The Effects of MP3 Compression on Emotional Characteristics in Musical Instruments

CHOI, Ga Lam

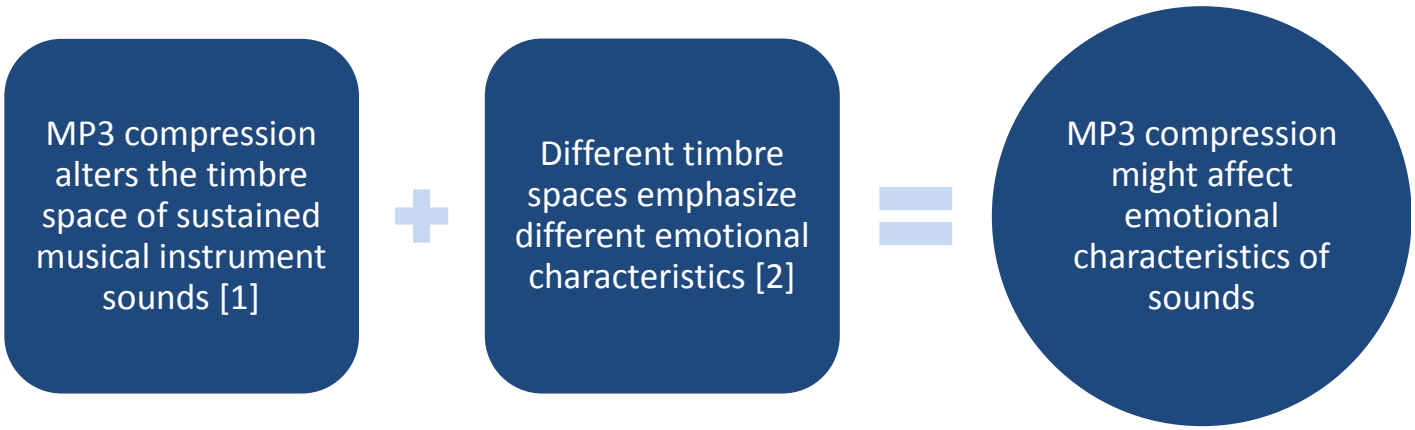
Prof. Andrew HORNER

Abstract

Previous research has shown that musical instrument sounds have strong emotional characteristics. Other research has shown that MP3 compression can change the timbre of musical instruments. This paper investigates the effect of MP3 compression on music emotion. We conducted listening tests to compare the effect of MP3 compression on the emotional characteristics of eight sustained instrument sounds compared pairwise. We compared these sounds over ten emotional categories. The results show that MP3 compression strengthened the emotional categories Sad, Scary, Shy, and Mysterious, and weakened Happy, Heroic, Comic, Romantic, and Calm. Interestingly, Angry was relatively unaffected by MP3 compression.

Objectives

This work is fundamental research in music emotion and timbre that will help audio engineers, musicians and listeners better understand the emotional effects of MP3 compression* on music.



It would be useful to know how MP3 compression affects the emotional characteristics of musical instruments. In particular, this research will tackle following questions:

- Generally, what are the emotional effects of MP3 compression? (Do all emotional characteristics decrease about equally with more compression?)
- Which emotional characteristics increase or decrease with more compression?
- Which emotional characteristics are unaffected by more compression?
- Which instruments change the most or the least with more compression?
- Which instruments are unaffected by more compression?

***MP3 Compression?** An audio coding format for digital audio which reduces the size of audio files by cutting out less audible parts of the sound.

[1] C. Lee, A. B. Horner and B. Wu, "The Effect of MP3 Compression on the Timbre Space of Sustained Musical Instrument Sounds," J. Audio Eng. Soc., vol. 61, pp. 840-849, 2013.

[2] B. Wu, C. W. Wun, C. Lee and A. Horner, "Investigating Correlation between Musical Timbres and Emotions," in *International Society for Music Information Retrieval Conference (ISMIR)*, Curitiba, Brazil, 2013.

Research Methodology

This research adopts a listening test from the research by Chau et al.[4], but using MP3 compressed stimuli.

8 Sustained Instrument Sounds

Bassoon	Clarinet
Flute	Horn
Oboe	Saxophone
Trumpet	Violin

- Fundamental frequencies close to Eb4
- Instrument sounds were compressed with three different bit rates (32, 56, and 112 Kbps) using LAME MP3 encoder [4]

10 Emotional Categories

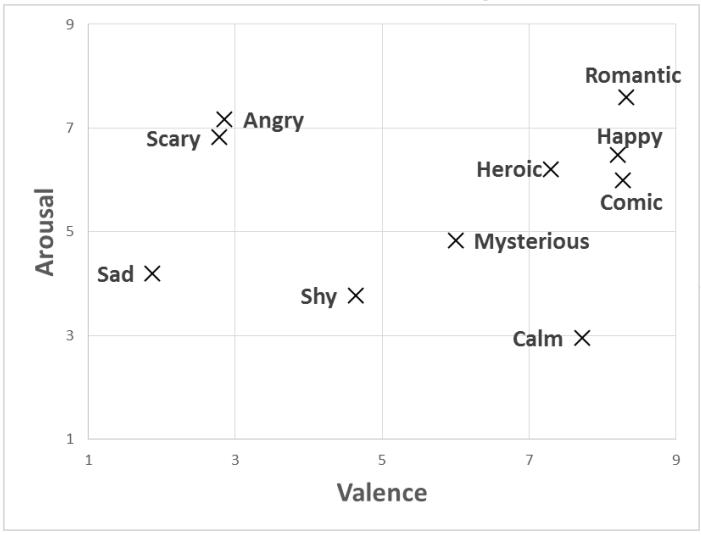


Figure 1. Distribution of the emotional characteristics in the dimensions valence and arousal [5].

Test Procedure

1. The subjects were provided with an instruction sheet containing definitions of the ten emotional categories from the Cambridge Academic Content Dictionary [6]

2. Every subject made pairwise comparisons (see the UI on the right)

   Each combination of two different compressions was presented for each instrument and emotion

   total P(4,2) × 8 instruments × 10 emotional categories = 960 trials

Test Duration: 2 hours including 5 minutes breaks every 30 minutes

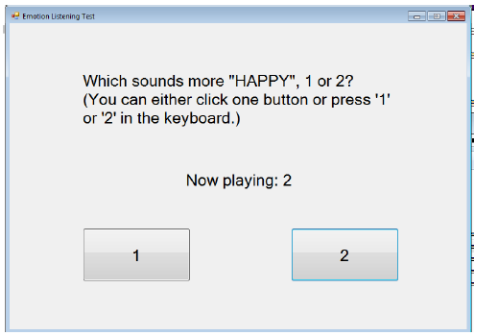


Figure 2. Listening test interface

Test Subjects

- 20 Undergraduate students from HKUST without any hearing problems

Test Environment

- Quiet room with 39dB SPL background noise level (mostly due to computers and ventilation)
- Noise further reduced with headphones with 78dB SPL output level

Equipment

- Sound Blaster X-Fi Xtreme Audio sound card (24 bits, max sampling rate 96kHz, 108dB S/N ratio)
- Sony MDR-7506 headphones

[4] C. J. Chau, B. Wu and A. Horner, "Timbre Features and Music Emotion in Plucked String, Mallet Percussion, and Keyboard Sounds," in International Computer Music Conference (ICMC), Athens, Greece, Sept 2014.

[5] 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.

[6] "Cambridge academic content dictionary," Cambridge University Press, [Online]. Available: http://dictionary.cambridge.org/dictionary/american-english.

Results

We ranked the compressed sounds by the number of positive votes they received for each instrument and emotion, and derived scale values using the Bradley-Terry-Luce (BTL) statistical model [7, 8]. For each instrument-emotion pair, the BTL scale values for original and three compressed sounds sum to 1. The BTL value for each sound is the probability that listeners will choose that compression rate when considering a certain instrument and emotion category. For example, if all four sounds (the original and three compressed sounds) are judged equally happy, the BTL scale values would be 1/4=0.25. We also derived the corresponding 95% confidence intervals for the compressed sounds using the method proposed by Bradley (1984) [7].

	112Kbps	**56Kbps**	**32Kbps**
Happy	1	3	8
Heroic	0	1	7
Romantic	1	0	6
Comic	0	2	5
Calm	2	2	8
Mysterious	0	2	6
Shy	1	0	8
Angry	1	0	1
Scary	0	2	7
Sad	0	1	8
Avg.	0.6	1.3	6.4

Table 1. The number of instruments that were significantly different from the original sound (i.e., the 95% confidence intervals of the original and compressed sounds did not overlap) for each compression rate and emotional category.

	Bs	Cl	Fl	**Hn**	Ob	Sx	Tp	Vn	Sum
Happy	2	3	1	**1**	2	1	1	1	12
Heroic	1	1	2	**0**	1	1	1	1	8
Romantic	1	1	1	**0**	1	1	1	1	7
Comic	1	1	1	**1**	1	1	1	1	8
Calm	1	2	1	**1**	2	3	1	1	12
Mysterious	1	0	1	**0**	2	1	1	1	7
Shy	1	1	1	**1**	1	2	1	1	9
Angry	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**1**
Scary	1	1	1	**0**	1	1	2	1	8
Sad	1	2	1	**1**	1	1	1	1	9
Sum	10	13	10	**5**	12	12	10	9	

Table 2. The number of compressed sounds that were significantly different from the original sound (i.e., the 95% confidence intervals of the original and compressed sounds did not overlap) for each instrument and emotional category.

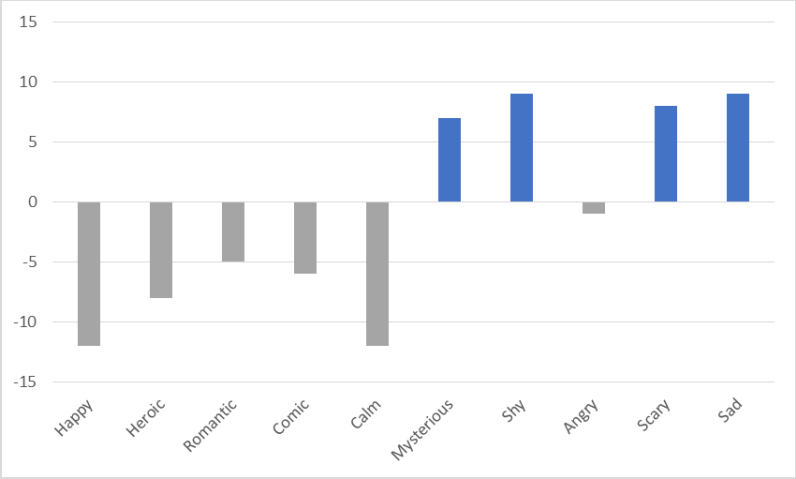


Figure 3. The number of significant differences between the original and compressed sounds, where strengthened emotional categories are positive, and weakened emotional categories are negative.

[7] 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(1), pp. 29-40, 2004.

[8] R. A. Bradley, "Paired Comparisons: Some Basic Procedures and Examples," Non-parametric Methods, vol. 4, pp. 299-326, 1984.

Discussion

Based on the Figures 3 and Tables 1 and 2, our main findings are as follows:

1. MP3 compression affected some instruments more and others less for different emotional characteristics. The clarinet, oboe, and saxophone were most affected, and the horn by far the least affected (see Table 3).

2. Negative and neutral emotional characteristics (Sad, Scary, Shy, and Mysterious) increased with more MP3 compression in the samples we tested (see Figure 4).

3. Positive emotional characteristics (Happy, Heroic, Romantic, Comic, and Calm) decreased with more MP3 compression in the samples we tested (see Figure 4).

4. Angry was relatively unaffected by MP3 compression for the rates we tested (see Figure 4).

Perhaps quantization jitter introduced into the amplitude envelopes by MP3 compression decreased positive emotional characteristics such as Happy, Heroic, Romantic, Comic and Calm while increasing others such as Mysterious by changing the quality of sounds to be somewhat different and unnatural.