# - USE OF KINECT V2 SENSOR IN UPPER EXTREMITY STROKE REHABILITATION -

By

KISHAK ZAKKA CINFWAT

GRADUATION REPORT

Submitted to
Hanze University of Applied Science Groningen

in partial fulfillment of the requirements
for the degree of

Fulltime Honours Master Sensor System Engineering

2016

ABSTRACT

USE OF KINECT V2 SENSOR IN UPPER EXTREMITY STROKE REHABILITATION

by

KISHAK ZAKKA CINFWAT

About 15 million suffer from strokes annually. Strokes are caused by cardiovascular problems, leading to paralysis or weakness of the extremities affecting patient independence thus requiring rehabilitation. Recently, virtual reality and other gaming systems have been developed and tested for functional rehabilitation and especially home use. This study examines the suitability of the Microsoft Kinect v2 in detecting Activities of Daily Living (ADL) within a hand function assessment procedure called the Southampton Hand Assessment Procedure (SHAP) using the adaptive boosting algorithm within the Kinect v2 SDK. Data from a group of 14 students was used. The results showed that the ADL tasks comprising; pouring from a carton, pouring from a jug, opening a glass jar, picking up coins, lifting a light tin and heavy jar across a barrier were detected with Root Mean Square confidence values below the minimum 0.95 recommended for deployment in an application or game. This implies further data would be required to train the detector for reliable detection. The use of mixed grasping strategies for some tasks may have affected the initial detection accuracy for the last two detectors. There is potential of the Kinect v2 in stroke rehabilitation and assessment.

DECLARATION

I hereby certify that this report constitutes my own product, that where the language of others is set forth, quotation marks so indicate, and that appropriate credit is given where I have used the language, ideas, expressions or writings of another.

I declare that the report describes original work that has not previously been presented for the award of any other degree of any institution.

Signed,

KISHAK ZAKKA CINFWAT

ACKNOWLEDGEMENTS

This endeavour would not have been possible but for many persons I have had the good fortune of meeting prior to and during the course of my studies at Hanze University of Applied Sciences. I would firstly like to thank the Dutch government for the StudNed scholarship that significantly reduced my tuition fees, without which the possibility of studying in the Netherlands would have been low.

I am equally grateful to my lecturers, thesis supervisors Corina Vogt and Ronald van Elburg, PhD. To my classmates, for the pleasant experience of studying together, my internship supervisors at the Centre for Human Movement Sciences, University Medical Centre Groningen (UMCG); Claudine Lamoth, PhD and Alessio Murgia, PhD for their support.

Similarly, I would have to specially acknowledge my dear wife, Gloria, I cannot thank you enough for holding the fort while I was away. And to my big boy - Mannminen, daddy will soon return from his wandering and hopefully stay a long time. To my parents; Mr. & Mrs. Zakka Cinfwat Dimka, I pray that God allows me the opportunity to honour you properly for your sacrifice through the years. To my siblings, the rest of the battalion, you guys are the best!

Certain individuals reminded me that family is made up of more than blood relations. For their encouragement and support, I would like to thank my colleagues; Engr. Oluwatosin Akinmade, Engr. Abubakar Yola, Mr. Peter Maschem, Engr. Emeso Ojo and many other friends like Mr. Andy Iza, Engr. Rotshak Gowon, Engr. Nathaniel Pali, Engr. Dauda Ishaya, Arc. Ibrahim Gado, Engr. Istifanus Lucas, Mr. Melvin Brown, Mr. Patrick Ndanema, Mr Chima Nwaorgu and Pastor Kennedy Mensah, you are all heaven sent.

I would also like to thank my friends, classmates and other students of the University of Groningen who took time to participate in my experiments. This work would not have been possible without you.

It is impossible to acknowledge everyone who helped along this stretch of the way. But, I am most grateful to God; who makes all things possible!

TABLE OF CONTENTS

**LIST OF TABLES**

**LIST OF FIGURES**

CHAPTER 1

# RATIONALE

This chapter introduces the causes of strokes, its global prevalence and highlights the role of rehabilitation as a key part of stroke management. It also examines current challenges in stroke rehabilitation. This leads to the research questions of the thesis project.

## 1.1 Introduction

According to the World Heart Federation (WHF), about 15 million people suffer from strokes annually with 5 million dying as a result, while 6 million end up with permanent disability globally (World Heart Federation 2016). The problem appears to be global, with prevalence of strokes spread across racial and geographic boundaries indicating a worldwide burden (Thrift et al. 2014). Although some researchers like Mackay et al. (2004) places Asian and black African populations as more predisposed to strokes, the global prevalence is significant.

In Europe and the USA, according to Vogiatzaki & Krukowski (2016), about 3% of the entire public health budget is expended on dealing with the effects of strokes. This is significant, considering the fact that there are other health concerns affecting the society and the fact that deaths from strokes are on the decline implying the number of stroke patients that survive is on the increase (Hughes et al. 2014; Thrift et al. 2014) thus, increasing the demand for specialized rehabilitation to regain full or partial independence.

## 1.2 Strokes

Strokes are caused by cessation of blood supply to the brain due to the narrowing, blockage or rupture of a blood vessel in the brain causing cerebral hypoxia (Waugh & Grant 2014). Sometimes, spontaneous intracranial haemorrhaging due to prolonged hypertension or other cardiovascular problems are also identified as causes of strokes (Waugh & Grant 2014).

Severe haemorrhaging often leads to death due to the increased Intra Cranial Pressure (ICP) which builds up from the lost blood within the cranium thereby damaging the brain cells further by exerting pressure on brain within the cranial cavity. The less severe variety, with less ICP often causes paralysis, loss of sensitivity, loss of speech, loss of vision and other symptoms that are often associated with strokes (Waugh & Grant 2014). There are also minor strokes called Transient Ischemic

**Met opmaak:** Inspringing: Eerste regel: 1,27 cm

Attacks (TIA) often caused by short cessation of blood flow to the brain but resolves within 24 hours. TIAs often signal a stroke or indicates a cardiovascular problem (Waugh & Grant 2014).

The typical physical presentation of strokes includes various forms of weakness, numbness, stiffness and paralysis of various types, ranging from dysphagia; affecting swallowing, general paralysis or hemiplegia; one-sided paralysis or hemiparesis; muscle weakness of one part of the body. Other presentations include foot drop, incontinence, seizures or epilepsy, spasticity etc. (Stroke 2016). There are other emotional and cognitive issues that occur after a stroke but the emphasis of this thesis and most physical rehabilitation efforts are on restoring limb function, which allows for re-integration into society by increasing functional independence.

### 1.3 Stroke Management

Strong et al (2007) posit that the reduction in mortality from strokes has led to an increase in persons requiring rehabilitation and admittance into long term care facilities. One of the major problems with strokes is the disability that often affects the survivor. This disability often creates further problems due to the loss of independence for the survivor coupled with the loss of self esteem, productivity loss in the working class stroke sufferer. This category of people would require some sort of continuous or routine care-giving depending on the severity of the stroke and disability.

According to Adams Jr. (2016), the management of a stroke from the onset of an attack involves emergency services where the victim is taken to a hospital where the suspicion of a stroke is confirmed, the cause of the stroke is evaluated through advanced imaging technologies and then treatments which could include surgery, aimed at controlling medical and neurological complications of the stroke are commenced. Thereafter, treatment therapies to prevent reoccurrences are initiated and after patient stabilization, rehabilitation activities are commenced to restore lost functionality (Vogiatzaki & Krukowski 2016; Wu et al. 2012). These rehabilitation activities comprise of disability assessment, goal setting, intervention or rehabilitation and reassessment in a cyclic fashion (Langhorne et al., 2011) as shown in figure 1.1 below.

**Met opmaak:** Inspringing: Eerste regel: 1,27 cm

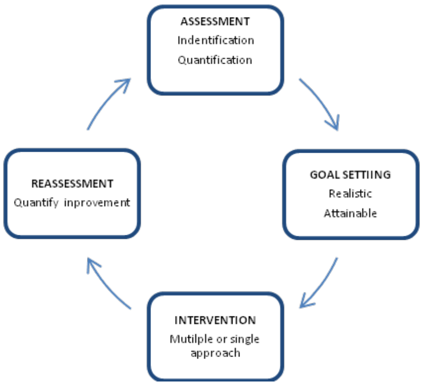


Figure 1.1 Stroke rehabilitation process

*1.3.1 Stroke rehabilitation*

In this report, the rehabilitation of motor capabilities is the primary interest. Stroke severity is often classified as mild, moderate and severe. This classification is based on the domains of neurologic impairment; the domains include; motor, sensory, vision, language, cognition and affect. While the severity of impairment is classified into 3 levels (Kelly-Hayes et al. 1998) and shown in the table 1.1 below:

*Table 1.1 Stroke Severity classifications*

| Severity of Impairment | |
|---|---|
| Level A: | Minimal or no neurological deficit due to stroke in the above domains. |
| Level B: | Mild/moderate deficit due to stroke in ≥ 1 domain. |
| Level C: | Severe deficit due to stroke in ≥ 1 domain. |

Source: Kelly-Hayes et al. (1998)

Strokes often disrupt the performance of Activities of daily Living (ADL) and the in ability to function is another means of stroke severity classification (Kelly-Hayes et al 1998). All patients presenting these varying levels of stroke impairment benefit from rehabilitation (Teasell & Hussein 2016; Nordin et al. 2012).

There is a difference in the duration of rehabilitation required to achieve a significant restoration of motor function and this is related to the level of severity of the stroke. According to Teasell & Hussein (2016), prior studies indicated that the intensity of the rehabilitation routines, in terms of time spent in therapy, showed benefit en route function restoration.

The success outcomes of rehabilitation are often measured by the restoration of functional independence, strength, sensory discrimination, and fine motor skills (Byl et al 2008). These

**Met opmaak:** Inspringing: Eerste regel: 1,27 cm

are the motor and sensory faculties that are affected by strokes and to achieve a restoration of the stroke survivor’s independence, the same must be significantly restored.

## 1.4 Activities of Daily Living (ADL)

According to the World Confederation for Physical Therapy (WCPT 2014, p.4) “Activities of daily living (ADL) — are the daily self-care activities required to function in the home and/or outdoor environment. They may be classified as basic or instrumental.” The basic ADL include; dressing, eating, mobility, toileting and hygiene while the instrumental Activities of Daily Living (IADL) include; shopping, housekeeping, managing finances, preparing meals and using transport (Kelly-Hayes et al. 1998). The basic ADL are shown in figure 1.2 below.

The loss of motor and other functions due to strokes often affects the individual’s ability to carry out activities of daily living (ADL). This leads to a significant dependence on others to function within society and often to survive (Loue & Sajatovic 2008). The aim of rehabilitation is often the restoration of the capability to carry out some, if not all of these ADL thereby significantly reducing the level of dependence of the stroke survivor on others for survival and as emphasized by Sveen et al. (2004), where the self facilitated participation in leisure activities appear to be a significant indicator of well-being after a stroke.

**Met opmaak:** Inspringing: Eerste regel: 1,27 cm

Figure 1.2 Basic Activities of Daily Living

ADL are typical areas of concern when determining disability from strokes and even other accidents involving trauma to the muscle, skeleton or nervous system. Disability tests, like the Wolf Motor Function Tests (WMFT) are often used determine the level of impairment in stroke patients and subsequent improvements from therapy can be tracked by repeated administration of the same test (Vogiatzaki & Krukowski 2016). Reaching and grasping impairments are often key areas of focus in rehabilitation because they impact significantly on self sustenance or ADL like eating, drinking, bathing and even movement in paraplegia.

### *1.4.1 Reaching and grasping in ADL*~~:~~

Reaching and grasping are key motor skills required for ADL involving the upper extremity (Levin et al 2015). It is said that motor control is made up of gross motor and fine motor skills. The detection of gross motor skills is significantly easier than that of fine motor skills. The objective of functional upper extremity rehabilitation often focuses on the restoration of both gross and fine motor skills which are useful for carrying out basic ADL not strength and range of motion (Thrasher et al 2008). Kim & Kim (2015) also highlight the importance of the hands in ADL.

## 1.5 Main Rehabilitation Approaches:

In the past, stroke rehabilitation emphasized strengthening, regaining of flexibility and motion of affected limbs via physical exercises. Dobkin (2009) indicated that even certain robot assisted therapies only succeeded in strengthening proximal muscles without functional improvements in the upper extremity for ADL. However, recent approaches are goal oriented i.e. targeted at achieving an ADL (Hochstenbach-Waelen & Seelen 2012). Here task specific exercise routines aimed at restoring specific skills are performed by the patient.

Rehabilitation motions and actions required to achieve a particular ADL are considered as functional rehabilitation. In rehabilitation efficacy assessment, performance of ADL is used to determine the effectiveness of rehabilitation therapy (Loue & Sajatovic 2008). Functional rehabilitation or task oriented rehabilitation approaches are sometimes augmented with functional electrical stimulation (FES) as reported by Freeman et al (2015) where low energy signals are applied to the limb being exercised in a bid to stimulate the nerves at the affected limb while the exercise routines are implemented.

Conclusions drawn from rehabilitation studies indicate that repetition, intensity and task oriented therapy are often beneficial to the patient undergoing therapy (Standen et al. 2015; Alankus et al. 2010). Also, studies have shown that a significant challenge in out-patient therapy continuation has been the sustenance of motivation and morale in the person undergoing rehabilitation (Hocine et al. 2015; Standen et al. 2015). According to Alankus et al. (2010) only about 31% of stroke patients accurately perform recommended therapeutic routines in out-patient scenarios. Attempts at solving these and the personnel shortage can be achieved through the deployment of assistive technology or automation of the rehabilitation therapy activities. Saposnik (2015) highlighted the challenges of conventional stroke therapy to include: Time and labour intensive procedures, restricted availability, adherence defaults, overall modest benefit, high cost and inconvenient travel requirements.

## 1.6 Technology in stroke rehabilitation

Rehabilitation of stroke patients is usually the task of physiotherapists and other specialists. The classic routines of physical and occupational therapy are labour intensive and time consuming. According to Zhou & Hu (2008), research into human motion capture with a view to applications in rehabilitation has been ongoing since the 1980's. In recent years, technological advances in pervasive computing have been applied to reduce the personnel requirements of rehabilitation through automation of processes like the use of rehabilitation robots (Loureiro et al. 2011), employing serious computer gaming together with other e-health paradigms to facilitate therapy and also track patient improvement remotely (Vogiatzaki & Krukowski 2016).

The recent technological advances in rehabilitation, especially with assistive technologies, apart from reducing or eliminating personnel demand, are systems capable of eliciting morale and motivation in the patient. Several of such systems have been developed and tested (Hocine et al. 2015). The application areas of these assistive technologies span; assessment, goal setting and therapy or intervention and then reassessment in a cyclic process (Langhorne et al., 2011). Sometimes a cocktail or multiple approaches are deployed at the same time for example robot assisted movement with Functional Electrical Stimulation (FES). Some possible pathways are shown in figure 1.3 below:

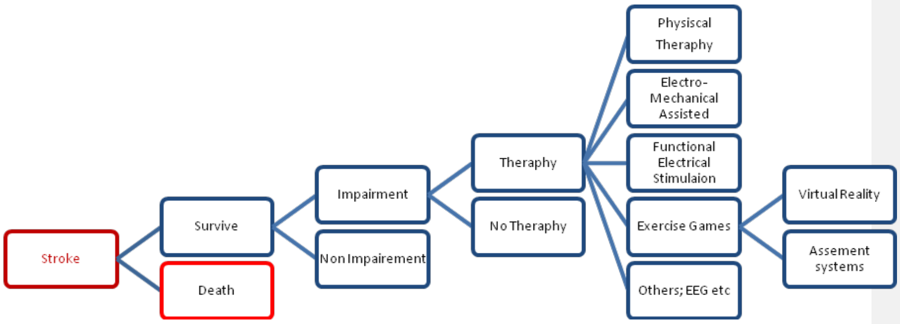


Figure 1.3 Pathways from Stroke to rehabilitation using assistive technologies

The modern rehabilitation systems are expected to be low cost when compared to out-patient therapy and therefore suitable for out-patient home use. The non-clinical setting application of therapy, according to Vogiatzaki & Krukowski (2016) is said to have facilitated quicker patient response to the therapy and recovery because of the familiar environment and perceived comfort of the home while the accuracy and frequency of execution of the prescribed routine can be monitored and recorded.

Despite the potential of these assistive technologies, their adoption for clinical and out-patient use has been slow. A mix of awareness, education, cost reduction and improvement in system design were presented as likely approaches to improve adoption of these assistive rehabilitation technologies (Hughes et al. 2014). According to Zhou & Hu (2008), the development of recovery or rehabilitation systems, six issues need to be taken into account: cost, size, weight, function, operation, and automation. This makes the selection of the sensor system required for rehabilitation a non-trivial task given that the cost, size, weight, function, operation and automation could be severely impacted by the sensor choice.

## 1.7 ATARGET

This thesis is based on an internship at the University Medical Centre, Groningen (UMCG) under the Adaptive Trainer using Augmented Reality Gaming for Exercise Therapy (ATARGET) project. The research project focuses on the use of Kinect v2 for the capture of upper extremity movements, analysing the sensor data and coupling to an augmented reality app as part of the development of an exercise gaming system. This system, when developed would aid in functional rehabilitation of stroke patients, athletes and others recovering from major injuries by being deployed as a Virtual Reality (VR) rehabilitation game.

A functional rehabilitation system, capable of measuring the functional movement of a patient undergoing, comparing against other records while facilitating the therapy would be a useful addition to the arsenal of clinicians involved in rehabilitation or physiotherapy which has become a significant healthcare requirement due to pathologies like strokes which are common in the large, global ageing population and in younger adults due to lifestyle choices or other unknown factors.

## 1.8 Stakeholders

Rehabilitation is often required for all survivors of strokes. This therapy often takes place in clinical settings and sometimes at home; in an out-patient setting, with familiar surroundings and surrounded by friends and family. Some of the immediate stakeholders involved in the project are presented in table 1.2 below.

*Table 1.2 Stakeholders of the ATARGET project at UMCG*

| | |
|---|---|
| UMCG | The UMCG provides the materials for this project and is the location where the stroke management often commences; from patient stabilization to rehabilitation. |
| Centre for Human Movement Sciences | Responsible for studies and research activities related to human motion and rehabilitation. The supervisors of the project are from the research centre. |
| Hanze University of Applied Sciences | Providing academic supervision and are therefore stakeholders. The supervisors are expected to provide coaching and necessary support where needed throughout this project. The assessors for this project are Corina Vogt and Ronald van Elburg. |
| Patients | Patients in this study comprises of individuals recovering from strokes and in need of upper extremity rehabilitation especially in an out-patient scenario. Others who could benefit from the systems could patients recovering from other upper extremity injuries. |

## 1.9 Research Question

In the light of the ATARGET project, an attempt to explore the Kinect v2 within stroke rehabilitation, a research question was developed:

*How well can the Kinect v2 detect Activity of Daily Living (ADL) motions when used in an upper extremity rehabilitation gaming or impairment assessment environment?*

- *How ADL related upper extremity movements are detected from Kinect v2 sensor data?*
- *How reliable will the results be?*

It is expected that the ability of the Kinect v2 to accurately detect the perfromance of ADL would aid in the automation of various aspects of rehabilitation as exercise games, functional assessment systems etc.

CHAPTER 2

**SITUATIONAL AND THEORETICAL ANALYSIS**

This chapter discusses the suitability of the Kinect sensors for rehabilitation systems especially considering the power of the Natural User Interface for health games, the power of Virtual Reality in rehabilitation, and the AI methods employed in systems involving the Kinect for activity detection, with the newest Kinect, Kinect v2 expected to enable the improved detection of both fine and gross motor skills.

## 2.1 Health Gaming in Stroke Rehabilitation

In recent years, significant efforts have been made to apply computer games, e-health and sensors to achieve less direct clinician intervention, reduce costs, sustain motivation, track progress and even dynamically alter the therapy using such systems (Vogiatzaki & Krukowski 2016; Freeman et al. 2015). Laver et al. (2013) published a review of tele-rehabilitation systems which employed information and communication technologies to facilitate remote rehabilitation monitoring.

Several rehabilitative gaming technologies have been applied in the past few years. These range from robot mediated gaming approaches (Freeman et al. 2015; Loureiro et al. 2011), Virtual reality based gaming systems (Standen et al. 2015), and augmented/virtual reality based games (Vogiatzaki & Krukowski 2016; Yeh et al. 2013) with adaptive attributes and sometimes dynamically difficulty adjustment (Hocine et al 2015).

### *2.1.1 Virtual reality in rehabilitation*

Games or applications considered as Virtual Reality (VR) games or applications, cover approaches to gaming or interaction within an artificial or surrogate 3D environment (Ojha, 1994). These games have been applied to rehabilitation and particularly to the rehabilitation of upper extremity conditions (Saposnik 2015; da Silva Cameirão et al 2011).

Levin et al. (2015) highlighted the use of Virtual Reality (VR) for post stroke recovery from upper limb disabilities such as arm paresis, citing the understanding that improvements may be achieved through sensory-motor learning and adaptive plasticity of the brain through the repetition of a variety of ADL related tasks in a visually rich training environment that incorporates cognitive challenges and VR is one method of creating such a training environment.

Similarly, according to several authors e.g. (Gatica-Rojas & Méndez-Rebolledo 2014; Lucca 2009), the choice of virtual reality environment for rehabilitation relies on the understanding that the motor imagery of the individual performing the rehabilitation task is capable of aiding the functional rearrangement of the damaged motor cortex i.e. the creation of alternate pathways for information and signal transfer within the brain

A simplified VR system is made up of input processors, simulation processors and rendering processors which interact with a world database. The rendering processor; which is the output stage is capable of producing haptic and other forms of feedback as shown in figure 2.1 below. It is important to note that the role of the database is integral to gesture recognition which is a key element of any software architecture of the virtual reality system (Maleshkov & Chotrov 2013). The gesture database references the specific objective of the interaction within the virtual or augmented environment. As shown by Yang & Kim (2002), performance evaluation, motion data processing, performance simulation, guidance and object control rely on a database which often includes gestures in a human immersive game environment.

**Met opmaak:** Inspringing: Eerste regel: 1,27 cm

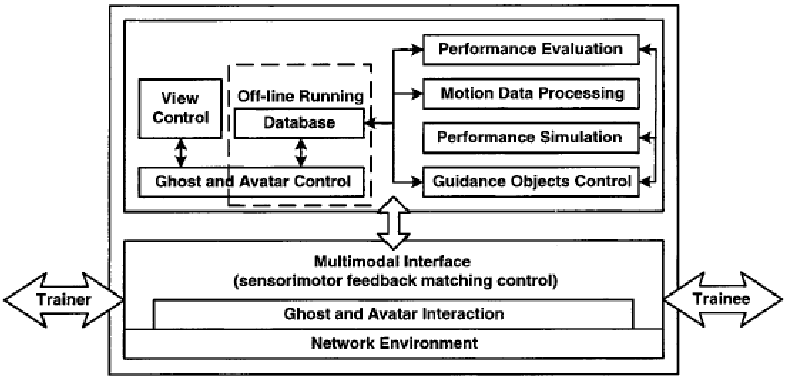


Figure 2.1 A possible VR training system architecture (Yang & Kim 2002)

Such VR or augmented reality rehabilitation systems are interactive, capable of providing audio, visual and haptic feedback while measuring the level of motor impairment and recovery as the system is used. The possibility of varying the environment could reduce boredom and facilitate motivation within the rehabilitation process (Morel et al 2015).

The older VR systems employed gloves embedded with sensors or some visual tracking on the gloves like colour bands to indicate the wrist and elbows in order to facilitate tracking of the patients' arms during exercise. With the advent of the Kinect with its skeletal tracking, the need for 'cyber' gloves has been eliminated for the Kinect VR rehabilitation systems e.g. SeeMe (Brontes Processing, 2014) which offers a virtual reality gaming system based on the Kinect V1 or V2 but focuses on range of motion, strength, movement quality, movement awareness and proprioception which could be of interest to upper extremity rehabilitation in stroke patients.

## 2.2 Sensors in Gaming and Rehabilitation.

To achieve accurate patient and object tracking in these e-health rehabilitation games, several sensors have been developed or adapted for the capture of the desired motion data of the affected limb. This data is needed for rehabilitation, assessment and monitoring (Mousavi et al 2014). Typically, the motion capture data supplied by the sensor facilitates computation for adaptive capabilities of the games.

Several sensor systems have been used in rehabilitation. These include accelerometers and gyroscopes embedded in a bracelet (Wu et al. 2014). Others like Hortal et al. (2013) discussed the use of electroencephalographic (EEG) electrode data to decode the movement velocity of the human upper limb to be applied in a future rehabilitation system. These examples show that several sensor technologies are being explored for their potential application in rehabilitation.

**Met opmaak:** Inspringing: Eerste regel: 1,27 cm

Camera systems for marker and marker-less motion tracking have been applied for stroke rehabilitation gaming and Alankus et al. (2010) describe a system employing the Nintendo Wii made by Sony Corporation of Japan and another system employing an ordinary web camera for use in rehabilitation games. Paquin (2014) also analysed a commercial Nintendo Wii based systems for fine motor recovery in chronic stroke in community level rehabilitation. These together with Kinect based systems provide an overview of the range of sensors deployed in rehabilitation and rehabilitation games.

### 2.3 Natural User Interfaces (NUIs) and Motion Capture

According to Liu (2010, p.1) Natural User Interface (NUI) "*...is an emerging computer interaction methodology which focuses on human abilities such as touch, vision, voice, motion and higher cognitive functions such as expression, perception and recall*"

In order to develop an electronic exercise gaming system or rehabilitation monitoring system, there is need to capture human motion data (Schwarz et al. 2012). Recent advances in sensors have led to the use of cameras and other marker-less systems which eliminate the need for cumbersome markers or suits. These new systems employ sensors that are non-intrusive on the user of the gaming system fitting the NUI paradigm.

One of the most common marker-less sensors in current use is the Microsoft Kinect. van Teijlingen et al. (2012) documented the Kinect sensor's rise to prominence among game application developers, researchers and hobbyist. They also highlighted the key advantages to include lower sensor cost, _unobtrusiveness, robustness despite lower fidelity when compared to industry gold standards for motion capture and activity detection.

### 2.4 Kinect Sensor

The Microsoft Kinect is a motion sensing camera initially designed as an accessory to the Microsoft gaming system; Xbox (Microsoft, 2015). The Kinect sensor as shown in figure 2.2 is equipped with an RGB camera, a depth sensing capability and a microphone array. In July, 2014 Kinect version 2 with a new System Development Kit (SDK) was released to the public. It boasted of significant improvements on the capabilities of the Kinect v1 and the existing SDKs.

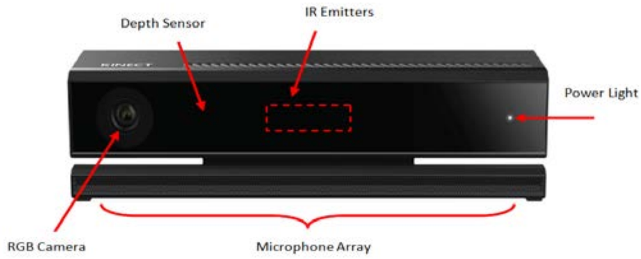


Figure 2.2 Front view of the Kinect v2 sensor.

The RGB camera and depth sensors together enable full-body 3-Dimension motion capture capabilities through skeletal tracking and thereby facilitate gesture recognition with further processing (Chang, Chen & Huang 2011). Schwarz et al. (2012) refer to the depth camera in the Kinect v1 as a Time of Flight (ToF) camera, which were hitherto expensive pieces of equipment but useful for various computer vision and image processing tasks.

Alabbasi et al. (2015) lists the major areas of Kinect sensor applications to include; retail marketing, healthcare, 3D-modeling and reconstruction, education, sign language recognition, robotics, control and natural language interface.



The Kinect sensor had already been adjudged to be sufficient for field use in ergonomic assessment of workplaces by Dutta (2012) when compared to a Vicon motion capture system which was used in the study as the gold standard. The root-mean-squared errors (SD) of 3-D markers were 0.0065 m (0.0048 m), 0.0109 m (0.0059 m), 0.0057 m (0.0042 m) in the *x*, *y* and *z* directions respectively. Similarly, Clark et al. (2012) showed the Kinect as having an acceptable validity as compared to a 3D camera employed as a reference motion capture device for postural assessment where they found the inter class correlation (ICC) difference to be 0.06 ± 0.05; range, 0.00 - 0.16.

Zhou et al. (2014) also showed that the Kinect is capable of being used to detect postures under scenarios where parts of the body are occluded and because of its depth camera system which based on Infrared emitted light, can work effectively under low illumination conditions.

*2.5.1 Skeletal tracking: Kinect v1 Vs Kinect v2*

One of the significant improvements in the Kinect v2 is the anatomically improved skeleton rendering compared to Kinect v1. This improves the sensor viability for clinical application and the inclusion of the hand tip and thumb joints which qualify the hand state of the skeleton in view. A comparison of the Kinect v1 and Kinect v2 rendered skeletons are shown in figure 2.3 below.

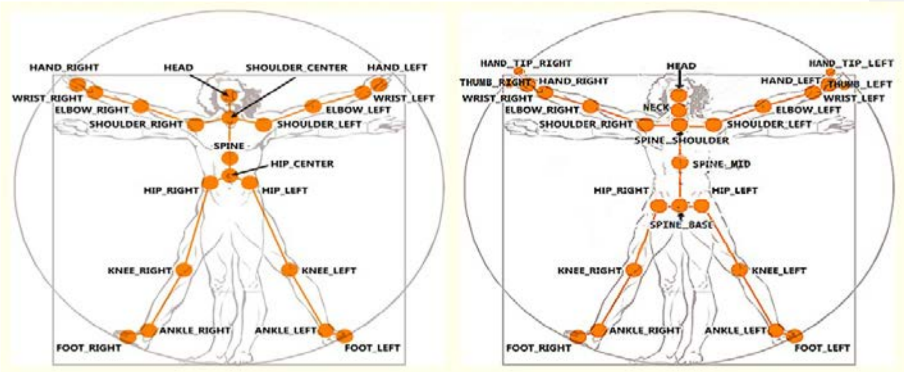


Figure 2.3 Kinect v1 and Kinect v2 skeleton positions relative to the human body (Valoriani & Giorio 2015)

The Kinect v2 has several data capture advantages over the Kinect V1 and these include an anatomically correct skeleton and the inclusion of hand tip and thumb joints which allow for the detection of more fine motor actions (Valoriani & Giorio 2015) than was possible with the Kinect v1. Apparently, the development of fine motor rehabilitation systems lags behind gross motor systems. Loureiro et al (2011) indicate that comparatively less work has been done on fine motor rehabilitation devices ans systems generally. In previous studies, comparing the Kinect v1 and other industrial skeletal tracking systems, the Kinect performed as well as industrial grade systems for gross motor detection with the main weakness being the detection of fine motor activity (Galna et. al 2014).

The Kinect v2 device improvements hold a significant promise for new and improved Kinect sensor applications in health, fitness, education and training, entertainment, gaming, movies, and communications (Alabbasi et al. 2015). The key improvements in the Kinect v2 sensor are summarized in table 2.1 below.

**Met opmaak:** Inspringing: Eerste regel: 1,27 cm

*Table 2.1 Key improvements in Kinect version 2.*

| Feature | Benefits |
|---|---|
| *Improved body tracking* | The enhanced fidelity of the depth camera, combined with improvements in the software, has led to a number of body tracking developments. The latest sensor tracks as many as six complete skeletons (compared to two with the original sensor), and 25 joints per person (compared to 20 joints with the Kinect v1 sensor). The tracked positions are more anatomically correct and stable and the range of tracking is broader. |
| *Depth sensing<br>512 x 424<br>30 Hz<br>FOV: 70 x 60<br>One mode: .5–4.5 meters* | With higher depth fidelity and a significantly improved noise floor, the sensor gives you improved 3D visualization, improved ability to see smaller objects and all objects more clearly, and improves the stability of body tracking. |
| *1080p color camera<br>30 Hz (15 Hz in low light)* | The color camera captures full, beautiful 1080p video that can be displayed in the same resolution as the viewing screen, allowing for a broad range of powerful scenarios. In addition to improving video communications and video analytics applications, this provides a stable input on which to build high quality, interactive applications. |
| *New active infrared (IR) capabilities<br>512 x 424<br>30 Hz* | Sensor has nighttime view capability, the new IR capabilities produce a lighting-independent view—and you can now use IR and color at the same time. |
| *Wider/expanded field of view* | A larger area of a scene to be captured by the camera. As a result, users can be closer to the camera and still in view, and the camera is effective over a larger total area. |

Source: Microsoft (2015)

A significant number of rehabilitation systems have the Kinect sensor as part of the architecture. Some of the Kinect systems were for physical rehabilitation e.g. (van Diest et al. 2013; Chang, Chen & Huang 2011) or neurological condition assessment, not necessarily related to strokes (Galna et al. 2014) but a significant number of studies were related to stroke conditions as presented in a review by Webster & Celik (2014) with over two dozen papers reviewed for Kinect based studies related to strokes.

The Kinect has been described as low cost sensor device, portable and simple to set up (Huang, Cheng, & Chiang 2013; Clark et al. 2012) thus, suitable for deployment in multiple applications within e-health paradigm. The problems of illumination change, background clutter, partial or full occlusion highlighted as a challenge of vision based hand gesture recognition systems highlighted by Zhu, Yang, & Yuan (2013) are mitigated by the Kinect sensor design.

**Met opmaak:** Inspringing: Eerste regel: 1,27 cm

2.6 Gesture Recognition from Kinect Sensor Data

The Kinect sensor provides depth sensing data used in developing exercise games, such applications draw strength in the developments of Natural User Interface (NUI) but the successful deployment of such systems relies on the robustness and execution speed of gesture recognition (Miranda et al. 2012) i.e. the realisation of real-time gesture recognition.

Several methods have been explored to realize gesture, posture or activity recognition as the case might be. To achieve gesture recognition with execution time variability, specified time constraint approaches were applied. Huang et al (2013) used Dynamic Time Wrapping (DTW) in a Kinect based dance assessment system to align timing and length differences between dance sequences.

**Met opmaak:** Inspringing: Eerste regel: 1,27 cm

For the actual gesture or activity recognition, Cottone et al (2014) applied a series of machine learning techniques first, by clustering the obtained joint data, then classifying the data employing a multiclass Support Vector Machine (SVM) then each activity was modeled as a sequence of known postures employing Hidden Markov Models (HMM). Earlier, Miranda (2012) had implemented a Kinect gesture recognition system employing a multi-class classifier Support Vector Machine and decision forests.

To achieve these results, several algorithms had to be developed and the skeletal joint data extracted and transformed to achieve recognition. This approach, though successful suffers from immediate applicability and would require advanced technical training to extract the skeletal joint data, develop and apply detection algorithms.

**Met opmaak:** Inspringing: Eerste regel: 1,27 cm

In the Kinect v2 SDK, a combination of heuristic recognition and machine learning methods are used in the 'Visual Gesture Builder' Application Program Interface (API) for gesture detection. The machine learning methods are ensemble methods namely; random forests and adaptive boosting (adaboost). According to Hastie et al. (2005) ensemble learning is based on building a prediction model based on the collective ability of simpler models. This significantly reduces the skill set required to develop gesture recognition.

In the Kinect v2 SDK, the ensemble methods of machine learning offered are found to be very useful, with adaboost determining if a gesture is being performed or not, while random forest determines the progress of a gesture being performed (Relyea & Marien 2013).

*2.6.1 Random Forests*

Random forests are based on decision trees. It is also an ensemble classification method. Given a collection of data and from the data, some subset is randomly selected with or without replacement. These subsets of data are employed to create trees. These collection of trees now make the forest. The randomly generated trees include some error but are capable of correctly predicting some events if placed in a committee (Hastie et al. 2005).

If given a training set X = {$x_1$, $x_2$, $x_{3....}x_n$} with responses Y = {$y_1$, $y_2$, $y_3$, . . ., $y_n$}, bagging implies a selection M times of a random sample with replacement from the given data set and fitting the trees to these samples.

For m = 1, 2, . . . , M;

1. Sample, with replacement, *n* training examples from *X*, *Y*; call these $X_m$, $Y_m$.
2. Train a decision or regression tree $f_m$ on $X_m$, $Y_m$.

After using the samples for training, predictions for unseen samples e.g. *x'* are made by averaging the predictions from all the individual regression trees from equation 2.1 below. Therefore, the majority vote from the tree generated from the random data results in a robust prediction of the dependent variable.

$$\hat{f} = \frac{1}{M}\sum_{m=1}^{M} \widehat{f_m}\,(\acute{x}) \qquad 2.1$$

*2.6.2 Adaptive Boosting (Adaboost)*

The idea behind boosting is based on the concept of 'weak classifiers' i.e. classifiers that are just slightly better than an unbiased coin toss for a two class problem with output Y as equation 2.2.

$$Y \in \{-1, 1\} \qquad 2.2$$

The weighted majority vote from the predictions of the weak classifiers results in a powerful predictor (Wang & Wang 2007; Hastie et al. 2005) as represented by equation 2.3 below.

$$G(x) = sign\left(\sum_{m=1}^{M} \alpha_m \; G_m(x)\right) \qquad 2.3$$

With α1, α2,…, αm calculated by theboosting algorithm and weighting the contribution of every $G_m(x)$ respectively. These weak classifiers and inaccurate rules sum up to become a highly robust classifier or prediction rule (Schapire 2013). The weak classfiers can be generated from neural networks, decision trees or any other classifiers for that matter. The adaboost is considered to be fast, simple to program, with versatile applications but it is also susceptible to uniform noise and could result in over fitting (Relyea & Marien 2013).

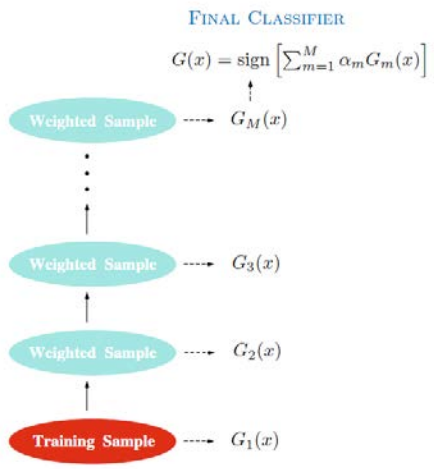


Figure 2.4 Schematic for Adaptive boosting (Hastie et al. 2005)

2.7 Features of Interest in Upper Extremity Rehabilitation

The challenge of detecting upper extremity motions using the Kinect required the use of body stream data particularly the skeletal tracking. The features of interest are the joint data for the upper extremity from the hip upwards. These are recorded and modeled as a skeleton on the depth data captured within a temporal sequence.

Specifically, the hand-tip, thumb, wrist, elbow and shoulder joints would sufficiently capture the gross and fine motor movements required for the performance of ADL. The Kinect v2 SDK already detects some inbuilt hand states like the open hand, closed hand or lasso which is represented by the closed hand with both the middle and index fingers pointing upward (Microsoft 2015).

CHAPTER 3

## CONCEPTUAL MODEL

This chapter presents the use of the Kinect sensor as a NUI for the marker-less capture of ADL data that could be used for upper extremity rehabilitation. The use of the Kinect v2 in developing gesture detection database based on hand function tests and common hand grasp patterns are discussed. Next, the role of the Kinect v2 Visual Gesture Builder API to facilitate the gesture detection is presented.

### 3.1 Kinect for Functional Rehabilitation

There have been a significant number of studies employing the Kinect as a sensor for stroke rehabilitation as reviewed by Webster & Celik (2014). However, there appears to be a shortage of recent reports based on the Kinect v2 although most of the papers do not specifically state the specific sensor version types, date of publication of the papers prior to 2014 often ruled out the use of the most recent Kinect sensor.

The promise of the Kinect v2 appears bright due to the enhanced skeletal data available about the hands, particularly the thumb and tip of the fingers. These additional skeletal data points would be useful in detecting fine motor skills often required for ADL. A particular interest will be placed on reaching and grasping because they embody a significant portion of the gross and fine motor skill respectively and all required in achieving basic ADL.

### 3.2 Functional Rehabilitation and ADL

The goal of functional rehabilitation is the restoration of basic ADL skills (WCPT, 2014) through physical exercise, gaming and other assistive technologies. Rehabilitation is made possible through gaming and other assistive technologies by the NUI the Kinect presents, the variety and the fidelity of the Kinect data streams. The ability to accurately determine therapeutic but functional motion prescribed would be beneficial in many respects.

An ADL detector could serve as evaluation data for clinicians to track the progress of a patient and the efficacy of prescribed routines in a simple e-health rehabilitation scenario by also employing the microphone array and the RGB camera over a network, connecting the clinician and patient in real-time.

**Met opmaak:** Inspringing: Eerste regel: 1,27 cm

The ADL database could be used within the gaming frame work to develop several interaction and adaptation strategies which truly makes the games automatically adjust difficulty to facilitate motivation.

The accurate detection of the motions commonly associated with ADL namely; reaching, grasping, transportation and releasing (Podobnik et al 2009; Štrbac et al. 2014; Barsotti et al. 2015), would be valuable for functional rehabilitation using assistive technologies.

### 3.3 Hand Function Tests

As discussed earlier, Langhorne et al (2011) highlighted the activities stroke management to include assessment, goal setting, intervention or rehabilitation and reassessment in a continuous fashion. The assessments are usually carried out in the form of hand or arm function tests. For example, Paquin (2014) evaluated stroke patients after a sustained use commercial video gaming on fine motor recovery in chronic stroke the tests used to test improvement were the Jebsen Hand Function Test (JHFT), Box and Blocks Test (BBT), Nine Hole Peg Test (NHPT), Stroke Impact Scale (SIS). All these tests, like the Wolf Motor Function Test (WMFT) involve reaching, grasping and transferring activities which are useful in determining level of motor impairment.

Concerning grasping, which is a key requirement for fine motor function, the common natural patterns include tip, tripod, power, lateral, spherical, extension etc (Light et al 2002) these grasp patterns as shown in Figure 3.1 below are important to the achievement of ADL.

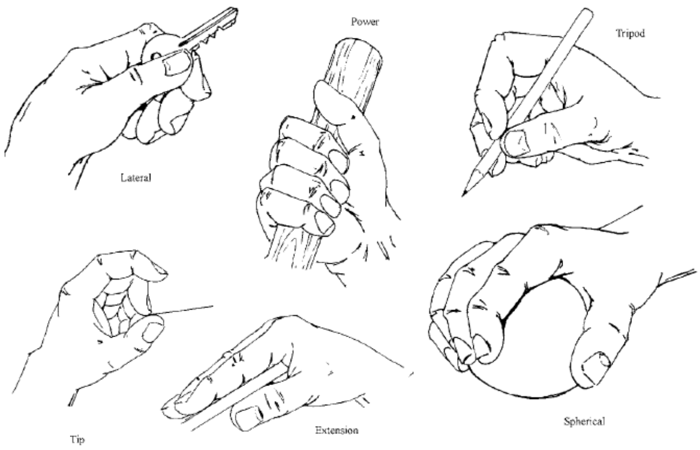


Figure 3.1 Hand grasping classification *(Light et al 2002)*

Furthermore, the table below associates some ADL with their natural or common grip classification (Kyberd et al 2009). These highlight the importance of reaching and grasping strategies for the performance of ADL.

*Table 3.1 Grip classifications for some common ADL tasks*

| NO. | TASK | NATURAL GRIP CLASSIFICATION |
|---|---|---|
| 1 | Pick up coins | Tip |
| 2 | Undo buttons | Tip/tripod |
| 3 | Simulate food cutting | Tripod/power |
| 4 | Simulate page turning | Extension |
| 5 | Remove jar lid | Spherical |

| | | |
|---|---|---|
| 6 | Pour water from jug | Lateral |
| 7 | Pour water from carton | Spherical |
| 8 | Move empty tin | Power/Spherical |
| 9 | Move full jar | Power/Spherical |

Source: Kyberd et al (2009)

*3.3.1 Southampton Hand assessment Procedure (SHAP)*

The Southampton Hand Assessment Procedure (SHAP) was originally designed to assess prosthetic arms use for ADL (SHAP Assessors Protocol, 2016). SHAP is now used to assess upper extremity impairment and track improvement after treatments of the upper extremity () and the kit available at the Centre of Human Movement Sciences, UMCG.

The test comprises activities involving the transfer of abstract and daily use objects of varying weights in a goal directed manner within a defined workspace while the task performance is self timed. Functional impairment is assessed based on the deviation of the time of task performance compared to the minimum normal task performance averages.

3.4 Conceptual Model

This model shows the use of the Kinect v2 and its associated SDK particularly the Gesture Builder Application Programming Interface (API). The envisaged system as shown in figure 3.1 would sense the motion made by the patient or user in a rehabilitation game scenario. Based on a database of desired manipulative gestures that can be performed within a virtual or augmented reality game, a verification of the real-time motion being performed matching a desired functional goal could be realised. This could be used then to track progress or systematically alters the next task to facilitate motivation or simply increments a counter to keep scores. The same approach could apply to an assessment system within a game or as a stand-alone VR based functional assessment system.

The database would be developed using the Kinect v2 data from ADL tasks performed form the SHAP the data would be collected, the training set tagged and fed to the adaboost algorithm for training, after training a different data set will be applied for analysis. Data used for analysis could be tagged and supplied for training while fresh data is captured for analysis. Until the desired level of detection is achieved. The process is illustrated in Figure 3.2 below.

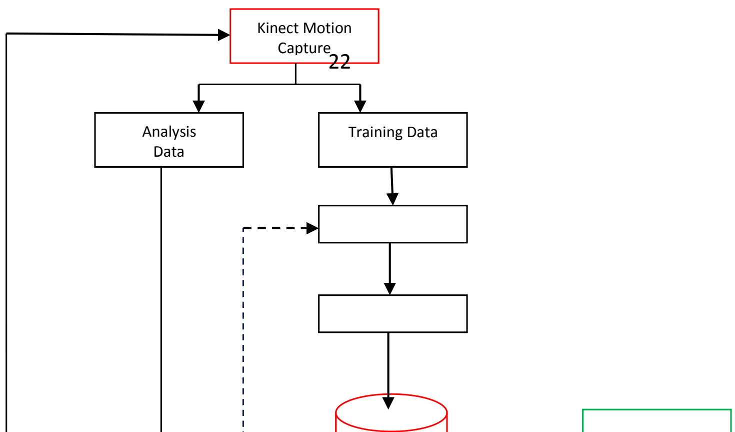

Figure 3.2 Conceptual model of a Kinect v2 gesture recognition database for a rehabilitation game or assessment scenario**.**

The development of a database for use within the rehabilitation game using the Kinect v2 is fairly straight forward. In the case of this project, the emphasis is in upper extremity rehabilitation therefore the manipulative or ADL gestures would be performed while seated.

The Kinect sensor would be employed to capture the motions of the subject during an activity. A subset of the ADL tasks in the SHAP would be utilized to develop and test the database. The desired gestures would be recorded by the Kinect v2, the range of motion of interest tagged and these tagged clips built into the detection database using the AI algorithms within the Visual Gesture Builder. Recorded untagged clips would be used to test the accuracy of detection of the desired motion. If the desired level as are not reached, more clips are captured, tagged and added to the solution. It is expected that the added clips with sufficient variation would improve the accuracy of the gesture database.

3.4 Hypothesis

In the light of the foregoing, the hypothesis to be tested is that, with sufficient data, the desired ADL represented by upper extremity data captured by the Kinect v2 would be detected reliably by the detector developed using the Visual Gesture Builder.

# CHAPTER 4

## RESEARCH DESIGN

This chapter describes the experiments carried out to generate data used to develop and test the performance of the Kinect v2 gesture database for the detection of some basic ADL motions involving reaching and grasping with the Kinect Visual Gesture Builder (VGB). It also describes experiment participants, ethical considerations and also indicates what data will be assessed to determine the success of failure of the ADL detector.

## 4.1 System Requirements

Microsoft recommends or requires the following computer hardware to effectively interface with the Kinect v2: A graphics processor ≥ DirectX 11, USB 3.0 and Windows 8 operating system and onwards.

*Table 4.1 Kinect v2 System Requirement*

| S/No | Hardware | Specification | Required | Recommended |
|---|---|---|---|---|
| 1 | Central Processing Unit (CPU) | i7 | | ✓ |
| 2 | Random Access Memory | 4GB | | ✓ |
| 3 | Graphics Processing Unit | DirectX 11 | ✓ | |
| 4 | Universal Serial Bus (USB) | 3.0 | ✓ | |
| 5 | Operating System (OS) | Windows 8.0/8.1/10 | ✓ | |

Source: Kaplan & Relyea (2013)

## 4.2 Experimental Setup

The proposed experiments will be based on the upper extremity actions or gestures of basic ADL i.e. skills which facilitate the realization of dressing, eating, mobility, toileting and hygiene as required in common hand function tests.

The assessment of improvement or responsiveness to therapy is determined by the repeated administering any standardized function test (Langhorne et al., 2011). Some of these function test elements could be used within the game for rehabilitation training and therefore used to automatically detect improvement by recording the performance data.

For example, picking and transferring a ball from a fixed position to another referenced position on a table, the subject would be required to sit on a chair facing a table with height around the level of the patient's arms and elbows at 90 degrees. The ball, cup or drinking bottle will be placed on a table in front of the patient at a full arms length along the anterior-posterior plane at an anthropometrically comfortable position. To investigate if the performance of an ADL task can be detected reliably by the Kinect v2, the Southampton Hand Assessment Procedure (SHAP) test (SHAP Assessors Protocol, 2016) was used. SHAP includes a battery of tests on abstract objects and objects of daily use. The ADL part of the test was selected as the areas of interest.


Figure 4.1 SHAP test kit opened.

An experimental protocol was developed, from the comprehensive SHAP test protocol and was used in the final data capture. The experiment only involved a subset of the ADL tasks from the comprehensive SHAP test. These include; pouring from a carton, lifting a heavy jar, pouring from a jug, picking up coins from a flat surface and lifting an empty tin. See Appendix A for the protocol.

As shown in figure 4.2, a chair would be required to allow the arm and hand to be placed at about the same level as the table surface, where the SHAP tray containing the object of the ADL task is placed on.

Figure 4.2 Setting for the SHAP ADL experiments

Three repetitions of the manipulative gesture were captured with the Kinect v2, for gesture recognition building and offline detection testing. The sensor was placed about 1.1 m above the ground level and test subject, placed about 2m from the sensor to locate the subject within the operating range of the Kinect v2 to maximize the accuracy of the data. Tao et al (2013) had indicated that the optimal position for the Kinect v1 in an upper extremity evaluation or detection scenario was between 1.45m and 1.75m from the user at 0.15m left or right.

For this data capture, the tray was placed 0.1m off the edge of the 0.75m high table. This region corresponds to the least depth accuracy error zone (< 2mm) in the vertical and horizontal planes of the Kinect v2 field of view as presented by Yang et al. (2015) are depicted in the hatched green areas. The zones and the respective capture accuracy are depicted in Figure 4.3 below.

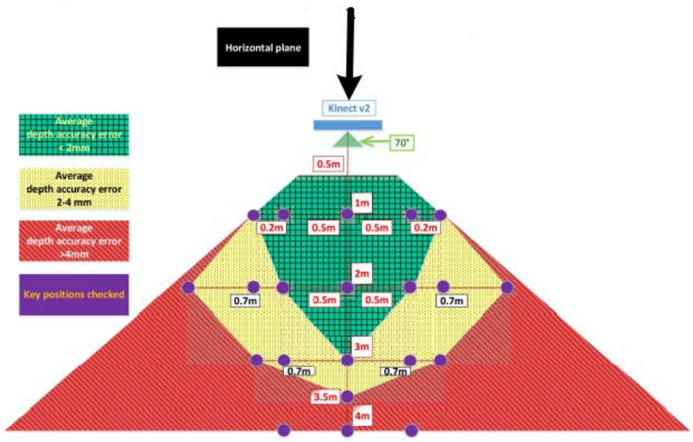


Figure 4.3 Accuracy error distribution of Kinect v2. *Yang et al (2015)*

## 4.3 Participants

Given the relatively short duration of the study and its exploratory nature, the data desired for use was to be made up of available healthy individuals aged between 18-40.

Conscious attempts were made to recruit participants across both genders, with a significant variation in morphological characteristics and age range to eliminate any bias, while enriching the training and testing data sets.

## 4.4 Data

The data captured involved Kinect v2 body stream clips of each of the participants performing the experiments which included the 6 selected ADL with 3 repetitions of each. The clips with the most reliably tracked skeletons per participant would be split into two groups with a ratio of 2/3 to 1/3 as suggested by Relyea & Marien (2013) implying about 67% used for training while 33% used to test the accuracy of the developed database for each of the selected gestures.

4.5 Using Visual Gesture Builder™

The gesture builder allows for the application of AI or machine learning to the recognition of gesture problems. Without serious consideration of computing resource constraints, the gesture builder, which is a new part of the Kinect v2 System Development Kit (SDK) allows for the gesture detection to become a data driven challenge and capable of detecting a gesture start and end as a discrete task or the progress in between the two or more points by using tagged captured data.

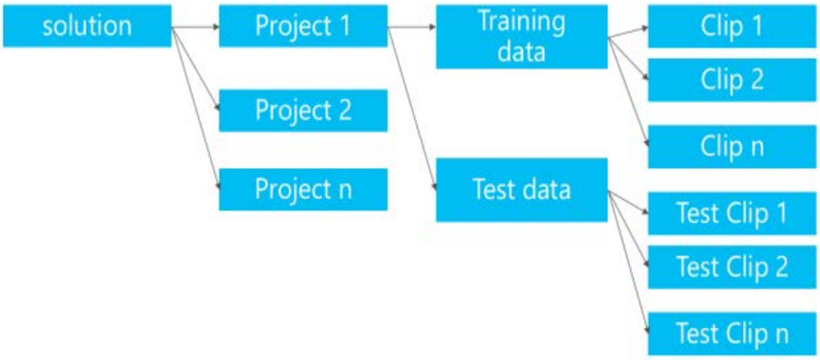


Figure 4.4 Visual Gesture Builder database building process

Figure 4.4 shows the process of using gesture builder for the detection of gestures. It also allows for testing and analysis of results with the option of a live preview. For each gesture or sets of gestures to be detected, the motion clips of the action are recorded and placed within the desired project. \thereafter, each clip in the training set is tagged by selecting the start to the end of the ADL motion. After tagging is complete, the project is compiled to train the detector, it reports on its choice of weak classifiers for the training set and then the detector is now available for detection of the desired motion.

With the testing or analysis component, data from different participants is loaded to the analysis heading. The clips are then analysed against the training set. The analysis automatically tags the test data. If the tagging is inaccurate, the test clips could be properly tagged and moved to the training group, compiled and re-tested. The test clip moved to training is replaced by a new clip for testing.

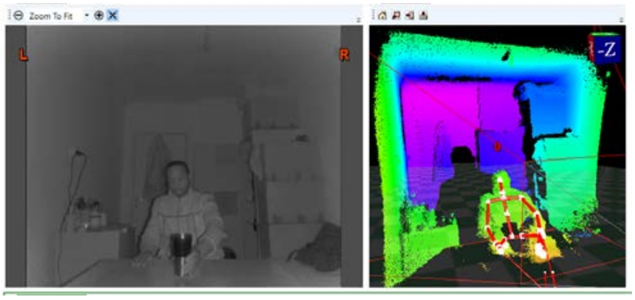

Figure 4.5 Visual Gesture Builder IR and Body Data Display

The VGB employs two algorithms; the adaboost for detecting discrete states in a Boolean fashion {True or False} and the random forest progress which is capable of detecting the progress of any captured motion, all these are methods of ensemble learning as discussed earlier. The overall workflow and the data utilization is shown in Figure 4.6.

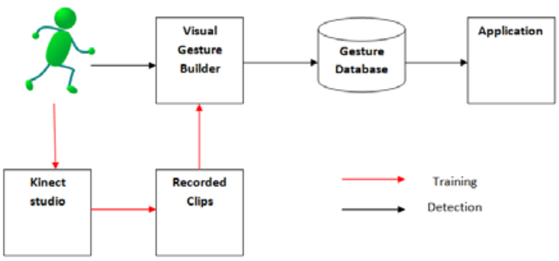


Figure 4.6 Kinect v2 visual gesture builder workflow.

4.6 Detection Assessment

The detection of the chosen ADL is based on the use of project analysis results of the VGB where the weighting factors of each final classifier will be recorded and finally the number of frames, the Root Mean Square (RMS) of the confidence values of detection will be displayed. This data would help in assessing the performance of machine learning algorithm (Fawcett, 2006). The closer the RMS value is to 1 or 100%, the better.

In a multiclass supervised learning scenario, involving multiple activities as done by Kim & Kim (2015), each activity detector or database was compared against all the activities captured in the study. Similarly, in this case, the emphasis is on the capability of the AI algorithms within the

**Met opmaak:** Inspringing: Eerste regel: 1,27 cm

Visual Gesture Builder to detect motor actions within an ADL function test scenario, with a restricted training dataset. The result will be a multiclass confusion matrix showing the performance of each gesture database against all the other ADL activities captured in the entire experiment. This is based on a suggestion by Luger (2005) for developing multiclass SVM. This will determine the specificity of each gesture detector within the developed database

A live verification of the database with the Kinect under data capture conditions will be carried out to determine the accuracy of the analysis results of the Visual Gesture Builder.

4.7 Ethical Consideration

As required by the UMCG, the study was approved by the local medical ethical committee (Approval reference: 2016.05.08_1) the study was carried out in agreement with the guidelines of the Helsinki protocol. Specific care was taken to ensure the consent of participants in the experiment was expressly obtained and documented by filling out a consent forms. The gender and age of the participants will be the data presented to forestall the possibility of using the data to easily identify the participant and the data will not be publicly presented in any person identifiable form except where permission is sought and granted. The consent form used is attached in Appendix B.

CHAPTER 5

**RESEARCH RESULTS**

This chapter presents the results of the experiment carried out with the SHAP kit and the level of detection achieved at the preliminary stage. It presents the performance of the adaboost algorithm and features utilized in training the detector for each of the ADL tasks selected.

5.1 Demographics

As at the time of reporting 15 participants mainly students of Hanze University and University of Groningen in the Netherlands 26.5 ± 4.7 years, 10 males and 5 females participated in the experiment, they all signed the consent forms as required by the ethics committee of the Human movement science department at the UMCG. One set of data was not used in the study because part of it was lost.

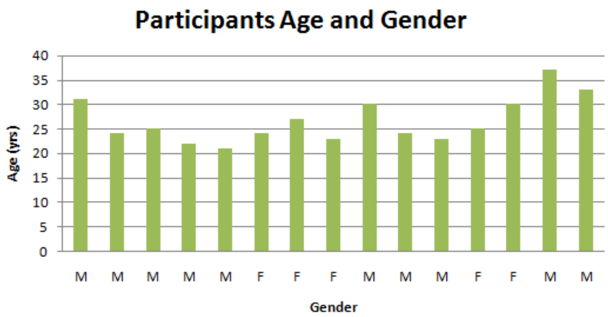


Figure 5.1 Age and Gender of experiment participants.

5.2 Data Management

The data collected was screened to ensure that the skeleton for each experiment tracked the depth image during ADL task completion. The data was coded and placed in different folders. Unsuitable samples; where the skeleton was not tracked or task execution was wrongly executed by the participant were clearly marked and not used for training or analysis.

5.3 Using Visual Gesture Builder (VGB)

The visual gesture builder API is very intuitive and requires more familiarity with simple settings which can be applied using the wizard. For the selected SHAP tasks, the settings of the training mask were as shown in Figure 5.2. The joints from the hip to the lower extremity were ignored. Furthermore, during training for each gesture, the left arm joint data was ignored except for the opening a lid task where the task required the use of both hands, as such, the left arm was not ignored for that case as shown in the project settings table 5.1 below.

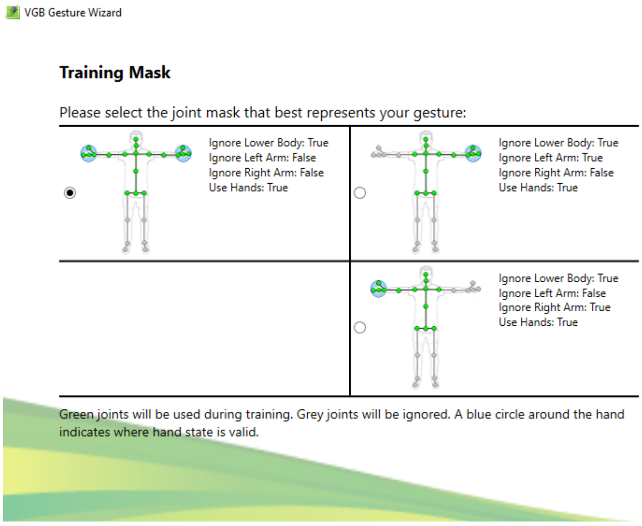


Figure 5.2 Joint mask selection on VGB wizard

*Table 5.1 VGB project settings for the 6 SHAP ADL Tasks*

| Name | Value | Carton Pour | Jar lifting | Jug pouring | Open Lid | Pick Coin | Tin Lifting |
|---|---|---|---|---|---|---|---|
| Accuracy Level | 0.95 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Number of Weak Classifiers at Runtime | 1000 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Filter Results | TRUE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Auto Find best Filtering Parameter | TRUE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Weight of False positives During Auto Find | 0.5 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Manual Filter Params: Num of Frames to Filter | 5 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Manual Filter Params: Threshold | 0.001 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Duplicate and Mirror Data during Training | TRUE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| % CPU for Training | 95 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Use Hands Data | TRUE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Ignore Left Arm | TRUE | ✓ | ✓ | ✓ | FALSE | ✓ | ✓ |
| Ignore Right Arm | FALSE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Ignore Lower body | TRUE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

For each of the data clips, the gesture desired was tagged after placement in the training folder of the

VGB. After all the loaded clips were tagged and verified, the gesture detector was built.

5.4 Adaboost Training and Analysis

Microsoft development and training teams recommend the use of 66 -70% of data samples for training and 30 – 33% for analysis. This ratio has been applied by Štrbac & Popović (2014) and Bhattacharya et al (2012) to Kinect detectors or classifiers. In this case 9 participants' data or skeleton tracked clips were used for training while the other 5 were used for analysis.  This corresponded to 64: 36 for training and analysis respectively.

The algorithm generated a pool of classifiers based on the 38 features. Microsoft does not provide definitions for these 38 features (Microsoft 2015b) attempts to define them except for the angles which are relative positions of the 25 joints in 3-D Cartesian plane would be speculative at best. A column plot of the number of weak classifiers, against the features for each of the selected SHAP ADL tasks is presented below:

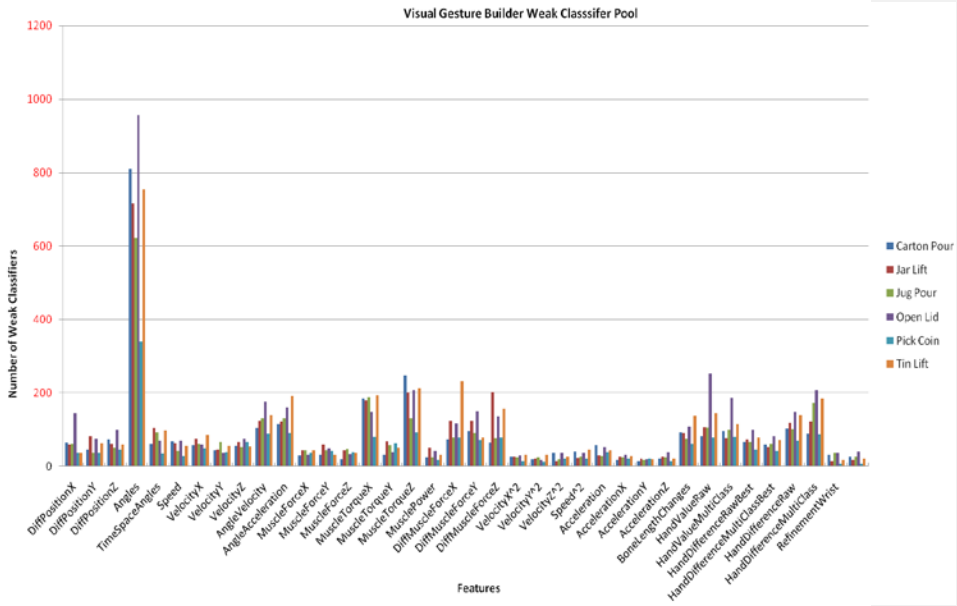


Figure 5.3 Plot of weak classifiers pool against features for selected SHAP tasks.

The feature with the highest number of weak classifiers for all the tasks was the skeletal angles with all except one ADL task i.e. the picking up coins, exceeding 600 weak classifiers generated during training. Subsequently, for 9 data sets submitted for training each unique ADL detector, the top 10 weak classifiers were produced and used by the adaboost algorithm for developing each detector. These top 10 weak classifiers and obtained results are presented below.

### *5.4.1 Pouring from a milk carton*

This activity is intended to mimic the pouring of a liquid from a carton into another vessel using one hand. For detecting this task, the top 10 weak classifiers generated for the task are shown below indicating the features of interest. The angles of the hand tip, hand and wrist of the right hand was used together with the muscle torque at the neck.

*Table 5.2 Top 10 Weak classifiers for Pouring from the milk carton*

| Features | Inferred Joint | fValue | Alpha |
|---|---|---|---|
| MuscleTorqueZ( Neck ) | Used | $\geq$ 0.099997 | 1.34668 |
| Angles( Head, ShoulderRight, HandRight ) | Rejected | < 94.000000 | 1.061935 |
| MuscleTorqueZ( Neck ) | Used | $\geq$ 0.299997 | 0.640294 |
| Angles( HandTipRightHandRightWristRight ) | Used | $\geq$ 104.000000 | 0.63746 |
| DiffMuscleForceY( ElbowRightWristRight ) | Rejected | $\geq$ -0.200000 | 0.635382 |
| DiffPositionZ( HipRightShoulderLeft ) | Used | < 0.000000 | 0.627732 |
| DiffPositionY( ElbowRightSpineMid ) | Used | $\geq$ 0.000000 | 0.552879 |
| DiffPositionZ( Head ShoulderLeft ) | Used | < 0.000000 | 0.542786 |
| Angles( Head ShoulderRightElbowRight ) | Used | < 116.000000 | 0.470378 |
| MuscleTorqueZ( HandRight ) | Used | $\geq$ 0.099997 | 0.464355 |

A screen capture of the VGB with training and analysis data is shown below. The training data set is indicated by the upper rectangle while the lower rectangle marks the test or analysis data. After analysing the clips presented to test the detector, the Root Mean Square Value for the confidence was 0.90 which was significant given that 1 is the maximum but less than the 0.95 required to guarantee robust detection.

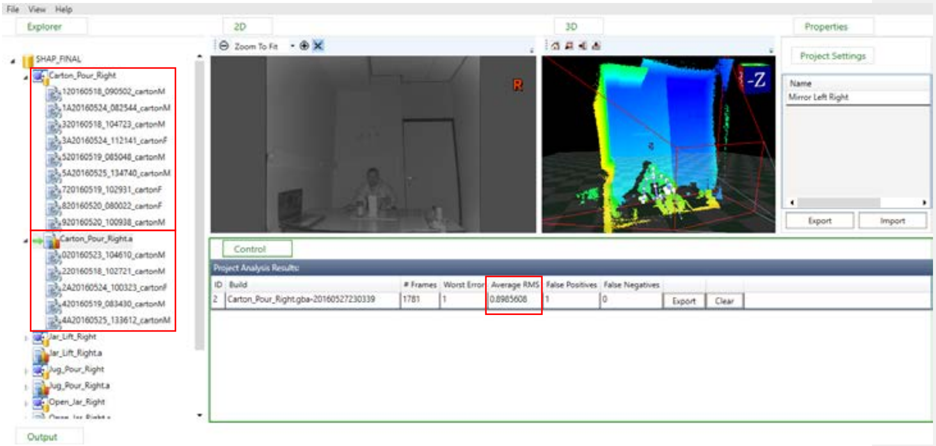

Figure 5.4 Screen shot of the VGB with carton pouring task training, analysis data and analysis result.

*5.4.2 Lifting jar across a barrier*

This task is aimed at assessing the ability of the subject in lifting relatively heavy object across a raised barrier and placed safely on the other side. The glass jar is filled with water to add to the weight of the otherwise empty jar. The top 10 weak classifiers for this ADL task are shown in the table below. The angles between the head, shoulder and right hand tip was part of the weak classifiers was used.

*Table 5.3 Top 10 Weak classifiers for lifting a filled jar across a barrier*

| Features | Inferred Joints | fValue | Alpha |
|---|---|---|---|
| Angles( Head, ShoulderRight, HandTipRight) | Used | < 100.0000 | 1.042535 |
| MuscleTorqueZ( Neck ) | Used | ≥ 0.199997 | 0.641804 |
| MuscleTorqueX( ShoulderLeft ) | Used | ≥-1.500003 | 0.632365 |
| DiffPositionY( ElbowRight, SpineMid ) | Used | ≥ 0.000000 | 0.623027 |
| Angles( Head, ShoulderRight, ThumbRight ) | Reject | ≥ 94.000000 | 0.411252 |
| MuscleTorqueX( HipRight ) | Reject | < 0.099997 | 0.410128 |
| MuscleTorqueY( ShoulderRight ) | Reject | ≥ 0.099997 | 0.406564 |
| VelocityZ( Head ) | Used | < -0.000003 | 0.371035 |
| DiffPositionY( HipLeft, SpineBase ) | Used | < 0.000000 | 0.365892 |
| Angles( Head, ShoulderRight, HandTipRight) | Reject | < 112.000000 | 0.360691 |

The figure below shows the screen shot with the training, analysis data and analysis result. The Root Mean Square Value of the confidence was 0.59 which is low and would require further training to improve the detector. The training, analysis and RMS results are marked in the red rectangles.

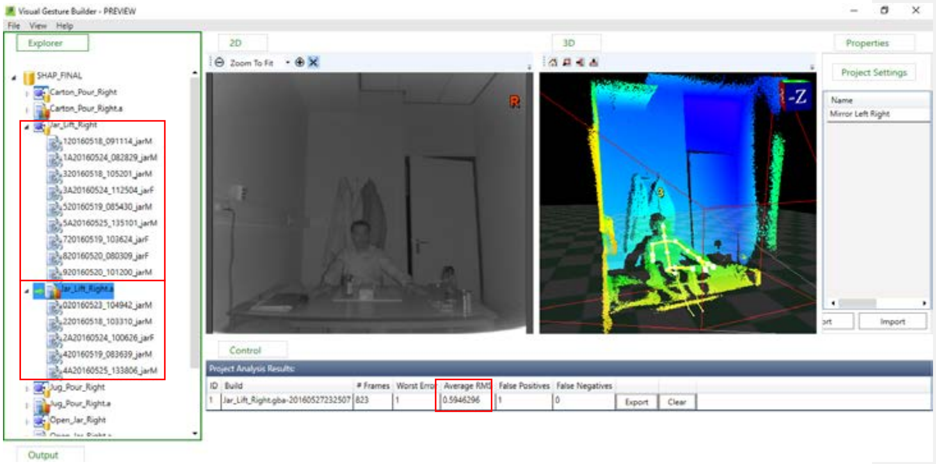

Figure 5.5 Screen shot of the VGB with jar lifting task training, analysis data and analysis result.

*5.4.3 Pouring from a jug*

This task is similar to the carton pouring task except that the grasping strategy for the jug differs from that of the carton. After training the tagged clips, the top 10 weak classifiers are shown in the table below. The differential position of the right hand and the tip was used while the angles of the thumb, hand and wrist was rejected after training at this point.

*Table 5.4 Top 10 Weak Classifiers for pouring from a jug*

| Features | Inferred Joints | fValue | Alpha |
|---|---|---|---|
| MuscleTorqueZ( Neck ) | Used | ≥ 0.099997 | 1.325125 |
| DiffPositionX( HipRight, HandTipRight ) | Used | ≥ 0.000000 | 1.289062 |
| Angles( ThumbRight, HandRight, WristRight ) | Rejected | ≥ 96.000000 | 0.677088 |
| Angles( Head, ShoulderRight, ElbowRight ) | Used | < 134.000000 | 0.660001 |
| DiffPositionZ( HipLeft, SpineMid ) | Used | ≥ 0.000000 | 0.440424 |
| MuscleTorqueX( ShoulderRight ) | Rejected | < -5.000000 | 0.421238 |
| MuscleTorqueZ( ShoulderLeft ) | Used | < -0.000003 | 0.39333 |
| MuscleTorqueZ( HandRight ) | Used | ≥ 0.099997 | 0.378896 |
| DiffPositionZ( HipLeft, HipRight ) | Used | ≥ 0.000000 | 0.366383 |
| Angles( SpineMid, Head, ShoulderLeft ) | Used | < 44.000000 | 0.362104 |

Like the carton pouring task, the detector had a significant Root Mean Square confidence which was 0.83 as shown on the screen capture below alongside the location of the training and analysis data sets. The RMS value is significant but below the suggested 0.95 required for deployment level detectors or databases. It is expected that with further data, the RMS will improve until it reaches the desired levels.

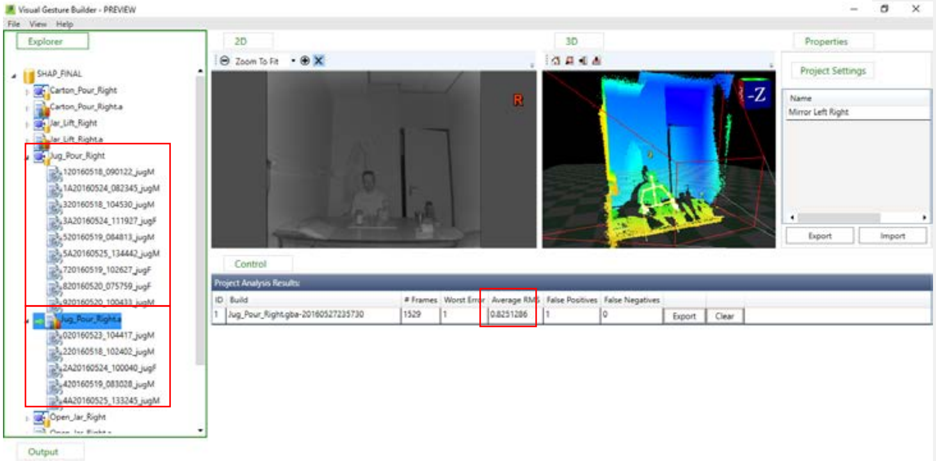


Figure 5.6 Screen shot of the VGB with jug pouring task training, analysis data and analysis result

*5.4.4 Opening a jar's lid*

This ADL task requires the use of two hands one to steady the jar, while the other hand which is being assessed opens the jar lid. After tagging the training data, the algorithm produced the top 10 weak classifiers below. In this detector the weak classifiers had significant left hand data. The only data used for the right hand was the wrist refinement.

*Table 5.5 Top 10 Weak Classifiers for opening a jar's lid*

| Feature | Inferred Joint | fValue | Alpha |
|---|---|---|---|
| DiffPositionY( HandRight, SpineMid ) | Rejected | ≥ 0.00000 | 0.857054 |
| DiffPositionX( HandLeft, ShoulderLeft ) | Used | ≥ 0.000000 | 0.773892 |
| RefinementWrist HAND: ( Right ) | Used | ≥ 0.010000 | 0.772312 |
| DiffPositionX( ElbowLeft, ShoulderLeft ) | Used | ≥ 0.000000 | 0.662979 |
| Angles( Head, ShoulderLeft, ElbowLeft ) | Used | < 138.000000 | 0.554752 |
| Angles( WristLeft, ElbowLeft, ShoulderLeft ) | Rejected | ≥ 96.000000 | 0.544208 |
| Angles( Head, ShoulderLeft, ElbowLeft ) | Used | < 142.000000 | 0.473123 |
| DiffPositionY( HandLeft, HandTipLeft ) | Rejected | < 0.000000 | 0.470146 |
| MuscleTorqueX( HandLeft ) | Used | < -0.300003 | 0.458506 |
| DiffPositionZ( Head, SpineBase ) | Used | ≥ 0.000000 | 0.440264 |

The Root Mean Square value of the confidence for the detector as measured by the test clips was 0.76

which is low. The mean RMS value, location of the training and analysis data sets are shown in the screen capture below. Further training would be required to improve the detector.

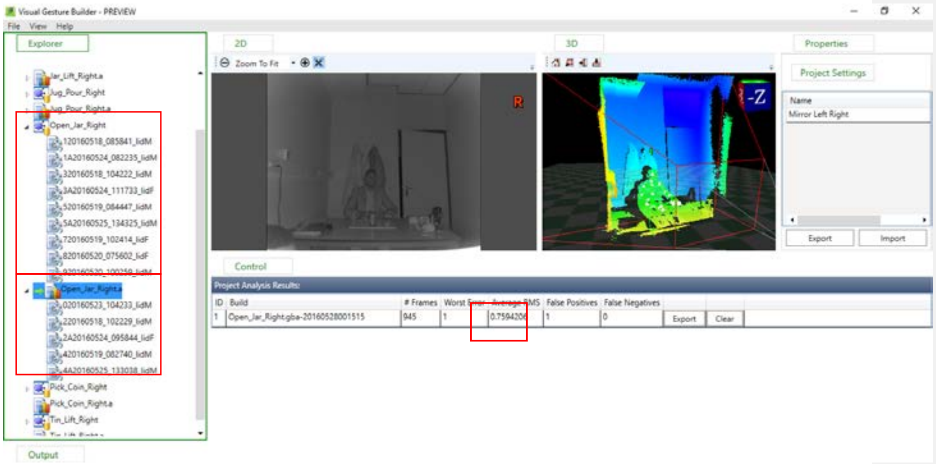


Figure 5.7 Screen shot of the VGB with Jar opening task training, analysis data and analysis result

*5.4.5 Picking up coins*

This task employs primarily the use of the fingers to retrieve four coins, 2, 2 pennies and 2, penny coins off the tray to an empty jar. After tagging the training data set, the training algorithm produced the top 10 weak classifiers below**.** The weak were dominated by the angles involving the hand tip, wrist refinement and the pre-built hand states, especially the open state.

*Table 5.6 Top 10 Weak Classifiers for picking up coins*

| Features | Inferred Joint | fValue | Alpha |
|---|---|---|---|
| Angles( Head, ShoulderRight, HandTipRight) | Used | < 106.0000 | 0.786924 |
| Angles( SpineMid, Head, HandTipRight ) | Used | < 44.0000 | 0.572314 |
| VelocityZ( HandRight ) | Used | < -0.100003 | 0.559919 |
| RefinementWrist HAND: ( Right ) | Used | ≥ 0.200000 | 0.453572 |
| RefinementWrist HAND: ( Right ) | Used | ≥ 0.020000 | 0.291309 |
| MusclePower( HandTipRight ) | Used | < 0.000000 | 0.290396 |
| VelocityY( ElbowRight ) | Rejected | ≥ -0.000003 | 0.281013 |
| HandDifferenceMultiClass( Right, OPEN, UNKNOWN) | Used | < 0.480000 | 0.269386 |
| HandValueMultiClass( Right, CLOSED ) | Rejected | < 0.400000 | 0.242479 |
| MuscleTorqueZ( ShoulderLeft ) | Used | < 1.199998 | 0.206369 |

The figure below shows the screen shot with the training, analysis data sets and analysis result. The Root Mean Square Value of the confidence was 0.74 which is low for detecting the coin picking task. This result would require improvement by supplying further tagged and assessment data clips.

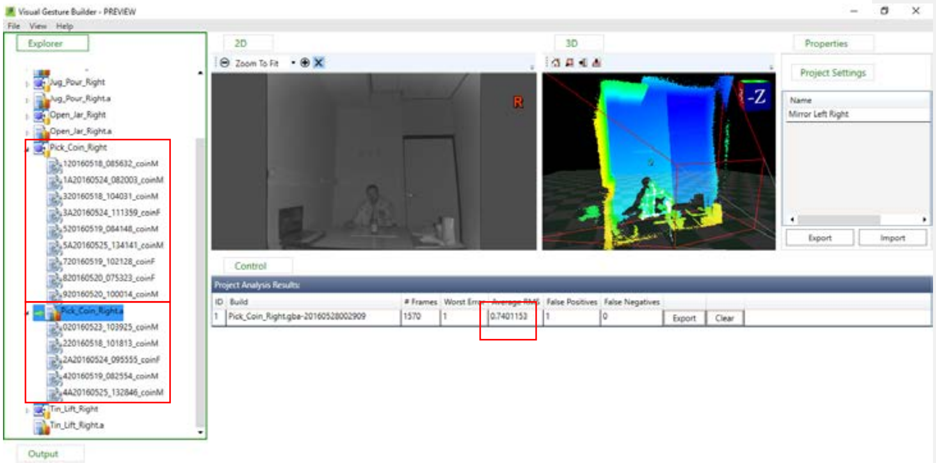


Figure 5.8 Screen shot of the VGB with coin picking task training, analysis data and analysis result

*5.4.6 Lifting an empty tin across a barrier*

This ADL task, while similar to the jar lifting task primarily differs in terms of the weight of the objects to be moved.  After tagging and training, the top 10 weak classifiers produced are shown below. The hand refinement and hand state values were some of the weak classifiers used.

*Table 5.7 Top 10 Weak Classifiers for lifting an empty tin across a barrier*

| Features | Inferred Joint | fValue | Alpha |
|---|---|---|---|
| DiffPositionY( WristRight, SpineMid ) | Used | ≥ 0.000000, | 1.110713 |
| Angles( Head, ShoulderRight, SpineShoulder ) | Used | ≥ 54.000000, | 0.655838 |
| MuscleTorqueZ( Neck ) | Used | ≥ 0.099997, | 0.651089 |
| DiffPositionX( Head, HipLeft ) | Used | < 0.000000, | 0.468952 |
| VelocityX( ShoulderRight ) | Used | ≥ -0.000003, | 0.463269 |
| DiffPositionZ( Head, HipRight ) | Used | ≥ 0.000000, | 0.437275 |
| Angles( HandRight, WristRight, ElbowRight ) | Rejected | < 84.000000, | 0.434455 |
| RefinementWrist HAND: ( Right ) | Used | ≥ 0.100000, | 0.424425 |
| DiffPositionY( HandRight, ElbowRight ) | Rejected | < 0.000000, | 0.421168 |
| HandDifferenceRaw( Right, LASSO, CLOSED ) | Used | < 0.480000, | 0.382012 |

The Root Mean Square value of the confidence for the detector as measured by the test clips was 0.64 which is low. The mean RMS value, location of the training and analysis data sets are shown in the screen capture below. This detector would require further training to achieve significant confidence levels.

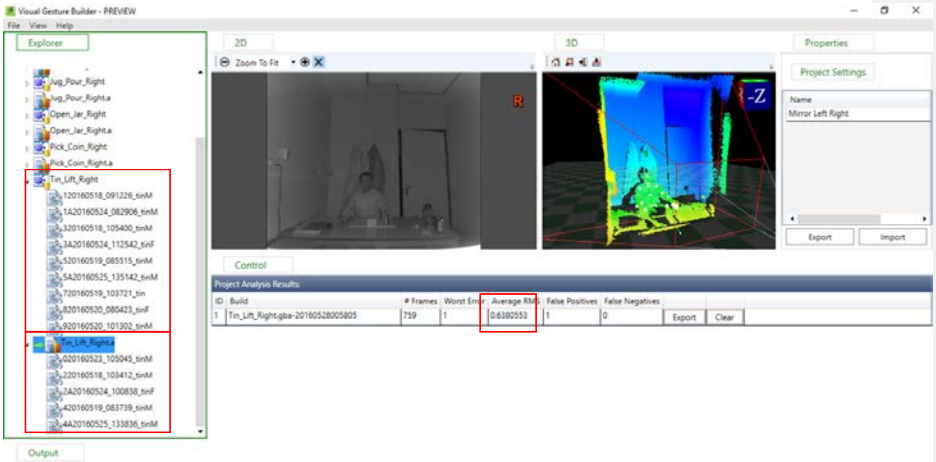


Figure 5.9 Screen shot of the VGB with Tin Lift task training, analysis data and analysis result

## 5.5 General detector comparison

To further test the system, live previews are possible with the Gesture Viewer; a preview of the solution is also possible allowing the Kinect v2 to display real-time the detection of which task was being carried out with the developed database as a reference. The ability of the detectors to discriminate between the tasks would be highly valued. Based on this idea, each task detector was tested with other task testing or analysis clips to investigate its discriminatory ability. The results of the tests are presented as a confusion matrix below.

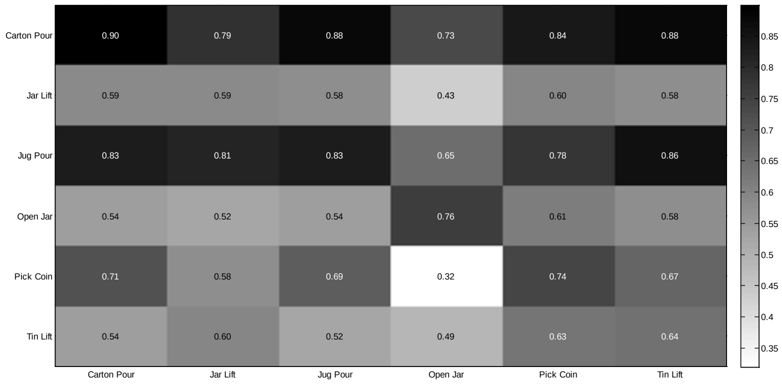


Figure 5.10 Multiclass confusion matrix for the selected SHAP tasks

The matrix in figure 5.9 shows the comparison of the detectors against test data for each ADL task in the experiment. For a successful system, the diagonal elements of matrix ought to have had the darkest shades corresponding to ≥ 0.95 while the off diagonal elements lightest. This would indicate accurate detection of gestures and discrimination between gestures. The ideal condition will be a diagonal matrix with sparse of diagonal elements.

At this iteration, only two detectors showed the necessary promise; carton pouring and jug pouring at 0.90 and 0.83 respectively while opening a jar and picking a coin show potential for improvements with subsequent training with more data.

The performance of the jar lifting task and the tin lifting task were initially confounding but, on close examination, it appears there was a failure on the part of the Supervising researcher to monitor the lifting strategy for the tin and jar lifting task therefore, a mix of strategies were used by the test subjects. A number of subjects used spherical grasp instead of the recommended power grasp. For the other tasks with relatively good average RMS confidence values, the grasp strategies were fairly consistent,

CHAPTER 6

## CONCLUSION AND RECOMMENDATIONS

This chapter highlights the key results of the study carried out. It discusses the results obtained in the study in relation to the research questions and hypothesis presented earlier in the report. The chapter also presents recommendations for further work.

### 6.1 Conclusion

The art of supervised learning or unsupervised learning requires features extraction from the data collected for intelligence extraction. The Kinect v2 sensor as shown in the results Figure 5.1 indicates that the detector mainly relies on the angles of the skeletal data among the 38 features that are available for use in developing the gesture detector.  The features appear intuitive but no formal definitions are provided in any publicly available documentation (Microsoft 2015) attempts to define them would involve significant speculation.

For each of the ADL tasks, a detector or database was developed. The general performance of the database was low except for the carton pouring  and the jug pouring tasks which when compared with the test data clips showed a RMS confidence of  0.90 and 0.83 for the tasks respectively, The next best RMS confidence were the jar opening and picking a coin tasks with  RMS confidence of 0.76 and 0.74 while the least detection confidence was by the apparently the simple task of lifting a water filled jar and empty tin across a barrier that was the least accurately detected for the given data set at 0.59 and

0.64 respectively.

These results are at a point in the data capture, training and testing iteration, it is expected that with further data capture, training and testing, the detection within class will improve while out of class detection reduces. In an activity detection scenario Kim & Kim (2015) obtained scores ranging from 0.85-0.96 in class detection and 0.06-0.11 between classes with many cases of outright discrimination resulting in a matrix with sparse off diagonal elements employing neural networks.

There appears to be a poor detection overall given that for product level development, it is recommended that confidence values for detection should be $\geq 0.95$. For the given data collected, the detectors show promise but fail to meet the required levels. However, with further data collection and further training and analysis iterations, the gesture detector would improve significantly and hopefully approach the required $\geq 0.95$.

A close observation of the data showed that the grasping strategy for some of the tasks like the jar and tin lifting task were carried out with mixed strategies, some subjects used a spherical grasp while a lateral grasp was expected, this was observed earlier but was allowed to see how well the detector will handle the mixed strategy employed by the subjects.

For the data collected, given the relatively short time of the study only 15 subjects were captured and only 14 subjects' data was used. This is low given the complexity of the tasks and the requirement of a 66:34 split between the training and assessment data sets, it is thus improper to move a poorly detected test case to training as it will render the ratios unbalanced unless there is extra data for re-assessment.

Another challenge is that the training mainly consisted of positively tagged data and all excluded or untagged was perceived as negative, this is a recommended practice, but to improve or modify the gesture detector, there is the option of recording a confounding gesture and tagging it as a negative gesture and supplied to training.

The top 10 weak classifier data for each of the gesture detector revealed the key features considered by the algorithm in arriving at the detection. The idea being that the weak classifiers will be improved with iterations of training. This weak classifier list for a highly reliable gesture detector is said to be a valuable tool for investigating the features required to detect the said gesture and therefore available for heuristic detection within the Kinect gesture detection development framework.

Live preview for each discrete gesture is possible i.e. at the project level while the whole set of gestures could be previewed together with each detector detecting its specific gesture. This idea leads to the development of the confusion matrix which was developed by testing the 6 ADL trained set with test data for other activities to assess the gesture detector specificity. The resulting matrix confirmed the similarity of the pouring task using the jug and the carton, while interestingly, the poorest detectors at 0.59 and 0.64 for the jar and tin lifting tasks appeared to detect the pouring tasks this obviously indicates the unreliability of the gesture detectors at that point.

Based on the results obtained so far, it is possible to conclude that the Kinect v2 sensor and its associated SDK, especially the Visual Gesture Builder has the potential for detecting ADL activities which can be used within rehabilitation games or assessment systems. To conclude affirmatively the RMS values for each detector would be above the 0.95 value for each detector to be used reliably. This recommended value of average RMS was not realised for any of the 6 detectors due to the shortage of time to acquire more data for training and analysis.

## 6.2 Recommendations

There needs to be more data captured for training and testing of the gesture detectors to achieve the required $\geq 0.95$ RMS confidence. In the case of this study, 20 subjects were initially envisaged but only 14 person data was used. Further data should be collected in an ongoing fashion to investigate the anticipated improvement of the gesture detectors without reference to the initial 20 person proposal.

It was observed that the objects of the SHAP kit occluded the skeleton in some cases and the suggested the possibility of missing details of the hand tip and thumb with the Kinect v2 placed frontally. It would be worthwhile investigating the placement of the Kinect overhead at a 60 degree angle or more to offer a better view of the skeletal features interest especially if fine motor detection is required. There are applications where Head and Shoulder Profile (HASP) from Kinect depth images have been used for human traffic detection and counting (Zhu & Wong 2013). It would be possible to utilize skeletal data from such a view for gesture assessment as it minimizes occlusions.

It is expected that if the gesture detection accuracy of $\geq 0.95$ is achieved for all tasks, an automation of the SHAP could be developed given that it will be possible to count the number of frames within the detected gesture block, as such, the time of completion of the task can be successfully calculated. This feature, available in the Kinect v2 SDK would make the test automation possible.

Similarly, it would also be possible to develop a virtual SHAP test given that with the actual SHAP test, some items in the kit are glassware liable to be dropped by patients. The shattering of the glassware could cause injury and unnecessary replacement costs. These could be mitigated by a VR SHAP testing system based on the Kinect V2.

## REFERENCES

Adams Jr, H.P., 2016. General Concepts: Management of Acute IschemicStroke. In IschemicStrokeTherapeutics (pp. 1-5). Springer International Publishing.

Alabbasi, H., Gradinaru, A., Moldoveanu, F. and Moldoveanu, A., 2015, November. Human motion tracking & evaluation using Kinect V2 sensor. In E-Health and Bioengineering Conference (EHB), 2015 (pp. 1-4). IEEE.

Alankus, G., Lazar, A., May, M. and Kelleher, C., 2010, April. Towards customizable games for stroke rehabilitation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 2113-2122). ACM.

Barsotti, M., Leonardis, D., Loconsole, C., Solazzi, M., Sotgiu, E., Procopio, C., Chisari, C., Bergamasco, M. and Frisoli, A., 2015, August. A full upper limb robotic exoskeleton for reaching and grasping rehabilitation triggered by MI-BCI. In Rehabilitation Robotics (ICORR), 2015 IEEE International Conference on (pp. 49-54). IEEE.

Bhattacharya, S., Czejdo, B. and Perez, N., 2012, November. Gesture classification with machine learning using kinect sensor data. In *Emerging Applications of Information Technology (EAIT), 2012 Third International Conference on* (pp. 348-351). IEEE.

Byl, N.N., Pitsch, E.A. and Abrams, G.M., 2008. Functional outcomes can vary by dose: learning-based sensorimotor training for patients stable poststroke. Neurorehabilitation and neural repair, 22(5), pp.494-504.

Chang, Y.J., Chen, S.F. and Huang, J.D., 2011. A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. Research in developmental disabilities, 32(6), pp.2566-2570.

Chedoke arm and hand activity inventory (CAHAI) 2016 Available from: http://www.cahai.ca/layout/content/CAHAI-Manual-English.pdf [24th April 2016]

Clark, R.A., Pua, Y.H., Fortin, K., Ritchie, C., Webster, K.E., Denehy, L. and Bryant, A.L., 2012. Validity of the Microsoft Kinect for assessment of postural control. Gait & posture, 36(3), pp.372-377.

Cottone, P., Maida, G. and Morana, M., 2014. User activity recognition via kinect in an ambient intelligence scenario. IERI Procedia, 7, pp.49-54.

da Silva Cameirão, M., Bermúdez i Badia, S., Duarte, E. and Verschure, P.F., 2011. Virtual reality based rehabilitation speeds up functional recovery of the upper extremities after stroke: a randomized controlled pilot study in the acute phase of stroke using the rehabilitation gaming system. Restorative neurology and neuroscience, 29(5), pp.287-298.

Dobkin, B.H., 2009. Progressive staging of pilot studies to improve phase III trials for motor interventions. Neurorehabilitation and Neural Repair, 23(3), pp.197-206.

Dutta, T., 2012. Evaluation of the Kinect™ sensor for 3-D kinematic measurement in the workplace. Applied ergonomics, 43(4), pp.645-649.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern recognition letters, 27(8), pp.861-874.

Freeman, C., Rogers, E., Burridge, J.H., Hughes, A.M. and Meadmore, K.L., 2015. Iterative Learning Control for Electrical Stimulation and Stroke Rehabilitation. Springer.

Galna, B., Barry, G., Jackson, D., Mhiripiri, D., Olivier, P. and Rochester, L., 2014. Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease. Gait & posture, 39(4), pp.1062-1068.

Gatica-Rojas, V. and Méndez-Rebolledo, G., 2014. Virtual reality interface devices in the reorganization of neural networks in the brain of patients with neurological diseases. Neural regeneration research, 9(8), p.888.

Hastie, T., Tibshirani, R., Friedman, J. and Franklin, J., 2005. The elements of statistical learning: data mining, inference and prediction. The Mathematical Intelligencer, 27(2), pp.83-85.

Hochstenbach-Waelen, A. and Seelen, H.A., 2012. Embracing change: practical and theoretical considerations for successful implementation of technology assisting upper limb training in stroke. Journal of neuroengineering and rehabilitation, 9(1), p.1.

Hocine, N., Gouaich, A., Cerri, S.A., Mottet, D., Froger, J. and Laffont, I., 2015. Adaptation in serious games for upper-limb rehabilitation: an approach to improve training outcomes. User Modeling and User-Adapted Interaction, 25(1), pp.65-98.

Hortal, E., Iáñez, E., Úbeda, A., Tornero, D. and Azorín, J.M., 2013. Decoding Upper Limb Movement Velocity for Stroke Rehabilitation. Converging Clinical and Engineering Research on Neurorehabilitation, p.415.

Huang, T.C., Cheng, Y.C. and Chiang, C.C., 2013. Automatic Dancing Assessment Using Kinect. In Advances in Intelligent Systems and Applications-Volume 2 (pp. 511-520). Springer Berlin Heidelberg.

Hughes, A.M., Burridge, J.H., Demain, S.H., Ellis-Hill, C., Meagher, C., Tedesco-Triccas, L., Turk, R. and Swain, I., 2014. Translation of evidence-based Assistive Technologies into stroke rehabilitation: users' perceptions of the barriers and opportunities. BMC health services research, 14(1), p.1.

Kaplan, J & Relyea, R. 2013. Kinect Data Sources and Programming Model Jump Start, Power Point Presentation, Microsoft Virtual Academy. Availablefrom: https://mva.microsoft.com/en-us/training-courses/programming-kinect-for-windows-v2-jump-start-9088?l=IwlH5Lf4_2604984382 [18th February 2016].

Kelly-Hayes, M., Robertson, J.T., Broderick, J.P., Duncan, P.W., Hershey, L.A., Roth, E.J., Thies, W.H. and Trombly, C.A., 1998. The American Heart Association stroke outcome classification: executive summary. Circulation, 97(24), pp.2474-2478.

Kim, H. and Kim, I., 2015. Human Activity Recognition as Time-Series Analysis. Mathematical Problems in Engineering, 2015.

Kyberd, P.J., Murgia, A., Gasson, M., Tjerks, T., Metcalf, C., Chappell, P.H., Warwick, K., Lawson, S.E. and Barnhill, T., 2009. Case studies to demonstrate the range of applications of the Southampton Hand Assessment Procedure. The British Journal of Occupational Therapy, 72(5), pp.212-218.

Langhorne, P., Bernhardt, J. and Kwakkel, G., 2011. Stroke rehabilitation. The Lancet, 377(9778), pp.1693-1702.

Laver, K.E., Schoene, D., Crotty, M., George, S., Lannin, N.A. and Sherrington, C., 2013. Telerehabilitation services for stroke. Cochrane Database Syst Rev, 12.

Levin, M.F., Magdalon, E.C., Michaelsen, S.M. and Quevedo, A.A., 2015. Quality of Grasping and the Role of Haptics in a 3-D Immersive Virtual Reality Environment in Individuals With Stroke. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 23(6), pp.1047-1055.

Light, C.M., Chappell, P.H. and Kyberd, P.J., 2002. Establishing a standardized clinical assessment tool of pathologic and prosthetic hand function: normative data, reliability, and validity. Archives of physical medicine and rehabilitation, 83(6), pp.776-783.

Liu, W., 2010. Natural user interface-next mainstream product user interface. In 2010 IEEE 11th International Conference on Computer-Aided Industrial Design&Conceptual Design 1.

Loue, S. and Sajatovic, M., 2008. Encyclopedia of aging and public health. Springer Science & Business Media.

Loureiro, R.C., Harwin, W.S., Nagai, K. and Johnson, M., 2011. Advances in upper limb stroke rehabilitation: a technology push. Medical & biological engineering & computing, 49(10), pp.1103-1118.

Lower, B. & Relyea, R., 2013. Programming Kinect for Windows v2 Jump Start, Power Point Presentation, Microsoft Virtual Academy. Available from: https://mva.microsoft.com/en-us/training-courses/programming-kinect-for-windows-v2-jump-start-9088?l=Uw33iKf4_2404984382. [18 February, 2016]

Lucca, L.F., 2009. Virtual reality and motor rehabilitation of the upper limb after stroke: a generation of progress?. Journal of rehabilitation medicine, 41(12), pp.1003-1006.

Luger, G.F., 2005. Artificial intelligence: structures and strategies for complex problem solving. Pearson education.

Mackay, J., Mensah, G.A., Mendis, S. andGreenlund, K., 2004. The atlas of heart disease and stroke. World Health Organization.

Maleshkov, S. and Chotrov, D., 2013. Affordable Virtual Reality System Architecture for Representation of Implicit Object Properties. arXiv preprint arXiv:1308.5843.

Microsoft 2015, Meet Kinect for Windows. Available from https://dev.windows.com/en-us/kinect [20th November, 2015]

Miranda, L., Vieira, T., Martinez, D., Lewiner, T., Vieira, A.W. and Campos, M.F., 2012, August. Real-time gesture recognition from depth data through key poses learning and decision forests. In Graphics, Patterns and Images (SIBGRAPI), 2012 25th SIBGRAPI Conference on (pp. 268-275). IEEE.

Morel, M., Bideau, B., Lardy, J. and Kulpa, R., 2015. Advantages and limitations of virtual reality for balance assessment and rehabilitation. Neurophysiologie Clinique/Clinical Neurophysiology, 45(4), pp.315-326.

Mousavi Hondori, H. and Khademi, M., 2014. A review on technical and clinical impact of microsoftkinect on physical therapy and rehabilitation. Journal of Medical Engineering, 2014.

Nordin, N.A.M., Aziz, N.A., Sulong, S. and Aljunid, S.M., 2012. Function and quality of life following stroke rehabilitation: have our stroke patients gained optimum recovery?. BMC Public Health, 12(2), p.1.

Ojha, A.K., 1994, April. An application of virtual reality in rehabilitation. In Southeastcon'94. Creative Technology Transfer-A Global Affair., Proceedings of the 1994 IEEE (pp. 4-6). IEEE.

Paquin, K.C., 2014. Effectiveness of commercial video gaming on fine motor recovery in chronic strke within community-level rehabilitation.

Podobnik, J., Mihelj, M. and Munih, M., 2009, June. Upper limb and grasp rehabilitation and evaluation of stroke patients using HenRiE device. In Virtual Rehabilitation International Conference, 2009 (pp. 173-178). IEEE.

Relyea, R., & Marien, J., 2013. Programming Kinect for Windows v2 Jump Start, Power Point Presentation, Microsoft Virtual Academy. Availablefrom: https://mva.microsoft.com/en-us/training-courses/programming-kinect-for-windows-v2-jump-start-9088?l=IwlH5Lf4_2604984382 [18th February 2016].

Schapire, R.E., 2013. Explaining adaboost. In Empirical inference (pp. 37-52). Springer Berlin Heidelberg.

Schwarz, L.A., Mkhitaryan, A., Mateus, D. and Navab, N., 2012. Human skeleton tracking from depth data using geodesic distances and optical flow. Image and Vision Computing, 30(3), pp.217-226.

SeeMe 2014: Brontes Processing Sp. z o.o.. A virtual reality rehabilitation systems by Available from: http://www.virtual-reality-rehabilitation.com [24th April 2016]

SHAP Assessors Protocol (2016) Accessedfrom: http://www.shap.ecs.soton.ac.uk/about-pubs.php on the 28th May 2016.

Standen, P.J., Threapleton, K., Connell, L., Richardson, A., Brown, D.J., Battersby, S., Sutton, C.J. and Platts, F., 2015. Patients' use of a home-based virtual reality system to provide rehabilitation of the upper limb following stroke. Physical therapy, 95(3), pp.350-359.

Štrbac, M., Kočović, S., Marković, M. and Popović, D.B., 2014. Microsoft kinect-based artificial perception system for control of functional electrical stimulation assisted grasping. BioMed research international, 2014.
Strbac, M.D. and Popovic, D.B., 2014, November. Computer vision with Microsoft Kinect for control of functional electrical stimulation: ANN classification of the grasping intentions. In Neural Network Applications in Electrical Engineering (NEUREL), 2014 12th Symposium on (pp. 153-156). IEEE.

Strong, K., Mathers, C. and Bonita, R., 2007. Preventing stroke: saving lives around the world. The Lancet Neurology, 6(2), pp.182-187.

Stroke, 2016, Post Stroke Conditions Available from: http://www.stroke.org/we-can-help/survivors/stroke-recovery/post-stroke-conditions/physical [22nd April 2016]

Sveen, U., Thommessen, B., Bautz-Holter, E., Wyller, T.B. and Laake, K., 2004. Well-being and instrumental activities of daily living after stroke. Clinical rehabilitation, 18(3), pp.267-274.

Tao, G., Archambault, P.S. and Levin, M., 2013, August. Evaluation of Kinect skeletal tracking in a virtual reality rehabilitation system for upper limb hemiparesis. In Virtual Rehabilitation (ICVR), 2013 International Conference on (pp. 164-165). IEEE.

Teasell, R. andHussein, N., 2016. General Concepts: Therapies for Rehabilitationand Recovery. In IschemicStrokeTherapeutics (pp. 195-201). Springer International Publishing.

Thrasher, T.A., Zivanovic, V., McIlroy, W. and Popovic, M.R., 2008. Rehabilitation of reaching and grasping function in severe hemiplegic patients using functional electrical stimulation therapy. Neurorehabilitation and neural repair, 22(6), pp.706-714.

Thrift, A.G., Cadilhac, D.A., Thayabaranathan, T., Howard, G., Howard, V.J., Rothwell, P.M. and Donnan, G.A., 2014. Global strokestatistics. International Journal of Stroke, 9(1), pp.6-18.

Valoriani, M., &Giorio, C., 2015. Develop Store Apps with Kinect for Windows v2, Power Point Presentation, .NETCAMPUS. Available from: http://www.slideshare.net/tinux/develop-store-apps-with-kinect-for-windows-v2 [18th February, 2016].

van Diest, M., Lamoth, C.J., Stegenga, J., Verkerke, G.J. and Postema, K., 2013. Exergaming for balance training of elderly: state of the art and future developments. Journal of neuroengineering and rehabilitation, 10(1), p.1.

van Teijlingen, W., van den Broek, E.L., Könemann, R. and Schavemaker, J.G., 2012. Towards sensing behavior using the Kinect.

Vogiatzaki, E. and Krukowski, A., 2016. Modern Stroke Rehabilitation through e-Health-based Entertainment. Springer.

Wang, C.C. and Wang, K.C., 2007. Hand Posture recognition using Adaboost with SIFT for human robot interaction. In Recent progress in robotics: viable robotic service to human (pp. 317-329). Springer Berlin Heidelberg.

World Confederation for Physical Therapy, 2014. World Confederation for Physical Therapy (WCPT) Glossary. Available from: http://wwwwcptorg/glossary [19th February, 2016]

World Heart Federation 2016, Stroke. Available from: http://www.world-heart-federation.org/cardiovascular-health/stroke/. [14 March 2016].

Waugh, A. and Grant, A., 2014. Ross & Wilson anatomyandphysiology in health andillness. Elsevier Health Sciences.

Wu, C.Y., Lin, K.C., Wolf, S.L. andRoby-Brami, A., 2012. Motor rehabilitationafterstroke. Stroke research and treatment, 2012.

Wu, E.H.K., Tseng, C.C., Yang, Y.Y., Cai, P.Y., Yen, S.S. and Chen, Y.W., 2014. Cross-Platform and Light-Weight Stroke Rehabilitation System for New Generation Pervasive Healthcare. In Advanced Technologies, Embedded and Multimedia for Human-centric Computing (pp. 1269-1277). Springer Netherlands.

Webster, D. and Celik, O., 2014. Systematic review of Kinect applications in elderly care and stroke rehabilitation. Journal of neuroengineering and rehabilitation, 11(1), p.1.

Yang, L., Zhang, L., Dong, H., Alelaiwi, A. and El Saddik, A., 2015. Evaluating and improving the depth accuracy of Kinect for Windows v2. Sensors Journal, IEEE, 15(8), pp.4275-4285.

Yang, U. and Kim, G.J., 2002. Implementation and evaluation of “just follow me”: An immersive, VR-based, motion-training system. Presence: Teleoperators and Virtual Environments, 11(3), pp.304-323.

Yeh, S.C., Lee, S.H., Wang, J.C., Chen, S., Chen, Y.T., Yang, Y.Y., Chen, H.R., Hung, Y.P., Rizzo, A. and Tsai, T.L., 2013. Stroke rehabilitation via a haptics-enhanced virtual reality system. In Advances in Intelligent Systems and Applications-Volume 2 (pp. 439-453). Springer Berlin Heidelberg.

Zhou, L., Liu, Z., Leung, H. and Shum, H.P., 2014, November. Posture reconstruction using Kinect with a probabilistic model. In Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology (pp. 117-125). ACM.

Zhou, H. and Hu, H., 2008. Human motion tracking for rehabilitation—A survey. Biomedical Signal Processing and Control, 3(1), pp.1-18.

Zhu, L. and Wong, K.H., 2013. Human tracking and counting using the KINECT range sensor based on Adaboost and Kalman filter. In Advances in Visual Computing (pp. 582-591). Springer Berlin Heidelberg.

Zhu, Y., Yang, Z. and Yuan, B., 2013, April. Vision based hand gesture recognition. In Service Sciences (ICSS), 2013 International Conference on (pp. 260-265). IEEE.

# APPENDIX A

## EXPERIMENTAL PROTOCOL FOR KINECT V2 VISUAL GESTURE BUILDER GESTURE DATABASE IN STROKE REHABILITATION USING THE SNAP ADL TOOLKIT

### 1. PURPOSE:

This experiment is aimed at generating data used in a gesture database which will be used to determine the accuracy and reliability of the Visual Gesture Builder database of the Kinect v2 in stroke rehabilitation.

The experiment would be based mainly on items from the Southampton Hand Assessment procedure (SNAP).

### 2. MATERIALS:

The materials required will include;

a) Kinect v2 sensor

b) Height adjustable chair

c) Height adjustable table

d) Southampton Hand Assessment Protocol (SHAP) testing kit.

e) Tripod for Kinect mounting

f) Measuring tape

g) Towel

**3. METHODS:**

Experimental setup:

a) Subject posture: seated in chair encouraging an erect posture and feet flat on the floor.

b) Height of table: variable height but at the level of the last costal rib.

c) Distance from table: Subjects elbow comes to the table edge.

d) Hands: resting on the table.

e) Kinect: Mounted on 60 cm above the height of the table and 150-200cm from subject.

It is expected that 20 individuals will perform 3 repetitions of each of the selected SHAP manipulative Activity of Daily Living (ADL) gestures to ensure sufficient training and testing data capture. The whole experiment will last 10 - 15 minutes per person, including the time taken to explain the procedures.

**4. EXPERIMENTS:**

In the normal SHAP experiments, the test subject self times, however, in this study, the objective is the collection of the skeletal data while the manipulative action is carried out therefore the timing aspect of the testing will not be carried out. A self paced speed will be recommended.

**A. Pick Up Coins**

- **Preparation:** Arrange the two 2p and two 1p coins in the designated areas on the board. Place the glass jar in the designated spot for this task with the lid removed.
- **Instruction**: Pick up each coin in turn by sliding the coin to the edge of the board using a tip or tripod grip and drop each coin into the glass jar. Move from right to left.
- **Repeat**: Reset the task for the participant.

**B. Remove Jar Lid**

- **Preparation:** The lid should be placed on the empty glass jar and tightened only with sufficient force as would be expected for everyday use/self storage. The jar should be placed in the designated area on the form board.

- **Instruction**: Both hands should be used for this task. Pick up the jar using a power grip with the non-dominant hand, undo the lid and return both the jar and the lid to the designated areas on the platform.
- **Repeat**: Reset the task for the participant.

**C. Pour Water From Jug**

- **Preparation:** Fill the glass jug with 100ml of water (100ml is marked on the jug). Place the jug in the designated area of the form board with the handle of the glass jug pointing the right for right-handed participants, and to the left for left-handed participants. Place the glass jar (without the lid) on the designated left area for right-handed participants and the right for left-handed participants.
- **Instruction**: Lift the glass jug by the handle using a lateral grip and pour the water into the glass jar.
- **Repeat**: Reset the task for the participant.

**D. Pour Water From Carton**

- **Preparation:** Empty the glass jar from the previous task and replace the jar in the same position on the form board. Fill the carton with 200ml of water (measured out in the glass jug). Place the carton in the designated area on the form board with the spout of the carton pointing toward to glass jar (according to the handedness defined for the previous task).
- **Instruction:** Pick up the carton using a power grip and show how to pour the water into the glass jar.
- **Repeat**: Reset the task for the participant.

**E. Move A Full Jar**

- **Preparation:** Fill the glass jar with water to the top of the label and tighten the lid. Place the jar in the designated area on the form board, on the left side of the board for right-handed participants and the right side of the board for left-handed participants. Place the empty carton lengthways along the middle of the form board (without obstructing the timer unit) to create a barrier.
- **Instruction:** Lift the jar over the carton using a power grip and place on the opposite side of the form board in the designated area.

- **Repeat**: Reset the task for the participant.

**F. Move An Empty Tin Can**

- **Preparation:** Place the empty tin (with the plastic lid on) in the same position on the board as defined for the jar in the previous task and keep the carton in the same position on the form board creating a barrier.
- **Instruction:** Lift the tin over the carton using the power grip and place on the opposite side of the form board in the designated area.
- **Repeat**: Reset the task for the participant.

## 5. CONTROLS:

In the development and testing of a Kinect Gesture Builder database, it is recommended to use 2/3 of the collected data for training and the 1/3 data for analysis. This grouping of data will be randomly done to eliminate any bias.

## 6. CONTENT OF THE SNAP KIT

| QUANTITY | ITEM | UTILIZED |
|---|---|---|
| 1 | Test case containing all SHAP equipment | |
| 1 | Backboard mounted in case with lock & key, door hand and zip | |
| 1 | SHAP form-board | ✓ |
| 1 | Foam insert containing all objects | |
| 1 | Timer unit | |
| 6 | Lightweight abstract objects | |
| 6 | Heavyweight abstract objects | |
| 1 | Lock and key mounted on backboard | |
| 1 | Zip mounted on backboard | |
| 4 | Coins (2 x 1p and 2 x 2p) | ✓ |
| 1 | Button board with 4 buttons attached | |
| 1 | Plasticine block | |
| 1 | Knife | |
| 1 | Note card | |
| 1 | Glass jar with lid | ✓ |
| 1 | Glass jug | ✓ |
| 1 | Cardboard juice carton | ✓ |
| 1 | Empty tin with plastic lid | ✓ |
| 1 | Door handle mounted on backboard | |
| 1 | Metal arrow unit | |
| 1 | Screwdriver | |

## APPENDIX B

## CONSENT TO PARTICIPATE IN AN EXPERIMENT (DUTCH)

TOESTEMMINGSVERKLARING

Voordeelnameaan het wetenschappelijkonderzoek:

**Titelonderzoek:Hoebetrouwbaarkan de Kinect v2 arm- enhandbewegingenherkennen?**

**Verantwoordelijkeonderzoeker:**Claudine Lamoth, AlessioMurgiaenKishakCinfwat

- Ik verklaar op een voor mij duidelijke wijze te zijn ingelicht over de aard, methode, doel en [indien aanwezig] de risico's en belasting van het onderzoek. Ik ben in de gelegenheid gesteld om vragen over het onderzoek te stellen en mijn vragen zijn naar tevredenheid beantwoord.
- Ik begrijp dat ik mijn deelname op ieder moment, om wat voor reden dan ook, mag en kan beëindigen zonder dat hieraan enige consequenties verbonden zijn.
- Ik weet dat de gegevens en resultaten van het onderzoek alleen geanonimiseerd en vertrouwelijk aan derden bekend gemaakt zullen worden.
- Ik geef toestemming dat bevoegde personen van het Centrum voor Bewegingswetenschappen

inzage kunnen krijgen in mijn gegevens en onderzoeksgegevens.

- Ik geef toestemming om de gegevens te verwerken voor de doeleinden zoals beschreven in de informatiebrief.
- **Ik stem toe met deelname aan het onderzoek.**

Naam :

Geboortedatum :

Handtekening : Datum

-----------------------------------------------------------------------------------------------------------

*In te vullen door onderzoeker:*

Ondergetekendeverklaartdat de hierbovengenoemdepersoonzowelschriftelijkalsmondeling over het bovenvermeldeonderzoekgeïnformeerd is. Hij/zijzalresterendevragen over het onderzoeknaarvermogenbeantwoorden. De deelnemerzal van eeneventuelevoortijdigebeëindiging van deelnameaanditonderzoekgeennadeligegevolgenondervinden.

Naam :

Functie :

Handtekening : Datum