Proceedings of the 2021 Design of Medical Devices Conference
DMD2021
April 12-15, 2021, Minneapolis, MN, USA

DMD2021-1067

QUANTIFYING POSTURAL INSTABILITY WITH POSE ESTIMATION SOFTWARE AND 3D DEPTH EXTRACTION

Cara Piazza BS
University of Minnesota
Minneapolis, MN

Joseph Schroeder BS
University of Minnesota
Minneapolis, MN

Chiahao Lu PhD
University of Minnesota
Minneapolis, MN

Arthur Erdman PhD
University of Minnesota
Minneapolis, MN

Matthew Johnson PhD
University of Minnesota
Minneapolis, MN

Scott E. Cooper MD PhD
University of Minnesota
Minneapolis, MN

ABSTRACT

Patients who suffer from Parkinson's Disease are more prone to postural instability, a major risk factor for falls. One of the most common clinical methods of gauging the severity of a patient's postural instability is with the retropulsion test [1], in which a clinician perturbs the balance of the patient and then rates their response to the perturbation. This test is subjective and largely based on the observations made by the clinician. In order to improve postural instability diagnosis and encourage more meaningful therapies for this cognitive-motor symptom, there is a clinical need to enable more objective, quantifiable approaches to measuring postural instability. In this paper, we describe a novel computational approach to quantifying the number, length, and trajectory of steps taken during a retropulsion test or other type of balance perturbation from a single camera facing the anterior side (front) of the subject. The computational framework involved first analyzing the video data using markerless pose estimation algorithms to track the movement of the subject's feet. These pixel data were then converted from 2D to 3D using calibrated transformation functions, and then analyzed for consistency when compared to the known step lengths. The testing data showed accurate step length estimation within 1 cm, which suggests this computational approach could have utility in a variety of clinical environments.

1. INTRODUCTION

One of the primary risk factors for falls in the United States is postural instability, which involves the loss of balance and postural reflexes. While loss of dynamic stability is common with age, disease such as Parkinson's disease (PD) greatly exacerbate the process. [2]. Over half off the estimated 930,000 people with PD in the United States fall with a high rate of reoccurrence [3] [4]. Engineering-based approaches to track fluctuations and progression of postural instability over time are notably lacking, which limits opportunities to develop intervening, correctional technologies.

Defining a subject's posture has posed many challenges due to its dynamic characteristics. Unlike basic mechanical systems, the biological system's intrinsic complexity is not easily defined by simple parameters. Therefore, it is difficult to provide an accurate, real-time diagnosis of postural instability. Consequently, current assessments of Parkinsonian postural instability are subjective. The most common assessment used, the retropulsion test, provides a single snapshot of posture but is highly qualitative and lacks repeatability and consistency. Also, it cannot be conducted without the presence of a trained physician.

In research settings, large and expensive equipment can be used to quantify postural instability by assessing joint angles and torques, body center of mass and muscle activation patterns. But, the high dimensionality of this data is time-consuming and computationally expensive to analyze. Due to the subjective nature of the current diagnosis methods for postural instability, there exists a need to reduce the complexity and increase the accessibility of methods of postural analysis.

This research was motivated by the need to provide a simple yet objective way to quantify postural responses. Often, a postural response results in a correction step (or steps) taken by the subject to regain balance. The main purpose of this study was to analyze postural response by extracting the step distance after a perturbation using a single video camera. To do this, postural movements were recorded using a single camera, and machine learning, optimization, and inverse mapping methods were used to quantify the responses by extracting step length. The goal, therefore, was to create a repeatable method in which postural movements, specifically backwards steps, could be quantified with minimal equipment making the approach more accessible within a broad range of clinical environments.

2. MATERIALS AND METHODS

For the experiment, a subject stood on a floor surface with known distances marked by lines of tape. The video camera was oriented to view the subject's frontal plane, and recorded a subject as they stepped backwards, in order to mimic a postural response to a perturbation. The subject's foot began at the starting tape and stepped backwards to the next tape (the ending tape). For each step it was indicated whether their foot landed on or off the ending tape (n=17 right foot; n=18 left foot). Video recordings were collected during the sessions. As a second validation, a subject took steps of varying lengths and a researcher measured each step the subject took using a tape measure. Video of this test was then analyzed in DLC and the video data was brought to 3D in MATLAB using the same methodology as with the previous validation.

From the video data, a machine learning markerless pose estimation software, DeepLabCut, was used to track and return the toe locations of the subject over time. The general workflow for this process can be seen in Figure 1. The protocol for the labeling was to mark the ankle and toes of both the subject's feet as well as three points which were used as known distances when calibrating. The toe pixel locations were mapped to 3D in order to extract the step length of the postural response. The toe time and 3D locations at the initiation and completion of the step were recorded.

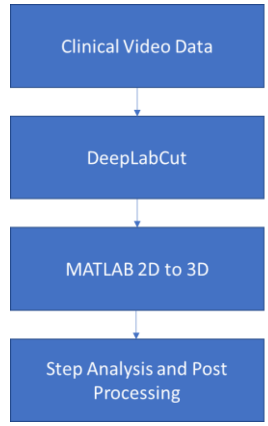

FIGURE 1: THIS FIGURE SHOWS THE PROCESS FLOW FOR TAKING THE RAW VIDEO TO 3D DATA AND FINALLY POST PROCESSING.

In order to take inversely map the 2D pixel data and to 3D, the pseudoinverse of the forward mapping matrix was developed as described below. The forward mapping matrix is a 3x4 matrix which can take 3D data and convert it into 2D pixel locations. The matrix consists of the camera's intrinsic properties (the intrinsic matrix), and the cameras location relative to the world origin (extrinsic matrix). Mapping from 3D to 2D provided a single, unique solution (Eq. 1). In Eq. 1, f_x and f_y indicate the focal lengths in the horizontal and vertical positions, and o_x and o_y indicate the pixel offset in the horizontal and vertical positions. The r and t values represent the rotation and translation of the camera with respect to the world origin. Eq. 2 shows the product of the intrinsic and extrinsic matrices solved with the resulting indices marked m_{11} through m_{34}. Since the forward mapping matrix is not square, it cannot be inverted in a classical manner, and therefore the inverse of this problem (i.e. inverse mapping from 2D to 3D) is much more complex. In order to overcome this barrier, the pseudoinverse of the mapping matrix was used (Eq. 3). The pseudoinverse is computed using singular value decomposition and computes a lest squares (best fit) solution to a system of linear equations that lacks a solution. The best fit solution for this scenario confined all solutions to the surface on which the subject was standing.

$$\begin{bmatrix} f_x & 0 & o_x & 0 \\ 0 & f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} y' \\ x' \\ z' \end{bmatrix} \sim \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} y' \\ x' \\ z' \end{bmatrix} \sim \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

$$pseudoinverse\left(\begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix}\right)\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3)$$

To properly calibrate the step extraction from 2D to 3D, known distances were used to determine the intrinsic and extrinsic matrices. These distances included one length in the mediolateral direction and one length in the anteroposterior direction. Combined, they allowed for a plane of the surface on which the subject was standing to be properly configured to calculate the 3D coordinates as shown in Eq. 3. This calculation was required for every additional trial measured to validate that the pseudoinverse was accurate to the recording taken. Small variations like camera zoom, camera angle and other factors can impact the transformation and throw off the results if not properly calibrated. By calibrating with these measurements, the L2 norm is minimized, and the pseudoinverse will force a unique solution. The camera used in this experiment was a Sony Model FDR-AX53 recording at 60fps for all trials and was fixed during recording to provide a consistent angle of measurement.

3. RESULTS AND DISCUSSION

All right and left backwards steps were analyzed. The raw data from these trials is shown in Fig. 1 for the anteroposterior plane, which is defined as the plane from the front to back of a subject. This data is directly taken from DeepLabCut and can be further processed based on the likelihood of each measurement provided by DeepLabCut. The steps were recorded, and the ending location of the foot was compared with the known location of the ending tape. The error between the known ending distance and measured ending distance was recorded.

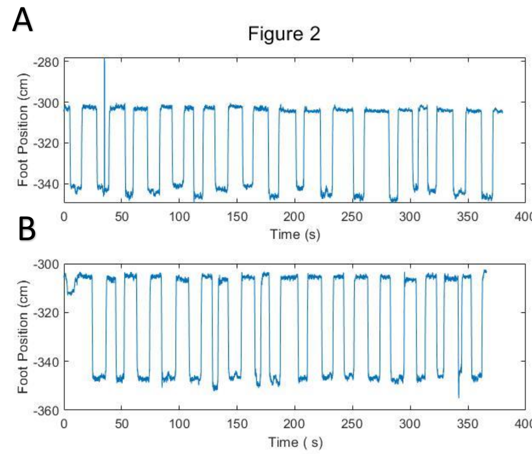

FIGURE 2: SHOWS THE RAW STEP DATA FROM DEEPLABCUT FOR THE RIGHT (A) AND LEFT (B) FOOT ANTEROPOSTERIOR STEPS. THE STEPS WERE SLIGHTLY VARIED IN LENGTH INTENTIONALLY WITH SOME STEPS NOT BEING TO THE MARKED LINE WHILE OTHERS WERE.

The results for step distances can be observed in Fig. 3 which displays the absolute step distance determined through our method. The starting location of the foot was determined by the plateau's average taken from the data shown in Fig. 2. The ending location of the foot was determined by the average of the following plateau. Any points that had a location probability less than 70% were neglected.

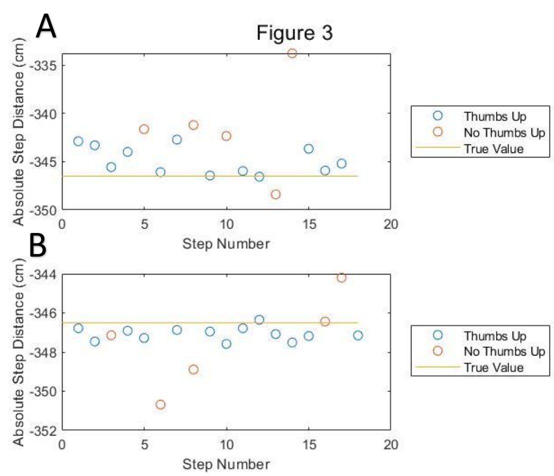

FIGURE 3: (A) SHOWS THE RIGHT FOOT ABSOLUTE STEP DATA. (B) SHOWS THE LEFT FOOT ABSOLUTE STEP DATA.

The data shown in Fig. 3 shows that the ending location of the steps varied for both trials with and without thumbs up. The first steps taken were those of the right foot (Fig. 3 A) and after that trial the subject took the steps with the left foot, (Fig. 3 B). The subject's steps became more consistent with the left foot. This can be seen as more steps align closely with the true value when a thumbs up is given, and the lack of a thumbs up align with steps that deviate further from the true value. The greater deviation seen in the right foot can be attributed to a lack of subject experience, rather than the method in which the distances were extracted. The accuracy of the left foot thumbs up steps can be attributed to subject learning. This shows that this method of data collection is a valid method within a margin of error of nearly one centimeter.

The data collected from the second validation, in which steps of varying length were taken, is shown below in Fig. 4. This data was also processed by removing the points of low confidence and then determining the measured distances.

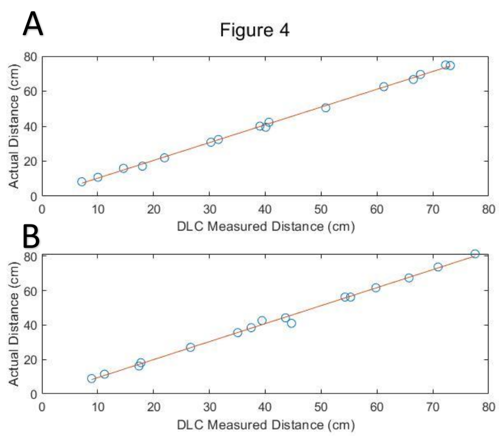

FIGURE 4: (A) SHOWS THE RIGHT FOOT DLC MEASURED DISTANCES VS THE ACTUAL DISTANCE ALONGSIDE THE LINEAR REGRESSION. (B) SHOWS THE LEFT FOOT DLC MEASURED DISTANCES VS THE ACTUAL DISTANCE ALONGSIDE THE LINEAR REGRESSION.

This second validation yielded more interesting results with the DLC measured step length consistently matching that of the actual step length in the trial. For both the right foot (Fig. 4 A) and the left foot (Fig. 4 B) the DLC value always fell within a close range with the max residual in the right foot at 2.654 cm and the max residual in the left foot at 3.752 cm. The linear regressions for the two feet showed the accuracy of the DLC measured data with the right foot being represented by y = 1.0159*x + 0.1316 with an R^2 value of 0.9986 and the left foot being represented by y = 1.0443*x – 0.9868 with an R^2 value of 0.9956.

There are various drawbacks to this current data analysis method. One drawback is due to the consistent identification of points on the subject's body during markerless pose estimation. Particularly darker clothes provide poor contrast, which leaves the researcher left to estimate the best approximation of a location on the subject's body. This can be mitigated by increasing the amount of labeling done to reduce noise, but this increase in time and computational effort is counter to the goal of this method to simplify the process of data analysis. However, upon labeling enough points, the data is cleaner and the markerless pose estimation can better approximate the subject's movement.

Despite this noise, there are many benefits to the markerless pose estimation method. DeepLabCut is itself highly customizable and can accommodate a range or points and scenarios. Unlike in a marker-based pose estimation scenario where the subject must where an identifiable marker, any video with sufficient quality can be used to accurately apply markerless pose estimation. This also enables researchers to retroactively use video taken for other studies and effectively track aspects of motion.

The next place to take this research is to further automate the data analysis to make this process easier and faster for the researcher. One area of automation that is important for the accuracy of the model is the optimization of the intrinsic and extrinsic matrices' variables. This process will be automated to allow for calibration to be automatic from test to test and then the step extraction can be fully automated once the raw data is exported from DeepLabCut. Additionally, this automation could enable the ability to change known distances from case to case and calibrate regardless of the scenario.

4. CONCLUSION

Throughout this data analysis, it has been shown that 3D step information can be extracted from a single video by using markerless pose estimation software. The transformation from 2D to 3D requires some assumptions be made but allows for an accurate approximation to be made on a known plane by using the pseudoinverse. This simple application, which utilizes a single camera, can also be applied to a home health setting. This would allow for evaluation of postural instability on a more consistent basis, which can allow for better patient insight and next steps in care. In the case of Parkinsonian patients this could help better inform physicians on the progression of their patients Parkinson’s with respect to postural instability. Upon proper calibration, this method has been shown to be accurate on its approximation of the 3d space using known distances, but further automation is needed to provide real time step length analysis.

While the labeling aspect of markerless pose estimation remains a time consuming and resource intensive step, it is possible the overall workflow of this process can further be optimized to allow for a faster video to data process. This includes automation in the determination of extrinsic and intrinsic parameters, automation in the step extraction and additionally automation in the post processing of the data. A major limitation to this method as opposed to the current standard of care is that these technologies are not real time and therefore the researcher would have to wait for the results upwards of days.

With the promising start to using this technology for pose estimation, it remains to be seen how this methodology will handle pathological pose data. The appeal of using this methodology for retrospective data collection and analysis is large and with a low barrier to entry it could prove to be a beneficial format for other researchers to incorporate into their toolbox. Additionally, with automation this methodology could prove to be useful in a clinical and diagnostic setting allowing a quantitative analysis of posture and gait.

ACKNOWLEDGEMENTS

We would like to acknowledge Dr. Scott Cooper, Dr. Matthew Johnson and Dr. Art Erdman for their support and interdisciplinary collaboration to make this research possible.

REFERENCES

[1] Nonnekes, J., Goselink, R., Weerdesteyn, V., & Bloem, B. R. (2015). The Retropulsion Test: A Good Evaluation of Postural Instability in Parkinson's Disease? Journal of Parkinson's Disease, 5(1), 43-47. doi:10.3233/jpd-140514

[2] Semyon M. Slobounov, Sarah A. Moss, Elena S. Slobounova, Karl M. Newell, Aging and Time to Instability in Posture, The Journals of Gerontology: Series A, Volume 53A, Issue 1, January 1998, Pages B71–B80, https://doi.org/10.1093/gerona/53A.1.B71

[3] Marras, C., Beck, J. C., Bower, J. H., Roberts, E., Ritz, B., Ross, G. W., . . . Tanner, C. (2018). Prevalence of Parkinson’s disease across North America. Npj Parkinson's Disease, 4(1). doi:10.1038/s41531-018-0058-0

[4] Pelicioni, P.H.S.; Menant, J.C.; Latt, M.D.; Lord, S.R. Falls in Parkinson’s Disease Subtypes: Risk Factors, Locations and Circumstances. Int. J. Environ. Res. Public Health 2019, 16, 2216.