Biometric Authentication via Finger Photoplethysmogram

Xiao Zhang
Tsinghua University
Haidian District, Beijing, China
zhangxiao14@mails.tsinghua.edu.cn

Zheng Qin
Tufts University
Medford, Massachusetts, USA
zheng.qin@tufts.edu

Yongqiang Lyu*
Tsinghua University
Haidian District, Beijing, China
luyq@tsinghua.edu.cn

ABSTRACT
With the advent of information society, identity recognition technology has become more and more important. Traditional biometric-based identity authentication technologies such as fingerprint, iris, and face have been widely used in various fields of society. However, such methods have more or less problems. This paper proposes a biometric identification technology based on photoplethysmogram (PPG) signal, which can achieve recognition accuracy about 90% via Support Vector Machine (SVM). Two kinds of identity authentication algorithms and optimization of one with Kernel Principal Component Analysis(KPCA) are introduced. Meanwhile, the False Reject (FR) and False Accept (FA) are low, which represents a good performance. Experimental results and simulation analysis show that the PPG-based technology is a reliable biometric technology with high security, and broad application prospects.

CCS Concepts
• **Security and privacy → Biometrics.**

Keywords
Finger photoplethysmogram; biometric recognition; multi-feature

1. INTRODUCTION
With the intensification of informationization, identity authentication has become an important part of our life. Traditional identity authentication technologies include resident identity cards, passports, magnetic cards, and account passwords. Although these technologies are widely used, they are more or less flawed, such as being easily lost, easily forgotten, easily falsified, easily stolen and attacked. However, biometrics-based identity authentication is in the ascendant. Different from the widely used systems such as fingerprint, face, voiceprint [1], and iris [2], the identity authentication based on PPG overcomes the limitations of traditional biometrics [3]. The content of this article is based on the PPG identity authentication. Compared with traditional biometric authentication methods, researches related to PPG are still in the beginning.

PPG is a simple, non-invasive, low-cost optical technique that can be used to detect changes in the volume of blood in tissue microvascular beds. This signal has been widely used in commercial fields and can be used to measure oxygen saturation, blood pressure, and cardiac output [4]. It can also be used to assess autonomic function and detect peripheral vascular disease. As a new biometric identification technology, PPG has been verified by studies for its universality, uniqueness, robustness and adaptability [5, 6]. The physiological record of PPG signal is more convenient for it can be acquired via wearable devices [7]. The PPG signal is a quasi-periodic signal and the waveform is divided into two parts: rising branch and falling branch. As shown in Fig. 1, the ascending branch represents the contraction of the heart, and the amount of blood in the arteries of the human body increases dramatically; the descending branch represents the diastole, and blood flows through the capillaries into the venous blood vessels resulting in a decrease in blood volume in the arteries. Besides, the beat pattern of the different moments of the same individual and the beat patterns of different individuals have a great correlation. This makes the PPG signal have a natural advantage in classification recognition [8].

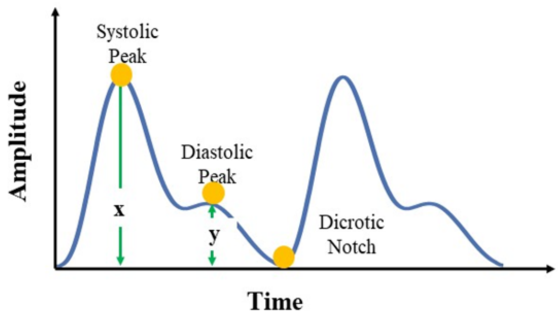


Figure 1. a typical waveform of the PPG.

The full text is arranged as follows. Some related concepts of PPG signal are present in Section 2. Section 3 introduces two PPG-based authentication algorithms while optimizing them. Section 4 introduces the contents of experimental collection, data preprocessing and result. Section 5 summarizes and looks forward to the above content.

2. RELATED WORK
Classification methods based on PPG waveforms and features are more and more popular. Bonissi et al. [5] proposed a baseline normalized method based on third-order highpass Butterworth filtering, while using the improved PanTompkins algorithm to achieve an EER of 9%. Kavsaolu et al. [6] proposed a PPG identification method based on features. They used FIR filters and extract time-domain features while ranking them based on distances. The recognition accuracy reaches 94.44% via KNN classification. Spachos et al. [8] also used KNN to do their research. Orjuela-Can et al. [9] used MLPs to learn the patterns of the time points near the starting point and the peak point of the PPG signal, and successfully achieved an accuracy of 98.1%.

CSAI '18, December 8–10, 2018, Shenzhen, China

ACM ISBN 978-1-4503-6606-9/18/12…$15.00
DOI: https://doi.org/10.1145/3297156.3297174

3. ALGORITHM AND OPTIMIZATION

3.1 PPG-45-based Authentication Algorithm

3.1.1 45 Features of PPG (PPG-45)

Features can be a good abstraction of the characteristics of the original data. In this experiment, researchers based on the research of Resit [6] and others, combined with the knowledge of the cognitive field, biomedicine, and biometric authentication, finally selected 45 features as basic data for identification. The 45 features include 39-time domain features and 6 frequency domain features.

3.1.2 Algorithm Flow

Let the total number of experimental individuals be n. Waveform period is T.

1) Calculate PPG-45
   The output is n sets of matrices (each matrix size is $T \times 45$).
2) Feature data processing
   Each matrix is averaged for every three adjacent row to obtain n sets of matrices (each matrix size is $[T/3] \times 45$). After that, the odd and even rows of each matrix are separated, and finally, n groups of matrices are obtained respectively as the training set and the test set (each matrix size is $[[T/3]/2] \times 45$).
3) SVM classification
   Since the data classification of this experiment is more than two types and it is linearly inseparable, a non-linear nuclear support vector machine is selected for classification during the application [9]. When classifying, the kernel function selects Gaussian kernel function. After the classification model was constructed, ten-fold cross validation and test set verification were used to obtain the final result.

3.2 Wave-based Authentication Algorithm

3.2.1 Algorithm Flow

Let the total number of experimental individuals be n. Waveform period is T.

1) One-dimensional interpolation
   Because the number of sample points in each cycle is inconsistent and subsequent processing cannot be performed, one-dimensional interpolation must be performed at the very beginning so that the number of sample points in each cycle is the same. Common interpolation methods include piecewise linear interpolation, cubic spline interpolation, and Lagrangian interpolation. Since the spline interpolation has many advantages compared to the other two, this experiment selects the cubic spline interpolation method, and the sample point for each cycle is 10 after the interpolation.
2) Sample data processing
   Each matrix is averaged for every three adjacent rows to obtain n sets of matrices (each matrix size is $[T/3] \times 10$). After that, the odd and even rows of each matrix are separated, and finally, n groups of matrices are obtained respectively as the training set and the test set (each matrix size is $[[T/3]/2] \times 10$).
3) SVM classification Be the same as PPG-45.

3.3 Optimize the Wave-based Algorithm

3.3.1 Kernel Principal Component Analysis(KPCA)

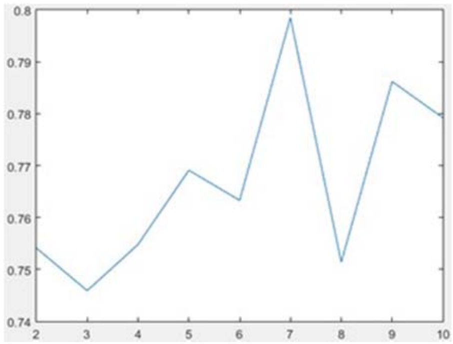
Principal component analysis (PCA) is a statistical method for transforming a set of potentially relevant variables into a set of linear uncorrelated variables through orthogonal transformation [10]. The transformed set of variables is called ingredient. In practical applications, there is a certain correlation between variables, that is, there is a certain overlap between the information reflected by this variable. PCA is to delete the correlation after projecting the variable, and keep the original information as much as possible.

Kernel Principal Component Analysis (KPCA) is a PCA-based algorithm [10]. Compared to PCA, KPCA can perform non-linear data reduction. The main idea of KPCA is to use nuclear techniques [11] to map eigenvectors in a linear space to high-dimensional space (Hilbert space) and then perform PCA processing in high-dimensional space. The purpose of mapping to high-dimensional space is to map data points that cannot be linearly classified in the low-dimensional space to high-dimensional space for linear classification.

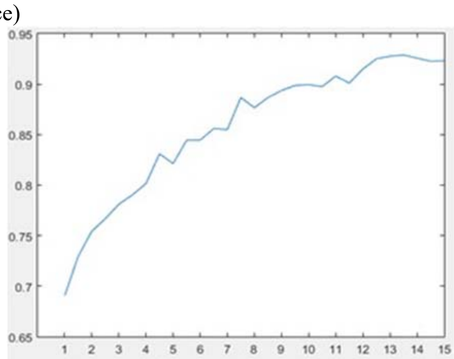
3.3.2 Parametric Simulation

The kernel function is a Gaussian radial kernel function. The simulation parameters are the dimensionality after dimension reduction and the variance of the kernel function.

1) Dimensionality after dimension reduction Fig.2.(a) shows that when the variance of the Gaussian radial kernel function is constant, the overall recognition rate changes with the dimensions after dimension reduction. From the figure, we can see that when the dimensionality is 7, the recognition rate reaches the highest.
2) The variance of the kernel function
   Fig.2.(b) shows that when the dimensionality after dimension reduction is 7, the overall recognition rate changes with the variance of the Gaussian radial kernel function. From the figure, the recognition rate reaches the highest when the variance is 13.5.



(a) Diagram of dimensionality and recognition rate (constant variance)



(b) Diagram of variance and recognition (constant dimensionality)

Figure 2. Parametric simulation.

4. EXPERIMENT AND DISCUSSION

4.1 Experimental Collection

This experiment uses a simple subtraction game. The rule of the game is to subtract 7 from 1000 one by one. The game does not have a repentance setting, that is, you cannot modify the answer whether you are right or wrong after each input. Each person to be tested needs to complete two sets of calculation games. After wearing the test equipment, everyone will be informed of the rules of the game and operational. Before the start of each test group, people are required to rest for 10 minutes in the experimental environment. Each set of computing game experiments lasts 5 minutes. The first two minutes are the preparation time, which means that the PPG signal of subject is collected but no game is played. After that, people would do the test. There will be a prompt tone during the phase transition. People must try their best to complete the test. The experiment cannot be abandoned after the experiment starts. After the experiment is completed, they will fill out a questionnaire. A total of 19 college students with similar educational experiences participated in the experiment. The experiment strictly followed the "Helsinki Declaration." Each person would be paid 100 yuan as compensation after the end of the experiment. In this experiment, the individual's PPG signal was collected using a transmission-type acquisition method. The collection frequency of equipment was 200 Hz, and the wearing part was a left index finger. The experiment was conducted in a quiet, electrically shielded, stable light source room with constant temperature and humidity.

4.2 Data Preprocessing

Due to the large interference in the environment during the acquisition, the PPG signal must be filtered firstly. Experiments show that there are three main types of interference in the PPG signal: baseline drift, motion artifact noise, and high-frequency random interference. Due to the large proportion of low-frequency interference, smooth filtering and band-pass filtering can be used for processing. Smoothing filtering is a low frequency enhanced spatial domain filtering technique. The key to smooth filtering is to select the neighborhood. The size of the neighborhood is directly related to the effect of smoothing. The larger the neighborhood is, the better the smoothness is. However, if the neighborhood is too large, the loss of the edge information will be larger and the output will be ambiguous. So we need to choose a proper size of the neighborhood. After testing, every 30 points as a neighborhood filter effect is relatively good. In addition, the experiment takes a 0.005-0.05 Hz 200-order band pass FIR filter for further process. Since the two algorithms used in the paper require the single cycle of PPG signal as data element, the PPG signal must be divided into discrete single-cycle waveforms. Due to the quasi-periodic nature of the PPG signal, a method for finding the neighbor maxima is used in segmentation.

4.3 Detection Error Tradeoff (DET)

As a learning and detection task, PPG-based identity authentication includes two possible error types: FR (False Rejection) and FA (False Acceptance). FR refers to an error that "must not be matched as a match." FA refers to an error that "must be matched as a mismatch." These two errors are contradictory, but they can be offset by changing the decision threshold [12]. In the two-classification problem, FRR (False Rejection Rate) and FAR (False Acceptance Rate) are indicators that correspond to these two errors respectively. Through these two indicators, we can do a preliminary analysis and judgment on the performance of a biometric system. As for the security of the system, if the FA is lower as the FR is larger, then the security of the system is higher. At the same time, as for the convenience, if the FR is lower as the FA is larger, the system is more convenient. Therefore, FR and FA are a group of interrelated and very useful judgments. As a traditional method, ROC (Receiver Operating Characteristic) plots the error rate under different decision thresholds. However, because the ROC curves are stacked at the origin of the coordinates, the visual effect is not good. Therefore, a new illustration, DET (Detection Error Tradeoff), is proposed. DET [11] is similar to ROC in principle, but its image is clearer than ROC, so this study uses DET to analyze system performance.

4.4 Result and Discussion

4.4.1 Recognition Rate and Time Analysis

The overall recognition rate of these three algorithms meets the requirements shown in Table 1. There is a large gap between the Model error rate and the Tenfold cross validation error rate, indicating a certain overfitting phenomenon. From the perspective of individual recognition rate, one or two individuals in the three algorithms have lower recognition rates. The improved algorithm has the highest recognition rate while the wave based algorithm has the lowest recognition rate. However, regarding running time, the wave-based algorithm runs faster, and the recognition rate does not drop much more than the other two. From this point, the wave-based algorithm is better. Since the KPCA analysis takes a very long time, the optimized algorithm is not suitable for practical application.

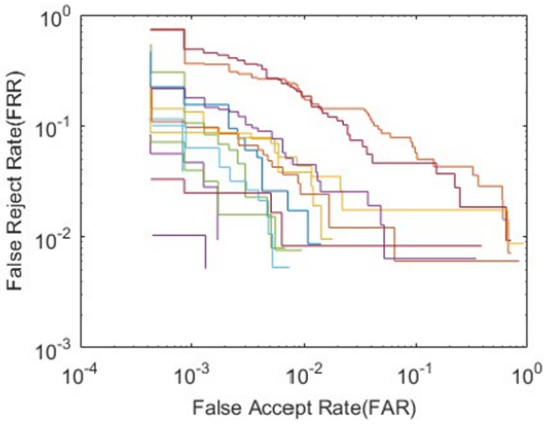
Table 1. Result of three algorithms

Algorithm	Model error rate	Tenfold cross validation error rate	Test set recognition rate	Running time
PPG-45	$4.0800 * 10^{-4}$	0.1032	91.15%	(135+97)s
Wave-based	0.0480	0.1485	87.70%	50s
Optimization	0.0046	0.0240	92.34%	Long

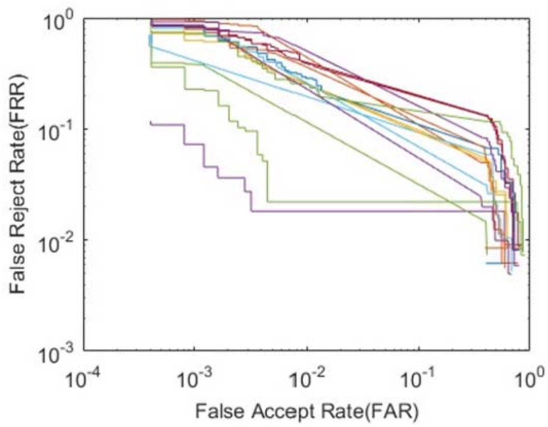
4.4.2 DET Analysis

There were 19 participants in this experiment, and the research belongs to a multiple classification issue. It is necessary to establish 19 kinds of binary models (for each individual classified as 'belonging to' or 'not belonging to') to perform DET analysis. Draw the DET curve for each model, as shown in Fig.3.

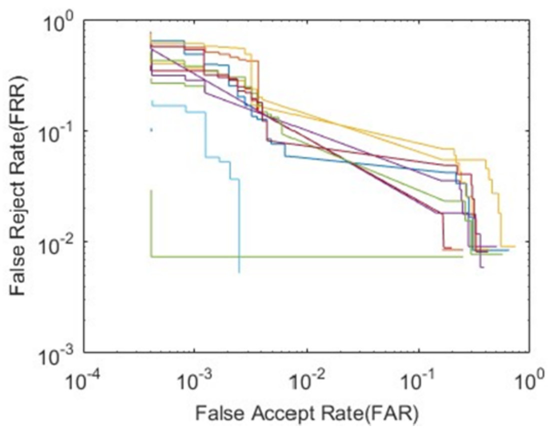
As can be seen from Fig.3, the curves of Algorithms 1 and 3 are close to the lower left corner, so the overall performance of the algorithm is high. The curves 1 and 3 are steep and close to the lower left corner, so the security is good; the curve of Algorithm 1 is closer to the horizontal axis, so the comfort is better. Comparing Algorithms 2 and 3, after the KPCA optimization, the DET curve is better close to the lower left corner, so KPCA has a certain degree of improvement in algorithm performance.



(a) Diagram of dimensionality and recognition rate (constant variance)



(b) Diagram of variance and recognition (constant dimensionality)



(c) Diagram of variance and recognition (constant dimensionality)

Figure 3. DET diagrams of three algorithms.

5. CONCLUSION

PPG-based biometric algorithms have a wide range of applications, but there are more or less certain security issues. Because the algorithm requires the individual to wear sensors for real-time acquisition, it imposes higher requirements on hardware performance and local security. However, compared to traditional biometric methods such as iris and fingerprint, PPG signal itself is difficult to copy and disguise. Therefore, safety has been greatly improved.

This thesis begins with the relevant concepts of PPG, combines specific experimental data and finally implements and optimizes the identity authentication algorithm based on PPG with two ideas. Although the recognition rate satisfies the requirement on the whole, the optimized algorithm has a slower simulation speed and is not suitable for real-time analysis, and some individual's recognition rate is higher than that of other individuals. In practical applications, the acquisition method is transmission-type acquisition, which is inconvenient. Therefore, we can now study reflective acquisition on the basis of the present.

6. ACKNOWLEDGMENTS

We gratefully acknowledge financial support from the Chinese National Key Research and Development Program under Grant No. 2017YFB0403404.

7. REFERENCES

[1] Matthew J Yuschik and Robert J Slezak. 2002. Voiceprint identification system.

[2] John Daugman. 2009. Chapter 25 - How Iris Recognition Works. In The Essential Guide to Image Processing, Al Bovik (Ed.).Academic Press, Boston, 715 – 739. https://doi.org/10.1016/ B978-0-12-374457-9.00025-1

[3] A. K. Jain, A. Ross, and S. Prabhakar. 2004. An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology 14, 1 (Jan 2004), 4–20. https://doi.org/10.1109/TCSVT.2003.818349

[4] M Elgendi. 2012. On the analysis of fingertip photoplethysmogram signals. Current Cardiology Reviews 8, 1 (2012), –.

[5] A. Bonissi, R. D. Labati, L. Perico, R. Sassi, F. Scotti, and L. Sparagino. 2013. A preliminary study on continuous authentication methods for photoplethysmographic biometrics. In 2013 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications. 28–33. https://doi.org/10.1109/BIOMS.2013.6656145

[6] A. Reit Kavsaolu, Kemal Polat, and M. Recep Bozkurt. 2014. A novel feature ranking algorithm for biometric recognition with PPG signals. Computers in Biology and Medicine 49 (2014), 1 – 14. https://doi.org/10.1016/j.compbiomed.2014.03.005

[7] T. Choudhary and M. S. Manikandan. 2016. Robust photoplethysmographic (PPG) based biometric authentication for wireless body area networks and m-health applications. In 2016 Twenty Second National Conference on Communication (NCC). 1–6. https://doi.org/10.1109/NCC.2016.7561152

[8] P. Spachos, Jiexin Gao, and D. Hatzinakos. 2011. Feasibility study of photoplethysmographic signals for biometric identification. In 2011 17th International Conference on Digital Signal Processing (DSP). 1–5. https://doi.org/10.1109/ICDSP.2011.6004938

[9] Alvaro D. Orjuela-Can˜o´n, Denis Delisle-Rodr´ıguez, Alberto Lo´pezDelis, Ramo´n Fernandez Vara-Prieto, and Manuel B. Cuadra-Sanz.2013. Onset and Peak Pattern Recognition on Photoplethysmographic Signals Using Neural Networks. In Proceedings, Part I, of the 18th Iberoamerican Congress on Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications Volume 8258 (CIARP 2013). Springer-Verlag, Berlin, Heidelberg, 543–550. https://doi.org/10.1007/978-3-642-41822-8 68

[10] Bernhard Sch¨olkopf, Alexander J. Smola, and Klaus-Robert Mu¨ller. 1999. Advances in Kernel Methods. MIT Press, Cambridge, MA, USA, Chapter Kernel Principal Component Analysis, 327–352. http://dl.acm.org/citation.cfm?id=299094.299113

[11] Mingxing He, Shi-Jinn Horng, Pingzhi Fan, Ray-Shine Run, RongJian Chen, Jui-Lin Lai, Muhammad Khurram Khan, and Kevin Octavius Sentosa. 2010. Performance Evaluation of Score Level Fusion in Multimodal Biometric Systems. Pattern Recogn. 43, 5 (May 2010), 1789–1800. https://doi.org/10.1016/j.patcog.2009.11.018

[12] J. Ortega-Garcia, J. Bigun, D. Reynolds, and J. GonzalezRodriguez. 2004. Authentication gets personal with biometrics. IEEE Signal Processing Magazine 21, 2 (Mar 2004), 50–62. https://doi.org/10.1109/MSP.2004.1276113