

Individual identification using electrocardiogram morphology

Antonio Fratini, Mario Sansone, Paolo Bifulco, Maria Romano, Alessandro Pepino and Mario Cesarelli
Dept. of Electrical Engineering and Information Technologies – University Federico II of Naples
Naples, Italy
a.fratini@unina.it

Giovanni D'Addio
S. Maugeri Foundation, Rehabilitation Institute of Telese
Telese Terme, Italy
gianni.daddio@fsm.it

Abstract— The use of electrocardiogram as biometric has raised attention in the last decade and a wide variety of ECG features were explored to verify the feasibility of such a signal. In this work the authors aim to describe a simple template based approach to the electrocardiographic biometric identification using the morphology of individual's heartbeat. The developed algorithm was tested on different recordings made available in the Physionet public database Fantasia: two different sets of heartbeats were extracted from individual recordings one was used for the template building while the second for the tests. The performances of the algorithm are encouraging with a true acceptance rate of 99.4%, however, the procedure needs to be tested on different recordings of the same individual, or during the course of a whole day or physical activity.

Keywords— *feature extraction, electrocardiogram morphology, individual identification*

I. INTRODUCTION

The Electrocardiogram (ECG) reflects the electrical activity of the heart, it is related to physiological factors such as its mass, conductivity and the activation times of its different areas; however, it also reflects its geometry and position within the thoracic cage. These peculiar aspects introduce a significant variability of the ECG between individuals, and this property can be used for the characterization of the subjects themselves [1, 2].

The idea of identifying subjects with ECG is relatively young, but it contains significant advantages. The main one is its robustness with respect to forgery or theft. In addition, the ECG is an indicator of life, in the sense that it indicates that the individual is alive and is actually the one who is requesting recognition. Previously proposed identification techniques are based on different characteristics of the ECG, in particular, there are different classes of parameters used in biometric identification: morphological, time-related, and other specific processing technique. The features related to morphology can be further divided in amplitude characteristics (peaks of single waves) [3-9] and signal specific trends (angles between the same waves). The temporal characteristics instead, are based on the extraction of intervals between the main waves; finally, other authors use coefficients obtained by means of signal

processing techniques such as correlation, Fourier Transform, Discrete Fourier Transform, Wavelet and others [10-13]. Many different classifiers have been used in the process of recognition: neural networks, Bayesian estimates, and features distances evaluation (e. g. Euclidean and Mahalanobis distance, k-NN) [14]. The performances of the studies are not easily comparable with each other because they are generally obtained on different subjects in different conditions: those studies show different results and the computational cost of the proposed systems varies in proportion to the performance improvement. Recognition rates generally decrease when recordings contain arrhythmias or varying heart rates. [1, 15] Moreover, most of the literature is based on fiducial point detection, which introduces errors related to the detection precision and the number of needed fiducials.

With this work the author aimed to design a recognition algorithm based on an analysis of the entire morphology of the heartbeat signal; in addition, the authors intended to use a simple template-based approach, using R-peak fiducial detection only as usually utilized in many electrocardiogram studies. [16]

The study was performed on 36 healthy subjects whose recordings were made available in the Physionet public database; two different sets of heartbeats were extracted from individual recordings one was used for the template building the second for the tests.

The performances of the algorithm are encouraging with a true acceptance rate of 99.4%, however, the procedure must be tested on different recordings of the same subject, and during the course of a whole day or physical activity.

II. METHODS

In order to avoid the need of precise detection of all of the individual ECG waves and, in turn, to reduce the errors related to sampling frequency or superimposed noise, we decided to use the entire ECG heartbeat morphology.

In particular, we only require R peak identification through established detection techniques; R peak was chosen as temporal marker of a heartbeat. In addition, the use of the

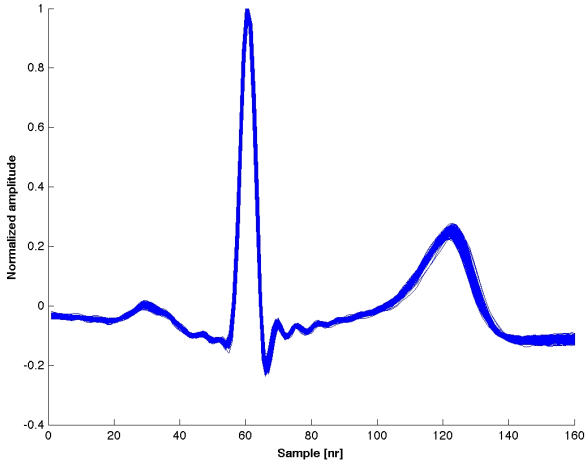


Fig. 1. Example of superposition of 300 heartbeats signals: the ECG beats belong to a single subject

entire ECG morphology raises the number of considered features assuring a better separability between individuals.

A. Signals

The electrocardiographic signals used in this study were taken from Physionet database; in detail, 36 individual recordings of the database "Fantasia" have been used: 18 tracks belong to young people between 21 and 34 years, including 9 women and 9 men, while the other 18 belong to elderly between 68 and 85 years, including 10 women and 8 men. Signals were all taken at rest while the subjects were in supine position watching Walt Disney movie "Fantasia" (1940) to maintain wakefulness and, at the same time, a mental state of calm and relaxation. Used Fantasia database signals includes ECG with beat annotations, they are around two hours long and all sampled at 250Hz. Signals were pre-filtered with a Hamming 30th order filter in a band of 0.5-40 Hz to maintain the peculiar characteristics of ECG while rejecting most of the noise and artifacts.

B. Template and test set construction

Two different time intervals were randomly selected for each signal: the first (template interval) of about five to six minutes of recording was used for the construction of the template while the second (test interval) of about four to five minutes was used for the construction of the test set; R waves peaks were then identified through the associated annotation files.

For each individual, three hundred ECG beats were extracted from the first interval while two hundred and five from the test interval; the duration of a single heartbeat was chosen equal to 640 ms, 240 ms before R peak and 400 ms after, to extract the entire beat morphology. However, since the sampling frequency of the signals used in this study is of 250Hz, the total number of samples for each single heartbeat is 160.

Each ECG beat was then normalized to its maximum amplitude and stored in a matrix; the inspection of heartbeats revealed that, for the same individual, they generally repeated

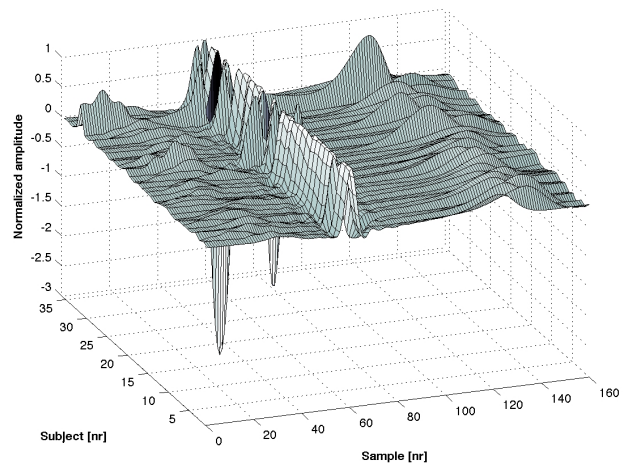


Fig. 2. Comparison of heartbeat from individuals: the figure contains mean heartbeats of all the subject included in this study

themselves in a very similar way (see figure 1) while almost all showed a considerable level of dissimilarity among individuals, as depicted in figure 2.

Mean and standard deviation were then computed for each of the one hundred and sixty samples of the heartbeats Fig. 2 were:

$$mean_heartbeat(i) = \frac{\sum_{j=1}^{300} normalized_heartbeat(i)}{300}$$

for $i = 1$ to 160

Comparison of heartbeat from different individuals: the figure as obtained considering all of the subject used in this study extracted from template interval, to obtain an index of data dispersion around the mean ECG beat signal. From these data, the templates were set up: in detail, for each of the sample of the ECG beat, template was identified as a domain of ± 3 standard deviations width around its mean; in this way it was possible to identify thirty six individual templates (see figure 3).

For each individual, two hundred test heartbeats were then obtained by averaging 5 heartbeats at a time, in the sequence of 205 heartbeats belonging to the test interval, sliding one beat to the next.

C. Identification procedure

The recognition phase was based on the computation of Euclidean distances between each sample of test heartbeat and template: if the test heartbeat sample is located within the template, that is its distance from the template mean is less than or equal to 3SD, then the sample is considered positive.

A score is assigned to each test heartbeat, it is incremented by one each time that a sample is considered positive; therefore, since the number of samples for single beat is equal to 160 the score can assume values in a range from 0 to 160 (all of the test heartbeat samples are within the template). The minimum score threshold to accept the identity of an individual

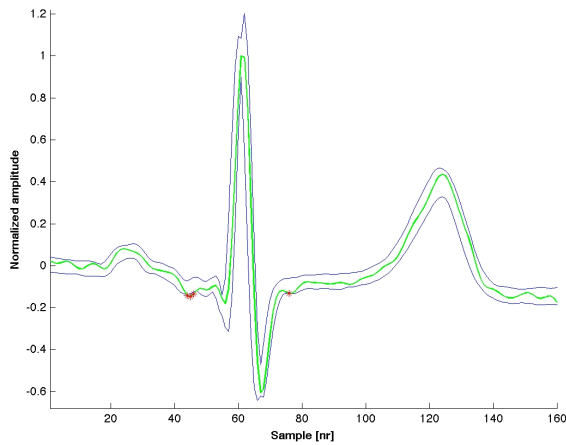


Fig. 3. Example of individual template: blue lines (superior and inferior) identify the template while green line represents a single individual heartbeat; red asterisk denote the samples outside the template

to be recognized was chosen from empirical observation equal to 150, that corresponds to 93.75% of coincidence between template and test heartbeat.

III. RESULTS

The performance of the system was tested using 200 test heartbeats for each subject (7200 total test heartbeats); they were then compared with all of the 36 subjects templates. In particular, with the identified minimum threshold (score equal to 150), the algorithm showed a true acceptance rate (TAR) of 99.37 % (7155 out of 7200), a false rejection rate of 0.6% (41 out of 7200) and a false acceptance rate of 0.05% (4 out of 7200).

IV. DISCUSSION

The choice of using the entire heartbeat morphology have been addressed with the intention of keeping low the computational complexity using only R peak detection. The results obtained confirmed the validity of the choice even if the number of features is still high (160), the algorithm is automatic and involves only mean and standard deviation evaluation. However, we have to consider the limitations of the results: the development of the algorithm involved the study of electrocardiographic recordings from individuals at rest and for a time not higher than two hours. It is therefore desirable to test the algorithm with 24 hours recordings and in a time period of months; it would also be interesting to verify its performance with recordings acquired before and after physical activity to take into account the effects of the change in heart rate in ECG morphology. During or after physical activity some ECG parameters [17] can result modified and this may complicate individuals' identification. Other difficulties can arise using different devices or sensors, particularly with dry electrodes [18-21] frequently used to allow more pervasive and user-friendly applications [22, 23]. Relative motion between patient and electrodes (but also if an individual is subjected to vibration for example in a vehicle [24]) can generate large

artifacts in bioelectric recordings that should be appropriately processed [25, 26]. Lastly, appropriate processing of the heart rate [27, 28] can contribute to identify possible arrhythmias that can hinder identification.

V. CONCLUSIONS AND FUTURE WORK

In this paper we presented a simple template-based approach for the biometric identification of individuals using heartbeat morphology. The algorithm was preliminary tested on different individual recordings showing encouraging results (TAR=99.4%); however, further analyses are necessary to assess algorithm performance in case of a wider number of individuals or in case of physical activity. A region of convergence should be evaluated in order to give more elements to sensitivity analysis. Additionally, algorithm performance will be compared with other algorithms present in literature by using the same electronic database (e.g. MIT/BIH Physionet); indeed, a wide range of performance is present in literature (generally between 80 and 100% of TAR).

ACKNOWLEDGMENT

The authors would like to thank eng. Flavia Tizzano for the useful discussions on the literature analysis of biometric use of ECG.

REFERENCES

- [1] D. F. Dickinson, "The normal ECG in childhood and adolescence," *Heart*, vol. 91, no. 12, pp. 1626–1630, Dec. 2005. [Online]. Available: <http://heart.bmj.com/cgi/doi/10.1136/hrt.2004.057307>
- [2] G. Wbbeler, M. Stavridis, D. Kreiseler, R.-D. Bousseljot, and C. Elster, "Verification of humans using the electrocardiogram," *Pattern Recognition Letters*, vol. 28, no. 10, pp. 1172–1175, Jul. 2007. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0167865507000463>
- [3] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG analysis: a new approach in human identification," *Instrumentation and Measurement, IEEE Transactions on*, vol. 50, no. 3, p. 808812, 2001. [Online]. Available: <http://ieeexplore.ieee.org/xpls/absall.jsp?arnumber=930458>
- [4] 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, Jan. 2005. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0031320304002419>
- [5] Y. Wang, F. Agraftioti, D. Hatzinakos, and K. N. Plataniotis, "Analysis of human electrocardiogram for biometric recognition," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, no. 1, p. 148658, 2008. [Online]. Available: <http://asp.eurasipjournals.com/content/2008/1/148658>
- [6] J. M. Irvine and S. A. Israel, "A sequential procedure for individual identity verification using ECG," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 1–14, 2009. [Online]. Available: <http://asp.eurasipjournals.com/content/2009/1/243215>
- [7] P. Sasikala and R. S. D. Wahidauanu, "Identification of individuals using electrocardiogram," *International Journal of Computer Science and Network Security*, vol. 10, no. 12, p. 147153, 2010. [Online]. Available: <http://paper.ijcsns.org/07book=201012=20101220:pdf>
- [8] Y. N. Singh, "Evaluation of electrocardiogram for biometric authentication," *Journal of Information Security*, vol. 03, no. 01, pp. 39–48, 2012. [Online]. Available: <http://www.scirp.org/journal/PaperDownload.aspx?DOI=10.4236/jis.2012.31005>
- [9] S. Zokaei and K. Faez, "Human identification based on ECG and palmprint," *International Journal of Electrical and Computer*

- Engineering (IJECE), vol. 2, no. 2, p. 261266, 2012. [Online]. Available: <http://iaesjournal.com/online/index.php/IJECE/article/view/292>
- [10] A. D. C. Chan, M. M. Hamdy, A. Badre, and V. Badee, "Wavelet distance measure for person identification using electrocardiograms," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 2, pp. 248–253, 2008. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4427376>
- [11] S.-C. Fang and H.-L. Chan, "Human identification by quantifying similarity and dissimilarity in electrocardiogram phase space," *Pattern Recognition*, vol. 42, no. 9, pp. 1824–1831, Sep. 2009. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0031320308004913>
- [12] Y. N. Singh and P. Gupta, "Correlation-based classification of heartbeats for individual identification," *Soft Computing*, vol. 15, no. 3, pp. 449–460, Nov. 2009. [Online]. Available: <http://www.springerlink.com/index/10.1007/s00500-009-0525-y>
- [13] M. M. Tawfik and H. S. T. Kamal, "Human identification using QT signal and QRS complex of the ECG," *The Online Journal on Electronics and Electrical Engineering (OJEEE)*, vol. 3, no. 1, p. 383387, 2011. [Online]. Available: <http://infomesr.org/attachments/24-084.pdf>
- [14] M. Patel, S. Lal, D. Kavanagh, and P. Rossiter, "Applying neural network analysis on heart rate variability data to assess driver fatigue," *Expert Systems with Applications*, vol. 38, no. 6, pp. 7235–7242, Jun. 2011. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0957417410013916>
- [15] F. Agraftoti and D. Hatzinakos, "ECG biometric analysis in cardiac irregularity conditions," *Signal, Image and Video Processing*, vol. 3, no. 4, pp. 329–343, Sep. 2008. [Online]. Available: <http://www.springerlink.com/index/10.1007/s11760-008-0073-4>
- [16] M. Cesarelli, P. Bifulco, and M. Bracale, "Evaluating time-varying heartrate variability power spectral density," *Engineering in Medicine and Biology Magazine, IEEE*, vol. 16, no. 6, pp. 76–79, nov.-dec. 1997
- [17] Gargiulo G, Bifulco P, Cesarelli M, Ruffo M, Romano M, Calvo RA, Jin C, van Schaik A. An ultra-high input impedance ECG amplifier for long-term monitoring of athletes. *Medical Devices (Auckl)*. 2010;3:1-9
- [18] Gargiulo G, Calvo RA, Bifulco P, Cesarelli M, Jin C, Mohamed A, van Schaik A. A new EEG recording system for passive dry electrodes. *Clin Neurophysiol*. 2010 May;121(5):686-93
- [19] Gargiulo GD, Shephard RW, Tapson J, McEwan A, Bifulco P, Cesarelli M, Jin C, Al-Ani A, Wang N, van Schaik A. Pregnancy detection and monitoring in cattle via combined fetus electrocardiogram and phonocardiogram signal processing. *BMC Veterinary Research*. September 2012; vol.8: n.164. (DOI: 10.1186/1746-6148-8-164)
- [20] Gargiulo, G., Bifulco, P., Calvo, R.A., Cesarelli, M., Jin, C., Van Schaik, A. A mobile EEG system with dry electrodes. *IEEE-BIOCAS Biomedical Circuits and Systems Conference, BIOCAS 2008*, 2008; art. no. 4696927, pp. 273-276 (DOI: 10.1109/BIOCAS.2008.4696927)
- [21] Gargiulo G, Bifulco P, McEwan A, Nasehi Tehrani J, Calvo RA, Romano M, Ruffo M, Shephard R, Cesarelli M, Jin C, Mohamed A, van Schaik A. Dry electrode bio-potential recordings. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Volume 2010, 2010, Pages 6493-6496*
- [22] Gargiulo GD, Armin M, McEwan AL, Bifulco P, Cesarelli M, Jin C, Ruffo M, Tapson J, van Schaik A. Investigating the role of combined acoustic-visual feedback in one-dimensional synchronous brain computer interfaces, a preliminary study. *Medical Devices: Evidence and Research*. 2012;5(1):81–88.
- [23] Gargiulo GD, McEwan AL, Bifulco P, Cesarelli M, Jin C, Tapson J, Thiagalingam A, Van Schaik A. Towards true unipolar bio-potential recording: A preliminary result for ECG. *Physiological Measurement*. January 2013; 34(1):N1-N7. (DOI: 10.1088/0967-3334/34/1/N1)
- [24] Fratini A, Cesarelli M, Bifulco P, Romano M. Relevance of motion artifact in electromyography recordings during vibration treatment. *J Electromyogr Kinesiol*. 2009 Aug;19(4):710-8.
- [25] Fratini A, La Gatta A, Bifulco P, Romano M, Cesarelli M. Muscle motion and EMG activity in vibration treatment. *Med Eng Phys*. 2009 Nov;31(9):1166-72.
- [26] Cesarelli M, Fratini A, Bifulco P, La Gatta A, Romano M, Pasquariello G. Analysis and modelling of muscles motion during whole body vibration. *Eurasip Journal on Advances in Signal Processing* 2010, art. no. 972353 (DOI: 10.1155/2010/972353).
- [27] Cesarelli M, Romano M, Ruffo M, Bifulco P, Pasquariello G, Fratini A. PSD modifications of FHRV due to interpolation and CTG storage rate. *Biomedical Signal Processing and Control*. 2011 Jul.;6(3): 225-230
- [28] Cesarelli M, Romano M, Bifulco P, Improta G, D'Addio G. An application of symbolic dynamics for FHRV assessment. *Studies in Health Technology and Informatics*. 2012;180: 123-127