1 Introduction: Social Signal Processing

Alessandro Vinciarelli

Introduction

Social signal processing (SSP) is the computing domain aimed at modeling, analysis, and synthesis of social signals in human–human and human–machine interactions (Pentland, 2007; Vinciarelli et al., 2008, 2012; Vinciarelli, Pantic, & Bourlard, 2009). According to different theoretic orientations, social signals can be defined in different ways, for example, "*acts or structures that influence the behavior or internal state of other individuals*" (Mehu & Scherer, 2012; italics in original), "*communicative or informative signals which . . . provide information about social facts*" (Poggi & D'Errico, 2012; italics in original), or "*actions whose function is to bring about some reaction or to engage in some process*" (Brunet & Cowie, 2012; italics in original). The definitions might appear different, but there seems to be consensus on at least three points.

- Social signals are *observable* behaviors that people display during social interactions.
- The social signals of an individual *A* produce changes in others (e.g., the others develop an impression or a belief about *A*, react to *A* with appropriate social signals, or coordinate their social signals with those of *A*).
- The changes produced by the social signals of *A* in others are not random, but follow *principles and laws*.

In a computing perspective, the observations above lead to the key idea that shapes the field of Social Signal Processing, namely that social signals are the *physical*, machine detectable trace of social and psychological phenomena not otherwise accessible to direct observation. In fact, SSP addresses the following three main problems.

- *Modeling*: identification of principles and laws that govern the use of social signals.
- *Analysis*: automatic detection and interpretation of social signals in terms of the principles and laws above.
- *Synthesis*: automatic generation of artificial social signals following the principles and laws above.

Correspondingly, this book is organized into four main sections of which the first three focus on the three problems outlined above while the fourth one introduces current applications of SSP technologies.

- *Part I Conceptual models of social signals*: this section covers definitions and models of social behaviour and social signals – the core concepts of SSP as researched in social psychology, cognitive sciences, evolutionary psychology, and anthropology.
- *Part II Machine analysis of social signals*: this section covers the technologies aimed at automatic detection of social signals apparent from face and facial behaviour, vocal expressions, gestures and body postures, proxemics, etc.
- *Part III Machine synthesis of social signals*: this section covers the technologies aimed at empowering artificial agents with the ability of displaying social signals, including expressive speech synthesis, facial animation, and dialogue management.
- *Part IV Applications of SSP*: this section covers the most important SSP applications domains, including socially intelligent surveillance, deception detection, healthcare, and multimedia indexing.

Every chapter is a survey aimed at beginners and experienced researchers in the field. For the former, the surveys will be a fundamental source of references and a starting point in the research on the topic. For the latter, the chapters will be a compendium of the large body of knowledge accumulated in SSP, informed by the critical views of some of the most influential researchers in the domain.

Part I Conceptual Models of Social Signals

Part I introduces social science perspectives on social signaling. Covered are theories and models related to the etiologies, form, and functions of social signals. The first chapter, "Biological and Social Signaling Systems" (Kory Floyd and Valerie Manusov), addresses the fundamental issue of nurture versus nature influences on social signals, focusing in particular on the interplay between innate biological processes and acquired components resulting from sociocultural processes. The next two chapters concern the horizontal versus vertical dimensions along which social messages are expressed and interpreted. The chapter, "Universal Dimensions of Social Signals: Warmth and Competence" (Susan Fiske and Cydney Dupree), surveys recent results on the perception of *warmth* and *competence*, the two dimensions along which people tend to assess unacquainted others in the earliest stages of an interaction. In particular, the chapter highlights that the two dimensions are *universal*, that is, they tend to appear in all situations and cultures. Judith Hall and Marianne Schmid Mast survey the use of social signals as a means to express social verticality – status and power differences between people belonging to the same social system – in the chapter entitled "The Vertical Dimension of Social Signaling."

The two chapters that follow concern the relationship between emotions and social signals. The fourth chapter, "Measuring Responses to Nonverbal Social Signals: Research on Affect Receiving Ability" (Ross Buck, Mike Miller and Stacie Renfro Powers), addresses the perception of emotions and affect that others display. In particular, the chapter focuses on pickup and processing of facial and bodily displays. It is complemented by the chapter authored by Klaus Scherer, Björn Schüller and Aaron Elkins,

"Computational Analysis of Vocal Expression of Affect Trends & Challenges," which focuses on the vocal expression of emotions. Furthermore, the chapter addresses the role that signal processing technologies can have in the investigation of social signals.

The role of social signals as a means to display identity and personality is the focus of "Self-presentation: Signaling Personal and Social Characteristics" (Mark R. Leary and Katrina P. Jongman-Sereno). In particular, this chapter analyses the considerable efforts that people make in order to lead others to treat them in desired ways. Finally, the last three chapters of Part I address phenomena that take place during the interaction between people. The chapter, "Interaction Coordination and Adaptation," by Judee Burgoon, Norah Dunbar, and Howard Giles focuses on the tendency of interacting people to mutually adapt their interaction styles or to adopt similar behavior patterns. Persuasion is at the core of the chapter authored by William Crano and Jason Siegel, "Social Signals and Persuasion," with particular attention to the effect of social signals on the credibility of a source. Finally, the last chapter of Part I, "Social Presence in CMC and VR" by Christine Rosakranse, Clifford Nass, and Soo Youn Oh, focuses on technology mediated interaction contexts and, in particular, on how to convey social presence when interaction is not face-to-face.

These Part I chapters supply essential context for conducting machine analysis of social signals. They identify the multitude of functions that given signals may perform and draw attention to the fact that many signals arise not from meanings that senders are attempting to convey but rather are a response to the displays of interlocutors and the jointly created exchange.

Part II Machine Analysis of Social Signals

The second part of the book deals with machine analysis of social signals. It represents a collection of surveys covering the state of the art in research and technology aimed at automatic detection of social signals.

The first two chapters deal with two of the most important sources of social signals, namely face and body. In "Facial Actions as Social Signals," Michel Valstar, Stefanos Zafeiriou, and Maja Pantic survey the past work in machine analysis of facial gestures (i.e., facial action units), which are the building blocks of all facial expressions, including the facial expressions typical of displays of social signals such as interest, mimicry, empathy, envy, and so on. Particular attention is paid to discussing automatic facial gesture recognition in unconstrained conditions and real-life situations. Ronald Poppe, the author of "Automatic Analysis of Bodily Social Signals," surveys the state of the art approaches and technologies for automatic recognition of social signals apparent from a human body's posture and movement. This includes interest detection in interactions with robot companions, detection of phenomena such as mimicry and turn taking, and deception detection.

The chapters following those mentioned above address the problem of using social signals as a means to infer people's characteristics. Personality traits profoundly influence one's displays of social signals and one's social interactions. For instance, it is

commonly known that extrovert people easily establish and have more pleasant social interactions than is the case with more introvert people. In "Computational Approaches for Personality Prediction," Bruno Lepri and Fabio Pianesi discuss two approaches to automatic prediction of one's personality. The first relies on automatic recognition of so-called distal cues (e.g., voice pitch) and learning which distal cues underlie which personality trait (extrovert, neurotic, agreeable, conscientious, open). The second approach to automatic personality prediction relies on one's profile and interactions in a social network such as Facebook. Attractiveness and likability affect social exchanges in very predictable ways. It is widely known, for example, that attractive people establish social interaction more easily than less attractive people. In "Automatic Analysis of Aesthetics: Human Beauty, Attractiveness, and Likability," Hatice Gunes and Björn Schüller survey the past work on automatic analysis of human attractiveness and likability based on audio and visual cues shown by the judged person.

The remaining chapters of Part II focus on phenomena that take place during social interactions. A large body of research in psychology points out that an individual's temporal coordination in social interactions has detrimental effects on the outcome of the interaction (e.g., whether one will feel liked or not, whether the outcome of negotiation will be positive or not, etc.). In "Interpersonal Synchrony: From Social Perception to Social Interaction," Mohamed Chetouani, Emilie Delaherche, Guillaume Dumas, and David Cohen focus on computational models of interpersonal synchrony and survey the automatic approaches to interpersonal synchrony assessment. Social emotions are defined as emotions that relate to interpersonal interactions, rather than to individual feelings (e.g., empathy, envy, shame, etc.). In "Automatic Analysis of Social Emotions," Hatice Gunes and Björn Schüller provide an overview of the past research on automatic recognition of social emotions from visual and audio cues. In "Social Signal Processing for Automatic Role Recognition," Alessandro Vinciarelli surveys the past work on this earliest research topic addressed by the SSP community – recognition of social roles (i.e., the position that someone holds in a given social context, such as "moderator" versus "discussion participant"). Particular attention is paid to open issues and challenges in this research field.

All previously mentioned approaches to automatic analysis of social signals build upon machine learning techniques to model latent and complex behavioral patterns, underpinning target social signals, from available data (i.e., audio, visual, multimodal observations of target social signals). In "Machine Learning Methods for Social Signal Processing," Ognjen Rudovic, Mihalis Nicolaou, and Vladimir Pavlovic focus on systematization, analysis, and discussion of recent trends in machine learning methods employed typically in SSP research.

Part III   Machine Synthesis of Social Signals

Part III includes surveys on some of the most important aspects of social signals synthesis, from the generation of artificial nonverbal cues, to the use of artificial cues to convey socially relevant information, to social robots.

The first two chapters address, respectively, speech synthesis and the generation of gestures and bodily movements. Kallirroi Georgila – author of "Speech Synthesis: State-of-the-art and Challenges for the Future" – describes state-of-the-art techniques for the generation of artificial speech and emphasizes in particular the synthesis of emotional and expressive speech through the use of paralanguage and nonverbal cues. Similarly, the authors of "Body Movements Generation for Virtual Characters and Social Robots" (Aryel Beck, Zerrin Yumak, and Nadia Magnenat-Thalmann) survey not only the technologies to synthesize nonverbal cues such as body posture, gestures, and gaze, but also the use of these cues when it comes to the communication of emotion and affect.

In the two chapters that follow those mentioned above, the authors address the problem of how to artificially generate social phenomena and, in particular, how to convey emotion and prosocial behavior. Marc Cavazza (author of "Approach and Dominance as Social Signals for Affective Interfaces") surveys the adoption of affective interfaces as a principled approach toward the improvement of the interactions between users and machines. Ketaki Shriram, Soon Youn Oh, and Jeremy Bailenson (authors of "Virtual Reality and Prosocial Behavior") survey the efforts aimed at promoting positive changes in behavior (e.g., increasing environment awareness or adopting healthier lifestyles) through the adoption of virtual spaces where it is possible to interact in a controlled setting, possibly including artificial characters.

The conclusive chapter of Part III, "Social Signal Processing in Robotics" by Maha Salem and Kerstin Dautenhahn, focuses on social robots, one of the most important forms of embodiment where the synthesis of social signals can play a crucial role in ensuring smooth, enjoyable, and effective interactions between humans and machines.

Part IV Applications of Social Signal Processing

The last part of the book deals with the applications of social signal processing. While being a relatively young domain (the very expression *social signal processing* was coined less than ten years ago), the methodologies produced in the field have been shown to be promising in a wide spectrum of application areas.

The first two chapters of this part show applications where the very analysis of social signals can serve practical purposes, namely surveillance and automatic understanding of group behavior. Dong Seon Cheng and Marco Cristani ("Social Signal Processing for Surveillance") show how the automatic analysis of social signals can improve current surveillance approaches that, typically, analyze human behavior without taking into account the peculiarities of social behavior. Daniel Gatica-Perez, Oya Aran, and Dinesh Jayagopi ("Analysis of Small Groups") survey efforts aimed at inferring the social phenomena taking place in small groups, such as social verticality, personality, group cohesion, and characterization. These efforts are beneficial in particular for applications aimed at making meetings effective and productive.

Another two chapters show the use of social signal processing methodologies as a support for multimedia indexing methodologies. The chapter "Multimedia Implicit Tagging" (Mohammad Soleymani and Maja Pantic) shows that capturing the reaction of a

user (e.g., laughter or sobbing) in front of a multimedia item (e.g., a video) provides information about the content of the item itself that can then be tagged with categories such as *funny* or *sad*. In a similar vein, Alessandro Vinciarelli ("Social Signal Processing for Conflict Analysis") shows that the detection of conflict can help to extract the most important moments in large repositories of political debates.

The last two chapters of this part target the adoption of social signal processing methodologies in two major application areas, that is, healthcare and deception detection. Mohamed Chetouani, Sofiane Boucenna, Laurence Chaby, Monique Plaza, and David Cohen ("Social Signal Processing and Socially Assistive Robotics in Developmental Disorders") show in particular that the analysis of social signals can help the detection of developmental problems in children that, in many cases, cannot even speak. Judee K. Burgoon, Dimitris Metaxas, Thirimachos Bourlai and Aaron Elkins ("Social Signals of Deception and Dishonesty") survey the progress on the possibility of developing technologies capable to identify people who lie.

Conclusions

This chapter provides a description of the book's organization and content. The goal is to allow the readers to identify chapters of interest quickly and easily and, at the same time, to develop awareness of the main problems and areas covered in social signal processing. The many authors involved in the book have made major efforts to combine rigour and depth with clarity and ease of access. This will hopefully make this book a valuable instrument for a wide spectrum of readers.

- *SSP beginners*: researchers starting their investigations in SSP will benefit from surveys because these provide an overview of the state-of-the-art perspectives, identify the most important challenges in the field, include rich bibliographies, and provide the right terminology.
- *SSP experts*: researchers knowledgeable in SSP can benefit from the surveys because these condensate, in a compact and concise form, a large body of knowledge typically scattered across multiple disciplines. Critical views of the authors could provide a fertile ground for discussion and, in turn, be an effective tool in pushing the limits of innovation in the field.
- *SSP teachers*: teachers will benefit from the material because it provides an introduction to the field and can be used as didactic material for students with different backgrounds and/or at different stages of their education. Furthermore, the material is organized in parts that correspond to the most natural structure of an SSP course.
- *SSP interested*: researchers and practitioners who are not active in the field, but are interested in the domain and research in the related areas (e.g., human behavior analysis) can benefit from the book because it provides a clear account of state-of-the-art challenges and opportunities in the field and a clear positioning of the SSP research with respect to the related areas. Furthermore, the book can be an excellent entry point to the SSP domain.

- *Graduate and undergraduate students*: students at all levels will benefit from the book because the material is introductory and provides a clear explanation of what the SSP domain is about. In this respect, the book can help the students to decide whether SSP actually fits their interests or not.
- *Industry experts*: industry practitioners (or observers) can benefit from the book because they can find in it an extensive overview of the state-of-the-art applications in a wide spectrum of topics of potential interest as well as an indication on the most important actors in the domain.

Like any vibrant research field, social signal processing keeps developing in both depth and breadth. New conceptual and methodological issues emerge with continuity, often inspired by new application domains. Correspondingly, the editors hope that the chapters of this book will not be considered as a static body of knowledge, but as a starting point toward new research and application avenues. The goal of this book is not to provide the conclusive word on social signal processing, but to allow any reader to quickly engage with novelties and progress that will hopefully come in the years after the publication of the volume.

References

Brunet, P. & Cowie, R. (2012). Towards a conceptual framework of research on social signal processing. *Journal of Multimodal User Interfaces, 6*(3–4), 101–115.

Mehu, M. & Scherer, K. (2012). A psycho-ethological approach to social signal processing. *Cognitive Processing, 13*(2), 397–414.

Pentland, A. (2007). Social signal processing. *IEEE Signal Processing Magazine, 24*(4), 108–111.

Poggi, I. & D'Errico, F. (2012). Social signals: A framework in terms of goals and beliefs. *Cognitive Processing, 13*(2), 427–445.

Vinciarelli, A., Pantic, M., & Bourlard, H. (2009). Social signal processing: Survey of an emerging domain. *Image and Vision Computing Journal, 27*(12), 1743–1759.

Vinciarelli, A., Pantic, M., Bourlard, H., & Pentland, A. (2008). Social signal processing: State-of-the-art and future perspectives of an emerging domain. *Proceedings of the ACM International Conference on Multimedia* (pp. 1061–1070). New York: Association for Computing Machinery.

Vinciarelli, A., Pantic, M., Heylen, D., Pelachaud, C., Poggi, I., D'Errico, F., & Schroeder, M. (2012). Bridging the gap between social animal and unsocial machine: A survey of social signal processing. *IEEE Transactions on Affective Computing, 3*(1), 69–87.