# QUANTIFYING DRUG-INDUCED DYSKINESIA USING CLINICAL VIDEOS OF PARKINSON'S DISEASE PATIENTS

By

Anusha Sathyanarayanan Rao

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Electrical Engineering

May, 2012

Nashville, Tennessee

Approved

Dr. Benoit M. Dawant

Dr. Robert E. Bodenheimer

Dr. Thomas L. Davis

Dr. A. B. Bonds

Dr. Mitchell Wilkes

This dissertation is dedicated to my beloved parents, Sathyanarayanan and Meenakshi and my dearest younger sister, Viju.

## ACKNOWLEDGEMENTS

I am extremely grateful to all my committee members for having supported my work to help shape it up as a doctoral dissertation. They have provided assistance at various stages and always me encouraged to successfully complete this dissertation.

First and foremost, I would like to thank my advisor, Dr. Benoit Dawant for accepting me as his student and giving me the freedom to choose my project of interest. Over the past several years, I have observed his dedication and sincerity towards research, his enthusiastic approach to collaborative interdisciplinary work and his attention to the progress of his students. He is definitely my role model for a career in academia. His class on Advanced Image Processing was one of the best I had at Vanderbilt University with very interesting projects and his teaching style is very effective. I can say without reservations that he is the best mentor and teacher I have ever had during my higher education. Be it discussions of research problems or trying out new ideas, reviews of papers or help with difficult concepts in image and signal processing, Dr. Dawant has always given his complete attention in spite of the several projects he works on in the lab. Discussing research with him is extremely exciting as his passion for it is infectious and I always come out of a meeting, brimming with several new ideas and the enthusiasm to implement them. He would always ascertain that I made progress and was motivated to proceed further with the newer ideas. His interest in his students does not merely stop at the professional level. He is a very kind, approachable, thoughtful and helpful person and I distinctly remember several occasions when he had personally offered his help during my health crises. I am extremely grateful to him for supporting me on his grant towards the end of my dissertation and I would

have to mention that he is probably the only person I know who always responds to emails so promptly, even during weekends!

My co-advisor Dr. Bobby Bodenheimer was instrumental in supporting my preliminary work in the motion capture lab at Belmont University and continued to support me on the Vanderbilt University Discovery grant for the initial work on my dissertation. He is very kind, extremely encouraging and supportive, and has advised me on several movement analysis concepts. I learned a lot of important concepts in his Graphics and Image Processing class which I eventually experimented with in my research. Due to his encouragement and assistance, I have definitely improved my proficiency in writing and now know the significance of preparing a well-organized 'Powerpoint' presentation. His attention to details, such as grammar, vocabulary and the significance of designing simple yet effective slides have had a lasting impact on me.

Dr. Thomas Davis, who is a neurologist and the out of department committee member, is in fact one of the strongest pillars on which this dissertation stands. Being a clinical application, I required significant assistance to understand the basics of the problem, the viability of solutions and the clinical impact the work can have. He has given more help than I can possibly imagine and am always grateful to him for the same. In spite of a busy clinic schedule, he has always agreed to meet me even on short notices and on weekends to discuss bottlenecks that we have had and constantly encouraged me to apply my work in other avenues such as autism, sleep disorders, etc. He is a very kind and approachable person and even non-technical conversations with him are as stimulating as the technical ones.

Dr. A.B. Bonds is one of the first few professors I started working with when I joined Vanderbilt University, in the capacity of his teaching assistant. Knowing my interest in teaching, he has constantly encouraged and motivated me to develop my skills in this area. Thanks to him, I was given the opportunity to be his TA for several semesters and also at his suggestion, I could gain experience as an instructor. I am always grateful to him for this support. He has also been very supportive towards my dissertation by suggesting several significant ideas during my proposal defense some of which are a part of this dissertation. My special thanks to Dr. Mitch Wilkes for supporting my work and giving me short but effective lectures on particle filters and other signal processing topics, some of which are a part of this dissertation.

Most of this dissertation is based on the dataset obtained from the Rush University Medical Center and I thank Dr. Christopher Goetz and Dr. Glenn Stebbins for the same and also for providing valuable comments on paper reviews. I would also like to thank the Department of Electrical Engineering – Dr. Daniel Fleetwood and Dr. Bharat Bhuva for supporting me as a TA over and beyond the typical period. Ms. Linda Koger, our administrative assistant, has been extremely helpful during my graduate study and has always ensured that my graduate school requirements are met with. My work was partially supported by the Vanderbilt University Discovery Grant and RO1 EB006136, for which I am extremely grateful. I have developed some special friendships during my stay at Vanderbilt; Linda and Sandy, our ex-administrative assistant, are definitely a part of it. I would like to thank my colleagues at the Medical Image Processing Lab for their friendship and support, especially Mr. Rui Li, who helped a lot with the image registration and tracking areas of the dissertation. My friends, Saumitra Vajandar, Sameer Mahajan, Devi Majumdar, Hari Timmaji, Sirisha Rangavajhala, and Vadiraj Hombal have been

my second family here in Nashville and I am always grateful for their love, friendship, acceptance, and encouragement.

Finally and most importantly, I am forever indebted to my parents for their utmost faith in my abilities and character and their unconditional love and support in all my endeavors, without which this dissertation would truly have not been possible. Words cannot express my gratitude and love towards my younger sister, who has been my only best friend through thick and thin and without her nurturing care and affection, I could not have overcome several obstacles in my life and proceeded forth towards this day. I thank the Almighty for giving me this wonderful family, this great opportunity to do a Ph.D. under the guidance of excellent advisors, and His ever generous support all my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

# CHAPTER I

# INTRODUCTION

## I. Significance

The Parkinson's disease foundation estimates that as many as one million individuals in the United States are afflicted with Parkinson's disease [1]. While approximately four percent of people with Parkinson's are diagnosed before the age of 50, incidence increases with age. The major symptoms of Parkinson's disease vary from person to person, but can include tremor, slowness and stopping of movements (bradykinesia and akinesia), limb stiffness or rigidity, and difficulties with gait and balance. The cause of the disease is unknown. Although there is presently no cure, treatment options include medication and surgery to manage the symptoms. Levodopa remains the most effective medication for Parkinson's disease. Unfortunately, with increased dosing and prolonged use of levodopa, patients experience other side effects including dyskinesia and "on-off" periods when the medication unpredictably starts or stops working [2]. Drug induced dyskinesia is characterized by hyperkinetic involuntary movements that may interfere with activities of daily living, cause functional impairments and eventually disable the patient as its severity worsens. Pharmacological and surgical treatments are constantly being developed to provide respite to patients.

The assessment of the severity of dyskinesia is essential to develop better therapies to treat it. Drug trials and surgical treatments for dyskinesia can be better developed and evaluated by the clinimetrics of dyskinesia. Several rating scales exist to rate the severity of dyskinesia. There are several disadvantages to these rating scales in general [3]. First,

they are subjective in nature and may not be reproducible, both between raters and by the same rater. Training raters can mitigate but not eliminate this effect. Second, questionnaires given out to patients for assessing dyskinesia may not represent accurate details of the severity of dyskinesia. Finally, most rating scales are based on a Likert-type 5 point scale ranging from 0 (no dyskinesia) to 5 (severe dyskinesia). The disadvantage of a 5 point discrete scale is the lack of resolution and hence the possibility of misclassifying patients with symptoms who fall in between two scores. Hence developing quantitative assessment techniques has been a research challenge over the last decade.

Several devices have been used in the recent past with reasonable amounts of success to quantify the severity of dyskinesia. The disadvantage of these device-based techniques is the use of expensive and dedicated devices that are accompanied by expensive and complex software. Specific data collection protocols have to be designed and wearing these devices may be inconvenient to the patients. Most of the quantitative techniques convert the continuous output of their analyses into discrete values to correlate their findings to the manual ratings of neurologists. Thus, in spite of having quantified dyskinesia successfully, these techniques may not be user friendly for the patients and neurologists. A low-cost quantitative technique that is based on widely available patient data and provides simple interpretations to the severity of dyskinesia would benefit neurologists. This dissertation discusses the development and validation of a continuous score that interprets the severity of dyskinesia using patient videos

Since Parkinson's disease is a movement disorder, the progression of the disease is assessed by evaluating the patient's motor skills periodically using the standardized Unified Parkinson's Disease Rating Scale (UPDRS) [49]. The UPDRS is a six part scale

in which part III focuses on the motor symptoms of Parkinson's disease and part IV is a dyskinesia questionnaire. Though the UPDRS assesses the severity of dyskinesia using a brief questionnaire, specialized dyskinesia rating sales have been developed that will be discussed in detail in the section, 'Clinimetrics of dyskinesia'. We mention the UPDRS examination here because our initial hypothesis of assessing dyskinesia using patient videos was tested using UPDRS evaluation videos. UPDRS evaluations are performed by trained raters or neurologists in a laboratory or a clinic, and are video recorded for retrospective analysis. These evaluation videos then become a part of the patient's clinical record. The most common way to use these videos is to perform retrospective subjective ratings.

Our goal was to use image and signal processing techniques on these patient videos to quantify the severity of dyskinesia. This problem is difficult, given the complexity in obtaining useful information from two dimensional videos frames as opposed to the direct quantitative measures such as acceleration, velocity, position coordinates, and displacements obtained from devices such as accelerometers, gyroscopes, etc. Tracking the patient's movement through video sequences and extracting meaningful features that represent the characteristics of dyskinesia is the first concern. The second issue is how to combine these features to develop a formula or a score that captures the severity of dyskinesia. The validation of such a score is the third issue since there is no gold standard available that can be used to compare with the quantitative score. Hence a validation protocol needs to be developed to assess the effectiveness of the score. By developing solutions to address these issues, this work will be the first to use widely available, cost

effective videos of Parkinson's disease patients to quantitatively assess the severity of dyskinesia.

## II. Background

Our work can be categorized under two topics: (1) video tracking and analysis, and (2) clinimetrics of dyskinesia. A brief review of published work in these areas is presented here with their relevance to the issues mentioned in the previous section.

### II. 1 Video tracking and analysis

Several surveys have been published in the area of human motion analysis and tracking [4, 5, 6, and 7]. These authors have presented several taxonomies of human tracking. The following figure represents the one such taxonomy.

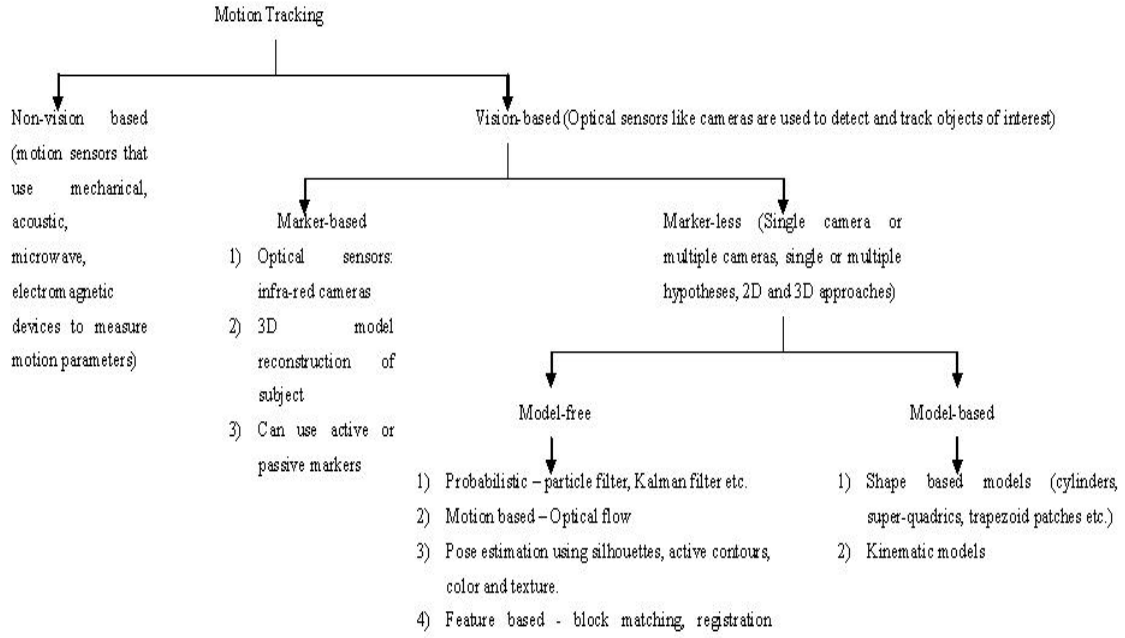


**Fig 1.1**. Motion tracking taxonomy.

Our work is an example of a vision-based, marker-less, model-free human motion tracking and estimation problem. We use an intensity-based non-rigid registration

algorithm to track points of interest in videos sequence and then proceed to use PCA to quantify the tracked motion data. Hence, we will focus predominantly on prior work in marker-less, model-free vision-based tracking methods and discuss the other aspects of tracking very briefly.

Non-vision based methods use motion sensors that are attached to or held by the human tracking target. These devices measure various kinetic and kinematic parameters that help track the subject's motion in space and time. A detailed review of different types of motion sensors and their applications can be found in [8]. Most of the prior work in dyskinesia quantification uses non-vision based methods and these are reviewed in the next section. Since it is not possible to specify the actual action of the subject at all instants of the sensor signals, a simultaneous video capture is also performed in some cases to aid the analysis. Zhou et al. give a detailed description of some of the recent non-vision based systems that are used for human motion tracking [5]. Non-vision based techniques have disadvantages owing to the use of dedicated devices which may be expensive, cumbersome, and not power-efficient. In addition to these disadvantages, these motion sensors can gather data only about the movement, but cannot provide any information about the shape of the object being tracked or about its environment.

Vision-based methods are called so because they use computational motion tracking techniques ostensibly similar to human vision, i.e., optical sensing techniques. Tracking can be performed with or without markers. Marker-based tracking techniques generally use infra-red cameras that can detect and track movements of passive reflective markers affixed to the subjects. To prevent occlusion problems, several cameras are used to capture the marker positions simultaneously and a 3D model of the subject's movement

is reconstructed. Marker-based motion capture has been successfully used in rehabilitation and gait analysis, sports, graphics, and animation. Internal motion models are used to compute the kinetic and kinematic parameters of movement of the subjects. Some examples of marker-based systems are Vicon [9], Qualisys [10], CODA [11]. The main disadvantages of these systems are their high cost, lack of portability, and the use of markers and special suits on subjects who may find it inconvenient.

Marker-less vision-based tracking techniques use video sequences obtained from cameras. These techniques can be used for 2D and 3D tracking applications. Model-based approaches use a-priori human body knowledge that is modeled using kinematic models or shape-based models. The challenge in these techniques is the development of a model with sufficient degrees of freedom to generate realistic kinematic movement. Poppe has described prior work using kinematic and shape models in his survey paper [12].

Our work tracks 2D video data using a model-free approach. These techniques do not use explicit a-priori body models to track subjects in video sequences. Since our dataset consists of uncalibrated videos of patients with different body types with possible presence of disabilities due to Parkinson's disease, developing a single body model including these shape constraints is not a tractable problem. The movement of the subject in different regions of interest is analyzed using image processing techniques without any prior information about the object shape. We have limited our review to a few examples in biomedical applications. Goffredo et al. used different types of pose estimation techniques to track human movement in video sequences. They used a silhouette-based human tracking technique on single camera video images to quantify human gait and researched its usability in clinical applications [13]. In their prior work, they combined

the classical ‘snakes’ technique [57] with neural networks as predictors to enhance the capability of silhouettes tracking techniques for applications in rehabilitation [14]. They also used a block matching technique to evaluate balance strategies in posturography [15]. Allin et al. similarly used the block matching technique to track movement of elderly subjects in uncalibrated videos to capture the postural sway parameters that distinguish normal and abnormal balance performance [16]. Chang et al. presented pioneering work in using single camera video images to quantify the abnormal gait and posture in Parkinson’s disease patients. They segmented patients from the video background to perform pose estimation by extracting features pertaining to distances and joint angles and use back propagation neural networks to perform classification between the normal and Parkinson’s disease group [17]. Similar work for gait and posture analysis was performed by Lee et al. [18] and Tan et al. [19]. Lee et al. used swing distances and joint angles as input features to general regression neural networks to classify normal and Parkinson’s disease gait [18]. Tan et al. focused more on the video tracking problem by developing better segmentation and image restoration techniques [19]. Sami et al. employed video tracking and analysis technique to study neonatal seizures. Motion detection and estimation is performed using optical flow computation. This step is followed by the use of adaptive block matching with Kalman filters to track multiple body parts in neonatal videos [20]. Emoto et al. have used particle filters to perform gait identification in humans under camera-view independent conditions [21]. Poppe presents several examples of color and texture-based motion tracking techniques that have been used in the past [12]. Additional work describing other areas of applications can be found

in the survey paper by Moeslund et al., which gives a detailed description of the various techniques used in tracking of humans in video sequences [4].

We have used an intensity-based non-rigid image registration algorithm to track dyskinetic movement. Non-rigid registration is a widely used medical image analysis technique. It could be point-based, intensity-based or surface-based registration technique. A detailed description of these methods and their various applications such as shape analysis, cardiac motion estimation, the development of medical atlases, image guided surgery, and study of disease progression can be found in Crum and Hill [22]. Lucas and Kanade developed a similar feature-based iterative image registration technique using spatial-intensity gradients in images and Newton-Raphson optimization techniques to track objects in stereo vision sequences [23]. Hager and Toyama developed a system called X Vision using the principle of image warping based on edge detection and sum-squared-difference optimization techniques to track facial features in video sequences [24]. Such sum-squared-difference based trackers have been used in the past for human tracking [25, 26]. Research in non-rigid image registration has been widely published and we have referenced only those studies that have applications in video tracking.

We now proceed to give a brief overview of our work in movement tracking using videos. We have used the intensity-based non-rigid image registration technique, Adaptive Bases algorithm (ABA) [27]. By employing non-rigid image registration on the entire image, we have omitted the segmentation or the pose estimation processes and track only certain points of interest on the human target as opposed to the entire human target. Tracking only these points of interests provides us with sufficient information and

makes this technique simple to use. We have not used any prior information or constraints regarding the shape or the position of the target in our method. Our technique uses adaptive radial basis functions to perform image warping and normalized mutual information as the similarity measure. ABA has been shown to be robust and accurate [27]. A possible disadvantage of this method may occur when using images that show high intensity homogeneities and exhibit occlusions and re-appearances of body parts. A detailed description of the registration algorithm and its application in tracking is discussed in chapter 2.

### II.2 Clinimetrics of dyskinesia

The clinimetrics of dyskinesia include qualitative and quantitative assessment of the severity of dyskinesia. Hoff et al. give a review of the techniques in the assessment of dyskinesia used one decade ago [28]. Though the subjective rating scales still remain in use widely to this day, several quantitative techniques have been developed that focus exclusively on the assessment of dyskinesia. A more recent review article by Kiejsers et al. discusses some of these techniques [29].

#### II.2.1 Rating Scales

Rating scales have been the most established and widely used means of assessment of severity of dyskinesia by clinicians for several decades. We present here a short review of the commonly used dyskinesia rating scales. Until recently, there was no standard dyskinesia rating scale such as the UPDRS for Parkinson's disease assessment. The key attributes of dyskinesia are described as anatomical distribution, phenomenology, duration, intensity, disability and patient perception. Different rating scales base their

dyskinesia rating on different attributes of dyskinesia. The Abnormal Involuntary Movement Scale (AIMS) is rated on different parts of the body thus emphasizing the anatomical distribution of dyskinesia and its intensity [30]. The Lang Fahn activities of daily living scale focuses on the intensity of dyskinesia experienced by the patients while performing regular day-to-day activities [31]. The most commonly used rating scale is the Rush Dyskinesia Rating Scale, which has undergone sufficient clinimetric testing [32]. This scale is used to perform objective assessments of the ability of the patient to perform tasks of daily living. Some of these scales are patient assessment questionnaires such as the Parkinson's disease Dyskinesia Rating Scale which may not be the most accurate technique to assess dyskinesia [33]. The Clinical Dyskinesia Rating Scale was developed based on anatomical distribution, disability and intensity of dyskinesia and is a commonly used rating scale [34]. The most recently developed scale is the Unified Dyskinesia Rating Scale (UDysRS) which is expected to become the standard dyskinesia rating scale equivalent to the UPDRS scale for Parkinson's disease symptoms [35]. The UDysRS is a five part scale that assesses: 1) general patient perceptions, 2) patient perception with Off Dystonia impact, 3) objective impairment based on anatomical distribution and intensity while performing four specific tasks, and 4) disability based on the Rush Dyskinesia Rating Scale. Finally a total objective score is computed based on the impairment and disability scores. The UDysRS combines several rating scales in such a way that all attributes of dyskinesia can be assessed using a single rating scale. The results of the clinimetric testing of this scale over a range of 70 patients indicated an inter-rater and intra-rater reliability with correlation coefficients ranging from 0.37 to 0.87 for various tasks. Further validation and development is underway to standardize the

UDysRS. Though neurologists extensively use these conventional rating scales, their disadvantages of subjectivity, low resolution and tedious rating instructions are compelling reasons for moving towards quantitative assessment techniques.

### II.2.2 Quantitative Assessment Techniques

Quantification of movement has been most commonly performed with the help of devices. Quantification of dyskinesia has been researched only in the past two decades and prior work in this area is limited. Burkhard et al. use solid state gyroscopes on the upper extremities of the patients to determine the severity of dyskinesia when they do not perform any voluntary tasks [36]. Kiejsers et al. used uni-axial accelerometers mounted on the upper and lower extremities of the most affected side. Parameters from the accelerometer data are used as inputs to a supervised neural network to detect and classify dyskinesia [37]. Subsequent research by Kiejsers et al. resulted in the development of an automatic method to assess dyskinesia in daily lives of Parkinson's disease patients and establish a set of movement parameters that can distinguish between voluntary and involuntary movements [38]. Similar studies using accelerometry data was published by Hoff et al. and Patel et al., Hoff et al. used bi-axial accelerometers and frequency spectrum analysis was performed on the recorded acceleration signals [39]. Patel et al. used tri-axial accelerometers and PCA was applied to determine the most useful features that can be extracted from the acceleration data. Clustering analysis was performed using these features to classify patients into varying severity ranges [40]. Liu et al. quantified dyskinesia in the arms using digitized spiral drawing tasks. The patients were asked to draw spirals on a digitized graphic tablet and the drawing velocity computed using the positions of the pen in the radial and tangential directions were

correlated with the amplitude of the dyskinetic movements [41]. Gour et al. used an electromagnetic motion tracker system to analyze patterns embedded in dyskinesia movements to determine if dyskinesia is truly random is nature [42]. Though quantification of dyskinesia was not the direct goal of this study, the authors have obtained results which interpret dyskinesia as not truly random movements but as having deterministic patterns. A more recent study was performed by Chung et al. to quantify dyskinesia with patient data collected using force plates. Patients were asked to suppress voluntary movements while standing on the force plates. The variations in the center of pressure were analyzed and correlated with the modified AIMS rating scale [43]. This technique is not effective when evaluating patients who are unable to stand on the force plate due to severe disability caused by Parkinson's disease.

In the initial development of our score, we have analyzed the tracked patient data using PCA [44]. PCA has been commonly used as a dimensionality reduction technique to determine the most significant features that best represent a high dimensional dataset in a lower dimensional space. We are not interested in reducing the dimensionality of our feature set. Instead, we want to use PCA to study the variances and co-variances of the tracked points as they move together in the video sequence. In motion analysis, Daffertshofer et al. have used PCA to study human movement coordination and variability in walking trials [45]. Beleznai et al. have used PCA to track local multiple objects in a moving video sequence and applied their algorithm on human movement tracking [46]. In this work, we have used eigenvalues and eigenvector of the covariance matrix to capture this variation. The severity score initially developed using eigenvalues is further modified to include more attributes of dyskinesia including randomness of

movement, chorea and dystonia. We have used some of the concepts in [45] in this context by analyzing the evolution of eigenmodes. We experimented with approximate entropy, frequency spectrum dispersion and speed of movement trajectories to include the effects of dystonia and chorea in the existing severity score. There have been very few studies reporting on the quantification of chorea and dystonia and the most relevant to our work is by Beuter et al. They quantify dystonia using a position transducer attached to the patient's left and right wrists. The patient is asked to perform the finger-to-nose task simultaneously with both hands. The finger-to-nose task requires the patient to touch their index fingers to the tip of the nose as fast as they can. The trajectories of this repetitive task are analyzed and features capturing the smoothness of the movement are used to quantify dyskinesia severity [47].

The clinimetrics of dyskinesia play an important role in the effective treatment of this motor dysfunction. Qualitative assessment provided by the rating scales are mostly preferred by neurologists because they are mostly based on visual observation of the patients. Since these observations are expressed through subjective interpretation of the attributes of dyskinesia as specified in the rating scales, they are rater-dependent. They are good initial assessment techniques, but fall short when a quantifiable measure or score is necessary to determine the efficacy in drug trials or surgical treatments. These quantitative measures will parameterize the attributes of dyskinesia qualitatively assessed in rating scales and yet be rater independent and precise.

## III. Overview Of This Dissertation

**Patient Data -** This dissertation was developed using patient data from two sources. The initial study utilized existing UPDRS videos of Vanderbilt University Medical Center's

(VUMC) patients and developed a severity score based on video tracking and PCA analysis. Due to several inadequacies in the quality of the videos, 35 patient videos and corresponding scores from the UDysRS trial study were utilized to perform further developmental and validation studies. These videos and the corresponding scores were acquired from Dr. Christopher Goetz and Dr. Glenn Stebbins at the Rush University Medical Center (RUMC), Chicago, IL [35].

The remainder of the dissertation is organized as follows. Chapter 2 discusses the necessity of the severity score, its development and validation using a small dataset of Parkinson's disease patients, including a control group of non-dyskinetic patients. This study was performed using the VUMC dataset. Chapter 3 validates the severity score on the RUMC dataset of 35 videos using a specially developed ranking protocol and reports on the intra-rater and inter-rater variability studies performed on this data using the ranking protocol. The robustness of the score is analyzed by evaluating its variations with the use of longer video sequences and changes in the position of the landmark points. Chapter 4 discusses the modification of the severity score to include chorea and dystonia and the validation study using the ranking protocol is repeated to show the improvement in the correlation between the severity score rankings and the neurologists' rankings and the severity score ranking and the UDysRS rankings. In Chapter 5, a summary of the contributions of this work is presented along with possible directions for future work in this area.

# CHAPTER II

## DEVELOPMENT OF SEVERITY SCORE

Rao AS, Bodenheimer RE, Davis TL, Li R, Voight C, Dawant BM, Quantifying Drug Induced Dyskinesia in Parkinson's disease Patients Using Standardized Videos, Proceedings of the30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2008; 1769-72.

### I. Introduction

Measurement of dyskinesia using devices such as accelerometers, gyroscopes or marker-based motion capture techniques; and the analyses of this data to develop mathematical parameters that effectively describe the various attributes of dyskinesia have been the conventional approaches to quantify dyskinesia objectively. Specifically, previous work includes the use of classification algorithms such as neural networks [37], expectation-maximization and clustering [40] on the data acquired with these sensors to detect and quantify dyskinesia in Parkinson's disease patients. These methods involve expensive, cumbersome devices that have to be worn or held by the patients and use complex data processing algorithms. The goal of this dissertation was to develop a completely non-invasive technique using patient videos, which are collected as a part of a Parkinson's disease patient's clinical record. We tested our hypothesis using patient videos which are routinely collected for UPDRS evaluations by the neurologists and may not be used primarily to evaluate dyskinesia. The UPDRS videos are captured with patients being in OFF medication and ON medication states. During the OFF medication state, the disease symptoms return and the severity of the disease and its progression can be evaluated.

During the ON medication state, the symptoms of Parkinson's disease are generally controlled, but the side effect of the medication, namely dyskinesia, surfaces and inhibits the patients from performing the UPDRS tasks efficiently. Using these ON medication videos of patients performing various tasks such as reading, pronation and supination of the wrist, finger-to-nose movements, opening and closing of fists, and heel tapping, we can detect and quantify the abnormality of the patient's movements. Dyskinesia is characterized by abnormal, random involuntary movements and we quantify these features by analyzing the trajectories of the movement of various body parts of the patient.

Though the characteristics of dyskinesia such as amplitude, velocity, frequency, and randomness have been quantified using time series analysis of movement trajectories [38], clustering analysis of features obtained from acceleration data [40] or by analyzing kinetic and kinematic parameters [39, 40, 41], these quantifying measures have been individually correlated to conventional rating scales that have been used as ground truths. We propose to combine our measures to form a single score, which collectively represent the attributes of dyskinesia, and compare it with neurologists' scores. We do not classify the patients under the conventional Likert scale pattern of 1 – 5, with 1 representing mild dyskinesia and 5, severe dyskinesia, but instead advocate the use of a continuous score. A continuous score provides a finer assessment of severity when compared to a rating scale in which several patients with varying degrees of dyskinesia may be accommodated into a single rating category. By knowing the range of scores obtained by our patients and their individual scores, one can determine the corresponding rating any patient would have received if evaluated using a conventional rating scale.

Since we have developed a continuous score, we have chosen to compare our scores to the neurologists' scores by rank order correlation. Since our pilot involves small sample set of patients, this validating technique was found to be appropriate. For larger datasets of patients, better validating techniques have been used as seen in the next chapter. This chapter focuses on the preliminary results obtained when attempting to analyze the severity of drug-induced dyskinesia using video based analysis. We have used an intensity-based non-rigid image registration algorithm to track patients' head and shoulders while they are performing the reading task in the UPDRS evaluation. The trajectories of the head and shoulder movement were analyzed using PCA [44] to determine movement parameters that could be used to quantify the severity of dyskinesia. These parameters were then combined to form the severity score, SVS. We rank the patients according to their SVS scores and compare our ranking to the rankings of an expert neurologist using Spearman rank order correlation [48]. A high correlation was observed between our rankings and the neurologist's rankings. Similar analysis was performed using a control group of non-dyskinetic Parkinson's disease patients to further establish the utility of our score in quantifying dyskinesia.

## II. Methods

### II.1 Patient data and pre-processing

UPDRS part III evaluations [49] are often video recorded as a part of a patient's clinical record. An ON medication segment and OFF medication segment is recorded to study the variability of motor symptoms of Parkinson's disease due to medication. Though the UPDRS part III examination does not include a dyskinesia rating, several

patients exhibit dyskinesia in the ON medication segment due to mental/physical exertion caused by the examination. The simplest task in the UPDRS part III motor examination is the reading task. The patient is asked to sit still and is given a few lines to read out loud and the rater then evaluates the patient's speech and expressions in this task. Since dyskinesia is initiated or worsens due to physical or mental exertion [2], patients affected by dyskinesia show involuntary movements of the upper and lower body while performing the reading task. Our aim is to detect and quantify these involuntary movements.

The patients were filmed for an average of ten seconds. A short video sequence of two seconds showing the patient performing the reading task was extracted from these 10s segments of 26 Parkinson's disease patients (13 dyskinetic (DP), 9 non-dyskinetic (NDP) and 4 non-dyskinetic with tremor (NDPT)). These short sequences were rated by an expert neurologist on a scale of 0 (absent) - 4 (severe) for dyskinesia severity. The dyskinetic patients were further ranked from 1(least) - 13 (most) based on their relative severity of dyskinesia. The mental exertion due to the speech task is known to initiate or worsen existing dyskinesia leading to involuntary movements of the head, shoulders and in facial expressions. Since we do not have the patient's entire body in view, neither the severity score nor the neurologist evaluates dyskinesia that may be present in the lower body. In this study we detect and analyze the severity of dyskinetic movement of the head and shoulders. Each sequence was extracted into 60 individual frames. All the frame sequences were normalized across patients to avoid scaling errors. Patient consent was obtained from all patients for using the videos for research studies and publication.

## II.2 Data analysis

### II.2.1 Registration and Tracking

The next step was to track the head and shoulders of the patient on each of the 60 video frames. After an initial attempt in using block matching techniques failed, a non-rigid image registration algorithm was used to perform motion tracking in the video frames. The adaptive bases algorithm (ABA) is an intensity-based non-rigid image registration algorithm [27]. In the following section, we briefly describe how ABA is used to track the patient's movements.

**Adaptive Bases Algorithm (ABA) – An overview**

Image registration is the process of deforming or morphing a source image to match a target image such that the variations in the images are minimized using a similarity measure. Image registration can be achieved by rigid or non-rigid transformations of the source image into the target image. Non-rigid transformations have higher degrees of freedom than rigid transformation, resulting in elastic deformations. Hence it is best suited to our application which involves, bending and stretching movements in patients. More details about non-rigid image registration and applications can be found in [22].

The ABA technique is intensity-based and uses compactly supported radial basis functions to model the deformation field and normalized mutual information as the similarity measure between the target image and the deformed image. It is a multi-resolution and multi-scale technique that follows a bottom-up approach of proceeding from coarse levels to finer levels, where each level is defined as a particular combination of resolution and scale. Deformation fields are first initialized at the lowest image

resolution and scale (fewer basis functions with wider regions of support). Within a given level, local deformations are performed by varying the basis function parameters such that each region of mis-registration is optimized before proceeding to the next level. The optimization process also integrates a constraint that ensures unnatural deformations do not occur by ensuring the Jacobian of the deformation field at every given level is positive definite. The final deformation field is the sum of the deformation fields obtained at each level of optimization. The process of using deformation field to track patient movements is described as follows.

**Tracking using ABA**

Each video frame is treated as an individual image and the change in the patient's positions from a given frame to the next is captured using ABA. For every frame $F_n$ as source image, the consecutive frame $F_{n+1}$ is the target image and the deformation field $DF_{n,n+1}$ is computed using ABA. This deformation field $DF_{n,n+1}$ gives the mapping of every point in $F_n$ to the corresponding point in $F_{n+1}$, thereby facilitating the tracking of desired points across the video sequence. Thus, starting with the first frame of the sequence, successive frames are registered to obtain a series of deformation fields relating a given frame to the next frame. The relation between the first frame and the final frame of the sequence is given by

$$F_n = \prod_{i=1}^{n} DF_{i,i+1}\, F_1$$

We were interested in capturing the movement of the head and shoulders in the frames and three landmark points were chosen for this purpose. The center of the forehead, right and left shoulders were selected as the landmark points. A higher number of landmark

points did not improve the tracking quality. Tracking the motion of these landmark points in video sequence is equivalent to tracking the motion of the head and shoulders in the video sequence. The three landmark points are chosen manually in the first video frame. Each frame is sequentially registered to the next, establishing a transformation of the landmark points from the given frame to the next. This process is repeated until the final video frame is reached, and the result is a set of image coordinates of the three landmark points in every frame of the video sequence. Figure 2.1 shows tracking of head and shoulders in a severely dyskinetic patient and Figure 2.2 shows the deformation fields computed between consecutive frames of a severely dyskinetic patient.

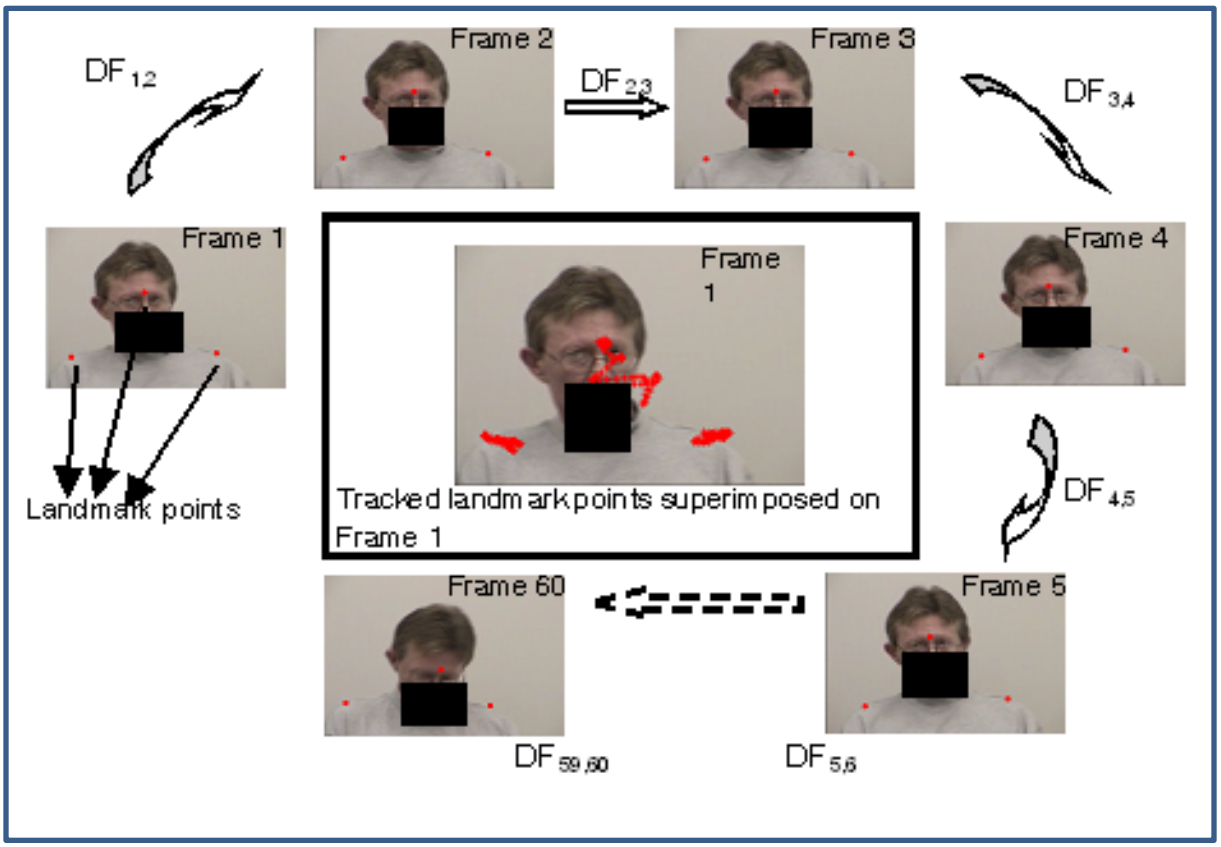


**Fig. 2.1** Graphical representation of the tracking of landmark points using ABA in a severely dyskinetic patient. The cluster of points in the center image represents the trajectory of motion of the landmark points at the head and shoulders.

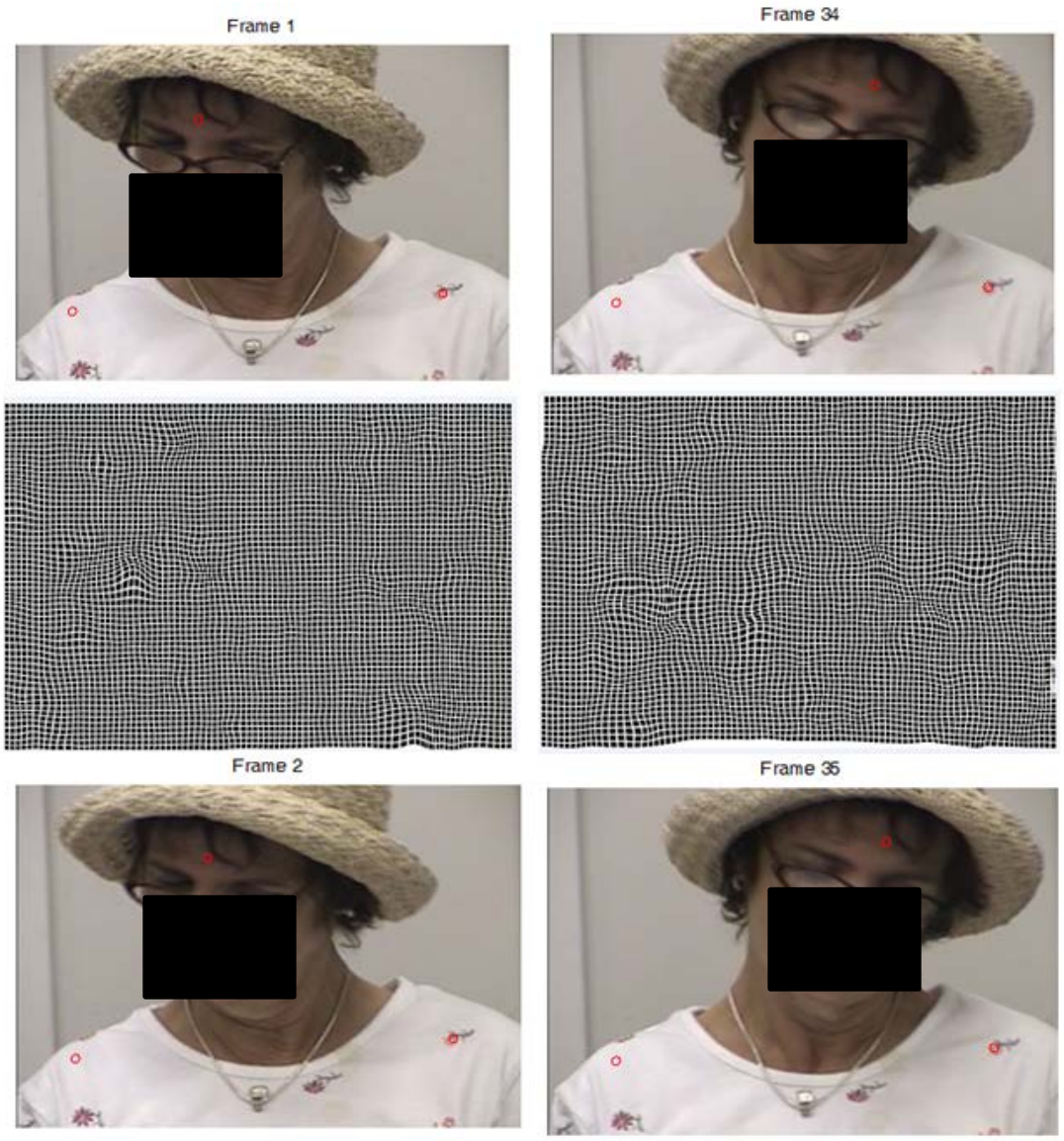


**Fig 2.2** Frame-by-frame tracking of head and shoulders in a severely dyskinetic patient. The deformation fields mapping the first frame to the next is shown superimposed over a grid.

By using non-rigid registration, we can automatically track the landmark points. The only manual step in this process is to pick the landmark points to be tracked in the first frame. Depending on the part of the body that needs to be evaluated for dyskinesia, these landmark points can be chosen accordingly. The quality of the tracking can be observed from the examples in Figures 2.1 and 2.3. We have performed only qualitative analysis

on the tracking results as small errors in tracking will not affect our analysis for large movements exhibited by dyskinetic patients, but may do so if there are extremely fast movements with low amplitude. Since ABA is an intensity-based algorithm, regions with homogenous intensity may not be correctly identified as mis-registered regions all the time, even where there is movement occurring in those areas. Thus, the quality of tracking may be reduced in these areas. When patients wear clothing without textural or color contrasts, such homogenous intensity areas are present and the movement of the landmark points, if chosen in those areas, in the successive frames may not be captured.

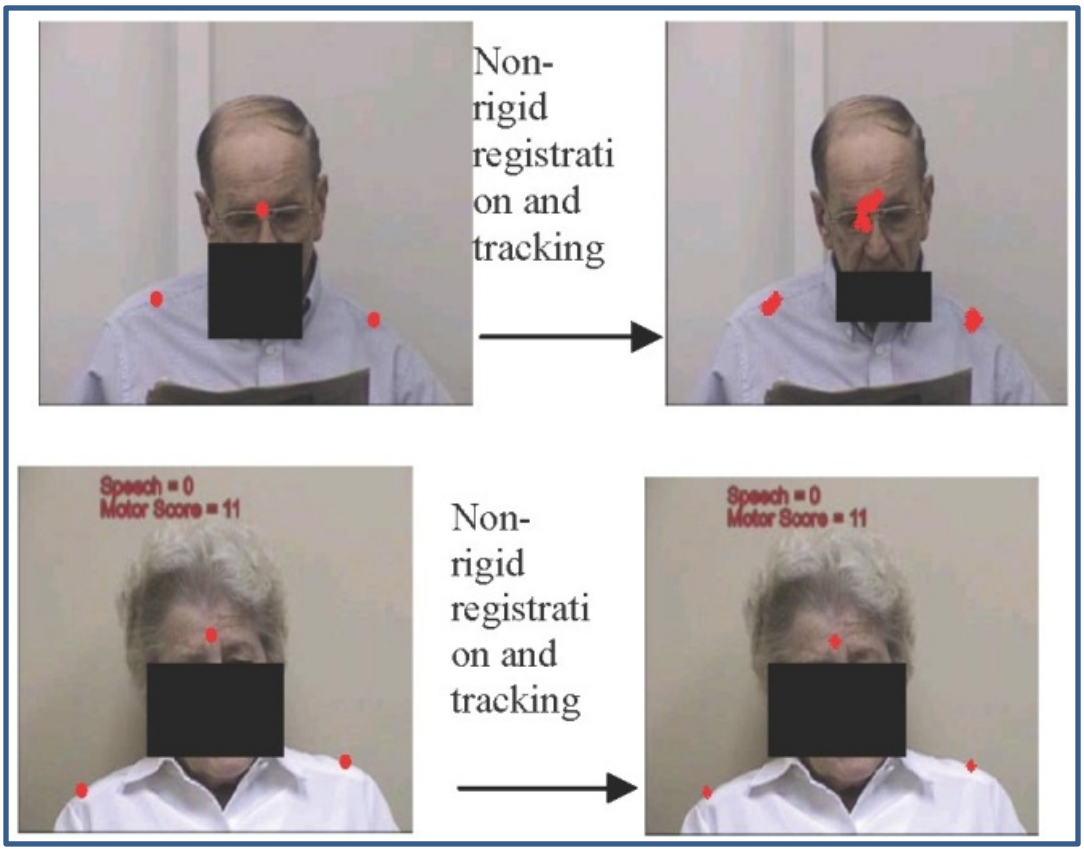


**Fig. 2.3** Tracking of head and shoulders in (top) a modereately dyskinetic and (bottom) a non-dyskinetic patient

The left image in Figures 2.3 (a) and 2.3 (b) shows the landmark points on the first frame of the patient. The right image shows the first frame with the landmark points computed

in subsequent frames superimposed on it. The differences in the trajectories of the landmark points for the DP and NDP patient shows the dyskinetic movement exhibited by the DP patient as compared to the absence of any movement in the NDP patient.

### II.2.2 Principal Component Analysis and Severity Score (SVS)

The result of registration and tracking of the landmark points is a set of image coordinates in every frame of the video sequence. This step is performed for all the 26 video sequences in this study. The goal is to establish a relation between the dyskinetic movements and the trajectory of the landmark points in the video sequence. Our analysis method mimics the strategy used by neurologists to rate the severity of dyskinesia. Neurologists look at multiple body parts simultaneously to determine the severity of dyskinesia. Similarly, we analyze the landmark points as a group instead of studying their individual trajectories. The motivation for this analysis was based on the concept of active shapes model introduced by Cootes et al. [50]. The principle behind the active shapes model is based on the point distribution model (PDM). Given a set of training shapes of an object in an image, its PDM is computed by defining a set of homologous 'landmark' points on the training shapes, computing their mean positions and determining the statistical variations in the positions of homologous landmark points from the mean positions. The eigenvalues and eigenvectors of the covariance matrix, obtained by PCA analysis of the training shapes, are used to capture these statistical variations across the training shapes. Each axis in the PC space corresponds to a mode of variation. Using the mean shape and weighted sum of most significant eigenvectors, new shapes can be constructed. Thus, shape synthesis was the main goal of the PDM.

We limit our analysis to studying the statistical variations in the set of training shapes. Using non-rigid registration for tracking, the three landmark points were automatically tracked, thus generating homologous points for every frame of the sequence. Our set of training shapes consisted of the set of frames in the video sequence, with each shape defined by the landmark points in the head and shoulders as seen in Figure 2.4. Procrustes method was not applied as all the video frames were already aligned [51].

Let

$$T = \begin{pmatrix} x_{11} & x_{12} & x_{13} & & & x_{1,N-1} & x_{1,N} \\ y_{11} & y_{12} & y_{13} & \dots\dots\dots\dots\dots\dots & & y_{1,N-1} & y_{1,N} \\ y_{M1} & y_{M2} & y_{M3} & & & y_{M,N-1} & y_{M,N} \end{pmatrix} \text{.....(2.1)}$$

represent the 2D coordinates of the 'M' landmark points in a video sequence of 'N' frames. T defines the set of training shapes.

Mean shape is computed by row averaging the coordinates to get $T^1$ as

$$T' = \begin{pmatrix} \bar{x}_1 & \bar{x}_2 & \dots\dots\dots & \bar{x}_{N-1} & \bar{x}_N \\ \bar{y}_1 & \bar{y}_2 & & \bar{y}_{N-1} & \bar{y}_N \end{pmatrix}, \text{.....(2.2)}$$

$$\bar{x}_i = \frac{1}{N} \sum_{k=1}^{N} x_{ik}\,, i = 1{:}\,M$$

$$\bar{y}_i = \frac{1}{N} \sum_{k=1}^{N} y_{ik}\,, i = 1{:}\,M$$

For each training shape, we calculated the demeaned shape such that

$$X = T - T' \text{.....(2.3)}$$

Next, the covariance matrix of X is computed as

$$COV = \frac{1}{N} \sum_{i=1}^{N} X_i\, X_i' \text{.....(2.4)}$$

Using singular value decomposition, the eigenvalues and eigenvectors of COV are computed as follows

$$\text{COV} = \Lambda_k \begin{pmatrix} 1 & \cdots & o \\ \vdots & 1 & \vdots \\ 0 & \cdots & 1 \end{pmatrix} V_k \ldots\ldots\ldots (2.5)$$

where $\Lambda_k$ are the eigenvalues and $V_k$ are the corresponding eigenvectors.

$V_k$ correspond to the principal components or modes of variation. $\Lambda_k$ represents the spread or variance along the corresponding mode or eigenvector $V_k$.

In our dataset, M = 6 corresponding to the 2D coordinates of three landmark points. A 6x6 covariance matrix was generated using (2.4) and the corresponding $\Lambda_k$ and $V_k$ were computed.

**Fig.2.4** Set of 60 training shapes defined by the head and shoulders of a severely dyskinetic patient.

The eigenvectors or the modes of variations, capture the way in which the landmark points move together across the frames. The eigenvalues corresponding to an eigenvector described the variance of the landmark points together along the direction of the eigenvector. We seek the smallest number of eigenvectors that capture 95% of the total

variance. These eigenvectors, usually fewer than six, are called the significant modes of variation as they capture the significant variations of the landmark points together. Thus, the mean positions, eigenvectors and eigenvalues together define a PDM. The application of PDM in [49] is to synthesize new positions of the landmark points. Since our goal is to analyze the movement and not synthesize new movement, PCA is used only to generate the eigenvalues and eigenvectors.

The next step of the data analysis was to correlate the PCA parameters to the attributes of dyskinesia. For our preliminary work we have hypothesized that the severity of dyskinesia is characterized by the amplitude and the directions of the random movements. Chapter 4 of this dissertation will discuss more complex elements of dyskinesia to develop a more accurate quantitative score. To correlate dyskinesia to the PCA parameters, we must interpret the PCA parameters generated from the dyskinetic movements.

- The eigenvalues capture the extent of movement along the corresponding mode of the variation. The sum of the eigenvalues, which provides the total variance in all the frames, is proportional to the amplitude of the total movement exhibited by the patient in the video sequence. Hence, large amplitudes of movement by the patient are characterized by high total variances.
- We cannot directly determine a PCA equivalent to the directions of movement since we have applied PCA to the entire dataset and obtained modes of variation for all the landmark points instead of directions of movement for individual landmark points. Thus the eigenvectors now represent the directions of movements of the landmark points together. These eigenvectors could indicate

movement of just one point, or of two points simultaneously or of all the three points simultaneously. If a patient exhibits movements in several directions, it results in a larger number of significant modes of variation required to capture the 95% of the variance. Thus the number of significant modes of variation can be partially used to represent the randomness of the directions of movement.

- To further understand the directions of movement, we proceed to study the fall off or the rate of decrease of eigenvalues along the modes of variation. Two patients could have similar total variance and number of significant modes of variation, but the distribution of the eigenvalues along the modes of variation could be different, thus differentiating the severity of dyskinesia in these patients. The distribution of the eigenvalues is an indication of the variation or the extent of movement along the corresponding modes of variation. A steep decrease of the eigenvalues indicates that most of the movement is concentrated in the first few modes, whereas a gradual decrease shows that more number of modes have significant variance. This gradual decrease implies that the patient exhibits significant movement in several directions which in turn indicates a more severe case of dyskinesia. To mathematically describe this parameter we have computed the standard deviation of the percentage contribution of the eigenvalues to the total variance. We call this parameter STDEV.

$$\text{STDEV} = \sqrt{\frac{\sum_{n=1}^{N}((\frac{EV_n}{TV}*100)-\{\frac{1}{N}\sum_{n=1}^{N}(\frac{EV_n}{TV}*100)\})^2}{N}}$$

where N is the number of eigenvalues, $EV_n$ is the $n^{th}$ eigenvalue, and TV is the total variance. A larger standard deviation indicates a steep rate of decrease

whereas a low standard deviation indicates a gradual rate of decrease of eigenvalues.

## III. Results

We estimate the relationship between the severity of dyskinesia and the PCA parameters using the following examples from the dataset of 13 dyskinetic patients with varying severity of dyskinesia. We show the significance of each of these parameters by using the rating of the neurologist as ground truth. Each of the two seconds video sequence is rated by an expert neurologist from 0 (no dyskinesia) to 4 (severe dyskinesia). The neurologist rated the videos based on the general definition of dyskinesia. We define this clinical rating as the clinical score. The following table shows the two PCA parameters (total variance [TV] and number of significant modes [NSM]) for each of the 13 dyskinetic patients along with the rating of the neurologist. Patients with similar clinical score and different clinical scores are analyzed to describe the significance of the PCA parameters and also derive a relationship between the parameters and the severity of dyskinesia.

| DP Patient number | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| TV | 26.1 | 36.08 | 45.08 | 213.45 | 305.11 | 66.16 | 113.7 |
| NSM | 2 | 2 | 3 | 1 | 1 | 3 | 3 |
| Neurologist Rating | 1 | 1 | 2 | 2 | 2 | 3 | 3 |
| DP Patient number | 8 | 9 | 10 | 11 | 12 | 13 | |
| TV | 191.07 | 156 | 587.03 | 610.88 | 711.97 | 1445.96 | |
| NSM | 2 | 2 | 1 | 1 | 2 | 3 | |
| Neurologist Rating | 3 | 3 | 3 | 3 | 4 | 4 | |

**Table 2.1**. Total variance (TV), number of significant modes (NSM) and neurologist rating for 13 DP patients

**(a) Patients with similar clinical score**

Patients 3, 4 and 5 were rated with a score of two by the neurologist. The mean of the total variance of these patients is 187.86 and the standard deviation is 131.93. Patient 3 requires three significant modes of variation, whereas patients 4 and 5 require only one significant mode of variation. Thus the clinical score of two for patient 3 is attributed to the small but significant movement along more modes of variation, whereas, for patients 4 and 5 it is attributed to the high variance but along only one mode of variation. The large variance in patients 4 and 5 was not significant enough to increase the severity score, thus accommodating all the three patients in the same score category. A similar trend is observed for patients 6, 7, 8, 9, 10, and 11 who have the same clinical score of three. Patients 6 and 7 with the lowest variance have the highest number of significant modes compared to the patients 10 and 11 who have the highest variance and the lowest number of significant modes. Thus the total variance and number of significant modes vary directly with the severity of dyskinesia.

**(b) Patients with different clinical score**

| DP Patient # | 3 | 5 | 6 |
|---|---|---|---|
| % contribution from EV 1 | 77.55 | 96.90 | 62.38 |
| % contribution from EV 2 | 11.66 | 2.28 | 24.47 |
| % contribution from EV 3 | 7.69 | 0.70 | 11.69 |

**Table 2.2** Percentage contribution of the first three eigenvalues (EV 1, EV 2, and EV 3) to the total variance for DP patients # 3, 5 and 6

Patient 5 has a higher variance than patient 6, but lower number of significant modes. Similarly, patient 3 has lower variance than patient 5 but the same number of significant modes as patient 6. But patients 3 and 5 have a clinical score of two and patient 6 has a score of three. The total variance of patient 3 and 6 are comparable and NSM are equal. But the spread of eigenvalues from Table 2 for patient 3 is less than that of patient 6. Eigenvalues distributed more evenly along the modes of variation implies significant movement along all the modes of variation. Thus we infer that the distribution of eigenvalues along the different modes is an important factor in determining the severity of dyskinesia. From the cases described above, we can hypothesize that:

Dyskinesia severity α Total variance (TV)

α Number of significant modes of variation (NSM)

α 1/distribution of eigenvalues along the modes of variation

A heuristic solution based on the above dependencies was developed to quantify the severity of dyskinesia. The distribution of eigenvalues along the modes of variation was computed as the standard deviation of the percentage contribution of the eigenvalues to the total variance (STDEV). A severity score (SVS) was formulated to rank the severity of the patients:

The score was normalized with respect to the number of frames in the video sequences to ensure valid results on longer sequences that contain more movement.

$$SVS = (TV * NSM) / (STDEV * NF)$$

where NF is the number of frames in the video sequences. In this study, the number of frames was constant for all patients.

### III.1 Validation of SVS using Neurologist Ranking

The severity score is a continuous variable and the clinical score by the neurologist is discrete, based on the 0 – 4 scale. Moreover, since several patients were given the same clinical score, the neurologist ranked the patients based on relative severity of dyskinesia from 1 (least severe) to 13(most severe). The dyskinetic patients were ranked based on the clinical score and the severity score SVS as seen in the table below.

| Patient Number | TV | NSM | STDEV | SVS | Ranking using SVS | Neurologist's Ranking |
|---|---|---|---|---|---|---|
| 1 | 26.10 | 2 | 27.53 | 1.90 | 1 | 1 |
| 2 | 36.08 | 2 | 33.35 | 2.16 | 2 | 2 |
| 3 | 45.00 | 3 | 30.16 | 4.48 | 3 | 3 |
| 4 | 213.45 | 1 | 39.90 | 5.35 | 4 | 4 |
| 5 | 305.11 | 1 | 39.32 | 7.76 | 5 | 5 |
| 6 | 66.16 | 3 | 24.34 | 8.15 | 6 | 7 |
| 7 | 113.70 | 3 | 31.70 | 10.76 | 8 | 8 |
| 8 | 191.07 | 2 | 28.66 | 13.33 | 9 | 9 |
| 9 | 155.99 | 2 | 36.69 | 8.50 | 7 | 6 |
| 10 | 587.03 | 1 | 38.03 | 15.43 | 11 | 11 |
| 11 | 610.88 | 1 | 39.74 | 15.37 | 10 | 10 |
| 12 | 711.97 | 2 | 35.57 | 40.03 | 12 | 12 |
| 13 | 1445.96 | 3 | 26.54 | 163.48 | 13 | 13 |

**Table 2.3**. STDEV, TV, NSM, ranking using SVS and ranking of the neurologist.
A comparison of the two rankings shows a high correlation with a Spearman rank order correlation of 0.99.

***Non-Dyskinetic Patients***

We have seen that the SVS can be used to measure the severity of dyskinesia quantitatively. As a part of our study, we also evaluated nine non-dyskinetic patients (NDP) and 4 non-dyskinetic patients with mild to severe tremor (NDPT). Tremor is often confused with mild dyskinesia. Tremor is a rhythmic involuntary motion whereas dyskinesia is a random involuntary motion. We also look at the two PCA parameters (TV and NSM) of the NDP and NDPT group as shown below: The NDP and NDPT patients were given a score of zero by the neurologist signifying the absence of dyskinesia.

| NDP Patient number | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| TV | 11.10 | 6.21 | 1.08 | 0.80 | 0.10 |
| NSM | 1 | 3 | 2 | 3 | 3 |
| NDP Patient number | 6 | 7 | 8 | 9 | |
| TV | 2.31 | 6.07 | 1.27 | 4.47 | |
| NSM | 2 | 2 | 1 | 2 | |

| NDPT Patient number | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| TV | 10.30 | 7.16 | 6.96 | 28.39 |
| NSM | 1 | 2 | 2 | 2 |

**Table 2.4**. TV and NSM for NDP and NDPT patients.

## IV. Discussion

To the best of our knowledge this is the first study that quantified dyskinesia using patient videos and PCA-based methods. We have shown that in our preliminary analyses; a single number, namely the severity score, SVS, can be used to quantify dyskinesia as

opposed to previous studies that have measured various attributes of dyskinesia and correlate them individually to rating scales. Our results indicate that the total variance, number of significant modes of variation and the standard deviation of the percentage contribution of the eigenvalues are suitable parameters to define the characteristics of dyskinesia. By combining these parameters suitably to form the SVS represented by a single number, efficacy studies in drug trials are easier to perform. If necessary, these individual parameters can also be assessed for each of the patients. The utility of our score is further tested on a larger number of datasets in Chapter 3. We have also shown the benefits of a continuous score over a discrete score and the possibility of finer classification of patient's dyskinesia severity.

The results of the analyses on the non-dyskinetic group of patients indicate that for most of the NDP and NDPT patients, the total variance was negligible compared to the DP patients. The high number of significant modes for NDP patients 2, 4 and 5 can be attributed to a small amount of noise due to the registration and tracking or the inability of the patient to sit completely still while performing the speech task. The small variance in these patients indicates that there is insignificant movement of the points, so large number of significant modes is still insufficient to classify them as dyskinetic patients. NDPT patient 4 had severe tremor. The total variance and number of significant modes of this patient is comparable to that of DP patient 1 from Table 1. But the percentage contribution of the first eigenvalue to the total variance for NDPT patient 4 is 92.2% whereas for DP patient 1 it is 68.51%. Tremor is a rhythmic motion that can be represented using a single mode. Though NDPT patient 4 has two significant modes, over 90% of the variation in movement is captured in the first mode itself. The other three

NDPT patients 1, 2 and 3 have smaller eigenvalues owing to their mild tremor. The significant variance of NDP Patient 1 is attributed to the downward movement of the camera as seen in Figure 2.5. Since the motion was only in the downward direction, there is only one significant mode of variation.

Some of the disadvantages of our score are that it cannot differentiate voluntary movement from dyskinetic movement and fails to give weightage to the most disabling type of dyskinesia – namely chorea and dystonia. Since this work has used the communication task as the task of interest for dyskinesia quantification, voluntary movements are rare and can be avoided by instructing the patients to control their voluntary movements. Parameters quantifying chorea and dystonia will be discussed in detail in Chapter 4.

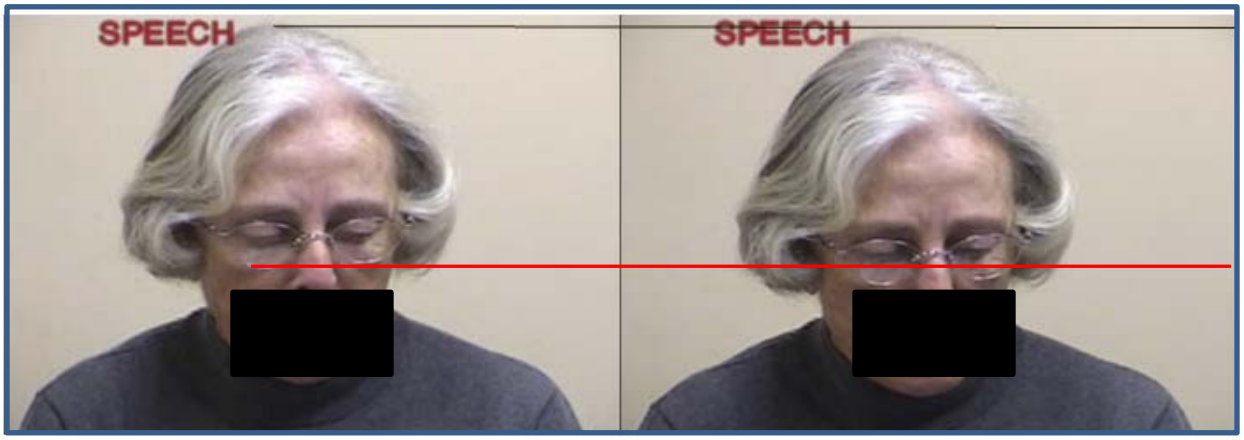


**Fig. 2.5**. First and last frame of a NDP patient # 1 with camera movement. Note the slight downward shift of the patient with respect to the background in the last frame compared to the first

# CHAPTER III

## VADIATING SEVERITY SCORE AND RATER VARIABILITY ANALYSIS

Rao AS, Dawant BM, Bodenheimer RE, Li R, Fang J, Phibbs F, Hedera P, Davis TL, Validating an objective video-Based dyskinesia severity score in Parkinson's disease patients, Parkinsonism and Related Disorders, In Review.

## I. Introduction

Levodopa therapy in Parkinson's disease patients results in drug- induced dyskinesia characterized by hyperkinetic involuntary movements that may often interfere with activities of daily living [1]. Despite current treatment measures, the disabling symptoms of dyskinesia continue to challenge the development of better pharmacological and surgical interventions. The efficacy of these treatments can be evaluated by reliable qualitative and quantitative clinical assessment of patients. In this context, rating scales have been the most established and widely used means of assessment of the severity of dyskinesia. The key attributes of dyskinesia include anatomical distribution, phenomenology, duration, intensity, disability, and patient perception [32]. Different scales base their severity ratings on different sets of attributes of dyskinesia. The most recently developed scale is the Unified Dyskinesia Rating Scale (UDysRS), which may become the standardized dyskinesia rating scale equivalent to the UPDRS scale for Parkinson's disease symptoms [35]. The UDysRS is a five part scale that assesses general patient perceptions, patient perception with Off Dystonia impact, objective impairment based on anatomical distribution and intensity while performing four specific tasks and the objective disability based on the Rush Dyskinesia Rating Scale. A total objective

score is computed based on the impairment and disability scores. The UDysRS is a combination of several rating scales in such a way that all attributes of dyskinesia are assessed using a single rating scale. The results of the clinimetric testing of this scale over a range of 70 patients indicated an inter-rater and intra-rater reliability with correlation coefficients ranging from 0.37 to 0.87 for various tasks. Further validation and responsiveness testing is underway.

We have already stated the disadvantages of rating scales and stressed the need for an objective quantitative measure of dyskinesia. All rating scales are based on a discrete five point scale. This lack of resolution leads to the possibility of misclassifying patients with symptoms that fall in between two rating intervals. These factors encourage the development of a quantitative measure that is based on a continuous scale. In chapter 1 we have discussed previous work in quantitative assessment techniques. In chapter 2, we have described our work, which is an example of a video-based, marker-less, model-free human motion tracking and estimation problem using the standardized UPDRS videos of Parkinson's disease patients which are generally a part of the patient's clinical records. A severity score (SVS) developed using this technique is assigned to each video sequence to describe the severity of dyskinesia exhibited by the patient [54]. Extensive validation studies were performed on SVS and we discuss these studies in detail in this chapter. We carried out three studies to this effect. In chapter 2, only a small cohort of patients was evaluated using SVS. The VUMC videos were of poor quality with lighting defects, noisy background and the distance between the camera and the patient constantly changing. Our goal was to evaluate SVS on a larger number of patients using longer video sequences. The UDysRS study [35] carried by Dr. Christopher Goetz of the Rush

University Medical Center (RUMC), utilized 70 patient videos evaluated by 20 neurologists all over the country. A sample of 35 videos from this study was used to validate SVS. Due to good quality of these videos, longer sequences could be extracted for analysis Though UDysRS is a five-point rating scale, the average of the total ratings of twenty neurologists resembles a continuous score and hence, for our comparison purposes, we will refer to them as UDysRS scores, which can be ranked similarly to the SVS scores. As a continuous variable, SVS could not be directly compared to the discrete rating scales. The ranking scheme used in chapter 2 could not be used for the larger set of patients. Hence, we developed a rating based ranking protocol that can be used by the neurologists' to rate and rank these patients. By comparing the neurologists' and UDysRS rankings to the SVS rankings, we validate the utility of SVS. The ranking protocol itself is validated by studying the rater variability while using the protocol. Additionally, we also observed the effect of converting SVS into a discrete rating and compared it with neurologists' ratings. Thus, we acquired the following parameters for each patient in the study: 1) neurologist ratings and rankings obtained using the new ranking protocol, 2) UDySRS rankings computed from the UDysRS score, and 3) SVS rankings and SVS ratings after the scores were thresholded. In order to analyze these parameters comprehensively, three studies were performed: (a) Validation of Ranking protocol: analysis of intra- and inter-neurologists agreement while using the ranking protocol; (b) Utility of SVS – comparing SVS vs. neurologists' and UDysRS rankings, evaluating the robustness of SVS to variation in landmark points and length of video sequences; and (c) Effect of ratings vs. rankings. Statistical analysis using Kendall's Tau-b correlation coefficient and intra-class correlation coefficient were performed to test our

hypotheses. Our results indicate that the ranking protocol is effective in validating SVS; and that SVS is a robust score that correlates moderately with neurologists' and the UDysRS rankings. We finally discuss the rationale of preferring a continuous score such as SVS over a discrete score.

## II. Methods

Our analysis used a dataset of 35 patient videos with varying dyskinesia severity obtained as part of the extensive clinimetric testing of the UDysRS [35]. The videos were captured in a controlled environment with plain backgrounds in a well-illuminated room with no occluding furniture. Details regarding the video protocol and the informed consent obtained from patients have been previously published. The patients were rated based on four tasks that are activities of daily living (ADL). Part IV UDysRS scores for each task was also available. Our task of interest was the communication task, where the patients were asked to read while seated on a chair. The communication task was the simplest task to track using our semi-automatic technique. Though speech disorders were primarily rated using this task, the patients also exhibited movement of the face, head, neck, hands and legs. Additionally, we observed that the patients with severely impaired speech also showed dyskinetic movements of these body parts. The average length of the communication task was one minute. A 10s excerpt from the middle of each 60s sequence was analyzed using our semi-automatic technique discussed in [54]. A segment from the middle of the video was chosen to avoid the effects of any starting and stopping movements. The patient's head, shoulders, chest, forearms, knees, feet, and the reading material were manually selected as seen in Figure 3.1 and semi-automatically tracked using the Adaptive Bases Algorithm (ABA), which is an intensity-based non-rigid image

registration algorithm [27]. Since it was not possible to track the forearms of the patients without occlusion problem, the reading material held in the patient's hands was tracked. The tracked anatomical points of interests were analyzed by applying PCA on the cluster of points from every frame of the video sequence as described in our prior work in [54]. A severity score was computed for each video sequence using the parameters obtained from the PCA analysis

$$SVS = (TV * NSM) / (STDEV * NF)$$

TV: total variance of all eigenmodes, where the total variance is the sum of the magnitude of all the eigenvalues.

NSM: Number of significant modes of variation, which defines the number of modes of variations that capture 90% of the variations in the patient movements

STDEV: standard deviation of the percentage contribution of the eigenvalues to the total variance, which represents the rate of fall of eigenvalues. A gradual fall of eigenvalues indicates complex patient movements in various directions and a steep fall indicates simple movements in fewer directions.

NF: Number of frames in the video sequence

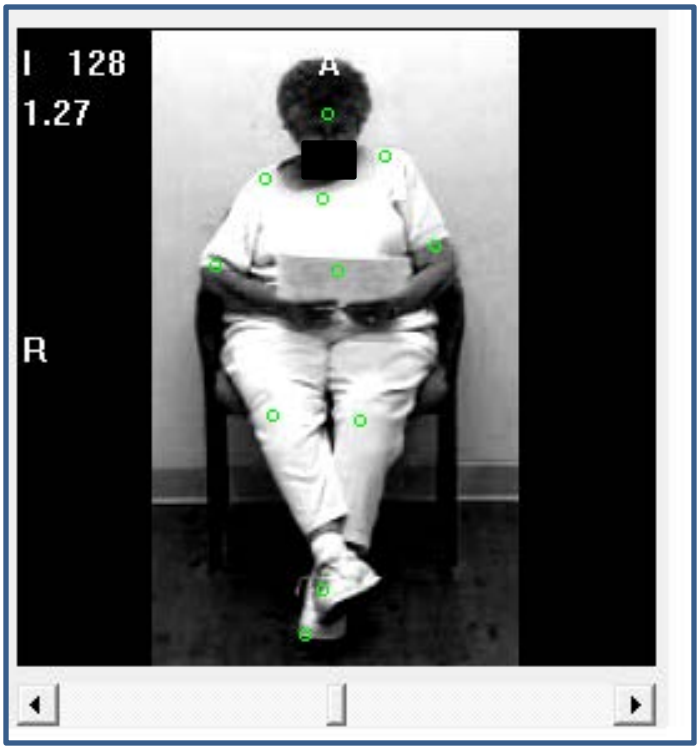


**Fig. 3.1**: Landmark points (in green circles) are chosen on the head, shoulders, chest, elbows, knees, feet and the reading material. These anchors are tracked using intensity-based registration.

The 10s sequences were ranked in the increasing order of SVS. These videos were also ranked based in the increasing order of the part IV UDysRS communication task scores. Four movement disorder neurologists, N1, N2, N3, and N4, from the Vanderbilt University Medical Center ranked the 35 ten seconds video sequences based on the clinical definition of dyskinesia. We developed a rating based ranking protocol to accomplish this task. The attributes used by the neurologists included amplitude and speed of dyskinetic movement, anatomical distribution of dyskinesia and the extent of disability seen in the patient. Dyskinesia in the patient's head, arms, trunk and legs was observed. The speech disability caused by dyskinesia was excluded from the protocol, but was accounted for in the UDysRS score. Each neurologist independently rated and

ranked the 35 video segments using this ranking protocol. Three sets of such ratings and rankings on the same dataset were obtained with a time lapse of one month between rankings to ensure the ratings and rankings were not voluntarily repeated. Thus, each neurologist had three sets of ratings and rankings – ratings: SET IR, SET IIR, and SET IIIR and rankings: SET Ir, SET IIr, and SET IIIr.

**II.1 Rating based ranking protocol**

(a) Ratings: The videos were first rated on a scale of one to four with one representing no dyskinesia, two - mild dyskinesia, three - moderate dyskinesia and four - severe dyskinesia.

(b) Rankings: The videos in each rating category, except the no dyskinesia category, were viewed simultaneously on a single screen and ranked according to increasing order of severity within that category.

(c) The first two and the final two videos in each rating category were then compared with the correspondingly ranked videos of the immediately next category to confirm that these rankings were still valid. Thus the neurologist could view videos across categories to finalize their ranks.

(d) In case of rank changes, steps (b) and (c) were repeated until ranks were finalized and the corresponding rating categories in step (a) were also modified to ensure coherence between ratings and ranking.

Each patient in the study had the following parameters: three sets of neurologist ratings and rankings, UDysRS ranking and SVS ranking.

### II.2 Data Analyses

Three studies were conducted from the data obtained using the above methods.

(a) Validation of ranking protocol: Evaluation of intra- and inter-neurologist ranking consistency.

(b) SVS Utility: (i) Robustness of SVS to longer video sequences and variations in landmark points, and (ii) evaluation of (SVS rankings vs. UDysRS rankings) and (SVS rankings vs. neurologists' rankings).

(c) Effect of ratings vs. rankings: Comparison of SVS ratings and rankings with neurologists' ratings and rankings.

The original rankings obtained from neurologists were modified as follows to permit statistical analysis because the number of non-dyskinetic patients (neurologist rating of 1) was different in both the inter- and intra-neurologist ratings. Hence the total number of patients ranked by each neurologist was not necessarily equal. To ensure statistical consistency in the analyses, for each set of rankings, two types of rank data sets were developed.

(1) **Type I**: All 35 video sequences were part of this dataset. A tied rank was assigned to non-dyskinetic patients such that its value is the average of the ranks the patients would have received if there were given distinct ranks [48]. This process ensured the maximum ranking in each ranking was 35, but the minimum rank would depend on the number of non-dyskinetic patients.

(2) **Type II:** Seven patients were consistently labeled non-dyskinetic by the senior neurologist in all the three rank sets. These patients were uniformly eliminated from the original rank sets of all the neurologists and the remaining 28 patients

were re-ranked keeping the order unchanged. UDysRS and SVS rankings were also modified accordingly.

Table 3.1 shows an example of Type I and Type II rankings of SET Ir of neurologists N1, N2, N3, and N4.

### II.2.1 Study I: Validation of ranking protocol: intra- and inter-neurologist agreement

Step (a) of the ranking protocol was based on the clinical definition of dyskinesia and not a specific rating scale. The ranking protocol was developed to facilitate the comparison of discrete neurologists' ratings in Step (a) to the continuous SVS score. By evaluating the intra- and inter-neurologist consistency in using the ranking protocol, the validity of the protocol can be determined. A high intra- and inter- neurologists consistency indicates that the protocol, based on clinical definition of dyskinesia, can be used to rank severity of dyskinesia by neurologists and in turn can be used to evaluate the utility of SVS. Independent analyses were performed to observe intra- and inter-neurologist agreement on the Type I and Type II ranking datasets. Kendall's tau-b correlation coefficient was computed pairwise between the four neurologists in each type to evaluate the inter-neurologist agreement [48]. Intra-class coefficient for each neurologist across SET Ir, SET IIr, and SET IIIr was computed to study the intra-neurologist agreement. [52].

| PN | N1 | N2 | N3 | N4 |
|---|---|---|---|---|
| 1 | 16 | 17 | 14 | 21 |
| 2 | 1 | 4 | 2 | 7 |
| 3 | 20 | 19 | 19 | 20 |
| 4 | 3 | 5 | 5 | 9 |
| 5 | 25 | 21 | 21 | 28 |
| 6 | 2 | 2 | 4 | 8 |
| 7 | 13 | 14 | 7 | 17 |
| 8 | 5 | 6 | 8 | 14 |
| 9 | 12 | 10 | 10 | 11 |
| 10 | 0 | 0 | 0 | 0 |
| 11 | 26 | 24 | 24 | 29 |
| 12 | 11 | 13 | 6 | 15 |
| 13 | 15 | 22 | 15 | 23 |
| 14 | 0 | 0 | 0 | 5 |
| 15 | 28 | 28 | 27 | 32 |
| 16 | 9 | 8 | 9 | 12 |
| 17 | 4 | 3 | 3 | 10 |
| 18 | 0 | 0 | 0 | 4 |
| 19 | 17 | 15 | 16 | 22 |
| 20 | 6 | 7 | 1 | 6 |
| 21 | 0 | 0 | 0 | 0 |
| 22 | 27 | 27 | 26 | 31 |
| 23 | 21 | 16 | 17 | 18 |
| 24 | 7 | 1 | 0 | 3 |
| 25 | 0 | 0 | 0 | 2 |
| 26 | 0 | 0 | 0 | 0 |
| 27 | 23 | 20 | 20 | 26 |
| 28 | 19 | 25 | 23 | 30 |
| 29 | 14 | 12 | 13 | 16 |
| 30 | 22 | 23 | 22 | 19 |
| 31 | 18 | 18 | 18 | 25 |
| 32 | 8 | 11 | 11 | 13 |
| 33 | 0 | 0 | 0 | 1 |
| 34 | 10 | 9 | 12 | 24 |
| 35 | 24 | 26 | 25 | 27 |

Raw rankings

| PN | N1 | N2 | N3 | N4 |
|---|---|---|---|---|
| 1 | 23 | 24 | 22 | 24 |
| 2 | 8 | 11 | 10 | 10 |
| 3 | 27 | 26 | 27 | 23 |
| 4 | 10 | 12 | 13 | 12 |
| 5 | 32 | 28 | 29 | 31 |
| 6 | 9 | 9 | 12 | 11 |
| 7 | 20 | 21 | 15 | 20 |
| 8 | 12 | 13 | 16 | 17 |
| 9 | 19 | 17 | 18 | 14 |
| 10 | 4 | 4 | 5 | 2 |
| 11 | 33 | 31 | 32 | 32 |
| 12 | 18 | 20 | 14 | 18 |
| 13 | 22 | 29 | 23 | 26 |
| 14 | 4 | 4 | 5 | 8 |
| 15 | 35 | 35 | 35 | 35 |
| 16 | 16 | 15 | 17 | 15 |
| 17 | 11 | 10 | 11 | 13 |
| 18 | 4 | 4 | 5 | 7 |
| 19 | 24 | 22 | 24 | 25 |
| 20 | 13 | 14 | 9 | 9 |
| 21 | 4 | 4 | 5 | 2 |
| 22 | 34 | 34 | 34 | 34 |
| 23 | 28 | 23 | 25 | 21 |
| 24 | 14 | 8 | 5 | 6 |
| 25 | 4 | 4 | 5 | 5 |
| 26 | 4 | 4 | 5 | 2 |
| 27 | 30 | 27 | 28 | 29 |
| 28 | 26 | 32 | 31 | 33 |
| 29 | 21 | 19 | 21 | 19 |
| 30 | 29 | 30 | 30 | 22 |
| 31 | 25 | 25 | 26 | 28 |
| 32 | 15 | 18 | 19 | 16 |
| 33 | 4 | 4 | 5 | 4 |
| 34 | 17 | 16 | 20 | 27 |
| 35 | 31 | 33 | 33 | 30 |

Type I Ranks

| PN | N1 | N2 | N3 | N4 |
|---|---|---|---|---|
| 1 | 16 | 17 | 15 | 17 |
| 2 | 1 | 4 | 3 | 3 |
| 3 | 20 | 19 | 20 | 16 |
| 4 | 3 | 5 | 6 | 5 |
| 5 | 25 | 21 | 22 | 24 |
| 6 | 2 | 2 | 5 | 4 |
| 7 | 13 | 14 | 8 | 13 |
| 8 | 5 | 6 | 9 | 10 |
| 9 | 12 | 10 | 11 | 7 |
| 11 | 26 | 24 | 25 | 25 |
| 12 | 11 | 13 | 7 | 11 |
| 13 | 15 | 22 | 16 | 19 |
| 15 | 28 | 28 | 28 | 28 |
| 16 | 9 | 8 | 10 | 8 |
| 17 | 4 | 3 | 4 | 6 |
| 19 | 17 | 15 | 17 | 18 |
| 20 | 6 | 7 | 2 | 2 |
| 22 | 27 | 27 | 27 | 27 |
| 23 | 21 | 16 | 18 | 14 |
| 24 | 7 | 1 | 1 | 1 |
| 27 | 23 | 20 | 21 | 22 |
| 28 | 19 | 25 | 24 | 26 |
| 29 | 14 | 12 | 14 | 12 |
| 30 | 22 | 23 | 23 | 15 |
| 31 | 18 | 18 | 19 | 21 |
| 32 | 8 | 11 | 12 | 9 |
| 34 | 10 | 9 | 13 | 20 |
| 35 | 24 | 26 | 26 | 23 |

Type II Ranks

**Table 3.1**. PN is the patient number and N1, N2, N3, and N4 represent the neurologists. The shaded rows in the raw and Type I rankings are the patients rated as non-dyskinetic by the respective neurologist.

### II.2.2 Study II: Validation of SVS Utility

#### II.2.2.a SVS vs. Neurologists, UDysRS

In Chapter 2 we have shown that the SVS is a good indicator of the severity of dyskinesia in a small sample set of patients. In this chapter we further test the validity of SVS, which is based on total movement of the patients, on a larger set of patient videos with a more stringent statistical measure of Kendalls Tau-b correlation coefficient. As opposed to the 2s video sequences we use 10s video sequences and track more number of body parts on the patients simultaneously. We compared the SVS scores to the UDysRS scores, which are used as a gold standard, by computing the Kendall's Tau-b correlation coefficient between SVS rankings and the UDysRS rankings. The SVS was also compared to neurologist's performance by computing the Kendall's Tau-b correlation coefficient between the SVS rankings and the neurologists' rankings. The statistical analysis was performed on Type I and Type II rankings of SET Ir, SET IIr and SET IIIr. A good correlation would indicate that the SVS can quantify dyskinesia as well as neurologists and can complement UDysRS scores by providing an objective dimension to it.

#### II.2.2.b Robustness of SVS

We study the robustness of SVS when using videos sequences longer than 10s and the effect of varying the positions of landmark points to be tracked. Patients in video sequences ranging in length from 20s to 30s were tracked using ABA and their SVS computed. The goal was to observe if 10s is a significant period of time to quantify dyskinesia and if there were significant changes in the SVS that affected the ranking of patients. The patients were generally filmed for one minute on the communication task.

**Variation in landmark points:**

Picking the landmark points is the only manual aspect of our quantifying technique which makes it semi-automatic. We have picked 11 landmark points – forehead, shoulders, chest, knees, feet and reading material as seen in Figure 3.1. Different neurologists' can pick these 11 points in the first frames with slight variations, paying attention to the fact that areas of intensity inhomogeneity are supportive of good tracking results. Landmark points were picked in three different trials with small variations in the positions as seen in Figure 3.2.

This study was done to emulate the scenario of three different neurologists picking the landmark points using a general guideline of their positions on each of the 35 patient video sequences. The SVS scores were computed for each trial set and the patients were ranked based on the SVS scores. The Kendall Tau-b correlation coefficient for each pair of rankings was computed. A high correlation coefficient would reflect the robustness of the SVS to the variation in the landmark points.

**SVS for longer video sequences**

In chapter 2, we quantified dyskinesia for short 2s video sequences. We wanted to observe if this short span of movement was sufficient to capture the actual severity of dyskinesia. In other words, even if longer video sequences were used, would the severity score remain the same? Two sets of analyses were performed to study this effect. Eight patients from the VUMC database were used as sample datasets covering a range of mild, moderate and severe dyskinesia. Since only 10s of usable video data was available in the

VUMC dataset, the tracking and SVS calculations were performed for a minimum of 2s and a maximum of 10s with an increment of 1s for each new trial.

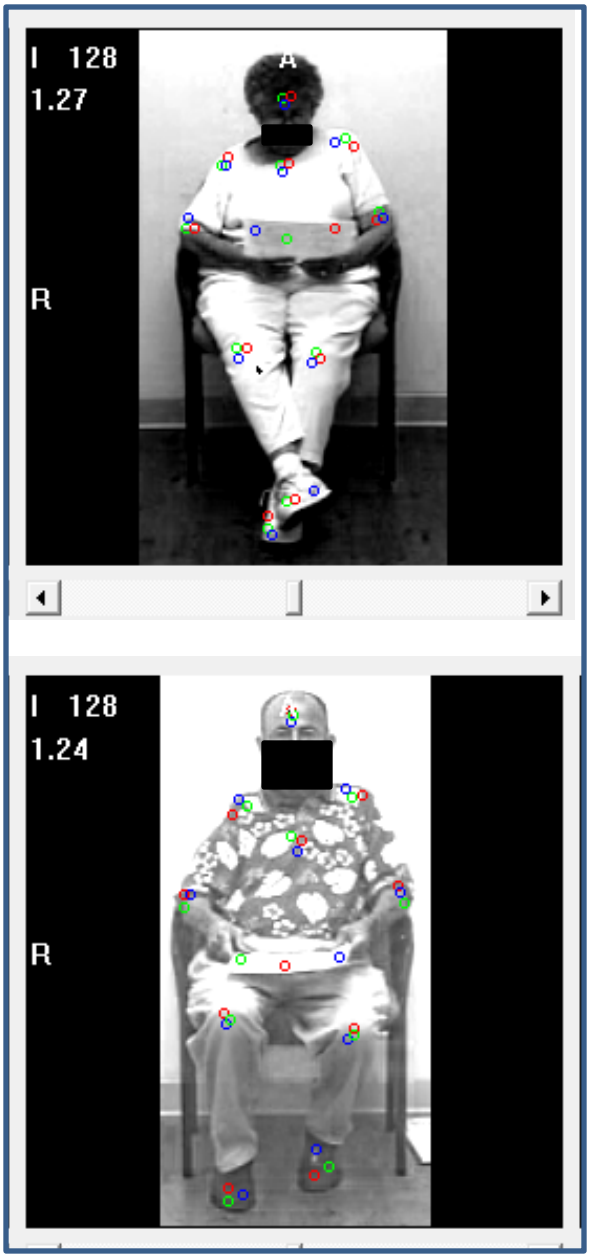


**Fig. 3.2**.Eleven landmark points plotted on the first frame in three different trials on two patientsThe three different trials are shown in three colors (red, green and blue circles)

As before, these videos focused only on the patient's upper body and hence only the head and shoulders were tracked and analyzed. We ensured that starting and stopping effects

were not included in both datasets. The trend of SVS values was observed with the increase in the length of the sequences. For a robust measure, we would expect the trend to plateau out beyond a particular time length making that a sufficient period to use for quantification purposes. Only a small sample of patients was used for this study as registration is a time consuming process. SVS is inherently dependent on the quality of tracking, poor tracking leading to incorrect measurement of movement and hence severity. Since tracking is registration based, and the accuracy of the points on the final frame is dependent not only the previous frame, but on all other frames preceding it. Once the landmark points drift in a given frame, it is difficult to bring them to the back to correct position without manual manipulation. Drifting can easily occur when intensity is homogenous in the vicinity of the tracked points, hence the requirement of picking points in areas of intensity inhomogeneity. Hence our goal was to observe the extent of variation in SVS when the quality of tracking declines.

### II.2.3 Study III: Effect of ratings vs. rankings

We propose that ranking the severity of dyskinesia within each rating category of mild, moderate or severe dyskinesia assists in quantifying the differences between patients at a finer level. It is easier for neurologists to use a discrete five point rating scale which allows them to assign more than one patient in a single category than to use a ranking system which compels them to observe differences in severity more closely in order to assign individual ranks to each patient. By thresholding the SVS appropriately and converting it to a discrete scale, we wanted to compare the discrete SVS ratings to the neurologists' discrete ratings in SET IR, SET IIR and SET IIIR. Based on the a-priori information that approximately an equal number of patients were present in each rating

category, the SVS was thresholded into four levels of severity – absent, mild, moderate and severe. Kendall Tau-b correlation coefficient was computed between these SVS ratings and the neurologists ratings. These correlation coefficients were compared to the correlation coefficients obtained in Study II which compares SVS rankings to neurologists' rankings. We would expect a higher correlation using the ratings than the rankings. In rankings, we compare the ordering of patients based on severity, whereas in ratings, groups of patients, irrespective of their ordering within the group, are compared. Our severity score SVS, in association with the ranking technique, can thus be used to complement rating scales to observe differences in dyskinesia severity in large patient databases.

## III. Results

### III.1 Study I: Validation of ranking protocol - intra- and inter-neurologist agreement

#### III.1.1 Intra- neurologist agreement

Intra-class correlation coefficient (ICC) was computed only for Type I rankings for each neurologist as this measure was used to study intra-neurologist consistency. Hence Type II rankings, which were developed mainly for inter-rater studies, would not apply in this experiment. High ICC values were observed as follows: N1 – 0.9525; N2 – 0.948; N3 – 0.9496; N4 – 0.9928 ($p <= 0.0001$).

A more detailed analysis was performed by computing pairwise Kendall tau-b values between the three sets of rankings for each neurologist and between the neurologists' rankings and the UDysRS rankings. Table 3.2 indicates that neurologist N1 and N4 are

more consistent across the three sets of rankings compared to neurologist N2 and N3 which was also observed from the ICCs.

| N1 | Set I | Set II | Set III |
|---|---|---|---|
| Set I | 1 | 0.9199 | 0.9129 |
| Set II | 0.9199 | 1 | 0.8815 |
| Set III | 0.9129 | 0.8815 | 1 |

| N2 | Set I | Set II | Set III |
|---|---|---|---|
| Set I | 1 | 0.849 | 0.8128 |
| Set II | 0.849 | 1 | 0.8017 |
| Set III | 0.8128 | 0.8017 | 1 |

| N3 | Set I | Set II | Set III |
|---|---|---|---|
| Set I | 1 | 0.8642 | 0.7941 |
| Set II | 0.8642 | 1 | 0.8258 |
| Set III | 0.7941 | 0.8258 | 1 |

| N4 | Set I | Set II | Set III |
|---|---|---|---|
| Set I | 1 | 0.9628 | 0.9899 |
| Set II | 0.9628 | 1 | 0.9595 |
| Set III | 0.9899 | 0.9595 | 1 |

**Table 3.2** Kendall tau-b Correlation matrix for set-wise intra-neurologist agreement.

### III.1.2 Inter-neurologist agreement

Though the neurologists independently ranked the videos based on the clinical definition of dyskinesia, high inter-neurologist agreement was observed between them. Type I and Type II rankings showed similar trends of agreement as seen in Table 3.3 and Table 3.4. Figures 3.3 and 3.4 are correlation plots of the values represented in Tables 3.3 and 3.4. Kendall tau-b values ranged from 0.7754 to 0.8746 for Type I rankings, and from 0.6931 to 0.8095 for Type II rankings. All tau values were statistically significant with $p <= 0.0001$.

| Type I, SET I | N1 | N2 | N3 | N4 |
|---|---|---|---|---|
| N1 | 1 | 0.8746 | 0.8642 | 0.7994 |
| N2 | 0.8746 | 1 | 0.8852 | 0.8337 |
| N3 | 0.8642 | 0.8852 | 1 | 0.8371 |
| N4 | 0.7994 | 0.8337 | 0.8371 | 1 |
| Type I, SET II | N1 | N2 | N3 | N4 |
| N1 | 1 | 0.8456 | 0.8223 | 0.8063 |
| N2 | 0.8456 | 1 | 0.8283 | 0.797 |
| N3 | 0.8223 | 0.8283 | 1 | 0.8097 |
| N4 | 0.8063 | 0.797 | 0.8097 | 1 |
| Type I, SET III | N1 | N2 | N3 | N4 |
| N1 | 1 | 0.844 | 0.8293 | 0.7857 |
| N2 | 0.844 | 1 | 0.799 | 0.8021 |
| N3 | 0.8293 | 0.799 | 1 | 0.7754 |
| N4 | 0.7857 | 0.8021 | 0.7754 | 1 |

**Table 3.3** Kendall Tau-b correlation matrix for Type I inter-neurologists agreement

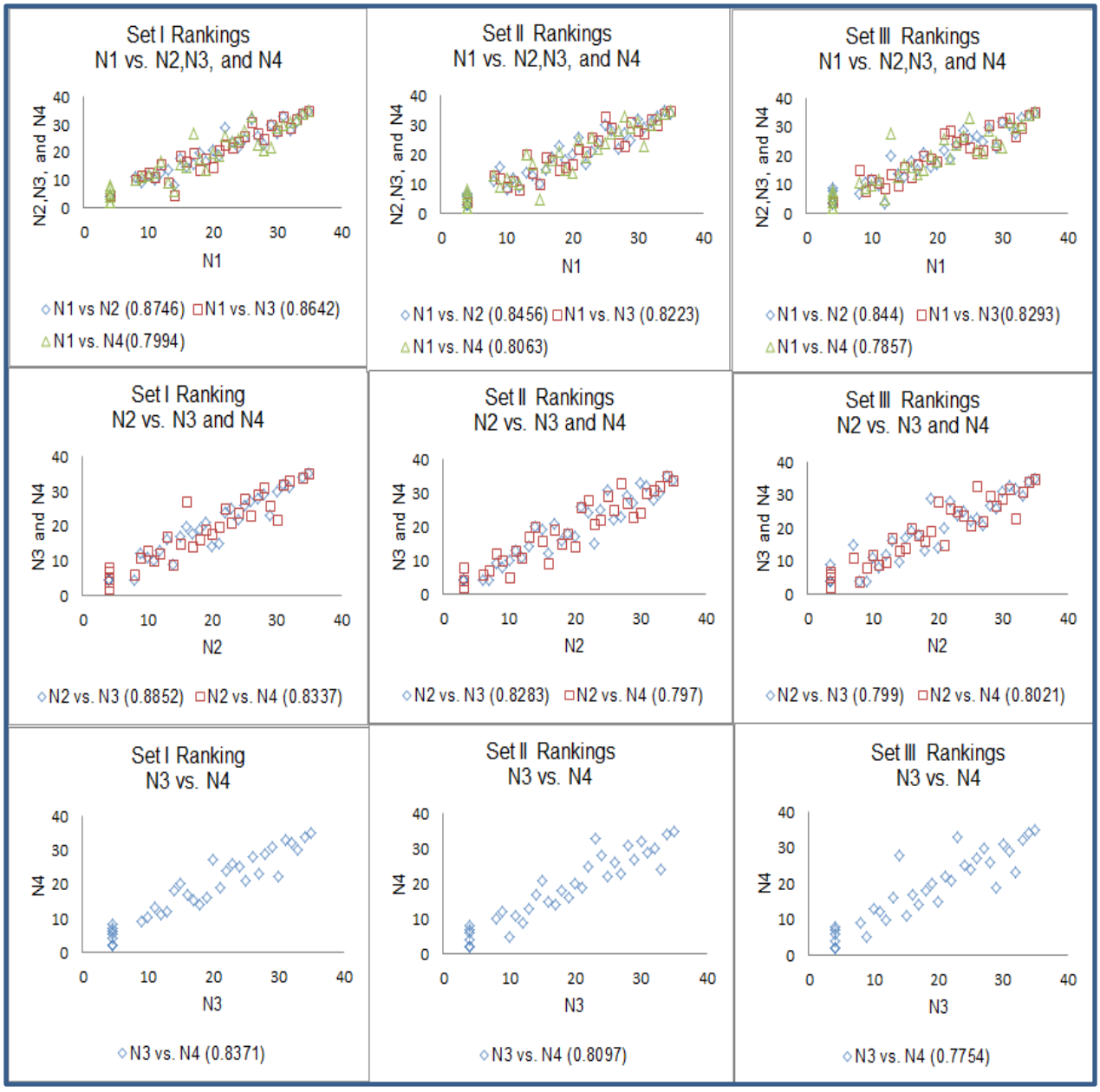


**Fig. 3.3.** Scatter plots showing inter-neurologist Type I ranking correlations. (First row) N1 vs. N2, N3 and N4; (second row) N2 vs. N3 and N4; (third row) N3 vs. N4 Kendall tau-b coefficient is indicated in parenthesis

| Type II, SET I | N1 | N2 | N3 | N4 |
|---|---|---|---|---|
| N1 | 1 | 0.8095 | 0.8042 | 0.7249 |
| N2 | 0.8095 | 1 | 0.836 | 0.7778 |
| N3 | 0.8042 | 0.836 | 1 | 0.7831 |
| N4 | 0.7249 | 0.7778 | 0.7831 | 1 |
| Type II, SET II | N1 | N2 | N3 | N4 |
| N1 | 1 | 0.7778 | 0.7302 | 0.7407 |
| N2 | 0.7778 | 1 | 0.7513 | 0.7196 |
| N3 | 0.7302 | 0.7513 | 1 | 0.746 |
| N4 | 0.7407 | 0.7196 | 0.746 | 1 |
| Type II, SET III | N1 | N2 | N3 | N4 |
| N1 | 1 | 0.8042 | 0.7407 | 0.709 |
| N2 | 0.8042 | 1 | 0.7345 | 0.735 |
| N3 | 0.7407 | 0.7345 | 1 | 0.6931 |
| N4 | 0.709 | 0.735 | 0.6931 | 1 |

**Table 3.4.** Kendall Tau-b correlation matrix for Type II inter-neurologist agreement

## III. 2 Study II: Validation of SVS Utility

### III.2.1 SVS vs. neurologists' and UDysRS rankings

Table 3.5 and Figure 3.5 show that Type I rankings across the three ranking sets exhibited higher Kendall tau-b values than Type II rankings. The tau values of the UDysRS rankings vs. SVS rankings and the neurologist's rankings vs. SVS rankings were comparable and indicated a moderate utility of the SVS as a dyskinesia quantifying score. The tau values ranged from 0.657166 to 0.7563 for Type I rankings, and from 0.5561 to 0.7143 for Type II rankings. All tau values were statistically significant with p <=0.0001.

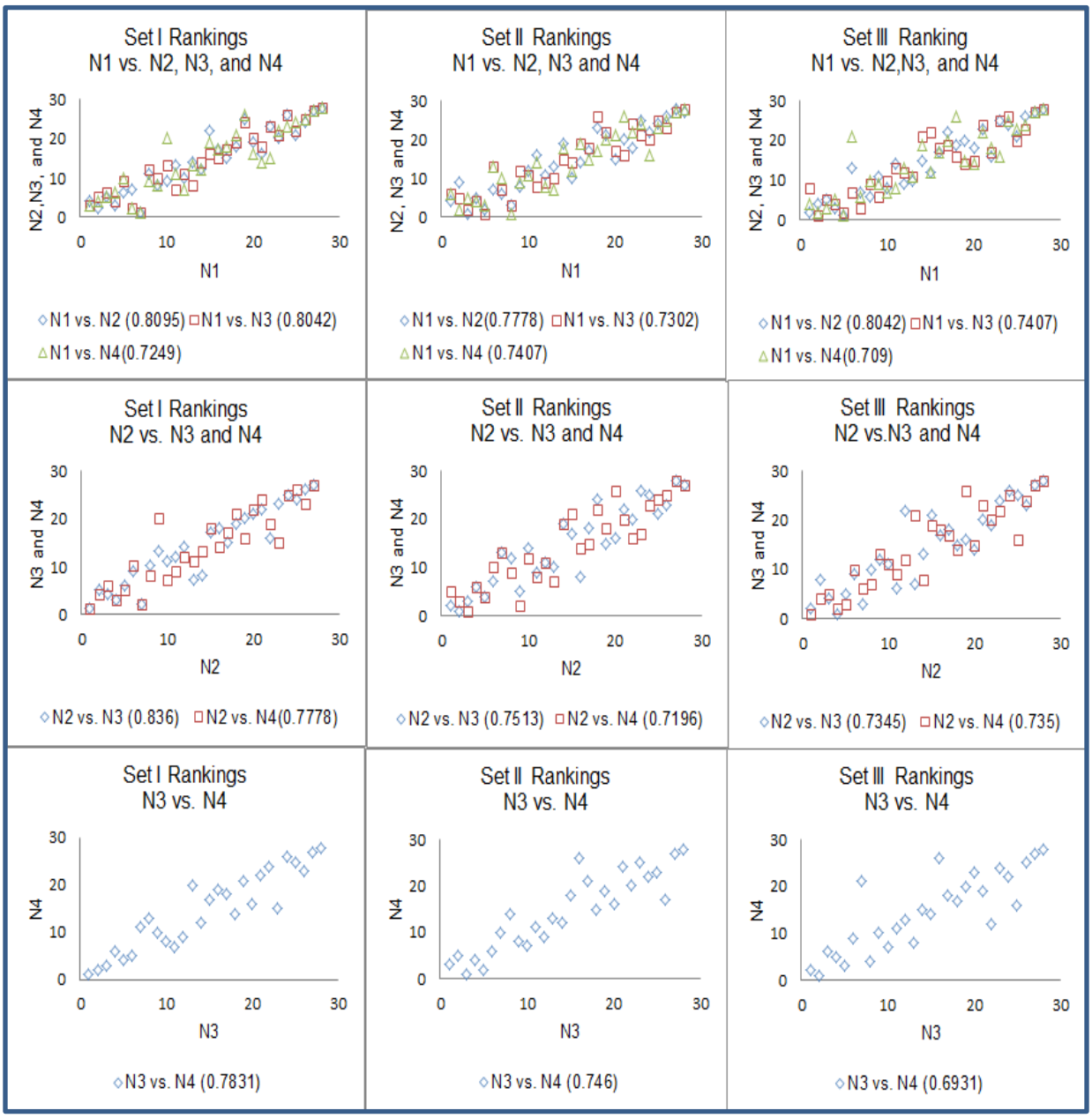


**Fig. 3.4.** Scatter plots showing inter-neurologist Type II ranking correlations. (First row) N1 vs. N2, N3 and N4; (second row) N2 vs. N3 and N4; (third row) N3 vs. N4 Kendall tau-b coefficient is indicated in parenthesis

| **Type I, SET I** | N1 | N2 | N3 | N4 | UDysRS | SVS |
|---|---|---|---|---|---|---|
| UDysRS | 0.7563 | 0.7632 | 0.7386 | 0.7313 | 1 | 0.6235 |
| SVS | 0.6879 | 0.64 | 0.6422 | 0.6066 | 0.6235 | 1 |
| **Type I, SET II** | N1 | N2 | N3 | N4 | UDysRS | SVS |
| UDysRS | 0.7461 | 0.7712 | 0.7563 | 0.7548 | 1 | 0.6235 |
| SVS | 0.6845 | 0.6797 | 0.6947 | 0.6234 | 0.6235 | 1 |
| **Type I, SET III** | N1 | N2 | N3 | N4 | UDysRS | SVS |
| UDysRS | 0.7871 | 0.7524 | 0.7495 | 0.7211 | 1 | 0.6235 |
| SVS | 0.6981 | 0.6503 | 0.6194 | 0.5965 | 0.6235 | 1 |
| | | | | | | |
| **Type II, SET I** | N1 | N2 | N3 | N4 | UDysRS | SVS |
| UDysRS | 0.6508 | 0.6614 | 0.6349 | 0.6296 | 1 | 0.5238 |
| SVS | 0.5979 | 0.5238 | 0.5397 | 0.4921 | 0.5238 | 1 |
| **Type II, SET II** | N1 | N2 | N3 | N4 | UDysRS | SVS |
| UDysRS | 0.6349 | 0.6984 | 0.6508 | 0.672 | 1 | 0.5238 |
| SVS | 0.5926 | 0.5714 | 0.6085 | 0.5238 | 0.5238 | 1 |
| **Type II, SET III** | N1 | N2 | N3 | N4 | UDysRS | SVS |
| UDysRS | 0.6984 | 0.6614 | 0.6402 | 0.619 | 1 | 0.5238 |
| SVS | 0.6138 | 0.5979 | 0.4921 | 0.4815 | 0.5238 | 1 |

**Table 3.5**: Kendall Tau-b correlation matrix for Type I and Type II SVS vs. neurologists SVS vs. UDysRS **Rankings**.

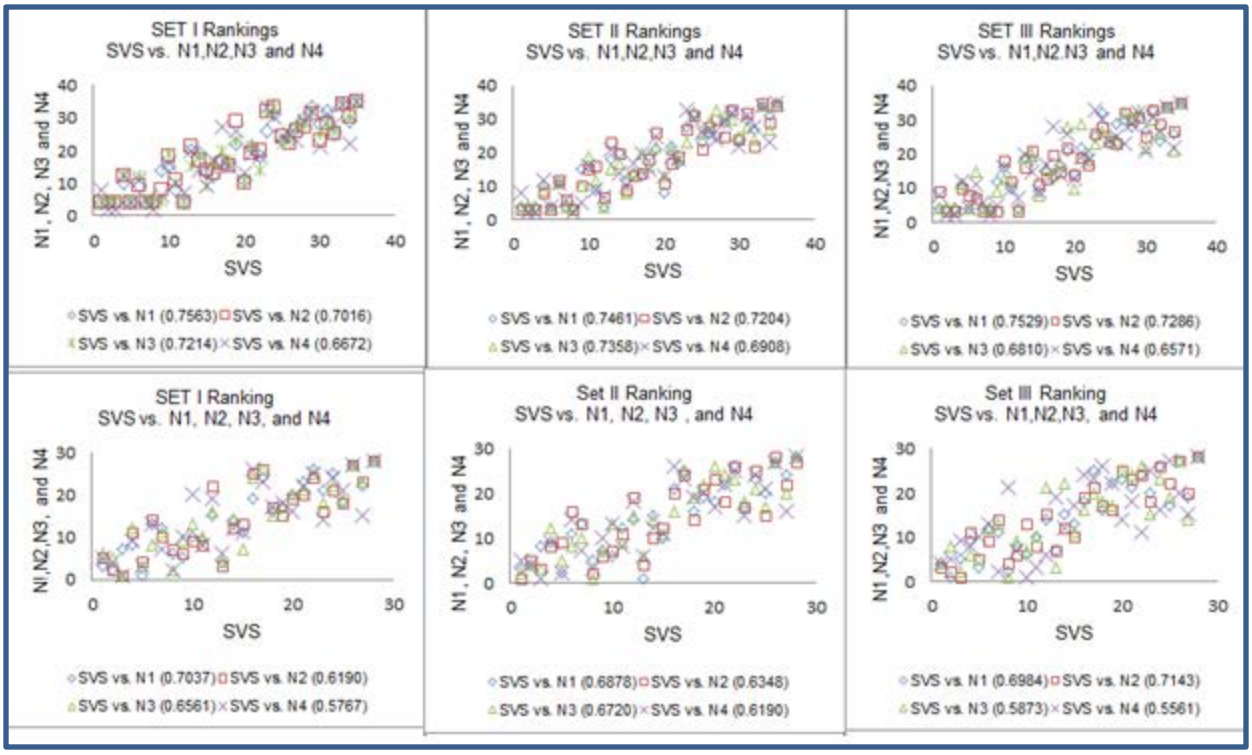


**Fig. 3.5** Scatter plots showing correlation between SVS and Type I neurologist' rankings (first row); SVS and Type II neurologists' rankings (second row); Kendall tau-b coefficient is indicated in parenthesis.

### III.2.2 Robustness of SVS

#### Variations in landmark points

From Table 3.6 we can see that SVS does not change significantly, but given that it is a continuous score, even small changes may lead to changing in the ranking order. Hence we compare the rankings of the score as opposed to the absolute score itself. The Kendall Tau-b correlation coefficients for the pairwise correlation between the rankings of each trial are seen in Table 3.7. In spite of being a conservative statistical measure, we do achieve high correlation between the rankings, which indicate the robustness of the score with respect to variations in the landmark points. Rows highlighted in Table 3.6 indicate

rankings which varying by more than an absolute difference of 4 ranks computed pairwise.

| PN | SVS - I | SVS - II | SVS -III | Rank I | Rank II | Rank III |
|---|---|---|---|---|---|---|
| 1 | 0.146949 | 0.113028 | 0.154349 | 23 | 21 | 23 |
| 2 | 0.1274 | 0.1404 | 0.1813 | **20** | **24** | **24** |
| 3 | 0.306498 | 0.32367 | 0.348641 | 31 | 32 | 32 |
| 4 | 0.010792 | 0.012624 | 0.007366 | 4 | 5 | 3 |
| 5 | 0.279027 | 0.273698 | 0.313936 | 30 | 31 | 31 |
| 6 | 0.012509 | 0.012509 | 0.010775 | 6 | 4 | 5 |
| 7 | 0.018397 | 0.02522 | 0.021519 | 11 | 12 | 10 |
| 8 | 0.050917 | 0.042751 | 0.046427 | 16 | 14 | 16 |
| 9 | 0.02149 | 0.022977 | 0.026919 | 12 | 11 | 12 |
| 10 | 0.011539 | 0.016827 | 0.021835 | **5** | **8** | **11** |
| 11 | 0.261305 | 0.24024 | 0.250256 | 29 | 28 | 30 |
| 12 | 0.132309 | 0.111467 | 0.111984 | 21 | 20 | 19 |
| 13 | 0.066233 | 0.067893 | 0.06744 | 18 | 18 | 18 |
| 14 | 0.001568 | 0.001968 | 0.00159 | 1 | 1 | 1 |
| 15 | 4.41795 | 3.951075 | 4.225237 | 35 | 35 | 35 |
| 16 | 0.0484 | 0.047213 | 0.040973 | 14 | 15 | 15 |
| 17 | 0.0657 | 0.0558 | 0.0296 | **17** | **17** | **13** |
| 18 | 0.0163 | 0.0162 | 0.0169 | 7 | 7 | 8 |
| 19 | 0.208416 | 0.196646 | 0.211009 | 26 | 25 | 27 |
| 20 | 0.048477 | 0.053404 | 0.054712 | 15 | 16 | 17 |
| 21 | 0.0074 | 0.0089 | 0.0086 | 2 | 3 | 4 |
| 22 | 0.696339 | 0.713673 | 0.578843 | 34 | 34 | 33 |
| 23 | 0.31232 | 0.268038 | 0.23405 | **32** | **30** | **28** |
| 24 | 0.089416 | 0.09219 | 0.132954 | 19 | 19 | 21 |
| 25 | 0.018267 | 0.018217 | 0.019524 | 10 | 9 | 9 |
| 26 | 0.007551 | 0.004943 | 0.006646 | 3 | 2 | 2 |
| 27 | 0.237275 | 0.245249 | 0.244092 | 27 | 29 | 29 |
| 28 | 0.16303 | 0.133933 | 0.124829 | **24** | **23** | **20** |
| 29 | 0.138021 | 0.114222 | 0.146501 | 22 | 22 | 22 |
| 30 | 0.618898 | 0.541843 | 0.727269 | 33 | 33 | 34 |
| 31 | 0.244618 | 0.218102 | 0.196831 | **28** | **27** | **25** |
| 32 | 0.0172 | 0.02 | 0.014 | 8 | 10 | 7 |
| 33 | 0.017923 | 0.012832 | 0.01335 | 9 | 6 | 6 |
| 34 | 0.044021 | 0.042402 | 0.036209 | 13 | 13 | 14 |
| 35 | 0.205827 | 0.216054 | 0.197191 | 25 | 26 | 26 |

**Table 3.6**. SVS scores and corresponding ranking for the three sets of trials.

| SVS | Trial 1 | Trial 2 | Trial 3 |
|---|---|---|---|
| **Trial 1** | 1 | 0.9294 | 0.8924 |
| **Trial 2** | 0.9294 | 1 | 0.916 |
| **Trial 3** | 0.8924 | 0.916 | 1 |

**Table 3.7.** Kendall Tau-b correlation matrix between the SVS rankings obtained from the three sets of landmark trials.

**Length of video sequences**

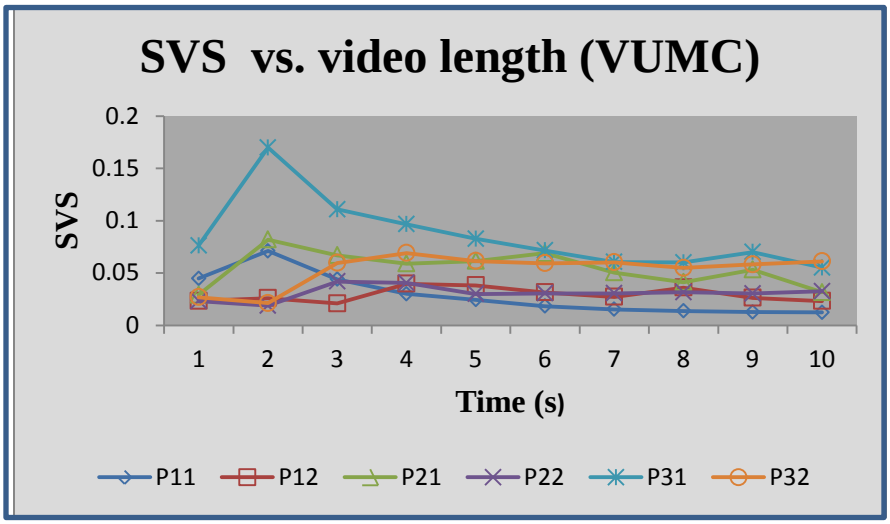


**Fig 3.6** SVS of 6 VUMC patients with varying lengths of video sequences, P11,P12: patients with dyskinesia rank of 1; P21,P22: rank 2, P31,P32: rank 3. All rankings are by expert neurologist.

**P41**

| Time (s) | SVS |
|---|---|
| 1 | 3.803667 |
| 2 | 1.471167 |
| 3 | 1.152444 |
| 4 | 0.837583 |
| 5 | 0.701133 |
| 6 | 0.926313 |
| 7 | 0.676476 |
| 8 | **0.59571** |
| 9 | **0.55248** |
| 10 | **0.53253** |

**P42**

| Time (s) | SVS |
|---|---|
| 1 | 5.5784733 |
| 2 | 3.0322217 |
| 3 | 1.9084667 |
| 4 | **1.499848** |
| 5 | **1.226791** |
| 5 and 1/3 | **1.137157** |

**Table 3.8** SVS values of two severely dyskinetic patients. P41, P42: rank 4

Figure 3.6 shows the variation of SVS with video sequences of different lengths. Due to the higher range of SVS values for the two severely dyskinetic patients, we have shown their SVS values in Table 3.8. Figure 3.6 indicates that beyond 7s, the SVS values tend to remain more or less constant and a similar effect is seen in Table 3.8. Sufficient data was not available for patient P42 to test with longer sequences.

### III.3 Study III: Effect of ratings vs. rankings

Table 3.9 shows that Kendall tau-b coefficients ranged from 0.6833 to 0.8519 for SVS ratings vs. neurologists' ratings which are higher when compared to the tau-b coefficients obtained in Study II while comparing rankings. These higher values indicate a more global agreement between the neurologists and the SVS.

| SVS | N1 | N2 | N3 | N4 |
|---|---|---|---|---|
| Set I | 0.8359 | 0.7369 | 0.7793 | 0.6833 |
| Set II | 0.7913 | 0.773 | 0.789 | 0.7155 |
| Set III | 0.8013 | 0.7002 | 0.8519 | 0.6833 |

**Table 3.9**: Kendall Tau-b correlation between SVS **ratings** and neurologists' **ratings**

## IV. Discussion

To the best of our knowledge this is the first study to report validation analyses on a severity score. SVS is a video-based objective measure of severity of dyskinesia. Previous studies on quantifying dyskinesia [37, 38, 39, and 40] have used classification based approaches to determine various metrics of dyskinesia using devices. We not only parameterize dyskinesia, but also combine these parameters mathematically to form the severity score, SVS. SVS can help neurologists evaluate and place dyskinetic patients on

a continuous scale and perform retrospective analyses on relative and absolute severity of dyskinesia in these patients.

A reliable and robust objective technique to quantify severity of dyskinesia has several potential advantages over clinical scales. In order to validate the utility of our score, we developed a ranking protocol which facilitates the comparison of our continuous score with the discrete ratings of neurologists. This chapter reports results on three studies that were performed to (1) validate the ranking protocol; (2) evaluate the robustness and utility of SVS; and (3) compare the utility of rankings over ratings. We have focused on quantifying the severity of dyskinesia while performing the communication task as it is easy to track using image registration without resorting to more complicated tracking algorithms needed for tasks such as buttoning a coat or drinking from a cup. The advantages of using our quantifying technique include the widespread availability of video recording equipment, its ease of use in clinical settings and the portability of the datasets for retrospective analyses and longitudinal studies

The advantage of using a continuous score over a discrete five point rating scale and hence the necessity to develop a ranking protocol was presented. The ranking protocol was designed based on the ratings assigned using the clinical definition of dyskinesia as opposed to a particular rating scale such as the CDRS or AIMS. Our results from Study I indicate that our ranking protocol was valid and produced high inter- and intra-neurologists' consistency while using it. Previous studies perform comparative analyses of their techniques with one of the standard rating scales. As mentioned earlier, all rating scales do not account for all attributes of dyskinesia. By using the general definition of dyskinesia as the basis of our ranking protocol, we are not biased towards any particular

attribute of dyskinesia. Our technique gives the participating neurologists the freedom to prioritize the different attributes including amplitude, velocity and directions of movement, degree of disability, anatomical distribution and type of dyskinesia. To further analyze the effect of this variable prioritization, three of the four neurologists were asked to record their priorities during their ranking process. The priorities seen in Table 3.10 indicate that our ranking protocol took most of the attributes of dyskinesia (as mentioned in Chapter 1) into account without the aid of a specific rating scale. The intensity of dyskinesia is a collective measure of amplitude, speed and general disability.

| Neurologist | Priority of Dyskinesia attributes |
|---|---|
| N1 | 1) Anatomical distribution; 2) Amplitide of movement; 3) Speed of movement |
| N2 | 1) Anatomical distribution; 2) Acceleration or sudden direction changes in movement; 3) Amplitude of movement; 4) general disability |
| N3 | 1) Anatomical distribution; 2)Speed of movement; 3) Amplitude; 4) bilateral dyskinesia |

**Table 3.10**: Priority of Dyskinesia attributes

We should state that the neurologists were asked not to determine the severity based on the most disabling type of dyskinesia – chorea, dystonia and tremor, but look more closely at the overall movement deficiencies. Hence the rankings of neurologist are inclusive of choreic and dystonic movement, but without explicit weightage given to either of them. Since SVS does not account for the most disabling type of dyskinesia or its severity, we believe that this protocol was most suited to our data.

Our results from Study II, based on amplitude and directions of movement, indicate that the SVS moderately correlates with the UDysRS and the neurologists' rankings. The lower correlation coefficients in Type II rankings can be attributed to two possibilities: (a) reduction in the sample size from 35 to 28 video sequences and (b) the number of dyskinetic patients removed from each set was based on the observation of only the most senior neurologist. Amplitude and directions of movement are necessary but not sufficient parameters to completely quantify dyskinesia. SVS was successfully tested on larger set of patient data with longer video sequences and simultaneous tracking of multiple body parts. The anatomical distribution of dyskinesia varies from person to person and we can see from Table 3.7 that neurologists consider this a significant parameter to evaluate the severity in patients. By performing simultaneous tracking and analysis of multiple body parts, we internally account for the anatomical distribution of dyskinesia in our severity score. This process emulates a neurologists' assessment more closely than previous studies that look at the movement data (obtained using devices) from multiple body parts individually. A significant point to note is that we had control data within our sample set in the form of seven non-dyskinetic patients. It can be seen that out of the seven patients, four of them had low severity scores, a trend similar to the 13 non-dyskinetic patients in chapter 2. The remaining three patients had the following issues – (a) one of them had a slightly higher rank since the patient exhibited voluntary motion using the limbs that were tracked and (b) the other two patients showed poor tracking quality with significant drifting of the landmark points. Table 3.11 shows the SVS and neurologists ranks of these 7 patients.

| PN | SVS rank | N1 | N2 | N3 | N4 | UDysRS |
|---|---|---|---|---|---|---|
| 10 | 12 | 4 | 4 | 4.5 | 2 | 6 |
| **14** | **2** | **4** | **4** | **4.5** | **8** | **7** |
| 18 | 9 | 4 | 4 | 4.5 | 7 | 3 |
| **21** | **5** | **4** | **4** | **4.5** | **2** | **1** |
| 25 | 10 | 4 | 4 | 4.5 | 5 | 2 |
| **26** | **3** | **4** | **4** | **4.5** | **2** | **4** |
| **33** | **7** | **4** | **4** | **4.5** | **4** | **5** |

**Table 3.11**. Ranks of seven non-dyskinetic patients. The highlighted rows indicate the four patients who are within the rank of 1 – 7 (corresponding to a tied rank of 4). Patient 18 exhibited voluntary motion that occluded a landmark point; and patients 25 and 18 showed drifting of landmark points owing to intensity homogeneity of their clothes.

SVS is inherently dependent upon tracking as the parameters of SVS are variance based and incorrect tracking can lead to inaccurate training shapes and inaccurate SVS scores. We have shown that SVS is robust to small variations to landmark points. Given the general guidelines for picking landmark points, the natural tendency is to pick these points on approximate joints especially for the shoulder, elbows, and knees. Variation can arise, when picking points on the forehead, reading material or feet. Our trials included such possible variations as seen in Figure 3.2. By looking at the tracking results of the patients highlighted in Table 3.6, we found that drifting of the landmark points had taken place due to: (a) point picked in areas such as broad horizontal stripes which can create small clusters of pixels with uniform intensity and (b), when trajectory of a given points crossed over the trajectory of other points. Thus, there are two main concerns in picking the landmark points – (a) maintain the sequence in which landmark points are picked; as different sequences can result in different training shapes for different patients, and (b) pick points on areas of intensity inhomogeneities ( this includes areas of contrast, texture differences or presence of buttons, pins etc.). Even though we pick points on the clothes of the patients as opposed to actual joints, which is the case for motion capture studies,

with good tracking, our technique is successful is capturing dyskinesia. Since we are looking at involuntary movement, cases of occlusion cannot be avoided and can be overcome by re-picking the landmark points on the affected frame(s).

We observed that tracking longer sequences using our image registration technique not only increases machine time, but also causes tracking quality to deteriorate; as drifts in landmark points are now propagated across more number of frames. This problem intensifies especially when tracking video sequences of severely dyskinetic patients or in patients wearing clothing without intensity variations. Ten seconds of video was found to be sufficient to capture the severity of dyskinesia without losing tracking quality. We do wish to specify that we did not evaluate the neurologists' rankings for longer video sequences and hence cannot attest to the fact that there will not be a change in the rankings when longer video sequences are rated and ranked. However, from our studies, we believe that these changes will be minimal and not affect the overall correlation between our score and the neurologists' evaluations.

Study III was performed to observe the effect of using SVS as a continuous variable as opposed to discretizing it to compare it directly with conventional discrete rating scales. A higher correlation between the SVS ratings and neurologists' ratings when compared to the correlation coefficients observed in Study II using Type I and Type II rankings indicate that SVS has a high global correlation with the neurologists as we were comparing groups of patients within rating categories irrespective of their ordering within the group. But in Study II, the ranking or the ordering of the patients within each group was compared thus requiring a more stringent quantifying technique which is capable of capturing the subtle differences between patients within that group. Hence we believe

that using a ranking technique can initiate the development of better quantifying measures. Two patients can be assigned the same category of severity using a rating scale, but not necessarily the same rank. This variation can occur as the neurologist is not restricted to the standard rating scale definition and may be inclined to assess the motor disability of one patient as more severe than the other, but not significant enough to assign the next higher rating on the scale. We do not reject the possibility of two patients receiving the same severity score and hence the same rank, but such an occurrence might be uncommon as SVS is a continuous variable.

An added advantage of using an objective score based ranking technique such as ours is its utility in handling large patient databases which get updated frequently with new patient records. If a new patient is added to the existing sample set, the current rankings and the SVS for the new patient can be used to determine the relative severity of the new patient with respect to the other patients with the SVS specifying the absolute dyskinesia severity, thus avoiding the manual re-ranking of the entire dataset by the neurologists. Thus the SVS and the corresponding rank can guide the neurologist in clinically rating the patient's dyskinesia severity based on their relative ranking to other patients. Further work is being done to improve and to automate the video tracking on longer video sequences and to include contributions from dystonia and chorea into the SVS. We believe that these factors will increase the correlation between the SVS and the neurologists' rankings. Such a severity score can be applied to other assessments tasks similar to the activities of daily living and used to quantify longer video sequences. Dyskinesia exhibits diurnal variations and our study is based on 10s video sequences of a specific task. Hence even though our semi-automatic technique is comparable to

conventional rating scales, which take such diurnal variations into account, it may not necessarily replace them.

# CHAPTER IV

# DEVELOPMENT OF SVSCD – CHARACERTIZING CHOREA AND DYSTONIA

## I. Introduction

One of the key attributes of dyskinesia is assessing the most disabling movements of the patients. Dyskinesia is the general terminology used for abnormal involuntary movements. Levodopa induced dyskinesia can appear as dystonia, chorea or tremor. The presence of any one or a combination of these types of movements indicates the side-effects of levodopa in a Parkinson's disease patient. In this chapter, we will discuss the characteristics of dystonia, chorea and tremor; develop parameters to quantify these movements, modify our severity score SVS developed in chapter 2 to include these parameters, and finally validate the new score.

Dystonia is a general term used for twisting movements which occur with sustained muscle contractions of opposing muscles [55]. These movements are repetitive in nature and severe cases of dystonia can exhibit sustained muscle contractions leading to abnormal body postures and hence, disability. Figure 5.1 shows an example of a Parkinson's disease patient with severe cervical dystonia. If dystonia occurs in a single body part, it is termed at focal dystonia. Typically focal dystonia affects the face, neck, voice and hand. When dystonia occurs in more than one body part, it is called generalized dystonia, which typically manifests itself in the trunk and legs. Initially, dystonia occurs with the presence of voluntary movements, but with progression of its severity, it can also occur in body parts that are at rest and not involved in the voluntary movements. This feature of dystonia further supports our technique in looking at multiple body parts of the

patients simultaneously to capture this inherent correlation in movement as opposed to individual analysis of body parts as prior research indicates. The key kinematic principle of dystonia we wish to capture is the presence of periodicity in the movements and reduced speed of movement due to sustained muscle contractions.

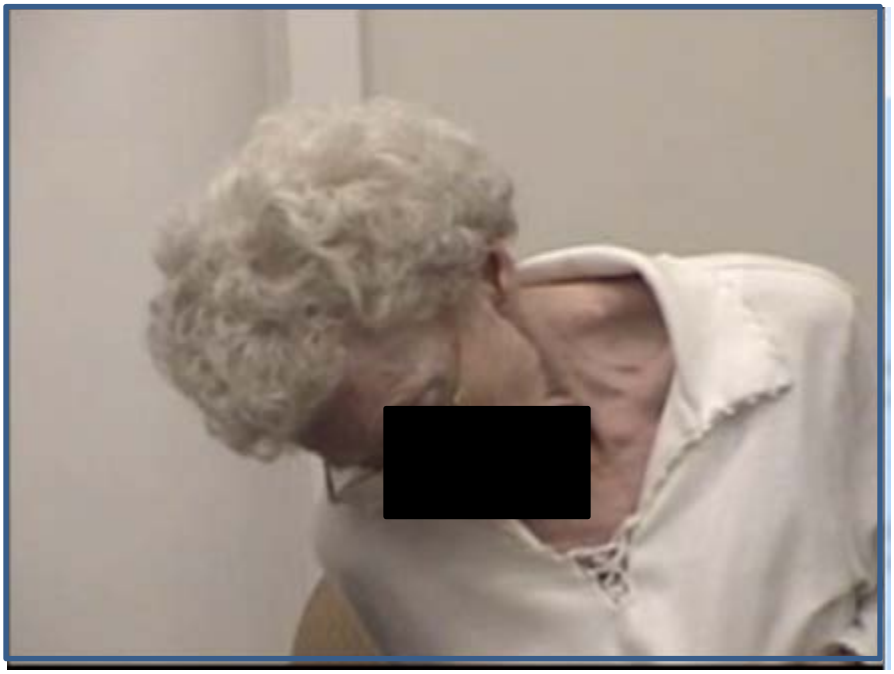

**Fig 4.1**. Parkinson's disease patient with severe cervical dystonia. This patient performed the entire communication task with this neck posture

Chorea, on the other hand, is exhibited as irregular, non-rhythmic and rapid involuntary movements which typically involve more than one body part. Severe chorea can cause violent movements in patients and interfere with almost all the activities of daily living. Figure 4.2 shows consecutive frames of a Parkinson's disease patient affected with severe chorea. As the severity of chorea increases, movements tend to have larger amplitudes and higher speed in more directions. Patients tend to hide their movements by forcefully holding back their limbs by gripping a chair tightly or by keeping knees and feet crossed

over one another, though they can do so only for short periods of time. It is hard to incorporate them into voluntary tasks as they are unpredictable and irregular. Higher speed of movement and the lack of periodicity are the parameters of interest in quantifying chorea since we have already incorporated amplitude and directions of movement in our score.

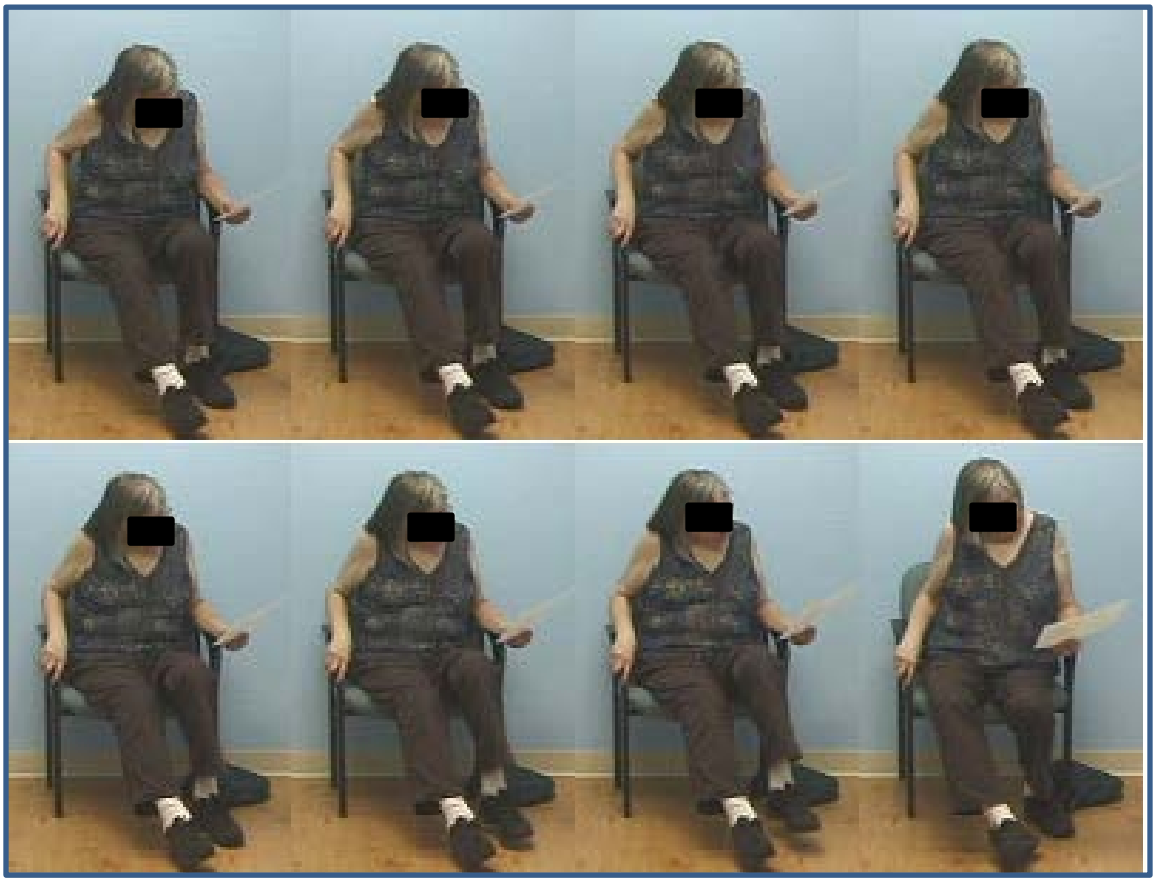

**Fig. 4.2**. Eight consecutive frames of a patient with severe chorea performing the communication task. Note the gripping of the chair with the right arm to provide stability. This patient was severely dyskinetic and she was falling off the chair.

Tremor is characterized by rhythmic oscillatory movements. It is usually produced by the alternate contraction of opposing muscles [55]. Resting tremor is a common symptom of Parkinson's disease, but can occur even in the ON medication state. It is classified as a

dyskinetic movement due to its abnormal and involuntary nature. Typically tremors are the easiest type of dyskinesia to quantify. We can quantify the severity of tremor by analyzing its amplitude and frequency components. Low amplitude - high frequency tremors are difficult to detect by video tracking unless high speed cameras and accurate tracking algorithms are used. The RUMC dataset contains two patients with this type of tremor and our technique could not track it successfully, especially if they occur in extremities such as fingers.

Even though, we have identified the features of interest in quantifying dystonia, chorea and tremor, it is important to note that these movements may not necessarily occur separately in patients. Most often, dyskinesia is exhibited as a combination of two or more of these movements. Hence quantifying each type individually is neither possible nor rational. While securing the ratings and rankings of the RUMC dataset by the four neurologists, we had not asked them to explicitly rate or rank dystonia, chorea and tremor, but use the general definition of dyskinesia which would effectively include the extent of disability caused by these movements. Hence our goal was to further modify our severity score to incorporate these parameters into the SVS and validate the new score using the techniques seen in chapter 3. Several studies have used entropy, frequency spectrum and time series analyses on various forms of motion data based on position, velocity and acceleration [53, 36, 42, 47]. These studies are performed on data collected on individual body parts. Daffertshofer et al. have used PCA in studying movement coordination in gait using similar techniques as ours, but their study was focused more on using PCA for synthesis than analysis [45]. We will use the PCA parameters obtained in chapter 2 in combination with frequency spectrum analysis to

develop features that differentiate dystonia, chorea and tremor movements embedded in the patient's dyskinetic movements. We also proceed to show which of these parameters are applicable to our dataset. A new score is developed using these parameters and a validation study is conducted by comparing the new SVS rankings to the neurologists' and UDysRS rankings. Our results indicate that the new SVS score captures the severity of dyskinesia more accurately and shows a higher correlation coefficient with the neurologists' and UDysRS rankings.

## II. Methods

The RUMC dataset used in chapter 3 provides us with longer video sequences of better image quality than the VUMC dataset. Hence we have used only this dataset to conduct our analyses. We used 35 ten seconds video sequences of patients with varying degrees of dyskinesia in our study. The data was tracked using ABA, an intensity-based image registration algorithm as in chapter 3. Since we have good tracking results for the communication task, we have used the same for analyses. We performed one study for testing our parameters and one validation study using only Type I ranks of the neurologists. The disadvantages of a smaller sample set due to the elimination of certain patient ranks as seen in chapter 3 prompted us to use only Type I rankings. The same eleven landmark points – forehead, shoulders, chest, elbows, knees, feet and reading material were tracked and their trajectories computed and stored as [x,y] position coordinates.

**II.1 Dystonia, Chorea and tremor parameters and development of SVSCD**

As in chapter 2, we have maintained our preference for using a combined trajectory approach as opposed to looking at the trajectories of the individual landmark points. Using the equations specified in chapter 2, PCA was performed on all the trajectories together and corresponding eigenvalues and eigenvectors computed. Daffertshofer et al. study the evolution of each mode of variation by projecting the set of training shapes, T, onto the eigenvector vector space spanned by $V_K$ as derived in equations (2.1) and (2.5)

$$E_k = V_K \, . T \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.1)$$

In [45], $E_k$ are defined as eigenmodes. For our analysis we use only a subset of $V_K$ in (4.1). This subset consists of eigenvectors corresponding to the most significant modes of variation. These eigenmodes are in decreasing order of the contribution to the total movement. For example, if a patient exhibits more chorea than dystonia, choreic movements are captured in the first few eigenmodes and vice versa. Instead of using the individual trajectories in T, we use the eigenmodes to analyze variations among patients. Frequency spectrum analysis was performed to determine the best suited parameter that captures the differences between dystonia, chorea and tremor.

The Fast Fourier Transform (FFT) of the eigenmodes was computed and the power spectra were analyzed. It has been reported in the literature that dyskinesia is concentrated in the 1 – 3 Hz band of frequencies [42, 53, 36, 37]. But this was observed in acceleration data and does not specify frequency bands corresponding to dystonia and chorea. Tremor is usually in the higher frequency ranges of 6 – 8Hz [28]. The power spectrum of each of the eigenmodes was analyzed in different frequency bands – B1: 0.5 – 1.5 Hz; B2: 1.5 – 3.5 Hz; B3: 3.5 – 10 Hz. The total power in the frequency range of

0.5 – 10 Hz was computed as TP. The distribution of power in these bands of interest was computed as

$$P_i = \frac{\text{Total power in Band } B_i}{TP} \text{ x } 100$$

where i = 1:3, denoting the three bands of interest.

Thus each mode had a set consisting of three ratios, P = {$P_1$, $P_2$ and $P_3$}. The power spectrum of all 35 patients were observed and analyzed. These observations and analyses will be discussed in the results section along with the basis for the new score called the SVSCD – SVS with chorea and dystonia parameters.

**II.2 Validation Study**

Since the new score was also a continuous variable, we followed a procedure similar to the validation study in chapter 3. The rankings for the 35 patients using SVSCD were computed. These rankings were compared to neurologists' and UDysRS rankings using the Kendall's Tau-b correlation coefficient. Only Type I rankings were used for the validation study. Instead of eliminating the non-dyskinetic patients as in the Type II rankings, a different strategy was used to create uniformity in comparing the tied neurologists' rankings to the tie-less SVSCD rankings. If the number of patients correctly identified as non-dyskinetic ( i.e. with low SVSCD scores) were comparable to the seven patients identified by senior neurologist N1, they were assigned tied ranks based on the number identified correctly. We argue that this strategy is rational as the ranking of non-dyskinetic patients is irrelevant and by assigning tied ranks, we can achieve a more meaningful correlation. We could not use this strategy in chapter 3 as only 50% of the

patients were correctly identified as non-dyskinetic, though the other truly non-dyskinetic patients had very low SVS scores.

## III. Results

We present our results in two sections – (a) Development of SVSCD, and (b) Validation of SVSCD.

### III.1 Development of SVSCD

There were four main observations noted from the power spectrum.

1) Non-dyskinetic patients normally exhibited fewer significant modes and a uniform distribution of power across B1, B2, and B3. In cases of tracking issues such as drifting, the power of the frequencies in the higher modes was significantly lower, thereby contributing negligibly to TP. Figure 4.3 shows the power spectrum of the six eigenmodes of patient 10, who is non-dyskinetic but had drifting issues during tracking.
2) Dyskinetic patients with predominantly choreic movements showed higher concentrations of frequencies in the B2 (1.5 – 3.5 Hz) band. We did not have a patient with only chorea. Hence, patient 32's movements were a good example of dyskinesia that did not show a lot of repetitive patterns. Note that in mode 1, the predominant frequency in band B2.
3) Dystonic patients, as seen in the example in Figure 4.5, exhibited frequencies concentrated in band B1 (0.5 – 1.5 Hz).
4) None of the patients exhibited significant frequency concentrations in band B3 (3.5 – 10 Hz) which includes tremor frequency band (6 – 8 Hz). Since our

tracking technique could not pick the tremor exhibited by two patients, our observation was valid. The power spectrum of these two patients showed only dystonic fetures. The power spectrum of the first two and the last eigenmode of one of the patients is shown in Figure 4.6 to illustrate our observation.

Most of the patients exhibited both chorea and dystonia in varying degrees. In order to observe quantifiable differences, the trend of the ratios P1, P2, and P3, was analyzed. The spread of powers P1, P2, and P3 for each mode differentiated patients as more choreic

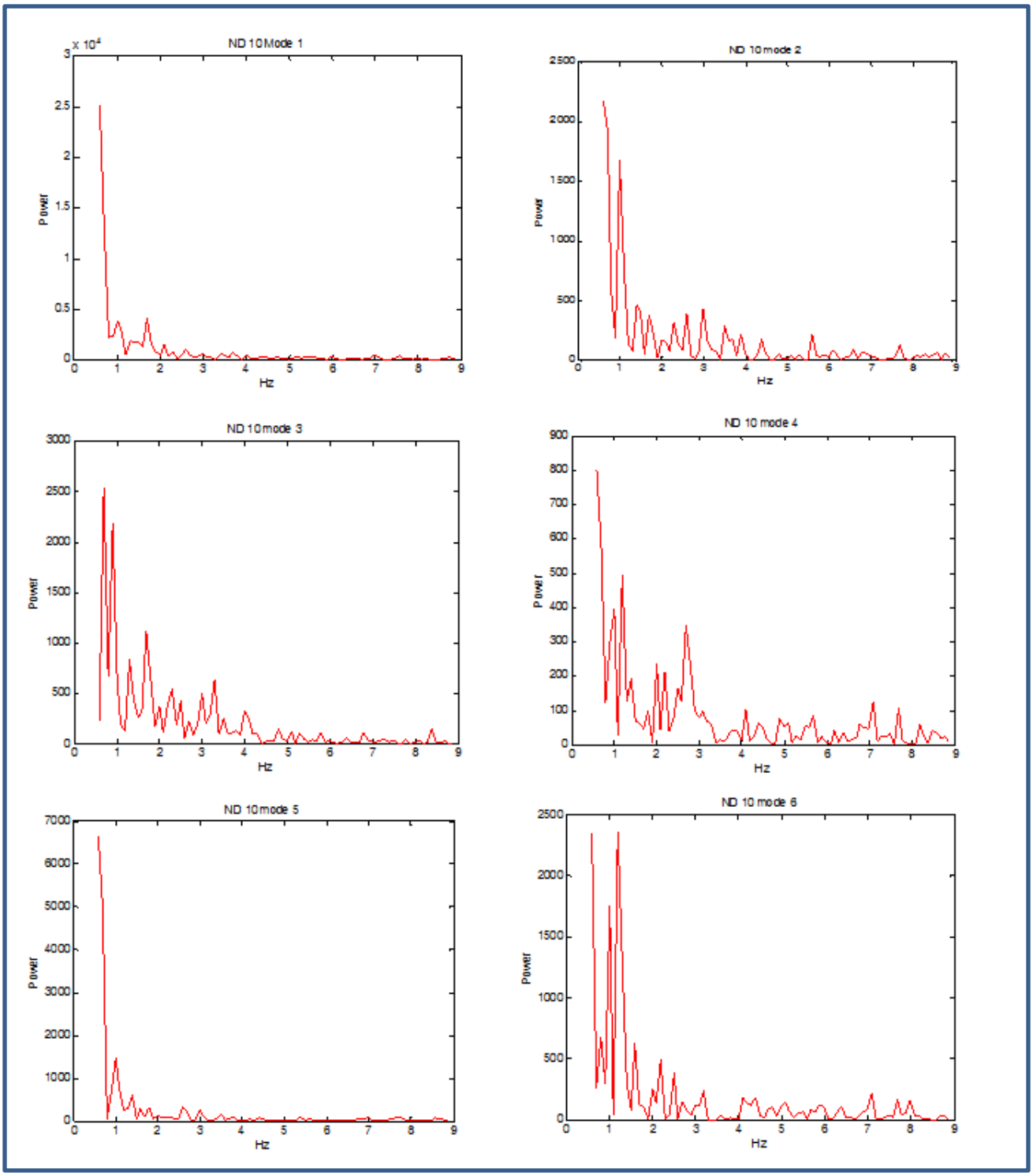


**Fig 4.3**. Power spectra of six eigenmodes of a non-dyskinetic patient with poor tracking. Note the low power of the frequencies in the higher modes and presence of frequency components > 3 Hz.

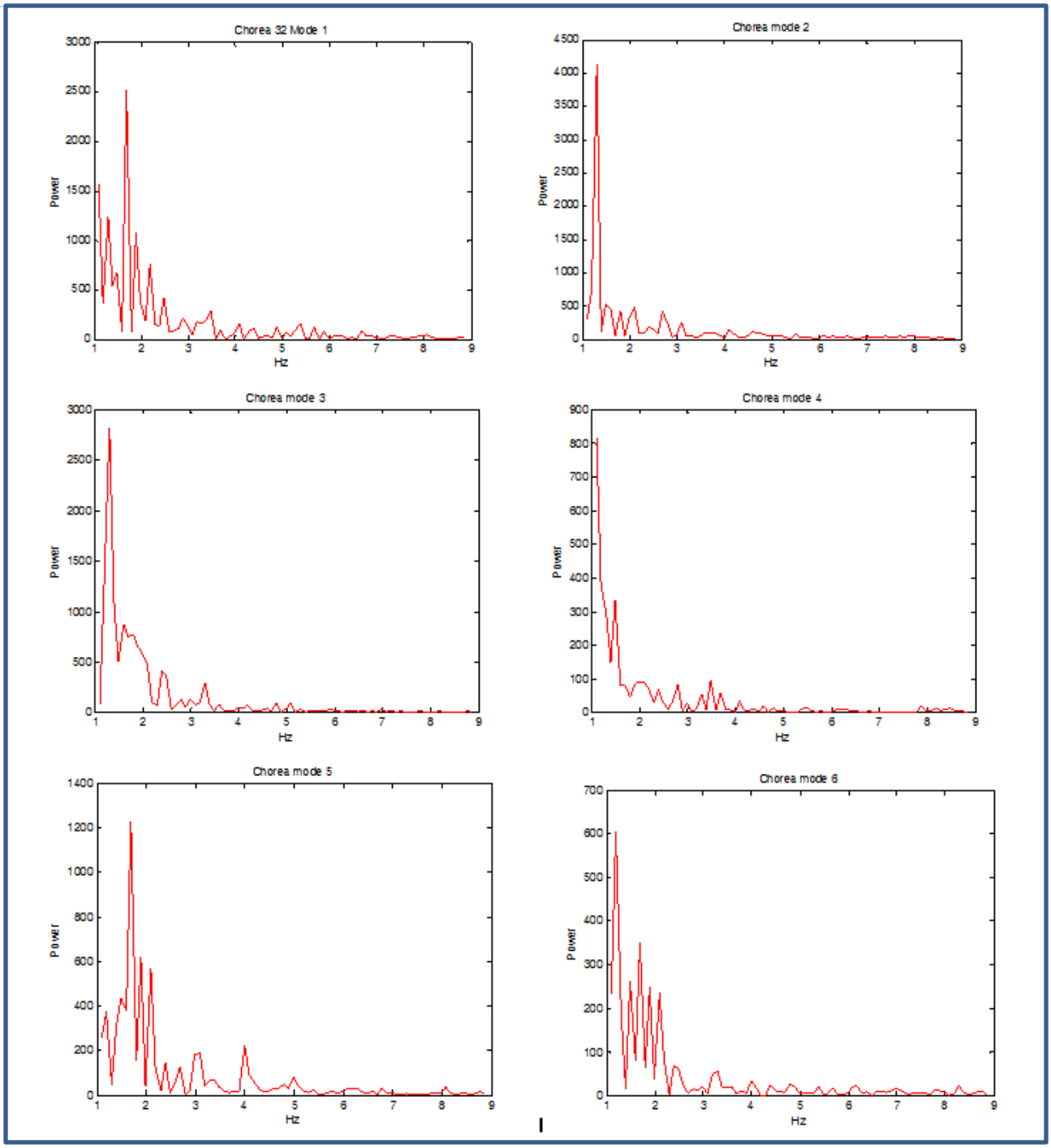


**Fig 4.4**. Power spectra of six eigenmodes of a patient who exhibited predominantly choreic movements. This patient did not show large amplitude of movements, but exhibited random uncorrelated movements of the head and feet explaining the presence of multiple frequency bands in the spectrum.

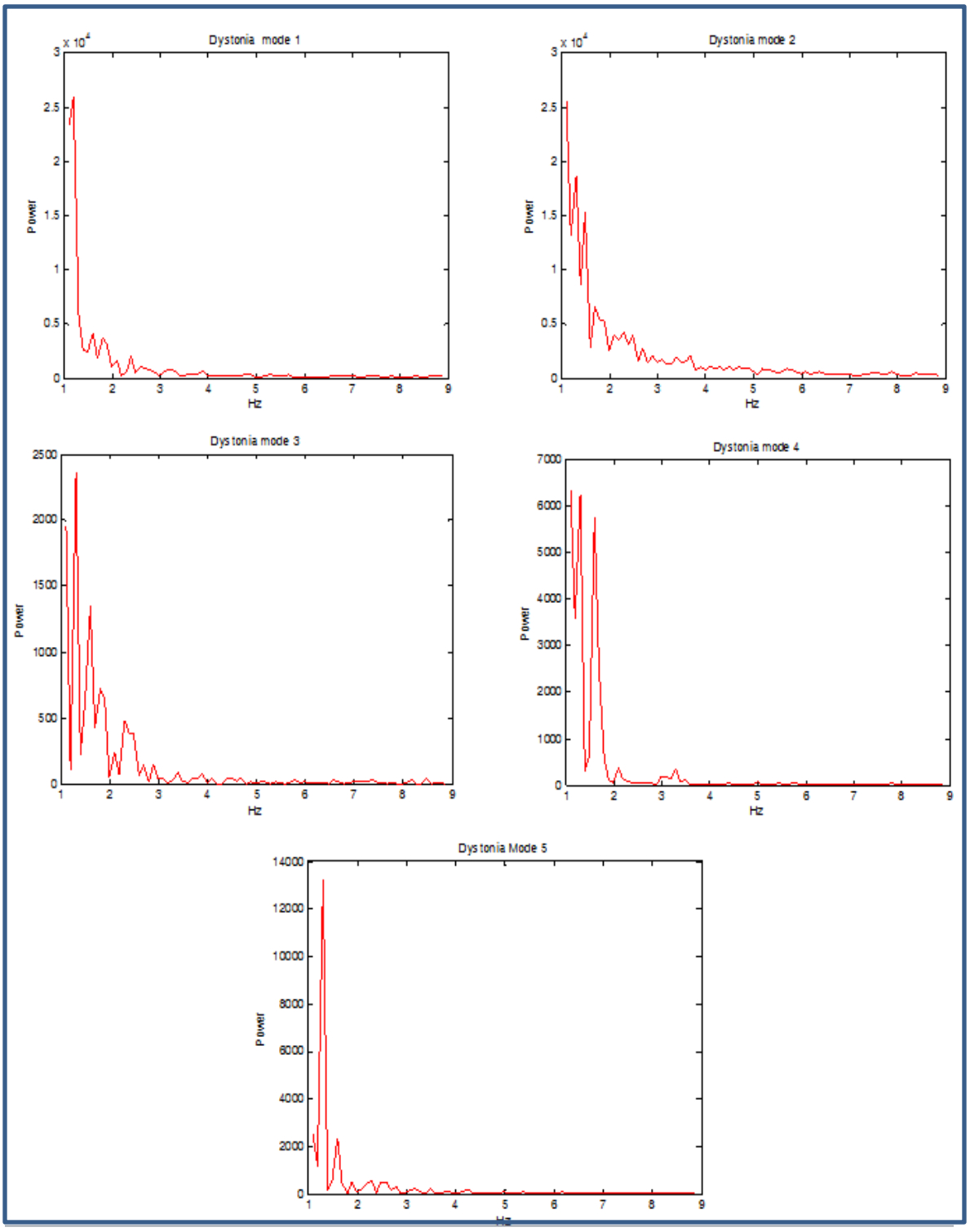


**Fig 4.5** Power spectra of five eigenmodes of a predominantly dystonic patient. Note the concentration of frequencies in band B1 in all of the modes.

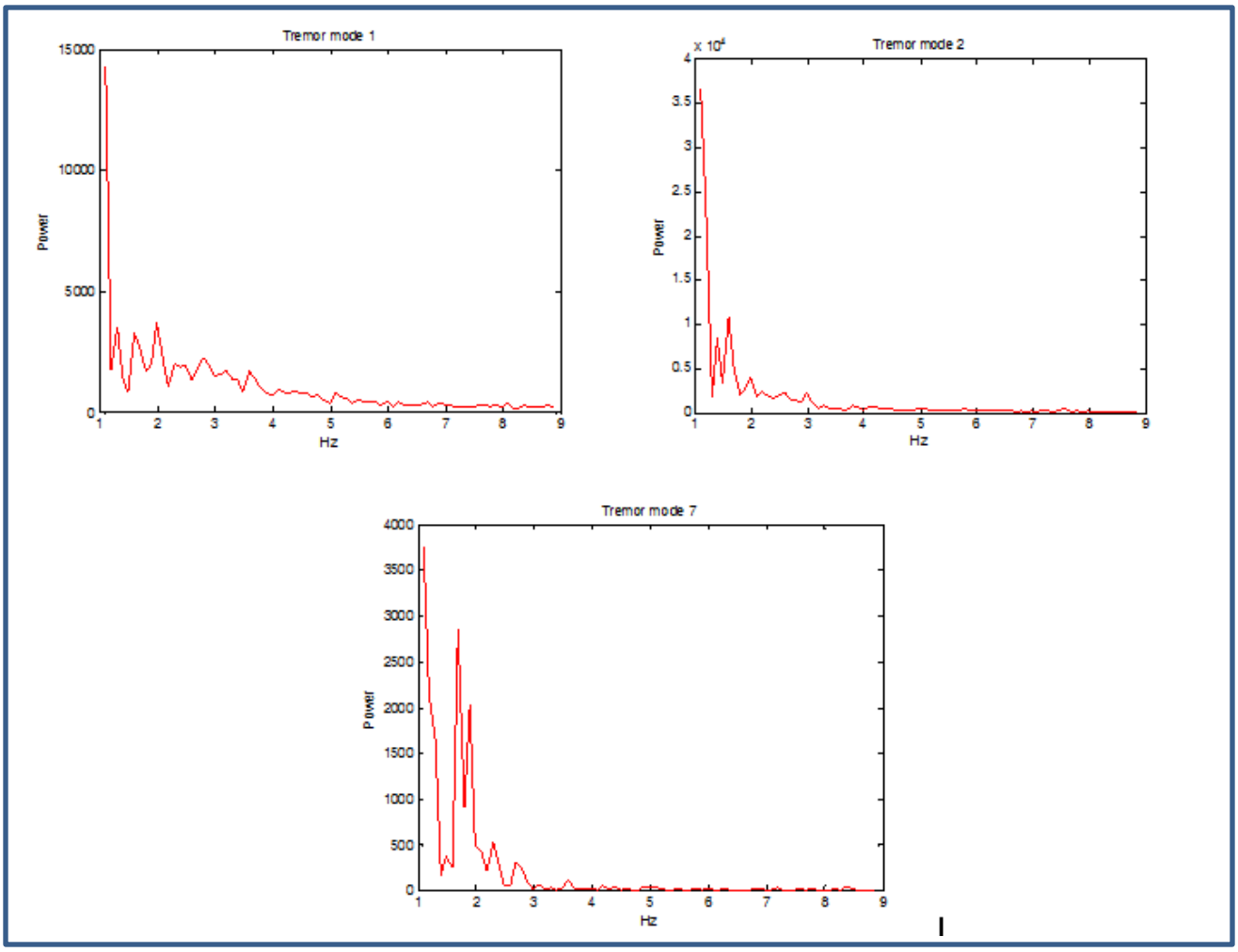


**Fig 4.6** Power spectra of modes 1, 2 and 7 shown for patient with tremor of hand.

than dystonic and vice versa. Table 4.1 shows the {P1, P2, P3} values for each of the modes for selected number of patients. To compute the spread of {P1, P2, P3}, their standard deviation was calculated and the average of the standard deviations for all the eigenmodes was computed and denoted as STDP. It was observed that the non-dyskinetic patients had the lowest STDP, followed by the choreic patients and the dystonic patients had the highest STDP. Table 4.2 shows the STDP values, for the patients shown in Table 4.1, illustrating this trend.

Non-Dyskinetic

| | EM1 | EM2 | EM3 | EM4 | EM5 | EM6 |
|---|---|---|---|---|---|---|
| P1 | 67.4081 | 58.6509 | 44.195 | 42.6595 | 79.8492 | 58.4399 |
| P2 | 19.1074 | 22.2533 | 36.9925 | 29.061 | 12.1743 | 18.4669 |
| P3 | 13.4846 | 19.0957 | 18.8125 | 28.2796 | 7.9765 | 23.0932 |

| | | | | | |
|---|---|---|---|---|---|
| P1 | 65.8373 | 81.5933 | 62.98 | 72.433 | 49.9371 |
| P2 | 22.5783 | 13.212 | 22.8396 | 16.9554 | 30.7113 |
| P3 | 11.5844 | 5.1947 | 14.1803 | 10.6116 | 19.3516 |

Predominantly Choreic

| | EM1 | EM2 | EM3 | EM4 | EM5 | EM6 |
|---|---|---|---|---|---|---|
| P1 | 72.1776 | 88.7431 | 77.1823 | 75.5495 | 71.5274 | 75.7619 |
| P2 | 20.9637 | 6.5965 | 19.0802 | 17.639 | 21.9106 | 17.5126 |
| P3 | 6.8588 | 4.6604 | 3.7375 | 6.8115 | 6.5621 | 6.7255 |

| | | | | | |
|---|---|---|---|---|---|
| P1 | 79.8278 | 77.863 | 86.1413 | 88.8784 | 81.7382 |
| P2 | 17.1705 | 14.2865 | 8.0716 | 7.9593 | 15.1759 |
| P3 | 3.0016 | 7.8505 | 5.7871 | 3.1623 | 3.086 |

Predominantly Dystonic

| | EM1 | EM2 | EM3 | EM4 | EM5 |
|---|---|---|---|---|---|
| P1 | 99.1889 | 86.6737 | 94.7352 | 94.9789 | 90.3039 |
| P2 | 0.5861 | 8.3617 | 4.4185 | 4.6408 | 7.6714 |
| P3 | 0.225 | 4.9645 | 0.8463 | 0.3803 | 2.0247 |

Combination of chorea and dystonia

| | EM1 | EM2 | EM3 | EM4 | EM5 | EM6 | EM7 |
|---|---|---|---|---|---|---|---|
| P1 | 72.3317 | 96.3721 | 98.5652 | 94.6404 | 97.0009 | 74.0713 | 90.5435 |
| P2 | 17.7174 | 2.7678 | 1.1834 | 3.8404 | 2.2132 | 23.4207 | 8.6212 |
| P3 | 9.9509 | 0.86 | 0.2514 | 1.5193 | 0.7859 | 2.508 | 0.8352 |

**Table 4.1** {P1, P2, P3} values for selected patients with specified types of dyskinesia, where EM denote eigenmodes.

Thus, we determined that STDP was low for non-dyskinetic patients and higher for dyskinetic patients, irrespective of the type of dyskinesia. Thus,

$$\text{Dyskinesia severity } \alpha \text{ STDP}$$

Hence, we modified our severity score SVS, which was

$$\text{SVS} = \frac{\text{TV x NSM}}{\text{STDEV}}$$

to include STDP. All the previous parameters were retained as they described only the amplitude and directions of movements. Based on the relationship between dyskinesia and STDP, we propose the new severity score, SVSCD, denoting SVS with chorea and dystonia, as:

$$\text{SVSCD} = \frac{\text{TV x NSMx STDP}}{\text{STDEV}}$$

Where TV: Total variance; NSM: number of significant modes of variation; STDP: average standard deviation of {P1, P2, P3} and STEV: standard deviation of the contribution of the eigenvalues to the total variance.

| Type | Case | STDP |
|---|---|---|
| Non-dyskinetic | 1 | 22.4995 |
| | 2 | 29.2349 |
| Choreic | 1 | 38.1936 |
| | 2 | 43.1691 |
| Dystonic | 1 | 51.8623 |
| Combination | 1 | 48.548 |

**Table 4.2** STDP values for the patients in Table 5.1

### III.2 Validation of SVSCD

Table 4.4 shows the STDP, SVSCD and the corresponding rankings; neurologists' and UDysRS rankings of the 35 RUMC patients. Kendall Tau-b correlation coefficients between the SVSCD rankings and the neurologists' and UDysRS rankings show considerable improvement from those observed using the SVS as seen in Table 4.3. All tau values were statistically significant with p <=0.0001.

| | SVSCD | N1 | N2 | N3 | N4 | UDysRS |
|---|---|---|---|---|---|---|
| **SVSCD** | 1 | 0.7734 | 0.6913 | 0.7007 | 0.647 | 0.637 |
| **N1** | 0.7734 | 1 | 0.8746 | 0.8642 | 0.7994 | 0.7563 |
| **N2** | 0.6913 | 0.8746 | 1 | 0.8852 | 0.8337 | 0.7632 |
| **N3** | 0.7007 | 0.8642 | 0.8852 | 1 | 0.8371 | 0.7386 |
| **N4** | 0.647 | 0.7994 | 0.8337 | 0.8371 | 1 | 0.7313 |
| **UDysRS** | 0.637 | 0.7563 | 0.7632 | 0.7386 | 0.7313 | 1 |

| | SVS | N1 | N2 | N3 | N4 | UDysRS |
|---|---|---|---|---|---|---|
| **SVS** | 1 | 0.7392 | 0.6434 | 0.6594 | 0.5965 | 0.6134 |
| **N1** | 0.7392 | 1 | 0.8746 | 0.8642 | 0.7994 | 0.7563 |
| **N2** | 0.6434 | 0.8746 | 1 | 0.8852 | 0.8337 | 0.7632 |
| **N3** | 0.6594 | 0.8642 | 0.8852 | 1 | 0.8371 | 0.7386 |
| **N4** | 0.5965 | 0.7994 | 0.8337 | 0.8371 | 1 | 0.7313 |
| **UDysRS** | 0.6134 | 0.7563 | 0.7632 | 0.7386 | 0.7313 | 1 |

**Table 4.3** Kendall Tau-b correlation matrix for SVS and SVSCD. The highlighted rows show the improvement with using SVSCD

| PN | STDP | SVSCD | SVSCD Rank | N1 | N2 | N3 | N4 | UDysRS |
|---|---|---|---|---|---|---|---|---|
| 1 | 45.96619 | 7.0948 | 25 | 23 | 24 | 22 | 24 | 24 |
| 2 | 30.7019 | 0.0834 | 2 | 8 | 11 | 10 | 10 | 15 |
| 3 | 43.16914 | 15.0505 | 32 | 27 | 26 | 27 | 23 | 30 |
| 4 | 32.49801 | 0.2394 | 6 | 10 | 12 | 13 | 12 | 12 |
| 5 | 44.84424 | 14.0782 | 31 | 32 | 28 | 29 | 31 | 31 |
| 6 | 34.22355 | 0.3687 | 8 | 9 | 9 | 12 | 11 | 10 |
| 7 | 40.29787 | 0.9953 | 13 | 20 | 21 | 15 | 20 | 18 |
| 8 | 38.71286 | 1.7973 | 17 | 12 | 13 | 16 | 17 | 27 |
| 9 | 46.96165 | 1.2642 | 14 | 19 | 17 | 18 | 14 | 16 |
| **10** | **22.49949** | **0.4913** | **9** | **4** | **4** | **4.5** | **2** | **6** |
| 11 | 51.65184 | 12.9262 | 30 | 33 | 31 | 32 | 32 | 33 |
| 12 | 37.23642 | 4.1699 | 20 | 18 | 20 | 14 | 18 | 22 |
| 13 | 32.44837 | 2.1883 | 19 | 22 | 29 | 23 | 26 | 21 |
| **14** | **24.5219** | **0.039** | **1** | **4** | **4** | **4.5** | **8** | **7** |
| 15 | 48.54796 | 205.1266 | 35 | 35 | 35 | 35 | 35 | 35 |
| 16 | 32.51816 | 1.3324 | 15 | 16 | 15 | 17 | 15 | 25 |
| 17 | 41.37078 | 0.5253 | 10 | 11 | 10 | 11 | 13 | 14 |
| **18** | **31.5318** | **0.1753** | **4** | **4** | **4** | **4.5** | **7** | **3** |
| 19 | 42.05907 | 8.8748 | 26 | 24 | 22 | 24 | 25 | 23 |
| 20 | 33.43582 | 1.8294 | 18 | 13 | 14 | 9 | 9 | 11 |
| **21** | **26.13197** | **0.2256** | **5** | **4** | **4** | **4.5** | **2** | **1** |
| 22 | 47.25518 | 27.3533 | 33 | 34 | 34 | 34 | 34 | 34 |
| 23 | 40.25066 | 6.0723 | 23 | 28 | 23 | 25 | 21 | 20 |
| 24 | 45.4451 | 6.0421 | 22 | 14 | 8 | 4.5 | 6 | 8 |
| **25** | **29.23488** | **0.5708** | **12** | **4** | **4** | **4.5** | **5** | **2** |
| **26** | **23.71956** | **0.1576** | **3** | **4** | **4** | **4.5** | **2** | **4** |
| 27 | 42.61222 | 10.4013 | 29 | 30 | 27 | 28 | 29 | 32 |
| 28 | 46.66236 | 5.8248 | 21 | 26 | 32 | 31 | 33 | 29 |
| 29 | 45.98504 | 6.7369 | 24 | 21 | 19 | 21 | 19 | 17 |
| 30 | 45.00643 | 32.7318 | 34 | 29 | 30 | 30 | 22 | 26 |
| 31 | 45.42109 | 8.9403 | 27 | 25 | 25 | 26 | 28 | 19 |
| 32 | 38.19355 | 0.5363 | 11 | 15 | 18 | 19 | 16 | 13 |
| **33** | **25.33456** | **0.3251** | **7** | **4** | **4** | **4.5** | **4** | **5** |
| 34 | 37.15962 | 1.3455 | 16 | 17 | 16 | 20 | 27 | 9 |
| 35 | 51.86232 | 10.2268 | 28 | 31 | 33 | 33 | 30 | 28 |

**Table 4.4**. STDP, SVSCD, SVSCD rankings and neurologists' and UDysRS rankings of 35 RUMC patients. Highlighted rows indicate non-dyskinetic patients as rated by senior neurologist

It must be noted that the SVS Kendall Tau-b coefficients reported in chapter 2 are different from the values reported in Table 4.3 This change is attributed to the fact that the two patients who exhibited voluntary movements were analyzed again after removing the landmark point causing occlusion or involved in the movement. Hence, the SVS correlated better, but still failed to recognize all the non-dyskinetic patients correctly as opposed to SVSCD, which could recognize five out of seven patients correctly. However, one of the patients, PN 2 in table 4.4, classified as mildly dyskinetic by the neurologists (rank 1 for the dyskinetic category, not including tied ranks) was misinterpreted as non-dyskinetic by the SVSCD. We do not assign a tied rank to this patient as that would bias the correlation coefficient towards the SVSCD. Instead, we remove him from the sample set. Table 4.5 shows the Kendall Tau-b correlation matrix showing a slight improvement in correlation coefficients compared to those observed in Table 4.3

| | **SVSCD** | **N1** | **N2** | **N3** | **N4** | **UDysRS** |
|---|---|---|---|---|---|---|
| **SVSCD** | 1 | 0.792 | 0.715 | 0.7215 | 0.6691 | 0.6781 |
| **N1** | 0.792 | 1 | 0.8778 | 0.863 | 0.7943 | 0.7667 |
| **N2** | 0.715 | 0.8778 | 1 | 0.8891 | 0.8344 | 0.7631 |
| **N3** | 0.7215 | 0.863 | 0.8891 | 1 | 0.827 | 0.7406 |
| **N4** | 0.6691 | 0.7943 | 0.8344 | 0.827 | 1 | 0.7328 |
| **UDysRS** | 0.6781 | 0.7667 | 0.7631 | 0.7406 | 0.7328 | 1 |

**Table 4.5** Kendall Tau-b correlation matrix patient PN 2 removed from the sample set. Thus our validation study indicates a moderate improvement in the capability of the severity score to quantify dyskinesia.

## IV. Discussion

We have proposed a new objective measure to quantify dyskinesia incorporating information about the type of dyskinesia – dystonia, chorea and tremor, based on frequency spectrum analysis. Consistent with the results reported in previous studies, we have also observed dyskinetic movements manifested in lower frequencies in the range of 0.5 – 3.5 Hz. All these studies look at the frequency spectrum of the kinematic signals obtained individually from different parts of the body as opposed to our study, which captures similar information from the composite signals obtained in the form of eigenmodes after PCA analysis on the set of training shapes [53, 42]. Thus, we eliminate the need to cross correlate the results obtained at one body part with those obtained at a different part as in [38].

We also conducted a study using approximate entropy as a parameter to quantify the differences between dystonia and chorea [56]. Entropy of a signal is defined as the amount of uncertainty in the given signal. In our application, this metric translates to estimating the randomness in the set of trajectories. Patients who are predominantly choreic would be expected to have highest entropy, followed by patients with dystonia who exhibit repetitive patterns, followed by patients with tremor having the lowest entropy. The two parameters required to compute approximate entropy of the eigenmodes are the approximate length of the pattern to be identified, and the similarity measure. Since most patients have a combination of dystonia, chorea and tremor, we found that approximate entropy did not provide consistent results in distinguishing dystonic and choreic patients. If our dataset had several cases of severely dystonic or choreic patients, this measure could have been used, but that would imply using a parameter that is highly

suited to a particular type of dataset. We believe that the STDP parameter is a general parameter that can be used in all datasets. Similar studies with a different dataset can be done to validate this claim. We did not validate SVSCD on the VUMC dataset due to lack of data; since we had determined in chapter 3 that at least 10s of video data showing the entire patients in the image is essential for capturing the actual severity of dyskinesia.

We do not explicitly quantify dystonia, chorea and tremor. We have developed SVSCD in such a way that it includes this information inherently. By this process, our final result is a single number that represents the severity of dyskinesia. STDP can be used as a metric independently to classify choreic and dystonic measurements. The 35 patients in the RUMC dataset can be re-ranked based on chorea and dystonia and their rankings can be correlated to the rankings obtained using STDP. We have not performed this validation; as it tedious and several patients have a combination of dystonia and chorea which might be difficult to distinguish visually. Due to lack of good tracking data, we have not shown any example of tremor being quantified using our score. Since our score is based on parameters measuring frequency dispersion, patients with tremor, that have been tracked successfully, should exhibit peaks in the range of 6 – 8 Hz.

Apart from repetitive movements, dystonia is also characterized by presence of sustained contractions and hence, twisted postures as seen in Figure 4.1. Simple parameters such as angles between the neck and the shoulders, knees and ankle can be analyzed to quantify such twisted postures. This analysis was not conducive to our data as it was difficult to judge in certain patients if they had voluntarily chosen to cross their feet or bend their knees outwards or if it was a dystonic posture. Hence, it becomes vital to design a study that includes conditions which clearly specify patients not to sit with crossed legs or read

with animated gestures of the head and hands. These voluntary movements, unless characterized using pose analysis methods, can be misinterpreted as dystonic movements by objective techniques such as ours. Trained neurologists can perceive these differences between a voluntary posture and a dystonic posture and this benefit exists for all cases of dyskinesia assessment. Hence, there is always a certain degree of compromise needed when quantifying movement that is involuntary and unpredictable. Our goal is to minimize this compromise; and hence, the improvement of the SVS to SVSCD with a future objective of quantifying voluntary movements.

# CHAPTER V

# DISCUSSION AND FUTURE WORK

## I. Discussion

This dissertation is a thorough study of the attributes of dyskinesia, the requirement of an objective measure to quantify this motor dysfunction and the development and validation of severity scores SVS and SVSCD, which assess the severity of dyskinesia using patient videos. Prior work in this area is limited and has focused mainly on establishing parameters that correlate well to various attributes of dyskinesia and in general to a rating scale. These parameters are predominantly kinetic and kinematic in nature and have been obtained using sensor-based devices. We have presented a low-cost, widely usable alternative technique to detect these parameters and combined them to form a severity score that can help the neurologists assess the severity of dyskinesia.

Chapter I discussed the various attributes of dyskinesia, its clinimetrics, advantages and disadvantages of the conventional dyskinesia rating scales. Hence the requirement of an objective dyskinesia rating technique is stressed. We mentioned the existing handful of objectively quantifying techniques, all of which use a sensor-based approach and emulate a discrete rating scale by classifying patients into categories of mild, moderate and severe dyskinesia. At this point, we emphasized the advantages of moving towards a continuous scale. We ended this chapter with a brief description of our technique, validation studies performed on it and an overview of the dissertation. There are three main contributions of this work and we discussed each of these contributions in detail in Chapters 2, 3 and 4.

In chapter 2, we have stressed three novel and significant features:-

1) Patient videos, which are part of patient clinical records, are used to perform movement analysis by tracking the patients' dyskinetic motion. Most Parkinson's disease patients are elderly people and it is not an easy task to recruit them for research studies that requires their participation, wearing devices and specialized motion capture suits. In contrast, we have proposed to extract equivalent information from existing/new patient videos, which makes our technique a patient friendly, low-cost option that allows retrospective analysis of large amounts of patient data without the need for setting up new studies. One advantage of sensor-based approaches is the availability of 3D movement data, whereas, our technique uses only 2D image data. We can capture side-ways movement, but not to and fro movement. We do not see this as a major disadvantage as dyskinetic movements are rarely uni-planar and are usually random involving significant sideways swaying. We believe that our score will not be largely affected by the depth information.
2) The adaptive bases algorithm, an intensity-based registration algorithm was used to track the movement of patient body parts. In our pilot study, we showed that the head and shoulders can be simultaneously tracked to quantify dyskinesia in the reading task of UPDRS patient videos. Good tracking results were observed for short 2s video sequences. This is equivalent to the position data one might obtain using a kinematic sensor. One of the disadvantages of registration based tracking is the machine time required to perform the registration.

3) The severity Score SVS was developed using PCA of the trajectories obtained by frame-by-frame tracking of the head and shoulders of the patients. The total variance, the number of significant modes of variation, and the standard deviation of the percentage contribution of the eigenvalues to the total variance represented the total amplitude of movement, directions of simultaneous movement of body parts and the distribution of movement along these random directions respectively.

Based on these preliminary results, we performed extensive validation studies that were discussed in Chapter 3. The SVS was tested on longer video sequences of 10s and multiple body parts comprising of heads, shoulders, chest, knees, feet, arms and the reading material were tracked successfully. By this process of simultaneous tracking, we emulated a neurologist's strategy to assess dyskinesia looking at the entire patient instead of focusing on individual body parts. Since the SVS is a continuous score, we developed a ranking protocol based on general dyskinesia ratings and compared the rankings of the SVS with the rankings of four expert neurologists and to the UDysRS rankings. We observed a moderate correlation between our semi-automatic score and manual rankings. These findings were based on the communication task videos obtained from the UDysRS trial studies at the Rush University medical center.

The tracking method was applied on drinking and walking tasks as well, but the results exhibited poor tracking quality and did not qualify for further testing of the score. Figure 5.1 shows an example of failed tracking in the drinking task. One of the possible directions of future work is to utilize a better image tracking method to track patients performing activities of daily living. Since these movements are complex, we would need

control data comprising of non-dyskinetic patients or normal people performing these tasks, to observe the differences between normal and pathological movements.

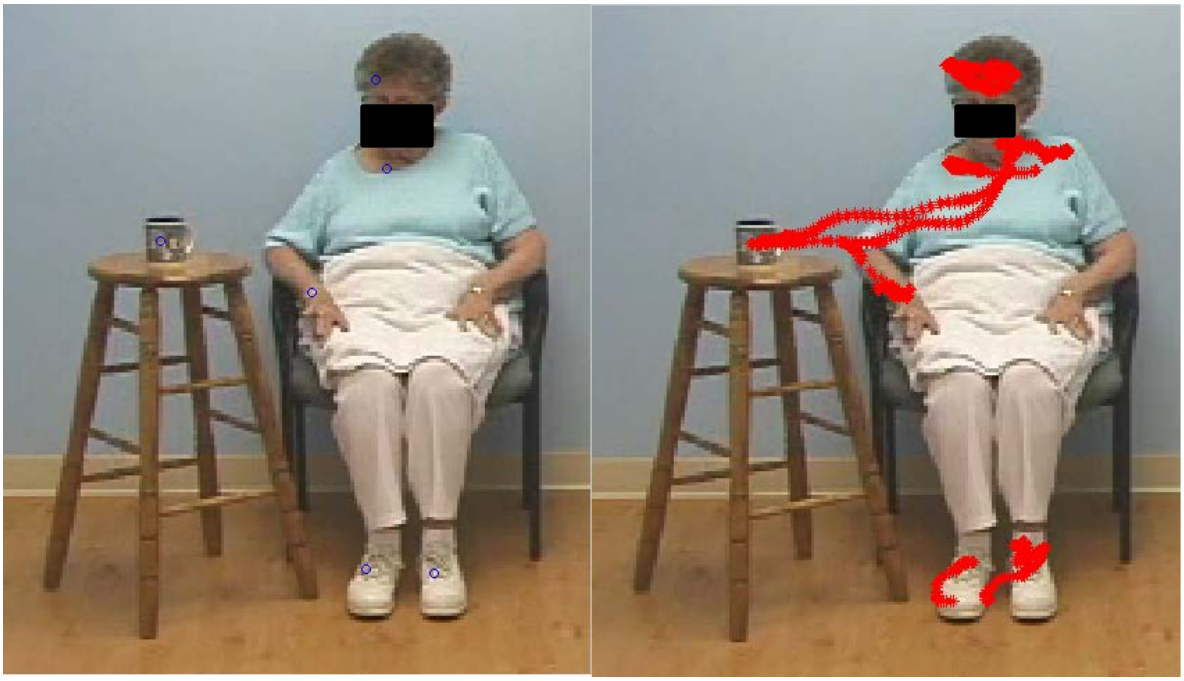

**Fig. 5.1** A moderately dyskinetic patient performing the drinking task. The first figure is the first frame with the landmark points in blue circle. The second frame shows the trajectories of these landmark points and due to occlusion of the landmark point on the chest by the point on the hand, the tracking fails.

We also validated the robustness of SVS by studying its variability with (a) variations in the landmark points which are picked manually on the first frame, and (b) longer video sequences between 2s – 10s. The SVS is reasonably robust to variations in the positions of landmark points and exhibited a high correlation coefficient of 0.9 between the three trials. The two main concerns in picking these landmark points is to maintain the sequence in which the landmark points are picked in all the trials in every patient; and to pick points in regions of high intensity inhomogeneity. We have shown that 10s of data is sufficient to capture the severity of dyskinesia.

We also validated our ranking protocol by performing intra- and inter-rater variability studies. A high agreement between and within the neurologists showed that our ranking protocol was appropriate for validating SVS. We would like to further emphasize the significance of our ranking protocol in future studies involving other tasks. Since we have already established the validity of the SVS score and the ranking protocol for the communication task, new patient data can be analyzed easily without re-ranking the new dataset. Given the SVS of a new patient, their relative position in the ranking can be determined and comparing their position to the existing patients in the dataset, a dyskinesia rating can be determined for them. Simple software could be written to perform this analysis automatically.

In chapter 2 and 3, our score was based on the movement amplitude and directions without any information on the type of dyskinesia – dystonia, chorea or tremor. In chapter 4, we revisited the SVS and experimented with power spectrum parameters of the eigenmodes obtained from the PCA analysis. Approximate entropy of the eigenmode time series was found to be redundant as both dystonia and chorea do not occur exclusively in all patients. To resolve these components from dyskinetic movements, we found frequency dispersion analysis beneficial and the SVS was modified based on these findings. The new SVS rankings showed considerable improvement when correlated with the neurologists' and UDysRS rankings. The frequency dispersion parameter gives additional information as to the presence of patterns in the movement trajectories. We have shown that this technique is sound and it follows results published previously on the frequency content of dyskinetic movement. Instead of performing analysis on individual

body parts, our technique is streamlined in the aspect that it looks at the net movement of all the body parts of the patients simultaneously.

One aspect of SVS that we have not discussed is its units. Based on its formula, the units of SVS is $pixel^4$ per Hz, where distance is measured in pixels and hence units of variance is $pixel^2$ and the units of STDP is $pixel^2$ per Hz. . The position coordinates are in image domain and since we use uncalibrated videos, we cannot accurately compute their real world equivalent. The software used to pick these points translated 1 pixel as 1mm x 1mm, but it would not be an accurate assumption without actually calibrating the cameras. Since all the videos are captured using the same camera, in the same room with the same background and approximately the same focal distance between camera and patients, our comparison across the RUMC dataset is valid. But if videos, for example from the VUMC dataset, are used; we would have to normalize them to make appropriate comparison. Since focal distances from the camera were not the same for all VUMC videos, by performing simulations, we determined the effect of scaling factors on SVS. It was observed that when a scaling factor was applied to the positions of the landmark points making them appear zoomed in (scaling factor > 1) or zoomed out (scaling factor < 1) , the eigenvalues were scaled by the square of the scaling factor. The distance between eyes is approximately similar for most people and we used this measure to determine the scaling factor in videos. One of the patients, P, was assigned to have a scaling factor of 1 and the distance between the eyes (measured in number of pixels) of all other patients was normalized to this patient. So any patient who was farther away the camera as compared to P, exhibited a zoomed out effect and their scaling factor was less than 1 and patients closer to the camera than P had a scaling factor greater than 1. A

similar normalizing procedure can be applied to compare SVS scores of videos obtained from different sources, thereby making our score readily usable for retrospective analysis of existing videos in several databases.

Though we have made significant progress in quantifying dyskinesia in patients using video data, there still exist certain areas which can be improved upon. We discuss these briefly in the following section.

## II. Future Work

Though the main focus of this work is to quantify dyskinesia, the parameters of the score are dependent on the tracking of the patients in the video sequences. Hence there are two parts to our approach – tracking and score development. Given that this is the first body of work using this approach in this field, we have made substantial progress in establishing a strong foundation to our approach. Our technique can be further improved in the following aspects.

### II.1 Tracking

We have used an intensity-based registration algorithm, ABA, for tracking the movement of patients in the consecutive frames. Our tracking fails in instances when there is insufficient intensity inhomogeneity on the patient's clothes and there is a tendency for the landmark points to slip away. One of the simplest solutions to continue using the registration technique is to apply stickers of contrasting colors on patients on the parts to be tracked. These stickers are easier to track given their color contrast. Stickers are non-invasive compared to markers or devices, but this technique can only be applied to new data. One of the biggest disadvantages to registration is the time taken to registration a

sequence of 300 frames. It takes an average of three hours to complete this process. Hence some of the more complex, model-free computer vision approach such as predictive filters, optical flow etc. can be applied to track the patients. We advocate model free approaches as dyskinesia severity is rated using various tasks, even though we have focused only on the communication task. Hence, a separate model would need to be developed for each new task which will make tracking a tedious process and less universally applicable. Better tracking results can also be obtained by using high speed cameras to collect patient data. We have used a regular commercially available camera with a recording rate of 30 frames per second. High speed cameras of up to 75,000 frames per second are available and will be especially useful in tracking patients with severe dyskinesia characterized by high speed movements. Since our goal was primarily to determine the attributes of dyskinesia and its kinematic equivalents, studying the effect of various tracking techniques was out of scope of this work. We do wish to add that by improving the tracking process, more accurate trajectories can be obtained and hence, a severity score that correlates better with neurologists' rankings.

**II.2 Score Development**

We have introduced a new way of quantifying dyskinesia, which comprehensively combines various attributes of dyskinesia instead of using individual parameters that can be correlated with rating scales. We have already stated the advantages of such a continuous score. Our score was developed and validated using the communication task. We have not looked at tasks which require voluntary actions to be performed.

Voluntary motion can be of two types – a sudden voluntary action performed during the task such as animated hand movement or toe tapping or repetitive voluntary actions such as finger tapping, heel tapping etc. The first type of movement in unpredictable and hence harder to quantify which is the reason most research studies specify that patients suppress all voluntary movements. The repetitive movements are easier to quantify provided they do not occur in the same band of frequencies as dyskinesia. As we have seen, dyskinesia manifests itself in the lower frequency bands of 0.5 – 3.5 Hz. and hence any repetitive motion performed at low frequencies can be misinterpreted as dystonia. This problem can occur in patients whose disability in performing the task due to Parkinson's disease may cause slowness in their movements. Gesture analysis and pose analysis can otherwise be used to quantify the voluntary movement. By including the quantification of voluntary motion, we feel that our severity score would be a comprehensive measure that can be used to quantify dyskinesia in patients successfully.

Sensor based approaches claim that their techniques can be used to capture diurnal changes in dyskinesia over 24 hour periods. Capturing and processing 24 hours of patient video is a tedious process and this may be one potential advantage of sensor techniques over ours. But we could still capture diurnal fluctuations of dyskinesia over different time periods in a day to study ON-OFF states. The accuracy of such studies remains to be seen as image quality is a limiting factor for the success of tracking and hence the score. Another interesting application of our work is studying movement inconsistencies in autism, palsy seizures etc. Though the pathologies are significantly different and the severity score may need to be modified, we believe that this idea could be used

successfully in these areas as well. Thus we have developed a novel technique to quantify drug-induced dyskinesia in Parkinson's disease patients using patient videos.

## REFERENCES


[1] http://www.pdf.org/en/about_pdf

[2] Rosal-Greif V L F, "Drug induced dyskinesias", American Journal of Nursing, 1982 Jan, Volume 82 No.I: 66- 69

[3] Hobart J C, Cano S J, Zajicek J P, Thompson A J , “Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations”, Lancet Neurology, December 2007, Volume 6 (12): 1094-105

[4] Moeslund T B , Hilton A, et al., "A survey of advances in vision-based human motion capture and analysis." Computer Vision and Image Understanding, 2006, Volume 104 (2): 90-126.

[5] Zhou H, and Hu T, "Human motion tracking for rehabilitation--A survey " Biomedical Signal Processing and Control, 2008, Volume 3 (1): 1-18

[6] Wang L, Hu W, Tan T, “Recent developments in human motion analysis”, Pattern Recognition, 2003, Volume 36 (3): 585 – 601

[7] Aggarwal J K, Cai Q, “Human motion analysis: a review”, Computer Vision and Image Understanding, 1999, Volume 73 (3): 428 – 440

[8] Welch G and Foxlin G, “Motion Tracking: No Silver Bullet, but a Respectable Arsenal”, IEEE Computer Graphics and Applications, special issue on Tracking, November/December 2002, Volume 22(6): 24–38

[9] http://www.vicon.com/

[10] http://www.qualisys.com/

[11] http://www.codamotion.com/

[12] Poppe R, "Vision-based human motion analysis: An overview " Computer Vision and Image Understanding, 2007 Volume 108(1-2): 4-18

[13] Goffredo M ; Schmid M ; Conforto S ; Carli M ; Neri A ; DýAlessio T, “Markerless Human Motion Analysis in Gauss–Laguerre Transform Domain: An Application to Sit-To-Stand in Young and Elderly People”, IEEE Transactions on Information Technology in Biomedicine, March 2009, Volume 13 (2): 207 – 216

[14] Goffredo M, Bernabucci I, Schmid M, Conforto S, “A neural tracking and motor control approach to improve rehabilitation of upper limb movements”, Journal of NeuroEngineering and Rehabilitation, 2008, Volume 5:1-5

[15] Goffredo M, Schmid M, Conforto S, D’Alessio T, "A markerless sub-pixel motion estimation technique to reconstruct kinematics and estimate the centre of mass in posturography " Medical engineering & physics 2006, Volume 28 (7): 719-726

[16] Allin, S , Beach C , Mitz A and Mihailidis A, “A Video based analysis of standing balance in a community center” IEEE Engineering in Medicine and Biology Society Conference (EMBC), August 2008: 4531-4534

[17] Chang R , Guant L, Burne J A, “A computer assisted image analysis system for diagnosing movement disorders”, Advanced Topics in Artificial Intelligence, 1997: 290-301

[18] Howard-Lee L G, and Lee I, "Video Analysis of Human Gait and Posture to Determine Neurological Disorders" EURASIP Journal on Image and Video Processing 2008

[19] Tan T , Guan L, Burne J, "A Real-time Image Analysis System for Computer-Assisted Diagnosis of Neurological Disorders " Real-Time Imaging 1999, Volume 5 (4):253-269

[20] Sami A, Karayiannis N B, Frost J D Jr., Wise M S, Mizrahi E M, “Automated tracking of multiple body parts in video recordings of neonatal seizures”, IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2004, Volume 1: 312-315

[21] Emoto M, Hayashi A, Suematsu N, Iwata K, “View Independent Gait Identification Using a Particle Filter”, IEEE Conference on Advanced Video and Signal Based Surveillance, November 2006: 74 – 80

[22] Crum W R T H , and Hill D L G, (2004) "Non-rigid image registration: theory and practice", British Journal of Radiology (Special Issue 2004) Volume 77: S140 - S153

[23] Lucas B D, and Kanade T, “An iterative image registration technique with an application to stereo vision”, Proceedings of 7th International Joint Conference on Artificial Intelligence, August 1981: 674-679

[24] Hager G, and Toyama K, “X Vision: Combining image warping and geometric constraints for fast visual tracking”, European Conference on Computer Vision, 2006: 507-517

[25] Hager G D, Dewan M, Stewart C V, “Multiple Kernel Tracking with SSD”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004, Volume 1: 790-797

[26] Zhu Q, Cheng K T, and Zhang H, “SSD tracking using dynamic template and log-polar transformation”, IEEE International Conference on Multimedia and Expo, June 2004, Volume 1: 723 – 726

[27] Rohde G K, Aldroubi A, Dawant B M, The Adaptive Bases Algorithm for Intensity-Based Non-rigid Image Registration, IEEE Transactions On Medical Imaging 2003; 22 (11): 1470-1479

[28] Hoff, J I , Wagemans, E A H , van Hilten, J J , “Accelerometric Assessment of Levodopa-Induced Dyskinesias in Parkinson’s Disease”, Movement Disorders, 2001, Volume 16 (1):58-61

[29] Keijsers, N L W , Horstink, M W I M , Gielen, C A M , “Online Monitoring of Dyskinesia in Patients with Parkinson’s Disease”, IEEE Engineering in Medicine and Biology Magazine, May/June 2003: 93-103

[30] Guy W , “AIMS: ECDEU assessment manual for psychopharmacology”, US Washington, DC, Government Printing Office, 1976

[31] Parkinson Study group, “Evaluation of dyskinesia in a pilot randomized placebo-controlled trial of ramacemide in advanced Parkinson’s disease, Neurology, 2001, Volume 58: 1660-1668

[32] Goetz C G , Stebbins G T, Shale H M, et al., “Utility of an objective dyskinesia rating scale for Parkinson’s disease inter- and intra-rater reliability assessment”, Movement Disorder, 1994, Volume 9: 390-394

[33] Katzenschlager R, Head J, Schrag A, Ben-Shlomo Y, Evans A, Lees A J, "Parkinson's disease research group of the united Kingdom, Fourteen year final report of the randomized PDRGUK trial comparing three initial treatments in Parkinson's disease", Neurology, 2008, Volume 71: 474-480

[34] Hagell P, Widner H, "Clinical Rating of Dyskinesias in Parkinson's disease: Use and Reliability of a New Rating Scale", Movement Disorders, 1999, Volume 14 (3): 448-455

[35] Goetz C G, Nutt J G, Stebbins G T, "The Unified Dyskinesia Rating Scale: Presentation and clinimetric profile", Movement Disorders, 2008, Volume 23 (16): 2398 - 2403

[36] Burkhard P R, Shale H, Langston J W, Tetrud J W, "Quantification of Dyskinesia in Parkinson's Disease:Validation of a Novel Instrumental Method", Movement Disorders, 1999, Volume 14 (5): 754–763

[37] Keijsers N W L , Horstink M W I M , Van Hilten J J , "Detection and Assessment of the Severity of Levodopa-Induced Dyskinesia in Patients with Parkinson's Disease by Neural Networks", Movement Disorders, 2000, Volume 15 (6): 1104-1111

[38] Keijsers N W L , Horstink M W I M , Gielen S C A M , "Assessment of the Severity of Levodopa-Induced Dyskinesia in Daily Life by Neural Networks", Movement Disorders, 2003, Volume 18 (1): 70-80

[39] Hoff, J I , Wagemans, E A H , Van Hilten, J J , "Accelerometric Assessment of Levodopa-Induced Dyskinesias in Parkinson's Disease", Movement Disorders, 2001, Volume 16 (1): 58-61

[40] Patel A, Sherrill D M, Hughes R, et al , "Analysis of the Severity of Dyskinesia in Patients with Parkinson's Disease via Wearable Sensors", Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks, 2006: 123 - 126

[41] Liu X , Carroll C B , Wang S , Zajicek J , Bain P G , "Quantifying Drug-induced Dyskinesias in the Arms Using Digitized Spiral-Drawing Tasks", Journal of Neuroscience Methods, 2005, Volume 144 (1): 47-52

[42] Gour J , Edwards R , Lemieux S , Ghassemi M , Jog M , Duval C , “Movement Patterns of Peak-Dose Levodopa-induced Dyskinesias in patients with Parkinson’s Disease”, Brain Research Bulletin, 2007, Volume 74 (1-3): 66-74

[43] Chung K A , Lobbs B M , Nutt J G , McNames J, Horak F, “Objective Measurement of Dyskinesia in Parkinson Disease using a Force Plate”, Movement Disorders, 2010, Volume 25 (5): 602 - 608

[44] IT Jolliffe, Principal Components Analysis New York: Springer- Verlag, 1986

[45] Daffertshofer A , Lamoth C J, Meijer O G, Beek P J, "PCA in studying coordination and variability: a tutorial" Clinical Biomechanics, 2004, Volume 19 (4): 415-428

[46] Beleznai C, Bischof B F H , "Multiple Object Tracking Using Local PCA " 18th International Conference on Pattern Recognition 2006, Volume 3:79-82

[47] Beuter A, Legros A, Cif L, Coubes P, “Quantifying motion inn dystonic syndromes: the bare essentials”, Journal of Clinical Neurophysiology, 2004, Volume 21 (3): 209 - 214

[48] Kendall M G, Rank Correlation Methods, Charles Griffin and Company Limited, London, 1948: 25-36

[49] http://www.mdvu.org/library/ratingscales/pd/updrs.pdf

[50] Cootes, T F , Taylor, C J , Cooper, D H , Graham, J , “Active Shape Models – Their Training and Application”, Computer Vision and Image Understanding, January 1995, Volume 61 (1): 38-59

[51] Dryden I L, and Mardia K V, Statistical Shape Analysis John Wiley & Sons, 1998

[52] Shrout P E, Fleiss J L, “Intraclass correlations: Uses in assessing rater reliability”, Psychology Bulletin 1979, Volume 86: 420-8

[53] Manson A J , Brown P, O’Sullivan J D , Asselman P, Buckwell D, Lees A J , “An ambulatory dyskinesia monitor”,Journal of Neurosurgery and Psychiatry, 2000, Volume 68: 196 – 210

[54] Rao A S, Bodenheimer R E, Davis T L, Li R, Voight C, Dawant B M, Quantifying Drug Induced Dyskinesia in Parkinson's disease Patients Using Standardized Videos, Proceedings of the30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2008: 1769-72.

[55] http://www.movement-disorders.org/learn/glossary.html

[56] http://www.physionet.org/physiotools/ApEn/

[57] Kass M, Witkin A, and Terzopoulos D, ``Snakes - Active Contour Models'', International Journal of Computer Vision, 1987, 1(4): 321-331