Quantifying Levodopa-Induced Dyskinesia Using Depth Camera

Maria Dyshel[1], David Arkadir[2], Hagai Bergman[2], and Daphna Weinshall[1]

[1]School of Computer Science and Engineering, Hebrew University, Jerusalem 91904, Israel
[2]Hadassah Medical School, Hebrew University, Jerusalem 91120, Israel

maria.dyshel@mail.huji.ac.il,arkadir@gmail.com,hagaibe@ekmd.huji.ac.il,daphna@cs.huji.ac.il

Abstract

We present a novel method to detect and assess the severity of Levodopa-Induced Dyskinesia (LID) in Parkinson's Disease (PD) patients, based on Microsoft Kinect recordings of the patients. Dyskinesia denotes involuntary movements induced by chronic treatment with levodopa in patients with PD. Detection and objective quantification of dyskinesia is essential for optimizing the medication regime and developing novel treatments for PD. We used Microsoft Kinect sensor to track limb and neck movements of a patient performing two motor tasks. Using a new motion segmentation algorithm, kinematic features were extracted from the videos and classified using Support Vector Machines (SVMs). The method was tested on 24 recordings of 9 PD patients, and achieved sensitivity of 0.82 at EER in overall dyskinesia detection. Moreover, it provided a numerical overall score for the severity of dyskinesia, which showed high correlation with the neurologist's assessment of the patient's state. The study shows that depth camera recordings can be used to monitor and grade the severity of levodopa-induced dyskinesia, and therefore can potentially provide valuable aid to clinicians and researchers.

1. Introduction

Parkinson's disease is a neurodegenerative disorder that manifests itself in various movement disorders, such as bradykinesia, tremor and rigidity. It is most often treated with dopamine precursor, levodopa; however, long-term treatment by levodopa often results in involuntary movements, called dyskinesia [15]. The pattern and severity level of levodopa-induced dyskinesia (LID) vary between different patients, and even for the same patient the severity of dyskinesia fluctuates substantially both on short and long timescale. Quantitative assessment of the dyskinetic condition will benefit diagnosis of the disease as well as treatment evaluation. Most clinical scales used nowadays are based on the patient's self-assessment and home diaries, as well as a clinician's observations. However, due to the fluctuations of LID in time, a single evaluation may not reflect the severity of LID experienced by the patient in everyday activities, and self-evaluation of the condition by the patient may be inaccurate. Therefore, an automated and objective procedure to quantify the severity of LID that could be carried out at home would aid the clinicians in monitoring and optimizing the patient's treatment, and possibly in developing new treatments for the disorder.

In recent years, several approaches have been tried to quantify LID in an automatic way. Wearable technology, such as gyroscopes and accelerometers have been used to quantify the patients' motion, both during performing of specialized tasks [11, 17, 10] and in everyday life [9, 12]. Diagnostic potential of other medical equipment, such as medical force plates [2], was also explored. Most of the earlier attempts either require that the patient wear cumbersome and expensive equipment or that dyskinesia assessment is performed in laboratory conditions, making daily monitoring of the patient's state impossible. Newer generations of wearable technology enable home monitoring of the patient; however, the amount and placement possibilities of the sensors on the patient's body are limited, which affects the quality of detection and characterization of the dyskinetic patterns. Furthermore, it was found that even though older patients have positive attitude towards technological developments designed to improve their healthcare, they are reluctant to use wearable technology [5].

Recent work of Rao *et al.* [13] proposed a vision-based method of LID quantification. In [13], the researchers use point distribution models (PDM) to track the face features of PD-affected patients on standardized videos taken during medical evaluation. A heuristic function is then devel-

oped based on the tracking data to quantify the severity of the dyskinesia, based on the number of principal components in the data and their relative weight. While providing non-intrusive method of dyskinesia assessment, the method requires standardized video recordings to function, meaning that its application requires that the patients arrive to the hospital for evaluation. Furthermore, this method uses advanced and computationally expensive computer vision techniques to acquire the raw data necessary for analysis, which may hinder its practical application.

Recent advances in computer-vision based home appliances allow us to acquire raw data of reasonable accuracy using a popular device that can be stationed in every home. Specifically, Microsoft Kinect is a motion-sensing device used to directly control video games through body movement. In addition to growing evidence of its potential in physical rehabilitation through gaming and interactive training [4], its possible applications in posture [3] and gait [6] analysis have also been subject to research in recent years. A study by Galna et al. [7] shows that the device can accurately measure the spatial characteristics of movement symptoms in patients with PD. Recent studies offer Kinect-based methods to detect and measure body tremor [16] and freeze of gait [1]; however, to our knowledge, the problem of LID has not been yet addressed.

In this paper we present a novel Kinect-based method to detect dyskinesia in PD patients and rate it according to an approved clinical scale. Ideally, our method will be used in a vision-based system to be stationed in the patient's home and used to monitor the patient's condition. Our method is based on motion detection and machine learning techniques, and therefore data is needed for the algorithm's design. In Section 2 we describe the experimental setup which was used for data collection. In Section 3 we describe our method, including the motion detection algorithm and the training of the final classifier and grader. In Section 4 we describe the empirical evaluation of the method, showing excellent agreement with the assessment provided by a clinician.

2. Methods: experimental setup

2.1. Technology overview

Microsoft Kinect^{TM} was developed as a motion-sensing attachment to the Xbox gaming console. It includes an RGB camera, an IR emitter and sensor, and a microphone array. Kinect produces a depth image using technology based on structured light technology [14]: the depth map is created by matching a pattern of laser IR dots to its projection on the scene in front of the camera. The depth channel is used by the sensor's software to extract the virtual skeleton, which consists of the positions of 20 anatomical landmarks of human body (which we will refer to as "joints" for simplicity)

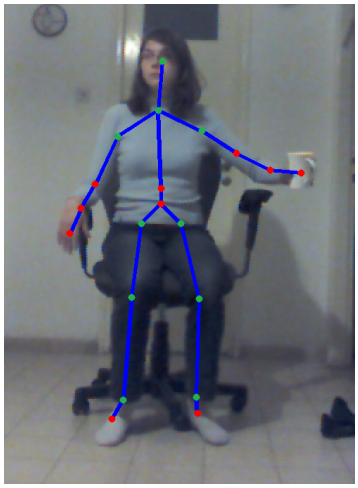

Figure 1: Output of the Kinect sensor. Joint positions are used for motion analysis, while the video recording serves for neurological diagnosis. The dots marked green signify the joints that were chosen to be used in the feature vector.

in three-dimensional space, at the rate of 30 fps and spatial resolution of 640x480 pixels [18]. We used Microsoft Kinect For Windows Standard Development Kit (SDK) to interface with the sensor and acquire the skeletal tracking data and video input. To reduce the noise in joint position tracking, we used the following noise filtering parameters of Kinect API: smoothing parameter 0.5, correction parameter 0.3, prediction parameter 0.5, jittering radius 0.01 and maximun deviation radius 0.04. Example of the Kinect output is shown in Fig. 1.

2.2. Data collection protocol

Nine patients (ages 54 to 72, including 5 women and 4 men), who were affected by Parkinson's disease with varying severity of LID (2 of the patients were not affected by LID), were recruited for our study and data collection. All of the subjects gave written informed consent for study participation in accordance with the Helsinki Declaration. Each of the patients was recorded 2 to 6 times, and for 5 patients, recordings in both dyskinetic and non-dyskinetic states were obtained. Overall, 24 recording sessions were collected and analyzed in this study. Each session included two conditions (sitting and standing, see below), giving us 48 recordings for further analysis.

The subjects were recorded at home at various hours during the day, and the recordings followed the protocol of Abnormal Involuntary Movement Scale (AIMS). During each recording session, the patients were asked to perform two motoric tasks, typically used during the standard Unified Parkinson's Disease Rating Scale (UDPRS) assessment: 1) tapping the thumb with an index finger, and 2) opening and closing the palm. In addition to recording the skeletal joint

positions as detected by Kinect, the RGB camera of the sensor was used to perform a video recording for neurological diagnosis.

Each of the movements was performed repeatedly with each hand over a period of 15 seconds, first in a sitting and then in a standing position, with at least 30 seconds lapse between the tasks. In addition, each patient was recorded while performing the activities required by the protocol of AIMS [8]: sitting with legs slightly apart; getting up from the chair; and walking a few steps. As assessing the severity of oral and facial movements was not part of the study, procedures meant to check the dental and oral state of the patient were omitted. The Kinect sensor was positioned at a distance of ˜2.3 m from the patient and height of ˜40-50 cm from the floor. A researcher stood beside the Kinect to demonstrate the movements and ensure the participant's safety.

Video recordings of tasks from each recording session were grouped into a video sequence together with the videos of the patient performing the additional tasks required by the AIMS protocol. The sequences, provided in random order, were then scored by a movement disorder neurologist using the AIMS grading system. Each video sequence was given 4 integer grades in the 0-4 range, based on the neurologist's impression from the state of the patient in the session. The grades quantified the severity of each of 4 categories of dyskinetic movement: upper limb movement (including finger or arm jerking movements), lower limb movement (including foot tapping or squirming), trunk movement (including neck twisting and trunk rocking), and overall movement. Note that only the recordings while performing the tasks were used as input to the algorithm.

3. Our method

We now describe our method, which included a number of steps: First, using the Kinect recordings, we obtained features based on the motion segmentation of each recording (Section 3.1) followed by noise reduction (Section 3.2). Next, we trained a classifier to detect LID (a binary task) and evaluate its severity; this was done using machine learning techniques, including feature selection (Section 3.3) and classifier training (Section 3.4).

3.1. Motion segmentation algorithm

The three-dimensional coordinates of 20 joints provided by the Kinect SDK (see Fig. 1) during the recordings of the 2 tasks described above were used as raw data for the algorithm. To analyze the motion of each joint, the time series of joint positions were partitioned into motion segments. Let $(X_1^j \dots X_n^j)$ denote the series of vector positions of joint j in frames $1 \dots n$. To partition the positions into segments $S_1^j \dots S_k^j$, the following procedure was used:

Algorithm 1

$S_1^j = [X_1^j]$ *% initialize the first segment*
k = 1
for i =1, $\dots$, $(n-1)$
  $V_i^j = X_{i+1}^j - X_i^j$
  if $\angle(V_i^j, V_{i-1}^j) > \frac{\pi}{2}, \dots, \angle(V_i^j, V_{i-d}^j) > \frac{\pi}{2}$
    $d = min\{5, |S_k^j| - 1\}$
    $S_k^j = S_k^j \cup [X_{i+1}^j]$ *% add position to existing segment*
  else
    $S_{k+1}^j = [X_{i+1}^j]$ *% start a new segment*
    $k = k + 1$

More specifically, the end of one motion segment and the start of the other is determined by the angles between the vector differences between adjacent points in the motion sequence. As long as the difference vector forms an obtuse angle with d previous difference vectors in the joint motion (we set $d = 5$), the points constituting it will belong to the same motion segment. If an acute angle (less than $\frac{\pi}{2}$) occurs between the difference vectors, signifying a change of direction, a new motion segment is started. This segmentation method helps to detect twitching, jerking movements with sudden change of directions that are typical in dyskinetic condition. We define the *length* of the segment to be the sum of the length of each vector difference between two adjacent joint positions in the segment.

3.2. Noise Reduction

Kinect joint detection is characterized by varying amounts of noise, mostly "jittering" noise. Even when we used enhanced smoothing parameters implemented by Kinect Software (described in technology overview), the deviation between the positions of the same joint in adjacent frames (0.03 seconds apart) resulting from the noise can sometimes reach a few centimeters. It was found that several parameters affect the amount and amplitude of the noise, such as:

- Lighting conditions - as natural light spectrum includes IR frequencies, daylight affects the quality of joint detection as well as the amount of noise in joint detection.
- Subject's clothing color - it was found that dark clothing hinders detection of the relevant joints.
- Subject's complexion and height can affect the quality of detection.
- Subject's posture - in general, most of the joints, especially spine and hip joints, are more noisy in sitting position, and the detection quality is worse than in standing position.

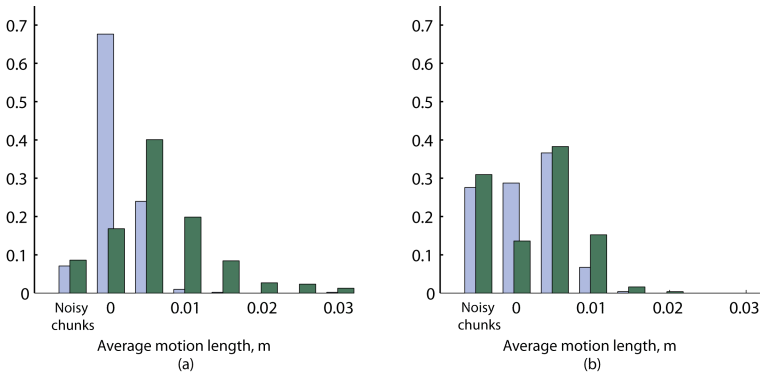

Figure 2: Choosing representative joints for movement representation: Normalized distributions of average motion lengths for (a) a representative joint (shoulder) (b) non-representative joint (elbow). The feature distribution is plotted in green for dyskinetic recordings and in blue for non-dyskinetic recordings.

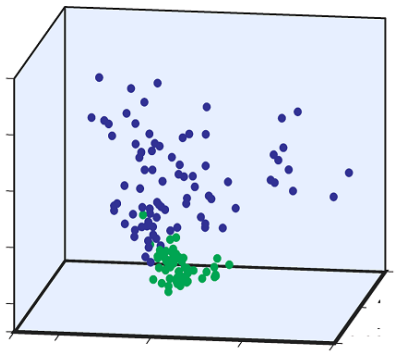

Figure 3: PCA representation of the feature vectors representing the chunks from dyskinetic (blue) and non-dyskinetic (green) videos of the same patient.

To deal with the noise, a number of techniques were used. Traditional jitter removal filtering techniques may introduce new motion segments into the series, affecting the segment length statistics, and we therefore sought another way to reduce noise. Jitter was removed by limiting the maximal distance between the joint positions in adjacent time frames. To remove the noisy fragments, all adjacent frames with difference above a threshold were marked as noisy frames, and the frames in their immediate proximity were weighted according to their distance to the noisy frames. All the frames with distance proximity grade above some threshold were deleted, and the edges of the removed fragments were marked so that they could be recognized by the motion segmentation algorithm. The segmentation algorithm was updated to end a motion segment and start a new one whenever it encountered this mark in the time series of joint positions.

3.3. Feature selection

Each video (450 frames long) was divided into 50-frame long chunks with overlap of 25 frames. We began by analyzing the movement of all joints except the 3 joints that participated directly in the performance of the task (those of the left or the right hand respectively). Thus the movement of each joint was segmented according to Algorithm 1, and then quantified by the average characteristic of these segments using 1 of 3 characterization methods. We use a single representation for both cases, when the task was performed by either the right or left hand. To accomplish this, each feature in this representation was labeled relatively to the hand that performed the task, such as 'same-side knee', 'opposing elbow' or 'opposing knee'.

Specifically, in each chunk each joint was assigned a list of motion segments. Subsequently, its representation was based on some average property of these segments, using one of the following methods:

- **Average motion length**: average segment length in the chunk.
- **Average motion speed:** average speed of the segment in the chnunk ($^{length}/_{num\ of\ frames * 0.03\,sec}$).
- **Distribution of quantized motion lengths:** the lengths of motion segments for each joint were quantized into three levels using k-means clustering. Each discriminative joint was represented in the vector by the normalized 3-bin distribution.

Next, we used the training data and discriminative feature selection in order to select the most discriminative joints which will represent each motion chunk. Specifically, for each joint we extracted the distribution of its chunk value (as computed above) based on all the recordings. Two distributions, represented by 30-bin histograms, were extracted, one for the dyskinetic state and one for the normal state (see examples in Fig. 2). The discriminative power of each joint was measured by the Earth Mover's Distance (EMD) between these two histograms. The procedure was performed separately for the two experimental conditions (sitting and standing), and the EMD distances of the histograms in the two conditions were added up to provide a discrimination grade per joint. Subsequently, the 10 joints with the highest discrimination grade were chosen as shown by green dots in Fig. 1. Example of motion length distribution in a discriminative joint vs. non-discriminative is shown in Fig. 2.

The feature selection process resulted in a 10-dimensional feature vector representing each chunk, where each of the chosen 10 joints was represented by a single number, calculated according to one of the three methods

listed above. Prior to classification, principal component analysis (PCA) was applied to reduce the dimensionality of the feature vector. To determine the number of dimensions, variance accumulation was applied, and it was found that in all cases 3 to 4 principal components accounted for 90% of the data variance; therefore, 4-dimensional PCA representations of the feature vectors described above were used for subsequent classification and grading. The separability of the PCA representations of chunks taken from dyskinetic and non-dyskinetic sessions can be seen on Fig. 3.

Joint selection was only used to assess the overall dyskinetic state of the patient. To assess the dyskinesia level in a restricted area (upper limbs, lower limbs, and the trunk), all the joints from the appropriate body part were used. For example, both knee, ankle and foot joints were used for lower limbs dyskinesia assessment, while the head, shoulders and hips were used for assessing trunk dyskinesia. Therefore, the classification and grading stages for the local dyskinesia assessment analysis involved 6-dimensional (for lower limb dyskinesia), 3-dimensional (for upper limb dyskinesia) and 7-dimensional (for trunk dyskinesia). Since the feature vectors in the local dyskinesia analysis were of low dimension to begin with, further dimensionality reduction by PCA was not required.

3.4. Classification and grading

Algorithm 2

Input 1: p - set of chunks from a pair of 1 dyskinetic and 1 non-dyskinetic recordings of the same patient
Input 2: $[r_1 \dots r_m]$ - sets of chunks from all other m recordings (m = 24 - 2n)
Output: grades = $[g_1 ... g_m]$ for every recording in Input 2.
Train $s = SVM\_RBF(p)$
Do for $j = 1, 2, \dots m$
  For each chunk c in r_m:
    Compute $s(c)$
  End
  $g_j = \frac{\#\{c|s(c)=True\}}{\#c}$
End

At the training stage, feature vectors representing chunks from recordings of the patients with both dyskinetic (with AIMS dyskinesia grade at least 2) and non-dyskinetic videos available were used for training a SVM classifier. We used soft-margin SVM with RBF kernel to separate between feature vectors representing recording chunks taken from dyskinetic and non-dyskinetic clips. We used the MATLAB Statistical toolbox to perform the classification.

The trained classifier was used to label the feature vectors representing the chunks of every recording in the data set not used for training. The grade of the recording was determined by computing the fraction of chunks classified as dyskinetic. The recording session received the average grade of all the recordings that were included in it (see Section 2.2). The computed grade was used for binary classification of the patient's condition, as well as for multi-level quantification of the dyskinesia severity.

4. Results

We used our data to test the ability of our method to accomplish one of two tasks: binary classification of the dyskinetic condition - present or not (Section 4.1), and quantified assessment of dyskinesia severity and its agreement with the clinician's evaluation (Section 4.1.1).

4.1. Binary classification of dyskinetic condition

The ability of our method to detect dyskinesia in a patient was checked by turning the neurologist grade into a binary grade - all recording sessions with grade above zero were assigned a value of 1. The procedure outlined in Section 3.4 was then executed to train the desired classifier. We used the 3 methods to characterize each movement segment as described in Section 3.3, and for comparison plotted the respective Receiver Operator Characteristics (ROC) curve. The Area Under the Curve (AUC) was used to measure the effectiveness of each method.

Specifically, to test our method's ability to detect dyskinesia in a patient (overall dyskinesia grade above 0), the algorithm was trained on recordings of a single patient, and then used to grade all the recordings in the database not used for training. We repeated this procedure for every patient for whom we had both dyskinetic and non-dyskinetic recordings available (resulting in 5 train-test sessions in total), each time using the data of a different patient to train the classifier. We then calculated the average AUC for each of the remaining recordings.

The results of the tests (averaged over the 5 repetitions) are shown in Table 1. The ROC curves of classifying the patient's state in 48 recordings (24 recording sessions with 2 postures each) using the three methods are presented in Fig. 4. These curves show nice agreement between the algorithm's classification decisions and the clinician's assessment. All 3 methods used to evaluate a movements segment show similar performance (with the *Motion length distribution* method performing slightly worse), achieving up to 0.85 sensitivity at Equal Error Rate point (EER).

	Average AUC	Average GCC
Average motion length	0.882	0.805
Motion length distribution	0.862	0.703
Average motion speed	0.906	0.789

Table 1: Average performance of the chosen features in assessing the overall classification success.

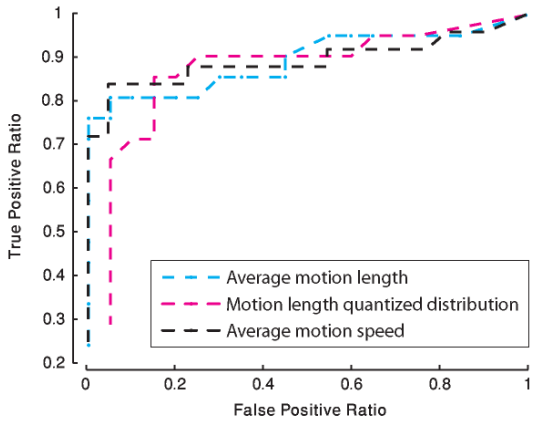

Figure 4: The ROC curves of dyskinesia detection using average motion length (blue), motion length bag of words (red) and average motion speed (black).

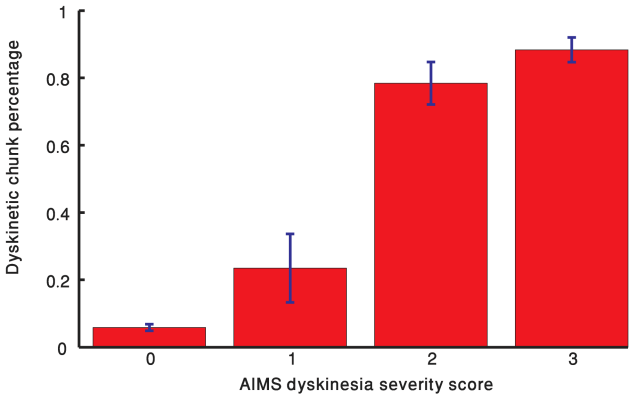

Figure 5: The correlation between the algorithm-produced grades and the grades given by the neurologist. The bars show the mean grade given by the algorithm to recordings with the appropriate AIMS grade, while the error bars show the variance.

4.1.1 Quantified assessment of dyskinesia severity

Since our method provides a quantitative measure of dyskinesia, we can use it to grade the severity of dyskinesia in each session. Fig. 5 shows the excellent qualitative agreement between the algorithm grades and the neurologist grades. We compared our measure to the grades given by the neurologist using the general correlation coefficient (GCC): Let a_i denote the grade assigned by the algorithm to session i, and let b_i denote the grade assigned by the neurologist. The general correlation coefficient (GCC) between the two grading scores is given by:

$$\Gamma = \frac{\sum_{i,j=1}^{n}(a_i - a_j)(b_i - b_j)}{\sum_{i,j=1}^{n}(a_i - a_j)^2 \sum_{i,j=1}^{n}(b_i - b_j)^2} \qquad (1)$$

The average GCC scores for the three tested methods are presented in Table 1. Due to the limited amount of available representatives of each dyskinesia severity class, further analysis of the grading process is left for future work.

4.1.2 Detection of dyskinesia in different parts of the body:

The ROC was calculated for every grade of the four grades of the AIMS protocol described in the previous section. For each grade, we chose the patients that displayed dyskinesia in the relevant region (with a grade of at least 2) for classifier training, and the classifier was then used to test all the recordings in the data set. 5 training sets of different patients were run to test lower limb and trunk dyskinesia detection; only 2 sets of training data showing patients with significant upper limb dyskinesia were available. The comparative results of using average motion length features to detect dyskinesia in upper limbs, lower limbs, and trunk, are presented in Fig. 6. Similar results were obtained for average motion speed features.

As can be seen in Fig. 6, the algorithm performs well in detecting trunk, shoulders and neck dyskinesia (sensitivity of 0.82 at EER), and slightly worse in detecting dyskinesia in lower limbs (sensitivity of 0.75 at EER). The algorithm's performance in detecting dyskinesia in upper limbs is much worse. This result may be expected, as hand joints detection was noted to be especially noisy during the recordings, mostly due to proximity of the hand to the body/chair which encumbered proper detection. In addition, as we examine only the motion of joints that do not participate in the tasks, which are performed by a hand, fewer joints are available for analysis in the case of the upper limbs.

4.1.3 Monitoring the state of a single patient:

To monitor the dyskinesia severity of a patient during the day, 6 recordings of the same patient were used. This patient showed a constant pattern of dyskinesia severity, with dyskinesia starting in the late morning hours, intensifying during the afternoon, and weakening towards the evening. The algorithm was trained on two training sets: the first set included two recordings of the patient - one taken in the morning, before the patient took the first medication of the day, and the other taken in the middle of the day, when the severity of dyskinesia was at its highest. The second set consisted of dyskinetic and non-dyskinetic recordings of 5 different patients. The algorithm was then used to grade all 6 recordings of the patient (for the recordings used in training, half of the recording was used for training and the other half for testing).

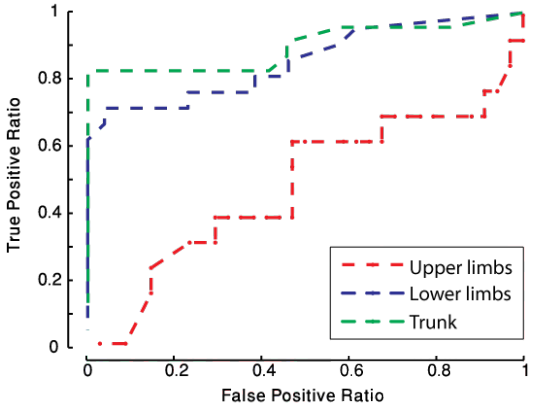

Figure 6: The ROC curves of dyskinesia detection in upper limbs (red), lower limbs (blue) and trunk (green), using average motion length feature.

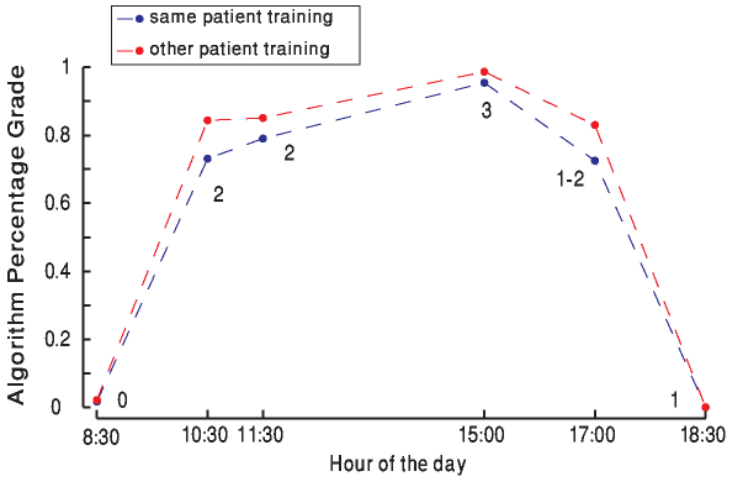

Figure 7: Results of monitoring the state of a single patient. The points represent the grades given by the algorithm to recordings of the patient taken during the day; the numbers near the points are the grades given by the neurologist

The grades of the recording given by the algorithm, compared to the grades given by the neurologist are presented in Fig. 7. The algorithm's grades presented in the figure show not only good correlation with the neurologist's grades, but also agree with the subjective perception of the patient regarding the severity of dyskinesia during the day. In addition, the recordings were ranked by the neurologist in pairwise comparisons of every two clips. Kendall's tau between the algorithm's ranking and the neurologist's ranking of the video is 0.6

5. Discussion

The results of the study show the ability of our method to accurately detect dykinesia in patients, and to give a severity assessment that correlates well with the assessment of a trained neurologist. Remarkably, the algorithm performs well even when trained on a very small set of one or two patients. Dyskinesia binary detection accuracy shown by the algorithm is similar to the accuracy of methods involving wearable accelerometers (for example, [10] reports reaching sensitivity of 0.73 with specificity of 1.0, as opposed to sensitivity 0.76 with specificity 1.0 reached in this study). However, in our case, the assessment procedure does not require the patient to use wearable devices, which may cause discomfort for the patient. On the other hand, using a camera for medical diagnosis can provide inaccurate results due to changes in the patient's movement pattern when observed. The ability of the algorithm to detect dyskinesia in various body parts varies according to the amount of noise affecting the Kinect's detection of those parts.

Still, the algorithm faces several challenges in achieving a reliable assessment of the patient's condition. The basic limitation of the system is the quality of joint detection by the Kinect camera and SDK. The Kinect sensor we used has limited efficiency in detecting small body movements, such as slight twitches and tremors, as well as finger movements, and can easily be mislead when the body parts are too close to each other. In addition, its facial tracking capabilities are currently very poor, so that it cannot currently be used to detect or rate dyskinetic condition of the face, mouth or teeth, even though they constitute an important part of the patient's state assessment. Newer generations of Kinect, which are becoming available, will likely enable more accurate and less noisy tracking of the patient's movement. The quality of dyskinesia detection and quantification using the next generation of Kinect is left for future work.

Though using motion segmentation and SVM classification yielded reasonably good performance, additional feature selection and classification techniques can be employed to improve the detection and quantification accuracy, as well as provide additional insights into the dyskinetic patients behavior. For example, using whole-body bag-of-words dictionary to detect more complex motion features can be used to detect more distinct features of the patient's movement, and therefore provide valuable insights for the treating physician. In addition, analyzing the visual and depth data acquired from the camera in conjunction with the skeletal data can significantly improve the assessment accuracy. A previous study showed the efficiency of neural networks [9] in assessing dyskinesia severity based on data acquired by accelerometers worn by the patient; a similar algorithm can be tested using our movement representation. These techniques may be experimented with in the future, as more data becomes available.

6. Summary

We have successfully showed that Microsoft Kinect can be used to detect dyskinesia and assess its severity of patients with Parkinson's disease, using very short assessment sessions which last around a minute long. We developed a machine learning algorithm that was able to grade the dyskinetic behavior both when trained on recordings by the same patient it was grading, or when trained on recordings of other patients. The algorithm achieved binary classification sensitivity of 0.82 at EER, and its severity grades were in good correlation with the clinician's assessments (correlation 0.8). Furthermore, we showed that the algorithm can be used to monitor the condition of PD patient during the day.

Acknowledgements

This work was supported in part by the Intel Collaborative Research Institute for Computational Intelligence (ICRI-CI), a grant from the Israel Science Foundation (ISF) to Prof. Daphna Weinshall, and by the Gatsby Charitable Foundations.

References

[1] A. Amini Maghsoud Bigy, K. Banitsas, A. Badii, and J. Cosmas. Recognition of postures and Freezing of Gait in Parkinson's disease patients using Microsoft Kinect sensor. In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 731–734, Apr. 2015.

[2] K. A. Chung, B. M. Lobb, J. G. Nutt, J. McNames, and F. Horak. Objective measurement of dyskinesia in parkinson disease using a force plate. *Movement Disorders*, 25(5):602–608, Apr. 2010.

[3] R. A. Clark, Y.-H. Pua, K. Fortin, C. Ritchie, K. E. Webster, L. Denehy, and A. L. Bryant. Validity of the microsoft kinect for assessment of postural control. *Gait & Posture*, 36(3):372–377, July 2012.

[4] A. Da Gama, P. Fallavollita, V. Teichrieb, and N. Navab. Motor Rehabilitation Using Kinect: A Systematic Review. *Games for Health Journal*, 4(2):123–135, Feb. 2015.

[5] G. Demiris, M. Rantz, M. Aud, K. Marek, H. Tyrer, M. Skubic, and A. Hussam. Older adults' attitudes towards and perceptions of "smart home" technologies: a pilot study. *Medical Informatics and Internet in Medicine*, 29(2):87–94, June 2004.

[6] M. Gabel, R. Gilad-Bachrach, E. Renshaw, and A. Schuster. Full body gait analysis with kinect. In *2012 Annual International Conference of IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1964–1967, Aug. 2012.

[7] B. Galna, G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, and L. Rochester. Accuracy of the microsoft kinect sensor for measuring movement in people with parkinson's disease. *Gait & Posture*, 39(4):1062–1068, Apr. 2014.

[8] W. Guy. *ECDEU assessment manual for psychopharmacology*. Rockville, Md. : U.S. Dept. of Health, 1976.

[9] N. L. W. Keijsers, M. W. I. M. Horstink, and S. C. A. M. Gielen. Automatic assessment of levodopa-induced dyskinesias in daily life by neural networks. *Movement Disorders*, 18(1):70–80, Jan. 2003.

[10] T. Mera, M. Burack, and J. Giuffrida. Quantitative assessment of levodopa-induced dyskinesia using automated motion sensing technology. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 154–157, Aug. 2012.

[11] S. Patel, D. Sherrill, R. Hughes, T. Hester, N. Huggins, T. Lie-Nemeth, D. Standaert, and P. Bonato. Analysis of the severity of dyskinesia in patients with parkinson's disease via wearable sensors. In *International Workshop on Wearable and Implantable Body Sensor Networks, 2006. BSN 2006*, pages 124–126, Apr. 2006.

[12] C. Pulliam, D. Heldman, E. Brokaw, M. Burack, and T. Mera. Continuous Motion Sensor Assessment of Parkinsons Disease During Activities of Daily Living (P1.196). *Neurology*, 84(14 Supplement):P1.196, Apr. 2015.

[13] A. Rao, R. Bodenheimer, T. L. Davis, R. Li, C. Voight, and B. M. Dawant. Quantifying drug induced dyskinesia in parkinson's disease patients using standardized videos. *Conference Proceedings of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2008:1769–1772, 2008.

[14] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *Proceedings of Computer Society Conference on Computer Vision and Pattern Recognition, 2003*, volume 1, pages I–195. IEEE, 2003.

[15] B. Thanvi, N. Lo, and T. Robinson. Levodopa-induced dyskinesia in parkinson's disease: clinical features, pathogenesis, prevention and treatment. *Postgraduate Medical Journal*, 83(980):384–388, June 2007.

[16] R. Torres, M. Huerta, R. Clotet, R. Gonzalez, L. E. Sanchez, D. Rivas, and M. Erazo. A Kinect Based Approach to Assist in the Diagnosis and Quantification of Parkinson's Disease. In A. Braidot and A. Hadad, editors, *VI Latin American Congress on Biomedical Engineering CLAIB 2014*, number 49 in IFMBE Proceedings, pages 461–464. Springer International Publishing, 2015.

[17] M. G. Tsipouras, A. T. Tzallas, G. Rigas, P. Bougia, D. I. Fotiadis, and S. Konitsiotis. Automated levodopa-induced dyskinesia assessment. *Conference Proceedings of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2010:2411–2414, 2010.

[18] Z. Zhang. Microsoft kinect sensor and its effect. *IEEE MultiMedia*, 19(2):4–10, Feb. 2012.