

Research article

Benny Liebold*, René Richter, Michael Teichmann, Fred H. Hamker, Peter Ohler

Human Capacities for Emotion Recognition and their Implications for Computer Vision

DOI 10.1515/icom-2015-0032

Abstract: Current models for automated emotion recognition are developed under the assumption that emotion expressions are distinct expression patterns for basic emotions. Thereby, these approaches fail to account for the emotional processes underlying emotion expressions. We review the literature on human emotion processing and suggest an alternative approach to affective computing. We postulate that the generalizability and robustness of these models can be greatly increased by three major steps: (1) modeling emotional processes as a necessary foundation of emotion recognition; (2) basing models of emotional processes on our knowledge about the human brain; (3) conceptualizing emotions based on appraisal processes and thus regarding emotion expressions as expressive behavior linked to these appraisals rather than fixed neuro-motor patterns. Since modeling emotional processes after neurobiological processes can be considered a long-term effort, we suggest that researchers should focus on early appraisals, which evaluate intrinsic stimulus properties with little higher cortical involvement. With this goal in mind, we focus on the amygdala and its neural connectivity pattern as a promising structure for early emotional processing. We derive a model for the amygdala-visual cortex circuit from the current state of neuroscientific research. This model is capable of conditioning visual stimuli with body reactions to enable rapid emotional processing of stimuli consistent with early stages of psychological appraisal theories. Additionally, amygdala activity can feed back to visual areas to modulate attention allocation according to the emotional relevance of a stimulus. The implications of the model considering other approaches to automated emotion recognition are discussed.

Keywords: Emotions in HCI, Emotion Recognition, Neural Networks

*Corresponding author: Benny Liebold, Institute for Media Research, Chemnitz University of Technology, Chemnitz, Germany, e-mail: benny.liebold@phil.tu-chemnitz.de

René Richter, Michael Teichmann, Fred H. Hamker, Peter Ohler: Chemnitz University of Technology, Chemnitz, Germany

1 Introduction

In recent years, research has been increasingly concerned with the role of emotions during the interaction between user and interactive systems under the phrase *affective computing* (Picard 1997). The fundamental idea is that by “incorporating” emotions into interactive systems the interaction process would lead to more appropriate system responses (Gratch and Marsella 2004) and thereby to responses of the system user that are comparable to real situations (de Melo, Carnevale, and Gratch 2012, Krämer et al. 2013). Following these functions, affective computing systems have to fulfill two requirements: The system (1) needs to perceive emotional reactions by the user to successfully incorporate this information into the interaction process and (2) it requires the ability to generate emotional processes itself. While these two requirements are often considered as separate functions of artificial cognitive systems, we argue that both are strongly interconnected and necessitate each other. Thus, any system fulfilling one of those functions would automatically have the ability to fulfill the other function as well. To go even further, emotional information processing cannot be clearly separated from other domains of information processing, because they are highly integrated (Pessoa 2008, 2012, Okon-Singer et al. 2015). This argument rests upon the premise that affective capabilities of interactive systems can only be modeled appropriately, if they are modeled after the human cognitive system (Vitay and Hamker 2011, Canamero 2005). The idea behind this statement is the question, why we should expect artificial systems to perform equally well in the domain of emotion recognition compared to humans, although they follow a completely different logic of information processing.

In this article we will review the human capacity to perceive and process emotional information and present implications for the design of affective interactive systems that are capable of perceiving emotions in their users. We discuss the social functions emotions can serve in HCI and thereby demonstrate the importance of the endeavor to model emotional processes apart from being important for emotion recognition. Since we argue that affective computing as a

basis for automated emotion recognition needs to resemble affective processes in humans, we review findings from psychology and neuroscience concerning the human capacity for emotional information processing. Therein, we focus on appraisal theories of emotion and the role of the amygdala in the early evaluation of intrinsic stimulus properties. To provide implications of how systems for automated affect recognition can benefit from this knowledge, we suggest a biologically inspired model of the amygdala. It is able to recognize appraisals of intrinsic stimulus properties in others as a first step towards biologically motivated models of emotion recognition. Finally, we give an overview about common methods for automated affect recognition and discuss shortcomings of these approaches concerning our suggested model and the properties of the human neural circuit.

2 Affective Computing and HCI

Affective computing has been widely discussed in the domain of embodied conversational agents (ECAs; e. g. Gratch and Marsella 2005, Becker-Asano and Wachsmuth 2009, Scherer, Bänziger and Roesch 2010), which are capable of natural interaction. But why should researchers want to consider emotional processes for their virtual agent architecture? Since ECAs aim to emulate human behavior, we have to consider the functions emotions serve in humans. The intrapersonal functions of emotions in terms of their contribution to evolutionary adaptiveness have been widely discussed (Frijda 1986, Lazarus 1991, Levenson 1999). Yet, these functions are only of secondary importance for the most part of HCI research. In line with the emotions as social information model (van Kleef 2010) emotional processing and expressions serve a number of interpersonal functions in human communication both on the side of the receiver and the communicator. First, they allow us to better understand our social environment by observing emotional reactions of others towards and understanding their appraisals of third entities (Manstead and Fischer 2001). Thereby, emotional expressions can help disambiguating social situations (van Kleef et al. 2011) and may also inform us about the affective value of novel entities, resulting in approach or avoidance behavior. Second, emotional expressions of others allow us to understand their intentions and predict their behavior (Phillips, Wellman and Spelke 2002). This theory of mind (Premack and Woodruff 1978, Wimmer 1983) is a fundamental ability that evolved out of environmental pressures towards bigger social groups and constantly updates presumed mental states of others.

Third, emotional expressions communicate information about relationships or social status (Fischer and Manstead 2008) and serve group and individual differentiation (Levenson 1999). As a consequence, this wide array of social functions renders emotional expressions important to better understand others and our mutual environment. ECAs might use the affective information from the user to understand their evaluation of situations and could thereby learn the affective value of novel stimuli, and vice versa. The absence of emotional expressions that would otherwise fulfill these functions can have detrimental effects on the interaction process. Thus, when technology is not capable of conveying affective information correctly, the user might be prone to misinterpret an ECA's behavior due to misperceptions of affective states (Liebold and Ohler 2013). This might at the same time result in unfavorable evaluations of the ECA. Qu and colleagues (Qu et al. 2014), for example, found that random expressive behavior of virtual agents during conversation decreased conversation dissatisfaction.

Following the social functions of emotion expressions, affective computing in interactive systems has been shown to have positive effects on the interaction process. In his model of social influence, Blascovich (Blascovich 2002, Guadagno et al. 2007) proposed that agency and behavioral realism are fundamental constituents of social influence of ECAs. Emotional processes are relevant for both dimensions, because emotionally driven behavior might increase behavioral realism and might thereby also facilitate agency beliefs, which are deduced from observed behavior. The increased social influence of ECAs with affective computing capabilities has been demonstrated in several studies: De Melo, Carnevale, and Gratch (2012) found that emotion displays of an ECA influence user decisions during virtual negotiations. Regarding the underlying mechanism, they showed that users perceived the ECA's emotion displays as indicators of its behavioral intentions. Regarding mimicking behavior, Krämer, Kopp, Becker-Asano, and Sommer (2013) reported that users smiled longer over the course of a communication episode, when the ECA smiled at times. At the same time explicit evaluations were not different from the non-smiling ECA. Other research on the link between affective computing and user interaction focused more on the perception of ECAs. Demeure, Niewondowski, and Pelachaud (2011) found that emotional reactions of ECAs in general led to higher evaluations of their competence, warmth, and believability. The effect was greatest for emotion displays that were both appropriate and plausible — yet, even inappropriate and implausible displays led to better evaluations. In another study addressing agent evaluation, Bosse, and Zwanenburg (2009)

reported better evaluations of ECAs that were capable of producing emotional reactions.

Apart from these results, it was also found that nonverbal behavior outweighs the positive effects of emotion displays during interaction processes (Cassell and Thorisson 1999, Krämer, Iurgel and Bente 2005). Yet, one should not conclude that researchers and developers should instead focus on aspects of nonverbal behavior completely. The hard distinction between emotion and other nonverbal behavior does not consider the fact that emotional processing can inform the selection of appropriate nonverbal behavior, such as nodding (Lee and Marsella 2010). Thus, by modeling emotional processes as a fundamental mental capacity ECAs may cover a wider range of interaction scenarios than pre-defined conversational functions (Schönbrodt and Asendorpf 2011). Some ECA scenarios even require affective computing, such as virtual patients, which aim to train novice therapists conversational and diagnostic skills (Kenny et al. 2007, 2008). Additionally, apart from specific interaction contexts, modeling emotional processes is important from a research perspective, because it allows emotion researchers to test their theoretical predictions and thus advances our understanding of emotional processes in human cognition.

A core component of affective ECAs is their ability to affectively interact with the user rather than merely to produce emotional reactions. To achieve this goal, they need to recognize the user's affective states in the course of a conversation. While several systems for automated emotion recognition have been presented (for a review see Zeng et al. 2009), these approaches usually consider emotion recognition as an independent system, which categorizes emotional displays either using handcrafted rule-based detection or some kind of machine-learning algorithm. Research in emotion psychology and on the neural basis of emotional processing, however, suggest that emotion recognition is a highly integrated process that relies on more fundamental capabilities of emotional processing. Therefore, any emotion recognition system aiming for human-like capabilities needs to reflect affective processes in humans requiring interdisciplinary cooperation between computer science, emotion psychology, and cognitive neuroscience.

3 Human Emotion Processing

A great number of theories have been presented in the field of emotion psychology to answer James (1884) question: What is an emotion? The presented theories differ in their

focus and the proposed mechanism, how emotions are elicited, and consequently in their definition of the construct of emotion. Although up to today no consensus about these questions has been reached, most researchers agree that emotions are syndromes of subjective feelings, psychophysiological responses, actions tendencies, and some form of cognitive processing (Frijda 1986, Oatley and Jenkins 1996, Izard 2010, Mulligan and Scherer 2012). Yet, the nature of the involved cognitive processes remains subject to debates. Some researchers argue that emotions are evolved affect patterns or basic emotions that met specific evolutionary pressures (Ekman 1972, Izard 1977, Plutchik 1980) and therefore exist as distinct neural activation patterns (e.g. Vytal and Hamann 2010). Others argue that emotions are a result of more fundamental cognitive processes (automated situation appraisals), which are by themselves evolutionary adaptive (Scherer 1984, Lazarus 1991). A third position is that different emotion words are merely a result of language development and categorization of affect and that emotions in principle only differ along either two or three fundamental dimensions, which are valence, arousal, and dominance with the first two being generally agreed upon in this perspective (Mehrabian and Russell 1974, Russell and Barrett 1999). A recent extensive meta-analysis on the neurophysiological evidence for these approaches was presented by Lindquist and colleagues (Lindquist et al. 2012). They found that the findings lend more support to appraisal theories than theories that assume distinct neural activation patterns for each emotion.

3.1 Appraisal in Affective Processing and Affect Recognition

While all three approaches provide a considerable body of empirical research, appraisal theories appear to be the most promising to cover human emotion processing in general for several reasons: (1) Emotions are directly derived from situation appraisals, which are — although usually not directly stated — fundamental components of other emotion theories as well, since situational information needs to be processed nonetheless. (2) Appraisal theories conceptualize emotion elicitation as processes highlighting temporal changes in emotional reactions with subsequent situation appraisals. (3) The appraisal processes during early stages are closely tied to basic human perception, which allows drawing links to involved mechanisms at a neural level. Thus, appraisal theories do not require the assumption that emotional reactions exist as distinct neural mechanisms as in the case of basic emotions, because they are a result of the involved appraisal processes.

One of the most influential appraisal theories was presented by Scherer (1984, 2001). In his component process model, he argues that emotions emerge from four subsequent groups of stimulus evaluation checks assessing relevance, implications, coping potential, and normative significance of an event. They directly serve proximal functions, such as directing attention, and triggering information search as well as identifying suited behavioral patterns, such as the fight-flight compound. If, for example, one of our friends insists on driving us home after having drunk too much, this sequence of appraisals can explain our emotional reactions. First, the event is personally relevant, because we did not expect to catch a ride that easy, being offered something is intrinsically pleasant, and it serves our goal to get home, which we would have needed to address nonetheless (positive surprise). Second, the consequences of our friend's offer are only that promising at first glance, as he appears to be rather drunk. Our friend consciously made the offer, although obviously not being able to drive a car. The consequences of driving a car in this state are worrisome to say the least and require our immediate response (fear). Third, we are sure that we can cope with the situation appropriately, because we are probably able to convince our friend, not to drive this evening — we just have to try hard enough (anger). Fourth, we realize that our friend's offer would not only harm herself, but also us and others, which is against societal and our own normative standards (blame). As outlined in this example, each of the four sequential appraisal check depends on several individual fundamental appraisals of a situation (Scherer 2001).

Over the course of these stimulus evaluation checks, reactions of the autonomic and somatic nervous system as well as the neuro-endocrine system are triggered. Scherer (2000) explains the emergence of emotional feelings by the fact that the involved systems become synchronized (compared to baseline activity) as a consequence of appraisal outcomes. This synchronization in combination with intrinsic valences attached to each appraisal check become conscious and are perceived as an emotional feeling (Scherer 2013). In essence, this explanation is a more sophisticated approach to James' (1884) notion that emotional reactions stem from interoceptive feedback. Since each sequential appraisal check provides efferent feedback to the involved systems, Scherer's approach accounts for temporal variations of emotional episodes and can account for a wider range of respective emotions. The component process model recently gathered an impressive body of empirical evidence regarding the nature and sequence of occurring appraisals not only

from self-report data, but also from analyses of facial expressions, responses of the autonomic nervous system, and data regarding brain topology and activity (for an overview see Scherer 2013).

To model emotion recognition, with human emotion processing in mind, it is necessary to identify links between Scherer's component process model and emotional displays. Scherer and Ellgring (2007a, 2007c) argued that sequential appraisal checks are linked to specific expressive patterns rather than expressions resulting from innate motor programs for basic emotions that are selected, when specific situations are encountered (Ekman 1972). Thus, modeling sequential appraisal checks instead of linking complete expressive patterns to elicited emotions would both allow creating more authentic virtual emotion displays and — perhaps more importantly — would allow the system to understand situational appraisals in others. The latter becomes possible, because the system is then able to detect expression components (e.g. the raise of eye-brows), from which specific appraisal processes, and subsequently emotions can be deducted. Contrary to this approach most systems for automated emotion recognition categorize emotions according to fixed expressive patterns — yet, some evidence for links between appraisal checks and specific expressions has been presented. This is especially true for appraisals in the early stage of emotional processing, such as appraisals of novelty, intrinsic pleasantness, and goal-conduciveness (for an overview see Scherer 2013). We believe that this link between basic emotional processing capabilities and the ability to recognize emotions in others is a key component of automated emotion recognition, since the mechanisms involved in situation appraisals have to be employed during emotion recognition as well.

At this point it is important to reconsider the functions emotion recognition serves in human cognition: We are not able to recognize emotional expressions merely to be informed about another person's emotional state, but to use this information to make inferences about their beliefs, desires, and intentions. In this context, de Melo, Carnevale, Read, and Gratch (2014) found that observers combine observed emotion expressions and social context into inferences about appraisals in others. While their findings support the notion that emotion recognition can be considered a process of appraisal recognition, it also stresses the importance of social context. The same expressive behavior in different social context can result in divergent or even detrimental inferences about the beliefs, desires and intentions of others. Thus, in the long run, situation appraisals need to be taken into account for automated

emotion recognition — or more precisely, for a cognitive model of theory of mind.

Since appraisal processes result from complex processes in the human brain, any model of emotion recognition needs to reflect these processes to some extent. One intriguing approach to this problem is to directly model the involved neural processes. Yet, it should be mentioned that modeling the neural basis of sequential situation appraisals can be considered a huge undertaking, since the involved mechanisms span from very fundamental mechanisms closely tied to perception to mechanisms requiring complex conscious processing in prefrontal cortex areas (Kane and Engle 2002). As a first step, researchers should focus on early appraisals that reflect highly automated processing and represent the foundation for every emotional process. These processes do not require attentional resources, but can result in attention being allocated to the emotional stimulus (Taylor and Fragopanagos 2005). According to the componential process model, these early processes include appraisals of novelty, intrinsic pleasantness, and goal-conduciveness for which the amygdala with its manifold connections to other brain areas has been suggested as the responsible neural substrate for the detection of emotional relevance (Sander, Grandjean and Scherer 2005).

3.2 Neural Basis of Early Affective Visual Processing

The processing of visual stimuli in the brain and the role of emotional cues has received ample attention in the field of neuroscience (for a review see Pessoa and Adolphs 2010). As a stimulus passes through several visual areas, neurons respond to more and more complex elements of the stimulus with each area (Creem and Proffitt 2001). The pre-processed information is forwarded to the amygdala (McDonald 1998, Sah et al. 2003), which is bidirectionally connected with several layers of the visual cortex (Catani et al. 2003, Gschwind et al. 2012).

The amygdala is believed to serve at least two central functions during affective processing. First, when a stimulus is presented several times before an emotion-laden situation then the amygdala learns to combine the received information from the visual cortex with the affective body reaction to the emotional situation in a similar fashion as in classical conditioning (Lin et al. 2010). As a result, the now emotion-laden visual stimulus directly influences the bodily reactions. This rapid modification of the autonomic and somatic nervous system based on conditioned

visual cues represents a core functionality of the amygdala (LeDoux et al. 1988, Hitchcock and Davis 1991, Shi and Davis 1999). It has been argued that these automated reactions represent the amygdala's role in the detection of emotionally arousing stimuli irrespective of their valence (Sander, Grandjean and Scherer 2005, Kensinger and Schacter 2006). The prefrontal cortex (PFC) on the other hand is believed to be responsible for the modulation of amygdala activity based on the evaluation of a stimulus. However, this likely important connection has not been satisfyingly defined (Shenhav and Greene 2014) and will not be specified in more detail. Still, these findings can be considered consistent with the early relevance appraisal of intrinsic stimulus properties of novelty, pleasantness, and goal-conduciveness (Scherer 2013).

A second suggested function of the amygdala is the guidance of attention to emotional stimuli (Holland and Gallagher 1999, Anderson and Phelps 2001). This is achieved via feedback connections to the visual cortex (Bonda 2000, Amaral, Behnia and Kelly 2003) that increase the activation of the neural pattern representing the affective stimulus and thereby influence the perception within the visual field (Schwabe et al. 2011). Further, recent research suggests that the amygdala receives information over several visual processing pathways in parallel and can simultaneously influence the processing in different stages due to bidirectional connections to the visual cortex (Pessoa and Adolphs 2010). This mechanism facilitates emotional guidance of vision and is important to cope with the complexity of our visual world. It enables us to concentrate on emotionally relevant stimuli while reducing the distraction by irrelevant stimuli.

The manifold connections of the amygdala to the visual cortex in combination with the conditioning of intrinsic stimulus properties also facilitate the immediate recognition of emotional signals in interpersonal communication. A possible mechanism can be seen in the concurrent perception of a conditioned stimulus and facial displays of others. Upon repeated presentation the amygdala associates the facial display with the triggered reaction to the arousing stimulus. Thus, the facial display becomes an emotional display itself, which can be considered as a mechanism for rapid affect recognition in early stages of affective processing.

3.3 A Brain-Inspired Model of Affect Recognition

Implementing these parallel and automatic processes into a machine recognition system will enable us to significantly increase the appropriateness of system responses

in human-machine interaction. Further, it can at the same time be a helpful approach to solve the bottleneck problem of processing all the information of the system's environment by focusing on the relevant information. Recently, fundamental progress has been made within the field of computational neuroscience to facilitate the implementation of such a circuit. Pessoa and Adolphs (2010) describe the neuroscientific foundations for a potential architecture of such a system. Other groups have examined the computational microcircuits for processing visual information and included attentional feedback (e. g. Beuth and Hamker 2015). This research on attention circuits in the visual system demonstrates how the visual cortex integrates feed-back signals to obtain attentional modulation for stimuli. Further, this can be used to integrate other kinds of sources for attentive feedback into the visual processing. Accordingly, in such a circuit the amygdala would play the role of a higher area receiving information from the cortex and modulating it reversely through its feedback.

To simulate the amygdala, it is important to differentiate between at least two emotionally relevant parts of the amygdala. First, the basolateral amygdala (BLA) receives input from various human senses (McDonald 1998, Sah et al. 2003). It can connect a salient stimulus with an already emotion-laden stimulus (Bergstrom et al. 2012) and feed back to the visual cortex (Amaral, Behnia and Kelly 2003, Vuilleumier et al. 2004). Second, the central amygdala (CE) receives information from BLA (Pape and Pare 2010) and modulates the current body state through its output connections (LeDoux et al. 1998). By this basic fragmentation, the amygdala converts the incoming information into a body reaction without adapting to a specific emotional state.

Based on the current understanding of the involvement of the amygdala into affective processes and its structural integration into the brain we propose a conceptualization for a model of visual emotion processing. Considering the previously given arguments, the model would be able to learn intrinsic stimulus properties, use these processes to understand facial expressions of emotion, and direct attention at emotionally relevant stimuli. The model should consist of the important ventral visual areas V1, V2, V4, and the inferior temporal cortex (IT), with their typical hierarchical order and bidirectional connections with each other. With this connectivity pattern the feed-back from higher areas can be propagated back to lower ones. The amygdala, on the other hand, consisting of BLA and CE should be bidirectionally connected to V4 and IT to modulate the visual processing. Furthermore, BLA should receive input from PFC to regulate BLA activity according to the valence of the stimulus, which is encoded by PFC based on its input from IT (see Figure 1).

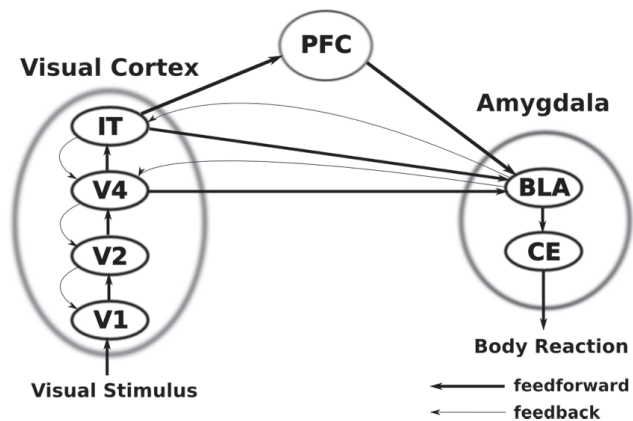


Figure 1: Conceptual structure for a hierarchical model of human-like emotion processing. BLA: basolateral amygdala, CE: central amygdala, IT: inferior temporal cortex, PFC: prefrontal cortex, V1, V2, V4: visual areas 1, 2, 4.

The model is capable of learning intrinsic stimulus properties and thus emotional relevance detection by conditioning the stimulus with body reactions in BLA. After several repetitions of the stimulus conditioning, the model will automatically increase arousal by triggering body reactions representing the appraisal of intrinsic pleasantness and goal-conductiveness. This mechanism also accounts for the detection of emotionally relevant facial features. As soon as the model has learned relevant facial features, the aforementioned appraisals would be automatically activated in tandem, which can be considered as early affect recognition. At the same time, BLA-activity is fed back to higher areas of the visual cortex to facilitate attention distribution towards the conditioned stimulus. This behavior is consistent with the novelty appraisal proposed by Scherer (2013). Further, the hedonistic value of the stimulus can be used by PFC to modulate amygdala activity. Thereby, positive evaluations of an event will inhibit amygdala activity.

4 Comparing Human Emotion Recognition and Current Computational Models

In order to give interactive systems the ability to interact with the user on an affective level, a broad variety of affect-recognition models have been developed over the last decade. The analysis of emotional facial expressions is the most common form of automated emotion recognition (Calvo and D'Mello 2010), since it is a very reliable emotional cue with a rich body of related research.

Nonetheless, by interpreting the prosody features of spoken words or the body posture of a person, similar information can be obtained and several models have been developed for these kinds of expressions (Zeng et al. 2009, Kleinsmith and Bianchi-Berthouze 2013, Ye and Fan 2014, Javier et al. 2015). While visual and auditory cues are usually the most important sources of information for human-human interaction, in case of HCI the sources can vary more widely. Data can also be acquired by measuring electrical brain activity with an electroencephalograph (EEG), brain imaging through functional magnetic resonance imaging (fMRI), or recording of vital signs, and be used in computational models to perceive the emotional state of the human counter-part (Calvo and D'Mello 2010). Yet, these approaches are based on information that are typically not available in HCI contexts and would greatly impair the usability of interactive systems. Thus, we focus on visual models and particularly models for facial expression recognition.

Models extracting emotions from human faces are numerous (Bartlett et al. 2006, Gunes and Piccardi 2007, Li et al. 2013). An extensive survey as well as an overview of related reviews has been presented by Zeng, Pantic, Roisman, and Huang (2009). Multiple computational models were found to perform reasonably well on images and video streams (range between 39.58 and 95 depending on the test setup). In most cases these models categorize emotions either by machine learning algorithms or rule-based approaches. Common methods are the tracking of facial points and relating them to facial action units (e.g. Kotsia et al. 2008, Valstar and Pantic 2012) or the extraction of shape features (e.g. Sarnarawickrame and Mindya 2013, Lozano-Monazor et al. 2014) as Gabor wavelets, followed by a classification, for example through support vector machines (e.g. Whitehill et al. 2009, Piatkowska and Martyna 2012, Lozano-Monazor et al. 2014). To deal not only with static expressions, some models use the output of the classifier as input to an hidden Markov model to analyze the temporal appearance of an expression (e.g. Barlett et al. 2003, Koelstra, Pantic and Patras 2010, Lahbiri et al. 2013). Even neural networks have been used for emotion recognition as alternative approach for extracting sophisticated shape features (e.g. Prevost, Belaroussi and Milgram 2006, Senechal, Prevost and Hanif 2010), as direct emotion classifiers (e.g. Perikos, Ziakopoulos and Hatzilygeroudis 2014), or even for the whole facial analysis (e.g. Kukla and Nowak 2015). Neural networks can also be used to integrate features from different information channels. Yet, the proposed models differ greatly in their classification of affective states and

approach to facial feature selection (between 2 and 9 emotion categories according to Zeng et al. 2009).

All these approaches differ substantially from an implementation based on neuroscientific evidence. Certainly, the visual cortex can be seen as a feature extracting stage. However, the amygdala is not merely a classifier to categorize the visual input. As previously explicated, the amygdala links visual information to body signals, whereby no classification in the sense of categorizing stimuli into discrete classes (e.g. emotions) has to be carried out. Consequently, there is no explicit decision for a particular kind of perceived emotion – rather, a stimulus invokes a previously learned response according to its intrinsic emotional valence. Since we follow a more fundamental approach where stimulus and response are associated through learning, visual features do not need to be mapped on distinct classes. Thus, whenever a learned stimulus is detected, a reaction is immediately triggered.

Whereas, current computational approaches remain limited to a few easily distinguishable emotional expressions that are usually considered to represent basic emotions, a neuroscience-inspired model of emotion recognition can account for more fundamental emotional processes. In our model, these processes are responsible for detecting emotional relevance (novelty and intrinsic pleasantness) of a stimulus as an early step in emotional processing without higher cognitive involvement. While it could be argued that this is essentially a step back from identifying more or less concrete emotional feelings, this approach has several advantages. It is inherently plausible, because the goal of all approaches is to mimic the human capacity to recognize emotions and the most plausible means is to mimic the human brain. Once such a model has been developed, it can be easily implemented together with other models of human cognition. The good compatibility of these models results from the fact that they account for universal functions of human cognition. They can complement each other, but do not fundamentally change with the context. As a consequence, they can also be applied in various contexts and are not limited to special cases, such as emotion recognition from facial features. In fact, the proposed amygdala model covers both the production and recognition of emotional processes and is independent of the modality, in which emotional displays of the interaction partner are encoded. Lastly, modeling emotion recognition as appraisal recognition accounts for the complex environmental conditions that lead to an emotional expression in others. An angry face is never merely an angry face, but an indicator of an arousing stimulus of negative and sometimes norm-violating quality, for which the other person is confident that she is able to overcome the problem. Understanding intrinsic stimulus

properties is the first step towards the necessary appraisals for these kinds of judgments.

Since the proposed processes need to be modeled as brain-inspired neural networks, this type of model comes in tandem with other beneficial features. First, the mechanisms to learn associations between stimuli and reactions are directly implemented into the model. Thus, the model will not have to reach a final state, which is then implemented into an interactive system — rather the model is still capable of learning associations after it has been implemented. Second, emotional processing actively influences feature extraction in the visual cortex. It directs attention to emotionally relevant features and entities, facilitating their processing (Adolphs and Spezio 2006). This mechanism should prove useful for HCI as well, when only limited resources for scene analysis are available and a more precise analysis of relevant features is necessary.

5 Conclusion and Challenges

Integrating emotional processes into HCI technology is an important research field to render interactive systems capable of more appropriate system responses to emotional user behavior. An important starting point for these endeavors is the recognition of affective user states. Although, numerous methods for the automated recognition of facial expressions have been presented, the methods cannot match the human ability to recognize emotions and incorporate them into interactions. As a first step to overcome these limitations, we suggest a substantially different approach to common emotion recognition systems. Instead of categorizing emotion indicators (i. e. facial features) into distinct emotion categories we suggest modeling emotion recognition after the human capacity for emotional processing in general. This is necessary, because the models for emotion recognition aim to mimic outcomes of human information processing and the involved mechanisms in humans are responsible for both emotion elicitation and recognition. This goal can be achieved by incorporating neuroscientific evidence on the visual cortex-amygdala circuit and related findings of emotion psychology. In the proposed model visual information is conditioned to automatically elicit body reactions and feedback to higher visual areas to enable affect-based attention allocation.

This approach has the advantage of not being limited to few distinct and complete facial expressions but tackling the problem at the level of single facial features. Furthermore, the association to body signals can be learnt during real interaction settings keeping the system flexible, when

employed in new contexts. The model in general is not restricted to be used only for emotion recognition, because it also generates visual output modulated by emotional relevance at the same time, which can be used for further processing and decision stages. These stages do not necessarily need to be modeled as neural systems considering the current lack of advanced neural models for higher cortical areas and processes, respectively. Yet, the resulting visual information and appraisals of our model can be integrated into existing computational architectures for recognition, decision-making, or higher-level emotional processing (e. g. Becker-Asano 2013) without strict adherence to the neurophysiological basis. This way, more complex affective processes can be modeled without the need of having a complete understanding of the interactions between involved brain areas — which we simply do not have at the moment. In such integrated systems, the emotional feedback in our system could, for example, still guide attention allocation to the relevant scene information, easing the bottleneck problem of scene analysis. Consequently, our proposal focuses on modeling early stages of emotional processes and should provide a sound basis for further efforts to model artificial emotions in line with automated emotion recognition and expression.

This approach leaves several aspects of human emotions and their technological implementation untouched. (1) Human emotions are oftentimes a result of more complex situational appraisals. However, these processes rely on more complex interactions within our neural system involving at least memory and conscious reasoning. Thus, we have focused on appraisals of intrinsic stimulus properties, accounting for emotional processes, upon which subsequent appraisals are based. These can only be modeled given that we have robust models of the necessary fundamentals. (2) Our approach focuses on appraisal theories. While compelling evidence for this viewpoint has been presented, not all emotion researchers agree that universal situation appraisals are the foundation of emotional processes. The basic emotion approach stating that emotions are based on evolved, but distinct, neural circuitry specific to each emotion has received reasonable attention both in psychology (e. g. Ekman 1972, Izard 1977, Picard 1997) and in the neurosciences (e. g. Phan et al. 2002, Murphy, Nimmo-Smith and Lawrence 2003, but see Lindquist et al. 2012). (3) Emotions are expressed as multimodal arrangements of several nonverbal channels including facial expressions, gestures, body posture, behavioral patterns, and tonal changes in the voice. Moreover, merely relying on facial expressions might lead to misclassifications, since social norms and deceptive intentions may lead to the suppression of emotional expressions. (4) It

remains uncertain, whether emotions occurring in natural settings always achieve the desired intensity threshold in their expression that is required to achieve robust tracking results. Automated emotion recognition algorithms usually perform better under lab conditions compared to real world encounters, where emotion expressions occur less prototypically. More so, expressions might as well be very subtle rendering it hard to classify them as significant indicators of emotional processes. Possibly, multimodal feature extraction might increase tracking reliability under these circumstances. (5) Since brain-inspired models of emotion recognition involve the simulation of complex models of neural networks, it remains a challenge to scale these models appropriately both to be capable of real-time processing and perhaps even to cover low-budget solutions with even less processing power.

Future work should further focus on implementing computer models for early stages of emotional processing. Although we believe in the benefits of this approach, first prototypes have to assess their ability for affective processing using standardized as well as natural datasets. As evidenced in this article, these problems need to be tackled by interdisciplinary research groups from psychology, cognitive neuroscience, and computer science to incorporate findings from various fields to create more lifelike and engaging interactive systems for HCI.

References

- Adolphs, R. and M. Spezio. 2006. Role of the amygdala in processing visual social stimuli. *Prog. Brain. Res.* 156: 363–378.
- Amaral, D. G., H. Behnia and J. L. Kelly. 2003. Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience*. 118(4): 1099–1120.
- Anderson, A. K. and E. A. Phelps. 2001. Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, 411(6835): 305–309. doi: 10.1038/35077083
- Barlett, M. S., G. Littlewort, P. Braathen, T. J. Sejnowski and J. R. Movellan. 2003. A prototype for automatic recognition of spontaneous facial actions. In: (S. Thrun and L. K. S. Saul, B., eds.), *Advances in Neural Information Processing Systems*. NIPS, Vancouver, CA, pp. 1271–1278.
- Bartlett, M. S., G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan. 2006. Fully automatic facial action recognition in spontaneous behavior. *7th International Conference on Automatic Face and Gesture Recognition (FG06)*.
- Becker-Asano, C. 2013. *WASABI for affect simulation in human-computer interaction: architecture description and example applications*. Paper presented at the Emotion Representations and Modelling for HCI Systems Workshop, Sydney, Australia.
- Becker-Asano, C. and I. Wachsmuth. 2009. Affective computing with primary and secondary emotions in a virtual human. *Auton. Agent Multi Agent Syst.* 20(1): 32–49. doi: 10.1007/s10458-009-9094-9
- Bergstrom, H. C., C. G. McDonald, S. Dey, H. Tang, R. G. Selwyn and L. R. Johnson. 2012. The structure of Pavlovian fear conditioning in the amygdala. *Brain Struct. Func.* 218(6): 1569–1589.
- Beuth, F. and F. H. Hamker. 2015. A mechanistic cortical microcircuit of attention for amplification, normalization and suppression. *Vision Res.* doi: 10.1016/j.visres.2015.04.004
- Blascovich, J. 2002. A theoretical model of social influence for increasing the utility of collaborative virtual environments. In: (W. Broll, C. Greenhalgh and E. F. Churchill, eds), *Collaborative virtual environments: Proceedings of the 4th international conference on collaborative virtual environments* ACM, New York, pp. 25–30.
- Bonda, E. 2000. Organization of connections of the basal and accessory basal nuclei in the monkey amygdala. *Eur. J. Neurosci.* 12(6): 1971–1992.
- Bosse, T. and E. Zwanenburg. 2009. There's always hope: enhancing agent believability through expectation-based emotions. 1–8. doi: 10.1109/acii.2009.5349424
- Calvo, R. A. and S. D'Mello. 2010. Affect detection: an interdisciplinary review of models, methods, and their applications. *IEEE Trans. Affect. Comput.* 1(1): 18–37.
- Canamero, L. 2005. Emotion understanding from the perspective of autonomous robots research. *Neural Netw.* 18(4): 445–455. doi: 10.1016/j.neunet.2005.03.003
- Cassell, J. and K. R. Thorisson. 1999. The power of a nod and a glance: envelope vs. emotional feedback in animated conversational agents. *Appl. Artif. Intell.* 13(4–5): 519–538. doi: 10.1080/088395199117360
- Catani, M., D. K. Jones, R. Donato and D. H. Ffytche. 2003. Occipito-temporal connections in the human brain. *Brain*. 126(9): 2093–2107.
- Creem, S. H. and D. R. Proffitt. 2001. Defining the cortical visual systems: “What”, “Where”, and “How”. *Acta Psychol.* 107(1–3): 43–68.
- de Melo, C. M., P. Carnevale and J. Gratch. (2012). The effect of virtual agents' emotion displays and appraisals on people's decision making in negotiation. In: (Y. Nakano, M. Neff, A. Paiva and M. Walker, eds) *Intelligent virtual agents, 12th International Conference, IVA 2012, Santa Cruz, CA, USA*. Springer, New York, pp. 53–66.
- de Melo, C. M., P. J. Carnevale, S. J. Read and J. Gratch. 2014. Reading people's minds from emotion expressions in interdependent decision making. *J. Pers. Soc. Psychol.* 106(1): 73–88. doi: 10.1037/a0034251
- Demeure, V., R. Niewiadomski and C. Pelachaud. 2011. How is believability of a virtual agent related to warmth, competence, personification, and embodiment? *Presence: teleoperators and virtual environments*. 20(5): 431–448. doi: 10.1162/PRES_a_00065
- Ekman, P. 1972. Universals and cultural differences in facial expression of emotion. In: (J. D. Cole, ed) *Nebraska symposium on motivation, 1971*. University of Nebraska Press, Lincoln, pp. 207–282.
- Fischer, A. H. and A. S. R. Manstead. 2008. Social functions of emotion. In: (M. Lewis, J. Haviland-Jones and L. Feldmann-Barrett, eds) *Handbook of emotions*, 3rd ed. Guilford Press, New York, pp. 456–468.

- Frijda, N. H. 1986. *The emotions. Studies in emotion and social interaction*. Cambridge University Press, Cambridge, MA.
- Gratch, J. and S. Marsella. 2004. A domain-independent framework for modeling emotion. *Cogn. Syst. Res.* 5(4): 269–306.
- Gratch, J. and S. Marsella. 2005. Lessons from emotion psychology for the design of lifelike characters. *Appl. Artif. Intell.* 19(3–4): 215–233. doi: 10.1080/08839510590910156
- Gschwind, M., G. Pourtois, S. Schwartz, D. Van De Ville and P. Vuilleumier. 2012. White-matter connectivity between face-responsive regions in the human brain. *Cereb. Cortex.* 22(7):1564–1576.
- Guadagno, R. E., J. Blascovich, J. N. Bailenson and C. McCall. 2007. Virtual humans and persuasion: the effects of agency and behavioral realism. *Media Psychol.* 10(1): 1–22. doi: 10.1080/15213260701300865
- Gunes, H. and M. Piccardi. 2007. Bi-modal emotion recognition from expressive face and body gestures. *J. Netw. Comput. Appl.* 30(4): 1334–1345.
- Hitchcock, J. M. and M. Davis. 1991. Efferent pathway of the amygdala involved in conditioned fear as measured with the fear-potentiated startle paradigm. *Behav. Neurosci.* 105(6): 826–842.
- Holland and Gallagher. 1999. Amygdala circuitry in attentional and representational processes. *Trends Cogn. Sci.* 3(2): 65–73.
- Izard, C. E. 1977. *Human emotions*. Plenum Press, New York.
- Izard, C. E. 2010. The many meanings/aspects of emotion: definitions, functions, activation, and regulation. *Emot. Rev.* 2(4): 363–370. doi: 10.1177/1754073910374661
- James, W. 1884. What is an emotion? *Mind.* 9(34): 188–205.
- Javier, G., D. Sundgren, R. Rahmani, A. Larsson, A. Moran and I. Bonet. 2015. Speech emotion recognition in emotional feedback for Human-Robot Interaction. *IJARAI.* 4(2).
- Kane, M. J. and R. W. Engle. 2002. The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective. *Psychon. Bull. Rev.* 9(4): 637–671. doi: 10.3758/bf03196323
- Kenny, P., T. D. Parsons, C. S. Pataki, M. Pato, C. ST-George, J. Sugar and A. A. Rizzo. 2008. Virtual justice: a PTSD virtual patient for clinical classroom training. *ARCTT.* 6: 111–116.
- Kenny, P., T. D. Parsons, J. Gratch, A. Leuski and A. A. Rizzo. 2007. Virtual patients for clinical therapist skills training. 4722: 197–210. doi: 10.1007/978-3-540-74997-4_19
- Kensinger, E. A. and D. L. Schacter. 2006. Processing emotional pictures and words: effects of valence and arousal. *Cogn. Affect. Behav. Neurosci.* 6(2): 110–126. doi: 10.3758/cabn.6.2.110
- Kleinsmith, A. and N. Bianchi-Berthouze. 2013. Affective body expression perception and recognition: a survey. *IEEE Trans. Affect. Comput.* 4(1): 15–33.
- Koelstra, S., M. Pantic and I. Patras. 2010. A dynamic texture-based approach to recognition of facial actions and their temporal models. *IEEE J. PAMI.* 32(11): 1940–1954.
- Kotsia, I., S. Zafeiriou, N. Nikolaidis and I. Pitas. 2008. Texture and shape information fusion for facial action unit recognition. *First International Conference on Advances in Computer-Human Interaction*.
- Krämer, N. C., I. A. Iurgel and G. Bente. 2005. *Emotion and motivation in embodied conversational agents*. Paper presented at the AISB'05 Convention, Symposium on Agents that Want and Like: Motivational and Emotional Roots of Cognition and Action, Hatfield, UK.
- Krämer, N. C., S. Kopp, C. Becker-Asano and N. Sommer. 2013. Smile and the world will smile with you — the effects of a virtual agent's smile on users' evaluation and behavior. *Int. J. Hum. Comput. Stud.* 71(3): 335–349. doi: 10.1016/j.ijhcs.2012.09.006
- Kukla, E. and P. Nowak. 2015. Facial emotion recognition based on cascade of neural networks. *Adv. Intel. Syst. Comput.* 67–78.
- Lahbiri, M., A. Fnaiech, M. Bouchouicha, M. Sayadi and P. Gorce. 2013. *Facial emotion recognition with the hidden Markov model. 2013 International Conference on Electrical Engineering and Software Applications*.
- Lazarus, R. S. 1991. *Emotion and adaption*. Oxford University Press, Oxford, UK.
- LeDoux, J. E., J. Iwata, P. Cicchetti and D. J. Reis. 1988. Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. *J. Neurosci.* 8(7): 2517–2529.
- Lee, J. and S. C. Marsella. 2010. Predicting speaker head nods and the effects of affective information. *IEEE Trans. Multimedia.* 12(6): 552–562. doi: 10.1109/tmm.2010.2051874
- Levenson, R. W. 1999. The intrapersonal functions of emotion. *Cogn. Emot.* 13(5): 481–504. doi: 10.1080/02699399379159
- Li, Y., S. Wang, Y. Zhao and Q. Ji. 2013. Simultaneous facial feature tracking and facial expression recognition. *IEEE Trans. Image Process.* 22(7): 2559–2573.
- Liebold, B. and P. Ohler. 2013. Multimodal emotion expressions of virtual agents. Mimic and vocal emotion expressions and their effects on emotion recognition. In: (T. Pun, C. Pelachaud and N. Sebe, eds) *2013 Humaine Association conference on Affective Computing and Intelligent Interaction, ACII 2013*. IEEE, Los Alamitos, CA, pp. 405–410.
- Lin, H.-C., S.-C. Mao, C.-L. Su and P.-W. Gean. 2010. Alterations of excitatory transmission in the lateral amygdala during expression and extinction of fear memory. *Int. J. Neuropsychopharmacol.* 13(3): 335–345.
- Lindquist, K. A., T. D. Wager, H. Kober, E. Bliss-Moreau and L. F. Barrett. 2012. The brain basis of emotion: a meta-analytic review. *Behav. Brain Sci.* 35(3): 121–143. doi: 10.1017/S0140525X11000446
- Lozano-Monazor, E., M. T. López, A. Fernández-Caballero and F. Vigo-Bustos. 2014. Facial expression recognition from webcam based on active shape models and support vector machines. *Lect. Notes Comput. Sc.* 147–154.
- Manstead, A. S. R. and A. H. Fischer. 2001. Social appraisal: the social world as object of and influence on appraisal processes. In: (K. R. Scherer, A. Schorr and T. Johnstone, eds) *Appraisal processes in emotion: theory, research, application*. Oxford University Press, New York, pp. 221–232.
- McDonald, A. J. 1998. Cortical pathways to the mammalian amygdala. *Prog. Neurobiol.* 55(3): 257–332.
- Mehrabian, A. and J. A. Russell. 1974. *An approach to environmental psychology*. MIT Press, Cambridge, MA.
- Mulligan, K. and K. R. Scherer. 2012. Toward a working definition of emotion. *Emot. Rev.* 4(4): 345–357. doi: 10.1177/1754073912445818

- Murphy, F. C., I. Nimmo-Smith and A. D. Lawrence. 2003. Functional neuroanatomy of emotions: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3(3): 207–233. doi: 10.3758/cabn.3.3.207
- Oatley, K. and J. M. Jenkins. 1996. *Understanding emotions*. Blackwell, Oxford, UK.
- Okon-Singer, H., T. Hendler, L. Pessoa and A. J. Shackman. 2015. The neurobiology of emotion-cognition interactions: fundamental questions and strategies for future research. *Front. Hum. Neurosci.* 9: 58. doi: 10.3389/fnhum.2015.00058
- Pape, H.-C. and D. Pare. 2010. Plastic synaptic networks of the amygdala for the acquisition, expression, and extinction of conditioned fear. *Physiol. Rev.* 90(2): 419–463.
- Perikos, I., E. Ziakopoulos and I. Hatzilygeroudis. 2014. Recognizing emotions from facial expressions using neural network. *IFIP AICT*. 236–245.
- Pessoa, L. 2008. On the relationship between emotion and cognition. *Nat. Rev. Neurosci.* 9(2): 148–158. doi: 10.1038/nrn2317
- Pessoa, L. 2012. Beyond brain regions: network perspective of cognition–emotion interactions. *Behav. Brain Sci.* 35(3): 158–159. doi: 10.1017/S0140525X11001567
- Pessoa, L. and R. Adolphs. 2010. Emotion processing and the amygdala: from a ‘low road’ to ‘many roads’ of evaluating biological significance. *Nat. Rev. Neurosci.* 11(11): 773–783. doi: 10.1038/nrn2920
- Phan, K. L., T. Wager, S. F. Taylor and I. Liberzon. 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*. 16(2): 331–348. doi: 10.1006/nimg.2002.1087
- Phillips, A. T., H. M. Wellman and E. S. Spelke. 2002. Infants’ ability to connect gaze and emotional expression to intentional action. *Cognition*. 85(1): 53–78. doi: 10.1016/S0010-0277(02)00073-2
- Piatkowska, E. and J. Martyna. 2012. Computer Recognition of Facial Expressions of Emotion. *Lect. Notes Comput. Sc.* 405–414.
- Picard, R. W. 1997. *Affective computing*. MIT Press, Cambridge, MA.
- Plutchik, R. 1980. *Emotion. A psychoevolutionary synthesis*. Harper & Row, New York.
- Premack, D. and G. Woodruff. 1978. Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1(04): 515–526. doi: 10.1017/S0140525X00076512
- Prevost, L., R. Belaroussi and M. Milgram. 2006. Multiple neural networks for facial feature localization in orientation-free face images. *Lect. Notes Comput. Sc.* 188–197.
- Qu, C., W.-P. Brinkman, Y. Ling, P. Wiggers and I. Heynderickx. 2014. Conversations with a virtual human: synthetic emotions and human responses. *Comput. Human Behav.* 34: 58–68. doi: 10.1016/j.chb.2014.01.033
- Russell, J. A. and L. F. Barrett. 1999. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *J. Pers. Soc. Psychol.* 76(5): 805–819. doi: 10.1037/0022-3514.76.5.805
- Sah, P., E. S. L. Faber, M. Lopez De Armentia and J. Power. 2003. The amygdaloid complex: anatomy and physiology. *Physiol. Rev.* 83(3): 803–834.
- Sander, D., D. Grandjean and K. R. Scherer. 2005. A systems approach to appraisal mechanisms in emotion. *Neural Netw.* 18(4): 317–352. doi: 10.1016/j.neunet.2005.03.001
- Sarnarawickrame, K. and S. Mindya. 2013. *Facial expression recognition using active shape models and support vector machines*. Paper presented at the Advances in ICT for Emerging Regions (ICTer).
- Scherer, K. R., T. Bänziger and E. B. Roesch (eds). (2010). *A blueprint for affective computing*. Oxford University Press, New York.
- Scherer, K. R. 1984. On the nature and function of emotion: a component process approach. In: (K. R. Scherer and P. Ekman, eds) *Approaches to emotion*. Erlbaum, Hillsdale, NJ, pp. 293–317.
- Scherer, K. R. 2000. Emotions as episodes of subsystem synchronization driven by nonlinear appraisal processes. In: (M. D. Lewis and I. Granic, eds) *Emotion, development, and self-organization: dynamic systems approaches to emotional development*. Cambridge University Press, New York, pp. 70–99.
- Scherer, K. R. 2001. Appraisal considered as a process of multi-level sequential checking. In: (K. R. Scherer, A. Schorr and T. Johnstone, eds) *Appraisal processes in emotion: theory, methods, research*. Oxford University Press, New York, NJ, pp. 92–120.
- Scherer, K. R. 2013. The nature and dynamics of relevance and valence appraisals: theoretical advances and recent evidence. *Emot. Rev.* 5(2): 150–162. doi: 10.1177/1754073912468166
- Scherer, K. R. and H. Ellgring. 2007a. Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion*, 7(1): 113–130. doi: 10.1037/1528-3542.7.1.113
- Scherer, K. R. and H. Ellgring. 2007c. Multimodal expression of emotion: affect programs or componential appraisal patterns? *Emotion*. 7(1): 158–171. doi: 10.1037/1528-3542.7.1.158
- Schwabe, L., C. J. Merz, B. Walter, D. Vaitl, O. T. Wolf and R. Stark. 2011. Emotional modulation of the attentional blink: the neural structures involved in capturing and holding attention. *Neuropsychologia*. 49(3): 416–425.
- Schönbrodt, F. D. and J. B. Asendorpf. 2011. The challenge of constructing psychologically believable agents. *J. Media Psychol.* 23(2): 100–107. doi: 10.1027/1864-1105/a000040
- Senechal, T., L. Prevost and S. M. Hanif. 2010. Neural network cascade for facial feature localization. *Lect. Notes Comput. Sc.* 141–148.
- Shenhav, A. and J. D. Greene. 2014. Integrative moral judgment: dissociating the roles of the amygdala and ventromedial prefrontal cortex. *J. Neurosci.* 34(13): 4741–4749.
- Shi, C. and M. Davis. 1999. Pain pathways involved in fear conditioning measured with fear-potentiated startle: lesion studies. *J. Neurosci.* 19(1): 420–430.
- Taylor, J. G. and N. F. Fragopanagos. 2005. The interaction of attention and emotion. *Neural Netw.* 18(4): 353–369. doi: 10.1016/j.neunet.2005.03.005
- Valstar, M. F. and M. Pantic. 2012. Fully automatic recognition of the temporal phases of facial actions. *IEEE J SMCB*. 42(1): 28–43.
- van Kleef, G. A., E. A. van Doorn, M. W. Heerdink and L. F. Koning. 2011. Emotion is for influence. *Eur. Rev. Soc. Psychol.* 22(1): 114–163. doi: 10.1080/10463283.2011.627192
- van Kleef, G. A. 2010. The emerging view of emotion as social information. *Soc. Personal. Psychol. Compass*. 4(5): 331–343. doi: 10.1111/j.1751-9004.2010.00262.x
- Vitay, J. and F. H. Hamker. 2011. A neuroscientific view on the role of emotions in behaving cognitive agents. *KI – Künstliche Intelligenz*, 25(3): 235–244. doi: 10.1007/s13218-011-0106-y

- Vuilleumier, P., M. P. Richardson, J. L. Armony, J. Driver and R. J. Dolan. 2004. Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nat. Neurosci.* 7(11): 1271–1278.
- Vytal, K. and S. Hamann. 2010. Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based meta-analysis. *J. Cogn. Neurosci.* 22(12): 2864–2885. doi: 10.1162/jocn.2009.21366
- Whitehill, J., G. Littlewort, I. Fasel, , Bartlett, M. and Movellan, J. 2009. Toward practical smile detection. *IEEE J PAMI.* 31(11): 2106–2111.
- Wimmer, H. 1983. Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition.* 13(1): 103–128. doi: 10.1016/0010-0277(83)90004-5
- Ye, W. and X. Fan. 2014. Bimodal emotion recognition from speech and text. *IJACSA.* 5(2).
- Zeng, Z., M. Pantic, G. I. Roisman and T. S. Huang. 2009. A survey of affect recognition methods: audio, visual, and spontaneous expressions. *IEEE Trans. Pattern. Anal. Mach. Intell.* 31(1): 39–58. doi: 10.1109/TPAMI.2008.52

Bionotes



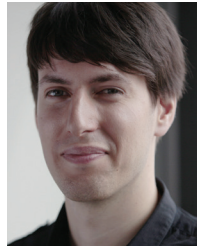
Benny Liebold
Chemnitz University of Technology,
Germany
benny.liebold@phil.tu-chemnitz.de

Benny Liebold is a researcher in the Institute for Media Research at Chemnitz University of Technology. His research focusses on the cognitive and emotional processing of virtual environments with an emphasis on the role of emotions in HCI, presence, game studies, and media effects in general, such as skill transfer and aggressive behavior.



René Richter
Chemnitz University of Technology,
Germany
rene.richter@cs.tu-chemnitz.de

Dipl. Inf. René Richter is a researcher in the department of computer science at Chemnitz University of Technology. His research field is computational neuroscience, in particular the influence of emotions on attention and how this interaction can be used in the context of HCI.



Michael Teichmann
Chemnitz University of Technology,
Germany
michael.teichmann@cs.tu-chemnitz.de

Dipl. Inf. Michael Teichmann works as a researcher in the department of computer science at Chemnitz University of Technology. His research focusses on neuroscientifically grounded computational models of the human visual cortex, in particular self-organization via neural plasticity in recurrent models of the ventral visual stream.



Fred H. Hamker
Chemnitz University of Technology,
Germany
fred.hamker@cs.tu-chemnitz.de

Fred H. Hamker is a professor of artificial intelligence in the department of computer science at Chemnitz University of Technology since 2009. He received his diploma in electrical engineering from the University of Paderborn in 1994 and his Ph. D. in computer science at Ilmenau University of Technology in 1999. He was a postdoc at the Goethe University Frankfurt and the California Institute of Technology (Pasadena, USA). In 2008 he received his *venia legendi* from the Department of Psychology at the University of Münster.



Peter Ohler
Chemnitz University of Technology,
Germany
peter.ohler@phil.tu-chemnitz.de

Peter Ohler is a professor of media psychology at the Institute for Media Research at Chemnitz University of Technology since 2002. He studied psychology at Saarland University and received his Ph. D. at TU Berlin in 1991. He had postdoc positions at TU Berlin and the University of Passau. In 2000 he received his *venia legendi* in psychology from TU Berlin. His research interests include the psychology of film, evolutionary psychology, cognitive science, and the psychology of play.