Optimizing the Hearing Aid Musical Experience

Eric Battenberg

May 18, 2007

Abstract

A novel approach to hearing aid fitting for music listening is presented. The method uses a subjective space to reduce the dimensionality of the hearing aid parameter space. This is accomplished using neural network regression to interpolate among hearing aid parameter settings. This listener-driven method provides not only a technique for optimal aid fitting, but also information on individual differences and the effects of gain compensation on different musical styles.

1 Introduction

Modern digital hearing aid technology that is heavily optimized for speech intelligibility often makes music unlistenable so that hearing aid wearers prefer to remove their hearing aids when listening to music. The effects of hearing aid processing on musical signals, and on the perception of music have received very little attention. There is no standard test of music perception, and to make the problem more difficult, different musical styles thrive in strikingly different acoustical environments. There have been some studies on the effect of reduced bandwidth on the perceived quality of music [1], but no systematic evaluation of the effects of dynamic range compression, the most ubiquitous form of gain compensation in digital hearing aids.

This paper presents a subjective listener-driven method for optimizing the parameter set of a digital hearing aid. A radial basis network is used as a regression method to interpolate a subspace of parameters settings. The listener navigates this subspace in real time using a two-dimensional interface and is able to quickly converge on his or her personal optimum for the parameter set.

2 Correcting Hearing Loss

2.1 Characteristics of Hearing Loss

Sensorineural hearing loss is the primary type of hearing loss that is treated by hearing aid devices. This type of loss results from damage or deterioration of the outer hair cells of the cochlea. Typical characteristics of sensorineural hearing loss include an increase in high frequency hearing threshold, a loss of automatic gain control in the inner ear, and a decrease in spectral and temporal resolution of the auditory system.

Compensating for a high frequency hearing threshold shift can be accomplished easily using linear frequency dependent amplification. This type of processing simply makes higher frequency sounds louder and does not solve the loss of automatic gain control in the inner ear. Typical loudness

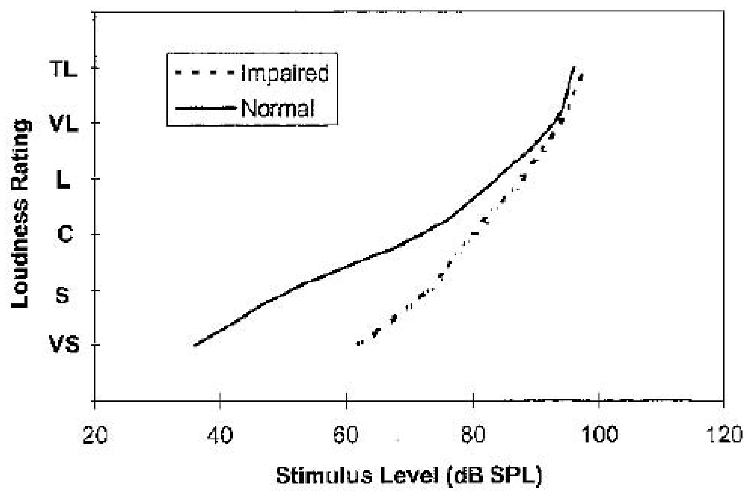

Figure 1: *Normal vs. Hearing Impaired Loudness Curve; VS - Very Soft, S - Soft, C - Comfortable, L - Loud, VL - Very Loud, TL - Too Loud [2]*

curves for individuals with normal hearing and hearing impairment are shown in Figure 1 [2]. The hearing impaired curve shows some loudness expansion, or steepening of the curve, compared to the normal curve. For the hearing impaired listener, this causes quiet sounds to be practically inaudible and loud sounds to become painfully loud.

The effect of linear amplification can be seen in Figure 2 [2]. The linearly amplified curve, though equal in perceived loudness to that perceived by someone with normal hearing at around 70 dB SPL, is too quiet at lower SPL values and much too loud at higher values.

2.2 Wideband Compression

The solution to this problem is non-linear amplification in the form of gain compression. The input/output curves of compressive amplification schemes with various compression ratios can be seen in Figure 3. The curve with a 1:1 ratio represents linear amplification, while at the extreme end we have limiting, which keeps the output under a certain level. Compression ratios used in hearing aids typically do not exceed 2.5:1.

In addition to choosing the optimal compression ratio for a particular user, compressors depend on other variable parameters. These include the compression threshold, the minimum input level to which compression is applied; the attack time, which determines how quickly gain is decreased when the signal level increases; and the release time, which determines how quickly the gain is increased when the signal level decreases. These are the basic parameters which characterize a wideband compression system. Other parameters found in digital hearing aids control noise reduction, feedback suppression, and directional microphone processing. Since advanced functions such as these are typically detrimental to music listening [3], we will continue to focus on optimizing the compression settings.

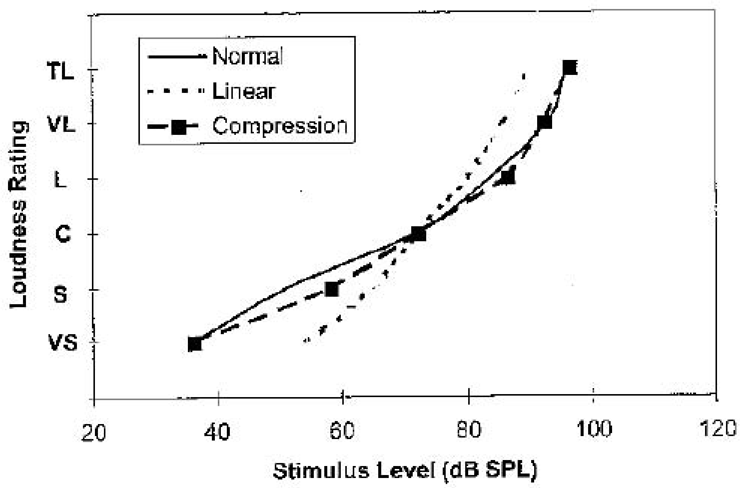

Figure 2: *Linear vs. Compressive Amplification [2]*

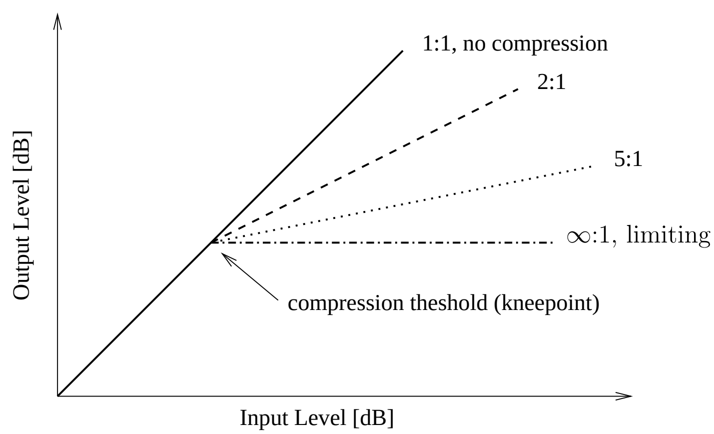

Figure 3: *Compression Curves of Various Ratios*

2.3 Multiband Compression

Because the loudness curves shown in Figure 1 typically vary with frequency, a wideband compressor operating at a single ratio does not adequately compensate for a loss of automatic gain control. In addition, the gain control of healthy outer hair cells operates in localized frequency regions. For these reasons, digital hearing aids have employed multiband compression systems, which filter a signal into several frequency bands and apply compression separately to each band.

Multiband compression is the primary means of compensating for sensorineural hearing loss in state-of-the art digital hearing aids. The problem faced by such a system is how one chooses a set of parameters for each band that optimally correct the distinct characteristics of an individual's hearing loss. The entire set of parameters can be thought of as a hearing aid "prescription". Solving such a large-dimensional optimization problem is no small feat, and it makes an optometrist's job seem trivial.

Currently, hearing aid parameters are chosen using extensive audiogram testing of a patient's hearing. Loudness curves at various frequencies can be drawn to characterize the specifics of the hearing loss, and then compression parameters can be set to attempt to correct the dysfunctional curves. Going beyond the audiogram fitting, many audiologists attempt to tweak the parameters to optimize speech intelligibility for the user. This process varies between audiologists, but combined with speech-specific improvements in digital hearing aids such as noise reduction, it usually provides satisfactory results. A variety of objective tests of speech comprehension are used to measure hearing aid efficacy [4] [5]; however, there is no standard metric for measuring a patient's perception of music. What follows is a novel approach to selecting compression parameter values to optimize a patient's music listening experience.

3 Dimensionality Reduction Via a Subjective Space Approach

3.1 Multidimensional Scaling

Characterizing perceptual dissimilarity as distance in a geometric representation has provided auditory researchers with a rich set of robust methods for studying the structure of perceptional attributes [6]. By viewing various stimuli in a two-dimensional plane with the distance between them representing their dissimilarity, one can visualize complex, high-dimensional quantities. The most well-known method for generating such a spatial representation is the multidimensional scaling (MDS) of pairwise dissimilarity judgments [7]. In the method, subjects rate the dissimilarity of pairs of stimuli. The MDS method treats the stimuli as points in a lower-dimensional space, in our case two-dimensional, and finds the spatial arrangement of the stimuli that maximizes the correlation between distances in the arrangement and subjective distances in the pairwise dissimilarity judgments.

3.2 Spatial Layout Technique

An alternative to the MDS method is to directly arrange the stimuli in a subjectively meaningful way on the two-dimensional space. This method has been found to be comparable in quality to MDS [8] [9], and in our case, much more efficient. This project includes the implementation of an interface, created in the Max/MSP audio programming environment [10], that allows the user to place various stimulus objects, each which represent a set of compression parameters, according to

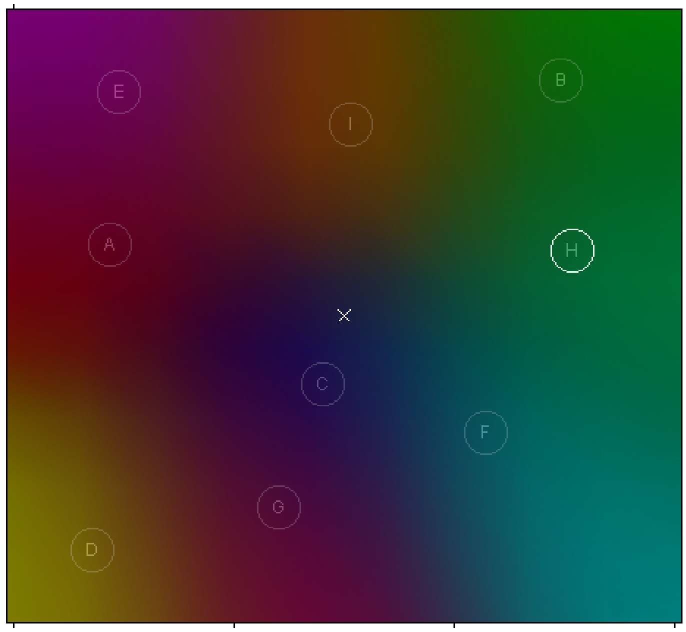

Figure 4: *Subspace Interface - The parameter sets are represented by clickable and moveable circles. The 'x' denotes the current interpolated location.*

these ideas of spatial dissimilarity. Within the interface, the object corresponding to each parameter set can be clicked on to hear how the parameters sound for a particular passage of music. Then the objects can be dragged to subjectively significant locations within the space. Once the objects are arranged, a type of interpolation can be used to fine tune the compression parameters. Section 4 describes how this interpolation is implemented.

3.3 Loudness Matching

One more consideration when navigating a subjective auditory space is that users will strongly prefer stimuli that sound louder to them [11]. In preliminary tests with the interface users preferred settings that sounded louder, regardless of the specifics of the compression parameters. To eliminate this dependence on subjective loudness, the system includes a loudness matching function, in which users equalize the overall loudness of each parameter set to the loudness of a reference setting by adjusting a final gain parameter. Loudness matching is still a complex problem when considering the diversity of hearing loss characteristics and how they affect loudness impression; however, the loudness matching function of this system seems to correct obvious loudness inequalities.

3.4 Subjective Space Navigation

After the initial parameter sets are equalized in loudness and their respective objects are placed subjectively within the space, the user can navigate the space using a cursor and explore a continuously varying spectrum of parameter settings interpolated from the surrounding objects. The instant visual and auditory feedback that a user receives from this process is especially helpful in directing the user toward their optimum parameter set within the space.

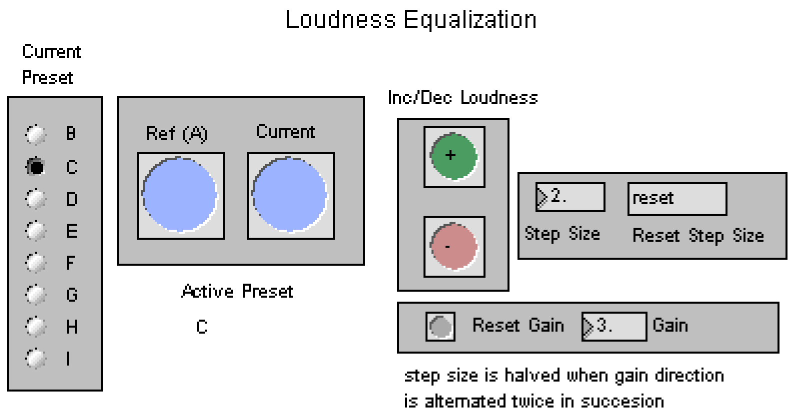

Figure 5: *Loudness Matching Interface - The user can alternate between the current parameter set and a reference and adjust the gain accordingly.*

The speed with which new preferred parameter sets can be discovered allows the user to easily select values for different musical styles and listening environments. The parameter set data provided by this procedure can also supply additional information on the specific characteristics of the patient's hearing loss.

4 Interpolation Using a Radial Basis Network

4.1 Network Architecture

Interpolation within the subspace is performed using a radial basis network [12] composed of a radial basis hidden layer and a linear output layer as shown in Figure 6. This simple two layer design is very effective in accomplishing our goal of parameter interpolation.

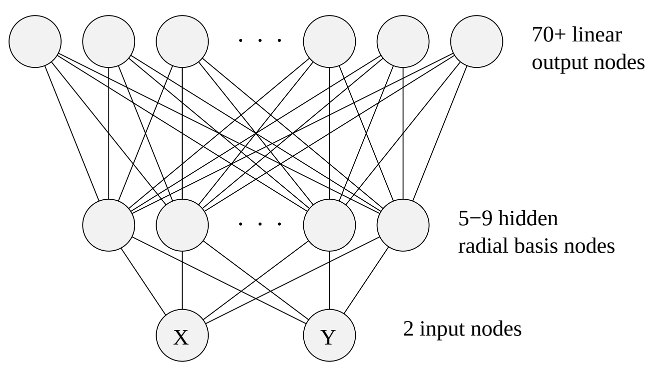

Figure 6: *Neural Network Architecture*

The specifics of the system are shown in Figure 7. To begin, the neural network takes the two-dimensional input vector and measures its distance from each of the q preset locations which are stored as the columns of a matrix $\mathbf{L}$. The output of this distance measure is a q-dimensional vector which is then scaled by a constant a and then passed through the Gaussian radial basis function.

The constant a affects the spread of the Gaussian function and ultimately controls the smoothness of the interpolation space. The output of the radial basis function is a q-dimensional vector of preset weights. For example, if the input location corresponds to one of the preset locations, then the weight corresponding to that preset would be 1. The radial basis weight vector is now the input to the linear output layer.

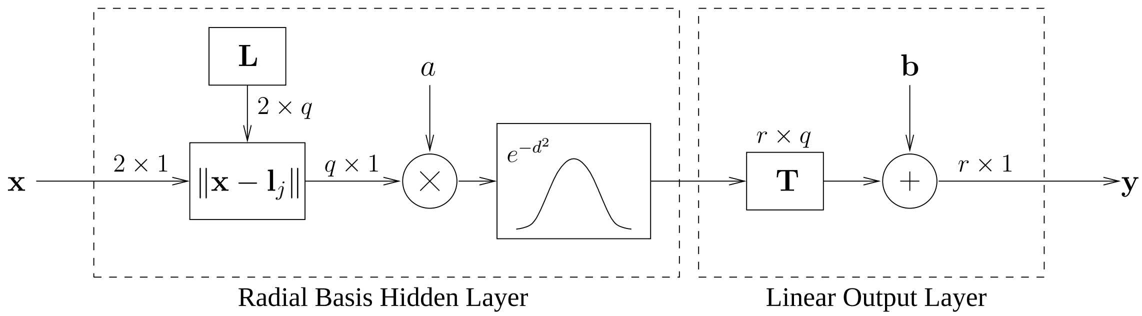

Figure 7: *System Diagram*

The linear layer consists of a mapping from the q-dimensional weight vector to the r-dimensional parameter space. This linear transformation is carried out using a matrix $\mathbf{T}$ that left multiplies the weight vector and a constant vector $\mathbf{b}$ which is summed with the resulting matrix product. If $\mathbf{w}$ is the weight vector and $\mathbf{y}$ the output vector, we have

$$\mathbf{y} = \mathbf{T}\mathbf{w} + \mathbf{b}. \tag{1}$$

4.2 Training

The training of the network is fairly simple and does not require any complex iterative algorithms. This allows the network to be retrained in real-time, so that the user can instantly experience the effects of moving presets within the space. The network is trained so that each preset location elicits an output equal to the exact parameter set corresponding to that preset.

The values that must be determined by training are the preset location matrix $\mathbf{L}$, the linear transformation matrix $\mathbf{T}$, and the vector $\mathbf{b}$. The matrix $\mathbf{L}$ is simply constructed by placing each two-dimensional preset location into a separate column of the matrix. The complexity of this process is trivial. The matrix $\mathbf{T}$ and vector $\mathbf{b}$ are chosen so that if the input location lies directly on a preset, then the output will be the parameters corresponding to that preset. To solve for these, we can set up a linear system of equations.

We can place $\mathbf{T}$ and $\mathbf{b}$ together in a matrix

$$\mathbf{T}' = [\mathbf{T}|\mathbf{b}]\,. \tag{2}$$

Then we place the weight vectors corresponding to each preset location into a matrix $\mathbf{W}$ and append a row vector of ones, $\mathbf{1}_{1\times q}$, so that

$$\mathbf{W}' = \left[\frac{\mathbf{W}}{\mathbf{1}_{1\times q}}\right]. \tag{3}$$

Let the matrix $\mathbf{P}$ be the target matrix composed of columns of the parameters corresponding to each preset. Now our linear system of equations can be represented by the single matrix equation

$$\mathbf{T}'\mathbf{W}' = \mathbf{P} \tag{4}$$

Because there are more degrees of freedom in the system than constraints, the system is underdetermined and has infinitely many solutions. We choose the solution, $\mathbf{T}'$ with the lowest norm by right multiplying by the pseudo-inverse of $\mathbf{W}'$. The solution with lowest norm was chosen to prevent the system from displaying erratic behavior and to keep any one weight from dominating the output.

After we have solved for $\mathbf{T}$ and $\mathbf{b}$, the training is complete. Compared to other neural network training procedures like back propagation, this method is extremely fast and still produces the desired results.

5 Conclusion

This project involved the development of a listener-driven interactive system for adjusting the high-dimensional parameter space of a multiband compressor system in a digital hearing aid. The three components of the system include loudness matching, subjective spatial layout, and nonlinear regression of the entire space. The real-time interactivity allows simple, intuitive exploration of a high-dimensional space of parameters. The hope is that this system can be used to help gather information about optimal corrections for patients with different types of hearing loss and as a tool in hearing aid fitting procedures.

In the near future, a hearing loss simulator will be inserted into the system, and a "Golden-Ears" experiment will be performed, in which users will attempt to correct their simulated hearing loss using the subjective subspace. Since this project is funded by Starkey Laboratories, a developer of hearing devices, the system will eventually be interfaced with actual hearing aids in tests that will determine its true efficacy.

References

[1] J. R. Franks, "Judgments of Hearing Aid Processed Music," *Ear and Hearing*, vol. 3, no. 1, pp. 18–23, 1982.

[2] B. Edwards, "Hearing aids and hearing impairment," in *Speech Processing in the Auditory System*, S. Greenberg, W. Ainsworth, A. N. Popper, and R. R. Fay, Eds. New York, NY: Springer-Verlag, 2004, pp. 339 – 421.

[3] M. Chasin and F. A. Russo, "Hearing Aids and Music," *Trends in Amplification*, vol. 8, no. 2, pp. 35–47, 2004. [Online]. Available: http://tia.sagepub.com/cgi/content/abstract/8/2/35

[4] "Methods for the calculation of the speech intelligibility index," American National Standards Institute, New York, NY, Tech. Rep. ANSI S3.2-1997, 1997.

[5] K. Kryter, "Methods for the Calculation and Use of the Articulation Index," *The Journal of the Acoustical Society of America*, vol. 34, p. 1689, 1962.

[6] R. Shepard, "Multidimensional Scaling, Tree-Fitting, and Clustering," *Science*, vol. 210, no. 4468, p. 390, 1980.

[7] I. Borg and P. Groenen, *Modern Multidimensional Scaling: Theory and Applications*. Springer, 1997.

[8] R. Goldstone, "An efficient method for obtaining similarity data," *Behavior Research Methods, Instruments, & Computers*, vol. 26, pp. 381–386, 1994.

[9] A. Momeni and D. Wessel, "Characterizing and controlling musical material intuitively with geometric models," *Proceedings of the 2003 conference on New interfaces for musical expression*, pp. 54–62, 2003.

[10] Cycling '74, "Max/MSP." [Online]. Available: http://www.cycling74.com/products/maxmsp

[11] H. Fastl, *Communication Acoustics*. Berlin, Germany: Springer, 2005, ch. Psycho-acoustics and sound quality, pp. 139–162.

[12] M. Buhmann, *Radial Basis Functions: theory and implementations*. Cambridge University Press, 2003.