

## **STEERING ENTROPY REVISITED**

Erwin R. Boer  
LUEBEC  
San Diego, California, USA  
E-mail: [erwinboer@san.rr.com](mailto:erwinboer@san.rr.com)

Michael E. Rakauskas, Nicholas J. Ward  
Dept. of Mechanical Engineering  
University of Minnesota  
Minneapolis, Minnesota, USA

Michael A. Goodrich  
Dept. of Electrical Engineering  
Brigham Young University  
Provo, Utah, USA

**Summary:** Drivers aim to maintain their vehicle within a number of individual situated safety margins. Safety margin violations are characterized by rapid strong corrective steering. Steering entropy was introduced to quantify drivers' efforts to maintain their lateral safety margins. In the original steering entropy, several computational assumptions were made. The objective is to scrutinize and motivate these choices and exemplify the effects of deviations from these choices with data from a driver distraction study. The new optimized algorithm is shown to yield significances where a number of classical metrics fail to find any significance. Its sensitivity is attributed to the fact that a number of observed changes in steering behavior all manifest in a widened steering prediction error distribution which the algorithm picks up sensitively with its log-based weighting of prediction error outliers and its use of a prediction filter that is maximally sensitive to the spectral characteristics of the baseline data.

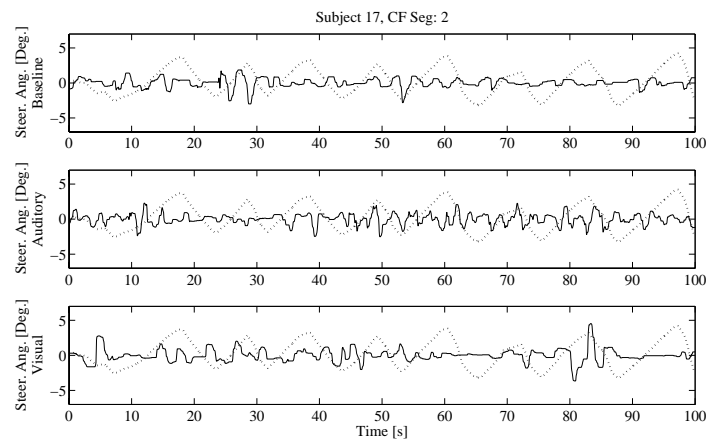
## **INTRODUCTION**

Degradations in drivers' ability to maintain safe operation of their vehicles manifests in changes in perception or control. These differences can be quantified most sensitively using drivers' control actions. Interpretation in terms of safety and performance requires sensing of the consequences of drivers' actions. It is important to recognize that: (i) a given safety or performance level can be achieved with a number of different behaviors, and (ii) drivers adapt their safety and performance criteria to cope with the total task demands. Lane keeping can be characterized as the process of maintaining time to line crossing within a self-imposed situated tolerance range. In response to changing task demands, drivers exhibit: (i) a change in the power spectral density of their steering control actions, and/or (ii) a change in tolerance margins. In general, as task demands increase, drivers exhibit an increase in steering power at certain frequencies and an increased number of safety margin violations. Steering entropy (SE) was developed to specifically quantify the increase in high frequency steering corrections that result after periods of diverted or reduced attention (i.e., in response to a perceived vehicle drift outside the acceptable tolerance margins that mounted during these periods of degraded information

pickup about the lateral vehicle state). Focus is directed to increasing the sensitivity of the SE metric to be able to better assess changes in steering behavior in response to changes task load.

## EXPERIMENTAL DATA

Steering data from a driver distraction study that took place in the UMN VESTR driving simulator was used (Rakauskas et al., 2005). In the section of data used in this study, subjects followed a lead vehicle for 120 s whose speed fluctuated roughly sinusoidally with a high (i.e., CF Seg. 2) cyclically switching frequency between 0.06 and 0.12Hz and an amplitude between 55 and 75 mph (see dotted lines in Figure 1). Twelve of the 48 subjects either performed no task (baseline), a hands-free cell phone conversation (audio task condition), or a set of common in-vehicle tasks (visual task condition). Steering profiles for all conditions for subject 17 are shown in Figure 1.



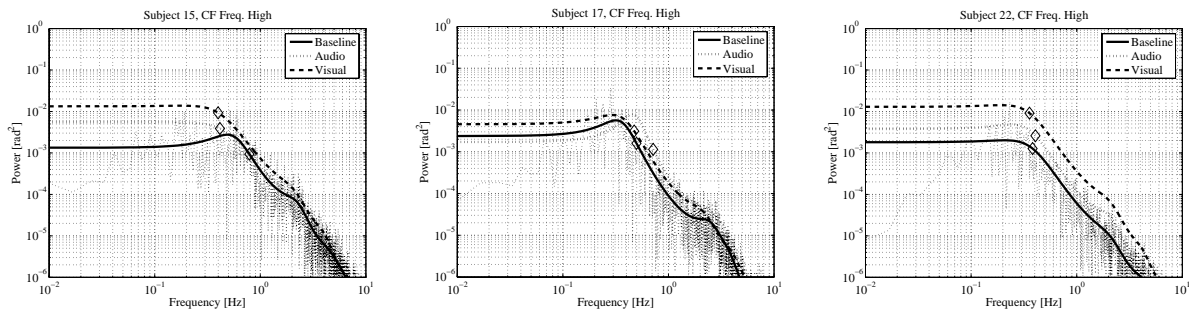
**Figure 1. Raw steering angle time series of subject 17 for baseline (top), cell phone task (middle), and in-vehicle tasks (bottom). Deviations from median vehicle speed [m/s] are shown as dotted lines.**

## SPECTRAL STEERING ANALYSIS

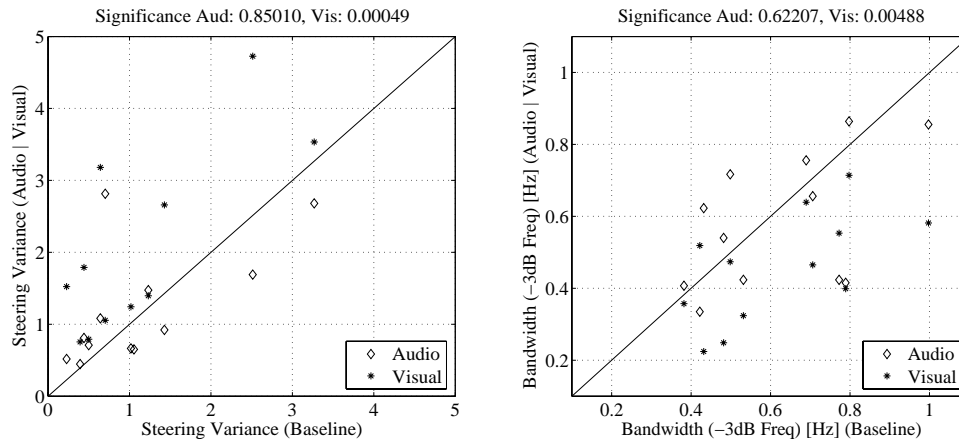
Steering data in Figure 1 shows several changes to in-vehicle tasks: (i) the cell phone task shows a higher frequency response (see discussion of power spectra in Figure 2 for alternative coping strategies/effects), (ii) the cell phone task shows fewer strong steering correction (possibly due to enhanced vigilance or relaxed safety margins), (iii) the visual tasks shows more strong steering actions (corrections), and (iv) the visual task also shows higher frequency steering actions.

The set of in-vehicle tasks results in an increase in steering power at all frequencies and is therefore easily differentiated from baseline driving by a number of steering performance assessment metrics; Figure 3 shows that steering variance increases and that bandwidth decreases for nearly all subjects. In contrast, the cell phone task does not yield significant effects with these metrics. However, as shown in Figure 7, the optimized SE algorithm does yield a significant effect for the cell phone task. The reason for this is that SE does not focus only on magnitude like steering variance nor only on frequency like bandwidth; it focuses essentially on the derivative (i.e., the prediction errors that result from a high-pass filtering of the steering data). Note that the derivative of a sinusoidal signal  $A \sin(2\pi ft)$  is  $2\pi f A \cos(2\pi ft)$ , whose power

increases with amplitude as well as frequency. This is important because drivers adapt or cope differently to in-vehicle tasks (particularly the cell phone task in this study). As Fig. 2 shows, three types of shifts in the power spectral density are observed in drivers' steering profile during the cell phone task (thick dotted lines); some subjects show primarily an increase at the low frequencies (e.g., subject 15), some primarily an increase in high frequencies (e.g., subject 17), and some subjects show an overall increase in power at all frequencies similar to what is observed in response to the visual tasks (e.g. subject 22). Because of these various coping strategies, magnitude-only or frequency-only metrics recognize too few subjects as having different steering profiles to yield overall significance. SE does not suffer the same fate (Figure 5 shows that the occurrence rate of high prediction errors increases for all three subjects in Figure 2). This shows that SE is a promising candidate to mine for differences, but that further spectral analyses should be conducted to assess the type of spectral change since these are associated with different types of coping strategies (e.g., vigilance increase, safety margin adaptation, stress, etc.).



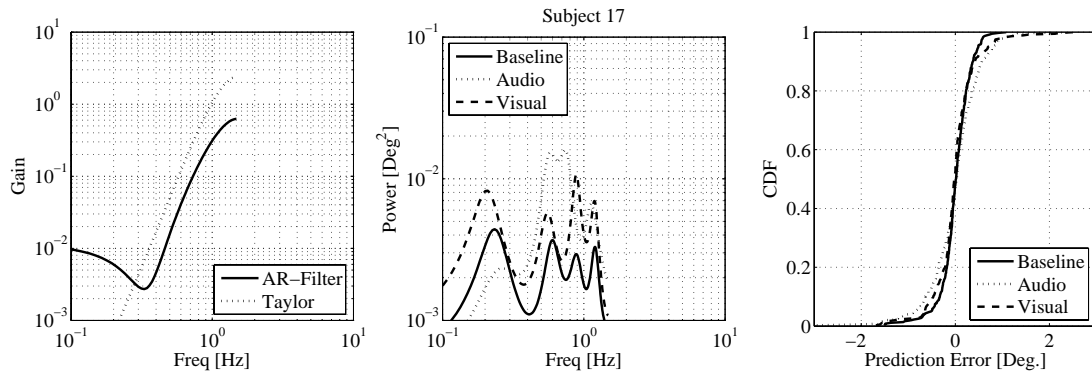
**Figure 2. Power spectral densities of the three task conditions for three representative subjects. The thin dotted line is the raw baseline power spectral density. The three thicker lines are 10<sup>th</sup> order Burg approximations for the three driving conditions; -3dB bandwidths (i.e. where power drops below 70.7% of the zero-frequency power) are indicated by diamonds.**



**Figure 3. Comparison between baseline (abscissa) and task condition (ordinate) on steering variance (left) and bandwidth (right). The significance values for the cell phone (audio) and visual tasks (visual) are shown in the panel titles. Each diamond or asterisk reflects one subject.**

The original SE method attributed its sensitivity primarily to high frequency corrective maneuvers. These corrective steering actions do indeed account for a small increase in the high frequency power (as observed in all subjects), however, the steering maneuver following these initial jerky corrective responses is of lower frequency and much longer duration. Thus, an increase in low frequency power is also expected. This is one of the reasons why the optimized SE algorithm yields superior sensitivity using a lower re-sampling frequency and a prediction filter with less low frequency attenuation than the original one, thereby increasing sensitivity to low frequency changes while maintaining sensitivity to high frequency changes in steering power.

The original SE algorithm was based on a Taylor expansion to generate prediction errors (PEs) from a steering profile sampled down to 7Hz. In Boer (2000), an alternative method was proposed that uses a PE generating “filter” derived from an AR-model of the reference baseline steering data. The advantage is that it is optimally tuned to the frequencies present in the baseline steering data, which depends on a number of factors such as driver skill, road curvature, speed, lateral disturbances, vehicle dynamics, etc. (Figure 2 demonstrates that subjects adopt different spectral signatures in their steering behavior). Because of the optimally frequency tuned AR-model, an SE method based on AR-models is also optimally sensitive to deviations from this baseline steering spectrum. The comparison between original Taylor expansion and AR-model shown in Figure 7 confirms that the AR-model approach yields higher significances, and in some cases yields significance where a Taylor-based expansion does not. Naturally this needs to be interpreted from the perspective that if a highly sensitive tool is needed to yield significance, the effect is probably small. However, in driving, it is very important to employ the most sensitive metrics available because a small change in driving behavior (e.g., reaction time) can have very large consequences when converted to annual fatalities.



**Figure 4. PE filters (left), the power spectral density of PEs (middle), and the PE distribution (right) for case-subject 17. The right two panels are based on the AR-model with a re-sampling frequency of 4Hz and an alpha of 0.2 (see below for details).**

The Taylor expansion and the AR-model derived PE filters are both high-pass filters (left panel, Figure 4). In fact, the AR-model-derived filter is an MA-filter whose MA coefficients are simply the AR coefficients of the model fit of the steering data (Boer, 2000). The 3<sup>rd</sup> order AR-model of the steering time series  $\{s_n\}$  is (2<sup>nd</sup> and 4<sup>th</sup> order yield similar SE sensitivity to task conditions):

$$s_n = -a_1 s_{n-1} - a_2 s_{n-2} - a_3 s_{n-3} + p e_n$$

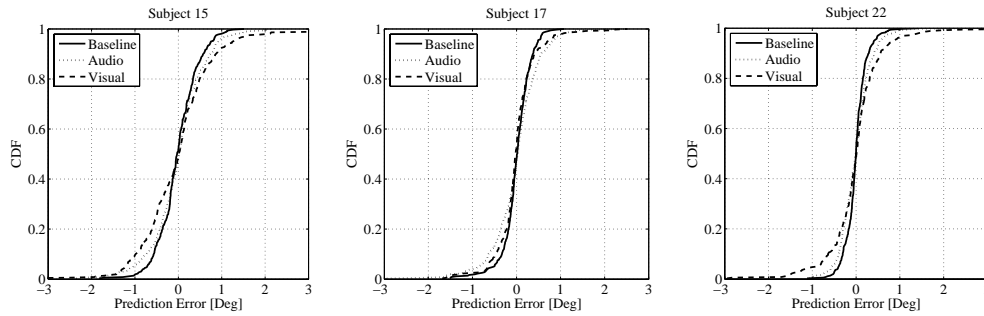
where the  $a_i$  coefficients are estimated with the Burg algorithm (Matlab Version 6.5, function “arburg”) and the PE time series  $\{pe_n\}$  is maximally white noise. This AR-filter is converted to the following MA PE generating filter (Boer, 2000):

$$pe_n = s_n + a_1 s_{n-1} + a_2 s_{n-2} + a_3 s_{n-3}$$

The corresponding 2<sup>nd</sup> order Taylor expansion prediction filter is (Nakayama et al., 1999):

$$pe_n = s_n - 2.5s_{n-1} + 2s_{n-2} - 0.5s_{n-3}$$

Fig. 4 shows the filter gains, the spectral density of the output PEs, and the PE distributions. The dip around 0.35Hz in the left panel of Figure 4 is reflected as a peak in the solid line in the center panel of the corresponding baseline power spectral density for subject 17. The PE distributions for subjects 15, 17, and 22 (also in Figure 2) are shown in Figure 5. Again, the PE distribution shows an increased number of high PEs compared to baseline for both task conditions even though the power spectral densities in Figure 2 do not show a consistent effect for the cell phone task.



**Figure 5. PE distributions from AR-model of subjects whose power spectral densities are shown in Figure 2. Adopted sampling frequency is 3Hz (i.e. the one deemed optimal).**

Prior to estimating the AR coefficients and applying the prediction filters, the steering data needs to be re-sampled down (see discussion of Figure 6 for choice of optimal re-sampling frequency  $f_s$ ). Prior to re-sampling, the steering data is filtered with a 5<sup>th</sup> order low-pass Butterworth filter (i.e. maximally flat in pass-band) with a cutoff frequency of  $3/7 f_s$  (just below Nyquist frequency). The AR-model is estimated on the first 60 s of the baseline data (reference). The baseline entropy is computed based on the second 60 s. This assures independence of the data used in the significance tests between baseline and task conditions. In this study, the AR-model-derived PE filter was estimated for each subject’s baseline reference data separately and used only on that subject’s unused baseline data and its condition trials to assess the shift in PE distributions from baseline to condition. (This differs from using all available baseline reference data to estimate a common baseline prediction filter.) The baseline PE filter is used to generate the baseline reference PEs  $pe_n^{bas,ref}$ , the baseline PEs  $pe_n^{bas}$ , as well as the task condition PEs

$pe_n^{cond_m}$  :

$$pe_n^{bas,ref} = s_n^{bas,ref} + a_1^{bas,ref} s_{n-1}^{bas,ref} + a_2^{bas,ref} s_{n-2}^{bas,ref} + a_3^{bas,ref} s_{n-3}^{bas,ref}$$

$$pe_n^{bas} = s_n^{bas} + a_1^{bas,ref} s_{n-1}^{bas,ref} + a_2^{bas,ref} s_{n-2}^{bas,ref} + a_3^{bas,ref} s_{n-3}^{bas,ref}$$

$$pe_n^{cond_m} = s_n^{cond_m} + a_1^{bas,ref} s_{n-1}^{bas,ref} + a_2^{bas,ref} s_{n-2}^{bas,ref} + a_3^{bas,ref} s_{n-3}^{bas,ref}$$

where subscript  $m$  denotes subject  $m$ , where superscript  $bas_m^{ref}$  refers to the first 60 seconds of subject  $m$ 's baseline data, and where  $bas_m$  refers to the second 60 seconds of the baseline data.

The PE filter is the main component of the SE algorithm. The rest is essentially a non-linear weighting of PEs by placing extra focus on extreme PEs as those are associated with the strongest and fastest steering corrections arguably associated with most taxing or critical situations. To measure the degree to which the PE distribution widens under different conditions, the probability density function of the baseline PEs is used. The PE distribution is approximated with  $2(M+1)$ -bins as follows. The set of lower bin bounds is

$\{-10e12, -(M)pe_\alpha, -(M-1)pe_\alpha, \dots, -pe_\alpha, 0, pe_\alpha, \dots, (M-1)pe_\alpha, (M)pe_\alpha\}$  and the set of upper bounds  $\{-Mpe_\alpha, -(M-1)pe_\alpha, \dots, -pe_\alpha, 0, pe_\alpha, 2pe_\alpha, \dots, (M-1)pe_\alpha, (M)pe_\alpha, 10e12\}$ , where the extreme range of  $-10e12$  and  $10e12$  assure that all PEs are captured, and  $pe_\alpha$  is defined as:

$$pe_\alpha = 0.5 \left( \left| \arg\{CDF(pe) = \alpha\} \right|_{pe} + \left| \arg\{CDF(pe) = 1 - \alpha\} \right|_{pe} \right)$$

which is essentially the absolute PE that is greater than a fraction  $(1 - 2\alpha)$  of all PEs. An alpha of 0.05 corresponds to the 90 percentile used in the original SE algorithm.

Next, the probability that a reference baseline PE falls in a bin is established. This is achieved by simply counting what fraction of reference PEs falls in a bin. This may leave some of the outer-bins empty. To avoid extremely low probabilities and thus extremely high "entropies," all bins with probabilities less than  $1.0e-3$  are replaced by  $1.0e-3$  (this is necessary to avoid that some PEs from non-reference baseline or condition data receive an excessively high sample entropy, thus assuring that the method does not become extremely sensitive to one or two outliers but requires a reasonable number of high PEs to substantially increase the entropy of the condition under investigation). Bins are indexed by  $k \in [1, K]$  where  $K = 2(M + 1)$ . The resulting discrete probability density function is denoted with  $P^{bas_m^{ref}}(pe_n)$ , where  $bas_m^{ref}$  refers again to subject  $m$ 's reference baseline data that was also used to establish the bins.

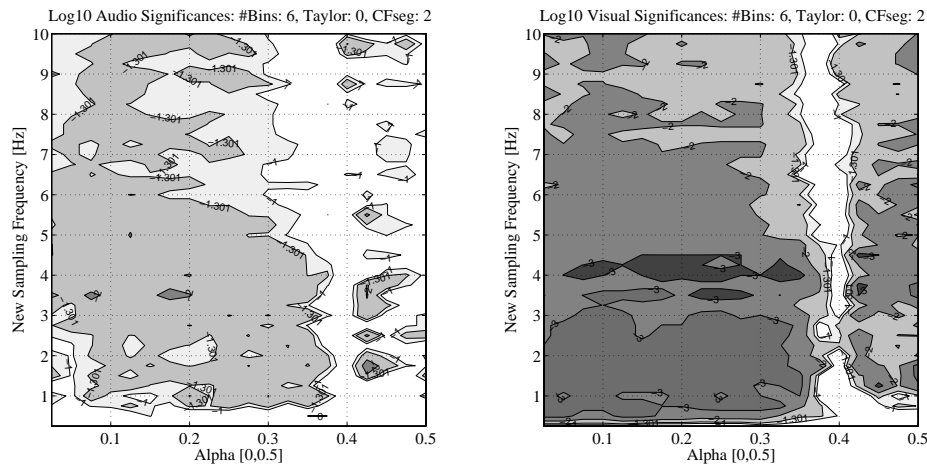
The entropy calculation of the binned PE distribution assigns high weight to outliers (i.e., those PEs that fall in low probability bins). The entropy of subject  $m$ 's second half baseline data as well as the cell phone (audio) and in-vehicle tasks (visual) condition are computed using:

$$H^{cond_m} = \frac{\sum_{n=1}^{N^{cond_m}} -\log_2(P^{bas_m^{ref}}(pe_n^{cond_m}))}{N^{cond_m}} = \sum_{k=1}^K \left\{ -\frac{N_k^{cond_m}}{N^{cond_m}} \log_2(P_k^{bas_m^{ref}}) \right\} = \sum_{k=1}^K \left\{ -P_k^{cond_m} \log_2(P_k^{bas_m^{ref}}) \right\}$$

where  $P_k^{bas_m^{ref}}$  is the probability associated with subject  $m$ 's bin  $k$ ,  $N_k^{cond_m}$  is the number of task condition PEs from subject  $m$  that fall in bin  $k$  of subject  $m$ 's discretized reference-baseline PE distribution. This equation is slightly different from the original one to yield higher sensitivity; the old equation only uses the bins from the baseline reference but not the baseline probabilities as argument to the  $\log_2$  function. Note that log base 2 is used in reference to information bits.

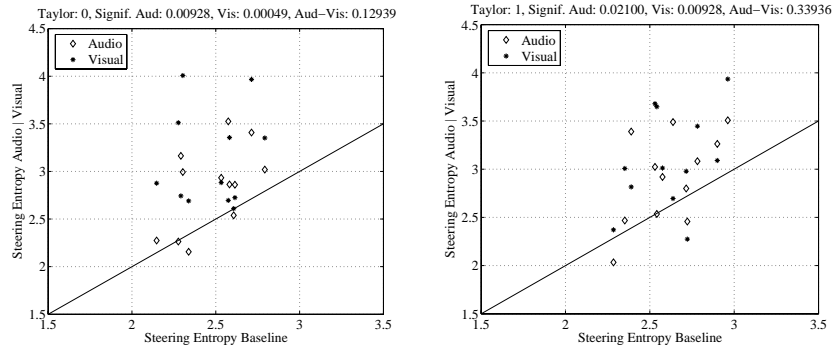
The free parameters of the new SE algorithm are now the re-sampling frequency, the alpha value, the number of bins, and whether a Taylor expansion or an AR-model-based prediction

filter is used. Their effect in terms of significances between baseline and condition per subject is shown in Figure 6, which shows results for the AR-model-based prediction filters. It is clear that a re-sampling frequency of 4Hz and an alpha of 0.2 yield a significance that is robust to changes in these parameter values (i.e., fall in the center of an area with similar significance values). Using 4-10 bins yields less significant results (i.e., the algorithm is less sensitive), but with 12 to 32 bins the same results are obtained. The recommended number of bins is set at  $K=14$  (i.e.,  $M=6$ ), to avoid that small data sets results in too many empty bins. The Taylor-based PE filter showed a slightly different pattern of sensitivities to the re-sampling frequency and alpha, but also gave the most sensitive and robust assessment of steering behavior changes with a re-sampling frequency of 4Hz and an alpha of 0.2. These differences are primarily due to the fact that the AR-Filter method weighs low frequencies more than Taylor, which greatly attenuates the low frequencies (e.g., left panel Fig. 4). Given that some subjects only show an increase in low frequency power during the cell phone task, the Taylor method is unable to differentiate their steering profiles from that of their baseline condition, thus yielding a lower overall significance between conditions across all 12 subjects. This is also reflected in the significances that both methods yield for the optimal parameter setting in Fig. 7.



**Figure 6. Effect of re-sampling frequency and alpha for the AR-model based PE method with #bins equal 14 (i.e.  $M=6$ ) on the log base 10 of the significance between baseline and cell phone (left) and between baseline and in-vehicle tasks (right). Note that the darkest area in the right panel has the same significance boundary (i.e., 0.001) as the one shade lighter (i.e., plotting code bug).**

The optimality of a 4Hz re-sampling frequency is attributed to the fact that the resulting frequency range (0-2Hz) spans drivers natural frequency range. It is clear from Fig. 7 that a very low re-sampling frequency eliminates high frequency effects from the analysis (some subjects only show a signature effect at high frequencies during cell phone task – Subj. 17 Fig. 2). On the other hand, a very high re-sampling focuses the AR-model too much on high frequencies that are beyond the natural operating range of the human driver, thereby yielding an overall reduction in sensitivity; furthermore, the power spectral densities shows very little differences at the highest frequencies thus yielding a PE distribution that focuses too much on the noise.



**Figure 7. Comparison between AR-filter (left) and Taylor (right) implementation of the SE method with optimal re-sampling frequency of 4Hz and an alpha of 0.2**

## CONCLUSIONS

A new steering entropy algorithm is recommended with a re-sample frequency of 4Hz, an AR-model-based prediction filter, and an alpha value of 0.2 rather than the original algorithm, which operated with a parameter setting of re-sampling frequency of 7Hz, a Taylor prediction filter, and an alpha of 0.05. The new settings yield significances where the old settings did not because they are more finely adapted to be sensitive to the different manners in which spectral changes in steering occur in response to changing driver task demands. It is recognized that optimality is conditioned on the dataset at hand, but that a high degree of confidence is attributed to these recommendations because the spectral shape of the observed steering data is very similar to those observed in other simulator and real-world data sets. Future studies will be subjected to the same in-depth analysis touched on in this paper to further improve our ability to sensitively assess drivers in their attempts to cope with the variety of self- and externally imposed driving demands.

## ACKNOWLEDGMENTS

The analyses in this paper were made possible by support from Nissan Motor Co. Ltd. The experimental data was generated at UMN with Grants from MnDOT and UMN's CTS.

## REFERENCES

- Boer, E. R. (2000). Behavioral Entropy as an Index of Workload. *Proceedings of the IEA 2000 / HFES 2000 Congress*.
- Nakayama, O., Futami, T., Nakamura, T., and Boer, E.R. (1999). *Development of a Steering Entropy Method for Evaluating Driver Workload*, SAE Technical Paper Series: #1999-01-0892. Presented at the International Congress and Exposition, Detroit, Michigan, March 1-4, 1999.
- Rakauskas, M.E., Ward, N.J., Bernat, E.M., Cadwallader, M., Patrick, C.J., de Waard, D. (2005). Psycho-physiological measures of driver distraction and workload while intoxicated. In, *Proceedings of the 3rd International Symposium on Human Factors in Driving Assessment, Training, and Vehicle Design*, June 27-30, 2005.