Thrilled to Bits: A Review of Emotional Intelligence in AI

Author
Tim van de Wiel
5972450

Bachelor's Thesis Artificial Intelligence
Utrecht University
7,5 ECTS

Supervisor
dr. B. M. Harvey

Second Reader
dr. D. Paperno

26th of June, 2019

“The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without any emotions.”

- Marvin Minsky, 1988

Abstract

Research on emotion technology in artificial intelligence is taking off, but there has been less attention on the possible applications of emotions and emotional intelligence in artificially intelligent systems. Reviewing contemporary literature, this thesis provides a comprehensive overview of emotion in humans, affective computing and modern emotion technologies. Additionally, recent literature is synthesized to provide key areas that could and should benefit from emotionally intelligent systems. It is concluded that especially in healthcare, education and customer service, future emotionally intelligent systems may thrive.

Keywords: emotion, emotional intelligence, emotion recognition, emotion expression, affective computing

Contents

Introduction

Chapter 1: Emotions in humans
1.1: Defining emotions
1.1.1 Feelings and moods
1.1.2 Emotion classification
1.2: Emotion and human behavior
1.2.1 Facial expression, physiology and vocal properties
1.2.2 Learning and memory
1.3: Emotional intelligence
1.4: Conclusion

Chapter 2: Contemporary emotion technology in AI
2.1: Affective computing
2.1.1 Recognizing and expressing emotions
2.1.2 Experiencing emotions
2.1.3 Emotional intelligence
2.2: Recognition of emotions
2.2.1 Speech
2.2.2 Facial expressions
2.2.3 Text
2.2.4 Other
2.3: Expression of emotions
2.3.1 Speech
2.3.2 Facial expressions
2.3.3 Non-human expressions
2.3.4 Other
2.4: Conclusion

Chapter 3: Applications of emotionally intelligent AI
3.1: Online
3.1.1 Chatbots
3.1.2 Education
3.2: Social Robots
3.2.1 Healthcare
3.2.2 Customer service
3.3: Conclusion

Conclusion

Introduction

Artificial intelligence is taking its place in the daily lives of most people. Be it through giving voice commands to Siri, receiving a tour from a robotic museum guide or being assisted by online chatbots, people interact with artificial intelligence regularly. Even when this is not made explicit, people often realize in their communicative efforts that they are not dealing with a human. This is because artificially intelligent systems lack one important attribute: emotion.

Multiple scholars recently noted the severe deficiency of research into how emotions should and could be implemented in intelligent systems (Fan et al, 2017; Franzoni et al., 2019). Even though emotion recognition and expression techniques seem to become more impressive and complex every day, little research has been done into how emotion can be applied in AI. This thesis aims to provide an overview on contemporary emotion technologies and their applications. It will answer the following question:

"To what extent can emotional intelligence be incorporated in artificially intelligent systems?"

Because a part of the field of artificial intelligence aims to mimic human cognition, it is important that there is a solid understanding of how the brain works. Therefore, this thesis will start off with trying to define human emotion in chapter 1. Also discussed is how emotions can affect human behavior and how human emotional intelligence is defined. Since human emotions are a very broad topic, the chapter will only concentrate on aspects that are relevant in the following chapters. Secondly, chapter 2 will cover contemporary theories, technologies and methods that encompass emotion in artificial intelligent system. It reviews the modern conventions in affective computing, and reports on relevant technologies in artificial intelligence, including but not limited to emotional speech recognition and facial expression generation. Lastly, chapter 3 will concisely address key areas that could benefit from emotional intelligent systems.

Chapter 1. Emotion in humans

To be able to discuss emotion and emotional intelligence in machines, it has to be clear what those terms mean. This first chapter aims to provide a short overview of emotion theory, emotion influence on human behavior and the definition of emotional intelligence in humans.

1.1 Defining emotions

Ever since the very first scientific attempt to define emotion by William James in 1884, countless arguments have been held on what is arguably one of the longest-lasting discussions in the scientific world (Akgün et al., 2010; Mulligan & Scherer, 2012). Carroll Izard (2010) carried out an experiment in which she interviewed thirty-five notable scientists from varying disciplines related to emotion research, and asked them to give their definition of emotion. The scientists were not only reluctant to define emotion, but their wildly different responses also showed that defining emotion is terribly complicated and open for different interpretations.

There are however some properties of emotion that are widely recognized. Emotion is an individual's reaction to events in one's environment or to their own actions. Emotion is dynamic; it evolves or devolves over time, but it does have a beginning and an end. Emotion consists of at least an individual subjective experience and a physical response, be it through a facial expression or an elevated heart rate. Related to this, emotion may prompt certain behavior (Mulligan & Scherer, 2012; Izard, 2010; Akgün et al, 2010; Adolphs & Anderson, 2018). The physical response could be further divided into a motor expression component and a neurophysiological component, just as the subjective experience could be divided into a motivational component and the actual subjective feeling (Scherer, 1984).

1.1.1 Feelings and moods

In everyday language, emotions, moods and feelings are often used interchangeably. However, scientifically, they are wholly different terms. Friedenberg & Silverman (2015) define feelings to be "the subjective experience of emotions" (p. 302). A feeling is what someone actually feels inside when experiencing an emotion. Others can deduce the emotion someone is going through based on looks and expressions, but will not be able to deduce feelings since they are entirely personal and subjective.

Moods are, very simply put, stretched out emotions. They are much longer in duration than emotions, but not as intense as emotions can be. Friedenberg & Silverman (2015) introduce the example of a student receiving a good grade on a test; it will put him in a 'good mood' for the rest of the day, while the actual emotion of happiness would probably only last a few minutes.

1.1.2 Emotion classification

The two leading theories on how to classify emotions are the basic emotion theory (BET), introduced by Paul Ekman in 1992, and the theory that emotions can be classified in terms of two or more dimensions (Cowen & Keltner, 2018; Mustafa et al., 2018; Dubois & Adolphs, 2015).

Ekman's now famous paper from 1992 introduced six basic emotions: anger, fear, sadness, enjoyment, disgust and surprise. To define these six cornerstones of human emotion, Ekman listed nine characteristics that distinguish them from each other and from other affective states, such as feelings and moods. For instance, basic emotions must have a quick onset, should have a distinctive physiology and also should have distinctive universal signals that can be understood in any culture. Five other emotions (contempt, shame, guilt, embarrassment and awe) were put forward at the time as possible candidates to be classified as basic emotions. Recently however, Keltner and colleagues (2019) synthesized academic work on BET and composed a list of twenty-four distinct emotional states, all modalities considered. This list includes emotions such as gratitude, relief, sympathy and triumph.

The other leading theory is one introduced by Russell (1980). Drawing on the 'emotion wheel' of Plutchik (1980), Russell proposes that all emotions can be defined along two dimensions, and can therefore be placed in a two-dimensional space (something he later came to call the Affect Grid (Russell et al., 1989)): that of pleasure-displeasure, or valence, and of high arousal-low arousal (Figure 1). They also suggest that these two concepts are never seen as independent, but instead form one emotion, just like hue and saturation form one color. For example, anger is high-arousal and low-valence, while enjoyment is high-arousal and high-valence.

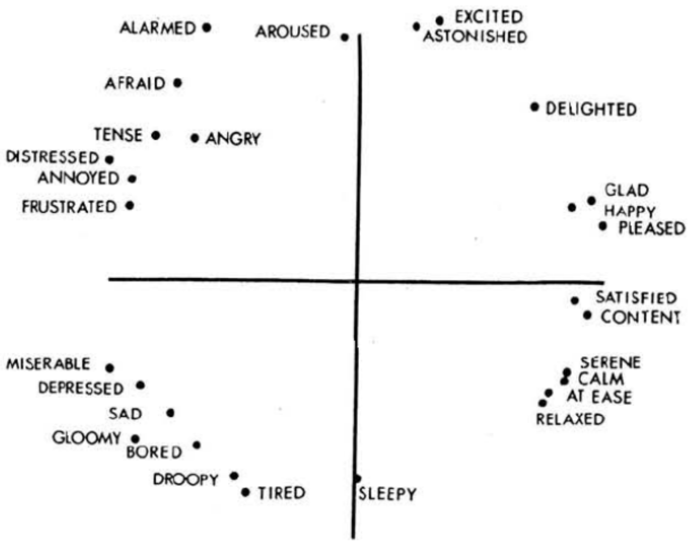

Figure 1. The arousal-valence model (adapted from Russell, 1980)

1.2 Emotion and human behavior

1.2.1 Facial expression, physiology and vocal properties

It falls outside the scope of this thesis to pick and defend a side in the debate on the categorization of emotions. However, the recognition and expression of different emotions is a substantial part of what is to come in this thesis. The facial, physiological and vocal properties of emotions are arguably the most defining and definitely the most used properties in emotion recognition and expression (Vinciarelli & Mohammadi, 2010; Mustafa, 2018). Therefore, some of the more defining facial, physiological and vocal properties of the six original basic emotions are presented in Table 1 below.

While physiological properties are without a doubt excellent indicators of certain emotional states, they can and will rarely be used. This is because measuring them can only be done in ways that are fairly intrusive to the user in question, which is often not practical in real-world scenarios. Nevertheless, to provide a full overview, they are included in Table 1.

Emotion	**Facial expression** (Keltner & Cordaro, 2015; Brave & Nass, 2002)	**Physiological signs** (Shu et al., 2018; Wibawa & Purnomo, 2016; Gilchrist et al., 2016; Siedlacka & Denson, 2019)	**Vocal properties** (Bachorowski, 1999; Scherer et al., 2003; Chong et al., 2018; Liu & Xu, 2016; Murray & Arnott, 1993)
Anger	Eyebrows curled down, eyes wide, hard stare, lips close together, dilated nostrils	Increased HR, increased SCR, increased RR	Higher fundamental frequency, higher intensity, slightly faster talking, tense articulation
Disgust	Eyes narrowed, lips separated, nose wrinkled, raised cheeks	Decreased HR, slightly increased RR	Lower frequency of first formant, slower talking, lower pitch average
Enjoyment (happiness)	Duchenne smile, eyes relaxed, raised cheeks	Decreased CO, increased RR	Higher fundamental frequency, higher intensity, slightly faster talking
Fear	Eyebrows and upper eyelid raised, lips separated and stretched, forehead wrinkled	Increased HR, decreased FT, increased SCR	Higher fundamental frequency, faster talking, higher pitch average
Sadness	Eyebrows up, lip corners down, lower lip raised	Lower level of oxygen in blood	Lower fundamental frequency, slower talking, lower intensity, slurring articulation
Surprise	Eyebrows and upper eyelid raised, mouth opened, jaw dropped, pupils dilated	Increased SCR	Raised pitch

Table 1. Ekman's six basic emotions along with some of their physical and physiological properties. HR = heart rate, SCR = skin conductance response, FT = finger temperature, CO = cardio output, RR = respiration rate

1.2.2 Learning and memory

Emotions can have a considerable impact on memory and, as a result, on learning as well. Several studies stress the existence of cognitive processes involved in memory which are enhanced by emotional arousal. People also pay more attention to stimuli that are emotionally loaded than neutral stimuli, thus consolidating them in the long-term memory (Kaplan et al., 2015; Tyng et al., 2017; Earles et al., 2015). This is not limited to episodic memory of events in the past, as Hostler et al., (2018) points out: prospective memory performance is also enhanced when presented with positively rated emotions. Gluck et al. (2016) also describe an experiment in which a group that was emotionally aroused during the experiment were better in remembering the events of a video shown during the experiment than a group who were not. This is however not the only reason that emotional events are often easily stored in and retrieved from memory. People think and talk more often about highly emotional events, probably because there is a high chance they are life-defining moments (e.g. hearing about the loss of a relative). By doing this, the original memory is strengthened and more easily remembered (Kaplan et al., 2015; Gluck et al., 2016).

Another effect of emotion on human memory is the so-called *mood congruency of memory* effect. People in a certain mood or emotional state, will find it easier to retrieve memories that match that same mood or emotional state (e.g. a sad person will remember something sad). This is especially apparent in clinically depressed patients, who will - already being in a sad mood - recall even more unhappy memories, which will lead to even more sadness, etcetera (Gluck et al., 2016). A closely related effect is that of *mood-dependent memory.* Here, the recall of memories will be better when the mood at recall matches the mood at the time of learning (e.g. a sad student will remember the material he studied when he was sad). The difference between the two effects is minimal, but both further underline the importance of emotion in learning (Friedenberg & Silverman, 2015; Lewis & Critchley, 2003).

1.3 Emotional intelligence

When interacting with other people, we often deal with the emotions of others. Being able to see the world through someone else's eyes is called having empathy. It is the "ability to imagine what someone else is thinking and feeling in a given situation" (Ioannidu & Konstantikaki, 2008, p. 2). Emotional intelligence is a step ahead of having empathy, and comprises of the ability to not only recognize and emphasize with the emotions of others, but also to understand and manage them. It is also about personal emotions - being able to cope with them and regulate them (Roth et al., 2018; Mattingly & Kraiger, 2019). Miners et al. (2018) add to that the "ability to reason accurately about emotions, and to use emotions and emotional knowledge to enhance thought" (p. 87). Having high emotional intelligence often also means an increase in the quality of social relationships (Mayer, 2004; Miners et al., 2018).

The above however is an example of one of two models of emotional intelligence. The ability-based model, as described above, states emotional intelligence as being a combination of abilities *and* behaviors that help someone in recognizing and managing their own and others' emotions. (Mattingly & Kraiger, 2019). The other side of the debate is called the 'mixed-model', which suggests that emotional intelligence is less of an ability, but consists of several competencies that together will help in successfully coping with different situations, such as emotionally dealing with others. Ioannidu & Konstantikaki (2008) composed a list these competencies, which includes being able to control own emotions and recognizing others' emotions.

However, Mattingly and Kraiger (2019) rightfully point out that this mixed-model approach is often criticized for being too broad and hard to validate. The argument against the ability-based model is somewhat similar, because it remains unclear whether emotional intelligence can actually be defined as just a set of abilities. Despite these criticisms, a better model is yet to be introduced, and both models provide useful quantifiable measures that could be implemented in computational systems.

1.4 Conclusion

Even though the concept of emotion cannot be defined in a single sentence or paragraph, its properties, symptoms and effects certainly can. The defining facial expressions, vocal properties and physiological signs of emotions are useful for recognizing and expressing emotions, something that will be discussed in the coming chapter. The effects of emotions on learning and memory could prove to be of value when examining possible applications of emotional intelligent systems. Empathy and emotional intelligence are more easily defined than emotion itself. Despite no agreement on what model of emotional intelligence is the most suitable one, both models lend themselves to be applied in the computing field. The next chapter explores whether this is actually possible.

Chapter 2. Contemporary emotion technology in AI

This second chapter aims to provide an overview of contemporary ideas and methods regarding emotion (technology) in AI. It will cover an introduction to and modern uses of affective computing, technologies for the recognition of emotions and the expression of emotions.

2.1 Affective computing

Affective computing, introduced by Rosalind Picard, is about "computing that relates to, arises from, or deliberately influences emotion or other affective phenomena" (Picard, 1999, p. 1). As Pestana et al. (2018) notes, the essence of affective computing is the recognition of and subsequent response to emotional interactions with a user. This can be achieved by measuring emotional signals from the user, such as the vocal and facial expressions dealt with in the first chapter of this thesis, and then adequately responding to those signals - be it by simply changing the interface in some way, or displaying an affective response.

2.1.1 Recognizing and expressing emotions

Picard (2000, p. 55) provides a list of criteria for a computer, or any computational device, to be able to recognize emotion. It comprises of six elements:

- input: it can receive a variety of signals, such as facial expressions or speech;
- pattern recognition: it can perform feature extraction and classification of those signals;
- reasoning: it can 'reason' about the input emotion based on knowledge of emotions and the pattern recognition results;
- learning: it can learn about the user and the users' emotional expressions;
- bias: its decisions can be influenced by its own emotional state;
- output: it can provide a result, which is naming the recognized emotion.

Obviously, some of these criteria are more important than others. Input and output are crucial for any kind of computing, while learning and bias are not essential for any system but are certainly present in any artificial intelligence.

Picard (2000, p. 60) also provides a list of criteria for the expression of emotions, which is equally important in affective computing:

- input: it can receive external or internal instructions, stating what emotion(s) to express;
- intentional and spontaneous pathways: it should be able to deliberately express an emotion, but it should also be able to let its outputs be influenced by an internal emotion;
- feedback: its emotion expression can affect the emotion, so a feedback loop must exist;
- bias-exclusion: by expressing a certain emotion, other emotions are less likely to be expressed;
- social display rules: the way it expresses its emotions should be defined by the relevant and present social norms;
- output: it should be able to express an emotion through visible or audible signs.

2.1.2 Experiencing emotions

The criteria for a system to recognize and express emotions might be clear, but what does such a device need to *have* emotion? Is it even possible for a machine, built out of technological parts, to experience feelings and moods? The following section aims to describe what it means for computers to truly have emotions.

Adapting a list from Picard (2000), who remains the leading figure in the field of affective computing, a computational device truly has emotion if it satisfies five conditions. The first of these conditions is that the device (or system) should have behavior that results, or at least appears to result, from the emotions that that system is currently experiencing. Secondly, the system should be able to give quick emotional responses to certain inputs. Related to this is the third condition: a system should be able to generate its emotions through their own (cognitive) reasoning about the situation, specifically about the relation to its “goals, standards, preferences and expectations” (Picard, 2000, p. 70). Through this reasoning, the system can infer that it has to express a certain emotion.

The fourth condition might be the hardest to meet, and is closest related to the many philosophical debates on the subject over the last few decades. Picard states that a system can only have emotions, when it can have emotional experiences. This means having cognitive awareness of its emotions, physiological awareness and subjective feelings. Cognitive awareness of its emotions can be as primitive as ‘knowing’ what emotional state it is in, and while physiological awareness is obviously unlike that of humans because of the difference in physiology, this can also be expressed in terms of awareness of inner temperature or voltage. Subjective feelings however cannot be that easily defined. They might not be like human subjective feelings, as it cannot be expected from a machine to feel what humans feel, since the aforementioned difference in physiology and cognition causes machines to have emotional experiences completely unlike what is known to humans.

Last but not least, the fifth condition states that the emotional system must be able to interact with other internal processes, just like humans do. As discussed in the first chapter of this thesis, emotions in humans can affect processes such as memory, learning and attention. This should be present in emotional systems as well.

However, Huang & Rust (2018) raise a very valid point against the necessity of these five conditions. As long as the system can *appear* to truly have emotions, does it really matter if it actually can have emotions? They draw the comparison to the Turing test, where a machine is considered to be intelligent if a human cannot tell the difference between the responses of the machine and another human. This seriously increases the chances of future systems to be classified as ‘having emotions’.

It is important to mention that the Turing test has never been passed thus far.

2.1.3 Emotional intelligence

As was mentioned in chapter 1, there are two main ways to define emotional intelligence: ability-based and the mixed-model approach. Both provide excellent criteria to determine whether or not any (artificial intelligent) system is also emotionally intelligent. The system at least needs to be able to recognize, express and (appear to) experience emotions. The obvious caveat to both approaches this is the previously mentioned difference between the subjective emotional experiences in humans and machines, in that those experiences cannot possibly be the same due to the fundamental differences between humans and machines. As a result, it cannot be expected from machines to understand human emotions the same way humans understand each other. As Picard (p. 80) phrases it, “the best empathy or understanding we can hope for is at the level of an outsider, who tries to understand, but who has never actually been in our shoes”.

Besides, recognizing, expressing and experiencing emotions alone does not make for a very high emotional intelligence. It also involves being able to reason about emotions, be it others’ emotions or personal emotions, and then using that reasoning to readjust to the situation and to guide further action. However, Fan et al. (2017) conducted an experiment in which they showed that there was no actual difference in the human perception of human and robotic agents in different social interactions, further solidifying the point Huang & Rust (2018) made about an emotional system only needing to appear emotional. It does not matter if the system actually extensively reasons about the world around it, as long as it appears to do so. It should be noted that this experiment only concerns human-robot interaction, not human-computer interaction as a whole.

From this it can be concluded that despite inevitable differences in the definition of emotional intelligence in humans and machines, human criteria can be used to gauge the emotional intelligence of machines. In the next part, modern emotion recognition and expression technologies will be covered.

2.2 Recognition of emotions

The first step in emotional intelligence is, as said before, being able to recognize emotions. In chapter 1, three important expressional properties of emotions were laid out: physiological signs, facial expressions and vocal properties. The following section will first address speech and facial recognition, after which it will shortly cover some other emotion recognition methods. Because of the limited scope of this thesis, it cannot and will not examine those technologies and their inner workings in detail, but rather provide a comprehensive overview.

2.2.1 Speech

Speech detection and recognition is nonintrusive, and numerous features of speech are associated with emotional states (also see Table 1). Some even call the recognition of emotion through speech the most effective way to recognize emotions (Mustafa et al., 2018). The features often used in emotional speech recognition can be divided into two groups: prosodic features and spectral features. Both of these groups belong to the category of ‘low-level descriptors’ (LLDs), as opposed to the category of functional features, which includes statistical derivatives of a speech signal such as the mean signal (Ali et al., 2017; Mustafa et al., 2018).

Prosodic features are the intonational and rhythmic features of speech, and are often considered to be the most useful features when it comes to emotional speech (Jurafsky & Martin, 2009; Li et al., 2015; Tahon et al., 2018). This includes features such as the intensity, duration, pitch, voice quality and the fundamental frequency of speech. Spectral features often provide complementary information to the prosodic features, as they contain the frequency contents of the speech signal (Wu et al., 2010). Because spectral features are often acquired over a very short period of time (usually 20-30 ms), they have to be used together with other speech features to be able to extract meaningful behavioral information from the speech signal (Ali et al., 2017; Wu et al., 2010).

These are clearly a lot of features, not all of which can be taken into account by an emotion recognizing system. It is therefore important to choose the right features for such a system, as this can influence the effectiveness and quality of recognition (Mustafa et al., 2018). After deciding on what features to use, and the actual extraction of the features, the machine should be able to use those extracted features to determine the emotional state expressed in the speech signal. Ali et al. (2017) identified the four most used classifiers in contemporary literature: (1) Quadratic Discriminant Analysis (QDA), (2) k-Nearest Neighbour (KNN), (3) Support Vector Machines (SVMs) and (4) neural networks.

Nassif et al. (2019) showed that using the latter, (deep) neural networks, can provide better results than most previously used methods. They can work with a large number of inputs, and can be either self-trained through a lot of training data, or it can be trained by humans (Ali et al., 2017). Even though they work better when confronted with isolated speech signals (such as single words or short sentences) instead of the more likely input of complicated sentences, neural networks can be used combined with another technology to make emotional speech recognition even more accurate. Note that a conventional emotional speech recognition system only processes the features of the sound of speech, and does not bother with the linguistic information embedded in speech (Mustafa et al., 2018). However, because certain linguistic information can correlate with certain features of speech (e.g. curse words are almost always angrily pronounced), neural networks can unintentionally learn to link specific words to emotional value anyway.

2.2.2 Facial expressions

The recognition of facial emotion expressions is, just like emotional speech recognition, nonintrusive and highly accurate (Jeon, 2017). People use the facial expressions of others daily to determine the emotional state of others, and it is one of the most important sources of information about others in interpersonal communication (Ko, 2018). Calvo & Nummenmaa (2015) showed that in humans, facial emotional expressions are perceived as categories, which is in accordance with Ekman's theory of distinct categories of emotion. It then makes sense for machines to do the same, instead of possibly measuring expressions on Russell's two dimensions.

As opposed to emotional speech recognizers, that can start processing features as soon as input is entered, conventional facial emotion recognizers (FER) first have to detect the actual face in the input. After this, the FERs operate in the same steps as conventional speech recognizers, which means that facial features are extracted from the input and the categorization of the expression according to pre-trained classifiers. The widely used system for facial features is the Facial Action Coding System (FACS). In the FACS, every single (observable) facial movement is classified as an Action Unit (AU), and corresponds to a specific muscle in the face. A facial expression is then formed by combining a select number of the 46 possible AUs (Mousavi et al., 2016; Tolba et al., 2018). However, since this system is based on analyzing static images, FACS will not be a very useful system when dealing with a continuous representation of a face. This is due to the binary nature of the AUs; a muscle can be either relaxed or contracted. Transitions in facial expressions cannot be adequately described because of this (Tolba et al., 2018). Another issue to overcome is that faces will not always be still and pointed right at a screen, nor will there be perfect lighting or neutral backgrounds (Jeon, 2017).

As Ko (2018) describes, the surge of research into neural networks have given rise to a new branch in expression recognition, which thus far yields better results than the conventional FERs. Aside from the improved accuracy and faces no longer having to be pointed forwards, neural networks also bring the advantage of not having to rely on pre-processing techniques and models, because they can learn directly from input images (Walecki et al., 2017). To overcome the shortcoming of neural networks in measurements over time, several hybrid approaches have been proposed, through combining conventional FERs with neural networks. Even these approaches have several limitations though, including the need for a lot of computing power, computing memory and large datasets. Nevertheless, hybrid approaches clearly perform better than all other options (Ko, 2018).

2.2.3 Text

Currently, the main form of communication between humans and computers is done by typing and reading text. Language in textual form could be emotionally charged, which then makes textual emotion recognition surely one of the most important ways for computers to recognize emotions (Daily et al., 2017). Recognizing emotion from text is often done using sentiment analysis tools, that integrate the fields of computational linguistics, text analysis and natural language processing. Conventional methods include keyword recognition (e.g. a curse word often involves an unhappy emotion) and machine learning. Using a machine learning algorithm with a large enough training set of previously categorized text will enable a system to learn the emotional value of certain words and their contexts (Cambria, 2016). Sentiment analysis however often does not go further than determining if a text is positively or negatively valenced. A major problem with keyword recognition, as Chatterjee et al. (2019) point out, is that a dictionary of all words possible linked with their associated emotion must be kept up to date, something which is unrealistic.

It is not easy to analyze natural language, let alone extracting the exact emotions that were expressed in a certain utterance. How text is interpreted relies heavily upon who is interpreting, and the absence of facial expressions, body language or speech intonation certainly does not help either (Kulkarni et al., 2018; Chatterjee et al., 2019). Kratzwald et al. (2018) mention that even though deep learning methods have become mainstream in facial and speech recognition, emotional text analysis using deep learning is still lagging behind. They demonstrate in a series of experiments that deep learning shows a substantial improvement in performance in natural language processing tasks (including sentiment analysis), compared to standard machine learning algorithms. Chatterjee et al. (2019) provide the same result, although they comment that deep learning models do take more time to train.

2.2.4 Other

There are other methods for measuring and recognizing emotions. As noted in chapter one, physiological signs such as heart rate or the skin conductance response prove to be very trustworthy indicators of emotional states. Since people come into contact with computing devices fairly often nowadays, it would be irrational to ignore physiology as a way to measure and recognize emotions. Measuring physiological features is however fairly intrusive, not to mention the practical issues that come with using physiology outside of the lab, e.g. the difference in human bodies, or environmental factors such as temperature which can influence several bodily functions. (Jeon, 2017). Affective wearables such as smartwatches or smart jewellery could prove to be a solution to the intrusiveness of measuring physiological features, but these wearables also suffer from practical drawbacks (Benyon, 2015; Costa et al., 2019).

Related to physiology, Daily et al. (2017) showed that the force people exert on objects (such as computer mouses, remote controls and gear sticks) correlate with their level of frustration. While this is not a way to recognize a distinct emotion, it can certainly help in combination with other measures.

2.3 Expression of emotions

After successfully recognizing emotions, an affective response must be generated. In this section, emotion expression in the form of speech, facial expressions and other expressive behaviors are covered.

2.3.1 Speech

The synthesis of emotional expressive speech is an important aspect of emotion expression in machines, just like it is an an important aspect of emotion recognition. Buchanan et al. (2018) found that most of the previous research on emotional expressive speech synthesis (ESS) has been directed at the modifying of basic prosodic speech features in synthesized speech output, such as pitch and loudness. This makes sense, since these are also thought to be the most useful speech features in emotion recognition (also see section 2.2.1 of this thesis). Voice quality is another important feature in emotion recognition and expression, but is significantly harder to modify than for example pitch or speech rate (Gobl & Ní Chasaide, 2003; Buchanan et al., 2018).

Human speech is incredibly varying: no single speaker has the same vocal tract, pronunciation style, and voice quality. Because of this, speech synthesizers often use a speech database containing speech from a single speaker. Some even eliminate the need for modifying the output speech by pre-recording different voice qualities when creating the database, enabling speech synthesizers to simulate emotion using only the available database (Kaliyev et al., 2018; Buchanan et al., 2018). However, as Burkhardt & Weiss (2018) note, only one emotional state is an over-simplification of most real world situations. No one ever only experiences one discrete emotion at a time, especially when considering ‘background’ emotional states such as moods. Being happy but bored is not the same as being sad and bored.

Comparable to speech recognition, (deep) neural networks have risen to become the mainstream technology in speech synthesis (Xue et al., 2018). They have shown to outperform most previous used technologies, including hidden Markov models, which was the leading technology for a while. It does not however remove the need for a large database containing high quality emotional speech though, and training a deep neural network on different speaking styles could prove to be an issue for the relationship between semantics and emotional expressivity (Tahon et al., 2018).

2.3.2 Facial expressions

Since facial expressions are crucial in interpersonal communication between humans, human-like facial expressions on emotional machines can further improve human-machine interaction. Being able to synthesize facial expressions further adds to a robot’s, or any other machine’s, capability to convey emotion and intention (Moosaei et al., 2015). Because of this, there has been a lot of interest in technologies that can synthesize facial expressions. Two main approaches have evolved from this interest. The first approach simply uses basic computer graphics techniques to transform input faces to convey the preferred expressions, while the second approach tries to build models that actually synthesize emotional faces (Song et al., 2018; Kollias et al., 2018).

Regarding the first approach, earlier works included generating new expressions by first creating highly detailed 3D models of ‘new’ faces, and then warping these to form an expression according to the corresponding facial features. The technique of composing new faces from other faces in existing databases also belongs to this first approach. The approach is often commended for resulting in realistic, high-quality images, but this comes with the cost of it being time-consuming and needing a lot of computing power. The second approach is not as expensive as the first approach, but often results in blurry or unrealistic images (Song et al., 2018).

It is worth noting that the advances in neural network research have also impacted facial expression synthesis. Previously, neural networks needed a large training set containing lots of pre-selected faces, rendering them inapplicable for situations where this cannot be realized (Kollias et al., 2018). However, very recently, Zakharov et al. (2019) presented a system that can synthesize facial expressions from only one pre-selected face (see Figure 2), opening up a whole new load of possibilities in the field of facial expression generation.

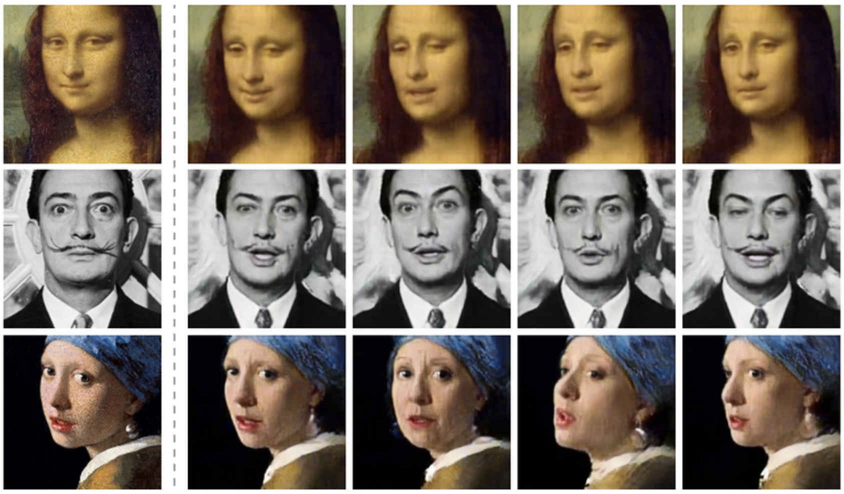

Figure 2. Generating multiple facial expressions from just one image. Adapted from Zakharov et al. (2019).

Being able to synthesize faces on a screen is one thing, but the advances in the field of physical robot’s facial expressions have been limited. According to Bartneck et al. (2008), a lot of attention in robotics has gone to the motor skills of robots - such as the ability to walk - which has resulted in a delay in the work on facial expressions of robots. Several obstacles in this field are hard to overcome: facial muscles are incredibly small and reactionary, which is hard to simulate. They also need to be able to be very precisely controlled, since facial expressions need to be accurate at low intensity levels as well. As Bartneck et al. (2008, p. 23) notes: “if the character managed to download a complete album of music it still did not save the world from global warming”.

The problem that all machines with facial expressions, be it virtual or physical robots, have to deal with, is the uncanny valley (Figure 3). Initially proposed by Masahiro Mori in 1970, it states that when the appearance of a robot or a virtual agent almost matches that of humans, but is still noticeably dissimilar, people will react very negatively towards it. Once the appearance is no longer distinguishable from real humans, no such effect is found and people react to it the same as to other humans (Zlotowski et al., 2018).

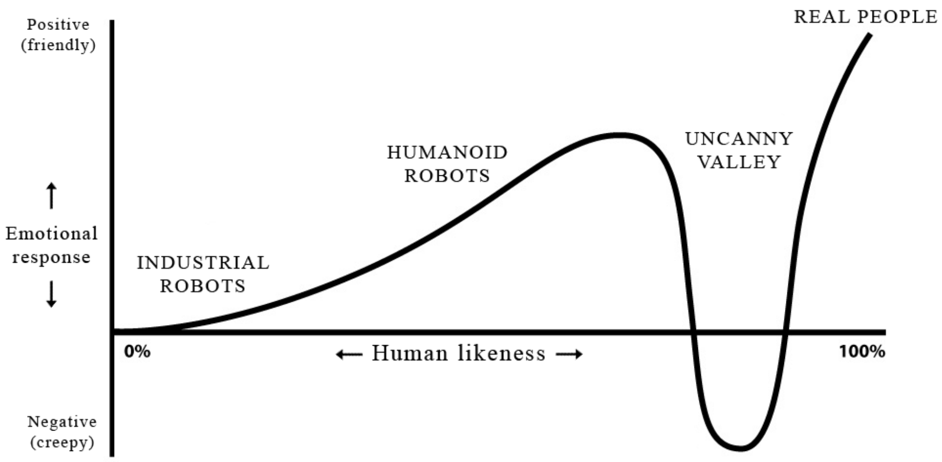

Figure 3. The uncanny valley.

2.3.3 Non-human expressions

The uncanny valley means that until human likeness of robots is practically 100%, robot developers deliberately avoid designing them like humans, and make sure they have distinct machine elements to them (Huang & Rust, 2018). One of the, if not the, most telling example of non-human emotion expression is the magic carpet in Disney's Aladdin. A carpet has no facial expression, no body, no voice - it cannot express emotion in remotely the same way humans do. However, by making use of the carpet's tassels, animators were able to bring it to life and convey to the public that it experiences emotions (Bennis & Biederman, 1998).

Another emotional expression to consider, that is somewhat speech-related, concerns computer sounds like most people are used to: beep sounds. Komatsu (2005) found that beeps with an increasing intonation (rising sound) are perceived as the computer being in some kind of disagreement, while a decreasing intonation (falling sound) was perceived to mean hesitation on the computer's side. Song & Yamada (2017) claim that disagreement and hesitation can be replaced with the emotions angry and sad, respectively.

Song & Yamada (2017) also found that color can have a significant effect on communicating emotions. For example, the color blue is heavily associated with sadness, while red is commonly recognized as anger. Fugate & Franco (2019) add that envy and jealousy are identified as being green, and happiness is tied with the color yellow. However, they also mention that color-emotion connections are not as specific as previously thought.

Fugate & Franco (2019) also mention that the color-emotion connections are likely to originate from language: when someone is 'seeing red', they are angry. Löffler et al. (2018) realized that many metaphors in language (not just the English language) involving emotion might be applied in robotics. Drawing on the work of Kövecses (2003), who composed a list of many emotional metaphors in English language, they propose some metaphors that could be used in robot design. For example, joy is described as being up ('cheering someone up') and warm ('feeling warm inside'). Anger on the other hand can be conceptualized as boiling hot ('boiling with anger'), while fear might be expressed as darkness ('overshadowed by fear').

While using the above non-human expressions of emotions are a good idea to prevent ending up in the uncanny valley, Cavallo et al. (2018) comment that through familiarisation of technology, the uncanny valley might level out completely and the academic field could shift their focus to advance work on humanoid robots.

2.3.4 Other

Interestingly, even just being aware of one's affective state reduces the influence of that affective state on behavior (Jeon, 2017). This effect is called the affect-as-information hypothesis, and could mean that just providing feedback to a user about their emotional state influences behavior. While this obviously is not like the other kinds of emotion expression covered in this section, it could still be of use in emotionally intelligent machines.

In the case of physical robots, Keltner et al. (2019) found that it is possible to accurately convey some emotions through very brief physical touches of the forearm. Among others, anger, fear and love were correctly interpreted by participants, while more complex emotions like embarrassment or awe were not. This could be of value for designing social comfort robots.

Emotional text synthesis does not nearly get as much attention as emotional speech synthesis or emotional text recognition, since it is not nearly as challenging to generate emotional text. According to Stark et al. (2018), a text can be made emotional simply through the addition of emotion words (e.g. 'I am angry'), emotional exclamations (e.g. 'yeah!') or through using words with the right emotional potential (e.g. 'family' has a positive emotional potential, 'betrayer' does not). Existing text generators have been providing impressive results for a long time now in several real world applications (Fedus et al., 2018; Tatalovic, 2018), and emotional text generation is no problem for the modern technologies. The difficulty lies in the recognition and subsequent processing of emotion in text (Asghar et al., 2017).

2.4 Conclusion

Despite fundamental differences between humans and machines that will never be overcome, this chapter has proved that one day, there might not be a difference anymore in the emotional functioning of humans and machines. It is probably too far fetched to say that machines ultimately will truly have and experience emotions, but like passing the Turing test means intelligence regardless of actually possessing intelligence, appearing to possess emotional intelligence can actually mean having emotional intelligence.

Research on emotion recognition and emotion expression by machines is at an all-time high, and results will only continue to improve. The potential of deep learning techniques is unparalleled, which opens the door to humanlike levels of recognition accuracy. And even though the uncanny valley still proves to be a problem in robotics, virtual face generation is climbing out of the valley towards 100% human likeness. The future is bright for affective computing.

Chapter 3. Applications of emotionally intelligent AI

Having discussed emotions in humans and how those emotions and emotional intelligence might be translated to machines, this chapter aims to discover in what fields, areas, jobs or services emotional intelligent machines could play a role in the (near) future. By reviewing recent literature, key areas are identified. A distinction is made between online virtual agents and physical robotic machines.

3.1 Online

Online emotional intelligent systems have one big advantage: they do not need to be embodied, and therefore escape all the problems that come with robotics. In this section, emotional intelligence in two key areas is discussed: chatbots, primarily in customer service, and education.

3.1.1 Chatbots

Chatbots, either online or running in software, are probably one of the earliest examples of artificial intelligence. ELIZA, introduced in 1966, already attempted to simulate a therapist providing emotional support. ELIZA was very primitive (although not for that time): she searched the input text for designated keywords, such as 'sad'. By then composing appropriate reactions through relevant linguistic rules and adding programmed words, she could respond in what seemed - sometimes - a genuine reaction (Liu & Sundar, 2018; Natale, 2019; also see figure 4).

```
YOU:   I am sad
ELIZA: I am sorry to hear that you are sad.
YOU:   Can you help me?
ELIZA: You like to think I help you -- don't you ?
YOU:   I do
ELIZA: You say you do ?
```

Figure 4. One of ELIZA's attempts at being a therapist.

Chatbots have obviously progressed since ELIZA. Moving on from simple rule-based responding to queries, they are now able to have complete conversations with people. While they have not yet passed the earlier mentioned Turing test, chatbots are widely used on the internet: in 2018, 36% of all brands online had implemented some kind of chatbot, and in 2020, 80% are expected to do so (Piccolo et al., 2018). Especially in marketing and customer service purposes, using chatbots is a popular approach to provide support to customers (Chaves & Gerosa, 2019). Bots that can recognize and express emotion can relieve customer service employees from some workload and thus lessen customer irritation. Some chatbots even mimic linguistic styles from the input of customers to improve customer relationships (Huang & Rust, 2018). Aside from customer service in companies, emotional chatbots can also play a role in government information and support (Androutsopoulou et al., 2018). They could answer questions from citizens or listen and process their complains, eliminating the need for human employees.

Another very important area where emotional chatbots could come into play is the medical field. Liu & Sundar (2018) describe that when users are searching for symptoms online and trying to diagnose themselves, they are often experiencing emotional distress through the anxiety and uncertainty about their symptoms. Medical advice that is accompanied by emotional support is more likely to be followed and will be rated more positively than neutral medical advice. Furthermore, disclosing symptoms to another can already reduce anxiety and improve the emotional experience a user is having - it being met with emotional support even further improves personal wellbeing (Ho et al., 2018). There are already chatbots available that can solicit trust and disclosure from users; another chatbot called Robin was created to support cyberbullying victims and has been critically acclaimed by anti-bullying experts (Liu & Sundar, 2018). It is clear that the potential for emotional intelligent systems in healthcare is high, but since healthcare normally takes place in a physical context, further discussion can be found in section 3.2.1.

To conclude, there is much potential for emotional chatbots in a diversity of fields. Since online chatbots are text-based, they will need to be able to recognize emotion from text. Based on the input, the chatbot must then not only formulate an appropriate answer, but also has to decide what emotion to express - if any. While users will know that they are conversing with a bot and not with a human, chatbots still need to adhere to conversational rules *and* select the correct emotion to express. Good emotional chatbots therefore will need high emotional intelligence, especially when it comes to being able to reason about personal and others' emotions. Emotional intelligence of that calibre might only be achieved when a bot is able to pass the Turing test, since such levels of intelligence are humanlike and thus require a bot's reactions to be indistinguishable from that of humans.

3.1.2 Education

As discussed in section 1.2.2 of this thesis, emotion can have considerable effect on learning and memory. In addition, a good teacher-student relationship improves self-esteem, confidence and intrinsic motivation (Guilherme, 2017). However, due to - among other things - a shortage of teachers, a lot of students do not get the individualized teaching they need to get the most out of their education. By employing emotional machines in the role of educators, either online or physical robots in the classroom, this could change, and emotion technology could thus have a large impact on future educational systems. Online teaching in itself can solve numerous problems, like providing education to students who cannot physically be present in schools. Online teaching being done by AI is an even bigger improvement, since it has the potential to solve the problem of oversized classes, and it can teach in places where no teachers may be available (such as third world countries). The improvement in the academic world could be enormous (Acemoglu & Restrepo, 2019). Governments and other educational companies seem to have realized this: almost a quarter of the 8 billion dollars of educational funding in the first half of 2018 was devoted to companies in AI or robotics (Afzal et al., 2019).

As an example of one of those applications, Afzal et al. (2019) developed an online tutor dubbed the Watson Tutor, to "promote student engagement with content, and self-reflection through the expression of a student's answers in their own words" (p. 46). It is designed to be able to do four things:

(1) ask questions, and if students cannot answer, provide them with hints, taking away any misconceptions;

(2) answer questions, and recommend relevant questions to ask;

(3) provide struggling students with multimedia content, something which can improve understanding of certain topics but is often overlooked by human educators;

and (4) test students by comparing and contrasting different parts of the material.

Especially for the first two components, it is necessary for the tutor to be able to recognize and express emotions. But unlike the chatbots discussed in section 3.1.1, this tutor will not converse on anything outside the course curriculum. This means that it does not need the high emotional intelligence allround chatbots need, and could maybe even function using pre-programmed responses to certain inputs - as long as it can accurately recognize input emotions. Afzal et al. do note that their tutor still has a long way to go, as almost half of the students that worked with the Watson Tutor expressed frustration about it not interpreting responses adequately. Eventual artificial intelligent educators, like Acemoglu & Restrepo (2019) suggested above, of course do need to be able to converse about topics outside the course curriculum. Those educators will need emotional intelligence that is equivalent to that of a human teacher - which again, probably involves needing to pass the Turing test. That is, they do not necessarily need to be indistinguishable from human teachers to students, but they need to be able to converse, explain, provide help and care for students exactly like human teachers can.

But assisting (or replacing) teachers is not the only influence emotional machines may have on education. Kaplan & Haenlein (2019) suggest that schools could use emotion recognition techniques on students, to analyze and test whether they are paying attention. Other applications in education they propose is reading emotions during tests to prevent and stop cheating, or to allow teachers to develop more effective teaching methods. These are obviously less complex systems than actual educators should be, since these only need some kind of emotion recognition technology, after which actual humans can interpret results and take appropriate actions. In this case, it really is not a case of whether it is possible, but whether it is accepted - constantly monitoring and reading students' emotions really touches the limits of what is ethical and what not.

3.2 Social Robots

Sometimes, online assistance is not enough. People might not have access to the internet all the time, or would rather interact with someone (or something) that they can put a face to. The presence of an actual head makes it easier to express emotions (Cavallo et al., 2018), although - as discussed in section 2.3.3 - emotions are just as easily expressed in non-human physical features. In this section, embodied emotional machines - also referred to as social robots - in two key areas are discussed: healthcare and customer service.

3.2.1 Healthcare

Healthcare is a very delicate field to work in. Professionals in the field are expected to be empathetically skilled, need to communicate well with patients and should generally be of high emotional intelligence (Huang & Rust, 2018). Is it then even feasible to employ social robots in healthcare? With the worldwide population ageing and the increasing healthcare costs that come with it, there might not even be a choice in the future. Social robots in healthcare can significantly reduce the workload of doctors and increase the amount of people that can be adequately helped and guided through their sickness (Miller & Polson, 2019). It might even allow more frequent home-visits by medical personnel, which is a welcome change, especially for the less mobile elderly.

However, to keep conversations and interaction natural, it is important for these social robots to be equipped with state-of-the-art emotion recognition and expression techniques. In the tricky environment of healthcare, it is important that social robots understand and can react to most, if not all, forms of human behavior. This requires an almost humanlike emotional intelligence that cannot be pre-programmed, unlike some earlier mentioned applications. It requires the robot to truly be able to reason about input and appropriate output. It is therefore unlikely that standalone robotic companions will arise in the near future. A more plausible option is a social robot that will 'only' assist human medical personnel.

In some cases, social robots might be an improvement over human medics though: for example, children with autism generally interact better with robots than with humans (Krakovsky, 2018; Cavallo et al., 2018). Some robots are designed to recognize the behavior the children are expressing, and reward appropriate behavior by performing enjoyable acts, like blowing bubbles. Others are made to function as a peer for the children, to provide a way for them to learn social skills. Because chronic or severe illnesses may prevent children from interacting with their friends and others, social robots could fill the gap by interacting with them and distracting them from the often intimidating medical procedures. In some cases, the social robots can also help gather data for medical purposes. Research has shown overwhelmingly positive reactions on many social robots in children healthcare, not only from children but also from medical staff, parents and others involved, and social robots are more effective than text interfaces or virtual robots (Dawe et al., 2019).

It is not weird that even non-autistic children and teens seem more comfortable when interacting with (robotic) devices than with adults: They have been interacting with the digital world for literally their entire lives (Björling et al., 2019). Several studies have found that social robots are more likely than adults to get children and teens to talk about personal information. This raises the question if social robots would be equally as useful when employed to help the elderly.

Recent work presents conflicting results when it comes to the attitude of the elderly towards social robots. Positive results emerge in Vitanza et al. (2019), with just the interaction process alone improving moods through reducing loneliness. Miller & Polson (2019) also report this effect, with one participant even describing the social robot like a friend. Niemelä & Melkas (2019) however describe more conservative reactions, with elderly not liking the idea of robots setting themselves up as being emotional and caring while they do not actually *feel,* and are therefore deceptive and not genuine in their interactions. Others express concern about losing human contact, especially if relatives assume the robot takes care of them (Wirtz et al., 2018). When or if social robots are eventually accepted by the elderly as friendly companions, they could be used to monitor not only moods, but also behavior and medical condition. Cavallo et al. (2018) even mentions that they could be used to measure the onset of cognitive degenerative diseases, like Alzheimer's.

It is clear that there are a lot of advantages to emotional and social robots in healthcare. They can keep patients of any age company, can monitor medical condition in unobtrusive ways and, in the case of children, can more easily obtain personal information than humans can. Still, because of the current limits of emotional intelligence in robots, it is unlikely that they will function on their own in the near future, let alone replace human medical personnel.

3.2.2 Customer service

Three years ago, in 2016, there were already more than ten thousand 'Pepper' service robots moving around the globe, in wildly different environments and functions: as a sales assistant in Nescafé stores, a waiter in Pizza Huts or offering food recommendations on airports (Mende et al., 2019). Pepper is a humanoid robot that communicates via speech and a screen on his chest. He is not emotional (yet), but he has proven the incredibly high potential of service robots in a multitude of applications by his versatile use.

Wirtz et al. (2018) describe that especially in ‘subordinate service roles’ (SSRs), such as waiters or cashiers, social robots may flourish. Since the employees in those roles often do not have great engagement or motivation, the social robots could provide better service, as they are not affected (or rather can be programmed to not be affected) by external factors and internal biases. Ivanov & Webster (2019) confirm these assumptions by reviewing studies on the application of social robots in hotels. Guests rate deploying the robots the highest when it comes to SSRs such as checking in, taking orders or room service, while they did not feel comfortable with social robots providing more humanlike services, like giving massages, babysitting or hairdressing. Bowen & Morosan (2018) also describe that social robots can provide a solution for institutions that lack the human resources to effectively provide help, data or other information to customers.

Different from healthcare social robots, these kind of robots do not need to be as emotionally intelligent. Of course, the technologies that should be used are still highly advanced; for example, a robot’s speech recognition in customer service must be accurate with many different speaking styles, accents and maybe even languages (Bowen & Morosan, 2018). And while companies or other institutions likely want their customers to stay on a predetermined path, users often want to be able to initiate (personal) conversations, suggest different topics or even want the robot itself to ask questions. Users “want to establish some kind of relationship with the system and a desire to share their opinions and thoughts with it” (Sono et al., 2019, p. 11), especially if the robot serves a role like that of a guide. Even though a social service robot does not have to be as watertight as a robot in healthcare, this obviously means that they too should be able to convey emotions and reason about how to act. Of course, there are gradations in this, as a check-in robot does not require the same emotional intelligence as a museum guide.

3.3 Conclusion

In this chapter, it has become clear that there is an abundance of possible applications for emotionally intelligent systems. In several fields, a system - be it a robot or an online interface - that is able to recognize and express emotions, and appears to understand them, can be a valuable asset to have.

Online, chatbots can assist users and potential customers where human employees cannot do that. Educational online systems can motivate students to learn, and try to get them in a better mood to better remember course material. In the physical world, social emotional robots are expected to play a big role in healthcare. From assisting the elderly with tasks that have become impossible for them to perform, to just providing them some company and offer a sympathetic ear. Children can find a friend in an emotional robot, and disclose information to them that they would not have disclosed to a medical professional. But social emotional robots can also be used in positions related to customer service, such as hotel clerks or museum guides. Numerous other applications of emotion technology or emotional intelligence in systems can be thought of. To name a few, personalized advertising (Kaplan & Haenlein, 2019; Huang & Rust, 2018), personal entertainment (Rincon et al., 2018; Gaggioli, 2018) and emotional gaming (Lara-Cabrero & Camacho, 2019; Frommel et al., 2018; Lara et al., 2018) are prevalent in recent literature. The potential for emotional technology is immense and the advances in the field of affective computing are very likely to improve the quality of life.

As with any new or upcoming technologies however, drawbacks are plenty. The most obvious one is that of privacy and security - any system with the capability to recognize emotions will have to store recordings somewhere. Through repeated interaction with a specific user, the system could learn a thing or two about them, resulting in even more (very private) data that needs to be stored safely and securely. Other concerns include the fear of social robots reducing human contact, or causing an increase in human unemployment.

Despite the issues surrounding the deployment of emotional intelligent systems, the possible advantages clearly far outweigh the possible negatives.

Conclusion

This thesis aimed to find the potential of artificial emotional intelligence. It set off by providing an overview of human emotion research, a field that still has numerous open ends and probably will have for a long time. This did not hinder the rise of affective computing however, discussed in chapter 2. It was made clear that even though it might never be possible to design systems that actually experience emotions the way humans do, it is sufficient to make them appear like they do. Emotion technology research is at an all time high with the recent application of deep learning techniques. Emotions can now be recognized with humanlike accuracy, and methods to express emotion are becoming equally advanced.

The bottleneck lies at connecting emotion recognition with emotion expression. For systems, having emotional intelligence means being able to decide what emotion to express based on what emotion was recognized and possible other factors, and there is still a lot of progress to be made in that field. In chapter 3, key areas that could profit from emotionally intelligent systems were discussed, along with the necessary perceived level of emotional intelligence. Some fundamental societal problems, such as the current lack of educators, or the ageing population and associated healthcare costs, could partly be solved by emotionally intelligent systems. The potential of emotional intelligence in machines has shown to be immense.

As mentioned in the introduction, research into emotions in artificial intelligence aside from recognition and expression is still limited. Future research needs to find a solid method to couple those standalone processes without the need for pre-programmed rules, as human behavior is too diverse in nature to fully grasp in rule-based emotion systems. State-of-the-art deep learning techniques in artificial intelligence can bring the field of affective computing to an unprecedented level, hopefully taking the next step towards full artificial emotional intelligence.

References

Acemoglu, D., & Restrepo, P. (2019). *The wrong kind of AI? Artificial Intelligence and the future of labor demand* (IZA Discussion Papers, No. 12292). Retrieved from https://www.econstor.eu/bitstream/10419/196790/1/dp12292.pdf

Adolphs, R., & Anderson, D. J. (2018). *The neuroscience of emotion: A new synthesis*. Princeton, NJ, United States of America: Princeton University Press.

Afzal, S., Dempsey, B., D'Helon, C., Mukhi, N., Pribic, M., Sickler, A., Strong, P., Vanchiswar, M. & Wilde, L. (2019). The personality of AI systems in education: Experiences with the Watson Tutor, a one-on-one virtual tutoring system. *Childhood Education, 95*(1), 44–52. https://doi.org/10.1080/00094056.2019.1565809

Akgün, M., Akıllı, G. K., & Çağıltay, K. (2010). Bringing affect to human computer interaction. In D. Gökcay, & G. Yildirim (Eds.), *Affective computing and interaction: Psychological, cognitive and neuroscientific perspectives* (pp. 308–324). Hershey, NY, United States of America: Information Science Reference.

Ali, M., Mosa, A. H., Machot, F. A., & Kyamakya, K. (2017). Emotion recognition involving physiological and speech signals: A comprehensive review. In K. Kyamakya, W. Mathis, R. Stoop, J. Chedjou, & Z. Li (Eds.), *Recent advances in nonlinear dynamics and synchronization* (pp. 287–302). https://doi.org/10.1007/978-3-319-58996-1_13

Androutsopoulou, A., Karacapilidis, N., Loukis, E., & Charalabidis, Y. (2018). Transforming the communication between citizens and government through AI-guided chatbots. *Government Information Quarterly, 36*(2), 358–367. https://doi.org/10.1016/j.giq.2018.10.001

Asghar, N., Poupart, P., Hoey, J., Jiang, X., & Mou, L. (2018). Affective neural response generation. In G. Pasi, B. Piwowarski, L. Azzopardi, & A. Hanbury (Eds.), *Advances in Information Retrieval. ECIR 2018: Lecture Notes in Computer Science, vol 10772* (pp. 154–166). https://doi.org/10.1007/978-3-319-76941-7_12

Bartneck, C., Lyons, M. J., & Saerbeck, M. (2008). The relationship between emotion models and artificial intelligence. In ORG. SAB (Ed.), *Proceedings of the Workshop on The Role Of Emotions In Adaptive Behaviour And Cognitive Robotics in affiliation with the 10th International Conference on Simulation of Adaptive Behavior: From Animals to Animates* (pp. 16–27). Osaka, Japan: SAB.

Bennis, W., & Biederman, P. W. (1998). *Organizing genius: The secrets of creative collaboration*. New York, NY, United States of America: Basic Books.

Benyon, D. (2015). *Designing interactive systems: A comprehensive guide to HCI, UX and interaction design* (3rd ed.). Harlow, United Kingdom: Pearson Education Limited.

Björling, E. A., Rose, E., Davidson, A., Ren, R., & Wong, D. (2019). Can we keep him forever? Teens' engagement and desire for emotional connection with a social robot. *International Journal of Social Robotics*. Advance online publication. https://doi.org/10.1007/s12369-019-00539-6

Bowen, J., & Morosan, C. (2018). Beware hospitality industry: The robots are coming. *Worldwide Hospitality and Tourism Themes, 10*(6), 726–733. https://doi.org/10.1108/whatt-07-2018-0045

Brave, S., & Nass, C. (2002). Emotion in human-computer interaction. In J. Jacko, & A. Sears (Eds.), *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (pp. 81–96). https://doi.org/10.1201/b10368-6

Buchanan, C. G., Aylett, M. P., & Braude, D. A. (2018). Adding personality to neutral speech synthesis voices. In A. Karpov, O. Jokisch, & R. Potapova (Eds.), *Speech and Computer: SPECOM 2018: Lecture Notes in Computer Science, vol 11096* (pp. 49–57). Cham, Switzerland: Springer Nature Switzerland.

Burkhardt, F., Weiss, B. (2018). Speech synthesizing simultaneous emotion-related states. In A. Karpov, O. Jokisch, & R. Potapova (Eds.), *Speech and Computer: SPECOM 2018: Lecture Notes in Computer Science, vol 11096* (pp. 49–57). Cham, Switzerland: Springer Nature Switzerland.

Calvo, M. G., & Nummenmaa, L. (2015). Perceptual and affective mechanisms in facial expression recognition: An integrative review. Cognition and Emotion, 30(6), 1081–1106. https://doi.org/10.1080/02699931.2015.1049124

Cambria, E. (2016). Affective computing and sentiment analysis. IEEE Intelligent Systems, 31(2), 102–107. https://doi.org/10.1109/mis.2016.31

Cavallo, F., Semeraro, F., Fiorini, L., Magyar, G., Sinčák, P., & Dario, P. (2018). Emotion modelling for social robotics applications: A review. *Journal of Bionic Engineering, 15*(2), 185–203. https://doi.org/10.1007/s42235-018-0015-y

Chatterjee, A., Gupta, U., Chinnakotla, M. K., Srikanth, R., Galley, M., & Agrawal, P. (2019). Understanding emotions in text using deep learning and big data. *Computers in Human Behavior, 93*, 309–317. https://doi.org/10.1016/j.chb.2018.12.029

Chaves, A. P., & Gerosa, M. (2019). *How should my chatbot interact? A survey on human-chatbot interaction design*. Retrieved from https://arxiv.org/pdf/1904.02743.pdf

Chong, C. S., Kim, J., & Davis, C. (2018). Disgust expressive speech: The acoustic consequences of the facial expression of emotion. *Speech Communication, 98*, 68–72. https://doi.org/10.1016/j.specom.2017.12.007

Costa, A., Rincon, J. A., Carrascosa, C., Julian, V., & Novais, P. (2019). Emotions detection on an ambient intelligent system using wearable devices. *Future Generation Computer Systems, 92*, 479–489. https://doi.org/10.1016/j.future.2018.03.038

Cowen, A. S., & Keltner, D. (2018). Clarifying the conceptualization, dimensionality, and structure of emotion: Response to Barrett and colleagues. *Trends in Cognitive Sciences, 22*(4), 274–276. https://doi.org/10.1016/j.tics.2018.02.003

Daily, S. B., James, M. T., Cherry, D., Porter, J. J., Darnell, S. S., Isaac, J., & Roy, T. (2017). Affective computing: Historical foundations, current applications, and future trends. In M. Jeon (Ed.), *Emotions and affect in human factors and human-computer interaction* (pp. 213–231). London, United Kingdom: Academic Press.

Dawe, J., Sutherland, C., Barco, A., & Broadbent, E. (2019). Can social robots help children in healthcare contexts? A scoping review. *BMJ Paediatrics Open, 3*(1), 1–16. https://doi.org/10.1136/bmjpo-2018-000371

Dubois, J., & Adolphs, R. (2015). Neuropsychology: How many emotions are there? *Current Biology, 25*(15), 669–672. https://doi.org/10.1016/j.cub.2015.06.037

Earles, J. L., Kersten, A. W., Vernon, L. L., & Starkings, R. (2015). Memory for positive, negative and neutral events in younger and older adults: Does emotion influence binding in event memory? *Cognition and Emotion, 30*(2), 378–388. https://doi.org/10.1080/02699931.2014.996530

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*(3-4), 169–200. https://doi.org/10.1080/02699939208411068

Fan, L., Scheutz, M., Lohani, M., McCoy, M., & Stokes, C. (2017). Do we need emotionally intelligent artificial agents? First results of human perceptions of emotional intelligence in humans compared to robots. In J. Beskow, C. Peters, G. Castellano, C. O'Sullivan, I. Leite, & S. Kopp (Eds.), *Intelligent virtual agents: IVA 2017: Lecture notes in Computer Science, vol 10498* (pp. 129–141). https://doi.org/10.1007/978-3-319-67401-8_15

Fedus, W., Goodfellow, I., & Dai, A. (2018). *MaskGAN: Better text generation via filling in the* ___. Paper presented at ICLR 2018, Vancouver, Canada. Retrieved from https://openreview.net/pdf?id=ByOExmWAb

Franzoni, V., Milani, A., Nardi, D., & Vallverdú, J. (2019). Emotional machines: The next revolution. *Web Intelligence, 17*(1), 1–7. https://doi.org/10.3233/WEB-190395

Friedenberg, J., & Silverman, G. (2015). *Cognitive science: An introduction to the study of mind* (3rd ed.). Thousand Oaks, CA, United States of America: SAGE Publications.

Frommel, J., Schrader, C., & Weber, M. (2018). Towards emotion-based adaptive games: Emotion recognition via input and performance features. In F. Mueller, D. Johnson, B. Schouten, Z. O. Toups, & P. Wyeth (Eds.), *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play* (pp. 173–185). https://doi.org/10.1145/3242671.3242672

Fugate, J. M. B., & Franco, C. L. (2019). What color is your anger? Assessing color-emotion pairings in English speakers. *Frontiers in Psychology, 10*(206), 1–17. https://doi.org/10.3389/fpsyg.2019.00206

Gaggioli, A. (2018). Virtual personal assistants: An emerging trend in artificial intelligence. *Cyberpsychology, Behavior, and Social Networking, 21*(12), 803–804. https://doi.org/10.1089/cyber.2018.29133.csi

Gilchrist, P. T., Vrinceanu, T., Béland, S., Bacon, S. L., & Ditto, B. (2016). Disgust stimuli reduce heart rate but do not contribute to vasovagal symptoms. *Journal of Behavior Therapy and Experimental Psychiatry, 51*, 116–122. https://doi.org/10.1016/j.jbtep.2016.01.005

Gluck, M. A., Mercado, E., & Myers, C. E. (2016). *Learning and memory: From brain to behavior* (3rd ed.). New York, NY, United States of America: MacMillan Learning.

Gobl, C., & Ní Chasaide, A. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication, 40*(1-2), 189–212. https://doi.org/10.1016/S0167-6393(02)00082-1

Guilherme, A. (2017). AI and education: the importance of teacher and student relations. *AI & SOCIETY, 34*(1), 47–54. https://doi.org/10.1007/s00146-017-0693-8

Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication, 68*(4), 712–733. https://doi.org/10.1093/joc/jqy026

Hostler, T. J., Wood, C., & Armitage, C. J. (2018). The influence of emotional cues on prospective memory: A systematic review with meta-analyses. *Cognition and Emotion, 32*(8), 1578–1596. https://doi.org/10.1080/02699931.2017.1423280

Huang, M., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research, 21*(2), 155–172. https://doi.org/10.1177/1094670517752459

Ioannidou, F., & Konstatikaki, V. (2008). Empathy and emotional intelligence: What is it really about. *International Journal of Caring Sciences, 1*(3), 118–123.

Ivanov, S., & Webster, C. (2019). Perceived appropriateness and intention to use service robots in tourism. In J. Pesonen, & J. Neidhardt (Eds.), *Information and communication technologies in tourism 2019: Proceedings of the international conference in Nicosia, Cyprus* (pp. 237–248). Cham, Switzerland: Springer Nature Switzerland.

Izard, C. E. (2010). The many meanings/aspects of emotion: Definitions, functions, activation, and regulation. *Emotion Review, 2*(4), 363–370. https://doi.org/10.1177/1754073910374661

Jeon, M. (2017). Emotions and affect in human factors and human-computer interaction: Taxonomy, theories, approaches and methods. In M. Jeon (Ed.), *Emotions and affect in human factors and human-computer interaction* (pp. 3–26). London, United Kingdom: Academic Press.

Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (2nd ed.). Upper Saddle River, NJ, United States of America: Pearson Education.

Kaliyev, A., Matveev, Y. N., Lyakso, E. E., & Rybin, S. V. (2018). Prosodic processing for the automatic synthesis of emotional Russian speech. In M. A. Nazarenko, S. V. Kruchinin, M. M. Yakhutlov, E. V. Bagrova, & V. N. Azarov (Eds.), *2018 IEEE International Conference on Quality Management, Transport and Information Security, Information Technologies (IT&QM&IS)* (pp. 653–655). Saint Petersburg, Russia: IEEE.

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of Artificial Intelligence. *Business Horizons, 62*(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004

Kaplan, R. L., Van Damme, I., Levine, L. J., & Loftus, E. F. (2015). Emotion and false memory. *Emotion Review, 8*(1), 8–13. https://doi.org/10.1177/1754073915601228

Keltner, D., & Cordaro, D.T. (2015). Understanding multimodal emotional expressions: Recent advances in Basic Emotion Theory. In A. Scarantino (Ed.), *Emotion Researcher*. Retrieved from| http://emotionresearcher.com/understanding-multimodal-emotional-expressions-recent-advances-in-basic-emotion-theory/

Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional expression: Advances in basic emotion theory. *Journal of Nonverbal Behavior*, 1–28. https://doi.org/10.1007/s10919-019-00293-3

Ko, B. (2018). A brief review of facial emotion recognition based on visual information. *Sensors, 18*(2), 401–421. https://doi.org/10.3390/s18020401

Kollias, D., Cheng, S., Pantic, M., & Zafeiriou, S. (2018). Photorealistic facial synthesis in the dimensional affect space. In L. Leal-Taixé, & S. Roth (Eds.), *Computer Vision: ECCV 2018 Workshops: Lecture Notes in Computer Science, vol 11130* (pp. 475–491). Cham, Switzerland: Springer Nature Switzerland.

Komatsu, T. (2005). Toward making humans empathize with artificial agents by means of subtle expressions. In J. Tao, T. Tan, & R. W. Picard (Eds.), *Affective Computing and Intelligent Interaction: First International Conference, ACII 2005* (pp. 458–465). Beijing, China: Springer.

Kövecses, Z. (2003). *Metaphor and emotion: Language, culture, and body in human feeling*. Cambridge, United Kingdom: Cambridge University Press.

Krakovsky, M. (2018). Artificial (emotional) intelligence. *Communications of the ACM, 61*(4), 18–19. https://doi.org/10.1145/3185521

Kratzwald, B., Ilić, S., Kraus, M., Feuerriegel, S., & Prendinger, H. (2018). Deep learning for affective computing: Text-based emotion recognition in decision support. *Decision Support Systems, 115*, 24–35. https://doi.org/10.1016/j.dss.2018.09.002

Kulkarni, P. V., Nagori, M. B., & Kshirsagar, V. P. (2018). An in-depth survey of techniques employed in construction of emotional lexicon. In S. Satapathy, & A. Joshi (Eds.), *Information and communication technology for intelligent systems. Smart innovation, systems and technologies, vol 106* (pp. 609–620). https://doi.org/10.1007/978-981-13-1742-2_61

Lara, C. A., Flores, J., Mitre-Hernandez, H., & Perez, H. (2018). Induction of emotional states in educational video games through a fuzzy control system. *IEEE Transactions on Affective Computing*, 3045, 1–12. https://doi.org/10.1109/taffc.2018.2840988

Lara-Cabrera, R., & Camacho, D. (2019). A taxonomy and state of the art revision on affective games. *Future Generation Computer Systems, 92*, 516–525. https://doi.org/10.1016/j.future.2017.12.056

Lewis, P. A., & Critchley, H. D. (2003). Mood-dependent memory. *Trends in Cognitive Sciences, 7*(10), 431–433. https://doi.org/10.1016/j.tics.2003.08.005

Li, Y., Chao, L., Liu, Y., Bao, W., & Tao, J. (2015). From simulated speech to natural speech, what are the robust features for emotion recognition? In ORG. Association for the Advancement of Affective Computing (Ed.), *2015 International conference on affective computing and intelligent interaction (ACII)* (pp. 368–373). https://doi.org/10.1109/ACII.2015.7344597

Liu, B., & Sundar, S. S. (2018). Should machines express sympathy and empathy? Experiments with a health advice chatbot. *Cyberpsychology, Behavior, and Social Networking, 21*(10), 625–636. https://doi.org/10.1089/cyber.2018.0110

Liu, X., & Xu, Y. (2016). Pitch perception and surprise in Mandarin Chinese: Evidence for parallel encoding via additive division of pitch range. *Tonal Aspects of Languages 2016*, 129–132. https://doi.org/10.21437/tal.2016-28

Löffler, D., Schmidt, N., & Tscharn, R. (2018). Multimodal expression of artificial emotion in social robots using color, motion and sound. In T. Kanda, S. Ŝabanović, G. Hoffman, & A. Tapus (Eds.), *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 334–343). https://doi.org/10.1145/3171221.3171261

Mattingly, V., & Kraiger, K. (2019). Can emotional intelligence be trained? A meta-analytical investigation. *Human Resource Management Review, 29*(2), 140–155. https://doi.org/10.1016/j.hrmr.2018.03.002

Mayer, J. D. (2004). What is emotional intelligence? *UNH Personality Lab, 8*. Retrieved May 19, 2019, from https://scholars.unh.edu/personality_lab/8

Mende, M., Scott, M. L., Van Doorn, J., Grewal, D., & Shanks, I. (2019). Service robots rising: How humanoid robots influence service experiences and elicit compensatory consumer responses. *Journal of Marketing Research.* Advance online publication. https://doi.org/10.1177/0022243718822827

Miller, E., & Polson, D. (2019). Apps, avatars, and robots: The future of mental healthcare. *Issues in Mental Health Nursing, 40*(3), 208–214. https://doi.org/10.1080/01612840.2018.1524535

Miners, C. T. H., Côté, S., & Lievens, F. (2018). Assessing the validity of emotional intelligence measures. *Emotion Review, 10*(1), 87–95. https://doi.org/10.1177/1754073917693688

Moosaei, M., Hayes, C. J., & Riek, L. D. (2015). Performing facial expression synthesis on robot faces: A real-time software system. In M. Salem, A. Weiss, P. Baxter, & K. Dautenhahn (Eds.), *4th International Symposium on New Frontiers in Human-Robot Interaction* (pp. 67–73). Canterbury, United Kingdom: AISB.

Mousavi, N., Siqueira, H., Barros, P., Fernandes, B., & Wermter, S. (2016). Understanding how deep neural networks learn face expressions. In ORG. Institute of Electrical and Electronic Engineers (Ed.), *Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN)* (pp. 227–234). https://doi.org/10.1109/IJCNN.2016.7727203

Mulligan, K., & Scherer, K. R. (2012). Toward a working definition of emotion. *Emotion Review, 4*(4), 345–357. https://doi.org/10.1177/1754073912445818

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America, 93*(2), 1097–1108. https://doi.org/10.1121/1.405558

Mustafa, M. B., Yusoof, M. A. M., Don, Z. M., & Malekzadeh, M. (2018). Speech emotion recognition research: An analysis of research focus. *International Journal of Speech Technology, 21*(1), 137–156. https://doi.org/10.1007/s10772-018-9493-x

Nassif, A. B., Shahin, I., Attili, I., Azzeh, M., & Shaalan, K. (2019). Speech recognition using deep neural networks: A systematic review. *IEEE Access, 7*, 19143–19165. https://doi.org/10.1109/access.2019.2896880

Natale, S. (2019). If software is narrative: Joseph Weizenbaum, artificial intelligence and the biographies of ELIZA. *New Media & Society, 21*(3), 712–728. https://doi.org/10.1177/1461444818804980

Niemelä, M., & Melkas, H. (2019). Robots as social and physical assistants in elderly care. In M. Toivonen, & E. Saari (Eds.), *Human-centered digitalization and services: Translational systems sciences, vol 19* (pp. 177–197). https://doi.org/10.1007/978-981-13-7725-9_10

Pestana, M. H., Wang, W. C., & Moutinho, L. (2018). The knowledge domain of affective computing: A scientometric review. In L. Moutinho, & M. Sokele (Eds.), *Innovative research methodologies in management: Volume II: Futures, biometrics and neuroscience research* (pp. 83–101). https://doi.org/10.1007/978-3-319-64400-4_4

Picard, R. W. (1999). Affective computing for HCI. In H. J. Bullinger, & J. Ziegler (Eds.), Proceedings of HCI International '99 8th International Conference on Human Computer Interaction: Ergonomics and User Interfaces (pp. 829–833). Mahwah, USA: Lawrence Erlbaum Associates.

Picard, R. W. (2000). *Affective computing*. Cambridge, MA, United States of America: MIT Press.

Piccolo, L. S. G., Mensio, M., & Alani, H. (2018). Chasing the chatbots. In S. S. Bodrunova, O. Koltsova, A. Følstad, H. Halpin, P. Kolozaridi, L. Yuldashev, A. Smoliarova & H. Niedermayer (Eds.), *Internet Science: INSCI 2018. Lecture Notes in Computer Science, vol 11551* (pp. 157–169). https://doi.org/10.1007/978-3-030-17705-8_14

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik, & H. Kellerman (Eds.), *Emotion: theory, research, and experience, Volume 1* (pp. 3–34). New York, NY, United States of America: Academic Press.

Rincon, J. A., Costa, A., Novais, P., Julian, V., & Carrascosa, C. (2018). A new emotional robot assistant that facilitates human interaction and persuasion. *Knowledge and Information Systems*. Advance online publication. https://doi.org/10.1007/s10115-018-1231-9

Roth, C. G., Eldin, K. W., Padmanabhan, V., & Friedman, E. M. (2018). Twelve tips for the introduction of emotional intelligence in medical education. *Medical Teacher*, 1–4. https://doi.org/10.1080/0142159x.2018.1481499

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

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect Grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 57*(3), 493–502. https://doi.org/10.1037/0022-3514.57.3.493

Scherer, K. R. (1984). Emotion as a multi-component process: A model and some cross-cultural data. In P. Shaver (Ed.), *Review of personality and social psychology, 5: Emotions, relationships, and health* (pp. 37–63). Beverly Hills, CA, United States of America: Sage.

Scherer, K. R., Johnstone, T., & Klasmeyer, G. (2003). Vocal expression of emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 433–456). Oxford, United Kingdom: Oxford University Press.

Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., Xu, X., Yang, X. (2018). A review of emotion recognition using physiological signals. *Sensors, 18*(7), 2074–2115. https://doi.org/10.3390/s18072074

Siedlecka, E., & Denson, T. F. (2019). Experimental methods for inducing basic emotions: A qualitative review. *Emotion Review, 11*(1), 87–97. https://doi.org/10.1177/1754073917749016

Song, L., Lu, Z., He, R., Sun, Z., & Tan, T. (2018). Geometry guided adversarial facial expression synthesis. In ORG. ACM (Ed.), *Proceedings of the 26th ACM international conference on Multimedia* (pp. 627–635). Seoul, Republic of Korea: ACM.

Song, S., & Yamada, S. (2017). Expressing emotions through color, sound, and vibration with an appearance-constrained social robot. In ORG. ACM (Ed.), *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction - HRI '17* (pp. 2–11). Vienna, Austria: ACM.

Sono, T., Satake, S., Kanda, T., & Imai, M. (2019). Walking partner robot chatting about scenery. *Advanced Robotics*. Advance online publication. https://doi.org/10.1080/01691864.2019.1610062

Stark, L., Brünken, R., & Park, B. (2018). Emotional text design in multimedia learning: A mixed-methods study using eye tracking. *Computers & Education, 120*, 185–196. https://doi.org/10.1016/j.compedu.2018.02.003

Tahon, M., Lecorve, G., & Lolive, D. (2018). Can we generate emotional pronunciations for expressive speech synthesis? *IEEE Transactions on Affective Computing*. https://doi.org/10.1109/taffc.2018.2828429

Tatalovic, M. (2018). AI writing bots are about to revolutionise science journalism: We must shape how this is done. *Journal of Science Communication, 17*(1), E1-7. https://doi.org/10.22323/2.17010501

Tolba, R. M., Al-Arif, T., & El Horbaty, E. M. (2018). Realistic facial animation review: Based on facial action coding system. *Egyptian Computer Science Journal, 42*(1). Retrieved from http://ecsjournal.org/Archive/Volume42/Issue1/6.pdf

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.01454

Vinciarelli, A., & Mohammadi, G. (2010). Towards a technology of nonverbal communication: Vocal behavior in social and affective phenomena. In D. Gökcay, & G. Yildirim (Eds.), *Affective computing and interaction: Psychological, cognitive and neuroscientific perspectives* (pp. 133–156). Hershey, NY, United States of America: Information Science Reference.

Vitanza, A., D'Onofrio, G., Ricciardi, F., Sancarlo, D., Greco, A., & Giuliani, F. (2019). Assistive robots for the elderly: Innovative tools to gather health relevant data. In S. Consoli, D. Reforgiato Recupero, & M. Petković (Eds.), *Data science for healthcare* (pp. 195–215). https://doi.org/10.1007/978-3-030-05249-2_7

Walecki, R., Rudovic, O., Pavlovic, V., Schuller, B., & Pantic, M. (2017). Deep structured learning for facial action unit intensity estimation. In ORG. Institute of Electrical and Electronic Engineers (Ed.), *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3405–3414). Honolulu, HI, United States of America: IEEE.

Wibawa, A. W., & Purnomo, M. H. (2016). Physiological pattern of basic human emotion state based on SPO2 sensor. *International Journal of Signal System Control and Engineering Application, 9*(3), 55–59. https://doi.org/10.3923/ijssceapp.2016.55.59

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: Service robots in the frontline. *Journal of Service Management, 29*(5), 907–931. https://doi.org/10.1108/josm-04-2018-0119

Wu, S., Falk, T. H., & Chan, W. (2010). Automatic speech emotion recognition using modulation spectral features. *Speech Communication, 53*(5), 768–785. https://doi.org/10.1016/j.specom.2010.08.013

Xue, L., Zhu, X., An, X., & Xie, L. (2018). A comparison of expressive speech synthesis approaches based on neural network. In ORG. ACM (Ed.), *The Joint Workshop of the 4th Workshop on Affective Social Multimedia Computing and first Multi-Modal Affective Computing of Large-Scale Multimedia Data Workshop (ASMMC-MMAC'18)* (pp. 15–20). Seoul, Republic of Korea: ACM.

Zakharov, E., Shysheya, A., Burkov, E., & Lempitsky, V. (2019). Few-shot adversarial learning of realistic neural talking head models. *Computer Vision and Pattern Recognition*. Advance online publication. Retrieved from https://arxiv.org/abs/1905.08233

Zlotowski, J. A., Sumioka, H., Nishio, S., Glas, D. F., Bartneck, C., & Ishiguro, H. (2018). Persistence of the uncanny valley. In H. Ishiguro, & F. Dalla Libera (Eds.), Geminoid Studies (pp. 163–187). Singapore, Singapore: Springer Nature Singapore.