Controller Design for Human-Robot Interaction

Eric Meisner Volkan Isler Jeff Trinkle

September 13, 2007

Department of Computer Science
Rensselaer Polytechnic Institute
110 8th Street Troy NY 12180

Abstract

Many robotics tasks require a robot to share the same workspace with humans. In such settings, it is important that the robot performs in such a way that does not cause distress to humans in the workspace. In this paper, we address the problem of designing robot controllers which minimize the stress caused by the robot while performing a given task. We present a novel, data-driven algorithm which computes human-friendly trajectories. The algorithm utilizes biofeedback measurements and combines a set of geometric controllers to achieve human friendliness. We evaluate the comfort level of the human using a Galvanic Skin Response (GSR) sensor. We present results from a human tracking task, in which the robot is required to stay within a specified distance without causing high stress values.

1 Introduction

Recent decades have witnessed tremendous success of robotics technologies in factory automation tasks. As the field of robotics continues to advance, robots will be utilized in many automation tasks involving humans. Examples of such tasks include robotic shopping carts [KG06] or guide-dogs for the visually impaired and robotic nurses that assist surgeons or care for the elderly [FND02] . In most of these tasks, it is important that the robots carry out their assignments in a "human-friendly" fashion. In other words, both physical and emotional well-being of the human should be assured by the robot's controller.

As an example, consider a scenario where a robotic dog, guiding a visually impaired person, must plan a path from point a to point b. Among the (infinitely) many paths

between a and b, what is the "best" path for guiding the person? Is it the shortest path or the one that maintains a minimum curvature? Is it a path that maintains a large clearance around obstacles to reduce the anxiety of the human? What is the right velocity and acceleration? Should the robot speak to the human or provide some other sort of sensorial input during the task? In what ways are humans sensitive to the path parameters and other environmental factors? Do they adapt easily over time? We believe that answering such questions will be crucial to determining the acceptability of robotic systems in tasks involving humans.

The primary objective of our research is to develop techniques for the design and implementation of "good" human-friendly robot controllers. In this work, we investigate the use of biofeedback techniques to accomplish these goals. We set out with the hypothesis that the human-friendliness of a controller can be quantified by measuring the physical manifestations of human emotions (*e.g.*, skin conductance and heart rate, called "sentic modulations" by Clynes [Cly97]) while s/he is interacting with the robot. According to Picard, emotion theorists suggest that one can provide input to the human visual, auditory, tactile, and other sensorial systems that affect the emotional state of the person, and hence their comfort level [Pic95].

The ability to gather biometric data opens up the possibility of using data-driven techniques to learn controllers that achieve desired comfort levels while simultaneously accomplishing a task. Even more, if an accurate mapping between emotional inputs and outputs can be identified, we should be able to design controllers that allow efficient task completion while also heightening the human comfort level, or at least improve the comfort level quickly after the task is completed. Obtaining such methods will be a significant step toward utilizing robots in important domains such as health-care and entertainment.

In this work, we demonstrate the feasibility of this approach. We focus on the problem of designing motion plans for a representative human tracking task. The robot's goal is to maintain proximity with the human without causing much stress – measured from a Galvanic Skin Response (GSR) sensor. Before we present technical details, we start with an overview of related work.

2 Related work

There are many factors complicating the design of human-friendly robots, not the least of which is the lack of agreement on a definition of human-friendliness of a robot. However, the consensus of emotion theorists is that the human sensory systems provide input to

human emotional states, also known as affective states, and that human emotions are expressed through the human physical system.

Clynes concludes that the emotional and physical systems of the human are inextricably coupled [Cly97]. Further, Clynes found that it was possible to reliably induce an emotional state through music and touch, and that people could quickly learn to express or display emotional states accurately by pressing on a tactile sensor. Thus, it is reasonable to hypothesize that the friendliest controllers will not only modulate the mechanical inputs to the human (*e.g.,* forces applied to joints), but also sensory inputs to the human (*e.g.,* facial expressions and voice inflections). Further, this modulation should be based on the physical and emotional state of the human. In the present work, rather than try to directly control the sensor inputs to the human, we focus in using the affective state of the human as evaluative feedback to our controller.

As a result of the above considerations, it is not surprising that a significant amount of work in Human-Robot Interaction (HRI) has focused on understanding the role of robot appearance and speech. One could classify these studies into two groups: open-loop or closed-loop. Closed-loop controllers receive feedback about the state of the system. This information is used to control the output of systems with unmodeled disturbances. Open-loop controllers do not use feedback, and rely entirely on the system model. The open-loop study described in [HRJ04] suggests that people prefer sharing responsibility with humanoids rather than machine-like robots. Roy *et al.* designed a spoken dialogue system for a nursing home assistant [RPT00]. The system achieves good performance with a representation that infers the user's intentions. Lesh *et al.* demonstrated how a plan recognition algorithm (which uses cues such as focus of attention in addition to speech) reduces the amount of communication required by the user [LRS99]. More recently Mitsunaga *et al.* demonstrated the use of gaze tracking to estimate the affective state during human-robot interaction [MSK+05]. Efforts in the closed-loop category are described in [ADOB97] and [Pen04]. Recent work has utilized biofeedback to understand a human's emotions, which can then be used to endow robots with "emotional intelligence." For example, in [RS05] the authors propose the display of emotion as a means of communication with the human. The work in [RSSK04] uses physiological data as a means of implicit human coordination in a human-robot coordination task. The work in [BA07] attempts to achieve non-verbal communication by managing personal space (proxemics) and body language. In [PCJ05, CP05] the authors define robot control strategies for embodied interaction based on human spacial behavior studies such as Hall's proxemics framework. In [RLS06], the authors demonstrate that a robot can use the stress of a human as input to a basketball game with a moving rim. Similarly,

in [BMA+01] the authors present a game where the players can control the speed of their racing "dragons" by lowering their own stress. During the game, stress values are obtained by taking Galvanic Skin Response (GSR) measurements. However, in all of this previous work, even though the controller utilizes biofeedback, it is hand-coded. That is, the design of the controller does not have the explicit objective of making it human-friendly.

Another important factor affecting the human-friendliness of a robot is *how* a robot accomplishes a given task. This controller design problem, which is the focus of the present work, has not received much attention in comparison to the factors discussed above. The reason perhaps lies in the difficulty of designing controllers for HRI. Note that HRI tasks usually take place in dynamic and complex environments where physical and emotional inputs are intermixed. Designing robust controllers for such environments is one of the most challenging and active research areas in robotics. Consequently, HRI researchers have restricted their investigations to instances where existing, well-established controllers can be utilized. For example, one of the state-of-the-art systems for assisting the visually-impaired works only in environments where existing localization and motion planning algorithms work robustly [KGNO06]. The present work differs from this line of work in terms of including an explicit human-friendliness objective into the design of the controllers.

While considerable work has been done in modeling affective state, there is no agreed-upon model. Picard's paper contains an excellent survey of the relevant emotion literature [Pic95]. It demonstrates the importance of emotion in rational decision making and perception, and discusses methods for observing and manipulating the user's emotional state. Both discrete and continuous models have been developed. Models generally employ from two to twenty basic states, with the most commonly agreed upon subset being fear, anger, sadness, and joy, but most researchers add a few more to these [Pic95]. For example, Plutchik uses eight basic states; the four just given plus disgust, acceptance, anticipation, and surprise [Plu80]. Continuous models can be obtained simply by associating an axis of an emotion space with each emotion of interest. For example, one could assign joy to an axis and consider negative values of joy as sorrow. In the emotion literature, a value from a continuous range of emotion is called the "valence." In this work we infer the affective state of the human through a one dimensional "comfort" parameter. We obtain the value of this parameter through skin resistance measurements. Under controlled settings, and in our specific experiment, the skin resistance is strongly correlated with the comfort level.

3 Motivation

In this section we describe a preliminary experiment which demonstrates the utility of using biometric feedback to design human friendly controllers. In this experiment we monitor the galvanic skin response of a human subject while walking a short path. In the mean time, a robot is tasked with traveling along a straight line to a goal location. The paths of robot and human are perpendicular and cross at the halfway point of the robots path and at 2/3 the length of the human's path. The problem is to find a motion strategy for the robot that causes low GSR response. The length of the human and robot paths are roughly 3.8 and 2.4 meters respectively. We record the GSR measurement of the human at 4 hertz for four different robot motion strategies, each of which reaches the goal while avoiding the human. The time units are .25 seconds each. The four controllers are described as follows:

- **GO**, which causes the robot to increase to its maximum velocity and maintain it until the goal is reached. The robot moves through the intersection point ahead of the human.
- **STOP** where the robot initially increases velocity, then slows and stops halfway between the start and the path intersection, waits for the human to pass the intersection point, and continues at maximum velocity to the goal.
- **WAIT** causes the robot to wait for 2 seconds before executing the strategy described in **GO**. The human moves past the intersection point before the robot reaches it.
- **SLOW** causes the robot to move at $\frac{1}{5}$ of its maximum velocity, starting immediately and ending at the goal. This allows the human to pass through the intersection point well before the robot reaches it.

The plots in figures 1 and 2 show the GSR reading of each of the four controllers, executed twice on the same individual, roughly 90 minutes apart. This is provided to illustrate that the change in response can in fact be attributed to the robot's actions and that the affect from the robot behavior is repeatable. During this experiment, we asked the test subject to take one breath and hold it for the length of the experiment. We have observed that some changes in GSR reading are correlated with breathing patterns, and we wish to eliminate the effects of varying breathing patterns on the observed response. We assume that the trials are short enough that this will not cause any adverse effects. In order to control for the affective state of the human at the start of a trial, we consider the

change in GSR from the start of the experiment, and give special attention to significant changes in response. GSR does not provide a direct mapping to comfort or stress, but by taking the aforementioned steps to control the environment, we can safely attribute the changes in GSR to the actions of the robot.

The response from the **GO** controller is shown in figure 1(a). The robot accelerates to .5 m/s, at a rate of .4 m/s^2, reaching the intersection point at $t = 8$ (i.e. 2 seconds). It begins slowing at $t = 12$, and reaches the goal and stops a $t = 16$. The response for both the first and second iterations of the GO controller show a significant increase in stress beginning at $t = 9$ and flattening near $t = 15$. Note that the shape of the response curves in 1(a) are similar even though the initial readings are quite different.

The response curves for the **STOP** controller are shown in 1(b). The robot accelerates to .3 m/s at a rate of .4 m/s^2, begins slowing at $t = 5$, and stops at $t = 7$. The robot then waits for .5 seconds as the human passes through the intersection point, then accelerates to .5 meters per second at .4 m/s^2, starting at $t = 10$ until $t = 15$. The robot begins slowing at $t = 18$, reaches the goal and stops at $t = 22$. For both plots in figure 1(b) , there is a minor increase and drop-off before $t = 7$, followed by a significant increase around $t = 10$. There is a general flattening of the response curve around $t = 18$. The test subject expressed that this was the least desirable controller of the four tested.

The **WAIT** controller causes the robot to do nothing until $t = 8$ when it begins accelerating, reaching full speed at $t = 12$. It start slowing at $t = 18$, stopping at $t = 22$. The robot passes through the intersection point slightly behind the human at $t = 14$. In both trials, there is an increase at $t = 14$ when the robot passes behind the human, and another at $t = 18$ when the robot begins to slow.

The **SLOW** controller, which moves the length of the path at .1 meters per second, crossing the intersection point at $t = 25$, gave the lowest stress response of all the controllers. This is evident when it is compared to the two baseline stress response plots in figure 3 .

Each of the four controllers has the same instrumental function, which is to make the robot reach the goal while avoiding the human. The results of these experiments confirm several concepts important to our work. First, is the idea that using geometric goals such as collision avoidance as design criteria for human-friendliness is not sufficient. This is demonstrated by the fact that each controller achieves collision avoidance but all have very different GSR profiles. The second important concept is that of repeatability. These preliminary experiments suggest that for an individual, repeating the same robot behavior, will evoke a repeatable stress response under controlled settings. This concept is essential to the problem of learning through repeated trials. Finally, the similarity of

the response from different trials with the same controller, and the correlation between robot actions and increased GSR shows that the robot behavior does in fact affect the stress level of the human.

4 Contributions

In the previous section we examined the problem of selecting the most human-friendly controller from a set of controllers that all perform the same instrumental task. In general, it is not possible to enumerate and evaluate all controllers for a particular task. For this reason we propose the following novel approach to creating human friendly controllers. Instead of utilizing existing controllers designed without regard for the human affective state, we utilize biofeedback as a means of evaluating the quality of controllers during the design process. Specifically, we incorporate GSR measurements as feedback in designing a controller for a human tracking task. In this task, the robot is charged with maintaining a specified minimum and maximum distance from the human. The goal of the robot is to move so that it maintains the appropriate distance while simultaneously minimizing the measured stress. This task is motivated by scenarios in which the robot may be required to stay close to the human. For example this may be done in order to track the human's gesture. A galvanic skin response (GSR) sensor is used to estimate the stress of the human participant. This sensor touches the human's skin at two points and measures the conductivity of the skin between the contact points. This measurement is typically used as a measure of the individual's excitement or relaxation. We assume that this is altered only by the interaction between human and robot. We formulate the tracking problem as a Markov Decision Process (MDP). We propose and implement a novel algorithm to compute policies for the MDP. The policies correspond to robot trajectories which simultaneously achieve the tracking objective and human friendliness.

Our contributions can be summarized as follows:

- A novel learning algorithm which utilizes biofeedback for designing human-friendly controllers
- An experimental setup for evaluating the robot controllers in a simple tracking task
- Evaluation of the algorithm on five test subjects.

5 Problem Formulation

We formulate the problem of computing human-friendly tracking controllers as that of computing an optimal policy for a Markov decision process (MDP). Recall that a finite MDP is a tuple $(S, A, \{P_{sa}\}, R, \gamma)$ where S is a finite set of states, A is a finite set of actions, P_{sa} is a probability distribution which returns the probability of transition to a particular state when action a is taken from state s, $R : S \rightarrow \mathbb{R}$ is the reward (or reinforcement) function, and $\gamma \in [0, 1)$ is the discounting factor, chosen to discount the value of future rewards.

The first step in building an MDP for the tracking task is to discretize the state space and action set of the robot. Let p_r denote the position of a reference point on the robot (given by its coordinates in the plane) and θ_r be the robot's orientation. The robot's state is then given by $x_r(t) = [p_r(t)\theta_r(t)]$. Its controllable input vector u_r consists of the robot's forward speed and turning angle. The human's state has two components: physical and emotional. To model the physical state of the human, let us assume that there is a single point of reference on the human. Let p_h and θ_h denote the position and orientation of this point of reference, respectively. The human's physical state is then given by $x_p = [p_h(t)\theta_h(t)]$. We assume the trajectory of the human is fixed, therefore x_p is given by a function of time $f(t)$. This is a strong assumption, however in practical applications, it may be possible to bound the speed of the human, and select controls which minimize the worst case output of the joint human-robot system. The affective state of the human at time t is denoted by $x_a(t)$. Let $g(x_p, x_a, x_r, u_r, t)$ be the general objective function which specifies the reward for a particular joint state of the human and the robot. We seek a control function $u_r(t)$ that maximizes

$$\sum_{t=0}^{t_f} g(x_p, x_r, x_a, u_r, t)$$

over the fixed time interval $[0, t_f]$. If the parameters of $g(\cdot)$ are known, then this strategy can be directly computed using dynamic programming and related techniques in the MDP framework. However, this requires a precise model of how human's stress level evolves as a function of all state values. Obtaining such a model is very difficult. A viable alternative is to use reinforcement learning techniques. Unfortunately existing algorithms such as value iteration require visiting all states of the MDP many times. This is very hard to achieve in human experiments as it too expensive in terms of time and energy. In the next section, we present a novel learning algorithm that is more appropriate for human-robot interaction experiments.

5.1 Learning Algorithm

In the previous section, we showed how the tracking problem can be modeled as an MDP. When the system dynamics (including the evolution of human's emotions) are known, an optimal or near-optimal policy for this MDP can be computed to design a human-friendly tracking algorithm. However, since we do not know the relationship between stress response and system state, it is not possible to directly solve for the controls $u_r(t)$.

At the heart of our approach lies the observation that for a fixed human trajectory and an initial position of the robot, what we are seeking is a robot trajectory, rather than a policy over the entire state space. To obtain such a trajectory, we first generate a number of candidate trajectories. These trajectories are obtained by computing optimal policies to various purely geometric objectives. The crucial point here is that these objectives do not involve human emotions and therefore corresponding policies can be computed off-line. We posit that a human-friendly robot trajectory can be obtained by combining pieces from the trajectories that correspond to geometric objectives. In other words, the set of geometric trajectories form a basis for computing a human-friendly trajectory. The basis trajectories are combined using biofeedback.

Our approach is as follows: We start by identifying a collection of Simple Geometric Objectives (SGOs). For the tracking problem, example SGOs maintain a minimum (or maximum) distance from the human. Suppose we have k SGOs and let π_i be the optimal policy for the i^{th} SGO. We will refer to π_i as *simple controllers*. Note that SGOs do not involve an emotional component. Therefore simple controllers can be computed directly. Our technique involves stitching π_i to obtain a human-friendly controller. At a high-level, the algorithm can be described as follows: We maintain an index $i \in [1, \ldots, k]$ of the most recently used simple controller. We also maintain a current controller π obtained by stitching π_i.

Algorithm 1 Policy Switch(set <controller> C)

```
1: i = 0
2: π = C[i]
3: response = execute( π )
4: while response > HIGH do
5:    select switching point s from response
6:    ++i
7:    π = stitch π and C[i] at s
8:    response = execute( π )
9: end while
```

Initially, we pick an arbitrary i, and define $\pi = \pi_i$ as the current controller. We

perform an experiment where we execute the current controller and record the stress response of the human subject. We then identify the minimum time index t_s during the execution of π where there is a significant stress increase. The time t_s marks a switching point. In the next trial, we modify π as follows. From t_0 to t_s, π uses the same strategy as the previous trial. At time step t_s, π switches to a simple controller chosen randomly that is not equal to π_i. After updating the current controller, we repeat the experiment until we obtain a trajectory where the stress level is expected to be low throughout the execution of the trajectory.

In the next section, we demonstrate the utility of this algorithm using real and synthetic experiments. As we will see, the algorithm computes human-friendly trajectories with a small number of trials.

6 Results

To demonstrate the utility of the learning algorithm, we first present results using a synthetic stress model of the human and several simple geometric controllers. We do this in order to demonstrate the feasibility of controller stitching and illustrate the desired result of the actual experiment. Next we present results from five human test subjects. We use the same simple controllers for both the simulation and experimental results. The controllers are designed to keep the robot within a specified minimum and maximum distance of the human. Each criterion specifies a target set in the state space. The goal of the robot is to follow a trajectory that will maximize its total reward, but it receives no reward for states that are not in this target set (i.e. too far or too close). By reducing the size of this target set, the robot will attempt to collect more large rewards as opposed to many small rewards.

The constraint options are as follows:

1. min distance 1 unit max distance 10 units
2. min distance 3 units max distance 5 units
3. min distance 1 unit max distance 7 units
4. min distance 3 units max distance 10 units

This set of controllers was computed off-line and chosen to provide robot motions with sufficient variety. Each controller corresponds to the optimal strategy for the particular geometric criterion. In computing the controllers, we fix the trajectory of the human and

assume that the human moves three times the speed of the robot, in order to approximate the relative speed of the robot hardware and human.

6.1 Synthetic Experiments

The plots in figure 5 show results from a simulated stress model. Here we assume that the stress level of the human is inversely proportional to the distance between the human and robot. In computing the robot strategy, we do not presume to know anything about the actual stress model of the human. The simulation shows the result of a series of policies that had the desired effect. There are four trials. Each trial is equivalent to one iteration of the while-loop in algorithm 1. For the first trial we use a single policy. We examine the stress response and identify a time to switch policies, switch time 1. For the second trial we use the same policy as the previous trial up to switch time 1, then switch to a different policy, and identify switch time 2. For the third trial, we execute policy 1 up to switch time 1, policy 2 up to switch time 2 and then execute a third policy. The fourth trial is identical to the third up to switch time 4, at which point a fourth and final policy is executed. At each successive trial the first significant peak in the GSR reading appears later in time. We expect to see a similar if not identical change in GSR readings for repeated portions of the robot trajectory, and changes in the response near the switching points. The resulting trajectories for executing each successive trial are shown in figure 6.

6.2 Real Experiments

Our experiments use the same four controllers as the synthetic experiments. In our experimental setup (Figure 4) the human approximately follows a trajectory. Controllers are pre-computed according to this trajectory. Subjects were instructed to walk slowly but were not given any queues to ensure exact timing. Previous experience and experiments show that giving queues to make the subject follow the trajectory exactly generates a significant amount of stress and makes it difficult measure the affect of the robot motion.

In our experiments, we used an Acroname Garcia to track the human (Figure 4(c)). A calibrated system of stereo cameras is used to track the human and robot to provide the controller with feedback. Stress was measured using the ThoughtStream Personal Biofeedback System [HC92]. The device measures galvanic skin response at 4 Hz, however we take readings at more incremental steps, depending on the actions of the robot, which are not atomic. The start state of the robot is consistent throughout the experiment. The data from the measurement device was sent to the controller from a portable

computer carried by the test subject. In order to keep from inducing stress, the computer, measurement device and all connection cables are packed neatly in an easily carried bag. Control inputs to the robot were also sent by wireless from the central computer. In order to control for reliability issues with the GSR sensor, we use a conservative thresholding filter to remove erroneous readings. These erroneous readings are mostly caused by the sensor breaking contact with the skin. This occurs infrequently and can be detected with high certainty (see figure 7).

We present results for five test subjects. In figures 8, 9, 10, 11, and 12, we give both the raw stress measurement as well as the change in stress during each trial. One trial uses the same test subject in four iterations of the while loop specified in algorithm 1. We assume that the trajectory of the human is the same for each trial. A trial consists of the following: select a policy and execute that policy. Determine the first switch time by examining the response. In the next trial run the same policy until the selected switch time, then execute a different policy. We examine the change in stress, relative to the measurement taken at t_0 in order to control for stress changes caused by previous trials.

For each subject we show two figures. The first shows the raw stress reading at each time step of each successive trial. The time of the new switching point for each new trial is illustrated by the change from regular to bold (Figures 12(a), 8(a), 9(a) and 10(a)). The second figure in each group shows the relative change in stress from the initial reading. (Figures 12(b), 8(b), 9(b) and 10(b)). In general, this is a more telling value, since it accounts for the affects of the previous trials.

The results for subject 1 (figures 8(a) and 8(b)), the first trial labeled C1, shows a spike between t_4 and t_5. Trial 2 (curve C2) switches to controller 2 at t_4 and the next significant increase is at t_8. Trial 3 (curve C3) shows a switch from controller 2 to 3 at t_6, which shows a minor improvement. Switching to controller 4 at t_8 (curve C4) shows a significant improvement, moving the next significant increase to t_{12}. For subject 2 (figures 9(a) and 9(b)), the first switching point (trial 2) occurs at t_{10}. Trials do not show consistency for identical portions of the robot trajectory. However, there is a significant change in the GRS response exactly at the first switching point, t_{10} of C2. Subject 3 (figures 10(a) and 10(b)) shows consistency across all 4 trials, but peaks are relatively small. There is consistency between trials C1 & C2, C3 & C4, but not between C2 & C3. There are marked changes at the first and third switching points (trials 2 and 3).

For subject 4 (figures 11(a) and 11(b)), there is consistency across all four trials. There are significant changes at the first two switching points, t_4 of C2, and t_3 of C3. Although the changes in stress across all 4 trials for subject 5 are small compared to

other subjects, there is no consistency for repeated portions of the robot trajectory. For this reason, we consider this experiment unsuccessful. One possible reason is that the speed of subject 5 was not consistent across the 4 trials.

We started with the assumption that by repeating a portion of a robot trajectory under identical conditions we would see approximately the same stress response from the human. Our observations can be summarized as follows:

- We can see a general consistency in the change in stress during repeated trials, particularly in the case of subjects 1 (figure 8(a)) and 4 (figure 11(a)). These similarities are also apparent in subjects 2 and 3 but to a lesser extent. This shows evidence for repeatability, which is essential to the success of our algorithm.

- Where the stress is consistent, between two trials, there is typically a change at the switching point. This is evident during trials 2, 3, and 4 of subject 1 (figure 8(a)), trial 2 of subject 2 (figure 9(a)) trials 2 and 4 of subject 3 (figure 10(a)), and trial 2 of subject 4 (figure 11(a)). This shows evidence that the measured affective state can be changed by altering the trajectory of the robot.

- Noteworthy is the fact that repeating a portion of a trajectory under the same conditions did not always evoke a similar response. This is particularly true in the case of subject (figure 11(a)), and trial 3 of subject 3. For the cases where this occurred (Subjects 3 and 5), the overall changes in stress is much smaller.

In general, the algorithm worked well in cases where the stress response is consistent for repeated portions of a trajectory. In cases where the algorithm did not work, we did not see consistency for the repeated portions of the trajectory. This suggests that emotional dynamics of some subjects change over time. Modeling this change, and designing controllers that account for it constitute our main future research efforts.

7 Conclusions

In this paper, we demonstrated the use of a galvanic skin response sensor as a tool for designing controllers for human robot interaction. Specifically, we presented an algorithm which yields human friendly controllers. The algorithm stitches together purely geometric controllers using biofeedback. In real experiments, we demonstrated evidence of repeatability in the stress patterns of individual test subjects as well as the ability to change these patterns by switching to different controllers.

In future work, we will investigate the use of biofeedback in controller design further by conducting larger, more controlled experiments, which involve more complex tasks for both human and robot. We will also explore the use of additional biofeedback sensors such as heart-rate sensors to design more robust controllers.

In the present work, we focused on generating robot trajectories – instead of polices across the entire state space. These trajectories can be considered as "expert demonstrations" and used to compute more general policies. The problem of computing a full policy from an expert's demonstrations is known as the *apprenticeship learning* problem [AN04, RA07]. Our future research includes investigating apprenticeship learning techniques to compute policies over the entire state-space.

References

[ADOB97] Gregory D. Abowd, Anind K. Dey, Robert Orr, and Jason A. Brotherton, *Context-awareness in wearable and ubiquitous computing*, ISWC, 1997, pp. 179–180.

[AN04] Pieter Abbeel and Andrew Y. Ng, *Apprenticeship learning via inverse reinforcement learning*, ICML '04: Proceedings of the twenty-first international conference on Machine learning (New York, NY, USA), ACM Press, 2004, p. 1.

[BA07] Andrew G. Brooks and Ronald C. Arkin, *Behavioral overlays for non-verbal communication expression on a humanoid robot*, Auton. Robots **22** (2007), no. 1, 55–74.

[BMA+01] D. Bersak, G. McDarby, N. Augenblick, P. McDarby, D. McDonnell, B. McDonald, and R. Karkun, *Intelligent biofeedback using an immersive competitive environment*, Ubicomp, 2001.

[Cly97] D.M. Clynes, *Sentics: The touch of emotions*, Anchor Press, 1997.

[CP05] H. I. Christensen and E. Pacchierotti, *Embodied social interaction for robots*, AISB-05 (Hertsfordshire) (K. Dautenhahn, ed.), April 2005, pp. 40–45.

[FND02] Terrence W Fong, Illah Nourbakhsh, and Kerstin Dautenhahn, *A survey of socially interactive robots: Concepts, design, and applications*, Tech. Report CMU-RI-TR-02-29, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, December 2002.

[HC92] Bruce Harrah-Conforth, *Accessing alternity: Neurotechnology and alternate states of consciousness*, 1992.

[HRJ04] P. J. Hinds, T. L. Roberts, and H. Jones, *Whose job is it anyway? a study of human-robot interaction in a collaborative task*, Human-Computer Interaction **19** (2004), 151–181.

[KG06] Vladimir A. Kulyukin and Chaitanya Gharpure, *Ergonomics-for-one in a robotic shopping cart for the blind*, HRI '06: Proceeding of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction (New York, NY, USA), ACM Press, 2006, pp. 142–149.

[KGNO06] V. Kulyukin, C. Gharpure, J Nicholson, and G. Osborne, *Robot-assisted wayfinding for the visually impaired in structured indoor environments*, Autonomous Robots (2006).

[LRS99] Neal Lesh, Charles Rich, and Candace L. Sidner, *Using plan recognition in human-computer collaboration*, UM '99: Proceedings of the seventh international conference on User modeling (Secaucus, NJ, USA), Springer-Verlag New York, Inc., 1999, pp. 23–32.

[MSK+05] N. Mitsunaga, C. Smith, T. Kanda, H. Ishiguro, and N. Hagita, *Robot behavior adaptation for human-robot interaction based on policy gradient reinforcement learning*, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2005), 2005, pp. 1594–1601.

[PCJ05] E. Pacchierotti, H. I. Christensen, and P Jensfelt, *Human-robot embodied interaction in hallway settings: a pilot user study*, IEEE ROMAN-05 (Hertsfordshire), August 2005, pp. 164–171.

[Pen04] Alex (Sandy) Pentland, *Healthwear: Medical technology becomes wearable*, Computer **37** (2004), no. 5, 42–49.

[Pic95] R. Picard, *Affective computing*, Tech. Report 321, MIT Media Laboratory, Perceptual Computing Section, November 1995.

[Plu80] R. Plutchik, *A general psychoevolutionary theory of emotion*, Emotion Research, Theory, and Experience (R. Plutchik and H. Kellerman, eds.), vol. 1, Theories of Emotion, Academic Press, 1980, pp. 3–33.

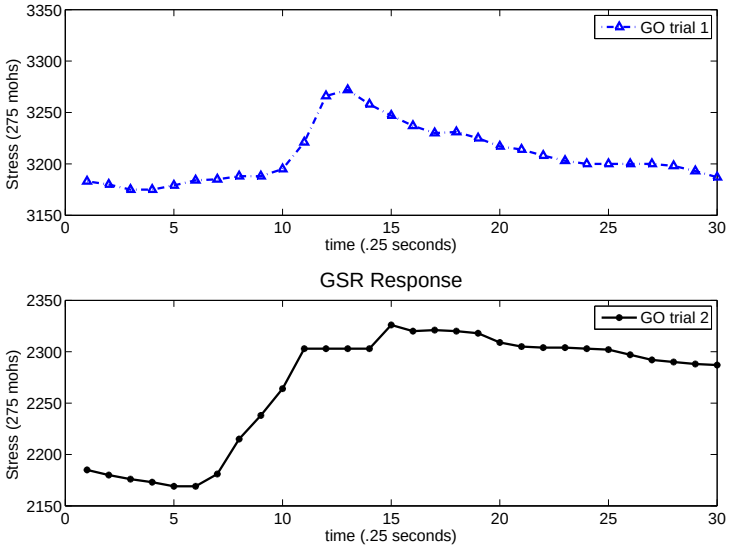
[RA07] D. Ramachandran and E. Amir, *Bayesian inverse reinforcement learning*, 20th International Joint Conference on Artificial Intelligence (IJCAI'07), 2007.

[RLS06] Pramila Rani, Changchun Liu, and Nilanjan Sarkar, *Affective feedback in closed loop human-robot interaction*, HRI '06: Proceeding of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction (New York, NY, USA), ACM Press, 2006, pp. 335–336.

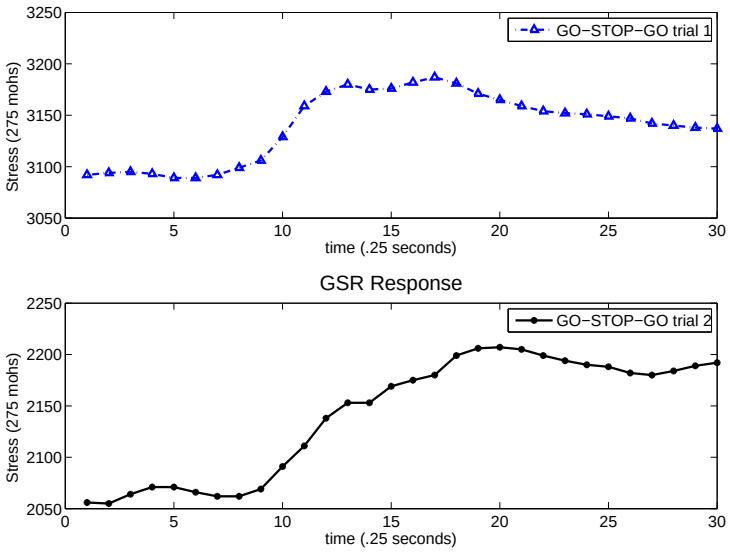
[RPT00] Nicholas Roy, Joelle Pineau, and Sebastian Thrun, *Spoken dialogue management using probabilistic reasoning*, ACL '00: Proceedings of the 38th Annual Meeting on Association for Computational Linguistics (Morristown, NJ, USA), Association for Computational Linguistics, 2000, pp. 93–100.

[RS05] P. Rani and N. Sarkar, *Making robots emotion-sensitive - preliminary experiments and results*, Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on, 2005, pp. 13–15.

[RSSK04] Pramila Rani, Nilanjan Sarkar, Craig A. Smith, and Leslie D. Kirby, *Anxiety detecting robotic systems - towards implicit human-robot collaboration*, Robotica **22** (2004), no. 1, 85–95.

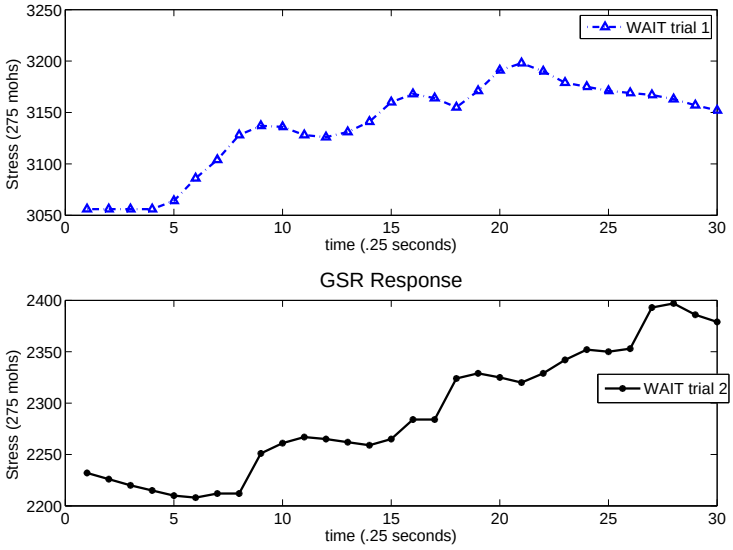


(a) GO

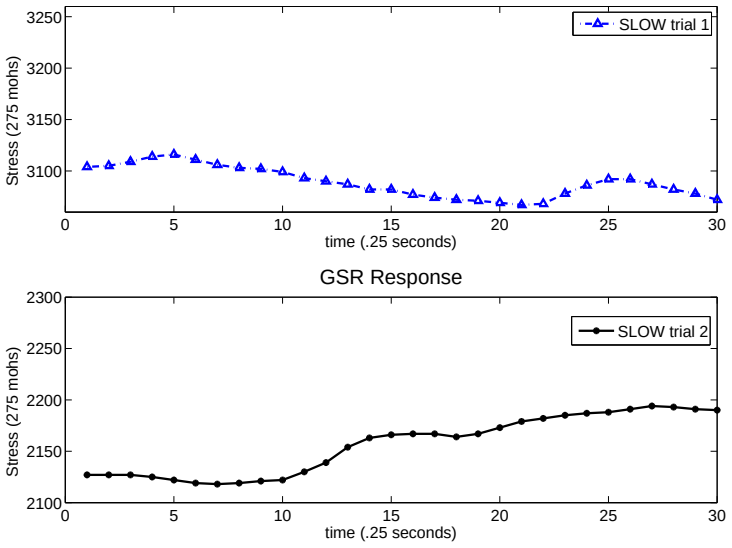


(b) STOP

Figure 1: Comparison of GSR response for two iterations of **GO** controller(1(a)) and two iterations of **STOP** (1(b))

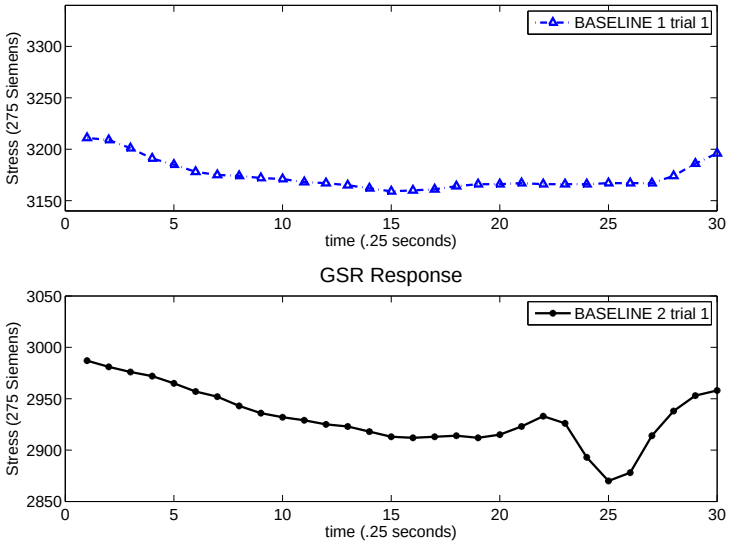


(a) WAIT

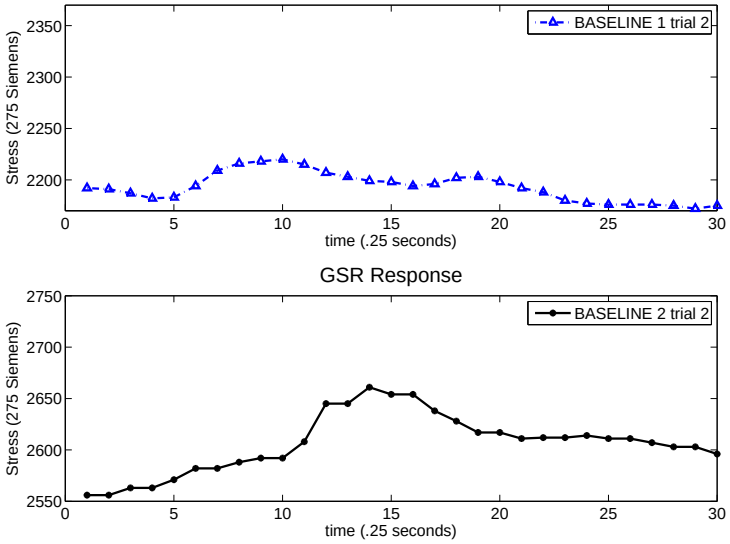


(b) SLOW

Figure 2: Comparison of GSR response for two iterations of **WAIT** controller(2(a)) and two iterations of **SLOW** (2(b))

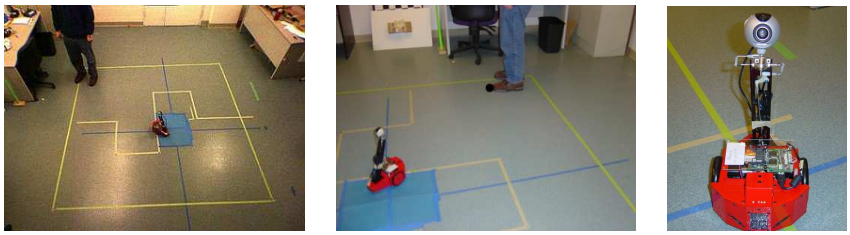


(a) Baseline1



(b) Baseline2

Figure 3: Baseline readings from before and after the two trials. Figure 3(a) shows the before and after baseline response in trial 1 and figure 3(b) shows the before and after baseline response for trial 2.



(a) (b) (c)

Figure 4: Experimental setup. The green square marks the human trajectory. The robot and human are tracked using a calibrated system of cameras. Figure 4(a) shows the view from the external cameras used to track the human and robot. Figure 4(c) shows the Acroname Garcia robot used in the experimental setup.

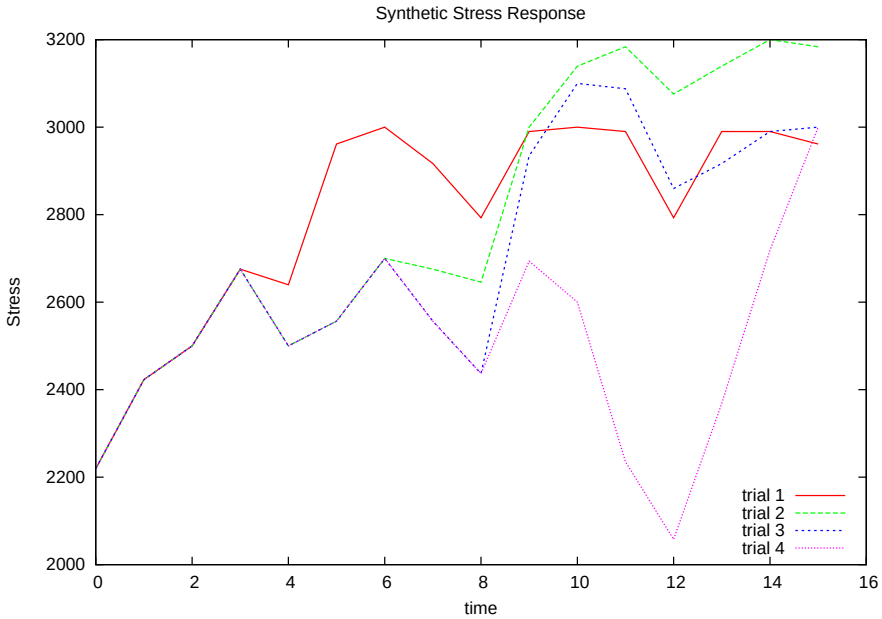
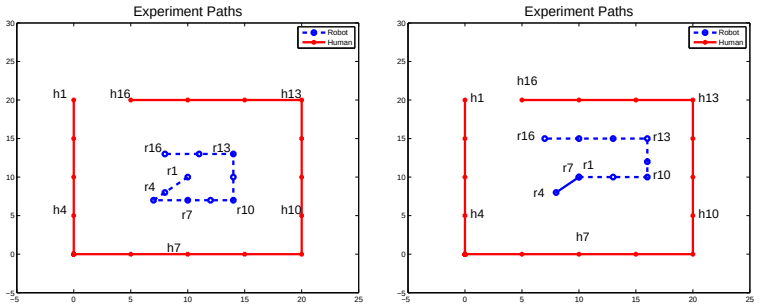


Figure 5: Results for four trials using synthetic stress model



(a) Control 1 (b) Control 2

(c) Control 3 (d) Control 4

Figure 6: Simulated robot and human trajectories. The trajectories produced by switching algorithm using the synthetic stress model (figure 5)

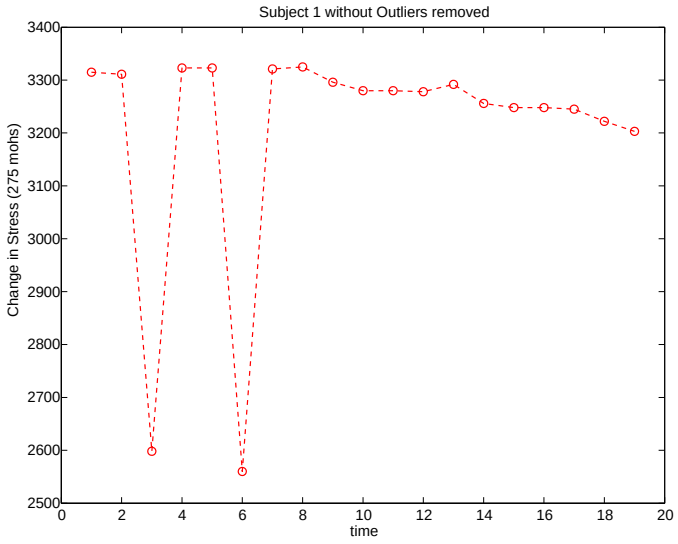
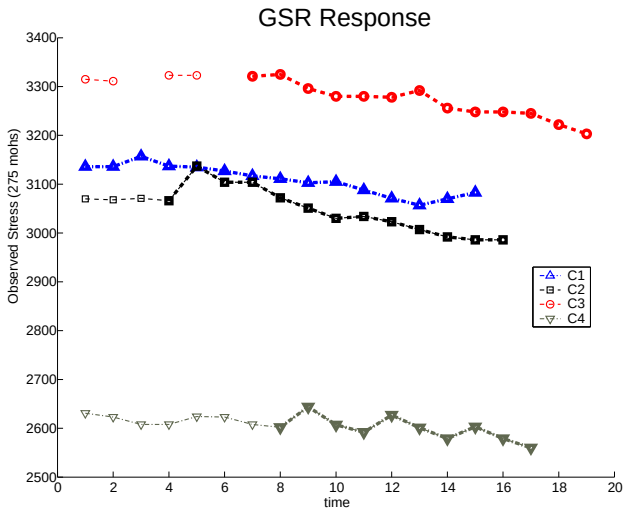
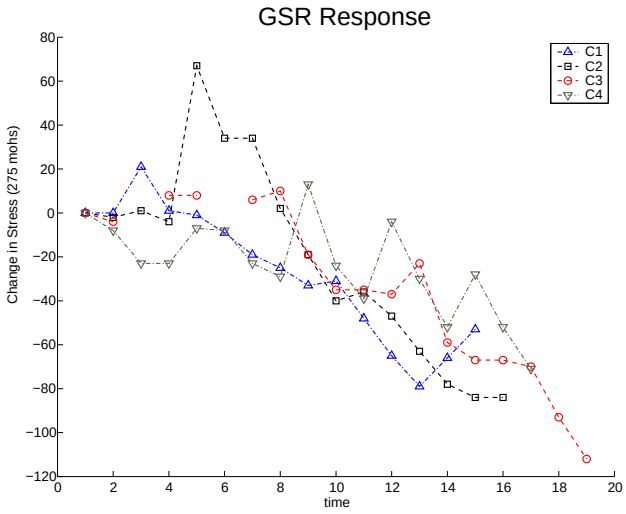


Figure 7: Stress response for trial 3 of subject 1 without outliers removed. Compare to figure 8 where readings at time t_3 and t_6 have been removed.

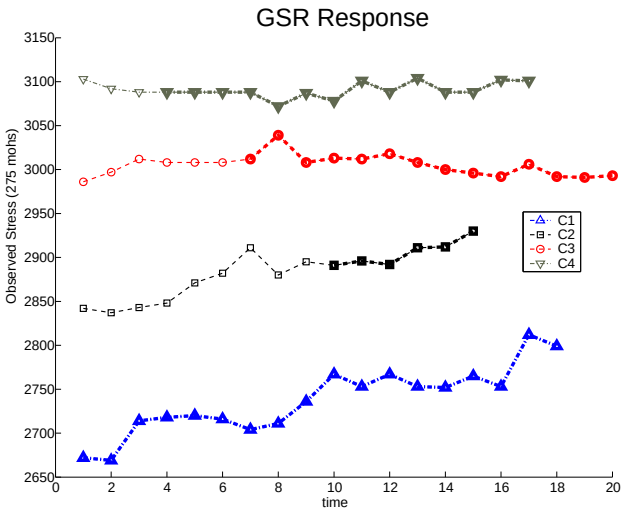


(a) Raw GSR

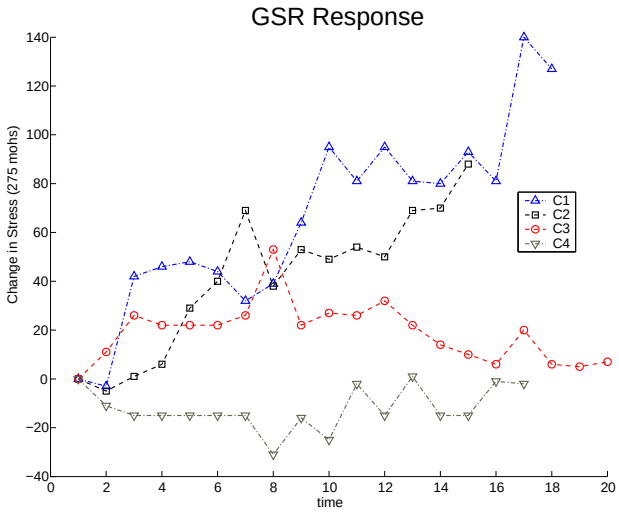


(b) Change in GSR

Figure 8: Results for Subject 1. 8(a) raw GSR readings and 8(b) change in GSR from initial value. Switching points are t_4, t_6, t_8. See text for further discussion.

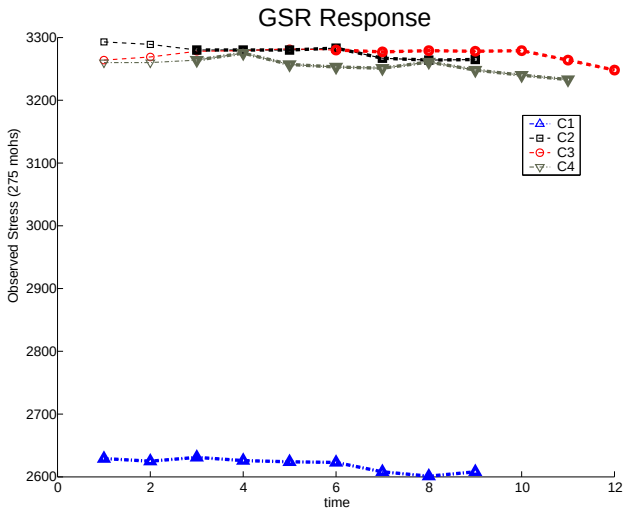


(a) Raw GSR

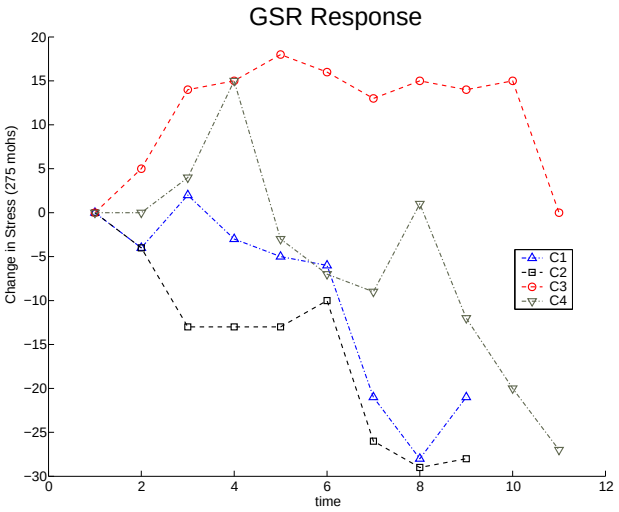


(b) Change in GSR

Figure 9: Results for Subject 2. 9(a) raw GSR readings and 9(b) change in GSR from initial value. Successful switching point at t_{10}.

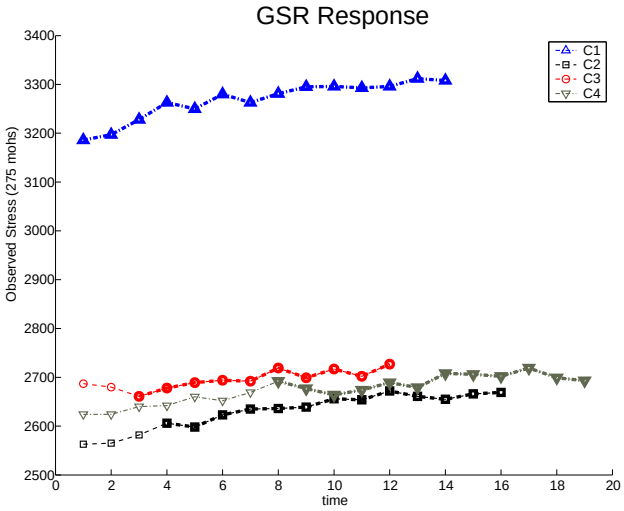


(a) Raw GSR

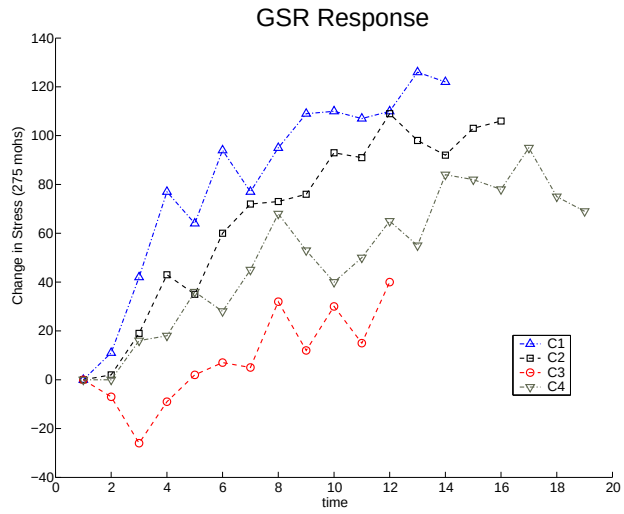


(b) Change in GSR

Figure 10: Results for Subject 3. 10(a) raw GSR readings and 10(b) change in GSR from initial value. There is consistency across all 4 trials, and significant changes at switching points in trials 2 and 4.

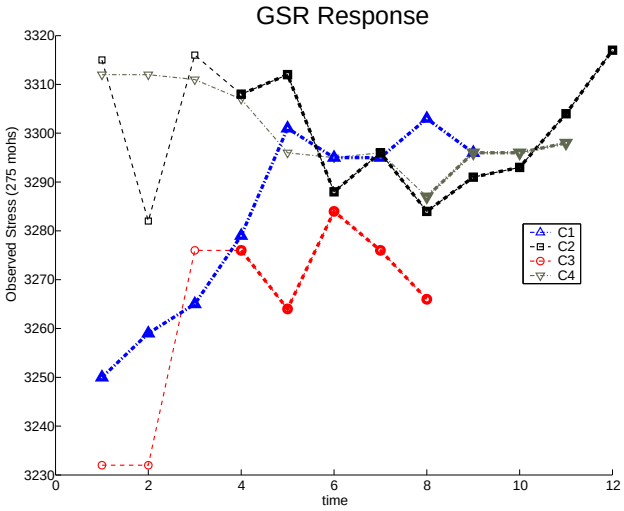


(a) Raw GSR

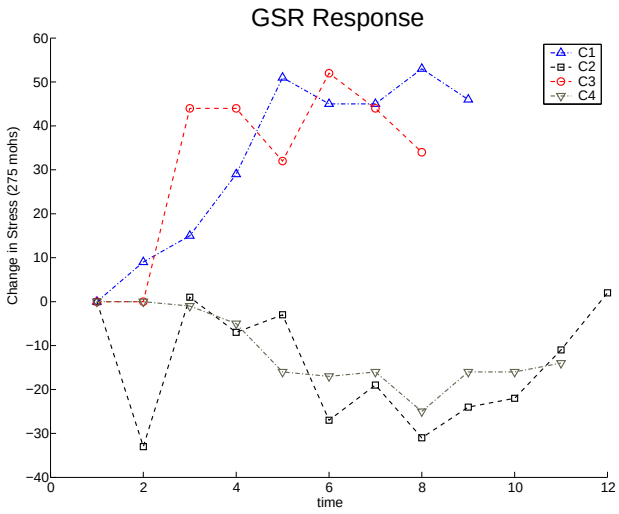


(b) Change in GSR

Figure 11: Results for Subject 4. 11(a) raw GSR readings and 11(b) change in GSR from initial value. Consistency across trials and significant changes at switching points for trials 2 and 3.



(a) Raw GSR



(b) Change in GSR

Figure 12: Results for Subject 5. 12(a) raw GSR readings and 12(b) change in GSR from initial value. Changes in GSR response for this subject are inconsistent but relatively small.