Journal of Neuroscience Methods 281 (2017) 7–20

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

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

Analysis of facial expressions in parkinson's disease through video-based automatic methods

Andrea Bandini [a,b,*], Silvia Orlandi [a], Hugo Jair Escalante [c], Fabio Giovannelli [d], Massimo Cincotta [d], Carlos A. Reyes-Garcia [c], Paola Vanni [d], Gaetano Zaccara [d], Claudia Manfredi [a]

[a] *Department of Information Engineering, Università degli Studi di Firenze, Via di S. Marta 3, 50139 Firenze, Italy*
[b] *Department of Electrical, Electronic and Information Engineering (DEI) "Guglielmo Marconi", Università di Bologna, Viale del Risorgimento 2, 40136, Bologna, Italy*
[c] *Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Computer Science Department, Luis Enrique Erro No. 1, Tonantzintla, Puebla 72840, Mexico*
[d] *Unit of Neurology, Florence Health Autority, Ospedale "Nuovo San Giovanni di Dio", Via Torregalli 3, Firenze, Italy*

HIGHLIGHTS

- Objectify facial hypomimia in Parkinson's disease through video analysis.
- Basic facial expressions of Patients and healthy subjects were investigated.
- The distance from a neutral baseline was computed to quantify expression changes.
- Parkinsonian patients reported on average lower amounts of facial mimicry.
- We demonstrated that contactless methods are suitable for objectifying hypomimia.

ARTICLE INFO

Article history:
Received 2 November 2016
Received in revised form 16 February 2017
Accepted 16 February 2017
Available online 20 February 2017

Keywords:
Parkinson's disease
Hypomimia
Automatic facial expression recognition
Facial mimicry
Contactless
Video-based

ABSTRACT

Background: The automatic analysis of facial expressions is an evolving field that finds several clinical applications. One of these applications is the study of facial bradykinesia in Parkinson's disease (PD), which is a major motor sign of this neurodegenerative illness. Facial bradykinesia consists in the reduction/loss of facial movements and emotional facial expressions called hypomimia.
New method: In this work we propose an automatic method for studying facial expressions in PD patients relying on video-based
Methods: 17 Parkinsonian patients and 17 healthy control subjects were asked to show basic facial expressions, upon request of the clinician and after the imitation of a visual cue on a screen. Through an existing face tracker, the Euclidean distance of the facial model from a neutral baseline was computed in order to quantify the changes in facial expressivity during the tasks. Moreover, an automatic facial expressions recognition algorithm was trained in order to study how PD expressions differed from the standard expressions.
Results: Results show that control subjects reported on average higher distances than PD patients along the tasks.
Comparison with existing methods: This confirms that control subjects show larger movements during both posed and imitated facial expressions. Moreover, our results demonstrate that anger and disgust are the two most impaired expressions in PD patients.
Conclusions: Contactless video-based systems can be important techniques for analyzing facial expressions also in rehabilitation, in particular speech therapy, where patients could get a definite advantage from a real-time feedback about the proper facial expressions/movements to perform.

* Corresponding author at: University Health Network – Toronto Rehabilitation Institute, 550 University Ave, Toronto, ON M5G 2A2, Canada.
E-mail addresses: andrea.bandini@uhn.ca (A. Bandini), silvia.orlandi@unifi.it (S. Orlandi), hugojair@inaoep.mx (H.J. Escalante), fabio.giovannelli@unifi.it (F. Giovannelli), massimo.cincotta@uslcentro.toscana.it (M. Cincotta), kargaxxi@inaoep.mx (C.A. Reyes-Garcia), paola.vanni@uslcentro.toscana.it (P. Vanni), gaetano.zaccara@uslcentro.toscana.it (G. Zaccara), claudia.manfredi@unifi.it (C. Manfredi).

http://dx.doi.org/10.1016/j.jneumeth.2017.02.006
0165-0270/

1. Introduction

One of the most common motor signs of Parkinson's disease (PD) is facial bradykinesia, consisting in the reduction and slowdown of facial movements that may affect both the upper and the lower part of the face (Bologna et al., 2013). Facial bradykinesia results in a reduction or loss of spontaneous facial movements and emotional facial expressions called hypomimia. PD face is often a "masked" or "poker" face, as if they had no interest in the surrounding environment (Tickle-Degnen and Doyle Lyons, 2004; Hemmesch et al., 2009). Facial expressions play a fundamental role in the communication of the emotional state; this process is composed by the "expression", where a specific set of features is produced, and the "perception" through which an observer receives these features and can infer the emotional state of a person (Scherer et al., 2011). However, if the pattern of expressed features is not explicit and does not vary among different emotional states, the underlying emotion cannot be reliably communicated and thus detected by the observer. This is the case of PD patients with facial hypomimia, where external observers may have difficulties in decoding the emotional state behind the "mask", with severe problems in social relationships (Tickle-Degnen and Doyle Lyons, 2004; Hemmesch et al., 2009). Moreover, facial bradykinesia affects all the functions that involve orofacial movements with severe problems of speech, swallowing, drooling, etc. (Bologna et al., 2013; Tjiaden, 2008).

The impairment in spontaneous facial expressions is consistent with the neuroanatomical evidence of damage to the extrapyramidal motor system. Moreover, PD patients may experience difficulties in performing voluntary facial expressions and facial movements (Simons et al., 2004; Simons et al., 2003; Jacobs et al., 1994; Madeley et al., 1995; Bowers et al., 2006). Although voluntary movements are generally less affected than automatic movements, also posed expressions appear to be impaired in PD (Jacobs et al., 1994; Madeley et al., 1995). The main signs of hypomimia in PD are (Bologna et al., 2013): wider palpebral distances that in conjunction with the reduced blink rate gives the impression of a staring expression; flattened nasolabial folds; unintentional lips separation (mouth opening).

Hypomimia may be improved by the dopaminergic therapy, while some studies (Weiss et al., 2010) demonstrated that neurosurgical treatments (in particular the deep brain stimulation – DBS) may worsen this sign.

Today, a big debate in the research community is focused on the nature of the hypomimia in PD patients. Some authors support the hypothesis that the loss of facial expressions could be due to impaired emotion recognition, with main deficits in the decoding of disgust, fear and sadness (Bologna et al., 2013). So, it is still unclear whether facial hypomimia is a pure motor disorder or is a secondary effect of expression recognition impairment.

In this work we propose an automatic method for studying facial expressions in PD patients relying on contactless video-based methods. In fact, applications in psychiatry and neurology (Hamm et al., 2011) demonstrated the high potentials of the algorithms for facial expression recognition. The use of contactless video-based systems like those described in other studies (Bandini et al., 2015a; Bandini et al., 2015b; Bandini et al., 2016) could be a valuable tool for analyzing facial expressions also in rehabilitation (in particular speech therapy), where patients could get a definite advantage from a real-time feedback about the proper facial expressions/movements to perform.

1.1. Automatic analysis of facial expressions

The development of automatic systems for facial expression recognition has become one of the most popular topics in computer science and artificial intelligence, founding a new research field called "affective computing" (Bettadapura, 2012). Examples of applications are: robotics, Human-Computer Interface, video games and entertainment, psychiatry (Hamm et al., 2011), neurology (Wu et al., 2014), automotive (Jabon et al., 2010), etc. These studies are the result of decades of research in psychology that over the 20th century tried to describe in an objective manner facial expressions related to emotional states.

Until a couple of decades ago, facial expression recognition was performed only by the perceptual evaluation of trained raters, who based their work on standards like the Facial Action Coding System (FACS) (Ekman and Friesen, 1977) that allows decomposing a facial expression into specific Action Units (AU). However, this field has significantly benefited from the latest improvements in computer vision and machine learning since the 90 s to date. In general, an automatic facial expression recognizer is composed by 3 parts (Bettadapura, 2012; Sariyanidi et al., 2015):

1. Face detection and tracking. In this first step, a face must be detected in the image or in a video frame. In case of video analysis the face must be tracked along the subsequent video frames, in order to keep track of facial changes over time.
2. Feature extraction. From the region of interest (ROI) detected in the previous step, some geometrical or appearance facial features are extracted. This step is essential to represent the face as a set of several components, thus focusing on those features that could bring information about expression variations. These features can be divided into two categories: permanent features (eyes, eyebrows, nose and lips) and transient features (facial lines and wrinkles that appear only in presence of particular expressions) (Bettadapura, 2012).
3. Expression recognition. After the face has been represented as a set of geometric and/or appearance features, these features are classified in order to predict the class they belong to (i.e., which expression or which combination of AUs). Usually, a classification algorithm is first trained on one or more databases of labeled expressions and then tested on facial expressions that do not belong to the training set.

Several approaches have been proposed to automatically recognize facial expressions. Usually, all of these methods follow the above scheme, but differ for the various algorithms used to address the three steps (Bettadapura, 2012; Sariyanidi et al., 2015; Zeng et al., 2009; Valstar et al., 2012).

Concerning face detection and tracking, one of the most used algorithms is the Active Appearance Model (AAM) (Sariyanidi et al., 2015; Cootes et al., 2001; Matthews and Baker, 2004). AAM builds statistical models of shape and appearance by means of the Principal Component Analysis (PCA) performed on a set of training faces (Cootes et al., 2001; Matthews and Baker, 2004). Other efficient face detection and tracking algorithms used for facial expressions recognition are the Supervised Descending Method (SDM) (Xiong and De la Torre, 2017), the constrained local model (CLM) (Cristinacce and Cootes, 2017) and the particle filtering (Patras and Pantic, 2017).

Considering the facial features used to predict a facial expression, the proposed systems can be divided into two classes: systems that use geometrical (or shape) features and systems that use appearance features (Valstar et al., 2012). Most of the geometrical-based systems exploit the facial points provided by the face tracker, since the tracked landmarks describe the shape of permanent facial features (eyebrows, eyes, nose, lips, face contour, etc.) (Bettadapura, 2012; Sariyanidi et al., 2015; Zeng et al., 2009; Soleymany et al., 2012; Pantic and Patras, 2005). Appearance-based systems, instead, describe facial expressions using the information of the image texture extracted with descriptors such as local binary pattern (LBP), Gabor filters, histogram of oriented gradients (HoG), etc. (Bettadapura, 2012; Sariyanidi et al., 2015; Zeng et al., 2009).

The aim is to describe texture variations that occur in the face while performing an expression, trying to catch characteristic signs such as wrinkle, furrows and facial lines.

Other important differences among the proposed approaches consist in the temporal representation of an expression that could be static or dynamic (Bettadapura, 2012; Sariyanidi et al., 2015). Static representation classifies facial expressions frame-by-frame; in contrast, dynamic (or spatio-temporal) representation considers the temporal dynamic of facial expressions, working on time windows of frames. The latter approach allows identifying the common known dynamics of facial expressions: onset-apex-offset (Sariyanidi et al., 2015). Of course, temporal or static representations give rise to different classification strategies. Commonly used classifiers for facial expressions are: Support Vector Machines (SVM), Adaboost, Neural Networks, Hidden Markov Models (HMM), etc. In particular, HMMs are often used when the temporal dynamic of the expression is required (Sariyanidi et al., 2015; Zeng et al., 2009).

Another important difference among the proposed approaches is the affective model. Some methods aim at classifying basic expressions (Zhao and Pietikäinen, 2009; Jeni et al., 2017), while other methods perform an AU recognition (Hamm et al., 2011; Pantic and Patras, 2005; Tian et al., 2001). Moreover, depending on the databases used to train the classifiers, spontaneous or posed expressions can be detected. The available databases may include posed expressions (Kanade et al., 2017; Gross et al., 2010; Langner et al., 2010), spontaneous expressions (McKeown et al., 2012; Bartlett et al., 2006), or both (Pantic et al., 2017; Lucey et al., 2017).

1.2. Automatic facial expression recognition in PD patients and other neurological disorders

Several works tried to identify the most impaired expressions in PD patients. However, as mentioned above, it is still unclear whether facial hypomimia is a pure motor disorder or it is caused by expression recognition impairment. Simons et al., 2004 (Simons et al., 2004) investigated facial expressivity in PD patients during several tasks: watching video clips, social interactions and expression posing. Results demonstrated that PD patients had a lower expressivity than healthy control (HC) subjects with problems in exhibiting some AUs: AU4 (brow lowerer), AU9 (nose wrinkler), AU10 (upper lip raiser), AU1 + 2 (brow raiser) and AU6 + 12 (check raise plus lip corner pull). Moreover, PD patients had difficulties in the imitation of happiness, surprise and disgust. The reduction of the overall expressivity was also demonstrated by Bowers et al., 2006 (Bowers et al., 2006), although the authors could not state whether some expressions were more impaired than others. The analysis of AUs in PD patients was also performed by other studies (Wu et al., 2014; Vinokurov et al., 2015). In (Wu et al., 2014) the authors combined facial Electromyography (EMG) and classifiers for AUs detection, finding expressive reduction along with the increase of the PD severity. Vinokurov et al. (2015), instead, used the AUs detected by means of a 3D depth camera to train an algorithm for the automatic identification of hypomimia in PD. Although a reduction of facial expressivity was demonstrated for PD patients, no further information was provided about the most impaired expressions and/or the most impaired facial movements.

Another study (Ricciardi et al., 2015) concerns the possible relationship between the reduced expressiveness in PD patients and the impairment in recognizing emotions. The authors demonstrated that PD patients had reduced facial expressiveness (both static and dynamic), with difficulties in acting particular expressions (happiness, surprise and sadness). Moreover, they reported impairment in recognizing disgust, fear and anger. However, from this work it is still not understood if voluntary acted expressions are impaired in PD patients.

Table 1
Characteristics of the two groups recruited for the experiments.

	PD patients		HC subjects	
	Mean	SD	Mean	SD
Age (years)	71.9	9.2	68.4	7.5
Male	13		6	
Female	4		11	
Disease duration (years)	8.2	5.0	–	
Hoehn & Yahr stage	2.1	0.4	–	
UPDRS motor score	17.5	10.3	–	

From these studies it is well accepted that PD patients have a reduced expressivity with respect to HC subjects. However, only few studies provided quantitative information about the impaired facial expressions and/or action units (Simons et al., 2004; Ricciardi et al., 2015). As described above, the implementation of automatic methods for facial expression recognition is a fairly unexplored field in PD patients with hypomimia, despite the huge developments of affective computing. Of the reported studies, only Wu et al., 2014 (Wu et al., 2014) made use of a fully automated facial expression recognizer. In particular, using appearance and geometrical facial features they were able to detect 11 AUs through a binary SVM for the recognition of each AU. Other studies (Bowers et al., 2006; Vinokurov et al., 2015) implemented semi-automatic methods. In (Bowers et al., 2006) a frame differencing algorithm was used to quantify the expression change along the recorded videos; however, it required the manual selection of some facial landmarks by the user. In (Vinokurov et al., 2015), the authors used commercially available software called faceshift (Anon, 2016a) that in conjunction with consumer depth cameras (like Microsoft Kinect) is able to track 3D facial landmarks and drive an avatar. However, this algorithm is person-specific and requires a calibration for each subject.

In order to overcome the aforementioned issues, we propose a fully automated video-based method to provide quantitative information about PD hypomimia.

2. Materials and methods

2.1. Subjects

Seventeen PD patients were recruited at the Department of Neurology of the Hospital “San Giovanni di Dio”, Firenze, Italy. Patients’ age ranged from 53 to 83 years (mean: 71.9 years; standard deviation – SD: 9.2 years). Thirteen patients were male, four were female. At the time of the experiment, disease duration ranged from 2 to 20 years (mean: 7.8 years, SD: 5.1 years). Before the experiment each patient underwent a neurological examination. The Hoehn and Yahr disease stage (Goetz et al., 2004) ranged from 1.5 to 3 (2.1 ± 0.4) and the UPDRS motor score (UPDRS part III (Fahn and Elton, 1987)) ranged from 6 to 43 (17.2 ± 10.1). All PD patients were under levodopa medication and were tested during their “on” state.

A group of seventeen healthy subjects was tested as control group (age: 53–84 years, mean: 68.8 years, standard deviation: 7.5 years), six male and eleven female. A summary of subjects’ characteristics is reported in Table 1, whilst a summary description of the syndrome for each patient is reported in Table 2. Signed informed consent was obtained from all the participants.

2.2. Experimental setup

Each subject was asked to perform the following tasks

- Displaying a neutral expression for at least 10 s;

Table 2
Syndrome description for each patient recruited in the study.

Patient	Age (Years)	Gender	Years of disease	H&Y stage	UPDRS motor score	Summary syndrome description
P1	67	F	6	2	14	Classic PD variant, mainly affecting the left side
P2	53	F	8	2	8	Classic PD variant, mainly affecting the left side
P3	69	M	14	2.5	17	Classic PD variant, mainly affecting the left side
P4	68	M	12	2.5	13	Classic PD variant, mainly affecting the right side
P5	56	M	15	2.5	29	Classic PD variant, mainly affecting the left side
P6	79	M	6	2	13	Tremor dominant PD
P7	74	M	7	2.5	27	Akinetic-rigid PD
P8	78	M	8	2	14	Tremor dominant PD, mainly affecting the right upper limb
P9	82	M	8	3	33	Akinetic-rigid PD
P10	81	M	2	2	11	Tremor dominant PD, mainly affecting the right side
P11	83	F	2	1.5	6	Classic PD variant, mainly affecting the right side
P12	68	M	6	2	9	Akinetic-rigid PD, mainly affecting the right side; early-onset freezing of gait
P13	78	M	3	1.5	14	Tremor dominant PD affecting the right upper limb
P14	79	F	20	2	16	Classic PD variant, mainly affecting the left side
P15	64	M	2	2.5	43	Classic PD variant, mainly affecting the right side
P16	63	M	10	2	8	Akinetic-rigid PD, mainly affecting the right side; marked motor fluctuations
P17	80	M	4	2	18	Classic PD variant, mainly affecting the left side

- Displaying basic expressions (happiness, anger, disgust and sadness) upon request of the clinician;
- Displaying basic expressions (happiness, anger, disgust and sadness) by imitating emotive faces shown on a screen (Fig. 1).

Thus, for each subject we obtained: 1 neutral video and 8 expressive videos (4 with acted/requested expressions and 4 with imitated expressions).

The subjects' face was recorded using the Microsoft Kinect for Windows sensor, as in (Bandini et al., 2016) for studying facial kinematics during a syllable repetition task. The Kinect sensor was placed in front of the subject's face at a distance between 0.5 and 0.7 m from the mouth and at a height close to that of the subject's eyes. However, unlike the experiments performed on speech articulatory movements (Bandini et al., 2016), here we used just the color stream (like a standard webcam), which makes the methodology broadly applicable. Videos were acquired at 640 × 480 pixels at 30 frames per second (fps). Color frames were recorded in 24-bit RGB images (8 bits per channel). Color stream was recorded and stored in avi files through the OpenNI (ver. 2.2) and OpenCV (ver. 2.4.9) libraries using a customized code written in C++ language.

The experiments were carried out in a quiet room of the "San Giovanni di Dio" hospital and subjects were required to stay seated during the test.

On the recorded videos, two tests were performed

- The first one consisted in the analysis of facial features with respect to the neutral baseline, in order to find the most discriminative features between HC subjects and PD patients, characterizing the deviant PD expressions, and giving an objective quantification of facial hypomimia in PD;
- The second test concerned the implementation of automatic facial expressions recognition algorithms, to study how the PD expressions differed from the standard expressions (in terms of classification results).

In the first case the aim was measuring the amount of facial movements in PD patients in order to find objective indices of facial mimicry in impaired expressions. However, since each person may show facial expressions in different ways we used the neutral videos to build a baseline for each subject, quantifying how the expressions (acted and imitated) differed from the personal baseline.

In the second case, the aim was that of assessing the intensity of facial expressions that PD patients were able to reach, with respect to HC subjects. This could provide useful information and real-time feedback for speech therapy, in order to elicit more enhanced facial movements, reaching a higher intensity of the expression.

2.3. *Analysis of expressive features with respect to the neutral baseline*

Each video was manually labeled, detecting the acted or the imitated expression (i.e. expected expression). For each PD patient/HC subject we got 9 videos: 1 for the neutral state, 4 for the posed expressions, 4 for the imitated expressions (with the exception of one patient that was not able to display sadness and disgust as requested by the clinician). The duration of these videos ranged between 3 s and 12 s (7.2 ± 2.4 s).

2.3.1. *Face tracking and facial features extraction*

The automatic identification of facial landmarks was performed through the Intraface tracking algorithm (Xiong and De la Torre, 2017), already used to study facial and articulatory movements in PD patients and HC subjects during various speech tasks (Bandini et al., 2015a; Bandini et al., 2015b; Bandini et al., 2016). Starting from the 49 facial landmarks tracked by this algorithm (Fig. 2), according to Soleymani et al., 2012 (Soleymany et al., 2012) we focused on the following 20 geometric features:

- Four eyebrows features: two angles between the horizontal line connecting the inner corners of the eyes (segment 23–26) and the line that connects inner and outer eyebrow (segment 2–5 for the right eyebrow and segment 6–9 for the left eyebrow); two vertical distances from the outer eyebrows (point 2 for the right eyebrow and point 9 for the left eyebrow) to the line that connects the inner corners of the eyes (segment 23–26).
- Ten eyes features: two distances between the outer eyes' corner and their upper eyelids (segment between point 20 and midpoint 21–22 for the right eye, segment between point 29 and midpoint

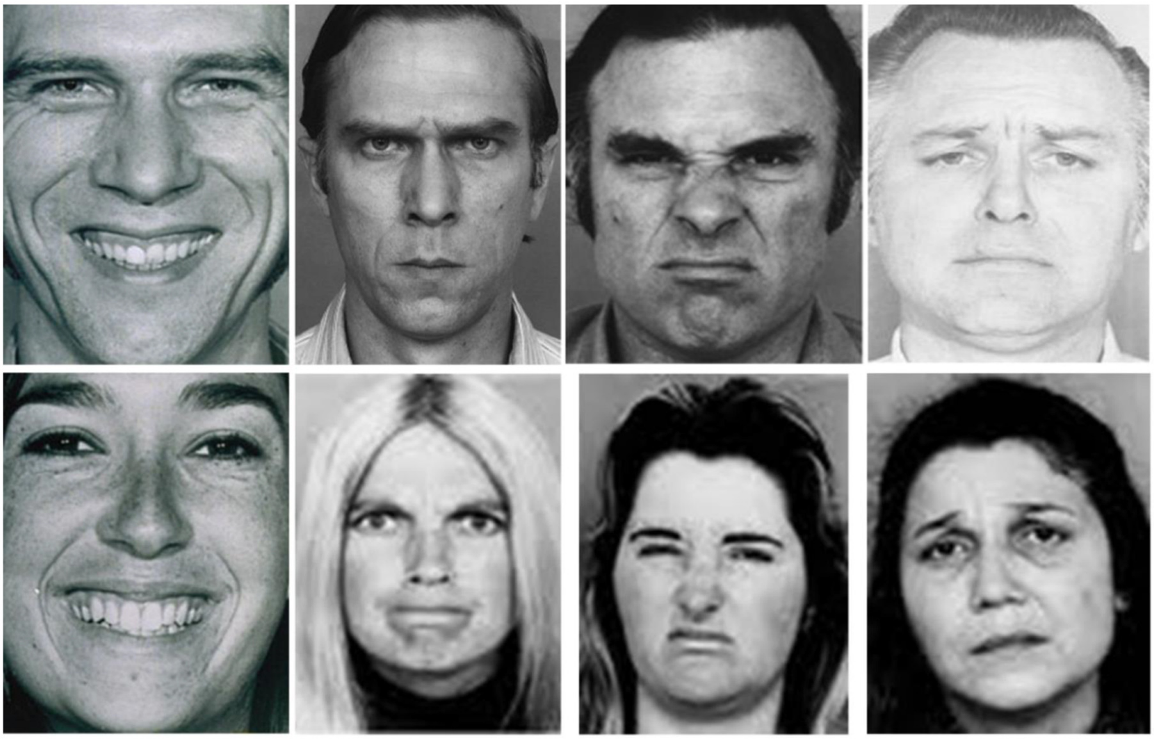

Fig. 1. Facial expression for the imitation task. From left to right: happiness, anger, disgust and sadness, for men (upper row) and women (lower row) (Ekman and Friesen, 1977).

27–28 for the left eye); two distances between the inner eyes' corner and their upper eyelid (segment between point 23 and midpoint 21–22 for the right eye, segment between point 26 and midpoint 27–28 for the left eye); two distances between the outer eyes' corner and their lower eyelids (segment between point 20 and midpoint 25–24 for the right eye, segment between point 29 and midpoint 31–30 for the left eye), distances between the inner eyes' corner and their lower eyelids (segment between point 23 and midpoint 25–24 for the right eye, segment between 26 and midpoint 31–30 for the left eye); vertical distances between the upper eyelids and the lower eyelids (segment between midpoint 21–22 and midpoint 25–24 for the right eye, segment between midpoint 27–28 and midpoint 31–30 for the left eye).
- Six mouth features: two distances between the upper lip and mouth corners (segments 32–35 and 35–38); two distances between the lower lip and mouth corners (segments 32–41 and 38–41); horizontal distance between the mouth corners (segment 32–38) and vertical distance between the upper and the lower lip (segment 35–41).

The line that connects the inner eye corners (segment 23–26, Fig. 2) was used as a reference line since the inner eye corners are stable facial points, i.e., changes in facial expression do not induce any changes in the position of these points. The aforementioned 20 features were extracted for all the video frames.

2.3.2. Baseline building

As PD patients may exhibit impairments of facial expressions/movements and since each person may show facial expressions in different ways, the building of a neutral baseline for each subject was necessary. For this step we considered only the neutral videos collected from each subject. Thus, each PD patient and HC subject had his/her own neutral baseline (i.e. a vector of 20 facial features), from which we computed the displacement of the facial features during the expressive tasks. The neutral baseline was built as follows:

- for each video frame of the neutral video the 49 facial landmarks were detected by means of the Intraface tracking algorithm;
- An average facial template from the facial landmarks vector was iteratively built by Procrustes analysis, as already done in (Hamm et al., 2011). For each video frame, each facial landmark was aligned to the template, updating the template by averaging the aligned landmarks;
- Once the average facial template was computed (i.e., a vector of 49 2D points), the 20 facial features were calculated as above.

2.3.3. Facial feature extraction from the expressive videos

After the set-up of the neutral baseline, the analysis was performed on the expressive videos (both acted and imitated expressions). Thus, for each subject and for each expressive video:

- the 49 2D facial points provided by the Intraface tracking algorithm were extracted from each video frame;
- For each frame the extracted facial model was aligned to the neutral template of the current subject by affine transformation (that includes rotations, translations, scaling and skewing), in order to suppress within-subject head pose variations. This affine transformation was estimated from 4 pairs of corresponding points in the facial model of the current frame and in the average neutral model: the two inner corners of the eyes, the nose tip and the point between the two eyes (points 23, 26, 14 and 11 in Fig. 2). These points were chosen because of their stability with respect to non-rigid facial movements (Pantic and Patras, 2005).
- Once the face model for the current video frame was aligned to the neutral template, the 20 facial features were extracted as above.

Thus, for each expressive video, a vector of 20 features for each frame was obtained.

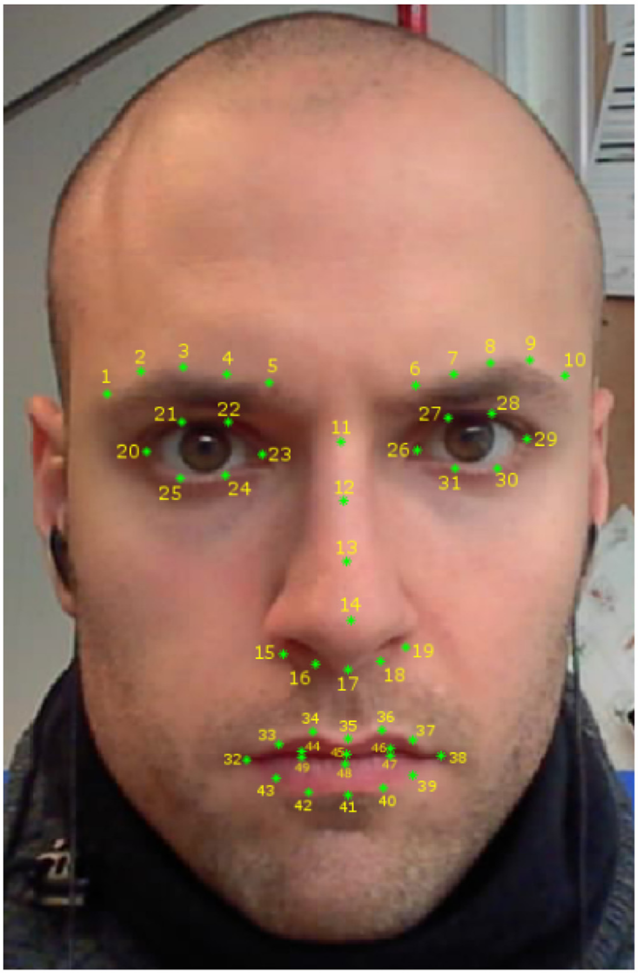

Fig. 2. Indexed Intraface points from which the facial features were computed.

2.3.4. Facial features analysis

Given the extracted parameters (i.e. facial features of the neutral baseline and facial features from the expressive videos), for each subject and for each expressive videos, the following processing was performed. For each video frame the Euclidean distance was computed between the facial features vector and the baseline features vector built for the current subject. This distance provides global information about the displacement of facial features from the neutral expression, during the displaying of different facial expressions. On the Euclidean distance vector, the following statistics were calculated along the duration of the expressive video: mean value, standard deviation, maximum value, minimum value and range (i.e., the difference between maximum and minimum values). Thus, for each subject and for each expressive video, we got five statistic features of the Euclidean distance between the 20 "expressive" facial features and the neutral baseline.

2.4. Automatic facial expression recognition

In this part of the study an automatic facial expression classifier was trained on different existing databases of facial expressions. Most of the available facial expressions databases are composed by videos and images taken from healthy subjects. Starting from this consideration, the aim was to measure the intensity of facial expressions that PD patients were able to reach, with respect to healthy control subjects that were assumed to express "standard" expressions. This measure was obtained through the prediction given by the classification algorithms. In this work supervised learning methods were implemented. Therefore the procedure consists of two steps:

- training phase – the classification algorithm was trained on labeled images obtained from reference databases;
- Test phase – the trained algorithm was tested on the subjects recruited in this study (both PD patients and HC subjects).

Video recordings considered for the test phase were the same used for the analysis of facial features with respect to the neutral baseline.

2.4.1. Training phase

The same 20 geometric features described in the previous section were used to train the classifier. This choice is reasonable in terms of dimensionality (20 features) and characterization of the face, as we can describe the behavior of the main components of the face (eyebrows, eyes and mouth). A Multi-class Support Vector Machine (SVM) was used. SVM is a supervised learning algorithm that divides a feature space into two classes through a linear separation (i.e., a hyperplane). This hyperplane is built exploiting the maximum separation margin between the two classes. One of the major advantages of SVM is its capability to deal also with problems in which features are not linearly separable. Through the so called "kernel trick" it is possible to map the original features into a higher-dimensional space in which features can be separable. Popular kernel classes are: polynomial functions and Gaussian radial basis function (as the one used in our work) (Hastie et al., 2009; Anon, 2016b; Anon, 2016c). One-versus-rest classifiers were built in order to perform a multiclass classification, training a single classifier per class. Thus, we trained five classifiers for the five classes considered in this study (neutral, happiness, anger, disgust, and sadness).

Two databases were considered for training: the Extended Cohn-Kanade database (CK+) and the Radboud Faces database (RaFD) (Langner et al., 2010; Lucey et al., 2017). The CK+ database contains facial expressions from 210 adults of different age (18–50 years old), culture and race. Participants performed sequences of facial movements, including single action units, combination of action units and posed expressions. Each sequence begins with a neutral expression. Of the whole database, 327 image sequences include a nominal emotional label manually checked from certified FACS coders, thus we assumed this set as reliably annotated. Afterwards, in order to enrich the database and make it more balanced, for each subject and for each expression sequence we considered the last 4 frames (i.e. the frames were the particular facial expression is more enhanced). Conversely, for each subject and for each sequence we chose the first frame to build the neutral expression class.

The RaFD database is composed by images of posed expressions obtained from 57 adults (Caucasian males and females and Moroccan Dutch males) and 10 Caucasian Dutch children (both boys and girls). Each subject showed eight facial expressions (neutral, anger, sadness, fear, disgust, surprise, happiness, and contempt, the same of the CK+ database) with three gaze directions and got from five different camera angles (in steps of 45°) (Langner et al., 2010). In this work, we excluded children images, taking into account only 3 viewpoints (45°, 135°, and 90° – frontal view), thus excluding profile views.

Samples from both databases were merged obtaining the following instances: 840 frames for neutral, 693 frames for anger, 749 frames for disgust, 789 frames for happiness, 625 frames for sadness, for a total of 3696 samples from both databases. Then, the Multi-Class SVM was trained on these samples. Our experiments were controlled and we knew the target expression that each sub-

Table 3
Confusion matrix, precision, recall and f-measure for the selected classifier. The multiclass accuracy is 88% (Neu – Neutral, Ang – Anger, Dis – Disgust, Hap – Happiness, Sad – Sadness).

		Predicted					Precision	Recall	F-Measure
		Neu	Ang	Dis	Hap	Sad			
Actual	Neu	714	37	41	4	44	0.83	0.85	0.84
	Ang	42	608	19	0	24	0.88	0.88	0.88
	Dis	9	3	736	1	0	0.82	0.98	0.90
	Hap	2	2	17	767	1	0.99	0.97	0.98
	Sad	93	39	81	1	411	0.86	0.66	0.74

ject had to reach during the tasks, therefore we forced the classifier to recognize only these five expressions.

Before training the classifier, a landmark template from training data (the samples from CK+ an RaFD databases) was built by Procrustes analysis. For each frontal training face, each facial landmark (extracted with Intraface) was aligned to the template, updating the template by averaging the aligned landmarks. Once the template was set up, each shape of the training set was aligned to the template by estimating the affine transformation. This step was essential to suppress within-subject head pose variations and inter-subject geometric differences. As already introduced in the previous section, the affine transformation was estimated from 4 pairs of corresponding points: the two inner corners of the eyes, the nose tip and the point between the two eyes (on the top of the nose). These points were chosen because of their stability with respect to non-rigid facial movements.

A10-fold cross-validation was performed to assess the accuracy of the classifier on the training set, obtaining an overall accuracy of 88%. In Table 3, confusion matrix and other measures of the classification performances (precision, recall and f-measure) are reported.

2.4.2. Test phase

Once the classifier has been trained, we used our database composed by HC subjects and PD patients recruited during the experiments as test set. Thus, for each subject and for each expressive video (8 videos with 4 imitated expressions and 4 acted expressions), we got the predicted facial expression label for each video frame. Since the Multi-Class SVM is a one-vs-rest classifier, each classifier produces the predicted class likelihoods by giving two scores: one corresponding to the negative class and one corresponding to the positive class. For instance, if the classifier is trained to recognize between happy and non-happy instances, the first score concerns the non-happy class, while the second score is related to the happy class. Thus, the predicted facial expression label was determined as the maximum among all the positive scores returned by the Multi-class SVM (winner-takes-all strategy).

3. Results

In this section we present the results concerning the two analyses performed. First we report the analysis of facial features with respect to the neutral baseline, characterizing the deviant PD expressions, and giving an objective quantification of facial hypomimia in PD. Afterwards, we report results of the automatic facial expressions recognition algorithm, to study how PD expressions differed from the standard expressions.

3.1. Analysis of expressive features with respect to the neutral baseline

For each subject the Euclidean distance between the expressive features and the corresponding neutral baseline of that subject was computed. As already mentioned in the previous section, for each expressive video, the following measures related to the distance were calculated: mean value, standard deviation, maximum value, minimum value and range. On average, HC subjects reported higher distances than PD patients along the whole tasks (12.68 ± 5.05 for HC subjects, 9.35 ± 3.85 for PD patients, $p < 0.00001$). Then, for each one of the 4 expressions, the following comparison on the aforementioned measures was performed: PD patients vs HC subjects during the acted expressions (PD_{act} vs HC_{act}), PD patients vs HC subjects during the imitated expressions (PD_{im} vs HC_{im}), PD patients during the acted expressions vs PD patients during the imitated expressions (PD_{act} vs PD_{im}) and HC subjects during the acted expressions vs HC subjects during the imitated expressions (HC_{act} vs HC_{im}). Results are reported in Tables 4–7. Within group differences (PD_{act} vs PD_{im} and HC_{act} vs HC_{im}) are reported in bold, while between-group statistically significant differences (PD_{act} vs HC_{act} and PD_{im} vs HC_{im}) are highlighted in grey. Both differences were evaluated through a two-tailed *t*-test. Differences were considered statistically significant for $p < 0.05$.

3.1.1. Comparison between groups (PD_{act} vs HC_{act} and PD_{im} vs HC_{im})

Significant differences were found in the imitation of anger. PD patients showed lower values of mean (8.44 ± 2.80 vs 12.50 ± 4.95, $p = 0.007$), maximum (13.47 ± 4.08 vs 17.40 ± 5.96, $p = 0.033$)

Table 4
Distance measures during the anger expression for PD patients and HC subjects. Results are reported for both tasks (acted and imitated expression). Statistically significant differences within groups (between acted and imitated) are reported in bold. Statistically significant differences between PD patients and HC subjects are highlighted in grey.

	Distance measures	HC subjects	PD patients
Acted	Mean	**10.13 ± 3.44**	8.05 ± 3.38
	SD	1.93 ± 0.96	2.11 ± 1.43
	Max	16.11 ± 6.20	14.27 ± 5.75
	Min	**4.94 ± 1.98**	4.03 ± 1.51
	Range	11.17 ± 6.08	10.24 ± 4.99
Imitated	Mean	**12.50 ± 4.95**	8.44 ± 2.80
	SD	1.65 ± 0.59	1.73 ± 1.05
	Max	17.40 ± 5.96	13.47 ± 4.08
	Min	**8.12 ± 4.71**	4.96 ± 2.20
	Range	9.28 ± 3.38	8.51 ± 4.38

Table 5
Distance measures during the disgust expression for PD patients and HC subjects. Results are reported for both tasks (acted and imitated expression). Statistically significant differences within groups (between acted and imitated) are reported in bold. Statistically significant differences between PD patients and HC subjects are highlighted in grey.

	Distance measures	HC subjects	PD patients
Acted	Mean	**12.77 ± 4.98**	9.38 ± 3.25
	SD	2.50 ± 1.15	2.31 ± 1.10
	Max	18.73 ± 6.84	15.62 ± 5.28
	Min	**6.78 ± 4.74**	4.96 ± 1.73
	Range	11.95 ± 5.61	10.65 ± 4.60
Imitated	Mean	**14.88 ± 5.22**	10.38 ± 4.79
	SD	2.11 ± 1.10	2.69 ± 3.19
	Max	20.24 ± 6.12	18.89 ± 8.17
	Min	**9.75 ± 5.28**	5.34 ± 2.31
	Range	10.49 ± 4.58	10.54 ± 8.25

Table 6
Distance measures during the happiness expression for PD patients and HC subjects. Results are reported for both tasks (acted and imitated expression). Statistically significant differences within groups (between acted and imitated) are reported in bold. Statistically significant differences between PD patients and HC subjects are highlighted in grey.

	Distance measures	HC subjects	PD patients
Acted	Mean	13.86 ± 5.86	10.94 ± 4.84
	SD	3.68 ± 2.34	2.81 ± 0.93
	Max	24.75 ± 9.95	17.80 ± 5.84
	Min	**6.17 ± 3.10**	5.45 ± 3.44
	Range	**18.57 ± 9.43**	12.35 ± 3.69
Imitated	Mean	14.17 ± 6.62	10.21 ±4.52
	SD	2.49 ± 1.29	2.42 ± 1.11
	Max	21.22 ± 6.85	16.01 ± 5.57
	Min	**8.54 ± 5.14**	5.26 ± 2.56
	Range	**12.69 ± 5.23**	10.75 ± 4.79

and minimum distance (4.96 ± 2.20 vs 8.12 ± 4.71, p = 0.020), as reported in Table 4.

During disgust, significant differences were found in both tasks (acted and imitated expressions) for mean distance, with lower values in PD patients (Table 5). During the acted expression of disgust, mean distance was 9.38 ± 3.25 for PD patients and 12.77 ± 4.98 for HC subjects (p = 0.027), while during the imitated disgust mean distance was 10.38 ± 4.79 for PD patients and 14.88 ± 5.22 for HC subjects (p = 0.014).

Concerning happiness, significant differences were found in both tasks (Table 6). During the acted expression PD patients showed reduced maximum (17.80 ± 5.84 for PD patients, 24.75 ± 9.95 for HC subjects, p = 0.020) and range values of distance (12.35 ± 3.69 for PD patients, 18.57 ± 9.43 for HC subjects, p = 0.019). During the imitation task, PD patients showed reduced maximum (16.01 ± 5.57 for PD patients, 21.22 ± 6.85 for HC subjects, p = 0.021) and minimum values of distance (5.26 ± 2.56 for PD patients, 8.54 ± 5.14 for HC subjects, p = 0.028).

Significant differences were also found during the display of sadness (Table 7). Concerning the acted expression, PD patients showed lower mean values of distance (8.76 ± 2.48 for PD patients, 11.14 ± 3.47 for HC subjects, p = 0.031). During the imitation task, PD patients showed lower mean (8.63 ± 2.94 for PD patients, 11.97 ± 4.13 for HC subjects, p = 0.011), maximum (13.73 ± 3.99 for PD patients, 16.91 ± 4.52 for HC subjects, p = 0.037) and minimum values of distance (4.99 ± 2.15 for PD patients, 7.45 ± 3.11 for HC subjects, p = 0.012).

3.1.2. Comparison within groups (PD_{act} vs PD_{im} and HC_{act} vs HC_{im})

During anger HC subjects showed an increase of the following parameters from the acted expression to the imitation task (Table 4): mean (10.13 ± 3.44 during acted anger, 12.50 ± 4.95 during imitated anger, p = 0.049) and minimum values of distance (4.94 ± 1.98 during acted anger, 8.12 ± 4.71 during imitated anger, p = 0.011). The same differences were also found during disgust (Table 5): HC subjects showed an increase (from the acted task to the imitated task) of mean (12.77 ± 4.98 during acted disgust, 14.88 ± 5.22 during imitated disgust, p = 0.0056) and minimum values of distance (6.78 ± 4.74 during acted disgust, 9.75 ± 5.28 during imitated disgust, p = 0.0015).

Concerning happiness, HC subjects showed an increase from the acted expression to the imitated expression in minimum value of distance (6.17 ± 3.10 during acted happiness, 8.54 ± 5.14 during

Table 7
Distance measures during the sadness expression for PD patients and HC subjects. Results are reported for both tasks (acted and imitated expression). Statistically significant differences within groups (between acted and imitated) are reported in bold. Statistically significant differences between PD patients and HC subjects are highlighted in grey.

	Distance measures	HC subjects	PD patients
Acted	Mean	11.14 ± 3.47	8.76 ± 2.48
	SD	2.37 ±1.39	1.91 ± 0.97
	Max	17.55 ± 6.28	13.89 ± 4.02
	Min	**5.54 ± 2.12**	4.92 ± 1.68
	Range	12.01 ±5.99	8.97 ± 3.87
Imitated	Mean	11.97 ± 4.13	8.63 ± 2.94
	SD	1.71 ± 0.65	1.98 ± 0.93
	Max	16.91 ± 4.52	13.73 ± 3.99
	Min	**7.45 ± 3.11**	4.99 ± 2.15
	Range	9.46 ± 3.36	8.74 ± 3.31

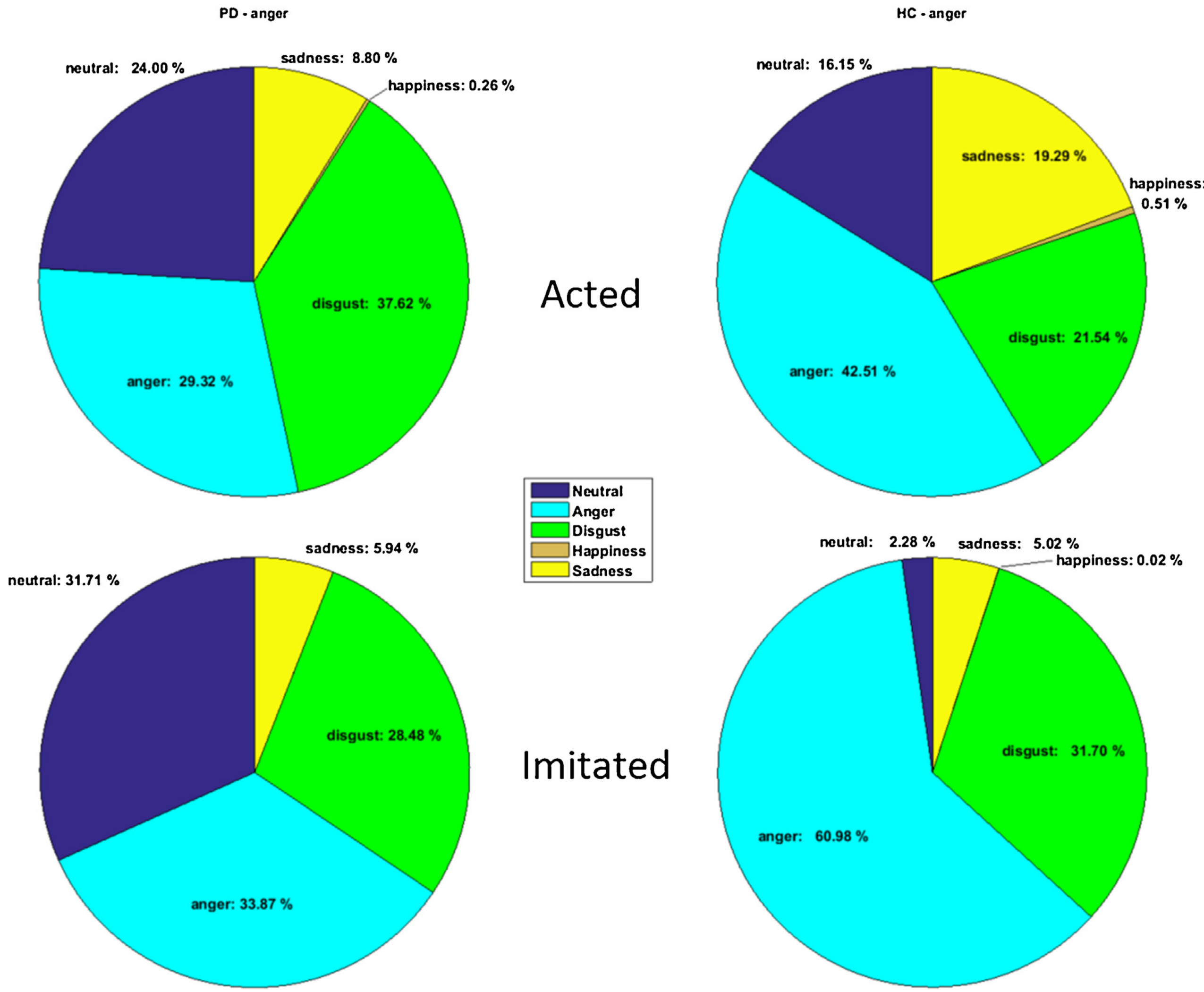

Fig. 3. Pie charts for anger: acted expressions (upper plots) and imitated expressions (lower plots). The two plots on the left concern PD patients, while those on the right concern HC subjects.

imitated happiness, $p = 0.039$) and a decrease of the range value of distance (18.57 ± 9.43 during acted happiness, 12.69 ± 5.23, $p = 0.028$), as reported in Table 6.

During sadness (Table 7), only a significant increase for HC subjects in the minimum value of distance from the acted to the imitated expression was found (5.54 ± 2.12 during acted sadness, 7.45 ± 3.11 during imitated sadness, $p = 0.005$).

No significant differences were found in the PD group between the acted expressions and the imitated expressions.

3.2. Automated facial expression recognition

For each expressive video the SVM-scores were extracted. The multiclass SVM provides 5 scores (1 for each class), thus for each video frame the maximum score was computed, retrieving the predicted class among: neutral, anger, disgust, happiness and sadness. This procedure provided the total number of frames assigned to a particular class for an expressive video. In order to test whether predictions match with the target expressions of a particular video, the total number of occurrences of each class was computed for both groups (PD patients and HC subjects) and for the acquisition tasks (acted and posed expressions). Afterwards, these values were normalized with respect to the overall number of video frames analyzed, in order to get a percentage of a particular expression. This percentage represents the frequency of displayed expression that occurred in a video.

For each one of the 4 target expressions (both acted and imitated) the pie plots of these percentages are shown in Figs. 3–6. For each target expression they provide a visual representation of the average expressions recognized by the classifier in a particular video/task.

3.2.1. Anger

Fig. 3 shows that HC subjects display the expression of anger better than PD patients during the acted task (42.51% vs 29.32%). In both groups, the other most frequent expression is disgust (21.54% in HC subjects and 37.62% in PD patients). Neutral expression is more evident in PD patients than in HC subjects (24.00% vs 16.15%), while the percentage of sadness exceeds 10% only in HC subjects (19.29%).

During the imitation task there is a noticeable increase of the target expression only in HC subjects (from 42.51% to 60.98%), while the increase of the anger percentage in PD patients is around 4% (from 29.32% to 33.87%). On the other hand, PD patients show an increase of the neutral expression (from 24.00% to 31.71%), while this expression almost disappears in HC subjects.

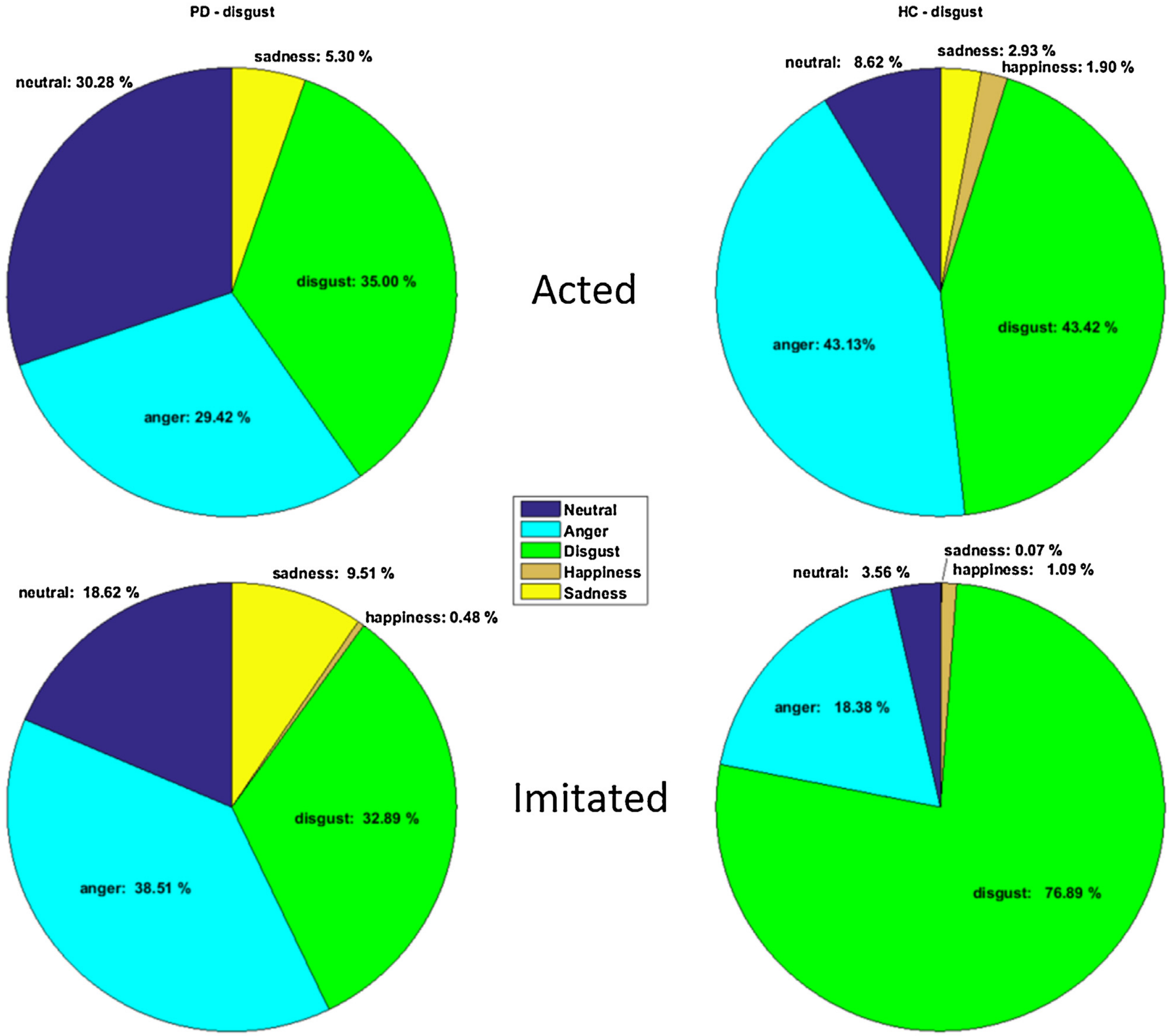

Fig. 4. Pie charts for disgust: acted expressions (upper plots) and imitated expressions (lower plots). The two plots on the left concern PD patients, while those on the right concern HC subjects.

3.2.2. *Disgust*

Fig. 4 shows that HC subjects display the expression of disgust better than PD patients during the acted task (43.42% vs 35.00%). In both groups, a large percentage of anger is present (43.13% in HC subjects and 29.42% in PD patients), with PD patients also showing higher percentage of neutral (30.28%). During the imitation task there is a large increase of the target expression only in HC subjects (from 43.42% to 76.89%), while in PD patients there is even a decrease of the disgust (from 35.00% to 32.89%) at the expense of an increase of the anger expression (from 29.42% to 38.51%). During the imitation PD patients still show a higher percentage of neutral than HC subjects (18.62% vs 3.56%).

3.2.3. *Happiness*

Fig. 5 shows that HC subjects display the expression of happiness better than PD patients during the acted task (31.58% vs 18.09%). However, in both groups, happiness is not the most frequent expression, since its percentage is 49.93% in HC subjects and 32.97% in PD patients. During the imitation, the percentage of happiness decreases in both groups (from 31.58% to 29.84% in HC subjects, from 18.09% to 10.00% in PD patients). Also in this case disgust is the most frequent expression (54.98% for HC subjects and 30.34% for PD patients), with an increase of neutral in PD patients from 25.19% to 29.44%.

3.2.4. *Sadness*

Fig. 6 shows that both groups have low percentage of sadness. This is the only case when the percentage of the target expression is higher in PD patients than in HC subjects (9.30% vs 5.50% during the acted expression, 11.81% vs 5.00% during the imitating task). In both groups, during the acted expression there are high percentages of disgust (44.58% in HC subjects and 27.99% in PD patients) and anger (42.77% in HC subjects and 31.74% in PD patients). During the acted expression there is a large increase of the anger percentage in HC subjects (from 42.77% to 59.95%), while PD patients show an increase of disgust (from 27.99% to 36.69%), with a consequent decrease of anger (from 31.74% to 28.82%) and neutral (from 30.98% to 22.68%). However, neutral still remains higher in PD patients than in HC subjects for both tasks.

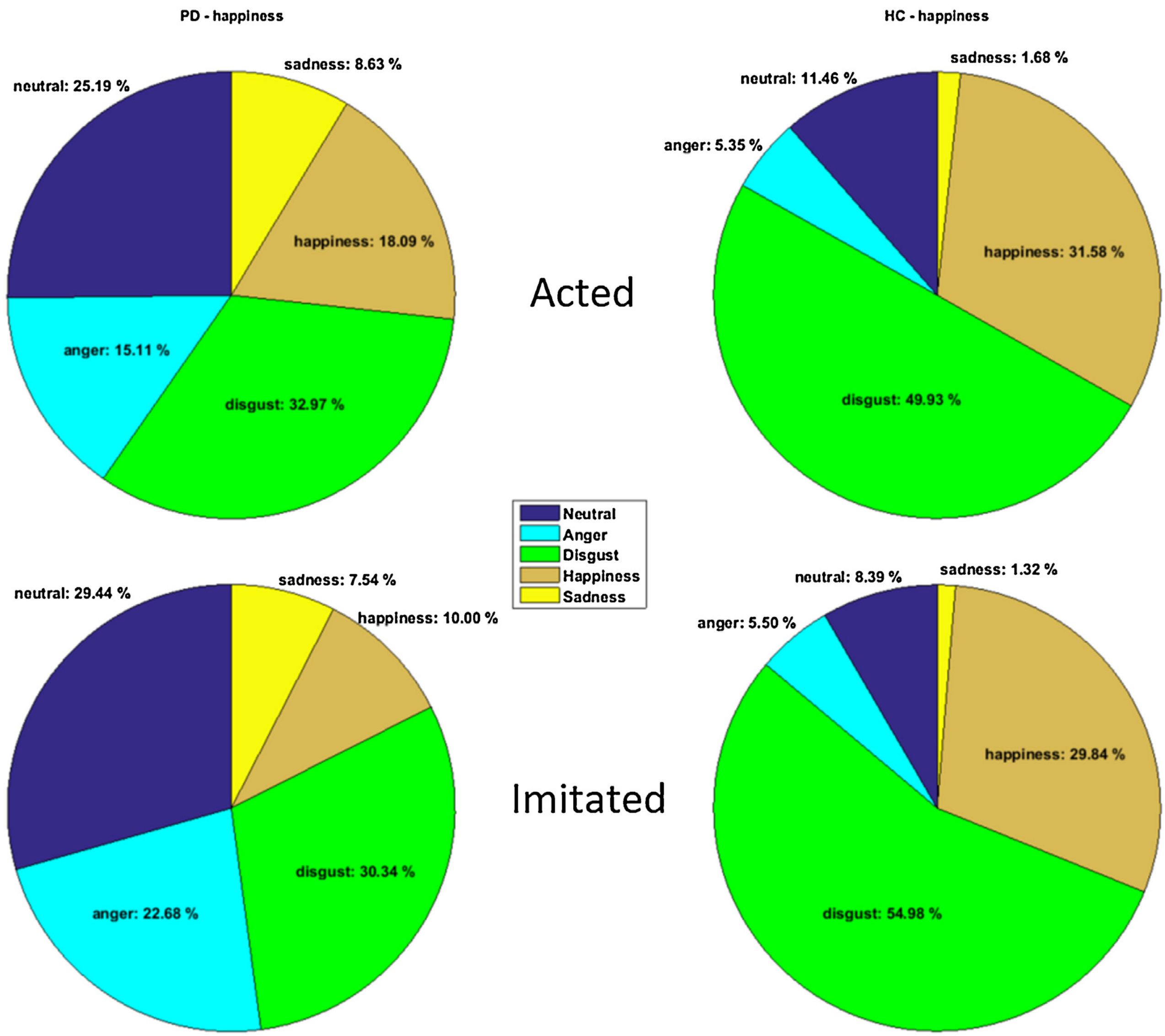

Fig. 5. Pie charts for happiness: acted expression (upper plots) and imitated expressions (lower plots). The two plots on the left concern PD patients, while those on the right concern HC subjects.

4. Discussion and conclusion

Considering PD patients, the percentage of target expression (that represents the frequency of displayed expression that occurred in a video) did not show major changes between acted and imitated tasks (Figs. 3–6). The percentage of neutral expression in Parkinsonian patients was always higher than HC subjects, during both tasks. Moreover, PD patients never showed an expression clearly prevalent over the others, as would be expected especially during the imitation. This suggests that PD patients had poor ability to show acted expressions both upon request and after the display of a visual aid for the imitation.

In contrast, HC subjects showed higher variations (between acted and imitated tasks) of the target expression than PD patients especially for anger and disgust. Moreover, during the imitation of anger and disgust (Fig. 3–4) the target expression greatly prevailed over the other expressions. This suggests that HC subjects were able to show a certain expression unambiguously upon request, but they clearly improved this ability when the visual aid for the imitation is shown. However, this result was not found during happiness and sadness (Figs. 5–6). Concerning happiness, the target expression remained stable between acted and imitated tasks and it was not the most frequent displayed expression (i.e. the highest expression in terms of percentage). In fact, a high percentage of disgust was visible during acted and imitated tasks (Fig. 5). Watching the video clips, we noticed that many HC subjects exaggerated this expression showing an excessive nose wrinkler, characteristic of disgust. Unlike PD patients, HC subjects did not show high percentages of neutral expression (Fig. 5); this means that they tried to exhibit happiness, but sometimes they failed in performing the right expression.

Sadness was the expression with the worst results in terms of target expression percentage. Both groups showed very low percentage of the target expression. PD patients always showed high percentages of neutral, disgust and anger. These three expressions fill almost the entire pie plot in Fig. 6. HC subjects showed similar results but with a clear reduction of the neutral expression. Thus, as in case of happiness, HC subjects tried to perform the expression but failed to do the right facial movements. In fact, after the experiment, the opinion of both groups was that it is very difficult to express sadness even when there is an image for the imitation.

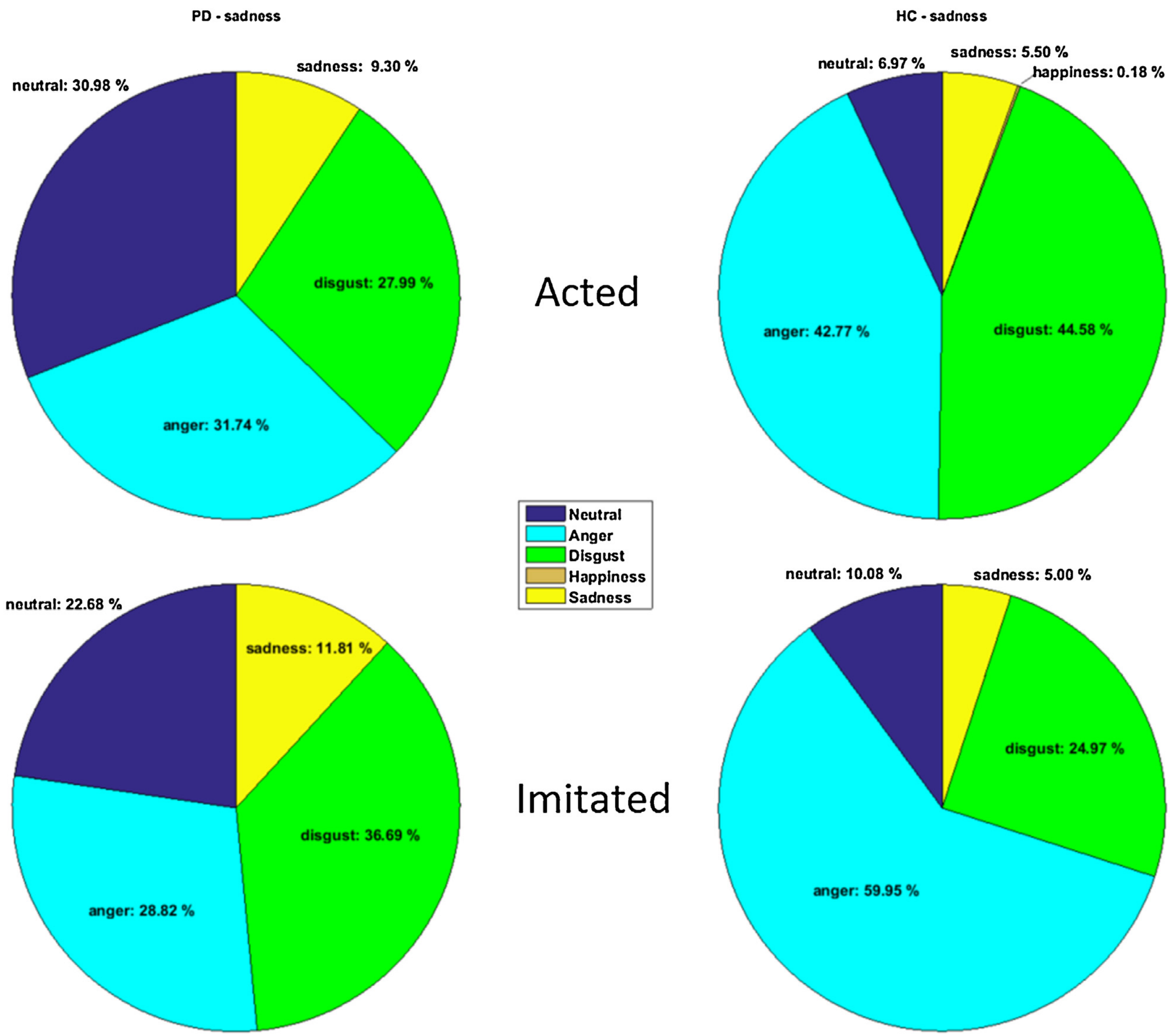

Fig. 6. Pie charts for sadness: acted expression (upper plots) and imitated expressions (lower plots). The two plots on the left concern PD patients, while those on the right concern HC subjects.

In this task both groups tend to assume different expressions, in particular anger and disgust.

The fact that HC subjects unsuccessfully tried to perform happiness and sadness is proven by the results in Tabs. 6 and 7. In fact, PD patients showed lower maximum and range values of distance during the acted happiness, lower maximum and minimum values of distance during the imitated happiness, lower mean value of distance during the acted sadness and lower mean, maximum and minimum values of distance during the imitated sadness.

Considering the Euclidean distance from the neutral baseline, PD patients never showed significant differences between acted and imitated expressions (Tables 4–7). This confirms the results of pie plots in Figs. 3–6, that is, the performances of PD patients in exhibiting a certain expression remain stable even after a visual aid has been shown for the imitation. In contrast, HC subjects showed an increase in mean distance between acted and imitated expressions, for both anger and disgust. This confirms that a higher expressivity is present during the imitation for HC subjects. Other significant variations were found in the minimum value of distance for all the expressions, although the decrease of the range value of distance during happiness indicates that HC subjects did not show big variations (between acted and imitated tasks) of the overall expressivity during this expression.

Considering only the acted expressions we can state that PD patients always showed a higher percentage of neutral than HC subjects (Figs. 3–6). Thus, the neutral expression was displayed more frequently in Parkinsonian patients. Distance results showed that no significant differences exist in the overall expressivity between HC subjects and PD patients during anger (Table 4), although the target expression is higher in HC subjects (Fig. 3). During acted disgust and acted sadness, PD patients showed a reduction of the mean distance with respect to HC subjects. Thus, the overall expressivity is reduced for PD patients. PD patients showed a lower expressivity also during the acted happiness, as suggested by a significant reduction of maximum and range values of distance.

The higher percentages of neutral expression in PD patients were confirmed also during the imitated task. The differences between PD patients and HC subjects were greatly enhanced during the imitation, since a significant reduction was detected in: mean distance during anger, disgust and sadness; maximum and minimum distance during anger, happiness and sadness. Thus, the reduction of the overall expressivity of PD patients is more enhanced during the imitated than the acted expressions.

These results showed that anger and disgust are the two expressions in which HC subjects showed a higher increase in the target expression during the imitation task. This means that HC subjects benefited from the displaying of the target expression in performing the imitation of those facial expressions. This was not true for PD patients, where the percentage of the target expressions (anger and disgust) remained stable or decreased from the acted task to the imitated task.

The Euclidean distance of the facial model from the neutral baseline was computed in order to quantify the amount of facial movements (and thus the overall expressivity) of both groups. On average, HC subjects reported higher distances than PD patients along the whole tasks; this confirms that HC subjects showed larger movements during both posed and imitated facial expressions.

From Tables 4–7 we noticed that PD patients did not show any significant difference between acted and imitated expressions in all the 4 tasks. This confirms what pie plots in Figs. 3–6 suggested, namely PD patients did not improve their facial mimicry even when they were asked to imitate another expression. The reduced Euclidean distance from the neutral baseline means that PD patients showed a global reduction of expressivity during the tasks. These results agree with other findings reported in literature (Simons et al., 2004; Bowers et al., 2006; Wu et al., 2014; Ricciardi et al., 2015). In particular, Simons et al. (2004) found a reduction of posed smiles and difficulties in imitating happiness, surprise and disgust in PD patients (facial movements from participants were coded by a certified FACS coder). The authors found that Parkinsonian patients had difficulties in exhibiting the following AUs during posed expressions: AU4 (brow lowerer), AU9 (nose wrinkler), AU10 (upper lip raiser), AU1 + 2 (brow raiser) and AU6 + 12 (cheek raise plus lip corner pull). In particular, AU4 is involved in sadness, fear and anger, AU9 is involved in disgust and AU6 + 12 are characteristic of happiness. In our case, we did not detect AUs, but the reduced percentage of disgust and anger in PD patients (Figs. 3–4), could be due to the impairment in exhibiting these movements: brow lowerer and nose wrinkler.

In all the recordings PD patients exhibited higher percentages of neutral with respect to HC subjects. This confirms the presence of a hypomimia that reduced the displaying of other facial expressions. In fact, even when HC subjects did not succeed in reaching the target expression (happiness – Fig. 5, sadness – Fig. 6), the neutral percentage was always lower if compared to PD patients.

As reported in Ricciardi et al., 2015 (Ricciardi et al., 2015), it is still unclear whether PD patients are unable to show posed expressions, since conflicting results are present in literature. Our results suggest that PD patients have difficulties in exhibiting posed expressions, if compared with HC subjects. However, the most evident result is that PD patients did not improve the ability to reach a requested expression even when a visual aid is proposed (in this case the imitation of an expression); in contrast, HC subjects showed this improvement, especially for anger and disgust.

As previously mentioned in the introduction, a still debated issue concerns the nature of the hypomimia in PD, with two main currents of thought: pure motor disorders or consequence of emotion recognition deficit. We believe that future works may address this issue by using such contactless methods in order to extract objective measures of facial mimicry in PD patients.

In this study we used a neutral baseline for each participant in order to quantify facial mimicry with respect to a reference that was specific for each subject, taking into account that each person may show facial expressions in different ways. The choice of a personal baseline is further corroborated by the fact that facial masking in PD may often create an impression of disinterest, anger, depression, etc. (Tickle-Degnen and Doyle Lyons, 2004). Thus, concerning the computation of the distance from the baseline, we were interested in how facial expressions (upon request or by imitation) were distant from a predefined reference.

Concerning the classification test, the algorithm for facial expression recognition was trained on databases of healthy subjects (CK+ and Radboud databases (Langner et al., 2010; Lucey et al., 2017)). These databases are extensively used in the literature of automatic facial expression recognition. We recognize that this might not be the best choice, since PD patients are characterized by alterations of facial mimicry. However, to our knowledge no databases of facial expressions from PD patients are currently available for research purposes. Future works should be addressed to the publication of this kind of databases, not only as far as PD is concerned, but also in other neuro-psychiatric disorders that induce alterations in facial expressions.

Limitations of this study are the heterogeneous PD sample and the exclusion of the expressions of fear and surprise. Patient and control samples should be more gender-balanced. However, PD is a neurodegenerative disease with higher prevalence in male population (Pringsheim et al., 2014) and this is reflected in the patients recruited in our study. Moreover, since most of the control subjects were the partners of the patients, it follows that the control group was composed by a high number of female.

Testing patients only in the "on-state" may be another limitation of the present study. However, as our video-based system was able to detect differences in facial expressions between PD patients and HC subjects during this phase, it is likely that these differences are more pronounced during the "off-state". This is reasonable, especially considering the beneficial effects of the dopaminergic therapy on hypomimia (Bologna et al., 2013).

More accurate results could be provided analyzing patients at different stages of the disease, in order to study facial hypomimia along the course of the disease. More information could be obtained by studying AUs, as already done in (Simons et al., 2004; Wu et al., 2014). In particular, the automatic detection of AUs could provide a meaningful aid to PD patients in carrying out therapeutic exercises that involve facial mimicry, with the aim of developing a tool for speech therapy and rehabilitation focusing mainly on those movements that seem to be more impaired in PD patients (as already demonstrated by (Simons et al., 2003)).

Another point that could be elucidated in future works could be the relationship of these methods with socio-cultural aspects. Actually, another still open debate in literature (in particular in psychology) concerns the cultural specificity in the expression and perception of emotions. Some authors (Jack et al., 2012) demonstrated that cultural differences may affect the performance of facial expressions recognition, because emotions are expressed somewhat differently across cultures. Nevertheless, there is evidence that among humans this might not be relevant (Soto and Levenson, 2009), so it is still difficult to draw firm conclusions about the expression of emotions across cultures and on the impact on facial expression recognition algorithms. However, considering that all PD patients and HC subjects recruited in this study constituted a culturally homogeneous group, the influence of cultural issue on the automatic expression recognition can be considered negligible. Moreover, this issue could be further investigated in the case of PD patients to check whether the socio-cultural differences may affect the disease and how.

A clear advantage of this study is the objectification of facial hypomimia in PD patients by means of a contactless technology. Wu et al., 2014 (Wu et al., 2014) already performed an analysis of facial mimicry in PD patients through AUs classification techniques, but differently from our work the authors recorded a group of PD patients monitored with contact sensors such as facial EMG and ECG. The automatic analysis of facial expressions is a continuously evolving field of research that finds several applications in medicine (Hamm et al., 2011). In particular, we believe that the system proposed here can be used to extract objective measures of facial mimicry representing an important means for analyzing facial expressions and movements also in a rehabilitation framework (in particular for speech therapy), whereby patients may obtain definite advantages about a real-time feedback of the right facial expressions/movements to perform.

Acknowledgements

This work was partially carried on under Project PGR00202 "Analysis and classification of voice and facial expression: application to neurological disorders in neonates and adults", Italian Ministry of Foreign Affairs – Progetti Grande Rilevanza Italy-Mexico.

References

http://www.faceshift.com. (Accessed 14 March 2016).
http://clopinet.com/isabelle/Projects/modelselect/MFAQ.html. (Accessed 13 March 2016).
http://scikit-learn.org/stable/supervised_learning.html. (Accessed 13 March 2016).
Bandini, A., Ouni, S., Cosi, P., et al., 2015a. Accuracy of a markerless acquisition technique for studying speech articulators. In: Proc. of INTERSPEECH, Dresden, Germany, pp. 2162–2166.
Bandini, A., Ouni, S., Orlandi, S., et al., 2015b. Evaluating a markerless method for studying articulatory movements: application to a syllable repetition task. In: Proc. of MAVEBA, Firenze, Italy, pp. 99–102.
Bandini, A., Orlandi, S., Giovannelli, F., et al., 2016. Markerless analysis of articulatory movements in patients with Parkinson's disease. J. Voice 30 (6) (766. e1-766. e11).
Bartlett, M., Littlewort, G., Frank, M., Lainscsek, C., Fasel, I., Movellan, J., 2006. Automatic recognition of facial actions in spontaneous expressions. J Multimed. 1 (6), 22–35.
V. Bettadapura, Face expression recognition and analysis: the state of the art. arXiv preprint arXiv:1203.6722 2012.
Bologna, M., Fabbrini, G., Marsili, L., et al., 2013. Facial bradykinesia. J. Neurol. Neurosur. Psychiatry 84, 681–685.
Bowers, D., Miller, K., Bosch, W., et al., 2006. Faces of emotion in Parkinsons disease: micro-expressivity and bradykinesia during voluntary facial expressions. J. Int. Neuropsych. Soc. 12, 765–773.
Cootes, T.F., Edwards, G.J., Taylor, C.J., 2001. Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. 23 (6), 681–685.
Cristinacce, D., Cootes, T., 2017. Feature detection and tracking with constrained local models. Sep 4–6, 2006; Edinburgh, UK In: Proceedings of the British Machine Vision Conference (BMVC), vol. 3, pp. 929–938.
Ekman, P., Friesen, W.V., 1977. Manual for the Facial Action Coding System. Consulting Psychologists Press, Palo Alto, CA.
Fahn, S., Elton, R., 1987. Members of the UPDRS development committee. In: Fahn, S., Marsden, C.D., Calne, D.B., Goldstein, M. (Eds.), Recent Development in Parkinson's Disease, vol. 2. Macmillan Health Care Information, Florham Park, NJ (153–163, 293–304).
Goetz, C.G., Poewe, W., Rascol, O., et al., 2004. Movement disorder society task force report on the Hoehn and Yahr staging scale: status and recommendations. Movement Disord. 19 (9), 1020–1028.
Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S., 2010. Mult. Image Vis. Comput. 28 (5), 807–813.
Hamm, J., Kohler, C.G., Gur, R.C., Verma, R., 2011. Automated facial action coding system for dynamic analysis of facial expressions in neuropsychiatric disorders. J. Neurosci. Meth. 200, 237–256.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning, 2nd ed. Springer, New York, NY.
Hemmesch, A.R., Tickle-Degnen, L., Zebrowitz, L.A., 2009. The influence of facial masking and sex on older adults' impressions of individuals with Parkinson's disease. Psychol. Aging 24 (3), 542–549.
Jabon, M., Bailenson, J., Pontikakis, E., Takayama, L., Nass, C., 2010. Facial expression analysis for predicting unsafe driving behavior. IEEE Pervasive Comput. 10 (4), 84–95.
Jack, R.E., Garrod, O.G.B., Yu, H., Caldara, R., Schyns, P.G., 2012. Facial expressions of emotion are not culturally universal. Proc. Natl. Acad. Sci. U. S. A. 109 (19), 7241–7244.
Jacobs, D.H., Shuren, J., Bowers, D., Heilman, K.M., 1994. Emotional facial imagery, perception, and expression in Parkinson's disease. Neurology 45, 1696–1702.
Jeni, L.A., Girard, J.M., Cohn, J.F., De La Torre, F., 2017. Continuous au intensity estimation using localized, sparse facial feature space. In: Proceedings of the 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Apr 22–26, 2013; Shanghai, China.
Kanade, T., Cohn, J.F., Tian, Y., 2017. Comprehensive database for facial expression analysis. In: Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition, Mar 28–30, 2000; Grenoble, France.
Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D.H.J., Hawk, S.T., van Knippenberg, A., 2010. Presentation and validation of the radboud faces database. Cognit. Emot. 24 (8), 1377–1388.
Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I., 2017. The extended Cohn-Kanade dataset (ck +): A complete dataset for action unit and emotion-specified expression. In: Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Jun 13–18, 2010; San Francisco, CA, USA.
Madeley, P., Ellis, A.W., Mindham, R.H.S., 1995. Facial expressions and Parkinson's disease. Behav. Neurol. 8, 115–119.
Matthews, I., Baker, S., 2004. Active appearance models revisited. Int. J. Comput. Vis. 60 (2), 135–164.
McKeown, G., Valstar, M., Cowie, R., Pantic, M., Schroder, M., 2012. The SEMAINE database: annotated multimodal records of emotionally colored conversations between a person and a limited agent. IEEE Trans. Affect Comput. 3 (1), 5–17.
Pantic, M., Patras, I., 2005. Detecting facial actions and their temporal segments in nearly frontal-view face image sequences. In: Proceedings of the 2005 IEEE International Conference on Systems, Man and Cybernetics (SMC), Oct 10–12, Waikoloa, HI, USA.
Pantic, M., Valstar, M., Rademaker, R., Maat, L., 2017. Web-based database for facial expression analysis. In: Proceedings of the 2005 IEEE International Conference on Multimedia and Expo (ICME 2005), Jul 6–8, 2005; Amsterdam, The Netherlands.
Patras, I., Pantic, M., 2017. Particle filtering with factorized likelihoods for tracking facial features. In: Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, May 19, 2004; Seoul, South Korea.
Pringsheim, T., Jette, N., Frolkis, A., Steeves, T.D.L., 2014. The prevalence of Parkinson's disease: a systematic review and meta-analysis. Movement Disord. 29 (13), 1538–1590.
Ricciardi, L., Bologna, M., Morgante, F., et al., 2015. Reduced facial expressiveness in Parkinson's disease: a pure motor disorder? J. Neurol. Sci. 358 (1-2), 125–130.
Sariyanidi, E., Gunes, H., Cavallaro, A., 2015. Automatic analysis of facial affect: a survey of registration, representation, and recognition. IEEE Trans. Pattern Anal. Mach. Intell. 37 (6), 1113–1133.
Scherer, K.R., Clark-Polner, E., Mortillaro, M., 2011. In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. Int. J. Psychol. 46 (6), 401–435.
Simons, G., Ellgring, H., Smith Pasqualini, M.C., 2003. Disturbance of spontaneous and posed facial expressions in Parkinson's disease. Cognit. Emot. 17, 759–778.
Simons, G., Smith Pasqualini, M.C., Reddy, V., Wood, J., 2004. Emotional and nonemotional facial expressions in people with Parkinson's disease. J. Int. Neuropsych. Soc. 10, 521–535.
Soleymany, M., Lichtenauer, J., Pun, T., Pantic, M., 2012. A multimodal database for affect recognition and implicit tagging. IEEE Trans. Affect Comput. 3 (1), 1–14.
Soto, J.A., Levenson, R.W., 2009. Emotion recognition across cultures: the influence of ethnicity on empathic accuracy and physiological linkage. Emotion 9 (6), 874–884.
Tian, Y.-L., Kanade, T., Cohn, J., 2001. Recognizing action units for facial expression analysis. IEEE Trans. Pattern Anal. Mach. Intell. 23 (2), 1–19.
Tickle-Degnen, L., Doyle Lyons, K., 2004. Practitioners' impressions of patients with Parkinson's disease: the social ecology of the expressive mask. So Sci Med. 58, 603–614.
Tjiaden, K., 2008. Speech and swallowing in Parkinson's disease. Top Geriatr. Rehabil. 24, 115–126.
Valstar, M.F., Mehu, M., Jiang, B., Pantic, M., Scherer, K., 2012. Meta-analysis of the first facial expression recognition challenge. IEEE Trans. Syst. Man Cybern. 42 (4), 966–979.
Vinokurov, N., Weinshall, D., Arkadir, D., Bergman, H., Linetsky, E., 2015. Quantifying hypomimia in parkinson patients using a depth camera. In: the 5th EAI International Symposium on Pervasive Computing Paradigms for Mental Health, Sep 24–25, Milan, Italy.
Weiss, D., Wächter, T., Breit, S., et al., 2010. Involuntary eyelid closure after STN-DBS: evidence for different pathophysiological entities. J. Neurol. Neurosurg. Psychiatry 81, 1002–1007.
Wu, P., Gonzalez, I., Patsis, G., et al., 2014. Objectifying facial expressivity assessment of Parkinson's patients: preliminary study. Comput. Math. Methods Med., 1–12.
Xiong, X., De la Torre, F., 2017. Supervised descent method and its applications to face alignment. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 23–28 2013; Portland, OR, USA, pp. 532–539.
Zeng, Z., Pantic, M., Roisman, G.I., Huang, T.S., 2009. A survey of affect recognition methods: audio, visual and spontaneous expressions. IEEE Trans. Pattern Anal. Mach. Intell. 31 (1), 39–58.
Zhao, G., Pietikäinen, M., 2009. Boosted multi-resolution spatiotemporal descriptors for facial expression recognition. Pattern Recogn. Lett. 30 (12), 1117–1127.