Biomedical Signal Processing and Control 31 (2017) 89–101

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Stress and anxiety detection using facial cues from videos

G. Giannakakis[a,*], M. Pediaditis[a], D. Manousos[a], E. Kazantzaki[a], F. Chiarugi[a], P.G. Simos[b,a], K. Marias[a], M. Tsiknakis[a,c]

[a] Foundation for Research and Technology—Hellas, Institute of Computer Science, Heraklion, Crete, Greece
[b] University of Crete, School of Medicine, Division of Psychiatry, Heraklion, Crete, Greece
[c] Technological Educational Institute of Crete, Department of Informatics Engineering, Heraklion, Crete, Greece

ARTICLE INFO

Article history:
Received 22 March 2016
Received in revised form 22 June 2016
Accepted 27 June 2016
Available online 29 July 2016

Keywords:
Facial cues
Emotion recognition
Anxiety
Stress
Blink rate
Head motion

ABSTRACT

This study develops a framework for the detection and analysis of stress/anxiety emotional states through video-recorded facial cues. A thorough experimental protocol was established to induce systematic variability in affective states (neutral, relaxed and stressed/anxious) through a variety of external and internal stressors. The analysis was focused mainly on non-voluntary and semi-voluntary facial cues in order to estimate the emotion representation more objectively. Features under investigation included eye-related events, mouth activity, head motion parameters and heart rate estimated through camera-based photoplethysmography. A feature selection procedure was employed to select the most robust features followed by classification schemes discriminating between stress/anxiety and neutral states with reference to a relaxed state in each experimental phase. In addition, a ranking transformation was proposed utilizing self reports in order to investigate the correlation of facial parameters with a participant perceived amount of stress/anxiety. The results indicated that, specific facial cues, derived from eye activity, mouth activity, head movements and camera based heart activity achieve good accuracy and are suitable as discriminative indicators of stress and anxiety.

1. Introduction

Stress and anxiety are common everyday life states of emotional strain that play a crucial role in the person's subjective quality of life. These states consist of several complementary and interacting components (i.e., cognitive, affective, central and peripheral physiological) [1] comprising the organism's response to changing internal and/or external conditions and demands [2]. Historically, the potential negative impact of stress/anxiety on body physiology and health has long been recognized [3]. Although stress and, particularly, anxiety are subjective, multifaceted phenomena, which are difficult to measure comprehensively through objective means, more recent research is beginning to throw light on the factors that determine the degree and type of stress/anxiety-related impact on personal health, including characteristics of the stressor itself, and various biological and psychological vulnerabilities. However, there is evidence that they bear both direct and long term consequences on the person's capacity to adapt to life events and function adequately, as well as to overall wellbeing [4].

Stress is often described as a complex psychological, physiological and behavioural state triggered upon perceiving a significant imbalance between the demands placed upon the person and their perceived capacity to meet those demands [5,6]. From an evolutionary standpoint, it is an adaptive process characterized by increased physiological and psychological arousal. The two key modes of the stress response ("fight" and "flight") were presumably evolved to enhance the survival capacity of the organism [1,7]. Nevertheless, prolonged stress can be associated with psychological and/or somatic disease [8]. Anxiety is the unpleasant mood characterized by thoughts of worry and fear, sometimes in the absence of real threat [9,10]. When anxiety is experienced frequently and at intensity levels that appear disproportional to the actual threat level, it can evolve to a broad range of disorders [11,12].

It should be noted that the terms stress and anxiety are often used interchangeably. Their main difference is that anxiety is usually a feeling not directly and apparently linked to external cues or objective threats. On the other hand, stress is an immediate response to daily demands and is considered to be more adaptive than anxiety. Nevertheless, anxiety and stress typically involve similar physical sensations, such as higher heart rate, sweaty palms

* Corresponding author at: Foundation for Research and Technology—Hellas, Institute of Computer Science, Computational BioMedicine Laboratory, N. Plastira 100, Vassilika Vouton, 70013 Iraklio, Crete, Greece.
E-mail addresses: ggian@ics.forth.gr (G. Giannakakis), mped@ics.forth.gr (M. Pediaditis), mandim@ics.forth.gr (D. Manousos), elenikaz@ics.forth.gr (E. Kazantzaki), chiarugi@ics.forth.gr (F. Chiarugi), akis.simos@gmail.com (P.G. Simos), kmarias@ics.forth.gr (K. Marias), tsiknaki@ics.forth.gr (M. Tsiknakis).

http://dx.doi.org/10.1016/j.bspc.2016.06.020
1746-8094/

Table 1
Categorization of facial features types connected with stress and anxiety.

Head	Eyes	Mouth	Gaze	Pupil
Head movement	Blink rate	Mouth shape	Saccadic eye movements	Pupil size variation
Skin colour	Eyelid response	Lip deformation	Gaze spatial distribution	Pupil ratio variation
Heart rate (facial PPG)	Eye aperture	Lip corner puller/depressor	Gaze direction	
	Eyebrow movements	Lip pressor		

and churning stomach [13], triggered by largely overlapping neuronal circuits [9] when the brain fails to distinguish the difference between a perceived and a real threat. These similarities extend to the facial expressions associated with each state and, accordingly, they were considered as a single state in the present study.

1.1. Physical, mental and cognitive effects of stress/anxiety

Stress and anxiety have impact both on physical and mental health [14]. They are also implicated in the onset and progression of immunological, cardiovascular, circulatory or neurodegenerative diseases [15]. Evidence from both animal experiments and human studies suggests that stress may attenuate the immune response and increase the risk for certain types of cancer [15].

Increased skeletal, smooth and cardiac muscle tension, gastric and bowel disturbances [13] are additional typical signs of stress and anxiety, which are linked to some of their most common symptoms and disorders, namely headache, hypertension, exaggeration of lower back and neck pain, and functional gastrointestinal disorders (such as irritable bowel syndrome) [4]. These signs are frequently accompanied by restlessness, irritability, and fatigue [12]. From a psychological perspective, prolonged stress and anxiety is often identified as a precipitating factor for depression and panic disorders and may further interfere with the person's functional capacity through impaired cognition (e.g., memory, attention [16] and decision making [17]).

Stress can be detected through biosignals that quantify physiological measures. Stress and anxiety affect significantly upper cognitive functions [18] and their effects can be observed through EEG recordings [19,20]. Mainly, stress is identified through arousal related EEG features such as asymmetry [20,21], beta/alpha ratio [22] or increased existence of beta rhythm [23]. Stress regulates active sweat glands [24] increasing the skin conductance during stress conditions. Thus, stress can be detected with the use of Galvanic Skin response (GSR) which has been adopted as a reliable psychophysical measure [25,26]. Breathing patterns are also correlated with emotional status and can be used for stress and anxiety detection [27]. Studies report that respiration rate increases significantly under stressful situations [28]. Additionally, EMG is a biosignal measuring muscle action potentials, where trapezius muscle behaviour is considered to be correlated with stress [29–31]. Finally, speech features are affected in stress conditions [32], the voice fundamental frequency being the most investigated in this research field [33,34].

1.2. Effects of anxiety/stress on the human face

An issue of great interest is the correspondence between information reflected in and conveyed by the human face and the person's concurrent emotional experience. Darwin argued that facial expressions are universal, i.e. most emotions are expressed in the same way on the human face regardless of race or culture [35]. There are several recent studies reporting findings that facial signs and expressions can provide insights into the analysis and classification of stress [34,36,37]. The main manifestations of anxiety on the human face involve the eyes (gaze distribution, blinking rate, pupil size variation), the mouth (mouth activity, lip deformations), the cheeks, as well as the behaviour of the head as a whole (head movements, head velocity). Additional facial signs related to anxiety may include a strained face, facial pallor and eyelid twitching [38]. In reviewing the relevant literature, facial features of potential value as signs of anxiety and stress states were identified (as listed in Table 1) and are briefly described in this section.

There have been reports that head movements can be used as a stress indicator, although their precise association has not yet been established. It has been reported that head movements during stressful conditions are more frequent [39], more rapid [37] and there is greater overall head motion [40]. In [41] head nods and shakes were employed among other features in order to discriminate complex emotional situations. Regarding the eye region, features like the blink rate, eye aperture, eyelid response, gaze distribution and variation in pupil size have been studied. Blinking can be voluntary but also as a reflex to external or internal stimuli and blinking rate typically increases with emotional arousal, including stress and anxiety levels [10,37,42,43]. The blinking rate is affected by various other states, such as lying [44], and disorders such as depression, Parkinson's disease [45] and schizophrenia [46]. It is also affected by environmental conditions such as humidity, temperature and lighting [47]. The percentage of eyelid closure induced by a light stimulus (increase in brightness) was significantly higher in a group of anxious persons when compared to the corresponding response of non-anxious individuals [48].

Similarly, gaze direction, gaze congruence and the size of the gaze-cuing effect are influenced by the level of anxiety or stress [49,50]. Persons with higher trait anxiety demonstrate greater gaze instability under both volitional and stimulus-driven fixations [51]. Moreover, high levels of anxiety were found to disrupt saccadic control in the antisaccade task [52]. Anxious people tend to be more attentive and to make more saccades toward images of fearful and angry faces than others [53,54]. Various studies have documented an association between pupil diameter and emotional, sexual or cognitive arousal. In addition, pupil dilation can be employed as an index of higher anxiety levels [55,56]. Pupillary response to negatively valenced images also tends to be higher among persons reporting higher overall levels of stress [57]. Pupil size may also increase in response to positive, as well as negative arousing sounds as compared to emotionally neutral ones [58].

There is also sufficient evidence in the research literature that mouth-related features, particularly lip movement, are affected by stress/anxiety conditions [37]. Asymmetric lip deformations have been highlighted as a characteristic of high stress levels [59]. In addition, it was found [60] that the frequency of mouth openings was inversely proportional to stress level, as indexed by higher cognitive workload. A technique designed to detect facial blushing has been reported and applied to video recording data [61] concluding that blushing is closely linked to anger-provoking situations and pallor to feelings of fear or embarrassment. The heart rate has also an effect on the human face, as facial skin hue varies in accordance to concurrent changes in blood volume transferred from the heart. There is the general notion in the literature that the heart rate increases during conditions of stress or anxiety [62–66] and that heart rate variability parameters differentiates stress and neutral states [66–68]. Finally, there are approaches that aim to detect stress through facial actions coding. The most widely used coding scheme in this context is the Emotional Facial Action Coding System (EMFACS) [69]. This coding was used to determine facial manifestations of anxiety [10], to classify (along with cortisol, cardiovascular measures) stress responses [70] and to investigate deficits in social anxiety [71].

1.3. *Procedural issues of the available literature*

Most of the relevant literature focuses on the automatic classification of basic emotions [72] based on the processing of facial expressions. Reports on stress detection are fewer, typically employing images [53,54,73], sounds [58], threats [51] or images mimicking emotions [74] as stressful stimuli. Results of some of these studies are often confounded by variability in stimulation conditions, e.g. videos with many colours and light intensity variations, which are known to affect the amplitude of pupillary responses. Perhaps more important with respect to the practical utility of automatic detection techniques is the validity of the experimental setup employed. Well-controlled laboratory conditions are necessary in the early phases of research in order to select pertinent facial features and optimize detection and classification algorithms. These settings are, however, far from real-life conditions in which anxiety and stress are affected by a multitude of situational factors, related to both context and cognitive-emotional processes taking place within the individual and determining the specific characteristics and quality of anxiety and accompanying feelings (fear, nervousness, apprehension, etc.) [75].

1.4. *Purpose of the paper*

The work reported in this manuscript focuses on the automated identification of a set of facial parameters (mainly semi- or/and non-voluntary) to be used as quantitative descriptive indices in terms of their ability to discriminate between neutral and stress/anxiety state. In support of this objective, a thorough experimental procedure has been established in order to induce affective states (neutral, relaxed and stressed/anxious) through a variety of external and internal stressors designed to simulate a wide range of conditions. An emotion recognition analysis workflow used is presented in Fig. 1.

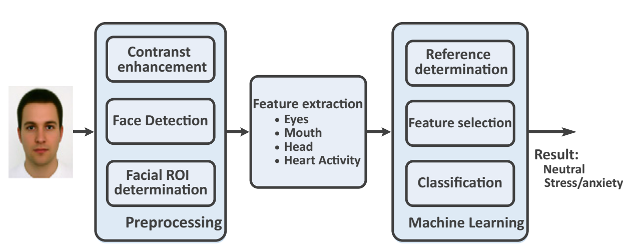


Fig. 1. Overview of the framework for the stress/anxiety detection through facial videos.

The advantages of the present study relate to the fact that (a) a variety of semi voluntary facial features are jointly used for the detection of anxiety and/or stress instead of the traditional facial expression analysis, (b) the experimental protocol consists of various phases/stressful stimuli covering different aspects of stress and anxiety, and (c) the facial features that are involved in stress/anxiety states for different stimuli are identified through machine learning techniques.

2. Methods

2.1. *Experimental setup*

Study participants were seated in front of a computer monitor while a camera was placed at a distance of approximately 50 cm with its field of view (FOV) able to cover properly the participant's face. At the start of the procedure, participants were fully informed about the details of the study protocol and the terms "anxiety" and "stress" were explained in detail. The experimental setup is shown in Fig. 2.

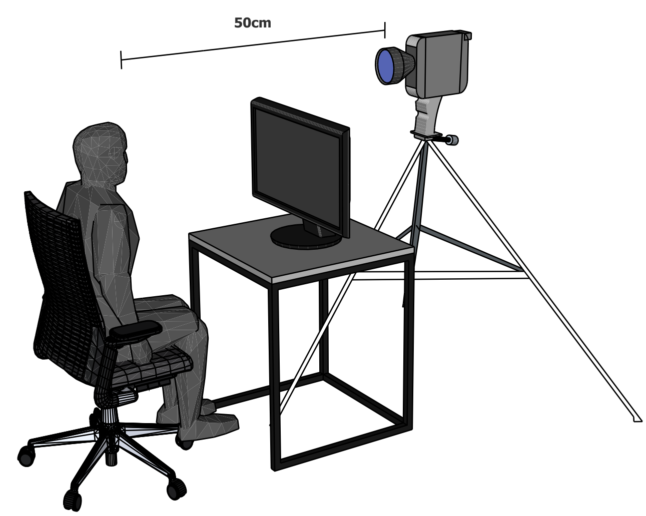


Fig. 2. Experimental setup for video acquisition.

The equipment included a colour video camera placed on a tripod behind and above the pc monitor. Diffused lightning was generated by two 500 W spotlights covered with paper. A checkerboard of 23-mm square size was used to calibrate the pixel size in calculating actual distances between facial features.

2.2. *Experimental procedure*

The ultimate objective of the experimental protocol was to induce affective states (neutral, relaxed and stressed/anxious) through a variety of external and internal stressors designed to simulate a wide range of everyday life conditions. A neutral (reference) condition was presented at the beginning of each phase of the experiment as listed in Table 2.

2.2.1. *Social exposure phase*

A social stressor condition was simulated by asking participants to describe themselves in a foreign language having its origin in the stress faced by an actor when he/she is at the stage. There is reported evidence that people talking a foreign language become more stressed [76]. A third condition in the exposure phase entailed reading aloud an 85 word emotionally neutral literary passage. The latter aimed to induce moderate levels of cognitive load and was similar to the self-describing task, in the sense that both required overt speech.

2.2.2. *Emotion recall phase*

Negative emotions that resemble acute stress and anxiety were elicited by asking participants to recall and relive a stressful and an anxiety eliciting event from their past life as if it was currently happening.

Table 2
Experimental tasks and conditions employed in the study.

	Experimental phase	Duration (min)	Affective State
Social Exposure			
1.1	Neutral (reference)	1	N
1.2	Self-describing	1	A
1.3	Text Reading	½	A
Emotional recall			
2.1	Neutral (reference)	½	N
2.2	Recall anxious event	½	A
2.3	Recall stressful event	½	A
Stressful images/Mental task			
3.1	Neutral/Stressful images	2	A
3.2	Mental task (Stroop)	2	A
Stressful videos			
4.1	Neutral (reference)	½	N
4.2	Calming video	2	R
4.3	Adventure video	2	A
4.4	Psychological pressure video	2	A

Note: Intended affective state (N: neutral, A: anxiety/stress, R: relaxed).

2.2.3. Stressful images/mental task phase

Neutral and unpleasant images from the International Affective Picture System (IAPS) [77] were used as affect generating stimuli. Stimuli included images having stressful content such as human pain, drug abuse, violence against women and men, armed children, etc. Each image was presented for 15 s.

Higher levels of cognitive load were induced through a modified Stroop colour-word reading task, requiring participants to read colour names (red, green, and blue) printed in incongruous ink (e.g., the word RED appearing in blue ink) [78]. In the present task, difficulty was increased by asking participants to first read each word and then name the colour of the word.

2.2.4. Stressful videos phase

Finally, three 2-min video segments were presented in attempting to induce low-intensity positive emotions (calming video), and stress/anxiety (action scene from an adventure film or scene involving heights to participants who indicated at least moderate levels of acrophobia, and a burglary/home invasion while the inhabitant is inside).

2.3. Data description

The participants in our study were 23 adults (16 men, 7 women) aged 45.1 ± 10.6 years. The study was approved by the North-West Tuscany ESTAV (Regional Health Service, Agency for the Technical-Administrative Services of Wide Area) Ethical Committee. All participants gave informed consent. The clinical protocol described in Section 2.2 was applied, which lead to the generation of 12 videos (3 neutral, 8 anxiety/stress and 1 relaxed states) for each participant. Videos were sampled at 50 fps with a resolution of 526 × 696 pixels. The duration of these videos ranged between 0.5, 1 and 2 min as described in Table 2. After the end of tasks 3.1, 4.2, 4.3, and 4.4 participants were asked to rate the level of stress they experienced during the preceding video/image on a scale of 1 (very relaxed) to 5 (very stressful).

2.4. Data preprocessing

Video pre-processing involved histogram equalization for contrast enhancement, face detection, and region of interest (ROI) determination.

2.4.1. Face ROI detection

Face detection is a fundamental step of facial analysis via image/video recordings. It is an essential pre-processing step for facial expression and face recognition, head pose estimation and tracking, face modelling and normalization [79,80]. Early face detection algorithms can be classified into knowledge based, feature invariant, template matching, and appearance based. These methods used different features, such as integrating edge, colour, size, and shape information and different modelling and classification tools, including mixture of Gaussians, shape templates, active shape models (ASM), eigenvectors, hidden Markov models (HMM), artificial neural networks (ANN), and support vector machines (SVM) [79]. It appears that the last category of appearance based methods were the most successful [80]. Modern approaches of face detection are based on the Viola and Jones algorithm [81] which is widely adopted in the vast majority of real-world applications. This algorithm combines simple Haar-like features, integral images, the *AdaBoost* algorithm, and the attention cascade to significantly reduce computational load. Several variants of the original Viola and Jones algorithm are presented in [80].

2.5. Active appearance models

Active appearance models (AAM) [82,83] have been widely applied towards emotion recognition and facial expression classification, as they provide a stable representation of the shape and appearance of the face. In the present work, the shape model was built as a parametric set of facial shapes, described as a set of L landmarks $\in R^2$ forming a vector of coordinates

$$S = [\{x_1, y_1\}, \{x_2, y_2\}, \ldots, \{x_L, y_L\}]^T$$

A mean model shape s_0 was formed by aligning face shapes through Generalized Procrustes Analysis (GPA) [84]. Alignment of two shapes S_1 and S_2 was performed by finding scale, rotation and translation parameters that minimize the Procrustes distance metric

$$d_{S_1,S_2} = \sqrt{\sum_{k=1}^{L} (x_{1k} - x_{2k})^2 + (y_{1k} - y_{2k})^2}$$

2

For every new alignment step the mean shape was re-computed and the shapes were aligned again to the new mean. This procedure was iterated until the mean shape did not change significantly. Then, Principal Components Analysis (PCA) was employed projecting data onto an orthonormal subspace in order to reduce data dimension. According to this procedure, a shape s can be expressed as

$$s = s_0 + \sum p_i s_i$$

where s_0 is the mean model shape and p_i the model shape parameters.

The *appearance model* was built as a parametric set of facial textures. A facial texture A of m pixels was represented by a vector of intensities g_i

$$A(x) = [g_1 g_2 \cdots g_m]^T \forall x \in s_0$$

Likewise the *shape model* was based on the mean appearance A_0 and the appearance images A_i which were computed by applying PCA to a set of shape normalized training images. Each training image was shape normalized by warping the training mesh onto the base mesh s_0 [83]. Following PCA, textures A_i can be expressed as

$$A(x) = A_0(x) + \sum \lambda_i A_i(x)$$

where $A_0(x)$ is the mean model appearance and λ_i the model appearance parameters.

When the model is created, it was fitted to new images or video sequences I by identifying the shape parameters p_i and appearance parameters λ_i that produced the most accurate fit. This nonlinear optimization problem aims to minimize the objective function

$$\sum_x [I(W(x;p)) - A_0(W(x;\Delta p))]^2 \forall x \in s_0$$

where W is the warping function. This minimization was performed using the inverse compositional image alignment algorithm [83].

2.6. Optical flow

In this work, dense optical flow was used for the estimation of mouth motion. The velocity vector for each pixel was calculated within the mouth ROI and the maximum velocity, or magnitude, in each frame was extracted. Optical flow was applied only on the Q channel of the YIQ transformed image, since the lips appear brighter in this channel [85]. For subsequent feature extraction, the maximum magnitude over the entire video was taken as the source signal.

Optical flow is a velocity field that transforms each of a series of images to the next image. In video recordings optical flow describes the relative motion between the camera and the scene. It works under two assumptions: The motion must be smooth in relation to the frame rate, and the brightness of moving pixels must be constant. For two frames, pixel $I(x,y,t)$ in the first frame moves by distance dx and dy in the subsequent frame that is taken after time dt. Accordingly, this pixel can be described as:

$$I(x, y, t) = I(x + dx, y + dy, t + dt)$$

In the current study, the velocity vector for each pixel was calculated by using dense optical flow as described by Farnebäck [86]. The procedure employs a two-frame motion estimation algorithm that initially applies quadratic polynomial expansion to both frames on a neighbourhood level. Subsequently, the displacement fields are estimated from the polynomial expansion coefficients. Consider the quadratic polynomial expansions of e.g. two frames in the form

$$f_1(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{A_1}\boldsymbol{x} + \boldsymbol{b_1}^T\boldsymbol{x} + c_1 \qquad f_2(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{A_2}\boldsymbol{x} + \boldsymbol{b_2}^T\boldsymbol{x} + c_2$$

that are related by a global displacement $\mathbf{d}$:

$$f_2(\mathbf{x}) = f_1(\mathbf{x} - \mathbf{d})$$

Solving for displacement $\mathbf{d}$ yields

$$\mathbf{d} = -\frac{1}{2}\mathbf{A}_1^{-1}(\mathbf{b}_2 - \mathbf{b}_1)$$

This means that $\mathbf{d}$ can be gained from the coefficients $\mathbf{A}_1$, $\mathbf{b}_1$ and $\mathbf{b}_2$, if $\mathbf{A}_1$ is non-singular. For a practical implementation the expansion coefficients ($\mathbf{A}_1$, $\mathbf{b}_1$, c_1 and $\mathbf{A}_2$, $\mathbf{b}_2$, c_2) are first calculated for both images and the approximation for $\mathbf{A}$ is made:

$$\mathbf{A} = \frac{\mathbf{A}_1 + \mathbf{A}_2}{2}$$

Based on the equations above and with

$$\Delta\mathbf{b} = -\frac{1}{2}(\mathbf{b}_2 - \mathbf{b}_1)$$

A primary constraint can be obtained:

$$\mathbf{A}d = \Delta\mathbf{b}$$

or

$$\mathbf{A}(\mathbf{x})\mathbf{d}(\mathbf{x}) = \Delta\mathbf{b}(\mathbf{x})$$

if the global displacement $\mathbf{d}$ is replaced by a spatially varying displacement field $\mathbf{d}(\mathbf{x})$. This equation is solved over a neighborhood of each pixel.

This method was selected because it produced the lowest error in a comparison to the most popular methods as reported in [86], as well as the fact that it produces a dense motion field that is able to capture the spatially varying motion patterns of the lips.

2.7. Camera based photoplethysmography

Camera based photoplethysmography (PPG) can be applied to facial video recordings in order to estimate the heart rate (HR) as described in [87] by measuring variations of facial skin hue associated with concurrent changes in blood volume in subcutaneous blood vessels during the cardiac cycle. The method determines a region of interest (ROI), which is reduced by 25% from each vertical edge of the whole face ROI. The intensities of the three colour channels (R, G, B) were spatially averaged in each frame to provide a time series of the mean ROI intensities in each channel. These time series were detrended and normalized. Then, Independent Component Analysis (ICA) decomposed the signal space into three independent signal components and the component with the lowest spectral flatness was selected as the most appropriate. The frequency corresponding to the spectral distribution peak, within the range of predefined frequencies, determined the heart rate. It should be noted, that quick body/head movements may lead to artifacts which deteriorate method's accurate estimation of heart rate. This limitation could be handled with improved camera characteristics or the usage of the ICA algorithm which enhances independent source signals and reduces motion artifacts [87,88].

3. Results

3.1. Facial feature/cue extraction

3.1.1. Eye related features

Eye related features were extracted using Active Appearance Models (AAM) [82] as described in Section 2.5. The delineation of the perimeter of each eyeball was segmented using 6 discrete landmarks as shown in the detail of left-hand panel. The 2-dimensional coordinates (x_i, y_i) of the 6 landmarks were used to compute the eye aperture

$$A = \frac{1}{2}\sum_{i=1}^{N} |x_i y_{i+1} - y_i x_{i+1}|$$

where $N = 6$ and the convention that $x_{N+1} = x_1 \quad y_{N+1} = y_1$

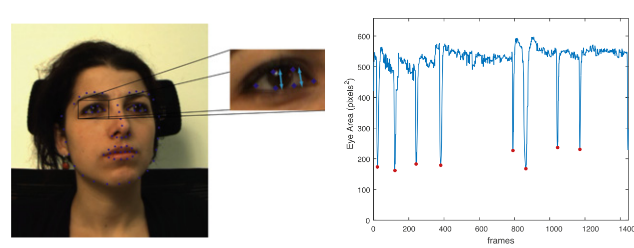
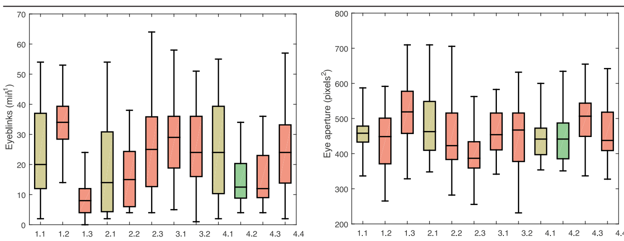


Fig. 3. (Left-hand panel) Eye aperture average distance calculation; (Right-hand panel) Time series of eye aperture and detection of blink occurrence (red circles) by the peak detection algorithm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

An example of an aperture time series is shown in Fig. 3. Time series were detrended and filtered in order to remove artefactual spikes introduced when the model occasionally loses its tracking and to smooth highly varying segments. The average aperture aggregated over the entire time mean value of this time series determines the mean eye aperture that served as a feature in subsequent analyses.

Table 3
Average (±s.d.) values of eye related features for each emotion elicitation task and the respective baseline and corresponding boxplots of eye related features.

Experimental phase	Eye Blink (blinks/min)	Eye aperture (pixels^2)
Social Exposure		
1.1 Neutral (reference)	23.9 ± 14.2	453.5 ± 56.6
1.2 Self-description speech	**35.0 ± 9.1***	441.8 ± 90.1
1.3 Text reading	**8.8 ± 5.5****	**513.6 ± 84.1***
Emotional recall		
2.1 Neutral (reference)	18.8 ± 13.9	493.0 ± 105.2
2.2 Recall anxious event	17.0 ± 11.1	457.8 ± 110.2
2.3 Recall stressful event	**25.9 ± 15.0****	**393.2 ± 77.0***
Stressful images/Mental task		
3.1 Neutral/Stressful images	**26.9 ± 13.2***	455.4 ± 53.1
3.2 Mental task (Stroop)	26.8 ± 14.0	424.1 ± 101.1
Stressful videos		
4.1 Neutral	24.6 ± 16.7	438.8 ± 62.4
4.2 Relaxing video (reference)	15.4 ± 8.6	453.6 ± 78.6
4.3 Adventure/heights video	15.3 ± 8.2	**493.7 ± 74.3***
4.4 Home invasion video	**23.7 ± 12.8****	447.5 ± 64.8



Colour coding of affect state: brown = neutral, red = stress/anxiety, green = relaxed
*p < 0.05.
**p < 0.01 for each task vs. the respective reference. In the stressful videos phase the reference was the relaxing video.

A second eye-related feature, the number of blinks per minute, was extracted from the same timeseries. Each eye blink can be seen as a sharp aperture decrease as shown in Fig. 3. Detection of possible troughs was performed through signal's derivative [89]. Then, peak detection criteria [90] were applied based on the parameters of peak SNR, first derivative amplitude (peak slope), peak distance, peak duration to filter small peaks and select only timeseries sharp decreases as shown in the right panel of Fig. 3. Especially, the minimum peak duration parameter was set to 100 msec (corresponding to 5 consecutive data points) given that an eye blink typically lasts between 100 and 400 msec.

Data normality was checked in all experimental phases and tasks. Extreme value analysis was applied to the data excluding one participant who, upon visual inspection of the video recordings showed stereotypical movements (tics) involving the upper facial musculature. Repeated measures of one-way ANOVA [91] was used to compare eye related features (blinks, aperture) across tasks separately for each experimental phase. The results are presented in Table 3.

Results indicate increased blink rate under stressful situations compared to the corresponding neutral or relaxed state. Specifically, statistically significant increased blink rate was found in the self-description ($p < 0.05$), the stressful event recall task ($p < 0.01$), and psychological pressure video viewing ($p < 0.01$). Stressful images viewing task was also associated with significantly increased blink rate ($p < 0.05$). On the other hand, eye blink frequency was reduced significantly ($p < 0.01$) during text reading, and increased during the Stroop colour-word task (both involved word reading). Thus it appears that increased cognitive effort per se was associated with increased blinking rate, whereas reading a continuous text for comprehension may require persistent attention and increased saccadic activity. Mean eye aperture was increased in stressful situations (text reading and adventure/heights video viewing).

3.1.2. Mouth related features

From the maximum magnitude signal in the mouth ROI (cf. Section 2.6) the following features were extracted: the variance of time intervals between adjacent peaks (VTI), the energy ratio (ENR) of the autocorrelation sequence which was calculated as the ratio of the energy contained by the last 75% of the samples of the autocorrelation sequence to the energy contained by the first 25%, the median, variance, skewness, kurtosis, and Shannon entropy in 10 equally sized bins. Features were not evaluated for tasks 1.2, 1.3, 3.1, 3.2 where mouth movements associated with speaking aloud would confound emotion-related activity. The results of the mouth related features are presented in Table 4.

Inspection of Table 4 reveals that systematic variation in several metrics was restricted to the emotion recall conditions (associated with increased VTI and skewness, and reduced entropy as compared to the neutral condition). VTI was used for estimating the periodicity of the movements based on the observation that rhythmic movements would produce variances close to zero. A higher VTI implies reduced rhythmicity of lip movements associated with increased levels of stress/anxiety. Increased skewness indicates higher asymmetry of the probability distribution, whereas entropy is an index of reduced disorder or spread within the energy of the signal. During watching stressful videos increased median and variance of the maximum magnitude were found indicating faster mouth movements.

Table 4
Average (±s.d.) of mouth related features for each emotion elicitation task and the respective neutral reference.

Experimental phase	VTI	ENR	Median	Variance	Skewness	Kurtosis	Shannon Entropy
Social Exposure							
1.1 Neutral (reference)	0.75 ± 0.14	2.97 ± 0.93	−0.008 ± 0.004	0.004 ± 0.002	1.52 ± 0.55	5.77 ± 4.34	3.09 ± 0.12
Emotional recall							
2.1 Neutral (reference)	0.72 ± 0.14	3.14 ± 0.84	−0.011 ± 0.006	0.005 ± 0.004	1.07 ± 0.43	2.24 ± 2.08	3.13 ± 0.09
2.2 Recall anxious event	**0.78 ± 0.14***	2.84 ± 0.58	−0.010 ± 0.007	0.005 ± 0.004	**1.39 ± 0.56***	3.56 ± 2.68	**3.02 ± 0.21***
2.3 Recall stressful event	0.76 ± 0.08	3.01 ± 0.57	−0.008 ± 0.003	0.004 ± 0.003	**1.29 ± 0.25***	2.99 ± 1.52	**3.05 ± 0.14***
Stressful video							
4.1 Neutral (reference)	0.78 ± 0.28	2.94 ± 0.79	−0.009 ± 0.004	0.007 ± 0.005	1.24 ± 0.86	2.84 ± 3.92	3.06 ± 0.20
4.2 Relaxing video	0.71 ± 0.19	2.49 ± 0.82	−0.013 ± 0.009	0.021 ± 0.021	0.89 ± 0.25	1.53 ± 1.08	3.10 ± 0.17
4.3 Adventure/heights video	0.74 ± 0.18	2.60 ± 0.67	−0.010 ± 0.006	**0.010 ± 0.009***	1.06 ± 0.37	2.50 ± 1.75	3.09 ± 0.15
4.4 Home invasion video	0.67 ± 0.12	2.71 ± 0.65	**−0.014 ± 0.006****	**0.010 ± 0.006***	1.07 ± 0.35	2.77 ± 2.52	3.15 ± 0.08

Note: Mouth features were not evaluated during phase 3 and tasks 1.2 and 1.3.
*p < 0.05.
**p < 0.01 for each task vs. the respective neutral reference.

3.1.3. Head movements

Active Appearance Models as described in Section 2.5 were used to determine five specific facial landmarks p_k, $k = 1, 2, \ldots, 5$ on the nose as a highly stable portion of the face. The distribution of these specific landmarks, shown in the left-hand panel of Fig. 4, were selected in order to represent all types of translation and rotation head movements in the coronal transverse planes. The 100th frame in the time series was used as reference and in each frame the Euclidean distance of these points from the corresponding points of the reference frame was calculated.

The total head movement was defined as the mean value of these distances across all frames N during each task

$$movm = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{5}\sum_{k=1}^{5}\|p_k - p_k^{ref}\|$$

where $p_k, k = 1, 2, \ldots, 5$ the landmark points and $\|.\|$ is the Euclidean norm

The speed of head motion was calculated by

$$speed = \frac{1}{N}\sum_{t=2}^{N}\frac{1}{5}\sum_{k=1}^{5}\|p_k(t) - p_k(t-1)\|$$

Table 5 demonstrates statistically significant increased head motion parameters (amplitude and velocity) during stressful states in each experimental phase and task. Increased amplitude of head motion and velocity compared to neutral state ($p < 0.01$) appears to be associated with speaking aloud (tasks 1.2, 1.3 and 3.2), and the most prominent increase was noted during the self-description speech. To a smaller degree, increased movement amplitude and velocity was found during silent recall of anxious and stressful, personal life event. During the Stroop colour-word test the movements were related in some way with the rhythmicity of words presentations and the rhythmicity of their answers. Increased head movement, without concurrent increase in velocity, was observed while watching video segments.

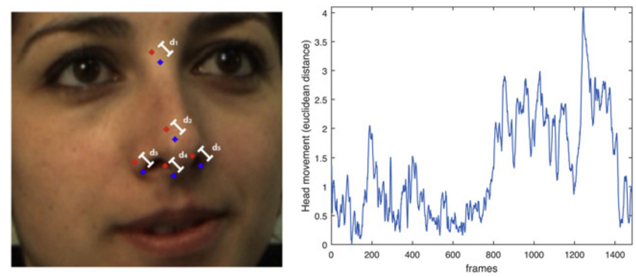
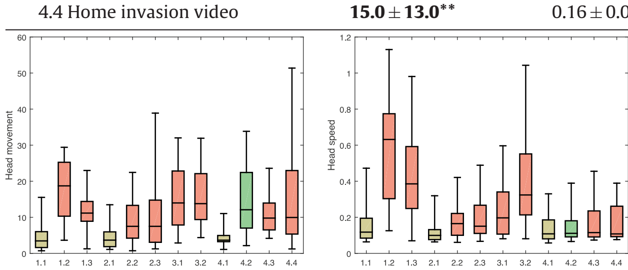


Fig. 4. (Left-hand panel) the positions of facial landmarks used for head movement analysis in the current frame are displayed by blue spots. The positions of corresponding landmarks during the respective reference period are shown in red. (Right-hand panel) the time series of estimated heal movement (task minus reference of 100th frame). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Average (±s.d.) values of head movement features for each emotion elicitation task and the respective neutral reference and corresponding boxplots.

Experimental phase	Head movement (pixels)	Head velocity (pixels/frame)
Social Exposure		
1.1 Neutral (reference)	4.8 ± 4.2	0.16 ± 0.11
1.2 Self-description speech	**17.8 ± 8.2****	**0.58 ± 0.28****
1.3 Text reading	**11.7 ± 4.4****	**0.44 ± 0.23****
Emotional recall		
2.1 Neutral (reference)	4.7 ± 3.4	0.12 ± 0.06
2.2 Recall anxious event	8.8 ± 5.8	**0.17 ± 0.09***
2.3 Recall stressful event	**9.6 ± 8.4***	**0.20 ± 0.12****
Stressful images/Mental task		
3.1 Neutral/stressful images	**15.8 ± 8.4****	**0.24 ± 0.14****
3.2 Mental task (Stroop)	**15.5 ± 8.1****	**0.40 ± 0.26****
Stressful videos		
4.1 Neutral (reference)	4.5 ± 2.5	0.13 ± 0.06
4.2 Relaxing video	13.9 ± 8.7	0.15 ± 0.09
4.3 Adventure/heights video	**10.6 ± 4.9****	0.16 ± 0.10
4.4 Home invasion video	**15.0 ± 13.0****	0.16 ± 0.09



Colour coding of affect state: brown = neutral, red = stress/anxiety, green = relaxed.
Colour coding of affect state: brown = neutral, red = stress/anxiety, green = relaxed.
*p < 0.05.
**p < 0.01 for each task vs. the respective neutral reference.

3.1.4. Heart rate from facial video

Following the methodology presented in Section 2.7 and splitting data into partially overlapping sliding windows of 30 s, heart rate estimates were obtained and are presented in Table 6.

As expected, heart rate increased when someone experiences anxiety or stress. The most apparent increases were found during the self-description, text reading and Stroop colour-word tasks as compared to the corresponding neutral reference period ($p < 0.01$), raising the question that heart rate may be affected by speech per se rather than, or in addition to, task-related elevated anxiety levels.

3.2. Relationship between self-reports and facial signs

The participants' self-reports of perceived stress after sessions 4.2, 4.3, and 4.4 were used to assess potential associations between stress and facial signs at the individual level. Average (±s.d.) ratings were 1.55 ± 0.89 (relaxing video), 3.50 ± 1.11 (adventure/heights video), and 2.68 ± 1.06 (home invasion video). One of the main problems in establishing associations relates to the different subjective scales implicitly applied by each person to mentally represent the entire range of experienced stress. Nonlinear models, such as pair-wise preference analysis, may be more appropriate to map individual physiological responses parameters to subjective stress ratings [92]. In the current study preference pairs were constructed consisting of a stress rating and a single feature value estimated for each of the four tasks listed above. Correlation values were calculated using the test statistic [93]

$$c(z) = \sum_{i=1}^{N}\{z_i/N\}$$

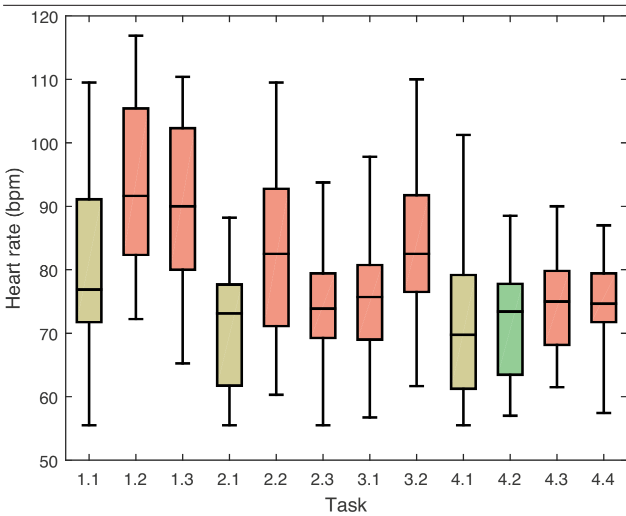
where z_i is 1 if there is an agreement between the rating difference (e.g. most stressful) and the corresponding difference in facial sign, and −1 if there is disagreement. N is the number of non-zero differences between self-reports and facial signs for each task. This statistic follows a binomial distribution

$$f(k, N, p) = \binom{N}{k} p^k (1-p)^{N-k}$$

where k is the number of agreements and p the a priori probability. The resulting binomially-distributed pairwise correlations for the stressful video phase are presented in Table 7, demonstrating significant, moderate-to-strong positive correlations between

Table 6
Average (±s.d.) heart rate values and corresponding boxplots for each emotion elicitation task and the respective neutral reference derived from facial video photoplethysmography.

Experimental phase	Heart Rate (bpm)
Social Exposure	
1.1 Neutral (reference)	81.0 ± 14.7
1.2 Self-description speech	**93.7** ± **12.8****
1.3 Text Reading	**89.7** ± **13.3***
Emotional recall	
2.1 Neutral (reference)	70.9 ± 8.9
2.2 Recall anxious event	**82.1** ± **12.9****
2.3 Recall stressful event	74.7 ± 9.5
Stressful images/Mental task	
3.1 Neutral/stressful images	74.7 ± 9.4
3.2 Mental task (Stroop)	**83.0** ± **10.8****
Stressful videos	
4.1 Neutral (reference)	72.0 ± 11.7
4.2 Relaxing video	72.2 ± 9.4
4.3 Adventure/heights video	74.8 ± 8.3
4.4 Home invasion video	74.7 ± 8.4



Colour coding of affect state: brown = neutral, red = stress/anxiety, green = relaxed.
*p < 0.05.
**p < 0.01 for each task vs. the respective neutral reference.

perceived stress levels and three facial features: higher blink rate, head movement velocity, and variance of time intervals between adjacent peaks (VTI) of mouth movements.

3.3. Neutral and stress/anxiety states reference determination

Initially, each measure obtained from individual participants during all tasks were referenced to the task 4.2, which corresponds to the relaxed state of each subject as a baseline of facial cues, and all subsequent feature analyses were conducted on the difference values

$$\hat{x} = x - x_{ref}$$

where x_{ref} is the baseline condition (task 4.2) for each feature. This transformation generates a common reference to each feature across subjects providing data normalization.

3.4. Feature selection and ranking

In principle, classification efficiency is directly related to the number of relevant, complementary features employed in the procedure (as compared to univariate approaches). However, when the number of features entered into the model exceeds a particular level, determined by the characteristics of the data set used, classification accuracy may actually deteriorate. In the present work,

Table 7
Pairwise rank correlations (c(z)) and corresponding *p*-values between stress self-ratings and recorded facial signs.

Facial features		Relaxing video paired with	
		Adventure/heights video	Home invasion video
Eye blinks	c(z)	0.100	**0.714****
	p	0.160	**0.006**
Eye aperture	c(z)	0.368	0
	p	0.052	0.209
Head movement	c(z)	−0.300	−0.067
	p	0.074	0.196
Head velocity	c(z)	**0.412***	0.333
	p	**0.033**	0.092
Mouth VTI	c(z)	**0.556***	0.200
	p	**0.012**	0.205
Mouth ENR	c(z)	0.176	0.200
	p	0.148	0.205
Mouth Median	c(z)	0.000	−0.273
	p	0.196	0.161
Mouth Skewness	c(z)	0.143	0.000
	p	0.183	0.246
Mouth Kurtosis	c(z)	0.200	0.000
	p	0.153	0.246
Shannon Entropy	c(z)	0.000	−0.200
	p	0.196	0.205
Heart Rate	c(z)	0.143	0.333
	p	0.140	0.092

Abbreviations; ENR: energy ratio; VTI: variance of time intervals between adjacent peaks.
*p < 0.05.
**p < 0.01.

feature selection was performed using sequential forward selection (SFS) and sequential backward selection (SBS) methods [94] to reduce the number of features to be used for automatic discrimination of facial recordings according to experimental task. The objective function evaluating feature subsets was the area under the ROC curve (AUC) [95] using the classifier under investigation and reflecting emotional states differences within each experimental phase. The initial feature set consisted of 12 features, and SFS proved more effective than SBS to minimize the objective function. Mouth-related features were not evaluated for data from phases 1 and 3 because they involve at least one speaking aloud task. This procedure identified the sets of three to six most pertinent features across experimental phases presented in Table 8. An example of the area under curve as a function of the selected number of features through SFS is presented in Fig. 5.

Table 8
Features selected from the sequential forward selection procedure.

Phase	Features Selected
1. Social Exposure	[head velocity, eye aperture, head movement, heart rate]
2. Emotional recall	[heart rate, mouth skewness, eye aperture, mouth Shannon entropy, mouth median, mouth variance, eye blinks]
3. Stressful images/ Mental task	[head velocity, head movement, eye blinks, heart rate]
4. Stressful Videos	[head movement, heart rate, eye aperture, eye blinks, mouth ENR, mouth skewness]

Abbreviation; ENR: energy ratio.

It should be stated that the effectiveness of a feature selection method depends on the evaluation metric, the classifier used for this metric, the distance metric and the selection method [96]. The most robust features are selected in most cases but different classification schemes may vary considering the selected least significant features as well as the overall order of selection.

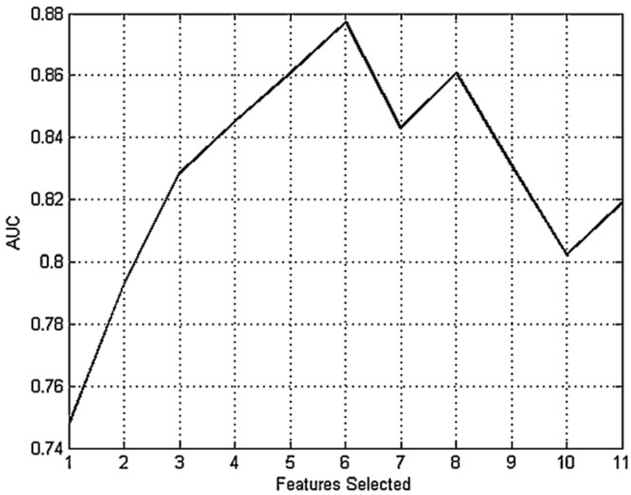


Fig. 5. Example of the area under curve as a function of the number of features selected from SFS for the stressful videos phase and for K-nn classifier.

3.5. Classification performance results

The system's classification procedure uses the feature set (referenced data) derived from the feature selection procedure for the data in each experimental phase (see Table 8). The classification methods that were used and tested are: K-nearest neighbours (K-nn), Generalized Likelihood Ratio, Support Vector Machines (SVM), Naïve Bayes classifier and AdaBoost classifier. Each classification method was tested in terms of classification accuracy, i.e. its capacity to discriminate between feature sets obtained during the neutral expression of each phase and feature sets obtained during the remaining, stress-anxiety inducing tasks. A 10-fold cross-validation was performed with the referred classifiers by randomly splitting the feature set into 10 mutually exclusive subsets (folds) of approximately equal size. In each case, the classifier was trained using 9 folds (90% of data) and tested on the remaining data (10% of data). The classification procedure was repeated 20 times and the average classification accuracy results are presented in Table 9

It can be observed that the classification results demonstrate good discrimination ability for the experimental phases performed. In each experiment phase, classification accuracy that range between 80% and 90% taking into account the most effective classifier. The best classification accuracy was presented in the social exposure phase using Adaboost classifier achieving accuracy of 91.68%. The phase with Stroop-colour test and stressful images (phase 3) appeared to be more consistent across classifiers tested.

4. Discussion

This study investigates the use of task elicited facial signs as indices of anxiety/stress in relation to neutral and relaxed states. Although there is much literature discussing recognition and analysis of the six basic emotions, i.e. anger, disgust, fear, happiness, sadness and surprise, considerably less research has focussed on stress and anxiety detection from facial videos. This can be justified partially by the fact that these states are considered as complex emotions that are linked to basic emotions (e.g. fear) making them more difficult to be interpreted in the human face. The sensitivity of specific measures of facial activity to situational stress/anxiety states was assessed in a step-wise manner, which entailed univariate analyses contrasting emotion elicitation tasks to their respective neutral reference conditions, followed by multivariate classification schemes. The ultimate goal of these analyses was to identify sets of facial features, representing involuntary or semi-voluntary facial activity that can reliably discriminate emotional from neutral states through unsupervised procedures.

The facial cues finally used in this study involved eye related features (blinks, eye aperture), mouth activity features (VTI, ENR, median, variance, skewness, kurtosis and Shannon entropy), head movement amplitude, head velocity and heart rate estimation derived from variations in facial skin colour. All these features provide a contactless approach of stress detection not interfering with human body in relation to other related studies employing semi-invasive measurements like ECG, EEG, galvanic skin response (GSR) and skin temperature. In addition, this approach incorporates information from a broad set of features derived from different face regions providing a holistic view of the problem under investigation. The sensitivity of these features was tested across different experimental phases employing a wide variety of stress/anxiety eliciting conditions. It was deduced that certain features are sensitive to stress/anxiety states across a wide range of eliciting conditions, whereas others display a more restricted, task-specific performance.

It can be concluded that eye blink rate increases during specific stress and anxiety conditions. In addition, head movement amplitude and velocity is also related to stress, in the form of small rapid movements. Regarding mouth activity, the median and temporal variability (VTI) of the maximum magnitude increased while watching a stressful video whereas Shannon entropy appears to decrease. Finally, increased heart rate was observed during stress and anxiety states, where the most apparent differences were noted in the social exposure phase (self-describing speech, text reading) and mental task (Stroop colour-word test). The feature selection procedure led to identifying specific features as being more robust in discriminating between stress/anxiety and a neutral emotional mode.

The classification accuracy reached good levels as compared to the results of the few related studies available [25,34,37,97,98], taking into account the complexity and situational specificity of these emotional states. Studies [37] and [97] used, among other features, facial features achieving 75%–88% and 90.5% classification accuracy respectively. Biosignals like Blood Volume Pulse (BVP), Galvanic Skin Response (GSR), Pupil Diameter (PD) and Skin Temperature (ST) are used also in literature for discriminating stress situation as in [25,98,99] achieving 82.8%, 90.1% and 90.1% classification accuracies respectively. To our knowledge, the best accuracy achieved is 97.4% in discriminating among 3 states of stress [30], but the techniques used (ECG, EMG, RR, GSR) are semi-invasive and therefore less convenient in real life applications. An intuitive review of stress

Table 9
Average stress classification accuracy results for each task.

	K-nn (%)	Generalized likelihood ratio (%)	Naive Bayes (%)	AdaBoost (%)	SVM (%)
1. Social Exposure	85.54	86.96	86.09	91.68	82.97
2. Emotional recall	88.70	72.39	73.26	81.03	65.82
3. Stressful images/Mental task	88.32	81.96	83.42	87.28	83.15
4. Stressful Videos	88.32	75.49	71.63	83.80	76.09

detection techniques with corresponding classification accuracies is presented in [34].

On the other hand, some limitations of the experimental procedure and analysis should be noted. The duration of facial recording may affect results, especially if it is very short and the consequent analysis of specific facial cues can be dubious. Based on our experiences, it is suggested that a video recording of at least 1 min duration could yield more reliable estimates and should be used for future recordings. Moreover, elicitation of anxiety and stress states took place in a controlled, laboratory environment somehow limiting the generalizability of results to real-life situations where the intensity and valence of stressors may vary more widely. Physical conditions during facial recordings are found to affect the sensitivity of certain features, such as pupil size variation or blink rate where global monitor luminance changes induced by alternating static or dynamic visual stimuli (video scenes) are affecting some physiological features independently of cognitive and emotional state variables. Individual differences in both static and dynamic aspects of facial features represent another source of unwanted variability in the association between preselected facial features and individual emotional states. In the present study, a person-specific, reference condition was used to scale each facial feature to a common baseline prior to classification. An active reference was selected (viewing a relaxing video), which was intended to elicit facial activity related to attention and visual/optokinetic tracking, without triggering intense emotions.

In conclusion, this study investigated facial cues implicated in stress and anxiety. A sound analytical framework has been developed for the analysis, recognition and classification for efficient stress/anxiety detection through facial video recordings, achieving good classification accuracy in relation to the neutral state.

Acknowledgments

This work was partially supported by the FP7 specific targeted research project SEMEOTICONS: SEMEiotic Oriented Technology for Individual's CardiOmetabolic risk self-assessmeNt and Self-monitoring under the Grant Agreement number 611516.

References

[1] N. Schneiderman, G. Ironson, S.D. Siegel, Stress and health: psychological, behavioral, and biological determinants, Ann. Rev. Clin. Psychol. 1 (2005) 607.
[2] H. Selye, The stress syndrome, AJN Am. J. Nursing 65 (1965) 97–99.
[3] L. Vitetta, B. Anton, F. Cortizo, A. Sali, Mind-Body medicine: stress and its impact on overall health and longevity, Ann. N. Y. Acad. Sci. 1057 (2005) 492–505.
[4] R., Donatelle (2013). My Health: An Outcomes Approach, Pearson Education.
[5] R.S. Lazarus, S. Folkman, Stress, Appraisal, and Coping, Pringer, New York, 1986.
[6] H. Selye, The Stress of Life, McGraw-Hill, New York, NY, US, 1956.
[7] T. Esch, G.L. Fricchione, G.B. Stefano, The therapeutic use of the relaxation response in stress-related diseases, Med. Sci. Monitor 9 (2003) RA23–RA34.
[8] S. Palmer, W. Dryden, Counselling for Stress Problems, SAGE Publications Ltd, 1995.
[9] L.M. Shin, I. Liberzon, The neurocircuitry of fear, stress, and anxiety disorders, Neuropsychopharmacology 35 (2010) 169–191.
[10] J.A. Harrigan, D.M.O. O'Connell, How do you look when feeling anxious? Facial displays of anxiety, Personality Individ. Differences 21 (1996) 205–212.
[11] Understanding anxiety disorders and effective treatment, American Psychological Association, http://www.apapracticecentral.org/outreach/anxiety-disorders.pdf, June 2010.
[12] M. Gelder, J. Juan, A. Nancy, New Oxford Textbook of Psychiatry, vol. 1 & 2, Oxford University Press, Oxford, 2004.
[13] L.A. Helgoe, L.R. Wilhelm, M.J. Kommor, The Anxiety Answer Book, Sourcebooks Inc., 2005.
[14] T. Esch, Health in stress: change in the stress concept and its significance for prevention, health and life style, Gesundheitswesen (Bundesverband der Arzte des Offentlichen Gesundheitsdienstes (Germany)) 64 (2002) 73–81.
[15] E.M.V. Reiche, S.O.V. Nunes, H.K. Morimoto, Stress, depression, the immune system, and cancer, Lancet Oncol. 5 (2004) 617–625.
[16] L. Schwabe, O.T. Wolf, M.S. Oitzl, Memory formation under stress: quantity and quality, Neurosci. Biobehav. Rev. 34 (2010) 584–591.
[17] E. Dias-Ferreira, J.C. Sousa, I. Melo, P. Morgado, A.R. Mesquita, J.J. Cerqueira, R.M. Costa, N. Sousa, Chronic stress causes frontostriatal reorganization and affects decision-making, Science 325 (2009) 621–625.
[18] B.S. McEwen, R.M. Sapolsky, Stress and cognitive function, Curr. Opin. Neurobiol. 5 (1995) 205–216.
[19] J. Alonso, S. Romero, M. Ballester, R. Antonijoan, M. Mañanas, Stress assessment based on EEG univariate features and functional connectivity measures, Physiol. Meas. 36 (2015) 1351.
[20] G. Giannakakis, D. Grigoriadis, M. Tsiknakis, Detection of stress/anxiety state from EEG features during video watching, 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2015) 6034–6037.
[21] R.S. Lewis, N.Y. Weekes, T.H. Wang, The effect of a naturalistic stressor on frontal EEG asymmetry, stress, and health, Biol. Psychol. 75 (2007) 239–247.
[22] K. Nidal, A.S. Malik, EEG/ERP Analysis: Methods and Applications, CRC Press, 2014.
[23] E., Hoffmann (2005). Brain Training Against Stress. Mental Fitness & Forsknings center ApS Symbion Science Park.
[24] W. Boucsein, Electrodermal Activity, Springer Science & Business Media, 2012.
[25] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Troster, U. Ehlert, Discriminating stress from cognitive load using a wearable EDA device, IEEE Trans. Inf. Technol. Biomed. 14 (2010) 410–417.
[26] A. de Santos Sierra, C.S. Ávila, J. Guerra Casanova, G.B. del Pozo, A stress-detection system based on physiological signals and fuzzy logic, IEEE Trans. Ind. Electron. 58 (2011) 4857–4865.
[27] Y. Masaoka, I. Homma, The effect of anticipatory anxiety on breathing and metabolism in humans, Respir. Physiol. 128 (2001) 171–177.
[28] K.B. Nilsen, T. Sand, L.J. Stovner, R.B. Leistad, R.H. Westgaard, Autonomic and muscular responses and recovery to one-hour laboratory mental stress in healthy subjects, BMC Musculoskelet Disord. 8 (2007) 1.
[29] J. Wijsman, B. Grundlehner, J. Penders, H. Hermens, Trapezius muscle EMG as predictor of mental stress, ACM Trans. Embedded Comput. Syst. (TECS) 12 (2013) 1–20.
[30] J.A. Healey, R.W. Picard, Detecting stress during real-world driving tasks using physiological sensors, IEEE Trans. Intell. Transp. Syst. 6 (2005) 156–166.
[31] P. Karthikeyan, M. Murugappan, S. Yaacob, EMG signal based human stress level classification using wavelet packet transform, in: Trends in Intelligent Robotics Automation, and Manufacturing, Springer, 2012, pp. 236–243.
[32] I.R. Murray, C. Baber, A. South, Towards a definition and working model of stress and its effects on speech, Speech Commun. 20 (1996) 3–12.
[33] P. Wittels, B. Johannes, R. Enne, K. Kirsch, H.-C. Gunga, Voice monitoring to measure emotional load during short-term stress, Eur. J. Appl. Physiol. 87 (2002) 278–282.
[34] N. Sharma, T. Gedeon, Objective measures, sensors and computational techniques for stress recognition and classification: a survey, Comput. Methods Prog. Biomed. 108 (2012) 1287–1301.
[35] C. Darwin, The Expression of the Emotions in Man and Animals, Oxford University Press, USA, 1998.
[36] N. Sharma, T. Gedeon, Modeling observer stress for typical real environments, Expert Syst. Appl. 41 (2014) 2231–2238.
[37] D.F. Dinges, R.L. Rider, J. Dorrian, E.L. McGlinchey, N.L. Rogers, Z. Cizman, S.K. Goldenstein, C. Vogler, S. Venkataraman, D.N. Metaxas, Optical computer recognition of facial expressions associated with stress induced by performance demands, Aviat. Space Environ. Med. 76 (2005) 172–182.
[38] M. Hamilton, The assessment of anxiety-States by rating, Br. J. Med. Psychol. 32 (1959) 50–55.
[39] W. Liao, W. Zhang, Z. Zhu, Q. Ji, A real-time human stress monitoring system using dynamic bayesian network, Computer Society Conference in Computer Vision and Pattern Recognition IEEE (2005), 70–70.
[40] U. Hadar, T. Steiner, E. Grant, F.C. Rose, Head movement correlates of juncture and stress at sentence level, Lang. Speech 26 (1983) 117–129.
[41] A. Adams, M. Mahmoud, T. Baltrusaitis, P. Robinson, Decoupling facial expressions and head motions in complex emotions, Affective Computing and Intelligent Interaction (ACII), 2015 International Conference On, IEEE (2015) 274–280.
[42] M. Argyle, M. Cook, Gaze and Mutual Gaze, Cambridge University Press, Cambridge, England, 1976.
[43] C.S. Harris, R.I. Thackray, R.W. Shoenber, Blink rate as a function of induced muscular tension and manifest anxiety, Percept. Motor Skill 22 (1966), 155-&.
[44] S. Leal, A. Vrij, Blinking during after lying, J. Nonverbal Behav. 32 (2008) 187–194.
[45] J.H. Mackintosh, R. Kumar, T. Kitamura, Blink rate in psychiatric-illness, Br. J. Psychiatry 143 (1983) 55–57.
[46] R.C.K. Chan, E.Y.H. Chen, Blink rate does matter: a study of blink rate, sustained attention, and neurological signs in schizophrenia, J. Nerv. Mental Dis. 192 (2004) 781–783.
[47] G.W. Ousler, K.W. Hagberg, M. Schindelar, D. Welch, M.B. Abelson, The ocular protection index, Cornea 27 (2008) 509–513.
[48] J.A. Taylor, The relationship of anxiety to the conditioned eyelid response, J. Exp. Psychol. 41 (1951) 81–92.
[49] E. Fox, A. Mathews, A.J. Calder, J. Yiend, Anxiety and sensitivity to gaze direction in emotionally expressive faces, Emotion 7 (2007) 478–486.
[50] J.P. Staab, The influence of anxiety on ocular motor control and gaze, Curr. Opin. Neurol. 27 (2014) 118–124.

[51] G. Laretzaki, S. Plainis, I. Vrettos, A. Chrisoulakis, I. Pallikaris, P. Bitsios, Threat and trait anxiety affect stability of gaze fixation, Biol. Psychol. 86 (2011) 330–336.
[52] N. Derakshan, T.L. Ansari, M. Hansard, L. Shoker, M.W. Eysenck, Anxiety, inhibition, efficiency, and effectiveness an investigation using the antisaccade task, Exp. Psychol. 56 (2009) 48–55.
[53] A. Holmes, A. Richards, S. Green, Anxiety and sensitivity to eye gaze in emotional faces, Brain Cognition 60 (2006) 282–294.
[54] K. Mogg, M. Garner, B.P. Bradley, Anxiety and orienting of gaze to angry and fearful faces, Biol. Psychol. 76 (2007) 163–169.
[55] M. Honma, Hyper-volume of eye-contact perception and social anxiety traits, Conscious. Cognition 22 (2013) 167–173.
[56] H.M. Simpson, F.M. Molloy, Effects of audience anxiety on pupil size, Psychophysiology 8 (1971), 491-&.
[57] M.O. Kimble, K. Fleming, C. Bandy, J. Kim, A. Zambetti, Eye tracking and visual attention to threating stimuli in veterans of the Iraq war, J. Anxiety Disorders 24 (2010) 293–299.
[58] T. Partala, V. Surakka, Pupil size variation as an indication of affective processing, Int. J. Hum.-Comput. Stud. 59 (2003) 185–198.
[59] D. Metaxas, S. Venkataraman, C. Vogler, Image-based stress recognition using a model-based dynamic face tracking system, in: Computational Science-ICCS 2004, Springer, 2004, pp. 813–821.
[60] W. Liao, W. Zhang, Z. Zhu, Q. Ji, A Decision Theoretic Model for Stress Recognition and User Assistance, AAAI, 2005, pp. 529–534.
[61] P. Kalra, N. Magnenat-Thalmann, Modeling of vascular expressions in facial animation, Computer Animation'94., Proceedings of IEEE (1994) 50–58 (201).
[62] J.M. Gorman, R.P. Sloan, Heart rate variability in depressive and anxiety disorders, Am. Heart J. 140 (2000) S77–S83.
[63] T.G. Vrijkotte, L.J. Van Doornen, E.J. De Geus, Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability, Hypertension 35 (2000) 880–886.
[64] C.D. Katsis, N.S. Katertsidis, D.I. Fotiadis, An integrated system based on physiological signals for the assessment of affective states in patients with anxiety disorders, Biomed. Signaling Process. 6 (2011) 261–268.
[65] J.A. Healey, R.W. Picard, Detecting stress during real-world driving tasks using physiological sensors, IEEE Trans. Intell. Transp. Syst. 6 (2005) 156–166.
[66] A. Steptoe, G. Willemsen, O. Natalie, L. Flower, V. Mohamed-Ali, Acute mental stress elicits delayed increases in circulating inflammatory cytokine levels, Clin. Sci. 101 (2001) 185–192.
[67] D. McDuff, S. Gontarek, R. Picard, Remote measurement of cognitive stress via heart rate variability, 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2014) 2957–2960.
[68] C. Schubert, M. Lambertz, R. Nelesen, W. Bardwell, J.-B. Choi, J. Dimsdale, Effects of stress on heart rate complexity—a comparison between short-term and chronic stress, Biol. Psychol. 80 (2009) 325–332.
[69] W.V. Friesen, P. Ekman, Emfacs-7: Emotional Facial Action Coding System, University of California at San Francisco, 1983, Unpublished manuscript.
[70] J.S. Lerner, R.E. Dahl, A.R. Hariri, S.E. Taylor, Facial expressions of emotion reveal neuroendocrine and cardiovascular stress responses, Biol. Psychiatry 61 (2007) 253–260.
[71] S. Melfsen, J. Osterlow, I. Florin, Deliberate emotional expressions of socially anxious children and their mothers, J. Anxiety Disorders 14 (2000) 249–261.
[72] P. Ekman, Are there basic emotions, Psychol. Rev. 99 (1992) 550–553.
[73] W. Liu-sheng, X. Yue, Attention bias during processing of facial expressions in trait anxiety: an eye-tracking study, Electronics and Optoelectronics (ICEOE), 2011 International Conference On, IEEE (2011), pp. V1-347-V341-350.
[74] A.M. Perkins, S.L. Inchley-Mort, A.D. Pickering, P.J. Corr, A.P. Burgess, A facial expression for anxiety, J. Pers. Soc. Psychol. 102 (2012) 910.
[75] J.R. Vanin, Overview of anxiety and the anxiety disorders, in: Anxiety Disorders, Springer, 2008, pp. 1–18.
[76] K. Hirokawa, A. Yagi, Y. Miyata, An examination of the effects of linguistic abilities on communication stress measured by blinking and heart rate, during a telephone situation, Soc. Behav. Personality 28 (2000) 343–353.
[77] P.J. Lang, M.M. Bradley, B.N. Cuthbert, International Affective Picture System (iaps): Technical Manual and Affective Ratings, NIMH Center for the Study of Emotion and Attention, 1997, pp. 39–58.
[78] J.R. Stroop, Studies of interference in serial verbal reactions, J. Exp. Psychol. 18 (1935) 643–662.
[79] M.H. Yang, D.J. Kriegman, N. Ahuja, Detecting faces in images: a survey, IEEE Trans. Pattern Anal. Mach. Intell. 24 (2002) 34–58.
[80] C. Zhang, Z. Zhang, A survey of recent advances in face detection, In: Tech. rep., Microsoft Research, 2010.
[81] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, Computer Vision and Pattern Recognition (CVPR) 2001 IEEE (2001) 511–518.
[82] T.F. Cootes, G.J. Edwards, C.J. Taylor, Active appearance models, IEEE Trans. Pattern Anal. Mach. Intell. 23 (2001) 681–685.
[83] I. Matthews, S. Baker, Active appearance models revisited, Int. J. Comput. Vision 60 (2004) 135–164.
[84] J.C. Gower, Generalized procrustes analysis, Psychometrika 40 (1975) 33–51.
[85] N. Thejaswi, S. Sengupta, Lip localization and viseme recognition from video sequences, in: Fourteenth National Conference on Communications, India, 2008.
[86] G. Farnebäck, Two-frame motion estimation based on polynomial expansion, in: Image Analysis, Springer, 2003, pp. 363–370.
[87] M. Poh, D. McDuff, R. Picard, Non-contact, automated cardiac pulse measurements using video imaging and blind source separation, Opt. Express 18 (2010) 10762–10774.
[88] J. Kranjec, S. Beguš, G. Geršak, J. Drnovšek, Non-contact heart rate and heart rate variability measurements: a review, Biomed. Signal Process. 13 (2014) 102–112.
[89] A. Felinger, Data Analysis and Signal Processing in Chromatography, Elsevier, 1998.
[90] C. Yang, Z. He, W. Yu, Comparison of public peak detection algorithms for MALDI mass spectrometry data analysis, BMC Bioinf. 10 (2009).
[91] E.R. Girden, ANOVA. Repeated Measures, Sage, 1992.
[92] C. Holmgard, G.N. Yannakakis, K.-I. Karstoft, H.S. Andersen, Stress detection for ptsd via the startlemart game, Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference On, IEEE (2013) 523–528.
[93] G.N. Yannakakis, J. Hallam, Ranking vs. preference: a comparative study of self-reporting, in: Affective Computing and Intelligent Interaction, Springer, 2011, pp. 437–446.
[94] P. Pudil, J. Novovičová, J. Kittler, Floating search methods in feature selection, Pattern Recogn. Lett. 15 (1994) 1119–1125.
[95] T. Fawcett, An introduction to ROC analysis, Pattern Recogn. Lett. 27 (2006) 861–874.
[96] M. Dash, H. Liu, Feature selection for classification, Intell. Data Anal. 1 (1997) 131–156.
[97] H. Gao, A. Yuce, J.-P. Thiran, Detecting emotional stress from facial expressions for driving safety, Image Processing (ICIP) 2014 IEEE International Conference On, IEEE (2014) 5961–5965.
[98] J. Zhai, A. Barreto, Stress recognition using non-invasive technology, FLAIRS Conference (2006) 395–401.
[99] A. Barreto, J. Zhai, M. Adjouadi, Non-intrusive physiological monitoring for automated stress detection in human-computer interaction, in: Human–Computer Interaction, Springer, 2007, pp. 29–38.