Effectiveness of Choosing Dissonant Combination of Tones for Multivariate Data Sonification

Yukio Horiguchi[(✉)], Moriyu Nakashima, Hiroaki Nakanishi, and Tetsuo Sawaragi

Department of Mechanical Engineering and Science, Kyoto University, Kyoto, Japan
{horiguchi,nakanishi,sawaragi}@me.kyoto-u.ac.jp, nakashima.moriyu.23z@st.kyoto-u.ac.jp

Abstract. In the present paper, additive syntheses of two tones are examined in order to look into what data-to-sound mapping design is better to acoustically communicate varying quantities of a multivariate system. The research focus is placed on the effectiveness of mappings for listeners to discriminate one variable component of sonified data from another. Dyads of pitches are tested through a couple of experiments where a special attention is paid to the "dissonance" of two tones. The first experiment investigates its effects on the perception of two static tones presented simultaneously. The second experiment investigates the effectiveness of choosing dissonant pitch combinations to sonify two quantities of a system being controlled manually. Experimental results demonstrate that the dissonance of tones can improve the perception of changes in individual variates out of a synthesized sound but that the frequency interval between two tones has stronger effects in dynamic situations.

Keywords: Tonal dissonance · Sonification · Parameter mapping · Auditory display

1 Introduction

Sonification is a subset of auditory display to convey information using non-speech audio [1–3]. By means of sonification, data relations are transformed into perceived relations in acoustic signals. While it has been successfully deployed in a wide range of application domains, general guidelines for sonification design are yet to be developed [2–4].

Parameter mapping is one of the most commonly used techniques to produce a sonification, which defines a set of mappings between data dimensions and acoustic dimensions (e.g., loudness, pitch, tempo, etc.) [1–4]. The authors have been doing empirical studies to examine parameter mappings to make audible the gap between desired and current state of dynamic systems being controlled. Various data-to-sound mappings were tested using a one-dimensional, invisible target tracking task [5, 6]. The experimental results clarified preferable polarity and scaling function of the mapping for a type of sonification that has a target to pursue. Reducing the sound intensity as the

gap decreases and making no sound at the goal state are found to be the most effective mapping among those that have intensity and pitch as acoustic parameters to use. These results are, however, limited to univariate systems. In order to extend the application domain, knowledge of designing sonification for multivariate systems needs to be explored.

In the present paper, additive syntheses of two tones are examined in order to look into what mapping design is better to communicate quantities of a multivariate system that changes dynamically. The research focus is placed on the effectiveness of data-to-sound mappings for listeners to discriminate one variate of sonified data from another. On the basis of empirical findings mentioned above, quantitative variables are mapped onto the amplitude of different tones so that the intensity of each tone represents the difference of the corresponding variable value from its reference. The pitch of tones is therefore the acoustic parameter to examine. Parameter mappings independent of the physical configuration of the task space are explored in this study.

Dyads of pitches are tested through a couple of experiments where a special attention is paid to consonance and dissonance of two tones. Consonance had been studied in the field of musical aesthetics and is a matter of tone quality playing an important role in determining a level of pleasantness/unpleasantness of sounds [7]. The tonal consonance is defined as the sensation of "clearness" while the dissonance is "turbidity". The consonance of a chord changes according to the frequency ratio of its components. Two tones sound consonant only if their frequency ratio is integer; otherwise they are dissonant. According to Stumpf's tonal fusion theory of consonance [8], two consonant tones may be perceptually fused into a single harmonic series. The research question here is whether or not dissonant combinations of tones can improve the listener's perception of changes in individual auditory components out of a synthesized sound.

2   Experiment 1: Discrimination of Static Tones

2.1   Experimental Setup and Procedure

The first experiment was carried out to examine the loudness perception of two *static* pure tones with different pitch levels. Participants were requested to judge the loudness of individual tones that were simultaneously presented to them.

The frequencies of the tones are defined as shown in Table 1. Frequency difference (3 levels) and consonance vs. dissonance (2 levels) are factors of the experiment. Combinations #1 through #3 consist of consonant tones with the 250, 500, and 750 Hz frequency difference, respectively. Combinations #4 through #6 consist of dissonant tones that have the same variations of the frequency difference.

The experiment employed a within-subject design. It consists of six sections that correspond to different tone combinations and each section is divided further into two blocks. In the first block, participants were exposed to individual tones of the combination separately to learn the reference sound level and the sound intensity variation range. The amplitude of each tone was set to 0.5, 0, and 1 by turns to illustrate to participants the intermediate, minimum (i.e., silent), and maximum sound levels of the tone, respectively.

Table 1. Combination of frequencies in Experiment 1

Combination #	Frequency of Tone 1 (Hz)	Frequency of Tone 2 (Hz)
1	750	500
2	1000	500
3	1500	500
4	763	513
5	1013	513
6	1513	513

In the second block, four static tones were presented in sequence that were composites of two pure tones of the same frequency combination but they were different combinations of amplitudes. Each time they heard a static tone, the participants were requested to tell their judgment of each constituent tone's loudness by placing a check mark on the scale shown in Fig. 1. Errors between presented and perceived sound level were evaluated for comparison. Twenty-four samples in total (6 frequency combinations × 4 amplitude combinations) were collected and the order of the test combinations was randomized across participants to balance the order effect.

Silent ├──────────────┼──────────────┤ Loud
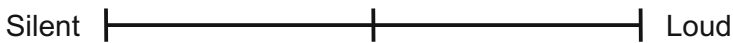
Fig. 1. Linear scale to answer the loudness of each tone

For this experiment, Max/MSP 5 was used for sound synthesis and auditory stimuli were presented though a BOSE Quiet Comfort 15 active noise canceling headphone.

2.2 Results

Ten undergraduate and graduate students (aged from 22 to 24 years) participated in the experiment from the department of mechanical engineering and science in Kyoto University. They are normal in hearing ability. The experimenter told them sufficient instructions about the purpose and procedure of the experiment as well as the designs of the auditory displays. All of them gave informed consent to the experiment.

Figure 2 presents a summary of the experimental results. A two-way ANOVA found no effect of the frequency difference but a significant effect of the consonance/dissonance factor ($p < 0.001$) onto the perception error. A post-hoc Tukey's test shows that, compared to the consonant combination, the perception error decreases by the dissonant combination of tones significantly when the frequency interval is 1,000 Hz ($p < 0.01$) and marginally significantly when the frequency difference is 500 Hz ($p = 0.070$). The perception error also decreases significantly with dissonant combinations as the frequency difference increases from 250 Hz to 1,000 Hz ($p < 0.05$).
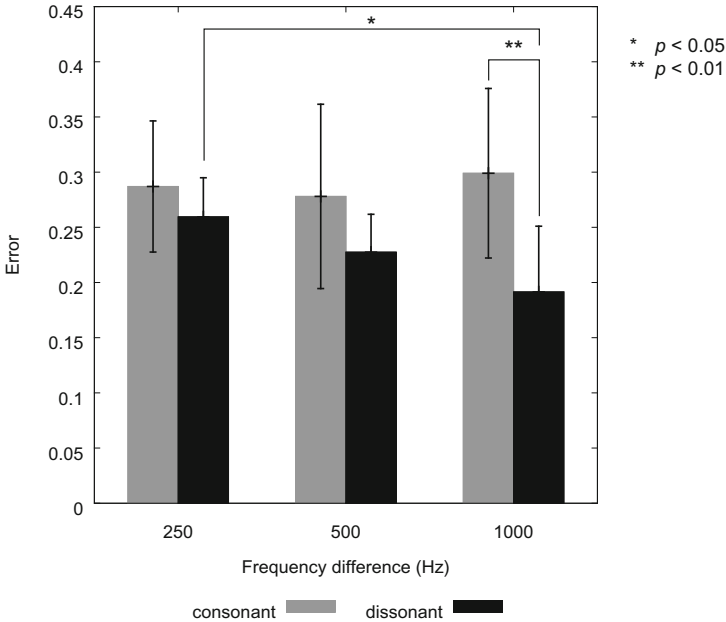
Fig. 2. Loudness perception error (mean + SD)

The above result indicates that the participants can achieve more accurate loudness perception when the two tones are given in a dissonant combination than in a consonant one. This effect becomes stronger when the tones have a larger difference in the frequency. On the other hand, some of the participants made comments that they could hardly measure how much amount of intensity difference the stimulus tones had relative to the reference sound level. They could only hear whether they sounded larger or smaller.

3   Experiment 2: Discrimination of Dynamic Tones for Control

3.1   Experimental Setup

In the second experiment, the same pitch combinations were tested using a kind of tracking task where the system being controlled had two dimensions of the state, one of which was controlled manually and the other was defined automatically. Both dimensions were sonified by different pure tones so that their intensity changed dynamically with the corresponding state variable. Human operators needed to specify the status of each variable from auditory signals for successful control of the system.

Figure 3 shows a visual representation of the experimental task. The two bars represent non-negative quantities to be displayed acoustically by the intensity of two tones. The operator's task is to make both values zero at the same instant. On the one
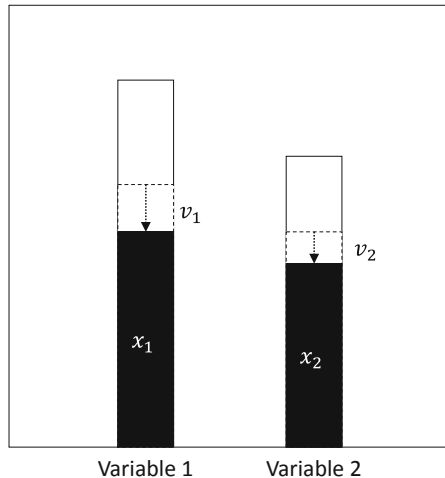
Fig. 3. Visual representation of the auditory tracking task

hand, Variable 1 corresponds to the system dimension the operator tries to control and its behavior is defined as a second order system. The input device is a joystick whose displacement is mapped onto the acceleration of the system being controlled. Variable 2, on the other hand, gives a reference signal to the operator. Its value automatically decreases at a constant rate that varies randomly from trail to trail. Participants were requested to control the position of the Variable 1 bar so that it reached zero exactly when the Variable 2 bar got vanished. The difference in the timing of reaching zero between the two variables is the main evaluation metric in this experiment.

On the basis of findings in our previous studies [5, 6], the data-to-sound mapping for each dimension was designed so that the loudness of a generated sound represents a quantity to be diminished. The amplitude of the tone is given by

$$A(d) = \text{C}\ d^{\frac{1}{2\log_{10} 2}} \tag{1}$$

where d represents the quantity of interest and C is a constant coefficient. Through this mapping, the sound intensity decreases as the quantity decreases and the "loudness" [9] of the sound changes proportional to the quantity.

The tested tones are defined in Table 2, which consist of the same pitch combinations as the first experiment. Frequency difference (3 levels), consonance vs. dissonance (2 levels), and higher vs. lower pitch for the control dimension (2 levels) were factors of this experiment. Combinations #1 through #6 consist of consonant tones while combinations #7 through #12 consist of dissonant tones. Within combinations #1 through #3 and #7 through #9, lower frequencies are assigned to the control dimension, i.e., Variable 1; otherwise, higher frequencies are assigned. The frequency difference of the two tones is chosen from 250, 500, and 750 Hz.

Same as Experiment 1, Max/MSP 5 was used for sound synthesis and auditory signals were presented though a BOSE Quiet Comfort 15 headphone. The sound

Table 2. Combination of frequencies in Experiment 2

Combination #	Frequency of Tone 1 (Hz)	Frequency of Tone 2 (Hz)
1	500	750
2	500	1000
3	500	1500
4	500	500
5	750	500
6	1000	500
7	513	763
8	513	1013
9	513	1513
10	763	513
11	1013	513
12	1513	513

pressure level of the generated sound ranged from 55 dB (no audio output) to 86 dB (the maximum output).

3.2   Procedure

The experiment employed a within-subject design and consisted of two sections. The first section is the practice session for participants to get familiar with the tracking task itself. In this section, the participants practiced the auditory tracking task with a visual aid for 3 min and then 15 trials with no aid. The visual representation of the tracking task, which was shown in Fig. 3, was provided as the visual aid.

The second section is divided into three blocks. The first block is the rehearsal session, in which the participants performed the auditory tracking task using a particular combination of tones. They rehearsed 6 trials without the visual aid to learn the tones to be tested in this section. At the beginning of this block, the experimenter told the participants which pitch in the combination was assigned to the control dimension. The second block is the recording session. The participants did 10 trials for the same configuration of the auditory display. The last block was prepared for collecting participants' evaluation of the mappings. The participants were requested to fill in a NASA-TLX subjective workload questionnaire sheet. The second section lasted until all of the tone combinations listed in Table 2 were tested.

One hundred and twenty trials per participant were recorded for performance evaluation. The order of the test combinations was randomized across participants to balance the order effect. Tracking error and workload measure are evaluation functions to compare different parameter mappings. The former aspect is evaluated by the amount of the time difference of reaching zero (time difference) and the remaining quantity when the other quantity has reached zero (position error). The NASA-TLX Weighted Work Load (WWL) score is used for the latter evaluation aspect.

3.3   Results

Twelve undergraduate and graduate students (aged from 22 to 31 years) participated in the experiment from the department of mechanical engineering and science in Kyoto University. All of them are normal in hearing ability and gave informed consent to the experiment.

Figures 4, 5 and 6 show experimental results described with respect to the time difference, position error, and WWL score, respectively. Both of the performance measures (Figs. 4 and 5) indicate that the listener can achieve more accurate control when the component tones are given with a larger difference in the pitch level. The more similar frequencies two tones have, the more difficult for listeners to discriminate them. This phenomenon, known as "auditory masking" (in the frequency domain) [10], obviously provides a solid rationale for the observed performance degradation with a smaller frequency difference. The WWL score (Fig. 6) shows the same trend, demonstrating that the participants' subjective evaluation of the task difficulty is congruent with their task performance.

Three-way ANOVAs with repeated measures found significant effects of the frequency difference ($p < 0.05$) onto all of three dependent measures. Post-hoc Tukey's tests show that there are significant mean differences in these measures between the frequency difference of 250 Hz and 1,000 Hz ($p < 0.05$). No effect was found with respect to the consonance/dissonance of the tone combination or the higher/lower pitch assignment to the control dimension. The mean difference in the time difference, however, approached significance between the consonant and the dissonant tone combination at the frequency difference of 500 Hz ($p = 0.088$). Although its effect is limited, there was a trend for performance improvement by choosing dissonant pitch combinations to sonify two quantities.
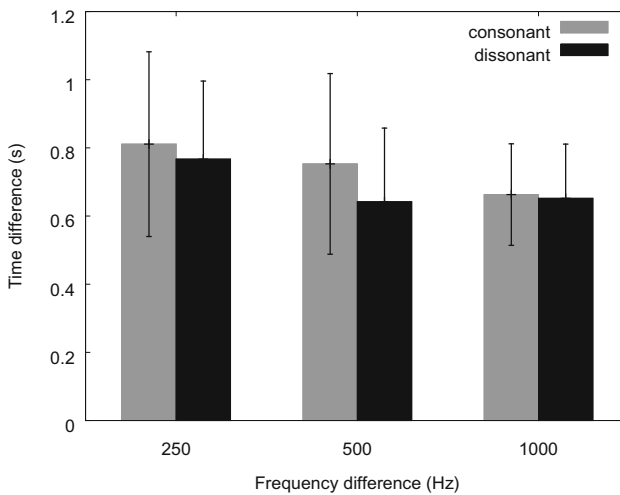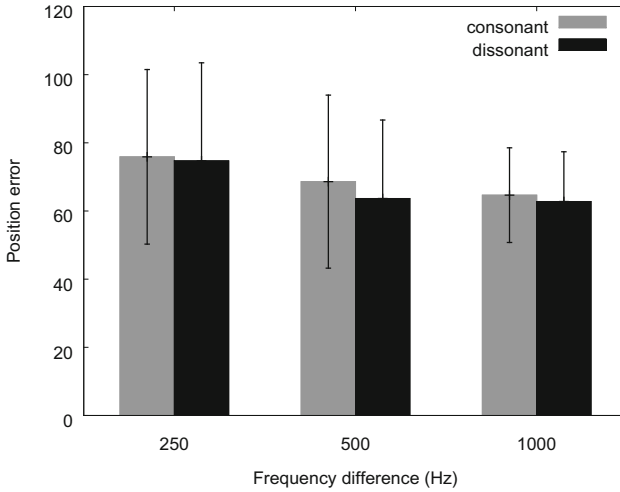

Fig. 4. Time difference (mean + SD)

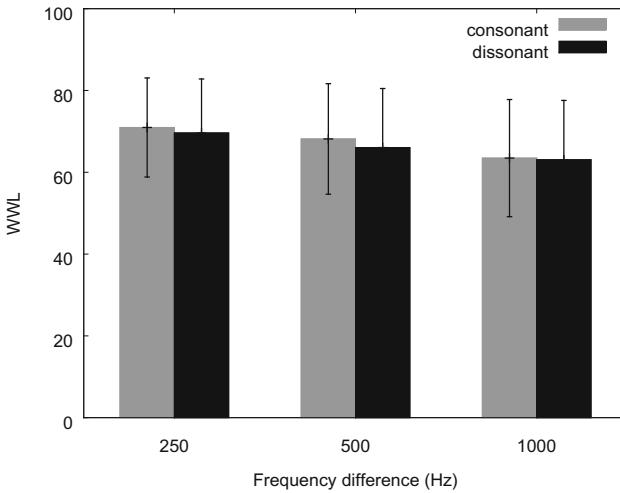**Fig. 5.** Position error (mean + SD)

Fig. 6. WWL score (mean + SD)

4   Conclusions

The overall results of the present study demonstrated that choosing dissonant pitch combinations can make a difference in sonification of multivariate data. The tonal dissonance is beneficial for listeners to differentiate one variate from another although the findings are limited to two variate systems.

The first experiment revealed that two tones in a dissonant combination are easier to recognize the intensity of each tone from their additive synthesis than those in a

consonant combination. On the other hand, the second experiment showed that this effect is limited to cases where the tones have no temporal variation in their intensity. Increasing the pitch interval of the tones has much stronger effects for dynamic situations than their tonal dissonance. The result implies that intensity variations of partials may cause substantial changes in tone quality. The second experiment, however, confirmed a trend for performance improvement by the tonal dissonance. The dissonance does not deteriorate the perception of varying quantities in a synthesized tone.

Because the experiments were designed to use equal frequency intervals, the most dissonant point [7] was not chosen to determine pitch combinations. The effectiveness therefore could be increased if the most dissonant combinations are employed for pitches to map. The more variates to sonify, the more difficult to determine pitch levels of individual quantities. Considering not only the frequency interval but the frequency ratio may help design decision making for such cases.

References

1. Kramer, G.: An introduction to auditory display. In: Kramer, G. (ed.) Auditory Display, Sonification, Audition, and Auditory Interfaces, pp. 1–7. Addison-Wesely, Boston (1994)
2. Eldridge, A.: Issues in auditory display. Artif. Life **12**(2), 259–274 (2006)
3. Peres, S.C., Best, V., Brock, D., Frauenberger, C., Hermann, T., Neuhoff, J.G., Nickerson, L.V., Shinn-Cunningham, B., Stockman, T.: Auditory interfaces. In: Kortum, P. (ed.) HCI Beyond the GUI: Design for Haptic, Speech, Olfactory, and Other Nontraditional Interfaces, pp. 147–195. Morgan Kaufmann, Burlington (2008)
4. Dubus, G., Bresin, R.: A systematic review of mapping strategies for the sonification of physical quantities. PLoS ONE **8**(12), e82491 (2013)
5. Horiguchi, Y., Yasuda, K., Nakanishi, H., Sawaragi, T.: Data-to-sound mapping to sonify ongoing system status in continuous manual control task. In: Proceedings of Human Factors and Ergonomics Society Annual Meeting, vol. 55, pp. 1235–1239 (2011)
6. Horiguchi, Y., Miyajima, K., Nakanishi, H., Sawaragi, T.: Parameter-mapping sonification for manual control task: timbre and intensity manipulation to sonify dynamic system state. In: Proceedings of the SICE Annual Conference 2013, pp. 2341–2346 (2013)
7. Kameoka, A., Kuriyagawa, M.: Consonance theory part I: consonance of dyads. J. Acoust. Soc. Am. **45**, 1451–1459 (1969)
8. DeWitt, L.A., Crowder, R.G.: Tonal fusion of consonant musical intervals: the oomph in Stumpf. Atten. Percep. Psychophy. **41**(1), 73–84 (1987)
9. Stevens, S.S.: The measurement of loudness. J. Acoust. Soc. Am. **27**(5), 815–829 (1955)
10. Moore, B.C.J.: An Introduction to the Psychology of Hearing, 3rd edn. Academic Press, Cambridge (1997)