33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Video Assessment of Finger Tapping for Parkinson's Disease and Other Movement Disorders

Kjersten Criss and James McNames

Abstract— **Functional motor impairment due to Parkinson's disease and other movement disorders are currently assessed with visual rating scales such as the Unified Parkinson's Disease Rating Scale (UPDRS). These methods rely on the subjective judgment of a rater to assign scores representing the extent of impairment while subjects perform prescribed activities. We describe a new model-based framework that uses statistical video processing to automatically track movement during prescribed activities. This approach has many advantages over traditional clinical rating scales. It can completely characterize movement during prescribed tasks over time objectively and precisely using hardware that is inexpensive and readily available. We demonstrate the potential of this framework with a simple statistical model applied to a paced finger tapping test. This technology could be deployed in a natural home environment for frequent assessments. This technology could ultimately improve both clinical practice and clinical trials.**

I. Introduction

A precise and accurate assessment of movement disorders is important for both clinical practice and clinical trials evaluating new therapies. In both cases, the impact of therapeutic changes is important. This is a challenging problem because for many movement disorders there is no directly biological measure, or biomarker, of the state of the disease or severity of functional impairment. In practice, the functional motor impairment is assessed with a visual exam by a neurologist or trained rater. In clinical trials, rating scales such as the Unified Parkinson's Disease Rating Scale (UPDRS) are used [1]. These scales require the subjects to complete a series of prescribed tasks which are observed by a rater. The rater then assigns a score to each task, often on a 4 or 5 point scale.

We propose a new framework for objectively assessing movement disorders based on a model-based statistical video processing algorithm. This type of objective measure has many potential advantages over a visual assessment by a rater. By eliminating the subjective judgment of the rater, this approach may have greater precision, less variability because it eliminates inter-rater and intra-rater variability, and more accuracy because it eliminates rater bias. It may also eliminate the cost normally required for raters to be trained, certified, and rate subjects. Increasingly, technologies are being deployed for home assessment, which can reduce cost and permit more frequent measurements which can be used to more precisely track the progression of a disease than infrequent visits to a clinical site required with rating scales. This increase in accuracy and measurement frequency could have a large impact on clinical trials and lead to fewer subjects, shorter trials, and substantial cost savings. This is potentially a low-cost approach because it does not require any specialized hardware and leverages the economy of scale that has driven down the cost of high-resolution web cameras.

This paper describes a method for assessing one of these tasks, finger tapping, with a model-based statistical video processing algorithm. We demonstrate the potential of this framework with a finger tapping test. This is an early and sensitive indicator of motor impairment caused by Parkinson's disease. This approach is novel because it is based on a statistical model that incorporates domain knowledge in the form of biomechanics and biodynamics, which we define as knowledge about physiology of movement.

Another video based effort to evaluate tapping tests utilizes the Virtual-Touchpad interface (VTP) [2]. The VTP is a mobile interface that instructs patients to perform the test, and includes a webcam to record the test, which can be evaluated in real time for a 15 fps video. The video is evaluated using contour algorithms to determine the type of gesture being made by the subject's hand (e.g. are the index finger and thumb apart or touching). The speed of tapping can be determined from this information. To score this type of test raters are asked to consider multiple factors. The VTP only considers the speed of tapping, while our method is able to address all by using a model based approach.

II. Algorithm Design

Finger tapping is one of the most sensitive tests in the 27-part motor UPDRS exam, which is often used as the primary endpoint in clinical trials. To determine the score assigned to this test, the rater is instructed to consider the tapping speed, tapping amplitude, number of hesitations and halts, and decrease in amplitude over the duration of the test. Our video tapping assessment is designed for the same test, but precisely quantifies the angles of each joint of the finger over the duration of the test. This permits us to precisely quantify all of the characteristics listed above.

The algorithm is divided into four components: the physical recording and camera settings, a mathematical model that relates the joint angles of interest to a video frame, a criterion

Kjersten Criss is a member of the Biomedical Signal Processing Laboratory in the Department of Electrical and Computer Engineering at Portland State University, Portland, Oregon, USA. Email: kjersten@pdx.edu (corresponding author).

James McNames is director of the Biomedical Signal Processing Laboratory. He is also professor and chair of the Department of Electrical and Computer Engineering at Portland State University, Portland, Oregon, USA. Email: mcnames@pdx.edu.

This work was partly supported by an undergraduate research grant from Intel and the Semiconductor Research Corporation education alliance.

for comparing the similarity of the model to each captured frame, and an optimization algorithm for minimizing the optimization criterion. Each of these components is described in detail in the following sections.

A. *Video Camera and Environmental Settings*

The most important characteristics of the video camera and environmental settings are sharpness of the image, frame rate, color difference between the hand and background, and masking of the fingers that were not of interest. During finger tapping subjects are typically instructed to tap with as much speed and amplitude as possible. For video processing it is important that the frame rate is high and blurring is minimized. This can be attained with good lighting to minimize the exposure time for each frame. To accurately compare the dichromatic hand model to an actual image of a hand, it is best for the average RGB color values of the hand and the background to be as different as possible. To accomplish the goals of having a crisp image and a large color difference, we used a black background that reflected little light and we illuminated the hand with white LED lights placed behind the camera. When it is not possible to constrain the background and lighting to produce the necessary difference in RGB color values, the video can be processed in HSV color space as long as there is a significant difference between the hue or saturation values of the hand and the background. The videos were recorded with a low enough exposure time so that little to no blurring could be seen in a video of a 1 Hz passed tapping. A small filming stage was created to provide a black background. The stage included a section to mask the other three fingers. In this study we used a Logitech (Fremont, CA) HD Pro Webcam C910. The videos had a resolution of 1280×720, and the frame rate was 30 fps.

B. *Model Design*

Our initial model consists of interconnected polygons to represent the relevant portions of the finger, thumb, and other parts of the hand. Three separate interconnected polygons were used to represent each phalanx of the index finger. Two additional polygons are used to represent the thumb, and a final polygon represents the remainder of the hand. The coordinates of each polygon forming the index finger can be rotated using rotational matrices and translations to model finger movement. In this example, the other polygons are static since the hand and thumb are fixed for the duration of the test. Thus, the state of the system and position of the polygons is completely determined by the three angles for each joint of the index finger, which we denote as $\theta_{1,n}$, $\theta_{2,n}$, and $\theta_{3,n}$ in order of nearest to farthest joints from the hand for the n^{th} frame. Figure 2 shows the joint angles and indicates the direction of positive rotation. Collectively we represent these as a vector

$$\theta_n = \begin{bmatrix}\theta_{1,n} & \theta_{2,n} & \theta_{3,n}\end{bmatrix}^{\mathrm{T}} \tag{1}$$

where $^{\mathrm{T}}$ represents the transpose operation. Figure 1 shows an example of the modeled image with the joint angles set to $\theta_n = 0°$.

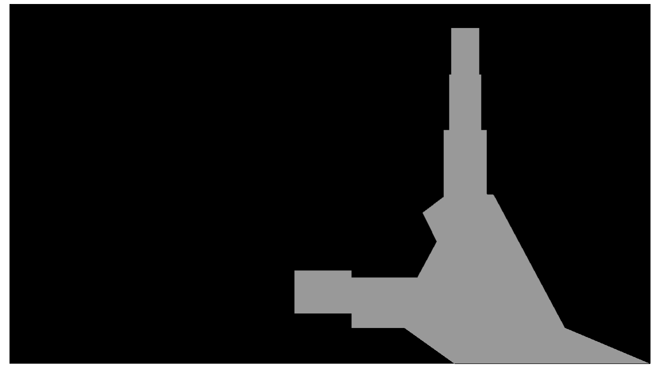


Fig. 1. This figure shows the model with the knuckle, proximal, and distal joint angles set to zero.

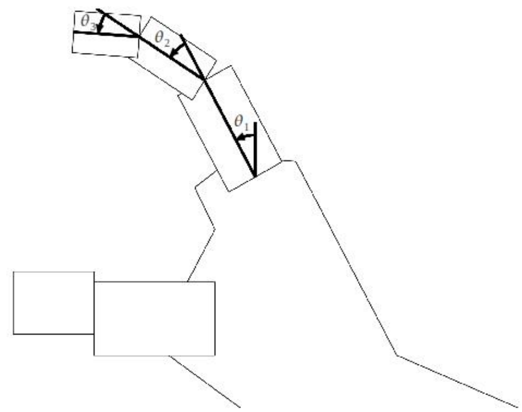


Fig. 2. This figure shows the model and demonstrates how the joint angles are defined.

For a given set of joint angles, the locations of the polygons is known. We have developed a fast, but straightforward algorithm, to convert the polygons into a matrix representing the pixelated RGB image with the same dimensions as the original video frames. For the examples shown in this paper the polygons are given a fixed color that was manually matched to the skin color of the video. The background was similarly selected in a manual fashion. However, it is possible to make the skin and background colors parameters that are concurrently estimated with the joint angles.

C. *Optimization Criterion*

For the sake of simplicity we used a statistical model of each frame as our polygon model with additive Gaussian noise,

$$f_n(i,j,k) = p_n(i,j,k;\theta_n) + e_n(i,j,k) \tag{2}$$

where $i \in \{1,2,3\}$ represents the RGB color index, j represents the row index, k represents the column index, n represents the frame index, $f_n((i,j,k;\theta_n)$ is the nth frame of video, $p_n(i,j,k;\theta_n)$ is the nth model frame, and $e_n(i,j,k)$ is the additive Gaussian noise with zero mean and unknown variance. We assume the noise between colors and pixels is independent and identically distributed.

Under these conditions, the maximum likelihood estimator is attained by minimizing the mean squared error (MSE),

$$\zeta(\theta_n) = \frac{1}{n_p} \sum_{i,j,k} \left[f_n(i,j,k) - p_n(i,j,k;\theta_n) \right]^2 \qquad (3)$$

where n_p is the total number of pixel RGB values in a single frame. If the RGB values of the hand and the background do not differ greatly, the MSE between the HSV values of the pixels can be used.

D. Optimization Algorithm

There are a wide variety of optimization algorithms that could be used to minimize the MSE. For this example, we chose to use the cyclic coordinate method [3]. This is a simple method that optimizes each angle individually with all of the others fixed. Once all of the angles are optimized, the method goes back to the first angle and repeats. Thus, the cyclic coordinate method essentially consists of a series of one dimensional optimizations that can be efficiently solved with a variety of line search algorithms. We chose to implement a Golden Section method line search to efficiently locate the minimum [3]. For the example in the next section, we stopped the line search when the difference between the upper and lower bounds of estimated angles was 0.1 radians ($5.73°$) and we stopped the cyclic coordinate method after four iterations.

Figure 4 shows three examples of the MSE versus the joint angles. These examples clearly illustrate that local minima are present, though at values near the global minima the error functions are locally convex. The Golden Section method searches for local minimum, but the local minimum in the error plots correspond to angles that cause a phalanx to completely overlap the hand or another phalanx, which is physically infeasible. The allowable angles for the model were limited to physically possible positions, eliminating the risk of erroneous solutions caused by local minimum.

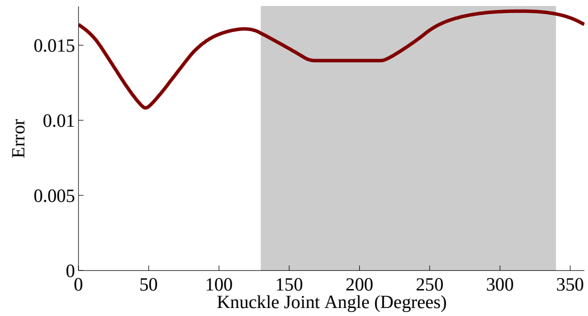
At 30 fps we do not expect the finger to change position substantially from one frame to the next. Thus, for each frame except for the first, we initialize the algorithm with the finger location estimated from the previous frame. For the first frame of video the proximal and distal joints angles were set to $180°$. For subsequent frames the solution was constrained to the following range of angles estimated in the previous frame

$$\theta_\ell(n-1) - 30° \leq \theta_\ell(n) \leq \theta_\ell(n) + 30° \qquad (4)$$

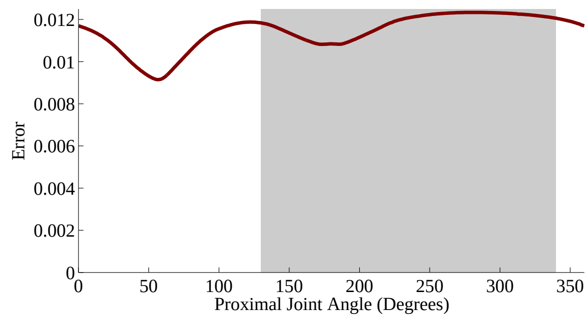
This knowledge of the finger dynamics could eventually be incorporated into the state dynamics of a state space model of the finger position, which could further improve the accuracy of the estimate.

III. Example Tests

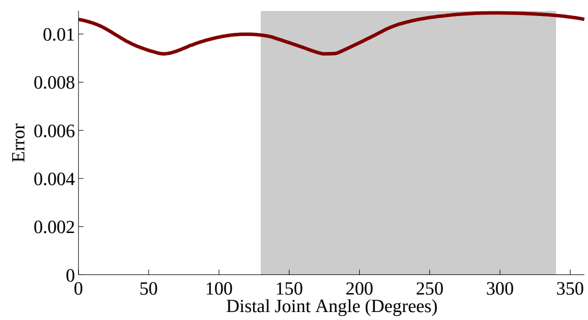
Figure 5 shows an example of the three estimated joint angles versus time during a 0.8 Hz paced tapping test. All of the information about the extent of impaired movement, including tapping rate, amplitude, hesitations, halts, and decrement in amplitude, can easily be estimated from these joint angles. There is noticeable noise in the distal joint, and to a lesser extent the proximal joint. This is to be expected because any error in the more proximal joint angles will result in misplacement of the more distal phalanges, which makes it more difficult to accurately estimate the distal joint angles.



(a) Knuckle (θ_1)



(b) Proximal (θ_2)



(c) Distal (θ_3)

Fig. 4. Example of the MSE versus the joint angles. The shaded region shows joint angles that are not physically feasible.

Figure 3 shows an example of the original frame and the modeled frame after optimization. The original frame was obtained during a paced tapping test and it illustrates that clear, crisp images could be obtained even during movement due to careful control of the recording environment. The accuracy of the estimated joint angles is immediately apparent through a visual comparison of the side-by-side images.

(a) Original Frame (b) Modeled Frame

Fig. 3. Example of an original video frame and the modeled video frame after optimization. This frame is part of a recording that was obtained during a paced tapping test.

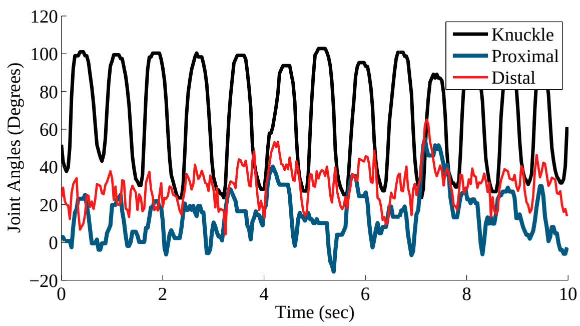


Fig. 5. Estimated joint angles during a paced tapping as a function of time.

IV. Discussion and Summary

In this paper we have introduced a new framework for the objective assessment of movement disorders. This has all of the advantages of other technologies that have been developed for this purpose that are essentially due to replacing the coarse subjective judgment of a rater with a carefully designed objective instrument [4]–[8]. However, this approach has several additional advantages. It does not require expensive or custom developed software. Rather, it simply requires a computer and web cam, both of which are low cost and widely available. It allows subjects to perform the task naturally in the same manner that was developed for movement disorders. Perhaps most importantly, the small size and ease of use makes it possible to perform the assessment in a natural home environment.

We have demonstrated this framework with a finger tapping test, which is a sensitive measure of functional motor impairment in Parkinson's disease. The model, optimization criterion, and optimization algorithm were relatively simple and could be improved in many ways. For example, a more flexible and precise model of the hand could be used. Other model parameters could be defined and jointly optimized. For example, the thumb position, hand color, background color, hand size, and hand position could be defined through additional parameters. Similarly, a more accurate noise model could be developed. A state space model could be defined to more accurately represent the dynamics of human movement and improve the estimation accuracy by using one of the state space tracking algorithms such as the unscented Kalman filter (UKF) or particle filters [9]. There are also a variety of heuristics that could be applied to reduce the computational requirements and reduce the overall processing duration. However, even without the many possible improvements, this relatively simple model was able to estimate the joint angles accurately.

References

[1] S. Fahn, R. L. Elton, and UPDRS Development Committee, "Unified Parkinsons disease rating scale," in *Recent developments in Parkinsons disease*, S. Fahn, C. D. Marsden, M. Goldstein, and D. B. Calne, Eds. Florham Park, N. J.: Macmillan Healthcare Information, 1987, vol. 2, pp. 153–163, 293–304.

[2] A. Kupryjanow, B. Kunka, and B. Kostek, "UPDRS tests for diagnosis of parkinson's disease employing virtual-touchpad," in *Database and Expert Systems Applications (DEXA), 2010 Workshop on*, 30 2010.

[3] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms.*, 2nd ed. John Wiley & Sons, Inc., 1993.

[4] S. K. Patrick, A. A. Denington, M. J. A. Gauthier, D. M. Gillard, and A. Prochazka, "Quantification of the UPDRS rigidity scale," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 9, no. 1, pp. 31–41, Mar 2001.

[5] A. S. Rao, R. E. 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*, vol. 2008, pp. 1769–1772, 2008. [Online]. Available: http://dx.doi.org/10.1109/IEMBS.2008.4649520

[6] J. M. Spyers-Ashby and M. J. Stokes, "Reliability of tremor measurements using a multidimensional electromagnetic sensor system," *Clinical Rehabilitation*, vol. 14, no. 4, pp. 425–432, Aug 2000.

[7] E. B. Montgomery, W. C. Koller, T. J. LaMantia, M. C. Newman, E. Swanson-Hyland, A. W. Kaszniak, and K. Lyons, "Early detection of probable idiopathic Parkinson's disease: I. development of a diagnostic test battery." *Movement Disorders*, vol. 15, no. 3, pp. 467–473, May 2000.

[8] E. L. Stegemöller, T. Simuni, and C. MacKinnon, "Effect of movement frequency on repetitive finger movements in patients with Parkinson's disease," *Movement Disorders*, vol. 24, no. 8, pp. 1162–1169, Jun 2009. [Online]. Available: http://dx.doi.org/10.1002/mds.22535

[9] O. Cappé, S. Godsill, and E. Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *Proceedings of the IEEE*, vol. 95, no. 5, pp. 899–924, May 2007.