USING PSYCHOACOUSTICAL MODELS FOR INFORMATION SONIFICATION

Sam Ferguson[†], *Densil Cabrera*[†], *Kirsty Beilharz*[*] *and Hong-Jun Song*[*]

[†]Acoustics Research Laboratory
[*]Key Centre of Design Computing and Cognition
Faculty of Architecture
The University of Sydney
`sferguson@arch.usyd.edu.au`

ABSTRACT

Psychoacoustical models provide algorithmic methods of estimating the perceptual sensation that will be caused by a given sound stimulus. Four primary psychoacoustical models are most often used: 'loudness', 'sharpness', 'roughness', and 'fluctuation strength', models for which have been presented by Zwicker and Fastl [1]. These four models have been used extensively for optimising product sound quality in industrial sound design applications. However, they also may be applied for auditory display purposes. This paper presents a method for their application and discusses effects and implications of using this method for designing auditory displays. This paper is primarily theoretical – however, sound examples of auditory graphing based on psychoacoustical models will be presented at the conference for discussion.

1. INTRODUCTION

The field of psychoacoustics aims to model parameters of auditory sensation in terms of physical signal parameters. Most famously, Zwicker and Fastl's monograph, 'Psychoacoustics: Facts and Models' [1], presents a suite of algorithms for calculating auditory sensations including loudness, sharpness, roughness, and fluctuation strength. These algorithms are implemented in 'sound quality' software, which is often used in industrial acoustics design (e.g. in optimizing the sound quality of appliances). Zwicker and Fastl give examples of how these algorithms offer a much more powerful understanding of auditory sensation than purely physical signal measurements. The application of psychoacoustical models extends to other fields – and can be used in music analysis [2, 3], voice quality analysis [4], audio system modeling [5], audio data rate compression [6], room acoustics, and indeed to auditory display. Within auditory display, psychoacoustical modeling could be used in a similar fashion to the sound quality field with model results being used to refine candidate sounds for a display under development, or to attempt to control for unintended perceptual effects created while using simple display parameters. Used in this manner, psychoacoustical modeling should be of substantial benefit in developing displays that are easy to interpret and pleasant to listen to. More ambitiously, psychoacoustical modeling could be applied directly in the generation of multidimensional auditory graphs. If the psychoacoustical models are truly reliable, then this approach to auditory graphing should succeed, notwithstanding cognitive limitations. The fact that this approach is not widely adopted can be attributed to several factors, most notably the difficulty in implementing psychoacoustical modeling and synthesis.

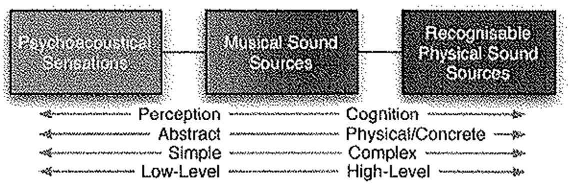


Figure 1: *Approaches to sonification can be conceptualised along some general axes.*

The idea of a 'perceptual sound space' for auditory display, as presented by Barrass, could be viewed as a precursor to the present project [7]. A pitch/timbre space derived from a combination of direct subjective testing and pre-existing psychoacoustical models was developed by Barrass, including the timbral dimension of sharpness. In a recent reflection on this work Barrass notes that perceptual scaling (at least in this sense) is rare in this field [8]. However, his approach has been influential as a foundation for the use of a comprehensive framework for aligning perceptual parameters to data attributes.

Neuhoff and Heller have investigated alternative methods of auditory graphing [9]. They favour the use of more complex acoustic features that map unambiguously to sound source characteristics, such as the use of footstep sounds to represent a Cartesian plane (eg. footstep speed and surface liquidity or solidity as the x and y axes). They suggest this due to the fact that the more common use of lower-level auditory dimensions (pitch, loudness and timbre) interact with each other to such a level that they often may confuse the user of the display. They argue that this technique allows the listener to develop robust mental models they can then associate with the target dimension.

Although the basis for their technique is clearly strong, it seems worthwhile to attempt to measure interaction directly and control it. However, to do this requires the assumption that lower-level auditory phenomena can be classified, generalised and modeled using the techniques that Zwicker and Fastl outline. Whilst the accuracy of their models is open to debate, and improvements to these models continue to be made, their thrust is to indeed classify and model (to a specified level of generality) the lower levels of auditory sensation. Therefore they may offer a solution to some of the interaction problems Neuhoff and Heller outline. Figure 1 illustrates some aspects of the contrast between their approach (on

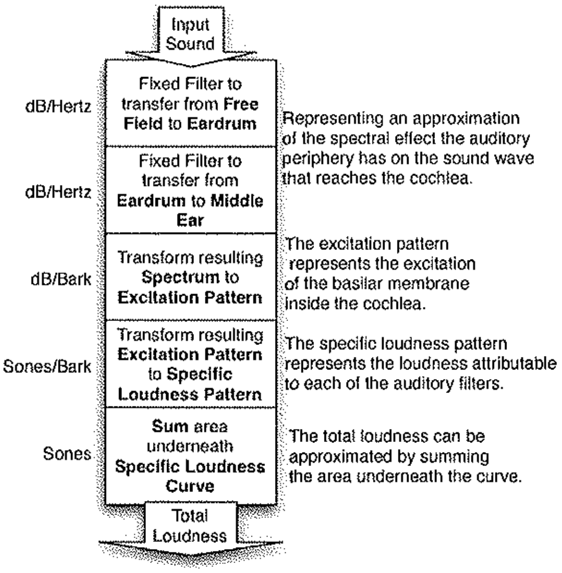


Figure 2: *The analysis stages in a typical loudness model.*

the right) and the approach of the present paper (left). Music-like sources, which are often used for auditory graphing, are between these extremes.

This paper is not about a particular auditory display, but examines the issues around the application of psychoacoustics to auditory graphing. It develops this idea through an overview of psychoacoustical modeling, and by way of example. The term 'psychoacoustics' is often used in the context of auditory display, with a wide range of meanings, including many aspects of auditory perception and cognition. The present paper is concerned with the part of psychoacoustics that develops and employs quantitative models of auditory perception. The scope of this paper is limited and is concerned primarily with an outline of the various psychoacoustical models that could be employed, a description of a method for applying of psychoacoustical models to auditory display, and a discussion of some of the issues and questions associated with this style of sonification.

2. PSYCHOACOUSTICAL MODELS

In this section we briefly review the psychoacoustical models of loudness, sharpness, roughness, fluctuation strength and pitch.

2.1. Loudness

The subjective impression of the 'loudness' of a sound has been extensively modelled by many researchers [10, 11, 12, 13, 14]. The models draw on data gained from subjective testing and from a physiological understanding of the auditory periphery. The unit of loudness, the *sone*, is a ratio scale referenced against the sensation produced by a 1 kHz sine tone with a sound pressure level of 40 dB.

Loudness is not generally regarded as a robust auditory scale for data representation. Reasons for this include: a) the complexity of loudness modelling (meaning that auditory display designers are rarely able to model it with any accuracy), b) the limited loudness range available for many auditory display contexts (especially if the display is to be reliably audible in a potentially noisy environment, yet of low annoyance), c) lack of ability to predict and calibrate the sound pressure level of a display in many contexts, d) experimental findings that loudness judgements are less reliable, or at least less sensitive, than other potential display parameters (eg pitch or pulse rate) [15], e) and that some even dispute the general validity of loudness modelling (for instance Warren [16]).

However, while it is likely that loudness is not an easily employed display parameter, it would be unwise to dismiss it entirely. Indeed, loudness is perhaps the most fundamental of auditory sensations and it is likely that existing auditory displays already manipulate loudness, even if they seek to maintain a constant acoustic output. This is due to loudness's inherent dependency on frequency content. Thus this manipulation of loudness may be a confound in the intended simplicity of a display.

The loudness model process (outlined in Figure 2) involves several stages, which often relate to the way in which the ear processes sound to stimulate the auditory nerve. The output of auditory models is used primarily in loudness modelling, but also in the other psychoacoustical models outlined below.

In terms of its relation to signal parameters, loudness is affected by the predominant frequency of a signal (eg. of a pure tone, tone component or noise band), the signal bandwidth, signal duration (in the case of brief signals, of less than 200 ms), amongst others. Arbitrary signals may be difficult to parameterise in these terms, and so are good candidates for full loudness modeling. In some circumstances, doubling or halving of loudness corresponds to a 10 dB gain interval – however, this rule of thumb does not apply to low frequency or quiet sounds. This complexity means that, while an objective correlate of loudness can be roughly estimated through weighted sound pressure level measurements, instances can occur where increased sound pressure level is accompanied by decreased loudness. Hence, signal gain may be used to roughly control loudness, but a loudness model is required if loudness is to be well controlled for a variety of signals.

A key concept in psychoacoustics is the specific loudness pattern, which can be crudely thought of as a psychoacoustical magnitude spectrum. More precisely defined, it is a representation of the amount of loudness attributable to auditory filters, which have characteristic frequencies from low to high, and is the penultimate stage in a loudness model. The auditory filter scale units are Barks (used in the Munich school of psychoacoustics, including by Zwicker and Fastl) or ERBs (used in the Cambridge school). The specific loudness pattern, especially in its time-varying form, is the basis for full models of loudness, sharpness and roughness, fluctuation strength and potentially other psychoacoustical parameters.

2.2. Sharpness

Sharpness (or brightness) is one of the most prominent features of timbre [17]. Models are based on the centroid, whether it be of the signal spectrum [18], or the specific loudness pattern [19, 1, 20]. Following Zwicker and Fastl, sharpness is modeled as a weighted centroid of the specific loudness pattern. Its unit is the *acum*, referenced to a band of noise 1 critical band wide, centered on 1 kHz at 60 dB. It is not independent of signal gain, because of increased auditory filter asymmetry for high sound pressure levels.

Sharpness is a musical term often used to describe a sound with a pitch slightly higher than a target pitch. However, in the present context the meaning is quite different, refering to the perception that the sound is 'sharp', 'harsh' or 'has an edge'. This is strongly related to the proportion of high frequency energy present in the sound, weighted towards energy in the region above 3 kHz.

Again, the dependency upon frequency content will have the implication that most displays will use signals of varying sharpness in a fairly uncontrolled way. For harmonic tones, sharpness can be roughly controlled through the distribution of the harmonic spectral envelope, although a sharpness model is required for scaling.

The model used by Zwicker and Fastl for calculating the sharpness of tones is summarised by Equation 1, where S is sharpness, N' is specific loudness, z is the bark scale of auditory filters (also know as critical band rate) and $g(z)$ is a weighting function that emphasises z for high critical band rates.

$$S = 0.11 \frac{\int_0^{24Bark} N' g(z) z \, dz}{\int_0^{24Bark} N' dz} \, acum. \tag{1}$$

2.3. Roughness

Roughness is a sensation caused by quite rapid amplitude modulation within auditory filters. This modulation can be caused by beats between two pure tone components, or by a signal with amplitude or frequency modulation. Beating within an auditory filter channel has been used to explain the acoustic component of tonal dissonance [21], and in this sense dissonance can be thought of as one form of roughness. The roughness of frequency-modulated tones is explained by the resulting amplitude modulation of auditory filters of fixed pass-band frequency. According to Zwicker and Fastl, maximum roughness is often achieved for modulation frequencies around 70 Hz (depending on the carrier frequency). Faster signal modulation rates have reduced modulation depth in auditory filters due to the ear's temporal integration. They present a set of simple models of roughness associated with a single known modulation frequency and carrier signal type. The unit of roughness is the *asper*, which is referenced to a 1 kHz tone at 60 dB with 100% amplitude modulation at 70 Hz.

The model presented by Zwicker and Fastl for calculating the roughness of modulated tones having a single modulation frequency is given in Equation 2, where R is roughness, f_{mod} is the modulation frequency, and $\triangle L_E$ is the excitation level within an audiitory filter. This uses the time-varying excitation pattern of the ear (similar to the specific loudness pattern, except that the magnitude is in decibels rather than sones/bark), with the difference between maximum and minimum excitation levels integrated across auditory filters used to determine roughness.

$$R = 0.3 \frac{f_{mod}}{kHz} \int_0^{24Bark} \frac{\triangle L_E(z) dz}{dB/Bark} \, asper. \tag{2}$$

Zwicker and Fastl note that it is advantageous to derive roughness from the specific loudness pattern (rather than the excitation pattern) and to take the correlation between the temporal envelopes across the auditory filter range into account. Such general models of roughness (eg Aures [22], Daniel & Weber's [23]) could be unnecessary for the the envisaged style of auditory graphing if the signals fit the criteria for Zwicker and Fastl's simpler model for tonal carriers and a single modulation frequency.

The roughness of auditory stimuli can be manipulated through control of modulation rate, modulation depth, and modulation type (frequency or amplitude modulation), as well as more complex influences such as loudness and carrier spectral content.

2.4. Fluctuation Strength

Fluctuation strength describes a sensation caused by relatively slow amplitude modulation within auditory filters with maximum sensitivity at around 4 Hz. It is easy to appreciate that the perception of the degree to which a sound is fluctuating should increase with the modulation frequency, but as the frequency increases from 4 Hz to 20 or 30 Hz the fluctuation becomes increasingly harder to track mentally, eventually merging into a constant (albeit rough) sound. As observed by Zwicker and Fastl, fluctuation strength appears to be related to speech, since its amplitude modulation spectrum also tends to peak in the vicinity of 4 Hz.

Zwicker and Fastl present a simple model of fluctuation strength, suitable for modulated tones, and one suitable for amplitude modulated broadband noise, both only accounting for a single modulation frequency. The unit of fluctuation strength is the *vacil*, referenced to a 60 dB 1 kHz pure tone 100% amplitude modulated at 4 Hz.

The model used by Zwicker and Fastl for calculating the fluctuation strength of tones is summarised by Equation 3. Like the simple roughness model, this fluctuation strength model is based on the fluctuation of the excitation pattern of the ear.

$$F = \frac{0.008 \int_0^{24Bark} (\triangle L/dB \; Bark) dz}{(f_{mod}/4Hz) + (4Hz/f_{mod})} \, vacil. \tag{3}$$

Zwicker and Fastl also refer to a general model of fluctuation strength, structured similarly to the general roughness model (based on the time-varying specific loudness pattern).

Fluctuation strength may provide some insight into pulse rate as a parameter for auditory alerts. Pulse rate is often used to indicate levels of urgency in auditory alert systems [24].

The fluctuation strength of auditory stimuli can be manipulated through control of modulation rate, modulation depth, and modulation type (frequency or amplitude modulation), as well as more complex influences such as loudness and carrier spectral content.

2.5. Pitch

A psychoacoustical pitch ratio scale is an elusive concept because of the complexity of pitch perception and cognition. Stevens and others developed two versions of the mel scale, based on magnitude judgments by listeners without musical training [25, 26]. However, Greenwood finds methodological flaws in the mel scale's derivation, and instead that frequency ratios form a better pitch scale [27]. Shepard describes some of the complexity of pitch structures for harmonic tones (such as pitch height, octave equivalence and the cycle of fifths) through multidimensional geometric figures [28]. In addition to complex structures of pitch height, pitch has the dimension of pitch strength, also known as 'tonalness' [29, 2].

The harmonic series is of great importance in pitch perception, and many pitched sounds in everyday experience exhibit harmonic spectra most notably the human voice. Even though many frequencies are present in a harmonic tone, usually the fundamental frequency dominates as the pitch percept. When the fundamental frequency is masked or completely absent, the fundamental as a pitch percept tends to remain. Models of pitch perception account for this phenomenon either explicitly through template matching [29] or implicitly through autocorrelation [30]. Parncutt has explained many aspects of musical harmony through this psychoacoustical process [2].

In auditory display applications such as auditory graphing, various approaches can be made to pitch encoding, including the frequency ratio scale (also known as logarithmic frequency distribution, exemplified by the musical chromatic scale), the diatonic major scale (which has the advantage of musical familiarity), the pentatonic scale (which has the advantage of relatively low dissonance for any tone combinations), and the linear frequency scale (which has the advantage of expressing low order integer ratios as familiar harmonic intervals). Unlike the other psychoacoustical scales referred to in this paper, such pitch scales are not necessarily psychological ratio scales, yet they can provide a range and level of precision in data representation unavailable for the other scales.

3. UBIQUITY OF PSYCHOACOUSTICAL SENSATION

Whilst these models can be complex to understand initially, the sensations they attempt to model are basic perceptions created to some degree by all sounds. This ubiquity may be a cause of some of the difficulties inherent in the usage of auditory displays. Whilst it is often assumed that a simple mapping to an aspect of sound such as fundamental frequency is relatively transparent, when we consider other aspects of sound such as sharpness or loudness, it becomes possible to see that confusion may be being caused by the implicit presentation of other sensations. In this situation it is up to the listener to develop the skills to respond to the perceptual parameter the sound designer intended to represent the data with, despite the concurrent unintended presentation of other perceptual parameters.

Figure 3 gives a visual explanation of three ways of dealing with this problem. In the first instance, we normalise other psychoacoustic sensations to control this arbitrary variance, in an attempt to reduce the confusion caused. In the second instance we manipulate this variance directly in an attempt to reinforce the representation parameter with psychoacoustic sensations that move in parallel. In the third we use the psychoacoustic sensations to represent separate data streams.

3.1. Sound Quality

The use of psychoacoustical models facilitates the prediction of subjective qualitative response to auditory stimuli. For example, Zwicker and Fastl present models of sensory pleasantness, based on a combination of psychoacoustical parameters (but dominated by sharpness). Such models are context-dependent, but could be developed for an auditory display application. Many studies in the field of sound quality indicate that psychoacoustical scales can better predict human evaluation of sound than physical signal measurements [1].

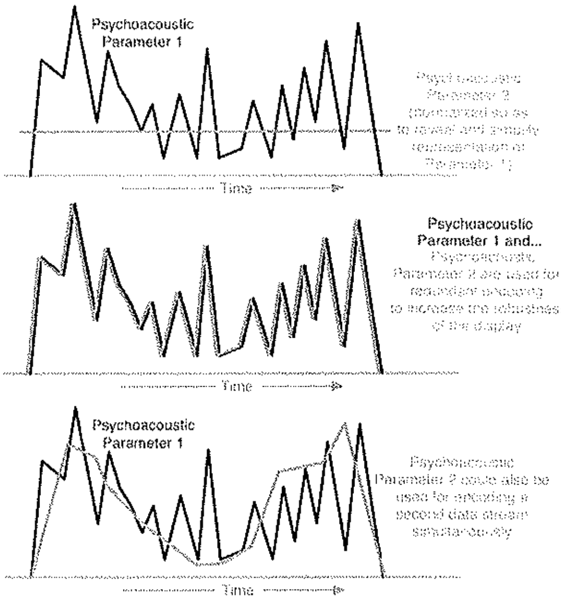


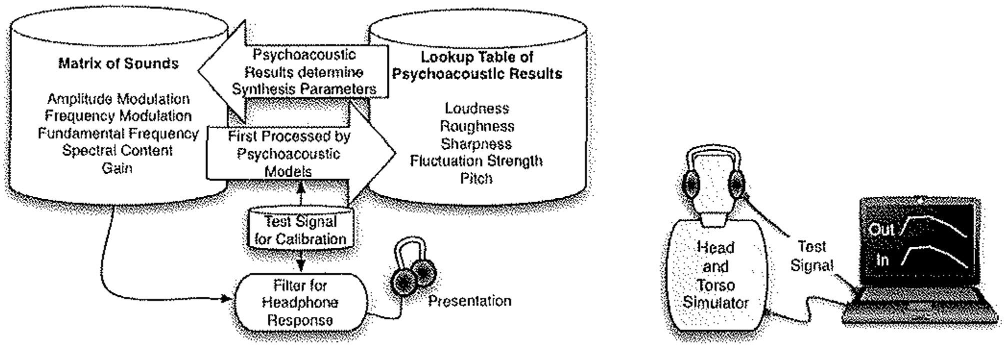
Figure 3: *Strategies for dealing with the existence of confounding psychoacoustical parameters in auditory displays.*

4. A METHOD FOR IMPLEMENTATION

These psychoacoustical models rate the quality of sounds using scales that represent particular aspects of human auditory perception. The algorithms involve several stages and can present identical results for different sound inputs. Therefore, they cannot be applied in reverse for synthesis, so the use of these models for auditory display can present a practical challenge.

Nevertheless, a method for implementing auditory displays need not approach synthesis by employing psychoacoustical models directly. A less direct approach is to generate a large matrix of sounds using traditional digital synthesis techniques, process these using psychoacoustical algorithms, and thereby generate a lookup table we can use to determine the appropriate digital synthesis parameters for any given set of psychoacoustical parameters (Figure 4(a)). Whilst this may seem a relatively trivial process, the implementation requires great care in preserving the orthogonality of the various psychoacoustic measures.

To generate a set of candidates for this analysis process simply requires a number of sounds that vary sufficiently for each psychoacoustic parameter to be in turn varied. This requires an understanding of to what, and over what range, the models are sensitive, as well as an understanding of the unintended effects manipulation of other parameters may induce. For instance, excessive frequency modulation may reduce the perception of tonalness in the produced sound. The manipulation of amplitude modulation, frequency modulation, fundamental frequency, spectral slope and gain allows us a range of control over the psychoacoustic parameters of loudness, sharpness, roughness, fluctuation strength and pitch perception. Table 1 describes possible synthesis parameters and their limits in more detail. Of course these parameters will



(a) *The method for implementation is to synthesise a large number of sounds that are systematically varied, analyse them using psychoacoustic models, and then resynthesise the appropriate sound needed to produce particular psychoacoustic sensations.*

(b) *Using a head and torso simulator is one method for developing the headphone filter.*

Figure 4: *A Method for Implementation*

not control the psychoacoustic parameters directly, they will only be used to vary these parameters significantly enough that a useful range of psychoacoustic results are produced.

There are other aspects of synthesis that need to be controlled for this scheme to avoid unintended auditory interaction. The use of amplitude decay envelopes, while popular for creating realistic sounds, also results in loudness (and indeed other) variance within each representative sound. For this particular application this is probably not appropriate, or at least it complicates matters to a degree not necessary at first. Therefore in our initial implementation the tones do not vary in amplitude over the length of time they sound.

Once the matrix has been devised and the samples analysed, the resulting psychoacoustic ratings can be used for synthesis. It is a straightforward process of scaling the data input to match the units of the particular model being employed, and then using the look up table to determine the correct set of synthesis parameters to use. The resulting tones, requiring a value for each of the five psychoacoustical parameters, are synthesised and then concatenated using a short cross-fade to avoid startling changes. The tone length should preferably not be shorter than the integration time of the ear ($\geqslant$ 125 ms). The output sound consists of concatenated complex tones incorporating variable fundamental frequency, gain, spectral slope, amplitude modulation and frequency modulation. The resulting auditory displays sound comparatively simple, but serve to demonstrate the potential of this technique.

A representation of how parameters for synthesis affect the psychoacoustical model results in the implemented matrix is presented in Figure 5 for a three dimensional model using pitch, loudness and sharpness.

4.1. Practicalities of Presentation

The models this auditory display method depend on are all dependent upon presentation level *and* spectral content. The linearity in terms of level and spectrum of the presentation method used is crucial in avoiding biasing the perception of any of these psychoacoustic parameters. Practically speaking this precludes the use of loudspeakers except in very controlled conditions: their frequency response and level at the listener position is heavily dependent on the acoustic response of the room and on the distance and location of the listening position.

The use of headphones with a known and preferably reasonably linear frequency response will provide a degree of insensitivity to listener movement. However, it is still necessary to calibrate the headphone output level with a known source. Due to the variability of spectral response in different headphone models it is probably better to use a broad-band noise source rather than the more common mid-frequency sine tone. By analysing and filtering recordings of this signal played through headphones placed on a Head and Torso Simulator we can be assured of a reasonable correspondence between the digital signals being analysed by the psychoacoustic models and the filtered sounds being received at the ears by the listener (Figure 4(b)).

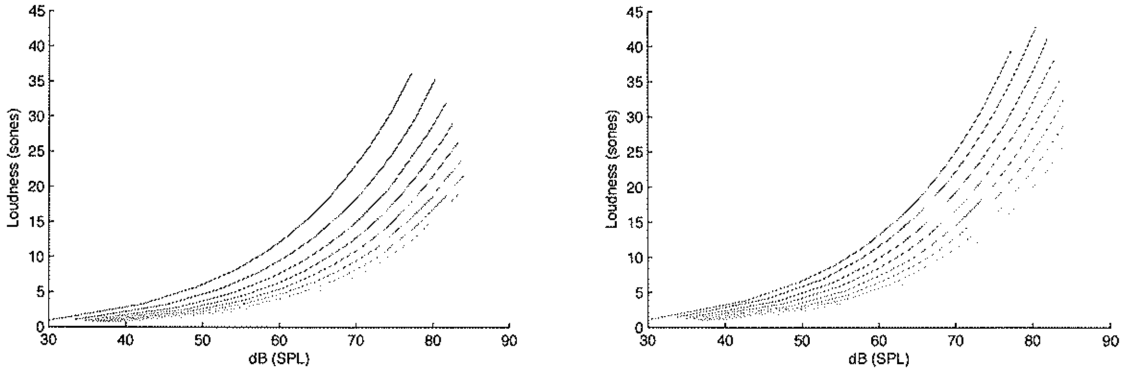
This requirement may seem extreme, and Flowers argues that for auditory graphing to become widespread it is best to use techniques that do not degrade when presented using audio systems of possibly lower quality and less predictability [15]. However, the preponderance of low quality presentation media does not limit Tufte from arguing strongly for the use of high resolution display methods in the visual domain [31]. Tufte simply dismisses the computer screen as a display method and demands high resolution printouts as the format necessary for graphical excellence. Whilst the abandonment of computer based graphical display methods is probably unlikely to occur in the near future, Tufte's unyielding approach towards graphical display could well be argued in the auditory domain also. With the widespread usage of digital playback mechanisms and in-ear headphones of known response, this requirement does not seem too difficult to approximate.

5. DISCUSSION

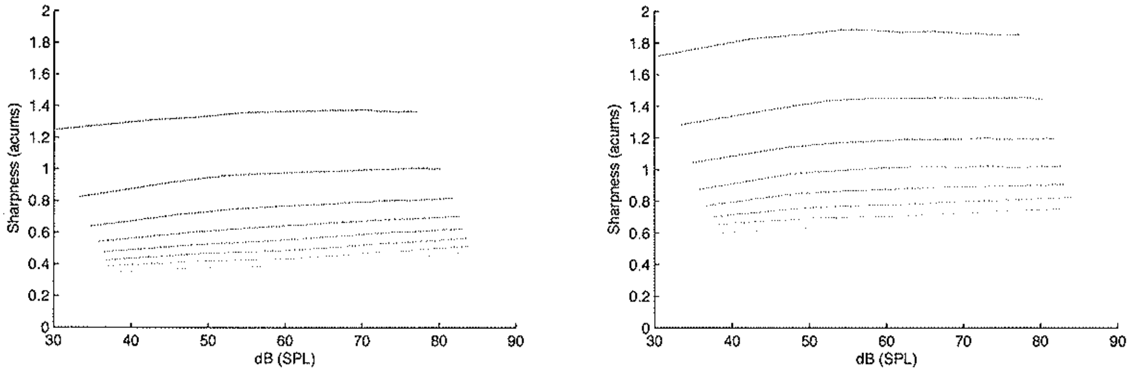
The sonification of information can be a complex process involving many levels of abstraction. Vickers has described how different types of mapping strategy may serve to obfuscate the intended data representation through using multiples levels of abstraction.

Table 1: *Manipulated synthesis parameters, and the psychoacoustic parameter over which they have maximum effect.*

Manipulation Method	**Range**	**Psychoacoustic Parameter**
Amplitude Modulation Index	$f_{mod} = 4$ Hz, modulation index varied	Fluctuation Strength
Frequency Modulation Index	$f_{mod} = 70$ Hz, modulation index varied	Roughness
Fundamental Frequency	220 – 880 Hz	Pitch
Gain	50 dB	Loudness
Spectral Slope	$-\infty$ to +6dB per octave	Sharpness



(a) *Varying level has a strong effect on loudness, but varying spectral slope also has an effect. A weaker effect of tone pitch is also seen.*



(b) *Sharpness is strongly affected by spectral content and the pitch of the tone synthesised, but is also weakly affected by the level of the tone.*

Figure 5: *Selected portions of a matrix with a total of 1800 candidates, which only varies pitch, loudness and sharpness. High frequency spectral content is represented by line saturation, with greater saturation representing greater high frequency content. 220 Hz tones are on the left marked in blue and 440 Hz tones are on the right marked in red.*

He cites a case of 'meta-meta-meta-abstraction' from the intended function to be represented, due to a transfer from a function to discrete tabulated values, to a visual graph, to quantised MIDI note numbers, to pitches that are audible [32]. In addition to those delineated by Vickers, an extra level of abstraction worth considering in order to '...match the signal to the final receiver system...' [33], is the auditory perception of sound.

The mappings of the psychoacoustical parameters of this study to data types has not been examined here, but is likely to be important in their effective use. The use of pitch, loudness and other signal variables has been investigated previously in this respect [34, 35]. For instance, mapping either size or temperature to a frequency scale results in different polarities being preferred [36, 37]. Similar investigations may be warranted for other psychoacoustical models. This may help in rendering displays that auditorally describe their data intuitively and implicitly.

Psychoacoustical models such as these are generally constructed through a compromise between magnitude estimation and magnitude production experiments. According to Marks, subjects tend to compress the response scale in both types of experiments, producing opposing distortions for the two experiment types (this is known as a 'regression effect') [38]. The auditory display application envisaged by this paper is purely magnitude estimation. Hence some compensatory expansion of psychoacoustical scales may yield some improvement in ratio scale representation for this application. This is one of a number of questions remaining to be investigated. These models have a basis different to that of most auditory display parameters; they have been developed through a subjective process that would seem to argue their appropriateness for auditory display. Walker suggests employing ecological theory integrated with psychoacoustics to explore dimensional interaction, which especially benefits high-dimensional display and complex sound presentation [39].

6. RESEARCH AGENDA

Clearly it is important to test the central conjecture of this paper, namely that an effective multidimensional auditory display can be developed from psychoacoustical models. Theoretically, the use of psychoacoustical parameters of sound for conveying data is supported by the methods by which these parameters were modeled initially. However, other considerations may surface when comparing auditory displays employing these methods with more traditional methods of mapping to sound.

Also, comparison between auditory displays attempting to ameliorate the effect of psychoacoustical parameters and those that do not should hopefully measure improvements in comprehension.

Certain psychoacoustical models may intelligently relate the auditory 'character' of a given sound parameter to a data source with a certain type or 'character'. It would be interesting to test these psychoacoustic parameters (especially sharpness, roughness and fluctuation strength) in this regard.

7. CONCLUSIONS

We will present examples of a multidimensional auditory display for graphing in the conference, based on some of the psychoacoustical models outlined here. These examples are developed essentially as an informal experiment, and we invite discussion about the effectiveness of such an approach. Limitations of the approach include the ability of listeners to attend to each psychoacoustical parameter, and the limitations of the model. Zwicker and Fastl do not claim that their models are always quantitatively accurate, especially in the case of their simple models for roughness and fluctuations strength. Alternative models exist for these, as well as for sharpness and loudness, which would yield different scaling. Nevertheless, psychoacoustical modeling can certainly contribute much to the field of auditory display, and we hope that this project stimulates further work in this area.

8. ACKNOWLEDGEMENTS

The authors thank the University of Salford for making available Matlab software for loudness and sharpness modeling. Ferguson's research was conducted with the assistance of an Australian Postgraduate Award and a Faculty of Architecture Departmental Top-up Scholarship. Beilharz's research is supported by a University of Sydney Bridging Support Grant in 2006.

9. REFERENCES

[1] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*. Berlin; New York: Springer, 1999.

[2] R. Parncutt, *Harmony: A Psychoacoustical Approach*. Berlin: Springer, 1989.

[3] D. Cabrera, "'PsySound': A computer program for the psychoacoustical analysis of music," *MikroPolyphonie*, vol. 5, 1999.

[4] R. Shrivastav, "The use of an auditory model in predicting perceptual ratings of breathy voice quality," *Journal of Voice*, vol. 17, no. 4, pp. 502–512, 2003.

[5] J. G. Beerends, A. P. Hekstra, A. W. Rix, and M. P. Hollier, "Perceptual Evaluation of Speech Quality (PESQ) The New ITU Standard for End-to-End Speech Quality Assessment Part II: Psychoacoustic Model," *Journal of The Audio Engineering Society*, vol. 50, no. 10, pp. 765–778, 2002.

[6] K. Brandenburg and M. Bosi, "Overview of MPEG audio: Current and future standards for low bit-rate audio coding ," *Journal of The Audio Engineering Society*, vol. 45, no. 1/2, pp. 4–21, 1997.

[7] S. Barrass, "A perceptual framework for the auditory display of scientific data," in *Proceedings of the International Conference on Auditory Display*, (Santa Fe), 1994.

[8] S. Barrass, "A comprehensive framework for auditory display," *ACM Transactions on Applied Perception (TAP)*, vol. 2, no. 4, pp. 389–402, 2005.

[9] J. G. Neuhoff and L. M. Heller, "One small step: Sound sources and events as the basis for auditory graphs," in *Proceedings of the Eleventh Meeting of the International Conference on Auditory Display*, (Limerick, Ireland), 2005.

[10] E. Zwicker and B. Scharf, "A model of loudness summation," *Psychological Review*, vol. 72, no. 1, pp. 3–26, 1965.

[11] S. S. Stevens, "The measurement of loudness," *Journal of the Acoustical Society of America*, vol. 27, no. 5, pp. 815–829, 1955.

[12] S. S. Stevens, "Perceived level of noise by mark VII and decibels (E)," *Journal of the Acoustical Society of America*, vol. 51, no. 2, pp. 575–601, 1972.

[13] B. C. J. Moore, B. R. Glasberg, and T. Baer, "A model for the prediction of thresholds, loudness, and partial loudness," *Journal of the Audio Engineering Society*, vol. 45, no. 4, pp. 224–240, 1997.

[14] B. R. Glasberg and B. C. J. Moore, "A model of loudness applicable to time-varying sounds," *Journal of the Audio Engineering Society*, vol. 50, no. 5, pp. 331–342, 2002.

[15] J. H. Flowers, "Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions," in *Eleventh Meeting of the International Conference on Auditory Display*, (Limerick, Ireland), 2005.

[16] R. M. Warren, *Auditory Perception: A New Analysis and Synthesis.*. Cambridge: Cambridge University Press, 1999.

[17] G. v. Bismark, "Timbre of steady sounds: A factorial investigation of its verbal attributes," *Acustica*, pp. 146–159, 1974.

[18] W. Lichte, "Attributes of complex tones," *Journal of Experimental Psychology*, vol. 28, no. 6, pp. 455–480, 1941.

[19] G. v. Bismark, "Sharpness as an attribute of the timbre of steady sounds," *Acustica*, vol. 30, pp. 159–172, 1974.

[20] W. v. Aures, "Berechnungsverfahren für den sensorischen Wohlklang beliebiger Schallsignale," *Acustica*, vol. 59, pp. 130–141, 1985.

[21] R. Plomp and W. J. M. Levelt, "Tonal consonance and critical bandwidth," *Journal of the Acoustical Society of America*, vol. 38, pp. 548–560, 1965.

[22] W. v. Aures, "Ein Berechnungsverfahren der Rauhigkeit," *Acustica*, vol. 58, pp. 268–281, 1985.

[23] P. Daniel and R. Weber, "Psychoacoustical roughness: implementation of an optimized model," *Acustica*, no. 83, pp. 113–123, 1997.

[24] E. J. Hellier, J. Edworthy, and I. Dennis, "Improving auditory warning design: Quantifying and predicting the effects of diffferent warning parameters on perceived urgency," *Human Factors*, vol. 35, no. 4, pp. 693–706, 1993.

[25] S. S. Stevens, J. Volkman, and E. Newman, "A scale for the measurement of the psychological magnitude of pitch," *Journal of the Acoustical Society of America*, vol. 8, pp. 185–190, 1937.

[26] S. S. Stevens and J. Volkman, "The relation of pitch to frequency," *American Journal of Psychology*, vol. 53, pp. 329–353, 1940.

[27] D. D. Greenwood, "The mel scale's disqualifying bias and a consistency of pitch-difference equisections in 1956 with equal cochlear distances and equal frequency ratios," *Hearing Research*, vol. 103, pp. 199–224, 1997.

[28] R. N. Shepard, "Geometrical approximations to the structure of musical pitch," *Psychological Review*, vol. 89, no. 4, pp. 305–333, 1982.

[29] E. Terhardt, G. Stoll, and M. Seewan, "Algorithm for extraction of pitch and pitch salience from complex tonal signals," *Journal of the Acoustical Society of America*, vol. 71, no. 3, pp. 679–688, 1982.

[30] J. C. R. Licklider, "A duplex theory of pitch perception," *Experientia*, vol. 7, no. 128-133, 1951.

[31] E. R. Tufte, *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire, Conn: Graphics Press, 1997.

[32] P. Vickers, "Whither and wherefore the auditory graph? Abstractions and aesthetics in auditory and sonified graphs," in *Eleventh Meeting of the International Conference on Auditory Display*, (Limerick, Ireland), 2005.

[33] E. Zwicker and U. T. Zwicker, "Audio engineering and psychoacoustics: Matching signals to the final receiver, the human auditory system," in *Readings in multimedia computing and networking*, pp. 11 – 22, Morgan Kaufmann, 2001.

[34] B. N. Walker and G. Kramer, "Mappings and metaphors in auditory displays: An experimental assessment," in *Proceedings of the International Conference on Auditory Display*, (Palo Alto, California), 1996.

[35] B. N. Walker, G. Kramer, and D. M. Lane, "Psychophysical scaling of sonification mappings," in *Proceedings of the International Conference on Auditory Display*, (Atlanta), 2000.

[36] B. N. Walker, "Magnitude Estimation of conceptual data dimensions for use in sonification," *Journal of Experimental Psychology: Applied*, no. 8, pp. 221–221, 2002.

[37] B. N. Walker, *Magnitude Estimation of Conceptual Data Dimensions for Use in Sonification*. Ph.D Thesis, Rice University, 2000.

[38] L. Marks, *The New Psychophysics*. New York: Academic Press, 1974.

[39] B. N. Walker and G. Kramer, "Ecological psychoacoustics and auditory displays: Hearing, grouping and meaning making," in *Ecological Psychoacoustics* (J. G. Neuhoff, ed.), New York: Academic Press, 2004.

ICAD 2006 @ Queen Mary, University of London

Proceedings of the 12th Meeting of the ICAD, 20-23 June 2006

Proceedings of ICAD 2006 at Queen Mary, University of London

The organising committee of ICAD 2006 would like to thank all authors for their contribution to these proceedings.

Editors: Tony Stockman, Louise Valgerður Nickerson, Christopher Frauenberger, Alistair D. N. Edwards and Derek Brock

Papers Chair: Alistair D. N. Edwards

Posters Chair: Derek Brock

Thinktank Chair: Paul Vickers

Concert Chair: Alberto de Campo

ISBN 0-902-23821-3 (CDROM)

ISBN 0-902-23820-5 (online)

Additional ICAD information and publications can be found at http://www.dcs.qmul.ac.uk/icad2006 or http://www.icad.org

Published by the Department of Computer Science, Queen Mary, University of London, UK.