Applicability of Lead V_2 ECG Measurements in Biometrics

H.H.P. Silva[1], H.F.S. Gamboa[1], A.L.N. Fred[1]
[1]Instituto de Telecomunicações, {hugo.silva, hgamboa, afred}@lx.it.pt
Av. Rovisco Pais, 1, 1049-001 Lisboa, Portugal

Abstract: **This paper presents statistical analysis results for lead V_2 ECG measurements, focusing its applicability to Human identification. We analyze the subject recognition rate provided by mean heartbeat waves, through contingency matrix analysis, and extend these results by evaluating the recognition rate using only a reduced number of mean ECG heartbeat waveforms.**

Introduction

Typically the ECG has been used for clinical diagnostics of the cardiac function [1, 2]; however, measurements are influenced by physiologic factors which include: skin conductance, genetic singularities, position, shape, and size of the heart, among others. Figure 1 illustrates a prototypical heartbeat waveform.

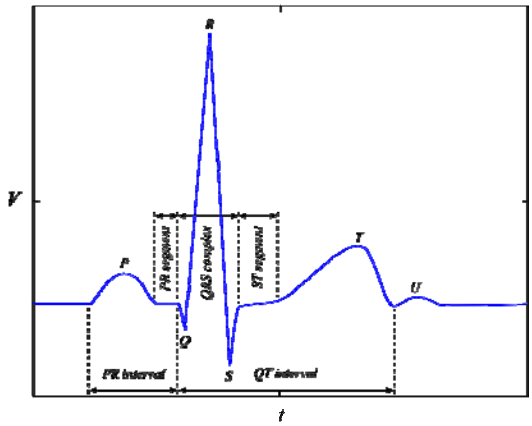

Figure 1. Prototypical ECG heartbeat waveform.

These singularities, which differ among individuals, make the ECG potentially interesting also for Human identification or identity verification purposes [3].

Acquisition Setup

Previous research work studying the ECG applicability for Human identification has typically explored readings from multiple leads [4], at rest [5], and at stress potentiating conditions [3]. Our work developed in the scope of a Human-computer interaction and multimodal electrophysiological signal acquisition setup, during the completion of a series of cognitive tasks, consisting of: an intelligence test, based on the Wisconsin sorting test, where the goal was to complete a logical sequence of four figures characterized by three properties (color, shape, and amount of symbols); a memory test, where the goal was to match hidden pairs of symbols; an association test, where the goal was to memorize a number of associations between abstract symbols and then reproduce them; a discovery test, where the goal was to identify an animal based on a singularity shown in a picture; and a concentration test inspired in a MENSA concentration test, were in two grids of 800 digits the goal was to mark every adjacent pair of digits that added 10.

A population of twenty five male and female volunteers, with a mean age of 23.4±2.5 years, participated in individual sessions where each was asked to complete the series of tasks, designed to take, in average, 30 minutes. The tasks were resolved using a computer, with which the subject merely interacted using the mouse as input device, and in a sitting position. ECG measurements were taken using a surface mount triode placed on the V_2 pre-cordial derivation.

Methodology

Each heartbeat waveform was sequentially segmented from the full recording, and after this, all individual waveforms were aligned by their *R* peaks. From the resulting collection of ECG heartbeat waveforms, the mean wave for groups of *10* heartbeat waveforms (without overlapping), was computed to minimize the effect of outliers. A labeled database was compiled, in which each pattern corresponds to a mean wave. For each mean waveform, the latency and amplitude for each of the *P-QRS-T* peaks were extracted, along with a sub-sampling of the waveform itself, providing a feature representation space of *53* features.

The ECG mean wave database is used for evaluation purposes; *50* data selection runs were performed, where in each run r, two mutually exclusive sets X_r, and Y_r are created, with respectively *10%* and *90%* of randomly selected patterns. For classification purposes, we use the k-NN decision rule with a Euclidean neighborhood metric [6]. A 1-NN neighborhood was adopted, since it is a particular case of the k-NN rule where the class w_x for a given pattern x is assigned as the class of the closest pattern from the training set X_r.

Results

Under the previously described framework, we computed the recognition error over *50* classification runs, where in each run individual mean waves from each test set Y_r were classified, using X_r as training set. We obtained an average *92±0.7%* subject recognition rate from a single mean wave. Figure 2 presents the contingency matrix for the subject classification error. The matrix diagonal corresponds to correct subject identification cases. As we can observe, the recognition rate is clearly above the *50%* threshold for all subjects. Therefore by majority voting the contingency matrix, a *100%* subject recognition rate is achieved.

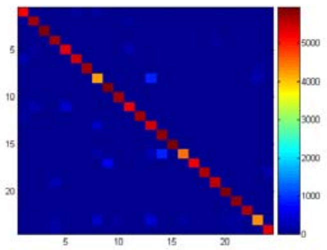

Figure 2. Contingency matrix for the recognition rate over *50* runs.

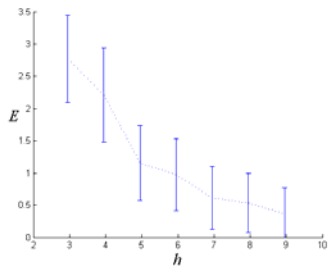

Figure 3. Sequential classifier combination results by combining a limited number *h* of mean waves.

To further improve the recognition rate from individual mean waves, we applied sequential classifier combination techniques to the problem. Instead of producing a decision based on a single mean wave, we consider a limited number of mean waves, classify each one of them individually using the 1-NN classifier, and apply majority voting to the group of individual decisions in order to obtain an overall subject identification decision. Figure 3 illustrates the evolution of the recognition error as the number *h* of individual mean waves is increased.

As we can observe, by combining the decisions of multiple mean waves, the recognition rate is highly improved. The highest recognition rate of *99.63±0.4%* was obtained for a set of *9* mean waves, which according to the adopted methodology, and in which each mean wave is computed from a group of *10* heartbeat waveforms, corresponds to approximately *90s* of signal acquisition (considering an average heart rate of *60bpm*).

Conclusions

In this paper we presented experimental results from a real world problem explored by our group, regarding Human identification using ECG signal. Preliminary results have shown that considering mean ECG waves corresponding to *10* heartbeat waveforms, it is possible to achieve *100%* recognition rates through contingency matrix analysis.

Furthermore, we evaluated the discriminative potential of groups of mean waves, by applying sequential classifier combination methods to the problem. Using this approach it was possible to obtain recognition rates above *99%*, from approximately *90s* of acquired signal.

Acknowledgments

This work was partially supported by the Portuguese Foundation for Science and Technology (FCT), Portuguese Ministry of Science and Technology, and FEDER, under grant POSI/EEA-SRI/61924/2004, and by the Institute for Systems and Technologies of Information, Control and Communication (INSTICC) and the Instituto de Telecomunicações (IT), Pólo de Lisboa, under grant P260.

References

[1] E. Chung. Pocket Guide to ECG diagnostics. Blackwell Publishing Professional, 2000.

[2] B. Lipman. ECG acessment and interpretation. F. A. Davis, 1994.

[3] S. Israel, J. Irvine, A. Cheng, M. Wiederhold, and B. Wiederhold. "ECG to identify individuals", Pattern Recognition, vol. 38, 2005, pp. 133-142.

[4] T. Shen, and W. Tompkins. "Biometric statistical study of one-lead ECG features and body mass index", Proceedings of the 27th Annual Conference of the IEEE Engineering in Medicine and Biology Society, 2005, pp. 1162-1165.

[5] L. Biel, O. Petterson, L. Phillipson, and P. Wide. "ECG analysis: A new approach in Human identification", IEEE Transactions on Instrumentation and Measurement, 2001, pp. 808-812.

[6] R. Duda, P. Hart, and D. Strork. Pattern classification 2nd ed. John Willey & Sons Inc, 2001.