Musical Timbre and Emotion: The Identification of Salient Timbral Features in Sustained Musical Instrument Tones Equalized in Attack Time and Spectral Centroid

Bin Wu[1], **Andrew Horner**[1], **Chung Lee**[2]

[1]Department of Computer Science and Engineering,
Hong Kong University of Science and Technology

[2]The Information Systems Technology and Design Pillar,
Singapore University of Technology and Design

{bwuaa,horner}@cse.ust.hk, im.lee.chung@gmail.com

ABSTRACT

Timbre and emotion are two of the most important aspects of musical sounds. Both are complex and multidimensional, and strongly interrelated. Previous research has identified many different timbral attributes, and shown that spectral centroid and attack time are the two most important dimensions of timbre. However, a consensus has not emerged about other dimensions. This study will attempt to identify the most perceptually relevant timbral attributes after spectral centroid and attack time. To do this, we will consider various sustained musical instrument tones where spectral centroid and attack time have been equalized. While most previous timbre studies have used discrimination and dissimilarity tests to understand timbre, researchers have begun using emotion tests recently. Previous studies have shown that attack and spectral centroid play an essential role in emotion perception, and they can be so strong that listeners do not notice other spectral features very much. Therefore, in this paper, to isolate the third most important timbre feature, we designed a subjective listening test using emotion responses for tones equalized in attack, decay, and spectral centroid. The results showed that the even/odd harmonic ratio is the most salient timbral feature after attack time and spectral centroid.

1. INTRODUCTION

Timbre is one of the most important aspects of musical sounds, yet it is also the least understood. It is often simply defined by what it is not: not pitch, not loudness, and not duration. For example, if a trumpet and clarinet both played A440Hz tones for 1s at the same loudness level, timbre is what would distinguish the two sounds. Timbre is known to be multidimensional, with attributes such as attack time, decay time, spectral centroid (i.e., brightness), and spectral irregularity to name a few.

Several previous timbre perception studies have shown

spectral centroid and attack time to be highly correlated with the two principal perceptual dimensions of timbre. Spectral centroid has been shown to be strongly correlated with one of the most prominent dimensions of timbre as derived by multidimensional scaling (MDS) experiments [1, 2, 3, 4, 5, 6, 7, 8].

Grey and Gordon [1, 9] derived three dimensions corresponding to spectral energy distribution, temporal synchronicity in the rise and decay of upper harmonics, and spectral fluctuation in the signal envelope. Iverson and Krumhansl [4] found spectral centroid and critical dynamic cues throughout the sound duration to be the salient dimensions. Krimphoff [10] found three dimensional correlates: (1) spectral centroid, (2) rise time, and (3) spectral flux corresponding to the standard deviation of the time-averaged spectral envelopes. More recently, Caclin *et al.* [8] found attack time, spectral centroid, and spectrum fine structure to be the major determinates of timbre through dissimilarity rating experiments. Spectral flux was found to be a less salient timbral attribute in this case.

While most researchers agree spectral centroid and attack time are the two most important timbral dimensions, no consensus has emerged about the best physical correlate for a third dimension of timbre. Lakatos and Beauchamp [7, 11, 12] suggested that if additional timbre dimensions exist, one strategy would be to first create stimuli with identical pitch, loudness, duration, spectral centroid, and rise time, but which are otherwise perceptually dissimilar. Then, potentially multidimensional scaling of listener dissimilarity data can reveal additional perceptual dimensions with strong correlations to particular physical measures. Following up this suggestion is the main focus of this paper.

While most previous timbre studies have used discrimination and dissimilarity to understand timbre, researchers have recently begun using emotion. Some previous studies have shown that emotion is closely related to timbre. Scherer and Oshinsky found that timbre is a salient factor in the rating of synthetic tones [13]. Peretz *et al.* showed that timbre speeds up discrimination of emotion categories [14]. Bigand *et al.* reported similar results in their study of emotion similarities between one-second musical excerpts [15]. It was also found that timbre is essential to musical genre recognition and discrimination [16, 17, 18]. Eerola

[19] carried out listening tests to investigate the correlation of emotion with temporal and spectral sound features. The study confirmed strong correlations between features such as attack time and brightness and the emotion dimensions valence and arousal for one-second isolated instrument tones. Valence and arousal are measures of how positive and energetic the music sounds [20]. Despite the widespread use of valence and arousal in music research, composers may find them rather vague and difficult to interpret for composition and arrangement, and limited in emotional nuance. Using a different approach than Eerola, Ellermeier *et al.* investigated the unpleasantness of environmental sounds using paired comparisons [21]. Emotion categories have been shown to be generally congruent with valence and arousal in music emotion research [22].

In our own previous study on emotion and timbre [23], to make the results intuitive and detailed for composers, listening test subjects compared tones in terms of emotion categories such as Happy and Sad. We also equalized the stimuli attacks and decays so that temporal features would not be factors. This modification allowed us to isolate the effects of spectral features such as spectral centroid. Average spectral centroid significantly correlated for all emotions, and a bigger surprise was that spectral centroid deviation significantly correlated for all emotions. This correlation was even stronger than average spectral centroid for most emotions. The only other correlation was spectral incoherence for two emotions.

Since average spectral centroid and spectral centroid deviation were so strong, listeners did not notice other spectral features much. This made us wonder: if we equalized average spectral centroid in the tones, would spectral incoherence be more significant? Would other spectral characteristics emerge as significant? To answer these questions, we conducted the follow-up experiment described in this paper using emotion responses for tones equalized in attack, decay, and spectral centroid.

2. LISTENING TEST

In our listening test, listeners compared pairs of eight instruments for eight emotions, using the tones that were equalized for attack, decay, and spectral centroid.

2.1 Stimuli

2.1.1 Prototype instrument sounds

The stimuli consisted of eight sustained wind and bowed string instrument tones: bassoon (Bs), clarinet (Cl), flute (Fl), horn (Hn), oboe (Ob), saxophone (Sx), trumpet (Tp), and violin (Vn). They were obtained from the McGill and Prosonus sample libraries, except for the trumpet, which had been recorded at the University of Illinois at Urbana-Champaign School of Music. All the tones were used in a discrimination test carried out by Horner *et al.* [24], six of them were also used by McAdams *et al.* [25], and all of them used our previous emotion-timbre test [23].

The tones were presented in their entirety. The tones were nearly harmonic and had fundamental frequencies close to

311.1 Hz (Eb4). The original fundamental frequencies deviated by up to 1 Hz (6 cents), and were synthesized by additive synthesis at 311.1 Hz.

Since loudness is potential factor in emotion, amplitude multipliers were determined by the Moore-Glasberg loudness program [26] to equalize loudness. Starting from a value of 1.0, an iterative procedure adjusted an amplitude multiplier until a standard loudness of 87.3 ± 0.1 phons was achieved.

2.2 Stimuli Analysis and Synthesis

2.2.1 Spectral Analysis Method

Instrument tones were analyzed using a phase-vocoder algorithm, which is different from most in that bin frequencies are aligned with the signal's harmonics (to obtain accurate harmonic amplitudes and optimize time resolution) [27]. The analysis method yields frequency deviations between harmonics of the analysis frequency and the corresponding frequencies of the input signal. The deviations are approximately harmonic relative to the fundamental and within $\pm 2\%$ of the corresponding harmonics of the analysis frequency. More details on the analysis process are given by Beauchamp [27].

2.2.2 Temporal Equalization

Temporal equalization was done in the frequency domain. Attacks and decays were first identified by inspection of the time-domain amplitude-vs.-time envelopes, and then harmonic amplitude envelopes corresponding to the attack, sustain, and decay were reinterpolated to achieve an attack time of 0.05s, a sustain time of 1.9s, and a decay time of 0.05s, for a total duration of 2.0s.

2.2.3 Spectral Centroid Equalization

Different from our previous study [23], we equalized the average spectral centroid of the the stimuli to see whether other significant features would emerge. Average spectral centroid was equalized for all eight instruments. The spectra of each instrument was modified to an average spectral centroid of 3.7, which was the mean average spectral centroid of the eight tones. This modification was accomplished by scaling each harmonic amplitude by its harmonic number raised to a to-be-determined power:

$$A_k(t) \leftarrow k^p A_k(t) \tag{1}$$

For each tone, starting with $p = 0$, p was iterated using Newton's method until an average spectral centroid was obtained within ± 0.1 of the 3.7 target value.

2.2.4 Resynthesis Method

Stimuli were resynthesized from the time-varying harmonic data using the well-known method of time-varying additive sinewave synthesis (oscillator method) [27] with frequency deviations set to zero.

2.3 Subjects

32 subjects without hearing problems were hired to take the listening test. They were undergraduate students and ranged in age from 19 to 24. Half of them had music training (that is, at least five years of practice on an instrument).

2.4 Emotion Categories

As in our previous study [23], the subjects compared the stimuli in terms of eight emotion categories: Happy, Sad, Heroic, Scary, Comic, Shy, Joyful, and Depressed. These terms were selected because we considered them the most salient and frequently expressed emotions in music, though there are certainly other important emotion categories in music (e.g., Romantic). In picking these eight emotion categories, we particularly had dramatic musical genres such as opera and musicals in mind, where there are typically heroes, villians, and comic-relief characters with music specifically representing each. Their ratings according to the Affective Norms for English Words [28] are shown in Figure 1 using the Valence-Arousal model. Happy, Joyful, Comic, and Heroic form one cluster and Sad and Depressed another.
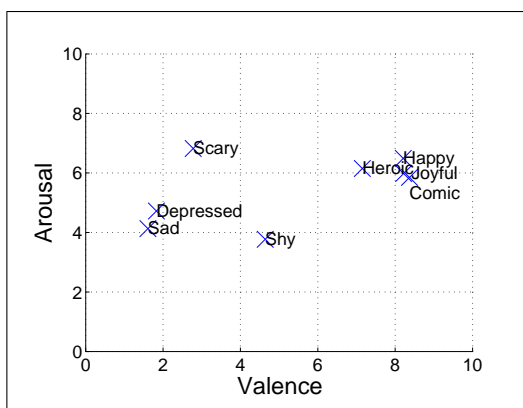

Figure 1. Russel's Valence-Arousal emotion model. Valence is how positive an emotion is. Arousal is how energetic an emotion is.

2.5 Listening Test Design

Every subject made pairwise comparisons of all eight instruments. During each trial, subjects heard a pair of tones from different instruments and were prompted to choose which tone more strongly aroused a given emotion. Each combination of two different instruments was presented in four trials for each emotion, and the listening test totaled $C_2^8 \times 4 \times 8 = 896$ trials. For each emotion, the overall trial presentation order was randomized (i.e., all the Happy comparisons were first in a random order, then all the Sad comparisons were second, ...).

Before the first trial, the subjects read online definitions of the emotion categories from the Cambridge Academic Content Dictionary [29]. The listening test took about 1.5 hours, with breaks every 30 minutes.

The subjects were seated in a "quiet room" with less than 40 dB SPL background noise level. Residual noise was

mostly due to computers and air conditioning. The noise level was further reduced with headphones. Sound signals were converted to analog by a Sound Blaster X-Fi Xtreme Audio sound card, and then presented through Sony MDR-7506 headphones at a level of approximately 78 dB SPL, as measured with a sound-level meter. The Sound Blaster DAC utilized 24 bits with a maximum sampling rate of 96 kHz and a 108 dB S/N ratio.

3. RESULTS

3.1 Quality of Responses

The subjects' responses were first screened for inconsistencies, and two outliers were filtered out. Consistency was defined based on the four comparisons of a pair of instruments A and B for a particular emotion as follows:

$$consistency_{A,B} = \frac{max(v_A, v_B)}{4} \qquad (2)$$

where v_A and v_B are the number of votes a subject gave to each of the two instruments. A consistency of 1 represents perfect consistency, whereas 0.5 represents approximately random guessing. The mean average consistency of all subjects was 0.76.

Predictably subjects were only fairly consistent because of the emotional ambiguities in the stimuli. We assessed the quality of responses further using a probabilistic approach which has been successful in image labeling [30]. We defined the probability of each subject being an outlier based on Whitehill's outlier coefficient. Whitehill *et al.* [30] used an expectation maximization algorithm to estimate each subject's outlier coefficient and the difficulty of evaluating each instance, as well as the labeling of each instance. Higher outlier coefficients mean that the subject is more likely an outlier, which consequently reduces the contribution of their vote toward the label. In our study, we verified that the two least consistent subjects had the highest outlier coefficients. Therefore, they were excluded from the results.

We measured the level of agreement among the remaining subjects with an overall Fleiss' Kappa statistic [31]. Fleiss' Kappa was 0.043, indicating a slight but statistically significant agreement among subjects. From this, we observed that subjects were self-consistent but less agreed in their responses than in our previous study [23] since the tones sounded more similar after spectral centroid equalization.

We also performed a χ^2 test [32] to evaluate whether the number of circular triads significantly deviated from the number to be expected by chance alone. This turned out to be insignificant for all subjects. The approximate likelihood ratio test [32] for significance of weak stochastic transitivity violations [33] was tested and showed no sigificance for all emotions.

3.2 Emotion Results

We ranked the spectral centroid equalized instrument tones by the number of positive votes they received for each
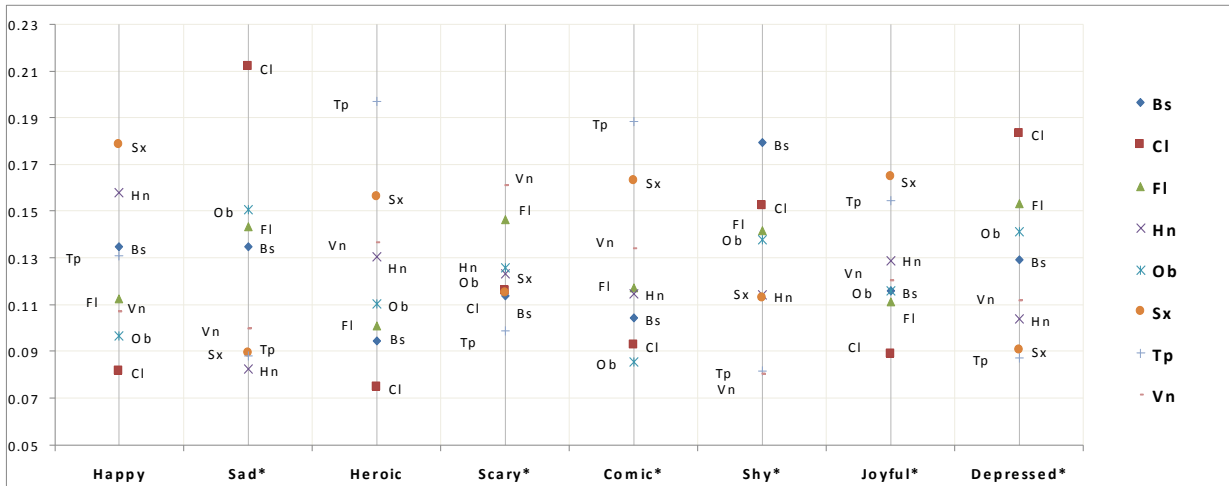
Figure 2. Bradley-Terry-Luce scale values of the spectral centroid equalized tones for each emotion.

emotion, and derived scale values using the Bradley-Terry-Luce (BTL) model [32, 34] as shown in Figure 2. The likelihood-ratio test showed that the BTL model describes the paired-comparisons well for all emotions. We observe that: 1) In general, the BTL scales of the spectral centroid equalized tones were much closer to one another compared to the original tones. The range of the scale considerably narrowed to between 0.07 and 0.23 (in the original tones it was 0.02 to 0.35). The narrower distribution of instruments indicates an increase in difficulty for listeners to make emotional distinctions between the spectral centroid equalized tones. 2) The ranking of the instruments was different than for the original tones. For example, the clarinet and flute were often highly ranked for sad emotions. Also, the horn and the violin were more neutral instruments, which contrasts with their distinctive Sad and Happy rankings respectively for the original tones. And surprisingly, the horn was the least Sad instrument. 3) At the same time, some instruments ranked similarly in both experiments. For example, the trumpet and saxophone were still among the most Happy and Joyful instruments, and the oboe was still ranked in the middle.

Figure 3 shows BTL scale values and the corresponding 95% confidence intervals of the instruments for each emotion. The confidence intervals cluster near the line of indifference since it was difficult for listeners to make emotional distinctions. Table 1 shows the spectral characteristics of the eight spectral centroid equalized tones (since average spectral centroid is equalized to 3.7 for all tones, it is omitted). Spectral centroid deviation was more uniform than in our previous study and near 1.0. This is a side-effect of spectral centroid equalization since deviations are all around the same equalized value of 3.7.

Table 2 shows Pearson correlation between emotion and the spectral features for spectral centroid equalized tones. Even/odd harmonic ratio significantly correlated with Happy, Sad, Joyful, and Depressed. Instruments that had extreme even/odd harmonic ratios exhibited clear patterns in the rankings. For example, the clarinet had the lowest even/odd harmonic ratios and the saxophone the

highest. The two instruments were consistently outliers in Figure 2 with opposite patterns. Table 2 also indicates that listeners found the trumpet and violin less shy than other instruments (i.e., their spectral centroid deviations were more than the other instruments).

4. DISCUSSION

These results and the results in our previous study [23] are consistent with Eerola's Valence-Arousal results [19]. Both indicate that musical instrument timbres carry cues about emotional expression that are easily and consistently recognized by listeners. Both show that spectral centroid/brightness is a significant component in music emotion. Beyond Eerola's findings, we have found that even/odd harmonic ratio is the most salient timbral feature after attack time and brightness.

For future work, it will be fascinating to see how emotion varies with pitch, dynamic level, brightness, and articulation. Do these parameters change emotion in a consistent way, or does it vary from instrument to instrument? We know that increased brightness makes a tone more dramatic (more happy or more angry), but is the effect more pronounced in some instruments than others? For example, if a happy instrument such as the violin is played softly with less brightness, is it still happier than a sad instrument such as the horn played loudly with maximum brightness? At what point are they equally happy? Can we equalize the instruments to equal happiness by simply adjusting brightness or other attributes? How do the happy spaces of the violin overlap with other instruments in terms of pitch, dynamic level, brightness, and articulation? In general, how does timbre space relate to emotional space?

Emotion gives us a fresh perspective on timbre, helping us to get a handle on its perceived dimensions. It gives us a focus for exploring its many aspects. Just as timbre is a multidimensional perceived space, emotion is an even higher-level multidimensional perceived space deeper inside the listener.
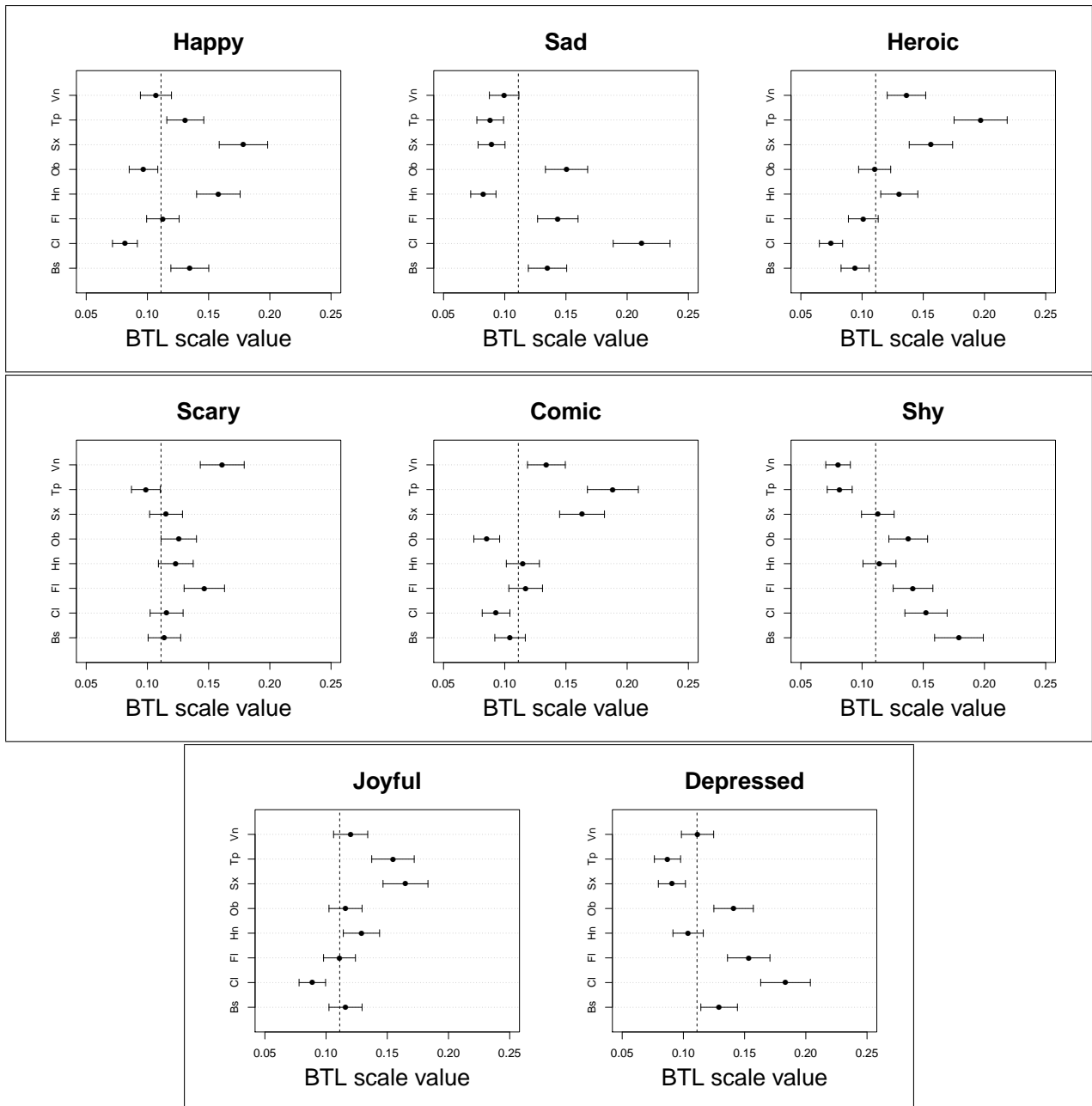
Figure 3. BTL scale values and the corresponding 95% confidence intervals of the spectral centroid equalized tones for each emotion. The dotted line represents no preference.

Features \ Emotion	Bs	Cl	Fl	Hn	Ob	Sx	Tp	Vn
Spectral Centroid Deviation	0.9954	1.0176	1.0614	1.0132	1.0178	1.018	1.1069	1.1408
Spectral Incoherence	0.0817	0.0399	0.1341	0.0345	0.0531	0.0979	0.0979	0.1099
Spectral Irregularity	0.0967	0.1817	0.1448	0.0635	0.1206	0.1947	0.0228	0.1206
Even/odd ratio	1.3246	0.177	0.9541	0.9685	0.456	1.7591	0.81	0.9566

Table 1. Spectral characteristics of the spectral centroid equalized tones.

Emotion Features	Happy	Sad	Heroic	Scary	Comic	Shy	Joyful	Depressed
Spectral Centroid Deviation	-0.2203	-0.3516	0.5243	0.4562	0.5386	**-0.7834**[**]	0.1824	-0.3149
Spectral Incoherence	0.1083	-0.298	0.31	0.4081	0.5046	-0.2665	0.3025	-0.2373
Spectral Irregularity	-0.13	0.499	-0.5082	0.2697	-0.3124	0.3419	-0.2543	0.4877
Even-odd ratio	**0.8596**[**]	**-0.6686**[*]	0.3785	-0.018	0.4869	-0.0963	**0.6879**[*]	**-0.6575**[*]

Table 2. Pearson correlation between emotion and spectral characteristics for spectral centroid equalized tones. **: $p < 0.05$; * : $0.05 < p < 0.1$.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] J. M. Grey and J. W. Gordon, "Perceptual Effects of Spectral Modifications on Musical Timbres," *Journal of the Acoustical Society of America*, vol. 63, p. 1493, 1978.

[2] D. L. Wessel, "Timbre Space as A Musical Control Structure," *Computer music journal*, pp. 45–52, 1979.

[3] C. L. Krumhansl, "Why is Musical Timbre So Hard to Understand," *Structure and Perception of Electroacoustic Sound and Music*, vol. 9, pp. 43–53, 1989.

[4] P. Iverson and C. L. Krumhansl, "Isolating the Dynamic Attributes of Musical Timbre," *The Journal of the Acoustical Society of America*, vol. 94, no. 5, pp. 2595–2603, 1993.

[5] J. Krimphoff, S. McAdams, and S. Winsberg, "Caractérisation du Timbre des Sons Complexes. II. Analyses Acoustiques et Quantification Psychophysique," *Le Journal de Physique IV*, vol. 4, no. C5, pp. C5–625, 1994.

[6] R. Kendall and E. Carterette, "Difference Thresholds for Timbre Related to Spectral Centroid," in *Proceedings of the 4th International Conference on Music Perception and Cognition, Montreal, Canada*, 1996, pp. 91–95.

[7] S. Lakatos, "A Common Perceptual Space for Harmonic and Percussive Timbres," *Perception & Psychophysics*, vol. 62, no. 7, pp. 1426–1439, 2000.

[8] A. Caclin, S. McAdams, B. K. Smith, and S. Winsberg, "Acoustic Correlates of Timbre Space Dimensions: A Confirmatory Study Using Synthetic Tones," *Journal of the Acoustical Society of America*, vol. 118, p. 471, 2005.

[9] J. M. Grey, "Multidimensional Perceptual Scaling of Musical Timbres," *The Journal of the Acoustical Society of America*, vol. 61, no. 5, pp. 1270–1277, 1977.

[10] J. Krimphoff, "Analyse Acoustique et Perception du Timbre," *unpublished DEA thesis, Université du Maine, Le Mans, France*, 1993.

[11] S. Lakatos and J. Beauchamp, "Extended Perceptual Spaces for Pitched and Percussive Timbres," *The Journal of the Acoustical Society of America*, vol. 107, no. 5, pp. 2882–2882, 2000.

[12] J. W. Beauchamp and S. Lakatos, "New Spectrotemporal Measures of Musical Instrument Sounds Used for a Study of Timbral Similarity of Rise-time and Centroid-normalized Musical Sounds," *Proc. 7th Int. Conf. Music Percept. Cognition*, pp. 592–595, 2002.

[13] K. R. Scherer and J. S. Oshinsky, "Cue Utilization in Emotion Attribution from Auditory Stimuli," *Motivation and Emotion*, vol. 1, no. 4, pp. 331–346, 1977.

[14] I. Peretz, L. Gagnon, and B. Bouchard, "Music and Emotion: Perceptual Determinants, Immediacy, and Isolation after Brain Damage," *Cognition*, vol. 68, no. 2, pp. 111–141, 1998.

[15] E. Bigand, S. Vieillard, F. Madurell, J. Marozeau, and A. Dacquet, "Multidimensional Scaling of Emotional Responses to Music: The Effect of Musical Expertise and of the Duration of the Excerpts," *Cognition and Emotion*, vol. 19, no. 8, pp. 1113–1139, 2005.

[16] J.-J. Aucouturier, F. Pachet, and M. Sandler, "The Way it Sounds: Timbre Models for Analysis and Retrieval of Music Signals," *IEEE Transactions on Multimedia*, vol. 7, no. 6, pp. 1028–1035, 2005.

[17] G. Tzanetakis and P. Cook, "Musical Genre Classification of Audio Signals," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 5, pp. 293–302, 2002.

[18] C. Baume, "Evaluation of Acoustic Features for Music Emotion Recognition," in *Audio Engineering Society Convention 134*. Audio Engineering Society, 2013.

[19] T. Eerola, R. Ferrer, and V. Alluri, "Timbre and Affect Dimensions: Evidence from Affect and Similarity Ratings and Acoustic Correlates of Isolated Instrument Sounds," *Music Perception*, vol. 30, no. 1, pp. 49–70, 2012.

[20] Y.-H. Yang, Y.-C. Lin, Y.-F. Su, and H. H. Chen, "A Regression Approach to Music Emotion Recognition," *IEEE TASLP*, vol. 16, no. 2, pp. 448–457, 2008.

[21] W. Ellermeier, M. Mader, and P. Daniel, "Scaling the Unpleasantness of Sounds According to the BTL

Model: Ratio-scale Representation and Psychoacoustical Analysis," *Acta Acustica United with Acustica*, vol. 90, no. 1, pp. 101–107, 2004.

[22] T. Eerola and J. K. Vuoskoski, "A Comparison of the Discrete and Dimensional Models of Emotion in Music," *Psychology of Music*, vol. 39, no. 1, pp. 18–49, 2011.

[23] B. Wu, S. Wun, C. Lee, and A. Horner, "Spectral Correlates in Emotion Labeling of Sustained Musical Instrument Tones," in *Proceedings of the 14th International Society for Music Information Retrieval Conference*, November 4-8 2013.

[24] A. Horner, J. Beauchamp, and R. So, "Detection of Random Alterations to Time-varying Musical Instrument Spectra," *Journal of the Acoustical Society of America*, vol. 116, pp. 1800–1810, 2004.

[25] S. McAdams, J. W. Beauchamp, and S. Meneguzzi, "Discrimination of Musical Instrument Sounds Resynthesized with Simplified Spectrotemporal Parameters," *Journal of the Acoustical Society of America*, vol. 105, p. 882, 1999.

[26] B. C. 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.

[27] J. W. Beauchamp, "Analysis and Synthesis of Musical Instrument Sounds," in *Analysis, Synthesis, and Perception of Musical Sounds*.   Springer, 2007, pp. 1–89.

[28] M. M. Bradley and P. J. Lang, "Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings," *Psychology*, no. C-1, pp. 1–45, 1999.

[29] "happy, sad, heroic, scary, comic, shy, joyful and depressed," *Cambridge Academic Content Dictionary*, 2013, online: http://goo.gl/v5xJZ (17 Feb 2013).

[30] J. Whitehill, P. Ruvolo, T. Wu, J. Bergsma, and J. Movellan, "Whose Vote Should Count More: Optimal Integration of Labels from Labelers of Unknown Expertise," *Advances in Neural Information Processing Systems*, vol. 22, no. 2035-2043, pp. 7–13, 2009.

[31] F. L. Joseph, "Measuring Nominal Scale Agreement among Many Raters," *Psychological Bulletin*, vol. 76, no. 5, pp. 378–382, 1971.

[32] 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.

[33] A. Tversky, "Intransitivity of Preferences." *Psychological Review*, vol. 76, no. 1, p. 31, 1969.

[34] R. A. Bradley, "Paired Comparisons: Some Basic Procedures and Examples," *Nonparametric Methods*, vol. 4, pp. 299–326, 1984.