LEARNING TO CONTROL A REVERBERATOR USING SUBJECTIVE PERCEPTUAL DESCRIPTORS

Zafar Rafii
EECS Department
Northwestern University
Evanston, IL, USA
ZafarRafii2011@u.northwestern.edu

Bryan Pardo
EECS Department
Northwestern University
Evanston, IL, USA
pardo@northwestern.edu

ABSTRACT

The complexity of existing tools for mastering audio can be daunting. Moreover, many people think about sound in individualistic terms (such as "boomy") that may not have clear mappings onto the controls of existing audio tools. We propose learning to map subjective audio descriptors, such as "boomy", onto measures of signal properties in order to build a simple controller that manipulates an audio reverberator in terms of a chosen descriptor. For example, "make the sound less boomy". In the learning process, a user is presented with a series of sounds altered in different ways by a reverberator and asked to rate how well each sound represents the audio concept. The system correlates these ratings with reverberator parameters to build a controller that manipulates reverberation in the user's terms. In this paper, we focus on developing the mapping between reverberator controls, measures of qualities of reverberation and user ratings. Results on 22 subjects show the system learns quickly (under 3 minutes of training per concept), predicts users responses well (mean correlation coefficient of system predictiveness 0.75) and meets users' expectations (average human rating of 7.4 out of 10).

1. INTRODUCTION

In recent decades, many audio production tools have been introduced to enhance and facilitate music creation. Often, these tools are complex and conceptualized in terms ("high cut","density") that are unfamiliar to many users. This makes learning these tools daunting, especially for inexperienced users.

One solution would be to redesign the standard interfaces to manipulate audio in terms of commonly used descriptors (e.g. "warm" or "enveloping"). This can be problematic, since the meanings of many words used to describe sound differ from person to person or between different groups [1]. For example, the audio signal properties associated with "warm" and "clear" have been shown to vary between English speakers from the UK and the US [2]. Since it may not be possible to create general controllers for terms whose meaning varies between groups, we propose mapping descriptive terms onto the controls for audio tools on a case-by-case basis.

While there has been much work on adaptive user interfaces [3], there has been relatively little on personalization of audio tools. A previous study showed success in personalizing an equalization tool [4]. Here, we propose to simplify and personalize the interface to one of the most widely applied classes of audio effect: *Reverberation*.

Reverberation is created by the reflections of a sound in an enclosed space causing a large number of echoes to build up and then slowly decay as the sound is absorbed by the walls and air [5]. The reflections modify the perception of the sound, its loudness, timbre and spatial characteristics [6]. Reverberation can be simulated using multiple feedback delay circuits to create a large, decaying series of "echoes" [7], and many reverberation tools have been built. Fig. 1 shows the interface of a typical reverberation tool. Note the 7 buttons and 14 sliders that control parameters (such as "density") whose meaning in this context are unfamiliar to the average person and to many musicians.

Figure 1. *Logic Audio*'s *Platinumverb* complex interface.

We propose a system that learns an audio concept a user has in mind ("boomy", for example) and builds a simple

reverberation controller to manipulate sound in terms of that descriptor. By automatically adapting the interface to an individual user's conceptual space, we hope to bypass the creative bottleneck caused by complex interfaces and individual differences in the meaning of descriptive terms.

The paper is organized as follows. The method used to map descriptive terms onto audio signal characteristics is described in Section 2. Section 3 presents the reverberation control used to perceptually alter sound. Experimental evaluation of the approach is described in Section 4. Finally, conclusions are given in Section 5.

2. LEARNING DESCRIPTIVE TERMS

We now give an overview of the process by which the system learns to build a controller that controls the reverberation of a signal in terms of a user-defined descriptive word.

2.1 The Training Process

In the training process, the user is presented with the *Perceptual Learner* interface shown in Fig. 2. The user selects a descriptive word (such as "boomy" or "church-like") to teach the system. The user is then presented with a series of audio examples generated from an original audio file and processed by the reverberator using a variety of reverberation settings. The reverberation settings used are chosen to explore the space of likely parameter settings for a digital reverberator, as described in Section 4.

The user moves a slider to rate each audio example on how well it represents the audio descriptor. Ratings range from 1 (captures the concept perfectly) to -1 (does not capture it at all). Training typically takes about 30 ratings (around two minutes for a five-second file). Fig. 3 illustrates the process.

Figure 2. Interface of the Perceptual Learner.

2.2 Mapping Signal Statistics to User Ratings

The system collects five impulse response measures (described in Section 3.2) for the reverberation applied to each example rated by the user. Once user ratings are collected, the system relates user ratings to each of the five measures using *linear regression*. This lets us build a model that predicts the expected user rating, given a reverberation impulse response signal characterized by these measures.

This mapping is used to build a controller that lets the user easily manipulate the audio in terms of the descriptor

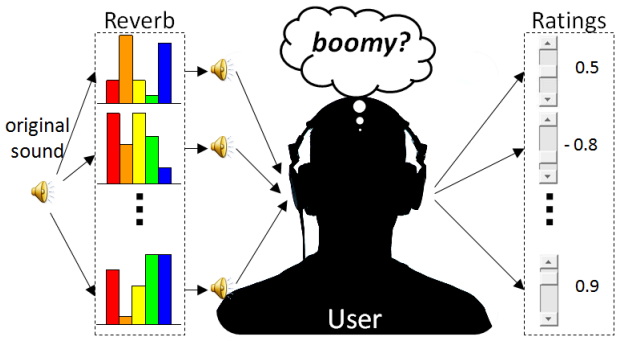

Figure 3. The training process: (1) audio examples are generated from an original sound using a reverberator set to a variety of parameter settings (5 control parameters shown by 5 different bars); (2) the user listen to the audio examples and uses a slider to rate how well each one fits the audio concept she/he has in mind.

(such as "boomy") using a simple slider as shown in Fig. 4. This slider affects all five reverberation measures in parallel, although not necessarily in the same direction. For example, "boomy" may be positively correlated with central time and negatively correlated with spectral centroid.

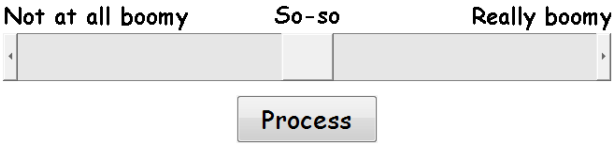

Figure 4. Interface of the Perceptual Controller.

3. THE REVERBERATION CONTROL

To build the new interface, we must map human feedback to reverberation controls. We do not, however, map user feedback directly to parameters for a specific reverberator, but onto measures of the reverberation (Section 3.2). This lets us use mappings learned using one reverberator to control another one, chosen later. The only requirement is that both reverberators have known mappings between control parameters and reverberation measures.

3.1 The Digital Reverberator

The approach we describe, while not tied to any particular reverberation approach, works best if the reverberator can generate a wide variety of impulse response functions on the fly. Thus, rather than use a convolution reverberator that selects from a fixed library of impulse responses, we have developed a digital stereo reverberation unit inspired by Moorer's work [8]. The reverberator, shown in Fig. 5, is easy to manipulate through the control parameters. The reverberation measures described in Section 3.2 can be derived easily as functions of those parameters. This is important for learning a mapping between human feedback and reverberator settings.

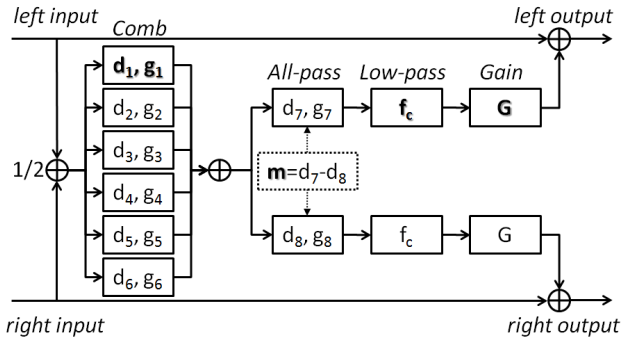

Figure 5. The digital stereo reverberation unit.

The reverberator uses six *comb* filters in parallel to simulate the complex modal response of a room by adding echoes together. Each comb filter is characterized by a delay factor d_k and a gain factor g_k (k=1..6). The delay values are distributed linearly over a ratio of 1:1.5 with a range between 10 and 100 msec, so that the delay of the first comb filter d_1, defined as the longest one, determines the other delays. The gain factor of the first comb filter g_1 has the smallest gain and has a range of values between 0 and 1. Although a comb filter gives a non-flat frequency response, a sufficient number of comb filters in parallel with equal values of *reverberation time* helps to reduce the spectral coloration.

An *all-pass* filter is added in series to increase the *echo density* produced by the comb filters without introducing spectral coloration, and doubled into two channels to simulate a more "natural sounding" reverberation in stereo. The all-pass filter is characterized by a delay factor d_a of 6 msec and a gain factor g_a fixed to $\frac{1}{\sqrt{2}}$. A small difference m is introduced between the delays to insure a difference between the channels, therefore the delays become $d_7 = d_a + \frac{m}{2}$ for the left channel and $d_8 = d_a - \frac{m}{2}$ for the right channel. The range of values for $m = d_7 - d_8$ is then defined between 0 and 12 msec. Note that to prevent exactly overlapping echoes, the delay values for the comb and the all-pass filters are set to the closest inferior prime number of samples.

To simulate air and walls absorption, a first-order low-pass filter of gain g_c defined from its cut-off frequency f_c is added at each channel [9]. f_c ranges between 0 and half of the frequency sampling f_s. Finally, a gain parameter G, whose range of values is between 0 and 1, controls the wet/dry effect. In summary, a total of only five independent parameters are needed to control the reverberator: $\boldsymbol{d}_1$, $\boldsymbol{g}_1$, $\boldsymbol{m}$, $\boldsymbol{f}_c$ and $\boldsymbol{G}$. The other parameters can be deduced from them according to the relations above.

3.2 The Reverberation Measures

We now define five measures commonly used to characterize reverberation and describe formulae to estimate values for these measures in terms of the parameters for our reverberator. For details on how we derive these formulae, we refer the reader to [10].

- *Reverberation Time* (T_{60}) is defined as the time in sec required for the reflections of a direct sound to decay by 60 dB below the level of the direct sound [5]. Based on the reverberation time of the comb filter and the other gains, we estimated the reverberation time of the whole reverberation unit as follows in Eq. 1.

$$T_{60} = \max_{k=1..6} \left(d_k \ \log \left(\frac{10^{-3}}{g_a \ (1 - g_c) \ G} \right) / \log g_k \right) \quad (1)$$

- *Echo Density* (D_t) is defined as the number of echoes per second at a time t. In practice, we computed the average echoes per second between time 0 and time t. We estimated the echo density of the whole reverberation unit at time t = 100 msec, as a combination of echo densities of the digital filters, as follows in Eq. 2.

$$D = \frac{t}{d_a} \sum_{k=1}^{6} \frac{1}{d_k} \quad (2)$$

- *Clarity* (C_t) describes the ratio in dB of the energies in the impulse response p before and after a given time t. It provides indication of how "clear" the sound is [11]. The definition of C_t in discrete time is given by Eq. 3.

$$C_t = 10 \log_{10} \left(\sum_{n=0}^{t} p^2[n] / \sum_{n=t}^{\infty} p^2[n] \right) \quad (3)$$

We estimated the clarity of the whole reverberation unit at t = 0, the arrival time of the direct sound, as shown in Eq. 4, assuming that the total energy of the reverberator is a linear combination of the energies of its filters.

$$C = -10 \log_{10} \left(G^2 \ \frac{1 - g_c}{1 + g_c} \ \sum_{k=1}^{6} \frac{{g_k}^2}{1 - {g_k}^2} \right) \quad (4)$$

- *Central Time* (T_C) is the "center of gravity" of the energy in the impulse response p, [11], defined in discrete time by Eq. 5.

$$T_C = \sum_{n=0}^{\infty} n p^2[n] / \sum_{n=0}^{\infty} p^2[n] \quad (5)$$

Based on the same assumption as for clarity, we estimated the central time of the whole reverberation unit as the combination of central times of the filters, as follows in Eq. 6.

$$T_C = \sum_{k=1}^{6} \frac{d_k {g_k}^2}{(1 - {g_k}^2)^2} / \sum_{k=1}^{6} \frac{{g_k}^2}{1 - {g_k}^2} \ + \ d_a \quad (6)$$

- *Spectral Centroid* (S_C) is the "center of gravity" of the energy in the magnitude spectrum P of the impulse response p, defined in discrete time by Eq. 7, where f_s is the sampling frequency.

$$S_C = \sum_{n=0}^{f_s/2} n P^2[n] / \sum_{n=0}^{f_s/2} P^2[n] \quad (7)$$

We estimated the spectral centroid of the whole reverberation unit from the characteristics of its low-pass filter, as

follows in Eq. 8.

$$S_C = \frac{\sum_{n=0}^{f_s/2} \frac{n}{1 + {g_c}^2 - 2\, g_c \cos(2\pi n)}}{\sum_{n=0}^{f_s/2} \frac{1}{1 + {g_c}^2 - 2\, g_c \cos(2\pi n)}} \quad (8)$$

Based on the relations between the parameters defined in section 3.1, the measures can be redefined as five functions of five independent parameters: $\boldsymbol{T}_{60}(\boldsymbol{d}_1, \boldsymbol{g}_1, \boldsymbol{f}_c, \boldsymbol{G})$, $\boldsymbol{D}(\boldsymbol{d}_1, \boldsymbol{m})$, $\boldsymbol{C}(\boldsymbol{g}_1, \boldsymbol{f}_c, \boldsymbol{G})$, $\boldsymbol{T}_C(\boldsymbol{d}_1, \boldsymbol{g}_1, \boldsymbol{m})$ and $\boldsymbol{S}_C(\boldsymbol{f}_c)$.

Note that these functions are not entirely invertible, especially for d_1 and g_1. When necessary, we estimate d_1 and g_1 from a reverberation measure by using tables of values.

4. EVALUATION

We have implemented the system in Matlab on a PC with an Intel Core2 Quad CPU of 2.66GHz and 6GB of RAM. The system was evaluated by 22 participants, 14 males and 8 females, between the ages of 18 and 29. All reported normal hearing and were native English speakers. 10 had a little or no musical background and 12 had a strong musical background i.e. practicing one or several instruments, more than 1 hour per week and for more than 10 years, or more than 6 hours per week and for more than 6 years.

All audio examples created were based on a 5.5 sec anechoic recording of an unaccompanied singing male sampled at 44,100 Hz. Prior to the study, a database of 1024 impulse response functions was generated using the reverberator described in Section 3.1. These impulse response functions were selected to evenly cover a range of the five reverberation measures (Section 3.2). The *Reverberation Time* ranged from 0.5 to 8 sec, the *Echo Density* from 500 to 10,000 echoes/sec, the *Clarity* from -20 to 10 dB, the *Central Time* from 0.01 to 0.5 sec, and the *Spectral Centroid* from 200 to 11,025 Hz (no low-pass filtering). These ranges were chosen by audio inspection so that they evenly cover a range of "good" values in the space of reverberation measures leading to natural sounding reverberation.

4.1 Experiment

Study participants were seated in a quiet room with a computer that controlled the experiment and recorded the responses. The stimuli were presented binaurally over headphones. Participants were allowed to adjust the sound level prior to starting the study. Prior to beginning the study, participants were quickly trained on the task. Each participant participated in a single one-hour session.

Each participant was asked to rate the same five descriptive words: *bright*, *clear* (two common audio descriptors), *boomy* (often related to reverberation), *church-like* and *bathroom-like* (related to models of space, respectively a church and a bathroom). These words were presented to each participant in a random order. For each descriptive word, the participant was asked to perform three tasks.

First, the participant was asked to rate a series of 60 audio examples. For each example the participant heard the audio modified by an impulse response function. The participant moved an on-screen slider (Fig. 2) to indicate the extent to which each sound exemplified the current word descriptor. Values ranged from 1 (captures the concept perfectly) to -1 (does not capture it at all). These 60 audio examples contained 35 examples chosen randomly from our database of 1024 examples. We then duplicated 25 of the 35 and added the duplicates to the set in random order, for a total of 60 examples. The 25-example duplicate set was used to measure consistency of user responses, while the 35-example training set was for system training. A previous study showed that around 25 examples are sufficient to model a user's preferences for an equalization controller [4], which is a closely related task.

Once the first task was completed, the system created a model of the effect of each reverberation measure on the user ratings, as described in Section 2.2. The data set used was the user ratings of the 35 non-duplicate examples in the first task. The new model was used to select a new set of audio examples. This set contained 11 audio examples chosen to evenly cover the range of user ratings from -1 to 1 (as predicted by the learned model) and 14 audio examples selected at random, for a total of 25. The participant was asked to rate the 25 new audio examples as she/he did in the first part.

Finally, the system used the learned model to build a slider (Fig. 4) that controls reverberation in terms of the learned descriptor. The controller mapped 11 audio examples chosen to evenly cover the range of user ratings from -1 to 1 onto slider positions. As the slider is moved to a new location, a different variant of the sound is played. This let the participant move the slider to change the degree of the effect. The user was asked to play with the controller for as long as neccesary to get a feel for how well it worked. The user was then asked to rate how well it manipulated the sound in terms of the learned descriptive word. *Human ratings* ranged from 0 (really bad) to 10 (really good).

4.2 Results

The average training time, over all the descriptive words and the participants, was 4 min 2 sec. Since only 35 of the 60 user-ratings in the first task were actually used for training the system, a model for a descriptive word was learned in only 2 min 20 sec of training (the mean time for a user to rate for 35 examples).

User consistency on a descriptive word was measured by computing the within-user correlation coefficient on rating the 25 pairs of duplicate examples in the first task. Average *user consistency* over all words and users was 0.65.

System predictiveness (how well the system learned) for a descriptive word was measured by computing the correlation coefficient between the user's observed ratings and the system's prediction of the user ratings on the second set of user-rated examples. *System predictiveness* was 0.75, averaged over all words and users.

System predictiveness was measured on a different data set than user consistency, so the results are not directly comparable. That said, the consistency of user ratings on

matched pairs of stimuli gives an indication that one might not be able to expect significantly better predictive power than that shown by our approach.

Average *human rating* over all words and users given to the final controller was 7.4 out of 10. This means that overall, the participants felt the system succeeded in providing a controller that lets the user manipulate the sound in terms of the descriptive words.

Mean correlation coefficients between user ratings and each of the five control measures (Section 3.2) used to generate the audio examples are shown in Tables 1 and 2. Table 1 shows values for the 10 participants with little or no musical background. Table 2 shows these values for the 12 participants with strong musical background.

	bright	*clear*	*boomy*	*bath'*	*church*
training time	4'07"	3'34"	4'35"	4'29"	4'57"
Reverb. Time	-0.02	0.19	-0.07	0.02	0.08
Echo Density	-0.01	-0.10	0.13	0.06	0.03
Clarity	**0.39**	**0.33**	0.08	0.06	0.16
Central Time	-0.45	-0.51	**0.28**	-0.22	**0.52**
Spec. Centroid	**0.36**	**0.38**	-0.23	0.08	0.02

Table 1. Average training time and correlation coefficients of the measures for the five descriptive words over the participants with little or no musical background.

	bright	*clear*	*boomy*	*bath'*	*church*
training time	3'14"	3'42"	4'17"	4'05"	3'36"
Reverb. Time	0.03	0.21	-0.04	-0.04	0.06
Echo Density	-0.02	-0.06	0.06	0.01	0.02
Clarity	0.29	**0.44**	0.14	-0.08	-0.03
Central Time	-0.17	-0.57	**0.46**	0.06	**0.70**
Spec. Centroid	**0.42**	0.29	-0.21	0.13	-0.03

Table 2. Average training time and correlation coefficients of the measures for the five descriptive words over the participants with strong musical background.

As we can see, participants with strong musical background completed the training more quickly. Both groups showed similar results for the *user consistency*, the *system predictiveness* and the *human ratings*. However there are relevant differences between these two groups in which signal measures most affect ratings of examples.

For both groups, *bright* and *clear* show overall a correlation with the *Clarity* and the *Spectral Centroid*, and a negative correlation with the *Central Time*. Table 1 indicates that participants with little or no musical background may have confounded *bright* and *clear*, while they seem distinct to people with strong musical background. Indeed, we should expect *bright* to be more correlated with the *Spectral Centroid* and *clear* with the *Clarity*, as shown in Table 2 by participants with strong musical background.

That said, *user consistency*, *system predictiveness* and *human ratings* are reasonably high on these words for both groups, even though the definitions of these words clearly vary between groups. These results indicate people with little musical experience can still define these terms with enough consistency for the system to model their preferences and provide a useful controller.

Boomy shows a significant correlation with the *Central Time* (in bold) and a negative correlation with the *Spectral Centroid*. Participants with strong musical background showed higher correlation with the *Central Time*. Furthermore, the distribution of the correlation coefficients of the measures for participants with strong musical background has a smaller standard deviation, which means that they showed a common understanding of the concept, while the standard deviation for participants with a little or no musical background is higher, especially for the *Central Time* and the *Spectral Centroid*, which means that the definition of the concept varied more greatly between them.

Table 3 highlights how well the system performs on a descriptive word where there was substantial disagreement between individuals. The table compares the correlation coefficients of the measures, the *system predictiveness* correlation coefficients, and the *human ratings* between four participants with little or no musical background for the descriptive word "boomy".

boomy	*user 11*	*user 12*	*user 13*	*user 22*
Reverb. Time	0.01	-0.04	-0.10	-0.18
Echo Density	**0.26**	-0.08	0.24	0.01
Clarity	-0.43	**0.10**	0.14	**0.36**
Central Time	-0.33	-0.17	**0.69**	*-0.32*
Spec. Centroid	*-0.74*	*-0.58*	*-0.15*	0.17
predictiveness	0.90	0.77	0.86	0.79
human ratings	7.0	10.0	8.0	8.0

Table 3. Comparison of the results between four participants with little or no musical background for *boomy* (the highest correlation coefficient is in bold and the highest negative correlation coefficient is in italic, for each user).

We can see that the correlation coefficients of the audio measures are very different from one participant to another, and yet, the *system predictiveness* and the *human ratings* are high. Again, this indicates our approach worked well to personalize a controller for each of these individuals, despite the variation in their personal definition of *boomy*.

Participants showed great variation in their responses to *bathroom-like* and distributions of the correlation coefficients between acoustic measures and user ratings show high standard deviation, especially for the *Clarity*, the *Central Time* and the *Spectral Centroid*. Table 4 compares the results for *bathroom-like* between four different participants: users 03 and 08 have a strong musical background, and users 12 and 13 have little or no musical background. Correlation coefficients of the measures are very different between participants, yet the *system predictiveness* and the *human ratings* are high.

Church-like shows overall a high correlation with the *Central Time* (in bold), especially for participants with a strong musical background. The distribution of the correlation coefficients of the measures show also significant standard deviation, especially for participants with little or no

bathroom-like	*user 03*	*user 08*	*user 12*	*user 13*
Reverb. Time	-0.14	0.13	0.05	0.07
Echo Density	0.10	-0.09	0.21	0.17
Clarity	-0.02	0.12	0.25	*-0.63*
Central Time	**0.78**	*-0.44*	0.01	**0.74**
Spec. Centroid	*-0.27*	**0.47**	**0.60**	-0.09
predictiveness	0.83	0.77	0.81	0.93
human ratings	7.0	8.0	10.0	7.0

Table 4. Comparison of the correlation coefficients of the measures between four different users for *bathroom-like*.

musical background. The same conclusions can be drawn here: participants have their own way of understanding the concept, and overall the system succeeds in grasping it to build a controller which meets participants' expectations.

Overall, *clear* shows the best mean results across all users: *user consistency*, 0.73, *system predictiveness*, 0.85, and *human rating*, 8.5. Overall, *bathroom-like* shows the worst results: *user consistency*, 0.62, *system predictiveness*, 0.62, and *human rating*, 6.8. Fig. 6 shows the distributions over all the participants of the *user consistency* and *system predictiveness* correlation coefficients, and the *human ratings* for *clear* and *bathroom-like*.

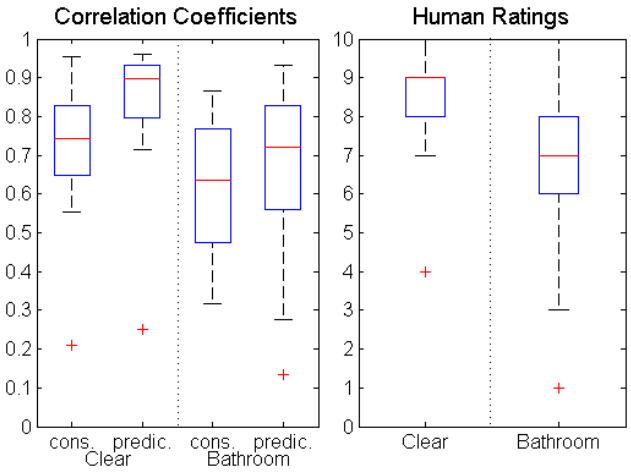

Figure 6. Left boxplot: distributions of *user consistency* and *system predictiveness* correlation coefficients for the best performing word: *clear* (left) and the worst performing word: *bathroom* (right) ; right boxplot: distributions of *human ratings* for *clear* (left) and *bathroom* (right).

5. CONCLUSION

A method for mapping subjective terms onto perceptual audio measures useful for digital reverberation control has been presented. This lets us build a simple controller to manipulate sound in terms of a subjective audio concept, bypassing the bottleneck of complex interfaces and individual differences in descriptive terms. The evaluation of our system showed that audio descriptors can be effectively and rapidly learned and controlled with this method.

Our study showed that people have different definitions of the same descriptor, and yet our system succeeds in learning an individual's concept so that people are satisfied with the final controller. This supports our contention that individualizing controllers is a useful approach.

There are a number of directions we expect to take in this work. We wish to conduct a more grounded psychoacoustic study to determine meaningful ranges for the set of reverberation measures. Finally, joint learning of controls for multiple audio effects (reverberation and equalization, for example) can be considered, to span a wider range of possible manipulations of sound. This work was supported by NSF grant number IIS-0757544.

6. REFERENCES

[1] Mihir Sarkar, Barry Vercoe, and Yang Yang. "Words that Describe Timbre, A Study of Auditory Perception Through Language," *Language and Music as Cognitive Systems Conference*, Cambridge, UK, May 2007.

[2] Alastair C. Disley and David M. Howard. "Spectral correlates of timbral semantics relating to the pipe organ," *Joint Baltic-Nordic Acoustics Meeting*, Mariehamn, Aland, Finland, 8-10 June 2004.

[3] Victor Alvarez-Cortes, Benjamin E. Zayas-Perez, Victor Huga Zarate-Silva, and Jorge A. Ramirez Uresti. "Current Trends in Adaptive User Interfaces: Challenges and Applications," *Electronics, Robotics and Automotive Mechanics Conference*, pp. 312-317, 2007.

[4] Andrew T. Sabin and Bryan Pardo. "Rapid learning of subjective preference in equalization," *125th Audio Engineering Society Convention*, San Francisco, CA, USA, 2-5 October 2008.

[5] Carl R. Nave. *HyperPhysics*, Georgia State University, Atlanta, GA, USA, 2006, http://hyperphysics.phy-astr.gsu.edu/hbase/hph.html.

[6] Pavel Zahorik. "Perceptual Scaling of Room Reverberation," *Journal of the Acoustical Society of America*, Vol. 15, No. B, pp. 2598-2598, 2001.

[7] Manfred R. Schroeder and Benjamin F. Logan. "'Colorless' Artificial Reverberation," *Journal of the Audio Engineering Society*, Vol. 9, No. 3, July 1961.

[8] James A. Moorer. "About This Reverberation Business," *Computer Music Journal*, July 1979.

[9] Fernando A. Beltrán, José R. Beltrán, Nicolas Holzem, and Adrian Gogu. "Matlab Implementation of Reverberation Algorithms," *Journal of New Music Research*, Vol. 31, No. 2, pp. 153-161, June 2002.

[10] Zafar Rafii and Bryan Pardo. "A Digital Reverberator controlled through Measures of the Reverberation," Northwestern University, EECS Department, Technical Report NWU-EECS-09-08, 2009.

[11] Fons Adriaensen. "Acoustical Impulse Response Measurement with ALIKI," *4th International Linux Audio Conference*, Karlsruhe, Germany, 27-30 April 2006.