Abstract

STEVENS, JASON MICHAEL. Actuation and Control Strategies for Miniature Surgical Robotic Systems. (Under the guidance of Dr. Gregory Buckner.)

Over the past 20 years, tremendous advancements have been made in the fields of minimally invasive surgery (MIS) and minimally invasive robotic assisted (MIRA) surgery. Benefits from MIS include reduced pain and trauma, reduced risks of post-operative complications, shorter recovery times, and more aesthetically pleasing results. MIRA approaches have extended the capabilities of MIS by introducing three-dimensional vision, eliminating tremors, and enabling the precise articulation of smaller instruments. These advancements come with their own drawbacks, however. Robotic systems used in MIRA procedures are large, costly, and do not offer the miniaturized articulation necessary to facilitate tremendous advancements in MIS.

This research tests the hypothesis that miniature actuation can overcome some of the limitations of current robotic systems by demonstrating accurate, repeatable control of a small end-effector. A 10X model of a two link surgical manipulator is developed, using antagonistic shape memory alloy (SMA) wires as actuators, to simulate motions of a surgical end-effector. Artificial neural networks (ANNs) are used in conjunction with real-time visual feedback to “learn” the inverse system dynamics and control the manipulator endpoint trajectory. Experimental results are presented for indirect, on-line learning and control. Manipulator tip trajectories are shown to be accurate and repeatable to within 0.5 mm. These results confirm that SMAs can be effective actuators for

miniature surgical robotic systems, and that intelligent control can be used to accurately control the trajectory of these systems.

Actuation and Control Strategies for Miniature Surgical Robotic Systems

by

Jason Michael Stevens

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of

Master of Science

Mechanical and Aerospace Engineering
Raleigh, North Carolina
July, 2002

APPROVED BY:

Dr. Edward Grant
(Advisory Committee Member)

Dr. M. K. Ramasubramanian
(Advisory Committee Member)

Dr. Gregory D. Buckner
(Chair of Graduate Advisory Committee)

Dedication

To God, who through His creation, inspires us to create.

Biography

I, Jason Michael Stevens, was born in Mt. Olive, North Carolina in 1980. My brother Matthew and I were raised by my parents David and Janie Stevens in rural Wayne county. I graduated from North Duplin High School in May of 1998. While at North Duplin I met my future wife and best friend Ashley Kelly. We will be married on August 3, 2002.

I never thought of becoming an engineer until the summer before I entered college. I had always planned on entering into an agricultural profession, but through the encouragement of my family and Doug Kornegay, an electrician I worked for that summer, I decided to enter engineering. I attended North Carolina State University, where I received my bachelor's degree in mechanical engineering in May of 2001. I began graduate study at N.C. State the summer of 2001. After completion of my master's degree, I plan to continue my education at N.C. State, pursuing a Ph.D. in mechanical engineering. I will continue conducting research in the field of medical instrumentation.

I have been interested in medical instrumentation since I began in engineering. I feel that God has given each person certain gifts in order that they may use those gifts to impact the lives of other people. I think that some of my gifts lie in the area of developing instrumentation and machinery that will directly improve the quality of life for people. I hope that someday, someone will live a happier, healthier life because of some aspect of one my research projects. At that point I could consider my work a success.

Acknowledgements

I would like to thank my wife to be, Ashley Kelly. I am truly blessed to have someone like her to offer tremendous love and support in all that I do. Thanks for putting up with a nerd like me.

I would like to thank my parents for the guidance, training, and encouragement they have given me through the years. Thank you to my mother for giving me all the hands-on science lessons as a small child and sparking my interest in scientific endeavors. Thank you to my father for teaching me hard work and patience.

I would like to thank Dr. Randolph Chitwood and Dr. Wiley Nifong of East Carolina University's Brody School of Medicine. Thank you for allowing me to get involved in this exciting area of research, and thanks for your attitude, enthusiasm, and excitement. Also, thank you to the Brody School of Medicine and the Doris Duke Foundation for providing the financial support to conduct this research.

Thank you to Robert Hughes and Keith Sanders for their assistance in my research. Thanks for your long hours, hard work, and most of all your friendship. The Mountain Dews were of great assistance too.

Finally, I would like to thank Dr. Gregory Buckner for all of his training, instruction, support, and friendship throughout my graduate study. It has been a pleasure to work for you.

Table of Contents

List of Tables.....	vii
List of Figures	viii
1. Introduction	1
1.1 Cardio-thoracic surgery: methods	1
1.2 Advancements in surgical techniques	2
1.3 Drawbacks of MIS techniques	3
1.4 Purpose.....	4
2. Background on MIRA Surgery	5
2.1 Advancements	5
2.2 Drawbacks of tele-operated systems	7
3. Miniature Actuation for Medical Applications	10
3.1 Candidate actuation technologies.....	10
4. SMA Background.....	13
4.1 SMA phase transformations.....	13
4.2 Stress- strain characteristics	15
4.3 Energy characteristics	16
4.4 Other considerations.....	19
5. Serial Manipulator Test Rig.....	20
5.1 Conceptual model.....	20
5.2 Experimental scale model	21
6. Manipulator Modeling and Kinematics.....	25
6.1 Modeling	25
6.2 Inverse kinematics.....	26
7. Manipulator Control: Process Overview.....	28
7.1 Hardware-software overview	28
8. Manipulator Control: Algorithm Development.....	32
8.1 Development of a linear SMA test rig	32
8.2 Controller development using the linear SMA test rig	35
8.3 Proportional Control with Varying Sample Rates.....	36
8.4 Intelligent Control	39
9. Manipulator Control: Experimental Evaluations	45
9.1 Training the neural network	45
10. Conclusion.....	56
11. Improvements and topics for further research.....	57
11.1 Improvements.....	57
11.2 Further Research	57
12. References	59

List of Tables

Table 1: Typical properties of candidate miniature actuators [[12], [16]]	11
Table 2: Experimental parameters for intelligent control of linear SMA test rig	43

List of Figures

Figure 1: The da Vinci surgical robot from Intuitive Surgical [15]	2
Figure 2: Laparoscopic surgical tools [11].....	6
Figure 3: Surgical console for Intuitive's da Vinci system [11]	7
Figure 4: Suturing with Intuitive's da Vinci robotic system [11]	8
Figure 5: Typical O.R. layout with robotic surgical system and console [11].....	9
Figure 6: SMA phase configurations [14].....	14
Figure 7: SMA phase transformation/actuation [14].....	14
Figure 8: Quasi-plastic stress/strain of SMAs	15
Figure 9: Pseudo-elastic stress/strain of SMAs	16
Figure 10: Helmholtz free energy diagram - low temperature state	17
Figure 11: Helmholtz free energy diagram - medium temperature state	17
Figure 12: Helmholtz free energy diagram - high temperature state	18
Figure 13: Gibbs free energy	19
Figure 14: Solid model of a conceptual, semi-autonomous, micro-robotic suturing device	21
Figure 15: CAD solid model of the 10X serial manipulator test rig	22
Figure 16: Photograph of the 10X serial manipulator	23
Figure 17: Photograph of the 10X serial manipulator and frame.....	23
Figure 18: Photograph of the complete 10X serial manipulator test rig	24
Figure 19: Manipulator schematic.....	25
Figure 20: Manipulator schematic with auxiliary variables.....	26
Figure 21: Hardware schematic.....	29
Figure 22: Electrical schematic of the current-amplifying transistor circuit	30
Figure 23: Flowchart of real-time image processing and control software.....	31
Figure 24: Functional schematic of the linear SMA test rig	33
Figure 25: Photograph of the linear SMA test rig	34
Figure 26: Open-loop step response of the linear SMA test rig	35
Figure 27: Tracking response of the linear SMA test rig: proportional control at 1,000 Hz sample rate.....	37
Figure 28: Tracking response of the linear SMA test rig: proportional control at 100 Hz sample rate.....	37
Figure 29: Tracking response of the linear SMA test rig: proportional control at 25 Hz sample rate.....	38
Figure 30: Tracking response of the linear SMA test rig: proportional control at 10 Hz sample rate.....	38
Figure 31: Tracking response of the linear SMA test rig: proportional control at 1 Hz sample rate.....	39
Figure 32: Functional schematic of the intelligent control system	40
Figure 33: Tracking response of the linear SMA test rig: untrained intelligent control at 1 Hz sample rate	42
Figure 34: Tracking response of the linear SMA test rig: trained intelligent control at 1 Hz sample rate	42
Figure 35: Tracking error cost function during on-line training: 0-1200 sec.....	43

Figure 36: Tracking response of the linear SMA test rig: trained intelligent control at 1 Hz sample rate	44
Figure 37: Tracking response of the linear SMA test rig: trained intelligent control vs. proportional control	44
Figure 38: Random input/output training data used for off-line training	46
Figure 39: Off-line training data for error cost functions	48
Figure 40: Tracking response of the manipulator tip: on-line training and control, with off-line pre-training	49
Figure 41: Error cost function for Link 1: on-line training and control, with off-line pre- training	50
Figure 42: Error cost function for Link 2: on-line training and control, with off-line pre- training	50
Figure 43: Weight vector w_1 : on-line training and control, with off-line pre-training	51
Figure 44: Weight vector w_2 : on-line training and control, with off-line pre-training	51
Figure 45: Tracking error cost function for Link 1 during on-line training: 0-800 sec	53
Figure 46: Tracking error cost function for Link 2 during on-line training: 0-800 sec	53
Figure 47: Weight vector w_1 : on-line training and control, with initial weights set to zero	54
Figure 48: Weight vector w_2 : on-line training and control, with initial weights set to zero	54
Figure 49: Tracking response of the manipulator tip: intelligent controller trained on-line	55

1. Introduction

Cardiovascular disease (CVD) is perhaps the most important medical issue in America today. CVD afflicts over 60 million people in the U.S. and is the leading cause of death, claiming more lives (2,600 per day) than the next six leading causes combined. Direct and indirect costs resulting from CVD are estimated to be \$329.2 billion per year [1]. One of the most common medical procedures associated with CVD is open-heart surgery, a technique used for a variety of procedures including coronary artery bypass grafts (CABGs), heart transplants, valve repairs, and valve replacements. In 1999 there were a total of 753,000 open-heart procedures in the U.S. [2].

1.1 Cardio-thoracic surgery: methods

The first attempts at open-heart surgery date back to World War II, when the invention of an effective heart-lung machine made this surgical innovation possible [20]. Open-heart surgery requires splitting and retracting the breastbone (a sternotomy), and enables direct visualization of the working area and manual manipulation of organs and instruments. Surgeons use their own hands in very natural and instinctive manner to sew, cut, and perform various procedures. The primary disadvantages of open-heart surgery are increased trauma and longer recovery periods. Most problems associated with this procedure stem from post-surgical complications and infections, not the procedure itself. The large incisions are painful, provide greater opportunity for infection, require longer recovery times for the patient, and are aesthetically displeasing. Another disadvantage of open-heart surgery is the surgeon's decreased dexterity in restricted spaces. Some

procedures require precise manipulation of tissue and sutures, and the human hand was not designed to effectively manipulate objects at this scale [4].

1.2 Advancements in surgical techniques

Recent advancements in the field have attempted to address the disadvantages of traditional open-heart surgery. The first series of advancements combined laparoscopic instruments with familiar surgical techniques, resulting in smaller incisions (partial sternotomies, transverse sternotomies, and mini-thoracotomies). Engineering advancements in port access technology allowed for even smaller incisions. The next major advancement was the development of video-scopic instrumentation: endoscopes that provided accurate visual feedback and served to enhance the surgeon's capabilities. Next, robotic manipulators were implemented to more accurately control the positioning of endoscopic equipment. The most recent advancement has been the development of tele-operated robots (Figure 1) that allow surgeons to control instruments from remote locations, resulting in minimally invasive, robot-assisted (MIRA) approaches [11].

Figure 1: The da Vinci surgical robot from Intuitive Surgical [15]

These techniques form what is collectively known as minimally invasive surgery (MIS). Many patients are opting for MIS because of the reduced pain, trauma, and recovery times. Minimally invasive procedures reduce the direct and indirect costs for patients because shorter recovery periods result in shorter hospital stays and shorter out of work periods [4].

1.3 Drawbacks of MIS techniques

Current MIS technology requires surgeons to acquire more training and achieve higher levels of surgical skills. Additionally, as these techniques become less invasive, visualization is reduced. With fully open procedures, the surgeon has a direct view of the total operating space. With MIS, the surgeon must rely on indirect endoscopic imaging, which can only provide a limited field of view [3].

Robotic surgical systems come with their own drawbacks. These systems tend to be very expensive, costing hospitals close to \$1 million per installation and requiring expensive service agreements. Due to the inherent complexity of robots, highly trained technicians assist in robotic procedures to handle any equipment-related problems that might arise during surgery. Robotic surgical systems also take up a large portion of the operating room space, requiring nearly 100 cubic feet of workspace for the robot and additional space for the operator console and control instrumentation.

The primary reason robotic surgical systems are so large, costly, and complex is related to the actuators used to control each axis. Although the instruments used to manipulate

tissue are typically less than 1 cm in diameter, the precision servomotors used to actuate the robot are orders of magnitude larger. The increased size of these systems reduces the maneuverability of instruments and visualization of the operating space.

1.4 Purpose

The purpose of this research is to test the hypothesis that miniature actuation can overcome some of the limitations of current surgical robotic systems, namely their size, cost, and complexity, by providing accurate, repeatable control of a small end-effector. This miniature actuation technology could enable future surgical robotic systems to be smaller, less invasive, and more immersive. This research focuses on the actuation and control of a miniature, two-link serial manipulator having potential applicability to a miniature robotic suturing device. The benefits of such a device would include reduced suturing times, reduced space requirements, and reduced surgical costs. The ultimate goal of this research is to provide enabling technology for semi-autonomous, surgical robotics with emphasis on minimally invasive, cardio-thoracic surgery.

2. Background on MIRA Surgery

2.1 Advancements

Tremendous advancements have been made in recent years in the field of minimally invasive surgery (MIS). Surgical techniques have progressed from traditional “open” approaches to laparoscopic approaches and, most recently, minimally invasive robot assisted (MIRA) surgery. MIS techniques have resulted in faster recovery times and reduced trauma for patients, and MIRA approaches are extending the capabilities of MIS by introducing three-dimensional vision, eliminating tremors, and enabling the complex articulation of very small instruments.

Laparoscopic surgery is performed through small ports made in the body instead of large incisions. Long, thin, manually operated instruments are inserted through these ports to cut, sew, and manipulate tissue (Figure 2). This type of operation is easier on the patient, results in faster recovery times, reduces chances of infection, and provides more aesthetically pleasing results. One primary drawback to laparoscopic surgery involves instrument manipulation. The long instruments amplify tremors and other movements of the surgeon because of the lever/fulcrum effect. This lever/fulcrum effect also results in transposition of motions. Motion in one direction outside the body results in motion in the opposite direction inside of the body [3]. Laparoscopic techniques require highly experienced surgeons that are capable of using these instruments precisely. Bailey et al. [3] states, “the skills necessary to perform safe laparoscopic or thoracoscopic surgery have not been routinely taught during most surgical training programs” [3]. Another

problem with laparoscopic surgery involves its limited field of view. Images from the endoscope are magnified and displayed on a two-dimensional monitor, resulting in a loss of depth perception. Bailey et al. [3] states that this loss of three-dimensional depth perception “severely hampers the surgeon’s ability to dissect tissue and identify important landmarks”.

Figure 2: Laparoscopic surgical tools [11]

MIRA surgical procedures are similar to those used in laparoscopic surgery, but the instruments are manipulated by a tele-robotic system that is controlled by the surgeon (Figure 1). The surgeon sits at a console where the operating field is displayed on a monitor (Figure 3). In some cases this console displays a 3-D image. The surgeon uses hand controls similar to joysticks that are located on the console. The surgeon’s movements are transmitted through the robot and result in instrument movements. Servomotors create control forces, and the power is transmitted from the servos down to the instrument tips. These systems allow for very precise control, tremor elimination, and scaled motion. They provide the benefits of laparoscopic surgery while giving the

surgeons increased dexterity precision and a more intuitive feel and control of the instruments.

Figure 3: Surgical console for Intuitive's da Vinci system [11]

2.2 Drawbacks of tele-operated systems

Suture management is an important issue for surgical robotic systems. For mitral valve procedures, suturing can consume the majority of the operating time [18]. A normal mitral valve repair involves about 15-20 sutures, each requiring 5-6 knots. Surgeons are typically very experienced and comfortable tying knots with their hands, but the tele-robotic system adds another level of complexity to this task. Surgeons must become very experienced on the system to perform this otherwise routine task, and even the most experienced surgeons spend a large amount of operating time simply tying knots. Figure 4 is a 10X photo of robotic manipulators placing a suture during a mitral valve repair

procedure. One can easily appreciate the complexity and reduced field of view compared to traditional “open” techniques.

Figure 4: Suturing with Intuitive’s da Vinci robotic system [11]

Overall system size is another drawback of current tele-operated systems. Current robotic systems use macroscopic, extra-corporeal actuators to control microscopic, endo-corporeal motions. The instruments range from 5 to 10mm in diameter and are typically 20mm in length, but the servomotors used to control these tools are orders of magnitude larger. The console and robot occupy a large portion of the operating room floor space. The operating console (Figure 3) takes up about as much space as a computer workstation, and the robot occupies approximately 10 square feet of floor space and 100 cubic feet of operating space. The manipulator robot must be situated beside the patient

and therefore occupies the majority of available space on one side of the patient. Figure 5 illustrates a typical operating room layout for a MIRA procedure.

Figure 5: Typical O.R. layout with robotic surgical system and console [11]

The cost of a single tele-operated robotic system can exceed \$1 million. Additional costs for these systems include training, instrument replacement, and technical support. The author recently investigated the causes of instrument fatigue failures associated with one of the leading surgical robotic systems. These failures resulted in the need for frequent replacement of expensive instruments. Technical support is typically required for tele-robotic procedures to supervise and correct any equipment-related problems. These factors contribute to the high cost of owning and maintaining a tele-robotic surgical system.

3. Miniature Actuation for Medical Applications

The potential advantages of miniature actuation for medical applications are numerous and significant. Although research and development in this area is still in its infancy, researchers have attempted to develop small instrumentation and actuation methods for medical use. One such device, a drug delivery system designed by Cao, et. al [7], makes use of micro-electromechanical systems (MEMS) and piezo-electric technology to drive miniature pumps and regulate dosages of medication. Another medical device benefiting from miniature actuation is the “inchworm endoscope”, which uses small shape memory alloy actuators to contract and elongate segments of a tubular body. These devices have been designed to inspect the inner lining of the colon [[22], [23]]. Another instrument developed by Piers, et al. [22] uses this inchworm concept with several modules connected in series. These modules behave as vertebrae, and each can be actuated separately to produce a snakelike effect.

3.1 Candidate actuation technologies

Possible candidates for miniature actuation of medical instruments included piezo-ceramics, magnetostrictive materials, shape-memory alloys (SMAs), conducting polymers, micro-electromechanical systems (MEMS), electrostatic devices, and micro-fluidic devices. Table 1 lists several main characteristics of these miniature actuation candidates.

Table 1: Typical properties of candidate miniature actuators [[12], [16]]

Actuator Type	Stroke	Response Time	Force/Torque	Driving Energy
Electromagnetic micromotor	rotating	100Hz	1mNm	Electrical
Electromagnetic solenoid	coil length	50Hz	10mN	Electrical
Electrostatic micromotor	rotating	1kHz	10pNm	Electrical
Quad type electrostatic	electrode gap	500Hz	10mN	Electrical
Comb type electrostatic	comb length	500Hz	0.2mN	Electrical
Stacked piezo	0.1% of length	100kHz	5kN/cm ²	Electrical
Bimorph-type piezo	1-5% of length	1kHz	10N	Electrical
Stacked optical PLZT	0.1% of length	1Hz	20kN/cm ²	UV Light
GMA actuator	0.2% of length	1MHz	1kN/cm ²	Magnetic
Wired SMA actuator	3-8% of length	5Hz	10kN/cm	Thermal
Polymer actuator	1-50% of volume	1-10Hz	1N	Chemical, Electrical

Piezo-electric materials produce small strains but very large forces. These actuators strain in the direction of an applied electric field, and have fast response times and very high resolution. Magnetostrictive materials similarly produce small strains, large forces, and have extremely fast response times. Magnetostrictive materials strain in the direction of an applied magnetic field. Shape memory alloys produce significantly larger strains, lower forces, and exhibit slower response times. SMAs are actuated by temperature changes in the material, most commonly applied by passing an electric current through the material. Conducting polymers similarly generate large strains, small forces, and have slow response times. They are driven by applied chemical or electric potentials. MEMS devices can generate small rotational torques and translational forces, and exhibit response times between SMAs and piezo-electrics. Electrostatic devices generate very small displacements and small forces. They require very large electric fields in order to operate. Micro-fluidic actuators operate similar to conventional hydraulic devices but on a much smaller scale. The forces they generate depend on the pressure within the lines, and the actuator stroke can vary depending on the design.

Shape memory alloy wires were determined to be the most appropriate actuator for this micromanipulator application because of their large strain rates and force capabilities. SMAs typically demonstrate 5-8% strains depending on the manufacturing processes, while magnetostrictive and piezo-electric materials are restricted to 0.1-0.2% strains [[12], [16]]. These materials are also attractive because of their high power density and low voltage/current operational requirements. SMAs are actuated via Joule heating, which can be accomplished with currents in the mA range and voltages below 10V [16]. Another beneficial characteristic of SMAs is their mechanical simplicity: the actuator consists of a wire that extends or contracts. No additional gearing, actuator housing, or moving mechanical parts are needed. This characteristic has the potential to simplify the design of a micromanipulator because SMAs perform in a similar manner to human muscles. The Nickel Titanium alloy selected for this application is non-corrosive, biologically compatible, and has been approved by the FDA for medical applications [12].

4. SMA Background

Shape memory alloys exhibit two main properties of interest: (1) the shape memory effect and (2) the super-elastic effect. The shape memory effect is a property of certain alloys exhibiting martensitic transformations. The alloy can be deformed in its low-temperature martensitic phase, but recovers its original shape upon heating to a critical reverse transformation temperature. The shape memory effect was discovered by Chang and Read in 1951 [8], who noticed this effect in a gold-cadmium alloy. In 1963 the properties of nickel-titanium alloys were publicized by Buehler, Gilfrich, and Wiley [6]. Other alloys exhibiting this effect include In-Tl, Cu-Zn, and Cu-Al-Ni. The shape memory effect makes these materials attractive as miniature actuators. These alloys can be deformed in their low-temperature martensitic phase, and heated to a high-temperature austenitic state causing them to return to their original shape. In the case of SMA wires, the material can be mechanically strained to elongate and, when heated, will constrict and return to a lower strain level (its original shape). Super-elasticity is another property of SMAs that permits large, nonlinear recoverable strain upon loading and unloading in the high-temperature austenitic state [20].

4.1 SMA phase transformations

The high-temperature phase of SMAs is an austenite lattice configuration, consisting of a face-centered cubic crystal structure that defines the “hardened” form of these materials. The low-temperature phase is a martensitic configuration resulting from a diffusionless transformation from austenite. The martensitic phase has two distinct configurations

depending on the direction of shear: “martensite plus” and “martensite minus” [14]. Figure 6 illustrates the crystal structures associated with these phase configurations. Figure 7 illustrates how these transformations affect material deformation. The material begins in an arbitrary low-temperature configuration. As stress is applied to the material, the low-temperature phase deforms and elongates. Some of these crystal sites even flip configurations so that the deformation aligns with the applied stress. Upon heating, the crystal structure transforms into austenite, causing the macroscopic deformation that results in the material contracting back to its original size. When the material cools, the crystal sites return to martensite plus or martensite minus.

Figure 6: SMA phase configurations [14]

Figure 7: SMA phase transformation/actuation [14]

4.2 Stress- strain characteristics

Shape memory alloys demonstrate highly nonlinear stress and strain characteristics. The low-temperature, martensitic phase characteristics are commonly referred to as “quasi-plastic” behavior. A typical stress-strain diagram for this phase, shown in Figure 8, reveals that the material can be deformed in the plastic region and retains this plastic deformation.

Figure 8: Quasi-plastic stress/strain of SMAs

Figure 9 illustrates the “pseudo-elastic” characteristics of the austenitic phase. This stress-strain plot is characterized by a linear region near the origin and nonlinear hysteresis loops away from the origin. In this phase the material is harder and has large nonlinear jumps in deformation for given stresses. The direction of each hysteresis loop depends on the previous stress-strain state. With increasing load, the material follows the upper portion of the right hysteresis loop. When unloading from this point, the material follows the lower portion of the right hysteresis loop. The material can also follow internal hysteresis loops inside the larger loops for intermediate loading configurations.

These hysteretic behaviors create challenging issues in controlling the actuator because there can be more than one strain for a given set of inputs.

Figure 9: Pseudo-elastic stress/strain of SMAs

For one actuation cycle, the alloy starts in the low-temperature quasi-plastic region, and as temperature increases the original hysteresis loop divides into two hysteresis loops, producing pseudo-elastic regions. These loops move away from the origin with increasing temperature.

4.3 Energy characteristics

Another way of examining the phase transformations of shape memory alloys is through the Helmholtz and Gibbs free energy relationships. The Helmholtz free energy $\psi(\epsilon)$ is used to describe the energy of an unloaded SMA, while the Gibbs free energy $G(\epsilon)$ is used to describe loaded cases. These diagrams have several wells that correspond to different phase equilibrium states, and the concavity of the wells indicate the energy barriers between phases.

Figures 10-12 show generic Helmholtz free energy diagrams for a shape memory wire at different temperatures. The wells of lower energy on each of these graphs represent equilibrium phases (both austenitic and martensitic). In Figure 10, the wells to the right and left of the origin represent the martensite plus and martensite minus phases below the transformation temperature, respectively. These two phases represent the only stable equilibrium points for this material. As temperature increases (Figure 11), a third well forms in the center representing the emerging stable austenite phase, but there remain significant energy barriers between the martensitic and austenitic wells. At high temperatures (Figure 12), the only stable energy state is in the austenitic phase [14].

Figure 10: Helmholtz free energy diagram - low temperature state

Figure 11: Helmholtz free energy diagram - medium temperature state

Figure 12: Helmholtz free energy diagram - high temperature state

These diagrams reveal the behavior of SMAs during reverse transformations (from high temperatures to low temperatures). As the SMA cools from the high-temperature state (Figure 12) to the medium-temperature state (Figure 11) it remains in the austenite phase because there is a slight energy barrier between the austenitic and martensitic wells. There are two ways to transform the material back into one of two martensitic phases: cool the material to a low-temperature state (Figure 10) where only martensite phases are stable or to add energy through an applied stress.

Figure 13 shows a generic Gibbs free energy diagram during arbitrary loading. The Gibbs free energy is simply the Helmholtz free energy with the product of stress times strain subtracted:

$$G(\epsilon) = \psi(\epsilon) - \sigma\epsilon \quad (4.1)$$

This modification gives the energy diagram the characteristic slope seen in Figure 13. Because this slope is dependent on the stress-strain product, increasing the applied stress

increases the slope of the Gibbs diagram, and serves to overcome the energy barriers that separate austenite and martensite phases. Thus, applied stresses eventually result in phase transformations that put the material in the lowest energy well.

Figure 13: Gibbs free energy

4.4 Other considerations

The response times for SMA materials are limited by cooling rates, which obviously depend on the temperature difference between the wire and its surroundings. Generally, materials with higher transition temperatures exhibit faster cooling rates and faster response times. Active cooling can be employed to significantly improve response times. Fatigue life is also an important design consideration for SMA materials. Although SMAs can achieve strains up to 8%, these strains dramatically reduce the fatigue life (tens of cycles). To realize a longer fatigue life (tens of thousands of cycles), strains not exceeding 4% are desirable. Another important design consideration is overheating. SMA materials heated above their annealing temperature (typically 540K) lose their shape memory effect [12].

5. Serial Manipulator Test Rig

The purpose of this research is to test the hypothesis that miniature actuation can overcome some of the limitations of current surgical robotic systems, namely their size, cost, and complexity, by providing accurate, repeatable control of a small end-effector. This research focuses on the experimental actuation and control of a miniature, two-link serial manipulator having potential applicability to a miniature robotic suturing device. This section details the design of a serial manipulator test rig to be used in experimental demonstrations.

5.1 Conceptual model

This research project was inspired by considering the potential of a semi-autonomous, micro-robotic suturing device that could replicate motions associated with pulling, repositioning, and holding soft tissue in a manner that would align the tissue for suturing. A biologically inspired architecture consisting of a frame, four serial manipulators, a positioning rod, and a suturing device was conceptualized (Figure 14). This micro-robot resembles a four-legged beetle, each leg being a two-link serial manipulator actuated by antagonistic shape memory alloy wires. These manipulators attach to shoulder joints that can be actuated to slide transversely along the body, giving additional degrees of freedom. This conceptual design includes an endoscopic positioning rod that attaches to the body, enabling surgeons to position the semi-autonomous instrument inside the patient. The positioning rod houses wiring and suturing materials to be delivered to the manipulator. The manipulator links are approximately 5mm in length, and the robot

frame is 8mm by 12mm. The size is such that the instrument could be inserted through the standard ports used in minimally invasive, cardio-thoracic procedures.

Figure 14: Solid model of a conceptual, semi-autonomous, micro-robotic suturing device

Figure 14 is a solid model of this conceptual suturing robot, created in SolidWorks™, showing the device above a generic depiction of tissue. Ideally, a surgeon would lower the robot into position above the desired suturing location. On command, the robot would align the tissue, place a suture (or staple), and release the tissue. The surgeon would then reposition the robot for the next suture.

5.2 Experimental scale model

To facilitate the experimental aspects of this research, a 10X scale model of a single manipulator was designed and fabricated. Figure 15 is a preliminary solid model of this

scaled manipulator design, and Figures 16 and 17 are photographs of the fabricated manipulator and test rig. The manipulator was mounted in an inverted fashion for the experiments. The frame and links were constructed from Delrin™, which is easily machinable, mechanically stiff, and electrically non-conductive. The frame was mounted to an aluminum base, and this base was mounted to a wooden platform that also supported the video camera equipment. The manipulator's shoulder joint was machined to slide along the frame's T-slot, allowing for an additional degree of freedom. The SMA wires were attached in antagonistic configurations similar to that of human muscles, as shown in Figures 16 and 17.

Figure 15: CAD solid model of the 10X serial manipulator test rig

Figure 16: Photograph of the 10X serial manipulator

Figure 17: Photograph of the 10X serial manipulator and frame

Figure 18 shows the complete serial manipulator test rig including the frame, base plate, wooden platform, and video camera. A white background was installed behind the test rig to aid in image processing. The video camera was mounted to optimize its field of view with respect to the manipulator's range. The aluminum mounting bracket below the video camera ensured that there was no relative motion between the manipulator frame and the camera.

Figure 18: Photograph of the complete 10X serial manipulator test rig

6. Manipulator Modeling and Kinematics

6.1 Modeling

Figure 19 is a schematic of the single manipulator described in Section 5 with joint angles and link lengths labeled. For serial manipulators of this type, forward kinematics can be used to transform any “joint space” description into a “Cartesian space” description [10]. In other words, forward kinematics provide the (x, y, z) coordinates of the endpoint (in fact any point on the manipulator) given the relative angular displacements $(\theta_1, \theta_2, \dots, \theta_n)$ of its joints. Similarly, inverse kinematics can be used to determine the relative joint angles required to reach a prescribed endpoint location. Once these desired angles are known, a control algorithm can extend or contract the SMA actuators to achieve these desired angles and consequently the desired endpoint location.

Figure 19: Manipulator schematic

6.2 Inverse kinematics

Various analytical and iterative techniques have been developed to analyze the inverse kinematics of generic manipulators, and many of these are described in Craig [10]. For the simple two-link serial manipulator of Figure 19, a straightforward geometrical approach can be applied to find angles θ_1 and θ_2 corresponding to endpoint location (x_e, y_e) . This approach uses the auxiliary variables defined in Figure 20.

Figure 20: Manipulator schematic with auxiliary variables

By geometric inspection, the forward kinematics can be written:

$$\begin{aligned} x_e &= L_1 \cos \theta_1 + L_2 \cos(\theta_1 + \theta_2) \\ y_e &= L_1 \sin \theta_1 + L_2 \sin(\theta_1 + \theta_2) \end{aligned} \tag{6.1}$$

Dimension c is simply the length of the vector from the origin to the end-effector:

$$c = x_e^2 + y_e^2 \quad (6.2)$$

Applying the “law of cosines” yields the following relationship:

$$c^2 = L_1^2 + L_2^2 - 2L_1L_2 \cos \delta \quad (6.3)$$

$$\cos \delta = \frac{c^2 - L_1^2 - L_2^2}{-2L_1L_2} \quad (6.4)$$

Joint angle θ_2 can be extracted:

$$\theta_2 = 180^\circ - \delta \quad (6.5)$$

θ_1 can be found using a similar approach:

$$\theta_e = \tan^{-1}(x_e, y_e)$$

$$\cos \beta = \frac{L_1^2 - L_2^2 - c^2}{-2L_1c}$$

$$\theta_1 = \theta_e - \beta \quad (6.6)$$

Physical measurements from the manipulator indicated that L_1 and L_2 were 50.55 mm and 57.79 mm, respectively.

7. Manipulator Control: Process Overview

Real-time feedback control is necessary to ensure precise positioning of the manipulator's end-effector, even in the presence of disturbances or unexpected loads. Given a desired tip trajectory (x_{ed}, y_{ed}) , the desired angular displacements of the manipulator's joints $(\theta_{1d}, \theta_{2d})$ can be calculated using the inverse kinematic relationships of (6.1)-(6.6). A video camera and real-time image processing can be used to determine the actual joint angles (θ_1, θ_2) , and a real-time control algorithm can be used to regulate currents in the antagonistic SMA cables to reduce or eliminate tracking errors. This section details the development of real-time image processing and control algorithms.

7.1 Hardware-software overview

The equipment used for the experimental aspects of this research, illustrated in Figure 21, consisted of a web camera, a personal computer with two multi-function data acquisition cards, transistor circuits and a power supply, and software for real-time image processing and control.

Figure 21: Hardware schematic

Because position feedback was essential for real-time control, a key factor in the test rig development was selecting appropriate sensors. Although rotary encoders on the manipulator joints would provide acceptable feedback, they were too large, expensive, and cumbersome for implementation on the micromanipulator. Of the various sensing technologies considered, a real-time vision system was determined to be the best choice for several reasons. Vision systems are non-contacting, do not interfere with manipulator dynamics, and have no moving or contacting parts that would add mass, stiffness, or damping. Additionally, endoscopic vision systems are standard equipment in current robotic surgical procedures, so it might be possible to implement this sensing technology with slight modifications to standard vision systems. The real-time vision system would be used to locate key points on the manipulator, namely the positions of the end-effector and joints, and from this information identify joint angles. Budgetary constraints resulted in the purchase of an *iBot Web Cam* from Orange Micro, Inc. This kit included a “firewire connection” camera and PCI card, and was purchased for \$149 (compared to approximately \$2000 for several commercially available image-processing kits).

Data acquisition hardware and software were used to generate the real-time control voltages that actuated the SMA cables. Two National Instruments PCI-6024E data acquisition cards, with termination boards and interconnecting cables, were installed to provide four real-time analog output control signals. These cards were connected to a custom current amplifying transistor circuit (Figure 22) that powered the four SMA actuators.

Figure 22: Electrical schematic of the current-amplifying transistor circuit

MATLABTM was selected for the real-time software platform. Custom software was written to process the video images, implement the control algorithm, and generate the control signals. The web camera generated still images and saved them as JPEG files every 1.0 seconds. The real-time image processing software loaded these images, filtered them, and executed an algorithm that located the centroids of the end-effector and second joint (in Cartesian coordinates). Using these centroid locations, an inverse kinematics algorithm determined joint angles. The control algorithm used the actual and desired

joint angles to determine the appropriate control action. Figure 23 illustrates the logic behind the real-time software.

Figure 23: Flowchart of real-time image processing and control software

8. Manipulator Control: Algorithm Development

Process nonlinearities and limited sample rates made the control of the manipulator test rig complex and challenging. The most significant nonlinearity was hysteresis in the SMA actuators (Figures 8 and 9), which depends on time, position, stress, and temperature. The manipulator kinematics (6.1)-(6.6), also highly nonlinear, exacerbated the problem. Another issue that proved challenging for controller was the limited (and sometimes variable) feedback sample rate (1.0 Hz), which proved to be much slower than the manipulator's dynamic time constant. The software drivers provided with the web camera were designed to convert live images to JPEG format at 1.0-second intervals, but delays were common. The final video processing algorithms made use of code that directly interfaced the web cam to MATLAB™ at a consistent 1.0 Hz rate.

This section outlines the development and experimental evaluations of real-time control algorithms for the manipulator test rig.

8.1 Development of a linear SMA test rig

A linear, single degree-of-freedom SMA test rig was designed and constructed to facilitate controller development. This test rig eliminated the kinematic nonlinearities present in the serial manipulator test rig, enabling the controller development to focus on hysteretic nonlinearities and sample rate issues. This test apparatus, illustrated schematically in Figure 24, consisted of an aluminum frame, two electrically insulated mounting brackets, and an electrically insulated mass constrained to translate in the

vertical direction. This mass was actuated by antagonistic, single-strand FlexinolTM cables of 0.1 mm diameter. To improve response times for this robot, wires with transition temperatures of 90.0 °C. were chosen. The displacement of the mass was measured using a Philtec RC-89 fiber optic displacement sensor, which had a measurement resolution of 142.0 nm and a bandwidth of 20.0 kHz. A small strip of reflective tape was applied to the sliding mass to improve sensor accuracy. The insulated brackets and the sliding mass were also constructed from DelrinTM. A photograph of the linear SMA test apparatus is presented in Figure 25.

Figure 24: Functional schematic of the linear SMA test rig

Figure 25: Photograph of the linear SMA test rig

The linear SMA test rig was powered using a simple transistor circuit (Figure 22) that regulated SMA currents. An antagonistic control approach was implemented, according to:

$$\begin{aligned} u_1 &= u_0 + u_c \\ u_2 &= u_0 - u_c \end{aligned} \tag{8.1}$$

Here u_1 and u_2 represent the upper and lower control voltages, respectively, and u_0 represents a bias voltage. The bias voltage was determined to be 3.0 V, half of the limiting voltage required to protect the SMA cables. The control voltage, u_c , was constrained to be in the range:

$$-3.0 \leq u_c \leq 3.0 \quad (8.2)$$

8.2 Controller development using the linear SMA test rig

Controller development began with an experimental determination of the system's step response. Figure 26 shows the open-loop response to a control voltage step from -3.0 V of 3.0 V at time $t = 0.0$ sec. This step response reveals first-order response characteristics with a time constant (time to reach 63% of steady-state) of approximately 0.06 sec. The peak-to-peak displacement is approximately 2.3 mm.

Figure 26: Open-loop step response of the linear SMA test rig

8.3 Proportional Control with Varying Sample Rates

Proportional+Integral+Derivative (PID) approaches have been used extensively to control SMA actuators [[17], [23]]. PID control has the advantages of being intuitive and easy to implement, but tends to deal poorly with the hysteretic characteristics of SMAs, especially at low sample rates. Nevertheless, a series of closed-loop control experiments using a simple proportional controller was conducted to study the effect of sample rate on tracking performance. The linear SMA test rig was capable of closed-loop control rates in excess of 1,000 Hz. Rates as low as 1 Hz were investigated, as this represented the update rate of the video system to be used on the manipulator.

Figures 27-31 show output tracking responses of the linear SMA test rig using proportional control with update rates ranging from 1,000 Hz to 1 Hz. For each case, the proportional gain was 30,000, and the desired output trajectory was a sinusoid:

$$y_r(t) = 0.00225 + 0.0005 \sin(t) \quad (m) \quad (8.3)$$

These tracking responses clearly show that proportional control performs well at high sampling rates, but the performance degrades rapidly as the sample period nears the system time constant. For sample rates as low as 100 Hz, the tracking performance is excellent. At 25 and 10 Hz, the tracking response deteriorates, becoming highly oscillatory with large tracking errors. At a sample rate of 1 Hz, the tracking performance is extremely poor; the measured response does not resemble the desired trajectory.

Figure 27: Tracking response of the linear SMA test rig: proportional control at 1,000 Hz sample rate

Figure 28: Tracking response of the linear SMA test rig: proportional control at 100 Hz sample rate

Figure 29: Tracking response of the linear SMA test rig: proportional control at 25 Hz sample rate

Figure 30: Tracking response of the linear SMA test rig: proportional control at 10 Hz sample rate

Figure 31: Tracking response of the linear SMA test rig: proportional control at 1 Hz sample rate

8.4 Intelligent Control

The poor tracking performance of proportional control at low sample rates prompted the development of a more sophisticated control algorithm, one that would incorporate a model of the SMA test rig dynamics and provide feed-forward control. Ideally, this controller would be nonlinear and time varying to address the characteristics of SMA dynamics. Various nonlinear models and control approaches have been presented by other researchers [9], but these typically require extensive effort to develop and implement, and results have been mixed.

An intelligent control system was developed to improve tracking performance at 1 Hz sample rates. This controller used an artificial neural network (ANN) to “learn” the

inverse dynamics of the linear SMA test rig on-line, as illustrated schematically in Figure 32.

Figure 32: Functional schematic of the intelligent control system

The ANN consisted of 121 2-dimensional radial basis functions (RBFs), with activation centers uniformly distributed over the normalized input domain $([0,1],[0,1])$. Inputs to the ANN were previous measured displacement $y(kT-1)$ and present desired displacement $y_r(kT)$, both linearly normalized. Output from the ANN was the open-loop control effort $u_c(kT)$ needed to achieve the desired displacement in one sample interval (1 sec). The ANN was trained on-line using standard backpropagation of error techniques [5]. Although it was possible to train the network in a “direct” fashion by comparing “predicted” control efforts with “actual” control efforts using archived data, an “indirect” approach was used. This approach used tracking errors to define a quadratic performance index:

$$J(e) = \frac{1}{2N} \sum_{k=1}^N (y_r(kT) - y(kT))^2 \quad (8.4)$$

Updating the network weights at each time step based on this cost function resulted in the following learning rule:

$$w = w - \eta \cdot (y_r(kT) - y(kT)) \cdot e^{-\left(\frac{(\bar{y}_r(kT) - c_1)^2 + (\bar{y}(kT-1) - c_2)^2}{s^2}\right)} \quad (8.5)$$

Here w represents the ANN weight vector, η is a positive learning rate that determines the rate of convergence, $\bar{y}_r(kT)$ is the normalized reference output, $\bar{y}(kT-1)$ is the delayed normalized measurement, c_1 and c_2 are vectors of RBF centers, and s is the scalar RBF spread.

Figures 33 and 34 show output tracking responses of the intelligent control system at the beginning and completion of on-line training, respectively. Network weights for this 1200-sec experiment were initialized to zero. Complete experimental parameters are compiled in Table 2. These responses show that the tracking performance improves dramatically as learning proceeds. Tracking errors after 1200 sec of training are quite small. Figure 35 shows how the error cost function defined in (8.4) reduces with training from $4.44 \cdot 10^{-7}$ to $4.81 \cdot 10^{-9}$. Figure 36 shows a 15-second tracking response the trained intelligent controller, in this case the desired trajectory has a frequency of 1.0 rad/s (0.16Hz). Figure 37 compares the tracking performance of the trained intelligent controller to the proportional controller ($K_p=30,000.0$) for the same desired trajectory. This figure leaves no doubt that the intelligent controller is far superior to proportional control at low sample rates, and has adequately “learned” the process nonlinearities.

Figure 33: Tracking response of the linear SMA test rig: untrained intelligent control at 1 Hz sample rate

Figure 34: Tracking response of the linear SMA test rig: trained intelligent control at 1 Hz sample rate

Table 2: Experimental parameters for intelligent control of linear SMA test rig

Duration of training	1200.0 sec
Initial weights, w	$[0,0,\dots,0]$
Activation centers, c_1	$[0.0,0.1,0.2,\dots,1.0]$
Activation spreads, s	0.3
Learning rate, η	20.0

Figure 35: Tracking error cost function during on-line training: 0-1200 sec

Figure 36: Tracking response of the linear SMA test rig: trained intelligent control at 1 Hz sample rate

Figure 37: Tracking response of the linear SMA test rig: trained intelligent control vs. proportional control
at 1 Hz sample rate

9. Manipulator Control: Experimental Evaluations

The development and experimental evaluation of an intelligent control strategy outlined in Section 8 utilized a linear, single degree-of-freedom test rig to isolate the kinematic nonlinearities of the serial manipulator. This enabled controller development to focus on the effects of SMA hysteresis and limited feedback sample rates. In this section, the intelligent control algorithms are implemented on the two-link serial manipulator and experimentally evaluated.

9.1 Training the neural network

The manipulator test rig (Figure 17) has two links, each with two antagonistic SMA cables. For this reason, the intelligent controller requires two artificial neural networks, each identical in structure to the network implemented on the linear SMA test rig. To speed up the training process, these neural networks were initially trained off-line using input-output data collected from the manipulator test rig. Random control voltages for each link were generated and applied to the antagonistic SMA actuators, according to:

$$\begin{aligned} u_{1a} &= u_{10} + u_{1c} \\ u_{1b} &= u_{10} - u_{1c} \\ u_{2a} &= u_{20} + u_{2c} \\ u_{2b} &= u_{20} - u_{2c} \end{aligned} \tag{9.1}$$

Here u_{1a} and u_{1b} represent the upper and lower control voltages for Link 1, respectively, and u_{10} represents the Link 1 bias voltage. Again the bias voltages for each link were set

to 3.0 V, half of the limiting voltage required to protect the SMA cables. The randomized control voltages, u_{1c} and u_{2c} , were constrained to be in the range:

$$\begin{aligned} -3.0 &\leq u_{1c} \leq 3.0 \\ -3.0 &\leq u_{2c} \leq 3.0 \end{aligned} \quad (9.2)$$

Input-output data corresponding to applied control voltages and actual joint angles were collected during a 500-second, open-loop experiments at 1-second intervals (Figure 38). This data was used to conduct off-line training of the artificial neural networks and obtain optimal weight vectors w_1 and w_2 , corresponding to the Link 1 and Link 2 ANNs, respectively.

Figure 38: Random input/output training data used for off-line training

A “backpropagation of error” training strategy almost identical to that presented in (8.4)-(8.5) was employed. One ANN learned the Link 1 control voltage (u_{1c}) needed to achieve a desired Link 1 angular displacement (θ_1) in one time step, while the other ANN learned the appropriate control voltage for Link 2. Both weight vectors were initialized with zeros, and 250 samples of input-output data were selected randomly from the original set for training. The remaining 250 samples were used to construct a testing data set, which validated that the ANNs were learning generalized information, not merely “memorizing” the training data. Both learning rates were set to 0.01, and training continued until the training dataset had been presented 1000 times. During this off-line training, all four error cost functions declined significantly, as shown in Figure 39.

Figure 39: Off-line training data for error cost functions

Following off-line training, the optimized weight vectors were implemented for on-line learning and control on the serial manipulator test rig. Both learning rates remained at 0.01, as the dynamics of SMA actuators are known to be time-varying. The desired end-effector trajectory was selected (after considerable experimentation) to be a smooth, closed-loop optimized for the manipulator's reachable workspace:

$$\begin{aligned} x_{ed}(t) &= -9.0 \cdot \cos(0.1t) \\ y_{ed}(t) &= 93.0 - 1.8 \cdot \cos(0.1t) + 2.0 \cdot \sin(0.1t) \cdot e^{-4.5 \cdot \cos(0.1t)} \end{aligned} \quad (9.3)$$

The desired and measured trajectories for a 1500-second, on-line training and control experiment are shown in Figure 40. Tracking for this test was excellent, as the peak tracking errors are below 0.5 mm. Figures 41 and 42 show plots of the error cost functions during this on-line experiment for Link 1 and Link 2, respectively. These cost functions peak and then level off quickly when on-line training begins (recall these networks were pre-trained off-line). These cost functions level off because the ANNs have adequately learned the deterministic features of the system dynamics. Figures 43 and 44 show plots of the weight vectors w_1 and w_2 during on-line training and control. These weight vectors are 121-element arrays that adapt and update based on system dynamics. Both plots show that the weights remain stable and relatively constant. As before, this is a good indication that the off-line training successfully learned the system inverse dynamics.

Figure 40: Tracking response of the manipulator tip: on-line training and control, with off-line pre-training

Figure 41: Error cost function for Link 1: on-line training and control, with off-line pre-training

Figure 42: Error cost function for Link 2: on-line training and control, with off-line pre-training

Figure 43: Weight vector w_1 : on-line training and control, with off-line pre-training

Figure 44: Weight vector w_2 : on-line training and control, with off-line pre-training

The next experiment involved on-line learning and control with weights initialized to zero; i.e. no off-line pre-training. Both learning rates were set to 0.01 as before. Figures 45 and 46 show plots of the error cost functions for Link 1 and Link 2, respectively, during the first 800 seconds of on-line learning and control. Unlike the previous case, these cost functions peak to large values, and continue declining for the remainder of the experiment. Although they never reach the steady-state levels seen in Figures 43 and 44, it is reasonable to expect that extended on-line learning would match the results of on-line learning with pre-training. Figures 47 and 48 show plots of the weight vectors w_1 and w_2 during on-line training and control. Notice that the weights begin at zero, but reach non-zero steady-state trajectories values 500.0 seconds. The weights for Link 1 lie within the range of -0.3 and $+0.3$, while the weights for Link 2 lie between -0.5 and $+0.5$. Figure 49 shows the end-effector tracking response following 700 seconds of on-line training. As before, the tracking performance is outstanding, with all tracking errors less than 0.5 mm.

Figure 45: Tracking error cost function for Link 1 during on-line training: 0-800 sec

Figure 46: Tracking error cost function for Link 2 during on-line training: 0-800 sec

Figure 47: Weight vector w_1 : on-line training and control, with initial weights set to zero

Figure 48: Weight vector w_2 : on-line training and control, with initial weights set to zero

Figure 49: Tracking response of the manipulator tip: intelligent controller trained on-line

10. Conclusion

This paper presents the need for miniature actuation in minimally invasive surgery robotics, and details the development of a 10X scale, two-link micromanipulator. This micromanipulator was actuated using antagonistic Nickel-Titanium shape memory alloy cables. Shape memory alloys were an attractive choice for miniature actuation because of their mechanical simplicity and their high power-to-weight ratio. Real-time measurement and control of the micromanipulator was demonstrated using a vision system and an intelligent controller that “learned” the inverse dynamics of the system. This intelligent controller was attractive because of its ability to learn the highly nonlinear and time-varying dynamics on-line.

Experimental results demonstrate that the manipulator can accurately and repeatably track an optimized trajectory to within 0.5 mm. These results validate the hypothesis that shape memory alloy wires are a promising approach to miniature actuation and warrant further research into their application as miniature surgical actuators.

This test rig performed well despite a sub-optimal design, fabrication imprecision, and limited experimental resources. With precision machining or rapid prototyping, backlash and stiction in the joints could be reduced and the accuracy of this manipulator could be improved. Further work could be done to optimize the SMA actuators, their placement, and the overall design of the manipulator. Finally, improving the bandwidth of the position sensing could dramatically improve performance.

11. Improvements and topics for further research

11.1 Improvements

The first area of improvement would be a redesign of the manipulator attachment points. Tensioning and slipping problems arose when attaching the SMA actuators and the electrical wires to the same post. A possible redesign would include busses or terminals to prevent the electrical connections from slipping while adjusting the tension on the SMA actuators. Another design improvement would maximize the amount of strain in the SMA for the given space would offer an increased range of motion. One possibility would be to use “accordion-like” designs for the actuators. This would increase the effective length of the actuator without increasing the length of the attachment points, thereby increasing the achievable displacement for a given amount of strain.

11.2 Further Research

Further testing should be conducted to determine the speeds ranges at which these actuators can be operated. The speed of actuation for this system was limited by the sample rate. With improved image processing and the addition of active cooling, significantly higher actuation frequencies could be tested. A more advanced vision system would add greater accuracy and repeatability. As shown in the 1-DOF system, excellent tracking results can be obtained with high sampling rates.

Fatigue life for the manipulators should also be studied to determine the effects of cyclic loading on the actuators. Some design factors may need to be revised, such as limiting strain to enhance fatigue life. The effects on the actuators such as changes in force or strain output should also be studied. A favorable advantage of the intelligent control system developed in this research is that it should adapt to such changes in actuator dynamics.

A broader area of research would investigate new methods for suturing, especially for mitral valve repair procedures. The current standard is for the surgeon to tie five or six surgeon's knots with a braided filament. This method is time consuming, but has been proven effective. It would be more advantageous to a miniature suturing system if the process was less complex. The use of staple-like sutures is one possibility. Another option would be to implement existing developmental technology based on the super-elastic effects of Nitinol staples.

12. References

- [1] American Heart Association. 2002 Heart and Stroke Statistical Update. Dallas, Texas: *American Heart Association*, 2001.
- [2] American Heart Association, “Open Heart Surgery Statistics”, <http://216.185.112.5/presenter.jhtml?identifier=4674>, 2002.
- [3] Bailey, R.W., and J.L. Flowers, Ed., *Complications of Laparoscopic Surgery*, Quality Medical Publishing Inc., St. Louis, MO, 1995.
- [4] Borst, C., “Operating on a Beating Heart”, *Scientific American*, October, pp. 2-7, 2000.
- [5] Buckner, G.D., K.T. Schuetze, and J.H. Beno, “Intelligent feedback linearization for active vehicle suspension control”, *ASME Journal of Dynamic Systems, Measurement, and Control*, vol. 123, no. 4, pp. 727-733, 2001.
- [6] Buehler, W.J., J.W. Gilfrich, and R.C. Wiley, “The effect of low temperature phase changes on mechanical properties of alloys near composition TiNi”, *Journal of Applied Physics*, 34, pp. 1475-1477, 1963.
- [7] Cao, L., S. Mantell, and D. Polla, “Design and simulation of an implantable medical drug delivery system using microelectromechanical systems technology,” *Sensors and Actuators*, 94, pp. 117-125, 2001.
- [8] Chang, L.C., and T.A. Read, *Trans. AIME*, 189, p. 47, 1951.
- [9] Choi, S.B., Y.M. Han, J.H. Kim, and C.C. Cheong, “Force tracking control of a flexible gripper featuring shape memory alloy actuators”, *Mechatronics*, 11, pp. 677-690, 2001.
- [10] Craig, J.J., *Introduction to Robotics, Mechanics and Control*, 2nd ed. Addison-Wesley, Reading, MA, 1989.
- [11] Felger, J.E., L. W. Nifong, and W. R. Chitwood, Jr, “The Evolution and Early Experience With Robot-assisted Mitral Valve Surgery”, *Current Surgery*, Vol. 58, Issue 6, pp. 570-575, Nov. 2001.
- [12] Gilbertson, R.G., *Muscle Wires Project Book*, 3rd ed, Mondo-tronics, San Anselmo, CA, 1994.
- [13] Hamilton/Columbus, A., “Forceps! Scalpel! Robot!”, *Time*, June 4 pp. 64-66, 2001

- [14] Heintze, O., and S. Seelecke, *SMA Simulation*, http://www.mae.ncsu.edu/homepages/seelecke/JavaSimWWW/Sma_Sim_home.html, June 2000.
- [15] Intuitive Surgical Website, Intuitive Surgical Inc., <http://www.intusurg.com>, 2002.
- [16] Ishihara, H., F. Arai, T. Fukuda, “Micro Mechatronics and Micro Actuators”, *IEEE/ASME Transactions on Mechatronics*, Volume: 1 Issue: 1, pp. 68 –79, 1996.
- [17] Kohl, M., B. Krevet, and E. Just, “SMA microgripper system”, *Sensors and Actuators*, A 3197, pp. 1-7, 2002.
- [18] Nifong, L. Wiley, Personal Interview, Aug. 9, 2001.
- [19] Noonan, D., “The Ultimate Remote Control”, *Newsweek*, June 25, pp. 71-75, 2001.
- [20] Open Heart.net, “In the Beginning”, <http://www.open-heart.net/Science/Index.cfm>, 2001.
- [21] Otsuka and C.M. Wayman, *Shape Memory Materials*, Cambridge University Press, Cambridge, UK, 1998.
- [22] Piers, J., D. Reynaerts, and H. Van Brussel, “Design of a shape memory actuated endoscopic tip”, *Sensors and Actuators*, A 70, pp. 135-140, 1998.
- [23] Reynaerts, D., H. Van Brussel, “Design of shape memory actuators”, *Mechatronics*, 8, pp. 635-656, 1998.
- [24] Reynaerts, D., J. Peirs, and H. Van Brussel, “Shape memory micro-actuation for gastro-intestinal intervention system”, *Sensors and Actuators*, 70, pp. 157-166, 1999.