THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Sensor-based Smart Recognition System For Wearable Devices

Weitao Xu

Master of Engineering, Shandong University

Bachelor of Engineering, Shandong University

A thesis submitted for the degree of Doctor of Philosophy at

The University of Queensland in 2016

School of Information Technology and Electrical Engineering

Abstract

With recent advances in wireless sensor networks and embedded computing technologies, on-body wearable devices such as smartphones and smart watches have become increasingly popular and play significant roles in our daily lives. The prevalence of smart wearable devices has sparked a new set of mobile computing applications that leverage the abundant information from sensors. In this thesis, I address three problems for wearable devices. These three problems correspond to three different recognition tasks: to help the user recognize a subject (face recognition), to assist the device to authenticate the identity of another device (device pairing) and to authenticate user (user authentication). The first problem is to implement a robust and efficient face recognition system on smart glass. The main challenge is the tension between the high computation requirements of accurate face recognition algorithms and the resource constraints of smart glasses. To address this challenge, I propose a robust and efficient sensor-assisted face recognition system on smart glasses by exploring the power of multimodal sensors including the camera and Inertial Measurement Unit (IMU) sensors. Extensive evaluation results show the proposed system improves recognition accuracy by up to 15% while achieving the same level of system overhead compared to the existing face recognition system (OpenCV algorithms) on smart glasses. The second problem is to generate a cryptographic key for legitimate devices so that devices on the same user's body can be paired together. I propose an automatic key generation system for wearable devices based on the user's unique gait. The intuition of the proposed key generation approach is that the devices on the same body experience similar motion signals that are produced by the unique walking pattern of the user. Therefore, the unique gait signal can be exploited as shared information to generate secret keys for all on-body devices. The evaluation results show that the proposed system can generate a common 128-bit key for two legitimate devices with 98.3% probability. The third problem is to develop a gait-based user authentication system by using Kinetic Energy Harvesting (KEH). The main feature of the proposed system is it utilizes the output voltage signal of the energy harvester to achieve gait recognition rather than the accelerometer. Compared with traditional accelerometer-based gait recognition system, the proposed system can reduce energy consumption by 78% while achieving comparable recognition accuracy.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

Journal papers

- **Weitao Xu**, Chitra Javali, Girish Revadigar, Chengwen Luo, Neil Bergmann, and Wen Hu. "Gait-Key: A Gait-based Shared Secret Key Generation Protocol for Wearable Devices". To appear in ACM Transactions on Sensor Networks (TOSN), accepted for publication in Nov 2016.

Conference Papers

- **Weitao Xu**, Guohao Lan, Qi Lin, Sara Khalifa, Neil Bergmann, Mahbub Hassan, Wen Hu. "KEH-Gait: Towards a Mobile Healthcare User Authentication System by Kinetic Energy Harvesting". To appear in the Proceedings of the Network and Distributed System Security Symposium 2017 (NDSS).
- **Weitao Xu**, Yiran Shen, Neil Bergmann, and Wen Hu. "Sensor-assisted Face Recognition System on Smart Glass via Multi-view Sparse Representation Classification". In Proceedings of the 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), pages 1-12, IEEE, 2016.
- **Weitao Xu**, Girish Revadigar, Chengwen Luo, Neil Bergmann, and Wen Hu. "Walkie-Talkie: Motion-Assisted Automatic Key Generation for Secure On-Body Device Communication". In Proceedings of the 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), pages 1-12, IEEE, 2016.
- Guohao Lan, **Weitao Xu**, Sara Khalifa, Mahbub Hassan, and Wen Hu. "Transportation Mode Detection Using Kinetic Energy Harvesting Wearables". In Proceedings of the 14th IEEE International Conference on Pervasive Computing and Communications (Percom), pages 1-4, IEEE, 2016.
- Girish Revadigar, Chitra Javali, **Weitao Xu**, Wen Hu, and Sanjay Jha. "Secure Key Generation and Distribution Protocol for Wearable Devices". In Proceedings of the 14th IEEE International Conference on Pervasive Computing and Communications (Percom), pages 1-4, IEEE, 2016.
- Yongtuo Zhang, Wen Hu, **Weitao Xu**, Hongkai Wen, and Chun Tung Chou. "NaviGlass: Indoor Localisation Using Smart Glasses". In Proceedings of the 14th International Conference on Embedded Wireless Systems and Networks (EWSN), pages 205-216, ACM, 2016.

Posters

- **Weitao Xu**, Yiran Shen, Neil Bergmann, and Wen Hu. “Poster Abstract: Robust and Efficient Sensor-assisted Face Recognition System on Smart Glass”. In Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems (SenSys), pages 379-380, ACM, 2015.
- **Weitao Xu**, “Mobile Applications Based on Smart Wearable Devices”. In Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems (SenSys), pages 505-506, ACM, 2015.
- Yiran Shen, Chengwen Luo, **Weitao Xu**, and Wen Hu. “Poster Abstract: An Online Approach for Gait Recognition on Smart Glasses”. In Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems (SenSys), pages 389-390, ACM, 2015.

Publications included in this thesis

"Sensor-assisted Face Recognition System on Smart Glass via Multi-view Sparse Representation Classification" – incorporated as the basis for Chapter 3.

Contributor	Statement of contribution
Author Weitao Xu (Candidate)	Designed experiments (80%) Wrote the paper (70%)
Author Yiran Shen	Designed experiments (20%) Wrote and edited paper (20%)
Author Wen Hu, Neil Bergmann	Wrote and edited paper (10%)

"Walkie-Talkie: Motion-Assisted Automatic Key Generation for Secure On-Body Device Communication" – incorporated as the basis for Chapter 4.

Contributor	Statement of contribution
Author Weitao Xu (Candidate)	Proposed algorithms (80%) Implemented system (80%) Wrote the paper (70%)
Author Girish Revadigar	Proposed algorithms (20%) Wrote and edited paper (10%)
Author Chengwen Luo	Implemented system (20%) Wrote and edited paper (10%)
Author Wen Hu, Neil Bergmann	Wrote and edited paper (10%)

"KEH-Gait: Towards a Mobile Healthcare User Authentication System by Kinetic Energy Harvesting" – incorporated as the basis for Chapter 5.

Contributor	Statement of contribution
Author Weitao Xu (Candidate)	Proposed algorithms (100%) Implemented system (80%) Wrote the paper (70%)
Author Guohao Lan, Qi Lin, Sara Khalifa	Implemented system (20%) Wrote and edited paper (20%)
Author Neil Bergmann, Mahbub Hassan, Wen Hu	Wrote and edited paper (10%)

Contributions by others to the thesis

No contributions by others.

Statement of parts of the thesis submitted to qualify for the award of another degree

None

Acknowledgements

Firstly, I would like to begin by thanking my supervisors, Prof. Neil Bergmann and Dr. Wen Hu, for the guidance, patience and encouragement I has been given throughout the PhD degree. Foremost, I would like to thank my principal supervisor Prof. Neil Bergmann from University of Queensland (UQ). Thanks for offering me the PhD position and supervision during the past three and half years. Prof. Neil Bergmann has deep insights into research questions and can always provide instructive suggestions. He is also a kind man in everyday life and I have a great time with him during the past years. I would also thank Dr. Wen Hu, who is my co-supervisor. Currently he is a senior lecturer in University of New South Wales (UNSW). He was more like a friend and encouraged me to overcome the challenging research questions.

Next, I would like to thank Dr. Yiran Shen. Yiran was my classmate when I was a undergraduate student. He helped me to start PhD study in Australia. We collaborated to conduct research projects, and he gave me a great deal of valuable help for my research. I would also like to thank Dr. Chengwen Luo. Chengwen was a Postdoc when I visited UNSW. Chengwen and I also worked together, and he helped me implement systems. I would also thank all the members I met in UQ, CSIRO, and UNSW. I enjoyed great experience with them. They are great people and always ready to help others. Specifically, I would thank some of my colleagues: Guohao Lan, Girish Revadigar, and Yongtuo Zhang. We worked together to overcome many challenging problems. I would also thank Prof. Mahbub Hassan and Prof. Chun Tung Chou. They gave me great help when I visited UNSW.

I would like to thank my family and my friends. Thank you to my wife, Susan, for her support and patience. Thank you for encouraging me in all of my pursuits and inspiring me to follow my dreams. Thank you to my parents for their love, support and encouragement.

Finally, I would like to thank the scholarship support from UQ, CSIRO and China Scholarship Council (CSC).

Keywords

wearable devices, smart glass, smart watch, face recognition, sparse representation, key generation, source separation, gait recognition, energy harvesting

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 080602, Computer-Human Interaction, 60%

ANZSRC code: 080502 Mobile Technologies, 20%

ANZSRC code: 080504 Ubiquitous Computing, 20%

Fields of Research (FoR) Classification

FoR code: 0806, Information Systems, 60%

FoR code: 0805, Distributed Computing, 40%

Contents

List of Tables

List of Figures

List of Abbreviations

The following table describes the meaning of various abbreviations and acronyms used throughout the thesis. The page on which each one is defined or first used is also given.

Abbreviation	Meaning	Page
ADC	Analog-to-Digital Converter	110
AES	Advanced Encryption Standard	60
BCs	Biometric Cryptosystems	22
BLE	Bluetooth Low Energy	5
BSS	Blind Source Separation	13
CPU	Central Processing Unit	1
CS	Compressive Sensing	9
CSI	Channel State Information	10
DH	Diffie-Hellman	4
ECC	Error Correcting Code	13
EEH	Electromagnetic Energy Harvester	7
EKF	Extended Kalman Filter	30
FPS	Frame Per Second	36
I2C	Inter-Integrated Circuit	101
ICA	Independent Component Analysis	13
IMDs	Implantable Medical Devices	5
IMU	Inertial Measurement Unit	3
IoT	Internet of Things	1
KEH	Kinetic Energy Harvesting	6
KNN	K-Nearest Neighbor	92
MAC	Message Authentication Code	64
MASO	Maximum Accuracy SamplingOptimization	3
MESO	Minimum Energy Sampling Optimization	3
MITM	Man-in-the-middle	76
MSE	Mean Square Error	26
MSSRC	Multi-Step Sparse Representation Classification	7
MVSRC	Multi-view Sparse Representation-based Classification	3
NB	Naive Bayes	92
OCR	optical character recognition	11
PEH	Piezoelectric Energy Harvester	6
PIN	Personal Identification Number	50
RSSI	Received Signal Strength Indicator	14
SNR	Signal to Noise Ratio	3
SRC	Sparse Representation-based Classification	3
SVM	Support Vector Machine	94
SVR	Support Vector Regression	3

Chapter 1

Introduction

Recent years have witnessed a remarkable growth in the number of smart wearable devices such as Apple Watch and Google Glass. It is estimated that there will be 20 billion Internet of Things (IoT) devices by the year 2020, and the majority of these will be wearable devices [99]. Much like the embedded systems they originate from, on-body IoT devices are equipped with a number of sensors which offer means to collect, store and distribute information. Therefore, smart wearable devices start to accumulate a great deal of sensitive data about their users, such as health data, emails, and locations. More importantly, the ubiquity of wearable devices has sparked a new set of mobile computing applications that significantly improve user experience such as object recognition [31], localization [147], and health monitoring [68].

Among these applications, one fundamental task is identity recognition. Identity recognition not only refers to helping users recognize a subject, but also it refers to assisting a device to recognize the identity of another device (i.e., device pairing) or user (i.e., user authentication). Although a large number of novel systems have been proposed on wearable devices [95, 145, 112, 149], the problem of robust and efficient recognition system has not been well studied for two main reasons. Firstly, smart mobile devices such as smartphone and smart watch, limited by their computational capacities (e.g., CPU, memory) and power supply, do not always meet the requirements of complex pattern recognition algorithms. For example, the battery life of smart glasses is limited by its size. It is reported that the fully charged battery on the Vuzix Smart Glass can last for one hour; however our practical experience shows that the battery would be completely drained within half an hour with display on, camera open and high CPU loading. Besides, constantly charging batteries is inconvenient and may introduce extra costs. Thus to improve energy efficiency and reduce computational cost is a crucial task for authentication

systems on resource-constrained mobile devices. Secondly, the majority of existing studies on recognition systems have used a very restrictive experimental setup where the performance evaluation was conducted on a dataset collected from a controlled laboratory environment. The performance of the system decreases rapidly when used in a real-world environment. With the pervasiveness of wearable devices in the wild, these is a need for robust authentication system in a realistic environment.

In this thesis, I will show how I address these challenges on wearable devices in three different systems. These three systems correspond to three different recognition tasks: to help the user recognize a subject (face recognition), to assist the device to authenticate the identity of another device (device pairing) and to authenticate the user (user authentication). Specifically, the three problems addressed in this thesis are:

- Robust and efficient face recognition system on smart glasses.
- Motion-assisted automatic device pairing system for wearable devices.
- Gait-based user authentication system using kinetic energy harvesting.

Each of these problems will be addressed in a chapter of this thesis. An overview of these three problems will now be presented.

1.1 Robust and Efficient Face Recognition System on Smart Glasses

Face recognition is a popular biometric-based authentication technique with applications ranging from surveillance to device unlocking to the organization of personal image collections. Face recognition has been extensively studied in the traditional computer vision community. However, as discussed in [124], most of the advanced face recognition methods fail on smart wearable devices because of the tension between high computation requirements and resource constraints. Recognition performance and energy consumption are two important factors that impede the prevalence of face recognition on smart glasses. Users expect real-time feedback with high accuracy in unconstrained real-world scenarios. Although cloud-based architectures can reduce the burden on smart glasses, the usability of the cloud-based recognition systems relies on the wireless connectivity. Besides, the qualities of the wireless connections will also

affect the cost of transmission significantly. According to the results reported in [28], the power consumption of wireless transmission on a smartphone (about 720mW) is approximately 12 times higher than that of CPU (about 60mW). Therefore, the *in-situ* approaches are preferable considering the cost of wireless transmission and the inconvenience of using wireless connections. This leads to the following research question:

Research Question 1: How the computational cost of face recognition system on smart glasses can be reduced, and how can face recognition performance on smart glasses be improved with Inertial Measurement Unit (IMU) sensor data?

In this thesis, I propose and implement a novel sensor-assisted face recognition system which runs locally on smart glasses by exploiting the information from both the camera and sensors on smart glasses to improve the recognition accuracy and reduce the energy consumption. The contributions of this study are threefold:

- I propose a novel face recognition algorithm called Multi-view Sparse Representation based Classification (MVSRC). It exploits the high agreement among the sparse representations of the face images from different view angles and applies a novel weighted Sparse Representation-based Classification (SRC) model to improve the Signal to Noise Ratio (SNR). The evaluation on several datasets show that MVSRC outperforms several state-of-the-arts multi-view face recognition algorithms.

- I propose a Support Vector Regression (SVR)-based estimation model to relate the recognition accuracy to the angle information obtained by IMU sensors. Then I design two sampling optimization approaches: Maximum Accuracy Sampling Optimization (MASO) and Minimum Energy Sampling Optimization (MESO) based on the estimation model to improve the efficiency of MVSRC while preserving its high recognition accuracy. MVSRC after sampling optimization is referred to fast-MVSRC.

- I implement a face recognition system based on the proposed methods on smart glasses and demonstrate that it significantly outperforms the existing *in-situ* face recognition algorithms on smart glasses. I also discuss the offloading approach and experimentally show that the cost of our system is in the same order of offloading to a nearby server (cloudlet).

This work will be described in detail in Chapter 3.

1.2 Motion-assisted Automatic Device Pairing System for Wearable Devices

With recent advances in wireless sensor networks and embedded computing technologies, wearable and implantable devices such as smartphones, smartwatches and pacemakers have become increasingly popular and play significant roles in our daily lives. For users, it is of potentially great value to associate a personal device with another device in a spontaneous manner. Pairing devices can be for the purpose of short-lived interactions, for example, file transfer and synchronization, or aimed at longer lived pairing, for example, pairing a smartphone with accessories. In current mobile systems, this is achieved by key exchange methods, which are either manual (e.g. typing in the key in a keypad) or exploit key-exchange algorithms. For the first case, a common mechanism for peer device authentication is Personal Identification Number (PIN) code entry by the user into the involved devices. However, a primary requirement for human-involved authentication is ease-of-use. Therefore, the human-involved authentication method is undesirable when users seek to engage in fast and short-lived authentications frequently. For the key-exchange algorithms, a common key exchange method is the Diffie-Hellman (DH) protocol which is used to distribute a symmetric key between two parties [41]. However, the premise of DH protocol is that two devices to be paired together are legitimate devices. It cannot be used to distinguish adversary devices from legitimate devices. This leads to the following research question:

Research Question 2: How can two legitimate devices belonging to the same user establish a secure communication channel in a user friendly manner?

In this work, I propose and implement a motion-assisted key generation technique to secure on-body device communication. The intuition of the proposed key generation approach is that the devices on the same body experience similar motion signals that are produced by the unique walking pattern of the user. Therefore, the unique gait signal can be exploited as shared information to generate secret keys for all on-body devices. The proposed approach enables unobtrusive establishment of secure communications between on-body devices. The main contributions of this study are threefold:

- **Source separation for body motion signal**: By using Blind Source Separation to separate motion signals generated from different body movements, e.g., gait and arm swing

motions, the proposed key generation approach achieves robust performance in generating keys for devices located at different body locations.

- **Shared key generation scheme**: I will present a novel, light-weight key generation scheme for on-body IoT devices based on body motion signals. I experimentally demonstrate that a common 128-bit key can be successfully generated by two independent wearable devices on the same body in 98.3% of the case, while the scheme also provides adequate security guarantees against impersonation attacks. By walking for 4.6s ($\approx$9 steps), the proposed key generation approach is able to generate a 128-bit key with entropy varying from 0.94 to 1 which demonstrates the high randomness of the keys.

- **System implementation**: I will illustrate the practicability of the proposed key generation approach by implementing the system in Bluetooth Low Energy (BLE) peripheral mode. I will investigate the system computation overhead and power consumption, and demonstrate the feasibility of the proposed scheme for contemporary on-body IoT devices.

This work will be described in detail in Chapter 4.

1.3 Gait-based User Authentication System Using Kinetic Energy Harvesting

With rapid advancements in embedded technology, wearable devices and Implantable Medical Devices (IMDs) have become an integral part of our everyday life. It is predicted that by 2025, the market for personal wearable devices will reach 70 billion dollars [5]. The major deployments of those devices are expected to be in health monitoring and medical assistance domains [11, 81]. Some popular wearable devices, such as Fitbit and Apple Watch, are already monitoring and storing a mass of sensitive health data about the user.

However, such wearable systems are vulnerable to impersonation attacks in which an adversary can easily distribute his device to other users so that data collected from these users can be claimed to be his own. In this way, the attacker can claim potential healthcare profits that are allocated to people with certain illnesses even though he may not have any illnesses [115]. To mitigate the risk of malicious attacks, most wearable devices rely on explicit manual entry of a secret PIN number. However, due to the small screens of wearable devices and frequent

unlocking requests, it is inconvenient for users to enter the keys manually. Furthermore, this method is not applicable when an adversary colludes with other users to spoof the healthcare company.

Gait recognition using wearable sensors, such as accelerometers, has emerged as one of the most promising solutions for user authentication. It offers several advantages over other biometrics especially when applied in wearable devices. Extensive previous studies have already demonstrated its feasibility in user authentication [49, 148], but they have also shown that continuous accelerometer sampling drains the battery quickly. A vision for wearable devices is to be battery-free (self-powered). A current trend in battery-free devices is to investigate kinetic energy harvesting (KEH) solutions to power the wearable devices [54, 135, 142, 101]. For example, AMPY [2] has released the world's first wearable motion-charger which can transform the kinetic energy from user's motion into battery power.

Motivated by this prospect, I propose gait recognition by simply observing the output voltages of KEH. The feasibility of the proposed idea is based on the observation that if humans have unique walking patterns, then the corresponding patterns of harvested power from KEH should be unique too. The proposed system offers several advantages. The major advantage of KEH-based gait recognition is the potential for significant power savings arising from not sampling an accelerometer at all. On the other hand, the output voltage can be used to charge the battery, thus further extending battery life. Finally, as energy harvesters will be integrated in wearable devices in the near future, the output voltage can be naturally utilized for authentication purposes without introducing extra sensors. This makes it a promising solution for light-weight authentication for wearable devices. To the best of our knowledge, this is the first work that proposes and experimentally validates the feasibility of gait recognition using KEH. This leads to the following research question:

Research Question 3: How does gait recognition using a Kinetic Energy Harvester compare to more conventional gait recognition using accelerometers in terms of accuracy, energy-efficiency and robustness to attack?

The main contributions of this work are as follows:

- I propose a novel gait-based user authentication system for mobile healthcare system, called KEH-Gait, which uses only KEH voltage as the source signal to achieve user authentication.

- I build two different KEH prototypes, one based on piezoelectric energy harvester (PEH)

and the other on electromagnetic energy harvester (EEH). Using these KEH devices, I evaluate gait recognition accuracy of KEH-Gait over 20 subjects. The evaluation results show that, with conventional classification techniques, which operate over single step, KEH-Gait achieves approximately 6% lower accuracy compared to accelerometer-based gait recognition.

- I demonstrate that authentication accuracy of KEH-Gait can be increased to that of accelerometer-based gait detection by employing a novel classification method, called Multi-Step Sparse Representation Classification (MSSRC), which efficiently fuses information from multiple steps.

- Finally, using measurements, I demonstrate that currently available microprocessors can read KEH voltage within 33 μs, which is two orders of magnitude faster than the time it takes to wakeup, interrogate and read acceleration values from typical 3-axis accelerometers. This means that with microprocessor duty cycling, KEH-Gait promises major energy savings over conventional accelerometer-based gait detection.

This work will be described in detail in Chapter 5.

1.4 Organization of The Thesis

The rest of this thesis starts with a literature review in Chapter 2, where I discuss Sparse Representation-based Classification and the state-of-the-arts of the three topics discussed in Sections 1.1 to 1.3. In Chapter 3 I provide the details of our face recognition system on smart glasses. Then in Chapter 4, I propose and implement an automatic device pairing system that can pair two legitimate devices on the same body automatically. Chapter 5 presents a gait-based user authentication system using KEH. Finally, I conclude the thesis in Chapter 6.

Chapter 2

Literature Review

This thesis addresses three challenges on smart wearable devices: robust and efficient face recognition system on smart glasses, motion-assisted automatic device pairing system for wearable devices, and gait-based user authentication system using kinetic energy harvesting. In this chapter, I first present the background of the Sparse Representation-based Classification (SRC) method, which is used in face recognition system on smart glass and user authentication system. After that, I will review the literature related to each of the three problems in Section 2.2 - Section 2.4, where each section addresses one problem.

2.1 Sparse Representation-based Classsification

2.1.1 Background on SRC

This section introduces the rand-SRC face recognition algorithm [138] and opti-SRC [124].

In [138], face recognition is cast as a sparse representation problem and is solved by a Sparse Representation Classifier. SRC is applied to solve the traditional linear equation: $y = Ax$, where $y \in \mathbb{R}^p$ is the test image vector which comes from concatenating the pixel values by rows or columns; $A \in \mathbb{R}^{p \times (N \cdot K)}$ is the dictionary consisting of K classes and each class contains N p-dimensional image vectors. With the knowledge of A and y, ℓ_1 optimization can be applied to solve the linear equation with the *sparse assumption*:

$$\hat{x} = \arg\min_x \|x\|_1 \quad \text{subject to } \|y - Ax\|_2 < \varepsilon \tag{2.1}$$

where ε is used to account for noise and the sparse assumption holds when the test image vector

can be represented by one of the classes in A. Due to the large dimensionality of the image vectors, solving Eq. (2.1) can be computationally intensive. Inspired by the recent information theory technique of Compressive Sensing (CS) [14, 26, 42], a random projection matrix $R \in \mathbb{R}^{m \times p}$ ($m \ll p$) can be applied to improve the efficiency of the ℓ_1 optimization. In particular, the projection matrices are randomly generated from Bernoulli or Gaussian distributions because of their information preserving properties:

$$\hat{x} = \arg\min_x \|x\|_1 \quad \text{subject to } \|Ry - RAx\|_2 < \varepsilon \tag{2.2}$$

After obtaining the *sparse* representation vector $\hat{x} \in \mathbb{R}^{N \cdot K}$, the class results can be determined by checking the *residuals* based on the Euclidean distance. The definition of the residual for class i is:

$$r_i(y) = \|y - A\delta_i(\hat{x})\|_2 \tag{2.3}$$

where $\delta_i(\hat{x}) \in \mathbb{R}^{N \cdot K}$ contains the coefficients related to class i only (the coefficients related to other classes are set to be zeros). Then the final result of the classification will be:

$$\hat{i} = \arg\min_{i=1,\ldots,K} r_i(y) \tag{2.4}$$

i.e., the right class produces the minimal residual.

To further improve the performance of SRC, [124] proposes a heuristic algorithm to find the optimal projection matrix instead of the random one. The classification accuracy is improved by up to 12% with the optimized projection matrix. They also improve the efficiency of SRC by casting the residual calculation to a significantly lower dimensionality by introducing the *compressed residual*:

$$r_i(y) = \|R_{opt}y - R_{opt}A\delta_i(\hat{x})\|_2 \tag{2.5}$$

where $R_{opt} \in \mathbb{R}^{m \times p}$ is the optimized projection matrix. The classification accuracy will be preserved at the significantly lower dimensionality as described in [124].

To conclude, the steps of the opti-SRC can be summarized as:

- Opti-SRC starts from building a dictionary A consisting of face images from different subjects. Then the optimized projection matrix is learned from dictionary A using the approach introduced in [124].

- Given a test image vector y, the coefficients vector is obtained by solving Eq. (2.2).
- The final classification result is determined by solving Eq. (2.4) which searches the minimum compressed residual obtained from Eq. (2.5).

2.1.2 Applications of SRC

SRC is an emerging classification method and has been successfully used in a variety of applications ranging from gait recognition [148], emotion recognition [36, 32], and image denoising [45]. Moreover, SRC has also been applied in sensor areas to solve a range of recognition tasks because it is known to be robust to noise. For example, Wei et al. [137] developed an acoustic classification system on wireless sensor networks by applying SRC to improve the recognition accuracy. Shen et al. [124] proposed opti-SRC by optimizing the random matrix used in SRC to increase the performance of face recognition system in smartphones. Several papers have exploited the sparsity of multiple measurements to improve the system performance. Misra et al. [100] used CS to compress GPS signals and exploits the information of various propagation paths to improve the SNR of GPS signals. In a recent work [136], the researchers improved activity classification accuracy by fusing several channel state information (CSI) vectors. In my study, I use CS to reduce the feature dimension of face images as will be described in Section 3.2.1.

2.2 Face Recognition on Mobile Devices

2.2.1 Face Recognition

Face recognition has been well researched in the computer vision community. It invokes new research challenges when used on smart devices. With the availability of OpenCV [8], many apps such as friends tagging have appeared on the app markets. There are three face recognition algorithms in OpenCV: EigenFaces [130], FisherFace [16] and LBPFace [10]. Although these methods can be used in real-time on smartphones, the recognition accuracies are unsatisfactory [124]. SRC [138] outperforms these three methods; however, it cannot provide consistent high recognition accuracy and is computationally intensive. To overcome its limitations, Shen et al. [124] proposed opti-SRC by optimizing the projection matrix to provide consistently better accuracy while solving the computation efficiency issue. Many efforts have also been made on

multi-view based face recognition to further improve recognition accuracy [144, 62, 109]. The advent of smart glasses makes face recognition easier to perform and more interactive for the user because of the first-person camera. In a recent study [56], a cloud-based system Gabriel was developed on Google Glass to provide cognitive assistance services, such as face recognition, object recognition and optical character recognition (OCR). Another face recognition application on Google Glass is NameTag [7], which allows users to capture face images and search the identities on social media sites, including more than 450,000 registered sex offenders. However, both Gabriel and NameTag cannot work without connecting to servers.

There have been several papers which propose sensor-assisted biometric authentication system. Biggio et al. [19] developed a multimodal system against spoofing attacks by fusing the information from the camera and fingerprint sensor. Chen et al. [30] proposed a face authentication system to prevent 2D media attacks and virtual camera attacks by utilizing motion sensors. Yang et al. [141] used motion sensors to compensate for the tilt of the smartphone for better face detection.

2.2.2 Applications on Smart Glasses

Smart glasses, e.g., Google Glass and Vuzix Smart Glasses, have attracted significant attention both from researchers and industrial communities since their introduction in 2013. Applications on smart glasses have advantages over other smart devices as smart glasses are equipped with a first-person camera which can be naturally used as a 'third eye' to deliver a significantly better user experience. For example, Yongtuo et al. [147] proposed an indoor localization system on smart glasses to help people navigate. Mayberry et al. [95] proposed iShadow to track the state of the eye through two cameras mounted on the eyeglass. Zhang et al. [145] presented a visual attention driven networking system based on smart glasses: iGaze, through which users can connect to a target by gazes. In the ThirdEye system [112], the smart glasses are used to track shoppers and infer the product layout by fusing video, WiFi, and inertial sensor data. Kiryong et al. [56] implemented a cloud-based cognitive assistance system on smart glasses to help people with cognitive impairment. Glimpse designed by Chen et al. [31] can provide continuous, real-time object recognition by using a cache scheme.

2.2.3 Summary of the Problem

The SRC face recognition algorithm has been shown to outperform some of state-of-the-art face recognition methods. However, SRC is a single image based face recognition method. It cannot be directly used in when the smart glasses capture a face image sequence. Besides, SRC is computationally intensive for resource-constrained mobile devices. Therefore, the computational efficiency of SRC for a face recognition system on smart glasses should be improved. Existing face recognition systems rely on clould-based architectures to reduce the computational burden for smart glasses. However, as discussed in Section 1.1, the usability of the cloud-based recognition systems relies on the wireless connectivity. Therefore, I aim to propose a robust and efficient *in-situ* face recognition system on smart glass. The solutions of these problems will be presented in Chapter 3.

2.3 Device Pairing System

2.3.1 Device Pairing System for Wearable Devices

Many techniques exist that could be used to generate a shared secret key between two parties by exploiting the wireless channel information. Some of the examples are physical layer characteristics based security mechanisms, e.g., Received Signal Strength Indicator (RSSI) have been proposed by researchers [117, 66, 125]. However, these schemes are suitable for wearable devices which are frequently exchanging wireless packets. The potential of using acceleration to generate a shared key has not been well explored in the literature. Bichler et al. [18] developed a method to generate a shared key based on acceleration data of shaking devices together.

There have been several previous works using accelerometers to determine whether the devices are worn on the same body. Cornelius and Kotz proposed to use coherence to analyze the similarity of acceleration signals from different devices, and then decide whether two devices are carried on the same body [35]. The idea of shaking two devices together to pair them was first proposed by Holmquist et al. [61]. Mayrhofer and Gellersen used the same technique but extended it to include secure authentication [96]. Hinckley developed a similar method to pair devices that uses bumping rather than shaking together [59]. These methods require the user to participate and shake/move the devices together, which is not suitable for many on-body devices such as a pacemaker.

2.3.2 Biometric-based Authentication System

Biometric recognition is the science of establishing the identity of a person using his/her anatomical and behavioral traits [65]. It has emerged as the most promising technique for recognizing individuals as it offers a number of advantages over traditional authentication methods such as passwords, plastic cards, and keys. Biometric recognition techniques include identification based on physiological characteristics (such as face, fingerprints, and voice) and behavioral traits (such as gait, signature and keystroke dynamics). Extensive research efforts have been devoted to this area. With the prevalence of smart devices, researchers now focus on how to authenticate the user of smart devices. Many biometric-based authentication systems have been proposed for smart devices, such as face recognition systems in smartphones [124, 30], gait recognition systems in smartphones [115, 107], user authentication based on keystroke analysis [22, 21], and gesture-based authentication systems for smart glasses [29].

Different from biometric recognition, biometric cryptosystems (BCS) were developed for the purpose of either securing a cryptographic key using biometric features or directly generating a cryptographic key from biometric features. BCS can be classified into two categories: key-binding schemes and key-generation schemes [113]. For key-binding schemes, helper data are obtained by binding a secret key to a biometric template. In key-generation schemes, keys are directly derived from a biometric template. Soutar et al. [127] proposed the first fingerprint based key-binding system Mytec2. After that, Juels and Wattenberg [70] proposed a fuzzy commitment scheme by combing error-correcting codes (ECC) and cryptography algorithms together. They later proposed one of the most popular BCSs called the fuzzy vault scheme [69]. A further study was conducted by Tong et al. [129], in which they proposed a fuzzy extractor to extract keys from biometric templates directly. State-of-the-art BCSs which were proposed previously mostly utilize physiological modalities such as iris [92] and fingerprint [83]. There are also some studies that use behavioral biometrics such as signature [90] and voice [27]. In a recent work [60], the researchers used gait to encrypt a cryptographic key through a fuzzy commitment scheme [70].

2.3.3 Blind Source Separation

Blind signal separation (BSS) refers to the separation of a set of source signals from a set of mixed signals, without any prior information (or with very little information) about the source signals or the mixing process. Independent Component Analysis (ICA) is one of the most

popular BSS techniques, and it aims to decompose a multivariate signal into independent non-Gaussian signals. ICA has been successfully applied in numerous areas such as biomedical signal processing [128] and speech separation [79]. De Moor et al. proposed to use ICA to decompose maternal and fetal Electrocardiogram (ECG) recorded simultaneously from cutaneous electrodes placed on the mother's abdomen and chest [37]. Other researchers have also applied ICA to remove artefacts from the Electroencephalogram (EEG) signals [128, 38]. Other examples in the biomedical area are the studies by Calhoun et al. and McKeown et al., in which ICA were applied to functional magnetic resonance imaging (fMRI) data to separate different active components [25, 97]. In the speech separation area, ICA is used for extracting the speech signals of interest from mixed signals [85]. The application of ICA on body sensor networks (BSN) is an emerging field. Lo et al. applied ICA on body sensor signals to separate different sources of movement due to running and respiration [87]. Atallah et al. used the ICA technique to detect walking gait impairment with an ear-worn sensor [13]. In a recent work [110], ICA was applied on an accelerometer sensor attached on the heel to distinguish the toe walking gait from normal gait in Idiopathic Toe Walker (ITW) children.

2.3.4 Summary of the Problem

Existing methods are not suitable for wearable devices because they either require the user to participate and shake/move the devices together, or they depend on a public key distribution architecture. Therefore, providing secure as well as efficient and user-friendly device pairing is a challenging task. Gait has not been well explored in BCS, therefore I propose a system to generate keys from gait signals. A significant challenge is the complex human motions involved in walking, and I address this problem by using BSS techniques. The proposed scheme can pair two devices on the same body automatically when the user is walking, thus improving the user experience significantly. The details of this work will be presented in Chapter 4.

2.4 Gait-based Authentication System Using Kinetic Energy Harvesting

2.4.1 Gait Recognition

Gait recognition has been well studied in the literature. From the way that gait data is collected, gait recognition can be categorized into three groups: vision based, floor sensor based, and wearable sensor based. In vision based gait recognition system, gait is captured from a remote distance using a video-camera. Then, video/image processing techniques are employed to extract gait features for further recognition. A large portion in the literature belong to this category [76, 71, 86, 57]. In floor sensor based gait recognition, sensors (e.g., force plates), which are usually installed under the floor, are used for capturing gait features, such as ground reaction force (GRF) [108] or heel-to-toe ratio [98].

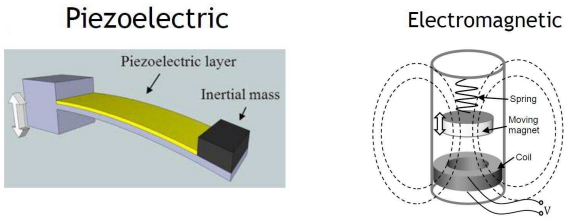
Compared with vision-based and other non-accelerometer based gait measurements, acceleration can reflect the dynamics of gait more directly and faithfully. For instance, accelerometer-based gait recognition does not suffer from the existing problems with vision-based methods, like occlusions, clutter, and viewpoint changes. Existing studies of wearable sensor-based gait recognition are mainly based on the use of body-worn accelerometers. Accelerometer-based gait recognition was first proposed by J.Mantyjarvi et al. [91] around 2005 and further developed by Gafurov et al. [48]. In the initial stages, dedicated accelerometers were used and worn on different body parts, such as lower leg [48], waist [91], hip [51], hip pocket, chest pocket and hand [133]. With the popularity of smartphones, researchers have proposed several methods to authenticate users by utilizing the built-in accelerometer [115, 40, 107, 39, 106, 75]. With the prevalence of wearable devices such as fitness trackers, researchers have proposed several gait-based authentication systems by utilizing the built-in accelerometer [89, 115, 107].

2.4.2 Kinetic Energy Harvesting (KEH)

Kinetic energy harvesting refers to the process by which energy is derived from kinetic energy, captured, and stored for wearable devices and wireless sensor networks. KEH has received growing attention over the last decade. The research motivation in this field is due to the reduced power requirement of small electronic components, such as the wireless sensor networks

used in passive and active monitoring applications. The three basic vibration-to-electric energy conversion mechanisms are the piezoelectric [88, 15], electromagnetic [44, 74, 116] and electrostatic [47, 33].

I built two energy harvesting devices based on piezoelectric and electromagnetic respectively. Therefore, I briefly describe PEH and EEH here to make the thesis self-contained. The piezoelectric effect converts mechanical strain into electric current or voltage. This strain can come from many different sources, such as human motions and low-frequency seismic vibrations. Figure 2.1(a) shows a basic design of PEH. Piezoelectric vibrational energy harvesters are usually inertial mass based devices, where a cantilever with a piezoelectric outer layer is excited into resonance by a vibration source at the root of the cantilever. The inertial mass is located on a vibrating host structure and the dynamic strain induced in the piezoelectric layer results in an alternating voltage output. Unlike piezoelectric harvesters, the basic principle of electromagnetic generators are based on Faraday's law of electromagnetic induction. As shown in Figure 2.1(b), the voltage, or electromotive force is generated when an electric conductor is moved through a magnetic field.



(a) (b)

Figure 2.1: Examples of two KEH devices: (a) PEH, and (b) EEH.

There have been extensive studies on wearable sensors. However, wearable sensors consume power and most existing wearable products are powered by batteries. Therefore, frequent recharging and replacement of the batteries are required, which is a major obstacle for achieving continuous gait recognition. To overcome this problem, researchers are investigating the use of the output signal from KEH to achieve a wide range of applications in activity tracking [72, 73] and health monitoring [77]. In [72, 73], the authors proposed the idea of using the energy harvesting power signal for human activities recognition. Their proposed system can achieve 83% accuracy for activity recognition. In [77], the authors conducted the first experimental study of using the output voltage signal from the PEH to estimate calorie expenditure of human activities. They have shown promising results from replacing accelerometers by KEH

for calorie expenditure. Following this trend, my proposed KEH-Gait utilizes the voltage signal generated by the kinetic energy harvester from walking to perform gait recognition. By doing so, KEH-Gait can reduce the power consumption of the gait recognition in the wearable device by not using the accelerometer.

2.4.3 Summary of the Problem

Most existing studies use accelerometers to record and analyze gait signals. The main problem is that continuous accelerometer sampling drains the battery quickly. High power consumption of accelerometer sampling, which is typically in the order of a few milliwatts, also makes it challenging to adopt gait-based user authentication in resource-constrained wearables. To overcome this problem, we explore the feasibility of of gait recognition using KEH. Compared to conventional accelerometer-based gait detection, KEH-Gait can reduce energy consumption by 78% while achieving comparable recognition accuracy. To the best of my knowledge, this is the first work that experimentally validates the feasibility of gait recognition using KEH, and the evaluation results show that the output voltage signal of energy harvester is a promising informative signal for wearable authentication system. The details of this work will be presented in Chapter 5.

Chapter 3

Sensor-assisted Face Recognition System on Smart Glasses

Chapter Summary: Face recognition is a hot research topic with a variety of application possibilities, including video surveillance and mobile payment. It has been well researched in the traditional computer vision community. However, new research issues arise when it comes to resource constrained devices, such as smart glasses, due to the high computation and energy requirements of the traditional accurate face recognition methods. This chapter describes a robust and efficient sensor-assisted face recognition system on smart glasses by exploring the power of multimodal sensors including the camera and Inertial Measurement Unit (IMU) sensors. The system is based on a novel face recognition algorithm, namely Multi-view Sparse Representation Classification (MVSRC), by exploiting the prolific information among multi-view face images. To improve the efficiency of MVSRC on smart glasses, we propose two novel sampling optimization strategies using less expensive inertial sensors. Our evaluations on public and private datasets show that the proposed method is up to 10% more accurate than the state-of-the-art multi-view face recognition methods while its computation cost is the same order as an efficient benchmark method (e.g., Eigenfaces). Finally, extensive real-world experiments show that our proposed system improves recognition accuracy by up to 15% while achieving the same level of system overhead to the existing face recognition system (OpenCV algorithms) on smart glasses.

3.1 Introduction

Face recognition has emerged as an active research area with numerous applications over the past decades. One of the typical applications of face recognition is to assist people in recognizing identities. The possibility of using wearable devices for deep cognitive assistance (e.g., offering hints for social interaction via real-time scene analysis) was first suggested nearly a decade ago [122, 123] and is becoming the focus of research with the advent of smart glasses such as Google Glass and Vuzix Smart Glass. Smart glasses have advantages over other smart devices as they are equipped with a first-person camera which can be naturally used as a 'third eye' to deliver a significantly better user experience for face recognition.

In this chapter, we aim to develop a robust and efficient face recognition system on smart glasses. Face recognition has been well researched in the computer vision community, yet there still remain many challenges on mobile devices. As discussed in [124], most of the advanced face recognition methods fail on portable smart devices because of the tension between high computation requirements and resource constraints. For instance, the battery life of smart glasses is limited by the battery size. It is reported that the fully charged battery on the Vuzix Smart Glass can last for one hour; however our practical experience shows that the battery would be completely drained within half an hour with display on, camera open and high CPU loading. Moreover, on smart devices, most of the applications involving face recognition are still using the inaccurate but efficient methods in the Open Source Computer Vision (OpenCV) library, e.g, Eigenfaces [130] proposed in 1991. Recently, Shen et al. [124] proposed a new face recognition system: opti-SRC, which is specifically designed for smart phones based on the sparse representation classification (SRC) algorithm [138]. However, as opti-SRC only applies to a single face image, it ignores the rich information enabled by the sensors (accelerometer, gyroscope, magnetometer, etc.) and video camera when used on smart glasses. This additional information may improve the performance of the recognition system and user experience significantly. There have been some recent face recognition systems implemented on smart glasses, e.g., Gabriel [56]. Gabriel shifts the computation burden to a cloudlet (local server) or cloud from the smart glasses while the smart glasses are only used for image capture and display of results. Gabriel provides assistance services to the users such as face recognition and object recognition. However, the usability of the cloud-based recognition systems relies on the wireless connectivity. The energy cost of wireless transmission depends greatly on the quality of

the wireless connection. Furthermore, according to the results reported in [28], the power consumption of wireless transmission on a smartphone(about 720mW) is approximately 12 times higher than that of CPU (about 60mW). Therefore, the *in-situ* approaches are preferable considering the relatively high cost of wireless transmission and the inconvenience of relying on wireless connections.

To overcome the challenges and exploit the useful information provided by smart glasses, we propose and implement a novel sensor-assisted face recognition system which runs locally on smart glasses. This exploits the information from both the camera and sensors on smart glasses to improve the recognition accuracy and reduce the energy consumption. The system recognizes the identities based on face image sequences collected from different view angles and utilizes the IMU sensors to improve its efficiency. To the best of our knowledge, our work is the first to consider *in-situ* face recognition on smart glasses by fusing IMU sensors. The proposed system presents a humble step forward for *in-situ* face recognition on smart glasses. The contributions of this chapter are threefold:

- We propose a novel face recognition algorithm called Multi-view Sparse Representation based Classification (MVSRC). It exploits the high agreement among the sparse representations of the face images from different view angles and applies a novel weighted SRC model to improve the Signal to Noise Ratio (SNR). Our evaluation on several datasets shows that MVSRC outperforms several state-of-the-art multi-view face recognition algorithms.

- We propose a Support Vector Regression (SVR)-based estimation model to relate the recognition accuracy to the angle information obtained by IMU sensors. Then we design two sampling optimization approaches: Maximum Accuracy Sampling Optimization (MASO) and Minimum Energy Sampling Optimization (MESO) based on the estimation model to improve the efficiency of MVSRC while preserving its high recognition accuracy. We refer to MVSRC after sampling optimization as fast-MVSRC.

- We implement a face recognition system based on the proposed methods on smart glasses and demonstrate that it significantly outperforms the existing *in-situ* face recognition algorithms on smart glasses. We also discuss the offloading approach and experimentally show that the cost of our system is in the same order of offloading to a nearby server (cloudlet).

The organization of this chapter as follows. We describe the system architecture in Section 3.2. In Section 3.3, we evaluate the performance of the proposed system on several datasets. We then implement the system on smart glasses and conduct real-world experiments to evaluate the system in Section 3.4. Finally, we discuss the feasibility of the system in Section 3.5 and summarize this chapter in Section 3.6.

3.2 System Architecture

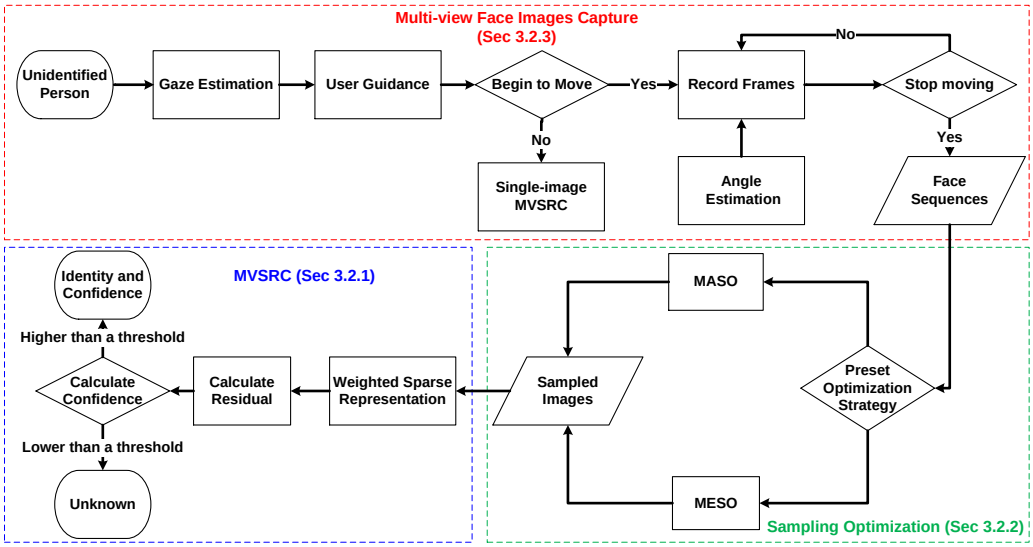


Figure 3.1: The processing pipeline of face recognition system

In this section, we will introduce the proposed system by walking through an example scenario and then describe the system architecture in detail.

Example Scenario. One day at a party, Tom wants to know the name of the man standing near him. Tom moves a few steps around the man and the smart glass pops up the name of Bob on the display. Then Tom says *hi* to Bob and they have a nice conversation.

System Overview. As shown in Figure 3.1, the face recognition system starts with acquiring face images when the user starts to move (i.e., walks a few steps around the subject). The angle information of the face images are estimated by the IMU sensors embedded on the smart glasses. Then the images and the associated angle labels are recorded and stored for further processing. After the user stops recording face images, the sampling optimization algorithm, which will be discussed in Section 3.2.2, will output a subset of face images based on user behavior. Finally, the MVSRC is applied on the samples (i.e. fast-MVSRC) to obtain the classification result and the smart glasses prompt the name on the display.

3.2.1 MVSRC

Multi-view Sparse Representation based Classification (we call it MVSRC for short) is built on single image approaches [138, 124]. The key assumption behind MVSRC is that face images obtained from different view angles tend to have a high agreement between the sparse representations because each of the face images from the same person should be linearly represented by the same class in the dictionary. Suppose we have acquired a set of M feasible face images from the camera. Following the single image approach described in Section 2.1.1, we can obtain a set of estimated coefficients vectors $\hat{X} = \{\hat{x}_1, \hat{x}_2, ..., \hat{x}_M\}$ by solving the ℓ_1 optimization problem for each of the face images. Theoretically, a precise sparse representation will only contain the non-zero entries at the locations related to the specific class. However, noise exists in the empirical estimations. Therefore, the estimated coefficients vector of the test image m can be expressed as:

$$\hat{x}_m = x + \varepsilon_m \tag{3.1}$$

where x is the theoretical sparse representation of the face image vector and ε_m is used to account for noise. The image vector could be misclassified due to low Signal to Noise Ratio (SNR). To enhance the SNR of the classification system, we propose a new sparse representation model by exploiting information from the multi-view face images. The new sparse representation model can be expressed as:

$$\hat{x}_{sum} = \sum_{m=1}^{M} \alpha_m \hat{x}_m \tag{3.2}$$

where α_m is the weight assigned to $\hat{x}_m$ based on the *Sparsity Concentration Index* (SCI) defined in [138]:

$$SCI(\hat{x}_m) = \frac{K \cdot max_j \|\delta_j(\hat{x}_m)\|_1 / \|\hat{x}_m\|_1 - 1}{K - 1} \in [0, 1] \tag{3.3}$$

SCI measures how concentrated the coefficients are in the dictionary. $SCI(\hat{x}_m) = 1$, if the test image can be strictly represented by a linear combination of images from only one class; and $SCI(\hat{x}_m) = 0$, if the coefficients are spread evenly over all classes. The weight of $\hat{x}_m$ is obtained by normalizing the SCIs among the multi-view face images:

$$\alpha_m = SCI(\hat{x}_m) / \sum_{n=1}^{M} SCI(\hat{x}_n) \tag{3.4}$$

In the new face recognition model, the SNR is improved in two aspects.

- As shown in Eq. (3.1), the estimated coefficients vector can be divided into the theoretical part (signal part) and noise part. The theoretical parts among the sparse representations of the multi-view face images from the same identity have a high agreement. However, due to the random nature of the noise, the agreement among the noise signals is low. It is straightforward to prove that the SNR of the face recognition system tends to be improved by summing up the coefficients vectors obtained from conducting sparse representation on different face images.

- Normalized weights assigned to each of the coefficients vectors are derived from their SCIs. SCI is designed to approximate the sparsity of the coefficient vectors. A higher SCI represents a more accurate approximation achieved by solving ℓ_1 optimization. Therefore, the coefficients vector with higher SNR will be assigned a relatively larger weight. Meanwhile a coefficients vector with a high noise level will be depressed by being assigned a smaller weight.

With the knowledge of $\hat{x}_{sum}$, the compressed residual of each class is computed as:

$$r_i(y_{sum}) = \|R_{opt}y_{sum} - R_{opt}A\delta_i(\hat{x}_{sum})\|_2 \tag{3.5}$$

where $y_{sum} = \sum_{m=1}^{M} \alpha_m y_m$ is the weighted summation of all the feasible face image vectors obtained by the glasses. Following the same approach in [138, 124], the final classification result is obtained by finding the minimal residual.

To recognize individuals that are not in the system, we adopt the same principle used in [124] by using confidence level defined as:

$$confidence = \left(\frac{1}{K}\sum_{i=1}^{K} r_i - \min_{i=1,\ldots,K} r_i\right) / \frac{1}{K}\sum_{i=1}^{K} r_i \tag{3.6}$$

The confidence level is in the range of $[0,1]$ and should be close to 1 if a subject is known by the recognition system; otherwise it will be close to 0. An appropriate threshold (0.2 in our system) can be chosen by a data-driven approach to make the recognition system robust to intruders. As a recognition system, our system will always provide a recognition result to the user. The confidence level is just used to indicate the credibility of the final result.

3.2.2 Optimized Sampling Strategy

Problem Statement

Considering the computation and energy consumption issues of the smart glasses, applying MVSRC straightforwardly on all of the M face images is not a desirable choice because it requires operating ℓ_1 optimizations M times. Evaluation in [124] shows that only a single ℓ_1 optimization takes almost 2/3 of the total computation time. Moreover, a large amount of redundant information exists among the adjacent frames as the face images with similar view angles contain large overlaps. This makes a downsampling strategy possible to improve the efficiency of MVSRC while preserving its accuracy.

To find the best sampling strategy, we propose two approaches to optimize the downsampling on the face images set with a predefined energy budget or an accuracy target respectively: the *Energy First* approach and *Accuracy First* approach.

Energy First. In this case, the energy consumption is considered as the first priority. The energy budget E_b is preset and we aim to find the optimal subset Ω_s of the face images set I to achieve the highest recognition accuracy A_c by solving the optimization problem below:

$$\Omega_s = \arg\max_{\Omega} A_c \quad \text{s.t. } E_{total} \leq E_b, \Omega \subseteq I \tag{3.7}$$

where Ω is one of the arbitrary subsets of I and E_{total} is the total energy consumption for face recognition.

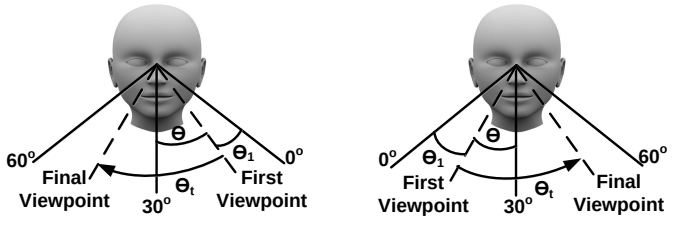
Accuracy First. In this case, the accuracy target A_t is regarded as the first priority. The total energy consumption E_{total} is minimized by solving the optimization problem while achieving the accuracy target:

$$\Omega_s = \arg\min_{\Omega} E_{total} \quad \text{s.t. } A_c \geq A_t, \Omega \subseteq I \tag{3.8}$$

To solve the above optimization problems, we start with analyzing the parameters affecting the face recognition accuracy. According to the processes of face recognition with smart glasses, we define a potential parameters list X and aim to relate this list to recognition accuracy by machine learning techniques. The parameters included in X in our system must satisfy two conditions: 1) it can be quantified and 2) it can be estimated by sensors on smart glasses. Using these two conditions, we build the list $X = (\theta_1, \theta_t, \theta_{s1}, \theta_{si}, n_s)$ consisting of the following parameter variables:

- θ_1: the view angle of the first recorded face image which is estimated by image processing method.
- θ_t: the total rotation angle displacement between the leftmost (rightmost) and rightmost (leftmost) face images in the yaw direction and is estimated by the IMU sensors.
- θ_{s1}: the view angle of the first face image in the chosen subset and is estimated by combining the result of θ_1 and analysis on IMU sensor readings.
- θ_{si}: the view angle interval among the face images in the chosen subset and is estimated by IMU sensors.
- n_s: the number of face images in the chosen subset.

The illustrative explanation of the parameter variables is shown in Figure 3.2 (θ is obtained by gaze estimation in Section 3.2.3). As the evaluation in Section 3.3.2, the feasible range of the view angle is between 30° to the left and 30° to the right of the frontal face respectively. The results are also consistent to the symmetric property of the human face. Therefore the original angle (0°) can be either the 30° view angle to the right (as shown in Figure 3.2(a)) or the 30° view angle to the left (as shown in Figure 3.2(b)). We choose the origin at the same side as the view angle of the first recorded face image.



(a) θ_1 at the left profile face (b) θ_1 at the right profile face

Figure 3.2: Angle coordinate system settings: θ is the relative view angle of the first face image to the frontal face, θ_1 is the view angle of the first face image to the origin ($\theta_1 = 30° - \theta$), θ_t is the rotation angle displacement between the first and last face images.

As the exact recognition accuracy cannot be computed without the knowledge of the groundtruth in a real world application, to estimate recognition accuracy, we model the correlation between the parameter variables and the recognition accuracy based on a novel Support Vector Regression (SVR)-based approach [126] to find the optimal observation of the parameters. The estimation model is learned offline and then used for *in-situ* accuracy estimation. We use our private

dataset (see Section 3.3 for the details of private dataset) which consists of 10 subjects to learn the estimation model. Each of the subject contributes 9 image sequences and each of the image sequences contains 61 face images from different view angles. In the following parts, we will describe how to build the estimation model.

SVR-based Estimation Model

We define the set of all the possible observations of X as $\{\chi_1, \chi_2, \chi_3, ..., \chi_N\}$ and the corresponding accuracies as $\{z_1, z_2, z_3, ..., z_N\}$. Each of the observations is related to a certain subset of the face images which is determined by the values of the parameters in X. With the information of the observations and the corresponding accuracies, we aim to find the function $f(\cdot)$ which best approximates the relation inherited between the input features X and it can be used later on to infer the accuracy z for a new input feature X. Specifically, the goal of regression is to find the function $f(\cdot)$ which relates the parameters list X to the recognition accuracy z with the deviation of at most ε:

$$Dev(z, f(X)) \leq \varepsilon \tag{3.9}$$

where $Dev(\cdot,\cdot)$ represents the deviation computation. We apply SVR [126] by using all the possible observations in the private dataset to find the function $f(X)$ and we use the Radial Basis Function (RBF) Kernel which is defined as:

$$k(x_i, x) = e^{-\gamma \|x_i - x\|^2} \tag{3.10}$$

where γ is a kernel parameter (0.01 in our experiment). For more details of SVR, readers are encouraged to refer to [126] for the step-by-step instructions.

To evaluate the estimation function, we divide the set of pairs of observations and the corresponding value of z into a training dataset $\{(\chi_{D,1}, z_{D,1}), (\chi_{D,2}, z_{D,2})..., (\chi_{D,L}, z_{D,L})\}$, and a test dataset $\{(\chi_{T,1}, z_{T,1}), (\chi_{T,2}, z_{T,2}),..., (\chi_{T,F}, z_{T,F})\}$, where $N = L + F$. The estimation function $f(X)$ is learned from applying SVR on the training dataset. Then we use the test dataset to evaluate the estimation performance. We compute the Mean Square Error (MSE) of $f(X)$ against the ground truth z via a 10-fold cross validation. Specifically, the dataset is equally divided into 10 sets. Then each of the 10 sets is selected as the test dataset (the remaining 9 sets act as training dataset) and the corresponding MSE is calculated. The final MSE over the 10 sets is only 0.0034 which indicates the function $f(X)$ learned from applying SVR-based approach on

Algorithm 1 Maximum Accuracy Sampling Optimization

1: Input: Estimation model f, total energy consumption E_{total}, energy budget E_b, angle of the first view θ_1, total rotation angle displacement θ_t, angle of the first sampled image θ_{s1}, interval between sampled images θ_{si}, number of sampled images n_s, maximum number of sampled images $N_{max} = [(min(\theta_1 + \theta_t, 60) - \theta_{s1})/\theta_{si}]$.
2: Initialization: allocate one empty list: Y, $m = 0$.
3: **for** $n_s = 1 : N_{max}$ **do**
4: **for** $\theta_{si} = 0 : \theta_t$ **do**
5: **for** $\theta_{s1} = \theta_1 : min(\theta_1 + \theta_t, 60)$ **do**
6: **if** $(E_{total} \leq E_b)$ **then**
7: $Y_m = f(\theta_1, \theta_t, \theta_{s1}, \theta_{si}, n_s)$
8: m++
9: **end if**
10: **end for**
11: **end for**
12: **end for**
13: $(\theta_{s1}, \theta_{si}, n_s) = \underset{\theta_{s1}, \theta_{si}, n_s}{\operatorname{argmax}} Y$
14: Output: $(\theta_{s1}, \theta_{si}, n_s)$

the training dataset can provide accurate estimation with the knowledge of the observations of the parameters list.

Sampling Optimization

With the knowledge of the estimation function, we propose two computationally efficient approaches to solve the optimization problems Eq. (3.7) and Eq. (3.8) respectively, i.e., Maximum Accuracy Sampling Optimization (MASO) and Minimum Energy Sampling Optimization (MESO). In the real application, θ_1 and θ_t are user-specific and determined before the sampling optimization stage. The optimization approaches are actually searching for the optimal observation of $(\theta_{s1}, \theta_{si}, n_s)$ under the predefined conditions (energy budget or accuracy target). The estimation function is used to efficiently approximate the recognition accuracy with the knowledge of the angle information (Line 7 in algorithm 1 and Line 6 in Algorithm 2).

MASO To solve the optimization problem Eq. (3.7), MASO finds the optimal observation of $(\theta_{s1}, \theta_{si}, n_s)$ (i.e., the sampling strategy) to achieve the highest recognition accuracy under the predefined energy budget as shown in Algorithm 1.

MESO MESO solves the optimization problem Eq. (3.8) by finding the optimal observation of $(\theta_{s1}, \theta_{si}, n_s)$ to minimize the energy consumption under the predefined accuracy target as shown in Algorithm 2.

Algorithm 2 Minimum Energy Sampling Optimization

Input: Estimation model f, total energy consumption E_{total}, accuracy target A_t, angle of the first view θ_1, total rotation angle displacement θ_t, angle of the first sampled image θ_{s1}, interval between sampled images θ_{si}, number of sampled images n_s, maximum number of sampled images $N_{max} = [(min(\theta_1 + \theta_t, 60) - \theta_{s1})/\theta_{si}]$.
Initialization: allocate one empty list: E, $m = 0$.
for $n_s = 1 : N_{max}$ **do**
    for $\theta_{si} = 0 : \theta_t$ **do**
        for $\theta_{s1} = \theta_1 : min(\theta_1 + \theta_t, 60)$ **do**
            if $(f(\theta_1, \theta_t, \theta_{s1}, \theta_{si}, n_s) \geq A_t)$ **then**
                $E_m = E_{total}$
                m++
            end if
        end for
    end for
end for
$(\theta_{s1}, \theta_{si}, n_s) = \underset{\theta_{s1}, \theta_{si}, n_s}{\operatorname{argmin}} E$
Output: $(\theta_{s1}, \theta_{si}, n_s)$

In Algorithm 1 and 2, the total energy consumption of the system can be expressed as:

$$E_{total} = T * (P_{base} + P_{dis} + P_{imu} + P_{cam}) + E_{cpu} \tag{3.11}$$

where T is the total operating time for classification; P_{base} denotes the baseline power consumption of the smart glass; P_{dis}, P_{imu} and P_{cam} are the power consumed by the display, IMU sensors, and camera respectively; E_{cpu} is the total energy consumption of CPU for classification which accounts for face detection, gaze estimation, ℓ_1 optimization, residual calculations and sampling optimization. E_{cpu} can be further split into a repeatable part and a one-time part. The one-time part consists of the energy consumption of gaze estimation, residual calculation and sampling optimization which are operated once only during classification, while the repeatable part includes face detection and ℓ_1 optimization which are operated on every sampled face image. Assuming m face images are sampled for classification, the total energy consumption can be further expressed as:

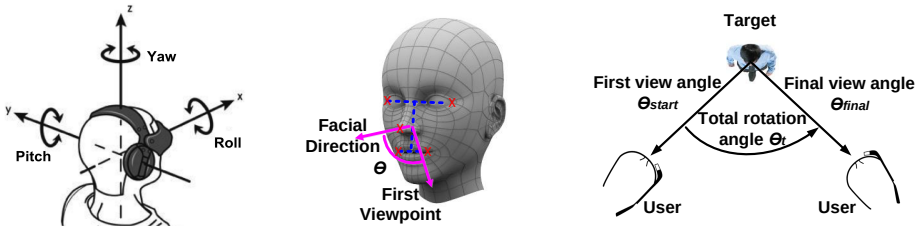
$$E_{total} = T * (P_{base} + P_{dis} + P_{imu} + P_{cam}) + m * E_u + E_1 \tag{3.12}$$

where E_u is unit energy consumption of the repeatable part on each image, E_1 is the energy consumption of the one-time part.

The optimization strategy and corresponding parameter (i.e., E_b in MASO and A_t in MESO) is customized by the user before recognition, and the optimization process is called online after the multi-view face images are obtained. As the results shown in Section 3.3.4, the sampling optimization component only takes less than 2.7% of the total computation time which suggests that our optimization method is computationally efficient and we also noticed that E_u (including face detection and ℓ_1 optimization) is the most energy-consuming component (238mJ), and the aim of the sampling optimization is to reduce the number of images (m in Eq 3.12) which is the proportional coefficient of E_u. Therefore, the proposed sampling strategies can reduce the computational cost of MVSRC significantly. In order to differentiate from MVSRC, we call MVSRC after sampling optimization as fast-MVSRC.

3.2.3 Sensors-assisted System

As described above, the sampling optimization process requires the angle information. The angle information is obtained by gaze estimation of the first face image and angle displacement estimation with IMU sensor readings.



(a) (b) (c)

Figure 3.3: (a) Head model of the user [12] (b) Gaze of the subject (c) Bird view of the recognition process.

Gaze Estimation

Gaze estimation is used to find the initial angle information θ_1 of the first image by the image processing technique proposed in [52]. The method uses the locations of the following five facial features: left and right eye far corners, left and right mouth corners and nose tip which are marked as red crosses in Figure 3.3(b). The angle θ between the view point of the first face image recorded and the frontal view is calculated by analyzing the relative positions of the five facial points. Then θ_1 in the view angle coordinate system can be obtained with the knowledge of θ ($\theta_1 = 30^\circ - \theta$ in our system). After obtaining the initial angle information, the view angles

of the face images recorded later can be calculated by accumulating the angle displacements along with θ_1 as reference.

Angle Displacement Estimation

From Figure 3.3(a) and Figure 3.3(c), we notice that the rotation angle is actually the angle change along yaw direction of the smart glass when the user moves around the subject. In practice, substantial pitch and roll rotations rarely occur. Moreover, the slight pitch rotation caused by the height difference between the subject and user is within the tolerance of the face recognition algorithms. One can estimate rotation angle by simply integrating gyroscope readings. However, the measurements from IMU sensors suffer from bias, noise and systematic errors (e.g., misalignment between the sensor axes and non-unit scale parameters) which lead to inaccurate orientation estimations [82]. To address this issue, we implement a sensor fusion algorithm to compensate for the weakness of each sensor by utilizing other sensors' information. Here we use quaternion-based Extended Kalman Filter (EKF) proposed in [121] to estimate the orientation of the smart glass. The EKF incorporates an in-line calibration procedure for modeling time-varying biases which may affect sensors like accelerometers and magnetometers, and a mechanism for adapting their measurement noise covariance matrix in the presence of motion and magnetic disturbances. Assume the output of EKF is quaternion $q = [w,x,y,z]^T$, we could compute the three Euler angles of head model in Figure 3.3(a) using the following equations:

$$\begin{bmatrix} \varphi \\ \psi \\ \theta \end{bmatrix} = \begin{bmatrix} atan2\left(2\left(wz+xy\right),1-2\left(x^2+z^2\right)\right) \\ asin\left(2\left(wx-yz\right)\right) \\ atan2\left(2\left(wy+xz\right),1-2\left(x^2+y^2\right)\right) \end{bmatrix} \tag{3.13}$$

where φ stands for roll, ψ represents pitch and θ represents yaw rotations respectively.

To improve user experience, the IMU sensor readings are used to automatically determine the start and end of the recognition process. From our observation, the gyroscope data along the yaw direction (perpendicular to the motion) exhibits large variations when the user moves during the recognition process. We first apply a low pass filter to filter out the small vibrations, then our system will start to record face images at θ_{start} when the gyroscope sensor reading along the yaw direction is larger than a threshold (0.15 rad/s in our system) and end the recording at θ_{final} when it is lower than the threshold in the sense that the user stops moving. The rotation angle can be simply obtained by $\theta_t = \left|\theta_{final} - \theta_{start}\right|$ as shown in Figure 3.3(c).

Another challenge is that the timestamps of sensors and video frames are usually not well synchronized [67]. Therefore, we apply the online calibration and synchronization method proposed in [67] to obtain the delay t_d between IMU sensors and camera, then t_d is in return used to synchronize the timestamps of sensor readings and images. For a full description of the EKF-based orientation estimation and synchronization, the reader is referred to [121] and [67] respectively.

After the user stops moving, θ_1, θ_t and face images associated with corresponding angle displacement are used in MASO or MESO which depends on user choice.

3.3 Evaluation

3.3.1 Goals, Metrics and Methodology

In this section, we will evaluate the performance of the proposed system via simulation. The goals of the simulation are fourfold: 1) to determine the choice of the key parameters including feasible range of views in the yaw direction and the number of projections used in MVSRC; 2) whether MVSRC outperforms the state-of-the-art face recognition methods in accuracy; 3) whether the proposed sampling strategies improve the efficiency of MVSRC while retaining high accuracy; and 4) to evaluate the angle estimation accuracy of IMU-based method and its impact on face recognition accuracy.

The evaluations are based on two datasets: Honda/UCSD video dataset (Honda/USCD) [78] and the private dataset we have collected with the smart glasses[1]. Honda/USCD video dataset is widely used for the evaluation of multi-view face recognition methods. It consists of 59 image sequences from 20 subjects recorded in different environments and each subject contributes at least two sequences. The number of frames of the sequences vary from 12 to 645. The angle information is not available in Honda/UCSD dataset, therefore we built our private dataset by obtaining both multi-view face images and their associated view angles. Our private dataset consists of 10 subjects (2 females and 8 males) aged from 24 to 43 with different skin tones. The face images are taken under 9 different categories by combining the different expressions (neutral, happy and sad) and locations (corridor, office and outdoor). The user wearing the smart glass records the video clips of the candidate to be recognized (suppose the candidate is

[1]Ethical approval for carrying out this experiment has been granted by the corresponding organization (Approval Number 2014000589)

just facing to the user) by moving around the subject from left to right (in yaw direction) with wide range. The flow of orientation information is obtained and synchronized with the video clips. Face regions are detected by a Viola-Jones face detector [134] and cropped to 48×48 gray-scale images. We then apply the method introduced in Section 3.2.3 to find the frame containing the face in frontal view angle. Finally we sub-sample the video clips by every 1° according to the associated angle displacement information until we reach 60° to both left and right direction. Therefore, for each video clip, we obtain a symmetric sequence of 121 face images with view angles from -60° (left) to $+60^\circ$ (right). A sequence of sample images is shown in Figure 3.4.

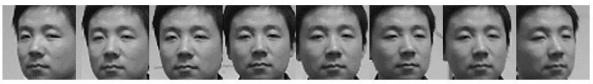


Figure 3.4: Examples of face images from private dataset.

As the proposed system is a multi-classification system, we focus on evaluating recognition accuracy rather than positive/false negative which are commonly used metrics in an authentication system. We determine the parameters (feasible angle range and number of projections) by gradually changing the parameters and the choices are made according to the evaluation results on the real datasets. In the evaluation of this section, the training set is derived from random selection as in [124]. We show the results of the average value and 95% confidence level of the performance metrics (accuracy, energy consumption and computation time) over 30 independent trails. The computation time and energy consumption are measured by running the system on Vuzix M100 Smart Glass.

3.3.2 Parameters Choice

In this section, we will determine the choice of the parameters including the feasible angle range and the number of projections applied in the system by evaluations on the real datasets.

Impact of View Angles

The view angles have substantial impact on the recognition accuracy. In this experiment, we evaluate the influence of different view angles on the recognition accuracy on the private dataset.

As described previously, we obtain a symmetric sequence of 121 face images with view angles from -60° (left) to $+60^\circ$ (right) for each video clip. We group the face images by the

view angles uniformly into 12 bins by every 10° and the frontal faces are picked up to form the 13th bin. Each bin represents an evaluation point. We calculate the recognition accuracy of each bin by using three single-image based face recognition methods: opti-SRC, rand-SRC and Eignfaces. We display the evaluation points at the medium degree of each bin (x-axis) in Figure 3.5(a). From the results, we observe that the recognition accuracy decreases when the view angles of the face images deviate from the frontal view angle and the recognition accuracy has dropped significantly when the view angle is over 30° apart from the frontal view. Therefore, we determine the feasible range of the view angles in our system as $[-30^\circ, 30^\circ]$ (the origin of the view angle is the frontal face). In addition, the work in [146, 132] also studies the effects of different poses on face recognition and their findings support our results. With this observation, we remove the images in our private dataset whose view angles are not in the feasible range. Therefore, the number of image in each sequence becomes to 61.

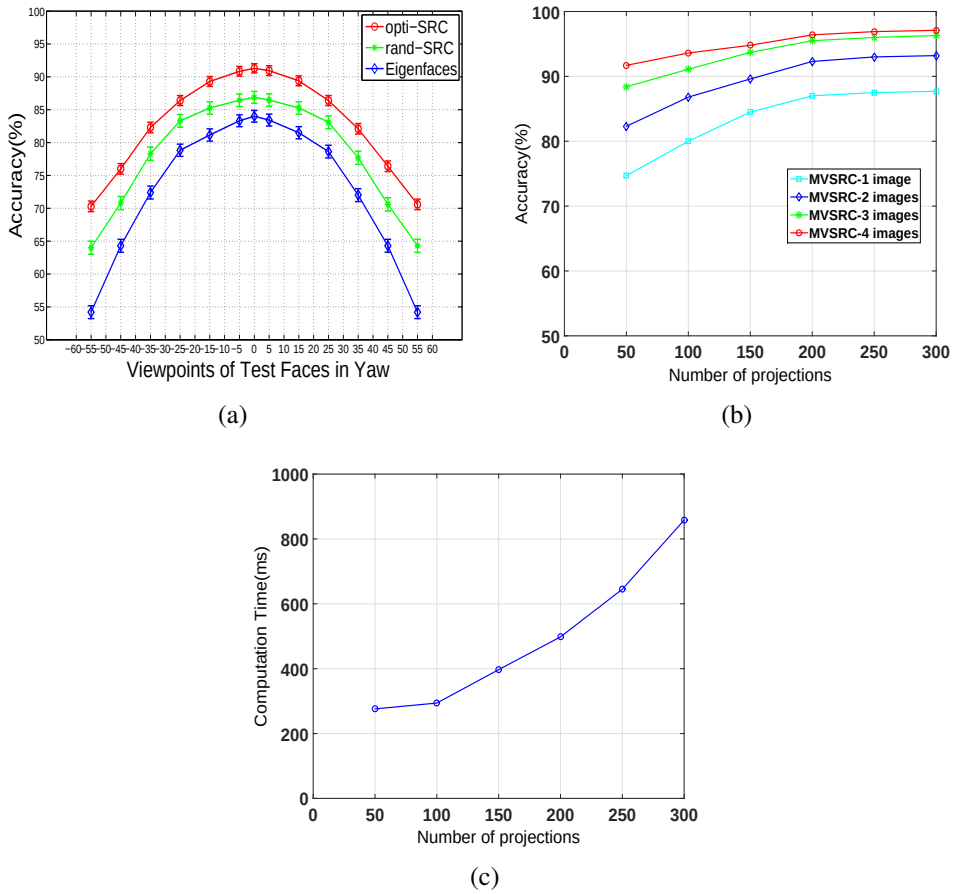


Figure 3.5: Experimental results of parameters choice.

Impact of Number of Projections

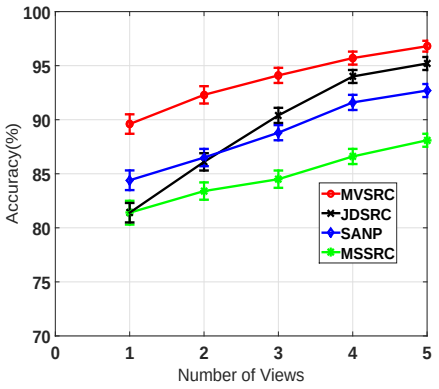
It is known that the recognition accuracy can be improved by increasing the number of projections or features. However, this also increases the computation cost significantly. To investigate the recognition accuracy on the number of projections, we evaluate the performance of MVSRC with different settings by varying the number of projections from 50 to 300. As MVSRC uses multiple face images to perform recognition, we calculate the accuracy of MVSRC with different number of *views* (the number of face images from different view angles for each classification) $n_{view} = 1, 2, 3, 4$. We group the face images of each image sequence in the test set into small subsets of n_{view} images and report the recognition accuracy of MVSRC on the small subsets of different sizes in Figure 3.5(b). We also evaluate the computation time of MVSRC with different number of projections. As the computation time of MVSRC is proportional to n_{view}, without loss of generality, we present the computation time of MVSRC when $n_{view} = 1$. From the results shown in Figure 3.5(b) and Figure 3.5(c), we find the growth of the recognition accuracy diminishes when the number of projections is above 200 while computation time keeps increasing substantially. Therefore, we choose the number of projections as 200.

3.3.3 Dataset Evaluation of MVSRC

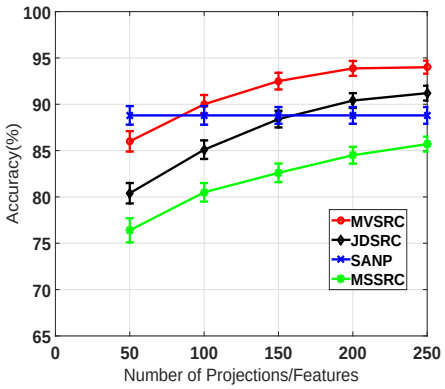
In this section, we compare MVSRC with several competing face recognition methods in the literature. Note that we do not consider angle displacement information in this section.

Comparison with State-of-the-Art

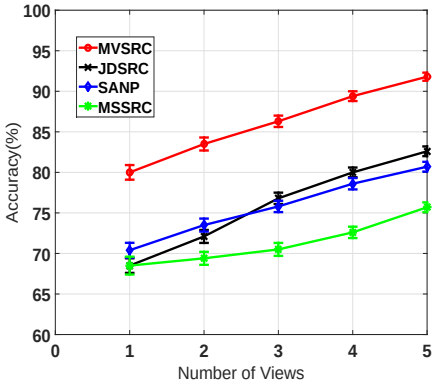
We compare MVSRC with three state-of-the-art multi-view face recognition methods, namely, JDSRC [144], SANP [62] and MSSRC [109]. We compute the recognition accuracy of different methods under different number of views (k) as well as different number of projections/features (d) on the private dataset and Honda/UCSD dataset respectively. For each dataset, we randomly choose 30 images from each subject to form the training set and the rest are used as the test set. We first evaluate the accuracy with different numbers of views from 1 to 5 by setting $d = 200$. We then evalulate the accuracy of different methods against the number of projections/features from 50 to 200 with $k = 3$. Figure 3.6(a) and Figure 3.6(b) plot the results of the private dataset. The results on Honda/UCSD dataset are shown in Figure 3.6(c) and Figure 3.6(d). Note that SANP is not a feature-based method, so the accuracy of SANP in Figure 3.6(b) and



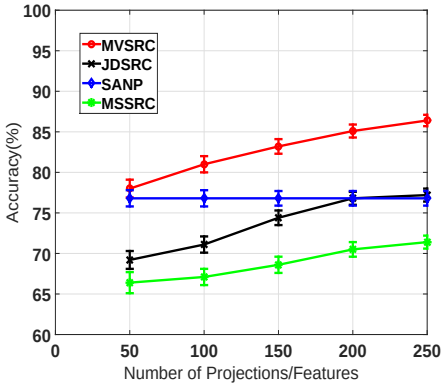
(a) MVSRC: private dataset



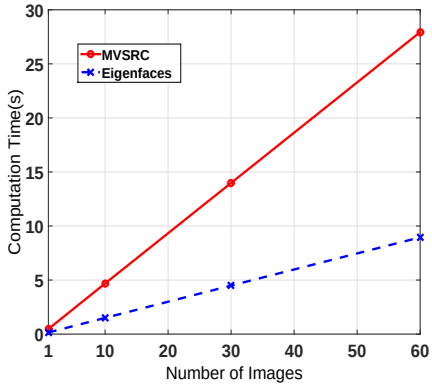
(b) MVSRC: private dataset



(c) MVSRC: Honda/UCSD dataset



(d) MVSRC: Honda/UCSD dataset



(e) MVSRC: Computation time

Figure 3.6: Evaluation results

Figure 3.6(d) is shown by a straight line.

From the results, we can see that MVSRC consistently achieves the best recognition accuracy and is up to 7% and 10% more accurate compared to the second best recognition method on the two datasets respectively. MVSRC, JDSRC and MSSRC are based on original SRC; however, we notice that MVSRC performs better than JDSRC and SANP when k=1. This is due to the fact that single-image MVSRC becomes opti-SRC and opti-SRC performs better than

SRC. We observed that the recognition accuracy of MSSRC is much lower than that reported in [109], because in [109], images are first eye-aligned using eye locations and normalized, then histogram equalization is performed, finally Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG) descriptors are extracted. However, these operations introduce extra system cost on smart glasses. In our evaluation, the face images are directly used for face recognition and no further pre-processing such as alignment or background removal is performed.

Computation Time Evaluation

The computation time is a crucial factor for face recognition systems because the users expect a real-time response. Eigenfaces is known to be efficient and is the most popular method used on resource-constrained devices. We use our private dataset to evaluate the computation time for MVSRC and Eigenfaces (with majority voting) on smart glasses with various sizes of image sequence from 1 to 60. The cost of the two methods is represented by the computation time used for one classification operation. The results in Figure 3.6(e) demonstrate that MVSRC requires significantly more computation time than Eigenfaces and the gap increases with the growth of the number of images used for recognition. However, we will show in the following section that the computation time of MVSRC can be reduced significantly while preserving high accuracy with the proposed sampling optimization method.

3.3.4 Dataset Evaluation of Optimization Strategies

To address the computation issue of MVSRC, we propose fast-MVSRC by combining MVSRC with the optimized sampling strategies described in Section 3.2.2. In this section, we start with some preliminary experiments to find the computation and energy cost information of our system. Then we evaluate the performance of fast-MVSRC using the proposed optimization algorithms and compare it with other common sampling strategies. Finally, we compare the recognition accuracy of fast-MVSRC with Eigenfaces under various computation cost on smart glass.

Preliminary Experiments

As the optimization methods proposed in Section 3.2.2 require the energy consumption information, we conduct preliminary experiments on the Vuzix Smart Glass to obtain the energy

consumption and computation time information. It is worth mentioning that the proposed system is not platform specific and is compatible with Google Glass as well.

In the preliminary experiments, we first evaluate the impact of image resolution and Frame Per Second (FPS) on system cost. Table 3.1 shows that the cost of face detection improves significantly with the increase of image resolution. Image downsampling reduces the system cost, however, it also leads to low recognition accuracy. Note that the recognition accuracy shown in Table 3.1 are the mean results of single-image MVSRC on private dataset without sampling optimization. The original image resolution of Vuzix Smart Glass is 432×240. As shown in Table 3.1, we found that the recognition accuracy drops significantly when the image is downsampled to 108x60 (4 times downsampling in both dimensions). Thus the raw image is downsampled to 216×120 (1/2 downsample) in the prototype. Table 3.2 illustrates the display power, camera power and mean angle estimation error under different FPS. To balance the system cost and the accuracy of angle estimation, we set FPS to 24 in the prototype system.

Table 3.1: System cost of face detection operation under different image resolution.

Resolution	Time(ms)	Energy(mJ)	Accuracy(%)
432*240	175	107	88.1
216*120	85	56	86.8
108*60	63	34	78.4

Table 3.2: Display power and camera power under different FPS

FPS	P_{dis}(mW)	P_{cam}(mW)	Angle Estimation error
27	265	177	1.9°
24	220	122	1.9°
20	204	112	2.9°
15	187	105	3.8°
10	162	92	5.7°

Table 3.3: Resource consumption on Vuzix Smart Glass

Operations	Time(ms)	Operations	Power(mW)			
Face Detection	85	Baseline	35			
Gaze Estimation	87	Display	220			
ℓ_1	350	Camera	122			
Residual	33					
	Energy(mJ)	Sampling rate	NORMAL	UI	GAME	FASTEST
E_u	238	Frequency(Hz)	5	15	50	205
E_1	62	IMU power	11	29	61	295

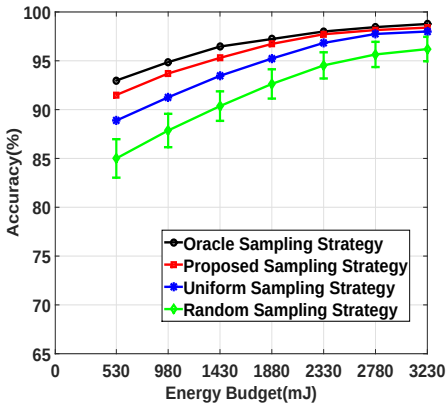
After image resolution and FPS are determined, we evaluate the resource consumption of

each component in the system. Table 3.3 shows the related specifications and resource consumptions on Vuzix Smart Glass. The computation time is obtained from the console of the Eclipse development environment and the energy consumption of each component is estimated by PowerTutor App (it was also used in [124]). The sampling rate of the IMU sensors can be set via Android API and is in four levels from low to high: NORMAL, UI, GAME and FASTEST. Considering both the energy consumption and the accuracy of the synchronization, we choose the sampling rate of the IMU sensors as GAME.

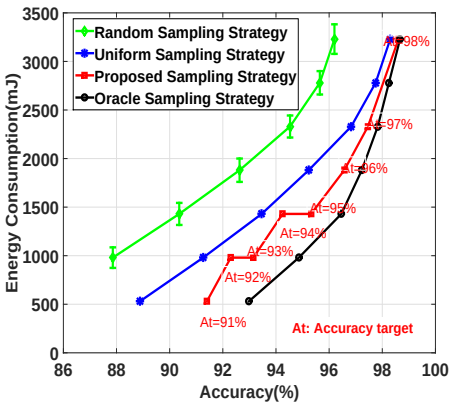
With the resource consumption information above, we evaluate the performance of fast-MVSRC with the two sampling optimization methods in the following sections.

Dataset Evaluation of MASO

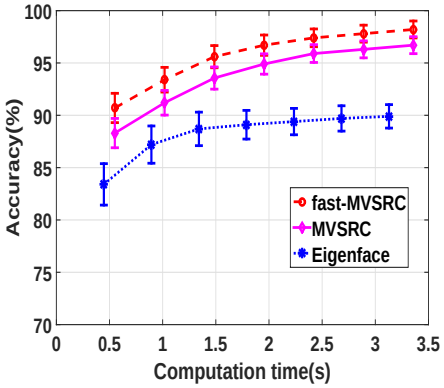
In this section, we compare the recognition accuracy of fast-MVSRC (MASO version) with different sampling strategies under different energy consumption budgets. The sampling strategies include the proposed algorithm MASO, random sampling strategy, uniform sampling strategy and oracle sampling strategy. For the random sampling strategy, we randomly choose a subset of the image sequences. The energy consumption of MVSRC with this subset should satisfy the budget. For the uniform sampling strategy, we divide the image sequence into uniform groups and select the face image in the middle of the group as the representative. We vary the energy budget from 530mJ to 3230mJ for each classification. We consider an offline oracle optimal strategy that provides an upper bound on recognition accuracy for a given energy budget. In terms of oracle sampling strategy, we calculate the recognition accuracy of MVSRC with all possible subsets to find the most accurate one. However it is not applicable for real-world applications as the recognition system cannot compute the recognition accuracy for each of the possible subsets without the knowledge of the groundtruth (the identity of each face image obtained manually). From the results in Figure 3.7(a), we can see that fast-MVSRC with MASO is comparable to the oracle sampling strategy, and achieves higher accuracy than the random and uniform approaches with the same energy budget. We also notice that our approach performs much better compared to random sampling strategy and uniform sampling strategy when the energy budget is limited which is often the case on smart glasses.



(a) MASO: Private Dataset



(b) MESO: Private Dataset



(c) Accuracy-Efficiency Curve

Figure 3.7: Evaluation results of optimization strategies.

Dataset Evaluation of MESO

Different from MASO, MESO aims to find the subset of face images to minimize the energy consumption on the premise of an accuracy target. We again compare fast-MVSRC (MESO version) with the random sampling, uniform sampling and oracle sampling strategy. The recognition accuracy target A_t varies from 91% to 98%. However, for random sampling strategy and uniform sampling strategy, we are not able to control the accuracy target because the accuracy estimation model is not applicable for these two strategies. We compute recognition accuracy of the random and uniform sampling strategies with different energy consumptions and the results are shown in Figure 3.7(b). We also consider an oracle optimal strategy that provides an lower bound on energy consumption to achieve a certain accuracy target. For a certain accuracy, the energy consumption of oracle sampling strategy is calculated offline by performing MVSRC on all possible subsets and choosing the one that consumes minimum energy. The accuracy target A_t of our method is labeled in the figure for reference, and we find that all the accuracy targets

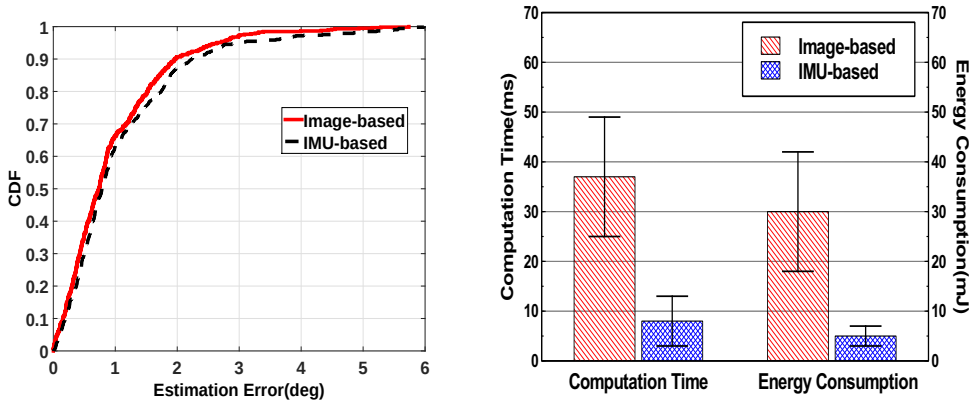
are achieved by fast-MVSRC with MESO. Figure 3.7(b) shows that performance of our approach is significantly closer to the oracle sampling strategy and saves up to 500mJ and 1200mJ per classification compared to the uniform and random sampling strategies respectively. We also notice that the energy consumption keeps almost the same as the accuracy target increases at some evaluation points. This is because the subsets with the same number of face images may produce different accuracy due to different view angle settings. The step behavior shown in Figure 3.7(b) may be due to insufficient dataset being used. We will study the step behavior in our future work.

Fast-MVSRC v.s. Eigenfaces

To demonstrate the effectiveness of fast-MVSRC, we compare the *Efficiency-Accuracy* performance of fast-MVSRC (MASO version), MVSRC and Eigenfaces. We define the *Efficiency-Accuracy* performance as the recognition accuracy with respect to the computation time. We calculate the accuracy of fast-MVSRC and Eigenfaces with majority voting data fusion for multiple images under different computation time. The computation time is varied by using different number of face images for classification. From the results in Figure 3.7(c), we can see the recognition accuracy of fast-MVSRC is up to 9% better than Eigenfaces under the same computation time. Fast-MVSRC performs better than MVSRC under the same computation time (i.e., the same number of images) because MASO chooses the optimal face images from the image sequence. Another important observation is that the growth of the recognition accuracy of fast-MVSRC diminishes when the computation time is around 1.4s–1.8s which indicates fast-MVSRC enables quick response. The results in Figure 3.6(e) and Figure 3.7(c) show that fast-MVSRC improves the efficiency of MVSRC significantly while preserving high recognition accuracy.

3.3.5 Evaluation of IMU-based Gaze Estimation

In this section, we evaluate the performance of IMU-based gaze estimation method and the impact of estimation error on the face recognition accuracy. We also evaluate the impact of total angle displacement on face recognition accuracy.



(a) (b)

Figure 3.8: Comparison with Image-based Gaze Estimation.

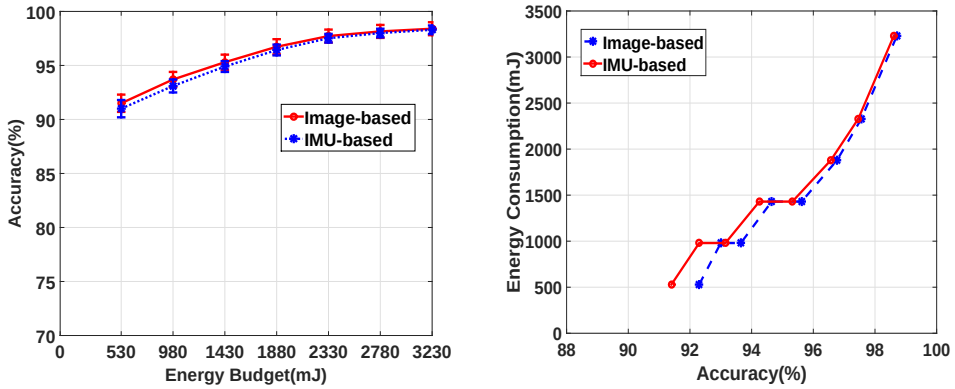
Comparison with Image-based Gaze Estimation

In this part, we compare the estimation accuracy and resource consumption of IMU-based gaze estimation used in our system and image-based gaze estimation proposed in [52]. The results in [52] show that the image-based method can achieve a mean angle estimation error of 2.5° and a maximum estimation error of 6° in 1000 samples of noisy face images. We randomly select 30 face images from each subject to form the comparison image set and use the angle information obtained in Section 3.3.2 as the corresponding estimated value of the IMU-based method. For image-based gaze estimation, we perform facial features detection first and use the method in [52] to estimate the gaze. Facial features are detected by a state-of-the-art facial landmark detector *flandmark* [131]. The ground truth are obtained by annotating facial features manually and then performing the method in [52]. From Figure 3.8(a) and Figure 3.8(b), we can see that our method reduces computation time by 65% and energy consumption by 78% respectively, while achieving comparable accuracy to the image-based gaze estimation method.

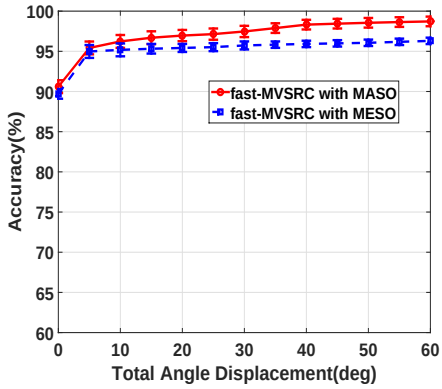
Impact of Angle Estimation Error

As shown in Figure 3.8(a), the estimation errors for most of the face images (over 95%) are within 3°. Therefore, it is important to know the impact of the estimation errors on the recognition accuracy.

We first evaluate the impact of estimation error on MASO. As described in Section 3.3.1, each subject in the private dataset has 9 image sequences collected in different categorizes. We randomly select 5 image sequences from each subject to form a training dataset and use the



(a) (b)



(c)

Figure 3.9: Evaluation results: (a) Impact of estimation error on MASO. (b) Impact of estimation error on MESO. (c) Impact of different total angle displacement.

rest sequences as testing data. We apply MASO on each testing sequence and obtain the angles of the sampled images. Then we select corresponding images in the test sequence according to their angle information. The angles of the testing images are obtained from two methods, i.e., IMU-based method and image-based method. We vary the energy budget from 530mJ to 3230mJ and calculate the corresponding recognition accuracy of IMU-based method and image-based method respectively. Then we conduct the same experiment procedures to evaluate the impact of estimation error on recognition accuracy by applying MESO. From Figure 3.9(a) and Figure 3.9(b), we can see that IMU-based method achieves comparable accuracy to image-based method. Therefore, we conclude that the minor errors introduced by the IMU based angle estimation will not have noticeable influence on the recognition accuracy (less than 1.3%).

Impact of Total Angle Displacement

As the total angle displacement between the starting view angle and the ending view angle varies in the practical use, we evaluate the impact of the total angle displacement on the recognition accuracy. In this experiment, we set $E_b = 2330mJ$ for MASO as from Figure 3.7(a) we notice that the accuracy levels off after $E_b \geq 2330mJ$. For MESO, we set $A_t = 95\%$ to correspond to the settings in Section 4.7. We gradually increase the total angle displacement from 0° to 60° by every 5° and the recognition accuracy is calculated. Figure 3.9(c) shows that the recognition accuracy increases with the growth of rotation angle displacement which suggests user can get higher accuracy if their motion covers larger view range. We also find that the total angle displacement for high recognition accuracy (95%) can be as small as 5° which enables a very short image collection phase (200ms in our user study).

3.4 Real-world Experiments

3.4.1 System Implementation

The prototype of our proposed face recognition system is implemented on Vuzix M100 smart glass[2]. The CPU is an OMAP4460 at 1.2GHz and the operating system is Android 4.0.4. It is equipped with a 5-megapixel camera and the images captured in our system are 216×120@24fps. We use the hardware face detection of OMAP for efficient operation and facial features (i.e. nose tip, eye outer corners and mouth corners) are detected by *flandmark* [131]. The efficient implement of ℓ_1 optimization algorithm ℓ_1-*Homotopy* [43] is used as [124], and its complexity is $O(k^3 + kmn)$, where k is the sparsity of the solution ($k \ll n$), m is the number of equations, and n is the number of unknowns, i.e., the number of columns in the training dictionary.

3.4.2 Experimental Description

We recruited 15 volunteers: 5 users and 10 subjects in the training set. The 10 subjects are the same as our private dataset which was collected under different environments. We select 30 face images from each subject to form the training set. Therefore, the training dictionary is a matrix of size 2304×300 (face image is resized to $48 \times 48 = 2304$). The experiments

[2]A video demonstration of the system can be found at the following URL:https://www.youtube.com/watch?v=lVRS4e3Glho

are conducted in two different locations, in an indoor office environment and outdoors. The lighting conditions are quite different for indoor (200-400 lux) and outdoor (over 1,000 lux) environments. Different lighting conditions are applied for indoor experiment (front-lighting, back-lighting and uniform lighting) and outdoor experiment (front-lighting and backlighting). Thus, the experiments are divided into 5 categories. For each category, users conducted two recognition attempts for each of the subjects. Therefore, we obtain 500 independent recognition results. The energy budget (E_b in MASO) is set as 550mJ (the actual energy consumption of our system was around 520mJ), and the accuracy target (A_t in MESO) is set as 95%.

We also implement the OpenCV face recognition algorithms (OpenCV-2.4.9) on Vuzix M100 smart glass as a benchmark. OpenCV provides three face recognition methods, namely EigenFaces [130], FisherFace [16] and LBPFace [10] in its library. In the experiments, we found that these three methods achieve comparable performance in terms of recognition accuracy and computational cost.

3.4.3 Experimental Results

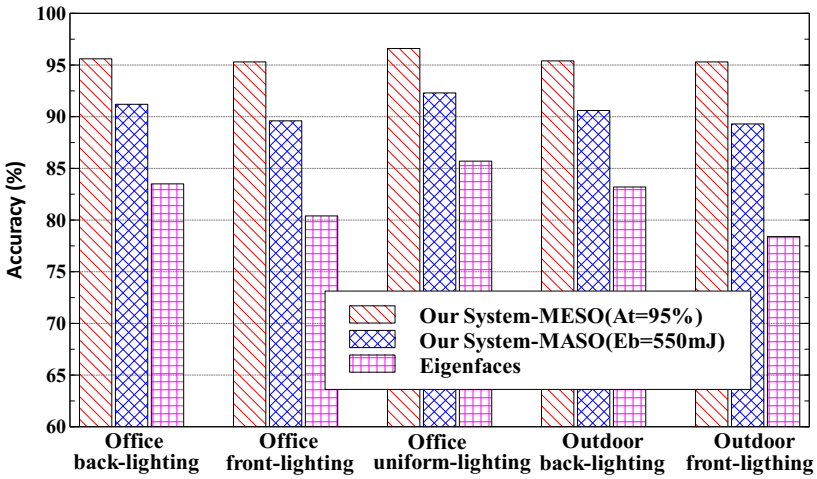


Figure 3.10: Recognition accuracy

The recognition accuracy of our system and Eigenfaces in different experimental categories are shown in Figure 3.10. The proposed system is very stable under different lighting conditions. When MASO is employed, our system outperforms Eigenfaces significantly in every experimental category and is up to 15% more accurate than Eigenfaces under outdoor front-lighting condition. When MESO is employed, the actual recognition accuracy of our system achieves the preset accuracy target in all lighting conditions. We also evaluate and compare the

system overhead of our system with Eigenfaces. From the results in Table 3.4, we can see that the cost of the proposed system (MASO) is in the same order of that of Eigenfaces. The resource consumption of the system when using MESO is higher than that of Eigenfaces because we set a high accuracy target.

Table 3.4: System overhead

Statistic	MASO	MESO	Eigenfaces
Computation Time	516-582ms	1470-1622ms	247-331ms
Energy Consumption	506-535mJ	1400-1466mJ	316-410mJ
Expected Battery Life	≈0.28hr	≈0.1hr	≈0.37hr
Memory Usage	55-64 MB		38-40 MB

3.5 Discussion

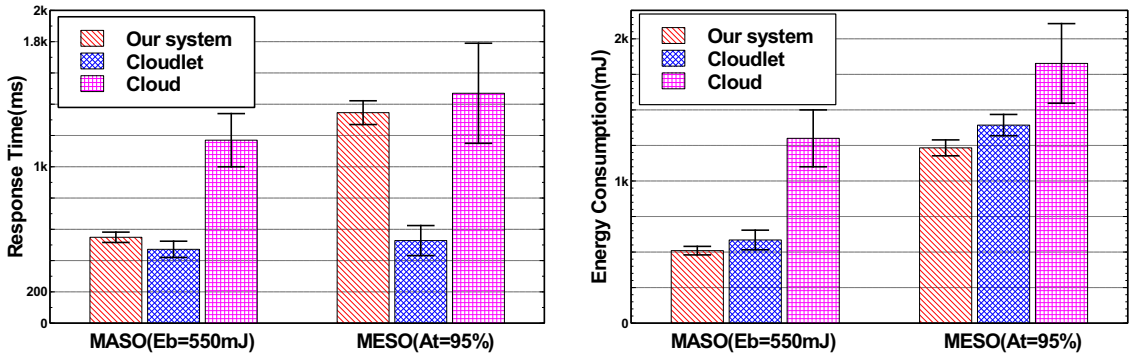
Feasibility The implementation of the system takes advantage of the following assumption: the subject to be recognized remains still and the user needs to move subtly to the left (right) of the target to capture multi-view images. Such assumption may cause inconvenience in practical scenarios. However, as the evaluation results in Section 3.3.5, a total angle displacement of 5° is sufficient to obtain a reliable recognition result (over 95%) and it only takes approximately 200ms for image collection. We believe it only requires small efforts of the user and subjects for normal cases. If the user remains static, single-image MVSRC will be adopted. However, if the user is willing to make extra small efforts with sacrificing user experience on one hand, the significant higher recognition accuracy will significantly improve user experience on the other hand. The proposed system provides such options to users. In practice, face recognition may be applied in a more complex scenario, such as when the user or subject is sitting. We defer face recognition in these scenarios to our future work.

Offload vs. In-situ Offloading computationally intensive operations from mobile devices to powerful infrastructure is a common strategy to reduce computation burden on resource constrained devices. In terms of offloading approaches, the smart glasses are used to capture images and perform sampling optimization, then the sampled images are transmitted to the server via a wireless network. Results obtained by running MVSRC on the server are sent back to the smart

Table 3.5: Server hardware specifications

Offload Strategy	CPU	RAM
Cloudlet	Intel Core 2.7Ghz 2cores	8GB
Cloud	Intel Xeon 2.5Ghz 1VCPU	1GB

glasses. We evaluated the response time and energy consumption of smart glasses by transmitting raw sampled images under two different offloading approaches: cloudlet and remote cloud. Hardware specifications of different offloading strategies are shown in Table 3.5. The cloudlet is implemented inside a Virtual Machine (VM) managed by Vmware Workstation on a Windows 7 host. We use Amazon EC2 VM instances located in Sydney as a remote cloud. The wireless network is based on a campus WiFi.



(a) Response time (b) Energy consumption

Figure 3.11: Comparison of different offloading approaches.

Figure 3.11 presents the response time and energy consumption of the smart glasses using different approaches. We can see a significant drop in both latency and energy consumption when switching from cloud to cloudlet. The performance of offloading to a remote cloud depends greatly on the network conditions. We also find that the cost of our proposed system (*in-situ*) is comparable to the offloading approach with cloudlet and is significantly lower than that with the remote cloud. It is worth mentioning that the energy consumption of the proposed system largely depends on the user-specified parameter settings of optimization strategy, i.e., E_b in MASO and A_t in MESO. Meanwhile, we also note that more advanced recognition methods such as 3D techniques can be achieved on a powerful server. However, offloading approaches require extra infrastructure and system cost. Furthermore, the proposed system has advantages over offloading strategies when the network is unavailable or in poor quality.

3.6 Summary of This Chapter

In this chapter, we explored the capability of smart glasses and proposed a novel face recognition system which utilizes the power of multimodal sensors. The proposed system improves recognition accuracy by combining multi-view face images and exploits prolific information from IMU sensors to reduce energy consumption. Specifically, we proposed a face recognition algorithm MVSRC which exploits prolific information from multi-view face images and weighted SRC to improve the recognition accuracy. Then we built a novel estimation model based on SVR, which utilizes the information from IMU sensors to improve the efficiency of MVSRC while preserving its high recognition accuracy. Extensive dataset based evaluations and real-world experiments demonstrate that our system is both accurate and efficient compared to the state-of-the-art.

Chapter 4

Motion-assisted Automatic Device Pairing System For Wearable Devices

Chapter Summary: Recent years have witnessed a remarkable growth in the number of smart wearable devices. For many of these devices, an important security issue is to establish an authenticated communication channel between legitimate devices to protect the subsequent communications. Due to the wireless nature of the communication and the extreme resource constraints of sensor devices, providing secure as well as efficient and user-friendly device pairing is a challenging task. Traditional solutions for device pairing mostly depend on key predistribution, which is unsuitable for wearable devices in many ways. In this chapter, we design *Walkie-Talkie*, a shared secret key generation scheme that allows two legitimate devices to establish a common cryptographic key by exploiting users' walking characteristics (gait). The intuition is that the sensors on different locations on the same body experience similar accelerometer signals when the user is walking. However, one main challenge is that the accelerometer also captures motion signals produced by other body parts (e.g., swinging arms). We address this issue by using a Blind Source Separation (BSS) technique to extract the informative signal produced by the unique gait patterns. Our experimental results show that Walkie-Talkie can generate a common 128-bit key for two legitimate devices with 98.3% probability. To demonstrate the feasibility, the proposed key generation scheme is implemented on modern smartphones. The evaluation results show that the proposed scheme can run in real-time on modern mobile devices and incurs low system overhead.

4.1 Introduction

During the past decade, the number of Internet of Things (IoT) devices introduced in the market has increased considerably. It is estimated that there will be 20 billion connected devices by the year 2020, and the majority of which are IoT and wearable devices [99]. With this trend, the number of connected devices per person rises dramatically. Much like the embedded systems they originate from, on-body IoT devices are equipped with a number of sensors which offer means to collect significant personal information and transmit the collected data to other personal devices. As such, secure data exchange among them becomes a significant problem. For example, smartphones need to frequently push notifications to devices such as smart watches, and read health-related sensor data from wearables or Implantable Medical Devices (IMDs). Since these devices usually contain sensitive private information, data sharing needs to be kept strictly among devices that belong to the same user (on the same body).

The wireless nature of the communication between these devices gives rise to security problems. A malicious external device can listen to the wireless communication between legitimate on-body devices and eavesdrop on private information about the user. To address this problem, conventional mechanisms rely on cryptographic keys to support the integrity and confidentiality of data communication. Specifically, two devices need to agree on a common secret key before communication, and then the established key can be used to encrypt/decrypt subsequent communications between these two parties. In dynamic mobile environments, devices need to perform peer-to-peer associations on-the-fly. However, a trusted authority for key management is not always available, making it difficult to distribute keys between legitimate devices.

In this chapter, we propose and implement a motion-assisted key generation technique for secure on-body device communication. The intuition of the proposed key generation approach is that the devices on the same body experience similar motion signals that are produced by the unique walking pattern of the user. Therefore, the unique gait signal can be exploited as shared information to generate secret keys for all on-body devices. This observation holds true for sensors on the main body trunk (i.e. waist, chest and head); however, for devices worn the wrist, the signals measured by on-broad sensors may be produced by other body motions (e.g., arm swing motion) primarily. Therefore, wearable sensor signals on the wrist is an aggregation of gait and arm swing signal. As a result, the most informative signals (e.g., caused by gait) for key generation may be overwhelmed by less informative signals (e.g., caused by arm swing motion). To address this issue, we use Blind Source Separation (BSS) technique to extract the

most informative signals from mixed signals recorded by the on-body devices. Since walking is a common daily activity, human gait can be automatically detected and measured in daily life without requiring the users to perform key generation explicitly. The proposed approach enables unobtrusive establishment of secure communications between on-body devices.

4.1.1 Motivation

This section discusses the benefits offered and applications enabled by the motion-assisted key generation technique proposed in this chapter.

- **On-body Authentication.** By allowing secure communication establishment only between legitimate on-body devices using the unique body motion signals, Walkie-Talkie enables on-body device authentication without any intrusive manual assistance. Unlike state-of-the-art biometric authentication methods that use face and fingerprints, Walkie-Talkie reduces expensive computation as well as the manual user input required by conventional authentication approaches. This makes it a promising technique for light-weight *continuous authentication* for on-body IoT devices. This feature is desirable especially for wearable and implantable devices, which are usually *small*, *sensor-equipped*, *produce sensitive private data*, and *require frequent authentication*.
- **Automatic Secure Pairing.** In mobile systems, device pairing is required to agree on common encryption schemes and encryption keys before communicating data. Currently, device pairing is achieved either through *explicit input* (e.g., entering the key manually on the device's screen) or sophisticated *peer-to-peer key-exchange algorithms*.

  For explicit input, some common mechanisms are a Personal Identification Number (PIN) code entry or pushing buttons on the devices to be paired. However, these manual approaches suffer from several limitations. First, the form factor of wearable devices are usually small, making it hard for users to enter the keys manually. Second, the number of pairings required is expected to grow considerably as IoT devices become increasingly pervasive. Consequently, explicit pairing places a large burden on device users and automatic pairing improves the user experience significantly. Another approach is through a peer-to-peer key-exchange algorithm. A popular key exchange algorithm is the Diffie-Hellman (DH) protocol [41], which is used to distribute symmetric keys between two parties. However, the DH protocol requires computationally intensive operations and a public key infrastructure, and is infeasible for resource-constrained wearable devices.

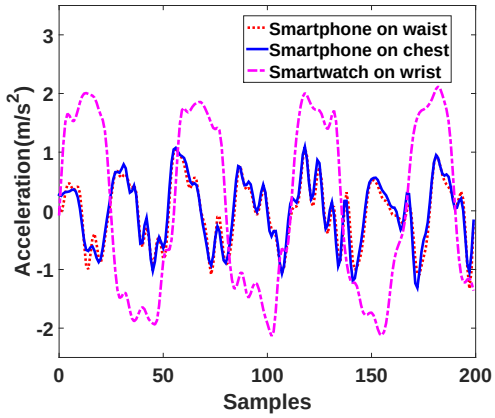


Figure 4.1: Acceleration signal in the gravity direction captured by devices located at different body locations when a user is walking.

- **Spontaneous Key Generation.** To reduce manual input, a user can choose to store the static keys on the device locally, e.g., user can pair two devices on their first use together and use the same key afterwards. However, a critical component of key management is key revocation which is used to revoke and update the secret key. Storing static keys locally poses significant security risks, especially when devices are only authorized to communicate temporarily for short-lived data exchange. So it is crucial that the keys are generated on-the-fly only when they are authorized to communicate.

4.1.2 Challenges and Contributions

Gait refers to an individual's unique walking pattern [105]. The gait signal produced when a user is walking serves as a valuable signal for key-generation for on-body devices, since the sensors on different body locations sense the same signal. The key idea of the proposed key-generation approach is based on this observation. However, due to the complexity of body movements, devices placed on different body locations will capture different acceleration signals due to the movement of other body parts (such as arms), and this becomes the key challenge when exploiting the common gait signal for key-generation.

Figure 4.1 plots the acceleration signal in the gravity direction captured by devices placed at different body locations when the user is walking. The acceleration readings on the body trunk (waist and chest) originate primarily from the walking action, and generate similar patterns. However, the sensors on the wrist capture the aggregated acceleration signal produced by both gait and arm swing. Thus the common motion signals (caused by gait) for key generation is overwhelmed by noise (caused by the arm swing motion). This makes it infeasible to use

the raw motion signals captured by the sensors to generate a common secret key directly. To address this challenge, Walkie-Talkie uses the Blind Source Separation technique described in Section 4.4 to separate the signals produced from gait and arm swing, and use the common gait signal to generate keys for secure communication for all on-body devices.

The second challenge is that the on-body devices are limited by their computational capacity and power supply. As described in [118], Implantable Medical Devices (IMDs) are long-lived devices and battery replacement requires surgical intervention. Therefore, the pairing protocol should be lightweight and energy-inexpensive. The proposed key generation scheme requires only lightweight signal processing techniques, Advanced Encryption Standard (AES) invocations and hash computations by the on-body devices.

To the best of our knowledge, this is the first work that exploits gait signals to achieve efficient key generation and secure communication establishment for devices placed at different body locations. Our main contributions are threefold:

- **Source separation for body motion signal**: By using Blind Source Separation to separate motion signals generated from different body movements, e.g., gait and arm swing motions, the proposed key generation approach achieves robust performance in generating keys for devices located at different body locations.

- **Shared key generation scheme**: We present a novel, light-weight key generation scheme for on-body IoT devices based on body motion signals. We experimentally demonstrate that a common 128-bit key can be successfully generated by two independent wearable devices on the same body in 98.3% of the cases, while the scheme also provides adequate security guarantees against impersonation attacks. By walking for 4.6s ($\approx$9 steps), the proposed key generation approach is able to generate a 128-bit key with entropy varying from 0.94 to 1 which demonstrates the high randomness of the keys.

- **System implementation**: We illustrate the practicability of the proposed key generation approach by implementing the system in Bluetooth Low Energy (BLE) peripheral mode. We report the system computation overhead and power consumption, and demonstrate the feasibility of the proposed scheme for contemporary on-body IoT devices.

The organization of this chapter is as follows. We introduce the user model and the adversary model in Section 4.2. We specify the design overview in Section 4.3, signal processing in Section 4.4, and key generation in Section 4.5 respectively. We then evaluate the performance

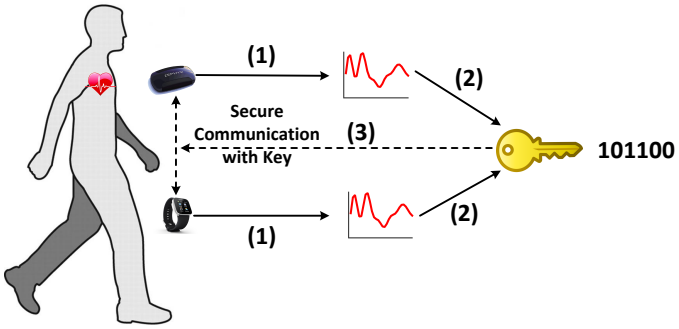


Figure 4.2: System overview: 1) Pacemaker and smart watch measure the similar gait signals simultaneously. 2) They use the gait signals to generate a shared secret key. 3) The key is then used to ensure the security of communication between two parties.

of the proposed scheme and analyze security issues in Section 4.6, and present the system implementation in Section 4.7. Finally Section 4.8 summarizes this chapter.

4.2 Model

Before discussing the framework of Walkie-Talkie, we first introduce the user model and the adversarial model.

4.2.1 User Model

We envision the use of Walkie-Talkie primarily for pairing wearable and implantable devices. Figure 4.2 illustrates a typical user model for on-body device communication in Walkie-Talkie. One morning, a user wants to pair his smart watch (Alice) with pacemaker (Bob) to read health information. The user launches Walkie-Talkie on the smart watch and walks several steps, and then both Alice and Bob generate a secret symmetric key by exploiting the measured gait signals during this period. The key is then used to encrypt/decrypt the messages between Alice and Bob.

4.2.2 Adversarial Model

To achieve secure communication, a common attack that needs to be addressed is the *impersonation attack*, in which an adversary (Eve) tries to impersonate a legitimate device to steal private information. We assume the presence of two types of impersonation attack during a key generation session: a passive eavesdropping adversary and an active spoofing attack. The passive

adversary knows the key generation mechanism and can eavesdrop on the messages exchanged between Alice and Bob during the key generation process. The active spoofing attacker tries to mimic the walking style of the genuine user to pair with one or both of the legitimate devices.

As discussed in [96], although the attacker can monitor messages exchanged between the legitimate devices, we assume that they can neither control the acceleration recorded locally by these devices nor perfectly estimate it, otherwise the protection of legitimate devices is impossible. We also assume that all the devices on the user's body are legitimate devices, i.e., an adversary cannot insert a device on the user to get the acceleration data. Another potential attack is replay attack, whereby a user may wear a wearable (provided by attacker) that records the movement pattern and the attacker later replays the signals to generate keys. Walkie-Talkie can address this type of attack because the keys are generated by the real-time gait signals, and the signals recorded a period of time ago are different from the current signal. Further potential threats include deriving the acceleration by studying a video of the target's gait through computer vision techniques. We believe this is a potential vulnerability of unknown severity and leave it as future work.

4.3 Design Overview

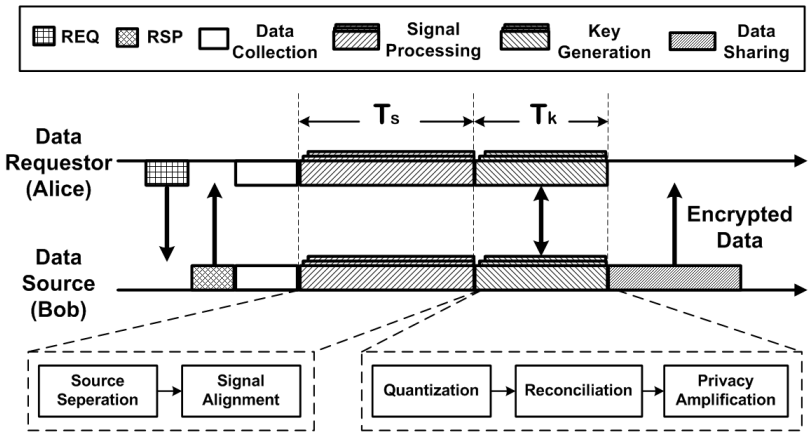


Figure 4.3: Flowchart of the key generation scheme.

Figure 4.3 shows the work-flow of Walkie-Talkie. Suppose Alice (e.g., smart watch) wants to read data from Bob (e.g., pacemaker). Alice first broadcasts a *REQ* request to Bob. After receiving the *REQ*, Bob replies with a *RSP* response. Then both Alice and Bob start to collect local motion sensor data and follow the steps shown in Figure 4.3 to generate a shared secret key. Finally, the key is used to encrypt/decrypt data to ensure secure communication between

Alice and Bob.

The key component of Walkie-Talkie consists of the following two steps:

- **Signal Processing** Signal processing consists of two steps: source separation and signal alignment. Source separation is performed on the acceleration data collected from the on-body devices to extract the signals produced by gait. As Alice and Bob sample acceleration data independently, we apply signal alignment to synchronize acceleration samples at Alice and Bob and transform the acceleration to the same body coordinate system to facilitate key generation.
- **Key Generation** The key generation component consists of three basic steps: quantization, reconciliation and privacy amplification. In quantization, the legitimate devices, Alice and Bob, convert acceleration samples into bits if they are both on the same body. In the reconciliation stage, Alice and Bob exchange error-correcting messages over a public channel that allows them to agree on an identical string of bits. However, the publicly exchanged messages reveal a certain amount of information about the bit strings to Eve. To address this issue, Alice and Bob diminish the partial information revealed to Eve by privacy amplification.

In the following sections, we will describe design details of each component. Table 4.1 summarizes the notation used in this chapter.

4.4 Signal Processing

4.4.1 ICA-based Source Separation

When an individual is walking, accelerometer recordings from one body location are typically a mixture of accelerations produced from multiple body locations (e.g., leg, waist, and arm). For wearable and implantable devices, most common locations are waist, chest, head and wrist. As described in Section 4.1.2, the sensors on the body trunk measure the motion signals produced by gait primarily. Therefore, the devices on the body trunk can exploit the acceleration readings directly to generate a key. However, sensors worn on the wrists capture signals from a combination of gait and arm swing motions. In order to exploit the useful signal (gait) to generate a key, we need to separate signals produced from leg motions (walking) and arm swing motions.

In this section, we apply the Independent Component Analysis (ICA) technique to separate signals from different body sources [64]. ICA is one of the most popular blind source separation

Table 4.1: A summary of the main symbol notations.

Symbol	Meaning
$Acc(t)$	raw linear acceleration data
A	mixing matrix
$S(t)$	independent components
W	unmixing matrix
$\tilde{S}(t)$	estimated independent components
$Acc^{'}(t)$	reconstructed acceleration
q_+, q_-	quantization boundaries (upper and lower)
L_{Alice}, L_{Bob}	index list of generated bits
$\tilde{L}$	common index list between L_{Alice} and L_{Bob}
$MAC(\cdot)$	message authentication code algorithm
K_{Alice}, K_{Bob}	generated key after quantization
$K^{'}_{Alice}, K^{'}_{Bob}$	generated key after reconciliation
$K^{''}_{Alice}, K^{''}_{Bob}$	final key after privacy amplification

(BSS) methods, which aims to separate the mixed signals into a set of independent sources given very little information (or no prior information) about the source signals. Before applying ICA, we first justify that on-body accelerometer satisfies the conditions for ICA. 1) The acceleration from the different sources is mixed linearly at each sensor location, as we record the linear acceleration along 3 channels of the accelerometer sensor for each location. 2) The acceleration of arm swing is independent from that originating from heel strike. As stated in [105], the movement patterns of various parts of the body are independent, and gait is the total pattern of movement when they are integrated together. 3) Time delays in signal transmission through the body are negligible. 4) There are fewer sources than mixtures. For each location, we attach a 3-channel accelerometer sensor, thus we have an observation of 3 channels and the signals are mainly from two sources: arm swing and walking. 5) Statistical distributions of the acceleration values produced by body movement are not Gaussian [63].

Suppose a smart watch is worn on one wrist of the user, and the measured linear accelerations by the built-in three channel accelerometer are $Acc(t)$. As the accelerometer signals recorded on the wrist are a mixture of the signal from leg and arm swing respectively, the ICA model of our problem can be written as:

$$Acc(t) = A \cdot S(t) \tag{4.1}$$

where A is the mixing matrix and $S(t)$ represents independent sources. Our aim is to find an unmixing matrix $W(W = A^{-1})$, so that we can calculate the estimated source signal $\tilde{S}(t)$ by:

$$\tilde{S}(t) = W \cdot Acc(t) = W \cdot A \cdot S(t) \tag{4.2}$$

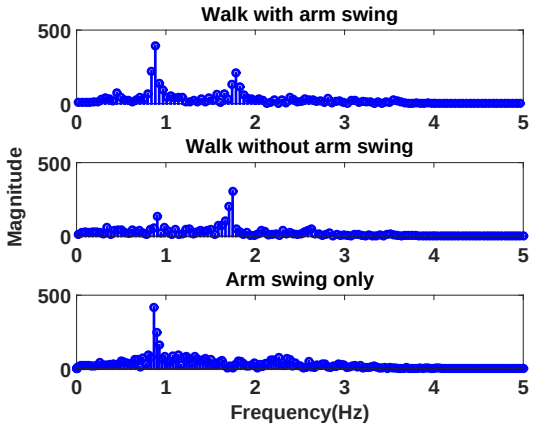


Figure 4.4: Frequency of different activities.

In the proposed system, we use FastICA (A fast fixed-point algorithm of independent component analysis) to solve the ICA model in Eq. 4.1, i.e., to estimate W. FastICA has been found to be 10-100 times faster than conventional gradient descent methods for ICA [63]. Therefore, FastICA is well suited for the resource-constrained on-body devices in this work.

After obtaining W, we obtain the estimated sources $\tilde{S}(t)$ by Eq. 4.2. In our problem, the rows of $Acc(t)$ are the linear acceleration values along three axes of the accelerometer. The acceleration signal without arm swing motion can be derived from $Acc^{'}(t) = W\bar{S}$, where $\bar{S}$ is the matrix of derived independent components with the row representing the arm swing set to zero. Assume the second ICA component represents the signal from arm swing. $\bar{S}$ can then be written as:

$$\bar{S} = \begin{bmatrix} \tilde{S}_{11} & \tilde{S}_{12} & \cdots & \tilde{S}_{1N} \\ 0 & 0 & \cdots & 0 \\ \tilde{S}_{31} & \tilde{S}_{32} & \cdots & \tilde{S}_{3N} \end{bmatrix} \tag{4.3}$$

where $\tilde{S}_{ij}(i, j = 1, \ldots, N)$ are the elements of matrix $\tilde{S}(t)$ and N is the number of acceleration samples. In the following section, we describe how we identify different motion components.

4.4.2 Identifying Motion Component

From the ICA model in Eq. 4.1, it can be seen that one cannot determine the order of the independent components, as a permutation matrix P and its inverse P^{-1} can be added in the model to yield $Acc(t) = AP^{-1}PS(t)$. The elements of $PS(t)$ are the original independent variables, but in a different order. The matrix AP^{-1} is therefore a new unknown mixing matrix, to be solved by the ICA algorithm. Furthermore, the order of components may also vary from one data segment to the next. Consequently, one has to depend on visual inspection of the ICA components

for further processing, a method which is not desirable for on-body sensors.

In practice, the separated components tend to have more distinctive properties than the original signals both in time and frequency domains. Figure 4.4 shows the frequency of walking while swinging an arm, walking without swinging an arm, and swinging an arm only. We notice that the dominant frequency of the signal from walking only is two times of that of arm swing signal. This is easy to understand because a gait cycle is composed of two steps and one arm swing cycle. Therefore, each step (left or right) registers as a strong repetitive acceleration signal and the signal is transmitted through the foot to the whole body. Due to the symmetry of the body, the signal produced by left and right step can be deemed to be same. However, the arm swing signal only repeats every two steps as the smart watch is worn on one wrist of the user. We use this observation to identify the signal from arm swing and foot. Specifically, after obtaining $\tilde{S}(t)$ by Eq. 4.2, we perform a Fast Fourier Transform (FFT) on the three independent components (ICs) in $\tilde{S}(t)$ (i.e., three rows of $\tilde{S}(t)$). Figure 4.5(d) illustrates the magnitude of the acceleration signals in three directions before ICA and after ICA. We can see that the original acceleration data contains signals from two frequencies primarily. The three separated independent components (ICs) present different frequency distributions. The frequencies of IC-2 are concentrated on the fundamental frequencies. As discussed above, the reconstructed signal without arm swing motion can be obtained by setting the second row of the matrix $\bar{S}$ to zero (see Eq. 4.3).

Figure 4.6 presents the acceleration in the gravity direction before and after source separation. We can see that the acceleration produced by walking is overwhelmed by arm swing in the raw acceleration signals. The acceleration after source separation is very similar to the readings on the chest, just the magnitude of the signal is reduced, because the signal produced from leg motion is attenuated through the body to the wrist. Note that one cannot simply apply a low-pass filter to filter out the signal produced by arm swing motion because the walking signal also contains a fundamental frequency component as shown in Figure 4.4.

4.4.3 Signal Alignment

The raw acceleration data cannot be used to generate the key directly as the accelerometer values are sensitive to sensor orientation and location. Additionally, different devices are usually not well time-synchronized which leads to the problem of signal synchronization. We address these two issues by temporal alignment and spatial alignment.

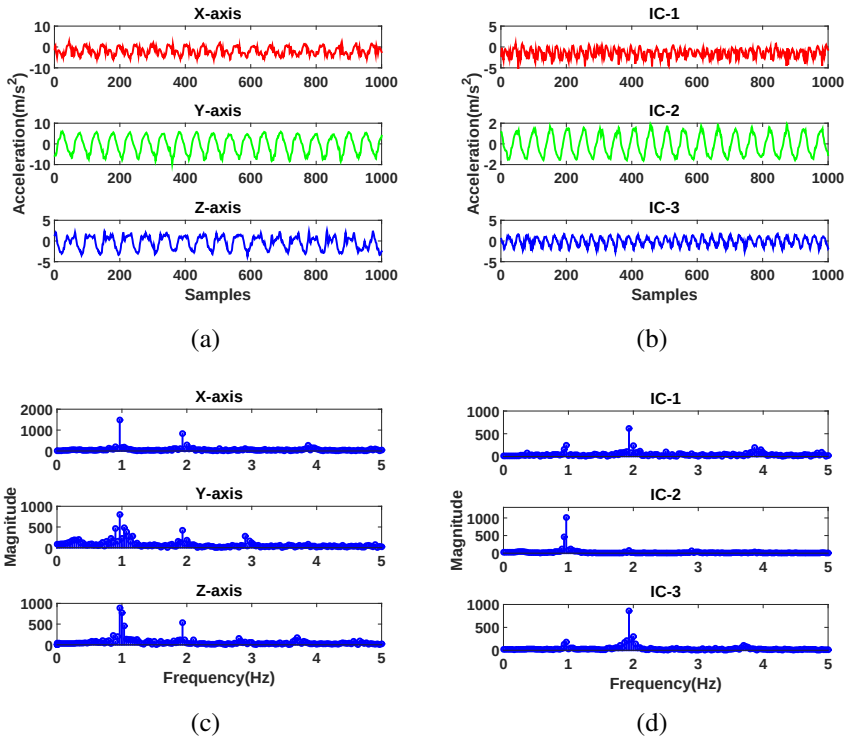


Figure 4.5: ICA results: (a) Raw acceleration $Acc(t)$. (b) Estimated independent components $\tilde{S}(t)$. (c) Frequency of raw acceleration. (d) Frequency of estimated independent components.

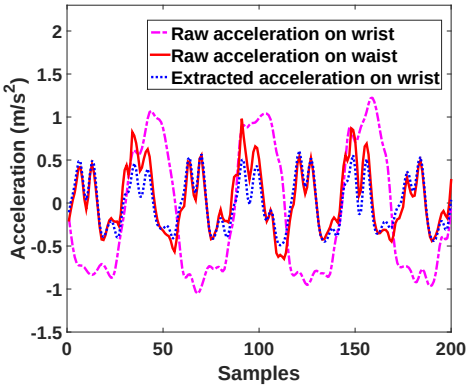


Figure 4.6: Comparison of raw signal and extracted signal.

Temporal Alignment

As devices sample acceleration values independently, temporal synchronization is required for key generation. In the system, we use an event-based approach in which devices detect the time point of a heel-strike event, and use this event as an anchor point. The intuition is that the acceleration values along gravity direction reach the peak simultaneously when the foot touches the ground, and time delays in signal transmission through the body are negligible. To detect heel-strike, we first apply a low-pass filter on acceleration along the gravity direction to

reduce noise. The cutoff frequency is chosen as 3Hz as the normal step frequency lies between 1.6-2.8 Hz [105]. Then the local maxima are detected to identify heel-strike events as shown in Figure 4.7.

Heel-strike events can be detected locally at each device without communication which eliminates the need for explicit synchronization between devices. When Alice receives a *RSP* from Bob, both of them reach to agreement to record acceleration from the next n_{start}-th heel-strike event and end recording at the subsequent n_{end}-th heel-strike event. The acceleration samples are then transformed to the body coordinate system as described in the following section.

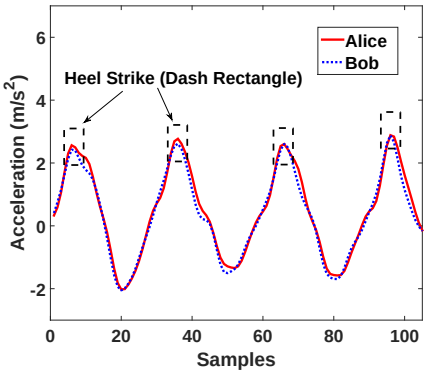


Figure 4.7: The peak of acceleration along the gravity direction indicates a heel-strike on the ground.

Spatial Alignment

Walking is inherently a three-dimensional movement. 3D acceleration data independently recorded at different locations lack spatial alignment and cannot be directly used to generate a shared secret key. We address this by transforming acceleration values of different devices to a common body reference coordinate system independent of orientation and location. Figure 4.8 illustrates the definition of the world coordinate system, the body reference coordinate system and the coordinate system of different devices. The world coordinate system is defined by North, East and the Down or gravity direction ($-$G). We refer to the device's local coordinate system as (X, Y, Z). The user plane of motion is defined as the Forward-Sideways plane which is perpendicular to gravity. Sideways points toward the right side of the user's forward direction.

Taking a smartphone as an example, assume the linear acceleration signals along three orthogonal directions of smartphone are Acc_x, Acc_y, and Acc_z respectively, the linear acceleration

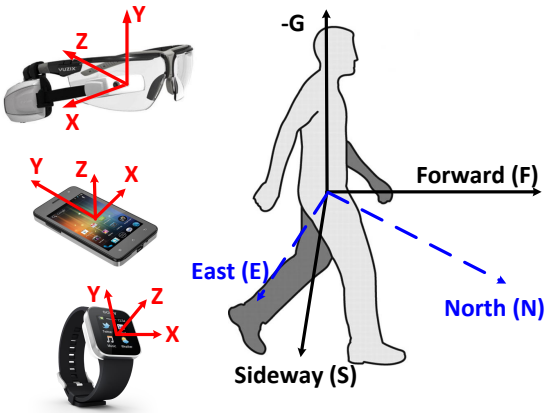
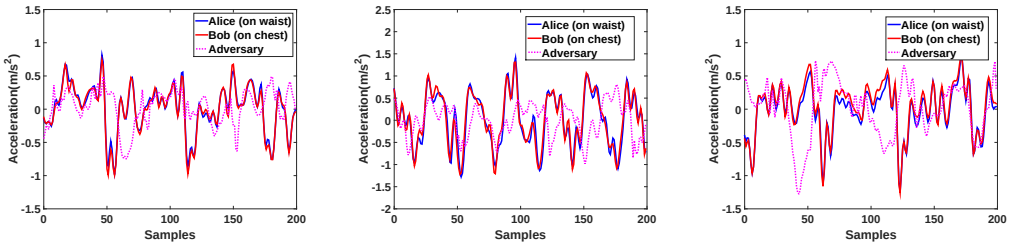


Figure 4.8: The different coordinate systems.



(a) Acceleration in the walking direction (b) Acceleration in the gravity direction (c) Acceleration in the sideways direction

Figure 4.9: Acceleration of two legitimate devices and an adversary device.

in the body reference system can be computed as:

$$\begin{bmatrix} Acc_G \\ Acc_F \\ Acc_S \end{bmatrix} = R_b^w \cdot R_w^d \cdot \begin{bmatrix} Acc_x \\ Acc_y \\ Acc_z \end{bmatrix} \tag{4.4}$$

where Acc_G, Acc_F, and Acc_S are linear accelerations along gravity direction, forward direction and sideways direction in the body reference system, R_b^w is the rotation matrix from the world coordinate system to the body coordinate system and can be computed by the method in [102]. R_w^d is the rotation matrix from the device coordinate system to the world coordinate system and can be obtained by the Android API. Note that the absolute walking direction of the user cannot be obtained accurately using a smartphone compass [119]. In Walkie-Talkie we don't have this problem because we consider the acceleration values only instead of walking direction. After obtaining the acceleration in the body coordinate system, we use Acc_G, Acc_F and Acc_S for key generation.

4.5 Key Generation

After source separation and signal alignment, we obtain acceleration values caused by gait along three directions: Acc_G, Acc_F and Acc_S. Figure 4.9 plots the acceleration of two legitimate devices and one adversary device in three directions. We can see that the devices on the same body follow the same pattern, however, the acceleration signal recorded by an adversary device significantly differs. This result is promising since our goal is to generate symmetric keys only for devices on the same body. The following key generation method is applied on two legitimate devices separately.

4.5.1 Multi-level Quantization

We perform filtering, and quantization for the acceleration values along the three directions separately. We first apply a low-pass filter for noise reduction. The cutoff frequency is chosen as 10Hz as the useful frequency of human motion lies below 10 Hz [80]. Note that the cutoff frequency of this low-pass filter is different from that used for heel-strike mentioned in Section 4.4.3. After filtering, the acceleration values are normalized to have zero-mean and unit length to alleviate the influence of different body locations. Then we extract multiple bits from the accelerometer signal samples by employing multi-level quantization technique [143]. More specifically, we segment the acceleration values with a moving window with no overlap (window size W). Thereafter, for each window, we generate bits by the following steps.

Determining the Upper Bound on the Number of Bits

The first step is to determine the maximum number of bits that can be assigned per sample. For a given window, we calculate the approximate entropy of samples by using the equation:

$$\mathscr{E} = -\sum_a p(a)\log_2 p(a), \tag{4.5}$$

where $p(a)$ is the probability of occurrence of acceleration sample a in a selected window. The upper bound of the quantization level m_{MAX} is calculated as $m_{MAX} \leq 2^{\mathscr{E}}$.

Determining the Quantization Intervals

In the next step, we calculate the size of each level in the quantization. In a multi-bit (i.e., m-ary) quantization, *guard bands* (g_i) are inserted between two consecutive levels (i.e., q_{i-1} and q_i) to increase the bit agreement ratio. The guard band samples are excluded during quantization and the remaining samples are encoded to bits according to their levels. The notation α represents the ratio of guard band to data, i.e., the excluded acceleration values in all the guard bands over the total number of samples. Each level in an m-ary quantization, with m being the number of levels, is represented by a number, i.e., level-0 to level-(m-1). The individual quantization intervals are calculated by the following equations:

$$I_0 = (q_0, q_1 - g_1], I_1 = (q_1, q_2 - g_2], \ldots, I_{m-1} = (q_{m-1}, q_m], \tag{4.6}$$

where q_0 is the minimum and q_m is the maximum value of acceleration samples in the window. For each level, we calculate the size of the quantization interval and guard band by the equations:

$$\int_{q_{i-1}}^{q_i - g_i} f_a da = \frac{1-\alpha}{m}, \int_{q_i - g_i}^{q_i} f_a da = \frac{\alpha}{m-1}, \tag{4.7}$$

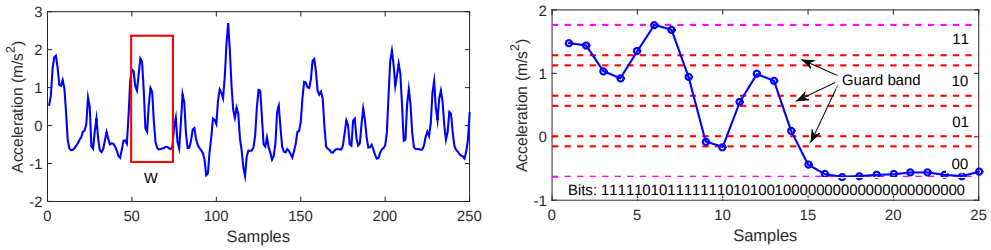
where $1 \leq i \leq m-1$. Each level in the quantization is assigned an n–bit code ($n = log_2 m$). We assign the bits to each level such that its decimal value denotes the index of the level. The secret bits are then extracted from acceleration samples based on their level in the quantization.

Extracting the final key

Similar to the single bit quantization, we perform quantization for the acceleration values along the three directions separately. Three separate bit streams K_G, K_F, K_S are extracted from Acc_G, Acc_F and Acc_S respectively, and the secret key for Alice is obtained by concatenating three bit streams together as $K_{Alice} = [K_G, K_F, K_S]$. The same quantization process is also performed by Bob independently to get K_{Bob}. Figure 4.10(a) plots the raw acceleration data recorded in an experiment, and Figure 4.10(b) illustrates the process of 2-ary quantization for a window.

4.5.2 Reconciliation

After quantization, each device ends up with a secret key string independently. However, there may be some bit mismatches due to noise and we often get $K_{Alice} \approx K_{Bob}$. The purpose of



(a) Raw acceleration data (b) Quantization

Figure 4.10: Illustration of quantization process for $W = 25$, $m = 4$ and $\alpha = 0.2$.

reconciliation is to correct the bit mismatches between Alice and Bob. In this system, we employ the Error Correction Code (ECC) [34] to reduce the bit mismatch rate.

Suppose the mismatching bits between Alice and Bob is $\varepsilon = K_{Alice} \oplus K_{Bob}$, and let $C(n,k)$ be an ECC that encodes k-bit message into n-bit code to resist r-bit random error. Function $f(\cdot)$ and $g(\cdot)$ denote the corresponding encoding function and decoding function. To start the reconciliation, Alice first computes the offset δ_{Alice} between K_{Alice} and its corresponding codeword as follows:

$$\delta_{Alice} = K_{Alice} \oplus f(g(K_{Alice})) \tag{4.8}$$

Then, Alice transmits δ_{Alice} to Bob via a public channel. Upon receiving δ_{Alice}, Bob can deduce K_{Alice} as follows:

$$K'_{Alice} = \delta_{Alice} \oplus f(g(K_{Bob} \oplus \delta_{Alice})) \tag{4.9}$$

If the mismatching rate ε is lower than the error-correcting ability of C, an appropriate error correction code C can be employed to ensure $K'_{Alice} = K_{Alice}$. Therefore, both Alice and Bob agree on the same key $K'_{Alice} = K_{Alice}$, and they use the key to encrypt/decrypt the communication between them.

Since Alice and Bob do not share an authenticated channel, Eve can impersonate Alice or Bob during the reconciliation process. Such an attack would allow Eve to insert her own fake messages, thus spoofing a legitimate device and disrupting the protocol without revealing his presence. To address this issue, we employ the message authentication code (MAC) method [94] to verify that the message has not been modified. Specifically, the MAC method contains the following three steps:

- To ensure the message δ_{Alice} is indeed sent from Alice, Alice sends a MAC message with δ_{Alice}, the overall message sent by Alice is $L_{Alice} = \{\delta_{Alice}, MAC(K_{Alice}, \delta_{Alice})\}$. After receiving L_{Alice}, Bob computes K'_{Alice} by Eq 4.9 and uses it for MAC verification. If Bob

obtains $MAC(K_{Alice}, \delta_{Alice}) \neq MAC(K'_{Alice}, \delta_{Alice})$, he can conclude that the message was not sent by Alice, indicating the presence of an adversary.

- If Bob does not detect the presence of an adversary, he computes δ_{Bob} and transmits the following message to Alice: $L_{Bob} = \{\delta_{Bob}, MAC(K_{Bob}, \delta_{Bob})\}$.
- Upon receiving L_{Bob}, Alice computes K'_{Bob} and uses it for MAC verification. If Alice obtains $MAC(K'_{Bob}, \delta_{Bob}) = MAC(K_{Bob}, \delta_{Bob})$, she can confirm that the message was indeed sent by Bob. Since Eve does not know the bits in K_{Bob} generated by Bob (he can just listen to the output of the $MAC(K_{Bob}, \delta_{Bob})$), modifying δ_{Bob} will fail the MAC verification at Alice.

Apart from verifying that the message has not been modified, the MAC verification also verifies whether Alice and Bob generate the same key. Because if $K'_{Alice} \neq K_{Alice}$, Bob cannot obtain $MAC(K'_{Alice}, \delta_{Alice}) = MAC(K_{Alice}, \delta_{Alice})$. In this case, the key generation process fails, and Bob will either notify Alice to restart the key generation process, or consider Alice as an unauthorized device and deny all Alice's consequent requests, depending on application requirements.

The reconciliation process not only reduces the mismatch rate between Alice and Bob, but also reveals partial information to an attacker, as δ_{Alice} is transmitted over a public channel and can be eavesdropped by an attacker. However, it can be theoretically proved that there are only $(n-k)$ bits of information leakage [93]. Moreover, since the secret key is derived from user's unique walking pattern, the attacker still cannot infer K_{Alice} by eavesdropping δ_{Alice}. To ensure there is no partial information leakage, we can further apply the privacy amplification technique described in the following part.

4.5.3 Privacy Amplification

After reconciliation, Alice and Bob agree on a common secret key as $K'_{Alice} = K_{Alice}$. Simply concatenating the bits generated from each time window does not necessarily produce a random secret key, as correlation between different steps may result in high correlation between key bits. Moreover, reconciliation leaks some information to an attacker. This issue can be addressed by privacy amplification techniques [17]. In the system, we use a bit-wise XOR function to combine keys generated from each direction and eliminate the correlation between them. Specifically, we interleave the bit streams from three directions in the time sequence and segment the concatenated keys into small windows with no overlap. Each window contains 30

bits which is close to the bits generated in a gait cycle duration as the evaluation results show in Section 4.6.5. Then we XOR two consecutive windows together to obtain the final key K''_{Alice}.

Another advantage of privacy amplification is that it diminishes the partial information revealed to Eve as discussed in [17]. In the reconciliation stage, Alice and Bob exchange messages over a public channel and the publicly exchanged messages reveal a certain amount of information about the bit strings to Eve. To reduce the impact of the revealed information, the privacy amplification significantly improves the randomness of the keys generated as the evaluation results in Section 4.6.5. Note that other privacy amplification methods such as a universal hash [17] can be employed to further enhance the randomness of the concatenated key. We refer the reader to [17] for more details.

After privacy amplification, the final key can be used by symmetric-key algorithms such as AES to ensure secure communication between Alice and Bob. If the length of final key is greater than 128 bits, the first 128 bits are used. If the key generation process fails, Alice will either notify Bob to restart the key generation process, or consider Bob as an unauthorized device and deny all Bob's consequent requests, depending on application requirements.

4.5.4 CIA Properties of Walkie-Talkie

As a security scheme, Walkie-Talkie achieves the CIA properties (confidentiality, integrity, and availability) by the following approaches:

- **Confidentiality**. Data confidentiality is the key focus and is achieved through encryption after key generation.
- **Integrity**. During key generation, integrity is achieved by the MAC; After key generation, with the key the data integrity can be easily achieved using any standard mechanisms such as hashing or checksumming and is beyond the scope of this study.
- **Availability (Anti-DoS attack)**. During key generation, to prevent the adversary from modifying messages to break the reconciliation between two legitimate devices, a MAC mechanism is used to ensure the integrity of the messages and to protect the availability of the key generation. After key generation, unauthorized communications can lead to Denial of Service (DoS) attacks, in which communications between legitimate devices are prevented and batteries are needlessly depleted [120]. To prevent such DoS attacks, Walkie-Talkie only allows authorised communications through authentication achieved by the key generation techniques.

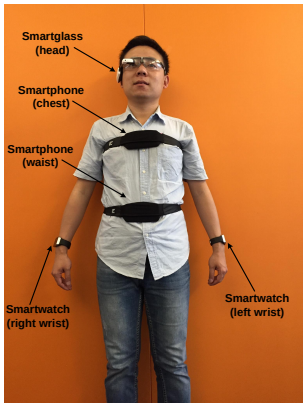


Figure 4.11: Body locations for data collection.

4.6 Evaluation

4.6.1 Goals, Metrics and Methodology

In this section, we evaluate the performance of the proposed key generation scheme. The goals of the evaluation are fourfold: 1) to determine the choice of the key parameters including the window size (W) and α in the quantization process as well as the sampling frequency (F_s); 2) to evaluate the impact of different components in the work-flow including ICA, reconciliation, and privacy amplification; 3) to evaluate the impact of different body locations on bit agreement rate including head, chest, waist, and wrist; 4) to evaluate the security of the scheme against various adversary attacks.

Data Collection. The dataset used to evaluate the performance of the proposed system consists of 20 subjects (14 males and 6 females)[1]. As shown in Figure 4.11, we collect acceleration data from the following body positions: head, chest, waist, and wrist. These positions represent the common locations of mobile devices and medical sensors (e.g., pacemaker). The sampling rate of all devices used in data collection is set to 100 Hz.

During the data collection phase, the participants were asked to wear mobile devices as shown in Figure 4.11 and walk for about 5 minutes at their normal speed (0.7-1.1m/s). The data collection was performed both indoors and outdoors to capture different terrains in practical scenarios. Note that we do not consider data collection on different days or different walking speeds (slow, normal and fast) as all the devices worn by the subject are measuring the same gait signal simultaneously. This is different to the data collection requirements in the study

[1]Ethical approval for carrying out this experiment has been granted by the corresponding organization (Approval Number HC15304)

of gait recognition. The detected peaks which indicate heel-strikes are used to synchronize acceleration samples recorded on different devices and segment steps. For each device attached on one subject, we break the continuous acceleration values into segments according to heel-strike points, each segment contains 10 steps. The segments are used to generate keys and evaluate the following metrics.

Metrics. For a shared key generation protocol, we focus on the following three evaluation metrics:

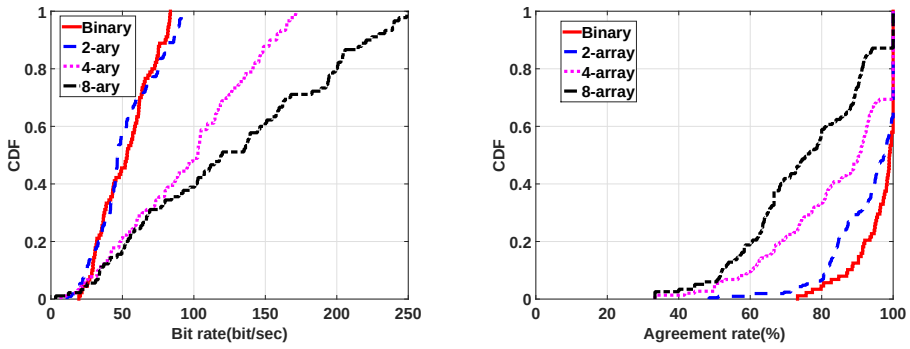
- **Bit agreement rate**: It represents the percentage of bits matching in the secret keys generated by two parties. This metric evaluates the potential of Alice and Bob agreeing on the same key.
- **Bit rate**: It denotes the average number of bits generated from the acceleration samples per unit time and is usually measured in bits per second (bps). This metric evaluates how fast Alice and Bob can generate shared secret bits.
- **Entropy**: It is the measure of uncertainty or randomness associated with the generated bit strings. Entropy of a binary bit string varies in the range $[0,1]$, and larger entropy indicates more randomness of the bit string.

We examine the impact of parameters on the generated key by a systematic exhaustive search. We vary the respective parameters within a dedicated range, i.e. $W = 5,10,\ldots,50$, $\alpha = 0,0.1,\ldots,1$, and $F_s = 10,20,30,50,100$. The goal of the exhaustive search is to find the optimal combinations which achieve good performance in both bit agreement rate and bit rate. After choosing the best combination $(W = 50, \alpha = 0.9, F_s = 50)$, we take turns to investigate the impact of each parameter on agreement rate and bit rate by fixing the other two parameters. Results are presented for the average values and 95% confidence levels of the performance metrics (bit agreement rate and bit rate).

4.6.2 Improvement of Multi-Level Quantization over Binary Quantization

Since m-ary quantization can be used to generate keys with more bits, we compare its performance with the binary quantization method used in Walkie-Talkie [139]. For evaluation purposes, we vary m from 2 to 8. Figure 4.12(a) plots the CDF of bit generation rate under

different methods. The "Binary" means the method used in Walkie-Talkie [139], and the others indicate the method described in Section 4.5.1. Compared with the binary quantization method, the higher level m-ary quantization can significantly increase the bit generation rate. Figure 4.12(b) is the CDF of the bit agreement rate between legitimate devices corresponding to the keys of Figure 4.12(a). Different to the bit generation rate, the bit agreement ratio decreases when higher quantization levels are used. This is because noise will produce more bit mismatches when quantization level increases. The experimental results suggest that multi-level quantization can significantly increase the bit rate while decreasing the bit agreement rate. We also tried quantization levels larger than 8, which yields even lower bit agreement ratio, so we limit our discussion to m =2, 4, and 8 in this study.



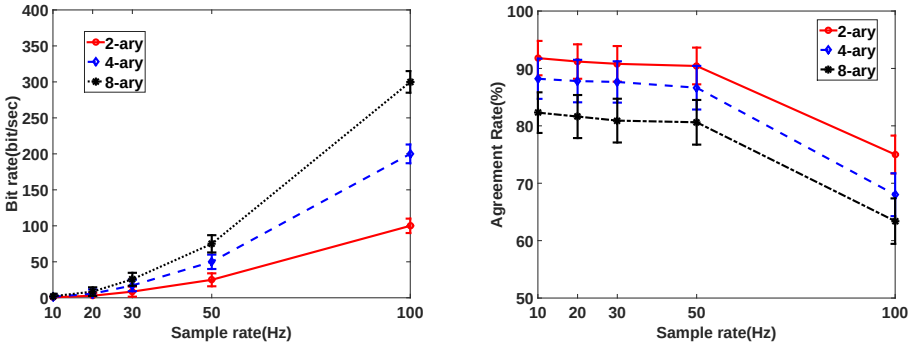
(a) CDF of bit generation rate (b) CDF of bit agreement rate

Figure 4.12: Binary quantization vs. m-ary quantization

4.6.3 Parameter Selection

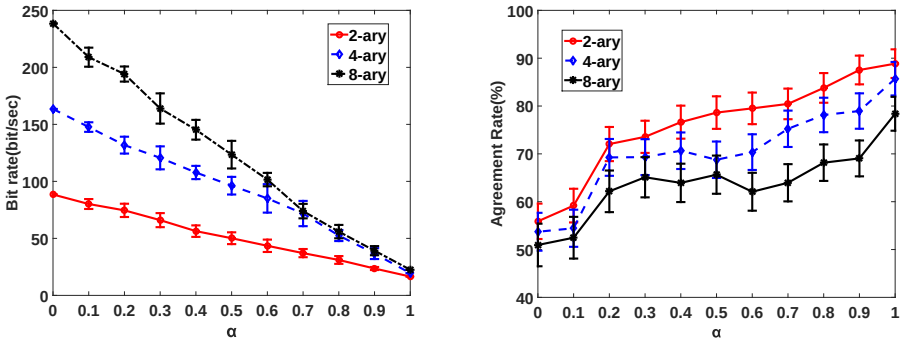
Impact of Sampling Rate

As mentioned above, the initial sampling rate is 100Hz. We evaluate the impact of different sampling rates on bit rate and bit agreement rate by downsampling F_s from 100Hz to 50Hz, 30Hz, 20hz and 10 Hz respectively. Figure 4.13(a) and Figure 4.13(b) show the impact of F_s on bit rate and bit agreement rate respectively. We can see that the agreement rate between legitimate devices varies inversely with sampling rate. The reason is that a higher sampling rate is able to record more acceleration values during the same period and thus improve bit rate; however, it reduces bit agreement as a higher sampling rate captures acceleration variation in more detail leading to less similarity between legitimate devices.



(a) Impact of F_s on bit rate (b) Impact of F_s on agreement rate

Figure 4.13: Impact of F_s.



(a) Impact of α on bit rate (b) Impact of α on agreement rate

Figure 4.14: Impact of α.

Impact of α

We evaluate the impact of α to explore the tradeoff between agreement rate and bit rate. Figure 4.14(a) shows that the bit rate decreases as α increases. This is because the parameter α in Eq. 4.7 decides the decision band to include or discard the acceleration measurements. A larger α means more acceleration readings are discarded. This reduces the length of generated keys and decreases the bit rate. On the other hand, as shown in Figure 4.14(b), the bit agreement rate increases with increasing α because more mismatches in the decision band are excluded.

Apart from sampling rate and α, we also investigated the impact of different window sizes when generating keys. We found that the moving window size W does not have much influence on the performance and a moving window with size of 50 is adequate for the proposed system.

4.6.4 Impact of Reconciliation

Reconciliation is used to correct errors between Alice's and Bob's keys. We examine the effectiveness of different ECC codes under different quantization levels. The candidate ECC codes are: *Hamming code*, *Golay Code*, *Reed-Solomon(RS) code*. Table 4.2 lists the parameters and properties of ECC codes used in our evaluation (code word length n, code length k, error-correcting ability r). Figure 4.15(a), Figure 4.15(b) and Figure 4.15(c) show the impact of ECC codes on the agreement rate under different quantization levels respectively. We can see a significant increase in the bit agreement rate after using the reconciliation technique. From the figures, we also find that a RS code with n=15, k=3 achieves the highest bit agreement rate. One drawback of the reconciliation process is that it reveals some information to attackers, this issue is solved by the privacy amplification process. As the results above, we choose RS(15,3) in our

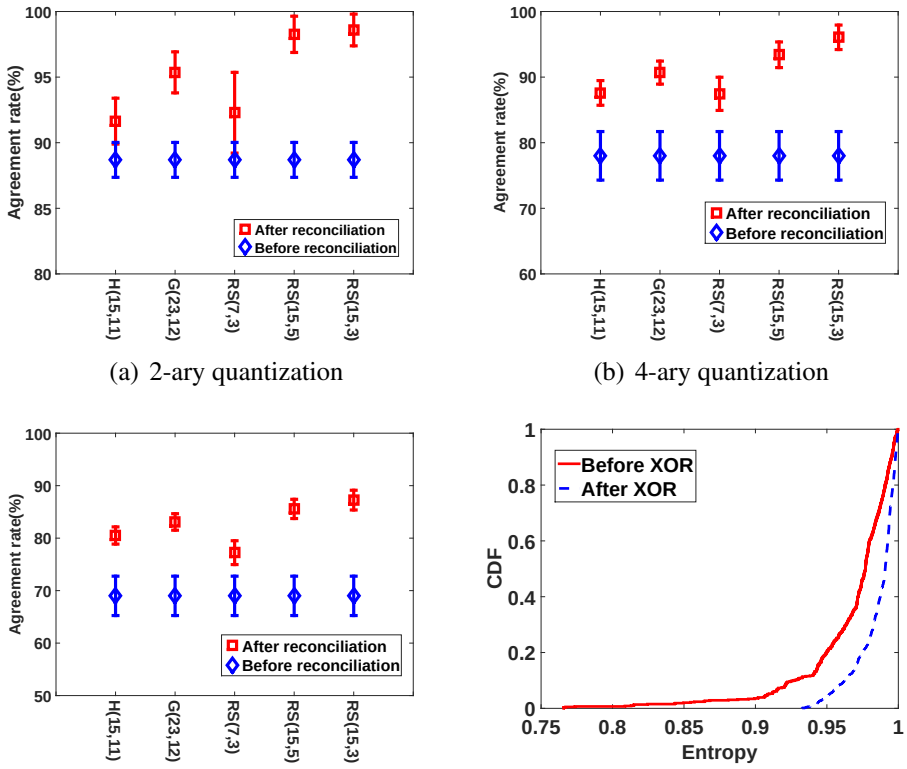
Table 4.2: Comparison of different ECCs

Code	n	k	r	Information Leakage
Hamming Code	15	11	1	0.27
Golay Code	23	12	3	0.48
RS(7,3)	7	3	2	0.57
RS(15,5)	15	5	5	0.67
RS(15,3)	15	3	6	0.8

system and use it for the rest of evaluation. After determining the ECC code, we examine the bit rate and match rate of different quantization levels. From Table 4.3, we can see that a fast key generation rate is at the expenses of bit agreement rate. Overall, 2-ary quantization is the best choice, and it can generate a common 128-bit key for two legitimate devices with 98.3% probability.

Table 4.3: Comparison of different quantization levels

	Bit rate (bit/sec)	Time to generate a 128-bit key	Probability of 100% match
2-ary quantization	28	4.6s	98.3%
4-ary quantization	37	3.5s	92.4%
8-ary quantization	43	3s	72.1%



(a) 2-ary quantization (b) 4-ary quantization

(c) 8-ary quantization (d) Impact of privacy amplification

Figure 4.15: Evaluation results.

4.6.5 Improvement of Key Randomness with Privacy Amplification

We now examine how the XOR function in privacy amplification helps to enhance the randomness of the final key. Figure 4.15(d) shows the entropy of the final key after privacy amplification. From the results, we can see that the distribution of entropy is closer to 1 after the XOR operation. We also notice that the entropy of the final keys varies from 0.94 to 1 which in turn indicates that the proposed method can extract secret keys with good entropy. Note that a drawback of using the XOR function is that the bit rate is reduced by a factor of 2 (we XOR two consecutive windows together). As the results in Table 4.3, the bit rate of 2-ary quantization can still achieve 28 bit/sec after privacy amplification.

4.6.6 Improvement of Bit Agreement Rate with ICA

We examine whether the application of ICA can improve the agreement rate. As ICA is applied on acceleration signals recorded from the smart watch only, we compute the bit agreement rate between keys generated from smart watch and devices placed at other locations by using raw acceleration values (without ICA) and extracted acceleration values (with ICA) respectively.

From the results in Figure 4.16, we can see a significant improvement in agreement rate after ICA. The maximum agreement rate of using raw acceleration values (without ICA) is near 50% which is like a random guess between 0 and 1. The results suggest that applying ICA can extract walking signals from arm swing signals effectively and thus improve the agreement rate significantly.

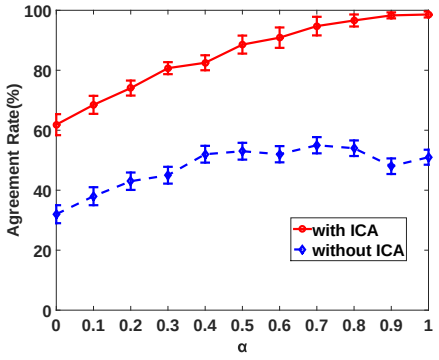


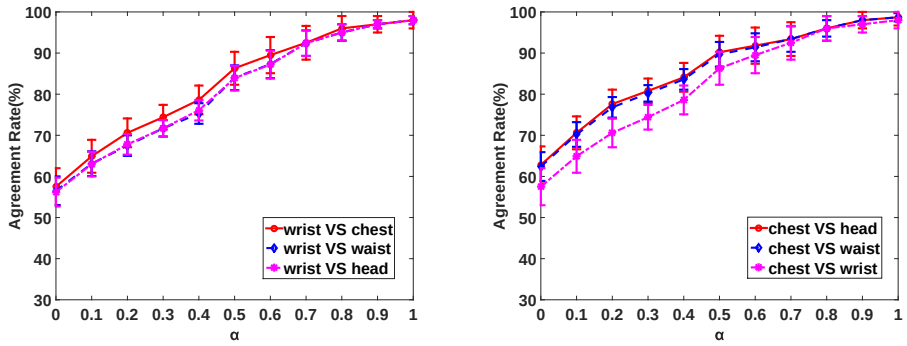
Figure 4.16: Impact of ICA

4.6.7 Bit Agreement Rate of Devices on Different Body Parts

We evaluate how well the proposed method performs for each body part: wrist, chest, waist, and head. For each body part, we compare the keys generated from other locations with the keys generated from this location. For example, in terms of wrist, we calculate the agreement rate by comparing the keys generated from wrist with keys generated from other locations (e.g., waist, chest, and head) respectively. As shown in Figure 4.17, we notice that the pairs of waist-to-chest and chest-to-head achieve the best agreement rate. This result is intuitive as sensors on the body trunk observe acceleration more similarly than sensors on the limbs.

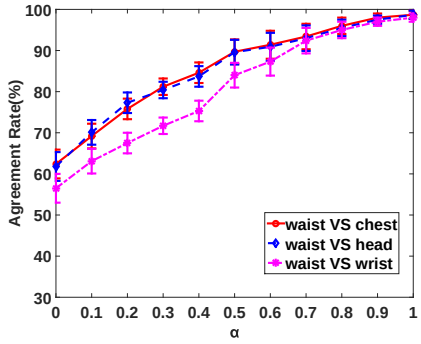
4.6.8 Randomness of the Final Key

Guaranteeing that the generated keys are random is crucial because they are intended for using as a cryptographic key. In order to validate the randomness of the final key, we apply the NIST suite of statistical tests [58] to all the keys generated from our dataset. The NIST statistical test gives the p-values of different random test processes, and the p-values indicate the probability that the key sequence is generated by a random process. Conventionally, if the p-value is less than 1%, the randomness hypothesis is rejected which means the key is not random. From

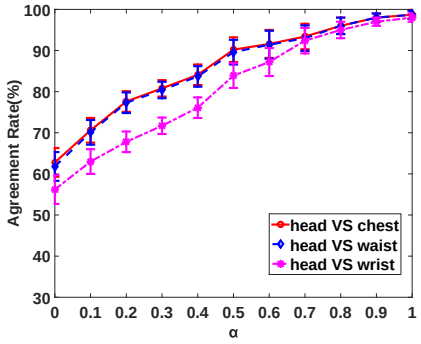


(a) Wrist VS other locations

(b) Chest VS other locations



(c) Waist VS other locations



(d) Head VS other locations

Figure 4.17: Bit agreement rate of different body parts.

Table 4.4, we can see that the p-values are all greater than 1% in the sense that the generated keys pass the randomness tests.

4.6.9 Security Analysis

We assume the presence of a passive adversary (eavesdropper) and an active attacker during an authentication session. The eavesdropper can listen to all the communications between Alice and Bob and knows the bit generation algorithm. The active attacker has complete communication control, i.e. can jam, forge and modify messages. Additionally, the adversary may mimic the walking style of the genuine user and start new protocol instances by injecting appropriate authentication request messages with multiple legitimate devices in parallel. We evaluate the robustness of the proposed system against the eavesdropper and active attacker by conducting the following two imposter attempt experiments.

- A passive impostor attempt is an attempt when an attacker tries to pair his device to a legitimate device by submitting his own walking signals.
- An active impostor attempt mimics the gait of the genuine user with the aim to pair with the devices of the genuine user.

Table 4.4: P-values of NIST Statistical Test.

NIST Test	p-value
Frequency	0.712248
FFT Test	0.557416
Longest Run	0.022491
Linear Complexity	0.380014
Block Frequency	0.978452
Cumulative Sums	0.986105
Approximate Entropy	0.996418
Non Overlapping Template	0.332475

Table 4.5: Mutual information among different devices

	Alice vs. Bob	Alice vs. Active attacker
Mutual Info. (bit)	1.42	0.21

The first experiment is conducted to evaluate the robustness to a passive imposter. For each location of one subject, we use the keys generated from the same location but from other subjects as passive imposter attempts. We then repeat this experiment by testing all the locations of the 20 subjects in the dataset. To evaluate the robustness against the second imposter attack scenario, we group the 20 subjects into 10 pairs. Each subject was told to mimic his/her partner's walking style and try to imitate him or her. Firstly, one participant of the pair acted as an attacker, the other one as a target, and then the roles were exchanged. The genders of the attacker and the target were the same. They observed the walking style of the target visually, which can be easily done in a real-life situation as gait cannot be hidden. Every attacker made 5 active impostor attempts by walking side-by-side the user. Figure 4.18 plots the bit agreement rate of passive imposter and active imposter, we find that the agreement rate of an active attacker is slightly higher than that of a passive attacker, but there is no regular pattern for agreement rate when α varies from 0 to 1. This phenomenon can be explained by two facts: first, the unique walking pattern of the genuine user is difficult to mimic, even an active attacker cannot produce similar walking patterns to the user. Even if an active attacker who can obtain approximately 50% agreement rate conducts brute-force attack, he still cannot guess the same key. Because the active attacker has no information about which bits are correct. Even a normal guesser can obtain 50% agreement rate as a cryptographic key contains 0 and 1 only. Therefore, he still needs 2^{128} attempts to guess the same 128-bit key which is infeasible in real-world scenarios. Moreover, the Reed-Solomon code may introduce more mismatching bits if the number of mismatching bits exceeds the correcting ability due to its nonlinear nature.

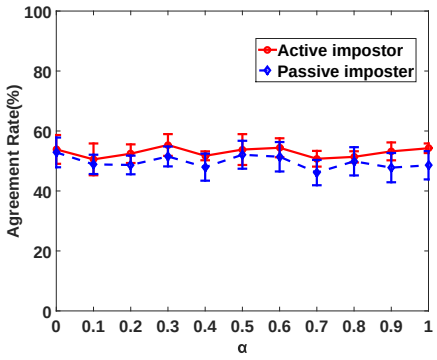


Figure 4.18: Agreement rate of impostors.

To further quantify the amount of information that can be inferred from mimicking the gait, we calculate pairwise mutual information among different devices in Table 4.5. We find that the legitimate devices on the same body can obtain 1.42 bits of information about the secret key. However, the active attacker, can only get 0.21 bits of information. This result suggests that the legitimate devices obtain 6 times more information about each other than the attacker.

The individual nature of walking gait provides our scheme security against passive eavesdroppers. Even if an active imposter can observe and try to mimic the walking style of the target, the results in Figure 4.18 show that he still cannot obtain a common secret key. However, an active attacker can impersonate Alice or Bob in the reconciliation stage and insert false values. Walkie-Talkie prevents such attack by the MAC method described in Section 4.5.2. A further concern to all key-agreement protocols is the man-in-the-middle attack (MITM). A MITM attack against our scheme rarely occurs as Alice and Bob exchange the offset (δ_{Alice} and δ_{Bob}) only instead of shared key during the reconciliation stage. Therefore, the shared key will not be compromised by MITM.

4.7 System Implementation

To validate the feasibility of the proposed key generation approach on wearable devices, we implemented the whole system using an Android OS application[2]. The system is implemented in Java and the implementation of FastICA is based on the Fastica Java library. The MAC algorithm described in Section 4.5.2 is implemented by keyed-hash message authentication code (HMAC-MD5). The sampling rate of the accelerometer is set as 50Hz and Bluetooth Low

[2]A video demonstration of the system can be found at the following URL: https://www.youtube.com/watch?v=YBFBJrNZy48

Energy (BLE) functionality is employed for wireless communication.

BLE is designed to provide significantly lower power consumption for devices with low power requirements. It introduces a new feature called *peripheral mode*, in which the data source can advertise and publish data without requiring to pair with the data requestor beforehand. BLE peripheral mode is designed for devices with resource constraints and which need to publish new data frequently. Therefore, we run the system in peripheral mode and advertise the data using broadcast packets. Bob organizes his data using *Generic Attribute Profile (GATT)* and encrypts the data to publish by AES. All the devices nearby including adversaries can receive the broadcast advertisements and read the public-available data from Bob. However, only Alice on the same body can generate the same key for data decryption. In this way, the private data is protected from reading by unauthorized devices.

Table 4.6 presents the system overhead (computation and energy consumption) of our system on a Moto E2 smartphone, which supports BLE peripheral mode. The computation time and energy consumption of each component are measured by averaging the results from running independent components separately and continuously for five minutes. Note that we do not consider the time for data collection (i.e., walking duration). The major components in Walkie-Talkie: the source separation (including ICA and component identification) and key generation take an average time of 108.3ms and 310.5ms respectively. When the scheme is fully employed, the computation time and energy consumption are 419.2ms and 198.5mJ respectively. The battery capacity of the Moto E2 smartphone is 2390 mAh (30.1 kJ), therefore, the energy cost of Walkie-Talkie amounts to 0.005% of the total capacity. We assume the smartphone with a targeted lifespan of one day which results in an energy budget of 1.25KJ per hour. To put this into perspective, with 5% of the budget per hour (62.5 J), Walkie-Talkie is capable of running approximately 317 times per hour, i.e., Walkie-Talkie can continuously run every 12 seconds. These results demonstrate that the proposed key generation approach has a low system overhead and can run in real-time on modern mobile devices. In addition, we also implement Diffie-Hellman (DH) protocol on the same smartphone and we find that Walkie-Talkie consumes approximately two times more energy than DH protocol.

Table 4.6: System overhead measured on Moto E2.

	Computation time (ms)	Energy consumption (mJ)
ICA	105.7	71.2
Component Identification	2.6	1.5
Key Generation	310.5	125.6
AES Encryption	0.2	0.1
AES Decryption	0.2	0.1
Total	419.2	198.5

4.8 Summary of This Chapter

In this chapter, we proposed and implemented a key generation approach that exploits the acceleration signals produced by gait to establish a common cryptographic key between two legitimate devices. By exploiting BSS and incorporating a multi-level quantization mechanism, Walkie-Talkie demonstrates superior effectiveness in performance. For example, when 2-ary quantization is employed, Walkie-Talkie can generate a common 128-bit key for two legitimate devices in 4.6s with 98.3% probability. Increasing quantization levels can improve the bit generation rate, but will decrease bit agreement rate. We also analyzed the security against various attackers. The proposed method obtains a security advantage from the fact that different people have distinctive walking styles. Finally, we prototyped the proposed scheme on a Motorola E2 smartphone to demonstrate the feasibility on contemporary mobile devices.

Although the evaluation results demonstrate the robustness and effectiveness of Walkie-Talkie, the current approach still has several limitations. First, Walkie-Talkie only addresses the problem of automatic pairing for devices located on the upper body of the user. For example, the acceleration sensors worn on the lower body of the user (e.g., leg) record half of gait cycle only, which makes key pairing significantly more challenging. Furthermore, we limit the positions of devices on the central body trunk in this study. There are more challenging scenarios in the real world, for example, the user may put his/her smartphone in the pocket or handbag. Some implantable devices such as pacemaker will be embedded inside user's body. The signals recorded in such positions will differ from those analysed in this study. In addition, Walkie-Talkie works for a pair of devices only because different pairs of devices will produce different mismatches. In many practical scenarios, it is essential to establish a common cryptographic key, known as a group secret key, among multiple smart devices belonging to a subject. These limitations are interesting research directions and we plan to study them in our future work.

Chapter 5

Gait-based User Authentication System Using Kinetic Energy Harvesting

Chapter Summary: Accelerometer-based gait recognition for mobile healthcare systems has became an attractive research topic in recent years. However, a major bottleneck of such a system is that it requires continuous sampling of the accelerometer, which reduces the battery life of wearable sensors. In this chapter, we present *KEH-Gait*, which advocates use of the output voltage signal from a kinetic energy harvester (KEH) as the source for gait recognition. *KEH-Gait* is motivated by the prospect of significant power saving by not having to sample the accelerometer at all. Indeed, our measurements show that, compared to conventional accelerometer-based gait detection, *KEH-Gait* can reduce energy consumption by 78.15%. The feasibility of *KEH-Gait* is based on the fact that human gait has distinctive movement patterns for different individuals, which is expected to leave distinctive patterns for KEH voltage as well. We evaluate the performance of *KEH-Gait* using two different types of KEH hardware on a data set of 20 subjects. Our experiments demonstrate that, although *KEH-Gait* yields slightly lower accuracy than accelerometer-based gait detection when a single step is used, the accuracy problem can be overcome by the proposed Multi-Step Sparse Representation Classification (MSSRC).

5.1 Introduction

Gait recognition using wearable sensors, such as accelerometers, has emerged as one of the most promising solutions for user authentication. Extensive previous studies have already demonstrated its feasibility in user authentication [49, 91, 148], but they have also shown that continuous accelerometer sampling drains the battery quickly. High power consumption of accelerometer sampling, which is typically in the order of a few milliwatts, also makes it challenging to adopt gait-based user authentication in resource-constrained wearables. Although power consumption may be not a big issue for wearables with large batteries such as smartphones, other wearables like IMDs suffer from short battery life because IMDs are long-lived devices and battery replacement requires surgical intervention [118].

A vision for wearable devices is to be battery-free (self-powered). A current trend in battery-free devices is to investigate kinetic energy harvesting (KEH) solutions to power the wearable devices [54, 135, 142, 101]. However, one fundamental problem in KEH is that the amount of power that can be practically harvested from human motion is insufficient to meet the power requirement of accelerometers for accurate activity recognition [72]. As reported in [54], the amount of power that can be harvested from human motion is only in the order of tens to hundreds of microwatts. This 2-3 orders of magnitude gap between power consumption and power harvesting is the biggest obstacle for realizing gait-based authentication in battery-less wearables. Although the power consumption of sensors has been reduced in recent years thanks to Ultra-Low-Power electronics, we believe in the near future energy harvesting will be used to augment or substitute batteries. For example, AMPY [2] has released the world's first wearable motion-charger which can transform the kinetic energy from user's motion into battery power. SOLEPOWER [9] produces smart boots that use user's steps to power embedded lights, sensors, and GPS. KINERGIZER has developed a small piezoelectric generator with the ability to harvest energy at low frequencies to produce as much as 200μW of power [6].

Motivated by this prospect, we propose gait recognition by simply observing the output voltages of KEH. The feasibility of the proposed idea is based on the observation that if humans have unique walking patterns, then the corresponding patterns of harvested power from KEH should be unique too. The proposed system offers several advantages. The major advantage of KEH-based gait recognition is the potential for significant power savings arising from not sampling an accelerometer at all. On the other hand, the output voltage can be used to charge the battery, thus further extending battery life. Finally, as the energy harvester will be

integrated in wearable devices in the near future, the output voltage can be naturally utilized for authentication purposes without introducing extra sensors. This makes it a promising solution for light-weight authentication for wearable devices. To the best of our knowledge, this is the first work that proposes and experimentally validates the feasibility of gait recognition using KEH.

The main contributions of this study are as follows:

- We propose a novel method of gait recognition, called KEH-Gait, which uses only KEH voltage as the source signal to achieve user authentication.

- We build two different KEH wearables, one based on a piezoelectric energy harvester (PEH) and the other on an electromagnetic energy harvester (EEH). Using these KEH devices, we evaluate gait recognition accuracy of KEH-Gait over 20 subjects. Our results show that, with conventional classification techniques, which operate over single step, KEH-Gait achieves approximately 6% lower accuracy compared to accelerometer-based gait recognition.

- We demonstrate that authentication accuracy of KEH-Gait can be increased to that of accelerometer-based gait detection by applying a novel classification method, called Multi-Step Sparse Representation Classification (MSSRC), which efficiently fuses information from multiple steps.

- Finally, using measurements, we demonstrate that microprocessors can read KEH voltage within 33 μs, which is two orders of magnitude faster than the time it takes to wakeup, interrogate and read acceleration values from typical 3-axis accelerometers. This means that with microprocessor duty cycling, KEH-Gait promises major energy savings over conventional accelerometer-based gait detection.

The rest of the chapter is structured as follows. Section 5.2 introduces trust models and attacker models of gait-based authentication system. Section 5.3 presents the system architecture of KEH-Gait. Prototyping of KEH wearables and gait data collection are described in Section 5.4. We present evaluation results in Section 5.5, and analyze power consumption in Section 5.6. We have a discussion of our work in Section 5.7 before concluding the chapter in Section 5.8.

5.2 Trust and Attack Models

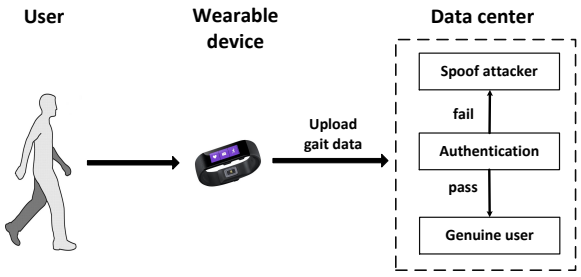


Figure 5.1: The overview of a typical healthcare monitoring system.

We envision the use of KEH-Gait primarily in resource-constrained healthcare monitoring wearable devices to authenticate the identity of the user to prevent spoofing attacks. KEH-Gait addresses the issue of short battery life by using an energy harvester to replace an accelerometer. In the near future, energy harvesters can even be integrated in the hardware system to achieve battery-free wearable devices. Figure 5.1 illustrates the work-flow of a typical healthcare monitoring system. In such a system, each user is given a unique user ID and a monitoring application which runs on a wearable device that can collect private sensor data and transmit them to the data center of a healthcare company. Before transmission, the device first collects gait data and transmits them to the sever. The server will then perform authentication to verify the user's identity by using the gait data. If the user passes authentication, the further private data like blood pressure or heart rate are then transmitted to the server. While if the user verification fails, i.e., the user spoofing attack is detected, the sensor data collected from this user's device will not be reported to the server. In the server, sensor data will be analyzed and processed by the healthcare company to derive user's physical and mental conditions. For instance, the measurements of heart-rate and blood pressure can be used to predict user's psychological conditions. A wide range of applications can also be enabled by such mobile healthcare systems and some examples are:

- Users' physical behaviors are often a reflection of physical and mental health and can be used by healthcare companies to facilitate early prediction of future health problems like depression [111].

- Health food companies can make advertisement by cooperating with healthcare related

applications such as "IDOMOVE"[1], e.g., providing discount coupons for users who walk more than 1hr a day.

For some applications, continuous authentication may be unnecessary. However, one-time validation of the user's identity is becoming insufficient for modern devices and applications that process sensitive data. A simple example is the mobile phone will lock the screen and demand users to enter their PIN every few minutes. Such situations might benefit from a seamless authentication approach that incorporates continuous verification of the user's identity. KEH-Gait leverages gait which is a common daily activity to provide unobtrusive and continuous authentication without user intervention. There are also many commercial products that provide biometrics-based continuous authentication systems such as BehavioSec [3] and Eyefluence [4]. It is worth mentioning that gait-based continuous authentication is not a general solution for all scenarios as people only walk 0.5-1 hour per day [1].

5.2.1 Trust Model

In this work, we assume the data collected by sensors built in the wearable devices are trustworthy. Also, our system trusts the communication channel between the wearable device and the healthcare company's server. We discuss the feasibility of our assumption as follows.

Tamper-resistant Sensor. An attack can physically accesses to the sensor or chipset and manipulate the recorded data. To make sure the device has not been modified, a healthcare company can apply tamper-resistant techniques [114]. As mentioned in [55], ARM TrustZone extension can also be used to ensure the integrity of the sensors [84].

Trusted Transmission. A man-in-the-middle(MITM) attack may occur when the device is communicating with the server. Therefore, the device and server should establish a secure communication channel. To address this attack, the healthcare company can install a digital certificate in the wearable device and the device will perform SSL authentication when communicating with the server.

5.2.2 Attack Model

The aforementioned mobile healthcare system is vulnerable to user spoofing attacks. For instance, an adversary can distribute his device to another person, and upload the data of that

[1]IDOMOVE: https://www.idomove.com/

person aiming to obtain healthcare benefits. Besides, multiple users may collude to launch user spoofing attacks to fool the mobile healthcare system. Therefore, the adversary model considered in this work focuses on impersonation attacks. We assume the presence of two types of impersonation attacks: a passive adversary and an active adversary. The passive adversary tries to spoof the healthcare system by using his own walking patterns. The active spoofing attacker knows the authentication scheme and will try his best to imitate the walking pattern of the genuine user to spoof the healthcare system.

The main goal of our system is to detect spoofing attacks. In fact, there are many other possible attacks to such healthcare system. We discuss these possible attacks and corresponding solutions. The first type of attacks we consider is replay attacks. In replay attacks, an adversary first records a measurement trace from another person. Then he replays the data trace to the monitoring device to fool the healthcare monitoring system. This attack can be easily detected as discussed in [55]. Although a MITM attack during communication between the device and server can be easily prevented, there is another type of MIMT in which an adversary may build a MITM monitor which bridges the user's skin and a wearable device. For example, once it detects a response message indicating healthy problems such as high blood pressure, it will manipulate the data and transmit the forged data to the server. This type of attack can be addressed by the scheme in [55]. Further potential threats include deriving the walking patterns by studying a video of the target's gait through computer vision techniques. We believe this is a potential vulnerability of unknown severity and leave it as future work.

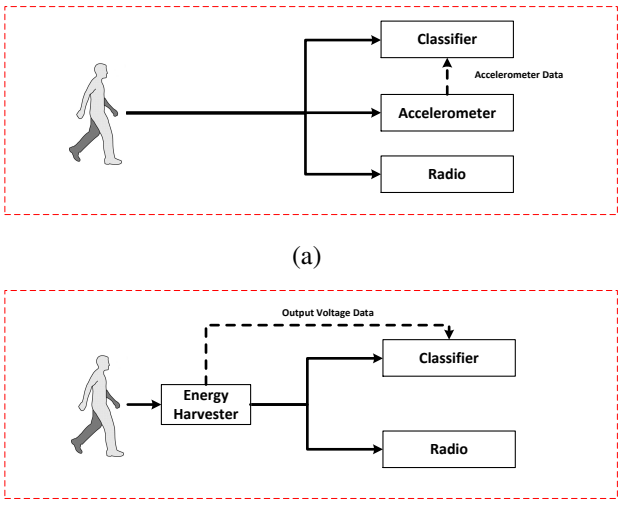


Figure 5.2: Gait recognition systems: (a) conventional accelerometer-based gait recognition and (b) KEH-Gait.

5.3 System Architecture of KEH-Gait

In this section, we discuss the proposed KEH-Gait framework in details. First, we compare KEH-Gait with traditional accelerometer based gait recognition systems. Figure 5.2(a) shows the pipeline of a traditional accelerometer-based gait recognition system, in which the accelerometer data are used to train a classifier for gait recognition. In contrast, as shown in Figure 5.2(b), the output voltage signal of the kinetic energy harvester will be exploited for gait recognition directly rather than powering the accelerometer. By not using the accelerometer, KEH-Gait can save the energy that is used to sample the accelerometer. The saved energy can be further used to power other components in the wearable device, such as the classifier and radio. The radio can be used to transmit the personal data to a base station or a server.

Figure 5.3 compares the output voltage signal from two types of energy harvester (EH) generated by two subjects when they are walking. These figures provide a clear visual confirmation that the voltage signal from the energy harvester contains personalized patterns generated by the subjects. This observation is promising as our goal is to recognize different subjects based on the output voltage signal of the EH when they are walking.

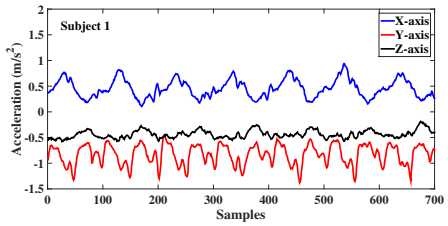
5.3.1 System Overview

As shown in Figure 5.4, the whole procedure of KEH-Gait consists of three parts: offline dictionary training, pre-processing of input signals, and classification.

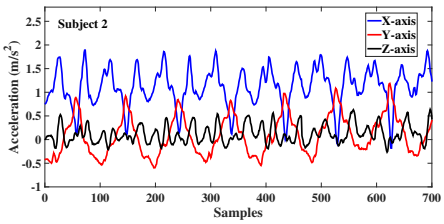
During the offline dictionary training phase, gait cycles are first segmented from the time series voltage signal and then interpolated into the same length. All detected cycles are passed to unusual cycles deletion to remove outliers of gait cycles. The obtained gait cycles are used to form the training dictionary A. After obtaining A, we apply the projection optimization algorithm in [124] to obtain a optimized projection matrix R_{opt}. Then the reduced training dictionary $\tilde{A} = R_{opt}A$ is used in the classifier as described in Section 2.1.1.

After the acquisition of the test signal, we again apply gait cycle segmentation and interpolation to obtain the gait cycles from the test signal. The same optimized projection matrix (as used for training) is used to reduce the dimension of the test signal and provide the measurement vector $\tilde{y}_i = R_{opt}y_i$, $i = 1, 2, \cdots, k$, and k is the number of obtained gait cycles.

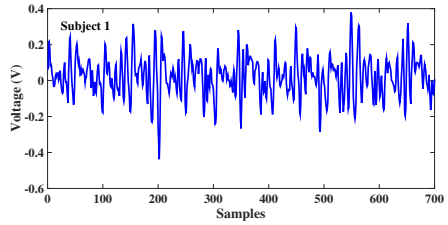
Now both the training dictionary $\tilde{A}$ and the measurements $\tilde{y}_i$ are passed to the classifier. The ℓ_1 classifier first finds the sparse coefficient vector x_i. Then the vectors of different gait cycles



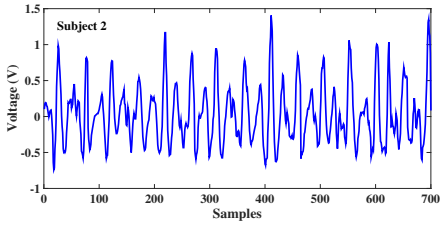
(a) Acceleration signal of subject 1



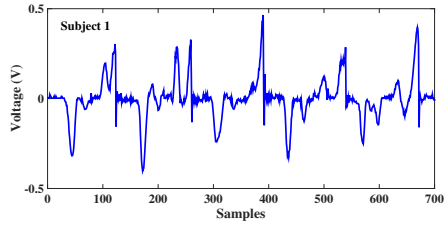
(b) Acceleration signal of subject 2



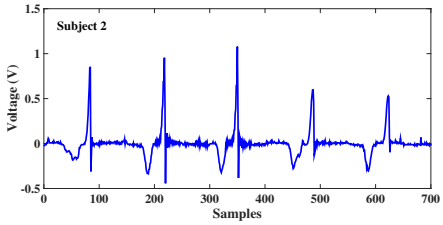
(c) Output voltage of subject 1 from a PEH device



(d) Output voltage of subject 2 from a PEH device



(e) Output voltage of subject 1 from an EEH device



(f) Output voltage of subject 1 from an EEH device

Figure 5.3: A comparison of the output voltage signal from different devices: (a) and (b) exhibit the acceleration signal from 3-axis accelerometer when two different subjects are walking; (c) and (d) plot the output voltage signal from a PEH device; (e) and (f) show the output voltage signal from an EEH device.

are fused based on a novel *sparse fusion* model, and the fused sparse vector is used to calculate the residuals. Finally, the identity is obtained by finding the minimal residual.

In the following sections, we detail the design of signal pre-processing, offline dictionary training, and classification in turn.

5.3.2 Signal Pre-processing

Gait Cycle Segmentation

In order to recognize a gait signal, it is essential that we separate the time series of walking periods into segments, such that each segment contains a complete gait cycle. The gait cycle can be obtained by combining two successive step cycles together as technically the gait cycle is across a *stride* (two steps). As mentioned in [20], typical step frequencies are around 1-2Hz, we apply a band-pass Butterworth filter [24] on the sampled data to eliminate out-band interference.

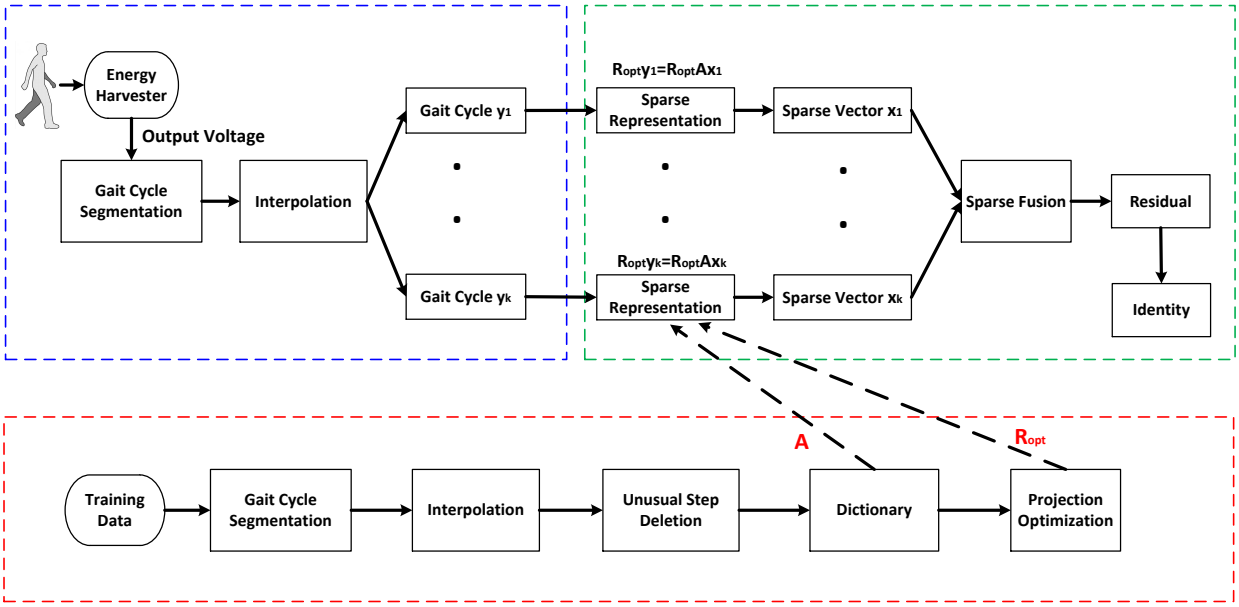


Figure 5.4: System flow chart of KEH-Gait

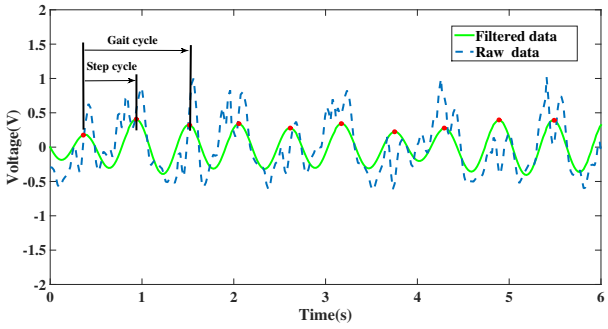


Figure 5.5: The time series of harvested energy: raw data (blue dash line), filtered data (green solid line).

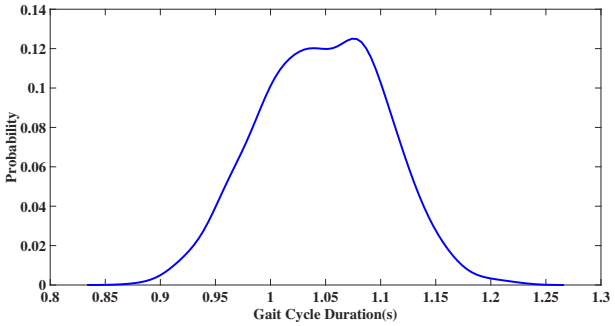


Figure 5.6: Distribution of cycle duration

The lower and upper cutoff frequency is set as 1Hz and 2Hz separately (filter order is 4). After filtering, the step cycles are separated by finding peaks associated with the heel strike as shown in Figure 5.5. Thereafter, the gait cycle is obtained by combining two consecutive step cycles together. After gait cycle extraction, the output voltage data are segmented into short gait

cycles based on the peak detection. Figure 5.6 presents the distribution of cycle duration (i.e. time length of stride) for 20 healthy subjects walking at their normal speed. We can see that most of the gait cycle ranges between 0.8-1.3s (80-130 samples at 100Hz sampling rate). These results in turn can be used to omit unusual gait cycles and exclude the cycles not produced by walking, i.e., the cycles which last less than 0.8s and exceed 1.3s are dropped.

Linear Interpolation

Detected cycles are normalized to equal length by linear interpolation because SRC requires vectors of equal length as input. As mentioned above, normal gait duration lies between 80 and 130 samples, we apply linear interpolation on the samples to ensure that they achieve the same length of 130 samples.

5.3.3 Offline Training

The training data are also passed to gait cycle segmentation and linear interpolation to obtain gait cycles with same length. In addition, we delete unusual cycles and optimize projection matrix to further improve recognition accuracy.

Deletion of Unusual Cycles

Unusual cycles caused by occasional abnormalities like temporary walking pauses or turning contains much noise that will deteriorate the recognition accuracy. Apart from deleting unusual cycles using cycle durations, the detected cycles are also passed to a function which further deletes unusual cycles. This function uses Dynamic Time Warping (DTW) distance scores to remove outliers from a set of cycles. Specifically, we first compute the DTW distance between the detected cycle and typical cycle. Thereafter, we delete unusual cycles by a simple threshold method, i.e., if the DTW distance of detected cycle and typical cycle is higher than a predefined value (12 in the proposed system), the detected cycle will be dropped. The typical cycle is the one which is assumed to represent the subject's gait signal. This is obtained by computing the the average of all cycles in the training data.

Projection Optimization

After unusual cycles removal, the remaining gait cycles obtained from training data are used to form the final training dictionary A. Motivated by a recent work [124], we apply the projection

matrix optimization method proposed in [124] to reduce the dimensionality of SRC while retaining the high classification accuracy. The projection matrix R_{opt} is learned from dictionary A based on Tabu search [53]. We refer the reader to [53] for more details.

5.3.4 MSSRC

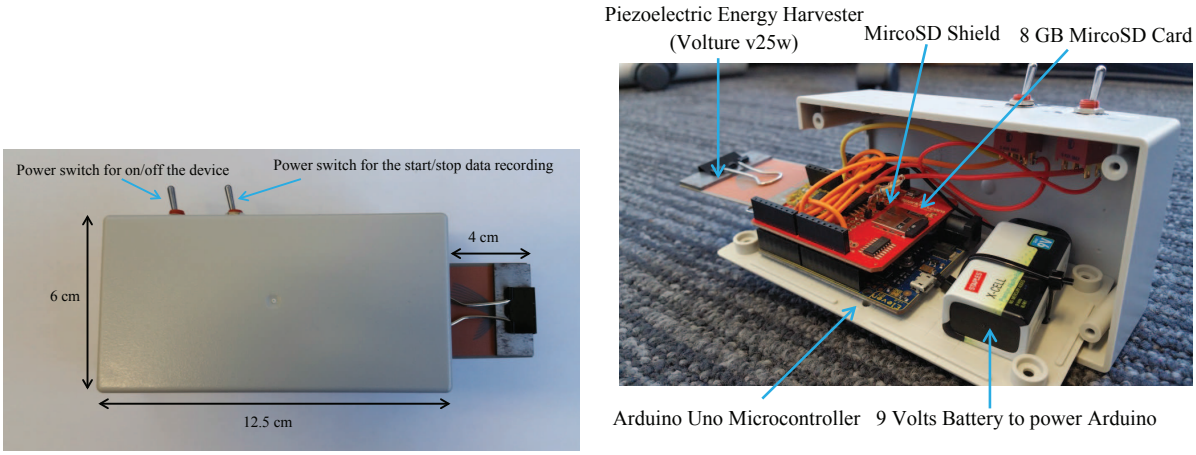
The MSSRC used in this study is similar to MVSRC which is mentioned in Section 3.2.1. Therefore, we briefly describe MSSRC and encourage the reader to refer to Section 3.2.1 for more details.

The key assumption behind MSSRC is that gait cycles obtained from consecutive gait cycles tend to have a high agreement on the sparse representations because each of the gait cycles from the same person should be linearly represented by the same class in the dictionary. Suppose we have acquired a set of M gait cycles (i.e., test vectors) from the test signal. We first calculate the *Sparsity Concentration Index* (SCI) of each test vector. Then we assign normalized weights to each of the test vectors. Finally, we can obtain the classification result by calculating the minimal residual using Eq 3.5. To identify whether the walker is the genuine user or imposter, we adopt the same principle in Section 3.2.1.

5.4 Hardware Platform and Data Collection

5.4.1 Proof-of-concept Prototype

PEH data logger. A data logger has been built to collect PEH voltage signals. The data logger includes a vibration energy harvesting product from the MIDÈ Technology called Volture, which implements the transducer to provide AC voltage as its output. Our hardware also includes a 3-axis accelerometer to record the acceleration signals, simultaneously with the voltage signal. An Arduino Uno has been used as a microcontroller device for sampling the data from the Volture. A sampling rate of 100Hz has been used for data collection. The sampled data has been saved on an 8GB microSD card which has been equipped to the Arduino using microSD shield. A nine-volt battery has been used to power the Arduino. To control the data collection, our data logger also includes two switches, one is an on/off switch and the other to control the start and stop of data logging. The Arduino measures voltage between 0 and 5 volts and



(a) (b)

Figure 5.7: PEH data logger: (a) the external appearance and (b) the internal details.

provides 10 bits of resolution (i.e., 1024 different values). Therefore, we calculated the corresponding output voltage from the measurements using the following formula $V = \frac{5*measurement}{1023}$. The hardware platform and the internal appearance of the data logger are shown in Figure 5.7.

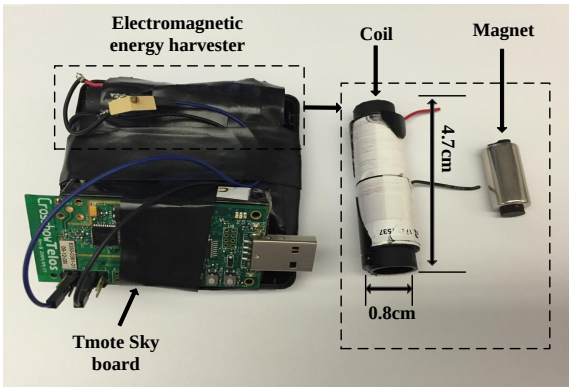


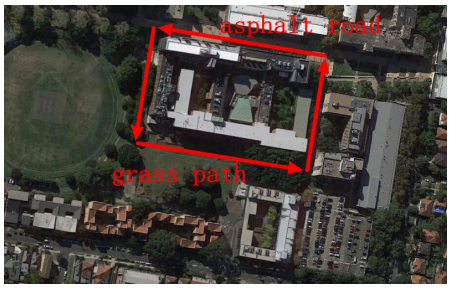
Figure 5.8: EEH data logger

EEH data logger. We also built an EEH data logger to collect voltage signals generated from an EEH device. The data logger contains a harvesting circuit, through which energy is generated by moving a magnet through an inductor. A Tmote sky board has been used as a microcontroller device for sampling the data from the inductor. A sampling rate of 100Hz has been used for data collection. The sampled data has been saved in the 48K Flash of the MSP430 microcontroller. Two AA batteries have been used to power the Tmote sky board. We use a button to control the data collection.

5.4.2 Data Collection

The dataset used to evaluate the performance of the proposed system consists of 20 healthy subjects (14 males and 6 females).During the data collection phase, the participants were asked to hold the data logger in one hand and walk at their normal speed (0.7-1.1m/s). The data collection is performed in several environments (indoor and outdoor) in order to capture the influence of different terrains. An illustration of indoor environment and outdoor environment is shown in Fig 5.9(a) and Fig 5.9(b). The terrain of the chosen outdoor environment varies including grass paths and asphalt roads. Each volunteer participated in two data collection sessions that was separated by one week. During each session, the participants were asked to hold the device (see Fig 5.9(c) and Fig 5.9(d)) and walked along the specific route shown in Figure 5.9(a) and Figure 5.9(b) for approximately 5 minutes. Based on the above description, the gait dataset is close to a realistic environment as it includes the natural gait changes over time and different environments (indoor and outdoor). In total, we have collected over 300

(a) Indoor experiment



(b) Outdoor experiment

(c) Holding PEH device

(d) Holding EEH device

Figure 5.9: The illustration of data collection.

seconds of samples for each subject from the EH devices as well as the accelerometer. We collect two voltage datasets by using the PEH and EEH devices, respectively, and perform gait cycle segmentation and unusual gait cycle deletion on both of the datasets, and finally we extract 200 gait cycles from each subject for evaluation. Although we collect data from the PEH and EEH prototype separately, the gait signals of the same subject can be deemed to be same as gait will not change in a short time.

As we collected data on the same day, the gait data from different devices only have slight differences because gait will not change in a short time. Note that gait signals can also be collected from other body locations such as the torso or thigh. In this study, we collect data from a device in the hand because the energy harvester will generate more energy, since energy output is based on kinetic motion. If the motion is larger the device will generate more energy. In addition, we do not separate gait signals from arm swing signals as we did in Chapter 4 when we performed authentication. This is because arm swing is also a weak biometric characteristic which can be used to improve authentication accuracy.

5.5 Evaluation

5.5.1 Goals, Metrics and Methodology

In this section, we evaluate the performance of the proposed system based on the collected dataset. The goals of the evaluation are threefold: 1) investigate the relation between recognition accuracy and sampling rate of accelerometer data; 2) compare the recognition accuracy of KEH-Gait with that of using accelerometer data; 3) compare the proposed classification method in KEH-Gait with several state-of-the-art classification algorithms.

In this work, we focus on the following three evaluation metrics:

- **Recognition accuracy**: it represents the percentage of correct classifications which is simply the number of true classifications over the total number of tests.
- **False positive rate (FPR)**: probability that the authentication system incorrectly accepts the access request by an imposter.
- **False negative rate (FNR)**: probability that the authentication system incorrectly rejects the access requests from the genuine users.

The recognition accuracy of KEH-Gait is obtained by using output voltage in one gait cycle as a test vector. For fair comparison, we perform the same signal processing and classification method on acceleration data. The only difference is the test vector is obtained by concatenating acceleration data along three axes in one gait cycle together. In the evaluation, we compare MSSRC with Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Naive Bayes (NB). The intuition for using SRC is that it has shown better performance than traditional classification methods (e.g., SVM and KNN) in recognition tasks such as face recognition [124, 140] and voice recognition [137]. SRC is known to be robust to noise because of its use of ℓ_1 optimization [124]. Thus, we use SRC in KEH-Gait and improve its performance by exploring the sparsity of testing vectors as discussed in Section 5.3.4. The parameters in SVM, KNN and NB are well tuned to give highest accuracy. For the KNN classifier we set the number of nearest neighbors as 10. For the SVM classifier, we choose a linear kernel function, and the soft margin constant is set to 10. We choose a normal Gaussian distribution for NB. For each classifier, we perform 10-fold cross-validation on the collected dataset. Specifically, we randomly split the dataset into 10 folds with equal size. Then, each fold is retained as the validation data for testing the classifier, and the remaining 9 folds are used as training data. The cross-validation process is then repeated 10 times, with each of the 10 folds used exactly once as the testing data. In the evaluation, we let k denote the number of gait cycles fused to perform classification and σ denote the compression rate. The compression rate means the number of projections/features over the dimension of the original feature vector. We plot the results of the average values and 95% confidence level of the recognition accuracy obtained from 10 folds cross-validation. The evaluation results of these experiments are presented in Section 5.5.4.

5.5.2 Recognition Accuracy v.s. Sampling Rate

In the first experiment, we evaluate the impact of sampling rate on the gait recognition accuracy of acceleration data. The goal is to investigate the relation between recognition accuracy and the consumed power of accelerometer, as the power consumption is directly related to the sampling rate. We use MSSRC as the classifier and calculate the recognition accuracy at different sampling rates by subsampling the acceleration data from 100Hz to 1Hz. As shown in Figure 5.10(a), the recognition accuracy increases with growing sampling rate. This is intuitive as the more measurements are sampled, the more information is available, and thus, enabling more accurate classification. However, the improvement diminishes after the sampling rate is greater

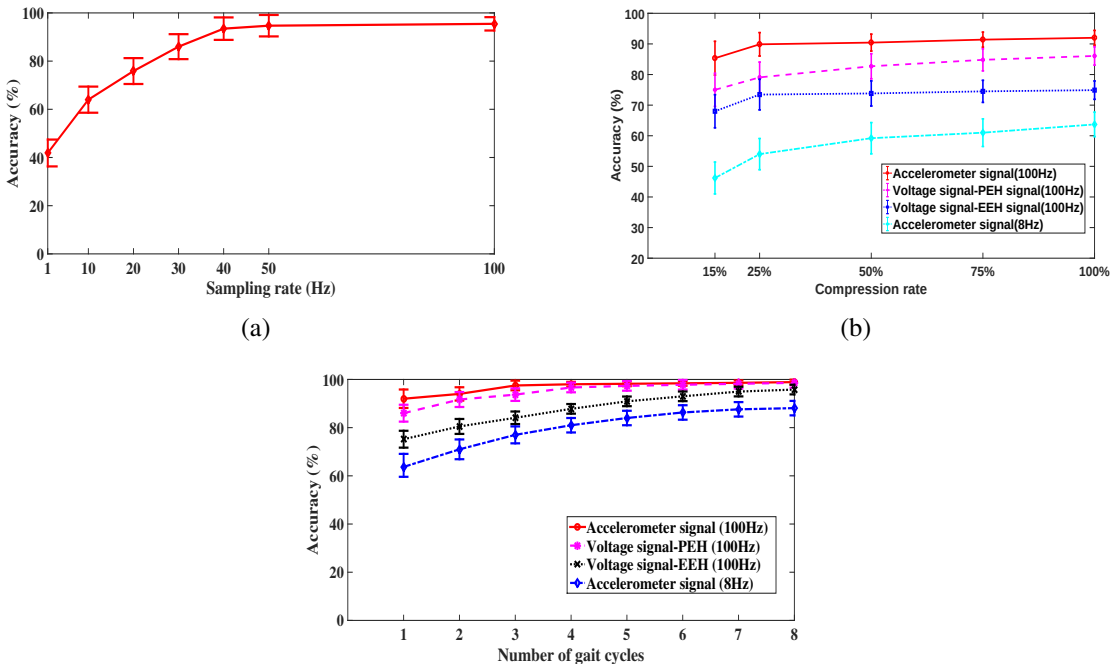


Figure 5.10: (a) Recognition accuracy vs sampling rate. (b) recognition accuracy under different compression rate when k=1. (c) recognition accuracy under different number of gait cycles when $\sigma = 75\%$.

than 40Hz. The results indicate that to achieve high recognition accuracy, a sampling rate of at least 40Hz is required. In the rest of the evaluation, we limit our discussion on sampling at 40Hz.

As we will discuss in Section 5.6.2, the power consumption of accelerometer-based system will increase significantly with the rising sampling frequency. Based on our measurement results, the accelerometer-based system consumes approximately $300\mu W$ at 40Hz to achieve accurate recognition. However, this consumption requirement is far beyond the actual power generated by the energy harvester (neither PEH, nor EEH). According to a recent theoretical study of energy harvesting from human activity [54], assuming 100% conversion efficiency, the power can be harvested from walking is only $155\mu W$. Unfortunately, in practical situations, according to our measurement results, the average power produced from walking is $19.17\mu W$ using EEH, and approximately $1\mu W$ using PEH which is not tuned specifically for human activity energy harvesting. In this case, due to the limited amount of power that is available to power the system, its sampling frequency will decrease below 40Hz. As a result, the recognition accuracy will dramatically decrease accordingly. The energy consumption of gait recognition is far beyond the energy that can be generated by the energy harvester; therefore, we cannot

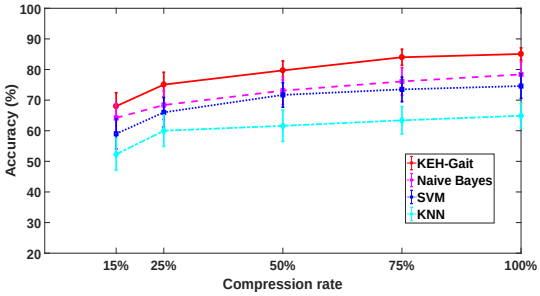
build a gait recognition system that powered by energy harvester only. The results highlight the necessity of using kinetic voltage signal to achieve gait recognition directly, instead of using the accelerometer signal. In the next subsection, we will show that the recognition accuracy of using kinetic voltage signal is comparable to that of using accelerometer data.

5.5.3 KEH-Gait v.s. Accelerometer-based System

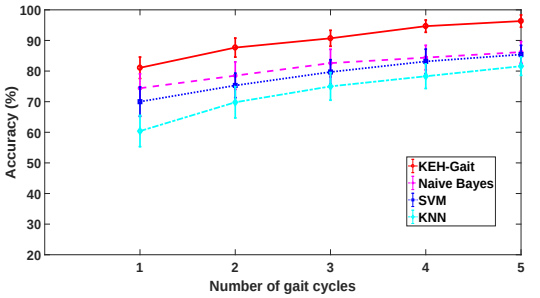
In this section, we investigate whether KEH-Gait can achieve comparable accuracy compared to accelerometer signals. When using accelerometer signal, we calculate the recognition accuracy at two different sampling rates: 1) raw sampling rate (100Hz) of the data logger; and 2) the highest achievable sampling rate of the accelerometer if it is powered by the energy harvester. From our dataset, the energy harvester can generate 19.17 μW on average from walking. Thus, according to the handbook of MPU9250 which is used in our prototypes, it can sample at most 8Hz if it is powered by the energy harvester.

In this experiment, we set $k = 1$ and calculate the recognition accuracy by varying the compression rate σ from 15% to 100%, and the results are plotted in Figure 5.10(b). We can see that the recognition accuracy of using voltage signal is significantly higher than that of using the accelerometer at a sampling rate of 8Hz. This suggests that the harvested power cannot support the accelerometer to sample at a high frequency which leads to low recognition accuracy; instead, using the voltage signal itself is able to achieve higher recognition accuracy. However, the recognition accuracy of using voltage signal is still approximately 6% (PEH) and 17% (EEH) below than that of using raw accelerometer signal when $\sigma = 100\%$.

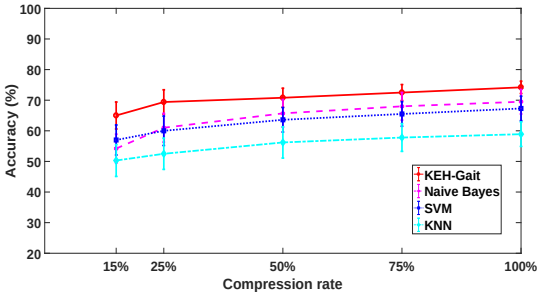
We now demonstrate that the recognition accuracy of using harvested power signal can be improved significantly by the proposed MSSRC, and it reaches a comparable recognition accuracy compared to using the raw accelerometer signal. In this experiment, we set $\sigma = 75\%$ as the accuracy improvement diminishes when the number of projections/features increased to 200 as shown in Figure 5.10(b). Then we calculate the recognition accuracy of KEH-Gait using accelerometer signal and voltage signal, while increasing k from 1 to 8. From the results in Figure 5.10(c), we notice that the recognition accuracy is improved significantly when more gait cycles are fused together. The result is intuitive as more information can be obtained to identify the subject by using more gait cycles. We also find that by using voltage signal of PEH, we can achieve a comparable accuracy compared to using raw accelerometer signal when $k = 8$, and the recognition accuracy of EEH is slightly lower (3%) than using raw accelerometer



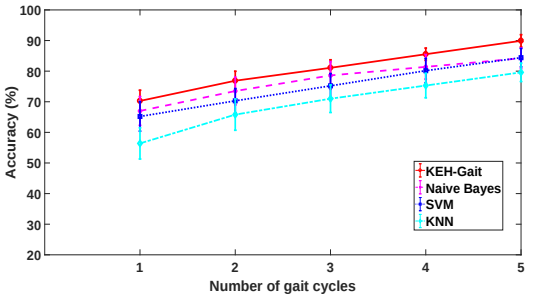
(a) PEH dataset with different compression rate ($k = 1$)



(b) PEH dataset with different k ($\sigma = 0.75$)



(c) EEH dataset with different compression rate ($k = 1$)



(d) EEH dataset with different k ($\sigma = 0.75$)

Figure 5.11: Comparison with other classification methods on two datasets (sample rate 40Hz).

signal. In the real applications, k can be tuned by the healthcare company to satisfy their own needs. For example, a larger k makes the system more secure from imposters while it sacrifices user experience because it will take more time to collect the required steps.

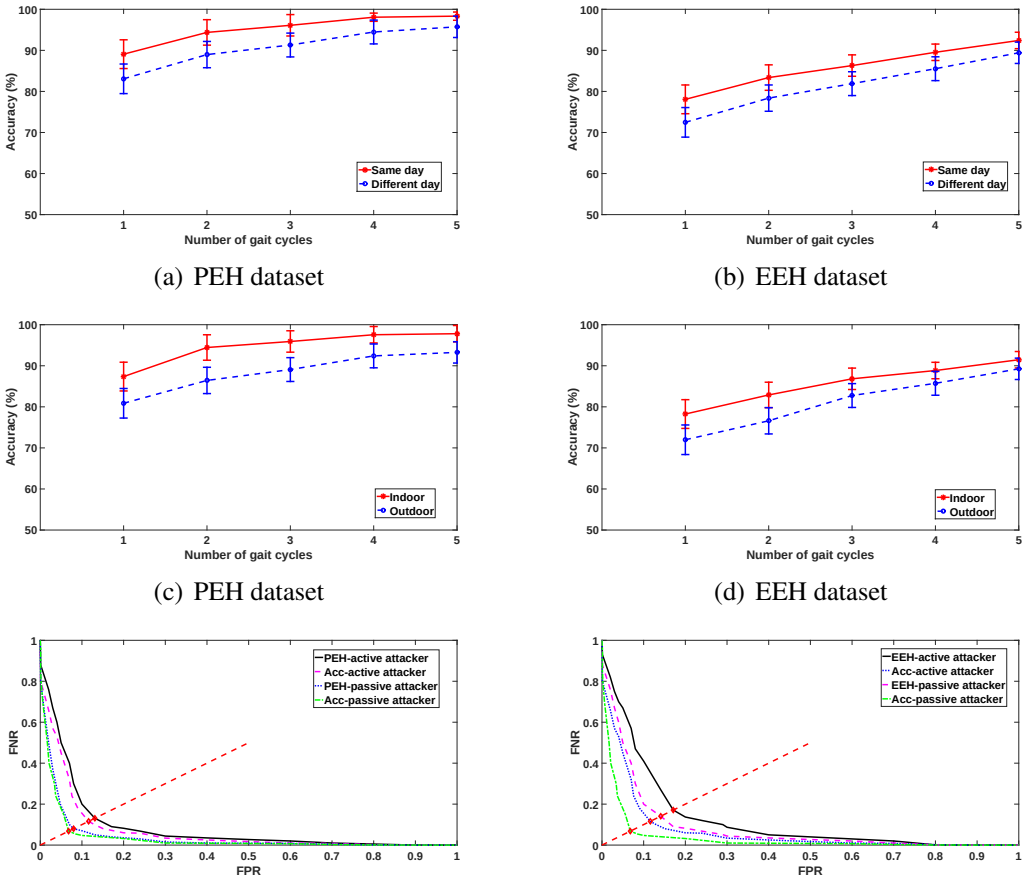
5.5.4 Comparison with Other Classification Methods

We now evaluate whether MSSRC outperforms other state-of-the-art classification algorithms. Specifically, we compare MSSRC with SVM, KNN, and NB. We perform comparison on two datasets separately.

Performance on PEH dataset. We follow the same experimental procedure in Section 5.5.3 to evaluate the recognition accuracy of different methods under different d (number of projections/features). From Figure 5.11(a), we find that KEH-Gait improves recognition accuracy by up to 7% compared to the second best classification method (i.e., NB). We further evaluate the recognition accuracy of SVM, KNN and NB by combining several gait cycles together. As KEH-Gait utilizes multiple gait cycles to find the final classification result, we apply the majority voting scheme to achieve a fair comparison. Specifically, we first obtain the identity of each gait cycle by using SVM, KNN and NB, then we apply majority voting scheme to combine the results together, the subject with the highest voting is declared to be the recognized person. Again, we set $\sigma = 75\%$ and calculate the recognition accuracy of different methods by varying k from 1 to 5 (number of gait cycles). From the results in Figure 5.11(b), we find that KEH-Gait consistently achieves the best performance and is up to 10% more accurate than the second best approach (i.e., NB). The improvement of MSSRC over other methods is because MSSRC exploits the sparsity of information from multiple gait cycles.

Performance on EEH dataset. We perform the same steps as above on the EEH dataset and plot the results in Figure 5.11(c) and Figure 5.11(d). The results show that KEH-Gait is 6% better than NB when $\sigma = 75\%$, $k = 1$, and 4% better than NB when $\sigma = 75\%$, $k = 5$. We also find that the overall performance on EEH dataset is lower than that on PEH dataset. We believe the drop on recognition accuracy is caused by the fact that the magnet is not sensitive to slight vibrations and motions.

The results in this section suggest that the proposed MSSRC in KEH-Gait can improve recognition accuracy significantly by fusing several steps together and it outperforms several state-of-the-art classification algorithms. Another straightforward method to apply SRC on multiple steps is to first apply SRC on each step and then obtain the final results by majority voting scheme. We found that MSSRC is approximately $3\% - 7\%$ more accurate than direct majority voting on our dataset since it exploits the sparsity information of multiple measurements.



(a) PEH dataset (b) EEH dataset

(c) PEH dataset (d) EEH dataset

(e) PEH dataset (f) EEH dataset

Figure 5.12: Evaluation results: (a)-(d) robustness to gait variations. (e)-(f) robustness against attackers.

5.5.5 Robustness to Gait Variations

To evaluate the robustness of KEH-Gait to gait variations, we conduct the following two experiments: different day evaluation and different environment evaluation. In this experiment, same day evaluation means the training set and test set are chosen from the sessions of the same day while different days evaluation chooses the sessions from two different days separated by 1 week. Similarly, in different environment evaluations, indoor evaluation means the training set and test set are chosen from indoor environment while outdoor evaluation chooses training data and test data from outdoor environment. We conduct this evaluation on PEH dataset and EEH dataset respectively. As the results in Figure 5.12(a) and Figure 5.12(b) show, the accuracy of different day is lower than the same day evaluation as the different days evaluation tends to produce more changes to gait. However, KEH-Gait can still achieve the accuracy of 95% and 89% on the two dataset respectively when more than 5 steps are used. This observation holds across the different evaluation environments. From Figure 5.12(c) and Figure 5.12(d), we can

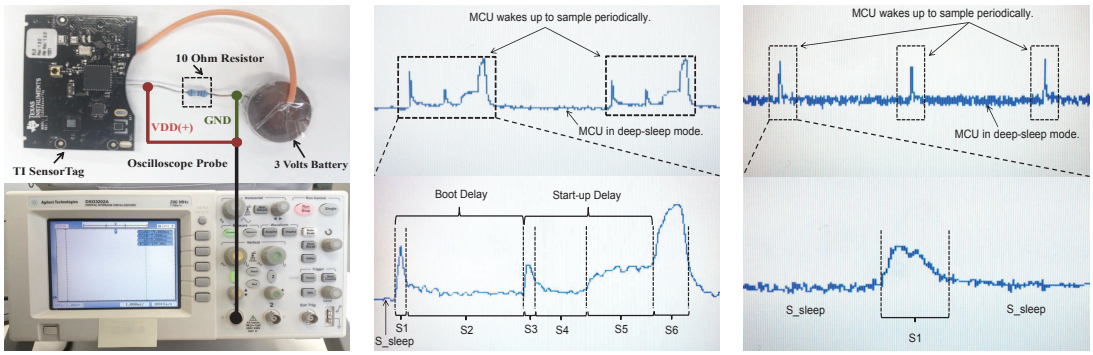
see that the outdoor environment achieves lower accuracy than the indoor environment because it includes several different terrains such as grass path and asphalt road. Gait changes can be caused many other factors such as speed and shoes etc.. We further discuss the influence of these factors in Section 5.7.2.

5.5.6 Robustness Against Attackers

As mentioned in Section 5.2, we assume the presence of a passive adversary and an active attacker during an authentication session. We evaluate the robustness of the proposed system against the eavesdropper and active attacker by conducting the following two imposter attempt experiments.

- A passive imposter attempt is an attempt when an imposter performs authentication using his own walking pattern. This attack happens when the genuine user passes his device to another person to spoof the healthcare system.
- An active imposter attempt means the imposter mimics the gait of the genuine user with the aim to spoof the healthcare system. This attack happens when the several users collude to fool the healthcare system.

The first experiment is conducted to evaluate the robustness to a passive imposter. In this experiment, we use the raw voltage signal from other subjects as passive imposter attempts. We then repeat this experiment by testing all the steps of the 20 subjects in the dataset. To evaluate the robustness against the second imposter attack scenario, we group the 20 subjects into 10 pairs. Each subject was told to mimic his/her partner's walking style and try to imitate him or her. Firstly, one participant of the pair acted as an imposter, the other one as a genuine user, and then the roles were exchanged. The genders of the imposter and the user were the same. They observed the walking style of the target visually, which can be easily done in a real-life situation as gait cannot be hidden. Every attacker made 5 active imposter attempts. The authentication accuracy is evaluated by FPR and FNR. In general, FPR relates to the security of the system, while FNR to the usability. An interesting point in the Decision Error Trade-off (DET) curve is the Equal Error Rate (EER) where FPR=FNR. For instance, an EER of 5% means that out of 100 genuine trials 5 are incorrectly rejected, and out of 100 imposter trials 5 are wrongfully accepted. We set $k = 5$ and vary the confidence threshold C to plot DET curve in Figure 5.12.



(a) Experimental setup for power measurement. (b) Profiling of accelerometer sampling. (c) Profiling of voltage sampling.

Figure 5.13: Measurement setup and results.

The results on two datasets are plotted in Figure 5.12(e) and Figure 5.12(f) respectively. The red dash line stands for the possible points where FPR is equal to FNR. The crossover (marked as a diamond) of the red dash line and FPR-FNR curve stands for the location of the EER. We notice that EER of KEH-Gait is 8.4% and 14.1% on the two datasets respectively, which means out of 100 passive imposter trials 8 are wrongfully accepted by using PEH and 14 are wrongfully accepted by using EEH. We also find that an imposter does benefit from mimicking the genuine user's walking style. The EER increases to 13.3% and 17.1% on the two datasets respectively. For the accelerometer-based system, the EER of a passive attacker and an active attacker are 6.8% and 11.6%, respectively. The results indicate that the PEH-based system can achieve comparable EER compared to the accelerometer-based system. The individual nature of walking gait provides our scheme security against impersonation attackers and the evaluation results are encouraging. The false negatives occur when the gait patterns of the imposter and user are close. This problem could be dealt with by using two factor authentication.

5.6 Power Consumption Profile

Battery lifetime is widely regarded as the major barrier to achieving long term human-centric sensing. Reducing system power consumption is a topic of very significant research efforts in both academic and industrial laboratories. In this section, we will conduct an extensive power consumption profiling of state-of-the-art wearable systems. Our study, which will be discussed in details later, shows that the proposed use of KEH power signal as the source for gait-recognition can greatly reduce the energy consumed by state-of-the-art accelerometer-based gait-recognition systems by 78.15%.

The energy consumption of our system consists of three parts: sensor sampling, memory reading/writing, and data transmission. We find that memory reading/writing consumes significantly less energy compared to the other two parts. A recent study [103] also investigates the energy consumption of different Random Access Memory (RAM) technologies, and their findings support our measurement results. According to their measurement, it only consumes 203pJ to write to (or read from) Static Random Access Memory (SRAM) which is used in SensorTag. That means if we collect 5s gait data at 40Hz, it only takes $5 \times 40 \times 203$ =40.6nJ to read or write data. Compared to the energy consumption of other parts, the energy consumed by SRAM is negligible. Therefore, we only consider the energy consumption of sensor sampling and data transmission in our evaluation.

5.6.1 Measurement Setup

The Texas Instrument SensorTag is selected as the target device, which is embedded with the ultra-low power ARM Cortex-M3 MCU that is widely used by today's mainstream wearable devices such as FitBit. The SensorTag is running with the Contiki 3.0 operating system. The experiment setup for the power measurement is shown in Figure 5.13(a). In order to capture both the average current and the time requirement for each sampling event, the Agilent DSO3202A oscilloscope is used. As shown in the figure, we connect the SensorTag with a 10Ω resistor in series and power it using a 3V coin battery. The oscilloscope probe is then connected across the resistor to measure the current going through.

5.6.2 Energy Consumption of Sensor Sampling

Power Consumption of Sampling Accelerometer

The SensorTag includes 9-axis digital MPU9250 motion sensor combining gyroscope, digital compass, and accelerometer. During the power measurements, we only enable the 3-axis accelerometer and leave all the other sensors turned off. The acceleration signal is sampled using the Inter-Integrated Circuit (I^2C) bus with a sampling frequency of 25Hz. Note that, it is also possible for the wearable devices to use analog accelerometers, which can be sampled through analog-to-digital converter (ADC) instead of I^2C bus. Sampling analog accelerometers could avoid power consumption and additional time requirement due to the I^2C bus, but at the expense of some processing costs in analog to digital converting. While it is not immediately obvious

whether analog accelerometer sampling would be less or more power consuming relative to the digital counterpart, a detailed measurement study [23] indicates that digital accelerometer is more power efficient than the comparable analog ones from the same manufacturers.

Table 5.1: States of accelerometer sampling, which takes 17.2ms in total and consumes 322μW.

State	**Description**	**Time** (ms)	**Power** (uW)
S1	MCU wakes up to boot accelerometer	0.6	768
S2	MCU sleep when accelerometer starts booting	7.2	72
S3	MCU walks up to initiate accelerometer	0.6	480
S4	MCU sleep when accelerometer starts initializing	3.2	72
S5	Accelerometer is turning on	4	480
S6	MCU walks up to sample accelerometer signal	1.6	1440
S_sleep	MCU in deep-sleep mode; accelerometer power-off	null	6

Figure 5.13(b) shows the details of accelerometer sampling energy profile. As shown, each accelerometer sampling event can be divided into six states. At the beginning of each event, the MCU is woken up by the software interrupt from the power-saving deep-sleep mode (S_sleep), and it boots the accelerometer (S1) before going back to sleep. During S2, the accelerometer starts to power up while the MCU is in sleep mode. Then, after one software clock tick (7.8 ms in Contiki OS), the MCU wakes up again by the interrupt to initialize the accelerometer (S3) and then goes back to sleep. The accelerometer starts initializing in S4 and turning on in S5. Finally, MCU wakes up in S6 to sample the acceleration signal and then goes back to deep-sleep again. The average power consumption and time requirement for each state are shown in Table 5.1.

Power Consumption of Sampling KEH

In this subsection, we investigate the power consumption in sampling the voltage signal of the power source. During the measurement, MCU is programmed to periodically sample the voltage of the lithium coin battery with 25Hz sampling rate. The MCU reads voltage signal through ADC. Figure 5.13(c) shows the details of voltage sampling. Similar to the accelerometer, the MCU goes back to deep-sleep mode after each sampling event. However, sampling the voltage takes only 0.6ms, which is much shorter than the 17.2ms required by the accelerometer sampling. This is because the MCU can read the voltage signal directly without having to prepare the hardware to be powered-up, and the voltage signal to be prepared by the power source. The details of power consumption and time duration for voltage sampling event are shown in Table 5.2.

Table 5.2: States of voltage sampling.

State	**Time** (ms)	**Power** (μW)
S1	0.6	480
S_sleep	null	6

Energy Consumption Comparison

We now compare the energy consumption of sampling accelerometer and KEH. In general, for the duty-cycled gait-recognition system, the average power consumption in data sampling, P_{sense}, can be obtained by the following equation:

$$P_{sense} = \begin{cases} \frac{T_S \times n}{1000} P_{sample} + (1 - \frac{T_S \times n}{1000}) P_{sleep} & \text{if } 0 \leq n \leq \frac{1000}{T_S}, \\ P_{sample} & \text{if } \frac{1000}{T_S} < n. \end{cases} \tag{5.1}$$

where, P_{sample} is the average power consumption in the sampling event (either sampling acceleration or KEH signal), and P_{sleep} is the average power consumption when the MCU is in deep-sleep mode (with all the other system components power-off). n is the sampling frequency, and T_S is the duration of time (in milliseconds) spent in a single sampling event. Based on the measurement results given in Table 5.1 and Table 5.2, we can obtain the average power consumption for the accelerometer sampling event which is 322μW with a time requirement of 17.2ms, and 480μW with a duration of 0.6ms for the KEH sampling event. Then, based on Equation 5.1, we get the power consumption in data sampling for both accelerometer-based and KEH-based gait-recognition systems with different sampling frequencies. The results are compared in Figure 5.14. It is clear to see that the proposed KEH-Gait achieves significant power saving in data sampling, comparing with the conventional accelerometer-based gait-recognition system. More specifically, given the analysis shown in Figure 5.10(a), a sampling rate higher than 40Hz is needed to achieve high recognition accuracy. With a 40Hz sampling frequency, in case of data sampling, KEH-Gait consumes 17.38μW, while the power consumption of the accelerometer-based system is 230.74μW.

As can be seen from Figure 5.10(c), to achieve the same recognition accuracy, it needs to collect 3 gait cycles for the accelerometer-based system and 5 gait cycles for the KEH-based system. If we assume one gait cycle takes 1s (the average time of one gait cycle is between 0.8s-1.2s), this results in 86.9μJ and 692.22μJ energy consumption in data sampling for KEH-Gait and accelerometer-based system, respectively.

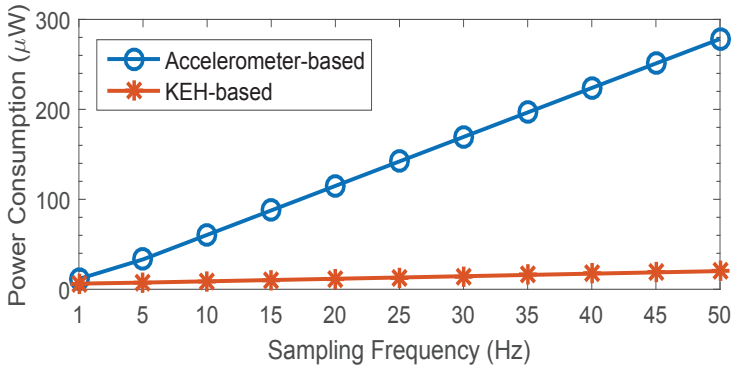


Figure 5.14: Power consumption comparison.

5.6.3 Energy Consumption of Data Transmission

Next, we evaluate the energy consumption of transmitting acceleration and KEH voltage data via Bluetooth. We conduct power measurement of the Bluetooth Low Energy (BLE) beacon using the embedded CC2650 wireless MCU in the SensorTag. With the 40Hz sampling rate and 75% compression rate, KEH-Gait generates 200 voltage samples every five seconds. This results in 300 bytes data to be transmitted in total (2 bytes for each of the 12-bits ADC voltage reading). This consumes an average power of 2.72mW with a transmission time of 52ms, which results in 106.08μJ of energy consumption. On the other hand, as 3-axis acceleration data is collected for 3s, it results in 540 bytes of data and the energy consumption of transmitting those data is 190.94μJ.

5.6.4 Total Energy Saving Analysis

After obtaining the energy consumption of sensor sampling and data transmission, we investigate the potential of KEH-Gait for energy saving. Based on the measured results, the energy consumption of KEH-Gait to complete one authentication is approximately 192.98μJ, which has reduced the energy consumption of the accelerometer-based system (883.16μJ) by 78.15%.

A recent study [23] tested the power consumption of six most commonly available accelerometers, and they found that when the sampling rate is 50 Hz, the mean power consumption of these accelerometers is 1542μW, and the minimum power consumption is 518μW. These accelerometers consume more power than the one used in our experiments. These results indicate that KEH-Gait is still superior to most commonly used accelerometers in terms of energy savings. On the other hand, the power consumption of accelerometers can be further reduced by

sensor batching technique. For example, ADXL345 can store 32 samples automatically without waking up the CPU. However, similar optimisation technique can also be integrated in the KEH-based system to reduce the system level energy consumption. In this way, the energy consumption of both accelerometer-based system and KEH-based system will be reduced. We defer the design optimization of buffer-enabled KEH-Gait to our future work.

5.7 Discussion

5.7.1 PEH v.s. EEH

In this study, we analyze the feasibility of using voltage signal generated from an energy harvester for gait recognition purposes. Specifically, we focus on two types of kinetic energy harvester: PEH and EEH. Our study demonstrates the harvested power signal caused by human gait motions can be used to identify different individuals. Table 5.3 summarizes a comparison between the PEH and EEH devices we used in this work. The first observation we can have is that the PEH we used achieves higher recognition accuracy and generates more energy than EEH when the user is holding the device in the hand and walking normally. The results can be explained by our observation that the EEH contains a heavy magnet which is not sensitive to weak vibrations and motions (compared in Figure 5.3(c) and 5.3(e)). This results in a roughly 10% difference in the recognition accuracy.

Table 5.3: Comparison between PEH and EEH used.

	Size (cm $\times$ cm $\times$ cm)	Weight (grams)	Accuracy (%)	Power (uW)	Cost (USD)
PEH	$4.6 \times 3.3 \times 0.1$	23.5	86.1	1	157
EEH	$4.7 \times 0.8 \times 0.8$	65	75.2	19.17	37.5

In addition to the system performance, another important characteristic in designing a wearable device is the form factor and weight. In case of the PEH device, we built it upon the Volture V25W PEH energy harvester with a 4.6cm $\times$ 3.3cm $\times$ 0.1cm form factor. And it can be further reduced to 2.2cm $\times$ 0.4cm $\times$ 0.1cm by exploiting smaller harvester products such as the PPA-1022. On the other hand, the EEH device requires large mass displacement to ensure the free movement of magnet which makes it difficult to reduce the form factor. Moreover, in order to generate more power from the PEH device, a 20 grams tip mass is attached to the PEH device and results in an overall weight of 23.5 grams. Fortunately, with current advancement in

PEH design, the overall weight of the PEH can be reduced to less than 10 grams without significantly sacrificing the output power. In comparison, the EEH device includes a heavy magnet and results in a weight approximately 65 grams in total. Given the above facts, we believe that PEH is more convenient to be embedded in future wearable devices that have strict constraints in size and weight.

Finally, the price of the PEH we used in our prototype is approximately 157USD (Volture V25W), while the cost of the EEH we used is 37.5USD. Although both of the prices can be largely reduced with a larger quantity of purchase, the cost of building the PEH device is higher than that of the EEH device.

5.7.2 Factors Affecting Gait Recognition

The most important factor that affects the accuracy of sensor-based gait recognition system is the position of the sensor. Previous studies have investigated gait recognition by attaching sensors on different positions of the body, such as lower leg [48], waist [91], hip [51] and hand [133]. They found that the lower body (e.g., ankle, thigh) achieves higher accuracy than upper body (e.g., arm and wrist) [148]. There are also many other factors that may impact the accuracy of a gait-based recognition system, such as shoes, clothes, walking speed and terrain. Previous studies have shown that the accuracy will decrease when the test and training samples of the person's walking are obtained using different shoe types and clothes [46]. Indeed, as shown in Section 5.5.5, the accuracy of KEH-Gait decreases when session 1 is used for training and session 2 is used for testing. The dataset used in the experiment is challenging as it includes the natural gait changes over time (two sessions separated by 1 week), as well as gait variations due to changing in clothes, terrain and shoes. However, KEH-Gait can still achieve an accuracy of 95% and 89% on the two datasets respectively by the proposed MSSRC. This in turn demonstrates the robustness of KEH-Gait to gait variations. The focus of our study is to demonstrate the feasibility of gait recognition using KEH and improve its performance. In fact, there has been several attempts to study the relationship between recognition performance and different factors [46, 104]. For example, in terms of walking speed, Muhammad and Claudia [104] found that normal walking has the best results and fast walking is slightly better than slow walking. As for different types of terrains, they reported that gravel walk has better results than grass and inclined walk. We encourage the reader to refer to [46, 50, 104] for more details.

5.8 Summary of This Chapter

In this chapter, we presented KEH-Gait, a kinetic energy harvesting signal based gait recognition system for user authentication. By not using the accelerometer, the proposed KEH-Gait eliminates the need for powering the accelerometer, making gait recognition practical for future self-powered devices. We also designed and implemented hardware platforms to collect voltage data from two types of kinetic energy harvesters (PEH and EEH). Evaluation results based on a data set of 20 subjects show that KEH-Gait is able to achieve comparable recognition accuracy compared to accelerometer based gait recognition system by the proposed MSSRC. Besides, KEH-Gait improves recognition accuracy by up to 10% compared to several state-of-the-art classification algorithms. More importantly, compared to conventional accelerometer-based gait detection, KEH-Gait can reduce energy consumption by 78.15%. To the best of our knowledge, this is the first work that experimentally validates the feasibility of gait recognition using KEH, and our results show that the output voltage signal of energy harvester is a promising informative signal for wearable authentication system. We also analyse and compare the two techniques used in our evaluation, we find that PEH is more convenient to be embedded in future wearable devices that have strict constraint in size and weight. However, compared to EEH, the disadvantages of PEH are its high cost and limited output power.

Chapter 6

Conclusions and Future Work

In this chapter, we first conclude the findings of this thesis. Then we finish the thesis by pointing out the future challenges with a short discussion of the potential solutions.

6.1 Conclusions

In recent years, with the rapid development of embedded technology, wearable devices such as smartphone and smart glass have penetrated to everyone's life. These smart devices usually come with a growing set of cheap powerful embedded sensors, such as accelerometer, digital compass, gyroscope, GPS, microphone, and camera. We believe that sensor-equipped mobile devices will revolutionize many sectors of our life, including healthcare, social networks, environmental monitoring, and transportation. Therefore, we focus on studying the recognition systems on these mobile devices and have addressed the following three research questions in this thesis:

1. **How the computational cost of face recognition system on smart glasses can be reduced, and how can face recognition performance on smart glasses be improved with Inertial Measurement Unit (IMU) sensor data?**

   In the first topic, we addressed the problem of assisting a user to recognize people with the help of smart glasses. Specifically, we proposed and implemented a novel sensor-assisted face recognition system which runs locally on smart glasses by exploiting the information from both the camera and sensors on smart glasses to improve the recognition accuracy and reduce the energy consumption. As shown in our evaluation, the proposed system

improves recognition accuracy by up to 15% compared with OpenCV face recognition methods while achieving a similar system cost.

2. **How can two legitimate devices belonging to the same user establish a secure communication channel in a user friendly manner?**

   In the second topic, we addressed the problem of pairing two wearable devices on the same body. The intuition of the proposed key generation approach is that the devices on the same body experience similar motion signals that are produced by the unique walking pattern of the user. Therefore, the unique gait signal can be exploited as shared information to generate secret keys for all on-body devices. We experimentally demonstrated that a common 128-bit key can be successfully generated by two independent wearable devices on the same body in 98.3% of cases, while the scheme also provides adequate security guarantees against impersonation attacks.

3. **How does gait recognition using a Kinetic Energy Harvester compare to more conventional gait recognition using accelerometers in terms of accuracy, energy-efficiency and robustness to attack?**

   In the last topic, we proposed a novel gait-based user authentication system by using the output voltage signal of the kinetic energy harvester. We first designed two prototypes by using energy harvesting techniques. Then we demonstrated the feasibility of achieving gait recognition using kinetic energy harvester. Compared to conventional accelerometer-based gait recognition system, the proposed system can reduce energy consumption by 78.15%.

6.2 Future Work

With the rapid development of wearable technology, wearable devices will play a significant role in our everyday life. With this trend, the wearable devices will spark a new set of applications and become a hot research topic. Although we have addressed three challenges, many problems still remain unexplored. Therefore, we point out several potential areas of future work.

6.2.1 Key Generation System for Multiple Devices

We have solved the problem of generating the same key for two wearable devices on the same user. However, the proposed system takes approximately 5s to generate a 128-bit key. We seek to improve the bit rate by using more complex key generation algorithms. Another problem is that the proposed system only works for two independent devices, and it fails to generate the same key for several devices. In the future, we also plan to study how to generate a group key for more devices on the same body by using the Fuzzy Vault scheme [69].

6.2.2 Context-aware Gait-based Authentication System

Although gait-based recognition has been well explored in the literature, there remain several challenges with wearable devices. Firstly, due to the high flexibility of arm motions, people could walk in a variety of ways (e.g., walk normally or walk upstairs). The walking pattern when the user is walking normally is clearly different from that of walking upstairs. The large majority of existing studies on accelerometer-based gait recognition have used a very restrictive experimental setup where the performance evaluation was conducted on a dataset collected from a controlled laboratory environment and the participants are asked to walk normally. As the pervasiveness of wearable devices in the wild, these is a need for robust and efficient authentication system in a realistic environment. One possible solution is to train different classification models for different activities. When the user is walking, the device will first detect the specific activity of the user (e.g., walking upstairs or walking while using a mobile phone). Then the classification is performed on the specific training model to improve classification accuracy. For example, if the user is walking upstairs, then the classification will be performed on the model trained from walking upstairs.

Bibliography

[1] http://eyefluence.com/.

[2] AMPY. http://www.getampy.com/ampy-move.html//.

[3] BehavioSec. https://www.behaviosec.com/.

[4] Eyefluence. http://eyefluence.com/.

[5] IoT (Internet of Things) for Residential Customers. http://www.navigantresearch.com/research/iot-internet-of-things-for-residential-customers/.

[6] KINERGIZER. http://kinergizer.com/.

[7] Nametag. http://www.nametag.ws/.

[8] OpenCV. http://opencv.org/.

[9] SOLEPOWER. http://www.solepowertech.com/#new-page//.

[10] Timo Ahonen, Abdenour Hadid, and Matti Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, 2006.

[11] Fahd Albinali, Stephen Intille, William Haskell, and Mary Rosenberger. Using wearable activity type detection to improve physical activity energy expenditure estimation. In *Proceedings of the International Conference on Ubiquitous Computing (Ubicomp)*, pages 311–320. ACM, 2010.

[12] Fernando Alonso-Martín, María Malfaz, João Sequeira, Javier F Gorostiza, and Miguel A Salichs. A multimodal emotion detection system during human–robot interaction. *Sensors*, 13(11):15549–15581, 2013.

[13] Louis Atallah, Omer Aziz, Benny Lo, and Guang-Zhong Yang. Detecting walking gait impairment with an ear-worn sensor. In *Proceedings of the 6th International Workshop on Wearable and Implantable Body Sensor Networks (BSN)*, pages 175–180. IEEE, 2009.

[14] Richard Baraniuk, Mark Davenport, Ronald DeVore, and Michael Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263, 2008.

[15] SP Beeby, MJ Tudor, and NM White. Energy harvesting vibration sources for microsystems applications. *Measurement science and technology*, 17(12):164–175, 2006.

[16] Peter N. Belhumeur, João P Hespanha, and David J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720, 1997.

[17] Charles H Bennett, Gilles Brassard, and Jean-Marc Robert. Privacy amplification by public discussion. *SIAM Journal on Computing*, 17(2):210–229, 1988.

[18] Daniel Bichler, Guido Stromberg, Mario Huemer, and Manuel Löw. Key generation based on acceleration data of shaking processes. In *Proceedings of International Conference on Ubiquitous Computing (Ubicomp)*, pages 304–317. ACM, 2007.

[19] Battista Biggio, Zahid Akhtar, Giorgio Fumera, Gian Luca Marcialis, and Fabio Roli. Security evaluation of biometric authentication systems under real spoofing attacks. *IET Biometrics*, 1(1):11–24, 2012.

[20] Agata Brajdic and Robert Harle. Walk detection and step counting on unconstrained smartphones. In *Proceedings of International Conference on Ubiquitous Computing (Ubicomp)*, pages 225–234. ACM, 2013.

[21] Arnaud Buchoux and N L Clarke. Deployment of keystroke analysis on a smartphone. In *Proceedings of Australian Information Security Management Conference*, pages 48–56, 2008.

[22] Attaullah Buriro, Bruno Crispo, Filippo Del Frari, and Konrad Wrona. Touchstroke: smartphone user authentication based on touch-typing biometrics. In *Proceedings of International Conference on Image Analysis and Processing*, pages 27–34. Springer, 2015.

[23] Felix Büsching, Ulf Kulau, Matthias Gietzelt, and Lars Wolf. Comparison and validation of capacitive accelerometers for health care applications. *Computer Methods and Programs in Biomedicine*, 106(2):79–88, 2012.

[24] Stephen Butterworth. On the theory of filter amplifiers. *Wireless Engineer*, 7(6):536–541, 1930.

[25] Vince D Calhoun, Jingyu Liu, and Tülay Adalı. A review of group ica for fmri data and ica for joint inference of imaging, genetic, and erp data. *Neuroimage*, 45(1):S163–S172, 2009.

[26] Emmanuel J Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.

[27] Brent Carrara and Carlisle Adams. You are the key: generating cryptographic keys from voice biometrics. In *Proceedings of the 8th Annual International Conference on Privacy Security and Trust (PST)*, pages 213–222. IEEE, 2010.

[28] Aaron Carroll and Gernot Heiser. An analysis of power consumption in a smartphone. In *Proceedigns of USENIX Annual Technical Conference*, volume 14. Boston, MA, 2010.

[29] Jagmohan Chauhan, Hassan Jameel Asghar, Anirban Mahanti, and Mohamed Ali Kaafar. Gesture-based continuous authentication for wearable devices: The smart glasses use case. In *Proceedings of International Conference on Applied Cryptography and Network Security*, pages 648–665. Springer, 2016.

[30] Shaxun Chen, Amit Pande, and Prasant Mohapatra. Sensor-assisted facial recognition: An enhanced bio-metric authentication system for smartphones. In *Proceedings of the 12th International Conference on Mobile Systems, Applications, and Services (Mobisys)*, pages 109–122. ACM, 2014.

[31] Tiffany Yu-Han Chen, Lenin Ravindranath, Shuo Deng, Paramvir Bahl, and Hari Balakrishnan. Glimpse: Continuous, real-time object recognition on mobile devices. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems (Sensys)*, pages 155–168. ACM, 2015.

[32] Sien W Chew, Rajib Rana, Patrick Lucey, Simon Lucey, and Sridha Sridharan. Sparse temporal representations for facial expression recognition. In *Proceedings of Pacific-Rim Symposium on Image and Video Technology*, pages 311–322. Springer, 2011.

[33] Anil K Chopra. *Dynamics of structures*, volume 3. Prentice Hall New Jersey, 1995.

[34] George C Clark Jr and J Bibb Cain. *Error-correction coding for digital communications*. Springer Science & Business Media, 2013.

[35] Cory T Cornelius and David F Kotz. Recognizing whether sensors are on the same body. *Pervasive and Mobile Computing*, 8(6):822–836, 2012.

[36] Shane F Cotter. Sparse representation for accurate classification of corrupted and occluded facial expressions. In *Proceedings of the 2010 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 838–841. IEEE, 2010.

[37] Bart De Moor, P De Gersem, Bart De Schutter, and W Favoreel. Daisy: A database for identification of systems. *JOURNAL A*, 38:4–5, 1997.

[38] Arnaud Delorme and Scott Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, 2004.

[39] Mohammad Derawi and Patrick Bours. Gait and activity recognition using commercial phones. *Computers & Security*, 39:137–144, 2013.

[40] Mohammad O Derawi, Claudia Nickel, Patrick Bours, and Christoph Busch. Unobtrusive user-authentication on mobile phones using biometric gait recognition. In *Proceedings of the 6th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)*, pages 306–311. IEEE, 2010.

[41] Whitfield Diffie and Martin E Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6):644–654, 1976.

[42] David L Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[43] David L Donoho and Yaakov Tsaig. Fast solution of-norm minimization problems when the solution may be sparse. *IEEE Transactions on Information Theory*, 54(11):4789–4812, 2008.

[44] Noël E Dutoit, Brian L Wardle, and Sang-Gook Kim. Design considerations for mems-scale piezoelectric mechanical vibration energy harvesters. *Integrated Ferroelectrics*, 71(1):121–160, 2005.

[45] Michael Elad and Michal Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image processing*, 15(12):3736–3745, 2006.

[46] Shuichi Enokida, Ryo Shimomoto, Tomohito Wada, and Toshiaki Ejima. A predictive model for gait recognition. In *Proceedings of the 2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, pages 1–6. IEEE, 2006.

[47] David J Ewins. *Modal testing: theory and practice*, volume 15. Research studies press Letchworth, 1984.

[48] Davrondzhon Gafurov, Kirsi Helkala, and Torkjel Søndrol. Biometric gait authentication using accelerometer sensor. *Journal of Computers*, 1(7):51–59, 2006.

[49] Davrondzhon Gafurov, Kirsi Helkala, and Torkjel Søndrol. Gait recognition using acceleration from mems. In *Proceedings of the First International Conference on Availability, Reliability and Security (ARES)*, pages 6–18. IEEE, 2006.

[50] Davrondzhon Gafurov and Einar Snekkenes. Gait recognition using wearable motion recording sensors. *EURASIP Journal on Advances in Signal Processing*, 2009:7–28, 2009.

[51] Davrondzhon Gafurov, Einar Snekkenes, and Tor Erik Buvarp. Robustness of biometric gait authentication against impersonation attack. In *Proceedings of OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*, pages 479–488. Springer, 2006.

[52] Andrew Gee and Roberto Cipolla. Determining the gaze of faces in images. *Image and Vision Computing*, 12(10):639–647, 1994.

[53] Fred Glover and Manuel Laguna. *Tabu Search*. Springer, 2013.

[54] Maria Gorlatova, John Sarik, Guy Grebla, Mina Cong, Ioannis Kymissis, and Gil Zussman. Movers and shakers: Kinetic energy harvesting for the internet of things. In *Proceedings of the ACM SIGMETRICS Performance Evaluation Review*, volume 42, pages 407–419. ACM, 2014.

[55] Le Guan, Jun Xu, Shuai Wang, Xinyu Xing, Lin Lin, Heqing Huang, Peng Liu, and Wenke Lee. From physical to cyber: Escalating protection for personalized auto insurance. In *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems (Sensys)*, pages 42–55. ACM, 2016.

[56] Kiryong Ha, Zhuo Chen, Wenlu Hu, Wolfgang Richter, Padmanabhan Pillai, and Mahadev Satyanarayanan. Towards wearable cognitive assistance. In *Proceedings of the 12th International Conference on Mobile Systems, Applications, and Services (Mobisys)*, pages 68–81. ACM, 2014.

[57] Jinguang Han and Bir Bhanu. Individual recognition using gait energy image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(2):316–322, 2006.

[58] Alan Heckert, James Dray, and San Vo. A statistical test suite for random and pseudorandom number generators for cryptographic applications. *NIST Special Publication*, 800:22, 2001. `http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-22r1a.pdf`.

[59] Ken Hinckley. Synchronous gestures for multiple persons and computers. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology*, pages 149–158. ACM, 2003.

[60] Thang Hoang and Deokjai Choi. Secure and privacy enhanced gait authentication on smart phone. *The Scientific World Journal*, 2014:254–278, 2014.

[61] Lars Erik Holmquist, Friedemann Mattern, Bernt Schiele, Petteri Alahuhta, Michael Beigl, and Hans-W Gellersen. Smart-its friends: A technique for users to easily establish connections between smart artefacts. In *Proceedings of International Conference on Ubiquitous Computing (Ubicomp)*, pages 116–122. ACM, 2001.

[62] Yiqun Hu, Ajmal S Mian, and Robyn Owens. Sparse approximated nearest points for image set classification. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 121–128. IEEE, 2011.

[63] Aapo Hyvärinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634, 1999.

[64] Aapo Hyvärinen, Juha Karhunen, and Erkki Oja. *Independent component analysis*, volume 46. John Wiley & Sons, 2004.

[65] Anil K Jain, Karthik Nandakumar, and Abhishek Nagar. Biometric template security. *EURASIP Journal on Advances in Signal Processing*, 8:113–134, 2008.

[66] Chitra Javali, Girish Revadigar, Lavy Libman, and Sanjay Jha. Seak: secure authentication and key generation protocol based on dual antennas for wireless body area networks. In *Proceedings of International Workshop on Radio Frequency Identification: Security and Privacy Issues*, pages 74–89. Springer, 2014.

[67] Chao Jia and Brian L Evans. Online calibration and synchronization of cellphone camera and gyroscope. In *Proceedings of 2013 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 731–734. IEEE, 2013.

[68] Zhenhua Jia, Musaab Alaziz, Xiang Chi, Richard E Howard, Yanyong Zhang, Pei Zhang, Wade Trappe, Anand Sivasubramaniam, and Ning An. Hb-phone: A bed-mounted geophone-based heartbeat monitoring system. In *Proceedings of the 15th International Conference on Information Processing in Sensor Networks (IPSN)*, pages 1–12. IEEE, 2016.

[69] Ari Juels and Madhu Sudan. A fuzzy vault scheme. *Designs, Codes and Cryptography*, 38(2):237–257, 2006.

[70] Ari Juels and Martin Wattenberg. A fuzzy commitment scheme. In *Proceedings of the 6th ACM conference on Computer and Communications Security (CCS)*, pages 28–36. ACM, 1999.

[71] Amit Kale, Aravind Sundaresan, AN Rajagopalan, Naresh P Cuntoor, Amit K Roy-Chowdhury, Volker Krüger, and Rama Chellappa. Identification of humans using gait. *IEEE Transactions on Image Processing*, 13(9):1163–1173, 2004.

[72] Sara Khalifa, Mahbub Hassan, and Aruna Seneviratne. Pervasive self-powered human activity recognition without the accelerometer. In *Proceedings of the International Conference on Pervasive Computing and Communications (Percom)*, pages 79–86. IEEE, 2015.

[73] Sara Khalifa, Mahbub Hassan, Aruna Seneviratne, and Sajal K. Das. Energy harvesting wearables for activity-aware services. *IEEE Internet Computing*, 19(5):8–16, 2015.

[74] Miso Kim, Mathias Hoegen, John Dugundji, and Brian L Wardle. Modeling and experimental verification of proof mass effects on vibration energy harvester performance. *Smart Materials and Structures*, 19(4):23–45, 2010.

[75] Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. Cell phone-based biometric identification. In *Proceedings of the 4th IEEE International Conference onBiometrics: Theory Applications and Systems (BTAS)*, pages 1–7. IEEE, 2010.

[76] Toby HW Lam and Raymond ST Lee. A new representation for human gait recognition: motion silhouettes image (msi). In *Proceedings of International Conference on Biometrics*, pages 612–618. Springer, 2006.

[77] Guohao Lan, Sara Khalifa, Mahbub Hassan, and Wen Hu. Estimating calorie expenditure from output voltage of piezoelectric energy harvester: an experimental feasibility study. In *Proceedings of the 10th EAI International Conference on Body Area Networks (BodyNets)*, pages 179–185. ICST, 2015.

[78] Kuang-Chih Lee, Jeffrey Ho, Ming-Hsuan Yang, and David Kriegman. Video-based face recognition using probabilistic appearance manifolds. In *Proceedings of 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages I–313. IEEE, 2003.

[79] Te-Won Lee, Michael S Lewicki, Mark Girolami, and Terrence J Sejnowski. Blind source separation of more sources than mixtures using overcomplete representations. *IEEE Signal Processing Letters*, 6(4):87–90, 1999.

[80] Jonathan Lester, Blake Hannaford, and Gaetano Borriello. Are you with me?-using accelerometers to determine if two devices are carried by the same person. In *Pervasive Computing*, pages 33–50. Springer, 2004.

[81] Jonathan Lester, Carl Hartung, Laura Pina, Ryan Libby, Gaetano Borriello, and Glen Duncan. Validated caloric expenditure estimation using a single body-worn sensor. In *Proceedings of the International Conference on Ubiquitous Computing (Ubicomp)*, pages 225–234. ACM, 2009.

[82] Mingyang Li, Hongsheng Yu, Xing Zheng, and Anastasios I Mourikis. High-fidelity sensor modeling and self-calibration in vision-aided inertial navigation. In *Proceedings of 2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 409–416. IEEE, 2014.

[83] Peng Li, Xin Yang, Hua Qiao, Kai Cao, Eryun Liu, and Jie Tian. An effective biometric cryptosystem combining fingerprints with error correction codes. *Expert Systems with Applications*, 39(7):6562–6574, 2012.

[84] He Liu, Stefan Saroiu, Alec Wolman, and Himanshu Raj. Software abstractions for trusted sensors. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services (Mobisys)*, pages 365–378. ACM, 2012.

[85] Junliang Liu, Fengqin Yu, and Ying Chen. Speech separation based on improved fast ica with kurtosis maximization of wavelet packet coefficients. In *New Perspectives in Information Systems and Technologies*, volume 6, pages 43–50. Springer, 2014.

[86] Zongyi Liu and Sudeep Sarkar. Improved gait recognition by gait dynamics normalization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(6):863–876, 2006.

[87] Benny Lo, Fani Deligianni, and G-Z Yang. Source recovery for body sensor network. In *Proceedings of the 3th International Workshop on Wearable and Implantable Body Sensor Networks (BSN)*, pages 4–12. IEEE, 2006.

[88] F Lu, HP Lee, and SP Lim. Modeling and analysis of micro piezoelectric power generators for micro-electromechanical-systems applications. *Smart Materials and Structures*, 13(1):57, 2003.

[89] Hong Lu, Jonathan Huang, Tanwistha Saha, and Lama Nachman. Unobtrusive gait verification for mobile phones. In *Proceedings of the International Symposium on Wearable Computers (ISWC)*, pages 91–98. ACM, 2014.

[90] Emanuele Maiorana. Biometric cryptosystem using function based on-line signature recognition. *Expert Systems with Applications*, 37(4):3454–3461, 2010.

[91] Jani Mäntyjärvi, Mikko Lindholm, Elena Vildjiounaite, Satu-Marja Mäkelä, and HA Ailisto. Identifying users of portable devices from gait pattern with accelerometers. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 2, pages 2–973. IEEE, 2005.

[92] R Alvarez Marino, F Hernandez Alvarez, and L Hernandez Encinas. A crypto-biometric scheme based on iris-templates with fuzzy extractors. *Information Sciences*, 195:91–102, 2012.

[93] Suhas Mathur, Robert Miller, Alexander Varshavsky, Wade Trappe, and Narayan Mandayam. Proximate: proximity-based secure pairing using ambient wireless signals. In *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services (Mobisys)*, pages 211–224. ACM, 2011.

[94] Suhas Mathur, Wade Trappe, Narayan Mandayam, Chunxuan Ye, and Alex Reznik. Radio-telepathy: extracting a secret key from an unauthenticated wireless channel. In *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking (Mobicom)*, pages 128–139. ACM, 2008.

[95] Addison Mayberry, Pan Hu, Benjamin Marlin, Christopher Salthouse, and Deepak Ganesan. ishadow: design of a wearable, real-time mobile gaze tracker. In *Proceedings of the 12th International Conference on Mobile Systems, Applications, and Services (Mobisys)*, pages 82–94. ACM, 2014.

[96] Rene Mayrhofer and Hans Gellersen. Shake well before use: Intuitive and secure pairing of mobile devices. *IEEE Transactions on Mobile Computing*, 8(6):792–806, 2009.

[97] Martin J McKeown, Terrence J Sejnowski, et al. Independent component analysis of fmri data: examining the assumptions. *Human Brain Mapping*, 6(5-6):368–372, 1998.

[98] Lee Middleton, Alex Buss, Alex Bazin, Mark S Nixon, et al. A floor sensor system for gait recognition. In *Proceedings of the 4th IEEE Workshop on Automatic Identification Advanced Technologies*, pages 171–176. IEEE, 2005.

[99] Peter Middleton, Peter Kjeldsen, and Jim Tully. Forecast: The internet of things, worldwide, 2013. *Gartner Research*, 2013.

[100] Prasant Kumar Misra, Wen Hu, Yuzhe Jin, Jie Liu, Amanda Souza de Paula, Niklas Wirstrom, and Thiemo Voigt. Energy efficient gps acquisition with sparse-gps. In *Proceedings of the 13th International Conference on Information Processing in Sensor Networks (IPSN)*, pages 155–166. IEEE Press, 2014.

[101] Paul D Mitcheson, Eric M Yeatman, G Kondala Rao, Andrew S Holmes, and Tim C Green. Energy harvesting from human and machine motion for wireless electronic devices. *Proceedings of the IEEE*, 96(9):1457–1486, 2008.

[102] Nesma Mohssen, Rana Momtaz, Heba Aly, and Moustafa Youssef. It's the human that matters: accurate user orientation estimation for mobile computing applications. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous)*, pages 70–79. ICST, 2014.

[103] Magnus Moreau. *Estimating the energy consumption of emerging random access memory technologies*. PhD thesis, Institutt for elektronikk og telekommunikasjon, 2013. `http://www.diva-portal.org/smash/get/diva2:649811/FULLTEXT01.pdf`.

[104] Muhammad Muaaz and Claudia Nickel. Influence of different walking speeds and surfaces on accelerometer-based biometric gait recognition. In *Proceedings of the 35th International Conference onTelecommunications and Signal Processing (TSP)*, pages 508–512. IEEE, 2012.

[105] M Pat Murray. Gait as a total pattern of movement: including a bibliography on gait. *American Journal of Physical Medicine & Rehabilitation*, 46(1):290–333, 1967.

[106] Claudia Nickel, Christoph Busch, Sathyanarayanan Rangarajan, and Manuel Möbius. Using hidden markov models for accelerometer-based biometric gait recognition. In *Proceedings of the 7th IEEE International Colloquium on Signal Processing and its Applications (CSPA)*, pages 58–63. IEEE, 2011.

[107] Claudia Nickel, Tobias Wirtl, and Christoph Busch. Authentication of smartphone users based on the way they walk using k-nn algorithm. In *Proceedings of the 8th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)*, pages 16–20. IEEE, 2012.

[108] Robert J Orr and Gregory D Abowd. The smart floor: a mechanism for natural user identification and tracking. In *CHI'00 extended abstracts on Human factors in computing systems*, pages 275–276. ACM, 2000.

[109] Enrique G Ortiz, Alan Wright, and Mubarak Shah. Face recognition in movie trailers via mean sequence sparse representation-based classification. In *Proceedings of the IEEE*

Conference on Computer Vision and Pattern Recognition (CVPR), pages 3531–3538. IEEE, 2013.

[110] Gita Pendharkar, Ganesh R Naik, and Hung T Nguyen. Using blind source separation on accelerometry data to analyze and distinguish the toe walking gait from normal gait in itw children. *Biomedical Signal Processing and Control*, 13:41–49, 2014.

[111] Mashfiqui Rabbi, Shahid Ali, Tanzeem Choudhury, and Ethan Berke. Passive and in-situ assessment of mental and physical well-being using mobile sensors. In *Proceedings of the 13th International Conference on Ubiquitous Computing (Ubicomp)*, pages 385–394. ACM, 2011.

[112] Swati Rallapalli, Aishwarya Ganesan, Krishna Kant Chintalapudi, Venkata N Padmanabhan, and Lili Qiu. Enabling physical analytics in retail stores using smart glasses. In *Proceedings of the 20th International Conference on Mobile Computing and Networking (Mobicom)*, pages 115–126. ACM, 2014.

[113] Christian Rathgeb and Andreas Uhl. A survey on biometric cryptosystems and cancelable biometrics. *EURASIP Journal on Information Security*, 2011(1):1, 2011.

[114] Srivaths Ravi, Anand Raghunathan, and Srimat Chakradhar. Tamper resistance mechanisms for secure embedded systems. In *Proceedings of the 17th International Conference on VLSI Design*, pages 605–611. IEEE, 2004.

[115] Yanzhi Ren, Yingying Chen, Mooi Choo Chuah, and Jie Yang. Smartphone based user verification leveraging gait recognition for mobile healthcare systems. In *Proceedings of the 10th Annual IEEE Communications Society Conference onSensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pages 149–157. IEEE, 2013.

[116] Jamil M Renno, Mohammed F Daqaq, and Daniel J Inman. On the optimal energy harvesting from a vibration source. *Journal of Sound and Vibration*, 320(1):386–405, 2009.

[117] Girish Revadigar, Chitra Javali, Wen Hu, and Sanjay Jha. Dlink: Dual link based radio frequency fingerprinting for wearable devices. In *Proceedings of 40th IEEE Conference on Local Computer Networks (LCN)*, pages 329–337. IEEE, 2015.

[118] Masoud Rostami, Ari Juels, and Farinaz Koushanfar. Heart-to-heart (h2h): authentication for implanted medical devices. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security (CCS)*, pages 1099–1112. ACM, 2013.

[119] Nirupam Roy, He Wang, and Romit Roy Choudhury. I am a smartphone and i can tell my user's walking direction. In *Proceedings of the 12th International Conference on Mobile Systems, Applications, and Services (Mobisys)*, pages 329–342. ACM, 2014.

[120] Michael Rushanan, Aviel D Rubin, Denis Foo Kune, and Colleen M Swanson. Sok: Security and privacy in implantable medical devices and body area networks. In *Proceedings of the 2014 IEEE Symposium on Security and Privacy*, pages 524–539. IEEE, 2014.

[121] Angelo M Sabatini. Quaternion-based extended kalman filter for determining orientation by inertial and magnetic sensing. *IEEE Transactions on Biomedical Engineering*, 53(7):1346–1356, 2006.

[122] M Satyanarayanan. From the editor in chief: Augmenting cognition. *IEEE Pervasive Computing*, 3(2):0004–5, 2004.

[123] Mahadev Satyanarayanan, Paramvir Bahl, Ramón Caceres, and Nigel Davies. The case for vm-based cloudlets in mobile computing. *IEEE Pervasive Computing*, 8(4):14–23, 2009.

[124] Yiran Shen, Wen Hu, Mingrui Yang, Bo Wei, Simon Lucey, and Chun Tung Chou. Face recognition on smartphones via optimised sparse representation classification. In *Proceedings of the 13th International Conference on Information Processing in Sensor Networks (IPSN)*, pages 237–248. IEEE, 2014.

[125] Lu Shi, Jiawei Yuan, Shucheng Yu, and Ming Li. Ask-ban: authenticated secret key extraction utilizing channel characteristics for body area networks. In *Proceedings of the 6th ACM conference on Security and Privacy in Wireless and Mobile Networks*, pages 155–166. ACM, 2013.

[126] Alex J Smola and Bernhard Schölkopf. A tutorial on support vector regression. *Statistics and computing*, 14(3):199–222, 2004.

[127] Colin Soutar, Danny Roberge, Alex Stoianov, Rene Gilroy, and Bhagavatula Vijaya Kumar. Biometric encryption: enrollment and verification procedures. In *Proceedings of SPIE Conference on Optical Pattern Recognition*, pages 24–35. International Society for Optics and Photonics, 1998.

[128] G Srivastava, S Crottaz-Herbette, KM Lau, GH Glover, and V Menon. Ica-based procedures for removing ballistocardiogram artifacts from eeg data acquired in the mri scanner. *Neuroimage*, 24(1):50–60, 2005.

[129] Valérie Viet Triem Tong, Hervé Sibert, Jérémy Lecoeur, and Marc Girault. Biometric fuzzy extractors made practical: a proposal based on fingercodes. In *Proceedings of International Conference on Biometrics (ICB)*, pages 604–613. Springer, 2007.

[130] Matthew Turk and Alex Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, pages 71–86, 1991.

[131] Michal Uřičář, Vojtěch Franc, and Václav Hlaváč. Detector of facial landmarks learned by the structured output svm. In *Proceedings of the 7th International Conference on Computer Vision Theory and Applications (VISAPP)*, pages 547–556. SciTePress — Science and Technology Publications, 2012.

[132] Ian Van der Linde and Tamara Watson. A combinatorial study of pose effects in unfamiliar face recognition. *Vision Research*, 50(5):522–533, 2010.

[133] Elena Vildjiounaite, Satu-Marja Mäkelä, Mikko Lindholm, Reima Riihimäki, Vesa Kyllönen, Jani Mäntyjärvi, and Heikki Ailisto. Unobtrusive multimodal biometrics for ensuring privacy and information security with personal devices. In *Pervasive Computing*, pages 187–201. Springer, 2006.

[134] Paul Viola and Michael J Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2004.

[135] Thomas Von Büren, Paul D Mitcheson, Tim C Green, Eric M Yeatman, Andrew S Holmes, and Gerhard Tröster. Optimization of inertial micropower generators for human walking motion. *IEEE Sensors Journal*, 6(1):28–38, 2006.

[136] Bo Wei, Wen Hu, Mingrui Yang, and Chun Tung Chou. Radio-based device-free activity recognition with radio frequency interference. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks (IPSN)*, pages 154–165. ACM, 2015.

[137] Bo Wei, Mingrui Yang, Yiran Shen, Rajib Rana, Chun Tung Chou, and Wen Hu. Real-time classification via sparse representation in acoustic sensor networks. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems (Sensys)*, page 21. ACM, 2013.

[138] John Wright, Allen Y Yang, Arvind Ganesh, S Shankar Sastry, and Yi Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(2):210–227, 2009.

[139] Weitao Xu, Girish Revadigar, Chengwen Luo, Neil Bergmann, and Wen Hu. Walkie-talkie: Motion-assisted automatic key generation for secure on-body device communication. In *Proceedings of 15th International Conference on Information Processing in Sensor Networks (IPSN)*, pages 1–12. IEEE, 2016.

[140] Weitao Xu, Yiran Shen, Neil Bergmann, and Wen Hu. Sensor-assisted face recognition system on smart glass via multi-view sparse representation classification. In *Proceedings of the International Conference on Information Processing in Sensor Networks (IPSN)*, pages 1–12. ACM, 2016.

[141] Xiaochao Yang, Chuang-Wen You, Hong Lu, Mu Lin, Nicholas D Lane, and Andrew T Campbell. Visage: A face interpretation engine for smartphone applications. In *Proceedings of International Conference on Mobile Computing, Applications, and Services (Mobicase)*, pages 149–168. Springer, 2013.

[142] Jaeseok Yun, Shwetak N Patel, Matthew S Reynolds, and Gregory D Abowd. Design and performance of an optimal inertial power harvester for human-powered devices. *IEEE Transactions on Mobile Computing*, 10(5):669–683, 2011.

[143] Kai Zeng, Daniel Wu, An Chan, and Prasant Mohapatra. Exploiting multiple-antenna diversity for shared secret key generation in wireless networks. In *Proceedings of the 29th International Conference on Information Communications (INFOCOM)*, pages 1837–1845. IEEE, 2010.

[144] Haichao Zhang, Nasser M Nasrabadi, Yanning Zhang, and Thomas S Huang. Joint dynamic sparse representation for multi-view face recognition. *Pattern Recognition*, 45(4):1290–1298, 2012.

[145] Lan Zhang, Xiang-Yang Li, Wenchao Huang, Kebin Liu, Shuwei Zong, Xuesi Jian, Puchun Feng, Taeho Jung, and Yunhao Liu. It starts with igaze: Visual attention driven networking with smart glasses. In *Proceedings of the 20th International Conference on Mobile Computing and Networking (Mobicom)*, pages 91–102. ACM, 2014.

[146] Xiaozheng Zhang and Yongsheng Gao. Face recognition across pose: A review. *Pattern Recognition*, 42(11):2876–2896, 2009.

[147] Yongtuo Zhang, Wen Hu, Weitao Xu, Hongkai Wen, and Chun Tung Chou. Naviglass: Indoor localisation using smart glasses. In *Proceedings of the 2016 International Conference on Embedded Wireless Systems and Networks (Sensys)*, pages 205–216. Junction Publishing, 2016.

[148] Yuting Zhang, Gang Pan, Kui Jia, Minlong Lu, Yueming Wang, and Zhaohui Wu. Accelerometer-based gait recognition by sparse representation of signature points with clusters. *IEEE Transactions on Cybernetics*, 45(9):1864–1875, 2015.

[149] Pengfei Zhou, Mo Li, and Guobin Shen. Use it free: instantly knowing your phone attitude. In *Proceedings of the 20th International Conference on Mobile Computing and Networking (Mobicom)*, pages 605–616. ACM, 2014.