TOWARDS IMPROVEMENT OF TRANSCRIPTION ACCURACY OF MIDI GUITAR BASED ON INTEGRATION WITH AUDIO SIGNAL PROCESSING

Masaki Otsuka
Nihon University
masaki@kthrlab.jp

Tetsuro Kitahara
Nihon University
kitahara@chs.nihon-u.ac.jp

ABSTRACT

We have been developing a method for improving the transcription accuracy of a MIDI guitar by integrating audio signal processing with the MIDI guitar's transcriptions. Here, we report the results of preliminary experiments.

1. INTRODUCTION

A MIDI guitar, which outputs the user's performance data in the MIDI format in real time, is useful for guitarists to compose musical pieces on a computer. However, a MIDI guitar causes some errors in transcription of the performance; in particular, missing notes that exist, outputting notes that do not exist, and fusing serial short notes often occur. To avoid such errors, we have been seeking to integrate audio signal processing, especially non-negative matrix factorization (NMF) [1], with transcriptions given by the MIDI guitar. In this paper, we compare transcriptions given by NMF and the MIDI guitar and discuss how the integration strategy is promising.

2. METHOD OF GUITAR PERFORMANCE TRANSCRIPTION USING NMF

NMF is a technique for separating a matrix V, representing a spectrogram, into the product of two matrices W and H, that is, $V \cong WH$, where W is a basis matrix and H is a gain matrix. The basis matrix consists of multiple basis vectors, each of which represents the spectrum of different notes, whereas the weight matrix represents the weight for each basis vector at each time.

This method can be applied after the spectrogram V is observed, in other words, after the performance is complete. However, usual MIDI guitars output the performance data in real time. Because we use NMF for improving their accuracy, the NMF also has to be performed in real time. To solve this problem, we propose a two-stage strategy. In the first stage, the user plays all chromatic notes successively using the same guitar as the one that he/she uses in the second stage. Once the performance is complete, NMF is applied and then the basis matrix is obtained. In the second stage, the user plays what is to be transcribed. While he/she plays, the weight matrix is estimated in real time using the basis matrix obtained in the first stage.

2.1 First Stage: Estimating Basis Matrix from Preliminary Performance

Once the user plays all chromatic notes successively for each string (we call this performance a preliminary performance), NMF is applied. We prepare 35 basis vectors although each string has 23 notes. After the NMF, pairs of basis vectors that have a high cosine similarity are searched and then merged.

2.2 Second Stage: Transcribing Main Performance

For every 100 ms, the power spectrum $\boldsymbol{v}_t$ (t: time) is obtained with the Fourier transform and then the gain vector $\boldsymbol{h}_t$ is calculated. The gain vector $\boldsymbol{h}_t$ is defined as $\boldsymbol{h}_t = W^{-1}\boldsymbol{v}_t$, where W is the basis matrix obtained in the first stage. Because W is not a square matrix in general, its inverse matrix cannot be calculated. We therefore use a pseudo-inverse matrix [2] instead.

After the gain vector $\boldsymbol{h}_t$ is calculated, a MIDI sequence is generated. When the value of a certain element of $\boldsymbol{h}_t$ is higher than a threshold (experimentally determined) though that of $\boldsymbol{h}_{t-1}$ is not higher than the same threshold, a MIDI Note On message for the corresponding note number is generated, and vice versa for a MIDI Note Off message.

3. EXPERIMENTAL RESULTS

We conducted an experiment on transcribing the same performance using a MIDI guitar and NMF. For the MIDI guitar, Stratcaster with Roland GK-3 and GI-10 was used. The BPMs of all performances were set to 120.

The results are shown in Figures 1–3. They are summarized as follows:

(1) The output of notes that do not exist

This type of error is mainly caused by double-pitch and half-pitch errors or brushing muted strings. The former is very common and inevitable in automatic music transcription, while the latter is difficult to avoid because the brushing of muted strings is commonly used together with chord stroking. This type of error was the most frequently shown for both the MIDI guitar and NMF, especially in

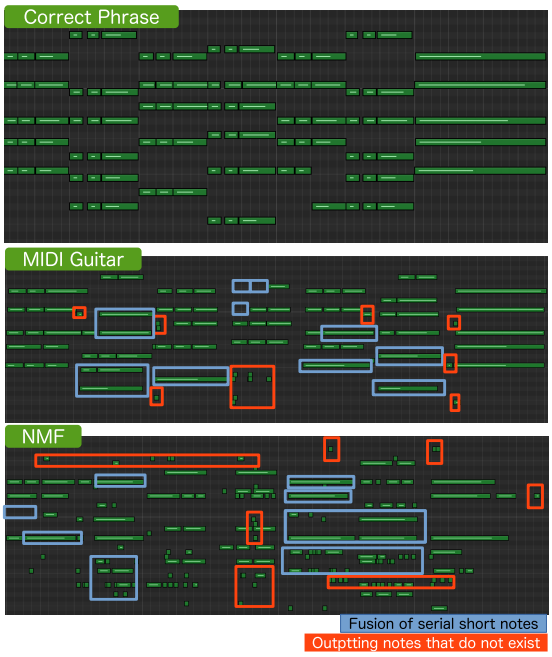

Figure 1. Transcriptions of a chord stroking using an open chord

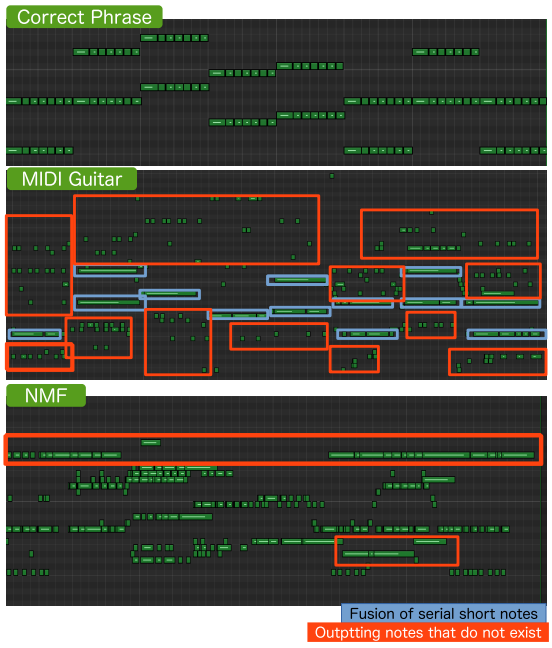

Figure 2. Transcriptions of a power chord with 16-th notes

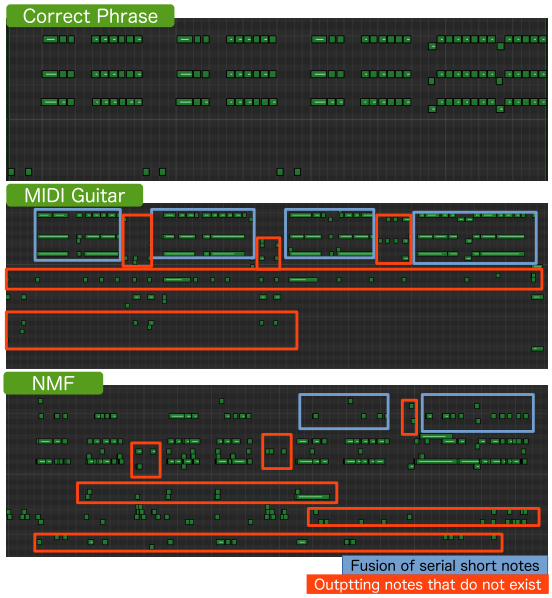

Figure 3. Transcriptions of a funk-style chord stroking

Figures 1–3, though NMF tended to output such errors a little more than the MIDI guitar.

(2) The missing of notes that exist

This type of error occurred especially in playing chord strokes. These errors were shown for both the MIDI guitar and NMF, although they were less frequent than the first type of errors.

(3) The fusion of serial short notes

This type of error occurred particularly with the MIDI guitar. For example, four successive sixteenth notes were output as one quarter note. Although these errors frequently occurred for the MIDI guitar, they were not so often seen in NMF. In particular, this type of error occurred in many parts in Figures 2 and 3 for the MIDI guitar.

4. CONCLUSION: TOWARDS INTEGRATION

In this paper, we showed our preliminary results of automatic guitar performance transcription using a MIDI guitar and NMF. In the next step, we will develop a method for integrating these two methods in order to improve the accuracy of the MIDI guitar, for example, based on the following strategies:

1. The notes output by both methods are adopted as is because they are considered highly reliable.
2. For the notes output by only one method, the reliabilities for these notes are calculated based on a certain criterion. Then, the notes with high reliabilities are adopted.
3. When a sequence of short notes with the same pitch is output by NMF although they are fused into one longer note by the MIDI guitar, the outputs by NMF are adopted because they are considered more reliable.

Acknowledgments: This work was supported by JSPS KAKENHI Grant Number 26240025.

5. REFERENCES

[1] D.D. Lee and H.S. Seung: "Learning the parts of objects with nonnegative matrix factorization", Nature, vol.401, pp.788-791, 1999.

[2] A. B. Israel and T.N.E. Greville: "Generalized Inverses: Theory and Applications", New-York, Springer, 2003.