9-29-2016 12:00 AM

Shape Memory Alloy Actuators and Sensors for Applications in Minimally Invasive Interventions

Amit Srivastava
The University of Western Ontario

Supervisor
Dr. Rajni V. Patel
The University of Western Ontario

Graduate Program in Electrical and Computer Engineering
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Engineering Science

Abstract

Reduced access size in minimally invasive surgery and therapy (MIST) poses several restriction on the design of the dexterous robotic instruments. The instruments should be developed that are slender enough to pass through the small sized incisions and able to effectively operate in a compact workspace. Most existing robotic instruments are operated by big actuators, located outside the patient's body, that transfer forces to the end effector via cables or magnetically controlled actuation mechanism. These instruments are certainly far from optimal in terms of their cost and the space they require in operating room. The lack of adequate sensing technologies make it very challenging to measure bending of the flexible instruments, and to measure tool-tissue contact forces of the both flexible and rigid instruments during MIST. Therefore, it requires the development of the cost effective miniature actuators and strain/force sensors. Having several unique features such as bio-compatibility, low cost, light weight, large actuation forces and electrical resistivity variations, the shape memory alloys (SMAs) show promising applications both as the actuators and strain sensors in MIST. However, highly nonlinear hysteretic behavior of the SMAs hinders their use as actuators. To overcome this problem, an adaptive artificial neural network (ANN) based Preisach model and a model predictive controller have been developed in this thesis to precisely control the output of the SMA actuators. A novel ultra thin strain sensor is also designed using a superelastic SMA wire, which can be used to measure strain and forces for many surgical and intervention instruments. A da Vinci® surgical instrument is sensorized with these sensors in order to validate their force sensing capability.

Keywords: Shape memory alloys (SMAs), hysteresis, artificial neural network (ANN), extended Kalman filter (EKF), Preisach model, strain sensor, force sensor, minimally invasive surgery and therapy (MIST).

Statement-of Co-Authorship

The thesis presented here has been written by Amit Srivastava under the supervision of Dr. Rajni V. Patel. Parts of the content of the thesis have been published in the proceedings of a refereed conference (mostly from Chapter 3), or will be submitted for publication shortly. The research published in the thesis has been primarily conducted by Amit Srivastava and the published paper has also been written by him. The research has been guided, supported and reviewed by Dr. Patel as the research supervisor.

The material presented in Chapter 2 is to be submitted as,

- Amit Srivastava, Christopher Ward, and Rajni V. Patel, "Adaptive Neural Preisach Model and Model Predictive Control of Shape Memory Alloy Actuators", (To be submitted to IEEE/ASME Transactions On Mechatronics)

  - A. Srivastava - developed the model, designed the control system and the experiments, designed the electrical and imaging part of the experimental set-up, collected data and analyzed the results and wrote the manuscript.

  - C. Ward - designed the mechanical part of the experimental set-up and reviewed the manuscript.

A version of the material presented in Chapter 2 has been published as

- A. Srivastava, R. Xu, A. Escoto, C. Ward, and R. V. Patel, "Design of an Ultra Thin Strain Sensor Using Superelastic Nitinol for Applications in Minimally Invasive Surgery", in *2016 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, Banff, Alberta, 12-15 July, 2016.

  - A. Srivastava - designed the strain sensor and strain gauge amplifier, analyze the sensor characteristics, calibrated and validated the strain sensors for use with the da Vinci surgical instrument, collected data and analyzed the results and wrote the manuscript.

– R. Xu - helped in designing the strain sensor, suggested sensorizing the da Vinci surgical instrument and helped in planning the experimental set-up for calibration and validation of the sensorized instrument. He also reviewed the manuscript.

– A. Escoto and C. Ward - helped in the design of the experimental set-up and reviewed the manuscript.

Acknowledgements

First of all, I would like to express my deepest gratitude to my supervisor, Dr. Rajni Patel, for giving me the opportunity to work with him. This thesis project would not have been possible without his consistent support, guidance and patience. It is impossible to adequately acknowledge him for the help and the suggestions that he provided me throughout my graduate studies. I would like to thank him from the bottom of my heart for being very kind to me. I feel myself fortunate to have him as my supervisor.

I would like to extend my thanks to my uncle Mohan for encouraging me to pursue my graduate studies here at the Western University. Without his motivation, I would have never thought of studying abroad. I am also very thankful to him and Aruna aunty for providing me with a home away from home.

I would like to thank all of my colleagues at CSTAR for their help and suggestions. In particular, I would like to thank Christopher Ward and Abelardo Escoto for providing their assistance throughout my stay at CSTAR. I am also very grateful to Ran Xu for helping me during the most struggling period of my research and providing me a different perspective to look at my research problems. My sincere thanks goes to Farokh Atashzar for never hesitating sharing his knowledge and valuable experiences with me which have always been very helpful to achieve my project goals. I would like to acknowledge Farshad Anooshahpour for having several good discussions with him related to the first part of my project. I also want to appreciate his helping nature and a good sense of humor. I am thankful to Dr. Peyman Yadmellat for proofreading some of the chapters of this thesis and providing me the feedback to improve them. I would also like to thank my friends Saeed, Aaron, Venky and Nikita for being very supportive throughout my time here.

Finally, last but by no means least, I wish to express my sincere gratitude to my parents and my two lovely sisters, Nammo and Shweta didi for their unconditional love, constant support and faith in me. I will be forever indebted to them for all the sacrifices they made to provide me a good life.

Dedicated to

My parents
and
Mohan uncle & *Aruna aunty*

Contents

List of Figures

List of Tables

List of Abbreviations

ANN	Artificial Neural Network
ANPM	Adaptive Neural Preisach Model
ANPMPC	Adaptive Neural Preisach Model Predictive Control
CTR	Concentric-Tube Robot
DAQ	Data Acquisition Card
DOF	Degrees of Freedom
EKF	Extended Kalman Filter
ER	Electrical Resistivity
FBG	Fiber Bragg Grating
FDA	U.S. Food and Drug Administration
FG	Field Generator
FOA	First Order Ascending
FOD	First Order Descending
FOR	First Order Reversal
GF	Gauge Factor
LMS	Least Mean Squares
LM	Levenberg-Marquardt
MIST	Minimally Invasive Surgery and Therapy
MLP	Multi-Layered Perceptron
MPC	Model Predictive Control
MRI	Magnetic Resonance Imaging
NOTES	Natural Orifice Transluminal Endoscopic Surgery
NPM	Neural Preisach Model
SMA	Shape Memory Alloy
SME	Shape Memory Effect
SNR	Signal-to-Noise ratio
SPLS	Single Port Laparoscopic Surgery
TSM	Tendon Sheath Mechanisms

TPM	Tendon Pulley Mechanisms
WCE	Wireless Capsule Endoscopy

Chapter 1

Introduction

1.1 Minimally Invasive Surgery and Therapy (MIST)

MIST involves procedures that require a few small incisions on a patient's body and specialized instruments inserted through these incisions in order to access the surgical site. An endoscope or some other form of imaging technique, such as ultrasound, magnetic resonance imaging (MRI) or X-Ray fluoroscopy, is also used to visualize the operating site during these procedures. The most commonly performed MIST involves needle and catheter based interventions. Over the past two decades, technological advances in MIST have led to the development of more innovative approaches for medical interventions, e.g, natural orifice transluminal endoscopic surgery (NOTES) [1, 2] and single port (incision) laparoscopic surgery (SPLS) [3, 4], with the aim of further reducing trauma to the patient. In comparison to open procedures, MIST offers several advantages to the patient, such as less trauma, quicker recovery, shorter hospital stay and better cosmesis [5, 6]. However, MIST has several drawbacks:

1. Since laparoscopic procedures are performed using long slender instruments through small access points, controlling the end effector of the instruments inside the patient's body becomes very counter-intuitive due to the fulcrum effect [7]. Because of the limited degrees of freedom of the instruments, dexterity is also impaired.

2. The organs cannot be directly manipulated by the clinician's hands. Therefore the sense of touch that is available during open surgery is lost. This means that the clinician cannot really feel how much force is being applied to the tissue. The ability to feel this force can help to improve the efficiency and accuracy in performing tasks such as suturing, blunt dissection and various needle and catheter based interventions.

1.2 Robot-Assisted MIST: Actuation and Sensing Technologies

Over the past two decades, it has been clinically shown and accepted that with robotic assistance in MIST, the problems stated above can be overcome. One such robotic surgical system that has revolutionized the way MIST is performed is the da Vinci® surgical system (Intuitive Surgical Inc., Mountain View, CA) [8]. This is a commercially available system that has been approved by the U.S. Food and Drug Administration (FDA) for several minimally invasive procedures. With robotic assistance, MIST can be performed with much higher accuracy and enhanced dexterity than with manual procedures. However, this brings with it several engineering design challenges for robotic researchers. The reduced access condition in MIST poses several restriction on the design of an appropriate robotic instrument. One major problem is that, in addition to high dexterity, the instruments designed for these procedures must be miniaturized which impedes the use of traditional actuation mechanisms to manipulate the end effector in a confined surgical environment. Incorporating strain and force sensing technology in these instruments is another big challenge. In order to measure the deflection or bending of the flexible instruments (such as needles, catheters and flexible endoscopes) during medical interventions, adequate strain sensors need to be developed which are small enough to pass through the access port, and are able to measure large strains. Similarly, miniature force sensor are needed that can measure tool-tissue contact forces in order to restore the sense of force feedback for the surgeons during the procedures. The sensors and the part of the actuators that remain inside the body during the procedure should also be bio-compatible and sterilizable if they are to be re-used.

The purpose of this section is to provide a brief overview of the state-of-the-art actuation and strain/force sensing mechanisms that are used in the instruments designed for various types of robot-assisted MIST. First, different kinds of actuation mechanisms that are used to manipulate the distal end of the instruments are discussed followed by a discussion of current strain/force sensing technologies. The reason that strain and force sensing technologies are discussed together is that in most cases, forces are calculated by measuring strains in some parts of the instruments; so the basic technology remains the same. Along with the working principle of these technology, examples of a few notable commercial products and research works are also presented in following sections.

1.2.1 State-of-the-Art Actuation Technologies

Le *et al.* [9] have reviewed the different actuation mechanism used in many surgical and diagnostic instruments except those required for needle-based intervention. In this section, a brief overview of some important actuation mechanisms including those for needle-based interventions are provided.

1.2.1.1 Tendon Based Power Transmission

Tendon (or cable) based power transmission is the most commonly implemented actuation mechanism to manipulate the end effector of the instruments used for MIST. In a tendon based force transmission mechanism, the instrument's end effector which remains inside the patient's body during the procedure, can be easily manipulated by actuators located far apart at the proximal end of the instrument and outside the patient's body. Cable driven actuation mechanisms can be divided into two groups: tendon-pulley mechanisms (TPM) and tendon-sheath mechanisms (TSM) [9]. In TPM, a cable is routed around at-least two pulleys: one at the distal end connected to the end effector, and another at the proximal end connected to the actuator (see Figure 1.1). Thus, the rotation of the actuator at the instrument's base can make the end effector to rotate in the same direction. The pulleys and the cables are self contained in a long and sufficiently thin shaft to pass through a small incision. The TPM is generally utilized in rigid surgical instruments, for example, laparoscopic instruments design for the da Vinci and the RAVEN surgical robots [10, 11]. The TSM on the other hand consists of a cable inside a hollow sheath (such as a hellical coil) where one end of the cable is connected to the distal end of the instrument. To actuate the end effector of the instruments, another end of the cable is pulled using an external actuator, thus, the cable slides inside the hollow sheath and forces are transmitted to the distal end of the sheath. This type of actuation mechanism is mostly utilized in flexible instruments such as flexible endoscopes and catheters. The two most notable flexible endoscopic robotic systems that use TSM for their instruments are ViaCath (Hansen Medical, Mountain View, CA, United States) [12] and MASTER (Nanyang Technological University, Singapore) [13]. The two commercially available robotic systems have been designed mainly for NOTES procedure. Artisan Extend^{TM} (Hansen Medical, Mountain View, CA, United States) is an example of an FDA approved robotic catheter that uses TSM to navigate it inside the blood vessels during the vascular procedures.

1.2.1.2 Magnetically Controlled Actuation

Magnetically controlled actuation is another very attractive way of actuation which has gained popularity in the last decade for use in medical interventions. Unlike the tendon driven actua-

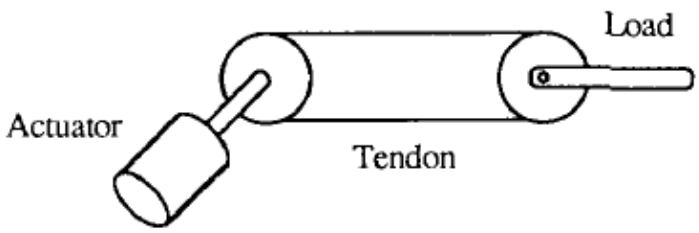

Figure 1.1: A tendon-pulley actuation mechanism [14].

tion mechanism (described in the previous section), a magnetically controlled actuation mechanism does not require any kind of mechanical link between the end-effector and the actuator. In this type of mechanism, a small magnetic component is usually embedded near/inside the end effector of the instrument. The position or/and the orientation of this magnetic component is then controlled by externally moving permanent magnets or electromagnets. An extensive review on magnetically controlled medical devices is given in [15]. Some notable examples of magnetically controlled robotic catheter systems are Niobe® magnetic navigation system (Stereotaxis, Inc., St. Louis, MO, United States) [16] and catheter guidance control and imaging (CGCI) system (Magnetecs Corp., Inglewood, CA) [17]. The Niobe® system uses two large permanent magnets, located on both sides of the patient's bed, to guide the catheter inside the heart. The catheter tip integrated with a small magnet is precisely oriented and positioned by rotating and moving the two permanent magnets mounted on robotic arms. The CGCI system on the other hand uses eight electromagnets, located around the patient's bed and mounted on four robotic arms, to control and navigate the catheter inside blood vessels. The Niobe® magnetic system has been approved by FDA for several cardiovascular interventional procedures, and is currently commercially available.

Another potential application of magnetically controlled actuation is in wireless capsule endoscopy (WCE) for visual examination of the gastrointestinal track (specially small intestine). The Niobe® system (mentioned above) has been proposed for WCE as well [18] and its performance has been evaluated through *in-vivo* experiments in [19].

1.2.1.3 Bevel-tip Flexible Needle Steering

Bevel-tip flexible needle steering is a type of actuation mechanism that is used in needle-based interventions where the needle has to continuously maneuver within tissues, such as during biopsy and therapy. In these kind of procedures, the needle has to be steered inside the tissue in order to avoid anatomical obstacles or to compensate for errors in reaching specific targets. A bevel-tip flexible needle (with or without a pre-bend near the tip) when inserted into tissue, bends in the direction of the tip asymmetry. Thus, by rotating and translating the base of the

needle, using rotational and linear stages, the needle can be steered to reach its target while avoiding collision with some anatomical obstacles [20]. This type of actuation mechanism depends on needle-tissue interaction forces, needle deflection and tissue deformation, which have been studied by several researchers, e.g., [20–23].

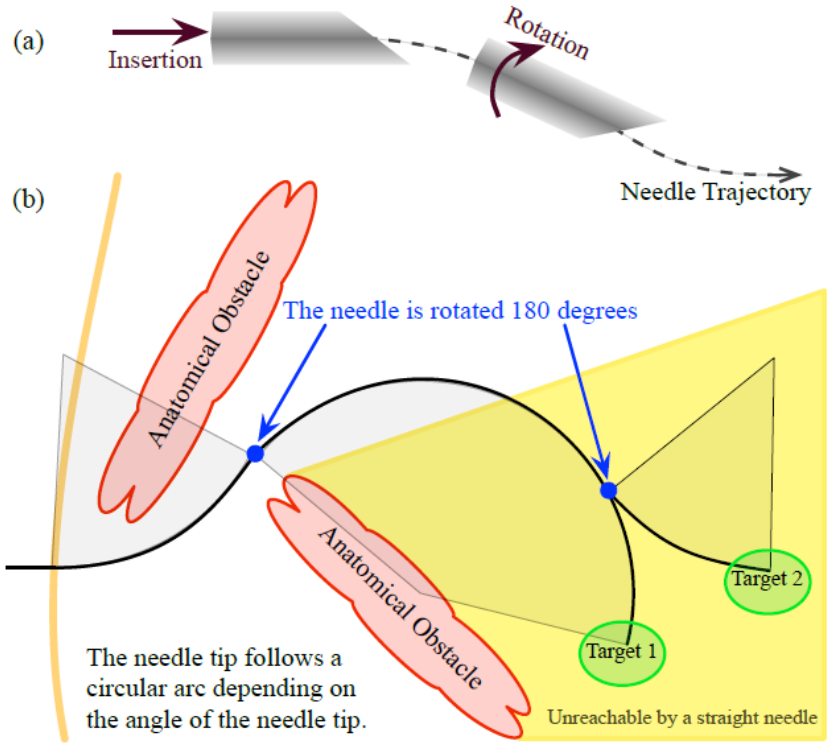

Figure 1.2: Bevel tip flexible needle steering. [22]

1.2.1.4 Concentric-Tube-Robot (CTR)

CTRs are a relatively new type of robotic systems that have been specifically designed for minimally invasive needle-based interventions [24]. The main motive behind developing these robots is to enhance the steering capability of the needles during needle-based interventions regardless of the type of tissue, i.e. soft tissue or hard tissue or fluid-filled body cavity. In order to achieve this goal, CTRs are comprised of two or more pre-curved superelastic metal/metal alloy tubes of different stiffness inserted concentrically (Figure 1.3). Thus, by inserting/retracting and rotating each tube with respect to the other using linear and rotary stages the needle can be steered to follow complex 3D trajectories [24–26].

It should be noted that all these mechanisms require heavy actuators at the base of the instruments or close outside the patient body. These actuators transmit power to the end-effector of the instrument (which remains inside the patient's body) via cables, thin flexible

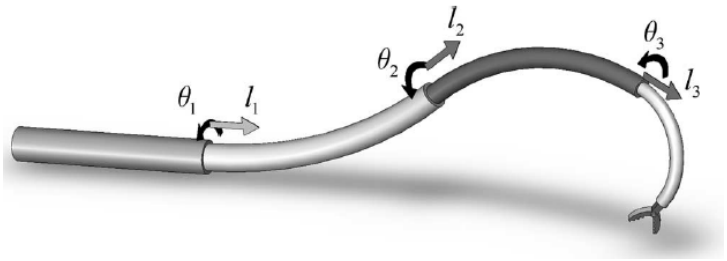

Figure 1.3: A concentric tube robot comprising of four concentric curved tubes [27].

needles or magnetically controlled actuation mechanisms in order to manipulate them inside the body. In most cases, for every additional degree of freedom of the end-effector inside the patient's body, the number of actuators outside the body has to be increased. Therefore, enhancing the dexterity of an instruments usually results in increased cost of the instrument and the space required for them. Although, these robotic instruments have greatly improved the way MIST was performed two decades ago, they are still nowhere close to their optimal design in terms of the cost and the space they acquire in the operating room.

1.2.2 State of the Art Strain/Force Sensing Technologies

A good literature survey of existing force sensing technologies is given in [28]. This section provides a brief overview of some important strain/force sensing technologies and a few examples.

1.2.2.1 Strain Gauges

Strain gauges are the most commonly used strain sensors that convert mechanical strain that is being applied to them into measurable electrical voltages. The working principle of the strain gauge is that their electrical resistance changes when strain is applied to them. This change in resistance can be measured by a quarter, half or full Wheatstone bridge circuit[29]. It is important to mention that if the resistance of a strain gauge is very low, even a small excitation voltage may cause significant temperature drift in the sensor output as a result of Joule heating [29]. Existing strain gauges can be divided into two main categories: metallic foil strain gauges and semiconductor strain gauges. Most of the metallic foil strain gauges have a maximum strain measurement range of 1.5% to 5%, whereas semiconductor strain gauges can measure strain up to only 0.5%. However, semiconducor strain gauges are 10 to 20 times more sensitive than metallic foil strain gauges. Therefore, semiconductor strain gauges are preferred over foil gauges in those applications where low measuring strains and high sensitivity are required. For example, in [30], a miniature force sensor is designed to measure three dimensional forces

acting at the tip of a hand held microsurgical instrument for retinal surgery. The overall force sensor mainly comprises two concentric cylinders joined with each other by 8 thin flexible beams in two layers (4 at each layer) and silicon strain gauge bonded to each of these beams (see Figure 1.4). Using Constantan (a foil) strain gauge, a laparoscopic instrument is sensorized in [31] in order to measure the three DOF force acting at the instrument tip, gripping force and torque about the shaft's axis during minimally invasive procedures. One major problem with the metallic-foil strain gauges is that they are very difficult to install on surgical instruments because they require proper treatment of the object's surface and specific tools. Therefore, considerable expertise is needed which increases the cost of using this technology despite the fact that strain gauges themselves are not very expensive.

Hammond *et al.* [32] have recently developed printing strain gauge technology that eliminates the installation problem of metallic-foil strain gauges. With this method, strain gauges can be directly printed on commercially available devices using a physical vapor deposition method. In [32], they printed a strain gauge on an 18 gauge biopsy needle to measure the tip deflection during the procedure. However, the sensitivity of this strain gauge was found to be very low.

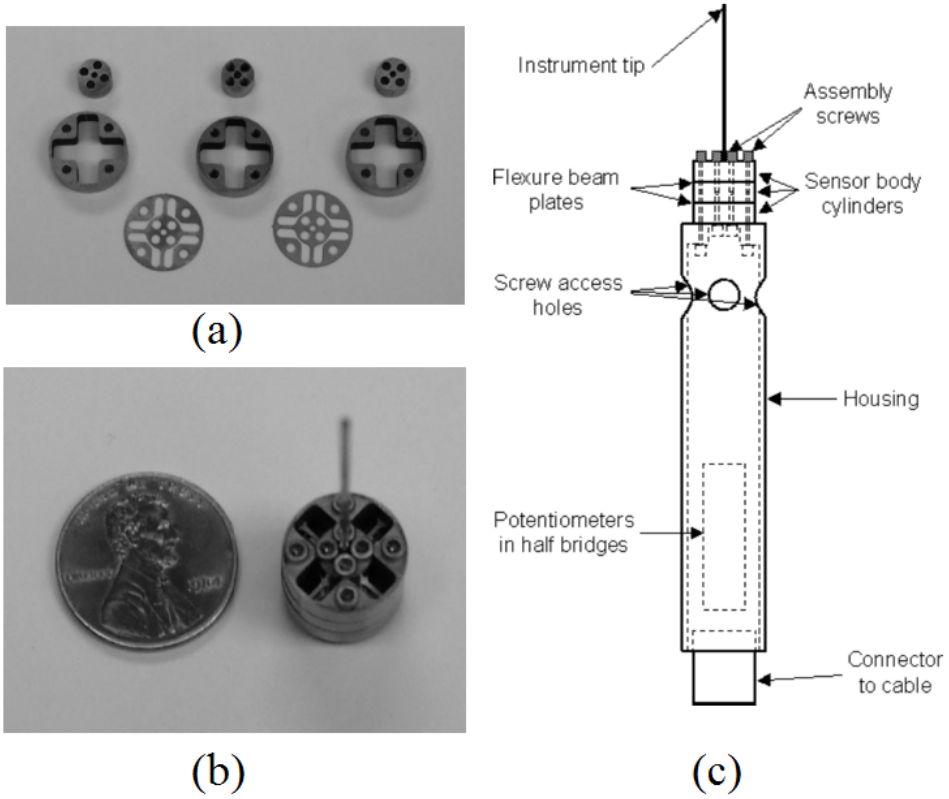

Figure 1.4: (a) Force sensor components, (b) assembled force sensor and (c) microsurgical instrument integrated with force sensor. [30]

1.2.2.2 Optical Sensors

The basic working principle of the optical sensing technology is the change in properties (such as intensity or wavelength) of the light being passed through the optical fiber in response to change in the physical properties (such as temperature or strain) of the fiber. Fiber Bragg Grating (FBG) sensors are a popular type of optical sensing technology. They are manufactured by inscribing periodic/aperiodic variations of refractive index into the fiber core. As a result, when an FBG sensor experience some strain, the wavelength of the reflected light is changed. By measuring the shift in the reflected wavelength, strain can be measured up to 1 μ strain of accuracy. Tacticath™ (Endosense™, SA, Geneva, Switzerland) [33] is an FDA approved ablation catheter that uses FBG sensors to measure contact forces applied by the catheter tip to the heart wall during a cardiac ablation procedure. The force sensor at the distal end of the catheter consists of a deformable body made of some elastic polymer and three optical fibres with FBG sensors on the deformable body (Figure 1.5) [34]. The force at the tip of the catheter causes strain in the optical fibres which in turn changes the wavelength of the reflected light by the FBG sensors. Thus, by measuring these change in wavelength both axial and lateral forces are computed. An application of the FBG sensors in needle-based interventions can be found in [35]. In [35], the FBG sensors are used to measure 3D deflection of an 18 gauge needle during MRI based interventions. Instead of embedding FBGs on a hollow outer needle, the authors modified the inner stylet of the needle by making grooves on it at 120° intervals along its axis, and then attached FBG sensors inside the three grooves. Assuming a cantilever beam model for the needle, the tip deflection is then calculated by measuring strain values from these sensors.

Although, FBG sensors can provide high accuracy and resolution of strain measurements, they are not desirable to be used for disposable instruments because of their high manufacturing cost. In addition, the interrogator system for measuring the wavelength reflected by them is also very expensive (around $20,000 CAD). Therefore, using FBG sensors for strain-measurement always results in a cost-ineffective instrument.

1.2.2.3 Electromagnetic Sensors

Electromagnetic sensors consist of at least two main components: a field generator (FG) and a sensing coil. The FG generates a position varying electromagnetic field which induces current in the sensing coil. The current induced in the sensing coil depends upon the angle and distance of the sensing coil from the FG. Thus, by measuring the electrical signals from the sensing coil, the relative position between the two is calculated. THERMOCOOL® SMARTTOUCH™ (Biosense Webster, Inc., Diamond Bar, CA, United States) is the first FDA approved commer-

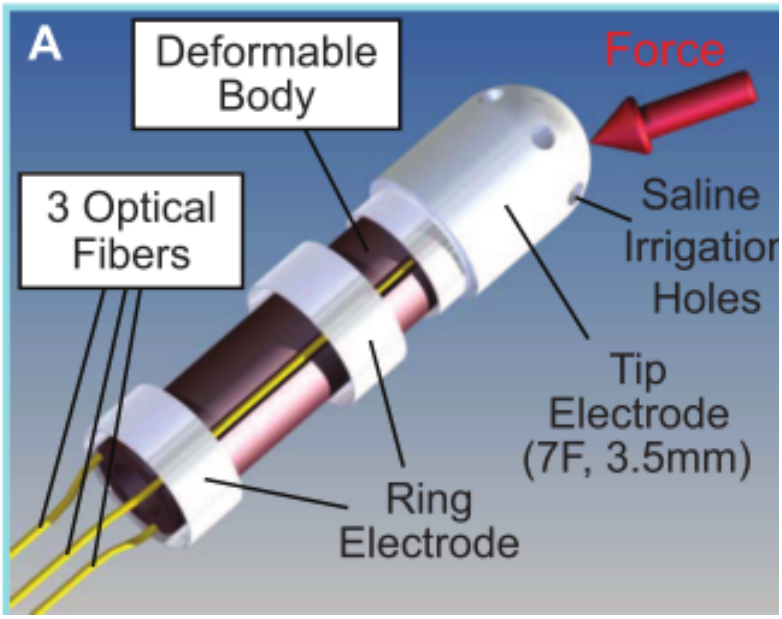

Figure 1.5: Tacticath^{TM} contact force catheter [34].

cially available cardiac ablation catheter [36] that has contact force sensor at the tip. The force sensor at the distal end of this catheter is comprised of a tiny spring, transmission coil and three location (receiver/sensing) coils (see Figure 1.6). The spring allows small bending in it when the catheter tip comes in contact with tissues. The transmission coil emits electromagnetic reference signals, and based on the signals received by the three location coils, bending in the spring is computed in order to estimate the magnitude and angle of the contact forces.

1.3 Shape Memory Alloys (SMAs)

SMAs are metallic alloys that exist in two stable phases: a high temperature Austenite phase, and a low temperature Martensite phase. A reversible solid-solid diffusionless transition between these two phases leads to some unique properties of SMAs [38], which are described below.

a) Shape Memory Effect (SME): The ability of SMAs to remember their original shape is known as SME. This property of SMAs can be illustrated by Figure 1.7. At high temperatures, SMAs exist in the Austenite phase which has a compact and uniform cubic crystal structure. Upon cooling them below the Austenite to Martensite transition temperature ($T_{A\rightarrow M}$), they are transformed into a relatively softer and easily deformable phase which is referred to as "twinned Martensite" phase. When the external stress is applied at this phase, the twinned Martensite phase undergoes some deformation. This deformed phase is known as "detwinned

Figure 1.6: THERMOCOOL® SMARTTOUCH™ catheter [37].

Martensite" phase. At this stage, if the SMAs are subjected to heating, a reverse transformation occurs above Martensite to Austenite transition temperature ($T_{M\rightarrow A}$), and they are transformed back to their parent Austenite phase, thus returning to their original shape.

b) Superelasticity/ pseudoelasticity: Superelasticity is another important property that results from stress induced Austenite to Martensite transformation of SMAs. At their high temperature Austenite phase when external load is applied to SMAs, they are directly transformed into the detwinned Martensite phase and show large deformations. But, they quickly return to their original shape as soon as the load is removed. This occurs without any application of heat (Figure 1.8). This property of SMAs is known as superelasticity or pseudoelasticity.

c) Electrical resistivity (ER) variations: Besides the SME and the superelasticity, SMAs also show very interesting behavior in their electrical resistance in response to the thermo-mechanical loadings. Several thermo-mechanical tests have been conducted in [39] to study the variations in the SMAs' ER when they are transformed from one phase to another. And, It is clearly evident from the literature that apart from having two different crystalline structures, SMAs also have noticeably different ER in their respective phases (Austenite and Martensite phase). In general, the ER in the Martensite phase is higher than that in the Austenite phase. Therefore, unlike other metals or metallic alloys, the electrical resistance of SMAs varies not only due to the direct effects of the stress and temperature but also due to stress and temperature induced Austenite-Martensite transformation.

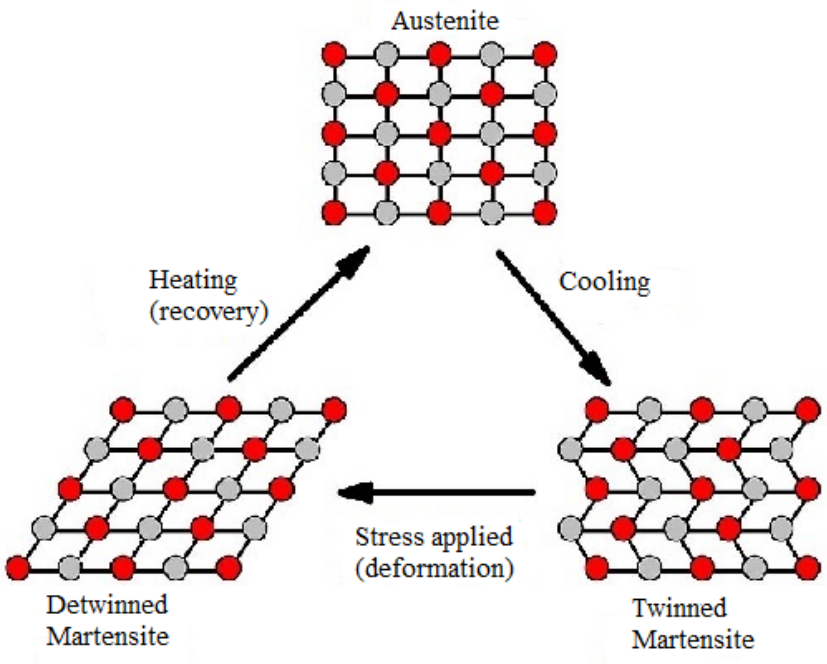

Figure 1.7: Shape memory effect.

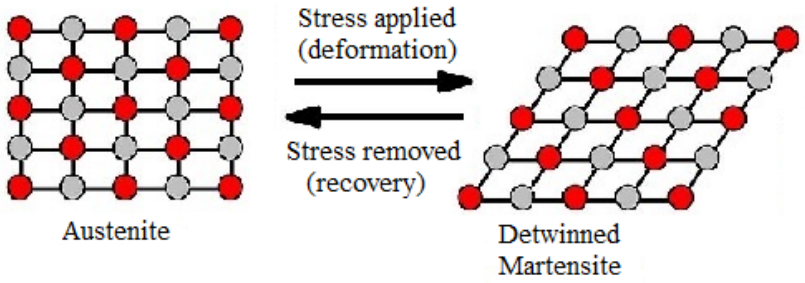

Figure 1.8: Superelastic/ pseudoelastic effect.

By utilizing the combinations of these properties, i.e. SME, superelasticity and ER variations, the SMAs have found a number of biomedical and non-biomedical applications, and have been made to function in many different ways. However, in the next few sections, only those minimally invasive biomedical applications are discussed where they are used either as actuators, as self sensing actuators or simply as strain sensors.

The property of remembering their original shape (i.e. the SME) enables SMAs to produce large force and strain when they are transformed from their deformed Martensite phase to the Austenite phase on application of heating. A typical 5 cm long Nitinol shape memory wire of 0.25 mm diameter can generate up to 8.9 N force over a 2 mm displacement [40] on heating. Thus, by controlling the temperature of SMAs, they can be used as actuators.

As mentioned earlier the SME in SMAs is the result of temperature induced Austenite-

Martensite transformation whereas superelasticity is the result of stress induced Austenite-Martensite transformation. Moreover, the two phases have different ER. Therefore, when the temperature induced Austenite-Martensite transformation occurs and the change in the resistance is used to estimate strain, the SMAs are know to work as the self sensing actuators. And, when the stress induced Austenite-Martensite transformation occurs and the resistance is used to estimate strain, the SMAs act purely as strain sensors.

These three SMA functions and the underlying mechanism are specified in Table 1.1.

Table 1.1: Different functions of SMAs

Function	**Underlying property**
Actuator	SME
Self sensing actuator	SME + ER variations
Strain sensor	Superelasticity + ER variations

1.3.1 SMAs as Actuators

SMA actuators have been extensively used in many surgical or diagnostic instruments. Some of these interesting applications are reviewed in this section.

1.3.1.1 Active Catheter

Several researchers have reported the use of SMAs for the development of active catheters [41–43]. Haga *et al.* [41] have designed a multilink active catheter using only three SMA extension-type spring actuators. The overall diameter of this catheter is 1.8 mm. The outer diameter of the SMA spring is 250 μm, and that of spring's wire is 50 μm. The three SMA springs are fixed around a thin silicone rubber tube at 120 degree. Then, the entire assembly is inserted inside an insulator-coated stainless steel coil. The coating is locally removed for making connections with SMA spring actuators and creating multiple joints. The stainless steel coil is provided ground. Thus, by supplying current to the center of each joint, they are individually controlled (Figure 1.9). In [43], a customized SMA actuator is laser machined from an SMA tubing for application in active catheter. The 1.5 mm long SMA actuator is reported to be able to contract upto 20% when actuated. Jayender *et al.* [42] developed an active catheter instrumented with SMA wire actuators. Three SMA actuators, at 120 degree apart from each other, were placed around the catheter to provide bending in all posible direction.

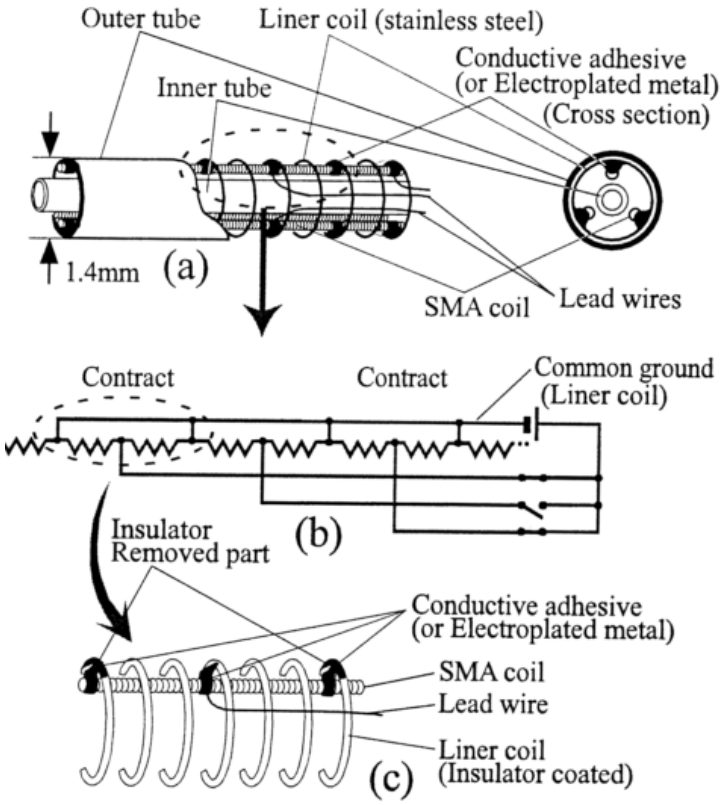

Figure 1.9: SMA springs actuated active catheter [41].

1.3.1.2 Active Needle

Ayvali *et al.* [44] have developed a discretely actuated cannula using a customized Nitinol shape memory actuators for applications in percutaneous and intravascular procedures. The steerable cannula has three straight segments of stainless steel which are then joined by two Nitinol wires (Figure 1.10). The Nitinol wires are annealed to achieve arc shape at high temperature Austenite phase. This design provides local actuation at each joint of the cannula to steer it and guide to the target location. The overall outer diameter of the final prototype is 3 mm.

1.3.1.3 Neurosurgical robot

Ho *et al.* have developed an MRI compatible SMA actuated neurosurgical robot in [45]. This robot has nine serially connected pivot joints where each joint is locally actuated by two antagonistic SMA wires. With some improvements, they redesigned the robot in [46]. In this design, instead of locally actuating using SMA wires, tendon-sheath mechanism is used and the free end of the two tendons were pulled by two antagonistically connected SMA springs. The main advantage of using tendon-sheath mechanism is that SMA actuator can be placed sufficiently

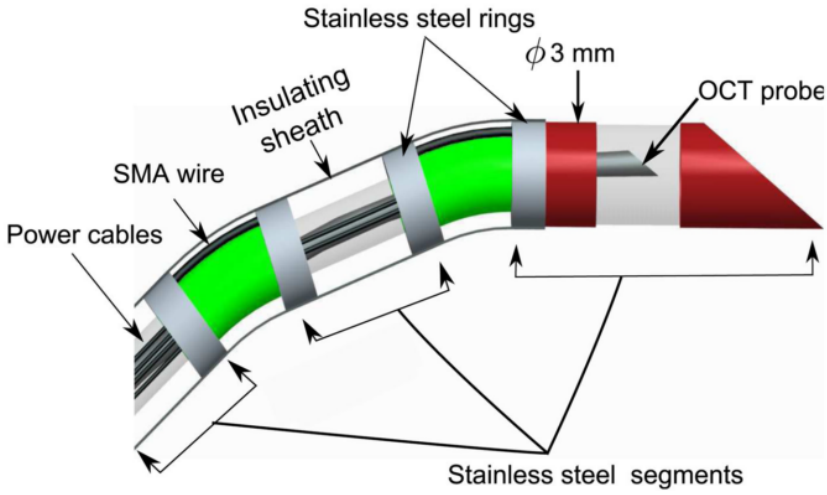

Figure 1.10: SMA actuated steerable cannula [44].

away from the actuating robot and outside the imaging region of MRI scanner. This helps in producing less artifacts in the MR images which is caused by energy dissipated by SMA wires. In addition, the current version of the robot is also shown to be capable of producing large range of motion and large output forces.

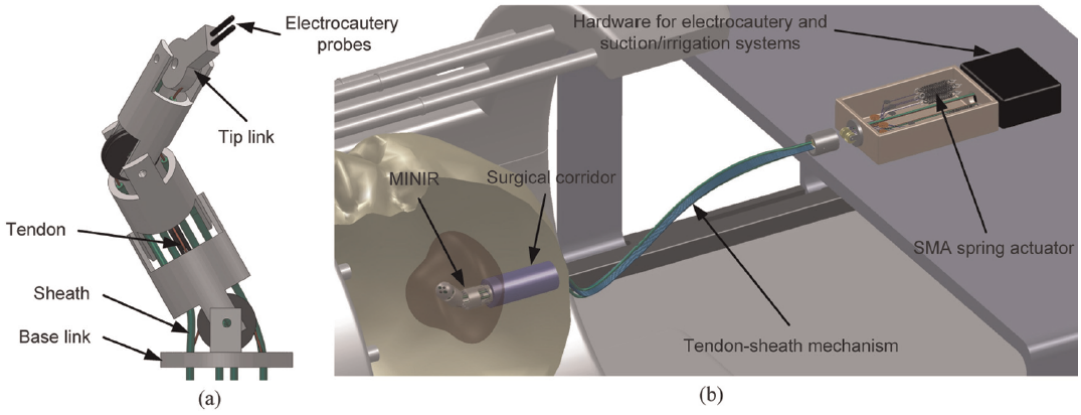

Figure 1.11: SMA actuated MRI compatible neurosurgical robot developed by Ho *et al.* [46].

1.3.2 SMAs as Self Sensing Actuators

As a self sensing actuators, SMAs have been used to design an active endoscope in [47], and an active laparoscopic instrument in [48], which are discussed as follows.

1.3.2.1 Active Endoscope

In [47], Ikuta *et al.* designed a 13 mm diameter active flexible endoscope using SMA spring actuators. The endoscope consists of five segments; four of them have flexibility in the same direction, and the one close to the tip has flexibility in an orthogonal direction. Each segment comprises a stainless steel coil and a pair of antagonistically connected SMA spring actuators (Figure 1.12). The outer diameter of the SMA spring actuators is 1 mm and the diameter of its wire is 0.2 mm. The bending of each segment is obtained by electrically heating the SMAs, while their resistance is used as a feedback to control it.

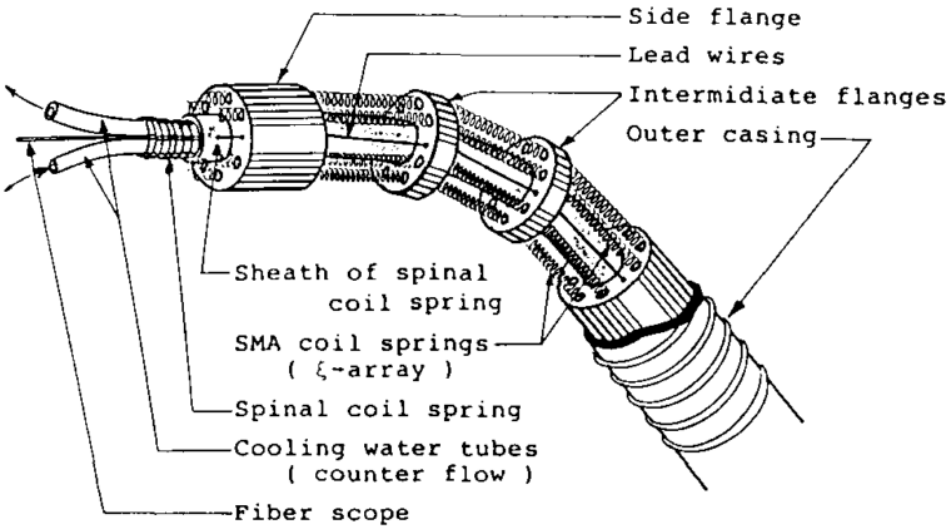

Figure 1.12: An active endoscope designe by Ikuta *et al.* [47]

1.3.2.2 Active Laparoscpic Instrument

In [48, 49], an SMA actuated laparoscopic surgical instrument, with 8 mm of outer diameter, is described. The end-effector of the instrument has three degrees of freedom (pitch, yaw and grip). The end-effector is composed of a ball-joint link, that can be rotated around two axes using 3 pairs of SMA wires, and another pair of SMA wires along with one spring is used to operate the forceps. Using SMA wires' own resistance as feedback, a closed-loop control is implemented to control the wrist motion [50] in all three degrees of freedom. Several in vitro and in vivo experiments are conducted to evaluate the device's performance such as maximum force, velocity and end-effector inclination. It is shown that the proposed design method is sufficient for the assistance task such as clamping, pulling and holding tissue.

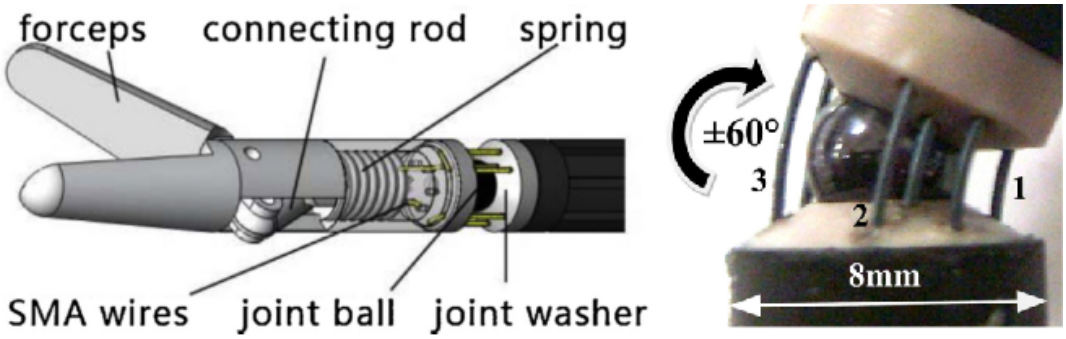

Figure 1.13: (a) SMA actuated laparoscopic forcep and (b) active joint [48, 49].

1.3.3 SMAs as Strain Sensors

A custom-tailored SMA sensor was proposed in [51, 52]. The sensor is made by laser machining a superelastic Nitinol tubing. The overall width of the sensor is reported to be approximately 700 μm. This sensor is designed to provide strain feedback in a catheter that is actuated by another custom-tailored SMA actuator previously designed by the authors in [43]. The sensor is shown to be able to accommodate upto 20% of strain. The gauge factor of this sensor is reported to be 6.9 which is around twice the gauge factor of metallic foil strain gauges, indicating higher sensitivity of their sensor. However, the overall resistance of their 1.5 mm long sensor is only 1.5 Ω which is significantly (around 100 times) lower than that of metallic foil strain gauges. As it mentioned earlier in section 1.2.2.1, due to such a low resistance of their sensor, the output is most likely to have very low SNR.

Schoor *et al.* [53] and Huff *et al.* [54] also proposed to use SMAs as strain sensors, and both have patented their designs. The sensor described in [53] is intended to use in those applications where high forces have to be detected, e.g., parachute canopies and seatbelts. In [54], a thin film SMA sensor is proposed for use as a strain gauge. The manufacturing process of a thin film SMA includes a vapor deposition method to deposit a thin layer of SMA on a polymer substrate. Maximum strain range of this sensor is reported to be only 1.2%. Since, a polymer substrate is used as a backing for thin film SMA, this sensor is most likely to have the same installation problems which are encountered when using metallic foil strain gauges (section 1.2.2.1).

1.4 Project Goals

The unique properties of SMAs (in particular Nitinol), such as biocompatibility, SME, superelasticity and ER variations make SMAs a promising candidate for use both as actuators and

strain sensors for applications in minimally invasive interventions. Therefore, the objective of reducing cost and design complexity of the actuation mechanism in existing robotic surgical instruments while cost-effectively providing force and/or strain feedback to the user (clinician), leads to the two main goals of this research project:

1. To develop a framework in order to control the output of the SMA actuators in real-time so that they can be used for position or force control. To achieve this goal, a model of strain in an SMA actuator with respect to input current is required. Then, using this model, a real-time control strategy needs be developed to appropriately control the SMA actuators.

2. To explore the use of SMAs as strain sensors for instrumentation of minimally invasive interventions. In order to achieve this goal, the use of superelastic SMAs as strain sensors need to be studied in order to develop a design that is simple and can be easily integrated on existing surgical instruments.

1.5 Significant Challenges

The main drawback of using SMA actuators is the nonlinear hysteretic behavior of strain with respect to input current which makes them difficult to control in real-time. In particular, for applications in minimally invasive interventions where accurate positioning of the instrument is required, this problem becomes challenging. Therefore, there is a need to develop an accurate model for SMA actuators which can be used to precisely control an SMA actuated end-effector. The Preisach model is a phenomenological model that can accurately describe the behavior of hysteresis in SMAs based on offline stored experimental data. However, using offline data to implement real-time control is not very convenient. Additionally, updating the offline data in order to adapt to changes in the operating condition is also challenging. Therefore, there is a need to develop an accurate model that can overcome these challenges.

Another major challenge is associated with the development of an SMA strain sensor suitable for thin, possibly flexible, surgical instruments. The sensor should be sensitive enough to accurately measure the forces, and have a large recoverable strain range to avoid saturation when measuring bending strain in flexible instruments. The sensor must also be far thinner than the instruments itself in order to accurately measure strain while maintaining the maximum allowable size of the instrument. In addition, due to low space constraints in MIST, issues in sensor installation and lead wire connections should also be properly addressed.

1.6 Contributions

The main contributions of this thesis can be summarized as follows.

1. A highly accurate adaptive ANN based Preisach model is proposed in this thesis. The proposed model can effectively deal with uncertainties and can be easily used in a real-time control. Thus, the proposed model overcomes the major drawbacks of the Preisach model while keeping its underlying characteristics. Furthermore, a novel control framework is also presented that uses this adaptive model and a model predictive control technique to control the SMA actuator.

2. A novel ultra thin strain sensor is also introduced in this thesis. With the advantage of thin size (15 μm), low cost, high sensitivity and large strain measurement range, this sensor provides a good practical alternative to current strain sensing technologies. The relatively higher sensitivity and larger strain measurement range of this sensor make it suitable for both rigid and flexible instruments. To the best of our knowledge, there is no commercially available strain sensor that is as thin as the sensor proposed in this thesis with similar sensitivity and range. The ultra thin size in addition to the designing simplicity of the sensor make it easy to install on almost all kinds of surgical instruments.

1.7 Outline of the Thesis

The thesis consists of four chapters.

Chapter 1 - Introduction

This chapter presents an overview of the state-of-the-art for actuation and sensing technologies used in current robotic surgical instruments, and discusses their major drawbacks. The unique properties of the SMAs in the context of specific applications are discussed that help in identifying SMAs as a potential alternative to existing actuation and sensing technologies.

Chapter 2 - Hysteresis Modeling and Control of Shape Memory Alloy Actuators

In this chapter, we developed a model and a control strategy to control the strain in an SMA actuator. For this purpose, the Preisach model is modified by using ANNs to replace the experimentally stored data required for the Preisach model with the ANN's weight parameters. Then, the extended Kalman filter (EKF) algorithm is used to update these weight parameters online in order to model the disturbances present in the system. Based on this adaptive model, a model predictive controller is implemented to control the strain in a shape memory Nitinol

wire actuator. The performance of the proposed model and the controller is verified through several experiments.

Chapter 3 - Design of an Ultra Thin Superelastic Shape Memory Alloy Strain Sensor

This chapter presents the design of a novel ultra thin strain sensor using a superelastic Nitinol wire. Before using Nitinol as a strain sensor, some important properties of Nitinol are discussed in the first section followed by the sensor selection criteria and a method for lead wire connection. Then, a method to use this sensor for measuring forces acting at the tip of a rigid surgical instrument is presented. To validate the force sensing capability of the proposed sensor, a da Vinci surgical instrument is sensorized using these sensors. Several experiments are performed to characterize different features such as force measurement accuracy, resolution and temperature drift.

Chapter 4 - Conclusion and Future Works

Chapter 4 provides concluding remarks followed by some suggestions for future works.

Bibliography

[1] T. Arulampalam, S. Paterson-Brown, A. Morris, and M. Parker, "Natural orifice transluminal endoscopic surgery," *The Annals of The Royal College of Surgeons of England*, vol. 91, no. 6, pp. 456–459, 2009.

[2] G. Sebastian, E. J. DeMaria, J. D. Reynolds, D. D. Portenier, and A. D. Pryor, "New developments in surgery: natural orifice transluminal endoscopic surgery (notes)," *Archives of Surgery*, vol. 142, no. 3, pp. 295–297, 2007.

[3] T. E. Udwadia *et al.*, "Single-incision laparoscopic surgery: An overview," *Journal of Minimal Access Surgery*, vol. 7, no. 1, p. 1, 2011.

[4] A. Cuschieri, "Single-incision laparoscopic surgery," *Journal of Minimal Access Surgery*, vol. 7, no. 1, p. 3, 2011.

[5] M. Ohgami, S. Ishii, Y. Arisawa, T. Ohmori, K. Noga, T. Furukawa, and M. Kitajima, "Scarless endoscopic thyroidectomy: breast approach for better cosmesis," *Surgical Laparoscopy Endoscopy & Percutaneous Techniques*, vol. 10, no. 1, pp. 1–4, 2000.

[6] E. Kilger, F. Weis, A. Goetz, L. Frey, K. Kesel, A. Schütz, P. Lamm, P. Überfuhr, A. Knoll, T. Felbinger, *et al.*, "Intensive care after minimally invasive and conventional coronary surgery: a prospective comparison," *Intensive Care Medicine*, vol. 27, no. 3, pp. 534–539, 2001.

[7] A. Gallagher, N. McClure, J. McGuigan, K. Ritchie, and N. Sheehy, "An ergonomic analysis of the fulcrum effect in the acquisition of endoscopic skills," *Endoscopy*, vol. 30, no. 07, pp. 617–620, 1998.

[8] "Intuitive Surgical Inc.". [Online]. Available: http://http://www.intuitivesurgical.com

[9] H. M. Le, T. N. Do, and S. J. Phee, "A survey on actuators-driven surgical robots," *Sensors and Actuators A: Physical*, 2016.

[10] T. Murphy and M. Nixon, "Surgical instrument wrist," Sept. 24 2013, US Patent 8,540,748.

[11] M. J. Lum, D. C. Friedman, G. Sankaranarayanan, H. King, K. Fodero, R. Leuschke, B. Hannaford, J. Rosen, and M. N. Sinanan, "The raven: Design and validation of a telesurgery system," *The International Journal of Robotics Research*, vol. 28, no. 9, pp. 1183–1197, 2009.

[12] D. J. Abbott, C. Becke, R. I. Rothstein, and W. J. Peine, "Design of an endoluminal notes robotic system," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2007, pp. 410–416.

[13] S. J. Phee, S. C. Low, V. Huynh, A. P. Kencana, Z. Sun, and K. Yang, "Master and slave transluminal endoscopic robot (master) for natural orifice transluminal endoscopic surgery (notes)," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2009, pp. 1192–1195.

[14] S. C. Jacobsen, H. Ko, E. K. Iversen, and C. C. Davis, "Control strategies for tendon-driven manipulators," *IEEE Control Systems Magazine*, vol. 10, no. 2, pp. 23–28, Feb 1990.

[15] L. Sliker, G. Ciuti, M. Rentschler, and A. Menciassi, "Magnetically driven medical devices: a review," *Expert Review of Medical Devices*, vol. 12, no. 6, pp. 737–752, 2015.

[16] "Stereotaxis, Inc.". [Online]. Available: http://www.stereotaxis.com

[17] "Magnetecs Corp.". [Online]. Available: http://www.magnetecs.com

[18] F. Carpi and C. Pappone, "Stereotaxis niobe magnetic navigation system for endocardial catheter ablation and gastrointestinal capsule endoscopy," *Expert Review of Medical Devices*, vol. 6, no. 5, pp. 487–498, 2009.

[19] F. Carpi, N. Kastelein, M. Talcott, and C. Pappone, "Magnetically controllable gastrointestinal steering of video capsules," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 231–234, 2011.

[20] N. Abolhassani, R. Patel, and M. Moallem, "Needle insertion into soft tissue: A survey," *Medical Engineering & Physics*, vol. 29, no. 4, pp. 413 – 431, 2007.

[21] R. J. Webster, J. S. Kim, N. J. Cowan, G. S. Chirikjian, and A. M. Okamura, "Nonholonomic modeling of needle steering," *The International Journal of Robotics Research*, vol. 25, no. 5-6, pp. 509–525, 2006.

[22] K. B. Reed, A. Majewicz, V. Kallem, R. Alterovitz, K. Goldberg, N. J. Cowan, and A. M. Okamura, "Robot-assisted needle steering," *IEEE Robotics & Automation Magazine*, vol. 18, no. 4, pp. 35–46, 2011.

[23] A. Asadian, M. R. Kermani, and R. V. Patel, "Robot-assisted needle steering using a control theoretic approach," *Journal of Intelligent & Robotic Systems*, vol. 62, no. 3, pp. 397–418, 2011.

[24] P. Sears and P. Dupont, "A steerable needle technology using curved concentric tubes," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2006, pp. 2850–2856.

[25] H. B. Gilbert, D. C. Rucker, and R. J. Webster III, "Concentric tube robots: The state of the art and future directions," in *Robotics Research*. Springer, 2016, pp. 253–269.

[26] R. Xu, A. Asadian, S. F. Atashzar, and R. V. Patel, "Real-time trajectory tracking for externally loaded concentric-tube robots," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 4374–4379.

[27] P. E. Dupont, J. Lock, B. Itkowitz, and E. Butler, "Design and control of concentric-tube robots," *IEEE Transactions on Robotics*, vol. 26, no. 2, pp. 209–225, 2010.

[28] A. Trejos, R. Patel, and M. Naish, "Force sensing and its application in minimally invasive surgery and therapy: a survey," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 224, no. 7, pp. 1435–1454, 2010.

[29] D. Patranabi, *Sensors and Transducers*, 2nd ed. PHI Learning Pvt. Ltd., 2003.

[30] P. Berkelman, L. Whitcomb, R. Taylor, and P. Jensen, "A miniature microsurgical instrument tip force sensor for enhanced force feedback during robot-assisted manipulation," *IEEE Transactions on Robotics and Automation*, vol. 19, no. 5, pp. 917–921, Oct. 2003.

[31] A. L. Trejos, R. V. Patel, M. D. Naish, A. C. Lyle, and C. M. Schlachta, "A sensorized instrument for skills assessment and training in minimally invasive surgery," *ASME. J. Med. Devices*, vol. 3, no. 4, pp. 041 002–12, Nov. 2009.

[32] F. L. Hammond, M. J. Smith, and R. J. Wood, "Estimating surgical needle deflection with printed strain gauges," in *the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2014.

[33] "St. Jude Medical, Inc.". [Online]. Available: http://www.sjm.com

[34] K. Yokoyama, H. Nakagawa, D. C. Shah, H. Lambert, G. Leo, N. Aeby, A. Ikeda, J. V. Pitha, T. Sharma, R. Lazzara, *et al.*, "Novel contact force sensor incorporated in irrigated radiofrequency ablation catheter predicts lesion size and incidence of steam pop and thrombus," *Circulation: Arrhythmia and Electrophysiology*, vol. 1, no. 5, pp. 354–362, 2008.

[35] Y.-L. Park, S. Elayaperumal, B. Daniel, S. C. Ryu, M. Shin, J. Savall, R. J. Black, B. Moslehi, and M. R. Cutkosky, "Real-time estimation of 3-d needle shape and deflection for mri-guided interventions," *IEEE/ASME Transactions On Mechatronics*, vol. 15, no. 6, pp. 906–915, 2010.

[36] "Biosense Webster, Inc.". [Online]. Available: http://www.sjm.com

[37] S. P. Page and M. Dhinoja, "Smarttouch^{TM} the emerging role of contact force technology in complex catheter ablation," *Arrhythmia & electrophysiology review*, vol. 1, no. 1, p. 59, 2012.

[38] E. Patoor, D. C. Lagoudas, P. B. Entchev, L. C. Brinson, and X. Gao, "Shape memory alloys, part i: General properties and modeling of single crystals," *Mechanics of Materials*, vol. 38, no. 5-6, pp. 391–429, 2006.

[39] V. Novák, P. Šittner, G. Dayananda, F. Braz-Fernandes, and K. Mahesh, "Electric resistance variation of niti shape memory alloy wires in thermomechanical tests: Experiments and simulation," *Materials Science and Engineering: A*, vol. 481, pp. 127–133, 2008.

[40] Dynalloy Inc., "Technical characteristics of flexinol actuator wires," *CA: Tustin*, 2011. [Online]. Available: http://www.dynalloy.com/pdfs/TCF1140.pdf

[41] Y. Haga, T. Mineta, and M. Esashi, "Multi-functional active catheter," *Sensors Update*, vol. 8, no. 1, pp. 147–186, 2000.

[42] J. Jayender, R. V. Patel, and S. Nikumb, "Robot-assisted active catheter insertion: algorithms and experiments," *The International Journal of Robotics Research*, vol. 28, no. 9, pp. 1101–1117, 2009.

[43] A. T. Tung, B.-H. Park, G. Niemeyer, and D. H. Liang, "Laser-machined shape memory alloy actuators for active catheters," *IEEE/ASME Transactions on Mechatronics*, vol. 12, no. 4, pp. 439–446, 2007.

[44] E. Ayvali, C.-P. Liang, M. Ho, Y. Chen, and J. P. Desai, "Towards a discretely actuated steerable cannula for diagnostic and therapeutic procedures," *The International Journal of Robotics Research*, vol. 31, no. 5, pp. 588–603, 2012.

[45] M. Ho, A. B. McMillan, J. M. Simard, R. Gullapalli, and J. P. Desai, "Toward a meso-scale sma-actuated mri-compatible neurosurgical robot," *IEEE Transactions on Robotics*, vol. 28, no. 1, pp. 213–222, Feb 2012.

[46] M. Ho, Y. Kim, S. S. Cheng, R. Gullapalli, and J. P. Desai, "Design, development, and evaluation of an mri-guided sma spring-actuated neurosurgical robot," *The International Journal of Robotics Research*, vol. 34, no. 8, pp. 1147–1163, 2015.

[47] K. Ikuta, M. Tsukamoto, and S. Hirose, "Shape memory alloy servo actuator system with electric resistance feedback and application for active endoscope," in *1988 IEEE International Conference on Robotics and Automation*. IEEE, 1988, pp. 427–430.

[48] X. Yuan, D. Liu, and M. Gong, "Design and research on a shape memory alloy-actuated single-port laparoscopic surgical robot," in *2014 IEEE International Conference on Mechatronics and Automation*, Aug 2014, pp. 1654–1658.

[49] Z. Y. Shi, D. Liu, and T. M. Wang, "A shape memory alloy-actuated surgical instrument with compact volume," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 10, no. 4, pp. 474–481, 2014.

[50] T.-M. Wang, Z.-Y. Shi, D. Liu, C. Ma, and Z.-H. Zhang, "An accurately controlled antagonistic shape memory alloy actuator with self-sensing," *Sensors*, vol. 12, no. 6, pp. 7682–7700, 2012.

[51] A. T. Tung, B.-H. Park, D. H. Liang, and G. Niemeyer, "Laser-machined shape memory alloy sensors for position feedback in active catheters," *Sensors and Actuators A: Physical*, vol. 147, no. 1, pp. 83–92, 2008.

[52] A. Tung, G. Niemeyer, D. Liang, B. Park, F. Prinz, and B. Hu, "Sensors and control for an interventional catheter," Jan. 8 2013, US Patent 8,347,738.

[53] M. van Schoor, A. Lengyel, G. Muller, and A. du Plessis, "Method and device for measuring strain using shape memory alloy materials," Apr. 22 2003, US Patent 6,550,341.

[54] M. Huff, W. Benard, F. Lisy, and T. Prince, "Method and sensor for detecting strain using shape memory alloys," Sept. 23 2003, US Patent 6,622,558.

Chapter 2

Hysteresis Modeling and Control of Shape Memory Alloy Actuator

2.1 Introduction

Due to their high power-to-weight ratio, low cost, silent actuation and biocompatibility, SMA actuators have been preferred over conventional actuators in many applications such as in aerospace [1], robotics [2, 3] and surgical tools [4–6]. The working principle of these actuators is the reversible diffusionless solid-state transition between a low temperature Martensite (M) phase and a high temperature Austenite (A) phase. At low temperatures, when SMAs are deformed with the help of some external means such as biased springs, they can regain their original shape when heated above the M to A transition temperature ($T_{M\rightarrow A}$). Further, when they are cooled below the A to M transformation temperature ($T_{A\rightarrow M}$), they are again easily deformed. This property enables SMAs to produce a large amount of force and strain (3-5%) on actuation. However, SMAs exhibit highly nonlinear temperature-strain behavior. In addition, the hysteresis caused by the difference between ($T_{M\rightarrow A}$) and ($T_{A\rightarrow M}$) makes them very difficult to use in real-time control applications. Therefore, the need arises for developing an accurate model that can be used to implement a real-time control system to control the strain in SMA actuators.

In the past, there have been many attempts to model the hysteresis behavior of SMA actuators ([7–11]). Among them, one of the most popular approaches is using Preisach model. In contrast to the models that are derived from first principles of physics, the Preisach model is a phenomenological approach that can accurately describe the hysteresis behavior (including minor loops in the input-output graph) of many physical systems (such as SMA actuators, piezoceramic actuators, cable-driven instruments and magnetostrictive actuators) without

providing physical insight into it ([9, 12, 13]). Thus, using the inverse Preisach model in a feedforward control loop, the output of a hysteretic system can be controlled [14]. However, this approach has very limited use because the Preisach model makes predictions based on the experimental measurements (first order reversal (FOR) curves) which are obtained offline. Therefore, apart from the complexity of finding the numerical inverse, it is also very difficult to update the tabulated experimental data that is needed to adapt to changes in the environmental conditions and modeling uncertainties [15, 16]. To overcome this problem, in [15], each of the FOR curves are approximated by a second order polynomial and the coefficients of the polynomials are stored in a table. These coefficients are then recursively updated using the least mean squares (LMS) algorithm to model the uncertainties in the system. The drawback of this approach is that the accuracy of the predicted output is limited by how good the polynomials fit the FOR curves. In [16], a single two-inputs (increasing temperature and decreasing temperature) one-output ANN with two hidden layers was used to approximate all of the FOR curves together. Although, open-loop results demonstrate the accuracy of this model, due to the high interconnections of neurons, it is difficult to identify which parameters should be updated in real-time when input is increasing or decreasing. If all of the weights are updated in case of monotonically changing inputs, the weights may converge to the values which could produce undesirable predictions when switching occurs.

In this paper, a new control framework for controlling the strain in the SMA actuators is proposed and experimentally validated. The numerical Preisach model described in [10] is modified so that it can be easily used in a real-time control system to determine the appropriate control signal as well as can adapt to changes in the operating conditions. In this regard, both the first order descending (FOD) and first order ascending (FOA) curves are used and each of these curves are approximated by an ANN. This modification not only improves the accuracy of the predictions but also makes the Preisach model useful for real-time control without changing the underlying properties of the Preisach model itself. The EKF algorithm is used to make the model adaptive by recursively updating the weight parameters of the ANN approximating the corresponding FOR curve. Furthermore, a model predictive controller is implemented that uses the temperature dynamics of the SMA actuator followed by the adaptive neural Preisach model (ANPM) to predict the future behavior of the SMA actuator. Then, the Levenberg-Marquardt (LM) algorithm iteratively computes the optimal control current to achieve the desired strain. Simulations and experimental results demonstrate the performance of the adaptive hysteresis modeling of an SMA wire actuator and the performance of the proposed controller to achieve the desire strain in the actuator.

2.2 Classical Preisach Model

This section describes the classical Preisach model. The major content and all the derivations here are adopted from the book "Mathematical Model of Hysteresis And Their Applications" written by Isaak Mayergoyz [10].

The classical Preisach model is described as an infinite number of parallely connected two-position relays with α and β being "up" and "down" switching values respectively [10]. Mathematically, the output ($f(t)$) of the Preisach model, at any time instant t, can be written as

$$f(t) = \iint_{\alpha \geqslant \beta} \mu(\alpha, \beta) \hat{\gamma}_{\alpha\beta}[u(t)] \, d\alpha \, d\beta \tag{2.1}$$

where $u(t)$ is the input and $\gamma_{\alpha\beta}$ ($\alpha \geqslant \beta$) is the output of the elemental relay operator which has a value of +1, when the input is greater than α, and 0 (-1 for some models), when the input is lower than β. For inputs between α and β, the value of $\gamma_{\alpha\beta}$ remains the same as in the previous state. $\mu(\alpha, \beta)$ is an arbitrary weight function, also known as the Preisach function or Preisach density function.

Geometrically, the Preisach model can be defined on a half plane ($\alpha \geqslant \beta$), where each point (α, β) in the plane corresponds to an elemental hysteresis operator with a specific Preisach function value ($\mu(\alpha, \beta)$). Consider a right angle triangle (Figure 2.1(a) and (b)) with its hypotenuse at line $\alpha = \beta$, and its vertex at (α_0, β_0) where α_0 and β_0 are the maximum and minimum value of the input ($u(t)$) that can be applied to the system. The value of $\mu(\alpha, \beta)$ outside of this triangle is assumed to be zero. At any time instant t, if the input ($u(t)$) is monotonically increased from β_0, a horizontal line $\alpha = u(t)$, in the triangle, starts moving upwards as shown in Figure 2.1(a), and all of the hysteresis operators with α value less than $u(t)$ are switched to the "up" position whereas all of the other hysteresis operators with α value greater that $u(t)$ remain in the "down" position. Now, after reaching a value α_1, if $u(t)$ starts monotonically decreasing, a vertical line $\beta = u(t)$ starts moving towards the left as shown in Figure 2.1(b). This results in all of the hysteresis operators with β value greater than $u(t)$ being turned back into the "down" position. When the input is increased again from the last minimum value β_1, the horizontal line $u(t)$ again moves upwards. Generalizing this pattern at any time instant t, the triangle is always divided into two subsets: $S^+(t)$, consisting of the points (α, β) for which $\gamma_{\alpha\beta}$ value is +1, and $S^0(t)$, consisting of the points (α, β) for which $\gamma_{\alpha\beta}$ value is 0, as shown in Figure (2.2). Thus, (2.1) can be rewritten as

$$f(t) = \iint_{S^+(t)} \mu(\alpha, \beta) \hat{\gamma}_{\alpha\beta}[u(t)] \, d\alpha \, d\beta + \iint_{S^0(t)} \mu(\alpha, \beta) \hat{\gamma}_{\alpha\beta}[u(t)] \, d\alpha \, d\beta \tag{2.2}$$

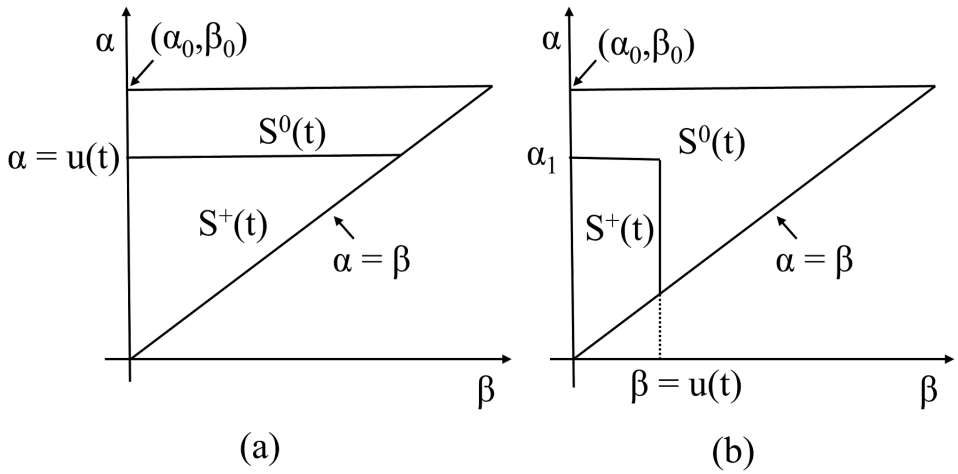

Figure 2.1: Preisach triangle and triangle subdivision due to (a) increasing input, (b) decreasing input.

where

$$\begin{aligned} \hat{\gamma}_{\alpha\beta}[u(t)] &= +1, \quad \text{if}(\alpha,\beta) \in S^{+}(t) \\ \text{and} \quad \hat{\gamma}_{\alpha\beta}[u(t)] &= 0, \quad \text{if}(\alpha,\beta) \in S^{0}(t) \end{aligned} \tag{2.3}$$

Therefore,

$$f(t) = \iint_{S^{+}(t)} \mu(\alpha,\beta)\, d\alpha\, d\beta \tag{2.4}$$

The vertices of the staircase formed by the horizontal and vertical lines (in Figure 2.2) represent the local extremums of the previous inputs.

From (2.4), it can be observed that $\mu(\alpha,\beta)$ is the only unknown to fully describe the hysteresis model of a system. Therefore, the problem of the hysteresis identification using Preisach model is in fact a problem of the Preisach function ($\mu(\alpha,\beta)$) identification, which needs to be determined experimentally. For this purpose, a basic approach, as described in [10], is to directly compute $\mu(\alpha,\beta)$ by differentiating (2.4) with respect to α and β, then, $\mu(\alpha,\beta)$ can be written as

$$\mu(\alpha,\beta) = \frac{d^2 f_{\alpha,\beta}}{d\alpha d\beta} \tag{2.5}$$

where $f_{\alpha,\beta}$ is the output of the hysteretic system along the FOR curves. It can be either FOD or FOA curves. These FOD (or FOA) curves are formed after the first reversal of the input from the major ascending (or major descending) curve as shown in Figure 2.3. Once ($\mu(\alpha,\beta)$) is identified, the output of the Preisach model can be predicted using (2.4). However, this

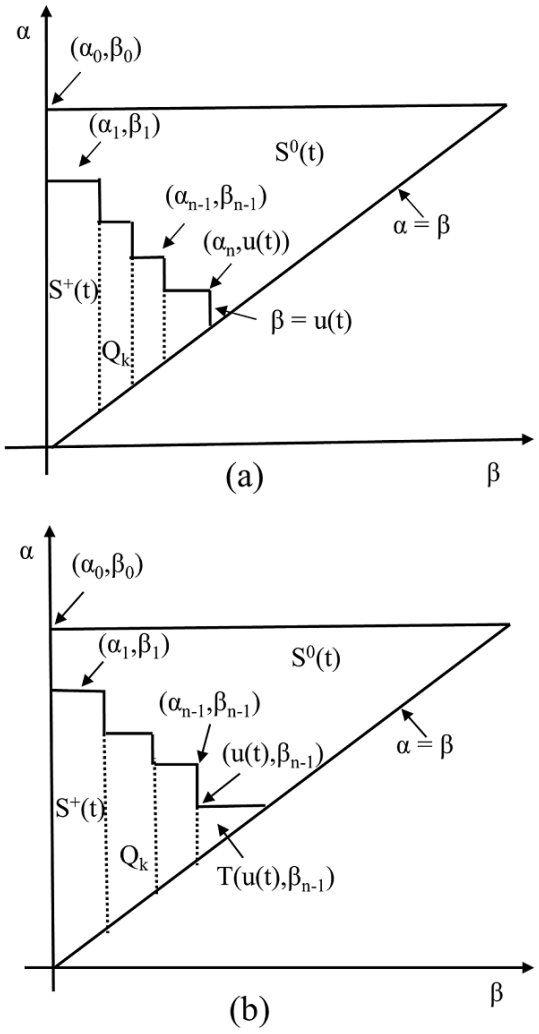

Figure 2.2: Subdevision of the Preisach triangle and the staircase formation due to the (a) decreasing input, and (b) increasing input.

approach suffers from two major limitations: first, it requires the numerical evaluation of the double integrals (2.4) which is a time consuming process and may significantly hinders the use of the Preisach model in real-time applications; and second, the double differentiation required for the identification of the Preisach function may amplify the noise inherently present in the experimental data. Therefore, to avoid these limitations, a numerical approach, also called a "numerical Preisach model" is derived in [10]. If we consider a function $F(\alpha_k, \beta_k)$, described by

$$F(\alpha_k, \beta_k) = f_{\alpha_k, \alpha_k} - f_{\alpha_k, \beta_k} \tag{2.6}$$

then, on a Preisach triangle, this function can be realized as the integral over the area of the

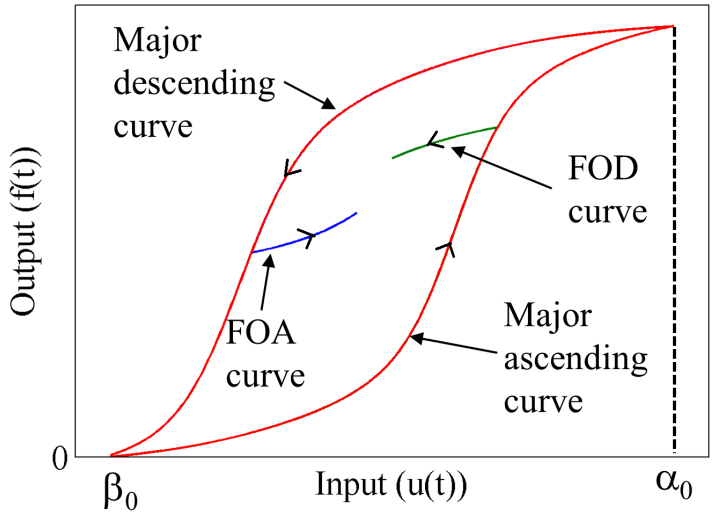

Figure 2.3: Major curves and first order reversal curves.

triangle $T(\alpha_k, \beta_k)$ (Figure 2.4), and can be expressed as follows.

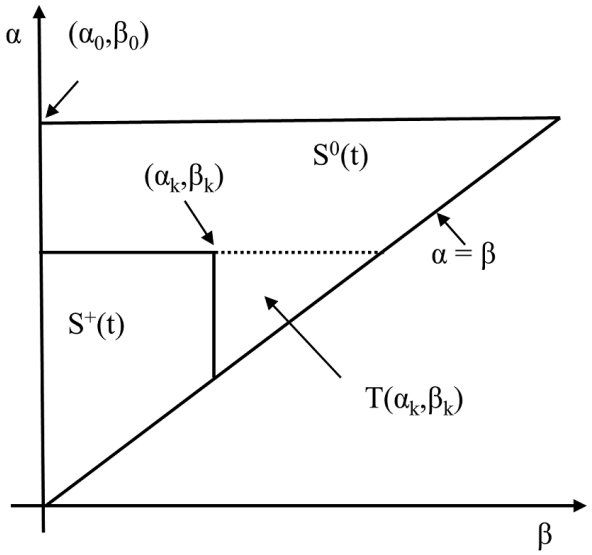

Figure 2.4: A sub-triangle inside a Preisach triangle.

$$F(\alpha_k, \beta_k) = \iint_{T(\alpha_k, \beta_k)} \mu(\alpha, \beta)\, d\alpha\, d\beta \tag{2.7}$$

Moreover, at any time instant t, if the input ($u(t)$) is decreasing, the output can be realized as

the integral over the area of the n trapezoids, formed in Figure 2.2 (a), and can be written as

$$f(t) = \sum_{k=1}^{n(t)} \iint_{Q_k} \mu(\alpha,\beta)\, d\alpha\, d\beta \tag{2.8}$$

Similarly, if the input ($u(t)$) is increasing, the output can be realized as the integral over the area of the $n-1$ trapezoids and one triangle $T(u(t),\beta_{n-1})$, formed in Figure 2.2 (b), and can be written as

$$f(t) = \sum_{k=1}^{n(t)-1} \iint_{Q_k} \mu(\alpha,\beta)\, d\alpha\, d\beta + \iint_{T(u(t),\beta_{k-1})} \mu(\alpha,\beta)\, d\alpha\, d\beta \tag{2.9}$$

Now, each trapezoid Q_k can be written as the difference of the two sub-triangles $T(\alpha_k,\beta_{k-1})$ and $T(\alpha_k,\beta_k)$.

$$\iint_{Q_k} \mu(\alpha,\beta)\, d\alpha\, d\beta = \iint_{T(\alpha_k,\beta_{k-1})} \mu(\alpha,\beta)\, d\alpha\, d\beta - \iint_{T(\alpha_k,\beta_k)} \mu(\alpha,\beta)\, d\alpha\, d\beta \tag{2.10}$$

Using (2.6), (2.7) and (2.10), the model output in (2.8) for a decreasing input can be rewritten as

$$f(t) = \sum_{k=1}^{n(t)-1} (F(\alpha_k,\beta_{k-1}) - F(\alpha_k,\beta_k)) + (F(\alpha_n,\beta_{n-1}) - F(\alpha_n,u(t))) \tag{2.11}$$

or

$$f(t) = \sum_{k=1}^{n(t)-1} (f_{\alpha_k,\beta_k} - f_{\alpha_k,\beta_{k-1}}) + (f_{\alpha_n,u(t)} - f_{\alpha_n,\beta_{n-1}}) \tag{2.12}$$

and the model output in (2.9) for an increasing input can be rewritten as

$$f(t) = \sum_{k=1}^{n(t)-1} (F(\alpha_k,\beta_{k-1}) - F(\alpha_k,\beta_k)) + F(u(t),\beta_{n-1}) \tag{2.13}$$

or

$$f(t) = \sum_{k=1}^{n(t)-1} (f_{\alpha_k,\beta_k} - f_{\alpha_k,\beta_{k-1}}) + (f_{u(t),u(t)} - f_{u(t),\beta_{n-1}}) \tag{2.14}$$

The two important properties that result from the Preisach model are the "congruency" property and the "wiping out" property [10]. According to the "congruency" property, when the input is varied back and forth between two fixed extremum values, the change in the output

remains the same and does not depend upon the past history of the extremum values. According to the "wiping out" property, when the input is increased, all of the vertices whose α value is lower than the current input, and while being decreased, all of the vertices whose β value is greater than the current input, are wiped out. These vertices no longer contribute to the prediction of the hysteresis output.

To implement the numerical Preisach model, the entire input range $[\beta_0, \alpha_0]$ is first divided into N levels. The data for the N FOR curves, f_{α_i,β_j}, such that $\{\beta_0 \leqslant \alpha_i \leqslant \alpha_0, i = 1 : N\}$ and $\{\beta_0 \leqslant \beta_j \leqslant \alpha_i, j = 1 : i\}$ are recorded and stored in a table. Then the output of the model is evaluated using (2.12) or (2.14). The value of f_{α_i,β_j} is obtained by looking up the table. If a point (α_i, β_j) coincides with the stored point in the table, that the value is directly used; otherwise, linear interpolation of the neighboring samples is performed to approximate this value.

2.3 An Artificial Neural Network (ANN)

An ANN is a type of machine learning method that can be used to approximate any smooth multidimensional nonlinear function. An ANN is described as a system of highly interconnected neurons with one or more layers. In general, a multilayer feedforward neural network, also referred to as a multilayer perceptron (MLP), has one input layer, one or more hidden layers and one output layer [17]. The input-layer neurons receive inputs and produce the same value at their output without any modifications; however, the hidden layer(s) and output layer neurons calculate the weighted sum of the values, received from the previous layer neurons, add a bias to it and then pass it through some predefined activation function (such as log-sigmoid, tangent-sigmoid or purely linear) to produce their output. Figure 2.5 shows the structure of a two-inputs (x_1 and x_2) and single-output (y) feedforward neural network with two hidden layers and having three neurons in each hidden layer. The neurons in the hidden layers have a log-sigmoid activation function and the neuron in the output layer has a linear activation function. For this neural network, the output of the first hidden layer neurons (Y_j^1) and second hidden layer neurons (Y_j^2) is written as

$$Y_j^1 = \frac{1}{1 + exp((\sum_{i=1}^{2} IW_{i,j} x_i) + IB_j)} \quad \text{for} \quad j = 1, 2, 3. \tag{2.15}$$

$$Y_j^2 = \frac{1}{1 + exp((\sum_{i=1}^{3} LW_{i,j} Y_i^1) + LB_j)} \quad \text{for} \quad j = 1, 2, 3. \tag{2.16}$$

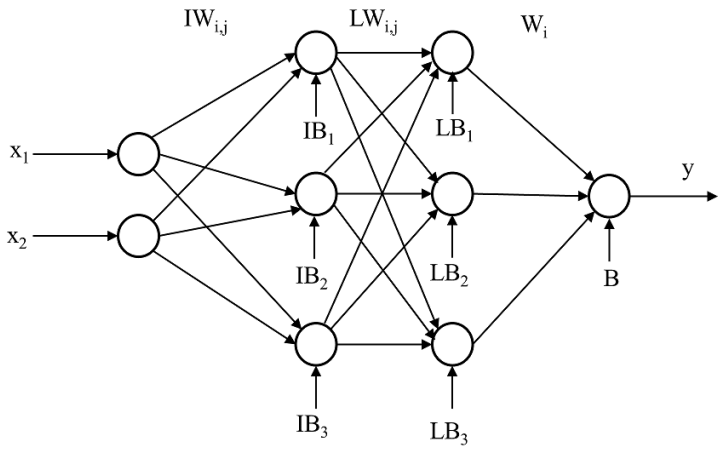

Figure 2.5: A static feedforward neural network.

where $IW_{i,j}$ denotes the weight connecting the output of the i^{th} neuron of the input layer to the input of the j^{th} neuron of the first hidden layer and IB_j denotes the bias applied to the j^{th} neuron of the first hidden layer. Similarly $LW_{i,j}$ and LB_j denote the weight connecting the output of the i^{th} neuron of the first hidden layer to the input of the j^{th} neuron of the second hidden layer, and the bias applied to the j^{th} neuron of the second hidden layer respectively.

The output of the neural network is given by

$$y = (\sum_{i=1}^{3} W_i Y_i^2) + B \tag{2.17}$$

where W_i and B denote the weight connecting the output of the i^{th} neuron of the second hidden layer to the input of the output neuron, and the bias applied to the output neuron respectively.

It should be noted that the number of input layer neurons is always equal to the number of inputs to the neural network and the number of the output layer neurons is always equal to the number of outputs of the neural network. Thus, to model an unknown input-output system using neural networks, the neural network is trained to find the optimum weights and bias values such that the mean-squared sum of the errors between the experimentally measured output and the ANN predicted output is minimized. The commonly used ANN training algorithms are the Levenberg-Marquardt (LM) algorithm, the Quasi-Newton algorithm and the Resilient backpropagation algorithm.

2.4 Hysteresis Modeling of the SMA Actuator

To model the strain in the SMA wire actuator with respect to the current flowing through it, this section first describes the temperature dynamics to simulate the temperature of the SMA. Then, the rate-independent hysteresis between the strain and the temperature of the SMA wire is modeled using the Neural Preisach Model (NPM).

2.4.1 Temperature versus current modeling

The strain in an SMA actuator directly depends on its temperature which in turn can be controlled by the process of Joule heating and natural convection. The temperature dynamics of the SMA wire actuator, used in this chapter, can be described as follows,

$$\dot{T} = \frac{u^2 R - hAl(T - T_a)}{mc_p} \tag{2.18}$$

where u is the current being passed through the SMA wire, R is the resistance, m is the mass, l is the length, c_p is the specific heat capacity, $A = \pi d$ is the circumferential area, d is the diameter, T is the temperature, T_a is the ambient temperature, and h is the cooling rate coefficient. To reflect the higher convection rate at higher temperatures, we model h as the quadratic polynomial of T as described in [8].

$$h = (h_0 + h_2 T^2) \tag{2.19}$$

All of these SMA parameters used in this work are specified in Table 2.1

Table 2.1: Specifications of the SMA Parameters

SMA Parameter	**Value**
Mass (m)	0.03166 g
Specific heat capacity (c_p)	0.8368 $J.g^{-1}K^{-1}$
Length (l)	0.1 m
Diameter (d)	0.254e-3 m
Ambient temperature (T_a)	21°C
Cooling rate coefficient (h_0)	55 $J.m^{-2}.s^{-1}.K^{-1}$
Cooling rate coefficient (h_2)	6e-5 $J.m^{-2}.s^{-1}.K^{-3}$

2.4.2 Strain versus temperature modeling

2.4.2.1 Neural Preisach Model (NPM)

To model the strain versus temperature hysteresis in the SMA wire actuator, the numerical Preisach model was modified in following aspects.

- Both, the FOD and FOA curves were used instead of only one of the two.
- Instead of storing actual sample values, the ANN is used to approximate each of the FOD and FOA curves, then the weights and bias values of the ANNs corresponding to FOD and FOA curves are stored in two separate tables.

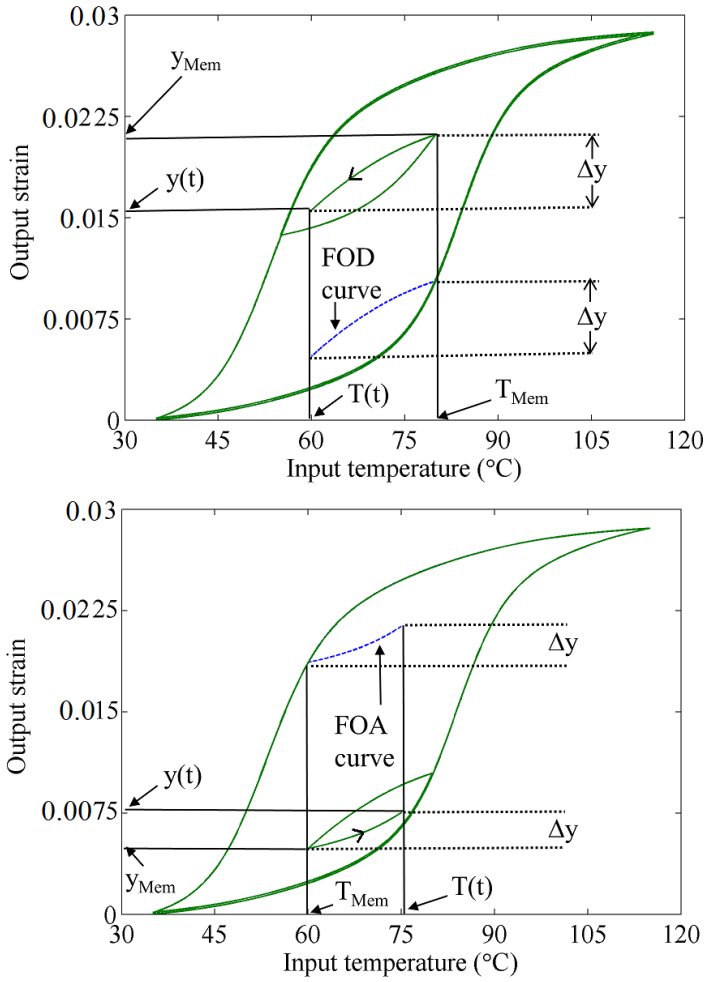

Figure 2.6: First order reversal curves.

From Figure 2.6, utilizing the “congruency” property of the Preisach model the prediction

equation of the strain $y_p(t)$ can be written in terms of FOD (or FOA) curves as follows,

$$y_p(t) = y_{Mem} + y^{NN}_{T_{Mem}}(T, W) - y^{NN}_{T_{Mem}}(T_{Mem}, W) \tag{2.20}$$

where T_{Mem} and y_{Mem} represent the temperature and strain memory at the local extremum respectively. $y^{NN}_{T_{Mem}}(T, W)$ represents the output value of the ANN approximated FOD (or FOA) curve, which starts from the temperature T_{Mem}, at current temperature T, and W denotes the weight parameters of the ANN. Note that the history of the extremums must be stored, and updated according to the "wipe out" property of the Preisach model.

To implement the NPM, the entire input range is first divided into N levels ($T_i, i = 1, 2...N$). The data for FOD and FOA curves that start from these N input values are recorded. A single-input and single-output feedforward ANN is trained offline for each of these FOR curves and the weight parameters of the ANNs corresponding to the FOD and FOA curves are stored in two separate tables. If the temperature at the local extremum (T_{Mem}) coincides with the sampled input value, the output can be predicted using (2.20). Most often it is not the case, therefore $y^{NN}_{T_{Mem}}(T)$ in (2.20) is calculated by linear interpolation of the two nearest FOR curves as follows,

$$y^{NN}_{T_{Mem}}(T) = c_1.y^{NN}_{T_1}(T, W_{c_1}) + c_2.y^{NN}_{T_2}(T, W_{c_2}) \tag{2.21}$$

$$c_1 = \frac{T_2 - T_{Mem}}{T_2 - T_1} \text{ and } c_2 = \frac{T_{Mem} - T_1}{T_2 - T_1} \tag{2.22}$$

where T_1 and T_2 are the two nearest input samples to T_{Mem}, W_{c_1} and W_{c_2} are the weight parameters of the two nearest FOR curves, and c_1 and c_2 are their weighting coefficients.

Thus, the NPM reduces the required memory resources because only a few weighting parameters need to be stored instead of all of the actual experimental measurements while also improving the accuracy of the predictions. Another advantage of this method is that an analytical expression is obtained which has a similar functional form for both decreasing and increasing inputs, that can easily be used in real-time control system using any nonlinear control method such as a nonlinear model predictive control as described in section 2.8.

2.4.2.2 Adaptive Neural Preisach Model (ANPM)

Since the ANNs used in NPM for the approximation of FOD and FOA curves, are trained offline, in order to model the uncertainties in real-time, an adaptive NPM is described in this section. The block diagram of the overall adaptive hysteresis modeling (current to strain) of the SMA actuator is shown in Figure 2.7, which consists of the temperature dynamics described in

section 2.4.1, prediction equation (2.20), a memory block; and the EKF algorithm block that updates the NPM model parameters. The memory block in the ANPM performs the following functions:

- Determines whether the SMA wire is heating or cooling by comparing its temperature at the current and previous state. Based on this, selects one of the two tables: one, that contains the weight parameters of the FOD curves, or other that contains the weight parameters of the FOA curves.
- Based on the "congruency" and "wiping out" property, the block stores and updates the local extrema, calculates T_{Mem}, and based on T_{Mem}, determines T_1, T_2, c_1, c_2, W_{c_1} and W_{c_2}.
- Updates the W_{c_1} and W_{c_2} values in the table which are returned from the EKF algorithm.

In [18], the EKF was used to train a multilayered neural network in real-time. In this work we also use the EKF to update the weights parameters of the ANN realized FOD and FOA curves in real-time. Writing the descrete state-space representation of the NPM as

$$W_c(k+1) = W_c(k) + \omega(k) \tag{2.23}$$

$$y_p(k) = c_1[y^{NN}_{T_1}(T, W_{c_1}(k)) - y^{NN}_{T_1}(T_{Mem}, W_{c_1}(k))] + c_2[y^{NN}_{T_2}(T, W_{c_2}(k)) - y^{NN}_{T_2}(T_{Mem}, W_{c_2}(k))] + y_{Mem} + \nu(k) \tag{2.24}$$

where

$$W_c(k) = \begin{bmatrix} W_{c_1}(k) \\ W_{c_2}(k) \end{bmatrix}$$

and $\omega(k)$ and $\nu(k)$ are zero mean-Gaussian noise with covariance matrices Q and R respectively. Then, using the formulation as described in [19], at each sampling instant k, the update equations for the ANPM weights can be written as

$$W^1_c(k) = W_c(k-1) \tag{2.25}$$

$$P^1(k) = P(k-1) + Q \tag{2.26}$$

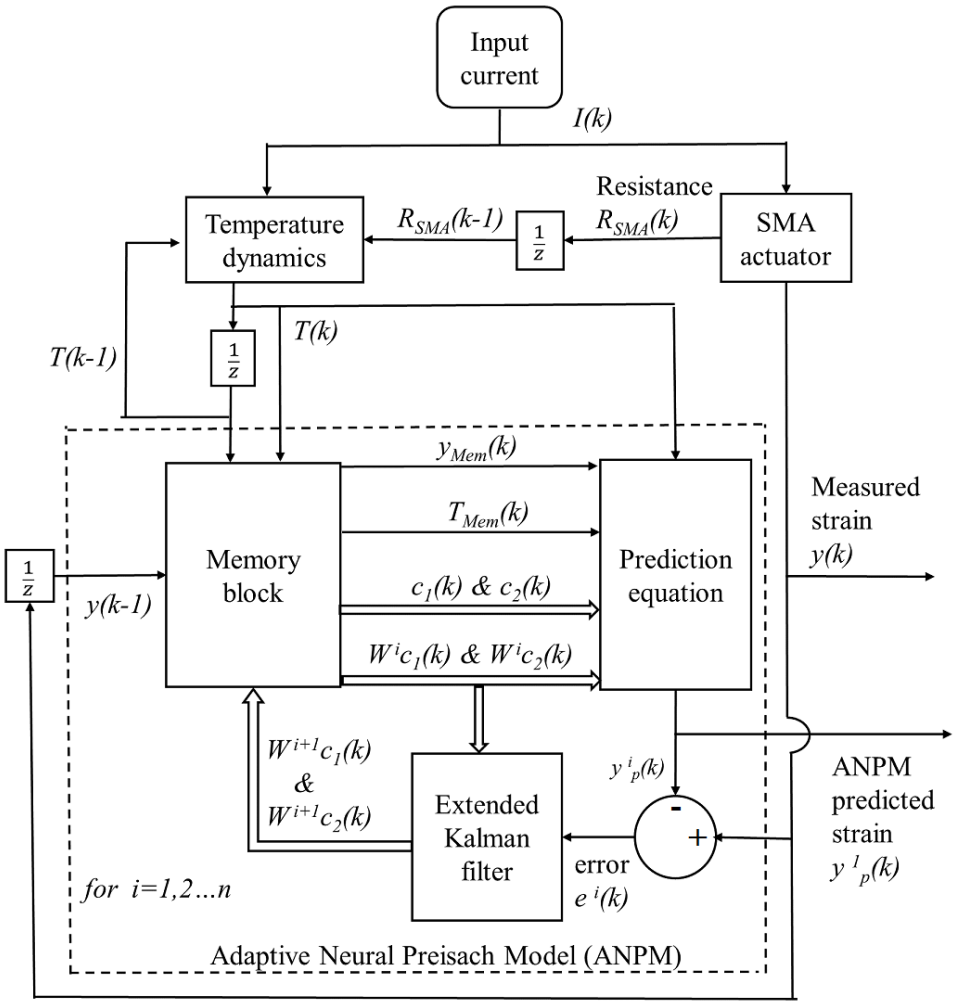

Figure 2.7: Block diagram of the adaptive current to strain hysteresis modeling of the SMA actuator using NPM.

for $\boldsymbol{i = 1, 2...n}$

$$K^i(k) = P^i(k)H^i(k)[(H^i(k))^T(k)P^i(k)H^i(k) + R]^{-1} \tag{2.27}$$

$$\begin{aligned} y^i_p(k) = c_1[y^{NN}_{T_1}(T, W^i_{c_1}(k)) - y^{NN}_{T_1}(T_{Mem}, W^i_{c_1}(k))] \\ +c_2[y^{NN}_{T_2}(T, W^i_{c_2}) - y^{NN}_{T_2}(T_{Mem}, W^i_{c_2})] + y_{Mem} \end{aligned} \tag{2.28}$$

$$W^{i+1}_c(k) = W^i_c(k) + K^i(k)[y(k) - y^i_p(k)] \tag{2.29}$$

$$P^{i+1}(k) = [I - K^i(k)H^i(k)^T]P^i(k) \tag{2.30}$$

$$P(k) = P^n(k) \tag{2.31}$$

$$W_c(k) = W_c^n(k) \tag{2.32}$$

where

$$H^i(k) = \Big[\frac{\partial y}{\partial W_c}\Big]_{W_c=W_c^i(k)} \tag{2.33}$$

P is the state error covariance matrix, K is the Kalman gain, y_p and y are the predicted and measured outputs respectively. The superscript i denotes the iteration number, and n denotes the total number of iterations for the weights update algorithm at each sampling instant.

2.5 Experimental Setup

The experimental setup used for the model identification, validation and real-time control is shown in Figure 2.8. In this setup a 10 cm long SMA wire with a diameter of 0.254 mm (Flexinol®, Dynalloy Inc., Irvine, CA, United States) was used. Electrical lead wires were attached to the ends of the SMA wire using screw clamps. One end of the wire was suspended on a wooden stand with the help of a custom made block. The other end was connected via a tether to a low friction slide to restrict the motion to the vertical direction. To isolate the SMA wire from the external environmental conditions, it was shielded with a cylindrical plastic tube. To measure the strain, a high resolution (1032 x 776) camera (Dragonfly®2 DR2-HICOL, Point Grey Research Inc., Richmond, BC, Canada) with an adjustable focal length lens (T4Z2813CS-IR, computar®, Cary, NC, US) was used to track the displacement of a marker attached to the front of the slide. The marker consists of the two black dots accurately printed 1 mm apart on a white paper. To increase the resolution of the measured displacement, the camera was placed as close as possible to the marker. To maximize the attainable frame rate, the camera was used in a custom (format_7) mode such that only the selected region of interest (1032 x 250) was read out. The slide was also attached to an extension spring that provides a recovering force to the SMA actuator as it cools. The tension of the spring was adjusted using a linear stage (TLSM050A, Zaber Technologies Inc., Vancouver, BC, Canada). A schematic diagram of the circuit to provide the current in the SMA actuator and to measure its resistance is shown in Figure 2.9. In this figure, a power amplifier (UPM 2405, Quanser Inc., Markham,

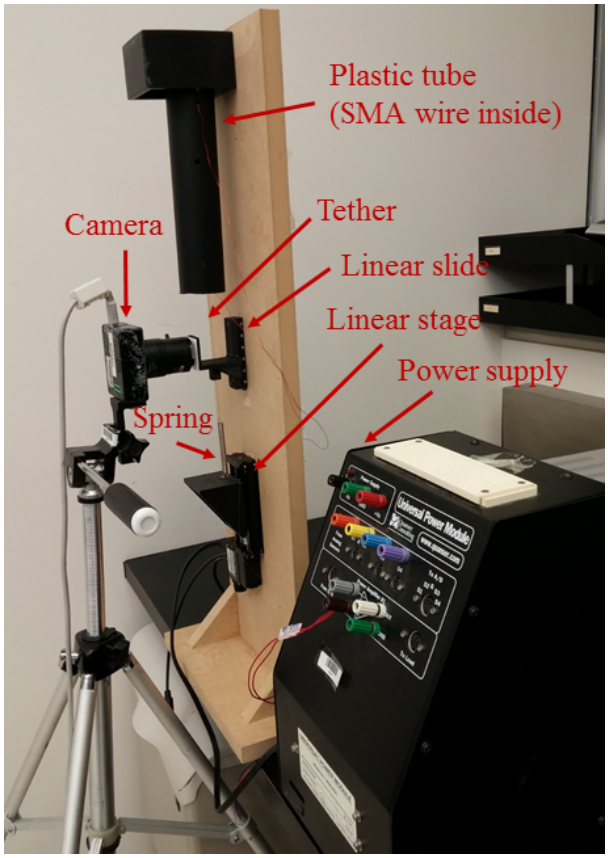

Figure 2.8: Experimental setup.

ON, Canada) was used in a voltage to current (V/I) converter configuration to provide the desired current from a digital to analog convertor of a data acquisition (DAQ) card (Sensoray 626, Sensoray Co. Inc., Chicago, IL, USA). The voltage across the SMA wire was read back through the analog to digital convertor of the DAQ card to measure the resistance of the wire using Ohm's law. A multithreaded application, using Microsoft® C++, MATLAB® and the QuaRC toolbox (Quanser Inc.), was developed to collect the camera and DAQ card readings and to provide the current control.

2.6 Strain Measurement Using Camera

To calculate the strain in the SMA wire, the displacement of the marker attached to the slide was measured and divided by the length of the wire. To measure the actual displacement of the marker (in mm) by measuring the change in the position of the marker in camera images (in pixels), the camera was first calibrated using OpenCV libraries [20] and the procedure described in [21]. In this method, the radial and tangential distortions present in the camera

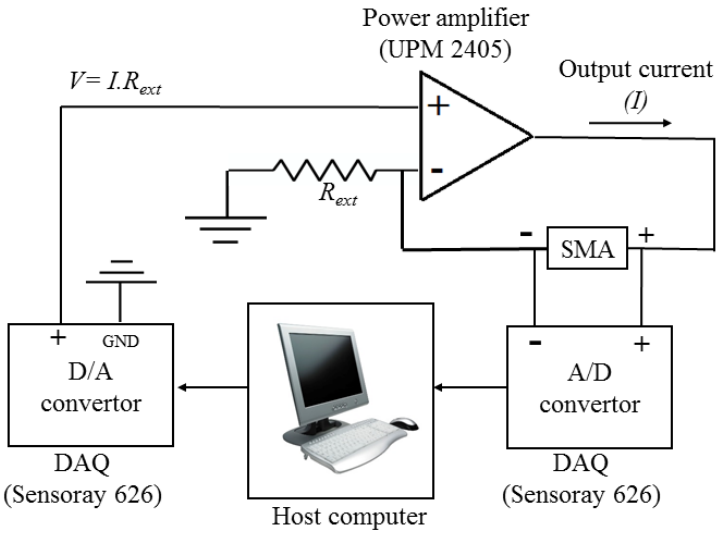

Figure 2.9: Schematic diagram of the circuit.

were modeled using following equations:

$$\begin{aligned} \hat{x} &= x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2\rho_1 xy + \rho_2(r^2 + 2x^2) \\ \hat{y} &= y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2\rho_2 xy + \rho_1(r^2 + 2y^2) \end{aligned} \tag{2.34}$$

where (x, y) is the coordinate of a point in the original (undistorted) image; $(\hat{x}, \hat{y})$ is the coordinate of the same point in the distorted image; $r = \sqrt{x^2 + y^2}$; k_1, k_2, k_3 are the radial distortion coefficients; and ρ_1, ρ_2 are the tangential distortion coefficients. The transformation of a point (X, Y, Z) of a 3D object in the object reference frame to the point (x, y) in its image in the camera reference frame can be described by following equation:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = s \underbrace{\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Intrinsic matrix}} \underbrace{\begin{bmatrix} R & \gamma \end{bmatrix}}_{\text{Extrinsic matrix}} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2.35}$$

where f_x and f_y are the focal length of the camera lens in the x and y direction of the camera reference frame; (c_x, c_y) are the coordinates of the principal point of the image; R and γ are a 3 x 3 rotation matrix and a 3 x 1 translation vector, respectively, that transform the coordinates of the points in the object reference frame to the camera reference frame.

Now, in order to estimate the intrinsic matrix, extrinsic matrix and the distortion coefficients, a planar 6 x 4 black-white chessboard pattern is captured with different nonparallel views. The corners of the chessboard pattern are detected for each of these views, and the unknown parameters are then computed using (2.34) and (2.35). It is important to mention

that the extrinsic matrix is different for different views whereas the intrinsic matrix and the distortion coefficients remain the same. Once all of the unknown parameters are identified, the captured images are undistorted using intrinsic matrix and distortion coefficients in real-time. (For more details, see [21–23]). The estimated intrinsic parameters and distortion coefficients are listed in Table 2.2 and Table 2.3 respectively.

Table 2.2: Camera Intrinsic Parameters

f_x	f_y	c_x	c_y
15.791 mm	15.782 mm	10.558 mm	1.863 mm

Table 2.3: Camera Distortion Coefficients

k_1	k_2	k_3	ρ_1	ρ_2
-0.26611	0.00868	0.08985	0.00027	0.00862

Now, to calculate the actual displacement of the marker (in mm), the camera was placed as close and as parallel as possible to the marker. After undistorting the image, the distance between two black dots in image pixels corresponds to 1 mm actual distance. Using this map, the displacement of the marker (in mm) is calculated by measuring the change in the position of one of the black dots in camera images (in pixels). To evaluate the accuracy of the measured displacement, the marker was attached to a linear stage. The linear stage was moved back and forth, and the displacement measured by the camera was compared with the actual change in the position of the linear stage as shown in Figure 2.10. It was found that the camera is able to accurately measure the actual displacement with an RMS error of 0.04 mm.

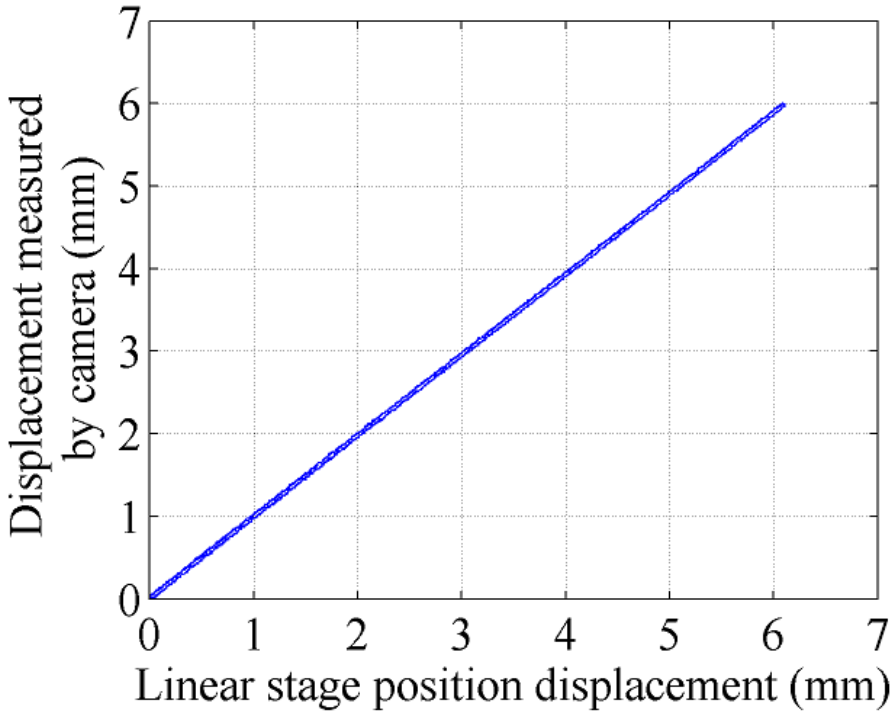

Figure 2.10: Validation of the displacement measurement using camera with respect to the linear stage position displacement

2.7 Model Identification and Validation

2.7.1 NPM Identification

Data for 13 FOD and 13 FOA curves were collected by applying the current at the rate of 5mA/s with minimum and maximum current of 150 mA and 500 mA respectively. Using (2.18), the temperature of the wire was simulated and a single-input, single-output feedforward ANN having one hidden layer was trained for each of these curves with temperature as an input and strain as an output. The hidden layer has two neurons which have a "log-sigmoid" activation function. The output layer neuron has a linear activation function. The network was then trained using the LM algorithm in MATLAB®. The weight parameters obtained after the training of the ANNs were stored in two different tables for FOD and FOA curves separately. Figure 2.11 shows good approximation of the experimentally obtained FOR curves with the ANN realized FOR curves.

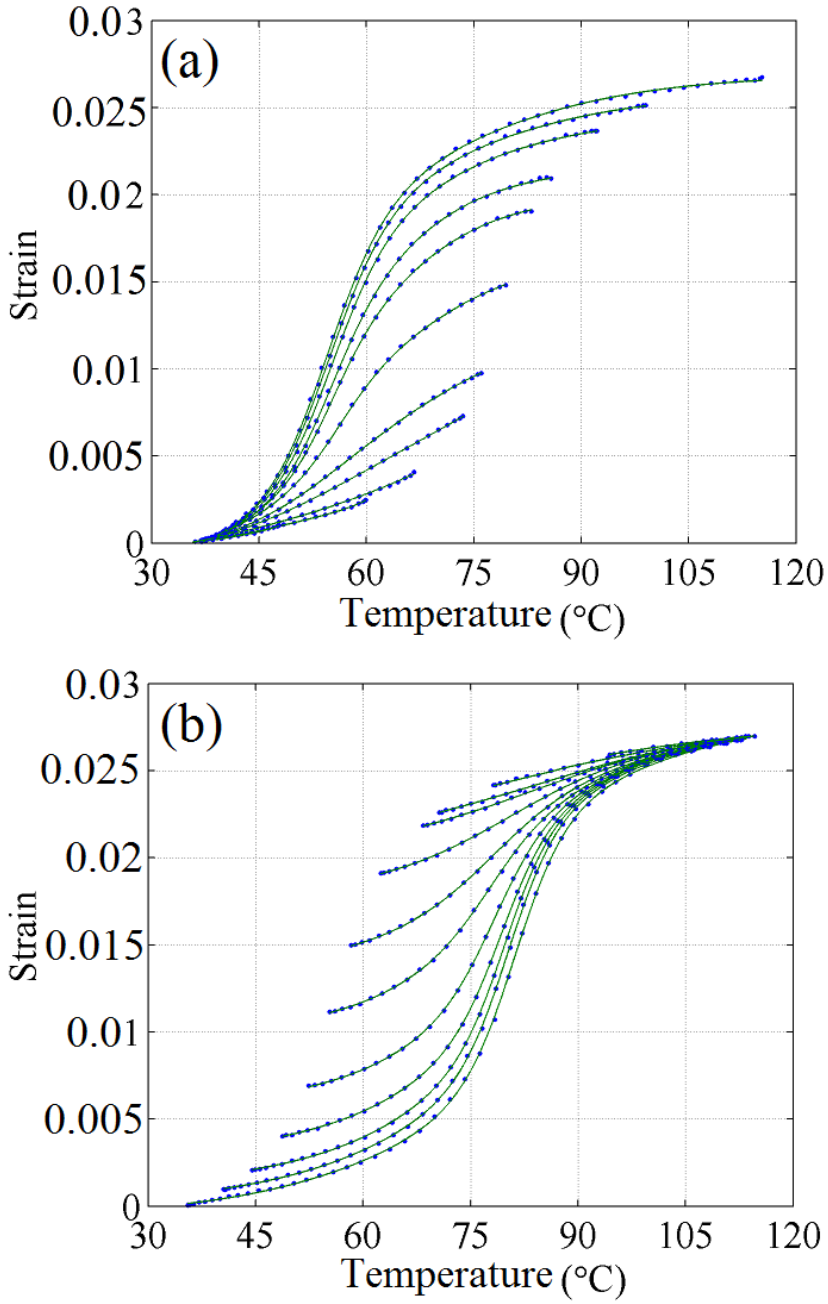

Figure 2.11: Experimental (dotted blue) and ANN approximated (solid green) (a) FOD and (b) FOA curves.

2.7.2 NPM and ANPM Validation

To evaluate the accuracy of the NPM and ANPM, random input currents with three different rates (5 mA/s, 10mA/s and 20 mA/s) were applied through the SMA wire (Figure 2.12). The covariance matrix of the process noise Q and the covarinace of the measurement noise R for

the EKF were chosen as:

$$Q = 0.1 * I_{14} \quad \text{and} \quad R = 0.05 \tag{2.36}$$

where I_{14} is a 14 x 14 identity matrix. The number of iterations (n), the EKF is applied to update the weight parameters was set 2.

The NPM predicted output when no feedback is available and ANPM predicted output when strain feedback is available was compared with the measured output as shown in Figure 2.13 and Figure 2.14, respectively. Although, NPM can accurately predict the output strain with an RMS error of 0.0085 strain, ANPM significantly improves the performance as is clear from Figure 2.14. The RMS error of ANPM predicted output is 0.0001 strain. The strain-temperature behavior of the ANPM and the measured data is also shown in Figure 2.15 which illustrates the hysteresis modeling performance of the proposed model.

To validate the ANPM capability to predict the output under varying environmental conditions, the plastic tube covering the SMA wire was removed that causes disturbances in the cooling rate due to uncontrollable air circulation in the environment. An increasing and decreasing the input current at a rate of 5 mA/s (Figure 2.15(a)) was then applied through the wire. Figure 2.15(b) shows how the disturbances in cooling rate affect the output strain; however, the ANPM is still able to predicted the output with good accuracy (0.00012 strain of RMS error).

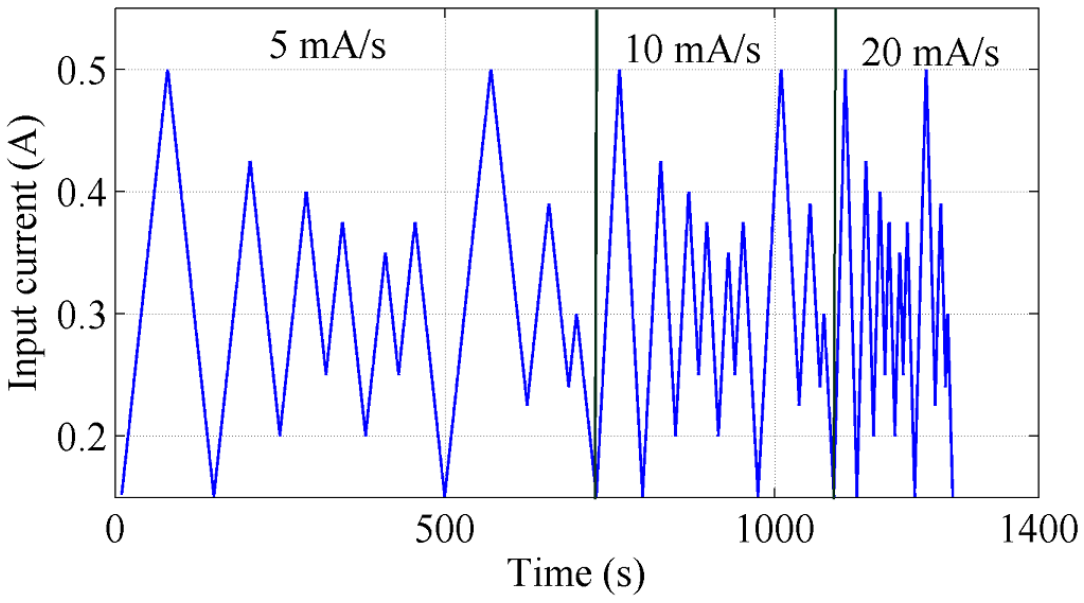

Figure 2.12: Input current for the validation of the NPM and the ANPM.

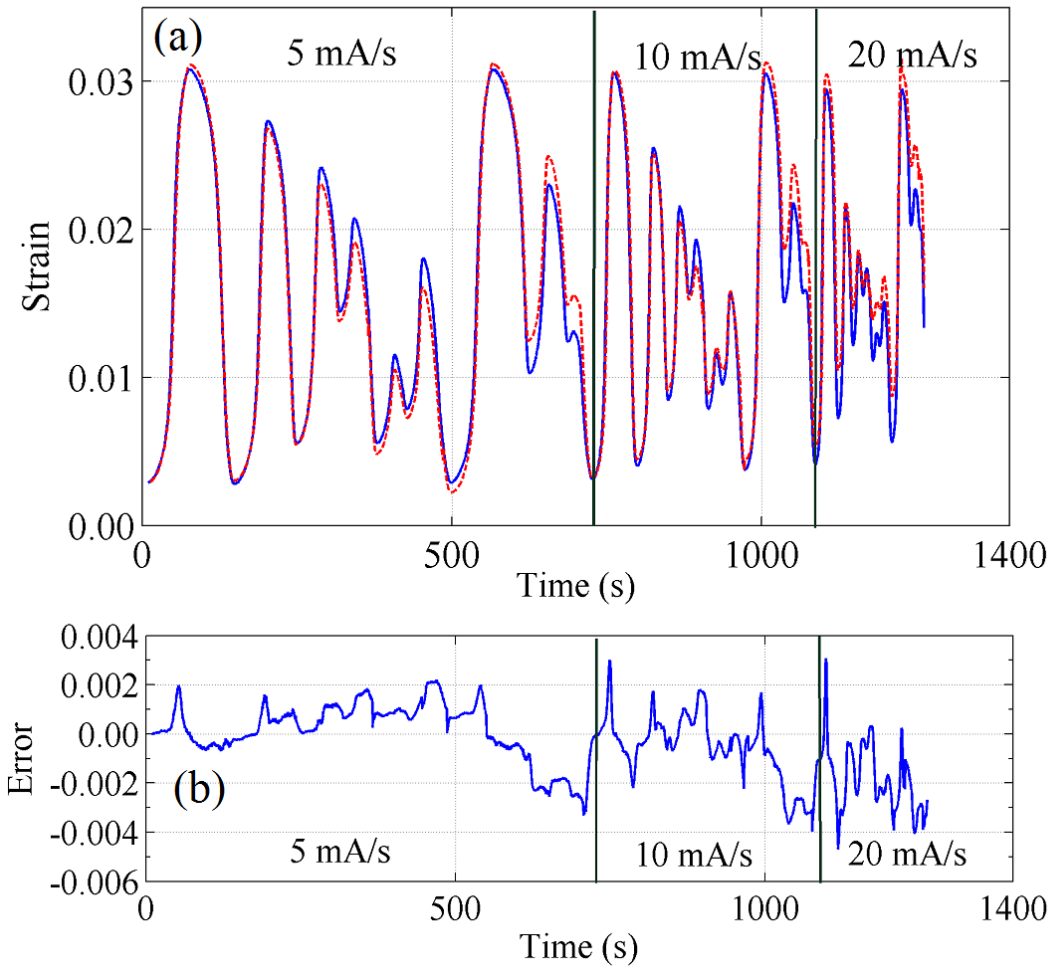

Figure 2.13: (a) Comparison of the NPM predicted strain (dashed red line) and the measured strain (solid blue line), (b) error in the predicted strain.

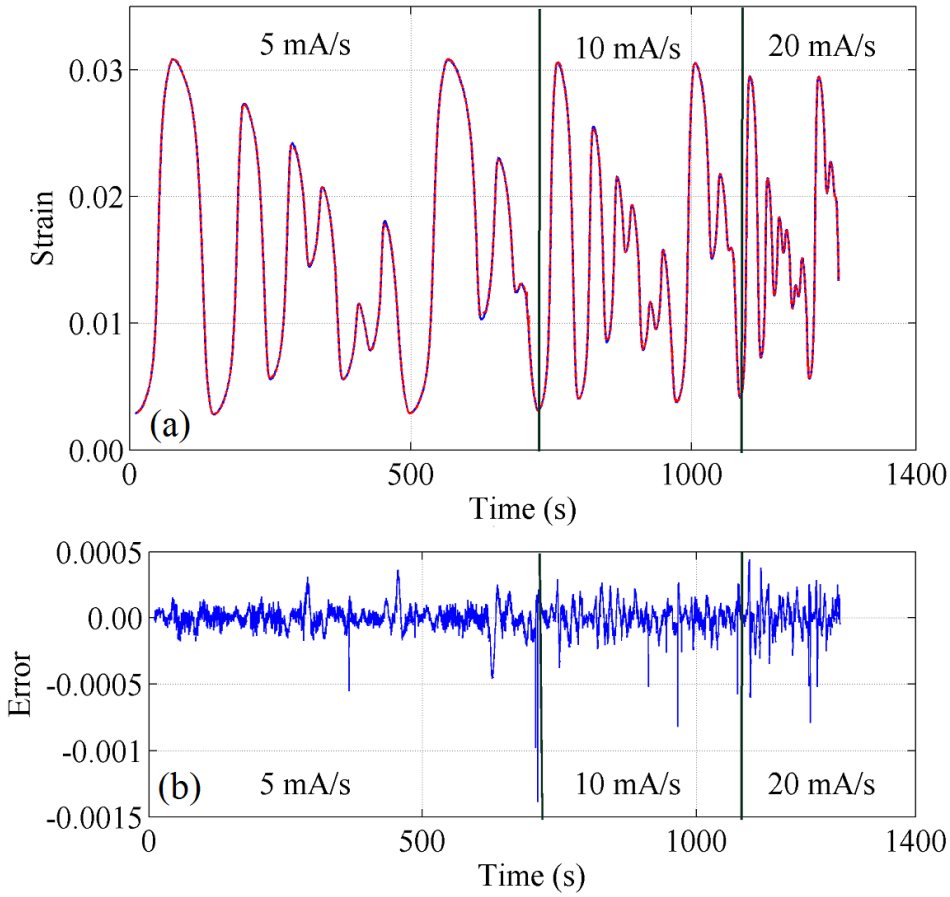

Figure 2.14: (a) Comparison of the ANPM predicted strain (dashed red line) and the measured strain (solid blue line), (b) error in the predicted strain.

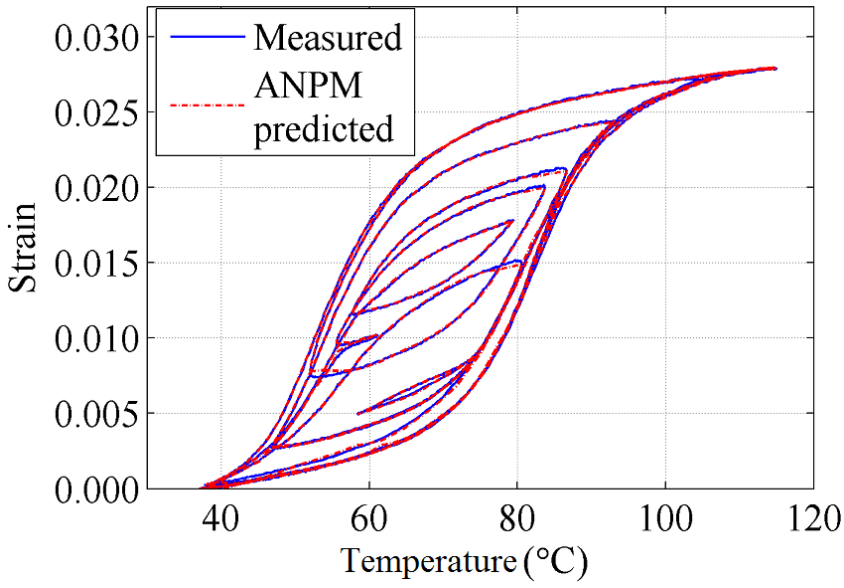

Figure 2.15: ANPM predicted strain with respect to the temperature in comparison to the experimental strain.

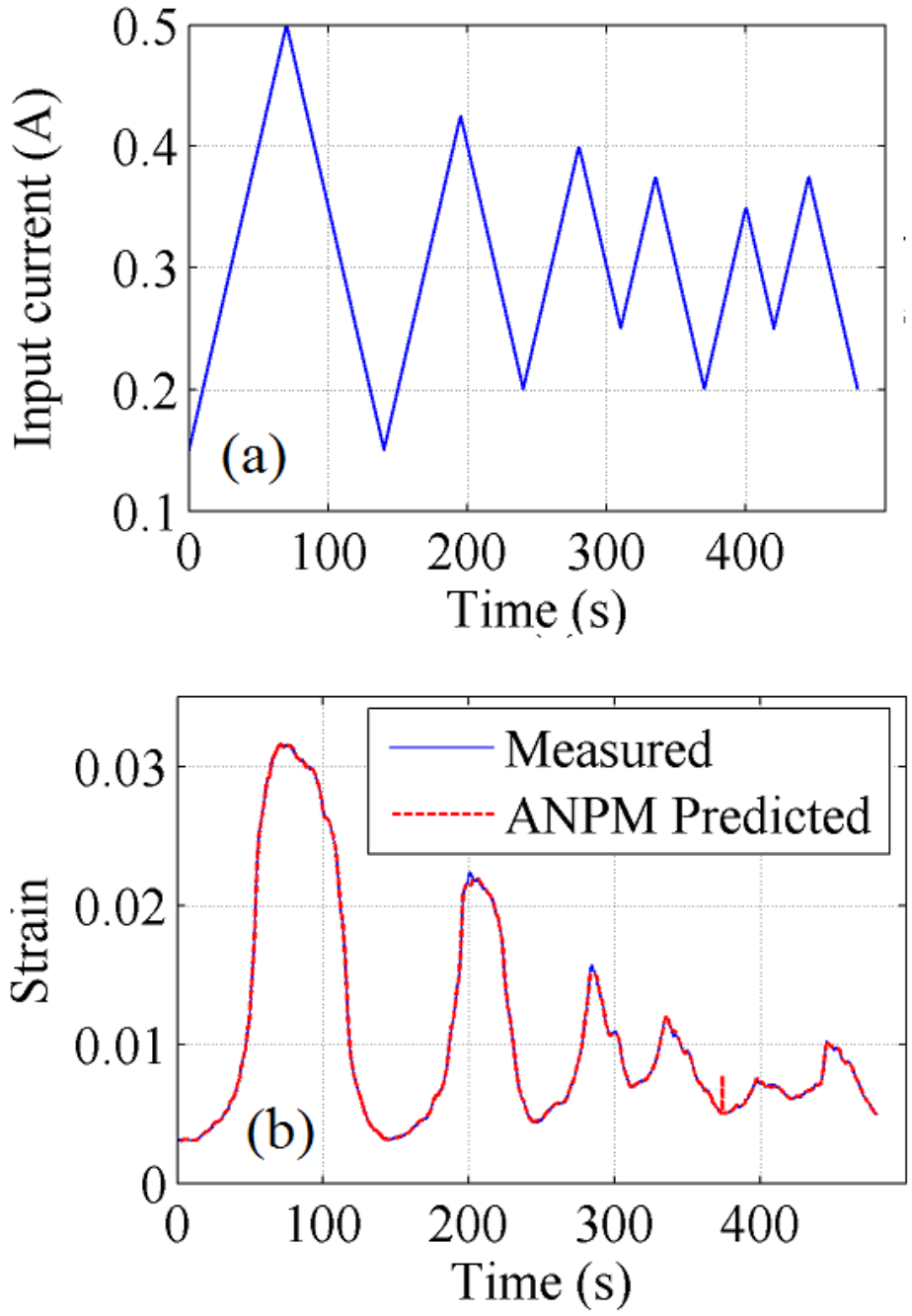

Figure 2.16: Open-loop response of the SMA actuator in the presence of uncertainties in cooling rate. (a) Applied input current, (b) measured and predicted strain.

2.8 Adaptive Neural Preisach Model Predictive Control

Model predictive control, also called as “receding horizon control”, is a type of adaptive control strategy which uses a plant’s dynamic model to predict its future behavior over a specified

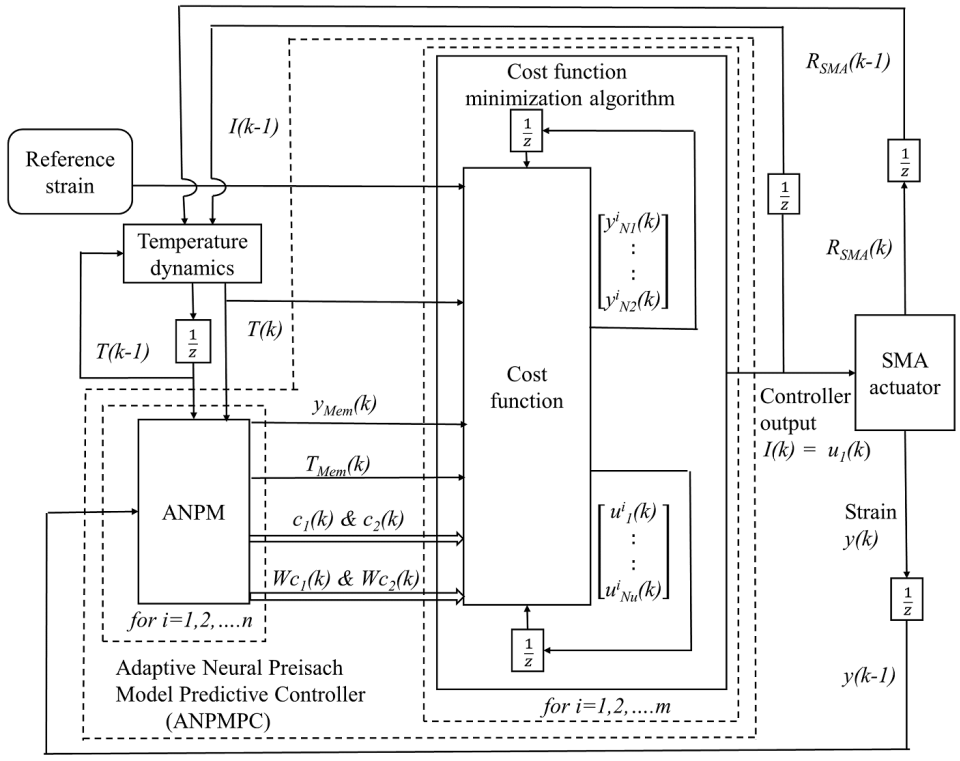

Figure 2.17: Block diagram of the ANPMPC.

time horizon, and then finds a control input that optimizes the plant's behavior over the same time horizon in order to achieve the desired output [24]. In [25], the angle of an SMA actuated robot arm with respect to the voltage across the SMA actuator was modeled using an ANN and a Neural model predictive controller was implemented to control the desired angle in simulations. In [26], it was shown that if a plant is subjected to parameter variations, the Neural model predictive control leads to a steady-state error in the output. To remedy this problem, a neural identification block is incorporated in the control scheme to update the weight parameters of the neural network model in real-time. A schematic of the Adaptive Neural Preisach Model Predictive Control (ANPMPC) used in this chapter is shown in Figure 2.17. The controller consists of two main blocks: an ANPM block that updates the weighting parameters of the NPM based on the predicted and the measured strain at previous sampling instant, and a cost function minimization block that determines the controller output by minimizing the cost function of the following form using a numerical optimization algorithm:

$$J = \sum_{j=N_1}^{N_2} [y_d(k) - y_p(k+j)]^2 + \sum_{j=0}^{N_u-1} [\lambda_j \Delta u(k+j)]^2 \tag{2.37}$$

$$\text{subject to the constraint,} \quad u_{min} \leqslant u(k+j) \leqslant u_{max}$$

where $\Delta u(k+j) = u(k+j) - u(k+j-1)$, λ_j is the weighting factor, N_1 is the minimum costing horizon, N_2 is the maximum cost horizon, N_u is the control horizon, $y_d(k)$ is the desired strain at a discrete time instant k, and $y_p(k+j)$ is the j-step ahead prediction of the strain which is calculated using the discrete-time state-space model ((2.38) and (2.39)) of the SMA actuator:

$$T(k+1) = T(k) + \frac{[u^2(k)R(k) - hA(T(k) - T_a)]T_s}{mc_p} \tag{2.38}$$

$$y_p(k) = c_1[y^{NN}_{T_1}(T(k), W_{c_1}) - y^{NN}_{T_1}(T_{Mem}, W_{c_1})] + c_2[y^{NN}_{T_2}(T(k), W_{c_2}) - y^{NN}_{T_2}(T_{Mem}, W_{c_2})] + y_{Mem} \tag{2.39}$$

where T_s is the sample time, and the values of c_1, c_2, W_{c_1}, W_{c_2}, T_{Mem} and y_{Mem} parameters in (2.39) are obtained from the ANPM block of the controller (Figure 2.17). In this chapter, the LM optimization algorithm is used to find the optimal control current that minimizes (2.37). At each sampling instant k, the algorithm iteratively updates the control input vector $U(k) = [u(k)\ u(k+1)\ ....\ u(k+N_u)]^T$ as follows,

$$U^{i+1}(k) = [(J^i(k))^T J^i(k) + \mu^i I]^{-1}(J^i(k))^T \boldsymbol{e}^i(k) + U^i(k) \tag{2.40}$$

$$\text{where} \quad \boldsymbol{e}^i(k) = \begin{bmatrix} y_d(k) - y^i_p(k+N_1) \\ \vdots \\ y_d(k) - y^i_p(k+N_2) \\ \Delta u^i(k) \\ \vdots \\ \Delta u^i(k+N_u) \end{bmatrix}$$

and $J^i(k)$ is the Jacobian matrix that contains the derivatives of $\boldsymbol{e}^i(k)$ with respect to the input vector $U^i(k)$ at the i^{th} iteration. μ^i is the LM parameter which is initialized to some constant μ_0. After every iteration, if the cost function is increased, it is multiplied by 10, and if the cost function is decreased, it is divided by 10. When the cost function reaches the minimum specified value, or when the maximum number of iterations (q) is reached, the first value of the control input vector is passed to the plant.

2.9 ANPMPC Simulations and Experiments

To investigate the performance of the proposed controller, simulations were first carried out in MATLAB®, then the experiments were performed using the setup described in section 2.5. The values of Q, R and n were chosen to be the same as in section 2.7.2. N_1, N_2, N_u, λ_j and μ_0 were chosen as 2, 3, 2, 0.025 and 10, respectively. The maximum number of iterations of the LM algorithm for searching the optimal control current was set to 6.

Figure 2.18 shows the simulation results of the sinusoidal reference input of 0.1 Hz at different DC offsets and Figure 2.19 shows the simulation results of the sinusoidal reference input of 0.05 Hz. Figure 2.20 shows the simulation results for different step reference inputs.

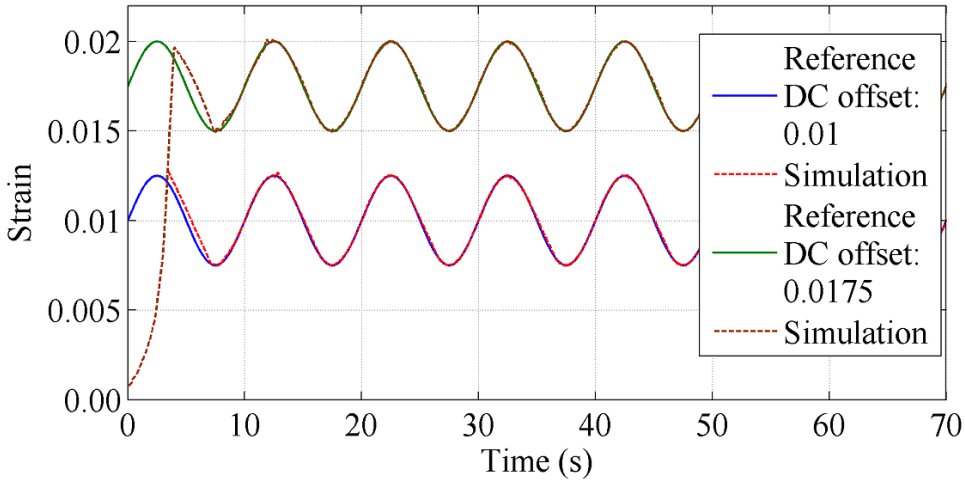

Figure 2.18: Simulation results: closed-loop control response to the 0.1 Hz sinusoidal reference input of different DC offset.

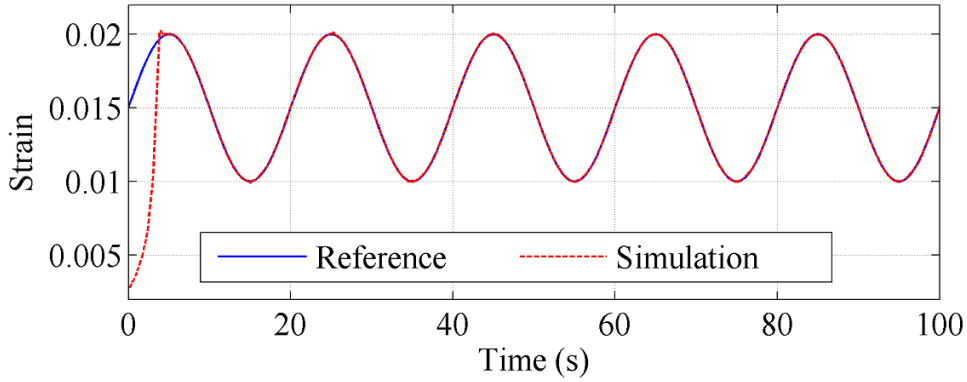

Figure 2.19: Simulation results: closed-loop control response to a 0.05 Hz sinusoidal reference input.

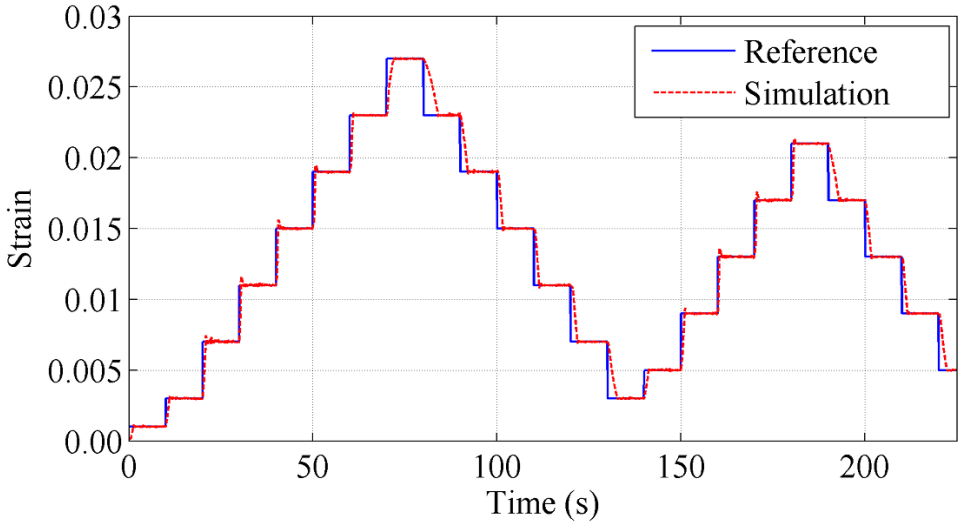

Figure 2.20: Simulation results: closed-loop control response to the step reference inputs.

Experimental verification of the ANPMPC was done at 30 ms sampling period. Figure 2.21 shows the closed-loop response of the SMA while tracking the step input reference. The figure shows acceptable performance of the controller with an RMS error of 0.0007 strain in the measured strain over the entire range of operation. The response of the SMA while tracking the sinusoidal reference input of 0.05 Hz is shown in Figure 2.22 and tracking the sinusoidal reference input of 0.1 Hz at different DC offsets is shown in Figure 2.23. To evaluate the performance of the controller under varying environmental conditions and in the presence of disturbances, the plastic tube covering the SMA wire was removed, and perturbations were introduced by changing the linear stage position. Removing the plastic tube covering the SMA wire causes disturbances in the cooling rate due to uncontrollable air circulation in the environment, and changing the linear stage position causes changes in the wire stress. Figure 2.24 shows the result of the closed-loop response to the sinusoidal reference input of 0.05 Hz frequency under these conditions. The controller is able to track the reference trajectory with an RMS error of 0.0012 strain.

2.10 Conclusion

This chapter presents a new control framework to control the strain in the SMA wire actuator using a model predictive control technique. For this purpose, strain versus current in the SMA actuator was modeled using a temperature dynamics and a modified Preisach model. In contrast to the classical Preisach model, an ANN based Preisach model is presented which utilizes both FOD and FOA curves. To model the uncertainties and parameter variations, the

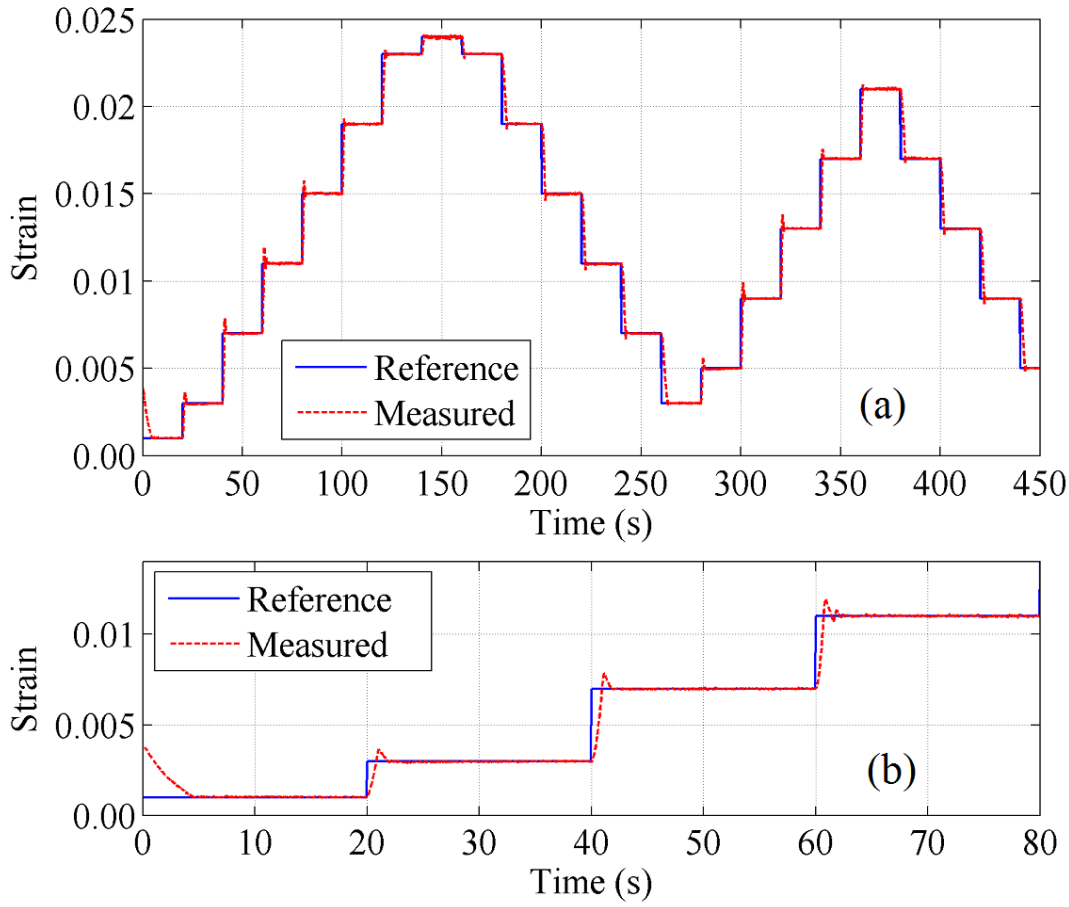

Figure 2.21: Experimental results: (a) closed-loop control response for step reference inputs (b) magnified view for the first 80 seconds.

EKF algorithm was used to update the weight parameters of the ANNs representing the appropriate FOR curves. Experiments were performed to evaluate the accuracy of the model predictions. Experiments were performed to investigate the performance of the controller that uses the temperature dynamics and the ANPM to predict the future behavior of the SMA, and the LM algorithm to optimize a cost function in order to achieve the desired strain. Results showed that the controller performed well while tracking various reference trajectories even in the presence of external disturbances. Although the proposed modeling method and control scheme is validated through simulations and experiments for the SMA actuator only, the same can be used for other hysteretic systems such as piezoelectric actuators, piezoceramic actuators and cable-driven instruments as well.

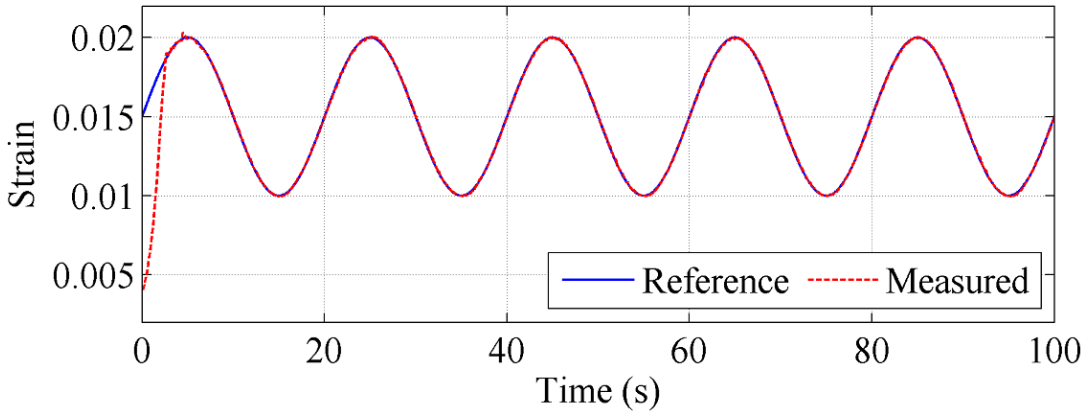

Figure 2.22: Experimental results: closed-loop control response for 0.05 Hz sinusoidal reference input.

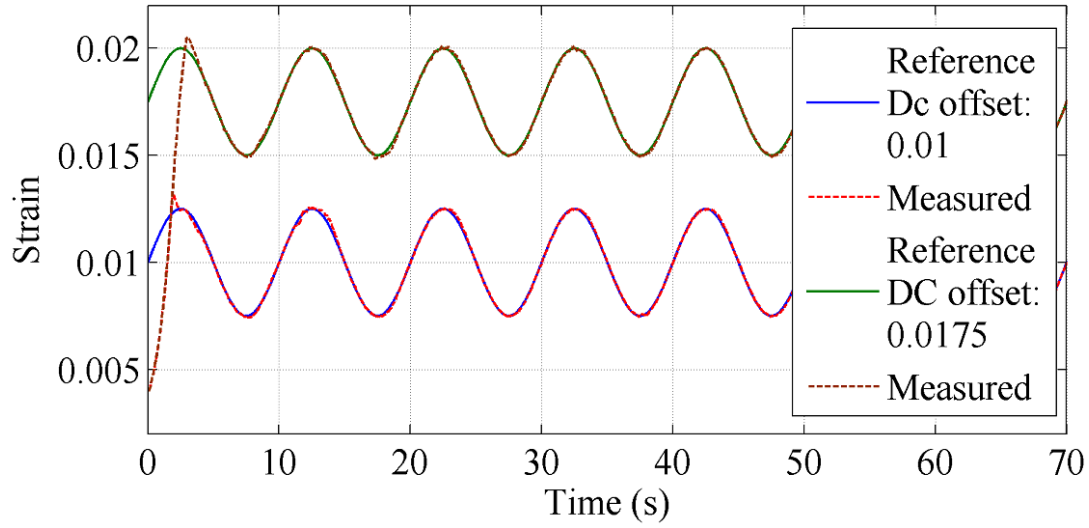

Figure 2.23: Experimental results: closed-loop control response for 0.1 Hz sinusoidal reference inputs at different DC offsets.

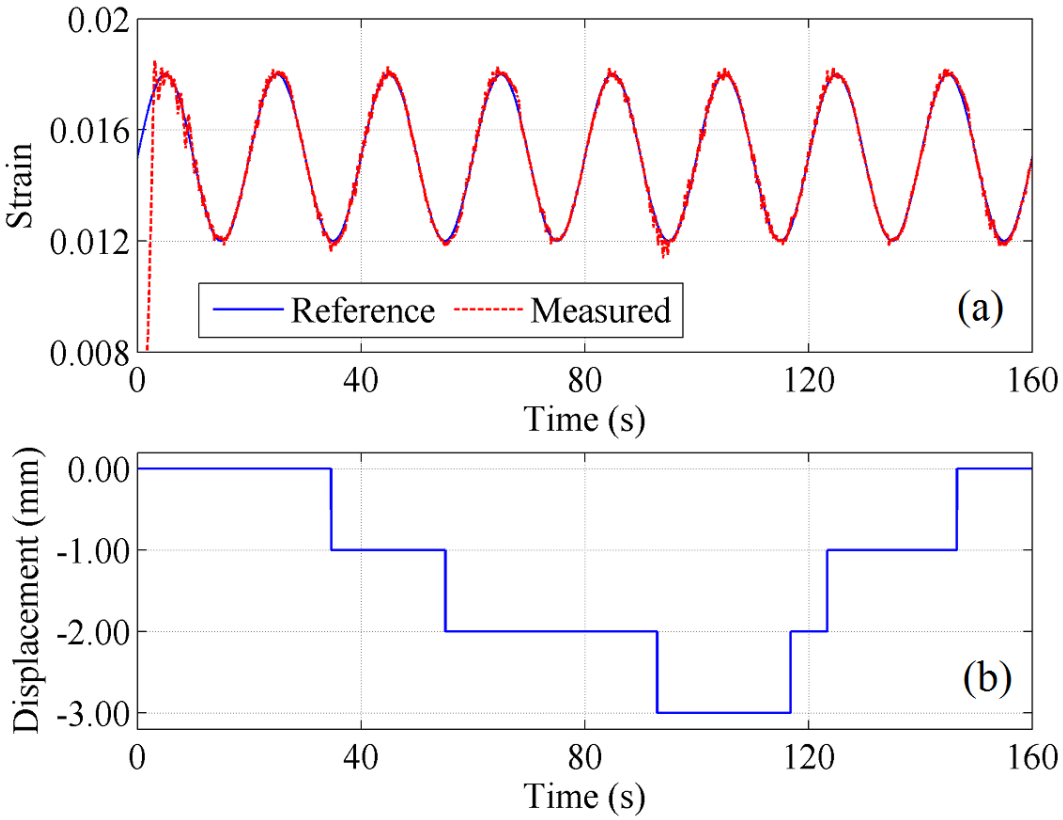

Figure 2.24: Experimental results: (a) Closed-loop response of the SMA to the sinusoidal reference input in the presence of cooling rate uncertainties and load disturbances, (b) displacement of the linear stage.

Bibliography

[1] D. J. Hartl and D. C. Lagoudas, "Aerospace applications of shape memory alloys," *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, vol. 221, no. 4, pp. 535–552, 2007.

[2] M. M. Ali and K. Takahata, "Frequency-controlled wireless shape-memory-alloy microactuators integrated using an electroplating bonding process," *Sensors and Actuators A: Physical*, vol. 163, no. 1, pp. 363 – 372, 2010.

[3] T. Tao, Y.-C. Liang, and M. Taya, "Bio-inspired actuating system for swimming using shape memory alloy composites," *International Journal of Automation and Computing*, vol. 3, no. 4, pp. 366–373, 2006.

[4] Y. Haga, Y. Tanahashi, and M. Esashi, "Small diameter active catheter using shape memory alloy," in *The Eleventh Annual International Workshop on Micro Electro Mechanical Systems, 1998. (MEMS 98)*, Jan 1998, pp. 419–424.

[5] Y. Nakamura, A. Matsui, T. Saito, and K. Yoshimoto, "Shape-memory-alloy active forceps for laparoscopic surgery," in *Proceedings of IEEE International Conference on Robotics and Automation, 1995.*, vol. 3, May 1995, pp. 2320–2327 vol.3.

[6] J. Jayender, R. Patel, and S. Nikumb, "Robot-assisted active catheter insertion: Algorithms and experiments," *The International Journal of Robotics Research*, vol. 28, no. 9, pp. 1101–1117, may 2009.

[7] K. Arai, S. Aramaki, and K. Yanagisawa, "Continuous system modeling of shape memory alloy (sma) for control analysis," in *Proceedings of the 5th International Symposium on Micro Machine and Human Science, 1994.* IEEE, 1994, p. 97.

[8] J. Jayender, R. V. Patel, S. Nikumb, and M. Ostojic, "Modelling and gain scheduled control of shape memory alloy actuators," in *2005 IEEE International Conference on Control Applications (CCA)*, Aug 2005, pp. 767–772.

[9] D. Hughes and J. T. Wen, "Preisach modeling of piezoceramic and shape memory alloy hysteresis," in *Proceedings of the 4th IEEE International Conference on Control Applications, 1995.*, Sep 1995, pp. 1086–1091.

[10] I. Mayergoyz, *Mathematical Model of Hysteresis And Their Applications*. Elsevier Science Inc., 2003.

[11] F. Auricchio and L. Petrini, "A three-dimensional model describing stress-temperature induced solid phase transformations: solution algorithm and boundary value problems," *International Journal for Numerical Methods in Engineering*, vol. 61, no. 6, pp. 807–836, Sep 2004.

[12] Z. Li, C. Y. Su, and T. Chai, "Compensation of hysteresis nonlinearity in magnetostrictive actuators with inverse multiplicative structure for preisach model," *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 2, pp. 613–619, April 2014.

[13] F. Anooshahpour, I. G. Polushin, and R. V. Patel, "Classical preisach model of hysteretic behavior in a da vinci instrument," *IEEE International Conference on Advanced Intelligent Mechatronics*, 2016.

[14] K. K. Ahn and N. B. Kha, "Improvement of the performance of hysteresis compensation in sma actuators by using inverse preisach model in closed loop control system," *Journal of Mechanical Science and Technology*, vol. 20, no. 5, pp. 634–642, 2006.

[15] V. T. Liu, W. C. Yang, and H. Y. Wing, "Realization of preisach model using adaptive polynomial approximation method," in *5th IEEE Conference on Industrial Electronics and Applications*, June 2010, pp. 1258–1263.

[16] M. R. Zakerzadeh, M. Firouzi, H. Sayyaadi, and S. B. Shouraki, "Hysteresis nonlinearity identification using new preisach model-based artificial neural network approach," *Journal of Applied Mathematics*, vol. 2011, pp. 1–22, 2011.

[17] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Prentice-Hall Inc., 1999.

[18] Y. Iiguni, H. Sakai, and H. Tokumaru, "A real-time learning algorithm for a multilayered neural network based on the extended kalman filter," *IEEE Transactions on Signal Processing*, vol. 40, no. 4, pp. 959–966, Apr 1992.

[19] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB*. Willy-IEEE Press, 2008.

[20] G. Bradski, "The opencv library," *Dr. Dobb's Journal of Software Tools*, 2000.

[21] G. Bradski and A. Kaehler, *Learning OpenCV*, 1st ed. OReilly Media, Inc., 2008.

[22] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proceedings of the 7th IEEE International Conference on Computer Vision*, 1999, pp. 666–673.

[23] ——, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[24] E. F. Camacho and C. Bordons, *Model Predictive Control*, 2nd ed. Springer-Verlag, 2004.

[25] N. Nikdel, P. Nikdel, M. A. Badamchizadeh, and I. Hassanzadeh, "Using neural network model predictive control for controlling shape memory alloy-based manipulator," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 3, pp. 1394–1401, March 2014.

[26] R. Hedjar, "Adaptive neural network model predictive control," *International Journal of Innovative Computing, Information and Control*, vol. 9, no. 3, pp. 1245–1257, 2013.

Chapter 3

Design of an Ultra Thin Superelastic Shape Memory Alloy Strain Sensor

3.1 Introduction

Force information plays a significant role in surgical procedures, for example, when performing tasks such as blunt dissection [1] and suturing [2, 3], and in general to prevent damage to tissue as a result of applying excessive force. This information is available to surgeons during open surgery as a result of their direct access to tissue. However, in minimally invasive surgery (MIS) which is performed with long slender instruments, either manually or with robotic assistance, this force information is very inaccurate (for manual MIS) or is not available (for current robotic MIS systems). Therefore for MIS, alternative means of obtaining the force of tool-tissue interaction is necessary. One approach is the sensorization of surgical tools with force/torque sensors. The resulting force information can then be presented to surgeons directly via a haptic interface or visually on a monitor as a force value.

Strain gauges and fibre optic sensors (such as Fiber Bragg Grating (FBG) sensors) have been used in recent research work to measure tool-tissue interaction forces in minimally invasive surgical tools [4]. Trejos *et al.* [5] sensorized a custom laparoscopic instrument using Constantan foil strain gauges to measure tool-tissue interaction forces and torques in 5 degrees of freedom (DOF). These strain gauges have a gauge factor (GF) of 2 and measure strain up to 3%; however, their installation is complex and they require frequent re-calibration. Recently Hammond *et al.* [6] developed a method to directly print strain gauges on commercially available medical devices. Their method does not require precise surface treatments that are required for conventional foil strain gauges; however the GF of this sensor was reported to be 1.16, which is not as high as commonly used metallic foil strain gauges, limiting its resolu-

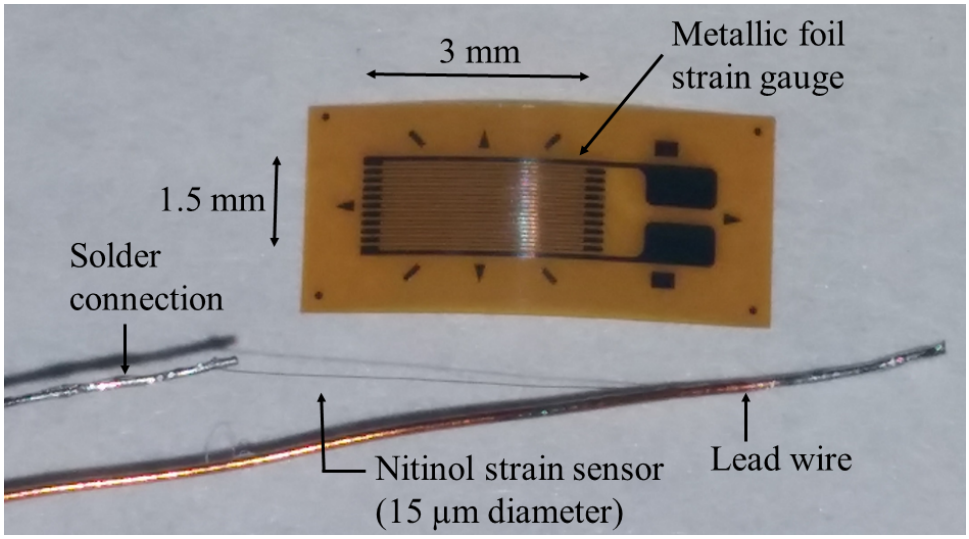

Figure 3.1: Nitinol strain sensor vs commercial metallic foil strain gauge.

tion and accuracy. In [7], 0.2 mm wide silicon strain gauges were used for the design of a microsurgical instrument tip force sensor for measuring contact forces during retinal surgery. Silicon strain gauges show much higher GF (50-200) than metallic foil gauges, but due to the very low measurement range (<0.5%), these strain gauges work better for applications where lower strains have to be measured and high sensitivity is required such as in [7]. FBG sensors are a better alternative to strain gauges in terms of sensitivity, magnetic resonance compatibility and low space requirement. These sensors were used in [8] to sensorize an arthroscopic tool, in [9] to measure lateral forces at the tip of a da Vinci surgical instrument, in [10] for force, torsion and shape sensing of a continuum robot, and in [11] to measure contact forces of an ablation catheter during radiofrequency ablation. However, the process of FBG inscription greatly increases the manufacturing cost of these sensors which hinders their use in commercial applications such as for da Vinci surgical tools, ablation catheters and surgical needles where the instruments are disposable or limited in the number of uses.

Strain sensors can also be made using SMAs. Schoor *et al.* [12] have a patent for an SMA strain gauge to measure large strains in fabrics (up to 20%). These gauges are relatively large and intended for high force applications (seatbelts and parachute canopies). In [13], Tung *et al.* fabricated a laser machined SMA sensor for position feedback in active catheters. Their sensor has a GF of 6.9 and can be elongated up to 30%; however, this sensor has only 1.5Ω of electrical resistance, which limits the maximum allowable voltage across the sensor in order to avoid thermal drift in the sensor output, thus, reducing the SNR of the measurements.

This chapter presents a novel-high performance ultra thin strain sensor (15 μm diameter) that uses a single superelastic Nitinol wire (Figure 3.1), and is suitable for use in a variety of

surgical instruments. In comparison to existing metallic foil strain gauges, the proposed sensor has higher sensitivity (GF = 3.56), larger measurement range (4.25%) and higher flexibility which make this sensor applicable to both rigid surgical instruments (such as laparoscopic tools) and flexible surgical instruments (such as catheters and surgical needles). As opposed to the SMA sensor, used in [13], which has a complex geometric structure, the simplicity of the sensor, described in this chapter, significantly reduces its manufacturing and installation cost. Therefore, this sensor has the advantage that it can be easily and cost effectively used in disposable surgical instruments where the use of expensive sensors such as FBGs may be difficult to justify. In this chapter, a da Vinci surgical instrument was sensorized using 8 of these sensors to demonstrate their force sensing capability in MIS. The design and validation of this technology is presented in the following sections.

3.2 Nitinol Sensor Design

3.2.1 Nickel-Titanium Alloy Properties

Nitinol is a Nickel-Titanium alloy that exhibits two unique properties: shape memory effect, and pseudoelasticy/superelasticy. Nitinol's shape memory characteristic allows deformation at a low temperature to be recovered when heated above its transition temperature. Above this transition temperature, Nitinol also exhibits pseudoelastic characteristics; allowing it to regain its original shape after being deformed with strains as high as 6-7%. Since its discovery in 1962, these characteristics have led to Nitinol being used extensively in many biomedical applications, such as active catheters, vascular stents and orthodontic wires [14].

3.2.2 Nitinol as a Strain Sensor

Two important characteristics of a strain sensor are its stretchability and sensitivity. These two characteristics for Nitinol are discussed below.

1) Stretchability: The Young's modulus for Nitinol is around 75 GPa, which when compared to the most common strain gauge material, Constantan, with a Young's modulus of 160 GPa suggests that Nitinol is flexible enough for use as a strain gauge. In addition, the pseudoelastic property of Nitinol enlarges the region of recoverable strain which makes Nitinol better than conventional strain gauges for those applications requiring measurement of higher strains.

2) Sensitivity: The sensitivity of a strain sensor is defined in terms of GF which is the ratio of the normalized change in resistance to strain in the sensor. If a Nitinol wire has length l and

cross-sectional area A, then the resistance R of wire is defined as

$$R = \frac{\rho l}{A} \tag{3.1}$$

where ρ is the resistivity of the wire. By taking derivatives on both sides (3.1) with respect to length, we can write ([15])

$$\frac{dR}{R} = \frac{d\rho}{\rho} + (1 + 2\nu)\epsilon \tag{3.2}$$

where ν is the Poisson's ratio and ϵ is strain in the wire. The Poisson's ratio of Nitinol is 0.33, so the GF of Nitinol is defined as

$$\text{GF} = \frac{dR/R}{\epsilon} = \frac{d\rho/\rho}{\epsilon} + 1.66 \tag{3.3}$$

Based on this, the GF also depends on the normalized change in resistivity with respect to strain, also known as the piezoresistive effect. It was found experimentally that the GF of Nitinol was 3.5, indicating that the resistivity of Nitinol increases when strain is increased. Hence, in terms of sensitivity, Nitinol performs better than most conventional strain gauges.

3.2.3 Sensor Selection Criteria

The most widely used method for measuring the output of strain gauges is using a Wheatstone bridge. The change in the output of a Wheatstone bridge with respect to strain is directly proportional to the excitation voltage (3.6). Therefore, the excitation voltage should be kept as high as possible to increase the signal-to-noise ratio (SNR) of the sensor; however, in practical applications the maximum excitation voltage is limited because Joule heating causes the sensor temperature to increase, resulting in a change in resistance and generating a drift in the sensor output. To reduce the Joule heating and increase the allowable excitation voltage, the resistance of the sensor should be maximized by selecting a long wire with a smaller diameter (3.1). Therefore, a superelastic Nitinol wire with a diameter of 15 μm (the thinnest currently available on the market) was used in this work.

3.2.4 Lead Wire Connection

The Nitinol wire is commonly connected to electrical lead wires using crimps, laser welding or soldering. Crimping requires an additional part which is not desirable in medical devices because it increases the size of the sensor. Laser welding can produce a clean connection, but in our experience, this is very difficult to achieve with Nitinol wire less than 100 μm in diameter.

Soldering Nitinol wires can also be challenging because the thin layer of titanium oxide on the surface of Nitinol makes it difficult to wet the surface with solder; however, good results can be obtained using a combination of solder (96.5 Sn/3.5 Ag) and flux (Flux #2, Indium, Chicago, IL, USA) as shown in [16].

3.2.5 Signal Conditioning

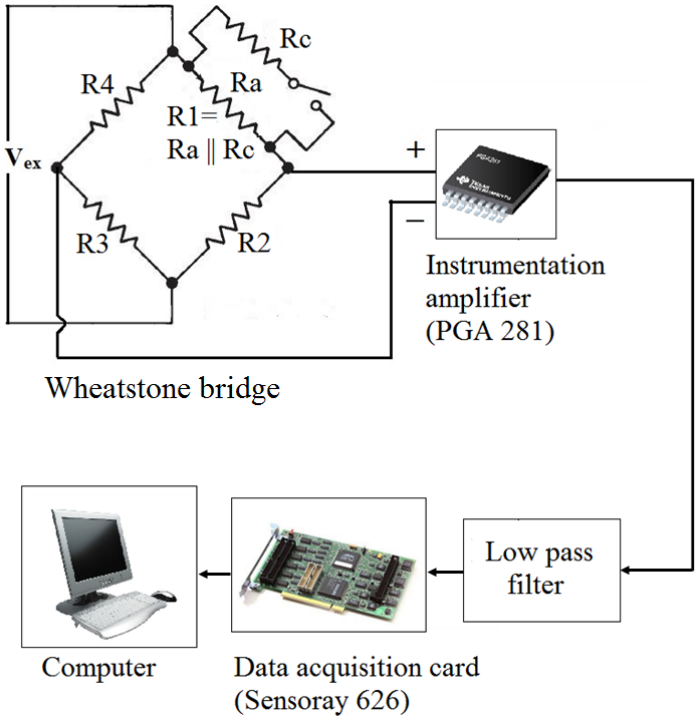

Figure 3.2: Overview of the sensor data flow.

A Wheatstone bridge is used to measure the voltage output of the Nitinol sensor (Figure 3.2). In order to balance the bridge, a trimmer potentiometer (*Rc*) is shunted with one of the fixed leg (*Ra*) of the Wheatstone bridge. The differential output of the Wheatstone bridge is then amplified by a programmable instrumentation amplifier (PGA 281, Texas Instruments, Dallas, TX, United States) with a gain (*G*) set to 176. To remove high frequency noise, the signal is passed through an RC low-pass filter with a 60 Hz cut-off frequency. Finally, the output signal is sent to the computer via an analog-to-digital converter of a data acquisition board (Sensoray 626, Sensoray Co. Inc., Chicago, IL, USA) at 250 Hz sampling rate.

3.3 Force Measurement for Surgical Instruments using Nitinol Sensors

To measure the lateral forces in the x and y directions acting at the distal end of a da Vinci instrument, Nitinol sensors were installed on the surface of the instrument shaft. The accuracy and SNR of the force measurements mainly depend on where the sensor is placed on the shaft. Therefore a model that relates the force at the distal end of the instrument to the sensor output is necessary to optimize the location of the sensors on the shaft. In the work described in this chapter, we have modeled the da Vinci shaft as a simply supported beam to find the strain on the surface (Figure 3.3). For the force (F_y) acting in the y direction at point P, the bending strain at some point B on the surface of the shaft can be written as

$$\epsilon_y = -\frac{F_y y r}{EI} \tag{3.4}$$

where r is the outer radius, y is the distance of point P from point B and EI is the flexural rigidity of the da Vinci shaft.

The two-DOF movement of the instrument's wrist causes change in the distance between point P, where the force is acting, and point B where the strain sensor is placed. To calculate the forces, invariant of the distance y, Blumenkranz *et al.* [17] proposed a method for measuring strains at two cross-sections. Therefore, we installed two aligned sets of Nitinol sensors at two cross-sections of the instrument shaft (Figure 3.3). Each cross-section has 4 Nitinol sensors, equally spaced around the circumference and parallel to the axis of the shaft. When the two cross-sections are separated by a distance l_y, F_y can be written as

$$F_y = \frac{(\epsilon_{1y} - \epsilon_{2y})(EI)}{r l_y} \tag{3.5}$$

where ϵ_{1y} and ϵ_{2y} are strain measured by sensors S^1_{y1} and S^2_{y1} respectively.

Sensors on the opposite sides of the shaft were connected in half Wheatstone bridge configurations. If the bridge is excited with a voltage V_{ex} and the strain (ϵ) experienced by two sensors is equal in magnitude, then the amplified output (V_{out}) of the half bridge can be written as

$$V_{out} = \frac{V_{ex} G F \epsilon}{1 + R_1/R_2} G \text{ or } \epsilon = \frac{V_{out}\,[1 + R_1/R_2]}{V_{ex} G F G} \tag{3.6}$$

Now, suppose that δ_V is the measurement noise in V_{out} and δ_ϵ is the noise in the calculated

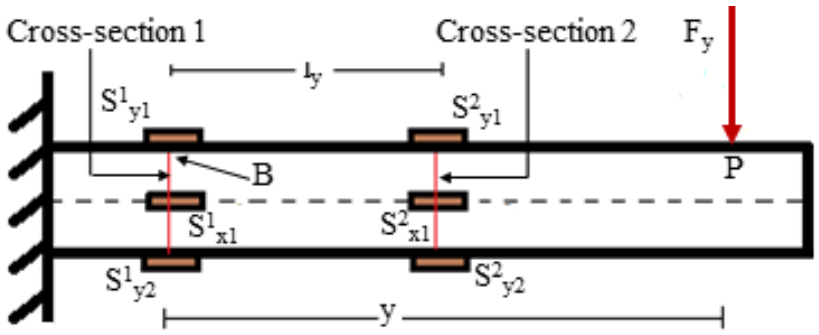

Figure 3.3: Cantilever model of a da Vinci shaft integrated with strain sensors at two cross-sections. Sensors opposite to S^1_{x1} and S^2_{x1} are not shown.

strain, such that

$$V_{measured} = V_{out} + \delta_V \text{ and } \epsilon_{measured} = \epsilon + \delta_\epsilon \tag{3.7}$$

then from (3.6),

$$\delta_\epsilon = \frac{\delta_V[1 + R_1/R_2]}{V_{ex}GFG} \tag{3.8}$$

and the force measured can be written as below.

$$F_{meassured} = \underbrace{\frac{(\epsilon_{1y} - \epsilon_{2y})(EI)}{rl_y}}_{\text{Signal}} + \underbrace{\frac{(\delta_{\epsilon 1y} - \delta_{\epsilon 2y})(EI)}{rl_y}}_{\text{Noise}} \tag{3.9}$$

From (3.9), the SNR would be

$$\text{SNR} = \frac{P_{Signal}}{P_{Noise}} = \frac{(\epsilon_{1y} - \epsilon_{2y})^2}{(\delta_{\epsilon 1y})^2 + (\delta_{\epsilon 2y})^2} \tag{3.10}$$

The denominator of (3.10) is independent of the sensor position. Hence, in order to maximize SNR, the numerator of (3.10) should be maximized, i.e., ϵ_{y1} should be large and ϵ_{y2} should be small. This suggests that the cross-section 1 should be selected as far as possible from the tip and cross-section 2 should be selected as close as possible to the tip. For this work, we have placed cross-section 1 at a distance of 80 mm from the instrument tip; however, to avoid frictional forces and torques created by the trocar and abdominal wall, this distance should be modified according to the application to ensure that cross-section 1 is always inside the body during use. Additionally, cross-section 2 should be set back from the instrument tip such that there are no forces acting on the instrument between cross-sections 1 and 2 (i.e. tool-tissue interaction should be distal to cross-section 2). For this work, we have placed cross-section 2

at a distance of 20 mm from the instrument tip.

3.4 Experiments

3.4.1 Gauge Factor Calculation

The GF was found experimentally using the setup shown in Figure 3.4. A 15 μm diameter Nitinol wire of length 40 mm was fixed at both ends using custom made clamps. The wire was then repeatedly stretched to a strain of 4.25% and relaxed using a linear stage (TLSM050A, Zaber Technologies Inc, Vancouver, BC, Canada) at a speed of 40 μm/s for 5 cycles. The resistance of the wire was computed continuously from the output of a quarter Wheatstone bridge.

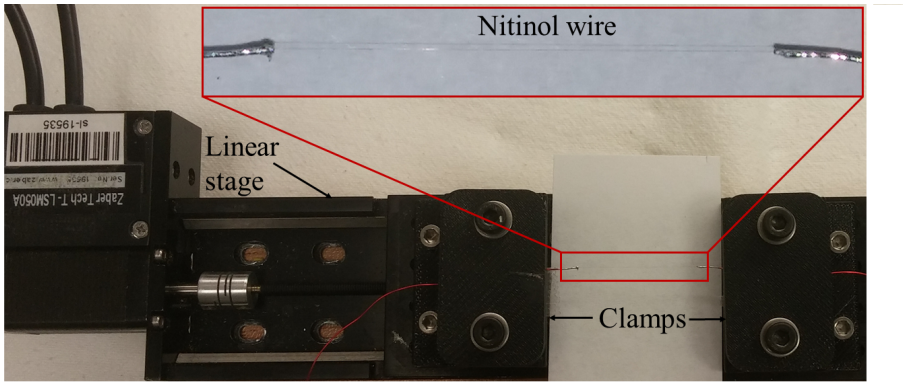

Figure 3.4: Experimental setup for calculating the gauge factor of Nitinol wire.

3.4.2 Sensor Calibration

Figure 3.5(a) shows the experimental setup used for calibrating the Nitinol sensors installed on the da Vinci instrument. In the setup, a force/torque (F/T) sensor (Nano-43, ATI Industrial Automation, Apex, NC, USA) was mounted on the distal end of the tool shaft using a custom made sensor interface. Nitinol sensor pairs at cross-section 1 were assumed to be perfectly aligned with sensor pairs at cross-section 2, while the pair S^1_{x1}, S^1_{x2} are at an angle of θ_x with the x-axis and the pair S^1_{y1}, S^1_{y2} at an angle of θ_y with the y-axis of the F/T sensor as shown in Figure 3.5(b). A transformation matrix (T_{2x4}) that relates voltage values from 4 half Wheatstone bridges to force values in the x and y direction can be written as (3.11), using (3.5), (3.6) and

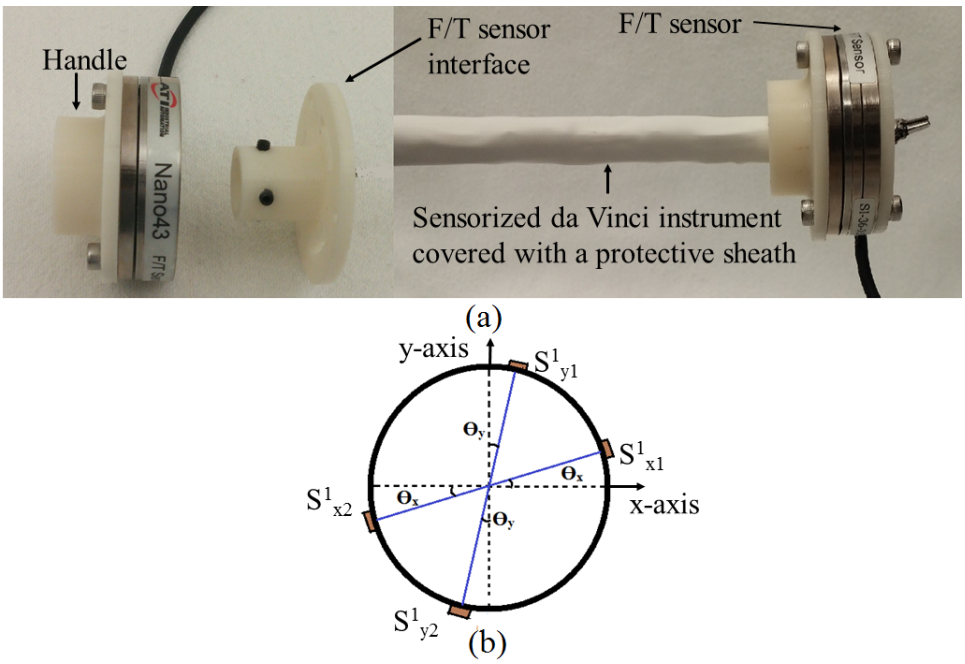

Figure 3.5: (a) Left: A F/T sensor with a handle and a sensor interface, right: F/T sensor mounted on a sensorized da Vinci instrument. (b) Cross-section of the sensorized da Vinci instrument shaft, showing Nitinol sensors orientation with respect to x-y axes of F/T sensor.

rotational transformations.

$$\begin{bmatrix} F_x \\ F_y \end{bmatrix} = T \begin{bmatrix} \Delta V_{x1} \\ \Delta V_{x2} \\ \Delta V_{y1} \\ \Delta V_{y2} \end{bmatrix} \tag{3.11}$$

where ΔV is the change in the sensor output voltage with respect to a no-load output voltage. For the calibration, random forces were applied to the handle connected to the F/T sensor. The f_x and f_y values from the F/T sensor (Figure 3.6(a)) and the voltage values from the Nitinol sensors (Figure 3.6(b)) were used to estimate the parameters of the transformation matrix using the *lsqnonlin* function in MATLAB®. The resulting transformation matrix parameters were:

$$T = \begin{bmatrix} 35.9399 & -40.3963 & -11.4245 & 12.7828 \\ -10.5195 & 11.8238 & -36.3261 & 40.6450 \end{bmatrix}$$

3.4.3 Sensor Performance Assessment

The performance of the sensorized da Vinci instrument was assessed experimentally as follows. (i) The lateral forces acting at the distal end of the instrument shaft were validated by using the same setup used for sensor calibration (Figure 3.5(a)). The handle attached to the F/T sensor

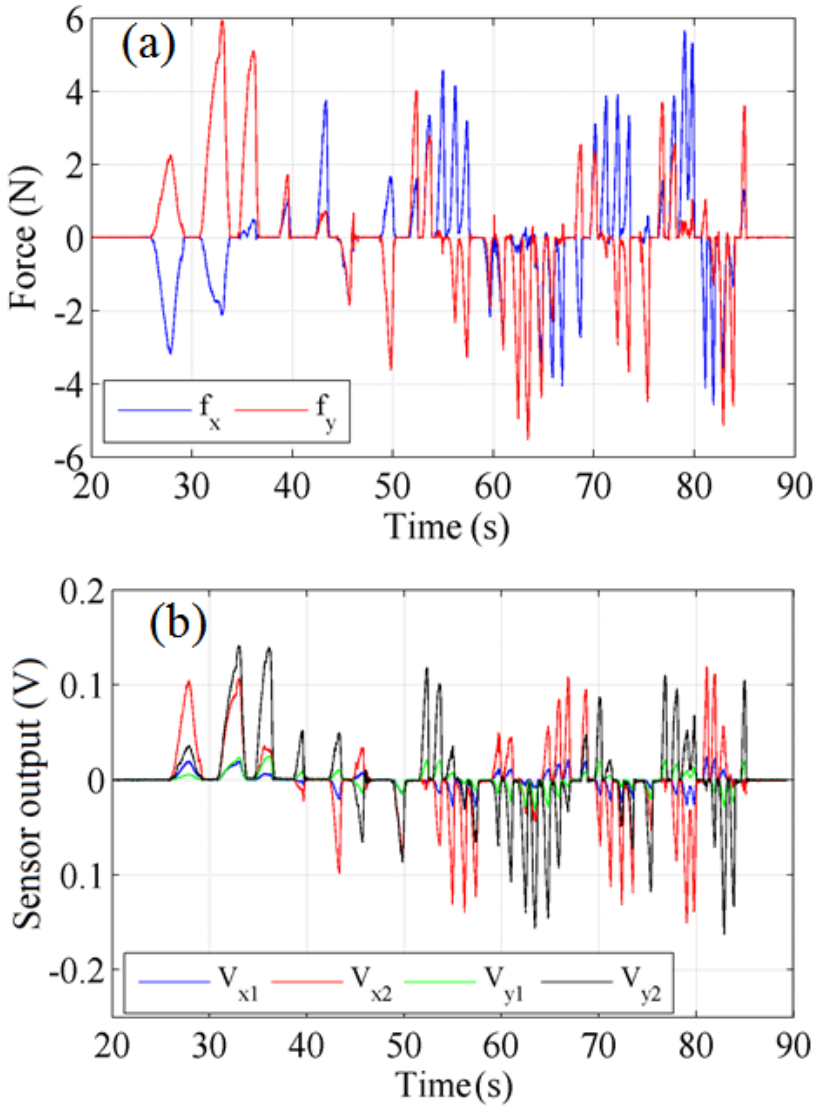

Figure 3.6: Data for calibration (a) the applied force, and (b) the sensor output.

was moved randomly in different directions and the forces measured by the F/T sensor and the Nitinol sensors in the *x* and *y* directions were compared.

(ii) The force acting on the tip of the instrument at different orientations were validated by suspending weights from the tip (0 to 475 g, in 25 g increments). This process was repeated for 6 different orientations of the tip (Figure 3.7). The resulting forces measured by the Nitinol sensors were compared with the theoretical applied forces.

(iii) The resolution of the Nitinol sensor was determined by hanging weights on the tip of the instrument (0 to 10 g, in 1 g increments) and observing the smallest detectable change in the signal. This resolution was then validated over a range of 0 to 50 g by hanging weights in 5 g increments.

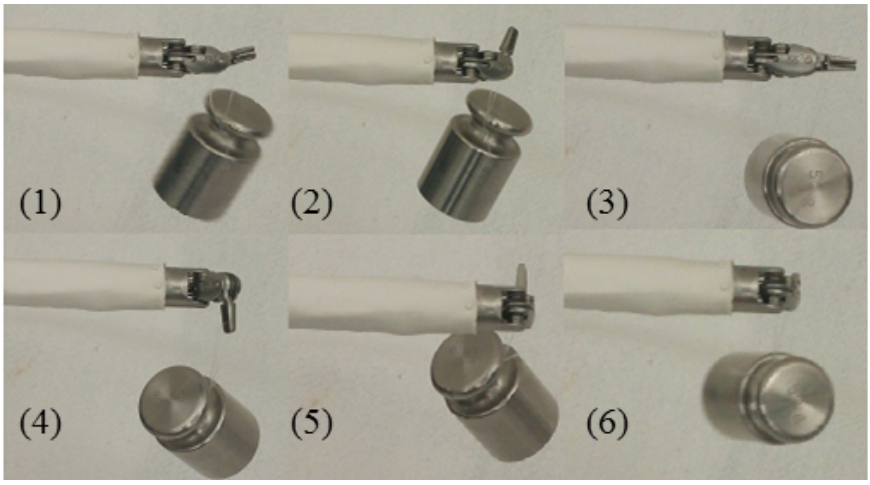

Figure 3.7: 6 different orientations of the tip used for validating the sensor measurement when forces are acting at different orientations of the tip.

(iv) To measure no-load signal drift, force measurements were recorded after the system had been turned off for 2 hours, and then again after the system had been running for 20 minutes.

3.5 Results

The normalized change in resistance versus strain for the Nitinol wire is plotted in Figure 3.8 for all five cycles. From (3.3), the GF of the Nitinol wire was found to be 3.56 by calculating the slope of the best linear fit of the plotted data. In addition, the sensor showed a linear response, good repeatability and negligible hysteresis. The sensor also performed well during measurement of the lateral forces acting near the distal end of the da Vinci instrument. Figure 3.9 shows the forces measured by the Nitinol sensors and a commercially available F/T sensor.

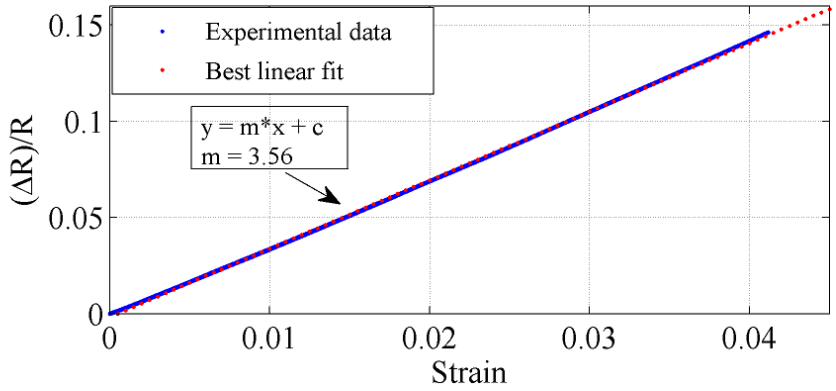

Figure 3.8: Normalized change in resistance with respect to strain in Nitinol wire (for 5 repetitive cycles).

Table 3.1: Comparison of Nitinol sensor with commercial strain sensors

Properties	**Nitinol sensor**	**Silicon strain gauge [18]**	**Metallic foil strain gauges [19]**			**FBG sensor**
			Constantan	**Isoelastic**	**Karma Alloy**	
Gauge factor	3.56	50-200	2.1	3.2	2.0	N/A
Maximum positive strain	4.25%	<0.5 %	3-5%	1.5% (Nonlinear above 0.5%)	1.5%	<1 %
Cost of 1 sensor	<2 $	10-25 $	10-25 $	10-25 $	10-25 $	100 $
Professional installation	Not required	Required	Required	Required	Required	Not required

The Nitinol sensors are able to accurately measure the two-DOF forces with an RMS error of 0.027 N in the x direction and 0.031 N in the y direction. Figure 3.10 shows the force measured by the Nitinol sensors when a series of weights are placed at the tip of the instrument for different orientations. The Nitinol sensors identified properly the applied forces independently of the tip orientation, with an RMS error of 0.032 N.

The force resolution of the sensorized tool was characterized based on the peak-peak noise present in the measured signal when a constant load is applied. Figure 3.11 shows force measurements from 0 to 50 g, in 5 g increments. Based on the experiments the resolution of the sensor was found to be at least 55 mN. Significant drift in the sensor force reading was observed when the system was just turned on (Figure 3.12(a)). The mean of the measured force for 1 second in the x and y directions drifted up to 91 mN and -76 mN after 500 seconds respectively. This drift occurs because the temperature of the sensors is increased due to Joule heating. After a few minutes, the sensor temperature becomes stable, and the signal drift is almost negligible as shown in Figure 3.12(b).

Table 3.1 shows a comparison between the Nitinol sensor under study and commercially available strain sensors such as silicon strain gauges (from Micron Instruments, Simi Valley, CA, USA), metallic foil strain gauges (from Vishay Precision Group, Raleigh, NC, USA) and FBG sensors (from Technica, Beijing, China). The Nitinol sensor has a better combination of high sensitivity and large strain measurement range than the compared metallic foil strain gauges. Unlike the metallic foil strain gauges, the Nitinol sensor does not require any backing material, which combined with its ultra thin structure significantly increases its flexibility. Therefore high flexibility in addition to high sensitivity and large strain range allows this sen-

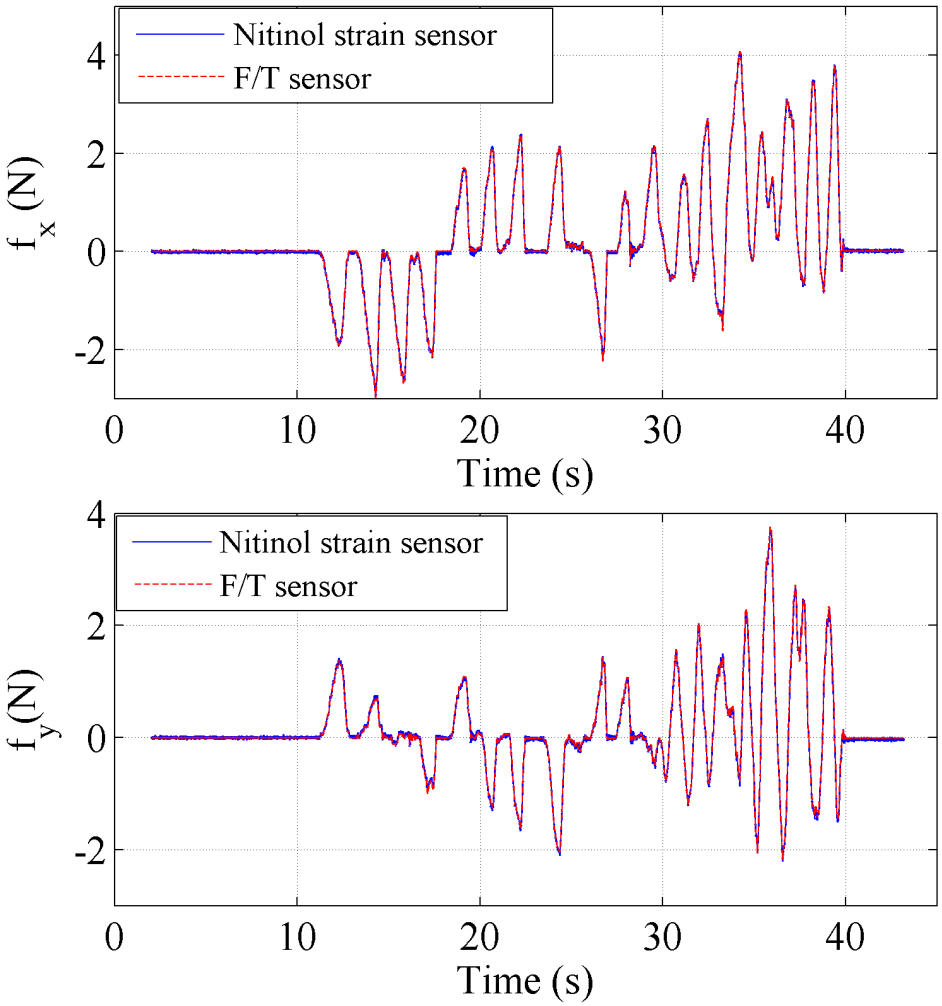

Figure 3.9: Force measured by F/T sensor versus force measured by Nitinol strain sensor.

sor to be used in both rigid and flexible surgical instruments. Although the sensitivity of the Nitinol sensor is lower than that of silicon strain gauges and FBG sensors, our results showed acceptable performance of the sensors for most of the surgical applications. Furthermore, the Nitinol sensor can be easily glued on objects without any sophisticated surface preparation whereas often a professional is required to install the silicon and metallic foil strain gauges on objects which significantly increases the overall cost of using these technologies Therefore, because of its low cost, relatively high sensitivity and large strain measurement range, the proposed Nitinol sensor is an attractive alternative to other strain sensors for variety of surgical instruments. It is worth mentioning, that in order to use this sensor on conductive surfaces, an insulation coating would be required.

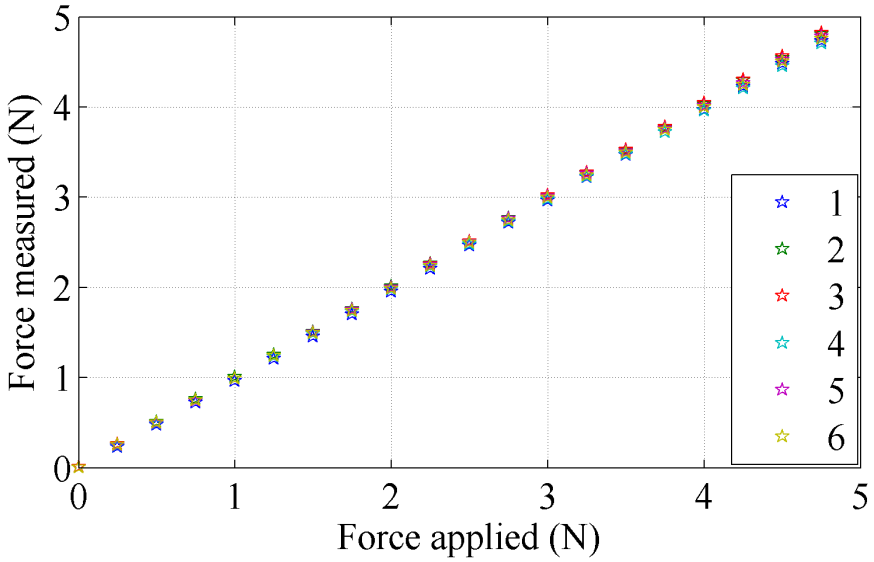

Figure 3.10: Resultant force measured by the Nitinol sensors for a series of weights placed at the tip for different orientations (see Figure 3.7).

3.6 Conclusion

This chapter presented a novel, ultra thin strain sensor made of superelastic Nitinol wire. These sensors were installed on a da Vinci surgical tool in order to demonstrate their capability of force sensing in a surgical instrument. The results showed that the Nitinol sensor technology performs well for rigid surgical instruments and has low cost and space requirements. Given the fact that Nitinol sensors also have large strain measurement range and good sensitivity, it can be concluded that they are excellent cost effective sensors for strain/force measurement of thin surgical instruments such as laparoscopes as well as more flexible instruments such as catheters and surgical needles.

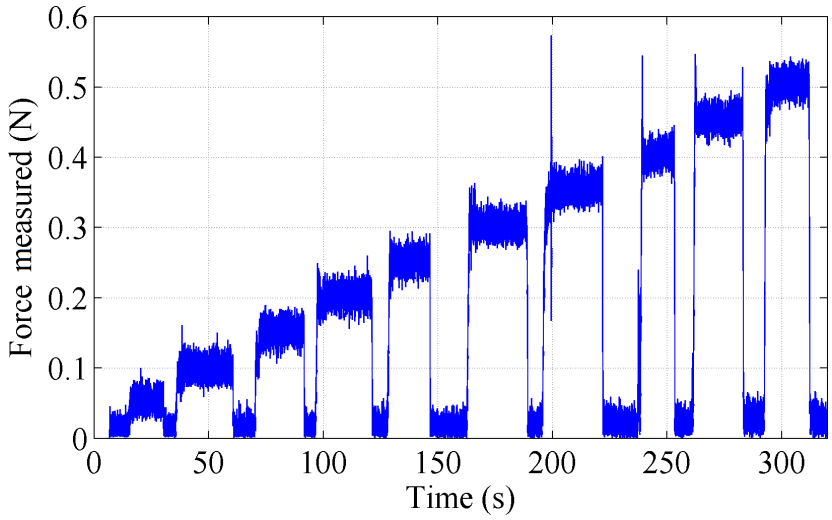

Figure 3.11: Nitinol sensor resolution characterization.

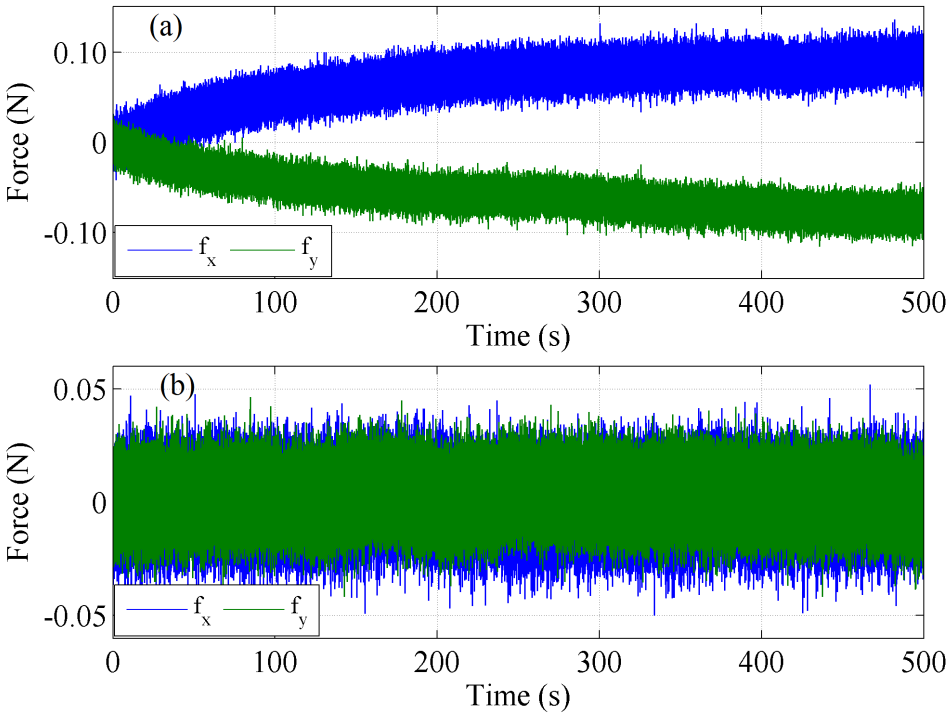

Figure 3.12: No-load signal drift in the force measurement, (a) immediately after turning on the sensor and (b) 20 minutes after turning on the sensor.

Bibliography

[1] C. Wagner, N. Stylopoulos, and R. Howe, "The role of force feedback in surgery: analysis of blunt dissection," in *Proceedings of the 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, Mar. 2002, pp. 68–74.

[2] M. Tavakoli, R. V. Patel, and M. Moallem, "Robotic suturing forces in the presence of haptic feedback and sensory substitution," in *Proceedings of 2005 IEEE Conference on Control Applications, 2005. CCA 2005.*, Aug. 2005, pp. 1–6.

[3] B. T. Bethea, A. M. Okamura, M. Kitagawa, T. P. Fitton, S. M. Cattaneo, V. L. Gott, W. A. Baumgartner, and D. D. Yuh, "Application of haptic feedback to robotic surgery," *Journal of Laparoendoscopic & Advanced Surgical Techniques*, vol. 14, no. 3, pp. 191–195, June 2004.

[4] A. L. Trejos, R. V. Patel, and M. D. Naish, "Force sensing and its application in minimally invasive surgery and therapy: A survey," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 224, no. 7, pp. 1435–1454, July 2010.

[5] A. L. Trejos, R. V. Patel, M. D. Naish, A. C. Lyle, and C. M. Schlachta, "A sensorized instrument for skills assessment and training in minimally invasive surgery," *ASME. J. Med. Devices*, vol. 3, no. 4, pp. 041 002–12, Nov. 2009.

[6] F. L. Hammond, M. J. Smith, and R. J. Wood, "Estimating surgical needle deflection with printed strain gauges," in *the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2014.

[7] P. Berkelman, L. Whitcomb, R. Taylor, and P. Jensen, "A miniature microsurgical instrument tip force sensor for enhanced force feedback during robot-assisted manipulation," *IEEE Transactions on Robotics and Automation*, vol. 19, no. 5, pp. 917–921, Oct. 2003.

[8] D. Yurkewich, A. Escoto, A. Trejos, M.-E. LeBel, R. Patel, and M. Naish, "Low-cost force-sensing arthroscopic tool using threaded fiber bragg grating sensors," in *5th IEEE*

RAS EMBS International Conference on the Biomedical Robotics and Biomechatronics, Aug. 2014, pp. 28–33.

[9] K. S. Shahzadaa, A. Yurkewich, R. Xu, and R. V. Patel, "Sensorization of a surgical robotic instrument for force sensing," in *SPIE BiOS. 2016 (In Press)*, 2016.

[10] R. Xu, A. Yurkewich, and R. V. Patel, "Curvature, torsion and force sensing in continuum robots using helically-wrapped fbg sensors," *IEEE Robotics and Automation Letters (In press)*, 2016.

[11] K. Yokoyama, H. Nakagawa, D. C. Shah, H. Lambert, G. Leo, N. Aeby, A. Ikeda, J. V. Pitha, T. Sharma, R. Lazzara, and W. M. Jackman, "Novel contact force sensor incorporated in irrigated radiofrequency ablation catheter predicts lesion size and incidence of steam pop and thrombus," *Circulation. Arrhythmia and electrophysiology*, vol. 1, no. 5, pp. 354–362, Dec. 2008.

[12] M. van Schoor, A. Lengyel, G. Muller, and A. du Plessis, "Method and device for measuring strain using shape memory alloy materials," U.S. Patent 6 550 341 B2, Apr. 22, 2003.

[13] A. T. Tung, B.-H. Park, D. H. Liang, and G. Niemeyer, "Laser-machined shape memory alloy sensors for position feedback in active catheters," *Sensors and Actuators A: Physical*, vol. 147, no. 1, pp. 83–92, Sept. 2008.

[14] L. Petrini and F. Migliavacca, "Biomedical applications of shape memory alloys," *Journal of Metallurgy*, vol. 2011, pp. 1–15, Sept. 2011.

[15] D. Patranabi, *Sensors and Transducers*, 2nd ed. PHI Learning Pvt. Ltd., 2003.

[16] "Soldering nitinol: Determining the best flux / solder combination," Fort Wayne Metals. [Online]. Available: www.fwmetals.com/default/assets/File/DesignofExperimentforSolderProject(1).pdf

[17] S. J. Blumenkranz and D. Q. Larkin, "Force and torque sensing for surgical instruments," U.S. Patent 8 945 095 B2, Feb. 3, 2015.

[18] "Semiconductor strain gages." [Online]. Available: https://www.microninstruments.com/descriptions/document/StrainGageVolume2014v1.pdf

[19] "Strain gage selection: Criteria, procedures, recommendations." [Online]. Available: http://www.vishaypg.com/docs/11055/tn505.pdf

Chapter 4

Conclusion and Future Work

This chapter presents the concluding remarks of this thesis research work and suggests possible area of future research on this topic based on the work presented in Chapters 2 and 3.

4.1 Conclusion

The research described in this thesis shows that SMAs have good potential for application both as actuators and strain sensors for instruments used in miniamally invasive interventions. When they are used as actuators, they help in significantly reducing the cost of the actuating the instrument while reducing the amount of space required to implement them. When they are used as strain sensors they provide a cost-effective solution to measure tool-tissue interaction forces and the large bending strains in flexible instruments. Therefore, in this thesis we have studied some of the major issues involved in using SMAs as actuators and strain sensors for applications in MIST.

In Chapter 2, our focus was developing an accurate model to precisely control SMAs when used as actuators. An adaptive model based model predictive controller was developed to control the strain in an SMA actuators. Several experiments were conducted in order to evaluate the performance of both the model and the controller. The results showed high accuracy of the model in describing the hysteretic behavior of SMAs in response to input current, and high accuracy of the controller in tracking various reference trajectories even in the presence of the disturbances.

Chapter 3 focused on designing a highly miniaturized strain sensor using superelastic Nitinol wire. The ultra thin strain sensor designed in this chapter comprised a single 15μm diameter of Nitinol wire and two lead wires connected to its two ends. The sensor was shown to have a GF of 3.5 which is better than most metallic foil strain gauges. In addition, the sensor also

had a large strain measurement range. These two properties make this sensor well suited for use in in both rigid and flexible surgical instruments. The sensor performance was validated by sensorizing a da Vinci surgical tool to measure the lateral forces acting near the distal end of the instrument. The results showed good accuracy and resolution of the force measurements.

4.2 Future Work

The research presented in this thesis has prepared the framework for implementing low-cost accurate sensors and actuators for the slender flexible instruments that are used in minimally invasive interventions. As part of the future work in this area, some important applications that are envisaged include sensing and actuation of the flexible instruments such as active catheters and endoscopes. The intention is to use SMA- based actuators and sensors in a feedback control stratgey to guide these instruments through the appropriate lumen. The sensors could be implemented near the instruments so as to measure the forces acting at the tip when it comes in contact with tissues. Using the model and control framework presented in Chapter 2, it should be posible to develop a closed-loop control scheme to control the tip contact force. This would be particularly useful applications such as catheter-based cardiac ablation where control of tip contact force plays an important role in the success of the treatment.

Another possible application of the technology resulting from this research is in the sensorizing of slender flexible needles to measure needle bending and tip deflection during needle-based intervention such as in biopsy and brachytherapy (localized radiotherapy for cancer treatment by implantation of small radioactive pellets in tumors). In this thesis, an application to sensorization of a da Vinci instrument was discussed. Along the same lines, conventional laparoscopic tools can also be sensorized with similar Nitinol SMAs. These can provide tool-tissue interaction forces during MIS.

Lastly, it should be noted that for clinical use of the instruments in the above mentioned applications, it will be necessary to first demonstrate that the SMA sensors and actuators would be able to withstand a conventional sterilization procedure used for surgical instruments. Therefore future work would address this issue from the point of view of development of appropriate packaging and implementation of sensors and actuators. Since the cost would be relatively low, it should be feasible to consider using this technology on single-use instruments.

Curriculum Vitae

Name:	Amit Srivastava
Post-Secondary Education and Degrees:	The University of Western Ontario London, Ontario, Canada 2013 - 2016 M.E.Sc (Electrical and Computer Engineering) Indian Institute of Technology Jodhpur Jodhpur, Rajasthan, India 2009 - 2013 B.Tech (Electrical Engineering)
Related Work Experience:	Teaching Assistant The University of Western Ontario 2013 - 2015 Research Assistant The University of Western Ontario 2013 - 2016

Publications:

A. Srivastava, R. Xu, A. Escoto, C. Ward, and R. V. Patel, "Design of an Ultra Thin Strain Sensor Using Superelastic Nitinol for Applications in Minimally Invasive Surgery", in *2016 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, Banff, Alberta, 12-15 July, 2016.