Cognitive Intelligence and Robotics

S. M. Mahbubur Rahman
Tamanna Howlader
Dimitrios Hatzinakos

Orthogonal Image Moments for Human-Centric Visual Pattern Recognition

Springer

Cognitive Intelligence and Robotics

Series Editors

Amit Konar, Department of Electronics and Telecommunication Engineering, Jadavpur University, Kolkata, India
Witold Pedrycz, Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada

Cognitive Intelligence refers to the natural intelligence of humans/animals involving the brain to serve the necessary biological functioning to perform an intelligent activity. Although tracing a hard boundary to distinguish intelligent activities from others remains controversial, most of the common behaviors/activities of living organisms that cannot be fully synthesized by artificial means are regarded as intelligent. Thus the act of natural sensing and perception, understanding of the environment and voluntary control of muscles, blood-flow rate, respiration rate, heartbeat, and sweating rate, which can be performed by lower level mammals, indeed, are intelligent. Besides the above, advanced mammals can perform more sophisticated cognitive tasks, including logical reasoning, learning and recognition and complex planning/coordination, none of which could be realized artificially to the level of a baby, and thus are regarded as cognitively intelligent.

The series aims at covering two important aspects of the brain science. First, it would attempt to uncover the mystery behind the biological basis of cognition with special emphasis on the decoding of stimulated brain signals/images. The coverage in this area includes neural basis of sensory perception, motor control, sensory-motor coordination and also understanding the biological basis of higher-level cognition, such as memory and learning, reasoning and complex planning. The second objective of the series is to publish brain-inspired models of learning, perception, memory and coordination for realization on robots to enable them to mimic the cognitive activities performed by the living creatures. These brain-inspired models of machine intelligence would supplement the behavioral counterparts, studied in traditional AI.

The series includes textbooks, monographs, contributed volumes and even selected conference proceedings.

More information about this series at http://www.springer.com/series/15488

S. M. Mahbubur Rahman ·
Tamanna Howlader · Dimitrios Hatzinakos

Orthogonal Image Moments for Human-Centric Visual Pattern Recognition

Springer

S. M. Mahbubur Rahman
Department of Electrical
and Electronic Engineering
Bangladesh University
of Engineering and Technology
Dhaka, Bangladesh

Tamanna Howlader
Institute of Statistical Research and Training
University of Dhaka
Dhaka, Bangladesh

Dimitrios Hatzinakos
Department of Electrical
and Computer Engineering
University of Toronto
Toronto, ON, Canada

ISSN 2520-1956 ISSN 2520-1964 (electronic)
Cognitive Intelligence and Robotics
ISBN 978-981-32-9944-3 ISBN 978-981-32-9945-0 (eBook)
https://doi.org/10.1007/978-981-32-9945-0

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

The introduction of orthogonal image moments more than thirty years ago was a milestone in pattern recognition. Since that time, mathematical properties of orthogonal moments have been thoroughly studied, new moments have been defined and innovative applications have been presented through the increasing number of research contributions in this field. In recent times, human-centric visual pattern recognition has become a trending topic due to its role in artificial intelligence, particularly in the applications of biometric recognition, affective computing, and human–computer interaction. Recent texts have focussed on the mathematical properties of moments and their invariants in pattern recognition and provide a cursory overview of related applications. In contrast, this book places emphasis on the use of orthogonal moments in solving specific problems arising in human-centric visual pattern recognition. It represents a compendium of research works that demonstrate the effectiveness of orthogonal moment-based features in face recognition, facial expression recognition, fingerprint classification, and iris recognition. It presents methods that address an unresolved issue in moment-based feature selection: how to decide the best candidate of higher-order moments to construct the feature vector. Furthermore, this book demonstrates the success of image moments in applications where other feature types have been prevalent, such as in the common problems of biometric recognition and affective computing.

In addition to offering new concepts that illustrate the use of statistical theories in moment-based methods, this book presents results implemented on recent databases involving challenging scenarios and provides comparisons with recent state-of-the-art methods. Conclusive remarks on the use of image moments in the practical problems of pattern recognition and future research directions are also given. The book will be of interest to researchers and graduate students working in the broad areas of computer vision and pattern recognition.

Dhaka, Bangladesh — S. M. Mahbubur Rahman
Dhaka, Bangladesh — Tamanna Howlader
Toronto, Canada — Dimitrios Hatzinakos
September 2018

Contents

About the Authors

S. M. Mahbubur Rahman received his Ph.D. in Electrical and Computer Engineering (ECE) from Concordia University, Montreal, QC, Canada, in 2009. Currently, he serves as a Professor at the Department of Electrical and Electronic Engineering (EEE), Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh. In 2017, he was a Visiting Research Professor of the EEE at the University of Liberal Arts Bangladesh. He was tenured at the University of Toronto as an NSERC (Natural Sciences and Engineering Research Council) Postdoctoral Fellow in 2012. He has a strong record of research in his area, which includes contributing to more than 50 publications in SCI-indexed journals and the peer-reviewed proceedings of international conferences. He has served as an Associate Editor for an SCI-indexed journal: Circuits, Systems, and Signal Processing, published by Springer Nature. His research interests are in the areas of biometric security systems, intelligent transportation systems, cognitive science, stereo vision, virtual reality, biomedical visualization, human–computer interaction, video surveillance, signal processing and communication systems.

Tamanna Howlader received her Ph.D. in Mathematics from Concordia University, Canada. Currently, she is a Professor of Applied Statistics at the Institute of Statistical Research and Training (ISRT), University of Dhaka, Bangladesh. She is a statistician who enjoys interdisciplinary research. The articles and book chapters that she has published demonstrate novel statistical applications in the areas of image processing, computer vision, pattern recognition and public health. Tamanna has received several prestigious awards including the Sydney R. Parker Best Paper Award from the Journal of Circuits, Systems and Signal Processing published by Springer Nature. She is a member of the International Statistical Institute.

Dimitrios Hatzinakos received his Ph.D. in Electrical Engineering from Northeastern University, Boston, MA, in 1990, and currently serves as a Professor at the Department of Electrical and Computer Engineering, University of Toronto (UofT), Toronto, Canada. He is the co-founder and since 2009 the Director and the

Chair of the management committee of the Identity, Privacy and Security Institute (IPSI) at the UofT. His research interests and expertise are in the areas of multimedia signal processing, multimedia security, multimedia communications and biometric systems. He is the author/co-author of more than 300 papers in technical journals and conference proceedings; he has contributed to 18 books, and he holds seven patents in his areas of interest. He is a Fellow of the IEEE, a Fellow of the Engineering Institute of Canada, and a member of the Professional Engineers of Ontario, and the Technical Chamber of Greece.

Chapter 1
Introduction

1.1 Introduction

We live in a world that is built upon patterns. What is a pattern? The Oxford dictionary defines a pattern as a repeated decorative design. In the language of pattern recognition, however, a pattern has been described as an entity that could be given a name [36]. Thus, the bird, boat, buildings, and people that we see in Fig. 1.1 are all examples of patterns. Recognizing patterns in the environment is one of the fundamental signs of intelligent behavior. A 3-year old child, for example, can discern the alphabets almost effortlessly, an eagle can spot its prey from a 1000 feet above the ground, and a carnivorous fish can capture a smaller fish camouflaged against the sand. Each of these examples illustrate the capacity to perceive order from disorder, which is not just a matter of visual skill but also a sign of intellectual skill that is essential for survival. Whether or not artificial systems could be infused with such intelligence is a question that has occupied the minds of scientists from as early as the 1950s. The quest for an answer stimulated intense research in the field of pattern recognition and its allied disciplines, namely, artificial intelligence, computer vision, and machine learning. These domains have a strong overlap with statistics, probability, computational geometry, and image processing. Over the past few decades, the development of exciting new methods coupled with the availability of high computational resources has led to some progress in the development of artificially intelligent machines capable of perceiving humans and emulating their actions.

1.2 Pattern Recognition: Mimicking the Human Visual System

Among the senses that living organisms possess, the most vital is vision. Vision provides several cues about an object including its motion, color, size or position that are crucial for survival. Machines with visual capabilities can be taught to perceive

S. M. M. Rahman et al., *Orthogonal Image Moments for Human-Centric Visual Pattern Recognition*, Cognitive Intelligence and Robotics,
https://doi.org/10.1007/978-981-32-9945-0_1

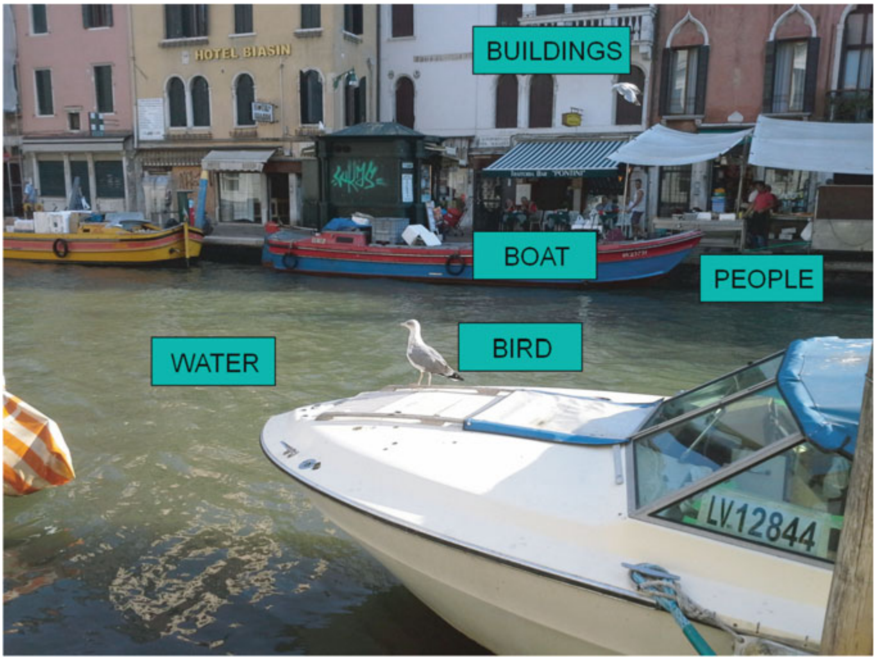

Fig. 1.1 Examples of patterns in an image

the environment and this makes them more useful in real-life applications. This book is concerned with visual pattern recognition, which may be described as the process of analyzing, describing, identifying and extracting the information contained in an image, video or other visual data. Visual patterns could be, for example, a fingerprint image, a handwritten cursive word, or a human face. The most efficient visual pattern recognition system is the human brain yet how it functions is still not fully understood. What follows is an oversimplified description of the mechanism by which the human visual system works. The human visual system mainly consists of two parts: the eyes and the visual cortex. The eyes act as image receptors, which capture light and convert it to signals that pass along optic nerve fibers to the visual cortex of the brain via the lateral geniculate nucleus (LGN). The visual cortex, also known as the image processing center of the brain, processes the signals received from the eyes and builds an internal picture of the scene being viewed. The cortex has columns of cells extending through six layers and arranged in particular directions that appear to be devoted to recognizing lines or bars of varying lengths and orientations. In this way, the cortex handles the processing of complex shapes and arrangements. It is believed that the brain processes visual information in a series of steps, each taking the output of the previous steps and building up progressively more complex impressions of the input [22]. A simplified version of the human visual system is shown in Fig. 1.2.

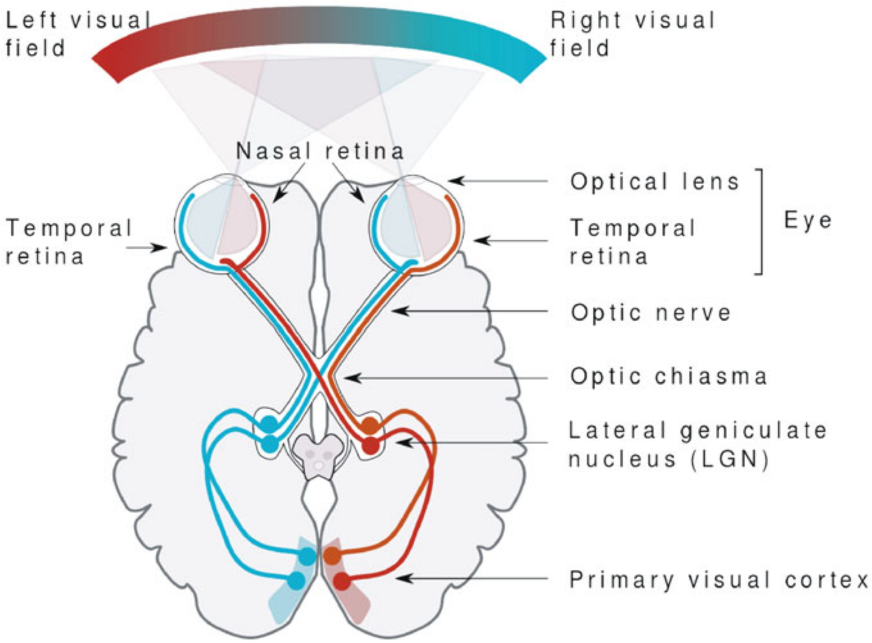

Fig. 1.2 General structure of human visual system. Reproduced from [1]

Since early times, engineers have worked with cognitive scientists to develop artificial visual pattern recognition systems. There has been measured success thanks to the availability of low-cost cameras, high-speed networks, huge-size storage capabilities, and increased computational power. In fact, many important image analysis tools such as edge detectors and neural networks have been inspired by human visual perception models. However, much work still needs to be done before artificial systems are able to match the capabilities of the human visual system, particularly in terms of accuracy. Nonetheless, visual pattern recognition systems are already being used for several purposes such as measuring size, localizing an object, estimating the pose, recognizing the activity or identity of a subject, target recognition in military applications, autonomous navigation, robotic assembly, industrial inspections, security and surveillance, biomedicine, agriculture, video gaming, virtual reality, and human–computer interaction (HCI) to name a few. In all these applications, the common challenge is the problem of recognizing complex patterns having high dimensional structure with arbitrary orientation, location, and scale.

Visual pattern recognition may be supervised or unsupervised. Figure 1.3 illustrates the difference between the two approaches. Supervised classification is characterized by a training data containing labeled samples and a test data set containing samples with unknown labels. The training data is used to learn the discriminatory features as well as to estimate the parameters of the classifier. The objective of supervised methods is to classify the samples in a test data set based on the learned

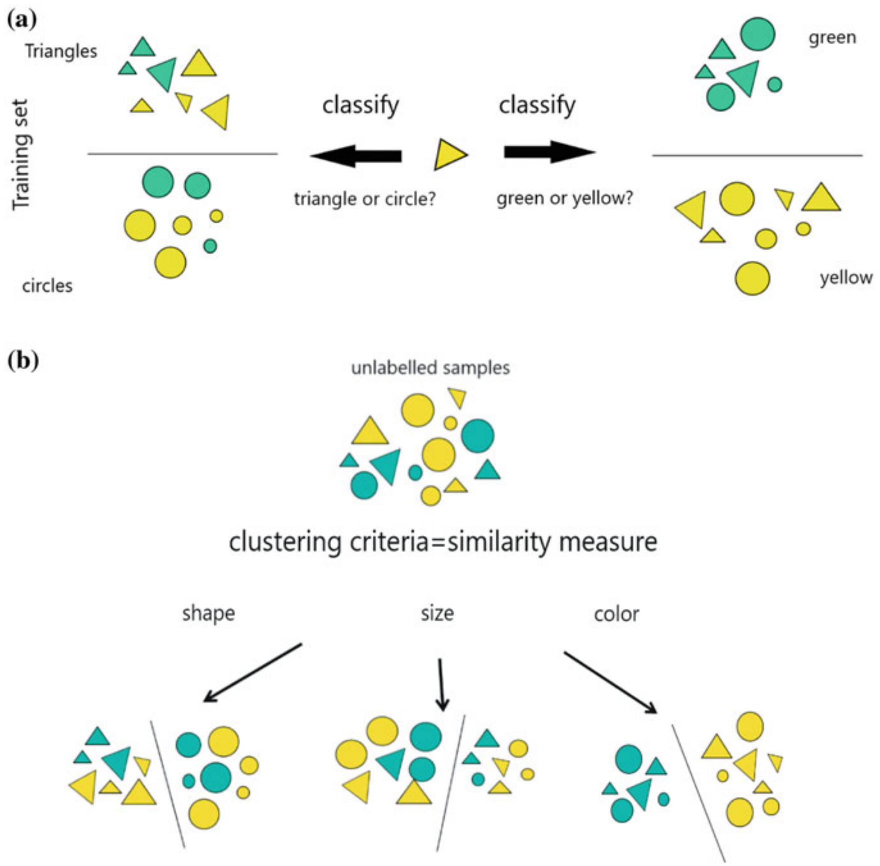

Fig. 1.3 An illustration comparing two types of classification. **a** Supervised classification uses labeled training samples to classify an object according to shape or color. **b** Unsupervised classification uses unlabeled samples and a similarity metric to group objects according to their similarity in shape, size, or color

features. In contrast, the unsupervised classification, also known as clustering, does not involve any training data with labeled samples. Rather, the objects in the entire data are grouped into homogeneous clusters based on some similarity criterion. In both approaches, the correct choice of features is important.

For most applications, the classification of visual patterns is supervised. Figure 1.4 shows the major components of a supervised visual pattern recognition system. These are data acquisition, preprocessing, feature extraction and representation, pattern classifier design, and performance evaluation. During data acquisition, different types of sensors and sensing modalities produce different types of images depending on the problem domain. For example, images may be 2D or 3D or they may be grayscale or color images, multispectral images, computed tomography (CT) scans, images pro-

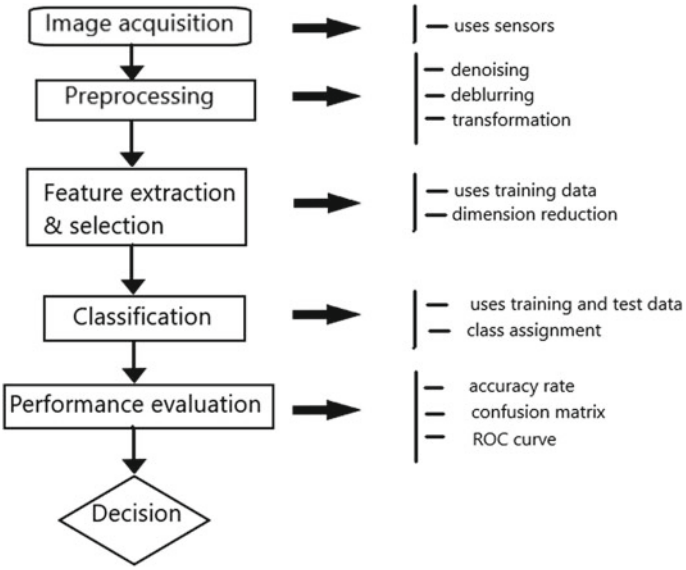

Fig. 1.4 Components of a supervised visual pattern recognition system

duced by magnetic resonance imaging (MRI), X-ray images, or microarray images. Once the image is produced, the next step is to remove irrelevant background or noise and perform image enhancement, if required. This step is known as preprocessing. At the heart of the pattern recognition system lies feature extraction and classification. The image will generally contain redundant information. The objective of feature extraction is to perform dimension reduction and discern properties or attributes of the patterns known as "features" that are discriminatory in nature and therefore useful for pattern classification. To put this in a more mathematical context, feature extraction aims to represent an object as a point in a finite-dimensional space known as the feature space such that different views of the object correspond to the same point in feature space [59]. Figure 1.5 illustrates the concept. A good feature algorithm achieves this without losing too much information about the object and may yield features that are less sensitive to noise and occlusion. Designing efficient features for feature extraction is the cornerstone of this book and will be discussed in sufficient detail from a certain perspective in the subsequent chapters.

The next step, classification, is essentially a decision-making process that involves finding the pattern class or pattern description of new unseen objects based on a set of training examples. Although it is generally agreed that there are four main approaches to pattern recognition, namely, template matching, statistical classification, syntactic or structural matching and neural networks, statistical classification and neural networks are the most widely used approaches [36]. The final step involves assessing the performance of the pattern recognition system, which often involves the use of