SAND2014-2352C

FEATURE-BASED MODEL VALIDATION

David J. Manko
dmanko@sandia.gov

Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-0557

ABSTRACT

Model validation has historically been conducted by defining and comparing quantities-of-interest (QoIs) which are selected for each application by expert opinion or correlation to damage/failure. Lack of understanding and variability of response can make damage/failure an unreliable measure of model validation. The novel approach taken in this work is to validate analytical models directly using the key response features identified by the subject matter expert (SME). Mature feature-based image analysis provides the technology for this new paradigm of model validation. Each problem class has its own set of essential features that must be carefully specified. Once identified, the features are extracted from the data set through segmentation. Image understanding is the process of identifying features in a data set while image matching is the process of correlating a specific feature to a data set. The degree in which a specific feature (e.g., initial impulsive response), segmented from an experimental data set, correlates to an analytical data set represents the degree of model validation. Feature-based model validation is applied to a single barge shock test where acceleration time histories are compared between 10 predictions and the experiment for eight locations on the barge. This example was previously evaluated using QoIs and a historical validation approach. Feature-based model validation mimics the SME approach to model validation and produces results for the presented example that are exactly consistent with previously published SME validation judgments.

BACKGROUND

Model validation has historically been conducted [1] by selecting and comparing quantities-of-interest (QoIs). Examples of QoIs are maximum acceleration at a point in a structure, root-mean-square (RMS) of response, and modal frequencies. A QoI is usually selected for a specific application based on precedent, expert opinion or correlation to damage/failure. The actual failure mechanism(s) may not be understood for a component or system such as voltage dropout of an integrated circuit board. Also, system variability may have a large effect on damage potential and occurrence, sometimes resulting in different failure modes. This lack of understanding and variability of response can make damage/failure an unreliable measure of model validation.

Selection of a QoI for a specific application may require extensive resources in testing and evaluation to demonstrate the suitability of the QoI for the range of system or component operating parameters (e.g., temperature, humidity). An example is the effort required to justify a fatigue damage indicator as a QoI for a high-performance aircraft. Another obstacle to model validation is the fact that input forcing functions are usually not quantified except in cases of precision testing. Attempts to compare output QoIs between model and experiment can be confounded by differences in the inputs for otherwise accurate analytical representations.

On the other hand, subject matter experts (SMEs) can identify the key “features” of a system or component response characteristics based on extensive observation similar to machine learning. Certain key features must agree (i.e., match) between the experimental and analytical results before the model is deemed adequate by the SME. For a candidate QoI to be suitable, it must correlate with these essential features of the response. It is difficult to encode expert judgment on an algorithmic basis due to many sources of spurious information such as temporal shifting, frequency distortion, nonlinearities, noise, and instrumentation error.

FEATURE-BASED IMAGE ANALYSIS

Feature-based image analysis [2] is the process of identifying and comparing key features in a data set. A feature is defined as a characteristic shape or construct. Typically, we think of the data set representing a visual image and a feature being a recognizable object. Segmentation is the process of separating features from the image background and from each other. At this stage the objects are not evaluated or categorized, they are simply extracted from a data set. Image understanding is the process of interpreting the segmented features to derive qualitative conclusions from the data set. Finally, image matching is the process of correlating a segmented specific feature to a data set.

We have all benefited from advances in the field of medical diagnostics that utilize feature-based image analysis technology. Medical imaging can generate vast amounts of data that must be evaluated by qualified medical personnel with important consequences. Computer-based diagnostic routines sift through the data to limit further examination to a manageable level. These diagnostic routines [3] utilize known features of the disease (e.g., tumor shape, coloring) under investigation to pare down the search space. The efficiency of medical diagnostic software is measured as the percentage of false positives.

PROPOSED APPROACH FOR MODEL VALIDATION

Even if it is understood and well known (e.g., repeatable fatigue failure at a joint), damage or failure may have too much variability to use as a measure of model validation. Selecting a QoI using expert opinion is an indirect approach that requires extensive testing to define the QoI limits of correlation and applicability. The approach pioneered in this report is to validate analytical models directly using the key response features identified by the SME. Feature-based image analysis provides the technology for this new paradigm of model validation as demonstrated in the following section.

Feature matching replicates the SME process of model validation where the SME emphasizes certain key features while ignoring non-essential information. Each problem class has its own set of essential features that must be carefully specified by interrogating the SME(s). The disparity of results confounds any attempts at point-by-point comparisons of analytical and experimental solutions analogous to region-based image matching [4].

Feature-based model validation utilizes existing segmentation and image understanding technology. If a feature can be identified then it can be matched (i.e., validated) in a data set. This approach implicitly includes the forcing function, in addition to the structural response, whereas frequency response functions (FRFs) [5] and modal assurance criteria (MAC) [6] only address the structural characteristics. The degree of validation can be tailored to the application because the degree of correlation or validation is quantified for the specified features. Different weighting factors can be applied to the different features in a data set according to their importance

APPLICATION OF FEATURE-BASED MODEL VALIDATION

Feature-based model validation is applied to an UNDEX problem of a barge shock test that was previously validated using standard, albeit windowed, QoIs [7]. The experimental data set for the single barge shock test consists of velocity measurements at eight gauge locations, A through F, around the barge. A bandpass 2-pole Bessel filter set at 0.25 and 250 Hz, respectively, was used when obtaining the experimental data. A corresponding set of 10 UNDEX simulations were conducted where the only model parameter that was changed was the charge density. The velocity results, both experimental and analytical, were numerically differentiated using a cubic polynomial that was least squares fit to a moving window of ten data points where the derivative of the cubic approximates the function derivative. This process is used to smooth experimental data that often has spikes or dropouts in the measured signals.

The simulations were conducted with minimal damping which produced significant high frequency content compared to the filtered experimental results. For consistency with the experimental data, the analytical results were filtered using a bandpass 2-pole Bessel filter set at 0.25 and 250 Hz, respectively. The procedure described in [8] was used for digital implementation of the analog Bessel filter and the resulting temporal lag (0.0005 sec) was negated during plotting of the results.

The Navy considers the velocity change of the initial impulse response as the important parameter for this class of problem. The step velocity has been used as a QoI for UNDEX events but this simple definition overlooks much of the response characteristics as discussed below. In this work, the initial impulse response is represented as a feature, not a QoI, and the model results are validated on this basis.

The first step after identifying the feature is to quantify the feature for subsequent identification and matching. The Point Distribution Model (PDM) and training shapes [9] are used to "educate" the software and quantify the initial impulse responses. To implement the PDM, landmark points are manually selected for a feature as shown for the experimental response of Gauge D in Figure 1. This process is repeated for the remaining experimental gauge responses to produce the training set in Figure 2. The landmark points of the training set are used to calculate a mean shape (shown in Figure 2 as the red curve) and principal components with corresponding eigenvalues to describe the feature variations about the mean shape. In this way, the PDM quantifies the feature for subsequent identification and matching.

Figure 1: Landmark Points for Experimental Response of Gauge D

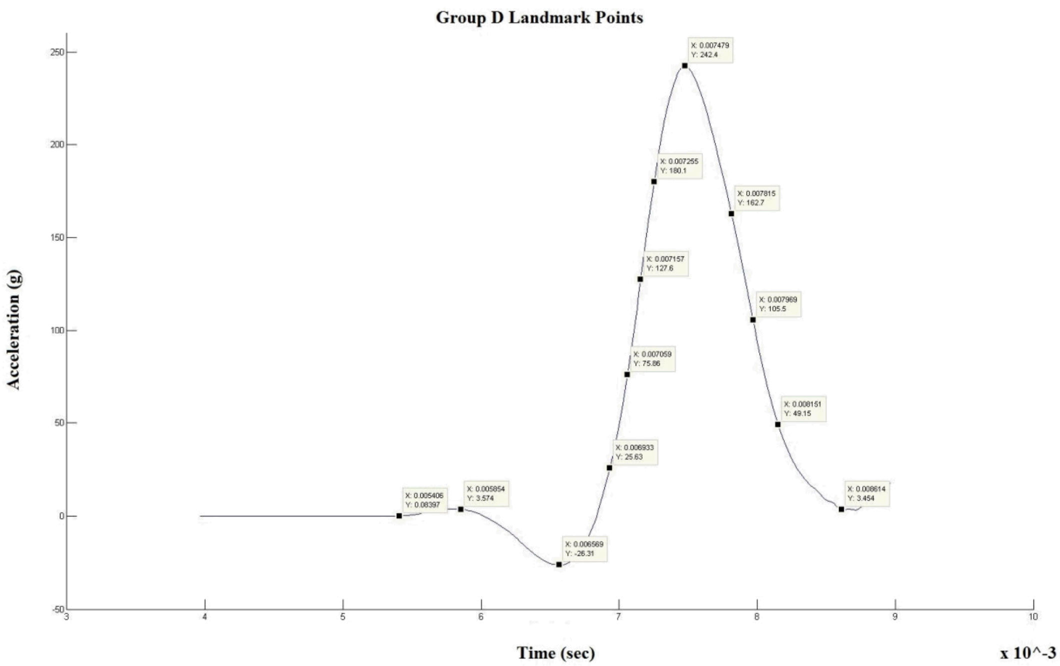
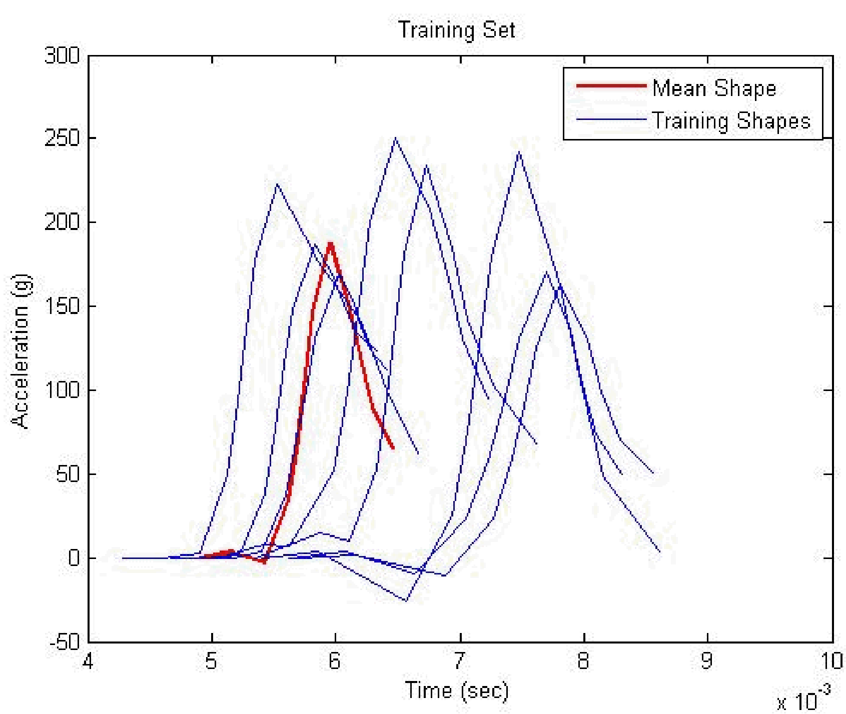


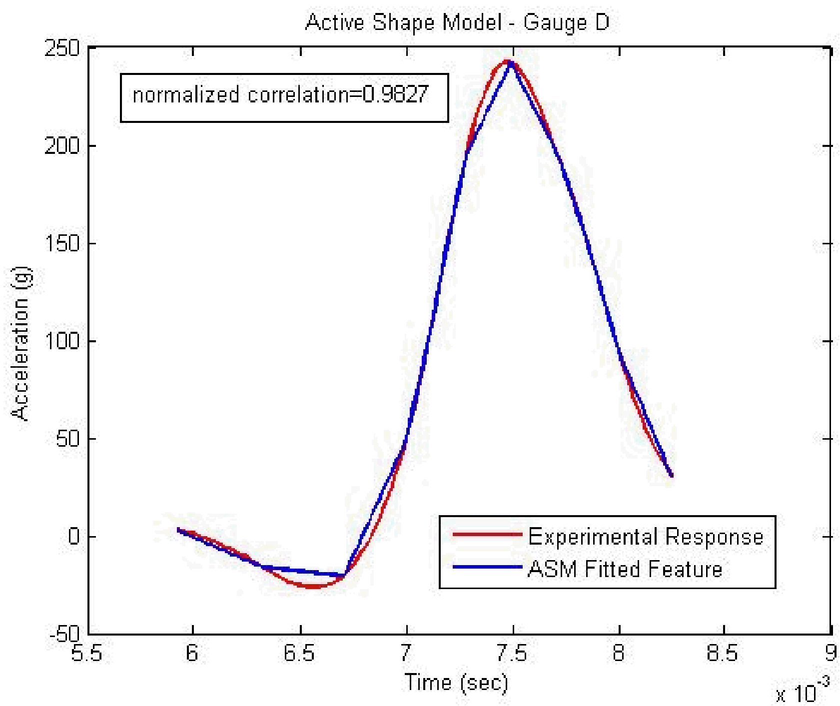
Figure 2: Training Set for Initial Impulsive Response Feature



The variability of the initial impulse responses can be observed in the training set above. Note that in some cases, the initial response is a negative acceleration which could indicate a horizontal measurement of an elevated structure undergoing a pendulum effect. The gauge locations are unknown for this test but there are clearly two different types of initial response attributed to vertical and lateral measurements. If more data were available, separate features could be defined for vertically and laterally measured initial responses. This discussion illustrates why a QoI cannot capture the intricacies of a system or component response in the same manner as a feature.

Once the features have been quantified using the PDM, the Active Shape Model (ASM) [9] is used to segment the features in the data set. The ASM morphs the mean shape of the feature within the constraints of the principal components to best fit the target data set. An example of applying the ASM to the initial impulse response feature for Gauge D is shown in Figure 3. The red curve is the original experimental response and the blue curve is the ASM fitted feature.

Figure 3: Active Shape Model Applied to Gauge D



The degree of feature matching is quantified by calculating the cross-correlation of the ASM fitted vector and the trial vector, and normalizing the product by the magnitude of the reference vector as follows:

$$Normalized\ Cross - Correlation = \frac{P^T \cdot Q}{R^T \cdot R}$$

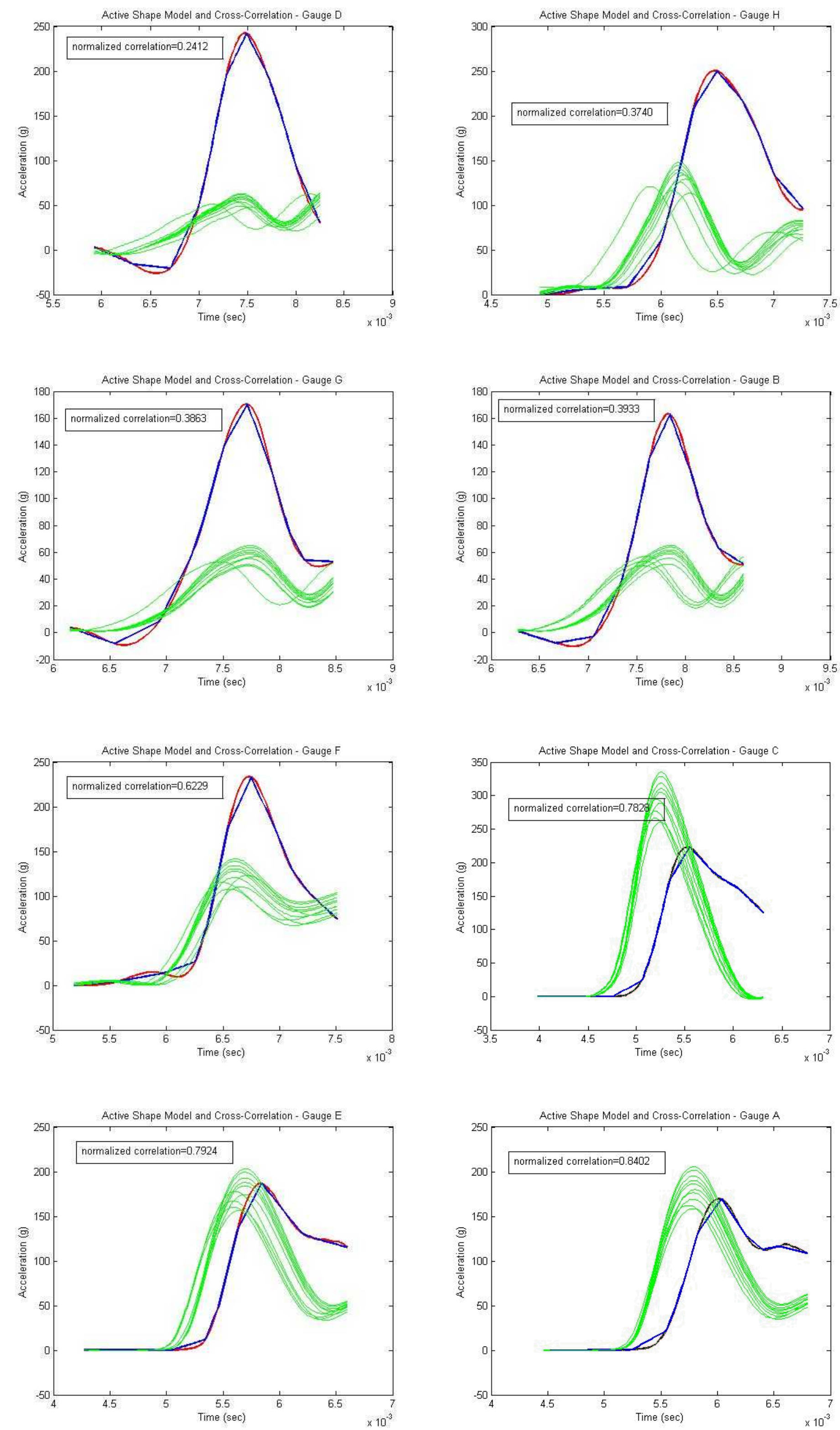
where P – ASM fitted vector,
Q – trial vector, and
R – reference vector.

Possible values for the above equation range from $-\infty$ to $+\infty$ with ideal correlation corresponding to a value of one. Therefore, model validation is quantified by how close the normalized cross-correlation approaches unity. Conversely, model accuracy decreases as normalized cross-correlation deviates from one.

The normalized cross-correlation value of 0.9827 was calculated for the ASM fitted feature (blue curve) and the experimental response (red curve) of Figure 3 where the experimental response was used for both the trial and reference vectors. This high degree of correlation is expected because the red curve is a member of the training set that was used for defining the initial impulse response feature. There are many other algorithms for calculating normalized cross-correlation [10] that may prove to be more suitable for this application.

The steps for feature-based model validation are to ASM fit the PDM-defined initial impulse response feature to the experimental response at a specific gauge location and then calculate the cross-correlation of the fitted feature to the calculated results at the same location. The results of these calculations are shown in Figures 4 where the subfigures are arranged in order of increasing correlation, or matching, for visual effect. The red curve is the original experimental response (reference vector in the above equation), the blue curve is the ASM fitted feature, and the green curves are the model results (trial vector in the above equation). The normalized cross-correlation values listed in the figures are the mean values for each set of ten simulation results. The rank ordering of the results for the feature-based model validation is exactly consistent with the SME judgments used for the previous QoI evaluations [7]. This example demonstrates that feature-based model validation replicates, in a quantifiable manner, the SME process of model validation.

Figure 4: Feature Matching Applied to the Eight Gauge Locations



The above example was conducted for proof-of-concept purposes only since it does not identify the best overall simulation, only the degree of correlation for the individual gauges. It is envisioned, that an actual implementation would combine all gauge contributions using weighted correlation to determine the simulation that best corresponds with experimental results. Weighting the individual gauge correlations would enable the emphasis of the most important response while minimizing the least important behaviors. Development is required to determine the best weighting function for a particular application.

A few observations can be made regarding this first application of feature-based model validation. Sufficient data must be available to define training sets for all relevant features. Knowledge retention is inherent in the definition of the training sets based on SME expertise. The approach is tolerant of noise in the data because only shapes characterized by training set are permissible. Uncertainty in the data is manifested as a spread in the eigenvalues plus, possibly, additional principal components. This latter point requires investigation.

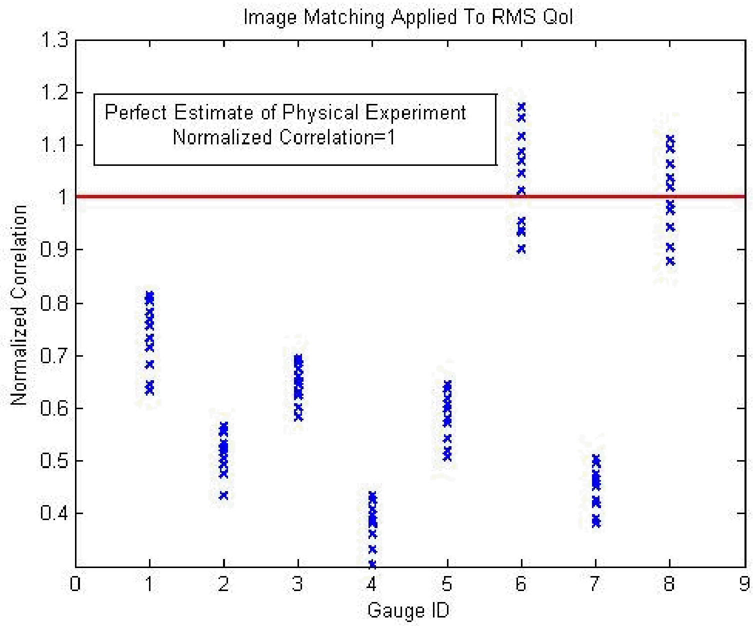
In the initial application described above, the ASM fitted feature was correlated to the analytical results to quantify the degree of validation. An alternate approach would be to ASM fit the features to both the experimental and analytical data, and the resulting eigenvalues could be used as a statistical means of quantifying validation [11]. Also, uncertainty could be propagated in this manner.

IMAGE MATCHING APPLIED TO QoIs

The preceding sections described feature-based model validation and its application. In this section, we explore the use of image processing techniques, particularly image matching, to quantify validation using QoIs. Definition and assessment of QoIs produces vector(s) that must be compared to reference vector(s) to conclude validation. Hypothesis testing has been used [12] to compare QoI vectors but subjectivity is still part of the process. That is, how many elements of a vector must agree before validation is concluded. Image matching [9] provides a methodology to objectively compare vectors and produce a single quantified value of correlation.

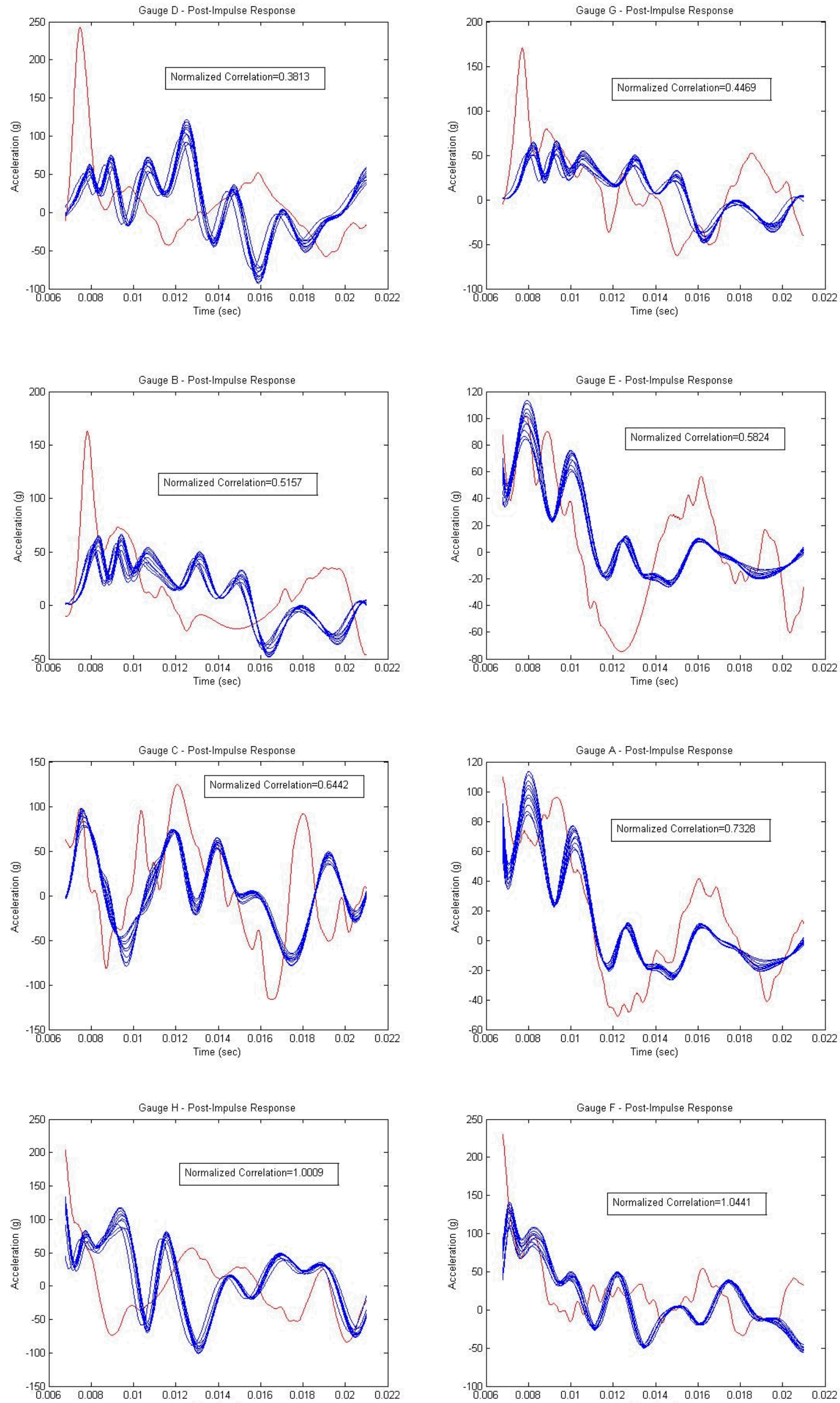
Windowed RMS time signal was identified [7] as the QoI that best correlated with the barge shock test described above. For this work, the windowed RMS QoI was calculated for the post-impulse response using a 1.5 msec interval width with no overlap and ten equally spaced Gaussian windows. The results of these calculations are ten element vectors for each data set evaluated. Image matching in the form of normalized cross-correlation described above is used to quantitatively compare the experimental and analytical data. The image matching results calculated for the eight gauge locations are collectively shown in Figure 5.

Figure 5: Image Matching Results for All Gauge Locations



The image matching results for the individual gauge locations are provided in Figure 6 where the subfigures are arranged in order of increasing correlation, or matching, for visual effect. The red curve is the original experimental response and the blue curves are the model results. The normalized cross-correlation values were calculated using the model results as the trial vectors and the experimental response for both the fitted and reference vectors. The calculated values listed in Figure 6 are the mean values for each set of ten simulation results. The rank ordering of these results is exactly consistent with the SME judgments used for the previous QoI evaluations [7].

Figure 6: Image Matching Results for Individual Gauge Locations



The correlation values for each gauge location must be combined in a weighted fashion for a comprehensive implementation. Development is required to determine best weighting function for a particular application. Normalized cross-correlation was used in this example but numerous alternative algorithms are available for the image matching calculations. Image matching provides a methodology to objectively quantify QoI comparisons thus supporting existing validation framework. The approach is still subjective, but it is less arbitrary than a test-of-hypothesis approach where image matching produces a single measure of validation versus a vector of passes and fails.

CONCLUSIONS

This report represents the first application of feature-based model validation which validates analytical models directly using the key response features identified by a subject matter expert (SME). Mature feature-based image analysis provides the technology for this new paradigm of model validation. It is much simpler to identify key features for a problem class than select a Quantity-of-Interest (QoI) that is suitable for the range of operational parameters. Feature-based model validation is applied to a single barge shock test where acceleration time histories are compared between 10 predictions and the experiment for eight locations on the barge. The Point Distribution Model (PDM) was used to quantify the feature for subsequent identification and matching while the Active Shape Model (ASM) was used to segment the features in the data set. Finally, the degree of matching or validation between the analytical and experimental results was quantified using normalized cross-correlation.

Feature-based model validation was shown to mimic the SME approach to model validation and produce results for the presented example that are exactly consistent with previously published SME validation judgments. As a further demonstration, image matching was applied to objectively quantify QoI comparisons thus supporting existing validation frameworks. The potential for feature-based model validation is extensive including any phenomenon that has identifiable features such as structural dynamics, solid mechanics, data evaluation, numerical convergence, etc. Feature matching can be used for system identification such as the characterization of nonlinear response. Also, an appropriate modeling approach could be identified by first conducting system identification analogous to medical screening. Ultimately, machine learning would be used to identify important features for a problem class provided sufficient data is available.

ACKNOWLEDGEMENT

The author would like to thank Dr. Abhinav Gupta of Carnegie Mellon University for his consultation and support of this work.

REFERENCES

[1] T. L. Paez and J. Red-Horse, "Structural Dynamics Challenge Problem: Summary," *Computer Methods in Applied Mechanics and Engineering*, 197 (2008), pp. 2660-2665.

[2] M. Lillholm, M. Nielsen and L. D. Griffin, "Feature-Based Image Analysis," *International Journal of Computer Vision* 52(2/3), pp. 73–95, 2003.

[3] N. Homma, "Pattern Recognition in Medical Image Diagnosis," *Pattern Recognition*, Chapter 18, InTech, DOI: 10.5772/7538, October 1, 2009.

[4] S. Gould, T. Gao and D. Koller, "Region-based Segmentation and Object Detection," *Advances in Neural Information Processing Systems* (NIPS 2009).

[5] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*, 4th Edition, Wiley-Interscience, New York, NY, February, 2010.

[6] R. J. Allemang, "The Modal Assurance Criterion –Twenty Years of Use and Abuse," *Sound and Vibration*, pp. 14-21, August, 2003.

[7] D. J. Manko and T. L. Paez, "Assessment of Validation Metrics for UNDEX Simulations," Presented at *83rd Shock and Vibration Symposium,* New Orleans, LA, November 4-8, 2012.

[8] http://unicorn.us.com/alex/2polefilters.html

[9] M. Sonka, V. Hlavac and R. Boyle, *Image Processing, Analysis, and Machine Vision, Third Edition*, Thomson Learning, Toronto, Canada, 2008.

[10] N. Roma, J. Santos-Victor and J. Tome, "A Comparative Analysis of Cross-Correlation Matching Algorithms Using a Pyramidal Resolution Approach," *Empirical Evaluation Methods in Computer Vision*, Chapter 6, World Scientific Publishing Co. Inc., Hackensack, NJ, January 1, 2002.

[11] J. Sun, J. He, N. Vlahopoulos and P. van Ast, "Model Update and Variability Assessment for Automotive Crash Simulations," *IMAC-XXV: Conference & Exposition on Structural Dynamics*, Orlando, FL, February 19-22, 2007.

[12] C. Scovel and I. Steinwart, "Hypothesis testing for validation and certification," *Journal of Complexity*, Submitted, 2010.