Cambridge University Press & Assessment
978-0-521-11596-4 — Multibiometrics for Human Identification
Edited by Bir Bhanu , Venu Govindaraju
Excerpt
More Information

Introduction

Biometric systems are increasingly being used in many applications that require positive identification of individuals for access authorization. Traditional biometric systems rely on a single-biometric, single-sensor paradigm for authentication or identification. For high-security and real-world requirements, this paradigm is inadequate when it comes to reliably providing a high level of accuracy performance. Multibiometrics, the technique of using multiple biometric modalities and sensors, promises to rise to the challenge of making biometric authentication truly robust and reliable. Moreover, using multiple biometrics enhances the coverage of the section of the population that is not able to provide any single biometrics. Multiple biometrics is naturally more robust against spoof attacks as well, since hackers have to contend with more than one biometrics. Further, fusing multibiometrics enables indexing of large databases for identification of individuals.

Compared with other books on the same topic, the key features of this book are the following:

1. It includes the entire gamut of multibiometrics topics, including multimodal, multisensory levels of fusion, multiple algorithms, and multiple data acquisition instances.
2. It includes chapters on the latest sensing devices for novel multibiometrics modalities, security assessment of multibiometrics systems and their dynamic management, and theoretically sound and novel approaches for fusion.
3. It provides information on publicly available multibiometrics databases and addresses research issues related to performance modeling, prediction, and validation of multibiometrics systems.

The various issues related to multibiometrics systems can be placed into the following five categories:

1. *Multimodal and multisensory biometric systems:* Examples of multimodal systems are face and gait, face and voice, face and ear, fingerprint and face, face and lip movements, pulse and heart rate, etc. Examples of multisensory systems include systems based on color video, infrared video, multispectral imagery, range, and acoustics.
2. *Fusion methods in multibiometric systems:* Fusion in multibiometrics systems occurs at pixel, feature, decision, rank and match score levels. It includes fusions of soft (gender, age, ethnicity, height, weight, etc.) and hard (face, fingerprint, ear, etc.) biometrics as well as the integration of contextual information with soft biometrics. It also includes multiple algorithms based on multiple methods for feature extraction, feature selection, and matching.
3. *Hybrid biometric systems:* These systems may include multiple samples and multiple instances of a biometric modality. Examples of such systems are face recognition with changing pose in video, multiple instances of fingerprints, multiview gait, and various other combinations of biometrics systems.
4. *Database and security of multibiometric systems:* Multimodal databases are required to develop and validate robust fusion algorithms. The security of such biometric systems is also important to guard them from hackers and protect the privacy of individuals.
5. *Performance of multibiometric systems:* Methods are needed for performance evaluation and prediction on large populations to establish a scientific foundation of multibiometric systems.

This edited book consists of 15 chapters distributed among the five categories described above. Four chapters discuss systems that combine multiple biometric modalities and/or multiple biometric sensors to build composite authentication systems. Three chapters focus on the strategies involved in fusing features and scores when using multiple modalities or sensors. Four chapters deal with hybrid systems that attempt to tap novel uses of current technologies for biometric authentication. Two chapters cover datasets and security concepts for biometrics. The final two chapters investigate aspects of measuring performance of systems used for biometric authentication.

The following section elaborates on the chapters falling into each of the five categories.

Part I: Multimodal and Multisensor Biometric Systems

Awareness in the biometrics community is growing regarding the benefits of using multiple modalities and sensors for authentication. Chapter 1 on multimodal

978-0-521-11596-4 — Multibiometrics for Human Identification
Edited by Bir Bhanu , Venu Govindaraju
Excerpt
More Information

modeling using face and ear biometrics for recognition uses a combination of two-dimensional (2D) face recognition and three-dimensional (3D) ear recognition. Models for the ear and face are extracted and enrolled in a database. Match scores for the two modalities are fused, and the system achieves a 100 percent rank-one identification rate. Chapter 2 on audiovisual speech synchrony detection tackles the problem of detecting synchrony between audio and video segments. The inadequacies of existing synchrony detection methods are highlighted, and a new approach that uses a time-evolution model is presented. The technique performs much better than the baseline set of existing techniques. Chapter 3 on multispectral contact-free palmprint recognition presents techniques that perform image acquisition, preprocessing, and image fusion for palmprint recognition using a contact-free multispectral sensor. Results are presented for performance on pixel- and score-level fusion that show that contact-free palmprint recognition can deliver high performance. Chapter 4 on beneath-the-skin face recognition presents novel feature extraction and matching techniques for thermal infrared imagery of the face. As opposed to traditional face recognition that captures a picture of the exterior of the face, thermal infrared imaging captures subsurface facial features that can be used for robust matching.

Part II: Fusion Methods in Multibiometric Systems

With multiple modalities and sensors comes the challenge of reliably combining the scores that are output by the various components. Chapter 5 considers the issues involved in fusing information from several biometric sources (including multiple modalities, multiple features, and multiple matchers) and presents a copula-based approach to fusion. A detailed primer on the statistical theory of copulas is also presented, and tests are conducted on the NIST datasets. Chapter 6 presents a feature-level fusion of face and fingerprint biometrics. The framework presented is based on a robust set of features extracted from a scale-space filtering of the raw face and fingerprint images. Results are presented on both synthetic and real datasets. Chapter 7 on adaptive biometric systems discusses the issue of the evolution of biometric templates over time and presents a template update model.

Part III: Hybrid Biometric Systems

Part III includes four chapters. Chapter 8 on a multiple projector-camera system for 3D gait recognition presents a system that can capture 3D human body measurements and applies these data to gait recognition. It utilizes 12 fast 3D cameras that can capture 1 million 3D points in 1 to 2 seconds. Chapter 9 on gait recognition presents a computational framework that integrates the physics of motion with the neurobiological basis for human perception. Results are reported on the USF gait dataset. Chapter 10 on face tracking and recognition in a camera network presents a technique that extracts rotation-invariant features from multiple calibrated-camera views of a face. The features are computed using spherical harmonics of a texture map, and they show good class-separation properties. Chapter 11 on a bidirectional relighting algorithm for 3D-aided 2D face recognition uses 3D data to perform robust 2D face recognition under severely degraded image acquisition conditions. A new 3D face dataset was collected for the purpose of testing the technique, which has been made available by the authors.

Part IV: Databases and Security

Chapter 12 describes the process of constructing a large multimodal dataset comprising of gait, ear, and semantic data. Semantic biometrics relies on descriptions provided manually using visual cues to define a biometric template for an individual. An analysis of the data collected is also presented. Chapter 13 presents dynamic multibiometrics security and the need for high-security application multibiometric systems that adaptively adjust to varying security requirements. A new score-level approach to ensure multibiometrics security is also presented.

Part V: Performance of Multibiometric Systems

To be useful in real-world scenarios, biometric systems need to operate under very high performance requirements. Chapter 14 presents two theoretical approaches to predict the performance of biometric fusion systems. These approaches allow for the selection of optimal combination of biometrics. Prediction tests are conducted on the publicly available XM2VTS and other multibiometrics databases. Finally, chapter 15 explores the problem of predicting (closed set) identification performance of biometric matchers in large-scale systems. Two major causes of prediction errors are identified. A novel score-resampling method that overcomes the binomial approximation effect is presented, and the score-mixing effect is reduced by using score selection based on identification trial statistics. Results showing the accuracy of the techniques are shown for the NIST biometric score dataset.

In summary, this edited book on *Multibiometrics for Human Identification* addresses a broad spectrum of research issues ranging from different sensing modes and modalities to fusion of biometrics samples and combination of algorithms. It covers publicly available multibiometrics databases and theoretical and empirical studies on sensor fusion techniques in the context of biometrics authentication, identification, and performance evaluation and prediction.

PART I

Multimodal and Multisensor Biometric Systems

1

Multimodal Ear and Face Modeling and Recognition

Steven Cadavid, Mohammad H. Mahoor, and Mohamed Abdel-Mottaleb

1.1 Introduction

Biometric systems deployed in current real-world applications are primarily unimodal – they depend on the evidence of a single biometric marker for personal identity authentication (e.g., ear or face). Unimodal biometrics are limited, because no single biometric is generally considered both sufficiently accurate and robust to hindrances caused by external factors (Ross and Jain 2004).

Some of the problems that these systems regularly contend with are the following: (1) Noise in the acquired data due to alterations in the biometric marker (e.g., surgically modified ear) or improperly maintained sensors. (2) Intraclass variations that may occur when a user interacts with the sensor (e.g., varying head pose) or with physiological transformations that take place with aging. (3) Interclass similarities, arising when a biometric database comprises a large number of users, which results in an overlap in the feature space of multiple users, requires an increased complexity to discriminate between the users. (4) Nonuniversality – the biometric system may not be able to acquire meaningful biometric data from a subset of users. For instance, in face biometrics, a face image may be blurred because of abrupt head movement or partially occluded because of off-axis pose. (5) Certain biometric markers are susceptible to spoof attacks – situations in which a user successfully masquerades as another by falsifying their biometric data.

Several of the limitations imposed by unimodal biometric systems can be overcome by incorporating multiple biometric markers for performing authentication. Such systems, known as multimodal biometric systems, are expected to be more reliable because of the presence of multiple (fairly) independent pieces of evidence (Kuncheva et al. 2000). These systems are capable of addressing the aforementioned shortcomings inherent to unimodal biometrics. For instance, the likelihood of acquiring viable biometric data increases with the number of sensed biometric markers. They also deter spoofing because it would be difficult for an impostor to spoof multiple biometric markers of a genuine user concurrently. However, the incorporation of multiple biometric markers can also lead to additional complexity in the design of a biometric system. For instance, a technique known as data fusion must be employed to integrate multiple pieces of evidence to infer identity. In this chapter, we present a method that fuses the three-dimensional (3D) ear and two-dimensional (2D) face modalities at the match score level. Fusion at this level has the advantage of utilizing as much information as possible from each biometric modality (Snelick et al. 2005).

Several motivations can be given for a multimodal ear and face biometric. First, the ear and face data can be captured using conventional cameras. Second, the data collection for face and ear is nonintrusive (i.e., requires no cooperation from the user). Third, the ear and face are in physical proximity, and when acquiring data of the ear (face), the face (ear) is frequently encountered as well. Often, in an image or video captured of a user's head, these two biometric markers are jointly present and are both available to a biometric system. Thus, a multimodal face and ear biometric system is more feasible than, say, a multimodal face and fingerprint biometric system.

For more than three decades, researchers have worked in the area of face recognition (Jain et al. 2007). Despite the efforts made in 2D and 3D face recognition, it is not yet ready for real-world applications as a unimodal biometric, system. Yet the face possesses several qualities that make it a preferred biometric, including being nonintrusive and containing salient features (e.g., eye and mouth corners).

The ear, conversely, is a relatively new area of biometric research. A few studies have been conducted using 2D data (image intensity) (Burge and Burger 1998, 2000; Chang et al. 2003; Abdel-Mottaleb and Zhou 2006; Yan and Bowyer 2007) and 3D shape data (Abdel-Mottaleb and Zhou 2006; Cadavid and Abdel-Mottaleb 2008). Initial case studies have suggested that the ear has sufficient unique features to allow a positive and passive identification of a subject (Ianarelli 1989). Furthermore, the ear is known to maintain a consistent structure throughout a subject's lifetime. Medical literature has shown proportional ear growth after the first four months of birth (Ianarelli 1989). Ears may be more reliable than faces, which research has shown are prone to erroneous identification because of the ability of a subject to change their facial expression or otherwise manipulate their visage. However, some drawbacks are inherent to ear biometrics. One such drawback, which poses difficulty to the feature extraction process, is occlusion due to hair or jewelery (e.g., earrings or the arm of a pair of eyeglasses).

Based on the above discussion, we present a multimodal ear and face biometric system. For the ear recognition component, first, a set of frames is extracted from a video clip. The ear region contained within each frame is localized and segmented. The 3D structure of each segmented ear region is then derived using a linearized Shape from Shading (SFS) technique (Tsai and Shah 1994), and each resulting model is globally aligned. The 3D model that exhibits the greatest overall similarity to the other models in the set is determined to be the most stable model in the set. This 3D model is stored in the database and utilized for 3D ear recognition.

For the face recognition component, we are inspired by our previous work in 2D face recognition using the Gabor filter component of our attributed relational graph method (Mahoor and Abdel-Mottaleb 2008; Mahoor et al. 2008). We utilize a set of Gabor filters to extract a suite of features from 2D frontal facial images. These features, termed attributes, are extracted at the location of facial landmarks, which have been extracted using the Active Shape Model (ASM) (Mahoor and Abdel-Mottaleb 2006). The attributes of probe images and gallery images are employed to compare facial images in the attribute space.

In this chapter, we present a method for fusing the ear and face biometrics at the match score level. At this level, we have the flexibility to fuse the match scores from various modalities upon their availability. First, the match scores of each modality are calculated. Second, the scores are normalized and subsequently combined using a weighted sum technique. The final decision for recognition of a probe face is made on the fused match score.

The remainder of this chapter is organized as follows: Section 1.2 discusses previous work in 3D ear recognition, 2D face recognition, and multimodal ear and face recognition. Section 1.3 presents our approach for 3D ear modeling and recognition from video sequences. Section 1.4 outlines our method for 2D face recognition using Gabor filters. Section 1.5 describes the technique for data fusion at the match score level. Section 1.6 shows the experimental results using the West Virginia University database to validate our algorithm and test the identification and verification performances. Last, conclusions and future work are given in Section 1.7. As a reference, a summary of the acronyms used in this chapter is provided in Table 1.2.

1.2 Related Work

In this section, we will briefly outline some of the prominent works in 3D ear recognition, 2D face recognition, and multimodal ear and face recognition. It is worth noting that a direct comparison between the performances of different systems is difficult and at times can be misleading because datasets may be of varying sizes, the image resolution and the amount of occlusion contained within the region-of-interest (ROI) may be different, and some may use a multi-image gallery for a subject whereas others use a single-image gallery.

1.2.1 3D Ear Recognition

Three-dimensional ear biometrics is a relatively new area of research. A few studies have been conducted, and most of the related work has been based on ear models acquired using 3D range scanners. To the best of our knowledge, we are the first to develop a 3D ear recognition system that obtains 3D ear structure from an uncalibrated video sequence. In this section, we will review the literature on 3D ear reconstruction from multiple views and 3D ear recognition.

Liu et al. (2006) describe a 3D ear reconstruction technique using multiple views. This method uses the fundamental matrix and motion estimation techniques to derive the 3D shape of the ear. The greatest difficulty with this approach is obtaining a set of reliable feature point correspondences because of the lack of texture on the ear surface. They first use the Harris corner criteria to detect salient features in each image and apply correlation matching. Then they use Random Sample Consensus (RANSAC) (Fischler and Bolles 1981) to eliminate outliers from the set of detected features. The authors report that automatically extracting feature points in this way yields poor results. Therefore, a semiautomatic approach is used that allows the user to manually relocate feature points that are poorly matched.

Bhanu and Chen (2003) present an ear recognition system using a local surface patch (LSP) representation and the Iterative Closest Point (ICP) algorithm. They used 3D ear range images obtained from the University of California at Riverside (UCR) dataset as well as the Notre Dame collection. The UCR collection comprises 155 subjects with 902 images, and the Notre Dame collection comprises 302 subjects. They report a rank-one recognition rate of 100%. Bhanu and Chen (2005, 2008) also developed an algorithm for ear matching by using a two-step ICP approach. The first step includes detecting and aligning the helixes of both the gallery and probe ear models. Second, a series of affine transformations are applied to the probe model to optimally align the two models. The root-mean-square distance (RMSD) is used to measure the accuracy of the alignment. The identity of the gallery model that has the smallest RMSD value to the probe model is declared the identity of the probe model. They report that out of a database of 30 subjects, 28 were correctly recognized. Bhanu and Chen (2004) also propose a method for detecting an ear region from a 3D range image. Their algorithm is based on a two-step