Athlete Identification using Acceleration and Electrocardiographic Measurements Recorded with a Wireless Body Sensor

Peter Christ[1], Felix Werner[1], Ulrich Rückert[1] and Jörg Mielebacher[2]

[1]*Cognitronics and Sensor Systems Group, CITEC, Bielefeld University, Bielefeld, Germany*

[2]*Mielebacher Informatiksysteme, Fahrenheitstr. 1, Bremen, Germany*

{pchrist, fwerner, rueckert}@cit-ec.uni-bielefeld.de, mail@mielebacher.de

Abstract: In this paper we propose a biometric method for identifying humans during walking and jogging. We use acceleration and electrocardiographic measurements recorded with a wireless body sensor attached to a chest strap. Our method does not require a particular acquisition setup. Information on the gait style and on the physiology is combined to identify a human despite severe motion related artefacts in the electrocardiograph and variations in the gait patterns. We propose to identify humans using features extracted in time and frequency domain and a standard classifier. With the collected data of 22 subjects on a treadmill at velocities from 3 to 9 km/h we obtained an accuracy of 98.1 %. The sensitivity of the identification ranged between 94.6 to 99.5 % for the different subjects and the specificity was higher than 99.7 %.

1 INTRODUCTION

The identification of humans is important for various applications such as surveillance systems, authorization checks at doors or electronic devices (e.g. computer, smartphone). A variety of biometric characteristics have been investigated such as information from fingerprint, iris and retina, human face, voice, gait or electrocardiograph.

Previous work has shown that discerning, reproducible information on the human is found in the ECG waveform, especially around the QRS complex (Conover, 2002; Chan et al., 2008). Moreover, biomechanical differences between the gait style of humans have been investigated and used for identification within video and acceleration sensor based applications (Gafurov et al., 2006; Nixon et al., 2006).

We propose a biometric measure combining both sources of information: characteristics in the electrocardiograph (ECG) waveform and the gait style. Unlike other applications, our approach focuses on the identification of athletes during physical exercise using a compact wireless body sensor (WBS) which is worn around the chest (see Figure 1). The WBS is typically used to measure the heart-rate and the body accelerations of athletes. Our identification method additionally utilizes the sensor measurements to identify the athlete, enabling an automatic annotation of

sensor data with the subject's identity. Our goal is to overcome the drawbacks of a manual annotation of measurements for applications in sports medicine and athlete training research. Furthermore, recognizing the subject allows to automatically load personal settings on the WBS or the sport equipment for a customized training. Our identification method is in particular interesting for a WBS which is used with several athletes of a mid-sized group.

Our identification method uses features in time and frequency domain to extract characteristics on the subject which are used as input to a classifier for identification. By combining information from gait and ECG we can successfully identify subjects despite of artefacts in the ECG caused by a slipping of the ECG electrodes and severe variations in the gait patterns between walking and jogging.

Previous work in this field focused on the identification of humans from either gait or ECG waveform characteristics. Mainly ECGs were used which were recorded at rest or with a clinical acquisition setup. The gait based identification was carried out for walking velocities.

Rong et. al (2007) proposed a method which uses measurements recorded during walking with an accelerometer located at the subject's waist. The method utilises a segmentation into gait cycles to extract gait patterns. Dynamic time warping is applied
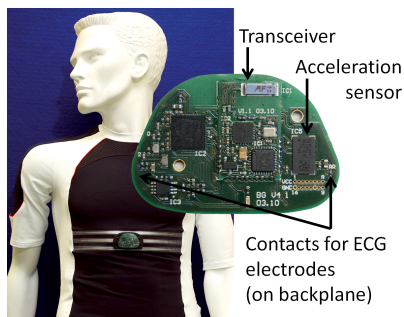
Figure 1: Our self-made wireless body sensor (WBS) and its integration into a chest strap. The WBS can measure a person's electrocardiograph (ECG) and accelerations of the body along three orthogonal axes.

to compensate natural changes in walking speed. The actual gait segment is then compared with a reference pattern of the subject and a 1-nearest neighbour classifier is used to recognize the subject. Ailisto et. al (2005) evaluated an accelerometer based identification based on similarities between gait segments to protect portable devices. Mäntyjärvi et. al (2005) evaluated a gait based identification for different walking velocities using correlation coefficients derived from a template comparison, frequency coefficients and a histogram based comparison. Gafurov et. al (2006) proposed two methods based on histogram similarity and gait cycle length to distinguish acceleration measurements recorded at the lower leg.

Several methods have been proposed to identify a human based on ECG measurements. Biel et. al (2001) used data from a standard 12-lead ECG recorded during rest to identify subjects using multivariate analysis. Furthermore, the study showed that identification is possible with even one-lead ECGs. Shen et. al (2002) also utilises data from one-lead ECGs to distinguish subjects using a template matching and a decision-based neural network. Chan et. al (2008) identifies subjects based on ECGs recorded within a non-clinical acquisition setup where the subjects were holding two electrodes on the pads of their thumbs. For classification, three qualitative measures were used: percent residual difference, correlation coefficient, and a novel distance measure based on wavelet transform.

This paper is organized as follows: Section 2 describes the identification of a subject based on acceleration and ECG measurements. Information on preprocessing, feature extraction and used classifiers is given. Section 3 explains the conducted experiment for data collection. Section 4 presents the experimental results of our identification method. The results are summarised and discussed in section 5 and a prospect on our future work is given.
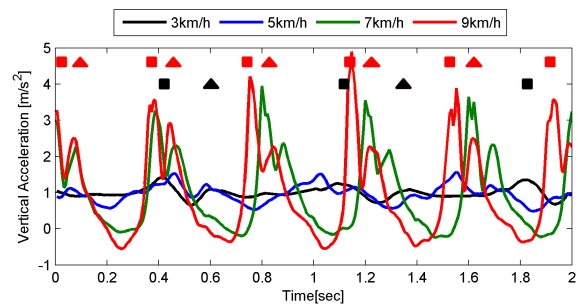

Figure 2: Vertical acceleration data of a subject walking and jogging at velocities from 3 to 9 km/h. Each stride is represented by two consecutive peaks which correspond to the heel strike (square) and the toe strike (triangle). These peaks are marked for 9 km/h (red) and for 3 km/h (black). Velocity can be increased with either longer strides (increase in signal amplitude) or a higher step frequency.
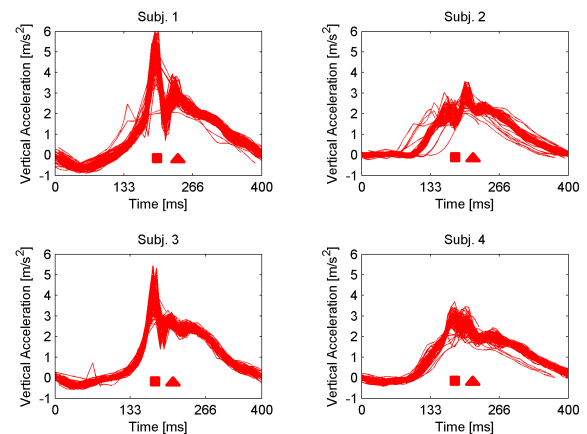

Figure 3: Alignment of 100 consecutive strides of four subjects jogging at 9 km/h. The vertical acceleration signals were automatically segmented into strides and crosscorrelation was used to align the strides. The peaks related to the heel strikes (square) and toe strikes (triangle) significantly differ in shape between the subjects.

2 IDENTIFICATION OF A SUBJECT

This section describes the identification of a subject based on gait style and ECG waveform characteristics. We describe the preprocessing of the signals, the feature extraction and the classifiers used for identification.

2.1 Gait Analysis for Identification

Previous work has shown that gait differs between humans and that the gait style is fairly stable for a subject (Bianchi et al., 1998; Nixon et al., 2006). Bianchi
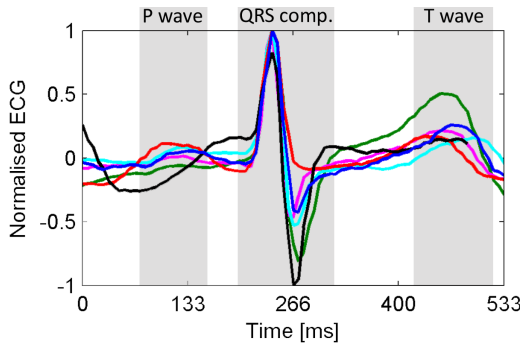
Figure 4: Comparison of heartbeat segments of six subjects (different colours). The DC-offset was removed and the heartbeat segments were aligned using cross-correlation. We use inter-subject variations in the ECG waveform to identify subjects.

et. al (1998) stated that the variability across humans depends on different kinematic strategies rather than on biomechanical characteristics. Their study showed that subjects are different in the ability of minimising energy oscillations of their body segments for transferring mechanical energy.

In order to measure these inter-subject differences, severe intra-subject variations in the gait patterns between walking and jogging have to be taken into account. The intra-subject variations are a result of an adaptation of the gait to achieve different velocities. The velocity of a person is described by stride length and stride frequency. According to Weyand et. al (2000), longer strides are achieved by applying greater support forces to the ground which significantly increases the amplitude of the vertical acceleration signal, whereas the step frequency changes frequency components of the signal.

Samples of vertical acceleration data of one subject walking and jogging at different velocities on a treadmill are shown in Figure 2. Strides are presented by two consecutive peaks corresponding to the heel and toe strikes. Significant changes in amplitude and an almost doubling of the step frequency can be observed between walking at 3 km/h and jogging at 9 km/h.

Despite of this intra-subject variability in the gait patterns, we observed inter-subject variations in acceleration signals recorded during walking and jogging (Christ et al., 2011). In particular, heel and toe strikes differ in the vertical acceleration signal's shape between subjects (see Figure 3). The peak acceleration of the heel strikes varies between the four subjects about 2 m/s^2.

2.2 ECG Analysis for Identification

Inter-subject variability is also found in the ECG's waveform. The variations depend on position, size and anatomy of the heart, age, sex, relative body weight, chest configuration and various other factors (Green et al., 1985; Simon and Eswaran, 1997). Figure 4 shows sample heartbeat segments from six subjects recorded with our WBS. The ECG reflects the electrical activity of the heart and consists of the P wave followed by the QRS complex and the T wave (Conover, 2002, chap. 2). Discerning information on the subjects is found in the QRS complex, the P and the T wave.

Chan et. al (2008) observed a high degree of reproducibility of information extracted from the QRS complex of a person through several sessions of recording. Furthermore, a higher identification accuracy was determined for the P wave than the T wave.

During physical exercise these characteristics can be superposed by motion related artefacts. These artefacts are caused by a slipping of the ECG electrodes and variations in the contact resistance during body movements (Christ et al., 2010). Figure 7 shows disturbances in the ECGs of two subjects recorded during jogging on a treadmill.

2.3 Preprocessing of Acceleration and ECG Signals

ECGs recorded with our WBSs showed hardware-related differences in the DC-offset making an ECG associable to a WBS. Furthermore, using textile ECG electrodes, the skin contact resistance decreases over time because of an increased transpiration which results in changes in the DC-offset. In order to avoid classification errors, we removed the DC-offset using a 4th-order high-pass butterworth filter with a cutoff frequency of $f_c = 0.67$ Hz. Additionally, we applied a low-pass filter with a cutoff frequency of $f_c = 40$ Hz to remove noise in the ECG signal.

With a decrease in skin contact resistance after a few minutes of exercise, we observed an increase in the ECG signal's amplitude which improved the signal-to-noise ratio. We normalised the signal's amplitude to assure that ECG segments are comparable. The results of the ECG preprocessing are shown in Figures 5 a and b.

For the frequency analysis of the acceleration measurements, we approximated the dynamic accelerations by applying a 4th-order butterworth high-pass filter with a cutoff frequency of $f_c = 0.1$ Hz to the magnitude of the acceleration vector $\mathbf{a} = (a_{AP}, a_{ML}, a_V)$; a_{AP} denotes anteroposterior accelera-

(a) ADC output and preprocessed ECG of subject 1.

(b) ADC output and preprocessed ECG of subject 2.

(c) Magnitude of the acceleration vector **a** and the offset reduction by the high-pass filter.
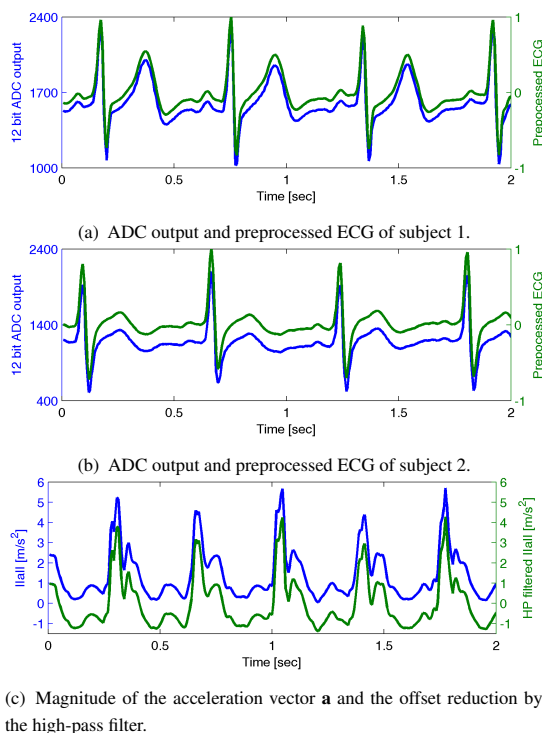
Figure 5: The 12 bit analog-to-digital converter (ADC) output and the preprocessed ECG in comparison. An offset of 300 between the ADC output of the two different subjects was removed by the preprocessing. In Figure 5 c the offset due to the static acceleration of gravity and a sensor-related zero-g-level offset are reduced after preprocessing.

2.4 Feature Extraction for Identification

In order to access characteristics of a subject in the acceleration and ECG measurements, we extracted features in the time and the frequency domain.

The features were calculated within a sliding window with no overlap and length N. Each window at time t consists of N measurements $x(t:t+N-1) = x(t), x(t+1), ..., x(t+N-1)$. We empirically determined an appropriate window length of two seconds ($N = 300$).

Time Domain Features. In the time-domain we calculated the variance, amplitude, mean and root mean square (RMS) along the three orthogonal axes a_{AP}, a_{ML} and a_V of the windowed acceleration signals.
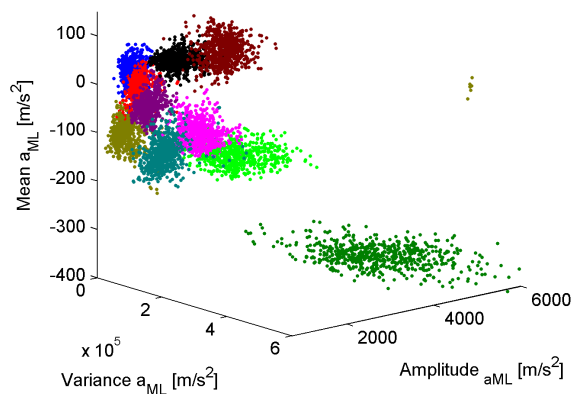

Figure 6: Visualisation of time domain features extracted from mediolateral accelerations a_{ML} of ten subjects at 9 km/h. Clusters are observable for the different subjects. In our feature selection we obtained a good identification performance based on the mean, the variance, the amplitude and the root-mean-square (RMS) features (see Table 3).
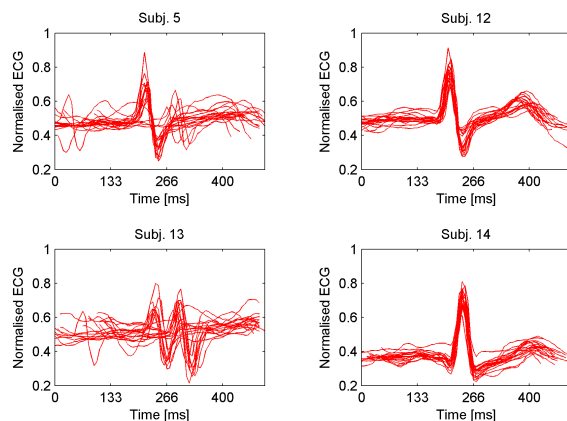

Figure 7: Alignment of 20 heartbeat segments of four subjects recorded during jogging on a treadmill. A correct placement of the chest strap is important for an identification based on a similarity measure between heartbeat segments. Motion related artefacts and poor skin contact can disturb the ECG-signal (see subjects 5 and 13).

The variance, mean and amplitude of a_{ML} are visualised in Figure 6. Discriminative clusters can be observed for the different subjects. From the ECG signal we calculated a feature measuring the closeness of an unknown heartbeat segment to five reference patterns stored for each subject. This step requires a segmentation of the ECG signal into heartbeats. We used a QRS detection based on the algorithm of (Afonso et al., 1999) in its implementation of Schloegl in the BioSig toolbox (Vidaurre et al., 2011). The five reference heartbeat segments were chosen randomly from the ECG data of each subject. However, we assured that only heartbeat segments without severe disturbances were chosen. For identification, an unknown segment x was aligned to each reference segment y
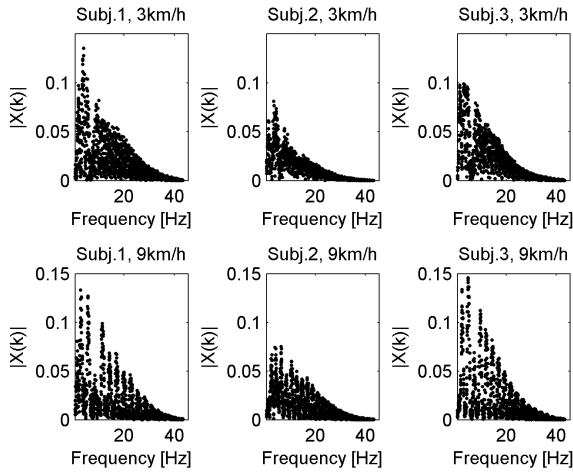
14

Figure 8: The FFT amplitude spectra of the ECG signals of three subjects during walking (3 km/h) and jogging (9 km/h). The amplitude spectra show differences between the subjects but also vary with the velocity.

using cross-correlation:

$$R_{xy}(m) = \frac{1}{N} \sum_{j=0}^{N-m-1} y(j+m)x(j) \qquad (1)$$

where N is the length of a segment and m the offset with $m = 0, 1, ..., 2N-1$. We calculated the Pearson's correlation coefficient as a measure of similarity between the two segments. The Pearson's correlation coefficient is defined as the covariance (cov) of the two segments divided by the product of their standard deviation σ:

$$r(x,y) = \frac{\text{cov}(x,y)}{\sigma_x \sigma_y} \qquad (2)$$

Figure 7 shows the alignment of 20 heartbeat segments of four subjects. The QRS-detection and the alignment are sensitive to motion-related artefacts (see subjects 5 and 13).

For heartbeat segments without major disturbances the alignment centred the segments around the QRS complex. The discerning information in this region of the ECG is fairly stable in relation to morphology changes in the ECG waveform during effort.

Frequency Domain Features. In the frequency domain we use the discrete Fourier transform (DFT) to extract frequency components of each window. The DFT is defined as:

$$X(k) = \sum_{j=t}^{t+K-1} x(j)e^{-i2\pi k \frac{j}{K}}, \quad k = 0, ..., K-1 \qquad (3)$$

where K is the number of outputs $X(k)$. We used a 512-point fast Fourier transform (FFT) algorithm to compute the DFT efficiently for our windows of the

length $N = 300$. Therefore, each window $\mathbf{x}(t:t+N-1)$ was padded with trailing zeros to the length of $K = 512$. Before calculating the FFT, a Hamming window function was applied to each window to reduce the spectral leakage.

Figure 8 shows the FFT amplitude spectra of ECGs of three subjects recorded during walking (3 km/h) and jogging (9 km/h). Despite of velocity related variations in the amplitude spectra, differences can be observed between the three subjects.

We calculated additional frequency domain features from the amplitude spectrum (FFT features): the variance, the mean, the Fourier coefficient with the highest amplitude and the Shannon entropy SE:

$$SE = -\sum_{k=0}^{K-1} |X(k)| \log_2(|X(k)|) \qquad (4)$$

where $X(k)$ is the output of the DFT of length K.

2.5 Methods for Classification

We used a standard classifier to identify the subject based on the extracted features. The identification performance was determined by evaluating three different classifiers: artificial neural network (ANN), support vector machine (SVM), and random forest (RF).

Artificial Neural Network (ANN). We used a feed-forward ANN with 25 neurons with tangent sigmoid activation functions in one hidden layer to associate the extracted features with the subjects' identities. The ANN was trained using back-propagation which is a supervised learning method (Han and Kamber, 2006). During training the prediction of the network is compared to the known target value (subject's identity) and the weights are modified to minimize the mean square error. These errors propagate backwards from the output layer to the hidden layer (Han and Kamber, 2006). The network was trained using the scaled conjugate gradient algorithm described in (Møller, 1993). The weights and bias values of the neurons were updated using a gradient descent with momentum.

Support Vector Machine (SVM). We used a ν-SVM (Schölkopf et al., 2000) with a sigmoid kernel in its implementation in the LIBSVM[1] (Chang and Lin, 2011). SVMs are fundamentally a two-class classifier. Various methods have been proposed how to use SVMs for multi-class problems (Bishop, 2006, chap.

---

[1] LIBSVM: library for support vector machines.

Table 1: Characteristics of the 22 subjects (15 men, 7 women) who participated in the data collection.

Characteristic	Mean $\pm$ SD	Range
Age (yr)	26.6 ± 4.0	18-33
Height (cm)	179.8 ± 9.6	160-198
Weight (kg)	76.7 ± 11.1	58-108

7). We used a one-against-one method which constructs $n(n-1)/2$ classifiers where n is the number of classes to distinguish. Each classifier is trained on tuples from two classes. A voting strategy is then applied to determine the winning class (Hsu and Lin, 2002).

Random Forest (RF). A random forest is a classifier consisting of a combination of tree predictors. The growth of each tree is governed by independently and identically distributed random vectors (Breiman, 2001). Each tree votes for one class and the class which occurs most frequently is the output of the classifier. RF classifiers are fast in the training phase and the training time is linear to the number of trees used. The testing of an unknown tuple is performed on each tree independently and is therefore parallelisable. We used a RF consisting of 100 trees, with each tree being constructed of ten randomly chosen features.

3 SUBJECTS AND DATA COLLECTION

Twenty-two healthy subjects (15 men, 7 women) volunteered to participate in the study. The subjects were informed verbally and in writing in advance and signed an informed consent document. With a questionnaire we additionally obtained information on age, height and weight (see Table 1).

The data was collected using the treadmills in the gymnasium of our university. Velocities between 3 to 9 km/h were chosen to cover slow, normal, and fast walking as well as jogging. The treadmill was set to no incline and the velocity was manually increased by 2 km/h every two minutes. This procedure was repeated twice for each subject in order to collect enough data.

The accelerations of the upper body and the ECG were recorded with a self-made WBS (see Figure 1). The WBS measures accelerations within a range of ± 6 m/s^2 along three orthogonally oriented axes using a commercial off-the-shelf accelerometer (ST LIS3LV02DL). The ECG is digitized using the analog-to-digital converter of a TI MSP430 micro-controller. Body accelerations and ECG were measured with a 150 Hz sampling rate and a 12 bit resolution (range 0 to 4095). The measurements were sent wirelessly to a nearby receiver for recording.

The subjects were given an explanation as how to place the chest strap with the WBS tightly around the chest. However, we didn't verify the correct placement of the WBS to assure real world conditions. Furthermore, no instructions were given on how to perform the exercise.

4 RESULTS

This section describes the evaluation of the athlete identification on data collected from 22 subjects during walking and jogging on a treadmill.

Evaluation Methods. All features were calculated on windows of acceleration and ECG measurements of two seconds. No overlap of the windows was chosen to ensure fully discriminative training and testing data. We concatenated features of two consecutive windows to have samples of four seconds of data to identify the subject. Our dataset consists of 4048 samples (184 samples per subject). The identification performance was determined using a ten-fold cross-validation. For evaluation, we used three statistical measures: sensitivity, specificity and accuracy. In order to calculate the statistics we obtained the number of true positive samples TP_i, true negative samples TN_i, false positive samples FP_i, and false negative samples FN_i from the classifier's output. For a class i the sensitivity R_i is defined as:

$$R_i = \frac{TP_i}{TP_i + FN_i} * 100 \qquad (5)$$

The sensitivity (also referred to as recall) measures the percentage of correctly classified positive samples in relation to all positive samples. For negative samples the specificity S_i is defined as:

$$S_i = \frac{TN_i}{TN_i + FP_i} * 100 \qquad (6)$$

We calculated the overall sensitivity $\overline{R}$ and the overall specificity $\overline{S}$ as a class-based weighted average. For our multi-class problem we refer to the overall sensitivity as the accuracy of the classifier:

$$ACC = \overline{R} = \sum_{i=1}^{n} p_i R_i \qquad (7)$$

where n denotes the number of classes and p_i the probability of the occurrence of the class in the test

Table 2: Accuracy (*ACC*) and overall specificity ($\overline{S}$) results determined with different classifiers on a feature space combining acceleration and ECG features (combination C8, see Table 3). The highest performance was obtained with the RF classifier.

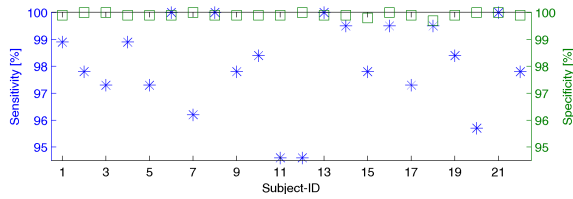| Classifier	*ACC*	$\overline{S}$
ANN	94.2 %	99.8 %
SVM	90.4 %	99.5 %
RF	98.1 %	99.9 %

Figure 9: Class-specific sensitivity (stars) and specificity (squares) results of the identification of the 22 subjects (RF classifier, feature combination *C*8). The sensitivity varied between 94.6 to 99.5 %. The specificity was over 99.7 % for all subjects.

data. In our test data the samples of the 22 classes are equally distributed ($p_i = 1/22, \forall i$). The overall specificity $\overline{S}$ is calculated accordingly. The optimum of the statistical measures is 100 %.

Results of the Athlete Identification. We determined the identification performance for three standard classifiers: ANN, SVM and RF. The classifiers and their parametrization are described in section 2.5. We achieved up to 98.1 % accuracy (see Table 2) with the RF classifier using a feature space combining acceleration and ECG based features. The lowest accuracy of 90.4 % was obtained with the SVM. For all three classifiers, we obtained an overall specificity $\overline{S}$ of more than 99 %.

The class-specific sensitivity (see Equation 5) of the identification varied between 94.6 to 99.5 % for the different subjects (RF classifier, see Figure 9). We observed only low deviations in the identification's specificity between the 22 subjects. A class-specific specificity (see Equation 6) of more than 99.7 % was achieved for all subjects.

We performed a feature selection using the ANN classifier to determine the impact of the different features and to identify combinations *C* with a high classification performance (see Table 3). We obtained a similar identification accuracy based on acceleration (86.6 %, *C*6) and ECG (84.8 %, *C*4) measurements. In combination, the accuracy improved to 94.2 % (*C*8).

The ECG contained more information on the subject in the frequency domain than the acceleration

Table 3: Identification accuracy of different feature combinations *C*. We obtained a similar accuracy with acceleration and ECG based features (see *C*4, *C*6). Combining both improved the accuracy (see *C*8). The feature selection was performed using the ANN classifier. * denotes the use of the average over the five correlation coefficients per subject.

C	Acceleration feat.			ECG feat.			*ACC*
	Time dom.	FFT coef.	FFT feat.	FFT coef.	FFT feat.	Corr. coef.	
*C*1	-	-	-	x	x	-	72.4 %
*C*2	-	-	-	-	-	x	80.3 %
*C*3	x	-	-	-	-	-	83.3 %
*C*4	-	-	-	x	x	x	84.8 %
*C*5	x	-	-	-	x	-	86.5 %
*C*6	x	x	x	-	-	-	86.6 %
*C*7	x	-	-	-	-	x	93.6 %
*C*8	x	x	x	x	x	x*	94.2 %

measurements (12.3 % higher accuracy). Frequencies of up to 10 Hz contained the most discriminant information of the acceleration measurements. A reduction of the frequency band from 40 to 10 Hz reduced the identification accuracy by only 3.8 %. For the ECG measurements, a reduction from 40 Hz to 15 Hz resulted in a 8.8 % lower accuracy. Overall, we obtained an accuracy of 72.4 % (*C*1) for features extracted from the ECG in the frequency domain.

We found that correlation coefficients describing the similarity between heartbeat segments provide useful insights to identify subjects (80.3 % accuracy, *C*2). To reduce the dimensionality of the feature space, we averaged the correlation coefficients corresponding to the five reference segments per subject. This averaging resulted in a 7.7 % lower accuracy. However, in combination with other features this difference was negligible (0.4 % for *C*8).

The time domain features calculated from the acceleration signals showed a good accuracy (83.3 %, *C*3). Additional information on the gait in the frequency domain improved the identification accuracy to 86.6 % (*C*6).

By combining the time domain features of the acceleration data with the correlation coefficients derived from the ECG, we achieved a high accuracy of 93.6 % (*C*7), which is only 0.6 % less than using the full feature set (*C*8).

For the time domain features extracted from the acceleration signals, we analysed the impact of the different acceleration axes on the subject's identification accuracy. The highest accuracy was obtained for the anteroposterior accelerations (a_{AP}). The mediolateral accelerations (a_{ML}) showed a 4.4 % and the vertical accelerations (a_V) a 16.2 % lower accuracy.

We additionally evaluated our approach using a hold-out validation for which the data set was split

Table 4: Identification accuracy (ACC) and equal error ratio (ERR) of the RF classifier obtained with a ten-fold cross-validation and a hold-out validation (percentage split: 66 % training, 34 % testing) on feature combination $C8$.

Validation	ACC	ERR
Ten-fold cross-validation	98.1 %	1.1 %
Hold-out validation	97.2 %	3.4 %

Table 5: Equal error ratio (ERR) of other gait based identification methods. N denotes the number of subjects who participated in the experiments.

	Velocities	N	ERR
Mäntyjärvi et. al (2005)	slow, normal and fast walking	36	7 %
Ailisto et. al (2005)	normal walking	36	6.4 %
Gafurov et. al (2006)	normal walking	21	5 %
Rong et. al (2007)	normal walking	21	5.6 %

in 66 % training data and 34 % testing data. A hold-out validation avoids temporal proximity between training and testing data and allows therefore a more accurate estimation of the generalization performance. We noted only a slight decrease in accuracy by 0.9 % for the RF classifier (see Table 4).

To estimate the impact of the number of subjects in the data set on the identification performance, we repeated our test with eleven out of the twenty-two subjects. For a group of eleven subjects which were randomly chosen, the overall accuracy could be improved to 99.3 % (RF classifier).

In order to compare our results with existing work, we additionally calculated the equal error rate (ERR) of the RF classifier on feature combination $C8$. The ERR is the rate at which both accept and reject errors are equal. For our dataset containing ECG and gait characteristics, we obtained an ERR of 1.1 % for the ten-fold cross-validation and 3.4 % for the hold-out validation. Compared to other approaches which are based on only gait characteristics our achieved ERR is lower (see Table 5). For a comparison of our approach with an ECG based identification we have chosen the method of Chan et. al (2008) because the results are also based on data from non-clinical ECGs. With 98.1 % our accuracy is higher than Chan et. al (2008) results (89 %). However, with an identification on ECG characteristics only, we obtained a lower accuracy (84.8 %, $C4$). Overall, our high performance is achieved by combining ECG and gait characteristics. We believe that motion related artefacts in the ECG, and a high variability in the gait patterns between changing from slow walking to jogging, reduce the identification performance when we use only one source of information.

5 DISCUSSION AND CONCLUSIONS

This paper is concerned with the identification of humans during walking and jogging using a single wireless body sensor module attached to a chest strap. Our approach focuses on recognising a human using a biometric measure based on the characteristics in the gait style and the ECG of the human and is hence independent of the used hardware. Thus, our system overcomes the drawbacks of an identification based on the WBS's serial number or an radio-frequency based identification (RFID) which recognises the hardware but not the subject itself.

We have collected data from 22 subjects on a treadmill at velocities from 3 to 9 km/h using a WBS attached to a chest strap. To assure real world conditions, no advice was given on how to perform the exercise and the correct placement of the chest strap was not verified. Despite severe variations in the gait patterns and motion-related artefacts in the ECG, which occur due to real world conditions and physical exercise, our method achieves up to 98 % accuracy.

We obtained a good identification accuracy for time domain features extracted from the acceleration signals. By using simple and low-dimensional features on the acceleration signal our method can potentially be implemented on computationally constrained platforms, such as a microcontroller on a WBS.

Our identification method can presumably not be extended to an unlimited number of subjects. The individual characteristics in the subject's ECG and gait patterns are extremely difficult to capture and may change over time because of an adaptation to physical exercise. However, we believe our method is well suited to provide an automatic annotation of sensor measurements from several WBSs with the subject's identity for use in sports medicine and athletic training research. Moreover, our method helps to customize a training session by loading personal settings of the recognized athlete on the WBS or other sport equipment.

Our future work includes the evaluation of the identification method within team sports. In particular, we want to recognize handball players in order to support a real-time vision-based tracking of these players.

ACKNOWLEDGEMENTS

REFERENCES

Afonso, V. X., Tompkins, W. J., Nguyen, T. Q., and Luo, S. (1999). ECG beat detection using filter banks. *Transactions on Biomedical Engineering*, 46(2):192–202.

Ailisto, H. J., Lindholm, M., Mantyjarvi, J., Vildjiounaite, E., and Makela, S. M. (2005). Identifying people from gait pattern with accelerometers. In *Society of Photo-Optical Instrumentation Engineers*, volume 5779, pages 7–14.

Bianchi, L., Angelini, D., and Lacquaniti, F. (1998). Individual characteristics of human walking mechanics. *Pflügers Archiv European Journal of Physiology*, 436:343–356.

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

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Chan, A. D. C., Hamdy, M. M., Badre, A., and Badee, V. (2008). Wavelet distance measure for person identification using electrocardiograms. *IEEE Transactions on Instrumentation and Measurement*, 57(2):248–253.

Chang, C.-C. and Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent System Technology*, 2:27:1–27:27.

Christ, P., Mielebacher, J., Haag, M., and Rückert, U. (2010). Detection of body movement and measurement of physiological stress with a mobile chest module in obesity prevention. In *Proc. of the 10th Australasian Conf. on Mathematics and Computers in Sport*, pages 67–74.

Christ, P., Werner, F., Rückert, U., and Mielebacher, J. (2011). An approach for determining linear velocities of athletes from acceleration measurements using a neural network. In *Proc. of the 6th IASTED Int. Conf. on Biomechanics*, pages 105–112. ACTA Press.

Conover, M. B. (2002). *Understanding electrocardiography*. Mosby.

Gafurov, D., Helkala, K., and Søndrol, T. (2006). Biometric gait authentication using accelerometer sensor. *Journal of Computers*, 1(7):51–59.

Green, L. S., Lux, R. L., Haws, C. W., Williams, R. R., Hunt, S. C., and Burgess, M. J. (1985). Effects of age, sex, and body habitus on QRS and ST-T potential maps of 1100 normal subjects. *Circulation*, 71(2):244–253.

Han, J. and Kamber, M. (2006). *Data mining: concepts and techniques*. The Morgan Kaufmann series in data management systems. Elsevier.

Hsu, C. W. and Lin, C. J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, 13(2):415–425.

Mäntyjärvi, J., Lindholm, M., Vildjiounaite, E., Mäkelä, S.-M., and Ailisto, H. A. (2005). Identifying users of portable devices from gait pattern with accelerometers. In *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, volume 2, pages ii/973–ii/976.

Møller, M. F. (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 6(4):525–533.

Nixon, M. S., Tan, T., and Chellappa, R. (2006). *Human identification based on gait*, volume 4. Springer-Verlag New York, Inc.

Rong, L., Jianzhong, Z., Ming, L., and Xiangfeng, H. (2007). A wearable acceleration sensor system for gait recognition. In *2nd IEEE Conf. on Industrial Electronics and Applications*, pages 2654–2659.

Schölkopf, B., Smola, A. J., Williamson, R. C., and Bartlett, P. L. (2000). New support vector algorithms. *Neural computation*, 12(5):1207–1245.

Shen, T. W., Tompkins, W. J., and Hu, Y. H. (2002). One-lead ECG for identity verification. In *Proc. of the 2nd IEEE Int. Joint Conf. on Engineering in Medicine and Biology Society*, volume 1, pages 62–63. IEEE.

Simon, B. P. and Eswaran, C. (1997). An ECG classifier designed using modified decision based neural networks. *Computers and Biomedical Research*, 30(4):257–272.

Vidaurre, C., Sander, T. H., and Schlögl, A. (2011). BioSig: The free and open source software library for biomedical signal processing. *Computational Intelligence and Neuroscience*, 2011:12.

Weyand, P. G., Sternlight, D. B., Bellizzi, M. J., and Wright, S. (2000). Faster top running speeds are achieved with greater ground forces not more rapid leg movements. *Journal of Applied Physiology*, 89(5):1991–1999.