A Framework Concept for Emotion Enriched Interfaces

Arne Bernin

HAL Id: hal-01556182
https://inria.hal.science/hal-01556182

Submitted on 4 Jul 2017

A Framework Concept for Emotion Enriched Interfaces

Arne Bernin

arne.bernin@haw-hamburg.de

Department Informatik, University of Applied Sciences Hamburg (HAW Hamburg), Germany
School of Computing,University of the West of Scotland (UWS), UK

Abstract. The work presented outlines my doctoral research in emotion and action based human computer interfaces, with the approach to create an open source framework. The idea is to find the suitable abstractions and interfaces to integrate different approaches to process input data to interpret user's emotions and actions. Fusing and comparing alternatives for processing will help to find the appropriate solution for a particular need for the next generation of interactive systems. Widening the sensing abilities of computers has the potential to add new degrees of freedom to the design of human computer interaction and ubiquitous entertainment computing.

1 Introduction

The level of interaction of today's computing systems lack various aspects that can be found in human-human communication. One major difference lies in ability to communicate on an emotional level. Combining emotions and actions shows great potential for new applications by leveraging of humanlike communication skills. This is particular relevant for communication centric applications such as gaming and learning. There are two fundamental designs to realize emotions in interactive Systems: 1. observing systems facilitate an emotional monologue i.e. provide no feedback. 2. Reactive systems have sensing and displaying abilities leading to an emotional dialog. The dialog may be facilitated by a combination of one or more modalities. Humans combine a variety of modalities as input channels for emotion recognition including gesture, posture, and prosody of speech, semantic, context and facial expressions.

Combining different data sources is also a common approach in computer science. Although there were some experiments in the field of combined channels for emotion recognition in computer science, the decision, which aggregation of modalities are more effective, is still an open research question. This decision also depends on the application and needs to be determined on the basis of a particular application scenario.

Currently there are many approaches in emotion interpretation and comparisons are rarely performed. Therefore, integrating established algorithms in a common framework enables to compare application specific differences and find opportunities for further research.

2 Framework

Current developments in entertainment computing go beyond the classic desktop scenario with one user in front of a display. Computer interaction has become ubiquitous. Therefore, the research scenarios should be developed in a ubiquitous context and laboratories, which we aim for in the livingplace hamburg, a smart home lab at the HAW Hamburg.[4]

The proposed framework enables to integrate different approaches in soft- and hardware and adds the ability to evaluate them on a scientific level. The requirements for the framework are:

1. Capability to log and replay all input data is needed for validation of an approach as well as comparison of different approaches with the same input data. Therefore, the complete data has to be written to data storage.
2. Processing of input data need to be performed at interactive rates. For this and to avoid bias due to fluctuating latencies, the framework has to be realtime orientated.
3. Concurrent processing of different modalities. The Framework needs the capacity to handle multiple input channels at the same time without fatal increase of processing time.
4. Facilities to analyze results for validation and comparison. In addition to logging of the input data, metrics of the system state, latencies, processing times and detection rate need to be saved for further analysis.
5. Easy exchange and integration of different processing components. This includes the requirements for an interface that works on different operating systems. Adapters for different programming languages are needed for flexibility.
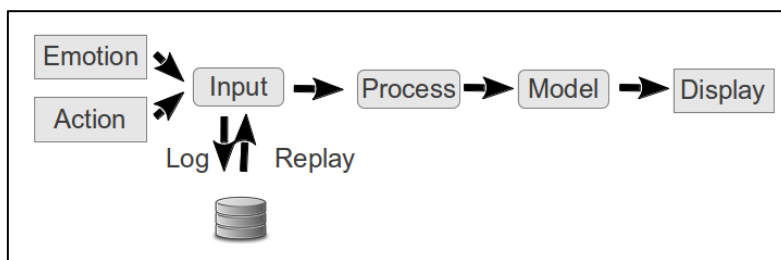

Figure 1: process flow

The flow model of processing is shown in Figure 1. The different components are:

Emotion: Sense describes the different input modalities as sensor input for different channels. Sense is gathered by observing the user with different technical systems. The user does not need to consciously act for expressing the sensed information as his current emotional state might be subconscious to him.

Action: Actions are performed by the user such as input to operate a game. Actions are actively and conscious expressions done by the user as part of his planned interaction.

Input: Input is an abstraction of Emotion and Action. Incoming data needs to be normalized for further processing. Input is also the step, where the data is first completely visible to the framework, so here is the place to log it.

Process: Input data needs to be filtered and processed. The aggregated Information is then used to update the emotional model.

Log/Replay: Capability to store all input data in a database for future evaluation.

Model: Emotional Model on Computational level. This model represents the current emotional state of the user as recognized by the system. It can also be used to validate state changes and avoid fluctuations caused by misattributed input data.

Display: Output modality. Appropriate Display is task for future research. The current interface for the livingplace as testing environment at the HAW is the central message broker.

3   Applications in Gaming

Sensing emotions opens a variety of applications for interactive systems. For my thesis work, i plan to evaluate my proposed framework in gaming scenarios. I envision two possible application domains in entertainment computing.

3.1   Serious Games

Serious games are used for learning. By Gamification of the learning context or by developing skills inside the game that also have a benefit in the real world. Emotion recognition has two major applications in the context of serious games: Feedback to the machine and feedback to the user. Feedback to the machine on an emotional level can be used to control the learning curve. Changing speed and level of presented facts in an automated learning system according to the speed of understanding can improve the efficiency of the learning process.[2] Provoking emotions can improve the user's motivation and memorization ability. Emotion recognition can be applied to verify the wanted effect. Further, emotional feedback to the user can improve can improve learning for public speaking and presentation skills.

3.2   Fun and Entertainment Games

There is an ongoing evolution in game controls from buttons and joysticks evolved to gestural input. In the future facial expressions and emotional state promise further enhancements. Expressed emotions can be transferred to avatar expression in game context. Like in serious games, Emotions can be applied to adapt the level of the game to user abilities and preferences. Current games use AI technologies, adding emotions have the potential to make them more believable.[1] Although, currently

unknown, using emotion recognition technologies could increase immersion and embodiment on the user side.

3.3 Contribution to Entertainment Computing

The main purpose for the planned framework is to evaluate different approaches of sensing emotions and corresponding models for research. Nevertheless, the framework and exemplary procedure can be applied to integrate emotional sensing and to design into entertainment computing and games. Our complete prototype system can be directly integrated in game engines as additional source of input.

4 Current and Future Work

The main focus of the current work is to develop the basics of the framework. Therefore, methods and technical components evaluation for realtime messaging systems, facial expressions and prosody of speech is in process, the procedure for framework development has been derived from this evaluations. Examples can be presented at the doctorial consortium. Suitable datasets for evaluation selected as well.

Future work includes the integration of a system for multi camera based determination of position and recognition of persons, currently developed at the HAW Hamburg. Integrating the results from the research done on terms of emotional surfaces at the HAW Hamburg are also part of future work. [3]

5 References

[1]   C. Becker-Asano. *WASABI: Affect Simulation for Agents with Believable Interactivity*. PhD thesis, Faculty of Technology, University of Bielefeld, 2008. IOS Press (DISKI 319).

[2]   S. D'Mello, T. Jackson, S. Craig, B. Morgan, P. Chipman, H. White, N. Person, B. Kort, R. El Kaliouby, R. Picard, and E. Al. AutoTutor detects and responds to learners affective and cognitive states. *Proceedings of the Workshop on Emotional and Cognitive issues in ITS in conjunction with the 9th International Conference on Intelligent Tutoring Systems*, 2008.

[3]   L. Müller, S. Keune, A. Bernin, and F. Vogt. Emotional interaction with surfaces - works of design and computing. *International Conference on Entertainment Computing*, 2012.

[4]   K. von Luck, G. Klemke, S. Gregor, M. A. Rahimi, and M. Vogt. Living place hamburg – a place for concepts of it based modern living. Technical report, Hamburg University of Applied Sciences, 2010.