Segmented Binary Control of Shape Memory Actuator Systems

by

Brian A. Selden

B.S., Mechanical Engineering
University of California, Berkeley, 2003

Submitted to the Department of Mechanical Engineering in Partial Fulfillment of the Requirements for the Degree of

Master of Science
at the
Massachusetts Institute of Technology

February 2005

Signature of Author:__ _____

Department of Mechanical Engineering
January 12, 2005

Certified by: ___ _____

Haruhiko H. Asada
Professor of Mechanical Engineering
Thesis Supervisor

Accepted by: _____ _____

Lallit Anand
Professor of Mechanical Engineering
Chairman, Department Committee on Graduate Students

Segmented Binary Control of Shape Memory Actuator Systems

by

Brian A. Selden

Submitted to the Department of Mechanical Engineering
on January 12, 2005 in Partial Fulfillment of the
Requirements for the Degree of Master of Science

ABSTRACT

A new approach to the design and control of shape memory alloy (SMA) actuators is presented. SMA wires are divided into many segments and their thermal states are controlled individually as a group of finite state machines. Instead of driving a current to the entire SMA wire and controlling the wire length based on the analogue strain-temperature characteristics, the new method controls the binary state (hot or cold) of individual segments and thereby the total displacement is proportional to the length of the heated segments, i.e. austenite phase. Although the thermo-mechanical properties of SMA are highly nonlinear and uncertain with a prominent hysteresis, Segmented Binary Control is robust and stable, providing characteristics similar to a stepping motor. However, the heating and cooling of each segment to its bi-stable states entail longer time and larger energy for transition. An efficient method for improving speed of response and power consumption is developed by exploiting the inherent hysteresis of SMA. Instead of keeping the extreme temperatures continuously, the temperatures return to intermediate "hold" temperatures closer to room temperature but sufficient to keep constant phase. Coordination of the multitude of segments having independent thermal states allows for faster response with little latency time even for thick SMA wires. Based on stress dependent thermo-mechanical characteristics, the hold temperature satisfying a given Stress Margin is obtained. The new control method is implemented using the Peltier effect thermoelectric devices for selective segment-by-segment heating and cooling. Experiments demonstrate effectiveness of the proposed method.

Thesis Supervisor: Harry H. Asada

Title: Ford Professor of Mechanical Engineering

Table of Contents

1 Introduction

Shape Memory Alloy (SMA) actuators produce the one of the highest stresses among all actuators that has ever been developed [DARPA and SRI International]. Its maximum actuation stress of over 200 MPa [1] is 570 times larger than the human muscle [2] and 25 times larger than the latest electroactive polymer actuators [3]. Its energy density of over 100 Joule/cm3 [1] is 100 times larger than that of piezoelectrics [4]. These high stress and energy density characteristics allow SMA actuators to be effectively used in various applications where space and weight constraints are critical design requirements. These include medical devices [5], robots [6], and smart structures [7].

Despite the tremendous actuation stress and energy density, SMA has highly complex nonlinear dynamics that limit applicability and utility to rather simple tasks. In the past few decades a number of research groups have modeled these complex SMA thermomechanical behaviors for position and force control. These include finite element methods based on the Galerkin method [8], Preisach approaches [9], and models based on thermodynamic principles and constitutive equations [10, 11]. Researchers also have attempted to compensate for these thermomechanical nonlinearities utilizing nonlinear control approaches: neural networks and a sliding mode based robust controller [12], neural fuzzy [13], dissipativity [14], and variable structure control [15]. Despite these valuable research efforts, control of SMA is still difficult. Fundamental control performance, e.g. speed of response and disturbance rejection, is still limited even when complex models and sophisticated controls are used.

An alternative to the design and control of SMA actuator systems is presented. There are two fundamental concepts. One is to utilize the hysteresis and the bi-stable nature of the nonlinearity inherent in SMA. Rather than coping with the highly nonlinear properties using complex dynamic compensation, we will exploit the nonlinearity in a new approach. The other is to divide a SMA wire into many segments and control each phase transition separately. Instead of controlling the phase transition of the entire SMA wire as a single plant as it is traditionally done, the phase transition is controlled segment

by segment. The resultant control scheme, called Segmented Binary Control (SBC), performs phase control of each segment in a binary mode with two temperature thresholds: one hot and one cold; each ensuring a complete transition to the austenite and martensite phase, respectively. The aggregate displacement of the entire SMA wire is proportional to the total length of the segments in the austenite phase. Since binary phase control exploits the saturation behavior and not the analogue range of the complex thermomechanical characteristics, the control of individual segments is straightforward, yet it brings about accurate discrete position control despite disturbances and changing loads.

SBC implementation requires locally heating and cooling SMA segments: thermoelectric devices, or Peltier modules, are a rational choice. Recently the SMA research community has applied these solid-state heat pumps successfully. In conjunction with resistance heating, the dynamic response of the system can be increased by a factor of 2 [16]. Another experiment achieved a cycling rate of 0.5 and 1.0 Hz at 2.5% at 1.3% strain, respectively [17].

In the following, the basic concept of the SBC approach is first described, followed by its improvements with regard to speed of response and energy consumption. The error modes are discussed and initial implementation and proof-of-concept prototyping are then described. Experimental data and performance tests will be presented to demonstrate advantages and to justify the concept.

2 Segmented Binary Control: Principle and Issues

Traditional SMA drive systems consist of heating the entire length of the SMA with electric current (resistance heating) and cooling with natural or forced convection. The entire SMA wire is controlled as a single plant: Figure 1a shows a schematic of this system. The fundamental difficulties of SMA control include:

- The phase transition diagram shows a prominent hysteresis with transitional regions that are both steep and nonlinear.
- The phase transition temperatures shift as the stress and environmental conditions change [18].
- The process is distributed and thereby phase transition is not uniform along the SMA wire.

2.1 Segmented Binary Control

As mentioned previously, these difficulties are managed with various nonlinear control methods with detailed models of the SMA phase transition. An alternative to these existing dynamic compensation methods is to independently control numerous SMA segments with a simple control law for each. To control the output displacement of a SMA wire, the existing methods use some **analogue** property of SMA, e.g. the ratio of austenite/martensite phase. In contrast, our approach, called Segmented Binary Control, is a **digital** approach based on the following concepts:

- **Segmentation**. Instead of treating the entire SMA wire as an aggregate single process, the wire is divided into a multitude of segments that are controlled separately, as shown in Figure 1b.
- **Binary Control.** The individual SMA phase transition is controlled in a binary manner. By manipulating every temperature to either the binary high or binary low temperature, T_H or T_C, each segment of a SMA wire will be in either the austenite phase ("ON") or the martensite phase ("OFF"), as shown in Figure 2.

Combining segmentation and binary control concepts creates a new digital control approach to SMA: Segmented Binary Control (SBC). For each segment turned "ON", it will contract some percentage of its length. The resultant displacement of the entire wire is determined by the addition of small displacements created at each individual segment. If all segments are equal length, by specifying the number, say n, of "ON" segments (austenite, hot segments), it will contract n units, where a unit depends on the length of the segment (see Figure 1b). Thus, SBC converts SMA to something synonymous with a stepping motor.

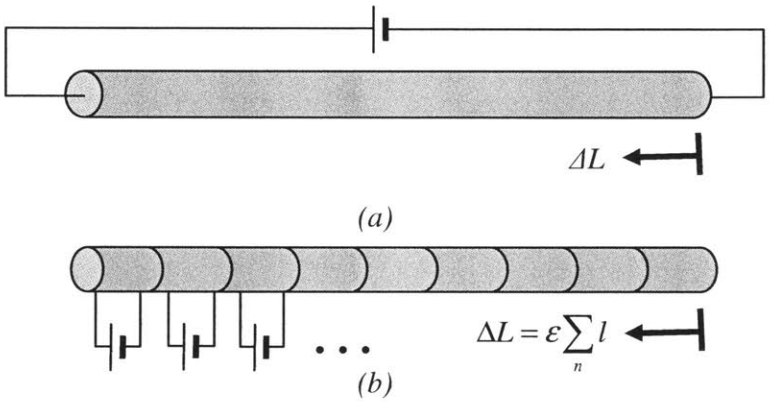

Figure 1. (a) Traditional SMA heating method,
(b) Segmented binary control with n segments turned "ON".

SBC uses two basic characteristics of SMA wires. First, the resultant motion, i.e. displacement at one end of the wire, is the integration of the strain along the whole wire. However, the strain does not have to be distributed uniformly as is traditionally done when the SMA wire acts as a single plant. Instead, like in SBC, the strain is generated at select segments and the resultant position is the summation of individual displacements. In other words, the same total displacement can be generated with selective heating and cooling of local segments as in heating the SMA as a single plant. Second, SBC uses SMA saturation behavior in phase transition, i.e. the two flat levels of strain in Figure 2. Instead of using the steep portions of the temperature-strain curves, which are depended on loading and environmental properties [18], SBC uses the bi-stable nature of the material. Like other digital devices SBC exploits the high fidelity and robustness of digital devices: it is insensitive to complex nonlinearity and varying hysteresis.

Furthermore, segmented control provides a multitude of degrees of freedom (DOF) in controlling the total displacement over a single SMA wire. The system is hyper-redundant, since individual segments can take diverse states to produce the same output displacement. Exploiting the multitude DOF create unique features for SMA, which would be unattainable with the traditional single process control. We will explore such unique features in the following sections.

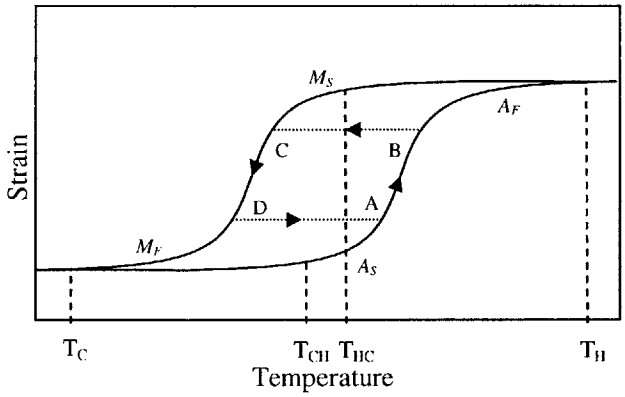

Figure 2. Strain-temperature characteristics and state transition threshold temperatures and intermediate hold temperatures.

Despite the unique features, Segmented Binary Control has a couple drawbacks that must be overcome. Two major problems with this approach are:

1. There are large latency times in state transitions.
2. Energy is needed to keep the SMA at either hot or cold state at all times.

The latency times are a consequence of the discrepancy between the SBC threshold temperatures (T_H and T_C) and the SMA transformation temperatures (A_S and M_S). As illustrated in Figure 3, the segment's temperature must be changed from T_C to A_S before any displacement is realized. This time is the latency time, L_H. After this latency period, there is a rise time until the transformation is completed and all the displacement is realized. Contraction begins at the A_S temperature, and is at the full extent at A_F ($<T_H$). There is a similar process for the opposite transition.

A finite amount of energy is needed for transition in order to change a segment's temperature from T_H to T_C, or vice versa. However, in order to maintain that temperature, heat must be pumped to counteract the losses to the environment. This becomes

increasingly more difficult as the difference between T_H and T_C deviate from room temperature.

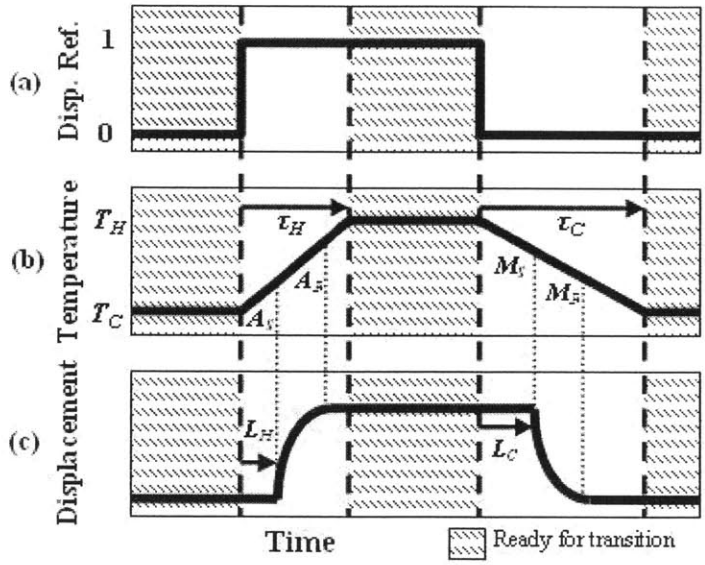

Figure 3. Latency times in phase transition: step responses of temperature and strain.

2.2 Hysteresis Loop Control

There is a simple, yet effective improvement to the latency times and energy inefficiency mentioned above. It entails incorporating two additional intermediate, user-defined temperatures that segments are maintained at until an input command is given to change states. The four user specified temperatures are:

- **Hot Temperature (T_H) :** The temperature that the SMA must reach temporarily to ensure the expected contraction. This is the same as in the basic SBC.
- **Hot to Cold Transition Temperature (T_{HC}) :** The temperature that the SMA must reach after reaching T_H before transitioning to T_C.
- **Cold Temperature (T_C) :** The temperature that the SMA must reach after T_{HC} to ensure the expected relaxation of the SMA. This is the same as in the basic SBC.
- **Cold to Hot Transition Temperature (T_{CH}) :** The temperature that the SMA must reach after reaching T_C before transitioning to T_H.

These "hold" temperatures are shown in Figure 2 along with the hot and cold threshold temperatures.

With these "hold" temperatures the control of each segment is performed in the following manner:

- When a segment is instructed to change to the "ON" state from the "OFF" state, heat the segment from the temperature T_{CH} to T_H and then immediately cool down to T_{HC}.
- When a segment is instructed to change to the "OFF" state from the "OFF" state, cool the segment from the temperature T_{HC} to T_C and then immediately heat up to T_{CH}.

This strategy utilizes the Schmidt-Trigger type bi-stable switching characteristics of SMA. Note that this strategy would not work if SMA had no hysteresis in its strain-temperature diagram. Because of the hysteresis inherent in SMA we do not have to keep the extreme temperatures, high or low, to maintain the current phase of a SMA segment. In contrast to the existing methods which strive to eliminate the effect of hysteresis, this control strategy exploits the hysteresis. We call this control method Hysteresis Loop Control (HLC) to differentiate it from the basic SBC.

Figure 4 shows a complete position cycle for a segment when HLC is used. Note that once the temperature has reached T_H and T_C, it immediately returns to the intermediate temperatures, T_{HC} and T_{CH}, respectively. Note also that the displacement begins to change soon after the command signal changes, since the intermediate temperatures, T_{HC} and T_{CH}, can be set close to the transition start temperatures A_S, and M_S, respectively. Therefore, the latency times, L_C and L_H, are significantly improved with the Hysteresis Loop Control.

Furthermore, HLC will save power since extreme temperatures, T_H and T_C, are needed only for a short time. Power is saved by reducing the extreme temperatures to the intermediate ones closer to room temperature and keeping them at those levels until commands are given to change states (assuming $T_C < T_{room} < T_H$). In particular, if both intermediate hold temperatures, T_{HC} and T_{CH}, are close to the ambient temperature, virtually no power is needed to keep SMA at the same state. Power is consumed only

when a transition is made. Consequently, the same output displacement may be held for a fixed load even after power has been shut off.

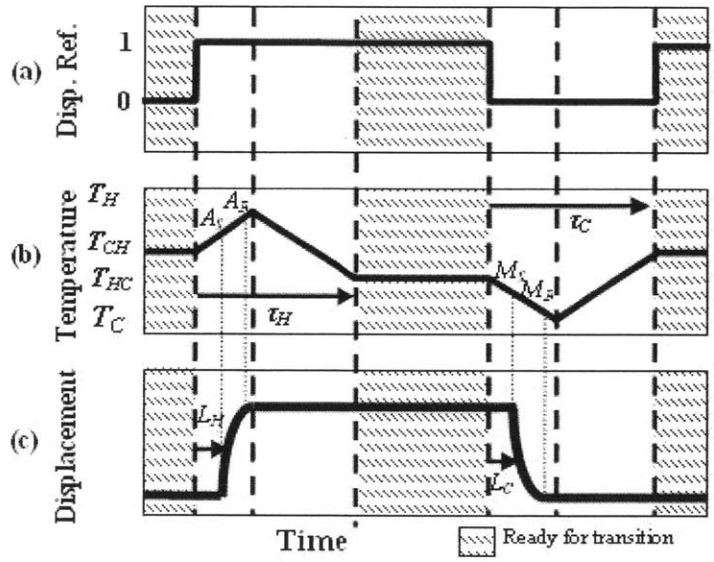

Figure 4. Comparison of latency time between Hysteresis Loop control and traditional analog control.

To fully exploit these features of the Hysteresis Loop Control, a few technical issues must be considered:

1. In the segmented control, a multitude of segments can produce the same output displacement. Those segments can be selectively activated to meet a control goal. Algorithms must be developed to fully exploit the extra freedom of the control system.
2. After making transitions to T_H and T_C the temperature must be brought back to the intermediate hold temperatures. There are delays associated with cooling/heating from the threshold temperatures (T_H and T_C) to the intermediate temperatures (T_{HC} and T_{CH}). These transition delays must be taken into account in control design.
3. The four temperatures for Hysteresis Loop Control (T_H, T_C, T_{HC}, and T_{CH}) must be characterized taking into account the load rejection of the system is lower when the hold temperatures are close to the hysteresis edges. The hold temperature must be selected properly so that it may be robust against load changes.

The following sections will address these issues in detail.

2.2.1 Zero Latency Control Using Inter-Segment Coordination

The Hysteresis Loop Control (HLC) can substantially eliminate the latency times for state transition. In HLC the temperature returns to an intermediate hold temperature, so that next phase transition can be executed without delay. Since the hold temperature is on the verge of state transition, the command for the next state transition will immediately produce displacement with little latency time. This will not only resolve the major drawback of large latency times in SBC but would also achieve even faster response than traditional analog controls do given appropriate materials.

Consider a SMA cycle of motion illustrated with broken lines on the diagram of temperature-strain characteristics in Figure 2. The trajectory starts from point A to point B along the upward hysteresis slope. In order to move backwards, the trajectory must be shifted to the downside slope CD. Before any tangible backward motion begins, the temperature must be lowered across the hysteresis BC, which entails a latency time. If the SMA acts as a single plant, these latency times are unavoidable and can only be reduced by increasing the heating and cooling rates. However, these latency times can be eliminated by using HLC and an algorithm for selectively activating a multitude of segments under HLC.

Figure 5 illustrates that, in the segmented control architecture, individual segments may take diverse states. Segment i in Figure 5, for example, has just made a phase transition to T_H and its temperature is being pulled back towards T_{HC}. Segment ii, on the other hand, has made a transition some time ago and the temperature has already changed to T_{HC}. It is this segment that can immediately execute the next command to transition down. For the cyclic trajectory illustrated in Figure 2, the backward motion may be generated by segments that are different from those responsible for the forward motion. Although those segments used for the forward motion are still at a high temperature when transitioning, other groups of segments, at the intermediate hold temperature (T_{HC}), are ready for the opposite phase transition. As shown in Figure 5, the backward motion can be created with no latency time by activating those segments at "ready state".

This selective activation is made possible only because of the segmented control architecture. The traditional approach based on a single state-transition model cannot eliminate the latency time, since it can control merely the single state that must transverse the wide hysteresis band in order to change the direction of motion. HLC with selective activation exploits the extra control freedom involved in the multitude of segments for overcoming the latency time problem.

The effect of this method will be more significant for thicker SMA wires, for which the traditional approach has particular difficultly in coping with the long latency time problem. It is expected that this selective activation method along with HLC will allow the use of thick SMA wires while minimizing latency times.

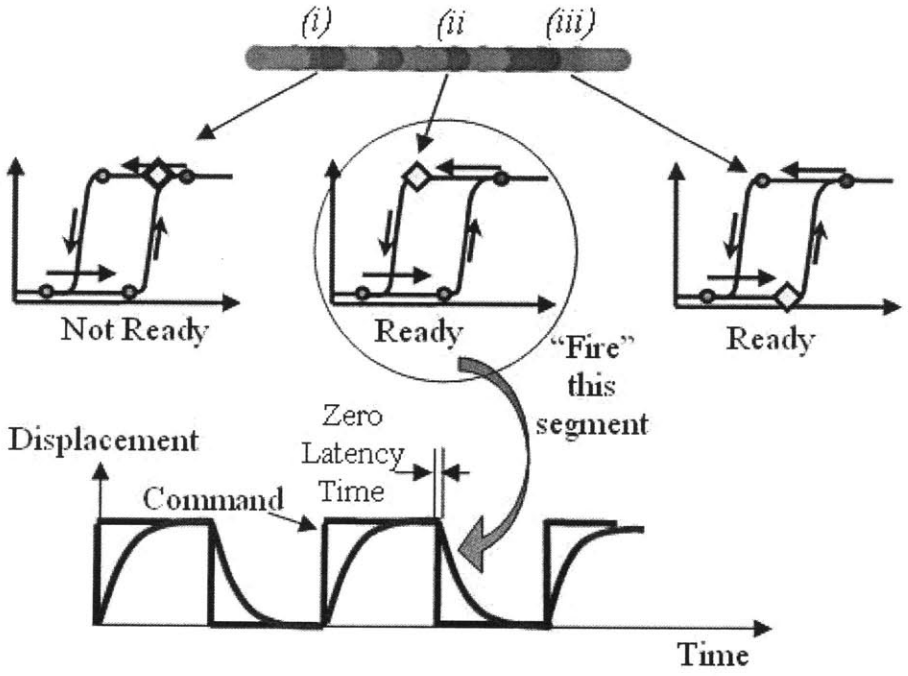

Figure 5. Zero latency control using multi-segment coordination.

Figure 6 shows the overall control system, consisting of global control and local segment control. The outer feedback loop is position feedback from a position transducer measuring the total displacement of the SMA wire. It is compared with a reference position and the difference is fed to the segment coordinator. A simple criterion can be used for the segment coordinator to select an optimal segment to trigger. For example,

among all the segments in "Ready T_{CH}" state the segment coordinator can select the one that has been in the ready state for the longest time. To implement this First-In-First-Out algorithm, the control system must keep track of transitions of individual segments. The transition history may be represented with a time tag attached to each segment indicating when a state transition command was given last time and how long it has been in the current state. Based on the time tag, the segment coordinator generates a trigger signal for the selected segment to initiate state transition. Each local segment control performs simple ON-OFF control based on local temperature measurement and the command from the segment coordinator. If the error of the global feedback is large, the segment coordinator can send triggering commands at a higher frequency. In other words, a type of pulse frequency modulation can be used for the segment coordinator in converting the analogue error signal to discrete commands. This frequency is upper-bounded by the heating and cooling time delays as well as the number of segments involved.

The two transition time delays arise due to the cooling/heating from the extreme temperatures to the intermediate temperatures. During this transition, a change state command is not valid to achieve the minimal latency times that this algorithm promotes. Two parameters aid in characterizing this phenomenon. The parameters τ_H and τ_C represent the minimum amount of time, from the instruction to change states until another command of state change is valid (See Figure 3 and Figure 4). The minimum time for one cycle for one segment is simply the sum of τ_H and τ_C with appropriate characteristic temperatures.

When a single segment has to alter its state between T_H and T_C, as in SBC, this period gives the upper-bound of the system bandwidth. The method developed to increase the bandwidth is to orchestrate the multitude of segments by shifting their phase angles.

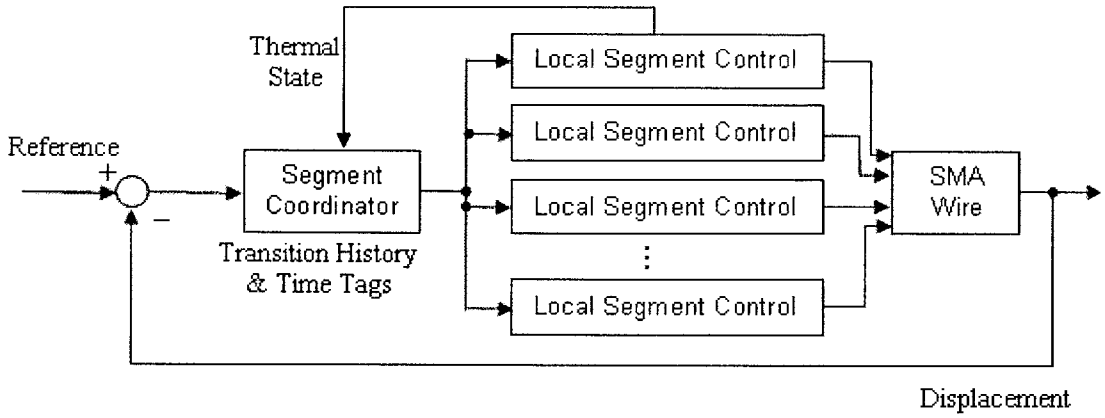

Figure 6. HLC system with segment coordinator and local segment controllers.

2.2.2 Stress Margin

Hysteresis Loop Control, with the four characteristic temperatures, can create very quick transitions by taking advantage of the hysteresis. After any segment transition, that segment's temperature is altered such that only a small temperature change will create the reverse transition. The closer the temperature is to the transition, the smaller the latency time before strain is realized. However, small latency times do not come without a cost: the system is now vulnerable to stress changes. An inherent property of SMA is that stress changes cause the temperature-strain hysteresis to shift [18]. If this shifting causes the operating temperature to move into the transitional areas of the hysteresis, SBC and HLC are no longer valid. In this section, we will obtain the way of selecting the intermediate temperatures so that unwanted phase changes do not occur despite a varying stress in a given range. Specifically we will obtain the "Stress Margin" to guarantee no phase transition occurs within a specified stress range.

Figure 7 illustrates how stress changes cause the hysteresis curve to shift. Whereas an increase in stress causes a shift to the right, a decrease shifts the hysteresis to the left. Any stress change causes one intermediate hold temperature (T_{HC} or T_{CH}) to be closer to the transition and the other to be farther. The stress margin can be quantified based on Figure 7. A practical method is to empirically determine the two extreme stress cases and use the applied loading method to obtain the appropriate information [19].

Given an appropriate stress range $[\sigma_{min}, \sigma_{max}]$ with:

1. $\sigma_{max} < 180$ MPa (the physical limit for the shape memory effect)
2. $\sigma_{min} > 30\%$ of maximum ≈ 54 MPa (the minimum bias force for full contraction).

The stress margin procedure is:

1. Find the temperature-strain hysteresis for σ_{max}. Determine a value of T_H that will complete a full contraction and T_{HC} sufficiently close to the transitional portion of the hysteresis.
2. Similarly, find the temperature-strain hysteresis for σ_{min}. Determine a value of T_C that will complete a full relaxation and T_{CH} sufficiently close to the transitional portion of the hysteresis.

These threshold temperatures and intermediate hold temperatures guarantee that unintended phase shifts will not occur for the stress range $\sigma_{min} < \sigma < \sigma_{max}$.

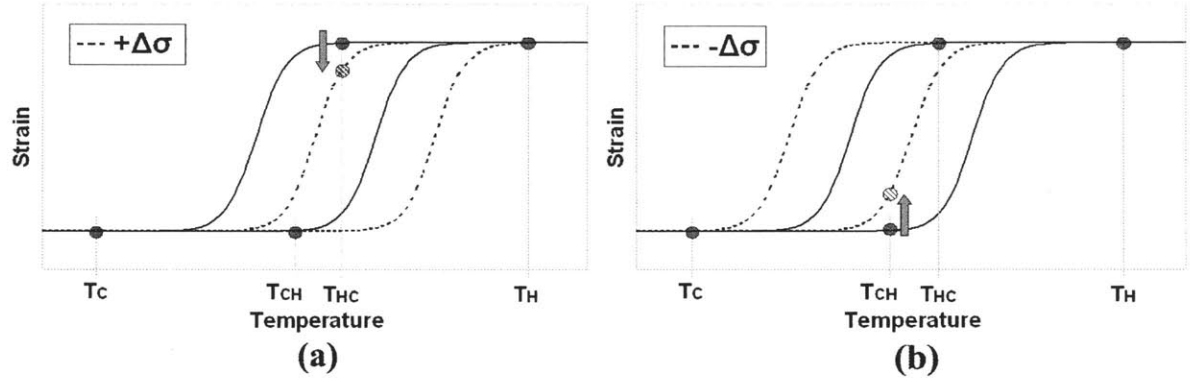

Figure 7. Partial transformations due to stress changes. These invalidate HCL.

The above empirical procedure can be replaced by an analytic method using stress rate c_m. The stress rate (c_m) quantifies the stress dependent shifting of hysteresis curves, or more specifically the shifting of the transition temperatures (A_f, M_S, etc.). It is documented that this property is constant for all moderate to high stresses [20] and has a value between 2.5–15 MPa°C^{-1} [21]. The hysteresis shifting is characterized by equation 1.

$$T_{shift} = \frac{\Delta\sigma}{c_m} \qquad (1)$$

Let T_H, T_C, T_{HC}, and T_{CH} be nominal threshold and intermediate hold temperatures for nominal stress $\bar{\sigma}$. In order to prevent the SMA from making any unintended phase shift induced by stress changes within a range: $\sigma_{\min} = \bar{\sigma} - \Delta\sigma < \sigma < \bar{\sigma} + \Delta\sigma = \sigma_{\max}$, the threshold temperatures and hold temperatures must be modified to:

$$\begin{aligned} T^*_H &= T_H + \frac{\Delta\sigma}{c_m}, \quad T^*_{CH} = T_{CH} + \frac{\Delta\sigma}{c_m} \\ T^*_C &= T_C - \frac{\Delta\sigma}{c_m}, \quad T^*_{HC} = T_{HC} - \frac{\Delta\sigma}{c_m} \end{aligned} \tag{2}$$

3 Sources of Error for Segmented Binary Control

Figure 8 below illustrates a series of hystereses for different loads. The discrepancies occur because the total strain is a sum of the shape memory effect, the compliance, thermal expansion, and many other factors.

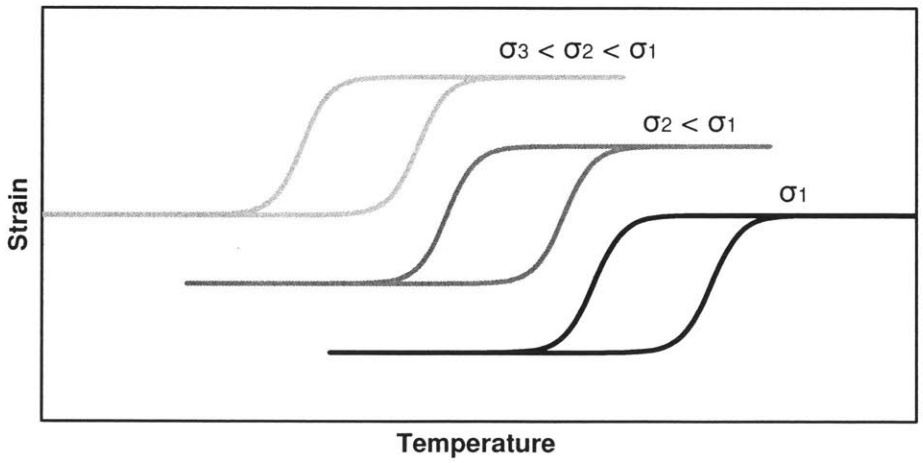

Figure 8. The effect of compliance and the shape memory effect on the temperature strain hysteresis.

The following equations give some mathematical foundation for the graphs above.

$\varepsilon_{total} = SME(\sigma,T) - C(\sigma,\%Aust,L,A) - TE(T,\%Aust)$, ε is positive for contraction.

SME = Shape Memory Effect

C = Compliance Strain

TE = Thermal Expansion

The shape memory effect has already been discussed sufficiently but the compliance, thermal expansion, and other factors are presented in the following sections.

3.1 Compliance of Segmented Binary Control compared to Analogue Control

One distinct advantage of Segmented Binary Control and Hysteresis Loop Control to analogue control is that it has a deterministic and steady compliance. Each segment is either martensite or austenite and the compliance is simply the Young's modulus of the respective phase. If there is a series of segments, they can be modeled as springs in series, each having one of two values. In addition, the characteristic temperatures can be set with the Stress Margin such that any load disturbance will not create unwanted partial transformations and thus the compliance does not change. These points are in direct contrast to analogue control.

Analogue control mediates the magnitude of strain by altering the composition (% austenite) of the SMA. It operates on the temperature hysteresis and only saturates at extreme cases. Therefore, the SMA is some combination of austenite and martensite and its compliance is a value in between the Young's modulus of the respective phases. However, the composition of the SMA wire is not easy to measure and this relationship is nonlinear. In addition, since the temperature hysteresis is stress dependent, there are more complexities.

The temperature hysteresis limits the strain within its bounds. Therefore, if the operating point was on this limiting edge and a change in stress caused the temperature hysteresis to shift (or change shape) in the correct direction, there will be significant unwanted strain. This is attributed mostly to the shifting hysteresis and not due to the compliance of the SMA. This particular argument is valid for a decrease in stress while on the heating edge or an increase in stress while on the cooling edge. This effect can be modeled as a comparatively small Young's modulus. The other cases are sufficiently modeled with the Young's modulus of the current composition assuming the Stress Margin procedure for selecting characteristic temperatures was used.

These compliances are summarized in the Figure 9 and quantified in Table 1.

a) Pure austenite. The compliance is that of austenite and does not change with stress.

b) Pure martensite. The compliance is that of martensite and does not change with stress.

c) Intermediate composition and on heating edge. For increases in stress, the compliance is some value between the Young's moduli of martensite and austenite. However, for decreasing stress the equivalent compliance is small due to the shifting temperature hysteresis and much unwanted strain incurs.

d) Intermediate composition and on cooling edge. For decreases in stress, the compliance is some value between the Young's moduli of martensite and austenite. However, for increasing stress the equivalent compliance is small due to the shifting temperature hysteresis and much unwanted strain incurs.

e) Pure austenite if the characteristic temperatures are set according to the Stress Margin procedure.

f) Pure martensite if the characteristic temperatures are set according to the Stress Margin procedure.

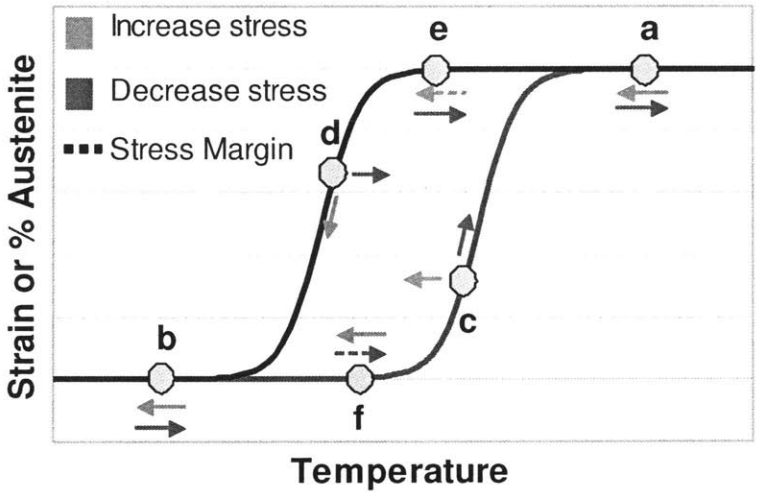

Figure 9. Shows the movement of different operating points on the temperature hysteresis based on the stress applied.

Table 1. Compliances for the different operating points on the temperature hysteresis.

	$+\Delta\sigma / \Delta\varepsilon$	$-\Delta\sigma / \Delta\varepsilon$
(a, e) Austenite	~ 83 GPa [21]	~ 83 GPa [21]
(b, f) Martensite [1]	~ 28 – 41 GPa [21]	~ 28 – 41 GPa [21]
(c) Intermediate Heating [1]	~ 28 – 83 GPa [2]	~ 0.3 – 9 GPa [3]
(d) Intermediate Cooling [1]	~ 0.3 – 9 GPa [3]	~ 28 – 83 GPa [2]

[1] Assume prior stress is sufficient to detwin martensite

[2] This takes a value between pure martensite and austenite that is dependent on the composition

[3] Calculated below

The unwanted strain due to temperature hysteresis shifting can be modeled as a small Young's modulus as listed in the table above. The following describes the calculation to derive these numbers. First, some relevant definitions and assumptions:

- Stress rate (c_m) [MPa $°C^{-1}$] –
    - Quantifies the stress dependent shifting of the hysteresis, or more specifically transition temperatures. Typical values are 2.5-15 MPa $°C^{-1}$ but for the Flexinol SMA wire used in experiment, the value was 7 MPa $°C^{-1}$.
- Skewness –
    - The temperature hysteresis width and shape. The hysteresis width is typically between 12 and 50°C and the difference between the martensite and austenite transformation temperatures ($M_S - M_F$ and $A_F - A_S$) is in the range of 10-25°C.
- Strain –
    - Maximum strain of Nitinol is 8 % although typical values are usually around 4 – 5 %.

Assumptions:

- Stress rate is constant.

- The "skewness" of the temperature hysteresis doesn't change with stress.
- Full contraction is achieved with the "reverse" transformation from M_S to M_F and subsequently fully relaxed with the "forward" transformation from A_S to A_F.
- The "reverse" and "forward" transformations have linear slope on the hysteresis.
- The "reverse" and "forward" transformations saturate.

The slope of the heating and cooling side of the temperature hysteresis is calculated (in units of °C unit strain $^{-1}$).

$$X_M[^\circ C] = \frac{M_S - M_F}{\varepsilon} \qquad (3)$$

$$X_A[^\circ C] = \frac{A_F - A_S}{\varepsilon} \qquad (4)$$

The $\pm \Delta\sigma / \Delta\varepsilon$ value is calculated from multiplying X_M and X_A by the stress rate, c_m. This value is the compliance.

$$\pm \frac{\Delta\sigma}{\Delta\varepsilon}[\text{Mpa}] = c_m\left[\frac{\text{Mpa}}{^\circ C}\right] * X_{M\ or\ A}[^\circ C] \qquad (5)$$

To find the minimum and maximum value of this compliance, use the values listed below in Table 2 and equation 5.

Table 2. Values used to determine the maximum and minimum compliance of the case of unwanted strain.

	Value for Max $\lvert\Delta\sigma / \Delta\varepsilon\rvert$	Value for Min $\lvert\Delta\sigma / \Delta\varepsilon\rvert$	Value for Flexinol $\lvert\Delta\sigma / \Delta\varepsilon\rvert$
$M_S - M_F$ or $A_F - A_S$ [°C]	25	10	25
ε [unitless]	0.04	0.08	0.05
c_m [MPa °C^{-1}]	15	2.5	7
$\pm \Delta\sigma/\Delta\varepsilon$ [GPa]	9	0.3	4

Whereas the compliance of pure austenite is stiff and has a Young's modulus of about 83 GPa, martensite is ductile with a value between 28 and 41 GPa. Segmented Binary Control and Hysteresis Loop Control are based upon negotiating solely between these two compliances. The entire system compliance is deterministic and can be modeled as a series of springs that can only take two values. In addition, it is insensitive to the direction of stress and whether it is heating or cooling. In comparison, the compliance of the analogue control case is sensitive to these. In some cases, the compliance is an intermediate value between martensite and austenite, between 28 and 83 GPa. However, in the other cases, the compliance acts **3 to 100 times lower**, with values of 0.3 to 9 GPa. This comparatively small compliance is the source of the unwanted strain.

3.2 Segmented Binary Control Spring Model

The overall compliance of Segmented Binary Control can be modeled as a series of linear springs as shown in Figure 10.

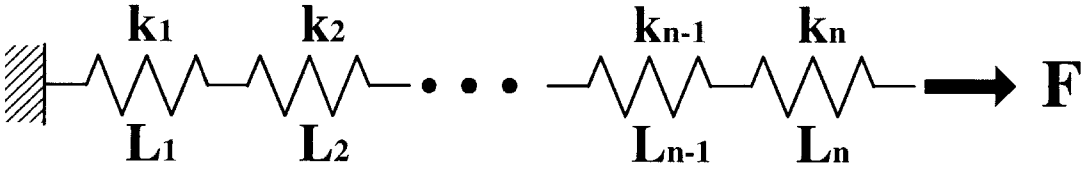

Figure 10. A model the compliance of SBC.

Each spring, assuming linear, has characteristics:

$$\begin{aligned} F &= k_1 x_1 \\ F &= k_2 x_2 \\ &\vdots \\ F &= k_n x_n \end{aligned} \qquad (6)$$

Using one spring constant for the entire system, k_{eff} :

$$F = k_{eff}(x_1 + x_2 + \cdots + x_n)$$

$$F = k_{eff}\left(\frac{F}{k_1} + \frac{F}{k_2} + \cdots + \frac{F}{k_n}\right)$$

$$k_{eff}^{-1} = \sum_{i=1}^{i=n} k_i^{-1} \qquad (7)$$

Using the material properties of SMA, we can find a value of a spring constant for a segment and an expression for k_{eff}:

$$k_i = \frac{EA}{L} \tag{8}$$

$$k_{eff}^{-1} = \sum_{i=1}^{i=n} \frac{L_i}{E_i A} = \frac{1}{A}\left(\sum_{i=1}^{i=n} \frac{L_i}{E_i}\right) \tag{9}$$

With the definition of strain, the compliance of the system can be determined.

$$\varepsilon = \frac{x_{total}}{L} = \frac{F}{k_{eff} L} = \frac{F}{AL}\left(\sum_{i=1}^{i=n} \frac{L_i}{E_i}\right) = \frac{\sigma}{L}\left(\sum_{i=1}^{i=n} \frac{L_i}{E_i}\right) \tag{10}$$

$$Compliance = \frac{\sigma}{\varepsilon} = L\left(\sum_{i=1}^{i=n} \frac{L_i}{E_i}\right)^{-1} \tag{11}$$

The 'n' segments can be lumped into two segments, one austenite and martensite.

$$Compliance = L\left(\frac{L_{aust}}{E_{aust}} + \frac{L_{mart}}{E_{mart}}\right)^{-1} \tag{12}$$

The compliance as a function of percent austenite is shown in Figure 11. The two curves are for values of martensite compliance are 28 and 41 GPa, respectively, and austenite compliance is 83 GPa. The true value of the compliance should be between the two curves.

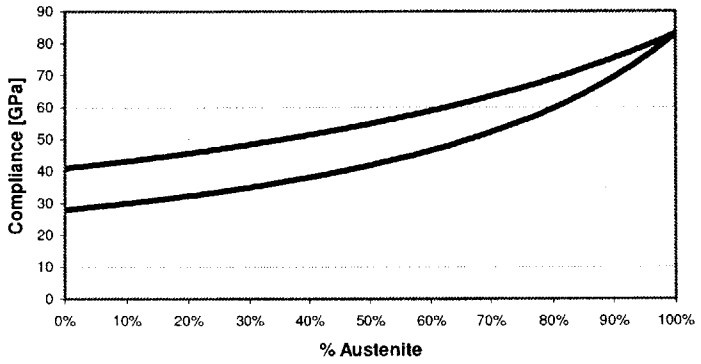

Figure 11. Percent austenite versus compliance of system.

This curve can be approximated by a third order system:

$$Compliance\,[GPa] = 55.603x^3 - 31.118x^2 + 30.029x + 27.382 \qquad \{0 \le x \le 1\}$$

The r^2 error of the third order equation is 0.9996.

3.3 Errors Due to SMA Compliance

The previous sections have discussed the variability of SMA compliance during operation of Segmented Binary Control. This section will examine the worst case errors of this compliance that is constantly changing.

There are two modes of the compliance error:

1. SMA phase stays constant and a load change occurs.
2. Load is constant and a phase change occurs.

The first error mode assumes the SMA to be either pure austenite or martensite and determines the strain error associated with a maximum stress applied. This error is a result of the elastic strain in the SMA. This is shown in equation 13.

$$\varepsilon_{error,\,\max,\,A\,or\,M} = \frac{\Delta\sigma_{\max}}{E_{A\,or\,M}} \qquad (13)$$

The second error mode assumes that the stress is constant and a phase change occurs. The error is a product of the changing Young's Modulus of the SMA. This error ignores the strain inherent to phase change.

$$\varepsilon_{error,\,\max} = \left|\varepsilon_M - \varepsilon_A\right| \qquad (14)$$

$$\varepsilon_{M\,or\,A} = \frac{\sigma}{E_{M\,or\,A}} \qquad (15)$$

$$\varepsilon_{error,\,\max} = \sigma_{\max}\left|E_M^{-1} - E_A^{-1}\right| \qquad (16)$$

To quantify these errors, use the values listed below.

Young's Modulus Martensite: 28 GPa

Young's Modulus Austenite: 83 GPa

Maximum Stress: 180 MPa

To compare the error strains in a relative sense, they are compared to the traditional 5% strain that is expected from the shape memory effect. The results are summarized in Table 3 below

Table 3. Maximum errors for the two error modes.

Error Mode	Phase	Stress	Maximum Error		Compared to 5% SMA strain
1	Constant	Changes	Austenite	0.22%	4.4%
			Martensite	0.65%	13%
2	Changes	Constant	0.43%		8.6%

The second error mode is not as important as the first. The second error mode occurs concomitantly with the shape memory effect. This error will be washed out by the uncertainty of the SMA strain.

3.4 Infinite Series Error

Since the heating element has a fixed width, it heats the SMA directly underneath (ignoring heat transfer between segments). As the temperature of the SMA increases past its transition temperatures, it begins to contract. As a result of this contraction, portions of the SMA that were not being heated are pulled into the area under the heating element and heated. This uncontracted material will contract when its temperature is raised high enough and the process is repeated. This infinite series of contractions results in an error of magnitude are shown in Table 4.

Assuming the magnitude of strain for an SMA wire saturates and has magnitude ε_{sat}, the total amount of strain is:

$$\varepsilon_{tot} = \sum_{i=1}^{\infty} (\varepsilon_{sat})^i = \frac{\varepsilon_{sat}}{1-\varepsilon_{sat}} \qquad (17)$$

Table 4. Errors due to infinite series of contractions.

ε_{sat} [%]	0	1	2	3	4	5	6	7	8
ε_{tot} [%]	0	1.01	2.04	3.09	4.17	5.26	6.38	7.53	8.7
% Error	0	0.99	1.96	2.91	4.08	4.94	5.96	7.04	8.05

Although an important effect, it is practically difficult to distinguish from the phase transformation of SMA.

3.5 Errors due to Coefficient of Thermal Expansion

Shape Memory Alloys, like all other metals, has a coefficient of thermal expansion. This effect counteracts the shape memory effect as the temperature is increased: whereas the shape memory effect causes contractions, the coefficient of thermal expansion will extensions in the material. As a result, there is an error due to this expansion coefficient. Table 5 below shows the maximum error for pure martensite \

$\alpha_{martensite} = 6.6 \times 10^{-6}\ °C^{-1}$

$\alpha_{austensite} = 11 \times 10^{-6}\ °C^{-1}$

Table 5. Errors due to thermal expansion.

Temperature (°C)	$\varepsilon_{martensite}$ [%]	% error[1]		$\varepsilon_{austensite}$ [%]	% error[1]
0	0	0		0	0
20	0.0132	0.264		0.022	0.44
40	0.0264	0.528		0.044	0.88
60	0.0396	0.792		0.066	1.32
80	0.0528	1.056		0.088	1.76
100	0.066	1.32		0.11	2.2
120	0.0792	1.584		0.132	2.64
140	0.0924	1.848		0.154	3.08

[1] Error is calculated assuming that 5% strain is desired.

This effect of this error is insignificant compared to the shape memory effect and it is also disguised in the strain temperature hysteresis.

3.6 Length Contraction Errors

Along a line of SMA segments, as any one segment in the train of segments is turned on, it contracts the SMA under that element. This contraction also pulls the SMA, from the heated segment to the free end, the same amount. Since the heating elements are static, this SMA movement causes the heating elements (from heated segment to free end) to cover a different portion of the SMA then it was previously. This will cause a

transient error if one of these elements was previously turned on and the shifting causes that segment to enter a zone that is turned off (and vice versa). However, at steady state this error is zero. A partial solution to this problem can be embedded into the architecture of segmentation and/or limiting the total length of the SMA wire. For example, put the largest segments closest to the free end and the smallest segments close to the fixed end.

For an extreme example, say a large segment (say 100 mm) is heated and the SMA contracts 5 mm (5%), the entire SMA, from this large segment to the free end, shifts 5 mm. If there were any segments of length 5mm, then they would not be covering the same part of the SMA as they were before. This is detrimental if the smaller segment is turned on and the new material is shifted in. There will be a transient error.

As mentioned previously, this error can be reduced in the segmentation architecture. By asserting a fixed percentage by which each Peltier segment allows to shift and an expected strain value for the segments, equation 18 will specify the minimal length of the segment. Specifically a shifting percentage of

0. The entire length of the segment is not allowed to shift out from underneath (say, 0mm shift for a 5mm segment). This value is impossible to reach practically.
1. The entire length of the segment is allowed to shift out from underneath (say, 5mm shift for a 5mm segment)

$$L_i > \frac{\varepsilon}{P} \sum_{n=1}^{i-1} L_n \qquad (18)$$

ε: Expected strain for the SMA

L_i: Length of each segment. L_1 is the element closest to the fixed end.

P: Allowable shifting of SMA.

For example, for the initial conditions ($L_1 = 1$, $\varepsilon = 0.05$, $P = 0.25$) the minimum L_2 is 0.2. If L_2 is decided to also be unit length ($L_1 = L_2$), L_3 must be at least 0.4. Accordingly, if all the segments were equal to unit length ($L_i = 1$), the maximum amount of segments is 6. For any value P, Table 6 shows the maximum number of segments, as calculated in the equation 19.

$$Maximum\ Segments = 1 + \frac{P}{\varepsilon} \qquad (19)$$

Table 6. The maximum number of segments of equal length segments such that the constraint P is satisfied. (ε=0.05)

P	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Max Segments	1	3	5	7	9	11	13	15	17	19	21

3.7 Length of Smallest Segment Based on Expected Stress Range

Since SMA has a high power to weight ratio, the strain due the compliance when acting against a maximal load is significant. An undesirable characteristic is if the strain due to the compliance (in the expected stress range) is larger than the strain that the smallest segment can produce. In this extreme scenario, by looking at the displacement, one does not know if there was a load change or the smallest element changed states. This concealment of information is undesirable if there is no force feedback and position control is important. However, a constraint can be placed on the segmentation architecture to alleviate this problem.

By limiting the smallest element to be sufficiently sized such that strain that it produces is greater than the strain that is produced by a change in load can improve the problem addressed earlier.

Parameters are.

$\Delta\sigma$: stress change, $\Delta\varepsilon$: strain change, ε_o: desired strain due to the shape memory effect, E_M: Young's modulus of martensite, ΔL: length change, L_o: total length of SMA, L_S: length of smallest segment

Due to Compliance

$$\Delta\varepsilon = \frac{\Delta\sigma}{E} \qquad (20)$$

$$\Delta L = L_o \cdot \Delta\varepsilon = L_o \frac{\Delta\sigma}{E} \qquad (21)$$

Due to the smallest element

$$\Delta L = L_S \cdot \varepsilon_o \qquad (22)$$

Equating the equations (21) and (22):

$$L_S \cdot \varepsilon_o = L_o \frac{\Delta\sigma}{E} \qquad (23)$$

$$L_S = L_o \frac{\Delta\sigma}{E \cdot \varepsilon_o} \qquad (24)$$

Given E = 28 GPa (worst case), $\varepsilon_o = 0.04$, and unit initial length, $L_o = 1$

$$L_S = \frac{\Delta\sigma}{1.12\ GPa} \qquad (25)$$

The results of this equation are shown in Table 7. SMA's operating range for stress is between 0 and 200 MPa.

Table 7. The smallest segment (normalized to the total length of SMA) such that strain that it produces is greater than the strain that is produced by a change in load as a function of stress.

$\Delta\sigma$ [MPa]	0	25	50	75	100	125	150	175	200
L_S [ ]	0.000	0.022	0.045	0.067	0.089	0.112	0.134	0.156	0.179

3.8 Heat Transfer Errors

Segmented Binary Control is based upon selectively heating and cooling individual segments under the assumptions:

- All segments are thermally isolated from each other
- Only the area under the heating devices change temperature (and therefore phase) and not adjacent areas.

The answer to this question is based on many factors including the cross section of the SMA, boundary conditions, and the thermal conductivity. The values of the thermal conductivity constant of various metals are listed in Table 8.

Table 8. Thermal conductivity constants for various metals, including Ni-Ti (SMA).

Material	Thermal Conductivity (W/m K)
Ni-Ti (martensite)	~9
Ni-Ti (austensite)	~18
Silver	406.0
Copper	385.0
Aluminum	205.0
Brass	109.0

Ni-Ti (SMA) has a low thermal conductivity: 10 to 20 times less conductive than aluminum. In addition, with HCL most segments are at T_{CH} and T_{HC} when they are not transitioning. These two temperatures tend to be quite close in practice and therefore there will not be much heat transfer between them.

4 Implementation

Segmented Binary Control as well as Hysteresis Loop Control are based upon segmenting shape memory alloys and manipulating each segment's local temperature. Each segment's temperature must be able to increase and decrease, possibly below ambient temperature, and be maintained at specific temperatures. A rational choice is an external heater and cooler: a thermoelectric module (TEM) or a Peltier element. Utilizing numerous Peltier junctions thermally in parallel, it has the capability of heating and cooling the same surface based on the polarity of the driving current. TEMs have additional benefits of being solid-state, compact, have no moving parts, and ease of temperature control.

The TEMs must be in close thermal contact with the SMA but also allow for mechanical movement. These two contradictory requirements are resolved by attaching a TEM to a cast acrylic substrate board with a groove for the SMA. This gap between the wire and each TEM is sufficient for the wire to move smoothly but also to ensure good heat transfer. A laser cutter (Epilog legend 24TT) was used to fabricate the acrylic substrate board with grooves. Furthermore, double-sided thermal tape is placed along the SMA and the substrate grooves for bonding the TEM and the acrylic board together. Thermal grease is applied to the excess space around the SMA to ensure sufficient thermal contact and heat sinks are attached to the outside surface of the TEMs. A thermocouple is used for temperature measurements. A schematic of this sandwich structure for heating and cooling one segment is shown in Figure 12.

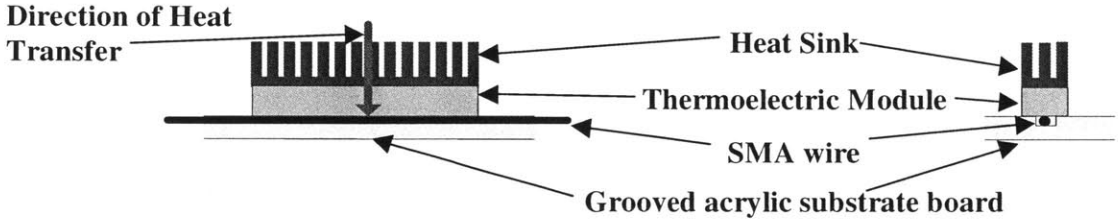

Figure 12. Schematic of one segment of an actuator array. Thermoelectric module on a grooved substrate board drives the segment of SMA actuator wire.

The experimental setup is shown in Figure 13. The system consists of one SMA wire, purchased from Dynalloy Inc., and eight heating/cooling TEM, with dimensions 30 mm by 5 mm, purchased from TE Technology. The SMA is fixed at one end and attached to springs on the other, creating the bias force. A linear potentiometer, from ETI systems, is attached to the end effector of the SMA for position measurement and a load cell, from Transducer Techniques, for force measurements.

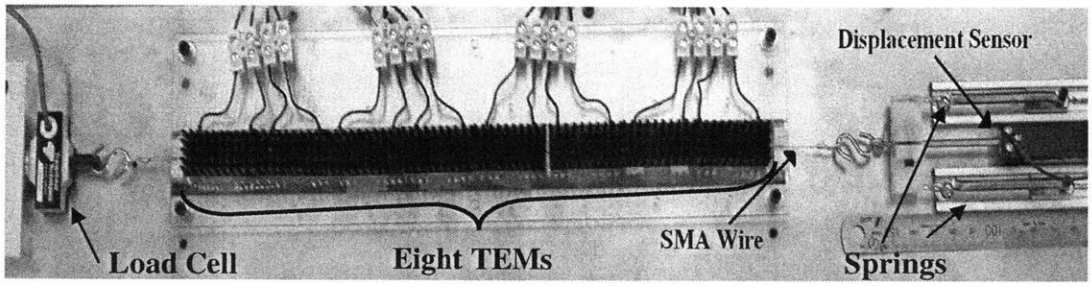

Figure 13. Experimental setup with eight segments.

Displacement and temperature measurements were obtained using a National Instruments Data Acquisition (DAQ) board, PCI-MIO-16E-1. Whereas the linear potentiometer created a signal directly connected into the DAQ, the eight thermocouple signals were processed in a cold junction compensation and amplifier chip, AD594CQ.

A graphic user interface (GUI) was used to display the measurements and control signals in real time. The software was designed in Visual Basic and utilized the NI Measurement Studio's active X controls. The control system consists of eight local ON-OFF controllers for the TEM and a global controller coordinating the individual segments. The block diagram of the controller has been shown in Figure 6.

A screen image of the control GUI indicating the temperatures, states of the individual segments, the total displacement of the entire SMA wire, and the control signal are shown in Figure 14. Based on the temperature measurement and the time history of commands given to each segment's local controller, the binary state of the segment is shown. The token moving around the discrete state network indicates whether the segment is ready for transition or still in transition to a new state. The segment coordinator selects which segment to fire in order to best follow the displacement command given to the system based on a First-In-First-Out algorithm. The time stamp

needed for selection is also shown in the GUI. In this figure, the first element has been ready to transition up for 13.4 seconds; the second has been ready to transition down for 23.7 seconds; and, the third and fourth are not ready for transitions and are in the cold states.

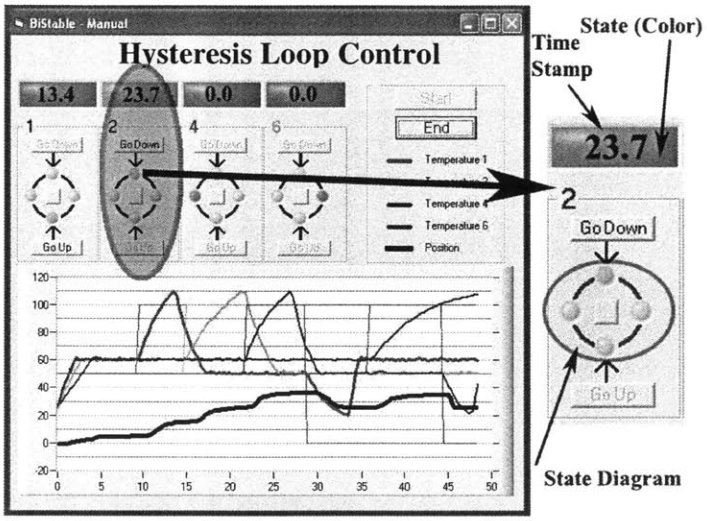

Figure 14. The GUI used for display and control for SBC and Hysteresis Loop Control.

The control signals from the software are interfaced to the hardware using another National Instruments board, PCI-6527. The digital outputs from this board are used as signals to the current drivers for the TEMs that consist of sixteen sets of one MOSFET and one relay, creating sixteen bipolar current sources.

5 Experimentation

Experiments have been conducted using the apparatus described above. Five major issues to be addressed through experiments are A) verification that SBC causes stepper motor-like behavior; B) a comparison of Hysteresis Loop Control and Segmented Binary Control; C) evaluation of inter-segment coordination for improving speed of response and bandwidth; D) verification of the stress margin algorithm and trade-offs between robustness and speed of response; and, E) HLC loading and errors.

5.1 Segmented Binary Control – SMA Stepping Motor

The ideal implementation of Segmented Binary Control would involve accurate discrete positioning regardless of outside influences. When a single heating element is turned "ON", by raising the temperature to the hot temperature threshold, a known and constant displacement should be achieved independent of the load. As an aggregate process, heating any number of heating elements would create a discrete position.

An experiment was designed to investigate the effect of loading on discrete positioning that SBC can provide. The five heating elements were sequentially heated to the hot temperature threshold temperature of 140°C and a position reading was recording. This experiment was performed for seven different loading conditions. The results are shown in Figure 15.

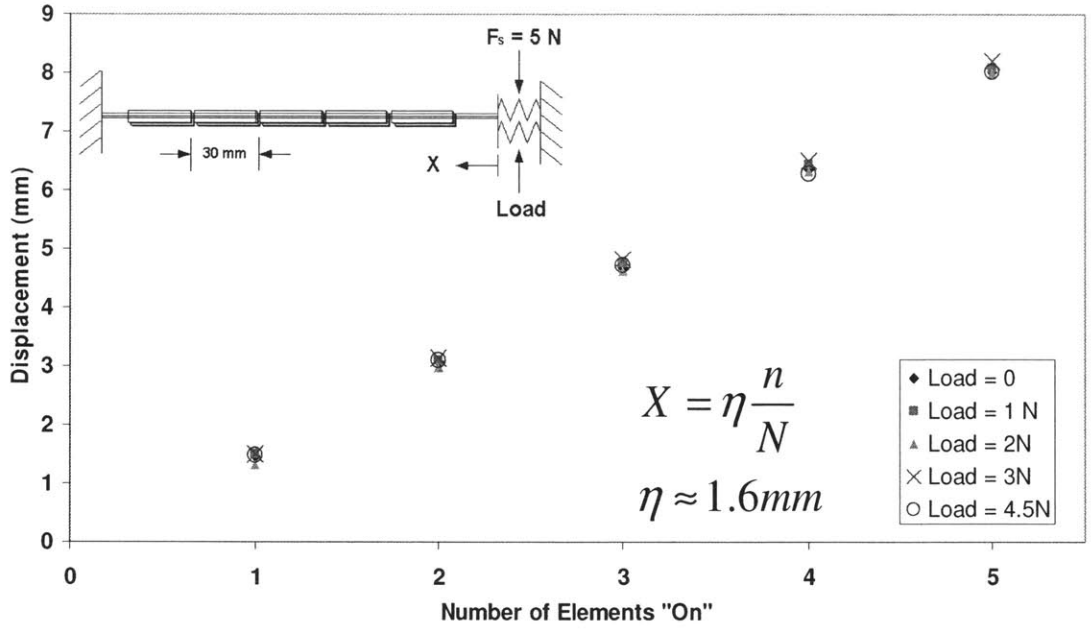

Figure 15. Linearity of SBC and HLC for various loads.

For bias forces (proportional to stress) of 550 to 1000 grams, the graph is essentially linear and has a slope of 1.6mm (5.3%) per element "ON". This means that discrete positioning can occur at multiples of 1.6mm and reasonable loading conditions with the current setup. There are minor differences in the points due to elastic strain. As an aggregate process, this resembles a stepper motor.

5.2 Hysteresis Loop Control and Segmented Binary Control

The HLC method has many advantages with respect to speed of response and repeatability compared to the original SBC. In particular, the latency times before strain is generated are significantly reduced. Figure 16 shows time responses of a single segment SMA using SBC and HLC. In these experiments, a SMA wire with diameter 0.254 mm and a A_S temperature of 70° C was used and was thermally cycled between 22°C (T_C) and 100°C (T_H). The top graphs show the segment temperature and the control command (on the right vertical axis), and the bottom graphs show the same control command and the resultant displacement of the SMA.

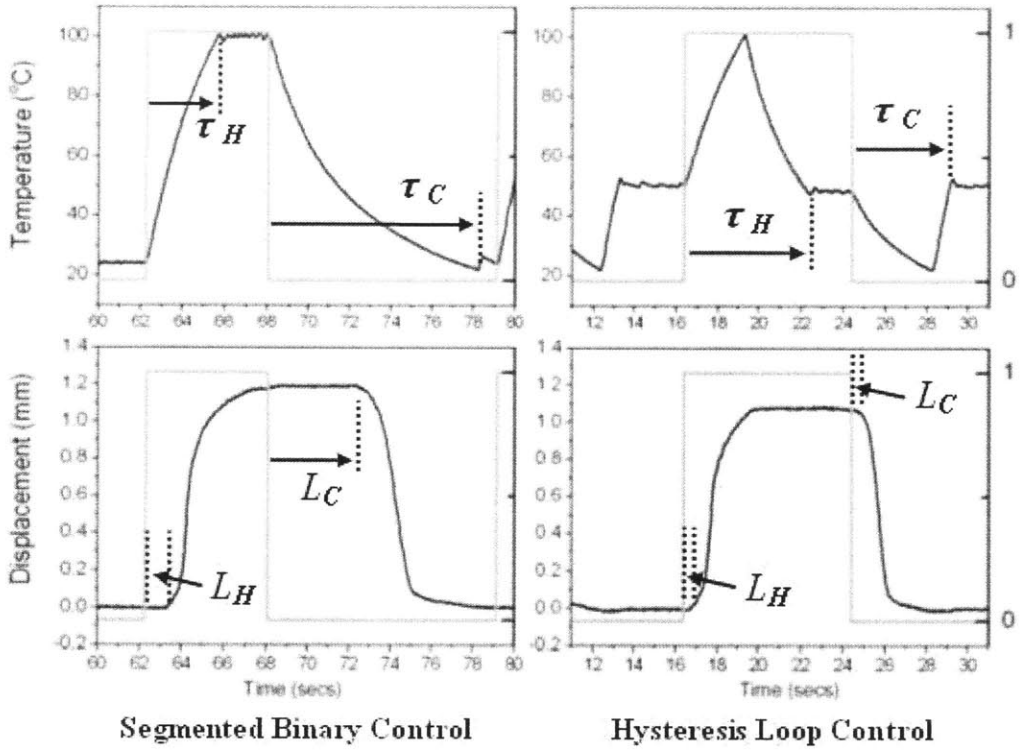

Figure 16. Control command, temperature, and displacement data using SBC and HLC.

The right two plots in Figure 16 are time profiles using HLC. In this experiment, the same diameter SMA wire was used as the previous experiment, 0.254 mm, and was thermally cycled with characteristic temperatures T_C, T_{CH}, T_{HC}, and T_H of 50°C, 100°C, 48°C, and 22°C, respectively. These temperatures were chosen to minimize the latency times. The additional intermediate temperatures, T_{HC} and T_{CH}, aid in reducing the latency times thus creating faster transitions after a control command. In this time response, the SMA displacement began to increase quickly having a latency time of only L_H = 0.2 second, and began to decrease with L_C = 0.3 second.

The temperature-time response, the top plot in Figure 16, indicates parameters τ_H and τ_C : the times needed for the segment to become "ready" for executing a next transition command. These time periods, τ_H and τ_C, represent inhibitory periods for the HLC control system being unable to respond to input commands without substantial latency times. In the above experiment the inhibitory periods are τ_H = 6.0 seconds and τ_C

= 4.6 seconds. A summary of the pertinent information from the SBC and HLC experiments is shown in Table 8.

Table 8. SBC versus Hysteresis Loop Control.

	SBC	HLC
L_H [sec]	1.1	0.2
L_C [sec]	4.5	0.3
τ_H [sec]	3.6	6.0
τ_C [sec]	10.2	4.6

5.3 Segment Coordination with Hysteresis Loop Control

As demonstrated in the previous experiments, Hysteresis Loop Control significantly reduces the latency times. However, the long inhibitory periods prevent the HLC control system from rapidly responding to alternating input commands. This drawback is particularly serious when tracking a square-wave trajectory as shown in Figure 5. The inter-segment coordination scheme described in the previous section can solve this problem.

Figure 17 shows an experimental comparison of square-wave trajectory controls with and without segment coordination. In both experiments, the control systems were commanded to move two points back and forth as many times as possible. Using segment coordination the HLC control system shown in the bottom plot can cycle through the trajectory 1.8 times more than that of the top plot without segment coordination. The one without segment coordination used a single segment that turns on and off consecutively, while the HLC with segment coordination used three segments, among which a segment ready for transition was selected and turned on or off. Both experiments used a thick SMA wire of 0.508 mm diameter with an A_S temperature of 90°C.

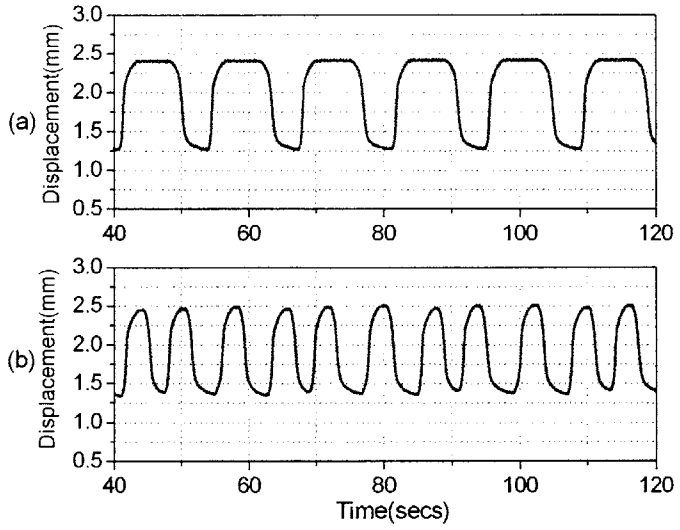

Figure 17. (a) SBC versus (b) coordinated HLC.

Figure 18 shows the detailed plots of the HLC control system using the First-In-First-Out segment coordination scheme. At the beginning Segment 1 is at T_{HC}, and Segment 2 and 3 are at T_{CH}. The temperature of Segment 2 starts to increase to T_H. As soon as Segment 2 reaches T_H, Segment 1, which is at T_{HC}, starts cooling down immediately. This coordination avoids the delay of cooling down Segment 2 from T_H to T_{HC}. When Segment 1 cools down to T_C, Segment 3, which is already at T_{CH}, transitions up. This coordination avoids the delay of heating Segment 1 from T_C to T_{CH}. A similar process can continue to repeat the displacement cycles.

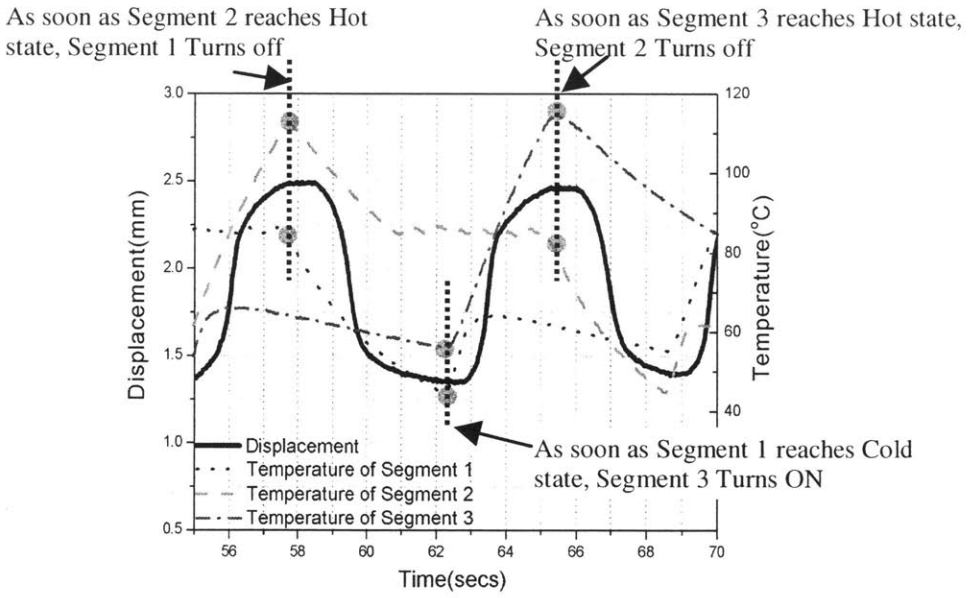

Figure 18. Detailed temperature and displacement graph of Hysteresis Loop Control experiment.

By consecutively coordinating the multiple segments the HLC control system can increase the bandwidth approximately 1.8 times greater than that without segment coordination. This increased bandwidth stems from the fact that only the slopes of the hysteresis loop must be traversed to complete a transition and not the width of the hysteresis, as illustrated in Figure 2.

5.4 Stress Margin with Revised Segmented Binary Control

The experiment described in Section 5.2 demonstrated the minimization of the latency times without considering varying loads for HLC. Following the procedure for tuning stress margins in Section 2.2.2 one can gain insights as to the tradeoff between minimizing the latency time and enhancing load disturbance rejection. The following experiment verifies improved disturbance rejection assured with stress margins.

Given minimum and maximum stresses of 70 and 130 MPa, respectively, a SMA wire of diameter 0.254 mm and an A_S temperature of 70°C was thermally cycled to generate the strain-temperature hysteresis curves under the maximum and minimum stresses. Then the characteristic temperatures were selected in accordance with the Stress

Margin procedure. These plots are shown in Figure 19 and the selected temperatures are listed in Table 9.

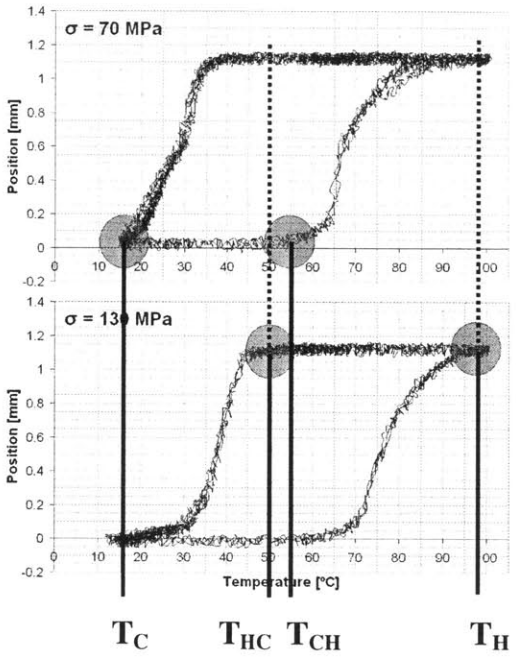

Figure 19. The temperature-position hysteresis for the maximum and minimum stress.

Table 9. The four characteristic temperatures for SBC, chosen from the hystereses above.

	T_C	T_{CH}	T_H	T_{HC}
Temperature	15 °C	55 °C	100 °C	50 °C

For each chosen characteristic temperature, marked with the circle in Figure 19, notice where that temperature exists on the temperature-strain plots shifted by the different stresses: each temperature holds without any danger creating unwanted partial transitions. However, there will be a greater latency time. Consider T_{HC}, chosen to be 50°C from the bottom graph of Figure 19 so that it is sufficiently near the transitional portion of the hysteresis curve for maximum stress of 130 MPa. Compare to the top graph with stress 70 MPa. In this graph, because of the stress induced shifting, the transitional portion of the hysteresis curve is around 40°C and the previously chosen T_{HC} is no longer

near the transitional hysteresis. Therefore given static characteristic temperatures, there will be a greater latency time for a stress of 70 MPa instead of 130 MPa. A similar argument can be made for T_{CH}.

Given the HLC characteristic temperatures listed in Table 9 (valid for a stress range between 70 and 130 MPa), HLC is implemented. Figure 20 shows the results. While the stress is changed from 130 MPa to 100 MPa and then 70 MPa, one SMA segment is cyclically turned "ON" and "OFF", and temperature and displacement data are taken. Compliance effects of the changing load are negated so that the constancy of the shape memory effect can be investigated.

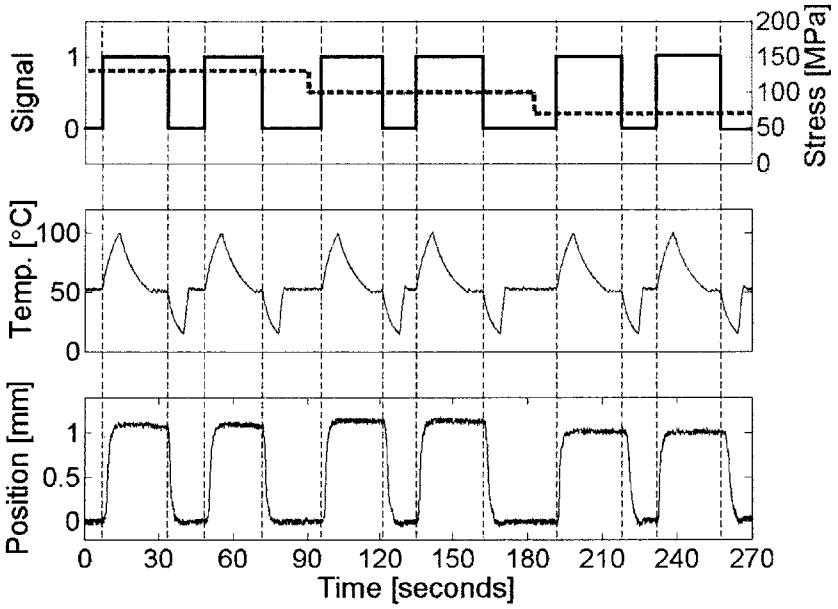

Figure 20. Experiment the constancy of the shape memory effect utilizing SBC for three different stresses.

Despite the changing stress, constancy of the shape memory effect remains. For stresses of 130, 100, and 70 MPa the average position is 1.1, 1.15, and 1.05 mm, respectively. Whereas the latency times associated with positive transitions (heating) ranges from 1 to 0.25 seconds, the latency times associated with negative transitions (cooling) ranges from 2 to 0.25 seconds. At any stress, there is a tradeoff between the two

latency times. At the extreme stress cases, whereas one latency time is minimized, the other is maximized. These results are summarized in Table 10.

Table 10. Summary of important data in Figure 15.

Stress [MPa]	130	100	70
Position [mm]	1.10	1.15	1.05
Latency time heating [sec]	1.0	0.5	0.25
Latency time cooling [sec]	0.25	1.0	2.0

5.5 Sources of Error

Figure 21 below shows experimental data shows this trend of vertical and horizontal shifting due to stress. In the experiment, a SMA of diameter 0.254 mm and A_S temperature of 70°C was loaded with an initial stress of 155 MPa and one of seven segments was thermally cycled three times. When the segment was subsequently turned off, the stress was reduced to 60 MPa and thermally cycled three additional times.

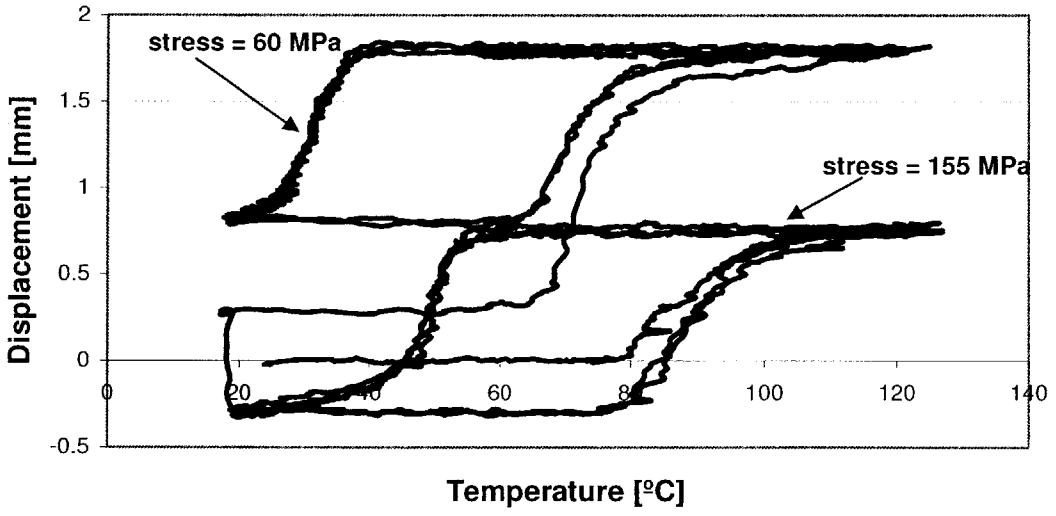

Figure 21. The SMA hysteresis for two different loading conditions.

Notice that the vertical discrepancy between the two hysteresis curves, the magnitude of the strain due to the SME, and hysteresis width are all nearly constant.

However, the steepness of the transitional sections does change noticeably. Also to note, the first cycles at both loads are significantly different than all other cycles and should be ignored in analysis.

By examining the figure above, it is obvious that the first cycle at a load is significantly different from the subsequent cycles, which follow a rigid temperature hysteresis. This occurs due to detwinning within the martensite phase. When changing loads, the martensite may be detwinned too much or too little for the new load. After the first cycle, the SMA finds its equilibrium. This leads to the conclusion that after every load change, the SMA must be reinitialized. This way, all subsequent cycles will act deterministically.

6 Discussion

The prototype system developed for experimental evaluation has demonstrated the feasibility and unique features of the proposed Hysteresis Loop Control and inter-segment coordination. There are, however, a few critical points needing further consideration.

One of the fundamental premises of the proposed method is that an SMA wire can be heated and cooled locally and selectively. Due to thermal conductivity along the wire, however, it might be troublesome to regulate adjacent segments at different temperatures. It turns out that this is not a fundamental limitation to this method for two reasons. One is that the thermal conductivity of SMA is very low, between 10 and 18 W/m K for Nickel-Titanium SMA [21]: more than ten times less conductive than aluminum. Therefore, it takes some time for one segment to influence its adjacent segments if they are separated with a gap, say 1 mm. Second, although segments are heated and cooled to extreme temperatures, T_H and T_C, they immediately return to the intermediate hold temperatures, T_{CH} and T_{HC}. In most cases, the two hold temperatures are very close to each other. Therefore, heat transfer between adjacent segments is negligibly small. This is another advantage of HLC over the basic SBC.

One potential limitation to the experimental apparatus built with Peltier effect TEDs is that segments of a SMA wire may shift to adjacent TED units, as the SMA wire shrinks and expands. This may cause some error when an adjacent segment at a very different thermal state is brought to the neighboring segment. The worst case scenario is that every pair of adjacent segments takes different thermal states. The effect is negligible at segments close to the fixed end of the SMA wire, but is most prominent at the other end. For the prototype built with 8 units of 30 mm long TEDs and a SMA wire of maximum 4 % strain, only 9 % of the SMA wire having different thermal states may be brought to adjacent TED units even for the worst case scenario. In average, the effect is a few percent. As the SMA wire length gets longer, this segment shift problem becomes more crucial. For long wire applications, design should be changed so that heating and cooling units may move together with the SMA wire.

The Hysteresis Loop Control of segmented SMA actuators has been implemented by using a standard SMA wire. These SMA actuator materials available on the market have been designed for conventional analogue controls, so that the linear range of the temperature-strain curves may be wide and that the hysteresis band be minimal. However, these characteristics are the exact opposite of what is desired in Hysteresis Loop Control. A highly nonlinear hysteresis with steep transition slopes is preferred to minimize the time to complete a transition.

Three characteristics of an ideal SMA actuator hysteresis for HCL are:

- Linear slope
- Moderate slope
- Minimal width

The linear slope is important for any kind of position control so that the amount of strain is predictable for a certain temperature and because linear control methods can be implemented.

A moderate hysteresis slope is needed to prevent any high accuracy temperature control. If the slope is too sensitive to temperature variations, control would be much more difficult. Consider the extreme case shown in second graph in Figure 22. The slope is so steep that temperature control, regardless of precision, could not mediate any intermediate strain.

The T-ε hysteresis width prevents quick changes of direction of strain. For example, if there is overshoot strain, the temperature must be reduced to the other side of the hysteresis to diminish the strain. Similarly, for a subsequent increase in strain, the SMA must be heated again, across the hysteresis. This hysteresis, in conjunction with slow cooling rates, is the limiting factor in SMA actuator bandwidth.

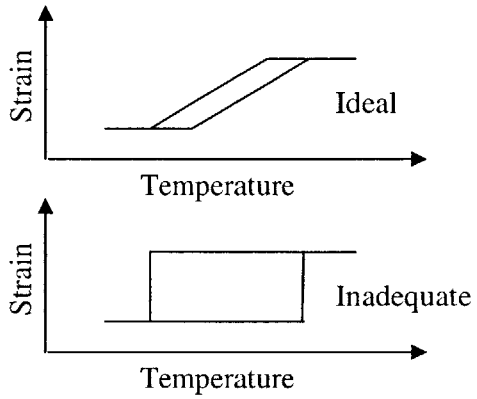

Figure 22. Schematics of two potential hysteresis curves.

For a general SMA actuator, the top figure is better in all aspects discussed above. However, the ideal SMA actuator characteristics for HLC have characteristics similar to the bottom graph of the above Figure 22: the curve that is insufficient for general SMA actuator design.

An ideal hysteresis curve for SBC is highly nonlinear: a stark contrast to the general case. This schematic and the corresponding SBC characteristic temperatures are shown in the Figure 23 below.

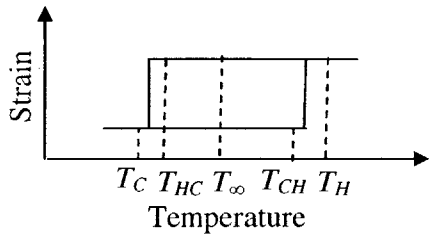

Figure 23. A desirable hysteresis curve for HLC.

At the core of the SBC and HLC ideal hysteresis loop is the minimization of energy: a minimum amount of energy is desired for every transition. This is easily accomplished if the temperature span for transitions is small and ambient temperature is at temperature contained inside the hysteresis. This way, after using only minimal energy for transitions, the temperature automatically begins to return to the hysteresis. In addition, not much energy is needed to return the temperature to T_{CH} and T_{HC}.

The desirable characteristics of an SBC SMA are:

1. Very steep transitions. This corresponds to $M_S \approx M_F$ and $A_S \approx A_F$ or $T_H - T_{CH} \approx Minimum$ and $T_{HC} - T_C \approx Minimum$.
2. Saturated strain at extreme high and low temperatures.
3. Symmetry between both transitions.
4. The midpoint of the hysteresis is at ambient temperature.
5. Low stress rate.
6. Martensite and Austenite phases similar Young's Modulus.

7 Conclusion

The segmented SMA actuator using Hysteresis Loop Control and inter-segment coordination has been presented. The new approach fundamentally differs from conventional single-process analogue controls in architecture, control methodology, and desired material properties. The major contributions are:

- New SMA actuator architecture based on segmentation and bi-stable digital control has been developed.
- Behavior similar to a stepping motor, discrete positioning, is observed.
- Exploiting the Schmidt-Trigger type bi-stable switching characteristics of SMA, the Hysteresis Loop Control method has been developed for improving speed of response and power saving.
- The First-in-First-out algorithm for coordinating a multitude of segments has been developed to minimize latency times in tracking alternating trajectories and thereby to improve the system bandwidth.
- Stress margins have been introduced to prevent individual segments from making unintended phase transitions due to stress changes.
- The segmented SMA architecture and HLC control with inter-segment coordination have been implemented by using Peltier effect thermoelectric devices for selective, local heating and cooling of individual SMA segments.
- The unique features of the proposed method have been demonstrated based on experiments using the prototype apparatus.

The effectiveness of the proposed method can be further improved once the properties of SMA are changed to be more nonlinear with steeper hysteresis curves. This will point in a new direction of SMA material development for the future. In addition, the proposed approach of segmented SMA actuators will be practically more useful if a methodology for reducing the number of local segment controllers is developed. As the number of actuator axes increases, such reduction of control complexity becomes more important. For a humanoid hand with a 10-axis SMA actuator array, the authors' group

has developed an effective methodology for reducing the number of local segment controllers [22].

References

[1] Hunter, I., S. Lafontaine, J. Hollerbach, and P. Hunter. 1991. "Fast Reversible NiTi Fibers for Use in MicroRobotics," Proc. 1991 IEEE Micro Electro Mechanical Systems—MEMS '91, Nara, Japan, pp. 166–170.B

[2] Hunter, I.W., and S. Lafontaine. 1992. "A Comparison of Muscle with Artificial Actuators," Technical Digest of the IEEE Solid-State Sensor and Actuator Workshop, Hilton Head, South Carolina, pp. 178–185.

[3] Pelrine, R. and R. Kornbluh, SRI International, 1999, by personal communication.

[4] Park, S., and T. Shrout. 1997. "Ultrahigh Strain and Piezoelectric Behavior in Relaxor Based Ferroelectric Single Crystals," J. Applied Physics, Vol. 82, pp. 1804–1811.

[5] Maeda, S., Abe, K., Yamamoto, K., Tohyama, O., Ito, H, "Active endoscope with SMA (shape memory alloy) coil springs," Proceedings of the IEEE Micro Electro Mechanical Systems, pp.290-295, 1996.

[6] Lee, Y., Kim, B., Lee, M., Park, J. "Locomotive mechanism design and fabrication of biomimetic micro robot using shape memory alloy," IEEE International Conference on Robotics and Automation, v 2004, n 5, 2004 IEEE International Conference on Robotics and Automation, 2004, pp. 5007-5012.

[7] Michaud, V., Schrooten, J., Parlinska, M., Gotthardt, R., Bidaux, J."Shape memory alloy wires turn composites into smart structures," The International Society for Optical Engineering, v 4698, 2002, pp. 406-415.

[8] Amalraj, "Finite-element modeling of phase transformation in shape memory alloy wires with variable material properties," Smart Mater. Struct. **9**, pp. 622-631, 2000.

[9] Khan, "Modeling of Shape Memory Alloy pseudoelastic spring elements using Preisach model for passive vibration isolation," Proceedings of SPIE - The International Society for Optical Engineering, v 4693, pp. 336-347, 2002.

[10] Benzaoui, "Experimental Study and Modeling of a TiNi Shape Memory Alloy Actuator," Journal of Intelligent Material Systems and Structures **8**, pp. 619-629, July 1997.

[11] D. Grant, "Accurate and Rapid Control of Shape Memory Alloy Actuator", Thesis of degree of PhD, McGill University, 1999.

[12] Song, "Precision tracking control of shape memory alloy actuators using neural networks and a sliding-mode based robust controller," Smart Mater. Struct. 12, pp. 223-231, 2003.

[13] Kumagai, "Neuro-fuzzy model based feedback controller for shape memory alloy actuators," Proceedings of SPIE, v 3984, pp. 291-9, 2000.

[14] Gorbet, "Dissipativity approach to stability of a shape memory alloy position control system," IEEE Transactions on Control Systems Technology, v 6, n 4, pp. 554-562, July 1998.

[15] Grant, "Variable structure control of shape memory alloy actuators," IEEE Control Systems Magazine, v 17, n 3, pp. 80-88, June 1997.

[16] Semenyuk, V., Stockholm, J., Musolff, A., Seelecke, S, "The Use of Thermoelectric Cooling for Shape Memory Wire Temperature Control," Proceedings of the ETS Workshop 98, 1998.

[17] Khan, M, Lagoudas, D., Rediniotis, O., "Thermoelectric SMA actuator : preliminary prototype testing," Proceedings of SPIE, v 5054, pp. 147-155, 2003.

[18] J.D. Harrison, "Measurable Change Concomitant with SME Transformation," Engineering Aspects of SMAs, eds. Duering et al., Butterworth, pp 106-209, 1990.

[19] Abel E., Luo H., Pridham M., and Slade A, "Issues concerning the measurement of transformation temperatures of NiTi alloys," Smart Materials and Structures 13, pp.1110-1117, 2004

[20] Ducrig, Engineering Aspects of SMA, eds. Duering et al., Butterworth, pp 16, 1990.

[21] Otsuka, K., Wayman, C., Shape Memory Materials, Cambridge University Press, pp 174.

[22] Cho, K.-J., and Asada, H., "Multi-Axis SMA Actuators Array driving Anthropomorphic Robot Hand", Accepted for publication in the proceedings of the 2005 IEEE International Conference on Robotics and Automation, 2005.