Universitat de Lleida
Escola Politècnica Superior

Màster en Interacció Persona-Ordinador

Master Thesis

Towards a lightweight adaptive multicomponential affective recognition system

Author
Andrés Navarro García

Director
Rosa María Gil Iranzo

September 2014

Abstract

As part of the Affective Computing research field, the development of automatic affective recognition systems can enhance human-computer interactions by allowing the creation of interfaces that react to the user's emotional state.

To that end, this Master Thesis brings affect recognition to nowadays most used human computer interface, mobile devices, by developing a facial expression recognition system able to perform detection under the difficult conditions of viewing angle and illumination that entails the interaction with a mobile device.

Moreover, this Master Thesis proposes to combine emotional features detected from expression with contextual information of the current situation, to infer a complex and extensive emotional state of the user. Thus, a cognitive computational model of emotion is defined that provides a multicomponential affective state of the user through the integration of the detected emotional features into appraisal processes. In order to account for individual differences in the emotional experience, these processes can be adapted to the culture and personality of the user.

Keywords: affective computing, affect recognition, facial expression recognition, computational model of emotion, appraisal

Contents

List of Figures

1. Introduction

1.1 Motivation and Objectives

Human Computer Interaction has rapidly evolved over the last decades. Starting from command line interfaces on the first computers, graphical user interfaces took over in the early 80s as a first attempt on making interfaces easier to use, leading to the currently existing wide number of interaction techniques, models and theories that aim to create a more usable user experience that focuses on the user's needs.

Any human interaction has a social and cultural base and the experience that results from the interaction reflects on the user emotional state. Hence, user's emotions should be explicitly considered in the design of any human computer interface. From an aesthetic perspective, the visual design of the interface can influence this emotional state. However, to include users' emotions as a fundamental part of the interaction, the interface should react to their emotional state changes, developing a system that adapts to the user's affective state.

As part of the Affective Computing field of research, which comprises among others the study of systems that recognize and react to the user's emotions, there is a need to provide methods and tools to develop such emotional adaptive systems.

Therefore, this work aims to create the required tools to develop a system able to recognize and react to the user's emotions that can be integrated seamlessly and unobtrusively in any wide used application. To that end, the objectives are the creation of an automatic recognition system with the following requirements:

- Recognize user's emotional state in real-time on mobile devices.
- Provide rich and extensive estimation of the user's emotional state.

In order to accomplish these objectives, estimation of the emotional state by the recognition of the facial expression is the followed approach because of not requiring additional devices, integrating seamlessly with any application and not distorting user experience.

The relation between facial expressions and emotions has been extensively studied. A long discussion has been present among emotional researches on whether facial expressions are the result of the underlying emotional processes or are evolved social signals. However, there is another approach that, based on more complex emotional models, considers facial expressions as serving both functions at the same time.

Most of the emotional expressions take place during a social interaction influenced by social emotional components, such as social regulation or social coordination. Hence, prototypical spontaneous expressions of an internal emotional state are less common than expressions where management of self-presentation and expression control by sociocultural rules processes are present. In order to discern between the different meanings that can be interpreted on a facial expression (i.e. discerning between cognitive difficulty in understanding, communicative signaling and emotional indication), affective recognition shouldn't be performed from isolated facial expressions. Therefore, integrating knowledge about social rules and context, as we humans do, will allow to estimate richer and more precise information about the emotions.

Cognitive theories of emotion state that emotions are elicited by the evaluation of situations and events. Among cognitive theories, appraisal theory has been largely developed, because its principles can be translated into a computational model. Several proposals can be found in the literature that suggest relationships between facial actions and appraisal dimensions (see (Kaiser & Wehrle, 2001) for more details). Considering facial expressions as indicators of certain appraisal processes, the representation of the facial expression can be integrated on a cognitive model that combines this expression features with contextual information into a multicomponential emotional state.

Social and cultural background of the interaction determines the emotional response to a situation or event and cannot be ruled out from the emotional recognition. Therefore, cultural and individual characteristics of the users should be modeled and integrated on the appraisal processes to include their influence to emotions, providing a cultural and personal adaptive affective recognition system.

1.2 Contributions

This Master Thesis presents an automatic facial expression recognition system that can perform the detection in the conditions present during an interaction with a mobile device, without requiring additional devices or distorting user experience.

Moreover, a cognitive computational model of emotion is proposed that can integrate emotional features resulted from the facial expression recognition and contextual information to estimate a multicomponential user's affective state that takes into consideration the individual characteristics of the user, such as culture and personality.

1.3 Master Thesis outline

The content of this document is divided into five chapters. This first chapter introduces the motivations and main contributions of the work. In Chapter 2 an overview of the state-of-the-art on affective recognition and computational modeling of emotion is presented. Chapter 3 details the developed facial expression recognition system, as well as its evaluation results and an enhancing proposal. Chapter 4 describes the proposed cognitive model and its individual adaptation. Finally, in Chapter 5 the conclusions are presented. Additionally, the list of references is included at the end.

2. Related Work

In this chapter, an overview of the state-of-the-art on affective recognition and computational modeling of emotion, as well as on multiculturalism and individual adaptation, is presented.

2.1 Affective computing

Affective computing aims to develop systems and devices that can recognize, interpret, process, and simulate human affects.

As R. Picard states in (Picard, 1995), affective computing allows the development of systems that are aware of users' emotions, bringing up plenty of possibilities to improve interaction with computers as we currently know.

Moreover, based on studies that prove emotions as a basic element on humans decision taking and problem solving abilities, R. Picard suggests that in order to develop intelligent systems not only capable to follow certain rules, but capable to make use of previous experiences and possess a sense of creativity, computers need to be emotional.

Having those emotional capabilities, computers would be able to react to social situations in an accepted manner, following a behavior that shows emotional intelligence (Goleman, 2006).

This emotional intelligence should take into account the user particularities, on an adaptation process that evaluates aspects like personality, social context and culture.

2.2 Theories of emotion

During history, a large number of psychological and philosophical emotion theories have been proposed (Strongman, 1978) without reaching a stage where a unified theory of the mind could be defined.

The fragmentation of the models, the ambiguities and the contrasting definitions, which are characteristics of a science as complex as psychology, are not only limited to emotional theories, but are also present in the definition of the emotion. Hence, currently there is not a generally accepted definition of emotion, not only as consequence of the historical diversity of linguistic terms used, but as the fundamental differences on the meaning of the emotion concept itself.

However, research on emotions can be performed without a precise definition of emotion, but with a working definition of emotion (R. Reisenzein, 2007) that defines the research domain by describing for example some examples of emotion. Hence, despite the differences on the nature or correct definition of emotion, emotions can be defined as those mental states of humans that are called by words such as “sadness”, “fear”, “joy”, and “anger”.

Although those working definitions of emotion can differ as much as their underlying theories of emotion, many of them share some principles. One shared principle is the concept of basic (or primary) and non-basic (or secondary) emotions, with different conceptions of basic emotions as biological primitive components of other emotions (Ortony & Turner, 1990) or as basic colors on a palette (K. R. Scherer, 1984b), but which can be described as the set of ontogenetically earlier types of emotions corresponding to the six basic emotions of Ekman et al. (Ekman, 1999), shown in Figure 1.

Figure 1: Ekman's six basic emotions[1]

The debate about the mixture of primary emotion compounds (Plutchik, 1980) leads to the concept of emotional dimensions. Dimensional theories postulate a core affect state as a single point along a number of continuous orthogonal dimensions. Hence, this primary emotional reaction, located in the dimensional space, is later differentiated and enriched by cognitive processing (Russell, 2003).

Most notably dimensional theories are the circumplex theory (Russell, 1980) shown in Figure 2, which places emotional state based on the two dimensions pleasure and arousal, and the PAD dimensional theory (Mehrabian & Russell, 1974) shown in Figure 3, which extends the circumplex axes with dominance dimension:

- Pleasure (or valence) dimension refers to the hedonic quality of the emotion, if it is either positive or negative.
- Arousal (or intensity, activation) dimension refers to the level of physiological arousal or neurological activation.

1 http://psych-your-mind.blogspot.com.es/2011/06/channels-of-emotion-not-just-in-face.html

- Dominance (or attention, control, power) dimension refers to the degree of control of the situation.

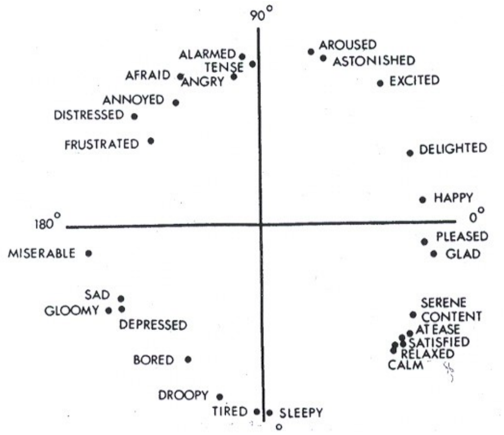

Figure 2: The circumplex model of core affect (Russell, 1980)

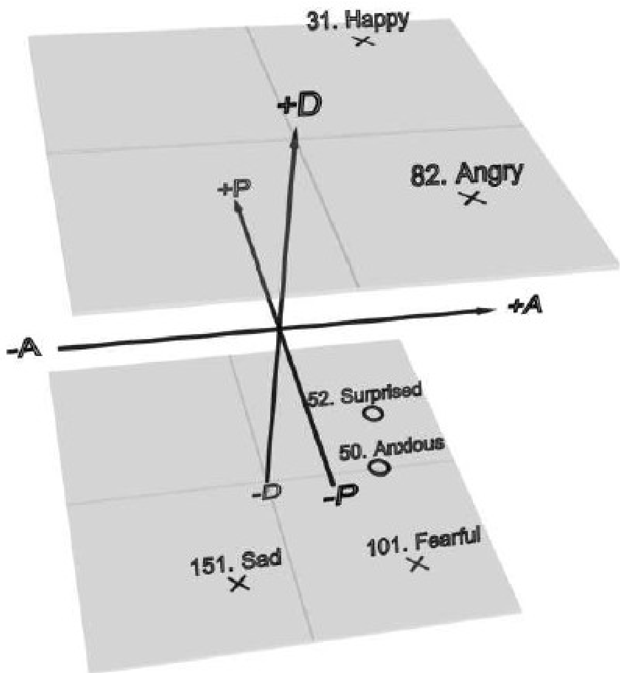

Figure 3: PAD dimensional space (Becker-Asano, 2008)

Dimensional theories present several drawbacks such as the lack of a functional perspective, the absence of a theoretical prediction of emotion differences and the lax of an explanatory mechanism (Grandjean, Sander, & Scherer, 2008).

Another shared principle is that emotions are reactions to the perception, imagination or thoughts of events or states of concerns. However, there is an opposite conception between cognitive and noncognitive theories about how is the process of emotion elicitation underlying those reactions. Cognitive theories propose that emotions require certain higher-order mental representations or even assume that emotions are a form of cognition. Noncognitive ones, propose that certain types of emotions are elicited via a more direct route based on sensory stimuli.

Appraisal theories of emotion consider that why a certain event results in certain emotional response can be explained by a subjective cognitive evaluation of events in relation to the agent's goals (I. Roseman & Smith, 2001), called appraisal. Hence, the basis of the theory is that the agent can evaluate a concept or situation with respect to the appraiser's beliefs, desires and intentions.

Main concepts on appraisal theories are the perception of the environment by the agent, the appraisal processes that evaluate that environment in terms of values on a set of measures called appraisal dimensions and the mediation that relates those appraisal values to the agent's emotional state (J Broekens, DeGroot, & Kosters, 2008).

One of those concepts in which appraisal theories differ from one another is the number and definition of those appraisal variables, and what is more important, the derivation from the variables to emotions. Scherer (K. R. Scherer & Schorr, 2001) categorizes appraisal approaches based on the nature of their underlying appraisal dimensions:

- Focus on the evaluation of the significance of events by a set of criteria.
- Focus on the causal attribution involved in emotion-antecedent appraisal.
- Focus on the evaluation of event relationship to agent's goals by specific patterns
- Focus on the semantics of emotion natural language.

For example, Roseman's theory of appraisal (I. J. Roseman, 1984) defines five components that influence emotion, shown in Figure 4: the *motivational state*, that evaluates the desirability of an event in terms of if it contains some important aspect, the *situational state*, that evaluates the desirability of the event, the *probability*, that evaluates how certain is the event, the *power*, that evaluates the perception of strength, and the *agency*, that evaluates whether the event was caused by the agent or by another.

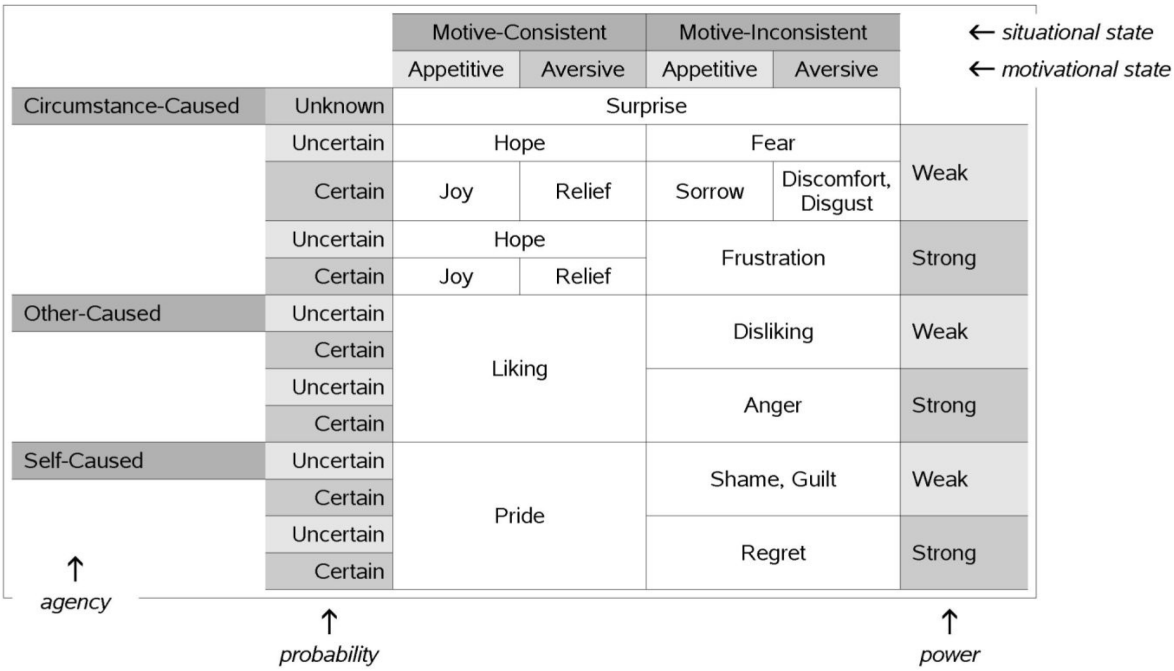

Figure 4 Appraisal components in Roseman's appraisal theory (Fieser & Dowden, 2014)

One of the most prominent theories that attempts to describe the underlying appraisal in humans is the “Component Process Model” from Scherer (K. R. Scherer, 1984a). Defining emotion as “an episode of inter- related, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism” (K. R. Scherer & Schorr, 2001), which are summarized in Table 1, and determining appraisal as a fixed sequence of stimulus evaluation checks (SECs):

1. Relevance check: evaluation of the novelty of the stimulus and its relevance to the agent’s goals.
2. Implication check: evaluation of the implications and consequences of the stimulus.
3. Coping potential check: evaluation of the ability to cope with the situation
4. Normative significance check: evaluation of the overall compatibility of the stimulus.

Emotion function	**Emotion component**	**Organismic subsystem**
Evaluation of objects and events	Cognitive	Information processing
System regulation	Peripheral efference	Support
Preparation and direction of action	Motivational component	Executive
Communication of reaction and behavioral intention	Motor expression	Action

Monitoring of internal state and organism-environment interaction	Subjective feeling	Monitor

Table 1: Relationships between the functions and components of emotion (K. R. Scherer & Schorr, 2001)

Another appraisal theory that has been widely used is the one proposed by Ortony, Clore & Collins (Ortony, Clore, & Collins, 1988), often referred as the OCC theory. Their theory categorizes emotions based on appraisal of valence and arousal by evaluating the desirability of an event, the praiseworthiness of an action and the like/dislike of an entity, as shown in Figure 5. The OCC-model presents a comprehensible and precise description of the emotion elicitation and there is a tendency to use it on the implementation of computational models because of its computational tractability.

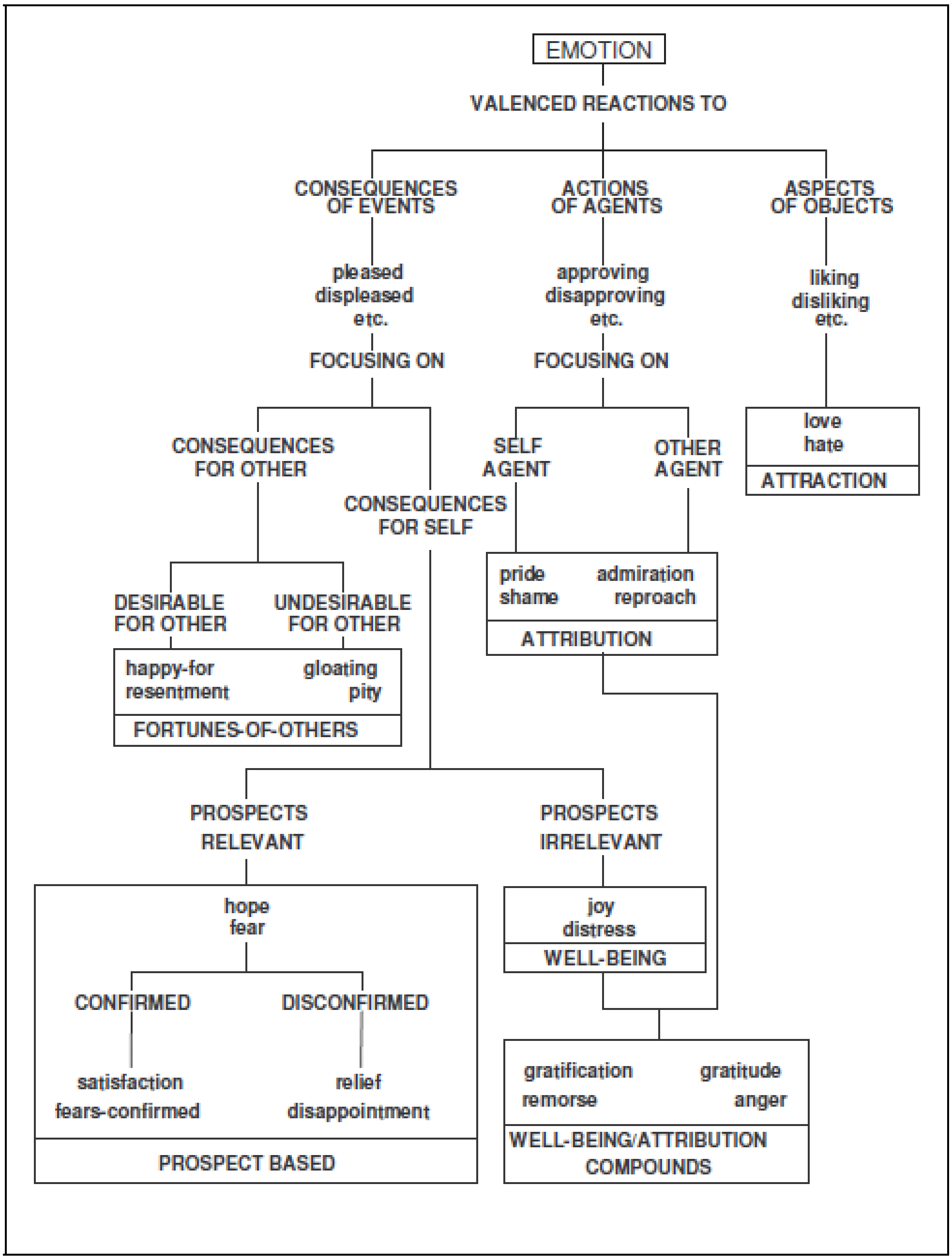

Figure 5: The OCC model of emotions (Ortony et al., 1988)

Other existing theories are based on anatomic approaches, which attempt to reconstruct neural links and processes by low-level sub-symbolic processes, on rational approaches, which attempt to integrate into a model of intelligence an abstraction of the emotional adaptive function, or communicative approaches, which emphasize the social-communicative function of emotion processes.

Finally, the effect of emotion on cognition can be determined as an affect-consequent system that can be can be categorized among behavior-consequent models, which map an agent's emotions state to certain action or expression, and cognitive-consequent models that determine how affect alters the cognitive processes (Lin, Spraragen, & Zyda, 2012). Examples of modeled effects are the modifications of attention or focus, the selection of goals and action choices, or emotional biases on decision of planning processes. A detailed comparison of effects modeled on different emotional models is presented in the work of Lin et al. (Lin et al., 2012).

2.3 Affective recognition

In order to develop a system that respond to user's affective state during the interaction with an interface, enabling the adaptation of it to enhance the quality of the interaction, it is required to estimate the user's affective state. Affective recognition has been in the last decades one of the most prominent research fields in Affective Computing.

Humans express emotions in multiple modalities:

- Our facial expression is greatly related to our emotional state.
- Despite the content of our speech, our voice contains several paralinguistic features related to our emotional state.
- Our body language and posture can also express our affective state.
- Several physiological signals have been studied to detect emotions, such as the galvanic skin conductance, heart beat,
- The text that we write or the content of our speech contain emotional attributes.

Therefore, affective recognition can be performed based on one of these modalities, or the information from more than one can be combined in a process called multimodal fusion, resulting in a more robust and accurate affective detection (Gilroy et al., 2009)

2.3.1 Facial expression recognition

One of the most used approaches in affective recognition is based on the concept that for each basic emotion there is a distinctive facial expression associated. Although an emotional theory that limits emotions to expressions is unable to describe several emotional components and to explain many emotional processes, there is evidence of universal facial expressions of emotion which has lead to the development of several affective recognition systems based on facial expression.

Automatic facial expression recognition presents several challenges. For example, the system must deal with variations in the scale and orientation of the face and with partially occluded faces by hair or glasses. However, affective recognition based on facial expressions present several advantages, such as being nonintrusive and not requiring complex or expensive hardware. In fact, most current used interfaces (e.g., smartphones, tablets, laptops and webcam equipped desktop computers) already contain a camera capable of recording the user.

Humans are capable or recognizing facial expressions even in hard situations, such as bad lighting conditions, far distance or partially occluded faces. Hence. Human visual system can be taken as a reference of what features should have a facial expression recognition system. Pantic and Rothkrantz (Pantic & Rothkrantz, 2000) define a set of 20 characteristics of an ideal system for facial expression analysis: "*1) Automatic facial image acquisition, 2) Subjects of any age, ethnicity and outlook, 3) Deals with variation in lighting, 4) Deals with partially occluded faces, 5) No special markers/make-up required, 6) Deals with rigid head motions, 7) Automatic face detection, 8) Automatic facial expression data extraction, 9) Deals with inaccurate facial expression data, 10) Automatic facial expression classification, 11) Distinguishes all possible expressions, 12) Deals with unilateral facial changes, 13) Obeys anatomical rules, 14) Distinguishes all 44 facial actions, 15) Quantifies facial actions codes, 16) # interpretation categories unlimited, 17) Features adaptive learning facility, 18) Assigns quantified interpretation labels, 19) Assigns multiple interpretation labels, 20) Features real-time processing*".

In order to perform the complex task of recognizing the facial expression, different steps must be followed. Fasel and Luettin (Fasel & Luettin, 2003) define a generic facial expression analysis framework composed of several stages, shown in Figure 6. The first stage, *face acquisition*, determines the location of the face in the image. Depending on the followed method, face can be normalized to address pose and illumination changes and can be segmented from the background. Second stage, *facial feature extraction*, extracts a representation of the facial features based on its deformation or motion. Final stage, *facial expression classification*, classifies extracted features into a coded facial expression.

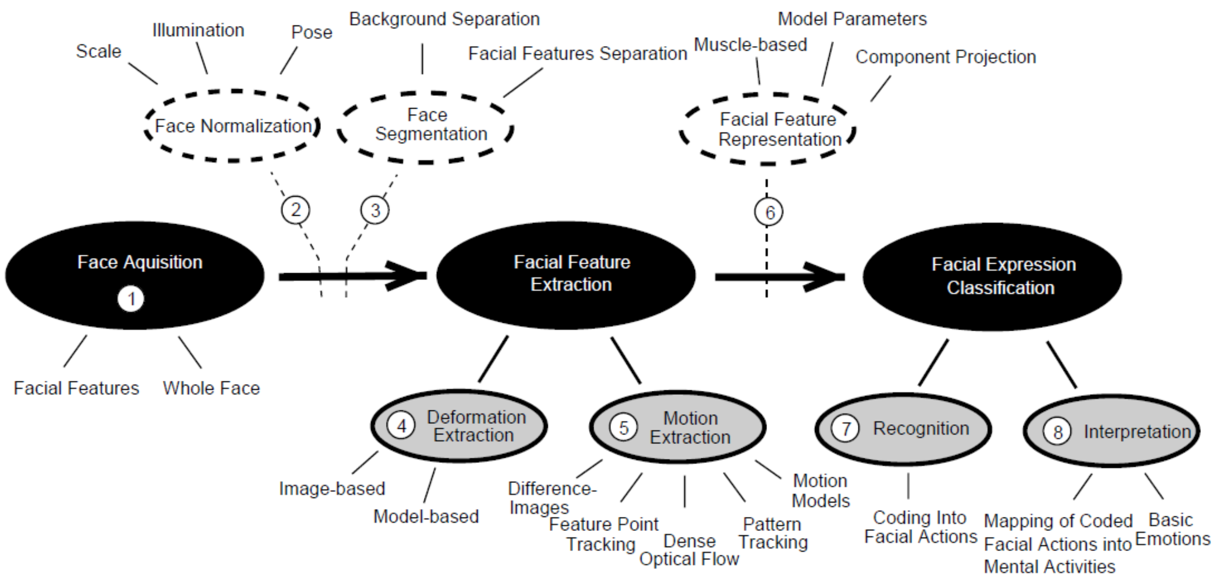

Figure 6 Generic facial expression analysis framework (Fasel & Luettin, 2003)

Multiple methods can be found in the literature to perform facial expression recognition. These methods can be classified in terms of the approaches followed at the different analysis steps. For example, feature extraction can be divided between local and holistic approaches, where the former focuses on areas that are prone to change and the latter processes the face as a whole unit, or between motion-based approaches that focus on facial changes and deformation approaches that focus on changes based on a defined model.

Many of these approaches rely on a face model. Compared to image-based methods, which are typically fast and simple, model-based methods may require to be manually constructed but have a more robust and precise performance. Face model can characterize the resulted appearance or the underlying muscle activities. Among appearance models, 2D models have a simpler mapping procedure since the recovery of the volumetric geometry is not needed, but are very limited on the camera viewing angle, which was a motivation to research on 3D methods that enable a broader range of camera viewing angles (Blanz, 2006).

Once the facial features have been extracted, they must be classified into a facial expression. The classification method can only consider the spatial configuration of the features or it can consider spatio-temporal information, allowing to model facial expression dynamics. For example, neural networks are often used in the first case, while Hidden Markov Models are commonly used in the second one.

2.4 Computational modeling of emotions

In order to develop an emotional adaptive system, a computational model of emotion that describes human emotion mechanisms is needed. Such model provides a complete and detailed definition of such processes which can be used to simulate and predict the emotion mechanisms as wells as to provide a detailed instantiation of an emotion theory or set of hypotheses (Joost Broekens, 2010).

2.4.1 Interdisciplinary exchange

The definition of a computational model of emotion is an interdisciplinary study between psychology and computer science. Although the former field is more interested in achieving a better theoretical understanding of human emotions, while the latter is more interested in models that can be applied on the development of artificial emotional mechanisms, there is a close relationship between both goals that has led to their cooperation. Therefore, the computational modeling of emotion provides valuable tools for different disciplines, resulting in great impact on different research fields.

In psychological research on emotion, computational models provide a method to evaluate human emotion theories and enforce a concrete and explicit formulation of the theories, which are typically described at an abstract level and without much detail.

In Artificial Intelligence, research in cognitive science and neuroscience has revealed that emotions play an important role in humans' cognitive processes, motivating the exploration of replicating emotional processes on computers to develop more intelligent and adaptive systems.

In Human Computer Interaction, research has revealed that users apply social rules to their interaction with computers, allowing the interaction to be considered as an interaction between humans (Reeves & Nass, 1996) and to consider relevant to this matter studies in social psychology communication (Moon & Nass, 1996). Hence, computational modeling of emotion has allowed the development of several applications that attempt to improve interaction between computer systems and human users by reacting to the perceived emotional state of the user.

2.4.2 History and evolution

Similar to the diversity of emotion theories, a wide variety of computational models have been developed. Since a computational model is a detailed instantiation of a theory, most of the computational models are based on the theories previously presented. However, many of those emotion theories lack of preciseness and completeness, which makes them very difficult to translate into a computer implementation. In Figure 7 a diagram of the evolution and relationships of computational models is shown.

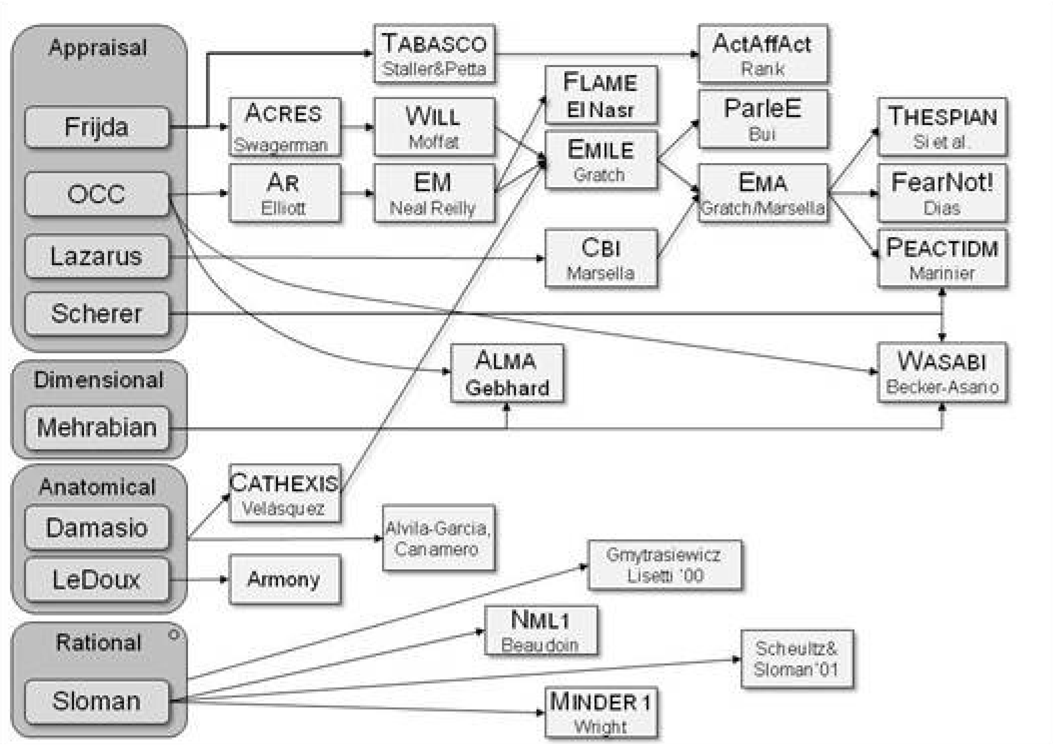

Figure 7: A history of computational models of emotion (Stacy Marsella, Gratch, & Petta, 2010)

Among cognitive computational models, appraisal theory tends to dominate them due to its emphasis on emotions as computable artifacts. In appraisal-based computational models, appraisal is the main process, which can be implemented by decision-theoretic plans, reactive plans, Markov-decision processes or detailed cognitive models (Stacy Marsella et al., 2010).

On the other hand, dimensional computational models focus on the structural and temporal dynamics of core affect which is modeled by a continuous time-varying process that places it at the dimensional space.

Furthermore, several approaches are based on both theories. For example, although not many dimensional approaches explore the relationship between core affect and cognition, some models incorporate appraisal as the trigger of core affect modification, instead of considering appraisal as an explanation of core affect (Stacy Marsella et al., 2010).

2.4.3 Emotional architectures

From the multiple emotion computational architectures that can be found in the literature, which were partially shown previously, the following ones are selected based on their prominence and their different and significant approaches.

FLAME: Fuzzy Logic Adaptive Model of Emotions

El-Nasr et. al. (El-Nasr, Yen, & Ioerger, 2000) present FLAME, an appraisal model based on OCC which uses fuzzy rules to evaluate the desirability of an event that affects agent's goals, following the emotional process shown in Figure 8. Dynamics of the emotional state are evaluated on a decay step and fed back to the system for the next iteration, influencing next states. Furthermore, mood is evaluated as the average of all emotions and used to define which emotion inhibits another when conflicting ones are present at the same time.

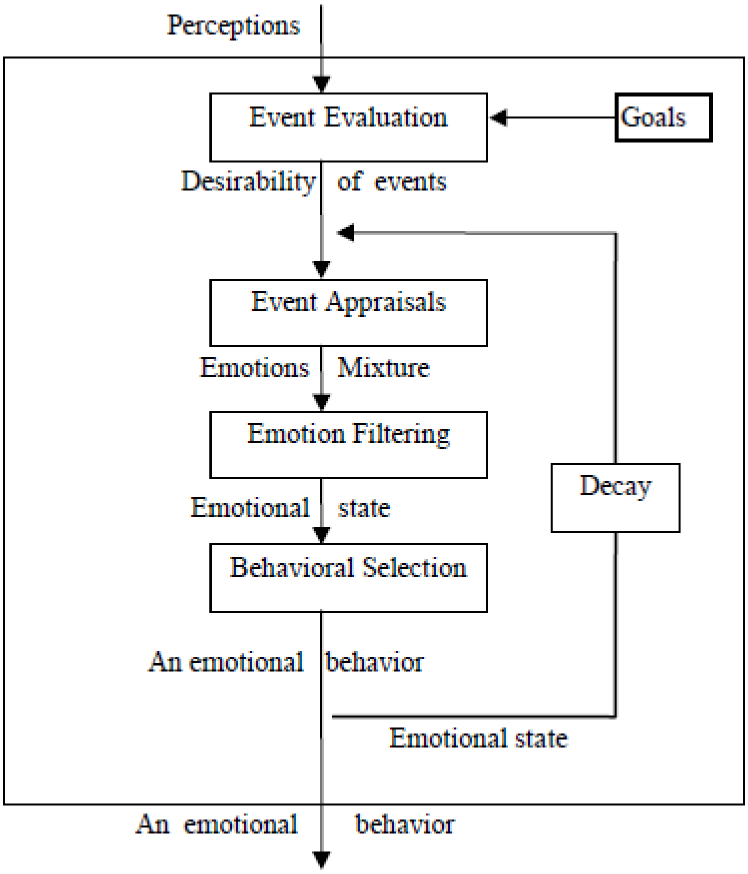

Figure 8: "Emotional process component" of FLAME (El-Nasr et al., 2000)

EMA

Marsella & Gratch (SC Marsella & Gratch, 2009) present EMA, a model based on Soar cognitive architecture that emphasizes dynamics of emotions and coping, motivated by Lazarus appraisal theory (Lazarus, 1994). Following their work on the Émile framework (Gratch & Marsella, 2004), a BDI-based approach that combines appraisal and coping processes, EMA uses appraisal frames associated with plan steps to derive emotions, as shown in Figure 9.

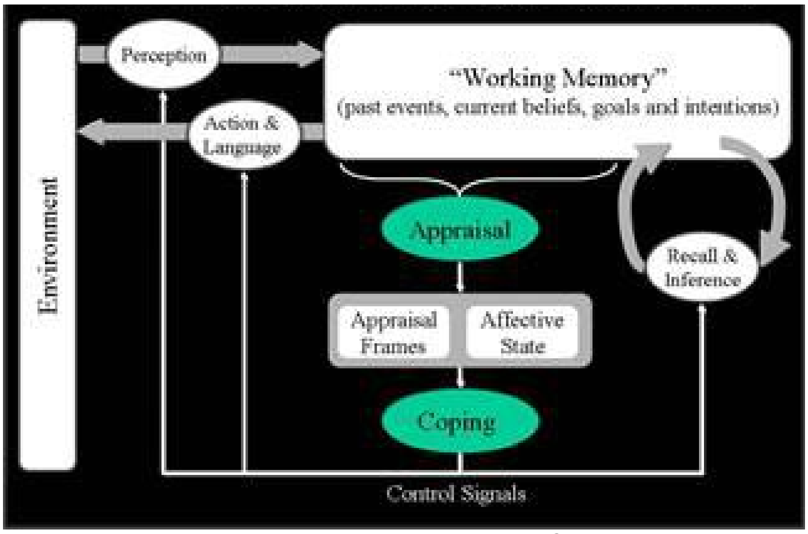

Figure 9: EMA process [2]

WASABI

Becker-Asano (Becker-Asano, 2008) present WASABI, a model that integrates emotion and cognition by combining a PAD-based model of core affect with an appraisal process based on Scherer's sequential-checking theory. Moreover, WASABI differentiates secondary emotions as a result of a conscious appraisal grounded on primary emotions, as shown in Figure 10.

2 http://people.ict.usc.edu/~gratch/

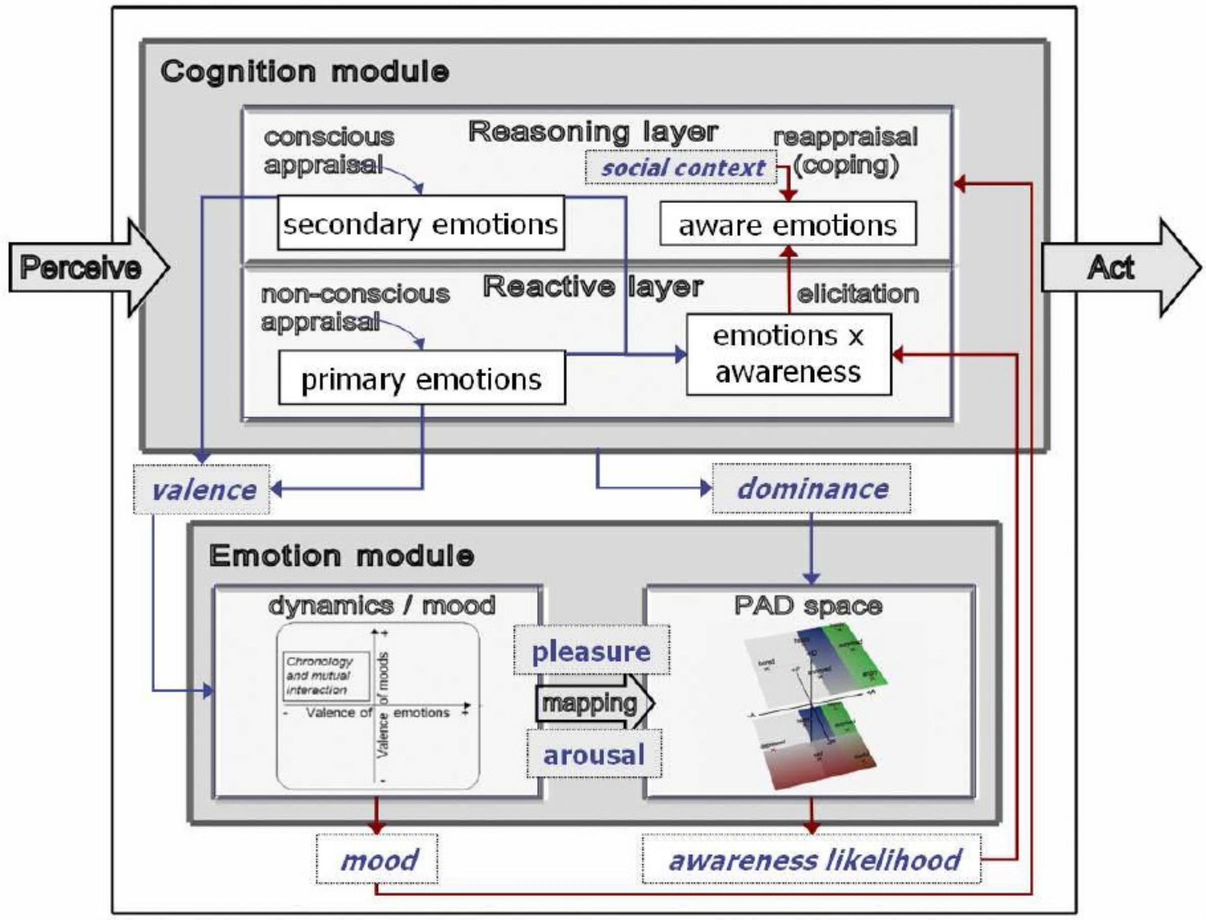

Figure 10: Interaction of cognition and emotion in WASABI (Becker-Asano, 2008)

2.4.4 Unifying approaches

As previously presented, there are numerous computational models of emotions, which are rarely cumulative and cannot easily be compared. However, several approaches have been presented that attempt to provide a common lexicon and a unified architecture.

Related to the definition of a unified architecture, Reisenzein et al. (Rainer Reisenzein et al., 2013) propose an intermediate step, the translation of psychological emotion theories into a precise but implementation-independent formalization. Such formalization can provide the benefit of resolving ambiguities and inconsistencies of the theory, resulting in a more precise, complete and consistent version. Such formalization can be detailed using the set theory language (J Broekens et al., 2008), shown in Figure 11, which provides a formal reconstruction of emotional theories, taking advantage of its expressiveness and the mathematical tools available for theory reconstruction.

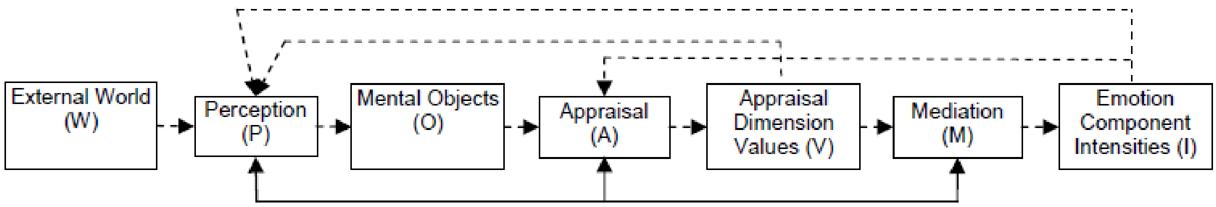

Figure 11: Set-theoretical formalization of appraisal theories (J Broekens et al., 2008)

Another approach towards a unified architecture, is the proposal from Marsella et al. (Stacy Marsella et al., 2010) of a compositional view of model building where emotional models can be assembled from sharable and changeable individual components or “sub-models”.

In Figure 12 the components are presented with their relationships derived from appraisal theory and following an information cycle, as defined by (Stacy Marsella et al., 2010):

- The **person-environment relationship** is a representation of the agent’s relationship with its environment that allows the agent to: “*derive the relationship between external events and the beliefs, desires and intentions of the agent or other significant entities in the environment*”.
- The **appraisal-derivation model** “*transforms some representation of the person-environment relationship into a set of appraisal variables*”.
- The **appraisal variables** “*correspond to the set of specific judgments that the agent can use to produce different emotional responses*”.
- The **affect-derivation** model “*maps between appraisal variables and an affective state, and specifies how an individual will react emotionally once a pattern of appraisals has been determined*”.
- The **affect-intensity model** “*specifies the strength of the emotional response resulting from a specific appraisal*”.
- **Emotion/Affect** “*is a representation of the agent’s current emotional state*”, based on a discrete emotion label or a point on a continuous dimensional space.
- The **affect-consequent model** “*maps affect into some behavioral of cognitive change*”.

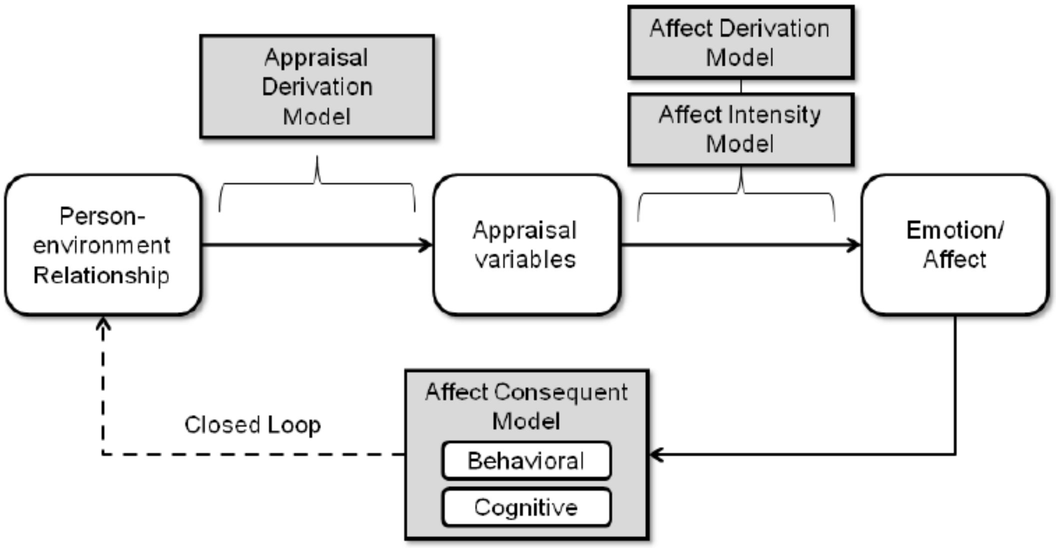

Figure 12: A component model view of computational appraisal models (Stacy Marsella et al., 2010)

2.5 Multiculturalism and personal adaptation

When designing an interface, one of the concepts that can improve its usability is the customization for the targeted user or group of users. One of the aspects that is mostly considered for this customization is the user or group culture. Several definitions can be found for the term culture. On human-computer interaction, Hofstede studies of cultural differences are commonly used, who stated that “culture is, to human collectivity, what personality is to the individual” and presented a five dimensions model of culture: short-term vs long-term orientation, femininity vs masculinity, power-distance, collectivism vs individualism and uncertainty avoidance.

However, identifying user culture brings new challenges. Users posses different cultures that makes them process the information in a different way (Nisbett, 2004). In order to design a usable interface, those cultural differences should be taken into account (Callahan, 2006). The process of adapting the interfaces to the different cultures is known as localization. Usually, localization is based on selecting certain countries of interest, which can be defined as internationalization.

Localization improves an interface usability in several aspects. For example, it enhances user performance and efficiency, due to the reduction of users cognitive workload and the avoidance of stress (Heimgärtner, Holzinger, & Adams, 2008).

Most of the strategies at interface localization make use of designing guidelines developed from theoretical studies or the evaluation with a reduced amount of users from a small number of countries. Moreover, these guidelines are mainly focused on visual aspects and do not study user work flow to fulfill a certain task.

On nowadays globalized world, it is very complex to define groups of users with similar characteristics based on nations, regions, teams, etc. Moreover, for certain characteristics each person can be considered to have a unique set. This leads us to another strategy to improve an interface usability and perception, the adaptation to the user's particularities.

When facing the challenge of adapting an interface to a particular user, most of the approaches are based on adjusting certain parameters for several groups of users. Classifying a user in one of the previously defined groups can result into an incorrect adaptation for one or several of the adaptation parameters. Hence, an approach that mitigates this grouping effect is to create adaptive interfaces which adapt dynamically and progressively to the user (Reinecke, 2010).

As Heimgärtner maintains in (Heimgärtner et al., 2008), system adaptation to individual expectations based on observation and analysis of user behavior brings a better alignment between system model and user's model, improving usability.

Individual adaptation can be performed in different steps. In (Benyon, 1993), focusing on cognitive and personality user characteristics, D. Benyon presents a process to develop adaptive interfaces where firstly user differences are assayed, secondly isolated from context and finally accommodated to the interface.

Several user characteristics can be considered when adapting an interface. One of them is the user personality, which refers to distinctive yet stable individual differences in behavior and cognition (Francis, Mehta, & Ram, 2009).

Adaptation to user personality can be performed following different approaches. For example, i n (Moon & Nass, 1996), from several studies about human preferences on personality similarity when addressing another human being, Moon and Nass support a similar effect when interacting with a computer. Hence, they defend that the social psychological rule called “law of attraction”, which states that people prefer to interact with other who share their personality type, can be applied to the interaction with computers. Furthermore, they prove the applicability of the gain-loss theory of attraction which states that people are more attracted to individuals who change from dissimilar personality behavior to a similar one, rather than a change in the opposite direction.

Personality also influences user attitude towards adaptation. As Goren-Bar et al. studied in (Goren-Bar, Graziola, Pianesi, & Zancanaro, 2006), personality traits related to the notion of control have an effect on the acceptance of the interface adaptivity, which depends on the type of adaptation, which they divided into four dimensions: location awareness, follow-ups, content adaptation to user interests and content adaptation to history of interaction. Moreover, they suggest that an adaptive interface should dynamically increase or decrease user control over the relevant dimensions of adaptivity based on the knowledge of the user's personality.

3. Facial expression recognition

The followed approach in recognizing the affective state of the user is based on facial expression.s As presented in the introduction chapter, the main objectives of the recognition system are having a good performance in mobile devices and providing a generic solution that does not require a specific training for each user.

3.1 Automatic system design

As described in Chapter 2.3.1 facial expression recognition can be analyzed as a multiple stages process where firstly the face is detected, secondly some facial features are extracted and finally a facial expression is classified. Multiple methods and approaches can be found in the literature for this different steps, but some of them are more appropriate than others for the purpose and requirements of the target system. Being targeted a mobile device, a low computational time and supporting a wide range of viewing angles and illumination conditions are required.

The detection of generic faces has greatly evolved over the last decade, motivated by its multiple applications. In terms of the training process required to function, detection systems can be classified between learning-free and learning-based approaches. Learning-free approaches rely on position, symmetry and edge shape of facial properties, without requiring a training process, but most of them do not provide a complete set of face features. On the other hand, learning-based approaches require a tedious learning phase where image conditions will affect detection performance.

The followed approach is based on Unzueta et. al. (Unzueta, Pimenta, Goenetxea, Santos, & Dornaika, 2014) proposal of a framework that presents both learning-free and learning-based advantages, not requiring a learning process and providing a rich set of facial features, with an accurate and efficient localization. This approach is based on a 3D appearance-based model consists in two steps: "*(1) the detection of facial features on the image and (2) the adjustment of the deformable 3D face model such that the projection of its vertices onto the 2D plane of the image matches the locations of the detected facial features.*"

The used 3D model is based on the Candide-3 model (Ahlberg, 2001), having several vertex simplifications and changes in the animation units to allow a better tracking through an OAM approach and a straightforward translation into facial action units, shown in Figure 13.

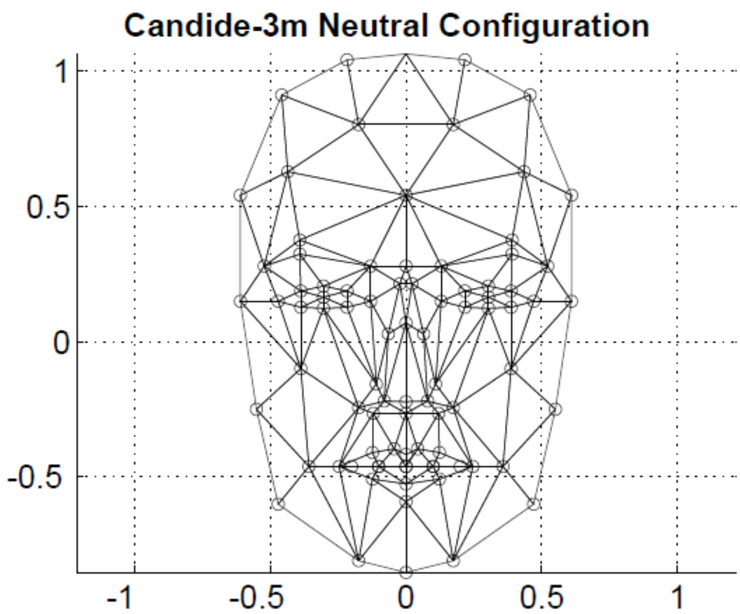

Figure 13 Geometry of the 3D face model (Unzueta et al., 2014)

The first step, the facial features detection, is implemented by an algorithm that detects fiducial facial points using smoothed gradient maps. Although it requires initial knowledge about how a face can be divided, a local image region gradient analysis does not require training with faces databases and due to its efficient estimation of points has a low computational cost.

The second step, the adjustment of the 3D model, is implemented by a fitting algorithm that locates and deforms the 3D face model to fit the detected points. This fitting process considers that a configuration of the 3D face model is determined by position, orientation, shape units (SUs), related to inter-person parameters such as eye width or separation distance, and animation units (AUs), related to expression. The estimation of such configuration is performed by iteratively minimizing the distance between the detected points and the projected model, in a method named deformable model backprojection.

Figure 14 Facial features detection procedure steps. From left to right and top down: (1) eye points detection, (2) eyebrow points detection, (3) mouth points detection, (4) nose points detection, (5) contour points detection and (6) face model fitting on the detected facial features. (Unzueta et al., 2014)

The defined AUs correspond to facial action units, allowing the direct translation of the resulted model configuration into facial action units, which serve as input for the emotional interpretation stage. Including head movement facial action units obtained from the estimated position and orientation, the resulted FACS action units are summarized in Table 2.

FACS AU number	FACS AU name
1	Inner Brow Raiser
2	Outer Brow Raiser
4	Brow Lowerer
5	Upper Lid Raiser
7	Lid Tightener
12	Lip Corner Puller
15	Lip Corner Depressor
16	Lower Lip Depressor
20	Lip Stretcher
23	Lip Tightener
26	Jaw Drop
27	Mouth Stretch
51	Head Turn Left
52	Head Turn Right
53	Head Up
54	Head Down
55	Head Tilt Left
56	Head Tilt Right
57	Head Forward
58	Head Back

Table 2 Detected FACS Action Units

Since the approach followed to provide a rich affective state is to include contextual information in the emotional interpretation process, as described in the introduction chapter, a very simple interpretation process has been implemented with the solely purpose of performing an initial evaluation of the facial expression recognition system.

The interpretation process follows a sign-based approach based on the coding scheme EMFACS[3], resulting in a discrete affective model composed by six basic emotions that are estimated by the combination of the action units described in Table 3.

3 http://face-and-emotion.com/dataface/facs/emfacs.jsp

Discrete Emotion	Action Units
Fear	1+2+4+5+7+20+26
Surprise	1+2+5+26
Sadness	1+4+15
Disgust	9+15+16
Anger	4+5+7+23
Happiness	6+12

Table 3 FACS action units based emotion estimation

3.2 Results

As defined in the objectives of this work, the main purpose of the facial expression recognition system is the detection of facial features during the interaction with a mobile device. This situation involves capturing the image of the face not from a frontal angle as well as different illumination conditions.

In order the evaluate the performance of the recognition system in different configurations of viewing angle and illumination, the CMU Pose, Illumination, and Expression (PIE) database (Sim, Baker, & Bsat, 2003) is used. The database contains 41.368 images of 68 people under 13 different poses, 43 different illumination conditions and 4 different expressions. In Figure 15 some examples of the database used in the evaluation are shown, with the adjusted 3D face model.

Figure 15 Examples of the 3D face model fitting process in the CMU PIE database

The followed approach, the combination of feature detection and the deformable backprojection (FFBP), is compared to another three alternatives: Holistic Approach (HA), Constrained Local Model (CLM) and Supervised Descent (SDM) (see (Unzueta et al., 2014) for more details). Fitting errors for the four approaches using the presented 3D face model are shown in Table 4. Obtained results have less error than HA and similar values to CLM. However, the followed approach does not depend on the quality of the trained model.

	Fitting error	
	Mean	**Stdev**
FFBP	14.93	7.35
HA	32.42	17.16
CLM	13.44	7..82
SDM	9.05	4.14

Table 4 Fitting errors comparison in the CMU PIE database

Since one of the main objectives is to develop a system capable of running on mobile devices, the approach has been evaluated in combination with OAM (FFBP-OAM) in a tracking situation on iOS and Android devices. From the results presented in Table 5 it can be seen that the computation time needed for the initial fitting and the following tracking are low enough to perform a real-time recognition.

Device	**Initialization (ms)**	**Frame tracking (ms)**
iPad 2	60	42
LG Nexus 5	52	26

Table 5 FFBP-OAM average computation times

However, during the evaluation it was found that there is a limitation on the out-of-plane rotations (-30 deg, 30 deg) and that when rapid movements occur, tracking degrades and a fitting reinitialization is needed. Automatic recovery from this situation was explored, but a different approach has been proposed and described in the following section.

Although a simple approach was followed to interpret facial features into an affective state, the emotional result is evaluated. Traditionally, the evaluation of affective recognition systems has been based on lab-controlled datasets where the users posed a particular emotion. Those sets with static backgrounds, controlled illumination and reduced user's head movement do not represent real-world situations.

Recent efforts on allowing researchers to extend and verify their methods on real-world data have resulted on several emotion databases representing real-world scenarios, such as the Acted Facial Expression In The Wild (AFEW), GENKI, Happy People Images (HAPPEI) or Static Facial Expressions In the Wild (SFEW) (Dhall, Goecke, Joshi, Wagner, & Gedeon, 2013). Hence, recognition for these real-world datasets is very challenging and even humans do not perform very well when validating the expressions without contextual information (Gehrig & Ekenel, 2013).

For the evaluation of the affective recognition system, the Acted Facial Expression In The Wild (AFEW) database (Dhall, Goecke, Lucey, & Gedeon, 2012) is used, and it is compared to some of the approaches that participated in the 2013 Emotion Recognition in The Wild Challenge and Workshop[4], of the ACM International Conference on Multimodal Interaction (ICMI 2013). Figure 16 shows some samples of the recognition results with the AFEW database.

Table 6 presents the overall accuracy on the recognition of the AFEW database for the presented approach (FFBP) and the challenge baseline based on LBP-TOP features (LBPTOP), a Convolutional neural Network approach (CN) (Kahou et al., 2013), a Bag of Words (BoW) approach (Sikka, Dykstra, Sathyanarayana, Littlewort, & Bartlett, 2013) and a Grassmannian Discriminant Analysis (GDA) approach (Liu, Wang, Huang, Shan, & Chen, 2013).

Method	**Overall accuracy**
LBPTOP	27,27%
GDA	30,81%
BoW	33,16%
CN	38,10%
FFBP	21,97%

Table 6 Accuracy comparison with the AFEW database

4 http://cs.anu.edu.au/few/emotiw.html

Figure 16 Recognition samples of the AFEW database

Analyzing the recognition results per emotion, shown in Table 7, neutral, angry and happy ones have a greater accuracy, while from the confusion matrix, presented in Table 8, it can be seen that neutral emotion is detected in many cases that correspond to other emotions. These results can be explained by the simple sign-based emotional interpretation approach that was used, where the emotions are detected if several action units are present. Hence, to be detected, facial expressions need to be a bit exaggerated, which results in the many neutral cases. Nevertheless, compared approaches do not perform recognition on real-time as the proposed approach is capable of, with the limited resources of a mobile device.

Emotion	**Accuracy**
Angry	30,51%
Disgust	6,00%
Fear	9,26%
Happy	29,03%
Neutral	49,09%
Sad	15,63%
Surprise	11,54%

Table 7 Recognition accuracy per emotion type

	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	18	1	3	6	15	12	4
Disgust	12	3	6	6	17	6	0
Fear	12	2	5	7	20	5	3
Happy	11	2	4	18	23	4	0
Neutral	9	3	4	7	27	2	3
Sad	15	3	3	6	22	10	5
Surprise	10	3	4	7	16	6	6

Table 8 Recognition confusion matrix

3.3 Future work

The dominant approach on 3D recognition of facial expressions is based on computing a time-varying description of facial actions and then to estimate from them the facial expression. This can be seen as performing a tracking process before performing the recognition process (F. Dornaika & Davoine, 2005). However, both processes are interrelated and performing both simultaneously can increase the reliability and robustness of the solution, allowing to perform facial expression recognition when partial occlusions, rapid movements or video discontinuities are present.

Taking into account the detected problems in the facial expression recognition system for high viewing camera angles and when fast movements and discontinuities occur, the implementation of a simultaneous facial action tracking and expression recognition system following Dornaika and Davoine proposal (Fadi Dornaika & Davoine, 2007) is planned.

The proposed approach considers 3D head pose independent of the facial actions and expressions, allowing to split the estimation of these parameters in two stages, as shown in Figure 17. In the first stage, 3D head pose is obtained by a deterministic registration method based on Online Appearance Models that uses a 3D model similar to the one used in the implemented system. In the second stage, the facial actions and the facial expression are simultaneously estimated using a stochastic framework based on a particle filter.

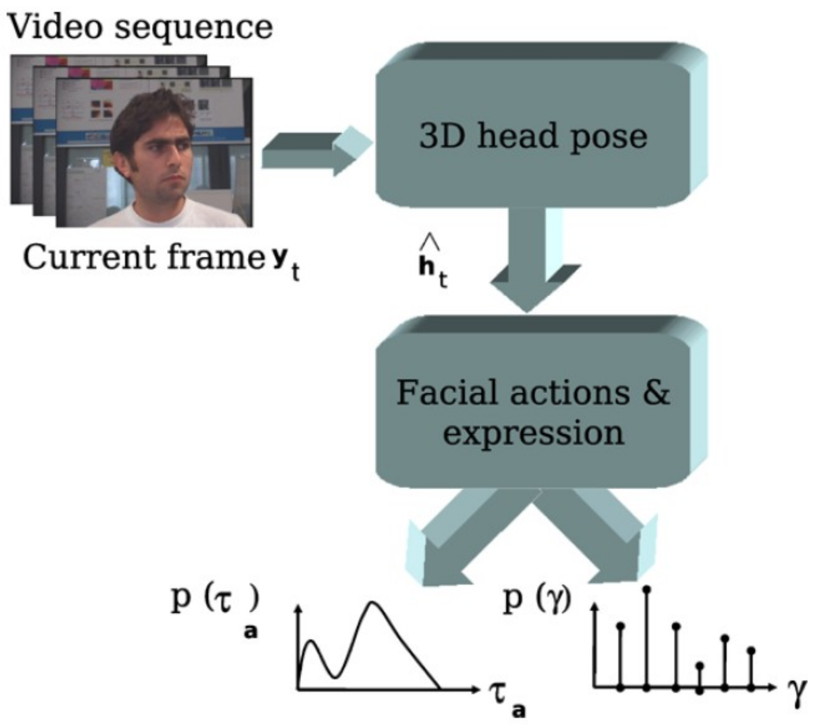

Figure 17 Proposed simultaneous approach (Fadi Dornaika & Davoine, 2007)

The approach was evaluated, assessing its fast recovery to rapid discontinuous facial movements, its great performance on low resolution videos and its robustness to lighting conditions (shown in Figure 18). Moreover, possible inaccuracies affecting the out-of-plane parameters associated with the 3D head pose have no impact on the stochastic tracking and recognition, having an acceptable estimation of the head pose in a wide interval of the out-of-plane rotations [-45 deg, 45 deg].

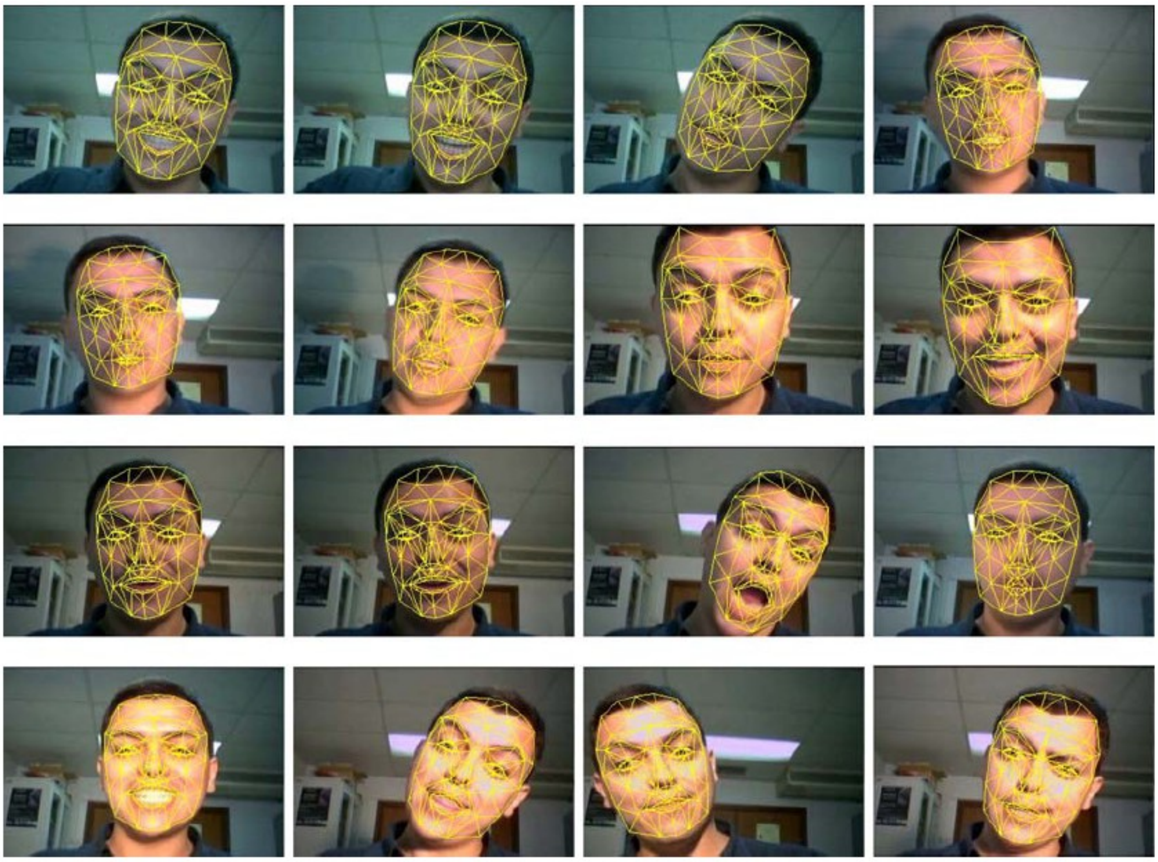

Figure 18 Simultaneous approach performance under illumination, pose and expression changes (Fadi Dornaika & Davoine, 2007)

4. Cognitive model for affective recognition

In this chapter, the proposed cognitive model that allows the integration of the detected facial expression features into an appraisal process is described.

4.1 Proposal design

There are mainly three emotional theories perspectives used on emotion recognition: discrete emotions, dimensional emotions and appraisals. Although discrete models have dominated emotion recognition for decades, systems based on the detection of prototypical emotion expressions are successful with intense and posed expressions, which rarely occur, but not on detecting spontaneous and enacted ones (K. Scherer & Ellgring, 2007). On the other hand, the representation of emotions in dimensional models as states on a multidimensional space seem too limited to capture the complexity of emotions, losing all the context and specifics of the emotion. Hence, appraisal models present an alternative that defines also the cognitive mechanisms of the emotion and which can combine dimensional models on the underlying elicitation of emotions and the discrete categorization of emotion qualities. Moreover, appraisal theories do not generally assume a unique relationship between a situation and the emotional response, being able to consider individual or cultural differences.

Among appraisal models, componential appraisal theories present several advantages such as: "*emotions are defined as complex, multicomponential and dynamic processes*"; allowing to make "*predictions about the determinants that elicit and differentiate emotions*", and defining a "*concrete mechanism underlying emotional response patterning*" (Grandjean et al., 2008).

Although a few appraisal-based approaches for emotion recognition can be found in the literature, due to the fact that "*this approach requires complex, multicomponential and sophisticated measurements of change*" (Gunes & Pantic, 2010), the proposed model includes an appraisal process to better satisfy affective applications requirements.

One of the proposals to use appraisal models in emotion recognition is the work of Mortillaro et al. In (Mortillaro, Meuleman, & Scherer, 2012), the authors propose a framework to create a link between models used for emotion recognition and emotion production. Based on the observation that humans infer appraisals from expressions and use them to establish an emotional state, they propose the use of appraisals as responses in the recognition model to feed the production model to estimate an emotion, as shown in Figure 19. Hence, they suggest to divide the emotion recognition process into the mapping of expressive features to appraisal variables and the mapping of appraisal variables to emotions, using the variables as an intermediate layer. Mortillaro et al. argue that their framework provide a number of benefits: 1) recognition of mixed (or multiple) emotions; 2) ability to integrate contextual information; and 3) facilitation of a finer level of interpretation.

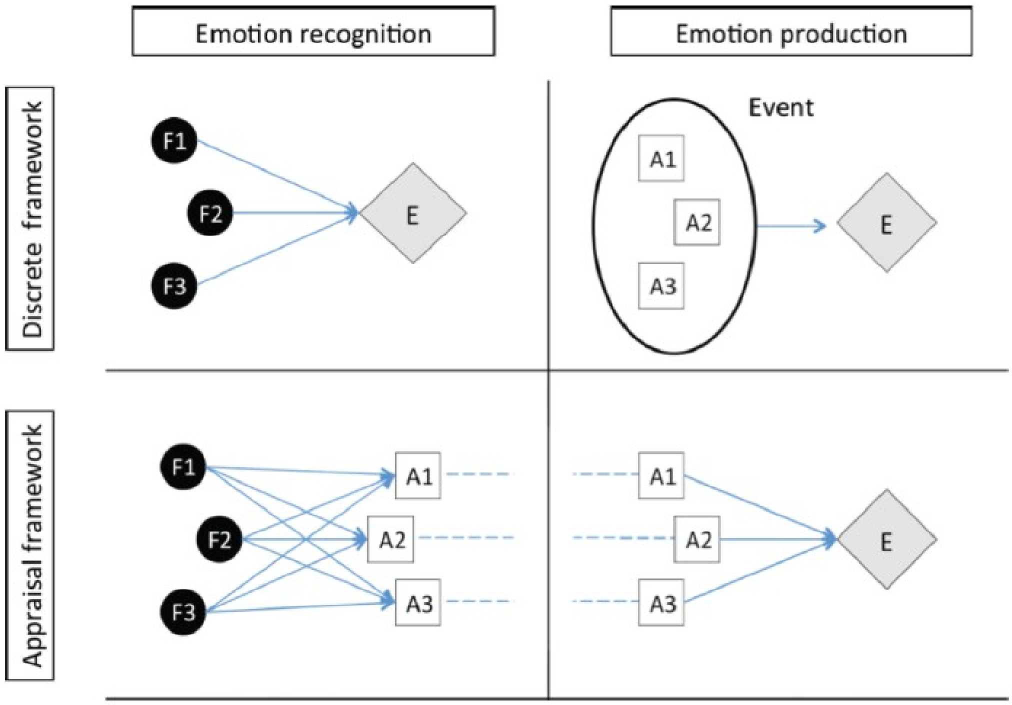

Figure 19: Discrete vs. appraisal framework for emotion recognition and production. Black circles denote expression variables (e.g., individual facial action units), white squares denote appraisal variables (e.g., goal obstructiveness, pleasantness), and grey diamonds denote emotion labels (e.g., anger, fear). (Mortillaro et al., 2012)

Based on Mortillaro et al. framework for integrating appraisal processes on emotion recognition and following unifying approaches presented in previous chapter, the proposed affective model is presented in Figure 20. Moreover, in order to mitigate contradictory differences between emotion theories, a dynamic system has been adopted where for example the dispute on whether emotion follows or precedes cognition has been solved by a loop paradigm.

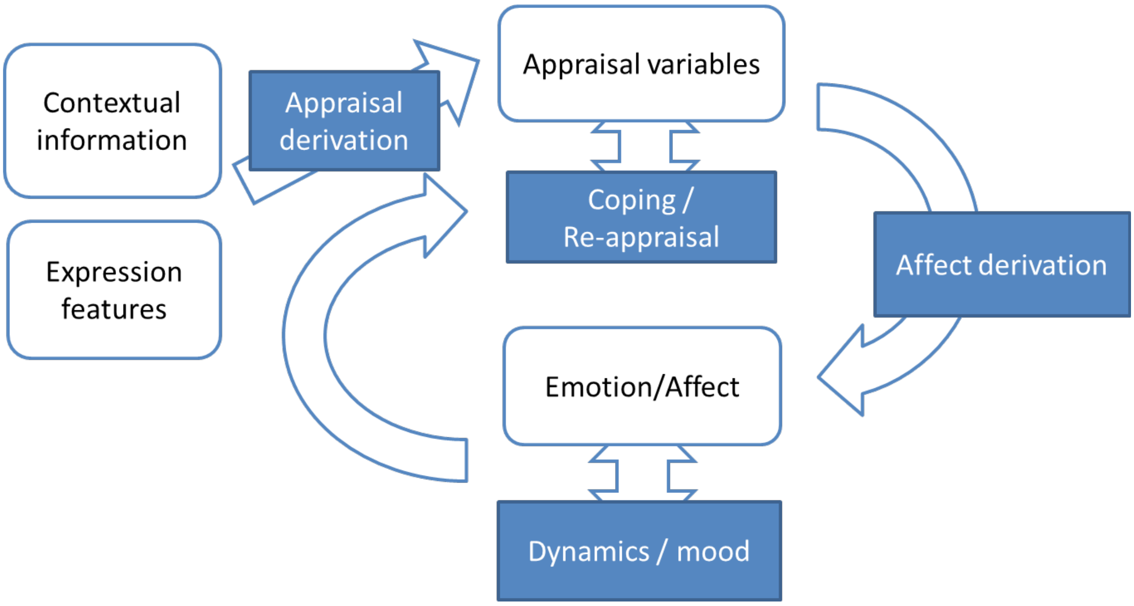

Figure 20: Affective model

In the proposed affective model a set of elements are present:

- The **contextual information** represents the perceived relationship between the user and the environment
- The **expression features** are the set of emotional variables that have been detected, which contain a description of the user's expressed affective state. This set can include facial action units, emotional speech descriptors or physiological signal values.
- The **appraisal variables** are the determined judgments that the system can use to infer the user's affective state.
- The **emotion/affect** is a representation of the user's current affective state, based on the model's working definition of emotion.

The closed-loop cycle of the proposed affective model is based on the following processes:

- **Appraisal derivation** translates the representation of the user-environment relationship and the emotional expressed features into the appraisal variables.
- **Affect derivation** maps appraisal variables to the user's affective state.
- **Dynamics/mood** simulates the dynamic evolution of the affective state over time, establishing the influence of mood.
- **Coping/re-appraisal** is the continuous re-appraisal process that implements coping strategies and closes the appraisal loop.

The user's affective state is represented on the basis of a working definition of emotion, which is composed by:

- **Core affect** as the subjective feeling located in a multi-dimensional space without representing the underlying cognitive process.
- **Full blown emotions** or prototypical emotional episodes, as a short, intensive, event triggered emotional burst.
- **Mood**, as a diffuse affective state of low intensity but relative long duration.

Considering the distinction between back-box models and process models (Wehrle & Scherer, 2001), where the former attempts to model the emotional mechanisms as accurately as possible without taking into account its interpretability, and the latter attempts to describe those mechanism in a meaningful way. The proposed affective model supports a hybrid solution where some of the previously describe processes could be based on a data-driven model that has been trained on empirically gathered data, and some others could be based on a theory driven one.

4.2 Cultural and personal adaptation

The emotional phenomena takes place in a social context (Calvo & D'Mello, 2010) where the individual characteristics determine the underlying process. In this chapter an approach to adapt the affective recognition system to cultural and personality differences of the user is proposed.

4.2.1 Culture and emotions

For many years, the research on the relationship between culture and emotions has been focused on basic emotions with the main consequence of considering universality and culture-specifity of emotions as mutually exclusive (Mesquita, Frijda, & Scherer, 1997). Hence, many efforts have been dedicated to study whether emotions are cross-culturally different or similar, focusing on their phenomenology from a concept of emotions as states rather than processes.

For example, wide evidence has been found on the universality of a set of facial expressions. However, cross-cultural studies have shown that the judgment of facial expression is also culturally dependent and partially influenced by learned display rules (Fasel & Luettin, 2003), that set cultural variations in the expressive behaviour.

Therefore, there is strong evidence of several universal aspects of emotions. They are present in all cultures, regardless of existing a linguistic term for them, and relevant events, in terms of the culture specific concerns, result on an emotional response.

On the other hand, there is also evidence on cross-cultural differences. For example, there is evidence for cultural differences in the elicitation of emotional events, where the same situations are interpreted differently because of different conditions or beliefs.

From a componential perspective of emotion, the universal contingencies hypothesis tries to explain cross-cultural differences through similarities in the appraisal process: "*if people from different cultures appraise a situation in the same way, they will experience the same emotion. If they experience a different emotion, it is because they have appraised the situation differently.*" (Mesquita & Ellsworth, 2001)

Although there is few evidence that cultural differences in emotions can be explained by differences in appraisals, some approaches can be found to integrate cultural differences in appraisal processes.

Some studies have provided evidence that several appraisal dimensions can be found across different cultures. However, classifying those dimensions into simple, which are more likely to be automatic and possibly subcortical (i.e. attention, valence and coping ability), and complex, more likely to be delayed and conscious (i.e. morality, fairness and agency), there is an hypothesis that relates simple appraisals with universality of emotions and complex appraisals with cross-cultural differences.

On the other hand, other studies propose the existence of different appraisal dimensions among cultures (i.e. interpersonal engagement in Japanese culture or esteem by others on Surinamese and Turkish (Mesquita & Ellsworth, 2001)). However, it is difficult to discern between the existence of a new appraisal dimension or a culturally higher prevalence of certain dimensions.

Another approach to explain cultural differences is to focus on the cultural salience, what a culture define as important, of events, emotions and appraisals. Hence, certain events can have culturally assigned meanings that cannot be unnoticed and will affect the emotional state of the individual, and even might be appraised in a culturally preconceived ways. Moreover, some emotions might be admired or despised by certain cultures, guiding the appraisal process and defining some appraisals more likely than others. Finally, some appraisals can be more emphasized than others, resulting in a higher importance of certain appraisal dimensions.

4.2.2 Personality and emotions

As well as there are cultural differences on how an emotion is elicited, individual differences influence the emotional phenomena. In fact, one of the original motivations underlying the development of appraisal theory was to account systematically for individual differences in emotional experience (Lazarus, 1994). However, although existing evidence of these individual differences, little effort has been dedicated.

The individual differences in the emotional experience can influence different appraisal components. For example, Kuppens et al. Suggest three different sources of individual differences: “*(1) individual differences in how a situation is appraised, (2) individual differences in how the appraisals are associated with emotional experience, (3) individual differences in emotional experience may arise from dispositional factors outside the ‘appraisal pathway.”* (Kuppens, Van Mechelen, & Rijmen, 2008)

Moreover, multiple individual differences can be considered to influence the emotional experience. Griner and Smith propose to integrate theoretical issues from the study of personality, including “*motivational constructs such as one's orientations, needs, values and goals*”. (Griner & Smith, 2000).

Considering the influence on personality and focusing on coping mechanisms, Li (Li, 2009) proposes to combine appraisal theories, that state that is the cognitive process responsible for coping, with diathesis-oriented theories, that state that individuals' personality traits influence coping.

In summary, individual differences, such as personality, should be part of the appraisal processes, integrated for example as parameters that modulate these processes.

4.2.3 Adaptive approach

In order to define the cultural and personality characteristics of the user, a model must be defined. As presented in Chapter 2.5, Hofstede cultural dimensions[5] are commonly used to state cultural differences. For defining the personality, several personality traits can be found in the literature. Among them, a very used set of traits are the Big Five personality traits (Costa, McCrae, & Psychological Assessment Resources, 1992). Thus, the resulting model is composed by the following dimensions:

- Cultural dimensions:
  - Power Distance Index (PDI): “*the degree to which the less powerful members of a society accept and except that power is distributed unequally*”.
  - Individualism vs Collectivism (IDV): "*individualism can be defined as a preference for a loosely-knit social framework in which individuals are expected to take care of only themselves and their immediate families*”, while “*collectivism represents a preference for a tightly-knit framework in society in which individuals can expect their relatives or members of a particular in-group to look after them*”
  - Masculinity vs Femininity (MAS): “*masculinity represents a preference in society for achievement, heroism, assertiveness and material rewards for success*”, as opposed to femininity that “*stands for a preference for cooperation, modesty, caring for the weak and quality of life*”.

5 http://geert-hofstede.com/dimensions.html

  - Uncertainty Avoidance Index (UAI): “*the degree to which the members of a society feel uncomfortable with uncertainty and ambiguity*”
  - Long Term Orientation vs Short Term Normative Orientation (LTO): “*societies with low score prefer to maintain time-honoured traditions and norms*”, whether those with high score “*take a more pragmatic approach*”.
  - Indulgence vs Restraint (IND): “*Indulgence stands for a society that allows relatively free gratification of basic and natural human drives related to enjoying life and having fun*”, as opposed to Restraint, that “*stands for a society that suppresses gratification of needs and regulates it by means of strict social norms*”.
- Personality dimensions:
  - Extraversion: this trait includes characteristics such as excitability, sociability, talkativeness, energy, surgency and emotional expressiveness.
  - Neuroticism: the extent to which individuals exhibit anxiety, depression, and irritability as well as act impulsively and unstable, and experience a sense of vulnerability.
  - Agreeableness: this trait includes attributes such as trust, altruistic and cooperative behavior, kindness, affection, straightforward and honest communication, as well as tender and sympathetic attitudes.
  - Conscientiousness: he degree to which individuals are competent, methodical, dutiful, disciplined, and with good impulse control and goal-directed behaviors.
  - Openness to experience: the extent to which individuals are open to fantasies, aesthetics, adventure, unusual ideas, as well as novel actions.

As previously presented, the cultural and personality differences, defined by the model dimensions, influence the appraisal processes. Following the proposed cognitive model, this influence is detailed for each of the appraisal process:

- Appraisal derivation: Interpretation of the events is influenced by the cultural differences. Therefore, cultural dimensions should be integrated in the stimulus evaluation process.
- Affect derivation: Cultural and personality dimensions should be translated into differences in the salience or accessibility of appraisal dimensions, resulting in different affective state for certain appraisals.
- Dynamics/mood: Personality has an important influence in the dynamics of the affective state, determining its evolution over time.
- Coping/re-appraisal: Cultural differences determine the prevalence of certain emotions, as well as the personality has a great influence in coping mechanism.

In summary, in order to adapt the cognitive model for affective recognition to the individual characteristics of the user, defined in this case as cultural and personality dimensions, a parametrized implementation is proposed, where each of the appraisal process is influenced by these dimensions.

5. Conclusions and Future Work

Although Human Computer Interaction has greatly evolved during the last decades, there is still a long path to go until we reach an interaction where the main principles and objectives are to provide emotional engagement, empathy and human connection, placing users' emotions in the centre of attention.

Together with a change on the design direction, there is a need of tools and guides for the development of emotional intelligent systems. In order to make applications capable of reacting to users' emotions, one of this needs is to develop affective recognition systems that can be integrated into our daily interactions.

This Master Thesis aims to take a few steps towards an affective recognition system that running on a mobile device can estimate a rich an extensive emotional state of the user.

The first step taken, and described in Chapter 3, is the development of a facial expression recognition system that can perform facial features detection in mobile devices. Due to the simple approach followed in the emotional feature interpretation, the evaluation under real-world situations resulted in a poor emotional recognition accuracy. However, being the actual objective of the recognition system, the feasibility of the system for detecting facial expressions was proven under severe pose and illumination conditions.

The second step, presented in Chapter 4, is the proposal of a cognitive model that combines emotional expression features and contextual information into a multicomponential affective state. Based on the extended among the research community appraisal theory, the model describes the different underlying emotional processes and their relationships needed to perform the affective recognition. Moreover, taking back the original motivation of appraisal theory to account for individual differences in the emotional experience, an adaptive approach is proposed to define and integrate the culture and personality of the user into the appraisal recognition processes.

In summary, as opposed to the dominating approaches in affective recognition that focus on obtaining the most precise and accurate information from a single emotional expression (i.e. detect a large number of emotions from facial expression) or combining multimodal expressions (i.e. combination of facial and speech emotional expression), the proposed approach focuses on what we humans do to estimate another person's emotional state, combine the expression information with the contextual information. By interpreting the detecting emotional features with the related situation and events, it will be possible to estimate a detailed emotional state of the user in spontaneous and non-posed experiences of the real life.

Many steps remain to achieve the proposed goal. From the several ones that can be foreseen, the following future work is proposed:

- Implementation of the facial expression recognition enhancing proposal to improve performance under the difficult conditions that entail the interaction with a mobile device.
- Implementation of an affective recognition system based on the proposed cognitive model that integrates detected facial expressions and contextual information into a multicomponential affective state.
- Culture and personality dimensions based parameterization of the affective recognition appraisal processes.

Bibliography

Ahlberg, J. (2001). Candide-3-an updated parameterised face, *1*, 1–16. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.33.5603

Becker-Asano, C. (2008). *WASABI: Affect simulation for agents with believable interactivity*. University of Bielefeld. Retrieved from http://www.becker-asano.de/Becker-Asano_WASABI_Thesis.pdf

Benyon, D. (1993). Accommodating Individual Differences through an Adaptive User Interface. *Differences*, *10*, 1–16. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Accommodating+Individual+Differences+through+an+Adaptive+User+Interface#0

Blanz, V. (2006). Face recognition based on a 3D morphable model. *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, 617–624. doi:10.1109/FGR.2006.42

Broekens, J. (2010). Modeling the Experience of Emotion. *International Journal of Synthetic Emotions*, *1*(1), 1–17. doi:10.4018/jse.2010101601

Broekens, J., DeGroot, D., & Kosters, W. (2008). Formal models of appraisal: Theory, specification, and computational model. *Cognitive Systems Research*, *31*(0). Retrieved from http://www.sciencedirect.com/science/article/pii/S1389041707000162

Callahan, E. (2006). Interface design and culture. *Annual Review of Information Science and Technology*, *39*(1), 255–310. doi:10.1002/aris.1440390114

Calvo, R., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *Affective Computing, IEEE Transactions …*, *1*(1), 18–37. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5520655

Costa, P. T., McCrae, R. R., & Psychological Assessment Resources, I. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI)*. Psychological Assessment Resources. Retrieved from http://books.google.es/books?id=mp3zNwAACAAJ

Dhall, A., Goecke, R., Joshi, J., Wagner, M., & Gedeon, T. (2013). Emotion recognition in the wild challenge 2013. *Proceedings of the 15th ACM on International Conference on Multimodal Interaction - ICMI '13*, 509–516. doi:10.1145/2522848.2531739

Dhall, A., Goecke, R., Lucey, S., & Gedeon, T. (2012). Collecting large, richly annotated facial-expression databases from movies. *IEEE Multimedia*, *6*(1), 1–14. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.407.4632

Dornaika, F., & Davoine, F. (2005). View and texture-independent facial expression recognition in videos using dynamic programming. In *IEEE International Conference on Image Processing 2005* (pp. II–1314). IEEE. doi:10.1109/ICIP.2005.1530305

Dornaika, F., & Davoine, F. (2007). Simultaneous Facial Action Tracking and Expression Recognition in the Presence of Head Motion. *International Journal of Computer Vision*, *76*(3), 257–281. doi:10.1007/s11263-007-0059-7

Ekman, P. (1999). Basic Emotions. In *Handbook of Cognition and Emotion* (pp. 45–60). John Wiley & Sons, Ltd. doi:10.1002/0470013494.ch3

El-Nasr, M., Yen, J., & Ioerger, T. (2000). Flame—fuzzy logic adaptive model of emotions. *Autonomous Agents and Multi-Agent Systems*, *3*(3), 219–257. Retrieved from http://link.springer.com/article/10.1023/A:1010030809960

Fasel, B., & Luettin, J. (2003). Automatic facial expression analysis: a survey. *Pattern Recognition*. Retrieved from http://www.sciencedirect.com/science/article/pii/S0031320302000523

Fieser, J., & Dowden, B. (2014). Theories of Emotion. *Internet Encyclopedia of Philosophy*. Retrieved from http://www.iep.utm.edu/

Francis, A. G., Mehta, M., & Ram, A. (2009). Emotional memory and adaptive personalities. *Handbook of Research on …*. Retrieved from http://www.igi-global.com/chapter/emotional-memory-adaptive-personalities/56197

Gehrig, T., & Ekenel, H. K. (2013). Why is facial expression analysis in the wild challenging? *Proceedings of the 2013 on Emotion Recognition in the Wild Challenge and Workshop - EmotiW '13*, 9–16. doi:10.1145/2531923.2531924

Gilroy, S. W., Cavazza, M., Niiranen, M., Andre, E., Vogt, T., Urbain, J., … Billinghurst, M. (2009). PAD-based multimodal affective fusion. In *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops* (pp. 1–8). IEEE. doi:10.1109/ACII.2009.5349552

Goleman, D. (2006). *Emotional intelligence*. Bantam Books. Retrieved from http://books.google.es/books?id=TQQlAQAAIAAJ

Goren-Bar, D., Graziola, I., Pianesi, F., & Zancanaro, M. (2006). The influence of personality factors on visitor attitudes towards adaptivity dimensions for mobile museum guides. *User Modeling and User-Adapted Interaction*, *16*(1), 31–62. doi:10.1007/s11257-006-9004-7

Grandjean, D., Sander, D., & Scherer, K. R. (2008). Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization. *Consciousness and Cognition*, *17*(2), 484–95. doi:10.1016/j.concog.2008.03.019

Gratch, J., & Marsella, S. (2004). A domain-independent framework for modeling emotion. *Cognitive Systems Research*, *5*(4), 269–306. Retrieved from http://www.sciencedirect.com/science/article/pii/S1389041704000142

Griner, L. a., & Smith, C. a. (2000). Contributions of Motivational Orientation to Appraisal and Emotion. *Personality and Social Psychology Bulletin*, *26*(6), 727–740. doi:10.1177/0146167200268007

Gunes, H., & Pantic, M. (2010). Automatic, Dimensional and Continuous Emotion Recognition. *International Journal of Synthetic Emotions*, *1*(1), 68–99. doi:10.4018/jse.2010101605

Heimgärtner, R., Holzinger, A., & Adams, R. (2008). From Cultural to Individual Adaptive End-User Interfaces: Helping People with Special Needs. In *Proceedings of the 11th international conference on Computers Helping People with Special Needs* (pp. 82–89). Berlin, Heidelberg: Springer-Verlag. doi:10.1007/978-3-540-70540-6_11

Kahou, S. E., Pal, C., Bouthillier, X., Froumenty, P., Memisevic, R., Vincent, P., … Bengio, Y. (2013). Combining Modality Specific Deep Neural Networks for Emotion Recognition in Video.

Kaiser, S., & Wehrle, T. (2001). Facial expressions as indicators of appraisal processes. *Appraisal Processes in Emotion: Theory, Methods, …*, 285–300. Retrieved from http://www.unige.ch/fapse/emotion/members/kaiser/article02.pdf

Kuppens, P., Van Mechelen, I., & Rijmen, F. (2008). Toward disentangling sources of individual differences in appraisal and anger. *Journal of Personality*, *76*(4), 969–1000. doi:10.1111/j.1467-6494.2008.00511.x

Lazarus, R. S. (1994). *Emotion and Adaptation*. Oxford University Press. Retrieved from http://books.google.es/books?id=tTdIlwpxtWsC

Li, M. (2009). Cognitive Appraisal and / or Personality Traits : Enhancing Active Coping in Two Types of Stressful Situations. *Cognition*. Retrieved from http://www.counseling.org/resources/library/VISTAS/2009-V-Online/Li.doc

Lin, J., Spraragen, M., & Zyda, M. (2012). Computational Models of Emotion and Cognition. *Advances in Cognitive Systems*, *2*, 59–76.

Liu, M., Wang, R., Huang, Z., Shan, S., & Chen, X. (2013). Partial least squares regression on grassmannian manifold for emotion recognition. *Proceedings of the 15th ACM on International Conference on Multimodal Interaction - ICMI '13*, 525–530. doi:10.1145/2522848.2531738

Marsella, S., & Gratch, J. (2009). EMA: A process model of appraisal dynamics. *Cognitive Systems Research*, *10*(2000), 70–90. Retrieved from http://www.sciencedirect.com/science/article/pii/S1389041708000314

Marsella, S., Gratch, J., & Petta, P. (2010). Computational models of emotion. *A Blueprint for Affective Computing (Series in Affective Science)*. Retrieved from http://books.google.com/books?hl=en&lr=&id=C2gLOQ105okC&oi=fnd&pg=PA21&dq=Computational+Models+of+Emotion&ots=-L9MZ7bkyR&sig=x9i7or9jlQISWSwipHtlE4LfXEM

Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. M.I.T. Press. Retrieved from http://books.google.es/books?id=EthOAAAAMAAJ

Mesquita, B., & Ellsworth, P. C. (2001). The Role of Culture in Appraisal. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal Processes in Emotion : Theory, Methods, Research: Theory, Methods, Research*. Oxford University Press.

Mesquita, B., Frijda, N. H., & Scherer, K. R. (1997). Culture and Emotion. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), *Handbook of Cross-cultural Psychology: Basic processes and human development* (pp. 255–297). Allyn and Bacon. Retrieved from http://books.google.es/books?id=tLvAmyvsU8UC

Moon, Y., & Nass, C. I. (1996). Adaptive agents and personality change: complementarity versus similarity as forms of adaptation. In *Conference Companion on Human Factors in Computing Systems* (pp. 287–288). New York, NY, USA: ACM. doi:10.1145/257089.257325

Mortillaro, M., Meuleman, B., & Scherer, K. R. (2012). Advocating a Componential Appraisal Model to Guide Emotion Recognition. *Int. J. Synth. Emot.*, *3*(1), 18–32. doi:10.4018/jse.2012010102

Nisbett, R. (2004). *The Geography of Thought: How Asians and Westerners Think Differently...and Why*. Free Press. Retrieved from http://books.google.es/books?id=525HX623L_cC

Ortony, A., Clore, G. L., & Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge University Press. Retrieved from http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/0521353645

Ortony, A., & Turner, T. (1990). What’s basic about basic emotions? *Psychological Review*, *97*(3), 315–331. Retrieved from http://psycnet.apa.org/journals/rev/97/3/315/

Pantic, M., & Rothkrantz, L. (2000). Automatic analysis of facial expressions: The state of the art. *… Analysis and Machine Intelligence, …*, *22*(12), 1424–1445. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=895976

Picard, R. W. (1995). Affective Computing.

Plutchik, R. (1980). *Emotion: A Psychoevolutionary Synthesis*. Harper and Row. Retrieved from http://books.google.es/books?id=G5t9AAAAMAAJ

Reeves, B., & Nass, C. (1996). *The media equation: how people treat computers, television, and new media like real people and places*. New York, NY, USA: Cambridge University Press.

Reinecke, K. (2010). *Culturally Adaptive User Interfaces*. Retrieved from http://www.ifi.uzh.ch/teaching/studiengaenge/phd/red-tape/DINF08_DissTitelPage.doc

Reisenzein, R. (2007). What is a definition of emotion? And are emotions mental-behavioral processes? *Social Science Information*, *46*(3), 424–428. doi:10.1177/05390184070460030110

Reisenzein, R., Hudlicka, E., Dastani, M., Gratch, J., Hindriks, K., Lorini, E., & Ch. Meyer, J.-J. (2013). Computational Modeling of Emotion: Toward Improving the Inter- and Intradisciplinary Exchange. *IEEE Transactions on Affective Computing*, *4*(3), 246–266. doi:10.1109/T-AFFC.2013.14

Roseman, I. J. (1984). Cognitive determinants of emotions: A structural theory. *Review of Personality and Social Psychology*, *5*(Emotions, relationships, and health), 11–36.

Roseman, I., & Smith, C. (2001). Appraisal theory: Overview, assumptions, varieties, controversies. In A. Schorr & K. Scherer (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 221–232). London: Oxford University Press. Retrieved from citeulike-article-id:1624042

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. doi:doi: 10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. doi:10.1037/0033-295X.110.1.145

Scherer, K., & Ellgring, H. (2007). Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion*, *7*(1), 113–130. Retrieved from citeulike-article-id:1623661

Scherer, K. R. (1984a). Emotion as a multicomponent process: A model and some cross-cultural data. *Review of Personality & Social Psychology*. Retrieved from http://psycnet.apa.org/psycinfo/1986-17269-001

Scherer, K. R. (1984b). On the nature and function of emotion: A component process approach. In K. R. Scherer & P. Ekman (Eds.), *Approaches To Emotion* (pp. 293–317). Hillsdale, NJ: Lawrence Erlbaum.

Scherer, K. R., & Schorr, A. (2001). Appraisal considered as a process of multilevel sequential checking. In T. Johnston (Ed.), *Appraisal processes in emotion: Theory, Methods, Research* (pp. 92–120). Oxford University Press. Retrieved from http://emotion-research.net/projects/humaine/ws/plenary2005/modeling/Scherer-01-Appraisal.pdf

Sikka, K., Dykstra, K., Sathyanarayana, S., Littlewort, G., & Bartlett, M. (2013). Multiple Kernel Learning for Emotion Recognition in the Wild. In *Proceedings of the 15th ACM on International Conference on Multimodal Interaction* (pp. 517–524). New York, NY, USA: ACM. doi:10.1145/2522848.2531741

Sim, T., Baker, S., & Bsat, M. (2003). The CMU Pose, Illumination, and Expression (PIE) database. In *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition* (pp. 53–58). IEEE. doi:10.1109/AFGR.2002.1004130

Strongman, K. T. (1978). *The psychology of emotion*. Wiley. Retrieved from http://books.google.es/books?id=taYQAQAAIAAJ

Unzueta, L., Pimenta, W., Goenetxea, J., Santos, L. P., & Dornaika, F. (2014). Efficient generic face model fitting to images and videos. *Image and Vision Computing*, *32*(5), 321–334. doi:10.1016/j.imavis.2014.02.006

Wehrle, T., & Scherer, K. (2001). Towards computational modeling of appraisal theories. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 350–365). Oxford University Press. Retrieved from http://www.unige.ch/fapse/emotion/members/kaiser/article03.pdf