Development of an Instrumented and Powered Exoskeleton for the Rehabilitation of the Hand

by

Puya Peter Abolfathi

BSc., BEng., Hons (Biomedical) , Flinders University, 1998

Thesis presented for the degree of Doctor of Philosophy

University of Sydney

School of Aerospace, Mechanical and Mechatronic Engineering

2007

Dedication

A mi amor Lorena, por darme tanto de ti.... gracias por todo....

Abstract

With improvements in actuation technology and sensory systems, it is becoming increasingly feasible to create powered exoskeletal garments that can assist with the movement of human limbs. This class of robotics referred to as human-machine interfaces will one day be used for the rehabilitation of paralysed, damaged or weak upper and lower extremities. The focus of this project was the development of an exoskeletal interface for the rehabilitation of the hands. A novel sensor was designed for use in such a device. The sensor uses simple optical mechanisms centred on a spring to measure force and position simultaneously. In addition, the sensor introduces an elastic element between the actuator and its corresponding hand joint. This will allow series elastic actuation (SEA) to improve control and safely of the system. The Hand Rehabilitation Device requires multiple actuators. To stay within volume and weight constraints, it is therefore imperative to reduce the size, mass and efficiency of each actuator without losing power. A method was devised that allows small efficient actuating subunits to work together and produce a combined collective output. This work summation method was successfully implemented with Shape Memory Alloy (SMA) based actuators. The actuation, sensory, control system and human-machine interface concepts proposed were evaluated together using a single-joint electromechanical harness. This experimental setup was used with volunteer subjects to assess the potentials of a full-hand device to be used for therapy, assessment and function of the hand. The Rehabilitation Glove aims to bring significant new benefits for improving hand function, an important aspect of human independence. Furthermore, the developments in this project may one day be used for other parts of the body helping bring human-machine interface technology into the fields of rehabilitation and therapy.

Acknowledgements

I would like to extend my gratitude to my PhD supervisor, Dr. Tim Scott for bringing me into the Quadriplegic Hand Research Unit and giving me this life changing opportunity. Thank you also Veronica Vare for your support, laughter and friendship.

Throughout the years in this project, I have been financially supported by many sources. I would like to extend my sincere appreciation in chronological order. In my first year I was supported by a scholarship from the QHRU and Dr. Scott's NH&MRC Development Grant. In my second year I was awarded University of Sydney UPA and a Centenary Foundation top-up. In subsequent years I was supported by the School of Aerospace, Mechanics and M'tronics Engineering and most recently funding from the Spinal Cord Injuries Unit Research Fund. In particular, I would like to thank Dr. Sue Rutkowski for her constant and kind support. In 2004, I was honoured to be the recipient of the Eureka Prize for Inspiring Science, sponsored by the British Council and the Australian Museum. Thanks to this prize, I was able to take an exciting study tour in the United Kingdom. This trip was instrumental in demonstrating the Rehabilitation Glove to potential collaborative partners and leading to the currently ongoing commercialisation of the technology. On this note, I would like to extend my appreciation to Dr. Barend Ter Haar and his team for believing in our project and for their solid support. Thank you also to Sandra See and Russel Carrington for their tireless efforts on the commercialisation front.

During the years spent on this project, a number of people have volunteered their time and hands courageously to help me with my contraptions, experiments, demonstrations and trials. Thank you all for your interest and vital participation.

This project and thesis would not have been completed without the constant technical (and emotional!) support and much needed advice of colleagues whose friendship I have thoroughly enjoyed. Firstly, I would like to convey my deepest gratitude to Rob Wilkins who has been a mentor in all things mechanical and with whom I share a love for ingenious mechanisms. I cannot thank enough the wonderful people at Biomedical Engineering particularly Bill Fisher, Clifford Stephenson, Fred Peterson, Mark Witcomb, Andrew Palmisano, Rodney Torrens, Mario Todaro, Van Doai, Michael Chang, Alan Steirn, Paul Holman and Sharon Leavers. David Roffe, Ray Jongs, Kevin Brigden and Eddy Jozefiak also deserve special mentions.

Finally, and with much emotion, I would like to thank my dear family and friends who really shared this entire journey with me. Maman and Baba, all my life you instilled in me the importance of education with so much dedication, love and patience. You are the best parents in the world. Pupak, my sister, you were there from the beginning and you are always there with an unconditional love that has supported me so much. Thank you to you and Behrad. During the term of this PhD, I was blessed with the birth of my beautiful nieces, Kiana and Elena and my beautiful nephew Evaristo, who constantly inspire me with the magic of life. My extended Uruguayan family have been a constant source of love and support. Gracias, queridos. I am blessed with the most wonderful and supportive friends. In the order of the longest number of hours spent talking about my PhD I would like to thank Simon, Margot, Brent, Carmine, Olivia, Basil, Geoff and of course Jill. You are like family to me.

Last and most importantly, I would like to thank the love of my life, Lorena Ressano Fernández without whom I certainly would not have made it this far. You have been my backbone, my inspiration, my teacher and my best friend. Te camelo.....

Intellectual Property Arising from this Work

Arising from the work presented in Chapter 6:

ABOLFATHI, PP; SCOTT, TRD; VARE, V; (2007) Force and Movement Summation Device. Prov. Australian Patent Application no. 2006900944 (filed 28 February 2007)

Containing material presented in Chapter 3 and Chapter 7:

ABOLFATHI, PP; SCOTT, TRD; VARE, V; WALLACE, GG; SPINKS GM and ZHOU, D (2004) Movement Facilitation Device. International Patent no. WO 2004/021936 A1

Communications

Due to the commercial nature of this project, a greater focus was placed on securing intellectual property than in the publication of the details of the technology developed. However, aspects of the project have been disclosed and presented in the following:

ABOLFATHI, PP; SCOTT, TRD; VARE, V; (2004) A novel rehabilitation glove with embedded artificial muscles enabling specific therapeutic and functional options. The 11th Meeting of the Combined Orthopaedic Associations – Sydney

ABOLFATHI, PP; SCOTT, TRD; VARE, V.(2004) Provision of Hand Opening and Closing to the Paralysed Upper Extremity of People with Spinal Cord Injury Using a Novel Instrumented Glove. XXIst Annual Scientific Research Meeting – Royal North Shore Hospital, Sydney

ABOLFATHI, PP; SCOTT, TRD; VARE, V; WALLACE, GG and SPINKS GM. (2003) New technology for rehabilitation of the hand using artificial muscles. XXth Annual Scientific Research Meeting – Royal North Shore Hospital, Sydney

Table of Contents

List of Figures

List of Tables

Glossary

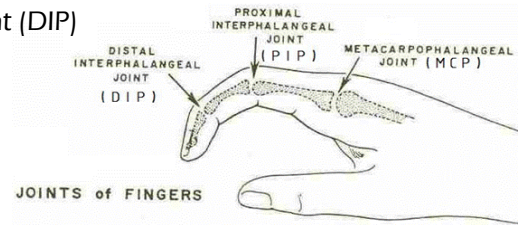
Active Movement – Movement produced by own muscle activity

AROM – Active Range of Motion

CPM – Continuous Passive Motion

EAP – Electro-active polymers

Finger joints and acronyms: Metacarpal Phalangeal joint (MCP), Proximal Inter-phalangeal joint (PIP), Distal Inter-phalangeal joint (DIP)



FLPC – Force Limited Position Control

FPT – Force-Position Transducer

FSR – Force Sensing Resistor

JCHAR – Joint characterisation (a term defined in this thesis)

LVDT – Linear Variable Differential Transformer

NI – National Instruments

Passive Movement – Movement produced by an external force

PLFC – Position Limited Force Control

PWM – Pulse Width Modulation

SCI – Spinal Cord Injury

SEA – Series Elastic Actuation

Smart materials (smart structures, intelligent materials) – terms used in the field of material science to refer to structures that can change shape (e.g. shorten) following the delivery of some form of energetic stimulus (e.g. heat or electricity) for use as a sensor, an actuator or both.

Stiction – A threshold of force caused by the static cohesion between two touching surfaces. Stiction must be overcome before relative movement between surfaces can occur. Stiction is a threshold, not a continuous force.

Tele-actuation – A type of actuation in which the actuator is remote to the end effector. Examples include cable-based actuator and Bowden cables.

TERT – Total End Range Time

TROM – Torque Range of Motion

Chapter 1 Introduction

A healthy working hand is most important in providing independence in an individual's life. The loss of hand function results in a severe compromise of the ability to feed and care for oneself, and limits one's participation in work, social and family life (Scott, Vare et al. 2003). There are many injuries or conditions that can result in loss of hand function. These include trauma, burns, formation of non-functional scar tissue following hand surgery, oedema, arthritis, stroke, cervical spinal cord injury and peripheral nerve injury.

"Stroke, traumatic brain and spinal cord injuries are a leading cause of disability and handicap in the industrialised world. Stroke patients are the largest group, with 750,000 individuals affected in the USA each year, the prevalence being 200-300 patients per 100,000 inhabitants (Hesse, Schmidt et al. 2003).

Following recent breakthroughs in material science, dynamic actuators have been produced that can contract in a similar way to muscles when given a stimulus. This project was inspired by the enormous potential of these "Artificial Muscles" as actuators in medical applications, opening the path to a new era of developments in healthcare technology.

The original and specific aim of this project was to create a portable, instrumented, electro-mechanical glove that can control the movement of hand joints for the provision of therapy, hand function and assessment. The main focus in this endeavour has been to investigate the use of new actuators, combined with novel sensors and mechanical components to produce a device likely to make a significant therapeutic impact for those suffering from loss of hand function. Broadly however,

the principles investigated in this project will also apply to the creation of powered exoskeletal human-machine interfaces that can work in parallel to human movement.

1.1 A brief history behind this project

In 1998 while listening to a radio program, Dr. Timothy Scott, director of the Quadriplegic Hand Research Unit at the RNSH, discovered that scientists at Wollongong University had began working on special polymers with muscle-like characteristics. Considering the enormous potential such material would present for the rehabilitation of damaged or paralysed hands (as well as other limbs), he immediately contacted the Intelligent Polymer Institute (IPI) where this research was being conducted to enquire about the technology. The result was a collaborative link between the IPI and the Quadriplegic Hand Research Unit and a provisional patent describing a Rehabilitation Glove with embedded artificial muscles to passively move the wearer's fingers in a controlled manner. The Rehabilitation Glove was envisioned as a device that could be used for therapy, function and assessment of the hand, promising a revolution in the field of hand therapy.

The initial support to design and develop such a device was provided by hospital grants, which allowed the first proof of concept. Given the early success, in 2002 the project was awarded an NH&MRC Development grant to bring the idea and technology to a commercial level. Throughout this time, the intelligent polymers of Wollongong University were being advanced for possible use as artificial muscles. However, as the technology was still a few years from fruition, alternative actuation technologies had to be incorporated in the Rehabilitation Glove to produce a working prototype.

The development of the Rehabilitation Glove was turned into a PhD project in 2003 and is the subject of this thesis. In 2004, the project was supported further by the award of the Australia Museum/British Council Eureka Prize for Inspiring Science.

Following the securing of additional patent applications and pursuit of commercial partnership by the Royal North Shore Hospital, a UK based medical device company agreed to licence the technology and bring the Rehabilitation Glove from a laboratory prototype to a commercial device for global use. It is hoped that the first Rehabilitation Glove will enter the market in 2009.

1.2 Problem statement

The goal of this project is to develop an external exoskeletal harness that is able to passively[*] mobilise the many joints of the hand in a safe and controlled manner. In essence, this device will encompass intricate mechanisms which must work parallel to human parts via a human-machine interface. To achieve this endeavour, a number of technical challenges must be overcome. These are listed below the broad objective.

- Broadly, the device must allow hybrid control of torque and angle of human joints via a safe and parallel[†] human machine interface. This will require:
    - The movement of multiple joints independently[‡] in a multi-link system using a portable harness.
    - Multiple actuators. Current technology is too bulky and inefficient to be able to fit multiple actuators within the limited space available around the hand and forearm.
    - Multiple torque and angle sensors.
    - Flexible /non-rigid coupling between human and machine parts.
    - A specialised control system that can combine torque control with angle control.

---

[*] Passive movement refers to movements caused by external forces (as opposed to active internal forces from muscles)

[†] Here, parallel means that both the human and the machine can work actively together.

[‡] This means that each joint can be controlled independent to the angle of its preceding or following joints in the multi-linked structure.

1.3 Layout of the thesis

The proposed technology can be separated in to the following components:

1. Actuator
2. Sensors – used for torque and angle feedback to the control system
3. Control system
4. Multi-linked exoskeletal human-machine interface

The interaction between the above components is illustrated in the block diagram in Figure 1.1.

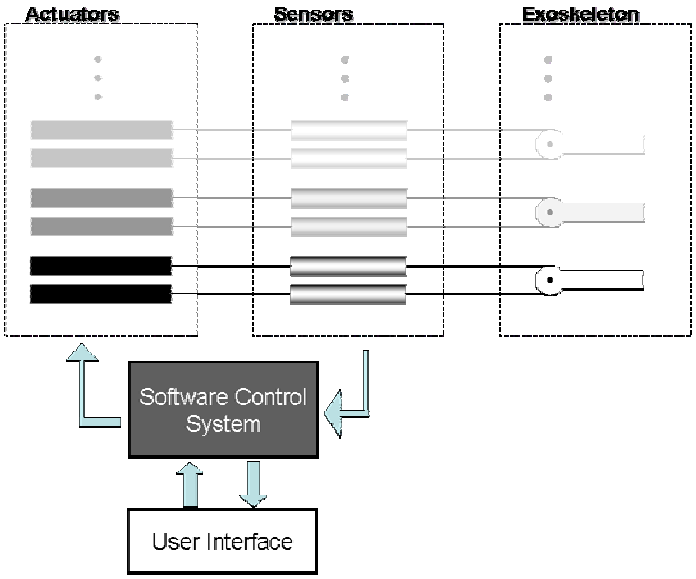


Figure 1.1 Components of the proposed system.

The thesis is divided into three main studies. Each study chapter is preceded by a chapter reviewing relevant literature and background.

Chapter 2 and Chapter 3, tackle the problem of providing torque and angle sensors appropriate for the proposed technology. In Chapter 3, a novel hybrid force-position

sensor is proposed. The sensor presented also inherently reduces the rigidity of the human machine interface through its series elastic component, an important development for safe and effective control.

Chapter 4 and Chapter 5 engage in the actuator problem. The focus of the review section (Chapter 4) and the study (Chapter 5) is the improvement of a genre of actuators based on special materials. The ultimate aim of this study is to create actuators that can contract and expand in muscle-like manner and exhibit comparable efficiencies. Such developments would allow multiple miniature actuators to work together within the given weight and volume constraints of the envisioned device.

Chapter 6 focuses on the current technology used for rehabilitation of hands. Chapter 7 aims to bring together concepts developed and presented in the previous chapters to evaluate the effectiveness of the proposed technology. Here, a limited (single joint) human–machine interface is tested with volunteer subjects. A hybrid torque-angle controlled algorithm is used to emulate potentially and clinically useful methods for assessment, therapy and function.

Although originally inspired by the advent of artificial muscle technology, in its final form, the Rehabilitation Glove may or may not incorporate such materials for actuators (it may instead utilise more standard motors). More importantly, the focus of this work is the development of a comprehensive therapy and function tool for the hand. This is an ambitious project that utilises many mechanical, electronic and software based solutions to consolidate various existing and new forms of therapy, functional aid and assessment into a single device.

Chapter 2 Control of Force and Position in Human-Machine Interfaces

2.1 Introduction

The advancement of material science and actuation technology is expected to lead to the rise of a new generation of low weight, high energy density, efficient actuators, whose size, function and effectiveness could compare with natural muscles (Hunter and Lafontaine 1992; Klute, Czerniecki et al. 2002; Bar-Cohen 2004; Madden, Vandesteeg et al. 2004; Baughman 2005; Spinks, Campbell et al. 2005; Ebron, Yang et al. 2006; Madden 2006). These actuators, collectively referred to as 'artificial muscles', may eventually be used to replace paralysed muscles in the human body to restore function. Potential beneficiaries include those with spinal cord injury, stroke, peripheral nerve damage, neurodegenerative disorders and cerebral palsy. Such actuators are not yet sufficiently advanced for implantation into the human body (see Chapter 4). It is likely that, as a milestone towards such an aim, these actuators will be placed externally in exoskeletal garments. This application will endeavour to restore function as well as provide new modes of therapy, dynamic splinting, rehabilitation, and enhancement of strength (Scott, Vare et al. 2003; Abolfathi, Scott et al. 2006).

In order for the effectiveness of the 'artificial muscle' type actuator to approach that of real muscle, it will be necessary for sophisticated means of powering and controlling of the actuator to be achieved (Jacobsen, Olivier et al. 2004). In Chapter 3, a novel, hybrid force and position transducer (FPT) developed as part of this study and suited to the control of artificial muscles and exoskeletal systems is evaluated.

The technology presented here also targets another field that is presently in the process of rapid advancement: Rehabilitation Robotics, also known as 'Assistive Robotics'. While inclusive of the exoskeletal components referred to above, this wide field also involves the application of human-controlled or autonomous robots to assist in medical rehabilitation and daily living. Rehabilitation robots can be used to enable severely disabled people to access and manipulate objects without assistance from another human (Prior and Warner; Hagan, Hillman et al. 1997). However, due to high cost, difficulty in control, safety concerns and flexibility issues, such technology is not widely used (Prior and Warner 1990; Zinn, Khatib et al. 2002). In any environment where machines interact with humans, there is a need to meet rigorous safety standards. This is achieved by using appropriate sensors, overlapping safety mechanisms and the implementation of robust control of the robots (Kawamura, Bagchi et al. 1995; Noritsugu and Tanaka 1997; Nef and Riener 2005).

2.1.1 Existing technology

There are currently a number of teams around the world working to provide safe, low cost and easy to use robotic devices to assist patients with therapeutic exercises and daily living tasks. With each of these groups, movement feedback and control of the robots has presented significant challenges (Prior and Warner 1990; Hagan, Hillman et al. 1997; Abolfathi, Scott et al. 2004; Jiping, Koeneman et al. 2005; Nef and Riener 2005; Romer, Stuyt et al. 2005).

The following sections give a brief overview of technologies and developments relevant to the ideas presented in this chapter.

2.1.1.1 Adaptation of traditional robotics to the rehabilitation field

Traditionally, robots have been most common in the manufacturing and industrial areas. In such settings, robotic operations normally involve repetitious tasks at high speed and high precision. Accurate position control is the predominant criterion of these operations. Conventional robots typically employ stiff mechanisms including highly geared electric motors, stepper motors, hydraulic actuators or pneumatic motors. Industrial applications are often in strictly controlled environments where this stiffness does not pose a problem (and is a desired property). But safe and effective interaction of machines with humans requires robotic components to deal flexibly with interfacing to human parts. Adaptation of robotic systems to human-centred application requires a shift of focus from position control, towards force control and ideally a combination of both parameters (Pratt and Williamson 1995; Noritsugu and Tanaka 1997; Nef and Riener 2005). Furthermore, multiple strategies need to be employed within their mechanical, electrical and software architecture to ensure a successful blend of performance and safety (Kawamura, Bagchi et al. 1995; Zinn, Khatib et al. 2002).

2.1.1.2 Series Elastic Actuation

The Series Elastic Actuator (SEA, described by Pratt and others in 1995) has emerged as promising mechanical design architecture for use in rehabilitation robotics systems. The SEA is composed of an elastic element placed in series between an actuator and the load. A force transducer coupled to the elastic component provides feedback to the control system. The control system moves the motor to reduce the error between the desired and the measured forces (Pratt and Krupp 2004). The main components of a Series Elastic Actuator are presented in Figure 2.1

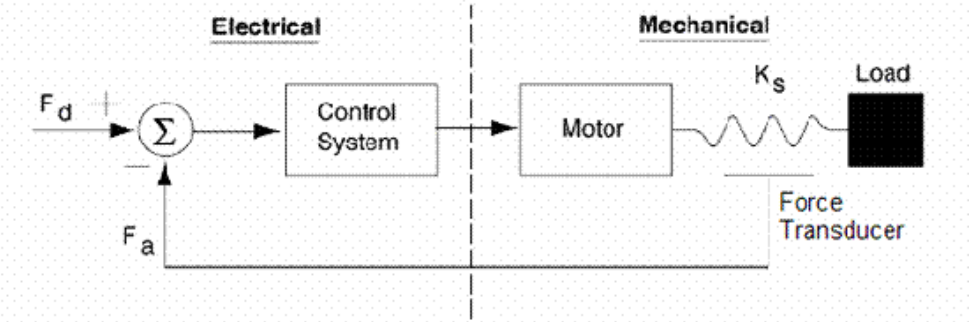


Figure 2.1 SEA Actuators. Schematic diagram showing components of a Series Elastic Actuator (adapted from Pratt and Krupp 2004).

Advantages of SEAs include:

1. Reliable force control with increased stability.
2. Simplicity
3. Robustness
4. Reduced reflection of motor inertia on the load (in contrast to stiff actuator-load coupling)
5. Reduced probability of damage to the motor during inadvertent contact with rigid obstacles
6. Potential for energy storage.

Disadvantages:

1. Low frequency bandwidth of force output (due to the elastic component acting as a mechanical low pass filter)
2. The reduction of stiffness may impede applications where stable position control is critical.

Since its conception the SEA has drawn wide interest from many areas. A number of actuators and robots, such as the MARIONET (Sulzer, Peshkin et al. 2005), COG (Brooks and Stein 1994), a one-D.O.F robot arm (Morita and Sugano 1995) and the Spring Turkey (Robinson, Pratt et al. 1999), use SEAs for control. The majority of SEA applications have been focused on creating tools for rehabilitation robotics and other human-machine uses.

The incorporation of an elastic component in the sensor described in Chapter 3 enables the utilisation of the principles of series elastic actuation in the design of the hand rehabilitation device.

2.1.1.3 Cable-drive actuation systems

There are many techniques for producing torque at a joint. In animal bodies, this is achieved by antagonistic musculo-tendon actuators (Caldwell and Tsagarakis, 2002). Traditional robots often use actuators placed directly on a joint. In contrast to direct-drive actuators, cable-drive systems are highly advantageous to human-machine interactions. Cable-drive systems are able to produce torques on multi-linked joints remotely without additional weight, inertia and end-effector volumes[*].

Lightweight, remotely actuated cables can span multiple joints if they pass through the centres of proximal joints[†] or if they are housed in sleeves (as in Bowden cables)[‡].

---

[*] End-effector refers to the end section of a multi-linked system where interaction with the external environment takes place (e.g. the gripper at the end of a robotic arm).

[†] To minimise torque on proximal joints and permit independent actuation of multi-link joint

[‡] Bowden cables are commonly used in prosthetic arms.

A number of cable-driven human-machine interactive devices have been developed. Examples include the MARIONET (Sulzer, Peshkin et al. 2005), UTAH-MIT hand (Jacobsen, Iversen et al. 1996), the STRING-MAN (Surdilovic, Bernhardt et al. 2003), the WAM (Salisbury, Eberman et al. 1989), the Phantom™[*], The SpringWalker™[†].

The following chapter presents the development and evaluation of a novel sensor, designed for use in human-machine interfaces.

[*] A commercial haptic device available from Sensable Inc. http://www.sensable.com

[†] Commercial device available from Applied Motion Inc. http://www.springwalker.com/

Chapter 3 Development of a Hybrid Force-Position Transducer

3.1 Introduction

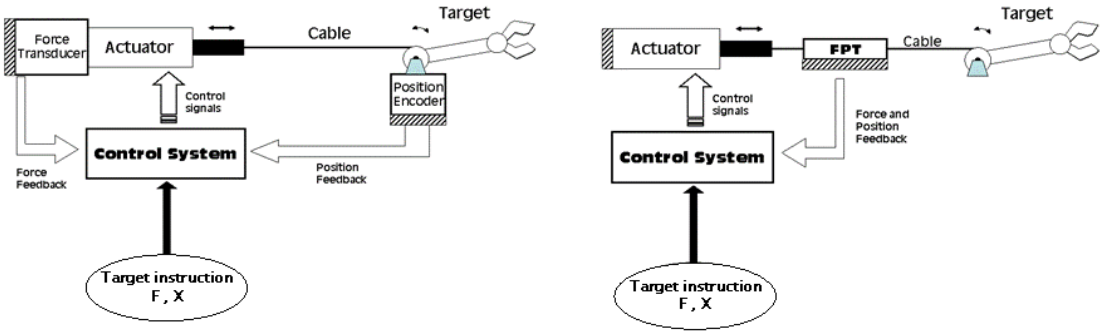
A novel hybrid sensor (referred to as the Force-Position Transducer or FPT) has been designed to provide position and force feedback for the control of actuator systems. Its operation is evaluated in detail. The FPT is particularly suited as a feedback sensor to indicate force (tension) and position for the control of cable-drive actuation systems. The compliant element present in the FPT will cause any cable-drive actuator to become a Series Elastic Actuator (SEA), an actuation system which is gaining increasing currency in rehabilitation robotics (Brooks and Stein 1994; Morita and Sugano 1995; Pratt and Williamson 1995; Robinson, Pratt et al. 1999; Sulzer, Peshkin et al. 2005). In this study, an FPT prototype was developed and evaluated for use in isotonic position control, isometric force control and combined force-position control. The results showed that the FPT is suited to simultaneous control of position and force in dynamic actuations involving non-controlled environments (such as those present with human-machine interfaces). The accuracy[*] of this early prototype was around ±2% (full-scale of 20mm) for position and ±10% f.s (full-scale of 4.9N) for force measurements. This is expected to improve significantly with future enhancements in manufacture of the device. Given the potential benefits of the FPT as a hybrid sensor and its inherent suitability to Series Elastic Actuation, the device is well positioned to play a role in future human-machine interfaces. Such systems include rehabilitation robots and exoskeletons for the enhancement of strength and provision of function and therapy.

---

[*] The evaluation of accuracy in this thesis has been presented in reference to full-scale for both force and position. Full-scale was considered more appropriate over standard deviation for reasons outlined in section 3.3.2.1 (page 25)

3.1.1.1 Traditional sensors used in control systems

In many typical force or position control systems, feedback information is obtained by placing individual force and/or position sensors directly on the load or on the actuator (see Figure 3.1). Typical position sensors include optical encoders, Hall-Effect sensors (for very small movements) and Linear Variable Differential Transformers (LVDTs). For force, sensors include Strain Gauges or Force Sensing Resistors (FSRs). The advantage of such sensors is the accuracy they provide (given stable conditions and proper calibration). The undesirable features common to many types of sensors are high complexity, high cost and need for signal conditioning. Although this may not be a significant impediment in some cases, for applications that require multiple sensors, the combined cost and complexity can become a considerable burden in the overall system (Prior and Warner 1990; Zinn, Khatib et al. 2002).



a) Typical Control System **b) FTP-Based Control System**

Figure 3.1 Typical and FTP based systems *a) The typical system used to control the movement and force provided to a target. The position is measured directly at the target load and the force is measured either at the target or at the actuator. The control system can be implemented in hardware or software and receives instruction for target position and force from an external source. b) The same control system algorithm receiving force and position feedback simultaneously from the hybrid Force-Position Transducer.*

3.1.2 Problem statement

There is a necessity for rehabilitation robotics to advance further and find wider use before they can bring benefits into people's lives (Kawamura, Bagchi et al. 1995; Miller 1998; Romer, Stuyt et al. 2005). Commercialisation is significantly impeded by the current cost of this technology (Prior and Warner 1990; Romer, Stuyt et al. 2005). The cost of rehabilitation technology is hoped to be decreased by the reduction of complexity without loss of control capabilities.

The FPT was initially developed to meet the requirements of a rehabilitation device created to provide therapy, assessment and function for the damaged or paralysed hand. As an upper-extremity exoskeleton, the "Rehabilitation Glove" was designed to be portable, affordable, cosmetically appealing and able to provide force and position control and feedback for all the joints in the hand (Abolfathi, Scott et al. 2006). The Rehabilitation Glove is based on a cable-drive actuation system that requires force and position control of up to 17 separate cables. The sensor requirements of the device are of the greatest challenges in the project. Following an early feasibility study, it was decided that currently available force and position sensors would over-complicate the device and its circuitry and make it too expensive as a rehabilitation product. The solution to the sensor challenge was the hybrid Force-Position Transducer (FPT) described in this paper.

3.1.3 Aims of the hybrid Force-Position Transducer (FPT)

The aim of the FPT project was to create an integrated device for use in cable-drive actuation with the following combined properties:

1. Measure the position of the load
2. Measure the tension in the cable
3. Provide a compliance in the cable for series elastic actuation

Furthermore, the FPT was designed to meet the following specific criteria:

- Introduce minimal friction on the cable and its actuation.
- Be rugged, reliable and provide reasonable accuracy and resolution for the safe and reliable control of hand joints.
- Be simple and cheap to produce and require minimal signal conditioning and circuitry.
- Consume minimal power.
- Able to be miniaturised to occupy minimal volume and weight in the exoskeleton.

3.1.4 The basic principles of the FPT

The FPT is placed in series with a cable-drive system between the actuator and the load. The central component of the FPT is a spring within a fixed chamber. Cables connect each end of the FPT spring with the actuator and load respectively. The spring, which is free to move with the cable, changes it position and length in response to displacement and tension of the cable respectively (see Figure 3.2).

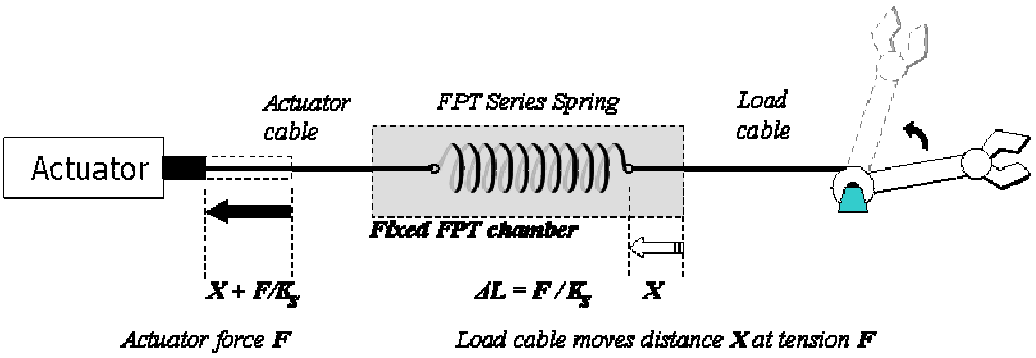


Figure 3.2 Basic principles of the FPT. The FPT spring with has an elastic coefficient K_S. When placed in series with the actuating cable the actuator must move by amount equal to the load's displacement plus F/K_s where F represents the cable tension.

The fixed FPT chamber contains novel optical sensors (see the following section) to measure the absolute positions of the two ends of the spring. The position of the load-facing end of the spring is equivalent to the position of the load. Cable tension is proportional to the length change of the FPT spring. This can be derived from knowing the positions of both ends of the spring. A detailed description of the basic principles of the FPT is presented in Appendix A.

3.2 Design, construction and calibration of the FPT

The greatest challenge in the development of the FPT was the implementation of an appropriate sensing mechanism to measure the positions of the spring-ends. Ideally, the sensing mechanism must have minimal contact with the moving components and be impervious to electromagnetic noise and ambient factors (such as temperature change). Following a feasibility study, a form of analogue optical sensing was conceived. Other methods considered but not selected due to the listed drawbacks included:

- Potentiometers – Involve high friction and display low accuracy.
- Hall Effect sensors – Suited only to short distances and require complicated circuitry.
- Linear Variable Differential Transformers – Though very accurate, are highly expensive, contain complicated circuitry and are affected by magnetic noise.
- Optical Encoders – Require pulse counting. Over time, pulse counts may be lost leading to accumulating errors.

The designed mechanism used polarising filters to attenuate light in proportion to displacement. The method was highly sensitive without the need for amplification.

In addition, the mechanism was designed to output large DC signals linearly proportional to the positions of each end of its spring. The concept required a miniature light source, a light receptor and a parallel pair of polarised film discs, acting to attenuate the light transmitted through them in proportion to linear movement of the spring. This was achieved by coupling the movement of the spring ends to rotation* of the polarised films.

3.2.1 Polarised disc method

Two polarised discs were made using photographic polarised filters cut into circles of 10mm diameter and held within acrylic (Perspex) ring frames (see Figure 3.3a). The rings were held concentrically and parallel to each other within chambers using ball bearings to reduce friction and allow rotation. The protruding pins were thus able to engage the cut away walls of an outer linear slider (referred to as the slider window). The discs were spring biased away from each other to the effect that their pins were forced to ride against opposite walls of the slider window at all times (see Figure 3.3b). By coating the pins with Teflon and minimising the bias spring coefficient (just enough to rotate the polarised discs), the friction between the pins and the slider window walls was curtailed. As the slider changed its position, each polarised disc rotated accordingly. This caused one disc to rotate relative to the other in a parallel plane and led to a phase difference in the orientation of the polarised films. The slider window was shaped in such a way that this phase difference was maintained between 0° and 90°. Further details of the polarised disc function are provided in Appendix A.

* This involved rotation of one film in respect to the other in parallel planes.

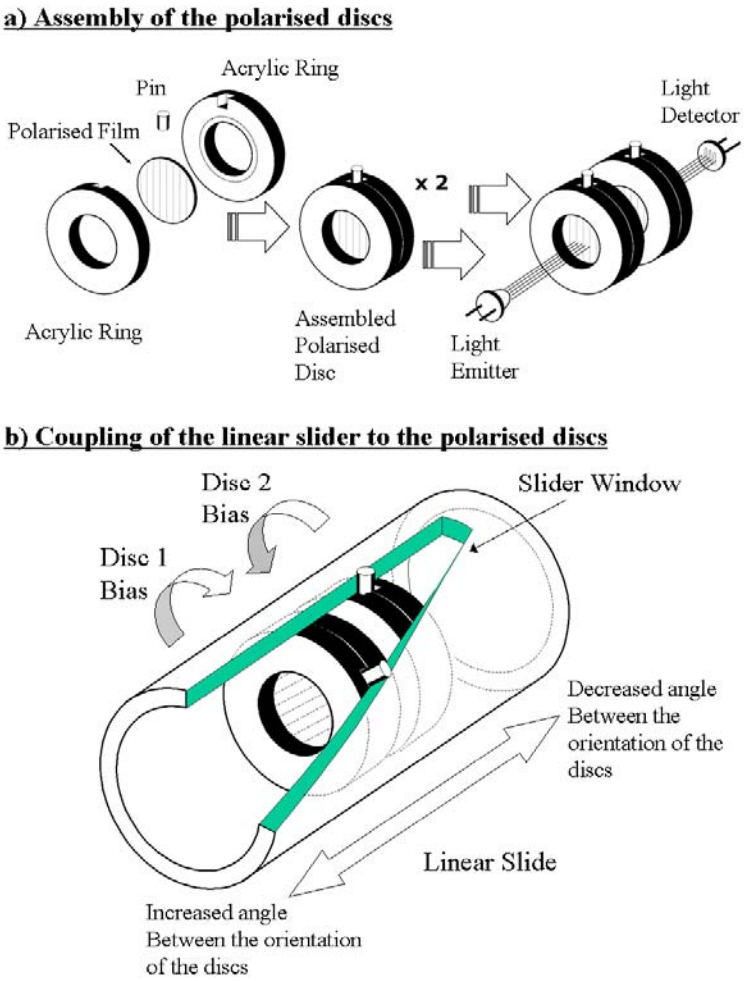


Figure 3.3 *Assembly and function of the optical sensing components.* **a)** *The polarised discs were assembled by gluing a circular polarised film and a protruding pin in between two acrylic rings. The polarised discs were free to rotate while light was passed through the pair and detected on the other side.* **b)** *Once assembled and held within a chamber using ball bearings, the pins of the polarised discs engaged with the cut away walls of an outer linear slider. The discs were spring biased away from each other to ride against opposite walls at all times. As the slider changed position, the correspondingly widening or narrowing opening caused the polarised discs to rotate relative to one another.*

The cut path was calculated based on the three dimensional geometries of the two disc pins interfacing with the slider window's edges. The calculated shape was dependent on the designated maximum travel as well as the diameter of the outer slider. A further complication arose due to the effect of the diameter of the pins (as the slider window walls were tangential to the pins). This was accommodated and two outer sliders were then constructed, one for each side of the FPT spring.

3.2.2 Assembly of the FPT

The evaluated FTP assembly (Figure 3.4) consisted of a cylindrical core shell containing the polarised discs and associated optical components. The inner-diameters of the two sliders were 0.1mm greater than the outer-diameter of the core shell allowing a smooth slide. Each outer slider also possessed a cross-pin that passed radially through a slot in the inner shell. The purpose of each pin was two-fold: Primarily each pin acted as the connection to the respective cable (either load-end cable or actuator–end cable). Secondly, the pins connected the sliders to opposite ends of the FPT spring[*]. Figure 3.4 shows a simplified depiction of the above elements. Figure A.5 in Appendix A presents the CAD drawings of the components and assemblies.

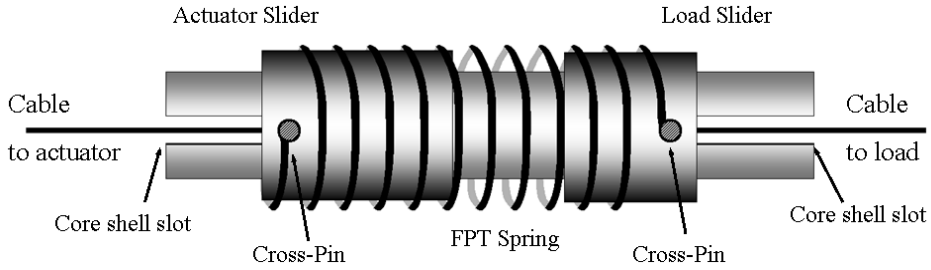


Figure 3.4 The final assembly of the FPT. The figure does not show the optical inner elements. Using this configuration, the FPT spring is free to slide and connected at each end to the actuator and load cables respectively. In addition each end of the spring is locked on to a linear outer slider. As the spring ends and the outer sliders move, changes occur in the inner optical components as described. These changes are measured using the output voltage of the respective phototransistors.

[*] Coupling was implemented by passing the spring over the sliders and connecting at each end.

3.2.2.1 Selection of the FPT spring

The FPT spring was selected by size and stiffness based on the length of the FPT, desired sensitivity and the actuator stroke length. A shorter spring can result in a shorter overall FPT length (see figure A.1 in Appendix A) but limits the spring's linear range of elasticity and hence its sensitivity. Conversely, a longer spring will increase the linear range and sensitivity of the spring but require a greater FPT length and actuator stroke length. The expected output stroke length in the presented work was 20mm. The maximum stroke length at the actuator end was 30mm. The additional 10mm represents the maximum length change of the spring. The spring coefficient was the selected based on the maximum tension specified (5N). The spring selected for the FPT prototype was 100mm in length and had an elastic coefficient of 0.63 Nmm^{-1}. This value was determined by calibration of the spring against hanging weights and displacement measured by a potentiometer[*]. Prior to calibration, the spring had been elastically pre-stretched by 15mm. This ensured that no plastic deformation would occur for subsequent stretches up to 15mm. The spring was then extended by up to 10mm repeatedly and retested. It was found to be linear within this range and elastic (i.e. no plastic deformations were observed). However, aside from the initial verification of the linearity of the FPT spring and its satisfaction of minimum stiffness, it was no longer required to calibrate the spring separately prior to the evaluation of the FPT. The calibration procedure of the FPT produces a number of constants which inherently incorporate the spring coefficient in them (see below).

[*] Honeywell potentiometer (MLT model, accurate within +/- 0.05mm)

3.2.3 Calibration of the FPT

There are two optical sensors (see Figure A.2 in Appendix A) in the FPT which must be calibrated. The sensors are designed to produce two voltage outputs linearly proportional to the positions of respective spring ends (see Appendix A). As the FPT is an integral, enclosed sensor within the actuation system, it may not be possible to calibrate both linear sliders against known positions. Instead, the FPT is calibrated using a single known position and a single known force (eg a hanging weight), both applied at the load end. This single point calibration will be beneficial in systems where multiple sensors are used.

3.2.3.1 Calibration Method

The calibration method involves three steps from which five voltages are measured and used to derive the calibration constants required for equations (3.1) and (3.2). These steps are detailed in Appendix A.

1. The FPT sliders are brought to minimum position (towards load direction). No force is applied.
2. The load end of the FPT is moved to a calibrated position. No force is applied.
3. With the load-end slider locked at the calibrated position, a calibrated force is applied at the actuator-end slider.

3.2.3.2 Deriving position and force (X and F) from voltage outputs (V_A and V_L)

The FPT output voltages are V_A and V_L measured at the phototransistors at the actuator and load ends of the FPT respectively (see the relationship between voltage output and slider position in equation (A.5) in Appendix A). The calibration process allows tension (F) and position (X) at the load end to be calculated from V_A and V_L by providing the constants for the following equations:

$$X = K_X \times V_L + C_X \quad (3.1)$$

$$F = K_{FA} \times V_A + K_{FL} \times V_L + C_F \quad (3.2)$$

Where K_X, C_X, K_{FA}, K_{FL}, and C_F are the derived calibration constants.

Appendix A explains how these constant are obtained. K_X, C_X are derived from equation (A.17) in Appendix A. K_{FA}, K_{FL}, and C_F are calculated from equation (A.16).

3.3 Methods – Testing the FPT

The constructed FPT body was fixed within a platform containing force and position sensors to test the function and accuracy of the device as a hybrid sensor (see Figure 3.5). A software-based control system was used to drive a stepper motor that was coupled to a worm drive linear stage to pull the actuator segment cable of the FPT. A Honeywell potentiometer (MLT model, accurate within +/- 0.05mm) was used as the reference position sensor. Its linear shaft was fixed to the load side slider (load-slider) of the FPT. The load cable was then configured in one of three ways as shown in Figure 3.5 a, b and c.

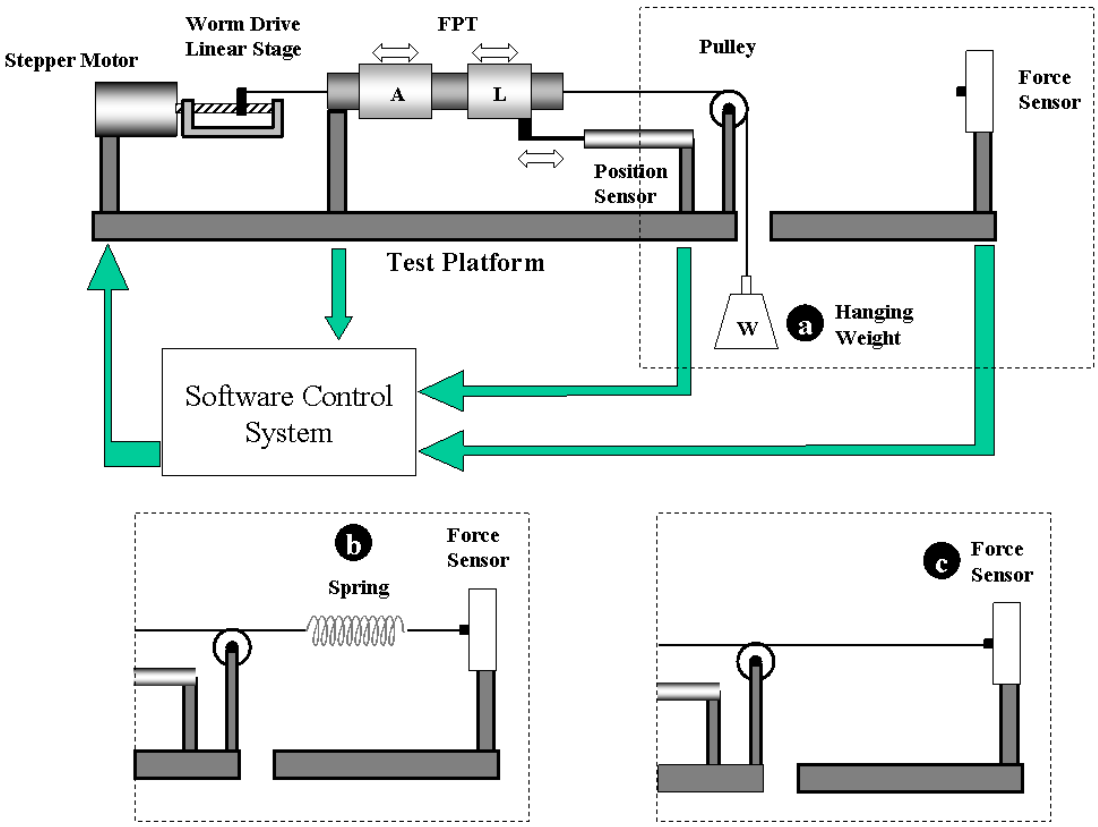


Figure 3.5 The FPT test platform.

The configurations were as follows:

a) The cable was attached to a hanging weight over a pulley.

b) The cable was connected to a spring attached to a perpendicular force sensor (15248A TOCO 4-element DC strain gauge balance circuit and amplifier).

c) The cable was fixed directly from the load-slider to the force sensor.

The software control was written in Labview™ and used a data acquisition card* to digitise and process the analogue inputs from the FPT, the potentiometer position sensor (0-0.5v) and the strain gauge force sensor (amplified voltage range was 0-1v

* National Instruments DAQ Card Model: PCIMIO 16XE-50

dc). The potentiometer was itself calibrated using a precision positioning system*. The strain gauge was also pre-calibrated against a series of known weights.

3.3.1 Test 1 – Linearity of the FPT outputs

Initially, the linearity of the slider outputs on each side of the FPT was tested against a positioning system (OWIS Precision Linear Stage LT 60-50) with a resolution better than 0.01mm.

1. The load-slider was connected to the moving stage of the positioning system and gradually moved from zero to its maximum travel (20mm). At incremental displacements of 0.5mm, the voltage outputs of the load-slider (V_L) were recorded and plotted against actual position (read from the positioning system)
2. The actuator-slider was connected to the moving stage of the positioning system and gradually moved from zero to its maximum travel (30mm). The incremental voltage outputs of the actuator-sliders (V_A) were plotted against positions set by the positioning system and analysed.

3.3.2 Test 2 – Accuracy of the FPT force and position (Fig. 3.5a)

The results of the first test showed that due to technical limitations in the precise construction of the FPT slider windows, the outputs were not linear. Compared to a linear fit, the slightly curved voltage-position and voltage-force relationships of the sensor were much better represented using polynomial fits. A 4^{th} order polynomial fit

* The potentiometer output was calibrated by characterising its travel against position set by the linear stage. The linear stage had an accuracy of 0.01mm where as the accuracy of the potentiometer was ±0.05mm.

was selected having been found to give better residuals than 2nd and 3rd orders and not as complex as 5th and 6th orders. This was implemented within the software based control system to obtain force and position measurements from the FPT, which were then compared to actual force and actual position (as provided by the reference sensors).

3.3.2.1 Evaluation of accuracy

During analysis, FPT position errors were measured as percentage of full-scale (20mm). Likewise, the FPT force errors were calculated as percentage of the full-scale force (5N). Full-scale was chosen over standard deviation as the more appropriate form of evaluating accuracy, because the absolute values of position and force are not as meaningful without reference to their ranges. For example a 0.5mm error would appear as a 50% error at 1mm and a 5% error at 10mm if the error is viewed as a percentage of absolute value. However, in position control, the absolute value is subjective to our point of reference and by itself meaningless. More interesting is how far we deviate from our desired value in proportion to our total capability. In the above example, a 0.5mm deviation would appear as 2.5% of the full-scale 20mm regardless of where it occurs. This is consistent with standard measurement practice (Carr and Brown, 2001).

3.3.2.2 Position accuracy

The load-slider, attached to a series of free weights (as in Figure 3.5a) was taken through its full range of movement several times while the system sampled the voltages. The program then calculated FPT positions at small intervals using the calibration coefficients and compared these with the actual positions measured by the potentiometer.

3.3.2.3 Force accuracy

While controlling position with 50g, 100g, 200g, 400g and 500g suspended masses, the FPT tension was simultaneously calculated using the pre-determined calibration coefficients and compared to the actual weights.

3.3.3 Test 3 – Isotonic Position control of a suspended mass (Fig. 3.5a)

A control system was implemented in the software using FPT position only as feedback. Two types of control were applied. First, different static positions were entered. The stepper motor was driven to bring the load-slider to the desired position. This was tested using several different position inputs with different loads (as in Figure 3.5a). Dynamic testing was implemented by imposing a sinusoidal* varying position at a variety of amplitudes, offsets and frequencies. The potentiometer was used to measure the accuracy of the control system in both modes.

3.3.4 Test 4 – Isometric Tension control of a cable (Fig. 3.5c)

The load cable was fixed directly onto the rigid force transducer to verify the accuracy of the force control system (see Figure 3.5c). As in position control, two modes were implemented, one of constant force at different values and one at cyclic transient values.

---

* The sinusoidal shape is useful for therapeutic movement control (see Chapter 7).

3.3.5 Test 5 – Position and force control of a dynamic element (Fig. 3.5b)

The final test incorporated the force and position control of a *dynamic load*; a system in which both position and force changed during movement. This is in contrast to the isotonic system, whereby force is static while movement changes, and, the isometric system in which only force changes. A complex dynamic system may be exemplified by the interaction of the human-machine interface where the human movement may impose external forces and movements on the control system. As a simplified representation of the dynamic system, a spring was placed in series between the load-slider cable and the fixed force sensor (see Figure 3.5b). The cables were tightened to the effect that at position 0 there was a small residual tension in the series cable. The spring used was examined prior to the test and found to have a spring constant of ~ 176 N/m.

Finally, two control systems were tested in which force control and position control were co-dependent. First, force control was applied while limited within position boundaries. In this thesis, this is defined as Position Limited Force Control (PLFC). Next, in a similar manner, position control was tested while limited to a maximum force. This is defined as Force Limited Position Control (FLPC).

3.4 Results

3.4.1 Test 1 Results – Linearity of the FPT outputs

The voltage output of the load-slider (V_L) against actual position (X_L) shows an output close to linear (a linear fit produces R^2 =0.9985). However, although repeatable and stable, the output is better represented by a 4^{th} order polynomial fit (produces

R^2 =0.9999, see Figure 3.6). The output characteristics of the actuator-slider are less linear (R^2 =0.9980). Again, a 4th order polynomial fit is a better representation (R^2 =0.9999). Based on the results from Test 1 and the polynomial best fits, a new method was subsequently developed to calibrate slider voltages vs. position on load and actuator sides.

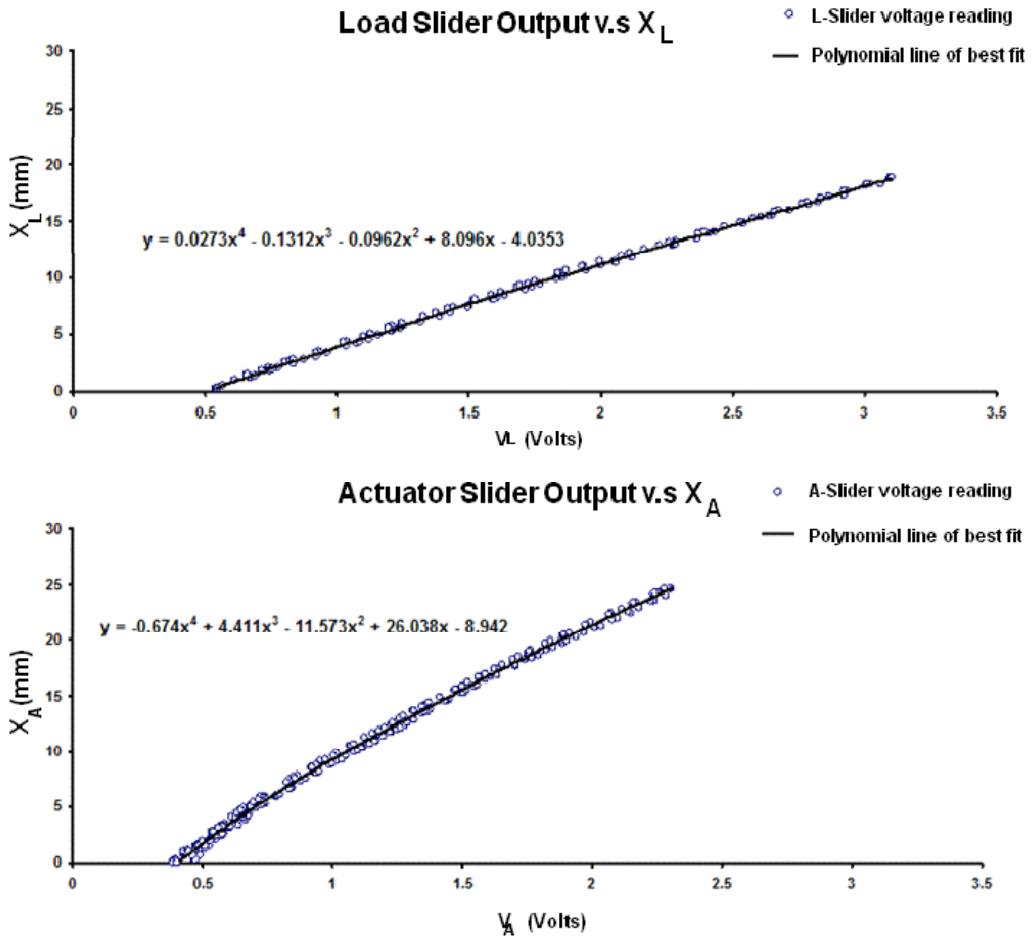


Figure 3.6 Slider voltage outputs v.s. positions. Graphs show the sampled voltage outputs of the actuator-slider and the load-slider, at position intervals. The load-slider has a range less than 20mm, while the actuator-slider needs a larger range due to the stretch of the spring. The load-slider is more linear than the actuator-slider, but neither can be represented by a linear equation with acceptable errors. The best approximations found were the 4th order polynomial functions shown on each graph.

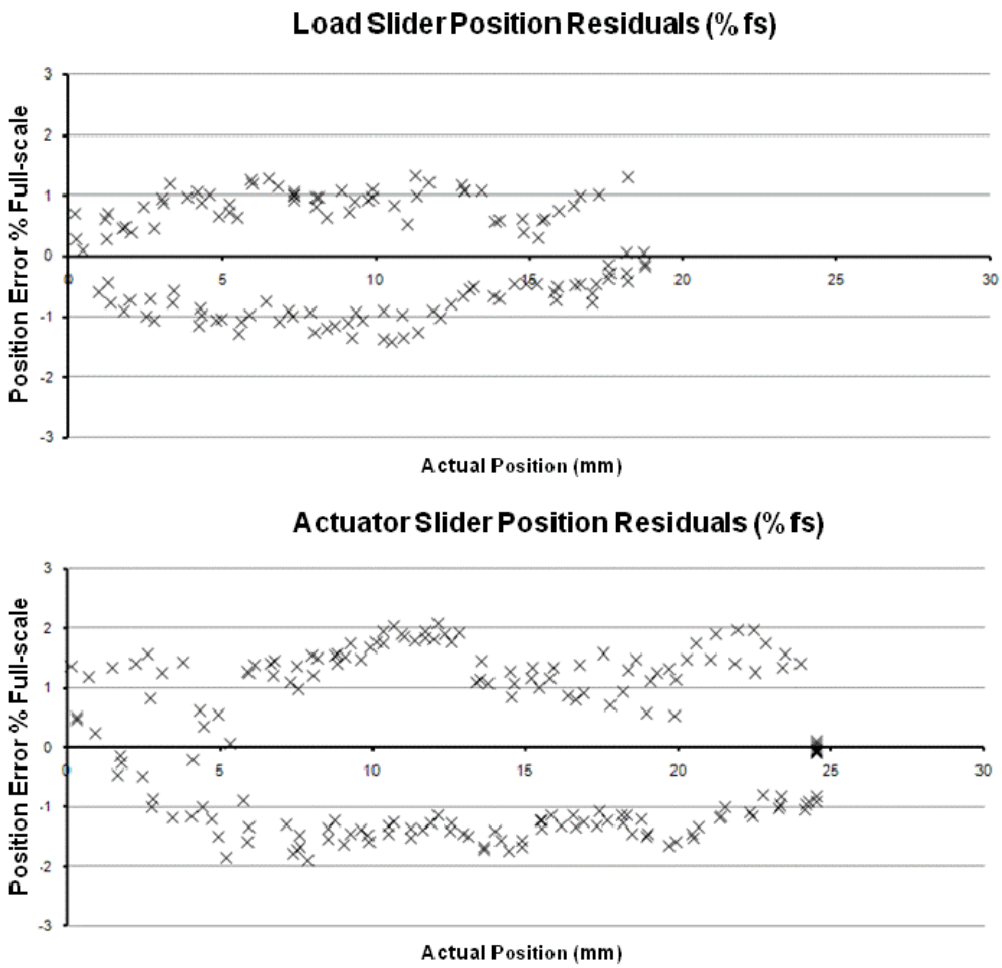


Figure 3.7 Actuator Slider and Load Slider residual plots. Graphs show residual plots terms of % error full-scale for each slider output respectively (i.e the difference between actual positions and positions calculated from the 4th order polynomial based calculations). Hysteresis is observed in both outputs.

3.4.1.1 Modification in Calibration Procedure

Software was used to fit polynomial lines to the plotted voltages and derive coefficients for the equations. The equations were next used to calculate position and force from the FPT voltages. Subsequently, the polynomial coefficients were used for all the tests. Due to the non-linearity of the prototype, the calibration procedure changed from those previously described in the design section to the following series of steps:

1. The load-slider was moved through its range and data points were collected to plot V_L against actual position (as measured by the potentiometer). This was performed by the motor and the polynomial of best fit was applied by the software to obtain the coefficients.
2. With F = 0, the actuator-slider was moved between X_{Amax} and [X_{Amax} - X_{Lmax}] (where X_{Lmax} is the load-slider's maximum position) and V_A was plotted against a derived X_A (as the reference position sensor was attached only to the load side). This gave enough data points to plot the upper section of the X_A range.
3. Finally maximum load was applied (4.9N) and the actuator-slider was moved again. This allowed the measurement of the constant K_s as well as the plotting of the lower positions of X_A.
4. X_A and X_L was calculated using V_A and V_L. K_S was also derived. Subsequently, X and F were able to be measured using the software and recorded in time-based samples during the experiment.

3.4.2 Test 2 Results – Accuracy of the FPT force and position

After the calibration of FPT tension and position, the accuracies of measurements were examined using the reference sensors.

3.4.2.1 Position Accuracy

Position accuracy was found to be better than ±0.4mm or ±2% of full-scale travel for a series of loads (applied as the hanging masses described in the Methods section). The position error changed with varying load and position. As each load was moved by the motor, the % error was recorded at regular intervals and plotted against

corresponding position. In order to examine the hysteresis errors, sampling was divided into two groups; the Push tests (movement away from the motor) and the Pull tests (towards the motor). The movements in the Push and Pull direction were repeated several times for each of the hanging masses 50g, 100g, 200g, 400g and 500g. These plots were respectively named P50 to P500. Figure 3.8 shows the % error versus position in the Pull and Push graphs.

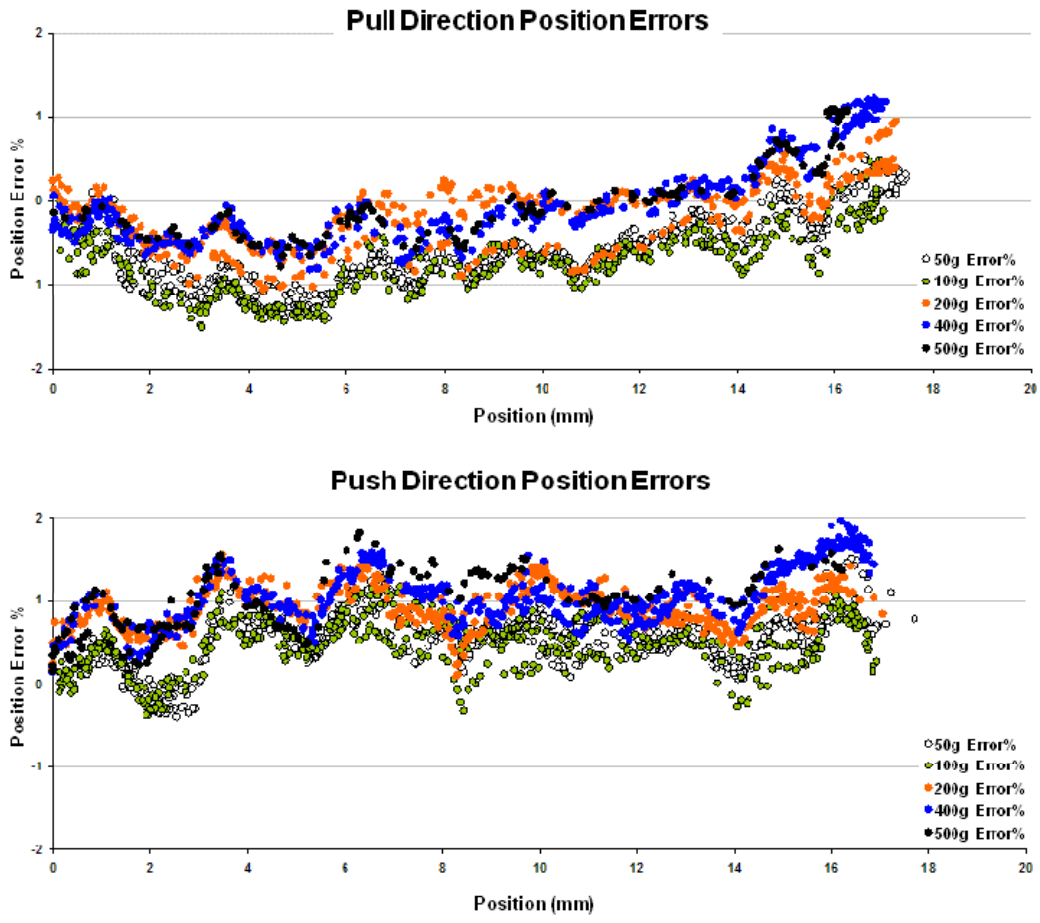


Figure 3.8 Position measurement errors for pull and push directions. Graphs show the error distribution in the Pull and Push directions for a range of loads (50g to 500g masses). The data shows that for a given load, a pattern is repeated in the distribution of the errors. However as the load is increased, the pattern seems to shift up.

3.4.2.2 Force Accuracy

Force accuracy was measured by fixing the load-slider to the reference force sensor using a rigid cable and pulling on the actuator-slider cable, thereby increasing the tension in the FPT (see Figure 3.5c). This was repeated with the load-slider at different starting points to see if the error distribution changed at varying positions. The force error was calculated as a % of the maximum load, 4.9N (in order to represent absolute error). The results of this test showed that force error is higher than position error for reasons discussed in section 3.5.2.2 (page 45). The force error plots show hysteresis. The graph in Figure 3.9 shows a variation in force error at a range of tensions applied isometrically at different fixed positions.

3.4.3 Test 3 Results – Isotonic Position control of a suspended mass

The results in this section are presented as time-based graphs showing constant or changing position inputs (desired position) and the response of the system measured by the reference position sensor. The control system was implemented by using the FPT position feedback (not shown in the graphs) to drive the stepper motor to reach the target position.

3.4.3.1 Constant position targets

Figure 3.10 shows the results of the position control test using constant positions as targets. The isotonic load during this test was 1.96N (200g hanging mass). Once the target was reached, the difference between the desired input and the actual position was measured using the reference sensor. Over 50 trials, the mean error was 1.41%

of full scale range of movement (20mm). Rise and set time of the response were limited to the motor speed (6mm/sec).

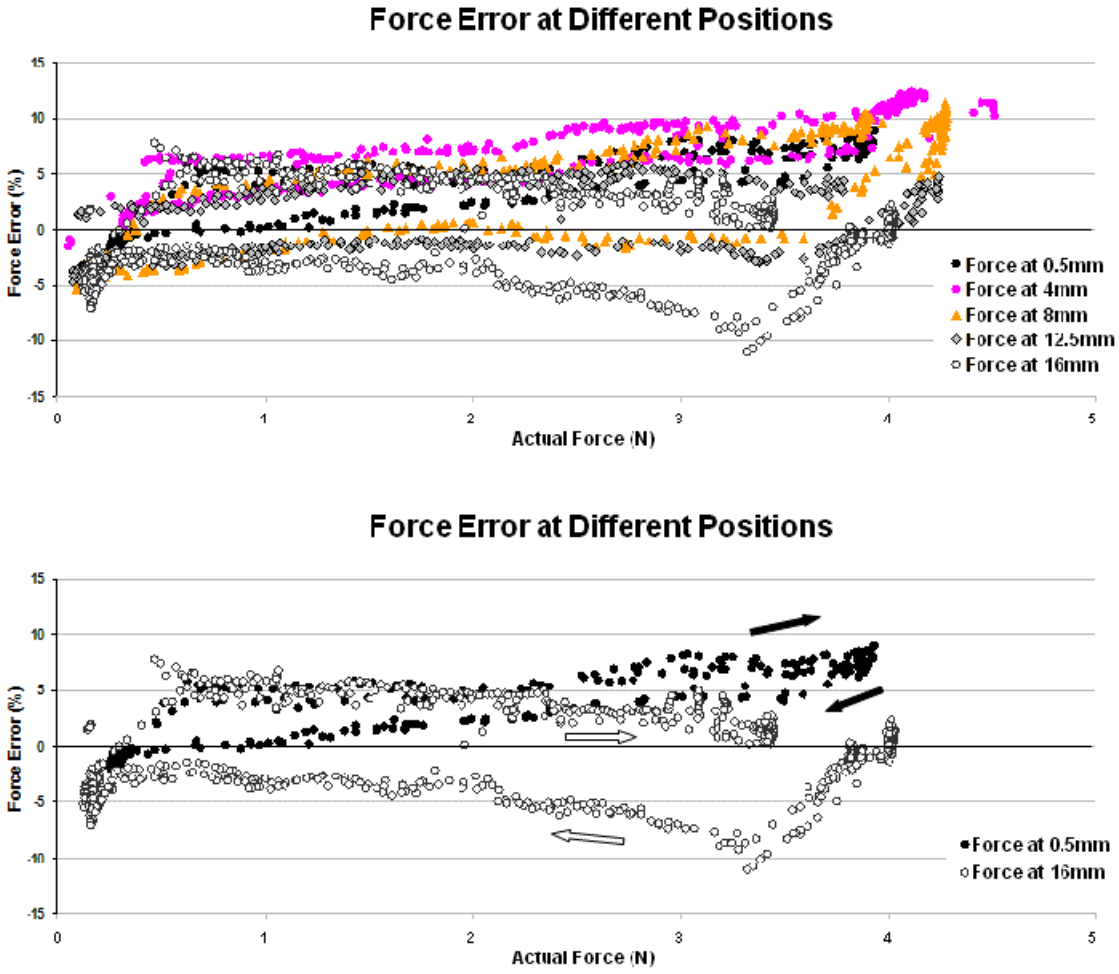


Figure 3.9 Force measurement errors. Graph show the shift in the force error distribution at a range of induced tensions applied at different fixed positions. For clarity, the lower graph isolates the distributions of the errors at 0.5mm and 16mm. The force error is positive when pulling and negative when releasing the cable as indicated by the arrows in the lower graph.

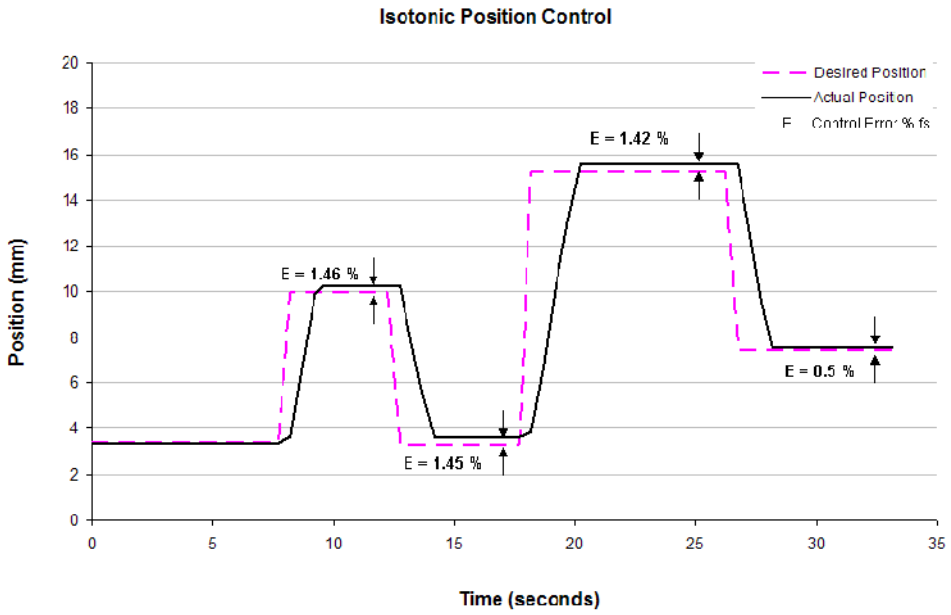


Figure 3.10 Isotonic position control. *The graph shows the results of a position control of a 200g hanging mass. Desired position was entered by the user at the computer interface. The software then used the FPT position as feedback to drive the motor to track the target. The actual position shown on the graph was measured by the reference sensor for verification.*

3.4.3.2 Cycling position control

Using the same isotonic load, sinusoidal position inputs were entered into the control system. The results of the test are presented in Figure 3.11 and Table 3.1. The differing inputs tested varied in amplitude, frequency and offset. The input target was followed by the motor using position feedback from the FPT.

The highest frequency attempted was 0.2Hz (5 second cycle length). The results showed that the system response lagged behind the target input by around 0.5 seconds due to the motor speed limit. The mean errors observed were mainly within ±2% with some peaks into 4% for the first three inputs. In the test with the 0.2Hz input however, the mean error exceeded 10% with some peaks beyond 20%.

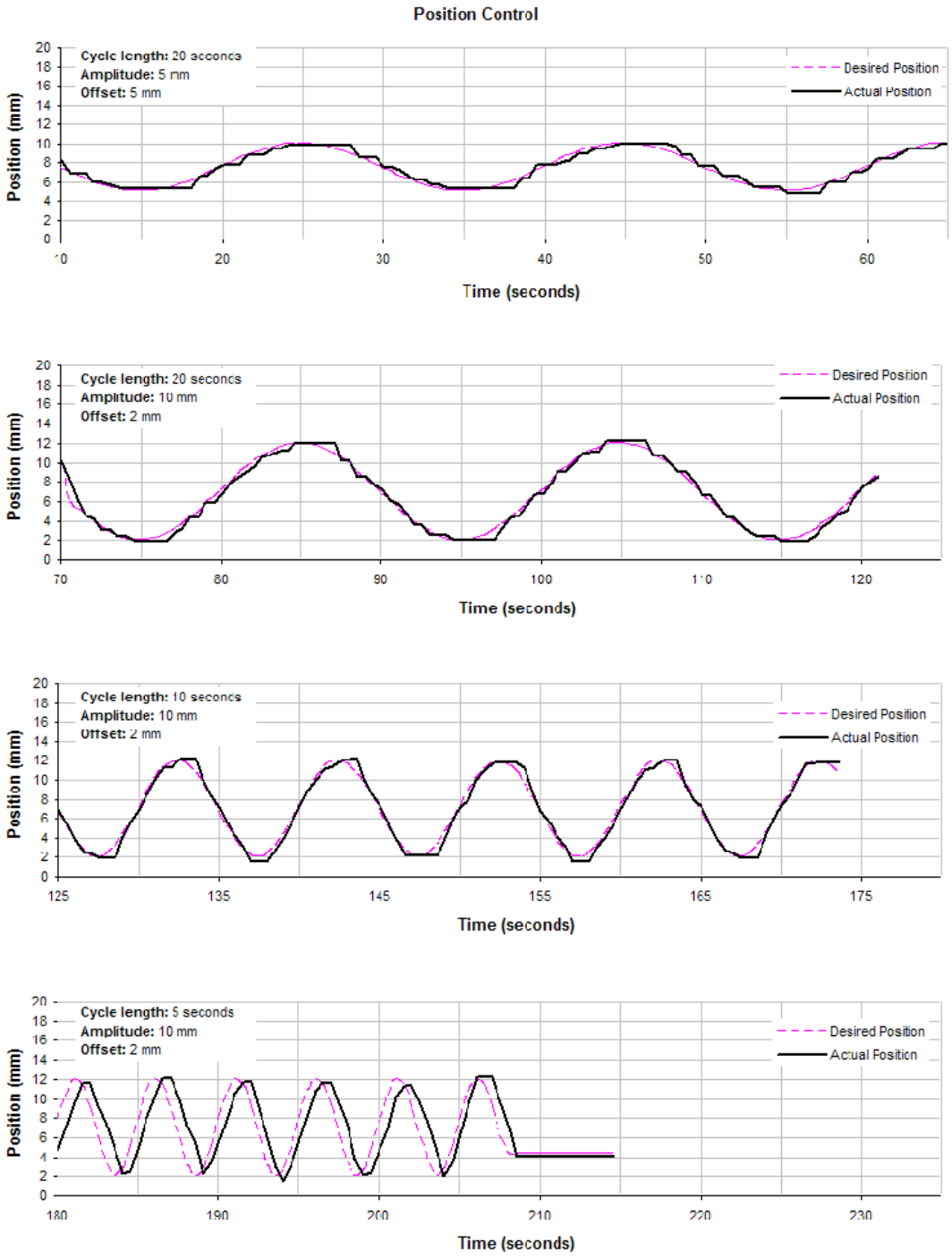


Figure 3.11 Cyclic, isotonic position control using the FPT for position feedback. *The graphs show desired position as the target input and actual position as measured by the reference position sensor. The graphs represent tests with inputs varied in their cycle lengths, amplitude and offset as indicated on each graph.*

Input	Cycle length	Amplitude	Offset	Mean error % Full scale	Lag (s)
1	20 sec	5mm	5mm	±1.53%	0.1
2	20 sec	10mm	2mm	±1.60%	0.1
3	10 sec	10mm	2mm	±2.02%	0.1
4	5 sec	10mm	2mm	±11.10%	0.5

Table 3.1 Result summary of cyclic, isotonic position control

3.4.4 Test 4 Results – Isometric Tension control of a cable

In this test, isometric forces produced by the system were compared with constant or changing tension targets (desired force). In the first trial within this test, a series of constant desired tensions (in Newtons) were entered and the response was recorded over time. The force control response performed notably more poorly than the position control. Subsequently, the input was changed to sinusoidal. A series of input amplitudes, offsets and frequencies were tested.

3.4.4.1 Constant tension targets

While the FPT load-slider was fixed to the reference force sensor, desired tensions (in Newtons) were entered into the control software. The motor pulled on the FPT spring until the required tensions (as measured by the FPT) were reached. The results of this test (Figure 3.12) show that the average control error remained within ±6%. As in the position control system, the algorithm was simple and used a tolerance window towards which the motor worked to reach. In tension control, this tolerance was 0.1N. Tolerances below this value caused the output to overshoot beyond the target value and cause instability (see Figure 3.12 bottom graph).

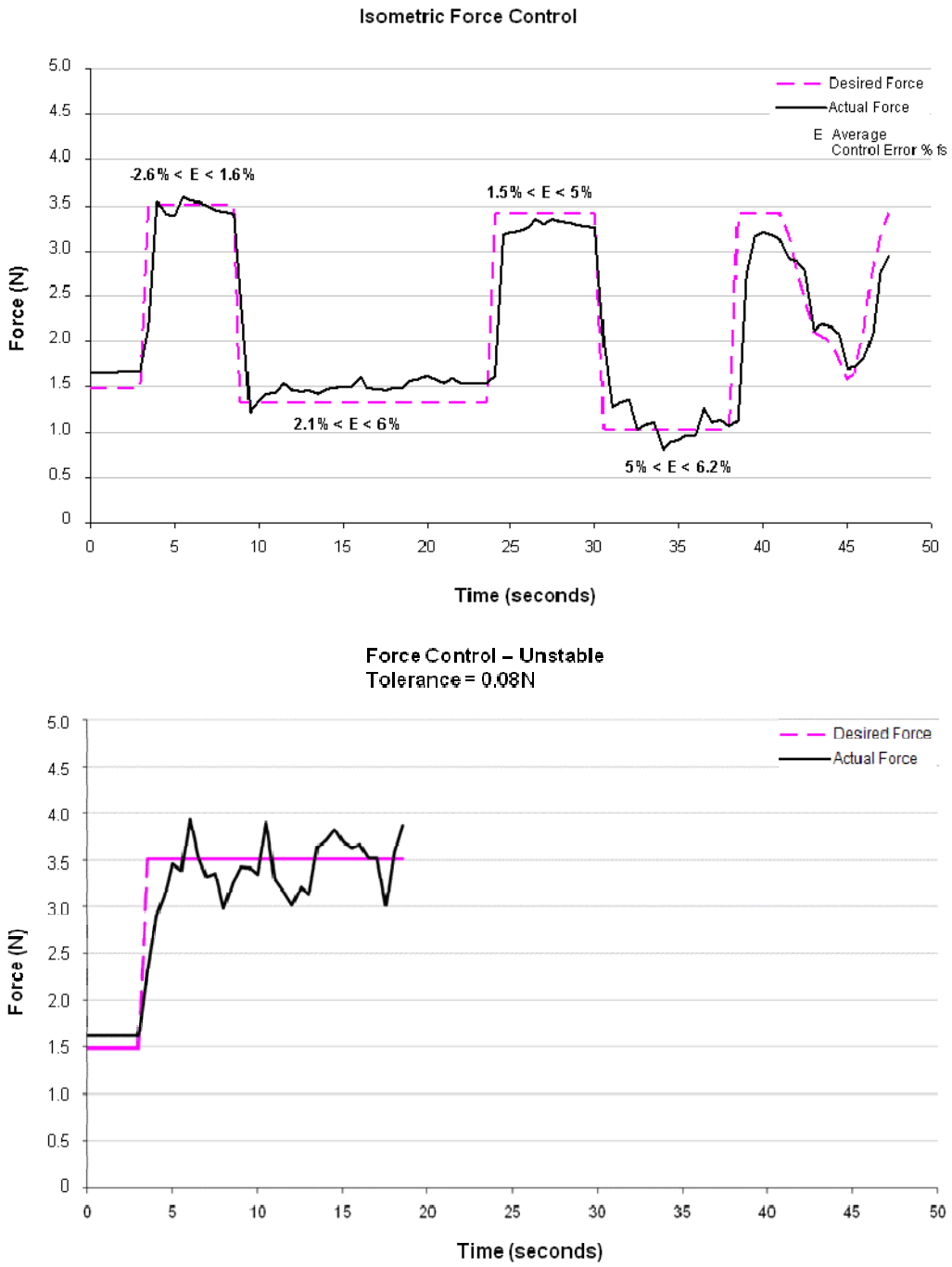


Figure 3.12 Isometric force control. *The graphs show the result of the isometric tension (force) test using constant values as targets. The response was found to be less accurate than observed with position. This is consistent with the FPT force accuracy which is lower than its position accuracy. The top graph shows the response using a tolerance window of 0.1N. Below this value, the control system became unstable as can be seen in the bottom graph (tolerance window of 0.08N).*

3.4.4.2 Cyclic tension control

The results in Figure 3.13 and Table 3.2 show that the system is capable of controlling tension according to a varying input. Different frequencies, amplitudes and offsets were applied to the control system. Inputs at frequencies within the speed capabilities of the motor were followed using the FPT as force feedback sensor. This was verified using the reference force sensor. The control error for all cyclic force tests remained within ±10% of full scale range.

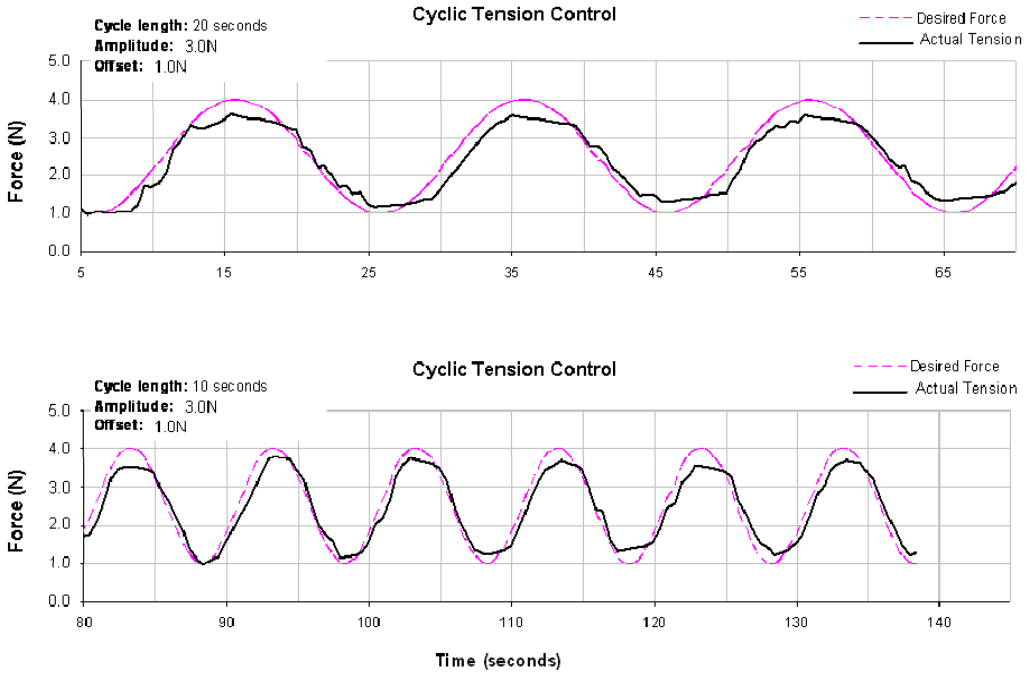


Figure 3.13 Cyclic tension control. The results of the isometric, cyclic tension control using the FPT for force feedback. The graph shows that the control system is driving the motor in order to follow the cyclic target input. Although not as accurate as the position control system, the force control is able to follow the target to within ±10% of full scale load.

Input	Cycle length	Amplitude	Offset	Mean error % Full scale	Lag (s)
1	20 sec	3 N	1 N	±9.4%	0.5
2	10 sec	3 N	1 N	±9.6%	0.5

Table 3.2 Result summary of cyclic, isometric tension control

3.4.5 Test 5 Results – Hybrid Position and force control of a dynamic element

In order to test the ability of the FPT to control a dynamic system, a spring was used. As the motor pulled back on the cable to reach a desired position, the tension in the system increased. In this mode, constant and cyclic control inputs were applied to both force and position.

3.4.5.1 Constant tension and position control tests

Using the same method as in the isometric and isotonic tests, tension and position were controlled during the application of a dynamic load. The results of the constant position and constant tension target inputs are shown in Figure 3.14. The graph shows that tension and actual position were co-dependent (due to the spring load). In the first graph, tension increases with position. In this test, a maximum tension of 4N was specified as a limit to the position control. The graph shows that the position target was followed by the control system. When the tension exceeded the limit, the motor stopped pulling on the cable. In practice, the actual tension at which this occurred was approximately 3.8N (due to the error in the FPT force reading). This was an example of *Force Limited Position Control (FLPC)*. Next, tension control was trialled with a position limit of 10mm. The lower graph in Figure 3.14 shows the results of this test. As tension increased, position also increased (again due to the presence of the spring load). During the tension control trial, whenever the position exceeded 10mm, the motor stopped. Subsequently when the tension target was brought back below the point at which the motor had halted, movement re-activated in the reverse direction. This was an example of *Position Limited Force Control (PLFC)*.

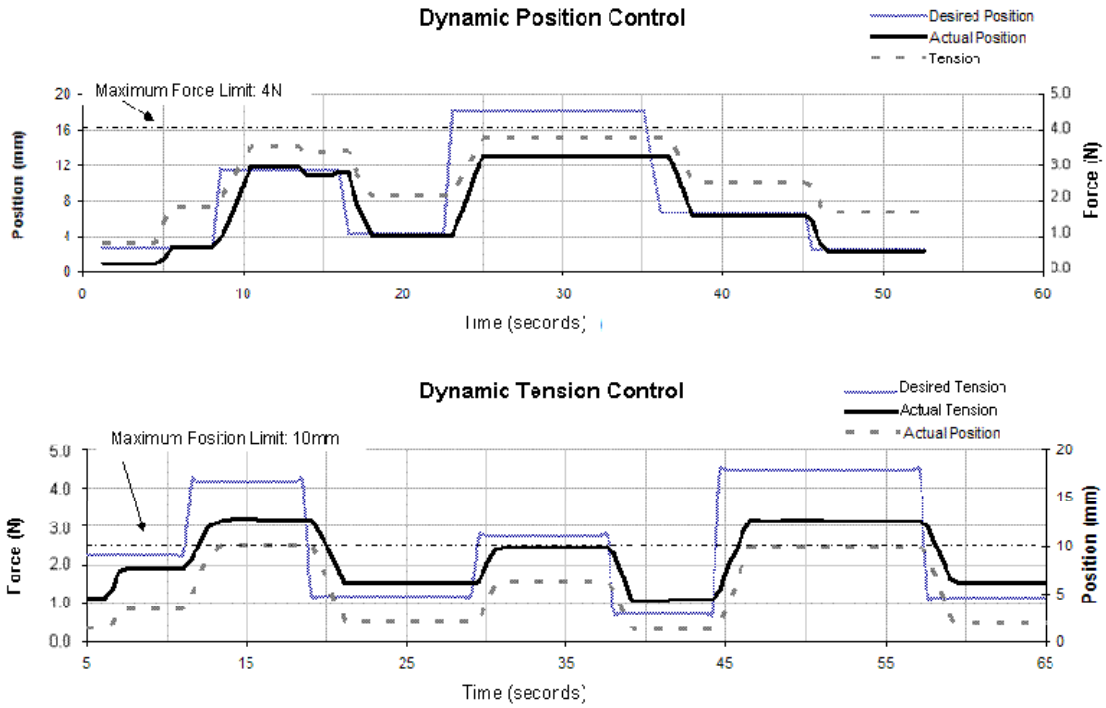


Figure 3.14 Dynamic position and tension control (constant). The top graph represents Force Limited Position Control (FLPC) applied to the spring load. The graph shows desired position, actual position and actual force. The set force limit is 4N. During position control, if the force measured by the FPT exceeds this limit, the motor ceases pulling the cable further, even while the target is higher. In the lower graph, tension (force) control is applied with a position limit of 10mm. As the force is increased the position also increases. When the position exceeds 10mm the motor halts until the desired force returns below the level which causes the position limit to be breached.

3.4.5.2 Dynamic cyclic position and tension control

The final sets of tests were cyclic FLPC and PLFC applied to the dynamic load.

Cycling tension and position are useful techniques used in rehabilitation therapy. An example of this is Continuous Passive Motion (CPM) therapy in which using a human-machine interface, a joint in the human body is gently and continuously moved between two predetermined limits (Le Stayo 1995; Adams and Thompson 1996). The cycle tests presented in Figure 3.15 show that while the position was cyclically controlled, a tension limit of 2.5N was enforced. Consequently when the tension in

the spring exceeded this limit (as measured by the FPT), the cycle operation was temporarily ceased until the target position reduced to a value that would lead to forces below this limitation.

The same process was introduced to the cyclic tension control of the spring load. During the cycling of tension, a position limitation of 4mm (from resting position) was applied. This caused the control system to stop the cycle at the moment where the position reading of the FPT reached 4mm. Again the motor resumed operation (in the reverse direction) when the target input reduced to value that would lead to a position less than 4mm.

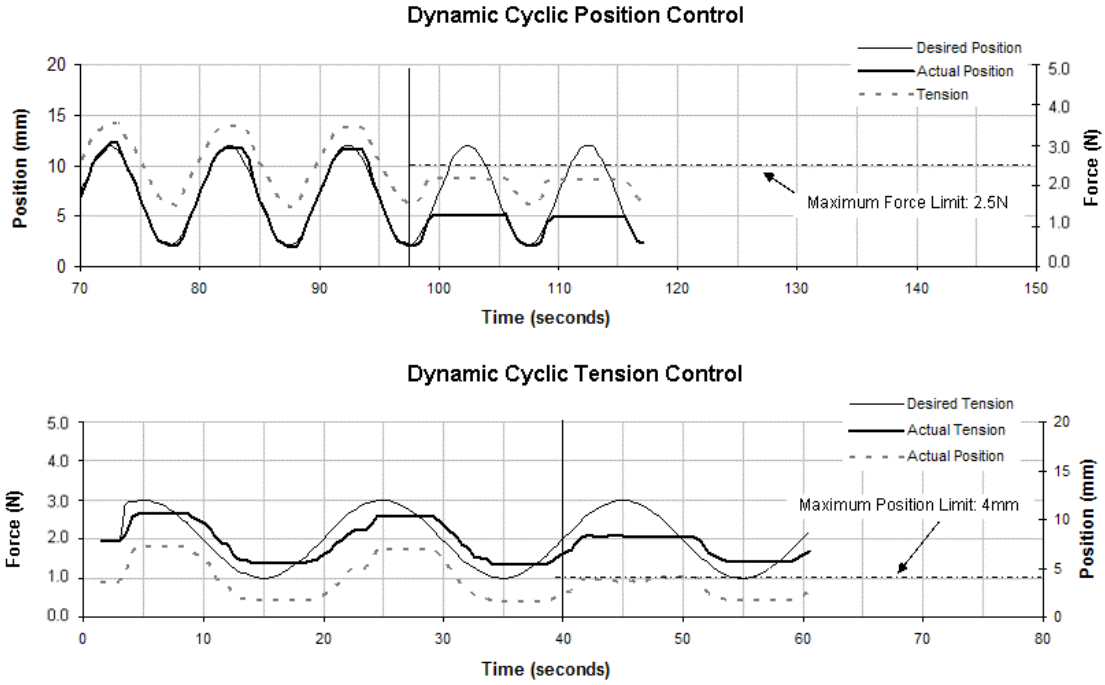


Figure 3.15 Dynamic position and tension control (dynamic). In the final series of tests, the tension and position cyclic control was applied to the external spring (dynamic load). During the position cycling test, a force limitation of 2.5N was imposed. Similarly, tension of the dynamic system was cycled repeatedly and following a number of cycles, a maximum position limit of 4mm was enforced to which the system responded appropriately. The solid vertical lines represent the points in time when the new control limits were imposed.

3.5 Discussions

3.5.1 Test 1 – Achieving linearity

During the modelling process, the slider window shapes were carefully designed to yield linear voltage outputs. The results however, still deviate from an acceptably straight line (see Figure 3.6). This is thought to relate to the limitations in manufacturing of the FPT. To achieve linearity, the tolerances in the dimensions of the window shapes, as well as the precision of the positions of the internal components must be better than the current prototype. Improvements would be well within the capability of an appropriate manufacturing facility. These should entail higher precision construction and better materials to improve linearity and reduce friction. The current prototype however, has been constructed within the confines of a standard hospital mechanical workshop and electronics laboratory.

3.5.2 Test 2 – Force and position accuracy

3.5.2.1 Hysteresis and load effects

It is evident from the graphs in Figure 3.8 that there is significant hysteresis between the Pull and Push directions. Additionally, there is a shift in the error distribution, with changing loads. The error distribution for each load has a defined shape or pattern. The repetition of these patterns within all loads suggests that they are caused by mechanical irregularities in the slider window walls (due to machining limitations in making the sliders). The random or noise components appear to be considerably smaller than the wider repeating patterns in the error distribution, the

hysteresis effect and the shifting distribution due to the changing loads. This shows that there is a potential to greatly reduce the error size by reducing the latter effects. The characteristic error patterns may be countered using a look-up table form of calibration (due to their predictability). Ideally however, by eliminating the error patterns through better manufacturing, the calibration of the FPT can be simplified greatly as described in section 3.2.3 (page 22)

The error shift can be viewed more clearly by isolating two different loads using the 50g and 500g hanging masses. Figure 3.16 shows that the shapes and hysteresis of both loads are very similar, but that there is a positive shift in the error distribution for the heavier load.

The hysteresis can be explained as before by the technical limitations in the construction of this first FPT prototype. This is expected to improve greatly with better manufacturing of the device. The error-distribution-shift with heavier loads is most likely due to the small elastic deformation experienced by the test platform causing a shift in the position of the reference sensor in respect to the FPT. Consequently this observed shift is not a real error of the FPT function but rather an effect of the small deformations in the test platform.

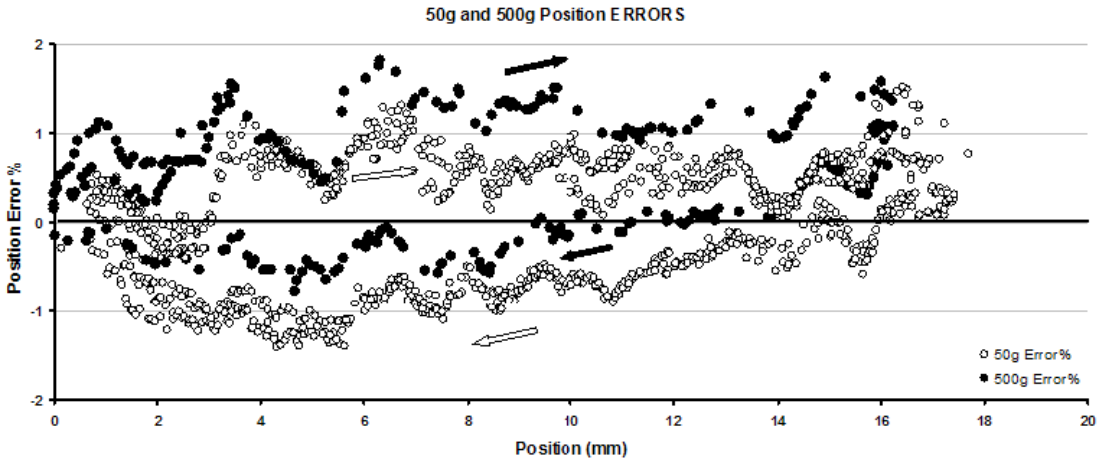


Figure 3.16 50g and 500g position errors. Graph shows the shift in the error distribution between 50g and 500g hanging masses.

3.5.2.2 Explanation of the large force error size

Generally, the force error is positive when pulling the cable and negative (or decreased) when releasing the cable. This is indicative of the friction in the system. The cable on the actuator side has extended the FPT spring but the friction has prevented the full force to be transmitted across the FPT to the reference force sensor (see figure 3.17) .

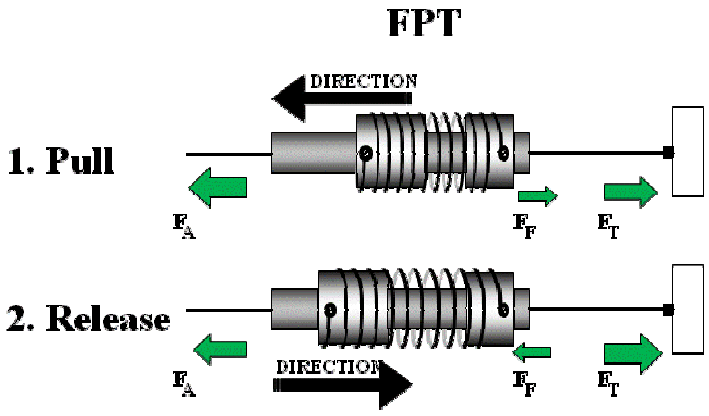


Figure 3.17 Contribution of friction to the error distribution. F_A: Force from actuator. F_F: Friction force at the load end. F_T: Tensile force in the cable leading to the reference transducer. When the motor is pulling on the cable, friction at the load end reduces some of the force transmitted to the reference sensor leading to a positive error. When releasing, the friction force acts in the opposite direction leading to a negative error.

It is also apparent that at lower values of position, the errors are shifted up. However, the hysteresis of these error points is not as great as those observed in the higher positions (see Figure 3.9).

It is to be expected that the FPT should exhibit larger force errors than position errors. Due to the design of the device, force is derived indirectly from two position measurements, accumulating errors from both. In addition, any friction in the FPT is expected to affect force measurements where it would not influence position measurements (as described previously).

Finally, the force calibration process itself may be a further contributing factor to the large errors. The actuator-slider is not calibrated directly. Rather, its calibration position is inferred from that of the load-slider (as described in the modified calibration section). During calibration of the upper part of the actuator-slider range, the load-slider is moved back and forth on the assumption that it is always at a constant distance away from the position of the actuator-slider (see steps 2 and 3 in the modified calibration section on page 30). This may introduce more error in the calibration than had the actuator-slider been calibrated directly. The lower position range of the actuator-slider is likely to exhibit even more error as it was calibrated by moving it through its range while lifting maximum weight. As before, there is an assumption that the spring will maintain a constant length throughout the movements. In practice friction and inertia in the system will compromise this process significantly. These factors may explain why tension at 16mm would have a wider error distribution than that found at 0.5mm: At 16mm more of the error prone lower range of the actuator-slider is used.

3.5.2.3 Proof of concept and improvements

Improvements in the construction of the FPT will lead to lower friction and tolerances and better rigidity in the placement of the components. It can be seen from the graphs that the random distribution of error is much lower in magnitude than the variation caused by hysteresis and suspected machining irregularities. Thus enhancements in the construction can be expected to greatly decrease calibration error, hysteresis and the overall error in force measurement.

While keeping the size of the errors in these tests in perspective, the results show that this first prototype functions as a proof of concept for the hybrid FPT presented in this report.

3.5.3 Test 3 – Isotonic position control

Figure 3.10 (constant position) and Figure 3.11 (cyclic position) showed that as the position input was changed, the motor drove the load immediately towards the target (desired position). However the motor speed limited the slope with which the target was followed.

The cyclic position control results showed that the system worked well for the given inputs though at times the movement was unsmooth. The unsmooth response of the system was due partially to the inaccuracy of the FPT but also in part due to the simplistic control algorithm that was implemented in the software. The control response was not proportional to the error. Rather, the algorithm specified that the motor should move towards the error until it reached the target within a certain tolerance. While within the tolerance window, the motor halted and reactivated only when the target exited the window. As a result, the movement appeared step-like. In the test results presented in Figure 3.11 the tolerance window size was ±0.8mm.

This value is ±4% of the full scale (20mm). The mean control errors however, were mostly around 2%. This is likely due to the following: While approaching a target value, the motor's momentum caused the final position to stabilise within the tolerance window between the boundary value (4%) and the desired target value.

3.5.4 Test 4 – Isometric tension control

As expected, the tension control performed poorly compared to position control. This was mainly due to the lower accuracy of force measurement of the FPT as well as the friction in the current prototype hindering smooth movement and force transmission in the system.

3.5.5 Test 5 – Hybrid position and tension control

As with previous tests, position control was more reliable than tension control. Although accuracy of hybrid control was limited to the precision of the FPT, this test was demonstrative of its potential.

3.5.6 Significance of results

The results of Test 3, Test 4 and Test 5 show conclusively that the FPT can be used as both a force sensor and a position sensor, within a control system requiring either of these as feedback. However, the most useful function of the FPT will be its ability to be used as a hybrid sensor to safely and effectively control force and movement in an environment that may involve unknown obstacles and independent passive and active elements. This is expected to have significant applications in human-machine interfacing.

An example may be a powered arm exoskeleton that could be used to aid flexing at the elbow (Charles Racine 2003). In a variety of scenarios, such as if the user resists the movement, if an obstacle is encountered along the way, or if the elbow reaches its limit, the exoskeleton would be able to respond appropriately. Using an FPT as the hybrid sensor for control of the actuator, limits can be set on either position, force (equivalent to tension or torque) or both. In the case of the user resisting movement, the system can be force limited so that if the force is exceeded, the direction of movement is automatically reversed. Similarly if the elbow reaches its physical limit, the control system can stop the motor. If an obstacle is encountered during the movement, the system can observe a set force limit, thereby protecting the user, the obstacle, as well as the system from damage. The elastic element in the FPT is highly useful in reducing the rigidity of the type found in industrial robotic systems (Pratt and Williamson 1995).

There are many other possibilities to utilise the FPT as a hybrid sensor. Current application includes the use of multiple FPTs within the Rehabilitation Glove to allow new forms of therapy, assessment and functional aid for the functionally compromised hand (Abolfathi, Scott et al. 2006).

3.5.6.1 Necessary improvements

The current prototype of the FPT has been constructed mostly by hand, within a facility with limited technical capabilities. This study has been useful as a proof of concept of the sensor and to show how it can be used. However, the sensor will require significant improvements in construction before it can be truly compared in accuracy, versatility, energy consumption, size, cost and performance with other sensors currently available. The anticipated improvements will include:

1. *Miniaturisation of the body of the sensor.* This will be mostly in diameter as the length will depend mainly on the actuator stroke, maximum force, resolution of the force and the FPT spring characteristics.
    - The minimum length of the sensor will be equal to the output stroke length (design dependent) + spring length + maximum spring extension + structural support dimensions. The spring length and maximum extension as well as FPT diameter are dependent on sensitivity, spring constant and spring range of linearity (as indicated in section 3.2.2.1).
    - Structural support can be improved by reducing dead-lengths at both ends of the FPT and reducing its cylindrical wall thickness. Currently this requires advanced manufacturing capabilities.
2. *Reduction of friction.* This can be achieved by improving the materials used, the precision of the machining and the finish of the surfaces. Friction is expected to be reduced greatly with future FPTs.
    - Materials such as Teflon can be used to coat the inner walls of the FPT shell to reduce friction between the spring and the FPT body.
    - Closer tolerances between the inner diameter of the FPT shell and the outer diameter of the spring body will ensure smoother travel and less vibration.
3. *Improvement to linearity.* The current prototype does not output the linear voltage response that it was designed to exhibit. However, it is understood that this problem related directly to the manufacturing quality. This aspect is expected to improve significantly, allowing the calibration method described for the linear system to be used. Having an accurate linear sensor means that less processing is required to obtain the position and force readings.

By improving the manufacturing of the FPT, the cost of the device can also be expected to decrease. The components used in the sensor will be inherently low cost and easy to machine given appropriate tools.

3.6 Conclusions

Currently multiple actuation systems are mostly restricted to industry-related robotics requiring complex but versatile high degree of movement operations. They utilise expensive, standardised actuators and sensors with few limitations imposed on energy consumption, volume and weight. In recent years, attention and activity has increased significantly in research related to the application of multi-actuation systems for non-industrial use. These research interests foresee a future in which multipurpose robots interact and function in the human world either as standalone units or as part of human-machine interfaces. Such a context calls for an entirely new approach in design of control systems and components. Accordingly, volume, weight, power consumption, simplicity, cost and the ability to react safely to an unpredictable, ever-changing environment are all valid concerns. For this reason, traditional methods employed in industrial settings are largely irrelevant and new appropriate measures are called for.

This chapter has presented the design of a novel hybrid force and position sensor applicable to effective and safe control of artificial muscles. The Force-Position Transducer (FPT) provides significant benefits including the simplification of multiple actuation systems requiring position and force feedback for each actuator. The hybrid sensor provides linear proportional signals without the need for amplification and complex signal processing. It is also an inherently robust, reliable and low cost device that can be placed in series between actuators and loaded segments.

In proposing the hybrid FPT, a first prototype was developed and tested as a proof of concept. This prototype utilised an optical-mechanical system in which movement caused the attenuation of light detected by phototransistors. The FPT was designed to produce two voltages directly proportional to the positions of the spring ends respectively. These signals are used to measure position and tension of the spring and hence the actuation cable.

In practice, the voltage-position relationships of the FPT were not linear. However, they were approximated following calibration to fourth order polynomial relationships. This lack of linearity was due to machining and construction methods and resources available.

Despite the lack of high precision, the FPT was used successfully in force and position control of a cable based actuation system. Furthermore, it was shown that using the FPT, force and position control algorithms can be made co-dependent. This will provide opportunities for safer and more effective function in human-machine interfaces.

Future work will require the FPT to be manufactured using more precise techniques to improve linearity and accuracy. Better materials and miniaturisation will also be necessary to make the novel hybrid sensor viable for use in actuation systems. Improvement of linearity using more sophisticated manufacturing techniques will streamline the calibration of the sensor (as described in section 3.2.3, page 22). Given the necessary enhancements, the FPT has the potential to play an active and significant role in future developments in artificial muscle technology and human-machine interfaces.

Chapter 4 A Review of Actuation Technology

4.1 Introduction

"*Actuators, the prime drive unit in any system (biological or mechanical), are responsible for transferring energy in its many forms into mechanical motion that permits interaction with the external environment*" (Caldwell, Tsagarakis et al. 2000).

Muscles are nature's elegant and effective solution to the problem of actuation. Muscles can be found with basically the same design, in a wide variety of living organisms, from the smallest insects such as mites to the largest animals such as elephants. However, in changed settings, muscles can be found to exhibit significant variation in performance (eg. force rise-time, contraction velocity, fatigue limits) (Hunter and Lafontaine 1992). Muscles perform several functions within the body. During locomotion, they animate the limbs in a controlled manner, within a variety of external environments. The main output variables within this control mechanism are contraction parameters (strain, tension, speed), stiffness and damping of each muscle body (Meijer, Rosenthal et al. 2001). Furthermore, muscles exhibit high energy densities and integrate actuation, support and fuel systems (Dickinson, Farley et al. 2000). In efficiency, muscles surpass internal combustion engines (Hunter and Lafontaine 1992) at around 40%.

Substantial effort is underway to develop actuators that can mimic muscle performance (Hunter and Lafontaine 1992; Inganäs and Lundström 1999; Meijer, Rosenthal et al. 2001; Caldwell and Tsagarakis 2002; Klute, Czerniecki et al. 2002; Del Cura, Cunha et al. 2003; Herr and Kornbluh 2004; Pratt and Krupp 2004; Baughman

2005; Ebron, Yang et al. 2006; Madden 2006). Recently, several emerging technologies have shown potential for creating viable artificial muscles. Of these, Electro-Active Polymers (Bay, West et al. 2003; Bar-Cohen 2004; Ebron, Yang et al. 2006; Yashin and Balazs 2006) and Shape Memory Alloys (Grant and Hayward 2000; Ebron, Yang et al. 2006; Madden 2006) have come closest to the performance and behaviour of muscles in actuation. However, currently no single material or system is available to compare with the biological muscle in all its properties. The effectiveness of muscle as a multi-faceted actuator coupled with its long cycle life-time[*] and its ability to self-repair, present an immensely challenging benchmark for human technology to aim for.

In 1999, Y. Bar-Cohen[†], a leading figure in the development of Electro-Active Polymer based artificial muscles, posed a challenge to the world-wide community[‡]: For scientists and engineers to develop a robotic arm that is actuated by artificial muscles to win an arm wrestling match against a human opponent. This endeavour has helped energise and focus artificial-muscle related research and development. While no winning artificial-muscle powered arm has been declared, the last three years have witnessed a number of earnest responses to the challenge[§].

---

[*] Cardiac muscles beat over 3 x 10^9 times in the life time of an average person (Hunter and Lafontaine 1992)

[†] Senior Research Scientist at NASA's Jet Propulsion Laboratory in California, USA

[‡] Bar-Cohen, Y., Ed. (1999). Proceedings of the SPIE's Electroactive Polymer Actuators and Devices (EAPAD) Conf., 6 th Smart Structures and Materials Symposium, Vol 3669.

[§] http://ndeaa.jpl.nasa.gov/nasa-nde/lommas/eap/EAP-armwrestling.htm

4.1.1 Why mimic muscles?

Since the beginning of technology, man has been searching for new ways to make things move. The question of *why mimic muscles* is related to another greater question: *why make actuators?* Actuators are used in almost every aspect of modern life, from car engines and camera zooms to industrial robots and artificial hearts. Actuators exist in many forms and shapes and each brings with it certain benefits at certain costs. The overwhelming majority of human-made actuation systems exist externally to the human body where there is greater room and freedom to tolerate the drawbacks of such technology. In recent years, there has been a significant trend in research aiming to create '*artificial muscles*', actuators that can work effectively and safely either within the body or in some manner interacting with it. In this setting, muscle-like technology would be of great value for implants and human assistive technology. Furthermore, such a technology could lead to the development of minimally invasive surgical and diagnostic tools in the form of highly dextrous catheters and endoscopes (Madden, Vandesteeg et al. 2004).

The design limitations for actuators working within or around the human body include:

- *Work (energy) density in respect to volume and weight*. Although the figures found in literature are varied, a typical skeletal muscle under normal conditions may exhibit energy densities of $40kJ/m^3$ and $38.5x10^{-3}kJ/kg$ (Madden, Vandesteeg et al. 2004). Artificial muscles should be within the vicinity of these figures (certainly within an order of magnitude).
- *Biocompatibility*. If used within the body, the artificial muscle must be biocompatible and safe from leakages, disintegration and must not pose a significant infection threat.

- *Safety*. As the actuators are designed to interact and work around the human body, safety is of prime concern. Inherent mechanical and electronic safeguards are of paramount importance.
- *Power supply and transmission*. Generally, actuators have been powered with a wide variety of sources including electrical, chemical, magnetic, thermal and mechanical. Methods of transmission of power to the actuators include using gears, cables and pneumatic and hydraulic systems. When used in and around the body, the power supply (e.g. battery, compressed air canister) and transmission system must not be excessively heavy, noisy, polluting, obstructive or potentially hazardous.

4.1.1.1 Relevant Characteristics of muscle

In 1922, A. V. Hill, one of the pioneers of muscle study and characterisation proposed an analogy between muscle's passive resistance and viscous shock absorbers (Hill 1953). Since Hill's early contributions, attempts at characterising the behaviour of muscles have led to several different mathematical models (Shadmehr and Arbib 1992; Durfee and Palmer 1994; Dickinson, Farley et al. 2000; Farahat and Herr 2005). These models are effective in predicting muscle behaviour because despite the variation in performance, the basic principles of skeletal muscles are highly consistent among different animals. A simple and widely used model is presented in Figure 4.1. The model consists of active, viscous and elastic elements placed in parallel followed by a series elastic component. In general, muscles can be considered as non-linear springs with changing resting length (Hunter and Lafontaine 1992; Shadmehr and Arbib 1992).

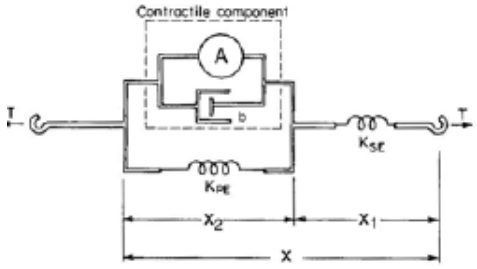


Figure 4.1 A model of skeletal muscles. *Adapted from Shadmehr and Arbib (1992). The muscle model places a contractile component in parallel to an elastic one (spring constant K_{PE}). The contractile components itself consists of an active element (A) in parallel to a viscous element (b). The parallel group is in series with an elastic component (spring constant K_{SE}).*

The output parameters of interest include:

- Length change (strain*) and rate of length change (speed)
- Muscle force (tension), maximum and continuous
- Stress (force per unit cross-sectional area)
- Stiffness (this parameter is complex and changes as a function of other values such as muscle length, rate of length change and force)
- Work density (normally based the per-unit-volume or per-unit-mass magnitude of mechanical work performed during one active cycle).
- Power density (equivalent to the speed and magnitude of work)

It must be noted that factors such as fatigue, long-term training or atrophy† of the muscle and activity of the nervous system can affect the above parameters.

In the following section, the output parameters of muscles measured under normal conditions will be compared with those of human-made actuators (see table in Appendix E for a numerical comparison of the various parameters).

---

* Strain is defined as the measure of the relative deformation during a distortion process. Specifically, Strain, ε = Length change ÷ Original length (Zhang 2001).

† wasting of the muscle due to long-term disuse

4.2 An overview of actuation technology

There is a lack of universally-agreed terminology to represent the type of actuators referred to in this thesis. These actuators function by exerting forces and causing displacement through shape change following their activation. "*The community as a whole has struggled with an appropriate definition of the technology, as evidenced by the wide variety of terms used in trying to describe it. Intelligent, smart, adaptive, active, sensory, metamorphic structures and/or materials and/or systems have all been used (many times interchangeably) to describe the research in the field*" (Spillman, Sirkis et al. 1996).

In order to maintain a consistent link to literature, the term "Smart Materials" will be used throughout this thesis[*]. A Smart Material may be defined as: A material with one or more properties that can be significantly altered in a controlled and reversible manner by external stimuli such as stress, temperature, light, moisture, pH, electric or magnetic fields. This property must change to such an effect that the material can be used as a sensor, an actuator or both. Examples include but are not limited to piezoelectric materials, shape memory alloys, electro-active polymers (EAP), rheological fluids, liquid crystal displays, thermochromic and photochromic materials (Janocha 1999).

[*] A simple comparison of width of use on the internet (using Google search engine) yields 1,090,000 466,000 and 173,000 hits for 'Smart Materials", 'Smart Structures' and 'Intelligent Materials' respectively. Within literature too, the term Smart Materials appears to be most widely used. One of the most relevant journals dealing with this material is called "Smart Materials and Structures".

As actuation systems continue to decrease in size and weight, so must their components. Traditional actuators such as electromagnetic, pneumatic and hydraulic motors undergo drastic reductions in power as they are scaled down. Increasingly, scientists and engineers are looking towards Smart Materials and nanotechnology to scale down actuators with minimal loss of energy density and power.

The materials field is vast and an in-depth review of all types of Smart Materials is beyond the scope of this study. However, a summary of the more relevant materials is presented here. The aim of this outline is to illustrate the technical challenges and potential benefits involved in the quest to create artificial muscles. One of the most promising materials, Shape Memory Alloys (SMA), will be the focus of this chapter and Chapter 5. Electro-Active Polymers are also of high interest and much of the work presented will apply equally to this class of material.

4.2.1 Electro-active polymers (EAP)

Within the last decade, new polymers have been developed to respond to electrical stimulation with significant shape change and in a controlled manner.

Electro-active polymers (EAPs) can be divided into two distinct groups: electronic and ionic. The first group are stimulated primarily by electric field or Coulomb forces. The latter, while still requiring electrical stimulation; involve mobility or diffusion of ions. The following is a list of some of the different polymer types in each category (Lewis, Moulton et al. 1997; Meijer, Rosenthal et al. 2001; Bar-Cohen 2002; Spinks, Liu et al. 2002; Ding, Liua et al. 2003; Smela 2003; Madden, Schmid et al. 2004; Baughman 2005).

- ELECTRONIC EAPS typically require high voltage electric fields at very low currents. In contrast to most other types of Smart Materials, Electronic EAPs produce impressive strains (up to 100% expansion) and practically useful forces at relatively high efficiencies. However, they are significantly restricted in usage due to their hysteresis and the requirement of high voltages (up to several kV).
    - *Ferroelectric Polymers* – Ferroelectric materials are analogous to ferromagnets. The application of an electric field aligns polarised domains causing a shape change that remains even after the field is removed.
    - *Dielectric Elastomer Actuators* – Actuation relies on the electrostatic attraction between conductive layers applied to two surfaces of elastomer films. The attraction induces large strains and large forces.
    - *Liquid Crystal Elastomers* (LCE) – LCE Actuators employ thermal or electrostatic energy to induce changes in the order and alignment of their liquid crystalline side chains. These changes in turn generate stresses in the polymer back bone resulting in actuation.
- IONIC EAPS rely on the transference movement of ions in and out of the polymer. This shift of ions is accompanied by structural changes resulting in expansions or contractions, shape changes that can be used for actuation.
    - *Ionic Polymer Gels (IPG)* – Composed of chainlike molecules dissolved in a solvent to form a semisolid, IPGs expand and contract with the movement of ions in and out of the gel. The ionic movement can be achieved using electrical stimulation.
    - *Ionic Polymer Metal Composites (IPMC)* – IPMCs use a wet polymer filling held between two opposing flexible metal electrodes. When

activated, ions are attracted by one electrode and repelled by the other. The ionic shift causes a reversible bending motion in the IPMC.

- *Conductive Polymers (CP)* – A Conductive Polymer works in similar ways to an IPMC but is itself used as an electrode. Two CP films are separated by an electrolyte. When the CP electrodes are stimulated with opposing charges, one of the conductive films expands while the other contracts. As with IPMCs, Conductive Polymers produce effective bending movements.
- *Carbon Nanotubes (CNT)* – CNTs are cylindrical structures composed of Carbon-Carbon bonds and a 60 atom-long circumference. One of the significant attributes of the nanotube is its strength. CNTs have a tensile strength 100 times higher than that of Steel. The Carbon-Carbon bond length can change when the ionic charge balance between the nanotubes and its surrounding electrolyte shifts. These shifts can produce actuations. However, while CNTs have shown potentially useful characteristics in the nano- and micro-scales, successful scaling into macroscopic levels has not been achieved.
- *Electro-rheological Fluids (ERF)* – ERFs undergo large changes to their viscosity when subjected to electric fields. The fluid is composed of particles (fine power) suspended in oil. Differences in the dielectric constants in the fluid and the particles are believed to cause the electro-rheological effect. In the presence of an electric field, the particles undergo induced dipole moments and form chains along the field lines. This changes the viscosity and yield stress of the ERF causing it to shift from a consistency of a liquid to

that of a gel. Potential applications of ERFs include shock absorbers, active dampers, clutches and variable flow pumps.

Because electrons move more easily than ions, electronic EAPs are faster than Ionic EAPs. Furthermore, the former have greater energy densities and can be operated in open air. Ionic EAPs on the other hand, must be bathed in liquid solvents. Electronic EAPs however, require much larger electric fields to cause them to activate. These fields (sometimes exceeding 150MV/m) can come close to breakdown levels for many polymers. Recent advances in the field have reduced the required field strengths for some electronic EAPs by almost an order of magnitude (Bar-Cohen 2004). Although this is a great improvement, the required fields are still too high for safe practical applications, particularly in close vicinity to humans.

Ionic EAPs are less efficient than their electronic counterparts but they react to voltages as low as 1-5 volts. In addition they are able to produce large strain bending motion. A general summary of advantages and limitations of both types of EAPs is presented (Madden, Vandesteeg et al. 2004):

- Advantages of electronic EAPs:
    - Large strains and forces
    - High bandwidth (fast rate of length change and high cycle frequency)
    - High energy and power densities
    - High efficiencies
    - Can function in open air
- Disadvantages of electronic EAPs:
    - Very high electric fields required albeit at low currents.
    - Require complex and large heavy accessory components such as high voltage transformers.

- Material breakdown likely at high voltages.
- Susceptible to fatigue

- Advantages of Ionic EAPs:
  - Require low input voltages and currents.
  - Suited for application requiring low force bending movement.
  - May one day be adapted to use fluids inside the body as electrolytes. paving a way for implantation of artificial muscles (Baughman 2005).
- Disadvantages of Ionic EAPs:
  - Low strains (mostly less than 1%) and forces
  - Slow
  - Low power output
  - Low efficiencies
  - Require encapsulation of electrolytic material
  - Subject to creep

4.2.2 Shape Memory Alloys (SMA)

Shape Memory Alloys were first discovered in 1951 when an alloy of gold and copper, after being deformed, was shown to return to its undeformed state once heated (Hunter and Lafontaine 1992). Since then, several other alloys have been discovered with similar properties. Of these, Nickel-Titanium (NiTi or Nitinol) has been the most widely studied due to its relatively large strains and low manufacturing costs. Many attempts have been made over decades of research to create artificial muscles using SMA wires or springs (Ikuta 1990; Kuribayashi 1991; Gorbet and Russell 1995; Grant and Hayward 1995; Lu, Grant et al. 1997; Schulz, Pylatiuk et al. 2001; Huang 2002; Ma and Song 2003; Nakamura, Saga et al. 2003; Kyu-Jin and Asada 2004; Bundhoo and Park 2005; Kyu-Jin, Rosemarin et al. 2006;

Lee, Son et al. 2006; Selden, Kyu-Jin et al. 2006). Since their discovery, SMAs have found uses in many applications including pipe fittings, valves, automobile fan clutches and thermal fuses (Grant 1999).

4.2.2.1 Mechanism

Otsuka and Shimizu (1970) showed that the 'shape memory effect' was caused by a structural phase transition from martensite to austenite. This phase shift accompanies thermodynamic changes. When the alloy is cool (below the phase transition temperature band), it is in the martensite phase and can be deformed. When heated above the phase transition temperature band, the austenite phase crystalline bonds re-align and the deformation is reversed (see Figure 4.2). In effect, the material 'remembers' its austenite shape. During the manufacturing process, the material's shape memory can be instilled using a 'training' process (Gilbertson 2000). Typically, SMA actuators are made as wires that can be stretched when cool. Heating the wire causes it to contract generating moderate strain (typically 3-5%) and significant stress (~200MPa) (Herr and Kornbluh 2004).

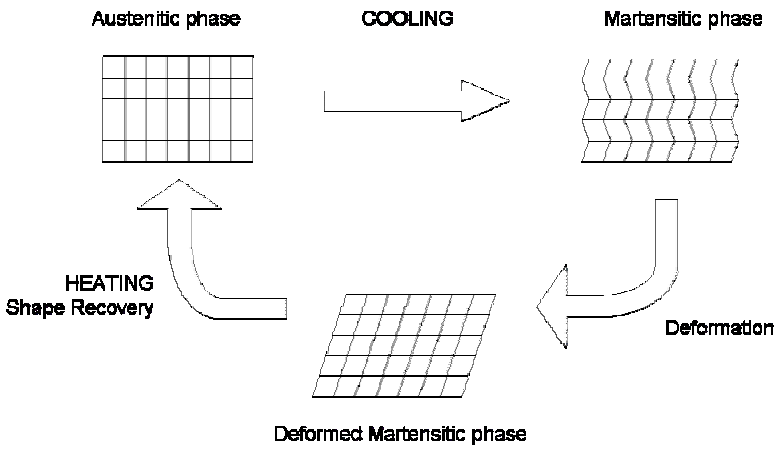


Figure 4.2 Heat induced phase changes of Shape Memory Alloys (adapted from Grant 1999).

4.2.2.2 One-way actuation:

Using SMA wires, actuation is achieved when, after heating, the wire contracts and works against a load. However, the shape change is not actively reversible. Hence a passive force must be applied to stretch the wire out before it can be re-activated. One method of doing so is to apply a force bias from an opposing spring or hanging weight (as will be demonstrated in this chapter). Another method is to use two SMA actuators in antagonistic arrangement. This method is analogous to antagonistic muscles.

4.2.2.3 Methods of Controlling Actuation of SMA

In 1984, Honma and others showed that SMA actuation can be controlled by electrical heating (due to the resistive nature of the material). Based on this work many groups began to apply the method to robotic applications (reviewed in Grant 1999). Currently, the most common approach to heating the material for actuation is electrical. However, other promising methods have been demonstrated such as Peltier heating and cooling (Selden, Kyu-Jin et al. 2006) and fuel-based systems (Ebron, Yang et al. 2006; Madden 2006; Jun, Rediniotis et al. 2007).

With electrical heating, several methods of controlling SMA-based actuators have been investigated. Most these can be divided into four categories:

1. *Linear control of current*: Linear proportional and proportional integral derivative (PID) control methods have been explored by many researchers (Ikuta 1990; Arai, Aramaki et al. 1994; Carpenter, Head et al. 1995; Gorbet and Wang 1995).

2. *Pulse Width Modulation (PWM)* (Hashimoto, Takeda et al. 1985; Kuribayashi 1986; Honma, Yoshiyuki et al. 1989; Gharaybeh and Burdea 1995; Ma and Song 2003).
3. *Segmental control:* More recently, researchers have explored methods of controlling discrete segments of SMA actuators using electrical taps (Lee, Son et al. 2006) and Peltier heaters (Selden, Kyu-Jin et al. 2004; Cho and Asada 2006; Selden, Kyu-Jin et al. 2006).
4. *Other methods:* Several other methods such as fuzzy logic (Nakazato, Kato et al. 1993), neural networks (Damle, Lashlee et al. 1994), feedback linearisation (Arai, Aramaki et al. 1995) and variable structure controllers have been applied (Grant and Hayward 2000; Elahinia, Koo et al. 2005).

Most of the abovementioned methods have shown some success in position and force control of SMA actuators. However, it is clear from the published works that complex control systems are required to control a highly non-linear actuator. The large range of factors affecting SMA behaviour (e.g ambient settings, stress, strain and material fatigue) hinder the robustness of the majority of controllers. SMA actuators are most suited to on/off type actuation and not proportional control. This consideration forms one of the bases of the work presented here.

4.2.2.4 Advantages of using SMAs as actuators:

There are many benefits in using Shape Memory Alloys as actuating materials. These include:

- Simplicity, compactness and low voltage requirement – SMAs are very simple to operate (particularly for on/off actuation) and can be effective actuators in wire form. Since they require very small voltages to operate

they can be safer to use than electronic EAPs. However, they do require high currents.

- Clean, silent, spark free operation – In contrast to many actuators such as electric motors and piezoelectric actuators, SMA actuators are silent when they operate. This can be highly advantageous for artificial muscle applications. They are also free of parts such as reduction gears and do not produce dust particles.
- High power/weight ratio and energy density – Comparing a large selection of actuation technologies, Ikuta (1990) concluded that SMA actuators featured the highest power to weight ratio at lower weights. The energy (work) density of SMAs is also very high; between 5,000 to 25,000 KJ/m^3 (Kornbluh, Pelrine et al. 1998; Roch, Bidaud et al. 2003; Madden, Vandesteeg et al. 2004). Muscles in comparison exhibit between 40-70 kJ/m^3 (Kornbluh, Pelrine et al. 1998; Madden, Vandesteeg et al. 2004).
- Long actuation life – When used within sustainable strain and stress limits, SMA actuators can be expected to last hundreds of thousands of operation cycles (see Table 4.1). However, this cycle life reduces dramatically with increase of work load strain and stress (Stalmans and Van-Humbeek 1995).
- Excellent corrosion resistance and biocompatibility (Stalmans and Van-Humbeek 1995).

4.2.2.5 Limitations of SMA actuators

Shape Memory Alloys are limited by a number of factors. Some of these are inherent to the material and may only be improved to a degree (such as inefficiency) while

other limitations are currently being addressed by many groups. The present thesis aims to contribute to this endeavour. The limitations include:

- Inefficiency – SMAs operate through heat and as such are limited by the Carnot efficiency model (Ikuta 1990) to at most 10%. However, under practical settings, SMA efficiency is typically below 1%. A significant amount of energy is lost as heat.
- Slow speed of activation - The speed of actuation is entirely dependent on the rate at which the temperature of the material can be changed. Typically, operation is faster for activation (heating) than it is for relaxation (cooling). This topic is further explored in the following section.
- Control difficulties - Hysteresis, nonlinearities, parameter uncertainties and un-modelled dynamics are among the factors that make accurate control of SMAs a challenge (Elahinia, Koo et al. 2005). The control method presented in this thesis aims to contribute considerably towards solving control problems.
- Lifetime dependency on work load – As noted above, excessive strains and stresses on SMA actuators can significantly reduce the functional lifetime of the material. Table 4.1 provides an example of the degree by which cycle life of SMAs can be affected by changes in applied work-load strains and stresses.

Cycles	Max. strain	Max. stress
10	8%	500 MPa
1000	4%	275 MPa
10,000	2%	140 MPa
100,000+	1%	70 MPa

Table 4.1 – Typical SMA actuator life cycles in respect to maximum strain and stress applied (Adapted from Stalmans and Van-Humbeek 1995).

Actuation speed of SMA material: One of the significant shortcomings of SMA-based actuators is their speed. The rate at which an SMA actuator can shift to austenite phase and return to martensite phase, depends on the thermal conductivity of the material. A variety of methods have been proposed to increase actuation speeds. However, most are associated with increasing complexity, size and weight and power consumption (Kuribayashi 1991) .

Typically, cooling time is several times slower than heating time (as observed in Figure 5.6, page 88). The differences in heating and cooling speeds of SMAs are highly dependent on several factors including:

- Diameter of the wire
- Method of activation (eg. DC current vs. pulse width modulation)
- Ambient temperature
- Medium surrounding the wire (eg. Air or liquid)
- Active cooling (eg air convection or Peltier active cooling)

Of the abovementioned factors, diameter decrease is most likely to increase actuation speed without significant drawback (Grant 1999). During cooling, heat transfer rates are highly dependent on surface area to volume ratios, a parameter inversely proportional to diameter. Figure 4.3 illustrates the relationship between actuation cycles per minute and SMA wire diameter (Gilbertson 2000) in still air.

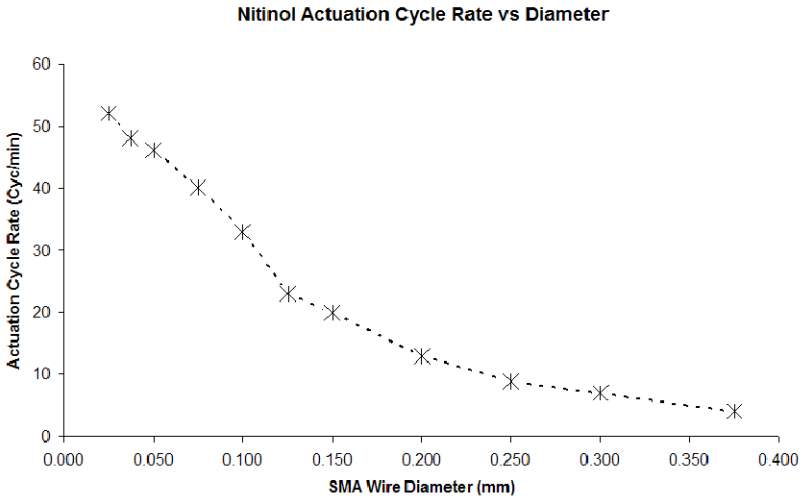


Figure 4.3 Relationship between actuation speed (cycles per minute) and SMA wire diameter (data obtained from Gilbertson 2000).

Though thinner SMA wires are faster, they are also much weaker. In fact, while cooling time has an inversely proportional relationship with cross-sectional diameter, actuation force is inversely proportional to its square (Gilbertson 2000). Therefore, doubling the speed means a quarter of the force. The loss of force can be recovered by increasing the number of actuators working together. One method is to place multiple actuator wires in parallel as a bundle. However, as will be shown in this experiment, bundling can be an inefficient method of combining work outputs of multiple actuators. In order to elucidate this, the theoretical boundaries of scaling will be determined using single wire performance.

Recent developments with fuel powered SMA actuators: Ebron and others (2006) coated the surface of Nitinol SMA with catalytic platinum. When exposed to a mixture of fuels and oxygen, the material was heated by oxidation on its surface (due to the presence of Pt catalyst). The system in effect converted chemical energy into thermal energy which powered the actuation of the SMA. Interruption of the fuel presence ceased the oxidation and allowed the wire to cool back into martensitic phase where it could be stretched back to its original length. The researchers

reported actuator stroke and power density comparable to those of natural muscle and generated stresses over a hundred times higher. Efficiency in this method of actuation is much higher than in electrical heating. This is because essentially all the chemical energy is converted directly into heat that is used to produce the actuation. The fuels tested were hydrogen, methanol and formic acid. The oxidant was either oxygen or air.

One of the benefits of fuel based actuation is the elimination of the energy storage problem. The main limitation with electrical power as the source of energy is that current storage technology is inefficient and has low energy density. Therefore, electrically powered actuators made equivalent to muscles in power output will require large and heavy batteries. The energy densities of hydrogen and methanol based fuels, like those of sugars, fats and gasoline are approximately two orders of magnitude higher than those of batteries. A further advantage of fuel circulation systems is the inherent option of liquid-cooling the SMA. Active cooling will increase the cyclic actuation speed of SMAs.

The main challenge with fuel based systems will be to create a safe circulation system that replaces the simple electrical connections that normally power actuators. To achieve this, pressures or diffusion gradients need to be generated to bring the fuel mixtures to the SMA for activation and to purge the fuels for relaxation (Madden 2006).

4.2.3 Other Actuation Technology

Notable actuation technologies that have been considered suitable actuators for use in human-machine interfaces include:

- Piezoelectric material. These crystals exhibit small strains in response to electrical charges. While, piezoelectric materials produce large forces (greater than 100 MPa), their small strains makes them unsuitable as macro-scale actuators (Del Cura, Cunha et al. 2003). More recently however, a class of piezo-based actuators have been developed that use vibrations, incremental and ratchet type movement to achieve macro actuation[*]. Although these technologies have tremendous potential due to their efficiencies and high energy densities, noise, complexity of circuitry and power consumption are some of the factors that need to be improved.
- Pneumatic Actuators – McKibben type and similar pneumatic actuators consist of braided flexible tubes which contract in length when they are inflated. Pneumatic actuators have been employed with promising results in a number of robotic applications (Nakamura, Saga et al. 2003; Kobayashi and Suzuki 2005). A significant limitation of McKibben type actuators is that they require compressed air to operate. Air compressors are bulky, noisy and are high energy consumers (thus requiring batteries as well). Pre-compressed air canisters are an alternative to compressor motors. However, compressed air storage is also bulky, heavy and must be refilled often to provide the required amount of energy.
- Electric motors – standard or miniature motors are currently the most common actuators. The advantage of using electric motors is that they

[*] Commercialised examples include the Squiggle Motor™, Piezo Legs™ and PiezoWave™

are easily available, reliable, and easy to control. Limitations include their requirement to be heavily geared to enable high actuation forces (electric motors are more efficient at high speeds). Geared motors are inefficient and exhibit large mechanical resistance (thereby stiffening the interface between actuator and load), a factor that makes them less suited for use as muscle-like actuators. In contrast, Series Elastic Actuators have shown some success in adapting geared motors to human-machine interfaces (see Chapter 2, section on SEA, page 9).

Chapter 5 A Method for Improving the Actuation Capabilities of Smart Materials for Use in Human-Machine Interfaces

5.1 Introduction

Despite their potentials, Smart Materials have not been utilised on a practical scale as reliable components of actuation systems. Following decades of research, scientists have shown functional properties of a range of materials that could equal or surpass muscles in actuation force and strain, efficiency, energy density and speed. However, the development of a single actuator that can compare with muscles in all of the above desired qualities continues to elude researchers. Thus, the development of mechanisms, systems and methods to exploit and fully optimise the potentials of Smart Materials is called for.

In this chapter, a work-summation concept is proposed to improve actuators for use in human-machine interfaces. The results of experiments applying the proposed methods to Shape Memory Alloy actuators (SMA) are presented and analysed. The aim of this study is to demonstrate that the work-summation system is an effective method for summing discrete actuations to produce a combined effort. Furthermore, it will be shown that in addition to its work-summation function, the method allows discrete stepwise control of SMA based actuators.

5.1.1 What are the potentials and challenges of current actuators?

When considering the properties and capabilities of actuators, it is important to make a distinction between potential and practical performance. Most actuators are complex systems made up of integrated active components. In an artificial muscle utilising 'smart' or active structures, the sub unit materials will exhibit properties different to those of the overall system. The theoretical boundary of work potential of such a system is a summation of the individual work-capability of each unit within. However, in practice, the overall work potential can never reach this boundary due to inevitable loss of energy within the system. Therefore, in effect there are two challenges ahead for designers of muscle-like actuators:

1. Improve the active/passive capabilities and properties of the unit structure (e.g. SMAs, Electro-Active Polymers, Carbon Nanotubes)
2. Minimise loss of energy when units are combined/scaled to amplify the system capability.

The first of these challenges is the focus of significant attention (Kornbluh, Pelrine et al. 1998; Inganäs and Lundström 1999; Pelrine, Kornbluh et al. 2000; Meijer, Rosenthal et al. 2001; Baughman, Zakhidov et al. 2002; Bay, West et al. 2003; Ding, Liua et al. 2003; Smela 2003; Spinks, Campbell et al. 2005) and will no doubt lead to impressive outcomes in the near future. The second challenge however, is one that can be tackled now using engineering principles and is the focus of this chapter.

5.1.2 The Work-summation Method

In this chapter a pulley-based work-summation system is proposed (see Appendix B for detailed description). This method is compared with an alternate form of multiple-actuator use, namely bundling. It is argued in this work that bundling is not an effective work-summation method. Furthermore, it will be shown that the pulley-based system satisfies the basic criteria of the work-summation system outlined below. In addition, the pulley-based system offers an advantageous form of actuation control with particular benefit to SMAs.

The development of a compact, efficient and effective work-summation system is necessary for the scaling of Smart Materials.

A work-summation system describes any mechanical assembly that can be used to combine the work output of discrete multiple actuators to yield a unified force and displacement. The functional criteria of such a system are as follows:

- The input work of each actuator unit is transmitted to the unified output with minimal loss of energy.
- Discrete function of each actuator contributes independently to the output.

5.1.3 Why is there need for work-summation?

The actuation output of Smart Materials can be scaled in many ways. Mechanical advantage (eg the form of fulcrum levers) can be used to amplify strain in exchange for force. Alternatively, output force may be increased at the cost of displacement. In order to amplify both force and displacement of the actuator, one method may be to

increase the size of the active material. Another solution is to recruit more actuator units to work together. The size-increase method is often counterproductive. Smart materials are relatively more effective in smaller sizes, volumes and diameters. For example, stress generated from a conducting polymer film of dimensions 50mm x 1mm x 0.05mm is ~5MPa (Spinks, Campbell et al. 2005), over 15 times greater than that of skeletal muscles (0.1 - 0.35 MPa). However, at this size, the polymer is capable of generating only 250mN at 1mm movement. Conducting polymers rely on the transference of ions across their surface area to shrink and expand the material. Hence, increasing the volume of the polymer will severely hinder the actuation as it would reduce the surface area to volume ratio. The most effective method of increasing work output would be to use multiple efficient polymer fibres together.

5.2 Methods and Materials

A series of test platforms were constructed and instrumented to measure and compare the actuation properties of four systems using Nitinol* SMA as the active material. These were:

1. Single-wire SMA referred to as 1-W
2. Four-wire SMA Bundle setup referred to as 4WB
3. Four-wire SMA Pulley setup referred to as 4WP
4. Five-Bit Binary SMA Actuator

Note: For the purpose of describing the methods and results of the experiments with references to the hanging masses used as loads, *gram* will be used as a unit of force (where 1 gram of force is taken to be equivalent to the force exerted by a hanging mass of 1 gram near the surface of the earth).

---

* Nitinol was selected as the test material as it was the most commonly and commercially available shape memory alloy (see section 4.4.2)

5.2.1 SMA single wire test (1-W)

This test involved the placement of a single Nitinol SMA wire of 90.7mm length and 150μm diameter within a test platform (see Figure 5.1). The wire was fixed at one end to a force transducer* via a non-compliant cable and connected at the other end to a hanging mass via a pulley. The connections were achieved using a 1mm diameter tubular copper crimp. The crimps both structurally fixed the SMA to the cables and electrically coupled each end to its respective electrode. The position of the hanging mass was measured using a high precision LVDT position transducer† with a full-scale range of 10mm (±0.1% linearity).

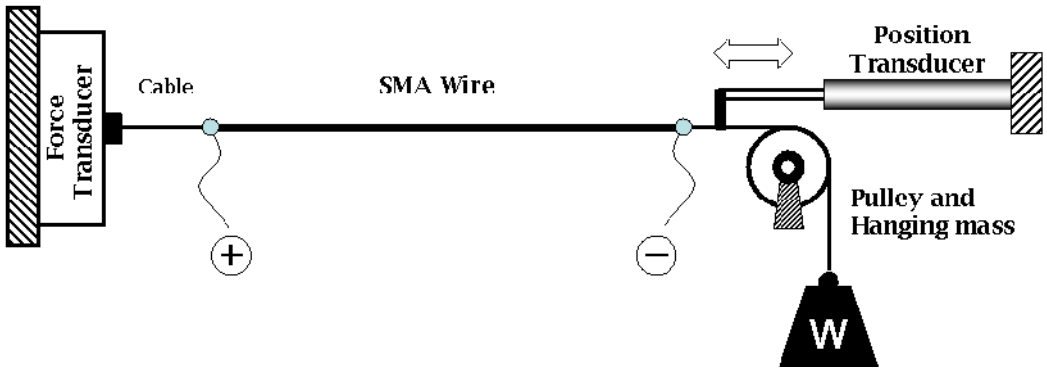


Figure 5.1 Schematic for the 1-W experiment.

5.2.1.1 Rated actuation properties of SMA wire

The following table presents the rated characteristics of the Nitinol SMA wire‡ used in this experiment (Gilbertson 2000).

---

* 15248A TACO 4-element DC strain gauge balance circuit and amplifier

† Macrosensors SE 750 – 250 General Purpose Single-Ended DC-LVDT Position Sensor

‡ Manufactured by Mondotronics Inc – Nitinol product name: Muscle Wire™

Type of wire	Flexinol 150 LT
Activation temperature	70˚
Wire resistance	50Ω/m
Recommended current (to heat wire)	400mA
Rated load limit (for maximum cycle life)	3.43N tension
Recommended deformation force (to stretch wire)	0.61N tension)
Recommended deformation	3-5%
Maximum cycle rate	20 cyc/min

Table 5.1 – Characteristics of Nitinol SMA material used

5.2.1.2 Control of actuation

Activation of the SMA was computer controlled and connected by an electrode at each end of the wire. Force and position measurements were continuously digitised and recorded by the control software written in Labview™. Upon activation, the system triggered a transistor-based 8-channel control circuit to deliver 400mA to the wire (see Appendix D for control circuit)[*]. The resulting heat forced the wire into austenite phase and a load bearing contraction (lifting the hanging mass). In a series of tests, the mass of the hanging weight was discretely increased: 75g, 100g, 150g, 200g, 250g, 275g, 300g, 325g, 375g, 400g, 450g, 475g and 500g. With each hanging mass, the SMA wire was activated and relaxed cyclically four times. The transient displacement and maximum contraction was detected and recorded for each weight and later averaged and plotted. The force transducer was used to verify the transmission of tension via the pulley and to observe transient force changes due to the inertia of the hanging mass.

[*]An 8-channel digital output, 16-channel analogue input data acquisition card was used: National Instruments DAQ Card Model: PCIMIO 16XE-50

5.2.2 SMA Four Wire Bundle Actuator Test (4WB)

Four Nitinol wires of length 92.0 ±0.5mm* were bundled at one end using a crimp which was then connected to a hanging mass and a position transducer (as in the 1-W setup). The opposite end of each wire was fixed to a dedicated force transducer. In order to implement parallel and close proximity of the four wires, direction changing pulleys were used to distribute the connecting cables to a larger space accommodating the placement of the force transducers (see Figure 5.2). The voltage signals from the position and force sensors were recorded throughout the experiment. The following series of hanging masses were applied as loads to the 4WB actuator: 100g, 200g, 350g, 400g, 500g, 800g, 1000g and1200g.

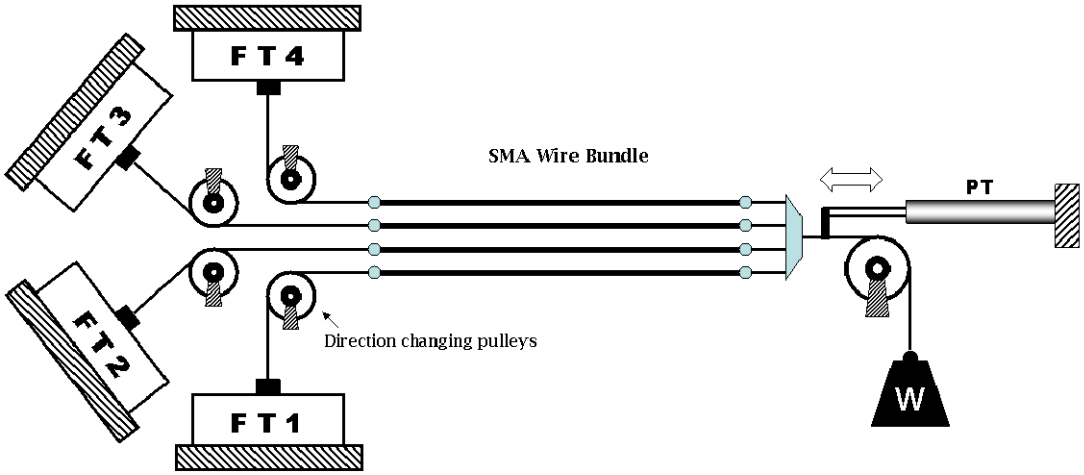


Figure 5.2 Schematic for the 4WB experiment. Direction changing pulleys were used to bring the connecting cables from the bundle of SMA wires to the distributed force transducers (FT1-FT4) perpendicularly.

* The lengths of the wires used in this test differed to that of 1-W experiment (90.7mm) due to the difficulties encountered in crimping the wires at precise points. This discrepancy however, is not expected to hinder the results as actuation displacements were measured as a percentage of wire length.

5.2.2.1 Control protocols

Power electrodes were connected to each end of each wire as in the 1-W setup. Each SMA wire was powered independently. Two open-looped control protocols were implemented for the combined actuation of the bundle.

1. Combined powering of all SMA wires
2. Individual and sequential activation of the wires

The individual activation test was applied only up to a hanging mass of 3.43N (350g mass). This ensured that no single wire underwent a tension greater than the recommended (rated) maximum (see table 5.1) prior to the completion of the combined activation tests. Changes in individual wire tension and combined displacement were recorded during the tests.

5.2.3 SMA Four Wire Pulley Actuator Test (4WP)

Four Nitinol wires of length 87.0 ±0.5mm* were connected to the four force transducers using the direction changing pulleys as described with the 4WB set up. At the opposite end, the wires were interconnected using free moving pulleys (implementing the work-summation principles). The wires were grouped in two pairs. In each pair the two wires were linked via a primary level pulley. The two primary pulleys were linked via a secondary level pulley producing a single mechanical output. The actuation output was connected to the hanging load and the position transducer as described in the previous tests (see Figure 5.3). The position and force sensors were monitored throughout the experiment. The

* As noted in section 5.2.2, the lengths of the wires used in this test differed to those of 1-W and the 4WB tests due to the difficulties encountered in crimping the wires at precise points.

following series of hanging masses were applied to load the 4WP actuator: 100g, 200g, 350g, 400g, 600g, 800g, 1000g, 1200g and1400g.

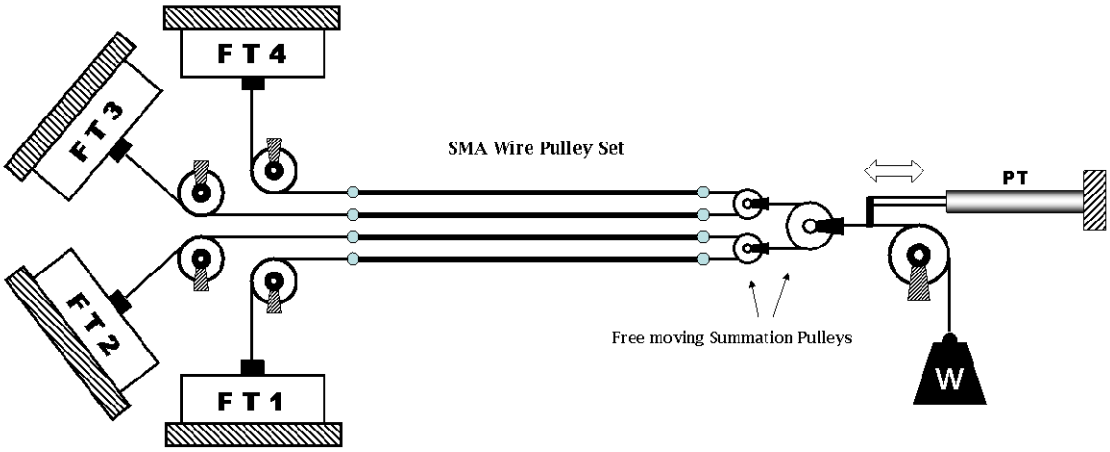


Figure 5.3 Schematic for the 4WP experiment. The fixed end of each wire was connected to a force transducer in a manner identical to the 4WB setup. The free ends of the wires were interconnected via work summation pulleys yielding a single actuation output.

5.2.3.1 Construction

The work-summation pulleys (see Appendix B) were constructed using low friction Teflon tube loops held within acrylic outer cages. These small and wheel-free pulleys allowed parallel cable separations of 4mm without the possibility of cable slippage (see Figure C.1in Appendix C).

5.2.3.2 Control protocols

Power supply electrodes were connected to each end of each wire as in the 1-W setup. Each SMA wire was powered independently. The same control protocols used with the 4WB actuator were repeated in this test.

1. Combined powering of all SMA wires
2. Individual and sequential activation of the wires

The distribution of the overall load between the four wires allowed the individual wire activation tests to be applied up to a load of 11.77N (using hanging mass of 1200g) before any individual wire tension exceeded the rated maximum (see the results section). Changes in individual wire tension and combined displacement were recorded during the tests.

5.2.4 SMA Five Bit Binary Actuator Test

In the 4WP test, a work-summation system was implemented using four identical SMA wires and three pulleys to produce a single combined output. This concept was expanded and refined in the 'five-bit Binary Actuator' test using thirteen pulleys, six SMA wires of alternate lengths, five power supply outputs and a 4:1 ratio wheel Figure 5.4). The design and construction of the actuator assembly are described in the following section.

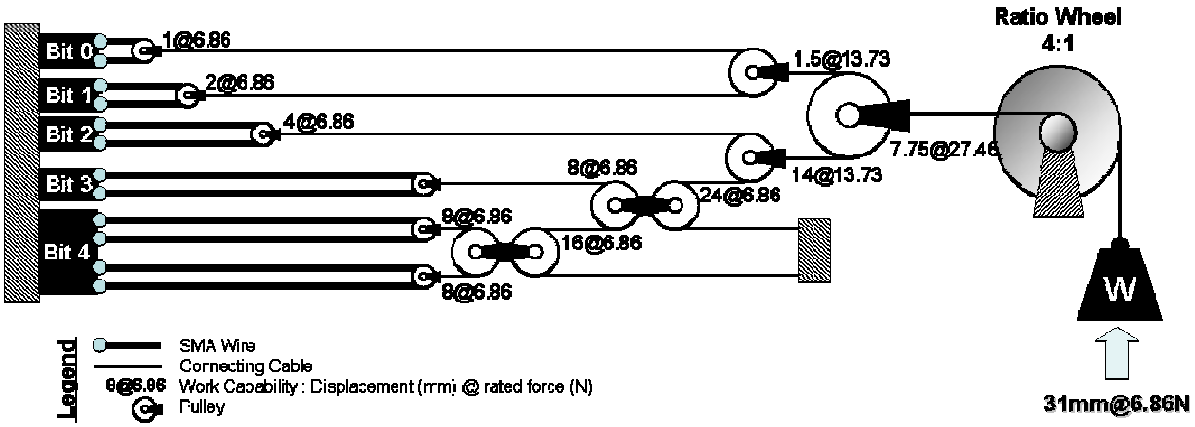


Figure 5.4 The Binary Actuator experimental setup schematic.
See Appendix C for photos of the actual setup.

5.2.4.1 Components and function of the Binary Actuator

Five power supply outputs (see Appendix D for control circuit) were independently activated using the control system. Each output powered a single 'bit' of the actuator. Bits zero to four each consisted of independently activated SMA-based units with discrete actuation outputs (see Figure 5.4). Each bit was designed to produce a work output equal to twice that of its preceding bit. A work-summation pulley network was implemented to allow each bit to contribute discretely to the Binary Actuator output. The actuator 'bits' were prepared as follows:

- **Bit-0**: A single Nitinol SMA wire (as described) of total length 40.0±0.5mm was looped around a pulley[*]. This yielded a theoretical actuation output of **1mm** at **6.86N** (700g) force (rated) at the final pulley output.
- **Bit-1**: A single Nitinol wire of length 80.0±0.5mm was looped around a pulley yielding an actuation output of **2mm** at **6.86N**. The outputs of Bit-0 and Bit-1 were added using a secondary level pulley.
- **Bit-2**: A single Nitinol wire of length 160.0±0.5mm was looped around a pulley yielding an actuation output of **4mm** at **6.86N**.
- **Bit-3**: A single Nitinol wire of length 320.0±0.5mm was looped around a pulley yielding an actuation output of **8mm** at **6.86N**.
- **Bit-4**: Two Nitinol wire pulley loops, each of length 160.0±0.5mm were looped around a secondary level pulley to produce an actuation output of 8mm at 13.73N. A reverse pulley with a cable loop fixed at one end was used to double the displacement to **16mm** and halve the force to **6.86N**.

[*] This primary level pulley served two purposes. Firstly it allowed the SMA wire to be powered at two fixed ends instead of free moving ends. Secondly, the pulley acted as mechanical leverage to increase actuation force. A 40mm Nitinol wire normally actuates 2mm at 3.43N (rated). However, when looped around the pulley, the output becomes 1mm at 6.86N.

- A series of summing and reverse pulleys were incorporated to produce a combined actuation output of 7.75mm at 27.46N (2800g). A 'ratio wheel' utilising an inner reel diameter of 5mm and an outer reel diameter of 20mm was placed in series with the actuation output cable. This allowed an amplification of displacement and reduction of force by a factor of four. The final actuation output of the assembly subsequent to the ratio wheel was calculated to **31mm** rated at **6.86N** (700g) force.

5.2.4.2 Control of Binary Actuator

Control of the Binary Actuator was implemented using software written in Labview™. A power circuit provided five-channels of current supply to the actuator bits. The control input was a binary value between 0-31. The control system powered the binary bits according to the control value. ie. Bit-0 corresponded to a value of 1, Bit-1 to 2, Bit2 to 4, Bit-3 to 8 and Bit-4 to 16. A control input of 19 for example, activated bits 0, 1 and 4 (1+2+16).

5.2.4.3 Open and closed loop control of the Binary Actuator

The Binary Actuator was initially tested using open-loop position control. Open-loop response was compared with an actuation model. Subsequently, the Binary Actuator was configured to work with the hybrid Force-Position Transducer (FPT) described in Chapter 3 for closed-loop force and position control. Reference position (Figure 5.5a) and force (Figure 5.5b) sensors were used to assess effectiveness of the control system working with the FPT.

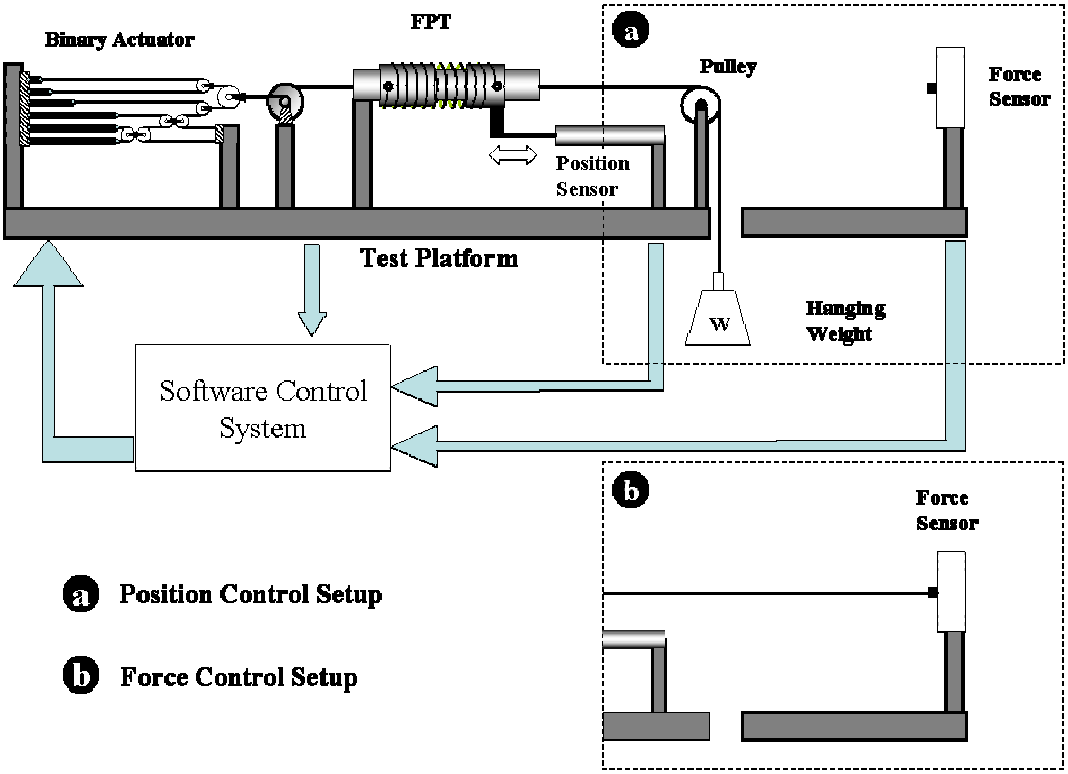


Figure 5.5 Closed loop control of the Binary Actuator. Using the setup described in Chapter 3, the Binary Actuator was tested with the FPT. Force and position control were assessed using reference sensors.

5.2.4.4 Non-cooled and active-cooled settings

When operated in still air, SMA wires heat to austenite phase much faster than they can cool down (see results and discussions section on actuation speed).

While testing the open-loop position control, the effects of active cooling using forced air (two 80mm computer cooling fans) were compared with non-cooled conditions. During the FPT-based closed-loop tests, active cooling was applied at all times.

5.3 Results

5.3.1 SMA single wire test results (1-W)

Figure 5.6 shows the actuation response of a single wire lifting a pre-load of 1.47N (150g). The SMA wire is shown exerting 5.88×10^{-3} Joules of work (4mm at 1.47N). To enable a comparison of mechanical work exerted by the alternate actuation configurations, work output is normalised to unit length of actuator. The wire lengths in the different setups vary slightly (as described in the Methods section). Hence, to normalise the results of the different setups, it is necessary to compare strain instead of absolute length change.

Strain is calculated based on displacement from a baseline (position of reference) to a stable position following activation. The rise and fall-times are inclusive of 2% - 98% of the total strain. Accordingly, the graph in Figure 5.6 shows a total normalised work of 6.47×10^{-2} J/m. This is calculated from the operation of the single wire of 90.7mm martensite length, lifting a 1.47N load (150g hanging mass), by 4mm.

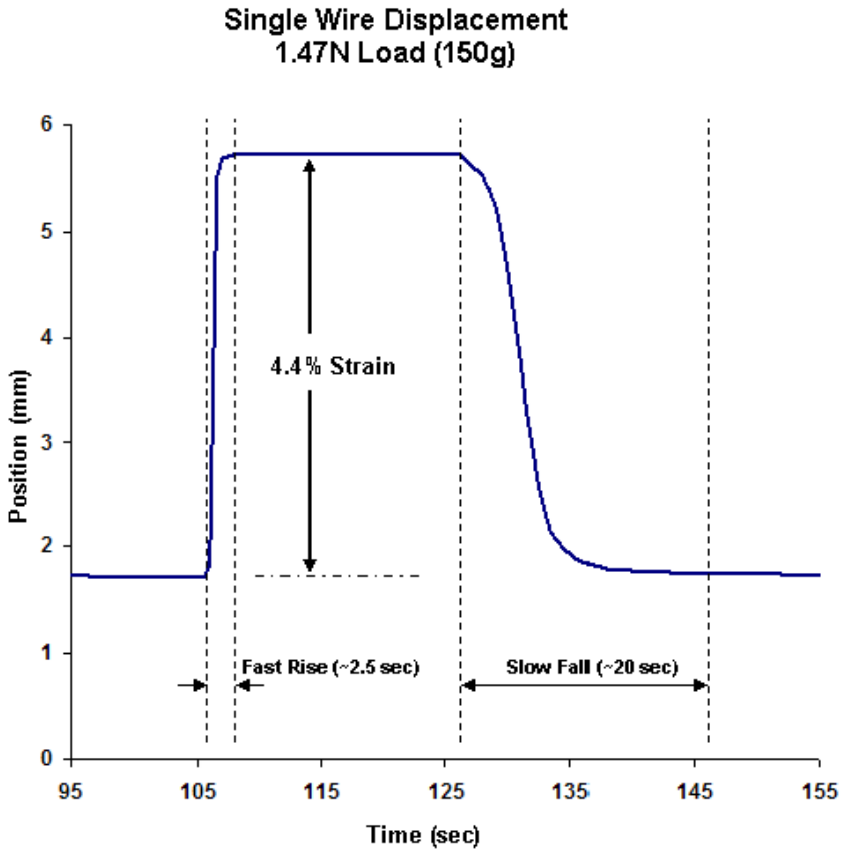
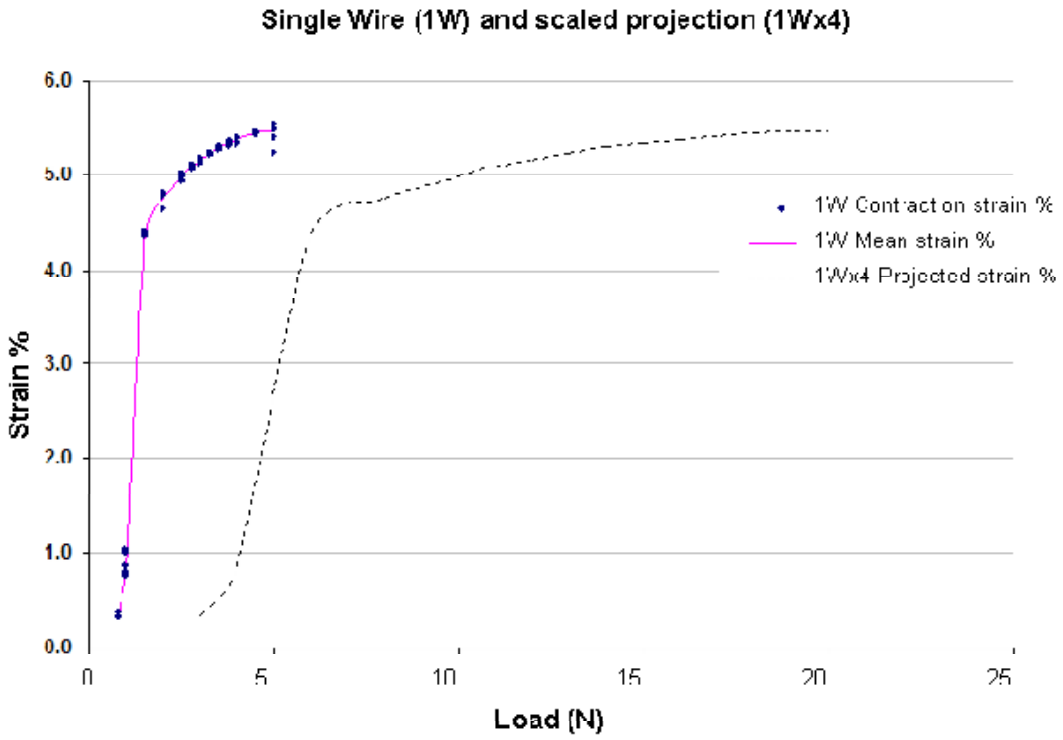


Figure 5.6 The characteristic actuation behaviour of a single Nitinol wire under load. Under non-cooled ambient settings, the activation (rise-time) of the actuation is over six times faster than relaxation (fall-time).

Figure 5.7 depicts single wire strains at different pre-loads. At 4.90N (500g), the wire failed (by breaking at the connection to the load cable). The recommended maximum load for continuous operation (millions of cycles) is 3.43N (350g, see Table 3.1). At this load, the SMA wire contracted by 5.3%.

Using the data obtained from the 1-W tests, a theoretical four-times-scaled curve was projected (1Wx4). This was achieved by multiplying actuated loads by a factor of four.



Load (N)	Mean Strain%	Standard Error n=8
0.74	0.33	0.019
0.98	0.87	0.050
1.47	4.38	0.005
1.96	4.77	0.027
2.45	4.99	0.012
2.70	5.09	0.006
2.94	5.17	0.005
3.19	5.23	0.003
3.47	5.30	0.004
3.68	5.34	0.005
3.92	5.37	0.009
4.41	5.46	0.005
4.90	5.46	0.036

Figure 5.7 Strain vs. load results for single SMA wire and its scaled projection. Graph shows the results of all trials, the mean strain of the single wire tests and the scaled projection. The table contains the mean strain and standard error for each hanging mass tested.

5.3.2 SMA four wire bundle test results (4WB)

The 4WB test was used to assess the actuation capacity of a bundled actuator while measuring individual wire forces.

5.3.2.1 4WB combined and sequential activation

Figure 5.8 contains graphs showing the actuator displacements and individual wire forces during combined (left side) and sequential activation (right side). The powering diagram in between the position and force graphs displays the activation state of each wire correspondingly. A total load of 350g is shown here.

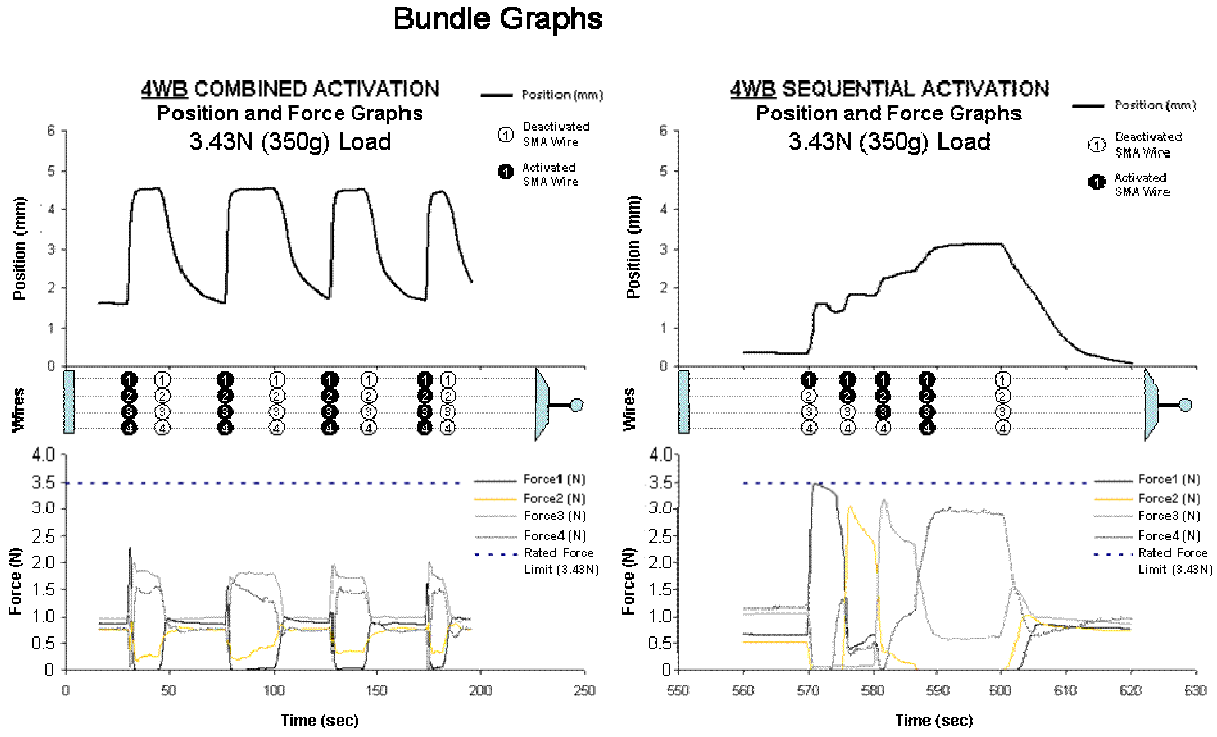


Figure 5.8 Combined and sequential activation of 4WB. Four-Wire Bundle (4WB) Actuator position and individual force graphs are shown in respect to SMA wire activation (middle diagrams). The graphs on the left represent combined wire activation. On the right, graphs show actuation when wires are activated sequentially.

5.3.2.2 4WB Actuator Strain

At 3.43N (350g) load and using combined activation, the 4WB actuator was able to achieve approximately 3% strain (in relation to the length of active portion of the bundle). Sequential activation of the four wires results in a gradual increase in the overall displacement up to 3% (see Figure 5.8 right). As with the 1-W experiment, the calculated strains are referenced to the baseline prior to activation. The baseline tended to shift slightly during each experiment and further as additional loads were applied (due to elastic strains).

Average strains were determined under the variety of loads described in the methods section. The results are depicted in Figure 5.10 (page 93). This graph shows a gradual increase in strains with increase in pre-load in a shape similar to that observed with 1-W. As expected, the 4WB curve is stretched out further to the right on the load axis. However, this projection is considerably less than that of the 1Wx4 curve.

5.3.3 SMA four wire pulley test (4WP)

The 4WP test was used to assess the actuation capacity of the pulley-based method of work-summation.

5.3.3.1 4WP combined and sequential activation

Figure 5.9 contains graphs showing the actuator displacements and individual wire forces during combined (left side) and sequential activation (right side). The powering diagram in between the position and force graphs displays the activation

state of each wire correspondingly. For direct comparisons with the equivalent 4WB test (refer to Figure 5.8) a total load of 3.43N (350g) is shown here.

5.3.3.2 4WP Actuator Strain

At 3.43N load and under combined activation, the 4WP actuator was able to reach an average strain of 3.76%. This was greater than the strain produced by the 4WB (3%) for the same load (over eight trials, P=0.03). A more comprehensive comparison of strain achieved by different actuators will be considered later. Average strains were determined under the variety of loads described in the methods section. The results are depicted in Figure 5.10.

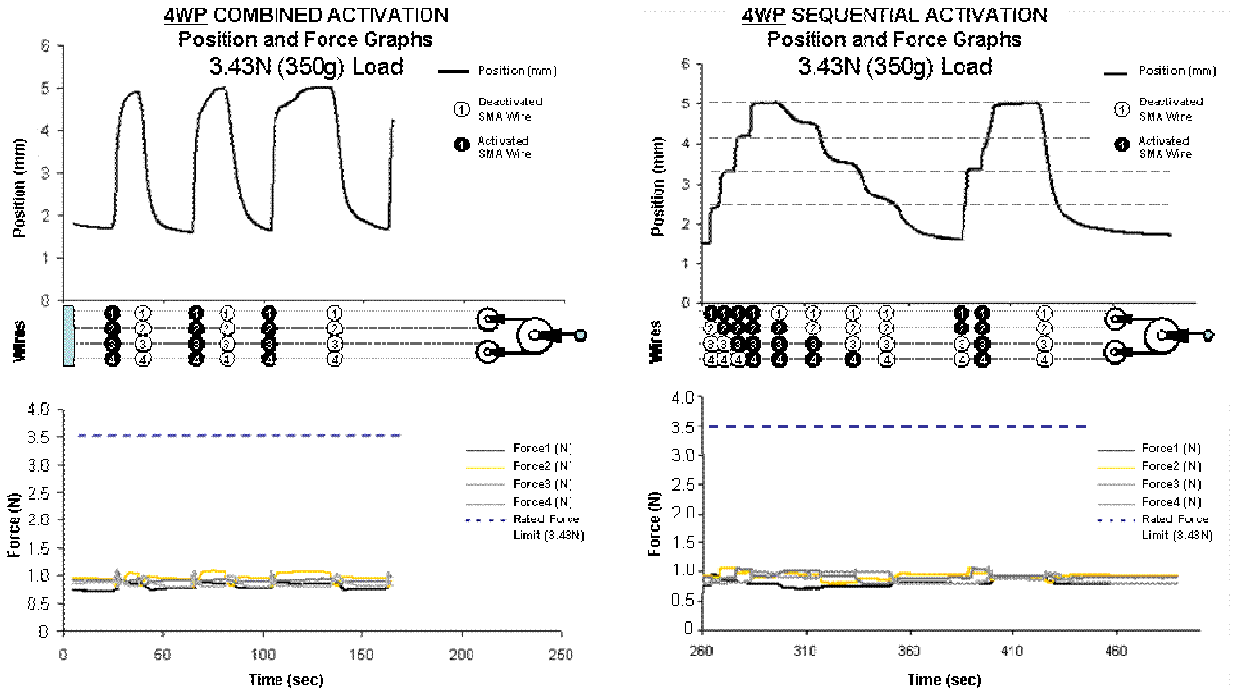
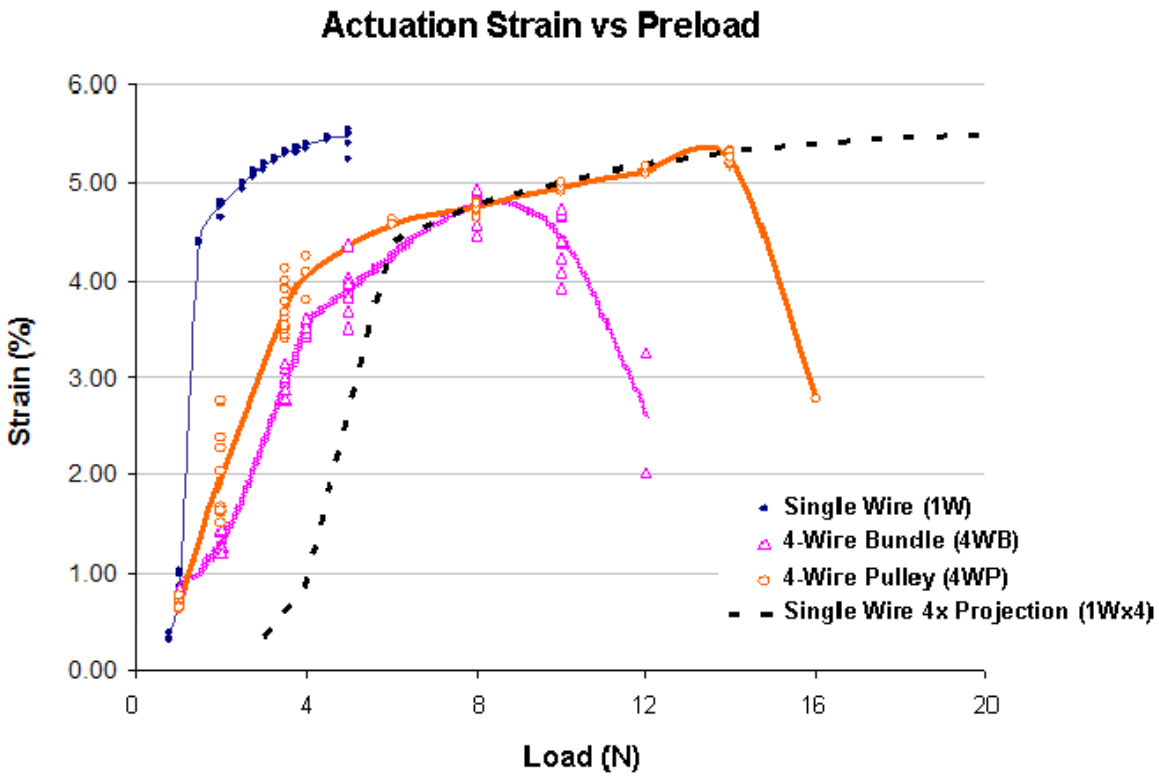


Figure 5.9 Combined and sequential activation of 4WP. Four-Wire Pulley (4WP) Actuator position and individual force graphs are shown in respect to SMA wire activation (middle diagrams). The graphs on the left represent combined wire activation. The graphs on the right show actuation when wires are activated in a variety of orders.



	4WB		4WP	
Load (N)	Mean Strain%	Standard Error n=8	Mean Strain%	Standard Error n=8
0.98	0.84	0.005	0.66	0.018
1.96	1.30	0.038	1.99	0.159
3.43	2.97	0.046	3.70	0.082
3.92	3.53	0.024	4.05	0.093
4.90	3.90	0.056		
5.88			4.57	0.013
7.84	4.79	0.063	4.75	0.019
9.81	4.43	0.112	4.95	0.024
11.77	2.64	0.612	5.12	0.029
13.73			5.23	0.021
15.69			2.78	

Figure 5.10 Comparison of results showing strain versus hanging mass loads for the alternate actuation systems under study. *The 1Wx4 projection was obtained by scaling the single wire data by a factor of four on the load axis. The solid lines represent averages.*

5.3.3.3 Wire force distributions across the loads

Even distribution of the total load among the wires is a key function of the work-summation system (represented by the 4WP actuator). By minimising the forces on each wire, the actuator is able to work with greater loads while maximising the cycle lifetime of the SMA material. Figure 5.11 contains individual wire force graphs for the 4WB and 4WP actuators. For each actuator, forces are shown during activation as well as relaxation. In addition, in each graph, the average as well as maximum forces are plotted. Using these plots, the sustainable load limits for each actuator were determined. The rated work zone was the range of loads up to and including the point at which any of the forces exceeded 3.43N (rated limit for an individual wire)

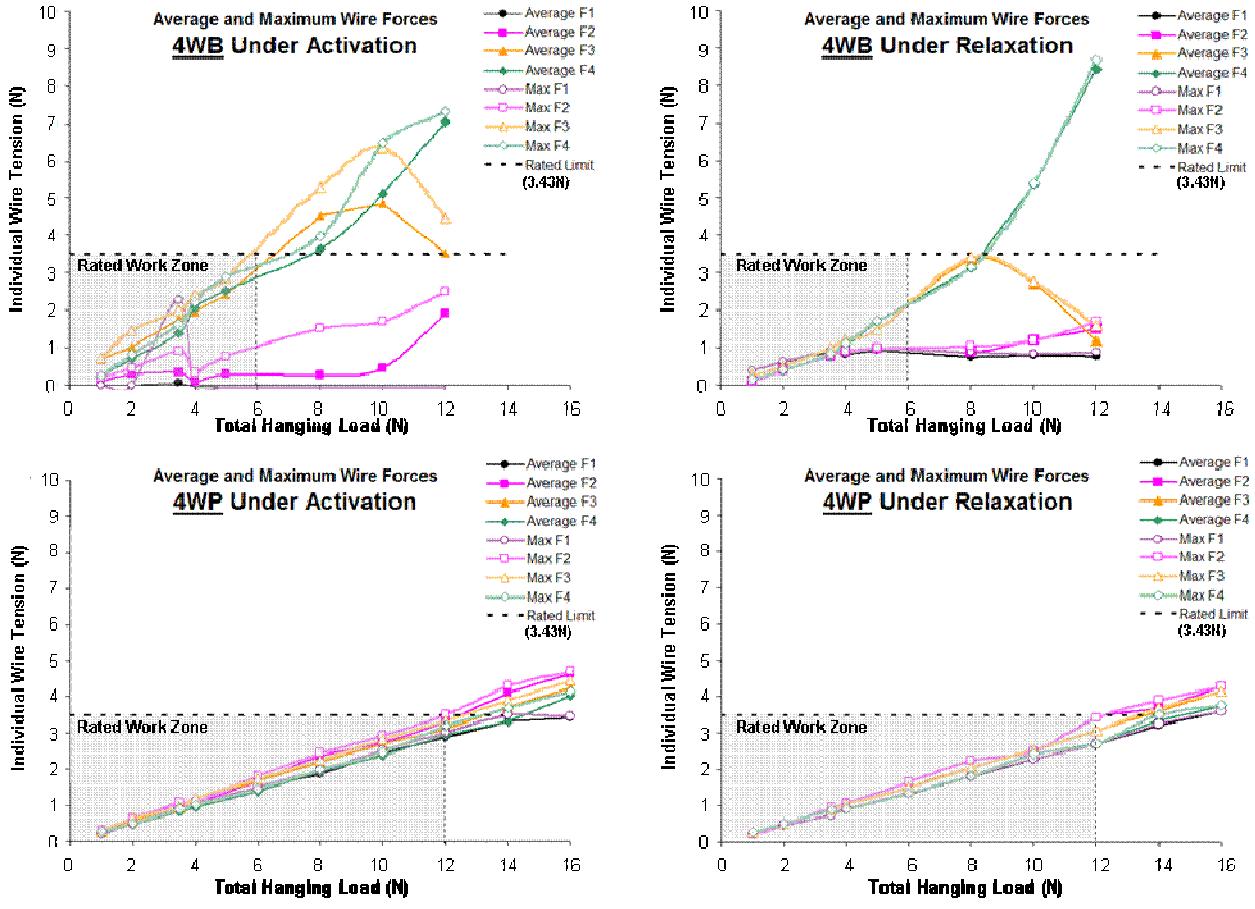
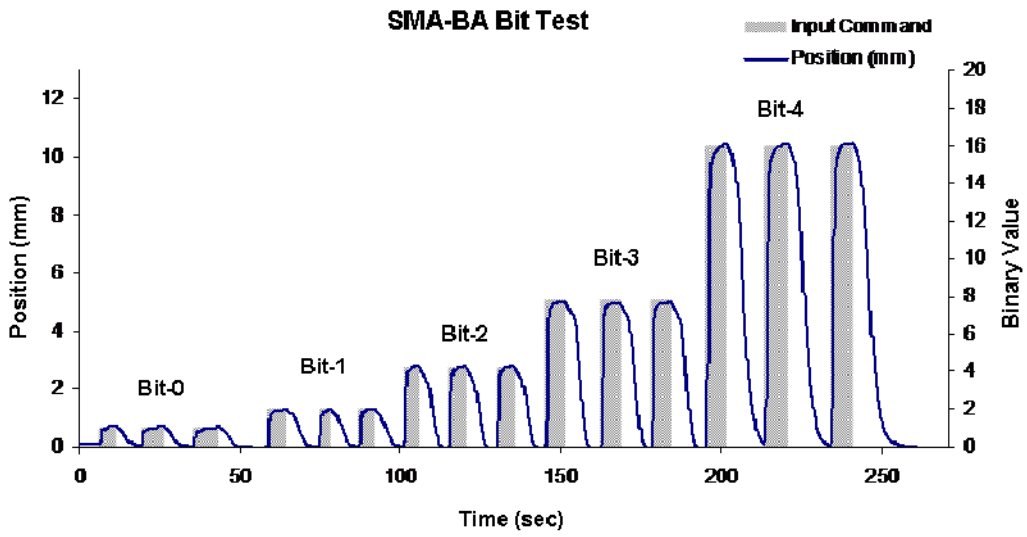


Figure 5.11 Average and maximum wire forces across a series of loads for the 4WB and 4WP actuators. The graphs on the left show forces during activation. The graphs on the right represent forces during relaxation.

5.3.4 SMA Binary Actuator (SMA-BA) results

Following assembly of the SMA-BA actuator, each bit was powered individually three times while the resulting displacement of a 6.86N load (700g hanging mass) was measured at the output. The results are presented in

Figure 5.12.



Binary Value	Mean Position (mm)	Standard Error n=3	Ratio to bit 1
1	0.69	0.014	1.00
2	1.35	0.004	1.95
4	2.83	0.007	4.09
8	5.51	0.012	7.96
16	11.14	0.011	16.08

Figure 5.12 Binary Actuator bit test. This graph shows the displacement responses of the SMA-BA system following the activation of individual bits. The displacements are consistently proportional to the bit values. However, the absolute displacements are lower than the binary values.

5.3.4.1 The bit-change artefact

The Bit test showed that the SMA-BA is capable of actuating with different sized steps. These steps may be in effect 'added' to produce a controlled position output. However, although each bit generates a stable and repeatable output, the step rise

and fall are different from each other and highly dependent on environmental conditions (e.g. ambient temperature and humidity). Mean rise-time was 1.96 ±0.18 secs (n=15) while average fall-time was found to be 14.5 secs ±0.45 secs (n=15). This anomaly causes an undesirable artefact in the actuation output whenever a bit rise coincides with another bit fall. Figure 5.13 demonstrates this occurrence using bit-0 and bit-1. Position is increased from 0 to 1mm and subsequently to 2mm. For the position output to increase by 1mm, bit-0 must be turned off while bit-1 is turned on. However, as the relaxation of bit-0 is slower than the activation of bit-1, the latter contributes with its maximum displacement before the former has stretched fully. The result is an overshoot that slowly reduces with the gradual relaxation of bit-0.

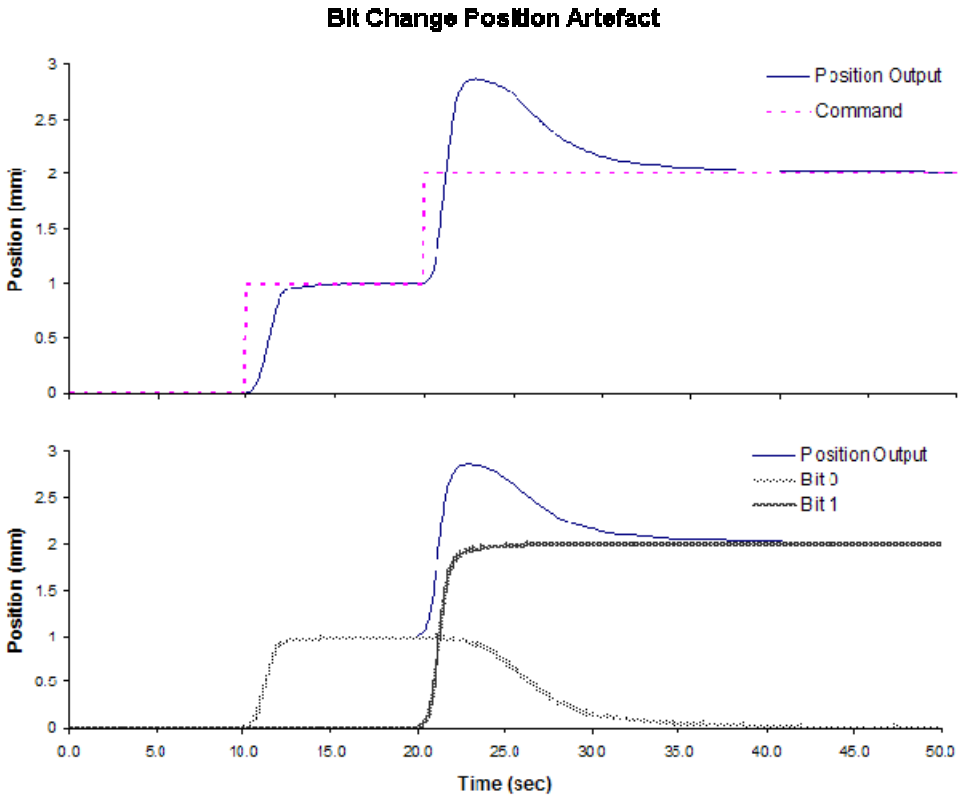


Figure 5.13 Bit change position artefact. The position output curve (upper graph) is a summation of the contributions of bit-0 and bit-1 (lower graph). The curve shows an overshoot when position is increased by one step. The underlying reason for the overshoot is apparent from the individual bit responses. Bit-0 is slower to relax than bit-1 is to activate. Hence, the displacement contribution of bit-1 manifests in the output before bit-0 has had time to fully relax.

Theoretically, the bit-change artefact can be eliminated by equalizing bit-rise and bit-fall durations. In practice this is not possible as heating and cooling mechanisms are often facilitated by different processes (e.g. heating by electric current and cooling by air convection). However, reducing the cooling time toward heating time will act to minimise the observed overshoot.

5.3.4.2 Comparing SMA-BA model with test rig response

A model, developed in software (Labview™), was used to simulate the response of the SMA-BA system. The model was prepared by recording the displacement-time relationship of a single wire (see Figure 5.6) during activation and relaxation. The recorded data was then used to predict the response of the SMA-BA by summing the displacement of five independent and parallel channels activated in binary increments. The displacement-time output of each channel was scaled by two in respect to its preceding channel. The model assumed a non-cooled system and predicted the displacement response of the system to a binary input increment from 0 to31. The simulation showed bit-change artefacts (Figure 5.14). The model response was compared with an actual response for an incremental control input.

Prior to testing the incremental response, the control input was calibrated. To calibrate, the maximum binary input '31' was entered (i.e all wires active) and the resulting displacement was measured. In the experiment depicted in Figure 5.14, the binary input 31 produced a displacement of 28.3mm. Subsequently, all binary inputs (0 to 31) were adjusted by a scale of 28.3/31. Next, the control input was incremented by one binary value every 10 seconds. The actuator displacement is superimposed with the simulated model and the adjusted calibration input. The results show an actual response consistent with predicted behaviour.

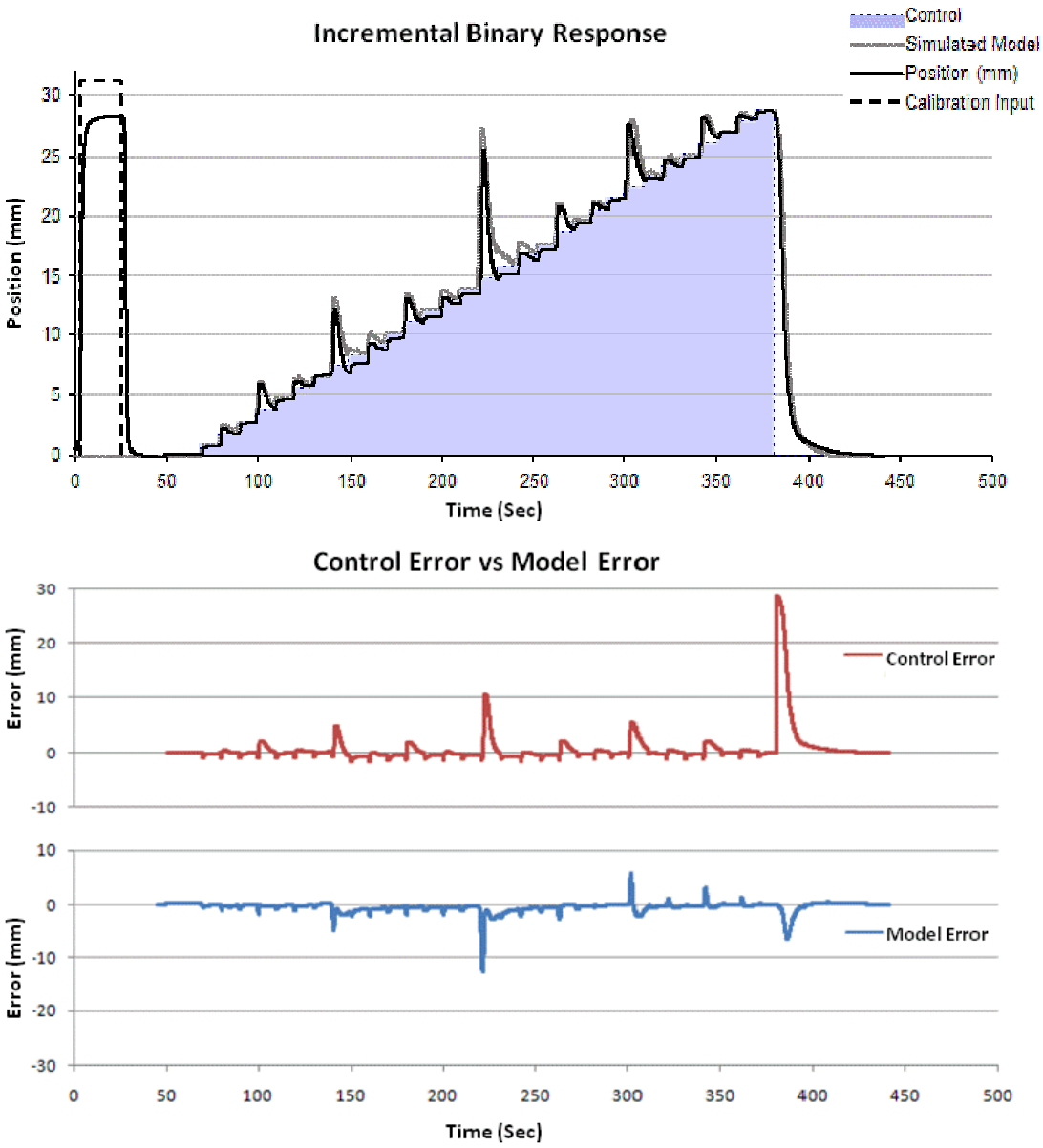


Figure 5.14 Comparison of SMA-BA model and experimental results for an incremental binary input. The calibration input is a step response of the maximum input (31). The resulting displacement is used to adjust the input values. The results show that the despite the bit-change overshoots, the actual response is consistently close to the control input and the simulated model. However, relatively large error peaks are evident in both control error and model error (bottom two graphs). In the control errors, these peaks are due to the bit-change overshoots. In the model errors however, the peaks are more related to the delay between predicted and actual behaviours.

5.3.4.3 Comparing SMA-BA non-cooled with active-cooled response

The incremental binary response established that the SMA-BA is predictable, reliable and controllable. The simulated model differed from the actual response with a mean error of ±2.48% of the full-scale (28.3mm). A significant limitation of the actuator however, is evidently the large bit-change overshoots that can introduce major errors in a closed-loop control system. Therefore, active-cooling was introduced to improve the system response.

Active cooling based on forced-air convection was implemented using mini-fans. The solution was effective in appreciably reducing the bit-change artefacts (see Figure 5.15). Mean difference (error) between command and actual position was reduced from ±4.23% to ±2.74% (of full-scale). Minor irregularities visibly remained due to imperfections* in the cooling provided.

Figure 5.16 overlays the displacement output of the non-cooled and active-cooled systems in response to the same arbitrary control input. The active-cooled result shows considerable improvement in respect to the non-cooled results with average errors of ±0.93% and ±9.92% respectively (% error based on a full scale of 28.3mm).

* Cooling was achieved by placing two mini-fans adjacent to the wires with a single fan speed. There was no temperature feedback and no fan speed control. Thus, the air flow above the wires could not be fully controlled.

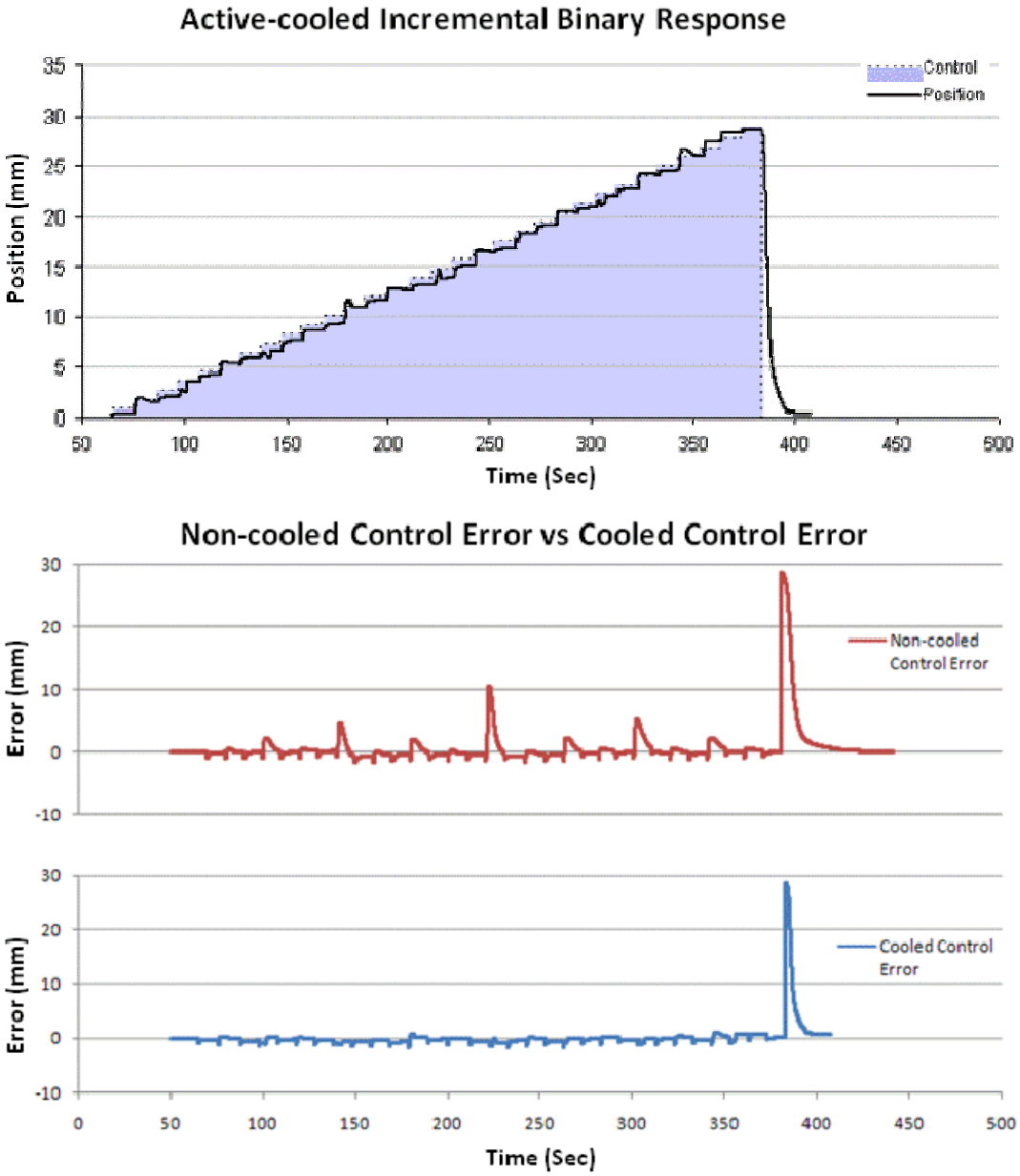


Figure 5.15 A repeat of the incremental binary response under active cooling. *While some irregularities remain in the stepwise response, the overshoots previously observed in the non-cooled system are now reduced. This reduction is demonstrated by the comparison of the non-cooled control error (as seen on figure 5.14) with the cooled controlled error (bottom two graphs). The final peak of the cooled control error is attributed to the final step from maximum position to zero. While the control signal changes instantly, the cooling and subsequent relaxation of all bits takes some time, even with active cooling.*

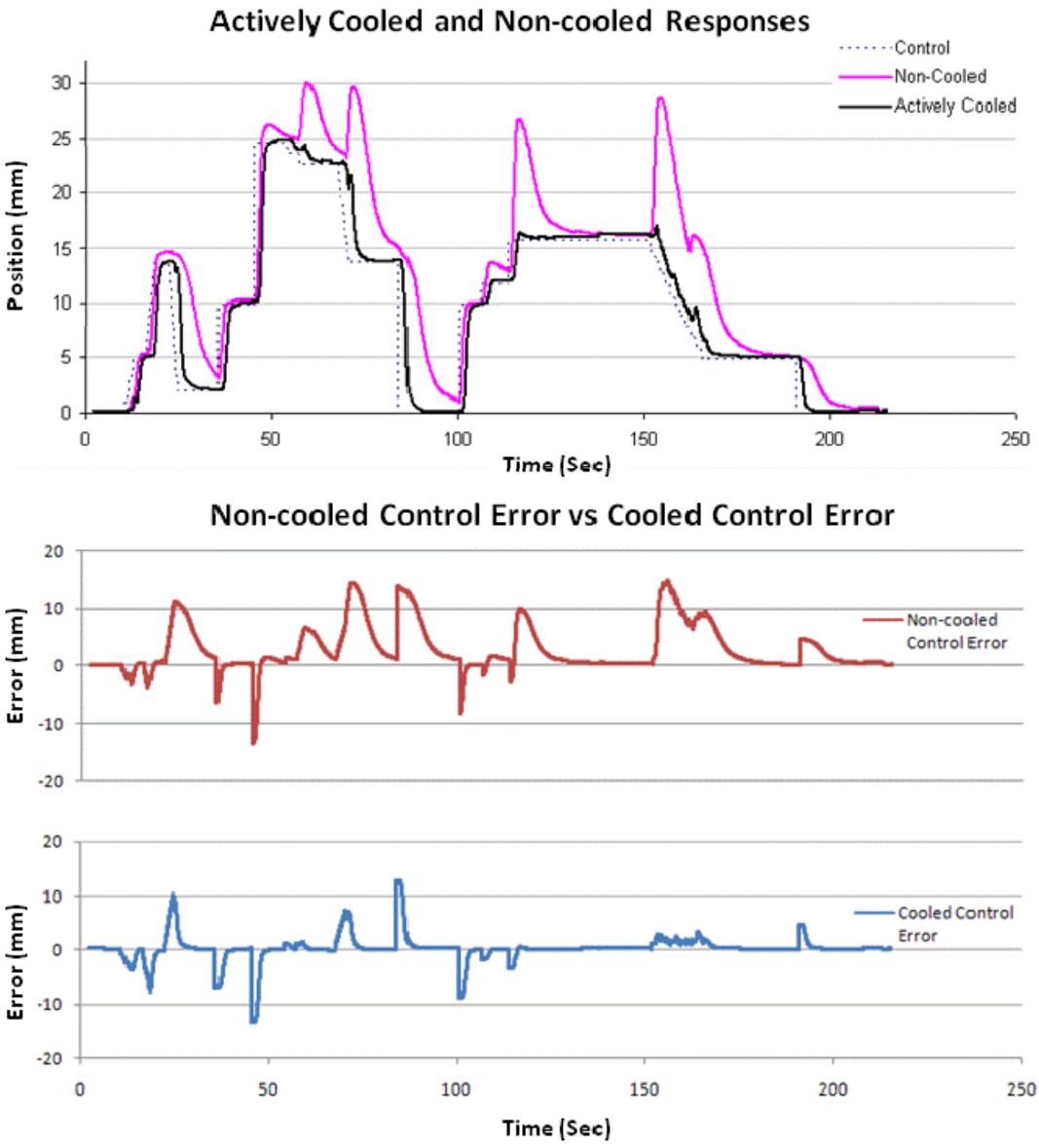


Figure 5.16 A comparison of the SMA-BA active-cooled and non-cooled systems using an identical arbitrary control input. *While the peak errors in both system are comparable (see bottom two graphs), the active-cooled system's average deviation from desired position was one order of magnitude better than that of the non-cooled system.*

5.3.4.4 Displacement vs. Binary Command response

To test the linearity of the SMA-BA system, the stable displacement responses were plotted against their respective binary commands. Each position value was obtained by entering the binary command and allowing the actuation response to fully settle (at least 10 seconds). The results in Figure 5.17 show a very linear response with an R^2 of 0.9988. A systematic variation is observed within these results with groups of readings appearing below the line of best fit and groups above. This is believed to be caused by slight variation in bit responses (e.g. contribution from bit 8 is slightly smaller than it should be) in addition to errors due to friction (causing some binary responses to be impeded prior to reaching predicted stable position).

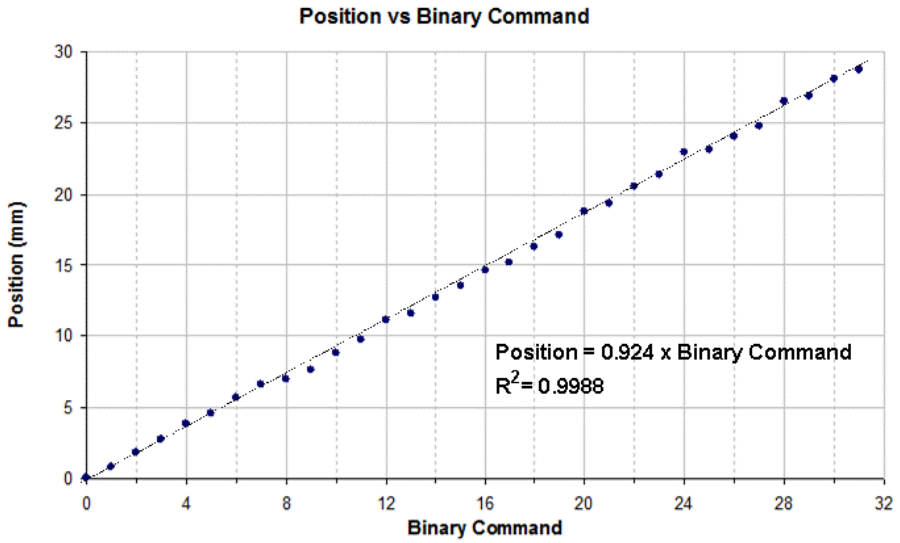


Figure 5.17 Stable position responses to binary commands from 0 to 31. The response is linear with an R^2 of 0.9988. The coefficient of the line of best fit was 0.924 meaning that the average strain of the wires for the applied load and ambient conditions was 4.62%.

5.3.5 Artificial Muscle Setup (active-cooled SMA-BA with FPT sensor)

In the final series of tests, the SMA-BA actuator was combined with the FPT sensor described in Chapter 3 to implement a closed-loop control system. Despite the linearity of the open-loop position response, closed loop control is more effective.

This is due to the fact that the response to binary command can vary with changing load or environmental conditions (e.g. temperature and humidity). Two modes of control were tested separately; position control and force control.

5.3.5.1 Position control using SMA-BA/FPT system

Position control was achieved with a 400g (3.92N load) hanging mass (see Figure 5.5a). An arbitrary position command was entered into the control system and the tracking response was recorded. Figure 5.18 shows the binary command changing as a function of the input command and the position feedback from the FPT. The resulting displacement of the actuator tracked the desired command closely with some error. An average lag of 2 ± 0.1sec and average position error of ±1mm (±3.2% full-scale) were measured (although peak errors of up to 6mm can be observed).

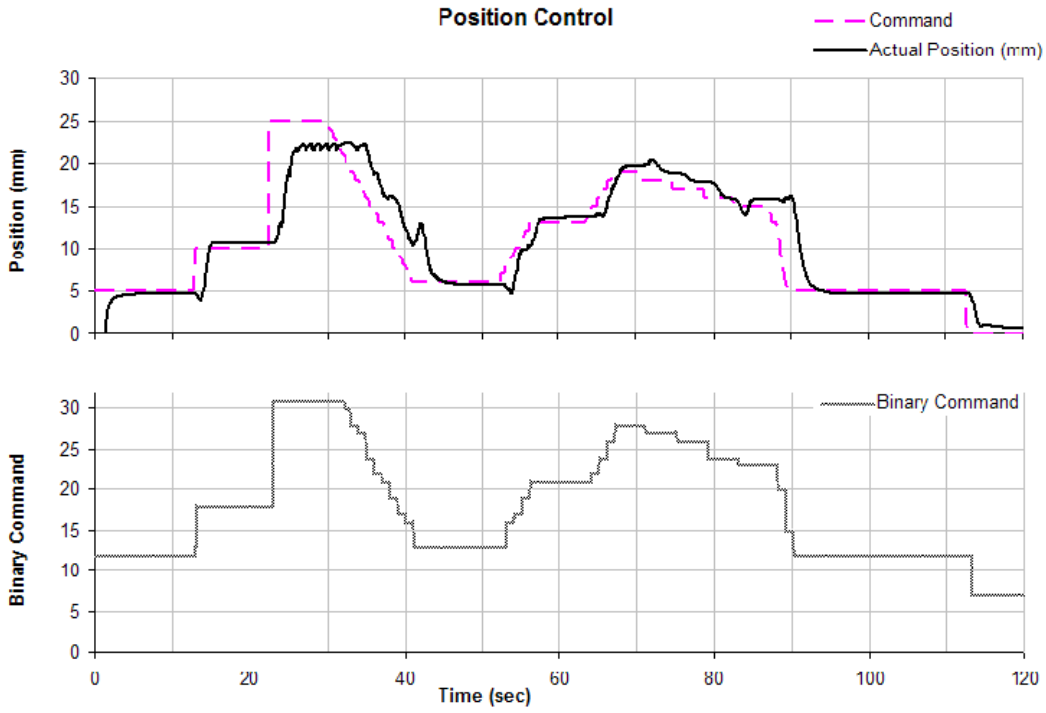


Figure 5.18 Isotonic position control implemented using the SMA-BA as the actuator and the FPT as feedback sensor. The top graph shows the tracking results measured by the LVDT position sensor. The bottom graph shows the changing binary input as a function of the command signal and the position feedback.

5.3.5.2 Force control using SMA-BA/FPT system

On average, the actual forces achieved were within ±0.29N (±6% full-scale 4.9N) of the target with a mean lag of 2 ± 0.1sec (see Figure 5.5b). However, peak force errors exceeding 0.69N were measured. The results are depicted in Figure 5.19.

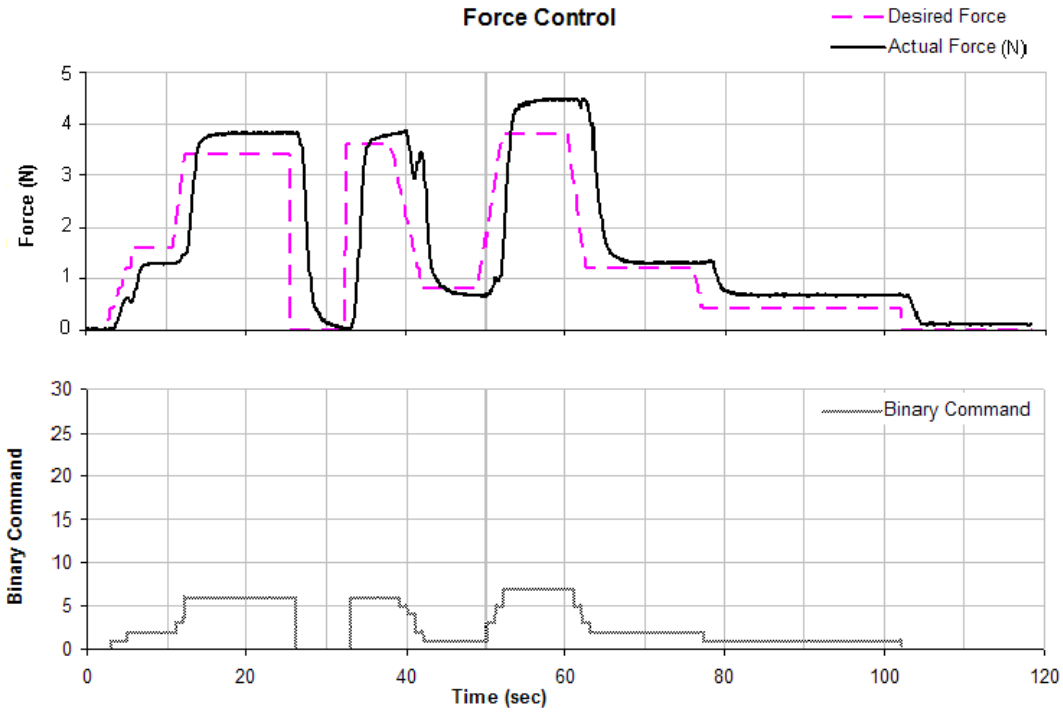


Figure 5.19 Isometric force control implemented using the SMA-BA as the actuator and the FPT as feedback sensor. *The top graph shows the tracking results. The bottom graph shows the changing binary input as a function of the command signal and the force feedback*

5.4 Discussion

The results of the 1-W, 4WB, and 4WP tests showed that the pulley system is a viable manifestation of the work-summation concept, successfully satisfying its objectives. The SMA Binary Actuator took the concept further showing that in addition to work-summation, the pulley method can improve control of SMA based actuation. The justifications for these conclusions are outlined in this section.

5.4.1 Variability of wire lengths between setups

Despite efforts to prepare each of the three setups (1-W, 4WB and 4WP) with identical length SMA wires, some variations eventuated. In order to compensate for these differences when comparing the results of alternate setups, strain percentages were considered in place of absolute displacements. Within each setup (WB and 4WP) however, maximum differences were able to be maintained less than ±0.6% of the total lengths.

5.4.2 Establishment of work output boundaries using the 1-W results

The single wire test was used to establish the practical and theoretical boundary of multi-wire work capacities. Table 5.1 presents the rated characteristics of the wire. Experimentally however, the strain of the material differs at different loads.

5.4.2.1 Single wire (1-W) actuation and scaled projection (1Wx4)

The 1-W curve in Figure 5.7 shows a low strain (less than 1%) for lighter preloads. This relates to the force required to stretch (deform) the SMA wire prior to being activated again. It was noted that 0.74N (75g) and 0.98N (100g) loads were not enough to stretch the wire significantly (as is evident in Figure 5.7). Nitinol can normally be stretched by a maximum of 8% of its austenite size[*]. As preload was increased, there was a gradual rise in strain. At 4.9N (500g), the wire failed (by breaking at the connection to the load cable).

The graph in Figure 5.7 shows that a four-fold scaling of a single wire actuator would allow actuation at four times the load range. The projected 1Wx4 curve was compared with the practical scaled actuators 4WB, the four wire bundle, and 4WP, the four wire pulley system (Figure 5.10). A detailed comparison of these alternate arrangements is found in section 5.4.5 (p. 111)

5.4.3 Bundling as a method for work-summation

The 4WB test represented a simplistic method used for scaling actuator work capacity, namely bundling (Mavroidis, Pfeiffer et al. 2002; Spinks, Campbell et al. 2005). Bundling involves placing parallel units of actuator material together and connecting the ends of all units to a single output. The problem with this method is that any slight variations in length of the individual units in parallel will be significant in relation to the overall shortening of the units (for example, a 50mm polypyrrole

[*] Although at 8% deformation, Nitinol can only actuate a few times before failing (Gilbertson, 2000).

fibre contracts no more than 1mm when activated). Some units in the bundle will be loose and others tight. Following activation, the loose units will tend to work mainly to reduce their slack while the tighter units will have to do the majority of the work (see Figure 5.20).

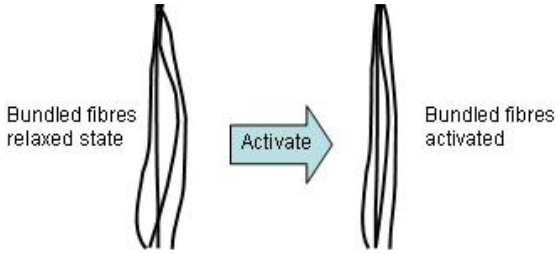


Figure 5.20 Activation of bundled actuators. When simply bundled together in this manner, the actuator forces do not sum effectively. The looser fibres tighten their slack while the tighter fibres have to carry the load.

5.4.3.1 Individual SMA wire contributions and tensions

Prior to the experiment, the bundle assembly was prepared with identically sized SMA wires (average difference in length were ±0.6% of total length). This was to ensure equal distribution of the overall load among the four wires. The combined-activation force graph (Figure 5.8, lower-left) shows relatively even forces prior to activation of the wires. However, upon powering, individual forces are seen to alter drastically. Wire-1 carries virtually no load during activation, while Wire-2 contributes less than 0.49N of force to the overall actuation. Wire-3 is seen to undertake the majority of the total load (around 1.77N) with Wire-4 carrying around 1.37N of force. Force graphs during sequential activation demonstrate much greater force variations. When activated individually, wires were seen to carry the entire load by causing other units to become slack. For a 3.43N load, sequential activation caused wires to reach the rated limit. Hence, beyond this load, sequential activation was not considered sustainable.

5.4.3.2 4WB Sustainable and maximum work limits

The 4WB actuator was tested with the series of hanging masses specified in the methods section. At 11.77N (1200g), the actuator failed following three cycles (wire-3 broke). While 11.77N was considered to be the maximum load for this bundle, a sustainable load limit was significantly lower, determined to be 5.88N (600g). In comparison, the sustainable limit for a single wire was designated to be 350g (based on manufacturer recommendations). It follows that a four-fold scaled actuator would ideally be rated at 13.73N (4x3.43N). The 4WB actuator failed at far below the ideal sustainable limit. More significantly, the actuator was able to work sustainably at only 43% of the ideal load. In the following sections, the work capacity of the bundle system will be further scrutinised and compared with the 1-W, 1Wx4 and 4WP actuators.

The sustainable load limit for the 4WB setup was determined to be at 5.88N (600g). This was determined by interpolating the load at which at least one of the wires would undergo a tension greater than 3.43N (see Figure 5.11 top graphs). Figure 5.21 supports this calculation by demonstrating that at loads greater than 5.88N (in this case 7.85N) some of the wires exceed the 3.43N limit during activation.

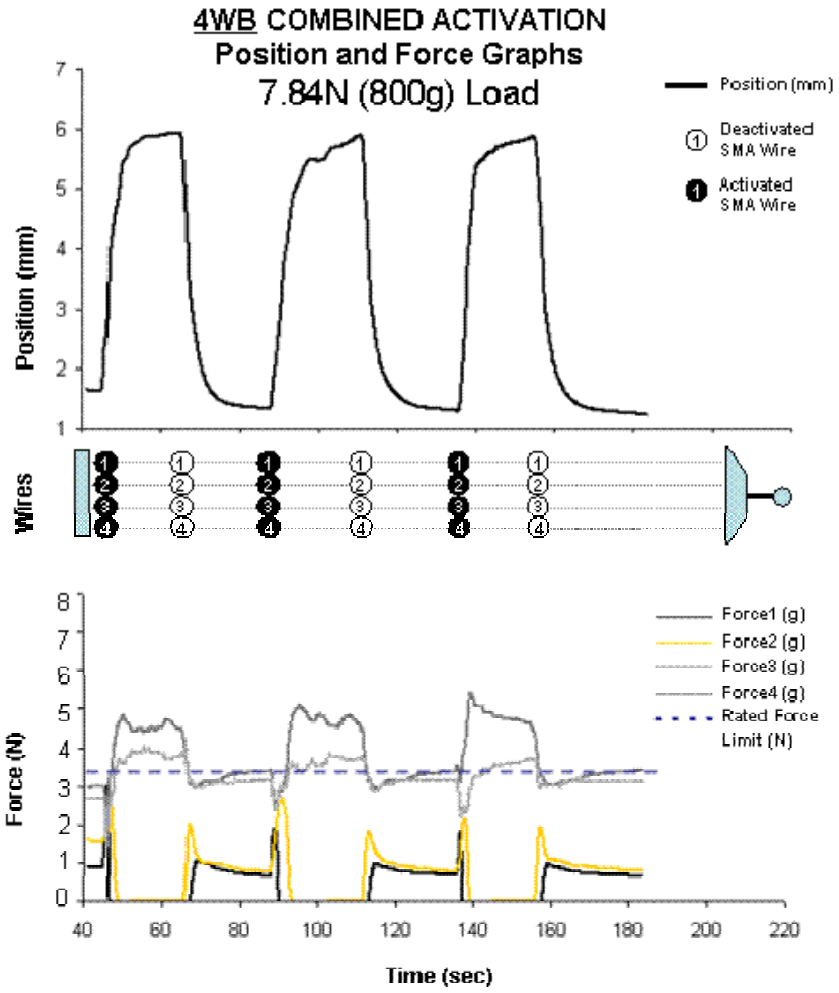


Figure 5.21 4WB Position and force graphs for 7.84N (800g) load.
Wires 3 and 4 reach forces of 3.92N and 4.90N.

5.4.4 Implementing work-summation using the pulley method

The 4WP actuator system was used to test the principles of the work-summation system proposed in this thesis. The principle aim of this design was to allow a scaling of actuation with minimal loss of work. However, an additional benefit of this method is evident in the results of the 4WP tests: the ability to actuate in discrete steps. This feature is further exploited in the Binary SMA Actuator to maximise resolution of displacement. As will be discussed, this is a significant and promising development for the control of SMA actuators. Moreover, the principles outlined also present opportunities for improvement of actuation for most types of Smart Materials.

5.4.4.1 Sequential activation

Using the 4WB setup, sequential activation produced an interesting result. The graph in Figure 5.9 shows step-wise displacement when the wires are powered sequentially. Sequential relaxation (de-activation) of the wires does not produce similarly shaped (sharp) steps as seen for activation. This is due to the slow cooling time of the wire in combination with slight friction in the pulleys. Sequential activation in pairs results in larger steps equivalent to double step sizes. This is consistent with expected behaviour.

5.4.4.2 Individual SMA wire contributions and tensions

While testing the 4WP actuator with the variety of prescribed loads, individual wire tensions were found to be relatively even (see force graphs of Figure 5.9 during both combined and sequential activation). The minor variations in forces observed are frictional, caused when wire moves within the Teflon pulleys. However, unlike the 4WB (see Figure 5.8) actuator, activation does not cause individual wire forces to vary significantly. This is particularly important for sequential activations. The implications are that an individual actuator unit can activate and contribute in work output independently without undergoing any additional stress.

5.4.4.3 4WP Sustainable and maximum work limits

The 4WP actuator failed while carrying a 15.69N (1600g) load. The maximum sustainable load however, was determined to be 11.77N (1200g, see section 5.3.3.3), 85% of the ideal load limit of 13.73N (obtained from the single wire characteristics scaled by four). This value was double the limit calculated for the 4WB actuator. 4WP position and force graphs for an 7.84N (800g) load are presented in Figure 5.22

for comparisons with equivalent tests using the 4WB actuator (Figure 5.21). In the 4WP graphs, individual wire tensions are shown to be well within the 3.43N (350g) limit for sustainable use (true for both combined and sequential activation). In the 4WB graphs, at least one of the wires exceeded forces of 4.90N.

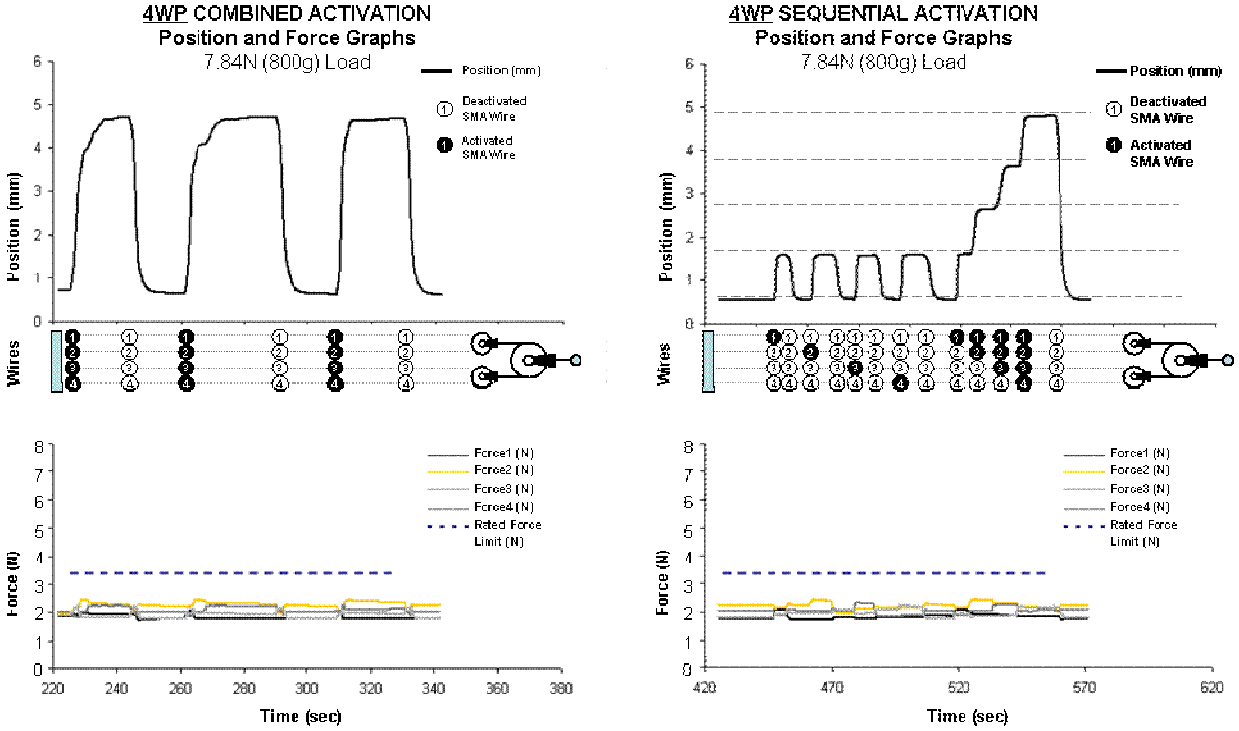


Figure 5.22 4WP Position and force graphs for 7.84N (800g) load. *Individual wire forces remain below the rated limit (3.43N) during all activations. In the sequential activation test (right) individual wires are shown to contribute equal displacement steps to the combined output.*

5.4.5 Comparisons of 1-W, 1Wx4, 4WB and 4WP

Figure 5.10 presents a comparison between the four-wire actuators (4WB and 4WP) and the four-fold scaled projection of the single wire (1Wx4). The graphs depict the strain outputs of each actuator type respectively, across the series of tested preloads. Evidently, neither 4WB, nor 4WP stretch across the load range as far as the single wire four-fold projection. Nevertheless, both 4WB and 4WP show the ability to actuate with lower loads than possible with 1Wx4. This lower load-range capability is

understood to be assisted by the uneven forces present within the wires in both 4W actuators. As some of the wires undergo slightly higher tensions than others, they are more likely to stretch. The steep slope at the lower loads of the 1-W curve shows that small variations in force cause large changes in strain. Within the 4WB and 4WP actuators, there are variations in the forces of each wire. Consequently, wires with higher tensions will produce larger strains and therefore contribute to a greater overall displacement.

5.4.5.1 Differences in wire tension distributions

It is clear from the results (see Figure 5.11) that the 4WP actuator is capable of maintaining even wire forces during activation as well as relaxation. Conversely, the bundle actuator is incapable of distributing the total load evenly. This shortcoming is more evident with larger loads. Furthermore, during activation, the individual forces tend to increase significantly in some of the 4WB wires. This phenomenon reduces the sustainable load limit for the 4WB method.

5.4.5.2 Work Zones as parameters for comparison

For another comparison of actuation capacity, the concept of 'work zone' is introduced here. For the purpose of analysis, a work zone refers to the collection of possible strain-load outputs where at each point, work = displacement x force. In the strain-load graph, work zones are represented as the areas under the curves. The work zones of interest are the areas up to and including the sustainable loads for each actuator type. These limits were determined to be 3.43N, 5.88N, 11.77N and 13.73N for the 1-W, 4WB, 4WP and 1Wx4 systems respectively. Figure 5.23 depicts the work zones areas of the compared systems within the sustainable limits. Ideally

the work zone area of a four-fold scaled system should be four times larger than the area of the single unit (as indicated by the 1-W and 1Wx4 areas). The results show that the 4WB area is significantly smaller than the ideal area. The 4WP area on the other hand appears comparable in size with the projection.

Although the 4WP sustainable load falls 1.96N short of the projection limit, the work zone of the former begins at much lighter loads. This feature effectively increases the total work zone area. Practically, this is a significant finding because the work zone of an actuator system can be easily shifted on the load axis using counter weights and springs[*]. The work zone area on the other hand, can only be increased by the contribution of additional actuation units. Equally, the work zone area can decrease as a result of inefficiencies leading to loss of energy.

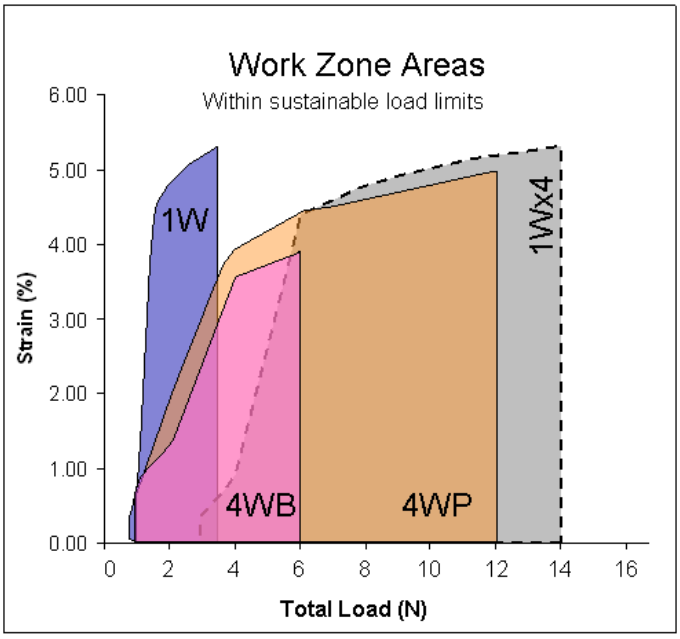


Figure 5.23 A comparison of work zone areas. Areas represent single wire (1-W), four wire bundle (4WB), four wire pulley (4WP) and single wire projection (1Wx4). The 4WB area is significantly smaller than the ideal area of 1Wx4. The 4WP area is comparable with that of the projection.

[*] A shift to the right on the load axis requires an agonist bias, whereas a left shift could be achieved with an antagonist bias.

A more conclusive comparison can be achieved by considering the absolute magnitudes of the work areas in question. These values were obtained by integrating the strain vs force curves for each actuator. The results are shown in Figure 5.24. For each actuator type, area size was calculated for total work maximum as well as total work rated.

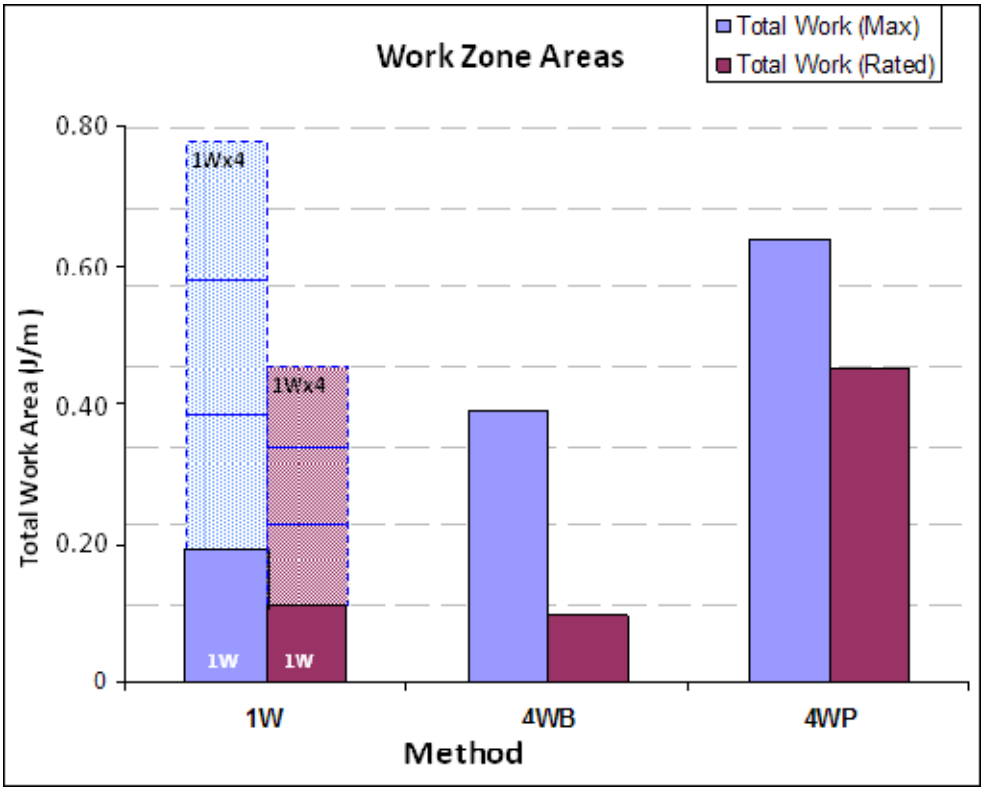


Figure 5.24 Magnitude comparison of Work zone areas. Absolute work zone areas were calculated for total work and rated work (i.e. work within sustainable load limits).

'Total Work Maximum' values refer to the total areas under the curves up to the points at which each actuator failed. Although these are interesting parameters, they do not have a great practical significance and are unreliable measurements*. 'Total

* This value is considered unreliable because it is highly dependent on the value of the load at which the actuator failed. Actuator failures are increasingly more likely as the load is increased. Nevertheless, they are unpredictable and random.

Work (Rated)' values are areas under each curve up to and including the determined sustainable limits. These values have the most significant practical implications.

Figure 5.24 shows that neither of the 4W actuators is capable of reaching the 1Wx4 projection in maximum work area (though 4WP performed significantly better than 4WB). However, within the more interesting 'rated' zones, 4WP matches the scaled projection. 4WB on the other hand exhibits a rated work potential even lower than the single wire.

5.4.5.3 Conclusions from the comparisons

The work-summation system proposed in this thesis was implemented using a simple pulley system. When tested against a theoretical benchmark, and compared with a more traditional bundling method, the pulley system showed;

- minimal loss of energy and higher efficiency than the bundle method
- significantly higher sustainable load limits than the bundle method
- similar total work zone area to that of the theoretical benchmark
- the ability to actuate in discrete steps, introducing a new method of control not possible with bundling

The main disadvantages of the pulley system are;

- Increase in the complexity and size of the total system. This can be minimised by miniaturisation of the pulley network. The control system can easily be integrated and miniaturised using chips

- Loss of some energy due to the friction* inherent in the pulleys. This may be minimised by advanced manufacturing techniques and the selection of improved material. It is possible that enclosed lubricated cable-tube systems or reliable miniaturised cable-wheel systems† could reduce friction further.

5.4.6 Binary Actuation

Figure 5.4 in the methods section described the design of the SMA Binary Actuator (SMA-BA). This novel design was implemented to exploit the step-wise behaviour of the work-summation system to permit segmental displacement. The 4WP actuator was shown to be operable in four equal steps. Hence, its position control had a resolution of [maximum strain] ÷ 4. If the 4WP could be re-arranged however, so that each wire could contribute twice the strain of its preceding unit, the resolution would be increased to [maximum strain] ÷15‡. Applying the same concept to a five wire actuator would increase the resolution to [maximum strain] ÷ 31, six wires to [maximum strain] ÷ 63, and so on.

The 5-bit SMA-BA was designed with the following aims:

- Scale-up SMA actuation capacity to produce displacements up to 30mm at a force output of 6.86N

---

*While friction was observed to contribute to the limitation of the pulley system, a full study of the effects of friction and research into material improvement was not within the scope of this project.

† It would be important to design a cable-wheel system with minimal possibility for the cable to fall off the wheel.

‡Number of steps = 1 + 2 + 4 + 8

- Allow open-loop control of displacement with a resolution of 1mm
- Allow closed-loop force and position control in conjunction with the FPT sensor (presented in Chapter 3).

5.4.6.1 Binary bit test – displacement contribution

The results of the binary bit test (see
Figure 5.12) showed that the displacement contribution from each bit was consistently proportional to the bit value. However, the absolute displacements were lower than the bit values. This was because when designing the actuator, the bit values were designated with the assumption that each wire would produce a strain of 5%. The actual strain from each bit however was closer to 3.2% (due to the lightness of the load applied).

5.4.6.2 Elimination of the bit-change artefact

The comparison of the cooled and non-cooled incremental binary responses showed that the bit-change artefact can be predicted (see Figure 5.14) and eliminated (see Figure 5.15). The cooling method used in these experiments was rudimentary. Nonetheless it was effective in significantly reducing the bit-change effects. Improvement of the cooling system to better match the heating and cooling rates of the SMA wires will further reduce the remaining irregularities (as seen in Figure 5.15). The detailed study of controlled cooling of the system was not within the scope of this study and is suggested for future investigation. Potential cooling mechanisms may include:

- Controlled air convection cooling with minimised turbulence in air flow. The fan speed should be controllable and temperature feedback should be incorporated in the control system.

- The SMAs may be placed in liquid coolants to allow faster dissipation of heat. However, this may also reduce efficiency due to loss of input energy.
- Peltier type active cooling elements may be incorporated in the control of temperature. Again, temperature feedback would be necessary for a closed loop control system.
- A combination of the above may be useful and effective and worth investigating in future work.

5.4.6.3 Closed-loop control of actuation force and position

Given that the SMA-BA has a nominal resolution of 1mm, the tracking results in Figure 5.18 are a compelling demonstration of the system's ability to provide closed-loop position control. Further improvement of the actuation resolution and speed and accuracy of the FPT feedback is necessary increase the accuracy of the control system.

Force control results were less accurate than the position control results (see Figure 5.19). The error was in the context of a force-output resolution of approximately 0.6N. This resolution relates to the position resolution of 1mm. The FPT spring has a spring coefficient of 0.61N/mm. Therefore, under isometric conditions (as illustrated in Figure 5.5B) every millimetre of actuator movement is equivalent to 0.61N force change.

5.4.7 Significance of results

The results show that the pulley-based actuation system is capable of implementing the work-summation system proposed in this chapter. The advantages of such a concept are evident:

- The work-summation system is an effective method for scaling up Smart Materials by efficiently combining multiple units. It allows the overall load to be distributed equally among the units and in doing so, minimises each unit's stress. At reduced stresses the cycle life of the materials are significantly lengthened. Consequently, a higher overall load may be sustainably actuated than is possible with bundling.
- The method allows multi-bit, stepwise actuation. This is of particular interest for control of SMAs as they are most suited to on/off type activations. Binary control of SMAs is much less complicated than linear control, pulse width modulation and non-linear control (see Introduction). In effect, the Binary-SMA functions similarly to the segmental controls proposed by Selden (2004) , Cho (2006) and Lee (2006). Selden and Cho used Peltier chips along the path of a single SMA wire. This method does not allow simultaneous scaling of multiple actuators and is highly inefficient in power consumption while adding excessive weight/volume to the actuator. Lee implemented segmental control of a single SMA actuator by introducing electrode taps along the wire. This method requires connecting several electrodes to a moving SMA wire and does not allow scaling. The 5-bit SMA-BA presented in this thesis actuated with a displacement of 30mm at up to 5.88N force. Using the protocols of the above authors, a single wire with equivalent capabilities would have to be 1.2 meters long (60ohms requiring 24 volts to heat up). However such a wire would not be able to actuate proportionally without the use of a

sophisticated control system involving Pulse Width Modulation (PWM), temperature control or current control.

Following Honma's early work* (1989), many groups became interested in using SMAs as artificial muscles. By the mid 1990's however, much of this interest declined due to perceived inherent inefficiencies of electrical heating, low speeds and difficulties in controlling the material. The work presented in this thesis contests this outlook on SMAs and aims to bring this material back into high relevance. Of the Smart Materials considered for actuation, SMA has one of the highest energy densities (see Appendix E). By exploiting this abundant property, efficiency, speed and controllability can be improved. The potential exists for further development of the ideas presented to create an integrated force/position controllable artificial muscle. Such a device may use SMAs (or other suitable Smart Materials) as active elements, the work summation method as its inner mechanism and the FPT for its feedback.

The proposed method offers solutions to longstanding challenges in scaling, improving and controlling Smart Materials. In particular relevance to SMAs, it is well timed to coincide with new developments in fuel-based actuators (Baughman 2005; Ebron, Yang et al. 2006; Jun, Rediniotis et al. 2007). Fuel driven SMA actuators are not just more energy efficient, but they are also faster. In the foreseeable future, a large number of small, thin, SMA units may be integrated using work-summation principles and incorporated into a fuel based system. Such an actuator would feature fast, high resolution actuation due to the large number of controllable bits. The increase of the resolution in conjunction with the decrease in cooling time will act to

* Honma showed that SMA actuators can be controlled by electrical heating

significantly reduce the undesired artefacts. In addition the overall work capacity will be brought closer to those of natural muscles.

5.4.7.1 Work in progress

Challenges ahead include the miniaturisation of the pulley network components and the reduction of friction to increase overall efficiency.

An alternative to the pulley-based method is the lever-based work-summation system (described in Appendix B). This method promises easier manufacture and miniaturisation. Levers are simpler components than pulleys and do not require cables to work. The lever method may provide benefits such as reduction of complexity, reduction of number of components and the ability to amplify force or movement directly by modifying lever length ratios (see Appendix B),

Currently a miniaturised, 32-unit lever-based system is under development for use with conducting polymers (see Figure 5.25).

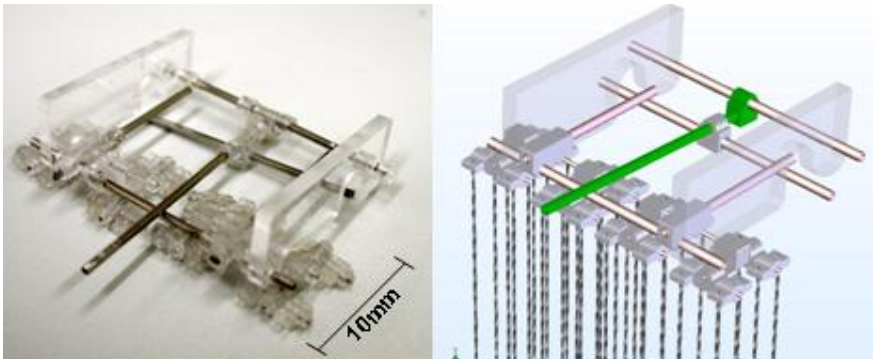


Figure 5.25 32 unit work summation unit based on the lever method. Right: 3-D design drawings. Left: Actual construction using tiny acrylic levers.

5.5 Concluding Remarks

The development of viable and economical muscle-like actuators for rehabilitation and enhancement of the body will be highly beneficial. Researchers around the world are working steadily toward the creation of artificial muscles. This collective work spanning several decades has tackled the problem from many angles. Consequently, the range of materials and methods under development is considerable. Many Smart Materials possess actuation potentials superseding equivalent properties in muscles. However, these promising potentials are typically observed only in laboratory settings within micro-level scales.

This chapter has proposed and evaluated a method for the summation of multiple, discrete work outputs. The methods described offer solutions to the problem of scaling-up Smart Material actuators. Instead of growing Smart Materials larger, many small-unit-actuators should be used in parallel and combined using the work-summation system. The results of this study have demonstrated the effectiveness of the hypothesis using a limited number of SMA actuators. Future studies will focus on larger numbers of actuator units and miniaturised components for the supporting structures. Such an endeavour will be a step closer to the long-awaited arrival of artificial muscles.