Harry Wechsler

Reliable Face Recognition Methods

System Design, Implementation and Evaluation

Springer

Reliable Face Recognition Methods

System Design, Implementation and Evaluation

Reliable Face Recognition Methods

System Design, Implementation and Evaluation

by

Harry Wechsler
George Mason University, USA

Springer

Harry Wechsler
George Mason University
Dept. of Computer Science
Fairfax, VA 22030
wechsler@gmu.edu

Library of Congress Control Number: 2006932031

**Reliable Face Recognition Methods:
System Design, Implementation and Evaluation**
by *Harry Wechsler*

ISBN-13: 978-0-387-22372-8
ISBN-10: 0-387-22372-X
e-ISBN-13: 978-0-387-38464-1
e-ISBN-10: 0-387-38464-2

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

With love to my children
Gabriela and Marc

Contents

Preface

Science is a way to teach how something gets to be known, what is known, to what extent things are known (for nothing is known absolutely), how to handle doubt and uncertainty, what the rules of evidence are, how to think about things so that judgments can be made, how to distinguish truth from fraud, and from show (Richard Feynman)

One of the grand challenges for computational intelligence is to understand how people process and recognize each other's face and to develop automated and reliable face recognition systems. This challenge underlies *biometrics*, the science of authenticating people by measuring their physical or external appearance and/or their behavioral or physiological traits. The physical and behavioral traits are not necessarily independent. The face we look at is a mix of both physical characteristics and emotive expressions. Face recognition has become a major biometric technology. Solving the face recognition problem will have a major scientific impact, as recognizing people is a first but critical step towards building intelligent machines that can function in human environments. "The ability to recognize living creatures in photographs or video clips is a critical enabling technology for a wide range of applications including defense, health care, human-computer interaction, image retrieval and data mining, industrial and personal robotics, surveillance and security, and transportation. Despite 40 years of research, however, today's recognition systems are still largely unable to handle the extraordinary wide range of appearances assumed by common objects [including faces] in typical images" (*Designing Tomorrow's Category - Level 3D Object Recognition Systems* [1]).

Biometrics has become the major component in the complex decision-making process associated with security applications. Key concerns related to accuracy and performance, benefits versus costs, information assurance, and security over privacy have surfaced and have yet to be resolved. Skepticism, the heart of scientific method, is needed to ferret out fact from fiction regarding what biometrics can actually do and to what extent. Advancing the field of biometrics for homeland security has taken on a sense of urgency in the post 9/11 world. Even though people can detect and identify faces with little or no effort, building an automated system for such purposes has proven elusive as reliable solutions have yet to be found. The all-encompassing *Face in a Crowd* biometric problem addresses both face detection and face recognition in cluttered environments. Biometric systems have to take into account the dynamic changes that affect the visual stimuli, including variability in the geometry of image formation, such as facial pose and distance from the camera, and illumination. Other factors that affect face recognition include facial expression due to emotion, occlusion and disguise, temporal changes and aging, and last but not least, the lack of adequate training data for learning how to represent and encode human faces.

A few major edited books treat face recognition. Among them are the first and seminal "*Face Recognition: From Theory to Applications*" (Wechsler et al., Springer, 1998), and most recently the "*Handbook of Face Recognition*" (Li and Jain, Springer, 2005). This book is the first to comprehensively address the face recognition problem in its entirety, while drawing inspiration and gaining new insights from complementary fields of endeavor, such as neurosciences, statistics, signal and image processing, computer vision, machine learning and pattern recognition, and statistical learning. The various chapters treat topics related to how people represent, process and/or respond to the human face, modeling and prediction, the face space, identification and verification, face detection, tracking and recognition, 3D, data fusion, denial and de-

[1] http://lear.inrialpes.fr/people/schmid/workshop.html

ception using occlusion and disguise, performance evaluation and error analysis, and finally, competing security and privacy considerations.

The underlying theme of the book is that the biometric inputs chart continuous and coherent space and time manifolds, which facilitate their recognition. Face recognition is dynamic rather than static. It continuously iterates, making specific interpretations and assigning confidence to them. Supporting and non-accidental evidence is accrued in an active fashion, leading to lower uncertainty in the recognition decisions made, and resolving ambiguity, if any. Integral to face recognition are advances in pattern recognition. Novel methods are proposed here to handle real life applications where variability, incomplete, noisy, distorted and/or disguised patterns are usually the norm rather than the exception. The overall goal of the book is *applied modern pattern recognition*, with the understanding that the novel methods described here apply to any objects. The face pattern is only one of the object patterns that surround us and need to be recognized. The scope for pattern recognition (Rosenfeld and Wechsler, 2000) is much wider here because among other things both training and testing can take place using incomplete or camouflaged/disguised patterns drawn from single and/or multiple image sets.

The emphasis throughout the book is on proper modeling and prediction. Gregory Chaitin, in the March 2006 issue of the Scientific American, recalls Gottfried Leibniz's 1685 philosophical essay *Discourse de métaphysique* (Discourse on Metaphysics). The essay discusses how one can distinguish between facts that can be described by some law and those that are lawless, irregular facts. Leibniz observed that "a theory has to be simpler than the data it explains, otherwise it does not explain anything. The concept of a law becomes vacuous if arbitrarily high mathematical complexity is permitted, because then one can always construct a law no matter how random and patternless the data really are." The corollary for Chaitin is that "a useful theory is a compression of the data; comprehension is compression." Modeling and prediction, the hallmarks of learning, can be implemented using novel methods driven by semi-supervised learning and transduction, exchangeability and rank order, and martingale. Overall, the book articulates new but promising directions for pattern recognition, while providing the motivation for innovative ways to approach the face recognition challenge.

The title of the book, *Reliable Face Recognition Methods*, refers to the normal expectation one has that face recognition should display robust performance despite suboptimal and/or adverse image acquisition conditions or lack of adequate training data. Even the top-ranked face recognition engines still reject legitimate subjects, while letting impostors pass through. "Reliable," throughout the book, means the ability to deploy consistent, dependable, large-scale and full-fledged operational biometric systems, which is the true hallmark of a mature technology. To that end, a large data base of facial images, such as FERET, is required to test and assess competing technologies. FERET, which was designed and developed at George Mason University under my direction, has become the standard data base used by researchers worldwide for R&D and benchmark studies on face recognition. Science needs to be replicated and tested for validation.

This book can serve both as an interdisciplinary text and as a research reference. Each chapter provides the background and impetus for understanding the problems discussed and the approach taken to solve them. The book can benefit advanced undergraduates ("senior") and graduates taking courses on pattern recognition or biometrics; scientists and practitioners interested in updating their knowledge; and government and industry executives charged with addressing ever-evolving biometric security requirements.

My gratitude goes to many people. Many thanks go to Professor Jack Sklansky, who introduced me to the field of pattern recognition. Much appreciation goes to my former doctoral students Srinivas Gutta, Shen-Shyang Ho, Jeffrey Huang, Fayin Li, and Albert Pujol, with whom I had many productive and rewarding collaborations. I am also grateful for the help and inspiration for this book from Josef Bigun, Vladimir Cherkassky, Clifford Claussen, Victor Chen, Stephen McKenna, Matt Matsuda, and

Barnabas Takacs. My thanks go also to the many people referenced throughout the book from whom I have drawn knowledge and motivation. From my brother Tobi, I learned to appreciate and love the visual arts, which led me to explore the scientific basis for perception-representation and interpretation. Thanks go also to my sister-in-law Nobuko for her friendship and kindness. Last but not least, heartfelt thanks go to my wife Michele for her encouragement and help with completing this book, and to my children Gabriela and Marc for the sparks in their eyes and their smiling faces.

Harry Wechsler
June 2006

1. Introduction

The first rule was never to accept anything as true unless I recognized it to be certainly and evidently such: that is, carefully to avoid all precipitation and prejudgment, and to include nothing in my conclusions unless it presented itself so clearly and distinctly to my mind that there was no reason or occasion to doubt it. The second was to divide each of the difficulties which I encountered into as many parts as possible, and as might be required for an easier solution. The third was to think in an orderly fashion when concerned with the search for truth, beginning with the things which were simplest and easiest to understand, and gradually and by degrees reaching toward more complex knowledge, even treating, as though ordered, materials which were not necessarily so. The last was, both in the process of searching and in reviewing when in difficulty, always to make enumerations so complete and reviews so general, that I would be certain that nothing was omitted.

From Discourse on Method and Meditations by Ren Descartes (1641) (translated by Laurence J. Lafleur and published by Liberal Arts Press, 1960)

Face recognition (Samal and Iyengar, 1992; Chellappa et al., 1995; Daugman, 1997; Jain et al., 1999; Zhao et al., 2003; Bolle et al., 2004; Li and Jain, 2005) has become a major biometric technology. Biometrics involve the automated identification or authentication from personal physical appearance or behavioral traits. Human physical appearance and/or behavioral characteristics are counted as biometrics as long as they satisfy requirements that include universality, distinctiveness or uniqueness, permanence or invariance, collectability, and acceptability (Clarke, 1994). The early belief in the uniqueness aspect of faces (to preempt forgeries) was one of the reasons behind their use, e.g., the face of Queen Victoria on the early stamps (Samal and Iyengar, 1992). Biometric systems, including face recognition systems, can be categorized according to their intended applications. According to Wayman (1999) a suitable self-evident taxonomy will include cooperative vs. non-cooperative, overt vs. covert, habituated vs. non-habituated, attended vs. non-attended, standard vs. non-standard operating environments, and public vs. private.

This book addresses the above taxonomy as it discusses face recognition along the complementary dimensions of science, (information) technology and engineering, culture and society, and visual arts. It is about science because it aims to understand and systematize the fundamental principles behind face processing. Face processing is an all-encompassing term that involves everything that facilitates face recognition, e.g., image capture, enrollment, and face detection and tracking. The scientific dimension is related to the basic research that supports technological progress. The book is about technology and engineering, because it deals with applied science and research aimed at practical ends, e.g., designing and building reliable face recognition systems. It is about culture and society because they affect the role the human face plays in our interactions. The book is also about the visual arts because the human figure has always occupied an important place in personal expression and contemplation. Art connects internal and external realities, provides for new perspectives of the world, and seeks for the ultimate truth and permanent essence embodied by fixed icons such as human faces and ideals. The arts activate abstraction and creativity and can stimulate innovative face recognition research, e.g., using the golden ratio template of human beauty for face recognition-by-parts (See Sect. 6.5 and 9.7). Last but not least, the book is about building completely automatic and full-fledged biometric systems that consider the full range of the face recognition sub-tasks, starting with data acquisition and enrollment and ending with different face authentication scenarios. The ever expanding scope of face recognition includes field operation rather than controlled in vitro lab conditions. This should lead to building search and analysis biometric video engines able to recognize people and/or interpret their activities and intentions from live-fed CCTV.

The book, multidisciplinary and syncretic, frames a modern research agenda for pattern recognition, in general, and face recognition, in particular. The modernity aspect refers to the scope of the enterprise. The book identifies real problems and motivates the need for large scale pattern recognition systems that can handle human diversity, temporal changes, and occlusion and disguise. The book, selective rather than encyclopedic, introduces new topics that require further investigation. It differentiates and motivates among existing problems and their proffered solutions, places emphasis on common threads, and focuses on what is most important. In particular, the book aims to fuse and reconcile the specific disciplines of image and signal processing, computer vision, machine learning and pattern recognition, while charting new but promising research directions.

Computer vision is about "computing properties of the 3D world from one or more digital images. As the name suggests, it involves computers interpreting images. Image analysis and/or understanding are synonyms for computer vision. Image processing and pattern recognition are disciplines related but not identical to computer vision. Image processing concerns image properties and image-to-image transformations, while pattern recognition [involves] recognizing and classifying objects using digital images" (Trucco and Verri, 1998). Learning, which is about generalization and prediction, lies at the interface between computer vision and pattern recognition. It plays a fundamental role in facilitating "the balance between internal representations and external regularities" (Nayar and Poggio, 1996). Face recognition requires new and robust learning paradigms. This includes 'good' classification methods that can work with only limited training data, which was acquired under fairly flexible and general assumptions. "The fewer assumptions a [computer vision] system imposes on its operational conditions, the more robust it is considered to be" (Moeslund and Granum, 2001).

The challenges confronting face recognition are many. First and foremost there is much variability in the image formation process that includes geometry, illumination, occlusion and disguise, and temporal changes (see Fig. 1.1). Even the faces of "identical" twins are different to some extent (see 1.1a). Biometrics in general, and face recognition, in particular, bear directly on the use of forensics in the courts of law. In a provocative *Science* editorial, titled "*Forensic Science: Oxymoron?*" Donald Kennedy, the Editor-in-Chief, makes the obvious point that the reliability of forensics "is unverified either by statistical models of [biometric] variation or by consistent data on error rates. Nor does the problem with forensic methods ends there. Processing and enhancement of such images could mislead jurors who believe they are seeing undoctored originals." Following the 1993 U.S. Supreme Court's Daubert case, the Court "did list several criteria for qualifying expert testimony: peer review, error rate, adequate testing, regular standards and techniques, and general acceptance" (Kennedy, 2003). Similar arguments apply to automatic face recognition and are considered throughout. Other factors adversely affecting face recognition include the lack of enough data to learn reliable and distinguishable face representations, and the large computational resources required to adequately process the biometric data. Comprehensive evaluations of the science underlying forensic techniques in general, and studies on the uniqueness of personal face signatures, in particular, are still lacking. The current Face Recognition Grand Challenge (FRGC) project (Phillips et al., 2005), administered by the US National Institute of Standards and Technology (NIST), aims for

98% average reliability at FAR = 0.1%, "a tough standard, but perhaps not tough enough to handle tens of millions of travelers per year", when one considers the false alarms. The scope for FRGC is relatively narrow compared to the earlier FERET and FRVT evaluations (see Sect. 13.4) because despite the relatively large corpus of data involved, the number of subjects enrolled, 275 and 675 for verl.0a and ver2.0, respectively, is only in the hundreds and thus much smaller than FRVT2002. FRGC functionality is further limited to verification compared to previous evaluations that also involved identification. Last but not least, the practicality of FRVT during both enrollment and testing is questionable due to its requirement for a large set of face images using different image capture methods.

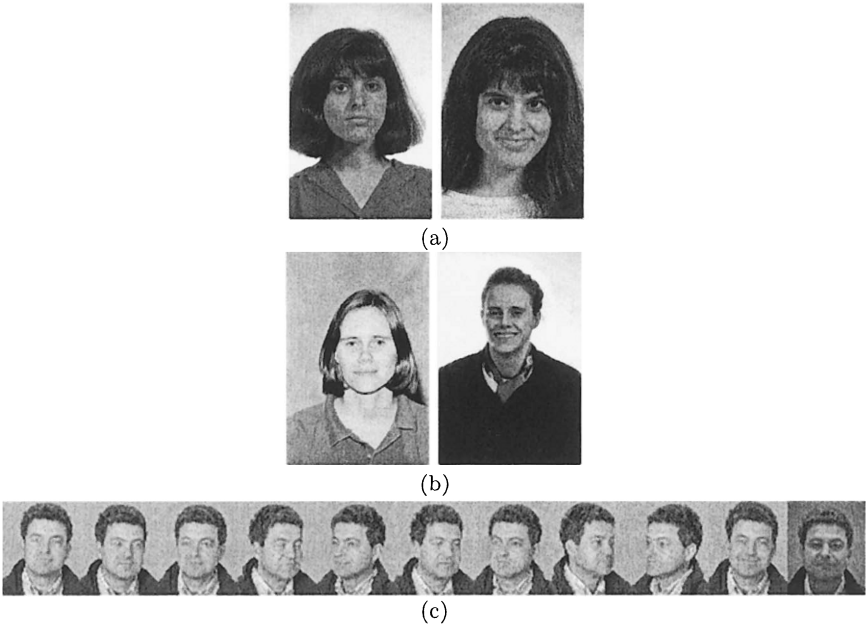

Fig. 1.1. Human Faces (from FERET). (a) Twins; (b) Temporal Variation; (c) Time Sequence Including Pose Variation.

The book is no panegyrics to some research utopia but rather an attempt to be as informative as possible, avoid heuristics, and last but not least cope with meaningful face recognition tasks (see Descartes' admonishments). The book is inclusive but in a comparative fashion in order to motivate and inspire. Hard or intractable problems, e.g., correspondence, segmentation (Gurari and Wechsler, 1982) and clutter, variability, and/or insufficient and/or missing information, are not avoided or glossed over. Folk wisdom chuckles that the difference between theory and practice finds no difference in theory but only in practice. Vapnik (2000) rightly points out, however, that there is nothing more practical than a good theory. The challenge for reliable face recognition is to show through fair and good experimentation that theory and practice are consistent.

The book is mostly about face recognition but it is quite relevant to categorization and recognition for science and technology in general. The driving and unifying force behind the proposed reconciliation among computer vision, machine learning, and pattern recognition, is the active, progressive and selective accrual of evidence needed to reduce uncertainty. Practical intelligence "modifies the stimulus situation as a part of the process of responding to it" (Vygotsky, 1976). Practical intelligence is actually much more than merely modifying or transforming the input. "For the young child, to think means to recall; but for the adolescent, to recall means to think" (Vygotsky, 1976). Connect "adolescent" to reliable face recognition engines, and connect "think" to reasoning and inference. Faces cannot be reliably located and identified from merely one still image. Rather than static inputs, the language of changes observed, their logical interrelations, and the plausible inferences or transformations over space and time underlie reliable face identification and authentication.

1.1 Tasks and Protocols

The generic (on-line) biometric system used herein for face recognition (see Fig. 1.2) will be referred to throughout the book. The **match** task evaluates to what extent the biometric **signatures** extracted from the unknown face exemplar(s) and the biometric signature(s) stored during the enrollment stage as reference **template(s)** are similar. The match score has to be compared against some a priori defined **threshold** value. Matching takes place against a single template (for **verification**), or against a list of candidate templates (for **identification**). Verification is also referred to as **authentication**. Identification is usually carried out using iterative verification and ranking. The face space, usually the basis needed to generate the templates, is derived using face images acquired ahead of time and independent of those that would be later on enrolled for training or queried on (see top of Fig. 1.2). The biometric templates encode the essential features of the face along the specific dimensions of the face space used. They are stored in some central data base but can be also carried by owners on a smart card.

Face recognition performance is still lacking. According to the December 6, 2003 issue of the Economist "governments are investing a lot of faith in biometric technology as a way to improve security. For the moment, this confidence is misplaced. Body-recognition technology is not the silver bullet many governments imagine it to be. Biometrics [are] too flaky to trust." The experience of the 2001 Super Bowl held in Tampa and the trial held at Boston's Logan International Airport in 2002, which exhibited a failure rate of 38.6% [while the false-positive rate exceeded 50%], are cases in point. Again according to the Economist "given the volume of air traffic, the incidence of false alarms will vastly outnumber the rare occasions when someone tries to subvert the system. The false alarms will either have to be ignored, rendering the system useless, or a time-consuming and expensive secondary-screening system will be needed." This book is about how to improve the state-of-the art for reliable face recognition.

Performance evaluation is an integral part of any serious effort to field reliable face recognition systems. **FERET** (Phillips et al., 1998) and **BANCA** (Bailly-Bailliere et al., 2003), the standard evaluation protocols in use today, are briefly described next. FERET starts by considering target (**gallery**) T and query (**probe**) Q sets. The output for FERET is a full (distance) matrix $S(q, t)$, which measures the **similarity** between each query face, $q \in Q$, and each target face, $t \in T$, pair. The nearest neighbor (NN) classifier authenticates then face images using the similarity scores recorded by S. The availability of the matrix S allows for different "virtual" experiments to be conducted when one selects the specific query P and gallery G as subsets of Q and T. Note that one can expand on the above model using data fusion when sets rather than singletons are matched, and both the query and the gallery sets are acquired using multimodal sensors.

The **closed (universe) set face recognition** model used by FERET for 1 : N identification, when each probe has always a mate in the gallery, is restrictive and does not reflect the true intricacies of positive and negative biometric enrollment and identification. Under positive enrollment, the client is authenticated to become eligible for "admission" or apprehended if found on some watch list, while under negative identification the biometric system has to determine that the client does not belong