

# Verification of humans using the electrocardiogram

Gerd Wübbeler \*, Manuel Stavridis, Dieter Kreiseler, Ralf-Dieter Bousseljot, Clemens Elster

*Physikalisch-Technische Bundesanstalt, Abbestr. 2-12, 10587 Berlin, Germany*

Received 7 October 2005; received in revised form 24 January 2007

Available online 8 February 2007

Communicated by G. Sanniti di Baja

## Abstract

A feasibility study on the potential of the electrocardiogram (ECG) for biometrical applications is presented. A test set of 234 ECG recordings from 74 subjects was compiled emulating a realistic scenario for ECG biometrics by using short measurements of 10 s length in combination with a practicable choice of ECG leads. The long-term stability of the individual ECG was investigated during time periods up to several years. Verification and identification was done by utilizing the heart vector and a simple distance measure. As a result, encouraging error rates were obtained; for verification, for instance, the achieved equal error rate was smaller than 3%.

© 2007 Elsevier B.V. All rights reserved.

**Keywords:** Biometrics; Electrocardiogram (ECG); Verification; Identification

## 1. Introduction

The performance of biometrical systems has been improved significantly during the last years (Yager and Amin, 2004; Wayman et al., 2005) and established approaches such as fingerprint or iris-scan are meanwhile tested in real-world applications. The operation of a biometrical system requires the compliance of different demands such as low error rates to achieve high security levels or the possibility of fake detection and testing for liveness (Ortega-Garcia et al., 2004). Moreover, the short- and long-term stability of the selected biometric feature or the user comfort during sampling have to be ensured. Hence the improvement of existing techniques as well as the development of new biometrical approaches, relying on yet unused physiological or behavioral features, is needed. The search for new biometric features is also of great interest due to the possibility to combine different and independent modalities to so-called multi-biometric systems thereby opening the way for further essential

improvements of the security level as well as better compliance of the large number of the partly diverging demands (Ross and Jain, 2003).

Recently, the use of the electrocardiogram (ECG) has been suggested as an additional tool for biometric applications (Biel et al., 2001). The ECG is long known for its diagnostic relevance (Chou, 1976). During the last decades there has been an increasing effort to develop computer based automatic diagnostics of the ECG (MacFarlane et al., 1990). In addition to its diagnostic value, the ECG is known to contain characteristics originating from the geometrical features of the individual body and heart that might be utilized for biometrical applications (Hoekema et al., 2001). A disadvantage of the ECG are inherent variations of different heartbeats of the same subject.

So far, only few studies treating the ECG for human identification can be found in literature (Biel et al., 2001; Kyoso and Uchiyama, 2001; Yi et al., 2003; Israel et al., 2003; Israel et al., 2005). Biel et al. (2001) used the diagnostic parameters recorded with a commercial ECG device to discriminate 20 subjects. In a recent study comprising 29 subjects, Israel et al. (2005) focused on more physiologically founded parameters describing the characteristic timings of the ECG signals. They also investigated the

\* Corresponding author. Tel.: +49 30 3481 7242; fax: +49 30 3481 7030.  
E-mail address: [gerd.wuebbeler@ptb.de](mailto:gerd.wuebbeler@ptb.de) (G. Wübbeler).

influence of different anxiety states on the identification of humans by their heartbeat characteristics.

In this work, we present a feasibility study which carries on quantifying the potential of the ECG for biometric person recognition. The main focus was laid on the assessment of real world performance. Consequently, three essential topics were dealt with: the long-term stability of the individual ECGs, a short recording time and the usage of easily applicable ECG leads hardly affected by actual positioning of the electrodes. While the first point is mandatory for the biometrical application of the ECG, the latter two are vital prerequisites for the implementation of a reliable and practicable device. Following this, the study was based on a test set compiled from ECG data of 74 subjects, repetitively recorded mainly at intervals of several months up to several years. The recording time was fixed to a length of 10 s; each recording session typically covered a set of 10 heart beats only. Almost position invariant measurements were achieved by recording ECG signals from three leads fixed to the extremities according to the Einthoven's triangular scheme (Nelson and Geselowitz, 1976). While in previous studies on ECG biometrics a set of features extracted from the ECG was utilized (Biel et al., 2001; Israel et al., 2005), in this work the pattern recognition task was performed by comparing the ECG traces themselves. Due to the short recording time no statistics of the individual heart beat were calculated; rather a single representative heart beat was selected automatically and used for ECG biometrics.

## 2. ECG data

The ECG data used in this work were taken from a large database containing more than 27000 ECG recordings, operated at the Physikalisch-Technische Bundesanstalt (PTB) for the development of computer based diagnostics (Bousseljot, 1994). The data were recorded with the subject in a supine position in a resting state. A test set was compiled from this database by selecting subjects which were repetitively recorded and, in addition, diagnosed as healthy by a physician. The test set contained 234 ECG recordings from 74 Caucasian subjects, 40 male and 34 female with an average age of 45.5 years ranging from 19 years to 86 years. The number of ECG recordings per subject varied from 2 up to 20 with most of the subjects recorded for 2 or 3 times (Table 1). The average time interval between any two recordings of the same subject was about 500 days (16.6 months). The heart rate within the 234 ECG recordings varied from 40 bpm (beats per minute) up to 90 bpm with intraindividual variations, observed between different recordings of the same subject, ranging from 0.4 bpm up to 31 bpm. Fig. 1 shows single heart beats from three ECG

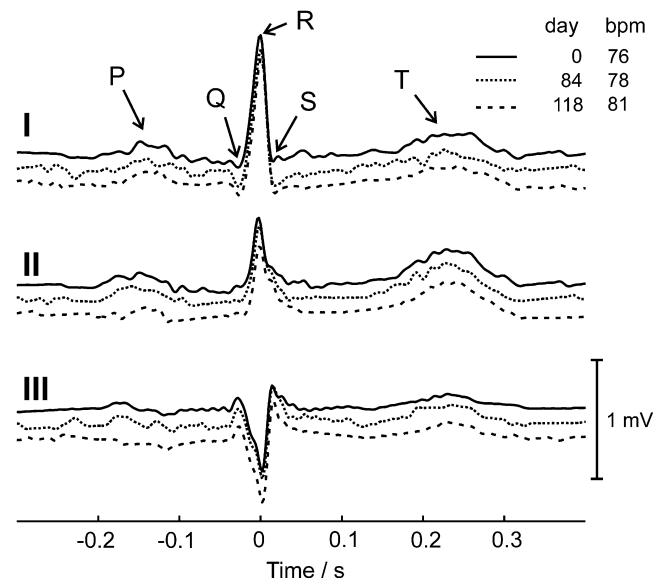


Fig. 1. Single heart beats from three ECG recordings of one subject performed during 4 months: ECG traces of the three Einthoven leads (I, II, III) are shown together with the recording day and heart rate in beat per minutes (bpm).

recordings of one subject performed during 4 months. Note the stability of the ECG traces during the QRS interval.

With regard to a practicable biometric system ECG traces from three channels, recorded according to Einthoven's triangular scheme (Nelson and Geselowitz, 1976), were used in this study. With the electrodes fixed to the left and right arm and to the left leg this configuration warrants fast and easily realizable recordings of ECG traces which are widely independent of the actual positioning of the electrodes (Schijvenaars et al., 1997). The sampling rate of the data used in this study was 500 Hz.

Standard signal processing techniques were applied for preprocessing, which were not optimized with respect to the extraction of biometric features. The ECG traces of 10 s length from the three recorded channels were baseline corrected by subtracting a moving median of 1 s width. Additionally, a low-pass filter was applied to each channel with a cut-off frequency of 75 Hz. Since the observed voltage levels as well as their channel-to-channel ratios presumably also contain discriminative features the data of the three channels were not normalized to a fixed amplitude (Hoekema et al., 2001). For beat detection, the positions of the R-peaks were determined from the absolute value of the low-pass filtered temporal derivative of each ECG trace using a threshold procedure.

## 3. Verification and identification

Since the number of recordings per subject in the test set varied largely (Table 1) special care was taken to perform the assessment of the biometrical potential of the ECG in a balanced manner. Therefore, repetitively two disjoint sets were drawn from the test set, each containing 74 ECGs

Table 1  
Distribution of the number of recordings per subject

Number of recordings	2	3	4	5	6	7	10	13	20
Number of subjects	44	16	5	3	1	2	1	1	1

MASIMO 1078

Masimo v. Apple

IPR2023-00745

representing each subject by one randomly selected recording. By analyzing the distances between each ECG in the first set and each ECG in the second set, error rates were calculated in a usual one-to-many, respectively one-to-one scenario with identification and verification being carried out by standard nearest neighbor and threshold schemes (Wayman et al., 2005). The error rates were stabilized to a level of about  $\pm 0.01\%$  by averaging over a large number ( $N = 10000$ ) of drawings of two disjoint sets.

The distance between two ECGs was determined utilizing one single heart beat of the two-dimensional time-dependent heart vector (see below) which is known as a characteristic of the ECG (MacFarlane and Veitch Lawrie, 1989). Shape and timing of the QRS-complex is known to undergo only weak variations with changing heart rates (Mücke, 1996); hence, we confined the trace of the heart vector to an interval of 100 ms length centered at the latency of the R-peak in order to obtain a biometric characteristic nearly invariant to the heart rate and, correspondingly, to the anxiety state. The interval length of 100 ms was not tuned with respect to the biometric task, rather it reflects the average length of a normal QRS interval.

The two-dimensional heart vector  $h(t) = (h_x(t), h_y(t))^T$  was determined from the three time-dependent preprocessed channels (Einthoven leads)  $I(t)$ ,  $II(t)$  and  $III(t)$  according to

$$\begin{pmatrix} h_x(t) \\ h_y(t) \end{pmatrix} = \begin{pmatrix} I(t) \\ [II(t) + III(t)]/\sqrt{3} \end{pmatrix}. \quad (1)$$

Fig. 2 shows traces of 100 ms length of the heart vector for three subjects, each recorded for three times (subject 1 from Fig. 1). Note the characteristic heart vector pattern of each subject, to a great extent being stable even during several years and, moreover, with changing heart rates.

The distance between two ECGs was calculated utilizing the trace of the two time-dependent heart vectors of a single selected heart beat as well as their first and second tem-

poral derivatives. The derivatives were determined for each component of the heart vector by differentiation of spline interpolants. The distance  $d$  was then calculated as

$$d = \min_{\Delta t} \sqrt{\{d_0^2(\Delta t) + d_1^2(\Delta t) + d_2^2(\Delta t)\}}, \quad (2)$$

where

$$d_v(\Delta t) = \frac{\sum_t \|h^v(t) - \tilde{h}^v(t - \Delta t)\|}{\sum_t \|h^v(t)\| + \sum_t \|\tilde{h}^v(t - \Delta t)\|}, \quad v = 0, 1, 2, \quad (3)$$

$h^v(t)$ ,  $\tilde{h}^v(t)$  denote the  $v$ th derivatives of the two heart vectors  $h(t)$ ,  $\tilde{h}(t)$  from the two ECGs to be compared, and  $\|\cdot\|$  stands for the Euclidean norm. The sum in (3) spans the range of an interval of 100 ms length centered at the position of the R-peak of a selected, single beat, and  $\Delta t$  accounts for a possible (small) shift between the two ECGs traces caused by inaccurate determination of the R-peak latencies. Beat selection was done as follows: For each ECG that beat was chosen for which the sum of the distances according to (2) to all other single beats of the *same* ECG recording is minimum.

#### 4. Results

Fig. 3 shows the empirical distributions of distances between ECGs from the same (genuine) or from different subjects (imposters). The two distributions are distinctly separated indicating a reasonable verification performance achievable by fixing the threshold to the abscissa of the intersection point of the two curves. The remaining overlap may partially be explained by the inherent variability of the individual heart beat leading to larger distances also for the genuine case. In addition, casual similarity of ECG patterns from different subjects cannot be ruled out.

From the distances the false matching rate (FMR) and false non-matching rate (FNMR) were calculated for various threshold values to generate the detection error trade-off (DET) curve of the verification process (Fig. 4). The error rates determined for verification are summarized in Table 2. Note that an equal error rate (EER) smaller than 3% was obtained.

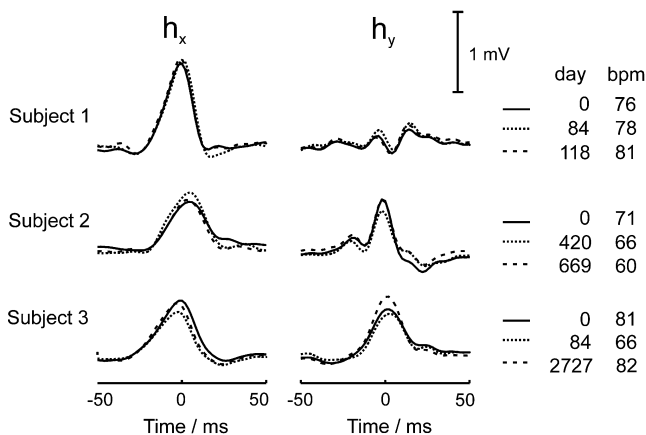


Fig. 2. Trace of the heart vector  $(h_x, h_y)^T$  during the QRS interval for three subjects (subject 1 from Fig. 1), each recorded for three times. Recording day and heart rate (bpm) are indicated.

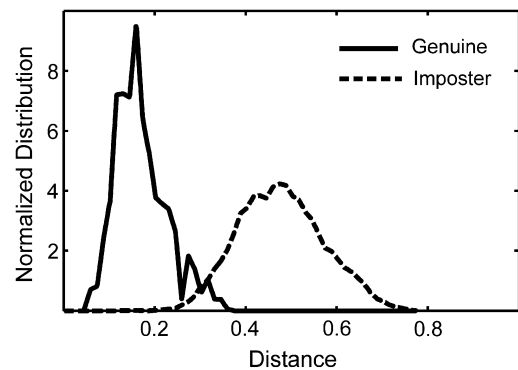


Fig. 3. Empirical distance distributions between the ECGs as calculated from the heart vector.

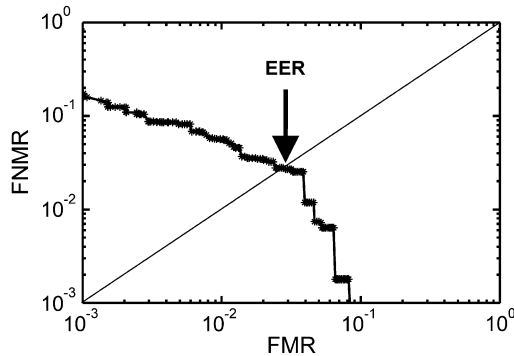


Fig. 4. Detection error trade-off (DET) curve for verification using ECG. The location of the equal error rate (EER) point is indicated.

Table 2

Verification error rates within the ECG test set

EER	FMR at FNMR = 10%	FMR at FNMR = 3%
2.8%	0.2%	2.5%

Table 3

Identification rates for the specified rank within the ECG test set

Rank 1	Rank 2	Rank 3
98.1%	98.8%	99.0%

In addition to verification, identification performance was determined by analyzing the nearest neighbors, found for each ECG according to the distance of corresponding heart vectors. In Table 3, the identification rates specified for finding the same subject at least at the indicated rank are presented.

The obtained results are encouraging and demonstrate the potential of heart beat signals for biometric applications. Further studies should address the optimization of the matching procedure in combination with larger data sets, additionally covering a realistic distribution of cardiac diseases. Likewise, if possible, the number of electrodes needed as well as the length of the recording time should be reduced. For the realization of an actual biometric system the development of easily applicable electrodes is needed to achieve a reasonable level of user comfort (cf. the EU-project HUMABIO, [www.humabio-eu.org](http://www.humabio-eu.org)). Ideally, such a system requires only to be touched by the hands or two fingers while still providing the required signal quality.

## 5. Summary and conclusion

This work presented a feasibility study to assess the real world performance of human verification and identification based on characteristics of the electrocardiogram (ECG). Particularly the long term stability of the individual ECG was focused on by taking data from 74 subjects repetitively

recorded during several months up to several years. A realistic and easily applicable protocol for ECG biometrics was emulated by using recordings of 10 s length taken only from the three Einthoven leads which are hardly affected by the actual positioning of the electrodes. Utilizing a distance measure between the ECGs which is based on the evolution of the corresponding heart vectors during the QRS interval, error rates were calculated with identification and verification being carried out by standard nearest neighbor and threshold schemes. Application to the test set yielded encouraging error rates; for verification an equal error rate of 2.8% was achieved and for identification a rate of 98.1% (first rank). The individual ECG might therefore be a useful candidate as an add-on to improve established biometrical systems.

## References

- Biel, L., Pettersson, O., Philipson, L., Wide, P., 2001. ECG analysis: A new approach in human identification. *IEEE Trans. Instrum. Meas.* 50 (3), 808–812.
- Bousseljot, R., 1994. Aufbau der EKG-Datenbank CARDIODAT der PTB (Setting up of the ECG database CARDIODAT of PTB). *Biomed. Tech.* 39 (Suppl. 1), 250–251.
- Chou, T.C., 1976. *Electrocardiography in Clinical Practice*. Grune and Stratton, New York.
- Hoekema, R., Uijen, G.J.H., van Oosterom, A., 2001. Geometrical aspect of the interindividual variability of multilead ECG recordings. *IEEE Trans. Biomed. Eng.* 48, 551–559.
- Israel, S.A., Scruggs, W.T., Worek, W.J., Irvine, J.M., 2003. Fusing face and ECG for personal identification. In: *Proc. 32nd Applied Imagery Pattern Recognition Workshop*, 2003, pp. 226–231.
- Israel, S.A., Irvine, J.M., Cheng, A., Wiederhold, M.D., Wiederhold, B.K., 2005. ECG to identify individuals. *Pattern Recognition* 38 (1), 133–142.
- Kyoso, M., Uchiyama, A., 2001. Development of an ECG identification system. In: *Proc. 23rd Ann. Internat. Conf. on the IEEE Engineering in Medicine and Biology Society*, vol. 4, pp. 3721–3723.
- MacFarlane, P.W., Veitch Lawrie, T.D., 1989. In: *Comprehensive Electrocardiology*, vol. 1. Pergamon Press, New York.
- MacFarlane, P.W., Devine, B., Latif, S., McLaughlin, S., Shoaib, D.B., Watts, M.P., 1990. Methodology of ECG interpretation in the Glasgow Program. *Methods Inf. Med.* 29, 354–361.
- Mücke, D., 1996. *Elektrokardiographie Systematisch*. Uni-Med Verlag AG, Lorch/Württemberg.
- Nelson, C.V., Geselowitz, D.B., 1976. *The Theoretical Basis of Electrocardiology*. Clarendon Press, Oxford.
- Ortega-Garcia, J., Bigun, J., Reynolds, D., Gonzalez-Rodriguez, J., 2004. Authentication gets personal with biometrics. *IEEE Signal Process. Mag.* 21, 50–62.
- Ross, A., Jain, A., 2003. Information fusion in biometrics. *Pattern Recognition Lett.* 24, 2115–2125.
- Schijvenaars, B.J., Kors, J.A., van Herpen, G., Kornreich, F., van Bommel, J.H., 1997. Effect of electrode positioning on ECG interpretation by computer. *J. Electrocardiol.* 30, 247–256.
- Wayman, J., Jain, A., Maltoni, D., Maio, D. (Eds.), 2005. *Biometric Systems Technology, Design and Performance Evaluation*. Springer, London.
- Yager, N., Amin, A., 2004. Fingerprint verification based on minutiae features: A review. *Pattern Anal. Appl.* 7, 94–113.
- Yi, W.J., Park, K.S., Jeong, D.U., 2003. Personal identification from ECG measured without body surface electrodes using probabilistic neural networks. In: *World Congress on Medical Physics and Biomedical Engineering*, August 2003, Sydney, Australia.