Modelling user affect and sentiment in intelligent user interfaces: a tutorial overview

Björn Schuller

Angaben zur Veröffentlichung / Publication details:

Schuller, Björn. 2015. "Modelling user affect and sentiment in intelligent user interfaces: a tutorial overview." In *Proceedings of the 20th International Conference on Intelligent User Interfaces - IUI '15, March 2015, Atlanta, GA, USA*, edited by Oliver Brdiczka, Polo Chau, Giuseppe Carenini, Shimei Pan, and Per Ola Kristensson, 443–46. New York, NY: ACM Press. https://doi.org/10.1145/2678025.2716265.

Modelling User Affect and Sentiment in Intelligent User Interfaces

[A Tutorial Overview]

Björn W. Schuller[1,2]
[1]Department of Computing
Imperial College London
SW7 2AZ London, United Kingdom
+44-207-59-48357
bjoern.schuller@imperial.ac.uk
[2]Chair of Complex Systems Engineering
University of Passau
94032 Passau, Germany

ABSTRACT

The computer-based automatic analysis of human sentiment, and affect are broadly expected to play a major role that will likely make 'that difference' in future Intelligent User Interfaces, as they bear the promise to lend interactive systems emotional intelligence. Such comprise intelligent digital games, e. g., for empowerment and inclusion, tutoring systems, information systems or virtual companions, e. g., in the car to name but a few. This tutorial aims to give a good introduction into the related fields of user Sentiment Analysis and user Affect Modelling. Its intention is to show the general technology, and its current reliability, the ways for technical integration and efficient embedding of solutions in a user interface context, and the latest trends in this young and ever emerging field. Emphasis is laid on highlighting the range of toolkits available at this moment with the aim of empowering one to immediately craft own solutions. This description contains the general motivation, goals, objectives, and topics.

Author Keywords

Sentiment Analysis, Opinion Mining, Affective Computing, Intelligent Interaction

ACM Classification Keywords

H.1.2 User / Machine systems: Human information processing; H.5.5 Sound and Music Computing: Methodologies and techniques, Modeling, Signal analysis, synthesis, and processing; I.2.10 Vision and scene understanding: Video analysis; I.5.2 Design Methodology: Feature evaluation and selection, Classifier design and evaluation; I.5.4 Applications: Signal processing

General Terms

Human Factors, Signal processing, Feature evaluation and selection, Classifier design and evaluation, Experimentation

IUI'15, Mar 29 - Apr 01, 2015, Atlanta, GA, USA
ACM 978-1-4503-3306-1/15/03.
http://dx.doi.org/10.1145/2678025.2716265

INTRODUCTION

The computer-based automatic analysis of human sentiment [1, 12, 4], and affect [20, 10] are researched since more than a decade and a half by now, and are recently reaching increasing maturity. In fact, they are expected to play a major role that will likely make 'that difference' in future intelligent User Interfaces (IUIs), as they bear the promise to lend them emotional intelligence: Interfaces that 'know' and can react appropriately to, e. g., the satisfaction or anger of their users can lead us away from the often prevailing connotation of 'cold' and 'mechanical' that the interfaces of the current generation are still partially faced with. The information is thereby increasingly accessed from multiple modalities [15, 17] – both, in affect recognition and sentiment analysis – and 'in the wild' [33, 9], thanks to the availability of increasingly large and realistic resources [24] and improved algorithms [21] including deep learning [26] and long-short-term memory architectures [14] and weakly supervised learning methods [25]. Besides a certain focus on analysis in research to date, a system that emulates emotion is likely to be perceived as even more 'emotionally intelligent' given that it manages to overcome the uncanny valley.

In this light, this tutorial[1] aims to give a good introduction into the fields of user Sentiment Analysis and user Affect Modelling. Its intention is to show the general technology, and its current reliability, the ways for technical integration and efficient embedding of solutions in a user interface context, and the latest trends in this young and ever emerging field. Going one step further with a 'dash of tea leaf reading' it aims to distill likely future directions in and for this exciting

[1]cf. http://www.openaudio.eu for additional information.

and potentially game-changing discipline. In particular, emphasis is laid on highlighting the range of toolkits available at this moment to enable those interested in enriching current user interfaces by according social skills to immediately craft their own solutions.

In the following, the motivation, goals and objectives, and content related to a tutorial on modelling user affect and sentiment in the context of intelligent interfaces will be laid out.

GENERAL MOTIVATION

Socially aware and competent systems in general tend to be perceived as 'intelligent' and the connection and relevance to *intelligent* User Interfaces appears thus straightforward: Next generation IUIs will likely be increasingly expected to be able to understand factors such as user affect and exploit the automatic analysis of their users' sentiment for adaptation or related purposes.

The fields of Affective Computing and Sentiment Analysis both offer suitable approaches and methods to assess human states and are recently gaining much attention such as in workshops (even commonly, cf., e. g., [22]) or special issues (cf., e. g., [2, 3]) on the research level, but also more and more in real-world products.

GOALS AND OBJECTIVES

In detail, the goals and ambitions of this tutorial are threefold:

1. Introducing the existing efforts and major accomplishments in automatic, dimensional, and continuous analysis of emotions from multiple cues and user modalities;
2. Demonstrating and discussing the practical aspects, available frameworks, tools, databases, and automatic analysers, that can be easily used by user interface researchers around the world;
3. Encouraging the integration of the recent developments in the field into intelligent user interfaces.

The need for a tutorial arises thus from the demand to prepare the user interface community for an ongoing change of IUIs becoming more and more emotion-aware.

TOPICS

The topics to be touched upon can be roughly divided into theoretical and practical aspects besides aiming to distill future directions. They will be laid out in the following.

Theoretical Aspects

The theoretical aspects include in particular the 'savoir faire' when it comes to representation of affect and sentiment, different modalities, data, and the actual recognition framework including the feature extraction and classification.

Setting

The requirements needed are best derived from the specific *use cases* such as video gaming [16, 13] or other IUIs. In addition, *definitions* of the most relevant terms are required. The setting is complemented by benefits and limitations of sentiment and affect modelling per se.

Affect and Sentiment Representation

A number of different representation forms exist, predominantly those based on some type of *categories* – even in ordinal problems such as sentiment, e. g., by sentiment polarity or three classes ('low', 'middle', and 'high'), etc. [5], or with multiple categories per instance such as in *complex models*, and *tagging*. Further popular representations include the *circumplex model*, the value-continuous *Power-Arousal-Valence model*, and the *appraisal model* with varying *benefits and limitations*.

Modalities and Cues

The major modalities for the measurement and synthesis in the domain of IUIs are *speech and language*, *visual signals*, and *bio signals* – the latter, however, rather for analysis so far. In addition, haptic interaction analysis has ever since been considered with lower attention. Besides a range of different (dis-)advantages, these different modalities are known to be highly synergistic in terms of affect modelling.

User Data

Data is one of the major bottlenecks in this field. Major attention is thus given to efficient procedures of *data acquisition and annotation*, such as by weakly supervised [28] or even unsupervised [34] learning approaches. As it is mostly the labels that are sparse rather than the data, *crowd-sourcing and cooperative learning* [35] are further popular recent assets to lead quickly to rich resources of annotated affective usage data. Similarly, *transfer learning* [7] can be used to adapt 'similar' data to the current IUI case. In fact, a number of *benchmarks and important databases* exist today and are worth discussing with respect to the different types of user data and the challenges involved with them (cf. also in relation with the next subsection). Arguably there is a wide range of sentiment data for reviews, but less annotated video data or bio signals. A dominating problem is usually the amount of noise.

Automatic Analysis and Prediction

As outlined above, a number of information-bearing signals and methods exist to analyse users' affective state, including most notably *speech emotion recognition* [31], *opinion and sentiment analysis* from text or also multiple modalities, or the interpretation of *visual signals, motion capture, and thermal imaging signals* [11], and *bio signals* which are best (all) combined in a *synergistic modality fusion*. A number of different *classification schemes* are available to obtain best results depending on the application context and can be optimised in the loop with an according *performance evaluation*. More recently, *distribution* [36] of a recognition system's entities is becoming of interest, to provide also access on mobile IUIs. Further, to provide more information to the IUI than just the 'best guess' of a user's state, *confidence measurement* [6] can give an additional information on how reliable this 'guess' likely is.

Practical Aspects

To set up an 'affective' IUI, one has the option to go with existing toolkits and even to test it against a number of standardised benchmarks that will be touched upon next.

Frameworks and Tools

Today, luckily a number of *tools for user sentiment and affect modelling* exist [27]. To name but a very few, these include the *openSMILE* [8] audio/visual feature extractor, which has been and still is the official baseline feature extractor of the AVEC, INTERSPEECH ComParE, and MAPTRAITS comparative research competitions in the field of affect analysis (cf. also below). A further interesting piece of software particularly suited for IUI design can be the publicly available *Semaine system* [18] that offers a virtual agent sensing and synthesising affect and behaviour. In particular, for machine learning, deep and recurrent neural networks as provided by the *CURRENNT toolkit* [29] are at the state-of-the-art. However, a broad range of machine learning toolkits such as WEKA [32] offer a broad selection of alternative intelligence algorithms. A range of *suitable and necessary and needed tools* include, e. g., such for enhancement of the signal of interest as by blind source separation [30].

Competitions and Data Presentations

Interestingly, in addition to the increasing availability of data and tools, there has been an increasing number of competitions in the field. The majority of these have dealt with the analysis of sentiment and affect rather than their synthesis. This is likely due to the comparative ease with which analysis measures can be objectively compared. An *overview on competitions* includes the INTERSPEECH *ComParE 2009 – 2015* competition series, the *2011 – 2014 Audio/visual Emotion Challenges* (AVEC), the Emotion Recognition In The Wild challenges, a number of tasks within the MediaEval series, and those dealing also with traits such as the personality of users, e. g., the *MAPTRAITS 2014 Challenge*. In fact, most of these allow free access to the data that has been used, enabling one to *experiencing the selected data* in practice.

Future Directions

The future potential in this field is still huge owing to its comparably young age of roughly a decade and a half and the need for further improvements. Some of the current trends include more integration of user traits [23], also to better understand the user state, and gathering experience in 'the wild', best with integration of contextual knowledge. However, many further issues such as language and cultural independence are still far from being solved. Similarly, standardisation for the embedding in user interfaces is still needed in many ways, but is making good progress, such as by recent recommendations and discussion groups of the W3C [19].

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Communitys Seventh Framework Programme under grant agreements No. 338164 (ERC Starting Grant iHEARu) and No. 289021 (STREP ASC-Inclusion) and from its Horizon 2020 Framework Programme under grant agreement No. 645094 (Sentiment Analysis in the Wild – SEWA).

REFERENCES

1. Balahur-Dobrescu, A., Taboada, M., and Schuller, B. W. *Computational Methods for Affect Detection from Natural Language*. Computational Social Sciences. Springer, 2015. to appear.
2. Cambria, E., Schuller, B., Liu, B., Wang, H., and Havasi, C. Guest Editor's Introduction: Knowledge-based Approaches to Concept-Level Sentiment Analysis. *IEEE Intelligent Systems Magazine 28*, 2 (March/April 2013), 12–14.
3. Cambria, E., Schuller, B., Liu, B., Wang, H., and Havasi, C. Guest Editor's Introduction: Statistcial Approaches to Concept-Level Sentiment Analysis. *IEEE Intelligent Systems Magazine 28*, 3 (May/June 2013), 6–9.
4. Cambria, E., Schuller, B., Xia, Y., and Havasi, C. New Avenues in Opinion Mining and Sentiment Analysis. *IEEE Intelligent Systems Magazine 28*, 2 (March/April 2013), 15–21.
5. de Groot, R. *Data Mining for Tweet Sentiment Classification: Twitter Sentiment Analysis*. LAP LAMBERT Academic Publishing, November 2012.
6. Deng, J., and Schuller, B. Confidence Measures in Speech Emotion Recognition Based on Semi-supervised Learning. In *Proceedings INTERSPEECH 2012*, ISCA (Portland, OR, September 2012).
7. Deng, J., Zhang, Z., Eyben, F., and Schuller, B. Autoencoder-based Unsupervised Domain Adaptation for Speech Emotion Recognition. *IEEE Signal Processing Letters* (2014).
8. Eyben, F., Weninger, F., Groß, F., and Schuller, B. Recent Developments in openSMILE, the Munich Open-Source Multimedia Feature Extractor. In *Proceedings ACM Multimedia*, ACM (Barcelona, Spain, October 2013), 835–838.
9. Gunes, H., and Schuller, B. Dimensional and Continuous Analysis of Emotions for Multimedia Applications: a Tutorial Overview. In *Proceedings ACM Multimedia*, ACM, ACM (Nara, Japan, October 2012). 2 pages.
10. Gunes, H., and Schuller, B. Categorical and Dimensional Affect Analysis in Continuous Input: Current Trends and Future Directions. *Image and Vision Computing, Special Issue on Affect Analysis in Continuous Input 31*, 2 (February 2013), 120–136.
11. Gunes, H., Schuller, B., Pantic, M., and Cowie, R. Emotion Representation, Analysis and Synthesis in Continuous Space: A Survey. In *Proceedings International Workshop on Emotion Synthesis, rePresentation, and Analysis in Continuous spacE, EmoSPACE 2011 at FG 2011*, IEEE (Santa Barbara, CA, March 2011), 827–834.
12. Liu, B. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press, April 2015.
13. Marchi, E., Zhang, Y., Eyben, F., Ringeval, F., and Schuller, B. Autism and Speech, Language, and

Emotion - a Survey. In *Evaluating the Role of Speech Technology in Medical Case Management*, H. Patil and M. Kulshreshtha, Eds. De Gruyter, Berlin, 2015. to appear.

14. Metallinou, A., Wöllmer, M., Katsamanis, A., Eyben, F., Schuller, B., and Narayanan, S. Context-Sensitive Learning for Enhanced Audiovisual Emotion Classification. *IEEE Transactions on Affective Computing 3*, 2 (April – June 2012), 184–198.
15. Morency, L.-P., Mihalcea, R., and Doshi, P. Towards multimodal sentiment analysis: Harvesting opinions from the web. In *Proceeding of the 13th ACM International Conference on Multimodal Interfaces (ICMI)*, ACM (New York, USA, 2011), 169–176.
16. Paletta, L., Schuller, B. W., Robinson, P., and Sabouret, N. IDGEI 2015: 3rd international workshop on intelligent digital games for empowerment and inclusion. In *Proceedings IUI 2015*, ACM, ACM (Atlanta, GA, March 2015). 2 pages.
17. Raaijmakers, S., Truong, K., and Wilson, T. Multimodal subjectivity analysis of multiparty conversation. In *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2008)* (Edinburgh, UK, 2008), 466–474.
18. Schröder, M., Cowie, R., Heylen, D., Pantic, M., Pelachaud, C., and Schuller, B. Towards responsive Sensitive Artificial Listeners. In *Proceedings 4th International Workshop on Human-Computer Conversation* (Bellagio, Italy, October 2008). 6 pages.
19. Schröder, M., Devillers, L., Karpouzis, K., Martin, J.-C., Pelachaud, C., Peter, C., Pirker, H., Schuller, B., Tao, J., and Wilson, I. What should a generic emotion markup language be able to represent? In *Proceedings ACII 2007, Lisbon, Portugal*, A. Paiva, R. W. Picard, and R. Prada, Eds., vol. 4738/2007 of *LNCS*. Springer, Berlin/Heidelberg, 2007, 440–451.
20. Schuller, B., and Batliner, A. *Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing*. Wiley, 2013.
21. Schuller, B., Batliner, A., Steidl, S., and Seppi, D. Recognising Realistic Emotions and Affect in Speech: State of the Art and Lessons Learnt from the First Challenge. *Speech Communication 53*, 9/10 (November/December 2011), 1062–1087.
22. Schuller, B., Buitelaar, P., Devillers, L., Pelachaud, C., Declerck, T., Batliner, A., Rosso, P., and Gaines, S., Eds. *Proceedings of the 5th International Workshop on Emotion Social Signals, Sentiment & Linked Open Data (ES^{3}LOD 2014)*, ELRA, ELRA (Reykjavik, Iceland, May 2014). Satellite of LREC 2014.
23. Schuller, B., Steidl, S., Batliner, A., Nöth, E., Vinciarelli, A., Burkhardt, F., van Son, R., Weninger, F., Eyben, F., Bocklet, T., Mohammadi, G., and Weiss, B. A Survey on Perceived Speaker Traits: Personality, Likability, Pathology, and the First Challenge. *Computer Speech and Language 29*, 1 (January 2015), 100–131.
24. Schuller, B., and Vasileios, V. Sentiment Analysis and Opinion Mining: On Optimal Parameters and Performances. *WIREs Data Mining and Knowledge Discovery* (2015). 8 pages, to appear.
25. Schuller, B., and Weninger, F. Human Affect Recognition – Audio-Based Methods. In *Wiley Encyclopedia of Electrical and Electronics Engineering*, J. G. Webster, Ed. John Wiley & Sons, New York, 2015. to appear.
26. Stuhlsatz, A., Meyer, C., Eyben, F., Zielke, T., Meier, G., and Schuller, B. Deep Neural Networks for Acoustic Emotion Recognition: Raising the Benchmarks. In *Proceedings ICASSP 2011*, IEEE (Prague, Czech Republic, May 2011), 5688–5691.
27. Valentine, S. *Sentiment Analysis 19 Success Secrets - 19 Most Asked Questions On Sentiment Analysis - What You Need To Know*. Emereo Publishing, February 2014.
28. Wan, X. Co-training for cross-lingual sentiment classification. In *Proceedings of ACL* (2009), 235–243.
29. Weninger, F., Bergmann, J., and Schuller, B. Introducing CURRENNT – the Munich Open-Source CUDA RecurREnt Neural Network Toolkit. *Journal of Machine Learning Research 15* (2014). 5 pages.
30. Weninger, F., and Schuller, B. Optimization and Parallelization of Monaural Source Separation Algorithms in the openBliSSART Toolkit. *Journal of Signal Processing Systems 69*, 3 (2012), 267–277.
31. Weninger, F., Wöllmer, M., and Schuller, B. Emotion Recognition in Naturalistic Speech and Language - A Survey. In *Emotion Recognition: A Pattern Analysis Approach*, A. Konar and A. Chakraborty, Eds., 1st ed. Wiley-Blackwell, December 2014.
32. Witten, I., and Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, San Francisco, 2005.
33. Wöllmer, M., Weninger, F., Knaup, T., Schuller, B., Sun, C., Sagae, K., and Morency, L.-P. YouTube Movie Reviews: Sentiment Analysis in an Audiovisual Context. *IEEE Intelligent Systems Magazine 28*, 3 (May/June 2013), 46–53.
34. Zagibalov, T., and Carroll, J. Unsupervised classification of sentiment and objectivity in chinese text. In *Proceedings of the International Joint Conference on Natural Language Processing* (2008), 304–311.
35. Zhang, Z., Coutinho, E., Deng, J., and Schuller, B. Cooperative Learning and its Application to Emotion Recognition from Speech. *IEEE Transactions on Audio, Speech and Language Processing* (2014).
36. Zhang, Z., Coutinho, E., Deng, J., and Schuller, B. Distributing Recognition in Computational Paralinguistics. *IEEE Transactions on Affective Computing 21*, 9 (2014), 1068–1072.