Proceedings of the American Control Conference
San Diego, California • June 1999

Control of Resistance Spot Welding

Justin Shriver
Huei Peng
S.Jack Hu
j.s.shriver@ieee.org

Abstract

This paper discusses recent progress in the development of a feedback controller for the resistance spot welding (RSW) process. Two key elements of our control architecture are operating condition detection and nugget growth estimation. In this paper the operating condition problem is addressed with respect to three well-defined operating conditions. We employ a principal component-based design for the operating condition detector. The nugget growth estimation problem is addressed by employing an affine input-output model and neural networks for local feedback.

1 Introduction

Our goal is to produce controllers to improve the reliability of the resistance spot welding process. Due to the extensive use of resistance spot welding, particularly in auto-bodies, even a small improvement would bring significant economic benefits. This large potential payoff has attracted a significant amount of research in both the resistance spot welding field in general and the specific subfield of resistance spot welding control. In order to improve on the current state of the art, we have undertaken a significant research program. A large component of the work is experimental which is briefly reviewed in this paper. In addition to the experimental work, we have produced both an operating condition detector and a simple control detailed in this paper.

2 Experimental Conditions

The data in this paper were all collected at the University of Michigan on an instrumented Taylor Winfield pedestal welder. For this round of experiments, we used the following parameters: electrodes CU-CR #7 A-nose from Nippert; a nominal welding force of 600 pounds, and a nominal current of 11kA. One should note that while the current was held constant, the welding times were varied from 2 to 24 cycles. All welds were performed on specially constructed coupons made from .8mm DS steel. The dimensions for the coupon were 140mm by 35mm. These coupons were assembled by placing a single anchor weld on each coupon, spaced 35mm from the test weld.

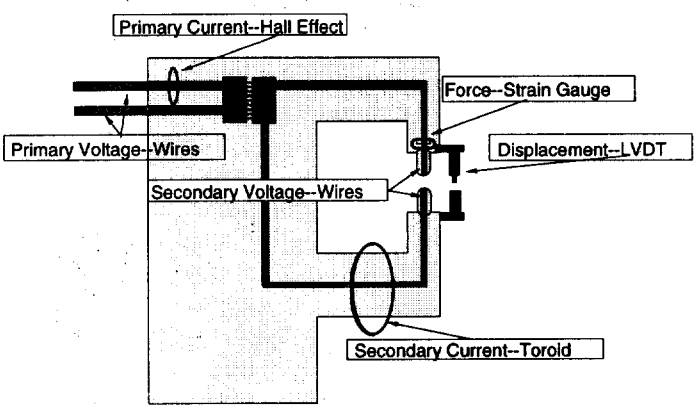

Figure 1: Instrumentation

One of the most important aspects is machine instrumentation. In this paper, we can only present a brief review of our instrumentation package. A schematic view of the instrumentation is shown in Fig. 1. The six instrumented signals are: primary voltage V_p, uncorrected secondary voltage V_{unc}, primary current I_p, toroid voltage V_{tor}, displacement D, and force F. From these signals we can derive the remaining signals of interest: corrected secondary voltage V_{sec}, secondary current I_{sec}, and dynamic resistance DR. The reason for selecting these signals is covered in [3], [2], along with many other issues related to the instrumentation. These signals are passed to a signal conditioning box and then to a commercial National Instruments AD board with a total bandwidth of 500kHz. The signals are saved, along with test information, by a custom Labview program. Another aspect of our experimental setup is apparatus for generating different operating (or fault) conditions on the welder in a reliable fashion. For this purpose, a unique fixture was created.

3 Nugget Growth Model

Any control implementation requires some sort of process model. For the resistance spot welding, the most direct way to approach the modeling problem is to start

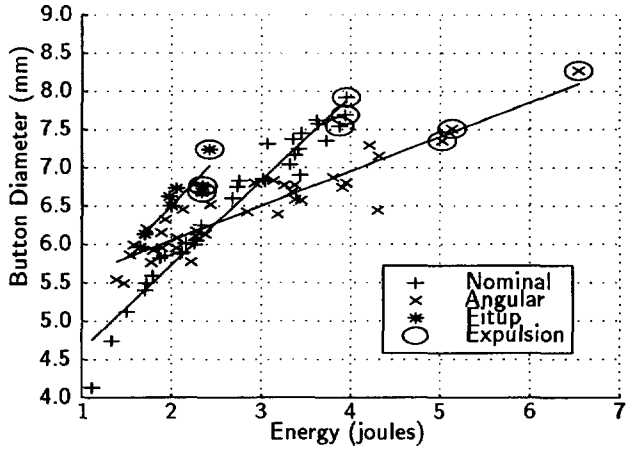

Figure 2: Nugget Growth by Operating Condition

with the governing partial differential equations. This yields two equations, an equation for the temperature dynamics

$$C\sigma\partial T/\partial t = \nabla \cdot (K\nabla T) + \rho_2\delta^2 \qquad (1)$$

and an equation for the voltage field

$$\nabla \cdot (1/\rho_1 \nabla V) = 0. \qquad (2)$$

The symbols have their standard meanings: ρ_1 is the electrical resistivity, V is the electrical potential, K is the thermal conductivity, T is the temperature, ρ_2 is the thermal resistivity, ∇ is the del operator, C is the specific heat, σ is mass density and δ is current density. For simple conditions, these equations admit analytical solutions. However, the conditions for this problem are not simple. The simplest reasonable assumption leaves ρ_1, ρ_2, V, K as functions of (x,y,z,t). In addition, heat is generated at two contact interfaces (electrode/workpiece,workpiece/workpiece) which require empirical modeling. The electrode/workpiece also has a similarly complex heat transfer property.

Having outlined the problems with a direct approach to the fundamental equations, one could hope to simplify the problem by identifying an input-output model. In employing such a technique, one hopes that complete knowledge of the state evolution is unnecessary. For this problem, a natural input-output pairing is energy (integral of input) and nugget diameter. Plotting the data from the three operating conditions together, along with lines representing the linear least squares fit of the data, reveals a clear affine pattern (Fig. 2).

The previous plot makes it clear that the model parameters change with operating condition. One should note that the parameters will likely change as the force, current, electrode force, or coupon dimensions are varied. However, we believe that for the range of parameters commonly used, the affine form of the model will be

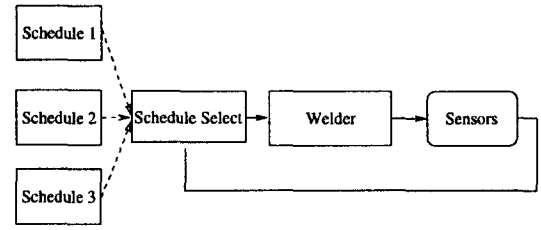

Figure 3: Open Loop Switched Control

sufficient. Most previous researchers have attempted to solve the problem independent of operating condition. From the plot it is clear that estimating nugget size and expulsion without knowledge of operating condition is difficult.

While an approximately linear input-output map is useful, it is not sufficient to implement a traditional feedback control design. The only control design available from the current models would be a switched set of open loop controllers (Fig. 3). We are limited to open loop control because there is no explicit relationship between the desired quantity and the measured outputs. To clarify this situation requires some additional notation: s_m represents the instrumented signals, s_{om} represents an instrumented signal(s) that can be related to the operating condition, o represents the operating condition, d represents the disturbance, u represents the controlled input, y is the output (related to the state), y_d represents the desired outputs (related to the state), y_o is the part of the output that can be measured, x represents the state, and x_d represents the states from which the desired outputs can be calculated. In this notation the outputs y need not be instrumented. The form of the controller we have designed is appropriate under the following assumptions:

$$o = h(s_{om}, u) + \epsilon(h(x, o, d) + j(x, u, o, d)) \qquad (3)$$

This first assumption is the key assumption: there is a measured signal(s) s_{om} that can be related to the operating condition o and the input u. Ideally, it would be just a function of the operating condition, but since the input is known, it also could be an invertible function of the input. This assumption is recognizable as the standard assumption for gain scheduling, which would be one way to view our switched control structure.

$$\dot{x_d} = g(u, o) + \epsilon(f(x, o, d) + g(x, u, o, d)) \qquad (4)$$

The second assumption is that evolution of the desired states is dominated by the input. Otherwise, a model of the general state evolution would be required. Note the desired states are those states that are sufficient to calculate the desired output.

$$y_d = h(x_d) + \epsilon(h(x, o, d) + j(x, u, o, d)) \qquad (5)$$

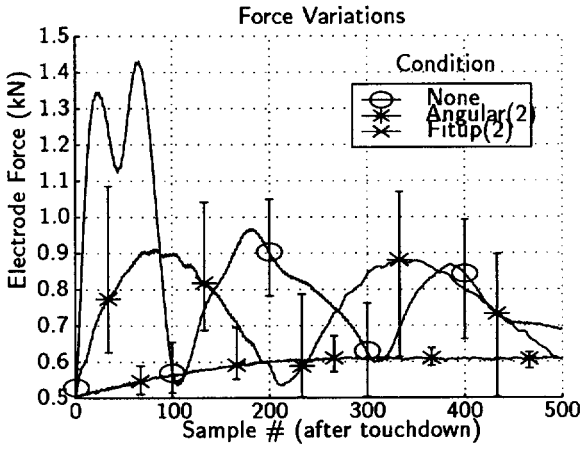

Figure 4: Force Variations

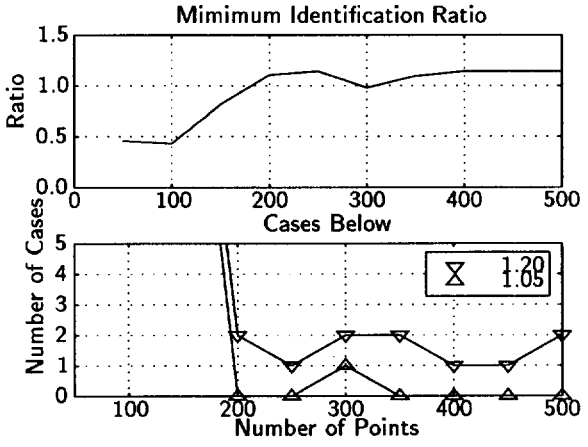

Figure 5: Determining Window Size

The third assumption is that the desired output, which is not instrumented, is dominated by its relationship to the desired states.

$$y_o = j(u) + \epsilon h(x_d, o, d) + j(x_d, u, o, d) \qquad (6)$$

The final assumption is that the outputs that are observed are not related to the state evolution. One should note that in the above equations ϵ is used to represent a "sufficiently small" quantity.

4 Operating Condition Detection

As outlined in the previous section, finding the relationship $o = h(s_{om}, u)$ is vital. Without this relationship the operating region cannot be detected and the switch to the appropriate controller cannot be performed. In our previous paper we reported good results by using the first principal component of dynamic resistance. For the current data set this identification algorithm proved ineffective. An alternative means to detect the operating condition, based on touchdown force profiles, was suggested by Lyndon Brown [1]. Plotting the force traces soon after impact reveals that the different operating conditions produce very different shapes (Fig. 4).

In implementing a principal component-based algorithm, there are four steps we must perform. First, we eliminate time jitter by triggering on a force level of .5kN. Second, we preprocess the force signal by normalizing it to have zero mean (not required but improves performance). Third, we select the number of components to use. This decision is simple because the first principal component carries two orders of magnitude more energy. Fourth, we select the length of the principal components. Selecting the length of the principal components is a more difficult task than the other three. To determine the appropriate length of

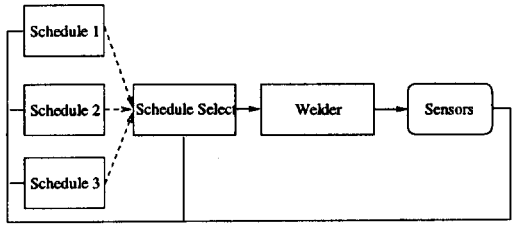

Figure 6: Closed Loop Switched Control

the principal components we employ two graphs (Fig. 5).The first shows the ratio between the power of the correct principal component and the signal and the maximum power between the remaining principal components and the signal. Ratios below one represent misidentification. The plot shows that as the length of the principal component under consideration is extended past two hundred points, there are no incorrect identifications. The lowest (worst) value of the identification ratio, above 200 points, is 1.17. However, simply picking long principal components is not the best strategy, as past 450 points the number of cases below 1.20 begins to increase.

5 Black Box Local Feedback

One of the assumptions made in our control design was that there is no observed output that could be related to the desired states. This assumption was formalized by stating that ϵ was small in Equation 6. However, what if ϵ is not small? If the form of h and j were known, then we could design a conventional observer. Without specific knowledge of h and j we could attempt a black box identification of the composite $h+j$. Given $h+j$ we could employ feedback throughout the process, resulting in the control structure shown in Fig. 6. However, for this process there are two primary complications in designing a feedback mechanism. First,

since weld diameter can only be determined by post-process destructive testing, how to construct the inputs for the black box is not clear. One could "join" the data (from different length experiments) together. This would give the measured outputs and the nugget size for every cycle. However, the data could only be "joined" if the disturbance d did not affect the desired output significantly, in which case feedback would not be necessary.

An alternate approach is to assume that each experiment represents a unique value for the disturbance and that d does affect the desired output. The disadvantage of this strategy is that each experiment can only provide one data point. Given this method to construct the input, the next step is to pick a method for approximating $h + j$. Due to the complexity of this process, we believe that a nonlinear black box technique will be required. In order to create the black box model, one of the most important steps is to identify which inputs are relevant. There is no systematic way to select inputs for the black box model; any decision is a matter of engineering insight. For four of the inputs we chose differences: max DR-initial DR, max DR-final DR, max D-initial D, and max D-final D. D and DR are displacement and dynamic resistance, respectively. Two additional inputs used were max DR and max D. For the seventh input, we used energy. Since the black box models are hard to verify, we would like to limit their impact on the nugget size estimator. One natural way is to use the black box technique to identify the error model between the linear model and the actual data rather than the model between the inputs and the outputs. Thus, we are creating a black box that takes seven inputs and estimates the error between the linear nugget size model and the actual nugget size.

Having selected the inputs and outputs for our black box model, the last remaining task is to attempt to create these models. For this purpose we will employ sigmoidal (specifically tansig) neural networks. To produce candidate neural networks, we use a highly randomized algorithm, the scope of which is beyond this paper. However, we can summarize by saying that we trained for 313,379 epochs (about 24 hours on a PII), which produced 1,266 candidate models with between 1 and 10 hidden neurons. These models were trained with 70% of the data and verified with the remaining 30%. One way to evaluate the performance is to plot the estimate from the linear estimator, black box plus the linear estimator, and the actual values. This is done for the nominal case in Fig. 7.

While the graphical approach is visually appealing, a more compact and precise representation can be arrived at by tabulating the results (Table 1). In this table nom represents the nominal operating condition, ang the angular fitup operating condition, fit the fitup

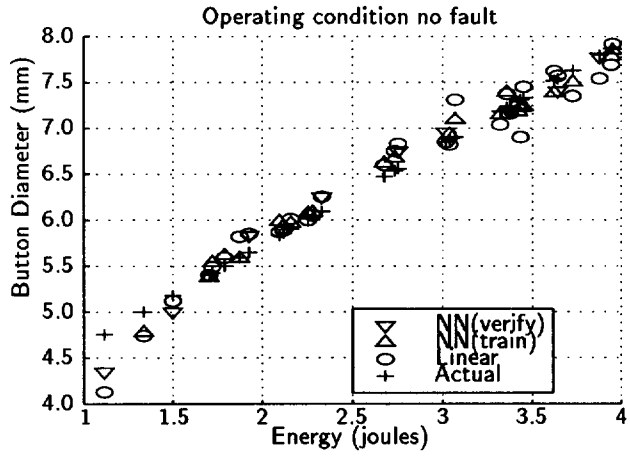

Figure 7: Estimated Nugget Growth (nominal operating condition)

Table 1: Performance for nominal operating condition

		Verify		Train		Whole	
Cond	Model	std	max	std	max	std	max
Nom	Lin	.27	.62	.18	.41	.21	.63
Nom	BB	.14	.24	.14	.32	.14	.32
Ang	Lin	.25	.65	.20	.38	.21	.65
Ang	BB	.24	.42	.11	.40	.15	.42
Fit	Lin	.19	.15	.18	.24	.17	.24
Fit	BB	.03	.12	.03	.04	.05	.12

operating condition, lin the linear model, and BB the black box model plus the linear model. We assess the error with respect to the infinity norm (or maximum error) and the standard deviation from the correct values std. Note that since the black box relationship cannot be easily validated, it is important to make sure that the performance on both the training and verification data is reasonable. The column marked whole represents the performance for the entire data set (training and verification). For any complex black box model, very high accuracy on the training data is possible. One only gains confidence in the model if the verification results are good, which is the case here.

6 Control Synthesis

Given both the linear input-output model, which provides a high gain open loop estimate of the nugget size, and the black box model, which provides a low gain closed loop estimate of nugget size, we have the two necessary components for our control design. The control strategy is to weld at a constant current (11kA) until the combined black box and linear model predict an adequate nugget size. Note that while this is a very simple strategy, it is well motivated. First, for the three operating conditions under consideration we

Table 2: Performance for various control strategies

			Small		Large	
Fault	Ctrl	Goal	#	err	#	err
Nom	All	6mm	3	.18mm	0	0mm
Nom	Lin	6mm	0	0mm	1	.25mm
Nom	BB	6mm	0	0mm	8	1.31mm
Ang	All	6mm	1	.23mm	8	.52mm
Ang	Lin	6mm	3	.23mm	3	.52mm
Ang	BB	6mm	3	.23mm	10	.52mm
Fit	All	6mm	0	0mm	0	0mm
Fit	Lin	6mm	0	0mm	0	0mm
Fit	BB	6mm	8	6mm	0	0mm

can achieve an adequate (6.5mm) weld by welding at 11kA. More importantly, because this process is nonlinear, one needs to verify the model over the entire range of possible inputs. Thus, an important goal of our control design method is to minimize the number of current levels used for welding.

In order to verify the control structure, we can compare the performance of our three estimators: a single linear model, a linear model for each region, a linear+ black box model for each region. Note that the linear models' performance will be closely tied to their performance in approximating the function, as these simply relate power to nugget size in a monotonic fashion. However, the black box estimator may not perform in line with its prediction error. This is because when used to predict nugget size, the model only needed to predict the last point. In the closed loop design, the model must predict all the intermediate points accurately, or the weld will be terminated at the wrong time. Another complication is that since the weld sizes are measured at the end of the weld, we can only approximate the error in most cases. However, as an measure of performance we report the number of welds that fall into two categories: undersized, and oversized along with approximate errors in Table 2. One should note that there are 32 welds for the nominal data set, 36 for the angular, and 8 for the fitup case. From this it is clear that the linear model for each operating region is more effective than the linear model for the entire data set. The comparison between the linear model for each region and the black box plus the linear region is less clear. For the nominal and angular faults both models perform about equally well in avoiding undersized welds, while the linear model is much better at avoiding oversized welds. However, a closer examination of the data shows that the black box model outperforms the linear model when estimating weld size at the end of the weld, exactly what the model was trained for. The problem comes in estimating intermediate weld sizes along longer welds. Thus the intermediate data points for some of the longer welds are outside of the range of the training data. The poor performance on these intermediate cycles suggests one of three possibilities: one, the linear correlation is sufficient and all additional variations are experimental noise; two, the data set was not rich enough to allow accurate information to be gathered for the intermediate lengths; three, the black box was ineffective. Due to the improvement in estimator accuracy that the black box model offered for the verification set, we believe the most likely possibility is the second. Thus, we are in the process of collecting a larger data set so we can determine the answer to this question. For the fitup fault the black box performs unacceptably. This is easily explained by the lack of data for this fault.

7 Conclusion

In this paper, we have presented a method for attacking problems of a particular class. This class is characterized by processes that do not have good structured models for the dynamics and for which there is no measurement available from which the desired output can be reproduced. If there were an accurate model, a model-based design would be expected to outperform the design we have described. Similarly, if there were a measured output that could be related to the desired output, a traditional observer-based design would be expected to predict the output more accurately than our technique. We believe the primary advantages of our control design are: it requires a relatively small set of data; it requires few assumptions on the actual plant dynamics; and it is relatively easy to design and implement. The small required data set is a particularly compelling advantage for nonlinear plants which require an extensive data collection program to cover a sufficient area to verify the feedback control.

8 Acknowledgments

We would like to acknowledge the support of the NIST-ATP program (95-02-0013 - Intelligent Resistance Welding) as well as the member companies of the Intelligent Resistance Spot Welding Program.

References

[1] Lyndon Brown. IRW 10th quarter report (internal). 1998.

[2] S. A. Gedeon, C. D. Sorensen, K. T. Ulrich, and T. W. Eagar. Measurement of dynamic electrical and mechanical properties of resistance spot welding. *Welding Journal*, 65(12):378s–385s, 1987.

[3] Justin Shriver, Huei Peng, and S. Jack Hu. Resistance spot welding: A neural network approach to modeling. Proceedings International Mechanical Engineering Conference and Exposition, 1998.