UNIVERSITÀ DEGLI STUDI DI MILANO

SCUOLA DI DOTTORATO IN INFORMATICA
DIPARTIMENTO DI INFORMATICA "GIOVANNI DEGLI ANTONI"
DOTTORATO DI RICERCA IN INFORMATICA, XXX CICLO

BIOMETRIC TECHNOLOGIES FOR AMBIENT INTELLIGENCE

INF/01 INFORMATICA

TESI DI DOTTORATO DI RICERCA DI
Abhinav Anand

RELATORE
Prof. Vincenzo Piuri

CORRELATORI
Prof. Fabio Scotti
Dr. Ruggero Donida Labati
Dr. Angelo Genovese

COORDINATORE DEL DOTTORATO
Prof. Paolo Boldi

A.A. 2016/2017

Dedicated to my brother **Kumar Ashish**

ACKNOWLEDGMENT

At the completion stage of my doctoral dissertation, it is one of the delights to look over the past journey and remember all the people who are responsible for effective completion of this long but fulfilling journey. This thesis appears in its current form due to the assistance and guidance of several people. I would, therefore, like to offer my sincere thanks to all of them.

I would first like to express my sincere gratitude to my supervisor, prof. Vincenzo Piuri for providing me the insight vision, warm encouragement, and thoughtful guidance at every stage of my research. Without his constant support, I could not have achieved my research goals and hope that I can in turn pass on the research values and the dreams that he has given to me.

Then, I would like to extend my gratitude to my co-supervisors Prof. Fabio Scotti, Dr. Ruggero Donida Labati and Dr. Angelo Genovese. They always encouraged me throughout my Ph.D., provided me the research insight and freedom to perceive my research goals. I appreciate their patience and assistance which helped me to overcome critical situations and finish this dissertation. Special thanks to Ruggero, who helped me to better express my thoughts, cultivate my research ideas and supported continuously to achieve it.

Furthermore, I would like to thank my lab mates from IEBIL laboratory, Enrique Munoz, and Gianluca Sforza for making it such a fun place to work and to collaborate. I am equally thankful to all my other colleagues from the Università degli Studi di Milano, who have helped me stay happy through these years. Specifically, I would like to thank Gerson Antunes Soares, Massimo Walter Rivolta, Ebadollah Kheirati Roonizi, Ruby Karmacharya, and Tewodros Mulugeta Dagnew for fruitful discussions and suggestions in research as well as in life. I would particularly like to thank my friend Ala Arman. You are like my brother who always helped me and stayed by my side. We really spent some quality time together. I am also grateful to my Indian friends with whom I enjoyed some good time together. Specifically, I would like to thank Colin, Vijendra, Naveen, Shakti, and Anudeep.

I would also like to thank staff member of Università degli Studi di Milano, who always assisted me with the bureaucracy and official works. Specifically, I would like to thank Lorena Sala, Claudia Piana, Danio Asinari, Daniela Mutti, and Mario Resmini.

A special thank goes to my dear friend Sunny Verma for his valuable support and fruitful discussion and arguments on various aspects of life, research, and literature.

Any work would not have been possible without the love and patience of my family. I warmly thank and appreciate my parents for their love and unconditional support in all aspects of my life. I would like to thank Amioy Bhaiya for his constant support and motivation throughout my career. You are the one who inspired me to pursue research. Then, I would like to thank my elder brother Kumar Ashish for his words of wisdom, encouragement, and unconditional love and support in my personal and professional life. You are the source of whatever I have achieved and your advice was always an added fillip. Finally, I would like to extend my sincere thanks to my other family members, Badi Mummy, my sister Arpna, brother-in-law Vikas Kaushal, sister-in-law Kannu Priya, and a special thank to my dearest BJ.

Finally, my greatest regard to the almighty God for the passion, strength, perseverance, and the resources to complete this study.

ABSTRACT

Ambient Intelligence (AmI) refers to an environment capable of recognizing and responding to the presence of different individuals in a seamless, unobtrusive and often invisible way. In this environment, people are surrounded by intelligent intuitive interfaces that are embedded in all kinds of objects. The goals of AmI are to provide greater user-friendliness, more efficient services support, user-empowerment, and support for human interactions. Examples of AmI scenarios are smart cities, smart homes, smart offices, and smart hospitals.

In AmI, biometric technologies represent enabling technologies to design personalized services for individuals or groups of people. Biometrics is the science of establishing the identity of an individual or a class of people based on the physical, or behavioral attributes of the person. Common applications include: security checks, border controls, access control to physical places, and authentication to electronic devices. In AmI, biometric technologies should work in uncontrolled and less-constrained conditions with respect to traditional biometric technologies. Furthermore, in many application scenarios, it could be required to adopt covert and non-cooperative technologies. In these non-ideal conditions, the biometric samples frequently present poor quality, and state-of-the-art biometric technologies can obtain unsatisfactory performance.

There are two possible ways to improve the applicability and diffusion of biometric technologies in AmI. The first one consists in designing novel biometric technologies robust to samples acquire in noisy and non-ideal conditions. The second one consists in designing novel multimodal biometric approaches that are able to take advantage from all the sensors placed in a generic environment in order to achieve high recognition accuracy and to permit to perform continuous or periodic authentications in an unobtrusive manner.

The first goal of this thesis is to design innovative less-constrained biometric systems, which are able to improve the quality of the human-machine interaction in different AmI environments with respect to the state-of-the-art technologies. The second goal is to design novel approaches to improve the applicability and integration of heterogeneous biometric systems in AmI. In particular, the thesis considers technologies based on fingerprint, face, voice, and multimodal biometrics.

This thesis presents the following innovative research studies:

- a method for text-independent speaker identification in AmI applications;
- a method for age estimation from non-ideal samples acquired in AmI scenarios;
- a privacy-compliant cohort normalization technique to increase the accuracy of already deployed biometric systems;
- a technology-independent multimodal fusion approach to combine heterogeneous traits in AmI scenarios;

- a multimodal continuous authentication approach for AmI applications.

The designed novel biometric technologies have been tested on different biometric datasets (both public and collected in our laboratory) simulating the acquisitions performed in AmI applications. Results proved the feasibility of the studied approaches and shown that the studied methods effectively increased the accuracy, applicability, and usability of biometric technologies in AmI with respect to the state-of-the-art.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

INTRODUCTION

In this chapter, we first present an introduction to ambient intelligence. Then, we describe biometric technologies for ambient intelligence. Subsequently, we discuss the objective of the thesis, the performed research, and the obtained results. Finally, the structure of the thesis is detailed.

1.1 AMBIENT INTELLIGENCE

Ambient Intelligence (AmI) refers to an environment capable of recognizing and responding to the presence of different individuals in a seamless, unobtrusive and often invisible way. In this environment, people are surrounded by intelligent intuitive interfaces that are embedded in all kinds of objects. Examples of AmI scenarios are smart cities, smart homes, smart offices, and smart hospitals.

The AmI technologies involve different areas of computer science, including: ubiquitous computing, sensors and networks, detection and tracking, human-centered approaches, and computational intelligence.

The goals of AmI are to provide greater user-friendliness, more efficient services, user-empowerment, and support for human interactions. In particular, AmI applications should select specific characteristics or features of the environment, which can be adjusted according to the preferences of the user inside the environment. AmI technologies can provide adaptation of certain features of the considered environment for allowing users to express their needs through various ways while interacting with the environment. These needs may be expressed through a set of commands provided by the user, or automatically estimated by the AmI through voluntary or non-voluntary actions of the user. For example, gestures and the emotional state can be used for environmental adaptability to facilitate the users with the preferred services.

In the current scenario, users can interact with the environment with expressed commands related to specific services. AmI technologies should be able to estimate and store the preferences of each single user unobtrusively so that the user and AmI should be able to access these preferences anywhere across different environments. These profiles should be unique for each user and selected by recognizing the individual. Furthermore, the adaptability of the infrastructure of AmI should consider the preferences of classes of users. For example, AmI technologies can customize the environment for children, or elder people.

To recognize the users or their characteristics, AmI technologies should use user-friendly and unobtrusive methods. In this context, biometric technologies represent enabling technologies to design personalized services for individuals or classes of people.

1.2 BIOMETRICS IN AMBIENT INTELLIGENCE

Biometrics is the science of establishing the identity of an individual or a class of people based on the physical, or behavioral attributes of the person. Biometric traits are the discriminative characteristics used to perform identity recognitions. Examples of physiological biometric traits are the fingerprint, face, iris, palmprint, and hand geometry. Examples of behavioral biometric traits are the voice, gait, keystroke and mouse dynamics. Soft biometric traits are characteristics that do not have sufficient discriminability to perform recognitions, but can help in profiling a user class. Examples of soft biometric traits are the age, gender, height, and weight. Common applications include: security checks, border controls, access control to physical places, and authentication to electronic devices.

Biometric technologies may allow designing adaptable infrastructure and intelligent support for AmI, based on the user-centric approach. In particular, biometric technologies can be considered in AmI to infer information on the user identity, which permits to select the preferences estimated for an individual or class of people and provide personalized services. Soft biometric information may allow for specifying the actions desired by a class of persons and customizing the infrastructure of AmI in order to provide class-specific services without identifying the individuals.

In AmI, biometric technologies should work in uncontrolled and less-constrained conditions with respect to traditional biometric technologies. Furthermore, in many application scenarios, it could be required to adopt covert and non-cooperative technologies. In these non-ideal conditions, the biometric samples frequently present poor quality, and state-of-the-art biometric technologies can obtain unsatisfactory performance.

The design of biometric technologies for AmI scenarios need to consider the specific features of each application. Table 1.1 shows a comparison between important features of biometric technologies for AmI with respect to the security and surveillance applications. From the table, it is possible to observe that the features of AmI applications are different from the ones of the security and surveillance applications. In particular, most of the biometric technologies designed for security applications require to perform identification and verification of the identities [1, 2]. In surveillance applications, biometric technologies have also been used for monitoring and tracking users

Table 1.1: Comparison of the features of ambient intelligence with other application scenarios.

Features	Security Applications	**Ambient Intelligence**	Surveillance Applications
Operational modalities	identification, verification [1, 2]	identification, verification, preference-estimation, facilitation, human-computer interaction [5, 6]	identification, behavior-detection, monitoring, tracking, watchlist [3, 4]
Overt/covert	mostly overt [7]	overt or covert	mostly covert [7, 3]
Most used traits	face, finger, iris [1]	face, voice, gait, soft biometrics [5, 6]	face, gait [3, 4]
Touchless/ Touch-based	touch-based or touchless [8]	mostly touchless [5, 6]	touchless [8]
Enrollment	require [1, 7]	required or not required	mostly not required

[3, 4]. Differently, in AmI applications, the primary tasks of biometric technologies is to facilitate users in different environments. The main characteristic that biometric technologies should have to be used in AmI scenarios are a simple human-computer interaction and, in many cases, the capability of working in covert conditions. For this reason, the mostly used biometric traits are the face, voice, gait, or soft biometrics [5, 6]. Furthermore, it is interesting to note that, many AmI scenarios do not require a preliminary enrollment of the users.

A way of increasing the accuracy of biometric systems in AmI scenarios characterized by non-ideal acquisition conditions is to fuse the information collected by different sensors in a multibiometric system. However, to the best of our knowledge, there are no studies in the literature for designing frameworks for biometric technologies to integrate heterogeneous traits, sensors, and environmental conditions typical of complex AmI. The current state-of-the-art for biometric technologies lacks in designing AmI technologies able to automatically recognize non-cooperative individuals using the widest set of information that can be obtained from the set of sensors deployed in a generic scenario. Moreover, user-friendly biometric technologies robust to non-ideal and less-constrained acquisitions should be further investigated to improve the accuracy and the quality of the human-computer interaction in AmI scenarios.

1.3 PERFORMED RESEARCH AND NOVELTIES

The thesis presents innovative less-constrained technologies able to increase the applicability of biometric systems in AmI and improve the quality of the human-computer interaction in different AmI scenarios. The realized approaches include biometric technologies based on less-constrained and non-cooperative acquisitions to facilitate the interaction between the users and the systems in AmI.

The first goal of this thesis is to design innovative less-constrained biometric systems, which are able to improve the quality of the human-machine interaction in different AmI scenarios with respect to the state-of-the-art technologies. The novelty of the approaches, with respect to the state-of-the-art, resides in the fact that we consider less-constrained acquisition scenarios and non-cooperative users to increases the applicability of biometric technologies in AmI. Differently, most of the methods in the literature require cooperative users to perform their recognition.

Novel feature extraction and matching techniques have been studied for unimodal biometric technologies. In particular, we studied text-independent speaker recognition methods to perform closed set identification by using a limited amount of computational resources and templates of small size, thus allowing to use in embedded architecture for the AmI applications.

A novel age estimation method has been designed to extract soft biometric information from non-ideal samples collected in AmI scenarios. The realized method is robust to non-ideal images and deal with samples affected by strong rotations.

The second goal of this thesis is to design novel approaches to improve the applicability and integration of heterogeneous biometric systems in AmI. The performed studies regard original methods for novel and comprehensive biometric systems able to deal with heterogeneous traits, sensors, and environmental conditions. Novel approaches have been studied to improve the applicability and integration of heterogeneous biometric systems in AmI. The realized approaches have been designed to improve the recognition accuracy of the already deployed biometric technologies in AmI.

In particular, adaptive cohort normalization methods have been studied to improve the recognition accuracy of the previously deployed biometric systems in AmI. The designed method can be applied in existing AmI applications in a privacy-compliant manner and without requiring hardware or software modifications.

Multimodal biometric systems have been studied to combine biometric information acquired from heterogeneous traits and sensors in AmI. Score-level fusion approaches based on fixed rules and matching score densities have been studied to improve the recognition accuracy of the multimodal biometric systems in AmI. Moreover, the designed multimodal biometric system is technology independent and based on privacy compliant training approaches.

A multimodal continuous authentication systems and adaptive fusion approaches have been studied for AmI applications. The studied dynamic and intelligent fusion approach allow to integrate the heterogeneous information available in terms of different biometric traits, data from multiple sensors, and quality of biometric samples.

The designed novel biometric technologies have been tested on different biometric datasets (both public and collected in our laboratory) simulating the acquisitions performed in AmI applications. Results proved the feasibility of the studied approaches and shown that the studied methods effectively increased the accuracy, applicability, and usability of biometric technologies in AmI with respect to the state-of-the-art.

1.4 STRUCTURE OF THE THESIS

The thesis is structured as follows:

- *Chapter* 2 presents an introduction to AmI and its application scenarios. In particular, it describes the most promising technologies, the techniques for human-computer interactions, current research trends, and challenges in AmI.
- *Chapter* 3 contains an introduction to biometric recognition, biometric modalities, and the general structure of biometric systems. A survey of the main biometric traits is presented, and the methodologies used for the evaluation of biometric systems are described. The applications and research trends in biometric recognition are also presented.
- *Chapter* 4 provides a literature review on biometrics for AmI. Techniques based on unimodal biometric technologies for AmI applications are first analyzed. Then multimodal fusion and score normalization approaches are discussed. The analysis of high-level design for biometric technologies for AmI and continuous authentication is then discussed. Finally, open problems and challenges are treated.
- *Chapter* 5 describes the performed research on biometric technologies for AmI, detailing the implemented methods designed for innovative and user-friendly biometric technologies to facilitate the human-computer interaction, and studied original methods for novel and comprehensive biometric systems to manage heterogeneous traits, sensors, and environmental conditions in AmI.
- *Chapter* 6 presents the performed experiments and obtained results related to the studied methods. In particular, this chapter analyzes the performance of the implemented techniques and compare the obtained results with the state of the art methods in the literature to assess the feasibility of the considered innovative methods.
- *Chapter* 7 summarizes the work and the obtained results, the originality of the contribution, and then presents a series of possible future works.
- *Appendix* A contains the list of peer-reviewed papers in which some of the ideas and significant results presented in this thesis have been published.

2

AMBIENT INTELLIGENCE AND ITS APPLICATION SCENARIOS

Ambient intelligence (AmI) is a multidisciplinary paradigm, in which, heterogeneous sensors connected through an unobtrusive network, operate collectively to facilitate the needs and requirements of the users residing inside the environment.

In this chapter, we first present an introduction to the AmI, its characteristics, and most promising technologies used in AmI. Then, we analyze the requirement of user's need in AmI and discuss AmI application scenarios. We also present a discussion on human-computer interaction in AmI. Finally, the challenges related to the ambient intelligence technologies are treated.

2.1 INTRODUCTION TO AMBIENT INTELLIGENCE

Ambient Intelligence (AmI) refers to an environment capable of recognizing and responding to the presence of different individuals in a seamless, unobtrusive and often invisible way. The concept of AmI has laid by Philips [9], which visualized the AmI as sensitive and responsive to the presence of the user and sympathetic to their needs [9].

The European Commission's Information Society Technologies Advisory Group (ISTAG) provided a broader and more formal definition describing AmI as seamless environment equipped with advanced and intelligent computing, aware of the specific requirements of the user inside the environment, adaptive to their needs, and responsive to their requirements in smart and intelligent manner [10]. AmI envisions people in the center of the system, and the priority is to bring more focus to the user than the technology itself.

AmI technologies can allow users to not only interact with the traditional devices such as keyboard, mouse, and screen, but also through their physiological, behavioral, and emotional attributes. Moreover, with the advancement in the sensing capabilities of devices, vision technologies, and computational intelligence techniques, it is possible to design less-constrained techniques for human-computer interaction in AmI .

2.1.1 CHARACTERISTICS OF AMBIENT INTELLIGENCE

AmI is a multidisciplinary concept, which congeries the technologies from different areas and possess different characteristics. The design and development of AmI need to consider the following characteristics of AmI technologies:

- *Unobtrusiveness*: refers to the ability of the technology to assist users in an unobtrusive manner as much as possible. The sensors, devices, and other hardware of the AmI technology should be invisible (or at least not easily detectable) to the users.

- *Context awareness*: implies that the system should be aware of the context and able to understand and respond to the requirements of the user without being intrusive. The term context can be defined as information that can be used to characterize the situation of an entity such as user, environment, objects, or all together [11]. AmI should be able to process the information necessary to estimate the preferences of the users, to provide proper facilitation.

- *connectedness*: refers to the state at which various components of the AmI systems are linked together. AmI consist of many sensors, actuators, hardware and software systems that should connect and interact with each other. These physical devices should be embedded in a network to present a distributed and ubiquitous system. The interoperability of these connected devices can be considered as an important criterion to design comprehensive systems for AmI.

- *Adaptability*: refers to the ability of the environment to respond to the user according to their preferences. AmI technologies should be able to learn from the feedback of the interactions, modify the actions accordingly, and respond with the services corresponding to the preferences of the user. To make the environment adaptive, It is important to consider the application context, learning ability of the technology, and the interaction history.

- *Personalized*: concerns with the customization of the components of AmI to provide services according to the individual preferences. AmI consider users in the center of the system. Different users may have different preferences, and the objective of AmI is to provide personalized services to them. For example, in a smart home, family members may have different choices for music, temperature, news, and diets. The AmI technology should be able to adjust the components of the environment to facilitate the users with the individual choice of service.

2.1.2 ENABLING TECHNOLOGIES FOR AMBIENT INTELLIGENCE

The enabling technologies for AmI can be considered from different areas of computer science. These techniques can be broadly divided into five different classes, as shown in Fig. 2.1.

1. *Ubiquitous computing*: implies a system which is spread throughout the environment. The advancement of ubiquitous computing provides smart sensors and objects to capture the contextual information from the users. One such computing technology is the radio frequency identification (RFID). RFID provides a contactless and automatic identification of objects by using a radio frequency of the tagged objects. RFID-based technologies have used in many AmI applications such as homes, shops, industries, and offices. Recently, many ubiquitous devices that take the remote commands from the users and provide ubiquitous services have been developed. Examples of such devices are Nest thermostat, Ubi, Amazon echo, and Apple Siri.

2. *Sensor networks and actuators*: AmI needs to use different sensors (and sensor networks) to capture information related to the various requirements of the users. The sensors acquire a certain behavior of the user inside the environment and pass to the system software, which response to the action based on the processed information. Sensors are designed to collect different information from the AmI regarding physical phenomenon (e.g., temperature, lighting, sound, positions, and directions), the behavioral phenomenon (e.g., emotion, gestures, diets, and allergies), and health conditions. The network of sensors allows to communicate and exchange information inside AmI. New sensors have been developed and integrated into the existing networks. The integration of new sensors in existing networks require a common protocol for wireless sensor network (e.g., Bluetooth, GPS, ZigBee, and UltraWideBand). The actuators are technologies that permit to perform a sequence of actions according to the data collected by the sensors. The examples of the actuators are switch-on and switch-off actions for lighting and thermostats. The Recent development of smart sensing devices made available new sensing technologies that act intelligently and automatically, such as vision sensors networks in surveillance and monitoring.

3. *Detection and tracking*: technologies in AmI should be able to detect, locate, and track the user continuously. Different technologies can be used to specify and locate the position of the user (e.g., GPS, RFID, and microchip). The detection and tracking can help in designing an environment with preventive measures for the users. For example, a smart and intelligent monitoring system can detect an abnormal behavior or walking pattern of the patients. It can prevent the danger of catastrophic events like fall or heart attack by sending the information to the nearest hospital or calling an ambulance. Moreover, the social and emotional status can be tracked and used to facilitate user with prioritized services.

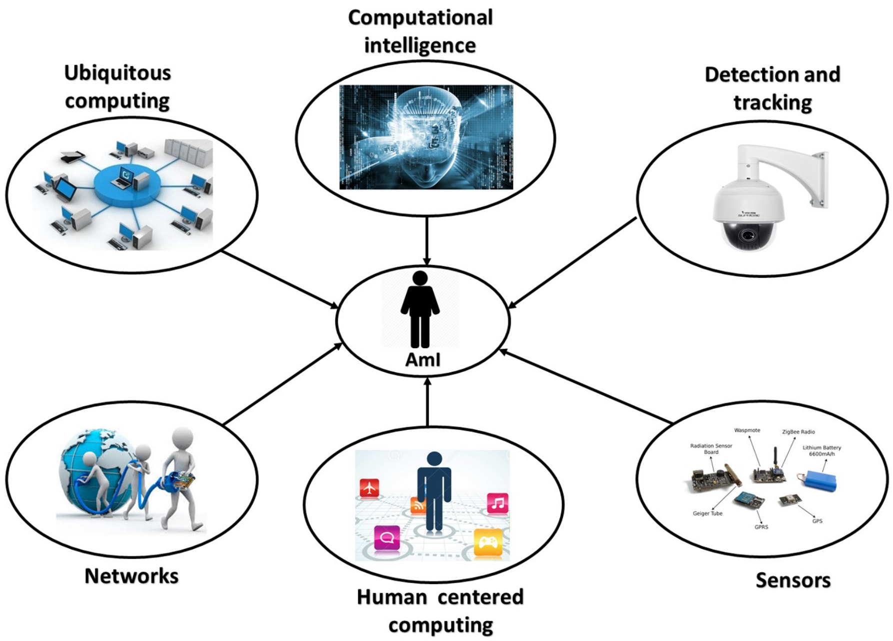


Figure 2.1: Components of ambient intelligence.

4. *Human centered computing*: AmI technologies are aimed to facilitate users according to their social, behavioral, and physiological interactions with the system. The human-centered computing uses the enabling technologies such as artificial intelligence, signal and image processing, and ubiquitous computing to design the smart environments to facilitate and support users day to day activities.

5. *Computational intelligence*: refers to the ability of the machines to learn and do the specific tasks according to the observed data. Techniques based on computational intelligence can be used to design an intelligent environment which can perform the desired tasks accurately and efficiently. These techniques allow the system to learn the specific behavior or pattern of the users, adapt and update itself during the system operation and provide low-cost solutions to design and maintain the system. Examples of computational intelligence techniques are

   - neural networks;
   - support vector machines;
   - fuzzy logics;
   - evolutionary computations, and
   - learning theory

   Some examples of the scenarios in which the computational intelligence techniques can be useful for AmI include

   - continuously learning and estimating the preference, requirements, and behavior of the user and adapting the system accordingly;

- locating, detecting, and tracking the users inside the environments;
- accurate diagnosis of the health conditions;
- smart visioning abilities to provide support for the older adults and patients, and
- intelligent and remote agents such as artificial robots and drones.

2.2 REQUIREMENT ANALYSIS IN AMBIENT INTELLIGENCE

To design an intelligent environment that can sense and respond smartly according to the needs of the user, it is important to analyze the list of needs and requirements of the user. The requirement analysis is a software engineering approach, which determines the specific needs, feature expectations, and needs of the considered user or group of users before designing the system. The list of requirements is the outcome of this process.

The requirement analysis of user's need in AmI includes the study and observations of different ways of interactions of the user with the environment. The design approaches of considered technologies for AmI must consider the possible needs of the users and the technology available to express those needs and different types of sensors (or combination of sensors) available to capture the expressions of the user knowingly and unknowingly.

2.2.1 REQUIREMENT FROM THE USER PERSPECTIVE

A list of possible requirements of the user in AmI is presented in the following.

1. *Service categorization*: refers to the different level of serves required to the user in AmI. In AmI scenarios, a secure authentication system may be required for accessing sensitive resources (e.g., email, login sessions bank transactions, etc.), a partial authentication system may be more suitable to provide access to special group members (e.g., entertainment, news, and other media related services), and other common services may be available to access without authentication. Moreover, the access modalities for some critical applications may differ for group of users.

2. *Usability of the system*: defined as the extent to which the considered technology can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in AmI.

3. *Voice interaction*: refers to the most natural and unconstrained HCI modalities. Voice-based interfaces are enabling technologies for AmI to take commands in truly unobtrusive and ubiquitous manner.

4. *Remote access*: of the services is an important requirement for the users. Systems should be able to operate from remote places as well as from the environment itself. For example, for domotic environments, the services should be managed

by the user from outside the home (e.g., lighting control, thermostat, locks, monitoring of child or old people).

5. *Preference-based services*: different users may have different preferences. To facilitate the users, the environment should understand and recognize its users, and the components of the AmI should be customized according to their preferences.

6. *Privacy requirements*: the privacy of the users inside AmI should not be compromised. Therefore, the technology which accesses user's sensitive information should be designed in a privacy-compliant manner.

7. *Special service in emergency*: AmI should be sensitive to the user information, especially regarding the health-related monitoring. In case of emergency situations, the environment should trigger an alarm, send messages to the owner, or take preventive actions to avoid dangerous, alarming, or uncommon events.

8. *Automatic Learning (Adaptability)*: systems should be able to automatically adapt themselves according to the preferences of the users. These systems should be based on intelligent and continuous learning techniques. For example, in domoic environment, adaptive methods are required for the services of temperature control, lighting control, working of electronic gadgets, working of the coffee machine, locking the main door, the announcement of things finished in kitchen, doctors appointment, schedulers list, and laundry services.

9. *Message broadcasting*: the technologies for AmI should be able to sense and track the presence of the user inside the environment and deliver the messages, mails, or other important notifications at the exact location of the users.

10. *Coupling of services*: refers to the array of sensors and actuators which anticipate and triggers the actions in reflection to the previous completed. For example, in smart homes, the coupling of services can be designed to make the environment automated and service-oriented to the users.

11. *Range of communication systems* : some services required a short range of communication (e.g., when a user wants to ask about some information of personal use or information like weather updates, traffic situation, web search, configuring a voice mail, and broadcasting a message in AmI). On the other hand, some services may require a long range of communication (e.g., asking general information, light controls, and heating controls.).

12. *Interaction and learning from user's input*: in AmI, the interaction between the human and the environments should be simplified and effective. In some applications, users prefer a system which takes input whenever a new event occurs. In other applications, it is required to design a system that does not ask any questions and automatically evaluates the behaviors (commonly termed as self-learning systems).

2.2.2 REQUIREMENTS FROM THE SYSTEM PERSPECTIVE

The requirements from the system perceptive are different from the requirements of the users. From the system's perspective, the main focus is on the interoperability, networking, throughput, reliability, and efficiency of the technologies used in the environments. According to the study performed in this work, the list of possible requirements are as follows.

1. *Complexity of the system*: the system should not be very complex; otherwise, the user may not be able to use it easily. The technologies for AmI need to be designed by considering their end users. The designed technologies should be characterized by easy procedures of installation, use, modification, and de-registration.

2. *Interoperability*: refers to the ability of the technologies to share and exchange information between different components. The new devices should be compatible with the existing networks and able to exchange the acquired information from other existing devices. Moreover, AmI systems should be able to support new technologies by adding new devices, subsystem, or software update.

3. *Switching technology*: AmI should provide the switching between different modalities to express different the needs and requirements accrediting to the situations. Switching technology should have the availability of options for selecting suitable modality for interaction (e.g., voice, gesture, touch, type).

4. *Distinction between the different groups of users*: the requirement of considering different age groups while designing AmI technologies is important in many scenarios. For example, in a typical smart home, the users are of different age groups, and the system should be designed to consider the distinct age groups while providing access to the resources.

5. *Reliability of the services*: the designed technologies should be characterized by reliable and consistent performance.

6. *Responsiveness*: actions taken by a user should be responded within a reasonable period. A delay in the response may affect the user's perception of the system. Examples of such cases are, taking actions in the emergency situations (e.g., in the event of fire, smoke, unusual behavior of elder persons, and intruder attacks. Moreover, the user should not be able to notice the transmission delay between two systems or a system and its subsystems.

7. *Accuracy and precision* : the system should be able to implement the requested actions with utmost accuracy and precision. If the designed system is not accurate, it can cause the devices to malfunction, which may cause user frustration.

8. *Portability*: to increase the portability of the system, it is necessary that the hardware units should be compact and lightweight.

9. *Customizability*: users may choose different combinations of devices to put under continuous monitoring and regulation. The design of the system should allow users to customize the devices according to their requirements.

10. *Energy efficiency*: AmI technologies keep the continuous listening devices running all the time. It is therefore required to consider devices that consume a low amount of electrical power.

11. *Multimodal interface*: in AmI, the entire user interface should be simple, self-explanatory and easy to learn and use. The system should provide multimodal interfaces which can combine different interaction modalities. The technology should be designed to operate with more than one modality (e.g., voice, gesture, touch, or image) and different modes depending on the type of application or services.

12. *Minimal interaction*: data should be displayed in a way that requires minimal interaction from the user. The system should require least user-machine interaction for a particular service.

13. *Safety of the user*: as the application deals with the remote handling of electrical appliances, ensuring the safety of the environment is of utmost importance. The process of distant regulation of a device should be safe enough for AmI, such that it does not cause electrical shocks, short-circuits, and cause fires.

14. *Language Independent*: the implementation of the system should be independent of any programming language. Each designer should be able to decide their application with the programming language of their choice.

15. *Group Identifier*: devices should be associated with different identifiers. This will enable devices to be addressed based on their groups instead of individuality. Moreover, it will also ensure the interoperability and scalability of the technology with large-scale systems.

16. *Inheritance*: it should be possible to refine dependency information and the group identifier of the device. For example, the group identifier "‘Kitchen"’ can be inherited from the another group identifier "‘ROOM"’.

17. *Prevent Undesirable Actions*: the system should be able to prevent the expected undesirable actions for a subsystems devices may be due to subsystem interactions.

18. *Authorization of Device*: only authorized devices or subsystems should be part of the system. It should be restricted for an intruder to add a subsystem to the system

19. *Discovery of new device*: a subsystem of the AmI systems should be able to publicize its existence and discover the other subsystems. For example, when a new subsystem is connected, it should automatically inform of its existence and look for other subsystems.

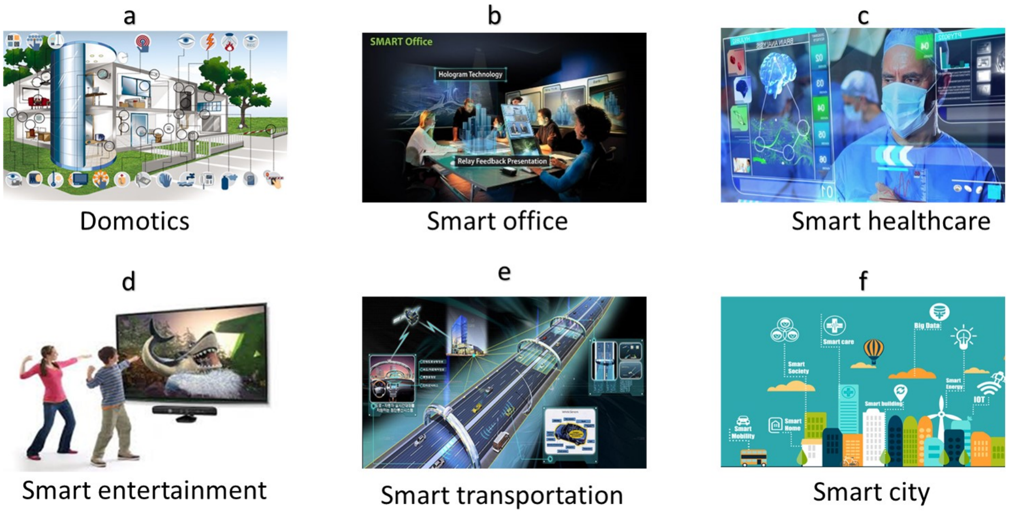


Figure 2.2: Applications of ambient intelligence

2.3 AMBIENT INTELLIGENCE APPLICATIONS

AmI technologies have been successfully deployed in many smart environments such as homes, offices, and transportation. There are many applications of AmI technologies which impact our daily lives. Fig. 2.2 shows examples of the real world applications of AmI technologies in domotic environments, offices, entertainment, healthcare, transportation, and smart cities.

2.3.1 DOMOTICS (SMART HOMES)

A growing area of research consists of the design of AmI technologies for the domotic environment. A domotic environment is a home with informatics, technologies, telematics, and robotics [12, 13]. Domotics defines the integration of technology and services inside the home to provide a better quality of life to its user. Other common names of domotics are smart homes or home automation. The focus of AmI in domotic environments is on smart devices and sensors to provide less-cooperative interactions and use of computational intelligence techniques for information retrieval, information processing, adaptive learning, and fast and reliable services to its residents [12, 14, 15, 16, 17, 18]. Examples of the smart devices developed for domotic environments are shown in Fig. 2.3. These devices include the thermostat for temperature control, continuous listening devices, smart cameras, locks, and interactive mirrors.

AmI technologies for domotics considers the design and development of smart virtual sensors that enable to estimate user's needs by evaluating her identity, physical characteristics or behaviors. Typical tasks of AmI technologies in domotic environments include locating and tracking users, listening and answering queries, turning on and off the lights, and media devices. Moreover, AmI systems are also considered for customizing the environments according to the preferences of the resident (e.g.,

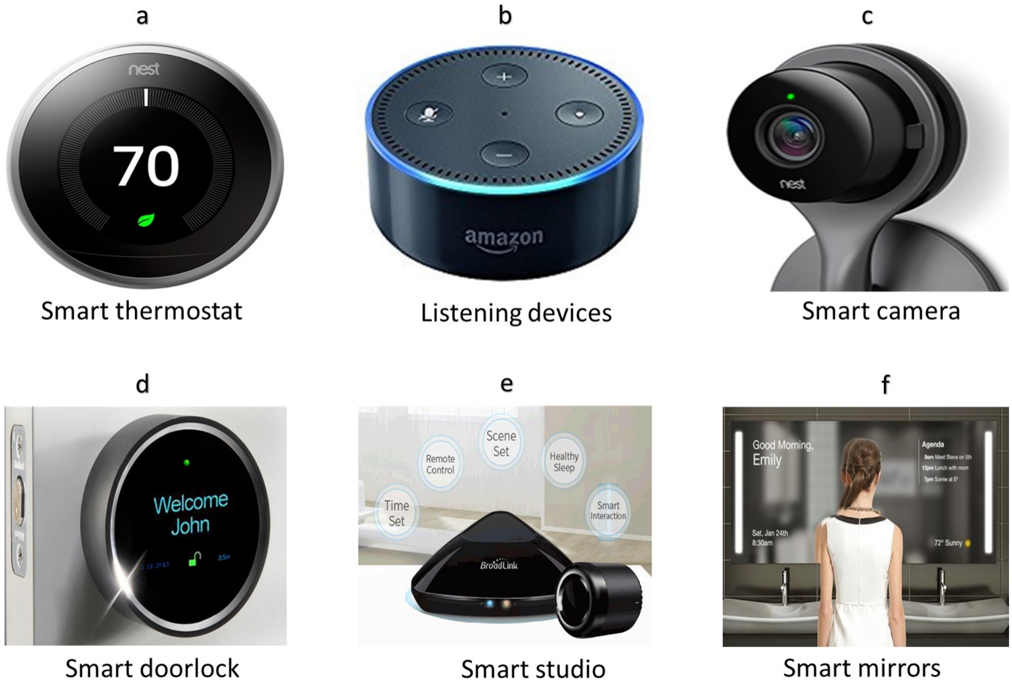


Figure 2.3: Smart devices and automation system designed for domotics.

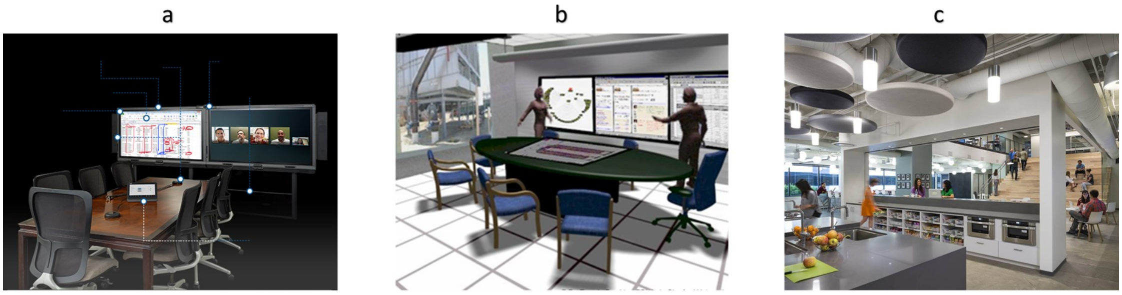
AC control, thermostat, broadcasting channels, and monitoring the health of people) [19, 12, 16, 20]. The Bill Gate's home is an ideal smart home where the user's preferences have been in the center of the design [21].

2.3.2 SMART OFFICES

Another important application of AmI technologies is in smart offices. A smart office is an environment which is adaptive to the user's preferences and capable of assisting in decision-making. The smart office environments use continuous and ubiquitous tracking to provide secure, and a reliable authentication and can manage the heterogeneous information collected from the environment available at each moment to facilitate the users [22, 23, 24]. A smart office is an intelligent environment equipped with smart sensing devices, actuators, ubiquitous computing technologies, secure mechanism to protect the personal information, and pervasive support systems [25, 26, 22]. Fig. 2.4 shows some examples of real-world projects using AmI technologies for the smart offices include Monica smart office [27], Standford's Interactive Workspace [28], and NIST smart space project [29].

2.3.3 SMART TRANSPORTATION

AmI technologies for smart transportation consist of the design of a fast, accurate, and secure system to make the public transport environment comfortable to the passen-



NIST smart space

Sandford interactive workspace

Monica smart office

Figure 2.4: Examples of smart offices.

gers, secure to the thefts, interactive and decision support systems for drivers, and smart sensors for information processing and communication with reasonably low operational cost [30]. AmI technologies can be used in transportation systems for tracking and locating vehicles, controlling automated doors and ports, efficient management of loading and unloading goods, smart traffic management, and identity recognition of the staff.

AmI technologies for the transportation systems include GPS trackers, RFID identifiers, and smart vision sensors. Advanced vision technologies have adopted in transportation systems for tracking the lanes and vehicles to alert the drivers. Interactive systems have been developed for analyzing the behavior of the drivers while they perform different actions, such as breaking, accelerating, and crossing lanes [31]. These observations are further used to assist drivers with interactive support in case of ambiguous driving patterns or emergency situations. Other examples of AmI technologies in transportation are gaze detection of drivers to analyze if she is tired or sleeping [32], continuous monitoring the state and mood of the driver to eventually provide support [33]. There has also been a Microsoft initiative for using AmI technologies in vehicle route planner [34, 35].

2.3.4 smart healthcare

AmI technologies are widely used to support and assist people at home and patients in hospitals to provide a continuous health monitoring and diagnosis. The advancement of the surveillance cameras and supporting technologies made it possible to assist mentally and physically challenged people to lead their lives independently. The primary tasks of AmI technologies in health care include activity recognition, monitoring, control of diets, anomaly detection, health diagnosis, automated support and interaction systems, and abnormal walking pattern detection [20]. Examples of application of AmI technologies in healthcare include:

- *Assistance*: AmI is also useful in assisting individuals with physical or memory-related problems to assist and remind users of their normal daily activities [36, 37]. Examples of such AmI technologies include personalized GPS and navigator to track the routes from anywhere to their home, and ambient media broadcasting systems to schedule meetings, appointments with doctors, regular visit for health monitoring, and diet-related assistance.

- *Home care:* In urban areas, most of the elderly population live solitary lives. "'Proactive Health Group"'is an initiative by Intel which aims to take care of the users living alone in their homes and improve the overall quality of life for them [38]. The focus of Intel's technology is on the social interactions to analyze the lifestyle of the individuals.

- *Comfort*: AmI systems can be adopted in hospitals and medical centers to ameliorate the user experience of patients coming to hospitals. An example of such technology is implemented in Chicago hospital children ward [39] for recognizing the patients from their RFID tag and customizing the components of the environment according to their preferences.

- *Extended-support*: AmI technologies can be used to extend the care and support to the older adults or patients in their homes. PathFinder project which connects hospitals to the smart homes [40, 41]. The smart healthcare facilities inside the houses have also advocated by AARP reports [42], which emphasized to extend the care and support for the patients in their homes. These technologies can also decrease the queues of patients in the hospitals, lower the burden of nursing staff, and improve the functionality of the medical centers [43].

2.3.5 SMART CITIES

AmI can help in designing a smart city with the aim of empowering the lives of its residents. A smart city is characterized by connected devices which work ubiquitously, smart cars with drive assistance, smart homes with automated lighting, temperature, thermostats, and security systems. Moreover, it also includes intelligent traffic management equipped with sensors and actuators, smart offices and meeting rooms, and other intelligent environments such as, shops, classrooms, and urban spaces. The focus of AmI in smart cities include the concept of consumerism, security, privacy, culture, and social aspects [44]. The vision of AmI for the smart city is still in developing phase and require innovative techniques to make the balance between the technological advancement of AmI and privacy issues related to the confidential information of people. Several countries are making efforts to bring the vision of smart city into reality. In Europe, USA and Canada, attempts have been made to build smart cities in accordance with the government initiatives [45]. Some initiatives and ongoing projects on smart cities have been presented in [46, 47, 48, 49, 50].

2.4 HUMAN-COMPUTER INTERACTION IN AMBIENT INTELLIGENCE

One of the main aims of AmI is to change the style and functioning of interaction between humans and the computers, commonly known as human-computer interaction (HCI). Traditional approaches of HCI include users interacting with computers or machines with limited capacities and are based on monitors, mouses, keyboards, and screen. These source of interactions are further constrained by fixed input-output devices and bandwidth allowed for the interactions. These approaches are not able to capture different needs, requirements, and expressions of the users while they are interacting. To simplify the HCI and improve the capabilities of the systems to understand and respond according to the human needs, the systems should be equipped with intelligent sensors and computational power.

The performed activities of the users with the software and hardware modules of the system can help in profiling users. Every user interacts with the system in a different style, and ability and a specific activity behavior can learn to extract information of the user. In particular, software or hardware interactions describe the behavior of direct or indirect interactions of the users with the system modules. These interactions provide information regarding the working pattern of the users and corresponding responses generated by the systems. These techniques include:

- *Graphical User Interface* provide a low-level interaction based information which allows extracting behavioral information of the user working on the system. This includes the pattern of mouse clicks, user-run commands, and keyboard activities.

- *Email behavior* can be used as peculiarities to characterize users. The patterns of mailing, such as length of the mails, email timing, the frequency of email checking, and management of mailbox can be extracted to profile individuals.

- *Audit logs* can provide behavior of the user while she is performing activities such as CPU usage, creation and deletion of files, frequent access to a certain directory, and performed activities on a system.

AmI technologies allow users to interact also with their physiological, behavioral, and emotional attributes. Ambient sensors and multimodal systems allow capturing a wide range of expressions of the users in a ubiquitous manner. Different modalities of interactions are needed to contextualize the possible interactions, which require multimodal interaction systems. Multimodal interface for HCI can accept inputs from the users in the form of voice commands, typing commands, valid gesture commands, facial expressions, and emotions. To capture these inputs, the infrastructure of AmI must be equipped with smart sensors such as video cameras, motion sensors, microphones, touch sensors, and other biometric technologies.

The technologies based on user-adaptive interfaces adopted for a simplified HCI in smart environments are:

- *Video technologies* aim to acquire the visual information such as the presence of the user, state of emotion, location, and contextual data. The preferences of the users are extracted from their commands or the non-voluntary actions.

- *Audio technologies* are used to take command from the users and respond to their queries. The devices based on smart voice technologies allow users to interact with the systems from anywhere and anytime. The continuous listening devices are developed which can understand the context of the conversation and respond to the user's queries.

- *Biometric technologies* can improve the interactions between the user and the systems in AmI. Biometric technologies can be considered as enabling technologies for AmI, which can offer user friendly, reliable, and secure HCI in AmI. The nature of biometric technologies allow to capture physical and behavioral attributes of the persons for active and passive interactions. In particular, biometric technologies can be adopted to estimate information from the users that permits to deliver ad-hoc services and facilities. For example, the soft biometric information such as age, gender, height, and emotional state of the users can be used either to automatically select specific classes of services or to design a personalized or priority based services to the users. Moreover, biometric technologies such as voice and face are unobtrusive and less-constrained HCI modalities which are suitable for the AmI.

2.5 CHALLENGES TO AMBIENT INTELLIGENCE

The futuristic vision of AmI which considers humans surrounded by smart sensors, intelligent networking, and ubiquitous computing devices is becoming a reality by increasing efforts to create smart environments [21, 27, 29, 51, 34, 38, 46, 50]. However, there are some challenges in AmI which need to be addressed.

- *Accurate context-aware technologies*: AmI technologies require processing a large amount of information in various complex environments. The context-aware technologies for AmI should be able to process the information necessary to estimate the preferences of the users. Moreover, these technologies should be accurate to provide decisions based on insufficient, incomplete and noisy data samples.

- *Interoperability between the devices*: The interoperability between the connected devices and new incoming devices is a challenge for AmI. To address this problem, AmI system should implement standardized interfaces for the systems and its components.

- *Conflicting requirements*: AmI technologies need to be customized to provide services in shared spaces with conflicting requirements (e.g., AC control in home or car, TV and media services at home). When more than one users are present in the environment, the AmI technologies should study the class (or group) preferences in place of personal preference of the users.

- *Privacy issues*: The privacy of the user's data is always in danger whenever the personal information regarding the user is accessed which provide the irrefutable and unique identity of the user. Several studies are performed and advocated the issues of misuse of the personal information of the data collected from the users [52, 53, 54]. The use of smart sensors and vision technologies, which continuously collecting sensible information from the user in a ubiquitous manner, raises the issues of privacy of the personal data [52]. As an example, in the development of the smart cities and collaborative environments, the vast amount of user's personal information may lead to putting user's personal life in trouble. In some cases, the privacy issues are agreeable and accepted by the citizens. The example is an initiative on the smart city in North Korea [55]. However, in most of the cases, privacy issues are seriously considered, and strict laws are imposed on the use of personal information. For example, in European projects for smart cities indicate and emphasize on the privacy regulations which restrict to access sensitive information of the users [10, 56, 57, 58].

- *Data security*: In addition to the privacy issues of AmI technologies, the risks related to the security issues are also demonstrated in the literature [59, 53]. The personal information accesses in AmI are distributed along different channels and networks. Security mechanism needs to be adopted to protect the collected data at each of the distributed networks to provide secure AmI systems. Attempts have been made to reduce the privacy and security issues in AmI [60, 61, 62]. Some studies focus on the security mechanism and encryption keys for the transaction of the personal information [62].

2.6 SUMMARY

Ambient intelligence refers to an environment capable of recognizing and responding to the presence of different individuals in a seamless, unobtrusive and often invisible way. AmI is user-centric and considers the users in the center of the system while designing the technologies. The basic foundation of AmI lies in the fact that the technologies should disappear into the environment to provide unobtrusive but ubiquitous and seamless services to its users. The underlying technologies should be aware of the context of the environment. AmI should be adaptive to the presence and preferences of its resident to provide the personalized services.

AmI is a multidisciplinary concept, and the enabling technologies for AmI systems can be considered from different areas of computer science. These enabling technologies include pervasive and ubiquitous computing, smart sensors and actuators, detection and tracking, human-centered computing, and computational intelligence.

The design approaches of considered technologies for AmI require analyzing the list of needs and requirements of the user. It is therefore important to perform requirement analysis of user's need in AmI to study and observe the possible ways of interactions of the user with the environment.

AmI technologies are successfully designed for many application scenarios and have shown its advantages in improving the quality of services and facilitation toward the users. Example of such application scenarios includes smart homes, smart offices, transportation, healthcare, entertainment, and smart cities. The underlying characteristics and enabling technologies of AmI help in designing an adaptive and personalized environment which can assist and support users in daily life activities and ameliorate the quality of life to the users.

The vision of AmI consists in the design of intelligent environments which can allow simplified and improved HCI. AmI technologies can allow users to communicate with the environment with their physiological, behavioral, and emotional attributes. The advancement in sensor technologies allows making AmI more sensible, intelligent, and less-cooperative in sensing the requirements of the users to improve the quality of the interaction between the users and the environment.

There are few challenges for AmI technologies which need to be addressed. AmI systems are consist of many heterogeneous sensors and computing devices. Advancement in the information technology brings more new devices to connect to the system. The interoperability between the connected devices and new incoming devices is a challenge for the AmI. Further, the use of smart sensors and vision technologies, which continuously collecting sensible and personal information from the user in a ubiquitous manner, raises the issues of privacy of the personal data. Moreover, the private information accesses in AmI are distributed along different channels and networks, which raises the risk of security of personal data of the users. It is important to protect the collected data at each of the distributed networks to provide secure AmI systems.

3

BIOMETRIC SYSTEMS

Biometric systems provide high accuracy in personal recognition, high level of security, user convenience, and resistance to the intrusion attacks. Biometric systems are gradually replacing other traditional security systems, such as passwords, keys, smart cards, and tokens.

In this chapter, we first describe biometric systems and their characteristics requirements. The structure of biometric systems and their operational modalities are presented. Further, different biometric traits, their characteristics, advantages, and disadvantages are discussed. Then, various techniques and protocols for the performance evaluation of the biometric systems, and their figures of merit are detailed. Finally, we outline the application of biometric systems, recent trends, and challenges.

3.1 OVERVIEW OF BIOMETRIC SYSTEMS

Biometrics is the science of establishing the identity of an individual or a class of people based on the physical, or behavioral attributes of the person. [1]. Physiological characteristics are related to the physical measurements of the human body, whereas, behavioral characteristics are related to a specific behavior of a human while performing some tasks, such as speaking, walking, or typing.

Traditional recognition systems based on passwords, tokens, keys, or smart cards impose to the users to carry or memorize a representation of his identity. These systems suffer from some drawbacks, as the password can be easily forgotten or hacked, tokens, keys, or smart cards can be lost, misplaced, or stolen. Hence, the level of security provided by the traditional recognition systems is not very reliable. The user can deny his identity by claiming that his identity token is stolen or forgotten. Also, it is possible that the user can hide his identity by presenting duplicate identity representations. Moreover, these systems are not much user convenient as users need to either always carry the identity documents or remember the passwords.

Biometric systems, on the other hand, provide more reliable, accurate, and user-friendly recognition systems. These physiological and behavioral characteristics are inherited by the individual and hence, it is difficult to forge or duplicate. Moreover, the unique nature of these characteristics makes the recognition system more robust against possible false claims of the user in order to hide her identity.

For these reasons, biometric systems are widely adopted in a number of applications such as civilian applications, security systems, access controls, border checks, criminal investigations, monitoring, forensics, and surveillance. Large-scale biometric systems are deployed in government applications such as Automated Fingerprint Identification System (AFIS) [63] and Automated Border Control (ABC) systems [64].

3.1.1 Characteristics requirement of biometric systems

Each biometric trait has its advantages and limitations, and the selection of the appropriate biometric trait depends on the considered application. No single biometric trait can satisfy the requirements imposed by all applications [65]. In order to be used for recognition systems, the biometric traits must satisfy the following characteristics [1]:

1. *Universality*: the biometric trait should be possessed by every individual accessing the application.

2. *Distinctiveness*: the underlying characteristics of the biometric traits should be sufficient to differentiate the individuals.

3. *Permanence*: the characteristics of the biometric trait should not change in time, at least for the operating period. Physiological traits are more stable and resistant to changes over a long period of time, whereas, traits associated with the behavior of the individual may change over the lifespan of the individual.

4. *Collectability*: it should be possible to measure the biometric trait quantitatively.

5. *Performance*: it refers to the accuracy, speed, and robustness of the considered biometric trait. The biometric system used in an application must be evaluated to ensure sufficient performance for the specific scenario.

6. *Acceptability*: it refers to the willingness of the users to cooperate with the system by presenting their biometric traits in order to be recognized.

7. *Circumvention*: it refers to the robustness of the biometric technologies against the fraudulent techniques.

3.2 Structure and operational modalities of biometric systems

Biometric recognition process can be divided into four steps. Details of these steps are presented below:

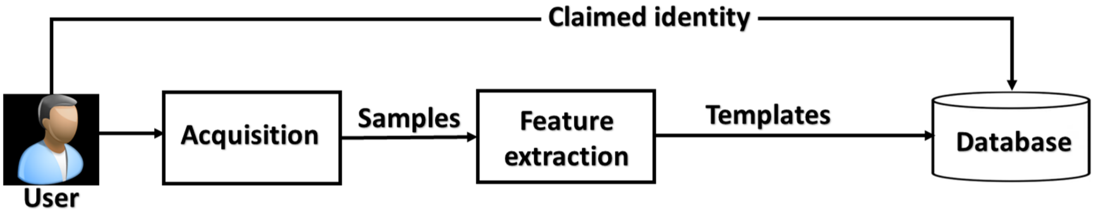


Figure 3.1: Schema of the enrollment step of biometric systems.

3.2.1 GENERAL STRUCTURE

The four basic modules of the typical biometric systems are:

1. *Acquisition*: according to the used biometric trait, a sensor is used to acquire the biometric characteristics from the user presented to the system and convert it to a digital form to be transferred to the next module. The captured biometric trait can be images, signal, frame sequence, and is typically referred as "‘sample"’.
2. *Feature Extraction*: this process involves extracting distinctive information from the captured raw samples and transforming it into a compact and effective representation (called "‘template"’) which is more compact and stable than the original sample. Templates can be composed of strings, pixels values, coordinate of particular points in the images, or signals.
3. *Enrollment*: the computed template of the user is stored in the database with associated identity. This step is known as enrollment. Fig. 3.1 shows the enrollment process of the biometric recognition systems.
4. *Matching*: the template of the user is compared with the templates stored in the database. The outcome of the template matching is a value called matching score. The template matching can be based on different metrics and can return a similarity or dissimilarity index.
5. *Decision*: the computed match score is used to provide the final decision of the biometric system, which is a boolean value representing the classes "‘accept"’ and "‘reject"’.

3.2.2 OPERATIONAL MODALITIES

The biometric systems can operate in two modalities: verification and identification.

- *Verification*: in the verification mode, the identity of the user is declared in advance through the user’s ID. The user’s biometric trait is acquired and converted into a template by using a feature extraction method. Then, the fresh computed template is compared with the reference template of the user stored in

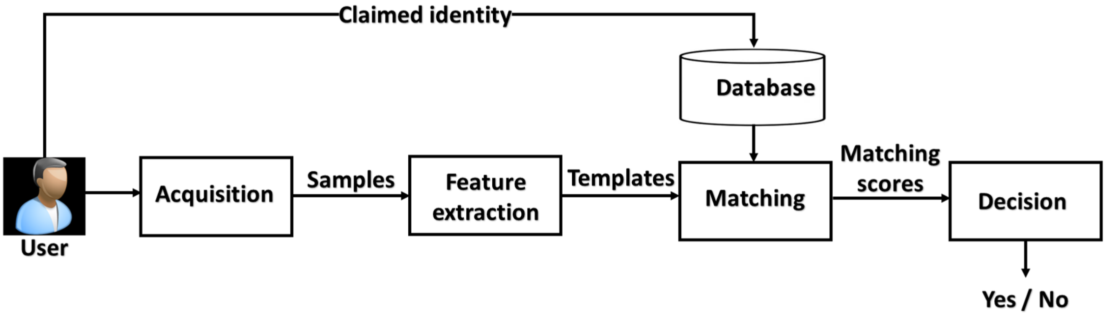


Figure 3.2: Schema of the verification mode of biometric systems.

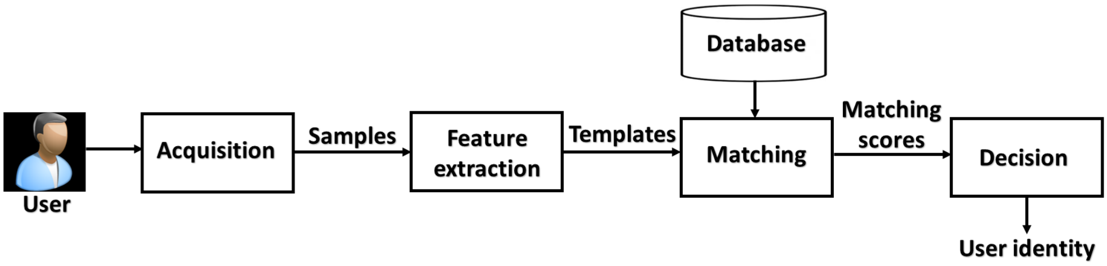


Figure 3.3: Schema of the identification mode of biometric systems.

the database, by using a suitable matching algorithm. In this case, a 1:1 matching is performed, and the access is granted to the user based on a selected threshold value. Fig. 3.2 shows the verification modality of the biometric recognition systems.

- *Identification*: in the identification mode, the biometric system has to establish the identity of the person by comparing the fresh template with all the templates stored in the database. In this case, 1:N matching are performed to search the identity of the person associated to the most similar template, based on the obtained matching scores. Fig. 3.3 shows the identification modality of the biometric recognition systems.

The term "‘recognition"’ is used when there is no need to make a distinction between verification and identification modes.

3.3 CHARACTERISTICS OF BIOMETRIC TRAITS

Biometric traits can be divided in to physiological and behavioral characteristics. Physiological traits are related to the physical measurements of the human body. Examples of physiological traits include the fingerprint [66], face [65], iris [67], palmprint [68], hand geometry [69], hand vein patterns [70], ear [71], DNA [72], and the ECG [73]. Fig. 3.4 shows the examples of the most used physiological biometric traits.

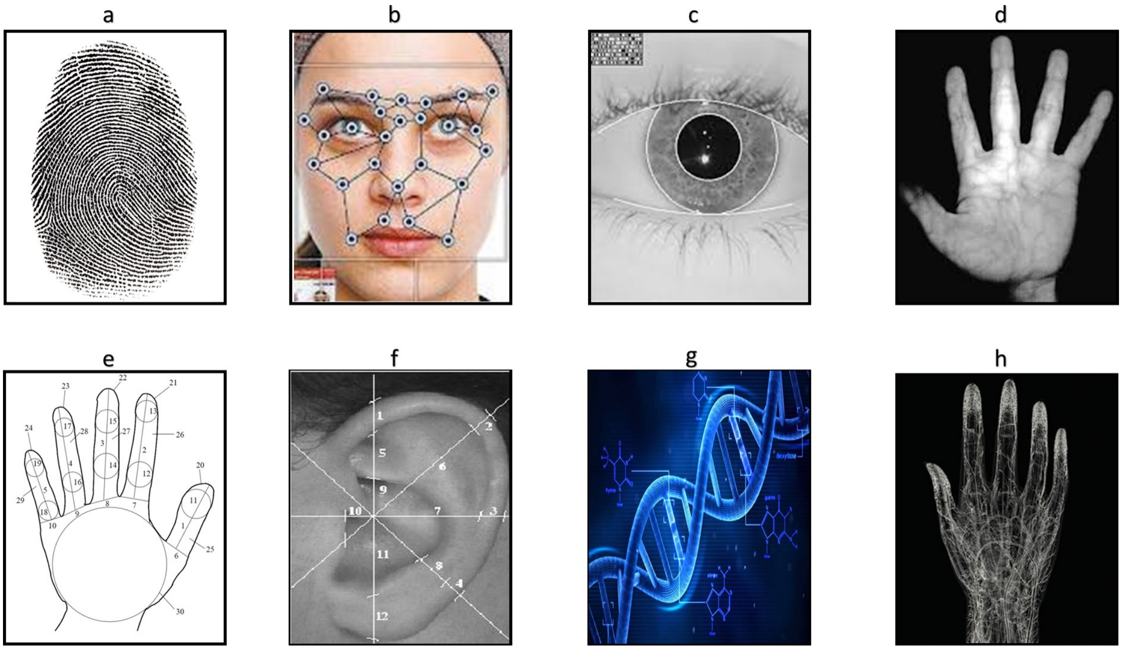


Figure 3.4: Examples of physiological biometric traits: (a) fingerprint, (b) face, (c) iris, (d) palmprint, (e) hand geometry, (f) ear shape, (g) DNA, and (h) vein patterns

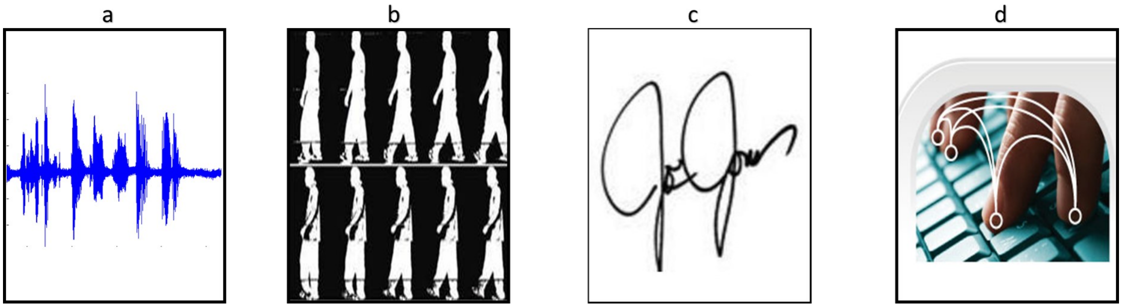


Figure 3.5: Examples of behavioral biometric traits: (a), voice, (b) gait, (c) signature, and (d) keystroke dynamics

On the other hand, behavioral traits are related to a specific behavior of a human. Examples of behavioral traits are voice [74], gait [75], signature [76], and keystroke dynamics [77]. Fig. 3.5 shows the example of most used behavioral biometric traits.

Both physiological and behavioral biometric traits have been studied for various applications. Innovative biometric traits are constantly investigated to improve the performance, speed, cost, or reduce the privacy risks of the biometric systems.

Other than the aforementioned physiological and behavioral biometric traits, soft biometric traits are studied in the literature in order to improve the accuracy of the biometric recognition systems or to perform less-cooperative recognitions. Soft biometric traits are defined as "’the characteristics that provide some information about the individual but lack the distinctiveness and permanence to sufficiently differentiate any two individuals"’ [78, 79]. This kind of traits cannot be used alone to recognize a person

Table 3.1: Characteristics of different biometric traits [1]

Traits	Univ.	Uniq.	Perm.	Coll.	Perf.	Acc.	Circ.
Face	H	L	M	H	L	H	L
Fingerprint	M	H	H	M	H	M	H
Hand geometry	M	M	M	H	M	M	M
Keystrokes	L	L	L	M	L	M	M
hand vein	M	M	M	M	M	M	H
Iris	H	H	H	M	H	L	H
Retinal scan	H	H	M	L	H	L	H
Signature	L	L	L	H	L	H	L
Voice	M	L	L	M	L	H	L
Face thermograms	H	H	L	H	M	H	H
Odor	H	H	H	L	L	M	L
DNA	H	H	H	L	H	L	L
Gait	M	L	L	H	L	H	M
Ear	M	M	H	M	M	H	M

Notes: Univ. = Universality; Uniq. = Uniqueness; Perm. = Permanence; Coll. = Collectability; Perf. = Performance; Acc. = Accuracy; Circ. = Circumvention; H = High; M = Medium; L = Low.

because these traits are not distinctive and reliable enough, and can be easily spoofed. However, the soft biometric information can be incorporated in the primary biometric systems in order to improve the performance of the recognition systems.

Soft biometric traits can be continuous or discrete [78]. Examples of continuous soft biometric traits includes height [80], weight [81], and other measurements of the body parts [78]. Examples of discrete soft biometric traits include age [82], gender [83], eye color [84], ethnicity [84], and skin and clothing color [78].

The biometric system based on a single biometric characteristic could not ensure a sufficient level of accuracy. To obtain a better performance in terms of accuracy, reliability, and security, it could be advantageous to combine multiple complementary biometric traits. The systems that use more than one biometric trait are known as multibiometric systems.

A comparison of the different characteristics of biometric traits is shown in the Table 3.1. A detailed overview of the physiological traits, behavioral traits, soft biometric traits, and multibiometric systems are presented below.

- **Physiological traits**

  The most used biometric trait is the fingerprint [66, 85] which is characterized by good performance in terms of accuracy and speed. Fingerprint recognition refers the pattern based matching of unique ridges, minutia points, and pores present in a finger. It is the most diffused biometric technology, is robust to aging, and offers high recognition accuracy [85]. However, fingerprint recognition technologies usually suffer from low user acceptability due to the high level of user cooperation required for capturing data. Typically biometric systems based on fingerprint are used for physical access control [66, 85], automated border control and e-gates [86], and forensic applications [66].

Other widely diffused biometric systems are based on the face characteristics [65]. Face recognition is the most natural method for human recognition. Face recognition systems can use either a captured still image or frame sequences acquired with low user cooperation. Face biometrics typically obtain less accuracy than the fingerprint biometrics. However, one of the important quality of face recognition is its high level of user acceptance. Face-based biometric systems are extensively used for security, forensics, and civilian applications [87, 65, 88]. Moreover, facial images can be used for analyzing expressions and emotions. For this reason, they are widely used in HCI [89, 65].

Biometric systems based on the iris are considered as the fastest and most accurate biometric systems [90, 91]. This is due to the fact that the iris shows very high discriminative characteristics. Iris recognition systems present high recognition accuracy and require low computational time [91, 92]. However, the acceptability of iris biometrics is very low due to the high level of required user cooperation. Moreover, the iris-based technologies are costly and can be perceived as dangerous for the health due to the used infrared illuminators. Iris-based biometric systems are typically used in scenarios that require high security and accuracy, such as automated border control [67]. Moreover, recent studies on iris recognition with mobile phones [93] could allow the diffusion of iris recognition systems for authentication in the smart phones.

Other biometric systems are based on hand characteristics, such as hand geometry [69], palmprint [68], and vein pattern [70]. Biometric systems based on hand geometry are characterized by low accuracy but have high user acceptance and low hardware costs. On the other hand, systems based on the palmprint and vein patterns usually present higher accuracy than that of the systems based on the hand geometry and also possess high user acceptance. Hand-based biometric systems are typically deployed in the application scenarios which does not require high security such as time and attendance systems [68, 69].

Some promising systems based on other physical traits use DNA [72] and ear shape [71]. Biometric systems based on DNA are the most accurate recognition systems. However, the recognition process is expensive and require long evaluation time. Biometric systems based on ear shape can obtain sufficient recognition accuracy for a variety of applications scenarios and have an advantage that it is possible to perform acquisitions in a contactless manner, even at long distances.

- **Behavioral traits**

  Biometric systems based on behavioral traits consider the characteristics of voice [94, 74], gait [95, 96], signature [76], and keystroke [77]. These systems are characterized by high user acceptability. However, they present a lower recognition accuracy compared to the physiological traits.

Voice recognition provides a true unobtrusive recognition technology. Voice recognition can be divided in two categories, namely: speaker recognition (which aims to recognize the user based on her voice) and speech recognition (which aims to recognize what is said). Speaker recognition systems can be classified into text-dependent and text-independent [94, 74]. The first class requires that the user enunciate a specific set of words, while the second class does not impose this limitation. Text-independent speaker recognition systems are usually less accurate than text-dependent systems. Moreover, voice recognition technologies are not very distinctive and usually affected from background noise and channel variations. Voice-based systems are typically diffused in many commercial applications, such as Apple Siri, Ubi and Amazon Echo, etc.

Systems based on the gait characteristics also suffer from low distinctiveness but possess sufficient discriminatory information to perform recognitions in certain applications. Gait-based systems are usually diffused in application scenarios which require low to medium security [75]. Other systems based on signature and keystroke are well accepted by the users but suffer from the fact that their characteristics changes over a period of time and are influenced by physical and emotional conditions of the users.

- **Soft biometric traits**

  Soft biometric traits, such as age, gender, height, skin and clothing color are also considered for recognition systems [78, 79, 80]. Although they do not allow to perform the univocal recognition of individuals, they can be used in combination with other physiological or behavioral traits to increase the recognition accuracy of the biometric systems. The advantage of using soft biometric information lies in the fact that they can be acquired without the cooperation of the user.

  Soft biometric information is typically used in application scenarios composed of a limited number of users, or to screen the large datasets in order to avoid unnecessary biometric comparisons. Moreover, soft biometric traits allow performing unobtrusive and continuous verifications [97].

3.4 MULTIBIOMETRIC SYSTEMS

In order to increase the recognition accuracy and reduce the effect of inter-class variation in biometric recognition systems, multibiometric systems are used. These systems combine the information from different biometric traits, multiple samples of the same trait, or different recognition algorithms [98, 99, 100, 101]. Multimodal systems can mitigate important problems of monomodal approaches, such as non-universality, high intra-class, and low inter-class variability. Multibiometric systems incorporate different kinds of modalities (hard or soft) with different characteristics. There are different methods for integrating biometric traits, such as:

- *Sensor-level fusion* [102]: where the raw biometric signals or images acquired from different sensors are fused. One example of this fusion level is combining face images acquired through different sensors: visible light camera and the infrared thermal camera.

- *Feature-level fusion* [102]: where different feature vectors are combined. The feature vectors can be either extracted from the same biometric trait through different algorithms or obtained using different biometric systems.

- *Score-level fusion* [103, 104]: where the matching scores computed from different biometric systems are combined. The matching scores of different biometric technologies can be combined using different fusion rules [104].

- *Rank-level fusion* [105]: where the matching scores computed from different biometric technologies are converted into a rank matrix by arranging in the decreasing order of confidence. This fusion technique can only be applied for identification systems. The rank matrices computed through different biometric matchers are fused using the rank-level fusion.

- *Decision level* [105]: where the fusion method combines the outputs of different biometric systems (matchers of classifiers) at decision level. This strategy can be used in both verification and identification modes. Most of the available commercial software provides the final decision as an output of the biometric system. Hence the decision level fusion can be appropriate in these scenarios.

Multibiometric systems can be categorized into two types on the basis of integration of the biometric traits: (1) fusion before the matching, which includes the sensor level and the feature level fusion methods and (2); fusion after the matching, which includes score level, rank level, and decision level fusion methods. The multibiometric systems obtain higher accuracy and are typically diffused in high-security application scenarios.

3.5 HUMAN-COMPUTER INTERACTION IN BIOMETRICS

The study presented in [106] argues that the biometric technologies for HCI can be divided into two broad classes: direct HCI-based biometric technologies and indirect HCI-based biometric technologies [107]. The first class of studies considers the direct interaction of humans with the systems using available sensors (mouse, keyboard, screen, muscle actions, behavioral and physiological biometrics). The other class of studies considers the indirect interaction of humans with the systems by observing their low-level behavioral actions. Example of this biometrics includes audit logs, system calls, graphical user interface, traffic of the network, registry access, and system calls [108]. This indirect biometric information can be acquired when the user performs some responsive actions unwittingly while interacting with the system.

The quality of HCI greatly influences the success of biometric technologies. Inconvenient interactions with the system can cause a performance degradation result in low social acceptance. In the context of AmI, the quality of HCI is of paramount importance because it can determine the success or the failure of technology. In the literature, there are different studies on techniques for evaluating and improving the quality of HCI in biometric systems, which can be of great help in designing biometric technologies for the domotic environment. These methods aim to analyze the interactions between human, sensors, and computers. Recent studies seek to improve the quality of HCI in biometrics without affecting the system performance [109, 110].

An important technique to evaluate and improve the quality of HCI is HBSI [111, 112]. This technique can be successfully used to analyze biometric recognition systems [112]. The main idea of this technique is that the numbers or Failure to Acquire (FTA) and Failure to Enroll (FTE) can be used to estimate usability metrics of the system. This method classifies erroneous presentations of the biometric trait to the sensor into three groups: Defective Interaction (DI), False Interaction (FI), and Concealed Interaction (CI). Correct presentations are classified into Failure to Detect (FTD) and Failure to Extract (FTX). Statistical analysis of these figures of merit allows estimating the HCI quality. With proper investigation, it is then possible to correct and improve the device and algorithms to avoid these errors [113].

There are also standards and guidelines aimed to improve HCI and usability of biometric systems. ISO 13407 guidelines [114] define the aspects on human-centered design for HCI system and suggest methods to incorporate usability into the design of biometric technologies. ISO [115] defines a way to test the performance of the biometric systems. ISO 9241-11 and ISO/IEC 14598-1 provide guidelines to be adopted for overall quality and usability of biometric technologies [116].

In the literature, there are also examples of studies on usability and user acceptance performed to different kinds of biometric technologies based on analysis of the system accuracy and questionnaires filled by the users [8, 117].

3.6 PERFORMANCE EVALUATION AND FIGURES OF MERIT OF BIOMETRIC SYSTEMS

Biometric technologies are widely used in many heterogeneous types of applications. Different applications require to evaluate and select the most suitable recognition technology suited for the considered applicative context. The performance of biometric systems can be analyzed by using different evaluation aspects and commonly adopted figures of merit.

The performance of the biometric systems can be analyzed by considering the following nine distinct aspects:

- *Accuracy* can be defined as the reliability of giving a correct recognition decision. In many high-security applications, like forensics and ABC e-gates [118, 119], this aspect is on high priority. Accuracy measures are described in more detail in Section 3.6.2.

- *Speed* measures the time taken in decision-making. Biometric identification systems require the use of faster algorithms with respect to verification systems. The speed of biometric systems should be determined by evaluating all the hardware and software modules composing the system. It is an important aspect to evaluate for online biometric systems.

- *Usability* defines by International Standard Organization (ISO) as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" [25]. The usability of the system can be evaluated by analyzing the acquisition time and a number of incorrectly captured samples. Poor usability decreases the accuracy of the biometric system and can cause user frustration, thus encouraging people not to buy and use the system.

- *Cost* measures the cost of the design, hardware, and software of the recognition algorithms. Cost of the sensors, devices, software, and computational architectures greatly influence the suitability of a biometric system. In general, low costs allows for a greater diffusion of a technology but frequently affects the system performance.

- *Security* measures the robustness of the system against possible attacks. In the context of biometrics, the term security ensures: authentication, data integrity, confidentiality, and non-repudiation [120]. It is continuous challenge to make the system more secure [121]. To make the biometric system more robust to attacks, it is important to investigate the robustness of the system against fake biometric traits, malicious software, and tampering with the features representation and stored templates.

- *Privacy* measures the capability of the system to prevent the identity theft and possible misuses of the biometric data. Security and privacy are two different concepts because the privacy protection is more restrictive than the security protection. Differently from security, privacy requires also the data protection [120, 122, 123]. The expectation from a reliable biometric recognition system is to provide data protection and an irrefutable proof of the identity of the user. However, it is an ongoing effort to make the system, more secure and less prone to misuse of private information. In biometric systems, a huge amount of personal information is collected by the sensors. The collected personal information can be either processed through the system or collected in the cloud. It is very important to have a safeguard on this information. Methods for template protection and matching in the encrypted domain [120, 122, 123, 124] should therefore be adopted to protect biometric data and sensible information of the users.

- *Scalability* of the system describes its flexibility in a high work-load situation (e.g., in the presence of big amounts of data). The scalability also defines how capable the system is to give the performance in situations when the work-load is increased. For example, a video surveillance system needs to process a large amount of data. A continuous listening device (e.g., Ubi, Amazon Echo, and Motox.) needs to capture and process a huge amount of data. In these situations, the biometric system should be highly scalable to process high work pressures.

To evaluate the scalability of the biometric systems, it is important to consider the network architecture, hardware architecture (e.g., sensors and CPU), and the performance of the recognition algorithms in terms of accuracy and speed.

- *Interoperability* represents the level of compatibility between different biometric technologies and systems. Biometric systems should be able to exchange the information between different components and modules. The biometric traits, sensors, evaluated systems, and the computational devices should be able to exchange data one to each other. In many applications, the interoperability is guaranteed by adopting standards.

- *Social acceptance* reflects the user's feeling and opinion regarding the biometric technology. It may vary within different users due to cultural differences, religious aspects, and usability aspects of the technology. The social acceptance is inter-related with the usability aspects. Low social acceptance can determine the failure of biometric systems.

3.6.1 EVALUATION PROTOCOLS AND STANDARDS

Biometric systems need to be evaluated in order to determine the feasibility of the system, select the appropriate biometric technology, and measure the performance of the considered biometric system in an applicative context. The evaluation protocol for biometric systems includes three strategies. The first is the technology evaluation, which determines the suitability of the biometric technology for the application. The second is the scenario evaluation, which evaluates the performance of the considered biometric technology for a particular application domain. The last is the operational evaluation, which is used to test the biometric system in a real environment. The applicability of the biometric technologies in the considered application scenario may be determined by analyzing the strengths and weaknesses obtained during these evaluations.

- *Technology evaluations* aim to analyze the performance of the biometric system, to measure its accuracy. This kind of evaluation is usually performed to determine the suitable choice of the biometric technology and methodology. Typically, technology evaluations are performed on public datasets, in order to compare the performances obtained by a biometric technique with the ones obtained by the other methods in the literature. A common approach to perform a technology evaluation consist of open competitions conducted by independent groups. Examples of technology evaluation competitions are Fingerprint Verification Competition [125], Face Recognition Technology evaluations [126] and the National Institute of Standards and Technology speaker recognition evaluations [127].

- *Scenario evaluations* measure the overall system performance of the biometric technology in a particular application domain. One of the main purposes of this kind of evaluation is to determine the applicability of biometric technology to obtain the performance requirements for the considered application. An example of scenario evaluation is the comparison of the performance achieved by using different biometric traits for access control systems in a specific application such as e-gates or laboratory access. Scenario evaluations can be performed by using different biometric traits acquired from the same individuals. To compensate the variations in biometric readings taken over a considered period, multiple samples of the same person can be used. Usually, the scenario evaluation is performed on databases composed of a large number of samples, to evaluate the performance with statistical significance. Since a scenario evaluation tests the complete system for a particular application condition, it is not always completely repeatable. A typical example of scenario evaluation is the UK Biometric Product Testing [128].

- *Operational evaluations* are used to test the selected biometric system in the real world scenarios. Operational evaluations are performed on all the possible target users of the system. This kind of evaluations is difficult or impossible to repeat. The objective of this kind of evaluations is to determine the impact of the applied biometric systems on the workflow of the considered application scenario and to analyze the advantages and disadvantages of the considered biometric applications.

3.6.2 FIGURES OF MERIT

The evaluation of the figures of merit of biometric systems is a stepwise procedure. The first step is to compute the matching scores by comparing the fresh acquired templates with the temples stored in the database. The next step is to compute the standard errors of the biometric systems using the computed matching scores. The obtained results allow computing the common figures of merit of the system.

Let us suppose, S_{ij} denotes the acquired j^{th} sample of the i^{th} individual, T_{ij} denotes the corresponding biometric template, n_i denotes total number of templates for an individual i, N is the total number of identities enrolled in the database, and $M(.)$ denotes the biometric matching function.

- **Matching operations**

  The matching scores are the results of comparing the fresh acquired template of an individual with the enrolled templates in the database. There are two types of matching functions described in the literature, known as symmetrical matching and asymmetrical matching [129, 130]. The matching function M(.) is called symmetrical if

  $$M(T_{ij}, T_{kl}) = M(T_{kl}, T_{ij}), (ij \neq kl) \tag{3.1}$$

  whereas the matching function M(.) is called asymmetrical if

  $$M(T_{ij}, T_{kl}) \neq M(T_{kl}, T_{ij}), (ij \neq kl) \tag{3.2}$$

The biometric templates T_{ij} created in the previous step is compared with the templates $T_{ik}(j < k < n_i)$. The resulting matching scores are stored in the "'genuine"' matching scores matrix represented by gms_{ijk}. The term "'genuine"' represent the case when the matching scores are obtained by comparing the templates of the same individual. The genuine matching scores corresponding to each user are computed and saved in the gms. The resulted score matrix is a square matrix, and in the case of symmetric matching functions, only the upper triangular matrix is computed. In the case of asymmetrical matching functions, the templates T_{ij} is compared with the templates $T_{ik}(j \leqslant n_i, K \neq i)$. The resulting matching score matrix is a square matrix but not symmetrical.

After computing the genuine matching scores, the next step is to compute the impostor matching scores. The term impostor represents the case when the matching score is obtained by comparing templates of different individuals. In case of symmetrical matching functions, each template $T_{i1}(i = 1, 2, ...N)$ is compared with the templates $T_{k1}(1 < k \leqslant N, k > i)$. The obtained matching scores are stored in the impostor matching score matrix ims_{ik}. The resulting score matrix is a square matrix, and in the case of symmetric matching functions, only upper triangular matrix is computed. In the case of asymmetrical matching functions, each template $T_{i1}(i = 1, 2, ...N)$ is compared with the templates $T_{k1}(1 < k \leqslant N, k \neq i)$. The resulting matching score matrix is a square matrix but not symmetrical. An example of score distribution curve for genuine and impostor matching scores is shown in Fig. 3.6.

The computation of the genuine matching scores (gms) and the impostor matching scores (ims) is often affected by the possible errors introduced by the enrollment step. These errors may result in failure to compute and store the biometric templates T_{ij} in the database. The error introduced in this case is denoted by the index REJ_{ENROLL}. The possible errors may be due to the following reasons: if an individual cannot interact correctly with the biometric sensor, or if the acquired sample of the individual is of poor quality or the algorithm crashes during the biometric processing. In particular, the proportion of individuals for which it is not possible to enroll the temples are refereed as Failure to Enroll Rate (FTER). Moreover, the proportion of enrollment attempts caused by the failure of the system to create samples with sufficient quality is referred as Failure to Accept rate (FTAR) [113, 117]. These errors cause the missing information in the score matrices gms and ims, and are accounted in REJ_{NGRA} and REJ_{NIRA} respectively.

- **Error rates**

  The computed genuine and impostor matching scores can be used to calculate the figures of merit for evaluating the accuracy of biometric systems [113]. These figures of merit are computed for the biometric systems allowing multiple attempts for acquiring multiple templates and considering the matching scores as a result of the comparison between the fresh acquired template with an enrolled template. In general, biometric systems can make two types of errors, namely, false

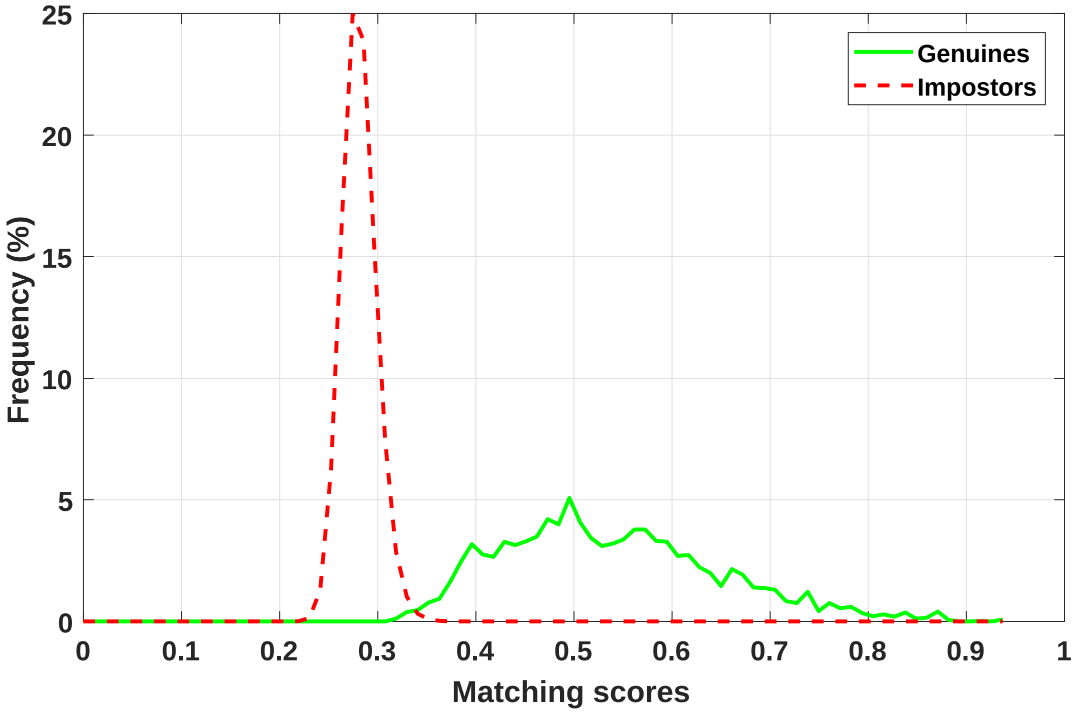


Figure 3.6: Examples of genuine and impostor scores distributions.

match rate (FMR(t)) and false non-match rate (FNMR(t)), which are functions of decision threshold t.

FMR represents the cases when the biometric samples from different individuals are incorrectly recognized as a match (false positive), which may occur due to large inter-user similarity. It refers to the proportion of matches between the templates of different individuals that are incorrectly recognized as a match. On the other hand, FNMR represents the case when the templates of the same individual are not recognized as matches (false negative), which may occur due to large intra-user variations. It refers to the proportion of matches between the templates of the same individual that are incorrectly not recognized as a match.

Both the errors FMR and FNMR are computed using the genuine and impostor matching scores by varying the threshold (typically ranging between 0 and 1, where 0 and 1 represent the minimum and maximum match score values of the genuine and impostor matching scores). The error metrics FMR and FNMR are computed as:

$$\mathrm{FMR}(t) = \frac{\mathrm{card}\{ims_{ik} | ims_{ik} \geqslant t\}}{\mathrm{NIRA}} \tag{3.3}$$

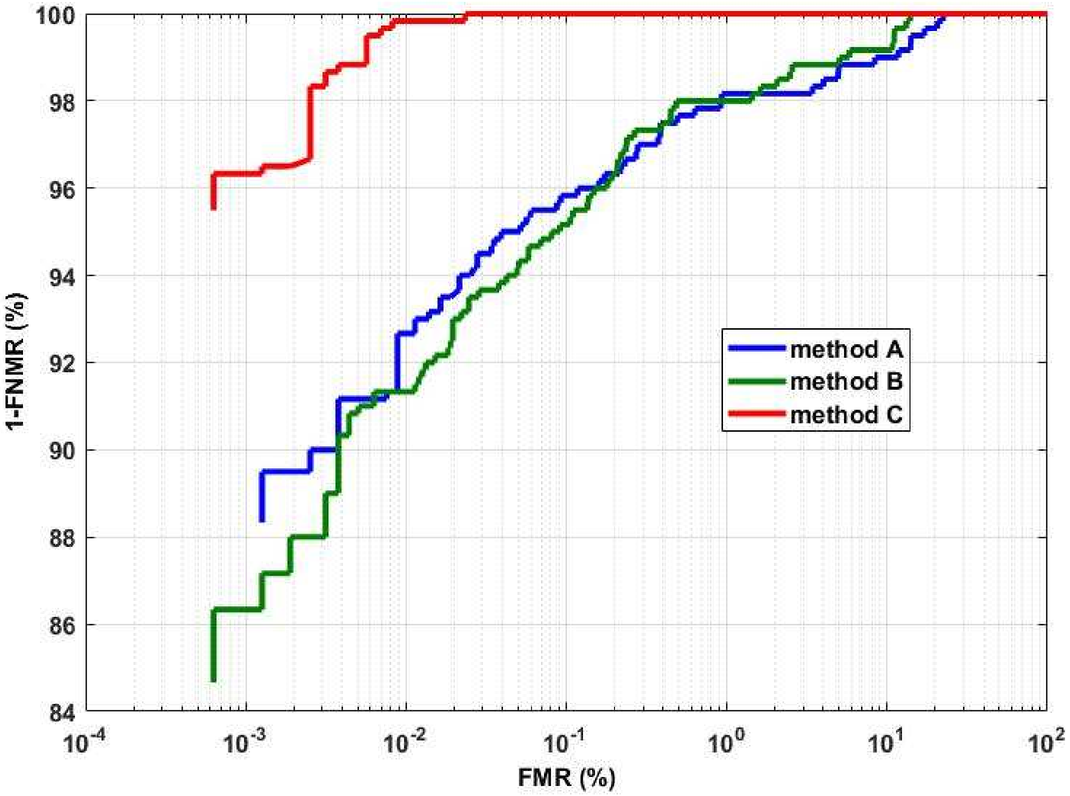


Figure 3.7: Example of ROC curves.

$$\mathrm{FNMR}(t) = \frac{\mathrm{card}\{gms_{ijk} | gms_{ijk} < t\}}{\mathrm{NGRA}} + \mathrm{REJ}_{\mathrm{NGRA}} \tag{3.4}$$

where card(.) represent the cardinality of the test.

In the context of biometric verification, the error rates FMR and FNMR are also known as false acceptance rate (FAR) and false rejection rate (FRR) respectively. FAR can be defined as the proportion of genuine matching scores that are less tan the threshold, whereas, FRR can be defined as the proportion of impostor matching scores that are greater than or equal to the threshold. It should be considered that error terms FAR and FRR are different from the FMR and FNM, as they are computed by considering the number of incorrect acquisitions. Often, the term genuine acceptance rate (GAR) is used to report the accuracy of the biometric verification systems, which can be defined as the proportion of the genuine matching scores that are greater than the threshold.

One of the most commonly used error terms is the equal error rate (EER), which can be defined as the operating point where FMR(t) = FNMR(t), for the corresponding threshold. A biometric system can be considered as accurate as lower the EER is. However, EER does not reflect all the system characteristics. Beside

EER, other commonly adopted performance criteria, proposed in FVC competitions [66], are:

- *ZeroFNMR*: defined as the lowest FMR at which the false rejection is zero
- *ZeroFMR*: defined as the lowest FNMR at which the false acceptance is zero

Both the error rates FMR and FNMR are complementary to each other, and it is not possible to minimize them simultaneously. When a decision threshold is selected to minimize one error rate, the other increases correspondingly. Two kinds of plots are normally used to show the characteristics of the error rates, the Receiver Operating Characteristics (ROC) curve and the Detection Error Tradeoff (DET) curve, which are widely used as a measure of the performance in biometric systems.

The ROC is used to plot the fraction of true positives versus the fraction of false positives, according to the threshold. In the context of biometric systems, FMR values are plotted on the x-axis and the (1-FNMR) values are plotted on the y-axis. ROC curves are plotted in linear, semi-logarithmic, or logarithmic scales. An example of the ROC curves is shown in Fig. 3.7. The DET is used to plot the fraction of false negatives versus the fraction of false positives, according to the threshold. In DET plot, the FMR values are plotted on x-axis and the FNMR values are plotted on y-axis. Normal and logarithmic scales are used to visualize this plot. An example of the DET curves is shown in Fig. 3.8. Both ROC and DET plots are used to directly compare the accuracy of biometric systems, given the fact that the errors must be evaluated for the same dataset.

In the context of the biometric identification, the accuracy is measured in terms of identification rate, which is the proportion of times in which the identity determined by the system is the true identity of the user.

The comparison of the fresh acquired template and the templates stored in the database (1:N matching) results in N matching scores. These matching scores are arranged in decreasing order of confidence to obtain N ranks corresponding to N matches. The best K matches are selected out of these N matching scores to represent the user identity. When the true identity is based on the best found match ($K = 1$), it is known as Rank-1 identification accuracy. In some applications, Rank-K identification accuracy is reported. In this case, the true identity of the user is expected to lie in the top K matches. The error plot commonly used for biometric identification systems is the Cumulative Match Characteristic (CMC) curve. It is computed by plotting ranks on the x-axis and identification accuracy on the y-axis. An example of the CMC curve is shown in Fig. 3.9.

- **Confidence bounds**

  The above-mentioned figures of merit are generally computed for datasets including a limited number of samples. The confidence bounds of the estimated accuracy need to be evaluated to obtain a generalization of the performance measures. Studies in the literature for the confidence bounds estimation can be categorized into parametric and non-parametric approaches.

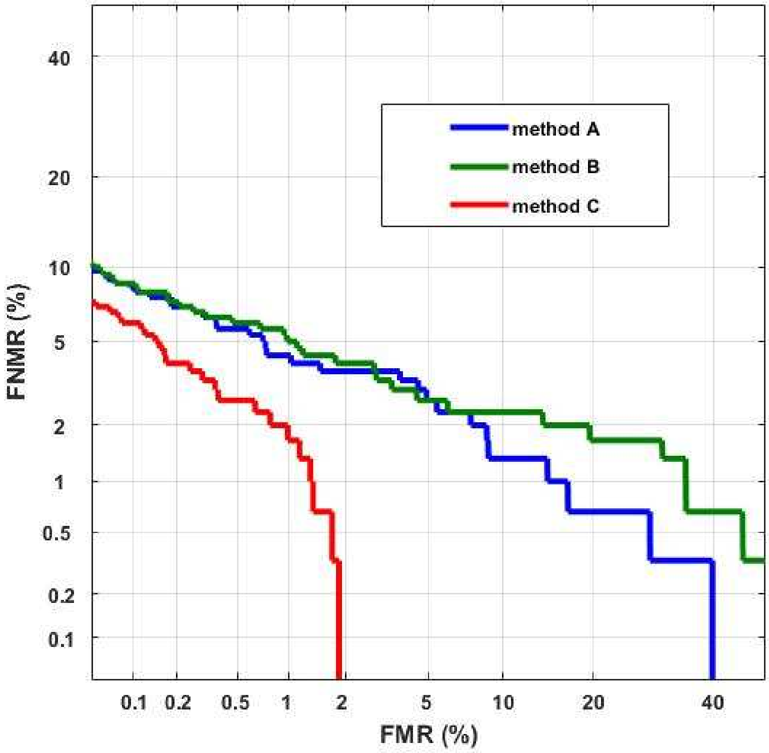


Figure 3.8: Example of DET curves.

In parametric approaches, the confidence of the obtained performance is estimated by using the probability distribution of the matching scores. In case of the datasets with a limited number of users, commonly used techniques are the Rule of 3 [131, 132] and the Rule of 30 [127], which compute the confidence interval of the accuracy by using statistical rule.

The Rule of 3 is used to compute the lowest error rate that can be statistically determined using N biometric comparisons. These error rate p is the error value for which the probability of 0 errors in N trials is equal to a fixed value (usually this value is 5%). It is possible to express the rule as

$p \approx 3/N$, with 95% confidence.

The Rule of 30 assumes that, to have a 90% confidence that the true error rate of the biometric system differs no more that ±30% from the computed error rate, at least 30 errors must be present in the system.

When the number of samples is sufficiently large, the technique based on the central limit theorem [133] can be used to obtain the confidence bounds of the estimated accuracy. The theorem implies the observed error rates should follow an approximately normal distribution. A $100(1-\alpha)\%$ confidence bounds of the

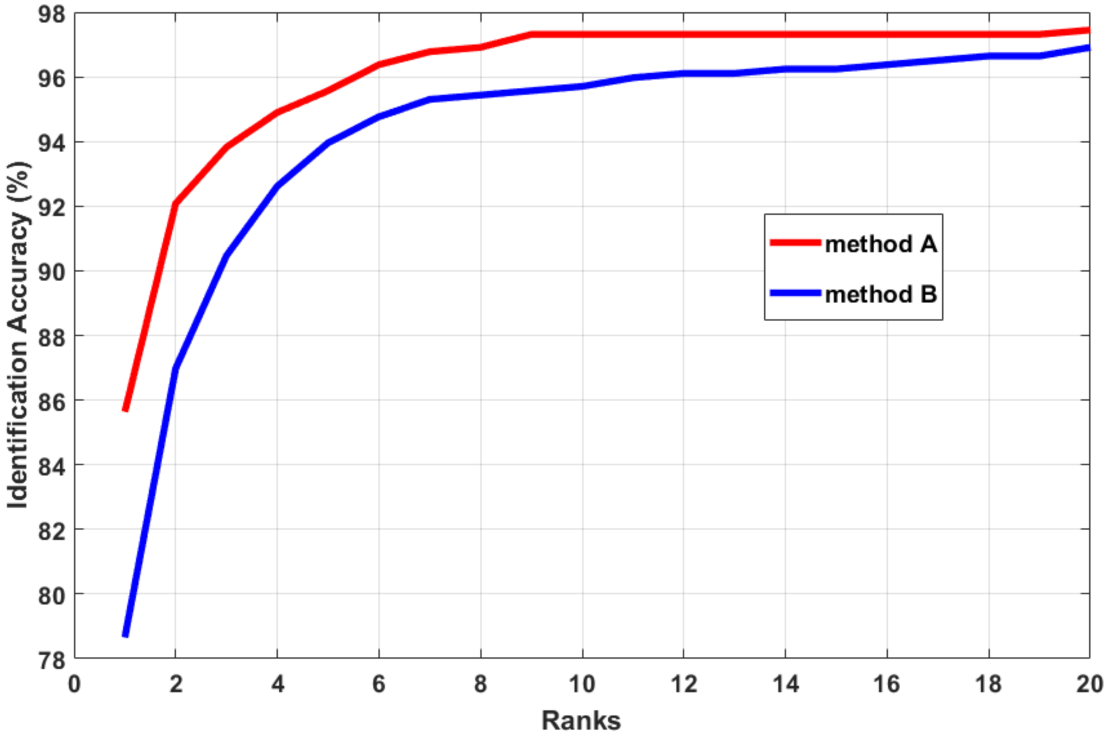


Figure 3.9: Example of CMC curves.

estimated error rates can be obtained under the assumption of normality, by using the formula:

$$\hat{p} \pm z(1 - \frac{\alpha}{2})\sqrt{\hat{V}(\hat{p})} \tag{3.5}$$

where, $\hat{p}$ is the error rate, $\hat{V}(\hat{p})$ is the estimated variance of the error rate, $z(.)$ is the inverse of the standard normal distribution.

In non-parametric approaches, the knowledge of the score distribution is not required. These approaches are used to estimate the confidence of the obtained accuracy in scenarios when the scores distribution is unknown or number of samples is too low to reliably estimate the confidence interval. A commonly adopted technique for the confidence estimation is the bootstrap [134, 135]. This method re-samples the matching scores several times. For each of these re-sampling, it computes the EER. The bootstrap method allows computing $100(1-\alpha)\%$ confidence bounds of the obtained accuracy with an upper and lower limits U and L respectively. The estimated confidence bounds for $\alpha/2$ of the bootstrap values will be lower than L, and $\alpha/2$ of the bootstrap values will be greater the U.

Some other techniques based on semi-parametric approaches for confidence estimation is presented in [136, 137].

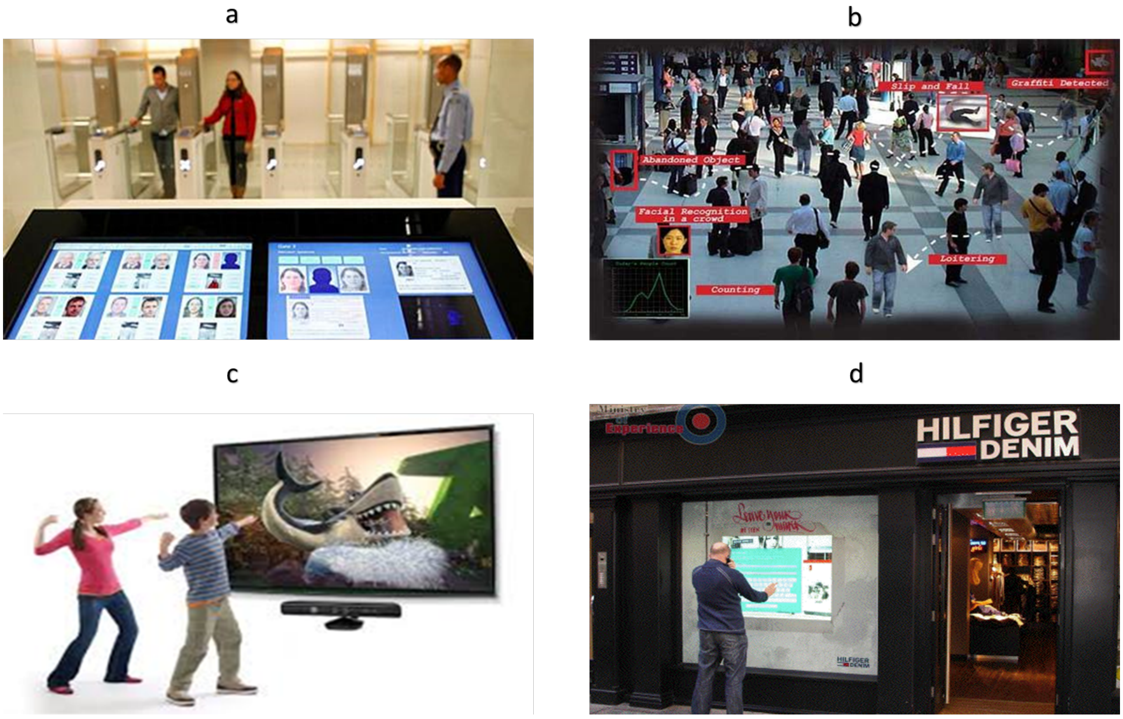


Figure 3.10: Applications of biometric systems in (a) access control, (b) surveillance, (c) entertainment, and (d) shops and malls

3.7 APPLICATIONS OF BIOMETRIC SYSTEMS

Biometric recognition systems have been successfully used in security, surveillance, medical, forensics, and civilian applications [1]. In particular, biometric systems are used in security areas for access control in restricted areas, military services, airports, and automated border controls [1, 118, 66]. In surveillance, it is used in suspicions behavior detection and monitoring, in medical it is used for performing diagnosis and medical analysis [3], in forensics often biometric is used in identifying or verifying the suspects [1], whereas in civilian applications it is used in e-commerce, smart homes, access system in classroom, and entertainment [1, 138]. Recent trends show that biometrics can be used to allow an effective human-computer interaction [106]. Moreover, biometrics can help in designing personalized or priority based services to the user without the need of establishing their identity [5]. An example of different applications of biometric recognition systems is shown in Fig. 3.10

3.8 RECENT TRENDS

The current research trends in biometric systems are detailed in this section.

- *Novel techniques for accuracy enhancement*: many works in the literature aim to improve the recognition accuracy of the biometric systems by investigating new techniques based on hardware and software [8, 1, 4, 66].

- *Managing security and privacy*: the security and privacy issues are prime concerns in the application of biometric recognition systems [139, 140]. Recent research have shown promising techniques to improve the privacy and security of the biometric data [120, 124, 123, 8, 141, 122, 142].

- *Improving the usability and acceptability*: the technological advancements should be accepted by the end users. There are studies in the literature for making biometric systems less-cooperative and to increase the social acceptance [143, 110, 25, 144, 117, 145]. Novel methods based on less-constrained acquisition techniques and non-cooperative users are increasingly studied, in order to improve the usability and acceptability of the biometric systems.

- *Multibiometric systems*: combining information from multiple biometric traits, samples, and sensors have shown promising results in improving the overall performance of the biometric systems with respect to the traditional unimodal systems [138, 146, 103, 147]. However, it is a complex task to integrate these heterogeneous information available in terms of different biometric data, and better techniques for efficient fusion strategies are constantly researched.

- *less-constrained and non-cooperative systems*: to make the biometric system less invasive and user-friendly, researchers have studied touchless and less-cooperative biometric technologies [148, 149, 150]. Moreover, new sensors are researched to acquire biometric data without the cooperation of the users [8].

- *Biometrics for facilitation*: recent studies show that biometrics have plenty of opportunities to provide much more than just recognition or identity management. Even in cases in which the accuracy of the identity establishment is not the primary aim, biometrics can be used to permit an effective and simplified human-computer interaction [5, 108, 151, 106, 152]. Moreover, researchers are studying novel techniques to use biometrics in designing priority based services to the users without the need of establishing her identity.

3.9 SUMMARY

Biometrics is the science of establishing the identity of an individual or a class of people based on the physical, or behavioral attributes of the person. These physiological and behavioral characteristics are unique for each individual and present an irrefutable representation of the associated identity. Biometric recognition systems use stepwise procedures (acquisition, feature extraction, matching, and decision) to establish the user identity. The main operational modalities of biometric systems are verification and identification. In case of verification, the system confirms the identity stated by the user, and in identification, the system establishes the true identity of the unknown user by using her biometric data.

Different biometric traits have different characteristics. Physiological traits represent physical attributes of the human body (e.g., face, iris, fingerprint), whereas behavioral traits represent a specific behavior of a human while performing some tasks (e.g., voice, gait, signature). Soft biometric traits represent some information about the individual (e.g., age, gender, height), but lack of distinctiveness and permanence, and hence may not be used alone for recognition purposes. However, soft biometric information can be used for the continuous authentication systems or to combine them with hard biometric traits to improve the accuracy of traditional biometric systems.

The design of biometric systems needs to consider the specific characteristics of each application scenario. In order to evaluate the biometric technologies for a particular application, different aspects need to be evaluated, such as: accuracy, speed, security, privacy, scalability, and interoperability. The error rates of the biometric systems usually depend on the inter-user similarities and intra-user variations in the identity matching. Further, the system errors such as FTER and FTAR describe the nature of the interaction between users and the biometric systems and need to be considered for effective implementation in practical problems. Moreover, the successful deployment of biometric systems in real-world applications also depends on other criteria, such as: usability, social acceptance, and cost.

Current studies explore new dimensions of applications for biometric systems, which is not only limited to the recognition of an individual but also provide facilitation in terms of personalized services in different environments. Among them, one of the most promising areas of study focuses on the use of biometric technologies for simplifying the human-computer interactions. Novel techniques are being researched for full or partial identity management, which can provide more anonymous recognition system and may find a reasonable balance between the privacy and security concerns. The studies are focused on the design of less-constrained and non-cooperative biometric technologies to improve the usability, applicability, and acceptability of biometric systems in different application scenarios.

4

BIOMETRIC TECHNOLOGIES FOR AMBIENT INTELLIGENCE

Biometric technologies are increasingly used to perform identity recognition in different scenarios that require high security and trust of the user identity. They are frequently used for applications such as access control systems, security checks, border controls, civilian applications, and surveillance. Biometric characteristics provide an irrefutable proof of the associated identity, which can be used to identify and verify the user uniquely. Biometrics can also be adopted to provide personalized services in user-friendly applications by inferring information on the user identity, which allows to select the needs and requirements of a person from a list of characteristics previously inferred by other adaptive services of an Ambient intelligence (AmI) scenario.

In this chapter, we first provide a survey on different biometric technologies used in AmI applications.Then, we present the multibiometric systems and score normalization approaches that can be used in AmI. The analysis of the high-level design of biometric technologies for AmI is also presented. Furthermore, the biometric technologies adopted for continuous authentication systems for AmI are detailed. Finally, the open problems and research challenges for biometric technologies in AmI are outlined.

4.1 MOST USED BIOMETRIC TRAITS IN AMBIENT INTELLIGENCE

There are few studies available in the literature focused on using biometrics for AmI. These studies consider biometric technologies to associate the estimated needs to the specific individual by evaluating their identity, physical characteristics, or behavior. A survey on different biometric traits for AmI has been presented in [5]. The described scenario presents different profiles for each identity (called partial identities) and biometric technologies that enable partial identity classifications such as gender, age, height, and other soft biometric traits. These partial identities enable the users to share different data according to various scenarios, thus providing control of data information to its owner. However, the presented study only considers the technological point of view.

In this section, we discuss the studies in the literature and commercial systems developed by using different biometric traits for AmI.

4.1.1 FACE

Face recognition is the most natural method for human recognition [1]. Face recognition systems can use either a set of captured still images or frame sequences acquired with low user cooperation. These biometric technologies are extensively used for security, forensics, and civilian applications.

One of the important quality of face recognition is its high level of user acceptance due to the low level of required user cooperation. Moreover, facial images can be used for analyzing expressions and emotion and are therefore widely used in AmI scenarios [153, 154].

Face-based systems have been developed for tracking and answering the users in smart environments [155, 15]. A multi-sensor based system is proposed to perform fusion using multiple face images to provide robust recognition system in ubiquitous computing environments [156]. A three-dimensional face recognition system has also been studied for user recognition in AmI [26].

Some studies dealing with face recognition in the domotic environment are presented in [157, 158]. The described systems can recognize and interact with the users without requiring to perform voluntary actions. In [158], the authors proposed a fast and low-cost embedded system for the domotic environment. The presented system (*HomeFace*) can deal with unconstrained facial acquisitions. Facial expression is also analyzed in domotic environments [15]. In this work, the domotic system can understand if the information presented to the user makes her understand the context or not by analyzing the face expression. Due to the increasing research on emotion vocabulary, the interest on facial emotion recognition for HCI in the domotic environment is constantly increasing [16].

Different face recognition methods have been studied for smart environments [159, 160, 161]. In particular, method studied in [159] used a simple neural net for face recognition for aligned and normalized face samples. The $eigenfaces$ based techniques are used to represent facial features. The work presented in [161] used the $eigenfaces$ with the residual error for face detection and recognition. Their method has achieved a reliable and real-time face recognition in a less-constrained environment. Moreover, different feature extraction methods have been studied for face recognition such as Canny filter[162], gradient analysis [163], Adaboost method [164], and LBP (local binary pattern) [165].

To evaluate the real applications of the face recognition methods with larger face dataset in minimally constrained environments, a FERET (face recognition technology) program is established [166]. The methods studied in [167] is based on Gabor jet features and template comparison of face image descriptors using a graph-matching pattern. A linear discriminant based techniques for face recognition have been presented in [168], whereas, a quadratic discriminant based face recognition technique has been presented in [169]. Some other methods based on deep neural networks for face recognition is also studied [170, 171]. Moreover, a commercial face recognition system has been proposed in [172]. Their method was based on a sparse variant of the eigenface transform for face representation and followed by neural networks to perform face recognition. Some other commercial systems developed for face recognition are reported in [173].

4.1.2 FINGERPRINT

Fingerprint recognition refers the pattern-based matching of unique ridges, minutia points, and pores present in the fingertip [174]. It is the most diffused biometric technology, is robust to aging, and offers high recognition accuracy [1, 66]. However, fingerprint recognition technologies usually suffer from low user acceptability due to the high level of user cooperation required for acquiring data [129].

Fingerprint recognition represents one of the most mature and accurate biometric technologies [175]. However, the recognition performance of fingerprint-based technologies can be negatively impacted by non-ideal conditions typical of AmI, such as dirt on the hands (e.g., after eating) or on the sensor (e.g., after multiple uses) [119], stress and lack of effective signaling in an unsupervised context [176].

Fingerprint recognition systems usually require additional sensors with respect to the ones present in the AmI. Nevertheless, fingerprint recognition has been successfully deployed in many commercial applications and more recently in mobile devices, as Apple iPhone 5s and Samsung Galaxy s5, thus making them reliable for authentication in smart environments such as domotic systems. Recent studies proposed to improve the user acceptance of fingerprint recognition in AmI by using touchless acquisition devices [85, 8] that could be more suitable for the smart environments with respect to traditional touch-based recognition methods.

Different fingerprint recognition methods are studied in the literature. Minutiae-based fingerprint recgonition methods are most studied and applied in the literature. This method searches the corresponding minutiae points in the considered feature sets of the fingerprint. Techniques for minutiae-based methods include Hough transform [177], relaxation [178], and energy minimization [179] are frequently used for fingerprint recognition. Methods are based on correlation-based techniques, which computes the matching scores between two fingerprint images, include cross-correlation values [180] and correlation between local regions [181]. Other commonly adopted fingerprint recognition methods include FingerCode template [182], scale-variant features (SHIFT) [183], and Gabor filters [184].

Recently, touchless fingerprint recognition methods are also studied in the literature, which are based on less-constrained acquisitions techniques [8, 85]. Some commercial software for fingerprint recognition include NIST BOZORTH3 [185], Dermalog [186, 187], and Neurotechnology VeriFinger [188].

4.1.3 VOICE

Voice recognition provides a true unobtrusive HCI technology. Voice-based systems can provide user recognition as well as unconstrained interactions between humans and the environment.

Speaker recognition systems estimate the identity of a person based on her speaking utterances [94, 189]. Speaker recognition is mostly used for authentication purposes [94]. Speaker recognition systems can be classified into text-dependent and text-independent [190, 191, 189]. The first class requires that the user enunciate a specific set of words, while the second class does not impose this limitation. Text-independent speaker recognition systems are usually less accurate than text-dependent systems. Nevertheless, text-independent systems are based on a natural and unconstrained HCI modality and are therefore suitable for user-friendly application scenarios.

In text-dependent speaker recognition systems, the phrases spoken is matched with the same enrolled phrase. These systems consider the feature dynamics of the words for identification. The most common modeling techniques for text-dependent speaker recognition are the Hidden Markov Models (HMM) [192] and Dynamic Time Warping (DTW) [193].

Text-independent systems pose no restrictions on the phrases spoken. Hence these systems do not consider the feature dynamics and process the feature vector as a bag of symbols. In this kind of systems, the speakers are frequently modeled by using the Gaussian Mixture Model (GMM) [190, 194] or Vector Quantization (VQ) [195]. GMM requires a large amount of training data to create the speaker model and to estimate a set of parameters (mean, variance, and weights related to each speaker). VQ clusters the speaker data by using k-means clustering. Each cluster is represented by a code that denotes the centroid of the clusters. The set of codes is known as codebooks, which are used to model the individuals. Universal Background model (UBM)[196, 197] is an another technique to model the speaker distribution, which is used for verification purpose. Usually, it uses a very large GMM trained to represent speaker-independent datasets. Other approaches in the literature use Support Vector Machines based on GMM [198] or Artificial Neural Networks [191].

The features corresponding to the speech signal represent differences between the vocal traits of sets of individuals, which are frequently described using the frequency spectrum of the signal [190, 194]. The Mel-Frequency Cepstral Coefficient (MFCC) is one of the mostly used feature extraction techniques [199, 190, 194, 189]. MFCC is a filterbank-based approach designed to resemble the human auditory frequency perception. Other feature extraction methods are: delta-MFCC and delta-delta MFCC [200], linear predictive cepstral coefficients [201], perceptual linear prediction [202], coefficients cepstral mean and variance normalization [203], relative spectral transform filtering [204], feature warping [193], i-vectors and super-vectors [192].

Speaker recognition systems based on deep learning have recently been proposed [205, 206]. The advantage of deep learning is that the system can learn discriminative features from the raw input signal. Studies have shown that deep learning can obtain better accuracy with respect to MFCC and GMM features [207, 208].

Methods in literature for text-independent speaker recognition frequently suffer from some drawbacks. The GMM method with MFCC features [190, 194] provides very reliable accuracy for speaker recognition, but it requires templates composed of a big number of features, which are difficult to store in low-cost hardware architectures. Deep learning shows improved accuracy, but it requires large amounts of training data, which increases the training time of identification applications based on single classifiers.

4.1.4 SOFT BIOMETRIC TRAITS

In AmI, the mostly used soft-biometric traits are extracted from face images, due to the simplicity and high acceptance of the acquisition process. Soft biometrics, such as age and gender play a major role in human-machine interaction to adjust the content presented to the user [78, 209]. Few noticeable applications of age and gender estimation can be considered in security and surveillance, health care systems, and entertainment applications. In particular, the age and gender estimation help in customizing the components of the AmI to provide facilitation in smart environments such as homes, offices, and smart cities [5]. However, their usage is still in the preliminary stage and requires more research for establishing them as a promising candidate for the interactions in AmI scenarios.

Gender classification has gained popularity over the years due to its application in face recognition [210] as well as in human-computer interaction [211]. One of the earliest work on gender classification using neural networks was proposed in [212]. Recently, many studies have focused on gender estimation from the facial images [213, 210, 83, 211]. In particular, a detailed analysis on the comparison of different methods and guidelines for gender classification is presented in [213]. Different features such as LBP, HOG, and BIF are also utilized for gender recognition [211]. Recently, A framework for real-time gender classification from video streams is proposed in [210].

Age estimation from face analysis is a widely studied problem. In the literature, there are techniques for age estimation [214], simulation of the aging process [215], and biometric recognition techniques designed to cope with images acquired at different ages [216]. The studies in the literature on age estimation methods can be divided into machine learning approaches based on feature sets designed to extract discriminative age characteristics and methods that directly infer knowledge from the samples by using deep neural networks. There are studies focusing on general texture features, such as: LBP (local binary patterns) features [217], Gabor features [218], and AAM (active appearance model) [215]. Other studies focus on computing novel features specifically designed for age estimation [219, 82]. The approach presented in [219] uses bio-inspired features, based on pyramids of Gabor filters. The method described in [220] uses the AAM to extract the regions of age local features. Since one of the main problems in designing age estimation methods is the fact that face datasets usually lack sufficient training data for many ages, recent techniques also propose ad-hoc learning methods able to exploit information from the ordinal relationship of the aging labels [221, 222].

Pattern recognition techniques based on deep learning have shown promising results in many real-world applications [223, 224]. The research community has also successfully designed and trained deep neural networks for age estimation [225, 226, 83, 227]. Most of these techniques are based on deep convolutional neural networks (CNNs) [225, 226]. Other methods are based on different approaches, such as the group-aware deep feature learning [227] or dropout-support vector machines [83]. These studies achieved high accuracy for heterogeneous datasets acquired in different application scenarios. However, the main drawback consists of the relevant amount of time needed to train deep neural networks in different application scenarios.

4.1.5 OTHER BIOMETRIC TRAITS

Apart from the above discussed biometric technologies, there are other biometric traits which could be used in AmI, such as gait, iris, palmprint, ear shape, hand geometry, and keystroke dynamics.

Gait recognition represents a promising technology for AmI due to its unobtrusive nature and the fact that it does not require user cooperation. Gait is a behavioral biometrics and can be captured through video sequences [96]. Although, the accuracy of gait recognition systems is lower than that of other more commonly used biometric traits [129]. The application of gait recognition in HCI for domotics requires further studies in real application scenarios.

There are biometric technologies based on different information extracted from eye images and frame sequences [228]. In the context of eye-based biometrics, recognition systems based on the iris are the most diffused ones [91]. This is because the iris shows very high discriminability, iris recognition systems present high recognition accuracy and require low computational time [129]. However, the acceptability of iris biometrics in AmI is very low due to the high level of required user cooperation. Moreover, most of the iris recognition systems use infrared illuminators, which can be perceived as dangerous for the health.

At the best of our knowledge, there are no studies on iris-based systems for AmI. However, recent studies on iris recognition performed at-a-distance, on the move, and in natural light conditions [92, 229] could increase the suitability of iris recognition for AmI applications. Moreover, recent studies on iris recognition with mobile phones [230, 93] could allow the diffusion of iris recognition systems for authentication in the domotic environment. Very recently, the first smartphones performing iris recognition have been launched in the market (e.g., Microsoft Lumia 950XL, Samsung S4, and Samsung Galaxy Tab7).

4.2 MULTIMODAL FUSION AND SCORE NORMALIZATION IN AMBIENT INTELLIGENCE

The main objective of using multimodal fusion in AmI is to increase the security while facilitating the services to the users. For this reason, it is important to rely on well-established and robust fusion techniques.

In AmI, it is needed to design a technology-neutral multimodal fusion method to compensate the biometric information coming from different software or hardware provided by various vendors. AmI is composed of many heterogeneous sensors combined to perform multiple tasks. The addition or removal of any particular sensor may cause the system to restart again from scratch. The technology-independent approach is required in AmI scenarios to manage the information acquired from the sensors. For this reason, it is necessary to design technology-neutral techniques that do not affect existing and proprietary biometric systems. Moreover, it is desirable to investigate a privacy-compliant training procedure, by using different datasets for training and test (including public datasets). This procedure reduces the privacy and data security issues in AmI.

Score level techniques are more suitable for AmI applications, which combine the matching scores obtained from different matching methods. In AmI context, these fusion methods offer a technology-neutral approach, which can favor the integration of the different modules of the biometric recognition process. In the literature, many score level methods have been proposed. However, not all of them are suitable for AmI applications. In particular, the application of learning-based methods that require the training of fusion models [104] is limited by privacy issues, since it is not always possible to use biometric data captured in AmI for training the models.

The application of multi-modal biometric recognition in AmI has been discussed in [231, 152]. In [232] a multimodal dialogue system is developed (*SmartKom*) for combining gesture, speech, and facial expressions. They showed that multimodal systems give better HCI mechanism compared to the single modalities. The work presented in [233] uses the speech, multi-modality, and visual cue acquisition, which enable the human-computer interaction inside the domotic environment. Multi-sensor based fusion approach is presented in [156] to perform fusion using multiple face images to provide robust recognition system in ubiquitous computing environments. Due to the nature of the AmI, which uses a wide set of heterogeneous sensors, multimodal HCI and multimodal biometrics represent challenging and promising research fields. Score level methods have been designed for smart environments by combining face and fingerprint in a hierarchical way [234].

Nonetheless, some techniques can be easily applied to AmI applications and that, to the best of our knowledge, have not been previously tested in real operational environments. These techniques include the well-known methods such as sum rule, product rule, maximum rule, minimum rule, and weighted sum rule.

Several works have demonstrated that the rule of the sum always helps in increasing the recognition accuracy [102].

Usually, the works proposing classifier-based methods require a training phase, use the same database to train and validate the technique, and only in some cases, the tests are performed using techniques such as cross-validation, which allows to avoid over-fitting and obtain realistic error estimations. However, this kind of approach is not directly applicable to AmI applications, because the possibility to store data needed by this operation is not common in real AmI applications. The likelihood ratio technique [104] offers a good alternative in this sense since it is a mature technique that relies on a simple robust model, Gaussian Mixture Models. Besides, it also permits to exploit quality scores.

Privacy protection represents another important challenge for the development of viable fusion methods for AmI. In some countries, the legal framework denies the possibility to store and disseminate biometric data obtained from government systems [235], resulting in a design problem from two perspectives. First, many advanced fusion techniques, such as classifier-based techniques [98], require a preliminary training to tune some of the parameters. In these cases, the larger the amount of data similar to the data that can be found in the operational environment, the more accurate the obtained model will be. However, it is difficult to obtain large quantities of biometric samples due to privacy limitations. Hence, the obtained models could not perform as expected, and it would be preferable to use techniques that only require simple training or that can be trained using public datasets. Second, the operational evaluation of the AmI system is more complex than the procedure used in other application scenarios, and the computation of typical figures of merit, such as FAR, FRR or ROC curves is also more challenging. It is, therefore, necessary to rely on evaluations carried out using public datasets or with internal testing procedures [236].

Several methods have been proposed in the literature to increase the recognition performance of biometric recognition technologies in already deployed systems, such as using a quality threshold to discard low-quality samples [175], fusing multiple images [237], or using multi-modal biometric systems [99]. However, these methods could decrease the throughput of AmI systems and reduce the user acceptance. Moreover, it is possible to use enhancement methods for low-quality images [238], but they need to be tuned according to the used acquisition sensor and feature extraction method, which is frequently manufactured by different producers.

Techniques based only on processing the matching scores resulting from identity comparisons can increase the recognition accuracy independently from the underlying hardware and software [239], and without requiring the user to be subject to multiple biometric acquisitions. These techniques are usually called score normalization methods and aim to increase the biometric recognition accuracy by better separating the genuine and impostor matching scores. They are based on the analysis of sets of genuine and impostor matching scores and can use statistical or computational intelligence approaches. Cohort normalization methods are score normalization approaches that use the matching scores obtained by comparing an input template with a set of cohort templates. Cohort templates are the templates in a biometric system other than the template of the claimed identity [240]. Cohort normalization has the advantage of making no assumptions on the nature of the biometric or the matcher [241], facilitating its application to different scenarios [242, 243] and sensors [244, 148, 245].

Fig. 4.1 shows the cohort score normalization process. A fresh acquired biometric sample (sample A) is compared with the reference sample (sample B) to produce fresh matching scores and also compared with the cohort samples stored in an external database to compute cohort scores. The fresh matching score is normalized using the normalization parameters estimated from the cohort scores.

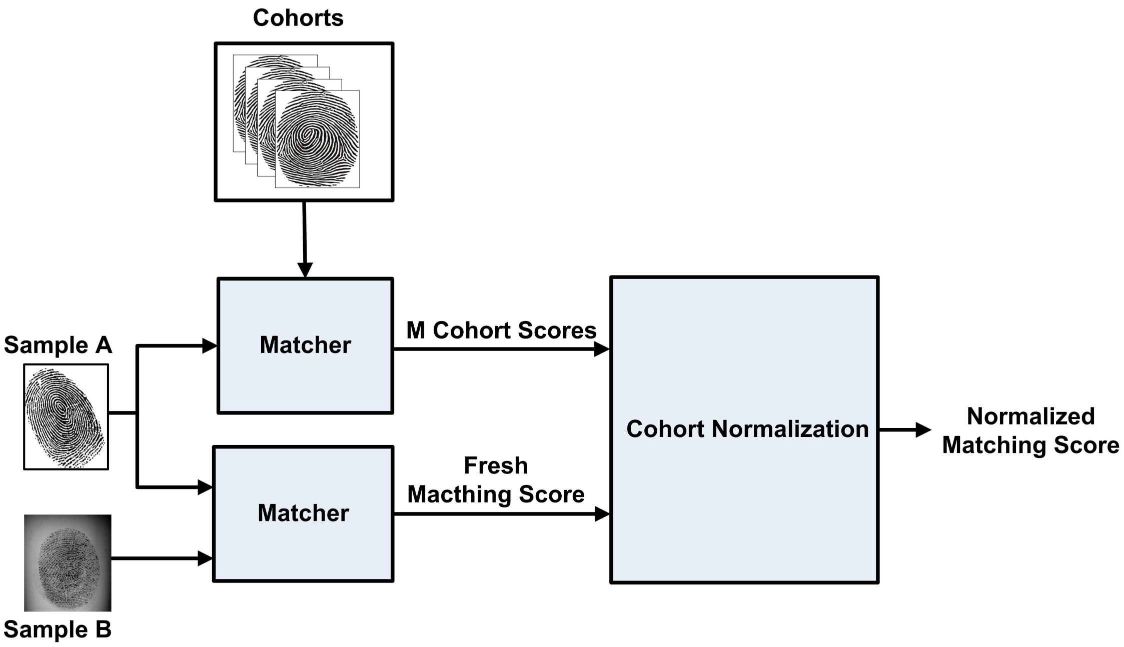


Figure 4.1: Cohort score normalization process

In the literature, there are different studies on cohort-based score normalization methods that aim to increase the accuracy of biometric recognition systems [246, 240], with applications to fingerprint [247], face [248], palmprint [249], as well as multimodal biometric systems [241].

Many methods analyze the cohort matching scores using algorithmic approaches to normalize the matching score computed from the fresh and the enrolled template (fresh score). The method described in [249] compares the fresh score with the highest cohort score for palmprint biometrics. The study described in [247] presents two techniques for fingerprint recognition: the first one normalizes the fresh score using the first and second order moments of the cohort scores, while the second one is the T-normalization. The method described in [241] computes the ratio between the fresh score and the maximum of the cohort scores to increase the accuracy of multimodal recognition systems. Another widely used statistical approach is the Z-normalization [250].

More complex methods train computational intelligence classifiers to learn the relation between fresh and cohort scores. This approach has been applied to fingerprint recognition by using the maximum of the cohort scores or the "'second best matching score"' together with feed forward neural networks[246]. Other methods use the whole set of cohort scores or a significant subset of it as input for a SVM classifier [243].

More recent approaches have shown that not only the most similar templates contain useful information, but also the most dissimilar ones can be exploited to increase the accuracy. For instance, the method described in [240] exploits this information by computing a polynomial regression of the cohort scores. In addition, the size, quality, and the number of users in a cohort set have a direct impact on the performance of the method [248]. The approach also provides a reference for tuning the parameters of cohort-based score normalization methods in the context of unconstrained face recognition.

However, the regulations of some countries pose strong restrictions on the use of biometric data captured for AmI applications [235], limiting the applicability of score normalization techniques. These regulations regard the type of data stored in the systems, limit the amount of usable information, and impose the use of well-known cryptographic algorithms (e.g., AES) that differ from template protection methods specifically designed for biometric systems [124, 251, 123]. In this context, most of the score normalization techniques, which perform multiple genuine identity comparisons, are not suitable because it is not always allowed to store additional data with respect to the biometric samples enrolled in the system. Also, cohort normalization methods, which do not consider genuine matching scores, need to be modified by including privacy-compliant procedures to be used in AmI applications.

4.3 HIGH-LEVEL DESIGN OF BIOMETRIC TECHNOLOGIES FOR AMBIENT INTELLIGENCE

The studied high-level design of biometric technologies for AmI applications consist of different modules that interact one to each other to facilitate the personalized services to the users residing in the environment. Fig. 4.2 shows the schema of the studied high-level design of biometric technologies for AmI. The users can interact with the AmI with voluntary or non-voluntary inputs. The biometric information is collected by the acquisition module, which can be composed of cameras, microphones, and other sensors. The acquired signals are then fed into the signal processing module, which processes the signal, extract meaningful information, and convert the raw signal into a discriminative and compact representation of the user's data (commonly referred as features or templates). The interaction module interacts with vocabulary library to link the processed signal with associated commands. The service selection module selects the service types requested by the user by interacting with the service library and provide the final desired personalized services to the users. The user is in the center of the system, and the objective of the design approach is to facilitate users with simplified human-computer interaction which results in the delivery of the requested services. The components of the studied high-level design are described below:

- *User's input*: The system can take input from the users in the form of voluntary or non-voluntary actions. The voluntary actions required user cooperation to express certain commands to interact with the environment. These commands can be expressed through biometric traits such as fingerprint, gestures, voice, iris, or other traits. Non-voluntary inputs can be acquired by using biometric traits such as the face, gait, gestures, voice, and emotions.

- *Sensor array*: The sensor array is composed of heterogeneous sensors and devices to capture the biometric traits of the users. Different biometric traits have different characteristics, which need to acquire and process differently. The sensor

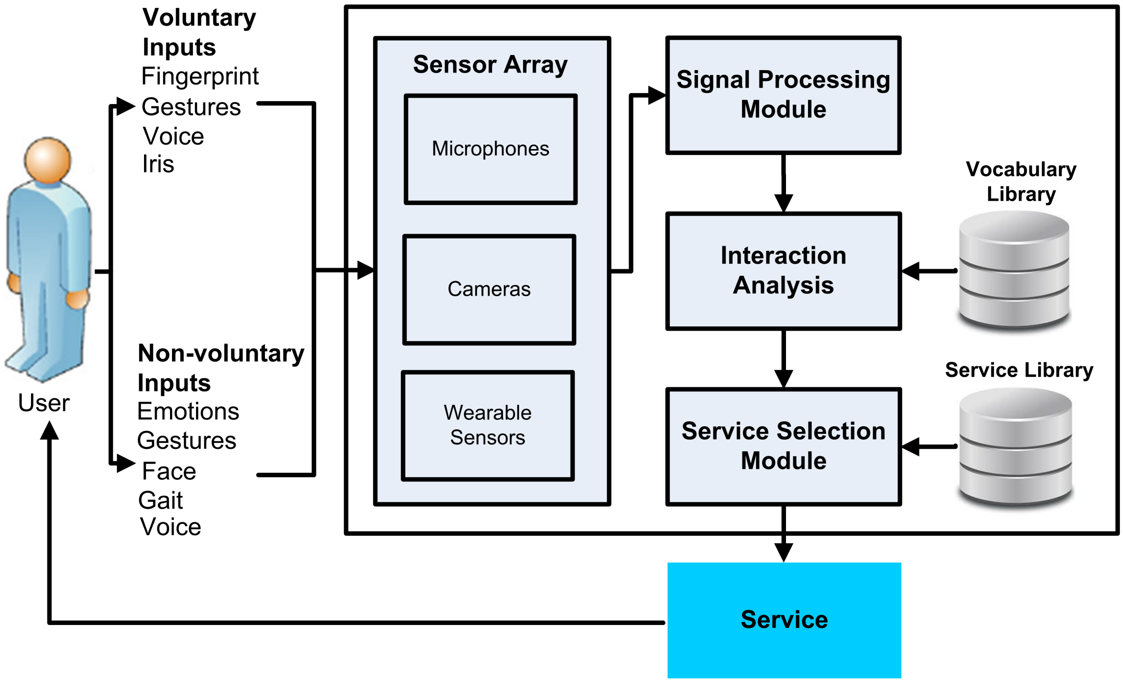


Figure 4.2: Studied high-level design of biometric technologies for AmI.

array includes cameras, microphones, and wearable sensors to acquire the heterogeneous biometric information from the user.

- *Signal processing module*: The acquired biometric traits are processed in the signal processing module. Different biometric signals (such as face, voice, finger, and gait) have distinct characteristics which require a different level of processing. This module converts the raw input biometric signals into corresponding templates. The templates are a compact representation of the signals which posses discriminative information.

- *Interaction analysis*: The interaction analysis module uses the computed biometric templates to interact with the system, to link the actions expressed by the user with a specific category of services. The different types of service information are stored in the service vocabulary. The interaction analysis model interacts with the service vocabulary to extract a list of possible services which can be linked to the input command provided by the user.

- *Service selection module*: The final aim of the studied system design is to select and facilitate the requested services to the users based on the provided input. The service selection module interacts with the service library and selects the desired services to the users. The outcome of this module gives the type of service requested and takes user's feedback regarding the services.

Considering the user-centric nature of AmI applications, the design methodologies for biometric technologies in AmI should present different characteristics with respect

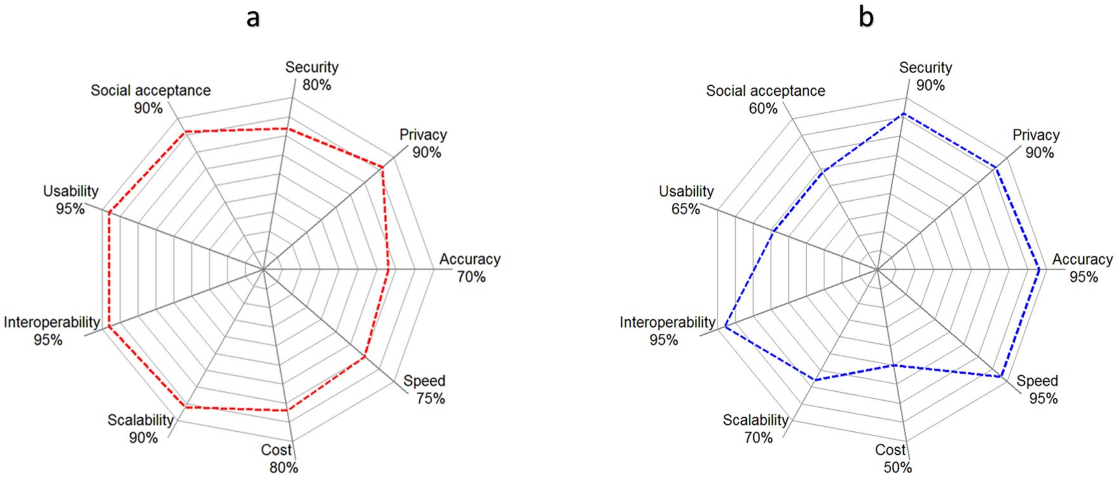


Figure 4.3: Comparison of various characteristics and evaluation aspects of biometric technology for: (a) AmI applications; (b) Security applications. The graphs represent the percentage of importance that we estimated for each of the nine aspects characterizing biometric applications [8]. This figure shows that biometric technologies for AmI present strong differences with respect to biometric technologies used for security applications. AmI is more user-centric and requires technologies with high usability and high user acceptance.

to that of the systems used in a vast majority of biometric applications. As described in chapter 3, it is possible to evaluate a biometric system by considering nine distinct aspects [8]: accuracy, speed, cost, scalability, interoperability, usability, social acceptance, security, and privacy. As an example, Fig. 4.3 shows two nine-dimensional graphs representing the priorities that we estimated for designing biometric systems for security applications security applications (e.g., ABC e-gates) and AmI.

4.4 CONTINUOUS AUTHENTICATION IN AMBIENT INTELLIGENCE

Most of the computer systems authenticate the user once during the initial login session. Once the user is authenticated positively, the access is granted to the user and assumed that the validity of the user is same during the session. This can be a critical flaw in the security of the systems as any impostor user can access the resources of the systems without the permission of the initial signed-in user. To deal with this problem, the authentication of the user needs to perform continuously based upon the activity user perform on the machine. This type of authentication is known as continuous authentication.

AmI is composed of heterogeneous sensors and capturing devices, which are used to acquire the information from the user in direct or indirect manner. The network of sensors used to obtain these information process large amount of biometric data. This large amount of biometric information allows designing continuous biometric authentication in AmI, which continuously tracks users and provides a secure and reliable ambient management systems for accessing resources. In AmI, continuous authentication assures the recognition of correct user accessing the services throughout the time [97, 252, 253]. Moreover, to customize the components of the environment according to the personal preferences of the users, it is required to understand and recognize that the user who is availing the services is the same or is changed over a period.

Continuous authentication systems for AmI must be designed for less-constrained acquisition scenarios and non-cooperative users. These systems should be able to estimate the biometric information from the user in transparent and unobtrusive manner. The use of hard biometrics, such as the fingerprint or iris, may not be suitable for AmI application because of the require user cooperation. Also, systems based on face biometrics may suffer from occlusions, pose variations, strong head rotations, or uncontrolled illumination conditions, which are common in AmI. The design methodology of continuous authentication systems for AmI applications should consider user-friendly and less-constrained biometric technologies to perform authentication without the cooperation of the users. It is required to investigate dynamic and intelligent fusion approaches to integrate the biometric information coming from different sensors with different sampling rate. The multimodal fusion strategies should be based on technology-independent methods. Moreover, the adaptive training approaches need to be considered while designing the continuous authentication system to minimize the effect of addition or removal of a particular sensor from the system.

There is a good deal of studies in the literature on continuous authentication using unimodal and multimodal biometrics. The early methods for continuous authentication using unimodal biometrics are proposed in [254, 255, 256, 257]. Since these methods use a single biometric technology, their methods are not able to authenticate the user and the continuous authentication system forced to shut down the console.

To deal with the uncertainties in the biometric measurement, multimodal continuous authentication systems are proposed [258, 259, 260, 97, 253, 261]. These methods show improvement in the performance of the continuous authentication systems compared to the unimodal systems.

In particular, a multimodal continuous authentication system is proposed in [258] by using face, fingerprint, and voice biometrics. The authors presented two important issues in continuous authentication systems, namely: the integration of biometric traits across modalities over time, and the determination of the authentication certainty even in the absence of biometric measurement. The authors used a score-level fusion based on the weighted sum of the scores computed from each modality. The weighting factor is chosen in such a manner that captures the reliability of the modality and decreases monotonically with the time. When the biometric data is absent, the authentication certainty must go down to maintain the security of the system, irrespective of the fact that the user is in front of the console or not. The method proposed in [259] integrates the face and the fingerprint traits during time by using hidden Markov models model (HMM) in a Bayesian framework. However, the inter-class and intra-class scores distributions are required to compute the reliability of each modality. The work proposed in [260] integrates the face and keystroke traits by using dynamic Bayesian networks (DBN) instead of HMM model. The advantages of using DBN instead of HMM model reside in the fact that it allows more hidden variables to capture contextual information, and in the fact that can integrate the modalities at score level as well as decision level.

In [97], face and soft biometric information, such as colors of user's clothing and facial skin, are used to monitor the user continuously. This method does not require inter-class and intra-class score distributions since no pre-registration is required. A decision level fusion approach is proposed in [261] to integrate three biometric traits (keystroke, face, and skin color) across modalities and over time using the weighted sum of the authentication scores provided by each individual authenticator.

Multimodal continuous authentication systems are also designed for smart phones to provide secure and reliable services to mobile users [252, 262, 263, 264, 265, 266, 267]. In particular, a multimodal continuous authentication system based on face images and touch gestures is proposed in [262]. The authors used a multivariate low-rank presentation method for combining face and touch gestures using feature level fusion. The method proposed in [263] integrates the face and speech modalities using score level fusion for continuous mobile authentication. The authors used LBP based method for face feature extraction and i-vector-based method for speech feature extraction. Other approaches use the fusion of behavioral and text-based modalities for continuous mobile authentication [264, 265, 266, 267, 268].

The multimodal continuous authentication systems based on face and fingerprint [253, 259, 269] require cooperative users to capture the biometric data which decreases the usability of the system. On the other hand, the authentication accuracy of the continuous authentication systems based on face and soft biometrics decreases when the face sample is absent or of poor quality.

Multimodal systems using face and voice biometrics can provide less-constrained and user-friendly systems for authentication. Face and voice are complementary biometric modalities which can be acquired without the cooperation of the user. There are few studies in literature exploring face and voice biometrics for continuous authentication systems [270, 258, 271, 263, 272]. Most of them are designed for either smart phones or protected internet services. There is only one study in the literature for multimodal continuous authentication systems using face and voice for general applications [258]. This study uses a simulated database for evaluating continuous authentication. Table 4.1 shows the comparison of different features of the studied methods in the literature for multimodal continuous authentication using face and voice which are close to the studied approach.

Other studies in the literature for multimodal continuous authentication systems using face and voice are tailored for specific applications such as smartphones or secure web services [270, 271, 263].

Multimodal fusion of biometric data coming from two data streams shows different characteristics. Each data stream may have been sampled at different frame rates, may have different length, possess different temporal characteristics, or less likely to be synchronized [273, 274, 275, 276, 277]. Example of such kind of asynchronous system is audio-visual systems, in which the video data and the audio data may have acquired with different non-constant sampling rates.

Table 4.1: Comparison of the studies in the literature for multimodal continuous authentication using face and voice

Ref	Traits	Type	App.	Fusion	DB	Cont.
[263]	face, and voice	cont.	mobile	logistic-regression	real, not public	not
[270]	face, voice, and touchanalytics	cont.	mobile	feature-concatenation	real, not public	yes
[258]	face, voice, and fingerprint	face: cont. voice: cont. finger: not cont.	general	Bayesian-approach	synthetic	yes
[272]	face and speech	continuous	general	Bayesian, maximum-likelihood	real, public	not

Ref. = Reference, Cont. = Continuous, App. = Application, DB = database

Different methods have been proposed in the literature to synchronize the data streams at the same sampling rate [273]. The most common strategy adopted in the literature is to synchronize the data at a regular interval or combine the modalities at the time instant when they are available in the system. There are also methods based on asynchronous hidden Markov model (AHMM) [276, 275], which are designed to deal with the asynchronous data streams by modeling joint probability distribution of asynchronous sequences of audio and video streams. Other common approaches include dynamic Bayesian networks [278], dynamic time warping [274], and correlation-based methods [277]. These methods have shown promising results in asynchronous multimodal systems. Nevertheless, finding the optimal time period required to obtain different data or best strategy to fuse the data streams are still a challenge and need further investigation.

To the best of our knowledge, there are only two studies in the literature which deals with a multibiometric asynchronous continuous authentication systems [279, 280]. A multimodal continuous speech recognition system based on a multi-stream approach is presented in [279]. This method performs an asynchronous modeling of acoustic and visual speech continuous speech recognition. A perceptual linear prediction based method is used to extract features from acoustic signals. An appearance-based model of the articulators is used from the mouth region to locate, track and recover visual speech features, such as lip tracking and lip shapes. To deal with the asynchronous data streams for multimodal fusion, visual vectors are artificially added (by copying frames). Moreover, in [280] the authors have proposed a novel multibiometric trust model to cope with the asynchronous data streams acquired from the face and keystroke dynamics. In their work, the keystroke is measured continuously, and the face recognition is performed periodically once in every one minute. The designed trust model deals with the cases when the minimum typing pattern or frontal faces are not available for a period, by introducing weights to the individual modality trust models.

Most of the studied methods for continuous authentication perform a one-time enrollment. However, there are few studies that continuously update the stored templates [281, 282, 283, 255, 257, 284, 285]. The idea of template update is to make the reference template relevant to the input templates acquired during the operational phase. In particular, the methods proposed in [281, 285] consider the criteria of change in illumination to update the stored template using image subtraction methods. In [282] a growing and sliding window template update method for keystroke biometric trait is presented. If the user is verified, the new template is inserted into the reference, and the oldest template is removed. However, it is necessary to recompute the model parameters whenever the new template is added in the reference. In [283] unsupervised template update strategies are proposed for managing the variability of the ECG signals. The authors presented a method to automatically update the database of biometric templates by analyzing the results obtained from the biometric queries. In literature, there are also studies on more complex template update strategies for periodically updating the biometric databases [286, 287].

4.5 OPEN PROBLEMS AND CHALLENGES

There are some open problems and challenges which need to be addressed for the technological advancement of AmI systems. Some of the research challenges for the application of biometric technologies in AmI are detailed in this section.

1. *Design issues*: There are no studies on a design methodology for biometric technologies in AmI. The design of biometric systems needs to consider the specific characteristics of each application scenario of AmI. Due to the user-centric nature of AmI applications, biometric systems for AmI should present different characteristics with respect to that of the systems applied in other scenarios such as security and surveillance. The design and implementation of biometric technologies require performing the evaluation of the requirements both from the point of view of users and the systems. It is needed to study the methodologies for formal analysis of application scenarios of AmI. Moreover, the evaluation of the applicability and usability of the biometric technologies in AmI are needed to be analyzed. The design methodology includes the system level and high-level design analysis of biometric applications for AmI, and a suitable representation model (such as UML diagrams).

2. *Integrating the information from multiple sources*: AmI is typically composed of a wide set of heterogeneous sensing devices to capture the personal, behavioral, and contextual information from the users. This heterogeneous information coming from different sensors show different characteristics. Each data stream may have been sampled at different frame rates, may have different length, possess different temporal characteristics, or in general less likely to be synchronized. There are some studies in the literature dealing with the asynchronous data. These methods have shown promising results in asynchronous multimodal systems, nevertheless, finding the optimal time period required to obtain different data

or best strategy to fuse the data streams are still a challenge and need further investigation.

3. *Technology-independent approaches*: In AmI, it is required to design a technology-neutral multimodal fusion method to compensate the biometric information coming from different software or hardware provided by different vendors. AmI is composed of many heterogeneous sensors combined to perform the tasks. The addition or removal of any particular sensor may cause the system to restart again from scratch. A technology-independent approach is required in AmI to manage the information acquired from the sensors. For this reason, it is necessary to design technology-neutral techniques that do not affect existing and proprietary biometric systems. Moreover, it is desirable to investigate privacy-compliant procedures.

4. *Less-constrained and non-cooperative systems*: Another aspect that should be considered to design biometric technologies for AmI is the fact that the technological development should be accepted by the end users. The issues related to the usability and the acceptability of the biometric systems need to be addressed, researched, and improved for the technological advancement in AmI. Most of the biometric technologies studied in literature for AmI require cooperative users to perform recognition, which decreases the usability and acceptability of biometric systems. For example, most of the implemented systems put a constrained on the user to place his face at a particular position and height to acquire good quality samples. The voice-based interactions in AmI are constrained by the distance of the user from the microphone, enunciating a fixed phrase, or in a particular style. Other systems based on iris, fingerprint, or hand and face gestures require active user cooperation. An open research area to design biometric systems for less-constrained acquisitions scenarios and non-cooperative users to increase the usability and acceptability of the biometric technologies in AmI applications.

4.6 SUMMARY

Biometrics has been widely used for identity recognition in various scenarios ranging from granting access at security checks to establishing the identity of the person in forensics. The robust and adaptive nature of biometrics makes it a suitable choice for security, civilian, forensics, and surveillance applications. Recent studies in the literature show that biometric technologies can also be used to provide personalized or priority based services to the user in different AmI.

Different biometric traits have been investigated for the application in AmI. The mostly used biometric traits for AmI include face, voice, gait, gestures, and soft biometric traits such as age, gender, height, and ethnicity. Apart from the unimodal biometric technologies, multimodal biometric systems have also been studied for AmI. AmI is composed of a wide set of heterogeneous sensors to acquire different information from the users. It is required to design novel systems to manage this heterogeneous information available regarding different biometric traits, data from multiple sensors, and quality of biometric samples.

The literature in AmI is prolific and growing rapidly. However, the literature lacks in the studies on how to design the environment intelligent in automatically adapting to the user's preferences. Moreover, it is also needed to design less-constrained and non-cooperative acquisition scenarios to increase the acceptability of biometric technologies in AmI. These issues need to be addressed, researched, and improved for the technological advancement of biometric systems in AmI.

Recent progress in biometric technologies has introduced contactless sensors which are capable of acquiring biometric information at a distance. These sensors allow performing ubiquitous and unobtrusive recognition in AmI. New techniques are studies to combine the various information acquired from different sensors. Nevertheless, it is still an open problem to manage the amount of information available in AmI. The asynchronicity of the data streams presents new challenges for multimodal fusion regarding different sampling rate, length variability, and different temporal characteristics. It is required to investigate dynamic and intelligent fusion approaches to combine the biometric information coming from different sensors with different sampling rate. The optimal strategy to combine these data streams are still a challenge and need further investigation. Additionally, the large amount of biometric information allows designing continuous biometric authentication in AmI, which continuously tracks users and provides a secure and reliable ambient management systems for accessing resources. The continuous authentication systems for AmI should be designed for less-constrained acquisition scenarios and non-cooperative users. Moreover, it is also required to investigate the adaptive training approaches for designing the continuous authentication systems to mitigate the effect of addition or removal of sensing devices from the system.

5

INNOVATIVE FRAMEWORKS FOR BIOMETRIC TECHNOLOGIES IN AMBIENT INTELLIGENCE

This chapter presents the studied innovative design approaches and methods for biometric technologies for AmI. These methods are based on less-constrained and non-cooperative acquisitions for improving the quality of human-computer interactions in biometric systems. Furthermore, the realized methods can manage heterogeneous traits, sensors, and environmental conditions, typical of AmI.

We first describe the studied methods for less-constrained human-machine interactions in biometric systems. These methods consist of novel feature extraction and matching techniques for unimodal biometric technologies in AmI. In particular, the studied methods for text-independent speaker recognition systems in AmI applications are first discussed. Then, the realized techniques for age estimation from non-ideal face images are detailed.

Then, we present the studied methods for novel and comprehensive systems for biometric recognition, able to deal with heterogeneous traits, sensors, and environmental conditions typical of complex AmI. The studied methods are based on adaptive training approaches, dynamic and intelligent fusion techniques, technology-independent and privacy-compliant approaches. In particular, the realized methods based on adaptive cohort normalization techniques for improving the recognition accuracy of the previously deployed biometric systems in AmI applications are discussed. The studied techniques for designing multibiometric systems in AmI applications are then detailed. Lastly, multimodal continuous authentication systems for AmI are presented.

5.1 USER-FRIENDLY AND LESS-CONSTRAINED TECHNOLOGIES FOR HCI IN BIOMETRIC SYSTEMS

This section presents the studied novel and user-friendly technologies for less-constrained HCI in biometric systems. These technologies are based on non-cooperative acquisitions for improving the quality of the interaction between the user and the environment.

The performed studied on less-constrained technologies are divided into two broad categories. The first category is based on methods for voice-based recognition for AmI applications. A novel feature extraction and classification method for text-independent speaker recognition systems are described. The considered system imposes no restriction on the spoken phrases and extracts the voice templates of the user which is composed of only 12 floating-points numbers. The second category is based on methods for human age estimation from facial images. The realized method estimates the age from non-ideal face images acquired with strong rotations and occlusions, which represent a typical scenario in AmI.

5.1.1 TEXT-INDEPENDENT SPEAKER RECOGNITION

Most of current biometric systems are designed for security applications [86, 288]. A growing research area consists of designing biometric technologies to improve the HCI in AmI. These technologies should be based on less-constrained technologies with respect to traditional biometric systems [244, 85, 289, 150]. In this context, voice recognition techniques are of paramount importance due to their high user acceptance amd low required cooperation.

Voice recognition provides a true unobtrusive HCI method. Voice recognition applications can be divided in two categories, namely: speaker recognition (which aims to recognize the user based on her voice) and speech recognition (which aims to recognize what is said). Speech recognition technologies are widely used in HCI for AmI [18, 290]. On the other hand, speaker recognition is mostly used for authentication purposes [94]. Studies in the literature use the speech-based interaction between human and computers to facilitate the users inside their home [291, 292, 232, 233]. The work in [292] presents an AmI based on speech and speaker recognition. The authors deployed the proposed technologies in the domotic system named STARHome, which is a functional prototype. Many commercial applications use speech recognition for HCI technologies designed for AmI, such as Apple Siri, Ubi and Amazon Echo, etc. The work proposed in [293] presents a list of commercial applications using speech for HCI in AmI.

Speaker recognition systems can be classified into text-dependent and text-independent [190, 191, 189]. The first class requires that the user enunciate a specific set of words, while the second class does not impose this limitation. Text-independent speaker recognition systems are usually less accurate than text-dependent systems. Nevertheless, text-independent systems are based on a natural and unconstrained HCI modality and are therefore suitable for user-friendly application scenarios.

In AmI, it is frequently needed to identify the users in relatively small closed-sets by using biometrics. There are two main categories of closed set identification systems: systems performing multiple identity comparisons, and systems that search the identity of the user by classifying a single biometric template. The first category has the advantage of being more scalable since the enrollment of new users does not require

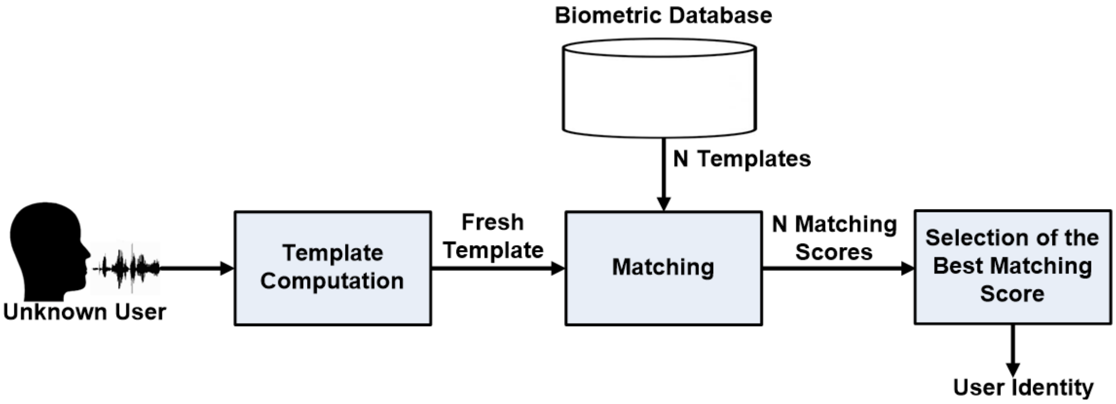


Figure 5.1: Schema of biometric identification systems performing multiple identity comparison.

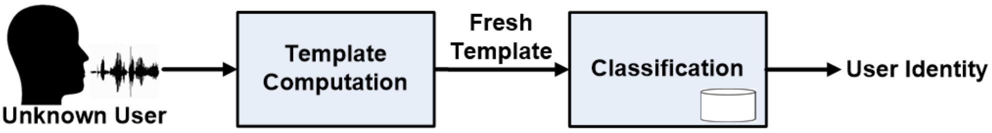


Figure 5.2: Schema of biometric identification systems that search the identity of the user by classifying a single biometric template.

to train any classifier. Nonetheless, having a biometric database of N identities, these systems need to compute N identity comparison to estimate the identity corresponding to the fresh template. Fig. 5.1 shows the schema of the identification systems performing multiple identity comparisons. Differently, the second category of identification systems estimates the final result by performing a single classification, thus requiring less computational time and resources. However, enrolling a new user requires to retrain the classifier. Fig. 5.2 shows the schema of the identification systems that search the identity of the user by classifying a single biometric template.

Biometric recognition algorithms for AmI are frequently deployed in embedded systems, characterized by reduced computational resources with respect to general purpose architectures. Therefore, these algorithms should be optimized in terms of computational time and memory. Specifically, identification applications should be based on fast feature extraction algorithms and use templates of limited size. The computational limitations also justify the choice of identification systems based on classifiers for a wide range of applications.

The studied text-independent speaker recognition method is designed to perform closed-set identification by using a limited amount of computational resources and templates of small size, thus allowing for its use in embedded architectures for AmI.

5.1.1.1 COMPUTATION OF INFORMATION-SET FEATURES

The studied feature extraction method for text-independent speaker recognition systems can be divided into three steps: : computation of MFCC, computation of information-set (ISF) features, and hierarchical classification. Fig. (Fig. 5.3) shows the schema of the studied feature extraction method.

- **MFCC feature extraction**

  MFCC features are widely used in the literature for text-independent speaker recognition [190, 189]. The computation of these features can be divided into the following tasks:

  - Framing and windowing: papers in the literature show that the speaker signal in small time duration windows is stationary and it is possible to extract reliable features in these windows. Hence, the signal is divided into windows of 20ms. The signal extracted from each window is called frame.
  - Computation of the DFT: to extract the spectral information from the signal of each window, we compute the energy available for each frequency band. Therefore, we compute the Discrete Fourier Transform (DFT), as follows:

    $$S_i(k) = \sum_{f=1}^{F} S_i(n)h(n)e^{-(j2\pi kn)/F}, 1 \leqslant k \leqslant K, \tag{5.1}$$

    where $h(n)$ is a sample analysis Hamming window, and K is the length of the DFT.
  - Computation of the Mel Filter Banks: to estimate the energy in different frequency regions, we use Mel Filter Banks [199]. These are triangular filter banks, non-linearly placed throughout the bandwidth and Mel scale. The Mel-spaced scale changes the signal from the frequency domain to the Mel-scale as follows:

    $$m = \ 2595 \log_{10}(1 + f/700) \tag{5.2}$$

    This bank of filters estimates the energy available for each of the frequency bands.
  - Computation of the Logarithm: after computing the Mel filter banks, we compute the logarithm of each filter bank. This task allows us to use the cepstral mean subtraction, which is a channel normalization technique.
  - Computation of the DCT: finally, we compute the Discrete Cosine Transform (DCT) of the filtered signal to estimate the cepstral coefficients. The resulting features are called Mel Frequency Cepstral Coefficients. In this work, we use 12 coefficients of the MFCC.

- **Information set features (ISF) computation**

  The concept of information set was introduced in [294] to enlarge the scope of a fuzzy set using the Hanman-Anirban entropy function [295]. The fuzzy set

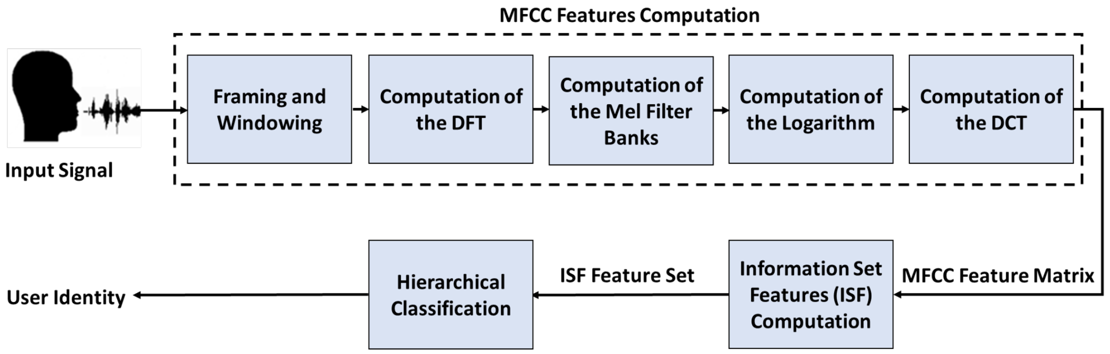

MFCC Features Computation
Input Signal
Framing and Windowing
Computation of the DFT
Computation of the Mel Filter Banks
Computation of the Logarithm
Computation of the DCT
MFCC Feature Matrix
Information Set Features (ISF) Computation
ISF Feature Set
Hierarchical Classification
User Identity

Figure 5.3: Schema of the proposed text-independent speaker recognition method.

theory considers only the value obtained by applying a membership function to a property, without taking into account the value of the property itself. Differently, an information set connects the attribute values and the fuzzified values by using empowered membership functions [296]. Feature extraction approaches based on the information set theory have been applied in biometric systems based on face [297] and ear [298].

Let us suppose a collection of values of an attribute $\Phi = \{\Phi_1, \Phi_2, ...\Phi_n\}$, an empowered membership function is defined as follows:

$$I_\Phi = \sum_i X_\Phi(\varphi_i) G_\Phi(\varphi_i), \tag{5.3}$$

where $G_\Phi(\varphi_i)$ is a gain function. $G_\Phi(\varphi_i)$ is computed as follows:

$$\begin{aligned} G_\Phi(\varphi_i) = \; & e^{-}[a_\Phi(x_\Phi(\varphi_i))^3 + b_\Phi(x_\Phi(\varphi_i))^2 \\ +c_\Phi(x_\Phi(\varphi_i)) + d_\Phi]^{\beta_\Phi}, \end{aligned} \tag{5.4}$$

where the parameters $(a_\Phi, b_\Phi, c_\Phi, d_\Phi, \beta_\Phi)$ are the real valued variables.

This formulation of entropy function can be modulated by selecting a suitable choice of parameters $(a_\Phi, b_\Phi, c_\Phi, d_\Phi, \beta_\Phi)$. As an example, using the variables ($a_\Phi = b_\Phi = 0$, $c_\Phi = 1/2\sigma_j$, $d_\Phi = -\mu_j/2\sigma_j$), we get the following function:

$$G_\Phi(\varphi_i) = e^{-[(x_\Phi(\varphi_i)-\mu_j)/2\sigma_j]^{\beta_\Phi}} \tag{5.5}$$

In this work, we apply the information set theory to reduce the size of the feature set. ISF enables to extract the cepstral as well the temporal possibilistic uncertainties from the MFCC features.

The MFCC feature matrix X is of dimension $(d \times m)$, where d is the number of cepstral coefficients and m is the number of frames. From each cepstral coefficient j of X, the proposed algorithm extracts the first and second order moments, creates a gain function G_j according to the extracted information, and computes

ISF value. The number of features composing the final ISF vector Y is equal to the number of cepstral coefficients d.

We compute every gain function as follows:

$$G_j = e^{-1/2[(X_{ij}-\mu_j)/\sigma_j]^2}, j = 1, 2, ...d, \tag{5.6}$$

where, μ_j and σ_j are the mean and variance of the cepstral coefficient j.

The ISF value for the cepstral coefficient j is then computed using the concept of empowered membership function, as follows:

$$Y_j = \sum_{i=1}^{m} \left(X_{ij} \cdot G_j\right). \tag{5.7}$$

5.1.1.2 HIERARCHICAL CLASSIFICATION

We use a hierarchical classification strategy to estimate the identity corresponding to the fresh template Y. Single classifiers may obtain unsatisfactory accuracy for problems with high numbers of classes involved [299]. To achieve higher accuracy, many studies in the literature use hierarchical classification approaches based on a pool of classifiers. There are different categories of strategies, including the flat classification approach, the local classifier approach, and the global classifier approach. In the studied method, we use a flat classification approach since it is one of the simplest and mostly used techniques in the literature.

Considering a biometric database composed of N enrolled identities, our method uses a pool of N binary classifiers and a score fusion strategy. Each classifier C_i considers the identity i as the positive class and returns a score value $s_i \in [0, 1]$. We use the following strategy:

$$\text{Identity} = \underset{i=1\ldots N}{\text{argmax}}(s_i). \tag{5.8}$$

We consider different types of classifiers: k-Nearest Neighbors (kNN) [300], Feedforward Neural Networks (FFNN) [301], and Support Vector Machines (SVM) [302]. More details on the learning strategies and configurations of the single classifiers are reported in the experimental result section (Section 6.1.1).

5.1.2 AGE ESTIMATION FROM FACE ANALYSIS

Face biometrics is the most natural method for human recognition because of its high level of user acceptance due to the low level of required user cooperation [1]. Face images can also be used to infer a wide set of soft biometric characteristics, such as: the emotional state, ethnicity, gender, and age. Among this set of characteristics, the automatic age estimation can be particularly important in different application scenarios, such as: security and defense scenarios, surveillance, health-care systems, en-

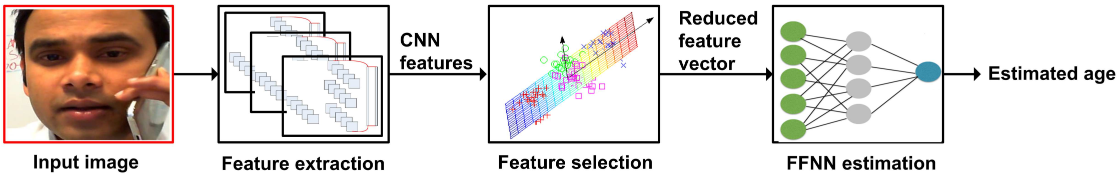


Figure 5.4: Schema of the proposed approach for age estimation.

tertainment, automated border controls, and human-machine interactions in ambient intelligence environments [79].

Recent works in the literature showed that deep learning techniques can achieve promising accuracy in age estimation [227, 83]. Moreover, the studies described in [303, 304, 305] analyzed the possibility of estimating the age by processing a face image using deep networks previously trained for face recognition. All of these methods apply fine tuning strategies to achieve accurate results on heterogeneous image datasets.

The studied approach presents a preliminary study on techniques to increase the accuracy of pre-trained deep networks without applying fine-tuning approaches. To the best of our knowledge, this work presents the first study on age estimation that uses previously trained convolutional neural networks (CNNs) without needing any training or fine tuning of the deep neural networks, thus considering CNNs trained for other applications as generic feature extractors. This approach has the advantage of simplifying the tuning task with respect to fine tuning techniques. The considered methods use heterogeneous networks trained using non-ideal samples to extract robust features from non-ideal face images. A feature level fusion is performed of the data computed by a set of CNNs, reduces the dimensionality of the obtained feature set, and estimates the age by using a feed-forward neural network (FFNN). To achieve robustness to non-ideal conditions, the dimensionality reduction method is tuned using public face datasets of poor-quality face images acquired in uncontrolled conditions.

The studied method estimates the age from a single face image by using multiple pre-trained deep networks that permit to robustly estimate features from non-ideal face images. The considered age estimation method can be divided into three main steps: feature extraction using pre-trained CNNs, dimensionality reduction, and age estimation using FFNNs. Fig. 5.4 shows the schema of the studied age estimation method.

5.1.2.1 FEATURE EXTRACTION USING CNN

We use one or more pre-trained CNNs to extract features from the input face images for age estimation. In this paper, we use VGG-Face CNN [170] and AlexNet CNN [171], although these two deep networks can be easily substituted.

VGG-Face CNN consists of 15 layers, trained on 2.6M facial images from 2622 individuals to perform face recognition. Each CNN block contains a linear operator followed by one or more non-linear layers, such as a rectification layer or max pooling. The first 12 such blocks are convolutional layers. The last 3 blocks are fully connected.

We used the second fully connected layer for feature extraction, obtaining 4096 dimensional feature sets.

AlexNet CNN consists of 8 layers, trained on 1.2M samples of ImageNet ILSVRC challenge dataset [306]. As in VGG-Face CNN, each CNN block contains a linear operator followed by one or more non-linear layers. The first 5 blocks are convolutional layers, while the last 3 are fully connected. We used the second fully connected layer for feature extraction, obtaining 4096-dimensional feature sets.

5.1.2.2 DIMENSIONALITY REDUCTION APPROACHES

First, we create a feature set of $8,192$ features by using the data extracted applying VGG-Face CNN and AlexNet CNN. Then, we reduce the dimensionality of this feature set by applying a dimensionality reduction technique. We analyzed the results of three strategies, namely: Statistical Dependency algorithm (SD) [307], Mutual Information (MI) [307], and Principal Component Analysis (PCA) [308]. To obtain an age estimation method that is easily applicable in different application scenarios, we choose to adopt an external image dataset to train only one time the parameters of the evaluated dimensionality reduction methods. In particular, to increase the robustness to rotations and non-ideal samples, we train the dimensionality reduction strategies using a public dataset of poor-quality face images acquired in uncontrolled conditions (WIKI Dataset [225]).

- Statistical Dependency (SD): This method aims to measure the statistical dependency of the features to its class labels. The feature set is first quantized in such a way that each bin contains roughly the same number of samples from the whole dataset. The statistical dependency between the feature values x and the age y, modeled as a class instead of a continuous value, is computed as:

$$SD = \sum_{x \in X} \sum_{y \in Y} p(x,y) \frac{p(x,y)}{p(x)p(y)} \tag{5.9}$$

  We calculate SD for each feature and select the features having the highest values as final feature set.

- Mutual Information (MI): MI quantifies the mutual dependency between the two random variables. This method takes into account that the statistical dependency between the features and the age labels may be affected by the highly informative level of quantization [307]. To reduce this effect, MI is computed between the features and the age labels as:

$$MI = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \left(\frac{p(x,y)}{p(x)p(y)} \right) \tag{5.10}$$

- Principal Component Analysis (PCA): PCA computes the Eigenvectors from the covariance matrix. The covariance matrix evaluates the amount by which neighboring features are related to each other. The Eigenvectors associated with the largest Eigenvalue are the ones that reflect the largest variance in the image.

5.1.2.3 AGE ESTIMATION USING FEED-FORWARD NEURAL NETWORK

The last step uses the features computed in the feature selection step to train a FFNN. We use a FFNN for regression to estimate a floating point number representing the age, as well as we apply a FFNN for classifying groups of ages. In particular, we use a single linear node for the output layer of the neural network and the scaled conjugate gradient backpropagation method [309] for training. The achieved results of the studied methods for age estimation is reported in the Section 6.1.2

5.2 METHODS TO IMPROVE DEPLOYED BIOMETRIC TECHNOLOGIES

This section presents the studied approaches for improving the recognition accuracy of the previously deployed biometric technologies in AmI. These studies are divided into three main categories.

The first category is based on the studied methods for adaptive cohort normalization techniques. These techniques are generic and can be applied in existing AmI applications without requiring hardware or software modifications.

The second category is based on the studied methods for combining multiple biometric technologies in AmI applications. These methods are technology-independent, which do not affect existing and proprietary biometric systems. Moreover, the realized system uses adaptive and privacy-compliant training approaches which reduce the privacy and data security issues in AmI.

The last category is based on the studied methods for multimodal continuous authentication systems and adaptive fusion approaches for AmI. The studied methods are designed for less-constrained and non-cooperative acquisition, which are able to deal with heterogeneous traits, sensors, and environmental conditions typical of complex AmI. A true multimodal audio-visual database is acquired in our laboratory simulating an AmI. The studied system handles the asynchronous behavior of the multimodal biometric data and deals with the uncertainties in the recognition scores of continuous authentication systems.

5.2.1 ADAPTIVE COHORT NORMALIZATION

Matching score normalization methods can improve the performance of biometric recognition in AmI applications and mitigate the effect of non-idealities typical of this scenario without modifying the existing biometric technologies [239, 99, 310, 311]. Several methods have been proposed in the literature to increase the recognition performance of biometric recognition technologies in already deployed systems, such as using a quality threshold to discard low-quality samples [175], fusing multiple images [237], or using multi-modal biometric systems [99]. However, these methods could decrease the throughput of AmI systems and reduce the user acceptance. Moreover, it is possible to use enhancement methods for low-quality images [238], but they need to be tuned according to the used acquisition sensor and feature extraction method, which are frequently manufactured by different producers.

Techniques based only on processing the matching scores resulting from identity comparisons can increase the recognition accuracy independently from the underlying hardware and software [239], and without requiring the user to be subject to multiple biometric acquisitions. These techniques are usually called score normalization methods and aim to increase the biometric recognition accuracy by better separating the genuine and impostor matching scores. They are based on the analysis of sets of genuine and impostor matching scores and can use statistical or computational intelligence approaches.

Cohort normalization methods are score normalization approaches that use the matching scores obtained by comparing an input template with a set of cohort templates. Cohort templates are the templates in a biometric system other than the template of the claimed identity [240]. Cohort normalization has the advantage of making no assumptions on the nature of the biometric or the matcher [241], facilitating its application to different scenarios [242, 243] and sensors [244, 148, 245].

However, the regulations of some countries pose strong restrictions on the use of biometric data captured for government applications [235], limiting the applicability of score normalization techniques. In this context, most of the score normalization techniques, which perform multiple genuine identity comparisons, are not suitable because it is not allowed to store additional data with respect to the biometric samples enrolled in the electronic documents or smart cards.

The studied cohort score normalization method is generic and does not make any assumptions about the nature of the biometric or the matcher, facilitating its application to different scenarios, and sensors. These methods are based on the privacy-compliant and adaptive normalization approaches for enhancing the biometric recognition accuracy in AmI systems. The adaptability of the analyzed cohort normalization method, which can be used for any biometric traits, makes it suitable for its application in AmI.

First, we present the studied cohort normalization methods for biometric technologies in AmI. Then, a case study for the fingerprint recognition systems is presented to demonstrate the applicability of the studied cohort normalization method for enhancing the recognition accuracy of the previously deployed biometric systems.

5.2.1.1 COHORT NORMALIZATION METHODS

In this section, we describe the studied approach for cohort score normalization in AmI systems. The approach has the features of being privacy-compliant and adaptive to different operational conditions (Fig. 5.5). To comply with privacy protection regulations on biometric data in AmI systems, the considered approach uses an external dataset of templates for computing the cohort matching scores. This dataset could be a public database or a dataset created and maintained secret by the vendor of the biometric technology.

The cohort score normalization procedure works as follows:

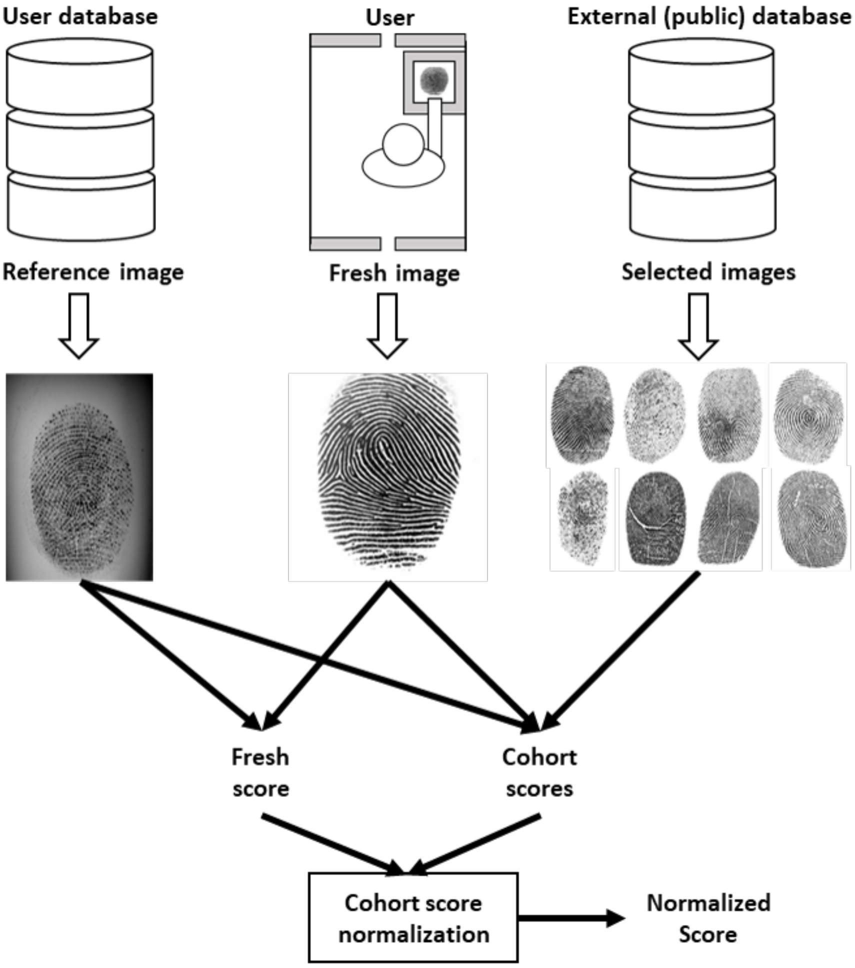


Figure 5.5: Outline of the proposed privacy-compliant cohort score normalization approach. To comply with privacy protection regulations on biometric data in AmI, our approach uses an external dataset of templates to compute the cohort matching scores.

1. For a fresh sample s_i and a sample stored in the biometric database s_j, an identity comparison is performed to simulate the biometric verification performed in the security applications, and the fresh matching score g_{ij} is computed as follows:

$$g_{ij} = match(s_i, s_j), \tag{5.11}$$

   where g_{ij} is the biometric similarity score obtained using the biometric matching function $(\cdot)$.

2. For each sample i, we extract a set E of n samples from an external database, and we compute the set M_i^C of impostor identity comparisons. The number of samples n is empirically selected as a tradeoff between accuracy and computational time. In studied approach, n is constant for each identity comparison g_{ij}. Each matching score m_k^C of M_i^C is computed as follows: $m_k^C = s | s = match(s_i, e_k)$, $e_k \subseteq E$, where e_k is the k-th sample in the external biometric database E.

Table 5.1: Cohort score normalization methods used in the studied method

Method	Description
Baseline	No normalization is performed
Max–rule [241]	Ratio of the raw score to the maximum of the cohort scores for each user.
T–norm [247]	The first and second order moments of the cohort scores are used to normalize the raw score.
SVM–all–cohorts	All the cohort scores for each user are used as input to a SVM classifier.
SVM–20–cohorts [243]	The 20 maximum cohort scores for each user are used as input to a SVM classifier.

3. The final normalized matching score m_{ij} is obtained by applying a cohort normalization method. Table 5.1 summarizes the cohort normalization methods that we considered in this work.

We chose the methods presented in Table 5.1 because they are well-known techniques in the literature. We did not use other well-known techniques that require to store additional information because they are not compliant with the privacy protection regulations imposed in AmI applications.

In the following, we briefly describe the implemented cohort normalization methods.

- The *Max–rule* normalization method [241] computes the ratio between the fresh matching score g_{ij} and the maximum score in the set of cohort scores M_i^C. After computing the set of cohort scores M_i^C from the external dataset E, the maximum of the cohort scores is used to normalize the fresh matching score g_{ij}. The final normalized matching score is computed as follows:

$$m_{MaxRule} = g/\max(m_1^C, \cdots, m_n^C). \tag{5.12}$$

- In the *T–norm* cohort normalization method [247], the first order moment $\mu^C = \mathbb{E}_{m_k^C \in M_i^C}[m_k^C]$ and second order moment $(\sigma^C)^2 = \mathbb{E}_{m_k^C \in M^C}[(m_k^C - \mu^C)^2]$ of the cohort set M^C are used to normalize the fresh matching score g. The final normalized matching score is computed as follows:

$$m_{TNorm} = [(g - \mu^C)/\sigma^C, \text{s.t. } \mu^C, \sigma^C \in M_i^C]. \tag{5.13}$$

- The method *SVM–all–cohorts* classifies a feature set obtained using the samples s_i and s_j in two classes: genuines ($m_{SVM} = 1$) and impostors ($m_{SVM} = 0$). For each identity comparison between the samples s_i and s_j, the SVM takes in input a feature vector x_l composed of the fresh score g_{ij} and the cohort scores M_i^C. The feature vector therefore consists of $n + 1$ values.

To comply with privacy laws, the classifier is trained using only data belonging to the external dataset E.

For each element of the training set, which represents a comparison between the biometric samples e_i and e_j pertaining to the external dataset E, the cohort set M_i^C is computed as the set of matching scores obtained from all the possible impostor comparisons between the sample e_j and the samples pertaining to E.

The training set is composed of $n \cdot (n-1)$ elements representing all the possible combinations of identity comparisons between the n samples belonging to the external dataset E. The number of samples n of E is equal to $x \cdot y$, where x is the number of individuals and y is the number of samples per individual. The training dataset is therefore composed of $x \cdot y$ genuines and $x^2(y-1)y$ impostors.

The application of SVM classifiers to the training dataset can obtain poor results because the class distribution is very imbalanced. To cope with this problem, we use an ensemble learning approach that combines the decisions of 25 SVM classifiers using a voting approach that chooses the most voted class, following the work in [312]. The set of impostor comparisons is divided into 25 subsets obtained by random sampling without substitution, where each subset contains $2 \cdot x \cdot y$ impostor comparisons. Each SVM is trained using the whole set of genuine comparisons combined with one of the impostor comparisons subsets.

- The method *SVM–20–cohorts* [243] is similar to SVM–all–cohorts but it uses a reduced feature set composed of g_{ij} and the 20 highest values of M_i^C. We chose to use 20 cohorts because studies in the literature show that this number allows obtaining a good tradeoff between accuracy and computational time [243].

5.2.1.2 CASE STUDY IN FINGERPRINT RECOGNITION

To demonstrate the applicability of the studied cohort score normalization approaches, we present a case study in the fingerprint recognition systems. The considered method is used to enhance the recognition accuracy of the previously deployed biometric systems based on fingerprint biometrics.

Fingerprint recognition represents one of the most mature and accurate biometric technologies [175]. However, the recognition performance of fingerprint-based technologies can be negatively impacted by non-ideal conditions typical of AmI, such as: dirt on the hands (e.g., after eating) or on the sensor (e.g., after multiple uses) [119], and lack of effective signaling in an unsupervised context [176].

In the presented case study, we studied a novel adaptive cohort normalization approach for AmI applications. First, the considered approach increased the accuracy of fingerprint recognition in our tests simulating AmI systems. Second, the approach considers privacy requirements imposed by current laws, using a privacy-compliant procedure that selects a limited number of cohorts from a fingerprint database captured in different conditions and containing different individuals (e.g., a public database). Third, we apply SVM for the score normalization in AmI systems. Moreover, to the best of our knowledge, it is the first work in the literature that uses external datasets for cohort normalization.

We performed a technological evaluation and a scenario evaluation by using data simulating different conditions. To simulate real scenarios, we used a commercial software for feature extraction and matching, which is currently adopted for fingerprint recognition. The detailed analysis of the obtained results is discussed in Section 6.2.1.

5.2.2 MULTIBIOMETRIC SYSTEMS FOR AMBIENT INTELLIGENCE

We also studied methods for combining the heterogeneous biometric information available in the AmI. The multibiometric systems integrate the information from different biometric systems (e.g., different biometric traits, multiple samples of same trait, different algorithms) at various levels. First, we present the studied technology-independent approaches for multimodal fusion in AmI. Then, the studied methods for the score-level fusion for multibiometric systems in AmI are detailed.

5.2.2.1 TECHNOLOGY-INDEPENDENT APPROACHES

AmI is composed of many heterogeneous sensing devices to acquire physiological, behavioral, and contextual information from the user. The collected information needs to be processed differently. Multibiometric systems in AmI applications require combining information acquired from different biometric traits. Each trait has its own characteristics, and require a distinct level of processing techniques. For example, the feature extraction process for physiological biometric traits such as fingerprint biometrics differs from the techniques used for face or iris images. Similarly, behavioral biometric traits such as voice, gait, or gestures possess distinct characteristics and require different feature extraction steps.

The traditional multibiometric systems integrate the biometric information coming from different sensors in technology-dependent manner. The addition or removal of any particular sensor may cause the system to restart again from scratch. To combine heterogeneous information from multiple biometric traits in AmI, it is important to fuse the traits in a technology-independent manner. The technology-independent approach considers a technology neutral multimodal fusion system which can integrate the information coming from multiple sources (or sensors) independently from their individual processing.

In the studied method for multibiometric systems for AmI applications, we consider a technology-independent approach in which different software and hardware provided by different vendors are used to perform multimodal fusion. From the discussion in the Section 4.2, we argued that the score-level fusion can be a suitable choice for multimodal systems in AmI, which combine the matching scores obtained from different matching methods. We studied a technology-independent score-level fusion method in AmI applications, by using the different combinations of recognition algorithms from different vendors, and analyzing the improvement in the recognition accuracy (in terms of EER and FMR_{1000}) with respect to using only the most accurate biometric trait (the fingerprint). In particular, we used the software Dermalog Fingercode3, Cognitec FaceVACS, Neurotechnology VeriFinger, and Neurotechnology VeriLook. These commercial software uses different image and signal processing tech-

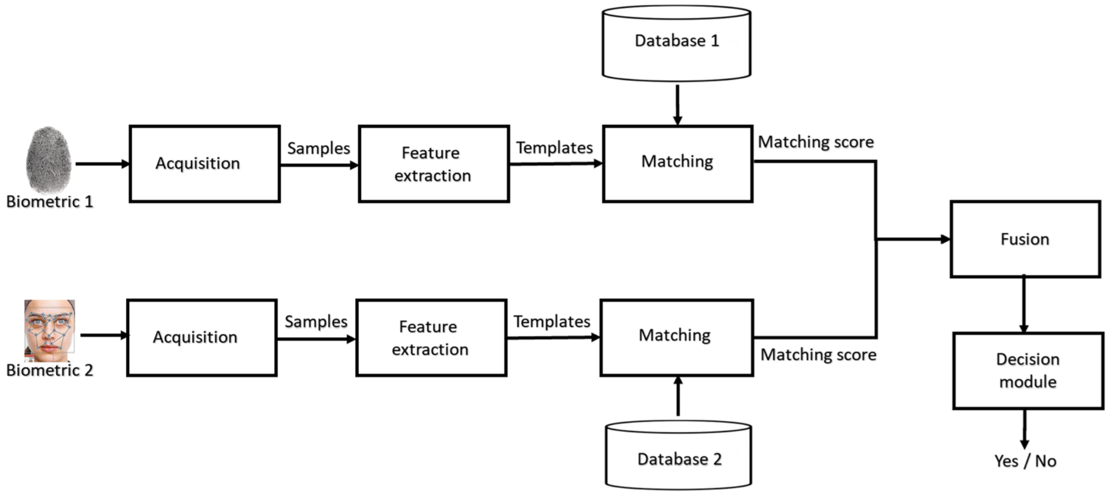


Figure 5.6: Proposed schema of multimodal fusion in AmI.

niques to compute the templates of the corresponding biometric technologies. The realized method integrates information acquired and processed by different software and hardware technologies to evaluate the multimodal biometric fusion in AmI applications. The proposed schema of the multibiometric system that performs the fusion after the matching at the score-level is shown in Fig. 5.6.

5.2.2.2 SCORE-LEVEL FUSION TECHNIQUES

The studied multimodal fusion methods are based on score-level fusion techniques. These techniques are technology-neutral and can be easily applied to AmI applications. First, we considered simple fusion approaches which do not require training. Let us suppose, s_i is the match scores of the biometric templates, and N is the total number of templates available in the system. The studied techniques include the well-known methods such as:

- *Sum rule*: $\sum_{i=1}^{N} s_i$
- *product rule*: $\prod_{i=1}^{N} s_i$
- *Maximum rule*: max $(s_1, s_2, ..., s_n)$
- *Minimum rule*: min $(s_1, s_2, ..., s_n)$
- *Weighted sum rule*: $\sum_{i=1}^{N} w_i s_i$

The sum rule is simple and effective score-level fusion technique, and have demonstrated in the literature that the rule of the sum always helps in increasing the recognition accuracy [102]. Usually, the works proposing classifier-based methods require a training phase, use the same database to train and validate the technique, and only in some cases, the tests are performed using techniques such as cross-validation, which allows to avoid over-fitting and obtain realistic error estimations. However, this kind of approach is not directly applicable to AmI, because the possibility to store data needed by this operation is not common in real AmI applications.

To comply with these constraints, we investigated a robust method for multimodal fusion based on Likelihood ratio [104]. The likelihood ratio technique offers a good alternative in this sense since it is a mature technique that relies on a simple robust model, Gaussian Mixture Models. In addition, it also permits to exploit quality scores of the acquired biometric samples. Incorporating the quality of the biometric samples can provide additional information regarding the computed biometric template, and hence, can improve the performance of multimodal fusion. For these reasons, besides the simple likelihood ratio method, we also studied the quality-based likelihood ratio technique for designing the fusion methods for AmI. The likelihood ratio and quality-based likelihood ratio fusion method can be computed by using the formulas:

- *Likelihood ratio*: $\frac{f_{gen}(s)}{f_{imp}(s)}$,

  where, $f_{gen}(s)$ and $f_{imp}(s)$ are the functions of genuine and impostor distributions, which can be computed as:

  $f_{gen}(s) = \sum_{j=1}^{M_{gen}} p_{gen,j} \phi^N(s; \mu_{gen,j}, \sum_{gen,j})$

  $f_{imp}(s) = \sum_{j=1}^{M_{imp}} p_{imp,j} \phi^N(s; \mu_{imp,j}, \sum_{imp,j})$

- *Quality-based likelihood ratio*: $\frac{f_{gen}(s,q)}{f_{imp}(s,q)}$,

  where, $f_{gen}(s, q)$ and $f_{imp}(s, q)$ are the functions of genuine and impostor distributions with its quality scores, which can be computed as:

  $f_{gen}(s, q) = \prod_{n=1}^{N} f_{gen,n}(s_n, q_n)$

  $f_{imp}(s, q) = \prod_{n=1}^{N} f_{imp,n}(s_n, q_n)$

We also investigated the multimodal fusion based on the Dempster-Shafer Theory (DST) [313, 314, 315]. The Dempster-Shafer theory is based on the theory of evidence, which combines evidence from different sources according to their degree of belief [316]. In DST-based fusion, the probability of a genuine enrollment attempt, the probability of an imposter enrollment attempt, and the probability of the uncertainty of the acquired data and/or classifier are considered when evaluating a biometric match.

Briefly, let us assume $m(G)$ denotes the probability of Genuines, $m(I)$ denotes the probability of impostors, and $m(U)$ represents the uncertainty when it is not possible to determine if the score belongs to the same user or a different user. The term $m(U)$ is an uncertainty function affecting the recognition accuracy of each biometric modality, and can be defined as:

$m_i(U) = \beta(\alpha(1 - Q_i) + (1 - \alpha)EER_i)$

where, Q_i is the normalized quality score and EER_i is the EER of the i^{th} modality. α is the weighting factor and β is the scaling parameter which can be learned from the training set ($0 < \beta < 1$).

The two probabilities $m(G)$ and $m(I)$ can be defined as:

$m_i(G) = S_i(1 - m(U)_i)$, and

$m_i(I) = (1 - S_i)(1 - m(U)_i)$

where, S_i is the normalized match score of the i^{th} modality.

The DST-based fusion method in case of bimodal systems can be computed by using the formula [314]:

- *Dempster-shafer theory*: $\frac{m_1(G)m_2(G)+m_1(G)m_2(U)+m_1(U)m_2(G)}{1-m_1(G)m_2(I)-m_1(I)m_2(G)}$

The DST-based fusion requires the normalization of the scores in the range of (0 to 1) by performing a preliminary normalization.

The considered methods for designing multibiometric systems in AmI applications are evaluated on public biometric databases simulating operational scenarios. The results of the studied score-level fusion techniques and privacy-compliant training approaches for the multimodal fusion systems in AmI applications are reported in detail in Section 6.2.2.

5.2.3 MULTIMODAL CONTINUOUS AUTHENTICATION

After designing the generic multimodal biometric systems for AmI applications, we further studied the methods for designing multimodal continuous authentication systems for AmI applications. The realized continuous authentication systems are based on face and voice biometrics. Face and voice are complementary biometric modalities which can be acquired without the cooperation of the user, and hence are suitable for AmI applications.

First, we present scenarios and the critical points in continuous authentication systems for AmI applications. Then, the asynchronous behavior of multimodal biometric systems in AmI is discussed. The feature extraction and matching of the face and voice biometric technologies are detailed. Finally, the studied methods for designing adaptive neural-based multimodal fusion for continuous authentication is outlined.

5.2.3.1 CONSIDERED SCENARIOS FOR CONTINUOUS AUTHENTICATION IN AMBIENT INTELLIGENCE

We consider an AmI with a single user, where, the user performs some activities (working with a laptop, talking on a smartphone, walking around, and playing games). While performing these activities, the user leaves her biometric data, which can help in providing a continuous authentication. An example of the considered scenarios of acquiring biometric information from the user while performing some activities is shown in Fig. 5.7.

The continuous authentication for AmI applications needs to be designed for less-constrained and non-cooperative acquisitions. Traditional methods of continuous authentication based on hard biometric technologies (e.g., fingerprint and iris biometrics) need cooperative users to perform acquisitions, which may not be suitable for AmI. On the other hand, methods based on soft biometric information (e.g., face and skin color) are usually dependent on the quality of facial features. If the face images are not acquired with good quality, the system may not provide a reliable recognition accuracy.

In AmI, the system should work in an unobtrusive manner, without requiring cooperation of the users. The designed system should not impose restrictions on the

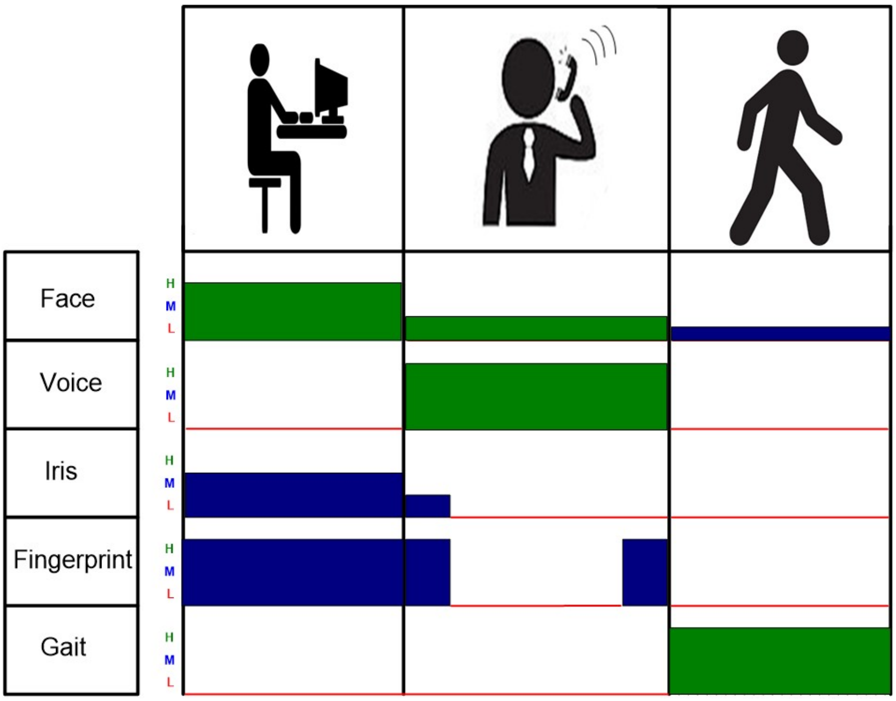


Figure 5.7: An example of the considered scenarios and scope of biometric technologies for continuous authentication in AmI.

users for acquiring her biometric data. Due to these reasons, we investigated multimodal continuous authentication systems for AmI applications based on user-friendly biometric technologies. The studied multimodal systems use face and voice biometrics. Face and voice are complementary biometric modalities which can be acquired without the cooperation of the user.

Most of the studies in the literature for multimodal continuous authentication using face and voice biometrics are usually tested on synthetic databases. These databases typically include frontal faces which may not be the real presentation of actual scenarios. For example, when the user is talking on the smartphone, there are occlusions and rotation of the face. In this interval of time, the match score obtained from the acquired face biometrics rapidly go down and may not be reliable enough for authentication. Nevertheless, the match score computed from the voice of the user at that given time can be used as reliable biometric information to perform authentication. Most of the databases in the literature do not show this "'critical points"' in continuous authentication systems.

To simulate this scenario, we acquired a multimodal database in our laboratory which shows the critical points in continuous authentication. Fig. 5.8 shows an example of the actual scenarios of AmI in which, the user is working on the console, taking calls, walking, and walk-out from the filed-view of the camera. During these activities, the behavior of the recognition scores of face and voice biometrics are shown in the figure. It is possible to observe that, in these critical points, it is required to investigate

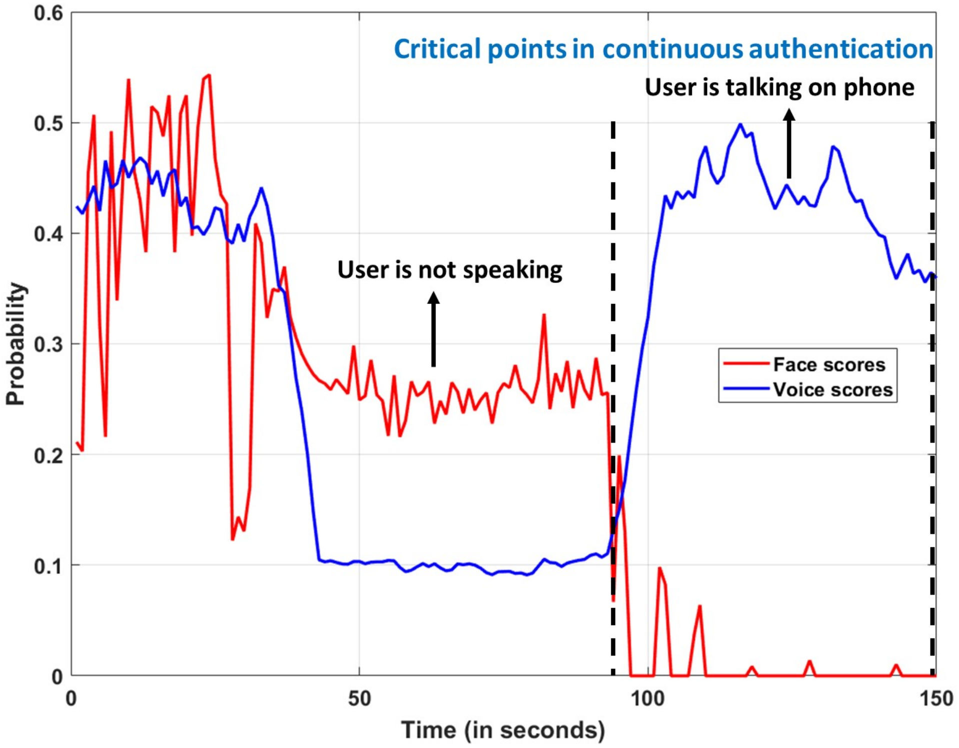


Figure 5.8: Example of critical points in continuous authentication: the face scores go rapidly down when the user starts talking on the phone, nevertheless the voice scores can be used to maintain the trust level of the multimodal system.

a suitable method to integrate the information available in the system to increase the trust level of the continuous authentication systems.

It is important to consider that, during the event, when the user is talking on the phone, it is not always possible to acquire the face images with good quality. It is a realistic assumption, which addresses the problem of missing information of biometric measurements in the certain interval of time in continuous authentication. Most of the approaches in the literature require frontal faces to compute recognition scores. However, in real scenarios, it is possible that the system may not be able to compute face templates when the user is rotated her face at the sharp degree of rotations while talking on the phone.

Traditional methods for multimodal fusion require the availability of the matching scores from all the used biometric modalities to perform the fusion. In cases when the biometric data of one modality is absent, the decision is based on the other available biometric information. In other words, the traditional fusion methods switch between multimodal and monomodal systems depending upon the availability of the acquired biometric samples. Hence, the traditional fusion methods do not provide continuous authentication using multimodal fusion.

The realized system deals with the uncertainties in the recognition scores of continuous authentication systems. The studied methods use computational intelligence techniques based on neural networks for designing multimodal fusion of biometric technologies in continuous authentication.

5.2.3.2 STUDIED ADAPTIVE MULTIMODAL FUSION

One of the important aspects of designing multimodal continuous authentication methods for AmI applications is to consider the asynchronous behavior of the acquired biometric information. AmI is composed of many heterogeneous sensors that collect different biometric samples. The biometric data coming from different streams show different characteristics. For example, the data streams in audio-visual systems consist of the face and voice biometrics, which may present different frame rates, different length, and possess different temporal characteristics.

In continuous authentication systems, the integration of the different biometric signals is performed periodically. The system evaluates the recognition scores of the system in a certain interval of time. However, due to the different sampling rate of face and voice biometric samples, it is difficult to consider the multimodal system in those intervals in continuous authentication systems.

The studied methods for multimodal continuous authentication are based on an adaptive fusion strategy which combines various biometric signals asynchronously.

5.2.3.3 PROPOSED APPROACH FOR CONTINUOUS AUTHENTICATION

The realized approach is based on neural networks for handling the asynchronous multimodal fusion. The schema of the designed multimodal continuous authentication system is shown in Fig. 5.9. The system has a video processing module, which processes the face images extracted from the video data. The audio processing module extracts the voice templates from the audio data. The studied adaptive fusion method based on neural networks perform the asynchronous multimodal fusion. The trust model is designed to track the behavior of the recognition scores in continuous authentication. If the recognition scores go below a certain threshold, the decision module locks the system to maintain the security of the system.

1. **Video processing module**

   The designed video processing module extracts face images from the video data and computes the face templates from the images. The feature extraction and matching of the face samples are performed using a commercial software Neurotechnology VeriLook. The software first computes the face templates of the samples extracted from the faces sequence. The computed templates are matched against the enrolled templates to compute the matching scores. The software computes the similarity scores s between the two templates T_i and T_j. The matching score matrix M is computed by comparing all the acquired face templates with the templates stored in the face database.

2. **Audio processing module**

   The audio processing module analyzes the voice signal. The voice features are extracted using the publicly available Microsoft speaker recognition (MSR) iden-

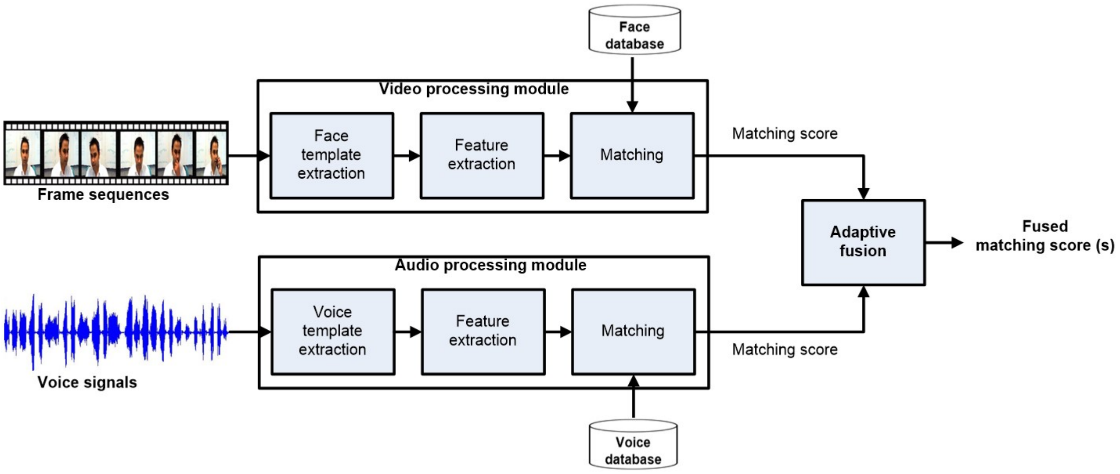


Figure 5.9: Schema of the studied multimodal continuous authentication for AmI applications.

tity toolbox version 1.0 [317]. This toolbox implements GMM-UBM method for text-independent speaker recognition systems, which is a technique to model the speaker distribution. Usually, it uses a very large GMM trained to represent speaker-independent datasets. The steps involved in voice feature extraction and matching are described below:

- *MFCC features*: First, the speech signal is pre-processed by applying a first order high-pass filter to boost high-frequency components to avoid spectrum tilt. Studies in the literature showed that the speaker signal in small time duration windows is stationary and it is possible to extract reliable features in these windows. Hence, the signal is divided into windows of 20ms. The next step is to calculate the power spectrum of each frame. We took clumps of periodogram bins and summed them up to evaluate how much energy exists in various frequency regions. This is performed by using Mel filter banks, which are triangular filter banks non-linearly placed throughout the bandwidth using the Mel scale. MFCC is a filter bank-based approach which is designed to resemble the human auditory frequency perceptron. This technique aims of extracting the low-frequency parts of the features.

  The implementation of the MFCC feature extraction from the raw input voice signal is the one described in Section 5.1.1.1. The template is computed for each of the voice signals.

- *Parameter tuning*: The parameters of this toolbox need to be tuned according to particular applications. There are two main parameters for the GMM-UBM method in this toolbox namely, a number of Gaussian mixtures, and relevance factor.
  - *Gaussian mixtures*: represent the probability distribution of observations in the overall population of the data. The number of mixtures provided is always in the power of 2 (2^m, where $m = 1, 2, 3,...$). Most of the works in the literature use 2048 mixtures for the GMM-UBM method. It is im-

portant to note that, for a classical GMM method, we need less mixture components (usually 8-32 mixtures), but in case of the GMM-UBM method, we need a large number of mixture components to construct a universal background model for overall speaker data. In this test, we used 2048 Gaussian mixtures.

  - *Relevance Factor*: plays an important role in minimizing the effect of unwanted session variability or channel and noise factors, in the MAP (maximum a posteriori) adaptation process. Most of the studies in the literature use the relevance factor in the range of $[8-20]$. We tested different configurations of the relevance factor values and found the best results using the value of the relevance factor is equal to 16.

- *Matching*: The computed fresh templates are matched against the enrolled templates to compute the matching scores. The software computes the similarity scores s between the two templates T_i and T_j. The matching score matrix M is computed by comparing all the computed voice templates with the templates stored in the voice database.

3. **Adaptive fusion method**

The studied method for adaptive multimodal fusion for continuous authentication is based on computational intelligence techniques. The designed method consider the voice signal to improve the recognition score of the continuous authentication system only when the user is talking on the phone. The fusion techniques are designed to handle the uncertainties of the biometric information in multimodal fusion.

- *Fusion with uncertainty in biometric data*: the uncertainty occurs in the biometric data when the system fails to compute the biometric template. In this situation, the representative biometric information is not available in the system, which results in the missing information. Fusion of data with missing templates can lead to force the system to take the decision on single available modality. The system based on traditional fusion rules may not be able to perform continuous multimodal fusion, rather, it performs switching between multimodal and monomodal systems. When both modalities are present, the decision of the system is based on multimodal fusion scores, whereas, if only one modality is available, the system decides on the available biometric data.

  We considered different techniques to replace missing values in input to the fusion module.

  - *mean of the matching scores*: in this case, the replaced value will represent the equal probability of being genuine or impostor user. The value is computed from the training set. The learning of neural network can consider this value as a pattern when the data is missing, and the fusion weights of neurons will be tuned accordingly.

- *minimum of the impostor value*: in this case, we are simulating an impostor attack. The replace value represents a weak impostor. This can be a suitable consideration to design a secure continuous authentication system. The value is computed from the training set.
- *high negative penalty*: in this case, we are helping to better configure the shape of the network when the input is not present in the monomodal trait. The replace value (a high negative value) make a pattern to the neural network to learn the fusion weights in cases when a trait is missing.

After replacing the scores of the missing data with the values mentioned above, a feed-forward neural network (FFNN) is trained using the training set. The FFNN is tested with different adopted strategies and compared with the traditional multimodal fusion techniques such as mean rule, minimum rule, and maximum rule. The considered configuration of FFNN and obtained results for multimodal fusion in continuous authentication is discussed in Section 6.2.3.

5.3 SUMMARY

Innovative methods for biometric technologies in AmI have been presented. The studied approaches are based on user-friendly biometric technologies to facilitate the human-computer interactions in AmI. The main novelties of the studied techniques are that they can deal with unconstrained acquisitions and non-cooperative users, and can manage heterogeneous traits, sensors, and environmental conditions, typical of AmI.

First, methods based on user-friendly technologies for designing less-constrained HCI in biometric systems have been described. The realized technologies are based on non-cooperative acquisition scenarios for improving the quality of interactions between the user and the environment. Moreover, the studied methods for novel feature extraction approaches and matching algorithms for unimodal biometric technologies in AmI applications have been presented.

In particular, innovative approaches for voice recognition systems for AmI applications have been described. The considered method has designed for novel feature extraction and classification for text-independent speaker recognition systems. The studied method imposes no restriction on the spoken phrases, uses a limited amount of computational resources, and computes templates of small size, allowing for its use in embedded architectures for AmI applications.

The studied methods for human age estimation from facial images have been described. The realized method uses pre-trained convolutional neural networks to estimate the age of a non-ideal face image acquired with strong rotations and occlusions, which represent the scenarios in AmI. The studied method for age estimation deal with less-constrained conditions and do not require complex training procedures.

Then, the studied methods for improving the recognition accuracy of the previously deployed biometric systems in AmI have been discussed.

In particular, the studied methods for adaptive cohort normalization have been first presented. The studied approaches use cohort normalization techniques in a privacy-compliant manner, to improve the performance of the existing biometric technologies in AmI. These methods are generic and can be applied in existing AmI applications without requiring hardware or software modifications. A case study in fingerprint recognition systems in AmI applications has been discussed.

Then, the realized methods for combining multiple biometric technologies in AmI applications have been presented. These methods are technology-independent, which do not affect existing and proprietary biometric systems in AmI. Moreover, the realized methods use privacy-compliant and adaptive training approaches which reduce the privacy and data security issues in AmI.

Finally, the studied methods for multimodal continuous authentication systems and adaptive fusion approaches for AmI have been described. The realized methods are based on dynamic and intelligent fusion systems. The studied methods are designed for less-constrained acquisition scenarios and non-cooperative users. The considered methods use computational intelligence techniques based on neural networks for designing novel multimodal fusion in continuous authentication systems. The realized approaches handle the asynchronous behavior of multimodal biometric data and deals with the uncertainties in the recognition scores of continuous authentication systems.

6

EXPERIMENTAL RESULTS

This chapter presents the results achieved by evaluating the studied approaches for biometric technologies in AmI applications (discussed in Chapter 5). The realized methods were evaluated using the testing procedures and figures of merit discussed in Section 3.6.

The realized methods were compared with state-of-the-art techniques. Several public biometric databases have been used to evaluate the considered methods practically. To evaluate the methods in real conditions, biometric samples of different users have been acquired in the laboratory by simulating the actual scenarios in AmI applications. Moreover, a real multimodal biometric database have been collected to test the studied multimodal continuous authentication systems in AmI.

In this chapter, we present the experimental procedures, evaluation protocols, considered datasets, and obtained results of the innovative methods described in Chapter 5. The structure of this chapter is as follows:

First, we present the evaluation methods and achieved results by the studied techniques for text-independent speaker recognition systems in AmI applications (Section 6.1.1). The used databases are presented. Then, the performance of the evaluated methods in different configurations are discussed. Finally, we compare the achieved results with that of state-of-the-art techniques .

Section 6.1.2 describes the experiments and results achieved by the realized methods for age estimation in AmI. The description of the collected biometric samples simulating AmI applications and used public datasets are presented. The configuration of the experiments and evaluation protocols for testing the pre-trained convolutional neural networks are discussed. Then, the considered feature selection and dimensionality reduction approaches are outlined. Finally, we compare the results obtained using public datasets and non-ideal samples acquired in our laboratory, with state-of-the-art techniques.

A discussion on the studied cohort normalization techniques for improving the recognition accuracy of the previously deployed biometric systems in AmI applications is also presented in Section 6.2.1. A case study in fingerprint recognition systems is described. The used datasets and evaluation strategies in a real operational environment are detailed. Then, the privacy-compliant approach of the testing procedure is outlined. A comparison of the obtained results with the traditional methods in the literature in terms of accuracy and time complexity is finally discussed.

Further, the studied methods for multibiometric systems in AmI based on score-level fusion are also analyzed in Section 6.2.2. The description of the used multimodal databases and experimental procedures are first detailed. Then, the considered privacy-compliant and technology-independent approaches are evaluated. Finally, the results of the realized score-level fusion techniques are discussed.

Lastly, the results of the techniques for multimodal continuous authentication for AmI applications are presented in Section 6.2.3. The evaluation procedure of the realized multimodal system for continuous authentication is detailed. The description of the protocol used for acquiring a true multimodal biometric database is detailed. Finally, the results obtained from the studied neural-based techniques for asynchronous multimodal fusion in continuous authentication are analyzed.

6.1 USER-FRIENDLY AND LESS-CONSTRAINED TECHNOLOGIES FOR HCI IN BIOMETRIC SYSTEMS

This section describes the performed evaluation of the studied methods for novel user-friendly and less-constrained technologies for HCI in AmI. In particular, the experimental procedures and achieved results of the studied methods for text-independent speaker recognition systems have been first presented. Then, the configurations of the experiments, evaluation techniques, and achieved results on the studied methods for age estimation from face analysis have been discussed.

6.1.1 TEXT-INDEPENDENT SPEAKER RECOGNITION

This section presents the experiments related to the studied methods based on text-independent speaker recognition systems designed for AmI applications in Section 5.1.1.

6.1.1.1 DATABASE DESCRIPTION

We evaluated the proposed method using a set of signals belonging to the Switchboard NIST 2003 SRE speaker database [318]. This database consists of 356 voice signals recorded on the telephone for a duration of 2 minutes, with a sampling rate of 8kHz at 16 bit. We extracted the speech signals of the 149 males of the training set. To create the samples for our tests, we divided the 2 minutes audio signals into five samples of 24 seconds each. In this manner, we obtained 745 samples (5 samples per individuals).

Table 6.1: Rank-1 Identification Accuracy achieved by the baseline method (MFCC+GMM) and the studied method for text-independent speaker recognition in different configurations

Methods	**Rank-1 Accuracy (%)**
ISF+KNN	64.43
ISF+FFNN	83.22
ISF+SVM	85.64
MFCC+GMM	78.66

6.1.1.2 EVALUATED METHODS AND CONFIGURATIONS

We used GMM with MFCC features (MFCC+GMM) as a baseline since it is a widely used method in the literature. To learn the parameters of GMM, we tested different numbers of mixtures (in the range of [1,...,32]) and we found the best configuration by using 16 Gaussian mixtures.

To evaluate the performance of proposed feature set (ISF), we used hierarchical classifiers based on three computational intelligence techniques, namely: kNN, FFNN, and SVM.

We tested kNN classifiers with k = 1, 3, and 5, and achieved the best results using k = 1 and Euclidean distance.

The considered configurations of FFNN are designed as follows. We used a single linear node for the output layer of the neural network. We tested different numbers of nodes with tan-sigmoidal transfer functions for the hidden layer (in the range of [1,...,100]) . We trained the neural networks using the Levenberg-Marquardt backpropagation algorithm with 500 epochs. We obtained the best results using 80 nodes in the hidden layer.

We tested three variants of SVM kernels: the linear kernel and two non-linear kernels: Gaussian kernel, and polynomial kernel of order 2. To learn the parameters of non-liner kernels, we optimized the value of σ in the range [0.1,...,3]. We achieved the best results using a Gaussian kernel with σ = 1.70.

6.1.1.3 ANALYSIS OF THE IDENTIFICATION PERFORMANCE

To evaluate the performance of the realized speaker recognition methods, we adopted an iterative-validation strategy. We performed 5 iterations. For each iteration, we randomly selected 4 samples per user to create the training set. The validation set was composed using the remaining sample for each individual.

We compared the results obtained using the designed ISF templates and different hierarchical classifiers with that achieved using MFCC features. Table 6.1 summarizes the results achieved by the evaluated methods. This table shows that the studied method increased the accuracy of the speaker recognition systems. The considered feature ex-

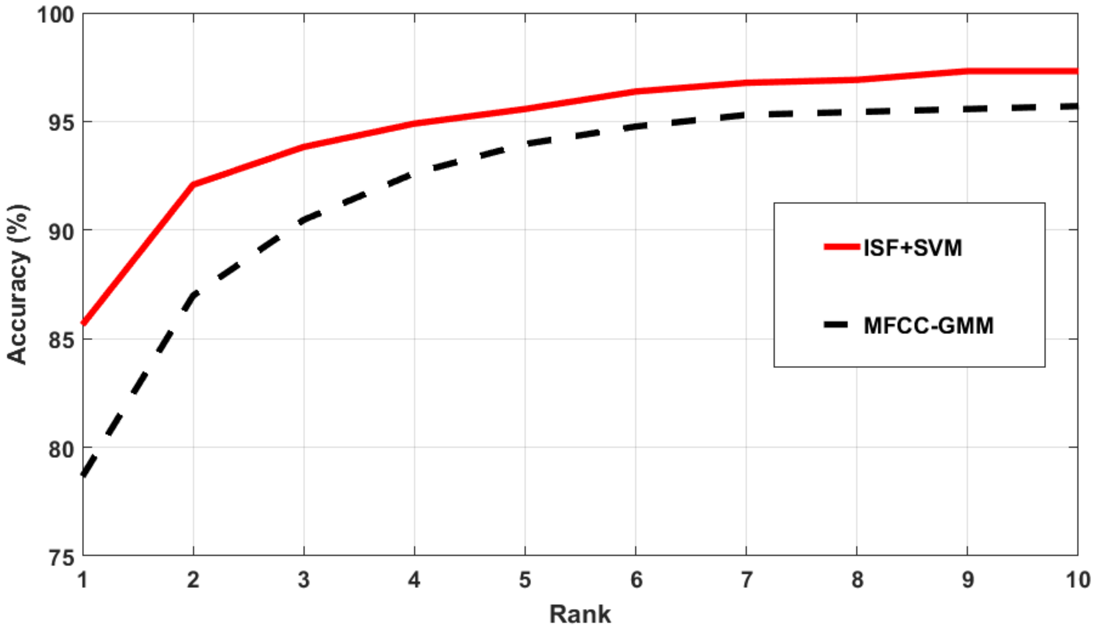


Figure 6.1: CMC curve achieved by the studied method for text-independent speaker recognition in its best configuration (ISF+SVM) and by the baseline method (MFCC+GMM). The realized method achieved better accuracy for each considered rank.

Table 6.2: Rank-1 Identification Accuracy achieved using different number of enrolled samples per user

Methods	Rank-1 Accuracy (%)			
	1 enrolled	**2 enrolled**	**3 enrolled**	**4 enrolled**
ISF+KNN	42.62	51.90	59.06	59.06
ISF+FFNN	48.20	69.60	71.80	80.50
ISF+SVM	65.94	74.72	80.54	84.56
MFCC+GMM	66.95	67.79	73.49	76.51

traction method in combination with a hierarchical classifier based on SVMs (ISF+SVM) achieved the best result on the considered dataset, with Rank-1 identification accuracy of 85.64%.

Fig. 6.1 shows the Cumulative Match Characteristic (CMC) curves obtained by comparing the baseline method (MFCC+GMM) and the studied method in its best configuration (ISF+SVM). Notably, the studied method achieved higher identification accuracy for all the considered ranks.

6.1.1.4 PERFORMANCE WITH REDUCED NUMBER OF ENROLLED SAMPLES

To investigate the application of the realized method with less enrolled samples, we performed an analysis by training the hierarchical classifiers with reduced numbers of enrolled samples per user.

To simulate this scenario, we adopted a 2 fold validation strategy. The first n samples per individual were selected as a training set. We tested four scenarios in which the classifiers were trained with 1 enrolled sample per user, 2 enrolled samples per user, 3 enrolled samples per user, and 4 enrolled samples per user respectively. The validation set was composed using the remaining samples per individual.

Table 6.2 reports the results obtained. In each evaluated scenario, the studied method achieved the best results. The studied method (ISF+SVM) achieved Rank-1 identification accuracy of 84.56% for 4 enrolled samples per user, 80.54% for 3 enrolled samples per user, 74.72% for 2 enrolled samples per user, and 65.94% for 1 enrolled sample per user. These results show that the realized speaker recognition system is capable of achieving better accuracy than the baseline method with a limited number of training samples.

6.1.2 AGE ESTIMATION FROM FACE ANALYSIS

In this section, we discuss the experimental setup, used datasets, and the obtained results by the studied methods based on age estimation using facial images in AmI applications. First, we describe the used public datasets and samples acquired in our laboratory simulating the conditions in AmI applications. In particular, we analyzed three test scenarios to evaluate the implemented techniques. First, we evaluated the performance of the studied feature selection methods. Second, we evaluated the performance of studied methods in general scenarios using a public dataset. Third, we evaluated the applicability of the realized method on images acquired in our laboratory simulating less-constrained and non-cooperative user scenarios. Finally, the analysis of the results obtained on public datasets and simulated less-constrained acquisition scenarios are presented.

6.1.2.1 DATABASE DESCRIPTION

We used sets of biometric data belonging both to public biometric databases and sets of samples acquired in our laboratory by simulating the acquisitions performed in less-constrained conditions. We tested our method using the following datasets:

- *Public datasets*: are used to train and compare the performance of the studied methods with respect to the state of the art techniques. In particular, two public datasets are used.
  - *WIKI Dataset*: we used this database [225] to estimate the parameters of feature selection methods and training the final models. It consists of $62,359$ images of popular celebrities from Wikipedia. Most of the images in the database display various appearances with respect to the poses and illumi-

Figure 6.2: Examples of cropped faces in WIKI dataset.

nation conditions. Some examples of faces in WIKI dataset is shown in Fig. 6.2.

- *Adience Benchmark Dataset*: to compare the classification accuracy of our method with state of the art techniques, we used the dataset Adience Benchmark of Unfiltered Faces for Gender and Age Classification [83]. This dataset consists of $26,000$ face images from $2,284$ individuals. The dataset is categorized into eight age groups: {[0, 2], [4, 6], [8, 13], [15, 20], [25, 32], [38, 43], [48, 53], [60, -]}. In our study, we assigned integer classes from 1 to 8 to the age groups sorted in increasing order. As described in [83], to process these samples using CNNs, we cropped the face region using the face detector proposed by Viola and Jones [319] and aligned the frontalization method described in [320]. Some examples of faces in WIKI dataset is shown in Fig. 6.3.

- *AmI-Face Dataset*: this dataset was created to evaluate the studied methods in non-ideal conditions. It consists of face images captured at various distances (25cm, 50cm, 100cm, and 150cm) from a webcam. To simulate a less-constrained acquisition scenario, we acquired faces at various degree of rotations:
  - images captured with the face at 0^o rotation (frontal face).
  - images captured with the face at 22.5^o rotation.
  - images captured with the face at 45^o rotation.
  - images captured with the face at 75^o rotation.
  - images captured with the face at 90^o rotation.

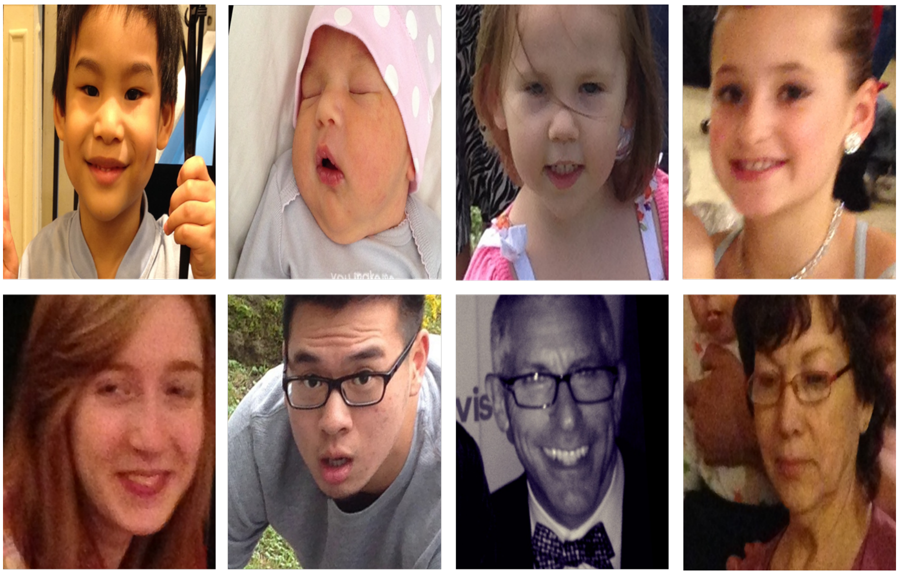


Figure 6.3: Examples of cropped faces in Adience benchmark dataset.

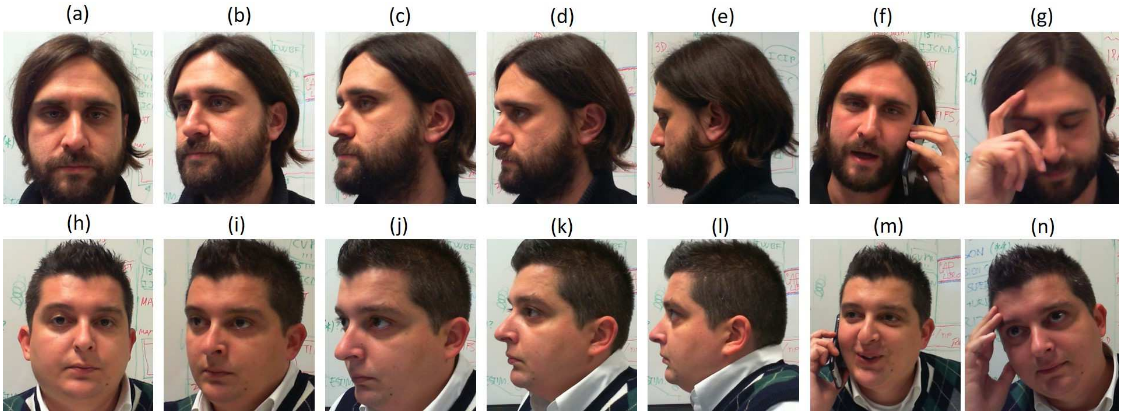


Figure 6.4: Examples of cropped faces in AmI-Face Dataset simulating less-constrained and non-cooperation scenarios including rotations: (a, h) frontal, (b, i) 22°, (c, j) 45°, (d, k) 75°, (e, l) 90°; and activities: (f, m) using cellphone, and (g, n) expression changes.

Moreover, two additional activity scenarios are included. In the first scenario, the images are acquired when the user is talking on the smart-phone, and in the second scenario, faces are acquired with different face expressions.

The dataset is composed of 4,535 face images of 16 individuals. The images were captured using a Microsoft LifeCam HD-3000. Fig. 6.4 shows some examples of the used face images.

6.1.2.2 CONFIGURATION OF THE EXPERIMENTS

A computational intelligence technique based on feed-forward neural networks (FFNN) was used to estimate the age from the face images. We tested different numbers of nodes with tan-sigmoidal transfer functions for the hidden layer of the FFNN regres-

Table 6.3: Comparison of the three feature selection methods in terms of MAE (in years)

Features	Test database											
	CNNi + PCA + FFNN				CNNi + MI + FFNN				CNNi + SD + FFNN			
	V	A	Mean	V+A	V	A	Mean	V+A	V	A	Mean	V+A
10	4.38	4.72	3.44	3.51	7.03	4.41	5.12	—	6.96	4.41	5.07	—
30	4.05	5.17	3.87	**3.30**	7.93	4.28	5.14	9.86	7.05	4.21	4.82	9.39
50	4.17	6.58	4.33	3.40	5.46	5.33	4.14	7.39	4.80	5.31	4.13	7.39
100	4.10	5.09	3.64	3.45	4.70	5.23	3.94	—	4.68	5.51	4.08	—
200	4.52	5.74	3.88	3.85	4.40	5.56	5.12	—	4.13	4.89	3.49	—
	V (4096 features)				A (4096 features)				V+A (8192 features)			
All	5.4				5.86				—			

We trained the feature reduction and the regression methods using WIKI Dataset, while we tested the performance of our method using AmI-Face dataset.
V = VGG Face CNN, A = AlexNet CNN, Mean = decision level fusion of V and A,
V+A = feature level fusion of V and A, — = the FFNN did not converge to a suitable model.

sion models (in the range of [1-100]). We imposed a maximum of 2000 training epochs. We obtained the best results using 15 nodes in the hidden layer. We evaluated the age estimation error in terms of MAE (Mean Absolute Error), which is the average of the absolute difference between the estimated age and the actual age. For age classification, we evaluated the exact classification error.

6.1.2.3 EVALUATION OF DIMENSIONALITY REDUCTION METHODS

We evaluated the accuracy of considered dimensionality reduction methods. To analyze the capability of the studied method to be applied in different scenarios without performing any training, we performed the dimensionality reduction and training of the regression methods using WIKI Dataset. We tested the obtained models on the AmI-Face dataset. In particular, we checked the performance of the method using the features obtained from VGG-Face CNN, AlexNet CNN, and different feature reduction strategies. We also compared the performance of the proposed feature level fusion with that of the single CNNs and of a method performing a decision level fusion (computed as the mean of the ages estimated using the CNNs singularly). Table 6.3 summarizes the obtained results.

Results show that, for all the tested configurations, the dimensionality reduction improved the estimation of the MAE with respect to the full feature set. In our tests, VGG-16 CNN performed better than AlexNet CNN, probably due to the fact that VGG-16 CNN has been designed for face recognition and Alexnet CNN has been designed for analyzing general images. The best performing feature set is the feature level fusion of VGG-Face CNN and AlexNet CNN using 30 features obtained after applying

Table 6.4: Performance of the studied method for age estimation from face analysis in less-constrained conditions in terms of MAE (in years)

Frontal	22°	45°	75°	90°	Cellp.	Expr.	DB
3.12	3.03	3.00	3.23	3.48	3.58	3.58	3.30

We trained the feature reduction and the regression methods using WIKI dataset, while we tested the performance of the studeid method using AmI-Face dataset.

Cellp. = cell phone, Expr. = expression changes, and DB = complete dataset.

PCA. Moreover, PCA is the dimensionality reduction method that achieved the best accuracy, probably due to its intrinsic capability of reducing noise by discarding the less significant eigenvectors. Another important observation is that fusion strategies always increased the performances of single CNNs. In particular, the feature level fusion obtained better results than the single CNNs and decision level fusion.

6.1.2.4 ANALYSIS OF THE PERFORMANCE FOR SIMULATED LESS-CONSTRAINED ACQUISITION SCENARIOS

We evaluated separately the accuracy of the best configuration of the realized method for each non-ideality of AmI-Face Dataset. Table 6.4 summarizes the achieved results. This table shows that the higher performance decreasing corresponded to the scenario in which the users were speaking with a smartphone and the scenario in which the users were appositely performing strong changes in their expression. In all these cases, MAE decreased of 0.46 years with respect to acquisitions performed with frontal face and neutral expression. This performance decreasing can be considered as satisfactory for age estimation methods working in unconstrained scenarios. It is also interesting to note that MAE decreased only of 0.46 years for acquisition performed with head rotations of 90° with respect to frontal acquisitions. This result is particularly promising since most of the methods in the literature are designed to work with frontal acquisitions and have not been evaluated with this kind of strong face rotations. The obtained performance suggest that groups of CNNs trained using big and heterogeneous datasets can extract discriminative features robust to a non-ideal application conditions.

6.1.2.5 COMPARISON WITH RECENT METHODS IN THE LITERATURE

To evaluate the performance of the realized method in general scenarios and compare our results with the state of the art methods, we used the face images of Adience Benchmark Dataset. We adopted the 5-fold cross-validation procedure suggested in [83] and evaluated the performance in term of exact classification accuracy (across all age groups).

Table 6.5 compares the accuracy of the best configuration of the studied method with that of state-of-the-art techniques for Adience Benchmark Dataset. For this scenario,

Table 6.5: Results in the literature for age classification on Adience benchmark dataset using classification accuracy

Method	Year	Description	Classification accuracy (%)
[83]	2014	Dropout SVM + LBP	45.10
[226]	2015	Deep CNN	50.70
[321]	2017	Deep CNN	61.30
[322]	2017	Deep features (DeepID2)	51.10
[323]	2017	Feedforward attention mechanism	61.80
[303]	2016	Pre-trained CNN + fine tuning	57.90
[304]	2016	Pre-trained CNN + fine tuning	64.00
[305]	2016	Pre-trained CNN + fine tuning	52.88
Proposed method for age estimation	2017	Multiple pre-trained CNNs	58.49

Tests performed using Adience benchmark dataset with 5-fold cross validation method. We used the parameters of PCA estimated using WIKI Dataset.

the best configuration consisted of the feature level fusion applied to a set of 500 values obtained by applying PCA to the feature sets extracted using VGG-16 CNN and AlexNet CNN.

Table 6.5 shows that the studied method obtained better or comparable classification accuracy with respect to state-of-the-art methods based on specifically designed features or on deep networks trained for the considered dataset. Moreover, it shows that the considered method achieved better accuracy with respect to most of the techniques based on the fine-tuning of pre-trained deep networks, thus proving the feasibility of using pre-trained CNNs as generic feature extractors for age estimation.

The evaluation of the studied method based on age estimation is performed by training it on a public dataset and testing on images acquired in a less-constrained scenario. The achieved results show that the studied method achieved satisfactory performance for non-ideal images acquired in unconstrained scenarios. Also, the accuracy of the realized age estimation method is compared with that of state-of-the-art techniques by using a challenging public dataset. The obtained results show that the studied method achieved better or comparable results with respect to the state of the art. Results also demonstrated that CNNs trained on general datasets can obtain satisfactory accuracy for different types of validation images. Furthermore, results proved that pre-trained deep networks can be considered as general feature extractors for age estimation, also without applying onerous fine-tuning techniques.

6.2 METHODS TO IMPROVE DEPLOYED BIOMETRIC TECHNOLOGIES

This section describes the performed evaluation of the studied methods for improving the recognition accuracy of previously deployed biometric technologies in AmI.

In particular, the validation strategies and achieved results of the realized methods for adaptive cohort normalization have been first presented. Then, the experimental procedures and results of the studied methods for multibiometric systems in AmI have been analyzed. Finally, the evaluation procedures and achieved results for the studied multimodal continuous authentication systems in AmI have been discussed.

6.2.1 ADAPTIVE COHORT NORMALIZATION

This section presents the experiments related to the studied methods for enhancing the performance of the previously deployed biometric technologies in AmI applications. The experimental procedures and obtained results of a case study in fingerprint recognition are described.

First, we describe the creation of the training and test datasets. Then, we illustrate the applicability of cohort normalization methods in a general scenario. Second, we evaluate the feasibility of the proposed approach for privacy-compliant scenarios. We simulate a deployment performance analysis using the procedure proposed by Frontex [324], which is an evaluation prodecure compliant with privacy protection regulations of European countries. We also present an analysis of the computational time required by the proposed approach. Other score normalization techniques are not compared with cohort normalization methods because techniques based on the analysis of genuine matching scores cannot be applied in many scenarios due to privacy protection regulations on the use of biometric data. Finally, the analysis of the obtained results is presented. The obtained accuracy of the studied methods is evaluated in terms of FMR (False Matching Rate) and FNMR (False Non-Matching Rate). As error measures, EER (Equal Error Rate), and FMR_{1000} (the higher FNMR for FMR $\leqslant 0.1\%$) [175] are considered, and the accuracy of biometric recognition techniques are evaluated by using ROC (Receiver Operating Characteristic) curves.

6.2.1.1 USED DATABASES

Fingerprint samples can be acquired in a wide variety of non-ideal situations, can present poor quality due to acquisition problems and can be captured using different acquisition sensors [176]. To simulate these problems and evaluate the performance of the proposed approach, we used several fingerprint datasets. The used data pertain both to public biometric databases and sets of samples acquired in our laboratory by simulating non-ideal scenarios. All the datasets include images captured with an optical sensor and at a resolution of 500 ppi, according to the ICAO specifications [325, 326].

- *Dataset–A (lab best-case scenario)*: this dataset simulates good-quality acquisitions. We created this dataset in our laboratory by acquiring biometric images using an optical four finger scanner and software currently adopted in real security applications [186]. In particular, we collected 1504 biometric samples from 188 fingers in two situations:

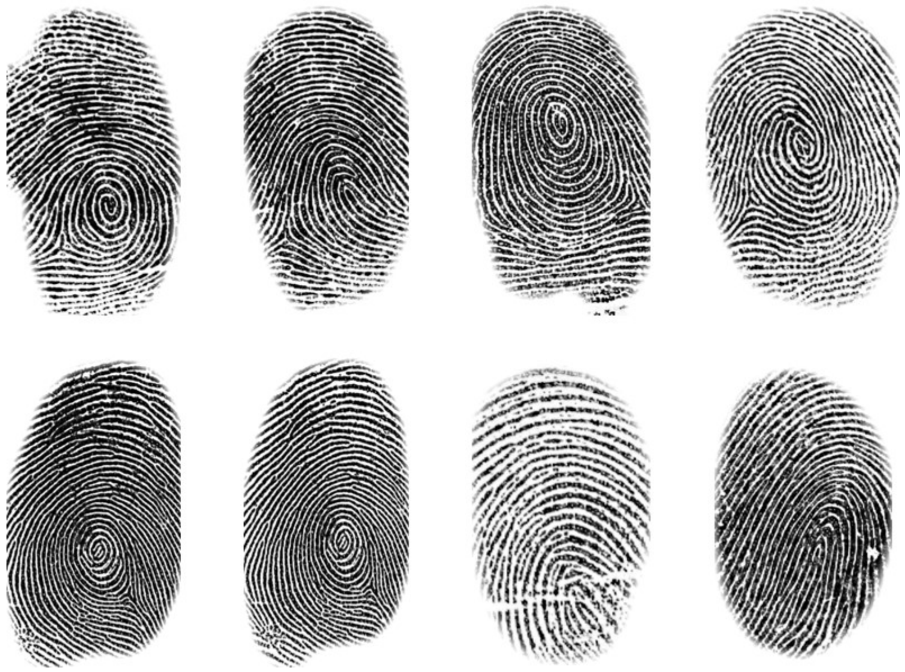


Figure 6.5: Examples of fingerprint images in Dataset–A: lab best-case scenario.

  - 752 images (4 samples per finger). The volunteers were asked to place their fingers on the sensor as they are, with no specific variations in behavioral or environmental conditions.
  - 752 images (4 samples per finger). The volunteers were asked to clean their fingers before performing biometric acquisitions.

  Examples of the fingerprint images in the best-case scenario are shown in Fig. 6.5

- *Dataset–B (lab worst-case scenario)*: this dataset simulates poor-quality acquisitions. We created this dataset in our laboratory by acquiring biometric images using an optical four finger scanner and software currently adopted in real security applications [186]. The images represent poor finger skin conditions or uncomfortable acquisition conditions [119]. In different scenarios the users can carry luggage and their fingers can be dirty after touching dusty, unclean surfaces (e.g., hand rails) or food covered in flour (e.g., donuts, croissants). To simulate these conditions, we collected 1504 biometric samples from 188 fingers in two situations:

  - 752 images (4 samples per finger). To simulate fingertips dirtied by touching dusty, unclean surfaces (e.g., hand rails) or food covered in flour (e.g., donuts, croissants), we acquired the fingerprint samples after dirtying the fingers with flour.
  - 376 images (2 samples per finger). To simulate the grease on the hands typically present after eating fast foods (e.g., sandwiches, mayonnaise, pizza) or using hand creams, we acquired the fingerprint samples after applying a hand cream.

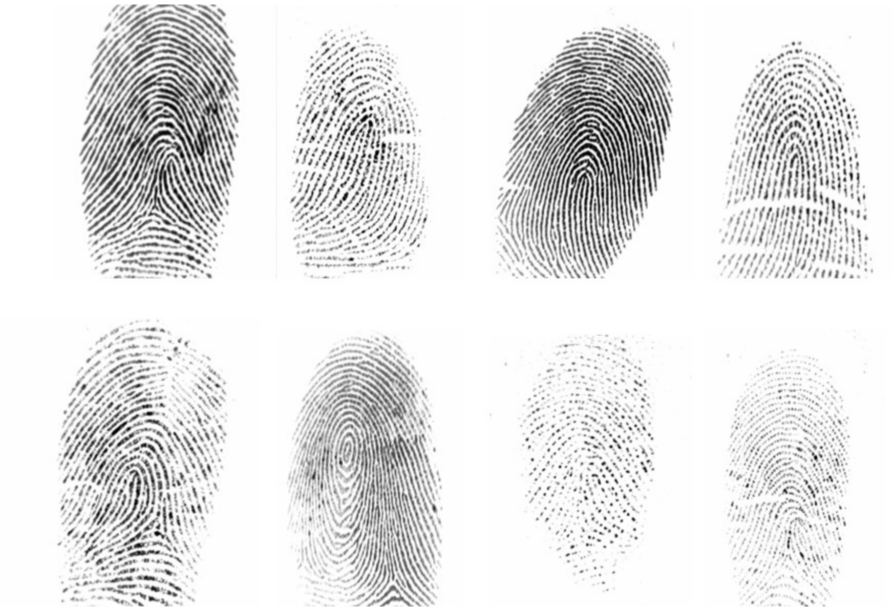


Figure 6.6: Examples of fingerprint images Dataset–B: lab worst-case scenario.

  - 376 images (2 samples per finger). To simulate uncomfortable conditions, we acquired the fingerprints while the users are holding a 4 kg bag on the same shoulder as the finger used for the acquisition.

  Examples of the fingerprint images in the worst-case scenario are shown in Fig. 6.6

- *Dataset–C*: this dataset is composed of fingerprint images collected from a greater number of individuals with respect to Dataset–A and Dataset–B. We used samples belonging to the CASIA Fingerprint Image Database Version 5.0 [327]. We extracted 2000 images by selecting the first two samples of the left and right indexes of all the 500 individuals of the CASIA database. We selected the two indexes of each individual because they are the two fingers most frequently enrolled in e-Passports [58]. Examples of the fingerprint images in Dataset-C is shown in Fig. 6.7

To compute the cohort scores in a manner compliant with privacy laws, we used a set of samples E corresponding to the public database FVC (Fingerprint Verification Database) 2002 DB1 [175], composed of fingerprint samples acquired using a legacy optical sensor with a resolution of 500 ppi. The set is composed of a total of $n = 800$ images acquired from $x = 100$ fingers ($y = 8$ samples per finger).

6.2.1.2 TEST 1: VALIDATION BASED ON A SINGLE DATASET

To prove the applicability of cohort normalization techniques in a general application scenario, we evaluated the performance of different methods by using an iterative validation procedure that uses the samples belonging to a biometric dataset for both the

Figure 6.7: Examples of fingerprint images in dataset-C: Casia Fingerprint Image Database Version 5.0.

training and testing processes [328]. In particular, for each considered dataset (Dataset–A, Dataset–B and Dataset–C), the samples of 50% of the fingers are used for computing the cohort scores, and the remaining samples are used for testing. We computed the training feature set needed by SVM classifiers from the partition used to compute the cohort scores. The evaluation is carried out 10 times, and the results are averaged. Similar procedures are widely used in the literature to validate score normalization methods [328, 243].

Table 6.6 summarizes the results achieved using the considered cohort score normalization methods and the described validation strategy based on a single biometric dataset. This table shows that cohort normalization methods increased the accuracy of the used fingerprint recognition software (baseline) for each considered dataset. Moreover, the performance improved in terms of EER as well as of FMR_{1000}. In particular, the methods based on SVM classifiers achieved the best accuracy for all the performed tests. This result could be due to the generalization capability of SVM classifiers, which allowed to achieve greater robustness to noisy data with respect to the other normalization methods. Nevertheless, the method T–Norm, which does not require a training step, achieved satisfactory results for all the performed tests. As an example, the method SVM–20–cohorts decreased the EER from 1.61% to 1.38% for Dataset–A, from 3.88% to 3.02% for Dataset–B, and from 3.61% to 3.07% for Dataset–C.

6.2.1.3 TEST 2: PRIVACY-COMPLIANT APPROACH

We evaluated the performance of the studied privacy-compliant approach using Dataset–A, Dataset–B and Dataset–C. For each user, the sample set E was used to compute the cohort scores and to train SVM classifiers. It is important to note that the obtained results are not directly comparable to those presented in the previous section because the evaluation procedure is different. In this case, the performance of each cohort normalization method is computed once using all the samples pertaining to the considered datasets.

Table 6.6: Results of the cohort normalization using the validation technique based on a single dataset (2-fold validation iterated 10 times)

Method	Test database (DB)					
	DB–A (lab best-case)		DB–B (lab worst-case)		DB–C (1000 fingers, public DB)	
	EER (%)	FMR$_{1000}$ (%)	EER (%)	FMR$_{1000}$ (%)	EER (%)	FMR$_{1000}$ (%)
Baseline (test 1)	1.61	1.86	3.88	7.31	3.61	7.42
Max–rule	1.46	1.77	3.39	6.54	3.43	6.40
T–norm	1.45	1.84	3.25	6.47	3.08	5.89
SVM–all–cohorts	1.38	1.80	3.16	7.14	3.23	6.96
SVM–20–cohorts	1.38	1.73	3.02	6.55	3.07	6.25

Table 6.7: Accuracy of the studied privacy-compliant approach for cohort normalization using different cohort normalization methods

Method	Test database (DB)					
	DB–A (lab best-case)		DB–B (lab worst-case)		DB–C (1000 fingers, public DB)	
	EER (%)	FMR$_{1000}$ (%)	EER (%)	FMR$_{1000}$ (%)	EER (%)	FMR$_{1000}$ (%)
Baseline (test 2)	1.49	1.71	3.97	7.62	3.59	7.50
Max–rule	1.31	1.63	3.57	6.90	3.27	6.98
T–norm	1.31	1.61	3.46	7.06	3.34	6.60
SVM–all–cohorts	1.22	1.61	3.34	6.87	3.40	6.85
SVM–20–cohorts	1.21	1.59	3.37	7.13	3.42	6.50

Table 6.7 summarizes the achieved results. This table shows that the studied approach increased the accuracy of the used fingerprint recognition software (baseline) for each implemented cohort normalization method in terms of both EER and FMR$_{1000}$. Also, in this case, SVM classifiers achieved the greatest performance improvements. In particular, SVM classifiers achieved the best FMR$_{1000}$ for each evaluated dataset. FMR$_{1000}$ is a particularly relevant figure of merit for evaluating the performance of biometric technologies for high-security applications. As an example, the method SVM–20–cohorts decreased the FMR$_{1000}$ from 1.71% to 1.59% for Dataset–A, the method SVM–all–cohorts decreased the FMR$_{1000}$ from 7.62% to 6.87% for Dataset–B, and the method SVM–20–cohorts decreased the FMR$_{1000}$ from 7.50% to 6.50% for Dataset–C.

Table 6.8: Accuracy of the studied privacy-compliant approach using different cohort normalization methods and the privacy-compliant test methodology proposed by Frontex [324]

Method	**Test database**	
	Dataset–C (1000 fingers, public dataset)	
	EER (%)	FMR_{1000} (%)
Baseline (test 3)	3.69	7.50
Max–rule	3.27	6.50
T–norm	3.17	5.70
SVM–all–cohorts	3.29	6.50
SVM–20–cohorts	3.18	6.00

Fig. 6.8 shows the ROC curves obtained by the studied approach using the cohort normalization methods that achieved the best performance in term of FMR_{1000} for Dataset–A, Fig. 6.9 for Dataset–B and Fig. 6.10 for Dataset–C. The ROC curves show that the realized method increased the accuracy of the used fingerprint recognition software for all the operational points of the biometric system by using SVM classifiers.

In order to prove the statistical significance of the results achieved by the considered approach with respect to the baseline method, we estimated the confidence bounds of the error rates achieved for each curve shown in Fig. 6.8 to Fig. 6.10 by using a method based on the central limit theorem [131] with 95% confidence limits. In particular, we observed that the confidence bounds estimated for the implemented method and for the baseline method present very limited overlapping regions ($FMR < 0.1\%$ for Dataset–A and Dataset–B and $FMR < 1\%$ for Dataset–C). These results confirm that the studied approach can increase the performance of a commercial fingerprint recognition technologies.

The investigated method and realized approach decreased the EER from $1,49\%$ to $1,21\%$ for Dataset–A. Therefore, the studied approach could reduce the number of identity recognition in cases of false non-matches of around 19% when applied to samples of good quality (Dataset–A).

6.2.1.4 TEST 3: PRIVACY-COMPLIANT TESTING

Since the considered approach could be applied in already deployed systems, we also tested its accuracy by simulating a scenario evaluation. This analysis is also useful to illustrate the process that should be followed to evaluate a real deployment, in which it is not possible to perform a performance evaluation using the mostly adopted strategies in the literature.

Scenario / operational evaluations of biometric technologies are difficult processes because privacy protection regulations impose strict limitations in storing samples and templates obtained from biometric documents as well from live acquisitions, thus making difficult to compute traditional figures of merit in an accurate manner.

To simulate a scenario evaluation, we used the privacy-compliant test methodology proposed by Frontex [324]. This procedure allows to store only the last 10 fresh templates to estimate the biometric recognition accuracy of a system. Moreover, it requires that each finger is presented only once to the system to obtain a single genuine score between the fresh sample and the one stored in the biometric document. This scenario evaluation procedure assumes that only genuine attempts of crossing the board are performed.

We simulated this test methodology by implementing a procedure that compares each biometric sample with the last 10 acquired fresh samples, using a first-in first-out (FIFO) structure. Moreover, we used a dataset with two samples per finger (Dataset–C) and considered the first sample as the one enrolled in the biometric document and the second sample as the fresh data. For each simulated access attempt, we computed a single genuine matching score and a maximum of more than 10% impostor identity comparisons (a maximum of 10 impostor matching scores obtained comparing the fresh sample and the samples stored in the FIFO structure, and a maximum of 10 impostor matching scores obtained comparing the sample enrolled in the biometric document and the samples stored in the FIFO structure).

We evaluated the performance of the realized approach using the external dataset E for each considered cohort normalization method. Table 6.8 summarizes the obtained results, confirming that the considered approach can increase the recognition accuracy of the fingerprint recognition software.

6.2.1.5 COMPUTATIONAL TIME

Since in many delpoyed systems it is not possible to store additional biometric data in electronic document or smart cards, the cohort scores should be computed for each access attempt. To validate the feasibility of the studied approach, we evaluated the computational time required by the used commercial matching software [187] and by the SVM classifiers for the computation of the normalized matching score. We performed this test using an Intel Xeon 3.6 GHz with 32 GB of RAM working with Windows 10 and Matlab R2015b. The time required to compute the cohort scores from E (800 identity comparisons) is 0.24 s. The classification time required by the Matlab toolbox for SVM is 0.02 s for SVM–20–cohorts and 0.342 s for SVM–all–cohorts. The obtained results suggest that the studied approach could be used in existing applications with satisfactory performance.

The performed evaluation and obtained results from the case study in fingerprint recognition techniques confirm that the studied method can increase the recognition accuracy of the previously deployed biometric systems in different applications. The considered adaptive cohort normalization technique is general and does not depend on a particular biometric technology. We presented the case study in fingerprint biometrics. However, this approach can be easily extended to any biometric technology. Moreover, the privacy-compliant testing procedure adopted for the evaluation of the

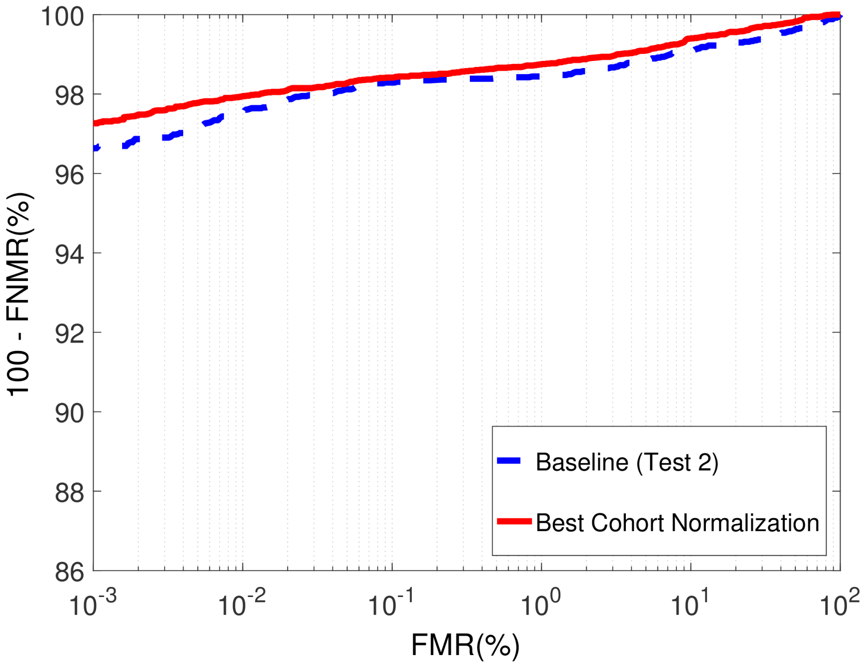


Figure 6.8: ROC curves of the studied privacy-compliant approach using the cohort normalization methods based on SVM–20–cohorts technique (that achieved the best performance in term of FMR_{1000}) for Dataset–A. The higher the values along the vertical axis $(100 - \text{FNMR}(\%))$ are, the better is the accuracy.

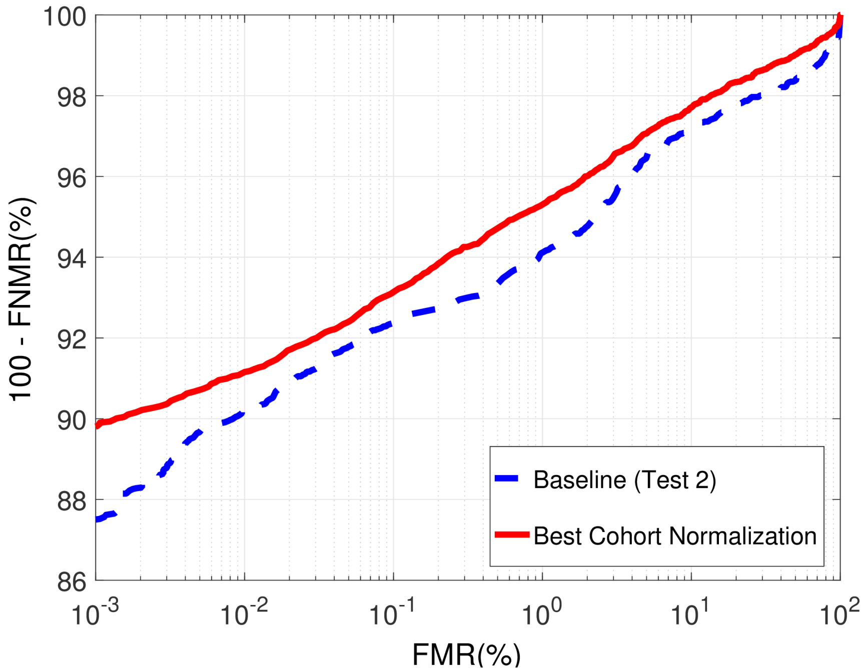


Figure 6.9: ROC curves of the studied privacy-compliant approach using the cohort normalization methods based on SVM–all–cohorts technique (that achieved the best performance in term of FMR_{1000}) for Dataset–B. The higher the values along the vertical axis $(100 - \text{FNMR}(\%))$ are, the better is the accuracy.

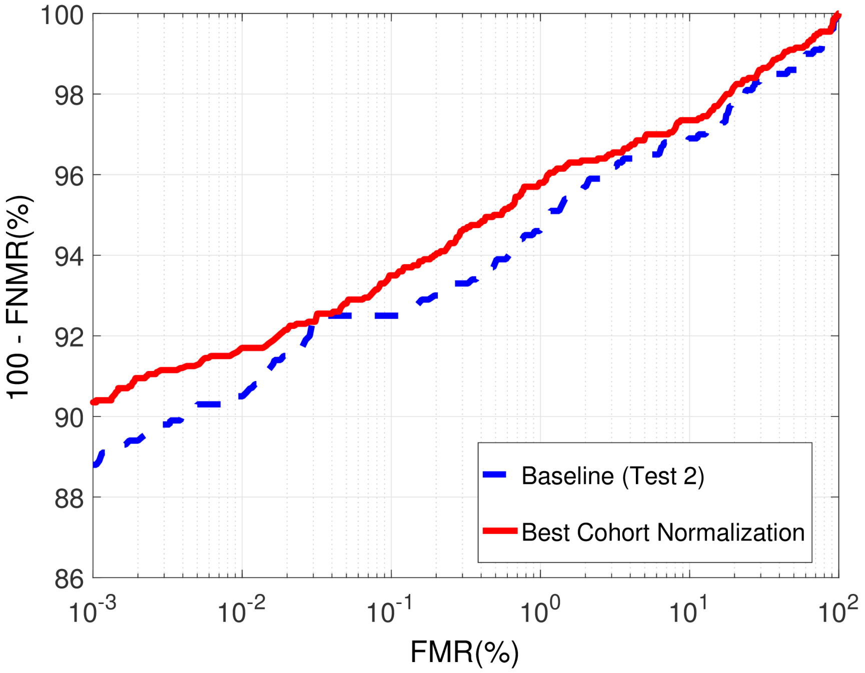


Figure 6.10: ROC curves of the studied privacy-compliant approach using the cohort normalization methods based on SVM–20–cohorts technique (that achieved the best performance in term of FMR_{1000}) for Dataset–C. The higher the values along the vertical axis $(100 - \text{FNMR}(\%))$ are, the better is the accuracy.

implemented techniques corroborate that the studied approach can perform biometric recognition with reduced privacy risks.

6.2.2 MULTIBIOMETRIC SYSTEMS FOR AMBIENT INTELLIGENCE

In this section, we discuss the experimental setup, used datasets, and the obtained results by the studied methods designed for multibiometric systems in AmI. First, we provide the description of the used public datasets and experimental procedures to evaluate the studied methods. The results obtained from the implemented fusion approaches are detailed. Then, the privacy-compliant fusion approaches and technology neutral evaluation of the realized methods are discussed. Finally, we present the analysis of the obtained results.

6.2.2.1 USED DATASETS

To evaluate the performance of multibiometric systems in AmI applications, we collected different datasets by considering biometric samples extracted from public biometric databases, simulating different conditions that can arise in AmI. We considered face and fingerprint databases since they are the most common biometric modalities used in similar applications [329]. All the images have been captured using acquisition devices similar to the ones used in controlled environments. In the case of face databases, we considered both ICAO compliant (good quality) images and non-ICAO compliant (medium-low quality) samples, to obtain a trade-off between the quality of images stored in the database and the quality of fresh images acquired in live environments. In particular, we considered the following databases:

- FEI Face Database [330], containing face images captured using a color camera with a uniform background. We selected 100 individuals, with 8 samples for each individual captured in the same session, for a total of 800 images. A subset of the images is ICAO compliant [331], thus allowing to simulate good quality images, while the rest of the images are more challenging. In particular, these images present challenging aspects that may appear in an AmI scenario, such as variations in the lighting or changes in pose and expression, which can simulate the possibility of a live acquisition in the smart environments where the person is not correctly following the acquisition protocol [332]. Also, even if a uniform background is not always present in real AmI systems, methods for face detection and segmentation have been proved to work also with unconstrained backgrounds in AmI scenarios [333, 334].

- AR Face Database [335], containing face images captured using a color camera with a uniform background. We selected 100 individuals, with 8 samples for each individual captured in two sessions taken 14 days apart, for a total of 800 images. Part of the database is ICAO compliant [331], as images stored in biometric databases should be, whereas other images present some challenges. In particular, this database allows us to simulate other conditions of AmI, such as when the biometric samples stored in the database have been captured before, and differences in make-up or hairstyle can be present during the deployment of the system.

- FVC (Fingerprint Verification Database) 2002 DB1 [175], containing fingerprint samples captured using a medium-quality, legacy optical sensor with a 13.2 × 25 mm sensing area and with 500 ppi resolution. The database is composed by 100 individuals, with 8 samples for each individual, for a total of 800 images. This database permits to simulate AmI applications with samples captured with old equipment.

- FVC (Fingerprint Verification Database) 2006 DB2 [336], containing fingerprint samples captured using a more recent medium-quality optical fingerprint acquisition sensor, with 17.8 × 25 mm sensing area and 500 ppi resolution. We selected 100 individuals, with 8 samples for each individual, for a total of 800 images. Differently, from the FVC 2002 DB1 database, the volunteers included also manual workers and older adults. Moreover, the acquisition procedure did not consider any constraint used for increasing the quality of the captured samples. This database allows simulating people with all kinds of ages, jobs, and familiarity with technology. Moreover, the fingerprint images captured in non-ideal situations simulate the possibility of people residing in AmI with fingers swollen, dirty, or greasy from the travel [119].

Then, using the four databases, we created two scenarios: *Scenario 1*, using FEI for face and FVC 2002 DB1 for fingerprint (Fig. 6.11); *Scenario 2*, using AR for face and FVC

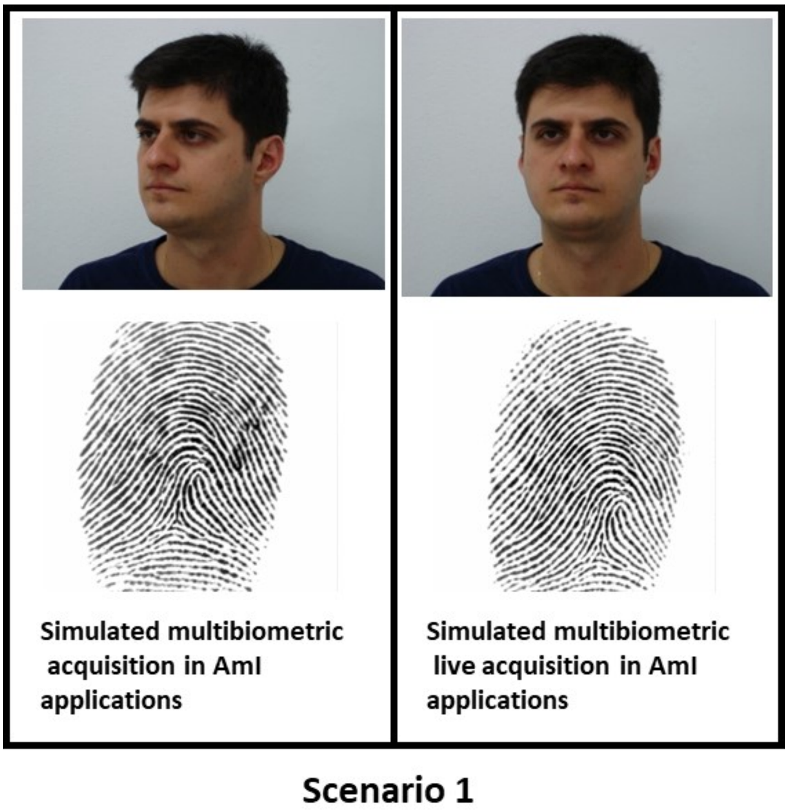

Simulated multibiometric acquisition in AmI applications
Simulated multibiometric live acquisition in AmI applications
Scenario 1

Figure 6.11: Simulated multibiometric systems for scenario 1.

2006 DB2 for fingerprint (Fig. 6.12). Moreover, to recreate the operational conditions of security applications, we applied the compression techniques described by the ICAO for storing biometric samples in the database [331]. In particular, we used the WSQ compression to produce fingerprint samples with $\approx 10\,\mathrm{kB}$ file size, and the JPG compression to produce face images with ≈ 90 pixels between the eyes and $\approx 15-20\,\mathrm{kB}$ file size.

6.2.2.2 experimental procedure

We used the biometric recognition software Cognitec FaceVACS v9.1.1.0 and Dermalog Fingercode3 v1.2.1613.13 to compute and match the templates from the face and fingerprint images, respectively. In both scenarios, we performed a scenario evaluation [337] for all the fingerprint and face databases, separately. For each database, the evaluation included 5600 genuine comparisons and 633600 impostor comparisons. We considered as error metrics the EER and the FMR_{1000} (the lowest FNMR for $\mathrm{FMR} \leqslant 0.1\%$).

Then, for each scenario, we performed the score-level fusion using the sum, product, max, min, weighted sum using Fisher's rule [338], NCW rule [147], MEW rule [339], OLD rule [339], likelihood ratio, and quality-based likelihood ratio [104].

The training of the likelihood ratio methods was performed on a random subset containing 50% of genuine scores and 50% of the impostor scores, and tested on the

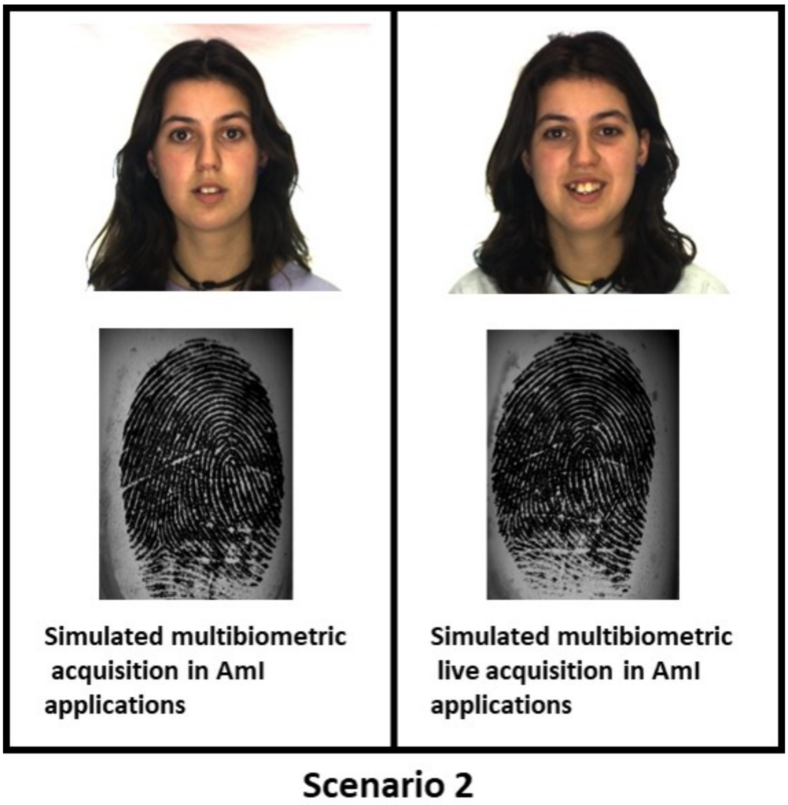

Simulated multibiometric acquisition in AmI applications
Simulated multibiometric live acquisition in AmI applications
Scenario 2

Figure 6.12: Simulated multibiometric systems for scenario 2.

remaining scores. The procedure was repeated 10 times, then the average FMR and FNMR were used to compute the error metrics [104]. We tested the privacy-compliant biometric fusion technique that can be applied in AmI systems by performing the training and the test using two different datasets. Moreover, a technology-neutral evaluation was performed by considering biometric recognition algorithms produced by different vendors.

6.2.2.3 RESULTS OF SCORE-LEVEL FUSION

The results for the Scenario 1 and Scenario 2 are reported in Table 6.9. In both scenarios, it is possible to observe that learning-based methods using the likelihood ratio obtained the best results in terms of EER and the FMR_{1000}, independently from the used normalization technique. Moreover, Table 6.9 shows that the sum rule allowed to obtain high accuracy in terms EER and FMR_{1000}, similar to the one obtained using learning-based methods, but required a preliminary Z-Score normalization to obtain the best results.

6.2.2.4 PRIVACY-COMPLIANT FUSION

To test the accuracy of the privacy-compliant score-level fusion, the scores obtained in Scenario 1 were used to train the likelihood ratio fusion model, which was then tested

Table 6.9: Score-level fusion results

		Scenario 1						Scenario 2					
		Normalization method						Normalization method					
Ref.	Fusion	No norm.		Min-max		Z-Score		No norm.		Min-max		Z-Score	
		EER (%)	FMR 1000 (%)	EER (%)	FMR 1000 (%)	EER (%)	FMR 1000 (%)	EER (%)	FMR 1000 (%)	EER (%)	FMR 1000 (%)	EER (%)	FMR 1000 (%)
-	Face	1.55	6.78	1.55	6.78	1.55	6.78	4.42	17.53	4.42	17.53	4.42	17.53
-	Fingerprint	1.14	1.71	1.14	1.71	1.14	1.71	0.82	1.07	0.82	1.07	0.82	1.07
-	Sum	1.00	1.60	0.17	0.42	0.09	0.07	0.64	1.07	0.28	0.60	0.21	0.21
-	Product	0.32	0.64	0.28	0.53	0.47	0.64	1.25	2.42	1.10	2.10	0.21	0.21
-	Max	1.14	1.71	0.45	6.75	0.21	0.39	0.82	1.07	0.57	1.46	0.37	0.75
-	Min	1.55	6.78	0.99	4.60	1.00	1.75	4.42	17.53	2.35	4.35	2.21	4.75
	Weighted sum												
[338]	Fisher	0.10	0.10 ^{a}	0.10	0.10 ^{b}	0.10	0.10 ^{c}	0.17	0.21 ^{d}	0.17	0.21 ^{e}	0.17	0.21 ^{f}
[147]	NCW	0.10	0.10 ^{g}	0.10	0.10 ^{h}	0.10	0.10 ^{i}	0.17	0.21 ^{j}	0.17	0.21 ^{k}	0.17	0.21 ^{l}
[339]	MEW	0.10	0.10 ^{m}	0.10	0.10 ^{n}	0.10	0.10 ^{o}	0.21	0.21 ^{p}	0.21	0.21 ^{q}	0.21	0.21 ^{r}
[339]	OLD	0.42	0.60 ^{s}	0.42	0.60 ^{t}	0.42	0.60 ^{u}	0.57	0.89 ^{x}	0.56	0.89 ^{y}	0.56	0.89 ^{z}
[314]	Dempster-Shafer theory*	-	-	0.11	0.16 $^{\nu}$	-	-	-	-	0.17	0.20 $^{\nu}$	-	-
[104]	Likelihood ratio	0.09	0.08	0.07	0.07	0.09	0.09	0.17	0.19	0.21	0.22	0.19	0.20
[104]	Quality-based likelihood ratio	0.07	0.07	0.07	0.07	0.07	0.07	0.13	0.14	0.11	0.12	0.10	0.10

* The Dempster-Shafer theory requires normalizing match scores between 0 and 1. Therefore, No norm. and Z-Score are not applicable for this fusion strategy.

$^{a}w = (0.98, 0.02)$; $^{b}w = (0.36, 0.64)$; $^{c}w = (0.49, 0.51)$; $^{d}w = (0.95, 0.05)$; $^{e}w = (0.25, 0.75)$; $^{f}w = (0.31, 0.69)$
$^{g}w = (0.28, 0.01)$; $^{h}w = (0.08, 0.21)$; $^{i}w = (0.11, 0.17)$; $^{j}w = (0.27, 0.01)$; $^{k}w = (0.07, 0.18)$; $^{l}w = (0.09, 0.16)$
$^{m}w = (0.28, 0.01)$; $^{n}w = (0.09, 0.19)$; $^{o}w = (0.13, 0.15)$; $^{p}w = (0.28, 0.01)$; $^{q}w = (0.11, 0.16)$; $^{r}w = (0.14, 0.14)$
$^{s}w = (0.26, 0.02)$; $^{t}w = (0.03, 0.26)$; $^{u}w = (0.04, 0.24)$; $^{x}w = (0.20, 0.08)$; $^{y}w = (0.01, 0.27)$; $^{z}w = (0.01, 0.26)$
$^{\nu}w = (0.3, 0.5)$

Table 6.10: Privacy-compliant score-level fusion results using the quality-based likelihood ratio

	Test scenario			
	Scenario 1		Scenario 2	
Train scenario	EER (%)	FMR_{1000} (%)	EER (%)	FMR_{1000} (%)
Scenario 1	0.07	0.07	0.16	0.18
Scenario 2	0.26	0.30	0.10	0.10

on the scores achieved in Scenario 2, and vice versa. A preliminary Z-Score normalization was used. The results are reported in Table 6.10, showing that the recognition accuracy was not significantly affected when the fusion model is trained on different datasets, thus demonstrating that it is possible to perform an off-line training of the fusion model in AmI applications even with data captured in a different context (e.g., public datasets).

Table 6.11: Improvement of EER and FMR_{1000} using the sum fusion, with respect to using only the most accurate biometric trait, in a technology-independent environment, for different combinations of recognition algorithms. Negative values correspond to increase in accuracy

Algorithm		Scenario 1		Scenario 2	
		ΔEER	ΔFMR_{1000}	ΔEER	ΔFMR_{1000}
Face	Fingerprint	(%)	(%)	(%)	(%)
Cognitec FaceVACS v9.1.1.0	Dermalog Fingercode3 v1.2.1613.13	-0.14	-0.11	-0.18	-0.00
Cognitec FaceVACS v9.1.1.0	Neurotechnology VeriFinger v6.0	-0.00	-0.00	-0.00 *	-0.00 *
Neurotechnology VeriLook v6.0	Dermalog Fingercode3 v1.2.1613.13	-0.75	-1.07	-0.70	-0.90
Neurotechnology VeriLook v6.0	Neurotechnology VeriFinger v6.0	-0.24	-0.26	-0.00 *	-0.00 *

* EER and FMR_{1000} were already equal to 0 using only the most accurate biometric trait

6.2.2.5 TECHNOLOGY-INDEPENDENT EVALUATION

In this section, we provide a technology-independent evaluation of the score-level fusion performance, by using the different combinations of recognition algorithms from different vendors, and analyzing the improvement in the EER and FMR_{1000} with respect to using only the most accurate biometric trait (the fingerprint). In particular, we used the software Dermalog Fingercode3, Cognitec FaceVACS, Neurotechnology VeriFinger, and Neurotechnology VeriLook. No previous normalizations were performed, and the sum rule was used as fusion method since it does not require any learning process. In all cases, it increased the accuracy of the recognition [102]. For each combination, we evaluated the differences ΔEER and ΔFMR_{1000} obtained by using the sum fusion strategy with respect to using the fingerprint, which were computed as follows:

$$\begin{aligned} \Delta EER &= EER_{sum} - EER_{finger} \quad ; \\ \Delta FMR_{1000} &= FMR_{1000sum} - FMR_{1000finger} \quad . \end{aligned} \tag{6.1}$$

The results are summarized in Table 6.11 for both Scenario 1 and Scenario 2, showing that in all cases the fusion allowed to obtain lesser or equal EER and FMR_{1000} with respect to using only the most accurate biometric trait, independently of the used recognition algorithm in the AmI context we simulated.

The presented experimental procedure, evaluation protocols and the obtained results analyze the feasibility of the studied multibiometric systems in AmI applications. The realized approach for combining multiple biometric traits in AmI is evaluated in general scenarios and also on the techniques based on technology-neutral approaches and privacy-compliant fusion.

The performed technology evaluation of the most commonly used score level fusion techniques shows that recent learning-based methods such as the likelihood ratio obtain the best accuracy. Moreover, the evaluation of the performance of a privacy-compliant score-level fusion using the likelihood ratio, demonstrate that fusion methods can be used to enhance the performance of multibiometric systems in AmI, even when actual data collected using AmI systems is not available.

The discussed technology-independent evaluation of the studied method for multimodal fusion in AmI proved that it is always possible to use score-level fusion to increase the recognition accuracy in multimodal AmI systems, independently of the used recognition algorithm.

6.2.3 MULTIMODAL CONTINUOUS AUTHENTICATION

This section presents the configuration of the experiments, used datasets, and obtained results for the methods for multimodal fusion in continuous authentication systems. First, we provide the description of the collected database. Then, the creation of training and test datasets are detailed. The configuration of the implemented method is outlined. Finally, the performance evaluation of the studied method and analysis of the obtained results are discussed.

6.2.3.1 COLLECTED DATABASE

To the best of our knowledge, there were no available any public multimodal datasets using face and voice biometrics, acquired in unconstrained conditions. To evaluate the studied methods for multimodal continuous authentication, we collected a true multimodal audio-visual dataset simulating less-constrained and non-cooperative acquisitions typical of AmI. It consisted of videos of 180 seconds collected from 12 individuals in our laboratory. The frame rate of the video files were 30 fps, and the sampling rate of the audio files were 44100 kHz. The size of each frame was 1280 × 720 pixel. The illumination conditions are kept similar to normal conditions in a room. The video and audio have been captured from a distance of 50 cm from the webcam using a Microsoft LifeCam HD-3000. The conditions are simulated when the user is sitting idle, working on the console, or talking on the smartphone. In particular, the images extracted in the period when the user is talking on the phone, show occlusions, sharp rotations and various poses. Examples of the images extracted from the database are shown in Fig 6.13. The figure shows different scenarios while the user is working in the office.

6.2.3.2 CREATION OF THE TRAINING AND TEST DATASETS

- *Enrollment set*: the first 20 seconds of each video have been used for creating the enrollment set. For making the face recognition system, a total of 600 face images have been extracted from the first 20 seconds of the video of each user. The best quality sample for each user is used to enroll the user.

  For the voice recognition system, a 20 second template is extracted for each user and used to enroll the user.

- *Test set*: to extract the face images for the test set, we used a sliding window method. After the first 20 seconds of the video, the 10 seconds samples have been extracted.

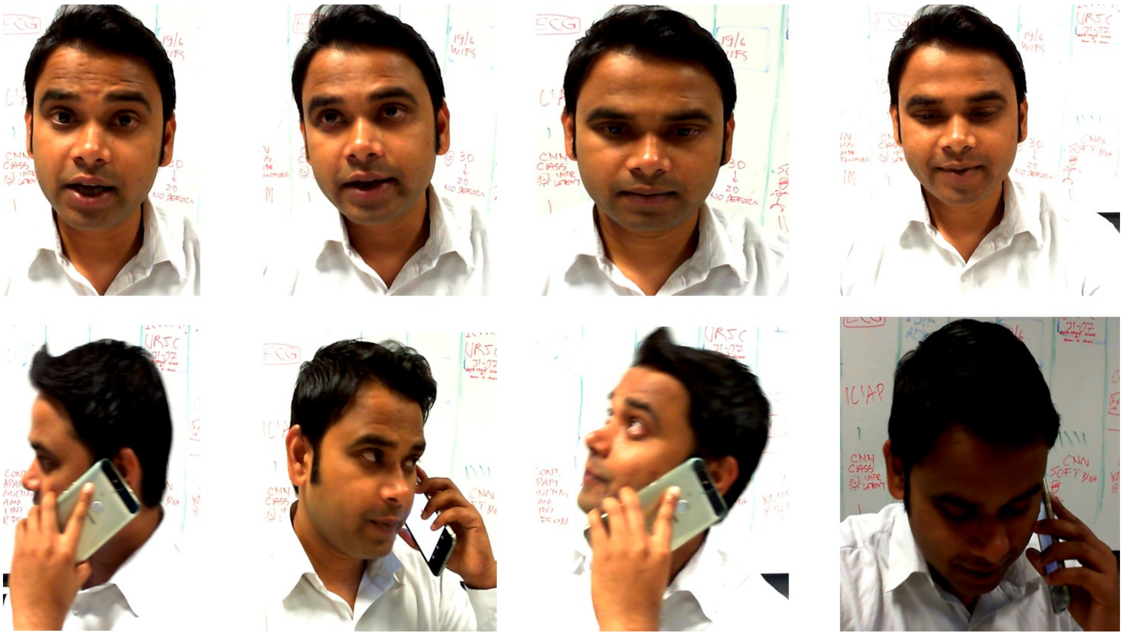


Figure 6.13: Examples of collected multimodal database for continuous authentication systems.

The face images have been selected every second after the 20 seconds of video used for the enrollment. The voice samples have a duration of 10 seconds and have been acquired every second. The acquisition of the voice samples simulates real conditions with a constant buffer of 10 seconds. The audio samples can present an overlap.

In total, we obtained 1800 samples from 12 individuals (150 samples per individual).

6.2.3.3 CONFIGURATION OF THE EXPERIMENTS

A computational intelligence technique based on feed-forward neural networks (FFNN) was used to perform asynchronous multimodal fusion. The considered configurations of FFNN are designed as follows. We used a single linear node for the output layer of the neural network. We tested different numbers of nodes with tan-sigmoidal transfer functions for the hidden layer of the FFNN (in the range from 1 to 10). We trained the neural networks using the Levenberg-Marquardt backpropagation algorithm with 500 epochs. We obtained the best results using 2 nodes in the hidden layer.

The evaluation is carried out 10 times, and the results are averaged. Similar procedures are widely used in the literature [328, 243]. For each iteration, we randomly selected 900 samples to create the training set. The validation set was composed using the remaining sample.

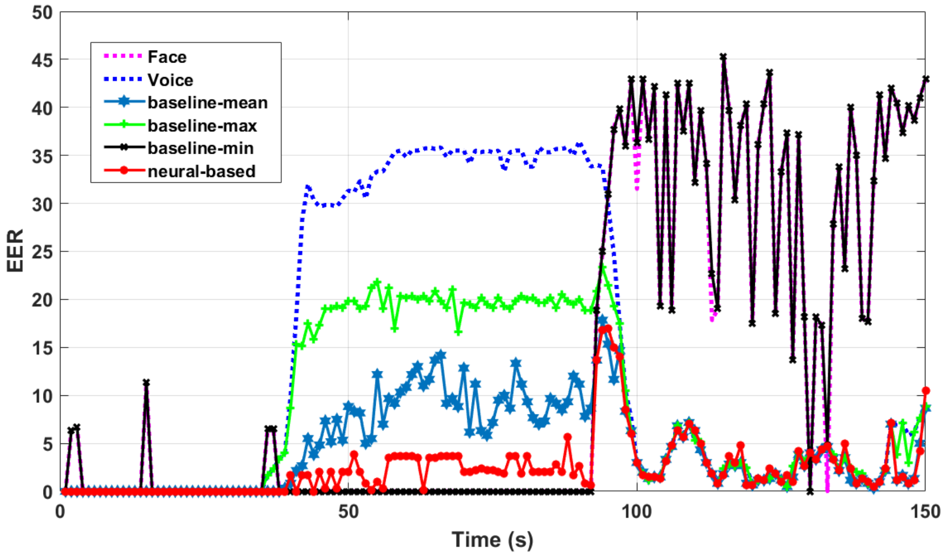


Figure 6.14: EER obtained at different time distances in seconds. In particular, the neural-based method has achieved lowest EER compared to other used fusion approaches. The obtained accuracy of the studied method for multimodal continuous authentication is measured as ERR, with mean equal to 2.38%, and the standard deviation equal to 3.10%

6.2.3.4 Performance evaluation

We evaluated the accuracy of the studied adaptive multimodal fusion techniques for continuous authentication systems. Fig. 6.14 reports the obtained results in terms of EER.

The achieved results of the studied neural-based multimodal fusion method have been compared with three baseline score-level fusion techniques, minimum rule, maximum rule, and mean rule. Fig. 6.14 shows the comparison between the baseline fusion techniques and the studied adaptive neural-based fusion technique. From the figure, it is possible to observe that, in the interval of critical points (from 45s to 150s) in continuous authentication systems, the realized technique able to improve the accuracy of the system, by obtaining lower EER. In particular, the considered technique is able to reduce the EER in the interval of critical points with respect to the compared techniques.

These results are preliminary and the proposed method should be evaluated using further data acquired in different scenarios. Nevertheless, these results are promising and confirm the feasibility of the method.

6.3 Summary

The experimental procedures, evaluation protocols and obtained results of the studied methods for biometric technologies in AmI applications have been described in this chapter.

The studied approaches based on user-friendly technologies for less-constrained human-machine interactions in biometric systems have been analyzed. The results achieved by using evaluation procedures are compared with state-of-the-art techniques. All the considered techniques have shown encouraging results. In particular, the results obtained for the studied methods for text-independent speaker recognition systems demonstrated that the proposed method reduced the size of the template with respect to traditional approaches based on GMM and achieved better identification accuracy. The studied method have been evaluated using reduced numbers of enrolled samples per individual. The achieved results showed that the realized method is capable of achieving better accuracy than the baseline.

The obtained results for the studied methods based on age estimation using non-ideal face images have been analyzed. The realized approach has been evaluated by training it on a public dataset and testing on images acquired in a less-constrained scenario. The analysis of the results showed that the studied method achieved satisfactory performance for non-ideal images acquired in unconstrained scenarios and obtained better or comparable results with respect to state-of-the-art methods. Moreover, the results also demonstrated that the considered technique trained on general datasets can obtain satisfactory accuracy for different types of validation images, thus, improving the generality of the system. Furthermore, results proved that pre-trained deep networks can be considered as general feature extractors for age estimation, also without applying onerous fine-tuning techniques.

Further, we have evaluated the studied methods for improving the recognition accuracy of the previously deployed biometric technologies in AmI.

First, the analysis have been performed on the studied methods for adaptive cohort normalization techniques. A technological and a scenario evaluation of the considered case study is performed by using biometric samples acquired by simulating different non-ideal conditions. For all the performed tests, the realized approach increased the accuracy of the existing biometric systems. The obtained results suggested that the studied techniques can be effectively applied in existing AmI applications in a privacy-compliant manner and without requiring hardware or software modifications.

Then, the studied methods for multibiometric systems for AmI applications have been evaluated. A privacy-compliant evaluation have been performed, which showed that the realized approach can increase the performance of the multibiometric systems in AmI applications, by complying the privacy issues of AmI. The technology-independent evaluation proved that it is always possible to use score-level fusion to increase the recognition accuracy of multimodal systems in AmI, independently of the used recognition algorithm.

Finally, the studied methods for multimodal continuous authentication systems for AmI applications have been analyzed. The considered methods have been tested on real multimodal database acquired in our laboratory simulating less-constrained and non-operative acquisition scenarios. The results showed that the realized approach can perform asynchronous multimodal fusion in a continuous manner. Moreover, the evaluation of the adaptive neural-based fusion method demonstrated that the realized approach achieved better performance with respect to the baseline fusion techniques in less-constrained scenarios, and were able to reduce the EER of the multimodal fusion in the interval of critical points in continuous authentication systems. The obtained results are preliminary and the studied method should be evaluated in different AmI using further data. Nevertheless, the achieved results are promising and confirm the feasibility of the studied method.

7

CONCLUSION AND FUTURE WORKS

7.1 CONCLUSION

The objectives of this thesis were the study and implementation of innovative approaches to improve the human-computer interaction in Ambient Intelligence (AmI) by using biometrics as enabling technologies to design personalized services for individuals or classes of people.

In this context, the first contribution of this thesis consists of proposing innovative less-constrained technologies able to increase the applicability of biometric systems in AmI and improve the quality of the human-computer interaction in different AmI scenarios, with respect to the state-of-the-art technologies. The realized approaches include biometric technologies based on less-constrained and non-cooperative acquisitions to facilitate the interaction between the users and the systems in AmI. With respect to the state-of-the-art, the novelty of the studied approaches resides in the fact that we considered less-constrained acquisition scenarios and non-cooperative users to increases the applicability of biometric technologies in AmI. The studied innovative methods for less-constrained biometric systems consist of a text-independent speaker identification system and a technique for age estimation based on facial images. These methods have been designed to be robust to different kinds of non-idealities typical of AmI scenarios and showed advantages with respect to state-of-the-art techniques. In particular, the realized text-independent speaker identification method achieved better accuracy with respect to traditional biometric systems for small databases of samples acquired in environmental conditions typical of AmI environments. Moreover, the size of the templates used by our method was significantly lower with respect to that of state-of-the-art methods, thus permitting its use in scenarios in which the biometric data are stored in electronic documents or smart cards. The implemented age estimation method based on deep learning techniques achieved satisfactory accuracy for non-ideal face images acquired in unconstrained scenarios typical of AmI applications.

Furthermore, this method was faster and easier to tune in different application conditions with respect to the approaches in the literature based on deep neural networks.

A further contribution consists of the study and realization of novel approaches to improve the applicability and integration of heterogeneous biometric systems in AmI. The performed studies regard original methods for novel and comprehensive biometric systems able to deal with heterogeneous traits, sensors, and environmental conditions. The novelty of the studied approaches consists of designing novel multimodal biometric approaches that take advantage from all the sensors placed in a generic environment in order to achieve high recognition accuracy and to permit to perform continuous or periodic authentications in an unobtrusive manner. The realized approaches have been designed to improve the recognition accuracy of the already deployed biometric technologies based on single and multiple biometric technologies. These approaches consist of a privacy-compliant cohort normalization technique, a technology-independent multimodal fusion strategy, and a multimodal continuous authentication system for AmI applications. These innovative approaches showed advantages with respect to state-of-the-art techniques. In particular, the cohort normalization approach achieved better accuracy with respect to baseline technologies by providing effective privacy protection mechanisms that permit its application in a wide set of countries. Furthermore, its performance was comparable to that of methods in the literature that do not consider privacy concerns. The implemented technology-independent multimodal fusion approach increased the accuracy of traditional biometric systems for low-quality samples typical of AmI scenarios. Moreover, a technology-independent evaluation of the considered method proved that the implemented score-level fusion approach increased the biometric accuracy of traditional systems for every evaluated combination of biometric recognition techniques and traits. The studied multimodal continuous authentication approach proved to effectively handle the asynchronous behavior of the multimodal biometric data and to be able to deal with the uncertainties in the recognition scores of continuous authentication systems. The obtained results were promising and confirmed the feasibility of the method. However, further studies should be performed to prove the effectiveness of this approach in different application scenarios.

The realized novel technologies have been tested on different biometric datasets (both public and collected in our laboratory) simulating acquisitions performed in AmI applications. The achieved results proved the feasibility of the studied approaches and shown that the studied methods effectively increased the accuracy, applicability, and usability of biometric technologies in AmI with respect to the state-of-the-art.

7.2 FUTURE WORKS

The thesis improved the state of the art for biometrics in AmI applications by proposing user-friendly and uncooperative technologies and by presenting novel approaches for increasing the accuracy of biometric systems already deployed in AmI scenarios. Nevertheless, the performed research could be expanded in several directions, including the design of novel techniques to extract and use soft biometric data, novel algorithms robust to non-ideal acquisitions for biometric recognition technologies usable in non-cooperative scenarios, techniques to select the more readable biometric traits according to the sensors deployed in a specific AmI environment, and approaches to integrate user-friendly methods for human-computer interaction with biometric technologies.

More details on possible future research topics are provided in the following:

- Novel methods to extract and process soft biometric information from non-ideal samples should be designed. These methods should extract information from multi-dimensional signals acquired in non-cooperative AmI scenarios. Moreover, soft biometric data should permit to classify the users according to sets of preferences related to physiological or cultural factors. Examples of multi-dimensional signals that could be used in this context are the face images, voice samples, and surveillance videos. Examples of soft biometric information are the gender, height, width, and ethnicity.

- Innovative methods for biometric recognition based on samples acquired in less-constrained conditions with respect to the literature should be designed to increase the possible application scenarios of biometric technologies in AmI. Examples of these biometric technologies are face recognition methods for mobile devices, technologies based on touchless fingerprint images, and biometric technologies for iris recognition based on samples acquire at a high distance from the sensor and in natural light conditions.

- Innovative techniques to select the most reliable biometric traits according to the sensors deployed in a specific AmI application should be studied. Novel techniques able to adapt a multibiometric system to the hardware configuration of different AmI scenarios should simplify the deployment of biometric technologies as well as increase the recognition accuracy of current technologies by considering the biggest amount of information available for each environment.

- Approaches to integrate user-friendly methods for human-computer interaction with biometric technologies should improve the usability of AmI technologies by providing integrated frameworks to evaluate different aspects of multi-dimensional sensors. Examples of technologies to be further studied are systems to analyze voice signals in terms of speaker as well as speech recognition, methods to infer the emotion and identity from face images, and techniques to infer the identity of the user from a frame sequence from the analysis of gestures.

- Biometric datasets composed of large numbers of samples, also acquired in AmI scenarios different from the ones considered in this thesis, should be collected. Examples of technologies to be further studied are face recognition systems based on samples of poor quality acquired from surveillance cameras and voice recognition systems based on samples acquired in noisy conditions such as background noise (television or radio sounds) or under stress conditions.

REFERENCES

[1] A. Jain, P. Flynn, and A. A. Ross, *Handbook of biometrics*. Springer Science & Business Media, 2007. *(Cited on pages XVII, 6, 7, 27, 28, 32, 46, 50, 51, and 72)*

[2] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on circuits and systems for video technology*, vol. 14, no. 1, pp. 4–20, 2004. *(Cited on pages 6 and 7)*

[3] J. Neves, F. Narducci, S. Barra, and H. Proença, "Biometric recognition in surveillance scenarios: a survey," *Artificial Intelligence Review*, vol. 46, no. 4, pp. 515–541, 2016. *(Cited on pages 7 and 46)*

[4] M. Tistarelli, S. Z. Li, and R. Chellappa, *Handbook of remote biometrics*. Springer, 2009, vol. 1. *(Cited on pages 7 and 46)*

[5] M. Tistarelli and B. Schouten, "Biometrics in ambient intelligence," *Journal of Ambient Intelligence and Humanized Computing*, vol. 2, no. 2, pp. 113–126, 2011. *(Cited on pages 7, 46, 47, 49, and 53)*

[6] V. Piuri, "Biometric technologies for ambient intelligence in the internet of things," in *Green Computing and Communications (GreenCom), 2013 IEEE and Internet of Things (iThings/CPSCom), Proceedings of the IEEE International Conference on IEEE Cyber, Physical and Social Computing*. IEEE, 2013, pp. lxxi–lxxii. *(Cited on page 7)*

[7] J. Wayman, A. Jain, D. Maltoni, and D. Maio, "An introduction to biometric authentication systems," *Biometric Systems*, pp. 1–20, 2005. *(Cited on page 7)*

[8] R. Donida Labati, V. Piuri, and F. Scotti, *Touchless fingerprint biometrics*. CRC Press, 2015. *(Cited on pages XIII, 7, 36, 46, 47, 51, and 60)*

[9] H. Bouma, "True visions–the emergence of ambient intelligence, by e. aarts, jl encarnação; 2006," *Gerontechnology*, vol. 6, no. 1, pp. 58–60, 2007. *(Cited on page 11)*

[10] K. Ducatel, M. Bogdanowicz, F. Scapolo, J. Leijten, and J.-C. Burgelman, "Ambient intelligence: From vision to reality," *IST Advisory Group Draft Report, European Commission*, 2003. *(Cited on pages 11 and 25)*

[11] A. K. Dey, "Understanding and using context," *Personal and ubiquitous computing*, vol. 5, no. 1, pp. 4–7, 2001. *(Cited on page 12)*

[12] J. Bravo, R. Hervás, and M. Rodriguez, *Ambient Assisted Living and Home Care*. Springer, 2012, vol. 7657. *(Cited on pages 19 and 20)*

[13] M. Friedewald, O. Da Costa, Y. Punie, P. Alahuhta, and S. Heinonen, "Perspectives of ambient intelligence in the home environment," *Telematics and informatics*, vol. 22, no. 3, pp. 221–238, 2005. *(Cited on page 19)*

[14] T. Kirishima, K. Sato, and K. Chihara, "Real-time gesture recognition by learning and selective control of visual interest points," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 351–364, 2005. *(Cited on page 19)*

[15] A. Pentland and T. Choudhury, "Face recognition for smart environments," *IEEE Transactions on Computer*, vol. 33, no. 2, pp. 50–55, 2000. *(Cited on pages 19 and 50)*

[16] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotion recognition in human-computer interaction," *IEEE Signal Processing Magazine*, vol. 18, no. 1, pp. 32–80, 2001. *(Cited on pages 19, 20, and 50)*

[17] A. C. Correia, L. Miranda, and H. Hornung, "Gesture-based interaction in domotic environments: state of the art and hci framework inspired by the diversity," in *Proceedings of the 14th International Conference on Human-Computer Interaction (INTERACT)*. Springer, 2013, pp. 300–317. *(Cited on page 19)*

[18] F. Portet, M. Vacher, C. Golanski, C. Roux, and B. Meillon, "Design and evaluation of a smart home voice interface for the elderly: acceptability and objection aspects," *Personal and Ubiquitous Computing*, vol. 17, no. 1, pp. 127–144, 2013. *(Cited on pages 19 and 68)*

[19] K.-N. Kim and R. Ramakrishna, "Vision-based eye-gaze tracking for human computer interface," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, vol. 2. IEEE, 1999, pp. 324–329. *(Cited on page 20)*

[20] G. Acampora, D. J. Cook, P. Rashidi, and A. V. Vasilakos, "A survey on ambient intelligence in healthcare," *Proceedings of the IEEE*, vol. 101, no. 12, pp. 2470–2494, 2013. *(Cited on pages 20 and 21)*

[21] B. Gates, N. Myhrvold, P. Rinearson, and D. Domonkos, "The road ahead," 1995. *(Cited on pages 20 and 24)*

[22] D. Cook and S. Das, *Smart environments: Technology, protocols and applications*. John Wiley & Sons, 2004, vol. 43. *(Cited on page 20)*

[23] I. Marsá Maestre, E. d. l. Hoz, B. Alarcos, and J. R. Velasco Pérez, "A hierarchical agent-based approach to security in smart offices," 2006. *(Cited on page 20)*

[24] D. Minder, P. J. Marrón, A. Lachenmann, and K. Rothermel, "Experimental construction of a meeting model for smart office environments," in *Proceedings of the Workshop on Real-World Wireless Sensor Networks, SICS Technical Report*, 2005, p. 09. *(Cited on page 20)*

[25] *International Organization for Standardization 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs): Part 11: Guidance on Usability*, 1998. *(Cited on pages 20, 37, and 47)*

[26] A. F. Abate, F. Narducci, and S. Ricciardi, "Biometrics empowered ambient intelligence environment," *Atti della Accademia Peloritana dei Pericolanti-Classe di Scienze Fisiche, Matematiche e Naturali*, vol. 93, no. 2, p. 4, 2015. *(Cited on pages 20 and 50)*

[27] C. Le Gal, J. Martin, A. Lux, and J. L. Crowley, "Smartoffice: Design of an intelligent environment," *IEEE Intelligent Systems*, vol. 16, no. 4, pp. 60–66, 2001. *(Cited on pages 20 and 24)*

[28] B. Johanson, A. Fox, and T. Winograd, "The interactive workspaces project: Experiences with ubiquitous computing rooms," *IEEE pervasive computing*, vol. 1, no. 2, pp. 67–74, 2002. *(Cited on page 20)*

[29] V. Stanford, J. Garofolo, O. Galibert, M. Michel, and C. Laprun, "The nist smart space and meeting room projects: signals, acquisition annotation, and metrics," in *Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP'03)*, vol. 4. IEEE, 2003, pp. IV–736. *(Cited on pages 20 and 24)*

[30] S. A. Velastin, B. A. Boghossian, B. P. Lo, J. Sun, and M. A. Vicencio-Silva, "Prismatica: toward ambient intelligence in public transport environments," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 35, no. 1, pp. 164–182, 2005. *(Cited on page 21)*

[31] A. Rakotonirainy and R. Tay, "In-vehicle ambient intelligent transport systems (i-vaits): Towards an integrated research," in *Proceedings of the 2004 7th IEEE International Conference on Intelligent Transportation Systems*. IEEE, 2004, pp. 648–651. *(Cited on page 21)*

[32] Y. Matsumotot, T. Ino, and T. Ogsawara, "Development of intelligent wheelchair system with face and gaze based interface," in *Proceedings of the 2001 10th IEEE International Workshop on Robot and Human Interactive Communication, 2001*. IEEE, 2001, pp. 262–267. *(Cited on page 21)*

[33] F. Lewis, "Smart environments: technologies, protocols, and applications," *Wireless Sensor Networks. John Wiley*, 2004. *(Cited on page 21)*

[34] J. Krumm and E. Horvitz, "Predestination: Inferring destinations from partial trajectories," *UbiComp 2006: Ubiquitous Computing*, pp. 243–260, 2006. *(Cited on pages 21 and 24)*

[35] J. Letchner, J. Krumm, and E. Horvitz, "Trip router with individualized preferences (trip): Incorporating personalization into route planning," in *Proceedings of the National Conference on Artificial Intelligence*, vol. 21, no. 2. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006, p. 1795. *(Cited on page 21)*

[36] E. Frank Lopresti, A. Mihailidis, and N. Kirsch, "Assistive technology for cognitive rehabilitation: State of the art," *Neuropsychological rehabilitation*, vol. 14, no. 1-2, pp. 5–39, 2004. *(Cited on page 21)*

[37] A. Mihailidis, J. C. Barbenel, and G. Fernie, "The efficacy of an intelligent cognitive orthosis to facilitate handwashing by persons with moderate to severe dementia," *Neuropsychological Rehabilitation*, vol. 14, no. 1-2, pp. 135–171, 2004. *(Cited on page 21)*

[38] S. Giroux and H. Pigot, *From Smart Homes to Smart Care: ICOST 2005, 3rd International Conference on Smart Homes and Health Telematics*. IOS Press, 2005, vol. 15. *(Cited on pages 22 and 24)*

[39] S. Marzano, "People as a source of breakthrough innovation," *Design Management Review*, vol. 16, no. 2, pp. 23–29, 2005. *(Cited on page 22)*

[40] "Ulster community and hospitals trust, interim poicy on research governance process, 2005." *(Cited on page 22)*

[41] D. J. Cook, J. C. Augusto, and V. R. Jakkula, "Ambient intelligence: Technologies, applications, and opportunities," *Pervasive and Mobile Computing*, vol. 5, no. 4, pp. 277–298, 2009. *(Cited on page 22)*

[42] M. Greenwald *et al.*, "These four walls... americans 45+ talk about home and community," *AARP, May*, 2003. *(Cited on page 22)*

[43] S. Mitchell, M. D. Spiteri, J. Bates, and G. Coulouris, "Context-aware multimedia computing in the intelligent hospital," in *Proceedings of the 2000 9th workshop on ACM SIGOPS European workshop: beyond the PC: new challenges for the operating system*. ACM, 2000, pp. 13–18. *(Cited on page 22)*

[44] M. Böhlen and H. Frei, "Ambient intelligence in the city overview and new perspectives," in *Handbook of ambient intelligence and smart environments*. Springer, 2010, pp. 911–938. *(Cited on page 22)*

[45] K. A. Paskaleva, "The smart city: A nexus for open innovation?" *Intelligent Buildings International*, vol. 3, no. 3, pp. 153–171, 2011. *(Cited on page 22)*

[46] "Hydra: A robust and self managing video sensing system for retrospective surveillance, defense intelligence agency dia-masint, pi surendar chandra, co-pi pat flynn, university at notre dam," [Online]. Available:, http://www.cse.nd.edu/ csesys/hydra/, accessed: August 21, 2017. *(Cited on pages 22 and 24)*

[47] "The disappearing computer, eu-funded initiative on the future and emerging technologies activity of the information society technologies (ist) research program," [Online]. Available:, http://www.disappearing-computer.net/, accessed: August 21, 2017. *(Cited on page 22)*

[48] A. Taylor, B. Donovan, Z. Foley-Fisher, and C. Strohecker, "Time, voice, and joyce," in *Proceedings of the 2004 1st ACM workshop on Story representation, mechanism and context*. ACM, 2004, pp. 67–70. *(Cited on page 22)*

[49] A. Galloway, "Intimations of everyday life: Ubiquitous computing and the city," *Cultural studies*, vol. 18, no. 2-3, pp. 384–408, 2004. *(Cited on page 22)*

[50] L. Gaye, L. E. Holmquist, and R. Mazé, "Sonic city: merging urban walkabouts with electronic music making," *Companion of UIST*, 2002. *(Cited on pages 22 and 24)*

[51] A. Waibel, T. Schultz, M. Bett, M. Denecke, R. Malkin, I. Rogina, R. Stiefelhagen, and J. Yang, "Smart: The smart meeting room task at isl," in *Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP'03)*, vol. 4. IEEE, 2003, pp. IV–752. *(Cited on page 24)*

[52] W. Weber, J. Rabaey, and E. H. Aarts, *Ambient intelligence*. Springer Science & Business Media, 2005. *(Cited on pages 24 and 25)*

[53] M. Friedewald, E. Vildjiounaite, Y. Punie, and D. Wright, "Privacy, identity and security in ambient intelligence: A scenario analysis," *Telematics and Informatics*, vol. 24, no. 1, pp. 15–29, 2007. *(Cited on pages 24 and 25)*

[54] P. Brey, "Freedom and privacy in ambient intelligence," *Ethics and Information Technology*, vol. 7, no. 3, pp. 157–166, 2005. *(Cited on page 24)*

[55] P. L. O Connell, "Korea's high-tech utopia, where everything is observed," *New York Times*, vol. 5, 2005. *(Cited on page 25)*

[56] Y. Punie, "The future of ambient intelligence in europe: the need for more everyday life," *Communications and Strategies*, vol. 57, no. 1, pp. 141–165, 2005. *(Cited on page 25)*

[57] A. Yannopoulos, V. Andronikou, and T. Varvarigou, "Behavioural biometric profiling and ambient intelligence," in *Profiling the European Citizen*. Springer, 2008, pp. 89–109. *(Cited on page 25)*

[58] Frontex Agency, "BIOPASS II Automated biometric border crossing systems based on electronic passports and facial recognition: RAPID and SmartGate," 2010. *(Cited on pages 25 and 103)*

[59] J. Delsing and P. Lindgren, "Sensor communication technology towards ambient intelligence," *Measurement Science and Technology*, vol. 16, no. 4, p. R37, 2005. *(Cited on page 25)*

[60] Y. Cho, S. Cho, D. Choi, S. Jin, K. Chung, and C. Park, "A location privacy protection mechanism for smart space," in *Proceedings of the 2003 International Workshop on Information Security Applications*. Springer, 2003, pp. 162–173. *(Cited on page 25)*

[61] E. Vildjiounaite, S.-M. Mäkelä, M. Lindholm, R. Riihimäki, V. Kyllönen, J. Mäntyjärvi, and H. Ailisto, "Unobtrusive multimodal biometrics for ensuring privacy and information security with personal devices," in *Proceedings of the International Conference on Pervasive Computing*. Springer, 2006, pp. 187–201. *(Cited on page 25)*

[62] F. Zhu, M. W. Mutka, and L. M. Ni, "The master key: A private authentication approach for pervasive computing environments," in *Proceedings of the 2006 4th IEEE International Conference on Pervasive Computing and Communications*. IEEE, 2006, pp. 10–pp. *(Cited on page 25)*

[63] P. Komarinski, *Automated fingerprint identification systems (AFIS)*. Academic Press, 2005. *(Cited on page 28)*

[64] "Automated border controls for europe-eu fp7 project," [Online]. Available:, http://abc4eu.com/, accessed: August 21, 2017. *(Cited on page 28)*

[65] A. K. Jain and S. Z. Li, *Handbook of face recognition*. Springer, 2011. *(Cited on pages 28, 30, and 33)*

[66] D. Maltoni, D. Maio, A. Jain, and S. Prabhakar, *Handbook of fingerprint recognition*. Springer Science & Business Media, 2009. *(Cited on pages 30, 32, 33, 43, 46, and 51)*

[67] J. Daugman, "Iris recognition at airports and border crossings," in *Encyclopedia of Biometrics, Second Edition*, 2015, pp. 998–1004. *(Cited on pages 30 and 33)*

[68] A. Kong, D. Zhang, and M. Kamel, "A survey of palmprint recognition," *Pattern Recognition*, vol. 42, no. 7, pp. 1408–1418, 2009. *(Cited on pages 30 and 33)*

[69] D. Sidlauskas and S. Tamer, "Hand geometry recognition," *Handbook of Biometrics*, pp. 91–107, 2008. *(Cited on pages 30 and 33)*

[70] L. Wang, G. Leedham, and S.-Y. Cho, "Infrared imaging of hand vein patterns for biometric purposes," *IET computer vision*, vol. 1, no. 3, pp. 113–122, 2007. *(Cited on pages 30 and 33)*

[71] D. J. Hurley, B. Arbab-Zavar, and M. S. Nixon, "The ear as a biometric," in *Handbook of biometrics*. Springer, 2008, pp. 131–150. *(Cited on pages 30 and 33)*

[72] T. Hicks and R. Coquoz, "Forensic dna evidence," *Encyclopedia of Biometrics*, pp. 716–723, 2015. *(Cited on pages 30 and 33)*

[73] I. Odinaka, P.-H. Lai, A. D. Kaplan, J. A. O'Sullivan, E. J. Sirevaag, and J. W. Rohrbaugh, "Ecg biometric recognition: A comparative analysis," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 6, pp. 1812–1824, 2012. *(Cited on page 30)*

[74] J. González-Rodríguez, D. T. Toledano, and J. Ortega-García, "Voice biometrics," in *Handbook of biometrics*. Springer, 2008, pp. 151–170. *(Cited on pages 31, 33, and 34)*

[75] R. Chellappa, A. Veeraraghavan, and N. Ramanathan, "Gait biometrics, overview," *Encyclopedia of Biometrics*, pp. 783–789, 2015. *(Cited on pages 31 and 34)*

[76] D. Impedovo and G. Pirlo, "Automatic signature verification: The state of the art," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 5, pp. 609–635, 2008. *(Cited on pages 31 and 33)*

[77] D. Shanmugapriya and G. Padmavathi, "A survey of biometric keystroke dynamics: Approaches, security and challenges," *arXiv preprint arXiv:0910.0817*, 2009. *(Cited on pages 31 and 33)*

[78] A. K. Jain, S. C. Dass, and K. Nandakumar, "Soft biometric traits for personal recognition systems," in *Biometric Authentication*. Springer, 2004, pp. 731–738. *(Cited on pages 31, 32, 34, and 53)*

[79] A. Dantcheva, P. Elia, and A. Ross, "What else does your biometric data reveal? a survey on soft biometrics," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 3, pp. 441–467, 2016. *(Cited on pages 31, 34, and 73)*

[80] S. Denman, A. Bialkowski, C. Fookes, and S. Sridharan, "Determining operational measures from multi-camera surveillance systems using soft biometrics," in *Proceedings of the 2011 8th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)*. IEEE, 2011, pp. 462–467. *(Cited on pages 32 and 34)*

[81] C. Velardo and J.-L. Dugelay, "Weight estimation from visual body appearance," in *Proceedings of the 2010 4th IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS),*. IEEE, 2010, pp. 1–6. *(Cited on page 32)*

[82] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 29, no. 12, pp. 2234–2240, 2007. *(Cited on pages 32 and 53)*

[83] E. Eidinger, R. Enbar, and T. Hassner, "Age and gender estimation of unfiltered faces," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 12, pp. 2170–2179, 2014. *(Cited on pages 32, 53, 54, 73, 96, 99, and 100)*

[84] A. Dantcheva, N. Erdogmus, and J.-L. Dugelay, "On the reliability of eye color as a soft biometric trait," in *Proceedings of the 2011 IEEE Workshop on Applications of Computer Vision (WACV)*. IEEE, 2011, pp. 227–231. *(Cited on page 32)*

[85] R. Donida Labati, A. Genovese, V. Piuri, and F. Scotti, "Toward unconstrained fingerprint recognition: a fully-touchless 3-d system based on two views on the move," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 2, pp. 202–219, February 2016. *(Cited on pages 32, 33, 51, and 68)*

[86] R. Donida Labati, A. Genovese, E. Muñoz, V. Piuri, F. Scotti, and G. Sforza, "Biometric recognition in automated border control: A survey," *ACM Computing Surveys (CSUR)*, vol. 49, no. 2, p. 24, 2016. *(Cited on pages 33 and 68)*

[87] M. Kosmerlj, T. Fladsrud, E. Hjelmas, and E. Snekkenes, "Face recognition issues in a border control environment," in *Advances in Biometrics*. Springer, 2005, vol. 3832, pp. 33–39. *(Cited on page 33)*

[88] C. Conde, I. Martín de Diego, and E. Cabello, "Face recognition in uncontrolled environments, experiments in an airport," in *Proceedings of the International Joint Conference on E-Business and Telecommunications (ICETE 2011) - Revised Selected*

Papers, M. S. Obaidat, J. L. Sevillano, and J. Filipe, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 20–32. *(Cited on page 33)*

[89] H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs, "A study of face recognition as people age," in *2007 IEEE 11th International Conference on Computer Vision*. IEEE, 2007, pp. 1–8. *(Cited on page 33)*

[90] R. P. Wildes, "Iris recognition: an emerging biometric technology," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1348–1363, 1997. *(Cited on page 33)*

[91] J. Daugman, "The importance of being random: statistical principles of iris recognition," *Pattern recognition*, vol. 36, no. 2, pp. 279–291, 2003. *(Cited on pages 33 and 54)*

[92] R. Donida Labati and F. Scotti, "Noisy iris segmentation with boundary regularization and reflections removal," *Image and Vision Computing, Iris Images Segmentation Special Issue*, vol. 28, no. 2, pp. 270 – 277, February 2010. *(Cited on pages 33 and 54)*

[93] S. Barra, A. Casanova, F. Narducci, and S. Ricciardi, "Ubiquitous iris recognition by means of mobile devices," *Pattern Recognition Letters*, vol. 57, pp. 66–73, 2015. *(Cited on pages 33 and 54)*

[94] G. R. Doddington, "Speaker recognition: Identifying people by their voices," *Proceedings of the IEEE*, vol. 73, no. 11, pp. 1651–1664, 1985. *(Cited on pages 33, 34, 52, and 68)*

[95] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette analysis-based gait recognition for human identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1505–1518, 2003. *(Cited on page 33)*

[96] J. E. Mason, I. Traoré, and I. Woungang, "Applications of gait biometrics," in *Machine Learning Techniques for Gait Biometric Recognition*. Springer, 2016, pp. 203–208. *(Cited on pages 33 and 54)*

[97] K. Niinuma, U. Park, and A. K. Jain, "Soft biometric traits for continuous user authentication," *IEEE Transactions on information forensics and security*, vol. 5, no. 4, pp. 771–780, 2010. *(Cited on pages 34, 61, and 62)*

[98] A. Ross, K. Nandakumar, and A. Jain, *Handbook of multibiometrics*. Springer, 2006, vol. 6. *(Cited on pages 34 and 56)*

[99] A. Ross, K. Nandakumar, and A. K. Jain, *Handbook of Multibiometrics*, ser. International Series on Biometrics. Springer, 2006, vol. 6. *(Cited on pages 34, 56, and 75)*

[100] A. Jain and A. Ross, "Learning user-specific parameters in a multibiometric system," in *Proceedings of International Conference on Image Processing (ICIP)*, vol. 1, 2002, pp. 57–60. *(Cited on page 34)*

[101] A. Ross, K. Nandakumar, and A. K. Jain, "Introduction to multibiometrics," in *Handbook of Biometrics*. Springer, 2008, pp. 271–292. *(Cited on page 34)*

[102] A. Ross and A. Jain, "Information fusion in biometrics," *Pattern Recognition Letters*, vol. 24, no. 13, pp. 2115–2125, 2003. *(Cited on pages 35, 55, 81, and 114)*

[103] M. He, S. Horng, P. Fan, R. Run, R. Chen, J. Lai, M. Khan, and K. Sentosa, "Performance evaluation of score level fusion in multimodal biometric systems," *Patt. Recogn.*, vol. 43, no. 5, pp. 1789–1800, 2010. *(Cited on pages 35 and 47)*

[104] K. Nandakumar, Yi Chen, S. Dass, and A. Jain, "Likelihood ratio-based biometric score fusion," *IEEE Transactions Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 342–347, 2008. *(Cited on pages 35, 55, 82, 111, 112, and 113)*

[105] A. Ross and A. Jain, "Information fusion in biometrics," *Pattern recognition letters*, vol. 24, no. 13, pp. 2115–2125, 2003. *(Cited on page 35)*

[106] R. V. Yampolskiy and V. Govindaraju, "Direct and indirect human computer interaction based biometrics," *Journal of computers*, vol. 2, no. 10, pp. 76–88, 2007. *(Cited on pages 35, 46, and 47)*

[107] R. V. Yampolskiy, "Human computer interaction based intrusion detection," in *Information Technology, 2007. ITNG'07. Fourth International Conference on*. IEEE, 2007, pp. 837–842. *(Cited on page 35)*

[108] H. Gamboa and A. Fred, "A behavioral biometric system based on human-computer interaction," in *Defense and Security*. International Society for Optics and Photonics, 2004, pp. 381–392. *(Cited on pages 35 and 47)*

[109] E. P. Kukula, S. J. Elliott, and V. G. Duffy, "The effects of human interaction on biometric system performance," in *Digital Human Modeling*. Springer, 2007, pp. 904–914. *(Cited on page 36)*

[110] A. S. Patrick, "Usability and acceptability of biometric security systems," in *Financial Cryptography*. Springer, 2004, p. 105. *(Cited on pages 36 and 47)*

[111] F. Tayyari and J. L. Smith, *Occupational ergonomics: principles and applications*. Chapman & Hall London, 1997. *(Cited on page 36)*

[112] E. P. Kukula, M. J. Sutton, and S. J. Elliott, "The human–biometric-sensor interaction evaluation method: biometric performance and usability measurements," *IEEE Transactions on Systems Instrumentation and Measurement*, vol. 59, no. 4, pp. 784–791, 2010. *(Cited on page 36)*

[113] A. Mansfield, "International Organization for Standardization/IEC 19795-1 biometric performance testing and reporting: Principles and framework, fdis ed., jtc1/sc37/working group 5, aug. 2005." *(Cited on pages 36 and 40)*

[114] "International Organization for Standardization, ISO 13407: Human-centred design processes for interactive systems," *ISO*, 1999. *(Cited on page 36)*

[115] "International Organization for Standardization, ISO/IEC JTC1 SC37: Standing dcoument 2, version 8, harmonized biometric vocablury," *ISO/IEC*, August 1997. *(Cited on page 36)*

[116] N. Bevan, "Human-computer interaction standards," *Proceedings of the 6th International Conference on Human Computer Interaction*, vol. 20, pp. 885–885, 1995. *(Cited on page 36)*

[117] M. F. Theofanos, B. Stanton, S. Orandi, R. Micheals, and N.-F. Zhang, *Usability Testing of Ten-prints Fingerprint Capture*. US Department of Commerce, Technology Administration, National Institute of Standards and Technology, 2007. *(Cited on pages 36, 40, and 47)*

[118] R. D. Labati, A. Genovese, E. Muñoz, V. Piuri, F. Scotti, and G. Sforza, "Automated border control systems: Biometric challenges and research trends," in *Proceedings of the International Conference on Information Systems Security*. Springer, 2015, pp. 11–20. *(Cited on pages 36 and 46)*

[119] R. Donida Labati, A. Genovese, E. Muñoz Ballester, V. Piuri, F. Scotti, and G. Sforza, "Automatic classification of acquisition problems affecting fingerprint images in automated border controls," in *Proceedings of the 2015 IEEE Symposium Series on Computational Intelligence*. IEEE, 2015, pp. 354–361. *(Cited on pages 36, 51, 79, 102, and 110)*

[120] R. Donida Labati, V. Piuri, and F. Scotti, "Biometric privacy protection: guidelines and technologies," in *Communications in Computer and Information Science*, M. S. Obaidat, J. Sevillano, and F. Joaquim, Eds. Springer, 2012, vol. 314, pp. 3–19. *(Cited on pages 37 and 47)*

[121] L. M. Mayron, Y. Hausawi, and G. S. Bahr, "Secure, usable biometric authentication systems," in *Proceedings of the International Conference on Universal Access in Human-Computer Interaction*. Springer, 2013, pp. 195–204. *(Cited on page 37)*

[122] T. Bianchi, R. Donida Labati, V. Piuri, A. Piva, F. Scotti, and S. Turchi, "Implementing fingercode-based identity matching in the encrypted domain," in *Proceedings of the 2010 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications*, September 2010, pp. 15–21. *(Cited on pages 37 and 47)*

[123] M. Barni, T. Bianchi, D. Catalano, M. D. Raimondo, R. Donida Labati, P. Failla, D. Fiore, R. Lazzeretti, V. Piuri, F. Scotti, and A. Piva, "A privacy-compliant fingerprint recognition system based on homomorphic encryption and fingercode templates," in *Proceedings of the 2010 IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS 2010)*, September 2010, pp. 1–7. *(Cited on pages 37, 47, and 58)*

[124] A. K. Jain, K. Nandakumar, and A. Nagar, "Biometric template security," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, p. 113, 2008. *(Cited on pages 37, 47, and 58)*

[125] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain, "Fvc2000: Fingerprint verification competition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 402–412, 2002. *(Cited on page 38)*

[126] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The feret evaluation methodology for face-recognition algorithms," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000. *(Cited on page 38)*

[127] G. R. Doddington, M. A. Przybocki, A. F. Martin, and D. A. Reynolds, "The nist speaker recognition evaluation–overview, methodology, systems, results, perspective," *Speech Communication*, vol. 31, no. 2, pp. 225–254, 2000. *(Cited on pages 38 and 44)*

[128] T. Mansfield, G. Kelly, D. Chandler, and J. Kane, "Biometric product testing final report," *Computing, National Physical Laboratory, Crown Copyright, UK*, 2001. *(Cited on page 39)*

[129] A. Jain, R. Bolle, and S. Pankanti, *Biometrics: personal identification in networked society*. Springer Science & Business Media, 2006, vol. 479. *(Cited on pages 39, 51, and 54)*

[130] M. Gamassi, M. Lazzaroni, M. Misino, V. Piuri, D. Sana, and F. Scotti, "Quality assessment of biometric systems: a comprehensive perspective based on accuracy and performance measurement," *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 4, pp. 1489–1496, 2005. *(Cited on page 39)*

[131] A. J. Mansfield and J. L. Wayman, "Best practices in testing and reporting performance of biometric devices," National Physical Laboratory, Tech. Rep., August 2002. *(Cited on pages 44 and 106)*

[132] B. D. Jovanovic and P. S. Levy, "A look at the rule of three," *The American Statistician*, vol. 51, no. 2, pp. 137–139, 1997. *(Cited on page 44)*

[133] G. W. Snedecor, *Statistical methods*. Iowa state college press, 1940. *(Cited on page 44)*

[134] R. Bolle, N. Ratha, and S. Pankanti, "Confidence interval measurement in performance analysis of biometrics systems using the bootstrap," in *Proceedings of the 2001 IEEE Workshop on Empirical Evaluation Methods in Computer Vision*, vol. 28, 2001. *(Cited on page 45)*

[135] N. Poh and S. Bengio, "Estimating the confidence interval of expected performance curve in biometric authentication using joint bootstrap," in *Proceedings of the 2007 IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP)*, vol. 2. IEEE, 2007, pp. II–137. *(Cited on page 45)*

[136] S. C. Dass, Y. Zhu, and A. K. Jain, "Validating a biometric authentication system: Sample size requirements," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 1902–1319, 2006. *(Cited on page 45)*

[137] N. Poh, A. Martin, and S. Bengio, "Performance generalization in biometric authentication using joint user-specific and sample bootstraps," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 3, pp. 492–498, 2007. *(Cited on page 45)*

[138] A. A. Ross, K. Nandakumar, and A. Jain, *Handbook of multibiometrics*. Springer Science & Business Media, 2006, vol. 6. *(Cited on pages 46 and 47)*

[139] EC, "Standards for security features and biometrics in passports and travel documents issued by Member States," Regulation No. 444/2009 amending Regulation No. 2252/2004, 2009, EU Commission. *(Cited on page 47)*

[140] A. Juels, D. Molnar, and D. Wagner, "Security and privacy issues in e-passports," in *Proceedings of the 2005 1st IEEE International Conference on Security and Privacy for Emerging Areas in Communications Networks, (SecureComm)*. IEEE, 2005, pp. 74–88. *(Cited on page 47)*

[141] L. Coventry, "Fingerprint authentication: The user experience," in *DIMACS workshop on usable privacy and security software*, 2004. *(Cited on page 47)*

[142] R. Donida Labati and F. Scotti, "Fingerprint," in *Encyclopedia of Cryptography and Security (2nd ed.)*, H. C. A. van Tilborg and S. Jajodia, Eds. Springer, 2011, pp. 460–465. *(Cited on page 47)*

[143] M. A. Sasse, "Usability and trust in information systems," 2005. *(Cited on page 47)*

[144] "Biometrics and usability: Efficiency, effectiveness, and user satisfaction,nist, 2009," [Online]. Available: http://www.nist.gov/itl/iad/vug/biousa/, accessed: February 14, 2016. *(Cited on page 47)*

[145] A. Krupp, C. Rathgeb, and C. Busch, "Social acceptance of biometric technologies in germany: A survey," in *Proceedings of International Conference on Biometrics Special Interest Group*. IEEE, 2013, pp. 1–5. *(Cited on page 47)*

[146] A. Ross and A. K. Jain, "Multimodal biometrics: An overview," in *Proceedings of 12th European Signal Processing Conference*. IEEE, 2004, pp. 1221–1224. *(Cited on page 47)*

[147] C. Chia, N. Sherkat, and L. Nolle, "Towards a best linear combination for multimodal biometric fusion," in *Proceedings of the 2010 20th IEEE International Conference on Pattern Recognition*, 2010, pp. 1176–1179. *(Cited on pages 47, 111, and 113)*

[148] R. Donida Labati, A. Genovese, V. Piuri, and F. Scotti, "Toward unconstrained fingerprint recognition: a fully-touchless 3-d system based on two views on the move," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 2, pp. 202–219, February 2016. *(Cited on pages 47, 56, and 76)*

[149] B. Kamgar-Parsi, W. Lawson, and B. Kamgar-Parsi, "Toward development of a face recognition system for watchlist surveillance," *IEEE Transactions on Pattern*

Analysis and Machine Intelligence, vol. 33, no. 10, pp. 1925–1937, 2011. *(Cited on page 47)*

[150] A. Genovese, V. Piuri, and F. Scotti, *Touchless palmprint recognition systems*. Springer, 2014, vol. 60. *(Cited on pages 47 and 68)*

[151] G. González, C. Angulo, B. López, and J. L. de la Rosa, "Smart user models: Modelling the humans in ambient recommender systems," in *Proceedings of the workshop on decentralized, agent based and social approaches to user modelling (DA-SUM)*. Citeseer, 2005, pp. 11–20. *(Cited on page 47)*

[152] A. Jaimes and N. Sebe, "Multimodal human–computer interaction: A survey," *Computer vision and image understanding*, vol. 108, no. 1, pp. 116–134, 2007. *(Cited on pages 47 and 55)*

[153] M. Pantic and L. J. Rothkrantz, "Automatic analysis of facial expressions: The state of the art," *IEEE Transactions on Pattern Analysis and Machine Intelligence,*, vol. 22, no. 12, pp. 1424–1445, 2000. *(Cited on page 50)*

[154] N. Sebe, I. Cohen, F. G. Cozman, T. Gevers, and T. S. Huang, "Learning probabilistic classifiers for human–computer interaction applications," *Multimedia Systems*, vol. 10, no. 6, pp. 484–498, 2005. *(Cited on page 50)*

[155] V. Menon, B. Jayaraman, and V. Govindaraju, "Biometrics driven smart environments: abstract framework and evaluation," in *International Conference on Ubiquitous Intelligence and Computing*. Springer, 2008, pp. 75–89. *(Cited on page 50)*

[156] Y. Gao, S. C. Hui, and A. C. M. Fong, "A multiview facial analysis technique for identity authentication," *IEEE Pervasive Computing*, vol. 2, no. 1, pp. 38–45, 2003. *(Cited on pages 50 and 55)*

[157] H. K. Ekenel, M. Fischer, H. Gao, L. Toth, and R. Stiefelhagen, "Face recognition for smart interactions," in *Proceedings of the 8th IEEE International Conference on Automatic Face & Gesture Recognition*. IEEE, 2008, pp. 1–2. *(Cited on page 50)*

[158] F. Zuo and P. H. De With, "Real-time embedded face recognition for smart home," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 1, pp. 183–190, 2005. *(Cited on page 50)*

[159] T. Kohonen, *Self-organization and associative memory*. Springer Science & Business Media, 2012, vol. 8. *(Cited on page 50)*

[160] M. Kirby and L. Sirovich, "Application of the karhunen-loeve procedure for the characterization of human faces," *IEEE Transactions on Pattern analysis and Machine intelligence*, vol. 12, no. 1, pp. 103–108, 1990. *(Cited on page 50)*

[161] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991. *(Cited on page 50)*

[162] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986. *(Cited on page 50)*

[163] T.-K. Kim and J. Kittler, "Locally linear discriminant analysis for multimodally distributed classes for face recognition with a single model image," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 3, pp. 318–327, 2005. *(Cited on page 50)*

[164] M. Viola and P. Viola, "Face recognition using boosted local features," in *Proceedings of the IEEE International Conference on Computer Vision*, 2003. *(Cited on page 50)*

[165] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," *Computer vision-eccv 2004*, pp. 469–481, 2004. *(Cited on page 50)*

[166] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss, "The feret database and evaluation procedure for face-recognition algorithms," *Image and vision computing*, vol. 16, no. 5, pp. 295–306, 1998. *(Cited on page 50)*

[167] L. Wiskott, N. Krüger, N. Kuiger, and C. Von Der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 775–779, 1997. *(Cited on page 50)*

[168] K. Etemad and R. Chellappa, "Discriminant analysis for recognition of human face images," *JOSA A*, vol. 14, no. 8, pp. 1724–1733, 1997. *(Cited on page 51)*

[169] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object representation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 696–710, 1997. *(Cited on page 51)*

[170] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *BMVC*, vol. 1, no. 3, 2015, p. 6. *(Cited on pages 51 and 73)*

[171] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105. *(Cited on pages 51 and 73)*

[172] P. S. Penev and J. J. Atick, "Local feature analysis: A general statistical theory for object representation," *Network: computation in neural systems*, vol. 7, no. 3, pp. 477–500, 1996. *(Cited on page 51)*

[173] D. O. Gorodnichy, E. Granger, and P. Radtke, "Survey of commercial technologies for face recognition in video," Canada Border Services Agency Ottawa ON Canada, Tech. Rep., 2014. *(Cited on page 51)*

[174] A. K. Jain, S. Prabhakar, and A. Ross, "Fingerprint matching: Data acquisition and performance evaluation," *Tech. Report MSU-CPS-99–14*, 1999. *(Cited on page 51)*

[175] D. Maltoni, D. Maio, A. Jain, and S. Prabhakar, *Handbook of fingerprint recognition*, 2nd ed. Springer Publishing Company, Incorporated, 2009. *(Cited on pages 51, 56, 75, 79, 101, 103, and 110)*

[176] C. Riley, G. Johnson, H. McCracken, and A. Al-Saffar, "Instruction, feedback and biometrics: The user interface for fingerprint authentication systems," *Human-Computer Interaction–INTERACT 2009*, pp. 293–305, 2009. *(Cited on pages 51, 79, and 101)*

[177] N. K. Ratha, K. Karu, S. Chen, and A. K. Jain, "A real-time matching system for large fingerprint databases," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 799–813, 1996. *(Cited on page 51)*

[178] S. Ranade and A. Rosenfeld, "Point pattern matching by relaxation," *Pattern recognition*, vol. 12, no. 4, pp. 269–275, 1980. *(Cited on page 51)*

[179] W. Sheng, G. Howells, M. Fairhurst, and F. Deravi, "A memetic fingerprint matching algorithm," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3, pp. 402–412, 2007. *(Cited on page 51)*

[180] B. R. Masters, R. C. Gonzalez, and R. Woods, "Digital image processing," *Journal of biomedical optics*, vol. 14, no. 2, p. 029901, 2009. *(Cited on page 51)*

[181] K. Nandakumar and A. K. Jain, "Local correlation-based fingerprint matching." in *ICVGIP*, 2004, pp. 503–508. *(Cited on page 51)*

[182] A. K. Jain, S. Prabhakar, L. Hong, and S. Pankanti, "Filterbank-based fingerprint matching," *IEEE transactions on Image Processing*, vol. 9, no. 5, pp. 846–859, 2000. *(Cited on page 51)*

[183] R. Zhou, S. Sin, D. Li, T. Isshiki, and H. Kunieda, "Adaptive sift-based algorithm for specific fingerprint verification," in *Hand-Based Biometrics (ICHB), 2011 International Conference on*. IEEE, 2011, pp. 1–6. *(Cited on page 51)*

[184] X. Liang, A. Bishnu, and T. Asano, "A robust fingerprint indexing scheme using minutia neighborhood structure and low-order delaunay triangles," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 4, pp. 721–733, 2007. *(Cited on page 51)*

[185] C. I. Watson, M. D. Garris, E. Tabassi, C. L. Wilson, R. M. McCabe, S. Janet, and K. Ko, "User's guide to nist biometric image software (nbis), national institute of standards and technology, 2006." *(Cited on page 51)*

[186] Dermalog, "LF10 Fingerprint Scanner." [Online]. Available: http://www.dermalog.com/en/products_solutions/fingerprintscanner/lf10.php *(Cited on pages 51, 101, and 102)*

[187] ——, "DERMALOG High Speed AFIS." [Online]. Available: http://www.dermalog.com/en/products_solutions/afis/ *(Cited on pages 51 and 107)*

[188] N. Technology, "VeriFinger, SDK Neuro Technology, 2010," accessed: September 14, 2017. [Online]. Available: http://www.neurotechnology.com/vf_sdk.html *(Cited on page 51)*

[189] D. A. Reynolds and R. C. Rose, "Robust text-independent speaker identification using gaussian mixture speaker models," *IEEE transactions on speech and audio processing*, vol. 3, no. 1, pp. 72–83, 1995. *(Cited on pages 52, 68, and 70)*

[190] T. Kinnunen and H. Li, "An overview of text-independent speaker recognition: From features to supervectors," *Speech communication*, vol. 52, no. 1, pp. 12–40, 2010. *(Cited on pages 52, 53, 68, and 70)*

[191] K. R. Farrell, R. J. Mammone, and K. T. Assaleh, "Speaker recognition using neural networks and conventional classifiers," *IEEE Transactions on speech and audio processing*, vol. 2, no. 1, pp. 194–205, 1994. *(Cited on pages 52 and 68)*

[192] H. Zeinali, H. Sameti, L. Burget, J. Černockỳ, N. Maghsoodi, and P. Matějka, "i-vector/hmm based text-dependent speaker verification system for reddots challenge," *Interspeech 2016*, pp. 440–444, 2016. *(Cited on page 52)*

[193] J. Pelecanos and S. Sridharan, "Feature warping for robust speaker verification," 2001. *(Cited on page 52)*

[194] D. Reynolds, "An overview of automatic speaker recognition," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)(S. 4072-4075)*, 2002. *(Cited on pages 52 and 53)*

[195] J. Pelecanos, S. Myers, S. Sridharan, and V. Chandran, "Vector quantization based gaussian modeling for speaker verification," in *Proceedings of the 2000 15th International Conference on Pattern Recognition*, vol. 3. IEEE, 2000, pp. 294–297. *(Cited on page 52)*

[196] T. Hasan and J. H. Hansen, "A study on universal background model training in speaker verification," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 7, pp. 1890–1899, 2011. *(Cited on page 52)*

[197] T. May, S. van de Par, and A. Kohlrausch, "Noise-robust speaker recognition combining missing data techniques and universal background modeling," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 1, pp. 108–121, 2012. *(Cited on page 52)*

[198] A. O. Hatch, A. Stolcke, and B. Peskin, "Combining feature sets with support vector machines: Application to speaker recognition," in *Proceedings of the 2005 IEEE Workshop on Automatic Speech Recognition and Understanding*. IEEE, 2005, pp. 75–79. *(Cited on page 52)*

[199] S. Davis and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," *IEEE transactions on acoustics, speech, and signal processing*, vol. 28, no. 4, pp. 357–366, 1980. *(Cited on pages 52 and 70)*

[200] K. Kumar, C. Kim, and R. M. Stern, "Delta-spectral cepstral coefficients for robust speech recognition," in *Proceedings of the 2011 IEEE International Conference on*

Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2011, pp. 4784–4787. *(Cited on page 52)*

[201] X. Huang, A. Acero, H.-W. Hon, and R. Foreword By-Reddy, *Spoken language processing: A guide to theory, algorithm, and system development*. Prentice hall PTR, 2001. *(Cited on page 52)*

[202] H. Hermansky, "Perceptual linear predictive (plp) analysis of speech," *the Journal of the Acoustical Society of America*, vol. 87, no. 4, pp. 1738–1752, 1990. *(Cited on page 52)*

[203] D. A. Reynolds, W. Campbell, T. Gleason, C. Quillen, D. Sturim, P. Torres-Carrasquillo, and A. Adami, "The 2004 mit lincoln laboratory speaker recognition system," in *Proceedings of the). IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1. IEEE, 2005, pp. I–177. *(Cited on page 52)*

[204] H. Hermansky and N. Morgan, "Rasta processing of speech," *IEEE transactions on speech and audio processing*, vol. 2, no. 4, pp. 578–589, 1994. *(Cited on page 52)*

[205] F. Richardson, D. Reynolds, and N. Dehak, "Deep neural network approaches to speaker and language recognition," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1671–1675, 2015. *(Cited on page 52)*

[206] E. Variani, X. Lei, E. McDermott, I. L. Moreno, and J. Gonzalez-Dominguez, "Deep neural networks for small footprint text-dependent speaker verification," in *Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 4052–4056. *(Cited on page 52)*

[207] J. Li, D. Yu, J.-T. Huang, and Y. Gong, "Improving wideband speech recognition using mixed-bandwidth training data in cd-dnn-hmm," in *Spoken Language Technology Workshop (SLT), 2012 IEEE*. IEEE, 2012, pp. 131–136. *(Cited on page 52)*

[208] L. Li, Y. Lin, Z. Zhang, and D. Wang, "Improved deep speaker feature learning for text-dependent speaker recognition," in *Proceedings of the2015 Asia-Pacific Annual Summit and Conference on Signal and Information Processing Association (APSIPA)*. IEEE, 2015, pp. 426–429. *(Cited on page 52)*

[209] D. Reid, S. Samangooei, C. Chen, M. Nixon, and A. Ross, "Soft biometrics for surveillance: an overview," *Machine learning: theory and applications*, pp. 327–352, 2013. *(Cited on page 53)*

[210] R. Azarmehr, R. Laganiere, W.-S. Lee, C. Xu, and D. Laroche, "Real-time embedded age and gender classification in unconstrained video," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2015, pp. 57–65. *(Cited on page 53)*

[211] G. Guo, C. Dyer, Y. Fu, and T. S. Huang, "Is gender recognition affected by age?" in *Proceedings of the 2009 12th IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE, 2009, pp. 2032–2039. *(Cited on page 53)*

[212] B. A. Golomb, D. T. Lawrence, and T. J. Sejnowski, "Sexnet: A neural network identifies sex from human faces," in *NIPS*, vol. 1, 1990, p. 2. *(Cited on page 53)*

[213] E. Mäkinen and R. Raisamo, "An experimental comparison of gender classification methods," *Pattern Recognition Letters*, vol. 29, no. 10, pp. 1544–1556, 2008. *(Cited on page 53)*

[214] T. Dhimar and K. Mistree, "Feature extraction for facial age estimation: A survey," in *Proceedings of the 2016 IEEE International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*. IEEE, 2016, pp. 2243–2248. *(Cited on page 53)*

[215] A. Lanitis, C. Taylor, and T. F. Cootes, "Toward automatic simulation of aging effects on face images," *IEEE Transactions Pattern Analysis and Machine Intelligence (PAMI)*, vol. 24, no. 4, pp. 442–455, 2002. *(Cited on page 53)*

[216] S. Sahni and S. Saxena, "Survey: Techniques for aging problems in face recognition," *MIT International Journal of Computer Science and Information Technology*, vol. 4, no. 2, pp. 82–88, 2014. *(Cited on page 53)*

[217] A. Gunay and V. Nabiyev, "Automatic age classification with lbp," in *Proceedings of the 2008 23rd IEEE International Symposium on Computer and Information Sciences, ISCIS'08*. IEEE, 2008, pp. 1–4. *(Cited on page 53)*

[218] F. Gao and H. Ai, "Face age classification on consumer images with gabor feature and fuzzy lda method," in *Proceedings of the International Conference on Biometrics*. Springer, 2009, pp. 132–141. *(Cited on page 53)*

[219] G. Guo, G. Mu, Y. Fu, and T. Huang, "Human age estimation using bio-inspired features," in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2009, pp. 112–119. *(Cited on page 53)*

[220] J.-D. Txia and C.-L. Huang, "Age estimation using aam and local facial features," in *Proceedings of the2009 5th IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. IEEE, 2009, pp. 885–888. *(Cited on page 53)*

[221] X. Geng, C. Yin, and Z. Zhou, "Facial age estimation by learning from label distributions," *IEEE Transactions Pattern Anal. Mach. Intell. (PAMI)*, vol. 35, no. 10, pp. 2401–2412, 2013. *(Cited on page 53)*

[222] S. Feng, C. Lang, J. Feng, T. Wang, and J. Luo, "Human facial age estimation by cost-sensitive label ranking and trace norm regularization," *IEEE Transactions on Multimedia*, vol. 19, no. 1, pp. 136–148, 2017. *(Cited on page 53)*

[223] R. D. Labati, A. Genovese, E. Muñoz, V. Piuri, and F. Scotti, "A novel pore extraction method for heterogeneous fingerprint images using convolutional neural networks," *Pattern Recognition Letters*, 2017. *(Cited on page 54)*

[224] A. Toshev and C. Szegedy, "Deeppose: Human pose estimation via deep neural networks," in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 1653–1660. *(Cited on page 54)*

[225] R. Rothe, R. Timofte, and L. V. Gool, "Dex: Deep expectation of apparent age from a single image," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2015, pp. 10–15. *(Cited on pages 54, 74, and 95)*

[226] G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2015, pp. 34–42. *(Cited on pages 54 and 100)*

[227] H. Liu, J. Lu, J. Feng, and J. Zhou, "Group-aware deep feature learning for facial age estimation," *Pattern Recognition*, vol. 66, pp. 82–94, 2017. *(Cited on pages 54 and 73)*

[228] Y. Zhang and Q. Ji, "Active and dynamic information fusion for facial expression understanding from image sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 699–714, 2005. *(Cited on page 54)*

[229] H. Proença and L. A. Alexandre, "The nice 1: Noisy iris challenge evaluation-part 1," in *Proceedings of 1st IEEE International Conference on Biometrics: Theory, Applications, and Systems*, 2007, pp. 1–4. *(Cited on page 54)*

[230] K. B. Raja, R. Raghavendra, V. K. Vemuri, and C. Busch, "Smartphone based visible iris recognition using deep sparse filtering," *Pattern Recognition Letters*, vol. 57, pp. 33–42, 2015. *(Cited on page 54)*

[231] S. Oviatt, "Multimodal interfaces," *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications*, vol. 14, pp. 286–304, 2003. *(Cited on page 55)*

[232] W. Wahlster, "Smartkom: Fusion and fission of speech, gestures, and facial expressions," in *Proceedings of the 1st International Workshop on Man-Machine Symbiotic Systems*, 2002, pp. 213–225. *(Cited on pages 55 and 68)*

[233] J. Kleindienst, T. Macek, L. Serédi, and J. Sedivy, "Vision-enhanced multimodal interactions in domotic environments," *IBM Tecnologías de Voz y Sistemas. República Checa*, pp. 1059–1064, 2004. *(Cited on pages 55 and 68)*

[234] D. Cuesta Cantarero, D. A. Pérez Herrero, and F. Martín Méndez, "A multimodal biometric fusion implementation for ABC systems," in *Proceedings of the European Intelligence and Security Informatics Conference (EISIC 2013)*, August 2013, pp. 277–280. *(Cited on page 55)*

[235] I. Iglezakis, "EU data protection legislation and case-law with regard to biometric applications," *Social Science Research Network*, 2013. *(Cited on pages 56, 58, and 76)*

[236] V. MacLeod and B. McLindin, "Methodology for the evaluation of an international airport Automated Border Control processing system," in *Innovations in Defence Support Systems -2*. Springer, 2011, vol. 338, pp. 115–145. *(Cited on page 56)*

[237] D. Lee, K. Choi, S. Lee, and J. Kim, "Fingerprint fusion based on minutiae and ridge for enrollment," in *Proceedings of the 4th International Conference on Audio- and Video-based Biometric Person Authentication*, 2003, pp. 478–485. *(Cited on pages 56 and 75)*

[238] J. Yang, N. Xiong, and A. V. Vasilakos, "Two-stage enhancement scheme for low-quality fingerprint images by learning from the images," *IEEE Transactions on Human-Machine Systems*, vol. 43, pp. 235–248, March 2013. *(Cited on pages 56 and 76)*

[239] N. Poh, J. Kittler, S. Marcel, D. Matrouf, and J. F. Bonastre, "Model and score adaptation for biometric systems: Coping with device interoperability and changing acquisition conditions," in *Proceedings of the 2010 20th International Conference on Pattern Recognition (ICPR 2010)*, August 2010, pp. 1229–1232. *(Cited on pages 56, 75, and 76)*

[240] A. Merati, N. Poh, and J. Kittler, "User-specific cohort selection and score normalization for biometric systems," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 4, pp. 1270–1277, 2012. *(Cited on pages 56, 57, and 76)*

[241] G. Aggarwal, N. K. Ratha, R. M. Bolle, and R. Chellappa, "Multi-biometric cohort analysis for biometric fusion," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*, March 2008, pp. 5224–5227. *(Cited on pages 56, 57, 76, and 78)*

[242] A. E. Rosenberg, J. DeLong, C. H. Lee, B. H. Juang, and F. K. Soong, "The use of cohort normalized scores for speaker verification," in *Proceedings of the Second International Conference on Spoken Language Processing (ICSLP 1992)*, October 1992. *(Cited on pages 56 and 76)*

[243] G. Aggarwal, N. K. Ratha, and R. M. Bolle, "Biometric verification: Looking beyond raw similarity scores," in *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW 2006)*, June 2006, pp. 31–31. *(Cited on pages 56, 57, 76, 78, 79, 104, and 116)*

[244] R. Donida Labati, V. Piuri, and F. Scotti, *Touchless Fingerprint Biometrics*. CRC Press, August 2015. *(Cited on pages 56, 68, and 76)*

[245] V. Piuri and F. Scotti, "Fingerprint biometrics via low-cost sensors and webcams," in *Proceedings of the 2008 IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS 2008)*, September 2008, pp. 1–6. *(Cited on pages 56 and 76)*

[246] S. Tulyakov, Z. Zhang, and V. Govindaraju, "Comparison of combination methods utilizing t-normalization and second best score model," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2008)*, June 2008, pp. 1–5. *(Cited on page 57)*

[247] N. Poh, A. Merati, and J. Kittler, "Making better biometric decisions with quality and cohort information: a case study in fingerprint verification," in *Proceedings of the 2009 17th European Signal Processing Conference (EUSIPCO 2009)*, August 2009, pp. 70–74. *(Cited on pages 57 and 78)*

[248] M. Tistarelli, Y. Sun, and N. Poh, "On the use of discriminative cohort score normalization for unconstrained face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 12, pp. 2063–2075, 2014. *(Cited on pages 57 and 58)*

[249] A. Kumar, "Incorporating cohort information for reliable palmprint authentication," in *Computer Vision, Graphics & Image Processing, 2008. ICVGIP'08. Sixth Indian Conference on*, December 2008, pp. 583–590. *(Cited on page 57)*

[250] M. Tistarelli and Y. Sun, "Cohort-based score normalization," in *Encyclopedia of Biometrics*, Z. S. Li and K. A. Jain, Eds. Springer US, 2014, pp. 1–9. *(Cited on page 57)*

[251] M. Barni, T. Bianchi, D. Catalano, M. Di Raimondo, R. Donida Labati, P. Failla, D. Fiore, R. Lazzeretti, V. Piuri, F. Scotti, and A. Piva, "Privacy-preserving fingercode authentication," in *Proceedings of the 2010 ACM Workshop on Multimedia and Security*, September 2010, pp. 231–240. *(Cited on page 58)*

[252] V. M. Patel, R. Chellappa, D. Chandra, and B. Barbello, "Continuous user authentication on mobile devices: Recent progress and remaining challenges," *IEEE Signal Processing Magazine*, vol. 33, no. 4, pp. 49–61, 2016. *(Cited on pages 61 and 62)*

[253] A. Azzini, S. Marrara, R. Sassi, and F. Scotti, "A fuzzy approach to multimodal biometric continuous authentication," *Fuzzy Optimization and Decision Making*, vol. 7, no. 3, pp. 243–256, 2008. *(Cited on pages 61 and 62)*

[254] M. E. Fathy, V. M. Patel, and R. Chellappa, "Face-based active authentication on mobile devices," in *Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015, pp. 1687–1691. *(Cited on page 61)*

[255] R. Donida Labati, R. Sassi, and F. Scotti, "Ecg biometric recognition: permanence analysis of qrs signals for 24 hours continuous authentication," in *Proceedings of the IEEE International Workshop on Information Forensics and Security (WIFS 2013)*, Guangzhou, China, November 2013. *(Cited on pages 61 and 64)*

[256] P. Bours and H. Barghouthi, "Continuous authentication using biometric keystroke dynamics," in *The Norwegian Information Security Conference (NISK)*, vol. 2009, 2009. *(Cited on page 61)*

[257] A. Bonissi, R. Donida Labati, L. Perico, R. Sassi, F. Scotti, and L. Sparagino, "A preliminary study on continuous authentication methods for photoplethysmographic biometrics," in *Proceedings of the 2013 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BioMS 2013)*, Napoli, Italy, September 2013, pp. 28–33. *(Cited on pages 61 and 64)*

[258] A. Altinok and M. Turk, "Temporal integration for continuous multimodal biometrics," in *Proceedings of the Workshop on Multimodal User Authentication*. Citeseer, 2003. *(Cited on pages 61, 62, and 63)*

[259] T. Sim, S. Zhang, R. Janakiraman, and S. Kumar, "Continuous verification using multimodal biometrics," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 4, pp. 687–700, 2007. *(Cited on pages 61 and 62)*

[260] J. Muncaster and M. Turk, "Continuous multimodal authentication using dynamic bayesian networks," in *Proceedings 2nd Workshop Multimodal User Authentication, Toulouse, France*, 2006. *(Cited on page 61)*

[261] C. Shen, H. Zhang, Z. Yang, and X. Guan, "Modeling multimodal biometric modalities for continuous user authentication," in *Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on*. IEEE, 2016, pp. 001 894–001 899. *(Cited on pages 61 and 62)*

[262] H. Zhang, V. M. Patel, and R. Chellappa, "Robust multimodal recognition via multitask multivariate low-rank representations," in *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*, vol. 1. IEEE, 2015, pp. 1–8. *(Cited on page 62)*

[263] C. McCool, S. Marcel, A. Hadid, M. Pietikäinen, P. Matejka, J. Cernockỳ, N. Poh, J. Kittler, A. Larcher, C. Levy *et al.*, "Bi-modal person recognition on a mobile phone: using mobile phone data," in *Multimedia and Expo Workshops (ICMEW), 2012 IEEE International Conference on*. IEEE, 2012, pp. 635–640. *(Cited on pages 62 and 63)*

[264] L. Fridman, S. Weber, R. Greenstadt, and M. Kam, "Active authentication on mobile devices via stylometry, application usage, web browsing, and gps location," *IEEE Systems Journal*, vol. 11, no. 2, pp. 513–521, 2017. *(Cited on page 62)*

[265] Z. Sitová, J. Šeděnka, Q. Yang, G. Peng, G. Zhou, P. Gasti, and K. S. Balagani, "Hmog: New behavioral biometric features for continuous authentication of smartphone users," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 5, pp. 877–892, 2016. *(Cited on page 62)*

[266] H. Saevanee, N. Clarke, S. Furnell, and V. Biscione, "Text-based active authentication for mobile devices," in *IFIP International Information Security Conference*. Springer, 2014, pp. 99–112. *(Cited on page 62)*

[267] W. Shi, J. Yang, Y. Jiang, F. Yang, and Y. Xiong, "Senguard: Passive user identification on smartphones using multiple sensors," in *Wireless and Mobile Computing, Networking and Communications (WiMob), 2011 IEEE 7th International Conference on*. IEEE, 2011, pp. 141–148. *(Cited on page 62)*

[268] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song, "Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication," *IEEE transactions on information forensics and security*, vol. 8, no. 1, pp. 136–148, 2013. *(Cited on page 62)*

[269] S. Kumar, T. Sim, R. Janakiraman, and S. Zhang, "Using continuous biometric verification to protect interactive login sessions," in *Computer Security Applications Conference, 21st Annual*. IEEE, 2005, pp. 10–pp. *(Cited on page 62)*

[270] O. Riva, C. Qin, K. Strauss, and D. Lymberopoulos, "Progressive authentication: Deciding when to authenticate on mobile phones." in *USENIX Security Symposium*, 2012, pp. 301–316. *(Cited on pages 62 and 63)*

[271] M. Muaaz, "A transparent and continuous biometric authentication framework for user-friendly secure mobile environments," 2013. *(Cited on page 62)*

[272] M. Soltane, N. Doghmane, and N. Guersi, "Face and speech based multi-modal biometric authentication," *International Journal of Advanced Science and Technology*, vol. 21, no. 6, pp. 41–56, 2010. *(Cited on pages 62 and 63)*

[273] P. K. Atrey, M. A. Hossain, A. El Saddik, and M. S. Kankanhalli, "Multimodal fusion for multimedia analysis: a survey," *Multimedia systems*, vol. 16, no. 6, pp. 345–379, 2010. *(Cited on pages 62 and 63)*

[274] M. Wöllmer, M. Al-Hames, F. Eyben, B. Schuller, and G. Rigoll, "A multidimensional dynamic time warping algorithm for efficient multimodal fusion of asynchronous data streams," *Neurocomputing*, vol. 73, no. 1, pp. 366–380, 2009. *(Cited on pages 62 and 63)*

[275] S. Bengio, "Multimodal authentication using asynchronous hmms," in *AVBPA*. Springer, 2003, pp. 770–777. *(Cited on pages 62 and 63)*

[276] ——, "Multimodal speech processing using asynchronous hidden markov models," *Information Fusion*, vol. 5, no. 2, pp. 81–89, 2004. *(Cited on pages 62 and 63)*

[277] H. Bredin and G. Chollet, "Audio-visual speech synchrony measure for talking-face identity verification," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 2. IEEE, 2007, pp. II–233. *(Cited on pages 62 and 63)*

[278] J. N. Gowdy, A. Subramanya, C. Bartels, and J. Bilmes, "Dbn based multi-stream models for audio-visual speech recognition," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on*, vol. 1. IEEE, 2004, pp. I–993. *(Cited on page 63)*

[279] S. Dupont and J. Luettin, "Using the multi-stream approach for continuous audio-visual speech recognition: Experiments on the m2vts database," in *Proceedings 5th International Conference on Spoken Language Processing*, vol. 4, no. EPFL-CONF-82480, 1998, pp. 1283–1286. *(Cited on page 63)*

[280] N. Damer, F. Maul, and C. Busch, "Multi-biometric continuous authentication: A trust model for an asynchronous system," in *Information Fusion (FUSION), 2016 19th International Conference on*. IEEE, 2016, pp. 2192–2199. *(Cited on page 63)*

[281] K. Niinuma and A. K. Jain, "Continuous user authentication using temporal information," in *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 2010, pp. 76 670L–76 670L. *(Cited on page 64)*

[282] A. Mhenni, C. Rosenberger, E. Cherrier, and N. E. B. Amara, "Keystroke template update with adapted thresholds," in *Advanced Technologies for Signal and Image Processing (ATSIP), 2016 2nd International Conference on*. IEEE, 2016, pp. 483–488. *(Cited on page 64)*

[283] R. Donida Labati, V. Piuri, R. Sassi, G. Sforza, and F. Scotti, "Adaptive ecg biometric recognition: a study on re-enrollment methods for qrs signals," in *Proceedings of the IEEE Workshop on Computational Intelligence in Biometrics and Identity Management (CIBIM 2014)*, Orlando, FL, USA, December 2014, pp. 30–37. *(Cited on page 64)*

[284] I. Brosso, A. La Neve, G. Bressan, and W. V. Ruggiero, "A continuous authentication system based on user behavior analysis," in *Availability, Reliability, and Security, 2010. ARES'10 International Conference on*. IEEE, 2010, pp. 380–385. *(Cited on page 64)*

[285] A. Prakash and R. Mukesh, "A biometric approach for continuous user authentication by fusing hard and soft traits." *IJ Network Security*, vol. 16, no. 1, pp. 65–70, 2014. *(Cited on page 64)*

[286] A. Rattani, B. Freni, G. L. Marcialis, and F. Roli, "Template update methods in adaptive biometric systems: a critical review," in *Proceedings of the International Conference on Biometrics*. Springer, 2009, pp. 847–856. *(Cited on page 64)*

[287] A. Rattani, G. L. Marcialis, and F. Roli, "Biometric system adaptation by self-update and graph-based techniques," *Journal of Visual Languages & Computing*, vol. 24, no. 1, pp. 1–9, 2013. *(Cited on page 64)*

[288] R. Donida Labati, A. Genovese, E. Muñoz Ballester, V. Piuri, F. Scotti, and G. Sforza, "Emerging biometric technologies for automated border control gates," in *Proceedings of the 13th International Conference on Pattern Recognition and Information Processing (PRIP 2016)*. IAPR, 2016, pp. 1–6. *(Cited on page 68)*

[289] R. Donida Labati and F. Scotti, "Noisy iris segmentation with boundary regularization and reflections removal," *Image and Vision Computing, Iris Images Segmentation Special Issue*, vol. 28, no. 2, pp. 270 – 277, February 2010. *(Cited on page 68)*

[290] J. F. Allen, D. K. Byron, M. Dzikovska, G. Ferguson, L. Galescu, and A. Stent, "Toward conversational human-computer interaction," *Artificial Intelligence magazine*, vol. 22, no. 4, pp. 27–37, 2001. *(Cited on page 68)*

[291] C. I. Nass and S. Brave, *Wired for speech: How voice activates and advances the human-computer relationship*. MIT press, 2005, vol. 9, no. 1. *(Cited on page 68)*

[292] K.-A. Lee, A. Larcher, H. Thai, B. Ma, and H. Li, "Joint application of speech and speaker recognition for automation and security in smart home," in *Proceedings of the 12th Annual Conference of the International Speech Communication Association*, 2011, pp. 3317–3318. *(Cited on page 68)*

[293] M. Vacher, B. Lecouteux, J. S. Romero, M. Ajili, F. Portet, and S. Rossato, "Speech and speaker recognition for home automation: Preliminary results," in *Proceedings of the 8th International Conference on Speech Technology and Human-Computer Dialogue (SpeD)*. IEEE, 2015, pp. 1–10. *(Cited on page 68)*

[294] M. Hanmandlu, "Information sets and information processing," *Defence Science Journal*, vol. 61, no. 5, p. 405, 2011. *(Cited on page 70)*

[295] F. Sayeed and M. Hanmandlu, "Properties of information sets and information processing with an application to face recognition," *Knowledge and Information Systems*, pp. 1–23, 2017. *(Cited on page 70)*

[296] M. Aggarwal and M. Hanmandlu, "Representing uncertainty with information sets," *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 1, pp. 1–15, 2016. *(Cited on page 71)*

[297] F. Sayeed and M. Hanmandlu, "Three information set-based feature types for the recognition of faces," *Signal, Image and Video Processing*, vol. 10, no. 2, pp. 327–334, 2016. *(Cited on page 71)*

[298] M. Hanmandlu *et al.*, "Robust ear based authentication using local principal independent components," *Expert Systems with Applications*, vol. 40, no. 16, pp. 6478–6490, 2013. *(Cited on page 71)*

[299] C. N. Silla Jr and A. A. Freitas, "A survey of hierarchical classification across different application domains," *Data Mining and Knowledge Discovery*, vol. 22, no. 1-2, pp. 31–72, 2011. *(Cited on page 72)*

[300] K. C. A. R.Webb, "Statistical pattern recognition, 3rd edition." *(Cited on page 72)*

[301] C. M. Bishop, *Neural networks for pattern recognition*. Oxford university press, 1995. *(Cited on page 72)*

[302] L. Wang, *Support vector machines: theory and applications*. Springer Science & Business Media, 2005, vol. 177. *(Cited on page 72)*

[303] G. Ozbulak, Y. Aytar, and H. K. Ekenel, "How transferable are cnn-based features for age and gender classification?" in *IEEE International Conference of the Biometrics Special Interest Group (BIOSIG), 2016*. IEEE, 2016, pp. 1–6. *(Cited on pages 73 and 100)*

[304] R. Rothe, R. Timofte, and L. Van Gool, "Deep expectation of real and apparent age from a single image without facial landmarks," *International Journal of Computer Vision*, pp. 1–14, 2016. *(Cited on pages 73 and 100)*

[305] J. C. Chen, A. Kumar, R. Ranjan, V. M. Patel, A. Alavi, and R. Chellappa, "A cascaded convolutional neural network for age estimation of unconstrained faces," in *Proceedings of the 8th IEEE Conference on Biometrics Theory, Applications and Systems*. IEEE, 2016, pp. 1–8. *(Cited on pages 73 and 100)*

[306] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, and M. Bernstein, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015. *(Cited on page 74)*

[307] J. Pohjalainen, O. Räsänen, and S. Kadioglu, "Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits," *Computer Speech & Language*, vol. 29, no. 1, pp. 145–171, 2015. *(Cited on page 74)*

[308] A. Malhi and R. X. Gao, "Pca-based feature selection scheme for machine defect classification," *IEEE Transactions Instrum. Meas.*, vol. 53, no. 6, pp. 1517–1525, 2004. *(Cited on page 74)*

[309] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural networks*, vol. 6, no. 4, pp. 525–533, 1993. *(Cited on page 75)*

[310] R. Auckenthaler, M. Carey, and H. Lloyd-Thomas, "Score normalization for text-independent speaker verification systems," *Digital Signal Processing*, vol. 10, pp. 42–54, 2000. *(Cited on page 75)*

[311] N. Poh, "User-specific score normalization and fusion for biometric person recognition," *Advanced Topics in Biometrics*, pp. 401–418, 2007. *(Cited on page 75)*

[312] R. Batuwita and V. Palade, "Class imbalance learning methods for support vector machines," *Imbalanced learning: Foundations, algorithms, and applications*, pp. 83–99, 2013. *(Cited on page 79)*

[313] G. Shafer, *A mathematical theory of evidence*. Princeton university press, 1976, vol. 42. *(Cited on page 82)*

[314] K. Nguyen, S. Denman, S. Sridharan, and C. Fookes, "Score-level multibiometric fusion based on dempster–shafer theory incorporating uncertainty factors," *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 1, pp. 132–140, 2015. *(Cited on pages 82, 83, and 113)*

[315] A. Kumar and Y. Zhou, "Human identification using finger images," *IEEE Transactions on image processing*, vol. 21, no. 4, pp. 2228–2244, 2012. *(Cited on page 82)*

[316] L. Xu, A. Krzyzak, and C. Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," *IEEE transactions on systems, man, and cybernetics*, vol. 22, no. 3, pp. 418–435, 1992. *(Cited on page 82)*

[317] S. Sadjadi, M. Slaney, and L. Heck, "Msr identity toolbox vs 1.0," *Microsoft Research.[Online] Available at http://research. microsoft. com*, 2013. *(Cited on page 87)*

[318] M. Przybocki and A. Martin, "The nist year 2003 speaker recognition evaluation plan," 2003. *(Cited on page 92)*

[319] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1. IEEE, 2001, pp. I–I. *(Cited on page 96)*

[320] T. Hassner, S. Harel, E. Paz, and R. Enbar, "Effective face frontalization in unconstrained images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 4295–4304. *(Cited on page 96)*

[321] A. Dehghan, E. G. Ortiz, G. Shu, and S. Z. Masood, "Dager: Deep age, gender and emotion recognition using convolutional neural network," *arXiv preprint arXiv:1702.04280*, 2017. *(Cited on page 100)*

[322] J. Huang, B. Li, J. Zhu, and J. Chen, "Age classification with deep learning face representation," *Multimedia Tools and Applications*, pp. 1–17, 2017. *(Cited on page 100)*

[323] P. Rodríguez, G. Cucurull, J. M. Gonfaus, F. X. Roca, and J. Gonzàlez, "Age and gender recognition in the wild with deep attention," *Pattern Recognition*, vol. 72, pp. 563–571, 2017. *(Cited on page 100)*

[324] F. Agency, *Best Practice Technical Guidelines for Automated Border Control (ABC) Systems*. European Agency for the Management of Operational Cooperation at the External Borders of the Member States of the European Union, 2016. *(Cited on pages XVII, 101, 106, and 107)*

[325] ICAO, "ICAO Doc 9303: Machine Readable Travel Documents (seventh edition) - Part 9: Deployment of biometric identification and electronic storage of data in eMRTDs," 2015. *(Cited on page 101)*

[326] ISO/IEC, "ISO/IEC 19794 (all parts): Biometric data interchange formats," 2011. *(Cited on page 101)*

[327] CASIA, "Fingerprint Image Database Version 5.0." [Online]. Available: http://biometrics.idealtest.org/dbDetailForUser.do?id=7 *(Cited on page 103)*

[328] K. Nandakumar, Y. Chen, S. C. Dass, and A. Jain, "Likelihood ratio-based biometric score fusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 342–347, February 2008. *(Cited on pages 104 and 116)*

[329] R. Donida Labati, A. Genovese, E. Muñoz, V. Piuri, F. Scotti, and G. Sforza, "Biometric recognition in Automated Border Control: a survey," *ACM Comp. Surv.*, vol. 49, no. 2, pp. 24:1–24:39, 2016. *(Cited on page 109)*

[330] E. Carlos and A. Gilson, "A new ranking method for principal components analysis and its application to face image analysis," *Image and Vision Computing.*, vol. 28, no. 6, pp. 902–913, 2010. *(Cited on page 110)*

[331] ICAO, “Doc 9303 - Machine Readable Travel Documents - Part 9,” 2015. *(Cited on pages 110 and 111)*

[332] J. Sanchez del Rio, C. Conde, A. Tsitiridis, J. Raul Gomez, I. Martin de Diego, and E. Cabello, “Face-based recognition systems in the ABC e-gates,” in *Proceedings of the 2015 9th Annual IEEE International Systems Conference (SysCon)*, 2015, pp. 340–346. *(Cited on page 110)*

[333] R. Raghavendra, K. Raja, B. Yang, and C. Busch, “Automatic face quality assessment from video using gray level co-occurrence matrix: an empirical study on Automatic Border Control system,” in *Proceedings of International Conference on Pattern Recognirion*, 2014, pp. 438–443. *(Cited on page 110)*

[334] R. Raghavendra and C. Busch, “Improved face recognition by combining information from multiple cameras in Automatic Border Control system,” in *Proceedings of the 2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2015, pp. 1–6. *(Cited on page 110)*

[335] A. Martinez and R. Benavente, “The AR face database,” The Ohio State University, Tech. Rep. CVC Technical Report 24, 1998. *(Cited on page 110)*

[336] R. Cappelli, M. Ferrara, A. Franco, and D. Maltoni, “Fingerprint verification competition 2006,” *Biometric Technology Today*, vol. 15, no. 7-8, pp. 7–9, 2007. *(Cited on page 110)*

[337] J. Jang and H. Kim, “Performance measures,” in *Encyclopedia of Biometrics*. Springer US, 2009, pp. 1062–1068. *(Cited on page 111)*

[338] S. Mika, G. Rätsch, J. Weston, B. Schölkopf, and K.-R. Müller, “Fisher discriminant analysis with kernels,” in *Proceedings of the 1999 IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing (NNSP)*, 1999, pp. 41–48. *(Cited on pages 111 and 113)*

[339] N. Damer, A. Opel, and A. Nouak, “Biometric source weighting in multi-biometric fusion: towards a generalized and robust solution,” in *Proceedings of the 2014 22nd European Signal Processing Conference (EUSIPCO)*, 2014, pp. 1382–1386. *(Cited on pages 111 and 113)*

A

PUBLICATIONS

Some ideas and significant results present in this thesis were published in:

1. *"Age Estimation Based on Face Images and Pre-trained Convolutional Neural Networks"*
   Abhinav Anand, R. Donida Labati, A. Genovese, Enrique Muñoz, V. Piuri,
   and F. Scotti

   Proceedings of the 2017 IEEE Symp. on Computational Intelligence in Biometrics and Identity Management (CIBIM 2017), November 27 to December 1, Honolulu, Hawaii, USA.

   ABSTRACT: Age estimation based on face images plays an important role in a wide range of scenarios, including security and defense applications, border control, human-machine interaction in ambient intelligence applications, and recognition based on soft biometric information. Recent methods based on deep learning have shown promising performance in this field. Most of these methods use deep networks specifically designed and trained to cope with this problem. There are also some studies that focus on applying deep networks pre-trained for face recognition, which perform a fine-tuning to achieve accurate results. Differently, in this paper, we propose a preliminary study on increasing the performance of pre-trained deep networks by applying post-processing strategies. The main advantage with respect to fine-tuning strategies consists of the simplicity and low computational cost of the post-processing step. To the best of our knowledge, this paper is the first study on age estimation that proposes the use of post-processing strategies for features extracted using pre-trained deep networks. Our method exploits a set of pre-trained Convolutional Neural Networks (CNNs) to extract features from the input face image. The method then performs a feature level fusion, reduces the dimensionality of the feature space, and estimates the age of the individual by using a Feed-Forward Neural Network (FFNN). We evaluated the performance of our method on a public dataset (Adience Benchmark of Unfiltered Faces for Gender and Age Classification) and on a dataset of non-ideal samples affected by controlled rotations, which we collected in our laboratory. Our age estimation method obtained better or comparable results with respect to state-of-the-art techniques and achieved satisfactory performance in non-ideal conditions. Results also showed that CNNs trained on general datasets can obtain satisfactory accuracy for different types of validation images, also without applying fine-tuning methods.

2. *"Text-Independent Speaker Recognition for Ambient Intelligence Applications by Using Information Set Features"*

   Abhinav Anand, R. Donida Labati, M. Hanmandlu, V. Piuri, and F. Scotti Proceedings of the 2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA 2017), Annecy, France, July 26-28, 2017.

   ABSTRACT: Biometric systems are enabling technologies for a wide set of applications in Ambient Intelligence (AmI) environments. In this context, speaker recognition techniques are of paramount importance due to their high user acceptance and low required cooperation. Typical applications of biometric recognition in AmI environments are identification techniques designed to recognize individuals in small datasets. Biometric recognition methods are frequently deployed on embedded hardware and therefore need to be optimized in terms of computational time as well as used memory. This paper presents a text-independent speaker recognition method particularly suitable for identification in AmI environments. The proposed method first computes the Mel Frequency Cepstral Coefficients (MFCC) and then creates Information Set Features (ISF) by applying a fuzzy logic approach. Finally, it estimates the user's identity by using a hierarchical classification technique based on computational intelligence. We evaluated the performance of the speaker recognition method using signals belonging to the NIST-2003 switchboard speaker database. The achieved results showed that the proposed method reduced the size of the template with respect to traditional approaches based on Gaussian Mixture Models (GMM) and achieved better identification accuracy.

3. *"Enhancing fingerprint biometrics in Automated Border Control with adaptive cohorts"*

   Abhinav Anand, R. Donida Labati, A. Genovese, Enrique Muñoz, V. Piuri,
   F. Scotti, and G. Sforza

   Proceedings of the 2016 IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA 2016), Athens, Greece, pp. 1-8, December 6-9, 2016

   ABSTRACT: Automated Border Control (ABC) systems are being increasingly used to perform a fast, accurate, and reliable verification of the travelers' identity. These systems use biometric technologies to verify the identity of the person crossing the border. In this context, fingerprint verification systems are widely adopted due to their high accuracy and user acceptance. Matching score normalization methods can improve the performance of fingerprint recognition in ABC systems and mitigate the effect of non-idealities typical of this scenario without modifying the existing biometric technologies. However, privacy protection regulations restrict the use of biometric data captured in ABC systems and can compromise the applicability of these techniques. Cohort score normalization methods based only on impostor scores provide a suitable solution, due to their limited use of sensible data and to their promising performance. In this paper, we propose a privacy-compliant and adaptive normalization approach for enhancing fingerprint recognition in ABC systems. The proposed approach computes cohort scores from an external public dataset and uses computational intelligence to learn and improve the matching score distribution. The use of a public dataset permits to apply cohort normalization strategies in contexts in which privacy protection regulations restrict the storage of biometric data. We performed a technological and a scenario evaluation using a commercial matcher currently adopted in real ABC systems and we used data simulating different conditions typical of ABC systems, obtaining encouraging results.

4. *"Enhancing the Performance of Multimodal Automated Border Control Systems"*

   Abhinav Anand, Ruggero Donida Labati, Angelo Genovese, Enrique Muñoz, Vincenzo Piuri, Fabio Scotti, and Gianluca Sforza

   Proceedings of the 15th IEEE Conference on Biometrics Special Interest Group (BIOSIG 2016), Darmstadt, Germany, pp. 1-5, September 21-23, 2016.

   ABSTRACT: Biometric recognition in Automated Border Control (ABC) systems is performed in response to an increased worldwide traffic, by automatically verifying the identity of the passenger during border crossing. Currently, ABC systems seldom use methods for multimodal biometric fusion, which have been proved to increase the recognition accuracy, due to technological and privacy limitations. This paper proposes a framework for the biometric fusion in ABC systems, with the features of being technology-neutral and privacy-compliant, by performing an analysis of the most suitable biometric fusion techniques for ABC systems and considering the current technical and legal limitations.