11

Biometric Applications of One-Dimensional Physiological Signals – Electrocardiograms

Jianchu Yao[1], Yongbo Wan[2] and Steve Warren[3]
[1]*East Carolina University, Greenville, NC,*
[2]*Oklahoma State University, Stillwater, OK,*
[3]*Kansas State University, Manhattan, KS,*
USA

1. Introduction

In the last decade, affordable computing power has become available on platforms intended for low-power, mobile applications. For example, Gumstix Overo products integrate WiFi/Bluetooth connectivity, microSD storage, and 600 MHz Texas Instruments OMAP (Open Multimedia Application Platform) 35xx processors with up to 256 MB of flash memory/SDRAM, offering laptop-like resources and performance in a form factor of a stick of gum (Chaoui, et al., 2001). This speaks to the potential for wearable, wireless medical devices based on such products to process signals on-board; functionality that previously required expensive, bulky, benchtop/bedside equipment. These computational capabilities also show promise for smart devices that implement context-based intelligence and on-device expert systems for clinical decision making. This move toward highly-capable and intelligent mobile medical devices drives the need for verification tools, data integrity checkers, and role-based security mechanisms that can also be implemented at the embedded level, since the potential usage scenarios and associated protection needs are numerous in comparison with legacy medical device applications in controlled hospital and home care settings. For example, many implementations of personal and body area networks have been developed to facilitate ambulatory monitoring of health status, where physiological parameters such as heart rate, heart activity, blood oxygen saturation, and respiration rate can be gathered through the use of mobile, wearable electrocardiographs and pulse oximeters (Jovanov, et al., 2009; Galeottei, et al., 2008; Chuo, et al., 2010). These data need to be authenticated and checked for integrity before they are stored in electronic patient records, which implies the need for "owner-aware" devices that verify user identity as part of the data acquisition process (Warren, et al., 2005; Warren & Jovanov, 2006). Many biometric authentication protocols are computationally intensive and can well-utilize the emerging computational capabilities of low-power mobile devices.

A broad range of biomedical data, from physiological signals/images to behavioral traits, has been explored for its biometric authentication and identification potential (Biel, et al., 2001; Chan, et al., 2008; Doi & Yamanaka, 2004; Duc, et al., 1997; Elsherief, et al., 2006; Faundez-Zanuy, 2005; Irvine, et al., 2001; Israel, et al., 2005; Shen, et al., 2002; Sullivan, et al., 2007; Yao & Wan, 2010; G. H. Zhang, et al., 2009). Image-based mechanisms that assess

fingerprints (Doi & Yamanaka, 2004; Matsumoto, et al., 2002), retinal patterns (Elsherief, et al., 2006), facial features (Daugman, 1998; Koh, et al., 1999; Philips, et al., 2003), and vein/palm structures (Doi & Yamanaka, 2004) are the leading biometric modalities used today for identity verification. Recently, one-dimensional physiological signals such as electrocardiograms (ECGs) (Agrafioti & Hatzinakos, 2008a, 2008b; Biel, et al., 2001; Israel, et al., 2005; Kyoso & Uchiyama, 2001; Micheli-Tzanakou, et al., 2009; Nasri, et al., 2009; Plataniotis, et al., 2006; Saechia, et al., 2005; Shen, et al., 2002; Singh & Gupta, 2008; Yao & Wan, 2008, 2010), photoplethysmograms (PPGs) (Ludeman & Chacon, 1995; Ludeman & Chacon, 1996; Love, et al., 1997; Ma et al., 2006; Bao et al., 2005; Gu, et al., 2003; Gu & Zhang, 2003; Wan, et al., 2007; Yao, et al., 2007) and electroencephalograms (EEGs) (Marcel & Millan, 2007) have garnered attention as promising biometric candidates based on the following thoughts:

- In many cases, these signals are already acquired and stored as part of the healthcare delivery process. It is therefore sensible to utilize them as identification attributes because no additional user action or data gathering is required, as is the case with most other biometric modalities. In other words, authentication, identification, or verification can occur behind the scenes even without subject awareness (Warren & Jovanov, 2006; Warren, et al., 2005). Further, no additional hardware would be required to implement this feature, which implies that biometric features can be added to the system without incurring significant additional device cost. The efficiency of this approach in terms of care delivery workflow, coupled with the ease of use and economic sensibility of such tools, should lead to increased technology acceptance by patients and providers.
- Since they represent an individual's underlying physiological status, these signals may be less sensitive to environmental factors that affect other biometric parameters, thereby avoiding substantial deteriorations in biometric performance when fully controlled environments are unobtainable. For example, environmental noise unavoidably interferes with voice-based biometric systems (Ming, et al., 2007). In such cases, costly environmental improvements are often required to ensure that identification systems work properly.
- The use of physiological signals for identification or verification can help to prevent failure to enroll (FTE) issues that may occur when a subject does not possess a particular biometric. Most current biometric approaches are affected by this. In fingerprint identification, for example, it has been estimated that fingerprints from up to 4% of the population cannot be used for identification purposes due to the poor quality of the fingerprint ridges (Jain, et al., 2004).
- Some current biometric approaches are subject to forgery. Artificial fingerprints, for example, can be constructed to circumvent a fingerprint verification system. Unlike most current biometric data, which are extracted from "surficial" parts of the human body, physiological signals represent core internal behavior and are hard to emulate with tissue phantoms (Jain, et al., 2004). This "innerness" makes the identification process less prone to forgery, preventing imposters from disguising their true identity by changing these metric patterns.

A number of research efforts have attempted to address these thoughts within the context of one-dimensional biomedical signals. This chapter provides an up-to-date review of this research, summarized from the following four perspectives: (1) the signals used for identification, with an emphasis on ECGs, (2) signal processing methods, (3) classification

methods, and (4) identification algorithm performance. As noted in (Matyas & Riha, 2003), "the main issue in biometric authentication systems is performance." High identification accuracy is critical for practical biometric technologies.

This chapter also presents some of the authors' research findings geared toward quantitatively evaluating the identification accuracy of ECG data as population size increases. It presents the design of a wearable electrocardiograph and the associated signal processing algorithms, followed by an assessment of identification-algorithm performance for this application. In the analysis, three distance measures were defined in the wavelet domain: the Distance of Discrete Wavelet Coefficients (DDWC), the Distance of Continuous Wavelet Coefficients (DCWC), and the Ratio of the volume of Intersection to the volume of Union (RItU). Evaluation results for all three distance measures demonstrated consistent declination as the population grew to a size of 50. Possible causes for this performance drop are discussed. These experiments also recognized that distinguishable information from these signals may not be as prevalent as the unique data acquired using more popular modalities. In other words, the number of possible combinations for the patterns of the statistical attributes that can be extracted from these signals is limited. Based on these findings, the chapter suggests scenarios that ECGs can be utilized as the sole modality for biometric purposes or those they can serve as a supplemental tool to other modalities.

2. ECG as a biometric modality: a systematic review

Researchers have recognized for some time that ECGs contain innate human attributes (i.e., they reflect the electro-myocytic properties of the heart), so it seems sensible that each individual's ECG may demonstrate his or her uniqueness and therefore be useful for identity verification. This section provides a detailed review of research geared toward the usefulness of ECGs as biometric indicators.

Since the first such effort was reported in 1999 (Biel, et al., 1999), approximately 20 different groups have researched this interesting area (Israel, et al., 2005; Chan, et al., 2008; Biel, et al., 2001; Yao & Wan, 2008; Zhang & Wei, 2006; Biel, et al., 1999; Boumbarov, et al., 2009; Chan, et al., 2006; Chiu, et al., 2008; Fatemian & Hatzinakos, 2009; Gahi, et al., 2008; Israel, et al., 2003; Kim, et al., 2005; Kyoso, 2003; Kyoso & Uchiyama, 2001; Micheli-Tzanakou, et al., 2009; Nasri, et al., 2009; Plataniotis, et al., 2006; Shen, et al., 2002; Sufi & Khalil, 2008; Wang, et al., 2006; Yao & Wan, 2010) – see Table 1. Most of these published works use a similar approach to present research findings. They first start with a brief description of the physiological origin of the ECG and its characteristics (e.g., P wave, QRS complex, and T wave) (Webster, 1998). They then present the three primary steps in a typical classification process—feature selection, pre-processing, and classification (which usually includes enrollment/training and identification/testing). Finally, they draw optimistic conclusions from the classification results. Despite this consistency in presentation format, the projects themselves differ with regard to (a) ECG data sources, (b) data collection processes, (c) classification feature selection, and (d) classification methods adopted to realize the final identification results. Each is detailed below:

- **ECG Data Sources:** Multiple groups experimentally acquired ECG data from volunteers (Biel, et al., 1999 2001; Chan, et al., 2006; Chan, et al., 2008; Gahi, et al., 2008; Israel, et al., 2005; Israel, et al., 2003; Kim, et al., 2005; Kyoso, 2003; Kyoso & Uchiyama, 2001; Sufi & Khalil, 2008; Wan & Yao, 2008; Wang, et al., 2006; Yao & Wan, 2008, 2010).

In these cases, experimental conditions were often well controlled: subjects were requested to rest for a period of time prior to data collection (Biel, et al., 2001; Chan, et al., 2008; Yao & Wan, 2008, 2010), with the exception of studies that examined the performance of the identification methods under conditions where heart rate was increased (Kim, et al., 2005). Some of these groups developed their own devices (Kyoso, 2003; Kyoso & Uchiyama, 2001; Wan, et al., 2007; Yao & Wan, 2008), while others collected data with off-the-shelf ECG products (Sufi & Khalil, 2008; Chan, et al., 2008). Other non-experimental data sources were employed, as in (Agrafioti & Hatzinakos, 2008a; Plataniotis, et al., 2006; Singh & Gupta, 2008; Agrafioti & Hatzinakos, 2008b), where ECG data were extracted from existing databases (e.g., PTB ("The PTB Diagnostic ECG Database") and MIT-BIH ("MIT-BIH Database Distribution"); both of these databases are available through the Internet for public research use.

- **Data Collection:** Some publications provide subject demographic data, including gender (Biel, et al., 2001; Kim, et al., 2005; Yao & Wan, 2008), age range (Biel, et al., 2001; Yao & Wan, 2008; Chan, et al., 2008; Kim, et al., 2005; Yao & Wan, 2010), and heart condition (Agrafioti & Hatzinakos, 2008a; Chiu, et al., 2008; Kim, et al., 2005; Plataniotis, et al., 2006; Singh & Gupta, 2008; Sufi & Khalil, 2008). However, few report complete demographic or health-condition information for participants. Furthermore, the time interval between subject enrollment and data collection, a critical element when determining the effectiveness of a biometric modality, is frequently overlooked (Chiu, et al., 2008; Gahi, et al., 2008; Israel, et al., 2005; Kim, et al., 2005; Kyoso, 2003; Plataniotis, et al., 2006; Saechia, et al., 2005; Sufi & Khalil, 2008). Even studies that record this information often mention it vaguely (Chan, et al., 2008; Fatemian & Hatzinakos, 2009; Singh & Gupta, 2008) (see Table 1).
- **Selection of Classification Features:** Most investigators assess time domain features (e.g., time intervals between P, Q, R, S, and T waves, along with their amplitudes) (Biel, et al., 2001; Boumbarov, et al., 2009; Gahi, et al., 2008; Israel, et al., 2005; Kyoso, 2003; Z. Zhang & Wei, 2006) and angle information (Singh & Gupta, 2008). Others believe that post-transform features are more distinctive and will therefore improve identification performance. For example, wavelet transformation was used in (Chan, et al., 2006; Chan, et al., 2008; Chiu, et al., 2008; Yao & Wan, 2008, 2010) to find the wavelet coefficients and distances in the wavelet domain that optimally quantify the similarity between two ECGs. Autocorrelation coefficients are a third type of statistical feature under investigation (Agrafioti & Hatzinakos, 2008a; Plataniotis, et al., 2006). In addition to these three types of analytical information, the appearance of the ECG waveforms was added as a classification feature in (Wang, et al., 2006). Finally, after recognizing the difficulties encountered when delineating ECG cycles, some investigators extracted classification features without the need to detect fiducial points (Plataniotis, et al., 2006; Agrafioti & Hatzinakos, 2008a), where the DCT (Discrete Cosine Transform) approach did not rely on the accurate location of each ECG cycle.
- **Classification Algorithms:** As in other pattern recognition domains, numerous classification algorithms have been created for human identification based on ECGs, where algorithm performance varies widely. While most of these approaches used variations of a "distance" concept (e.g., Euclidean distance (Israel, et al., 2003; Plataniotis, et al., 2006) or Mahalanobis' distance (Kyoso, 2003; Kim, et al., 2005) to quantify the similarities between the unknown data and the waveforms enrolled in the

database, the classification algorithms they adopted were different, including (a) classic linear discriminate analysis (LDA) (Agrafioti & Hatzinakos, 2008a; Chan, et al., 2008; Kim, et al., 2005; Kyoso, 2003; Wang, et al., 2006), (b) neural networks (Saechia, et al., 2005; Shen, et al., 2002; Wan & Yao, 2008; Boumbarov, et al., 2009), and/or (c) voting after initial results were available from the first classification level (Israel, et al., 2005; Agrafioti & Hatzinakos, 2008b). Rather than use the intuitive distance concept, other investigators employed a sequential approach that employs a hidden Markov model (Boumbarov, et al., 2009) and a probabilistic, Bayesian-theorem-based approach (Z. Zhang & Wei, 2006). Both approaches obtained results comparable to distance-based methods.

Group	Year	Subjects	Time Span	Success Rate
Biel	2001	20	6 weeks	90-100%
Kyoso	2003	9	N/A	Wide range
Shen	2002	20	N/A	80-95%
Israel	2005	29	N/A	100%
Kim	2005	10	N/A	N/A
Saechia	2005	N/A	N/A	97%
Zhang	2006	502 records	Same datasets	82-97%
Wang	2006	13	A few years	84.6%
Plataniotis	2006	14	N/A	92.8-100%
Chan	2008	50	> 1 day	95%
Yao	2008	20	Hours to weeks	91.5%
Agrafioti	2008	27	Mixed length	96%-100%
Agrafioti	2008	14	A few years	85.6%-100%
Sufi	2008	15	N/A	93-95%
Gahi	2008	16	N/A	100%
Singh	2008	25	Same time or unclear	98.5-99%
Chiu	2008	45	N/A	100% and 81%

Table 1. Summary of research on ECG analysis as a biometric modality

- **Other Endeavors:** Unlike most of the aforementioned research, which sought better identification rates, other investigators wished to improve computational efficiency. They tried to reduce the number of necessary features by (1) selecting the most meaningful features after observing how each feature changed the classification results (Biel, et al., 2001; Agrafioti & Hatzinakos, 2008b) or (2) using methods such as principle component analysis (PCA) (Yao & Wan, 2008; Sufi & Khalil, 2008; Z. Zhang & Wei, 2006; Yao & Wan, 2010).

3. ECG as a bioidentification modality: performance evaluation

This section presents the authors' recent research on the performance and limitations of ECGs as biometric indicators, as well as other potential application fields.

3.1 Data acquisition

ECG data for this study were collected with an "in-house" device (see Fig. 1), and MATLAB scripts processed these data using wavelet-based approaches. Data collection and pre-processing details were described in (Wan & Yao, 2008; Yao & Wan, 2010). Thirty participants (26 males and 4 females) with ages ranging from 18 to 51 years were recruited for data collection. A total of 121 datasets were collected from these subjects, where each subject participated in multiple ($2 \le N_i \le 5$) data collection sessions; consecutive sessions were a few weeks apart.

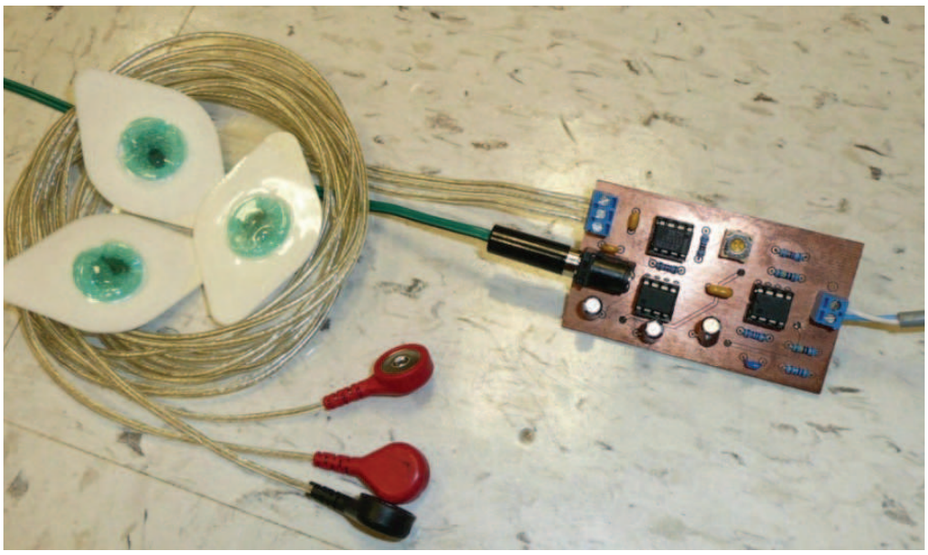


Fig. 1. An "in-house" ECG module for bioidentification data acquisition

3.2 Signal preparation and feature extraction

As specified in (Wan & Yao, 2008; Yao & Wan, 2010), the raw ECG signals were pre-processed to remove signal noise, detect R waves, and normalize each signal to a pre-defined length and amplitude range. Specifically, two major noise sources (low frequency signal drifts at around 0.06 Hz and higher frequency signal spikes at 60 Hz) were first filtered with "hard thresholding" after a scale 12 Daubechies's db6 wavelet transform was applied to all of the heart beat cycles. Detailed wavelet parts at scales 2, 3, and 4 were reconstructed so that the R peaks could be located as the fiducial points to identify ECG cycles. Identified ECG cycles were interpolated to a pre-defined length for the convenience of future steps. Sixty consistent heartbeat cycles from each of the datasets were selected and their amplitudes were normalized to the range of [-1, 1]. In this step, data consistency was examined by calculating the Euclidean distance between the mean of each cycle and the mean of all cycles.

A wavelet transform (Yao & Wan, 2008), similar to (Chan, et al., 2008; Chiu, et al., 2008), was applied to each processed time-domain signal, and then wavelet coefficients were calculated

for each cardiac cycle within that signal. Depending on the continuous or discrete wavelet transform applied, the number of coefficients varied, as discussed earlier when the specific measures were introduced. Six sets of wavelet coefficients (corresponding to sixty heart beats) were saved for each of the 121 ECG datasets. From this point on, the wavelet coefficients served as the "statistical features" and were manipulated for subsequent classification decisions.

Out of the N_i coefficient sets obtained from each subject, one coefficient set (corresponding to one heart interval) was enrolled in the database, creating a database of 30 coefficient sets. The other N_i-*1* coefficient sets (corresponding to N_i-*1* heart intervals of the same subject) were used for classification tests: 121-30 = 91 coefficient sets.

3.3 Measures of signal similarity/difference

The goal of this exercise was to explore identification-algorithm performance as a function of test population size. Three distance measures were utilized to represent the level of similarity between the unknown wavelet coefficient set and the enrolled coefficient sets: (1) Distance of Discrete Wavelet Coefficients (DDWC), (2) Distance of Continuous Wavelet Coefficients (DCWC), and (3) Ratio of Intersection to Union of continuous wavelet coefficients (RItU). The following paragraphs describe the three distance measures in detail.

3.3.1 Distance of Discrete Wavelet Coefficients (DDWC)

The wavelet distance proposed in (Chan, et al., 2008) was examined first (it was referred as WDIST in (Chan, et al., 2008)). This distance is notated here as DDWC to distinguish it from the distance obtained from a continuous wavelet transform. In this case, coefficients from a discrete wavelet transform were utilized for distance measure calculations. The DDWC is defined by

$$DDWC_n = \sum_{p=1}^{P}\sum_{q=1}^{Q} \frac{\left| c_0^{p,q} - c_n^{p,q} \right|}{max(\left| c_0^{p,q} \right|, T.H.)} \tag{1}$$

where $c_0^{p,q}$ is the q^{th} wavelet coefficient at the p^{th} scale of the unknown coefficient set; $c_n^{p,q}$ is the q^{th} wavelet coefficient at the p^{th} scale of the enrolled coefficient set; P is the number of scales of the wavelet transform; and Q is the number of coefficients at a specific scale. *T.H.* is a pre-selected normalization constant. To obtain the DDWC measure, a scale 6, Bior1.1 wavelet transform (Mallat, 1999) was applied to the pre-processed ECGs, yielding coefficient structures of 256 elements. The 'Bior1.1' basis function belongs to the Biorthogonal Wavelet Pairs wavelet family. The orthogonal discrete wavelet transform functions have excellent localization properties in both the time and frequency domains (Kharate, et al., 2007), and the coefficients obtained contain distinctive information. Note that the basis function chosen here is different from that in (Chan, et al., 2008), which used a db3 function.

3.3.2 Distances of Continuous Wavelet Coefficients (DCWC)

The discrete wavelet transform is usually implemented as a dyadic-orthogonal transform where a signal can be presented as a combination of elements in the orthogonal basis set without information redundancy. The continuous wavelet transform decomposes time-domain signals into temporal-spectral components with continuous scale factors and

translation parameters. The coefficients obtained from a continuous wavelet transform depict the detailed, smooth transitions of the signal energy distribution along the time and frequency dimensions (see Fig. 2).

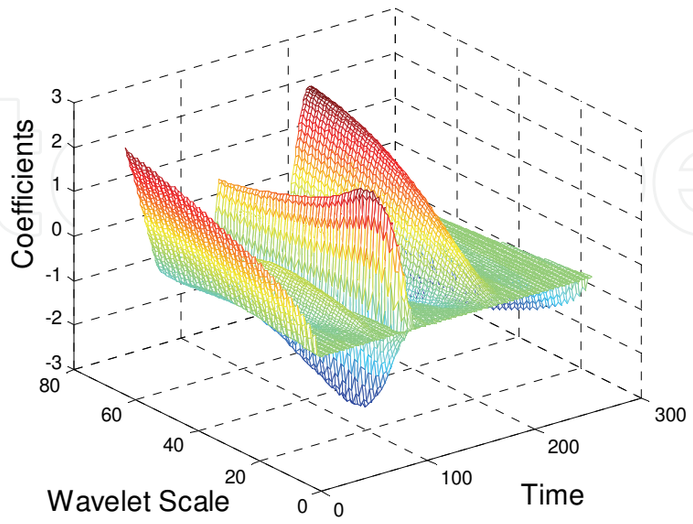


Fig. 2. A 3-D coefficient surface obtained from a continuous wavelet transform

Assuming that the inclusion of the smooth transition of coefficients at different scales will yield better identification results, a distance of continuous wavelet coefficients (DCWC) is defined by

$$DCWC_n = \frac{\sum_{p=1}^{P}\sum_{q=1}^{Q}\left|c_0^{p,q} - c_n^{p,q}\right|}{max\left(ABS(C_0)\right)} \tag{2}$$

where $c_0^{p,q}$ is the q^{th} wavelet coefficient at the p^{th} scale of the unknown coefficient set; $c_n^{p,q}$ is the q^{th} wavelet coefficient at the p^{th} scale of the enrolled coefficient set; $P = 64$ is the number of scales of the wavelet transform; and $Q = 256$ is the number of coefficients at a given scale. In this experiment, the continuous wavelet transform used the same basis function, Bior1.1, as was used in the discrete wavelet transform. Note that the denominator in Eq. (2) contains the maximum of the absolute value of the coefficients of the unknown subject. Experiments showed that this normalization could obtain better classification results than using the denominator in Eq. (1) and avoided the process of finding the threshold.

3.3.3 Ratio of Volume of Intersection to Volume of Union (RItU)

The waveform coefficients, when plotted as a mesh, form a 3-dimensional spatial surface as shown in Fig. 2. A more intuitive way to quantify the similarity of two signals is the ratio of the volume under the intersection of the two signals to the volume under the union of the two signals (see Fig. 3 for a graphical depiction of the intersection and union of two 2-D curves). The more two compared signals differ, the smaller the ratio. When the two signals are identical, the ratio is 1, and when the two signals do not overlap, the ratio is 0, implying that they are separate from each other and that the similarity between them is minimal. In addition to taking into account the distance between two coefficient sets, as in the other two

measures, the RItU measure also considers the coefficient locations over the temporal and frequency dimensions. This volume ratio is mathematically defined as

$$RItU = \cap\left(C^T, C^E\right) / \cup(C^T, C^E) \tag{3}$$

where C^T is the coefficient set to be tested, and C^E is one of the coefficient sets enrolled in the database. The intersection and union of the two coefficient sets are further defined by

$$\cap\left(C^T, C^E\right) = \sum_{p=1}^{P}\sum_{q=1}^{Q} c_i,$$

where
$$c_i = \begin{cases} min(ABS(c_0^{p,q}), ABS(c_n^{p,q})), & SIGN(c_0^{p,q}) = SIGN(c_n^{p,q}) \\ 0, & SIGN(c_0^{p,q}) \neq SIGN(c_n^{p,q}) \end{cases} \tag{4}$$

and

$$\cup\left(C^T, C^E\right) = \sum_{p=1}^{P}\sum_{q=1}^{Q} c_u, \ where\, c_u = \max(ABS(c_0^{p,q}), ABS(c_n^{p,q})) \tag{5}$$

where the notation follows from Eq. (2) because the spatial surfaces used to calculate RItU were also determined from continuous wavelet transforms.

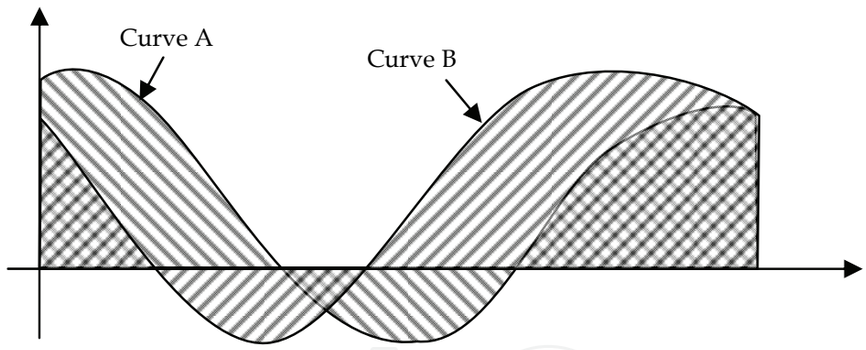


Fig. 3. The intersection (gridded area) and union (all shaded areas) of two curves

3.4 Evaluation of identification performance changes as population size increases

The classification method used in this experiment finds the distances (D) (as defined in the previous section) from the to-be-tested coefficient set $C_j^T \ (1 \le j \le 91)$ to the coefficient sets $C_k^E \ (1 \le k \le 30)$ enrolled in the database and uses these distances as the quantitative measure of signal difference/similarity. After all of the distances are compared, C_j^T is classified to the closest enrolled subject S_i. i.e., the unknown coefficient set:

$$C_j^T \to S_i \text{, where } i = arg\ i\ min D_{ij} \tag{6}$$

To evaluate the deterioration in accuracy as the test population size increases, a varied number (5, 10, 15, 20, 25, and 30) of subject waveforms were tested with the wavelet distance approach using the three difference/similarity measures introduced above. Coefficient sets

stored for testing purposes were randomly selected to perform identification tests. When a certain number (again, 5, 10, 15, 20, 25, or 30) of subject waveforms were tested, the total number of coefficient sets selected for testing could vary since the number of coefficient sets N_i for subject t could be different. The identification accuracy rate (AR) is defined as

$$AR = D_S / D_T \tag{7}$$

where D_S is the number of coefficient sets that have been successfully identified and D_T is the total number of coefficient sets selected for testing.

Repeated random sub-sampling was implemented to eliminate possible classification biases. A total of 20 trials with randomly selected unknown datasets were conducted for each case with a specific subject number (5, 10, 15, 20, and 25); only one test was conducted for the 30-subject case since all of the subjects were examined. In each trial, wavelet coefficient sets were selected randomly from those set aside for testing and classified according to the three measures. The average accuracy and standard deviation for all trials, using the three difference/similarity measures, was examined to analyze the biometric performance trend.

4. Experimental results

Fig. 4 illustrates identification performances when the three distance definitions, DDWC, DCWC, and RItU are utilized to measure subject similarity/difference. Comparing the three approaches, it is obvious that DDWC outperforms the other two distance measures. The latter two methods (DCWC and RItU) generate similar results, where the accuracy rate from the DCWC method is slightly higher than the RItU method. More importantly, these plots demonstrate that the classification accuracies for all three measures decline consistently by 12% as the number of test subjects increases from 5 to 30. Note also that, as the number of subjects grows, the standard deviation of the accuracy rate decreases (e.g., the DDWC method yields standard deviations of 6.6 and 2.3 for 5 and 25 subjects, respectively). This is true because more repeated datasets (and therefore a larger percentage of subjects) existed when larger numbers of subjects were incorporated.

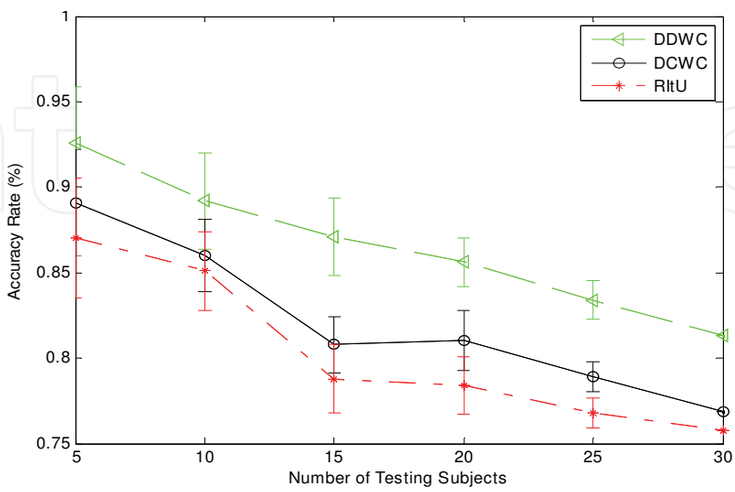


Fig. 4. Identification accuracy rate with the three difference/similarity measures versus the number of test subjects

5. Discussion

Possible causes for this performance drop can be identified. First, since signals such as ECGs are collected to provide patient health status and diagnose suspected illness, the signals are expected to show variations over time. Even for individuals whose health status does not change significantly over a period of time, normal circadian rhythms (on both a cycle-to-cycle basis and over a longer time interval) coupled with changes in stress level, emotions, and activity will in aggregate create variations that a robust identification algorithm must be able to tolerate. For some physiological signals, the detection environment may be another critical factor. PPGs, for example, which are based on light intensity either transmitted through or reflected by tissue, are extremely sensitive to motion artifacts, in spite of multiple existing approaches to help remediate these artifacts. ECGs are also sensitive to motion artifact and can be easily corrupted by electromagnetic interference that exists in most mobile patient environments. Without compensation, such variations and artifacts ultimately make one-dimensional signals less than ideal for identification or verification. More consistent attributes uniquely associated with patients are then desired.

These experiments also recognized that distinguishable information from these signals may not be as rich as the unique data acquired using popularly adopted modalities. In other words, the number of possible combinations for the patterns of the statistical attributes that can be extracted from these signals is limited. As the number of subjects increases beyond a certain number (20 to 30 in this case), the likelihood of having subjects whose signals are very similar increases significantly.

Therefore, despite the advantages that one-dimensional physiological signals may hold with respect to biometric identity assessment, performance assessments from previous research remind one that caution is required when such data are utilized for identification, especially when the subject population is large. The authors believe that these signals hold clear potential for this purpose, with the following qualifiers:

- The class of one-dimensional signals discussed here should be used with caution as a sole source of blind identification, primarily due to the less than desired uniqueness of the signal shapes and their time-dependent variations. However, when these signals are used as supplemental traits combined with other biometrics (e.g., in a data fusion approach), they are desirable due to the natural physical coupling between these various signal modalities, which is expected to improve the overall performance of the affiliated identification algorithms.
- While these nontraditional biometric modalities may not offer sufficient identification accuracy as required for legitimate authentication (i.e., where the goal is to identify an unknown subject given a large number of existing datasets), they may be better suited for individual verification, where the newly gathered signal is only compared to a recent set of data, with the assumption that the subject's identity is already inferred. Current verification processes (e.g., the two-stage process that requires something you have and something you know) usually seek information such as a password/passphrase, date of birth, home address, mother's maiden name, etc. A verification approach with one-dimensional signals, such as the ones proposed here, circumvents this process by employing non-transferrable datasets already native to the user.
- As the demand for long term state-of-health monitoring increases, medical sensors implemented on personal, wearable, or implanted platforms demand strict rules of

engagement to improve system interconnectivity and reliability so that they can be seamlessly woven into the user environment without requiring additional user intervention. Owner-aware sensors, or devices that recognize their owners based on an assessment of the data sets acquired from those individuals, are an appealing idea because they bolster security in the environment, minimizing their impact on normal human behavior, and increase the viability of the monitoring, diagnosis, and treatment process.

- Although the performance of these identification algorithms requires improvements for large populations, some of these one-dimensional signals do offer fairly accurate classifications when the subject population is relatively small. This points to their feasibility for environments such as homecare settings or community health centers, both of which are vital to an aging population. In these applications, health data could be constantly or periodically collected, so identification performance deteriorations caused by long-term signal alterations are expected to be minimized. Indeed, these signal alterations may themselves provide trend data as an additional means to distinguish individuals.

6. Conclusion

This chapter recognizes several important questions that arise upon completion of a comprehensive review of existing research work that explores the possibility of using ECGs as waveforms for human identification. It answers one of these questions by investigating how identification performance changes as a function of subject population size. Using three wavelet coefficient-based distances to measure the similarity/difference between unknown datasets and those in a database, consistent performance trends were obtained from the three discrimination cases, confirming that accuracy declines as the population grows. This finding is a reminder that, although ECG-based authentication holds potential for applications where ECG data have already been acquired or stored, caution is needed when the population size is large.

7. References

Agrafioti, F., & Hatzinakos, D. (2008a). ECG Based Recognition Using Second Order Statistics. *The 3rd International Symposium on Communications, Control and Signal Processing*, 2008.

Agrafioti, F., & Hatzinakos, D. (2008b). Fusion of ECG sources for human identification. *The 3rd International Symposium on Communications, Control and Signal Processing*, 2008.

Bao, B., Zhang, Y., & Shen, L. (2005). Physiological Signal Based Entity Authentication for Body Area Sensor Networks and Mobile Healthcare Systems. *Proceedings of the 27th Annual IEEE Engineering in Medicine and Biology Conference*, 2455-2458.

Biel, L., Pettersson, O., Philipson, L., & Wide, P. (1999). ECG analysis: a new approach in human identification. *Proceedings of the 16th IEEE Instrumentation and Measurement Technology Conference*, 1999.

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

Boumbarov, O., Velchev, Y., & Sokolov, S. (2009). ECG personal identification in subspaces using radial basis neural networks. *IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, 2009.

Chan, A. D. C., Hamdy, M. M., Badre, A., & Badee, V. (2006). Person Identification using Electrocardiograms. *Canadian Conference on Electrical and Computer Engineering*, 2006.

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

Chaoui, J., Cyr, K., Gregorio, S. d., Giacalone, J.-P., Webb, J., & Masse, Y. (2001). Open multimedia application platform: enabling multimedia applications in third generation wireless terminals through a combined RISC/DSP architecture. *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '01)*.

Chiu, C.-C., Chuang, C.-M., & Hsu, C.-Y. (2008). A Novel Personal Identity Verification Approach Using a Discrete Wavelet Transform of the ECG Signal. *International Conference on Multimedia and Ubiquitous Engineering*, 2008.

Chuo, Y., Marzencki, M., Hung, B., Jaggernauth, C., Tavakolian, K., Lin, P., et al. (2010). Mechanically Flexible Wireless Multisensor Platform for Human Physical Activity and Vitals Monitoring. *IEEE Transactions on Biomedical Circuits and Systems*, 4(5), 281.

Daugman, J. (1998). Phenotypic Versus Genotypic Approaches to Face Recognition *Face Recognition: From Theory to Applications* (pp. 108–123): Springer-Verlag.

Doi, J., & Yamanaka, M. (2004). Biometric authentication using finger and palmar creases. *2004 IEEE Symposium on Virtual Environments, Human-Computer Interfaces and Measurement System*.

Duc, B., Bigun, E. S., Bigun, J., Maitre, G., & Fischer, S. (1997). Fusion of audio and video information for multi-modal person authentication. *Pattern Recognition Letters*, 18(9), 835–843.

Elsherief, S. M., Allam, M. E., & Fakhr, M. W. (2006). Biometric Personal Identification Based on Iris Recognition. *2006 International Conference on Computer Engineering and Systems*.

Fatemian, S. Z., & Hatzinakos, D. (2009). A new ECG feature extractor for biometric recognition. *16th International Conference on Digital Signal Processing*, 2009.

Faundez-Zanuy, M. (2005). Biometric verification of humans by means of hand geometry. *39th Annual 2005 International Carnahan Conference on Security Technology*.

Gahi, Y. L., M.; Zoglat, A., Guennoun, M., Kapralos, B., & El-Khatib, K. (2008). Biometric Identification System Based on Electrocardiogram Data. *The New Technologies, Mobility and Security*, 2008.

Galeottei, L., Paoletti, M., & Marchesi, C. (2008). Development of a low cost wearable prototype for long-term vital signs monitoring based on embedded integrated wireless module. *Computers in Cardiology*, 905 - 908.

Gu, Y., Zhang, Y., & Zhang, Y. T. (2003). A Novel Biometric Approach in Human Verification by Photoplethysmographic Signals. *4th Annual IEEE Conference on Information Technology Applications in Biomedicine*, UK.

Gu, Y., & Zhang, Y. T. (2003). Photoplethysmographic authentication through fuzzy logic. *The IEEE EMBS Asian-Pacific Conference on Biomedical Engineering*.

Irvine, J., Wiederhold, B. K., Gavshon, L. W., Israel, S., McGehee, S. B., Meyer, R., et al. (2001). Heart rate variability: A new biometric for human identification. *Proceedings of International Conference on Artificial Intelligence*, 1106–1111.

Israel, S., Irvine, J. M., Cheng, A., D.Wiederhold, M., & K.Wiederhold, B. (2005). ECG to identify individuals. *Pattern Recognition Society*, 38, 133-142.

Israel, S., Scruggs, W. T., Worek, W. J., & Irvine, J. M. (2003). Fusing Face and ECG for Personal Identification. *Proceedings of the 32nd Applied Imagery Pattern Recognition Workshop (AIPR)*.

Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 4- 20.

Jovanov, E., Poon, C., Yang, G.-Z., & Zhang, Y. T. (2009). Guest Editorial Body Sensor Networks: From Theory to Emerging Applications. *IEEE Transactions on Information Technology in Biomedicine*, 13(6), 859 - 863.

Kharate, G. K., Ghatol, A. A., & Rege, P. P. (2007). Selection of mother wavelet for image compression on basis of image. *IEEE Signal Processing, Communications and Networking*, Anna University, Chennai, India.

Kim, K.-S., Yoon, T.-H., Lee, J.-W., Kim, D.-J., & Koo, H.-S. (2005). A Robust Human Identification by Normalized Time-Domain Features of Electrocardiogram. *27th Annual IEEE International Conference of the Engineering in Medicine and Biology Society*.

Koh, L. H., Ranganath, S., Lee, M. W., & Venkatesth, Y. V. (1999). An integrated face detection and recognition system. *International Conference on Image Analysis and Processing*.

Kyoso, M. (2003). A technique for avoiding false acceptance in ECG identification. *Proceedings of IEEE EMBS Asian-Pacific Conf. Biomedical Engineering*, 190–191.

Kyoso, M., & Uchiyama, A. (2001). Development of an ECG identification system. *Proceedings of the 23rd IEEE Engineering in Medicine and Biology Confonerence*, 4, 3721-3723.

Love, J., Warren, S., Laguna, G., & Miller, T. (1997). Personal Status Monitor, SAND97-0418, DOE Category UC-706, February 1997, 199 pages.

Ludeman, L. & Chacon, M. (1995). Evaluation of Blood Pulse Signature for Potential Application in a Multisensor Biometric Identity Verification System. Final Report, Sandia National Laboratories, Contract No. AM-5506 Amendment No. 1, September 14, 1995, 16 pages.

Ludeman, L. & Chacon, M. (1996). Use of Blood Pulse Signature for Identity Verification. *Proceedings of the 7th International Conference on Signal Processing Applications & Technology*, October 7-10, 1996, Boston, Massachusetts, USA, pp. 1608-1612.

Ma, T., Zhang, Y., & 7hang, Y. (2006). *Biometrics, Wyle Encyclopedia of Biomedical Engineering*, April 14, 2006, pp. 1-13, http://onlinelibrary.wiley.com/doi/10.1002/9780471740360.ebs0164/pdf.

Mallat, S. (1999). *A wavelet tour of signal processing* (2 ed.): Academic Press.

Marcel, S., & Millan, J. D. R. (2007). Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation. *IEEE Transactions on Information Technology in Biomedicine Pattern Analysis and Machine Intelligence*, 29(4), 743 - 752.

Matsumoto, T., Matsumoto, H., Yamada, K., & Hoshino, S. (2002). *Impact of artificial gummy fingers on fingerprint systems.* Paper presented at the Proc. SPIE.

Matyas, V. J., & Riha, Z. (2003). Toward reliable user authentication through biometrics. *IEEE Security Privacy*, 1(3), 45–49.

Micheli-Tzanakou, E., Plataniotis, K., & Boulgouris, N. (2009). Electrocardiogram (ECG) Biometric for Robust Identification and Secure Communication. *Biometrics : Theory, Methods, and Applications* (pp. 383 - 427).

Ming, J., Hazen, T. J., Glass, J. R., & Reynolds, D. A. (2007). Robust speaker recognition in noisy conditions. *IEEE Transactions on Information Technology in Biomedicine Audio, Speech, and Language Processing*, 15(5), 1711 - 1723.

MIT-BIH Database Distribution. from http://ecg.mit.edu/ Harvard-MIT Division of Health Sciences and Technology.

Nasri, B., Guennoun, M., & El-Khatib, K. (2009). Using ECG as a measure in biometric identification systems. *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*.

Philips, P. J., Grother, P., Micheals, R. J., Blackburn, D. M., Tabassi, E., & Bone, J. M. (2003). Face recognition vendor test 2002: Overview and summary, from http://www.frvt.org/FRVT2002/documents.htm

Plataniotis, K. N., Hatzinakos, D., & Lee, J. K. M. (2006). ECG Biometric Recognition Without Fiducial Detection. *2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*.

The PTB Diagnostic ECG Database. from http://www.physionet.org/physiobank/database/ptbdb/, Physikalisch-Technische Bundesanstalt (PTB)

Saechia, S., Koseeyaporn, J., & Wardkein, P. (2005). Human Identification System Based ECG Signal. *TENCON 2005*, 2005 IEEE Region 10.

Shen, T. W., Tompkins, W. J., & Hu, Y. H. (2002). One-lead ECG for identity verification. *Proceedings of the 2nd Joint EMBS/BMES Conference*, 62–63.

Singh, Y. N., & Gupta, P. (2008). ECG to Individual Identification. *2nd IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2008.

Sufi, F., & Khalil, I. (2008). An automated patient authentication system for remote telecardiology. *International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, 2008.

Sullivan, T. J., Deiss, S. R., & Cauwenberghs, G. (2007). A low-noise, non-contact EEG/ECG sensor. *IEEE Biomedical Circuits and Systems Conference*, 2007, Montreal, Canada.

Wan, Y., Sun, X., & Yao, J. (2007, Oct. 17-20). Design of a photoplethysmographic sensor for biometric identification. *International Conference on Control, Automation and Systems 2007*, Seoul, Korea.

Wan, Y., & Yao, J. (2008, 22-24 October). A neural network to identify human subjects with electrocardiogram signals. *The World Congress on Engineering and Computer Science (WCECS) 2008 of International Association of Engineers (IAENG)*, San Francisco, CA.

Wang, Y., Plataniotis, K. N., & Hatzinakos, D. (2006). Integrating Analytic and Appearance Attributes for Human Identification from ECG Signals. *2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*.

Warren, S., & Jovanov, E. (2006). The need for rules of engagement applied to wireless body area networks. *3rd IEEE Consumer Communications and Networking Conference*, 2006 (CCNC 2006).

Warren, S., Lebak, J., Yao, J., Creekmore, J., Milenkovic, A., & Jovanov, E. (2005). Interoperability and Security in Wireless Body Area Network Infrastructures. *27th Annual International Conference of the Engineering in Medicine and Biology Society*, 2005. IEEE-EMBS 2005. .

Webster, J. G. (1998). *Medical instrumentation: Application and design*: Wiley.

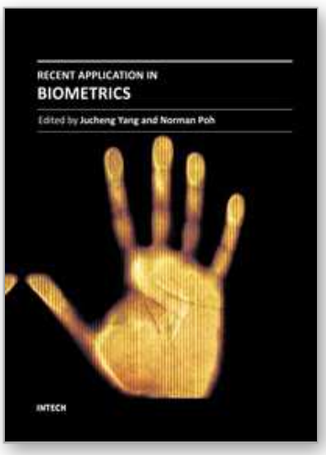
Yao, J., Sun, X., & Wan, Y. (2007, 23rd - 26th August). A pilot study on using derivatives of photoplethysmographic signals as a biometric identifier. *29th International Conference of the IEEE Engineering in Medicine and Biology*, Lyon, France.

Yao, J., & Wan, Y. (2008, June 1-3). A wavelet method for biometric identification using wearable ECG sensors. *5th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2008),* Hong Kong.

Yao, J., & Wan, Y. (2010). Improving Computing Efficiency of a Wavelet Method Using ECG as a Biometric Modality. *International Journal of Computer and Network Security*, 2(1), 15-20.

Zhang, G. H., Poon, C. C. Y., & Zhang, Y. T. (2009). A biometrics based security solution for encryption and authentication in tele-healthcare systems. *2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies*, 2009 (ISABEL 2009).

Zhang, Z., & Wei, D. (2006). A new ECG identification method using Bayes' Theorem. *TENCON IEEE Region 10 Conference*, Hong Kong.



Recent Application in Biometrics

Edited by Dr. Jucheng Yang

ISBN 978-953-307-488-7
Hard cover, 302 pages
Publisher InTech
Published online 27, July, 2011
Published in print edition July, 2011

In the recent years, a number of recognition and authentication systems based on biometric measurements have been proposed. Algorithms and sensors have been developed to acquire and process many different biometric traits. Moreover, the biometric technology is being used in novel ways, with potential commercial and practical implications to our daily activities. The key objective of the book is to provide a collection of comprehensive references on some recent theoretical development as well as novel applications in biometrics. The topics covered in this book reflect well both aspects of development. They include biometric sample quality, privacy preserving and cancellable biometrics, contactless biometrics, novel and unconventional biometrics, and the technical challenges in implementing the technology in portable devices. The book consists of 15 chapters. It is divided into four sections, namely, biometric applications on mobile platforms, cancelable biometrics, biometric encryption, and other applications. The book was reviewed by editors Dr. Jucheng Yang and Dr. Norman Poh. We deeply appreciate the efforts of our guest editors: Dr. Girija Chetty, Dr. Loris Nanni, Dr. Jianjiang Feng, Dr. Dongsun Park and Dr. Sook Yoon, as well as a number of anonymous reviewers.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Jianchu Yao, Yongbo Wan and Steve Warren (2011). Biometric Applications of One-Dimensional Physiological Signals - Electrocardiograms, Recent Application in Biometrics, Dr. Jucheng Yang (Ed.), ISBN: 978-953-307-488-7, InTech, Available from: http://www.intechopen.com/books/recent-application-in-biometrics/biometric-applications-of-one-dimensional-physiological-signals-electrocardiograms

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen