Software Architecture for Smart Emotion Recognition and Regulation of the Ageing Adult

José Carlos Castillo [1,2,*]

Phone +34 967 599200
Email jccmontoya@gmail.com

Álvaro Castro-González [2]

Antonio Fernández-Caballero [1]

José Miguel Latorre [3]

José Manuel Pastor [4]

Alicia Fernández-Sotos [5,6]

Miguel A. Salichs [2]

[1] Instituto de Investigación en Informática de Albacete, Universidad de Castilla-La Mancha, 02071 Albacete, Spain

[2] Robotics Lab, Universidad Carlos III de Madrid, 28911 Madrid, Spain

[3] Instituto de Investigación en Discapacidades Neurológicas, Universidad de Castilla-La Mancha, 02071 Albacete, Spain

[4] Instituto de Tecnologías Audiovisuales, Universidad de Castilla-La Mancha, 16071 Cuenca, Spain

[5] Facultad de Educación de Albacete, Universidad de Castilla-La Mancha, 02071 Albacete, Spain

[6] Conservatorio Profesional de Música Maestro Gómez Villa, 30530 Cieza, Murcia, Spain

Abstract

This paper introduces the architecture of an emotion-aware ambient intelligent and gerontechnological project named “Improvement of the Elderly Quality of Life and Care through Smart Emotion Regulation”. The objective of the proposal is to find solutions for improving the quality of life and care of the elderly who can or want to continue living at home by using emotion regulation techniques. A series of sensors is used for monitoring the elderlies’ facial and gestural expression, activity and behaviour, as well as relevant physiological data. This way the older people’s emotions are inferred and recognized. Music, colour and light are the stimulating means to regulate their emotions towards a positive and pleasant mood. Then, the paper proposes a gerontechnological software architecture that enables real-time, continuous monitoring of the elderly and provides the best-tailored reactions of the ambience in order to regulate the older person’s emotions towards a positive mood. After describing the benefits of the approach for emotion recognition and regulation in the elderly, the eight levels that compose the architecture are described.

AQ1

AQ2

Keywords

Emotion recognition
Emotion regulation
Ambient intelligence
Gerontechnology
Software architecture

Introduction

Population ageing is a recent concern in developed countries due to the decreasing birth rates and the higher life expectancy. This phenomenon relates to a shift in the structure of the population ages, where basically those groups with greater ages grow, while younger population is reduced. Although improvements in nutrition and health care enable ageing with high quality of life, there are cases in which some health-related issues require close and personalized supervision. Indeed, Ambient Assisted Living (AAL) approaches are gaining importance in recent years because they can provide a personalized support to older people, mainly at home. We believe that perceiving and enhancing the quality of life of the elderly who live at home is possible through automatic emotion recognition and regulation. Some insights in these topics are presented in Sect. 2 from a psychological perspective.

From an individual perspective, quality of life can be considered in terms of well-being. It includes emotional (self-esteem, emotional intelligence, mindset), social (friends, family, community) and physical (health, physical safety) aspects in a person's life. Indeed, it has been largely studied that positive emotional states promote healthy perceptions, beliefs and physical well-being [38]. Now, quality of life and care (QL&C) assesses how welfare is affected due to illness or disability. These aspects can only be assessed through proper monitoring of the individuals in the environment. Usually, elder people tend to live at their own home rather than choosing other options. This leads to many challenges in order to adapt some environments to provide safety and home care to the elder, particularly if the elder is in trouble. In recent years, Information Technology and Communications (ICT) has been playing an active role towards mitigating the most common problems found at home (e.g. [4, 9, 12]).

This paper presents the software architecture of a system to guide the ageing adults towards a positive mood. In this field, many terms have been used in the literature with no clear distinction among them: emotion, mood, feelings or emotional state, among others. For the purpose of this work, we will not differentiate between them and use them as synonyms. The study of the subtle differences among these terms is out of the scope of this paper.

Furthermore, the system, described in Sect. 3, aims at finding solutions to improve the quality of life and care of ageing adults who can or want to keep living at home. The approach uses advanced tools and techniques of ICT supplemented with expert knowledge based on experimental techniques from psychology, neurobiology and music about the regulation of emotions.

Computer tools and techniques used in this project are directly linked to the concepts of Ambient Intelligence (AmI). Indeed, we are convinced that AmI is a well-suited area for the design, implementation and deployment of global systems that seek to provide effective solutions to real problems in our society. The general objective of this project results in some specific goals, such as (1) to analyse the emotional states and the regulation techniques based on the knowledge provided by subject matter experts, (2) to monitor and recognize emotions in Ambient Intelligence environments, (3) to regulate emotions through adapting the environment, (4) to construct an intelligent emotion regulation system and (5) to validate the emotion regulation system based on the expert knowledge. The expected project outcome is an AAL system to raise the ageing adult's QL&C.

In this article, we introduce a gerontechnological software architecture able to deal with the problem of detecting and regulating emotions through ambient intelligent which is described in Sect. 4. The proposal is prepared to accept any sensor and actuator for continuously monitoring of the elder, providing the best-tailored reactions to regulate their emotional state towards a positive mood. Section 5 presents the limitations of the system and some general issues that must be considered. Finally, the conclusions are summarized in Sect. 6.

Fundamentals of Emotion Recognition and Regulation

For emotion interpretation and regulation, we heavily rely on Ambient Intelligence (AmI) [37] which proposes the creation of intelligent environments to suit the needs, tastes and interests of people that live inside them. It is worth highlighting that AmI integrates three major research areas: sensing of the environment and/or people, application of computing technologies, and acting on the real and/or virtual environment. Moreover, the relationship between emotions and AmI is called "emotion-aware AmI" (AmE) [50], which could be defined as emotion conscious AmI. AmE exploits concepts from psychology and social sciences to adequately analyse the state of the individual and enrich the contextual information. AmE achieves this objective by extending the AmI devices through a collection of improved sensors able to recognize human emotions from facial expressions [46] and human behaviours, such as hand gestures, body movements and speech [42, 48]. The exploitation of these new emotion context-aware devices allows systems to deliver a highly personalized and dedicated collection of services to support users and improve their personal care.

Automatic monitoring of emotional states is a valuable tool in the areas of Health Sciences and Rehabilitation, Clinical Psychology, Psychiatry and Gerontology. Current research in wireless area networks (WSNs) [5] and body area networks [20] enables the inclusion of advanced monitoring devices. Thus, energy-efficient small size devices are being dedicated to body feature measurements. It is important to consider a multisensory approach, which combines multiple sources of information presented by various sensors to generate a more accurate and robust interpretation of the environment [35]. The availability of new types of sensors for monitoring tasks poses new challenges in multisensory data fusion. Now, it is possible to construct models of the environment and to diagnose situations through the analysis of a sequence of sensory information.

Ultimately, we are talking about semantic interpretation from multi-sensed and video-controlled environments [11, 16]. Such interpretation is defined as the need to recognize situations, activities and interactions between different actors involved [13]. In these environments in which information comes from different sensors or a combination of them, the problem is to link the physical signals received with the interpretation of their meaning.

In our case, the goal is to develop a software architecture to recognize the activity and mood of older people living alone in their homes. Indeed, we believe that the ability to monitor changes in the emotional state of people in their own context allows implementing regulatory strategies for reducing negative affect. Despite the fact that emotion interpretation in humans has traditionally been an area of interest in disciplines such as psychology and sociology, there is a lack of real applications that relate emotion to human behaviour. This fact is mainly due to the large number of theoretical models proposed and the complexity of human emotions [27]. Often, emotions appear as an affective phenomenon, i.e. as a state of mind, an interpersonal attitude or a personality trait. Moreover, other authors identify emotions by intensity and duration and in combination with other emotions to result into more complex ones [41].

Existing emotion recognition technologies are divided into four major categories depending on what kind of data is analysed: physiological signals, facial expressions, behaviour or voice [24], [21]. Physiological emotion recognition shows acceptable performance but has some critical weaknesses that prevent its widespread use; they are obtrusive to users and need special equipment or devices [36]. The capture of physiological data is increasingly done through body sensor networks. These allow continuous measurement in the daily life of a person of physiological parameters such as heart rate, muscle tension, skin conductance, breathing rate and so on. When combined with contextual information extracted from the environment through WSNs, such as the ambient temperature or humidity, these parameters can be used to infer emotions [31]. Facial analysis includes a number of processing steps that attempt to detect or track the face, to locate characteristic facial regions such as eyes, mouth and nose on it, to extract and follow the movement of facial features. This is one of the least intrusive processes of automatic emotion recognition. In recent years, various methods have been developed for extracting and analysing the facial features from a description of facial expressions. The Facial Action Coding System [7] encodes all possible facial expressions as unitary actions (UAs) that may occur individually or in combination. In fact, facial expressions associated with emotions are generally described as a set of UAs [45]. Let us highlight the FaceSense research project [10] developed by the Affective Computing group of Massachusetts Institute of Technology. Speech signal conveys a large amount of information being also a nonintrusive technique for emotion recognition. Two main problems are addressed: (1) finding the set of features in the speech signal that are most significant in conveying emotions and (2) finding the best classification algorithm that can indicate emotional expression, based on the extracted features [6].

Emotion regulation refers to a set of processes that either stop the emotion from emerging or prevent it from being expressed once it is triggered [18]. Bottom-up emotion generation refers to the elicitation of emotion by the presentation of a stimulus that is thought to have simple physical properties that are inherently emotional [30]. Antecedent emotion regulation strategies apply, while the emotion is still unfolding and has not reached its peak. In our case, we are interested in response-focused emotion regulation, which tries to aim at altering and controlling the experiential, behavioural and physiological response to the fully established emotion. There is a great deal of work based on the use of music [1], colour [8] and lighting [34] in emotion regulation.

Recently, different innovative strategies have been developed to improve mood through external stimulation. In the field of gerontology, some techniques have been used to aid reminiscence and life review that claim to help the elderly to remember autobiographical events in an organized manner by using auditory stimuli such as music and visual stimuli such as photographs or videos. Specifically, “Life Review Based on Specific Positive Events” (ReVISEP) [42] is a technique aimed at the recovery of especially positive memories occurred throughout life. The main effect of this technique is a better mood and increased life satisfaction, both in older people with and without major depression.

For a long period in the history of psychological research, emotion and cognition were studied independently as if they were irrelevant to one another. However, after the appearance of a pioneering work [40], renewed interest in the study of the relationship between cognition and emotion has led to the development of a wide range of techniques that allow temporary induction of different mood, both positive and negative. The first modern technique is the Velten mood induction procedure [47]. Other developed techniques are induction through music [32], film sequences [19] and autobiographical memory [3]. It is considered more useful to combine two or more induction techniques simultaneously, since multiple inductions contribute additively to mood [2]. In the same way that emotional states are induced for research, it is also possible to induce emotional states in order to regulate emotions.

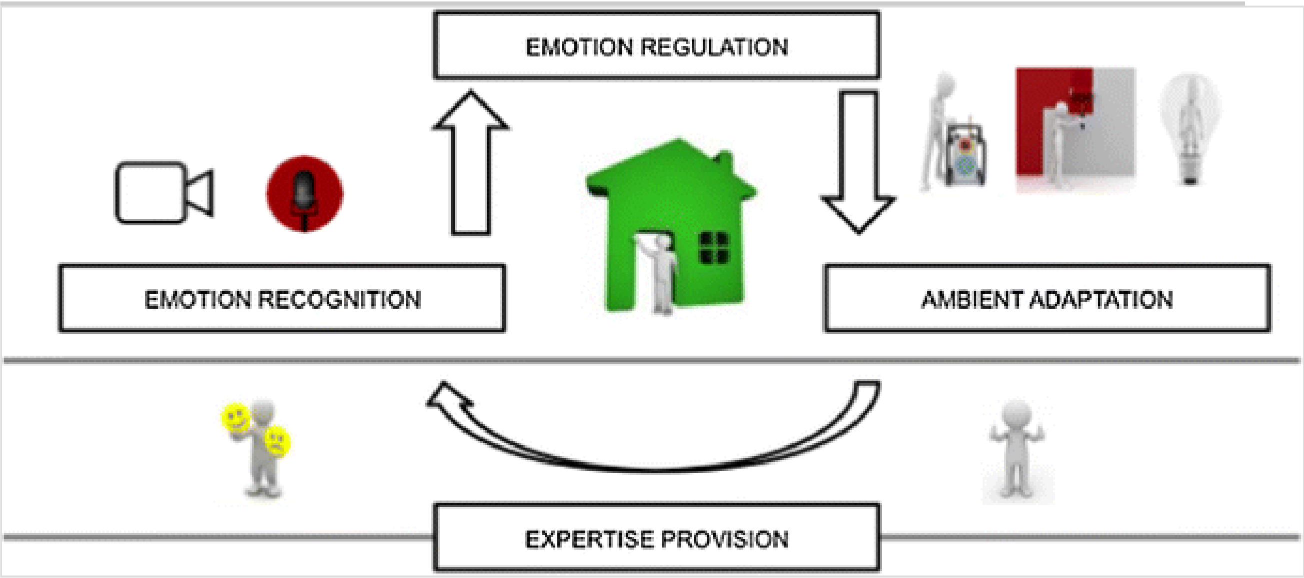
The most studied field is the induction of positive emotional states through music [22], which has been tested with some success in various mental health conditions such as depression [25]. Emotional responses to music are born of different musical features, including mode, harmonic complexity, tempo and intensity [26]. Although there are not many works with elderlies, we consider feasible to translate the principles of emotion regulation through music to ageing adults. One of the most used sets of stimuli in experimental research on emotions is the International Affective Picture System (IAPS) [23]. The IAPS is a collection of more than 1000 colour photographs that represent objects, people, landscapes and everyday situations. It allows accurate selection of the stimuli according to their position in the affective space defined by the dimensions of valence, arousal and dominance.

A New AmE Approach for Emotion Recognition and Regulation

In this paper, we propose a new system to improve the elder's QL&C living at their home. The system is based on emotion recognition and regulation considering the expertise of a multidisciplinary group of subject matter experts. The general overview of the proposed system is shown in Fig. 1 .

Fig. 1

General layout of the Smart Emotion Recognition and Regulation project



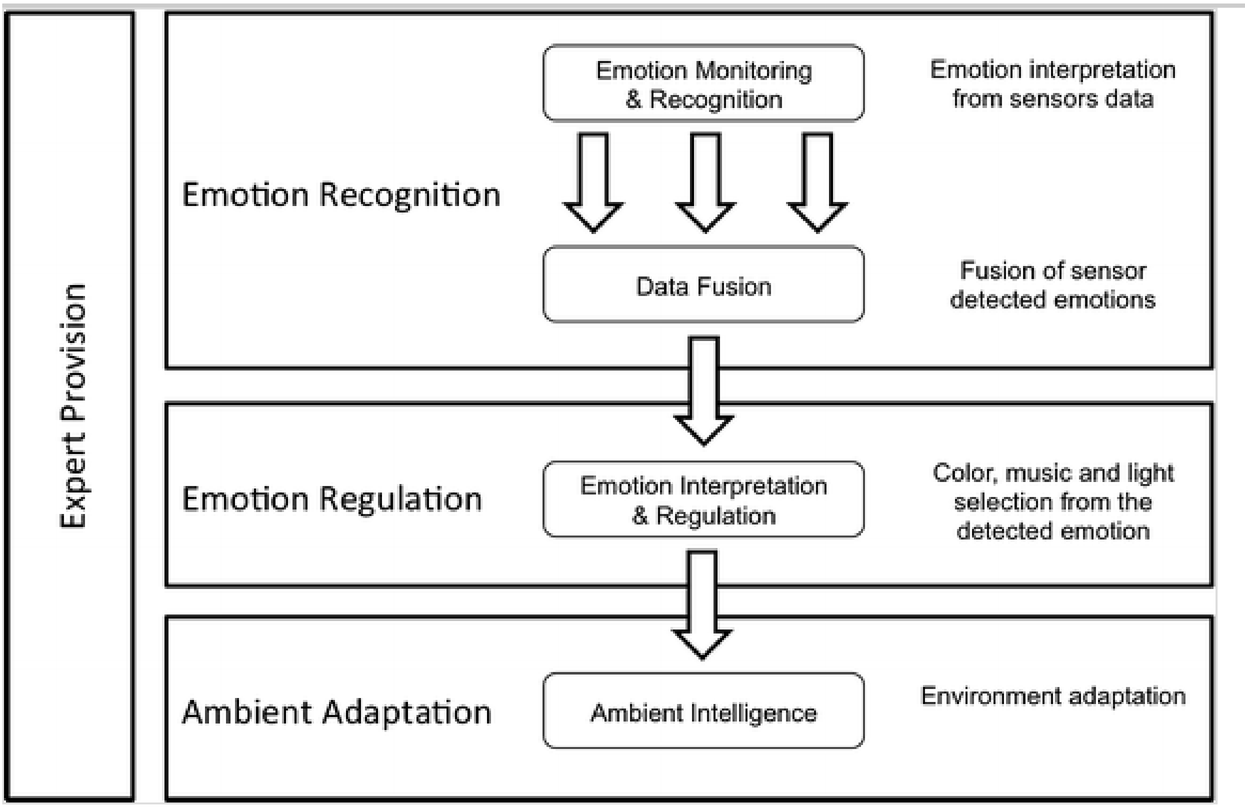
As depicted in Fig. 1 , the four main modules of the smart ambient are "Emotion Recognition", "Emotion Regulation", "Ambient Adaptation" and "Expertise Provision". Each one of these modules is in charge of performing different tasks.

Emotion Recognition

This module encompasses two main tasks: "Emotion Monitoring & Recognition" and "Data Fusion" (see Fig. 2). The former is devoted to information acquisition and processing from the different sensors placed in the environment to recognize elder's emotions: camera(s) to acquire the elder's face and gestures, camera(s) to track the elder's behaviour, microphone(s) for emotional speech recognition, and body sensors to acquire physiological signals. Reduced sensor awareness is crucial not only to increase elder acceptance but also to get unbiased emotion recognition. For this reason, our goal is to miniaturize the needed body sensors to turn them as less cumbersome as possible from the elder's perspective. With the information acquired from the sensors, face detection and gesture detection as well as speech recognition are crucial steps towards obtaining an emotion interpretation in terms of positive/negative mood. The activity recognition evaluates the user's behaviours as normal or abnormal, and the affect recognition is obtained from the analysis of the physiological data. Some technical details on the ongoing work have been recently presented [15, 28, 29].

Fig. 2

Main tasks of the system



The previous task generates a linguistic label related to an emotion for each sensory source, namely ambient information, face, gesture recognition, speech recognition, activity detection and physiological data. This single information sources might detect different emotions (even in contradiction with others). Thus, a higher-level task, "Data Fusion", is crucial to achieve a sensible output. Indeed, a big deal of work is performed in data fusion, as emotion detection and regulation are only viable if all the data recorded from all the sensors provide enough consensual evidence on the emotion detected. More information on the data fusion architecture is available in a recent article [43].

Emotion Regulation

The next module is "Emotion Regulation" (middle block of Fig. 2). Its input is the detected emotion as obtained from "Data Fusion", and its objective is to provide the best-suited conditions of music performance, as well as colour and light to attain the desired emotion in the elderly. Following the idea presented by Bachovik et al. [2], we propose to simultaneously use multiple induction techniques to improve the regulation of the elder's mood.

The task present in this module, "Emotion Interpretation and Regulation", is in charge of analysing the emotion recognized by the previous module. Experts play a fundamental role in this task, as it is necessary to have a precise insight of how stimuli affect emotions in order to decide the emotion-tailored music, colour and lighting conditions for each situation.

Here, the system automatically decides the proper reaction in response to the detected emotion. The knowledge that is used to make these kinds of decisions is provided by subject matter experts. Decision-making can vary from one user to other and also in accordance with the evolution of the elder. Then, it is important to provide an easy way to update the knowledge where the decisions are based on. Some initial results of the proposal have been published so far in relation to music [14] and colour and light [33, 44].

Ambient Adaptation

"Ambient Adaptation" uses the emotion-tailored music, colour and lighting conditions as an input to build up an AmI system that integrates the whole process (see bottom part of Fig. 2). "Ambient Intelligence" creates an intelligent system capable of adapting the ambience towards regulating emotions. Apart from interacting with the actuators in the environment, the "Ambient Intelligence" task is able to transmit alarms to physicians, care providers and relatives in case a potentially dangerous situation is detected.

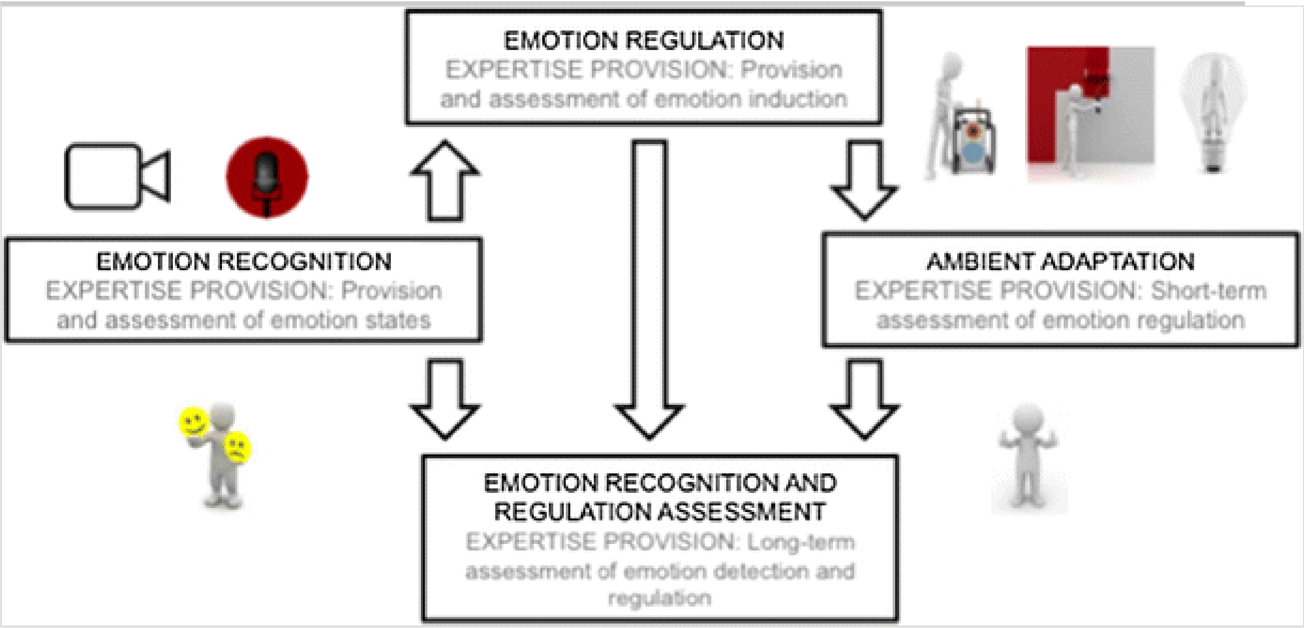
The implementation of this task involves controlling several domotic technologies (e.g. colour led, shutters or piped music). This task includes the development of software to control the different elements altering the ambient, also referred as actuators. Consequently, the "Ambient Adaptation" module works as a bridge to translate the commands generated from the "Emotion Regulation" to different technologies involved.

Expertise Provision

“Expertise Provision” is a module that incorporates all the roles that the subject matter experts play in the system (see Fig. 3). Subject matter experts are present in all modules and they play a transversal role providing the expert knowledge that is needed for the success of the system. Experts are involved in all the main tasks that are present in several parts of the system, and they differ depending on the module where the expert takes action.

Fig. 3

Expertise provision



1. “Emotion Recognition” module. Psychologists have a vast experience in assessment of emotional states (e.g. emotion or mood) based on physiological signals and the observation of the individual’s behaviour [20], [41], [46]. Experts have to transfer this knowledge to the engineers who translate it into algorithms. Therefore, the algorithms to evaluate the sensorial data for emotion recognition rely on the expert knowledge.

2. “Emotion Regulation” module. Physicians and psychologists decide how to act in the ambient to induce particular emotional reactions in the ageing adult. They provide the guidelines to adapt the environment based on the recognized emotional states.

3. “Ambient Adaptation”. After ambient adaptation takes effect, the

experts evaluate the immediate (short-term) elder's reaction, for example, assessing the elders' behaviour, while a relaxing music is played.

4. "Emotion Recognition and Regulation Assessment". The results of the system have to be evaluated in the long term. Experts will analyse historical data collected along the continuous operation of the system. They will ultimately decide the performance of the system based on the results of the analysis.

We are happy to count on a multidisciplinary team composed of computer scientists, electrical engineers and, as subject matter experts, psychologists, neurobiologists, physicians and musicians.

Software Architecture for Emotion Recognition and Regulation

Nowadays, there is a need for architectural solutions to develop complex software systems. In this sense, a set of software blocks is essential so that the programmers can use, extend or modify them to fit specific applications. This mechanism allows the decomposition of an application into a set of independent modules describing a set of interfaces in order to ease communications. Such flexible approaches enable the quick creation and deployment of applications. Under this scheme, components are defined as reusable and independent blocks that are combined with other components to build a specific system. From a developer perspective, building an application consists of assembling some existing components and a few of their own.

Among the number of advantages, the biggest one is that a modular architecture reduces the time in developing any software as it provides a set of blocks that users can directly use or improve. This means that users are abstracted from the lower level implementation and focus the effort on programming their required module. The main disadvantage of such approaches is the development complexity. For small projects, it is faster to code directly rather than to implement a whole architectural infrastructure. In contrast, for big projects, like the one proposed here, these solutions are an appropriate option.

Real-time emotion recognition and regulation involves a wide diversity of technologies. This results in a complex system that requires a powerful, flexible software/hardware framework. Designing a new solution presents crucial challenges:

- *Deciding the purpose of the system* This simple decision involves several issues related to the scope of the problem to be solved. For instance, a solution for web development has nothing to do with the one suggested in this paper, so the requirements of each one must be carefully analysed before taking further steps.

- *Assessing the levels* After having confirmed the purpose, it is necessary to establish a division to allow a structured development. Finding an optimal decomposition involves both a deep analysis of the problem and subject matter experts.

- *Determining the grain size* This problem is related to the levels one. At an early design stage, the functionalities included in each level have to be clearly defined at design time.

- *Avoiding being restrictive* Finally, the system must be adaptable to a growing problem, but not too wide to lose its purpose. This is an important trade-off that must also be considered at the design stage.

After considering these aspects, several works have presented solutions for emotion recognition. A system called Smart Sensor Integration considerably boosts the development of multimodal online emotion recognition systems. Furthermore, Gehrig and Ekenel [17] present a common architecture for real-time action unit detection and emotion recognition. For these tasks, they employ a local appearance-based face representation approach using discrete cosine transform. Santos et al. [39] come up with a proposal that performed real-time stress detection based only in two physiological signals: heart rate and galvanic skin response.

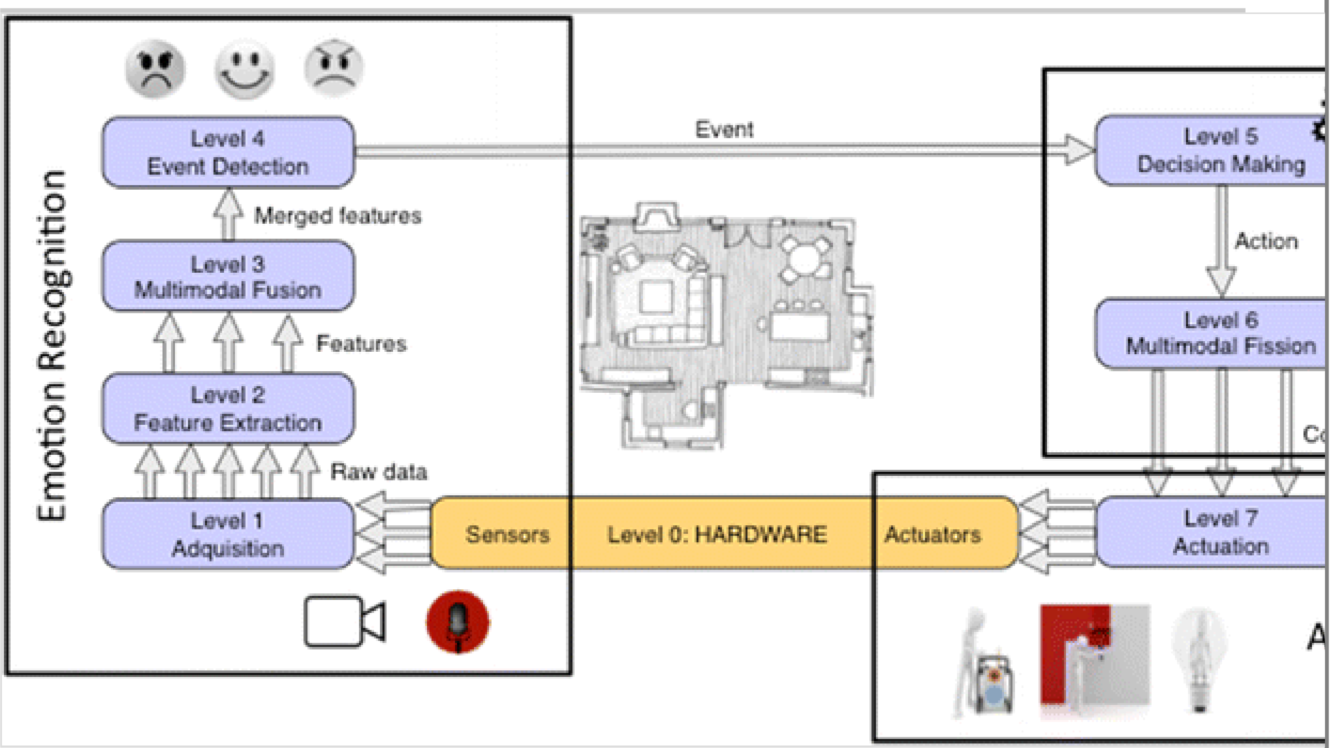
In relation to emotion regulation, the literature does not offer many software architectures. Emotion regulation is mostly inspired by works from the area of psychology. Psychologists have explored traditional frameworks (defined as a set of concepts, practices, and criteria to deal with a particular kind of problem) for diagnosis and treatment of deceases. An example is the work presented by Werner and Gross, which provides a framework for researchers and clinicians interested in understanding the role of emotion regulation processes in Psychopathology [49]. We will base on these works to provide a proper reaction of the ambience. Lastly, another paper provides a framework for researchers and clinicians interested in understanding the role of emotion regulation processes in Psychopathology [46].

Levels of the Software Architecture

In this section, we deal with the organization of the software modules involved in the proposed smart ambient system. We suggest a hierarchy of software levels intended to deal with the problems that commonly arise in real environments (hardware abstraction, noise in the data, or real-time processing). Next, the software architecture describing the functions performed at each level is presented. Each level is defined by the information that it consumes (receives), how these incoming data are transformed, and the resulting data it provides (sends) to the next level. The data flows are detailed for each level in the next subsections. Figure 4 shows the most outstanding elements to be considered in the proposed architecture for emotion recognition and regulation.

Fig. 4

Software Architecture for Emotion Recognition and Regulation



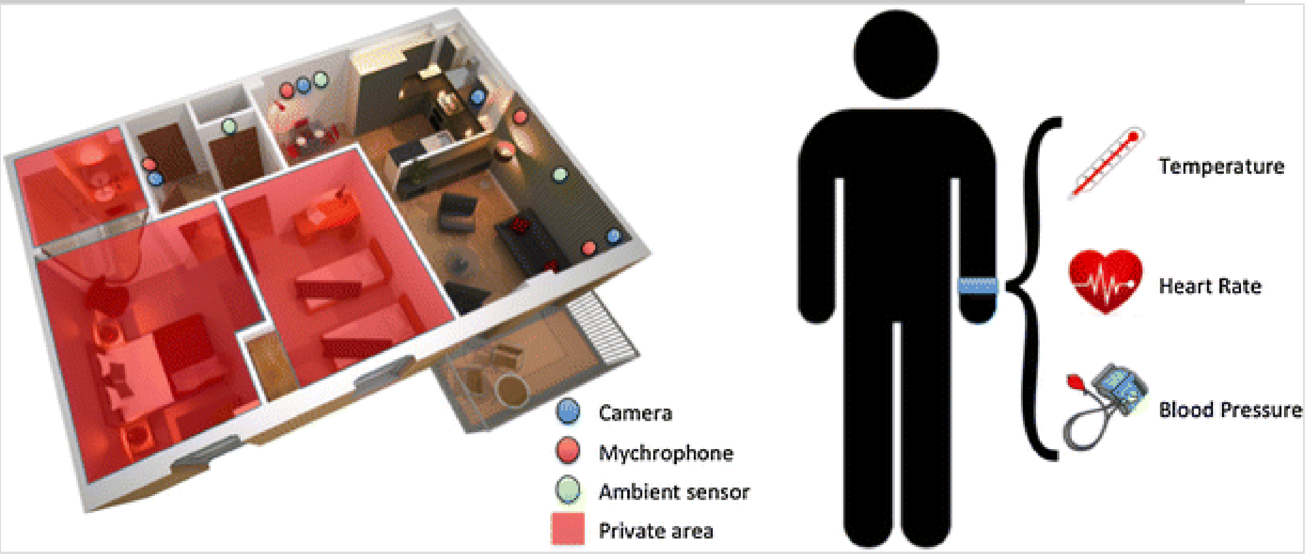
Level 0: Hardware—Sensors and Actuators

The proposed software architecture is composed of software elements located at different levels. However, when dealing with data coming from the real world, also referred to as the environment, the hardware elements need to be placed at the lowest level. In our case, this is Level 0. This layer represents the hardware components in charge of perceiving the environment and performing actions. No software module is considered here, although software from other levels directly interacts with it.

Sensors are devices that capture data from the environment and convert the information into electronic format. They can be roughly classified according to the information they provide, among others, visual data (RGB cameras, IR cameras or depth sensors), audio (microphones), biometric data (blood pressure, perspiration, breathing rate and so on) and ambient conditions (temperature, humidity, light). An example of a possible sensor deployment is shown in Fig. 5 .

Fig. 5

Conceptual description of the sensor deployment: *Left*, ambient sensors, the *red shaded* areas are considered as private (e.g. bathroom or bedroom); hence, no sensors are placed there. *Right*, body sensors, a wristband is a low intrusive option for sensor placement



AQ3

Actuators are devices able to alter the environment after receiving instructions and transforming them into actions. Some actuators considered in this project are as follows: projectors, light bulbs and sound systems.

Level 1: Acquisition

Software at this level handles the communication with the sensors. For each sensor, there is a module in this level that understands the data received. Since this is the lowest processing level, it does not consume data from other levels. It reads the low-level, raw data from the sensors and provides it to the next levels.

Level 2: Feature Extraction

This is one of the key levels. Here the raw data provided by the sensors are interpreted by different algorithms. This means that the input data are transformed into more informative, numerical features that summarize the raw data. Each feature represents an important characteristic of the data that is of special interest.

From a computer vision perspective, this level is in charge of detecting, recognizing and tracking the ageing adult in his/her home. In this sense, features may be understood as the track and shape of the elderly. Regarding speech recognition, some features coming from the elder's utterances can be extracted, such as the sentiment (positive, negative, neutral), the gender of the speaker, or the topics involved within the speech. So, this level consumes raw data and provides features to the upper level that will have to interpret them.

Level 3: Multimodal Fusion

A system operating in the real world should not rely completely on single or separate information sources due to problems such as noise, false positives or occlusions, and a fusion step is crucial in order to reduce the uncertainty associated with the extracted features. Indeed, feature extraction is performed over data coming from very different sources. In our scenario, we have information coming from cameras, wearable sensors sending biometric data, microphones and others. Different features are extracted for each kind of data. All these heterogeneous features are combined in order to provide higher-level abstraction of information. This is known as multimodal fusion.

In our particular problem, the recognition of the elder's emotion in the scenario is crucial. For this task, our architecture uses features related to the users' speech, their blood pressure, and their face. The fusion of all these features results on richer information at the time of assessing the elder’s emotion. This level consumes the features extracted in Level 2 and provides a combination of features in a time window. It is important to consider the time constrains in order to determine how often the fusion process is performed. For example, if it is detected that the elder is talking about a sad event, how long do we have to consider this feature? Therefore, the frequency of the fusion process will determine the lifespan of the features. The fusion process gives temporal coherence to the features.

Level 4: Event Detection

This level receives a set of features and processes them. The result is a high-level description of the situation (e.g. the elder is happy). In our problem, the emotion analysis logic is at this level. Its inputs are the array of features sent by the multimodal fusion. The output is the events related to the emotion of the elder. This means that the emotion recognition algorithm corresponds here. Moreover, our architecture is not limited to detect the ageing adult emotion. The algorithms intended to detect other events have to be implemented at this level, and the inputs would be the same.

Level 5: Decision-Making

Here is where the *intelligence* of the system lies: considering the events published by the Level 4 as well as other external factors such as the ambient information (e.g. temperature, humidity) provided by the Level 0, the system reacts and decides the action to execute. For example, considering that our system reacts to the elder’s mood, once this is detected, the system will vary the illumination, move the shutters, or change the music. This level consumes events and provides actions to be executed by the different actuators in the scenario.

Level 6: Multimodal Fission

This level defines the particular result of an action. A multimodal fission component accepts actions and splits them into several commands. Each command is directed to a different actuator in the environment. For instance, if the action is to create a relaxing environment, the multimodal fission will launch commands addressed to dim the lights, project warm colours and play pleasant music.

The fission process is tuned according to the user's likes. Hence, some information associated with the elderlies is welcome. Knowing some of their preferences, some environmental conditions, such as the background music and the projections, are selected accordingly. This results in a faster achievement of the desired effect. This level consumes actions from the decision-making level and provides commands to the proper actuators.

Level 7: Actuation

This level is the counterpart of Level 1. Here the modules communicating with the actuators send the proper commands. There is one module in charge of managing each actuator. The modules at this level consume commands published by the Multimodal Fission level. There is no information sent to other levels, but they write low-level instructions (or primitives) to hardware components. As an example, a text-to-speech module receives the command *say hello* that is transformed into low-level instructions that synthesize "hello" through speakers.

Discussion

The project presented in this paper aims to improve the elders' QL&C by leading them towards positive moods. Considering the current state-of-the-art, it is not possible to build a reliable system to quantify a person's QL&C. However, it is not possible to automatically assess the benefit that such an AmE system will cause in the elder's daily life. Consequently, we need subject matter experts to assess the performance of our system. These experts are in charge of providing the initial expert knowledge to the system. Moreover, they have to evaluate the elders' mood in order to adjust the system when the results are not satisfactory. So, the success of the project is highly dependent of the involvement and participation of subject matter experts (psychologists or physicians).

The emotion recognition and regulation system requires a fine-tuning of the parameters and configurations. That is, for example, the proper coloured lights or sounds that result in a positive mood can vary from one elder to other. Therefore, the system has to be adjusted for each senior according to his/her preferences. This implies a previous analysis of the participants to obtain as many details as possible. This results in the elaboration of individual profiles with personal and private data. These profiles have to be carefully handled according to the current regulation and ethical rules. Therefore, private data are not disseminated, and participants are thoroughly informed about how their data will be treated and used.

In line with the ethical aspects, monitoring a person in his/her home must be considered a critical issue. The use of heterogeneous sensors to monitor an elder generates large volumes of sensitive data (e.g. images or audio recordings). These sensitive data are considered as private data and are treated similarly to the personal profiles mentioned above. Moreover, some areas such as the bathroom or bedroom are considered as especially private, and therefore, the system would not be deployed there. The location where the system is installed is conceived and agreed with the participants, also taking into account that sensors are placed in strategic positions seeking good coverage of the monitored area. The devices to be installed will be selected considering their impact in the environment. Consequently, we aim at installing nonintrusive devices with a friendly appearance.

It is a complex matter to ensure high performance when detecting the emotional state of elder people from a system endowed with multiple information sources that acquire ambient data. Moreover, as the system is designed to be as less intrusive as possible, it is expected that single sensors will get high-quality information just during short periods of time. If we consider face detection, the algorithms will perform better when the user faces a camera at a determined range. For this reason, we propose the user of multisensory fusion and long-term analysis. Therefore, a person mood is not considered as an instantaneous phenomenon but the result of the analysis of the sensor information along a period of time combined with expert knowledge. We believe that this is the main strength of the system, constant monitoring and analysis to allow practitioners to assess the evolution of elder’s mood through long periods of time.

Besides, we believe that the success of the project is closely related to the attitude of the ageing adults towards the system. If elders feel uncomfortable at their own home, it will be almost impossible to induce a positive mood. The use of intrusive sensors and the deployment of noticeable sensors in the house do not help. As a result, the use of reduced body sensors and their proper placement in the house are major issues that should not be underestimated. Actually, we are studying the possibility of using body sensors that look like jewels or to integrate other sensors into the furniture or structure of the house.

Conclusion

This paper has described a project named "Improvement of the Elderly Quality of Life and Care through Smart Emotion Regulation". The objective of the project is to find solutions for improving the quality of life and care of the elderly who can or want to continue living at home by using emotion regulation techniques. For this sake, we have introduced the four main modules that make up the smart ambient system. These are "Emotion Recognition", "Emotion Regulation", "Ambient Adaptation" and "Expertise Provision". Heterogeneous vision and body sensors are used for monitoring the elderly and detecting his/her emotions, whilst music, colour and light are the simulating means to regulate the emotions towards a positive and pleasant mood.

This paper has also justified the need of a specific software architecture for emotion recognition and regulation tasks. Besides, we have introduced a new gerontechnological approach for monitoring the elderly at home. In first place, the goal is to detect the elder's emotions by analysing their physiological signals, facial expression, behaviour and voice. Then, the system provides the best-tailored reactions of the ambience to regulate the elder's emotions towards a positive emotion. The current state-of-the-art in emotion regulation through music, colour and light is used with the final goal of enhancing the quality of life of elder people living alone at their homes. After describing the benefits of our automated solutions for emotion recognition and regulation in the elderly, we have detailed the eight levels of the proposed architecture for implementing the smart ambient system.

AQ4

Acknowledgments

This work was partially supported by Spanish Ministerio de Economía y Competitividad/FEDER under TIN2013-47074-C2-1-R grant. José Carlos Castillo was partially supported by a grant from Iceland, Liechtenstein and Norway through the EEA Financial Mechanism, operated by Universidad Complutense de Madrid.

Compliance with Ethical Standards

Conflict of Interest José Carlos Castillo, Álvaro Castro-González, Antonio Fernández-Caballero, José Miguel Latorre, José Manuel Pastor, Alicia Fernández-Sotos and Miguel A. Salichs declare that they have no conflict of interest.

Informed Consent All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008 (5). Additional informed consent was obtained from all patients for which identifying information is included in this article.

Human and Animal Rights This article does not contain any studies with human or animal subjects performed by the any of the authors.

References

1. Bachorik JP, Bangert M, Loui P, Larke K, Berger J, Roew R, et al. Emotion in motion: investigating the time-course of emotional judgments of musical stimuli. Music Percept. 2009;26:355–64.

2. Bower GH. Mood and memory. Am Psychol. 1981;36:129–48.

3. Brewer D, Doughtie EB. Induction of mood and mood shift. J Clin Psychol. 1980;36:215–26.

4. Castillo JC, Fernández-Caballero A, Castro-González Á, Salichs MA, López MT. A framework for recognizing and regulating emotions in the elderly. Lect Notes Comput Sci. 2014;8868:320–7.

5. Chen M, Gonzalez S, Vasilakos A, Cao H, Leung VCM. Body area networks: a survey. Mob Netw Appl. 2011;16:171–93.

6. Devillers L, Lamel L. Emotion detection in a task-oriented dialogs. In: Proceedings of IEEE international conference on multimedia & expo, Baltimore, MD, USA; July 2003, vol 3, pp 549–552.

7. Ekman P, Friesen WV, Hager JC. Facial action coding system [e-book]. Salt Lake City: Research Nexus; 2002.

8. Elliot AJ, Maier MA. Color and psychological functioning. Curr Dir Psychol Sci. 2007;16:250–4.

9. Esposito A, Fortunati L, Lugano G. Modeling emotion, behavior and context in socially believable robots and ICT interfaces. Cogn Comput. 2014;6:623–7.

10. FaceSense. 2013. http://affect.media.mit.edu/projects.php?id=2523. Accessed 29 July 2015.

11. Fernández-Caballero A, López MT, Saiz-Valverde S. Dynamic stereoscopic selective visual attention (DSSVA): integrating motion and shape with depth in video segmentation. Expert Syst Appl. 2008;34:1394–402.

12. Fernández-Caballero A, Latorre JM, Pastor JM, Fernández-Sotos A. Improvement of the elderly quality of life and care through smart emotion regulation. Lect Notes Comput Sci. 2014;8868:348–55.

13. Fernández-Caballero A, Castillo JC, López MT, Serrano-Cuerda J, Sokolova MV. INT3-Horus framework for multispectrum activity interpretation in intelligent environments. Expert Syst Appl. 2013;40:6715–27.

14. Fernández-Sotos A, Fernández-Caballero A, Latorre JM. Elicitation of emotions through music: the influence of note value. Artif Comput Biol Med 2015:488–497.

AQ5

15. García-Rodríguez C, Martínez-Tomás R, Cuadra-Troncoso JM, Rincón M, Fernández-Caballero A. A simulation tool for monitoring elderly who suffer from disorientation in a smart home. Expert Syst. 2015;32:676–87.

16. Gascueña JM, Fernández-Caballero A. On the use of agent technology in intelligent, multi-sensory and distributed surveillance. Knowl Eng Rev. 2011;26:191–208.

17. Gehrig T, Ekenel HK. A common framework for real-time emotion

recognition and facial action unit detection. In: Proceedings of IEEE computer society conference on computer vision and pattern recognition workshops, Colorado Springs, USA, June 2011; pp. 1–6.

18. Gross JJ, Barrett LF. Emotion generation and emotion regulation: one or two depends on your point of view. Emot Rev. 2011;3:8–16.

19. Gross JJ, Levenson RW. Emotion elicitation using films. Cogn Emot. 1995;9:87–108.

20. Jovanov E, Milenkovic A. Body area networks for ubiquitous healthcare applications: opportunities and challenges. J Med Syst. 2011;35:1245–54.

21. Kim JB, ParkJS OhYW. Speaker-characterized emotion recognition using online and iterative speaker adaptation. Cogn Comput. 2012;4:398–408.

22. Kim J, André E. Emotion recognition based on physiological changes in music listening. IEEE Trans Pattern Anal Mach Intell. 2008;30:2067–83.

23. Lang PJ, Bradley MM, Cuthbert BN. International affective picture system (IAPS): affective ratings of pictures and instruction manual. Technical report A–8, University of Florida, USA, 2009.

24. Lee H, Choi YS, Lee S, Park IP. Towards unobtrusive emotion recognition for affective social communication. In: Proceedings of IEEE consumer communications and networking conference, Las Vegas, NV, USA, January 2012; pp. 260–264.

25. Livingstone SR, Palmer C, Schubert E. Emotion response to musical repetition. Emotion. 2012;12:552–67.

26. Livingstone RS, Thompson WF. The emergence of music from the theory of mind. Musicae Scientiae. 2009;17:83–115.

27. López JM, Gil R, García R, Cearreta I, Garay N. Towards an ontology for describing emotions. Lect Notes Comput Sci.

2008;5288:96–104.

28. Lozano-Monasor E, López MT, Fernández-Caballero A, Vigo-Bustos F. Facial expression recognition from webcam based on active shape models and support vector machines. Ambient Assist Living Daily Act 2014:147–154.

29. Martínez-Rodrigo A, Zangróniz R, Pastor JM, Latorre JM, Fernández-Caballero A. Emotion detection in ageing adults from physiological sensors. Ambient Intell Softw Appl 2015:253–261.

30. McRae K, Misra S, Prasad AK, Pereira SC, Gross JJ. Bottom-up and top-down emotion generation: implications for emotion regulation. Soc Cogn Affect Neurosci. 2012;7:253–62.

31. Medjahed H, Istrate D, Boudy J, Baldinger JL. A pervasive multi-sensor data fusion for smart home healthcare monitoring. In: Proceedings of IEEE international conference in fuzzy systems, Taipei, Taiwan, June 2011; pp. 1466–1473.

32. Niedenthal PM, Halberstadt JB, Setterlund MC. Being happy and seing happy: emotional state mediates visual word recognition. Cogn Emot. 1997;11:403–32.

33. Ortiz-García-Cervigón V, Sokolova MV, García-Muñoz RM, Fernández-Caballero A. Led strips for color- and illumination-based emotion regulation at home. Ambient assisted living. Development and testing of ICT-based solutions in real life situations 2015 (in press).
AQ6

34. Pail G, Huf W, Pjrek E, Winkler D, Willeit M, Praschak-Rieder N, et al. Bright-light therapy in the treatment of mood disorders. Neuropsychobiology. 2011;64:152–62.

35. Pavón J, Gómez-Sanz JJ, Fernández-Caballero A, Valencia-Jiménez JJ. Development of intelligent multi-sensor surveillance systems with agents. Robot Auton Syst. 2008;55:892–903.

36. Picard RW, Vyzas E, Healey J. Toward machine emotional

intelligence: analysis of affective physiological state. IEEE Trans Patterns Anal Mach Intell. 2001;23:1175–91.

37. Sadri F. Ambient intelligence: a survey. ACM Comput Surv (CSUR). 2011;43(4):36.

38. Salovey P, Rothman AJ, Detweiler JB, Steward WT. Emotional states and physical health. Am Psychol. 2000;55:110–21.

39. de Santos A, Sánchez C, Guerra J, Bailador G. A stress-detection system based on physiological signals and fuzzy logic. IEEE Trans Industr Electron. 2011;58:4857–65.

40. Schachter S, Singer JE. Cognitive, social and physiological determinants of emotional state. Psychol Rev. 1962;69:379–99.

41. Scherer K. What are emotions? And how can they be measured? Soc Sci Inf. 2005;44:695–729.

42. Silva PRD, Osano M, Marasinghe A, Madurapperuma AP. Towards recognizing emotion with affective dimensions through body gestures. In: Proceedings of 7th IEEE international conference on automatic face and gesture recognition, Southampton, United Kingdom, April 2006; pp. 269–274.

43. Sokolova MV, Fernández-Caballero A, López MT, Martínez-Rodrigo A, Zangróniz R, Pastor JM. A distributed architecture for multimodal emotion identification. Trends Pract Appl Agents Multi Agent Syst Sustain 2015:125–132.

44. Sokolova MV, Fernández-Caballero A, Ros L, Fernández-Aguilar L, Latorre JM. Experimentation on emotion regulation with single-colored images. Ambient assisted living. Development and testing of ICT-based solutions in real life situations 2015 (in press).

45. Soleymani M, Lichtenauer J, Pun T, Pantic M. A multi-modal affective database for affect recognition and implicit tagging. IEEE Trans Affect Comput. 2012;3:42–55.

46. Susskind JM, Littlewort G, Bartlett MS, Movellan J, Anderson AK. Human and computer recognition of facial expressions of emotion. Neuropsychologia. 2007;45:152–62.

47. Velten E. A laboratory task for induction of mood states. Behav Res Ther. 1968;6:473–82.

48. Vogt T, André E, Wagner J. Automatic recognition of emotions from speech: a review of the literature and recommendations for practical realisation. Lect Notes Comput Sci. 2008;4868:75–91.

49. Wagner J, André E, Jung F. Smart sensor integration: a framework for multimodal emotion recognition in real-time. In: Proceedings of 3rd international conference on affective computing and intelligent interaction and workshops, Amsterdam, The Netherlands, September 2009; pp. 1–8.

50. Zhou J, Yu C, Riekki J, Kärkkäinen E. AmE framework: a model for emotion aware ambient intelligence. In: Proceedings of the second international conference on affective computing and intelligent interaction (ACII2007): Doctoral Consortium, September 2007.