Photoplethysmogram Based Biometric Identification Incorporating Different Age and Gender Group

Siti Nurfarah Ain Mohd Azam, Khairul Azami Sidek and Ahmad Fadzil Ismail
Department of Electrical and Computer Engineering, Faculty of Engineering, International Islamic University Malaysia, 53100 Jalan Gombak, Kuala Lumpur, Malaysia. azami@iium.edu.my

Abstract—Biometric is the authentication and identification of a person by measuring or estimating their physiological characteristics. First generation biometric such as fingerprint, signature and voice have drawback and easily can be duplicated which lead to serious identity theft crime. Therefore, second generation of biometric was introduced by using bio-signal. This study evaluates the possibility of applying PPG as biometric identification system incorporating different age, gender group, and time variability. A total of 36 subjects were involved in this study consists of 18 males and 18 females for age difference and gender analysis. The PPG signals were taken in resting state by using pulse oximeter. The PPG signal was differentiated twice in order to form APG signal. These signals then undergo preprocessing and the segmentation process was done by using MATLAB. The highest peaks from the signal was used as reference point to determine the appropriate distance for one cycle of both signal. Then, the signals were classified by four commonly used classifiers which are Bayes Network, Naïve Bayes, Multilayer Perceptron, and Radial Basis Function. The outcome from this study suggested the accuracy up to 100% for different age group, 91.11% for female subjects and 95% for male subjects.

Index Terms—Acceleraton Plethysmogram; Biometric; Bio-Signal; Photoplethysmogram.

I. INTRODUCTION
It is very crucial and important today that every person’s identity and privacy is being protected. The reliability of traditional security system have drawbacks and limitations. Therefore, biometric recognition system was introduced. Traditional authentication system were based on user’s knowledge, such as passwords or personal identification number (PIN), that can be forgotten, smartcards and cardkey can be lost and misplaced which indicates the weakness of the traditional systems.

On the other hand, biometric can provide solution for better security system and has been recognize as one of the most reliable technologies for future human identification and verification [15]. It is believed that biometric can solve many of the security issues and have better potential in replacing the traditional security methods.

Biometrics is the authentication and identification of a person by measuring or estimating their physiological characteristics [13]. According to International Organization for Standardization (ISO), biometric is defined as “the automated recognition of individual based on their behavioral and biological characteristics” [9].

Numerous biometric measures have been studied for identification purposes, however, these method are not the same level of system complexity, cost and accuracy. Thus, this study will propose the use of PPG signals as biometric identification system incorporating different age and gender group. PPG signals are low in deployment cost, easy to use, smaller in size and conveniently can be used to various parts of human body such as finger, ear lobe, wrist or arm [15].

Photoplethysmogram (PPG) was previously implemented to measure the oxygen saturation, blood pressure, cardiac output, and for evaluating autonomic functions. It is a promising technology due to its simplicity, low cost and non-invasiveness. In recent years, PPG signals have been used for biometric recognition. However, to the best of our knowledge, there have been little research on the implementation of PPG signals as a biometric recognition system incorporating different age and gender group. The feasibility of applying PPG signals as a biological discriminant has been preliminary studied. However, the underlying issues that governs a practical biometric system have not been properly addressed. The objective of this work is to increase user acceptability of PPG based biometric identification by incorporating gender and age variability. Previous works have been more focused on person identification using normal subjects without considering conditions which could promote fluctuation of PPG signals. Therefore, in this study, we will investigate the possibility of categorizing individual built upon their PPG morphological signal. This technique will improve the current identification system by providing a complement which will be able to reduce cases involving identity crime.

A. Photoplethysmogram (PPG)
A photoplethysmograph is a device to optically obtain a volumetric measurement of an organ. PPG signals is obtained using pulse oximeters to monitor the blood pressure which are non-invasive devices. Figure 1 shows how the PPG is measured using pulse oximeter and Easy Pulse sensor and Figure 2 shows the acquired signal [12].

The heart is pumping the blood to all areas of the human body by arteries. Blood pressure is pushed against the walls of the arteries. Systolic pressure is where the maximum blood pressure occurs when the heart is pumping in and diastolic pressure is at the lowest blood pressure when the heart is resting. Figure 3 shows the location of the systolic and diastolic peaks in the PPG signals. The signals imitate the fluctuations in blood volume which is triggered by blood vessel expansion and contraction [11].
Figure 1: Fingertip is attached with pulse oximeter to monitor the blood pressure [1]

Figure 2: The acquired PPG signal.

Figure 3: Systolic and diastolic peaks of PPG signals [2]

APG is the second derivative of PPG signal as shown in Figure 4 and it is the acceleration of blood which occurs inside the finger.

Figure 4: The APG signal waveform [2]

The advantage of APG as compared to PPG signal is that it consists of 5 peaks which could increase the accuracy of biometric identification system. APG signal consists of four systolic and one diastolic waves. The waveforms are the a-wave (early systolic wave), b-wave (early systolic negative wave), c-wave (late systolic re-increasing wave), d-wave (late systolic re-decreasing wave) and e-wave (early diastolic positive wave). Every wave position indicates the closure of aortic valve and blood flow which can be used to define and observe heart function [3].

The remaining sections in this paper are structured as follows; the next section review the related biometric identification system based on PPG. Later, Section III, elaborates more on the technique of the study which comprises of data collection, pre-processing, segmentation and classification mechanism. In Section IV, the performance of the proposed system is discussed. Lastly, in Section V, the study is concluded based on the experimentation and outcomes from the earlier section.

II. LITERATURE REVIEW

Current biometric identification system which mostly employs the use of fingerprint scan have the possibility to be impersonated. Singh et al. in [16] proposed the possibility of using PPG as a secondary identification feature to be used with the fingerprint scanner. The morphological signal of PPG differs from one subject to another however depends on the location of the pulse oximeter when the waveform is recorded. Therefore, the technique of acquiring PPG signals need to be standardized. The data is measured by taking the reading of each finger profiles with respect to other finger profiles of the same and different subject. The study found out that, the correlation of the pulse profile of the same subject with the same finger is always higher when compared to the same finger with different subject. Besides that, it is observed the autocorrelation of the same subject with the same finger is mostly higher when compared with the same subject with different finger.

Many researchers recommend the use of PPG signals together with other biometric components to enhance the security of the system. Unseen internal organs seem to be more reliable than visible body parts as the vulnerability and the possibility of falsity become the common weakness of conventional biometric. Therefore, Salanke et al., in [11] also proposed the use of PPG signals as the complementary feature to current identification system. The PPG signals were recorded with 37 MHz sampling frequency. The experiment was conducted with 9 healthy subjects with the average age of 34. For each subject, 8 samples were acquired where 4 were in relaxed state and the remaining 4 with stressed condition. The study aims to identify the changes in PPG signals in different states. The results show that, all subjects produced identical waveforms during relaxed conditions but when distressed, the waveform generated is not identical due to motion artifact. Mahalanobis distance is used to identify the subjects by generating a matching score by the selected features resulting in the smallest distance which was considered as a match. However, the study only focuses on healthy subjects which at some conditions, this technique may not be applicable to people with heart abnormalities.

One of the obstacles faced when recording the PPG signals is that there are chances that the signals is being contaminated by motion artifact due to the movement of the finger connected to the pulse oximeter and surrounding light
variation because of the environment. Thus, in another study, Salanke et al., in [17] suggests the method to extract features in order to offer better differentiation among subjects by applying Fourier series analysis and Semi Discrete Decomposition. The normal PPG signal was used for the training purpose and the PPG with motion artifacts were used for the testing samples after going through pre-processing. Matching score was generated by the selected feature then Euclidean’s distance was estimated and compared with the set of samples stored. The results found out that, after applying Fourier series analysis and Semi Discrete Decomposition, it gives better outcome in classifying individuals. However, the study was limited to only healthy subjects and does not state the effect of this method for different category of ages. The shape and contour of PPG signal is too simple and identify the phase change thus make it difficult to analyze. Therefore, Sidek et al., [3] suggests an alternative method that will assists the original PPG signal analysis which is its second derivative also known as APG signal. APG signal contain far more information compared to PPG signal as APG has 3 peaks more than PPG does. The PPG data were taken from public online database with 10 subjects involved. The duration of each signals are 10 seconds with sampling frequency of 125 Hz. The signals were then differentiated in order to obtain APG signal. The available noises from the data then undergo pre-processing and were filtered out by using Butterworth low pass filter. The signals after that segmented by identifying the highest peaks. The same amount of points were selected from the peak to the left and the right of reference point to complete one cycle of APG signal. Next, the signals were normalized. Later, the data were classified by using commonly used classifier Bayes Network and K-Nearest Neighbor (KNN). The study discovered that, when compared with PPG signal, APG gave higher accuracy. However, there was no other parameter and variable used for this study which might be the factor that will affect the performance accuracy. In the past years, there have been many different types of biomedical signals introduced as identification system. It is proven that, these signals have the ability to identify individuals through different mechanisms. APG is believed to offer better accuracy as it provides more distinctive information which can be applied for biometric identification system. Thus, in other study by Sidek et al., in [10], a Cardioid based graph method by applying APG signal for biometric identification purposes was proposed. The data were collected from MIMIC II Waveform, Version 3, Part 2 and 3, an online public repository available in Physionet [4]. Each recording have 10 seconds of PPG signal with the sampling rate of 125 Hz. In order to obtain APG signal, the PPG is differentiated twice and then the signals were pre-processed by using low pass filter to reduce the noise. The APG is then segmented by identifying one complete cycle of the signal. Then, the Cardioid loop was formed. The center and the Euclidean distance of the Cardioid were then calculated. For classification, the data were classified by using Multilayer Perceptron (MLP) and Naïve Bayes. The results showed that, APG gave higher percentage accuracy when compared with PPG signal thus conclude that APG is better than PPG for biometric identification purposes. However, this study did not consider any external factor such as age of the subjects, health conditions, and gender.

Nevertheless, most of the related works only focuses on how to obtain high accuracy of the system but neglecting some important variables that need to be considered. Relative to these findings, different age group effect on the robustness, reliability and accuracy of subject recognition is a substantially fundamental issue for PPG biometric as applications are becoming more diversified. Thus, the robustness of PPG samples towards different age and gender group is a pertinent area which is a uniquely under-researched.

III. METHODOLOGY

Figure-5 summarizes the proposed identification system which consists of the data collection, pre-processing, segmentation, signal differentiation, and classification stages. Each stage will be elaborated further in the next sub-sections.

![Proposed Identification System](image)

A. Signal Acquisition

A total of 36 subjects participated in the experiment involving 18 males and 18 females with different age divided into 9 groups as shown in Table 1. The PPG signals were measured by using Easy Pulse sensor. PPG signals of 5 different cycles were collected from each subject with sampling rate of 360 Hz. Each sample of the PPG signal was recorded for 15 seconds.

<table>
<thead>
<tr>
<th>Group</th>
<th>Age range (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>6-10</td>
</tr>
<tr>
<td>B</td>
<td>11-15</td>
</tr>
<tr>
<td>C</td>
<td>16-20</td>
</tr>
<tr>
<td>D</td>
<td>21-25</td>
</tr>
<tr>
<td>E</td>
<td>26-30</td>
</tr>
<tr>
<td>F</td>
<td>31-35</td>
</tr>
<tr>
<td>G</td>
<td>36-40</td>
</tr>
<tr>
<td>H</td>
<td>41-45</td>
</tr>
<tr>
<td>I</td>
<td>46-50</td>
</tr>
</tbody>
</table>

B. Pre-processing

The raw PPG signals contain unwanted signals such as noise and baseline wandering which is produced from the environment factor during the PPG data collection procedure. For the pre-processing stage, Butterworth filter is used to filter out these unwanted signals to obtain as accurate data as possible. The filter is a form of signal processing technique which intends to gain as flat a frequency response as possible in the pass band and also denoted to as a maximally flat magnitude filter. The choice of Butterworth as a filter is because the advantage is it results in smooth, monotonically decreasing frequency response.

C. Feature Extraction

The segmentation process of the PPG signal is done by using MATLAB. One complete cycle of PPG signal is segmented by part through using specific algorithm. The procedure begins by detecting the systolic peak and making it the pivot since it signifies to the highest and most clear peak in a PPG signal. The highest peak of the signal is set as...
the reference point as illustrated in Figure 7. Based on the reference point, 40 points were selected to the left and 60 points to the right of the waveform.

![Reference point](image)

**Figure 6:** Segmentation of PPG signal.

The steps is repeated in different time instances to collect more PPG signals for every subject that would characterize enrolment and recognition datasets. The points taken will be used for classification later. This procedure is the same for the APG signal. When the APG signal is acquired, the highest peak becomes the reference point and from the point, 20 points were selected to the left and 40 points to the right of the waveform.

**D. Signal Differentiation**

Second order differentiation is performed after acquiring PPG data to produce the APG signal. In the preliminary stage, the PPG is denoted as \( x \) is differentiated to obtain the differentiated values denoted as \( y \) shown in Equation 1.

\[
y(m) = x(n-1) - x(n)
\]  

(1)

where \( m = 1, 2, 3, \ldots, (N-1) \) and \( y(m) \) is the differentiated PPG signals.

\[
\frac{d^2 y}{dx^2} = \frac{\partial}{\partial x} \left( \frac{\partial y}{\partial x} \right)
\]  

(2)

APG signal is produced when the second order derivative is performed as shown in Equation (2).

**E. Classification**

For classification purpose, four different types of classifier will be used in this study that are Bayes Network, Naïve Bayes, Multilayer Perceptron and Radial Basis Function to evaluate the accuracy of proposed system. The signals will be classified by using open source software named Weka [5].

**IV. EXPERIMENTATION AND RESULTS**

Figures 7 to 21 illustrate the steps of the feature extraction stage. Figure 7 to 9 illustrate the PPG and APG signals for three different subjects with various age group and dissimilar gender. Subject 14 (Group D) and Subject 25 (Group G) are males with age 21 and 36 years old whereas Subject 19 (Group E) is a female with age of 26 years old. Figures 7(a), 8(a), and 9(a) show the raw PPG signal of Subject 14, Subject 25 and Subject 19.

![PPG and APG signals](image)

**Figure 7:** (a) Raw PPG signal for Subject 14, (b) Filtered PPG signal for Subject 14, and (c) APG signal for Subject 14.

![PPG and APG signals](image)

**Figure 8:** (a) Raw PPG signal for Subject 25, (b) Filtered PPG signal for Subject 25, and (c) APG signal for Subject 25.

![PPG and APG signals](image)

**Figure 9:** (a) Raw PPG signal for Subject 19, (b) Filtered PPG signal for Subject 19, and (c) APG signal for Subject 19.

Figure 7(b), 8(b) and (b) show the filtered PPG signal of each subjects. The purpose of filtering is to obtain as flat magnitude frequency as possible and also to remove unwanted noise that might affect the preciseness of the experiment. Next, after the filtering process was done, the PPG signal is differentiated twice in order to produce APG signal as shown in Figure 7(c), 8(c) and 9(c). After that, the reference point (marked by the red circles) for the PPG and APG signal were determine by identifying the highest peak from each signal. This step is shown in Figure 10, 11, and 12 for PPG signal and Figure 13, 14 and 15 for APG signal.

![Reference point](image)

**Figure 10:** Reference point of PPG signal for Subject 14.
Each of the signals was segmented together as described in Figure 16 to 21. These signals were overlapped with each other in order to observe the pattern.
The PPG and APG signals were separated into two categories which are age and gender. Figure 22 and 23 show the classification results for PPG and APG signal based on age group. Figure 24 and 25 illustrate the classification outcomes for PPG and APG waveform subject to gender.

Figure 22: Classification accuracy of PPG signal based on age group.

Figure 23: Classification accuracy of APG signal based on age group.

Figure 22 describes the classification accuracy of PPG signal based on different age group by applying BN, NB, MLP and RBF network as classifiers. It can be observed that, the accuracy of groups A, B, G and H are slightly lower when compared to groups C, D, E, F and I. Groups C, D, F, and I gave the accuracy as high as 100% for all four type of classifiers.

Groups A and B were individuals with age of 6 to 15 years old where they are children and young teenagers. The results show that, the accuracy of PPG signal was inconsistent. However, they manage to achieve the accuracy up to 95%. The classification accuracy of their PPG signal might be affected by their skin. Children and early teenagers have thinner skin [6] compared to adult thus could affect the accuracy of the PPG readings as the signals are obtained using pulse oximeters where the device is attached onto the index finger to monitor the blood pressure. In addition, their epidermis is thinner and under-keratinized as compared with adults. Groups C, D, E, and F were individuals with age of 16 to 35 years old where they are late teenagers and adults. The classification accuracy within these groups was constant and stable. The results might be influenced by the heart respiratory rate as adult’s heart rate is lower and calmer.

Besides that, groups G, H and I were individuals with age of 36 to 50 years old where they are classed as middle adulthood. However, based on the results, only groups G and H were less constant in accuracy. As human grows older, middle age adults experience numerous physical changes which leads to cultivation of wrinkles, depletion of organs functions and also the decrease of the lung and heart capabilities [7]. Thus, the preciseness of the PPG signal might be affected due to the changes experienced by middle age adults. Therefore, it can be concluded that, PPG signal based biometric identification system is most suitable to be implemented for person from groups C, D, E and F as their PPG signals are stable and more consistent as compared to other groups.

Figure 23 demonstrates the classification accuracy of APG signal based on different age group. From the table and figure, it is clearly seen that the results are inconsistent for each group and classifier. Only Group I gave 100% accuracy and Group B gave constant results. Nevertheless, the results obtained are more than 90% accuracy which means that the APG signal is reliable to be applied for person recognition.

The inconsistency in the classification accuracy might be due to the number of peaks of APG signal. Some of the peaks are visible and some are not visible thus affecting the preciseness of the outcomes. Based on the figure, it can be concluded that, for APG signal, groups B and I are most suitable to be applied as biometric identification system because of the stability in classification accuracy.

On the other hand, for the APG to be applied as biometric identification method, it can be suggested that MLP and RBF networks can be applied as they give higher accuracy when compared to other type of classifier. MLP has the ability to derive from incomplete data and can be used to extract patterns. The ability of MLP to learn makes it very flexible and powerful. RBF network are able to give high accuracy because it is more intuitive compared to other type of neural network and much easier to be designed and trained. Therefore, this method can be applied for person recognition even with different age group.

Based on the results, it shows that both PPG and APG are reliable as biometric identification system despite of various ages. The PPG signal gave better accuracy for groups C, D, E, and F thus it is suggested that this technique should be applied for individuals within this age group in order to obtain high precision. For APG signal, MLP and RBF should be implemented as the accuracy for both classifiers are better than NB and BN.
V. CONCLUSION

A new potential substrate named BV board was developed. The characteristics of electrical properties were measured, analyzed and presented in this paper. Furthermore, the obtained results for BV board were compared with familiar existing substrate which is a FR-4 substrate. The results show BV board has more consistencies as compared to the FR-4 substrate. It is suggested to use BV in microwave application such as designing a patch antenna. For further study, other characteristics such as strength, breakdown voltage and water absorption of BV board will be evaluated. This is important in order to determine prospect of the board to be used in other applications such as insulation of electrical devices. In addition, a patch antenna would be developed onto this BV substrate in order to observe the real performance of this substrate in microwave application.

ACKNOWLEDGMENT

We would like to acknowledge International Islamic University Malaysia for providing us the Research Initiative Grant Scheme to fund this research.

REFERENCES


