

# Psychoacoustic Aids for the Musician's Exploration of New Material

by David L. Wessel and Bennett Smith

Initially published in: *Proceedings of the ICMC 1977*, pp. 190-192;  
<http://hdl.handle.net/2027/spo.bbp2372.1977.015>

Original PDF with new page layout.

## Comment by Miller Puckette

In this paper Wessel and Smith make an early attempt to treat acoustic timbre as a musical quantity on a par with pitch and duration. To do this they propose to lay a coordinate system on the world of musical

sounds so that relationships between them can be developed as part of the work of composing a piece of music. This is especially important to electronic musicians since, for the first time historically, we have had the power to manipulate sounds directly as part of the composition process, and so we have naturally sought ways to "organize sound".

## Post scriptum by Bennett Keith Smith

Working with David introduced me to multidimensional scaling and the world of psychoacoustics and I'm still exploring that rich world. Thank you, David, for your vision.

\*\*\*\*\*

At IRCAM we have developed facilities that enable the musician to explore and characterize the perceptual structure inherent in a body of sound material. The primary goal of these facilities is to provide the musician with essential psychoacoustic information about the perceptual behavior of a selected body of material.

In a typical application of the programs the composer selects the sonic objects, selects from a set of programs that run different kinds of psychoacoustical experiments, and gets in return for his patient listening and carefully considered judgments various representations of the underlying perceptual structure of the material. These representations may suggest ways for the composer to link the elements of the material which are not immediately obvious from only an acoustical characterization and/or casual listening. The representations can in turn serve as input data structures for compositional procedures that facilitate the further study of the material in a variety of controlled musical contexts.

In this talk we shall describe examples of such programs, the theoretical ideas behind them, and the actual application of them to problems of composition.

In designing the programs we had the following criteria in mind:

1. A musician with limited knowledge of the experimental procedures used in psychoacoustics should be able to design and implement a psychoacoustical experiment relevant to a particular musical problem.

2. The presentation of sounds and the recording of the subjective judgments in the experiment should be completely automated. The experiments should be fast-paced with a minimum of tedium.

3. Following the listening and judgment session the subjective data should be analyzed immediately and a variety of intuitively graspable representations of the results should be accessible at the musician's interface to the computer.

4. With the representations of the results in view the musician should be able to hear the sounds under study so as to allow for aural consideration, interpretation, and validation of the results.

5. The representations of the perceptual structure in a set of material should also be linked to the sound synthesis programs so as to permit rapid specification of a compositional idea in terms of the perceptual structure obtained from the experiment. The idea is to enable further study of the material in controlled musical contexts.

#### CONTRASTS:

We will now describe a collection of programs called CONTRASTS that run and analyze exploratory experiments involving judgments of perceptual contrast between two sounds. CONTRASTS can be used to run the now familiar multidimensional scaling and clustering experiments that John Grey and David Wessel and their students have applied to the study of timbre.

In a typical application, a set of from 10 to 30 sounds is selected and all the possible pairs of the sounds are formed. These pairs are sequenced in a random order and each pair is presented to the listener upon demand and remains available for repeated listening until a judgment of perceptual contrast is decided upon. This judgment is entered by the listener and a new pair is made available. After completion of the all the pairs a matrix of contrast judgments is formed and submitted automatically to a variety of data analysis programs. The data analysis programs are of two basic types, those like KYST and INDSICAL that provide spatial representations and those like HICLUST and ADCLUS that provide "feature-based" representations.

Multidimensional scaling programs provide spatial configurations in which the sounds are represented as points. The scaling programs seek a good statistical match between the distances in the spatial configuration and the magnitudes of the corresponding contrast judgments. These geometric representations have considerable intuitive appeal in that they are readily interpretable in terms of regions, directions, and dimensions.

#### ----- Slide -----

Clustering programs provide cluster and hierarchical tree representations of the stimuli. Such representations are based upon the notion that each stimulus consists of a set of "features" and that the perceptual contrast is a function of the number of features two stimuli have in common as well as the number of features that differ between them. These feature based models are able to account of some aspects of the contrast judgments that the geometric models cannot, like asymmetry in contrast, the behavior of prototypical stimuli, and certain kinds of context effects.

#### ----- Slide -----

There are significant differences between the spatial and "feature" approaches and for some problems the two approaches compliment each other quite nicely. When using the geometric approach one usually seeks a representation in terms of a few well behaved quantitative dimensions. For example, in the timbre space work a "spectral-pitch" or "brightness" dimension associated with the center or average of the spectral energy distribution has been consistently obtained. And

#### ----- Slide and Tape -----

indeed, direct manipulation of this property has proven to be compositionally potent. However, other properties of sounds or patterns of sounds might be better characterized in terms of many qualitative features

#### ----- Slide and Tape -----

rather than in terms of a few quantitative dimensions. Examples of such features include the presence or absence of noise in the initial attack transient and the presence

or absence of beats or vibrato. In our work we have found it informative to use both approaches in combination. We typically superimpose the clustering solutions upon the spatial configurations.

#### ----- Slide -----

At this point in the talk it seems appropriate to give a kind of running description of CONTRASTS at work.

#### ----- Coordinated Audio and Slide Sequence -----

---- An example of the commentary that might go with the slides is given below.

One is seated before a terminal, in this case a Datamedia that communicates with IRCAM's PDP-10. First a body of material is selected. To accomplish this an all digital recording and editing facility is available for the "live" sounds as well as most of the well known sound synthesis and processing programs. In collecting or synthesizing the material for study care must be taken to normalize the sounds with respect to the dimensions or properties that are not to be involved in the investigation. In the case of timbre, normalization is accomplished by equalizing the subjective duration, pitch, and loudness. For a given set of material one then produces between 20 and 30 examples each stored as a separate sound file on the PDP-10's disk. A list is then supplied as input to a program that runs the contrast judgment session. But before actually making the judgments it is worth while exploring the sounds with a program called KEYS.

KEYS takes the sound file list and displays it at the terminal, associating a letter of the alphabet with each sound. The listener need only press the right key to hear the desired sound.

After familiarizing oneself with the general character of the material the rather long but really not so bad judgment session is run. The number of pairs that are to be judged is  $n(n-1)/2$ , where  $n$  is the number of items in the set. So for 25 sounds 300 judgments are required. With sounds of less than a half second in duration and a little determination the 300 judgments can be completed in under 45 minutes.

Upon completion of the judgments the matrix of contrast judgments like the one on the slide I'll show is shipped off to the various programs for analysis. And within less than two minutes the solutions from the analysis programs are available for display on the CRT. ---like the one on the next slide--- Hard copy is available either from a Versatec or a line printer. With the graphic and cluster representations in front of you, you can now run the KEYS program again and explore and aurally validate or unvalidate the representations. ---tape demo with the slides to show keys in action--- The same alphabetic characters are used in all the representations and in KEYS as well, so one has little difficulty keeping track of what sound goes with a point or element in the representations.

At this point if you are sufficiently satisfied with the perceptual potency of the representations you can proceed to use them as data structures for compositional algorithms. With a feature recently incorporated into the MUSIC program by Leland Smith one can read in a number of different files and modify them with signal processing procedures that use both the optimized unit generators like envelope generators and filters and ALGOL like language that operates within the MUSIC program. Standard manipulations are those that change the pitch and amplitude envelope. In the examples a number of different timbral trajectories are examined. In the first two the same melodic and rhythmic material is used. In the first a nearly continuous path is traveled. In the second, very large timbral intervals are used between the adjacent notes. ---play examples and point to the trajectories shown on the slide---

In the next set of examples a direction in the timbre space is correlated in various ways with the pitch structure. In the first example the "spectral-pitch" axis and musical pitches are positively correlated, that is, as the pitch gets higher the timbre gets brighter in the next example a negative correlation is used so that as the pitch increases the "brightness" or "spectral-pitch" decreases.

In the next set of examples the "feature" characterizations are used as a basis for the manipulations. Following this will be a series of examples showing the various effects, like auditory stream segregation, when the timbral interval is manipulated. Examples of transpositions will also be given. The material used in the examples will consist of the complete set of Grey-Gordon instrumental timbres and interpolations among them, vocal sounds, inharmonic sounds, and wind instrument multiphonics.

The examples will be selected to demonstrate how the representations aid in characterization of the material in musical contexts.