2014 IEEE 28-th Convention of Electrical and Electronics Engineers in Israel

Automatic Assessment of Parkinson's Disease From Natural Hands Movements Using 3D Depth Sensor

Ben Dror, Eilon Yanai, Alex Frid, Nimrod Peleg
Technion, IIT
Haifa, Israel

Nadav Goldenthal
Technion, IIT and Microsoft Israel
Haifa, Israel

Ilana Shlezinger
Rambam Health Care Campus
Haifa, Israel

Hagit Hel-Or, Shmuel Raz
University of Haifa,
Haifa, Israel

***Abstract*— Parkinson's Disease (PD) is a degenerative disease of the central nervous system with a profound effect on the motor system. Symptoms include slowness of movement, rigidity of motion and in some patients, tremor. The severity of the disease is quantified using the Unified Parkinson Disease Rating Scale (UPDRS) which is a subjective scale performed and scored by physicians. In this work, we present an automated, objective quantitative analysis of four UPDRS motor examinations of Hand Movement and Finger Taps.**

For this purpose, a non-invasive system for recording and analysis of fine motor skills of hands was developed. The system is based on a simple low-cost depth acquisition sensor, similar to the second generation of Microsoft's Kinect sensor, and novel recursive self-correcting hand tracking algorithm. The system allows patients to perform test tasks in a natural and unhindered manner.

The evaluation of the system was carried out on PD patients and controls. Machine Learning based classification was performed on the acquired data, followed by a decision making scheme.

***Index Terms*—Machine Learning, Support Vector Machine (SVM) classification, Parkinson's Disease, UPDRS.**

I. Introduction

Parkinson's disease (PD) is a neurodegenerative disorder characterized by "masklike" facial features, bradykinesia, tremor at rest, and muscle rigidity. PD affects approximately 1%-3% of individuals over the age of 65. The number of patients suffering from PD is increasing; partly in association with increased life expectancy [1]. PD causes a social and financial burden. In the USA, the annual total cost associated with PD was estimated to be 23 billion US$ [2].

The development of new treatments for PD is hampered by the subjectivity of treatment efficacy assessment, which is done today by using the Unified Parkinson Disease Rating Scale (UPDRS) [3]. This tool is based on a score derived from the neurological evaluation performed by the treating physician and is a subjective score. Thus, the UPDRS lacks a high degree of objectivity, impartiality and sensitivity. Moreover, intermittent hospital monitoring, which provides only a narrow window into the health of a person, might miss trends that can lead to early detection of a problem.

Several studies have proposed methods for automating the UPDRS test. These have focused on methods that use wearable devices and sensors such as wireless accelerometers [4], wearable triaxial gyroscopes [5], virtual touchpads [6], [7] or EOG signal analysis (i.e. analysis of the electric signals originating in the eye) [8]. Or remote methods, such as [9], [10], but with many restrictions on patient movements and complex setups, which makes it less approachable for daily use or for unexperienced personnel. Furthermore, none of these studies analyzed movement and motor performance of specific UPDRS motor tasks. And thus, to date, cannot be adopted in clinical practice.

In this work a passive, non-invasive and relatively non restricting system that allows an examinee to perform several UPDRS motion tests in his natural environment, without being confined to a small region in space, is developed using a 3D depth sensor. Specifically, motor features of PD as expressed by Hand Movements, Finger Taps and Rapid Alternating Movements of Hands (questions 23-25 of the motor UPDRS) were acquired. Hands tracking and feature extraction algorithms were used to obtain objective and quantitative measures from the acquired data and Machine Learning techniques were used to classify patients into healthy and PD groups.

II. Systems and Methods

A. Settings and Data Acquisition Procedure

For data acquisition, a Microsoft 3D camera sensor based on Time of Flight (TOF) technology was used. The data from this camera acquired in 4 matrices of 986 by 274 pixels with a rate of 24 frames per second. The first three containing aligned Cartesian coordinates with 1mm resolution, and the fourth matrix represents reflection intensity (of Near Infra-Red spectrum) that gives a user a detailed grayscale image. The advantages of this sensor are its high resolution, relatively fast frame rate and in its native perfect synchronization between the spatial and intensity data.

During the experiment, a total of 13 patients were recorded (8 PD and 5 Healthy), between the ages of 50 to 75, performing each of the four tests. All the patients were checked by physician in order to determine the severity of the disease. The acquisition of Finger Taps and Hand Movements information was performed by requesting the subjects to perform the UPDRS Hand Movement tasks:

1. *Finger Taps* – patient taps thumb with index finger of same hand in rapid succession.
2. *Hand Movements* – patient opens and closes hand in rapid succession.
3. *Rapid Alternating Movements of Hands* (two conditions) – patient performs pronation-supination movements of hands, vertically and horizontally, with as large an amplitude as possible, with both hands simultaneously.

B. Hand Motion Tracking Algorithm

Most of the hands and skeleton tracking algorithms were initially designed for gestures understanding in native user interfaces, and for the gaming industry. This makes them less suitable for use by the elderly, or people with different disabilities, due to different motion-prediction and signal filtering techniques. For this reason, we developed a new robust algorithm for hand tracking, in order to constrain the examinee as little as possible. The only assumption for the proposed hand tracking algorithm is that the examinee is located in front of the camera sensor and within its coverage area.

The proposed hand motion tracking algorithm consists of four main steps: (1) hand palm detection step; (2) hand palm tracking step; (3) center of palm determination step; and (4) self-correction step. These steps are described in detail below.

1) Hand Palm Detection Algorithm

The main stages of this part can be seen in Fig. 1. The algorithm starts by performing background segmentation using depth thresholds, then the center of body mass is identified using clustering and depth gradients detection is performed using a difference filter in the horizontal direction (first and second stages in Fig. 1). Later, an "Edge to Edge" distances map is constructed, in order to segment out the torso and find the head location, describing the distances between the start and end edges in each row (see step three and four in Fig. 1). This process is performed similarly to the one described in [11]. After the main body parts are segmented out, two thresholding procedures (coarse and fine) are performed in order to focus on the palm region (as shown in last two steps in Fig. 1). If after the final step the palm area could not be found, the process is repeated on the following frame.

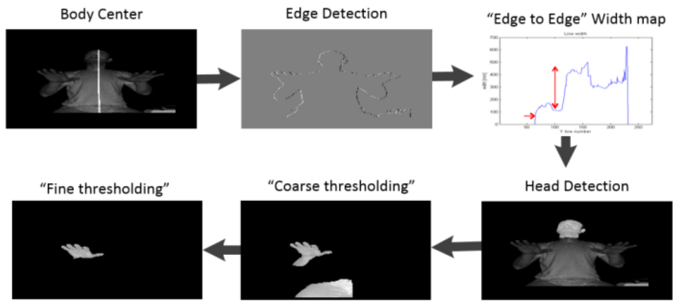

Fig. 1. Palm detection algorithm graphical description. Starting from body center determination, followed by edge detection on depth data, then "Edge to Edge" map creation and head and torso detection and removal. Finished by coarse and fine thresholding, leaving up only the palm displayed.

2) Palm Tracking Step

Once the hand palm is found, its center coordinates are calculated (explained in detail in the next section) and used as a predictor to the next frame's Region of Interest (ROI) location (i.e. regions which are left after completion of the segmentation process). This way, the following frames do not need the hand palm detection step. This method is used recursively, until the algorithm fails to detect connected elements in the tracking area, or has reached the last frame. If the ROI was not found in previous frames, the method will be implemented backwards recursively, until the algorithm fails, or succeeds to find the palms in all missing frames.

3) Palm Center Determination

First a binary segmentation inside ROI is performed to find the palm inner contour, then the center of the palm is defined as the center of the biggest circle that fits inside the contour, which is a variation of an idea first described in [12]. The procedure of finding the biggest circle is implemented by a "growing from inside" procedure.

4) Tracking Algorithm Termination Terms

In case of a wrong identification of a palm center (for example due to confusion with other body parts), the tracking must break and a correct ROI must be re-found. Most of the wrong identification cases can be identified by the following termination steps:

Solidity (S): A measure that can be calculated on defined image region. Solidity is defined as the portion of the pixels in the convex hull that are also in the region. It can be calculated as the ratio between region area, and the convex area. While the head, knees and torso are characterized by high (bigger than 0.92) values of Solidity, the palm is usually characterized with smaller values (0.7 when opened to 0.95 when closed).

Normalized Area (NA): Computed by multiplying the number of pixels the object consists by the squared distance of this object from the sensor.

The advantage of this factor is that it can be considered as distance invariant. Intuitively, a palm is smaller than other body organs, and when it is closed it is significantly smaller (as can be seen in Fig. 2).

Solidity and Normalized Area (SNA): A Combination of the two above factors: if an object is both big and has high solidity, then it is probably not a palm, as can be seen in Fig. 2.

Stability (St): Can be defined as measure of change of the *SNA* components and center of palm. In case of a static position, these components do not change their values, which may indicate a misidentification of the palm and a correction of previous frames can be issued.

If one of the Termination Terms is satisfied, the Tracking recursion will be terminated, and Palm Detection Algorithm will be called again.

C. Algorithm Quality Assurance

In order to validate the results of the palm center detection and tracking algorithms, a comparison to manually marked frames from several positions, recordings and examinees was performed. The success rate of the algorithm stands on 88%, while most of the errors are from false negative group (9%), i.e. frames with hand palm that the algorithm decided to omit. The false negative group is not expected to influence the classification since they will be ignored at features extraction phase.

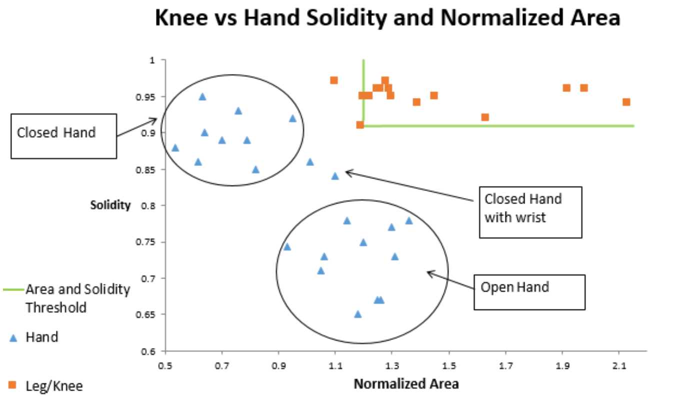

Fig. 2. Solidity parameter versus the normalized area. The triangles represents the open and close hand states, and the square dots - those of leg or knee. It can be seen clearly that by using the combination of those two parameters a tracking mistake between patient's knee and hand can be identified.

D. Feature Extraction

Each motion test performed by an examinee was divided into time frames of 4 seconds each, with 30% overlap. From each time frame a set of spatial and temporal features was then extracted separately for each hand.

The following features were extracted and saved as the features vector from every time frame:

1) *Average Absolute Speed and Variance*: computed separately for each of the three (X, Y, Z) axes and for the total speed.
2) *Average Absolute Acceleration and Variance*: computed separately for each of the three (X, Y, Z) axes and for the total acceleration.
3) *Average Absolute Speed above Threshold*: This feature differs from the first feature by neglecting frames where the palm is static.
4) *Average of 5 Maximal Absolute Speeds*: computed for the three axes (X, Y, Z). The reason for averaging 5 maximal speeds instead of just picking maximal speed is to reduce singular values effect.
5) *Average of 5 Maximal Absolute Accelerations*: computed separately for each of the axes (X, Y, Z).
6) *Main Motion Frequency and its Amplitude*: The frequency with the biggest amplitude under FFT (Fast Fourier Transform). Computed separately for each of the axes (X, Y, Z). This feature gives a measure of how distinct the main frequency is.
7) *Number of Typical Frequencies*: for each axis of the coordinates' FFT with an amplitude of more than 70% from the maximal amplitude.

E. Training the Support Vector Machine

For the classification between the Healthy and PD patients, a Support Vector Machine (SVM) with Radial Basis Function (RBF) as its kernel, that is implemented in libsvm package [13] is used. The RBF kernel requires optimization of two real valued parameters, namely *C*, and γ. While the γ parameter defines how far the influence of a single training example reaches, the *C* parameter trades off the misclassification of training examples against the simplicity of the decision surface.

First, all the feature vectors were scaled to values between zero and one. Then, in order to find the optimal *C* and γ parameters for the classification, a grid search procedure is carried out in two stages. In the first stage, an exponential scale is used for the parameters, and in the second stage, around the parameter-values that led to the best classification, a finer exponential scale was used to refine the results. To generalize the classification results and increase the reliability of the system, a cross validation technique is applied for each *C* and γ selection using the *Leave-one-Out* methodology. In addition, in order to avoid an imbalanced training data set, the smaller group in the training data set is oversampled till both groups are equally sized. A pseudo code of the method appears in Fig. 3.

1. Normalize the data set
2. For Each *C*, γ*:*
   - 2.1. Cross Validate using leave-one-out.
     - 2.1.1. Train and test the SVM.
     - 2.1.2. Store the success rate.
   - 2.2. Compute the average success rate.
   - 2.3. Update the best *C* and γ if needed.
   - 2.4. Return to 2.1 with next *C*, γ.
3. Choose *C*, γ with best average success rate, and perform step (2) using fine scale around the selected parameters.

Fig. 3. Pseudo code for the SVM training procedure

F. Decision Making

Because the classification stage is performed on four seconds time frames and separately for each hand, an addition of a high level decision making process is needed. Thus, the final classification is performed in two steps, at first each hand performance is classified using a majority vote over the time frames. Then, for the final classification, the worst-case decision is chosen (i.e., if at least one hand action was classified as PD, the examinee will be classified as having a PD).

The full proposed scheme is summarized in Fig. 4 below.

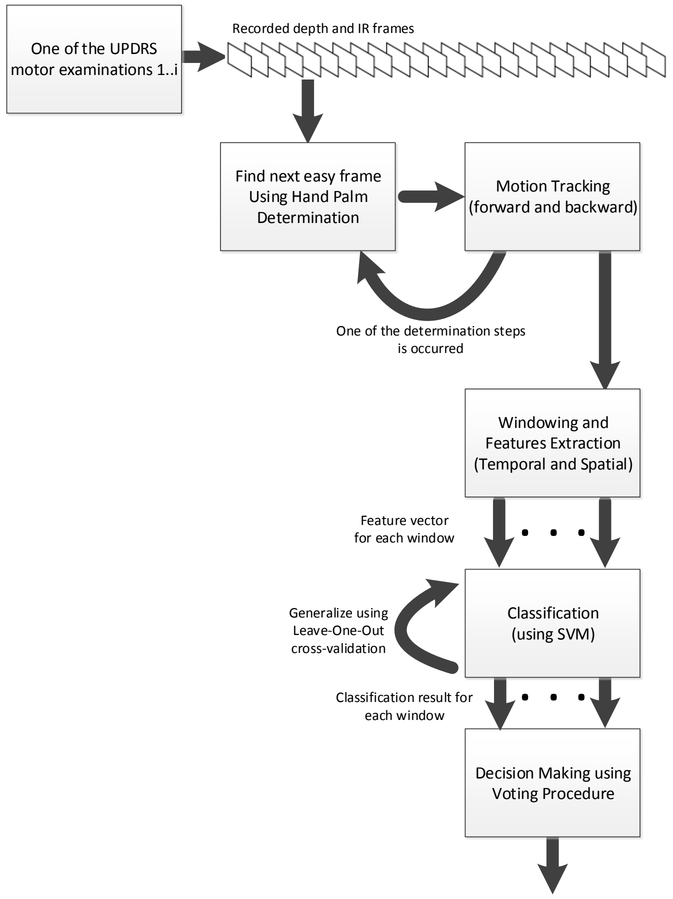

Fig. 4. Block diagram of the proposed scheme. Starting from 3D image acquisition through Hand Palm determination, motion tracking, followed by feature extraction, Classification and the Decision making processes.

G. Post-Processing Stage

As a Post-Processing stage, a method for analysis and exploring most significant features was implemented. The motivation for this stage is double: first, to improve the classification by removing unimportant features, and second, to provide some feedback to the physician about the most important features.

The features in the feature-vectors described earlier were first divided into smaller groups, set by their intuitive relation (i.e., velocity components, variances, etc.) Then, the Machine Learning phase was performed on all permutations of the groups. The combination of parameters with best classification results was then selected.

III. Results

For the evaluation of the proposed system, a total of 13 patients were used (8 having PD and 5 Healthy), all the patients (Healthy and PD) were examined by a physician before recordings. The results for the four UPDRS tests are summarized in the following confusion matrices in terms of binary classification. Table I and III presents results of the Hand Movement test using all 32 features and only the best subset, respectively. Table II and IV represents the first condition of the Rapid Alternation of Hands test, in the same manner. Table V represents the classification results of finger taps test. Interestingly, it can be seen that the classification using smaller subsets of feature vectors has better results.

TABLE I. Confusion Matrix for Hand Movement Test using all features

	Healthy	PD
Healthy	**80%**	20%
PD	20%	**80%**

TABLE II. Confusion Matrix for Rapid Alternating Hands Test – first condition using all features

	Healthy	PD
Healthy	**75%**	25%
PD	0%	**100%**

TABLE III. Confusion Matrix for Hand Movement Test using only best predictors

	Healthy	PD
Healthy	**80%**	20%
PD	0%	**100%**

TABLE IV. Confusion Matrix for Rapid Alternating Hands Test - first condition using only best predictors

	Healthy	PD
Healthy	**100%**	0%
PD	0%	**100%**

TABLE V. Confusion Matrix for Finger Taps Test

	Healthy	PD
Healthy	**100%**	0%
PD	0%	**100%**

IV. Discussion

We present a new diagnostic tool that has the ability to monitor PD patients as well as extending to other monitoring applications providing objective and reproducible quantitative outputs. Our approach is non-invasive and can be applied at home settings, possibly even over the web. Extension of the tool can be seen to take over the position of the UPDRS scales used today, which are subjective measures. Beside the application itself, this method shows much promise for general machine perception of human conditions. As a further point, the general setup of this work seems appropriate for application to a wide range of neurological diseases and states.

References

[1] C. B. Levine, K. R. Fahrbach, A. D. Siderowf, R. P. Estok, V. M. Ludensky, and S. D. Ross, “Diagnosis and treatment of Parkinson’s disease: a systematic review of the literature,” *Evid. Rep. Technol. Assess. (Summ.)*, no. 57, pp. 1–4, May 2003.

[2] D. M. Huse, K. Schulman, L. Orsini, J. Castelli-Haley, S. Kennedy, and G. Lenhart, “Burden of illness in Parkinson’s disease,” *Mov. Disord.*, vol. 20, no. 11, pp. 1449–1454, 2005.

[3] S. Fahn, R. Elton, and UPDRS Development Committee, “Unified Parkinson’s disease rating scale,” *Recent Dev. Park. Dis.*, vol. 2, pp. 153–163, 1987.

[4] R. LeMoyne, T. Mastroianni, M. Cozza, C. Coroian, and W. Grundfest, “Implementation of an iPhone for characterizing Parkinson’s disease tremor through a wireless accelerometer application,” in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2010, pp. 4954–4958.

[5] J. D. Hoffman and J. McNames, “Objective measure of upper extremity motor impairment in Parkinson’s disease with inertial sensors,” in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society,EMBC*, 2011, pp. 4378–4381.

[6] A. Kupryjanow, B. Kunka, and B. Kostek, “UPDRS Tests for Diagnosis of Parkinson’s Disease Employing Virtual-Touchpad,” in *2010 Workshop on Database and Expert Systems Applications (DEXA)*, 2010, pp. 132–136.

[7] T. E. Eichhorn, T. Gasser, N. Mai, C. Marquardt, G. Arnold, J. Schwarz, and W. H. Oertel, “Computational analysis of open loop handwriting movements in Parkinson’s disease: A rapid method to detect dopamimetic effects,” *Mov. Disord.*, vol. 11, no. 3, pp. 289–297, 1996.

[8] J. A. Christensen, R. Frandsen, J. Kempfner, L. Arvastson, S. R. Christensen, P. Jennum, and H. B. D. Sorensen, “Separation of Parkinson’s patients in early and mature stages from control subjects using one EOG channel,” in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2012, pp. 2941–2944.

[9] T. Vaserman, T. Shraga, A. Frid, H. Hel-Or, S. Raz, and I. Schlesinger, “Automatic Quantitative Assessment of Hand Movements and Finger Taps in Parkinson’s Disease,” presented at the The 12th Bar-Ilan Symposium on the Foundations of Artificial Intelligence (BISFAI), Bar-Ilan University, Ramat Gan, Israel, 2013.

[10] C.-W. Cho, W.-H. Chao, S.-H. Lin, and Y.-Y. Chen, “A vision-based analysis system for gait recognition in patients with Parkinson’s disease,” *Expert Syst. Appl.*, vol. 36, no. 3, Part 2, pp. 7033–7039, Apr. 2009.

[11] X. Liu and K. Fujimura, “Hand gesture recognition using depth data,” in *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings*, 2004, pp. 529–534.

[12] F. T. Cerezo, “3D Hand and Finger Recognition using Kinect,” Universidad de Granada (UGR), Spain, unpublished, 2011.

[13] C.-C. Chang and C.-J. Lin, “LIBSVM: A library for support vector machines,” *ACM Trans Intell Syst Technol*, vol. 2, no. 3, pp. 27:1–27:27, May 2011.