

MUSIC INFORMATION PROCESSING SYSTEM AND ITS APPLICATIONS TO COMPARATIVE MUSICOLOGY

Yasuaki Nakamura and Seiji Inokuchi
Department of Control Engineering
Faculty of Engineering Science
Osaka University
Toyonaka, Osaka, Japan

In the application of digital data processing techniques to comparative musicology, it is an interesting subject to determine the type of scale in folk music and to compare the similarities in melody and rhythm between different musics. This paper describes two subjects: data entry technique for the realization of a music database and its application to comparative musicology. The former deals with dictation of melody sound. The latter concerns a method to find what scale is included in the melody, using knowledge in the modern ethno-musicology, and describes a certain measure of similarity of music.

1. INTRODUCTION

Digital data processing techniques gave the ethno-musicologists a large capacity of data processing ability never before obtainable [1]. This paper deals with data entry method for realization of a music database and its applications to comparative musicology.

It is difficult to compare the music information by listening to the melody or by watching the note. It is troublesome, still more, to search what scale is included in the melody and to detect the similarity between two different melodies.

Figure 1 shows the music information processing system which has been developed for this research. Musical information is expressed by two types of pattern: one is audible pattern -sound- and another is visible pattern -music note-. Though it has analogy with the language, which has speech pattern and character, it is difficult to dictate the sound of melody and to read

the music note. This system has two data entry functions: dictation of melody and recognition of music note. A synthesizer and a display are used for monitoring the music.

2. DATA ENTRY AND DATABASE

2.1 Melody Analyser

In the early period, punched card or tape was used for comparative musicology [1]. In Japan there are a large number of folk musics which have never been collected before. So we have developed a melody analyser, which detects the pitch and the length of tone, for dictating folk music.

Pitch interval can be identified by detecting the fundamental frequency of sound. But the frequency difference ratio between the adjacent tones (semi-tone) is very small ($2^{1/12}=0.059$) and the duration of one tone may be very short, which requires that the measurement of the fundamental frequency should be done accurately during a short period [2].

As it is only fundamental frequency to be detected for pitch identification, a band-pass filter of Butterworth type ranging from 131 Hz (name of pitch: C₂) to 1047 Hz (C₅) is used. Sound signal is sampled with the relatively low sampling rate (6kHz).

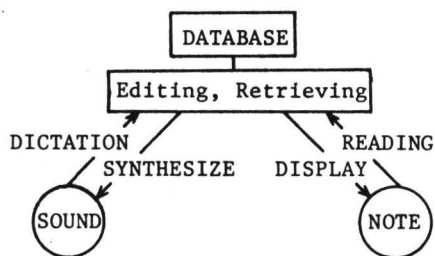


Fig.1 Music Information Processing System

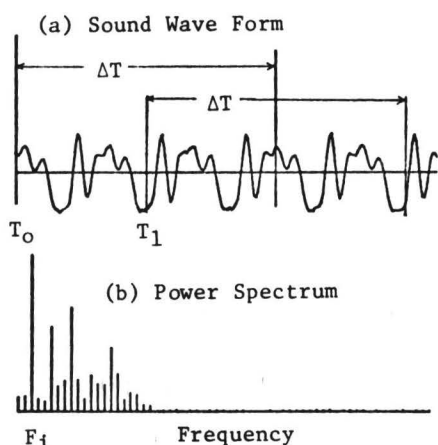


Fig.2 Frequency Analysis

For the identification of pitch, we transform the sound signal (a) in Fig.2 from T_0 to $T_0+\Delta T$ into the power spectrum (b) using the FFT algorithm. In the discrete frequency analysis, the frequency difference ΔF between adjacent spectra, which is the inverse of data length ΔT , is too large. In order to obtain the accurate frequency F_p , we compensate spectral peak frequency F_1 with the phase change as follows:

$$F_p = F_1 + \left(\frac{\theta_1 - \theta_0}{2\pi(T_1 - T_0)} \right)$$

where θ_1 and θ_0 are phases of F_1 in the interval $[T_1, T_1+\Delta T]$ and $[T_0, T_0+\Delta T]$, respectively. Pitch P is expressed as follows:

$$P = 12 \cdot \log_2(F_p/F_r)$$

where F_r is reference frequency.

The pitch obtained above has unstable frequency at the transition of tone. As the pitch of melody changes stepwise, a median filter gives good results rather than a moving averaging filter.

In case of vocal sound, F_p may vary over the range of semi-tone during the same tone because of the vibrato and the pitch scale may shift little by little for lack of absolute pitch. For the correction of them, the thresholding with hysteresis is carried out and the pitch is identified as the accumulation of the quantized pitch difference between the adjacent tones.

The segmentation is done at the time when one of the following conditions is satisfied: (1) Pitch changes stepwise, (2) Correlation of spectra between adjacent intervals becomes smaller than the predetermined value and (3) Power increases abruptly. Segmented length is quantized by using

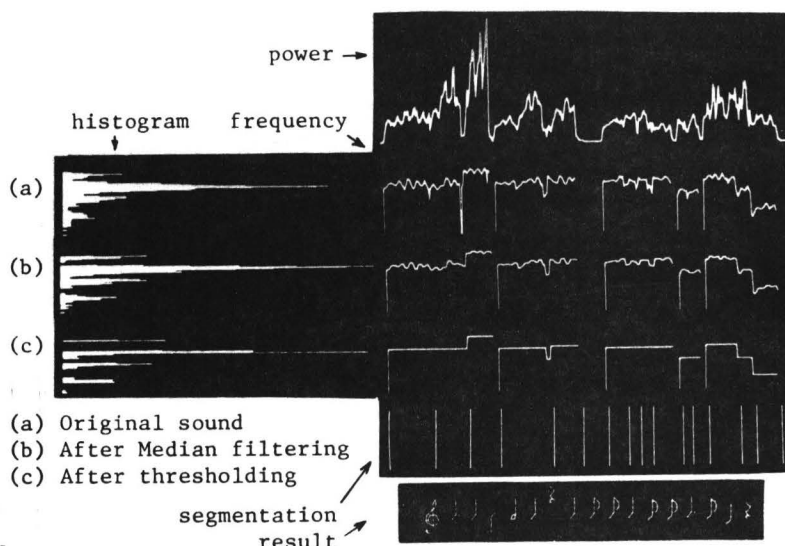


Fig.3 Dictation of Melody sound

its histogram partition. Figure 3 shows an example of dictation of a Japanese nursery song.

2.2 Music Database

Music data are encoded into a three-number code as shown in Fig.4. The m , n and l represent pitch name in octave, octave position and note length, respectively. The m and n are obtained as:

$$m = \text{Mod}_{12}[P] \quad \text{and} \quad n = [P/12]$$

Notation of the music note is not unique but the pitch code in database described above is unique [3].

Music data can be entered into the database not only by the melody analyser but also by an optical note reader or a data tablet.

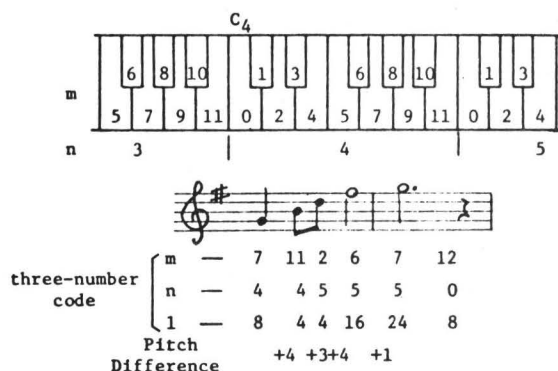


Fig.4 Data Format

3. APPLICATIONS TO COMPARATIVE MUSICOLOGY

3.1 Determination of Scale Type

In the past, the scale of Japanese folk music was classified into Ritsu, Ryo, Yo and In scales, as the European music is classified into major scale and minor one. These scales, named after their origin, were partially illogical. In the modern Japanese ethno-musicology, Koizumi[4] claims that Japanese folk music should be categorized into Minyo, Miyakobushi, Ritsu and Ryukyu scales as shown in Fig.5, where they are shown in the form of pitch difference sequence. And he also proposed that the type of scale should be determined by the type of tetrachord included in melody and the position of nuclear notes.

(1) Minyo scale	<u>+3</u> <u>+2</u> +2 <u>+3</u> <u>+2</u>
(2) Miyakobushi scale	<u>+1</u> <u>+4</u> +2 <u>+1</u> <u>+4</u>
(3) Ritsu scale	<u>+2</u> <u>+3</u> +2 <u>+2</u> <u>+3</u>
(4) Ryukyu scale	<u>+4</u> <u>+1</u> +2 <u>+4</u> <u>+1</u>

Fig.5 Scale Type of Japanese Folk Music
(Underlines indicate Tetrachord)

Tetrachords are the fundamental units to compose a scale and a melody, as structured with two Tetrachords conjuncted.

Nuclear notes, which mean the important notes of melody, have the following features:

(i) Basic note of tetrachord, (ii) Median note of three adjacent notes, (iii) Very often final tone of melody and (iv) Upper note in two-tone melody.

This knowledge makes it possible to categorize the type of scale by [1] Finding tetrachord and making its histogram and [2] Finding nuclear note and scoring it.

3.2 Similarity of Melody

The melody may be considered as a finite Markov chain process, in which the state k_i is the pitch note in scale. The pitch transition probability p_{ij} , the conditional probability that the pitch k_i is followed by k_j , has the feature of the melody.

$$P = \begin{matrix} & k_1 & k_2 & \dots & k_n \\ \begin{matrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{matrix} & \begin{pmatrix} p_{11}, & p_{12}, & & p_{1n} \\ p_{21}, & p_{22}, & & p_{2n} \\ \vdots & & & \\ p_{n1}, & p_{n2}, & & p_{nn} \end{pmatrix} \end{matrix}$$

The weight of pitch note in scale can be expressed by its steady state probability W , which can be solved by the following equations.

$$W = W \cdot P \quad \text{and} \quad \sum_{i=1}^n w_i = 1$$

The weight of pitch is effective to express the similarity between different melodies belonging to the same scale class.

3.3 Complexity of Rhythm and Melody

Let's assume the rhythm ratio R_i as:

$$R_i = \log_2(l_i/l_{i+1}) = \log_2(l_i) - \log_2(l_{i+1})$$

Its standard deviation σ_R indicates the complexity of rhythm.

$$\sigma_R = (1/n-1) \sqrt{\sum_{i=1}^n R_i^2}$$

The root mean square of pitch difference in melody can indicate the complexity of melody:

$$d_m = \sqrt{(1/n) \sum_{i=1}^n (P_i - P_{i+1})^2}$$

where P_i is the pitch of note $(m + nx12)$.

On (σ_R, d_m) plane, horizontal and vertical axes indicate the complexity of rhythm and melody, respectively.

4. CONCLUSION

A method for dictating music sound was proposed for the realization of a music database. Regarding the applications of a Japanese folk music database, this paper described an algorithm for determining the type of scale using the knowledge about the modern ethno-musicology. And some measures of similarity in melody and rhythm were proposed for the comparison of folk music.

REFERENCES

- [1] W.H. Davidow Ed. "Music by Computer", from 1966-FJCC.
- [2] H.B. Lincoln Ed. "The Computer and Music", Cornell University Press, 1970.
- [3] H.C. Longuet Higgins, "On Interpreting Bach" Machine Intelligence 6, Edinburgh, 1971.
- [4] F. Koizumi, "Research on Japanese Traditional Music", Ongaku-no-Tomo Sha, (in Japanese)