MULTI-SCALE SIGNAL FEATURE PROCESSING FOR AUTOMATIC, OBJECTIVE VEHICLE NOISE AND VIBRATION QUALITY ANALYSIS

David A. Whitney and Bruce Broder

TASC
55 Walkers Brook Drive
Reading, MA 01867-3297
dawhitney@tasc.com

ABSTRACT

This paper describes a new signal processing technique for understanding the dynamics of time-varying signals in vehicles: *Hyperstate analysis.* Vehicle noise and vibration are examples of randomly-varying transient or non-stationary signals that are not effectively analyzed with classical spectral analysis techniques. By the use of nested Hidden Markov Models, Hyperstate analysis explicitly identifies transient and nonstationary behavior on many time scales for better signal discrimination. It uses a probability-based framework that allows for automated, objective classification of noisy signals. The technique is applied to engine starting sequences from different types of vehicles. This work demonstrates that Hyperstate analysis discerns similarities and differences in randomly-varying signals of this type, and can perform effective automatic, objective classification and signal decomposition for NVH (noise, vibration, and harshness) studies.

1. INTRODUCTION

This paper describes a new signal processing technique for understanding the dynamics of time-varying signals in vehicles: *Hyperstate analysis.* Starting sequences, blower motor squeaks, instrument panel rattles, trasmission gear noise, wiper motor noise, and power accessory noise are all examples of transient or non-stationary vehicle noise that cannot be effectively analyzed with classical spectral analysis techniques. Time-frequency representations of signals like the Wigner-Ville and Choi-Williams distributions can provide better information about time-varying signals, but can be difficult to interpret for complex signals, and can explicitly describe variation for only a single time scale. Hyperstate analysis explicitly identifies dynamic features on many time scales for better signal discrimination, and uses a probability-based framework that allows for automated objective classification of noisy signals. The probability framework is more powerful than a simple pattern or template matching approach because it accounts for random variations in signatures from vehicle to vehicle and over time.

A Hyperstate framework can serve as the core automated signal processing technique that is combined with psychoacoustic analysis to replace subjective human jury evaluations with an objective, automated quality evaluation. Hyperstate techniques can also be used to determine the sensitivity of passengers to various sounds or vibrations having complicated multi-level patterns and for the diagnosis of mechanical problems.

2. THEORETICAL BACKGROUND

Hyperstate analysis was developed to address the general problem of classification and estimation of nonstationary stochastic processes. The Hyperstate solution is given by multiresolution stochastic modeling and associated nonlinear optimal filtering. This new technique provides dynamic models for discrete or symbolic data. Dynamics of the signal are modeled on mutliple time or space scales (resolutions). At each scale, banks of discrete Hidden Markov Models (HMM) are used to represent alternative process dynamics. A hierarchy of nested HMMs results, and statistical likelihoods for these nested models are computed rigorously using nonlinear filtering generalizations of the Kalman filter described in [4].

The Hyperstate model structure is depicted in Figure 1. Figure 1a shows an example of a standard discrete Markov chain with four states ($A_{1,k}, B_{1,k}, A_{2,k}, B_{2,k}$). The model in the example permits transitions between any two states every T_k time units according to the transition probabilities specified by the links. All of the dynamics in the model take place on the same time scale. In many physical systems, however, different dynamics occur on different time scales. For example, Figure 1a might actually be used to represent a composite system formed from two linked submodels: one consisting of states $(A_{1,k}, B_{1,k})$, the other of states $(A_{2,k}, B_{2,k})$. Transitions among states *within* a submodel may occur at the rate T_k, while transitions *between* the sub-models occur at the (slower) rate T_{k+1}. While the model illustrated in Figure 1a does allow dynamics on both the T_k and T_{k+1} time scales to be described, it does so only *implicitly*, and can lead to a much more complex, less descriptive, and less computationally efficient model than is required to describe the actual physical process.

Hyperstate models *explicitly* represent dynamics on multiple time scales, as illustrated in Figure 1b. The two sub-models are now represented by two states, $A_{1,k+1}$ and $B_{1,k+1}$, in the Hyperstate hierarchy: $A_{1,k+1} \Leftarrow (A_{1,k}, B_{1,k})$, and $B_{1,k+1} \Leftarrow (A_{2,k}, B_{2,k})$. Transitions between the sub-

0-7803-2431-5/95 $4.00 © 1995 IEEE

model states $A_{1,k+1}$ and $B_{1,k+1}$ are explicitly represented as occuring on the (longer) time scale T_{k+1}. These new states are "hyperstates", since they represent HMMs themselves, not simple states. Hierarchies with arbitrary numbers of layers of nested HMMs corresponding to arbitrary dynamic resolutions $(\dots, T_k, T_{k+1}, \dots)$ can be specified.

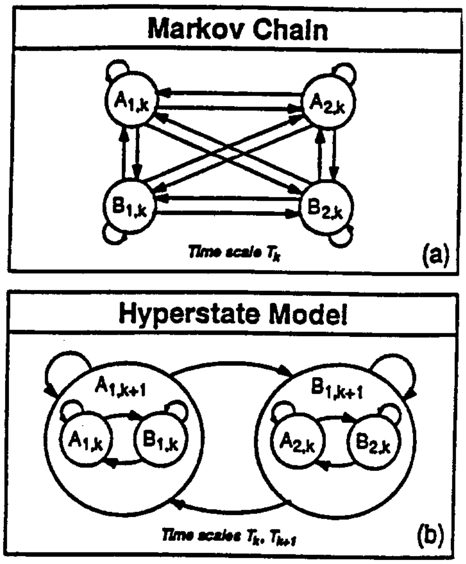

Figure 1: (a) Conventional and (b) Hyperstate Markov Models

The Hyperstate modeling and filtering technique has several important features:

- Nonstationary, time-varying signatures are modeled
- Dynamics on multiple time scales are explicitly described
- Rigorous probability measures are produced for signature discrimination
- A new computationally tractable algorithm results – constraints on state transition paths yield a simpler structure
- The filtering is optimal with respect to the modeled hierarchy
- The explicit identification of system states at multiple resolutions provides added insight into signal features – dynamics on different time scales are simultaneously modeled
- High-fidelity classification and estimation are achieved by propagation of full likelihood information across resolutions – no intermediate decision thresholds that could result in information loss are used
- Rigorous data fusion involving numeric and symbolic information at different resolutions can be performed.

Hyperstate modeling is a new general technique for rigorous *multiresolution* probabilistic modeling of *nonlinear* stochastic processes using a hierarchy of discrete HMM models. Single-resolution HMM models have been widely used in speech processing [2]. A much less general two-stage hierarchy for word recognition has also been suggested [3]. Techniques for multiscale signal processing using linear models driven by Gaussian noise have also been proposed [1].

3. ALGORITHM DESCRIPTION

Processing starts at the 0-th Resolution by applying a signal-to-symbol transform to provide a discrete label (symbol) sequence and its associated likelihoods that characterize the stochastic process. These are then processed by a hierarchy of nonlinear filters. Each filter's dynamics are determined by a distinct HMM. Multiple HMMs can exist at each resolution level.

The algorithm hierarchy is described by resolutions $i = 1, 2, \dots, R$, with corresponding time scales $1 = T_0 < T_1 <, \dots, < T_{R-1}$. Resolution i has models $j = 1, 2, \dots, J_i$. Using notations and conventions based on [4], the observed output at time t of a HMM model is y_t^{ij}, and the corresponding hidden state variable is x_t^{ij}. Φ^{ij} and H^{ij} are the state-transition and state-to-output observation matrices of HMM model j at resolution i. The (m,n) entry of Φ^{ij} is the conditional probability $Pr(x_{t+1}^{ij} = m \mid x_t^{ij} = n)$. The (m,n) entry of H^{ij} is $Pr(y_t^{ij} = m \mid x_t^{ij} = n)$. The vector of likelihoods of the states for model j in resolution level i at time k is $\mathcal{L}_k^{ij}$. The likelihood of model j in resolution level i at time k based on all data at resolutions 0 through i is λ_k^{ij}. The vector of likelihoods of all models at resolution i is $\lambda_k^i = (\lambda_k^{i1}, \dots, \lambda_k^{iJ_i})^T$.

The Hyperstate algorithm flow is given in Table 1.

4. ANALYZING VEHICLE NOISE AND VIBRATION

Hyperstate filtering techniques are demonstrated on an acoustic data set for passenger vehicle engine starting sequences involved in two experiments. In the first experiment, each of four different vehicle types (A, B, C, and D) was started a given distance from a reflective surface and the acoustic signatures were digitized and recorded. The second experiment recorded data for the same vehicles positioned farther from the surface.

In the present analysis, the four starting sequence signatures from the first experiment are used to develop the Hyperstate models and filters. The model structures are then fixed and the second set of signatures is processed. The classification performance on the second data set is then determined. The limited amount of training data (one example per vehicle) and the fact that the second experiment has slightly different test conditions than the first make this is a difficult classification problem. This data set is very limited, having only one recorded example for each scenario and vehicle. Nevertheless, the Hyperstate analysis is successful in characterizing the different features of each vehicle.

The starting sequence is approximately 3 seconds in duration, and four time resolutions were chosen: $T_0 = 3$ msec, $T_1 = 30$ msec, $T_2 = 300$ msec, $T_3 = 3$ sec for filtering Stages 0, 1, 2 and 3. A single channel of stereo data collected by a head recording device, downsampled from 44.1 kHz to 22.05 kHz, was used. Then a state-space model for

TABLE 1
Hyperstate Algorithm

For resolutions $i = 1, 2, \ldots, R$

Initialize prior state distributions $\mathcal{L}_{0(+)}^{ij}$
for all models $j = 1, 2, \ldots, J_i$

- **INPUT**
  Model likelihoods λ_k^{i-1} from resolution $i-1$ at rate T_{i-1} for times $k = T_{i-1}, 2T_{i-1}, \ldots, T_i \pmod{T_i}$[1].

- **NONLINEAR FILTERING FOR STATE LIKELIHOODS**

  For times $k = T_{i-1}, 2T_{i-1}, \ldots, T_i \pmod{T_i}$
  For models $j = 1, 2, \ldots, J_i$
  Extrapolate state likelihoods
  Update state likelihoods
  Next model
  Next time

- **FORM MODEL LIKELIHOOD**

  For models $j = 1, 2, \ldots, J_i$ at times $T_i \pmod{T_i}$
  Compute model likelihood
  Compute new prior distribution
  Next model

- **OUTPUT**
  Model likelihoods λ_t^i for resolution i at rate T_i, for times $t = T_i, 2T_i, \ldots$

Next resolution

[1]The term "(mod T_i)" signifies that the steps for resolution $i-1$ are repeated for successive time intervals of length T_i. This provides outputs at times $T_i, 2T_i, \ldots$ that are inputs to the resolution i processing.

the average acoustic signature of one of the four vehicles is computed. This model serves as a *whitening filter* for all the data so that the analysis corresponds to looking at departures from the "average" signature which will distinguish between vehicles. The complete Hyperstate analysis hierarchy is depicted in Figure 2.

After the whitening procedure, the data are fed into the Stage 0 bank of filters. The filters are matched to a selected set of spectral shapes corresponding to possible departures from the white (average spectral signature) noise over a 3 msec interval. The filters consist of an all-pass filter (for the white noise model), a low-pass filter with a cutoff frequency of 500 Hz, two band-pass filters with passbands of [500 Hz, 2 kHz] and [2 kHz, 8 kHz], and a high-pass filter with a corner frequency of 8 kHz. These filters are chosen using psychoacoustic information and approximately correspond to an equal division of the frequency domain with respect to regions of maximum sensitivity for human perception. The likelihoods of the five spectral models (denoted W, L, B1, B2, and H) are passed on to the Stage 1 processing.

With inputs at 3 msec intervals and outputs at 30 msec intervals, the Stage 1 Hyperstate models describe patterns in the spectral shapes of Stage 0 over a ten time-step interval. Over such short intervals, it is reasonable to consider patterns of two dominant spectral shapes from the set of five possible labels. Thus, twenty-five (5^2) two-state Markov chains are used to model the temporal dynamics between any pair of the four Stage 0 models. The likelihoods of these models are computed and used as inputs to the Stage 2 processing.

Stage 2 combines ten outputs (at 30 msec time-steps) from Stage 1 to produce one output every 300 msec. Over these intervals it is assumed that patterns of four spectral shapes, described as pairs of Markov models from Stage 1, characterize the transients. Since there are twenty-five models at Stage 1, 625 (25^2) two-state Markov chains are used to model the temporal pattern at Stage 2. The likelihoods of the models are computed as outlined in Table 1.

Based on the Stages 1 and 2 outputs from the *first* experiment, Stage 3 Hyperstate models were designed to classify these data into one of the four vehicle types, A through D. The four Hidden Markov Models at Stage 3 each had an observation space of dimension 625 corresponding to the Stage 2 outputs. The classification probabilties output at Stage 3 after each of the four signatures from the first experiment were processed through the entire Hyperstate hierarchy are shown in Table 2. A column in this table corresponds to processing a single starting sequence through each of the models for the vehicle types. For example, Column B indicates the probabilities corresponding to each Hyperstate model after processing the starting sequence from vehicle B. Using this data, the model for B is found to have a probability of 0.91 while the model for C has probability 0.09. As expected, each signature is correctly classified with classification probabilities near one — in this case, the classifier is being tested against the same data used to develop it.

The Hyperstate architecture developed above to match the data from the first experiment was used to process the data from the second experiment. These data were sequestered and not used to develop the models. The classification probabilities are shown in Table 3.

TABLE 2
Classification Probabilities
First Experiment Models and Data

Training Experiment

MODEL TYPE	DATA TYPE A	B	C	D
A	1.00	0.00	0.00	0.00
B	0.00	0.91	0.00	0.00
C	0.00	0.09	1.00	0.00
D	0.00	0.00	0.00	1.00

Note that vehicle A is correctly classified using a maximum *a posteriori* probability criterion, even though the

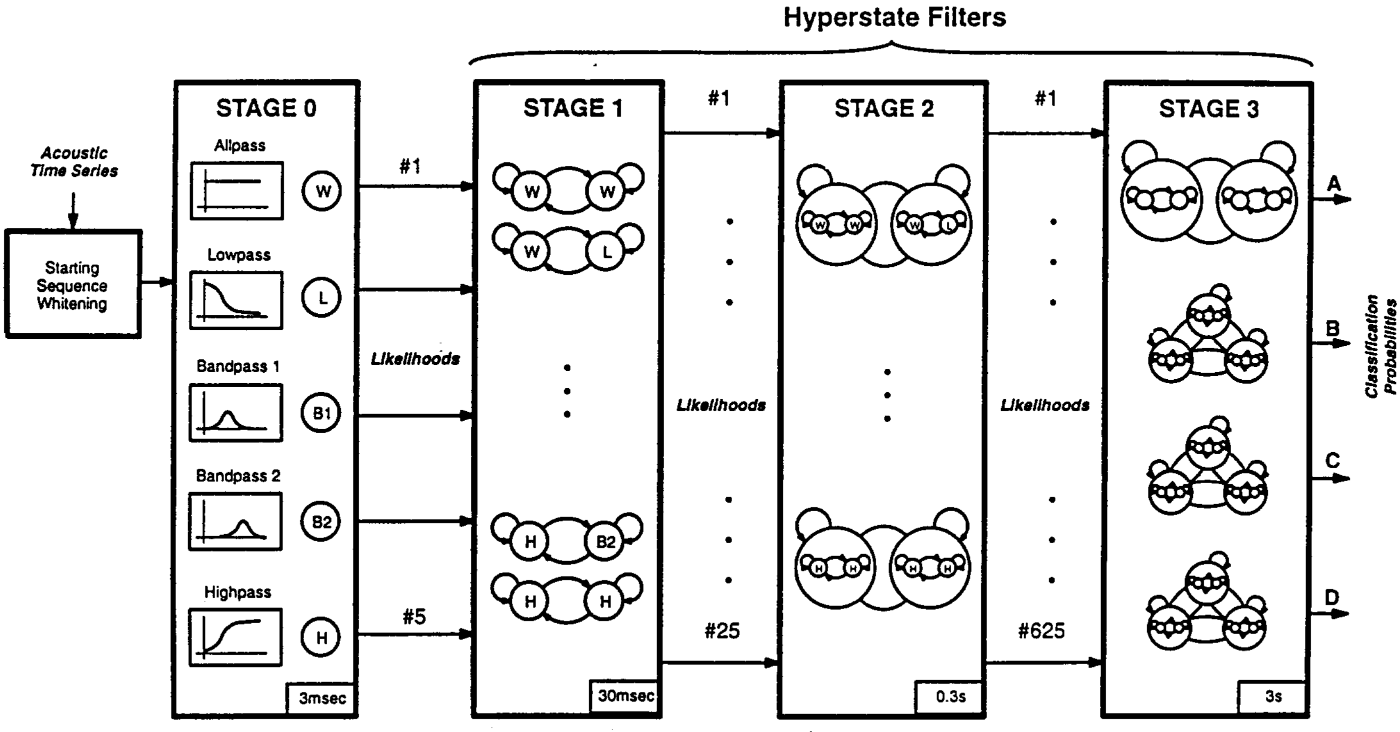

Figure 2: Hyperstate Filtering Architecture for Staring Sequence Data

TABLE 3
Classification Probabilities
First Experiment Models and Second Experiment Data

Testing Experiment

MODEL TYPE	DATA TYPE			
	A	B	C	D
A	0.45	0.00	0.00	0.00
B	0.00	0.99	0.00	0.96
C	0.37	0.01	1.00	0.04
D	0.18	0.00	0.00	0.00

probability is only 0.45. Vehicles B and C are correctly classified with very high classification probabilities. Vehicle D is incorrectly classified as type B. Considering that only a *single* sample of training data was available, and that the training and sequestered data exhibited different vehicle–wall reflection patterns, the results are quite good.

5. SUMMARY

Hyperstate analysis is a new framework for modeling and filtering of nonstationary stochastic processes that is especially well-suited for many NVH problems. It defines a rigorous way to compute probabilities for discrete, nonlinear, nonstationary stochastic processes. The hierarchical Hyperstate formulation provides a computationally efficient, mathematically optimal framework for nonlinear filtering of multiresolution processes. The ability to explicitly model and estimate nonlinear system dynamics on different resolution scales can significantly improve signature discrimination and characterization performance. It can also provide significant insights into correlating mechanical sources and perceived sound quality, as well as providing malfunction diagnostics. Its capability for automatic classification can make it useful for replacing or supplementing human jury evaluations. Hyperstate analysis is a general technique with diverse applications, and it has been successfully applied to the acoustic classification of vehicle starting sequences.

6. REFERENCES

[1] Basseville, M., *et al.*, "Modeling and estimation of multiresolution stochastic processes", *IEEE Trans. Inform. Thy.*, March 1992, Vol. 38, No.2, pp. 766-84.

[2] Juang, B.H. and Rabiner, L.R.," Hidden Markov models for speech recognition", *Technometrics*, August 1991, Vol. 33, No. 3, pp. 251-72.

[3] Lee, C.H., Juang, B.H., Soong, F.K., and Rabiner, L.R., "Word recognition using whole word and subword models", *Proceedings of ICASSP*, 1989, pp. 683-86.

[4] White, J.V., "Modeling and filtering for discretely-valued time series," Chapter 10 in *Bayesian Analysis of Time Series and Dynamic Models*, edited by J.C. Spall, Marcel Dekker, Inc., New York, 1988.