to full amplitude coincidence of all pattern samples;
- EQU<100%, DIFF>0% scores the percentage of the non coinciding samples and the accumulated amplitude differences for them.

Figures 1,2 illustrate the normalized 2D space of the QRS patterns in 15-lead ECG and the calculated EQU, DIFF scores for two scenarios:

(1) One subject is compared to the same subject when its ECG is taken after >1year (Figure 1): EQU=67.8±23.6% (range: 29-100%); DIFF=6.1±6.9% (range: 0-19.1%).

(2) The first subject is compared to a different subject (Figure 2): EQU=41.9±20.7% (range: 13-69%); DIFF=14±8.8% (range: 3.4-32.7%).

IV. Results

The total database statistical distributions in Table 1 show significantly higher EQU and lower DIFF scores in scenario 1 vs. 2 for all 15-leads, $p<0.05$.

Linear Discriminant Analysis (LDA) over EQU and DIFF scores is applied to estimate the potential for person verification of each lead (Fig. 3) and the different lead systems: limb, chest, orthogonal (Table 2).

The following performance indices are considered in the person verification task:

- Sensitivity (Se) scoring the correct verification rate comparing equal subjects, N=316 cases:

$$Se = \frac{\text{Total Nb Correct Verifications}}{N} * 100\ (\%)$$

- Specificity (Sp) scoring the correct rejection rate comparing all different subjects N*(N-1)=99540 cases:

$$Sp = \frac{\text{Total Nb Correct Rejections}}{N*(N-1)} * 100\ (\%)$$

The histograms (Fig.4) give a hint about the range of EQU, DIFF thresholds that provides the top performance found in limb+chest leads (Sp/Se=92.9/92.1%): >92% of equal subjects have EQU>75% or DIFF≤3%; >95% of different subjects have EQU≤80% or DIFF>2%.

Table 1. Distribution of EQU and DIFF scores in 15-lead ECG, estimated for the two groups of equal and different subjects, presented as: mean value ± std (10-90 percentile range). The index MDstd, calculated as the difference between means normalized to the mean standard deviation in the two groups is used to rate the leads with the most separable distributions (bolded).

ECG Leads		EQU score (%) Equal subjects 316 cases	EQU score (%) Different subjects 99540 cases	DIFF score (%) Equal subjects 316 cases	DIFF score (%) Different subjects 99540 cases	MDstd for EQU/DIFF
Limb	I	90.8±13.2 (70-100)	67.7±15.9 (46-88)	1.2±2.1 (0-3.6)	5.5±3.5 (1.4-10.2)	1.59/ 1.51
	II	**89.2±14.3** (68-100)	**58.2±17.3** (36-81)	**1.5±2.4** (0-4.4)	**7.9±4.3** (2.7-13.7)	**1.97/ 1.89**
	III	65.6±24.2 (31-98)	34.0±16.5 (14-56)	8.2±7.8 (0.2-21.3)	20.7±8.6 (8.7-31.6)	1.55/ 1.53
	aVR	**93.7±11.4** (77-100)	**68.3±16.4** (46-89)	**0.8±1.7** (0-3.1)	**5.3±3.3** (1.2-9.9)	**1.83/ 1.78**
	aVL	67.6±23.7 (33-99)	36.9±16.7 (16-59)	7.7±7.5 (0.1-19.4)	19.2±8.4 (7.8-29.9)	1.52/ 1.45
	aVF	77.6±22.0 (44-100)	44.9±18.3 (22-70)	4.1±5.0 (0-11.3)	12.9±6.8 (5.0-22.5)	1.62/ 1.50
Chest	V1	77.7±22.9 (42-100)	49.5±19.5 (25-76)	3.7±5.1 (0-9.9)	10.2±5.8 (3.5-17.9)	1.33/ 1.20
	V2	69.9±22.6 (38-100)	43.3±17.9 (21-68)	6.6±6.4 (0-15.9)	14.3±6.4 (6.2-23.1)	1.31/ 1.20
	V3	66.3±22.1 (32-96)	43.4±17.8 (21-67)	8.0±6.7 (0.5-17.9)	15.5±6.8 (6.8-24.5)	1.15/ 1.10
	V4	77.0±21.0 (46-100)	53.0±19.4 (27-79)	4.7±5.4 (0-11.7)	11.4±6.6 (3.7-20.9)	1.19/ 1.12
	V5	**86.5±16.0** (64-100)	**61.6±18.8** (36-86)	2.4±3.7 (0-6.5)	7.9±5.2 (2.0-14.8)	**1.43**/ 1.24
	V6	86.6±15.1 (63-100)	64.8±18.0 (40-88)	2.1±2.8 (0-6.2)	6.3±4.0 (1.5-11.8)	1.32/ 1.22
Orthogonal	X	86.7±15.1 (63-100)	65.0±18.0 (46-89)	2.1±2.8 (0-6.2)	6.3±4.0 (1.5-11.8)	1.32/ 1.22
	Y	**77.7±21.9** (43-100)	**45.0±18.3** (40-88)	**4.1±5.0** (0-11.2)	**12.9±6.8** (4.9-22.5)	**1.63/ 1.50**
	Z	70.5±22.3 (37-100)	44.1±18.0 (21-69)	6.4±6.3 (0-15.5)	13.9±6.4 (5.9-22.8)	1.31/ 1.18
Limb		**88.2±9.3** (75-97)	**61.4±11.2** (47-76)	**1.0±1.7** (0-3.2)	**7.2±3.1** (3.4-11.3)	**2.61/ 2.58**
Chest		87.2±9.5 (71-97)	63.4±12.6 (46-80)	1.0±1.9 (0-4.1)	6.4±3.4 (2.3-11.0)	2.15/ 2.04
Orthog.		85.3±11.0 (69-98)	59.5±12.6 (43-76)	1.6±2.1 (0-4.9)	7.4±3.5 (3.1-12.1)	2.19/ 2.07

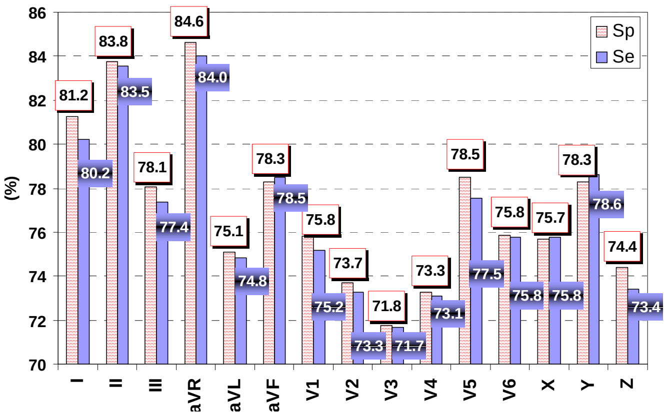

Fig. 3. Mean lead performance for person verification estimated for EQU and DIFF scores by LDA.

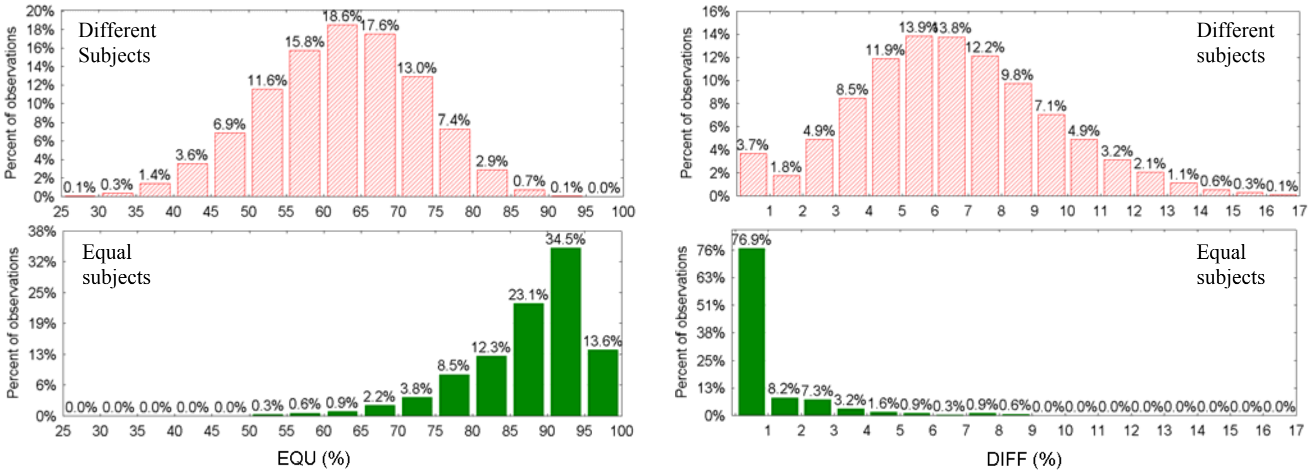

Fig. 4. Histograms of EQU and DIFF scores estimated as a summary for all leads in the top-rated Limb+Chest lead system.

TABLE 2. BEST PERFORMANCE OF DIFFERENT LEAD SYSTEMS FOR PERSON VERIFICATION ESTIMATED FOR EQU, DIFF BY LDA.

ECG leads	Sp (%) 99540 cases	Se (%) 316 cases
Limb (DIFF)	92.1	89.6
Chest (EQU)	90.1	85.4
Orthogonal (EQU)	86.6	83.5
Limb+Chest (EQU and DIFF)	**92.9**	**92.1**
Limb+Chest+Orthogonal	92.2	90.8

IV. DISCUSSION AND CONCLUSIONS

This study presents a simple method for evaluation of equality and difference of 15-lead QRS patterns that are observed between two recordings of different and equal subjects. The temporal alignment and the normalization for reducing the contribution of the height and width variability of QRS patterns in different recordings is of crucial importance for the correct EQU/DIFF measure.

The study over a large-sized database with 316 subjects provides unbiased person verification from the real-case scenario. The leads with the best separable statistical distributions are bolded in Table 1: II, aVR for limb leads; V5 for chest leads; Y for orthogonal leads; Limb over chest and orthogonal leads. LDA performance of different lead systems shows (Table 2):

(1) Limb leads have the biggest potential for person verification by DIFF score (Sp/Se=92.1/89.6%) with the top-3 rated limb leads (Fig.3): aVR (84.6/84%), II (83.8/83.5%), I (81.2/80.2%).

(2) Chest leads are the second rated for person verification by EQU score (90.1/85.4%) with the top-3 rated Chest leads (Fig.3): V5 (78.5/77.5%), V6 (75.8/75.8%), V1 (75.8/75.2%). The misplacement of the intra-subject lead positions may play deteriorating role for the total performance loss of all chest leads, with the most prominent negative influence in V3,V4,V2.

(3) Synthesized orthogonal leads have the least contribution to the person verification (86.6/83.5%), with the top rated lead Y (78.3/78.6%) in Fig.3.

(4) The combination limb+chest leads provides the best accuracy for person verification (92.9/92.1%), obtained with the common evaluation of EQU and DIFF. Further estimation of orthogonal leads deteriorates the results.

ACKNOWLEDGEMENTS

This study was supported by the Bulgarian Scientific Research Fund, grant T02/11.

REFERENCES

[1] L. Biel, O. Pettersson, L. Philipson, P. Wide. *ECG analysis: a new approach in human identification*, IEEE Transactions on Instrumentation and Measurement, Vol. 50(3), pp. 808-812, 2001.

[2] S. Israel, J. Irvine, A. Cheng, M. Wiederhold, B. Wiederhold. *ECG to identify individuals*, Pattern Recognition, Vol. 38, pp. 133-142, 2005.

[3] Y. Wang, F. Agrafioti, D. Hatzinakos, K. Plataniotis. *Analysis of human electrocardiogram for biometric recognition*, EURASIP Journal on Advances in Signal Processing. article ID 148658, 2008.

[4] MIT-BIH Normal Sinus Rhythm database, available at: http://www.physionet.org/physiobank/database/nsrdb/

[5] PTB database, available at: http://www.physionet.org/physiobank/database/ptbdb/

[6] K. Plataniotis, D. Hatzinakos, J. Lee. *ECG biometric recognition without fiducial detection*, Proc. on Biometrics Symposium BSYM'06, Baltimore, Sept. 19-21, 2006.

[7] F. Agrafioti, D. Hatzinakos. *ECG biometric analysis in cardiac irregularity conditions*, Signal Image and Video Processing, Vol. 3, pp. 329-343, 2009.

[8] MIT-BIH Arrhythmia database, available at: http://www.physionet.org/physiobank/database/mitdb/

[9] F. Poree, A. Gallix, G. Carrault. *Biometric identification of individuals based on the ECG. Which conditions?*, Computers in Cardiology, Vol. 38, pp. 761-764, 2011.

[10] G. Wubbeler, M. Stavridis, D. Kreiseler, R. Bousseljot, C. Elster. *Verification of humans using the electrocardiogram*, Pattern Recognition Letters, Vol. 28, pp. 1172-1175, 2007.

[11] K. Sidek, I. Khalil, M. Smolen. *ECG biometric recognition in different physiological conditions using robust normalized QRS complexes*, Computers in Cardiology Vol. 39, pp. 97-100, 2012.

[12] Z. Zhao, L. Yang, D. Chen, Y. Luo. *A human ECG identification system based on ensemble empirical mode decomposition*, Sensors, Vol. 13, pp. 6832-6864, 2013.

[13] MIT-BIH ST change database, available at: http://www.physionet.org/physiobank/database/stdb/

[14] Long-term ST database, available at: http://www.physionet.org/physiobank/database/ltstdb/