Liveness Detection and Automatic Template Updating using Fusion of ECG and Fingerprint

Majid Komeili, Narges Armanfard, Dimitrios Hatzinakos

***Abstract*—Fingerprint has been extensively used for biometric recognition around the world. However, fingerprints are not secrets and an adversary can synthesis a fake finger to spoof the biometric system. The mainstream of the current fingerprint spoof detection methods are basically binary classifier trained on some real and fake samples. While they perform well on detecting fake samples created by using the same methods used for training, their performance degrades when encountering fake samples created by a novel spoofing method. In this paper, we approach the problem from a different perspective by incorporating ECG. Compare with the conventional biometrics, stealing someone's ECG is far more difficult if not impossible. Considering that ECG is a vital signal and motivated by its inherent liveness, we propose to combine it with a fingerprint liveness detection algorithm. The combination is natural as both ECG and fingerprint can be captured from fingertips. In the proposed framework, ECG and fingerprint are combined not only for authentication purpose but also for liveness detection. We also examine automatic template updating using ECG and fingerprint. In addition, we propose a stopping criterion that reduces the average waiting time for signal acquisition. We have performed extensive experiments on LivDet2015 database which is presently the latest available liveness detection database and compare the proposed method with six liveness detection methods as well as twelve participants of LivDet2015 competition. The proposed system has achieved a liveness detection EER of 4.2% incorporating only 5 seconds of ECG. By extending the recording time to 30 seconds, liveness detection EER reduces to 2.6% which is about 4 times better than the best of six comparison methods. This is also about 2 times better than the best results achieved by participants of LivDet2015 competition.**

***Index Terms*—Electrocardiogram, Fingerprint, Liveness Detection, Biometric, template updating.**

I. Introduction

BIOMETRIC systems have been deployed around the world and have been extensively used in the past decades. However the potential of fooling or spoofing this technology is widely admitted. Nowadays biometric spoof detection is an active research area and there has been a lot of efforts towards a promising approach to ensure the presence of a real legitimate user.

There are different attack points in a biometric system. The first vulnerable point is the sensor used in the biometric system. Biometric systems and in particular fingerprint can be spoofed by presenting synthetic samples to the sensor e.g. gummy fingers that have fingerprint impressions. There has been a huge literature on other vulnerabilities that for example bypass feature extraction or matcher, or manipulate database or communication channel. However, in such cases, some information about the system such as feature extraction, matcher, database and/or physical access to some of those components is necessary. In contrast, fooling the sensor using a fake biometric sample does not need any specific information about internal mechanism of the biometric system. In addition, sensor level attacks are in analog domain and hence many solutions such as cryptography and watermarking that are in digital domain are not useful. This highlights the importance of developing biometric spoofing countermeasures to classify an input sample as live or fake that is focus of this study. Investigating other types of attacks is out of scope of this study.

Liveness detection has been an active area in the past decade and numerous approaches have been proposed in the literature to solve this problem. Considering the results reported in LivDet2009 [1], LivDet2011 [2], LivDet2013 [3] and LivDet2015 [4], liveness detection is still an open problem and performance of the existing approaches does not satisfy requirements of many practical applications [5]. The mainstream of the current approaches use some training samples artificially created via certain spoofing process and work well on test samples created by the same process involved in training, but their performance on a novel type of spoof is questionable. In practice, the way that a fake biometric is fabricated is unknown.

In this study we approach the problem from a different perspective by incorporating electrocardiogram (ECG). ECG is among the newer additions to the biometric family and unlike the conventional biometrics such as fingerprint, iris and face, ECG is a vital signal and presence of the ECG automatically ensures the liveness [6]–[12]. In addition, conventional biometrics can be easily stolen from people. For example, fingerprints may be left behind whenever we touch a glass surface. Even iris images can be captured from a few meters distance. Not to mention face images which can be captured from a longer distance. However, compared to the conventional biometrics, if not impossible it is far more difficult to steal someone's ECG. Beside its advantages, accuracy of ECG is not as good as some other mature biometrics such as fingerprint [13], [14]. Therefore, we seek to fuse ECG and fingerprint to improve the recognition rate and more importantly the liveness detection performance.

Fusion of ECG and fingerprint is not a new idea. In [15] and [16] it has been suggested to combine ECG with other biometrics such as face and fingerprint to get a better recog-

M. Komeili and D. Hatzinakos are with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ontario, Canada. E-mail: {mkomeili, dimitris}@ece.utoronto.ca

N. Armanfard is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada. E-mail: armanfn@mcmaster.ca

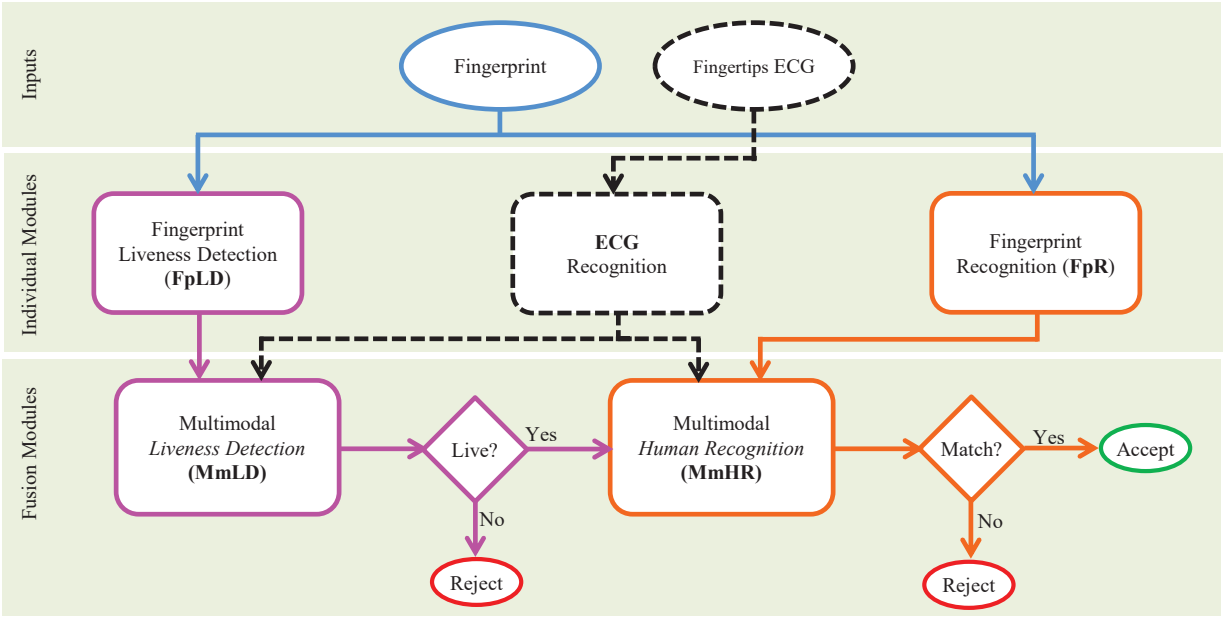

Fig. 1. Block diagram of the proposed method.

nition rate. But, all these works were restricted to analyzing recognition rate and have failed to consider spoof attacks. The improved verification rate in the conventional fusion approach is because when input samples of one trait have poor quality and hence less informative, the other trait will help the system to still identify the user. However, this opens up the possibility of spoofing because such system may accept a fake copy of an authentic fingerprint even if ECG does not match. This contradicts the main motivation in utilizing ECG which is liveness detection. This issue has been overlooked in the previous works [15], [16].

In this paper, we offer an alternative to the conventional fusion of ECG and fingerprint by proposing to fuse ECG with fingerprint for liveness detection. To this end, we combine ECG recognition score with fingerprint liveness detection score instead of fingerprint recognition score. This greatly improves the accuracy of liveness detection task. Figure 1 shows an overview of the proposed framework which will be explained in detail in section III. In this context, we use terms "fingerprint liveness detection" and "fingerprint recognition" to indicate two blocks in the proposed system as shown in Figure 1. Our ECG signals are collected from fingertips as shown in Figure 2. Majority of previous works on ECG recognition have been based on signals collected form chest area [17]–[21] or lower rib cage [22] and only a few works have been done based on fingertip ECG signals, e.g. [23], [24]. Fingertip ECG has two advantages: first, it eliminates the need for user to undress for electrode placement; second, it makes the fingerprint a natural choice to be fused with ECG. The main contributions of this paper are as follows:

- In order to get the most out of ECG, we fuse it with a fingerprint liveness detection method for liveness detection purpose and also fuse it with a fingerprint recognition method for recognition purpose. Although the latter has been previously investigated, e.g. in [15] and [16], to the best of our knowledge, the former has not been explored in the previous literature.
- In addition, the proposed system is capable of automatically adapting ECG and fingerprint templates to operational data. Since ECG is a time dependent signal and its waveform might be affected by factors like diet and emotion, template updating is crucial to maintain the performance of the system in long term without requiring to re-enroll or retrain the system from scratch and to the best of our knowledge this has not been investigated in the literature.
- Another shortcoming of the previous works on ECG [7], [22], [25]–[31] is lack of a proper stopping criterion to limit the length of recording sessions. Previous works usually limits the length of sessions by fixing the number of recorded heartbeats to a predefined threshold. Therefore, the number of recorded heartbeats is the same for all subjects which does not consider the fact that some subjects have a very stable ECG and do not need as many samples as other subjects with less stable ECG. In this paper, we present an easy-to-compute yet effective criterion based on local averaging and correlation that measures heartbeat consistency (HC) in successive heartbeats.

The rest of this paper is organized as follows: Section II briefly reviews the previous works on ECG recognition, fusion of ECG and fingerprint as well as fingerprint liveness detection. Section III provides an overview of the proposed approach. Datasets used in the experiments are described in section IV. Details of the proposed approach including ECG recognition, fusion of ECG and fingerprint and template updating are explained in section V, VI and VII respectively. Computational costs are discussed in section VIII and section IX concludes the paper.

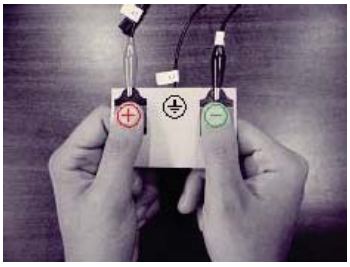

Fig. 2. Recording ECG signals from fingertips.

II. Related work

A. ECG Recognition

Previous works on ECG biometric can be categorized in fiducial-based and nonfiducial-based approaches. Fiducial based approaches rely on some points of a heartbeat such as onset and end of each wave. This requires P, R and T waves to be located and features such as peaks, slope, radius of curvature and area be computed in a region surrounding each of P, R, and T waves [19]–[21]. Detection of such characteristic points, however, may not be always possible due to noise. Therefore, following [25], we do not consider fiducial dependent approaches.

On the other side, nonfiducial based approaches are holistic and consider ECG signal as a set of heartbeats or just a time series without segmenting it to heartbeats. For example, in [18] signals are segmented in overlapping windows and autocorrelation features are extracted and linear discriminant analysis (LDA) is used for dimension reduction. In [22], short time Fourier transform (STFT) features are extracted and after a feature selection, log-likelihood ratio is used for classification. In [17], sparse coefficients of an over-complete dictionary is used as features. Max-pooling is used for aggregation of samples to construct templates. In the following section, previous works on fusion of ECG and fingerprint will be discussed.

B. Fusion of ECG and Fingerprint

While multimodal biometric systems based on conventional traits such as face and fingerprint have been extensively investigated in the literature [32], [33], there exist only a few works about a multimodal biometric system that includes ECG. In [16] the idea of securing handheld devices and fingerprint readers with ECG biometrics is pointed out and a biometric system based on ECG signals collected from fingertips is investigated but no experimental result on fusion of ECG and fingerprint was reported. In [15], fusion of face, fingerprint and ECG is studied. They used an ECG dataset of 78 subjects acquired from European ST-T Database, MIT-BIH Normal Sinus Rhythm Database, MIT-BIH Arrhythmia Database and QT Database of PhysioBank [34]. A fiducial based method previously presented in [35] were used for ECG recognition. It involves extracting fiducial features related to various intervals, amplitude and angles. For fingerprint and face recognition, match scores provided in NIST-BSSR1 [36] were used. BSSR1 has match scores for two faces and one fingerprint matcher. Since ECG signals used in their work were acquired from chest area, its fusion with fingerprint is of little value in real-world scenarios because the user needs to undress for electrode placement. Further, the ECG recognition method that was used is fiducial-based and, as suggested in [25], is not appropriate for real-world scenarios where accurate detection of characteristic points may not be possible due to noise. Moreover, they performed fusion only to achieve a better recognition rate. They indeed failed to benefit from the main advantage of ECG, i.e. its inherent liveness detection. However, our work does not have aforementioned limitations. Our ECG signals are collected from fingertips. More importantly, in addition to fusion of ECG and fingerprint for human recognition, for the first time, we perform fusion of ECG and fingerprint for liveness detection. In the following section, we briefly explain the previous works on fingerprint liveness detection.

C. Fingerprint Liveness Detection

Liveness detection methods can be divided into two main categories: hardware or software-based. Hardware-based techniques add some specific components to the capture device that can look for particular properties such as fingerprint sweat or blood pressure. In software-based solutions, detection is performed by processing the obtained image. In [37], morphology-based and perspiration-based features are jointly considered. Pore perspiration is investigated in [38]–[40]. In [41] the difference between quality of fake and real samples is used to detect fake samples. This is realized through analyzing 25 image quality measures extracted from samples. There are other works that have successfully applied local descriptors for liveness detection. Local Binary Pattern (LBP) and its extensions have been used for liveness detection in [42], [43]. LBP is a texture descriptor based on intensity difference between a pixel and its neighboring pixels [44]. In [45], a method based on Local Phase Quantization (LPQ) was proposed. Similar to LBP, LPQ works on patches but instead of gradient it computes phase information by computing short time Fourier transform. Phase information are then decorrelated and uniformly quantized. Another descriptor that has been used for liveness detection is Weber Local Descriptor WLD [46]. Binarized Statistical Image Features (BSIF) [47] was also tested for fingerprint liveness detection. It is based on binarizing the response to a set of filters. Filters are not fixed and are learnt using independent component analysis (ICA). Recently a method known as Local Contrast Phase Descriptor (LCPD) is presented in [48] that takes into account a spatial-domain component inspired by WLD and a phase component inspired by LPQ. There are some other works [49], [50] that are based on deep neural networks. Such methods learn the features from training data. To reduce the number of training data required to train such networks, pre-trained deep networks have been used in [49] and [50].

To improve the performance of liveness detection, there has been a few works on fusion of liveness scores with conventional biometrics such as fingerprint and face [37], [51]. Since these biometrics are easy to spoof, when involved in the fusion process, make the security of the resulting

system questionable. In the 2-dimensional space defined by liveness and recognition scores, distribution of samples from different spoofs can greatly vary [51]. Therefore, any decision boundary in the aforementioned 2-dimensional space may fail when encountering a novel spoof with a different distribution. Recently, a method based on 1-median filtering is presented in [52]. As an alternative to sum rule as a conventional fusion rules, 1-median filtering is used for a mutibiometric problem with 5 traits. However, 1-median filtering requires enough number of traits to determine the median and it cannot be used in our scenario that has 2 traits, i.e. ECG and fingerprint.

In spite of all efforts in this field, accuracy of current liveness detection methods does not satisfy the requirement of many practical applications. All these methods suffer from significant performance variation when encountering different spoofs especially when encountering novel spoofs that have not been seen during training [5]. A detailed survey of liveness detection approaches can be found in [53] and [54].

III. Overview of the Proposed Approach

A. ECG Recognition

ECG is a non-stationary signal and factors like diet, emotion and heart rate affect its waveform. We consider a scenario such that subjects are enrolled in one recording session and tested on another session which is at least one week apart. Under this scenario, while enrollment is restricted to only one session, we can benefit from an auxiliary dataset of some generic subjects for which multiple sessions are available. By looking into multiple sessions of the auxiliary dataset, we select a subset of features that are more stable across different sessions. After performing feature selection on the auxiliary dataset, templates (i.e. classifiers) of biometric system's users (i.e. test dataset) are constructed by using only the selected features. Auxiliary dataset is dedicated to feature selection and there is no overlap between subjects of auxiliary dataset and actual biometric system's users involved in enrollment and testing.

B. Fusion of ECG and Fingerprint

In general, a fingerprint biometric system has a fingerprint liveness detection (FpLD) module and a fingerprint recognition (FpR) module. A desired system should have a good recognition rate as well as a good liveness detection performance. Block diagram of the proposed framework is shown in Figure 1.

FpLD is a binary classifier (i.e. live/fake) that computes liveness score for a test fingerprint sample. ECG recognition can also be treated as a binary classifier (i.e. genuine/impostor) that computes an authentication score for a test ECG signal. ECG can be fused with FpLD to form a Multimodal Liveness Detection (MmLD) block. The use of ECG for fingerprint liveness detection is motivated by the assumption that ECG is hard to spoof compare with the conventional biometrics and presence of an authentic ECG inherently implies liveness of the subject. FpR block can also be treated as a binary classifier (i.e. genuine/impostor) such that given a fingerprint sample, it provides an authentication score that can be further fused with ECG score to form a Multimodal Human Recognition (MmHR) block. While the MmLD block does the liveness detection task and aims to reject the spoof attempts, the second block (i.e. MmHR) does the recognition task and aims to reject impostor attempts.

C. Template Updating

ECG and fingerprint samples that their scores are greater than some updating thresholds can be added to previously existing training samples to update the templates. This process is known as template updating or adaption [55]. The updating thresholds are usually different than the acceptance thresholds and are set to zero false acceptance rate point to ensure that only live and genuine samples are selected for template updating.

IV. Datasets

We use ECG database collected in our lab (BioSec) in University of Toronto [56]. ECG signals were recorded using Vernier EKG sensor and Go!Link interface [57] with 12-bit resolution and sampling rate of 200 Hz using three dry AgCl electrodes from fingertips as shown in Figure 2. There are 82 subjects that have 2 or more ECG recordings in sitting posture. 46 out of 82 subjects have exactly 5 sessions. Follow-up sessions are collected over a six-month period. We divide the database into 2 parts. Aforementioned 46 subjects are used for enrollment and testing (i.e. testing dataset) and the remaining 36 subjects are used as an *auxiliary dataset* for feature selection that will be described in section V-B.

We use LivDet2015 fingerprint database [4] which includes 4 datasets corresponding to 4 different scanners: CrossMatch, DigitalPersona, GreenBit and Biometrika. Among them, CrossMatch is excluded from our experiments due to small number of spoof samples per finger that makes it unsuitable for template updating. Characteristics of these datasets are shown in Table I.

The three datasets that are considered in this study contain live and fake samples that are constructed through 6 different spoofs including Ecoflex, Gelatine, Latex, WoodGlue, Liquid Ecoflex and RTV. Each dataset originally comes in two sets: training and testing. The training set is the same as testing set but does not include Liquid Ecoflex and RTV spoofs. In this study, to emphasize on the more challenging case of detecting unknown spoofs, one experiment is performed for every spoof such that samples of that spoof are omitted from the training set. Thus, type of spoof is always unknown in all 6 experiments of each scanner.

A chimeric dataset is constructed by combining the above ECG and fingerprint datasets. Each subject in ECG dataset is paired with a subject in the fingerprint dataset in a random way. In this way, a chimeric dataset of 46 unique subjects is generated. All experiments are repeated 50 times to cope with the randomness in constructing the chimeric dataset and the average results are reported. We consider each finger as a subject due to the limited number of actual subjects. Note that not all fingers have fake samples. Also, not all subjects have samples of all 10 fingers. For each subject 4 live and 3 fake

TABLE I. Characteristics of LiveDet2015 [4] datasets used in our experiments.

Scanner	Model	Resolu. (dpi)	Size (px)	Format
DigitalPersona	U.are.U5160	500	252×324	PNG
GreenBit	DactyScan26	500	500×500	PNG
Biometrika	HiScan-PRO	1000	1000×1000	BMP

samples are used: 1 live sample for enrollment and 3 live and 3 fake samples for subsequent test sessions. The number of test sessions, 3, is due to restriction of the number of fake and live samples that belong to the same subject in LivDet2015 database.

As explained above, the ECG dataset has 5 sessions. But, the fingerprint datasets have only 4 sessions. We pick the 2nd session for training and 3rd, 4th and 5th sessions for testing. We ignore the first session and train on the second session because the second session is less noisy that makes it a better choice for enrollment and constructing templates.

V. ECG Recognition

A. Description of Features

Different types of features have been employed in the literature to represent ECG signals. Unlike many of previous works that rely on only one type of feature, we concatenate different types of features and construct a fairly comprehensive feature vector. We use short-time Fourier transform with Hamming window of the length 16 with step size of 13 computed over a 1 second window centered at R peak. Continuous wavelet transform with 32 scales and Daubechies 5 as mother wavelet is computed on a 1 second window centered at R peak. Mean of power, standard deviation of power, maximum amplitude, standard deviation of amplitude, kurtosis and skewness are computed on a 2-second window centered at R peak on the following frequency bands: 8-13Hz, 13-18Hz, 18-25Hz, 25-30Hz, 30-35Hz, 35-50Hz. The signal's amplitude itself is also considered. This gives a total of 7198 features. We use z-score normalization for every feature, so that features have zero mean and unit variance after normalization.

B. Feature Selection

Let $\mathcal{X} = \{\mathbf{X}_1, \ldots, \mathbf{X}_N\}$ be the auxiliary dataset consisting of N subjects and assume that there are M_i different sessions available for the i-th subject i.e. $\mathbf{X}_i = \{\mathbf{X}_{i,1}, \ldots, \mathbf{X}_{i,M_i}\}$, where $\mathbf{X}_{i,j}$ consists of samples of j-th session of i-th subject. Weight of l-th feature consists of two terms. The first term $\mathbf{w}_1(l)$ encourages class separability and is defined as follows:

$$\mathbf{w}_1(l) = \frac{1}{N} \sum_{i=1}^{N} d\left(f(\mathbf{X_i}(l)), f(\mathcal{X}(l))\right) \tag{1}$$

where $d(\cdot)$ is the symmetric Kullback-Leibler divergence and $f(\cdot)$ denotes probability density function (pdf). $f(\mathbf{X_i}(l))$ is pdf of l-th feature computed over all samples of i-th subject and $f(\mathcal{X}(l))$ is pdf of l-th feature computed over all samples. We consider normal distribution and use maximum likelihood estimates, i.e. sample means and variances to estimate the symmetric Kullback-Leibler divergence as follows:

$$d(f_1, f_2) = \frac{\sigma_1^2 + (\mu_1 - \mu_2)^2}{2\sigma_2^2} + \frac{\sigma_2^2 + (\mu_1 - \mu_2)^2}{2\sigma_1^2} - 1 \tag{2}$$

where $f_1 = \mathcal{N}\left(\mu_1, \sigma_1^2\right)$ and $f_2 = \mathcal{N}\left(\mu_2, \sigma_2^2\right)$ are two distributions whose distance is to be computed. $\mathbf{w}_1$ is large when the overall distribution of every subject is different from the overall distribution of all subjects. However, $\mathbf{w}_1$ only considers overall class distribution of a subject and disregards the session distributions of that subject. To address this limitation, we define the second term, $\mathbf{w}_2$, as follows:

$$\mathbf{w}_2(l) = \frac{1}{\sum_{i=1}^{N} M_i} \sum_{i=1}^{N} \sum_{j=1}^{M_i} d\left(f(\mathbf{X}_{i,j}(l)), f(\mathbf{X_i}(l))\right) \tag{3}$$

The first summation is over all subjects and the second summation is over all sessions and denominator is a normalization factor. $\mathbf{w}_2$ encourages the stability across multiple sessions and is smaller when a session can represent the actual distribution of the corresponding subject. Finally, considering both $\mathbf{w}_1$ and $\mathbf{w}_2$, weight of the l-th feature, $\mathbf{w}(l)$, is defined as follows:

$$\mathbf{w}(l) = \lambda \mathbf{w}_1(l) - (1 - \lambda)\, \mathbf{w}_2(l) \tag{4}$$

where λ is a parameter that controls the trade-off between class separability and across session stability. Features can be selected by comparing their weights with a threshold. We can also sort features according to their weights and pick top features. A proper value for λ can be determined using cross validation. In our experiments, λ and the number of selected features are set to 0.3 and 400 respectively. The proposed feature selection method has been presented in our previous study in [58].

C. Stopping Criteria

Previous works on ECG recognition [22], [25]–[31] suffer from being restricted to have a predefined recording length. Motivated by the fact that some subjects have a very repetitive ECG and does not need as many heartbeats as some others, we define heartbeat consistency (HC) as a stopping criterion as follows:

$$HC(n) = corr\left((\mathbf{b}_n + \mathbf{b}_{n+1}), (\mathbf{b}_{n+2} + \mathbf{b}_{n+3})\right) \tag{5}$$

where $corr$ is correlation function and $\mathbf{b}_n$ is n^{th} heartbeat amplitude. We keep recording until HC exceeds a predefined threshold or recording time reaches 30 seconds, whichever satisfied first. Therefore, a session can take at most 30 seconds. As shown in Figure 3, increasing HC threshold increases the average recording time because a larger HC threshold is harder to satisfy. For example, recording times for different sessions of the subjects for HC=0.92 are shown in Figure 4. This is corresponding to an average recording time of about 10 seconds. For each of 46 subjects, recording time of 3 test sessions are shown with circles. The average length of sessions are also provided as a solid red line.

The assumption behind HC is that in general recording more ECG samples in a session improves the accuracy. This

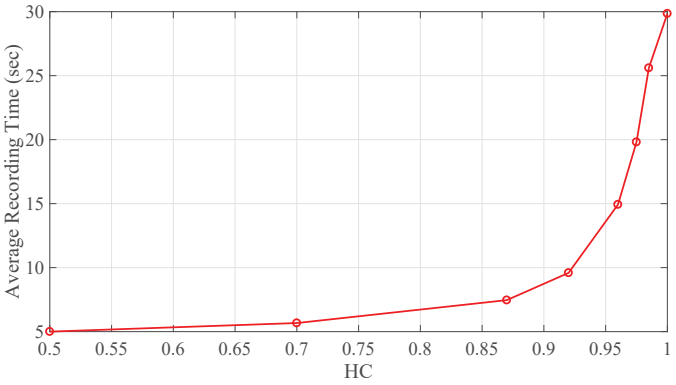

Fig. 3. Average recording time versus HC.

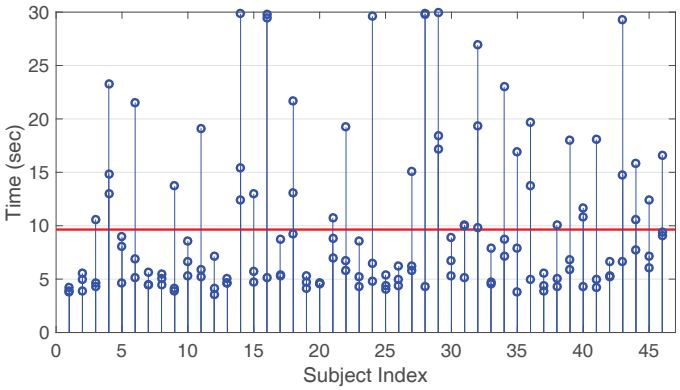

Fig. 4. Recording time for different subjects for HC=0.92. For each subjects 3 values corresponding to 3 test sessions are marked with circles. The average recording time is also shown as a solid red line.

assumption is experimentally validated in the next section (–see Figure 5) which is in line with the body of the literature that have investigated the effect of the number of samples on accuracy of ECG biometric recognition (e.g. [22], [25]–[31]). HC makes the system more convenient (in terms of acquisition time) for majority of the subjects. A more stable ECG naturally implies a higher degree of redundancy among heartbeats. If there is diversity in the input signal, HC lets the system to capture more samples (by extending the recording time) which usually results in a more reliable decision. For example, as shown in Figure 4, when the average recording time is about 10 seconds (i.e. HC=0.92), majority of the cases (i.e. about 60%) experience a shorter acquisition time of about 5 seconds. That means HC has been able to cut the acquisition time in half for majority of the subjects.

HC as defined in (5) is based on 4 heartbeats. Increasing the number of heartbeats involved in computing HC, will increase the acquisition time for all subjects which is not desired. Another drawback is that it degrades the local behavior of HC. For example, assume that we have a stable ECG signal with a distortion around the 10th heartbeat. If HC is defined over 4 heartbeats (i.e. correlation between averages of the first 2 heartbeats and the second 2 heartbeats as in (5), it stops as soon as we get 4 heartbeats. However, if HC be defined over for example 10 consecutive heartbeats (i.e. correlation between averages of the first 5 heartbeats and the second 5 heartbeats) we need to keep recording until 20th heartbeat to get rid of the effect of 10th heartbeat (i.e. a poor local behavior). Therefore, increasing the number of heartbeats involved in computing HC is against the local behavior of HC and is not desired. On the other side, we need at least 4 heartbeats to perform averaging and correlation. Averaging is a simple yet effective and widely used noise reduction method and makes the correlation results more reliable. Therefore, we consider 4 heartbeats in computing HC as in (5).

Any R peak appeared during the first second of recording cannot be used because some of the aforementioned features require a 2-second window centered at R peak. Moreover, by definition HC needs at least 4 full heartbeats and that limits the minimum recording time to about 5 seconds.

D. Classification

We use linear SVM for classification and since it is a binary classifier, a one-versus-all strategy is adopted. The score corresponding to a session (i.e. s) is determined as weighted sum of its heartbeats scores (i.e. s_i) as follows:

$$s = \frac{\sum_{i=1}^{K} w_i s_i}{\sum_{j=1}^{K} w_j} \tag{6}$$

Note that length of sessions are determined by HC and K is not fixed and above summation is indeed over all heartbeats of each session. Denominator is just a normalization factor and weights w_i are defined as:

$$w_i = \frac{1}{1 + \exp(-\frac{s_i - \mu}{\sigma})} \tag{7}$$

where s_i denotes score of i-th heartbeat and μ and σ are mean and standard deviation of the scores in the corresponding session. The final decision for a session can be made by comparing its score against an acceptance threshold.

Such definition of weights (after normalization) can be interpreted as probabilities assigned to heartbeats where higher probabilities are assigned to the heartbeats with larger scores. This makes the final decision less sensitive to outliers because outliers are artifacts that do not look like a regular heartbeat and hence usually do not match templates and their weights in (6) are small. Heartbeats with a very small HC value are usually very irregular due to muscle or electrode movements. Therefore, heartbeats with HC value below 0.5 are considered as outlier and discarded. Note that this does not require further computation because HC values for the heartbeats that precede the stopping point are already computed.

We have also considered the outlier removal methods presented in [59] (i.e. DMEAN and DBSCAN). DMEAN method detects outliers based on their distance to mean template (i.e. mean of all heartbeats) and requires computing some statistics such as mean, standard deviation, min and max over entire session. DBSCAN method is based on clustering all samples of a session and detects samples not belonging to any cluster as outlier. Since in [59] DMEAN outperformed DBSCAN in all experiments, we only consider DMEAN method as an alternative to our outlier removal method based on HC values.

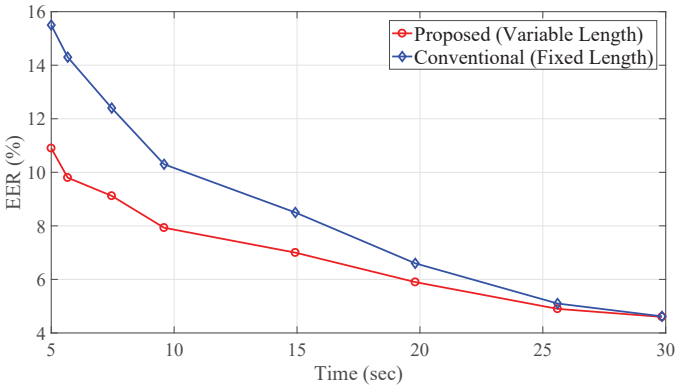

Fig. 5. ECG recognition performance. EER of ECG recognition using the proposed stopping criteria HC is compared with ECG recognition without HC stopping criteria, i.e. same recording length for all sessions.

Note that the results reported in [59] for outlier removal are based on estimating mean template over 2-minute sessions but we apply DMEAN only to the portion of the signal that precedes the stopping point. Nonetheless, both methods yield similar performance while our method performed slightly better (i.e. 0.4%) over different recording lengths. This can be attributed to the fact that even if the outlier removal module fails to detect a few artifacts, the effect of those artifacts will be mitigated by the weighting strategy in (6) –i.e. less sensitivity to outliers.

We force the enrollment session to include a minimum of 22 heartbeats. This sets the minimum length of an enrollment session to about 20 seconds. Figure 5 shows EER of ECG versus average length of ECG recording. Different recording length is achieved through changing the HC threshold as suggested in Figure 3. It can be seen that increasing the recording time improves the performance of ECG recognition. For comparison, EER of the conventional approach that has a fixed session length is also presented in Figure 5. It can be seen that the proposed stopping criterion effectively reduces the recording time needed to achieve a desired EER. As the recording time approaches 30 seconds, HC criterion is dominated by the 30-second limit and converges to the conventional approach. Recording time beyond 30 seconds is not considered in our experiments because it is of little value in a real-world scenario. However, increasing the maximum recording time to 1 minute leads to an average EER of 3.8% and further increasing the maximum recording time to 2 minutes did not improve the performance anymore. Since not all subjects have recordings longer than 2 minutes recording, we did not go beyond 2 minutes. Results in Figure 5 are computed over 46 subjects. Considering that there are 3 test sessions per subject, the number of positive and negative trials is 46×3 and $46 \times 45 \times 3$ respectively.

VI. Fusion of ECG and Fingerprint

In this section, we first investigate the fusion of ECG and fingerprint for liveness detection purpose (i.e. MmLD block). Then, we investigate the fusion of ECG and fingerprint for human recognition purpose (i.e. MmHR block).

A. Multimodal Liveness Detection (MmLD)

We examine 6 different fingerprint liveness detection algorithms including LCPD[1] [48], BSIF[2] [47], IQA[3] [41], LBP[4] [44], LPQ[5] [45] and WLD[6] [46]. Performance of these methods are provided in the right side of Table II, III and IV corresponding to 3 different scanners. It can be seen that on average LCPD outperforms the other liveness detection methods. Therefore, we use it in the subsequent experiments and fuse it with ECG.

The resulting multimodal liveness detection module, referred as MmLD, can be realized through different fusion rules. We consider weighted sum, product and maximum rules. In the weighted sum rule, weights can be determined by using an evaluation set. However, a simple approach to determine the weights is to choose them proportional to the EER of the individual traits. In our experiments, since EER of ECG is roughly half of EER of LCPD, we set the weight of ECG and LCPD to $2/3$ and $1/3$, respectively. As shown in Table II, III and IV, the sum rule outperforms the other fusion rules. Therefore, we use it in the subsequent experiments. It can be seen that the proposed method (i.e. fusion of ECG and LCPD) performs significantly better than the comparison methods. Results in Table II, III and IV are computed over 46 chimeric subjects for which there are samples for 3 test sessions. Therefore, there are 46×3 positive trials and 46×3 negative trials (spoof trials) in each of 50 runs. Note that since there are more subjects in the fingerprint dataset than ECG dataset, we randomly pick 46 fingerprint subjects in each run and report the average and standard deviation values computed over 50 runs.

We also compare the performance of the proposed method with participants of LivDet2015 competition [4]. LivDet2015 test set includes 6 different spoofs, among them LiquidEcoflex and RTV are not in the training set. Since they have similar training set we can test on both at the same time. This is the same as the protocol used in LiveDet2015 competition for unknown spoofs except that instead of using all test samples of LiquidEcoflex and RTV (i.e. 500 fake and 1000 live samples) at once, we use 276 spoof and 276 live samples (i.e. $46 \times 3 \times 2$) in each run. Note that since we repeat the experiments 50 times, all samples are indeed involved in our experiments. Table V shows the half total error rate (HTER) for participants of LivDet2015 competition on aforementioned unknown spoofs as well as EER of the proposed method for various ECG lengths. For example ECGFP-5 denotes the proposed method with 5 seconds of ECG. Results for LivDet2015 participants in Table V are reported from [4]. It can be seen that the proposed method (i.e. fusion of the ECG and LCPD) performs significantly better than the state-of-the-art methods in LivDet2015 competition.

Note that performance of LCPD used in our method is in the range of top performers of LivDet2015 competition. If instead of LCPD, we used a better algorithm, we could perform fusion

[1] http://www.grip.unina.it/web-download.html
[2] http://www.ee.oulu.fi/ jkannala/bsif/bsif.html
[3] Codes for IQA method was obtained directly from the author.
[4] http://www.cse.oulu.fi/wsgi/CMV/Downloads/LBPSoftware
[5] http://www.cse.oulu.fi/wsgi/CMV/Downloads/LPQMatlab
[6] http://www.cse.oulu.fi/wsgi/CMV/Research/

TABLE II. EER (in percent) of MmLD (fusion of ECG and LCPD) and six comparison methods on DigitalPersona dataset. Standard deviations are in parentheses.

	Fusion of ECG and LCPD									No Fusion					
	Sum			product			Max			LCPD	BSIF	IQA	LBP	LPQ	WLD
Time (sec)	5	10	30	5	10	30	5	10	30	-					
Ecoflex	2.5 (0.9)	2.2 (0.7)	1.9 (0.8)	6.4 (0.9)	4.1 (0.8)	3.6 (0.8)	4.0 (1.1)	3.9 (1.0)	3.9 (1.0)	6.0 (1.1)	15.4 (2.1)	21.7 (2.2)	8.0 (1.3)	10.5 (1.7)	19.2 (2.6)
Gelatine	6.3 (1.3)	4.9 (1.2)	4.1 (1.1)	8.8 (1.1)	5.5 (0.7)	4.1 (0.7)	16.9 (1.9)	16.9 (2.0)	16.1 (1.7)	17.4 (1.7)	23.3 (2.3)	22.0 (2.0)	14.1 (1.6)	16.5 (1.4)	30.3 (2.6)
Latex	2.7 (0.7)	2.3 (0.7)	2.0 (0.9)	4.4 (1.5)	2.9 (0.9)	2.5 (0.6)	3.9 (1.0)	3.9 (1.0)	3.8 (1.0)	4.9 (0.8)	11.9 (1.2)	9.4 (1.7)	6.9 (1.1)	11.4 (1.6)	14.8 (1.9)
LiqEcoflex	7.1 (2.1)	4.9 (1.3)	3.7 (1.1)	8.0 (1.5)	5.6 (1.1)	4.6 (1.1)	15.5 (2.0)	15.6 (2.0)	15.4 (2.0)	16.7 (2.1)	17.8 (1.9)	29.7 (2.7)	22.2 (1.8)	21.0 (2.3)	27.1 (2.5)
RTV	4.0 (0.9)	3.1 (1.2)	2.6 (0.6)	6.6 (0.8)	4.2 (0.7)	3.6 (0.8)	6.1 (1.3)	6.1 (1.5)	6.0 (1.4)	8.0 (1.3)	15.0 (1.7)	21.2 (2.1)	10.3 (1.3)	16.5 (1.8)	19.7 (2.4)
WoodGlue	10 (1.5)	7.0 (1.3)	4.9 (1.0)	10.2 (1.1)	6.7 (1.5)	5.3 (1.4)	25.2 (1.8)	25.3 (1.8)	25.3 (1.8)	27.3 (1.6)	32.4 (2.4)	45.8 (2.4)	27.7 (2.8)	39.2 (2.6)	42.1 (2.8)
Average	**5.4** (0.7)	**4.1** (0.5)	**3.2** (0.4)	7.4 (0.6)	4.8 (0.5)	3.9 (0.5)	11.9 (0.6)	11.9 (0.6)	11.8 (0.7)	13.4 (0.5)	19.3 (0.9)	25.0 (1.2)	14.9 (0.8)	19.2 (1.1)	25.5 (1.4)

TABLE III. EER (in percent) of MmLD (fusion of ECG and LCPD) and six comparison methods on GreenBit dataset. Standard deviations are in parentheses.

	Fusion of ECG and LCPD									No Fusion					
	Sum			product			Max			LCPD	BSIF	IQA	LBP	LPQ	WLD
Time (sec)	5	10	30	5	10	30	5	10	30	-					
Ecoflex	1.5 (1.0)	1.2 (0.8)	0.8 (0.7)	3.7 (1.4)	2.5 (0.8)	2.0 (0.6)	2.1 (0.8)	2.1 (0.9)	2.1 (1.0)	3.7 (0.5)	7.8 (1.3)	5.8 (1.6)	9.7 (1.5)	4.9 (0.8)	11.8 (2.1)
Gelatine	4.1 (1.4)	3.5 (0.9)	3.0 (1.1)	7.5 (1.1)	4.6 (1.0)	3.7 (0.9)	7.9 (1.4)	7.9 (1.3)	7.9 (1.4)	9.3 (1.3)	13.0 (2.1)	17.0 (2.4)	14.7 (2.1)	9.1 (1.2)	19.6 (2.3)
Latex	1.6 (0.8)	1.3 (0.8)	1.0 (0.8)	3.1 (1.4)	2.0 (1.0)	1.6 (0.5)	1.8 (0.7)	1.3 (0.7)	1.2 (0.7)	3.3 (0.7)	6.4 (1.5)	9.0 (1.8)	11.5 (1.8)	9.3 (1.4)	15.4 (1.9)
LiqEcoflex	3.9 (1.2)	3.2 (1.2)	2.7 (1.2)	8.2 (1.1)	4.9 (0.8)	3.8 (0.4)	8.0 (1.3)	8.1 (1.3)	8.0 (1.3)	8.7 (1.4)	13.1 (1.9)	15.8 (2.2)	13.3 (1.9)	10.6 (1.6)	24.2 (2.4)
RTV	1.7 (0.9)	1.3 (0.8)	1.0 (0.8)	5.3 (1.1)	3.2 (0.7)	2.6 (0.5)	2.6 (0.8)	2.4 (0.7)	2.5 (0.7)	4.2 (0.7)	5.9 (1.8)	8.4 (1.5)	11.5 (1.2)	6.1 (1.5)	18.1 (2.5)
WoodGlue	9.1 (1.4)	6.0 (1.7)	3.7 (0.8)	10.3 (0.7)	6.5 (1.3)	4.1 (0.9)	17.6 (2.3)	17.7 (2.3)	17.7 (2.3)	18.9 (2.4)	18.0 (2.3)	22.4 (2.7)	17.6 (2.3)	21.6 (2.0)	35.9 (2.9)
Average	**3.7** (0.7)	**2.7** (0.6)	**2.0** (0.5)	6.4 (0.5)	4.0 (0.5)	3.0 (0.4)	6.7 (0.6)	6.6 (0.6)	6.6 (0.6)	8.0 (0.5)	10.7 (0.9)	13.1 (1.1)	13.1 (0.9)	10.3 (0.7)	20.8 (1.3)

with that algorithm and get even better results. In addition, although in this study we investigate fusion of ECG and a software-based liveness detection method, a hardware-based method can also be used in the same way because the proposed system performs the fusion on score level. In addition to 12 participants listed in Table V, one complete fingerprint system is also submitted to LivDet2015 competition. Participants were provided with three spoof receipts and their submitted fingerprint system were tested on those spoof types as well as two unknown spoof materials. HTER of the submitted system was 8% as reported in [4] which is by far behind the proposed method.

In addition to the final report of LivDet2015 competition [4], recently organizers of LivDet series have presented a summary of LivDet competitions in [60]. They have selected 3 out of 12 algorithms submitted to LivDet2015 competition and reported rate of misclassified fake fingerprints (ferrfake) when rate of misclassified live fingerprints (ferrlive) is 1%. This represents the percent of spoof attacks that have been able to fool the system when only 1% of live attempts are mistakenly rejected. These results are reported from [60] in Table VI. For comparison, ferrfake of the MmLD (fusion of ECG and LCPD) is also presented. The average ferrfake of these methods at ferrlive=1% is about 45%. This indicates the poor performance of the state-of-the-art liveness detection methods that is also pointed out in [60]. It can be seen that the performance of the proposed method is by far better than the participants of the LivDet2015 competition. Positive and negative trials used in Table VI are the same as Table V explained before –i.e. 276 positive and 276 negative trials per experiment.

B. Multimodal Human Recognition (MmHR)

In this section, we investigate fusion of ECG and fingerprint for human recognition purpose. To this end, we use the fingerprint recognition software from NIST Biometric Image Software (NBIS 5.0.0) [61]. NBIS detects minutiae by an algorithm called MINDTCT and computes fingerprint matching

TABLE IV. EER (in percent) of MmLD (fusion of ECG and LCPD) and six comparison methods on Biometrika dataset. Standard deviations are in parentheses.

	Fusion of ECG and LCPD									No Fusion					
	Sum			product			Max			LCPD	BSIF	IQA	LBP	LPQ	WLD
Time (sec)	5	10	30	5	10	30	5	10	30	-					
Ecoflex	2.7 (1.2)	2.3 (1.1)	1.9 (1.1)	6.2 (1.4)	4.2 (1.1)	3.2 (1.1)	4.3 (1.5)	3.7 (1.4)	3.7 (1.0)	5.4 (1.2)	7.1 (1.1)	12.5 (1.8)	3.0 (1.6)	11.7 (2.1)	15.7 (2.1)
Gelatine	5.0 (1.7)	4.5 (1.7)	3.7 (1.4)	8.8 (1.5)	6.2 (1.2)	5.6 (1.0)	9.4 (1.6)	10.0 (1.6)	9.9 (1.5)	10.9 (2.0)	11.8 (1.7)	20.2 (2.2)	11.0 (2.4)	15.3 (1.9)	24.9 (2.4)
Latex	1.9 (1.0)	1.4 (0.8)	1.2 (0.8)	3.2 (1.1)	2.2 (0.8)	1.6 (0.7)	2.5 (0.9)	2.5 (0.9)	2.0 (0.8)	3.1 (1.6)	5.9 (1.3)	16.2 (2.0)	8.4 (1.6)	19.5 (2.3)	16.9 (2.1)
LiqEcoflex	4.3 (1.6)	3.8 (1.4)	3.0 (1.2)	8.3 (1.3)	5.5 (1.3)	4.9 (1.3)	7.6 (1.7)	7.7 (1.9)	7.4 (1.7)	9.5 (1.5)	12.6 (1.8)	25.5 (2.7)	10.9 (1.8)	27.2 (2.5)	28.4 (2.6)
RTV	4.7 (1.7)	4.0 (1.4)	3.3 (1.3)	8.6 (1.4)	5.8 (1.1)	4.8 (1.1)	8.3 (1.3)	8.3 (1.3)	8.2 (1.4)	10.2 (1.3)	8.7 (1.8)	15.6 (2.2)	8.4 (1.2)	24.1 (2.3)	26.7 (2.8)
WoodGlue	8.6 (1.8)	6.4 (1.5)	4.5 (1.4)	10.5 (1.0)	6.9 (1.6)	4.7 (1.3)	17.0 (2.2)	17.0 (2.2)	17.1 (2.2)	15.6 (2.1)	16.0 (1.9)	24.7 (2.3)	19.9 (2.4)	46.9 (2.8)	32.8 (2.6)
Average	**4.5** (0.4)	**3.7** (0.5)	**2.9** (0.5)	7.6 (0.8)	5.1 (0.6)	4.1 (0.6)	8.2 (0.5)	8.2 (0.6)	8.0 (0.5)	9.1 (0.7)	10.3 (0.8)	19.1 (1.1)	10.3 (0.9)	24.1 (0.9)	24.2 (1.3)

TABLE V. Comparison with LivDet2015 competition. HTER (in percent) is reported for all participants of LivDet2015 competition from [4]. EER (in percent) for the MmLD (fusion of ECG and LCPD) for different ECG lengths is also presented. To be consistent with LivDet2015 competition, only LiquidEcoflex and RTV are considered as unknown spoofs. Standard deviations are in parentheses.

		Dig.Persona	GreenBit	Biometrika	Average
LivDet2015 Competition [4]	COPILHA	24.2	30.6	32.7	29.1
	CSI	26.2	20.3	17.1	21.2
	CSI MM	26.4	14.4	11.2	17.3
	hbirkholz	13.7	11.0	7.1	10.6
	hectorn	19.3	12.0	15.0	15.4
	anonym	18.5	11.4	10.8	13.5
	Jinglian	15.1	7.7	5.3	9.3
	UFPE I	23.7	28.0	41.6	31.1
	UFPE II	24.7	18.6	32.7	25.3
	nogueira	7.1	5.5	7.2	6.6
	titanz	12.6	10.5	7.6	10.2
	unina	18.2	5.3	6.2	9.9
Proposed method	ECGFP-5	5.0 (1.5)	2.6 (0.9)	4.9 (1.0)	**4.2** (0.7)
	ECGFP-7.5	4.2 (1.3)	2.2 (0.9)	4.7 (1.1)	**3.7** (0.7)
	ECGFP-10	3.4 (1.0)	2.1 (0.8)	4.3 (0.9)	**3.3** (0.5)
	ECGFP-30	2.6 (0.8)	1.8 (0.8)	3.5 (0.7)	**2.6** (0.5)

TABLE VI. ferrfake (in percent) at ferrlive=1%. Results for 3 participants of LivDet2015 is reported from [60]. ferrfake of the MmLD (fusion of ECG and LCPD) for different ECG lengths is also presented. To be consistent with LivDet2015 competition, only LiquidEcoflex and RTV are considered as unknown spoofs. Standard deviations are in parentheses.

		Dig.Persona	GreenBit	Biometrika	Average
LivDet2015 Competition [60]	unina	52.2	53.2	19.8	41.7
	nogueira	25.3	23.4	20.4	23.0
	anonym	84.0	75.9	57.1	72.3
Proposed method	ECGFP-5	18.0 (1.1)	11.0 (0.7)	20.7 (1.4)	**16.6** (0.6)
	ECGFP-7.5	13.9 (1.1)	7.8 (0.9)	18.7 (1.1)	**13.4** (0.6)
	ECGFP-10	10.3 (1.0)	4.6 (0.9)	14.3 (1.3)	**9.7** (0.6)
	ECGFP-30	4.5 (0.9)	3.0 (0.7)	5.9 (0.9)	**4.5** (0.5)

scores using a matcher called Bozorth3. MINDTCT performs image binarization, minutiae detection, false minutiae removal, neighbor ridges counting and minutiae quality assessment, and generates a list consists of location, orientation, type and quality of the detected minutiae to be used by Bozorth.

TABLE VII. EER (in percent) of MmHR (fusion of ECG and Bozorth) for different fusion rules and different values of HC threshold. Standard deviations are in parentheses.

	Sum			product			Max			Bozorth (no fusion)
Time (sec)	7.5	10	25	7.5	10	25	7.5	10	25	
DigitalPersona	4.0 (0.4)	3.4 (0.3)	**2.1** (0.3)	6.2 (0.3)	4.3 (0.3)	2.8 (0.2)	4.1 (0.5)	3.8 (0.3)	2.5 (0.2)	8.6 (0.6)
GreenBit	2.6 (0.2)	2.2 (0.2)	**1.3** (0.2)	5.3 (0.3)	3.4 (0.2)	2.2 (0.2)	2.9 (0.2)	2.7 (0.2)	1.9 (0.2)	2.7 (0.4)
Biometrika	4.6 (0.4)	3.9 (0.3)	**2.6** (0.3)	7.3 (0.3)	5.3 (0.3)	3.8 (0.3)	5.3 (0.4)	4.7 (0.3)	3.3 (0.3)	15.1 (0.7)

Bozorth is a minutiae-based fingerprint matching method invariant to rotation and translation. It considers the location and orientation of the top 150 high quality minutiae and computes the matching score. EER of Bozorth on 3 different datasets is shown in the last column of Table VII.

We consider three different rules for fusion of ECG and fingerprint recognition, i.e. sum, product and maximum rules and results are presented in Table VII. Since the performance of Bozorth and ECG recognition modules are on average in the same range, we use equal weights for both traits in sum rule. Considering that we have 46 subjects for which there are 3 test session available, the number of positive and negative (zero-effort) trials in Table VII is respectively 46×3 and $46\times\ 45\times3$ for each spoof in each run.

Comparing the results of different scanners in Table VII, one may observe that while Bozorth gives an EER as good as 2.7% on GreenBit scanner, its performance degrades to 15.1% on Biometrika scanner. On the other side, performance of the multimodal system is less sensitive to the scanner, i.e. a better generalization on different scanners. We use the sum rule for fusion of ECG and Bozorth in the subsequent experiments due to its better performance compare with product and max rules.

VII. Automatic Updating of ECG and Fingerprint Templates

Automatically updating the ECG templates allows to model the intra-class variation of ECG over different sessions. So the biometric system can adapt to temporal variations of ECG signal across different sessions. If the system ensures the liveness and genuineness in a trial, then the provided ECG and fingerprint samples can be safely added to the corresponding templates. To prevent adaption using impostor samples, zero false acceptance rate for both liveness detection and human recognition tasks is necessary. This can be achieved in two different ways. First, the fusion score of MmLD and MmHR blocks, i.e. S_{MmLD} and S_{MmHR} can be compared against updating thresholds:

$(S_{MmLD} > T^U_{MmLD})\&(S_{MmHR} > T^U_{MmHR})$

where T^U_{MmLD} and T^U_{MmHR} are thresholds corresponding to zero false acceptance rate operating point of MmLD and MmHR blocks, respectively and & is logical AND operator. Those subjects that pass the MmLD block are fed to the MmHR block and those that pass that block too are selected for updating.

In the second approach, each of the ECG, FpLD and FpR blocks are set to zero false acceptance operating point. Considering that inside MmLD (MmHR) block, ECG and FpLD (FpR) are in parallel configuration, the following criterion can be used to achieve the overall zero false acceptance rate:

$((S_{ECG} > T^U_{ECG})|(S_{FpLD} > T^U_{FpLD}))\&((S_{ECG} > T^U_{ECG})|(S_{FpR} > T^U_{FpR}))$

where S_{ECG}, S_{FpLD} and S_{FpR} are the scores and T^U_{ECG}, T^U_{FpLD} and T^U_{FpR} are the updating thresholds corresponding to zero false acceptance rate operating point of ECG, FpLD and FpR blocks, respectively and | is logical OR operator. Both approaches ensure that fake samples as well as impostor attempts are rejected and will not contribute to template updating process. We use the second criterion because it is independent of the fusion rule inside MmLD and MmHR blocks. Queries that satisfy above criterion are selected for template updating.

Updating a multimodal template includes updating a fingerprint template and an ECG template. In order to update an ECG template, we add the new ECG samples to the previously existing training samples of that subject and re-train the SVM model. In order to update a fingerprint template, we add the new fingerprint sample to the set of previously existing samples of that subject. Since Bozorth is a matcher, if a fingerprint template contains more than one sample, we compute the match scores against each of the samples in the template and consider the average score as the matching score assigned to that trial.

Note that automatic template updating using a single biometric may not be effective because it tends to add only those samples that are very similar to the existing samples (i.e. lack of diversity). However, in the proposed system it makes sense to perform template updating because a multimodal system can potentially add more diverse samples to the templates. For example, if with high confidence fingerprint is live and matches the corresponding template, we can update both ECG and fingerprint templates. This implies that more diverse ECG samples can be added to the ECG template and the resulting ECG samples can better reflect the within class variations of the corresponding subject. Likewise, if with high confidence ECG matches the corresponding template, both ECG and the fingerprint templates are updated. This implies that the new fingerprint sample can potentially improve the diversity for the corresponding fingerprint template to better reflect the within class variations. We set the minimum length of the first test session to 12 heartbeats, i.e. about 10 seconds. This allows a more effective template updating specially when HC threshold is set to a small value. However, the length of the subsequent test sessions, i.e. second and third sessions, are determined solely by HC criterion and the 30-second upper limit.

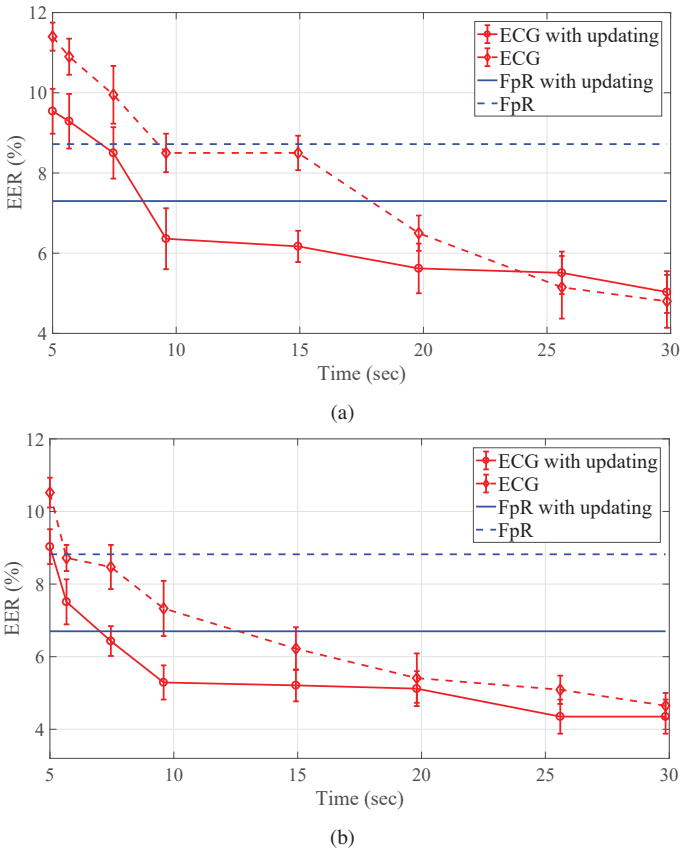

Fig. 6. EER of ECG and fingerprint recognition with and without template updating: (a) 2nd test session. (b) 3rd test session. Results are averaged over 3 datasets. Results with (without) template updating are shown with solid (dashed) lines. Bars represent standard deviations. In the 2nd test session, standard deviation for fingerprint recognition with (without) template updating is 0.4 (0.6). Likewise, in the 3rd test session, standard deviation for fingerprint recognition with (without) template updating is 0.5 (0.7).

We consider 2 types of attack: zero-effort attack and spoof attack. In the zero-effort attack, intruder provides his own fingerprint sample but claims to be someone else, hence the term zero-effort. In the spoof attack, intruder provides a fake fingerprint sample of the claimed identity. In both cases, ECG signal is also recorded. For each of 46 chimeric subjects there are 3 fake and 3 live samples corresponding to 3 rounds of testing. Therefore, in each testing round there are 46 positive, 46 spoof and 46×45 zero-effort trials. However, for computational simplicity we only consider 500 zero-effort

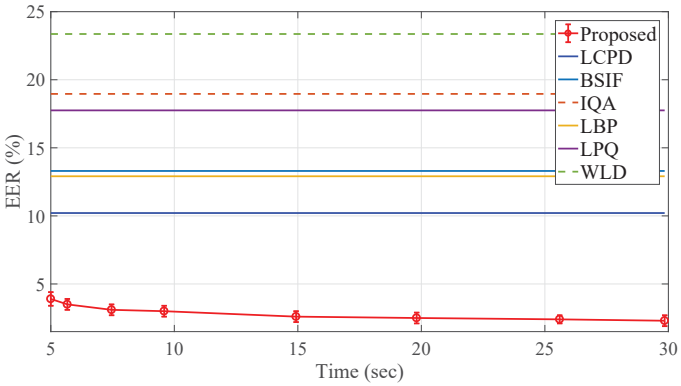

Fig. 7. EER of the MmLD (fusion of ECG and LCPD) with template updating is compared with conventional liveness detection methods in the 3rd test session. Bars represent standard deviations. Standard deviations of LCPD, BSIF, IQA, LBP, LPQ and WLD are respectively 0.5, 0.4, 0.6, 0.7, 0.5 and 0.8.

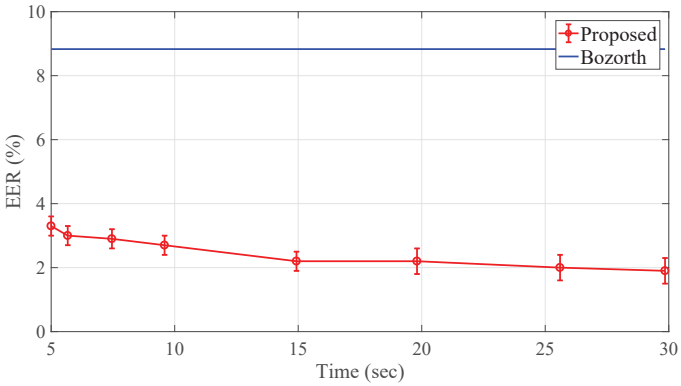

Fig. 8. EER of MmHR (fusion of ECG and Bozorth) with template updating is compared with Bozorth in the 3rd test session. Bars represent standard deviations. Standard deviation of Bozorth is 0.5.

trials.

At the end of the first test session, on average about 79% of chimeric subjects are selected for template updating. Likewise, at the end of the second test session about 76% are selected. In Figure 6, the effect of template updating on ECG and fingerprint recognition is shown for 2nd and 3rd test sessions. It can be seen that the performance of both ECG recognition and fingerprint recognition using Bozorth improves due to template updating. The recognition rates reported in Figure 6 are based on 46 positive and 500 negative (zero-effort) trials for each type of spoof in each dataset. The results for different types of spoofs and datasets are averaged and the entire process is repeated 50 times. The average and standard deviation are reported.

Figure 7 shows EER of the MmLD block in the third test session after performing automatic temple updating in the first and second test sessions. For comparison, EER of other methods are also shown as horizontal lines. It can be seen that the proposed approach performs significantly better than comparison methods for a wide range of ECG lengths. In computing the liveness detection results in Figure 7, 46 positive and 46 negative (i.e. spoof) trials have been used for each type of spoof in each dataset. The results for different types of spoofs and datasets are averaged and the entire process is repeated 50 times and average and standard deviation are reported.

Figure 8 compares the recognition rate of the proposed system with Bozorth in the third test session. It can be seen that the proposed method performs significantly better than Bozorth. In computing the recognition rates in Figure 8, 46 positive and 500 negative (i.e. zero-effort) trials has been used for each type of spoof in each dataset. The results for different types of spoofs and datasets are averaged and the entire process is repeated 50 times as explained before and average and standard deviation are reported. Note that the main focus of this study is liveness detection rather than recognition rate. We do not emphasis on fingerprint recognition performance and use Bozorth only as an average commercial fingerprint recognition system. If we had a more accurate and complicated fingerprint recognition algorithm, the results after fusion with ECG could be even better. Therefore, we have considered a more challenging case. This is important because it demonstrates that two average low cost commercial fingerprint and ECG biometric systems with mediocre individual performances can be combined to provide significantly better performance. Note that our ECG signals are recorded using Vernier [57] which is a low cost nonprofessional commercially available hardware designed for educational use. Nevertheless, we have demonstrated promising results under such practical setup. One may get even better results by using a more complicated fingerprint recognition method or more sophisticated ECG acquisition hardware. But for the sake better generalization on real-world scenarios, we did not seek that direction.

VIII. Computational Costs

Experiments are performed in MATLAB on a desktop with an Intel core i7-3770 CPU and 16GB RAM. CPU-time for computing the match score between two fingerprints is about 380 msec. The CPU time for feature extraction with HC threshold of 0.92 (i.e. 10 seconds of ECG on average) is about 630 msec. Extracting LCPD features takes about 5 seconds per image. The time for testing against an SVM model is very small and can be neglected. Training one SVM for ECG recognition (HC threshold of 0.92) takes about 300 msec. Training one SVM for fingerprint liveness detection takes about 6 seconds. The feature selection part is very fast and takes less than a second.

IX. Conclusion

ECG can be recorded from fingertips. Therefore, fingerprint is the natural choice to be fused with ECG. On the other side fingerprint is vulnerable to spoof attacks and ECG has inherent liveness detection. This paper presented a unified approach for fusion of fingerprint and ECG that fills the gap between these two sides. To get the most out of ECG, the proposed system fuses ECG with a conventional fingerprint liveness detection method for a better liveness detection performance, and also fuses it with a fingerprint recognition method for a

better recognition rate. Extensive experiments on LiVDet2015 database with 6 different spoofs, 3 different scanners, 6 liveness detection algorithms and 3 different fusion methods have demonstrated that ECG significantly improves the performance in both tasks, i.e. liveness detection and human recognition. In addition, the proposed system automatically performs template updating, so that the performance of biometric system can be maintained in long term without manual re-training or re-enrollment. We also compared the proposed system with the results reported in LivDet2015 competition which is the latest competition in this field and demonstrated that it outperforms all participants of that competition.

ACKNOWLEDGMENT

The authors would like to thank Dr. Nasir Memon for his valuable suggestions. This research was supported by Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

[1] G. L. Marcialis, A. Lewicke, B. Tan, P. Coli, D. Grimberg, A. Congiu, A. Tidu, F. Roli, and S. Schuckers, "First international fingerprint liveness detection competitionlivdet 2009," in *Image Analysis and Processing–ICIAP 2009*. Springer, 2009, pp. 12–23.

[2] D. Yambay, L. Ghiani, P. Denti, G. L. Marcialis, F. Roli, and S. Schuckers, "Livdet 2011fingerprint liveness detection competition 2011," in *Biometrics (ICB), 2012 5th IAPR International Conference on*. IEEE, 2012, pp. 208–215.

[3] L. Ghiani, D. Yambay, V. Mura, S. Tocco, G. L. Marcialis, F. Roli, and S. Schuckcrs, "Livdet 2013 fingerprint liveness detection competition 2013," in *Biometrics (ICB), 2013 International Conference on*. IEEE, 2013, pp. 1–6.

[4] V. Mura, L. Ghiani, G. L. Marcialis, F. Roli, D. A. Yambay, and S. A. Schuckers, "Livdet 2015 fingerprint liveness detection competition 2015," in *Biometrics Theory, Applications and Systems (BTAS), 2015 IEEE 7th International Conference on*. IEEE, 2015, pp. 1–6.

[5] C. Sousedik and C. Busch, "Presentation attack detection methods for fingerprint recognition systems: a survey," *IET Biometrics*, vol. 3, no. 4, pp. 219–233, 2014.

[6] S.-C. Fang and H.-L. Chan, "Human identification by quantifying similarity and dissimilarity in electrocardiogram phase space," *Pattern Recognition*, vol. 42, no. 9, pp. 1824–1831, 2009.

[7] J. M. Irvine, S. A. Israel, W. Todd Scruggs, and W. J. Worek, "eigenpulse: Robust human identification from cardiovascular function," *Pattern Recognition*, vol. 41, no. 11, pp. 3427–3435, 2008.

[8] S.-C. Fang and H.-L. Chan, "Qrs detection-free electrocardiogram biometrics in the reconstructed phase space," *Pattern Recognition Letters*, vol. 34, no. 5, pp. 595–602, 2013.

[9] N. Karimian, Z. Guo, M. Tehranipoor, and D. Forte, "Highly reliable key generation from electrocardiogram ECG." *IEEE transactions on biomedical engineering*, 2016.

[10] A. D. Chan, M. M. Hamdy, A. Badre, and V. Badee, "Wavelet distance measure for person identification using electrocardiograms," *Instrumentation and Measurement, IEEE Transactions on*, vol. 57, no. 2, pp. 248–253, 2008.

[11] C. Ye, M. T. Coimbra, and B. V. Kumar, "Investigation of human identification using two-lead electrocardiogram ECG signals," in *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*. IEEE, 2010, pp. 1–8.

[12] S. Z. Fatemian, F. Agrafioti, and D. Hatzinakos, "Heartid: Cardiac biometric recognition," in *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*. IEEE, 2010, pp. 1–5.

[13] Y. Su, J. Feng, and J. Zhou, "Fingerprint indexing with pose constraint," *Pattern Recognition*, vol. 54, pp. 1–13, 2016.

[14] H.-W. Jung and J.-H. Lee, "Noisy and incomplete fingerprint classification using local ridge distribution models," *Pattern Recognition*, vol. 48, no. 2, pp. 473–484, 2015.

[15] Y. N. Singh, S. K. Singh, and P. Gupta, "Fusion of electrocardiogram with unobtrusive biometrics: An efficient individual authentication system," *Pattern Recognition Letters*, vol. 33, no. 14, pp. 1932–1941, 2012.

[16] C. X. Zhao, T. Wysocki, F. Agrafioti, and D. Hatzinakos, "Securing handheld devices and fingerprint readers with ECG biometrics," in *Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on*. IEEE, 2012, pp. 150–155.

[17] J. Wang, M. She, S. Nahavandi, and A. Kouzani, "Human identification from ECG signals via sparse representation of local segments," *Signal Processing Letters, IEEE*, vol. 20, no. 10, pp. 937–940, 2013.

[18] F. Agrafioti and D. Hatzinakos, "ECG based recognition using second order statistics," in *Communication Networks and Services Research Conference, 2008. CNSR 2008. 6th Annual*. IEEE, 2008, pp. 82–87.

[19] A. Kaveh and W. Chung, "Temporal and spectral features of single lead ECG for human identification," in *Biometric Measurements and Systems for Security and Medical Applications (BIOMS), 2013 IEEE Workshop on*. IEEE, 2013, pp. 17–21.

[20] M. Tantawi, K. Revett, A. Salem, and M. F. Tolba, "Fiducial feature reduction analysis for electrocardiogram ECG based biometric recognition," *Journal of Intelligent Information Systems*, vol. 40, no. 1, pp. 17–39, 2013.

[21] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold, "ECG to identify individuals," *Pattern recognition*, vol. 38, no. 1, pp. 133–142, 2005.

[22] I. Odinaka, P.-H. Lai, A. D. Kaplan, J. A. O'Sullivan, E. J. Sirevaag, S. D. Kristjansson, A. K. Sheffield, and J. W. Rohrbaugh, "ECG biometrics: A robust short-time frequency analysis," in *Information Forensics and Security (WIFS), 2010 IEEE International Workshop on*. IEEE, 2010, pp. 1–6.

[23] A. Lourenço, H. Silva, and A. Fred, "Unveiling the biometric potential of finger-based ECG signals," *Computational intelligence and neuroscience*, vol. 2011, p. 5, 2011.

[24] H. P. Da Silva, A. Fred, A. Lourenco, and A. K. Jain, "Finger ECG signal for user authentication: Usability and performance," in *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*. IEEE, 2013, pp. 1–8.

[25] I. Odinaka, P.-H. Lai, A. D. Kaplan, J. A. O'Sullivan, E. J. Sirevaag, and J. W. Rohrbaugh, "ECG biometric recognition: A comparative analysis," *Information Forensics and Security, IEEE Transactions on*, vol. 7, no. 6, pp. 1812–1824, 2012.

[26] M. Hejazi, S. Al-Haddad, Y. P. Singh, S. J. Hashim, and A. F. A. Aziz, "ECG biometric authentication based on non-fiducial approach using kernel methods," *Digital Signal Processing*, vol. 52, pp. 72–86, 2016.

[27] Y. N. Singh, "Human recognition using fisher s discriminant analysis of heartbeat interval features and ECG morphology," *Neurocomputing*, vol. 167, pp. 322–335, 2015.

[28] E. J. d. S. Luz, D. Menotti, and W. R. Schwartz, "Evaluating the use of ECG signal in low frequencies as a biometry," *Expert Systems with Applications*, vol. 41, no. 5, pp. 2309–2315, 2014.

[29] K. A. Sidek, I. Khalil, and H. F. Jelinek, "ECG biometric with abnormal cardiac conditions in remote monitoring system," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 11, pp. 1498–1509, 2014.

[30] I. Odinaka, J. A. OSullivan, E. J. Sirevaag, and J. W. Rohrbaugh, "Cardiovascular biometrics: combining mechanical and electrical signals," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 1, pp. 16–27, 2015.

[31] S. Wahabi, S. Pouryayevali, S. Hari, and D. Hatzinakos, "On evaluating ECG biometric systems: Session-dependence and body posture," *Information Forensics and Security, IEEE Transactions on*, vol. 9, no. 11, pp. 2002–2013, 2014.

[32] H. Mehrotra, R. Singh, M. Vatsa, and B. Majhi, "Incremental granular relevance vector machine: A case study in multimodal biometrics," *Pattern Recognition*, vol. 56, pp. 63–76, 2016.

[33] S. Bharadwaj, H. S. Bhatt, R. Singh, M. Vatsa, and A. Noore, "Qfuse: Online learning framework for adaptive biometric system," *Pattern Recognition*, vol. 48, no. 11, pp. 3428–3439, 2015.

[34] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000 (June 13), circulation Electronic Pages: http://circ.ahajournals.org/cgi/content/full/101/23/e215 PMID:1085218; doi: 10.1161/01.CIR.101.23.e215.

[35] Y. N. Singh and P. Gupta, "Biometrics method for human identification using electrocardiogram," in *International Conference on Biometrics*. Springer, 2009, pp. 1270–1279.

[36] National institude of standards and technology. biometric score set - release 1 (bssr1). http://www.nist.gov/itl/iad/ig/biometricscores.cfm. Last accessed 21 September 2016.

[37] E. Marasco, Y. Ding, and A. Ross, "Combining match scores with liveness values in a fingerprint verification system," in *Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on*. IEEE, 2012, pp. 418–425.

[38] R. Derakhshani, S. A. Schuckers, L. A. Hornak, and L. O'Gorman, "Determination of vitality from a non-invasive biomedical measurement for use in fingerprint scanners," *Pattern recognition*, vol. 36, no. 2, pp. 383–396, 2003.

[39] S. T. Parthasaradhi, R. Derakhshani, L. A. Hornak, and S. A. Schuckers, "Time-series detection of perspiration as a liveness test in fingerprint devices," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 35, no. 3, pp. 335–343, 2005.

[40] G. L. Marcialis, F. Roli, and A. Tidu, "Analysis of fingerprint pores for vitality detection," in *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, 2010, pp. 1289–1292.

[41] J. Galbally, S. Marcel, and J. Fierrez, "Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition," *IEEE transactions on image processing*, vol. 23, no. 2, pp. 710–724, 2014.

[42] S. B. Nikam and S. Agarwal, "Texture and wavelet-based spoof fingerprint detection for fingerprint biometric systems," in *2008 First International Conference on Emerging Trends in Engineering and Technology*. IEEE, 2008, pp. 675–680.

[43] Z. He, Z. Sun, T. Tan, and Z. Wei, "Efficient iris spoof detection via boosted local binary patterns," in *International Conference on Biometrics*. Springer, 2009, pp. 1080–1090.

[44] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[45] V. Ojansivu, E. Rahtu, and J. Heikkilä, "Rotation invariant blur insensitive texture analysis using local phase quantization," in *International Conference on Pattern Recognition*, vol. 4, 2008.

[46] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikainen, X. Chen, and W. Gao, "Wld: A robust local image descriptor," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 9, pp. 1705–1720, 2010.

[47] L. Ghiani, A. Hadid, G. L. Marcialis, and F. Roli, "Fingerprint liveness detection using binarized statistical image features," in *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*. IEEE, 2013, pp. 1–6.

[48] D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva, "Local contrast phase descriptor for fingerprint liveness detection," *Pattern Recognition*, vol. 48, no. 4, pp. 1050–1058, 2015.

[49] E. Marasco, P. Wild, and B. Cukic, "Robust and interoperable fingerprint spoof detection via convolutional neural networks," in *Technologies for Homeland Security (HST), 2016 IEEE Symposium on*. IEEE, 2016, pp. 1–6.

[50] R. F. Nogueira, R. de Alencar Lotufo, and R. C. Machado, "Fingerprint liveness detection using convolutional neural networks," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 6, pp. 1206–1213, 2016.

[51] N. Poh, R. Wong, and G.-L. Marcialis, "Toward an attack-sensitive tamper-resistant biometric recognition with a symmetric matcher: a fingerprint case study," in *2014 IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM)*. IEEE, 2014, pp. 175–180.

[52] P. Wild, P. Radu, L. Chen, and J. Ferryman, "Robust multimodal face and fingerprint fusion in the presence of spoofing attacks," *Pattern Recognition*, vol. 50, pp. 17–25, 2016.

[53] D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva, "An investigation of local descriptors for biometric spoofing detection," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 849–863, 2015.

[54] E. Marasco and A. Ross, "A survey on antispoofing schemes for fingerprint recognition systems," *ACM Computing Surveys (CSUR)*, vol. 47, no. 2, p. 28, 2015.

[55] A. Rattani, E. Granger, and F. Roli, *Adaptive Biometric Systems*. Springer, 2015.

[56] Medical biometric databases. http://www.comm.utoronto.ca/~biometrics/databases. Last accessed 21 September 2016.

[57] Vernier ekg sensor. www.vernier.com. Last accessed 21 September 2016.

[58] M. Komeili, N. Armanfard, D. Hatzinakos, and A. Venetsanopoulos, "Feature selection from multisession electrocardiogram signals for identity verification," in *Electrical and Computer Engineering (CCECE), 2015 IEEE 28th Canadian Conference on*. IEEE, 2015, pp. 603–608.

[59] A. Lourenço, H. Silva, C. Carreiras *et al.*, "Outlier detection in non-intrusive ECG biometric system," in *International Conference Image Analysis and Recognition*. Springer, 2013, pp. 43–52.

[60] L. Ghiani, D. A. Yambay, V. Mura, G. L. Marcialis, F. Roli, and S. A. Schuckers, "Review of the fingerprint liveness detection (livdet) competition series: 2009 to 2015," *Image and Vision Computing*, 2016.

[61] C. I. Watson, M. D. Garris, E. Tabassi, C. L. Wilson, R. M. Mccabe, S. Janet, and K. Ko, "User's guide to nist biometric image software (nbis)," 2007.