

Preface

Overview and Goals

Human behavior is complex, but not random. Computer analysis of human behavior in its multiple scales and settings leads to a steady influx of new applications in diverse domains like human-computer interaction, affective computing, social signal processing, and ambient intelligence. We envision seamlessly integrated plug and play devices that can be used to endow a given environment with an awareness of the physical, functional, temporal and social organization of its internal domestic dynamics, as well as the personalities and social relationships of its inhabitants, providing a vast array of new services for ordinary people. We picture intuitive tools for social scientists, psychologists and doctors to observe and quantify human behavior. We desire realistic virtual agents and engaging robots that can analyze and properly respond to social context. We seek intelligent algorithms to process vast collections of multimedia information to retrieve relevant material in response to semantic queries.

The realization of all these tools and systems requires a fundamental grasp of the key issues, as well as knowledge and experience over the computational tools. The goal of this book is to provide a solid foundation toward achieving this.

Most significantly, the focus of the book is in advanced pattern recognition techniques to automatically interpret complex behavioral patterns generated when humans interact with machines or with others. This is a challenging problem where many issues are still open, including the joint modeling of behavioral cues taking place at different time scales, the inherent uncertainty of machine detectable evidences of human behavior, the mutual influence of people involved in interactions, the presence of long term dependencies in observations extracted from human behavior, and the important role of dynamics in human behavior understanding.

The contiguity of these problems with the field of pattern recognition is straightforward. The editors of the present volume (together with Alessandro Vinciarelli from the Univ. of Glasgow and Nicu Sebe from the Univ. of Trento) organized the First Workshop of Human Behavior Understanding as a satellite workshop to Int. Conference on Pattern Recognition (ICPR) in 2010, with wide attendance. Similarly,

the Human Communicative Behaviour Analysis Workshop is held as a satellite to IEEE's Computer Vision and Pattern Recognition (CVPR) Conference since 2008. The topics presented in the book are actively researched in the pattern recognition community.

Target Audience

The book is planned as a graduate textbook/guidebook on computer analysis of human behavior. Starting from the preliminaries, it covers major aspects concisely, introduces some of the most frequently used techniques and algorithms in detail, and provides examples of real applications. Each chapter is a stand-alone treatment of a relevant subject, discussing key issues, classic and recent approaches, as well as open questions and future research directions. Since the subject matter is very broad, we have restricted the number of chapters to ensure that the book can be covered in one semester and focused on providing both a good background and a comprehensive vision of the field. Each chapter is supplemented with educational material, including chapter summary, glossary, questions, and online lecture slides to help the instructor.

Organization of the Book

We have divided the book into four parts. The first part, called "*The Tools of the Trade*", is a selection of basic topics that the reader will repeatedly come across in human behavior analysis. These chapters are all written in an intuitive and didactic way, and pave the way for more advanced discussions.

The first of these four chapters introduces Bayesian methods for behavior analysis, because there are numerous uncertainties in measuring and assessing behavior, and one also needs to deal with idiosyncrasies and inconsistencies. In particular, Gaussian processes and Dirichlet processes are covered in this chapter.

Almost all human behavior is temporal, but the time scale may range from milliseconds (e.g. the movement of a facial muscle) to hours (e.g. sleep cycles) or months (e.g. habits). The temporal dimension is the sine qua non of human behavior, and subsequently, the second chapter introduces basic methods for temporal analysis, including Hidden Markov Models, Conditional Random Fields, and variants thereof. This chapter also gives a concise introduction to graphical models, their factorization and how to perform inference in graphical models. Subsequently, it complements the first chapter nicely.

The third chapter discusses how we can detect and track humans by computer vision methods to understand their actions, and is a prerequisite for most material presented in Parts II and III. The visual modality can provide a system with high-dimensional and dynamic data, creating some of the most formidable challenges in behavior analysis, and most analysis pipelines start with detection and tracking.

The fourth chapter is an introduction to computational visual attention, and explains how humans make sense of the immensely rich perceptual input (*the firehose of experience*, in the memorable expression of Ben Kuijpers), as well as how computer systems can mimic the process of attention to reduce their computation load, especially in terms of bottom-up (data driven) and top-down (semantically driven) approaches. The VOCUS attention system is described in detail, and applications are given from the field of mobile robotics.

Two major application areas of human behavior analysis are activity recognition, and analysis of social signals, including those that pertain to affect. The second part of the book is devoted to “*Analysis of Activities*”, whereas the third part deals with social and affective behavior.

The first chapter in the second part (i.e. Chap. 5) is on gait and posture analysis, which is a relevant problem for clinical applications, surveillance, and ergonomics. In this chapter the reader will be introduced to a host of sensors (like gyroscopes and accelerometers) that can be used to measure useful physiological and movement related signals. Most of these sensors are easily integrated into smart phones, creating a huge potential for mobile phone applications (and games) that are based on human behavior analysis.

The second part continues with Chap. 6 on hand gesture analysis. Hand gestures can be used to define natural interfaces in human-computer interaction, but they are also rich sources of social and contextual cues during conversations. This chapter builds heavily on temporal analysis and tracking material of the first part.

Automatic analysis of complex human behavior is, as we mentioned earlier, one of the grand challenges of multimedia retrieval. Some behaviors are simple (e.g. walking) and can be detected by looking at simple cues. Some behaviors are complex (e.g. flirting) and require extensive knowledge and processing of context. Chapter 7 is on semantics of human behavior in image sequences, and discusses how environment influences the perceived activities, and how bottom-up and top-down approaches can be integrated for recognizing events.

With the third part of the book, “*Social and Affective Behaviors*”, we move toward applications where pattern recognition and machine learning methods need to be complemented with psychological background knowledge and models. Social behaviors constitute a major research area, largely overlapping with the field of affective computing. This part opens up with a psychological treatise on social signals, written in a very accessible way for an audience with mainly computer science background. A wide range of signals like dominance, persuasion, shame, pride, and enthusiasm are introduced and discussed in Chap. 8.

Poggi and D’Errico define a *social signal* as a communicative or informative signal that, either directly or indirectly, conveys information about social actions, social interactions, social emotions, social attitudes and social relationships (2010). The audio modality is more frequently used for processing such signals compared to the visual modality. Chapter 9 is an extensive and technical discussion of voice and speech analysis for assessing human communicative behavior.

Chapter 10 is on analysis of affect from combined audiovisual cues. It discusses how affective signals can be measured and evaluated in a continuous manner, as

well as how to perform multimodal fusion. While the discussions of data annotation and experimental design in Chap. 10 pertain mainly to affective displays, identified issues and challenges are valid for almost the entire range of human behavior analysis.

The last chapter of the third part of the book, Chap. 11, discusses social interactions and group dynamics. Four case studies from a meeting scenario are presented, where the authors combine audiovisual cues to estimate the most and least dominant person, emerging leaders, and functional roles in the group, as well as the group dynamics as a whole.

The fourth part of the book is devoted to “*Selected Applications*” from three different research fields (ambient intelligence, biometrics, and gaming, respectively) that reflect the diversity and scope of behavioral cues and their usage.

Chapter 12 describes a vision of ambient assisted living, where a smart environment monitors the activities of its inhabitants for health care purposes. The aging population of developed countries call for technologies to allow elderly to remain longer in their home environments, giving them a higher quality of life, as well as reducing the costs of health care. Smart monitoring tools, provided that they deal with acceptance and privacy issues properly, are of great value.

Behavioral biometrics is the identification of a person via behavioral cues. Chapter 13 surveys this new field, and reveals that an astonishing number of behavioral cues, measured directly or indirectly, can be used to verify the identity of a person. This is a joint achievement of improvements in pattern recognition methods, as well as sensor technologies.

Finally, Chap. 14 deals with games, which are major economic drivers behind computer scientific research. Games do not only serve entertainment; there are games to exercise the body and mind, and sometimes a game is the means to a completely different end. Take for instance the robotics community, which uses robot soccer as a driver behind great advances in robot mechanics, coordination, planning, and a host of other challenges. This is an application area where both real and virtual human behavior can be analyzed for improving engagement and interaction, as well as for teaching computers and robots skills at human level.

Taken together, these chapters cover most of the field of human behavior analysis. It is possible to include many more tools and applications, as the field is positioned at a confluence of many different and mature research areas. Nonetheless, we hope that this collection will be a useful teaching tool for initiating newcomers, as well as a timely reference work that sums up recent research in this advancing area.

Amsterdam, The Netherlands

Albert Ali Salah
Theo Gevers



<http://www.springer.com/978-0-85729-993-2>

Computer Analysis of Human Behavior

Salah, A.A.; Gevers, T. (Eds.)

2011, XVI, 412 p., Hardcover

ISBN: 978-0-85729-993-2