Affective Man-Machine Interface: Unveiling Human Emotions through Biosignals

Egon L. van den Broek[1], Viliam Lisý[2], Joris H. Janssen[3,4],
Joyce H.D.M. Westerink[3], Marleen H. Schut[5], and Kees Tuinenbreijer[5]

[1] Center for Telematics and Information Technology (CTIT), University of Twente
P.O. Box 217, 7500 AE Enschede, The Netherlands
vandenbroek@acm.org
[2] Agent Technology Center, Dept. of Cybernetics, FEE, Czech Technical University
Technická 2, 16627 Praha 6, Czech Republic
viliam.lisy@agents.felk.cvut.cz
[3] User Experience Group, Philips Research
High Tech Campus 34, 5656 AE Eindhoven, The Netherlands
{joris.h.janssen,joyce.westerink}@philips.com
[4] Dept. of Human Technology Interaction, Eindhoven University of Technology
P.O. Box 513, 5600 MB, Eindhoven, The Netherlands
j.h.janssen@tue.nl
[5] Philips Consumer Lifestyle Advanced Technology
High Tech Campus 37, 5656 AE Eindhoven, The Netherlands
{marleen.schut,kees.tuinenbreijer}@philips.com

Abstract. As is known for centuries, humans exhibit an electrical profile. This profile is altered through various psychological and physiological processes, which can be measured through biosignals; e.g., electromyography (EMG) and electrodermal activity (EDA). These biosignals can reveal our emotions and, as such, can serve as an advanced man-machine interface (MMI) for empathic consumer products. However, such a MMI requires the correct classification of biosignals to emotion classes. This chapter starts with an introduction on biosignals for emotion detection. Next, a state-of-the-art review is presented on automatic emotion classification. Moreover, guidelines are presented for affective MMI. Subsequently, a research is presented that explores the use of EDA and three facial EMG signals to determine neutral, positive, negative, and mixed emotions, using recordings of 21 people. A range of techniques is tested, which resulted in a generic framework for automated emotion classification with up to 61.31% correct classification of the four emotion classes, without the need of personal profiles. Among various other directives for future research, the results emphasize the need for parallel processing of multiple biosignals.

That men are machines (whatever else they may be) has long been suspected; but not till our generation have men fairly felt in concrete just what wonderful psycho-neuro-physical mechanisms they are.

William James (1893; 1842 – 1910)

A. Fred, J. Filipe, and H. Gamboa (Eds.): BIOSTEC 2009, CCIS 52, pp. 21–47, 2010.

1 Introduction

Despite the early work of William James and others before him, it took more than a century before emotions became widely acknowledged and embraced by science and engineering. However, currently it is generally accepted that emotions cannot be ignored; they influence us, be it consciously or unconsciously, in a wide variety of ways [1]. Let us briefly denote four issues on how emotions influence our lives:

- long term physical well-being; e.g., Repetitive Strain Injury (RSI) [2], cardiovascular issues [3,4], and our immune system [5,6];
- physiological reactions / biosignals; e.g., crucial in communication [7,8,9,10];
- cognitive processes; e.g., perceiving, memory, reasoning [8,11]; and
- behavior; e.g., facial expressions [7,8,12].

As is illustrated by the three ways emotions influence us, we are (indeed) *psycho-neuro-physical mechanisms* [13,14], who both send and perceive biosignals that can be captured; e.g., electromyography (EMG), electrocardiography (ECG), and electrodermal activity (EDA). See Table 1 for an overview. These biosignals can reveal a plethora of characteristics of people; e.g., workload, attention, and emotions.

In this chapter, we will focus on biosignals that have shown to indicate people's emotional state. Biosignals form a promising alternative for emotion recognition compared to:

- facial expressions assessed through computer vision techniques [12,15,16]: recording and processing is notoriously problematic [16],
- movement analysis [15,17]: often simply not feasible in practice, and
- speech processing [12,18,19]: speech is often either absent or suffering from severe distortions in many real-world applications.

Moreover, biosignals have the advantage that they are free from social masking and have the potential of being measured by non-invasive sensors, making them suited for a wide range of applications [20,21]. Hence, such biosignals can act as a very useful interface between man and machine; e.g., computers or consumer products such as a mp3-player [22]. Such an advanced Man-Machine Interface (MMI) would provide machines with empathic abilities, capable of coping with the denoted issues.

In comparison to other indicators, biosignals have a number of methodological advantages as well. First of all, traditional emotion research uses interviews, questionnaires, and expert opinions. These, however, can only reveal subjective feelings, are very limited in explaining, and do not allow real time measurements: they can only be used before or after emotions are experienced [7,8,10]. Second, the recent progress in brain imaging techniques enables the inspection of brain activity while experiencing emotions; e.g., EEG and fMRI [11,28]. Although EEG techniques are slowly brought to practice; e.g., Brain Computer Interfacing (BCI) [29,30], these techniques are still very obtrusive. Hence, they are not usable in real world situations; e.g., for the integration in consumer products. As a way between these two research methods, psychophysiological (or bio)signals can be used [7,8,10,14]. These are not, or at least less, obtrusive, can be recorded and processed real time, are rich sources of information, and are relatively cheap to apply.

Table 1. An overview of common biosignals/ physiological signals and their features, as used for emotion analysis and classification

physiology	features	unit	remark
cardiovascular activity [23]	heart rate (HR)	beats / min	
through ECG or BVP	SD IBIs	s	heart rate variability (HRV) index
	RMSSD IBIs	s	heart rate variability (HRV) index
	LF power (0.05Hz - 0.15Hz)	ms^2	sympathetic activity
	HF power (0.15Hz - 0.40Hz)	ms^2	parasympathetic activity
	VLF power ($<$ 0.05Hz)	ms^2	
	LF / HF		
	pulse transit time (PTT)	ms	
electrodermal activity (EDA) [24]	mean, SD SCL	μS	tonic sympathetic activity
	number of SCRs		rate phasic activity
	SCR amplitude	μS	phasic activity
	SCR 1/2 recovery time	s	
	SCR rise time	s	
skin temperature [25]	mean, SD temp	$^oC(F)$	
respiration [25]	respiration rate		
	amplitude respirations		
	respiratory sinus arrythmia		
muscle activity	mean, SD corrugator supercilii	μV	frowning
through EMG [26, 27]	mean, SD zygomaticus major	μV	smiling
	mean, SD upper trapezius	μV	
	mean, SD inter-blink interval	ms	

Legend: ECG: electrocardiogram; BVP: blood volume pulse; EMG: electromyogram; IBI: inter-beat interval; LF: low frequency; HF: high frequency; VLF: very low frequency; SCL: skin conductance level; SCR: skin conductance response; SD: standard deviation; RMSSD: root mean sum of square differences. See also Fig. 3 for plots of the three facial EMG signals and the EDA signal.

A number of prerequisites should be taken into account when using either traditional methods (e.g., questionnaires), brain imaging techniques, or biosignals to infer people's emotional state. In Van den Broek et al. (2009), these are denoted for affective signal processing (ASP); however, most of them also hold for brain imaging, BCI, and traditional methods. The prerequisites include:

1. the validity of the research employed,
2. triangulation; i.e., using multiple information sources (e.g., biosignals) and/or analysis techniques, and
3. inclusion and exploitation of signal processing knowledge (e.g., determine the Nyquist frequencies of biosignals for emotion classification).

For a discussion on these topics, we refer to Van den Broek et al. (2009). Let us now assume that all prerequisites can be satisfied. Then, it is feasible to classify the biosignals in terms of emotions. In bringing biosignals-based emotion recognition to products, self-calibrating, and automatic classification is essential to make it useful for Artificial Intelligence (AI) [1,31], Ambient Intelligence (AmI) [20,32], MMI [7,33], and robotics [34,35].

In the pursuit toward empathic technology, we will describe our work on the automatic classification of biosignals. In the next section, we provide an overview of previous work. Section 3 provides an introduction to the classification techniques employed. Subsequently, in Sect. 4, we present the experiment in which we used four biosignals signals: three facial EMGs and EDA. After that, in Sect. 5, we will briefly introduce the preprocessing techniques employed. This is followed by Sect. 6 in which the classification results are presented. In Sect. 7, we reflect on our work and critically review it. Finally, in Sect. 8 we end with drawing the main conclusions.

2 Background

A broad range of biosignals are used in affective sciences; see Table 1. To enable processing of the signals, in most cases comprehensive sets of features have to be identified for each biosignal; see also Table 2. To extract these features, affective signals are processed in the time (e.g., statistical moments), frequency (e.g., Fourier), time-frequency (e.g., wavelets), or power domain (e.g., periodogram and autoregression) [36]. In Table 1, we provide a brief overview of the biosignals most often applied, including their best known features, with reference to their physiological source. In the next paragraph, we describe the signals and their psychological counterparts.

First, electrocardiogram (ECG; measured with electrodes on the chest) and blood volume pulse (BVP; measured with infra-red light around the finger or ear) can be used to derive heart beats. The main feature extracted from these heart beats is heart rate (HR; i.e., the number of beats per minute). HR is, however, not very useful in discriminating emotions as it is innervated by many different processes. Instead, the heart rate variability (HRV) provides better emotion information. HRV is more constant in situations where you are happy and relaxed, whereas it shows high variability in more stressful situations [20,55,56]. Second, respiration is often measured with a gauge band around the chest. Respiration rate and amplitude mediate the HRV and are, therefore,

Table 2. An overview of 20 studies on automatic classification of emotions, using biosignals / physiological signals

information source	year	signals	participants	number of features	selection / reduction	classifiers	target	classification result
[37] Sinha & Parsons	1996	$\mathcal{M}$	27	18		LDA	2 emotions	86%
[9] Picard et al.	2001	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{R}$,$\mathcal{M}$	1	40	SFS, Fisher	LDA	8 emotions	81%
[38] Scheirer et al.	2002	$\mathcal{C}$,$\mathcal{E}$	24	5	Viterbi	HMM	2 frustrations	64%
[39] Nasoz et al.	2003	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{S}$	31	3		k-NN, LDA	6 emotions	69%
[40] Takahashi	2003	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{B}$	12	18		SVM	6 emotions	42%
[41] Haag et al.	2004	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{S}$,$\mathcal{M}$,$\mathcal{R}$	1	13		MLP	valence / arousal	64 – 97%
[42] Kim et al.	2004	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{S}$	175	10		SVM	3 emotions	78%
[43] Lisetti & Nasoz	2004	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{S}$	29	12		k-NN, LDA, MLP	6 emotions	84%
[44] Wagner et al.	2005	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{R}$,$\mathcal{M}$	1	32	SFS, Fisher	k-NN, LDA, MLP	4 emotions	92%
[45] Yoo et al.	2005	$\mathcal{C}$,$\mathcal{E}$	6	5		MLP	4 emotions	80%
[46] Choi & Woo	2005	$\mathcal{E}$	1	3	PCA	MLP	4 emotions	75%
[47] Healey & Picard	2005	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{R}$,$\mathcal{M}$	9	22	Fisher	LDA	3 stress levels	97%
[34] Liu et al.	2006	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{M}$,$\mathcal{S}$	14	35		RT	3 anxiety levels	70%
[48] Rani et al.	2006	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{S}$,$\mathcal{M}$,$\mathcal{P}$	15	46		k-NN, SVM, RT, BN	3 emotions	86%
[49] Zhai & Barreto	2006	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{S}$,$\mathcal{P}$	32	11		SVM	2 stress levels	90%
[50] Jones & Troen	2007	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{R}$	13	11		ANN	5 arousal levels	31 / 62%
						5 valence levels	26 / 57%	
[51] Leon et al.	2007	$\mathcal{C}$,$\mathcal{E}$	8	5	DBI	AANN	3 emotions	71%
[52] Liu et al.	2008	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{S}$,$\mathcal{M}$	6	35		SVM	3 affect states	83%
[53] Katsis et al.	2008	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{M}$,$\mathcal{R}$	10	15		SVM, ANFIS	4 affect states	79%
[54] Yannakakis & Hallam	2008	$\mathcal{C}$,$\mathcal{E}$	72	20	ANOVA	SVM, MLP	2 fun levels	70%
[33] Kim & André	2008	$\mathcal{C}$,$\mathcal{E}$,$\mathcal{M}$,$\mathcal{R}$	3	110	SBS	LDA, DC	4 emotions	70 / 95%

Signals: C: cardiovascular activity; E: electrodermal activity; R: respiration; M: electromyogram; B: electroencephalogram; S: skin temperature; P: pupil diameter. classifiers: MLP: MultiLayer Perceptron; HMM: Hidden Markov Model; RT: Regression Tree; BN: Bayesian Network; ANN: Artificial Neural Network; AANN: Auto-Associative Neural Network; SVM: Support Vector Machine; LDA: Linear Discriminant Analysis; k-NN: k-Nearest Neighbors; ANFIS: Adaptive Neuro-Fuzzy Inference System; DBI: Davies-Bouldin Index; PCA: Principal Component Analysis; SFS: Sequential Forward Selection; SBS: Sequential Backward Selection; DC: Dichotomous Classification.

often used in combination, which is called respiratory sinus arrhythmia (RSA) [25]. RSA is primarily responsive to relaxation and emotion regulation [57]. Third, electrodermal activity (EDA) measures the skin conductance of the hands or foots. This is primarily a response to increases in arousal. Beside the general skin conductance level (SCL), typical peaks in the signal, called skin conductance responses (SCRs), can be extracted. These responses are more event related and are valuable when looking at short timescales. Fourth, skin temperature, measured at the finger, is also responsive to increases in arousal, but does not have the typical response peaks as EDA has. Finally, electromyogram (EMG) measures muscle tension. In relation to emotions, this is most often applied in the face, where it can measure smiling and frowning [58,59].

When processing such biosignals some general issues have to be taken in consideration:

1. Biosignals are typically derived through non-invasive methods to determine changes in physiology [21] and, as such, are indirect measures. Hence, a delay between the actual physiological change and the recorded change in the biosignal has to be taken into account.
2. Physiological sensors are unreliable; e.g., they are sensitive to movement artifacts and to differences in bodily position.
3. Some sensors are obtrusive, preventing their integration in real world applications [20,22].
4. Biosignals are influenced by (the interaction among) a variety of factors [36,60]. Some of these sources are located internally (e.g., a thought) and some are among the broad range of possible external factors (e.g., a signal outside). This makes affective signals inherently noisy, which is most prominent in real world applications.
5. Physiological changes can evolve in a matter of milliseconds, seconds, minutes or even longer. Some changes hold for only a brief moment, while others can even be permanent. Although seldom reported, the expected time windows of change are of interest [20,22]. In particular since changes can add to each other, even when having a different origin.
6. Biosignals have large individual differences. On the one hand, this calls for methods and models tailored to the individual. It has been shown that personal approaches increase the performance of ASP [20,33,50]. On the other hand, generic features are of the utmost importance. Not in all situations, a system or product can be calibrated. Moreover, directing the quest too fast towards people's personal profiles could diminish the interest in generic features and, consequently, limit the progress in research towards them.

The features obtained from the biosignals (see Table 1) can be fed to pattern recognition methods (see Table 2); cf. [29]. These can be classified as: template matching, syntactic or structural matching, and statistical classification; e.g., artificial neural networks (ANN). The former two are not or seldom used in ASP, most ASP schemes use the latter.

Statistical pattern recognition distinguishes supervised and unsupervised (e.g., clustering) pattern recognition; i.e., respectively, with or without a set of (labeled) training data [61,62,63]. With unsupervised pattern recognition, the distance / similarity measure used and the algorithm applied to generate the clusters are key elements.

Supervised pattern recognition relies on learning from a set of examples (i.e., the training set). Statistical pattern recognition uses input features, a discriminant function (or network function for ANN) to recognize the features, and an error criterion in its classification process.

In the field of ASP, several studies have been conducted, using a broad range of signals, features, and classifiers; see Table 2 for an overview. Nonetheless, both the recognition performance and the number of emotions that the classifiers were able to discriminate are disappointing. Moreover, comparing the different studies is problematic because of:

1. The different settings the studies were applied in, ranging from controlled lab studies to real world testing;
2. The type of emotion triggers used;
3. The number of target states to be discriminated; and
4. The signals and features employed.

To conclude, there is a lack of general standards, low prediction accuracy, and inconsistent results. However, for affective MMI to come to fruition, it is eminent to deal with these issues. This illustrates the need for a well documented general framework. In this chapter, we set out to initiate its development, explore various possibilities, and apply it on a data set that will be introduced in the next section.

3 Techniques for Classification

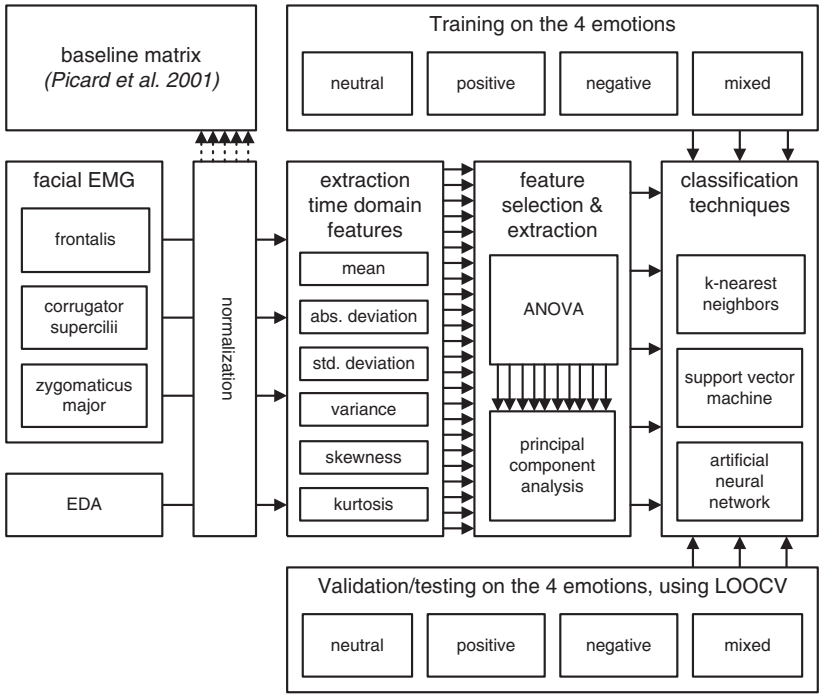
In this section, we briefly introduce the techniques used in the research conducted, for those readers who are not familiar with them. Figure 1 presents the complete processing scheme of this research. The core of processing scheme consists of three phases, in which various techniques were applied.

First, analysis of variance (ANOVA) and principal component analysis (PCA) are introduced that enabled the selection of a subset of features for the classification of the emotions. Second, the classification was done using k-nearest neighbors (k-NN), support vector machines (SVM), and artificial neural networks (ANN), which will be briefly introduced later in this section. Third and last, the classifiers were evaluated using leave-one-out cross validation (LOOCV), which will be introduced at the end of this section.

3.1 Analysis of Variance (ANOVA)

Analysis of variance (ANOVA) is a statistical test to determine whether or not there is a significant difference between the means of several data sets. ANOVA examines the variance of data set means compared to within class variance of the data sets themselves. As such, ANOVA can be considered as an extension of the t-test, which can only be applied on one or two data sets. We will sketch the main idea here. For a more detailed explanation, we refer to Chapter 6 of [64].

ANOVA assumes that the data sets are independent and randomly chosen from a normal distribution. Moreover, it assumes that all data sets are equally distributed. These



Legend: EMG: electromyography EDA: electrodermal activity; ANOVA of variance; LOOCV: leave-one-out cross validation.

Fig. 1. The complete processing scheme, as applied in the current research

assumptions usually hold with empirical data. Moreover, the test is fairly robust against limited violations.

Assume we have D data sets. For each data set d, the sum t_d and mean $\bar{s_d}$ of all samples are defined as:

$$t_d = \sum_{i=0}^{S-1} x_{id} \quad \text{and} \quad \bar{s_d} = \frac{t_d}{s_d}$$

where x_{id} denotes one data sample and s_d denotes the number of samples of data set d. Subsequently, the grand sum T and the total number of data samples S can be defined as:

$$T = \sum_{d=0}^{D-1} t_d \quad \text{and} \quad S = \sum_{d=0}^{D-1} s_d.$$

The total sum of squares SS (i.e., the quadratic deviation from the mean) can be written as the sum of two independent components:

$$SS_H = \sum_{d=0}^{D-1} \frac{t_d^2}{s_d^2} - \frac{T^2}{S} \quad \text{and} \quad SS_E = \sum_{d=0}^{D-1} \sum_{i=0}^{S-1} x_{id}^2 - \sum_{d=0}^{D-1} \frac{t_d^2}{s_d^2},$$

where indices $_H$ and $_E$ denote hypothesis and error, as is tradition in social sciences. Together with S and D, these components define the ANOVA statistic:

$$F(D-1, S-D) = \frac{S-D}{D-1} \cdot \frac{SS_H}{SS_E},$$

where $D-1$ and $S-D$ can be considered as the degrees of freedom.

The hypothesis that all data sets were drawn from the same distribution is violated if

$$F_\alpha(D-1, S-D) < F(D-1, S-D),$$

where F_α denotes the ANOVA statistic that accompanies chance level α, considered to be acceptable. Often α is chosen as either 0.05, 0.01, or 0.001. If $\alpha < 0.05$ the data sets are assumed to be different.

3.2 Principal Component Analysis (PCA)

Through principal component analysis (PCA), the dimensionality of a data set of interrelated variables can be reduced, preserving its variation as much as possible. The variables are transformed to a new set of uncorrelated but ordered variables: the principal components. The first principal component represents, as much as possible, the variance of the original variables. Each succeeding component represents the remaining variance, as much as possible. For a brief introduction on PCA, we refer to Chapter 12 of [64].

Suppose we have a set of data, each represented as a vector x, which consists of n variables. Then, the principal components are defined as a linear combination $\alpha \cdot x$ of the variables of x that preserves the maximum of the (remaining) variance, denoted as:

$$\alpha \cdot x = \sum_{i=0}^{n-1} \alpha_i x_i,$$

where $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{n-1})^T$. The variance covered by each principal component $\alpha \cdot x$ is defined as:

$$var(\alpha \cdot x) = \alpha \cdot C\alpha,$$

where C is the covariance matrix of x.

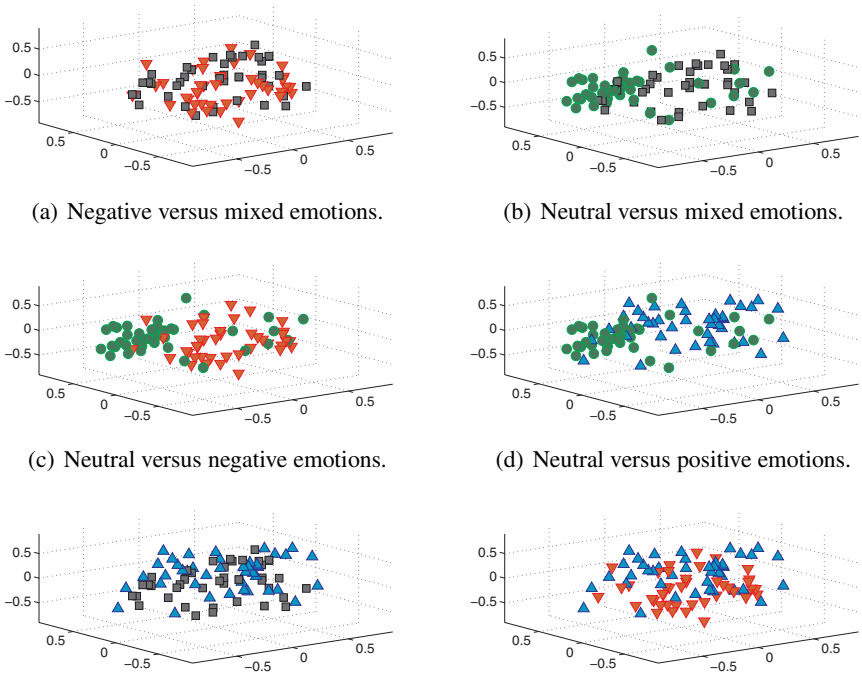
To find all principal components, we need to find the maximized $var(\alpha \cdot x)$ for them. Hereby, the constraint $\alpha \cdot \alpha = 1$ has to be taken into account. The standard approach to do so is the technique of Lagrange multipliers. We maximize

$$\alpha \cdot C\alpha - \lambda\left(\sum_{i=0}^{n-1} \alpha_i^2 - 1\right) = \alpha \cdot C\alpha - \lambda(\alpha \cdot \alpha - 1),$$

where λ is a Lagrange multiplier. Subsequently, we can derive that λ is an eigenvalue of C and α is its corresponding eigenvector.

Once obtained the vectors α, a transformation can be made that maps all data x to its principal components:

$$x \rightarrow (\alpha_0 \cdot x, \alpha_1 \cdot x, \ldots, \alpha_{n-1} \cdot x)$$



(a) Negative versus mixed emotions.

(b) Neutral versus mixed emotions.

(c) Neutral versus negative emotions.

(d) Neutral versus positive emotions.

(e) Positive versus mixed emotions.

(f) Positive versus negative emotions.

Fig. 2. Visualization of the first three principal components of all six possible combinations of two emotion classes. The emotion classes are plotted per two to facilitate the visual inspection.The plots illustrate how difficult it is to separate even two emotion classes, where separating four emotion classes is the aim. However, note that the emotion category neutral can be best separated from the other three categories: mixed, negative, and positive emotions, as is illustrated in b), c), and d).

Note that the principal components are sensitive to scaling. In order to tackle this problem, the components can be derived from the correlation matrix instead of the covariance matrix. This is equivalent to extracting the principal components in the described way after normalization of the original data set to unit variance.

PCA is also often applied for data inspection through visualization, where the principal components are chosen along the figure's axes. Figure 2 presents such a visualization: for each set of two emotion classes, of the total of four, a plot denoting the first three principal components is presented.

3.3 k-Nearest Neighbors (k-NN)

k-nearest neighbors (k-NN) is a very intuitive, simple, and often applied machine learning algorithm. It requires only a set of labeled examples (i.e., data vectors), which form the training set.

Now, let us assume that we have a training set x^l and a set of class labels C. Then, each new vector x_i from the data set is classified as follows:

1. Identify k vectors from x^l that are closest to vector x_i, according to a metric of choice; e.g., city block, Euclidean, or Mahalanobis distance.
2. Class c_i that should be assigned to vector x_i is determined by:

$$c_i = \underset{c \in C}{\operatorname{argmax}} \sum_{i=0}^{k-1} w_i \gamma(c, c_i^l),$$

where $\gamma(.)$ denotes a boolean function that returns 1 when $c = c_i^l$ and 0 otherwise and

$$w_i = \begin{cases} 1 & \text{if } \delta(x_i, x_i^l) = 0; \\ \frac{1}{d(x_i, x_i^l)^2} & \text{if } \delta(x_i, x_i^l) \neq 0, \end{cases}$$

where $\delta(.)$ denotes the distance between vectors x_i and x_i^l. Note that, if preferred, the factor weight can be simply eliminated by putting $w_i = 1$.

3. If there is a tie of two or more classes $c \in C$, vector x_i is randomly assigned to one of these classes.

The algorithm presented applies to k-NN for weighted, discrete classifications, as will be applied in the current research. However, a simple adaptation can be made to the algorithm, which enables continuous classifications. For more information on these and other issues, we refer to the various freely available tutorials and introductions that have been written on k-NN.

3.4 Support Vector Machine (SVM)

Using a suitable kernel function, a support vector machine (SVM) ensures the division of a set of data into two classes, with respect to the shape of the classifier and misclassification of the training samples. The main idea of SVM can be best explained with the example of a binary linear classifier.

Let us define our data set as:

$$D = \left\{(x_i, c_i) | x_i \in \mathbb{R}^d, c_i \in \{-1, 1\}\right\} \text{ for } i = 0, 1, \ldots, N-1,$$

where x_i is a vector with dimensionality d from the data set, which has size N. c_i is the class to which x_i belongs. To separate two classes, we need to formulate a separating hyperplane $w \cdot x = b$, where w is a normal vector of length 1, x is a feature vector, and b is a constant.

In practice, it is often not possible to find such a linear classifier. In this case, the problem can be generalized. Then, we need to find w and b so that we can optimize

$$c_i(w \cdot x_i + b) \leq \xi_i,$$

where ξ_i represents the deviation (or error) from the linearly separable case.

To determine an optimal plane, the sum of ξ_i must be minimized. The minimization of this parameter can be solved by Lagrange multipliers α_i. From the derivation of this

method, it is possible to see that often most of the α_is are equal to 0. The remaining relevant subset of the training data x is denoted as the support vectors. Subsequently, the classification is performed as:

$$f(x) = \text{sgn}\left(\sum_{i=0}^{S-1} c_i \alpha_i x \cdot x_i + b\right),$$

where S denotes the number of support vectors.

For a non-linear classification problem, we can replace the dot product by a non-linear kernel function. This enables the interpretation of algorithms geometrically in feature spaces non-linearly related to the input space and combine statistics and geometry. A kernel can be viewed as a (non-linear) similarity measure and induce representations of the data in a linear space. Moreover, the kernel implicitly determines the function class, which is used for learning [63].

The SVM introduced here classified samples in two classes. In the case of multiple classes, two approaches are common: 1) for each class, a classifier can be build that separates that class from the other data and 2) for each pair of classes, classifiers can be build. With both cases, voting paradigms are used to assign the data samples x_i to classes c_i. For more information on SVM, [62,63] can be consulted.

3.5 Artificial Neural Networks (ANN)

Artificial neural networks (ANN) are inspired by their biological counterparts. Often, ANN are claimed to have a similar behavior as biological neural networks. Although ANN share several features with biological neural networks (e.g., noise tolerance), this claim is hardly justified; e.g., a human brain consists of roughly 10^{11} brain cells, where an ANN consists of only a few dozens of units.

Nevertheless, ANN have proved their use for a range of pattern recognition and machine learning applications.

Moreover, ANN have a solid theoretical basis [61,62].

ANN consist of a layer of input units, one or more layers of hidden units, and a layer of output units. These units are connected with a weight w_{ij}, which determines the transfer of unit u_i to unit u_j. The activation level of a unit u_j is defined as:

$$a_j(t+1) = f(a_j(t), i_j(t)),$$

where t denotes time, $f(.)$ is the activation function that determines the new activation based on the current state $a(t)$ and its effective input, defined as:

$$i_j(t) = \sum_{i=0}^{U_j-1} a_i(t) w_{ij}(t) + \tau_j(t),$$

where $\tau_j(t)$ is a certain bias or offset and U_j denotes the number of units from which a unit u_j can receive input. Note that at the input layer of a ANN, the input comes from the environment; then, i is the environment instead of another unit.

On its own, each neuron of an ANN can only perform a simple task. In contrast, a network of units can approximate any function. Moreover, ANN cannot only process input, they can also learn from their input, either supervised or unsupervised. Although various learning rules have been introduced for ANN, most can be considered as being derived from Hebb's classic learning rule:

$$\Delta w_{ij} = \eta a_i a_j,$$

which defines the modification of the weight of connection (u_i, u_j). η is a positive constant. Its rationale is that w_{ij} should be increased with the simultaneous activation of both units and the other way around.

Various ANN topologies have been introduced. The most important ones are recurrent and feed-forward networks, whose units respectively do and do not form a directed cycle through feedback connections. In the current research, a feed-forward network is applied: the classic multilayer perceptron (MLP), as is more often used for emotion recognition purposes; see also Table 2. It incorporated the often adopted sigmoid-shaped function applied to $f(.)$:

$$\frac{1}{1 + e^{-a_j}}$$

Throughout the 60 years of their existence, a broad plethora of ANN have been presented, varying on a range of aspects; e.g., their topology, learning rules, and the choice of either synchronous or asynchronously updating of its units. More information on ANN can be found in various introductions on ANN.

3.6 Leave-One-Out Cross Validation (LOOCV)

Assume we have a classifier that is trained, using a part of the available data set: the training data. The training process optimizes the parameters of a classifier to make it fit the training data. To validate the classifier's performance, an independent sample of the same data set has to be used [61,62].

Cross validation deviates from the general validation scheme since it enables the validation of a classifier without the need of an explicit validation set. As such, it optimizes the size of the data set that can be used as training data.

Various methods of cross validation have been introduced. In this section, we will introduce leave-one-out cross validation (LOOCV), a frequently used method to determine the performance of classifiers. LOOCV is typically useful and, consequently, used in the analysis of (very) small data sets. It has been shown that LOOCV provides an almost unbiased estimate of the true generalization ability of a classifier. As such, it provides a good model selection criterion.

Assume we have a classifier (e.g., k-NN, a SVM, or an ANN) of which we want to verify its performance on a particular data set. This data set contains (partly) data samples x_i with known correct classifications c_i^l. Then, classifier's performance can be determined through LOOCV, as follows:

1. $\forall_i$ train a classifier C_i with the complete data set x, except x_i.
2. $\forall_i$ classify data sample x_i to a class c_i, using classifier C_i.

3. Compute the average error of the classifier through

$$\mathcal{E} = \frac{1}{D} \underset{c \in C}{\operatorname{argmax}} \sum_{i=0}^{D-1} \gamma(c_i, c_i^l),$$

where D denotes the number of data samples and $\gamma(.)$ denotes a boolean function, which returns 1 if $c_i = c_i^l$ and 0 otherwise. Note that $\frac{1}{D}$ can be omitted from the formula if no comparisons are made between data sets (with different sizes).

Instead of one data sample x_i, this validation scheme also allows a subset of the data to be put aside. Such a subset can, for example, consist of all data gathered of one person. This enables an accurate estimation of the classification error $\mathcal{E}$ on this unknown person.

The processing scheme as presented here can be adapted in various ways. For example, in addition to the boolean function $\gamma(.)$, a weight function could be used that expresses the resemblance between classes. Hence, not all misclassifications would be judged similarly.

All results reported in this chapter are determined through LOOCV, if not specified in another way. For more information on cross validation, LOOCV in particular, we refer to [62].

4 Recording Emotions

We conducted an experiment in which the subjects' emotions were elicited, using film fragments that are known to be powerful in eliciting emotions in laboratory settings; see also [58,59,65]. As biosignals, facial EMG and EDA were recorded. These are known to reflect emotions [66]; see also both Table 1 and Table 2. The research in which the data was gathered is already thoroughly documented in both [58] and [59]. Therefore, we will only provide a brief summary of it.

4.1 Participants

In the experiment, 24 subjects (20 females) participated (average age 43 years). Mainly females were solicited to participate since we expected a more and stronger facial emotion expression of females [67]. Consequently, a relative small number of males participated. The biosignal recordings of three subjects either failed or were distorted. Hence, the signals of 21 subjects remained for classification purposes.

4.2 Equipment and Materials

We selected 8 film fragments (120 sec. each) for their emotional content. For specifications of these film fragments, see [58,59]. The 8 film fragments were categorized as being neutral or triggering positive, negative, or mixed (i.e., simultaneous negative and positive; [68]) emotions; hence, 2 film fragments per emotion category. This categorization was founded on Russell's valence-arousal model, introduced in [69]. Note that the existence of mixed emotions, the way to determine them, and the method to analyze ratings of the possible mixed emotions is still a topic of debate; e.g., [20,22,68].

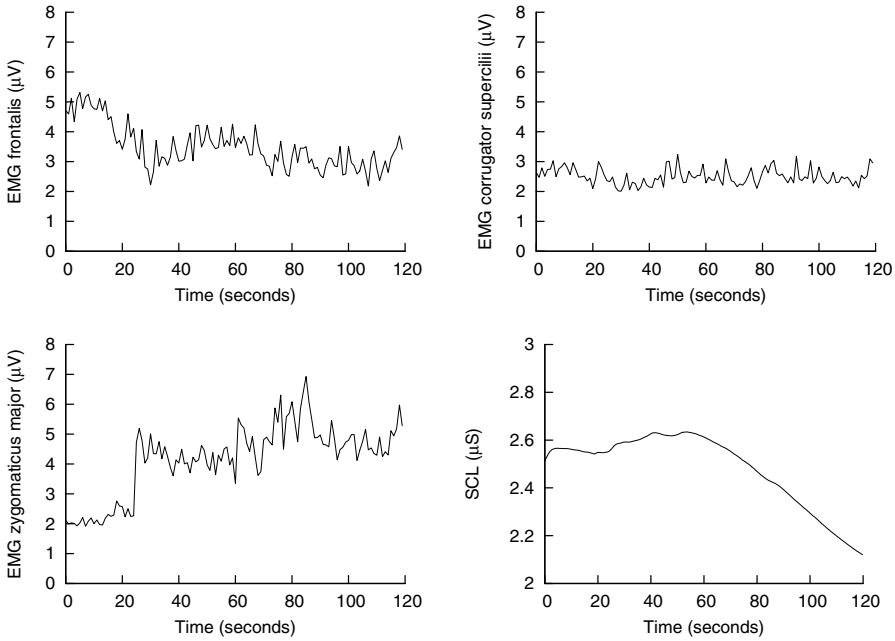


Fig. 3. Samples of the electromyography (EMG) in μV of the frontalis, the corrugator supercilii, and the zygomaticus major as well as of the electrodermal activity (EDA) in μV, denoted by the skin conductance level (SCL). All these signals were recorded in parallel, with the same person.

A TMS International Porti5-16/ASD system was used for the biosignal recordings, which was connected to a PC with TMS Portilab software[1]. Three facial EMGs were recorded: the right corrugator supercilii, the left zygomaticus major, and the left frontalis muscle. The EMG signals were high-pass filtered at 20 Hz, rectified by taking the absolute difference of the two electrodes, and average filtered with a time constant of 0.2 sec. The EDA was recorded using two active skin conductivity electrodes and average filtering with a time constant of about 2 sec. See Fig. 3 for samples of the three EMG signals and the EDA signal.

4.3 Procedure

After the participant was seated, the electrodes were attached and the recording equipment was checked. The 8 film fragments were presented to the participant in pseudo-random order. A plain blue screen was shown between the fragments for 120 seconds. This assured that the biosignals returned to their baseline level, before the next film fragment was presented.

After the viewing session, the electrodes were removed. Next, the participants answered a few questions regarding the film fragments viewed. To jog their memory, representative print-outs of each fragment were provided.

[1] URL of TMS Portilab software: http://www.tmsi.com/

5 Preprocessing

The quest towards self-calibrating algorithms for consumer products and for AmI and AI purposes gave some constraints to processing the signals. For example, no advanced filters should be needed, the algorithms should be noise-resistant, and should (preferably) also be able to handle corrupt data. Therefore, we chose to refrain from advanced preprocessing schemes and applied basic preprocessing. Figure 1 presents the complete processing scheme as applied in the current research.

5.1 Normalization

Humans are known for their rich variety in all aspects, this is no different for their biosignals. In developing generic classifiers, this required the normalization of the signals. This was expected to boost its performance significantly [48].

For each person, for all his signals, and for all their features separately, the following normalization was applied:

$$x_n = \frac{x_i - \bar{x}}{\sigma},$$

where x_n is the normalized value, x_i the recorded value, $\bar{x}$ the global mean, and σ the standard deviation.

Normalization of data (e.g., signals) has been broadly discussed. This has resulted in a variety of normalization functions; e.g., see [24,61,62].

5.2 Baseline Matrix

In their seminal article, Picard, Vyzas, and Healey (2001) introduced a baseline matrix for processing biosignals for emotion recognition. They suggested that this could tackle problems due to variation both within (e.g., inter day differences) and between participants. Regrettably, Picard et al. (2001) did not provide evidence for its working.

The baseline matrix requires biosignals recordings while people are in a neutral state. Regrettably, such recordings were not available. Alternatively, one of both available neutral film fragments was chosen [58,59].

In line with Picard et al. (2001), the input data was augmented with the baseline values of the same data set. A maximum performance improvement of 1.5% was achieved, using a k-NN classifier. Therefore, the baseline matrix was excluded in the final processing pipeline.

5.3 Feature Selection

To achieve good classification results with pattern recognition and machine learning, the set of input features is crucial. This is no different with classifying emotions [7,8,10]. As was denoted in Sect. 2, biosignals can be processed in the time, frequency, time-frequency, and power domain.

For EMG and EDA signals, the time domain is most often employed for feature extraction; see also Table 1. Consequently, we have chosen to explore a range of features from the time domain: mean, absolute deviation, standard deviation (SD), variance, skewness, and kurtosis. Among these are frequently used features (i.e., mean and SD)

Table 3. The best feature subsets from the time domain, for k-nearest neighbor (k-NN) classifier with Euclidean metric. They were determined by analysis of variance (ANOVA), using normalization per signal per participant. EDA denotes the electrodermal activity or skin conductance level.

feature	EDA	facial electromyography (EMG)		
		frontalis	corrugator supercilii	zygomaticus
mean				o
absolute deviation				o
standard deviation (SD)		o		o
variance		o		o
skewness	o		o	o
kurtosis			o	

Table 4. The recognition precision of the k-nearest neighbors (k-NN) classifier, with $k = 8$ and the Euclidean metric. The influence of three factors is shown: 1) normalization, 2) analysis of variance (ANOVA) feature selection (FS), and 3) Principal Component Analysis (PCA) transform.

normalization	no *fs*	ANOVA *fs* (10 features)	ANOVA *fs* & PCA (5 components)
no	45.54%		
yes	54.07%	60.71%	60.80%

and rarely used, but promising, features (i.e., skewness and kurtosis) [58,59]; see also Table 3.

To define an optimal set of features, a criterion function should be defined. However, no such criterion function was available in our case. Thus, an exhaustive search in all possible subsets of input features (i.e., 2^{24}) was required to guarantee an optimal set [70]. To limit this enormous search space, an ANOVA-based heuristic search was applied.

For both the normalizations, we performed feature selection based on ANOVAs. We selected the features with ANOVA $\alpha \leq 0.001$ (see also Sect. 3), as this led to the best precision. The features selected for each of the biosignals are presented in Table 3.

The last step of preprocessing was PCA; see also Sect. 3. The improvement of the PCA was small compared to feature selection solely. However, it was positive for normalization; see also Table 4. Figure 2 presents for each set of two emotion classes, of the total of four, a plot denoting the first three principal components. As such, the six resulting plots illustrate the complexity of separating the emotion classes.

6 Classification Results

This section reports the results of the three classification techniques applied: k-nearest neighbors (k-NN), support vector machines (SVM), and artificial neural networks

Table 5. Confusion matrix of the k-NN classifier of EDA and EMG signals for the best reported input preprocessing, with a cityblock metric and $k = 8$

		real			
		neutral	positive	mixed	negative
classified	neutral	71.43%	19.05%	9.52%	14.29%
	positive	9.52%	57.14%	9.52%	21.43%
	mixed	4.76%	4.76%	64.29%	11.90%
	negative	14.29%	19.05%	16.67%	52.38%

(ANN); see also Sect. 3. In all cases, the features extracted from the biosignals were used to classify participants' neutral, positive, negative, or mixed state of emotion; see also Fig. 2. For the complete processing scheme, we refer to Fig. 1.

6.1 k-Nearest Neighbors (k-NN)

For our experiments, we have used MATLAB[2] and a k-NN implementation, based on SOM Toolbox 2.0[3]. Besides the classification algorithm described in Sect. 3.3, we have used a modified version, more suitable for calculating the recognition rates. Its output was not the resulting class, but a probability of classification to each of the classes. This means that if there is a single winning class, the output is 100% for the winning class and 0% for all the other classes. If there is a tie of multiple classes, the output is divided among them and 0% is provided to the rest. All the recognition rates of the k-NN classifier reported in the current study were obtained by using this modified algorithm.

A correct metric is a crucial part of a k-NN classifier. A variety of metrics provided by the `pdist` function in MATLAB[2] was applied. Different feature subsets appeared to be optimal for different classes. Rani et al. (2006) denoted the same issue in their empirical review; cf. Table 3. The results of the best preprocessed input with respect to the four emotion classes (i.e., neutral, positive, negative, and mixed) is 61.31%, with a cityblock metric and $k = 8$; cf. Table 4.

Probability tables for the different classifications given a known emotion category are quite easy to obtain. They can be derived from confusion matrices of the classifiers by transforming the frequencies to probabilities. Table 5 presents the confusion matrix of the k-NN classifier used in this research, with a cityblock metric and $k = 8$.

6.2 Support Vector Machines (SVM)

We have used MATLAB[2] environment and a SVM and kernel methods toolbox[4], for experimenting with SVMs. We used input enhanced with the best preprocessing, described in the previous section. It was optimized for the k-NN classifier; however, we

[2] MATLAB online: http://www.mathworks.com/products/matlab/

[3] The MATLAB SOM Toolbox 2.0 is available through: http://www.cis.hut.fi/projects/somtoolbox

[4] The SVM and kernel methods toolbox is available through: http://asi.insa-rouen.fr/enseignants/ arakotom/toolbox/

expected it to be a good input also for more complex classifiers, including SVM. This assumption was supported by several tests with various normalizations. Hence, the signals were normalized per person, see also Sect. 5. After feature selection, the first 5 principal components from the PCA transformation were selected, see also Sect. 3.

The kernel function of SVM characterizes the shapes of possible subsets of inputs classified into one category [63]. Being SVM's similarity measure, the kernel function is the most important part of an SVM; see also Sect. 3. We applied both a polynomial kernel, with dimensionality d, defined as:

$$k_P(x_i, x^l) = \left(x_i \cdot x^l\right)^d$$

and a Gaussian (or radial basis function) kernel, defined as:

$$k_G(x_i, x^l) = \exp\left(-\frac{|x_i - x^l|^2}{2\sigma^2}\right),$$

where x_i is a feature vector that has to classified and x^l is a feature vector assigned to a class (i.e., the training sample) [63].

A Gaussian kernel ($\sigma = 0.7$) performed best with 60.71% correct classification. However, a polynomial kernel with $d = 1$ had a similar classification performance (58.93%). All the results were slightly worse than with the k-NN classifier.

6.3 Artificial Neural Networks (ANN)

We have used a multi-layer perceptron (MLP) trained by a back-propagation algorithm that was implemented in the neural network toolbox of MATLAB[2]; see also Sect. 3. It used gradient descent with moment and adaptive training parameter. We have tried to recognize only the inputs that performed best with the k-NN classifier.

In order to assess what topology of ANN was most suitable for the task, we conducted small experiments with both 1 and 2 hidden layers. In both cases, we did try 5 to 16 neurons within each hidden layer. All of the possible $12 + 12 \times 12$ topologies were trained, each with 150 cycles and tested using LOOCV.

The experiments using various network topologies supported the claim from [71] that bigger ANN do not always tend to over fit the data. The extra neurons were simply not used in the training process. Consequently, the bigger networks showed good generalization capabilities but did not outperform the smaller ones. A MLP with 1 hidden layer of 12 neurons showed to be the optimal topology.

An alternative method for stopping the adaptation of the ANN is using validation data. For this reason, the data set was split into 3 parts: 1 subject for testing, 3 subjects for validation, and 17 subjects for training. The testing subject was completely removed from the training process at the beginning. The network was trained using 17 randomly chosen training subjects. At the end of each training iteration, the network was tested on the 3 validation subjects.

This procedure led to a 56.19% correct classification of the four emotion classes.

6.4 Reflection on the Results

Throughout the last decade, various studies have been presented with similar aims. Some of these studies reported good results on the automatic classification of biosignals that should unveil people's emotions; see Table 2. For example, Picard et al. (2001) reports 81% correct classification on the emotions of one subject [9]. Haag et al. (2004) reports 64%–97% correct classification, using a band function with bandwidth 10% and 20%. This study was conducted on one subject. This study reports promising results but also lacks the necessary details needed for its replication [41]. More recently, Kim and André (2008) reported a recognition accuracy of 95% and 70% for subject-dependent and subject-independent classification. Their study included three subjects [33].

In comparison with [9,41,33], this research incorporated data of a large number of people (i.e., 21), with the aim to develop a generic processing framework. At first glance, with average recognition rates of 60.71% for SVM and 61.31% for k-NN and only 56.19% for ANN, its success is questionable. However, the classification rates differ among the four emotion categories, as is shown in Table 5, which presents the confusion matrix of the results of the k-NN classifier. Neutral emotional states are recognized best, with a classification rate of 71.43%. Negative emotional states are the most complex to distinguish from the other three emotion categories, as is marked by its 52.38% correct classification rate. The complexity of separating the four emotion classes from each other is illustrated in Fig. 2.

Taking in consideration the generic processing pipeline (see also Fig. 1) and the limitations of other comparable research (cf. Table 2), the results reported in this chapter should be judged as (at least) reasonably good. Moreover, a broad range of improvements are possible. One of them would be to question the need of identifying specific emotions, using biosignals for MMI. Hence, the use of alternative, rather rough categorizations, as used in the current research, should be further explored.

With pattern recognition and machine learning, preprocessing of the data is crucial. This phase could also be improved for the biosignals used in the current study. First of all, we think that the feature selection based on an ANOVA was not sufficient for more complex classifiers such as neural networks. The ANOVA tests gathered the centers of random distributions that would generate the data of different categories; hereby assuming that their variances were the same. However, a negative result of this test is not enough to decide that a feature did not contain any information. As an alternative for feature selection, the k-NN classifier could be extended by a metric that would weigh the features, instead of omitting the confusing or less informative features.

Taken it all together, the quest towards affective MMI continues. Although the results presented are good compared to related work, it is hard to estimate whether or not the classification performance is sufficient for embedding of affective MMI in real world applications. However, the future is promising with the rapidly increasing amount of resources allocated for affective MMI and the range of improvements that are possible. This assures that the performance on classification of emotions will achieve the necessary further improvements.

7 Discussion

This chapter has positioned *men as machines* in the sense that they are *psycho-neuro-physical mechanisms* [13]. It has to be said that this is a far from new position; it is already known for centuries, although it was rarely exploited in application oriented research. However, in the last decade interest has increased and subareas evolved that utilized this knowledge. This chapter concerns one of them: affective MMI; or as Picard (1997) coined it: affective computing.

A literature overview is provided of the work done so far, see also Table 1 and Table 2. In addition, some guidelines on affective MMI are provided; see Sects. 1 and 2. To enable the recognition of these emotions, they had to be classified. Therefore, a brief description was provided of the classification techniques used (Sect. 3). Next, a study is introduced in which three EMG signals and people's EDA were measured (see also Fig. 3), while being exposed to emotion inducing film fragments; see Sect. 4. See Fig. 1 for an overview of the processing scheme applied in the current research. Subsequently, preprocessing and the automatic classification of biosignals, using the four emotion categories, were presented in Sect. 5 and Sect. 6.

Also in this research, the differences among participants became apparent. They can be denoted on four levels; see also Sect. 1. People have different physiological reactions on the same emotions and that people experience different emotions with the same stimuli (e.g., music or films). Moreover, these four levels interact [7,8,14]. Although our aim was to develop a generic model, one could question whether or not this can be realized. Various attempts have been made to determine people's personal biosignals-profile; e.g., [9,14,33,48]. However, no generally accepted standard has been developed so far.

In pursuit to generic affective MMI processing schemes, the notion of time should be taken into consideration, as was already denoted in Sect. 2. This can help to distinguish between emotions, moods, and personality [20,72,73]:

1. Emotion: A short reaction (i.e., a matter of seconds) to the perception of a specific (external or internal) event, accompanied by mental, behavioral, and physiological changes [7,10].
2. Moods: A long lasting state, gradually changing, in terms of minutes, hours, or even longer. They are experienced without concurrent awareness of their origin and are not object related. Moods do not directly affect actions; however, they do influence our behavior indirectly [7,10,74].
3. Personality: People's distinctive traits and behavioral and emotional characteristics. For example, introvert and extrovert persons express their emotions in distinct ways. Additionally, also self-reports and physiological indicators / biosignals will be influenced by people's personality trait [19,75].

With respect to processing the biosignals, the current research could be extended by a more detailed exploration of the time windows; e.g., with a span of 10 seconds [7,8,10,22]. Then, data from different time frames can be combined and different, better suitable normalizations could be applied to create new features. For example, information concerning the behavior of the physiological signals could be more informative than only the integral features from a large time window. Studying short time

frames could also provide a better understanding on the relation between emotions and their physiological correlates / biosignals, see also Table 1.

Other more practical considerations should also be noted. The advances made in wearable computing and sensors facilitates (affective) MMI; e.g., [21]. Last years, various prototypes have been developed, which enable the recording of physiological signals; e.g., [76]. This enables the recordings of various biosignals in parallel. In this way, an even higher probability of correct interpretation can be achieved [7,8,20].

Affective MMI can extent consumer products [22]. For example, a mp3-player could sense its listener's emotions and either provide suggestions for other music or automatically adapt its playing list to these emotions. In addition, various other applications have been proposed, mockups have been presented, and implementations have been made. Three examples of these are clothes with wearable computing, games that tweak its behavior and presentation depending on your emotions, and lighting that reacts on or adapts to your mood.

Affective signal processing (ASP) could possibly bring salvation to AI [1,20]. With understanding and sensing emotions, true AI is possibly (and finally) within reach. Current progress in biomedical and electrical engineering provide the means to conduct affective MMI in an unobtrusive manner and, consequently, gain knowledge about our natural behavior, a prerequisite for modeling it. As AI's natural successor, for AmI [20], even more than for AI, emotions play a crucial role in making it a success. Since AmI was coined by Emile Aarts [32], this has been widely acknowledged and repeatedly stressed; e.g., [20,32].

An extension of MMI is human-robot interaction. With robotics, embodiment is a key factor. Potentially, robots are able to enrich their communication substantially through showing some empathy from time to time. As with AI and AmI, this requires sensing and classification of emotions, as can be conveniently done through biosignals [34,35].

Of interest for affective MMI are also the developments in brain-computer interfacing (BCI) [29,30]. In time, affective BCI will possibly become within science's reach. Affective BCI, but also BCI in general, could advance AI, AmI, and human-robot interaction. Slowly this becomes acknowledged, as is illustrated by a workshop on affective BCI, as was held at the IEEE 2009 International Conference on Affective Computing and Intelligent Interaction[5]. With affective BCI, again both its scientific foundation and its applications will be of interest.

Without any doubt affective MMI has a broad range of applications and can help in making various areas more successful. Taking it all together, the results gathered in this research are promising. However, the correct classification rate is below that what is needed for reliable affective MMI in practice. Providing the range of factors that can be improved, one should expect that the performance can be boosted substantially. That this is not already achieved is not a good sign; perhaps, still some essential mistakes are made. One of the mistakes could be the computationally driven approach. A processing scheme that is founded on or at least inspired by knowledge from both biology, in particular physiology, and psychology could possibly be more fruitful . . .

[5] The IEEE 2009 International Conference on Affective Computing and Intelligent Interaction: http://www.acii2009.nl/

8 Conclusions

Affective MMI through biosignals is perhaps the ultimate blend of biomedical engineering, psychophysiology, and AI. However, in its pursuit, various other disciplines (e.g., electrical engineering and psychology) should not be disregarded. In parallel, affective MMI promotes the quest towards its scientific foundation and screams for its application [7,8,10]. As such, it is next generation science and engineering, which truly bridges the gap between man and machine.

As can be derived from this chapter, still various hurdles have to be taken in the development of a generic, self-calibrating, biosignal-driven classification framework for affective MMI. The research and the directives denoted here could help in taking some of these hurdles. When the remaining ones will also be taken; then, in time, the common denominators of people's biosignals can be determined and their relation with experienced emotions can be further specified. This would mark a new, biosignal-driven, era of advanced, affective MMI.

Acknowledgements. The authors thank Leon van den Broek (Radboud University Nijmegen, The Netherlands / University of Utrecht, The Netherlands), Frans van der Sluis (University of Twente, The Netherlands), and Marco Tiemann (Philips Research, The Netherlands) for their reviews of this book chapter. Furthermore, we thank the editors for inviting us to write a chapter for their book.

References

1. Picard, R.W.: Affective Computing. MIT Press, Boston (1997)
2. van Tulder, M., Malmivaara, A., Koes, B.: Repetitive strain injury. The Lancet 369(9575), 1815–1822 (2007)
3. Schuler, J.L.H., O'Brien, W.H.: Cardiovascular recovery from stress and hypertension factors: A meta-analytic view. Psychophysiology 34(6), 649–659 (1997)
4. Frederickson, B.L., Manusco, R.A., Branigan, C., Tugade, M.M.: The undoing effect of positive emotions. Motivation and Emotion 24(4), 237–257 (2000)
5. Ader, R., Cohen, N., Felten, D.: Psychoneuroimmunology: Interactions between the nervous system and the immune system. The Lancet 345(8942), 99–103 (1995)
6. Solomon, G.F., Amkraut, A.A., Kasper, P.: Immunity, emotions, and stress with special reference to the mechanisms of stress effects on the immune system. Psychotherapy and Psychosomatics 23(1-6), 209–217 (1974)
7. Fairclough, S.H.: Fundamentals of physiological computing. Interacting with Computers 21(1-2), 133–145 (2009)
8. Mauss, I.B., Robinson, M.D.: Measures of emotion: A review. Cognition and Emotion 23(2), 209–237 (2009)
9. Picard, R.W., Vyzas, E., Healey, J.: Toward machine emotional intelligence: Analysis of affective physiological state. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(10), 1175–1191 (2001)
10. van den Broek, E.L., Janssen, J.H., Westerink, J.H.D.M., Healey, J.A.: Prerequisits for Affective Signal Processing (ASP). In: Encarnação, P., Veloso, A. (eds.) Biosignals 2009: Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing, Porto – Portugal, pp. 426–433 (2009)

11. Critchley, H.D., Elliott, R., Mathias, C.J., Dolan, R.J.: Neural activity relating to generation and representation of galvanic skin conductance responses: A functional magnetic resonance imaging study. The Journal of Neuroscience 20(8), 3033–3040 (2000)
12. Zeng, Z., Pantic, M., Roisman, G.I., Huang, T.S.: A survey of affect recognition methods: Audio, visual, and spontaneous expressions. IEEE Transactions on Pattern Analysis and Machine Intelligence 31(1), 39–58 (2009)
13. James, W.: Review: La pathologie des emotions by Ch. Féré. The Philosophical Review 2(3), 333–336 (1893)
14. Marwitz, M., Stemmler, G.: On the status of individual response specificity. Psychophysiology 35(1), 1–15 (1998)
15. Gunes, H., Piccardi, M.: Automatic temporal segment detection and affect recognition from face and body display. IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics 39(1), 64–84 (2009)
16. Whitehill, J., Littlewort, G., Fasel, I., Bartlett, M., Movellan, J.: Towards practical smile detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 31(11), 2106–2111 (2009)
17. Daly, A.: Movement analysis: Piecing together the puzzle. TDR – The Drama Review: A Journal of Performance Studies 32(4), 40–52 (1988)
18. Ververidis, D., Kotropoulos, C.: Emotional speech recognition: Resources, features, and methods. Speech Communication 48(9), 1162–1181 (2006)
19. Van den Broek, E.L.: Emotional Prosody Measurement (EPM): A voice-based evaluation method for psychological therapy effectiveness. Studies in Health Technology and Informatics (Medical and Care Compunetics) 103, 118–125 (2004)
20. van den Broek, E.L., Schut, M.H., Westerink, J.H.D.M., Tuinenbreijer, K.: Unobtrusive Sensing of Emotions (USE). Journal of Ambient Intelligence and Smart Environments 1(3), 287–299 (2009)
21. Gamboa, H., Silva, F., Silva, H., Falcão, R.: PLUX – Biosignals Acquisition and Processing (2010), `http://www.plux.info` (Last accessed January 30, 2010)
22. van den Broek, E.L., Westerink, J.H.D.M.: Considerations for emotion-aware consumer products. Applied Ergonomics 40(6), 1055–1064 (2009)
23. Berntson, G.G., Bigger, J.T., Eckberg, D.L., Grossman, P., Kaufmann, P.G., Malik, M., Nagaraja, H.N., Porges, S.W., Saul, J.P., Stone, P.H., van der Molen, M.W.: Heart rate variability: Origins, methods, and interpretive caveats. Psychophysiology 34(6), 623–648 (1997)
24. Boucsein, W.: Electrodermal activity. Plenum Press, New York (1992)
25. Grossman, P., Taylor, E.W.: Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. Biological Psychology 74(2), 263–285 (2007)
26. Fridlund, A.J., Cacioppo, J.T.: Guidelines for human electromyographic research. Psychophysiology 23(5), 567–589 (1986)
27. Reaz, M.B.I., Hussain, M.S., Mohd-Yasin, F.: Techniques of EMG signal analysis: detection, processing, classification and applications. Biological Procedures Online 8(1), 11–35 (2006)
28. Grandjean, D., Scherer, K.R.: Unpacking the cognitive architecture of emotion processes. Emotion 8(3), 341–351 (2008)
29. Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., Arnaldi, B.: A review of classification algorithms for EEG-based brain-computer interfaces. Journal of Neural Engineering 4(2), R1–R13 (2007)
30. Bimber, O.: Brain-Computer Interfaces. IEEE Computer 41(10) (2008); [special issue]
31. Minsky, M.: The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind. Simon & Schuster, New York (2006)
32. Aarts, E.: Ambient intelligence: Vision of our future. IEEE Multimedia 11(1), 12–19 (2004)

33. Kim, J., André, E.: Emotion recognition based on physiological changes in music listening. IEEE Transactions on Pattern Analysis and Machine Intelligence 30(12), 2067–2083 (2008)
34. Liu, C., Rani, P., Sarkar, N.: Human-robot interaction using affective cues. In: Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2006), Hatfield, UK, pp. 285–290. IEEE Computer Society, Los Alamitos (2006)
35. Rani, P., Sims, J., Brackin, R., Sarkar, N.: Online stress detection using psychophysiological signals for implicit human-robot cooperation. Robotica 20(6), 673–685 (2002)
36. Cacioppo, J.T., Tassinary, L.G., Berntson, G.: Handbook of Psychophysiology, 3rd edn. Cambridge University Press, New York (2007)
37. Sinha, R., Parsons, O.A.: Multivariate response patterning of fear. Cognition and Emotion 10(2), 173–198 (1996)
38. Scheirer, J., Fernandez, R., Klein, J., Picard, R.W.: Frustrating the user on purpose: A step toward building an affective computer. Interacting with Computers 14(2), 93–118 (2002)
39. Nasoz, F., Alvarez, K., Lisetti, C.L., Finkelstein, N.: Emotion recognition from physiological signals for presence technologies. International Journal of Cognition, Technology and Work 6(1), 4–14 (2003)
40. Takahashi, K.: Remarks on emotion recognition from bio-potential signals. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Palmerston North, New Zealand, October 5-8, vol. 2, pp. 1655–1659 (2003)
41. Haag, A., Goronzy, S., Schaich, P., Williams, J.: Emotion recognition using bio-sensors: First steps towards an automatic system. In: André, E., Dybkjær, L., Minker, W., Heisterkamp, P. (eds.) ADS 2004. LNCS (LNAI), vol. 3068, pp. 36–48. Springer, Heidelberg (2004)
42. Kim, K.H., Bang, S.W., Kim, S.R.: Emotion recognition system using short-term monitoring of physiological signals. Medical & Biological Engineering & Computing 42(3), 419–427 (2004)
43. Lisetti, C.L., Nasoz, F.: Using noninvasive wearable computers to recognize human emotions from physiological signals. EURASIP Journal on Applied Signal Processing 2004(11), 1672–1687 (2004)
44. Wagner, J., Kim, J., André, E.: From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification. In: Proceedings of the IEEE International Conference on Multimedia and Expo. (ICME), Amsterdam, The Netherlands, July 6-8, pp. 940–943 (2005)
45. Yoo, S.K., Lee, C.K., Park, J.Y., Kim, N.H., Lee, B.C., Jeong, K.S.: Neural network based emotion estimation using heart rate variability and skin resistance. In: Wang, L., Chen, K., S. Ong, Y. (eds.) ICNC 2005. LNCS, vol. 3610, pp. 818–824. Springer, Heidelberg (2005)
46. Choi, A., Woo, W.: Physiological sensing and feature extraction for emotion recognition by exploiting acupuncture spots. In: Tao, J., Tan, T., Picard, R.W. (eds.) ACII 2005. LNCS, vol. 3784, pp. 590–597. Springer, Heidelberg (2005)
47. Healey, J.A., Picard, R.W.: Detecting stress during real-world driving tasks using physiological sensors. IEEE Transactions on Intelligent Transportation Systems 6(2), 156–166 (2005)
48. Rani, P., Liu, C., Sarkar, N., Vanman, E.: An empirical study of machine learning techniques for affect recognition in human-robot interaction. Pattern Analysis & Applications 9(1), 58–69 (2006)
49. Zhai, J., Barreto, A.: Stress detection in computer users through noninvasive monitoring of physiological signals. Biomedical Science Instrumentation 42, 495–500 (2006)
50. Jones, C.M., Troen, T.: Biometric valence and arousal recognition. In: Thomas, B.H. (ed.) Proceedings of the Australasian Computer-Human Interaction Conference (OzCHI), Adelaide, Australia, pp. 191–194 (2007)
51. Leon, E., Clarke, G., Callaghan, V., Sepulveda, F.: A user-independent real-time emotion recognition system for software agents in domestic environments. Engineering Applications of Artificial Intelligence 20(3), 337–345 (2007)

52. Liu, C., Conn, K., Sarkar, N., Stone, W.: Physiology-based affect recognition for computer-assisted intervention of children with Autism Spectrum Disorder. International Journal of Human-Computer Studies 66(9), 662–677 (2008)
53. Katsis, C.D., Katertsidis, N., Ganiatsas, G., Fotiadis, D.I.: Toward emotion recognition in car-racing drivers: A biosignal processing approach. IEEE Transactions on Systems, Man, and Cybernetics–Part A: Systems and Humans 38(3), 502–512 (2008)
54. Yannakakis, G.N., Hallam, J.: Entertainment modeling through physiology in physical play. International Journal of Human-Computer Studies 66(10), 741–755 (2008)
55. Task Force: Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. European Heart Journal 17(3), 354–381 (1996)
56. Ravenswaaij-Arts, C.M.A.V., Kollee, L.A.A., Hopman, J.C.W., Stoelinga, G.B.A., Geijn, H.P.: Heart rate variability. Annals of Internal Medicine 118(6), 436–447 (1993)
57. Butler, E.A., Wilhelm, F.H., Gross, J.J.: Respiratory sinus arrhythmia, emotion, and emotion regulation during social interaction. Psychophysiology 43(6), 612–622 (2006)
58. van den Broek, E.L., Schut, M.H., Westerink, J.H.D.M., van Herk, J., Tuinenbreijer, K.: Computing emotion awareness through facial electromyography. In: Huang, T.S., Sebe, N., Lew, M., Pavlović, V., Kölsch, M., Galata, A., Kisačanin, B. (eds.) ECCV 2006 Workshop on HCI. LNCS, vol. 3979, pp. 52–63. Springer, Heidelberg (2006)
59. Westerink, J.H.D.M., van den Broek, E.L., Schut, M.H., van Herk, J., Tuinenbreijer, K.: 14. In: Computing emotion awareness through galvanic skin response and facial electromyography. Philips Research Book Series, vol. 8, pp. 137–150. Springer, Dordrecht (2008)
60. Cacioppo, J., Tassinary, L.: Inferring psychological significance from physiological signals. American Psychologist 45(1), 16–28 (1990)
61. Mitchell, T.M.: Machine Learning. The McGraw-Hill Companies, Inc., Columbus (1997)
62. Bishop, C.M.: Pattern Recognition and Machine Learning. Information Science and Statistics. Springer Science+Business Media, LLC, New York (2006)
63. Schölkopf, B., Smola, A.J.: Learning with kernels: Support Vector Machines, Regularization, Optimization, and Beyond. In: Adaptive Computation and Machine Learning. The MIT Press, Cambridge (2002)
64. Rencher, A.C.: Methods of Multivariate Analysis, 2nd edn. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., New York (2002)
65. Rottenberg, J., Ray, R.R., Gross, J.J.: 1. In: Emotion elicitation using films, pp. 9–28. Oxford University Press, New York (2007)
66. Kreibig, S.D., Wilhelm, F.H., Roth, W.T., Gross, J.J.: Cardiovascular, electrodermal, and respiratory response patterns to fear- and sadness-inducing films. Psychophysiology 44(5), 787–806 (2007)
67. Kring, A.M., Gordon, A.H.: Sex differences in emotion: Expression, experience, and physiology. Journal of Personality and Social Psychology 74(3), 686–703 (1998)
68. Carrera, P., Oceja, L.: Drawing mixed emotions: Sequential or simultaneous experiences? Cognition & Emotion 21(2), 422–441 (2007)
69. Russell, J.A.: A circumplex model of affect. Journal of Personality and Social Psychology 39(6), 1161–1178 (1980)
70. Cover, T.M., van Campenhout, J.M.: On the possible orderings in the measurement selection problem. IEEE Transactions on Systems, Man, and Cybernetics SMC-7(9), 657–661 (1977)
71. Lawrence, S., Giles, C.L., Tsoi, A.: What size neural network gives optimal generalization? Convergence properties of backpropagation. Technical Report UMIACS-TR-96-22 and CS-TR-3617 (April 1996)
72. Barrett, L.F.: Valence as a basic building block of emotional life. Journal of Research in Personality 40, 35–55 (2006)

73. Russel, J.A., Barrett, L.F.: Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. Journal of Personality and Social Psychology 26(5), 805–819 (1999)
74. Gendolla, G.H.E.: On the impact of mood on behavior: An integrative theory and a review. Review of General Psychology 4(4), 378–408 (2000)
75. Cooper, C.L., Pervin, L.A.: Personality: Critical concepts in psychology, 1st edn. Critical concepts in psychology. Routledge, New York (1998)
76. Lukowicz, P.: Wearable computing and artificial intelligence for healthcare applications. Artificial Intelligence in Medicine 42(2), 95–98 (2008)