Research Article

Objectifying Facial Expressivity Assessment of Parkinson's Patients: Preliminary Study

Peng Wu,[1] Isabel Gonzalez,[1] Georgios Patsis,[1] Dongmei Jiang,[2] Hichem Sahli,[1] Eric Kerckhofs,[3] and Marie Vandekerckhove[4]

[1] *Department of Electronics and Informatics, Vrije Universiteit Brussel, 1050 Brussels, Belgium*
[2] *Shaanxi Provincial Key Lab on Speech and Image Information Processing, Northwestern Polytechnical University, Xi'an, China*
[3] *Department of Physical Therapy, Vrije Universiteit Brussel, 1050 Brussels, Belgium*
[4] *Department of Experimental and Applied Psychology, Vrije Universiteit Brussel, 1050 Brussels, Belgium*

Correspondence should be addressed to Peng Wu; pwu@etro.vub.ac.be

Patients with Parkinson's disease (PD) can exhibit a reduction of spontaneous facial expression, designated as "facial masking," a symptom in which facial muscles become rigid. To improve clinical assessment of facial expressivity of PD, this work attempts to quantify the dynamic facial expressivity (facial activity) of PD by automatically recognizing facial action units (AUs) and estimating their intensity. Spontaneous facial expressivity was assessed by comparing 7 PD patients with 8 control participants. To voluntarily produce spontaneous facial expressions that resemble those typically triggered by emotions, six emotions (amusement, sadness, anger, disgust, surprise, and fear) were elicited using movie clips. During the movie clips, physiological signals (facial electromyography (EMG) and electrocardiogram (ECG)) and frontal face video of the participants were recorded. The participants were asked to report on their emotional states throughout the experiment. We first examined the effectiveness of the emotion manipulation by evaluating the participant's self-reports. Disgust-induced emotions were significantly higher than the other emotions. Thus we focused on the analysis of the recorded data during watching disgust movie clips. The proposed facial expressivity assessment approach captured differences in facial expressivity between PD patients and controls. Also differences between PD patients with different progression of Parkinson's disease have been observed.

1. Introduction

One of the manifestations of Parkinson's disease (PD) is the gradual loss of facial mobility and "mask-like" appearance. Katsikitis and Pilowsky (1988) [1] stated that PD patients were rated as significantly less expressive than an aphasic and control group, on a task designed to assess spontaneous facial expression. In addition, the spontaneous smiles of PD patients are often perceived to be "unfelt," because of the lack of accompanying cheek raises [2]. Jacobs et al. [3] confirmed that PD patients show reduced intensity of emotional facial expression compared to the controls. In order to assess facial expressivity, most research relies on subjective coding of the implied researchers, as in aforementioned studies. Tickle-Degnen and Lyons [4] found that decreased facial expressivity

correlated with self-reports of PD patients as well as the Unified Parkinson's Disease Rating Scale (UPDRS) [5]. PD patients, who rated their ability to facially express emotions as severely affected, did demonstrate less facial expressivity.

In this paper, we investigate automatic measurements of facial expressivity from video recorded PD patients and control populations. To the best of our knowledge, in actual research, few attempts have been made for designing a computer-based quantitative analysis of facial expressivity of PD patient. To analyze whether Parkinson's disease affected voluntary expression of facial emotions, Bowers et al. [6] videotaped PD patients and healthy control participants while they made voluntary facial expression (happy, sad, fear, anger, disgust, and surprise). In their approach, the amount of facial movements change and timing have been quantified by

estimating an *entropy score* plotted over time. The entropy is a measure of pixel intensity change that occurred over the face as it moved during expression. Also they computed the time it took an expression to reach its peak entropy value from the onset of each trial. Using both measures, entropy and time, the authors demonstrated that less movements occurred over the face of PD patients when they were asked to mimic a target expression relative to control.

Despite its good results, the above described amount of facial movements does not directly relate to measuring facial muscles activity. Indeed, facial expressions are generated by contractions of facial muscles, which lead to subtle changes in the area of the eyelids, eye brows, nose, lips, and skin texture, often revealed by wrinkles and bulges. To measure these subtle changes, Ekman and Friesen [7] developed the Facial Action Coding System (FACS). FACS is a human-observer-based system designed to detect subtle changes in facial features and describes facial expressions by action units (AUs). AUs are anatomically related to contraction of specific facial muscles. They can occur either singly or in combinations. Simons et al. in [2] used the FACS to analyze facial expressivity of PD patients versus control. In their study, odor was used as a means of inducing facial expression. Certified FACS coders annotated the facial expressions. A total facial activity (TFA) measure, estimated as the total number of displayed AUs, was used to assess facial expressions. The authors demonstrated that the TFA measure revealed that compared to controls, PD patients have reduced level of facial activity in reaction to unpleasant odors.

Estimating only the number of displayed AUs does not completely capture the dimensions of facial masking that are present in PD and defined in the *Interpersonal Communication Rating Protocol*-Parkinson's Disease Version (ICRP-IEB) [9]. The ICRP-IEB facial expressivity is based on the FACS. It defines expressivity in terms of (i) *frequency*, that is, how often a behavior or movement occurs, (ii) *duration*, that is, how long a behavior or movement lasts, and (iii) *intensity or degree* being the strength, force, or level/amount of emotion or movement.

In this study, we propose a system that (i) automatically detects faces in a video stream, (ii) codes each frame with respect to 11 action units, and (iii) estimates a facial expressivity as function of frequency, duration, and intensity of AUs. Although there is already a substantial literature on automatic expression and action unit recognition, it still remains an active area of study due to the challenging nature of the problem [10]. Moreover, in contrast to AU detection, there is scarce work in the literature on AU intensity estimation. The proposed facial expressivity quantity makes use of the output of a previously developed automatic AU recognition system [11] based on support vector machines (SVM) and AdaBoost. To determine the AU intensity, the resulting AU distance measure, from the AU SVM classifier, is mapped to the estimates of probability using the Platt scaling algorithm. Platt scaling [12] refers to a technique whereby a score-to-probability calibration curve is calculated using the training set. Frame-by-frame intensity measurements are then used

TABLE 1: The selected movie clips listed with their sources.

Emotion	Excerpt's source	
	#1	#2
Amusement	Benny and Joone	The god father
Sadness	An officer and a gentleman	Up
Surprise	Capricorn one	Sea of love
Anger	Witness	Gandhi
Disgust	Pink flamingos	Trainspotting
Fear	Silence of the lambs	The shining
Neutral	Colour bar patterns	Hannah and her sisters

FIGURE 1: Emotion elicitation protocol (SR indicates self-report).

to estimate facial expression dynamics which were previously intractable by human coding.

2. Methods

2.1. Pilot Study. Our study aims to quantify facial expressivity dynamics of PD patients. Thus, gathering usable qualitative emotional data is the essential step prior to the analysis of facial behaviors. To voluntarily produce spontaneous facial expressions that resemble those typically triggered by emotions, in our study, six emotions (amusement, sadness, anger, disgust, surprise, and fear) were elicited using movie clips. During the movie clips, physiological signals and frontal face video of the participants were recorded. Fifteen participants, 7 PD patients and 8 healthy control persons, participated in the pilot study. After each movie, the participants were asked to rate the intensity of dominant emotions they experienced while watching the movie clips. The evaluation of the participant's self-reports showed that the disgust-induced emotion was significantly higher than the other emotions. Thus we focused on the analysis of the recorded data during watching disgust movie clips.

2.2. Emotion Induction Protocol. Based on the studies of Gross and Levenson [13], Hagemann et al. [14], Lisetti and Nasoz [15], Hewig et al. [16], Westerink et al. [17], and Schaefer et al. [18], we composed a set of movie clips [19]. For each emotion (amusement, sadness, anger, disgust, surprise, fear, and neutral) two excerpts were used, as listed in Table 1. In the sequel, we will refer to move clips as *excerpt#i*; $i = 1$, 2. For example, the two surprise movie clips will be denoted as *surprise*#1 and *surprise*#2.

The used protocol is depicted in Figure 1. The participants were told to watch 2 training movie clips and the 14 movie clips of Table 1. The movie clips were shown randomly, 2 successive ones with different emotions. After each video clip the participants filled in a questionnaire (i.e., self-report) with their emotion-feeling (amusement, sadness, anger, disgust, surprise, fear, and neutral) and rated the strength of their responses using a 7-point scale.

FIGURE 2: Experimental setup.

FIGURE 3: Locations of the EMG electrodes.

The data recording took place in a sound-isolated lab under standardized lighting condition. The movie clips were watched by the participants while sitting on a sofa that was facing a screen with dimensions 1.25 by 1.25 m (see Figure 2).

2.3. Participants. This pilot study considered seven PD patients (3 men and 4 women, between the ages of 47 and 76 years, durations of PD ranged from 1.5 to 13 years) and eight control participants (5 men and 3 women, between the age of 27 and 57 years). The control participants were recruited from the VUB ETRO department without specific characteristics. The PD patients were selected with the help of the Flemish Parkinson League. The study was approved by the committees on human research of the respective institutions and was completed in accordance with the Helsinki Declaration.

Based on the medical dossier as well as the feedback from the PD patients, we defined three PD categories. One patient had a very light form of PD and therefore was classified as the least severe case (denoted as LP). Another patient had the most severe form of PD of the whole group (MP). The rest of PD patients were situated between those two extremes; we denoted them by intermediate PD (IP).

2.4. Data Acquisition. During the movie clips, physiological signals and frontal face video of the participants were recorded. As physiological signal we recorded electromyography (EMG) and electrocardiogram (ECG). All the data channels (i.e., EMG, ECG, and videotape) were synchronized.

The ECG measures the activity of heart contractions. The physical action of the heart is induced by a local periodic electrical stimulation, and as a result a change in potential of 1.0-2.0 μV is measured during a cardiac cycle between two surface electrodes [20]. Heart rate (HR) and heart rate variability (HRV) are the cardiovascular response features most often reported as indicators of emotions [21]. In this study, the ECG was measured at 512 Hz using a Shimmer.

The EMG measures the frequency of muscle tension and contraction. Within each period of interest, the root mean square (RMS) and absolute mean value (AMV) are commonly used as features [22]. In this study, two facial muscles were measured, at 2000 Hz using the EMG Biomonitor
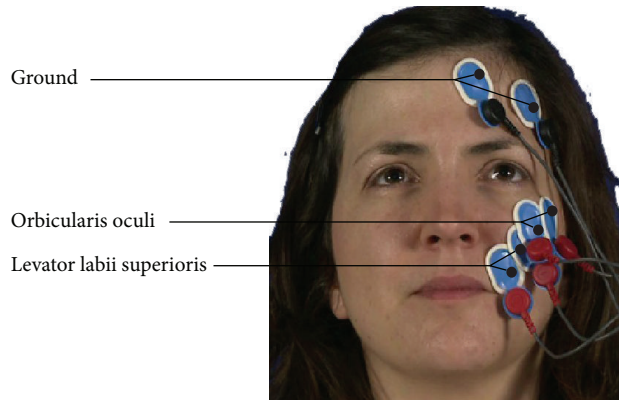
ME6000, namely, the levator labii superioris (LLS) and the orbicularis oculi (OO), as illustrated in Figure 3.

The frontal face video of the participants was recorded for the purpose of quantifying facial expressivity. Following the (ICRP-IEB) [9], we selected 11 action units (AU1, AU2, AU4, AU6, AU7, AU9, AU12, AU20, AU23, AU25, and AU27) among the 41 ones defined by FACS. Table 2 lists the used AUs, along with their associated facial muscles, as well as the ones defined by the measured EMG.

Our purpose, being the development of a nonobtrusive approach for facial expressivity assessment by automatically recognizing facial action units (AUs) and estimating their intensity, the ECG and EMG measurements have been included in our experimental protocol to quantitatively asses the emotional manipulation for inducing facial expressions and confirm facial muscle activity when expressing disgust expression. As our objective did not aim at physiologically investigating the effect of Parkinson's on facial EMG, only the LLS and OO were measured by the EMG as a complementary information to the considered AUs (see Table 2).

2.5. Data Preparation. It is an accepted practice to code a proportion of the observations or "thin-slices" of recorded data as a representation of the analyzed behavior [23]. In our experiments, 30 s data (ECG, EMG, and video record) extracts were selected, corresponding to the last 30 s segments of the shown video clips. Intuitively, one may think that longer segments of expressive behavior in persons with PD would give more information and thus increase the accuracy of analysis. However, Ambady and Rosenthal [24] found that the judgment performance did not significantly improve when using 5 minutes slices versus 30 s slices. This was also confirmed in [25, 26]. Note that baseline data were also selected from the stimuli neutral#1.

2.6. Physiological Data Processing. Physiological signals need to be preprocessed prior to feature extraction in order to remove noise and enhance signal-to-noise ratio (SNR) [27, 28]. In this work, we make use of a newly developed signal denoising approach based on the empirical mode decomposition (EMD) [8]. Different from state-of-the-art

TABLE 2: FACS AUs and related muscles.

AU	FACS name	Facial muscle	Videotape	EMG
AU1	Inner brow raiser	Frontalis (pars medialis)	X	—
AU2	Outer brow raiser	Frontalis (pars lateralis)	X	—
AU4	Brow lowerer	Depressor glabellae, depressor supercilii, and CS	X	—
AU6	Cheek raiser	OO (pars orbitalis)	X	X
AU7	Lid tightener	OO (pars palpebralis)	X	X
AU9	Nose wrinkler	LLSAN	X	—
AU10	Upper lip raiser	LLS, caput infraorbitalis	—	X
AU12	Lip corner puller	Zygomaticus Major	X	—
AU20	Lip stretcher	Risorius	X	—
AU23	Lip tightener	Orbicularis Oris	X	—
AU25	Lips part	Depressor labii inferioris	X	—
AU27	Mouth stretch	Pterygoids, digastric	X	—
AU45	Blink	Contraction OO	—	X
AU46	Wink	OO	—	X
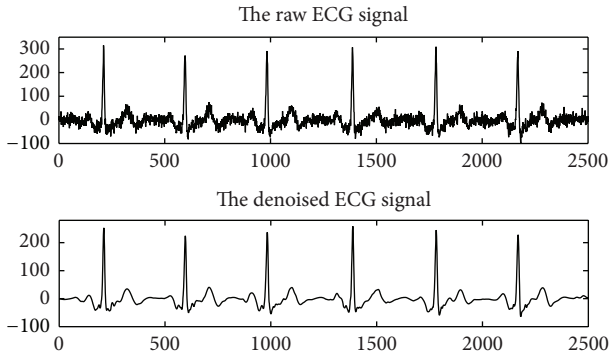
FIGURE 4: Example of physiological signal (ECG) denoising using the method proposed in [8].

FIGURE 5: Change points detection.

methods [27–33], our approach estimates the noise level of each intrinsic mode functions (IMFs), rather than estimating the noise level of all IMFs using Donoho's strategy [34], prior to the reconstruction of the signal using the thresholded IMFs. Please refer to [8] for more details. Figure 4 illustrates the denoising results.

2.6.1. Electrocardiogram (ECG). Heart rate (HR) and heart rate variability (HRV) are the cardiovascular response features most often reported as indicators of emotion [21]. HR is computed using the time difference between two consecutive detected R peaks of the QRS complexes (i.e., RR interval) and is expressed in beats per minute. HRV is the variation of beat-to-beat HR. Thus, the first step in extracting HR and HRV starts from the exact detection of R peaks in the QRS complex. Thus the detection of QRS complex, in particular R peak detection, is the basis for ECG processing and analysis. Many approaches for R peaks detection have been proposed [35]. However, most of them are off-line and targeting the noiseless signal, which do not meet the requirements of many real-time applications. To
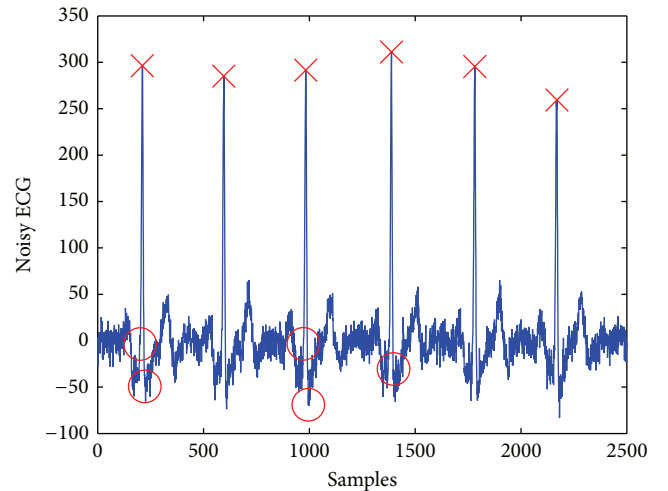
overcome this problem, in [8], we proposed an approach based on a change point detection (CPD) algorithm for event detection in time series [36] that minimizes the error in fitting a predefined function using maximum likelihood. In our current implementation polynomial fitting functions of degree 1 have been selected empirically. An example of results is illustrated in Figure 5 which detects all R picks (cross) and some irrelevant change points (circle), which can be filtered out using a predetermined threshold. Once the R peaks were detected, HR and HRV features are estimated as the difference between the HR (HRV) estimated from the considered 30 s window (of the stimuli) and the one estimated from the neutral 30 s.

2.6.2. Electromyogram (EMG). The absolute mean value (AMV) is the most commonly used feature for identifying

the strength of muscular contraction [22, 37], defined over the length, N, of the signal $x(t)$ as follows:

$$\text{AMV} = \frac{1}{N}\sum_{t=1}^{N}|x(t)|. \tag{1}$$

The AMV value during the considered 30 s window was expressed as a percentage of the mean amplitude during the 30 s neutral baseline. This percentage score was computed to standardize the widely different EMG amplitudes of individuals and thus to enable comparison between individuals and groups.

2.7. Design and Statistical Analysis. Preliminary analysis were performed to check the validity of the acquired data.

(1) Manipulation Check. We calculated the descriptive statistics (the mean and standard deviation) of self-reported emotional ratings to check if the participants' emotional experience was successfully manipulated. Then, Wilcoxon rank sum tests were performed to compare the self-report between groups (PD and control) and between the two clips of the same emotion.

(2) Analysis of Physiological Parameters. Physiological variables were tested for univariate significant differences, via repeated-measures analysis of variance (ANOVA), between groups (PD versus control) and between the disgust#1, disgust#2, and neutral#2 stimulants. Group and stimuli are the between-subjects and within-subject factors, respectively. Results were considered statistically significant at $P < .05$.

2.8. Facial Action Units Recognition. In this study we make use of an automatic facial action units recognition system developed at our department [11]. This system allows context-independent recognition of the following action units: AU1, AU2, AU4, AU6, AU7, AU9, AU12, AU20, AU23, AU25, and AU27. The overall recognition scheme is depicted in Figure 6.

The head and facial features were tracked using the constrained shape tracking approach of [38]. This approach allows an automatic detection of the head and the tracking of a shape model composed of 83 landmarks. After the facial components have been tracked in each frame, both *geometry-based* features and *appearance-based* features are combined and fed to the AdaBoost (adaptive boosting) algorithm for feature selection. Finally, for each action unit, we used a binary support vector machine (SVM) for context-independent classification. Our system was trained and tested on the Kanade et al. DFAT-504 dataset [39]. The database consists of 486 sequences of facial displays that are produced by 98 university students from 18 to 30 years old, of which 65% is female. All sequences are annotated by certified FACS coders, start with a neutral face, and end with the apex of the expression. SVM [11] has been proven to be powerful and robust tools for AU classification.

A test sample (a new image) $\mathbf{z}$ can be classified according to the following decision function:

$$D(\mathbf{z}) = \text{sign}(h(\mathbf{z})). \tag{2}$$

Table 3: Facial expressivity items defined in ICRP-IEB [9] and used AUs.

Item	Gestalt degree	Related AUs
(1) Active expressivity in face	Intensity	11 AUs
(2) Eyebrows raising	Intensity + frequency	AU1 and AU2
(3) Eyebrows pulling together	Intensity + frequency	AU4
(4) Blinking	Frequency	AU45
(5) Cheek raising	Intensity + frequency	AU6
(6) Lip corner puller	Intensity + frequency	AU12

The output $h(\mathbf{z})$ of a SVM is a distance measure between a test pattern and the separating hyperplane defined by the support vectors. The test sample is classified to the positive class (the AU to which it was trained) if $D(\mathbf{z}) = +1$ and is classified to the negative class if $D(\mathbf{z}) = -1$.

Unfortunately, we cannot use directly the output of a SVM as a probability. There is no clear relationship with the posterior class probability $P(y = +1 \mid \mathbf{z})$ that the pattern $\mathbf{z}$ belongs to the class $y = +1$. Platt [12] proposed an estimate for this probability by fitting the SVM output $h(\mathbf{z})$ with a sigmoid function as follows:

$$P(y = +1 \mid \mathbf{z}) = \frac{1}{1 + \exp(Ah(\mathbf{z}) + B)}. \tag{3}$$

The parameters A and B are found using maximum likelihood estimation from the training set. The above equation has a mapping range in $[0; 1]$.

2.9. Facial Expressivity. In this work we follow the recommendation of the *Interpersonal Communication Rating Protocol*-Parkinson's Disease Version (ICRP-IEB) [9], where the expressivity, based on the FACS, has been defined in terms of (i) *frequency*, that is, how often a behavior or movement occurs, (ii) *duration*, that is, how long a behavior or movement lasts, and (iii) *intensity or degree* being the strength, force, or level/amount of emotion or movement. In ICRP-IEB facial expressivity is coded according to the gestalt degree of intensity/duration/frequency of 7 types of facial expressive behavior (items), depicted in recorded videos of PDs, using a 5-point Likert type scale: 1 = low (with low to no movement or change or infrequent), 2 = fairly low, 3 = medium, 4 = fairly high, and 5 = high (very frequent, very active). Table 3, lists 6 of the ICRP-IEB facial expressivity items and the corresponding AUs used in this study for estimating them. The last item is related to active mouth closure during speech, which was not used in our model, as the participants were not asked to speak during the experiments.

As our current AU recognition system considers only 11 AUs (AU1, AU2, AU4, AU6, AU7, AU9, AU12, AU20, AU23, AU25, and AU27), we did not consider blinking (AU45) in our facial expression formulation.
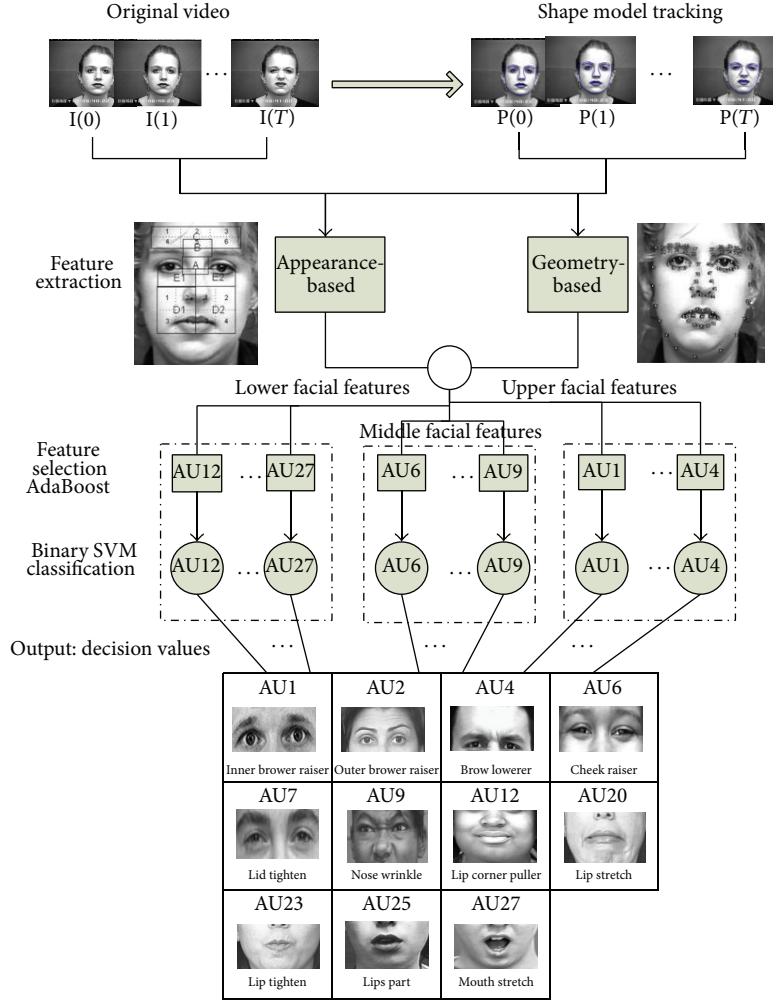
FIGURE 6: The AU recognition system.

Following the criteria of the ICRP-IEB, we defined the facial expressivity of a participant as follows:

$$\text{EFE} = \text{AMC}_e + \text{AI}_e - \left(\text{AMC}_n + \text{AI}_n\right), \qquad (4)$$

where AMC_e and AMC_n are the amount of movement changes during the disgust and neutral facial expression, respectively. AI_e and AI_n are the intensity of the displayed AUs during the disgust and neutral facial expression, respectively.

It has to be recalled that for all these estimations only the considered 30 s windows are used. The quantities AMC and AI refer to frequency and intensity, respectively. Their detailed formulation is given in the following sections. In order to assess the effectiveness of the proposed formulation we also compared it to the following definition of facial expression, where the baseline of neutral emotion was not considered. Consider the following:

$$\text{FE} = \text{AMC}_e + \text{AI}_e. \qquad (5)$$

(1) Intensity of Displayed AUs. It has been shown in [40] that the output margin of the learned SVM classifiers contained

information about expression intensity. Later Savran et al. [41] estimated the AU intensity levels, using logistic regression on SVM scores. In this work, we propose using Platt's probability, given by (3), as action unit intensity at frame t, $I_t(\text{AU}i) = P_{\text{AU}i}^t$, with $P_{\text{AU}i}^t = P(c = \text{AU}i \mid \mathbf{z}_t)$, is estimated using (3). The $I_t(\text{AU}i)$ time series is then smoothed using a Gaussian filter. We denote by $\tilde{I}_t(\text{AU}i)$ the smoothed value. Figure 7 plots, for a participant, the smoothed intensity of the facial AU7, also illustrating its different temporal segments neutral, onset, apex, and offset.

Having defined the AU intensity, the intensity of displayed AUs during a voluntary facial expression (disgust or neutral) is given by

$$\text{AI} = \sum_{i \in \text{DAUs}} \frac{1}{N_i'} \sum_{t \in T_i} \tilde{I}_t(\text{AU}i), \qquad (6)$$

where DAUs is the set of displayed (recognized) facial action units during the considered 30 s window, T_i is the set of
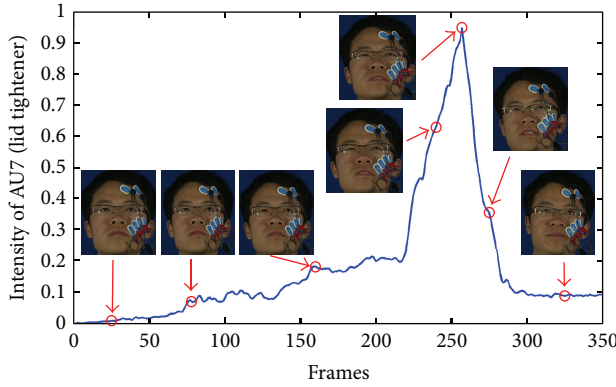
FIGURE 7: Example of AU intensity.

frames where AUi is active, and N_i' is the cardinal of T_i, being the number of frames where AUi is active.

(2) Amount of Movement Change. The amount of movement change during a voluntary facial expression (disgust or neutral) is given by

$$\text{AMC} = \sum_{i \in \text{DAUs}} \frac{1}{N_i'} \sum_{t \in T_i} \left| \tilde{I}_t \left(\text{AU}i \right) - \tilde{I}_{t+1} \left(\text{AU}i \right) \right| \qquad (7)$$

with DAUs, N_i', and T_i as defined above.

3. Results and Discussion

3.1. Manipulation Check. Participants' emotion was successfully elicited using the movie clips of disgust. For both disgust#1 and disgust#2, most participants (14/15) self-reported the target emotion as their dominated emotion. For the other video clips the reported emotions are as follows: amusement#1 (9/15), amusement#2 (12/15), surprise#1 (11/15), surprise#2 (2/15), anger#1 (6/15), anger#2 (9/15), fear#1 (10/15), fear#2 (11/15), neutral#1 (13/15), and neutral#2 (13/15). Annotation of the video records, using the Anvil annotation tool [42], further confirmed that the movie clips of disgust induced the most reliable emotional data. Therefore, for further analysis, we decided to use only the data recorded during watching the disgust movie clips.

The Wilcoxon rank sum tests were implemented on the 14 disgusting self-reports. Results showed that there was no significant difference in self-reported emotional ratings between disgust#1 ($M = 6.50$, SD = .76) and disgust#2 ($M = 6.64$, SD = .63) and between PD group ($M = 6.51$, SD = .65) and control group ($M = 6.64$, SD = .74). This is what we expected since Vicente et al. [43] also reported that PD patients at different stages of the disease did not significantly differ from the controls in the self-reported emotional experience to presented movie clips.

3.2. Univariate Analysis of Group Effects. The EMG data from 3 participants (1 control and 2 PD) and the ECG data from 10 participants (4 control and 6 PD) were discarded because of not being well recorded due to sensor malfunctions. Results

of the repeated-measures ANOVAs on single physiological variables showed that significant main effect of group on the LLS activity ($F(1, 12) = 3.38$, $P = .09$), the OO activity ($F(1, 12) = 1.92$, $P = .19$), HR ($F(1, 3) = .93$, $P = .41$), and HRV ($F(1, 3) = .53$, $P = .83$) was not found. Table 4 shows the descriptive data for control and PD groups during exposure to the three movie clips disgust#1, disgust#2, and neutral#2 and the tests of significance comparison between groups using Wilcoxon rank sum tests. The Wilcoxon rank sum tests revealed that

(1) comparable levels of baseline activity in control and PD groups over both the LLS and the OO were found;

(2) although disgust#1 elicited more muscle activity over the LLS and the OO for control group than for PD group, this difference did not reach statistical significance;

(3) disgust#2 elicited significantly more LLS activity for control than for PD.

These results indicated that PD displayed less muscle activity over the LLS when expressing disgust than control. In addition, disgust#2 induced more muscle activity over the LLS and the OO than disgust#1, which is consistent with the self-report and may be due to the fact that disgust#2 is slightly more disgusting than disgust#1.

3.3. Univariate Analysis of Stimuli Effects. For LLS, we found an effect of the stimuli on muscle activity ($F(2, 24) = 9.47$, $P = .001$). Post-hoc tests indicated significant differences between disgust#1 and neutral#2 ($P = .01$) and between disgust#2 and neutral#2 ($P = .01$). Both disgust#1 ($M = 2.82$, SD = 1.92) and disgust#2 ($M = 5.37$, SD = 5.43) elicited more muscle activity than neutral#2 ($M = 1.10$, SD = .38). This main effect was qualified by a stimuli × group interaction ($F(2, 24) = 4.17$, $P = .028$), which was consistent with the results of the Wilcoxon rank sum tests which compared the physiological responses between groups.

For OO, we also found an effect of the stimuli on muscle activity ($F(2, 24) = 5.45$, $P = .012$). Post-hoc tests indicated significant difference only between disgust#1 and neutral#2 ($P = .002$). The disgust#1 ($M = 2.32$, SD = 1.57) elicited more muscle activity than neutral#2 ($M = 1.35$, SD = 1.11). No significant stimuli × group interaction effect ($F(2, 24) = 1.77$, $P = .192$) was found.

We expected that the disgust clips elicited significantly more muscle activity over LLS than the neutral clip, because normally LLS is involved in producing the disgust facial expression [44]. The fact that OO was also significantly different was probably due to the following:

(1) crosstalk [45], the LLS, and the OO lie in the vicinity of each other;

(2) the disgust#1 elicited is not only a disgust but also a bit of amusement by the funny motion of the character and the background music so that the OO was also involved in producing the facial expression of amusement.

TABLE 4: Statistical summary of physiological measures.

Variable	Stimuli	Control		PD		Sig.
		Mean	SD	Mean	SD	
LLS	Neutral#2	1.04	.34	1.16	.43	.71
	Disgust#1	3.36	2.04	2.27	1.77	.26
	Disgust#2	8.04	6.49	2.71	2.25	.04*
OO	Neutral#2	1.35	1.11	.94	.54	.81
	Disgust#1	2.82	1.77	1.81	1.28	.32
	Disgust#2	5.20	5.71	2.07	1.48	.13
HR	Neutral#2	1.72	2.77	2.76	—	—
	Disgust#1	−.53	2.64	5.12	—	—
	Disgust#2	2.22	5.53	5.65	—	—
HRV	Neutral#2	1.53	12.96	−7.65	—	—
	Disgust#1	8.01	11.45	26.34	—	—
	Disgust#2	14.86	35.64	14.92	—	—

*The two groups are significantly different; that is, $P < .05$.

"—" We did not compare the cardiac responses (i.e., HR and HRV) between groups, because due to technical failures we lost the ECG data for 10 participants and thus only 5 participants (4 controls and 1 PD) completed the ECG recording.

TABLE 5: Facial expressivity assessment based on different methods.

Var.	C*			LP*			IP*			MP*		
	D#1*	D#2*	N#2*	D#1*	D#2*	N#2*	D#1*	D#2*	N#2*	D#1*	D#2*	N#2*
TFA	8	8	4	5	5	4	5	6	4	4	—	5
AMC◊	13.9	27.7	10.8	6.7	6.5	4.3	8.2	6.2	6.9	3.9	—	8.2
AI	5.4	5.1	1.8	2.8	2.8	2.2	3.8	3.9	3.1	2.2	—	3.3
FE◊	72.2	75.1	24.6	20.9	18.7	14.2	44.4	35.6	26.2	13.5	—	32.2

*D#1, D#2, N#2, C, LP, IP, and MP denote disgust#1, disgust#2, neutral#2, the control, the least, intermediate, and most severely form of Parkinson's patients, respectively.

◊presented in percentages.

"—" The face of the MP while watching D#2 was blocked by his hand.

Note that disgust#2 ($M = 3.64$, SD $= 4.32$) elicited more muscle activity over OO (because of crosstalk) than neutral#2 ($M = 1.14$, SD $= .87$); however, unlike disgust#1, the difference did not reach statistical significance (because disgust#2 did not elicit amusement at all) which can also be interpreted.

Moreover, the main effect of stimuli on cardiac parameters for both HR ($F(2, 6) = .37$, $P = .70$) and HRV ($F(2, 6) = 0.84$, $P = .48$) was not found, which is not consistent with what we expected: unchanged [46, 47] or increased [48, 49] HR and increased HRV [21, 47]. This may be due to the fact that we did not have enough recorded ECG data for a statistical analysis.

3.4. Qualitative Analysis of Facial Expressivity. To qualitatively analyze facial expressivity we used the total facial activity TFA measure of [2, 50], being the total number of displayed AUs in response to the stimuli. Compared to the control (C), the PD groups (LP, IP, and MP) showed the attenuation of their facial activities (variable TFA) while watching disgusting movie clips (see Table 5). However, comparable TFA was found while watching neutral movie clips.

Visual inspection of the of the displayed AUs, as illustrated in Figure 8, shows that AU1, AU2, AU6, AU9, and AU45 occurred more frequently for C, except for IP, who produced more frequently AU6 than C. A possible explanation is that the IP cannot deactivate AU6 even during neutral state. As shown in Figure 9, during neutral state, PD patients produced more continuous active action units, such as AU4 and AU25 (for all PD), AU6 (for IP), and AU9 (for MP). Moreover, certain AU combinations were considered likely to signify "disgust" on the basis of Ekman and Friesen's description of emotions [7]. "Disgust" AUs were considered as the combination of eyebrows lowerer (AU4), cheek raiser (AU6), and nose wrinkler (AU9). Because cheek raiser is very difficult to produce on demand without including other AUs, especially the lid tightener (AU7) [51], lid tightener was also taken into consideration; that is, the expected AU pattern of "disgust" was the combination of AU4, AU6, AU7, and AU9. Consistent with what we expected, C displayed 98 and 87 "disgust" frames while watching disgust#1 and disgust#2, respectively. The PD goup (LP, IP, and MP) did not display any "disgust" frames, only for IP who had 4 "disgust" frames while watching disgust#1. Furthermore, the IP displayed quite few frames with combinations of cheek raiser and lid tightener during neutral state. Instead,
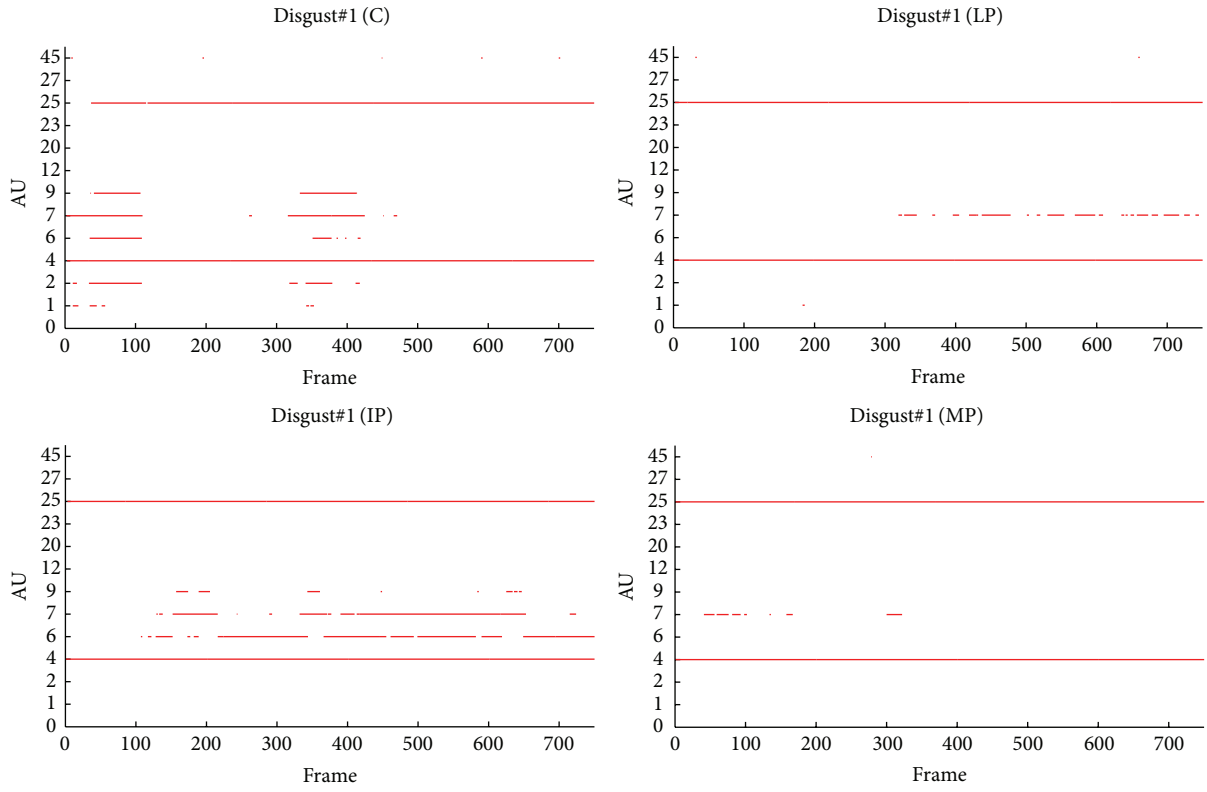
Disgust#1 (C)

Disgust#1 (LP)

Disgust#1 (IP)

Disgust#1 (MP)

FIGURE 8: The facial activities while watching disgust#1.

N(C)

N(LP)

N(IP)

N(MP)
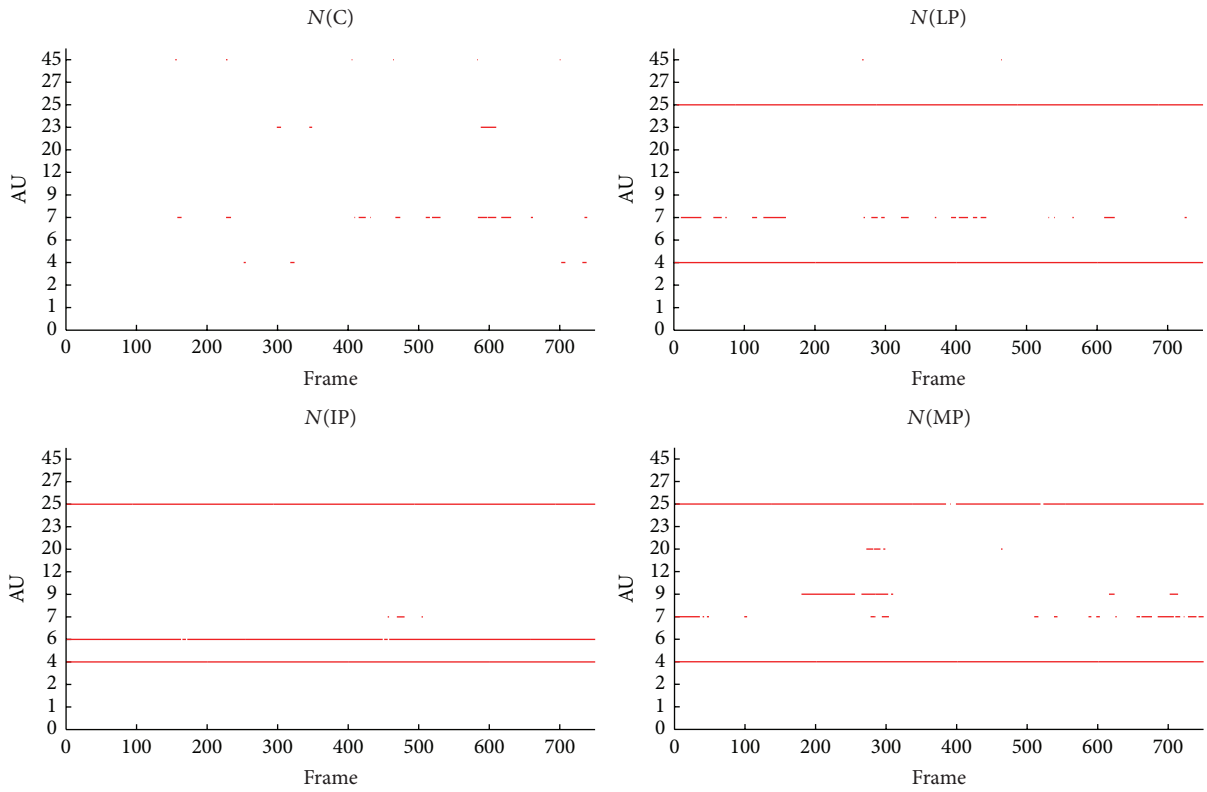
FIGURE 9: The facial activities while watching neutral#2.
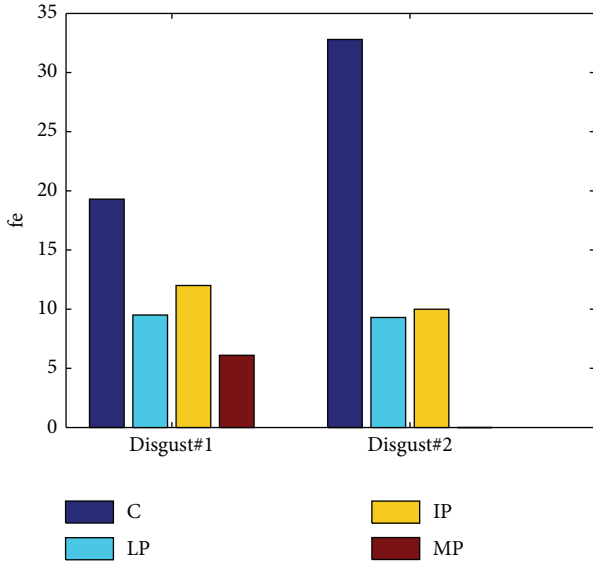
FIGURE 10: The quantified facial expressivity.

FIGURE 11: The quantified facial expressivity based on the improved function.

the cheek raiser alone was displayed mostly (see Figure 9), which indicates that IP patients have control problems for facial muscles combinations.

Analyzing the defined quantities AMC, AI, and FE using the segments with active AUs, as it can be seen from Table 5, the control C had higher facial expressivity than LP, IP, and MP during the disgusting state while MP performed the lowest facial expressivity. More specifically, compared to C, PD (LP, IP, and MP) had smaller values of variables AMC, AI, and FE while watching disgusting movie clips, especially for brow lower (AU4), nose wrinkler (AU9), and blink (AU45). C tended to show similar intensities of brow lower (variable e) with bigger variance (variables AMC and FE). In addition, C showed higher intensities of nose wrinkler with larger variance.

3.5. Quantitative Analysis of Facial Expressivity. Figure 10 depicts the values of the facial expressivity equation (5) for the participants. We did not assess the facial expressivity of the MP patient during disgust#2, as he had his hand on front of his face during the experiments. As it can be seen, a significant difference between C and PD patients is present. C got the highest score while the lowest score was obtained by the MP patient. However, the score of the IP patient was slightly higher than the LP one, which is due to the fact that the LP and IP expressed "facial masking" in different ways: the LP patient showed the attenuation of the intensities of facial movements. On the contrary, the IP patient produced high intensities of facial movements not only in his emotional state but also during neutral state; that is, he cannot relax the muscles well. In order to take both kinds of "facial masking" into account, we computed the facial expressivity using (4). The results are shown in Figure 11. As it can be seen, the proposed facial expressivity allows distinguishing between control and PD patients. Moreover the facial expressivity decreases along with the increase of PD severity.
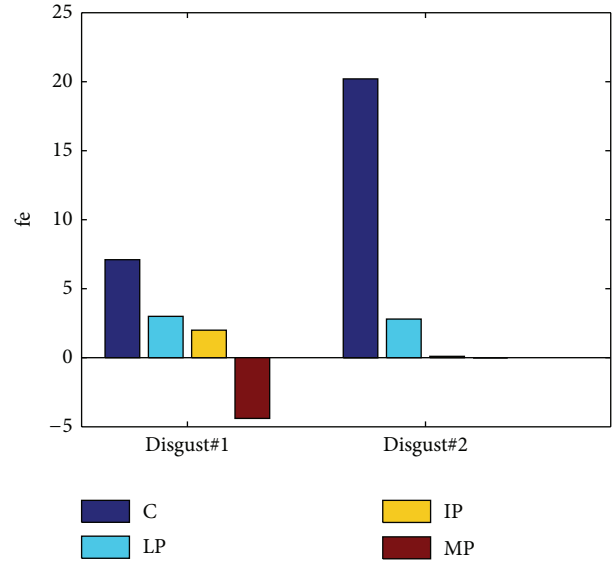
4. Conclusions

This study investigated the phenomenon of facial masking in Parkinson's patients. We designed an automated and objective method to assess the facial expressivity of PD patients. The proposed approach follows the methodology of the *Interpersonal Communication Rating Protocol*-Parkinson's Disease Version (ICRP-IEB) [9]. In this study, based on the Facial Action Coding System (FACS), we proposed a methodology that (i) automatically detects faces in a video stream, (ii) codes each frame with respect to 11 action units, and (iii) estimates a facial expressivity as function of *frequency*, that is, how often AUs occur, *duration*, that is, how long AUs last, and *intensity* being the strength of AUs.

Although, the proposed facial expressivity assessment approach has been evaluated in a limited number of subjects, it allows capturing differences in facial expressivity between control participants and Parkinson's patients. Moreover, facial expressivity differences between PD patients with different progression of Parkinson's disease have been assessed. The proposed method can capture these differences and give a more accurate assessment of facial expressivity for Parkinson's patients than what traditional observer based ratings allow for. This confirms that our approach can be used for clinical assessment of facial expressivity in PD. Indeed, nonverbal signals contribute significantly to interpersonal communication. Facial expressivity, a major source of nonverbal information, is compromised in Parkinson's disease. The resulting disconnect between subjective feeling and objective facial affect can lead people to form negative and inaccurate impressions of people with PD with respect to their personality, feelings, and intelligence. Assessing in an objective way the facial expressivity limitation of PD would allow developing personalized therapy to benefit facial expressivity in PD. Indeed Parkinson's is a progressive disease, which means that the symptoms will get worse as time

goes on. Using the proposed assessment approach would allow regular facial expressivity assessment by therapist and clinicians to explore treatment options.

Future work will (i) improve the AU recognition system and extend it to more AUs and (ii) consider clinical usage of the proposed approach in a statistically significant PD patients population.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.
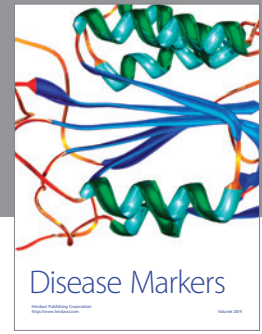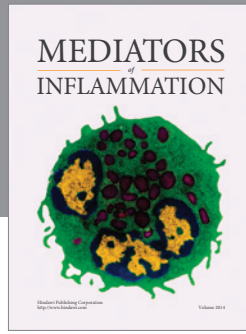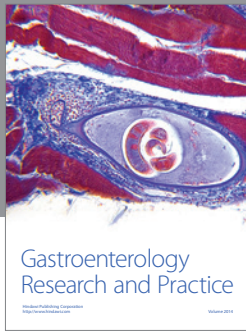
Acknowledgments

References

[1] M. Katsikitis and I. Pilowsky, "A study of facial expression in Parkinson's disease using a novel microcomputer-based method," *Journal of Neurology Neurosurgery and Psychiatry*, vol. 51, no. 3, pp. 362–366, 1988.

[2] G. Simons, H. Ellgring, and M. C. Smith Pasqualini, "Disturbance of spontaneous and posed facial expressions in Parkinson's disease," *Cognition and Emotion*, vol. 17, no. 5, pp. 759–778, 2003.

[3] D. H. Jacobs, J. Shuren, D. Bowers, and K. M. Heilman, "Emotional facial imagery, perception, and expression in Parkinson's disease," *Neurology*, vol. 45, no. 9, pp. 1696–1702, 1995.

[4] L. Tickle-Degnen and K. D. Lyons, "Practitioners' impressions of patients with Parkinson's disease: the social ecology of the expressive mask," *Social Science & Medicine*, vol. 58, no. 3, pp. 603–614, 2004.

[5] J. J. van Hilten, A. D. van der Zwan, A. H. Zwinderman, and R. A. C. Roos, "Rating impairment and disability in Parkinson's disease: evaluation of the unified Parkinson's disease rating scale," *Movement Disorders*, vol. 9, no. 1, pp. 84–88, 1994.

[6] D. Bowers, K. Miller, W. Bosch et al., "Faces of emotion in Parkinsons disease: micro-expressivity and bradykinesia during voluntary facial expressions," *Journal of the International Neuropsychological Society*, vol. 12, no. 6, pp. 765–773, 2006.

[7] P. Ekman and W. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*, Consulting Psychologists Press, Palo Alto, Calif, USA, 1978.

[8] P. Wu, D. Jiang, and H. Sahli, "Physiological signal processing for emotional feature extraction," in *Proceedings of the International Conference on Physiological Computing Systems*, pp. 40–47, 2014.

[9] L. Tickle-Degnen, *The Interpersonal Communication Rating Protocol: A Manual for Measuring Individual Expressive Behavior*, 2010.

[10] G. Sandbach, S. Zafeiriou, M. Pantic, and L. Yin, "Static and dynamic 3D facial expression recognition: a comprehensive survey," *Image and Vision Computing*, vol. 30, no. 10, pp. 683–697, 2012.

[11] I. Gonzalez, H. Sahli, V. Enescu, and W. Verhelst, "Context-independent facial action unit recognition using shape and gabor phase information," in *Proceedings of the 4th International Conference on Affective Computing and Intelligent Interaction (ACII '11)*, vol. 1, pp. 548–557, Springer, Memphis, Tenn, USA, October 2011.

[12] J. C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," in *Advances in Large Margin Classifiers*, A. J. Smola, P. Bartlett, B. Schoelkopf, and D. Schuurmans, Eds., pp. 61–74, The MIT Press, 1999.

[13] J. J. Gross and R. W. Levenson, "Emotion elicitation using films," *Cognition and Emotion*, vol. 9, no. 1, pp. 87–108, 1995.

[14] D. Hagemann, E. Naumann, S. Maier, G. Becker, A. Lürken, and D. Bartussek, "The assessment of affective reactivity using films: validity, reliability and sex differences," *Personality and Individual Differences*, vol. 26, no. 4, pp. 627–639, 1999.

[15] C. L. Lisetti and F. Nasoz, "Using noninvasive wearable computers to recognize human emotions from physiological signals," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, Article ID 929414, pp. 1672–1687, 2004.

[16] J. Hewig, D. Hagemann, J. Seifert, M. Gollwitzer, E. Naumann, and D. Bartussek, "A revised film set for the induction of basic emotions," *Cognition and Emotion*, vol. 19, no. 7, pp. 1095–1109, 2005.

[17] J. H. Westerink, E. L. van den Broek, M. H. Schut, J. van Herk, and K. Tuinenbreijer, "Probing experience," in *Assessment of User Emotions and Behaviour to Development of Products*, T. J. O. W. F. P. Joyce, H. D. M. Westerink, M. Ouwerkerk, and B. de Ruyter, Eds., Springer, New York, NY, USA, 2008.

[18] A. Schaefer, F. Nils, P. Philippot, and X. Sanchez, "Assessing the effectiveness of a large database of emotion-eliciting films: a new tool for emotion researchers," *Cognition and Emotion*, vol. 24, no. 7, pp. 1153–1172, 2010.

[19] F. Verbraeck, *Objectifying human facial expressions for clinical applications [M.S. thesis]*, Vrije Universiteit, Brussel, Belgium, 2012.

[20] O. Alzoubi, *Automatic affect detection from physiological signals: practical issues [Ph.D. thesis]*, The University of Sydney, New South Wales, Australia, 2012.

[21] S. D. Kreibig, "Autonomic nervous system activity in emotion: a review," *Biological Psychology*, vol. 84, no. 3, pp. 394–421, 2010.

[22] F. D. Farfán, J. C. Politti, and C. J. Felice, "Evaluation of emg processing techniques using information theory," *BioMedical Engineering Online*, vol. 9, article 72, 2010.

[23] N. Ambady, F. J. Bernieri, and J. A. Richeson, "Toward a histology of social behavior: judgmental accuracy from thin slices of the behavioral stream," *Advances in Experimental Social Psychology*, vol. 32, pp. 201–271, 2000.

[24] N. Ambady and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: a meta-analysis," *Psychological Bulletin*, vol. 111, no. 2, pp. 256–274, 1992.

[25] K. D. Lyons and L. Tickle-Degnen, "Reliability and validity of a videotape method to describe expressive behavior in persons with Parkinson's disease," *The American Journal of Occupational Therapy*, vol. 59, no. 1, pp. 41–49, 2005.

[26] N. A. Murphy, "Using thin slices for behavioral coding," *Journal of Nonverbal Behavior*, vol. 29, no. 4, pp. 235–246, 2005.

[27] A. O. Andrade, S. Nasuto, P. Kyberd, C. M. Sweeney-Reed, and F. R. van Kanijn, "EMG signal filtering based on empirical mode decomposition," *Biomedical Signal Processing and Control*, vol. 1, no. 1, pp. 44–55, 2006.

[28] M. Blanco-Velasco, B. Weng, and K. E. Barner, "ECG signal denoising and baseline wander correction based on the empirical mode decomposition," *Computers in Biology and Medicine*, vol. 38, no. 1, pp. 1–13, 2008.

[29] A. O. Boudraa, J. C. Cexus, and Z. Saidi, "Emd-based signal noise reduction," *Signal Processing*, vol. 1, pp. 33–37, 2005.

[30] T. Jing-tian, Z. Qing, T. Yan, L. Bin, and Z. Xiao-kai, "Hilbert-Huang transform for ECG de-noising," in *Proceedings of the 1st International Conference on Bioinformatics and Biomedical Engineering*, pp. 664–667, July 2007.

[31] A. Karagiannis and P. Constantinou, "Noise components identification in biomedical signals based on empirical mode decomposition," in *Proceedings of the 9th International Conference on Information Technology and Applications in Biomedicine (ITAB '09)*, pp. 1–4, November 2009.

[32] Y. Kopsinis and S. McLaughlin, "Development of EMD-based denoising methods inspired by wavelet thresholding," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1351–1362, 2009.

[33] F. Agrafioti, D. Hatzinakos, and A. K. Anderson, "ECG pattern analysis for emotion detection," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 102–115, 2012.

[34] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.

[35] B.-U. Köhler, C. Hennig, and R. Orglmeister, "The principles of software QRS detection," *IEEE Engineering in Medicine and Biology Magazine*, vol. 21, no. 1, pp. 42–57, 2002.

[36] V. Guralnik and J. Srivastava, "Event detection from time series data," in *Knowledge Discovery and Data Mining*, pp. 33–42, 1999.

[37] M. Hamedi, S.-H. Salleh, and T. T. Swee, "Surface electromyography-based facial expression recognition in Bi-polar configuration," *Journal of Computer Science*, vol. 7, no. 9, pp. 1407–1415, 2011.

[38] Y. Hou, H. Sahli, R. Ilse, Y. Zhang, and R. Zhao, "Robust shape-based head tracking," in *Advanced Concepts for Intelligent Vision Systems*, vol. 4678 of *Lecture Notes in Computer Science*, pp. 340–351, Springer, 2007.

[39] T. Kanade, J. F. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," in *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 46–53, Grenoble, France, 2000.

[40] M. S. Bartlett, G. C. Littlewort, M. G. Frank, C. Lainscsek, I. R. Fasel, and J. R. Movellan, "Automatic recognition of facial actions in spontaneous expressions," *Journal of Multimedia*, vol. 1, no. 6, pp. 22–35, 2006.

[41] A. Savran, B. Sankur, and M. Taha Bilge, "Regression-based intensity estimation of facial action units," *Image and Vision Computing*, vol. 30, no. 10, pp. 774–784, 2012.

[42] Anvil, the video annotation tool, http://www.anvil-soft-ware.org/.

[43] S. Vicente, J. Péron, I. Biseul et al., "Subjective emotional experience at different stages of Parkinson's disease," *Journal of the Neurological Sciences*, vol. 310, no. 1-2, pp. 241–247, 2011.

[44] A. van Boxtel, "Facial emg as a tool for inferring affective states," in *Proceedings of Measuring Behavior 2010*, A. J. Spink, F. Grieco, O. Krips, L. Loijens, L. Noldus, and P. Zimmeran, Eds., pp. 104–108, Noldus Information Technology, Wageningen, The Netherlands, 2010.

[45] C.-N. Huang, C.-H. Chen, and H.-Y. Chung, "The review of applications and measurements in facial electromyography," *Journal of Medical and Biological Engineering*, vol. 25, no. 1, pp. 15–20, 2005.

[46] R. W. Levenson, P. Ekman, K. Heider, and W. V. Friesen, "Emotion and autonomic nervous system activity in the minangkabau of west sumatra," *Journal of Personality and Social Psychology*, vol. 62, no. 6, pp. 972–988, 1992.

[47] S. Rohrmann and H. Hopp, "Cardiovascular indicators of disgust," *International Journal of Psychophysiology*, vol. 68, no. 3, pp. 201–208, 2008.

[48] O. Alaoui-Ismaïli, O. Robin, H. Rada, A. Dittmar, and E. Vernet-Maury, "Basic emotions evoked by odorants: comparison between autonomic responses and self-evaluation," *Physiology and Behavior*, vol. 62, no. 4, pp. 713–720, 1997.

[49] J. Gruber, S. L. Johnson, C. Oveis, and D. Keltner, "Risk for mania and positive emotional responding: too much of a good thing?" *Emotion*, vol. 8, no. 1, pp. 23–33, 2008.

[50] W. Gaebel and W. Wölwer, "Facial expressivity in the course of schizophrenia and depression," *European Archives of Psychiatry and Clinical Neuroscience*, vol. 254, no. 5, pp. 335–342, 2004.

[51] P. Ekman, W. Friesen, and J. Hager, *Facial Action Coding System: The Manual*, 2002.

The Scientific World Journal

Gastroenterology Research and Practice

MEDIATORS of INFLAMMATION

Journal of Diabetes Research

Disease Markers

Journal of Immunology Research

International Journal of Endocrinology

PPAR Research

Hindawi

Submit your manuscripts at
http://www.hindawi.com

BioMed Research International

Journal of Ophthalmology

Stem Cells International

eCAM
Evidence-Based Complementary and Alternative Medicine

Journal of Obesity

Journal of Oncology

Computational and Mathematical Methods in Medicine

Behavioural Neurology

Parkinson's Disease

AIDS Research and Treatment

Oxidative Medicine and Cellular Longevity