Pop Music Visualization Based on Acoustic Features and Chord Progression Patterns Applying Dual Scatterplots

Misa Uehara
Ochanomizu University
misa@itolab.is.ocha.ac.jp

Takayuki Itoh
Ochanomizu University
itot@is.ocha.ac.jp

ABSTRACT

Visualization is an extremely useful tool to understand similarity among large number of tunes, or relationships of individual characteristics among artists, effectively in a short time. We expect chord progressions are beneficial in addition to acoustic features to understand the relationships among tunes; however, there have been few studies on visualization of music collections with the chord progression data. In this paper, we present a technique for integrated visualization of chord progression, meta information and acoustic features in collections of large number of tunes. This technique firstly calculates the acoustic feature values of the given set of tunes. At the same time, the technique collates typical chord progression patterns from the chord progressions of the tunes given as sequences of characters, and records which patterns are used in the tunes. Our implementation visualizes the above information applying the dual scatterplots, where one of the scatterplots arranges tunes based on their acoustic features, and the other figures co-occurrences among chord progression and meta information. In this paper, we introduce the experiment with tunes of 20 Japanese pop musicians using our visualization technique.

1. INTRODUCTION

Thanks to the enlargement of storage of music players, now we can bring a large number of tunes in our daily life. Also, recent on-line music delivery services enabled us easier to find favorite tunes without visiting off-line music stores. On the other hand, it is not always easy to quickly understand which tunes are preferable for users. Though various music recommendation techniques have been developed, it is not still easy for users to understand how the recommended tunes are estimated as preferable for them. We expect visualization is an effective and intuitive approach to make users overview and understand the relevancy or similarity of particular tunes or artists easily.

There are many factors which affect user's preference or impression of pop music. Acoustic features and chord progression are typical musical factors, while visual factors such as fashions and verbal factors such as lyrics are

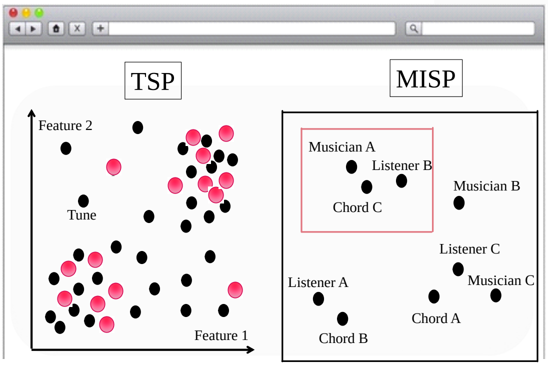

Figure 1. Illustration of the structure of the presented visualization technique. It features the tune scatterplot (TSP) in the left side of the window, and the meta information scatterplot (MISP) in the right side. When a user interactively selects a set of meta information in MISP as enclosed by a pink rectangle, dots corresponding to the selected meta information are colored in TSP.

also important for the preferences and impressions of the tunes. This paper presents our visualization technique for overview and exploration of relevancy among acoustic features, chord progression patterns, and meta information of pop tunes. The presented visualization technique displays the following two scatterplots side-by-side, as illustrated in Figure 1:

Tune scatterplot (TSP): The scatterplot visualizing a set of tunes with their acoustic feature values. Tunes correspond to dots one-by-one in this scatterplot.

Meta Information Scatterplot (MISP): The scatterplot visualizing the co-occurrence of meta information including artist names, preferred tunes of particular listeners, and chord progression patterns. Meta information corresponds to dots one-by-one in this scatterplot.

Two of the acoustic features are automatically or manually assigned to the two axes of TSP, and the dots corresponding to the tunes are displayed in TSP. Meanwhile, a dimension reduction scheme is applied to the set of meta information including chord progression patterns to place them as dots in MISP.

We suppose the following operation scenario. Users firstly look at MISP, and then interactively select a set of closely displayed dots corresponding to well-correlated meta information (e.g. artist name) and chord progression patterns. The technique assigns independent colors to each of the selected dots in MISP At the same time, tunes in TSP corresponding to the manually selected dots in MISP are also colored. Interactively selecting the acoustic features which are to be assigned to the axes of TSP, we can discover relationships among meta information, chord progression patterns, and acoustic features.

We expect this visualization technique can be used for various purposes. Students majoring in pop music analysis can study the trends related to acoustic features and chord progression patterns. Consumers can interactively explore their favorite tunes. Marketing experts can discuss how to optimize the music recommendation services by observing the visualization results.

2. RELATED WORK ON MUSIC VISUALIZATION

Visualization is a useful tool to briefly understand the contents of music, and actually several survey or tutorial presentations have been presented [2] [6]. Most of music visualization techniques can be divided into the following two categories: visualizing the detail of one tune [12] [5], and visualizing collections of large number of tunes. This paper discusses the latter category of visualization techniques.

There have been a lot of techniques for visualizing collections of large number of tunes focusing on similarity of their acoustic features. Pampalk [9] represented acoustic similarity of sets of tunes by applying self organizing map (SOM) and a graphical metaphor of islands. Goto et al. [3] represented the sets of tunes as moving objects so that users can interactively catch and play the interested tunes. Leitich et al. [8] applied GeoSOM to display a set of tunes onto a sphere based on the similarity of spectrum descriptor of the tunes. Kusama et al. [7] applied an abstract image generation technique and a zooming user interface to intuitively explore the hierarchically structured tunes. Acoustic similarity is also applied to visualize the similarity of artists [10] in addition to above studies to represent the similarity of tunes. These techniques apply acoustic analysis to organize the sets of tunes; however, it is difficult to directly read the acoustic features from their representation. It is easier to directly read the acoustic features if users can select important features and assign them to axes in the display spaces [11] [13]. However, these studies did not apply the knowledge related to chord progressions.

There have been several studies on visualization of music structure of a single tune with chord information [1] [4]. On the other hand, there have been few studies on visualization of large number of tunes applying combination of acoustic features and chord information.

3. PRESENTED VISUALIZATION TECHNIQUE

This section describes the processing flow of the presented visualization technique. The technique consists of four technical components: acoustic feature calculation, chord progression pattern matching, scatterplot construction, and interaction.

3.1 Acoustic Feature

We suppose to calculate the acoustic feature values of the given set of tunes as a preprocessing. Currently we apply MIRtoolbox [14] to calculate the following feature values.

RMS energy is the root-mean-square of the acoustic energy. This value tends to be higher while applying of recent pop, rock, or electric music, because their acoustic power is controlled as nearly constant by electric effects such as compressor and limiter. On the other hand, this value tends to be lower while applying to ballads, classical music, and other non-electric music, because their acoustic power varies along their developments. Consequently, this value is useful to divide the tunes according to their genres or instruments.

Tempo can be calculated from the cyclic patterns of power peak or harmony change. We believe tempo is important information to estimate the preference of music listeners.

Brightness is the ratio of acoustic energy of 1500Hz or higher frequency, which is mainly brought from overtones of instruments. This value is useful to divide tunes according to orchestration or recording settings: it tends to be higher if instruments which sound rich overtones (e.g. violin, saxophone, and cymbal) are effectively used by the arrangements of the tunes.

Mode is the ratio of time occupied by major or minor harmonies. This value is useful to divide enjoyable and sad sounds of the music.

Spectral irregularity is the degree of variation of the successive peaks of the spectrum. This value is useful to measure the dynamics of music.

Inharmonicity is the amount of energy outside the ideal harmonic series. This value is useful to divide traditional and modern music, because inharmonic tones are relatively often used by modern classical music, jazz, and contemporary pop music.

3.2 Chord Progression Pattern Matching

We suppose that chord progression information is provided as sequences of characters for each of given tunes. Currently we use the chord progression database for Japanese pop music on the Web [15]. Our implementation then transposes chord progressions of all the tunes to C-major or A-minor as a preprocessing.

The technique also supposes that several typical chord progression patterns are provided. Table 1 shows examples of the typical chord progression patterns used in many Japanese pop songs. Our implementation collates the prepared typical patterns with the chord progression of the tunes, and records which patterns are used in the tunes.

3.3 Scatterplot construction

The presented visualization technique supposes the set of tunes as $T = \{t_1, t_2, ..., t_m\}$, where t_i is the i-th tune, and m is the total number of tunes. A tune t_i has the values $t_i = \{f_{i1}, f_{i2}, ..., f_{in_F}, c_{i1}, c_{i2}, ..., c_{in_C}\}$, where f_{ij} is the

Table 1. Examples of typical chord progression patterns.

1	C F G
2	F G7 Em Am
3	Am F G C
4	Am Dm G Am
5	C Am F G7
6	F G Am Am
7	C Am Dm G7
8	Am Em F G7
9	C G Am Em F C F G

j-th acoustic feature value, c_{ij} is the j-th meta information value, n_F is the number of acoustic feature values, and n_C is the number of meta information values. Meta information value is a boolean variable regarding various attributes such as chord progression pattern, artist name, and preference of a listener. The corresponding value will be true if the tune contains the specific chord progression pattern, or if the specific listener prefers the tune.

TSP displays the set of tunes as m dots. The positions of the dots are calculated when the two acoustic features are assigned to the horizontal and vertical axes.

MISP represents the relevancy among meta information values as n_C dots. The technique calculates the distances between arbitrary pairs of meta information values, and applies a dimension reduction scheme to calculate the positions of the dots corresponding to the meta information values. Our current implementation calculates the distance between the u-th and v-th meta information values as $d_{uv} = 1 - m_{uv}/m$, where m_{uv} denotes the number of tunes which both u-th and v-th values are true. Then, it simply applies multidimensional scaling (MDS) to calculate the positions. This scatterplot can represent various trends of meta information, including co-occurrence of multiple chord progression patterns, and preference of chord progression patterns of artists or listeners.

3.4 Interaction

Our implementation features TSP in the left side, and MISP in the right side of the window. It also features the following interaction mechanisms.

Selection of meta information. This implementation provides an interaction to drag the pointing devices in MISP to select dots corresponding to the meta information. It assigns independent colors to the selected dots in MISP, and to the dots corresponding to the tunes in TSP.

Suppose two dots in MISP corresponding to the u-th and v-th meta information are selected by the drag operation, and blue and red are assigned to the two dots respectively. This implementation then assigns blue or red to the dots in TSP corresponding to tunes whose c_{iu} and/or c_{iv} values are true. If both c_{iu} and c_{iv} values of a particular tune are true, TSP displays the dot corresponding to this tune as combination to two hemi-circles painted in blue and red.

Our current implementation limits the number of selected dots 5 or smaller.

Selection of acoustic features. Our implementation features GUI widget buttons to select two acoustic features to be assigned to horizontal and vertical axes of TSP. Users can freely and interactively observe the relationships between meta information selected in MISP and arbitrary pairs of acoustic features.

The visualization technique also features a method for automatic selection of the acoustic features for the axes of TSP. When a user selects an arbitrary set of meta information by the drag operation in MISP, the technique evaluates the visualization results of TSP for each pair of acoustic features, and automatically applies the pair of acoustic features which brings the best visualization result. Here, we suppose it is more meaningful if colored dots concentrate at the particular portions in TSP, because such visualization results bring clearer knowledge on relationships between meta information and acoustic features. Based on this supposition, the technique calculates the entropy of the colored dots in the display space. The technique divides the display space into l_D subspaces, and count the number of the dots which a specific combination of colors are assigned. It then calculate the entropy of the dots $H = -\sum_i^{l_D} p_i$, where p_i denotes the ratio of the number of the dots in the i-th subspace. Here, we need to calculate this entropy H for each combination of the assigned colors. If MISP assigns blue and red to dots, we calculate the entropy for blue, red, and blue+red dots. In other words, we calculate the entropy for $2^{l_C} - 1$ times, if the number of assigned colors is l_C. The technique calculates the sum of the entropy $sumH = \sum^{2^{l_C}-1} H$ for each pair of acoustic features, and finally applies the pair of acoustic features which bring the smallest $sumH$ value.

4. EXAMPLES

This section introduces our experiment using the presented visualization technique. We applied 100 Japanese pop tunes including 5 tunes for each of 20 musicians.

Figure 2 shows a snapshot of our implementation. MISP displays 29 dots corresponding to 20 musicians and 9 chord progression patterns in this example. When a user drags the cursor on MISP, the implementation assigns colors to the dots which are close to the trajectory of the drag operation. Simultaneously, TSP colors the dots corresponding to the meta information (musicians or chord progression patterns) dragged on TISP. Thanks to this mechanism, users can interactively select the set of interested (and well correlated) meta information, and visually observe the relationship between the selected meta information and acoustic features.

Figure 3 shows a close up view of MISP displaying the 29 dots. Here we could observe several reasonable correlations indicated as (a), (b), and (c). Figure 3(a) depicts that the musician Tetsuya Komuro often uses the chord progression patterns “3” and “6”. Actually the chord progression pattern “3” is famously called “Komuro chord progression” by Japanese pop fans. Figure 3(b) depicts that the musicians Yumi Matsutoya, Kazumasa Oda, and Aiko commonly used similar chord progression patterns. We later found that these musicians actually used many of the patterns shown in Table 1. Figure 3(c) depicts that the mu-

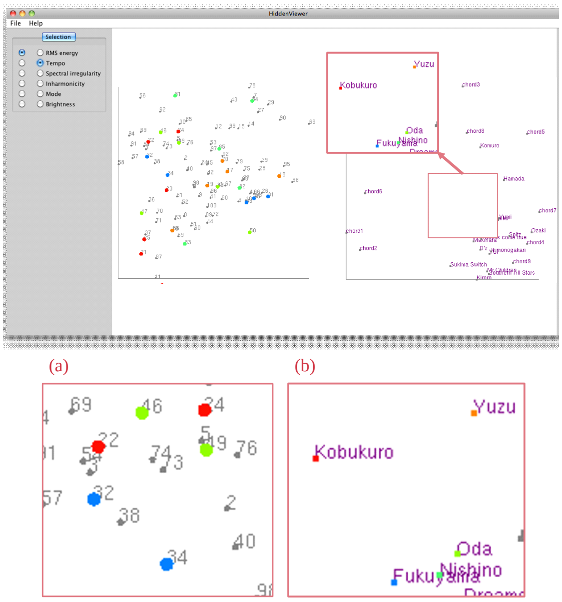

Figure 2. Snapshot of out implementation. When a user drags the cursor on MISP, the dots close to the trajectory of the drag operation are colored. Also, the dots in TSP corresponding to the meta information colored in MISP are also colored. (a) shows the colored dots in TSP. (b) shows the interactively selected dots in MISP.

sicians Mr. Children and Southern All Stars use similar chord progression patterns. We later found that they often used patterns “1” and “2”.

Figure 4 shows an interesting trend discovered during the interactive operations. Here, a user selected two dots corresponding to chord progression patterns “1” and “2” by a drag operation in MISP. These dots are colored in red and orange respectively, as indicated in Figure 4(b). At the same time, the dots corresponding to the tunes using the patterns “1” and “2” are also colored in red or orange in TSP. While dots colored in either red or orange are well scattered, dots colored in both red and orange are concentrated in the upper-left region in TSP, as indicated in Figure 4(a). We visually discovered an association rule that the tunes including both chord progression patterns “1” and “2” tend to have smaller RMS energy and larger Tempo values.

Figure 5 shows another trend discovered during the interactive operations. Here, a user selected four dots corresponding to a chord progression pattern “6”, and three artist name (Kobukuro, Masaharu Fukuyama, and Sukima Switch) during the drag operation in MISP. These dots are colored in red, orange, bright green, and cyan, respectively, as indicated in Figure 5(b). At the same time, the dots corresponding to the tunes played by one of the above artists and used the chord progression pattern “6” are also colored in red or orange in TSP. While dots colored in red are well scattered, dots colored in both red and one of other colors are concentrated in the particular region in TSP, as indicated in Figure 5(a). We found that these three artists used the same chord progression pattern to their tunes which

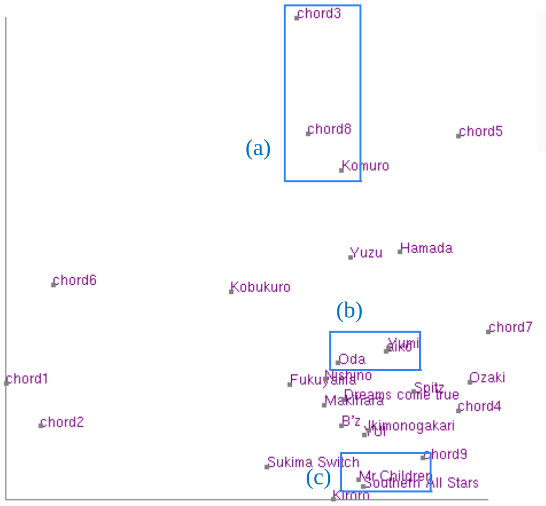

Figure 3. Close up view of MISP. We could observe several reasonable correlations among musicians or chord progression patterns.

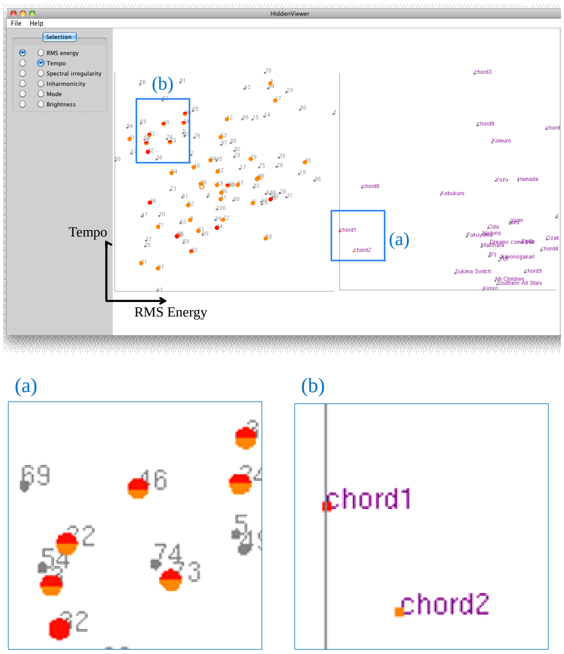

Figure 4. Example of an interesting trend. (a) Tunes which contain both the selected patterns concentrated in the upper-left region in TSP. (b) A user interactively selected two dots corresponding to chord progression patterns “1” and “2”.

have similar acoustic features.

We would like to observe more associations between meta information and acoustic features using this visualization technique, and discuss what kinds of chord progressions and acoustic features are coupled while composing and arranging pop music.

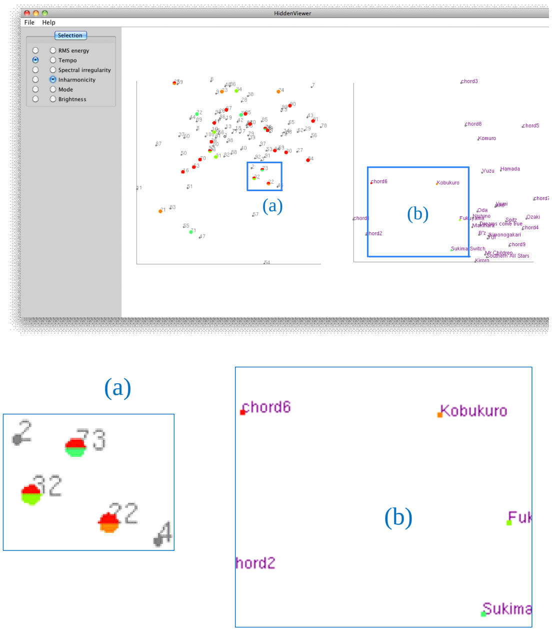

Figure 5. Example of an interesting trend. (a) Tunes which contain two of the selected meta information concentrated in the particular region in TSP. (b) A user interactively selected four dots corresponding to a chord progression pattern and three artist names.

5. CONCLUSIONS

This paper presented a visualization technique featuring dual scatterplots. One of the scatterplots (called TSP) represents the distribution of tunes with their acoustic features, while the other (called MISP) visualizes the correlations among meta information of the tunes. The technique provides an interaction mechanism to select interested set of meta information in MISP, and represent the distribution of the selected meta information in TSP. The paper introduced several examples of visualization results with Japanese pop songs to demonstrate the effectiveness of the presented technique.

The following are our potential future issues.

Listener preferences as meta information. Our current dataset only contains artist names and chord progression patterns as meta information. On the other hand, relationships between listeners' preferences and other meta information or acoustic features are also interesting and worth to be visualized. We would like to hear favorite tunes from experimental users of this technique, add the information to our current dataset, and observe the visualization result again.

Extension of chord progression pattern extraction. Our current implementation on chord progression pattern matching is too naive. There are many chord progressions which are seemingly different but theoretically similar; however, our current implementation does not recognize such patterns. We would like to extend the implementation to extract patterns more flexibly. Also, we would like to extract usage of tensions. It is often observed that specific tensions are used by specific composers or specific genre of tunes. It is also an important factor to discover the characteristics of tunes or artists.

Improvement of visual representation. Our current implementation of scatterplots just assigns colors to meta information in the order of the drag operations. In other words, the coloring mechanism does not have particular semantics. We would like to improve the mechanism so that users can understand the relationships between the selected meta information more intuitively. Also, we would like to test with other dimensionality reduction schemes to MISP. Our current implementation just applies a classical MDS. We observed inconsistent layout results, where dots corresponding to correlated meta information are distantly placed while dots corresponding to less correlated meta information are closely placed. We expect other dimensionality reduction schemes will improve the results.

Scalability test. Our current dataset is too small and therefore the examples shown in this paper does not demonstrate the scalability of the presented technique. Also, correlations among meta information in the examples are not reliable because our dataset only contains 5 tunes for each artist. We would like to extend the dataset and test the scalability of the visualization technique.

Acknowledgments

We appreciate J-Total Music for their approval of our usage of the chord progression published on the Web for our academic research purpose.

6. REFERENCES

[1] P. Ciuha, B. Klemenc, D. Solina, Visualization of Concurrent Tones in Music with Colours, ACM International Conference on Multimedia, pp. 1677-1680, 2010.

[2] J. Donaldson, P. Lamere, Using Visualization for Music Discovery, International Symposium on Music Information Retrieval, 2009.

[3] M. Goto, T. Goto, Musicream: New Music Playback Interface for Streaming, Sticking, Sorting, and Recalling Musical Pieces, 6th International Society for Music Information Retrieval, pp. 404-411, 2005.

[4] M. Goto, K. Yoshii, H. Fujihara, M. Mauch, T. Nakano, Songle: A Web Service for Active Music Listening Improved by User Contributions, 12th International Society for Music Information Retrieval, pp. 311-316, 2011.

[5] A. Hayashi, T. Itoh, M. Matsubara, Colorscore - Visualization and Condensation of Structure of Classical Music, Knowledge Visualization Currents: from Text to Art to Culture, Springer Edit Volume, ISBN-978-1-4471-4302-4, 2012.

[6] J. Holm, Visualizing Music Collections Based on Metadata: Concepts, User Studies and Design Implications, Tampere University of Technology, 2012.

[7] K. Kusama, T. Itoh, Abstract Picture Generation and Zooming User Interface for Intuitive Music Browsing, Springer Multimedia Tools and Applications, Vol. 73, No. 2, pp. 995-1010, 2014.

[8] S. Leitich, M. Topf, Globe of Music - Music Library Visualization Using GeoSOM, International Conference on Music Information Retrieval, pp. 167-170, 2007.

[9] E. Pampalk, Islands of Music: Analysis, Organization, and Visualization of Music Archives, Master Thesis, Vienna University of Technology, 2001.

[10] E, Pampalk, M. Goto, Musicrainbow: A New User Interface to Discover Artists Using Audio-based Similarity and Web-based Labeling, International Conference on Music Information Retrieval, pp. 367-770, 2006.

[11] Y. Saito, T. Itoh, MusiCube: A Visual Music Recommendation System featuring Interactive Evolutionary Computing, Visual Information Communication and Interaction Symposium (VINCI'11), 2011.

[12] H.-H. Wu, J. P. Bello, Audio-Based Music Visualization for Music Structure Analysis, Sount and Music Computing, 2010.

[13] J. Zhu, Perceptual Visualization of a Music Collection, IEEE International Conference on Multimedia and Expo, 1058-1061, 2005.

[14] O. Lartillot, MIRtoolbox, available from http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox

[15] J-Total music, http://music.j-total.net/index.html