Dengpan Mou

Machine-based Intelligent Face Recognition

基于机器的智能人脸识别

高等教育出版社
HIGHER EDUCATION PRESS

Dengpan Mou

Machine-based Intelligent Face Recognition

基于机器的智能人脸识别

With 58 figures

高等教育出版社
HIGHER EDUCATION PRESS

Author

Dengpan Mou
Harman/Becker Automotive Systems GmbH
Becker-Goering-Strasse 16
D-76307, Karlsbad
Germany
E-mail: dengpan.mou@harman.com

图书在版编目（CIP）数据

基于机器的智能人脸识别=Machine-based Intelligent Face Recognition：英文/牟登攀著. —北京：高等教育出版社，2010.1
ISBN 978-7-04-022355-2

Ⅰ.基… Ⅱ.牟… Ⅲ.面-图像识别-智能系统-研究-英文 Ⅳ.TP391.41

中国版本图书馆CIP数据核字（2009）第130939号

策划编辑	刘 英	**责任编辑**	刘 英	**封面设计**	张 楠
版式设计	马敬茹	**责任校对**	王效珍	**责任印制**	陈伟光

出版发行	高等教育出版社	**购书热线**	010-58581118
社 址	北京市西城区德外大街4号	**咨询电话**	400-810-0598
邮政编码	100120	**网 址**	http://www.hep.edu.cn
总 机	010-58581000		http://www.hep.com.cn
		网上订购	http://www.landraco.com
经 销	蓝色畅想图书发行有限公司		http://www.landraco.com.cn
印 刷	涿州市星河印刷有限公司	**畅想教育**	http://www.widedu.com
开 本	787×1092 1/16	**版 次**	2010年1月第1版
印 张	11.75	**印 次**	2010年1月第1次印刷
字 数	200 000	**定 价**	32.00元

物料号 22355-00

Sales only inside the mainland of China
仅限中国大陆地区销售

Preface

We can't solve the problems by using the same kind of thinking we used when we created them.
Albert Einstein (1879-1955)

State-of-the-art machine-based face recognition technology, although booming since last decades, is still suffering a lot from critical research challenges, such as the lack of fundamental intelligence, the difficulties of running completely automatically and unsupervisedly without separate training, and the typical failures of dealing with free face pose variations, etc. Those limitations greatly hinder the wide applications it could have had. This book is the first to discuss the general engineering methods of imitating intelligent human brains for video-based face recognition. The advances and evidences from the cognitive science research are introduced in this book, which further strengthen our thoughts and proposals to achieve such a fundamental intelligence in machine vision.

Regarding intelligence, we have defined two directions. The first effort is to simulate the ability of self-learning, self-matching and self-updating. This side of intelligence can be detailed into the following features: the whole recognition procedure is running in an unsupervised, automatic, non-invasive, and self-updated way. It is important to note that, the fully automatic procedure is a generalized face recognition procedure, which includes the task of enrollment (training) and updating as well. However, those steps are typically separate and supervised in machine learning, and therefore missing the essentials of intelligence.

The other main focus of the book is to explore the novel ways on how to implement the high-level analysis in machine-based face recognition, to simu-

late the process in human brains. Through high-level analysis, it is possible to combine multiple available methods, which include conventional machine learning algorithms, image processing approaches, predefined rules, video context, temporal and spatial correlations and even logic deduction. The fusion of multiple approaches contributes significantly to the improved face recognition performance.

Experiments are made through long-term (over years) constructed sequences with more than 30 specific subjects and more than 20 faces from TV news channels. The evaluation results demonstrate the robustness of the proposals in unconstraint video scenarios.

The objective of the author is to provide this book for scientists, researchers and students in the areas of machine-based face recognition. The fundamentals and research backgrounds are provided, aiming to help the beginners to quickly step into the field. Introduction and analysis of the state-of-the-art technology can assist experts to easily keep up to date with the world-wide overview.

The author does hope that the proposals to achieve both intelligence and robustness could be somehow helpful for other researchers, to finally popularize the technology, and to pervasively apply it for designing general machine learning algorithms.

Dengpan Mou
Villingen-Schwenningen, Germany
April 2009

Acknowledgements

Many research results of this book derive from my PhD research in the University of Ulm, Germany, during 2000-2004. Therefore, my first and deepest gratitude is to my PhD advisor, Prof. Dr.-Ing. Albrecht Rothermel, for his invaluable supervision and support. I am deeply grateful to Dr. Rainer Schweer and his group from Deutsche Thomson-Brandt GmbH for stimulating discussions and financially supporting me with such a promising research direction. I would like to express my warmest thanks to Prof. Dr. Heiko Neumann, for his earnest review and comments to shape my PhD thesis.

In writing this book, I have been fortunate to benefit from valuable research discussions with a large number of research groups through my technical reports in many different universities and conferences during the last several years. Although there is no space to mention the hundreds of researchers and scientists for their helpful comments, I would like to take this opportunity to generally thank them all.

My deep acknowledgements go to the following previous colleagues and friends: Dr.-Ing. Ralf Altherr, Dr.-Ing. Markus Buck, Dr.-Ing. Sviatoslav Bulach, Markus Bschorr, Zhen Chang, Richard Geißler, Christian Günter, Roland Hacker, Frank Hagmeyer, Stefan Hirsch, Dr.-Ing. Thomas Kumpf, Martin Lallinger, Dr.-Ing. Roland Lares, Cheng Miao, Dr.-Ing. Oliver Pfänder, Dr.-Ing. Ivan Perisa, Markus Prokein, Xavier Queffelec, Dr.-Ing. Wolfgang Schlecker, Dr.-Ing. Karsten Schmidt, Ralf Schreier, Walter Schweigart, Lei Wang, Lin Wang, Yi Wang, and Chi Zhang for their generous help and contributions to the face databases.

I would like to thank my colleagues at Harman/Becker Automotive Systems GmbH: Michael Maier and Christopher Schmidtmann, who broaden my vision with their extensive industrial research and development experiences.

I am so grateful to my editor, Dr. Liu Ying, from Higher Education Press,

for her great encouragement and generous supports over the years.

In particular, I thank my dearest parents and wife Xiaoying Ge for their endless love and support. Without these emotional and mental supports, the completion of the book could not be possible. This book is dedicated to them.

Contents

List of Figures

List of Tables

1 Introduction

Abstract Face recognition is quite intuitive for most human beings while rather too complex for a machine vision system to be pervasively applied. As the starting point of this book, we are discussing the fundamental pros and cons of state-of-the-art machine vision system versus its human being counterpart. Inspired from the comparison, in the second part of this chapter, a machine-based intelligent face recognition system is briefly introduced. With the high intelligence and high recognition robustness, the system could be widely used for prospective applications such as future smart home, automotive and children mobile phones. In the third part, the discussion of those applications is aiming at broading the visions of the readers. Finally, the outline of the rest part of this book is given.

1.1 Face Recognition—Machine Versus Human

Face recognition has been studied for many years and its attraction grows rapidly due to the wide and promising applications, such as daily identification systems (e.g. automatic banking, access control, computer log-in, etc.), in communication systems (e.g. teleconference and video-phone, etc.), in public security systems (e.g. criminal identification, digital driver license, etc.) and in law enforcement systems.

Face recognition seems instinctive for human beings, but it is really a tough and complex task for a machine-based system. Over the decades, scientists in cognitive and neuroscience are always trying to explore how human

beings recognize faces and why we are in general good at recognizing faces. The recent advances of their research are enlightening [1], but their contributions to the mathematical models and engineering solutions for a machine vision system are still far from enough. Therefore, researchers from the computer science are constructing vast numbers of mathematical models and dedicated algorithms, which may cover the field of artificial intelligence, machine learning, image processing and even video signal processing, etc.

Machine vision systems could have several major advantages over humans. It can have a huge storage medium to deal with much larger amount of people. Furthermore, a machine is not easy to be tired and can work 24 hours a day without any problems. It is always expected to replace the human resources to significantly lower the cost. More importantly, the use of machine can keep the privacy. It is always fair and can 100% follow the predefined rules. For example, suppose that a completely automatic system is required to alarm strangers. A machine will not keep the information of known people and therefore is not violating their privacy. But a human supervisor cannot 100% objectively manage it during the monitoring. And in this case, probably all people under supervision would prefer a machine than a human supervisor.

The most recent face recognition tests [1, 2] claim that, in a particular case, when non-familiar faces (frontal) are tested with significant illumination variations, machine-based face recognition algorithms could be superior to human brains. Nevertheless, in terms of the general recognition precision, the most successful state-of-the-art face recognition technology is still not able to compete with the level of human systems. From the engineering point of view, there is too little attention to the research on face recognition by imitating a human being in a fundamental way, which could combine every means for recognition, not necessarily based on pure psychophysical/neurobiological science or pure mathematical models. For example, it is just a piece of cake for a nine-year old child to recognize people when they turn their heads from frontal to profile views in a video sequence, but can lead to failures for most of current face recognition systems. The reason behind it might be that, in this special case, a child applies his/her multiple approaches including "image processing", video context, logic deduction, experiences, etc. to recognize the people, while those machine-based systems are merely using the mathematical-based image processing methods to calculate the correlation.

Another crucial drawback of most current machine recognition tech-

niques is their lack of intelligence. Many systems are not able to memorize the faces by themselves without the help of a human supervisor or the cooperation of the users. Any new user is supposed to follow some instructions to be enrolled into the databases. The instructions are normally from a human supervisor who is also required to select and update the databases. They are hence invasive to the users and are difficult to be applied in many areas like consumer electronics. Quite recently, there are some advances [3, 4, 5] in the research of building an automatic face recognition system. But their assumptions restrict the real applications. The most critical one is that they normally assume only one person existing in a video sequence, which greatly decreases the complexity of the automatic procedure. With the same previous example, the child has no difficulty in self-learning and memorizing unknown faces, identifying known faces even if several people existing with free behaviors. But that would make a typical error for those systems.

Robustness is the most concerned question for the researchers. Recent face recognition surveys [6, 7] reveal that lighting changes, indoor/outdoor changes, pose variations and elapsed time databases are the critical parameters which greatly influence the performance of a face recognition system. But the effect from those parameters is significantly database dependent. If a database has already enrolled different mugshots under various environments and can update with recent views, state-of-the-art face recognition techniques can produce robust enough results.

Inspired from the above, we define an ideal machine-based face recognition system, which is unique due to the following features:

- Self-learning: completely automatic and unsupervised;
- Non-invasive;
- Robust: in unconstrained environment, against pose and lighting changes, occlusion and aging problems.

In this book, we are exploring the ways and algorithms to approach such an ideal system, mainly but not limited to the application of consumer electronics.

1.2 Proposed Approach

The rule-based algorithms for computer vision and pattern recognition used to

be popular more than 15 years ago. For example, to detect faces, people used eyes, noses, mouths and other features to define the corresponding rules. When there are several facial features detected, and the location of the features agrees with some predefined rules, a face is expected to exist. Although intuitive, these rules are not robust enough to deal with partial occlusions, pose and other frequent day-to-day face changes when applied in images. Learn-based methods therefore appeal more attractions. Nearly all the best face detection and recognition algorithms from the state-of-the-art are learn-based. However, the machine learning method generally suffers from requiring huge numbers of examples for training which makes it difficult to build an automatic system without any human supervision or user's cooperation. In this book, we explore the ways to combine some general rules as major methods together with learn-based methods to achieve such a system for face recognition in video.

For the convenience of describing the proposed approaches, we list the overview of the functional blocks of the automatic system [8] in Fig. 1.1.

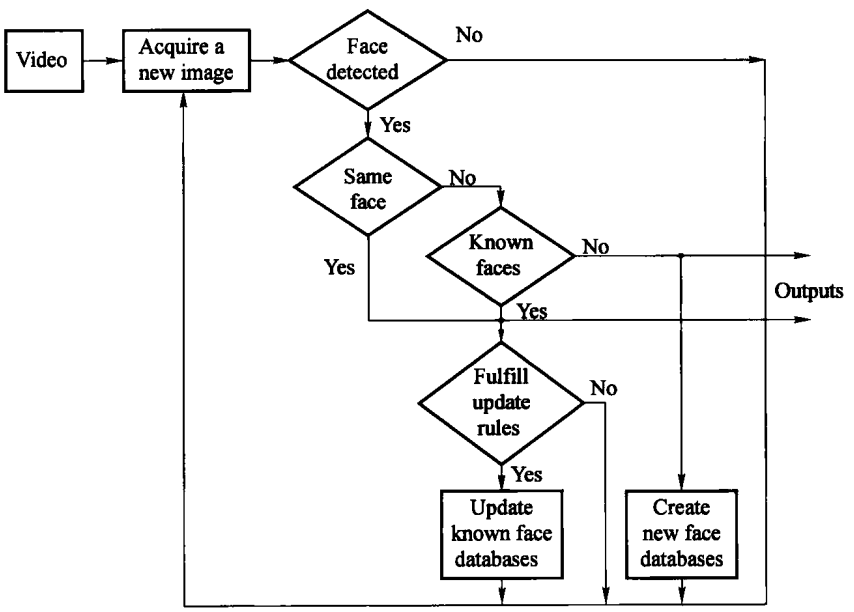

Fig. 1.1 Overview of the System Functional Blocks

Images are continuously acquired from a video source. An image-based face detector and a novel temporal-based face tracker are included to detect whether there are faces in the current image. Any typical image-based face de-