

Investigating How Convolution Reverberation Effects the Space of Instrument Emotional Characteristics

Ronald K. Mo

Department of Computer Science
and Engineering, HKUST
ronmo@cse.ust.hk

Richard H. Y. So

Department of Industrial Engineering
and Decision Analytics, HKUST
rhyso@ust.hk

Andrew Horner

Department of Computer Science
and Engineering, HKUST
horner@cse.ust.hk

ABSTRACT

Even though emotional characteristics in musical instruments can be significantly changed with reverberation [1, 2, 3, 4, 5, 6, 7, 8], it has been shown that these changes in character are relatively uniform and instrument independent for parametric artificial reverberation (artificial reverberation) [9, 10]. This paper investigates whether they are also relatively uniform for convolution reverberation. We compared eight sustained instrument tones with different hall impulse responses over ten emotional characteristics. The results show a similarly remarkable consistency in listener rankings, with strong correlations ranging from 90 to 96% —the same range found for artificial reverberation. These results indicate that the underlying instrument space does not change much with convolution reverberation in terms of emotional characteristics, and that each instrument has a particular footprint of emotional characteristics. In fact, the footprints are basically the same for both reverberation, and our results show a correlation of 88% between them.

1. INTRODUCTION

Researchers have considered various relationships between timbre (sound color) and music emotion, and have found that different instruments have different timbral and emotional characteristics. By changing the pitch and dynamics, the timbre and emotional characteristics also change. These characteristics are further modified by the performance environment —by the amount and length of reverberation in the space, which smears the temporal and spectral envelopes and changes the emotional character of the sound. The same idea holds when artificial reverberation is added as a post-process. Readers may refer to [3] for a summarization.

Our recent work has shown that the emotional characteristics of musical instruments are significantly changed with artificial reverberation [3, 6]. We found that reverberation length and amount had a strongly significant effect on the emotional characteristics Mysterious and Romantic and a medium effect on Heroic, Sad, and Scary. Reverberation had a mild effect on Happy and relatively little effect on Shy. Interestingly, anechoic tones were judged most

Comic.

We followed up our initial artificial reverberation study with a corresponding study on convolution reverberation [7]. We compared musical instruments over five hall impulse responses ranging from the 1-second Royal National Theatre to the 5-second King’s College Chapel. The results showed that convolution reverberation had more pronounced effects on the emotional characteristics compared to artificial reverberation. While the results were more pronounced for convolution reverberation compared to artificial reverberation, there was also a strikingly strong agreement in their results, and correlation coefficient between them was 0.74 over all emotional characteristics. This strong correlation indicates that reverberation time has a remarkably consistent effect on the emotional characteristics regardless of whether using convolution or artificial reverberation.

We continued our study of artificial reverberation with an investigation of whether the changes in emotional characteristics are relatively uniform or instrument-dependent [10, 9]. We wanted to know whether artificial reverberation changes the underlying instrument space. We found that the changes in the emotional characteristics were remarkably consistent among the instruments with different reverberation amounts and lengths. In particular, the instruments clustered into two fairly distinct groups: those where the high-arousal emotional characteristics were strong (e.g., oboe, trumpet, and violin), and those where the low-arousal characteristics were strong (e.g., bassoon, clarinet, flute, and horn). The saxophone was an outlier, and was uniquely somewhat strong for most emotional characteristics.

This uniformity of changes in the emotional characteristics was in fact contrasting to our initial artificial reverberation study [3, 6]. The strong distinct changes found in our initial study led us to expect some instrument-dependencies in the follow-up study, which used exactly the same tones. Nevertheless, there is no contradiction in their results: artificial reverberation distinctly changes the character of the sound, but does so in a uniform way across the instruments.

In any case, it would be useful to have the results for both convolution and artificial reverberation, since they are both widely used for acoustic musical instruments, and yet they are fundamentally different.

Therefore, the main goal of our current study is to investigate whether changes in emotional characteristics are relatively uniform or instrument dependent with convolution reverberation. The results were amazingly uniform in our previous study of artificial reverberation, and we want to know if they are basically uniform for convolution re-

Copyright: ©2018 Ronald K. Mo et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

verberation as well. Moreover, the question whether reverberation is uniform or instrument-dependent sheds light on other aspects of music emotion and timbre. Certainly each instrument has a distinct timbre. The spectral and temporal envelopes of the a particular instrument are different depending on the room reverberation due to smearing, but the instrument identity remains unchanged. In light of this, are there distinctive emotional characteristics that identify each instrument? If so, it helps explain why performers can practice in small rehearsal rooms and reasonably predict the emotional blends and balances between the instruments even when the final performance is in a large concert hall. If not, then in each performance environment the instruments will assume different characters, which helps explain their rich versatility. Either way, the results will deepen our understanding of these issues.

2. METHODOLOGY

For this investigation, we conducted a listening test to investigate whether convolution reverberation changes the emotional characteristics of instruments uniformly or in an instrument-dependent way. To easily compare and contrast the convolution and artificial reverberation results, we conducted the listening test for convolution reverberation in the same way as we did for artificial reverberation [9, 6]. Listeners compared the instruments pairwise for each hall and emotional characteristic. Below are some the main points, especially the differences from the artificial reverberation test.

We tested sustained musical instruments representing the wind and bowed string families obtained from the University of Iowa Musical Instrument Samples [11]. They included the bassoon (bs), clarinet (cl), flute (fl), horn (hn), oboe (ob), saxophone (sx), trumpet (tp), and violin (vn). These sounds were all recorded in an anechoic chamber, and were thus free from reverberation. These instruments were nearly harmonic and had fundamental frequencies close to Eb4 (311.1 Hz).

In addition to the anechoic sounds, we compared sounds that had been processed with a convolution reverberation with reverberation lengths of approximately 1s and 2s, which according to Hidaka and Beranek [12] and Beranek [13] typically correspond to small and large concert halls. To do this, we selected several representative hall convolution reverberations based on the impulse responses in Altiverb [14]. We measured their reverberation lengths based on their reverberation time RT_{60} , and picked those that most closely matched the reverberation times we tested in our previous study of artificial reverberation. We also included the cathedral impulse response of King’s College Chapel with a 5.44 second reverberation time to determine the effects for a more extreme case.

For this study we tested the ten emotional categories: Angry, Calm, Comic, Happy, Heroic, Mysterious, Romantic, Sad, Scary, and Shy. Some choices of emotional characteristics are fairly universal and occur in many previous studies roughly corresponding to the four quadrants of the Valence-Arousal plane (e.g., Happy, Sad, Scary/Fear/Angry, Tender/Calm/Romantic), but there are lots of variations beyond that [15]. For this study, we used the same categories that we used in our previous research on musical in-

struments and reverberation. Just before the listening test, subjects read online definitions of the emotional categories used in our experiment, which were taken from the Cambridge Academic Content Dictionary[16].

To limit the length of the test and to minimize listener fatigue, we divided listeners into two groups, where each group heard a different set of five emotional categories. We hired 62 subjects to take the listening test, and since they each heard half the emotional categories, 31 subjects compared each emotional category. All subjects were fluent in English. They were all undergraduate students at the Hong Kong University of Science and Technology where all courses are taught in English. Subjects were not musical experts (e.g., recording engineers, professional musicians, or music conservatory students) but average attentive listeners. Among the 62 subjects, there were 38 males and 24 females. The subjects ranged in age from 18 to 23. In terms of musical experience, 31 subjects had some experience playing an instrument (an average of 5.8 years), and 31 subjects did not have experience playing an instrument. In recruiting the subjects, all indicated they had no known hearing problems.

In the listening tests, subjects heard paired comparisons between the instruments for the each hall and emotional category. During each trial, subjects heard a pair of sounds and were prompted to choose which more strongly aroused a given emotional category. Subjects did not need to remember all of the tones, just the two in each comparison. Each combination of two different instruments were presented for each of the eight instruments over the six halls and ten emotional categories, and the complete listening test totalled $C_2^8 \times 6 \times 10 = 1680$ trials (840 trials per listener since we divided the task into two groups). For each listener, the overall trial presentation order was randomized to average out effects due to learning or fatigue. For the two sounds A and B, they heard AB where the order of A and B was random for each comparison. The listening test took about 95 minutes, with forced short breaks about every 30 minutes.

3. RESULTS

For our listening test, listeners compared each pair of instruments for each hall and emotional characteristic. Based on the listening test data of 31 subjects for each group, we derived scale values using the Bradley-Terry-Luce (BTL) statistical model [17, 18]. The BTL value for each instrument is the probability that listeners will choose that instrument when considering a certain hall and emotion category. For example, if all eight instruments were judged equally Happy, the BTL scale values would be $1/8 = 0.125$. Figures 1 to 4 show the BTL scale values and the corresponding 95% confidence intervals for each emotional category.

Though there are differences between the individual graphs in each of Figures 1 - 4, there are striking similarities. For example, with Angry in Figure 1, the trumpet is usually first, the oboe second, and clarinet last. Extending this idea, we observe that the trumpet or oboe were ranked highest for nearly all halls in the high-arousal categories (e.g., Angry and Scary). The clarinet and horn were ranked highest for the low-arousal categories (e.g., Calm and Ro-

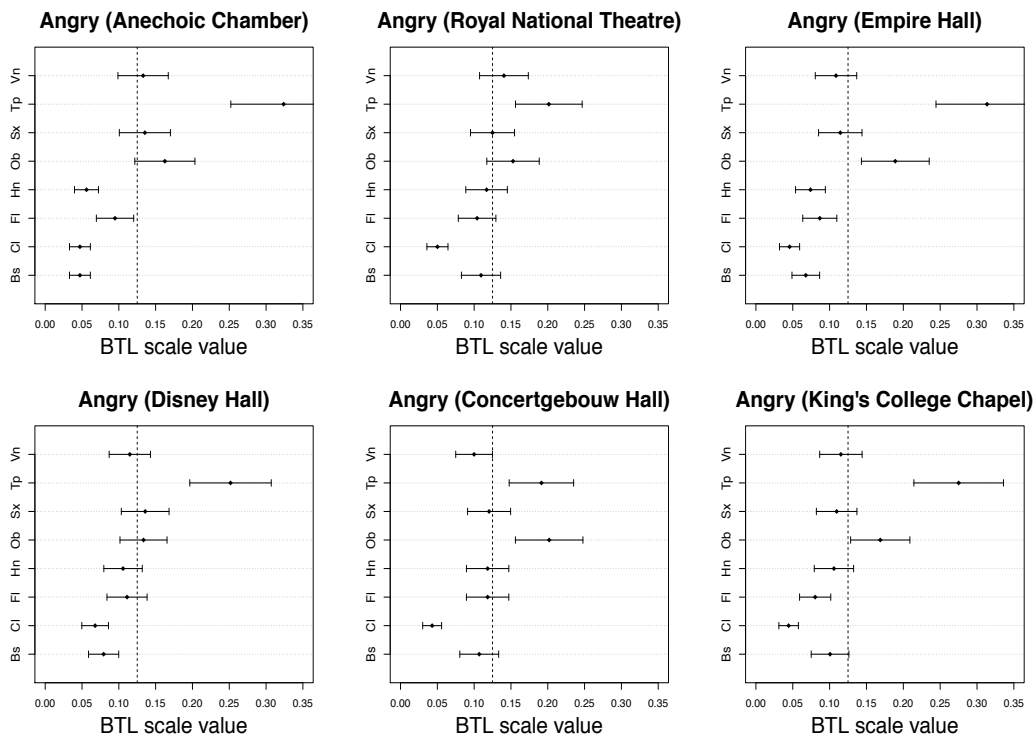


Figure 1. BTL scale values and the corresponding 95% confidence intervals with different halls for the category Angry.

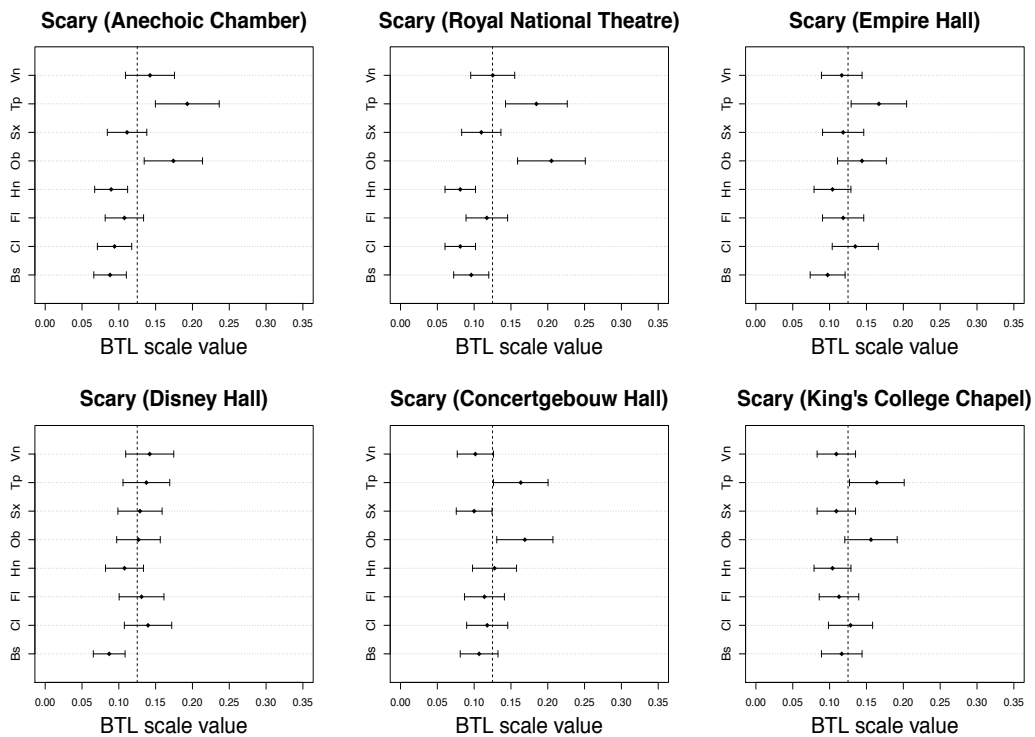


Figure 2. BTL scale values and the corresponding 95% confidence intervals with different halls for Scary.

matic). Scary consistently had the narrowest range among all halls.

We wanted to determine the number of times each instrument was significantly greater than the other seven instruments for each hall and emotional characteristic. Therefore, we used Paired t-tests and Pearson correlation to an-

alyze our data. Table 1 gives a summary of our analysis. It shows the number of times each instrument was significantly greater than the other seven instruments over all six halls for each emotional characteristic. The maximum possible value is 42 and the minimum possible value is 0. For example, for Angry the trumpet was statistically

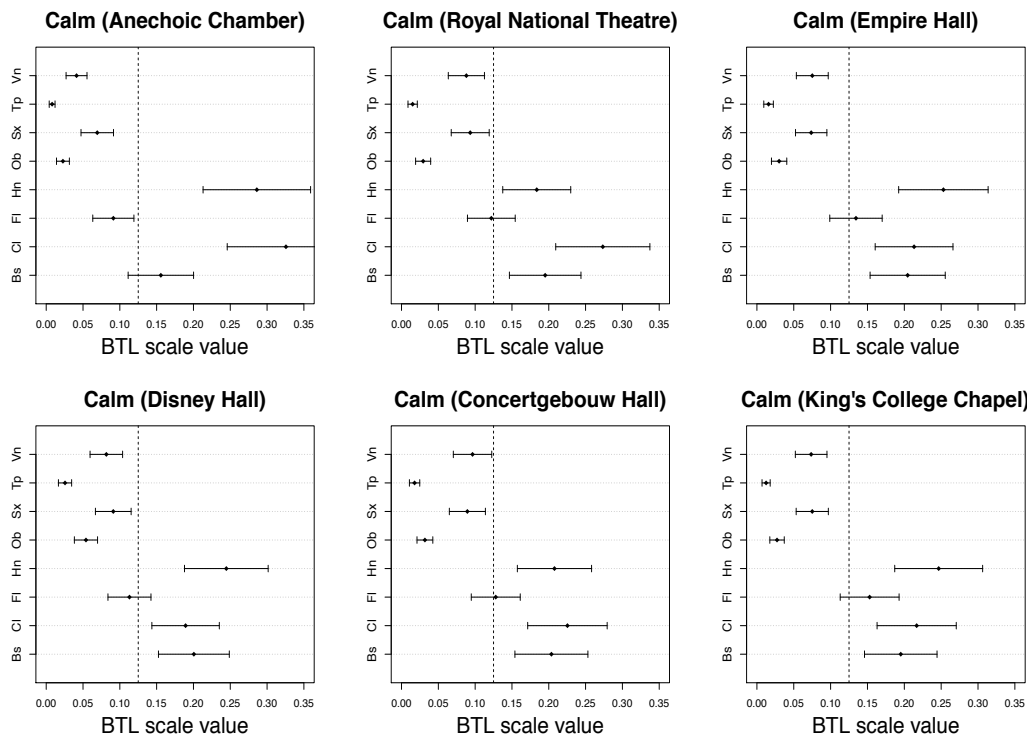


Figure 3. BTL scale values and the corresponding 95% confidence intervals with different halls for Calm.

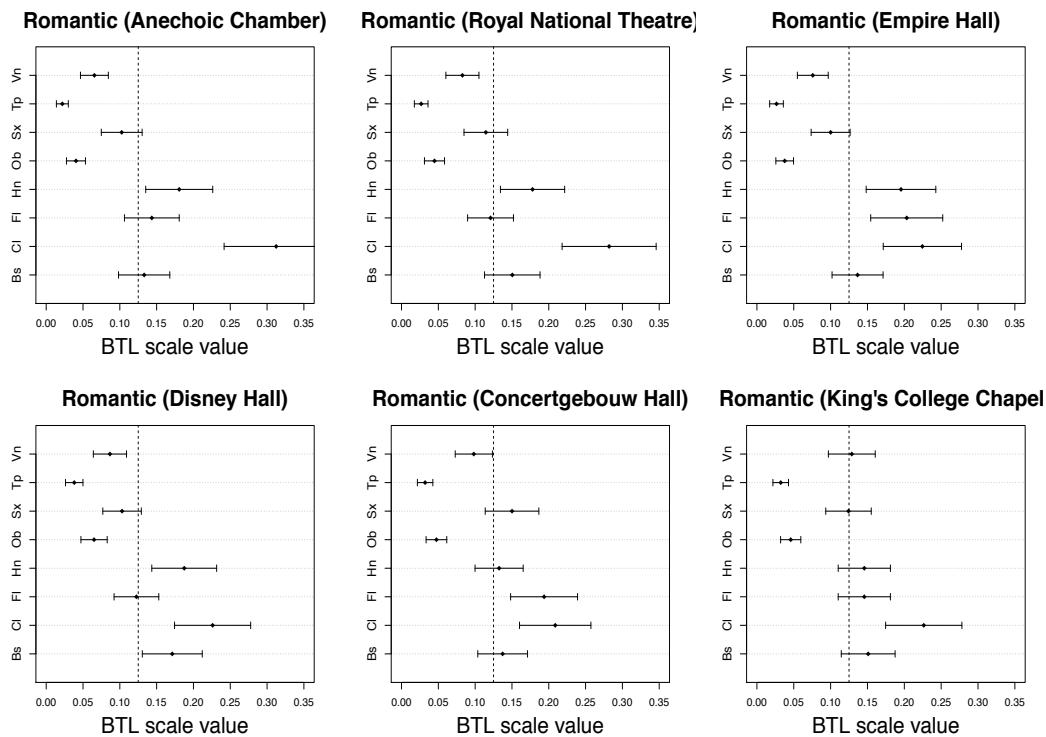


Figure 4. BTL scale values and the corresponding 95% confidence intervals with different halls for Romantic.

significantly greater than all the other seven instruments most of the time, and its value is 34. Table 1 makes it obvious that the trumpet was ranked the highest for Angry, Comic, Happy, and Heroic, while it was a close second to the oboe for Scary. The clarinet was ranked highest for Calm, Mysterious, and Romantic, while the horn

was highest for Sad, and they were tied for Shy. In general, the trumpet, oboe, saxophone, and violin were strong for the high-arousal characteristics, and the clarinet, horn, bassoon, and flute were strong for low-arousal characteristics. These two groups represent the high-and low-arousal timbres, though Table 1 also indicates that there is some

	Bs	Cl	Fl	Hn	Ob	Sx	Tp	Vn
Happy	4	3	11	0	28	16	34	18
Heroic	5	0	4	1	15	11	34	10
Comic	3	2	12	0	30	20	37	22
Sad	24	27	21	34	3	9	0	9
Scary	0	2	3	0	12	0	10	1
Shy	25	33	21	33	5	12	0	11
Romantic	17	32	17	19	3	15	0	11
Mysterious	15	31	18	30	3	8	0	5
Angry	3	0	7	3	20	10	34	8
Calm	29	30	17	29	6	13	0	11

Table 1. How often each instrument was statistically significantly greater than the others over the six halls. The maximum possible value is 42 and the minimum possible value is 0. The maximum for each emotional characteristic is shown in bold.

Reverb Types	Pearson Correlation
ANE & RNT	0.950
ANE & EMP	0.949
ANE & DNH	0.937
ANE & CGH	0.929
ANE & KCC	0.933
RNT & EMP	0.936
RNT & DNH	0.917
RNT & CGH	0.944
RNT & KCC	0.944
EMP & DNH	0.935
EMP & CGH	0.948
EMP & KCC	0.962
DNH & CGH	0.915
DNH & KCC	0.919
CGH & KCC	0.962

Table 2. Pearson correlation between the different halls based on the listener voting data.

crossover as well.

We also wanted to determine how similar were results for the different halls. Therefore, we ran correlations for the voting data (i.e., the number of positive votes received by each instrument for emotional category and hall). In all cases, the correlations were statistically significant (at the $p < 0.0001$ level) and very strong, ranging from 90 to 96%, indicating a near-linear relationship and a very high level of agreement. In particular, Table 2 shows Pearson correlation between the different halls based on the voting data.

4. CONCLUSION

This section discusses the results for our experiment on convolution reverberation. It compares contrasts these results with those found in our previous study of artificial reverberation [10, 9]. It also considers the implications of the results in wider applications and research.

4.1 Comparison of Results for Artificial and Convolution Reverberation

We were curious to find out to what degree the results for artificial and convolution reverberation were in agreement. To answer this question, we correlated the BTL rankings for our artificial and convolution reverberation experiments. Eight emotional categories were tested in both experiments, and five reverberation times were tested in both experiments. All the correlations were statistically significant (at the $p < 0.0001$ level), and seven out of eight categories were near 80% or more, indicating a very high level of agreement. The overall correlation was also very strong at about 85%. We also correlated the number of significant differences from both experiments. Once again, Scary was not significantly correlated by this measure, but the other categories were near 90%. This indicates a very strong agreement in the patterns of significant differences, meaning the instrument footprints of emotional characteristics strongly match in artificial and convolution reverberation, except for Scary.

4.2 Implications of the Results

The results in this paper and in our previous study of artificial reverberation [10, 9] have shown that emotional characteristics are changed uniformly by artificial reverberation in about the same way for different instruments. For example, added reverberation might bring out certain characteristics, but in a uniform way for the instruments, and not some instruments more than others.

At this point, one may wonder whether there are other halls that we did not test where the results might be instrument-dependent. It is possible. Perhaps in cathedrals with long reverberation lengths of 10 seconds or more the excessive smearing of the temporal and spectral envelopes effects instruments with strong isolated resonances more than other instruments with smoother spectral envelopes. More generally, perhaps there are non-music halls (e.g., stadiums, arenas, lecture theaters) where undesirable acoustic features (e.g., standing waves) result in instrument dependencies in the relative emotional characteristics. Perhaps the instrument-independent behavior of concert halls is what helps distinguish a good music venue from a poor one. Further work will be needed to confirm this, but it was remarkable how uniform the instruments were within the concert halls and anechoic chamber we tested. In any case, it will be interesting to see whether these same overall results hold for other instruments, pitches, and dynamics, as we only tested Eb4-*forte* tones for eight instruments. It will also be interesting to see whether these results hold for other types of reverberation such as plate reverberation.

More broadly, perhaps the relatively consistent ranking of emotional characteristics between the instruments is what helps us to identify each instrument regardless of room reverberation. The columns of Table 1 represent the overall footprints of the emotional characteristics for each instrument for our Eb4-*forte* tones. The instruments clustered into two fairly distinct groups: those where the positive energetic emotional characteristics were strong (e.g., trumpet, oboe, saxophone, violin), and those where the low-arousal characteristics were strong (e.g., clarinet, horn, bassoon, flute). Looking in more detail, the trumpet, oboe,

saxophone, and violin had similar deep footprints, but the trumpet's footprints were deepest, the oboe second deepest, and saxophone and violin similar but shallower. In the same way, the clarinet and horn had similar footprints, and the bassoon and flute were shallower. For Scary, the oboe and trumpet were uniquely strong. The saxophone and violin had the most even distribution, with medium values for most emotional categories.

This work also has implications for music emotion research of single musical instrument tones, where most studies do not explicitly state whether the tones are anechoic or with light reverberation, and assume it does not matter too much. The results of this study suggest that this is a somewhat safe assumption if the relative emotional characteristics between instruments are the main consideration. Of course, if changes in timbre or absolute emotional characteristics are the main consideration, reverberation can indeed make a difference. In any case, it is useful to know which situations are relatively safe and which can be problematic.

5. REFERENCES

- [1] D. Västfjäll, P. Larsson, and M. Kleiner, "Emotion and auditory virtual environments: affect-based judgments of music reproduced with virtual reverberation times," *CyberPsychology & Behavior*, vol. 5, no. 1, pp. 19–32, 2002.
- [2] A. Tajadura-Jiménez, P. Larsson, A. Väljamäe, D. Västfjäll, and M. Kleiner, "When room size matters: acoustic influences on emotional responses to sounds," *Emotion*, vol. 10, no. 3, pp. 416–422, 2010.
- [3] R. Mo, B. Wu, and A. Horner, "The Effects of Reverberation on the Emotional Characteristics of Musical Instruments," *J. Audio Eng. Soc.*, vol. 63, no. 12, pp. 966–979, 2016. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=18055>
- [4] R. Mo, G. L. Choi, C. Lee, and A. Horner, "The Effects of MP3 Compression on Perceived Emotional Characteristics in Musical Instruments," *Journal of the Audio Engineering Society*, vol. 64, no. 11, pp. 858–867, 2016.
- [5] —, "The Effects of MP3 Compression on Emotional Characteristics," in *42nd International Computer Music Conference, Utrecht, The Netherlands*, 2016, p. 411.
- [6] R. Mo, B. Wu, and A. Horner, "The Effects of Reverberation Time and Amount on the Emotional Characteristics," in *42nd International Computer Music Conference, Utrecht, The Netherlands*, 2016, p. 12.
- [7] R. K. Mo and A. Horner, "The Effects of Convolution Reverberation on the Emotional Characteristics," in *43rd International Computer Music Conference, Shanghai, China*, 2017.
- [8] D. Williams, "On the Affective Potential of the Recorded Voice," *JOURNAL OF THE AUDIO ENGINEERING SOCIETY*, vol. 64, no. 6, pp. 429–437, 2016.
- [9] R. K. Mo, R. H. So, and A. Horner, "How Does Parametric Reverberation Change the Space of Instrument Emotional Characteristics?" in *43rd International Computer Music Conference, Shanghai, China*, 2017.
- [10] R. Mo, R. H. So, and A. Horner, "An Investigation into How Reverberation Effects the Space of Instrument Emotional Characteristics," *Journal of the Audio Engineering Society*, vol. 64, no. 12, pp. 988–1002, 2016.
- [11] "University of Iowa Musical Instrument Samples," *University of Iowa*, 2004, <http://theremin.music.uiowa.edu/MIS.html>.
- [12] T. Hidaka and L. L. Beranek, "Objective and subjective evaluations of twenty-three opera houses in Europe, Japan, and the Americas," *The Journal of the Acoustical Society of America*, vol. 107, no. 1, pp. 368–383, 2000.
- [13] L. Beranek, *Concert halls and opera houses: music, acoustics, and architecture*. Springer Science & Business Media, 2004.
- [14] "Altiverb 7," *Audio Ease*, 2016, <https://www.audioease.com/altiverb/>.
- [15] P. N. Juslin and J. Sloboda, *Handbook of music and emotion: Theory, research, applications*. Oxford University Press, 1993.
- [16] Cambridge Academic Content Dictionary. [Online]. Available: <http://dictionary.cambridge.org/dictionary/american-english>
- [17] R. A. Bradley, "14 Paired comparisons: Some basic procedures and examples," *Nonparametric Methods*, vol. 4, pp. 299–326, 1984.
- [18] F. Wickelmaier and C. Schmid, "A Matlab Function to Estimate Choice Model Parameters from Paired-comparison Data," *Behavior Research Methods, Instruments, and Computers*, vol. 36, no. 1, pp. 29–40, 2004.