Imperial College of Science, Technology and Medicine
Hamlyn Centre

A 2D Based Quantitative Assessment Framework for Upper Limb Functionality

Liall Arafa

Supervised by Professor Guang-Zhong Yang, Dr. Fani Deligianni, Dr. Jindong Liu

A thesis submitted in partial fulfilment of the requirements for the degree of MRes in Medical Robotics and Image Guided Intervention and for the Diploma of Imperial College
September 2018

Abstract

The increase in the world elderly population paired with a global nursing shortage has motivated the research and development of quantitative assessment for rehabilitation. In spite of existing assessment technology, there is yet to be a method of quantitative assessment of upper limb functionality that has been universally adopted, with most clinicians using traditional and subjective means. In order to mitigate this lack of adoption, this thesis presents a 2D Based Quantitative Assessment Framework for Upper Limb Functionality aimed at providing effective quantitative assessment while exploiting the ease of use of 2D data collection. The framework is based on clinical standards, extracting functionality related features from 2D video of subjects completing movement exercises in accordance with an experimental set up. Those features are used to train support vector regression models to learn quantitative assessment functions which generate an assessment score given a set of features.

Word Count- 13,587

Acknowledgements

I would like to express my gratitude to:

- Professor Guang-Zhong Yang
- Dr.Fani Deligianni
- Dr.Jindong Liu
- Dr.Soma Banarjee and her Team at the Hyper Acute Stroke Ward at Charing Cross Hospital
- Mohammed AbdelAziz
- Daniel Freer
- Rejin Varghese

Dedication

To my Momma Lina Sawan, my Rock Grisilda Grizzly Duli and my Sensei Miso.

‘It’s too late to be ready.’

Master Dōgen

May the technology we create always be toward the enlightenment of humanity

Contents

Fugl-Meyer Analysis - **FMA**

Wolf Motor Function Test - **WMFT**

Action research Arm Test - **ARAT**

Quantitative Assessment Function- **QAF**

Support Vector Machine- **SVM**

Support Vector Regression- **SVR**

Sequence Minimal Optimisation- **SMO**

Chapter 1

Introduction

The quality of patient-centred care in global health care facilities is currently at threat due to a number of significant strains on them. The first strain is caused by the dramatic increase in the world elderly population, which is growing at a faster rate than all young age groups, [1] and is expected to more than double by 2050, rising from 962 million in 2017 compared to an estimated 2.1 billion estimated in 2050. The second strain is the global nursing shortage, where demands for nurses are significantly higher than the supply of qualified nurses. Thirdly, global health care facilities are striving to improve the safety and efficacy of treatments, while driving down costs. The use of assistive robots and novel technologies in health care settings is one of the most important applications of robots, and are being developed with the aim of solving these problems [27] , while improving upon existing patient care techniques in terms of efficacy, availability and cost. The implementation of technology based rehabilitation systems such as robotics [19], [25] and virtual/augmented reality solutions for home upper limb rehabilitation[20][5][6][41][9] and other wearables [37] for patients who have suffered stroke or other neurological conditions, has vast potential to benefit patients.The effects of stroke are a major cause of functional disability and often lead to a loss of independence of daily living and of important tasks, and post-stroke recovery using existing systems remains poor. Stroke usually leads to the loss of functionality in the arm and leg on either the left or right side of the body. Efforts of developing hands free solutions for stroke recovery are motivated by the poor recovery rate of upper limb functions post-stroke, the public health burden of the increasing numbers of patient population requiring rehabilitation, and the personalised and long-term treatment that assistive technologies can provide. Much research is being developed into the development of such systems, and early clinical trials show promising results for upper limb recovery when compared to conventional therapies [40], [18], [41].

1.1 Existing Upper Limb Rehabilitation Technology

The most popular stroke technology in terms of research and adoption by clinicians is rehabilitation technology, which aids the patient in some form with exercises that restore lost motor functionality.An overview of the features, strengths and weaknesses of the rehabilitation landscape can be used in order to guide the development of a novel, effective system.

1.1.1 Contact Based Systems

Contact-focused assistive robotic (AR) rehabilitation systems are devices which use physical interaction as the primary means to perform rehabilitation. Some contact-focused AR devices are already utilised within clinical practice as well as clinical evaluation[34][4]. There are various physical approaches to restore the functionality of the upper extremity, such as orthoses, functional electrical stimulation, and physical therapy. AR devices can be categorised by their mechanical designs and the assistance that they provide. In terms of mechanical designs, devices are usually designed in end-effector based or exoskeleton based setups. Contact-based AR rehabilitation can provide feedback on the patients progress in a variety of ways including visual, tactile, via audio, and in the form of electrical stimulation. These are used to guide the patient during exercise to adhere to a pre-set trajectory or movement. Although the utilisation of contact-based AR has potential benefits, the outcome of contact-based AR devices in terms of upper limb recovery in clinical practice is not as positive as expected [38] [31], and the issue of maximising the number of repetitions while maximising the patients attention and effort still remains. In addition, many of contact-based AR systems are not easily portable, requiring a complex setup and large area for operation. In terms of cost-effectiveness, the results of the first large randomised multicentre in which training with the MIT-Manus AR system was compared with the intensive therapist-provided therapy showed that the cost of both treatments were similar [23]. However as development in robotic technology progresses, cost of equipment may decrease.

1.1.2 Contact-Free Systems and Pose Estimation

Most contact free tracking systems for the human body solve the problem of pose estimation. In computer vision and robotics, a common task is to identify a specific object in an image and calculate its position and orientation relative to a coordinate system. This information can be utilised in a variety of ways, including robot navigation and object classification. The combination of position and orientation of an object is considered its pose. Human pose estimation has many uses including in animation, interactive art, augmented reality,video games and fitness uses. In the medical field, human body pose estimation provides a wealth of information, providing physicians help in the diagnosis, treatment planning and progress evaluation. For example, motion analysis is used by orthopedists to assess the status of diseases affecting the skeletal system, and neurologists analyze the movements of epilepsy patients to identify the cortical locations that cause epileptic seizures.[35] Pose estimation can be used for quantifying patient movement by identifying the skeleton of a patient frame by frame, and identifying the coordinates of the key points that compose that skeleton.

Rehabilitation systems that utilise contact-free technologies have some significant advantages over contact-based AR, such as maximising patient engagement, low costs, and small or ubiquitous devices. In addition, contact-free systems can generally be placed in the home or hospital more easily than some large contact-based assistive robotic systems, with home- based solutions allowing patients to continue their exercise without the need of a physiotherapist. This takes some pressure off the public health care system as a whole. The need for a robotic peripheral is removed, potentially further driving down costs and the complexity of setting up such a device. Since the rehabilitation of the patient is not physically assisted by any robotic device, contact-free systems must be able to accurately identify, track the position and orientation of the patients upper limbs within an acceptable range, and post-process data in order to provide accurate feedback to a patient and ensure that their rehabilitation treatment is being adhered to. Contact-free rehabilitation systems often combine sensing technology with interactive gaming or virtual/augmented reality environments[30], which has the potential to enhance the patient engagement and motivation [33], while also allowing for proper visualisation of the routine they are currently doing versus the one prescribed to them. Motion tracking for upper limbs can be implemented in a variety of ways.

1.1.3 Non-Visual Tracking Systems

Sensors can be utilised within non-visual tracking systems in order to collect movement information.[37] They are commonly categorised as mechanical, inertial, radio, or microwave and magnetic based [45]. These sensors can be wearable or be embedded in some hand-held peripheral or remote. Some systems use a combination of sensors. One significant benefit of non-vision based tracking systems is that they do not need a constant line of sight from sensor to detector in order to track, thus can be more robust in terms of position when based in a variable home environment. Inertial sensors such as accelerometers and gyroscopes can be used to implement non-vision based tracking for an easy to use and cost-efficient way for full-body human motion detection. Inertial sensors have already been implemented in a number of monitoring applications[24] [46][44], and accelerometers and IMU sensors are the most frequently used technology within feedback sensors from recent wearable upper limb rehabilitation systems [11]. Some systems use the built-in accelerometers within a smartphone such as a system developed by Ferreira et al [14]. The motion captured by the inertial sensors can be of high sensitivity and large capture areas. The motion capture data is transmitted wirelessly to a machine for further processing or visualisation. However, the position and angle of an inertial sensor cannot be correctly determined, due to an effect known as integration drift, and designing drift-free inertial systems is an area of current research. Magnetic motion tracking systems are widely utilised in virtual reality applications as they have numerous benefits such as small size, high sampling rate, and lack of occlusion from non-metallic objects placed between the sensor and receiver. In terms of upper limb rehabilitation, wrist and hand monitoring systems [15] and serious virtual reality games [5] have been paired with magnetic sensors. However, magnetic sensors inherent weaknesses such as latency and jitter which reduces accuracy of tracking, where the reduction of both is an active research area. Other non-vision systems that use wearable sensors include a project that was part of the myHeart initiative that utilised elastomer strain gauge sensors which were to be embedded in garments to be worn by the patient[16][17]. The garment was especially realised for measuring a specific set of upper-limb rehabilitation exercises, with the sensing element placed on the arm, forearm, shoulder and chest. Although, wearable sensors have the advantage of avoiding the line of sight requirement, they also require a patient to set up sensors in the correct position and place in order for accurate motion tracking assessment. Considering that patients will have a disability, this may prove difficult or impossible. In addition, the extra work of placing sensors multiple times a day in order for assessment may contribute to a lack of motivation to exercise, and result in patients not adhering to their prescribed regime.

1.2 Visual Tracking Systems

The implementation of computer vision techniques in tracking systems is common, and has a its own set of advantages and challenges. Since the limbs are tracked visually, the line of sight between the patient and camera must remain clear. Visual based tracking systems are commonly integrated with virtual/augmented reality video games, in an effort to provide implicit rehabilitation to patients while providing an appealing and engaging task. Studies show that the training using video games had a positive effect on motor function and patent motivation, and may be an effective intervention for home-based rehabilitation[41], [6], [9], [6], [20]. Visual based tracking systems can either use mounted markers which are detected or natural features within an image for marker free tracking.

Marker based tracking systems can be passive, active or hybrid in style- a passive system uses a number of markers that do not generate any light, whereas active markers produce light to be detected. Optical marker systems are among the most accurate motion capture technologies, and are used extensively in VFX and the film industry, such as the Vicon motion capture systems. They are also used in medical science. The Vicon system has been used to calculate joint centres and segment orientations by optimising skeletal parameters for gait analysis [13]. Passive optical markers consist of retro-reflective material which reflects incoming IR light back into the light source. Passive optical markers can include infrared reflecting markers. Active optical markers are infrared light emitting diodes (LEDs), and require a wire or battery to operate and emit IR light directly. Optical marker tracking systems are usually large, expensive, and require extensive set-up and calibration. Therefore, they are not ideal for hospital or home environments. More practical markers for vision based home rehabilitation include patterned markers such as ArUco or Hiro markers, commonly used in augmented reality applications. These are cheap to produce and only requires a video camera and the appropriate software to detect. This is commonly detected with a webcam or smartphone. AR markers have unique patterns embedded within their border, which can be used to identify them. Research done by Burke et al [6] uses AR markers attached to an object that the patient will interact with as part of various augmented reality games which aim to implicitly allow for upper limb exercises . The position and orientation of the marker object is tracked, and the system can then augment the captured image of the real environment with computer generated graphics to present a variety of game or task-drive scenarios to the user. However, the actual position and orientation of the patients upper limbs is unknown, and the only patient movement that can be assumed is the movement of the marker. This limits the types of exercises that can be done and the level of feedback detail which can be given to the patient. Although marker based tracking is often computationally simpler and can be more robust to accuracy errors than marker-less tracking, systems that use markers have the same time-consuming and potentially difficult task of the patient having to place markers, as evident with other wearable sensors.

Marker-less tracking systems uses natural features within an image and depth information in order to track upper limb movement, removing the need for a physical marker to be placed within the scene. Thus, they are found to be more appealing, comfortable, enjoyable and more intuitive to use. The most frequently used system in marker less tracking systems for upper limb rehabilitation is the Microsoft Kinect [18].

The Kinect is an infrared motion capture pose estimation device used for interactive computer games, and enables users to control and interact with the virtual reality environment via the RGB camera and depth sensor without need for a remote controller, thus enabling exercise at home without the assistance of a therapist. The Kinect V2 is optimised for 0.5m -4.5m which is able to capture the full body action with all patient joints. Vision-based body tracking systems rely on a depth-sensing camera which records a 3D position rather than a colour. In the Kinect, the depth sensor consists of an infrared projector and a CMOS infrared video camera, which is mounted on the front of the device. The infrared projector emits a speckle pattern of infrared dots in the cameras view. The depth sensor technology enables the Kinect to create a 3D map and a randomised decision forest consisting of three trees, each trained by 300, 000 images, and the Kinect is able to specify body joints positions in 30 frames per second. The skeletal frame of a person is inferred from a series of depth images over time. A statistical inference algorithm estimates key points in the body, such as the head and the hand, and connects them within a skeleton [38]. The skeletal inference relies on machine leaning techniques that fit the depth image sequence that was captured to a set of pre-computed examples in which the depth image sequence and skeleton position are known. The example data set is key to the accuracy of the inference. In the Kinect, the algorithm is based on 100, 000 examples of young, healthy adults, instead of older people or those with abnormal movement patterns. Given the training set, this may mean that the joints of upper limb rehabilitation will not be as accurately determined as with healthy adults, but whether there is a significant disparity has yet to be studied. In addition to the depth camera, the Kinect has an RGB camera, which provides colour data for processing and analysis [26]. Since the Kinect was primarily built for the Xbox gaming systems, it is commonly integrated with ease into serious upper limb rehabilitation video games in addition to un-augmented monitoring and feedback applications. The main advantages of the Kinect include that :

- It is easily portable and can easily be integrated into the hospital or home
- The accuracy of the Kinect is strong [41]
- It is commonly integrated with engaging tasks such as serious games.

The main disadvantages of the Kinect for use in general upper limb rehabilitation systems are:

- The majority of evaluation studies on the Kinect are focused on sets of gross movements that are advantageous for the Kinect, instead of evaluations based on more realistic and/or specific diagnostic movements, so evaluation on its use for upper limb rehabilitation specifically remains limited
- Kinect is unable to accurately assess internal joint rotations of the shoulder and uses a less clinically viable single point estimation, thus the Kinect has yet to be shown to be clinically viable for identification of the shoulder
- Fine motor motions are not captured by the Kinect alone and thus rehabilitation goals that include fine motor movements cannot be captured by the Kinect alone
- Kinect requires a TV screen for the patient to view, a computer or Xbox console and the Kinect peripheral itself. Thus, the Kinect cannot process data that is pre-recorded, and pose estimation is done live. This means that pose estimation can only be done by centres that own the Kinect device and the rest of its setup, creating a large restriction on the number of clinician's able to utilise the technology.

1.2.1 2D Based Rehabilitation Technology

The need for peripherals such as the Kinect or more complex multiple camera setups serve as a barrier to wide-spread adoption and utilisation of vision-based rehabilitation technology. The relative low cost and ease of use of visual-based marker free systems can be expanded upon with the use of technology that only relies on 2D input from a single standard RGB camera. This would further decrease the barrier to being practically adopted in the public health care system, and overcome resistance by clinicians. Given that in the current age, the use of smartphones and tablets with 2D video recording capabilities is extremely common, the potential amount of data that can be captured and used in order for rehabilitation uses is a figurative goldmine, yet technology that can exploit it has yet to be developed and adopted.

The OpenPose library is capable of the real-time multi-person keypoint detection and multi-threading with state-of-the-art accuracy, with a run-time that is invariant to the number of people in the image. It does not require a depth camera for pose estimation. Instead, it relies entirely on 2D input video. OpenPose represents the first real-time system to jointly detect human body, face and hand key points on single images. The library is built upon a convolutional neural network (CNN) which allows for multi-Person 15 or 18 keypoint body pose estimation and rendering. An initial version that allows for multi-person 2x21 keypoint hand estimation and rendering, although running time linearly depends on number of people in the image. The overall pipeline of the system is:

1. The method takes the entire image as the input for a two-branch CNN to jointly predict confidence maps for body part detection and part affinity fields for parts associations.
2. The parsing step performs a set of bipartite matchings to associate body parts candidates.
3. The full body poses for all the people in the image is then assembled.

Although the ease of 2D rehabilitation data collection has enormous potential for overcoming the adoption barrier compared to traditional 3D methods, there are drawbacks. As OpenPose is based on 2D data, the lack of the Z axis means that an entire dimension of pose motion is lost, with the distance from the camera and occluding movements significantly reducing the amount of data that is possible to analyse. Thus, using a 3D system such as the Kinect yields in more potential data that can be exploited and a more true to the real-world subject movement.In order to circumvent the limitations of 2D data, subject's must be framed in such a way that limits data loss and emphasises the desired movement. In addition, the suitability of Kinect for upper limb rehabilitation has been shown [41], whereas the suitability of OpenPose has yet to be studied. However, given that research shows that general 2D video input is viable for accurate tracking and assessment of the upper limbs, [11][10] , it is reasonable to assume that a sufficiently-well developed 2D system using OpenPose would be able to minimise some of the drawbacks of losing the 3rd dimension, while still retaining the important advantage of requiring very little from clinician's in terms of data collection.

1.3 Thesis Motivation: Quantifying Assessment

1.3.1 Observations At the Stroke Unit and Clinical Motivation

In order to gain insight into the reality of how rehabilitation technology was practically applied to patients by staff in a public hospital setting, I shadowed Dr.Soma Banarjee and her team of doctors and physiotherapists for a number of days spread over the course of a few months at the Stroke Unit at Charing Cross Hospital . I followed doctors as they completed rounds of the ward, interacting and assessing patients who had suffered strokes of varying severity. For patients that were able to talk, typical encounters between doctors and patients went as follows- once the doctor and patient had small talk and the patient was comfortable, the doctor would ask the patient about the symptoms they were suffering from. This included on what parts of the body did the stroke affect and to what extent.Then the patient would be asked to complete a number of exercises in order for the doctor to assess the severity of the stroke. For example, a patient who had lost functionality in her left foot was asked to first raise her right non-paretic foot and hold it in the air for a number of seconds, before attempting the same with the paretic left foot. Similarly, a patient who had lost functionality in the right arm was asked to raise the non-parietic arm and then attempt the exercise with the paretic arm. A patient's progress was based on their ability to perform a functional task with their paretic limb as well as being able to perform it with their non-paretic limb. By comparing the difference of ability in completing the exercise between the paretic and non-paretic limbs, the doctors were able to identify what level of severity the patient had suffered from. I also shadowed a number of physiotherapists as they spent time completing rehabilitation tasks with patients. Physiotherapists would spend around 20-30 minutes with a patient, and complete a variety of exercises to restore functions in the affected limbs. For patient's which suffered from upper limb effects, these included the patient sitting in front of a mirror and completing active exercises assisted by the physiotherapist range which included the motions of the flexion/extension at the gleno/humeral joint and Scapula protaction/retractions. Exercises were tailored to the patient's level of comprehension and disability in order to maximise recovery efficiency.

A patient's rehabilitation progress and goals were all in context of the functional tasks that they were able to complete. Thus, there was a lack of quantitative data relating to their condition or progress for both staff and patient. Functional tasks were exercises that a person will almost definitely encounter in day to living, such as reaching and grabbing for a cup, reaching and lifting objects and placing them on an elevated box, and placing their arms on a table.Each patient had a unique set of tasks that they should work towards in the short term (next few weeks) and in the long term (next few months). For example, a patient that had lost limb functionality in the right arm and had issues with self-balance, was given the short term goal of being able to sit up and support themselves without the use of his arms, and given the long term goal of being able to independently walk around outside near the hospital's entrance.
When my time at the hospital was concluding, I interviewed a small number of physiotherapists on how they used rehabilitation technology to aid them in rehabilitation.I asked them on how they used the rehabilitation technology (such as the cheap rehabilitation robotics paired with serious games) that the ward had. The physiotherapists agreed that the technology was barely utilised as it was time-consuming to teach patients how to use them (if they were able to at all), and given the time-limited and pressurised environment of the hospital they were seen as not worth the benefit. Thus, traditional means were heavily relied on. I also asked about how frequently traditional upper limb functionality measures such as Fugl-Meyer Analysis (FMA), Action Research Arm Test (ARAT) and the Wolf Motor Function Test (WMFT) were used. The physiotherapists agreed that although useful, these kind of measures were unable to be conducted regularly in the context of the time-sensitive, pressurised environment of the hospital. This then prompted me to ask about whether technology that could give quick, easy to understand, quantitative data on a patient's exercise and progress would be useful. They agreed that this would be useful, yet they were not aware of such a technology being adopted or developed.

1.3.2 Quantification of Upper Limb Functionality

Standard assessments for upper limb functionality in which a doctor or physiotherapist observes a patient carry out some exercises or functional task with both paretic and non-paretic limbs are still the most widely used in practice. In medical standard practice tests such as the FMA , WMFT, and ARAT, patients are tasked with a variety of functional exercises based on every tasks, including reaching, grasping, gripping, pinching. Some measures such as FMA also test for the sensitivity of the paretic limb in response to touch. By comparing the function of the non-paretic limb with the paretic limb, the physiotherapist can give a score depending on the test chosen. Although these methods are medical standard there remains major issues:

1. The clinical evaluation is time-costly and requires physical involvement with the physiotherapist potentially leading to the evaluation not being completed successfully
2. The assessment is subjective and may differ from physiotherapist to physiotherapist
3. There is no quantitative data on a patient in terms of how their limbs move at the nose, neck, shoulders, elbows and wrists, meaning that general healthy models for a particular exercise cannot be made
4. There is no quantitative data on tracking of a patient's progress.

While there is extensive research in the sector of rehabilitation, there remains much less research on technology that is able to quantify patients limb functionality, effectively, cheaply and easily. This was evident in both a literature review and time spent observing at a stroke unit. By utilising technology that can efficiently, accurately and cheaply quantify patient upper limb functionality, many of the problems that arise with the use of standard upper limb functionality tests can be avoided. Quantitative assessment allows for clinician's to more objectively and accurately assess a patient's functionality.

Research shows that the increased use of kinematics in the clinics is desirable, but it will require a simpler method and analysis model, which needs to be reliable, easy to use, allow manageable data analysis, provide clear results and meet the clinical questions of concern.[3]The development of quantification systems would solve the major issues of subjectivity of the physiotherapist/doctor and the lack of quantitative data on a patient's progress in context of the functional tasks they are performing. In recent years there has been research towards creating both quantitative and autonomous analysis of post-stroke upper limb function by pairing data captured from computer vision, IMUs and other sensors with machine learning algorithms. Related research can be compared in the tables below:

Paper	**A Computational Framework for Quantitative Evaluation of Movement during Rehabilitation** [12]
Key Contributions	1. Aims to create a framework for a novel, more finely tuned quantitative clinical assessment scale where 0 indicates unimpaired and 1 indicates a severe impairment as opposed to traditional clinical observation. 2. Mathematically defines the problem of quantitative evaluation of movement 3. Present the framework of learning the evaluation function from clinical observations. 4. Apply the evaluation framework on a specific reach and grasp task 5. Show the experimental results. Categorises evaluation metrics by Temporal Profile, Targeting, Trajectory and Velocity Profile
Strengths	Validated using stroke patients and an quantitative assessment scale based on the Wolf Motor Function(WMTT) Test. Mathematically defines the evaluation function by finding the optimal weights and normalisation functions such that the difference between the computational evaluations of the movement sand clinical observation is minimised. Has good results on applying the framework for the reach and grasp task Larger number of patients being experimented on
Weaknesses	Relies on Optotrak, unwieldy and expensive for the hospital/home. Searching for the optimal solution is difficult because the number of potential normalisation factor is large, the movement contains a high number of dimensions (kinematic variables), The framework is only tested and validated with a single reach and grab movement exercise instead of a range of movement exercises.
Extra Notes	The therapists in this paper compare two movement videos and determines which movement they believe to have a greater impairment

Paper	**Component-level tuning of kinematic features from composite therapist impressions of movement quality** [39]
Key Contributions	1. Using Chen et al's framework [12] Proposes component-level kinematic features for movement quality assessment of wrist movement 2. Proposes a generic framework for tuning the thresholds and weights associated with each of these kinematics features using movement quality labels provided by therapists The Kinematic features include : a) Trajectory error b) Velocity Profile deviation c) Jerkiness d) Segmentation
Strengths	Mathematically defines the kinematic features that define movement quality well Use optimisation theory in order to estimate the threshold values for each kinematic feature The choice of kinematic features are based on clinical studies which relate multi-joint characteristics to their kinematic characteristics
Weaknesses	Uses a single marker on the wrist, meaning that shoulder data was not captured. This possibly caused the weak correlation between segmentation and therapist rating, as the segmentation feature requires both wrist and shoulder data Uses linear model which only looks at the mean of the dependant variable, sensitive to outliers, and relies on data being independent Complex movements (e.g lift and transport an object) are not accounted for.
Extra Notes	Used 10 Participants for data collection, with videos of different tasks recorded and paired with an assessment score from a therapist.

Paper	**Assessing Upper Extremity Motor Function in Practice of Virtual Activities of Daily Living** [2]
Key Contributions	Describes the design of a low-cost virtual world-based system for practice of meaningful activities that incorporate specific functional movements, and investigates the concurrent validity of motor performance metrics generated during virtual Assistive Daily Living practices. Paper focuses on a hypothesised correlation between the Virtual Occupational Therapy Assistant and the accepted clinical measure (the WMFT). Motor function related metrics considered in this study include: a) Duration, sub-task completion time in seconds b) Normalised speed (NS) percent- mean speed achieved divided by peak speed during performance of each task. Describes the design of a low-cost virtual world-based system for practice of meaningful activities that incorporate specific functional movements, and investigates the concurrent validity of motor performance metrics generated during virtual Assistive Daily Living practices. Paper focuses on a hypothesized correlation between the Virtual Occupational Therapy Assistant and the accepted clinical measure (the WMFT). Motor function related metrics considered in this study include: a) Duration, sub-task completion time in seconds b) Normalised speed (NS) percent- mean speed achieved divided by peak speed during performance of each task.
Strengths	Upper limb joint angles, rates and accelerations are clearly defined as a series of Euler angle rotations applied sequentially about Axes 1-4 in the shoulder, employing dot notation to indicate differentiation with respect to time. Splits main tasks into a variety of subtasks and scores each subtask in comparison to different metrics, giving an overall more accurate performance evaluation for each main task.
Weaknesses	Although reasoning is justified in this paper, the motor function metrics chosen are limited only the two highest metrics correlated with UE evaluation (according to one study) are chosen. These are normalised speed and movement arrest period ratio. This means the quality of the evaluation is completely based on these two metrics since the assessment of motor function takes into many different parameters.
Extra Notes	Uses a Kinect sensor, kinematic pose estimation algorithms, and game engine technology Kinematic Pose estimation algorithm are based on an adaptation of an unscented Kalman filter, the UKF based solution to the inverse kinematic problem produces angles and angular rates defining shoulder and arm motion The tracking filter enforces realistic arm kinematics and joint angle constraints, and handles noisy measurements and sensor dropout.

Paper	**Automated analysis and quantification of human mobility using a depth sensor**[22]
Key Contributions	Proposes non-invasive framework to recognise, assess, and quantify the mobility of participants. The system acquires motion capture (MoCap) from the single Kinect Sensor, and decompose it into a set of novel joint-group features. Analysis techniques are then employed to provide joint feedback highlighting the state of mobility. To identify and label the MoCap data, a method for automating the ground truth labelling of MoCap that is free from human bias or interpretation. Composed of three main parts: A) Feature Encoding using metrics B) Motion identification using machine learning classifiers C) Motion Analysis using multiple standard vector models In order to describe postural change over time, takes into account Euler Angles, Euclidean Distance, Body Lean Angle, Centre of Mass
Strengths	Uses machine learning in order to gain more robust classification of mobility. The framework is divided into two aspects: first, offline training of multiple machine learning classifiers based on an exemplar-based pose selection and second, online detection and identification of motions in real time. List Multiple machine learning classifiers utilised, including finetuning/parameter selection methodology including: a) SVM b) Random Forests c) ANN d) GRBM e) Adaptive Boosting f) LPoost g) RUSBoost h) Total Boost i) Bagging
Weaknesses	No clinical validation.
Extra Notes	Uses Kinect Assesses mobility for entire skeleton (not just upper limb).

Paper	**Kinect based system and artificial neural networks classifiers for physiotherapy assessment** [8]
Key Contributions	Proposes a system to unite the Kinect 360 and Artificial Neural Network (ANN) to aid in quantifying the patients performance ,by detecting human poses and gestures. The authors design a feature vector based on joint groups. The first group is composed of the torso joints; with the second group the remaining joints. The vector is computed by extracting the associated angles between joints. The work employs a multilevel ANN that decomposes each joint into a separate model. This allows the recognition of complex motion sequences and assesses their correctness in relation to a predefined model. Vectors are created between first and second joint orders in order to find angles Defines a gesture as a sequence of detected poses.
Strengths	Uses ANN in order to gain a robust method of gesture recognition.
Weaknesses	While the user is performing the gestures, some poses experience an unstable classification. This can lead to a correct gesture followed by an incorrect one or vice versa. Limited novelty there exists human pose estimation libraries based on machine learning, and this does not show significant improvements Does not do any quantitative evaluation itself.
Extra Notes	Separates joints directly connected to the torso (referring to them as first level joints), whereas joints directly connected to the first level joints such as wrists and ankles are referred to as second level joints.

Paper	**Automated assessment of upper extremity movement impairment due to stroke**[29]
Key Contributions	Developed an automated method of measuring the quality of movement in clinically -relevant terms from low -cost motion capture. Unconstrained movements of upper extremity were performed by people with chronic hemiparesis and recorded by standard and low -cost motion capture systems. Quantitative scores derived from motion capture were compared to qualitative clinical scores produced by trained human raters. The study results have shown that using low -cost motion capture with an automated scoring algorithm is a feasible method to assess objectively upper -arm impairment post stroke.
Strengths	Estimates the minimal number of movement repetitions needed in order to get a sufficiently precise motion capture with the Kinect system. This was done by bootstrapping the data in several steps to estimating the errors of averaging one, two, three repetitions of the same movement.
Weaknesses	Takes into account the joint angles of both the paretic and non -paretic arm, deconstructing both into principal components. Those principal components of the paretic and non - paretic arms are compared using a coefficient of determination. This measure constitutes a quantitative sore of impairment. However, this quantitative score of impairment is not widely spread in the clinical world such as WFMT,ARAT or FMA, and is solely based on four joint angles, significantly limiting its effectiveness.
Extra Notes	Uses Kinect and Impulse Patients were tracked by both the Impulse motion capture system and the Kinect simultaneously ,with both data being temporally aligned.

Paper	**Automated Evaluation of Upper -Limb Motor Function Impairment Using Fugl -Meyer Assessment**[21]
Key Contributions	This paper proposes a novel automated Fugl Meyer Analysis (FMA) system to overcome the limitations of the FMA. For automation, they used Kinect v2 and force sensing resistor sensors owing to their convenient installation as compared with body -worn sensors. Based on the linguistic guideline of the FMA, a rule -based binary logic classification algorithm was developed to assign FMA scores using the extracted features obtained from the sensors. It completes 26 out of 33 FMA tests. Features extracted include joint angles, segment rotation, landmark position and open/close hands. Clinician FMA scores compared with FMA automated scores.
Strengths	Takes into account the impracticality of collecting large sets of data from patients and aims at rule -based binary logic classification which may be more practical. Gives patients real time skeletisation feedback. There is a high correlation between clinician FMA and automated FMA. 79 percent of the FMA tests were automated through optimised sensor selection, and approximately 90 percent scoring accuracy was achieved by employing a rule -based binary logic classification algorithm without learning procedures. The proposed system can reduce a clinicians required time for the FMA by more than 85 percent, which would contribute to frequent evaluation of upper -limb motor function and improvement in upper-limb intervention for rehabilitation.
Weaknesses	Requires two sensors motion tracking errors in Kinect V2 due to few subjects severe hand contracture and results in extraction of inaccurate features on the forearm pronation/supination. The threshold used to generate the binary variables were disagreed upon tug levels were bot consistent for each subject although the system used the same grasp threshold Clinicians variable scoring due to synergistic patterns led to errors in the tests.
Extra Notes	Uses Kinect V2 and secondary grip sensor based on rule - based bi nary logic classification algorithm as opposed to machine learning, as states that it is difficult and impractical to collect data from numerous patients with various motor abilities only for machine learning.

1.3.3 Addressing the Research Gap

The framework described in this thesis is based on developing a framework for quantitative assessment of upper limb functionality. Unlike previous research which uses 3D pose estimation systems such as the Kinect, this framework aims to exploit the ease in which capturing 2D data can be collected, in an effort to increase the likelihood that clinicians will adopt the system and continue to generate increasingly accurate models.

The framework operates in the 2D context by modelling the Quantitative Assessment Function (QAF) as a linear combination of non-linear functionality related features. Using these QAFs, clinicians will be able to provide more finely scaled and accurate clinical assessmen depending on the movement exercise being assessed . The upper limbs in the context of this thesis are defined as the region extended from the deltoid region up to hand, including the arm axilla and shoulder. In addition to the upper limbs, the centre of the face at the nose and the base of the neck are modelled, as they are mechanically linked to the upper limbs and movements of these points influence movements in the limbs. In this iteration of the framework, the hands are not currently modelled.

Chapter 2

Basing the Framework on Clinical Standards

2.1 Introduction

It is essential that movement quantification in the 2D context must be based on clinical standards and mathematically defined. The 3D computational assessment frameworks proposed and tested in Chen et al's paper [12] is based on creating a novel, more finely scaled standardised assessment scale than current medical standard measurements, and is tested using 3D technologies. It is done by calculating an evaluation function that corresponds with a clinician's assessment of a patient's upper limb functionality. The 2D framework described in this thesis is based on quantifying patient data in the 2D context in order to build quantitative general models and the relationship between different functionality-related features. Although these two frameworks have different approaches and aims, this 2D framework uses the foundation set in the 3D computational assessment framework and builds upon it.

2.2 Defining the Model Function For Each Feature

The solution to the problem of general movement quantification is derived by determining the QAF for movement exercises which are used to assess patients. The general form of a QAF is defined below:

$$y_i = \sum_{i=1,j=1}^{N,M} w_i^j[\phi_i^j(x(K)_i^j - x(L)_i^j] + ... + w_i^M[\phi_i^M(x(K)_i^M - x(L)_i^M]$$

Where:

y_i is the quantitative assessment score

x_i is a non-normalised feature

φ_i is the normalisation function

w_i is the weight

K,L is the movement pair of the paretic and non-paretic limb

N is the number of data points

M is the number of features included in the model

As shown mathematically, the modelling function is a linear combination sum of the weights and the normalised difference in value of a functionality-related feature between a movement pair. The value of assessment score y^i, will always range within 0 and 1, where 0 indicates unimpaired functionality and 1 indicates a severe loss of functionality.

2.2.1 2D Quantitative Assessment Scale

The optimal model function for the quantitative score will give an assessment value for a particular movement such that the difference with clinical observations $z^{k,l}$ are minimal. This is defined as :

$$min \sum_{K,L} L\left(\sum_{i=1}^{N} w_i^j[\phi_i^j(x(K)_i^j - x(L)_i^j] + ... + w_i^M[\phi_i^M(x(K)_i^M - x(L)_i^M], z_{k,l}\right)$$

Where :

$z_{K,L}$ is the clinical observations of the exercise for both limbs

L() is a loss function

The inputs in order to calculate the model function are the non-normalised functionality-related features $x_1 ... x_N$. These features are mostly kinematic features which are calculated from the exercise data of the patient, while others are functionality indicative measures such as the difference in time taken to complete both trials and the clinician's assessment. Kinematic features can include speed, angles, jerk, segmentation, distance of the paretic trajectory from the reference non-paretic trajectory. Kinematic features can be calculated at key points which represent skeletal movement, such as the nose, neck, shoulders, elbows, and wrist.

The value of $\varphi_i(w_i)$is between 0 and 1, where 0 indicates the ideal profile and 1 indicates the maximum deviation from the ideal profile. This is modelled on the clinical observations that a patient with two fully functional limbs should complete movement tasks with equivalent feature profiles, while an impaired limb will have different levels of deviations from the ideal movement of the unimpaired limb depending on severity of the impairment. The highest level of impairment should have the maximum deviation from the ideal profile. The normalisation function φ

takes the lowest difference between profiles as the ideal movement 0 and the highest deviation as 1.

A difference in this 2D framework to Chen et al's 3D assessment framework is that negative weights are included in the calculation of the model function , whereas in the 3D assessment framework they were set to non-negative ($w_i \geq 0$). Chen et al justify that since $\varphi(x_i)$ is ordered in terms of quality of movement components, w_i should be non-negative. Since there are no pr-established QAF models in the 2D context, it is important negative weights are not excluded and is a property of the data that is reflected in the Support Vector Regression (SVR) model. The impact of these negative weights will be negative on the model, and instead of contributing to the model in calculating a given value, it will contribute in not learning that value. Later on negative weights can be filtered out, but at this stage it is more important to understand the relationship of the quantitative score and the functionality - related features.

In order to find the QAFs for a particular movement the practical steps are follows:

1. Select the optimal rehabilitation assessment test and then the movement exercises that can effectively represent the trajectory of both the paretic and non-paretic limbs. Record video of subjects carrying out these movement exercises in the 2D domain.
2. Use pose estimation in order to find the coordinates of patient's key points during the exercise
3. Using the key point coordinates select and extract functionality-related features from the exercise
4. Create profiles of the functionality-related features for both non-paretic and paretic limb, and calculate the differences in the profiles point by point
5. Find the normalisation function for each difference in kinematic feature
6. Use Support Vector Regression to calculate the weights and the QAF for each movement exercise

2.3 Choosing the Optimal Functionality Assessment to Base Quantification Scale

The limb functionality assessment that would serve as the basis for this quantitative framework was selected in accordance to the criteria that:

- A currently used medical standard assessment scale should be selected
- The movements for both limbs should be clear and visible in 2D
- That the clinician will not be able to have physical contact with the patient

The lack of physical access ruled out FMA, while ARAT and WMFT could be done on visual inspection alone. Between them, there showed a high concurrent validity and correlation between both tests and both show excellent inter-and intra-observer reliability.[28] Ultimately, the choice of assessment to use was the WMFT as the ARAT is usually performed with a more complex practical set up, including a specialised hollow box. I felt that this specialised set up may make it more difficult for clinicians and could occlude limb movements, whereas WMFT exercise practical set ups are relatively simple and require only non-specialised household items.The equipment used in the WMFT depends on the exercise; it ranges from a variety of household items including tables, chairs, boxes, cans, pencils,paperclips and towels. The WMFT quantifies a patient's upper extremity through timed functional tasks. The most widely used version of the WMFT consists of 17 items - The first 6 items involve timed functional tasks, items 7-14 are measures of strength, and the remaining 9 items consist of analysing movement quality when completing various tasks. The examiner should first test the non-paretic limb, then the paretic limb. Some exercises are specially selected in that they should be performed as quickly as possible truncated at 120 seconds, assuming that an individual who could not complete a task within that time would be unable to do so if given unlimited time.[42]The WMFT clinical assessment scale ranges from 1-6, where 1 indicates a complete lack of motor functionality indicates a complete lack of functionality, and 6 a functionality indistinguishable from the non-paretic limb.

These are formally defined as:

- WMFT1-"Does not attempt with UE being tested"
- WMFT2-"UE being tested does not participate functionally; however, an attempt is made to use the UE. In unilateral tasks, the UE not being tested may be used to move the UE being tested".
- WMFT3-"Does attempt, but requires assistance of the UE not being tested for minor readjustments or change of position, or requires more than 2 attempts to complete, or accomplishes very slowly. In bilateral tasks, the UE being tested may serve only as a helper".
- WMFT4-"Does attempt, but movement is influenced to some degree by synergy or is performed slowly or with effort".
- WMFT5-"Does attempt; movement is similar to the non-affected side but slightly slower; may lack precision, fine coordination or fluidity".
- WMFT6-"Does attempt, movement appears to be normal".

[43]

This framework adopts the quantitative scale introduced by Chen et al [12]- ranges from 0 denoting unimpaired function, 0.25 denoting mild functionality loss, 0.5 at moderate loss, 0.75 at serious loss , and 1 denoting severe loss deviation, increasing in increments of 0.05. 0.5 denotes a moderate level of functionality loss. This scale is simple enough for clinicians to adopt easily, while still being fine enough to remain informative.

2.4 Selecting Movement Exercises

The appropriate WMFT exercises were chosen and framed in such a way that the both limbs could be clearly seen in the 2D context. Since hands and fine motor function of the hand and fingers are not currently modelled in this framework, exercises that include finger functionality such as picking up paper clips with a pincer grip are not used.

The exercises chosen were:
a) Lift Can- the subject attempts to lift a can and bring it close to his/her lips with a cylindrical grasp.
b) Forearm to box - The client attempts to place the forearm on the box by abduction of the shoulder
c) Weight to box: subject attempts to place a weighted object on the box placed on the table top.

2.5 Utilising OpenPose

Given the accuracy and popularity of OpenPose, it was adopted as the means of subject pose estimation in this framework. [?][7] Once the clinical assessment scale and movement exercises was chosen, pose estimation was applied to 2D video of patient's completing the movement exercises.

OpenPose identifies a skeleton for each person in an image or single frame of a video, as a collection of key points, seen in Figure 2.1. For each key point, it can output the X,Y coordinates in terms of pixel position, as well as a confidence score of that key point's position.

OpenPose pose estimation data can be written a variety of formats including JSON files, images and video. This framework is primarily based on the JSON output of OpenPose. For each frame of an input video of a patient completing a movement exercise, OpenPose outputs a JSON file. Each JSON file contains the number of skeletons in a particular frame, and the X,Y position data and confidence score for each key point.

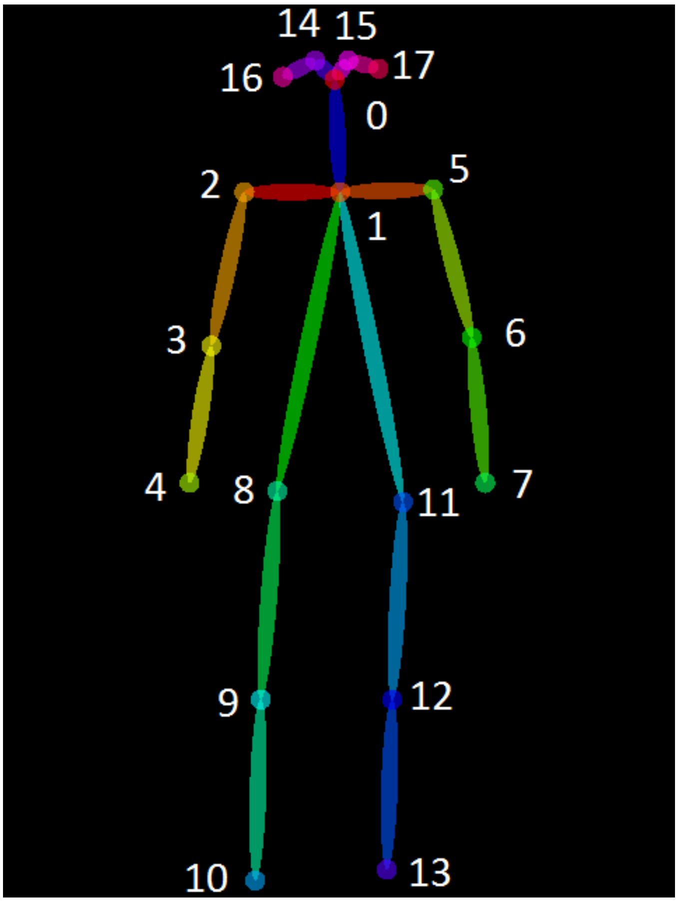


Figure 2.1: Front view of an OpenPose Skeleton in COCO Format. The position of the key-points are defined as follows: 0-Nose, 1-Neck, 2-Right Shoulder, 3-Right Elbow,4-Right Wrist,5-Left Shoulder,6-Left Elbow,7-Left Wrist, 8-Right Thigh, 9-Right Knee,10-Right Foot,11-Left Thigh,12-Left Knee, 13-Left Foot, 14- Right Eye, 15- Left Eye, 16- Right Ear, 17 - Left Ear

Although this framework processes the whole skeleton output, only the upper limbs are considered,calculating the features of key points 0,1,2,3,4,5,6,7. For those that wish to adapt the framework for the quantification of some other movement, different sections of the skeleton can be selected by a simple change in the source code.

2.6 Experimental Set-Up

The camera should view the patient from the side at roughly 90 degrees, so as limit information lose to the non-existent z axis of the 2D video, and the camera should remain in the same position at all times. It should placed approximately a metre off the ground in order to achieve full body capture while reducing angles. The framing of the subject in this manner is consistent with previous research demonstrating the sufficient accuracy of conventional 2D video for quantifying the upper limb kinematics to measure hand activity level in repetitive motion occupational tasks. [10] Each subject completed the same movement exercise once per limb, with the limb that is completing the movement exercise always closest to the camera and non-occluded by any object. Boxes were placed in front of the subject. If the subject was seated or standing, this remained the same for movement exercises of both limbs, and if seated the chair should not change heights. This is to reduce discrepancy between key point coordinates of both limbs due to non-functionality related occurrences. No truncation was necessary, as all exercises were below the 120s limit.

The movement exercise for each limb had its own video. The trajectories of each movement exercise were captured in their entirety and not edited in any way between the start and end of the movement exercise. The reason for the entire motion trajectory being captured was two fold. The first reason is that there remains a lack of studies addressing more complex, everyday tasks with ecological validity, and thus more kinematics investigating purposeful everyday tasks are warranted. [3] By including the entire movement trajectory and not isolating different components of it, a more holistic metric profile can be derived from any functionality-related feature that better represents the entire motion. The start and end of the movement exercise were clear,distinct identifiable points- for all movement exercises recorded, subjects started and ended with their hands flat on their knees. This was to limit the amount of frames where the movement was not being enacted, which would shift the profile of any given functionality-related feature, with the profiles for each limb being shifted to different extents. The camera used was the Samsung mobile S6 camera, in order to collect and test data captured by a commonly used mobile .

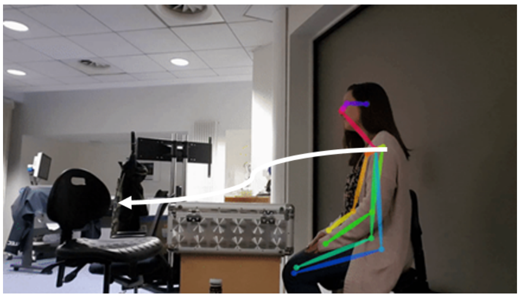
2.7 Results

Videos of healthy patient movement exercises for the lift can, forearm to box and weight to box were recorded following the experimental set-up. Three healthy patients were used for each movement exercise. In addition to the healthy patients, simulated stroke trials where a WMFT2 loss of functionality in one limb for each exercise was recorded. These simulated stroke exercises were used as placeholders for clinical trials and real patient data. The stroke level simulated in the video was based on observations of stroke patients in hospital and video evidence online. The clip for each limb was input to OpenPose which generated a JSON File for each frame. Images representing the OpenPose skeletons overlaid on patients starting the movement exercise in both the healthy trials and simulated stroke trial are shown in Figure 2.2. The pixel dimensions of the camera used was 1280 x 720 pixels.

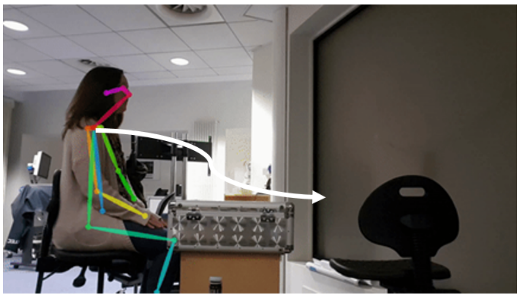
2.8 Discussion

Advantages

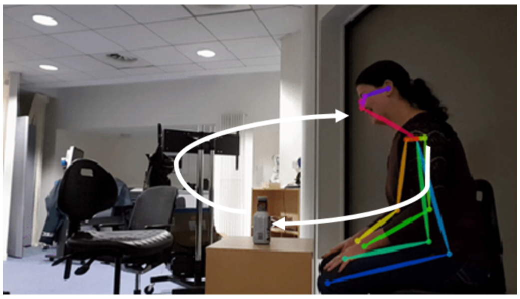
Modelling the framework on clinical standards means that the QAF learnt for each exercise would have direct use to clinicians, aiding them in accurately assessing patient functionality. Another advantage of this framework is that the experimental set-up for collection of patient data is very simple and does not require special equipment, other than tables, chairs and a camera, which should encourage busy clinician's to collect more data. In addition, videos of subjects can be recorded at any time and anywhere, without needing to be recorded in front of a particular system, such as is done with the Kinect.



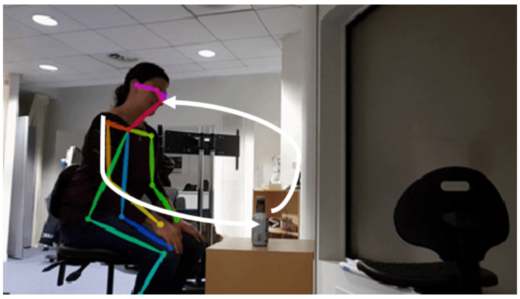
(a) Forearm to Box: Non-Paretic Limb



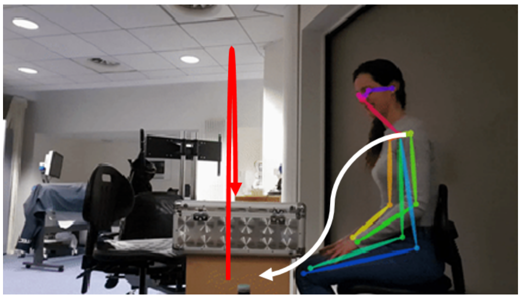
(b) Forearm to Box: Paretic Limb



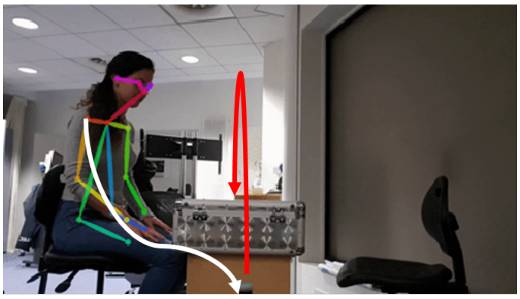
(c) Drink Cup Exercise: Non-Paretic Limb



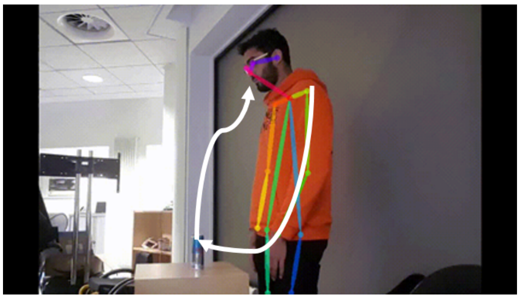
(d) Drink Cup Exercise: Paretic Limb



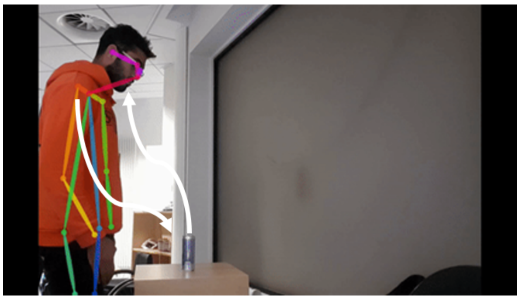
(e) Lift Weight Exercise: Non-Paretic Limb



(f) Lift Weight Exercise: Paretic Limb



(g) Simulated Stroke Cup Exercise: Non-Paretic



(h) Simulated Stroke Paretic Exercise: Paretic

Figure 2.2: OpenPose on Subjects Carrying Out Movement Exercises

Disadvantages

The normalisation function φdoes not use a range of values which it considers ideal and maximum deviated, only the lowest and highest differences between profiles. This can cause some inaccuracy because movements of the different limbs will not be exactly the same in terms of trajectory. Physiotherapist labelled video of patients with stroke compared to healthy patients are needed in order to compare the two and create a range of ideal movement and deviated movement.

Another issue is that the patient videos must be run through OpenPose, and the JSON files for each frame evaluated. Clinicians may need to learn how to use the system as there is currently no GUI implementation. In addition, in order to utilise OpenPose's full functionality using the Nvidia parallel computing platform CUDA and it's associated GPU-accelerated library of primitives for deep neural networks (cuDNN), clinician's should have access to a computer with a GPU. However, only one GPU capable computer is needed in order to process many patients.

On processing some of the subject videos through OpenPose, the frame rate dropped significantly, from standard 40fps to around 9. This low efficiency frame rate is highly likely to contribute to jitter of the movement trajectory, decreasing the accuracy of the data captured.

Problems

Initially I encountered difficulties in the building of OpenPose. However, after following instructions online and trying different builds, this issue did not arise again.

Future work

Exercises that tested fine motor function of the fingers, such as flipping cards or picking up paper clips were not used as hands, fingers and thumbs are not modelled in the iteration of this framework. This is because they require a separate pose estimation model than that of the base skeleton ranging from the centre of the face at the nose through to the wrist. Once this 2D framework is clinically validated for the upper limbs, the inclusion of the hands and exercises for fine motor function are the next natural step. In addition, the simulation stroke movement exercise for the cup exercise and it's clinical assessment score of WMFT2 is not valid and exists only as a proof of concept while clinical trails pend approval. Real videos of patient's completing movement exercises that follow the experimental setup, and a trained physiotherapist which labels that video with a clinical assessment must be collected in order to validate the model function of the clinical assessment for a given movement exercise. OpenPose developers are consistently improving their model to run at more optimised speeds, and in future more efficient models should be utilised in order to ensue the smoothness of movement trajectories.

Summary

Basing the framework on clinical standards means that results are immediately translatable to clinical work. The advantages of basing the framework on clinical data and the simple and reproducible experiment set-up outweigh the disadvantages of the need to learn OpenPose and have access to a GPU-capable computer. Future work will include reducing errors by using margins incorporating ideal and deviated ranges and ensuring that frame rate during processing is consistently at an acceptable standard, as well as implementing fine motor function in order to build a more comprehensive model of the upper limbs.

Chapter 3

Functionality-Related Feature Extraction and Comparison

3.1 Introduction

Given the position data of each key point of a subject while they complete a movement exercise, features which relate to the functionality of the upper limb can be extracted. For each frame, these features are extracted and a profile for the entire movement exercise generated. The profiles of both limbs are compared and the difference between them at 100 points calculated. These differences provide information on the level of a subject's motor functionality. These differences are stored and form the basis of the data of which the model functions of the features will be derived.

3.2 Selecting Features

The selection of what features to use in order to calculate the model function for a movement task is non-trivial, as it is essentially breaking down an entire movement into its constituent elements and selecting the features of that movement that relate to upper limb functionality. The 2D context poses further restrictions on what kind of features are possible to extract, limiting the potential information that can be used. Most functionality-related features are kinematic ones, although there exists other non-kinematic indicative features. The features extracted in this framework are:

1. The speed of the key points at the nose, neck, shoulders, elbows and wrist
2. The jerk at the key points at the nose, neck, shoulder, elbows, and wrist
3. The angle formed at the elbow
4. The difference in the trial size for both paretic and non-paretic trials

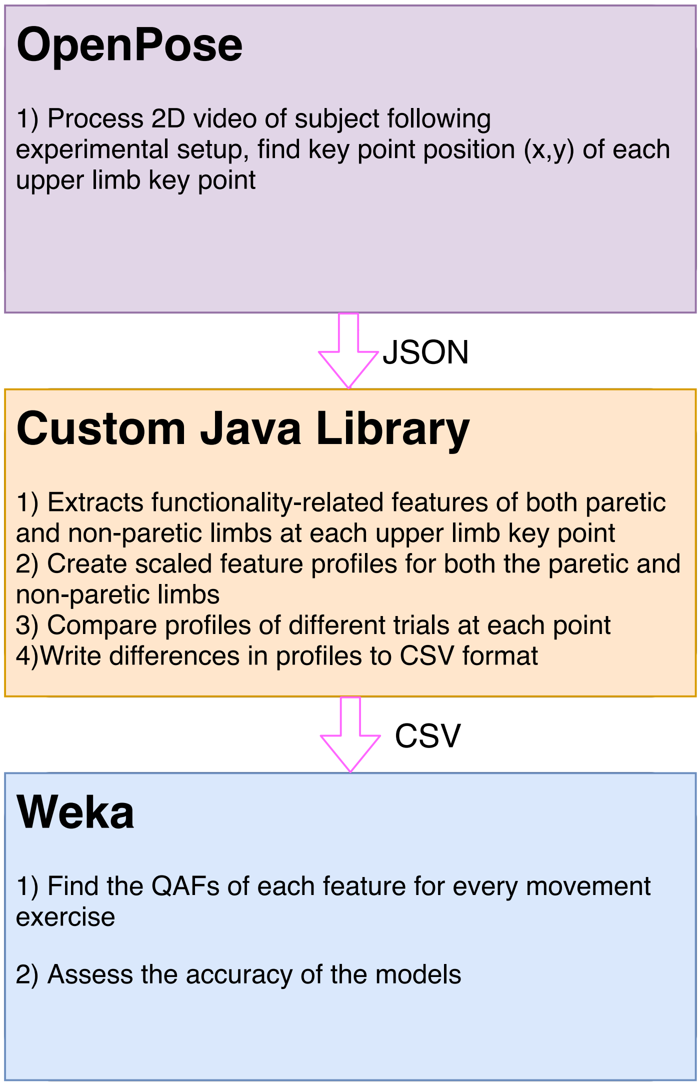
The features selected for this framework are based on previous research [39], which utilises trajectory error, velocity profile deviation, jerkiness and segmentation. Segmentation of a movement occurs when the opening of the elbow does not synchronise with the shoulder moving forward. In the current iteration of this framework, segmentation is not included. This is because the calculation of segmentation requires the projection of the 3D trajectory onto the X-Z and Y-Z plane, which is not possible given the 2D context. The addition of the trial size feature was due to clinical observations that paretic and non-paretic limbs complete the same movement exercise with varying time dependant on the level of impairment.[36] The magnitude distance was calculated rather than the difference in terms of X,Y coordinates to abstract the clinician from viewing pixel coordinates and simplify results. Given the side view framing of subjects, the only angle that is possible to extract with relative accuracy is the angle formed at the elbow.

3.3 System Pipeline and Java Implementation

In order to extract the functionality-related features from the OpenPose output JSON files, a novel Java library was written. The library reads in all the OpenPose JSON coordinate files of both limb trials. For each frame, a skeleton object was defined, with all the coordinates of the key points which compose it. For each key point in both limbs, all the features were calculated. The difference in trial size is calculated, after which the library scales the different length trials equally and calculates a profile for each feature. These scaled profiles are then compared point by point at 100 equally spaced points, and the difference between them and written in a CSV file, in preparation for SVM classification and the learning of the model functions. The only input need from the clinician/user is the path locations to the input JSON files and the desired output folder path for the CSV files.

The library is open source and provides a platform for clinicians and researchers to use in order to alter or build upon for their own 2D based functionality-related feature experiments, or when building their own 2D general quantitative models. Graph functionality in the code also allows for immediate display of the feature profiles of different limbs on the same graph, which visually can provide graphical indication of the difference in a feature between the two limbs to a clinician.

The library was implemented in Java as it is the language that runs on the most devices worldwide. By writing in Java, it is a step towards processing of patient data on mobile devices providing additional simplification and ease of use for the clinician.


OpenPose
1) Process 2D video of subject following experimental setup, find key point position (x,y) of each upper limb key point
JSON
Custom Java Library
1) Extracts functionality-related features of both paretic and non-paretic limbs at each upper limb key point
2) Create scaled feature profiles for both the paretic and non-paretic limbs
3) Compare profiles of different trials at each point
4)Write differences in profiles to CSV format
CSV
Weka
1) Find the QAFs of each feature for every movement exercise
2) Assess the accuracy of the models

3.4 Calculating Features

Using position coordinates of a key point (x, y), the kinematic features were calculated at each frame, f. For every frame the angle at the elbow, θ, was calculated (degrees). Between every 2 frames, the speed (pixel/frame) was calculated, and between every 4 frames the jerk (pixel/frame^3) was calculated. The feature equations are defined as:

$$speed(f) = \frac{x(f), y(f)}{f}$$

$$jerk(f) = \frac{x(f), y(f)}{f^3}$$

$$\theta(f) = arccos \frac{uv}{\mid u \mid \cdot |v|}$$

Where:

- The units of the x,y coordinates are in terms of pixels
- u denotes the vector from the shoulder pointing to the elbow
- v denotes the vector from the elbow pointing to the wrist

This is shown in Figure 3.1.

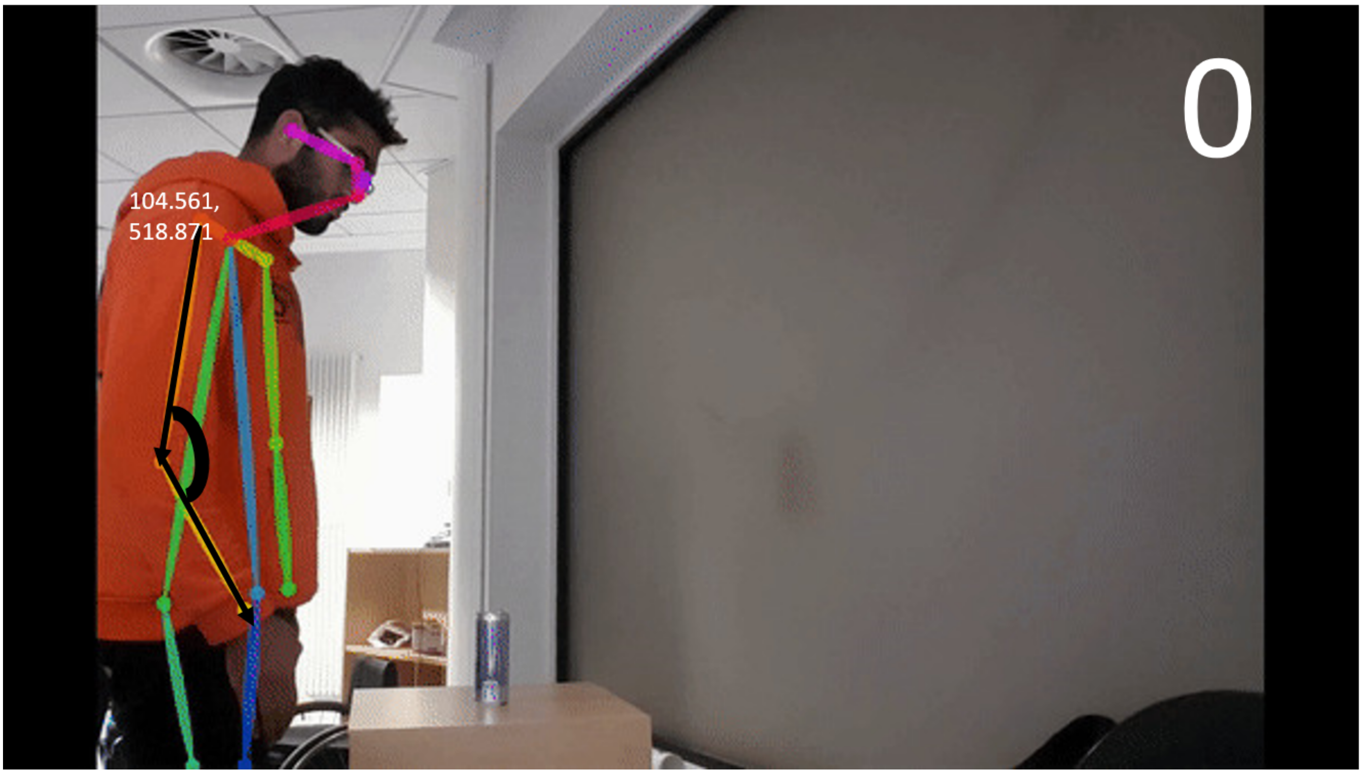
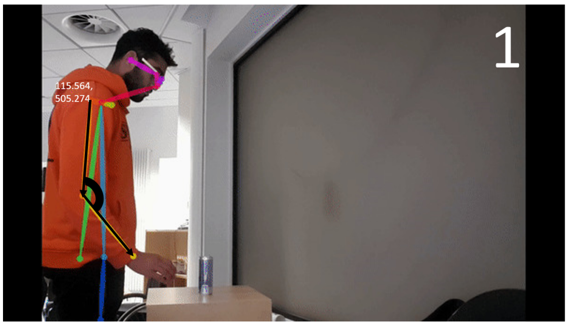
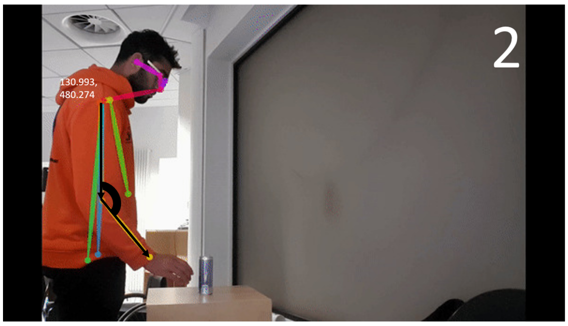


Figure 3.1: A single frame containing (X,Y)coordinates of the shoulder key point , and the angle at the elbow, θ

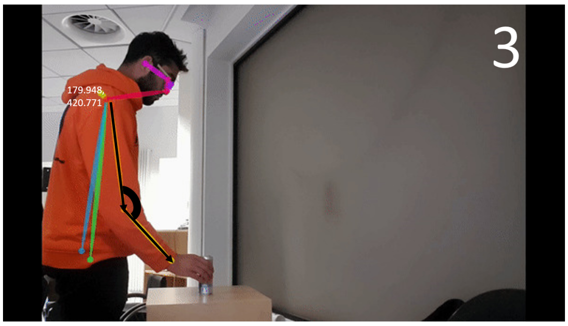
There is a loss of 2 frames at the beginning and end of the movement exercise due to the jerk being the third derivative of the x,y position. The loss of these data points should be negligible as they do not occur during the movement trajectory and affect profiles of both limbs equally.



(a) Speed : pixel/frame



(b) Acceleration: pixel/frame^2



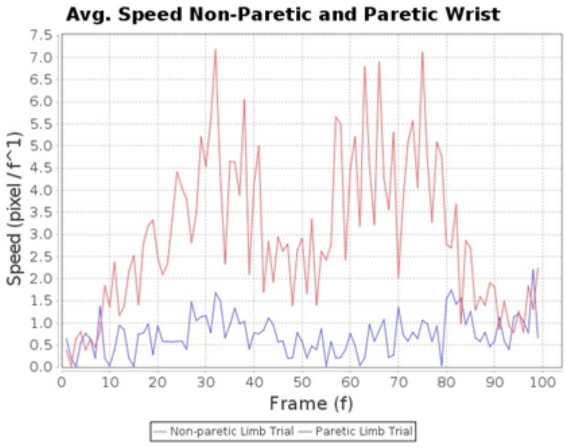
(c) Jerk: pixel/frame^3

Figure 3.2: Calculation of Features

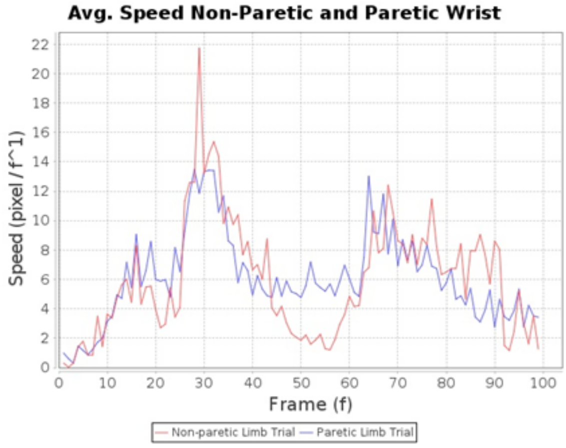
3.5 Calculation and Comparison of Feature Profiles

Once all the features were calculated at each frame for both limbs, they were scaled to the same size, in order to give comparable profiles and minimise displacement of the trajectory to different regions by a difference in trial length. This gives a more accurate comparison, where the same regions on both profiles represent the same sections of the subject's movement trajectory. The graphical representation of the types of movement profiles calculated for the different limbs are shown in Figure 3.3.

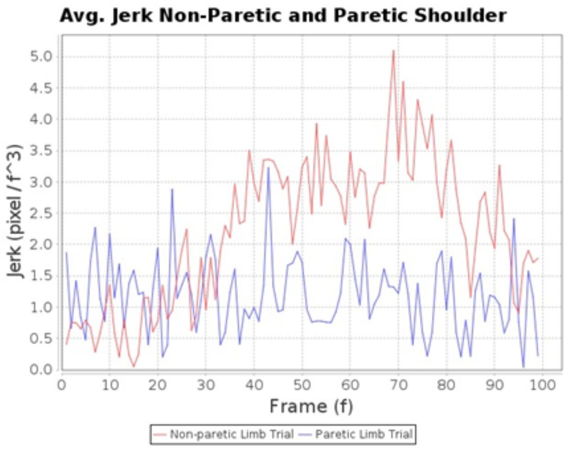
The feature profiles of the simulated stroke limb and the healthy limb are much more disparate than the feature profiles of the healthy left and right limbs in the healthy trial.



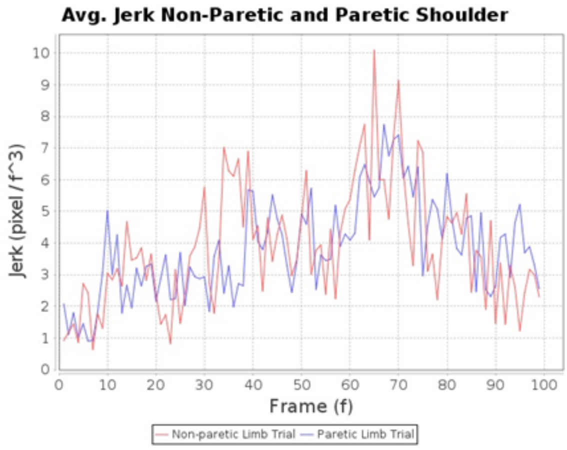
(a) Speed Profiles. Healthy Wrist in Red, Sim. Stroke Wrist in Blue



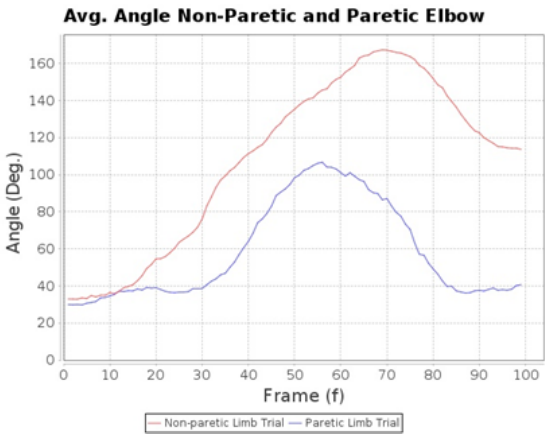
(b) Speed Profiles. Healthy Left Wrist in Red, Healthy Right Wrist in Blue



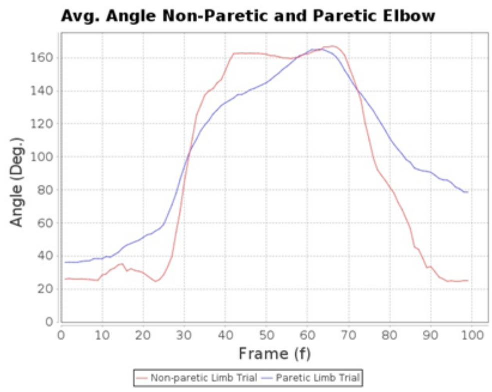
(c) Jerk Profiles. Healthy Left Shoulder in Red, Sim. Stroke Right Shoulder in Blue



(d) Jerk Profiles. Healthy Left Shoulder in Red, Healthy Right Shoulder in Blue



(e) Angle Profiles. Healthy Left Elbow in Red, Sim.Stroke Right Elbow in Blue



(f) Angle Profiles. Healthy Left Elbow in Red, Healthy Right Elbow in Blue

Figure 3.3: Examples of Feature Profiles of Both Limbs

3.6 Results

For each feature in both limbs, the difference between the feature profiles of both limbs are calculated at 100 points. Given that the average trial size was around 300 frames for a healthy patient, 100 points was chosen as a median point choice as a high number points would mean the profiles were to be compared on a very high scale, increasing the likely hood of comparing sudden spikes in a profile found on a fine scale, causing inaccuracies. Too low a scale and the trends of the profile would not be followed. For each exercise, the differences in all feature profiles were stored in CSV format in preparation for use in regression.These values form the basis for calculating the general quantitative model functions. In addition, the metric graphs produced by the library can be viewed directly by the clinician to view the disparity between the feature profiles of both limbs for a given movement exercise.

3.7 Discussion

Advantages

Features are selected on existing clinical basis related to limb functionality. In addition, scaled profiles are created in order to compare trends describing the same motion in both trials, reducing errors caused by the displacement of profiles from each other. The framework provides a simple open-source Java library for clinicians and researchers to utilise, change or build upon to fit individual needs and experiments . It is easy to operate, and only requires the input paths of the JSON file locations and the desired output folder of the CSV files to run.The code is available for cloning on GitHub.

Disadvantages

For a single subject, the optimal variance in the difference between the metric profiles for a given functionality-related feature should remain low- this axiom is derived from the observation that a single subjects motor functionality will remain the same in a single trial. However, there remains a high amount of variance and spiking in the feature profiles due to in accurate jitter of key points between frames and incorrect key point estimates. The extent to which these errors affect the QAFs are described in section IV.As mentioned in part II, the frame rate occasionally dropped to low single digits while processing, likely contributing further to reduced accuracy, yet the extent of its effects has yet to be studied.

Another potential source of error is that the difference in value change of the Z-axis or in angle is not accounted for via normalisation between different trials. This means that even small changes of distance or angles will affect the proportion of pixels movement. The extent of effect of the pixel position of the values in relation to the distance and angle of the subject on the data must be calculated in order to build a normalisation function. Changes in depth and angle is even more prevalent when generating models of multiple subjects, as different subjects will never sit or stand in exactly the same fashion as each other.

Some functionality-related clinical features such as segmentation was not used as the 2D context poses further restrictions on what kind of features possible to extract rather than 3D. The magnitude distance of the paretic trajectory from the reference non-paretic trajectory is not currently calculated for the key points at the nose,neck, shoulders, elbows and the wrist as there currently no consistent reference point to be able to mirror that reference trajectory to the paretic arm, losing potentially valuable information.

Problems

The challenge of writing the 2D feature extractor and comparison library in Java from scratch was very time consuming . However, with much time investment and effort it was built,debugged and tested.

Future Work

Low pose estimation scores, or a change in position of a given key point between two frames which is extremely high should be filtered out, reducing jitter and misidentifying limbs.

The utilisation of physical markers of known dimensions on screen will allow for the z-axis to be recovered using 2D data. In addition, by taking a one-of fixed measurement of the upper limb (such as the fore-arm) the angle that the subject has shifted in each frame can be calculated. The implementation of these two features will allow for normalisation of different subjects, resulting in more accurate QAFs.

Working with trained clinician can identify and add more functionality-related features, to build more advanced and accurate general quantitative models.

In terms of user experience, an application that allows for on board processing, GUI, OpenPose and regression integration will further encouraging clinicians to adopt and use this framework. Since the library is already written in Java, it can be ported and built upon to suit mobile devices.

Summary

Un-smooth movement trajectories due to variety of factors including jitter, mis-identification, slow frame rate and lack of normalisation in termsof the Z axis and the angle cause the propagation of errors. These must be filtered out using thresholds of change between frames and pose estimation scores respectively. In addition, normalisation measures such as physical marker on screen and knowledge of a subject's limb length will allow for regaining depth and angle, which will improve QAF accuracy. In order to mitigate these problems, the Java custom library provides clinicians and researchers with a solid starting platform to change and build their own models and feature extractors. The novelty of the library is a benefit to the motor function quantification community.

Chapter 4

Learning the Quantitative Assessment Functions

4.1 Introduction

Using the data on the difference in profiles for each functionality-related feature, the QAFs for each movement exercise for can be calculated using supervised learning. The class of supervised learning algorithms utilised in this framework in order to derive the quantitative model functions are Support Vector Regression models (SVRs). Since the magnitudes of weights correspond to the importance of each feature in classifying data points with larger weights indicating more relevance to a given feature, by applying multiclass SVRs and calculating these weights for every functionality related feature and th clinician's assessment, the QAF can be learnt. In addition, the Root Mean Squared Error is calculated in order to identify the accuracy of the SVR.

4.2 Support Vector Regression

Support Vector Machines (SVMs) is a popular algorithm used to analyse data for classification analysis. They are inherently binary discriminitive classifiers formally defined by a separating hyperplane. Given labelled training data, the algorithm outputs an optimal hyperplane which categorises new examples. In 2D space, this line divides the plane in two sections in which each class lays either side, where negative scoring functions as classed as belonging to y=-1, and positive scoring functions are classed as belonging to y=1. The extension of SVM for regression is called Support Vector Regression (or SVR). SVR works by finding a line of best fit that minimises the error of a cost function. Non-linear multi-class values such as those in this framework will have non-linear separations that depend on the polynomial used. These separations are drawn using an optimisation process that only considers those data instances in the training closest to the line with the minimum cost. SVR was chosen over other models such as linear regression as it the flexible curves of separations that SVR can produce with a non-linear kernel is better suited for learning the boundaries of the non-linear kinematic features.

The SVM regression algorithm utilised in this framework is the popular Sequential Minimisation Optimisation(SMO). SMO implements John Plat's sequential minimal optimisation algorithm for training a support vector classifier. As SMO spends most of its time calculating the decision function, rather than performing quadratic programming, it can exploit data sets which contain a significant number of zero elements. These datasets are called sparse. SMO performs particularly well on sparse data sets. [32] the robustness of SMO against sparse datatests is a valuable asset. Given that the 2D context of the input data imposes limitation on functionality-related features data can be extracted, depending on the movement exercise and quantitative model that a clinician wishes to derive, sparse data may occur. Thus, this framework accounts for sparse data via SMO.

4.3 SVR Implementation

The SVR was implemented in Weka, as its GUI and visualisation tools allowed for rapid prototyping of different SVR models. Clinician's and researchers who want to create their own 2D quantitative models can utilise Weka without needing to implement code. However, the CSV output provided by the custom Java is a standard file format and can be input to whatever machine learning implementation is desired.

A multi-class SVR with a polykernel was run for each functionality related feature for both healthy and simulated stroke trials, as well as the clinical assessment.The polykernel performed better than the RBFKernel when compared in an RMSE Error T Test, with the polykernel gaining a lower RMSE value(0.43 vs 0.90 on the drink cup data).Thus, it was chosen as the kernel for learning the QAFs. The complexity parameter C, controls how flexible the process for drawing the hyperline can be. This was set to 1000 when learning the QAFs , as it gave similar results to C values of higher exponents on the same data set, while running at a reasonable speed.

In order to assess the accuracy of each SVR classifier, The RMSE was calculated. For each movement exercise,the magnitude of the weights for every other feature was calculated. The normalisation function φimplemented was to take the lowest difference between profiles as the ideal movement 0 and the highest deviation as 1. Using the magnitude of the weights, the QAF of each feature for every movement exercise was learnt. The most reliable way to evaluate the SVM algorithm for each feature was by utilising stratified 10- fold cross evaluation.

4.4 Results

The QAFs that were learned for each movement exercise is found below.

Where

- y_i is the quantitative assessment score ranging from 0 to 1
- d_i represents the difference in non-normalised features for a movement pair K,L (x_i(K) -x_i(L))

4.4.1 Forearm to Box Exercise Quantitative Assessment Function

y_i=

$0.0012[\phi_i^{noseS}(d_i^{noseS})]$

$+ 0.0019[\phi_i^{neckS}(d_i^{neckS})]$

$- 0.0016[\phi_i^{shoulderS}(d_i^{shoulderSpeed})]$

$- 0.0056[\phi_i^{elbowS}(d_i^{elbowS}]$

$- 0.0147[\phi_i^{wristS}(d_i^{wristS})]$

$+ 0.0297[\phi_i^{elbowA}(d_i^{elbowA})]$

$- 0.0012[\phi_i^{neckJerk}(d_i^{neckJerk})]$

$- 0.0003[\phi_i^{shoulderJerk}(d_i^{shoulderJerk})]$

$- 0.0031[\phi_i^{elbowJerk}(d_i^{elbowJerk})]$

$- 0.0032[\phi_i^{wristJerk}(d_i^{wristJerk})]$

$+ 0.995[\phi_i^{trialSize}(d_i^{trialSize})]$

$- 0.0002$

4.4.2 Drink Cup Exercise Quantitative Assessment Function

y_i=

$-0.0009[\phi_i^{noseS}(d_i^{noseS})]$

$+ 0.0011[\phi_i^{neckS}(d_i^{neckS})]$

$- 0.0032[\phi_i^{shoulderS}(d_i^{shoulderSpeed})]$

$- 0.0013[\phi_i^{elbowS}(d_i^{elbowS}]$

$+ 0.0026[\phi_i^{wristS}(d_i^{wristS})]$

$- 0.0001[\phi_i^{elbowA}(d_i^{elbowA})$

$+ 0.0002[\phi_i^{noseJerk}(d_i^{noseJerk})]$

$- 0.0002[\phi_i^{neckJerk}(d_i^{neckJerk})]$

$- 0.0063[\phi_i^{shoulderJerk}(d_i^{shoulderJerk})$

$+ 0.0024[\phi_i^{elbowJerk}(d_i^{elbowJerk})]$

$+ 0.0076[\phi_i^{wristJerk}(d_i^{wristJerk})]$

$+ 0.9983[\phi_i^{trialSize}(d_i^{trialSize})]$

$- 0.0022$

4.4.3 Lift Weight Exercise Quantitative Assessment Function

y_i=

$0.0052[\phi_i^{noseS}(d_i^{noseS})]$

$+ 0.0005[\phi_i^{neckS}(d_i^{neckS})]$

$+ 0.0057[\phi_i^{shoulderS}(d_i^{shoulderSpeed})]$

$+ 0.0026[\phi_i^{elbowS}(d_i^{elbowS}]$

$+ 0.0099[\phi_i^{wristS}(d_i^{wristS})]$

$+ 0.0034[\phi_i^{elbowA}(d_i^{elbowA})$

$- 0.0082[\phi_i^{noseJerk}(d_i^{noseJerk})]$

$+ 0.0018[\phi_i^{neckJerk}(d_i^{neckJerk})]$

$+ 0.0024[\phi_i^{shoulderJerk}(d_i^{shoulderJerk})$

$+ 0.001[\phi_i^{elbowJerk}(d_i^{elbowJerk})]$

$+ 0.002[\phi_i^{wristJerk}(d_i^{wristJerk})]$

$+ 0.9962[\phi_i^{trialSize}(d_i^{trialSize})]$

$+ 0.0072$

4.5 Discussion

4.5.1 Evaluating the Accuracy of the Model Functions

Using the Root Mean Square Error and comparing it to the variance of the data set, we can draw general assumptions on the accuracy of the SVRs and thus the accuracy of the model functions derived from them. There does not exist formal thresholds for what accuracy levels are regarded as acceptable and what are not- the RMSE in relation to the variance of the data provides an indicator to its accuracy. The utilisation of the SVR indicated that the large variability in the the kinematic feature data sets of many due to jitter and mis-identification, has a significant negative impact on the accuracy of the SVR, achieving a lower accuracy than desired for creating the general QAF of many features for clinical use. Thus, in order to create these accurate models, jitter and spikes must be filtered, in addition to the implementation of normalised models and an investigation to how that affects the accuracy of the classifier. The RMSE results for the QAF and other other accuracy metrics can be found in the appendix.

For the QAFs, an RMSE of 0-0.05 would be acceptable asthe assessment score falls in between 0 and 1, with increments of 0.05. Although the QAFs generated have RMSEs within this optimal range, this is not due to a balanced dependence over all features - the weights attributed to all kinematic features are relatively small. The reason that the QAFs have a high accuracy is due to the inclusion of the difference in trial size feature, of which is very heavily weighted. This feature had less variability in data and had a negative linear relationship with the score as the difference in trial size between the paretic limb trial and the paretic limb trial decreases, the score increased. Because the QAFs are so heavily weighted on the trial size, assessment scores are dependant on that sole variable, with assessment scores ranging at either the 0 or 1 boundary. The variance in the data set should be decreased before expecting the clinical assessment QAF to correlate with clinical observations.

Advantages

The advantages of calculating QAFs for each movement exercise is that the same quantitative scoring system can be applied to multiple movements, standardising the score for the clinician while simultaneously providing unique assessment depending on the exercise. In addition, given that the framework out puts a standard CSV file of the differences in kinematic features, clinicians and researchers can optimise run and different regression algorithms in order to suit their clinical needs.

Disadavantages

As explained in section III, The RMSE is unacceptably high for many kinematic features due to a variety of accuracy reducing aspects. This meant that the QAFs learnt were overly weighted on trial size rather than the kinematic features extracted from the movements, and assessment scores given were usually weighted at 0 or 1, losing the fine-scale value which is desired.

Problems

The challenge of framing the clinical problem of quantifying an assessment score into a machine learning problem required me to delve more deeply into how different types of classifiers and regression models work, their advantages and disadvantages, in order to find the right solution.

Future Work

Mitigating these accuracy reducing aspects should improve the quality of the QAFs learnt by weighing a higher importance on kinematic features than is present currently. In future different types of models, kernels and C values will be optimised further to create maximum hyperlines which better do regression values.

Summary

In summary, due to propagating inaccuracies in kinematic features, the QAFs learned weighed too heavily on trial size rather than on those features, and thus its ability assess on the finer scale was reduced. However, a framework was provided in which these problems can be solved with general ease.

Chapter 5

Conclusion

5.1 How Can These Results Be Used by the Scientific Community

Clinicians and researchers can apply and build upon this framework in order to derive their own 2D based QAFs for other assessment scales, and in order to generate better QAFs of different assessment scales and exercises. In addition, they can follow the experimental setup proposed to use in their own experiments. The novel open source 2D feature extractor library in Java can be utilised in order to effectively extract features, compare profiles, write CSV files of the data extracted, or generate clinically relevant metric profile graphs from 2D data. In addition, they can evaluate and optimise the QAFs learnt in this framework in their own experiments and research.

5.2 Limitations and Future Work

Developing in the 2D domain came with many challenges, including effects of jitter of key points between frames, inaccurate pose estimation, and a loss of data in the Z axis. These inaccuracies propagated to the reduce the ability of the QAFs to provide accurate quantitative clincial scores. In addition, low frame rates experienced when running OpenPose may have contributed to inaccuracy. Movement profiles should be smoothed by filtering out jitter between frames by setting a threshold of change between frames that is deemed reasonable, and ignoring values that exceed it. Mis-identification can be dealt with by utilising pose estimation scores that are output along with key point coordinates, and choosing to ignore key point positions of low confidence. Also, different types of computer system setups can be experimented with in order to find an optimal consistent framerate. In addition, the effect of changes in distance and angle must be studied on the kinematic features must be studied. In order to improve accuracy, a normalisation method using a physical marker of known dimensions (such as a Hiro or Aruco Marker) can be placed on screen, and using computer vision techniques the Z axis can be recovered to a degree. In order to normalise angles, a one-time measurement of a subjects fore arm length can be done, and by comparing the known length to the length present in the frame, information on the subject's angle can be inferred. Also, ranges of idealised deviated motions can be identified with the aid of a physiotherapist learnt as a feature in order to inform the QAFs on movement regions rather than solely on individual points. Finally, in terms of implementation, an app can be developed using the custom Java library in order to increase ease of use, and lower the barrier of clinician adoption.

5.3 Summary of Thesis Achievements

1. Identified a research gap in quantifying upper limb functionality using 2D in order to exploit the immense amount of potential data and lower the barrier to clinical adoption by means by both literature review and first hand-experience observing at a stroke word

2. Developed a framework based on clinical standards so that progress made would be directly applicable to the clinical setting

3. Developed an experimental setup in which to optimise the quality of data given the challenge of 2D data collection

4. Collected both healthy data and simulated stroke data to make up the data set from a variety of subjects to act as a placeholder for clinical trials

5. Wrote an open source 2D kinematic feature extraction library from the ground up in Java in order to extract kinematic features, create comparable feature profiles, calculate the differences, and write them into standard CSV format

6. Learnt the quantitative assessment function of three clinically standard exercises using machine learning techniques

7. Identified sources of error with the current iteration of the framework and proposed solutions through future work

Bibliography

[1] World Population Prospects The 2017 Revision. Technical report.

[2] Richard J. Adams, Matthew D. Lichter, Eileen T. Krepkovich, Allison Ellington, Marga White, and Paul T. Diamond. Assessing Upper Extremity Motor Function in Practice of Virtual Activities of Daily Living. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2015.

[3] Margit Alt Murphy and Charlotte K Häger. Kinematic analysis of the upper extremity after stroke how far have we reached and what have we grasped? *Physical Therapy Reviews*, 20(3):137–155, 6 2015.

[4] Bambi R Brewer, Sharon K McDowell, and Lise C Worthen-Chaudhari. Poststroke Upper Extremity Rehabilitation: A Review of Robotic Systems and Clinical Results. *Topics in Stroke Rehabilitation*, 14(6):22–44, 12 2007.

[5] J W Burke, M McNeill, D Charles, P Morrow, J Crosbie, and S McDonough. Serious Games for Upper Limb Rehabilitation Following Stroke. In *2009 Conference in Games and Virtual Worlds for Serious Applications*, pages 103–110, 2009.

[6] J W Burke, M D J McNeill, D K Charles, P J Morrow, J H Crosbie, and S M McDonough. Augmented Reality Games for Upper-Limb Stroke Rehabilitation. In *2010 Second International Conference on Games and Virtual Worlds for Serious Applications*, pages 75–78, 2010.

[7] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *CoRR*, abs/1611.0, 2016.

[8] Francisco Cary, Octavian Postolache, and Pedro Silva Girão. Kinect based system and Artificial Neural Networks classifiers for physiotherapy assessment. In *IEEE MeMeA 2014*

- *IEEE International Symposium on Medical Measurements and Applications, Proceedings*, 2014.

[9] Chien-Yen Chang, B Lange, Mi Zhang, S Koenig, P Requejo, N Somboon, A A Sawchuk, and A A Rizzo. Towards pervasive physical rehabilitation using Microsoft Kinect. In *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*, pages 159–162, 2012.

[10] Chia-Hsiung Chen, David P Azari, Yu Hen Hu, Mary J Lindstrom, Darryl Thelen, Thomas Y Yen, and Robert G Radwin. The accuracy of conventional 2D video for quantifying upper limb kinematics in repetitive motion occupational tasks. *Ergonomics*, 58(12):2057–2066, 12 2015.

[11] Chia-Hsiung Chen, Yu Hen Hu, Thomas Y Yen, and Robert G Radwin. Automated Video Exposure Assessment of Repetitive Hand Activity Level for a Load Transfer Task. *Human factors : the journal of the Human Factors and Ergonomics Society.*, 55(2):298–308, 2013.

[12] Yinpeng Chen, Margaret Duff, Nicole Lehrer, Hari Sundaram, Jiping He, Steven L. Wolf, Thanassis Rikakis, Tuan D. Pham, Xiaobo Zhou, Hiroshi Tanaka, Mayumi Oyama-Higa, Xiaoyi Jiang, Changming Sun, Jeanne Kowalski, and Xiuping Jia. A Computational Framework for Quantitative Evaluation of Movement during Rehabilitation. pages 317–326, 2011.

[13] Roy B Davis, Sylvia unpuu, Dennis Tyburski, and James R Gage. A gait analysis data collection and reduction technique. *Human Movement Science*, 10(5):575–587, 1991.

[14] Carlos Ferreira, Vnia Guimarães, Antnio Santos, and Ins Sousa. Gamification of Stroke Rehabilitation Exercises Using a Smartphone. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*, PervasiveHealth '14, pages 282–285, ICST, Brussels, Belgium, Belgium, 2014. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

[15] N Friedman, J B Rowe, D J Reinkensmeyer, and M Bachman. The Manumeter: A Wearable Device for Monitoring Daily Use of the Wrist and Fingers. *IEEE Journal of Biomedical and Health Informatics*, 18(6):1804–1812, 2014.

[16] T Giorgino, P Tormene, F Lorussi, D De Rossi, and S Quaglini. Sensor Evaluation for Wearable Strain Gauges in Neurological Rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 17(4):409–415, 2009.

[17] T Giorgino, P Tormene, G Maggioni, C Pistarini, and S Quaglini. Wireless Support to Poststroke Rehabilitation: MyHeart's Neurological Rehabilitation Concept. *IEEE Transactions on Information Technology in Biomedicine*, 13(6):1012–1018, 2009.

[18] Els Knippenberg, Jonas Verbrugghe, Ilse Lamers, Steven Palmaers, Annick Timmermans, and Annemie Spooren. Markerless motion capture systems as training device in neurological rehabilitation: a systematic review of their use, application, target population and efficacy. *Journal of NeuroEngineering and Rehabilitation*, 14(1):61, 2017.

[19] Gert Kwakkel, Boudewijn J Kollen, and Hermano I Krebs. Effects of Robot-Assisted Therapy on Upper Limb Recovery After Stroke: A Systematic Review. *Neurorehabilitation and Neural Repair*, 22(2):111–121, 9 2007.

[20] GyuChang Lee. Effects of Training Using Video Games on the Muscle Strength, Muscle Tone, and Activities of Daily Living of Chronic Stroke Patients. *Journal of Physical Therapy Science*, 25(5):595–597, 2013.

[21] S Lee, Y S Lee, and J Kim. Automated Evaluation of Upper-Limb Motor Function Impairment Using Fugl-Meyer Assessment. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(1):125–134, 2018.

[22] Daniel Leightley, Jamie S. McPhee, and Moi Hoon Yap. Automated Analysis and Quantification of Human Mobility Using a Depth Sensor. *IEEE Journal of Biomedical and Health Informatics*, 2017.

[23] Albert C Lo, Peter D Guarino, Lorie G Richards, Jodie K Haselkorn, George F Wittenberg, Daniel G Federman, Robert J Ringer, Todd H Wagner, Hermano I Krebs, Bruce T Volpe, Christopher T Bever, Dawn M Bravata, Pamela W Duncan, Barbara H Corn, Alysia D Maffucci, Stephen E Nadeau, Susan S Conroy, Janet M Powell, Grant D Huang, and Peter Peduzzi. Robot-Assisted Therapy for Long-Term Upper-Limb Impairment after Stroke. *New England Journal of Medicine*, 362(19):1772–1783, 5 2010.

[24] H J Luinge and P H Veltink. Measuring orientation of human body segments using miniature gyroscopes and accelerometers. *Medical and Biological Engineering and Computing*, 43(2):273–282, 2005.

[25] Peter S. Lum, Charles G. Burgar, Peggy C. Shor, Matra Majmundar, and Machiel Van der Loos. Robot-assisted movement training compared with conventional therapy techniques for the rehabilitation of upper-limb motor function after stroke. *Archives of Physical Medicine and Rehabilitation*, 83(7):952–959, 2 2002.

[26] Amir Mobini, Saeed Behzadipour, and Mahmoud Saadat Foumani. Accuracy of Kinects skeleton tracking for upper body rehabilitation applications. *Disability and Rehabilitation: Assistive Technology*, 9(4):344–352, 7 2014.

[27] Goldie Nejat, Yiyuan Sun, and Mary Nies. Assistive Robots in Health Care Settings. *Home Health Care Management & Practice*, 21(3):177–187, 7 2008.

[28] Rinske Nijland, Erwin van Wegen, Jeanine Verbunt, Renske Wijk, Joost Kordelaar, and Gert Kwakkel. *A comparison of two validated tests for upper limb function after stroke: The Wolf Motor Function Test and the Action Research Arm Test*, volume 42. 7 2010.

[29] Erienne V Olesh, Sergiy Yakovenko, and Valeriya Gritsenko. Automated Assessment of Upper Extremity Movement Impairment due to Stroke. *PLoS ONE*, 9(8):e104487, 2014.

[30] Shyamal Patel, Hyung Park, Paolo Bonato, Leighton Chan, and Mary Rodgers. A review of wearable sensors and systems with application in rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 9(1):21, 2012.

[31] James Patton, Steven L Smal, and William Zev Rymer. Functional Restoration for the Stroke Survivor: Informing the Efforts of Engineers. *Topics in stroke rehabilitation*, 15(6):521–541, 2008.

[32] John Platt. Fast Training of Support Vector Machines Using Sequential Minimal Optimization. MIT Press, 1 1998.

[33] David J Reinkensmeyer and Michael L Boninger. Technologies and combination therapies for enhancing movement training for people with a disability. *Journal of NeuroEngineering and Rehabilitation*, 9(1):17, 2012.

[34] R Riener, T Nef, and G Colombo. Robot-aided neurorehabilitation of the upper extremities. *Medical and Biological Engineering and Computing*, 43(1):2–10, 2005.

[35] Loren Arthur Schwarz, Diana Mateus, and Nassir Navab. Discriminative Human Full-body Pose Estimation from Wearable Inertial Sensor Data. In *Proceedings of the 2009 International Conference on Modelling the Physiological Human*, 3DPH'09, pages 159–172, Berlin, Heidelberg, 2009. Springer-Verlag.

[36] Nicolas Schweighofer, Cheol E Han, Steven L Wolf, Michael A Arbib, and Carolee J Winstein. A Functional Threshold for Long-Term Use of Hand and Arm Function Can Be Determined: Predictions From a Computational Model and Supporting Data From the Extremity Constraint-Induced Therapy Evaluation (EXCITE) Trial. *Physical Therapy*, 89(12):1327–1336, 12 2009.

[37] Rejin John Varghese, Daniel Freer, Fani Deligianni, Jindong Liu, and Guang-Zhong Yang. WEARABLE ROBOTICS FOR UPPER-LIMB REHABILITATION AND ASSISTANCE: A REVIEW OF THE STATE-OF-THE-ART, CHALLENGES AND FUTURE RESEARCH. Technical report.

[38] Janne M Veerbeek, Anneli C Langbroek-Amersfoort, Erwin E H van Wegen, Carel G M Meskers, and Gert Kwakkel. Effects of Robot-Assisted Therapy for the Upper Limb After Stroke: A Systematic Review and Meta-analysis. *Neurorehabilitation and Neural Repair*, 31(2):107–121, 9 2016.

[39] Vinay Venkataraman, Pavan Turaga, Michael Baran, Nicole Lehrer, Tingfang Du, Long Cheng, Thanassis Rikakis, and Steven L. Wolf. Component-level tuning of kinematic features from composite therapist impressions of movement quality. *IEEE Journal of Biomedical and Health Informatics*, 2016.

[40] Qi Wang, Panos Markopoulos, Bin Yu, Wei Chen, and Annick Timmermans. Interactive wearable systems for upper body rehabilitation: a systematic review. *Journal of NeuroEngineering and Rehabilitation*, 14(1):20, 2017.

[41] David Webster and Ozkan Celik. Systematic review of Kinect applications in elderly care and stroke rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 11(1):108, 2014.

[42] Jill Whitall, Douglas N. Savin, Michelle Harris-Love, and Sandra McCombe Waller. Psychometric Properties of a Modified Wolf Motor Function Test for People With Mild and Moderate Upper-Extremity Hemiparesis. *Archives of Physical Medicine and Rehabilitation*, 87(5):656–660, 5 2006.

[43] Steven L Wolf, Paul A Thompson, David M Morris, Dorian K Rose, Carolee J Winstein, Edward Taub, Carol Giuliani, and Sonya L Pearson. The EXCITE Trial: Attributes of the Wolf Motor Function Test in Patients with Subacute Stroke. *Neurorehabilitation and Neural Repair*, 19(3):194–205, 9 2005.

[44] Huiyu Zhou and Huosheng Hu. Inertial motion tracking of human arm movements in stroke rehabilitation. In *IEEE International Conference Mechatronics and Automation, 2005*, volume 3, pages 1306–1311, 2005.

[45] Huiyu Zhou and Huosheng Hu. Human motion tracking for rehabilitationA survey. *Biomedical Signal Processing and Control*, 3(1):1–18, 2008.

[46] Huiyu Zhou, Thomas Stone, Huosheng Hu, and Nigel Harris. Use of multiple wearable inertial sensors in upper limb motion tracking. *Medical Engineering and Physics*, 30(1):123–133, 2 2008.

Chapter 6

Appendix

Weight	Score
noseSpeed	-0.0009
neckSpeed	0.0011
shoulderSpeed	-0.0032
elbowSpeed	-0.0013
wristSpeed	0.0026
elbowAngle	-0.0001
nose Jerk	0.0002
neckJerk	-0.0002
shoulderJerk	-0.0063
elbowJerk	0.0024
wristJerk	0.0076
trialsize	0.9983
c	-0.0022
Correlation coefficient	0.9999
Mean absolute error	0.0029
Root mean squared error	0.0038
Relative absolute error	1.20%
Root relative squared error	1.09%

Figure 6.1: Drink Cup QAF Results

Weight	Score
noseSpeed	0.0052
neckSpeed	0.0005
shoulderSpeed	0.0057
elbowSpeed	0.0026
wristSpeed	0.0099
elbowAngle	0.0034
nose Jerk	-0.0082
neckJerk	0.0018
shoulderJerk	0.0024
elbowJerk	0.001
wristJerk	0.002
trialsize	0.9962
c	-0.0072
Correlation coefficient	0.9999
Mean absolute error	0.0039
Root mean squared error	0.0058
Relative absolute error	1.05%
Root relative squared error	1.34%

Figure 6.2: Lift Weight QAF Results

Weight	Score
noseSpeed	0.0012
neckSpeed	0.0019
shoulderSpeed	-0.0016
elbowSpeed	-0.0056
wristSpeed	-0.0147
elbowAngle	0.0297
nose Jerk	0
neckJerk	-0.0012
shoulderJerk	-0.0003
elbowJerk	-0.0031
wristJerk	-0.0032
trialsize	0.995
Correlation coefficient	0.9964
Mean absolute error	0.0289
Root mean squared error	0.0398
Relative absolute error	7.70%
Root relative squared erroı	9.18%

Figure 6.3: Forearm to Box QAF Results