A study on sound source identification using information of Amplitude modulation

Isamu KAWAHIRA

School of Information Science,
Japan Advanced Institute of Science and Technology

February 15, 2000

Keywords: amplitude modulation, musical instrument tone, modulation - carrier frequency map, matching, sound source identification .

1 Introduction

Recently, some computational models based on auditory scene analysis (ASA) have been proposed. All these segragation models used amplitude (or power) supectrum as acoustic features. Thus, these models can not segregate mixed signals in same frequency region completely. Sound source segragation is only half of the auditory scene analysis problem. The other half is sound source identification. In comparison with sound source segragation, relatively limited amount of work has been reported on this problem.

Most naturally occurring sounds show complex amplitude envelope patterns reflecting characteristics of their source. Namely, this is an amplitude modulation (AM). The analysis of AM plays important roles in detection, separation and identification of sound sources.

This paper uses a modulation - carrier frequency map to detect modulation frequency. A modulation - carrier frequency map is shown in figure 1 (left panel). For example, even thoughsignal component and noise component are overlap in the frequency domain, difference of amplitude modulation can be detected and signal component and noise component can be segregated by the modulation - carrier frequency map (figure 1:right panel).

This paper proposes a method for identification musical instrument tone performed by matching method between modulation-carrier frequency detection map and sound source feature database.

2 Detection of modulation - carrier frequency map

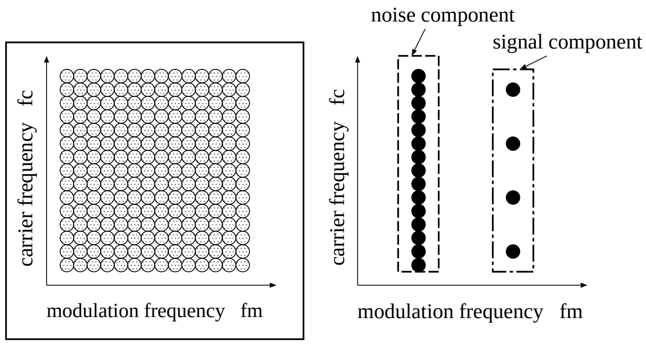

Figure 1: modulation-carrier frequency map(left), example of detection(right)

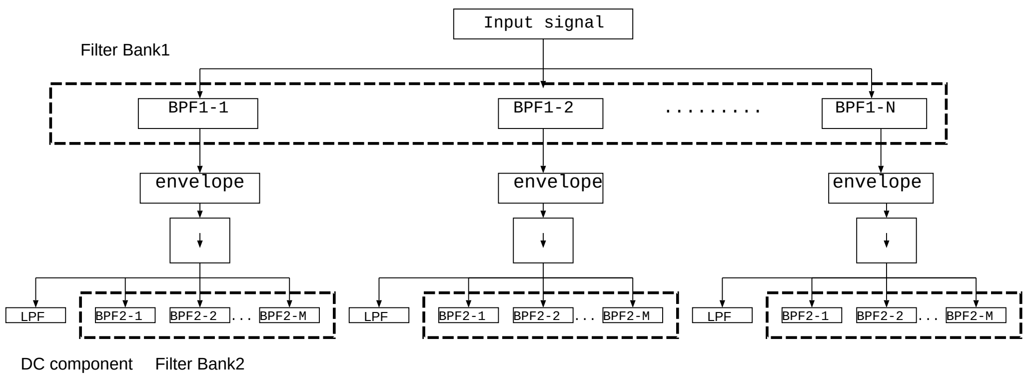

Figure 2: modulation-frequency map model

Figure 2 shows a model of the modulation - carrier frequency map. The modulation-carrier frequency map is detected as follows:

- 1. Input signal passes Filter bank 1 to separate into each carrier frequency.
- 2. Instantaneous amplitude envelope is obtained from output of Filter bank 1.
- 3. Down sampling is performed.
- 4.
  - 4-1. DC component is obtained using LPF.
  - 4-2. AM component is obtained passing through Filter bank 2.

Filter bank 1 is a bank of BPFs to filter carrier frequency. Filter bank 2 is a bank of BPFs to filter modulation frequency.

DC component includes Low frequency component and DC because LPF is used. Both Filterbank 1 and Filterbank 2 are constructed by wavelet filterbank. The basis of wavelet filterbank is a Gabor function.

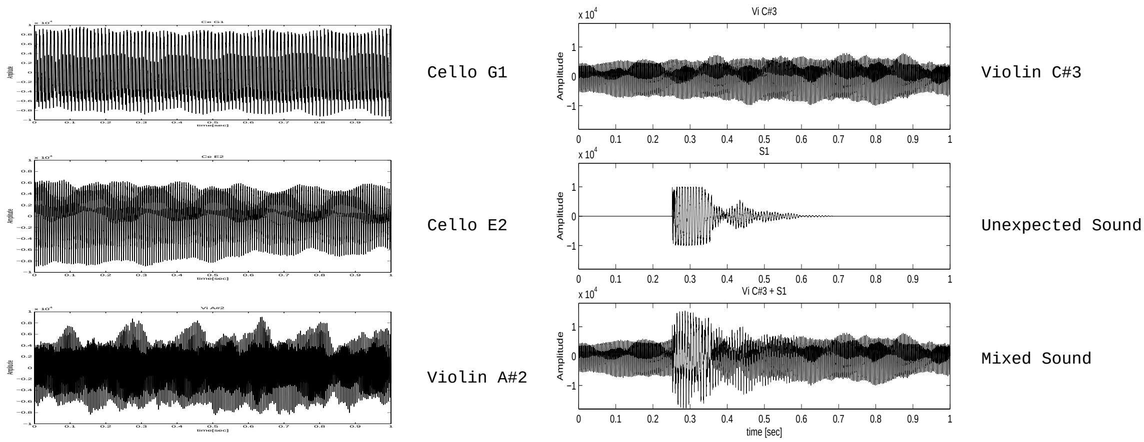

Figure 3: experimental signals

Experimental signals are shown in figure 3. The left panel signals are chello G1 (upper), Chello E2 (middle), and Violin A♯2 (lower). The right panel signals are the mixed signal (Violin C♯3 and a sudden noise).

The detected results are shown in figure 4.

The result of musical instrument tone shows that the fundamental frequency of each scale can be obtained on the high modulation frequency region in the modulation - carrier frequency map and this region varies violently.

Compared to the high modulation frequency region in the modulation - carrier frequency map, the whole region of carrier frequency axis (f_c) and region of modulation frequency axis (f_m) from 1 to 40 (64×40) is almost constant.

Therefore, the almost constant region is registered for Data Base. A matching method is performed using this region.

3 Matching for sound source identification

Sound source identification performed using two different types of matching method. It consist of the whole matching and the partial matching. The purpose of the whole matching is that evaluation measure is comfirmed for identification. The purpose of the partial matching is that influence is examined by sudden noise.

For the whole matching, 64×40 region which explaind in section 2 treated as one vector, and similarity is calculated using this vector. In this case, similarity of the same scale is over 0.9 and similarity of mixed signal is not over 0.9. From this result, if similarity for the whole matching is over 0.9, this note can be identified.

For the partial matching, 1×40 region which is divided as carrier frequency axis (f_c) treated as one vector, and using this vector similarity is calculated. In this case, a rate of similarity over 0.9 is nearly 100 % for the same scale and instrument. When sudden noise is exposed, similarity of each region is under 0.9. From this result, compared with

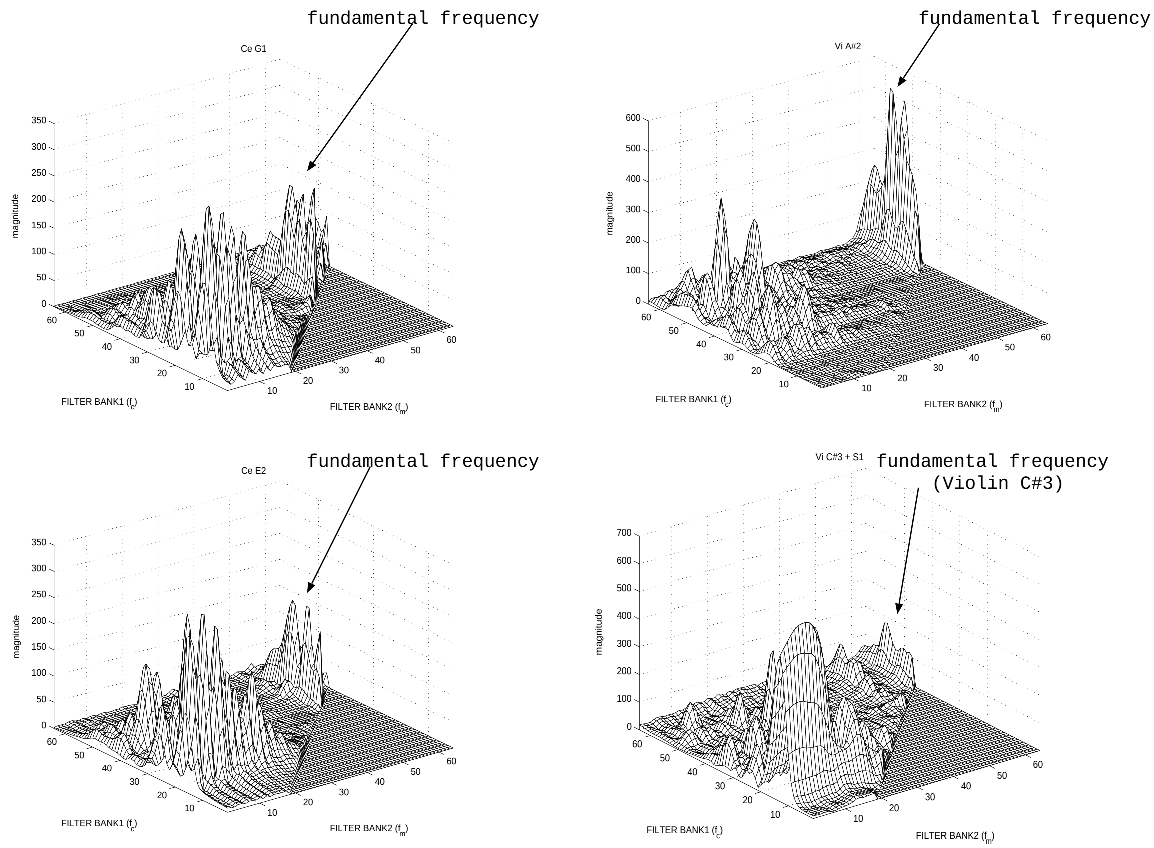

Figure 4: results of modulation-carrier frequency map

similarity of the same scale and mixed signal, information of influence by sudden noise obtains roughly.

The similarity is calculated as follows :

$$s^{(l)} = \cos\theta^{(l)} = \frac{(\boldsymbol{x}, \boldsymbol{r}^{(l)})}{\|\boldsymbol{x}\| \cdot \|\boldsymbol{r}^{(l)}\|} \qquad (0 \leq s^{(l)} \leq 1)$$

$\boldsymbol{x}$: input map
$\boldsymbol{r}^{(l)}$: data base map
$s^{(l)}$: similarity $(1 \leq l \leq 12)$

Among carrier frequency axis (f_c) from 33 to 64, similarity over 0.9 through the all time means that musical sound component exists on the map, influence by sudden noise is not in the region. Fig.5 is a result of Cello G2 by the partial matching.

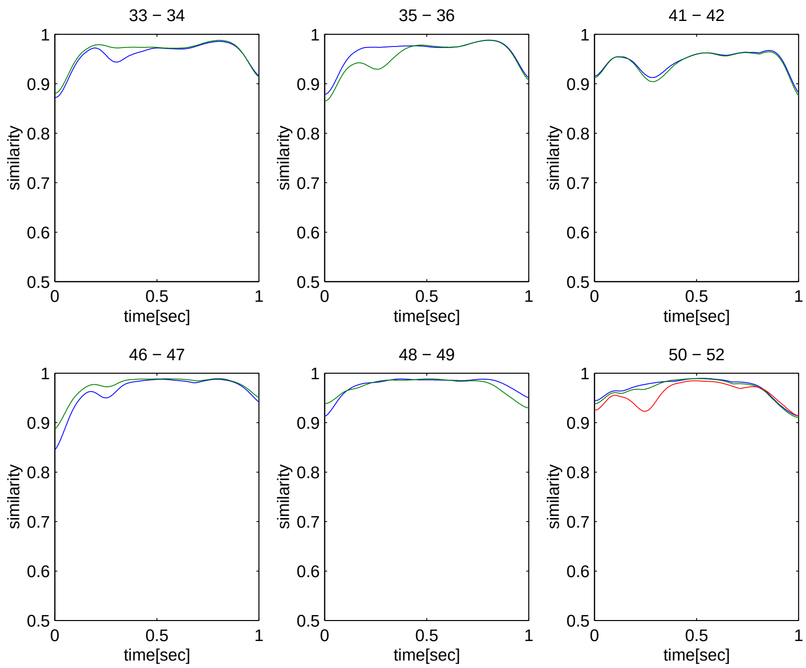

Figure 5: results of partial matching

4 Conclusion

This paper proposed a method for identification musical instrument tone performed by matching method between the modulation-carrier frequency detection map and the sound source feature database. The results show that proposed method can identify a scale of musical instrument tone when it is searched for region of modulation frequency axis (f_m) and can confirm influence for sudden noise on the modulation - carrier frequency map.