Microtonal Music Dataset v1

Tatsunori Hirai[1], Lamo Nagasaka[1], and Takuya Kato[1,2] *

[1] Komazawa University
[2] ExaWizards Inc.
thirai@komazawa-u.ac.jp

Abstract. In this paper, we propose a microtonal music dataset, comprising musical compositions that utilize microtones, tones with intervals that are more refined than those found in the 12 equal temperament. As part of the Microtonal music dataset v1, we present 100 manually created microtonal music pieces, along with their characteristics and statistical information. Furthermore, we will discuss the potential for future music information processing research that can be realized using the microtonal music dataset.

Keywords: Microtonal music; microtone; dataset

1 Introduction

The recent advancements in generative AI technology are progressing at an astonishing speed, and the distinctions between human-composed music and AI-generated music are becoming increasingly blurred. As a result, we are reaching a point where music can be generated with a single click, reducing the need for human composition, especially for music for trivial purposes.

Microtonal music is cited as one of the musical expressions that necessitates tools for supporting expression. Microtonal music refers to music that uses microtones, pitches that do not conform to the 12 equal temperament,which is difficult to perform with many traditional instruments. Composing microtonal music is challenging even for people with experience in composing conventional music. We believe that expanding human expressive capabilities through AI assistance, especially for music that are difficult to perform or compose within current frameworks, can contribute to the development of musical culture. Therefore, in this study, we propose a microtonal music dataset to accelerate research on the technology capable of handling microtonal music.

If technology capable of handling microtones is realized, it could enable support such as redesigning the piano roll according to the temperament inferred from the microtones input by the user[1], or providing accompaniment to the microtone melodies composed by the user. These tasks are currently challenging even for humans, and assistance through technologies such as AI is effective.

* This work was supported by JSPS KAKENHI Grant Number JP23K17023.

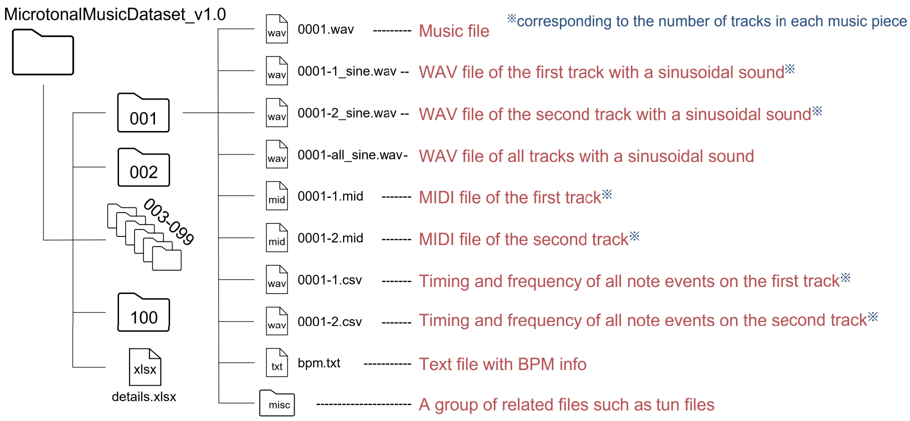

Fig. 1. Directory structure of the dataset.

In this study, as a first step towards realizing technology capable of handling such microtones, we propose a dataset composed of 100 manually composed microtonal music pieces.

2 Related Work

The RWC Music Dataset, consisting of 315 songs and 50 types of instrumental sounds, has significantly impacted music information processing research. This dataset avoid copyright issues in research, and continues to be influential in the field.

The JSB Chorales Dataset, which digitizes 382 four-part chorales composed by Bach, has been utilized in many studies on music generation technologies[3], [4]. On the other hand, no dataset related to microtonal music has been proposed so far.

3 Design of Microtonal Music Dataset v1

3.1 Structures of the Dataset

The structure of the Microtonal Music Dataset directory is illustrated in Fig 1. The data for each piece is stored in folders with IDs 001 to 100. Each folder contains the music file, MIDI data for each track that composes the piece, wav data written with sine waves for those tracks, a CSV file recording the frequency, onset, and offset of each note for every track, and a text file recording the BPM data of the piece. Since MIDI data cannot record the exact frequency of microtones, the frequency information is included in the CSV file. The misc folder includes items such as tuning files (.tun) loaded into the software synthesizer for playing microtones. The file details.xlsx consolidates various information, including the statistical data of the pieces comprising the dataset and the temperament information.

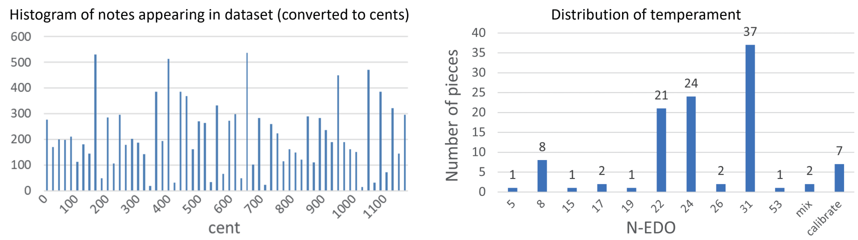

Fig. 2. Histogram of notes in cent (left) and distribution of temperament in the dataset (right).

3.2 Dataset Creation Method

The pieces comprising the dataset were produced using the DAW software Studio One v5.5.2 by one of the authors. To create pieces including microtones in Studio One, we used the software synthesizer Vital, which can play microtones by loading tuning files, and SimpleMicrotonalSynth, which allows for a variety of selectable microtonal tuning options. Each piece was created either by inputting one microtone at a time or by using a MIDI controller for real-time input.

The tuning files for loading into Vital were created using a web page called ScaleWorkShop. We also used microtones expressed by tuning the synthesizer in cent units.

4 Statistis of the Dataset

Here we describe the characteristics of the dataset. Piece lengths average 22.7 seconds, ranging from 6.0 to 74.0 seconds, with an average BPM of 130.4, between 80 and 180.

Fig. 2 (left) illustrates the histogram of notes appearing in the dataset, converted into cents. Here, we set 261.626Hz as the 0-cent reference point and, by utilizing octave equivalence, we convert all notes to frequencies within the same octave before calculating their values in cents. In fig. 2 (left), bins at multiples of 100 represent the notes in 12 equal temperament. The fact that these bins do not show particularly high values indicates that no specific 12 equal temperament notes are being used extensively. Conversely, the infrequent use of sounds in certain frequency bands, such as those around 340 cents and 1040 cents, is intriguing. Despite being a microtonal music dataset, many pieces also include tones from 12 equal temperament. Notably, in 24 equal temperament, half the tones align with the 12 equal temperament. In the pieces, the proportion of 12 equal temperament notes ranged from 0% to 67.1%, with an average of 21.4% across the dataset. Considering the familiarity of the music, a certain degree of use of the 12 equal temperament notes is allowed.

In the dataset, 525 distinct pitches appear, including 474 types of microtones and 51 tones of the 12 equal temperament. These microtones could be candidates for pitches in microtonal music generation models using this dataset.

This dataset includes microtonal music based on N-equal temperament (N-EDO) other than 12, music created by uniformly shifting tunings from specific temperament by X cents (calibrate), and their combinations (mix). Fig. 2 (right) shows the distribution of temperament that make up this dataset. As a temperament that is easy for the composer to create music, scales that felt harmonically familiar were frequently adopted, particularly pieces with the 22, 24, and 31 equal temperaments. In the future, we plan to enhance the diversity of temperaments.

5 Potential Uses of the Dataset

The primary envisioned use of this dataset is for machine learning models. In deep music generation models based on 12 equal temperament, the model is constructed on the assumption that the input and output data are in 12 equal temperament. When extending this to microtonal music, it is anticipated that simply changing the input and output layers would not be sufficient. By utilizing this dataset, it becomes possible to further conduct research into models that can handle microtones.

Additionally, it can be utilized as test data to further generalize conventional music recognition techniques. In music analysis, concepts that presuppose 12 equal temperament, such as chromagrams, are sometimes used; however, these cannot be applied to microtonal music. We expect this dataset to be valuable for developing how conventional techniques, such as pitch recognition and chord recognition, can be generalized to microtones.

6 Conclusion

In this paper, we introduced a dataset titled Microtonal Music Dataset v1, consisting of 100 short pieces of music that include microtones. By advancing research based on this dataset, we believe that current music information processing techniques can be extended to include microtones, and ultimately, this could lead to the application of generative AI technology to enhance human musical expression.

In the current dataset, 100 pieces of microtonal music were created, but because the range of sounds that microtones encompass is diverse, there are plans to increase both the number of pieces and the diversity of the music in upcoming versions such as version 2 and beyond. Specifically, in order to enable the conversion of music in 12 equal temperament into microtonal music, we would like to increase the data of microtonal pieces that are paired with 12 equal temperament music. By doing so, we believe it will be possible to microtonalize existing pieces in the 12 equal temperament and significantly increase the size of the dataset. Additionally, by exploring methods of data augmentation, we plan to develop this dataset into a resource that is adequately applicable to deep learning techniques, which are indispensable for large-scale data.

References

1. Hirai, T.: Redesigning a Piano Roll: A Melody Input Interface That Can Play Microtones with an Arbitrary Number of Keys. SMC, pp.1–7 (2022)
2. Goto, M.: Development of the RWC music database. ICA, Vol. 1, pp.553–556 (2004)
3. Allan, M., and Williams, C.: Harmonising chorales by probabilistic inference. NIPS 17 (2004)
4. Boulanger-Lewandowski, N., et al.: Modeling temporal dependencies in high-dimensional sequences: Application to polyphonic music generation and transcription. ICML, pp.1881–1888 (2012)