CREATING REPRESENTATIVE CURVES FROM MULTIPLE TIME HISTORIES OF VEHICLE, ATD AND BIOMECHANICS TESTS

Guy S. Nusholtz
Timothy P. Hsu
Yibing Shi
Sadegh Babaii Kochekseraii
Manuel Alejandro Gracián Luna
Chrysler LLC
USA
Paper Number 09-0249

ABSTRACT

When test responses from specimens such as Post-Mortem Human Surrogates (PMHS), anthropomorphic test devices (ATD) or vehicle crash events are "perfectly repeatable," the response in terms of transducer time histories is similar and the output from any one of the tests can be used to represent any other test. However, if there is test-to-test variability, the underlying fundamental response as obtained by the transducer time history is not determined by a single test and methods are needed that can use multiple tests to reduce the inherent error. This paper will explore, using different transducer time histories from PMHS, ATD and vehicle tests, the effect of signal alignment and signal "shape" on the results from signal addition. New procedures for transducer time history alignment and signal addition will be introduced and discussed, and different methods of obtaining the underlying response will be evaluated.

INTRODUCTION

If measurements subject to random variation about some nominal or "true" value, there is potential to better understand the nominal performance with repeated measurements. For data sets in which each measurement is a single scalar value and multiple measurements are independent, the central mean theorem implies that the mean should be a better estimate of the true value of the measurement than any of the individual measurements. Comparisons of two or more different measurement data sets can be accomplished by comparing the means. However, it is not clear that this approach is valid for comparisons of different sets of finite duration time history measurement, such as: acceleration, force or displacement time history obtained from a human surrogate test or the load time history from a barrier load cell array in a vehicle crash.

Although addition of scalar data is straightforward, the addition of finite time histories is not; for example, defining the numerical procedures such as alignment, individual or accumulative durations, and magnitude of the time histories, to name a few, is subject to interpretation and different definitions could result in different end points. Consequently, there exist a large number of possible methods of signal addition resulting in no unique "best" average signal. Nonetheless, there have been several attempts to combine time history signals to obtain an "average" or "representative" time history [1,2,3,4].

This paper presents two different methods for obtaining a representative time history or "representative curve" (RC) of finite duration time history signals: The first (Procedure A) considers both the shape and magnitude of the time history and the resulting representative signal is constructed by weighing each of the signals by its magnitude; the second (Procedure B) considers only the shape of the signal and the resulting representative signal is constructed by weighing each of the signals equally. In both procedures the signals are shifted to minimize the difference between them and they are then combined. Using the same signals these two procedures can produce different RCs depending on the nature of the signals used in the construction.

Signal Alignment and Representative Curve (RC) Generation

In many cases, signals from a test series taken under the same test conditions do not duplicate well. Many techniques are available to build a RC out of the group. Very often, alignment is necessary to position the signals in time to obtain meaningful results. Figures 1-3 show the different means resulted from the same signals with different alignment schemes. The shapes and curves are different. The magnitude may also be different.

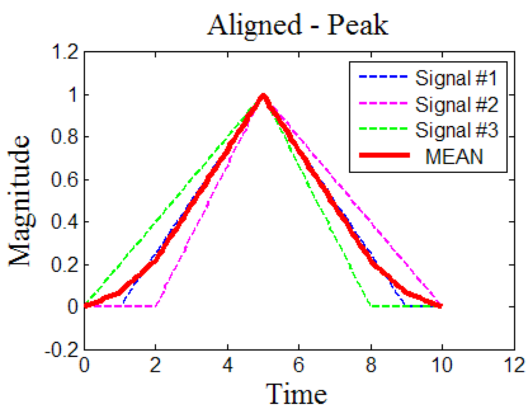

Figure 1 - Mean with Signals Aligned at Peak

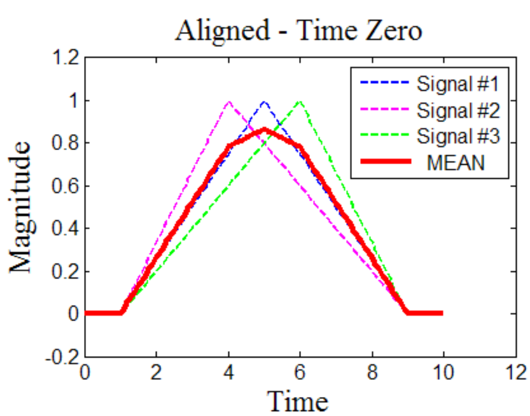

Figure 2 - Mean with Signals Aligned at Time Zero

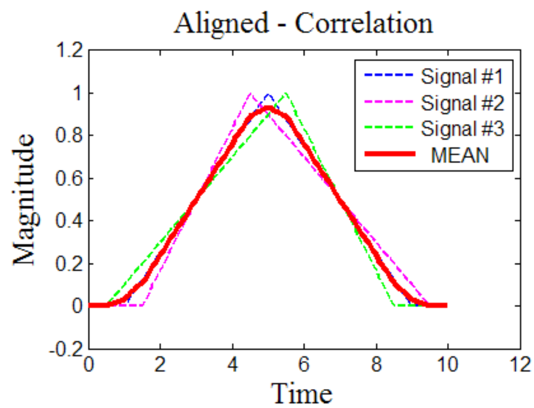

Figure 3 - Mean with Signals Aligned based on Maximum Cross-Correlation

Creating an RC by aligning the time histories so that they can be added is not always trivial. In many cases, adding the signals, after they have been aligned, will distort the "underlying response" and reduce the value of the resulting average: the representative curve is not representative of the curves used to construct it. The key question as to the usefulness of the information in the RC is whether the differences in the signals are dominated by random variations or due to deterministic changes. If they are deterministic, then the RC may be an artifact of the process used and not representative of the underlying response. This question will not be answered in this paper. Instead it will be assumed that there is a fundamental basis to attribute the variation to randomness.

If the variations can be attributed to randomness then statistically speaking, the standard deviation of all the signals can be obtained, and minimizing the covariance or maximizing the correlation will give the best results for alignment. The correlations at the aligned state can be used as an assessment of the quality of the agreement between the signals with emphasis on the “phasing” component of the agreement. It may be supplemented by a measure similar to variance, but normalized at each time step by the mean value of the signals (the coefficient of variation) to give another evaluation of the agreement that emphasizes more on the "magnitude" component of the agreement.

On the other hand, if the difference among the signals is determined to be dominated by deterministic changes among tests, then the above approach does not address the nature of the problem. In this case, the goal would be to identify a master curve that entails the response characteristics of the system, and different curves are then to be "scaled" back to this master curve. An example of this is the acceleration response curves of crash tests of the same vehicle under different velocities, where a second order differential equation can be utilized to model the behavior and "scale" the set of signals. In more general cases, the task will essentially be a system identification problem to define the fundamental characteristics of the system.

What criteria should be used to align signals and to judge the quality of created RCs needs to be decided first. For alignment, commonly used tools are: "eyeballing," "time zero," minimum variance, and maximum correlation. When digital data are absent and correlations are low, especially with old data (non-digital) and different lab facilities, eyeballing presents itself to be the preferred choice. Time zero has the advantage of aligning the event in time, an example being vehicle crash signals in which distinct time zero information is available. However, in many cases, due to vehicle build variation and other confounding factors, the first mode frequencies are often quite different causing the overlaid signals to be inconsistent with the time integrals (as required by conservation of momentum). The variance and the correlation approach, on the other hand, often yield similar time shifts. The starting times do not always line up; however, aligning in many cases ensures a consistent RC.

The following presents two statistical, correlation based methods that build upon the work incorporated into ISO9790 [1] and the Maltese methods [3]. One notable difference is that the current methods do not generate acceptance corridors. Instead they examine and compare the magnitude, shape, and phase of the curves to determine the level of similarity.

Procedure A (Maximal Correlation and Normalization) - Methodology and Characteristics

"Phase," "shape," and "magnitude" are three concepts that have been defined and used in previous studies [4]. Procedure A uses these to establish an RC from multiple time histories which are assumed to have independent random phase, shape, and magnitude variations.

Phase Alignment

With a set of n time history responses $R_i(t)$ ($i=1, 2, ..., n$) for phase alignment, since absolute time is immaterial, without loss of generality, the time for the first response is picked as the absolute time. There are then only $n-1$ time shifts to be found per some requirement. These are denoted as h_i ($i=2, 3, ..., n$).

The coefficient of correlation is used as a measure of the phase agreement between a pair of similar signals,

$$c_{ij} = c_{ij}(h_i, h_j),$$

where $h_1=0$. It is noted that $c_{ij}=1$, when $i=j$.

At this point, a measure is needed that collectively gauges the quality of the matrix $[c_{ij}]$. The most straightforward summary measure would be the sum of all its elements. Based on this, the following normalized alignment measure C is constructed:

$$C(h_2, h_3, ..., h_n) = \frac{1}{n(n-1)}\left(\left(\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}(h_i, h_j)\right) - n\right).$$

Note that $-1 \le C \le 1$. (Since the sign of c_{ij} is significant, the above uses the actual value instead of the absolute value or the square of c_{ij}).

The measure C is a gauge of the quality of the collection of the time shifts. It is a function of the $n-1$ shifts. Maximizing C with respect to these shifts will determine the optimal collective phase agreement. In this study, the unconstrained nonlinear optimization routine in Matlab® was used with minor modifications to avoid local trapping associated with discrete signals.

Shape Extraction

For each of the n phase-shifted responses $X_i(t)=R_i(t-h_i)$, its normalized response is defined to be,

$$x_i = \frac{X_i}{\|X_i\|},$$

where $\|X_i\|$ is a norm defined as

$$\|X_i\| = \sqrt{\int X_i^2 dt}.$$

The integration is used here for ease of expression, and it is to be interpreted as summation if the time histories are treated as discrete signals. The integral, as all others throughout this paper, has limits of (-∞, +∞). All time histories here are assumed to be bounded (i.e., the norm exists). This condition is automatically satisfied by impact test signals which start and end at zero magnitude.

The following time history *y* is defined as the shape representation of the set of time histories:

$$y = \frac{\frac{\sum x_i}{n}}{\left\|\frac{\sum x_i}{n}\right\|} = \frac{\sum x_i}{\left\|\sum x_i\right\|}.$$

In other words, *y* is the normalized version of the average of the normalized responses. The average of its correlation with each of the original signals is found as:

$$p = \frac{1}{n}\sum p_i = \frac{1}{n}\sum \int x_i y dt = \int \frac{\sum x_i}{n} y dt = \left\|\frac{\sum x_i}{n}\right\|,$$

which is the norm of the average of the signals.

p is named the "shape similarity factor" of the original set of signals, as it reflects the overall shape similarity quality based on all the signals.

A special property of *p* is:

$$p = \left\|\frac{\sum x_i}{n}\right\| = \frac{1}{n}\sqrt{\int \left(\sum x_i\right)^2 dt} \le \frac{1}{n}\sum \int x_i^2 dt = \frac{1}{n}\sum 1 = 1,$$

or, $0 \le p \le 1$. The inequality in the above relationship is based on the Minkowski's inequality which basically says that the norm of the sum is no more than the sum of the norms; and the last equality in the expression is because x_i is already normalized.

Magnitude Scaling

The normalized optimal shape *y* established above needs to be scaled back to the physical measurement space to carry an appropriate magnitude. Given that each signal has a magnitude factor, assuming it is randomly distributed, then its sample average is an unbiased estimate of the mean of the magnitude. Therefore, the final representative curve is:

$$Y = \left(\frac{1}{n}\sum\|X_i\|\right) y = \left(\frac{1}{n}\sum\|X_i\|\right)\frac{\sum x_i}{\left\|\sum x_i\right\|}.$$

<u>Procedure B (Mean-To-Mean Approach) - Methodology and Characteristics</u>

The Mean-To-Mean (MTM) methodology is based on a number of available statistical and numerical analysis methods. The major ones are the normalized cross correlation assessment known as cross correlation coefficient of a pair of signals [5]. An approach using an iterative improvement of solution of non-linear equations is also implemented in the procedure (Appendix A).

For the set of signals to be aligned using the cross correlation coefficient, two signals in the group that are most correlated are identified. The pair is aligned using maximum cross correlation process and its sample means calculated. The mean is grouped with the rest in the signal set again replacing the two most correlated signals. All the signal subsets associated with that group pair should be shifted based on the alignment of the pair. This process continues until all signals in the set are aligned using the same procedure.

Additional optimization steps are incorporated in the MTM algorithm, including a prescreening process to identify signal pairs with mutual maximal cross correlation coefficients (CCC). The process is as follows: for a signal set with n signals, CCCs between each signal and another signal in the set are calculated. For each signal, there will be n-1 CCCs. The maximal CCC for each signal is identified. The maximal CCCs for all signals are listed according to their values, from maximum to minimum. Signal pairs with mutual maximal CCC are taken out and put in separate groups. This is a way of identifying the signals with the most influence early in the alignment process and at the same time reducing the effects of any individual signals on the overall performance of the alignment process.

A numerical procedure is generated based on this algorithm. The key element in this algorithm is to evaluate only two signals at a time.

To explain the methodology, an example is shown here with 6 signals: A, B, C, D, E and F.

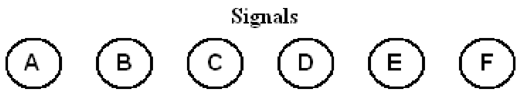

Figure 4 – Time Domain Signals

First, the cross correlation coefficients are obtained with respect to each other in order to identify the pair of signals with the highest cross correlation coefficient. The pair is aligned based on the maximal CCC.

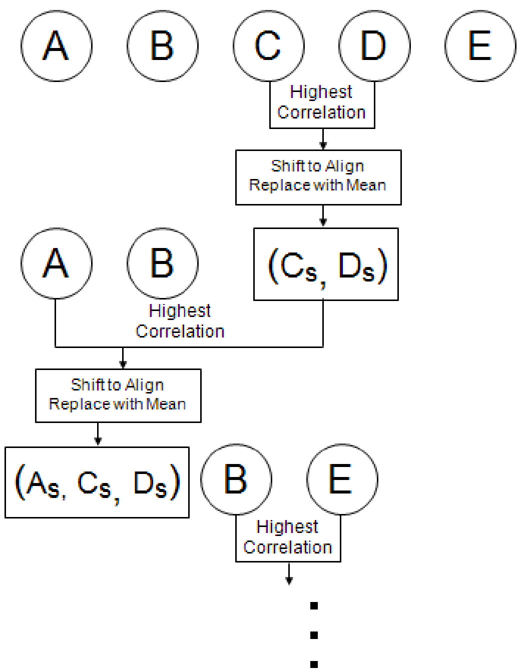

Figure 5 - Procedures for Signal Alignment

Suppose that signals C and D have the highest cross correlation coefficient. They are aligned based on time lag of the maximum cross correlation and their mean obtained as follows:

$$Mean = \frac{C_s + D_s}{2}$$

where C_S and D_S are the shifted signals of C and D. The mean, Mean, then replaces C_S and D_S in the subsequent analysis. The whole process is repeated until a final mean is obtained.

$$Mean = \frac{A_s + B_s + C_s + D_s + E_s + F_s}{2}$$

Further improvement of the final mean or the representative curve is achieved with additional iterations of the process as follows,

- Obtain initial solution
- Repeat the alignment process
- Subtract the error from the solution
- Obtain the improved solution
- Repeat until convergence achieved

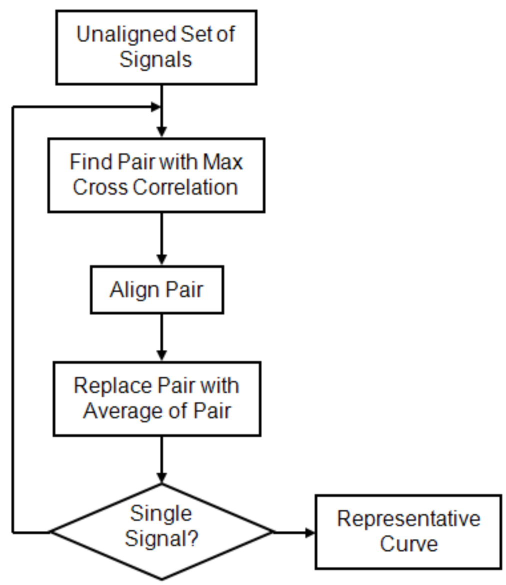

Figure 6 – Flowchart for Signal Alignment

Examples

The methods discussed have application limitations. A variety of data sets, taken from NHTSA' database, has been selected to provide some examples of its range of applicability. The specific units used in the graphs and tables shown are purposely left out, they are for illustrative purposes only and not for direct comparison to real test events. Time is plotted as steps depending on the sampling rate used and cannot be directly related to real time.

PMHS Tests:

PMHS tests are used to characterize the response of the human body to impact. Similar tests carried out on different PHMSs in different labs can result in signals with marked contrasts, creating a challenge to aggregate such contrasting signals and obtain a unique representative signal for the set.

Figure 7 shows the original PMHS data that serve as a base for both methodologies. Figures 8-10 show the results from Method A, Method B and their comparison. Table 1 shows the time shifts (in number of time steps) using the alignment schemes of Method A and Method B.

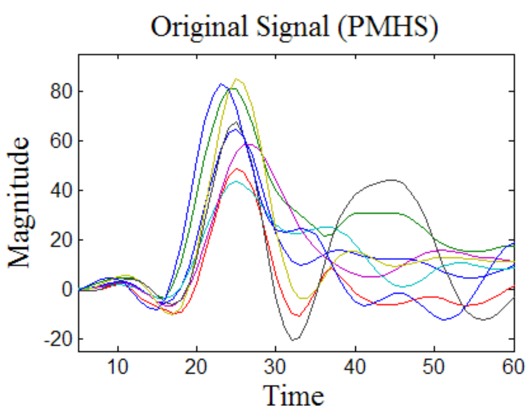

Figure 7 - Original PMHS Signals

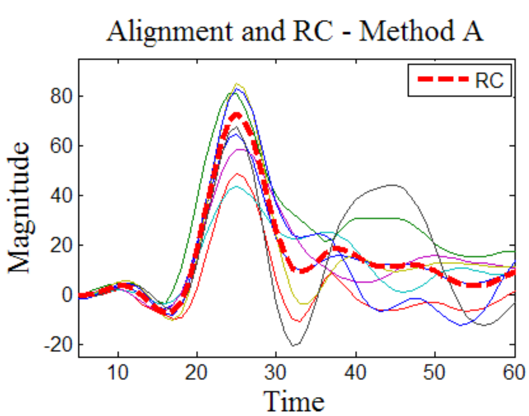

Figure 8 - PMHS Signals Processed (Method A)

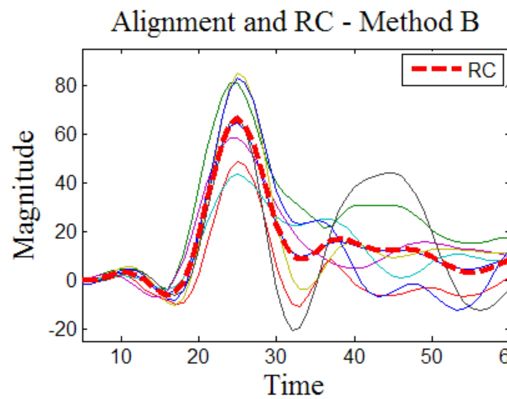

Figure 9 - PMHS Signals Processed (Method B)

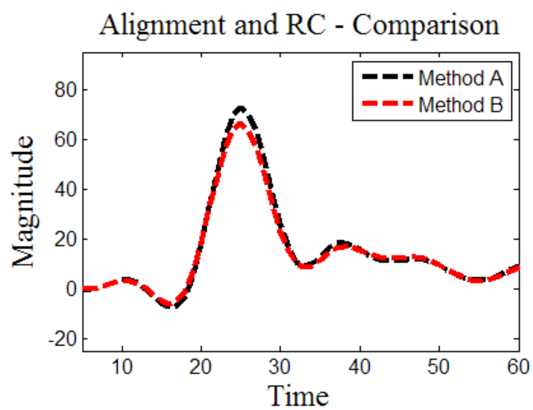

Figure 10 - PMHS Signals Processed (Overlay)

Signal ID	Method A	Method B	Difference (A vs. B)
1	-98	-98	0
2	-98	-98	0
3	-98	-98	0
4	-98	-98	0
5	-99	-100	-1
6	-98	-98	0
7	-98	-98	0
8	-96	-96	0

Table 1 - Time Shifts Comparison (PMHS)

Vehicle Crash (NCAP) Tests:

Vehicles available in NHTSA crash database [6] are classified into compacts cars, sedans, SUVs, minivans and trucks. Frontal rigid barrier forces from NCAP tests were downloaded from the database and summed over the total number of cells in the rigid barrier to obtain the total force of impact for each test. Method A and Method B are used to align and extract a representative curve for the set. Figures 11-14 and Table 2 show the results from the study.

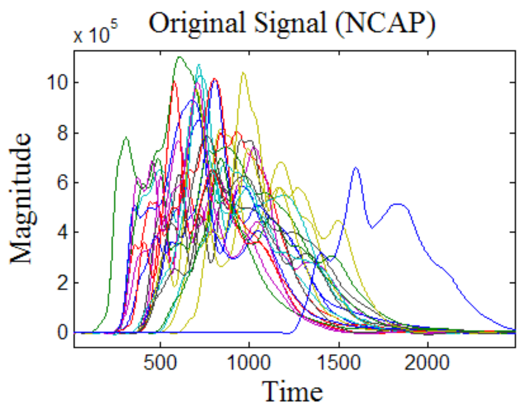

Figure 11 - Original NCAP Signals

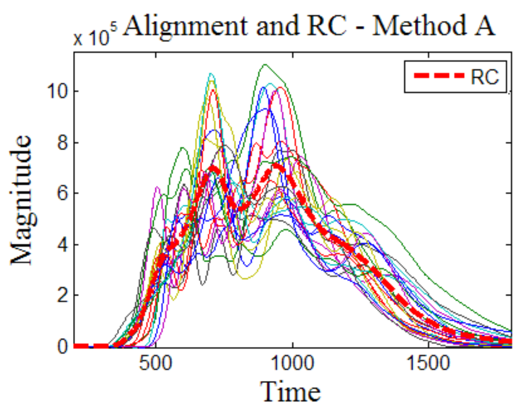

Figure 12 - NCAP Signals Processed (Method A)

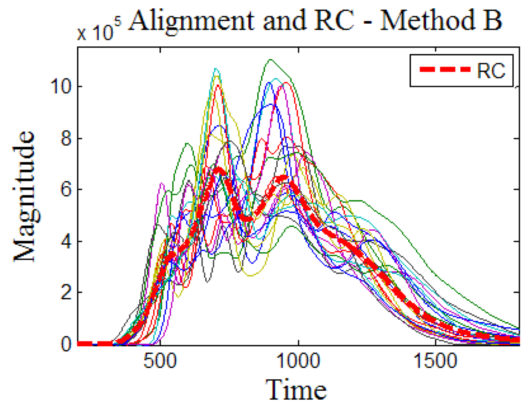

Figure 13 - NCAP Signals Processed (Method B)

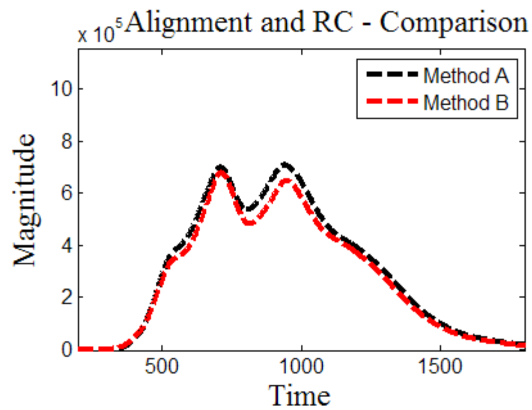

Figure 14 - NCAP Signals Processed (Overlay)

Signal ID	Method A	Method B	Difference (A vs. B)
1	867	867	0
2	1161	1166	5
3	898	898	0
4	1062	1064	2
5	873	872	-1
6	711	708	-3
7	867	869	2
8	957	965	8
9	993	996	3
10	1019	1022	3
11	718	717	-1
12	1000	1000	0
13	853	848	-5
14	805	804	-1
15	1089	1098	9
16	725	722	-3
17	999	997	-2
18	856	852	-4
19	1097	1097	0
20	607	604	-3
21	859	860	1
22	1	0	-1

Table 2 - Time Shifts Comparison (NCAP)

CONCLUSIONS

Two signal alignment methods are presented and used to analyze different types of time domain data. One scheme aligns the signals based on the cross correlation coefficients and normalizes the signals to form a representative curve (RC). The other aligns the signals based on cross correlations and then averages the signals.

The methods are aimed at minimizing the differences between the resultant RC and the signals used to generate the RC. Assuming that the variations from test to test for the transducer time histories are the result of randomness and not deterministic, these methods may be useful for obtaining the underlying response characteristic. The representative curve obtained from these methods may be used for different types of analysis such as determining the biofidelity metrics for ATD design, comparing different ATD responses under similar impact conditions and analysis of different vehicle crash characteristics.

REFERENCES

[1] ISO/TR9790. (1999) Road Vehicles Anthropomorphic Side Impact Dummy - Lateral Impact Response Requirements to Assess the Bio-fidelity of the Dummy. International Standards Organization, American National Standards Institute, NY, NY

[2] Hsu, T. P, Nusholtz, G. S. (2005) Considerations of Bio-fidelity Corridors for Lateral Impacts, SAE Technical Paper Series 2005-01-0308, April 11, 2005.

[3] Maltese, M. R.; Eppinger, R. H.; Rhule, H. H.; Donnelly, B. R.; Pintar, F. A.; Yoganandan, N. (2002) Response corridors of human surrogates in lateral impacts. National Highway Traffic Safety Administration, Washington, D.C./ Wisconsin Medical College, Milwaukee/ Veterans Affairs Medical Center, Milwaukee, Wisc. 31 p. Stapp Car Crash Journal, Vol. 46, Nov. 2002, p. 321-351.

[4] Xu, L.; Agaram, V.; Rouhana, S.; Hultman, R. W.; Kostyniuk, G.W.; McCleary, J.; Mertz, H.; Nusholtz, G.S.; and Scherer, R.. (2000) Repeatability Evaluation of the Pre-Prototype NHTSA Advanced Dummy Compared to the Hybrid III, SAE Technical Paper Series 2000-01-0165, March 6-9, 2000

[5] Ifeachor, E. C.; and Jervis B. W. (1993), Digital Signal Processing: A Practical Approach, Addison-Wesley.

[6] http://www-nrd.nhtsa.dot.gov/database/aspx/vehdb/queryvehicle.aspx

Appendix A: MTM Flowchart