

The PARK Framework for Automated Analysis of Parkinson's Disease Characteristics

RAINA LANGEVIN, MOHAMMAD RAFAYET ALI, and TAYLAN SEN, University of Rochester, USA
CHRISTOPHER SNYDER, TAYLOR MYERS, and E. RAY DORSEY, University of Rochester Medical Center, USA
MOHAMMED (EHSAN) HOQUE, University of Rochester, USA

There are about 900,000 people with Parkinson's disease (PD) in the United States. Even though there are benefits of early treatment, unfortunately, over 40% of individuals with PD over 65 years old do not see a neurologist. It is often very difficult for these individuals to get to a physician's office for diagnosis and subsequent monitoring. To address this problem, we present PARK, Parkinson's Analysis with Remote Kinetic-tasks. PARK instructs and guides users through six motor tasks and one audio task selected from the standardized MDS-UPDRS rating scale and records their performance via webcam. An initial experiment was conducted with 127 participants with PD and 127 age-matched controls, in which a total of 1,778 video recordings were collected. 90.6% of the PD participants agreed that PARK was easy to use, and 93.7% mentioned that they would use the system in the future. We explored objective differences between those with and without PD. A novel motion feature based on the Fast Fourier Transform (FFT) of optical flow in a region of interest was designed to quantify these differences in the collected video recordings. Additionally, we found that facial action unit AU4 (brow lowerer) was expressed significantly more often, while AU12 (lip corner puller) was expressed less often in various tasks for participants with PD.

CCS Concepts: • **Human-centered computing** → *Ubiquitous and mobile computing*;

Additional Key Words and Phrases: Parkinson's disease, Remote, Accessible, MDS-UPDRS

ACM Reference Format:

Raina Langevin, Mohammad Rafayet Ali, Taylan Sen, Christopher Snyder, Taylor Myers, E. Ray Dorsey, and Mohammed (Ehsan) Hoque. 2019. The PARK Framework for Automated Analysis of Parkinson's Disease Characteristics. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 2, Article 54 (June 2019), 22 pages. <https://doi.org/10.1145/3328925>

1 INTRODUCTION

Over 10 million people in the world are living with Parkinson's disease (PD) [32], a disease which causes over 29,000 deaths per year [17] and affects around 900,000 people [42] in the United States alone. Shockingly, even in developed nations like the United States, an estimated 40% of people aged 65 or older living with PD do not receive care from a neurologist [60]. A large problem in both Western and Asian countries is that many individuals, around 15% or higher, living with PD have not been diagnosed [44]. Even when diagnosed, PD requires a continuous process of trying and adjusting medications [57]. A simple online PD symptom screening

Authors' addresses: Raina Langevin, rlangevi@u.rochester.edu; Mohammad Rafayet Ali, mali7@cs.rochester.edu; Taylan Sen, tsen@cs.rochester.edu, University of Rochester, Computer Science, Rochester, NY, 14260, USA; Christopher Snyder, chris.snyder@chert.rochester.edu; Taylor Myers, taylor.myers@chert.rochester.edu; E. Ray Dorsey, ray.dorsey@chert.rochester.edu, University of Rochester Medical Center, Department of Neurology, Rochester, NY, 14260, USA; Mohammed (Ehsan) Hoque, mehoque@cs.rochester.edu, University of Rochester, Computer Science, Rochester, NY, 14260, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, or post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2019 Association for Computing Machinery.
2474-9567/2019/6-ART54 \$15.00
<https://doi.org/10.1145/3328925>

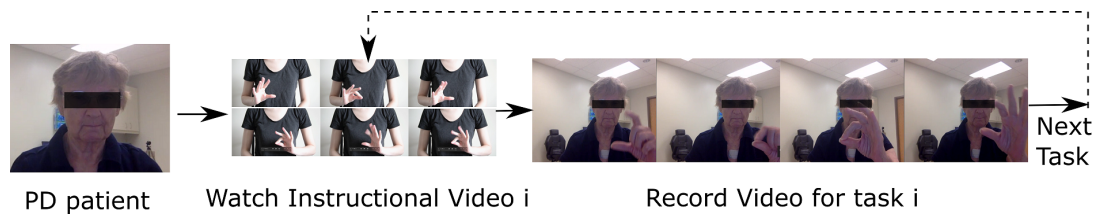


Fig. 1. Users watch a video for each of the tasks. The PArk framework then records the video of the user locally. After uploading the recorded video to the server, PArk takes the user to the next task.

and assessment application has the potential to identify thousands of undiagnosed people living with PD and additionally provide individuals being treated with an unobtrusive and objective method for assessment.

However, for several reasons, the development of such a system is not a trivial task. There is currently no diagnostic test for detecting PD. The symptoms of PD persist over a long period of time and tend to slowly worsen over several years, making it difficult to identify [31]. These symptoms include upper-limb tremor, rigidity, bradykinesia, and others. Diagnosis is primarily done through observation and is limited by the potential subjectivity of the individual conducting the screening. The most commonly used PD drug, levodopa [30], typically loses its responsiveness over time, requiring continual monitoring of symptoms and adjustment of the medication [57]. Travel to a medical facility may be particularly burdensome to individuals with PD whose mobility is significantly compromised. Additionally, it may be hard for neurologists, Parkinson's nurses or other trained individuals to compare a patient's status when only meeting with them once a month or quarter [23]. Further, evaluating individuals with PD in a clinical environment may yield assessments that are significantly different than their real-world situation. With 89% of adults in the United States using the Internet, either with home broadband access or smartphone use [18], an online system has the potential to reach a large population for screening.

This paper discusses the web application framework, titled Parkinson's Analysis with Remote Kinetic-tasks (PAReK), which collects videos of participants performing vocal and motor exercises and extracts a novel motion feature for the objective assessment of PD symptoms. PAReK is designed for ease of use by individuals with PD. The development of the system was informed by expert panels and visits to the clinic where participants with PD tested PAReK and provided feedback. The tasks were selected from the Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS), which is the current standard for assessing PD [29]. An initial experiment involving a total of 127 users with PD and 127 age-matched users was conducted. Each participant performed one assessment session that was composed of seven video-recorded tasks taking about 15 minutes (total of 1,778 recordings). To differentiate the users with and without PD we designed a motion metric feature from the frequency spectrum of the optical flow of the region of interest in the video frames. The feature identifies the relative speed of the motor exercise and further characterizes the participants with PD with significance ($p < 0.05$). Additionally, facial action units were extracted using the off-the-shelf feature extraction tool Openface [7]. These features revealed that average eye brow-lowering muscle contraction was higher, while maximum smile intensity was lower among participants with PD.

In this paper, we make the following contributions:

- Development of an online framework which explores the feasibility of the remote assessment of PD using a standard webcam and microphone.
- Design and validation of a novel video analysis algorithm to measure a motion metric which identifies differences between participants with and without PD.

Individuals at varying stages in their disease progression were recruited, and the majority were in treatment. The goal of this work is to detect differences between the two groups in the selected measurements, and to test the usability of the system.

2 BACKGROUND

2.1 Traditional Diagnosis of Parkinson's Disease

In the health care field, trained individuals commonly determine the severity of PD using the Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS), a scale with 65 items answered on a five-point scale and a total range of 0-260 points [33]. This rating scale was developed as a revision of the original Unified Parkinson's Disease Rating Scale [29]. In the motor examination of the MDS-UPDRS, trained individuals are instructed to observe a series of physical exercises and "rate what they see" based on a set of evaluation criteria. For each exercise, they assign an integer score to the observed motor movement from 0-4 (normal - severe).

Applications that provide objective monitoring of PD are an important solution to the limitations of in-person ratings. There is a need for more sensitive motor assessments that decrease the inter-rater variability between trained individuals. For those with PD in the United States, medication costs them, on average, \$2,500 a year [32]. Without precise feedback, individuals with PD could be spending money on medication that might not work for them. In countries where medical costs are not paid out of pocket, travel and time costs are still a factor, and the medical expenses remain a burden for government health care. The subjective rating of symptoms not only has financial implications, but it can cause over-medication which leads to negative side effects. Furthermore, the majority of seniors do not own a smartphone, while half of seniors have high-speed internet at home, [3] and it is likely that a web application could serve the needs of this population group. An online framework is a feasible way to collect videos and provide assessments [20].

2.2 Computer-Aided Systems and Methods

Past work has identified novel approaches to differentiate between healthy individuals and those with PD. Audio recordings of phonations have been used to distinguish between these two groups [39]. Individuals living with PD have been found to have speech that is softer, less distinct, and has less rhythm, stress and intonation [41]. Several studies have investigated the characteristics in facial expressions of those with PD. The low intensity of facial expressions associated with PD has been described as a "mask"-like expression [53] and has been found in both posed and spontaneous expressions [8, 35, 61]. This lack of facial expressivity is consistent with the neurological damage of PD [8]. Additionally, specific characteristics of PD facial abnormalities include a wider palpebral fissure (i.e. eyelids wide open), a reduced blink rate, non-distinct nasolabial folds, and an open mouth [13].

An experimental practice in the diagnosis of PD is the use of smartphone applications and wearable technologies, such as ParkNosis [48] and PERFORM [54]. These systems provide the convenience of being able to receive feedback at home without visiting the doctor's office. Self-monitoring health technologies, such as smartphone applications, have been successful in differentiating between individuals with PD and those from an age-matched control group [5, 6], determining and validating severity scores [46, 62] and detecting responses to medication using quantified measures [63]. Various methods have been proposed for gathering data from remote interfaces [27]. In one smartphone application, participants performed five tasks that assessed voice, posture, gait, finger tapping and response time [6]. The capabilities of wearable sensors have been leveraged for PD diagnosis [16], and often studies have employed accelerometers to identify and measure tremors. For example, the use of gyroscopes to detect and objectively quantify tremors and bradykinesia was validated on a set of individuals with and without PD [49]. Outside the use of sensors, [55, 56] analyzed the rhythmic movements of body parts from video data using the time-domain and frequency domain spectrum, producing results comparable to accelerometer

generated data. It has been found to be feasible to engage a large number of individuals with PD using remote monitoring [15, 46, 63].

The “digital divide” serves as a geographic and social barrier that limits the access of older groups, and groups with more chronic conditions, to telemedicine [21, 22]. The majority of seniors do not own a smartphone [11]. The use of a smartphone or wristband may require visits to the hospital where they are given and equipped with these technologies. Additionally, seniors, over the age of 65 and living in the United States, who have less education and a lower household income are much less likely to use technology [3].

Even when access is provided, prior research has identified that older individuals still have difficulty with online technology [10] and can benefit from customized interfaces [2]. In terms of providing care via video-conferencing, instead of smartphone-based collection, it has been found to be feasible. But there can be issues with the sound and quality of videos if there is a low connection speed [1]. Despite this, individuals with PD may prefer virtual visits to in-person appointments [9]. Video-conferencing with individuals with PD in their homes was found to improve convenience, comfort and saved time [9]. To our knowledge, there is no direct comparison of the experience of individuals with PD with smartphones versus video-calling applications, like Skype. But past work has indicated that web-based applications are a viable and well-received solution [20]. This is a relatively new field and it could be beneficial to explore multiple solutions, both web-based and smartphone-based.

3 PARK SYSTEM

In this section, we describe the development process for the system. First, we discuss the tasks we selected from the motor exercises in Part III of the MDS-UPDRS. We then describe the online interface we built and the design considerations. Lastly, we describe a set of survey questions we used to gather information from participants about their experience using the system.

3.1 Task Selection

The MDS-UPDRS combines a quantitative and qualitative approach to assessing PD. Part I and II examine the non-motor and motor experiences of daily living, Part III contains the motor examination, and Part IV examines motor complications. From the motor assessment in Part III of the MDS-UPDRS, we have selected six tasks for the PARK system. These six exercises are related to the severity of PD motor symptoms, such as upper-limb tremor, rigidity, and bradykinesia. Despite the subjectivity of the MDS-UPDRS, [43] determined severity levels of PD based on MDS-UPDRS scores from all the sections.

In order to analyze facial expressions, while both talking and not talking, two separate tasks (Task 1 and Task 2) were introduced. In the MDS-UPDRS, speech is rated by observing unstructured conversation. In contrast, we chose to have participants read a specific sentence to standardize the task for future analysis. Task 7, the vocal task, is not from the MDS-UPDRS, but it has been used in well-known PD research [38]. The following motor tasks were chosen considering the feasibility of self-recording with a webcam.

- (1) Task 1: Resting Facial Expression
 - (a) Requirements: Participant is recorded sitting still for 10 seconds.
 - (b) Evaluation criteria: Observe eye-blink frequency, masked faces (stiff facial expression), and parting of lips.
- (2) Task 2: Speech
 - (a) Requirements: Participant is recorded while reciting the sentence: *“the quick brown fox jumps over the lazy dog. The dog wakes up and follows the fox into the forest, but again the quick brown fox jumps over the lazy dog.”*
 - (b) Evaluation criteria: Observe volume, modulation, clarity, slurring, repetition of syllables, and rapid speech.

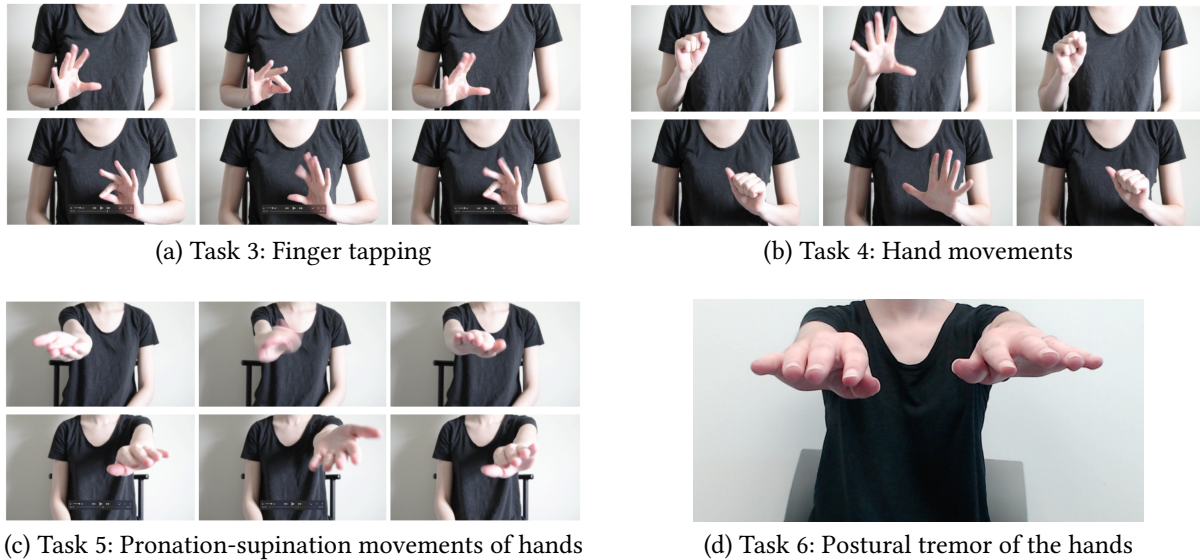


Fig. 2. Screenshots of selected motor exercises from the MDS-UPDRS.

(3) Task 3: Finger Tapping

- (a) Requirements: Participant taps index finger and thumb together 10 times, done for both hands.
- (b) Evaluation criteria: Observe speed, amplitude, decrementing amplitude, hesitations, and halts. Depicted in Figure 2a.

(4) Task 4: Hand Movements

- (a) Requirements: Participant keeps arm bent and makes a tight fist, then opens the hand 10 times as fully and as quickly as possible, done for both hands.
- (b) Evaluation criteria: Observe speed, amplitude, decrementing amplitude, hesitations, and halts. Depicted in Figure 2b.

(5) Task 5: Pronation-supination Movements of Hands

- (a) Requirements: Participant extends an arm out in front of their body with their palm down, then they turn the palm up and down 10 times as fast and as fully as possible, done for both arms.
- (b) Evaluation criteria: Observe speed, amplitude, decrementing amplitude, hesitations, and halts. Depicted in Figure 2c.

(6) Task 6: Postural Tremor of the Hands

- (a) Requirements: Participant extends both arms out in front of their body with their palms down, for 10 seconds.
- (b) Evaluation criteria: Observe the highest tremor amplitude for each hand. Depicted in Figure 2d.

(7) Task 7: Voice task

- (a) Requirements: Participant is recorded saying “ahh” for as long as possible.
- (b) Evaluation criteria: Observe vocal tremors and breathiness. This is a sustained phonation, when a single vowel is produced and the pitch is held for as long as possible.

3.2 User Interface Design

PARK is designed for people with PD and those at risk of the disease. In designing PARK, an expert panel of four neurologists at the University of Rochester Medical Center was consulted. One neurologist walked through the prototype at different stages of development in several meetings. During open discussions, the neurologist noted what was working well and potential improvements. Additional feedback on the selection and modification of the tasks was gathered remotely via email and in the clinic from the remaining neurologists.

In the initial stages of the system's development, we gained important insights from two pairs of participants with PD and their caregivers during clinic visits. This work was done in collaboration with the neurologists and a research coordinator. During two clinic visits, both the person with PD and their caregiver were given a short introduction about the project after their appointment concluded. They were then observed as they navigated through the system. When they had a question or were stuck, we stepped in to explain. After they finished the survey, we openly discussed their impressions of the system. Their responses were recorded in writing immediately following the visit. It was observed that one of the participants had trouble with their eyesight and difficulty concentrating. Though the instructional video said to stop at 10 taps in Task 3, they had to be told verbally when to stop. Another participant mentioned that the video instructions were not necessary because they were familiar with the tasks, though they acknowledged that the videos helped them focus and they might not have paid attention otherwise. The conclusion from this feedback was to continue using instructional videos and to make each step of the process more visible and clear. Following the clinic visits, we developed a set of guidelines for the design of PARK.

As those diagnosed with PD tend to be part of an older demographic, we reviewed relevant literature and found that our guidelines were in line with past work on design for older adults. In a study of 125 websites offering health resources, [10] evaluated the ability of the homepage designs to meet the needs of older users based on guidelines set by the National Institute on Aging and the National Library of Medicine [45]. As humans age, they experience a decline in vision, cognition and physical skills and [10] proposed that all these should be addressed by web design. Many of their suggestions are in line with the established W3C guidelines [4] regarding text, color and contrast, and multimedia captions. The W3C guidelines additionally supplemented that it is important to limit distractions, time restraints and mouse use, and instead promote keyboard use.

Building on this knowledge, our objective was to design an accessible and convenient interface for people with PD and older adults. Based on our discussions with neurologists and literature review the following guidelines were determined.

- The tasks should not require help from another person or caregiver.
- The tasks should not require the users to stand up or walk.
- The interface should give clear instructions on how to complete the tasks.
- The interface should include both audio and visual instructions.
- The interface should allow the users to easily navigate with few clicks.
- The interface should allow the users to retake their video.

Using these guidelines, sans serif fonts and large buttons were selected, subtitles were added to the videos, and black text on an off-white background was used for visual clarity and simplicity. We used Arial font for the main content of the website, as well as Questrial and Helvetica Neue fonts for page titles, navigation and footer content. The font sizes were at least 16px, which is the equivalent of 12pt. For clear cognitive navigation, a contact email was listed at the bottom of each page to provide assistance. There were only two pages listed in the navigation bar: the homepage and the about page. The about page served as the help feature and contained contact information, a privacy statement, details about the study, and a link to download Google Chrome.

Our goal was to have a simple and clear interface where a user would be taken step by step through the voice and motor exercises. Following each instruction page, the user is required to watch a video of an individual

completing the exercise. Once the user has finished watching the example video, they can begin recording themselves (see Figure 1). To reduce difficulties that may be caused by motor impairment, users were able to progress through the pages either with a mouse click or the space bar. Once the user has completed all of the exercises, they are taken to a demographic survey and feedback form with Google's Speech Recognition API enabled. The web application is hosted on a remote server and utilizes the RecordRTC JavaScript library [36] for uploading submitted videos. At this point in time, participants do not receive feedback on their videos. Figure 3 shows a vision of a future PARK use case with a flowchart of the system.

3.3 Survey

We administered a survey at the end of the video collection in the PARK system. The survey consisted of demographic questions, as well as ten feedback questions. For the feedback survey questions, participants responded on a five-level Likert scale (i.e., 1 = "strongly disagree," 2 = "disagree," 3 = "neither," 4 = "agree," and 5 = "strongly agree").

Table 1. Feedback survey questions.

1. The web application was easy to use.
2. I understood all the instructions for performing the tasks.
3. I would like to use the system in the future.
4. I had no problem browsing through the system.
5. The videos shown in the interface were helpful.
6. I knew how to perform the tests beforehand.

Table 2. Feedback discussion questions.

1. What did you like most about the web application?
2. What did you like least about the web application?
3. What did you like the most about this study?
4. What suggestions would you make to improve this study?

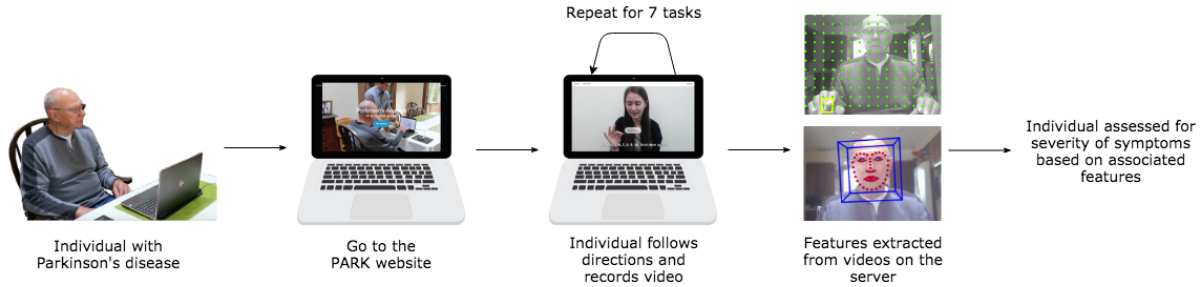


Fig. 3. Vision of PARK System flowchart.

4 AUTOMATED ASSESSMENT ENGINE

An automated assessment engine was developed to collect video data and to extract features potentially corresponding to motor symptoms of PD. We based the development of our assessment methods off of past researchers' success in demonstrating differences in facial expressions and motion features.

4.1 Facial Features

Through the use of automated facial expression extraction, we sought to quantify the differences in facial expression for each of the tasks between individuals diagnosed with PD and the control group. The Facial Action Coding System is a well established system of objectively representing the magnitude of facial contractions for a

Table 3. OpenFace Facial Action Coding System (FACS) features.

| AU | Description | Expected Observation | AU | Description | Expected Observation |
|------|---------------------|-------------------------------------------|------|----------------------|-------------------------------|
| AU1 | Inner brow raiser | decrease | AU14 | Dimpler | less spontaneous smiling |
| AU2 | Outer brow raiser | decrease | AU15 | Lip corner depressor | - |
| AU4 | Brow lowerer | increase | AU17 | Chin raiser | - |
| AU5 | Upper lid raiser | - | AU20 | Lip stretcher | - |
| AU6 | Cheek raiser | decrease | AU23 | Lip stretcher | lip tightener |
| AU7 | Upper lid tightener | - | AU25 | Lips part | lips parted greater |
| AU9 | Nose wrinkler | decrease | AU26 | Jaw drop | - |
| AU10 | Upper lip raiser | decrease; less prominent nasolabial folds | AU28 | Lip suck | - |
| AU12 | Lip corner puller | less spontaneous smiling | AU45 | Blink | decreased eye-blink frequency |

number of dimensions called facial action units (“AUs”) [25]. Each facial action unit roughly corresponds to a major facial muscle group. We used the OpenFace tool version 2 [7] for automatically extracting the commonly used facial action units shown in Table 3. Openface has been shown to have good AU level correlations when run on a set of expert labeled public facial expression databases, with an average correlation of 0.73 over all AUs [7]. We configured OpenFace to run on our PARK server and extract AU features from the webcam videos as they are uploaded.

As discussed in the Background section and summarized in Table 3 the “Expected Observation” column, prior research has identified several ways in which facial expressions in individuals with PD deviate from healthy individuals. Facial features were analyzed in every task since we conjectured that the facial features would be affected differently by the given tasks. Each of the tasks evaluates a different type of manifestation of motion disorder. The physical exertion levels evaluated include the following: at rest, static exertion (hold still), and dynamic exertion for fine movement, (index finger-thumb), medium-scale movement (hand), and gross motor movement (arm). Facial expressions were analyzed in each of these scenarios since our common experience found that individuals often make a non-neutral facial expression when they concentrate on a task.

4.2 Motion Features

In addition to facial expression analysis, we sought to measure differences in hand motion between the group with PD and the control group in Tasks 3-6. Due to the limited video quality of the data, it was not possible to track individual finger movements reliably. More specifically, some of the videos involved individuals who moved their hands into and out of the camera view. Even though the video was configured to capture high quality frames, due to limited recording bandwidth on participant machines or poor lighting, the video surrounding moving fingers was often blurred. Participants were asked to perform 10 cycles per hand in each task, however, in many cases slightly more or less cycles were performed. In summary, the algorithm we developed was based on actual data gathered in the wild in contrast to idealistic laboratory conditions. We focused on the development of a motion feature with the following characteristics:

- The metric should not highly depend on the number of motion cycles performed (i.e., the number of finger taps).
- The metric should not drastically suffer if the individual’s hand temporarily moved off the screen.

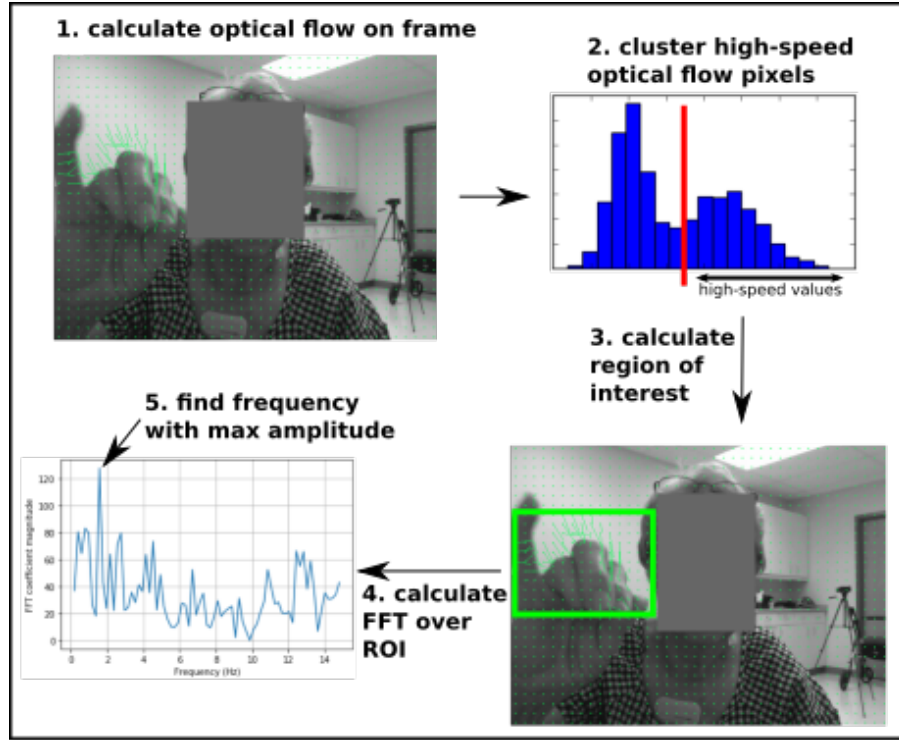


Fig. 4. Algorithmic steps for calculation of motion magnitude metric.

- The metric should not deteriorate with slight to medium differences in hand orientation or distance relative to the webcam.
- The metric should not be corrupted from nonperiodic background movement.
- The metric should produce discernible differences between the group with PD and the control group.

After experimentation, we developed a methodology to estimate the maximal motion frequency in a region of interest. Our method combines the computer vision method known as optical flow with motion frequency analysis. To this end, we implemented the Farnebäck algorithm in OpenCV version 3.1.0 in C++. Optical flow involves calculating the frame to frame direction and magnitude of movement. In performing this calculation, the video is divided into a grid of separate regions. The direction and magnitude of movement is then calculated separately for each region for successive frames [28]. In addition to optical flow, we used Fourier analysis to determine the frequency of extracted motions in the participant videos [52]. The steps for calculating the maximal motion frequency (“motion metric”) for each participant video are detailed in Algorithm 1 and shown in the first three steps of Fig. 4.

Starting with our video files, which were of a fixed resolution (640×480), we applied the steps in Fig. 4 to yield a single scalar motion metric value for each video file. As shown in step 1, optical flow is calculated for each grid point (i.e., green dots in Fig. 4-1) between two successive video frames. In summary, the x and y motion at each grid point is estimated between the previous and current frames. In step 2, the motion at each grid point is clustered into a low and high speed group using k-means clustering. In step 3, the region of interest is calculated as the minimum sized rectangular bounding box that contains all of the grid points in the high speed cluster. The

Algorithm 1: Process video to return the motion metric, a .**Result:** File containing average motion magnitude per frame-----
Definitions $frame$ = matrix of video pixels F = optical flow vector matrix for $I \times J$ regions M = magnitude scalar matrix for $I \times J$ regions L = cluster assignments of $I \times J$ elements in M by k-means = 0,1 B = bounding box of region of interest A = array of motion per frame, in the region of interest

START

 $t = 0$; $last\ frame$ = capture next video frame; $frame$ = capture next video frame;**repeat** F = calculate Farnebäck Optical Flow ($last\ frame$, $frame$); initialize M sizeof F ; **forall** $f_{i,j} \in F$ **do** $m_{i,j} = \sqrt{f_{i,j}.x^2 + f_{i,j}.y^2}$; **end** L = k-means (M , $k=2$), where,

cluster 0 = lower magnitude;

cluster 1 = larger magnitude;

 B = minimum rectangular subset of $[I \times J]$ which includes all of cluster 1 elements; $A[t] = \frac{\sum_{m_{i,j} \in B} m_{i,j}}{|B|}$; $last\ frame = frame$; $t = t + 1$;**until** no frame exists to read; $a = \operatorname{argmax}(FFT(A))$;-----
OUTPUTmotion metric, a .

average motion of all the grid points in the bounding box is calculated, yielding a single value for the current frame. This process is repeated for all the frames to produce a time series signal of average speeds in the bounding box of each frame. In step 4, the Fourier transform is taken of the speed vs. time signal. In step 5, the frequency of motion which has the maximum energy is identified, and is returned as the motion metric for the given video file.

The variables used in Algorithm 1 are defined as follows. F is the optical flow matrix that is unique to each frame with $I \times J$ grid points. Each grid point $f_{i,j}$ has an associated x and y value representing horizontal and vertical components of the estimated motion. M is a matrix of scalar magnitudes (i.e. speed) of each grid point with identical dimensions to F . L is a matrix of the same dimensions as F that has values of 0 and 1, representing the grid points' cluster assignments. B represents the minimum bounding box of grid points in the high magnitude

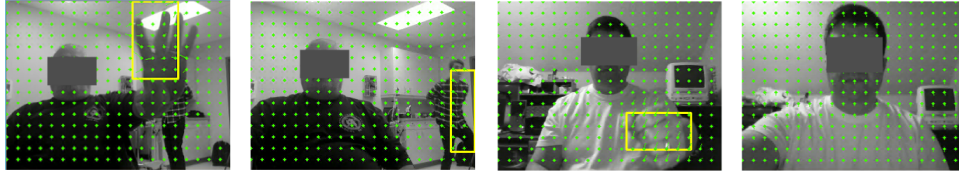


Fig. 5. Participants with background motion. On the left, the participant's hand is detected during the task, and after a person is detected in the background. On the right, the participant's hand is detected, and after they adjust their webcam with no detection.

cluster. A is the motion magnitude time series that represents each frame's estimated motion magnitude in the region of interest.

Algorithm 1 starts with the time $t = 0$, *lastframe* capturing the first frame, and *frame* capturing the second video frame. The optical flow between *lastframe* and *frame* is calculated and loaded into F . For each coordinate (i.e. grid point i,j), we calculate the optical flow vector magnitude and load it into M . These values represent the estimated speed of motion for each grid point. All of the grid points are then clustered, according to their magnitudes in M , into either a low magnitude cluster or a high magnitude cluster. The cluster assignments are calculated using the k-means algorithm and stored in Boolean matrix L . The boundary that encompasses all the regions in the high speed optical flow cluster group is calculated to define our region of interest as bounding box B . B includes all elements in the high magnitude cluster and likely some elements in the low magnitude cluster. The average magnitude over all elements in bounding box B is then calculated and stored in $A[t]$. This process is repeated for each subsequent frame.

In calculating the boundary box, k-means clustering was used as a more robust method than a static speed threshold. More specifically, a static threshold would not be able to as robustly handle situations in which a participant's hand changes in distance from the camera. The k-means algorithm sets the speed threshold dynamically by automatically adjusting for different participant positioning.

The boundary box normally included the regions that we were interested in, namely the hands and arms. But, in some cases (e.g., if someone was adjusting their webcam), the whole body would be included. Additionally, if there was background motion, it would sometimes be included in the boundary box as well. This occurred when the motion was significant and not distant from the webcam. In Figure 5, the boundary box surrounds the hand motion as the task is performed. In the case of the first participant with PD, on the left, a person is walking in the background and is included once they begin to quickly move off screen. On the right, the second participant with PD adjusts their webcam after the task, but it is done slowly and the motion is not included.

After Algorithm 1 has completed and returns A , the time series of average motion magnitudes inside the region of interest, the Fourier transform is taken. This transformation was then normalized by the length of the transformed values in order to put all data points into a consistent frequency domain. From the frequency domain signal, the frequency with maximum amplitude is returned as the motion metric for the given video. These steps were performed in order to obtain a metric focusing on dynamic movement characteristics in the region of maximum movement that is resilient to spurious (i.e., less energetic) movements.

5 ANALYSIS AND RESULTS

In this section we present the results from the survey and the extracted features. We present initial findings on facial and motion features that are expressed differently for the two participant groups in specific tasks.

Table 4. Baseline Characteristics.

| Characteristic | No. (%) | |
|------------------------------------------------------------------------|-----------------------------------------------|--------------------------------------------------|
| | Participants With Parkinson's Disease (n=127) | Participants Without Parkinson's Disease (n=127) |
| Age, mean (SD) | 62.4 (7.4) | 62.1 (7.2) |
| Women | 60 (47.2%) | 85 (66.9%) |
| White race | 117 (92.1%) | 106 (83.5%) |
| College graduate | 110 (86.6%) | 88 (69.3%) |
| Previously used video calling ^a | 96 (81.4%) | 79 (75.2%) |
| Previously used automated PD analysis | 51 (40.2%) | NA |
| Taking levodopa | 72 (56.7%) | NA |
| Taking sinemet | 18 (14.2%) | NA |
| Years since diagnosis ^b , mean (SD) | 6.9 (5.7) | NA |
| MDS-UPDRS score, mean (SD) Modified motor ^c (range 0-24) | 9.9 (3.2) | NA |

^a 96 out of 118 participants with PD, and 79 out of 105 participants without PD that answered this question.

^b Excluding one participant that did not correctly enter their year of diagnosis.

^c Including motor assessment of Part III: 3.1 Speech, 3.2 Facial Expression, 3.4 Finger Tapping, 3.5 Hand Movements, 3.6 Pronation-supination Movements of Hands, and 3.15 Postural Tremor of Hands for 20 randomized participants with PD.

5.1 Data Collection

There was a total of 127 participants with PD and 127 age-matched participants. Their characteristics are shown in Table 4. Participants were recruited in-person during scheduled clinic visits, at PD support group meetings, and online. The study had IRB approval and, if they selected to, participants could receive \$50 Amazon gift cards for their involvement.

We recruited three medical doctors with MDS-UPDRS certification to rate the videos. Two of them rated the videos of 10 randomized participants and the other rated a different set of 10 random participants to determine MDS-UPDRS scores. The certified raters conducted a motor assessment from Part III: 3.1 Speech, 3.2 Facial Expression, 3.4 Finger Tapping, 3.5 Hand Movements, 3.6 Pronation-supination Movements of Hands, and 3.15 Postural Tremor of Hands, of the MDS-UPDRS. Accordingly, each video was rated on an integer scale from 0 to 4. Since it would be time intensive to rate all the videos for 127 participants with PD, only 20 randomly selected participants with PD were chosen for their videos to be rated. We did not evaluate the Hoehn & Yahr scores [34] because measuring balance impairment in-person is integral to the rating, and this was not feasible in our study.

5.2 Facial Feature Results

We compared an age-matched control group (n=127) to a group of participants with PD (n=127) for our analysis of facial features over all the video frames for each task. Three participants were left out of the facial feature analysis due to low confidence scores in OpenFace for all of their video frames (i.e., using an OpenFace confidence level of 75%). Analysis of our data revealed two differences between the groups with and without PD. Shown in Table 5 are the differences in AU expression frequencies between the groups among the tasks sorted by increasing

Table 5. Observed differences between the groups with and without PD.

| AU | task # | PD mean | Control mean | std. dev. | p-value | corrected p-value (x224) |
|-----------|--------|---------|--------------|-----------|----------|--------------------------|
| AU 4 avg | task 1 | 1.23 | 0.841 | 0.782 | 5.51 e-5 | 0.01 |
| AU 4 avg | task 5 | 1.108 | 0.842 | 0.731 | 0.005 | > 1 |
| AU 4 avg | task 3 | 1.177 | 0.954 | 0.789 | 0.008 | > 1 |
| AU 4 avg | task 4 | 1.147 | 0.905 | 0.763 | 0.009 | > 1 |
| AU 4 avg | task 2 | 1.220 | 0.993 | 0.788 | 0.020 | > 1 |
| AU 4 avg | task 6 | 1.078 | 0.857 | 0.772 | 0.023 | > 1 |
| AU 4 avg | task 7 | 1.270 | 1.117 | 0.844 | 0.163 | > 1 |
| AU 12 max | task 5 | 0.646 | 0.935 | 0.640 | 2.29 e-4 | 0.05 |
| AU 12 max | task 7 | 0.456 | 0.813 | 0.762 | 0.001 | > 1 |
| AU 12 max | task 6 | 0.439 | 0.646 | 0.595 | 0.003 | > 1 |
| AU 12 max | task 4 | 0.527 | 0.806 | 0.687 | 0.003 | > 1 |
| AU 12 max | task 1 | 0.351 | 0.525 | 0.600 | 0.016 | > 1 |
| AU 12 max | task 3 | 0.421 | 0.607 | 0.628 | 0.036 | > 1 |
| AU 12 max | task 2 | 0.535 | 0.686 | 0.617 | 0.037 | > 1 |

Mann-Whitney U-test p-value. In the “corrected p-value” column are the p-values that have the most aggressive Bonferroni multiplier applied to consider the effects of multiple hypothesis testing. We used a correction factor of 224 ($(\# \text{ AUs}) \times (\# \text{ tasks}) \times (\# \text{ metrics}) = 17 \times 7 \times 2$). As shown in Table 5, the average AU4 (“brow lowerer” muscles, corrugator supercilii and depressor supercilii) shows a substantial difference between the groups, with the group with PD demonstrating significantly greater average expression of AU4. This difference was most pronounced in Task 1 (Resting Facial Expression). The Mann-Whitney U-test p-value is 0.012 after Bonferroni correction. While Tasks 2-6 also show greater AU4 averages, their level of difference does not meet significance at 95% when using Bonferroni correction.

In Table 5, the group with PD had lower maximum AU12 (lip corner puller) on average during Task 5 (Pronation-supination Movements of Hands). This effect is also shown, but to a lesser degree, in the other tasks. None of the other facial expression action units listed in Table 3 showed statistically significant differences (at 95% confidence) between the group with PD and control group.

5.3 Motion Feature Results

The Tasks 3, 4, 5, and 6 are related to hand movements. All 127 participants with PD and 127 age-matched control participants were included in this analysis. Figure 6 compares differences in the motion metric between the groups with and without PD. All tasks, except for Task 6 (Postural Tremor of the Hands), show significant differences.

In the other three tasks, the group with PD had a mean frequency that is lower than the control group ($p < 0.05$). All the following values represent the *mean(std)* of people with PD vs. the *mean(std)* of the controls: Task 3 0.11(0.13) vs. 0.13(0.14), Task 4 0.12(0.13) vs. 0.16(0.15), and Task 5 0.10(0.11) vs. 0.14(0.15). This may indicate that the participants with PD have relatively slower hand movements than the participants without PD. It is possible that Task 6 (0.06(0.09) vs. 0.06(0.10)) is not significant because the hands tend to be too close to the camera and fill the entire screen, or go outside of the frame. One neurologist who rated the videos said that some of the exercises, specifically noting Task 6, were tough to see on screen because of these tendencies. Additionally, Task

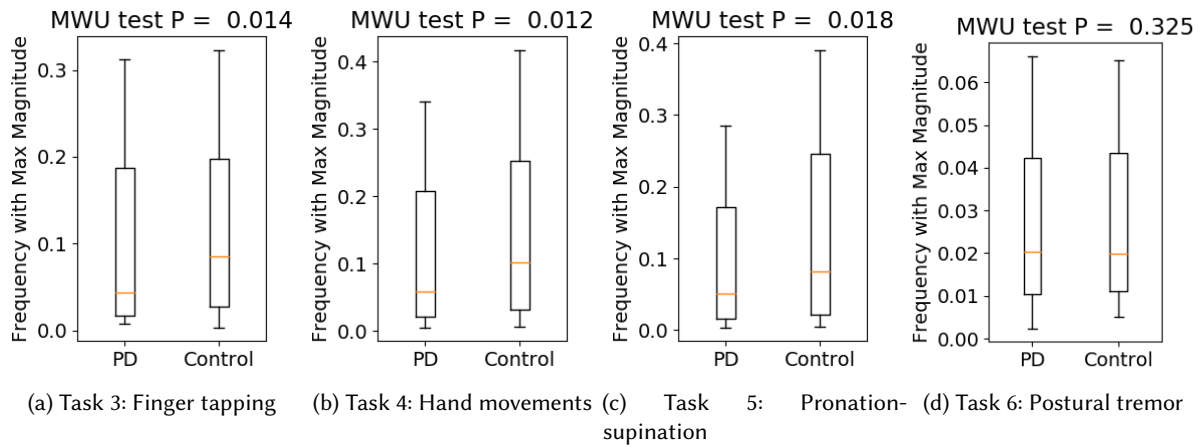


Fig. 6. Boxplot of the frequencies with maximum magnitude

Table 6. Pearson correlation coefficients of task average MDS-UPDRS rating and motion metric.

| | Task 3 | Task 4 | Task 5 | Task 6 |
|------------------|--------|--------|--------|--------|
| MDS-UPDRS Rating | 0.313 | 0.070 | 0.235 | -0.235 |

6 is different than the other three hand movement tasks since it requires the participants to hold their hands still. This fundamental difference in the nature of the tasks may have been reflected in the motion metric.

In past work, [59] were able to generate scores for hypomimia that were correlated to UPDRS ratings of PD symptoms. Similarly, to validate our motion feature, we calculated the Pearson correlation coefficients for the motion metrics and neurologist given MDS-UPDRS scores for Tasks 3, 4, 5, and 6. We have selected the videos of 20 random participants for each of the tasks and had three MDS-UPDRS trained medical doctors assign scores. The videos of the first 10 participants were rated by two doctors and the average score was used for the correlation analysis. The remaining 10 participants were rated by one doctor. For Task 6, only 19 participants were analyzed as one participant was not rated because they performed the task incorrectly. The coefficients are shown in Table 6. While the motion metrics from Tasks 3, 4, and 5 have a positive correlation with the MDS-UPDRS scores, there is a negative correlation with Task 6, in which individuals are holding their arms stationary. As a reference, the correlation between the ratings of the two doctors was 0.75. Additionally, the average Cohen's Kappa values between the two doctors was 0.37 (Task 1 = 0.125; Task 2 = 0.0741; Task 3 = 0.375; Task 4 = 0.697; Task 5 = 0.1667; Task 6 = 0.8039). It should be noted that the correlation and Kappa measures were calculated on a small subset of the dataset (i.e., only 10 videos) and thus might not generalize well.

As a validation of the resilience of the motion metric, we created 30 test videos of a single individual performing Task 3 (Finger Tapping) with various simulated "quality issues." More specifically, five test videos were recorded for each of the six scenarios outlined by our characteristics to evaluate Algorithm 1's performance. Namely, this includes Test 1 (baseline video with 10 taps per hand), Test 2 (five taps), Test 3 (fifteen taps), Test 4 (each hand moves off screen for one second), Test 5 (increased distance of 2 ft. from the webcam), and Test 6 (presence of person walking in background). The variance of the motion metrics in Table 7 is $8.4e-3$ which provides a measurement of the resilience of the motion metric in various unideal scenarios.

Table 7. Motion metric of the finger tapping task on different scenarios.

| Test | Motion Metrics | | | | |
|--------------------|----------------|--------|--------|--------|--------|
| | Test 1 | Test 2 | Test 3 | Test 4 | Test 5 |
| Baseline | 0.109 | 0.031 | 0.033 | 0.048 | 0.034 |
| Five taps | 0.024 | 0.244 | 0.049 | 0.045 | 0.045 |
| Fifteen taps | 0.024 | 0.029 | 0.028 | 0.014 | 0.238 |
| Hand off screen | 0.061 | 0.031 | 0.076 | 0.074 | 0.031 |
| Increased distance | 0.037 | 0.053 | 0.018 | 0.018 | 0.018 |
| Noisy background | 0.188 | 0.065 | 0.032 | 0.452 | 0.034 |

5.4 Survey Results

In their survey ratings in Figure 7, 90.6% of the participants with PD agreed that PARK was easy to use, and 93.7% agreed that they would use the system in the future. 93.7% found the use of instructional videos helpful and 94.5% agreed that they understood all the instructions. To this end, many of the participants with PD were already familiar with the MDS-UPDRS exercises (82.7%). For those who were not familiar with the tasks (22 participants with PD), 16 of them agreed that it was easy to use and that they understood the PARK instructions. There were 16 of participants with PD (12.6%) that did run into some sort of problem when using the system. For example, they had problems with set up (e.g., they had to change their browser from Internet Explorer to Chrome) (n=4), the interface itself (e.g., the survey did not highlight which specific question they missed, or they could not read text on the screen)(n=2), camera adjustment (n=3), difficulty typing (n=1), and slow internet connection (n=1). There were some who agreed that they ran into problems, but they did not provide a written response (n=5).

For the written discussion responses, the majority of participants with PD opted to give more detailed feedback. Written responses were obtained from 112 of the 127 participants with PD. A large percentage, 50 out of the 112, mentioned that they liked the system's "ease of use." In responses that did not directly discuss ease of use, other topics included the enjoyment of using the technology (n=6), the "ability to do it at home" and by themselves (n=21), and the desire to use the system more often to track their symptoms (n=6).

For the thematic analysis, the 112 written responses of participants with PD were used as content. We approached our thematic analysis by looking at the explicit content and the assumed reality of the data set. We determined 15 labels that were used to code the responses: personal assistance, unfamiliar technology, enjoyable technology, helping research, desire for options, privacy, physical difficulty, home use, self-paced, insecurity, easy, quick, curiosity, understandable, interface problems. Based on this coded dataset, we defined six themes:

- (1) Technological issues and limitations: In the prototyping stage, topics about improvements to the web framework were discussed after participants completed the study. In this initial stage, a participant might need help orienting themselves in the recording frame, or with moving closer and farther from the camera depending on the task at hand. One participant made a suggestion to limit scrolling. A technological limitation with prototyping was that participants were using our laptops, which in most cases were Macs, and were unfamiliar with how to navigate. For example, one participant mentioned that they disliked "using a mac" and another wanted help with "how to understand and use the computer technology if unfamiliar e.g. apple." This issue became less of a concern outside of prototyping because participants were using their personal computers. But it is important to note that participants still ran into trouble with various issues when we deployed the study online (n=13). For example, some participants had difficulty when enabling their microphone or when their default browser did not support video recording capabilities.
- (2) Ease of use: The ease of use of PARK was the most common theme (n=50). Participants found that it was easy and convenient because the instructional videos were straightforward and short, and the standardized

tasks were familiar. As one participant noted, it was “quick and to the point. videos were not too long.” The text instructions were aided by visual examples of the tasks. It is also likely that many wrote that it was “easy” because they could do it in any environment and did not have to travel to meet with a trained professional. Also, it was a only a short selection of tasks from the MDS-UPDRS.

- (3) Personal agency: There was a number of participants (n=26) who appreciated that they had more control over where and when they participated. They could record task videos at home, at their own pace, and did not have to travel to the doctor’s office. It appeared that participants were interested in additional personal agency within the framework as well. They were curious about the research direction and how the study would measure changes in symptoms. They looked forward to continuous use of the system. One participant said that, “I hope I can use some day to help evaluate how medicines are working. Instruction video optional if able to use on a regular basis.” They expressed a desire to skip over the video instructions if they were able to use the system more often.
- (4) Supportive environment: Conversely, for those who participated in the clinic or in a PD support group, the help and personal assistance was appreciated (n=2). One participant commented that the “people were nice to work with.” Participants were always positive about their experiences when support was present. Another participant said that the exercises were simple, but it took some getting used to. They completed the tasks at home and had physical difficulty with the system, saying that they “had to sit to[o] far away from computer so my arms got tired.” There is less emphasis on personal agency in cases like this one (n=4). Perhaps, as a future use, it could be beneficial to optionally introduce the PARK system in the clinic setting with the support of a trained administrator. After being familiarized, the user could then use it at home or another location following the visit.
- (5) Comfort on camera: A less prominent theme was how comfortable participants felt recording videos of themselves (n=4). Some felt comfortable and trusted the security of the data, and one liked that it was a private and non-invasive process. Another enjoyed using the technology and said, “It’s fun to watch myself though I didn’t get my hands in the video very well.” On the other hand, one participant said they disliked “seeing myself on camera,” and seemed to feel less confident and anxious about self-recording.
- (6) Internal incentives: When participants touched on what they liked about PARK, and the study’s strengths, they expressed inner motivations for their involvement (n=13). In one response, an individual said, “Happy to participate and move you forward. I like the concept of technology use,” and another said they liked “playing with the computer to provide feedback for research.” The enjoyment of using new technology and the desire to help were often tied together. Participants appreciated that the study was helping to further PD research because its purpose was aimed towards a cause important to them. Interestingly, those who showed internal incentives were likely to show more than one.

6 DISCUSSION

Why does the group with PD show lower maximum AU12 expression? This finding is consistent with several prior studies and serves to validate the overall utility of the PARK system. Generally, facial expressions of those with PD are marked by a decreased level of movement in the lower face [14, 50] with some exceptions including AU10 (“upper lip raiser”) [50]. [14] found decreased levels of expressivity in individuals with PD for several action units, including AUs 6, 9, 10, 12, and 15. Other observed behavior in individuals with PD includes less spontaneous smiling [51] and lips parted while the mouth is at rest. [47] found patients with PD were significantly less able to express posed happiness. Similarly, [8] found a significant reduction in mean facial expression distance between participants with and without PD during acted expressions of happiness. Smiling is often associated with AUs 6, 7, and 12. Therefore, we would expect to see lower values for these AUs. Part of the reason why the difference in AU12 is significant only in Task 5 could be due to the relative magnitude of physical exertion in each of the

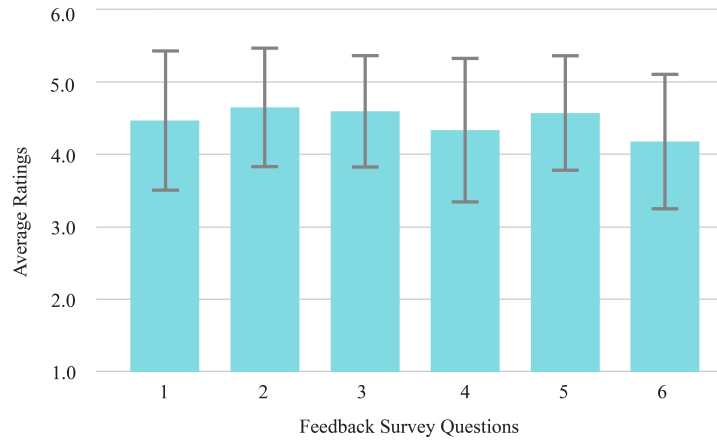


Fig. 7. Average ratings and standard deviation of the survey responses of participants with PD. The questions are available in Table 1.

tasks. Task 5 is the most physically demanding in that it requires the participant to move their arm (the other tasks involve only hand or finger movement). It is not uncommon for facial expressions to occur during exertion. Overall, expressivity levels are generally lower among individuals with PD, however, some AUs, such as AU4 and AU25, are more heavily expressed during resting activity [61].

Why does the group with PD show higher average levels of AU4 expression? The muscles associated with AU4 include the corrugator supercilii, which pulls the brow downwards and towards the center of the face, as well as the depressor supercilii, which pulls the brow directly downwards [24]. The expression of AU4 has been demonstrated to be strongly correlated with negative valence and inversely related to positive valence [37]. In this context, valence refers to composite self-reported positive and negative affective reactions. As PD is associated with distress, it is not surprising that we would find higher levels of AU4 expression in the group with PD during the resting task. Though significant differences in AUs 4 and 26 were not observed in [14], their study was concerned with posed facial expressions. While [14] asked participants to mimic the six basic emotions, and measured the expression of AU4 during anger and sadness, this work is concerned with the resting facial expressions of participants. It is important to note that our data showed that the difference in AU4 expression between the group with PD and the control group is strongest in Task 1 where participants remained at rest. However, in all the tasks, average AU4 expression level was higher in the group with PD. The reason the difference is significant only in Task 1 may be due to the increased variance in facial expression levels in other tasks. In all tasks other than Task 1, individuals are either moving or speaking. Perhaps the resting task is least affected by "noise" from physical exertion and movement. It is reasonable that different individuals tend to react differently during active tasks (i.e. there is only one way to truly stand still, but many ways to move). The sample variances directly affect the p-value calculation, driving the p-value higher when the variance within the groups is larger.

Why does the group with PD show a lower average motion metric level? We observed that participants with PD had a low average frequency in their motion features compared to their age-matched control. The lower motion metric level is consistent with the general decrease in muscle control associated with PD. This might be related to the freezing of gait [58], prolonged reaction time [12, 26], or other movement disorders [19] among the participants with PD. Our motion metric is calculated from the optical flow of the movements in the video frames. This might be a reason why we observe the lower motion metric among the PD participants. One of the

original intentions in the design of the motion metric was that the motion metric produces discernible differences between the two groups. From a statistical significance perspective, our results show that this objective was achieved on a group level. However, the variability of the motion metric is too high to enable the use of the motion metric independently as a diagnostic measure for individual evaluation of PD.

How robust is the motion metric in different conditions? We tested our motion metric in six different corner conditions. Those conditions were performed by the same person to verify if the motion metric is affected by the video conditions. In the baseline test, the variance of Tests 2-5 is $4.5e-5$, but for Tests 1-5, when Test 1 is included, it is $8.8e-4$. Test 1 has a value of 0.109 which is roughly three times larger than the other tests. It appears that in 20% of the cases untargeted motion is picked up by the algorithm. We acknowledge this as a limitation of the motion metric. In regards to the other scenarios, noisy background has the largest effect on the motion metric with a variance of $2.5e-2$ in this condition. While the variance of the motion metric is small enough to observe statistically significant differences between the group with PD and the control group, the variance is too large to apply the motion metric as a diagnostic tool for individuals. In summary, while the motion metric is resilient enough to show group-level differences, it is not robust enough to enable individual-level diagnosis of PD.

Overall, the low variance ($8.4e-3$) in the motion metric on different conditions indicates that the metric is dependent on the persons' movement characteristics and not the video conditions, excluding background motion. Two major steps of the motion metric calculations might explain this resilience. First, the motion metric is calculated based on a region of interest (i.e., where most movement is visible). Second, the motion metric is calculated in the frequency domain, which captures the prominent movement in the video frame. Though there can be movements in the head or other body parts, after performing the Fourier transform we were able to isolate these movements easily.

We found that the correlations between the MDS-UPDRS ratings and the motion metric were low (<0.4). It is possible that the motion metric is not holistically representative of PD motor symptoms. Since the frequency measurement captures only one aspect of the MDS-UPDRS rating, which is also based on amplitude, hesitations and halts, this might explain why the correlations were not as strong.

How was the PARK system received by participants? Based on this early study, participants were hopeful about the feedback the web application could provide in the future. They appreciated the personal agency it afforded them, yet they also valued the in-person relationships developed at the clinic. In regards to the user interface, they wanted more control to move through the pages. Some participants who were familiar with the motor exercises wanted to skip the videos and the instructions. However, we made it mandatory to watch the videos keeping in mind that people with PD show cognitive impairment [40]. Because there might be users with reduced visual and cognitive abilities, video instructions were preferable to stand-alone written instructions. This method was intended to make it easier for users to follow along and feel that using PARK is similar to a doctor's visit where someone is demonstrating the exercises. In the future, participants may not be using a personal computer or they may be sharing a monitor with other participants in a group screening. To limit confusion, it is important that an individual is there to assist them and to administer the study because they might be unfamiliar with the operating system. To further follow the guidelines outlined for older adults, we found that there could be improvements in time performance, for users on very slow Internet connections, and translations for non-English speaking users [10].

6.1 Limitations

At this time, participants do not receive feedback on their symptoms, and this system has not been validated as an assessment tool for measuring symptoms of PD. Because the frequency measurement captures only one aspect of the MDS-UPDRS rating, it is important to note that the PARK system tasks do not address all PD motor symptoms. If a person with PD experiences halting in the motor tasks, this would not be recognized. PARK

would be useful for those affected specifically by upper-limb tremor, rigidity or bradykinesia. While statistically significant differences have been identified in the median feature values between the group with PD and the control group, due to the high variance in each of the relevant features, accurate individual predictions cannot be made.

A limitation with the video and audio recording capability of the RecordRTC JavaScript library is that it is only available for Firefox, Chrome, and Opera. If a user accessed PARK through another incompatible browser, they would be directed to a page with instructions on downloading Chrome. However the situation can be inconvenient and confusing for users. For example, one individual did not know how to access Chrome, and incorrectly tried to change their settings in Safari to have Google as their search engine as a solution. This is an issue that should be addressed in future work. We also acknowledge that if a participant has a slow network connection, they will experience a longer wait time for their videos to upload which can diminish usability for older participants.

It should be noted that we collected vowel phonations in this study, but we did not analyze this data for the paper. Additionally, a limitation of the task selection is that the development of Tasks 1 and 2 make the videos more difficult for trained individuals to rate.

6.2 Future Work

The data obtained by the PARK system identifies significant differences between the group with PD and the control group. These findings, consistent with past research, demonstrate that PARK is a viable system for remotely gathering data useful for detecting PD. It should be noted that while our participants did not take part in frequent re-evaluations with PARK over the course of treatment or a long duration, we have demonstrated that participants are capable of remotely accessing the PARK system and completing the online tasks without our help. It follows that an online framework could easily be used for frequent monitoring without the constraints of scheduling clinicians or requiring screened individuals to travel into a health care facility. Our tool is not validated either as a diagnosis nor assessment tool, and the features examined have been evaluated from a statistical rather than classification analysis. It should be clearly stated that the correlation of MDS-UPDRS scores with our motion metric is exploratory and was only done on a subset of the data, and is thus not yet ready for clinical trials.

In future work, we plan to more closely track disease progression through home monitoring. This could be possible by enabling participants to set up accounts to track their progress and receive an assessment score. Currently, the features have not been evaluated in relation to specific drug treatment, or on-off states, as this information has not been collected. But, eventually, we hope to monitor the severity of symptoms through feature extraction for this purpose. If symptoms can become quantifiable using the discussed features, this technology could be used to help clinicians find the optimal dose and schedule for symptom modifying treatments. With the ever rising costs of drug development, an objective PD assessment system may help improve the evaluation of new PD drugs by making high frequency online assessments of PD easy and cost effective to implement.

7 CONCLUSION

A web application that indicates the presence of Parkinson's disease (PD) and provides an analysis of symptoms could be of great benefit to individuals with PD. It is hard to determine the presence of PD in its beginning stages, which is why early detection is important. The PARK web application could be helpful to individuals, who for a number of reasons are unable to or do not visit the hospital or clinic, in order to alert them of PD symptoms. Based on our results, the use of videos, or potentially images, to determine the presence of a significant furrowed brow (AU4), the absence of smiling (AU12), and the frequencies of hand motion could be new methods for indicating symptoms of PD. For a definite diagnosis, it is recommended that a person visits a neurologist or another trained individual to receive a holistic diagnosis that takes into consideration their medical history. An online framework

that assesses features of PD could be introduced during a clinic visit to initially supplement the tool with personal support. Additionally, for undiagnosed individuals who are curious about the presence of PD symptoms, it could serve as an important starting point.

ACKNOWLEDGMENTS

This work was in part supported by National Science Foundation Award IIS-1750380, Microsoft Azure for Research grant, and the National Institute of Neurological Disorders And Stroke of the National Institutes of Health under Award Number P50NS108676. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

- [1] Amir Abdulahi, Michael T Bull, Kristin C Darwin, Venayak Venkataraman, Matthew J Grana, E Ray Dorsey, and Kevin M Biglan. A feasibility study of conducting the montreal cognitive assessment remotely in individuals with movement disorders. *Health Informatics Journal*, 22(2):304–311, 2016. doi: 10.1177/1460458214556373. PMID: 25391849.
- [2] Mohammad Rafayet Ali, Kimberly Van Orden, Kimberly Parkhurst, Shuyang Liu, Viet-Duy Nguyen, Paul Duberstein, and M Ehsan Hoque. Aging and engaging: A social conversational skills training program for older adults. In *23rd International Conference on Intelligent User Interfaces*, pages 55–66. ACM, 2018.
- [3] Monica Anderson and Andrew Perrin. Technology use among seniors, May 2017. URL <http://www.pewinternet.org/2017/05/17/technology-use-among-seniors/>.
- [4] Andrew Arch, Shadi Abou-Zahra, and W3C Web Accessibility Initiative (WAI). Developing websites for older people: How web content accessibility guidelines (wcag) 2.0 applies, January 2018. URL <https://www.w3.org/WAI/older-users/developing/>.
- [5] S. Arora, V. Venkataraman, A. Zhan, S. Donohue, K.M. Biglan, E.R. Dorsey, and M.A. Little. Detecting and monitoring the symptoms of parkinson’s disease using smartphones: A pilot study. *Parkinsonism & Related Disorders*, 21(6):650 – 653, 2015. ISSN 1353-8020. doi: <https://doi.org/10.1016/j.parkreldis.2015.02.026>. URL <http://www.sciencedirect.com/science/article/pii/S1353802015000814>.
- [6] S. Arora, V. Venkataraman, A. Zhan, S. Donohue, K.M. Biglan, E.R. Dorsey, and M.A. Little. Detecting and monitoring the symptoms of parkinson’s disease using smartphones: A pilot study. *Parkinsonism and Related Disorders*, 21(6):650–653, June 2015. doi: 10.1016/j.parkreldis.2015.02.026.
- [7] Tadas Baltrušaitis, Peter Robinson, and Louis-Philippe Morency. Openface: an open source facial behavior analysis toolkit. In *IEEE Winter Conf. Applicat. of Comput. Vision*, 2016. doi: 10.1109/WACV.2016.7477553. URL <http://ieeexplore.ieee.org/document/7477553/>.
- [8] Andrea Bandini, Silvia Orlandi, Hugo Jair Escalante, Fabio Giovannelli, Massimo Cincotta, Carlos A Reyes-Garcia, Paola Vanni, Gaetano Zaccara, and Claudia Manfredi. Analysis of facial expressions in parkinson’s disease through video-based automatic methods. *Journal of neuroscience methods*, 281:7–20, 2017.
- [9] CA Beck, DB Beran, KM Biglan, and et al. National randomized controlled trial of virtual house calls for parkinson disease. *Neurology*, 89(11):1152–1161, September 2017. doi: 10.1212/WNL.0000000000004357.
- [10] Shirley Ann Becker. A study of web usability for older adults seeking online health resources. *ACM Trans. Comput.-Hum. Interact.*, 11(4): 387–406, December 2004. ISSN 1073-0516. doi: 10.1145/1035575.1035578. URL <http://doi.acm.org/10.1145/1035575.1035578>.
- [11] A. Berenguer, J. Goncalves, S. Hosio, D. Ferreira, T. Anagnostopoulos, and V. Kostakos. Are smartphones ubiquitous?: An in-depth survey of smartphone adoption by seniors. *IEEE Consumer Electronics Magazine*, 6(1):104–110, Jan 2017. ISSN 2162-2248. doi: 10.1109/MCE.2016.2614524.
- [12] EL Berry, RI Nicolson, JK Foster, Marlene Behrmann, and HJ Sagar. Slowing of reaction time in parkinsons disease: the involvement of the frontal lobes. *Neuropsychologia*, 37(7):787–795, 1999.
- [13] Matteo Bologna, Giovanni Fabbrini, Luca Marsili, Giovanni Defazio, Philip D Thompson, and Alfredo Berardelli. Facial bradykinesia. *J Neurol Neurosurg Psychiatry*, 84(6):681–685, 2013.
- [14] Matteo Bologna, Isabella Berardelli, Giulia Paparella, Luca Marsili, Lucia Ricciardi, Giovanni Fabbrini, and Alfredo Berardelli. Altered kinematics of facial emotion expression and emotion recognition deficits are unrelated in parkinson’s disease. *Frontiers in neurology*, 7: 230, 2016. ISSN 1664-2295. doi: 10.3389/fneur.2016.00230. URL <http://europepmc.org/articles/PMC5155007>.
- [15] Brian M. Bot, Christine Suver, Elias Chaibub Neto, Michael Kellen, Arno Klein, Christopher Bare, Megan Doerr, Abhishek Pratap, John Wilbanks, E. Ray Dorsey, Stephen H. Friend, and Andrew D. Trister. The mpower study, parkinson disease mobile data collected using researchkit. *Scientific Data*, 3(160011), March 2016. doi: 10.1038/sdata.2016.11.
- [16] Abdul Haleem Butt, Erika Rovini, Dario Esposito, Giuseppe Rossi, Carlo Maremmanni, and Filippo Cavallo. Biomechanical parameter assessment for classification of parkinson’s disease on clinical scale. *International Journal of Distributed Sensor Networks*, 13(5): 1550147717707417, 2017. doi: 10.1177/1550147717707417.

- [17] CDC. National center for health statistics: Parkinson's disease mortality by state, 2016. URL https://www.cdc.gov/nchs/pressroom/sosmap/parkinsons_disease_mortality/parkinsons_disease.htm.
- [18] Pew Research Center. Internet/broadband fact sheet, February 2018. URL <http://www.pewinternet.org/fact-sheet/internet-broadband/>.
- [19] AR Cools, J Horstink van den Bercken, MW Horstink, KP Van Spaendonck, and HJ Berger. Cognitive and motor shifting aptitude disorder in parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 47(5):443–453, 1984.
- [20] E. Dorsey, V. Venkataraman, M.J. Grana, and et al. Randomized controlled clinical trial of "virtual house calls" for parkinson disease. *JAMA Neurology*, 70(5):565–570, 2013. doi: 10.1001/jamaneurol.2013.123.
- [21] E. Ray Dorsey and Eric J. Topol. State of telehealth. *New England Journal of Medicine*, 375(2):154–161, 2016. doi: 10.1056/NEJMra1601705. PMID: 27410924.
- [22] E. Ray Dorsey, Alistair M. Glidden, Melissa R. Holloway, Gretchen Lano Birbeck, and Lee Schwamm. Teleneurology and mobile technologies: the future of neurological care. *Nature Reviews Neurology*, 14:285–297, 2018.
- [23] E.R. Dorsey, S. Papapetropoulos, M. Xiong, and K. Kiebertz. The first frontier: Digital biomarkers for neurodegenerative disorders. *Digital Biomarkers*, 1(1):6–13, 2017.
- [24] Paul Ekman and Wallace V Friesen. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *semiotica*, 1(1):49–98, 1969.
- [25] Paul Ekman and Wallace V Friesen. *Manual for the facial action coding system*. Consulting Psychologists Press, 1978.
- [26] EV Evarts, H Teräsväin, and DB Calne. Reaction time in parkinson's disease. *Brain: a journal of neurology*, 104(Pt 1):167–186, 1981.
- [27] Joan M. Fallon, James J. Fallon, Matthew Heil, and Stephen J. Weiss. Systems and methods employing remote data gathering and monitoring for diagnosing, staging, and treatment of parkinsons disease, movement and neurological disorders, and chronic pain. United States Patent Application, July 2010.
- [28] Gunnar Farneback. Two-frame motion estimation based on polynomial expansion. In *Proceedings of the 13th Scandinavian Conference on Image Analysis, SCLIA'03*, pages 363–370, Berlin, Heidelberg, 2003. Springer-Verlag. ISBN 3-540-40601-8. URL <http://dl.acm.org/citation.cfm?id=1763974.1764031>.
- [29] Movement Disorder Society Task Force. The unified parkinson's disease rating scale (updrs): Status and recommendations. *Movement Disorders*, 18(7):738–750, July 2003. doi: 10.1002/mds.10473.
- [30] Parkinson's Foundation. Levodopa, 2018. URL <http://parkinson.org/Understanding-Parkinsons/Treatment/Prescription-Medications/Levodopa>.
- [31] Parkinson's Foundation. What is parkinson's?, 2018. URL <http://parkinson.org/understanding-parkinsons/what-is-parkinsons>.
- [32] Parkinson's Foundation. Statistics, 2018. URL <http://parkinson.org/Understanding-Parkinsons/Causes-and-Statistics/Statistics>.
- [33] Christopher G. Goetz, Barbara C. Tilley, Stephanie R. Shaftman, Glenn T. Stebbins, Stanley Fahn, Pablo Martinez-Martin, Werner Poewe, Cristina Sampaio, Matthew B. Stern, Richard Dodel, Bruno Dubois, Robert Holloway, Joseph Jankovic, Jaime Kulisevsky, Anthony E. Lang, Andrew Lees, Sue Leurgans, Peter A. LeWitt, David Nyenhuis, C. Warren Olanow, Olivier Rascol, Anette Schrag, Jeanne A. Teresi, Jacobus J. van Hilten, and Nancy LaPelle. Movement disorder society-sponsored revision of the unified parkinson's disease rating scale (mds-updrs): Scale presentation and clinimetric testing results. *Movement Disorders*, 23(15):2129–2170, 2008. doi: 10.1002/mds.22340.
- [34] Margaret M Hoehn, Melvin D Yahr, et al. Parkinsonism: onset, progression, and mortality. *Neurology*, 50(2):318–318, 1998.
- [35] Daniel H Jacobs, Jeffrey Shuren, Dawn Bowers, and Kenneth M Heilman. Emotional facial imagery, perception, and expression in parkinson's disease. *Neurology*, 45(9):1696–1702, 1995.
- [36] Muaz Khan. Recordrtc: Webrtc javascript library for audio+video+screen recording, June 2017. URL <https://github.com/muaz-khan/RecordRTC>.
- [37] Jeff T Larsen, Catherine J Norris, and John T Cacioppo. Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, 40(5):776–785, 2003.
- [38] Max A. Little, Patrick E. McSharry, Eric J. Hunter, Jennifer Spielman, and Lorraine O. Ramig. Suitability of dysphonia measurements for telemonitoring of parkinson's disease. *IEEE Trans. Biomed. Eng.*, 56(4):1015–1022, April 2009. doi: 10.1109/TBME.2008.2005954. URL <http://ieeexplore.ieee.org/document/4636708/>.
- [39] Max A Little, Patrick E McSharry, Eric J Hunter, Jennifer Spielman, Lorraine O Ramig, et al. Suitability of dysphonia measurements for telemonitoring of parkinson's disease. *IEEE transactions on biomedical engineering*, 56(4):1015–1022, 2009.
- [40] Irene Litvan, Jennifer G Goldman, Alexander I Tröster, Ben A Schmand, Daniel Weintraub, Ronald C Petersen, Brit Mollenhauer, Charles H Adler, Karen Marder, Caroline H Williams-Gray, et al. Diagnostic criteria for mild cognitive impairment in parkinson's disease: Movement disorder society task force guidelines. *Movement disorders*, 27(3):349–356, 2012.
- [41] GRAEME Macphee. Diagnosis and differential diagnosis of parkinson's disease. *Parkinson's disease in the older patient*, pages 41–75, 2008.
- [42] C. Marras, J. C. Beck, J. H. Bower, E. Roberts, B. Ritz, G. W. Ross, R. D. Abbott, R. Savica, S. K. Van Den Eeden, A. W. Willis, and CM Tanner on behalf of the Parkinson's Foundation P4 Group. Prevalence of parkinson's disease across north america. *npj Parkinson's Disease*, 4(21):1–7, 2018. doi: 10.1038/s41531-018-0058-0.

- [43] Pablo Martínez-Martín, Carmen Rodríguez-Blázquez, Mario Alvarez, Tomoko Arakaki, Víctor Campos Arillo, Pedro Pedro Chaná, William Fernández, Nélida Garretto, Juan Carlos Martínez-Castrillo, Mayela Rodríguez-Violante, Marcos Serrano-Dueñas, Diego Ballesteros, Jose Manuel Rojo-Abuin, Kallol Ray Chaudhuri, and Marcelo Merello. Parkinson's disease severity levels and mds-unified parkinson's disease rating scale. *Parkinsonism and Related Disorders*, 21(1):50–54, 2015.
- [44] Weerasak Muangpaisan, Hiroyuki Hori, and Carol Brayne. Systematic review of the prevalence and incidence of parkinson's disease in asia. *Journal of Epidemiology*, 19(6):281–293, 2009.
- [45] NIA and NLM. Making your web site senior-friendly: A checklist, 2002. URL <https://www.nlm.nih.gov/pubs/checklist.pdf>.
- [46] Thanneer M. Perumal, Megharyam Tummacherla, Phil Snyder, Elias Chaibub Neto, E. Ray Dorsey, Lara Mangravite, and Larsson Omberg. Remote assessment, in real-world setting, of tremor severity in parkinson's disease patients using smartphone inertial sensors. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, UbiComp '18, pages 215–218, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5966-5. doi: 10.1145/3267305.3267612. URL <http://doi.acm.org/10.1145/3267305.3267612>.
- [47] L. Ricciardi, M. Bologna, F. Morgante, D. Ricciardi, B. Morabito, D. Volpe, D. Martino, A. Tessitore, M. Pomponi, A.R. Bentivoglio, R. Bernabei, and A. Fasano. Reduced facial expressiveness in parkinson's disease: A pure motor disorder? *J. Neurol. Sci.*, 358(1-2): 125–130, 2015.
- [48] Abdulwahab Sahyoun, Karim Chehab, Osama Al-Madani, Fadi Aloul, and Assim Sagahyroon. Parknos: Diagnosing parkinson's disease using mobile phones. In *IEEE 18th Int. Conf. e-Health Networking, Applicat. and Services*, pages 387–392, 2016. doi: 10.1109/HealthCom.2016.7749491. URL <http://ieeexplore.ieee.org/document/7749491/>.
- [49] Arash Salarian, Heike Rusmann, Christian Wider, Pierre R. Burkhard, Franios J. G. Vingerhoets, and Kamiar Aminian. Quantification of tremor and bradykinesia in parkinson's disease using a novel ambulatory monitoring system. *IEEE Trans. Biomed. Eng.*, 54(2):313–322, February 2007. doi: 10.1109/TBME.2006.886670. URL <http://ieeexplore.ieee.org/document/4067166/>.
- [50] Gwenda Simons, Heiner Ellgring, and Marcia Smith Pasqualini. Disturbance of spontaneous and posed facial expressions in parkinson's disease. *Cognition and Emotion*, 17(5):759–778, 2003. doi: 10.1080/02699930302280.
- [51] Marcia C. Smith, Melissa K. Smith, and Heiner Ellgring. Spontaneous and posed facial expression in parkinson's disease. *Journal of the International Neuropsychological Society*, 2(5):383–391, 1996. doi: 10.1017/S1355617700001454.
- [52] Elias M Stein and Guido Weiss. *Introduction to Fourier analysis on Euclidean spaces (PMS-32)*, volume 32. Princeton university press, 2016.
- [53] Linda Tickle-Degnen and Kathleen Doyle Lyons. Practitioners' impressions of patients with parkinson's disease: the social ecology of the expressive mask. *Social Science & Medicine*, 58(3):603–614, 2004.
- [54] Alexandros T. Tzallas, Markos G. Tsipouras, Georgios Rigas, Dimitrios G. Tsalikakis, Evaggelos C. Karvounis, Maria Chondrogiorgi, Fotis Psomadellis, Jorge Cancela, Matteo Pastorino, Maria Teresa Arredondo Waldmeyer, Spiros Konitsiotis, and Dimitrios I. Fotiadis. Perform: A system for monitoring, assessment and management of patients with parkinson's disease. *Sensors (Basel)*, 14(11):21329–21357, November 2014. doi: 10.3390/s141121329.
- [55] Zdenka Uhríková and Václav Hlaváč. Periodic motion detection on patient with motion disorders. In *Computer-Based Medical Systems, 2008. CBMS'08. 21st IEEE International Symposium on*, pages 90–92. IEEE, 2008.
- [56] Zdenka Uhríková, Otakar Šprdlík, Václav Hlaváč, and Evžen Růžička. Action tremor analysis from ordinary video sequence. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 6123–6126. IEEE, 2009.
- [57] San Francisco University of California. Parkinson's disease clinic and research center: Parkinson's disease medications, 2014. URL <http://pdcenter.neurology.ucsf.edu/patients-guide/parkinson%E2%80%99s-disease-medications>.
- [58] Sarah Vercruyse, Joke Spildooren, Elke Heremans, Jochen Vandenbossche, Oron Levin, Nicole Wenderoth, Stephan P Swinnen, Luc Janssens, Wim Vandenberghe, and Alice Nieuwboer. Freezing in parkinson's disease: a spatiotemporal motor disorder beyond gait. *Movement Disorders*, 27(2):254–263, 2012.
- [59] Nomi Vinokurov, David Arkadir, Eduard Linetsky, Hagai Bergman, and Daphna Weinshall. Quantifying hypomimia in parkinson patients using a depth camera. In Silvia Serino, Aleksandar Matic, Dimitris Giakoumis, Guillaume Lopez, and Pietro Cipresso, editors, *Pervasive Computing Paradigms for Mental Health*, pages 63–71, Cham, 2016. Springer International Publishing. ISBN 978-3-319-32270-4.
- [60] A.W. Willis, M. Schootman, B.A. Evanoff, J.S. Perlmutter, and B.A. Racette. Neurologist care in parkinson disease. *Neurology*, 77(9): 851–857, 2011. ISSN 0028-3878. doi: 10.1212/WNL.0b013e31822c9123. URL <http://n.neurology.org/content/77/9/851>.
- [61] Peng Wu, Isabel Valle González, Dongmei Jiang, and Hichem Sahli. Objectifying facial expressivity assessment of parkinson's patients. 2015.
- [62] A. Zhan, S. Mohan, C. Tarolli, and et al. Using smartphones and machine learning to quantify parkinson disease severity: The mobile parkinson disease score. *JAMA Neurology*, 2018. doi: 10.1001/jamaneurol.2018.0809.
- [63] Andong Zhan, Max A. Little, Denzil A. Harris, Solomon O. Abiola, E. Ray Dorsey, Suchi Saria, and Andreas Terzis. High frequency remote monitoring of parkinson's disease via smartphone: Platform overview and medication response detection. *CoRR*, abs/1601.00960, 2016.

Received November 2018; revised February 2019; accepted April 2019