Mem. Fac. Eng., Osaka City Univ., Vol. 43, pp.37−41 (2002)

Musical Instrument Identification by Analyzing Frequency Spectrum

Kunihiro YASUDA*, Kouichi AKIYAMA** and Hiromitsu HAMA***

(Received September 30, 2002)

Synopsis

In this paper, we propose a new method that identifies a musical instrument by analyzing the frequency spectrum of the limited monophonic sound. The timbre of instruments depends on the harmonic construction, that is, the ratio of spectrum elements to the fundamental frequency. We also calculate features of the spectrum (centroid, smoothness, etc.). Therefore, in analyzing the spectrum by various methods, we try to find characteristics of the timbre. In practical, concerning the timbre of electric bass and organ, we clarify relations between the harmonic constructions and the features of the frequency spectrum.

KEYWORDS: instrument timbre, frequency spectrum, characteristic of timbre

1. Introduction

Recently, there are many research works so that the computer may perceive a music as a human[1) 2)]. In this paper, particularly, we try instrument timbre recognition. Humans can identify easily instruments on listening a music performed by many instruments, even if the person haven't learned specially. For example, if two notes played at the same loudness and pitch by different instruments, they can be distinguished easily[3)]. So the instrument timbre is supposed the difference of sound sonority except for loudness and pitch information. Then, we can consider that the harmonic construction of the frequency spectrum plays the vital role. Generally, the feature of the harmonic construction appears on power ratio between the fundamental frequency and harmonic elements. We also pay attention to even and odd harmonic elements, and analyze the relation between harmonic elements. Then, we try to carry out characteristics of timbre, and to analyze the frequency spectrum using the calculation method for spectrum features, proposed by McAdams.[4)]

2. Analyzing Frequency Spectrum

Firstly, let $x(n)$ an input signal. Then we obtain the frequency spectrum $X(k)$ by FFT (Fast Fourier Transform) for an input signal $x(n)$.

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j(2\pi/N)kn}, \quad k = 0,1,2,\cdots,N-1. \tag{1}$$

Next, we obtain the peak elements $h(i)$ over a threshold level in $X(k)$ as harmonic ones, where $i\ (=1,2,3,\cdots)$ is the degree. And when $h(i)$ is the maximum, $h(i) = h_{mas}$ is the fundamental frequency element. Then, we calculate the power ratio $r(j)$ of each $h(i)$ for the fundamental frequency element $h_{\max}$. A power ratio $r(i)$ is defined as follows:

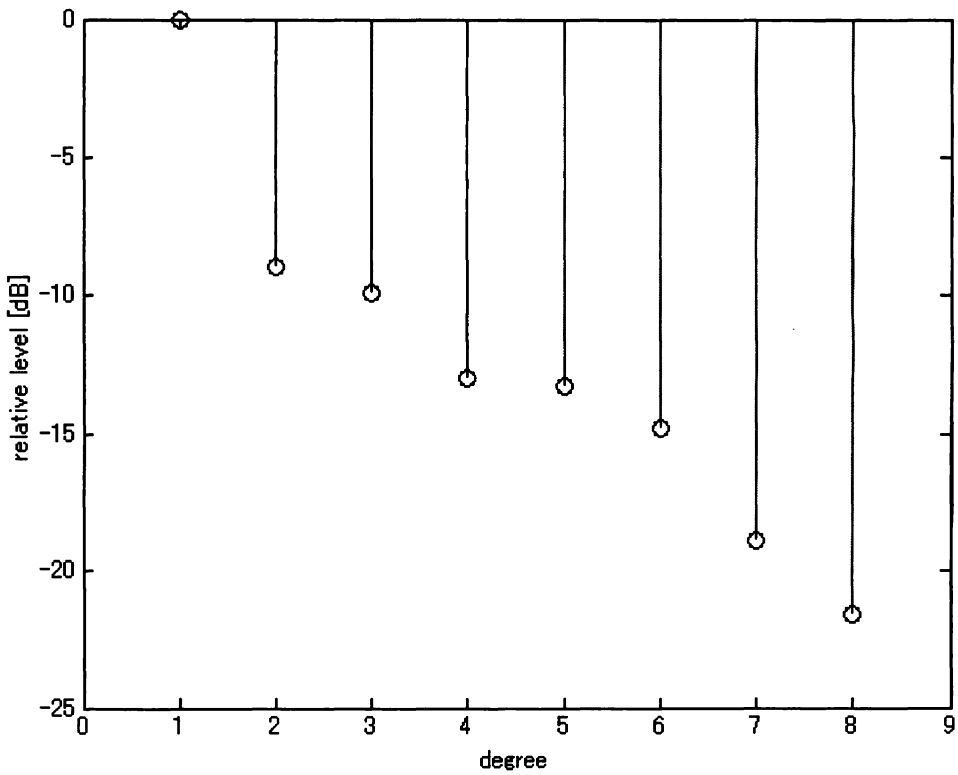
$$r(i) = 10 \log_{10} \left(h(i)/h_{\max}\right). \tag{2}$$

In this paper, we deal with the ratios of electric bass and organ in instruments as a model data. Each ratio model is the mean values of ratios which we calculated for five pitches (C, D, E, F, and G) of electric bass and organ, shown in **Fig.1.**

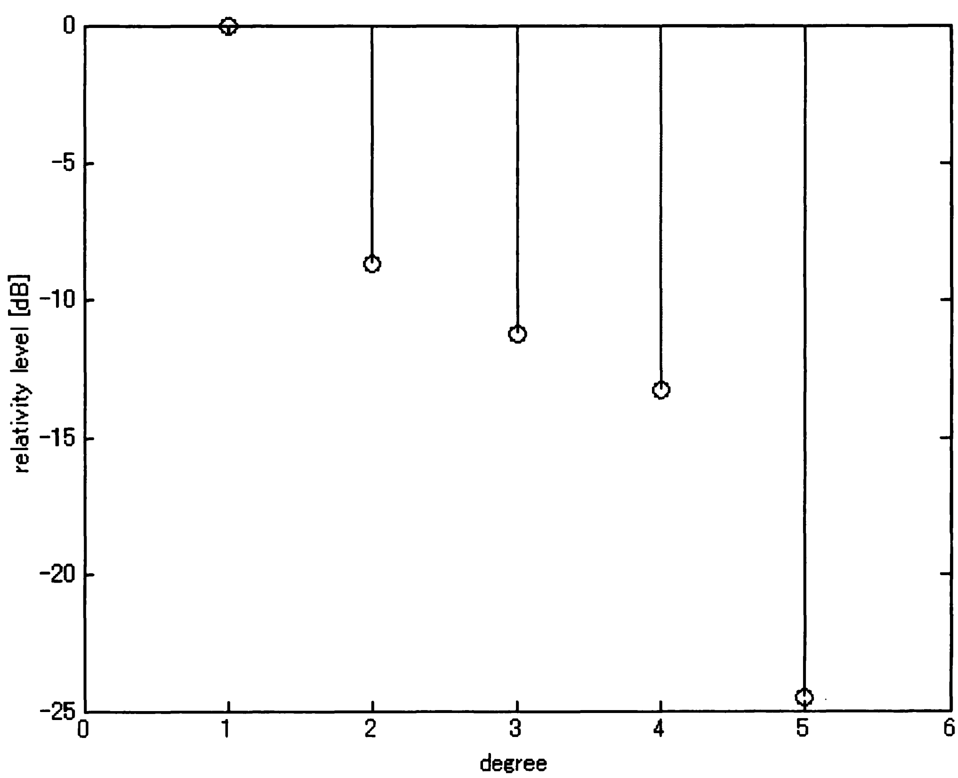
* Student, Master Course of Department of Physical Electronics and Informatics

** Student, Doctor Course of Department of Physical Electronics and Informatics

*** Professor, Department of Physical Electronics and Informatics



(a) an electric bass.



(b) an organ.
Fig.1 Harmonic element power ratio.

Secondly, we introduce the calculation method for spectrum features. This method deals with 4 parameters, that is,

LAT = (log attack time), the time it takes to progress from a threshold level to the maximum in the rms amplitude envelope.

SC = (spectral centroid), the center of gravity of the long-term amplitude spectrum.

SS = (spectral smoothness), related to the degree of amplitude difference between next partials in the spectrum calculated over the duration of the tone.

SF = (spectral flux), a measure of the degree of variation of the spectrum over time.

They are defined as:

$$LAT = \log_{10}\left(t_{\max} - t_{threshold}\right), \tag{3}$$

$$SC = \frac{1}{T}\int_0^T B(t)dt \quad with\ B(t) = \left[\frac{\sum_{k=1}^{N} kA_k(t)}{\sum_{k=1}^{N} A_k(t)}\right], \tag{4}$$

$$SS = \sum_{k=1}^{N}\left|20\log(A_k) - \frac{20\log(A_{k-1}) + 20\log(A_k) + 20\log(A_{k+1})}{3}\right|, \tag{5}$$

$$SF = \frac{1}{M}\sum_{p=1}^{M}\left|r_{p,p-1}\right| \quad with\ M = \frac{T}{\Delta t} and\ \Delta t = 16m\sec, \tag{6}$$

where, $t_{\max}$ = (the instant in time at which the rms amplitude envelope attains its maximum), $t_{threshold}$ = (the time at which the envelope exceeds a threshold value), T = (the total duration of the sound), t = (the begin time of the sliding short-term Fourier analysis window), A_k = (the amplitude of partial k), N = (the total number of partials) and $r_{p,p-1}$ = (the Pearson product-moment correlation coefficient between the amplitude spectra at times t_p and t_{p-1}).

3. Experiment and Discussion

Here, we made an experiment to analyze the frequency spectrum. We prepared five instruments sound sources (sampling frequency = 441000Hz, 16bit, mono), electric bass, organ, piano, bass drum and mallets. First, we selected two of five instruments in pairs and sounded each instrument at left and right speaker, respectively. Then, we recorded the sound by two instruments as a monophonic sound. The experimental environment is shown in **Fig.2** and the pairs of five instruments are shown in **Table 1.** Next, we calculated the frequency spectrum of the recorded sound using FFT. And then we compared the frequency spectrum with the model frequency spectrum, normalized five spectrums of electric bass and organ, respectively, using power ratio, LAT, SC, SS, and SF.

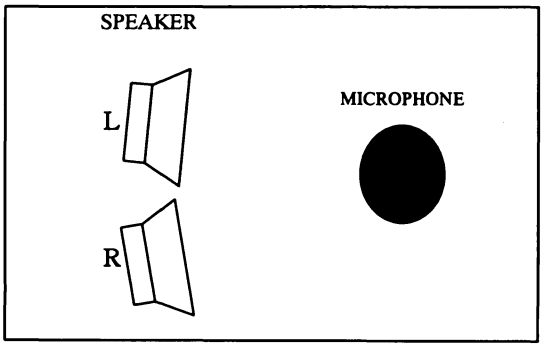


Fig.2 Experimental environment.

Table 1 Pair of instruments.

Left Speaker	Right Speaker
Electric Bass (pitch C)	Organ (pitch C)
Electric Bass (pitch D)	Organ (pitch F)
Electric Bass (pitch C)	Bass Drum
Electric Bass (pitch D)	Mallets
Piano (pitch C)	Organ (pitch E)
Mallets	Organ (pitch F)
Mallets	Piano (pitch C)
Piano (pitch C)	Bass Drum

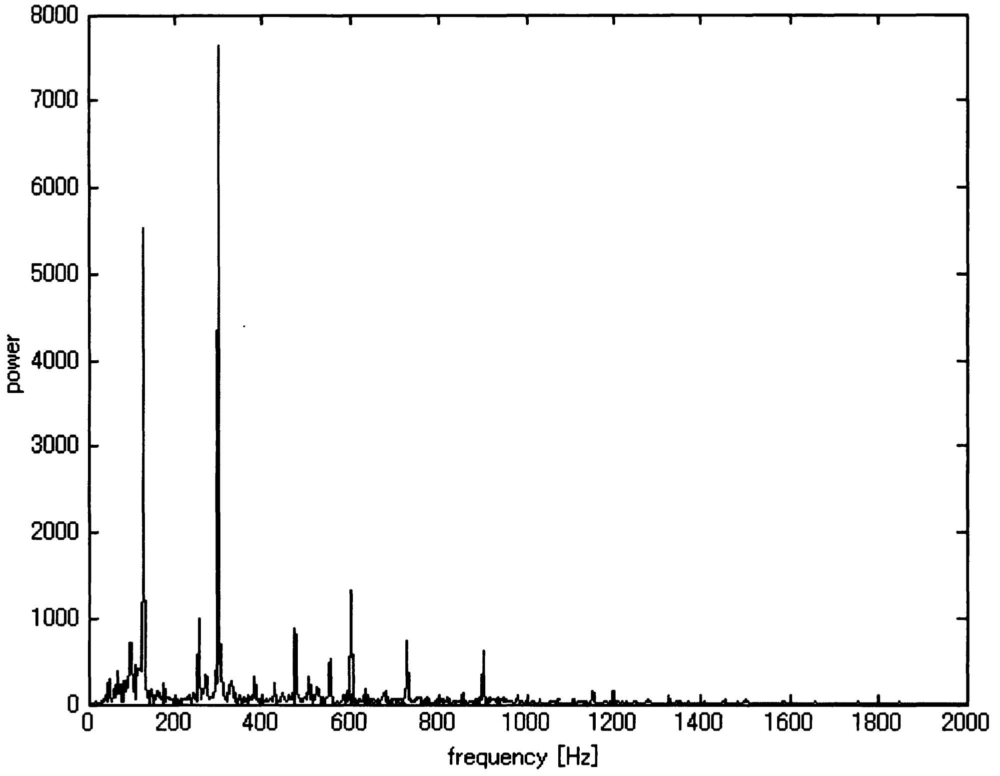
This experiment resulted in as follows. In electric bass and organ sound, the frequency spectrum was quite similar to the model data. However, since spectrum elements were overlapping when the fundamental frequency is same pitch or overtone, it was difficult to detect the spectrum elements. So, we need to propose the more accurate detection method. In electric bass and other instruments, since electric bass's frequency spectrum is precise and bass drum's is not, we could easily identify electric bass. But, electric bass and mallets, the frequency spectrum of the recorded sound was not similar to that of electric bass. We thought that the reason was for them to influence each other. In organ and others, the frequency spectrum is quite similar to organ's one. We supposed that the spectrum was characteristics of keyboard instruments. Case of mallets, organ wasn't almost involved from mallets. So, the model spectrum of organ was similar to the recorded sound. In other instruments' sound, piano's spectrum also was similar to organ's one. As a whole, concerning electric bass and organ, we would get good results. **Fig.3** shows frequency spectrum in electric bass (pitch D) and organ (pitch F), and piano (pitch C) and organ (pitch E).

4.Conclusion

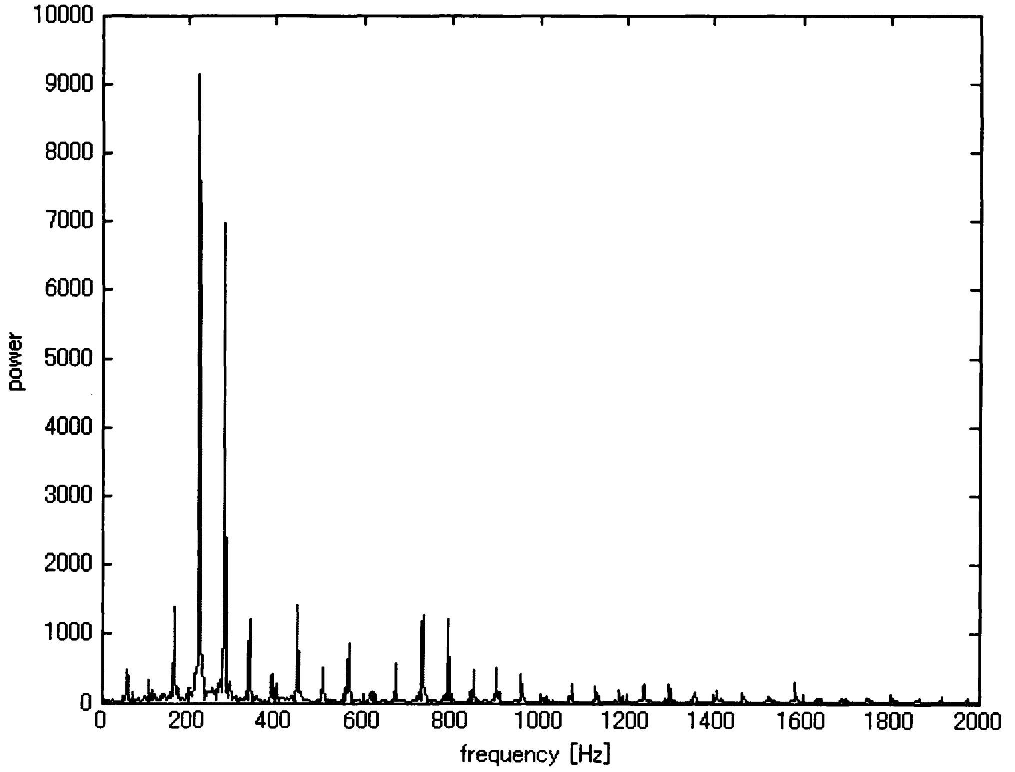
In this paper, we proposed a new method to identify instruments by analyzing the frequency spectrum. Using the power ratio of harmonic elements for the fundamental frequency element and spectrum features (LAT, SC, SS and SF), we could confirm the possibility to identify instruments in the limited monophonic sound. However, we worked only with a small number of instruments and an experimental environment. So, in near future, we will make clear how different instruments are influenced from each other, with more instruments, more microphones and so on. Then we will try to develop the more effective and robust method to find harmonic elements in the frequency spectrum. Additionally, we will analyze the frequency spectrum in terms of the relation between elements at even degrees and odd degrees.

References

1) Stephen W. Hainsworth and Malcolm D. Macleod, "Automatic Bass Line Transcription from Polyphonic Music", Signal Processing Laboratory, Department of Engineering, 2000.
2) Paul Walmsley, "Signal Separation of Musical Instruments", Signal Processing and Communications Laboratory, 1997.
3) Yoshinori Ando, "Research for Timbre of Instruments" , Chuoukouronsya, 1978.
4) Stephen McAdams, "Perspectives on the Contribution of Timbre to Musical Structure", Computer Music Journal, 23:3, pp. 85-102, 1999.



(a) electric bass and organ.



(b) piano and organ.

Fig.3 Frequency spectrum.