UNIVERSITY OF CALIFORNIA, SAN DIEGO

Validating Objective Metric for Improving Assessment of Spasticity

A thesis submitted in partial satisfaction of the requirements of the degree Master of Science

in

Electrical Engineering (Signal and Image Processing)

by

Fei Deng

Committee in charge:

Truong Nguyen, Chair
Harinath Garudadri
Ramesh Rao
Michael Yip

2017

The Thesis of Fei Deng is approved and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____
Chair

University of California, San Diego

2017

Table of Contents

List of Figure

List of Tables

Acknowledgements

Chapter 1, in part, has been submitted for publication of An Instrumented Glove for Improving Spasticity, 2016, Saisri Padmaja Jonnalagedda, Fei Deng.

Chapter 2, in full, is a reprint of submitted publication An Instrumented Glove for Improving Spasticity, 2016, Jonnalagedda, Padmaja, Fei Deng, Kyle Douglas, Leanne Chukoskie, Michael Yip, Tse Nga Ng, Truong Nguyen, Andrew Skalsky, and Harinath Garudadri.

Chapter 4, in full is being prepare for publication An Instrumented Glove for Augmenting Spasticity Assessment with Objective Metrics, Saisri Padmaja Jonnalagedda, Fei Deng.

ABSTRACT OF THE THESIS

Validating Objective Metric for Improving Assessment of Spasticity

by

Fei Deng

Master of Science in Electrical Engineering (Signal and Image Processing)

University of California, San Diego, 2017

Professor Truong Nguyen, Chair

An instrumented glove worn by doctor to improve subjective assessment of spasticity with an objective metric in order to improve inter- and intra- reliability. The glove contains force sensor array and an inertial measurement unit (IMU). A mock patient is constructed with a mechanism to adjust the arm stiffness, a loadcell and a gyroscope in order to measure the work done to move the arm. The mock patient is the ground truth for validating the proposed objective metric. The force measured by the loadcell in moving the arm of the mock patient and the force estimated by the glove in moving the arm has correlation of 0.83 with untrained users. With

experts trained un spasticity assessment, the correlation is 0.84. The mock patient validated the proposed objective metric, since measurement from the loadcell and the estimation from the glove show a consistent inter- and intra- reliability. Such result show promise of the proposed objective metric improves the assessment of spasticity.

Introduction:

Spasticity is a neuromuscular consequence of brain damage which will result stroke, tumor, or cerebral palsy. Patient with spasticity will suffer under involuntary muscle tone or muscle stiffness that causes resistance of movement. Spasticity is assessed under involuntary motion by feeling the resistance or muscle tone. The existing assessment methods or assessment tools are not able to produce reliable result. The most common diagnose metric is a subjective metric which the doctor score the condition of the patient based on feeling. This diagnose procedure can result large variation between different doctors. Moreover, inadequate dosage is not able to affect condition of the patient, and over-dosage will result severe side effects such as seizures, blurred vision, and severe rashes. Thus, we designed an objective metric for improving spasticity assessment, and designed validation tools to verify the reliability of the objective metric.

Chapter 1: Prior Work and Invented Method

1.1 Prior Work:

There are many methods and off-the-shell products have been used to diagnose spasticity. They can be categorized into three classes: clinical methods, biomechanical tools, and neuron-physiological assessment tools.

1.1.1 Clinical Methods:

Ashworth and Modified Ashworth Scale (MAS) is the most widely used method for diagnose spasticity [5-6]. MAS is a 6-point scale as shown in table 1. Based on the score, the clinic will determine the dosage of drug. The MAS is a highly subjective measure, which is hard to give a reliable and repeatable score [7-8]. The benefit of the MAS is its simplicity. However, MAS is highly subjective and has high inter-rater and intra-rater variability.

Table 1.1 Modified Ashworth Scale

Score	
0	No increase in muscle tone
1	Slight increase in muscle tone, manifested by a catch and release or by minimal resistance at the end of the range of motion (ROM) when the affected part(s) is moved in flexion or extension
1+	Slight increase in muscle tone, manifested by a catch, followed by minimal resistance throughout the remainder (less than half) of the ROM
2	More marked increase in muscle tone through most of the ROM, but affected part(s) easily moved
3	Considerable increase in muscle tone, passive movement difficult
4	Affected part(s) rigid in flexion or extension

Tardieu and Modified Tardieu Scale (MTS) measure the angle of the passive range under slow velocity and the angle of the catch at high velocity [15]. MTS considers the velocity aspect of spasticity. Moreover, it has a better intra-rater and inter-rater reliability compare to the MAS scale [16]. However, MTS is still a subjective measurement, since the rater gives different score before and after the training.

Other clinical methods such as Hypertonia Assessment Tool, Composite Spasticity Scale, Gross Motor Function Classification System, and King's Hypertonicity Scale were invented but no widely used in diagnose spasticity [19-21].

1.1.2 Biomechanical Tools

Myotonometer is able to measure the amount of muscle-tissue displacement per unit force applied [25-27]. However, it requires a voluntary motion from the patient which would have variance between different patients. Moreover, patients with serious injury or disable may not able to produce such voluntary motion. Wartenberg pendulum test can record the activity of the leg during swing. However, it can only work for the leg, and it is not suitable for other part of the body. Dynamometry can record the force and velocity during passively stretched motion. There is an experiment tries to correlate the torque during motion with the slope of work methods of spasticity. However, the standard clinical assessment tool for spasticity is the MAS scale, and there is no experiment performed to correlate the result from dynamometry and its corresponding MAS value. Inertial sensor can record the angular position and angular velocity during the muscle flexion and extension. However, the amount of force used to perform such motion is an

important aspect for spasticity assessment, and such sensor is not recording it. Stiffness tool with robotic-assisted gait orthosis is implemented for measure the stiffness by produce torque to knee and hip. However, there is only one experiment tried to correlate the measurement with GMFCS, instead of MAS scale.

1.1.3 Neuron-Physiological Assessment Tools

Some of the neuron-physiological assessment tools such as electromyography, tonic stretch reflex testing, and H-reflex were related to spasticity assessment [22-23]. First, like some of the biomechanical tools described in section 1.1.2, it requires voluntary motion by patients. Moreover, there is no defined translation from the measurement in the neuron-physiological testing to level of spasticity.

1.2 Method:

1.2.1 Instrumental Glove

The approach to improve spasticity assessment is construct an objective metric for diagnose spasticity. The objective metric is achieved using an instrumental glove as shown in Figure 1.1. The glove would be worn by the doctor to record the motion of the doctor during the clinical trial. The glove is integrated with force-dependent resistive sensor array (from Tekscan, [41]) and an inertial measurement unit (IMU) consisting an accelerometer, gyroscope and a magnetometer [42]. We did an experiment to compare the effect of using golf glove verses baseball glove. Golf glove is thinner compare to the baseball glove, but it is not robust enough to hold the force sensor array attached to it

Another experiment is done for decide wither the force sensor array should attach to the glove using double sided tape or glue. However, the glue can result permanent damage to the force sensor array. Thus, the resulting design is attaching the force sensor array on a baseball glove with double sided tape.

Figure 1.1: Force sensor and IMU on the glove.

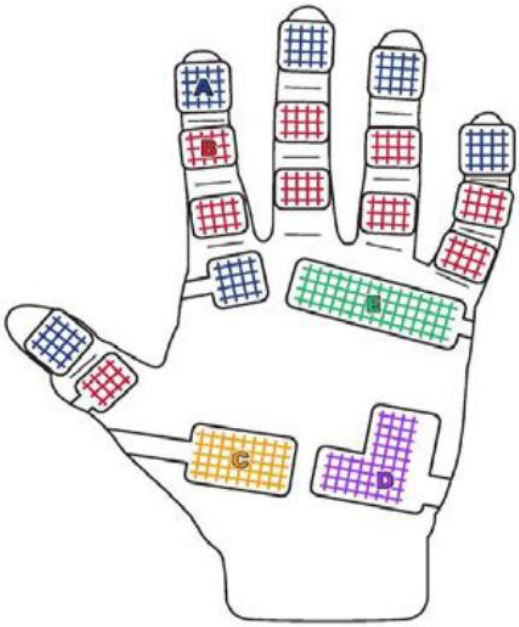

Figure 1.2 Cartoon of the sensing regions and sensing units

The force sensor contains 18 regions and 349 sensing units as shown in Figure 1.2. Each sensing unit can record the applied force (Newtons) on it. The resolution of the output for each sensing unit is 8-bits, and the data is collected in 100Hz. The IMU is attached on the back of the glove, and only the gyroscope is being used in this research to record the angular velocity during the motion. During the experiment, doctor will wear the glove to perform cycles of movement with the patient.

After the glove was constructed, one experiment has done with two doctor worn the glove and tested it across five patients with spasticity. However, the agreement between two doctors was only 27% for the MAS rating. Thus, it is impossible to use the doctor's rating to validate the estimation of the glove. Due to this fact, we created a mock patient as the ground truth in order to validate the performance of the instrumental glove.

1.2.2 Mock Patient

After the development of the glove, we cannot validate the glove with doctors since the score of the doctors is not reliable. Thus, it is necessary to develop another metric to validate the accuracy and reliability of the glove. The metric developed for validation of the glove is a mock patient as shown in Figure 3. The mock patient it is built to simulate the muscle stiffness of the real patient. The arm has a lever connected to a disc clamped by a 5"C-clamp with stationary-bike brake pads. The level of stiffness of the mock patient can be adjusted by changing the pressure from the braking pads applied to the C-clamp. Using this structure, the mock patient can simulate the muscle stiffness of the arm.

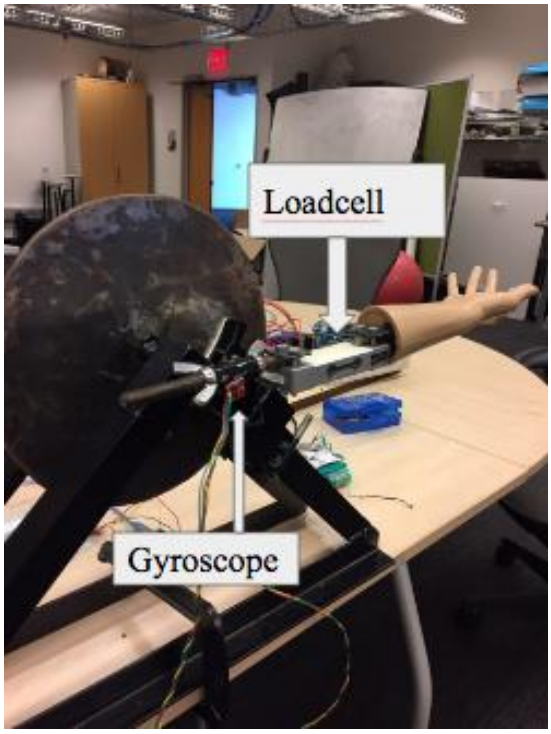

Figure 1.3 Model of the mock patient

The mock patient consist two sensors: loadcell and gyroscope. The loadcell is a wheatstone bridge that output voltage proportional to the force applied. The range of the output voltage is in the order of millivolt. Thus, the HX711 chip is used for as both amplifier and ADC [43]. The chip can amplify the signal into the range of 0-5V, and it can take samples with 80Hz. The data is recorded using Arduino Uno. The gyroscope used in the mock patient is ITG3200 [43]. It can communicate with Arduino Uno under I2C interface. It will record the angular velocity of row, pitch, and yaw in 80Hz. The data is also recorded using Arduino Uno. In the computer, the Arduino Uno is controlled using a python script.

1.2.3 Implement the System for Data Extraction

As discussed in section 1.2.1 and section 1.2.2, the mock patient contains two sensors: loadcell and gyroscope. Moreover, the instrumental glove also contains two sensors: force-sensor array and gyroscope. Since all the sensor is off-the-shell products, each sensor will collect data with different duration and sampling frequency. Data string from each sensor will store into its own file, and there is a time stamp stored for each sample. Based on the time stamp from each sensor, the system will extract the overlapping portion of the data from each sensor, and synchronize them.

Moreover, since the loadcell and gyroscope in the mock patient collect data in 80Hz, and the force sensor and the IMU on the glove collect data in 100Hz, all the data would be down-sample to 20Hz.

1.2.4 Data Processing [46]:

The data collected from force sensor array on the glove and the loadcell on mock patient is shown in Figure 4. The top plot is the data from the force sensor array on the glove. As mentioned in section 1.2.1, the glove has 349 sensing units. The plot was generated by taking the sum of the force sensed by all 349 sensing units. The bottom plot in Figure 4 is the data from loadcell. The data collected from IMU on the glove and the gyroscope on the mock patient is shown in Figure 5.

After obtain the data from both the glove and mock patient, the data need to be process as described by Jonnalagedda. For the glove, a force index is generated from the force sensor array for each trial. A velocity index is also generated from the IMU on the glove. The same process also used with the data from the loadcell and data from the gyroscope on the mock patient. The data collection and final result is described in chapter 2.

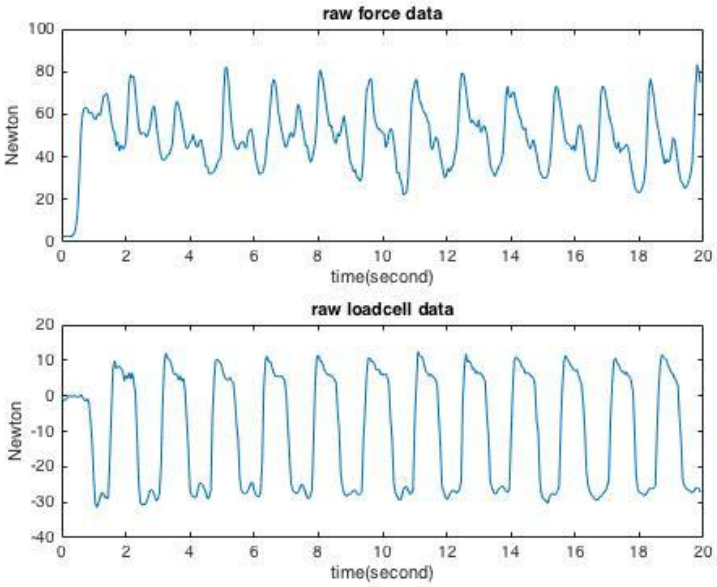

Figure 1.4 Raw force data from the force sensor array (Top)
Raw data collected from the loadcell on the mock patient (Bottom)

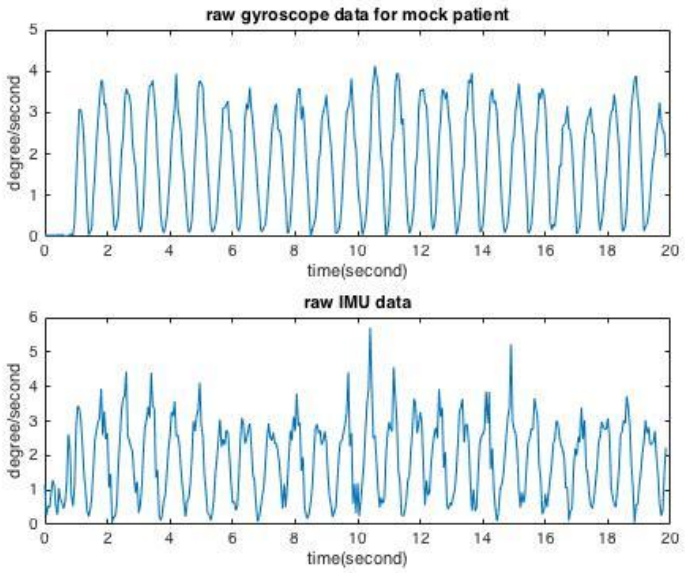

Figure 1.5 Raw data from gyroscope on the mock patient (Top)
Raw data from the IMU on the glove (Bottom)

Acknowledgement

Chapter 1, in part, has been submitted for publication of An Instrumented Glove for Improving Spasticity, 2016, Saisri Padmaja Jonnalagedda, Fei Deng.

Chapter 2: An Instrumented Glove for Improving Spasticity Assessment

Padmaja Jonnalagedda, Fei Deng, Kyle Douglas, Leanne Chukoskie, Michael Yip-IEEE Member, Tse Nga Ng-IEEE Member, Truong Nguyen-IEEE Member, Andrew Skalsky, Harinath Garudadri-EMBS Member

2.1 Abstract

An instrumented glove worn by caregivers that can augment subjective assessments of spasticity with an objective, repeatable metric with reduced inter- and intra- rater variabilityand improved resolution over existing standards is highly desirable. We present the design and preliminary results of such a system using commercial, off the shelf (COTS) components. The glove includes spatially-resolved, force-dependent resistive sensor elements and an inertial measurement unit. We developed a mock patient that is equipped with a mechanism to adjust the arm stiffness, a load-cell and a potentiometer to measure the work done to move the arm. The mock patient provides ground truth to validate the proposed concept. We report the power measured by the sensors in the mock patient to move the arm and the power estimated by the glove in moving the arm and show Pearson correlation coefficient of 0.64. We observe that raw sensor data and instrumentation errors contributed to significant outliers in these experiments. Initial assessments by clinician show promise of the proposed approach to improve spasticity assessment. Future work includes improvements to instrumentation and further clinical evaluations.

2.2 Introduction

Spasticity is a debilitating condition and the most common physical symptom of

acquired brain injury, stroke, or other neuro-muscular disorders such as cerebral palsy, which affects764,000 people and is diagnosed in two to three live births out of every 1,000 in the United States. Patients with spasticity are unable to produce smooth and fluid limb movements due to the imbalance of signals from the brain and spinal cord to the muscles. The pharmaceutical industry spends billions of dollars developing drugs to relieve spasticity, but these efforts are stymied by the lack of repeatable, objective metrics to quantify the outcomes [1-3]; excessive dosage of drugs to treat spasticity can cause severe side effects such as such as seizures, blurred vision, and severe rashes, while inadequate dosage is ineffective at treating spasticity.

Table 2.1 Modified Ashworth Scale

Score	
0	No increase in muscle tone
1	Slight increase in muscle tone, manifested by a catch and release or by minimal resistance at the end of the range of motion (ROM) when the affected part(s) is moved in flexion or extension
1+	Slight increase in muscle tone, manifested by a catch, followed by minimal resistance throughout the remainder (less than half) of the ROM
2	More marked increase in muscle tone through most of the ROM, but affected part(s) easily moved
3	Considerable increase in muscle tone, passive movement difficult
4	Affected part(s) rigid in flexion or extension

The current benchmark for assessing spasticity is the 6-point modified Ashworth score (MAS) shown in Table 1 [4,5].There are several limitations to this MAS, including poor inter-rater reliability and poor sensitivity to changes in spasticity [6-8]. An approach that allows reproducible assessment with improved resolution is urgently needed to monitor patient progress under medication and eliminate negative reactions.

This research is aimed at improving spasticity assessment by augmenting MAS with an objective, repeatable measure that shows finer level of details than MAS and has

reduced variability in intra-rater and inter-rater scores. In Section 2.3, we present prior efforts to improve spasticity assessment. In Section 2.4, we present the development of an instrumented glove that senses pressure and hand motion during spasticity assessment. Since MAS is a highly subjective rating, we initially lacked a reliable criterion measure for verifying the glove measurements. To overcome repeatability issues, we present in Section 2.5 the development of a mock-patient that was used to generate a "ground truth" criterion metric for validating objective scores from the glove. We present experimental results in Section 2.6. In Section 2.7, we discuss sources of errors in the instrumentation, present future work and conclusions.

2.3 Prior Work

Many researchers have taken different approaches to address the lack of quantitative assessment of spasticity. Wearable devices [17, 18, 20] and EMG sensors [19] have been deployed on patients to detect spasticity symptoms, but the drawback is that such devices can be inconvenient and uncomfortable for the patient.

Studies using electromyography (EMG) sensors [9, 19, 21]were carried out on patients with spasticity to characterize the patients' muscle tones under flexion and extension. Wu et al.[9] measured the catch angle reliably by determining the instantaneous velocity and the time derivative of torque. Research by Park et al. [10] also targeted measurement of catch angle and elbow range of motion. Both of the above studies were focused on identifying the presence/absence of a catch phase for correlation to a MAS score between 1 and 2, but these studies did not provide a continuous scale to quantify the different levels of severity.

The lack of a quantitative scale for spasticity was addressed by development of

musculoskeletal models [11] or haptic simulations [12] to determine key physical parameters that contribute to spasticity. One of the most common models is the Haptic Elbow Spasticity Simulator (HESS) [5], [6], [7],in which the properties of spasticity are simulated with the muscle resistance as torque and the catch phase as an impulse. Development of the HESS simulator mainly benefits the doctors as they can practice MAS assessments without requiring actual patients. Their research focused on modeling of spasticity and emphasized on the factors that characterized each MAS level. Alternatively, a mathematical model by Zakaria et al. [16] formulated the resistance as torque and accounted for additional parameters such as the angular velocity, modulus of elasticity etc. The above models have yet to be translated into physical tests that can be implemented on patients to track the spectrum of spasticity conditions.

2.4 Instrumented Glove

Our approach to improve spasticity assessment is an instrumented glove worn by the doctor during patient evaluation. We integrated a spatially-resolved, force-dependent resistive sensor array (by Tekscan, [22]) and an inertial measurement unit (IMU) consisting an accelerometer, gyroscope and a magnetometer [23].

The force sensor on the glove measures the contact force being applied to move a patient's limb. The level of muscular resistance to motion indicates severity of spasticity. Figure 1shows the force sensor integrated on to a golf glove. It has 18 sensing regions, with a total of 349 sensing elements that output a voltage proportional to the applied force. The raw output is a spatial map of 8-bit values for each sensing element. The data was collected at 20Hz. For our analysis, we used the sum of the output of all the sensing elements.

During the experiment, the researchers wore the glove and performed cycles of movement with the patient, such as elbow flexion and extension, and the sensor recorded the force F (Newtons) versus time as shown in Fig. 2.2.A. The IMU is attached to the back of the glove as shown in Figure 2.1 (right). It is used to characterize the hand maneuvers during clinical assessment of spasticity. In this work, we use only the gyroscope data to estimate the power needed to manipulate a limb.

The IMU data is collected at 20 Hz. The angular velocity v from gyroscope is converted to linear velocity at the location of the grip in the mock patient. The gyroscope data in a typical maneuver is shown in Figure 2.2B. We estimate the power to move the patient's limb as F*v.

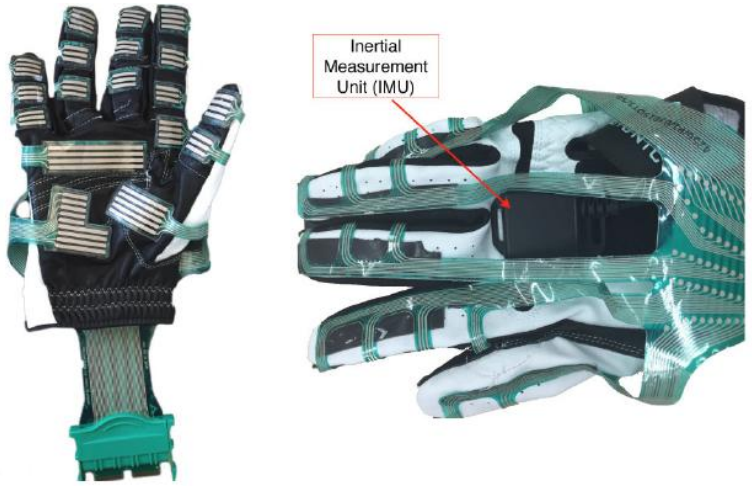

Figure 2.1: Instrumented Glove. The IMU in installed on the back of the glove

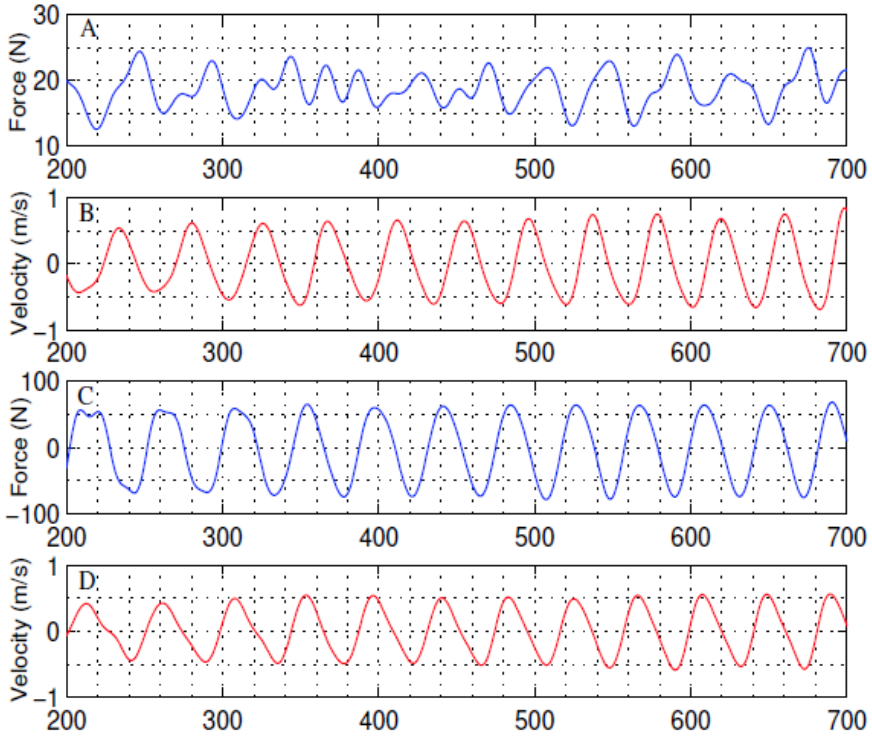

Figure 2.2: Low pass filtered raw sensor data. A. Total pressure from glove sensors over time. B. Linear velocity from gyroscope data. C. Force from load cell data. D. Linear velocity from differentiated potentiometer data. The positive half cycle corresponds to flexion and the negative half cycle corresponds to extension.

In our initial study, five individuals with cerebral palsy volunteered to participate in this study. Participants and/or their parents provided informed consent as per the UCSD Human Subjects Internal Review Board regulations. Participants engaged in a modified Ashworth scale assessment with two physicians well-trained in this methodology (AS and his colleague) and then again by the same two physicians while wearing the spasticity measurement device. These data were collected in UCSD's Research on Autism and Development Laboratory.

In this experiment, there was substantial inter-rater variability resulting in only 27% agreement in MAS values. Consequently, we were not able to use these data to validate the estimates from the glove sensors. To mitigate this, we created a mock patient capable

of generating criterion metric(ground truth) that can be used to validate the objective numbers estimated from the glove sensors.

2.5 Mock Patient's arm structure

The mock patient has an arm structure as shown in Fig. 2.3. The arm has a lever connected to a disc clamped by a 5" C-clamp with stationary-bike brake pads, such that the resistance can be changed manually. The arm has an embedded load cell (model HX711 [24]) that senses the dead weight *m* due to the resistance set by the clamp. We compute the force to overcome this resistance as $F = ma$, where $a$is standard gravity, 9.8 m/s2. We use the term "preset resistance on the mock patient" to denote the force required to move the arm. The units are Newtons. The mock patient also has a potentiometer [25] to sense the angular velocity *v* during flexion and extension. We use this to measure the power as $F \times v$, in N-m/s. In our experiments, we measure the power from the mock patient sensors and use it compare with the power estimated from the sensors in the glove worn by the rater.

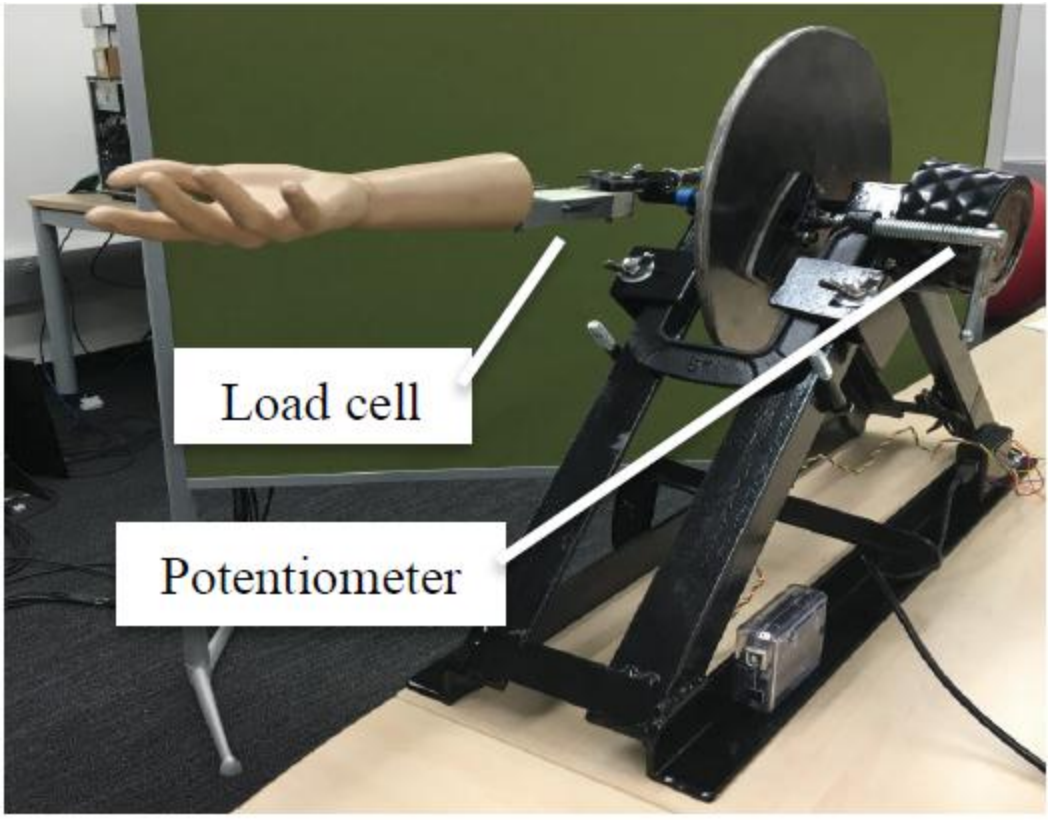

Figure 2.3: Model of the mock patient

2.6 Result

We investigated the agreement between measured power from the mock patient and estimated power from the glove. We focused on MAS values of 1+, 2 and 3 in this study. The values of 0 and 4 are easy to assess since they correspond to normal tone and rigid limbs, respectively. Similarly, a value of 1 is also easy to assess since it is characterized by catch and release. A well-trained physician in spasticity assessment (AS)tested the mock patient and identified the range of to be 20—90 Newtons for MAS values of 1+, 2 and 3. Spasticity is a highly velocity driven response [1], [2], [8], [14]. For both glove and mock patient, we converted the angular velocity to linear velocity (see Figure 2.1) and estimate the power to move the patient's limb as $F*v$. Here, we present experimental results for two trials by 4 researchers.

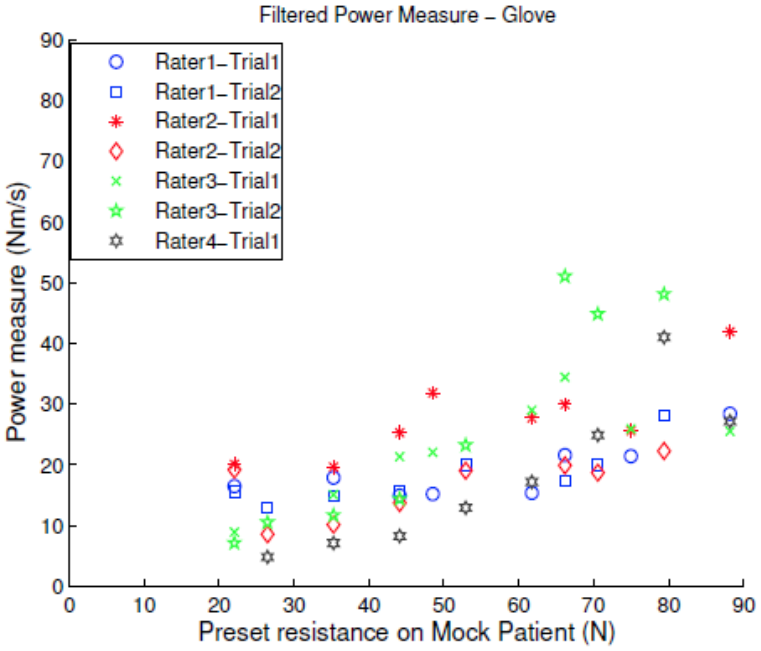

Figure 2.4 Power measurements for the instrumented glove

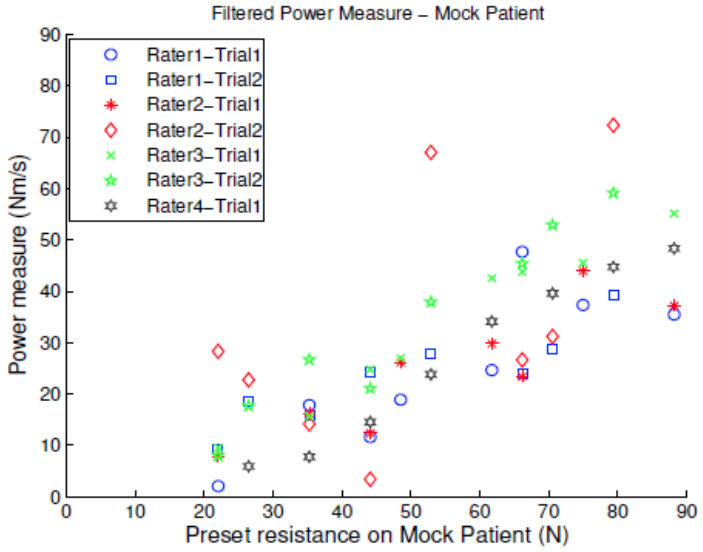

Figure 2.5 Power measurements for the instrumented the mock patient

Figures 2.4 and 2.5 show the power measured from the glove sensors and mock patient sensors, respectively, for different preset resistances on the mock patient. Note that while there are outliers in both cases, the mock patient data shows better agreement

with the preset resistances, compared to that of the glove. From Figure 2.1, it can be seen that the force data from glove does not follow the cyclical nature of other sensors. Figure 2.6 shows the power measured from the mock patient sensors versus power measured from the glove sensors. We note that there are bias and variability issues in all these experiments. The Pearson correlation coefficient between the mock patient and the glove was 0.64. When we compute the agreement between the mock patient and glove for flexion and extension independently, the Pearson coefficients were 0.64 and 0.57 respectively. The experimenters gripped the mock patient at the wrist – flexion involved in pushing the mock patient arm, while extension involved pulling it.

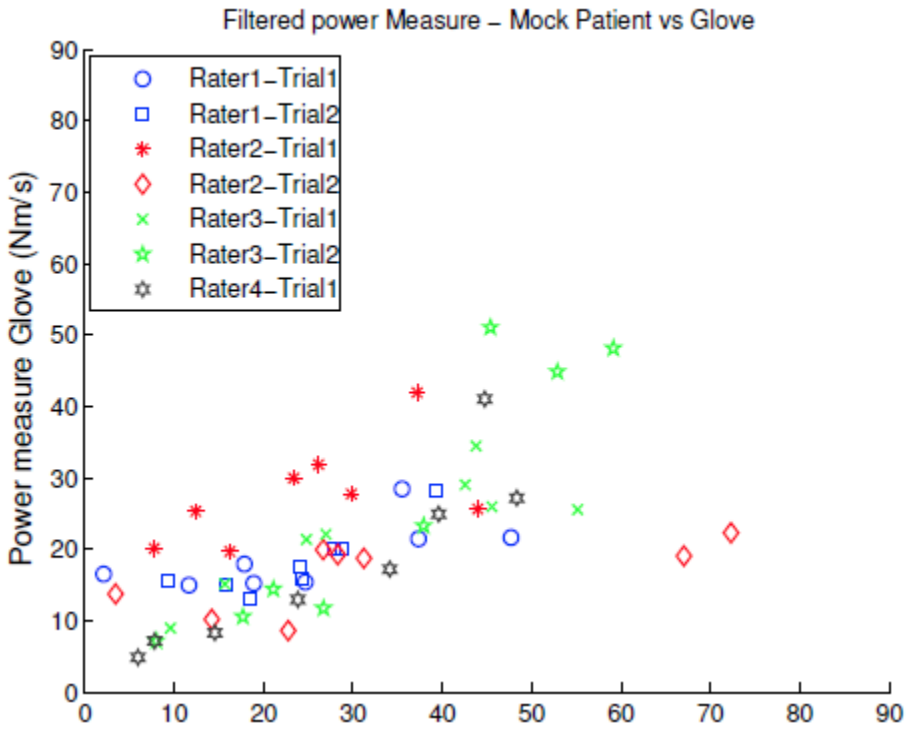

Figure 2.6 Instrumented glove versus mock patient

We performed another experiment with a physician (AS) performing MAS assessment for various resistance settings of the mock patient, as shown in Figure 2.7. The physician did not know the resistance setting so that he could provide an unbiased assessment. This shows the promise of improving MAS ratings resolution with the instrumented glove.

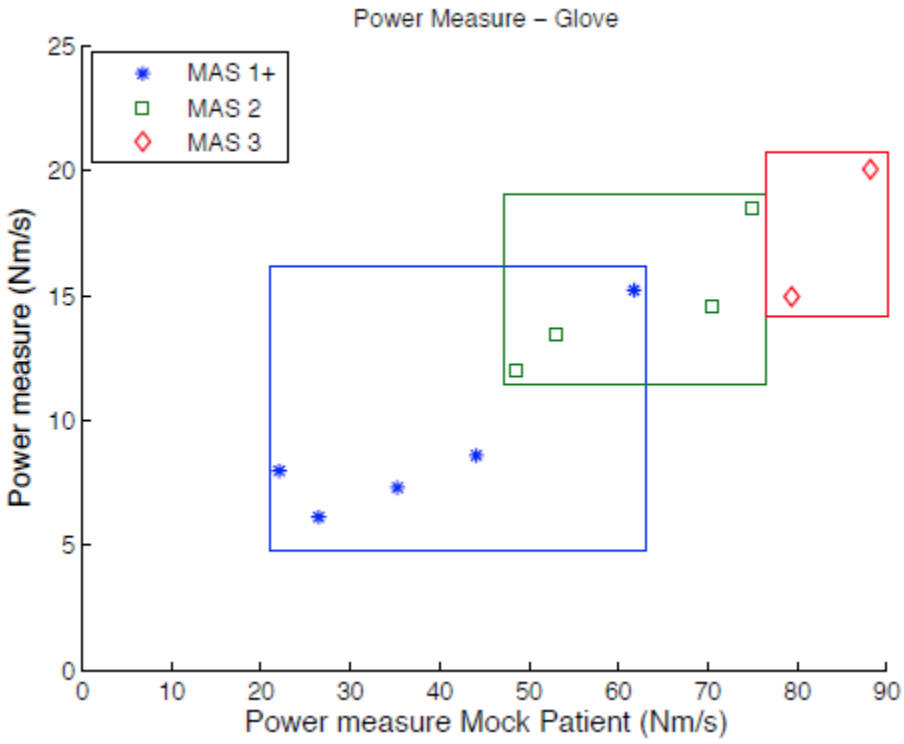

Figure 2.7 Power estimates for MAS value

2.7 Discussion and Future Work

There are some sources of error such as grip variation, posture, etc. that could introduce certain bias and also result in outliers in the measurements. In addition, we observed certain errors in the pressure sensor, similar to other researchers ([29] reported up to 34% errors). Further, our COTS instrumentation used different clock domains for the potentiometer, load cell, pressure sensor and the gyroscope. This resulted in

significant drift in the alignment between pressure and gyroscope data; load cell and potentiometer data during each experiment. Future work must address (i) improvements in sensor reliability (ii) custom hardware to acquire glove sensor data with a common clock and mock patient sensor data with a common clock (iii) further testing by doctors to understand the statistical validity of results shown in Figure 7.

2.8 Conclusion

Spasticity is a debilitating neurological, musculer-skeletal condition, affecting people with CP, TBI, stroke, etc. This research addresses development of an instrumented glove to be worn by doctors while performing MAS assessment, a gold standard in current standard of care for diagnosis and treatment of spasticity. We presented a design of the glove based on COTS components. In order to develop an objective metric from the glove measurements, we presented the development of a mock-patient arm with adjustable resistance to motion and sensors to report the load and angular displacement. We presented power (N-m/s) measured at the mock patient and estimated by the glove for various stiffness values that correspond to MAS values of 1+, 2 and 3. Our results demonstrate that the instrumented glove has a correlation of 0.64 with the mock patient. Preliminary assessment by a physician demonstrates that an objective metric based on measured power has improved resolution over MAS. Future work will include improvements to sensors, custom hardware to mitigate clock issues and additional characterization in clinical settings.

2.9 Acknowledgements

We thank Arsh Buch, Leon Nguyen, Phillip Duong for developing the instrumented glove and Steven Apodaca, mBenjamin Hobbs, Oscar Guerrero, Jacob

Rozelle, Amy Teshima for developing the mock patient as part of ECE 191, a UCSD project based undergraduate course.

2.10 Reference

[1] Skalsky AJ, Vadivelu S, Dalal P Spasticity, In: Chapman's Orthopedic Surgery. 4th edition, M Chapman ed., Jaypee Medical Publishers, Philadelphia, PA USA (In press).

[2] Skalsky AJ, Fournier CM. Intrathecal Baclofen Bolus Dosing and Catheter Tip Placement in Pediatric Tone Management. Phys Med Rehabil Clin N Am. 2015 Feb; 26(1):89-93. doi: 10.1016/j.pmr.2014.09.011. Review. PubMed PMID: 25479782.

[3] Deon LL, Gaebler-Spira D. Assessment and treatment of movement disorders in children with cerebral palsy. Orthop Clin North Am. 2010 Oct;41(4):507-17.

[4] Ashworth B. Preliminary trial of carisoprodol in multiple sclerosis. Practitioner 1964;192:540–2.

[5] Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. *Phys Ther* 1987;67:206–7.

[6] Pandyan AD, Johnson GR, Price CI, et al. A review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity. Clin Rehabil1999;13:373–83.

[7] Fosang AL1, Galea MP, McCoy AT, Reddihough DS, Story I. Measures of muscle and joint performance in the lower limb of children with cerebral palsy. Dev Med Child Neurol. 2003 Oct; 45(10):664-70.

[8] JFM Fleuren *et al*. Stop using the Ashworth Scale for the assessment of spasticity. J. *Neurol Neurosurg Psychiatry* 2010;**81**:46 52 doi:10.1136/jnnp.2009.177071.

[9] Wu, Yi-Ning, Hyung Soon Park, Yupeng Ren, Deborah Gaebler-Spira, Jia-Jin Chen, and Li-Qun Zhang. "Measurement of elbow spasticity in stroke patients using a manual spasticity evaluator." In *Engineering inMedicine and Biology Society, 2006. EMBS'06. 28th AnnualInternational Conference of the IEEE*, pp. 3974-3977. IEEE, 2006.

[10] Park, Hyung-Soon, Qiyu Peng, and Li-Qun Zhang. "A portable telerehabilitation system for remote evaluations of impaired elbows in neurological disorders." *IEEE Transactions on Neural Systems andRehabilitation Engineering* 16, no. 3 (2008): 245-254.

[11] Xin, Zhang, Norimasa Adachi, and Kazunori Hase. "A computed model for quantitative evaluation of spasticity in man." In *Computational Engineering in Systems Applications, IMACSMulticonference on*, vol. 2, pp. 1844-1847. IEEE, 2006.

[12] Grow, David I., Mengnan Wu, Michael J. Locastro, Sugandha K. Arora, Amy J. Bastian, and Allison M. Okamura. "Haptic simulation of elbow joint spasticity." In *2008 Symposium on Haptic Interfaces for VirtualEnvironment and Teleoperator Systems*, pp. 475-476. IEEE, 2008.

[13] Park, Hyung-Soon, Jonghyun Kim, and Diane L. Damiano. "Development of a haptic elbow spasticity simulator (HESS) for improving accuracy and reliability of clinical assessment of spasticity." *IEEE transactions on neural systems and rehabilitationengineering* 20, no. 3 (2012): 361-370.

[14] Park, Hyung-Soon, Jonghyun Kim, and Diane L. Damiano. "Haptic recreation of elbow spasticity." In *2011 IEEE International Conferenceon Rehabilitation Robotics*, pp. 1-6. IEEE, 2011.

[15] Kim, Jonghyun, Hyung-Soon Park, and Diane L. Damiano. "Accuracy and reliability of haptic spasticity assessment using HESS (haptic elbow spasticity simulator)." In *2011 Annual International Conference of theIEEE Engineering in Medicine and Biology Society*, pp. 8527-8530. IEEE, 2011.

[16] Zakaria, Noor Ayuni Che, Takashi Komeda, Cheng Yee Low, Fazah Akhtar Hanapiah, and Kaoru Inoue. "Spasticity mathematical modelling in compliance with modified Ashworth scale and modified Tardieu scales." In *Control, Automation and Systems (ICCAS), 201515th International Conference on*, pp. 1893-1897. IEEE, 2015.

[17] Ishikawa, Shun, Shogo Okamoto, Kaoru Isogai, Yasuhiro Akiyama, Naomi Yanagihara, and Yoji Yamada. "Wearable dummy to simulate joint impairment: severity-based assessment of simulated spasticity of knee joint." In *System Integration (SII), 2013 IEEE/SICE InternationalSymposium on*, pp. 300-305. IEEE, 2013.

[18] Ferreira, João, Vitor Moreira, José Machado, and Filomena Soares. "Improved biomedical device for spasticity quantification." In *Bioengineering (ENBENG), 2013 IEEE 3rd Portuguese Meeting in*, pp. 1-4. IEEE, 2013.

[19] Misgeld, Berno JE, Markus Lüken, Daniel Heitzmann, Sebastian I. Wolf, and Steffen Leonhardt. "Body-Sensor-Network-Based Spasticity Detection." *IEEE journal of biomedical and health informatics* 20, no. 3 (2016): 748-755.

[20] Peng, Qiyu, Parag Shah, Ruud W. Selles, Deborah J. Gaebler-Spira, and Li-Qun Zhang. "Measurement of ankle spasticity in children with cerebral palsy using a manual spasticity evaluator." In *Engineering inMedicine and Biology Society, 2004. IEMBS'04*

26th AnnualInternational Conference of the IEEE, vol. 2, pp. 4896-4899. IEEE, 2004.

[21] Sherwood, Arthur M., M. M. Priebe, and W. B. McKay. "Quantification of surface electromyographic recordings for assessment of spasticity." In *Engineering in Medicine and Biology Society, 1996. BridgingDisciplines for Biomedicine. Proceedings of the 18th AnnualInternational Conference of the IEEE*, vol. 2, pp. 597-598. IEEE.

[22] https://www.tekscan.com/products-solutions/systems/grip-system

[23] http://www.motionnode.com

[24] https://www.sparkfun.com/products/13230

[25] http://sensing.honeywell.com/honeywell-sensing-potentiometers-lineguide-007067-1-en.pdf

[26] Orfanidis, Sophocles J. *Introduction to Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1996.

[27] J. Lance, Spasticity: Disordered Motor Control, R. Feldman, R. Young, and W. Koella, Eds. Year Book Medical Publishers, 1980.

[28] R. W. Bohannon and M. B. Smith, “Interrater reliability of a modified Ashworth Scale of muscle spasticity,” Phys. Ther., vol. 67, no. 2, pp. 206–7, Feb. 1987.

[29] Wettenschwiler, Patrick D., Rolf Stämpfli, Silvio Lorenzetti, Stephen J. Ferguson, René M. Rossi, and Simon Annaheim. "How reliable are pressure measurements with Tekscan sensors on the body surface of human subjects wearing load carriage systems?" *International Journalof Industrial Ergonomics* 49 (2015): 60-67.

Acknowledgements

Chapter 2, in full, is a reprint of submitted publication An Instrumented Glove for Improving Spasticity, 2016, Jonnalagedda, Padmaja, Fei Deng, Kyle Douglas, Leanne Chukoskie, Michael Yip, Tse Nga Ng, Truong Nguyen, Andrew Skalsky, and Harinath Garudadri.

Chapter 3: Improve the System with Neural Network

As shown in section 2.8, the correlation between the instrumental glove and the mock patient is 0.64. This correlation can be improved with neural network by taking the advantage that the force sensor array on the glove contains spatial information which does not take into account by summing them up.

3.1 Neural Network

The force sensor array on the glove contains 349 sensing elements, and every sensing element record the data at 100Hz. Both the force sensor array and loadcell data would be down-sample to 20Hz. The neural network can map the data from the force sensor array to the data from loadcell.

3.1.1 Loss Function

Let $X = [x_1, x_2, \dots, x_N]$ as the input of the Neural Network where N is the number of sample, and let $Y = [y_1, y_2, \dots, y_N]$ be the output of the Neural Network. In this case, the data from force sensor array is feed into the network, and the network should have output match the data from the loadcell. Since the neural network is performing regression instead of classification, 2-norm loss function is used to evaluate the accuracy. Given the target $T = [t_1, t_2, \dots, t_N]$ and the output of the neural network Y, the loss L is computed as:

$$L = \frac{1}{N}\sum_{n=1}^{N} \|t_n - y_n\|^2$$

3.1.2 The Structure of the Neural Network

The structure of the neural network is shown in Figure 3.1. The neural network has an input layer, one hidden layer and an output layer. There are 100 neurons in the hidden

The neural network has an input layer, one hidden layer and an output layer. There are 100 neurons in the hidden. Let w_{ij} be the weighting matrix between the input layer and hidden layer, and w_{jk} be the weighting matrix between the hidden layer and output layer. Moreover, since the value of the loadcell data is usually in the range of -30 to 30,$tanh$ function is used as the activation function. $tanh$ function can map value to the range of $-1\ to\ 1$, so the loadcell data need to be normalize before training. As discussed in the beginning of section 3.1, the force sensor array has dimension of 349 and there is bias term need to be add to the input, so $dim(x_n) = 350$ and $x_{350} = 1$ for all n.

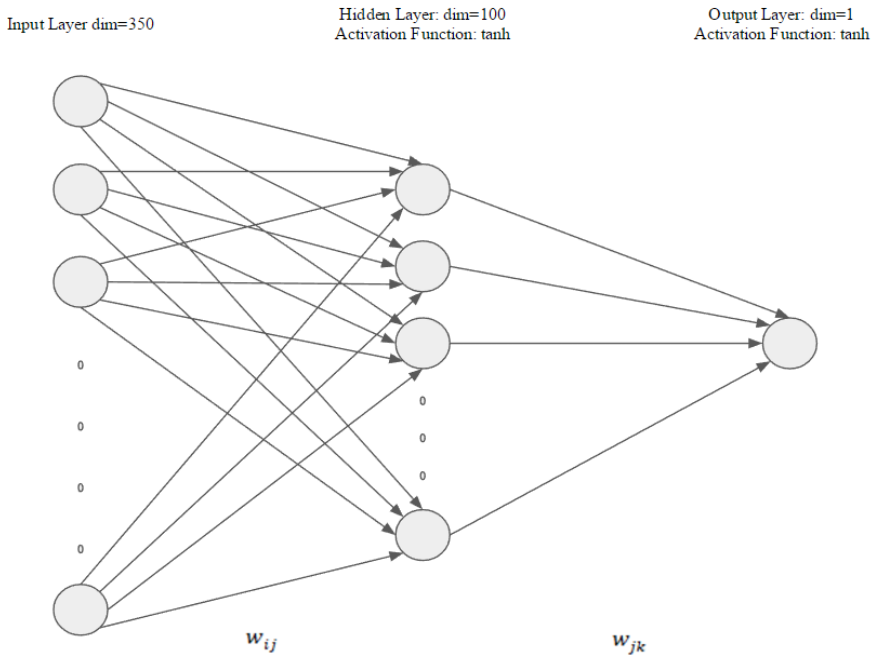

Figure 3.1 The structure of the neural network

Moreover, since the hidden layer has 100 neurons, the size of w_{ij} is 100×350 and the size of w_{jk} is 1×100. Overall, the output of the neural network is calculated as:

$$h_n = tanh(w_{ij}\, x_n)$$

$$y_n = tanh(w_{jk}\, h_n)$$

where h_n is the output of the hidden layer.

3.1.3 Stochastic Gradient Descent

The neural network was trained using stochastic gradient descent. The training dataset was separated into many batches of 100 samples, and the neural network was trained with respect to one batch at a time.

As discussed previously, the neural network is trained in order to minimize the 2-norm loss function. The gradient can be computed as following:

$$\frac{dL}{dw_{jk}} = \frac{dL}{dy_n}\frac{dy_n}{dw_{jk}} = 2(t_n - y_n)\frac{dy_n}{dw_{jk}} = 2(t_n - y_n)\left(1 - tanh(w_{jk}\, h_n)^2\right) h_n$$

$$\delta_j = 2(t_n - y_n)(1 - tanh(w_{jk}\, h_n)^2)$$

Thus:

$$\frac{dL}{dw_{jk}} = \delta_j\, h_n$$

$$\frac{dL}{dw_{ij}} = \frac{dL}{dy_n}\frac{dy_n}{dh_n}\frac{dh_n}{dw_{ij}}$$

$$\frac{dL}{dw_{ij}} = \delta_j\, w_{jk}\,(1 - tanh(w_{ij}\, x_n)^2)x_n$$

After every training epoch, the weighting matrix w_{ij} and w_{jk} can be update as following:

$$w_{ij} = w_{ij} - \eta_1 \frac{dL}{dw_{ij}}$$

$$w_{jk} = w_{jk} - \eta_2 \frac{dL}{dw_{jk}}$$

where η_1 and η_2 are the step size.

3.1.4 Regularization:

In order to prevent the neural network from over-fitting, the loss function should be modified to have a regularization term. The loss function as shown in section 3.3.1 is modified as following

$$L = \frac{1}{N}\sum_{n=1}^{N} \|t_n - y_n\|^2 + \frac{\mu_1}{2}\|w_{ij}\|^2 + \frac{\mu_2}{2}\|w_{jk}\|^2$$

where μ_1 and μ_2 are step size.

The goal of adding the norm of the weights to the loss function is preventing the weight reach extreme value during training. Accordingly the gradient as described in section 3.3.3 also modified as following:

$$\frac{dL}{dw_{ij}} = \delta_j w_{jk} \left(1 - tanh\left(w_{ij} x_n\right)^2\right) x_n + \mu_1 w_{ij}$$

$$\frac{dL}{dw_{jk}} = \delta_j h_n + \mu_2 w_{jk}$$

3.2 Result of the Neural Network

Two sets of data were collected during the study. One set contains 15 data sets from 10 non-clinician raters, and the other contains 6 data sets from 6 clinician raters. Each data set has 6 trials which the mock patient was set to between 20-90 Newton. Each

trial has 20 seconds and since the data is processing at 20Hz, there are 400 samples for each trial. During the training process for both clinician dataset and non-clinician dataset, data from one rater was put into the testing set, and data from all the other raters were put into the training set. This training method can validate the inter-rater reliability of the neural network.

The example of the output of the neural network is plotted in Figure 3.2, 3.2(a) is the plot by summing the reading from all the force sensor array. 3.2(b) is the plot of the output of the neural network, and 3.2(c) is the loadcell reading, which is the target of the neural network.

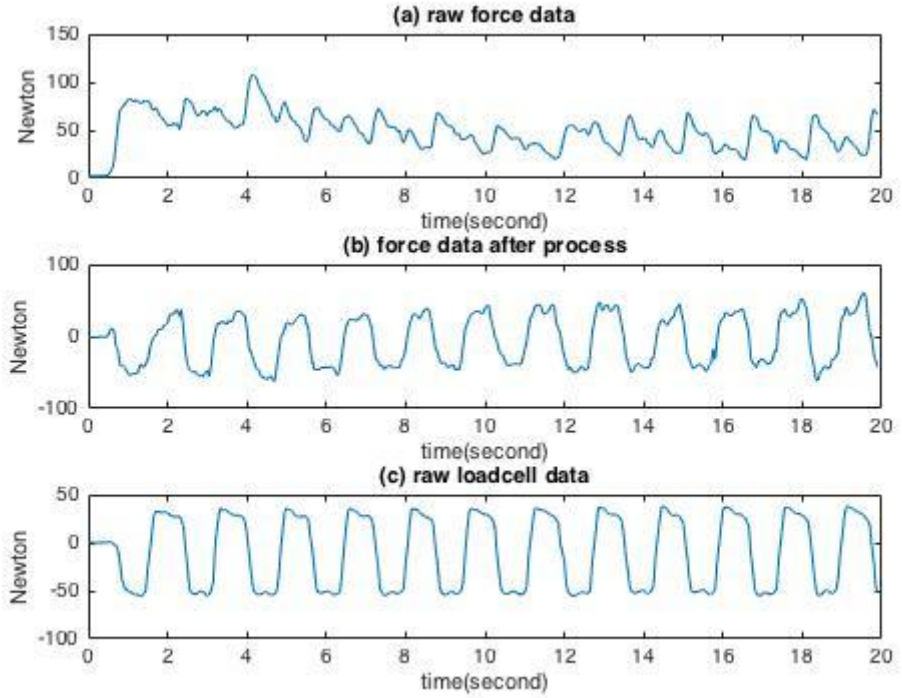

Figure 3.2 (a) Raw force data from the force sensor (b) Force data after the process (c) Raw loadcell data

Since the neural network is performing regression, one way to verify the performance of the neural network is compute the mean and standard deviation of the

error, which is the difference between the output of the neural network and the target for the testing set. Figure 3.3 and Figure 3.4 shows the mean and the standard deviation of the error.

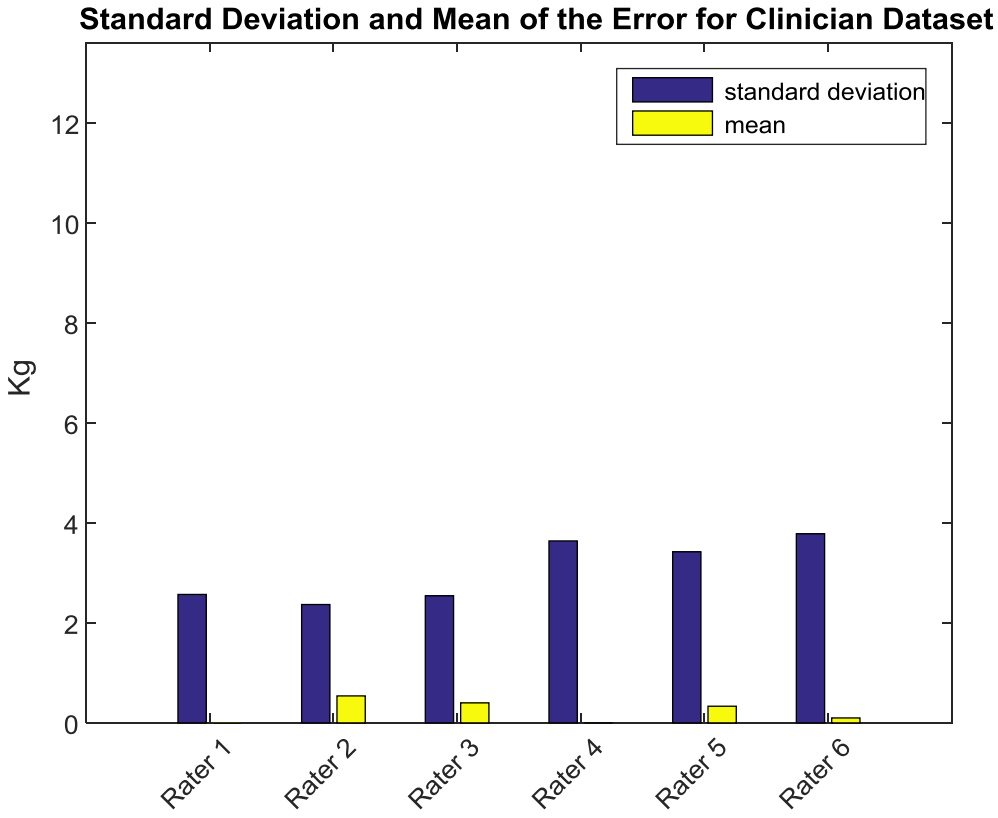

Figure 3.3 The mean and standard deviation of the difference between the output of the neural network and the target for clinician dataset

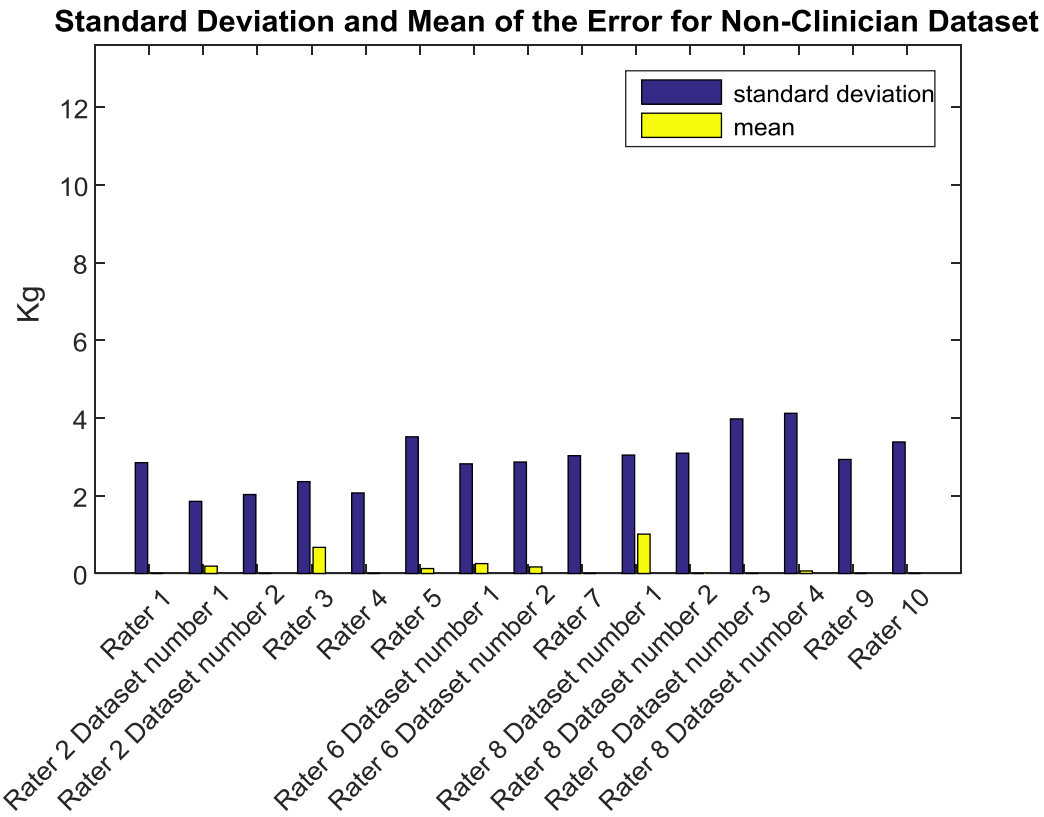

Figure 3.4 The mean and standard deviation of the difference between the output of the neural network and the target for non-clinician dataset

As shown in Figure 3.3 and Figure 3.4, the mean is very low across all the rater. Thus, all the neural networks have small bias. The upper bound of the range of the loadcell reading is 90 Newton, and most of the neural networks have standard deviation around 20 Newton. This is an acceptable result, since the final goal of applying the neural network is to increase the correlation between data from the force sensor array and the data from the loadcell.

Observing the performance of correlation of the neural network is to calculate the correlation between the output of the neural network and the target. The correlation is defined as Pearson correlation coefficient (PCC), which represent linear correlation between two variables.

Given the output of the neural network $Y = [y_1, y_2, \dots, y_N]$ and the target $T = [t_1, t_2, \dots, t_N]$ The Pearson correlation coefficient ρ can be computed as following:

$$\rho = \frac{\sum_{i=1}^{N}(y_i - \mu_y)(t_i - \mu_t)}{\sigma_y \sigma_t}$$

where μ_y is the mean of Y, μ_t is the mean of T, σ_y is the standard deviation of Y, and σ_t is the standard deviation of T.

The Pearson correlation coefficient ρ is in the range of $[-1,1]$, where two variables are completely correlated if $\rho = 1$. $\rho < 0$ when two variables are reverse correlated, and $\rho = 0$ when two variables are not correlated with each other.

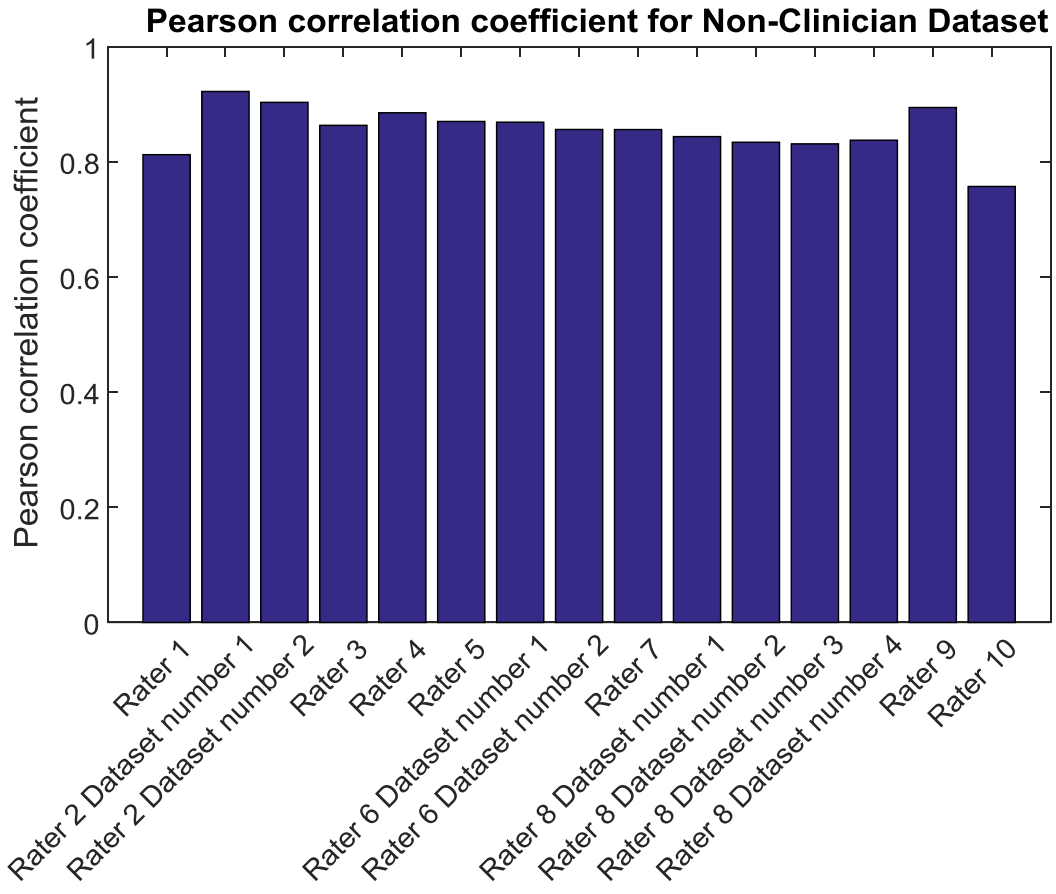

Figure 3.5 The correlation between the output of the neural network and the target for non-clinician dataset

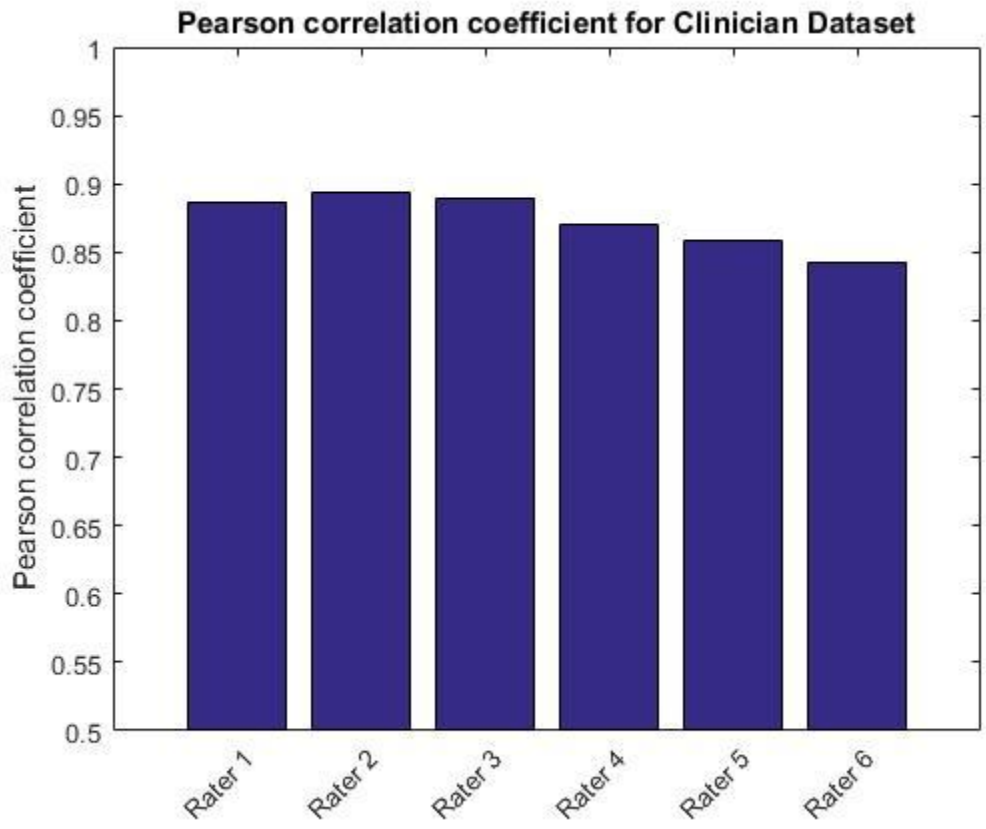

Figure 3.6 The correlation between the output of the neural network and the target for clinician dataset

As observed in Figure 3.5 and Figure 3.6, there is a high linear correlation between the output of the neural network and the target. Thus, a high correlation is expected for the force index from the output of the neural network and the loadcell. The Pearson correlation coefficient for data from all raters in the non-clinician set is 0.83. The correlation is 0.84 for the clinician set.

3.3 The Result of Work Calculation

Two force indexes would be computed, after map the data from the force sensor array on the glove to the loadcell data. One is compute using data from loadcell, and the other is computed using the output of the neural network. The computational procedure is

the same for both, and the goal is to reach a high correlation between the weight index from the output and the force index from the loadcell.

3.3.1 Compute the Weight Index [45]

The force index is computed as described by Jonnalagedda. Given the data from the loadcell $T = [t_1, t_2, \dots, t_N]$, the weight index is computed as following:

$$f_T = \sqrt{\sum_{i=1}^{N} t_i^2}$$

The same procedure is used to compute the weight index from the output of the neural network. Given the output of the neural network $Y = [y_1, y_2, \dots, y_N]$

$$f_Y = \sqrt{\sum_{i=1}^{N} y_i^2}$$

After obtained the weight index of the data from the loadcell and the glove, an linear conversion would be made to convert it into Kg as described be Jonnalagedd

3.3.2 Correlation between the Weight Index

The Pearson correlation coefficient (PCC) was computed after obtain the force indexes from every trial and each rater. As shown in Figure 3.7 and Figure 3.8, the Pearson correlation coefficient between the force index calculated from the loadcell data and the instrumental glove data after the neural network is 0.88 for the non-clinician dataset. Moreover, the PCC is 0.89 for the clinician dataset. Thus, the neural network

successfully correlated the data from the force sensor array on the glove and the data from the loadcell.

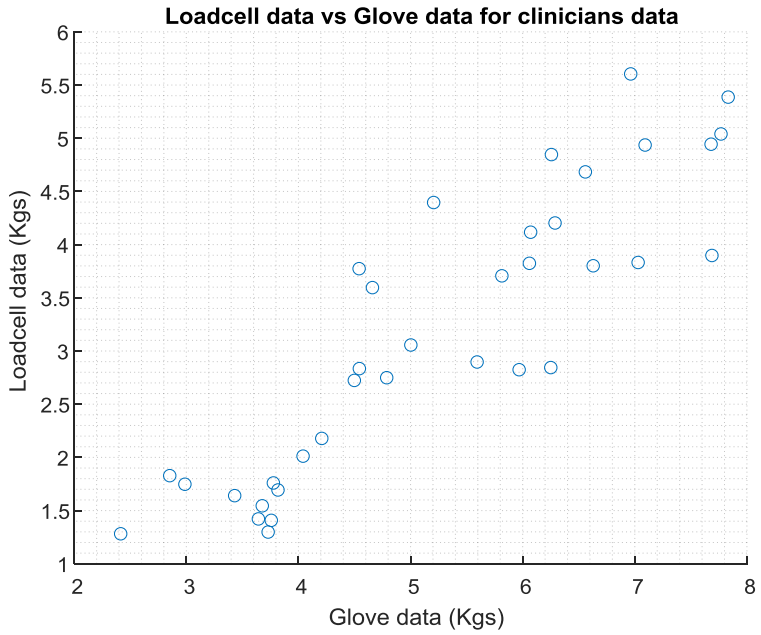

Figure 3.7 Loadcell data vs. data from force sensor array on the glove. The correlation is 89%

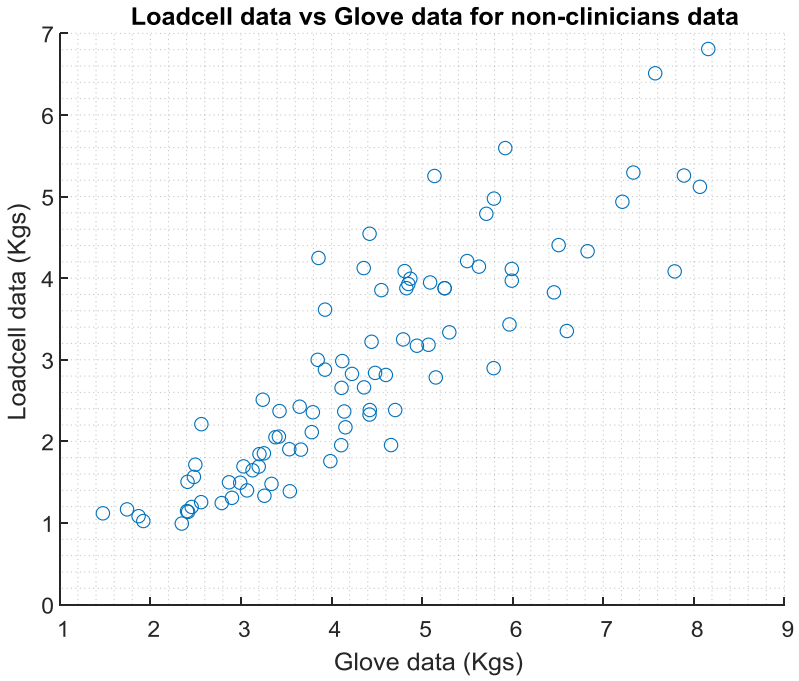

Figure 3.8 Loadcell data vs. data from force sensor array on the glove. The correlation is 88%

Chapter 4: An Instrumented Glove for Augmenting Spasticity Assessment with Objective Metrics

4.1 Abstract

In this contribution, we propose an instrumented glove worn by experts to augment subjective assessments of spasticity with an objective, repeatable metric with reduced inter- and intra- rater variability and improved resolution over current best practices. We present the system design and validation using commercial, off the shelf (COTS) components. The glove includes spatially-resolved, force-dependent resistive sensor elements and an inertial measurement unit (IMU). We describe development of a mock patient equipped with a mechanism to adjust the arm stiffness, a load-cell and an IMU to measure the work done to move the arm. The mock patient provides ground truth to validate the proposed concept. We report the power measured by the sensors in the mock patient to move the arm and the power estimated by the glove in moving the arm and show Pearson correlation coefficient of 0.9 with untrained users. With experts trained in spasticity assessment, the correlation was 0.7 and 0.8 with and without outliers, respectively. We identify the sources of errors during expert assessment trails and the limitations of the COTS realization of the glove and the mock patient. We conclude with recommendations for improving the glove electronics, mock patient realization and guidelines for experts to incorporate limitations of electronics in the proposed system to improve spasticity assessment and patient care.

4.2 Introduction

Spasticity is a neuro-muscular disorder characterized by an increase in muscle-tone or stiffness of the limbs. It often occurs in patients with problems like Cerebral Palsy

Multiple Sclerosis (MS), Traumatic Brain Injury (TBI), Stroke, Spinal Cord Injury (SCI), Paralysis etc. It is typically caused by damage in the part of brain and/or spinal cord which is responsible for motor control. It is estimated that spasticity affects more than 12 million people around the world. About 80 percent of people with cerebral palsy (CP) and Multiple Sclerosis (MS) have spasticity (of varying degree). Since an estimated 500,000 people in the United States suffer with a form of CP, this means about 400,000 people suffer from some degree of spasticity. Similar statistics for MS show an estimated 400,000 people in the United States with MS and hence 320,000 people with some degree of MS-related spasticity [1]. The pharmaceutical industry spends billions of dollars developing drugs to relieve spasticity, but these efforts are stymied by the lack of repeatable, objective metrics to quantify the outcomes [2-4]; excessive dosage of drugs to treat spasticity can cause severe side effects such as such as seizures, blurred vision, and severe rashes, while inadequate dosage is ineffective at treating spasticity.

There are many methods to assess spasticity, the most commonly used metric being the Modified Ashworth Scale (MAS) [5-6]. The methods and metrics for spasticity assessment lack in being repeatable, consistent or objective [7-8]. This results in inaccurate prescription of treatment which is either inadequate or copious to the patients resulting in either no relief or seizures, fits etc. All the methods and their effect on assessing spasticity and its treatment are described in the next section.

The current assessment methods also require high level of medical expertise for the clinicians assessing. The most common tool, the MAS is so because of its simplicity and ease of use. However, at the same time, MAS is highly inconsistent as a metric and has high inter and intra rater variability.

The spasticity assessments are done weeks and months apart and during the said time, the condition of the patient can massively vary and the medication may not be appropriate. The account of the patient and their family members as to how the patients are feeling also affects the treatment.

Due to the above said reasons, there is a high need for a repeatable, objective and consistent metric that can be employed with ease of use across clinicians or raters of varied medical expertise is highly desirable. The research of this paper focuses on the same. For the aforementioned purpose, an instrumented glove has been developed which has an array of sensors mounted on it to sense force and motion of arm that give an objective rating of the extent of spasticity. A "mock patient" has been developed to achieve the replicability of MAS ratings to serve as the "ground truth" for validating the glove.

The paper is organized as follows: Section 4.3 describes the currently present metrics for spasticity assessment and prior research to address the lack of repeatable assessment metrics. Section 4.4 describes the development of the instrumented glove and the mock patient. Section 4.5 details the experimental protocol for data collection, description of clinical trials and algorithms that go into calculating the metric. Section 4.6 presents the results from the experimental data and the algorithms from Section 4.5. Section 4.7 is the conclusion and the future scope.

4.3 Prior Work

There are many methods to diagnose spasticity. There are clinical scales, which basically are based on a doctor's "feel" of the patients' stiffness. Therefore, these methods are very subjective. Clinical methods of assessment include:

1. Ashworth and Modified Ashworth Scale: MAS is the most widely used metric on account of its simplicity. MAS is a highly subjective rating [7, 9-10]. It has high inter- and intra-rater variability [11-12]. It has also been claimed that MAS does not consider the velocity aspect and only captures resistance to passive movement [13-14]. It does not distinguish between neural and non-neural causes of resistance [13]. Considerable research has been put into understanding spastic models, yet none address developing an objective metric.
2. Tardieu and Modified Tardieu Scales: In MTS, the angle for catch (using goniometers) at high velocity stretch and the angle for full passive range at slow velocity stretch responses are measured [15]. Thus, it considers the velocity aspect of spasticity. It is suggested as the more appropriate metric over MAS because of this [15]. The MTS performs better in case of intra and inter-rater reliability than MAS [16]. It's inter-rater reliability is still not very good [16] [17]. Even though it is closer to actual description of spasticity given by Lance [18], the MTS is still subjective in nature. This is proven by change in variability (both) before and after training of raters. It is less popular than MAS because MAS is simpler.
3. Hypertonia Assessment Tool [19]
4. Composite Spasticity Scale [20]
5. Gross Motor Function Classification System - Expanded & Revised (GMFCS - E&R) [21]
6. King's Hypertonicity Scale

Secondly, there are neuro-physiological assessment tools which are inclusive of the neurological aspect of spasticity. There methods don't always correlate to the actual level

of spasticity even though the measurement correlation is usually high. These methods also often rely on voluntary motion by patients which is an undesirable property in assessment as the patients may or may not move to their full extent and this might cause them inconvenience. Some neuro-physiological assessment tools are as follows:

1. Electromyography
2. Tonic stretch reflex testing
3. H-reflex

The neuro-physiological tests use some sensors to get measurements. There is some consistent disadvantage that all these tools display. They instrument the patient and they do not have a defined translation to extent of spasticity. All the neurophysiological tools and their variants are not commonly used since literature does not back these methods up with a direct correlation to level of spasticity [22-23]. None of these methods correlate to spasticity levels and merely give measurements of passive reflex threshold, velocities and stretch angles

The third type of assessment tools are biomechanical tools. These are machines or use some mechanical tools to assess spasticity. Some of these methods are:

1. Myotonometer
2. Wartenberg Pendulum Test
3. Three-dimensional pendulum test
4. Dynamometry
5. Measures using goniometry
6. Inertial sensors
7. Stiffness tool with robotic-assisted gait orthosis

These methods either get too bulky for the patient or reply on the voluntary motion of the patients which is not reliable [24]. Some studies also mention a not so significant correlation with clinical scales [25-27].

Many researchers have taken different approaches to address the lack of quantitative assessment of spasticity. Wearable devices [28-30] and EMG sensors [31] have been deployed on patients to detect spasticity symptoms, but the drawback is that such devices can be inconvenient and uncomfortable for the patient. Studies using electromyography (EMG) sensors [31, 32] were carried out on patients with spasticity to characterize the patients' muscle tones under flexion and extension. Wu et al. [33] measured the catch angle reliably by determining the instantaneous velocity and the time derivative of torque. Research by Park et al. [34] also targeted measurement of catch angle and elbow range of motion. Both the above studies were focused on identifying the presence/absence of a catch phase for correlation to a MAS score between 1 and 2, but these studies did not provide a continuous scale to quantify the different levels of severity. The lack of a quantitative scale for spasticity was addressed by development of musculoskeletal models [35] or haptic simulations [36] to determine key physical parameters that contribute to spasticity. One of the most common models is the Haptic Elbow Spasticity Simulator (HESS) [37-39], in which the properties of spasticity are simulated with the muscle resistance as torque and the catch phase as an impulse. Development of the HESS simulator mainly benefits the doctors as they can practice MAS assessments without requiring actual patients. Their research focused on modeling of spasticity and emphasized on the factors that characterized each MAS level. Alternatively, a mathematical model by Zakaria et al. [40] formulated the resistance as torque and

accounted for additional parameters such as the angular velocity, modulus of elasticity etc. The above models have yet to be translated into physical tests that can be implemented on patients to track the spectrum of spasticity conditions.

4.4 Experimental Setup

The experimental setup consists of two parts: a) the instrumented glove and the b) mock patient. The instrumented glove is intended to be worn by the raters/clinicians who assess the patients. The sensors on the glove would then give an estimate of the extent of spasticity. We have decided to instrument the raters instead of the patients for the following reasons:

1. It is more convenient for the patients to not wear instruments or sensors as seen from previous studies in section II
2. Considering the doctor-patient ratio, it makes more financial sense to instrument the doctors

The mock patient is a validating ground truth for the glove. This is used to simulate consistent conditions for the glove to test.

A. Instrumented Glove:

Our approach to improve spasticity assessment is an instrumented glove worn by the doctor during patient evaluation. We integrated a spatially-resolved, force dependent resistive sensor array (by Tekscan, [41]) and an inertial measurement unit (IMU) consisting an accelerometer, gyroscope and a magnetometer [42]. The force sensor on the glove measures the contact force being applied to move a patient's limb. The level of muscular resistance to motion indicates severity of spasticity. Figure 4.1 (right) shows the force sensor integrated on to a golf glove. It has 18 sensing regions, with a total of 349

sensing elements that output a voltage proportional to the applied force. The raw output is a spatial map of 8-bit values for each sensing element. The data was collected at 20Hz. For our analysis, we used the sum of the output of all the sensing elements. During the experiment, the researchers wore the glove and performed cycles of movement with the patient, such as elbow flexion and extension. The IMU is attached to the back of the glove as shown in Figure 4.1 (left). It is used to characterize the hand maneuvers during clinical assessment of spasticity. In this work, we use only the gyroscope data to estimate the power needed to manipulate a limb. The IMU data is collected at 20 Hz. The angular velocity v from gyroscope is converted to linear velocity at the location of the grip in the mock patient. We estimate the power to move the patient's limb as F*v. In our initial study, five individuals with cerebral palsy volunteered to participate in this study. Participants and/or their parents provided informed consent as per the UCSD Human Subjects Internal Review Board regulations. Participants engaged in a modified Ashworth scale assessment with two physicians well-trained in this methodology (AS and his colleague) and then again by the same two physicians while wearing the spasticity measurement device. These data were collected in UCSD's Research on Autism and Development Laboratory. In this experiment, there was substantial inter-rater variability resulting in only 27% agreement in MAS values. Consequently, we were not able to use these data to validate the estimates from the glove sensors. To mitigate this, we created a mock patient capable of generating criterion metric (ground truth) that can be used to validate the objective numbers estimated from the glove sensors.

Figure 4.1: Instrumented glove and IMU

B. Mock Patient:

The mock patient has an arm structure as shown in Fig. 4.2. The arm has a lever connected to a disc clamped by a 5”C-clamp with stationary-bike brake pads, such that the resistance can be changed manually. The arm has an embedded load cell (model HX711 [43]) that senses the dead weight m due to the resistance set by the clamp. We compute the force to overcome this resistance as F = m*a, where a is the differential of the velocity found by gyroscope data. We use the term “preset resistance on the mock patient” to denote the force required to move the arm. The units are Newtons. The mock patient also has a gyroscope [44] to sense the angular velocity v during flexion and extension. We use this to measure the power as F*v, in N-m/s. In our experiments, we measure the power from the mock patient sensors and use it compare with the power estimated from the sensors in the glove worn by the rater.

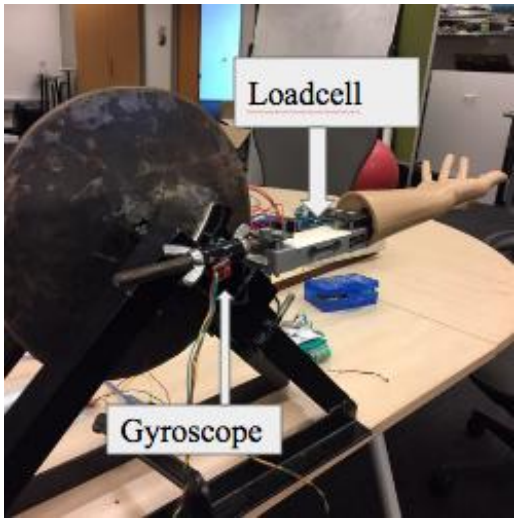

Figure 4.2: Mock patient with loadcell and gyroscope

4.5 Experiments and Algorithms

Two sets of data collection were done. One has 15 datasets from 10 non-clinician raters from a MAS range of 1+ to 3 (the MAS range settings we suggested by an expert - AS). The second data collection was done among 6 clinicians from various affiliations [47].

A. Experimental protocol

The raters should hold the mock patient arm parallel to the wrist with thumb on the top side of them arm. In that position, they should do multiple flexion and extension maneuvers for a 20 second duration. This counts as on trial. The raters do this for multiple weight settings. In this experiment, there are 6 weight settings at 3 pound increments from 5 to 20 pounds. All 6 trials count as one set. All the clinician and non-clinician raters did these sets for the purpose of this experiment.

For each of the trials, there are four data streams collected from 4 sensors: glove pressure sensors, loadcell, gyroscope on the glove and gyroscope on the mock patient.

B. Algorithm

We get 4 sets of data from the entire setup. Force data and gyroscope data from both the mock patient and the glove. The consistent metric, as mentioned above, is power. However, certain pre-processing steps need to be followed to obtain meaningful data information. The block diagram in Fig 4.3 explains the algorithm in use.

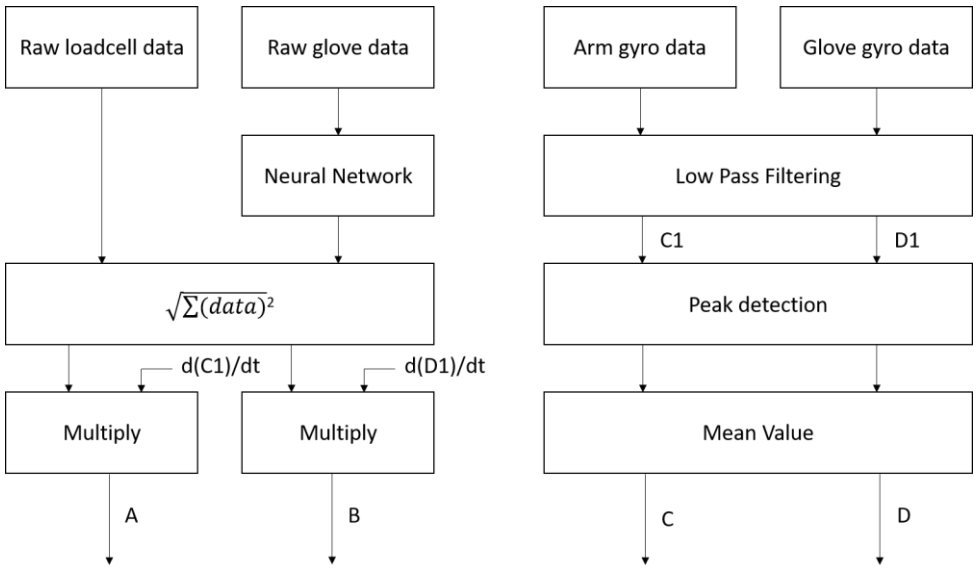

Figure 4.3: Block diagram of algorithm

The data from glove pressure sensors has the highest amount of error among the four sensors. A neural network has been employed to remodel this data without the error terms using data from loadcell. The description of neural network is mentioned in the next section. The square root of sum of squared of the data is found for both glove and loadcell data. This assess the frequency content of the signal. Alternate intuition is to find the square root of sum of data FFT squared. This intuition is also due to assessment of energy content of the data. Both these yield the same result owing to the Parseval's theorem. For the gyroscope data from both the mock patient and glove, we do FFT based

low pass filtering and find the peak values. The median of these peaks is considered as the value of speed in computing power. Thus, finally, the product A*C for glove and B*D for mock patient give the power expended in the maneuvers (since F=m*a and a=dv/dt). The analysis in [46] mentioned drift in signals as a major source of error. This algorithm aggregates the effect of drift and thus gives better result.

Figure 4.4 shows the glove force vs loadcell force measure data. Similarly, figure 4.5 shows the glove and mock patient gyroscope data waveforms. The actual force in Newtons is found by multiplying the glove (or loadcell) data with its corresponding acceleration found using gyroscope data.

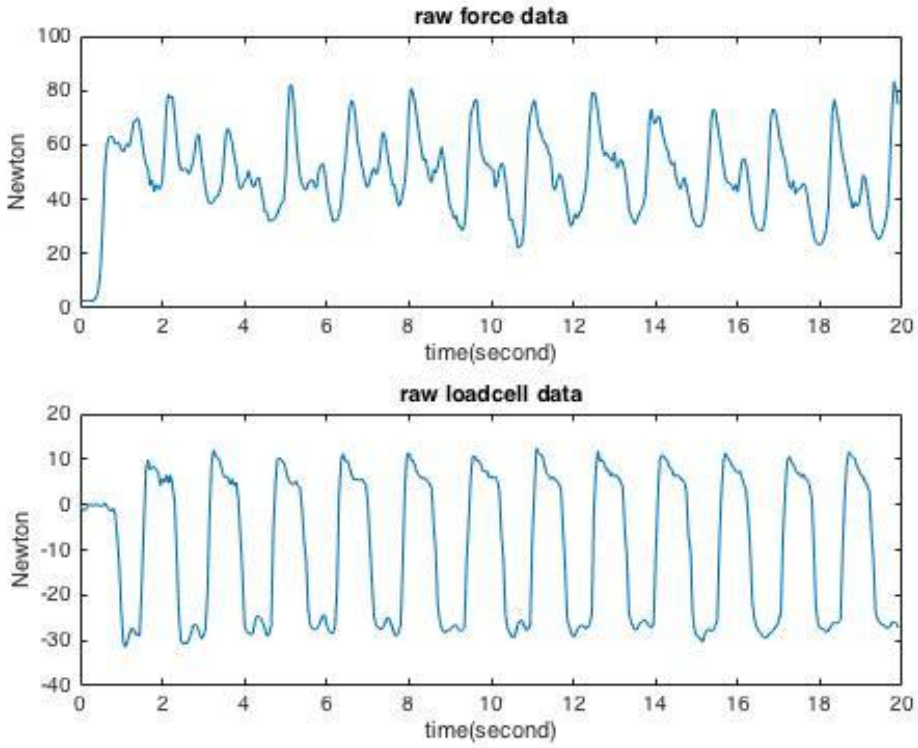

Figure 4.4: Glove and loadcell force data

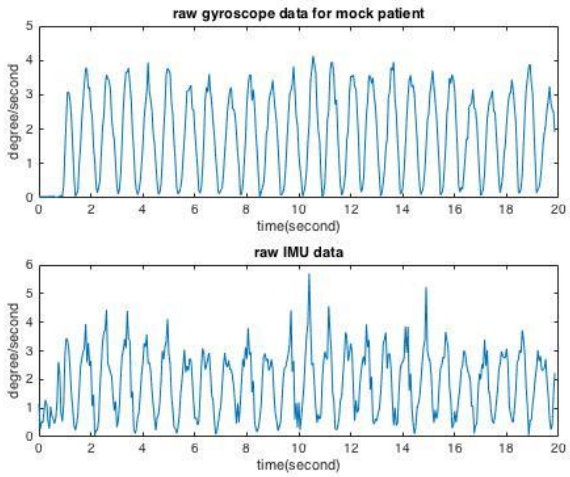

Figure 4.5: Mock patient gyroscope data (top) and glove gyroscope data (bottom)

C. The Neural Network

The instrumented glove contains a total of 349 sensing elements and records data at 20Hz. Thus, there are 349 dimensions for each sample. However, since each rater has different gripping and hand size, simply take the sum of the output of all the sensing elements will not match the loadcell reading in the mock patient. Even with same rate, there are disturbance can come from changing of gripping during the trial. Thus, it requires a robust approach to map the glove data to the loadcell reading. Since the dimension of the glove data is much larger than the dimension of the loadcell reading, the mapping can be solved using a neural network.

The neural network contains an input layer, one hidden layer, and one output layer, and there are 100 neurons in the hidden layer, and 1 neuron in the output layer. The activation function between each layer is the tanh function. Moreover, the network is trained using stochastic gradient descent with regularization factor. Since the neural network needs to perform regression, the loss function is 2-norm loss.

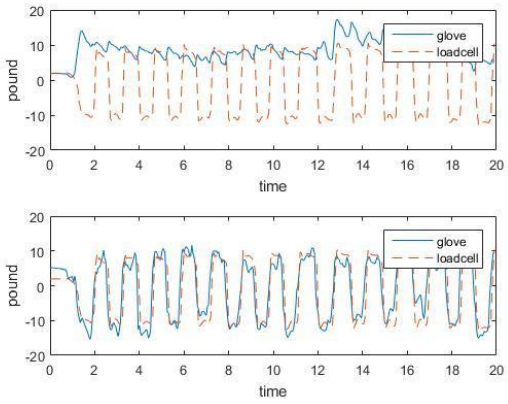

Figure 4.6: Glove vs loadcell before and after NN

During the training process for every rater in non-clinician and clinician datasets, one rater's data was put into the testing set and all of the other rater's data was put into training set. This training process can make sure that this approach can be generalized across different rater.

In the first plot of Figure 4.6, the glove data is simply generated by taking the sum of the output of the sensing elements. The glove data in the second plot is processed using the neural network. It is obvious that the glove data in the second plot is more correlated with the loadcell reading.

4.6 Results

For the data collected from clinicians and non-clinicians, power expended is calculated as explained in section IV.B. This section shows the results thus obtained.

For the non-clinician data, the correlation between glove and loadcell force (A vs B on Figure 3) is shown in Figure 4.7. The Pearson correlation coefficient obtained is 88%.

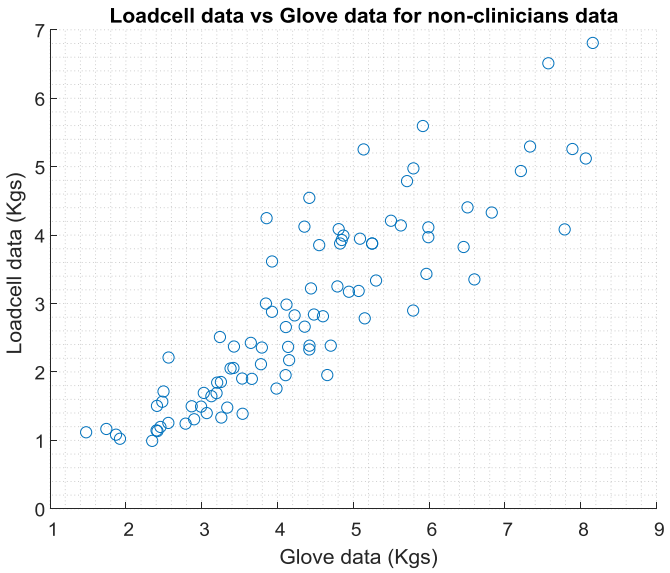

Figure 4.7: Loadcell data vs final glove data for non-clinicians' data. The correlation obtained is 88%

The correlation between mock patient and glove gyroscope data (C vs D on Figure 3) is give in Figure 4.8. The Pearson correlation coefficient obtained is 83%.

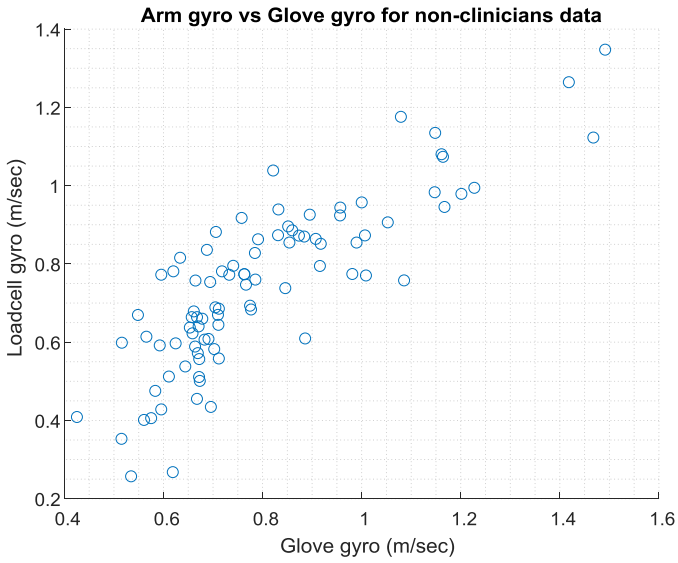

Figure 4.8: Arm gyro data vs glove gyro data for non-clinicians' data. The correlation obtained is 83%

The result (A*C vs B*D in Figure 3) for the non-clinicians' data is given below in Figure 4.9. The result and Pearson correlation coefficient obtained is 90%.

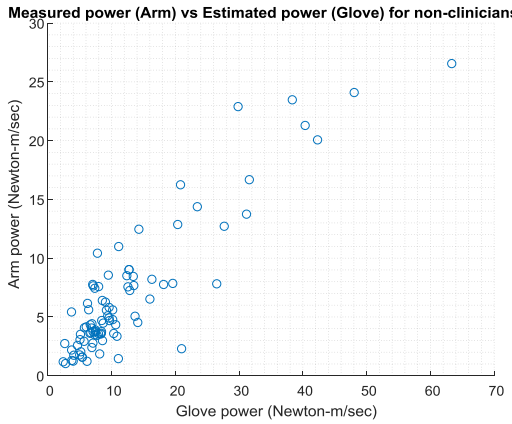

Figure 4.9: Final measure power (Arm) vs estimated power (Glove) in non-clinicians' data. The correlation coefficient obtained is 90%

To further evaluate how the algorithm performs across different raters, the variation of final correlation between mock patient (measured) power and glove (estimated) power across raters is shown in Figure 4.10.

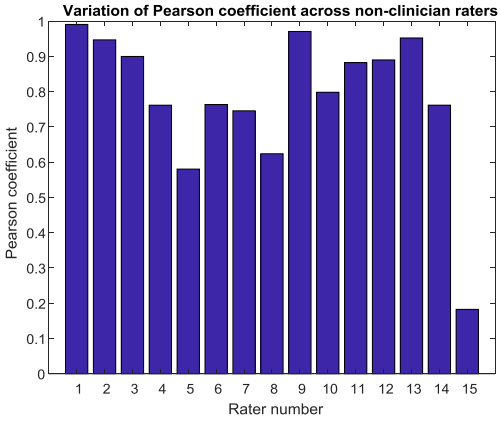

Figure 4.10: Variation of Pearson correlation coefficient (between final measured arm and estimated glove power) across non-clinicians

Similarly, to investigate how the algorithm performs with varying weight settings across all raters, the said correlation is plotted for different weights across all raters in Figure 4.11.

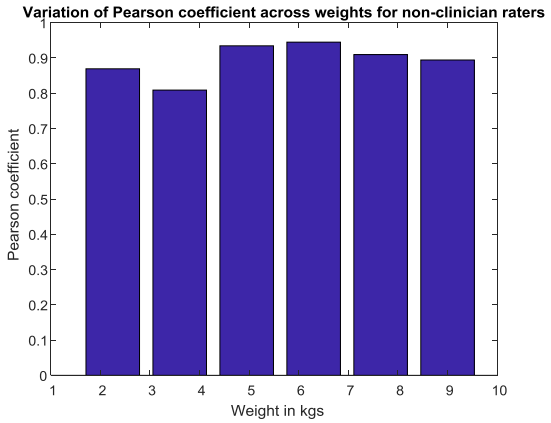

Figure 4.11: Variation of Pearson correlation coefficient (between final measured arm and estimated glove power) among all non-clinicians across different weight settings

The similar results are show in the following figures for the clinician datasets. It is noteworthy that with the non-clinician data, the experiment protocol was followed as instructed to the non-clinician raters. Thus, the results for non-clinician raters is under more controlled environment as compared to the clinicians' data where some bias was introduced due to highly varying grip (as compared to what was mentioned in section IV.A) and left-handed doctors using a right handed-glove.

The correlation between loadcell data and glove data is shown in Figure 4.12. The correlation coefficient is found to be 89%.

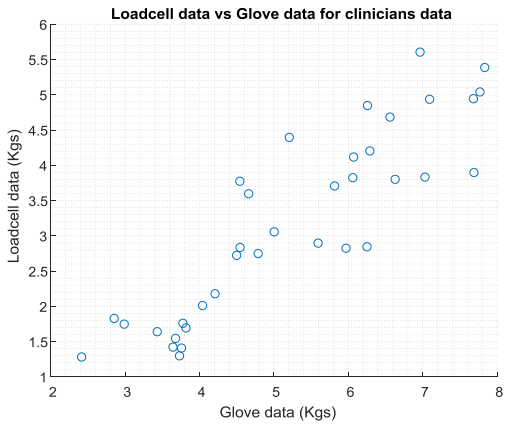

Figure 4.12: Loadcell data vs final glove data for clinicians' data. The correlation obtained is 89%

The correlation between mock patient gyro data and glove gyro data is shown in Figure 4.13. The correlation coefficient is found to be 71%.

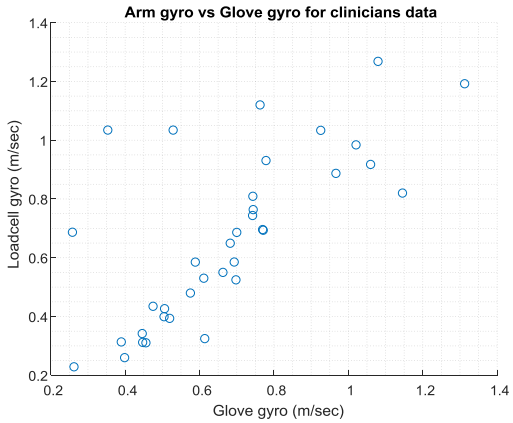

Figure 4.13: Arm gyro data vs glove gyro data for clinicians' data. The correlation obtained is 71%

The correlation between final measured power and estimated power is shown in Figure 4.14. The correlation coefficient is found to be 74%.

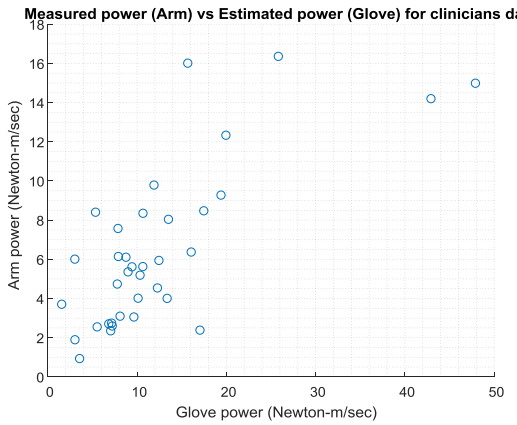

Figure 4.14: Final measure power (Arm) vs estimated power (Glove) in clinicians' data. The correlation coefficient obtained is 74%

The variations of correlation between measured and estimated power across different raters and different weights are shown in Figure 4.15 and Figure 4.16 respectively.

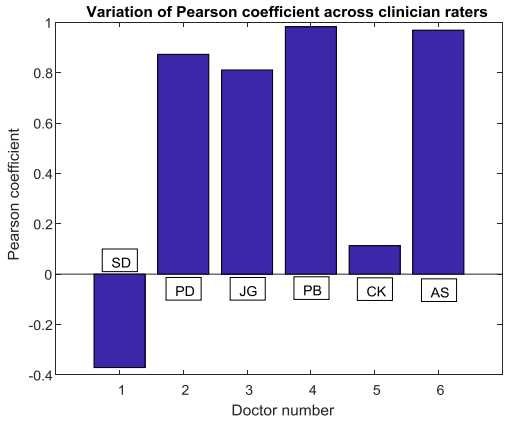

Figure 4.15: Variation of Pearson correlation coefficient (between final measured arm and estimated glove power) across clinicians in descending order

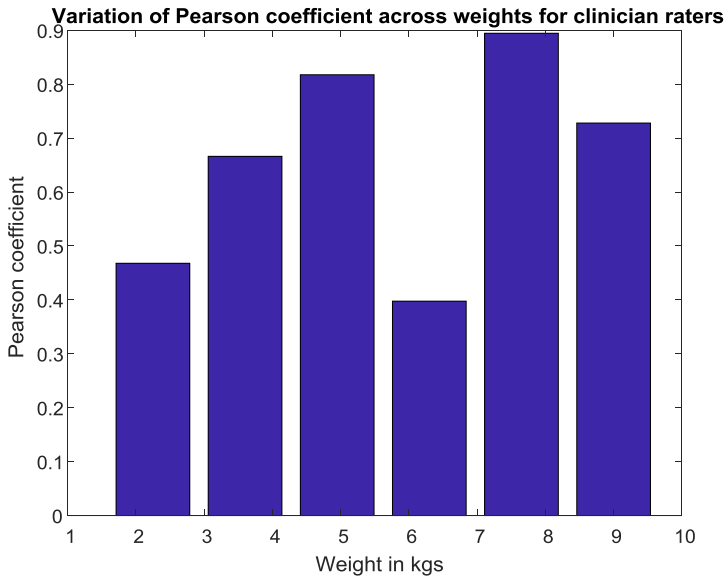

Figure 4.16: Variation of Pearson correlation coefficient (between final measured arm and estimated glove power) among all clinicians across different weight settings

For various weight settings, the MAS value as assigned by the clinicians for various weight settings is shown below.

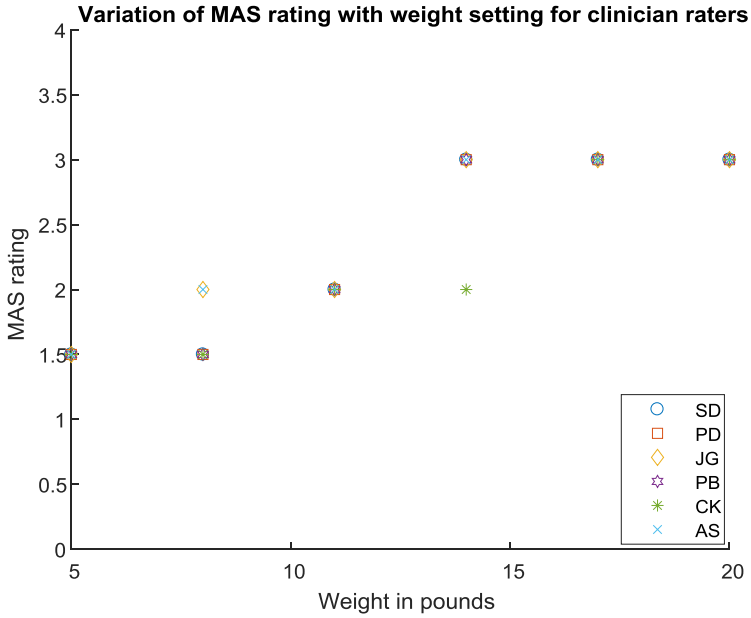

Figure 4.17: Variation of MAS rating by clinicians for varying weight settings

4.7 Conclusion and Future Scope

In conclusion, we see that the glove gives very reliable variation and correlates to the ground truth. In cases where the correlation falls, the raters are either left handed (PD, PB and JG) or had high grip variations. As can be seen in Figure 4.10, the correlation is very stable across various raters, thus showing very positive signs for mitigation of inter-rater variability which is a huge concern in the other subjective metrics. By comparing Figures 4.9, 4.14 and 4.17; we see that there is a definite correlation between the estimated power and the MAS rating. Thus, we conclude that this estimate shows positive signs of being consistent unlike clinical tools in Section II. It also shows that it can correlate to MAS unlike the neuro-physiological tools. By the consistency across weights in Figure 4.11, we can also conclude that it has the potential to be a repeatable metric. Thus, with some more improvement, this can be a repeatable, consistent and objective metric with a definitive mapping to standard spasticity measures. This glove needs to be only worn for assessment and thus does not require any clinical expertise on the rater's part.

For the future developments in this research, we aim to make the glove robust against grip variations. We also aim to improve the current mock patient to include high variety of spasticity profiles based on real patient data. As can be seen in Figure 4.17, the mock patient is repeatable for weight settings and thus can be used to train inexperienced clinicians in spasticity assessment. We are experimenting with the resolution of the glove sensors in order to print our own flexible force sensors instead of the COTS sensors which have been established to have considerable variance [45] (up to 34%). Even though current algorithms mitigate drift effects, to allows for higher flexibility with

sampling and processing, we would like to get all the sensors on a common clock. All these steps are essentially to improve sensor reliability and to mitigate grip issues.

Reference: refer to Chapter 6

Acknowledgements

Chapter 4, in full is being prepare for publication An Instrumented Glove for Augmenting Spasticity Assessment with Objective Metrics, Saisri Padmaja Jonnalagedda, Fei Deng.

Chapter 5: Conclusion and Future Work

In conclusion, the instrumental glove has a reliable correlation with the mock patient. The neural network gives a consistent mapping from the force sensor array to the loadcell. Even through output after the mapping has variance compare to the target, the correlation improved by increasing the number of samples in the experimental trial. With the improvement from the neural network, the mock patient validates the instrumental glove as an objective metric for improving assessment of spasticity.

For the future work, the mock patient can be improved to have a better simulation to the real patient. The current mock patient is using breaking pad to create friction as described in section 1.2.2. In the future, the mock patient will be constructed using a motor instead of the breaking pads. The resistance force can be generated using the motor. The resistance force provided by the breaking pads can decrease when the usage is increased, and only one resistance force can be setup during one trial. The motor can provide more consistent resistance force, since the usage will not change the performance of the motor too much. Moreover, since the motor is controlled by a computer program, it can provide different scale of resistance force during one trial to have a better simulation of the real patient. Thus, with the mock patient which is constructed with a motor, it can provide better validation of the objective metric.

Chapter 6: Reference

[1] https://www.aans.org/Patients/Neurosurgical-Conditions-and-Treatments/Spasticity

[2] Skalsky AJ, Vadivelu S, Dalal P Spasticity, In: Chapman's Orthopedic Surgery. 4th edition, M Chapman ed., Jaypee Medical Publishers, Philadelphia, PA USA (In press)

[3] Skalsky AJ, Fournier CM. Intrathecal Baclofen Bolus Dosing and Catheter Tip Placement in Pediatric Tone Management. Phys Med Rehabil Clin N Am. 2015 Feb; 26(1):89-93. doi: 10.1016/j.pmr.2014.09.011. Review. PubMed PMID: 25479782.

[4] Deon LL, Gaebler-Spira D. Assessment and treatment of movement disorders in children with cerebral palsy. Orthop Clin North Am. 2010 Oct;41(4):507-17.

[5] Ashworth B. Preliminary trial of carisoprodol in multiple sclerosis. Practitioner 1964;192:540–2

[6] Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. Phys Ther 1987;67:206–7

[7] Pandyan AD, Johnson GR, Price CI, et al. A review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity. Clin Rehabil1999;13:373–83.

[8] JFM Fleuren et al. Stop using the Ashworth Scale for the assessment of spasticity. J. Neurol Neurosurg Psychiatry 2010;81:46- 52 doi:10.1136/jnnp.2009.177071.

[9] Clopton N, Dutton J, Featherston T, Grigsby A, Mobley J, Melvin J. Interrater and intrarater reliability of the Modified Ashworth Scale in children with hypertonia. Pediatr Phys Ther 2005; 17: 268–274

[10] Ansari, N.N., Naghdi, S., Arab, T.K. and Jalaie, S., 2008. The interrater and intrarater reliability of the Modified Ashworth Scale in the assessment of muscle spasticity: limb and muscle group effect. *NeuroRehabilitation*, *23*(3), pp.231-237.

[11] Mehrholz J, Wagner K, Meissner D, et al. Reliability of the Modified Tardieu Scale and the Modified Ashworth Scale in adult patients with severe brain injury: a comparison study. Clin Rehabil. 2005;19:751-759

[12] Mutlu, Akmer, Ayse Livanelioglu, and Mintaze Kerem Gunel. "Reliability of Ashworth and Modified Ashworth scales in children with spastic cerebral palsy." *BMC Musculoskeletal Disorders* 9, no. 1 (2008): 44.

[13] Alhusaini AA, Dean CM, Crosbie J, Shepherd RB, Lewis J. Evaluation of spasticity in children with cerebral palsy using Ashworth and Tardieu Scales compared with laboratory measures. J Child Neurol 2010; 25: 1242–1247.

[14] Kumar RT, Pandyan AD, Sharma AK. Biomechanical measurement of post-stroke spasticity. Age Ageing. 2006;35:371-375

[15] Haugh AB, Pandyan AD, Johnson GR. A systematic review of the Tardieu Scale for the measurement of spasticity. Disabil Rehabil. 2006;28:899-907.

[16] Mehrholz J, Wagner K, Meissner D, et al. Reliability of the Modified Tardieu Scale and the Modified Ashworth Scale in adult patients with severe brain injury: a comparison study. Clin Rehabil. 2005;19:751-759

[17] Mackey, Anna H., Sharon E. Walt, Glenis Lobb, and N. Susan Stott. "Intraobserver reliability of the modified Tardieu scale in the upper limb of children with hemiplegia." Developmental Medicine & Child Neurology 46, no. 4 (2004): 267-272.

[18] Lance JW. Pathophysiology of spasticity and clinical experience with Baclofen. In Lance JW. Feldman RG, Young RR, Koella WP, editors. Spasticity: Disordered Motor Control. Year Book: Chicago; 1980; 185 – 204. (line 786 – 788)

[19] Jethwa A, Mink J, Macarthur C, Knights S, Fehlings T, Fehlings D. Development of the Hypertonia Assessment Tool (HAT): a discriminative tool for hypertonia in children. Dev Med Child Neurol 2010; 52: e83–e87.

[20] http://www.strokengine.ca/psycho/csi_psycho/

[21] https://www.canchild.ca/en/resources/42-gross-motor-function-classification-system-expanded-revised-gmfcs-e-r

[22] Poon, Dora MY, and CHRISTINA WY HUI-CHAN. "Hyperactive stretch reflexes, co-contraction, and muscle weakness in children with cerebral palsy." *Developmental Medicine & Child Neurology* 51, no. 2 (2009): 128-135.

[23] Jobin A, Levin MF. Regulation of stretch reflex threshold in elbow flexors in children with cerebral palsy: a new measure of spasticity. Dev Med Child Neurol 2000; 42: 531–540.

[24] http://journals.humankinetics.com/doi/abs/10.1123/jsr.2014-0271

[25] Li, Xiaoyan, Henry Shin, Sheng Li, and Ping Zhou. "Assessing muscle spasticity with Myotonometric and passive stretch measurements: validity of the Myotonometer." Scientific Reports 7 (2017).

[26] Rydahl, Sarah J., and Brenda J. Brouwer. "Ankle stiffness and tissue compliance in stroke survivors: a validation of Myotonometer measurements." *Archives of physical medicine and rehabilitation* 85, no. 10 (2004): 1631-1637

[27] Leonard, Charles T., Jason S. Brown, Timothy R. Price, Susan A. Queen, and Eugene L. Mikhailenok. "Comparison of surface electromyography and myotonometric measurements during voluntary isometric contractions." *Journal of Electromyography and Kinesiology* 14, no. 6 (2004): 709-714.

[28] Ishikawa, Shun, Shogo Okamoto, Kaoru Isogai, Yasuhiro Akiyama, Naomi Yanagihara, and Yoji Yamada. "Wearable dummy to simulate joint impairment: severity-based assessment of simulated spasticity of knee joint." In System Integration (SII), 2013 IEEE/SICE International Symposium on, pp. 300-305. IEEE, 2013.

[29] Ferreira, João, Vitor Moreira, José Machado, and Filomena Soares. "Improved biomedical device for spasticity quantification." In Bioengineering (ENBENG), 2013 IEEE 3rd Portuguese Meeting in, pp. 1-4. IEEE, 2013.

[30] Peng, Qiyu, Parag Shah, Ruud W. Selles, Deborah J. Gaebler-Spira, and Li-Qun Zhang. "Measurement of ankle spasticity in children with cerebral palsy using a manual spasticity evaluator." In Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE, vol. 2, pp. 4896-4899. IEEE, 2004

[31] Misgeld, Berno JE, Markus Lüken, Daniel Heitzmann, Sebastian I. Wolf, and Steffen Leonhardt. "Body-Sensor-Network-Based Spasticity Detection." IEEE journal of biomedical and health informatics 20, no. 3 (2016): 748-755.

[32] Sherwood, Arthur M., M. M. Priebe, and W. B. McKay. "Quantification of surface electromyographic recordings for assessment of spasticity." In Engineering in Medicine and Biology Society, 1996. Bridging Disciplines for Biomedicine. Proceedings of the 18th Annual International Conference of the IEEE, vol. 2, pp. 597-598. IEEE

[33] Wu, Yi-Ning, Hyung Soon Park, Yupeng Ren, Deborah Gaebler-Spira, Jia-Jin Chen, and Li-Qun Zhang. "Measurement of elbow spasticity in stroke patients using a manual spasticity evaluator." In Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE, pp. 3974-3977. IEEE, 2006.

[34] Park, Hyung-Soon, Qiyu Peng, and Li-Qun Zhang. "A portable telerehabilitation system for remote evaluations of impaired elbows in neurological disorders." IEEE Transactions on Neural Systems and Rehabilitation Engineering 16, no. 3 (2008): 245-254.

[35] Xin, Zhang, Norimasa Adachi, and Kazunori Hase. "A computed model for quantitative evaluation of spasticity in man." In Computational Engineering in Systems Applications, IMACS Multiconference on, vol. 2, pp. 1844-1847. IEEE, 2006.

[36] Grow, David I., Mengnan Wu, Michael J. Locastro, Sugandha K. Arora, Amy J. Bastian, and Allison M. Okamura. "Haptic simulation of elbow joint spasticity." In 2008 Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, pp. 475-476. IEEE, 2008.

[37] Park, Hyung-Soon, Jonghyun Kim, and Diane L. Damiano. "Development of a haptic elbow spasticity simulator (HESS) for improving accuracy and reliability of clinical assessment of spasticity." IEEE transactions on neural systems and rehabilitation engineering 20, no. 3 (2012): 361-370.

[38] Park, Hyung-Soon, Jonghyun Kim, and Diane L. Damiano. "Haptic recreation of elbow spasticity." In 2011 IEEE International Conference on Rehabilitation Robotics, pp. 1-6. IEEE, 2011.

[39] Kim, Jonghyun, Hyung-Soon Park, and Diane L. Damiano. "Accuracy and reliability of haptic spasticity assessment using HESS (haptic elbow spasticity simulator)." In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 8527-8530. IEEE, 2011.

[40] Zakaria, Noor Ayuni Che, Takashi Komeda, Cheng Yee Low, Fazah Akhtar Hanapiah, and Kaoru Inoue. "Spasticity mathematical modelling in compliance with modified Ashworth scale and modified Tardieu scales." In Control, Automation and Systems (ICCAS), 2015 15th International Conference on, pp. 1893-1897. IEEE, 2015.

[41] https://www.tekscan.com/products-solutions/systems/grip-system

[42] http://www.motionnode.com

[43] https://www.sparkfun.com/products/13230

[44] http://sensing.honeywell.com/honeywell-sensing-potentiometers-lineguide-007067-1-en.pdf Wettenschwiler, Patrick D., Rolf Stämpfli, Silvio Lorenzetti, Stephen J. Ferguson, René M. Rossi, and Simon Annaheim. "How reliable arepressure measurements with Tekscan sensors on the body surface of human subjects wearing load carriage systems?" International Journal of Industrial Ergonomics 49 (2015): 60-67

[45] Saisri Jonnalagedda, Augmenting Subjective Assessments with Objective Metrics for Spasticity Assessment