A Regional Classification of Japanese Folk Songs

— Classification based on transition probabilities of tetrachords —

Akihiro KAWASE*, Akifumi TOKOSUMI*

**University of Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan*
Tel : +81-3-5734-2680/Fax : +81-5734-3618, E-mail : { kawase, akt } @valdes.titech.ac.jp

Abstract: This study focuses on the melodies of Japanese folk songs, which were created by anonymous nonprofessional musicians and transmitted orally. We extract the musical notes from a collection of catalogued texts for Japanese folk songs. These are converted to a MIDI channel message sequence, and are classified according by Japanese regions in order to execute a principal components analysis in terms of pitch transitions based on units from the tetrachord theory, rather than octave units. The tetrachord is a tonal system, established by the Japanese musicologist Fumio Koizumi. It has a distinctive character that is completely dissimilar from Western music theory. Being constructed throughout the common practice period of music history (those periods identified as Baroque, Classical and Romantic in Western Europe), the characteristics of the melodies in Japanese folk songs represents a tonal system that derived from and formed by chant, antiphonal singing and vocalization in a primitive society, which is not observed in harmonically developed and accentuated Western music. The results indicate that Japanese folk songs that there are striking differences in terms of the tetrachord theory classification according regional location.

Keywords: Folk song, Tetrachord theory, Tonal system, Principal components analysis

1 INTRODUCTION

Jan Beran (2004) suggests that notes (or rather pitches) that differ by an octave can be considered as being equivalent with respect to their harmonic meaning, and so there are only twelve different intervals. Dealing with the set $\mathbf{Z}_{12} = \{0, 1, \ldots, 11\}$, the sum of two elements x, $y \in \mathbf{Z}_{12}$, $z = x + y$ is interpreted as the interval resulting from increasing the interval x by the interval y. For example, an instrument tuned according to an equal temperament has only twelve different notes that can be represented as the modulo 12 [1].

Wei Chai and Barry Vercoe (2001) classified Irish, German and Austrian folk songs using hidden Markov models. They represented the melodies in several ways; including an absolute pitch representation, as vectors for modulo 12 pitches; an absolute pitch with a duration representation, as vectors for the modulo 12 pitches together with durations; an interval representation, as a sequence of intervals with 27 symbols indicating -13 to $+13$ semitones; and a contour representation, as a sequence of intervals with only five interval classes [2].

While pitches that differ by an octave have been considered as being equivalent, with many studies of music classification based on that point of view, the assumption may lead to incorrect conclusions.

This is because, according to studies of phonetics, the source of sound for the human voice is the vibration of the vocal folds, with a filter consisting of the resonant cavities forming by the vocal tract, including the larynx, the pharynx, the mouth, and the nasal cavity. The pitch of a human voice varies due to the thickness, the tensile strength and the lengths of the vocal folds. Thus, there is a range within which it is possible to produce the human voice [3, 4].

This study focuses on the melodies of endangered Japanese folk songs, which were created by anonymous nonprofessional musicians and orally transmitted down from a primitive society. The aim of the study is to conduct a regional classification of Japanese folk songs by applying Fumio Koizumi (小泉文夫)'s tetrachord theory which is not based on the octave unit.

2 TETRACHORD THEORY

Influenced by the methods of Western comparative musicology, the Japanese musicologist Fumio Koizumi (1957) conceived of a scale based not on the octave unit but rather on the interval of a perfect fourth, and has developed his tetrachord theory to account for Japanese folk songs[5]. The tetrachord is a unit consisting of two stable outlining tones called the "nuclear tones", and one unstable intermediate tone located between the nuclear tones.

The term tetrachord derives from ancient Greek music theory which literally means four strings; it is a series of four tones filling the interval of a perfect fourth. Although Koizumi's tetrachord theory has only one intermediate tone, its two outer nuclear tones that form the same interval as the ancient Greek tetrachord[6].

Depending on the position of the intermediate tone, four different tetrachords can be formed (Table 2.1).

Table 2.1 : Koizumi's four basic tetrachords

types	names	pitch intervals		
1	*min'yō*	minor 3rd	+	major 2nd
2	*miyakobushi*	minor 2nd	+	major 3rd
3	*ritsu*	major 2nd	+	minor 3rd
4	*ryūkyū*	major 3rd	+	minor 2nd

For example, with three pitches outlining the interval of a fourth (for instance, A and D) and an intermediate tone dividing the fourth, there is (type 1) a minor third and a major second (with C), called the *min'yō* (民謡) tetrachord; (type 2) a minor second and a major third (with B$^\flat$), called the *miyakobushi* (都節) tetrachord; (type 3) a major second and a minor third (with B), called the *ritsu* (律) tetrachord; (type 4) a major third and a minor second (with C$^\sharp$), called the *ryūkyū* (琉球) tetrachord.

Each tetrachord was named after one of four Japanese pentatonic scales, which are conjunctions of two identical tetrachords (Figure 2.1). However, pentatonic scales are combinations of tetrachords, we do not analyze the Japanese folk songs from the point of view of musical scales. Because each tone has a completely different function from European scales, Japanese scales do not have a clear concept of the octave. For example, the most widely known anhemitonic major pentatonic scale (with semitones; e.g., C-D-E-G-A-C) is not actually one of the basic scales of Japanese music.

3 MATERIALS AND METHODS

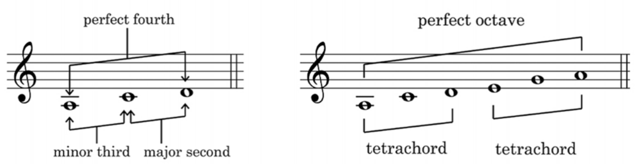
3.1 Overview of data

This study randomly sampled 20% of the music corpora included in the "*Nihon min'yō taikan* (日本民謡大観)"[7], which has 74,047 tones from 1,201 Japanese folk song pieces for 46 prefectures of Japan. "*Nihon min'yō taikan*" is a large-scale publication by the NHK (Japanese Broadcasting Corporation, 1944-1993) which is based on an idea developed by the pioneering folklorist Kunio Yanagita (柳田國男) and the musicologist Kashō Machida (町田桂聲) [6].

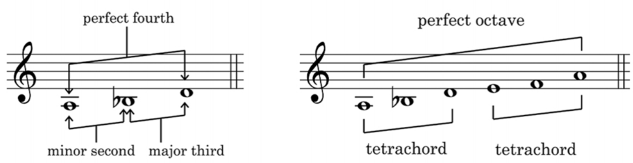
All scores are written (taken down and transposed to) with either three-flat key signatures (Cm key ↔ E$^\flat$ key) or no sharps/flats (Am key ↔ C key). These are based on two basic types of pentatonic scales, one with a semitone (the *in* scale (陰音階), Figure3.1) and another without a semitone (the *yō* scale (陽音階), Figure 3.2). Therefore, it is less to comprehend the absolute pitch than to follow the relative pitch.

Figure 3.1 : Score written with *in* scale

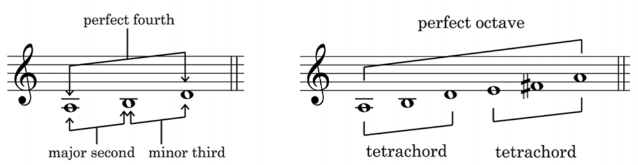
Figure 3.2 : Score written with *yō* scale



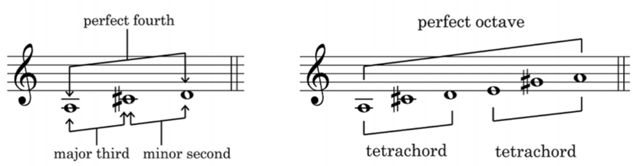
(type 1) : *min'yō* tetrachord and *min'yō* scale



(type 2) : *miyakobushi* tetrachord and *miyakobushi* scale



(type 3) : *ritsu* tetrachord and *ritsu* scale



(type 4) : *ryūkyū* tetrachord and *ryūkyū* scale

Figure 2.1 : Koizumi's four basic tetrachords and their combinations

3.2 Generating sequence of notes

In order to digitize the Japanese folk song pieces, we generate a sequence of notes as follows.

1. Create a Bitmap image file by scanning the music score.
2. Convert the Bitmap image file into a Standard MIDI File (SMF).
3. For a given SMF, extract pitch information from note events and sort the list

$$X = (x_1,\ x_2, \ldots,\ x_k, \ldots,\ x_e) \tag{1}$$

in ascending order, where $x_k (k \in \mathbf{N})$ are the notes for number k, and e is the number of elements in a sequence X.

An entire score is thus converted into a sequence of notes (symbolized as tones, e.g., D-4, C-4, D-4, E-4, G-4, ...).

3.3 Digitizing sequence of notes

We digitize the sequence of notes by the MIDI Tuning Standard

$$\text{Pitch } = 69 + 12 \log_2 \left(\frac{f}{440} \right) . \tag{2}$$

The MIDI Tuning Standard is a specification of musical pitch in order to provide compatible tuning with a range of reference frequencies among MIDI devices, where the pitch normally associated with A-5 (tuned to approximately $f = 440\,\text{Hz}$) is pitch 69. Accordingly, each number is semitonically assigned to notes on the keyboard (Figure 3.3).

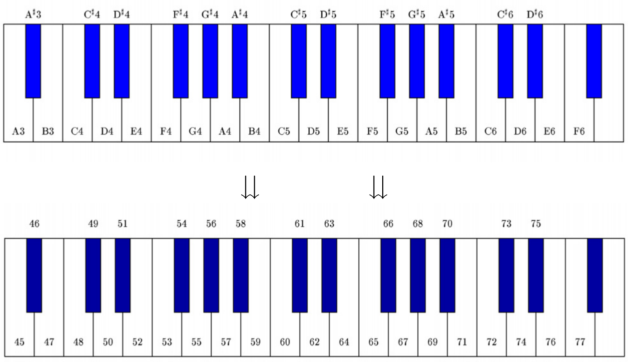


Figure 3.3 : Pitch values of MIDI Tuning Standard

The sequence of notes X yielded by equation (1) was digitized into a pitch sequence

$$S = (s_1,\ s_2, \ldots,\ s_k, \ldots,\ s_e) \tag{3}$$

from equation (2), where $s_k (k \in \mathbf{N})$ are the values for number k, and e is the number of elements in a sequence S.

3.4 Probability calculation

In the pitch sequence given by equation (3), we regard $i \in \mathbf{Z}$ as the pitch interval from one pitch to another in one step, $j \in \mathbf{Z}$ as the pitch interval from one step to another at two steps (Table 3.1). i and j can be written as

$$i = s_{k+1} - s_k \ , \ j = s_{k+2} - s_k \ .$$

Table 3.1 : Corresponding pitch intervals i and j

$\lvert i \rvert$ or $\lvert j \rvert$	pitch intervals
0	perfect first
1	minor second
2	major second
3	minor third
4	major third
5	perfect fourth
7	perfect fifth
12	perfect octave

We estimate $P_1(i)$, $P_2(j)$, the frequencies of transition to i, j by the following formulae,

$$P_1(i) = \frac{n(i)}{\displaystyle\sum_i n(i)} \quad , \quad P_2(j) = \frac{n(j)}{\displaystyle\sum_j n(j)} \ ,$$

where $n(i)$ and $n(j)$ are

$$n(i) = \sum_{k=1}^{e-1} \mathbf{1}\{i = s_{k+1} - s_k\} \ ,$$

$$n(j) = \sum_{k=1}^{e-2} \mathbf{1}\{j = s_{k+2} - s_k\} \ ,$$

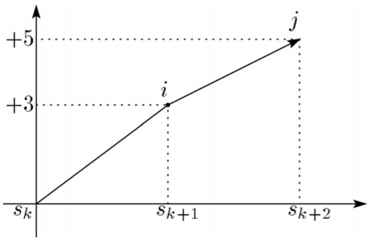
the total amounts of pitch transitions for i and j.

3.5 Transition probabilities of tetrachords

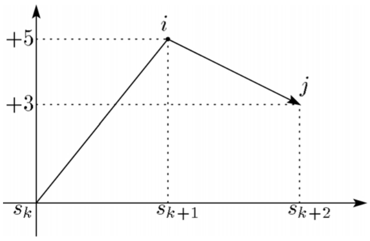
Considering the above, the probabilities of tetrachords can be obtained by multiplications of six transition patterns described in the next two pages.

First we calculate four transition probabilities of tetrachords for every prefecture separately, and create a four-dimensional data with 46 samples. Then create a four-dimensional data by dividing all 46 samples into eight traditional regions of Japan (Figure 3.4). In order to identify patterns in the data, and to highlight their similarities and differences, we applied a principal components analysis (PCA) to the data.

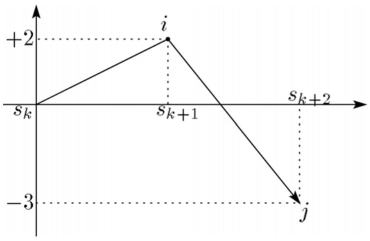
***min'yō* tetrachord**



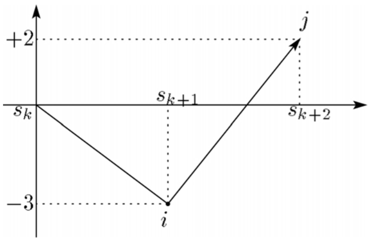
$: P_1(+3) \cdot P_2(+5)$



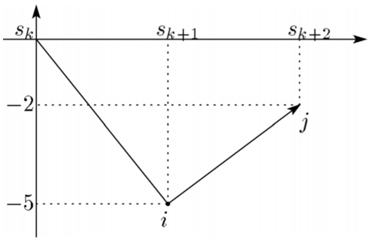
$: P_1(+5) \cdot P_2(+3)$



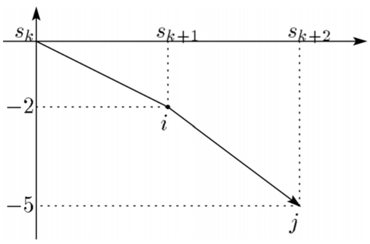
$: P_1(+2) \cdot P_2(-3)$



$: P_1(-3) \cdot P_2(+2)$



$: P_1(-5) \cdot P_2(-2)$

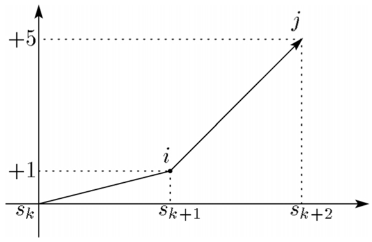


$: P_1(-2) \cdot P_2(-5)$

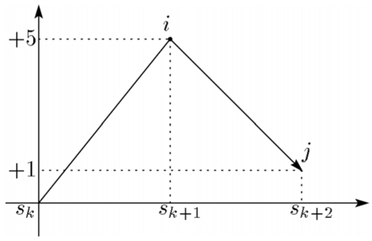
+)

$$P_1(+3) \cdot P_2(+5) + P_1(+5) \cdot P_2(+3)$$
$$+P_1(+2) \cdot P_2(-3) + P_1(-3) \cdot P_2(+2)$$
$$+P_1(-5) \cdot P_2(-2) + P_1(-2) \cdot P_2(-5)$$

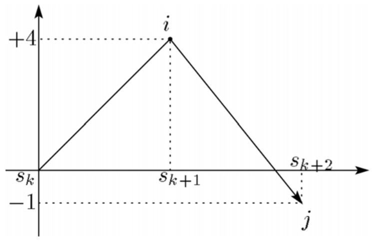
***miyakobushi* tetrachord**



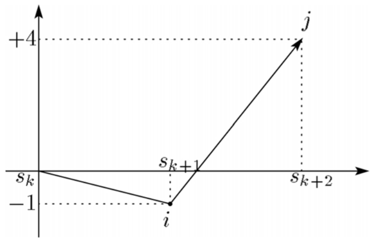
$: P_1(+1) \cdot P_2(+5)$



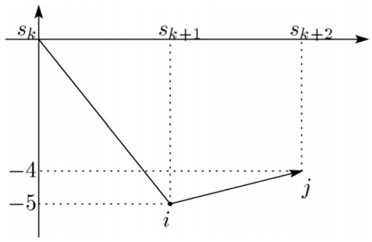
$: P_1(+5) \cdot P_2(+1)$



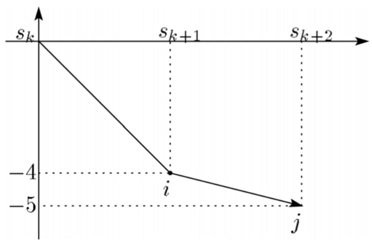
$: P_1(+4) \cdot P_2(-1)$



$: P_1(-1) \cdot P_2(+4)$



$: P_1(-5) \cdot P_2(-4)$

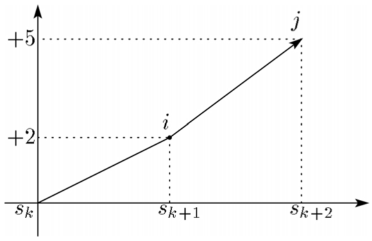


$: P_1(-4) \cdot P_2(-5)$

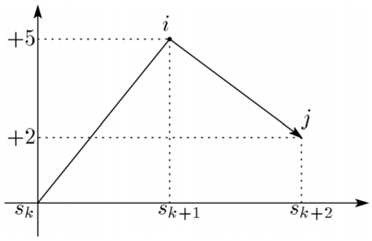
+)

$$P_1(+1) \cdot P_2(+5) + P_1(+5) \cdot P_2(+1)$$
$$+P_1(+4) \cdot P_2(-1) + P_1(-1) \cdot P_2(+4)$$
$$+P_1(-5) \cdot P_2(-4) + P_1(-4) \cdot P_2(-5)$$

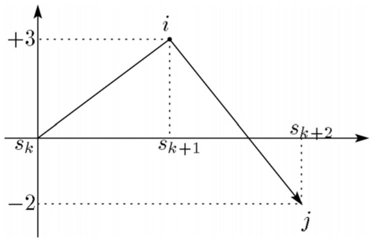
***ritsu* tetrachord**



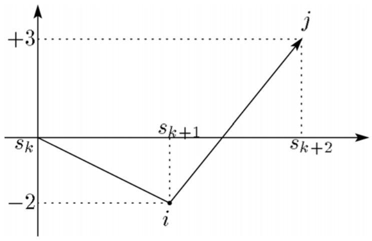
: $P_1(+2) \cdot P_2(+5)$



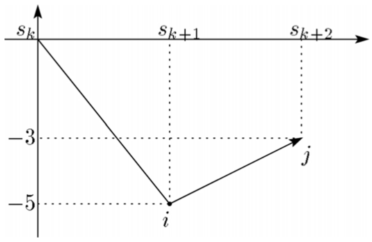
: $P_1(+5) \cdot P_2(+2)$



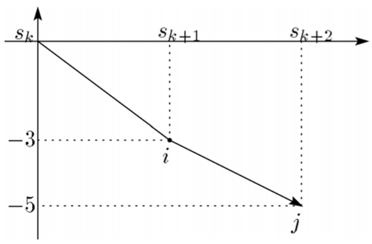
: $P_1(+3) \cdot P_2(-2)$



: $P_1(-2) \cdot P_2(+3)$



: $P_1(-5) \cdot P_2(-3)$

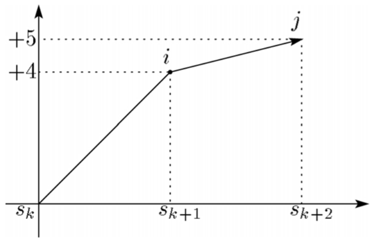


: $P_1(-3) \cdot P_2(-5)$

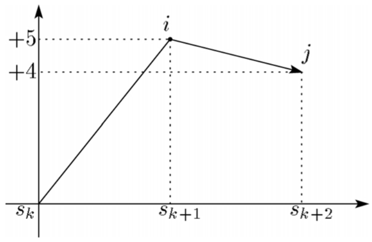
+)

$$P_1(+2) \cdot P_2(+5) + P_1(+5) \cdot P_2(+2)$$
$$+P_1(+3) \cdot P_2(-2) + P_1(-2) \cdot P_2(+3)$$
$$+P_1(-5) \cdot P_2(-3) + P_1(-3) \cdot P_2(-5)$$

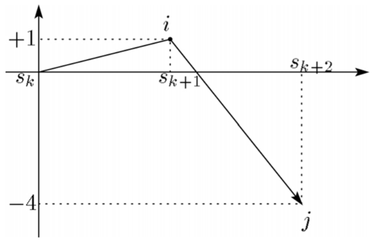
***ryūkyū* tetrachord**



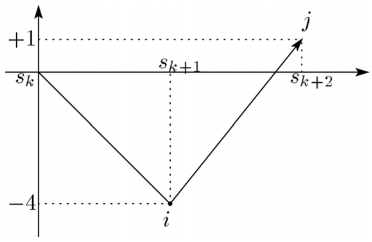
: $P_1(+4) \cdot P_2(+5)$



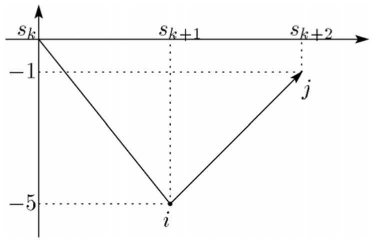
: $P_1(+5) \cdot P_2(+4)$



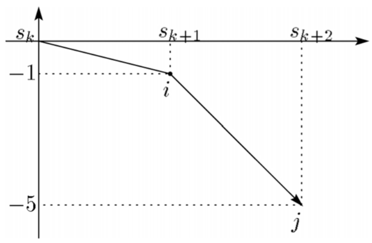
: $P_1(+1) \cdot P_2(-4)$



: $P_1(-4) \cdot P_2(+1)$



: $P_1(-5) \cdot P_2(-1)$



: $P_1(-1) \cdot P_2(-5)$

+)

$$P_1(+4) \cdot P_2(+5) + P_1(+5) \cdot P_2(+4)$$
$$+P_1(+1) \cdot P_2(-4) + P_1(-4) \cdot P_2(+1)$$
$$+P_1(-5) \cdot P_2(-1) + P_1(-1) \cdot P_2(-5)$$

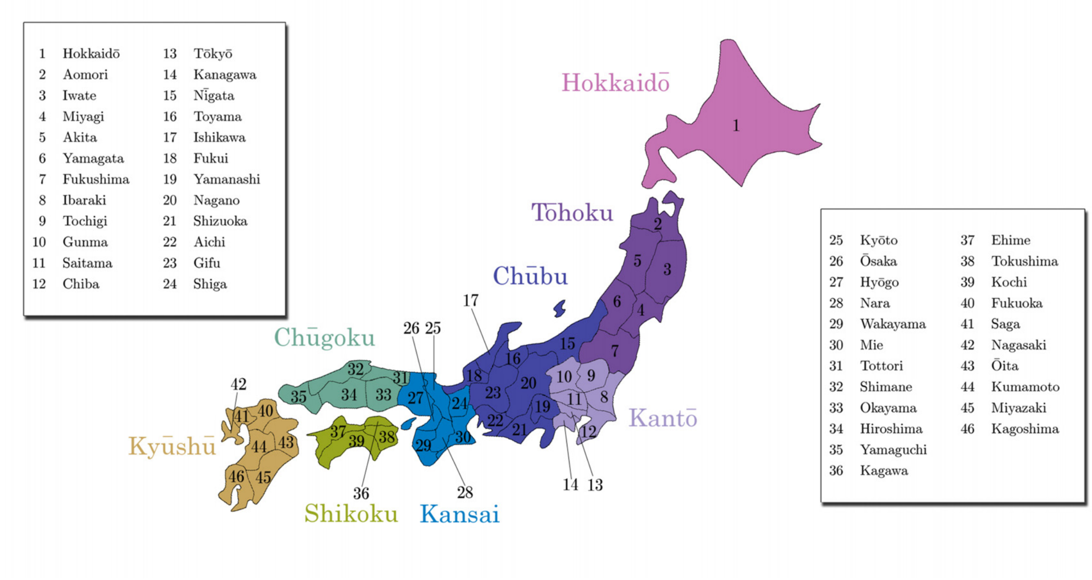


Figure 3.4 : Prefectures and regions of Japan

4 Results

4.1 Frequency of transition

The relative frequencies for i and j are shown in Figure 4.1 and 4.2.

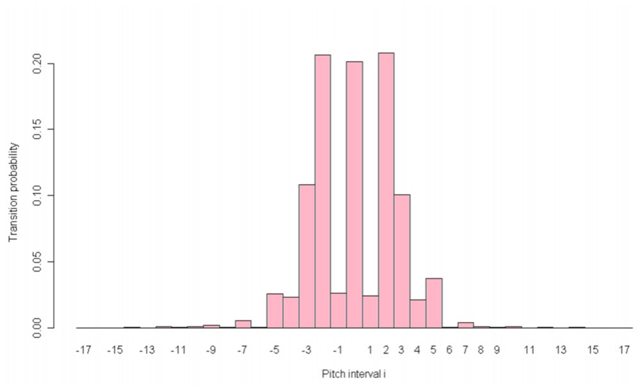


Figure 4.1 : Frequencies for interval i

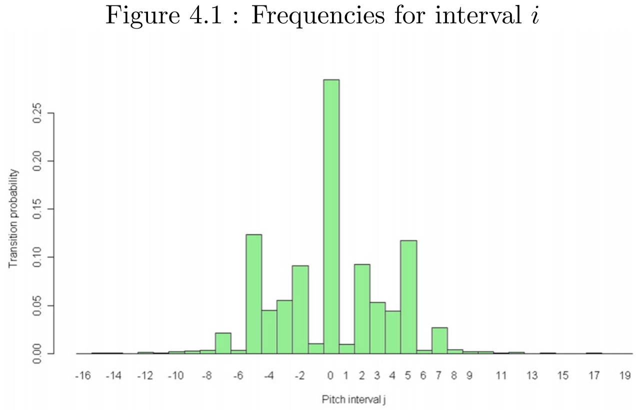


Figure 4.2 : Frequencies for interval j

For the interval i, the graph shows that the frequencies are high in the order of "±2", "0", "±3", and form a symmetric pattern with respect to $i = 0$. This implies that pitch transitions occur almost equally in both ascending and descending directions. Approximately 98.3% of the frequency is maintained to between $-5 \leq i \leq +5$; that is, the interval of a perfect fourth pitch.

For the interval j, similar to interval i, the graph shows that frequencies are extremely high at "0", then "±5", and form a symmetric pattern with respect to $j = 0$. Approximately 92.7% of the frequency is maintained to between $-5 \leq j \leq +5$.

Accordingly, this observation confirms the qualitative results of Minao Shibata (柴田南雄)'s hypothesis[8].

4.2 Principal components analysis

The implementation of the PCA is summarized in this section.

■ Prefectures

As shown in Table 4.1, the component loadings for the first two principal components of each prefecture explain more than 97% of the variability. The component loadings provide the coefficients of the principal components and give an indication of their contributions.

Table 4.1 : Component loadings (prefectures)

variable	first CPA axis	second CPA axis
min'yō	0.9232018	−0.3316606
miyakobushi	−0.9183890	−0.3851199
ritsu	0.8547430	−0.4913735
ryūkyū	−0.9042785	−0.4119282
eigenvalue	3.2440453	0.6694489
variance	81.10%	16.74%

In the first column, the absolute value of all four variables are approximately 0.9, where the *miyakobushi* and the *ryūkyū* represent a negative quantity, while the *min'yō* and the *ritsu* represent a positive quantity. This indicates that the profile of the first PCA axis z_1 is the relative pitch intervals between the nuclear tone and its intermediate tone, or, in other words, the differences in patterns of transition from the nuclear tone. Thus, as the value increases, the adjacent intermediate tone forming the nuclear tone tends to form a major second interval, and as the value decreases, it tends to form a minor second interval.

In the second column, all four component loadings represent a negative quantity, and have almost the same weight. This indicates that the profile of the second PCA axis z_2 is the persuasiveness of the tetrachord. Thus, as the value increases, the inclination of a pitch transition loses its persuasiveness, and as the value decreases, it enhances its persuasiveness.

Finding that the first two components accounted for most of the statistical variance, we can display the relevant structures in Table4.1 by plotting the first two columns against each other (Figure 4.3) We can now see that indeed there is a strong contrast between *miyakobushi* (都節), *ryūkyū* (琉球) and *min'yō* (民謡), *ritsu* (律).

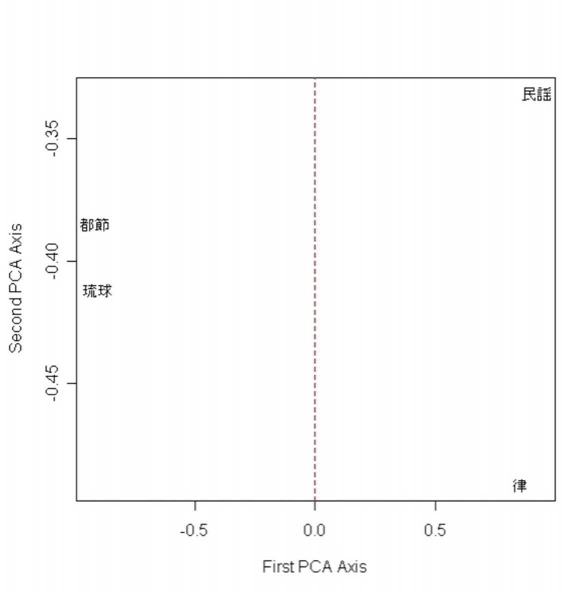


Figure 4.3 : Plot of the first two component loadings (prefectures)

The corresponding scores, z_1 and z_2 for each sample are plotted in a two-dimensional space to complete the PCA (Figure4.4).

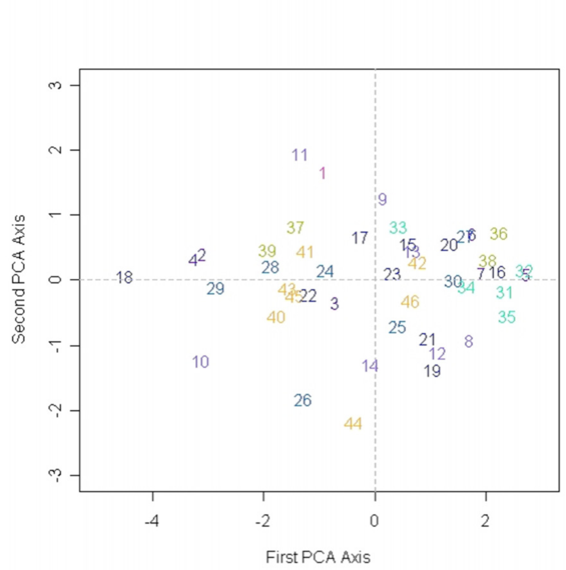


Figure 4.4 : Plot of the first two component scores (prefectures)

The result of a hierarchical cluster analysis, the dendrogram shown in Figure 4.5 is the result of applying the Ward method clustering to the corresponding scores.

If we look for a height where there are five vertical lines and trace the lines down to the individuals, the partition corresponding to five clusters gives the cluster as $C_1 = \{$2-Aomori, 4-Miyagi, 10-Gunma, 18-Fukui and 29-Wakayama$\}$; $C_2 = \{$1-Hokkaidō, 3-Iwate, 11-Saitama, 22-Aichi, 24-Shiga, 28-Nara, 37-Ehime, 39-Kōchi, 40-Fukuoka, 41-Saga, 43-Ōita and 45-Miyazaki$\}$; $C_3 = \{$5-Akita, 6-Yamagata, 7-Fukushima, 16-Toyama, 20-Nagano, 27-Hyōgo, 30-Mie, 31-Tottori, 32-Shimane, 34-Hiroshima, 35-Yamaguchi, 36-Kagawa and 38-Tokushima$\}$; $C_4 = \{$9-Tochigi, 13-Tōkyō, 15-Nīgata, 17-Ishikawa, 23-Gifu, 33-Okayama, 42-Nagasaki and 46-Kagoshima$\}$; $C_5 = \{$8-Ibaraki, 12-Chiba, 14-Kanagawa, 19-Yamanashi, 21-Shizuoka, 25-Kyōto, 26-Ōsaka and 44-Kumamoto$\}$.

We can see that groups for C_1 and C_2 lie at one extreme, with the group for C_3 at the other extreme, while C_4 and C_5 are all close to the origin. This indicates (1) that in C_1, the intermediate tone tends to form a minor second interval, has no remarkable feature to consist a tetrachord; (2) in C_2, the intermediate tone also tends to form a minor second interval with the adjacent nuclear tone, gives a weak explanation for consisting a tetrachord; (3) in C_3, the intermediate tone tends to form a major second interval, has no remarkable feature to consist a tetrachord; (4) in C_4, the intermediate tone remains unbiased, gives a weak explanation for consisting a tetrachord; and (5) similarly, in C_5, the intermediate tone remains unbiased, but gives a strong explanation for consisting a tetrachord.

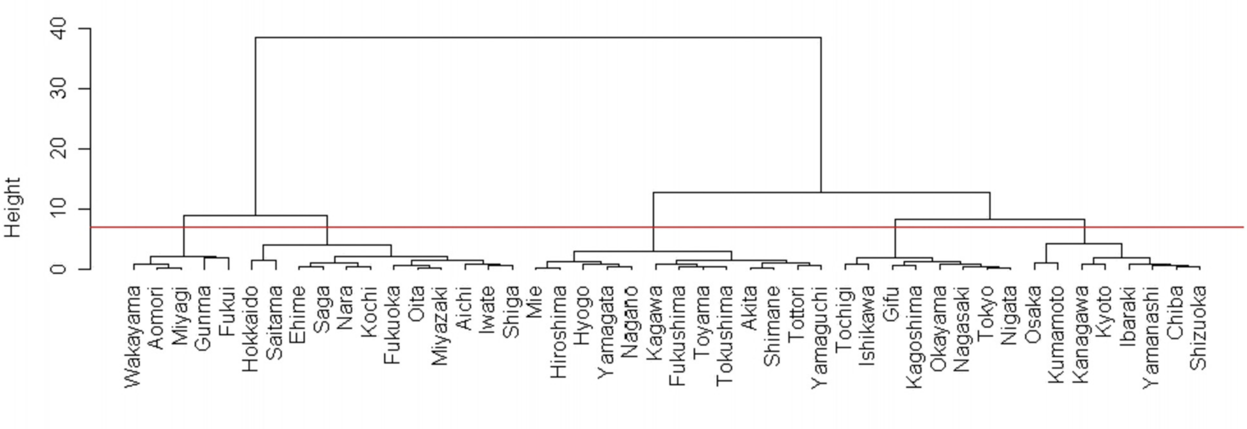


Figure 4.5 : Dendrogram based on component scores using Ward method

■ Regions

As shown in Table 4.2, the component loadings for the first two principal components of each region explain more than 95% of the variability. The component loadings provide the coefficients of the principal components and give an indication of their contributions.

Table 4.2 : Component loadings (regions)

variable	first PCA axis	second PCA axis
min'yō	0.8982806	−0.3801175
miyakobushi	−0.7362723	−0.6638233
ritsu	0.7109653	−0.6867505
ryūkyū	−0.9039719	−0.3771724
eigenvalue	2.67164177	1.19903594
variance	66.79%	29.98%

In the first column, the absolute value of all four variables are approximately 0.8, where the *miyakobushi* and the *ryūkyū* represent a negative quantity, while the *min'yō* and the *ritsu* represent a positive quantity. This indicates that the profile of the first PCA axis z_1 is the relative pitch intervals between the nuclear tone and its intermediate tone. Thus, as the value increases, the adjacent intermediate tone forming the nuclear tone tends to form a major second interval, and as the value decreases, it tends to form a minor second interval.

In the second column, all four component loadings represent a negative quantity, and have almost the same weight. This indicates that the profile of the second PCA axis z_2 is the persuasiveness of the tetrachord. Thus, as the value increases, the inclination of a pitch transition loses its persuasiveness, and as the value decreases, it enhances its persuasiveness.

Finding that the first two components accounted for most of the statistical variance, we can display the relevant structures in Table4.2 by plotting the first two columns against each other (Figure 4.6) We can now see that there is a strong contrast between *miyakobushi* (都節), *ryūkyū* (琉球) and *min'yō* (民謡), *ritsu* (律).

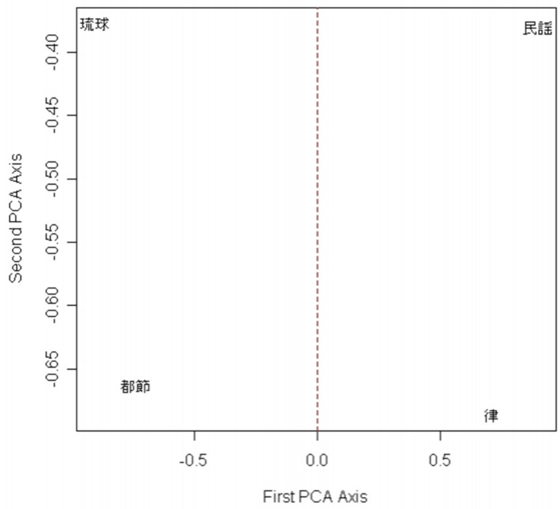


Figure 4.6 : Plot of the first two component loadings (regions)

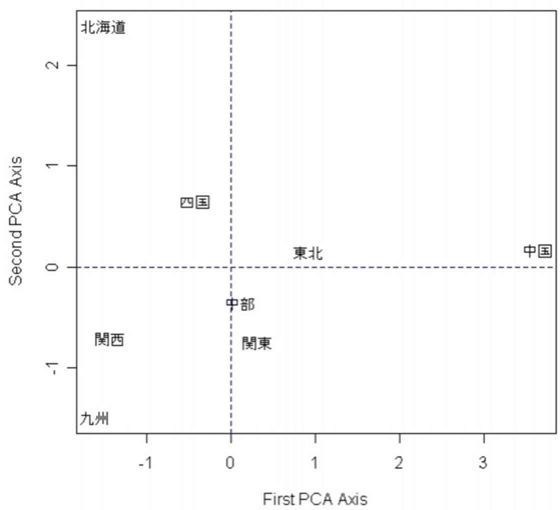


Figure 4.7 : Plot of the first two component scores (regions)

The corresponding scores, z_1 and z_2 for each sample are plotted in a two-dimensional space to complete the PCA (Figure4.7)

We can see that samples for Hokkaidō (北海道), Kansai (関西) and Kyūshū (九州) lie at one extreme, with the sample for Chūgoku (中国) at the other extreme, while Shikoku (四国), Tōhoku (東北), Kantō (関東) and Chūbu (中部) are all close to the origin. This indicates (1) that in Hokkaidō, the intermediate tone tends to form a minor second interval with the adjacent nuclear tone, gives a weak explanation for consisting a tetrachord; (2) that in Kansai and Kyūshū, the intermediate tone also tends to form a minor second interval, but gives a strong explanation for consisting a tetrachord; (3) in Chūgoku, the intermediate tone strongly tends to form a major second interval, has no remarkable feature to consist a tetrachord; (4) in Shikoku and Tōhoku, the intermediate tone remains unbiased, gives a weak explanation for consisting a tetrachord; and (5) similarly, in Kantō and Chūbu, the intermediate tone remains unbiased, but gives a strong explanation for consisting a tetrachord.

5 DISCUSSION

This study has focused on the melodies of Japanese folk songs, and has classified them according to region by executing principal components analysis in terms of pitch transitions based but on a unit of Koizumi's tetrachord theory, rather than on the octave unit.

Although regions were not classified according to geographical factors or cultural background, we were able to successively classify the melodies into two basic groups according to the behavior of the intermediate tone.

According to Gerald Groemer (1994), amateur singers of Japanese folk songs tend to identify and label their songs with highly specific names, often indicating a song's function or its place of origin (e.g., *Taue uta* (田植唄) 'rice-planting song', *Ise ondo* (伊勢音頭) 'song from *Ise*') [6]. The most widely used classification scheme, which we have also used in the present study, was established by the folklorist Kunio Yanagita and the musicologist Kashō Machida. As the song pieces that we randomly sampled cover a wide range of cultural assets from short children's song to long dance songs, it would seem to be worthwhile to analyze the pieces in terms of their original functions or specific rural localities.

References

[1] Jan Beran, "Statistics in Musicology ", Chapman and Hall, New York(2003).

[2] Wei Chai and Barry Vercoe, "Folk Music Classification Using Hidden Markov Models", *Proceedings of International Conference on Artificial Intelligence* (2001).

[3] Richard E. Berg and David G. Stork, "The Physics of Sound", Prentice Hall, New Jersey(1982).

[4] Alexander Wood, "The Physics of Music", Chapman and Hall, New York(1981).

[5] Fumio Koizumi, " *Nihon dento ongaku no kenkyu 1* (日本傳統音楽の研究 1) Studies on Traditional Music of Japan 1", Ongaku no Tomosha, Tokyo(1965).

[6] Robert C. Provine, Yoshihiko Tokumaru and J. Lawrence Witzleben, "The Garland Encyclopedia of Wold Music *Volume 7* East Asia: China, Japan and Korea", Routledge, New York(2002).

[7] Nihon Hoso Kyokai, " *Nihon min'yo taikan* (日本民謡大観) Anthology of Japanese Folk Songs, 12 Volumes", Nihon Hoso Shuppan Kyokai, Tokyo(1944-1993).

[8] Minao Shibata, " *Ongaku no gaikotsu no hanashi* (音楽の骸骨のはなし) On the Skeletal Structure of Music ", Ongaku no Tomosha, Tokyo(1978).