ACTIVE SENSORY TUNING OF WINDNOISE USING A GENETIC ALGORITHM

Andrew Sterian, Paul Runkle, and Gregory H. Wakefield
Department of Electrical Engineering and Computer Science
The University of Michigan, Ann Arbor, Michigan 48109-2122

ABSTRACT

The subjective tuning of multidimensional systems is considered. Listener preference for certain attributes of sound over others involves multiple options, costs, and payoffs, and may require complex strategies to arrive at an optimal solution. When the complexity of such strategies becomes excessive, listeners adopt simpler strategies that may lead to poor, but achievable, solutions. Active sensory tuning (AST) is aimed at providing the listener an efficient search strategy that yields good solutions within reasonable time bounds. Within an engineering context, AST provides a direct link between the human user and the engineer's design parameters whereby the user can tune the design to their desired goal. This contrasts with current, though much simpler, techniques whereby the user can rank their preference but the engineer must interpret such rankings with respect to the selected parameters.

1. INTRODUCTION

The present paper discusses the application of AST to the subjective characterization of the spectral shape of windnoise. A genetic algorithm is proposed as an aid in this description which is otherwise a very difficult task for human users. The genetic algorithm is a randomized search technique requiring no mathematical problem description and is robust to noise. This report proposes a variation of the genetic algorithm that incorporates human feedback in the search through a multidimensional parameter space. Experimental results are presented in which listeners tune the spectral shape of windnoise within a constrained parameter space to their preferred value. Application of this technique to the general problem of tuning multi-parameter digital signal processing systems by humans is discussed.

1.1. Windnoise

Listener preference for certain attributes of automotive sounds, such as windnoise, involves many factors, costs, and payoffs, all of which are difficult to model in order to arrive analytically at a satisfactory windnoise spectrum. Though hard to quantify and equally hard to relate back to the physical design variables, sound quality factors play a considerable role in the customer's judgment of a vehicle. Therefore, research in the automotive industry has focussed on implementing methods by which to measure customer response to various engineering design variables that affect the sound quality of the vehicle, and to relate these measurements back to the design variables.

Experimental measurement of listener preference is hampered by the often large number of design variables that must be considered. A typical factorial experimental design requires an unacceptably large number of measurements for human sensory scaling. It is sometimes possible to reduce the number of measurements through careful selection of what conditions are of greatest importance, but this is not always possible.

Relating listener preference to the physical design variables also can be hampered by the degree to which customer response can be quantitatively represented. For example, stating that the sound is too "boomy", while descriptively accurate, may not convey to the design engineer what physical variables involved in generating the sound must be adjusted to minimize this objection.

1.2. Active Sensory Tuning

Data from the experimental approach outlined above not only suggest which physical variables, or their values, are important, but also which degrade the perceived quality of the sound even further. In this sense, the approach provides the engineer with too much information; only options that improve the quality are typically of interest. Alternatively, the approach provides the engineer with too little information because it is often necessary to infer the best design parameters from the data.

An alternative approach is suggested by the fact that the actual goal of the experiment is to learn much less about the design space, but to acquire this knowledge through a direct mapping of the responses to the physical variables of interest. Such an approach falls within a formalism we call active sensory tuning.

A tuning problem differs from the scaling problem above in that the human user is only required to "make it better" rather than indicate why one option is better or worse than another. Common tuning problems occur in one-dimensional regulators, such as the temperature of bathwater, the loudness of a radio, or the speed of a car. In more complicated tuning problems, such as the fitting of eyeglasses or hearing aids, a prescriptive "best solution" is obtained through measuring the human's response on a battery of tests and inferring from these tests the best choice of the physical parameters. An active form of this sensory tuning replaces

 0-7803-2431-5/95 $4.00

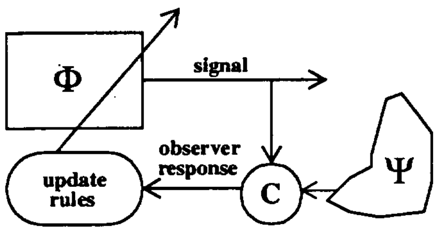

Figure 1. Block diagram of an active sensory tuning system.

the battery of prescriptive tests with an iterative search by which an algorithm uses the human response to manipulate the system parameters to yield a better, if not best, solution.

Accordingly, temperature regulation of bathwater is solved typically by use of active sensory tuning as is perceptual equalization of a stereo system. Of these two tasks, however, the one-dimensional search along the temperature variable is deemed "easy" whereas the multi-dimensional search over all spectral shapes is deemed "difficult", and, indeed, many users of equalizers give up long before they have achieved a desirable fit that matches the music, the room, and the loudspeaker/stereo system.

As shown in Figure 1, we model active sensory tuning as a problem of adaptive signal processing. That is, tuning, like adaptive signal processing, is viewed as an iterative process in which the state of the system Φ is modified over a number of observations according to an error signal C and update rules. In both cases, the update rules are designed to drive the average state of the system to Φ^*

$$\Phi^* = \operatorname{argmin} C(\Phi;\Psi) \tag{1}$$

Unlike standard adaptive signal processing, however, the error signal C is generated by a human user based on some perceptual parameters Ψ. These differences impact significantly on the selection and design of the update rules since, in general, neither C nor Ψ are known analytically.

The type of control (or error) signal generated by the human user in the feedback loop imposes a strong constraint on the design of the update rules, as does the limited number of observations that are possible in any practical tuning system. The former factor is often discussed in the psychometric literature as an issue of numerical scale: in providing numbers as a measure of the merit of a given solution, the mathematics of standard adaptive signal processing requires that this numerical feedback satisfy the properties of either an interval or ratio scale. However, a body of psychometric evidence supports the hypothesis that human observers utilize weaker scales when using numbers to quantify their perceptual response. In the worst case, the observer may use numbers to categorize their response, from which little information beyond set membership of potential solutions can be inferred. In contrast with this categorical scale, ordinal scaling by the human observer provides a strictly ordered number system from which directionality (better/worse) can be inferred, but ratio (twice/half) and absolute differences are not known directly. This fundamental difference in the nature of the "numbers" provided by a human user in quantifying their response to options is one of the primary challenges in designing an active sensory tuning system.

The second challenge is that human users can handle a modest number of iterations in searching for an ideal solution. As opposed to the standard optimization problem encountered in adaptive signal processing, "finite" in this case means on the order of a few hundred trials rather than "less than infinity". From an analytical standpoint, this suggests that the rules should take into account the initialization of the feedback system as well as the update rules. It also effectively imposes a limit on the degree to which analytic tools based on asymptotic analysis can be applied in treating the problem at hand.

We have explored several solutions to the active sensory tuning problem. One class of solutions is based on a modification of the LMS algorithm and has been shown to work well when a deterministic signal is to be extracted from an additive noise. In the presence of other contaminating deterministic signals, however, or for perceptual tuning problems in which the tuning surface is multimodal, this approach has been shown to fail [3]. The present solution is a response to the inherent specificity of the LMS approach and reflects an approach that is much more robust to poorly defined error surfaces and feedback measures.

2. GENETIC ALGORITHM FOR ACTIVE SENSORY TUNING

A genetic algorithm (GA) is an adaptive system that attempts to find good solutions to multidimensional optimization problems. The GA is loosely based on the principles of natural selection and evolution as found in nature. Good solutions to the optimization problem are found by repeatedly forcing a population of candidate solutions to compete for the right to survive over a large number of generations. By making the chance of survival at each generation proportional to a candidate solution's fitness in solving the problem (natural selection), the population increases in overall fitness after some generations pass (evolution). While the GA does not guarantee that the optimal solution will be found, in practice the algorithm yields good results for a wide range of problems [1].

Genetic algorithms have achieved prominence recently as being good alternatives to traditional optimization methods. Generally, genetic algorithms are more robust at handling problems with incomplete mathematical descriptions, noisy evaluation mechanisms, and noisy or time-varying solution spaces. Subjective tuning exhibits all of these properties; human observers are unable to provide mathematical descriptions of their internal listening processes, their evaluation of auditory stimuli with respect to a given criterion is noisy, and they may in fact change their perception of what sounds good and what sounds bad over time. The claims that genetic algorithms make of robustness and utility under these difficult conditions make them worthy of study as aids in subjective tuning.

2.1. Genetic Algorithm Nomenclature

An optimization algorithm seeks to maximize (or minimize) some objective function $F:S \rightarrow R$ which maps elements w of the function domain S onto the real domain. Genetic algorithms operate on a collection of entities, known as *chromosomes*, that encode elements $w \in S$. This collection of chromosomes is known as the *population*. Each chromosome describes an element $w \in S$ using a *bit string* structure. The mapping from elements of S to bit strings is problem-specific but, when elements w consist solely of numeric components, this mapping can be a simple concatenation of the binary representation of each component. For example, if S is Z^3 (where Z represents the domain of 8-bit integers) then any $w = (x, y, z)$ can be represented by the 24-bit bit string comprised of the concatenation of the binary representations for x, y, and z. When a component of w can be identified with a string of bits in the bit string, then this string of bits is called a *gene*.

2.2. Genetic Algorithm Mechanics

The genetic algorithm works as follows. At the beginning of the algorithm, the first population is initialized with random elements, w_i for $i = 1 \ldots N$ where N is the size of the population. Then, members of the population are selectively chosen (based on their fitness $F(w)$) to provide genetic material (i.e., to act as *parents*) for a new set of points in S (i.e., the *offspring*). Genetic material consists simply of bits from a member's bit string representation. This new generation of points replaces the current population and the process repeats.

Since offspring are produced from parents which are judged to be more fit than the remainder of the population, the offspring will likely be of high relative fitness since they inherit the genetic makeup of their parents. In successive generations, the fitness of the population should steadily increase.

The two primary operations which transfer genetic material from parents to offspring are called *crossover* and *mutation*. Crossover uses two parent members to produce one (or more) offspring by combining the genetic material of the parents. In the simplest form of the operator, a cut point is chosen at a random position along the length of the two parent bit strings. The offspring is formed by taking the bits from the beginning of the string to the cut point from the first parent, and completing the string with bits from the cut point to the end of the string from the second parent. Two offspring are easily formed by swapping the roles of the parents. The mutation operator simply changes random bits in the bit string of a population member.

In keeping with the evolutionary paradigm, crossover is compared to sexual reproduction, since two parents combine genetic material to form an offspring different from the parents, and mutation is compared to asexual reproduction, since only one parent is involved and the only possibility of having an offspring different from the parent is by random genetic mutation.

There are many implementation variations of genetic algorithms but all embody the same principles of randomized search using crossover and mutation operators (and perhaps others as well including problem-specific operators [1]).

The success of genetic algorithms has been attributed to their combination of rapid searching of the function space (exploration) coupled with their use of memory to find new, better solutions from previous solutions (exploitation). Goldberg [2] proposed that the GA combines short, highly fit building blocks (small substrings within a chromosome) into larger blocks of even higher fitness. The term *implicit parallelism* is used [1] to describe how the GA simultaneously searches multiple regions of the function space. Even if the reason for the success of GA's is not completely understood, there is ample empirical evidence that supports the use of the GA in optimization problems.

2.3. Equalizer Model

Before discussing the operation of the genetic algorithm in this study, we first describe the implementation of the equalizer model and the mapping of chromosomes to equalizer gains. The design of the equalizer was driven by the desire to span the range of turbulence known to strongly influence subjective preference and constrained by the availability of real-time computing power. Also, the length of the chromosomes in the genetic algorithm is dependent upon how many bands are chosen in the equalizer as well as the gain resolution of each band. The desire to keep the system dimensionality low was also a driving factor in the design.

The equalizer implementation consists of 6 frequency bands with center frequencies chosen to be linearly spaced below 1000 Hz and logarithmically spaced above 1000 Hz (100-400, 400-700, 700-1000, 1000-2000, 2000-4000, 4000-8000 Hz). This frequency assignment mirrors the width of critical bands of the human auditory system which are of fairly constant width below 1000 Hz and grow logarithmically above 1000 Hz. Each frequency band was implemented with an IIR filter designed to have 0.1 dB ripple in the pass band and at least 60 dB attenuation in the stop band.

The genetic algorithm operates on bit strings that represent an encoding of the equalizer gains. Each gain is represented by four consecutive bits in the bit string allowing for 16 possible gain values. Each increment in this 4-bit value represents a 2 dB change in the equalizer gain. We identify each 4-bit string in a chromosome as a gene. There are 6 genes in each chromosome (corresponding to the 6 frequency bands) for a total bit string length of 24. The leftmost gene in the chromosome represents frequency band 1 and the rightmost gene represents frequency band 6.

3. EXPERIMENT AND RESULTS

The search space was restricted to lie within the class of all valid windnoise spectra by introducing a monotonicity constraint on the chromosome mapping. The mapping of genes to filter gains was preceded by a step in which the gene values were implicitly sorted in non-increasing order. Thus, the lowest frequency band was always assigned the highest gain and the highest frequency band, the lowest gain. Crossover and mutation probabilities were chosen based on optimizing the expected number of iterations to convergence according to Monte Carlo simulations of the experiment. Subjects performed a paired-comparison task to rank-order the chromosomes based on preference. Fitness values were assigned to chromosomes based upon a linear mapping of the rank.

The three authors participated as subjects in the tuning task. All had hearing thresholds that were within the norms established for the ASPEN laboratory at the University of Michigan. At the beginning of each tuning session, twelve chromosomes were drawn at random and white noise signals were filtered by the equalizer set to each of these chromosomes through the monotonicity-mapping described above. The subject's task was to listen to a pair of 1-second long filtered noises and indicate which of the two best exemplified an "ideal" windnoise signal. The search was terminated when the population of chromosomes contained fairly homogeneous genetic material, according to a preset criterion. In all, each subject completed a minimum of three sessions.

The table below presents the best solution as tuned by each subject for each run. In general, up to 700 trials and 39 generations were necessary to obtain convergence, but the average number of trials was more on the order of 325 for 20 generations. S1 and S3 exhibit greater preference for windnoise dominated by little high-frequency energy whereas S2 exhibits a more gentle rolloff in the preferred windnoise spectrum. The individual differences noted are similar to those observed in other experiments in which windnoise preference has been measured using much more intensive data collection. A key result is that these differences agree with these other measures and require far less time to measure using the current technique.

Table 1: Relative gains [dB] of equalizer bandpass filters [kHz] for windnoise preference

Sub-Run	0.1-0.4	0.4-0.7	0.7-1	1-2	2-4	4-8
S1-1	28	2	2	0	0	0
S1-2	30	12	8	6	0	0
S1-3	30	4	0	0	0	0
S2-1	30	18	14	10	4	2
S2-2	30	10	8	6	4	0
S2-3	30	10	8	2	2	0
S3-1	30	30	4	2	0	0
S3-2	30	20	8	2	0	0
S3-3	30	6	6	2	0	0

4. CONCLUSIONS

The incorporation of human response in a standard psychophysical paradigm joined with the genetic algorithm has been shown to be an effective tool in the subjective tuning of a multidimensional system. Future work includes the study of other optimization methods, including hybrid algorithms, and determining feasible bounds on system dimensionality.

[Work supported by grants from the Ford Motor Company and the National Institutes of Health and an NSERC Scholarship to the second author.]

5. REFERENCES

[1] Beasley, D., Bull, D.R., Martin, R.R. (1993). An Overview of Genetic Algorithms: Parts 1 and 2, *University Computing* **15** (2,4).

[2] Goldberg, D.E. (1989). "Genetic Algorithms in Search, Optimization, and Machine Learning," Addison-Wesley, Reading MA.

[3] Runkle, P.R. (1992). "Active sensory tuning using the Pseudo-Signed LMS algorithm," unpublished Directed Study Report, EECS:Systems, The University of Michigan.