Descriptive Linear Modeling of Steady-State Visual Evoked Response

by

William H. Levison
BBN Laboratories Incorporated
10 Moulton St.
Cambridge, MA 02238

Andrew M. Junker
AFAMRL

Kevin Kenner
Synergy, Inc.

Proceedings of the Twenty-First Annual Conference on Manual Control
June 17-19, 1985
Ohio State University, Columbus, Ohio

ABSTRACT

AFAMRL is currently conducting a study to explore use of the steady-state visual-evoked electrocortical response as an indicator of cognitive task loading. Application of linear descriptive modeling to steady-state visual evoked response (VER) data obtained in the AFAMRL study is summarized in this paper. Two aspects of linear modeling are reviewed: (1) "unwrapping" the phase-shift portion of the frequency response, and (2) parsimonious characterization of task-loading effects in terms of changes in model parameters. Model-based phase unwrapping appears to be most reliable in applications -- such as manual control -- where theoretical models are available. Linear descriptive modeling of the VER has not yet been shown to provide consistent and readily interpretable results.

INTRODUCTION

Considerable effort has been devoted in recent years to the development of reliable metrics for pilot workload. Such metrics could be of value in the areas of cockpit design, pilot training, and flight operations. A measurement technique suitable for in-flight application could potentially warn of impending performance degradation and thereby allow timely remedial action. Assessment of workload in both simulated and operational flight tasks would enhance the identification of workload "bottlenecks", provide additional data for the evaluation of the crew/system interface, and, in general, provide information necessary for maintaining task workload within desired limits throughout a given mission.

Various studies have been undertaken in recent years to develop reliable metrics of pilot workload, including subjective estimates, primary and secondary task measures, and physiologic measures. Exploration of physiologic measures has been motivated by the desire to obtain one or more measures that are non-interfering with the primary mission and are not likely to be biased by the subject's preference for a given man/machine interface or his unwillingness to admit that a particular task is difficult.
AFAMRL is currently conducting a study to explore use of the steady-state visual-evoked electrocortical response as an indicator of cognitive task loading [1]. This paper summarizes the results to date of an effort to characterize the visual evoked response (VER) via linear descriptive modeling. Two applications of linear modeling are reviewed. Part I describes methods for "unwrapping" the phase-shift portion of the frequency response, an issue of concern when analyzing behavioral as well as physiological response. The central issue of this paper -- characterization of task-loading effects in terms of changes in model parameters -- is addressed in part II.

As of the writing of this paper, characterization of task loading effects is still in progress. Part II of this paper is consequently written in the style of a progress report.

PART I: PHASE UNWRAPPING

Nature of the Problem

To obtain the plots of amplitude-ratio ("gain") and phase-shift that are commonly used to characterize the response of linear systems, one typically employs the following procedure:

1. Compute Fourier transforms of the "input" and "output" time histories.
2. Divide Fourier coefficients (or cross-power spectral quantities) at frequencies of interest to obtain estimates of the frequency response as complex numbers.
3. Perform an appropriate nonlinear transformation to express the frequency response in terms of gain and phase-shift.

Various averaging techniques may be performed to enhance the reliability of the results as discussed in [2].

Procedures of this sort necessarily yield somewhat ambiguous phase-shift estimates, because phase repeats every 360 degrees. For example, a negative real number can be considered to have a phase shift of +180 degrees, -180 degrees, -540 degrees, etc. Therefore, we can shift any phase estimate by an integral multiple of plus or minus 360 degrees (one "cycle") and not be at variance with the data. In general, the frequency analysis scheme described so far must be accompanied by a procedure for "unwrapping" the phase in a meaningful way. Otherwise, the frequency shaping of the phase response will have a sawtooth appearance, since Fourier analysis schemes can only identify phase shift within a single cycle (typically, -180 to 180 degrees).

Techniques for Unwrapping the Phase Shift

Certain assumptions must be made in order to derive a method for unwrapping the phase. In the case of manual control data, we usually assume that phase varies relatively smoothly with frequency. That is, we assume that the frequencies at which we obtain frequency-response estimates are sufficiently close together so that successive phase estimates are unlikely to differ by more than 180 degrees. We simply unwrap the phase by adjusting the phase at each measurement frequency by the number of cycles required so that it does not differ from the preceding (in frequency)
estimate by more than 180 degrees. We also assume a reference point for the phase obtained at the lowest measurement frequency — usually 0 or -180 degrees.

The assumption of a smoothly-varying phase response is not always justified, however. For example, unless the frequency-response measurements are finely quantized in frequency space, a highly-resonant system (especially one that is accompanied by significant pure delay) may well exhibit sharp changes in phase-shift in the region of the resonance.

If we wish to avoid the constraint that successive phase measurements differ by less than 180 degrees, we must assume that the phase and gain curves are related to each other in an orderly manner, and we must have a quantitative understanding of the analytic constraints (typically, a linear model) on the gain and phase curves. In this case, the experimental phase-shift is unwrapped with respect to a model-generated baseline.

Although we do not generally recommend that one "adjust the data to fit the model", such adjustments are entirely legitimate provided they are integral multiples of 360 degrees.

In general, the use of a model to unwrap the phase curve implies a model-matching exercise: a single iterative procedure is employed to jointly select parameters to best characterize the data and to unwrap the phase. Ideally, the model used for this purpose is a "theoretical model", i.e., one that is expected on theoretical grounds to provide a good match to the data. Otherwise, a "descriptive" model may be employed which, while having no theoretical justification, is of a form that generates the type of qualitative frequency dependencies exhibited by the data.

The following procedure is suggested for unwrapping the phase via model analysis:

1. Use a theoretical model if one is available. Otherwise, select the least complex descriptive analytic model that seems likely to provide an acceptable match to the data.

2. For theoretical modeling, select an initial set of model parameters based on theoretical considerations or on previous modeling results. For descriptive modeling, important features of the frequency response may be analyzed to provide a reasonable initial parameter selection.

3. Using the current model parameters, predict gain and phase at each measurement frequency.

4. Readjust the experimental phase shift at each frequency, where necessary, by an integral multiple of 360 degrees until the experimental phase estimate is within 180 degrees of the corresponding model prediction.

5. Using an appropriate adjustment scheme and matching criteria, readjust independent model parameters to improve the match to the data.

6. Iterate on steps 3-5 until the matching criteria are satisfied. The resulting adjusted experimental phase curve is substituted for the sawtooth curve originally yielded by the Fourier analysis scheme.

This procedure is based on the assumption that frequency response data are to be matched. Other techniques for parameter adjustment might be employed if modeling is to be applied instead to the relevant time histories.
The validity of this procedure can be judged in a particular application in terms of the resulting model match. If a good qualitative match is obtained to both the gain and phase curves (note: experimental gain is not adjusted), then the resulting adjustments to the phase curve can be accepted as valid; otherwise the phase curve should be unwrapped using another model form.

**Application of Model-Based Phase Unwrapping**

Application of the model-based technique described above is demonstrated for both manual control and physiological response data. A theoretical model is used for the manual control data, whereas a linear descriptive model is employed for the physiologic data.

**Manual Control Example**

Figure 1a shows frequency-response data obtained in a recent simulation of an F-14 performing a steady-state gunsight tracking task [3]. The data points related to phase shift show sharp positive jumps at around 1 and 11 rad/sec because of the -180 and +180 degree boundaries on the Fourier analyzer.

Because these data were obtained in a tracking task employing a known task environment using linearizable vehicle dynamics, the optimal control model (OCM) for piloted systems was used to unwrap the phase. No model-matching was employed; rather, a single prediction of pilot response behavior was generated using pilot-related model parameters typical of those found to match human operator behavior in previous studies. The phase-shift curve was then used as a point-by-point baseline for unwrapping the experimental phase data. As shown in Figure 1b, the initial selection of model parameters gave a qualitatively good match to the data; there was no need to improve the model-match, via parameter adjustment, in order to demonstrate the validity of the unwrapped phase curve.

For this particular data set, the same phase unwrapping is generated by simply assuming that consecutive data points do not differ by more than 180 degrees. Nevertheless, in general, the results are more compelling if they are shown to be consistent with reasonable analytical constraints.

**Application to Visual Evoked Response**

At present, theoretical models of the type available for manual control do not exist for the visual evoked electrocortical response (VER). Unlike the manual control task, where a specific response strategy can usually be derived for accomplishing well-defined control objectives (particularly in a laboratory setting), the VER is not known to have a similar teleological foundation. Unless one is using the VER for biofeedback in a control loop, it is not clear why the electrocortical potentials recorded from the scalp should bear any particular relationship to the visual stimulus. Thus, to the extent that we rely on model analysis to unwrap the steady-state VER phase data, we must currently use descriptive models.

Figure 2a shows the average gain and phase data obtained from a single subject in an ongoing AFAMRL study of steady-state VER. (The details of this experiment are
Figure 1. **Pilot** Frequency Response, Simulated F-14 Gunsight Tracking Task

briefly summarized later in this paper and in more detail by Junker et al [1]). The unmodified phase curve shows upward-directed discontinuities at around 8, 15, and 20 Hz.

Because the gain curve has the general appearance of a second-order resonant lowpass filter, a linear model of the following form was employed to unwrap the phase:

\[
F(s) = \frac{K \omega_0 e^{-sT}}{s^2 + 2\zeta \omega_0 s + \omega_0^2}
\]

where the four independent model parameters are the asymptotic low-frequency gain \(K\), the natural frequency \(\omega_0\), the damping-ratio \(\zeta\), and the pure time delay \(T\). (The frequency variable "s" is not a model parameter.)
An initial selection of parameters was based on the apparent resonance frequency, the asymptotic low-frequency gain, and the difference between maximum and low-frequency gains. In addition, the monotonic and relatively sharp negative increase in phase shift with frequency suggested the presence of a pure delay term, which was also included in the model. The initial estimate of the delay was chosen on the basis of the slope of the phase curve after a preliminary unwrapping in which a 180-degree difference limitation was imposed.

A scalar model-matching error was defined as the rms difference between model predictions and experimental data, weighted inversely by the standard errors of the experimental data. (The unwrapped phase estimates were used for this computation.) Best-fitting model parameter values were identified using a quasi-Newton gradient search scheme similar to that employed previously in manual control studies [4,5].
Because the lowest measurement frequency was relatively large (5 Hz, compared to 0.15 rad/sec for the tracking data), we could not rely on the data of Figure 2a to determine the asymptotic zero-frequency phase shift. It was not obvious whether the asymptotic frequency would be referenced to 0 degrees (implying a positive low-frequency model gain), or -180 degrees (implying a negative gain). Accordingly, model analysis was performed with both positive and negative gains, and results were accepted from the model yielding the smallest matching error. ("Gain" here refers to the scale factor parameter K, specified as a real number, not the amplitude-ratio portion of the frequency response, which is specified in logarithmic units.)

Analysis with the negative gain yielded a substantially lower matching error; the resulting phase curve is shown in Figure 2b. The relatively good qualitative match to the data suggests that the phase curve is likely to be valid, with the possible exception of the phase at the highest measurement frequency.

Application of the same model form to another VER data set is shown in Figure 3 for both positive model gain (Fig. 3a) and negative model gain (Fig. 3b). For this data set, the two model-matches yielded nearly identical matching errors, but the unwrapped phase curves differed by 360 degrees. Apparently, the -180 degree phase shift imposed on the model predictions by the negative gain shifted the predicted phase response sufficiently to require an extra 360 degrees of unwrapping in order to minimize model-data differences.

Because we have no theoretical basis for determining the asymptotic low-frequency phase shift (equivalently, the sign of the model gain parameter), and because the qualitative matches to the data sets are equally good (though different in detail), the two phase curves must be considered equally valid. Thus, the phase unwrapping remains to some extent ambiguous when a second-order resonant loss-pass filter is adopted as the model form. Other model forms might provide unambiguous results, but that would have to be determined from trial and error.

PART II: LINEAR MODELING OF STEADY-STATE VER

Background

Prior research has indicated that recorded scalp electrical potentials respond, to some extent, in a manner linearly related to the visual stimulus. There is, in addition, a strong nonlinear component of the response, plus a substantial amount of unrelated ongoing electrical activity that is present. Under proper stimulus conditions, the linear component of the response is large enough to allow its estimation with reasonable statistical confidence. Thus, this electrophysiological system lends itself to the analytical techniques employed in pilot/vehicle analysis -- i.e., to the measurement of describing function and remnant -- as has been demonstrated above in Part I. The focus of the ongoing research, to which this paper is addressed, is to determine whether such measures are sensitive to workload and other forms of stress.

As noted earlier, we cannot define a "purpose" for the visual evoked response, in the sense that we can for control response in a well-defined tracking task. Not only do we lack a theoretical model for what the evoked response ought to be, there is no obvious functional relation between the response (electrical potentials measured at the scalp) and the demands of the "task" (which may be no more specific than to attend to or fixate on the stimulus). Therefore, our basis for interpreting visually evoked response is not as solid as our basis for interpreting manual control response, and intra- and inter-subject variability tends to be substantially greater than with
manual control response behavior. The averaging technique described elsewhere in these Proceedings by Levison [2] were developed largely to deal with this variability.

A number of research efforts have focused on obtaining a frequency-response description of the VER [6-9]. In what is perhaps the most comprehensive effort to date, Spekreijse [9] measured the VER using inputs consisting of single sinusoids (as opposed to a sum-of-sinusoids), or single sinusoids plus Gaussian noise. His work focused a great deal on characterizing the nonlinear aspects of the response. On the basis of numerous sub-experiments, Spekreijse concluded that nonlinear response components in the VER were due largely to memoryless rectification and saturation nonlinearities and that these nonlinearities were located prior to the "cortical selective process". If this model is correct, then nonlinear VER components are not influenced by the operator's cognitive state, and we are justified in characterizing task-related VER changes in terms of quasi-linear model parameters even though the VER may contain significant nonlinear response components.

Figure 3. Visual Evoked Response, Example 2
Second-order Lowpass Resonant Filter
More recently, Junker and Peio [10] obtained steady-state evoked responses to sum-of-sinusoids visual stimuli. They found that, although the nature of the frequency response varied from subject-to-subject, it appeared to be relatively stable for a given subject across replications, and to be influenced by the task environment. Preliminary analysis of their data revealed that, for at least some of the data sets, the frequency response could be reasonably well characterized by a second-order linear descriptive model.

Experiments

Details of the VER experiment are provided in a companion by Junker et al. [1]. A brief overview is given here.

Electrocortical response was recorded from subjects exposed to spatially uniform light stimulus modulated by a complex sum of sinusoids. Ten sinusoidal components of uniform amplitude and random phasing were used, with component frequencies ranging from 6.25 to 21.75 Hz.

Three task loading conditions, provided in a balanced order, were explored: (a) no explicit task, other than attending to the flashing lights, (b) a first-order manual tracking task, and (c) a grammatical reasoning task. Analysis techniques similar to those applied extensively to manual control analysis were employed here to obtain the frequency response characteristics of the VER. Response metrics consisted of amplitude ratio ("gain") and phase shift, measured at stimulus frequencies, and "remnant" (response components at other than input frequencies) averaged over 1-Hz "windows" centered about each input frequency. Only the gain and phase data are considered here.

Data from seven subjects were considered statistically reliable and were made available for model analysis. Each VER frequency response considered in this paper represents the average of from six to eight 40-second segments of electrocortical recordings. Averaging was performed as described by Levison [2].

Model Analysis

Model analysis was performed as described in Part I. The objectives of this analysis were to unwrap the phase to aid in overall interpretation of the frequency response, and to determine whether or not the independent model parameters would provide a parsimonious and consistent characterization of task-loading effects.

As noted above, preliminary results led us to believe that a lowpass filter of the type defined in Equation 1 would characterize the steady-state visual evoked response at least for a portion of the subject population. Data from all seven subjects were initially modeled in this manner. Positive and negative gains were tested, and whichever sign yielded the smallest matching error was included in the parameterization for a given data set.

Application of the second-order model did not yield consistently useful results, either for phase unwrapping or for interpretation of the evoked response. The resonant lowpass filter provided a good qualitative match to only a portion of the data sets; for data where the match was not qualitatively acceptable, the validity of the resulting phase curve had to be questioned.
The best-fitting model parameters did not reveal a consistent trend with task loading, and they tended to vary over wide ranges from one data set to the next. Nearly as many data sets were best matched with a positive model gain parameter as with a negative gain. This result implies that the polarity of the recording electrodes was changed from one condition to the next — a notion at variance with the experimental procedures followed in this study.

Even where a good qualitative match was obtained, the resulting model parameters were often inconsistent with the assumption of a stable linear system. For example, the model fits shown in Figures 3a and 3b were obtained with negative damping ratios — a characteristic of a system whose oscillatory response grows exponentially with time. Such a result is inconsistent with electrocortical responses obtained with transient stimuli. When subsequent model analysis was performed with the constraint that the damping-ratio and natural-frequency parameters remain positive, substantially greater matching errors were obtained in most cases.

Inspection of the data (specifically, the gain curves) suggested that other model forms would more closely resemble the frequency dependency of the data. Figure 4 shows an example of a data set matched with the following fourth-order bandpass filter:

\[
F(s) = K \cdot \frac{s^2}{s^2 + 2\zeta \omega_1 s + \omega_1^2} \cdot \frac{\omega_2^2 e^{-sT}}{s^2 + 2\zeta \omega_2 s + \omega_2^2}
\]

This model also has four independent parameters: gain, two natural frequencies, and delay. (The damping ratios were fixed at 0.707.)

By constraining the two frequency parameters to be positive, we were able to characterize the data with a stable linear system. Analysis with this model form was not conducted on a large scale, however, because of the sensitivity of the results to the initial parameter selection — a situation not uncommon when employing gradient search schemes.

The difficulty of obtaining a consistent model-based characterization of the steady-state VER is indicated by inspection of the gain curves shown for two test subjects in Figure 5. For the baseline (no-task) condition, the data for Subject 2 (Fig. 5a) resemble the frequency response of a resonant lowpass filter, whereas the data for Subject 3 (Fig. 5d) resemble an inverted "v" and are perhaps modeled by a tuned bandpass filter. (The data shown in Figure 5a were used for the demonstration of phase unwrapping in Figure 2.)

The curves for the tracking condition (Figures 5b and d) show no consistent effects of task loading: the data from Subject 2 reveal regions of diminished response, whereas the data from Subject 3 show less of a qualitative change from the baseline. For the grammatical reasoning condition, however, both subjects showed gain response curves that appeared to vary less with frequency than the baseline.
The trends revealed in Figure 5 suggested the hypothesis that the gain response is "flatter" for the reasoning task than for the baseline condition. Accordingly, data from the first three test subjects providing complete data sets (Subjects 2, 3, and 5) were modeled with a simple gain/delay model of the form:
Figure 3. Effects of Task Loading on Visual Evoked Gain Response
where \( K \) and \( T \) are the "gain" and delay parameters, respectively. The reasoning behind this test was that, if the flat-response hypothesis were true, this model form would yield lowest matching errors for the grammatical reasoning condition.

Figure 6 (bottom graph) shows that, for the three subjects tested (time did not permit testing of the entire data base), the gain/delay model yielded the lowest matching error for the reasoning task, thereby providing some quantitative support for the qualitative trend suggested above. Testing of the remaining data is required to explore the generality of the hypothesis. Visual inspection of the frequency response yielded by the other subjects (not shown here) suggests that this trend will not hold for the entire subject population.

The top two graphs of Figure 6 show that task loading conditions did not have a consistent effect on the gain and delay parameters across the three subjects. This simple model form, then, appears to be of use only for testing some very general data trends -- not for parameterizing the VER in a meaningful way.

DISCUSSION

The use of a model to unwrap the phase-shift response is not uncommon, but it is usually informal and implicit. Typically, the individual performing the analysis has an expectation of what the frequency dependency should be, based on previous experience with similar systems, and unwraps the data according to a qualitative "mental model". What we have done here is to suggest that the procedure be made more explicit with the use of a specific mathematical model, with a combined procedure of phase unwrapping and parameter adjustment if need be. Provided a suitable model structure is available, with a solid basis for initial parameter selection, such a procedure provides a means for automated phase unwrapping.

Although preliminary results encouraged the application of linear descriptive models of the VER, modeling of this form has not been demonstrated so far to be a reliable method for characterizing task loading effects. Although model forms can be found to provide a reasonable qualitative match to the data, the appropriate model form appears to vary across subjects and sometimes across tasks, parameter variations do not follow a clear trend, and model parameter values are not always consistent with a stable response mechanism.

It is tempting to conclude that the relative lack of modeling success (in terms of our stated goals) is due, in part, to the fact that we are attempting to model a nonlinear response mechanism with a linear model. We do not think this is a major factor. However nonlinear the VER might be, it does contain a measurable and
Figure 6. Effects of Task Loading on Parameters of a Gain/Delay Model
generally statistically reliable linear response component. If task loading were to change the response behavior in a consistent manner, we would expect the linear response component to change in a consistent manner.

It is possible that we have not explored the appropriate model forms. To the extent that model analysis is pursued during the remainder of this study, model forms that have a structure based more on theoretical considerations [11,12,13] will be explored. Another avenue to be explored is the effect of task loading on the variability of the VER, rather than the mean [14].

A more likely source of the difficulty is that there is no "reason" for the electrocortical potentials to exhibit a particular pattern, in terms of what the subject is trying to accomplish. To create a situation closer to that of manual control tasks, where generation of a particular response behavior can aid the achievement of task-related goals imposed upon the test subject, it is anticipated that the AFAMRL study will explore the use of the evoked response in a continuous control task employing biofeedback. A task environment of this sort is expected to reduce the variability of the VER and make it more sensitive to task loading. The use of the VER as an "unobtrusive" measure of task loading may be compromised, however, as the VER will now be a component of a secondary task competing for attention with the primary cognitive (or psychomotor) task.

Inspection of the available data base suggests that there may be important intersubject differences in terms of the linear response behavior. Thus, while not yielding a consistent index of task loading, linear analysis may prove viable as a means for characterizing subject differences. It remains to be established whether such differences, if found to be statistically significant, relate in a consistent manner to behavioral aspects of interest, and not simply to physical characteristics such as differences in the shape of the skull.

Finally, we note that the "remnant" (background EEG) remains to be analyzed. Although the effects of task loading and individual differences appear to be smaller for the remnant than for the main curve, it is possible that remnant changes are statistically more significant.

ACKNOWLEDGEMENT

The work reported herein was supported by the Air Force under contract number F33615-84-C-0515.

REFERENCES


